PHYSICAL REVIEW A **67**, 010101(R) (2003)

Compatibility of quantum states

David Poulin* and Robin Blume-Kohout†
Theoretical Division, Los Alamos National Laboratory, MS-B210, Los Alamos, New Mexico 87545
(Received 16 May 2002; published 15 January 2003)

We introduce a measure of compatibility between quantum states—the likelihood that two density matrices describe the same object. Our measure is motivated by two elementary requirements, which lead to a natural definition. We list some properties of this measure, and discuss its relation to the problem of combining two observers' states of knowledge.

DOI: 10.1103/PhysRevA.67.010101 PACS number(s): 03.65.Ta, 02.50.−r, 03.67.−a

The quantum superposition principle induces a qualitative difference between classical and quantum states of knowledge. The state of a quantum system can be fully specified, yet not predict with certainty the outcome of a measurement—a state of affairs which has only the observers' ignorance as classical analog. In quantum mechanics, incomplete knowledge is represented by a mixed density matrix—though a mixed density matrix does not necessarily reflect ignorance—which corresponds imperfectly to a classical distribution; the ''quantum uncertainty'' of pure states combines with the ''classical uncertainty'' of a distribution to yield an object, which can be represented by different decompositions or realizations.

The fidelity of two quantum states [1] ρ_A and ρ_B,

$$F(\rho_A,\rho_B)=\mathrm{Tr}\{\sqrt{(\rho_A)^{1/2}\rho_B(\rho_A)^{1/2}}\} \tag{1}$$

(or more precisely F^2), measures the likelihood that various measurements made on the two states will obtain the same result. Thus, fidelity is a measure of similarity between states, which does not distinguish between classical and quantum uncertainty.

In this paper, we introduce *compatibility*, a measure similar to fidelity, but which compares two observers' states of knowledge, not the results of the measurements which they could do. We want the compatibility to measure classical admixture, while treating different pure states as fundamentally different: if two observers claim to have complete knowledge of a system, their descriptions had better agreed completely [8]. Hence, a compatibility measure $C(\rho_A,\rho_B)$ should satisfy the two following requirements.

(1) When $[\rho_A,\rho_B]=0$ (classical mixture) the compatibility should be equal to the fidelity.

(2) The compatibility of incompatible states should be 0.

While our first requirement should be transparent, the second sounds tautological, and requires further explanation. Consider two observers (Alice and Bob) whose respective states of knowledge are described by ρ_A and ρ_B. (Throughout this paper, we use subscript k to designate either A or B.) Brun, Finkelstein, and Mermin [2] defined Alice's and Bob's descriptions to be *compatible* if and only if they *could* be describing the same physical system. They then addressed the following question: under what conditions are ρ_A and ρ_B compatible? Their answer is quite simple: ρ_A and ρ_B are compatible if and only if the intersection of their supports, $\mathcal{S}=\mathcal{S}(\rho_A)\cap\mathcal{S}(\rho_B)$, is nonempty. The support $\mathcal{S}(\rho)$ of a density matrix ρ is the complement of its null space $\mathcal{N}(\rho)$; to obtain the projector $P_{\mathcal{S}(\rho)}$ onto $\mathcal{S}(\rho)$, diagonalize ρ and replace each nonzero eigenvalue with 1. Thus, ρ_A and ρ_B are compatible if $P_{\mathcal{S}(\rho_A)}P_{\mathcal{S}(\rho_B)}$ has at least one unit eigenvalue. In other words, two states of knowledge are incompatible if between them they rule out all possible pure states.

With this definition, state $\rho_A=|0\rangle\langle 0|$ is compatible with state $\rho_B=\epsilon|0\rangle\langle 0|+(1-\epsilon)|1\rangle\langle 1|$ and state $\rho_B'=(1-\epsilon)|0\rangle\langle 0|+\epsilon|1\rangle\langle 1|$ as long as $0<\epsilon<1$. Nevertheless, as $\epsilon\to 0$, it is clear that the compatibility of ρ_A and ρ_B should vanish while that of ρ_A and ρ_B' should approach unity. The definition of Ref. [2] makes no distinction between these two cases and this is what originally motivated the present work.

Now that requirement (2) has been clarified, we can proceed with the definition of the compatibility measure.

