arXiv:2306.03869v3 [quant-ph] 29 Dec 2025

Connecting classical finite exchangeability to quantum theory and indistinguishability

Alessio Benavoli[a,b], Alessandro Facchini[c,d], Marco Zaffalon[c]

[a] *School of Computer Science and Statistics, Trinity College Dublin, Dublin, Ireland*
[b] *Trinity Quantum Alliance, Trinity College Dublin, Unit 16 Trinity Technology and Enterprise Centre, Dublin, Ireland*
[c] *SUPSI, DTI, Istituto Dalle Molle di Studi sull'Intelligenza Artificiale (USI-SUPSI), Lugano, Switzerland*
[d] *Management in Networked and Digital Societies (MINDS) Department, Kozminski University, \city {Warsaw} \country {Poland}*

Abstract

Exchangeability is a fundamental concept in probability theory and statistics. It allows to model situations where the order of observations does not matter. The classical de Finetti's theorem provides a representation of infinitely exchangeable sequences of random variables as mixtures of independent and identically distributed variables. The quantum de Finetti theorem extends this result to symmetric quantum states on tensor product Hilbert spaces. It is well known that both theorems do not hold for finitely exchangeable sequences. The aim of this work is to investigate two lesser-known representation theorems, which were developed in classical probability theory to extend de Finetti's theorem to finitely exchangeable sequences by using quasi-probabilities and quasi-expectations. With the aid of these theorems, we illustrate how a de Finetti-like representation theorem for finitely exchangeable sequences can be formulated through a mathematical representation which is formally equivalent to quantum theory (with boson-symmetric density matrices). We then show a promising application of this connection to the challenge of defining entanglement for indistinguishable bosons.

Keywords: Exchangeability, de Finetti, quantum de Finetti, bosons

1. Introduction

Let $(t_1, t_2, \ldots, t_r)$ be a sequence of r random variables, each variable taking values in a finite possibility space Ω, and let $P(t_1, t_2, \ldots, t_r)$ be their joint

probability distribution. When $P(t_1, t_2, \ldots, t_r)$ is symmetric to permutations of the labels of the variables t_i, the sequence is called finitely exchangeable with respect to P. The classical de Finetti theorem states that, whenever the considered sequence is finitely exchangeable with respect to P and for every $s > r$ there is a finitely exchangeable sequence $(t_1, t_2, \ldots, t_s)$ with respect to a joint probability P' such that P coincides with the marginal of P' on $(t_1, t_2, \ldots, t_r)$, then $P(t_1, t_2, \ldots, t_r)$ can be represented as:

$$\begin{aligned} P(t_1, t_2, \ldots, t_r) &= E[P(t_1)P(t_2)\cdots P(t_r)] \\ &:= \int P(t_1)P(t_2)\cdots P(t_r)dq(P), \end{aligned} \tag{1}$$

with q a probability measure on the marginal probability distribution $P(t)$ (of one variable).

One well known issue with this theorem is that it necessarily requires an infinite sequence extending $(t_1, t_2, \ldots, t_r)$ for which exchangeability holds for each segment longer than r. Without such requirement, and thus by sticking to a finitely exchangeable sequence only, the representation of the operator E in Eq. (1) as a probabilistic mixture of identical product distributions fails in general. However, if the finitely exchangeable sequence of r variables is part of a finite but large enough finitely excheangable sequence of $s >> r$ variables, the joint distribution of the initial r random variables can be well-approximately represented in a de Finetti-like manner. Much of the work on finite exchangeability in classical probability has actually focused on deriving analytical bounds for the error of this approximation [1, 2].

A quantum analogue of de Finetti's theorem has been derived in [3–8]. Widely used in quantum information theory, it states the following. Let $\rho^{(r)}$ be a density matrix on the r-tensor product Hilbert space $\mathcal{H}^{\otimes r}$ (if $\dim(\mathcal{H}) = d$, then $\rho^{(r)}$ is a $d^r \times d^r$ complex Hermitian positive semi-definite matrix with trace one). Whenever $\rho^{(r)}$ is symmetric, that is it is such that $\Pi_{sym}\rho^{(r)}\Pi_{sym}^{\dagger} = \rho^{(r)}$ with the matrix Π_{sym} being the so-called 'symmetriser', and there is a symmetric density matrix $\rho^{(s)}$ on $\mathcal{H}^{\otimes s}$ such that $\rho^{(r)} = Tr_{s-r}(\rho^{(s)})$ for every $s > r$, where Tr_{s-r} denotes the partial-trace (which performs a marginalisation operation over the subset of the $s - r$ subsystems), then it holds that

$$\rho^r = \int (\sigma^{\otimes s})dq(\sigma), \tag{2}$$

with σ a density matrix in $\mathcal{H}$, $\sigma^{\otimes s}$ denotes the tensor product of σ with itself s-times, and q a probability measure on the set of all density matrices in $\mathcal{H}$. The research on analytical bounds has also been extended to the quantum setting [5, 6] with applications to quantum cryptography [9, 10], emergence of classical reality [11] and, more recently, the black hole information puzzle [12].

In this paper we investigate two other representation theorems that extend the classical de Finetti's theorem so as to cover also the case of finitely exchangeable sequences. Both modify the definition of operator E in Eq. (1). The first solution keeps the same representation as in the original theorem, except that q is a normalised signed-measure (a quasi-probability) [13–15] (which also holds true in the quantum de-Finetti case [6, 16, 17]). The second solution instead employs a *quasi*-expectation operator applied to certain polynomial functions [14, 15, 18–20]. Our objective is to illustrate, with the aid of the above results, how de Finetti-like representation theorems for finitely exchangeable systems can be formulated using a mathematical representation which is formally equivalent to quantum theory (QT for short). This objective is achieved in two steps. First, we will use polynomials of complex conjugate variables to re-derive the extension of the classical de Finetti's theorem based on *quasi*-expectation operators but expressed with the same mathematical objects as in QT. Second, we will demonstrate that such *quasi*-expectation operators are equivalent to boson-symmetric density matrices. In doing so, we will be able to simulate finite exchangeable sequences using quantum experiments with bosons. In summary, we will show that a de Finetti-like representation of finite exchangeability leads to a representation theorem which encompasses both classical and quantum theories as special cases. We will then discuss how such representation is related to previous results [21–23] and to other formulations of QT involving quasi-probability densities and expectations. Finally, we will use this representation to provide fresh insights into the challenge of defining entanglement for indistinguishable bosons. There are at least five main different definitions of entanglement for indistinguishable bosons in literature. Recently, [24] attempted to systematise these different definitions by proposing a set of properties that should be used to characterise entanglement for indistinguishable particles. Through our connection between QT and classical finitely exchangeable probabilities, we will show that the definitions of two properties given in [24] lead to clearly interpretable contradictions when applied to classical finitely exchangeable probabilities. For instance, one of these properties, *local operators*, would im-

ply that the uniform distribution is finitely but not infinitely exchangeable, which is clearly false. We then use this analysis to reformulate this property in a way that avoids this issue.
To provide a clearer and more practical understanding of the different representation theorems, in what follows we use a dice rolling experiment as a case study. These examples will later help to clarify the application of exchangeability to the problem of defining entanglement for indistinguishable bosons.

2. The classical de Finetti's representation theorem

Consider a dice whose possibility space is given by its six faces:

$$\Omega := \Big\{ \; , \; , \; , \; , \; , \; \Big\}.$$

Let us assume that the upward-facing side of the dice is the one that is facing the reader. By $t_1, t_2, \ldots, t_r$ we denote the results of r-rolls of the dice. For example, for $r = 3$, we can have the three outcomes:

$$\underset{t_1}{\quad} \quad \underset{t_2}{\quad} \quad \underset{t_3}{\quad}$$

that is $t_1 = 2$, $t_2 = 3$, and $t_3 = 3$.

Definition 1. *Let $t_1, t_2, \ldots, t_r$ a sequence of variables and P a joint probability. The sequence is said to be finitely exchangeable with respect to P, if*

$$P(t_1, t_2, \ldots, t_r) = P(t_{\pi(1)}, t_{\pi(2)}, \ldots, t_{\pi(r)}),$$

for any permutation π of the index sequence $(1, 2, \ldots, r)$.

Considering the previous example. To say that the three rolls are finitely exchangeable means that

$$P\left(\; \right) = P\left(\; \right) = P\left(\; \right).$$

Definition 2. *The sequence is called infinitely exchangeable (with respect to P) if it is finitely exchangeable (with respect to P) and if for any $s > r$, there is a distribution P' with respect to which $t_1, t_2, \ldots, t_s$ is finitely exchangeable and such that*

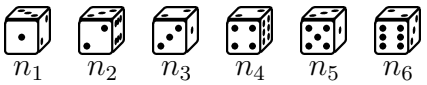
$$P(t_1, \ldots, t_r) = \sum_{t_{r+1}, \ldots, t_s} P'(t_1, \ldots, t_r, t_{r+1}, \ldots, t_s).$$

Stated otherwise, an infinitely exchangeable sequence can be understood as a subsequence

$$\ldots, t_{k-1}, t_k, \overbrace{t_{k+1}, t_{k+2}, \ldots, t_{k+n-1}, t_{k+r}}^{\text{subsequence of length } r}, t_{k+r+1}, \ldots$$

of an infinite sequence of random variables whose joint distribution is invariant under permutations, meaning that their order is irrelevant.

Given r exchangeable rolls $t_1, t_2, \ldots, t_r$, their output is fully characterised by the counts:



where n_i denotes the number of times the dice landed on i-th face in the r-rolls. Based on this fact, de Finetti proved his famous *representation theorem.*[1]

Proposition 1 ([25])**.** *Let Ω be the possibility spaces given by the six faces of a dice, and let $(t_1, \ldots, t_r)$ be a sequence of random variables that is infinitely exchangeable with probability measure P. Then there exists a distribution function q such that*

$$P(t_1, \ldots, t_r) = \int_\Theta \theta_1^{n_1} \theta_2^{n_2} \cdots \theta_5^{n_5} \theta_6^{n_6} dq(\boldsymbol{\theta}), \tag{3}$$

where $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_6]^\top$ are the probabilities of the corresponding faces and whose values belong to the possibility space:

$$\Theta = \left\{ \boldsymbol{\theta} \in \mathbb{R}^6 : \theta_i \geq 0, \ \sum_{i=1}^{6} \theta_i = 1 \right\},$$

[1]De Finetti actually proved his theorem for the binary case, a coin. His result can easily be extended to the dice.

and $n_i = \sum_{j=1}^{r} \delta_i(t_j)$*, where* $\delta_a(b)$ *is the Kronecker delta*[2]*, that is* n_i *is the number of occurrences of the i-th face in the sequence* $t_1, \ldots, t_r$ *(this means that* $\sum_{i=1}^{6} n_i = r$*).*

In the context of Bayesian statistics, the theorem states that any infinitely exchangeable sequence can be seen as generated by first selecting a probability distribution $\boldsymbol{\theta}$ from some prior q (probability distribution on the space of probability distributions) and then letting the t_i to be i.i.d. with common distribution $\boldsymbol{\theta}$. For a selected $\boldsymbol{\theta}$, this for instance means that $P(t_1 = 2, t_2 = 3, t_3 = 3) = \theta_2\theta_3^2$ (the product comes from the independence assumption).

As mentioned in Section 1, one known issue with the above theorem is that the assumption of an infinite exchangeable sequence is necessary for the specific representation to hold [26, 27]. As a counter-example, consider:

$$\begin{aligned} &P(t_1 = i, t_2 = i) = 0, \\ &P(t_1 = i, t_2 = j) = P(t_1 = j, t_2 = i) = \frac{1}{30} \end{aligned} \tag{4}$$

for all $i \neq j = 1, 2, \ldots, 6$. The variables t_1, t_2 are finitely exchangeable. Assume that a representation like (3) holds. Then

$$0 = P(t_1 = i, t_2 = i) = \int_{\Theta} \theta_i^2 dq(\boldsymbol{\theta})$$

for all i, would imply that q puts mass 1 at all the $\theta_i = 0$, which is impossible. This means that, although the P in (4) is a valid probability, we cannot express it as (3). In other words, exchangeability is not equivalent to independence.

Examples of situations that are truly finitely exchangeable were already discussed by de Finetti in [26]. They can be illustrated by the act of selecting items from an urn without replacement.

Example 1. *Suppose we have an urn containing one black (B) ball and one white (W) ball. Draw out balls, one at a time and without replacement, and note the colour. Then we have that*

[2]The Kronecker delta is defined as follows $\delta_a(b) = 1$ if $a = b$ and zero otherwise.

$$
\begin{aligned}
&P(t_1 = B, t_2 = B) = 0,\\
&P(t_1 = B, t_2 = W) = P(t_1 = W, t_2 = B) = 0.5,\\
&P(t_1 = W, t_2 = W) = 0,
\end{aligned}
$$

which is finitely but not infinitely exchangeable.

In QT, we can rule out this type of urn-model dependence by choosing the direction of the measurement randomly and independently for each particle. This is done in Bell's experiments. This comment is relevant for the relationship between exchangeability and entanglement, which we discuss in Section 5.

Kendall [28] proved that every finite system of exchangeable events is indeed equivalent to a random sampling scheme without replacement, where the number of items in the sampling has an arbitrary distribution.

While it is evident that de Finetti's theorem can be invalid for finite exchangeable sequences, for a large enough integer $s > r$, the distribution of the initial r random variables of an exchangeable vector of length s can be approximately represented as in Eq. (3). As previously mentioned, much of the work on finite exchangeability has focused on deriving analytical bounds for the error of this approximation [1, 2]. To clarify the meaning of the term 'bound' in this context, we use the dice-rolling example to discuss the general problem of bounding a linear function of event probabilities, where the underlying probability measure P_r is a probability distribution of a finitely exchangeable sequence $t_1, \ldots, t_r$. In the rest of the manuscript, the subscript r in P_r will be used to indicate the length of the sequence $t_1, \ldots, t_r$.

Example 2. *Consider $r = 2$ (two rolls) and assume, for example, we aim to find a lower bound for the following quantity*

$$
P_2\left(\square\,\square\right) - P_2\left(\square\,\square\right) + P_2\left(\square\,\square\right), \tag{5}
$$

where P_2 is a probability distribution of the finitely exchangeable sequence t_1, t_2 (two rolls). In other words, among all probability measures P_2 we aim to find the one which minimises (5)*. We will focus only on the lower bound (minimisation) as an example.*

Therefore, our goal is to compute:

$$v_2 = \min_{P_2} \left(P_2 \left(\boxdot \boxdot \right) - P_2 \left(\boxdot \boxdot \right) + P_2 \left(\boxdot \boxdot \right) \right).$$

It is not difficult to see that the minimum is $v_2 = -0.5$ *and the finitely exchangeable* P_2 *which achieves it is given by setting* $P_2 \left(\boxdot \boxdot \right) = P_2 \left(\boxdot \boxdot \right) = 0.5$ *and zero otherwise.*

Now assume that these two rolls are part of a finitely exchangeable sequence of three rolls, and let P_3 *be the relative finitely exchangeable probability. In this case, we aim to solve the problem*

$$v_3 = \min_{P_3} \left(P_3 \left(\boxdot \boxdot \right) - P_3 \left(\boxdot \boxdot \right) + P_3 \left(\boxdot \boxdot \right) \right),$$

where $P_3(t_1, t_2)$ *is computed from* $P_3(t_1, t_2, t_3)$ *via marginalisation, that is* $P_3(t_1 = a, t_2 = b) = \sum_{i=1}^{6} P_3(t_1 = a, t_2 = b, t_3 = i)$ *for every* $a, b \in \{1, 2, \ldots, 6\}$*. Note that, since* P_3 *is the probability distribution of the exchangeable sequence* t_1, t_2, t_3 *then* $P_3(t_1 = a, t_2 = b, t_3 = i) = P_3(t_1 = b, t_2 = a, t_3 = i) = P_3(t_1 = i, t_2 = b, t_3 = a)$ *and so on for all the* $3!$ *permutations. In this case, computing* v_3 *is more challenging because exchangeability imposes many constraints on* P_3*, such as:*

$$P_3 \left(\boxdot \boxdot \boxdot \right) = P_3 \left(\boxdot \boxdot \boxdot \right) = P_3 \left(\boxdot \boxdot \boxdot \right).$$

In particular, this means that, in this case, to compute v_3 *we cannot set* $P_3 \left(\boxdot \boxdot \right)$ *equal to zero. Indeed, since its value depends on the value of* $P_3 \left(\boxdot \boxdot \boxdot \right)$ *via marginalisation, then if* $P_3 \left(\boxdot \boxdot \right)$ *were zero, than also* $P_3 \left(\boxdot \boxdot \right)$ *would be zero. We will later show that in this case the solution of the minimisation problem is* $v_3 \approx -0.167$*.*

We can also show that when the two rolls are part of an infinitely exchangeable sequence of rolls it holds that

$$\begin{aligned} v_\infty &= \inf_{P_\infty} \left(P_\infty \left(\boxdot \boxdot \right) - P_\infty \left(\boxdot \boxdot \right) + P_\infty \left(\boxdot \boxdot \right) \right) \\ &= 0. \end{aligned}$$

In summary, we can compute lower bounds of $P_n\left(\boxdot\,\boxdot\right) - P_n\left(\boxdot\,\boxdot\right) + P_n\left(\boxdot\,\boxdot\right)$ *for any* $n \geq 2$ *and these bounds satisfy the inequalities*

$$-0.5 = v_2 < v_3 < \cdots < v_\infty = 0.$$

This means that as n *increases,* v_n *approaches the value computed using the probability* P_∞ *of an infinitely exchangeable sequence. In Section 4, we will discuss a general numerical algorithm to compute a lower bound of any linear function of event probabilities.*

It is important to emphasize that the inequalities $v_2 < v_3 < \cdots < v_\infty$ *do not hold for every linear function of event probabilities. For example, if we consider the lower bound* $w_n = \min_{P_n}\left(P_n\left(\boxdot\,\boxdot\right) + P_n\left(\boxdot\,\boxdot\right)\right)$ *then* $w_2 = w_3 = \cdots = w_\infty = 0$. *The linear function of event probabilities in* (5) *acts as a "witness". Specifically, it can be used (by solving the minimisation problem) to distinguish the probability distribution of a finitely exchangeable sequence from that of an infinitely exchangeable one.*

Why are these bounds important? First, statistical problems and experiments involve exchangeable sequences that are always finite. It is not possible to perform an infinite number of trials. Second, de Finetti's theorem provides a useful and simple representation of the mechanism generating the data (which, in turn, can be inverted to estimate $\boldsymbol{\theta}$ from the counts $n_1, \ldots, n_6$ via Bayes' rule). Consequently, practitioners are still inclined to use de Finetti's theorem as an approximation for the mechanism generating the finite exchangeable sequence and use these bounds to assess the error of the approximation for a particular inference.[3] These facts motivate the search for representation theorems that apply to both the finitely and the infinitely exchangeable case.

