The de Finetti theorem for test spaces

Jonathan Barrett[1] and Matthew Leifer[2,3]

[1] H H Wills Physics Laboratory, University of Bristol,
Tyndall Avenue, Bristol BS8 1TL, UK
[2] Institute for Quantum Computing, University of Waterloo,
200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada
[3] Perimeter Institute for Theoretical Physics, 31 Caroline Street North,
Waterloo, Ontario, N2L 2Y5, Canada
E-mail: j.barrett@bristol.ac.uk and matt@mattleifer.info

Abstract. We prove a de Finetti theorem for exchangeable sequences of states on *test spaces*, where a test space is a generalization of the sample space of classical probability theory and the Hilbert space of quantum theory. The standard classical and quantum de Finetti theorems are obtained as special cases. By working in a test space framework, the common features that are responsible for the existence of these theorems are elucidated. In addition, the test space framework is general enough to imply a de Finetti theorem for classical processes. We conclude by discussing the ways in which our assumptions may fail, leading to probabilistic models that do not have a de Finetti theorem.

Contents

1. Introduction

There are many scenarios involving probabilistic reasoning about a large number of systems, where it is important that these systems can be regarded as identical and independent. Classical parameter estimation and quantum tomography, in which a source is calibrated by measuring separately a number of systems it has produced, provide excellent examples. More generally, much of experimental science involves repetition of an experiment, with conclusions drawn from observed relative frequencies, where the conclusions are only valid under the assumptions of identity and independence of the separate trials. A problem that arises, therefore, is what justifies these assumptions? De Finetti theorems are designed to answer this question.

De Finetti theorems have a fundamental significance for the subjective Bayesian interpretation of probabilities. They enable the subjective Bayesian to explain why a rational agent treats a sequence of trials as identical and independent, without appealing to a notion of objective chance, or to unknown probabilities. In addition to this fundamental significance, they are also important technical tools. For example, quantum de Finetti theorems have been applied to various problems in quantum information theory, including proofs of the security of quantum key distribution [1], algorithms for deciding the separability of bipartite quantum states [2, 3] and global optimization of the maximum output purity of quantum channels [4].

In this paper, we provide a de Finetti theorem for nonsignalling states on *test spaces*. A test space is a generalization of the sample space of classical probability theory and the Hilbert space of quantum theory. The de Finetti theorem for test spaces includes the classical and quantum theorems as special cases. It also implies a de Finetti theorem for classical processes, which can also be viewed as a theorem about states in theories that exhibit 'superquantum' correlations, as have been studied recently in quantum information [5]–[17].

The remainder of this paper is structured as follows. In section 2, the classical and quantum de Finetti theorems are reviewed and their relevance for Bayesian statistics is discussed. In section 3, the test space framework is introduced and its connection to the convex sets

framework used in [18, 19] is explained. Section 4 states the de Finetti theorem for test spaces and section 5 outlines some of its consequences, including the classical and quantum theorems and the theorem for classical processes. Section 6 discusses the role of the various assumptions of the test space framework and outlines some more general scenarios in which the theorem fails. Section 7 concludes.

2. The classical and quantum de Finetti theorems

De Finetti introduced his theorem in the context of a subjective Bayesian approach to probability theory [20, 21]. In this approach, probabilities are not defined as limiting relative frequencies, nor as objective properties of the physical world. Instead they are measures of the degrees of belief of a decision-making agent. An immediate question is, why should degrees of belief be represented by real numbers obeying the Kolmogorov axioms? De Finetti's answer is provided by his famous Dutch Book argument [20, 21][4]. But a second question is, why do the usual rules of statistical inference apply to these quantities? In particular, how and why should relative frequencies be used to update probability assignments?

Consider an experiment in which a trial with d possible outcomes is repeated n times. Under a standard sort of analysis, the trials are first judged to be independent and identically distributed, so that the probability of getting the outcome sequence $x_1, \ldots, x_n$ is given by $P^n(x_1, \ldots, x_n) = p(x_1) \times \cdots \times p(x_n)$, for some 'unknown' probability distribution p. The distribution p is a parameter to be estimated.

To the subjective Bayesian, however, this is problematic since no sense can be given to an 'unknown' probability. De Finetti provides an alternative analysis. If the trial can in principle be repeated an arbitrary number of times, then the joint distribution over outcome sequences P^n should be defined for any n, so consider an infinite sequence of distributions $P^1, P^2, \ldots$. Suppose that for the first n trials, the agent is indifferent as to whether any further trials are actually performed or not, and that the agent is also indifferent as to the order in which the outcomes are reported. This suggests that the sequence $P^1, P^2, \ldots$ should satisfy the following.

Definition 1. P^n is symmetric *if and only if it is invariant under permutations of the n tests. That is, $P^n(x_1, \ldots, x_n) = P^n(x_{\pi(1)}, \ldots, x_{\pi(n)})$ for all permutations $\pi : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$.*

Definition 2. *The sequence $P^1, P^2, \ldots$ is* exchangeable *if and only if*

1. $\forall n$, P^n is symmetric,

2. $\forall n, \forall x_1, \ldots, x_n$,
   $P^n(x_1, \ldots, x_n) = \sum_{x_{n+1}=1}^{d} P^{n+1}(x_1, \ldots, x_n, x_{n+1})$.

Theorem 1 (de Finetti's representation theorem [21, 24]). *If the sequence $P^1, P^2, \ldots$ is exchangeable, then P^n can be written in the form*

$$P^n(x_1, \ldots, x_n) = \int_{\Delta_d} \mathrm{d}\mu(p) \; p(x_1) \cdots p(x_n), \tag{1}$$

where Δ_d is the set of all probability distributions over the outcomes $\{1, \ldots, d\}$, μ is a probability measure on Δ_d, μ is independent of n and μ is unique.

[4] More sophisticated arguments, framed in terms of decision theory, are advocated by many contemporary subjective Bayesians [22, 23].

This is the classical de Finetti theorem for infinite sequences and a finite number of outcomes. It shows that if the sequence $P^1, P^2, \ldots$ is exchangeable, then P^n can be written *as if* it were generated via a probability distribution μ over unknown probabilities p. In particular, if the first $m < n$ trials are performed, and standard Bayesian updating applied directly to the joint distribution P^n, one finds that the posterior probability for the remaining $n - m$ trials is given by

$$P^n(x_{m+1}, \ldots, x_n | x_1, \ldots, x_m) = \int_{\Delta_d} \mathrm{d}\mu(p | x_1, \ldots, x_m) \, p(x_{m+1}) \cdots p(x_n), \tag{2}$$

where μ is updated as if Bayesian conditioning had been performed on an unknown parameter p directly:

$$\mathrm{d}\mu(p | x_1, \ldots, x_m) = \frac{\mathrm{d}\mu(p) \times p(x_1) \times \cdots \times p(x_m)}{P^n(x_1, \ldots, x_m)}.$$

The quantum de Finetti theorem is a generalization of the classical theorem. It was first presented in [25, 26], and a simpler proof given in [27, 28]. Whereas the classical theorem concerned the outcome probabilities of a test that could be repeated an arbitrarily large number of times, the quantum theorem concerns the joint state of an arbitrarily large number of quantum systems. Suppose that each of these systems is associated with a d-dimensional Hilbert space H_d. The joint state of n systems is then a density operator on the tensor product Hilbert space $H_d^{\otimes n}$. Exchangeability is defined for a sequence of states $\omega^1, \omega^2, \ldots$, where ω^n is a state of n systems.