Definition 1. Let $\mathfrak{B}_0(\mathcal{H})$ be the set of all density matrices on Hilbert space $\mathcal{H}$. For $\rho\in\mathfrak{B}_0(\mathcal{H})$, define $\mathcal{P}(\rho)$ as the set of realizations of ρ:

$$\mathcal{P}(\rho)=\left\{P:\int_{\mathfrak{B}_0(\mathcal{H})}P(\sigma)\sigma d\sigma=\rho\right\}, \tag{2}$$

where the P are probability distributions over $\mathfrak{B}_0(\mathcal{H})$. Then, the compatibility of ρ_A and $\rho_B\in\mathfrak{B}_0(\mathcal{H})$ is defined as

$$C(\rho_A,\rho_B)=\max_{\substack{P_A\in\mathcal{P}(\rho_A)\\P_B\in\mathcal{P}(\rho_B)}}\int_{\mathfrak{B}_0(\mathcal{H})}\sqrt{P_A(\sigma)P_B(\sigma)}d\sigma, \tag{3}$$

the integral representing the classical fidelity $F(P_A,P_B)$ (or statistical overlap) of two classical distributions P_A and P_B.

Lemma 1. All distributions $P\in\mathcal{P}(\rho)$ must vanish outside $\mathfrak{B}_0(\rho)$: the set of density matrices with support restricted to $\mathcal{S}(\rho)$.

Proof. Let $P(\sigma)$ be a realization of ρ. We can separate ρ into two parts:

*Present address: Physics Department and IQC, University of Waterloo, Waterloo, Canada N2L 3G1.
Email address: dpoulin@iqc.ca
†Email address: rbk@socrates.Berkeley.edu

$$\rho = \int_{\mathfrak{B}_0^{\perp}(\rho)} P_k(\sigma)\sigma \; d\sigma + \int_{\mathfrak{B}_0(\rho)} P_k(\sigma)\sigma \; d\sigma \tag{4}$$

$$= p\rho' + (1-p)\rho'', \tag{5}$$

where ρ', by definition, has support on $\mathcal{N}(\rho)$, ρ'' has support strictly on $\mathcal{S}(\rho)$ and $p = \int_{\mathfrak{B}_0^{\perp}(\rho)} P_k(\sigma)\, d\sigma$. If $p \neq 0$, there exists $\psi \in \mathcal{N}(\rho)$ such that ρ does not annihilate ψ. This contradicts the definition of $\mathcal{N}(\rho)$, so we conclude that $p=0$ and therefore P is restricted to $\mathfrak{B}_0(\rho)$.

Theorem 1. Definition 1 satisfies both of our requirements.

Proof.

(1) If ρ_A commutes with ρ_B, then they have orthogonal decompositions onto the same set of pure states: $\rho_A = \Sigma_i a_i |\phi_i\rangle\langle\phi_i|$; $\rho_B = \Sigma_i b_i |\phi_i\rangle\langle\phi_i|$. Thus $C(\rho_A,\rho_B) \geqslant \Sigma_i \sqrt{a_i b_i} = F(\rho_A,\rho_B)$. Later (see property 4) we show that $C(\rho_A,\rho_B) \leqslant F(\rho_A,\rho_B)$; therefore for commuting density matrices $C(\rho_A,\rho_B) = F(\rho_A,\rho_B)$.

(2) If ρ_A and ρ_B are incompatible, their supports are disjoint, which implies that $P_A(\sigma)$ and $P_B(\sigma)$ are restricted to disjoint sets, implying that $C(\rho_A,\rho_B)=0$.

Note that this measure is not the only one that satisfies our two requirements. For example, define

$$D_n(\rho_A,\rho_B) = Tr\{[(\rho_A)^{1/2n}(\rho_B)^{1/n}(\rho_A)^{1/2n}]^n\}. \tag{6}$$

Clearly, $D_n(\rho_a,\rho_B) = F(\rho_A,\rho_B)$ when $n=1$ or for any n when $[\rho_A,\rho_B]=0$, so it satisfies our first requirement. For the second requirement, notice that

$$\lim_{n\to\infty} D_n(\rho_A,\rho_B) \leqslant \lim_{n\to\infty} Tr\{[P_{\mathcal{S}(\rho_A)} P_{\mathcal{S}(\rho_B)}]^n\}, \tag{7}$$

which is 0 if $P_{\mathcal{S}(\rho_B)} P_{\mathcal{S}(\rho_B)}$ has no unit eigenvalue, i.e., if ρ_A and ρ_B are incompatible. Therefore, $D(\rho_A,\rho_B) = \lim_{n\to\infty} D_n(\rho_A,\rho_B)$ is a valid measure of compatibility.

