Contextuality Scenarios Arising from Networks of Stochastic Processes

Rodrigo Iglesias
Instituto de Matemática de Bahía Blanca, CONICET
Departamento de Matemática, Universidad Nacional del Sur
Bahía Blanca, Argentina.

Fernando Tohmé
Instituto de Matemática de Bahía Blanca, CONICET
Departamento de Economía, Universidad Nacional del Sur
Bahía Blanca, Argentina.

Marcelo Auday
Instituto de Investigaciones Económicas y Sociales del Sur, CONICET
Departamento de Humanidades, Universidad Nacional del Sur
Bahía Blanca, Argentina.

Abstract

An *empirical model* is a generalization of a *probability space*. It consists of a simplicial complex of subsets of a class $\mathcal{X}$ of random variables such that each simplex has an associated probability distribution. The ensuing marginalizations are coherent, in the sense that the distribution on a face of a simplex coincides with the marginal of the distribution over the entire simplex.

An empirical model is said *contextual* if its distributions cannot be obtained marginalizing a joint distribution over $\mathcal{X}$. Contextual empirical models arise naturally in quantum theory, giving rise to some of its counter-intuitive statistical consequences.

In this paper we present a different and classical source of contextual empirical models: the interaction among many stochastic processes. We attach an empirical model to the ensuing network in which each node represents an open stochastic process with input and output random variables. The statistical behavior of the network in the long run makes the empirical model generically contextual.

1 Introduction

Let $\mathcal{X} = \{X_1, X_2, X_3\}$ be a set of random variables. A probability distribution of the joint random variable (X_1, X_2, X_3) gives rise to a family of marginal distributions of (X_1, X_2), (X_2, X_3), (X_3, X_1), X_1, X_2 and X_3. This family of distributions is said *consistent* in the sense that the distribution of X_i is the marginal of (X_j, X_i) as well as of (X_i, X_k).

Not all consistent families of distributions on (X_1, X_2), (X_2, X_3) and (X_3, X_1) arise as marginalizations of a joint distribution over (X_1, X_2, X_3):

Example 1. *Let X_1, X_2, X_3 be random variables with values in $\{0, 1\}$ and let*

$$
\begin{aligned}
P(X_1 = 0 \wedge X_2 = 0) = P(X_1 = 1 \wedge X_2 = 1) = 1/2 \\
P(X_2 = 0 \wedge X_3 = 0) = P(X_2 = 1 \wedge X_3 = 1) = 1/2 \\
P(X_3 = 0 \wedge X_1 = 1) = P(X_3 = 1 \wedge X_1 = 0) = 1/2
\end{aligned}
$$

This is a consistent family of distributions in the above sense but there is no probability assignment $P(X_1 = 0 \wedge X_2 = 0 \wedge X_3 = 0)$ of which they are its marginals.

This observation goes back to the seminal work of George Boole [4], who studied the conditions that a set of probabilities of logically related events must satisfy. Similar behaviors arise in other cases (most notably in quantum mechanics). An encompassing mathematical framework for their analysis is that of **empirical models** [2].

An empirical model consists of a set $\mathcal{X}$ of random variables $X_1, ..., X_n$, a family of subsets of $\mathcal{X}$ constituting an abstract simplicial complex, and a family of probability distributions, one for each simplex, satisfying a consistency condition. Namely, that if $\mathcal{U}' \subseteq \mathcal{U}$ are two simplices of the complex, the distribution attached to $\mathcal{U}'$ is identical to the marginal on $\mathcal{U}'$ of the distribution over $\mathcal{U}$.

Given a particular empirical model a fundamental question is whether there exists a joint distribution probability of $(X_1, ..., X_n)$ such that all the distributions over subsets of variables are obtained as its marginalization. If such a joint distribution exists, the family of consistent distributions over the subsets is said *extendable* [10]. If, on the contrary, the joint distribution does not exist, the empirical model is said to be *contextual*.

Vorobev [10] gave a combinatorial characterization of those simplicial complexes for which any family of consistent distributions is extendable. The failure in satisfying some of the assumptions that lead to this result allow the emergence of contextuality. This case, more interesting than the extendable setting, motivates the generalization of probability spaces to empirical models.

The main motivation for considering this generalization of the notion of probability space comes from quantum theory. Many contextual empirical models have been shown to agree with the statistical predictions of quantum mechanics concerning certain experimental designs [1]. According to quantum mechanics there are physically realizable families of consistent distributions of random variables which cannot be obtained as the marginals of the joint distribution of all the variables.

The formalism of quantum theory has also been used to describe probabilities in macroscopic natural phenomena beyond quantum mechanics. For an extensive developement in this direction see [6].

We analyze here a well motivated mathematical model, other than quantum theory, leading also to contextual empirical models. This alternative framework is classical, in the sense that any physical instantiation does not require quantum phenomena. It involves the interactions among many stochastic processes. In particular, our construction can be seen as a vast generalization of Example 1.

The basic unit in our framework is an *open stochastic process*. It models a device that receives the values of a set of input variables and –depending on the values of a set of internal values– generates the values of a set of output variables according to a probability distribution encoded in a stochastic matrix. Pictoricallly, we represent an open stochastic process as in Figure 1.

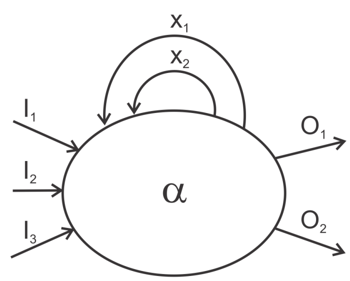

Figure 1: α represents an open stochastic process. Arrows I_1, I_2, I_3 represent the input variables, X_1, X_2 the internal variables and O_1, O_2 the output variables.

Open stochastic processes can be composed to build more complex stochastic processes connecting some of the output variables with some of the input variables of another process , as illustrated in Figure 2.

Under this representation our construction shows close resemblances to tensor networks and the operation of tensor contraction developed in the community of quantum computational complexity [8] ,[11]. We can provide a graphical represen-

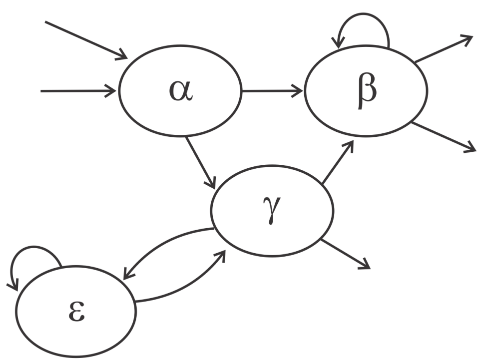

Figure 2: A network of stochastic processes. Each arrow represents a variable. An output variable of one process can become an input variable of another one.

tation of the network of stochastic processes depicting each process as a node and each variable as an arrow. The resulting network yields a single large process, as represented in Figure 3.