3. First representation

A fact about finite exchangeability that has not received equal attention is that a de Finetti-like representation actually holds in the finite case if we

[3]In the previous example, the inference is the computation of $P_n\left(\boxdot\,\boxdot\right) - P_n\left(\boxdot\,\boxdot\right) + P_n\left(\boxdot\,\boxdot\right)$.

drop the non-negativity condition for q. The result was already known to de Finetti [29]. Explicitly derived in [13, 14] for some special cases, it was generalised by [15].

Proposition 2 ([15]). *Given a sequence of finitely exchangeable variables $t_1, t_2, \ldots, t_r$, there exists a signed measure ν, satisfying $\int_\Theta d\nu(\boldsymbol{\theta}) = 1$, such that:*

$$P(t_1, \ldots, t_r) = \int_\Theta \theta_1^{n_1}\theta_2^{n_2}\cdots\theta_5^{n_5}\theta_6^{n_6} d\nu(\boldsymbol{\theta}). \tag{6}$$

Signed measure means that ν can have negative values. Note that, in Eq. (6), the only valid signed-measures ν are those which give rise to well-defined (non-negative and normalised) probabilities $P(t_1, t_2, \ldots, t_r)$.

By comparing the two representations (3) and (6), it becomes apparent that there exists a significant difference that may not be immediately noticeable.

Any probability measure q originates, via Eq. (3), into a valid (non-negative and normalised) probability $P(t_1, \ldots, t_r)$, that is for any $[n_1, \ldots, n_6]$ such that $\sum_{i=1}^6 n_i = r$, we have that:

$$\int_\Theta \theta_1^{n_1}\theta_2^{n_2}\cdots\theta_5^{n_5}\theta_6^{n_6} dq(\boldsymbol{\theta}) \geq 0, \tag{7}$$

$$\sum_{[n_1', \ldots, n_6']:\sum_{i=1}^6 n_i' = r} b_{n_1' n_2' \ldots n_6'} \int_\Theta \theta_1^{n_1'}\theta_2^{n_2'}\cdots\theta_5^{n_5'}\theta_6^{n_6'} dq(\boldsymbol{\theta}) = 1, \tag{8}$$

where $b_{n_1 n_2 \ldots n_6} = r!/(n_1!n_2!n_3!n_4!n_5!n_6!)$ is the multinomial coefficient. Note that, (7)–(8) imply that $P(t_1, \ldots, t_r) \geq 0$ and $\sum_{t_1,\ldots,t_r} P(t_1, \ldots, t_r) = 1$. Indeed, consider $P(t_1, \ldots, t_r)$ and let $n_i = \sum_{j=1}^r \delta_i(t_j)$ the number of occurrences of the i-th face in the sequence $t_1, \ldots, t_r$. From (3), we have that $P(t_1, \ldots, t_r) = \int_\Theta \theta_1^{n_1}\theta_2^{n_2}\cdots\theta_5^{n_5}\theta_6^{n_6} dq(\boldsymbol{\theta})$ (which is nonnegative). This equality holds for any permutation $t_{\pi(1)}, t_{\pi(2)}, \ldots, t_{\pi(r)}$, as they all have the same amount of counts $n_1, n_2, \ldots, n_6$. There are $r!/(n_1!n_2!n_3!n_4!n_5!n_6!)$ of these permutations. Therefore, we have that

$$\sum_{t_1,\ldots,t_r} P(t_1, \ldots, t_r) = \sum_{\mathbf{n}=[n_1', \ldots, n_6']:\sum_i n_i' = r} b_{n_1' n_2' \ldots n_6'} \int_\Theta \theta_1^{n_1'}\theta_2^{n_2'}\cdots\theta_5^{n_5'}\theta_6^{n_6'} dq(\boldsymbol{\theta}) = 1$$

(normalised). Notice that not all signed measures ν that integrate to one define valid probabilities $P(t_1, \ldots, t_r)$ via Eq. (6). In some cases, these measures

may result in a negative value for $P(t_1, \ldots, t_r)$. This is why this representation is not used in practice. Similarly, this is why in QT we avoid using 'negative probabilities', despite the well-known fact that such probabilities could enable a hidden variable representation, at least mathematically.

For finite exchangeability, Jaynes' work [14] is interesting because it contains a method to define valid P via Eq. (6) using Bernstein (or Legendre) polynomials. A more recent approach is discussed in the next section.

4. Second representation

Reasoning about finitely exchangeable variables t_i can be reduced to reasoning about polynomials of count vectors, known as Bernstein polynomials [14, 15, 18–20]. Working with this polynomial representation automatically guarantees that finite exchangeability is satisfied and that P are valid probabilities, without having to deal directly with the signed measure ν.

The *multivariate Bernstein polynomials* are polynomials of the form

$$\theta_1^{n_1}\theta_2^{n_2}\cdots\theta_5^{n_5}\,(1-\theta_1-\cdots-\theta_5)^{n_6} \;\text{ with }\; \sum_{i=1}^{6} n_i = r, \tag{9}$$

where r is the number of rolls. They satisfy many useful properties, such as forming a base for the linear space of all polynomials whose degree is at most r and forming a partition of unity. To simplify the notation, we will denote $1-\theta_1-\cdots-\theta_5$ as θ_6, so that $\sum_{i=1}^{6}\theta_i = 1$.

The fundamental concept behind the new representation theorem that focuses on Bernstein polynomials involves working directly with the linear operator L defined as

$$L(\ldots) := \int_\Theta (\ldots)\, d\nu(\boldsymbol{\theta}), \tag{10}$$

and determining the specific properties that L must adhere to in order to produce valid probabilities.

More precisely, let L be a *linear operator* from the space $\mathcal{B}_r$ of polynomials of degree at most r to $\mathbb{R}$. Our goal is to write any finitely exchangeable $P(t_1, t_2, \ldots, t_r)$ as $L(\theta_1^{n_1}\theta_2^{n_2}\cdots\theta_6^{n_6})$, where $\sum_{i=1}^{6} n_i = r$. In the aim of determining the appropriate properties of L, we first introduce the closed convex cone

$$\mathcal{B}_r^+ = \left\{ \sum_{\mathbf{n}=[n_1,\ldots,n_6]:\sum_i n_i = r} u_{\mathbf{n}}\,\theta_1^{n_1}\theta_2^{n_2}\cdots\theta_6^{n_6} : u_{\mathbf{n}} \in \mathbb{R}^+ \right\}, \tag{11}$$

which includes all non-negative linear combinations of Bernstein's monomials of degree r, as well as all constant polynomials with value $c \geq 0$. This latter property holds because

$$c = c(\theta_1 + \theta_2 + \cdots + \theta_6)^r = c \sum_{\mathbf{n}=[n_1,\ldots,n_6]:\sum_i n_i = r} b_{n_1 n_2 \ldots n_6} \theta_1^{n_1} \theta_2^{n_2} \cdots \theta_6^{n_6}, \tag{12}$$

and the coefficients $c\, b_{n_1 n_2 \ldots n_6}$ are positive. In the following, we will use g to denote generic polynomials of the vector of variables $\boldsymbol{\theta}$ and c to represent the constant polynomial with value c, meaning $g(\boldsymbol{\theta}) = c$ for all $\boldsymbol{\theta}$.

Proposition 3 ([20])**.** *Given a sequence of finitely exchangeable variables $t_1, t_2, \ldots, t_r$, there exists a linear operator $L : \mathcal{B}_r \to \mathbb{R}$ satisfying*

$$L(g) \geq \sup c \ \textit{s.t.}\ g - c \in \mathcal{B}_r^+, \tag{13}$$

for all polynomials $g \in \mathcal{B}_r$ and where c is the constant polynomial of value c, such that:

$$P(t_1, \ldots, t_r) = L(\theta_1^{n_1} \theta_2^{n_2} \cdots \theta_5^{n_5} \theta_6^{n_6}), \tag{14}$$

where $n_i = \sum_{j=1}^{r} \delta_i(t_j)$ denotes the number of occurrences of the i-th face in the sequence $t_1, \ldots, t_r$.

With the expression in Eq. (13), we mean that $L(g) \geq \gamma$ where $\gamma = \{\sup c \mid g - c \in \mathcal{B}_r^+\}$. Note that, the supremum in (13) is taken over the value c, which defines the constant polynomial c. We call a linear operator L satisfying (13) a *quasi-expectation operator*. The general definition of *quasi-expectation operator*, along with some of its main properties, are outlined in Appendix A. Proposition 3 states that that any probability distribution P of an exchangeable sequence $t_1, \ldots, t_r$ can be defined as a quasi-expectation of a polynomial of the variables $\theta_1, \ldots, \theta_6$. Since the polynomial $\theta_1^{n_1} \theta_2^{n_2} \cdots \theta_5^{n_5} \theta_6^{n_6}$ only depends on the sequence $t_1, \ldots, t_r$ through the number of occurrences n_i, the probability P defined by (14) is by definition finitely exchangeable. The condition (13) ensures that the quasi-expectation operator L defines valid (non-negative and normalised) probabilities P via (14). In fact, note that, for $g = c$ (the constant polynomial), the condition (13) is equal to $L(c) \geq c$. Similarly for $g = -c$, we have that $L(-c) \geq -c$. Since L is linear, we have that $L(-c) = -L(c) \geq -c$ and so $L(c) \leq c$. The two inequalities imply that

$L(c) = c$. For $c = 1$, we have that $L(1) = 1$. Hence, from (12) for $c = 1$ and by linearity of L, we can derive that

$$\sum_{\mathbf{n}=[n_1,\ldots,n_6]:\sum_i n_i=r} b_{n_1 n_2 \ldots n_6} L\left(\theta_1^{n_1}\theta_2^{n_2}\cdots\theta_6^{n_6}\right) = L(1) = 1.$$

Consider now $P(t_1,\ldots,t_r)$ and let $n_i = \sum_{j=1}^{r} \delta_i(t_j)$ the number of occurrences of the i-th face in the sequence $t_1,\ldots,t_r$. From (14), we have that $P(t_1,\ldots,t_r) = L(\theta_1^{n_1}\theta_2^{n_2}\cdots\theta_5^{n_5}\theta_6^{n_6})$. This equality holds for any permutation $t_{\pi(1)}, t_{\pi(2)},\ldots,t_{\pi(r)}$ as they all have the same amount of counts $n_1, n_2,\ldots,n_6$. There are $r!/(n_1!n_2!n_3!n_4!n_5!n_6!)$ of these permutations. Therefore, we have that

$$\sum_{t_1,\ldots,t_r} P(t_1,\ldots,t_r) = \sum_{\mathbf{n}=[n_1,\ldots,n_6]:\sum_i n_i=r} b_{n_1 n_2 \ldots n_6} L\left(\theta_1^{n_1}\theta_2^{n_2}\cdots\theta_6^{n_6}\right) = L(1) = 1$$

(normalised). It remains to show that $P(t_1,\ldots,t_r) > 0$ for every $t_1,\ldots,t_r$. Note that, for $g(\boldsymbol{\theta}) = \theta_1^{n_1}\theta_2^{n_2}\cdots\theta_5^{n_5}\theta_6^{n_6}$, the supremum value of c such that the polynomial $g - c \in \mathcal{B}_r^+$ is $c = 0$. From (13), it follows that $L(g) \geq 0$ (non-negative) for every $g(\boldsymbol{\theta}) = \theta_1^{n_1}\theta_2^{n_2}\cdots\theta_5^{n_5}\theta_6^{n_6}$.

Remark 1. *It is important to note that L is defined on $\mathcal{B}_r$, which is the space of polynomials of degree up to r, but since Bernstein polynomials of degree r form a basis for the space of polynomials of degree up to r, L is completely determined once we specify its value $L(g)$ for every Bernstein polynomial g of degree r.*

Example 3. *Consider the case $r = 2$ and a probability P defined via a linear operator L such that $L(\theta_1\theta_2) = 0.5$ and $L(\theta_i\theta_j) = 0$ for any other monomial of degree two in $\mathcal{B}_2$. It is easy to check that L satisfies* (13) *and, therefore, defines a non-negative and normalised finitely exchangeable P. From* (14)*, it follows that:*

$$\begin{aligned}
&P(t_1 = 1, t_2 = 1) = L(\theta_1^2) = 0,\\
&P(t_1 = 1, t_2 = 2) = P(t_1 = 2, t_2 = 1) = L(\theta_1\theta_2) = 0.5,\\
&P(t_1 = 2, t_2 = 2) = L(\theta_2^2) = 0,
\end{aligned}$$

and zero otherwise.

We will now use this representation through the quasi-expectation operator to determine bounds for linear functions of event probabilities. In particular, in the next example, we will exploit (13) to compute a lower bound for $L(g)$ for $g(\boldsymbol{\theta}) = \theta_1^2 - \theta_1\theta_2 + \theta^2$.

Example 4. *Let us consider again the problem discussed in Example 2, with $r = 2$. Assume we aim to find the finitely exchangeable probability P_2 which minimises*

$$v_2 = \min_{P_2} \Big(P_2 \left(\boxdot \, \boxdot \right) - P_2 \left(\boxdot \, \boxdot \right) + P_2 \left(\boxdot \, \boxdot \right) \Big).$$

Using the operator L, this quantity can be rewritten as

$$\begin{aligned} v_2 &= \min_{L_2} L_2(\theta_1^2 - \theta_1\theta_2 + \theta_2^2) \\ &= \min_{L_2} \Big(L_2(\theta_1^2) - L_2(\theta_1\theta_2) + L_2(\theta_2^2) \Big) \end{aligned}$$

where L_2 indicates the linear operator corresponding to P_2.