Definition 1 can also be generalized to

$$E_\alpha(\rho_A,\rho_B) = \max_{\substack{P_A \in \mathcal{P}(\rho_A) \\ P_B \in \mathcal{P}(\rho_B)}} \int_{\mathfrak{B}_0(\mathcal{H})} [P_A(\sigma)]^\alpha [P_B(\sigma)]^{1-\alpha} d\sigma, \tag{8}$$

$0<\alpha<1$, which is the Rényi overlap of P_A and P_B, the fidelity corresponding to the special case $\alpha = 1/2$. This definition allows for an asymmetry between Alice and Bob, which can be useful when one of the participant is more trustworthy than the other.

Although these alternative definitions offer some interesting features, we shall concentrate on Definition 1 in the following. [Labels D and E indicate that the results also hold for measure $D(\rho_A,\rho_B)$ and $E_\alpha(\rho_A,\rho_B)$, respectively, the proofs are given for $C(\rho_A,\rho_B)$ only.]

Theorem 2 (E). To compute the compatibility of two states, it is sufficient to maximize over pure state realizations. In other words,

$$C(\rho_A,\rho_B) = \max_{\substack{Q_A \in \mathcal{Q}(\rho_A) \\ Q_B \in \mathcal{Q}(\rho_B)}} \int_{\mathfrak{B}_0^1(\mathcal{H})} \sqrt{Q_A(\psi) Q_B(\psi)}\, d\psi, \tag{9}$$

where $\mathfrak{B}_0^1(\mathcal{H})$ is the set of all pure states in $\mathcal{H}$ and $\mathcal{Q}(\rho)$ is the set of pure state realizations of ρ:

$$\mathcal{Q}(\rho) = \left\{ Q : \int_{\mathfrak{B}_0^1(\mathcal{H})} Q(\psi) |\psi\rangle\langle\psi| d\psi = \rho \right\}, \tag{10}$$

Q are probability distributions on $\mathfrak{B}_0^1(\mathcal{H})$.

Proof. Choose a standard pure state decomposition for $\sigma \in \mathfrak{B}_0(\mathcal{H})$, $\sigma = \int_{\mathfrak{B}_0^1(\mathcal{H})} f_\sigma(\psi) |\psi\rangle\langle\psi| d\psi$ (e.g., eigendecomposition). Then

$$\begin{aligned} &\int_{\mathfrak{B}_0(\mathcal{H})} \sqrt{P_A(\sigma) P_B(\sigma)}\, d\sigma \\ &\quad = \int_{\mathfrak{B}_0(\mathcal{H})} \int_{\mathfrak{B}_0^1(\mathcal{H})} \sqrt{P_A(\sigma) f_\sigma(\psi) P_B(\sigma) f_\sigma(\psi)}\, d\sigma d\psi \\ &\quad \leqslant \int_{\mathfrak{B}_0^1(\mathcal{H})} \sqrt{Q_A(\psi) Q_B(\psi)}\, d\psi, \end{aligned} \tag{11}$$

since fidelity can only increase under the marginalization $Q_k(\psi) = \int_{\mathfrak{B}_0(\mathcal{H})} P_k(\sigma) f_\sigma(\psi)\, d\sigma$.

Theorem 3. When one of the two states is pure (say ρ_B), $C(\rho_A,\rho_B) = \sqrt{p}$ where p is given by

$$p = \min_{q \in [0,1]} \{ q : \det_{\mathcal{S}(\rho_A)} \{\rho_A - q\rho_B\} = 0 \}, \tag{12}$$

if ρ_B lies within $\mathcal{S}(\rho_A)$ and $p=0$ otherwise.

Proof. There is a unique realization for ρ_B: $P_B(\sigma) = \delta(\sigma - \rho_B)$. The maximum value of q for which we can write $\rho_A = q\rho_B + (1-q)\sigma$ (with σ a valid density matrix) is p. The result follows.

Theorem 4 (E). Any local maximum of $F(P_A,P_B)$ over $\mathcal{P}(\rho_A) \otimes \mathcal{P}(\rho_B)$ is a global maximum.

Proof. Fidelity is a concave function: $F(\lambda P_A + [1-\lambda]P_A', P_B) \geqslant \lambda F(P_A,P_B) + [1-\lambda] F(P_A',P_B)$. The sets $\mathcal{P}(\rho_A)$ and $\mathcal{P}(\rho_B)$ are convex: any convex combination of valid probability distributions of mean ρ is also a valid probability distribution with mean ρ. The result follows automatically.

We now give a list of properties of the compatibility measure.

