Quantum Bayes rule

Rüdiger Schack,[1] Todd A. Brun,[2] and Carlton M. Caves[3]

[1]*Department of Mathematics, Royal Holloway, University of London, Egham, Surrey TW20 0EX, United Kingdom*
[2]*Physics Department, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213*
[3]*Center for Advanced Studies, Department of Physics and Astronomy, University of New Mexico, Albuquerque, New Mexico 87131*

We state a quantum version of the Bayes rule for statistical inference and give a simple general derivation within the framework of generalized measurements. The rule can be applied to measurements on N copies of a system if the initial state of the N copies is *exchangeable*. As an illustration, we apply the rule to N qubits. Finally, we show that quantum state estimates derived via the principle of maximum entropy are fundamentally different from those obtained via the quantum Bayes rule.

During the last decade, interest in Bayesian methods of statistical inference has increased considerably [1,2]. At the heart of the Bayesian approach is Bayes's rule, which indicates how to update a state of knowledge in the light of new data. The simplest form of the rule is

$$p(H|D) = \frac{p(D|H)p(H)}{p(D)}, \quad (1)$$

where $p(D|H)$ is the probability for the data D given a hypothesis H, $p(H)$ is the *prior* probability that the hypothesis is true, $p(H|D)$ is the *posterior* probability that the hypothesis is true given the data, and $p(D) = \Sigma_H p(D|H)p(H)$ is the probability for the data averaged over all hypotheses. The conceptual simplicity of the Bayes rule is a major strength of the Bayesian approach.

The Bayes rule relies, however, on the classical assumption that acquiring data D does not alter the probability distribution in any way other than eliminating contrary hypotheses. In quantum mechanics this need not be true, as Heisenberg's famous uncertainty principle illustrates.

Problems of statistical inference and state estimation are of central importance in quantum information theory. After the early pioneering work on quantum inference [3–6] and quantum state tomography [7–9], a large amount of work has been done on the subject (see, e.g., Refs. [10–19]). In many of the cited papers, a quantum version of the Bayes rule is used either implicitly or explicitly, yet these papers also illustrate that there remains considerable confusion about the exact circumstances in which the Bayes rule applies to quantum systems. Jones [6] has derived a quantum Bayes rule for pure states only. In this paper, we derive a general rule, valid both for pure and mixed states, and give a precise condition for its validity.

We consider the following general inference problem. Let $\mathcal{H}$ be the Hilbert space of a quantum system. The Hilbert space of N copies of the system is given by the N-fold tensor product $\mathcal{H}^{\otimes N}$. Suppose one is given a (prior) state $\hat{\rho}^{(M+N)}$ on $\mathcal{H}^{\otimes(M+N)}$ and the results of measurements on M subsystems. The task is to find the (posterior) state of the remaining N subsystems conditioned on the measurement results. The problem is, in principle, completely solved by the theory of generalized measurements [20], which prescribes the state of

the total system after the measurement. There is no room in quantum theory for an additional independent inference principle; any inference rule must be derivable from the basic theory.

The rule for assigning a post-measurement state is provided by the formalism of generalized measurements [20]. An arbitrary measurement on M subsystems is described by a set of completely positive, trace-decreasing operations, $\{\mathcal{F}_k\}$, which act on the selected M subsystems. The action of such a set of operations on a system in state $\hat{\rho} = \hat{\rho}^{(M+N)}$ can be written as

$$\mathcal{F}_k(\hat{\rho}) = \sum_l \hat{A}_{kl} \hat{\rho} \hat{A}_{kl}^\dagger, \quad (2)$$

where the operators $\hat{A}_{kl}$ obey

$$\sum_{kl} \hat{A}_{kl}^\dagger \hat{A}_{kl} = \hat{I}. \quad (3)$$

Defining $\hat{E}_k = \Sigma_l \hat{A}_{kl}^\dagger \hat{A}_{kl}$, condition (3) takes the form $\Sigma_k \hat{E}_k = \hat{I}$, which implies that the set $\{\hat{E}_k\}$ forms a positive operator valued measure (POVM).

