Cryptographic Distinguishability Measures for Quantum-Mechanical States

Christopher A. Fuchs and Jeroen van de Graaf

***Abstract*—This paper, mostly expository in nature, surveys four measures of distinguishability for quantum-mechanical states. This is done from the point of view of the cryptographer with a particular eye on applications in quantum cryptography. Each of the measures considered is rooted in an analogous classical measure of distinguishability for probability distributions: namely, the probability of an identification error, the Kolmogorov distance, the Bhattacharyya coefficient, and the Shannon distinguishability (as defined through mutual information). These measures have a long history of use in statistical pattern recognition and classical cryptography. We obtain several inequalities that relate the quantum distinguishability measures to each other, one of which may be crucial for proving the security of quantum cryptographic key distribution. In another vein, these measures and their connecting inequalities are used to define a *single* notion of cryptographic exponential indistinguishability for two families of quantum states. This is a tool that may prove useful in the analysis of various quantum-cryptographic protocols.**

***Index Terms*— Bhattacharyya coefficient, distinguishability of quantum states, exponential indistinguishability, Kolmogorov distance, probability of error, quantum cryptography, Shannon distinguishability.**

I. Introduction

THE field of quantum cryptography is built around the singular idea that physical information carriers are always quantum-mechanical. When this idea is taken seriously, new possibilities open up within cryptography that could not have been dreamt of before. The most successful example of this so far has been quantum-cryptographic key distribution. For this task, quantum mechanics supplies a method of key distribution for which the security against eavesdropping can be assured by physical law itself. This is significant because the legitimate communicators then need make no assumptions about the computational power of their opponent.

Common to all quantum-cryptographic problems is the way information is encoded into quantum systems, namely, through their quantum-mechanical states. For instance, a 0 might be encoded into a system by preparing it in a state ρ_0, and a 1 might likewise be encoded by preparing it in a state ρ_1. The choice of the particular states in the encoding will generally determine not only the ease of information retrieval by the legitimate users, but also the inaccessibility of that information to a hostile opponent. Therefore, if one wants to model and analyze the cryptographic security of quantum protocols, one of the most basic questions to be answered is the following. What does it mean for two quantum states to be "close" to each other or "far" apart? Giving an answer to this question is the subject of this paper. That is, we shall be concerned with defining and relating various notions of "distance" between two quantum states.

Formally a quantum state is nothing more than a square matrix of complex numbers that satisfies a certain set of supplementary properties. Because of this, any of the notions of distance between matrices that can be found in the mathematical literature would do for a quick fix. However, we adhere to one overriding criterion for the "distance" measures considered here. The only physical means available with which to distinguish two quantum states is that specified by the general notion of a quantum-mechanical measurement. Since the outcomes of such a measurement are necessarily indeterministic and statistical, only measures of "distance" that bear some relation to statistical-hypothesis testing will be considered. For this reason, we prefer to call the measures considered herein *distinguishability measures* rather than "distances."

In this paper, we discuss four notions of distinguishability that are of particular interest to cryptography: the probability of an identification error, the Kolmogorov distance (which turns out to be related to the standard trace-norm distance), the Bhattacharyya coefficient (which turns out to be related to Uhlmann's "transition probability"), and the Shannon distinguishability (which is defined in terms of the optimal mutual information obtainable about a state's identity). Each of these four distinguishability measures is, as advertised, a generalization of a distinguishability measure between two probability distributions.

Basing the quantum notions of distinguishability upon classical measures in this way has the added bonus of easily leading to various inequalities between the four measures. In particular, we establish a simple connection between the probability of error and the trace-norm distance. Moreover, we derive a very simple upper bound on the Shannon distinguishability as a function of the trace-norm distance $\mathrm{SD}(\rho_0, \rho_1) \leq$

Manuscript received March 11, 1998; revised January 12, 1999. The work of C. A. Fuchs was supported by a Lee A. DuBridge Fellowship and by DARPA through the Quantum Information and Computing (QUIC) Institute administered by the U.S. Army Research Office. The work of J. van de Graaf was supported by Natural Sciences and Research Council of Canada and FCAR, Québec, Canada..

C. A. Fuchs is with the Bridge Laboratory of Physics 12-33, California Institute of Technology, Pasadena, CA 91125 USA.

J. van de Graaf was with the Laboratoire d'Informatique Théorique et Quantique, Université de Montréal, Montreal, Que., Canada. He is now with CENAPAD, the Center of High Performance Computing, Universidade Federal de Minas Gerais, Belo Horizonte, 31270-010 Brazil.