By (13)*, knowing that $\theta_6 := 1 - \theta_1 - \cdots - \theta_5$, the value v_2 can be determined by solving the following optimisation problem:*[4]

$$v_2 = \max c \quad s.t. \quad \theta_1^2 - \theta_1\theta_2 + \theta_2^2 - c \in \mathcal{B}_2^+, \tag{15}$$

where

$$\begin{aligned} \mathcal{B}_2^+ = \{ & u_{200000}\theta_1^2 + u_{110000}\theta_1\theta_2 + u_{101000}\theta_1\theta_3 + \\ & u_{100100}\theta_1\theta_4 + u_{100010}\theta_1\theta_5 \\ & + u_{100001}\theta_1(1 - \theta_1 - \cdots - \theta_5) \\ & + u_{020000}\theta_2^2 + \cdots + u_{000011}\theta_5(1 - \theta_1 - \cdots - \theta_5) \\ & + u_{000002}(1 - \theta_1 - \cdots - \theta_5)^2 : u_{[n_1,n_2,\ldots,n_6]} \geq 0 \}. \end{aligned} \tag{16}$$

For a particular c, the polynomial $\theta_1^2 - \theta_1\theta_2 + \theta_2^2 - c$ belongs to $\mathcal{B}_2^+$ if we can find coefficients $u_{200000}, u_{110000}, \ldots, u_{000002} \geq 0$ such that

[4]The supremum can be attained in this case and, therefore, we have replaced sup with max.

$$
\begin{aligned}
&\theta_1^2 - \theta_1\theta_2 + \theta_2^2 - c = u_{200000}\theta_1^2 + u_{110000}\theta_1\theta_2 \\
&+ u_{101000}\theta_1\theta_3 + u_{100100}\theta_1\theta_4 + u_{100010}\theta_1\theta_5 \\
&+ u_{100001}\theta_1(1 - \theta_1 - \cdots - \theta_5) \\
&+ u_{020000}\theta_2^2 + \cdots + u_{000011}\theta_5(1 - \theta_1 - \cdots - \theta_5) \\
&+ u_{000002}(1 - \theta_1 - \cdots - \theta_5)^2.
\end{aligned} \tag{17}
$$

Therefore, the solution of (15) *can be computed by solving the following linear programming problem:*

$$
\begin{aligned}
&\max_{c\in\mathbb{R}, u_{[n_1,n_2,\ldots,n_6]}\in\mathbb{R}^+} c \\
&s.t. \\
&- u_{000002} + u_{100001} - u_{200000} + 1 = 0 \\
&- 2u_{000002} + u_{010001} + u_{100001} - u_{110000} - 1 = 0 \\
&- u_{000002} + u_{010001} - u_{020000} + 1 = 0 \\
&- c - u_{000002} = 0 \\
&- 2u_{000002} - u_{101000} + u_{100001} + u_{001001} = 0 \\
&\quad \texttt{additional 21 constraints}
\end{aligned} \tag{18}
$$

where the equality constraints have been obtained by equating the coefficients of the monomials in (17)*. For instance, consider the constant term in the r.h.s. of* (17)*: that is* u_{000002}*. It must be equal to the constant term* $-c$*. Therefore, we have that*

$$-c - u_{000002} = 0.$$

Similarly, consider the coefficient of the monomial θ_1^2 *in the r.h.s. of* (17)*: this is* $u_{200000} - u_{100001} + u_{000002}$*. It must be equal to the coefficient of the monomial* θ_1^2 *in* $\theta_1^2 - \theta_1\theta_2 + \theta_2^2 - c$*, which is* 1*, that is*

$$1 - (u_{200000} - u_{100001} + u_{000002}) = 0.$$

The last specified constraint in Eq. (18) *is relative to the monomial* $\theta_1\theta_3$*: since* $\theta_1\theta_3$ *is not in* $\theta_1^2 - \theta_1\theta_2 + \theta_2^2 - c$*, its coefficients are equated to zero. In total, there are 26 constraints.*

The solution of Eq. (15) *can be computed using any linear programming solver and is*

$$v_2 = -0.5.$$

The problem

$$v_3 = \min_{P_3} \Big(P_3\left(\boxdot\,\boxdot\right) - P_3\left(\boxdot\,\boxdot\right) + P_3\left(\boxdot\,\boxdot\right) \Big),$$

where $P_3(t_1 = a, t_2 = b) = \sum_{i=1}^{6} P_3(t_1 = a, t_2 = b, t_3 = i)$ *and* P_3 *is the probability distribution of the finitely exchangeable sequence* t_1, t_2, t_3*, can be also solved via linear programming. In this case, we only need to express the above marginalisation in terms of the* L_3 *operator, for instance:*

$$\begin{aligned} P_3\left(\boxdot\,\boxdot\right) &= L_3\big(\theta_1^2(\theta_1 + \theta_2 + \cdots + \theta_6)\big) \\ P_3\left(\boxdot\,\boxdot\right) &= L_3\big(\theta_1\theta_2(\theta_1 + \theta_2 + \cdots + \theta_6)\big) \\ P_3\left(\boxdot\,\boxdot\right) &= L_3\big(\theta_2^2(\theta_1 + \theta_2 + \cdots + \theta_6)\big) \end{aligned} \tag{19}$$

where we have exploited that $\theta_1 + \theta_2 + \cdots + \theta_6 = 1$*. We can thus compute marginals on the result of two rolls by using a linear operator which is only defined on polynomials of degree 3, an operation that, as clarified in the next section, corresponds to the operation* ***partial trace*** *in QT. The solution is*

$$v_3 \approx -0.167.$$

Note that, in the infinitely exchangeable case, the operator $L_\infty(\dots) := \int_\Theta (\dots)\, dq(\boldsymbol{\theta})$ *applied to the polynomial* $\theta_1^2 - \theta_1\theta_2 + \theta_2^2$*, satisfies*

$$\begin{aligned} v_\infty &= \inf_{L_\infty} L_\infty(\theta_1^2 - \theta_1\theta_2 + \theta_2^2) \\ &= \inf_{L_\infty} L_\infty(\theta_1^2) - L_\infty(\theta_1\theta_2) + L_\infty(\theta_2^2) = 0. \end{aligned}$$

Observe that, $v_\infty = 0$ *because* $\theta_1^2 - \theta_1\theta_2 + \theta_2^2$ *is nonnegative and its infimum is* 0*.*

Figure 1 shows the value of $v_n = \min L_n(g)$ *for* $g(\boldsymbol{\theta}) = \theta_1^2 - \theta_1\theta_2 + \theta_2^2$ *and* $n = 2, \dots, 6$*, where the subscript in* L_n *indicates the linear operator corresponding to* P_n*, that is*

$$v_n = \min_{P_n} \Big(P_n\left(\boxdot\,\boxdot\right) - P_n\left(\boxdot\,\boxdot\right) + P_n\left(\boxdot\,\boxdot\right) \Big)$$

as a function of n*. Notice that the gap between the value computed using de*

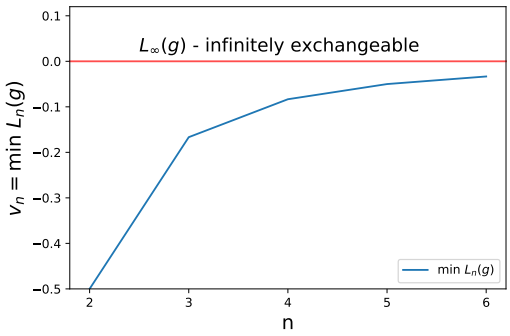


Figure 1: Value of v_n as a function of n. The red line corresponds to v_∞.

Finetti's theorem (red line) and the value computed for finitely exchangeable probabilities reduces as of the number of rolls n increases. This implies that, as the total number of rolls increases, the worst-case marginal expectation for two rolls, which are part of a longer finitely exchangeable sequence of additional $0, 1, 2, \ldots$ *rolls, converges to the classical probability (infinitely exchangeable) value. We can refer to this phenomenon as the 'emergence of classical probability'. We will discuss it again in the next Section.*

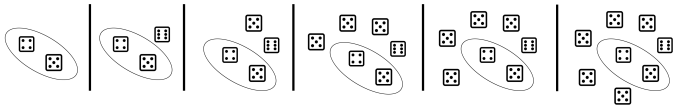


Figure 2: The worst-case marginal expectation for two rolls, which are part of a longer finitely exchangeable sequence of additional $0, 1, 2, \ldots$ rolls, converges to the classical probability value.

5. Making the connection with QT

In this section, we will strengthen the connection between finite exchangeability and finite dimensional QT.

Let $\mathbf{x}$ be a complex vector in $\mathbb{C}^d$ such that $\mathbf{x}^\dagger \mathbf{x} = 1$. We can express the probabilities of the six faces ($d = 6$) of the dice as:

$$\boldsymbol{\theta} = [\theta_1, \ldots, \theta_6] = [x_1^\dagger x_1, \ldots, x_6^\dagger x_6].$$

where † denotes conjugate transpose. Moreover, notice that we can represent the monomials of the Bernstein polynomials with a specific degree r as $\otimes_r \boldsymbol{\theta}$,

where $\otimes$ is the Kronecker product. It follows that any polynomial in $\mathcal{B}_r$ can equivalently be written as

$$g(\theta) = (\otimes_r \mathbf{x})^\dagger D (\otimes_r \mathbf{x}), \tag{20}$$

for some diagonal Hermitian matrix D.

Example 5. *The polynomial $\theta_1^2 - \theta_1\theta_2 + \theta_2^2$ can be written as*

$$\begin{aligned} g(\mathbf{x}^\dagger, \mathbf{x}) &= (x_1^\dagger x_1)^2 - (x_1^\dagger x_1)(x_2^\dagger x_2) + (x_2^\dagger x_2)^2 \\ &= (\otimes_2 \mathbf{x})^\dagger D (\otimes_2 \mathbf{x}) \end{aligned}$$

where D is a diagonal matrix ($diag(D) = [D_1, D_2, \ldots, D_{36}]^\top$) with all zero entries expect $D_1 = 1 = D_8$ and $D_2 = D_7 = -0.5$. Note that, to emphasise that the polynomials we will consider in this section are function of a complex vector $\mathbf{x}$ and its conjugate-transpose $\mathbf{x}^\dagger$, we will denote them as $g(\mathbf{x}^\dagger, \mathbf{x})$.

We now introduce a new linear operator $\widehat{L} : \mathcal{S}_r \to \mathbb{R}$, where

$$\mathcal{S}_r := \{(\otimes_r \mathbf{x})^\dagger G (\otimes_r \mathbf{x}) : G \text{ is Hermitian}\},$$

is the vector space of polynomials of degree $2r$ of the variables $x_i, x_i^\dagger$ for $i = 1, \ldots, d$. Note that, $\mathcal{S}_r$ includes the polynomials of the form (20).

From linearity of the trace $Tr(\cdot)$ and of $\widehat{L}$, we have:

$$\begin{aligned} \widehat{L}(g) &= \widehat{L}((\otimes_r \mathbf{x})^\dagger D (\otimes_r \mathbf{x})) \\ &= Tr\Big(D\widehat{L}\Big((\otimes_r \mathbf{x})(\otimes_r \mathbf{x})^\dagger\Big)\Big). \end{aligned} \tag{21}$$

Note that, $\widehat{L}\Big((\otimes_r \mathbf{x})(\otimes_r \mathbf{x})^\dagger\Big)$ is a Hermitian matrix of dimension $d^r \times d^r$. Example for $r = 2, d = 6$:

$$\widehat{L}\Big((\otimes_2 \mathbf{x})(\otimes_2 \mathbf{x})^\dagger\Big) = \widehat{L} \begin{pmatrix} (x_1^2)^\dagger x_1^2 & (x_1^2)^\dagger x_1 x_2 & \ldots & (x_1^2)^\dagger x_6^2 \\ (x_1 x_2)^\dagger x_1^2 & (x_1 x_2)^\dagger x_1 x_2 & \ldots & (x_1 x_2)^\dagger x_6^2 \\ (x_1 x_3)^\dagger x_1^2 & (x_1 x_3)^\dagger x_1 x_2 & \ldots & (x_1 x_3)^\dagger x_6^2 \\ \vdots & \vdots & \vdots & \vdots \\ (x_6^2)^\dagger x_1^2 & (x_6^2)^\dagger x_1 x_2 & \ldots & (x_6^2)^\dagger x_6^2 \end{pmatrix}, \tag{22}$$

where $\widehat{L}$ is applied element-wise. Observe that the above matrix has additional symmetries (besides being Hermitian), which follow by the symmetries of $(\otimes_r \mathbf{x})(\otimes_r \mathbf{x})^\dagger$. We can then prove the following representation theorem.

Theorem 4. *Consider a linear operator $\widehat{L} : \mathcal{S}_r \to \mathbb{R}$ satisfying*

$$\widehat{L}(g) \geq \sup c \text{ s.t. } g - c \in \mathcal{S}_r^+, \tag{23}$$

for all polynomials $g \in \mathcal{S}_r$, where c is the constant polynomial of value[5] c and $\mathcal{S}_r^+ = \{(\otimes_r \mathbf{x})^\dagger H (\otimes_r \mathbf{x}) : H \text{ is Hermitian and PSD}\}$ with PSD meaning Positive Semi-Definite. We can then prove that

1. *If $\widehat{L}$ satisfies (23) for each polynomial $g \in \mathcal{S}_r$, then the matrix $M = \widehat{L}\Big((\otimes_r \mathbf{x})(\otimes_r \mathbf{x})^\dagger\Big)$ is PSD with trace one.*
2. *$P(t_1, \ldots, t_r) = \widehat{L}((\otimes_r \mathbf{x})^\dagger D^{(t_{1:r})} (\otimes_r \mathbf{x})) = Tr(D^{(t_{1:r})} M)$ — where $D^{(t_{1:r})}$ is a diagonal Hermitian matrix depending on $t_1, \ldots, t_r$,*
   and $M = \widehat{L}\Big((\otimes_r \mathbf{x})(\otimes_r \mathbf{x})^\dagger\Big)$ — is a probability (nonnegative and normalised) and is symmetric to permutations of the labels of the variables t_i.

We have constructed a quasi-expectation operator $\widehat{L}$ on a space of polynomials of complex variables, which is completely determined by trace-one PSD matrices M (PSD with trace one and having the same symmetries as the matrix $(\otimes_r \mathbf{x})(\otimes_r \mathbf{x})^\dagger$). Moreover, this operator, when applied to polynomials $(\otimes_r \mathbf{x})^\dagger D^{(t_{1:r})} (\otimes_r \mathbf{x})$, defines a probability P which is symmetric to permutations of the labels of the variables t_i.

Example 6. *Consider the case $d = 2$ (a coin) and $r = 3$. Then, for instance, we have that*

$$\begin{aligned}
p(t_1 = 1, t_2 = 1, t_3 = 2) &= (\otimes_3 \mathbf{x})^\dagger D^{(112)} (\otimes_3 \mathbf{x}) = (x_1^\dagger x_1)^2 (x_2^\dagger x_2), \\
p(t_1 = 1, t_2 = 2, t_3 = 1) &= (\otimes_3 \mathbf{x})^\dagger D^{(121)} (\otimes_3 \mathbf{x}) = (x_1^\dagger x_1)^2 (x_2^\dagger x_2), \\
p(t_1 = 2, t_2 = 1, t_3 = 1) &= (\otimes_3 \mathbf{x})^\dagger D^{(211)} (\otimes_3 \mathbf{x}) = (x_1^\dagger x_1)^2 (x_2^\dagger x_2),
\end{aligned}$$

with $\mathbf{x} = [x_1, x_2]^\top$ and $D^{(112)} = diag(0,1,0,0,0,0,0,0)$,
$D^{(121)} = diag(0,0,1,0,0,0,0,0)$, $D^{(211)} = diag(0,0,0,0,1,0,0,0)$.

[5] This can be obtained by choosing $H = cI$ where I is the identity matrix.

For the second statement in Theorem 4, note that $Tr(DM) = Tr(D\,\mathrm{diag}(M))$ holds for any diagonal matrix D. Hence Theorem 4 uses some redundant quantities through $\widehat{L}\Big((\otimes_r \mathbf{x})(\otimes_r \mathbf{x})^\dagger\Big)$. This withstanding, it enables another representation result that links finite exchangeability to QT, as discussed next.

First, we recall some useful definitions. Let $\mathcal{H}^{\otimes r}$ be the compound Hilbert space of r particles. A fundamental postulate of QT for bosons is that the quantum states are symmetric upon exchange of the subsystems, respectively. This postulate is formulated mathematically by introducing a projection operator Π_{sym} from the tensor product space to this symmetric subspace. The projector is called *symmetriser* and is defined as [30, B-2-c]:

$$\Pi_{sym} := \frac{1}{N!}\sum_i \Pi_i, \tag{24}$$

where the summations are performed over the $N!$ permutation matrices, Π_i, associated to the permutations of the labels of the particles. Then the postulate states that a density matrix ρ of r particles is said to be *Bose-symmetric* if:

$$\rho = \Pi_{sym}\rho\Pi_{sym}.$$

As briefly derived in Appendix B.2, this definition is equivalent to the more commonly employed one (for instance in [3, 6]), namely, $\rho = \Pi_i\rho = \rho\Pi_i^\dagger$ for every permutation matrix Π_i.