If the measurement result is k, the (normalized) state of all $M+N$ systems after the measurement is

$$\hat{\rho}_k^{(M+N)} = \frac{1}{p_k} \mathcal{F}_k(\hat{\rho}^{(M+N)}), \quad (4)$$

where $\mathcal{F}_k(\hat{\rho}^{(M+N)})$, given in Eq. (2), is the unnormalized state conditioned on measurement outcome k, and the normalization is provided by the probability for outcome k,

$$p_k = \text{Tr}[\mathcal{F}_k(\hat{\rho}^{(M+N)})] = \text{Tr}(\hat{E}_k \hat{\rho}^{(M+N)}). \quad (5)$$

Since we are assuming that only the first M systems are measured, the operators in Eq. (2) take the form $\hat{A}_{kl} \otimes \hat{I}$, where the $\hat{A}_{kl}$ are arbitrary operators acting on the selected M subsystems and the $\hat{I}$ is the identity on the remaining N. The probabilities p_k can thus be rewritten as $p_k = \text{Tr}[(\hat{E}_k \otimes \hat{I})\hat{\rho}^{(M+N)}] = \text{Tr}_M(\hat{E}_k \hat{\rho}^{(M)})$, where $\hat{\rho}^{(M)}$ is the prior mar-

ginal density operator of the measured subsystems and Tr_M denotes a trace over the measured subsystems.

Performing a partial trace over the selected M subsystems yields the posterior state of the remaining N subsystems,

$$\hat{\rho}_k^{(N)} = \mathrm{Tr}_M(\hat{\rho}_k^{(M+N)}). \tag{6}$$

An exact quantum analogue of the classical Bayes rule would write this posterior state as a mixture in which the updating as a consequence of obtaining result k (the "data") would appear in the probabilities in the mixture, but not in the density operators that contribute to the mixture. Classically it is possible to obtain information about a system without disturbing it, while quantum mechanically it is not; hence, Eq. (6) generally includes both updating due to the information acquired and due to the disturbing effects of the measurement. The quantum updating takes the form of the classical Bayes rule only if the measured and unmeasured systems in Eqs. (5) and (6) are initially *unentangled*, so that the measurement does not disturb the unmeasured systems.

Notice also that for a product prior,

$$\hat{\rho}^{(M+N)} = \hat{\rho}_0^{\otimes(M+N)} \equiv \hat{\rho}_0 \otimes \cdots \otimes \hat{\rho}_0, \tag{7}$$

where $\hat{\rho}_0$ is some state on $\mathcal{H}$ and the tensor product contains $M+N$ terms, the posterior state is $\hat{\rho}_k^{(N)} = \hat{\rho}_0^{\otimes N}$, irrespective of the measurement result. No learning from data is possible for product priors. This shows in particular that the totally mixed state for $M+N$ subsystems, which is both a product state and the state of maximum entropy on $\mathcal{H}^{\otimes(M+N)}$, does not allow learning from measured data.

In many practical situations, one can restrict attention to prior states of the form

$$\hat{\rho}^{(N)} = \int d\hat{\rho}\, p(\hat{\rho})\hat{\rho}^{\otimes N}, \tag{8}$$

where $d\hat{\rho}$ is a measure on density operator space and $p(\hat{\rho})$ is a normalized generating function, $\int d\hat{\rho}\, p(\hat{\rho}) = 1$. Prior states of the form (8) arise, e.g., if each subsystem is prepared in the same, unknown way, as in quantum state tomography. A state of N subsystems, $\hat{\rho}^{(N)}$, can be expressed in the form (8), if and only if it is *exchangeable*, i.e., if (i) it is invariant under permutations of the subsystems and (ii) for any $M > 0$, there is a state $\hat{\rho}^{(N+M)}$ of $N+M$ subsystems that is invariant under permutations of the subsystems and that satisfies $\hat{\rho}^{(N)} = \mathrm{Tr}_M(\hat{\rho}^{(N+M)})$ [21,22]. The expansion (8) is then unique. This is the quantum version of the fundamental representation theorem due to de Finetti [23]; for an elementary proof of the quantum theorem, see Ref. [24].