Communicated by D. Stinson, Associate Editor for Complexity and Cryptography.

Publisher Item Identifier S 0018-9448(99)03178-8.

$\frac{1}{2}\operatorname{Tr}|\rho_0-\rho_1|$. (The usefulness of this particular form for the bound was realized while one of the authors was working on [1], where it is used to prove security of quantum key distribution for a general class of attacks.) Similarly, we can bound the quantum Shannon distinguishability by functions of the quantum Bhattacharrya coefficient.

In another connection, we consider an application of these inequalities to protocol design. In the design of cryptographic protocols, one often defines a *family* of protocols parameterized by a *security parameter*, n—where this number denotes the length of some string, the number of rounds, the number of photons, etc. Typically, the design of a good protocol requires that the probability of cheating for each participant vanishes exponentially fast, i.e., is of the order $O(\varepsilon^n)$, for ε between 0 and 1. As an example, one technique is to compare the protocol implementation (the family of protocols) with the *ideal protocol specification* and to prove that these two become exponentially indistinguishable[1] [2], [3].

To move this line of thought into the quantum regime, it is natural to consider two families of quantum states parameterized by n and to require that the distinguishability between the two families vanishes exponentially fast. *A priori*, this exponential convergence could depend upon which distinguishability measure is chosen—after all, the quantum-mechanical measurements optimal for each distinguishability measure can be quite different. However, with the newly derived inequalities in hand, it is an easy matter to show that exponential indistinguishability with respect to one measure implies exponential indistinguishability with respect to each of the other four measures. In other words, these four notions are equivalent, and it is legitimate to speak of a single, unified *exponential indistinguishability* for two families of quantum states.

The contribution of this paper is threefold. In the first place, even though some of the quantum inequalities derived here are minor extensions of classical inequalities that have been known for some time, many of the classical inequalities are scattered throughout the literature in fields of research fairly remote from the present one. Furthermore, though elements of this work can also be found in [4], there is presently no paper that gives a systematic overview of quantum distinguishability measures from the cryptographer's point of view. In the second place, some of the inequalities in Section VI are new, even within the classical regime. In the third place, a canonical definition for quantum exponential indistinguishability is obtained. The applications of this notion may be manifold within quantum cryptography.

The structure of the paper is as follows. In the following section we review a small bit of standard probability theory, mainly to introduce the setting and notation. Section III discusses density matrices and measurements, showing how the combination of the two notions leads to a probability distribution. In Section IV, we discuss four measures of distinguishability, first for classical probability distrubitions, then for quantum-mechanical states. In Section V, we discuss several inequalities, again both classically and quantum mechanically. In Section VI these inequalities are applied to proving a theorem about exponential indistinguishability. Section VII discusses an application of this notion—in particular, we give a simple proof of a theorem in [5] that the Shannon distinguishability of the parity (i.e., the overall exclusive–or) of a quantum-bit string decreases exponentially with the length of the string. Moreover, the range of applicability of the theorem is strengthened in the process.

This paper is aimed primarily at an audience of computer scientists, at cryptographers in particular, with some small background knowledge of quantum mechanics. Readers needing a more systematic introduction to the requisite quantum theory should consult Hughes [6] or Isham [7], for instance. A very brief introduction can be found in the appendix of [8].

[1]This notion is more commonly called *statistical indistinguishability* in the cryptographic literature. However, since the word "statistical" is likely to already be overused in this paper, we prefer "exponential."

II. Probability Distributions

Let X_0 be a stochastic variable over a finite set $\mathcal{X}$. Then we can define $p_0(x) \stackrel{\text{def}}{=} \operatorname{Prob}[X_0 = x]$, so X_0 induces a probability distribution p_0 over $\mathcal{X}$. Let p_1 be defined likewise. Of course, $\Sigma_{x\in X}\, p_t(x) = 1$ for $t = 0, 1$. After relabeling the outcomes $x_1, x_2, x_3, \cdots, x_m$ to $1, 2, 3, \cdots, m$ we get

		$x=1$	$x=2$	$x=3$	$\cdots$	$x=m$
X_0	$\pi_0=\frac{1}{2}$	$p_0(1)$	$p_0(2)$	$p_0(3)$	$\cdots$	$p_0(m)$
X_1	$\pi_1=\frac{1}{2}$	$p_1(1)$	$p_1(2)$	$p_1(3)$	$\cdots$	$p_1(m)$

Here π_0 and π_1 are the *a priori* probabilities of the two stochastic variables; they sum up to 1. Throughout this paper we take $\pi_0 = \pi_1 = \frac{1}{2}$. (Even though much of our analysis could be extended to the case $\pi_0 \neq \pi_1 \neq \frac{1}{2}$, it seems not too relevant for the questions addressed here.) Two distributions are *equivalent* (i.e., *indistinguishable*) if $p_0(x) = p_1(x)$ for all $x \in \mathcal{X}$, and they are *orthogonal* (i.e., *maximally indistinguishable*) if there exists no x for which both $p_0(x)$ and $p_1(x)$ are nonzero.