Property 1 (D). $C(\rho_A,\rho_B)$ is symmetric.

Property 2 (D and E). Compatibility is invariant under unitary transformation: $C(U\rho_A U^\dagger, U\rho_B U^\dagger) = C(\rho_A,\rho_B)$.

Property 3 (D and E). For pure states $C(\psi_A,\psi_B) = 1$ if and only if $|\langle\psi_A|\psi_B\rangle|^2 = 1$ and 0 otherwise.

Property 4 (D). (Upper bound) $C(\rho_A,\rho_C) \leqslant F(\rho_A,\rho_B)$.

Property 5 (D and E). $F(\rho_A,\rho_B) = 0 \Rightarrow C(\rho_A,\rho_B) = 0$ and $F(\rho_A,\rho_B) = 1 \Leftrightarrow C(\rho_A,\rho_B) = 1 \Leftrightarrow \rho_A = \rho_B$.

Property 6 (Lower bound) $C(\rho_A,\rho_B) \geqslant r\sqrt{\epsilon_A \epsilon_B}$, where ϵ_k is the greatest value of q for which one can write $\rho_k = (q/r)P_{\mathcal{S}} + (1-q)\sigma$ with σ being a valid density matrix,

see Eq. (12), and $r=\mathrm{Tr}\{P_{\mathcal{S}}\}$ is the dimension of $\mathcal{S}=\mathcal{S}(\rho_A)\cap\mathcal{S}(\rho_B)$. (For compatible states, $\epsilon_k\geqslant\lambda_k^0$, the smallest nonzero eigenvalue of ρ_k.)

Property 7 (E). (Multiplicativity) $C(\rho_A\otimes\rho_A',\rho_B\otimes\rho_B')\geqslant C(\rho_A,\rho_B)C(\rho_A',\rho_B')$.

Proofs. Properties 1, 2, and 3 are straightforward from Definition 1.

Property 4: Assume that $Q_k(\psi)$ are the optimal distributions given by Theorem 2. We choose $\vec{x}\in\mathbb{R}^{(2N-2)}$ (where N is the dimension of $\mathcal{H}$) as a parametrization for $\mathfrak{B}_0^1(\mathcal{H})$: $\psi=\psi(\vec{x})$, and construct the purifications

$$|\Psi_k\rangle=\int\sqrt{Q_k(\psi(\vec{x}))}|\psi(\vec{x})\rangle\otimes|\vec{x}\rangle d\vec{x}, \tag{13}$$

where $\vec{x}$ is now treated as a quantum continuous variable $\langle\vec{x}|\vec{x}'\rangle=\delta(\vec{x}-\vec{x}')$ (e.g., position of an particle in a N-dimensional box). Then $C(\rho_A,\rho_B)=\langle\Psi_A|\Psi_B\rangle\leqslant F(\rho_A,\rho_B)$, since the fidelity is the maximum of this quantity over all purifications.

This proof introduces an interesting distinction between fidelity and compatibility. Fidelity is the optimal inner product between all purifications of ρ_A and ρ_B. On the other hand, compatibility involves purifications of a very special kind [Eq. (13)]. All that is needed to transform compatibility into fidelity is to replace Eq. (13) by

$$|\Psi_A\rangle=\int\sqrt{Q_A(\psi(\vec{x}))}|\psi(\vec{x})\rangle\otimes U_A|\vec{x}\rangle d\vec{x},$$

$$|\Psi_B\rangle=\int\sqrt{Q_B(\psi(\vec{x}))}|\psi(\vec{x})\rangle\otimes U_B|\vec{x}\rangle d\vec{x} \tag{14}$$

for arbitrary unitary operators U_A and U_B.

Property 5 follows from $F(\rho_A,\rho_B)=1\Leftrightarrow\rho_A=\rho_B$, requirement (1), and property 4.

Property 6: We can choose a distribution where ρ_k has probability $r\epsilon_k$ at the point $\sigma=P_{\mathcal{S}}/r$.

Property 7: The product of the optimal distributions for $C(\rho_A,\rho_B)$ and $C(\rho_A',\rho_B')$ are valid distributions over the combined Hilbert space but might not be optimal. We do not know if this inequality can be reduced to an equality. In other words, it is possible that the optimal distribution for $\rho_A\otimes\rho_A'$ and $\rho_B\otimes\rho_B'$ involves nonproduct states.