Example 7. *Assume that $dim(\mathcal{H}) = d = 2$ and $r = 2$. In this case, the symmetriser is equal to:*

$$\Pi_{sym} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 \\ 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

It can be verified that, for instance, the following density matrices

$$\frac{1}{2}\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \qquad \frac{1}{2}\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{25}$$

are all Bose-symmetric, as they satisfy $\rho = \Pi_{sym}\rho\Pi_{sym}$. *Let us write a general density matrix as:*

$$\rho = \begin{bmatrix} \rho_{11} & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{12}^\dagger & \rho_{22} & \rho_{23} & \rho_{24} \\ \rho_{13}^\dagger & \rho_{23}^\dagger & \rho_{33} & \rho_{34} \\ \rho_{14}^\dagger & \rho_{24}^\dagger & \rho_{34}^\dagger & \rho_{44} \end{bmatrix}, \tag{26}$$

which must also be Hermitian with trace-one (the elements in the diagonal are nonnegative real numbers and sum up to one), then

$$\Pi_{sym}\rho\Pi_{sym} = \begin{bmatrix} \rho_{11} & \frac{1}{2}(\rho_{12}+\rho_{13}) & \frac{1}{2}(\rho_{12}+\rho_{13}) & \rho_{14} \\ \frac{1}{2}(\rho_{12}^\dagger+\rho_{13}^\dagger) & \frac{\rho_{22}+\rho_{33}+\rho_{23}+\rho_{23}^\dagger}{4} & \frac{\rho_{22}+\rho_{33}+\rho_{23}+\rho_{23}^\dagger}{4} & \frac{1}{2}(\rho_{24}+\rho_{34}) \\ \frac{1}{2}(\rho_{12}^\dagger+\rho_{13}^\dagger) & \frac{\rho_{22}+\rho_{33}+\rho_{23}+\rho_{23}^\dagger}{4} & \frac{\rho_{22}+\rho_{33}+\rho_{23}+\rho_{23}^\dagger}{4} & \frac{1}{2}(\rho_{24}+\rho_{34}) \\ \rho_{14}^\dagger & \frac{1}{2}(\rho_{24}^\dagger+\rho_{34}^\dagger) & \frac{1}{2}(\rho_{24}^\dagger+\rho_{34}^\dagger) & \rho_{44} \end{bmatrix}. \tag{27}$$

We can then define the conditions for a generic matrix ρ *to be Bose-symmetric by equating the elements of the two matrices in* (26) *and* (27). *For instance, for the element in the first-row and second-column, this leads to* $\rho_{12} = \frac{1}{2}(\rho_{12} + \rho_{13})$, *which implies that* $\rho_{12} = \rho_{13}$. *Similarly, we can derive that* $\rho_{24} = \rho_{34}$. *For the four central elements, we have that* $\rho_{22} = \rho_{23} = \rho_{33} = \frac{\rho_{22}+\rho_{33}+\rho_{23}+\rho_{23}^\dagger}{4}$, *which is a real number. Therefore, a generic* 4×4 *Bose-symmetric matrix must have these symmetries:*

$$\rho = \begin{bmatrix} \rho_{11} & \rho_{12} & \rho_{12} & \rho_{14} \\ \rho_{12}^\dagger & \rho_{22} & \rho_{22} & \rho_{24} \\ \rho_{12}^\dagger & \rho_{22} & \rho_{22} & \rho_{24} \\ \rho_{14}^\dagger & \rho_{24}^\dagger & \rho_{24}^\dagger & \rho_{44} \end{bmatrix}, \tag{28}$$

where $\rho_{ii} \in \mathbb{R}$ *and* $\rho_{11} + 2\rho_{22} + \rho_{44} = 1$.

We can in general prove the following equivalence result.

Theorem 5. *The following sets are equal:*

$$\mathcal{S}_1 = \{\rho : \ \rho = \Pi_{sym}\rho\Pi_{sym}, \ \rho \succeq 0, \ Tr(\rho) = 1\}, \tag{29}$$

$$\mathcal{S}_2 = \{M : \ M = \widehat{L}((\otimes_r \mathbf{x})(\otimes_r \mathbf{x})^\dagger)\}, \tag{30}$$

where ρ is the density matrix of r-indistinguishable bosons (each one being a d-level quantum system), and $\rho \succeq 0$ means that the matrix ρ is positive semi-definite.

This results states that boson-symmetric density matrices and the set of matrices M obtained by applying the operator $\widehat{L}$ to the matrix $(\otimes_r \mathbf{x})(\otimes_r \mathbf{x})^\dagger$ of r-th degree polynomials in the variables $x_i, x_i^\dagger$ for $i = 1, \ldots, d$ are equivalent. Eq. (30) is obtained by applying a simple change of variables to the classical probability models for finitely exchangeable probabilities, and actually offers an alternative representation of the second-quantisation formalism in QT, which is of interest in its own right (as discussed in the proof of Theorem 5 in Appendix B.3).

Example 8. *For the case $dim(\mathcal{H}) = d = 2$ and $r = 2$, this equivalence can be seen by comparing the matrix in* (28) *with*

$$\widehat{L}\Big((\otimes_2 \mathbf{x})(\otimes_2 \mathbf{x})^\dagger\Big) = \widehat{L}\begin{pmatrix} (x_1^2)^\dagger x_1^2 & (x_1x_2)^\dagger x_1^2 & (x_1x_2)^\dagger x_1^2 & (x_2^2)^\dagger x_1^2 \\ (x_1^2)^\dagger x_1x_2 & (x_1x_2)^\dagger x_1x_2 & (x_1x_2)^\dagger x_1x_2 & (x_2^2)^\dagger x_1x_2 \\ (x_1^2)^\dagger x_1x_2 & (x_1x_2)^\dagger x_1x_2 & (x_1x_2)^\dagger x_1x_2 & (x_2^2)^\dagger x_1x_2 \\ (x_1^2)^\dagger x_2^2 & (x_1x_2)^\dagger x_2^2 & (x_1x_2)^\dagger x_2^2 & (x_2^2)^\dagger x_2^2 \end{pmatrix}, \tag{31}$$

computed for $\mathbf{x} = [x_1, x_2]^\top$. The two matrices have the same symmetries. It can also be noted that, in both matrices, the four central elements are real numbers and the diagonal elements sum up to one, as $(x_1^\dagger x_1)^2 + 2(x_1x_2)^\dagger x_1x_2 + (x_2^\dagger x_2)^2 = \theta_1^2 + 2\theta_1\theta_2 + \theta_2^2 = (\theta_1 + \theta_2)^2 = 1$ and $\widehat{L}(1) = 1$.

From Theorems 4–5, the following result follows directly.

Corollary 1. *The operator defined by $\widehat{L}(g) := Tr(GM)$ for all $g \in \mathcal{S}_r$, with M being (i) PSD; (ii) trace one; and (iii) having the same symmetries as $(\otimes_r \mathbf{x})(\otimes_r \mathbf{x})^\dagger$; satisfies* (23) *for each $g \in \mathcal{S}_r$.*

Remark 2.

- *In [21, 22], we derived a representation for indistinguishable particles for bosons using exchangeability and the symmetries* $\rho = \Pi_{sym}\rho\Pi_{sym}$*, but not in the same way as the de Finetti's representation. In the de Finetti's representation of the probability distribution for an exchangeable sequence, the only relevant information about the sequences is the number of occurrences of an event (such as the number of rolls* n_i *of the i-th face). This naturally leads to a sort of second-order quantisation representation, where the variables (e.g.,* $\theta_1, \ldots, \theta_6$*) appear exponentiated (e.g.,* $\theta_1^{n_1}, \ldots, \theta_6^{n_6}$*) in the polynomial term, such as in Eq.* (3)*. In this paper, we have shown that we can provide a similar representation (with exponentiated variables* $\mathbf{x}, \mathbf{x}^\dagger$*) for QT (bosons), which is equivalent to* $\rho = \Pi_{sym}\rho\Pi_{sym}$*, as proven in Theorem 5.*

- *An approach based on a similar exponentiated polynomial representation was proposed in [23] to define a criterion to determine if a density matrix of distinguishable particles is entangled or not. This approach is known as the Doherty-Parrilo-Spedalieri (DPS) hierarchy. For a bipartite density matrix* ρ^{AB}*, the level-k DPS relaxation determines whether there exists an extension* $\tilde{\rho}^{A_1 \ldots A_k B_1 \ldots B_k}$ *that satisfies the following conditions: (i) it is invariant under left or right multiplication by any permutation of the A or B subsystems, and (ii) it remains positive semidefinite under partial transposition of any subset of the systems. The latter condition is commonly referred to as Positivity under Partial Transpose (PPT). Such a level-k extension exists if and only if* ρ^{AB} *is not entangled [23, Th. 1]. Notably, finding such an extension* $\tilde{\rho}$ *reduces to solving k semidefinite programming problems or, dually, k sum-of-squares optimisation problems [31]. Their main result [23, Th. 1] is proven using the quantum de Finetti theorem. Unlike this method, which serves as a mathematical optimisation tool to determine whether a density matrix is not entangled, our main results (Theorem 4 and 5) provides a representation theorem for boson-symmetric density matrices. However, the condition* (23) *that defines valid linear operators* $\widehat{L}$ *could also be used to define a hierarchy of optimisation problems to verify if the density matrix of n indistinguishable particles is not entangled, using [23, Th. 1].*

- *Theorem 5 demonstrates that any finite system of exchangeable events is equivalent to a quantum experiment involving indistinguishable bosons measured using projective measurements along the coordinate axes. This implies that we can build a quantum experiment whose probability of the measurement outcomes coincides with the probability of a finitely exchangeable sequence of dice rolling. This means that quantum experiments provide an alternative way to the urn-model discussed in Example 1 to realise a finitely exchangeable probability. To illustrate this observation, we provide an example.*

Example 9. *We can compute* $v_n = \min_{P_n} P_n\left(\boxdot\,\boxdot\right) - P_n\left(\boxdot\,\boxdot\right) + P_n\left(\boxdot\,\boxdot\right)$ *using* n *indistinguishable bosons. Let* $\rho^{(n)}$ *denote the density matrix of* n *6-level bosons, and* $\Pi^{(n)}_{sym}$ *be the symmetriser for the* n *bosons. Note that, we consider 6-level bosons because we are focusing on a dice rolling experiment. Here* n *corresponds to the number of rolls (*$n = r$*). From Theorems 4–5, we can rewrite* v_n *as* $v_n = \min_{\rho^{(n)}} Tr(D^{[n]}\rho^{(n)})$ *with* $\rho^{(n)} = \Pi^{(n)}_{sym}\rho^{(n)}\Pi^{(n)}_{sym}$, $Tr(\rho^{(n)}) = 1$ *and* $\rho^{(n)} \succeq 0$. *All the superscripts are just indexes. Note that, the superscript in* $D^{[n]}$ *denotes the number of bosons (the number of rolls in the dice experiment). For instance, for* $n = 2$, $D^{[2]}$ *can be written as an algebraic sum of the matrices* $D^{(t_1t_2)}$. *In particular, we have that* $D^{[2]} = D^{(t_1=1,t_2=1)} - 0.5D^{(t_1=1,t_2=2)} - 0.5D^{(t_1=2,t_2=1)} + D^{(t_2=2,t_2=2)}$ *and it is a diagonal matrix with all zero entries except* $D^{[2]}_1 = 1 = D^{[2]}_8$ *and* $D^{[2]}_2 = D^{[2]}_7 = -0.5$ *(*$D^{[2]}$ *is a physical-observable operator, as it is symmetric to permutations). This holds because the diagonal of* $\rho^{(2)}$ *corresponds to the monomials* $[\theta_1^2, \theta_1\theta_2, \ldots, \theta_2\theta_1, \theta_2^2, \ldots, \theta_6^2]$. *Any* $D^{[n]}$ *(for* $n > 2$*) can be defined in a similar way exploiting the 'partial trace'-style operations illustrated in Eq.* (19).

The value of v_n *is shown in Figure 3 as a function of* n. *Note that, in this case, in order to compute* $v_n = \min_{\rho^{(n)}} Tr(D^{[n]}\rho^{(n)})$, *we would in principle need to solve a semidefinite programming problem due to the constraint* $\rho^{(n)} \succeq 0$ *(it must be a PSD matrix). However, since* $D^{[n]}$ *is diagonal, this problem reduces to a linear programming problem.*

The values of v_n *in Figure 3 coincide exactly with the values in Figure 1.*

This means that the worst-case marginal expectation for an operator $D^{[n]}$ *on two bosons, which are part of a longer sequence of additional 0, 1, 2,... bosons, converges to the classical probability value as the total number of bosons increases, see Figure 4. In QT, we would refer to this phenomenon as the*

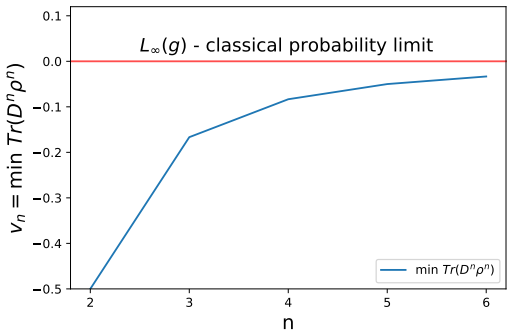


Figure 3: Value of v_n as a function of n. The red line corresponds to v_∞.

'emergence of classical reality'. In probability theory, this is the convergence from finite to infinite exchangeability.

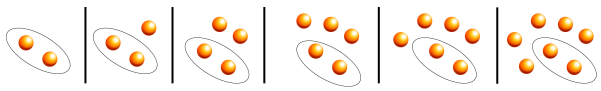


Figure 4: The worst-case marginal expectation for two bosons, which are part of a longer indistinguishable sequence of additional $0, 1, 2, \ldots$ bosons, converges to the classical probability value.

The example above illustrates the following connection between finite exchangeability and entanglement. To understand it, for simplicity, we consider the case when $r = d = 2$, which corresponds to a coin tossing experiment with two coins. By Theorem 5, the finitely exchangeable sequence t_1, t_2 with respect to the probability $P_2(t_1 = H, t_2 = T) = P_2(t_1 = T, t_2 = H) = 0.5$ (H=Heads, T=Tails) can be represented as

$$\begin{aligned} \rho &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 \\ 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ &= M = \widehat{L} \begin{pmatrix} (x_1^2)^\dagger x_1^2 & (x_1^2)^\dagger x_1 x_2 & (x_1^2)^\dagger x_1 x_2 & (x_1^2)^\dagger x_2 x_2 \\ (x_1 x_2)^\dagger x_1^2 & (x_1 x_2)^\dagger x_1 x_2 & (x_1 x_2)^\dagger x_1 x_2 & (x_1 x_2)^\dagger x_2 x_2 \\ (x_1 x_2)^\dagger x_1^2 & (x_1 x_2)^\dagger x_1 x_2 & (x_1 x_2)^\dagger x_1 x_2 & (x_1 x_2)^\dagger x_2 x_2 \\ (x_2^2)^\dagger x_1^2 & (x_2^2)^\dagger x_1 x_2 & (x_2^2)^\dagger x_1 x_2 & (x_2^2)^\dagger x_2 x_2 \end{pmatrix} \end{aligned} \tag{32}$$

where $\widehat{L}((x_1 x_2)^\dagger x_1 x_2) = 0.5$ and zero otherwise.

The diagonal elements of ρ include the probabilities $\rho_{11} = P_2(t_1 = H, t_2 = H) = 0$, $\rho_{22} = P_2(t_1 = H, t_2 = T) = 0.5$, $\rho_{33} = P_2(t_1 = T, t_2 = H) = 0.5$ and $\rho_{44} = P_2(t_1 = T, t_2 = T) = 0$. The symmetry condition $\rho = \Pi_{sym}\rho\Pi_{sym}$ implies that the off-diagonal elements ρ_{23}, ρ_{32} are equal to 0.5. In fact, when considering the matrix defined by $\widehat{L}$, the four elements $M_{22}, M_{23}, M_{32}, M_{33}$ are equal: by definition, the operator $\widehat{L}$ applies to the same monomial $x_1^\dagger x_2^\dagger x_1 x_2$.

In Example 9, we then chose a Hermitian operator $D^{[2]}$, which for the two coins case reduces to $D^{[2]} = \text{diag}([1, -0.5, -0.5, 1])$, and results in the value $Tr(D^{[2]}\rho) = -0.5$. This value is the same value obtained by solving $v_2 = \min_{P_2} P_2(H, H) - P_2(H, T) + P_2(T, T) = 0 - 0.5 + 0 = -0.5$.

In Example 5, we observed that the coefficients in the diagonal of $D^{[2]}$ give rise to the polynomial $\theta_1^2 - \theta_1\theta_2 + \theta_2^2$. This polynomial is always nonnegative and, therefore, can serve as a *finite exchangeability witness*. In fact, we used the inequality $-0.5 = v_2 < v_\infty = 0$ to prove that P_2 is finitely exchangeable. The reader may thus ask: is $D^{[2]}$ also an entanglement witness for ρ, as defined in Eq. (32)?[6] Several nonequivalent ways of identification and quantification of the entanglement for bosons have been proposed, see [24] for a recent review. One of the main notions of entanglement for bosons states that non-entangled states are equivalent to *simple symmetric tensors* [33–51]. Following [24], we refer to this definition of entanglement for identical particles as *Entanglement-I*. In the case of two bosons, this definition implies that a pure state is non-entangled if and only if it is a tensor product of two identical one-particle states [40]. Notice that this definition is consistent with the model introduced through the operator $\widehat{L}$ in Theorem 4. Indeed, a pure[7] non-entangled density matrix is simply the matrix $M = (\otimes_r \hat{\mathbf{x}})(\otimes_r \hat{\mathbf{x}})^\dagger$ for some $\hat{\mathbf{x}} \in \mathbb{C}^d$ with $\hat{\mathbf{x}}^\dagger\hat{\mathbf{x}} = 1$, corresponding to the product of identical one-particle states (each state corresponding to $\hat{\mathbf{x}}$). In the dice experiment, this means that we have complete knowledge of the probabilities of each face of the dice. Therefore, $D^{[2]} = \text{diag}([1, -0.5, -0.5, 1])$ is a witness for this definition of entanglement. Indeed, observe that the polynomial $(\otimes_2 \mathbf{x})^\dagger D^{[2]} (\otimes_2 \mathbf{x})$

[6] In QT, in the case of distinguishable particles, witnesses are commonly [32] used to differentiate entangled density matrices from separable (non-entangled) ones, which are compatible with classical probability theory.