We focus on the case in which this large process is closed, meaning that the resulting network has no input or output variables and thus all the variables are internal. In this sense, the large closed stochastic process can be seen as a large Markov chain.

We can analyze the statistical behavior of this Markov chain at the local level –i.e. at each node of the network–, namely the probabilistic features of the relation between the input and output variables of each individual open stochastic process. More specifically, we consider the probability $P^n_t(i, j)$ –as time goes to infinity– of the event that at time t the vector of input variables of node n has value i and at $t+1$ its vector of output variables gets value j. We obtain, in this way, a distribution of possible input-output values at each node. Our main result, Theorem 2 indicates that this family of distributions constitutes an empirical model.

This paper is organized as follows. Section 2 presents the basic elements in the description of empirical models. Section 3 describes the structure and behavior of networks of stochastic processes while Section 4 recasts this information in the framework of empirical models. Section 5 illustrates how a contextual empirical model can be defined upon a network of stochastic processes. Section 6 shows that the contextuality arising from our model can be strictly stronger than that obtained from quantum systems. Section 7 concludes the paper.

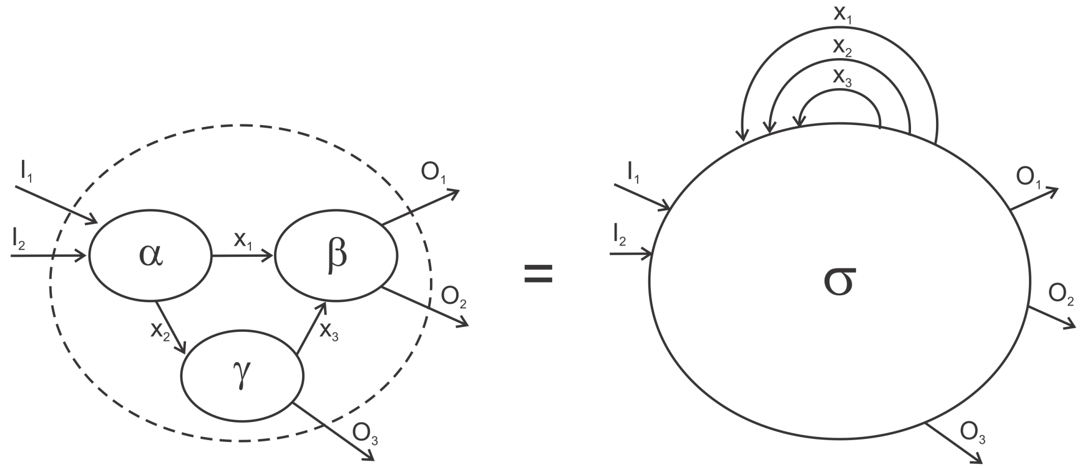

Figure 3: A network composed by stochastic processes α, β, γ that yields a single open stochastic process σ. Any variable that is simultaneously input of a node and output of another, becomes an internal variable of the resulting global process σ.

2 Empirical models

The notion of empirical model was defined in [2] as a formal framework in which the weird predictions of quantum mechanics can be made sense. A previous concept, useful for the definition of empirical models, is that of *measurement scenario* consisting of a finite set $\mathcal{X}$, a finite set $\mathcal{O}_X$ for each $X \in \mathcal{X}$, and a family $\mathcal{M}$ of subsets of $\mathcal{X}$ that covers $\mathcal{X}$ and such that if $\mathcal{U}' \subseteq \mathcal{U}$ and $\mathcal{U} \in \mathcal{M}$ then $\mathcal{U}' \in \mathcal{M}$.

The elements of $\mathcal{X}$ are called the *observables* or *variables*. When an observable X is measured, an *outcome* from the set $\mathcal{O}_X$ is obtained. Each subset $\mathcal{U}$ in the family $\mathcal{M}$ is called a *context* and it represents a subset of observables that can be jointly measured. We say that $\mathcal{U}$ is a *maximal context* if it is not properly contained in another context.

Let $\mathcal{U}$ be a subset of $\mathcal{X}$. A section s over $\mathcal{U}$ is an element of the cartesian product

$$\mathcal{E}(\mathcal{U}) = \prod_{X \in \mathcal{U}} \mathcal{O}_X$$

A *section* over $\mathcal{U}$ represents the outcome of the joint measurement of the variables in $\mathcal{U}$. If $\mathcal{U}' \subseteq \mathcal{U}$ there exists a natural restriction map $|_{\mathcal{U}'} : \mathcal{E}(\mathcal{U}) \to \mathcal{E}(\mathcal{U}')$

$$s \to s|_{\mathcal{U}'}$$

Let us represent a *probability distribution over* $\mathcal{U}$ as a formal linear combination

$$\pi = \sum_{s \in \mathcal{E}(U)} \pi_s \; s$$

such that π_s is a non-negative real number for each s and $\sum_{s \in \mathcal{E}(\mathcal{U})} \pi_s = 1$. The number π_s represents the probability of obtaining the outcome s when the variables in $\mathcal{U}$ are measured. Let $\mathcal{D}(\mathcal{U})$ be the convex set of all probability distributions over $\mathcal{U}$.

The restriction map $|_{\mathcal{U}'} : \mathcal{E}(\mathcal{U}) \to \mathcal{E}(\mathcal{U}')$ induces a map $|_{\mathcal{U}'} : \mathcal{D}(\mathcal{U}) \to \mathcal{D}(\mathcal{U}')$ by linear extension. Namely, if $\pi = \sum_{s \in \mathcal{E}(\mathcal{U})} \pi_s s$ then

$$\pi|_{\mathcal{U}'} \;=\; \sum_{s \in \mathcal{E}(\mathcal{U})} \pi_s \; s|_{\mathcal{U}'} \;=\; \sum_{s' \in \mathcal{E}(\mathcal{U}')} \left(\sum_{s|_{\mathcal{U}'} = s'} \pi_s \right) \; s' \tag{1}$$

In other words, if π is a probability distribution over a set $\mathcal{U}$ of random variables and $\mathcal{U}'$ is a subset of $\mathcal{U}$, then $\pi|_{\mathcal{U}'}$ is the *marginal distribution* of π when restricted to the variables in $\mathcal{U}'$.