Definition 3. ω^n is symmetric *if and only if it is invariant under permutations of the n systems, i.e.*

$$\mathrm{Tr}\left(Q_1 \otimes \cdots \otimes Q_n \, \omega^n \right) = \mathrm{Tr}\left(Q_{\pi(1)} \otimes \cdots \otimes Q_{\pi(n)} \, \omega^n \right),$$

for all permutations $\pi : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$, and for any projection operators $Q_1, \ldots, Q_n$.

Note that this is equivalent to requiring that $\omega^n = S_\pi \omega^n S_\pi^\dagger$ for all permutations π, where S_π is the operator that permutes the n systems according to π.

Definition 4. *The sequence $\omega^1, \omega^2, \ldots$ is* exchangeable *if and only if*

1. *$\forall n, \omega^n$ is symmetric,*
2. *$\forall n, \omega^n = \mathrm{Tr}_{n+1}(\omega^{n+1})$, where Tr_{n+1} is the partial trace over the $n + 1$st system.*

Theorem 2 (the quantum de Finetti theorem). *If the sequence $\omega^1, \omega^2, \ldots$ is exchangeable, then ω^n can be written in the form*

$$\omega^n = \int_\Omega \mathrm{d}\mu(\omega) \, \omega \otimes \cdots \otimes \omega, \tag{3}$$

where Ω is the set of density operators on H_d, μ is a probability measure on Ω, μ is independent of n, and μ is unique.

From a fundamental point of view, the quantum de Finetti theorem is particularly important for those approaches to quantum theory that take a subjective view of the quantum state [29]–[33]. These approaches are closely related to the Bayesian view of probabilities. A quantum state is taken to represent the degrees of belief of an agent, where these might be beliefs about the potential outcomes of measurements. In this case, the notion of an unknown quantum state, so prevalent in the literature, becomes problematic. The quantum de Finetti theorem shows how to dispense with this notion, at least in some situations.

Both of the theorems just presented assume finite sample spaces (or finite-dimensional Hilbert spaces), and concern a number of trials or systems that tends to infinity. Both have been generalized in a number of ways. Classical theorems for a finite number of trials are discussed in [34]–[36] and quantum theorems for a finite number of systems in [1], [37]–[40]. In addition to the quantum de Finetti theorem, there is also a de Finetti theorem for quantum operations [28, 41], for representations of unitary groups [40, 42] and for unitarily invariant quantum states [43].

3. Test spaces

The classical de Finetti theorem involves probabilities of outcome sequences for a test that can in principle be repeated an arbitrarily large number of times. The quantum de Finetti theorem involves the joint quantum state of a number of quantum systems that can in principle be arbitrarily large. Both of these can be viewed as special cases of a more general scenario. Suppose that a single system is associated with a number of different possible tests, which are mutually exclusive in the sense that only one can be performed at a time. The classical case is recovered when there is in fact only one such test, and the quantum case when the tests correspond to the different possible measurements on a quantum system. A state is an assignment of probabilities to the outcomes of all the different possible tests.

Assume further that given n systems, a test for each system can be independently chosen, and that a joint state is an assignment of probabilities to outcome sequences for each possible sequence of tests. Given all this, it is possible to define exchangeability for a sequence of states and to prove a de Finetti representation theorem. These ideas are formalized in this section, with the representation theorem given in the next.

3.1. Single systems

The technical notion that we use to describe a single system is that of a *test space*. Test spaces were introduced with the explicit purpose of describing probabilistic models more general than classical and quantum theory, but including both as special cases [44, 45].

Definition 5. *A* test space *consists of a pair* (E, S), *with* E *a set, and* S *a set of countable subsets of* E *that covers* E.

The idea is that each element of E is a possible outcome of a test. Each set $s \in S$ corresponds to a possible test, with the elements of s being the outcomes of that test. The sets in S may overlap, thus the definition of a test space is designed to allow for the possibility that outcomes
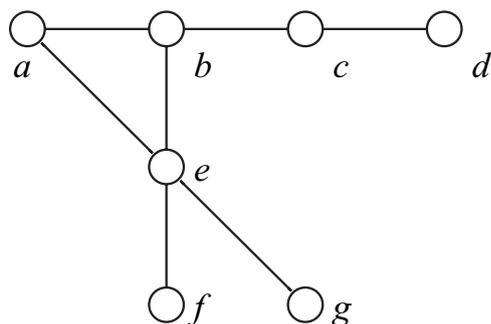
Figure 1. A Greechie diagram of the test space $(\{a, b, c, d, e, f, g\}, \{\{a, b, c, d\}, \{a, e, g\}, \{b, e, f\}\})$.
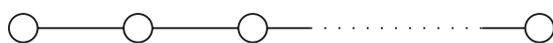
Figure 2. Greechie diagram for classical probability theory over a finite set.

of two different tests are identified. The sets E and S themselves may have any cardinality, but the definition stipulates that the outcomes of any particular test are countable.

For finite E, a test space can be conveniently summarized by a Greechie diagram [46]. Each element of E is represented by a circle, and tests are represented by connecting the corresponding set of circles with a continuous line. Examples of Greechie diagrams are given in figures 1–3.

Definition 6. *A* state *on a test space is a map* $\omega : E \to [0, 1]$ *satisfying* $\sum_{e \in s} \omega(e) = 1 \; \forall s \in S$.

A state defines probabilities for the outcomes of each test such that (i) these probabilities sum to 1 for each test, and (ii) if an outcome appears in more than one test it gets the same probability in each case. Given a test space $\mathcal{A}$, write the set of all possible states $\Omega(\mathcal{A})$. Note that it is easy to construct test spaces for which $\Omega(\mathcal{A})$ is the empty set, or for which a particular outcome has probability 0 in all states, or which may for similar reasons be judged unsatisfactory. Extra assumptions would rule these out, but here there is no need.