The significance of part (ii) of the definition of exchangeability given above is illustrated by the GHZ state $\hat{\rho}_{\mathrm{GHZ}} = |\psi_{\mathrm{GHZ}}\rangle\langle\psi_{\mathrm{GHZ}}|$, where $|\psi_{\mathrm{GHZ}}\rangle = (|000\rangle + |111\rangle)/\sqrt{2}$. This three-particle state is invariant under permutations of the three subsystems, but it cannot be obtained by a partial trace from a four-particle state that is invariant under permutations

of all four particles. The GHZ state is thus not exchangeable, in accordance with the fact that it cannot be written in the form (8).

If the condition of exchangeability is fulfilled, the question of finding a suitable prior state reduces to finding a suitable prior measure $p(\hat{\rho})d\rho$ in the expansion (8). Much work has been done on suitable prior measures on density operator space (see, e.g., Refs. [12,25–27]). As in the classical theory of inference [1], there exists no unique choice of prior measure; different kinds of prior information lead to different prior measures.

The rule of inference, however, becomes extremely simple if the prior state is of the form (8). In this case, we show below that if a measurement performed on the first subsystem yields result k, the posterior state of the remaining $N-1$ subsystems is given by

$$\hat{\rho}_k^{(N-1)} = \int d\hat{\rho}\, p(\hat{\rho}|k)\hat{\rho}^{\otimes(N-1)}, \tag{9}$$

where

$$p(\hat{\rho}|k) = \frac{p(k|\hat{\rho})p(\hat{\rho})}{p_k}. \tag{10}$$

Here $p(k|\hat{\rho}) = \mathrm{Tr}(\hat{E}_k\hat{\rho})$ is the probability of obtaining the measurement result k for a single subsystem, given that the state of the single subsystem is $\hat{\rho}$, and $p_k = \int d\hat{\rho}\, p(k|\hat{\rho})p(\hat{\rho})$ is the average probability of obtaining k. This is the quantum Bayes rule; it is completely analogous to the classical rule (1).

In the case that the integration in Eq. (8) is restricted to pure states, the rule (10) has been derived by Jones [6] and applied to purifications of mixed states by Bužek *et al.* [14]. Tarrach and Vidal [15] have used Eq. (10) to find optimal measurements on N copies of system, identically prepared in an unknown mixed state by some preparation device. To our knowledge, Eq. (10) has not been derived in the general context considered here.

If measurements are performed on several subsystems individually, the rule (10) can be simply iterated. The situation considered here, where measurements are done one subsystem at a time, is in practice the most important, but it is straightforward to generalize the rule to the case of collective measurements on several subsystems.

Strictly speaking, the generating function $p(\hat{\rho})$ should not be called a probability—after all, a mixed state $\hat{\rho}$ is itself a summary of incomplete knowledge about a subsystem. Nevertheless, the content of the quantum Bayes rule (10) is that the functions $p(\hat{\rho})$ and $p(\hat{\rho}|k)$ can be used as if they were a prior probability and a conditional posterior probability for density operators. This interpretation is obviously appropriate in the case that the exchangeable state (8) is known to have arisen from an experiment in which each subsystem is prepared in the same unknown state, with $p(\hat{\rho})$ being the probability that this unknown state is $\hat{\rho}$.

To derive the rule (10), we denote by $\{\mathcal{F}_k\}$ the set of completely positive, trace-decreasing operations, which de-

scribe the measurement on the first subsystem. The result of the measurement is k with probability

$$p_k = \text{Tr}[\mathcal{F}_k(\hat{\rho}^{(N)})] = \int d\hat{\rho} \, p(k|\hat{\rho}) p(\hat{\rho}). \qquad (11)$$