Let $\mathcal{P}$ be a correspondence that assigns a probability distribution $\pi_{\mathcal{U}} \in \mathcal{D}(\mathcal{U})$ to each context $\mathcal{U} \in \mathcal{M}$. We say that the correspondence $\mathcal{P}$ is a *no-signalling empirical model* for $\mathcal{M}$ if for every $\mathcal{U}' \subseteq \mathcal{U}$ with $\mathcal{U} \in \mathcal{M}$ we have the compatibility condition

$$\pi_{\mathcal{U}}|_{\mathcal{U}'} = \pi_{\mathcal{U}'}$$

For short, we refer to no-signalling empirical models simply as empirical models. We say that a given empirical model is *non-contextual* if there exist a probability distribution $\pi_{\mathcal{X}}$ over $\mathcal{X}$ –the set of all the observables– such that for all $\mathcal{U} \in \mathcal{M}$

$$\pi_{\mathcal{U}} = \pi_{\mathcal{X}}|_{\mathcal{U}}$$

If such a global distribution fails to exist we say that the empirical model is *contextual*. As shown in Example 1, contextual empirical models can be easily constructed, by just postulating adequate distributions over subsets of variables. But random mechanisms generating this kind of behaviors are not pervasive. As noted above, quantum mechanics is a source of instances, but classical systems supporting them are much more rare. In the next section we present one, a network of stochastic processes.

3 Networks of stochastic processes

The representation of quantum processes by means of non-contextual empirical models relies on the properties of non-classical correlations among particles. A classical analogy involves the interaction among different random-generating subsystems. We will show how the system obtained through those interactions, represented as a network of stochastic processes, can be seen as an empirical model.

3.1 Stochastic processes

Intuitively, a process is a device or agent that receives as input the values of a set of variables and –depending on these values and its internal state– generates the values of another set of output variables. If this process is not deterministic and follow a probability distribution, we say that it constitutes a stochastic process.

More precisely, a *stochastic process* consists of a set of input variables $I_1, ..., I_n$, a set of internal variables $X_1, ..., X_m$ and a set of output variables $O_1, ..., O_r$, related by a stochastic matrix σ. We denote the entries of this matrix by

$$\sigma \left(\begin{array}{cccccc|cccccc} I_1 & ... & I_n & X_1 & ... & X_m & X_1 & \dots & X_m & O_1 & \dots & O_r \\ i'_1 & \dots & i'_n & x'_1 & \dots & x'_m & x_1 & \dots & x_m & o_1 & \dots & o_r \end{array} \right)$$

where this number represents the probability that, if at time t we have $I_1 = i_1, ..., I_n = i_n$ and $X_1 = x'_1, ..., X_m = x'_m$, then at time $t+1$ we have $X_1 = x_1, ..., X_m = x_m$ and $O_1 = o_1, ..., O_r = o_r$. Note that the rows of the stochastic matrix σ are indexed by tuples of input and internal variables, $(i'_1, ..., i'_n, x'_1, ..., x'_m)$, while the columns by tuples of internal and output variables, $(x_1, ..., x_m, o_1, ..., o_r)$.

This notation allows to distinguish easily the internal variables as those which appear simultaneously at both sides of the vertical line. The input variables, in turn, appear only at the left side of the vertical line but not at the right. The output variables, in turn, only appear at the right side of the vertical line but not at the left.

3.2 Composition of stochastic processes

We say that a stochastic process is *closed* if all its variables are internal. Otherwise, we say it is *open*. The importance of open processes is that they can be combined to yield new processes by connecting output variables of some process to input variables of another. Let us see how this *composition* works in an example, which is illustrated in Figure 4.

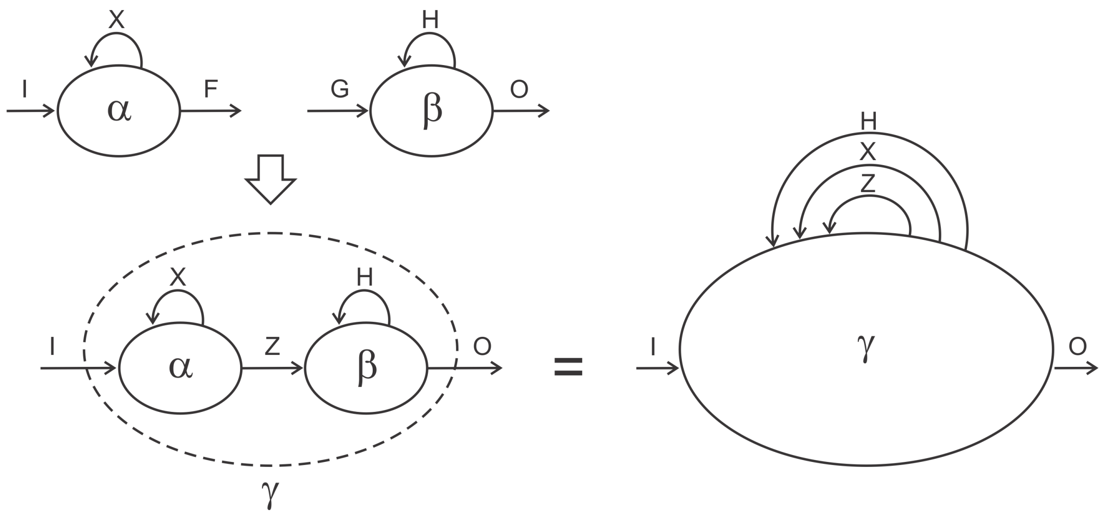

Figure 4: Processes α and β are composed by connecting the output variable F of α with the input variable G of β. The resulting process γ has one input variable I, one output variable O and three internal variables X, Z and H.

Let α and β be two open stochastic processes given by

$$\alpha \begin{pmatrix} I & X & \bigg| & X & F \\ i' & x' & \bigg| & x & f \end{pmatrix} \quad \text{and} \quad \beta \begin{pmatrix} G & H & \bigg| & H & O \\ g' & h' & \bigg| & h & o \end{pmatrix}$$

respectively. If we connect the output variable F of α to the input variable G of β we obtain the stochastic process γ whose coefficients are given by the product of coefficients of the α and β matrices:

$$\gamma \begin{pmatrix} I & X & Z & H & \bigg| & X & Z & H & O \\ i' & x' & z' & h' & \bigg| & x & z & h & o \end{pmatrix} = \alpha \begin{pmatrix} I & X & \bigg| & X & Z \\ i' & x' & \bigg| & x & z \end{pmatrix} \beta \begin{pmatrix} Z & H & \bigg| & H & O \\ z' & h' & \bigg| & h & o \end{pmatrix}$$

A key feature in this composition is that the output variable F of α and the input variable G of β become jointly a single internal variable, Z, of γ.