The set of possible state spaces of a test space is generic, in the sense that every finite-dimensional compact convex set arises as the set of states for some test space[5]. As noted above, discrete classical probability theory is recovered when there is only one test, i.e. when $\mathcal{A}$ is of the form $(E, \{E\})$ (see figure 2). It is also possible to recover classical probability theory over an arbitrary measurable set [48]. Quantum theory with projective measurements is recovered when $\mathcal{A} = (\mathcal{P}(H), \mathcal{M}(H))$, with $\mathcal{P}(H)$ the set of projection operators on a Hilbert space H and $\mathcal{M}(H)$ the set of projective decompositions of the identity. In this case, Gleason's theorem [49]

---

[5] This statement extends to infinite dimensions. Precisely: every compact convex subset of a locally convex topological vector space is affinely homeomorphic to the set of all states on some test space. In [47], this was proved for the state space of an orthomodular lattice, but the set of all finite ortho-partitions of unity of the lattice is a test space that has the same state space. We thank A Wilce for telling us about this result.
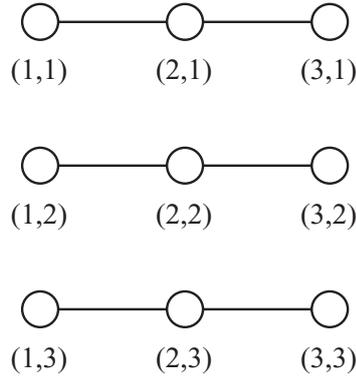
Figure 3. The test space for a process with three inputs and three outputs.

implies that each state corresponds to a density operator ρ, with the probability assigned to projector P given by $\mathrm{Tr}(\rho P)$[6].

It will be useful to define arbitrary linear combinations of states. Given states $\omega_1, \ldots, \omega_k \in \Omega(\mathcal{A})$, the linear combination $v = \sum_i r_i \omega_i$, for real r_i, is defined as a map $E \to \mathfrak{R}$ that satisfies

$$v(e) = \sum_i r_i \omega_i(e) \quad \forall e \in E. \tag{4}$$

The set of all linear combinations of states is a vector space denoted $V(\mathcal{A})$. In the case of quantum theory, for example, $V(\mathcal{A})$ is the real vector space of Hermitian operators on the Hilbert space. Clearly, $\Omega(\mathcal{A})$ is a convex subset of $V(\mathcal{A})$. Since $\Omega(\mathcal{A})$ by definition spans $V(\mathcal{A})$, they have equal dimension. Importantly, from hereon we assume the following.

Assumption 1. $V(\mathcal{A})$ *is finite dimensional.*

Finally, let $V^*(\mathcal{A})$ be the vector space dual to $V(\mathcal{A})$, that is the set of all linear maps $V(\mathcal{A}) \to \mathfrak{R}$. Note that each outcome of a test, that is each $e \in E$, can be uniquely identified with a map $\tilde{e} \in V^*(\mathcal{A})$ such that $\tilde{e}(\omega) = \omega(e) \; \forall \omega \in \Omega(\mathcal{A})$. Under this identification one can write $\omega(e)$ and $e(\omega)$ interchangeably, according to whether states are viewed as assigning probabilities to outcomes of tests, or vice versa. The set E can be viewed as a subset of $V^*(\mathcal{A})$, and it is easy to see that the span of E is equal to $V^*(\mathcal{A})$.

Before moving on to composite systems, we briefly note that there is an alternative approach to operational probabilistic theories that has recently been used to investigate the information theoretic properties of such theories [6, 18, 19]. In this approach one starts with a compact convex set Ω, to be interpreted as a space of states, and defines measurement outcomes

---

[6] This holds provided the dimension of the Hilbert space is $\geqslant 3$. A generalization of Gleason's theorem to encompass positive operator valued (POV) measurements does hold for dimension 2 [50, 51], but the formalism of test spaces is not general enough to encompass these measurements. This is because a POV decomposition of the identity can contain multiple instances of the same term, such as $\{I/2, I/2\}$, and states are constrained to assign the same probability to each instance. In the test space formalism, different outcomes of the same test are always considered distinct. For precisely this reason, generalizations of test spaces known as *effect-test spaces* have been studied that do encompass quantum POV measurements [52, 53]. However, there is no real need to consider them here because we are primarily concerned with properties of states, and the set of possible state spaces of an effect-test space is no more general than that of a test space.

to be the set of affine functionals $f : \Omega \to [0, 1]$. The present work could easily have been formulated in this framework, but it would be odd to do so from a subjective Bayesian point of view. If the states are supposed to represent degrees of belief then it makes sense to start with the objects that they are degrees of belief about, i.e. the tests, rather than the states themselves. There is no loss of generality in working with test spaces, since the result of [47] implies that any compact convex set can arise as the state space of a test space in the finite-dimensional case.

3.2. Multi-partite systems

In order to consider multi-partite systems, one needs to consider the composition of test spaces and the definition of joint states. Suppose that two systems A and B are associated with test spaces $\mathcal{A} = (E, S)$ and $\mathcal{B} = (F, T)$. If the combined system is regarded as a system in and of itself, then it too should be associated with a test space. But how is this constructed, and how is it related to $\mathcal{A}$ and $\mathcal{B}$? Nothing that has been said with respect to single systems implies a unique answer, so further assumptions are needed.

Suppose that given separate systems A and B, it is possible to perform any test s on system A simultaneously with any test t on system B. A joint state assigns probabilities to pairs (e, f) of outcomes. In particular, if $e \in s$ and $e \in s'$, then the probability of obtaining (e, f) does not depend on whether the tests performed are s and t, or s' and t. Similarly if $f \in t$ and $f \in t'$. Suppose further that a specification of the joint probability for all outcome pairs serves to define the joint state uniquely.

This motivates the following.

Definition 7. *Given two test spaces, $\mathcal{A} = (E, S)$ and $\mathcal{B} = (F, T)$, the* Cartesian product*, $\mathcal{A} \times \mathcal{B}$, is a new test space whose set of outcomes is the set theoretic Cartesian product $E \times F$, and whose set of tests is $\{s \times t \,|\, s \in S, t \in T\}$, where $s \times t$ is again the set theoretic Cartesian product.*

From definition 6, a state on $\mathcal{A} \times \mathcal{B}$ is a map $E \times F \to [0, 1]$, with probabilities summing to 1 for each pair of tests (s, t).

Definition 8. *Consider a bipartite system corresponding to a test space $\mathcal{A} \times \mathcal{B}$. A state $\omega \in \Omega(\mathcal{A} \times \mathcal{B})$ is* nonsignalling *if and only if*

$$\sum_{f \in t} \omega(e, f) = \sum_{f \in t'} \omega(e, f) \quad \forall e, t, t', \tag{5}$$

$$\sum_{e \in s} \omega(e, f) = \sum_{e \in s'} \omega(e, f) \quad \forall f, s, s'. \tag{6}$$

If a state is nonsignalling, then the marginal probability of obtaining outcome e for test s does not depend on which $\mathcal{B}$ test is performed. This means that a marginal state $\omega_A \in \Omega(\mathcal{A})$, can be defined such that

$$\omega_A(e) = \sum_{f \in t} \omega(e, f), \tag{7}$$

where the right-hand side does not depend on the choice of t. Similarly, one can define a marginal $\omega_B \in \Omega(\mathcal{B})$.