If the measurement result is k, the state of all N subsystems after the measurement is

$$\hat{\rho}_k^{(N)} = \frac{1}{p_k} \int d\hat{\rho} \, p(\hat{\rho}) \mathcal{F}_k(\hat{\rho}) \otimes \hat{\rho}^{\otimes(N-1)}, \qquad (12)$$

where, by a slight abuse of notation, we denote by $\mathcal{F}_k(\hat{\rho})$ the unrenormalized state of a single subsystem with premeasurement state $\hat{\rho}$ conditioned on the measurement result k. A partial trace over the first subsystem gives the state of the remaining $N-1$ subsystems

$$\hat{\rho}_k^{(N-1)} = \text{Tr}_1(\hat{\rho}_k^{(N)})$$

$$= \frac{1}{p_k} \int d\hat{\rho} \, p(\hat{\rho}) \text{Tr}[\mathcal{F}_k(\hat{\rho})] \hat{\rho}^{\otimes(N-1)}$$

$$= \frac{1}{p_k} \int d\hat{\rho} \, p(\hat{\rho}) \text{Tr}(\hat{E}_k \hat{\rho}) \hat{\rho}^{\otimes(N-1)}$$

$$= \int d\hat{\rho} \frac{p(\hat{\rho}) p(k|\hat{\rho})}{p_k} \hat{\rho}^{\otimes(N-1)}$$

$$= \int d\hat{\rho} \, p(\hat{\rho}|k) \hat{\rho}^{\otimes(N-1)}. \qquad (13)$$

In the last line we have substituted $p(\hat{\rho}|k)$ for the right-hand side of Eq. (10). This completes the derivation.

We now illustrate the quantum Bayes rule for a system of $M+N$ qubits, for which the Hilbert space $\mathcal{H}$ of each subsystem is two dimensional. An arbitrary exchangeable state of $M+N$ qubits can be written in the form

$$\hat{\rho}^{(M+N)} = \int dx \, dy \, dz \, p(x,y,z) \hat{\rho}_{x,y,z}^{\otimes(M+N)}, \qquad (14)$$

where $\hat{\rho}_{x,y,z} = \frac{1}{2}(\hat{I} + x\hat{\sigma}_x + y\hat{\sigma}_y + z\hat{\sigma}_z)$ and the integrals range over the volume of the sphere of radius 1. Here $\hat{\sigma}_x$, $\hat{\sigma}_y$, and $\hat{\sigma}_z$ are the Pauli operators, and $\hat{I}$ denotes the unit operator.

Now assume that $\hat{\sigma}_z$ measurements are performed on M qubits. The probability of obtaining result ± 1, given state $\hat{\rho}_{x,y,z}$, in a $\hat{\sigma}_z$ measurement on a single qubit, is

$$p(\pm 1|\hat{\rho}_{x,y,z}) = \frac{1}{2}(1 \pm z). \qquad (15)$$

If the M measurements of $\hat{\sigma}_z$ yield M_+ results of $+1$ and M_- results of -1, where $M_+ + M_- = M$, then the state of the remaining N qubits is

$$\hat{\rho}_{M_+,M_-}^{(N)} = \int dx \, dy \, dz \, p(x,y,z|M_+,M_-) \hat{\rho}_{x,y,z}^{\otimes N}, \qquad (16)$$

where

$$p(x,y,z|M_+,M_-) = \mathcal{N} p(x,y,z) \left(\frac{1+z}{2}\right)^{M_+} \left(\frac{1-z}{2}\right)^{M_-}, \qquad (17)$$

$\mathcal{N}$ being a normalization factor. In the limit $M \to \infty$, assuming $(M_+ - M_-)/M \to E_z$, we obtain

$$p(x,y,z|M_+,M_-) \to p(x,y|E_z) \delta(z - E_z), \qquad (18)$$

where $p(x,y|E_z) = p(x,y,E_z)/\int dx \, dy \, p(x,y,E_z)$ is the prior conditional probability for x and y, given that $z = E_z$. Equation (18) expresses clearly the gain in information about z.