3.3 The network of processes

Given two stochastic processes α and β we say that α *provides* β if there exists a variable X which is at the same time an output variable of α and an input variable of β. A *reciprocity* is a pair (α, β) of stochastic processes such that α provides β and

β provides α (see Figure 5). In particular, (α, α) is a reciprocity if and only if α has at least one internal variable.

Figure 5: Left: processes α and β form a reciprocity. Right: a network without reciprocities.

We define a *network* $\mathcal{N}$ as a family $\{\sigma_i\}_{i=1,\ldots,n}$ of stochastic processes yielding a stochastic process σ obtained through the composition of members of $\{\sigma_i\}_{i=1,\ldots,n}$. The elements of this family are called the *nodes* of the network and the variables of σ are called the *arrows*. σ is called the *global process*. We say that $\mathcal{N}$ is a *network without reciprocities* if no pair (σ_i, σ_j), for $i \neq j$, is a reciprocity (see Figure 5). We say that a network is *closed* if all the variables of the global process σ are internal (see Figure 6).

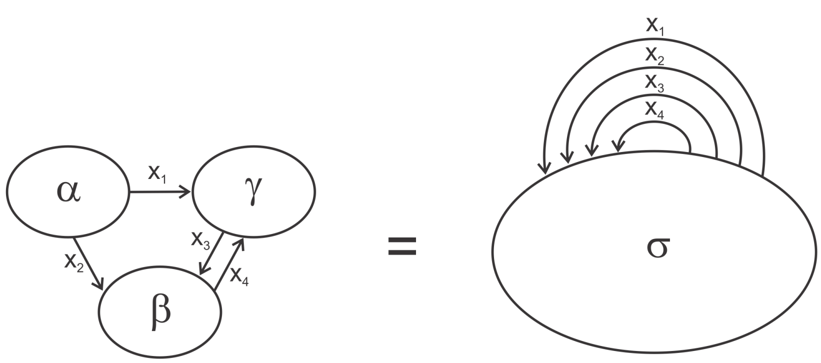

Figure 6: A closed network. All the variables become internal variables of the global process σ.

It is useful to depict a network as an oriented graph with multiple arrows where each node represents a process σ_i and each arrow represents a variable. The input variables of the process σ_i are represented by the incoming arrows of the node while its output variables are seen as arrows leaving the node. The input variables of the global process σ are represented by arrows with no initial node. In turn, the output variables of σ are arrows without terminal node. The internal variables of σ are represented by those arrows connecting two nodes of $\mathcal{N}$.

3.4 Dynamics of closed stochastic processes

Every closed stochastic process gives rise to a Markov chain. Let σ be a closed stochastic process with (internal) variables $X_1, ..., X_n$. Let

$$\pi_t \begin{pmatrix} X_1 & \dots & X_n \\ x_1 & \dots & x_n \end{pmatrix}$$

be the probability that at time t we have $X_1 = x_1, ..., X_n = x_n$. Then the evolution of this probability distribution is determined by

$$\pi_{t+1} \begin{pmatrix} X_1 & \dots & X_n \\ x_1 & \dots & x_n \end{pmatrix} = \sum_{x'_1 \dots x'_n} \sigma \left(\begin{matrix} X_1 & \dots & X_n \\ x'_1 & \dots & x'_n \end{matrix} \,\middle|\, \begin{matrix} X_1 & \dots & X_n \\ x_1 & \dots & x_n \end{matrix} \right) \pi_t \begin{pmatrix} X_1 & \dots & X_n \\ x'_1 & \dots & x'_n \end{pmatrix}$$

For short, we write this equation as

$$\pi_{t+1} \begin{pmatrix} X \\ x \end{pmatrix} = \sum_{x'} \sigma \left(\begin{matrix} X \\ x' \end{matrix} \,\middle|\, \begin{matrix} X \\ x \end{matrix} \right) \pi_t \begin{pmatrix} X \\ x' \end{pmatrix}$$

If we consider the stochastic matrix σ as a linear transformation, we see that it transforms the convex set of all probability distributions into itself. Then, by the Brouwer's fixed point theorem[1], there exists at least one probability distribution fixed by σ. Such a distribution is called a *stationary distribution*.[2]

4 The empirical model associated to a network

From now on, let $\mathcal{N}$ be a closed network without reciprocities with a global process σ. Given a node β of the network $\mathcal{N}$, let $\mathcal{I}_\beta = \{I_1, ..., I_n\}$ and $\mathcal{O}_\beta = \{O_1, ..., O_m\}$ be

[1] *"For any continuous mapping $f : \Delta \to \Delta$, with Δ a compact and convex set, there exists $\bar{x} \in \Delta$ such that $f(\bar{x}) = \bar{x}$."* In our case the mapping is given by σ, and Δ is the compact and convex set of probability distributions on the internal variables.

[2] According to the Convergence Theorem of finite Markov chains (Theorem 4.9 of [7]), the *irreducibility* and *aperiodicity* of σ are sufficient conditions for the uniqueness of a stationary distribution ω, which is obtained as the limit

$$\omega = \lim_{t \to \infty} \pi_t$$

respectively the set of input and output variables of β. Note that β has no internal variables since $\mathcal{N}$ lacks reciprocities. We will define a probability distribution δ_β over $\mathcal{I}_\beta \cup \mathcal{O}_\beta$. If the global process σ has a unique stationary distribution, the coefficient

$$\delta_\beta \begin{pmatrix} I_1 & ... & I_n & O_1 & ... & O_m \\ i_1 & ... & i_n & o_1 & ... & o_m \end{pmatrix}$$

will be interpreted as the limit –when $t \to \infty$– of the probability that at time t we have $I_1 = i_1, ..., I_n = i_n$ and at time $t+1$ we have $O_1 = o_1, ..., O_m = o_m$.