Definition 9. *Given $\omega_A \in \Omega(\mathcal{A})$ and $\omega_B \in \Omega(\mathcal{B})$, the* direct product, $\omega_A \otimes \omega_B \in \Omega(\mathcal{A} \times \mathcal{B})$, *is defined so that*

$$(\omega_A \otimes \omega_B)(e, f) = \omega_A(e)\omega_B(f) \quad \forall (e, f) \in E \times F. \tag{8}$$

The definition of the Cartesian product of test spaces is valid for any pair of test spaces, including the case in which one of them is itself a Cartesian product. Considering three test spaces, $\mathcal{A} = (E, S)$, $\mathcal{B} = (F, T)$ and $\mathcal{C} = (G, U)$, it is easy to see that $(\mathcal{A} \times \mathcal{B}) \times \mathcal{C}$ is isomorphic to $\mathcal{A} \times (\mathcal{B} \times \mathcal{C})$. Thus one can simply write $\mathcal{A} \times \mathcal{B} \times \mathcal{C}$, with states on $\mathcal{A} \times \mathcal{B} \times \mathcal{C}$ identified with maps $E \times F \times G \to [0, 1]$. The product $\mathcal{A} \times \mathcal{A} \times \cdots \times \mathcal{A}$, where there are n terms in the decomposition, can be written $\mathcal{A}^{\times n}$.

The notion of a nonsignalling state has been defined with respect to bipartite decompositions. It extends readily to the case of an n-fold product.

Definition 10. *Consider a product of test spaces, $\mathcal{A}_1 \times \cdots \times \mathcal{A}_n$. A state $\omega \in \Omega(\mathcal{A}_1 \times \cdots \times \mathcal{A}_n)$ is n-fold nonsignalling if and only if ω is nonsignalling with respect to every bipartite split. More formally, let α be a subset of $\{1, \ldots, n\}$, let $|\alpha| = k$, and write $\alpha = \{i_1, \ldots, i_k\}$. Then ω is n-fold nonsignalling iff*

$$\sum_{e_{i_1} \in t_{i_1}} \cdots \sum_{e_{i_k} \in t_{i_k}} \omega(e_1, \ldots, e_n) = \sum_{e_{i_1} \in t'_{i_1}} \cdots \sum_{e_{i_k} \in t'_{i_k}} \omega(e_1, \ldots, e_n),$$

for all α, for all e_j with $j \notin \alpha$, and for all tests $t_{i_1}, \ldots, t_{i_k}$ and $t'_{i_1}, \ldots, t'_{i_k}$.

Finally, this will be useful.

Lemma 1. *Consider the test space $\mathcal{A}_1 \times \cdots \times \mathcal{A}_n$. The direct product states, of the form $\omega_1 \otimes \cdots \otimes \omega_n$, span a subspace of $V(\mathcal{A}_1 \times \cdots \times \mathcal{A}_n)$, and the subspace can be identified with $V(\mathcal{A}_1) \otimes \cdots \otimes V(\mathcal{A}_n)$. If a joint state ω is n-fold nonsignalling, then $\omega \in V(\mathcal{A}_1) \otimes \cdots \otimes V(\mathcal{A}_n)$, i.e. ω can be written as a linear combination of direct products.*

Proof. Begin with the $n = 2$ case. The tensor product $V(\mathcal{A}_1) \otimes V(\mathcal{A}_2)$ can be defined as the set of bilinear maps $V^*(\mathcal{A}_1) \times V^*(\mathcal{A}_2) \to \Re$. Any direct product $\omega_1 \otimes \omega_2$ defines such a map via $(a, b) \to a(\omega_1)b(\omega_2)$, and it is straightforward that these span $V(\mathcal{A}_1) \otimes V(\mathcal{A}_2)$. Now consider a nonsignalling joint state ω. The fact that ω is nonsignalling permits the definition of the marginal state ω_1. Define the conditional state $\omega_{2|e}$ such that $\omega_{2|e}(f)$ is the probability of outcome f on system 2, given that outcome e was obtained on system 1. Thus $\omega(e, f) = \omega_1(e)\omega_{2|e}(f)$, for all $(e, f) \in E \times F$. Note that $\omega_{2|e} \in \Omega(\mathcal{A}_2)$. Now suppose that f and $\{g_i\}$ are elements of F such that, considered as elements of $V^*(\mathcal{A}_2)$, $f = \sum_i r_i g_i$. Thus $\omega_2(f) = \sum_i r_i \omega_2(g_i)$ for all $\omega_2 \in \Omega(\mathcal{A}_2)$. Then $\omega(e, f) = \omega_1(e)\omega_{2|e}(f) = \sum_i r_i \omega_1(e)\omega_{2|e}(g_i) = \sum_i r_i \omega(e, g_i)$. A value $\omega(e, b)$ can now be defined for arbitrary $b \in V^*(\mathcal{A}_2)$ by linear extension. Similar reasoning concludes that $\omega(e, b)$ is linear in the first argument, hence can be extended to $\omega(a, b)$ for $a \in V^*(\mathcal{A}_1)$. So ω defines a bilinear map $V^*(\mathcal{A}_1) \times V^*(\mathcal{A}_2) \to \Re$ as required. The extension to general n is straightforward. $\qquad\square$

Note that if we have two quantum test spaces $\mathcal{A} = (\mathcal{P}(H_A), \mathcal{M}(H_B))$ and $\mathcal{B} = (\mathcal{P}(H_B), \mathcal{M}(H_B))$ then the space of nonsignalling states on $\mathcal{A} \times \mathcal{B}$ is larger than the state space of $\mathcal{AB} = (\mathcal{P}(H_A \otimes H_B), \mathcal{M}(H_A \otimes H_B))$. To see this, recall that the nonsignalling state space

of $\mathcal{A} \times \mathcal{B}$ only has to be positive for all possible choices of local measurements on $\mathcal{A}$ and $\mathcal{B}$, whereas the test space $\mathcal{AB}$ includes joint measurements, such as the Bell measurement for example. Thus, the criteria to be a state on $\mathcal{AB}$ are more restrictive than those for nonsignalling states on $\mathcal{A} \times \mathcal{B}$. Indeed, if we take a state on $\mathcal{AB}$ and perform a positive, but not completely positive, map on system $\mathcal{A}$ then the result is still a valid state on $\mathcal{A} \times \mathcal{B}$, but not on $\mathcal{AB}$ in general, e.g. consider performing a partial transpose on a Bell state. Nevertheless, the state space of $\mathcal{AB}$ is still a convex subset of the nonsignalling states on $\mathcal{A} \times \mathcal{B}$, which is enough to apply our theorem. More generally, one might want to consider rules for composing subsystems that yield a convex subset of the nonsignalling states on the Cartesian product for arbitrary test spaces.

4. A de Finetti theorem for test spaces

Given a system associated with a test space $\mathcal{A} = (E, S)$, suppose that n copies are associated with the product $\mathcal{A}^{\times n}$. Given an infinite sequence of states $\omega^1, \omega^2, \ldots$ where $\omega^n \in \Omega(\mathcal{A}^{\times n})$, it is possible to define symmetry and exchangeability in a manner similar to the classical and quantum cases. The main difference is that here, the definition of exchangeability involves the extra condition that the states are nonsignalling.

Definition 11. *A state $\omega^n \in \Omega(\mathcal{A}^{\times n})$ is* symmetric *if and only if it is invariant under permutations of the n systems. That is,*

$$\omega^n(e_1, \ldots, e_n) = \omega^n(e_{\pi(1)}, \ldots, e_{\pi(n)}),$$

for all permutations $\pi : (1, \ldots, n) \to (1, \ldots, n)$.