[7] A pure (or atomic) probability measure, corresponding to a non-entangled case, is one that assigns a precise probability distribution concentrated on a single outcome $\mathbf{x}$. In other words, it represents complete certainty about $\mathbf{x}$, as the probability measure cannot be decomposed into a convex combination of other measures.

is nonnegative. Since $Tr(D^{[2]}\rho) = -0.5 < 0$, this means that ρ in Eq. (32) is entangled.

Remark 3. *To clarify this point, consider the polynomial*

$$(\otimes_2 \mathbf{x})^\dagger D^{[2]} (\otimes_2 \mathbf{x}) = \theta_1^2 - \theta_1\theta_2 + \theta_2^2.$$

This polynomial is nonnegative for every probabilities $[\theta_1, \theta_2]$ we choose for the faces of the coin. According to the Entanglement-I definition, the only pure non-entangled states are those where $[\theta_1, \theta_2]$ are fully known. In other words, in this pure case, an operator $\widehat{L}$ is not entangled if $\widehat{L}\left((\otimes_2 \mathbf{x})(\otimes_r \mathbf{x})^\dagger\right) = (\otimes_2 \hat{\mathbf{x}})(\otimes_2 \hat{\mathbf{x}})^\dagger$ for some $\hat{\mathbf{x}} \in \mathbb{C}^2$ with $\hat{\mathbf{x}}^\dagger \hat{\mathbf{x}} = 1$. A pure state $\hat{\mathbf{x}}$ attains the infimum of $(\otimes_2 \mathbf{x})^\dagger D^{[2]} (\otimes_2 \mathbf{x})$, since any mixture of pure states yields a higher value. Therefore, since (i) $(\otimes_2 \mathbf{x})^\dagger D^{[2]} (\otimes_2 \mathbf{x}) \geq 0$; (ii) the operator $\widehat{L}$ defined as $\widehat{L}\left((\otimes_2 \mathbf{x})(\otimes_2 \mathbf{x})^\dagger\right) = \rho$, with ρ in Eq. (32), is such that $\widehat{L}\left((\otimes_2 \mathbf{x})^\dagger D^{[2]} (\otimes_2 \mathbf{x})\right) < 0$; this means that this operator $\widehat{L}$ must be Entangled-I (it cannot be written as $\rho = \sum_i w_i (\otimes_2 \hat{\mathbf{x}}_i)(\otimes_2 \hat{\mathbf{x}}_i)^\dagger$ for any probabilistic mixture of $\hat{\mathbf{x}}_i$).

Entanglement-I has been challenged as a definition of entanglement for indistinguishable particles [24]. We will discuss these arguments in the next Section 6.

6. Application

For indistinguishable particles (IPs) systems, many (of the standard equivalent) notions of entanglement are ambiguous. The reason is that the symmetries resulting from indistinguishability have mathematically the same structure as a conventional superposition. Therefore, symmetries generate a sort of *structural entanglement*, which is simply determined by the particles fundamental indistinguishability. For instance, the density matrices in (25) are Bose-symmetric, and therefore valid density matrices for indistinguishable bosons. Their structure is constrained by this symmetry. However, in the case of distinguishable particles, both matrices correspond to density matrices of two entangled particles. This creates an ambiguity between the symmetries arising from indistinguishability and the structural features that give rise to entanglement in the distinguishable case. This has given rise to different viewpoints about physical meaning and assessment of entanglement for IPs. There at least five main different definitions of entanglement for IPs

in literature – we refer the reader to [24] for an overview, and a survey of the literature on the argument.

Recently, [24] attempted to systematise these different definitions by proposing a set of properties that should be used to characterise entanglement for IPs. These properties have been defined by trying to generalise properties that hold for entanglement of distinguishable particles.

The proposed properties are: (i) *local operators:* entanglement for IPs must correspond to the presence of non-local correlations; (ii) *effective distinguishability:* the entanglement theory developed for IPs must reduce to the one for distinguishable particles when applied to IPs that have effectively become distinguishable by "freezing" suitable degrees of freedom; (iii) *information processing resources:* local operators acting on separable states must not enhance the performances of informational tasks with respect to classical ones. It is important to note that [24] does not relate entanglement to the violation of Bell-type inequalities (that is, through Dutch-book arguments).

In [24], it is claimed that 'Entanglement-I' violates the properties *local operators* and *effective distinguishability*. We will use the analogy between finite exchangeability and quantum theory developed in this paper to highlight some issues in the definitions of *local operators* used in [24], and explain the argument about *effective distinguishability*. In particular, we will show that the definitions given in [24] lead to contradictions when applied (through our connection between QT for bosons and classical finite exchangeability) to classical finitely exchangeable probabilities. We then propose a modification of these properties, which avoids this issue.

6.1. Local operators

We start first with the argument about *local operators* [24, Sec. 3.1.1]. Consider two single-particle operators (Hermitian matrices) O_1 for particle 1 and O_2 for particle 2. We know that

$$O_1 \otimes O_2 = (O_1 \otimes I_2)(I_1 \otimes O_2),$$

where I_i is the identity-matrix for the i-th particle. For distinguishable particles, it is well known that a bipartite density matrix ρ_{12} is separable if and only if, for all O_1, O_2:

$$Tr((O_1 \otimes O_2)\,\rho_{12}) = \sum_{j=1}^{m} p_j Tr(O_1 \rho_j^{(1)}) Tr(O_2 \rho_j^{(2)}), \tag{33}$$

for some probabilities p_j, $\sum_{j=1}^{m} p_j = 1$, and single-particle density matrices $\rho_j^{(1)}$ and, respectively $\rho_j^{(2)}$. For a pure density matrix ρ_{12}, the above equation implies that:

$$Tr((O_1 \otimes O_2)\rho_{12}) - Tr((O_1 \otimes I_2)\,\rho_{12})Tr((I_1 \otimes O_2)\,\rho_{12}) = 0. \tag{34}$$

The extension of this definition to IPs is quite challenging because it requires that each particle is individually addressable. For IPs, $O_1 \otimes O_2$, $O_1 \otimes I_2$ and $I_1 \otimes O_2$ are not physical observable in general, because they are not symmetric.

To account for this issue, [24, Sec. 3.1.1] considers two commuting operators O_1, O_2 and define their symmetrised versions $\mathcal{P}(O_1, I), \mathcal{P}(I, O_2)$, as follows:

$$\mathcal{P}(O, I) = \mathcal{P}(I, O) = O \otimes I + I \otimes O. \tag{35}$$

Commutativity of O_1, O_2 implies that they have a common eigenbasis $\{|e_\lambda\rangle\}_\lambda$ corresponding to eigenvalues $\{o_\lambda^{(1)}\}_\lambda$ for O_1, and $\{o_\lambda^{(2)}\}_\lambda$ for O_2. Now consider a non-entangled-I state of the form $|\psi\rangle \otimes |\psi\rangle$ with

$$|\psi\rangle = \frac{|e_\lambda\rangle + |e_\kappa\rangle}{\sqrt{2}}, \tag{36}$$

and $\lambda \neq \kappa$, where $|\psi\rangle$ denotes a column-vector in $\mathbb{C}^2$, $\langle\psi|$ denotes its transpose; each vector $|\psi\rangle$ has unit-norm, that is $\langle\psi|\psi\rangle = 1$.

Then, the analogous (according to the definition in [24]) of (34) for a non-entangled-I state becomes:

$$\begin{aligned}
&\langle\psi\otimes\psi|\mathcal{P}(O_1, I)\mathcal{P}(O_2, I)\,|\psi\otimes\psi\rangle - \\
&\langle\psi\otimes\psi|\mathcal{P}(O_1, I)\,|\psi\otimes\psi\rangle\,\langle\psi\otimes\psi|\mathcal{P}(O_2, I)\,|\psi\otimes\psi\rangle \\
&= \frac{1}{2}\big(o_\lambda^{(1)} - o_\kappa^{(1)}\big)\big(o_\lambda^{(2)} - o_\kappa^{(2)}\big)\ .
\end{aligned} \tag{37}$$

The above difference is zero if and only if either $o_\lambda^{(1)} = o_\kappa^{(1)}$ or $o_\lambda^{(2)} = o_\kappa^{(2)}$, which is not true in general for generic commuting O_1, O_2. According to the definition of *local operators* in [24], this means that, in general, non-entangled-I states do not satisfy the factorisation of expectations for (symmetric) local operators, which (in the distinguishable particles case) highlights the presence of non-local correlation (entanglement).

6.1.1. Analogy with finite/infinite exchangeability

Hereafter, by using the analogy between finite exchangeability and QT, we challenge this conclusion and, in particular, the definition (35) of local operators.

Consider a coin tossing experiment with two coins ($r = d = 2$). We choose the following diagonal matrices to define polynomials over the probabilities of the coin sides.

$$O_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad O_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \tag{38}$$

They commute in the sense that $O_1O_2 - O_2O_1 = 0$. These diagonal matrices are defined for the single coins, we then extend them to two coins using the approach described in [24, Sec. 3.1.1]. This results in the following symmetrised versions:

$$\begin{aligned} \mathcal{P}(O_1, I) &= \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\ \mathcal{P}(I, O_2) &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \\ \mathcal{P}(O_1, I)\mathcal{P}(I, O_2) &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \end{aligned}$$

which are all diagonal operators and, therefore, they are valid observables for the tossing of two finitely exchangeable coins. Indeed, these operators define

the following polynomials

$$\begin{aligned}
g_1(\mathbf{x}^\dagger, \mathbf{x}) &= (\otimes_2 \mathbf{x})^\dagger \mathcal{P}(O_1, I)(\otimes_2 \mathbf{x}) \\
&= 2(x_1^\dagger x_1)^2 + 2(x_1^\dagger x_1)(x_2^\dagger x_2) && (39) \\
&= 2\theta_1^2 + 2\theta_1\theta_2 && (40) \\
g_2(\mathbf{x}^\dagger, \mathbf{x}) &= (\otimes_2 \mathbf{x})^\dagger \mathcal{P}(I, O_2)(\otimes_2 \mathbf{x}) \\
&= 2(x_1^\dagger x_1)(x_2^\dagger x_2) + 2(x_2^\dagger x_2)^2 && (41) \\
&= 2\theta_1\theta_2 + 2\theta_2^2 && (42) \\
g_{12}(\mathbf{x}^\dagger, \mathbf{x}) &= (\otimes_2 \mathbf{x})^\dagger \mathcal{P}(O_1, I)\mathcal{P}(I, O_2)(\otimes_2 \mathbf{x}) \\
&= 2(x_1^\dagger x_1)(x_2^\dagger x_2) && (43) \\
&= 2\theta_1\theta_2. && (44)
\end{aligned}$$

For a pure state (that is, for a choice of the probability of Heads θ_1 and Tails θ_2 of the coin), the *local operators* condition (37) requires that:

$$\begin{aligned}
&2\theta_1\theta_2 - (2\theta_1^2 + 2\theta_1\theta_2)(2\theta_1\theta_2 + 2\theta_2^2) \\
&= -2\theta_1(1-\theta_1) = 0,
\end{aligned}$$

(we have exploited that $\theta_2 = 1 - \theta_1$), which is zero if and only if either $\theta_1 = 0$ or $\theta_1 = 1$. The *local operators* condition (37) is therefore violated by the independent tossing of two identical coins with $\theta_1 \in (0, 1)$. For instance, for $\theta_1 = 0.5$, the difference is equal to $-2\theta_1(1-\theta_1) = -0.5$, which coincides with the value computed by using (37) (for O_1, O_2 defined in (38)). This shows that the argument about *local operators* is problematic. If it were true – by exploiting the connection between entangled states being equivalent to finite exchangeability and non-entangled states being equivalent to infinite exchangeability – this would imply that the uniform probability $P(t_1 = H, t_2 = H) = \theta_1^2 = 0.25$, $P(t_1 = H, t_2 = T) = P(t_1 = T, t_2 = H) = \theta_1\theta_2 = 0.25$ and $P(t_1 = T, t_2 = T) = (1-\theta_1)(1-\theta_2) = 0.25$ is finitely (entangled) but not infinitely exchangeable (not entangled), which is clearly false.

The key issue here is the definition of local operators, and we will use the analogy with exchangeability to properly define them. From the results in Section 5, it is immediate verify that, for any an observable G,

$$\widehat{L}\Big((\otimes_r \mathbf{x})^\dagger G(\otimes_r \mathbf{x})\Big) = \widehat{L}\Big((\otimes_r \mathbf{x})^\dagger \tfrac{\Pi_i^\dagger G \Pi_j + \Pi_j^\dagger G \Pi_i}{2}(\otimes_r \mathbf{x})\Big),$$

for all $\Pi_i, \Pi_j \in \mathbb{P}_2$, where $\mathbb{P}_2$ denotes the collection of all permutation matrices of two particles, and

$$\widehat{L}\Big((\otimes_r \mathbf{x})^\dagger G(\otimes_r \mathbf{x})\Big) = \widehat{L}\Big((\otimes_r \mathbf{x})^\dagger \Pi_{sym} G \Pi_{sym}(\otimes_r \mathbf{x})\Big).$$

This holds because $\Pi_{sym}(\otimes_r \mathbf{x}) = (\otimes_r \mathbf{x})$ from Theorem 5. The symmetries in exchangeability (and, analogously, indistinguishability) thus define *equivalence classes* of observables, consisting of all permuted gambles with the same expectation as $Tr(GM)$, where $M = \widehat{L}\Big((\otimes_r \mathbf{x})(\otimes_r \mathbf{x})^\dagger\Big)$. This equivalence class contains physical observables, that is operators that are symmetric to permutations. We can thence calculate the expectation value of a non-physical observable G by simply choosing any physical observable in the equivalence class $\{G\}^{eq}$ and computing its expectation value. It is immediate to show that

$$\frac{O_1 \otimes I + I \otimes O_1}{2} \in \{O_1 \otimes I\}^{eq} \tag{45}$$

$$\frac{I \otimes O_2 + O_2 \otimes I}{2} \in \{I \otimes O_2\}^{eq} \tag{46}$$

$$\frac{O_1 \otimes O_2 + O_2 \otimes O_1}{2} \in \{O_1 \otimes O_2\}^{eq} \tag{47}$$

In fact, notice for instance that

$$\Pi_{sym} \frac{O_1 \otimes O_2 + O_2 \otimes O_1}{2} \Pi_{sym} = \Pi_{sym}(O_1 \otimes O_2)\Pi_{sym}$$

for each O_1, O_2. Therefore, it holds that

$$\begin{aligned}
&(\otimes_2 \mathbf{x})^\dagger \tfrac{O_1 \otimes O_2 + O_2 \otimes O_1}{2} (\otimes_2 \mathbf{x}) - \\
&(\otimes_2 \mathbf{x})^\dagger \tfrac{O_1 \otimes I + I \otimes O_1}{2} (\otimes_2 \mathbf{x}) \\
&\cdot (\otimes_2 \mathbf{x})^\dagger \tfrac{I \otimes O_2 + O_2 \otimes I}{2} (\otimes_2 \mathbf{x}) \\
&= (\otimes_2 \mathbf{x})^\dagger (O_1 \otimes O_2)(\otimes_2 \mathbf{x}) - \\
&(\otimes_2 \mathbf{x})^\dagger (O_1 \otimes I)(\otimes_2 \mathbf{x}) \\
&\cdot (\otimes_2 \mathbf{x})^\dagger (I \otimes O_2)(\otimes_2 \mathbf{x}) = 0,
\end{aligned}$$

or, using the bra-ket notation:

$$\begin{aligned}
&\langle\psi\otimes\psi|\,\tfrac{O_1\otimes O_2+O_2\otimes O_1}{2}\,|\psi\otimes\psi\rangle -\\
&\langle\psi\otimes\psi|\,\tfrac{O_1\otimes I+I\otimes O_1}{2}\,|\psi\otimes\psi\rangle\\
&\cdot\,\langle\psi\otimes\psi|\,\tfrac{I\otimes O_2+O_2\otimes I}{2}\,|\psi\otimes\psi\rangle\\
&=\langle\psi\otimes\psi|\,(O_1\otimes O_2)\,|\psi\otimes\psi\rangle -\\
&\langle\psi\otimes\psi|\,(O_1\otimes I)\,|\psi\otimes\psi\rangle\\
&\cdot\,\langle\psi\otimes\psi|\,(I\otimes O_2)\,|\psi\otimes\psi\rangle = 0.
\end{aligned}$$

This shows that the difference in (48) is zero when we account for the constant 1/2 in the computations. To sum up, entangled-I states satisfy the factorisation of expectations for (symmetric) local operators, once we define these operators within the equivalence class of the local operators for distinguishable particles.