This distribution does not depend on the uniqueness of the stationary distribution ω of the global process σ. Let $\omega|_{\mathcal{I}_\beta}$ be the marginal of ω over the subset of input variables $\mathcal{I}_\beta$. The distribution δ_β is defined by

$$\delta_\beta \begin{pmatrix} I_1 & \dots & O_m \\ i_1 & \dots & o_m \end{pmatrix} = \beta \left(\begin{matrix} I_1 & \dots & I_n \\ i_1 & \dots & i_n \end{matrix} \,\middle|\, \begin{matrix} O_1 & \dots & O_m \\ o_1 & \dots & o_m \end{matrix} \right) \omega|_{\mathcal{I}_\beta} \begin{pmatrix} I_1 & \dots & I_n \\ i_1 & \dots & i_n \end{pmatrix}$$

Theorem 1. *Let $\mathcal{N}$ be a closed network without reciprocities. Let ω be a stationary distribution of the entire process of $\mathcal{N}$. Let β be a node of $\mathcal{N}$ and let $\mathcal{I}_\beta$, $\mathcal{O}_\beta$ and δ_β be defined as above. Then the marginal distributions of δ_β over $\mathcal{I}_\beta$ and $\mathcal{O}_\beta$ respectively are given by*

$$\delta_\beta|_{\mathcal{I}_\beta} = \omega|_{\mathcal{I}_\beta} \qquad\qquad \delta_\beta|_{\mathcal{O}_\beta} = \omega|_{\mathcal{O}_\beta}$$

Proof. Let I be the joint of the variables $I_1, ..., I_n$ and let O be the joint of the variables $O_1, ..., O_m$. The first equality is obtained using the fact that node β can be identified with its stochastic matrix:

$$\delta_\beta|_{\mathcal{I}_\beta} \begin{pmatrix} I \\ i \end{pmatrix} = \sum_o \delta_\beta \begin{pmatrix} I & O \\ i & o \end{pmatrix} \tag{2}$$

$$= \sum_o \beta \left(\begin{matrix} I \\ i \end{matrix} \,\middle|\, \begin{matrix} O \\ o \end{matrix} \right) \omega|_{\mathcal{I}_\beta} \begin{pmatrix} I \\ i \end{pmatrix} \tag{3}$$

$$= \omega|_{\mathcal{I}_\beta} \begin{pmatrix} I \\ i \end{pmatrix} \tag{4}$$

For the second equality, let R be the joint of the variables of the global process σ which are not contained in $\mathcal{I}_\beta \cup \mathcal{O}_\beta$. Let θ be the process obtained by restricting the network $\mathcal{N}$ to the nodes other than β (see Figure 7).

This means that the entire process σ is factorized as

$$\sigma \left(\begin{matrix} I & R & O \\ i' & r' & o' \end{matrix} \,\middle|\, \begin{matrix} I & R & O \\ i & r & o \end{matrix} \right) = \beta \left(\begin{matrix} I \\ i' \end{matrix} \,\middle|\, \begin{matrix} O \\ o \end{matrix} \right) \theta \left(\begin{matrix} O & R \\ o' & r' \end{matrix} \,\middle|\, \begin{matrix} R & I \\ r & i \end{matrix} \right)$$

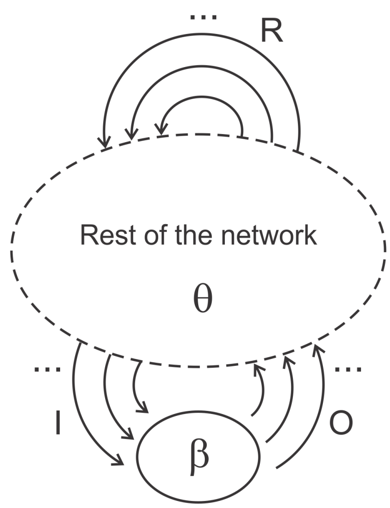

Figure 7: I is the joint of all the input variables of β, while O is the joint of all its output variables. R is the joint of the rest of the variables in network $\mathcal{N}$. Process θ is the restriction of $\mathcal{N}$ to the nodes that are not β.

We will use the fact that ω is a stationary distribution of the process σ:

$$\omega\begin{pmatrix} I & R & O \\ i & r & o \end{pmatrix} = \sum_{i',r',o'} \sigma\left(\begin{matrix} I & R & O \\ i' & r' & o' \end{matrix} \middle| \begin{matrix} I & R & O \\ i & r & o \end{matrix}\right) \omega\begin{pmatrix} I & R & O \\ i' & r' & o' \end{pmatrix}$$

Then

$$
\begin{aligned}
\delta_\beta|_{\mathcal{O}_\beta}\begin{pmatrix} O \\ o \end{pmatrix} &= \sum_{i'} \delta_\beta \begin{pmatrix} I & O \\ i' & o \end{pmatrix} && (5)\\
&= \sum_{i'} \beta \left(\begin{matrix} I \\ i' \end{matrix} \;\middle|\; \begin{matrix} O \\ o \end{matrix}\right) \omega|_{\mathcal{I}} \begin{pmatrix} I \\ i' \end{pmatrix} && (6)\\
&= \sum_{i'} \beta \left(\begin{matrix} I \\ i' \end{matrix} \;\middle|\; \begin{matrix} O \\ o \end{matrix}\right) \sum_{r',o'} \omega \begin{pmatrix} I & R & O \\ i' & r' & o' \end{pmatrix} && (7)\\
&= \sum_{i'} \beta \left(\begin{matrix} I \\ i' \end{matrix} \;\middle|\; \begin{matrix} O \\ o \end{matrix}\right) \sum_{r',o'} \sum_{i,r} \theta \left(\begin{matrix} O & R \\ o' & r' \end{matrix} \;\middle|\; \begin{matrix} R & I \\ r & i \end{matrix}\right) \omega \begin{pmatrix} I & R & O \\ i' & r' & o' \end{pmatrix} && (8)\\
&= \sum_{i,r} \sum_{i',r',o'} \beta \left(\begin{matrix} I \\ i' \end{matrix} \;\middle|\; \begin{matrix} O \\ o \end{matrix}\right) \theta \left(\begin{matrix} O & R \\ o' & r' \end{matrix} \;\middle|\; \begin{matrix} R & I \\ r & i \end{matrix}\right) \omega \begin{pmatrix} I & R & O \\ i' & r' & o' \end{pmatrix} && (9)\\
&= \sum_{i,r} \sum_{i',r',o'} \sigma \left(\begin{matrix} I & R & O \\ i' & r' & o' \end{matrix} \;\middle|\; \begin{matrix} I & R & O \\ i & r & o \end{matrix}\right) \omega \begin{pmatrix} I & R & O \\ i' & r' & o' \end{pmatrix} && (10)\\
&= \sum_{i,r} \omega \begin{pmatrix} I & R & O \\ i & r & o \end{pmatrix} && (11)\\
&= \omega|_{\mathcal{O}_\beta} \begin{pmatrix} O \\ o \end{pmatrix} && (12)
\end{aligned}
$$

□

Given a network without reciprocities $\mathcal{N}$, we attach to this network a measurement scenario as follows. The set of variables of this measurement scenario is the set of all the variables of the global process σ. To each node β in the network $\mathcal{N}$ corresponds a maximal context, $\mathcal{I}_\beta \cup \mathcal{O}_\beta$, which we denote by $\mathcal{U}_\beta$. They exhaust the set of maximal contexts of the measurement scenario.