Definition 12. *A sequence of states $\omega^1, \omega^2, \ldots$ where $\omega^n \in \Omega(\mathcal{A}^{\times n})$ is* exchangeable *if and only if*

1. *$\forall n \; \omega^n$ is symmetric,*
2. *$\forall n \; \omega^n$ is n-fold nonsignalling,*
3. *$\omega^n(e_1, \ldots, e_n) = \sum_{e_{n+1} \in s} \omega^{n+1}(e_1, \ldots, e_n, e_{n+1})$.*

Theorem 3 (the de Finetti theorem for test spaces). *Suppose that the sequence $\omega^1, \omega^2, \ldots$ where $\omega^n \in \Omega(\mathcal{A}^{\times n})$, is exchangeable. Then ω^n can be written in the form*

$$\omega^n = \int_{\Omega(\mathcal{A})} d\mu(\omega) \; \omega \otimes \cdots \otimes \omega, \tag{9}$$

where μ is a probability measure on $\Omega(\mathcal{A})$, μ is independent of n, and μ is unique.

Proof. The proof is an adaptation of the proof of the quantum de Finetti theorem due to Caves *et al* [27]. Recall that in quantum theory, an informationally complete measurement is a POV measurement such that if the outcome probabilities are all known, then the state is determined uniquely. The strategy of Caves *et al* is to generate a classical distribution by considering an informationally complete POV measurement performed separately on each quantum system. Applying the classical de Finetti theorem to the distribution of outcome sequences allows the form of the quantum state to be inferred.

In our context, the test space $\mathcal{A}$ need not include a test that is informationally complete for the state space Ω. But for the purposes of proof, this does not matter. All that is needed is a corresponding mathematical construction.

Lemma 2. *There exists a set $M = \{a_1, \ldots, a_d\}$, where $a_i \in V^*(\mathcal{A})$ and d is the dimension of $V^*(\mathcal{A})$, such that*

1. *M is a basis for $V^*(\mathcal{A})$, i.e., the a_i are linearly independent,*
2. *$0 \leqslant a_i(\omega) \leqslant 1 \; \forall \omega \in \Omega(\mathcal{A})$,*
3. *$\sum_i a_i(\omega) = 1 \; \forall \omega \in \Omega(\mathcal{A})$.*

Proof. This result is not new. It is also used in [18], and we give the same proof. A more general version is proven in [54]. Let u be the unique vector in $V^*(\mathcal{A})$ such that $u(\omega) = 1$ for all $\omega \in \Omega(\mathcal{A})$. Consider an arbitrary basis $\{b_1, \ldots, b_d\}$ of $V^*(\mathcal{A})$. Apply an invertible linear transformation to obtain $\{\tilde{b}_1, \ldots, \tilde{b}_d\}$, where $\sum_i \tilde{b}_i = u$. Some of the $\tilde{b}_i$ may be negative on some states ω. Define a constant c as the minimum value of $\tilde{b}_i(\omega)$, where the minimum is taken over all i and all $\omega \in \Omega(\mathcal{A})$. Then $\tilde{b}_i - cu$ is positive and nonzero. It follows that $\sum_i (\tilde{b}_i - cu) = (1 - dc)u$, with $1 - dc > 0$. Define $a_i = (\tilde{b}_i - cu)/(1 - dc)$. $\qquad \square$

The set M will play the role of an informationally complete measurement. The linear independence of the a_i means that a state ω is determined uniquely by the values $a_i(\omega)$. The idea now is that given a nonsignalling $\omega^n \in \Omega(\mathcal{A}^{\times n})$, one can at least imagine a measurement corresponding to M performed separately on each system. The probability of obtaining an outcome sequence $(a_{i_1}, \ldots, a_{i_n})$ for a direct product state $\omega_1 \otimes \cdots \otimes \omega_n$ is defined as $a_{i_1}(\omega_1) \cdots a_{i_n}(\omega_n)$. Recalling lemma 1, according to which a nonsignalling state can be written as a linear combination of direct product states, the probability of the sequence $(a_{i_1}, \ldots, a_{i_n})$ for an arbitrary nonsignalling state can be defined by linear extension. Note that this way, the sequence $(a_{i_1}, \ldots, a_{i_n})$ corresponds to a vector $a_{i_1} \otimes \cdots \otimes a_{i_n} \in (V^*(\mathcal{A}))^{\otimes n}$, such that the probability is given by $(a_{i_1} \otimes \cdots \otimes a_{i_n})(\omega^n)$. The vectors $a_{i_1} \otimes \cdots \otimes a_{i_n}$ are linearly independent and span the tensor product space $(V^*(\mathcal{A}))^{\otimes n}$. This means that the joint measurement $M^{\times n}$ is informationally complete for the nonsignalling n-partite system.

For simplicity, write the probability of the outcome sequence $(a_{i_1}, \ldots, a_{i_n})$ as $P^n(i_1, \ldots, i_n)$ so that

$$P^n(i_1, \ldots, i_n) = (a_{i_1} \otimes \cdots \otimes a_{i_n})(\omega^n). \tag{10}$$

If ω^n is symmetric, then so is P^n, since

$$\begin{aligned}
P^n(i_{\pi(1)}, \ldots, i_{\pi(n)}) &= (a_{i_{\pi(1)}} \otimes \cdots \otimes a_{i_{\pi(n)}})(\omega^n) \\
&= (a_{i_1} \otimes \cdots \otimes a_{i_n})(\Pi(\omega^n)) \\
&= (a_{i_1} \otimes \cdots \otimes a_{i_n})(\omega^n) \\
&= P^n(i_1, \ldots, i_n),
\end{aligned} \tag{11}$$

for any permutation π.

Further, if the sequence $\omega^1, \omega^2, \ldots$ is exchangeable, then so is the sequence $P^1, P^2, \ldots$, since it is symmetric and

$$\begin{aligned}
\sum_{i_{n+1}} P^{n+1}(i_1, \ldots, i_n, i_{n+1}) &= \sum_{i_{n+1}} (a_{i_1} \otimes \cdots \otimes a_{i_n} \otimes a_{i_{n+1}})(\omega^{n+1}) \\
&= (a_{i_1} \otimes \cdots \otimes a_{i_n})(\omega^n) \\
&= P^n(i_1, \ldots, i_n).
\end{aligned} \tag{12}$$

IOP Institute of Physics Φ DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

Now apply the classical de Finetti theorem (theorem 1) to the sequence $P^1, P^2, \ldots$ to obtain

$$P^n(i_1, \ldots, i_n) = \int_{\Delta_d} \mathrm{d}\mu(p) \, p(a_{i_1}) \cdots p(a_{i_n}), \tag{13}$$

where Δ_d is the set of probability distributions over $\{a_1, \ldots, a_d\}$.