It is worth mentioning that no smooth function of the compatibility satisfying $f(C)=1\Leftrightarrow C=0$ and $f(C)=0\Leftrightarrow C=1$ can be used to build a metric on $\mathfrak{B}_0(\mathcal{H})$. This is best illustrated by the following two-dimensional example. Assume that states ρ_+ and ρ_- are pure, derived from $|\psi_\pm\rangle=\cos\epsilon|0\rangle\pm\sin\epsilon|1\rangle$, and $\rho_0=(1-\epsilon)|0\rangle\langle0|+\epsilon|1\rangle\langle1|$, where $\epsilon\to0$. One can easily verify that $C(\rho_+,\rho_-)=0$ and $C(\rho_+,\rho_0)=C(\rho_-,\rho_0)=1-O(\epsilon)$ so $f(C(\rho_+,\rho_-))=1>f(C(\rho_+,\rho_0))+f(C(\rho_-,\rho_0))\to0$ as $\epsilon\to0$. This is in contrast with classical distributions: when $[\rho_A,\rho_B]=0$, $\cos^{-1}F(\sqrt{\rho_A}\rho_B)$ is a valid distance measure [3].

Measurement. Suppose that Alice and Bob acquire their knowledge of ρ_A and ρ_B through measurement. These states will always be compatible: incompatible knowledge acquired through measurement would indicate an inconsistency in quantum theory [9]. For example, they can each be given many copies of a quantum system in state ρ of which they initially have no knowledge except the dimension. They carry out independent measurements on those copies and, with the help of Bayesian rules, update their description of the system (see Ref. [4], and references therein). As mentioned earlier, their descriptions will always be compatible. Nevertheless, a low compatibility could result as a consequence of one of the following situations: (i) they were given copies of different states, i.e., the promise of identical systems was broken; (ii) their measurement apparatus is miscalibrated; or (iii) they are in a very improbable branch of the universe.

These eventualities cannot be detected by the fidelity of ρ_A and ρ_B. For example, suppose that, for a two-level system,

$$\rho_A=(1-\epsilon)|0\rangle\langle0|+\frac{\epsilon}{2}\mathbb{1},$$

$$\rho_B=(1-\epsilon)|+\rangle\langle+|+\frac{\epsilon}{2}\mathbb{1}, \tag{15}$$

where $|+\rangle=(1/\sqrt{2})(|0\rangle+|1\rangle)$. As the observers' knowledge becomes more and more accurate ($\epsilon\to0$), the compatibility goes to 0, indicating one of the three situations listed above. On the other hand, fidelity saturates at $F^2=1/2$, which is the same as if both Alice and Bob had a vague knowledge of the state, e.g.,

$$\rho_A=\left(\frac{1}{2}+a\right)|0\rangle\langle0|+\left(\frac{1}{2}-a\right)|1\rangle\langle1|,$$

$$\rho_B=\left(\frac{1}{2}-a\right)|0\rangle\langle0|+\left(\frac{1}{2}+a\right)|1\rangle\langle1|, \tag{16}$$

with $a=\sqrt{2}/4$. This clearly illustrates the fact that fidelity makes no distinction between classical and quantum uncertainty.

Combining knowledge. Now, assume that Alice and Bob want to pool their information. If $C(\rho_A,\rho_B)=0$ (which cannot result from measurement), their "knowledge" is contradictory. When $C(\rho_A,\rho_B)>0$, however, they can combine their states of knowledge to get a new density matrix ρ_{AB}. This issue has recently been studied by Jacobs [5] but with the only conclusion that ρ_{AB} should lie in $\mathcal{S}(\rho_A)\cap\mathcal{S}(\rho_B)$.

We propose that the state obtained from combining two states of knowledge should be the one which is maximally compatible with both of them. This requires a definition of three-way compatibility:

$$C(\rho_A,\rho_B,\rho_C)=\max_{\substack{P_A\in\mathcal{P}(\rho_A)\\P_B\in\mathcal{P}(\rho_B)\\P_C\in\mathcal{P}(\rho_C)}}\int_{\mathfrak{B}_0(\mathcal{H})}\sqrt[3]{P_A(\sigma)P_B(\sigma)P_C(\sigma)}d\sigma. \tag{17}$$