6.2. Effective distinguishability

As done in [24], we consider now particles described by an "external" degree of freedom with values L, R, and an "internal" degree of freedom with values $\uparrow,\downarrow$. Therefore, the single-particle Hilbert space is spanned by the orthogonal states $|L,\uparrow\rangle, |L,\downarrow\rangle, |R,\uparrow\rangle, |R,\downarrow\rangle$. Among the vectors of the two-particle symmetrized Hilbert space, we focus on those of the form [24]:

$$\left|\phi_1^{\text{ind}}\right\rangle = \frac{1}{\sqrt{2}}\Big[\,|L,\uparrow\rangle\otimes|R,\downarrow\rangle + |R,\downarrow\rangle\otimes|L,\uparrow\rangle\,\Big]. \tag{48}$$

Suppose L and R are orthogonal single-particle states, such as when particles are in separate spatial regions (left and right), so that $\langle L\,|R\rangle = 0$. Then the external degree of freedom, which can be identified as position, can be used to distinguish the particles, as if they were labelled. Indeed, by measuring and then fixing the spatial degree of freedom, one can assign the particle in the region L as particle 1, and the particle in R, as particle 2.

The label attribution is incompatible with particle identity. However, with respect to the observables $\mathcal{P}(P_L\otimes O_1, P_R\otimes O_2)$, where $P_{L,R}$ are the projectors $|L\rangle\langle L|$ and $|R\rangle\langle R|$ onto states localised in the left, respectively right region, while O_1 and O_2 are arbitrary observables relative to the internal degree of freedom, we cannot distinguish the states satisfying Eq. (48) from those of the form

$$\left|\phi_1^{\text{dist}}\right\rangle = |L,\uparrow\rangle\otimes|R,\downarrow\rangle\,. \tag{49}$$

In other words, the mean values of the observables $\mathcal{P}(P_L \otimes O_1, P_R \otimes O_2)$ satisfy the equalities:

$$\begin{aligned}
&\left\langle \phi_1^{\text{ind}} \right| \mathcal{P}\big(P_L \otimes O_1, P_R \otimes O_2\big) \left| \phi_1^{\text{ind}} \right\rangle = \\
&= \langle L, \uparrow | \, P_L \otimes O_1 \, | L, \uparrow \rangle \, \langle R, \downarrow | \, P_R \otimes O_2 \, | R, \downarrow \rangle \\
&= \left\langle \phi_1^{\text{dist}} \right| \big(P_L \otimes O_1\big) \otimes \big(P_R \otimes O_2\big) \left| \phi_1^{\text{dist}} \right\rangle \ . \qquad (50)
\end{aligned}$$

The same equality is also true for

$$\left| \phi_2^{\text{dist}} \right\rangle = |R, \downarrow\rangle \otimes |L, \uparrow\rangle \, , \qquad (51)$$

as

$$\begin{aligned}
&\left\langle \phi_1^{\text{ind}} \right| \mathcal{P}\big(P_L \otimes O_1, P_R \otimes O_2\big) \left| \phi_1^{\text{ind}} \right\rangle = \\
&= \left\langle \phi_2^{\text{dist}} \right| \big(P_R \otimes O_2\big) \otimes \big(P_L \otimes O_1\big) \left| \phi_2^{\text{dist}} \right\rangle , \qquad (52)
\end{aligned}$$

where $\left| \phi_1^{\text{dist}} \right\rangle , \left| \phi_2^{\text{dist}} \right\rangle$ are the "non-symmetric" parts of $\left| \phi_1^{\text{ind}} \right\rangle$.

According to [52], this is an argument against Entanglement-I. The reason is that one can construct two-particle states of the form (48), therefore Entangled-I, that, by using spatial orthogonality, can be made effectively correspond to distinguishable particle separable states, such as those satisfying Eq. (49) (or Eq. (52)). Therefore, [52] claims that distinguishable particle-entanglement cannot be recovered from Entanglement-I.

6.2.1. Analogy with finite/infinite exchangeability

Hereafter, we leverage again the similarities between finite exchangeability and indistinguishability to clarify this counter-example. The derivations presented in Section 5 need a modification to account for the inherent partial exchangeability of this example.

We make the analogy using a pair of coins, and we assume that the possibility space is given by:

$$\Big\{ (\mathrm{L})(\uparrow), (\mathrm{L})(\downarrow), (\mathrm{R})(\uparrow), (\mathrm{R})(\downarrow) \Big\}$$

We have two types of coins, one with faces L/R and the other with faces $\uparrow$ / $\downarrow$. In this case, by $t_1, t_2, \ldots, t_r$, we denote the results of r-tosses of the pair, such as for example:

$$\underset{t_1}{(\mathrm{R})(\uparrow)} \quad \underset{t_2}{(\mathrm{L})(\uparrow)} \quad \underset{t_3}{(\mathrm{L})(\uparrow)}.$$

We assume that the sequence $t_1, t_2, \ldots, t_r$ is finitely exchangeable. For two tosses of the pair, this for instance means that

$$P\left(\text{(R)(}\uparrow\text{)}, \text{(L)(}\uparrow\text{)}\right) = P\left(\text{(L)(}\uparrow\text{)}, \text{(R)(}\uparrow\text{)}\right),$$
$$P\left(\text{(R)(}\uparrow\text{)}, \text{(L)(}\downarrow\text{)}\right) = P\left(\text{(L)(}\downarrow\text{)}, \text{(R)(}\uparrow\text{)}\right).$$

Technically, this is a definition of *finite partial exchangeability* as proposed by de Finetti in [53]. Noticed that the definition assumes the pair being exchangeable, not the single coins. This means that in general:

$$P\left(\text{(R)(}\uparrow\text{)}, \text{(L)(}\downarrow\text{)}\right) \neq P\left(\text{(L)(}\uparrow\text{)}, \text{(R)(}\downarrow\text{)}\right), \tag{53}$$

where we have exchanged only the first coin in the pair. We then adapt Proposition 2 as follows. Let θ_L, θ_R be the probability for face-L and face-R and $\theta_\uparrow, \theta_\downarrow$ the probability for face-$\uparrow$ and face-$\downarrow$. Then we have that:

$$\begin{aligned} &P(t_1, \ldots, t_r) \\ &= \int_\Theta (\theta_L\theta_\uparrow)^{n_1}(\theta_L\theta_\downarrow)^{n_2}(\theta_R\theta_\uparrow)^{n_3}(\theta_R\theta_\downarrow)^{n_4} d\nu(\boldsymbol{\theta}), \end{aligned} \tag{54}$$

where n_1 is the number of times the pair landed on $L, \uparrow$, n_2 is the number of times the pair landed on $L, \downarrow$ and so on.

To make the connection with QT in the context under consideration, we introduce a complex vector $\mathbf{x} = [x_{11}, x_{12}, x_{21}, x_{22}]^\top$ in $\mathbb{C}^4$, such that $\mathbf{x}^\dagger\mathbf{x} = 1$. We can then express the probabilities of the four outcomes of the two coins as:

$$[\theta_L\theta_\uparrow, \theta_L\theta_\downarrow, \theta_R\theta_\uparrow, \theta_R\theta_\downarrow] = [x_{11}^\dagger x_{11}, x_{12}^\dagger x_{12}, x_{21}^\dagger x_{21}, x_{22}^\dagger x_{22}].$$

and consider $r = 2$ rolls.

Observe that, expressed in the "language of coins", the assumption that the particles are in separate spatial regions (left and right) means that $\widehat{L}(\theta_L\theta_\uparrow\theta_L\theta_\uparrow) = \widehat{L}((x_{11}^\dagger x_{11})^2) = 0$, and also $\widehat{L}((x_{11}^\dagger x_{11})(x_{12}^\dagger x_{12})), \widehat{L}((x_{12}^\dagger x_{12})^2)$ are all zeros (similarly for the quasi-expectations obtained replacing θ_L by θ_R). In other words, we assign zero quasi-expectation to the cases in which the coin with face L/R lands on the same side, either L or R. There remain only eight cases:

$$\left(\text{(L)(}\uparrow\text{)}, \text{(R)(}\uparrow\text{)}\right), \left(\text{(L)(}\downarrow\text{)}, \text{(R)(}\downarrow\text{)}\right), \left(\text{(L)(}\uparrow\text{)}, \text{(R)(}\downarrow\text{)}\right), \left(\text{(L)(}\downarrow\text{)}, \text{(R)(}\uparrow\text{)}\right),$$
$$\left(\text{(R)(}\uparrow\text{)}, \text{(L)(}\uparrow\text{)}\right), \left(\text{(R)(}\downarrow\text{)}, \text{(L)(}\downarrow\text{)}\right), \left(\text{(R)(}\downarrow\text{)}, \text{(L)(}\uparrow\text{)}\right), \left(\text{(R)(}\uparrow\text{)}, \text{(L)(}\downarrow\text{)}\right).$$

The elements within the same column are exchangeable. By leveraging the spatial degrees of freedom to distinguish the two pairs of coins, one effectively uses the L/R outcomes as contextual variables to label the coin showing the $\uparrow$ or $\downarrow$ face as "left" or "right". In other words, the outcome L/R is used to select one of the rows above. Since the probabilities associated with elements in each row are not finitely exchangeable (as noted in (53)), they are not subject to any exchangeability constraints.

Therefore, the argument in [24] against "Entanglement I" (that is, that one can construct two-particle states Entangled-I such that, by exploiting spagtial orthogonality, they effectively correspond to separable (non-entangled) states of distinguishable particles) appears rather weak. As we have shown, exchangeability imposes constraints only on the columns, not on the rows. Hence, there is no reason why "Entanglement I" (finite exchangeability) should constrain the rows.

However, we agree with [24] that, the equivalence in Theorem 5 solely provides a mathematical representation of indistinguishability and, as such, we should not use it directly to define a notion of entanglement. Entanglement should not depend on the definition of objective reality that the modeller has in their mind. However, we believe that this mathematical representation (and the analogy between classical finite exchangeability and quantum theory) can help to better understand and compare the various definitions of entanglement for indistinguishable particles.

7. Conclusions

We have established a link between the mathematical representation of finite and infinite exchangeability in classical probability and the mathematical model of indistinguishable bosons in QT. We have demonstrated the parallelism between approaches from two 'independent' research communities: probability theorists and quantum physicists. Both communities have developed related mathematical models, with similar structure and purpose, to tackle analogous problems, such as negative probabilities, symmetries and bounds with respect to the classical probability limit. They both have also obtained finite dimensional representations, via $L, \widehat{L}$ and ρ, for computing probabilities instead of relying on infinite dimensional objects such as the signed measure ν.

We believe that what above points to a deeper and fundamental connection between the two areas. For instance, the representations in Propositions

2 and 3 can actually be seen as specific instances of representations in Generalised Probabilistic Theory [54–82]. Moreover, as discussed in Section 6 (Application), the strong affinities between the two perspectives can help to clarify the definition of entanglement for indistinguishable particles. As future work, in addition to digging into the aforementioned two research avenues, we plan to explore how the interpretation of density matrices in QT as quasi-expectations might yield new insights into quantum information theory. For example, we believe that this interpretation and connection could prove useful in comprehending measures of uncertainty developed in quantum information theory, such as Von Neumann's entropy and quantum discord, and why, surprisingly in certain instances (quantum discord for separable density matrices), they deviate from their classical counterparts. Finally, we intend to investigate the feasibility of extending our results to fermions.

Appendix A. Quasi-expectation operator

Consider a vector of variables $\mathbf{x}$ taking value in the possibility space Ω, and a vector space $\mathcal{F}$ of real-valued bounded functions on $\mathbf{x}$ including the constants. Let $\mathcal{F}^+$ be the closed convex cone [8] of nonnegative functions in $\mathcal{F}$.

An expectation operator is defined as follows.

Definition 3 ([83, Sec. 2.8.4])**.** *Let $\widehat{L}$ be a linear functional $\widehat{L} : \mathcal{F} \to \mathbb{R}$. $\widehat{L}$ is an **expectation operator** if it satisfies the following property:*

$$\widehat{L}(g) \geq \sup c \ \textit{s.t.}\ g - c \in \mathcal{F}^+, \tag{A}$$

for every $g \in \mathcal{F}$ and c is the constant function of value c.

It can be easily verified that (A) is equivalent to:

$$L(g) \geq \inf_{\mathbf{x}\in\Omega} g(\mathbf{x}). \tag{A.1}$$

In the following, to simplify the notation, we simply write $\inf g$ instead of $\inf_{\mathbf{x}\in\Omega} g(\mathbf{x})$. Linearity and (A.1) are the two properties that define a classical expectation operator. Indeed, from these two properties, we can derive that

[8] A subset $\mathcal{C}$ of a real-vector space $\mathcal{F}$ is a cone if for each $f \in \mathcal{F}$ and positive scalar $\alpha > 0$, the element αf is in $\mathcal{C}$. A cone $\mathcal{C}$ is a convex cone if $\alpha f + \beta g$ belongs to $\mathcal{C}$, for any scalars $\alpha, \beta > 0$ and $f, g \in \mathcal{F}$.

- $L(0) = 0$;
- $0 \overset{A}{=} L(0) = L(g - g) \overset{linearity}{=} L(g) + L(-g)$ and so $L(g) = -L(-g)$;

which, together with $L(-g) \geq \inf -g = -\sup g$, leads to

$$\inf g \leq L(g) \leq \sup g. \tag{A.2}$$

This means that $L(g)$ is a 'weighted-average' of g: the weights being the probability measure associated to the expectation operator; note in fact that $\inf g \leq \int_\Omega g\, dp \leq \sup g$ for any probability measure p. This formulation in terms of probabilities is not necessary. Indeed, we can more generally work with expectation operators.

A quasi-expectation operator is instead defined as follows.

Definition 4. *Let $\widehat{L}$ be a linear functional $\widehat{L} : \mathcal{F} \to \mathbb{R}$ and $\mathcal{C}^+$ be a closed convex cone (including the constants) such that $\mathcal{C}^+ \subset \mathcal{F}^+$. We call $\hat{L}$ a* ***quasi-expectation*** *operator (QEO) if it satisfies*

$$\widehat{L}(g) \geq \sup c \ \textit{s.t.}\ g - c \in \mathcal{C}^+, \tag{A*}$$

for every $g \in \mathcal{F}$.

Let $\underline{c}_g$ be equal to the supremum value of c such that $g - c \in \mathcal{C}^+$. It can be verified that (A*) implies:

$$\widetilde{L}(g) \geq \underline{c}_g, \ \text{where} \inf_{\mathbf{x}\in\Omega} g(\mathbf{x}) \geq \underline{c}_g. \tag{A.3}$$

A QEO (conservatively) relaxes property (A) by providing a lower bound $\underline{c}_g$ to $\inf g$.

From property (A*), similarly to what was done for expectation operators, we can derive

$$\underline{c}_g \leq \widetilde{L}(g) \leq \overline{c}_g, \tag{A.4}$$

with $\overline{c}_g$ be equal to the infimum value of c such that $c - g \in \mathcal{C}^+$, where

$$\underline{c}_g \leq \inf g \leq \sup g \leq \overline{c}_g. \tag{A.5}$$

Notice that since the external inequalities of Eq. (A.5) can be strict for some g, we cannot in general define a QEO as an integral with respect to a probability measure and, therefore, a QEO cannot be a 'weighted average'.

In other words, a QEO is not a classical expectation operator. In general, in order to write a QEO as an integral and satisfy (A.5), we need to introduce some negative values:

$$\widehat{L}(g) = \int_{\Omega} g \, d\nu,$$

where ν is a signed-measure.

Appendix B. Proofs

Appendix B.1. Proof of Theorem 4

We will first prove the first statement.