For each such distribution p, there is a unique $\omega_p \in V(\mathcal{A})$ such that $a_i(\omega_p) = p(a_i)$ (where uniqueness follows from the linear independence of the a_i). Thus equation (13) can be rewritten as

$$(a_{i_1} \otimes \cdots \otimes a_{i_n})(\omega^n) = \int_{\Delta_d} \mathrm{d}\mu(p) \, a_{i_1}(\omega_p) \cdots a_{i_n}(\omega_p). \tag{14}$$

Since the joint measurement is informationally complete for the nonsignalling n-partite states, this implies

$$\omega^n = \int_{\Delta_d} \mathrm{d}\mu(p) \, \omega_p \otimes \cdots \otimes \omega_p. \tag{15}$$

This is not quite sufficient to establish theorem 3. The right-hand side of equation (15) is an integral over all $\omega_p \in V(\mathcal{A})$ satisfying $0 \leqslant a_i(\omega_p) \leqslant 1$ and $\sum_i a_i(\omega_p) = 1$. That is, it is an integral over all ω_p that return sensible probabilities for the imaginary informationally complete measurement. But in general, there are ω_p satisfying these conditions that are not valid states because they return a value < 0 for some element of the test space. It remains to show that the integral can be restricted to those $\omega_p \in \Omega(\mathcal{A})$.

To this end, consider an $\omega_p \in V(\mathcal{A})$ such that $0 \leqslant a_i(\omega_p) \leqslant 1$ and $\sum_i a_i(\omega_p) = 1$, but $\omega_p(e) < 0$ for some $e \in E$. It must be the case that $e \in s$ for some test s. Let the other elements of s be $\{f_1, \ldots, f_k\}$ and note that $\sum_i \omega_p(f_i) > 1$. There must exist an $\epsilon > 0$ and a neighborhood N of ω_p in $V(\mathcal{A})$ such that $\sum_i \omega(f_i) > 1 + \epsilon$ for all $\omega \in N$. Let $\tilde{N}$ be the subset of Δ_d such that $p(a_i) = a_i(\omega)$ for some $\omega \in N$. Expression (15) holds for any n, so suppose that n is even and that the test s is performed on each of the n systems. The probability that outcome e is never obtained is defined by ω^n and is given by

$$\int_{\Delta_d} \mathrm{d}\mu(p) \left(\sum_i \omega_p(f_i) \right)^n = \int_{\Delta_d \setminus \tilde{N}} \mathrm{d}\mu(p) \left(\sum_i \omega_p(f_i) \right)^n + \int_{\tilde{N}} \mathrm{d}\mu(p) \left(\sum_i \omega_p(f_i) \right)^n$$

$$> \int_{\tilde{N}} \mathrm{d}\mu(p) \left(\sum_i \omega_p(f_i) \right)^n$$

$$> (1 + \epsilon)^n \int_{\tilde{N}} \mathrm{d}\mu(p), \tag{16}$$

where we have used the fact that the first term after the equality is > 0 if n is even. For large enough n this expression is > 1 unless $\mu(\tilde{N}) = 0$. This holds for any $\omega_p \notin \Omega(\mathcal{A})$. It follows that, with a suitable redefinition of μ,

$$\omega^n = \int_{\Omega(\mathcal{A})} \mathrm{d}\mu(\omega) \, \omega \otimes \cdots \otimes \omega. \tag{17}$$

$\square$

5. Consequences of theorem 3

The de Finetti theorem for test spaces is rather general. Very little goes into the definition of a test space itself. In fact, in the finite-dimensional case we are considering, the state space of an individual system may be an arbitrary compact convex set. Thus, the most substantive assumptions that go into the theorem are those that concern how test spaces combine when joint systems are considered.

In the classical case, the test space $(E, \{E\})$ contains a single test and the state space for an individual system is Δ_d, where d is the number of elements of E. In this case, the nonsignalling condition is redundant because there is just a single test on each system and hence no freedom to choose alternative measurements. The Cartesian product of classical test spaces corresponds to the usual Cartesian product of sample spaces and so theorem 3 reduces to theorem 1 straightforwardly. The quantum case is a little more subtle because the state of n quantum systems belongs to the state space of $(\mathcal{P}(H^{\otimes n}), \mathcal{M}(H^{\otimes n}))$, rather than the state space of the Cartesian product of n test spaces of the form $(\mathcal{P}(H), \mathcal{M}(H))$. Nevertheless, a quantum state defined by a density operator on $H^{\otimes n}$ is uniquely specified by the probabilities for measurement outcomes of the form $Q_1 \otimes Q_2 \otimes \cdots \otimes Q_n$, where the Q_i are projection operators. It follows that the states on $(\mathcal{P}(H^{\otimes n}), \mathcal{M}(H^{\otimes n}))$ can be identified with a subset of the nonsignalling states on $(\mathcal{P}(H), \mathcal{M}(H))^{\times n}$. Hence theorem 3 implies the quantum de Finetti theorem.

As a further illustration of the generality of the result, we note that a theorem for classical processes, or conditional probabilities, can also be viewed as a special case of theorem 3. A process C can be thought of as taking an input Y into an output X, where Y takes values in $\{1, \ldots, k\}$ and X takes values in $\{1, \ldots, d\}$. The process can be defined as the set of conditional probabilities of the form $P(X = x | Y = y)$ (abbreviated $P(x|y)$). Clearly, the set of all such C can be regarded as the set of states on a test space (E, S), where E consists of ordered pairs with $E = \{(x, y)\}_{x=1,\ldots,d, y=1,\ldots,k}$ and $S = \{\{(x, 1)\}_{x=1}^d, \ldots, \{(x, k)\}_{x=1}^d\}$, as illustrated for the $d = k = 3$ case in figure 3. The state space of this test space is isomorphic to the set of conditional probability distributions via the identification $\omega((x, y)) = P(x|y)$.

More generally, a process C^n takes inputs $Y_1, \ldots, Y_n$ into outputs $X_1, \ldots, X_n$, where each Y_i takes values in $\{1, \ldots, k\}$, and each X_i takes values in $\{1, \ldots, d\}$. Such a C^n can be defined as the set of conditional probabilities of the form $P^n(x_1, \ldots, x_n | y_1, \ldots, y_n)$, and can also be identified with a state on the Cartesian product $(E, S)^{\times n}$. The n-fold nonsignalling and symmetry conditions can be defined for C^n exactly as they are in definitions 10 and 11. Exchangeability for a sequence $C^1, C^2, \ldots$ can be defined exactly as in definition 12. Theorem 3 becomes the following theorem for classical processes.

Theorem 4. *If the sequence C^1, C^2, $\ldots$ is exchangeable, then the conditional probabilities defining C^n can be written in the form*

$$P(x_1, \ldots, x_n | y_1, \ldots, y_n) = \int_\Omega \mathrm{d}\mu(p) \, p(x_1|y_1) \cdots p(x_n|y_n), \tag{18}$$

where Ω is the set of processes that take a single input X into a single output Y, μ is a probability measure on Ω, μ is independent of n, and μ is unique.

A couple of remarks regarding theorem 4 might be helpful.