Consider, for example, a reflection-invariant prior, i.e., a prior satisfying $p(x,y,z) = p(-x,y,z) = p(x,-y,z) = p(x,y,-z)$. An isotropic prior, for which $p(x,y,z) = p(\sqrt{x^2+y^2+z^2})$, is a special case of this sort of prior. For a reflection-invariant prior, the marginal state for a single subsystem before any measurements is the maximally mixed state $\hat{\rho}^{(1)} = \frac{1}{2}\hat{I}$. After M measurements of $\hat{\sigma}_z$, in the limit $M \to \infty$, the posterior state for a single additional subsystem is

$$\hat{\rho}_{E_z}^{(1)} = \frac{1}{2}(\hat{I} + E_z \hat{\sigma}_z), \qquad (19)$$

which is the state obtained in Ref. [14] for the case of an isotropic prior. Our analysis puts this in a clear perspective: the data dictate the expectation value $\langle \hat{\sigma}_z \rangle = E_z$ for the state (19); for a reflection-invariant prior, the $\hat{\sigma}_z$ measurements tell one nothing about the direction of the spin in the x-y plane, so $\hat{\sigma}_x$ and $\hat{\sigma}_y$ retain the zero expectation values that apply to the prior marginal state of a single subsystem.

It is important to note that the state $\hat{\rho}_{E_z}^{(1)}$ does not allow one to make predictions about frequencies in future repeated measurements of, e.g., the observable $\hat{\sigma}_x$. Although $\text{Tr}(\hat{\rho}_{E_z}^{(1)} \hat{\sigma}_x) = 0$, it would be wrong to predict that the frequency of the outcome $+1$ in a large number of future $\hat{\sigma}_x$ measurements will be close to 1/2. The correct prediction for future $\hat{\sigma}_x$ measurements follows from the full state $\hat{\rho}^{(N)}$ with the limiting posterior (18); for the probability of obtaining N_+ results of $+1$ and N_- results of -1 in N measurements of $\hat{\sigma}_x$, we get

$$p(N_+,N_-|E_z) = \frac{N!}{N_+!N_-!} \int dx \, dy \, p(x,y|E_z)$$

$$\times \left(\frac{1+x}{2}\right)^{N_+} \left(\frac{1-x}{2}\right)^{N_-}. \qquad (20)$$

Only in the extreme case that the prior has the special form $p(x,y,z) = p(y,z)\delta(x)$ does the probability (20) become identical to the prediction $P(N_+,N_-) = 2^{-N} N!/N_+!N_-!$ that would follow from assigning the product state $\hat{\rho}_{E_z}^{(1)\otimes N}$ to

the N subsystems. It is clear that this prediction is not implied by the $\hat{\sigma}_z$ measurement data and is therefore unwarranted unless there is additional prior information.

The marginal state $\hat{\rho}_{E_z}^{(1)}$ of Eq. (19) can also be derived from the principle of maximum entropy [28,29]. If all that is known about the state $\hat{\rho}$ of some system is the expectation value of one or several observables, the maximum-entropy principle assigns the state $\hat{\rho}$ that maximizes the von Neumann entropy subject to the constraints given by the expectation values (see Ref. [30] for a derivation of the maximum-entropy principle in the quantum case).

In the example above, the maximum-entropy assignment following from the constraint $\langle\hat{\sigma}_z\rangle=E_z$ for a single subsystem, is identical to the marginal state (19). This identity has also been noted by Bužek *et al.* [14], who state that '' . . . as soon as the number of measurements becomes large then [the] Bayesian inference scheme becomes equal to the reconstruction scheme based on the Jaynes principle of maximum entropy . . . .'' This statement is misleading, however, since the equality holds only for the marginal state of a single subsystem (and even then only for special priors, such as a reflection-invariant prior). Unlike the full state $\hat{\rho}_{M_+,M_-}^{(N)}$ in Eq. (16), found via the Bayes rule, the single-subsystem state $\hat{\rho}_{E_z}^{(1)}$ derived via the maximum-entropy principle does not allow one to make predictions for measurements on more than one subsystem.

On the other hand, applying the maximum-entropy principle directly to N subsystems, fails for the following reason, well known from classical probability theory [31,32]. Maximizing the von Neumann entropy of $\hat{\rho}^{(N)}$ subject to the constraint that $\langle\hat{\sigma}_z\rangle=E_z$ for each subsystem, yields the product state $\hat{\rho}_{\mathrm{MAXENT}}^{(N)}=\hat{\rho}_{E_z}^{(1)\otimes N}$. As discussed above, this state assignment is unwarranted because it leads to predictions for, say, future $\hat{\sigma}_x$ measurements, which are in no way implied by the constraint on $\langle\hat{\sigma}_z\rangle$. Furthermore, any product state assignment precludes learning from subsequent measurements, even though that should be possible, as was discussed in the paragraph after Eq. (7).