Theorem 2. *Let $\mathcal{N}$ be a closed network without reciprocities with global process σ. Let ω be a stationary distribution of σ. Let $\mathcal{P}$ be the correspondence that –for each node β in the network $\mathcal{N}$– assigns the distribution δ_β to the maximal context $\mathcal{U}_\beta$. Then $\mathcal{P}$ is a no-signalling empirical model.*

Proof. It is sufficient to prove that given two nodes α and β of the network we have

$$\delta_\alpha|_{\mathcal{U}_\alpha \cap \mathcal{U}_\beta} = \delta_\beta|_{\mathcal{U}_\alpha \cap \mathcal{U}_\beta}$$

Since the network has no reciprocities, we can assume without loss of generality that all the arrows in $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$ go from α to β, that is, the variables in $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$ are output variables of α and at the same time they are input variables of β. This means that

$$\mathcal{U}_\alpha \cap \mathcal{U}_\beta = \mathcal{O}_\alpha \cap \mathcal{I}_\beta$$

By Theorem 1 we have

$$\begin{aligned}
\delta_\alpha|_{\mathcal{U}_\alpha \cap \mathcal{U}_\beta} &= \delta_\alpha|_{\mathcal{O}_\alpha}|_{\mathcal{U}_\alpha \cap \mathcal{U}_\beta} && (13)\\
&= \omega|_{\mathcal{O}_\alpha}|_{\mathcal{U}_\alpha \cap \mathcal{U}_\beta} && (14)\\
&= \omega|_{\mathcal{U}_\alpha \cap \mathcal{U}_\beta} && (15)\\
&= \omega|_{\mathcal{I}_\beta}|_{\mathcal{U}_\alpha \cap \mathcal{U}_\beta} && (16)\\
&= \delta_\beta|_{\mathcal{I}_\beta}|_{\mathcal{U}_\alpha \cap \mathcal{U}_\beta} && (17)\\
&= \delta_\beta|_{\mathcal{U}_\alpha \cap \mathcal{U}_\beta} && (18)
\end{aligned}$$

□

5 Contextuality in networks of stochastic processes

Theorem 2 establishes that any stationary distribution in a closed network of stochastic processes without reciprocities gives rise to an empirical model. To determine whether such model is contextual or not requires further conditions. The main result in [10] indicates that an empirical model will be non-contextual if the simplicial complex over its variables is *regular*.[3] This implies, in our setting, that a necessary condition for the extendability of the model is that the probability distributions over *all* the combinations of variables in the empirical model must be consistent. But in the network $\mathcal{N}$, supporting the empirical model, the requirement of consistency is imposed only on the internal distributions of the nodes and on those corresponding to connections among them. Furthermore, since the stochastic processes at the nodes are independent of each other, the distributions that matter for the definition of the empirical model have a natural local nature.[4] Example 1 illustrates that the focus

[3] A simplicial complex S is *regular* iff it belongs to the smallest class of simplicial complexes $\mathcal{S}$ such that: (a) includes the class of all the proper subsets of vertices of cardinality n, for every possible n; and (b) given $T \in \mathcal{S}$, $\hat{T} \in T$ and a set of vertices A that does not belong to T, if we add all the subsets of $\hat{T} \cup A$ to T, we obtain another simplicial complex in $\mathcal{S}$ [10](def. 2.1).

[4] A similar argument for the origin of contextuality in a combinatorial representation of quantum systems is presented in [3].

on only some joint distributions on the set of variables (instead of over all possible combinations) can be the source of contextuality.

We will in fact show that distributions like those of Example 1 can be commonplace in our framework. This indicates that the empirical models supported by networks of stochastic processes are not extendable in general. That is, in a very natural sense we can say that contextuality arises as a generic property of our empirical models.

Let α, β and γ be three open stochastic processes, each one with just one pair consisting of an input and an output variable, both assumed to be boolean. That is, taking values in $\{T, F\}$. The corresponding matrices are:

$$\alpha\begin{pmatrix} X' & \bigg| & Y \\ x' & & y \end{pmatrix} = \begin{cases} 1 & \text{if } \; y = \neg x' \\ 0 & \text{otherwise} \end{cases}$$

$$\beta\begin{pmatrix} Y' & \bigg| & Z \\ y' & & z \end{pmatrix} = \begin{cases} 1 & \text{if } \; z = y' \\ 0 & \text{otherwise} \end{cases}$$

$$\gamma\begin{pmatrix} Z' & \bigg| & X \\ z' & & x \end{pmatrix} = \begin{cases} 1 & \text{if } \; x = z' \\ 0 & \text{otherwise} \end{cases}$$

Now we compose these processes by connecting the output variable Y of α with the input variable Y' of β, as well as Z of β with Z' of γ and X of γ to X' of α. The result is a closed network as shown in Figure 8.

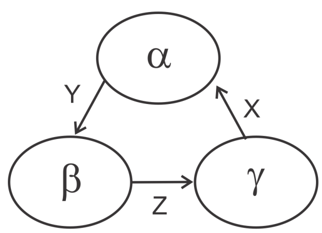

Figure 8: Graphical representation of the example.

The global process σ attached to this network is given by

$$\sigma\begin{pmatrix} X & Y & Z & \bigg| & X & Y & Z \\ x' & y' & z' & & x & y & z \end{pmatrix} = \alpha\begin{pmatrix} X & \bigg| & Y \\ x' & & y \end{pmatrix} \;\; \beta\begin{pmatrix} Y & \bigg| & Z \\ y' & & z \end{pmatrix} \;\; \gamma\begin{pmatrix} Z & \bigg| & X \\ z' & & x \end{pmatrix}$$

that is,

$$\sigma\begin{pmatrix} X & Y & Z & \bigg| & X & Y & Z \\ x' & y' & z' & & x & y & z \end{pmatrix} = \begin{cases} 1 & \text{if } \; (y = \neg x') \wedge (z = y') \wedge (x = z') \\ 0 & \text{otherwise} \end{cases}$$

The state space of the network is given by the eight possible values of the joint variable (X, Y, Z). The dynamics determined by σ can follow two cycles, one of which goes through the states (x, y, z) such that $(x = y) \wedge (z = \neg x)$ and the other through those such that $(x = z) \wedge (y = \neg x)$.