Firstly, we described theorem 4 as a special case of theorem 3. It is worth noting the sense in which it is strictly less general. After all, any state on any test space could be thought of as

a classical process, which takes an input (choice of test) into an output (outcome of test). The point is that, with this identification applied to a generic test space $\mathcal{A}$, not all sets of conditional probabilities $p(x|y)$ will correspond to valid states on $\mathcal{A}$. It is only if $\mathcal{A}$ has the special feature that the tests are nonoverlapping that this will be the case. The integral in equation (18) ranges over all processes of the form $p(x|y)$, whereas in equation (9), it is important that the integral ranges only over $\Omega(\mathcal{A})$.

Secondly, the de Finetti theorem for classical processes can be viewed as a de Finetti theorem for *states* in a theory that is nonclassical and nonquantum, but admits the most general type of correlations compatible with the nonsignalling requirement. This theory is discussed in [6], where it is called generalized nonsignalling theory. It admits superquantum correlations, which have been discussed in the quantum information literature under the name Popescu–Rohrlich, or nonlocal, boxes. Of course, as in the quantum case, exchangeable states do not actually exhibit such correlations, since the de Finetti theorem shows that they are separable.

6. When does a de Finetti-type theorem not hold?

The de Finetti theorem for test spaces holds thanks to a number of assumptions concerning how systems combine to make joint systems. One of these is the nonsignalling condition for joint states. Others are encoded in the formal definition of a Cartesian product of test spaces. It is interesting to see what happens when these assumptions are relaxed, so in this section we present a number of cases where the theorem fails.

6.1. The nonsignalling condition

First, the assumption that the joint states are nonsignalling is crucial not only in the proof, but also in the very definition of exchangeability. In general, if a state $\omega \in \Omega(\mathcal{A}_1 \times \mathcal{A}_2)$ is signalling, then it is not possible to define marginal states ω_1 and ω_2. But exchangeability requires that, given the $n+1$st state in the sequence, the marginal state of the first n systems should be defined and should equal the nth state. In fact, if a state $\omega \in \Omega(\mathcal{A}_1 \times \mathcal{A}_2)$ is only signalling in one direction then one of the marginals can be defined. For example, if probabilities of outcomes for system 2 depend on which test was performed on system 1, but not vice versa, then ω_1 is well defined. But such a state is not symmetric, thus could not form part of an exchangeable sequence. Arguably, the possibility of performing tests on one system that do not affect the other is part and parcel of what we mean when we speak of separate systems (or separate trials).

6.2. Simultaneous measurements

Implicit in the definition of the Cartesian product of test spaces is the idea that a test on one system can be regarded as simultaneous with a test on the other system. One can certainly imagine rules for combining systems where this is not the case. As a very simple example, consider two classical bits which have combined in the following strange manner. If bit 1 is measured before bit 2, then the bits are found to be 00 or 11 with equal probability. On the other hand, if bit 2 is measured before bit 1, then the outcomes are 01 and 10 with equal probability. Note that a suitable nonsignalling condition is satisfied and it is possible to define marginal

states for these bits. With more complicated test spaces one can construct examples like this which are also symmetric[7]. We leave open the status of the de Finetti theorem in such cases.

6.3. Extra degrees of freedom

Another assumption that is implicit in the Cartesian product of test spaces is that the joint state of two systems is completely specified by the probabilities for the joint outcomes (e, f) of each pair of local tests (s, t).

One can construct theories in which a joint state does indeed determine such probabilities, but is not completely specified by them. There are extra degrees of freedom, bound up in the two systems, which are inaccessible unless some kind of joint operation involving both systems at once is performed.

As discussed in [27], a clear example of this is provided by a modification of quantum theory in which real Hilbert spaces are used rather than complex Hilbert spaces. States and observables correspond to real symmetric, rather than complex Hermitian, operators. For a two-dimensional system (a *rebit*), there are measurements corresponding to x- and z-spin but not y-spin. In this case, if σ_y is the usual Pauli matrix, then $1/4(I \otimes I + \sigma_y \otimes \sigma_y)$ is an allowed state of two rebits. But the only way it can be distinguished from $1/4(I \otimes I)$ is if a joint observable such as $\sigma_y \otimes \sigma_y$ is measured. Note that in real quantum theory, this really is a joint observable: it cannot be measured via a separate σ_y measurement on each system.

In [27] it is shown explicitly that the de Finetti theorem fails in real quantum theory. If ω^n is defined by

$$\omega^n = \frac{1}{2}\left(\frac{I + \sigma_y}{2}\right)^{\otimes n} + \frac{1}{2}\left(\frac{I - \sigma_y}{2}\right)^{\otimes n}, \tag{19}$$

then it is real and symmetric, and the sequence $\omega^1, \omega^2, \ldots$ is exchangeable. But by the de Finetti theorem for complex quantum theory, the right-hand side of equation (19) is the unique de Finetti representation for this sequence.

7. Conclusion

In this paper, an infinite de Finetti theorem for test spaces has been presented, which generalizes both the classical and quantum de Finetti theorems. To illustrate the generality of the result, we have shown that a de Finetti theorem for classical processes, which may also be interpreted as a de Finetti theorem for nonlocal boxes, follows as a special case.

From a practical point of view, proving theorems for test spaces, rather than just for quantum theory, confers significant advantages. Not only do we achieve a unification of the classical and quantum results, but we also obtain results that apply to essentially arbitrary convex sets. This is potentially relevant when technological limitations prevent the preparation

---

[7] Consider the test space $\mathcal{A}$, with outcomes $E = \{a, b, c, d\}$, and two tests corresponding to $s1 = \{a, b\}$, and $s2 = \{c, d\}$. Suppose that there are two systems, A and B, each described by the test space $\mathcal{A}$, and that they have combined as follows. If the test $s1$ is performed on both systems, the outcomes are completely random and uncorrelated. If the test $s2$ is performed on both systems, then the outcomes are completely random and uncorrelated. On the other hand, if the test $s1$ is performed on either one of the systems, followed by $s2$ on the other, then the joint outcomes are ac or bd with equal probability. If the test $s2$ is performed on either one of the systems, followed by $s1$ on the other, then the joint outcomes are ad or bc with equal probability. It is clear that this peculiar bipartite system does not allow signalling and is also invariant under a permutation of the two systems.

of arbitrary quantum states of certain systems, so that there is an effective restriction to a convex subset.

From a foundational point of view, this work can be seen as part of a project of understanding what is responsible for the enhanced information processing power of quantum theory, and for the project of deriving quantum theory from information theoretic axioms. In particular, if one adopts a subjective Bayesian approach to probability, it might be desirable to impose the requirement that, in any reasonable theory, one should be able to make sense of the idea of reconstructing an unknown state of a system by making repeated measurements on identically prepared copies. Having a de Finetti theorem for test spaces means that this does indeed make sense for theories in this framework, and that the existing approaches to Bayesian state tomography in quantum theory would generalize straightforwardly.

There are various directions for future work. It would be useful to produce a finite de Finetti theorem for test spaces. It would also be useful to establish whether a (finite or infinite) de Finetti theorem holds without the assumption made here of finite dimensionality of state spaces. Finally, as discussed in section 6.2, it might be interesting to explore the status of de Finetti-type theorems in cases where systems combine in nonstandard ways.