About the trace, by exploiting linearity of $\widehat{L}$, note that

$$Tr(M) = \widehat{L}\Big((\otimes_r \mathbf{x})^\dagger I (\otimes_r \mathbf{x})\Big)$$

where I is the identity matrix. Remember that, by an abuse of notation, by c we denote the constant polynomial of value c. For $g(\mathbf{x}, \mathbf{x}^\dagger) = (\otimes_r \mathbf{x})^\dagger I (\otimes_r \mathbf{x})$, we have that $g - c$ is equal to $(1-c)(\otimes_r \mathbf{x})^\dagger I (\otimes_r \mathbf{x})$. This polynomial belongs to $\mathcal{S}_r^+$ provided that the matrix $(1-c)I$ is PSD. Therefore, $\sup c$, such that $(1-c)I$ is PSD, is equal to $c = 1$. Similarly, we can consider the polynomial $-g - c$, with g as before, and note that $\sup c$ is equal to -1. Therefore, we have shown that

$$\widehat{L}(g) \geq 1, \quad \widehat{L}(-g) = -\widehat{L}(g) \geq -1$$

which implies that $\widehat{L}(g) = 1$. For the proof about M being PSD, note that for any polynomial $(\otimes_r \mathbf{x})^\dagger G (\otimes_r \mathbf{x})$ with G PSD, we have that $\widehat{L}(g) \geq 0$. This implies that $Tr(GM) \geq 0$ for all PSD matrices G. From a well-known result in linear algebra this implies that M is PSD.

For the second part of the theorem, we first notice that $(\otimes_r \mathbf{x})^\dagger D (\otimes_r \mathbf{x}) = \mathbf{1}^\top D(\otimes_r \boldsymbol{\theta})$, where $\mathbf{1}$ is a column vector of ones of suitable dimension. From (14), this allows us to rewrite

$$P(t_1, \ldots, t_r) = L(\theta_1^{n_1} \theta_2^{n_2} \cdots \theta_5^{n_5} \theta_6^{n_6}) = L(\mathbf{1}^\top D(\otimes_r \boldsymbol{\theta})).$$

We will now prove that, for each fixed L, there exists one $\widehat{L}$ such that $L(\mathbf{1}^\top D(\otimes_r \boldsymbol{\theta})) = \widehat{L}((\otimes_r \mathbf{x})^\dagger D(\otimes_r \mathbf{x}))$ for each diagonal D. This shows that we can represent any finitely exchangeable P via $\widehat{L}$ and that, moreover, the relative P is nonnegative and normalised.

The result is obtained by showing that, for any polynomial $g(\otimes_r \boldsymbol{\theta}) = \mathbf{1}^\top D(\otimes_r \boldsymbol{\theta}) = g(\mathbf{x}, \mathbf{x}^\dagger) = (\otimes_r \mathbf{x})^\dagger D(\otimes_r \mathbf{x})$ with D diagonal, having defined

$$L(g) \geq \sup c_1 \text{ s.t. } g - c_1 \in \mathcal{B}_r^+, \tag{B.1}$$

and

$$\widehat{L}(g) \geq \sup c_2 \text{ s.t. } g - c_2 \in \mathcal{S}_r^+, \tag{B.2}$$

we have that $c_1 = c_2$.

This is immediate, because $g - c_2 = (\otimes_r \mathbf{x})^\dagger (D - c_2 I)(\otimes_r \mathbf{x})$. This polynomial is in the cone of $\mathcal{S}_r^+$ when the coefficients $diag(D - c_2 I)$ are all nonnegative. The same condition, being nonnegative, holds for the coefficients of the polynomial in $\boldsymbol{\theta}$. Note that, we can write any polynomial $g(\otimes_r \boldsymbol{\theta})$ as $\mathbf{1}^\top D(\otimes_r \boldsymbol{\theta})$, where $\mathbf{1}$ is the column vector of ones. Therefore, $g - c_1 = \mathbf{1}^\top (D - c_1 I)(\otimes_r \boldsymbol{\theta})$. Since $\mathbf{1}^\top (D - c_1 I) = diag(D - c_1 I)$, the polynomial $g - c_1$ belongs to the cone $\mathcal{B}_r^+$ whenever the coefficients $diag(D - c_1 I)$ are all nonnegative, equivalently when $D - c_1 I$ is PSD. Then the statement follows by convexity of the cones $\mathcal{B}_r^+, \mathcal{S}_r^+$ and Proposition 3.

Appendix B.2. Equivalent definitions of Bose-symmetric density matrix

Proposition 6. *The following definitions of Bose-symmetric density matrix ρ (for r particles) are equivalent:*

$$\rho = \Pi_{sym} \rho \Pi_{sym} \tag{B.3}$$

$$\rho = \Pi_i \rho = \rho \Pi_i^\dagger \text{ for each permutation matrix } \Pi_i. \tag{B.4}$$

We start the proof by showing that condition (B.4) implies condition (B.3). From $\rho = \Pi_i \rho$ for each Π_i, we obtain that $\rho = \frac{1}{N!} \sum_i \Pi_i \rho$. Then, by the definition (24) stating that $\Pi_{sym} := \frac{1}{N!} \sum_i \Pi_i$ it follows that

$$\rho = \Pi_{sym} \rho. \tag{B.5}$$

Similarly, from $\rho = \rho \Pi_i^\dagger$ and the fact that $(\cdot)^\dagger$ distributes over sum, for each Π_i, we get that $\rho = \rho(\frac{1}{N!} \sum_i \Pi_i^\dagger) = \rho(\frac{1}{N!} \sum_i \Pi_i)^\dagger$. By applying definition (24) and since $\Pi_{sym}^\dagger = \Pi_{sym}$ (see for instance [30, B-2-c]), from (B.4), we have that

$$\rho = \rho \Pi_{sym}. \tag{B.6}$$

From equation(B.5), by multiplying both sides by Π_{sym} from the right, we obtain that $\rho\Pi_{sym} = \Pi_{sym}\rho\Pi_{sym}$. Finally, using Eq. (B.6), we obtain that:

$$\rho = \Pi_{sym}\rho\Pi_{sym}.$$

To verify that condition (B.3) implies condition (B.4), let us assume the former. But since $\Pi_i\Pi_{sym} = \Pi_{sym}$ (see for instance [30, B-2-c]), $\Pi_{sym}^\dagger = \Pi_{sym}$ and therefore $\Pi_{sym} = (\Pi_i\Pi_{sym})^\dagger = \Pi_{sym}^\dagger\Pi_i^\dagger = \Pi_{sym}\Pi_i^\dagger$, it is immediate to derive that

$$\Pi_i\rho = \Pi_i\Pi_{sym}\rho\Pi_{sym} = \Pi_{sym}\rho\Pi_{sym} = \rho,$$

and that

$$\rho\Pi_i^\dagger = \Pi_{sym}\rho\Pi_{sym}\Pi_i^\dagger = \Pi_{sym}\rho\Pi_{sym} = \rho.$$

Appendix B.3. Proof Theorem 5

In order to prove Theorem 5, we need to introduce some notation and state a lemma.

First, we recall some results from [84]. They will allow us to rewrite any density matrix of r-particles as $\rho = \widehat{L}((\mathbf{z}_1 \otimes \mathbf{z}_2 \otimes \cdots \otimes \mathbf{z}_r)(\mathbf{z}_1 \otimes \mathbf{z}_2 \otimes \cdots \otimes \mathbf{z}_r)^\dagger)$ where $\mathbf{z}_i \in \mathbb{C}^n$ (n=2 for a qubit) such that $\mathbf{z}_i^\dagger\mathbf{z}_i = 1$, where $\widehat{L}$ is a linear operator from the monomials represented by the elements of the matrix $(\mathbf{z}_1 \otimes \mathbf{z}_2 \otimes \cdots \otimes \mathbf{z}_r)(\mathbf{z}_1 \otimes \mathbf{z}_2 \otimes \cdots \otimes \mathbf{z}_r)^\dagger$ to $\mathbb{C}$, such that the matrix $\widehat{L}((\mathbf{z}_1 \otimes \mathbf{z}_2 \otimes \cdots \otimes \mathbf{z}_r)(\mathbf{z}_1 \otimes \mathbf{z}_2 \otimes \cdots \otimes \mathbf{z}_r)^\dagger)$ is PSD and with trace one ($\widehat{L}$ preserves the constants).

We can equivalently rewrite the elements of ρ_{ij} as $(y_{\boldsymbol{\alpha},\boldsymbol{\beta}})_{\boldsymbol{\alpha},\boldsymbol{\beta}} := (\widehat{L}(\mathbf{z}^{\boldsymbol{\alpha}}(\mathbf{z}^\dagger)^{\boldsymbol{\beta}}))_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ where $\mathbf{z} := [\mathbf{z}_1^\top, \mathbf{z}_2^\top, \ldots, \mathbf{z}_r^\top]^\top$ and $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are a binary encoding of the indexes i and, respectively j. In particular, we have $\boldsymbol{\alpha} := [\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r]$ where $\boldsymbol{\alpha}_k$ is a vector in $\mathbb{N}^n$ which has all zero entries expect one element which is equal to one. The notation $\mathbf{z}^{\boldsymbol{\alpha}}$ must be interpreted as $\prod_{i=1}^r \prod_{j=1}^n z_{ij}^{\alpha_{ij}}$.

Example 10. *For instance, for $n = 2$ and $r = 2$ (two particles), we have that $\mathbf{z} := [z_{11}, z_{12}, z_{21}, z_{22}]^\top$. We can then rewrite for instance ρ_{11} as $\widehat{L}(z_{11}z_{21}(z_{11}z_{21})^\dagger) = \widehat{L}(\mathbf{z}^{\boldsymbol{\alpha}}(\mathbf{z}^\dagger)^{\boldsymbol{\beta}}) = \widehat{L}(\mathbf{z}^{[1,0,1,0]}(\mathbf{z}^\dagger)^{[1,0,1,0]})$, where $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2]$ with $\boldsymbol{\alpha}_1 = [1, 0]$ and $\boldsymbol{\alpha}_2 = [1, 0]$, and $\boldsymbol{\beta} = \boldsymbol{\alpha}$. Note that, we have*

$$\mathbf{z}^{[1,0,1,0]} = z_{11}^1 z_{12}^0 z_{21}^1 z_{22}^0 = z_{11}z_{21}$$
$$(\mathbf{z}^\dagger)^{[1,0,1,0]} = z_{11}^\dagger z_{21}^\dagger.$$

Introducing the notation $(y_{\boldsymbol{\alpha},\boldsymbol{\beta}})_{\boldsymbol{\alpha},\boldsymbol{\beta}} := (\widehat{L}(\mathbf{z}^{\boldsymbol{\alpha}}(\mathbf{z}^\dagger)^{\boldsymbol{\beta}}))_{\boldsymbol{\alpha},\boldsymbol{\beta}}$*, we have that the elements of the matrix* ρ*, for the example* $n = r = 2$*, can be denoted as follows:*

$$\rho = \begin{bmatrix} y_{[1,0,1,0],[1,0,1,0]} & y_{[1,0,1,0],[1,0,0,1]} & y_{[1,0,1,0],[0,1,1,0]} & y_{[1,0,1,0],[0,1,0,1]} \\ y_{[1,0,0,1],[1,0,1,0]} & y_{[1,0,0,1],[1,0,0,1]} & y_{[1,0,0,1],[0,1,1,0]} & y_{[1,0,0,1],[0,1,0,1]} \\ y_{[0,1,1,0],[1,0,1,0]} & y_{[0,1,1,0],[1,0,0,1]} & y_{[0,1,1,0],[0,1,1,0]} & y_{[0,1,1,0],[0,1,0,1]} \\ y_{[0,1,0,1],[1,0,1,0]} & y_{[0,1,0,1],[1,0,0,1]} & y_{[0,1,0,1],[0,1,1,0]} & y_{[0,1,0,1],[0,1,0,1]} \end{bmatrix}.$$

where $y_{[1,0,1,0],[1,0,1,0]} = \widehat{L}(z_{11}z_{21}(z_{11}z_{21})^\dagger)$ *and so on.*

If ρ is the density matrix of r indistinguishable bosons, then it has to satisfy the constraint $\Pi_{sym}\rho\Pi_{sym} = \rho$. We can then prove the following lemma.

Lemma 1. *Let* $\rho_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \widehat{L}(\mathbf{z}^{\boldsymbol{\alpha}}(\mathbf{z}^\dagger)^{\boldsymbol{\beta}}) = y_{\boldsymbol{\alpha},\boldsymbol{\beta}}$*. If* ρ *satisfies* $\Pi_{sym}\rho\Pi_{sym} = \rho$ *then for all the indexes* $\boldsymbol{\alpha}, \boldsymbol{\beta}$*, the elements* $\rho_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ *have to satisfy the following constraints:*

$$\begin{aligned} &y_{[\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_r],[\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_r]} = y_{[\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_r],[\boldsymbol{\beta}_{\pi(1)},\ldots,\boldsymbol{\beta}_{\pi(r)}]} \\ &= y_{[\boldsymbol{\alpha}_{\pi(1)},\ldots,\boldsymbol{\alpha}_{\pi(r)}],[\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_r]} = y_{[\boldsymbol{\alpha}_{\pi(1)},\ldots,\boldsymbol{\alpha}_{\pi(r)}],[\boldsymbol{\beta}_{\pi(1)},\ldots,\boldsymbol{\beta}_{\pi(r)}]} \end{aligned}$$

for any permutation π *of the labels of the variables.*

Proof. This follow from the following known property of the symmetriser:

$$\Pi_{sym}P_\pi = P_\pi\Pi_{sym} = \Pi_{sym}$$

where P_π is any permutation matrix (defined by the permutation π of the labels) associated with the r particles. Note that, previously in the manuscript, we denoted permutation matrices with Π_i instead of P_π. Here we use P_π to state that this is the permutation matrix relative to the label permutation π. Given that $\rho = \Pi_{sym}\rho\Pi_{sym}$, this means that

$$\begin{aligned} P_\pi\rho &= P_\pi\Pi_{sym}\rho\Pi_{sym} = \Pi_{sym}\rho\Pi_{sym}P_\pi \\ &= P_\pi\Pi_{sym}\rho\Pi_{sym}P_\pi = \Pi_{sym}\rho\Pi_{sym} = \rho. \end{aligned}$$

Therefore, by linearity of $\widehat{L}$, we have that

$$\begin{aligned} \rho &:= \widehat{L}((\mathbf{z}_1 \otimes \mathbf{z}_2 \otimes \cdots \otimes \mathbf{z}_r)(\mathbf{z}_1 \otimes \mathbf{z}_2 \otimes \cdots \otimes \mathbf{z}_r)^\dagger) \\ &= P_\pi\widehat{L}((\mathbf{z}_1 \otimes \mathbf{z}_2 \otimes \cdots \otimes \mathbf{z}_r)(\mathbf{z}_1 \otimes \mathbf{z}_2 \otimes \cdots \otimes \mathbf{z}_r)^\dagger) \\ &= \widehat{L}(P_\pi(\mathbf{z}_1 \otimes \mathbf{z}_2 \otimes \cdots \otimes \mathbf{z}_r)(\mathbf{z}_1 \otimes \mathbf{z}_2 \otimes \cdots \otimes \mathbf{z}_r)^\dagger) \\ &= \widehat{L}((\mathbf{z}_1 \otimes \mathbf{z}_2 \otimes \cdots \otimes \mathbf{z}_r)(\mathbf{z}_1 \otimes \mathbf{z}_2 \otimes \cdots \otimes \mathbf{z}_r)^\dagger P_\pi) \\ &= \widehat{L}(P_\pi(\mathbf{z}_1 \otimes \mathbf{z}_2 \otimes \cdots \otimes \mathbf{z}_r)(\mathbf{z}_1 \otimes \mathbf{z}_2 \otimes \cdots \otimes \mathbf{z}_r)^\dagger P_\pi) \end{aligned}$$

for every P_π. Note that, the permutation matrix rearranges the $\boldsymbol{\alpha}$-th element of the vector $\mathbf{z}_1 \otimes \mathbf{z}_2 \otimes \cdots \otimes \mathbf{z}_r$ according to the following rule:

$$P_\pi(\mathbf{z}_1 \otimes \cdots \otimes \mathbf{z}_r)_{[\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_r]} = (\mathbf{z}_1 \otimes \cdots \otimes \mathbf{z}_r)_{[\boldsymbol{\alpha}_{\pi(1)},\ldots,\boldsymbol{\alpha}_{\pi(r)}]}$$
$$= \mathbf{z}^{[\boldsymbol{\alpha}_{\pi(1)},\ldots,\boldsymbol{\alpha}_{\pi(r)}]}$$

which ends the proof. □

For the example in (B.7), the results of the above lemma implies for instance that

$$y_{[1,0,0,1],[1,0,0,1]} = y_{[1,0,0,1],[0,1,1,0]}$$
$$= y_{[0,1,1,0],[1,0,0,1]} = y_{[0,1,1,0],[0,1,1,0]}.$$

We are now ready to prove Theorem 5. First we prove that $\mathcal{S}_2 \subseteq \mathcal{S}_1$. Note that

$$\Pi_{sym} M \Pi_{sym} = \widehat{L}\Big(\Pi_{sym}(\otimes_r \mathbf{x})(\otimes_r \mathbf{x})^\dagger \Pi_{sym}^\dagger\Big)$$
$$= \widehat{L}\Big((\otimes_r \mathbf{x})(\otimes_r \mathbf{x})^\dagger\Big) = M, \tag{B.7}$$

where we have exploited (24) to derive $\Pi_{sym}(\otimes_r \mathbf{x}) = \otimes_r \mathbf{x}$. Note in fact that, $\otimes_r \mathbf{x}$ is clearly symmetric to permutations, being the r-times tensor product of $\mathbf{x}$.