If the measurements on individual subsystems correspond to an informationally complete POVM [20] or if they contain sequences of measurements of a tomographically complete set of observables [14], the posterior probability on density operators approaches a δ function in the limit of many measurements. This is the case of quantum state tomography [7–9], which can thus be viewed as a special case of quantum Bayesian inference. In this limit, the exact form of the prior probability on density operators becomes irrelevant. In all other situations, however, there will be some dependence on this prior.

[1] J. M. Bernardo and A. F. M. Smith, *Bayesian Theory* (Wiley, Chichester, 1994).

[2] D. Malakoff, Science **286**, 1460 (1999).

[3] S. Sýkora, J. Stat. Phys. **11**, 17 (1974).

[4] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).

[5] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).

[6] K. R. W. Jones, Ann. Phys. (N.Y.) **207**, 140 (1991).

[7] K. Vogel and H. Risken, Phys. Rev. A **40**, 2847 (1989).

[8] D.T. Smithey, M. Beck, M.G. Raymer, and A. Faridani, Phys. Rev. Lett. **70**, 1244 (1993).

[9] U. Leonhardt, Phys. Rev. Lett. **74**, 4101 (1995).

[10] S. Massar and S. Popescu, Phys. Rev. Lett. **74**, 1259 (1995).

[11] Z. Hradil, Phys. Rev. A **55**, R1561 (1997).

[12] P.B. Slater, J. Math. Phys. **38**, 2274 (1997).

[13] R. Derka, V. Bužek, and A.K. Ekert, Phys. Rev. Lett. **80**, 1571 (1998).

[14] V. Bužek, R. Derka, G. Adam, and P. L. Knight, Ann. Phys. (N.Y.) **266**, 454 (1998).

[15] R. Tarrach and G. Vidal, Phys. Rev. A **60**, 3339 (1999).

[16] K. Banaszek, G.M. D'Ariano, M.G.A. Paris, and M.F. Sacchi, Phys. Rev. A **61**, 10 304 (2000).

[17] R.D. Gill and S. Massar, Phys. Rev. A **61**, 42 312 (2000).

[18] J.C. Lemm, J. Uhlig, and A. Weiguny, Phys. Rev. Lett. **84**, 2068 (2000).

[19] M. Paini, e-print quant-ph/0002078.

[20] K. Kraus, *States, Effects, and Operations. Fundamental Notions of Quantum Theory*, Lecture Notes in Physics Vol. 190 (Springer-Verlag, Berlin, 1983).

[21] E. Størmer, J. Funct. Anal. **3**, 48 (1969).

[22] R.L. Hudson and G.R. Moody, Z. Wahrscheinlichkeitstheor. Verwandte Geb. **33**, 343 (1976).

[23] B. de Finetti, *Theory of Probability* (Wiley, New York, 1990).

[24] C. M. Caves, C. A. Fuchs, and R. Schack, e-print quant-ph/0104088.

[25] D. Bures, Trans. Am. Math. Soc. **135**, 199 (1969).

[26] S.L. Braunstein and C.M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).

[27] K. Życzkowski, P. Horodecki, A. Sanpera, and M. Lewenstein, Phys. Rev. A **58**, 883 (1998).

[28] E.T. Jaynes, Phys. Rev. **106**, 620 (1957).

[29] E.T. Jaynes, Phys. Rev. **108**, 171 (1957).

[30] R. Balian and N. L. Balazs, Ann. Phys. (N.Y.) **179**, 97 (1987).

[31] E. T. Jaynes, in *Maximum Entropy and Bayesian Methods in Applied Statistics*, edited by J. H. Justice (Cambridge University, Cambridge, 1986), p. 26.

[32] J. Skilling, in *Maximum Entropy and Bayesian Methods*, edited by J. Skilling (Kluwer Academic, Dordrecht, The Netherlands, 1989), p. 45.