The process σ yields an infinite family of stationary distributions. Let us focus on the following one:

$$\omega \begin{pmatrix} X & Y & Z \\ x & y & z \end{pmatrix} = \begin{cases} 0 & \text{if } \ (x = y) \wedge (z = \neg x) \\ 1/6 & \text{otherwise} \end{cases}$$

We can easily check that ω is indeed stationary.[5] Note that the marginals $\omega|_X$, $\omega|_Y$ and $\omega|_Z$ are uniform distributions, that is

$$\omega|_X \begin{pmatrix} X \\ x \end{pmatrix} = 1/2 \ \text{ for all } \ x$$

$$\omega|_Y \begin{pmatrix} Y \\ y \end{pmatrix} = 1/2 \ \text{ for all } \ y$$

$$\omega|_Z \begin{pmatrix} Z \\ z \end{pmatrix} = 1/2 \ \text{ for all } \ z$$

Then, according to the construction leading to Theorem 2, the empirical model attached to the network is given by:[6]

$$\delta_\alpha \begin{pmatrix} X & Y \\ x & y \end{pmatrix} = \alpha \left(\begin{matrix} X \\ x \end{matrix} \,\middle|\, \begin{matrix} Y \\ y \end{matrix} \right) \omega|_X \begin{pmatrix} X \\ x \end{pmatrix} = \begin{cases} 1/2 & \text{if } \ x = \neg y \\ 0 & \text{otherwise} \end{cases}$$

$$\delta_\beta \begin{pmatrix} Y & Z \\ y & z \end{pmatrix} = \beta \left(\begin{matrix} Y \\ y \end{matrix} \,\middle|\, \begin{matrix} Z \\ z \end{matrix} \right) \omega|_Y \begin{pmatrix} Y \\ y \end{pmatrix} = \begin{cases} 1/2 & \text{if } \ y = z \\ 0 & \text{otherwise} \end{cases}$$

[5] That is, ω is a fixed point of the mapping given by

$$\sum_{x',y',z'} \sigma \left(\begin{matrix} X & Y & Z \\ x' & y' & z' \end{matrix} \,\middle|\, \begin{matrix} X & Y & Z \\ x & y & z \end{matrix} \right) \omega \begin{pmatrix} X & Y & Z \\ x' & y' & z' \end{pmatrix}$$

[6] Note that, in general, the distributions δ_α,δ_β and δ_γ do not coincide with $\omega|_{X,Y}$, $\omega|_{Y,Z}$ and $\omega|_{Z,X}$ respectively. In this case we have, for instance that (analogously for (Y, Z) and (Z, X))

$$\omega|_{X,Y} \begin{pmatrix} X & Y \\ x & y \end{pmatrix} = \begin{cases} 1/6 & \text{if } \ x = y \\ 1/3 & \text{otherwise} \end{cases}$$

$$\delta_\gamma \begin{pmatrix} Z & X \\ z & x \end{pmatrix} = \quad \gamma \left(\begin{array}{c|c} Z & X \\ z & x \end{array}\right) \quad \omega|_Z \begin{pmatrix} Z \\ z \end{pmatrix} = \begin{cases} 1/2 & \text{if } \ z = x \\ 0 & \text{otherwise} \end{cases}$$

These distributions over (X, Y), (Y, Z) and (Z, X) are the same as those of Example 1.

6 Comparison to quantum contextuality

The previous example suggests that the networks of processes lead to empirical models that exhibit a strong form of contextuality. As a way of measure this contextuality, in this section we consider a familiar quantum scenario –the Clauser-Horne-Shimony-Holt (CHSH) scenario– and construct a simple network of processes based on this scenario. We show that the empirical model arising from this network is a Popescu-Rohrlich (PR) box, which is well known to be a no-signalling model that maximally violates the Tsirelson's bound. This shows that the contextuality allowed by our construction is not restricted to quantum contextuality, but it corresponds to the contextuality of generalised probabilistic theories (see for example [5]).

The CHSH scenario consists of four observables A_1, A_2, B_1 and B_2 toghether with four contexts (A_1, B_1), (A_2, B_1), (A_1, B_2) and (A_2, B_2). The set of outcomes of each variable is $\{0, 1\}$. The PR box is defined by

$$\begin{aligned} P(A_1 = 0 \wedge B_1 = 1) = P(A_1 = 1 \wedge B_1 = 0) = 1/2 \\ P(A_1 = 0 \wedge B_2 = 0) = P(A_1 = 1 \wedge B_2 = 1) = 1/2 \\ P(A_2 = 0 \wedge B_1 = 0) = P(A_2 = 1 \wedge B_1 = 1) = 1/2 \\ P(A_2 = 0 \wedge B_2 = 0) = P(A_2 = 1 \wedge B_2 = 1) = 1/2 \end{aligned}$$

This empirical model can be realized within our framework as follows. We attach a stochastic process α_{ij} to each context (A_i, B_j). The set of outcomes of each variable is $\{0, 1\}$. The matrices defining these stochastic processes are given by

$$\alpha_{11} \left(\begin{array}{c|c} A_1' & B_1 \\ a_1' & b_1 \end{array}\right) = \begin{cases} 1 & \text{if } \ a_1' = \neg b_1 \\ 0 & \text{otherwise} \end{cases}$$

$$\alpha_{21} \left(\begin{array}{c|c} B_1' & A_2 \\ b_1' & a_2 \end{array}\right) = \begin{cases} 1 & \text{if } \ b_1 = a_2' \\ 0 & \text{otherwise} \end{cases}$$

$$\alpha_{22} \left(\begin{array}{c|c} A_2' & B_2 \\ a_2' & b_2 \end{array}\right) = \begin{cases} 1 & \text{if } \ a_2' = b_2 \\ 0 & \text{otherwise} \end{cases}$$

$$\alpha_{12}\begin{pmatrix} B_2' & \bigg| & A_1 \\ b_2' & & a_1 \end{pmatrix} = \begin{cases} 1 & \text{if } b_2' = \neg a_1 \\ 0 & \text{otherwise} \end{cases}$$