We will now prove that $\mathcal{S}_1 \subseteq \mathcal{S}_2$. This part is based on the same insight that motivates the second quantisation formalism in QT. We start from Lemma 1 and note that in

$$y_{[\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_r],[\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_r]} = y_{[\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_r],[\boldsymbol{\beta}_{\pi(1)},\ldots,\boldsymbol{\beta}_{\pi(r)}]}$$
$$= y_{[\boldsymbol{\alpha}_{\pi(1)},\ldots,\boldsymbol{\alpha}_{\pi(r)}],[\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_r]} = y_{[\boldsymbol{\alpha}_{\pi(1)},\ldots,\boldsymbol{\alpha}_{\pi(r)}],[\boldsymbol{\beta}_{\pi(1)},\ldots,\boldsymbol{\beta}_{\pi(r)}]}$$

the permutation acts on the indexing vectors. Each $\boldsymbol{\alpha}_i$ is an indicator vector which denotes a component of the vector $\mathbf{z}_i$ and the permutation simply swaps this indicator vectors. We now introduce the occupation number vector $[k_1, k_2, \ldots, k_n]$ which, for a given $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$, counts the number of variables $\mathbf{z}_i$ that are in their state 1, state 2,..., or state n. All the vectors $[\boldsymbol{\alpha}_{\pi(1)}, \ldots, \boldsymbol{\alpha}_{\pi(r)}]$ have the same occupation number. For instance, for the previous example $y_{[1,0,0,1],[1,0,0,1]} = y_{[1,0,0,1],[0,1,1,0]} = y_{[0,1,1,0],[1,0,0,1]} = y_{[0,1,1,0],[0,1,1,0]}$. The occupation number for $[1,0,0,1]$ is $[1,1]$ because variable 1 is in state 1 and variable 2 in state 2. $[0,1,1,0]$ has the same occupation number. Therefore, we can write $y_{[1,0,0,1],[1,0,0,1]} = y_{[1,0,0,1],[0,1,1,0]} = y_{[0,1,1,0],[1,0,0,1]} =$

$y_{[0,1,1,0],[0,1,1,0]} \to y_{[k_1,k_2][k_1',k_2']}$ with $k_1, k_2; k_1', k_2' = 1$, where we used $'$ to denote the occupation number relative to $\boldsymbol{\beta}$. The representation of $y_{[1,0,1,0],[1,0,1,0]}$ using the occupation number as index is $y_{[2,0],[2,0]}$.

We conclude the proof by noticing that we can represent the elements of $\widehat{L}((\otimes_r \mathbf{x})(\otimes_r \mathbf{x})^\dagger)$ as $(\widehat{L}(\mathbf{x}^{\mathbf{k}}(\mathbf{x}^\dagger)^{\mathbf{k}'}))_{\mathbf{k},\mathbf{k}'} = y_{\mathbf{k},\mathbf{k}'}$, where $\mathbf{k} = [k_1, \ldots, k_n]$ is such that each $k_i \in \mathbb{N}$ denotes the degree of the i-th variable in the monomial. For instance, $\widehat{L}(x_1 x_2 (x_1 x_2)^\dagger) = y_{[1,1],[1,1]}$ and $\widehat{L}(x_1^2 (x_1^2)^\dagger) = y_{[2,0],[2,0]}$.

References

[1] P. Diaconis, Finite forms of de finetti's theorem on exchangeability, Synthese 36 (1977) 271–281.

[2] P. Diaconis, D. Freedman, Finite exchangeable sequences, The Annals of Probability (1980) 745–764.

[3] R. L. Hudson, G. R. Moody, Locally normal symmetric states and an analogue of de finetti's theorem, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 33 (1976) 343–351.

[4] C. M. Caves, C. A. Fuchs, R. Schack, Unknown quantum states: the quantum de finetti representation, Journal of Mathematical Physics 43 (2002) 4537–4559.

[5] R. Renner, Symmetry of large physical systems implies independence of subsystems, Nature Physics 3 (2007) 645–649.

[6] M. Christandl, R. König, G. Mitchison, R. Renner, One-and-a-half quantum de finetti theorems, Communications in mathematical physics 273 (2007) 473–498.

[7] R. König, R. Renner, A de finetti representation for finite symmetric quantum states, Journal of Mathematical physics 46 (2005) 122108.

[8] R. Koenig, G. Mitchison, A most compendious and facile quantum de finetti theorem, Journal of Mathematical Physics 50 (2009) 012105.

[9] R. Renner, Security of quantum key distribution, International Journal of Quantum Information 6 (2008) 1–127.

[10] H.-Y. Su, Improved quantum de finetti theorem and its application in quantum key distribution, Physical Review A 101 (2020) 042310.

[11] F. G. Brandao, M. Piani, P. Horodecki, Generic emergence of classical features in quantum darwinism, Nature communications 6 (2015) 7908.

[12] R. Renner, J. Wang, The black hole information puzzle and the quantum de finetti theorem, arXiv preprint arXiv:2110.14653 (2021).

[13] C. Dellacherie, P.-A. Meyer, Probabilities and potential. b, volume 72 of north-holland mathematics studies, 1982.

[14] E. Jaynes, Some applications and extensions of the de finetti representation theorem, Bayesian Inference and Decision Techniques with Applications: Essays in Honor of Bruno de Finetti. North-Holland Publishers (1982).

[15] G. J. Kerns, G. J. Székely, De Finetti's theorem for abstract finite exchangeable sequences, Journal of Theoretical Probability 19 (2006) 589–608.

[16] Z. Ammari, F. Nier, Mean field limit for bosons and infinite dimensional phase-space analysis, Annales Henri Poincaré 9 (2008) 1503–1574.

[17] M. Lewin, P. T. Nam, N. Rougerie, Derivation of nonlinear gibbs measures from many-body quantum mechanics, Journal of Statistical Physics 157 (2014) 745–779.

[18] Gert de Cooman and Enrique Miranda, Symmetry of models versus models of symmetry, Probability and Inference: Essays in Honor of Henry E. Kyburg, Jr., eds. William Harper and Gregory Wheeler, pp. 67-149, King's College Publications, London, 2007.

[19] G. De Cooman, E. Quaeghebeur, Exchangeability and sets of desirable gambles, International Journal of Approximate Reasoning 53 (2012) 363–395.

[20] A. Benavoli, A. Facchini, M. Zaffalon, Bernstein's socks, polynomial-time provable coherence and entanglement, in: J. D. Bock, C. de Campos, G. de Cooman, E. Quaeghebeur, G. Wheeler (Eds.), ISIPTA ;'19: Proceedings of the Eleventh International Symposium on Imprecise

Probability: Theories and Applications, PJMLR, JMLR, Ghent, Belgium, 2019.

[21] A. Benavoli, A. Facchini, M. Zaffalon, Quantum indistinguishability through exchangeable desirable gambles, in: ISIPTA'21 Int. Symposium on Imprecise Probability: Theories and Applications, PJMLR, 2021.

[22] A. Benavoli, A. Facchini, M. Zaffalon, Quantum indistinguishability through exchangeability, International Journal of Approximate Reasoning 151 (2022) 389–412.

[23] A. C. Doherty, P. A. Parrilo, F. M. Spedalieri, Complete family of separability criteria, Physical Review A 69 (2004) 022308.

[24] F. Benatti, R. Floreanini, F. Franchini, U. Marzolino, Entanglement in indistinguishable particle systems, Physics Reports 878 (2020) 1–27.

[25] B. de Finetti, La prévision: ses lois logiques, ses sources subjectives, Annales de l'Institut Henri Poincaré 7 (1937) 1–68.

[26] B. De Finetti, Funzione caratteristica di un fenomeno aleatorio, in: Memorie della R. Accademia dei Lincei 4, 1930, pp. 86–133.

[27] P. Diaconis, D. Freedman, Finite exchangeable sequences, The Annals of Probability 8 (1980) 745–764.

[28] D. G. Kendall, On finite and infinite sequences of exchangeable events., Technical Report, Johns Hopkins univ Baltimore md dept of statistics, 1966.

[29] B. de Finetti, Sulla proprietà conglomerativa delle probabilità subordinate, Rendiconti del Reale Instituto Lombardo 63 (1930) 414–418.

[30] C. Cohen-Tannoudji, B. Diu, F. Laloë, Quantum Mechanics, Volume 2: Angular Momentum, Spin, and Approximation Methods, John Wiley & Sons, Germany, 2020.

[31] J. B. Lasserre, Moments, positive polynomials and their applications, volume 1, World Scientific, 2009.

[32] R. Horodecki, P. Horodecki, M. Horodecki, K. Horodecki, Quantum entanglement, Reviews of modern physics 81 (2009) 865.

[33] R. Paškauskas, L. You, Quantum correlations in two-boson wave functions, Physical Review A 64 (2001) 042310.

[34] K. Eckert, J. Schliemann, D. Bruß, M. Lewenstein, Quantum correlations in systems of indistinguishable particles, Annals of physics 299 (2002) 88–127.

[35] X. Wang, B. C. Sanders, Canonical entanglement for two indistinguishable particles, Journal of Physics A: Mathematical and General 38 (2005) L67.

[36] D. Zhou, Correlation in quantum states of two identical particles, Europhysics Letters 88 (2009) 10006.

[37] M. Kuś, I. Bengtsson, "classical" quantum states, Physical Review A 80 (2009) 022319.

[38] M. Kotowski, M. Kotowski, M. Kuś, Universal nonlinear entanglement witnesses, Physical Review A 81 (2010) 062318.

[39] J. Grabowski, M. Kuś, G. Marmo, Segre maps and entanglement for multipartite systems of indistinguishable particles, Journal of Physics A: Mathematical and Theoretical 45 (2012) 105301.

[40] J. Grabowski, M. Kuś, G. Marmo, Entanglement for multipartite systems of indistinguishable particles, Journal of Physics A: Mathematical and Theoretical 44 (2011) 175302.

[41] A. Sawicki, M. Oszmaniec, M. Kuś, Critical sets of the total variance can detect all stochastic local operations and classical communication classes of multiparticle entanglement, Physical Review A 86 (2012) 040304.

[42] P. Hyllus, L. Pezzé, A. Smerzi, G. Tóth, Entanglement and extreme spin squeezing for a fluctuating number of indistinguishable particles, Physical Review A 86 (2012) 012337.

[43] M. Oszmaniec, M. Kuś, Universal framework for entanglement detection, Physical Review A 88 (2013) 052328.

[44] T. Wasak, P. Szańkowski, P. Ziń, M. Trippenbach, J. Chwedeńczuk, Cauchy-schwarz inequality and particle entanglement, Physical Review A 90 (2014) 033616.

[45] M. Oszmaniec, M. Kuś, Fraction of isospectral states exhibiting quantum correlations, Physical Review A 90 (2014) 010302.

[46] A. Sawicki, M. Oszmaniec, M. Kuś, Convexity of momentum map, morse index, and quantum entanglement, Reviews in Mathematical Physics 26 (2014) 1450004.

[47] G. Rigolin, Entanglement, identical particles and the uncertainty principle, Communications in Theoretical Physics 66 (2016) 201.

[48] G. García-Calderón, R. Romo, Buildup of symmetrization entanglement for the nonescape probability of two identical particles, Physical Review A 96 (2017) 062124.

[49] M. Karczewski, S.-Y. Lee, J. Ryu, Z. Lasmar, D. Kaszlikowski, P. Kurzyński, Sculpting out quantum correlations with bosonic subtraction, Physical Review A 100 (2019) 033828.

[50] B. Morris, B. Yadin, M. Fadel, T. Zibold, P. Treutlein, G. Adesso, Entanglement between identical particles is a useful and consistent resource, Physical Review X 10 (2020) 041012.

[51] P. Kunkel, M. Prüfer, H. Strobel, D. Linnemann, A. Frölian, T. Gasenzer, M. Gärttner, M. K. Oberthaler, Spatially distributed multipartite entanglement enables epr steering of atomic clouds, Science 360 (2018) 413–416.

[52] F. Benatti, R. Floreanini, K. Titimbo, Entanglement of identical particles, Open Systems & Information Dynamics 21 (2014) 1440003.

[53] B. D. Finetti, On the condition of partial exchangeability (1980) 193–206. doi:doi:10.1525/9780520318328-005.

[54] G. Birkhoff, J. Von Neumann, The logic of quantum mechanics, Annals of mathematics (1936) 823–843.

[55] G. W. Mackey, Mathematical foundations of quantum mechanics, Courier Corporation, California, US, 2013.

[56] J. M. Jauch, C. Piron, Can hidden variables be excluded in quantum mechanics, Helv. Phys. Acta 36 (1963) 827–837.

[57] L. Hardy, Foliable operational structures for general probabilistic theories, Deep Beauty: Understanding the Quantum World through Mathematical Innovation; Halvorson, H., Ed (2011) 409.

[58] L. Hardy, Quantum theory from five reasonable axioms, arXiv preprint quant-ph/0101012 (2001).

[59] J. Barrett, Information processing in generalized probabilistic theories, Physical Review A 75 (2007) 032304.

[60] G. Chiribella, G. M. D'Ariano, P. Perinotti, Probabilistic theories with purification, Physical Review A 81 (2010) 062348.

[61] H. Barnum, A. Wilce, Information processing in convex operational theories, Electronic Notes in Theoretical Computer Science 270 (2011) 3–15.

[62] W. Van Dam, Implausible consequences of superstrong nonlocality, arXiv preprint quant-ph/0501159 (2005).

[63] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, M. Żukowski, Information causality as a physical principle, Nature 461 (2009) 1101.

[64] B. Dakic, C. Brukner, Quantum theory and beyond: Is entanglement special?, arXiv preprint arXiv:0911.0695 (2009).

[65] C. A. Fuchs, Quantum mechanics as quantum information (and only a little more), arXiv preprint quant-ph/0205039 (2002).

[66] G. Brassard, Is information the key?, Nature Physics 1 (2005) 2.

[67] M. P. Mueller, L. Masanes, Information-theoretic postulates for quantum theory, in: Quantum Theory: Informational Foundations and Foils, Springer, 2016, pp. 139–170.

[68] B. Coecke, R. W. Spekkens, Picturing classical and quantum bayesian inference, Synthese 186 (2012) 651–696.

[69] C. M. Caves, C. A. Fuchs, R. Schack, Unknown quantum states: the quantum de Finetti representation, Journal of Mathematical Physics 43 (2002) 4537–4559.

[70] D. Appleby, Facts, values and quanta, Foundations of Physics 35 (2005) 627–668.

[71] D. Appleby, Probabilities are single-case or nothing, Optics and spectroscopy 99 (2005) 447–456.

[72] C. G. Timpson, Quantum Bayesianism: a study, Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 39 (2008) 579–609.

[73] C. A. Fuchs, R. Schack, Quantum-Bayesian coherence, Reviews of Modern Physics 85 (2013) 1693.

[74] C. A. Fuchs, R. Schack, A quantum-bayesian route to quantum-state space, Foundations of Physics 41 (2011) 345–356.

[75] N. D. Mermin, Physics: Qbism puts the scientist back into science, Nature 507 (2014) 421–423.

[76] I. Pitowsky, Betting on the outcomes of measurements: a bayesian theory of quantum probability, Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 34 (2003) 395–414.

[77] I. Pitowsky, Physical Theory and its Interpretation: Essays in Honor of Jeffrey Bub, Springer Netherlands, Dordrecht, 2006, pp. 213–240.

[78] A. Benavoli, A. Facchini, M. Zaffalon, Quantum mechanics: The Bayesian theory generalized to the space of Hermitian matrices, Physical Review A 94 (2016) 042106.

[79] A. Benavoli, A. Facchini, M. Zaffalon, A Gleason-type theorem for any dimension based on a gambling formulation of Quantum Mechanics, Foundations of Physics 47 (2017) 991–1002.

[80] S. Popescu, D. Rohrlich, Causality and nonlocality as axioms for quantum mechanics, in: Causality and Locality in Modern Physics, Springer, 1998, pp. 383–389.

[81] M. Navascués, H. Wunderlich, A glance beyond the quantum model, in: Proceedings of the Royal Society of London A: Mathematical, Physical

and Engineering Sciences, volume 466, The Royal Society, 2010, pp. 881–890.

[82] M. Plávala, General probabilistic theories: An introduction, arXiv preprint arXiv:2103.07469 (2021).

[83] P. Walley, Statistical Reasoning with Imprecise Probabilities, Chapman and Hall, New York, 1991.

[84] A. Benavoli, A. Facchini, M. Zaffalon, The weirdness theorem and the origin of quantum paradoxes, Foundations of Physics 51 (2021).