Note added. Related results have recently been obtained by Christandl and Toner [55], who derive a de Finetti theorem for classical processes, analogous to theorem 4 of the present work, but extended to the finite case.

Acknowledgments

References

[1] Renner R 2005 *PhD Thesis* Swiss Federal Institute of Technology (arXiv:quant-ph/0512258)
[2] Doherty A C, Parrilo P A and Spedalieri F M 2004 *Phys. Rev.* A **69** 022308 (arXiv:quant-ph/0308032)
[3] Ioannou L M 2007 *Quantum Inf. Comput.* **7** 335 (arXiv:quant-ph/0603199)
[4] Audenaert K M R 2004 *Proc. MTNS2004* ed B De Moor, B Motmans, J Wilems, P Van Dooren and V Blondel (arXiv:quant-ph/0402076)
[5] Barrett J, Linden N, Massar S, Pironio S, Popescu S and Roberts D 2005 *Phys. Rev.* A **71** 022101
[6] Barrett J 2007 *Phys. Rev.* A **75** 032304 (arXiv:quant-ph/0508211)
[7] Barrett J and Pironio S 2005 *Phys. Rev. Lett.* **95** 140401 (arXiv:quant-ph/0506180)
[8] Brassard G, Buhrman H, Linden N, Méthot A A, Tapp A and Unger F 2006 *Phys. Rev. Lett.* **96** 250401 (arXiv:quant-ph/0508042)
[9] Broadbent A and Méthot A A 2006 *Theor. Comput. Sci.* **358** 3 (arXiv:quant-ph/0504136)
[10] Buhrman H, Christandl M, Unger F, Wehner S and Winter A 2006 *Proc. R. Soc.* A **462** 1919 (arXiv:quant-ph/0504133)
[11] van Dam W 2000 *PhD Thesis* University of Oxford
[12] van Dam W 2005 arXiv:quant-ph/0501159

[13] Jones N S and Masanes Ll 2005 *Phys. Rev.* A **72** 052312 (arXiv:quant-ph/0506182)

[14] Khalfi L A and Tsirelson B S 1985 *Symp. on the Foundations of Modern Physics* ed P Lahti and P Mittelstaedt (Singapore: World Scientific) pp 441–60

[15] Popescu S and Rohrlich D 1994 *Found. Phys.* **24** 379

[16] Short A J, Gisin N and Popescu S 2006 *Quantum Inf. Process.* **5** 131 (arXiv:quant-ph/0504134)

[17] Short A J, Popescu S and Gisin N 2006 *Phys. Rev.* A **73** 012101 (arXiv:quant-ph/0508120)

[18] Barnum H, Barrett J, Leifer M and Wilce A 2006 arXiv:quant-ph/0611295

[19] Barnum H, Barrett J, Leifer M and Wilce A 2007 *Phys. Rev. Lett.* **99** 240501 (arXiv:0707.0620v1)

[20] de Finetti B 1993 *Probabilità e Induzione—Induction and Probability* Biblioteca di STATISTICA (Bologna: CLUEB), a collection of de Finetti's original papers and English translations

[21] de Finetti B 1990 *Theory of Probability: A Critical Introductory Treatment* vols 1 and 2 (New York: Wiley) [Engl. Transl. by A Smith]

[22] Savage L J 1972 *The Foundation of Statistics* (New York: Dover)

[23] Bernardo J M and Smith A F M 2000 *Bayesian Theory* (*Wiley Series in Probability and Statistics*) (New York: Wiley)

[24] Hewitt E and Savage L J 1955 *Trans. Am. Math. Soc.* **80** 470

[25] Hudson R L and Moody G R 1976 *Z. Wahrscheinlichkeitstheor. Verwandte Geb.* **33** 343

[26] Hudson R L 1981 *Found. Phys.* **11** 805

[27] Caves C M, Fuchs C A and Schack R 2002 *J. Math. Phys.* **43** 4537 (arXiv:quant-ph/0104088)

[28] Fuchs C A and Schack R 2004 *Quantum State Estimation (Lecture Notes in Physics* vol 649) ed M G A Paris and J Řeháček (Berlin: Springer) chapter 5, pp 147–82 (arXiv:quant-ph/0404156)

[29] Caves C M, Fuchs C A and Schack R 2002 *Phys. Rev.* A **65** 022305 (arXiv:quant-ph/0106133)

[30] Fuchs C A 2002 arXiv:quant-ph/0205039

[31] Fuchs C A 2003 *J. Mod. Opt.* **50** 987

[32] Pitowsky I 2003 *Stud. Hist. Phil. Mod. Phys.* **34** 395 (arXiv:quant-ph/0208121)

[33] Caves C M, Fuchs C A and Schack R 2006 arXiv:quant-ph/0608190

[34] Diaconis P 1977 *Synthese* **36** 271

[35] Diaconis P and Freedman D 1980 *Ann. Probab.* **8** 745

[36] Kendall D G 1967 *Stud. Sci. Math. Hung.* **2** 319

[37] Koenig R and Renner R 2005 *J. Math. Phys.* **46** 122108 (arXiv:quant-ph/0410229)

[38] D'Cruz C, Osborne T J and Schack R 2007 *Phys. Rev. Lett.* **98** 160406 (arXiv:quant-ph/0606139)

[39] Renner R 2007 arXiv:quant-ph/0703069

[40] Koenig R and Mitchison G 2009 *J. Math. Phys.* **50** 012105 (arXiv:quant-ph/0703210)

[41] Fuchs C A, Schack R and Scudo P F 2004 *Phys. Rev.* A **69** 062305 (arXiv:quant-ph/0307198)

[42] Christandl M, Koenig R, Mitchison G and Renner R 2007 *Commun. Math. Phys.* **273** 473–98 (arXiv:quant-ph/0602130)

[43] Mitchison G 2007 arXiv:quant-ph/0701064

[44] Foulis D and Randall C H 1972 *J. Math. Phys.* **13** 1667

[45] Randall C H and Foulis D 1973 *J. Math. Phys.* **14** 1472

[46] Greechie R J 1969 *Carribean J. Sci. Math.* **1** 15–26

[47] Shultz F W 1974 *J. Comb. Theory* A **17** 317

[48] Wilce A 2000 *Current Research in Operational Quantum Logic* ed B Coecke, D Moore and A Wilce (Dordrecht: Kluwer) pp 81–114

[49] Gleason A M 1957 *J. Math. Mech.* **6** 885

[50] Busch P 2003 *Phys. Rev. Lett.* **91** 120403 (arXiv:quant-ph/9909073)

[51] Caves C M, Fuchs C A, Manne K and Renes J M 2004 *Found. Phys.* **34** 193 (arXiv:quant-ph/0306179)

[52] Pulmannová S and Wilce A 1995 *Int. J. Theor. Phys.* **34** 1689

[53] Gudder S 1997 *Found. Phys.* **27** 287

[54] Singer M and Stulpe W 1992 *J. Math. Phys.* **33** 131

[55] Christandl M and Toner B 2007 arXiv:0712.0916