Hence, our rule for combining states of knowledge reads

$$\rho_{AB} = \arg[\max_{\rho} C(\rho_A, \rho_B, \rho)]; \quad (18)$$

in the eventuality that the maximum over ρ is not unique, one can discriminate with a maximum entropy $S(\rho)$ criteria which is well motivated in the current context. For any fixed P_A and P_B, the P_C that optimizes Eq. (17) is proportional to the geometric average of P_A and P_B. Therefore, defining $\tilde{P}_A$ and $\tilde{P}_B$ as the distributions that optimized Eq. (3), we get

$$\rho_{AB} = \int_{\mathfrak{B}_0(\mathcal{H})} P_{AB}(\sigma)\sigma \; d\sigma, \quad (19)$$

where $P_{AB} = \sqrt{\tilde{P}_A \tilde{P}_B}/C(\rho_A, \rho_B)$. Furthermore, there is a simple relation between the optimal three-way compatibility and the compatibility of the two original descriptions: $C(\rho_A, \rho_B, \rho_{AB})^3 = C(\rho_A, \rho_B)^2$.

Knowledge. Knowledge of a quantum system can take many forms; as Bennett expresses it, ''It is possible to *know* or *possess* a quantum state in infinitely many physically inequivalent ways, ranging from complete classical knowledge, through possession of a single specimen of the state, to weaker and less compactly embodiable forms such as the ability to simulate the outcome of a single POVM (positive operator valued measure) measurement on the state [6].''

The compatibility measurement of Eq. (3) is meaningful when we consider classical description of the quantum states; the quantum fidelity [Eq. (1)] corresponds to a situation where single specimens of the quantum states are available (respectively ''knowledge of the quantum'' and ''quantum knowledge''). One can define compatibility measurements according to the *type of knowledge* one is dealing with. For example, we can define the compatibility between a state ρ and an ensemble $\{q_j, \sigma_j\}$ as $\max_{\mathcal{P}(\rho)} F(P,Q)$, where $Q(\sigma) = \Sigma_j q_j \delta(\sigma - \sigma_j)$. While the pure state $|+\rangle = (1/\sqrt{2})(|0\rangle + |1\rangle)$ is compatible with the ensemble $E_1 = \{(1, p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1|)\}$, it is incompatible with the ensemble $E_2 = \{(p, |0\rangle\langle 0|), ((1-p), |1\rangle\langle 1|)\}$, even if they are realizations of the same state.

An ensemble embodies more knowledge than its associated (average) state. In our prescription for combining knowledge, we have assumed that all of Alice's and Bob's knowledge was encapsulated in their respective density matrices. Note that all knowledge can be represented in this form by including ancillary systems [e.g., Eq. (13)].

Suppose, instead, that both Alice's and Bob's states of knowledge are represented by the ensemble E_1. Obviously, their combined density matrix should be $\rho_{AB1} = p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1|$. On the other hand, when both their states of knowledge are E_2, Bayesian rules would suggest that their combined state should be $\rho_{AB2} = p^2|0\rangle\langle 0| + (1-p)^2|1\rangle\langle 1|$ (with proper normalization)—but this assumes that their knowledge was acquired *independently* [5]. If their knowledge came from a redundant source, the Bayesian rule would then yield state ρ_{AB1}, as would our prescription.

Hence, this illustrates that our rule for combining states of knowledge assumes no more information than what is encapsulated in the density matrices. Furthermore, it can quite simply be adapted to different forms of knowledge, either through the use of ancillary systems or of generalized compatibility measures.

The authors would like to acknowledge Howard Barnum, Kurt Jacobs, Harold Ollivier, and Wojciech Zurek for discussions on this subject. R.B.K. would also like to thank Todd Brun for discussions and inspiration.

[1] R. Jozsa, J. Mod. Opt. **41**, 12 (1994); **41**, 2315 (1994).
[2] T.A. Brun, J. Finkelstein, and N.D. Mermin, Phys. Rev. A **65**, 032315 (2002).
[3] C. A. Fuchs, Ph.D. thesis, University of New Mexico, Albuquerque, 1995 (unpublished).
[4] R. Schack, T.A. Brun, and C.M. Caves, Phys. Rev. A **64**, 014305 (2001).
[5] K. Jacobs, Quantum Info. Proc. **1**, 73 (2002).
[6] C. H. Bennett, Fourth Workshop on Quantum Information Processing, Amsterdam, 2001 (unpublished).
[7] C.M. Caves, C.A. Fuchs, and R. Schack, e-print quant-ph/0206110.
[8] Caves, Fuchs, and Schack have recently challenged this statement [7]. The distinction between our studies leading to this disagreement can be summarized by ''state of belief vs state of knowledge.''
[9] Incompatible knowledge could emerge as a consequence of the finite accuracy of the measurement apparatus: nevertheless, such limitations should be taken into account in the state estimation.