As in the example of the previous section, we compose these processes to build a closed network. Namely, we connect the variable A_i' with A_i and the variable B_i' with B_i, for $i = 1, 2$. The resulting global process σ is given by

$$\sigma\begin{pmatrix} A_1 & B_1 & A_2 & B_2 & \bigg| & A_1 & B_1 & A_2 & B_2 \\ a_1' & b_1' & a_2' & b_2' & & a_1 & b_1 & a_2 & b_2 \end{pmatrix} =$$

$$\alpha_{11}\begin{pmatrix} A_1 & \bigg| & B_1 \\ a_1' & & b_1 \end{pmatrix} \quad \alpha_{21}\begin{pmatrix} B_1 & \bigg| & A_2 \\ b_1' & & a_2 \end{pmatrix} \quad \alpha_{22}\begin{pmatrix} A_2 & \bigg| & B_2 \\ a_2' & & b_2 \end{pmatrix} \alpha_{12}\begin{pmatrix} B_2 & \bigg| & A_1 \\ b_2' & & a_1 \end{pmatrix} =$$

$$= \begin{cases} 1 & \text{if } (a_1' = \neg b_1) \wedge (b_1' = a_2) \wedge (a_2' = b_2) \wedge (b_2' = a_1) \\ 0 & \text{otherwise} \end{cases}$$

Since this is a permutation matrix, then the uniform distribution on the state space of the joint variable (A_1, B_1, A_2, B_2) is stationary, so we define

$$\omega\begin{pmatrix} A_1 & B_1 & A_2 & B_2 \\ a_1 & b_1 & a_2 & b_2 \end{pmatrix} = 1/16$$

for all a_1, b_1, a_2, b_2. The marginals $\omega|_{A_1}$, $\omega|_{B_1}$, $\omega|_{A_2}$ and $\omega|_{B_2}$ are uniform distributions. Then the empirical model attached to the network is given by

$$\delta_{\alpha_{11}}\begin{pmatrix} A_1 & B_1 \\ a_1 & b_1 \end{pmatrix} = \alpha_{11}\begin{pmatrix} A_1 & \bigg| & B_1 \\ a_1 & & b_1 \end{pmatrix} \quad \omega|_{A_1}\begin{pmatrix} A_1 \\ a_1 \end{pmatrix} = \begin{cases} 1/2 & \text{if } a_1 = \neg b_1 \\ 0 & \text{otherwise} \end{cases}$$

$$\delta_{\alpha_{21}}\begin{pmatrix} B_1 & A_2 \\ b_1 & a_2 \end{pmatrix} = \alpha_{21}\begin{pmatrix} B_1 & \bigg| & A_2 \\ b_1 & & a_2 \end{pmatrix} \quad \omega|_{B_1}\begin{pmatrix} B_1 \\ b_1 \end{pmatrix} = \begin{cases} 1/2 & \text{if } b_1 = a_2 \\ 0 & \text{otherwise} \end{cases}$$

$$\delta_{\alpha_{22}}\begin{pmatrix} A_2 & B_2 \\ a_2 & b_2 \end{pmatrix} = \alpha_{22}\begin{pmatrix} A_2 & \bigg| & B_2 \\ a_2 & & b_2 \end{pmatrix} \quad \omega|_{A_2}\begin{pmatrix} A_2 \\ a_2 \end{pmatrix} = \begin{cases} 1/2 & \text{if } a_2 = b_2 \\ 0 & \text{otherwise} \end{cases}$$

$$\delta_{\alpha_{12}}\begin{pmatrix} B_2 & A_1 \\ b_2 & a_1 \end{pmatrix} = \alpha_{12}\begin{pmatrix} B_2 & \bigg| & A_1 \\ b_2 & & a_1 \end{pmatrix} \quad \omega|_{B_2}\begin{pmatrix} B_2 \\ b_2 \end{pmatrix} = \begin{cases} 1/2 & \text{if } b_2 = a_1 \\ 0 & \text{otherwise} \end{cases}$$

which coincides with the PR box.

7 Conclusions

While the result presented in the previous sections is just an instance, we can hint at its generality. Networks are combinations of arbitrary stochastic processes and the variables of non directly connected nodes can be distributed independently.

The previously known sources of contextuality were of quantum nature. In this paper we introduced a classical model, namely that of a network of stochastic processes. The ensuing structure shares with quantum systems the property of contextuality but without resorting to their physical features.

References

[1] Abramsky, Samson, Rui Soares Barbosa, Kohei Kishida, Raymond Lal and Shane Mansfield. "Contextuality, Cohomology and Paradox". arXiv preprint arXiv:1502.03097 (2015).

[2] Abramsky, Samson and Adam Brandenburger. "The Sheaf-Theoretic Structure of Non-Locality and Contextuality". New Journal of Physics 13.11 (2011): 113036.

[3] Acín, Antonio, Tobias Fritz, Anthony Leverrier and Ana B. Sainz. "A Combinatorial Approach to Nonlocality and Contextuality". Communications in Mathematical Physics 334 (2015): 533–628.

[4] Boole, George. An Investigation of the Laws of Thought: on Which are Founded the Mathematical Theories of Logic and Probabilities. Dover Publications, 1854.

[5] Cabello, Adán, Simone Severini, and Andreas Winter. "Graph-theoretic approach to quantum correlations." Physical review letters 112.4 (2014): 040401.

[6] Khrennikov, Andrei Y. "Ubiquitous quantum structure". Springer, 2014.

[7] Levin, David A., Yuval Peres and Elizabeth L. Wilmer. "Markov Chains and Mixing Time". American Mathematical Society, 2009.

[8] Markov, Igor L., and Yaoyun Shi. "Simulating Quantum Computation by Contracting Tensor Networks." SIAM Journal on Computing 38.3 (2008): 963-981.

[9] Pitowsky, Itamar. "George Boole's Conditions of Possible Experience and the Quantum Puzzle." British Journal for the Philosophy of Science (1994): 95-125.

[10] Vorobev, N. N., "Consistent Families of Measures and their Extensions", Theory of Probability and Applications, vol. 7, 1962, p. 147-163.

[11] Wood, Christopher J., Jacob D. Biamonte, and David G. Cory. "Tensor Networks and Graphical Calculus for Open Quantum Systems." Quantum Information and Computation 15.9 10 (2015): 759-811.