Observe that $p_t(x)$ denotes the conditional probability that $X = x$ *given* that $T = t$, written as $\operatorname{Prob}[X = x|T = t]$. So the joint probability is half that value

$$\operatorname{Prob}[X = x \wedge T = t] = \operatorname{Prob}[T = t]\operatorname{Prob}[X = x|T = t] \tag{1}$$

$$= \pi_t p_t(x) \tag{2}$$

$$= \tfrac{1}{2}\, p_t(x). \tag{3}$$

We define the conditional probability $r_t(x) := \operatorname{Prob}[T = t|X = x]$, and the probability that $X = x$ *regardless* of t, that is, $p(x) := \operatorname{Prob}[X = x]$. Using Bayes' Theorem we get

$$r_t(x) = \operatorname{Prob}[T = t|X = x] \tag{4}$$

$$= \operatorname{Prob}[T = t]\operatorname{Prob}[X = x|T = t]/\operatorname{Prob}[X = x] \tag{5}$$

$$= \tfrac{1}{2}\, p_t(x)/p(x.) \tag{6}$$

Observe that $r_0(x)+r_1(x) = 1$ for all x. Using $p(x)$ and $r_t(x)$

we can represent the situation also in the following way:

		$x=1$	$x=2$	$x=3$	$\cdots$	$x=m$
X		$p(1)$	$p(2)$	$p(3)$	$\cdots$	$p(m)$
X_0	$\pi_0=\frac{1}{2}$	$r_0(1)$	$r_0(2)$	$r_0(3)$	$\cdots$	$r_0(m)$
X_1	$\pi_1=\frac{1}{2}$	$r_1(1)$	$r_1(2)$	$r_1(3)$	$\cdots$	$r_1(m)$

III. Density Matrices and Measurements

Recall that a quantum state is said to be a *pure* state if there exists some (fine-grained) measurement that can confirm this fact with probability 1. A pure state can be represented by a normalized vector $|\psi\rangle$ in an N-dimensional Hilbert space, i.e., a complex vector space with inner product. Alternatively, it can be represented by a projection operator $|\psi\rangle\langle\psi|$ onto the rays associated with those vectors. In this paper N is always taken to be finite.

Now consider the following preparation of a quantum system: **A** flips a fair coin and, depending upon the outcome, sends one of two different pure states $|\psi_0\rangle$ or $|\psi_1\rangle$ to **B**. Then the "pureness" of the quantum state is "diluted" by the classical uncertainty about the resulting coin flip. In this case, no deterministic fine-grained measurement generally exists for identifying **A**'s exact preparation, and the quantum state is said to be a *mixed* state. **B**'s knowledge of the system—that is, the source from which he draws his predictions about any potential measurement outcomes—can now no longer be represented by a vector in a Hilbert space. Rather, it must be described by a *density operator* or *density matrix*[2] formed from a statistical average of the projectors associated with **A**'s possible fine-grained preparations.

Definitions 1. (See for Instance [9], [7], [10]): A *density matrix* ρ is an $N\times N$ matrix with unit trace that is Hermitian (i.e., $\rho=\rho^\dagger$) and positive semi-definite (i.e., $\langle\psi|\rho|\psi\rangle\geq 0$ for all $\psi\in\mathcal{H}$).

Example: Consider the case where **A** prepares either a horizontally or a vertically polarized photon. We can choose a basis such that $|\mathsf{H}\rangle=\binom{1}{0}$ and $|\mathsf{V}\rangle=\binom{0}{1}$. Then **A**'s preparation is perceived by **B** as the mixed state

$$\frac{1}{2}|\mathsf{H}\rangle\langle\mathsf{H}|+\frac{1}{2}|\mathsf{V}\rangle\langle\mathsf{V}|=\frac{1}{2}\begin{pmatrix}1&0\\0&0\end{pmatrix}+\frac{1}{2}\begin{pmatrix}0&0\\0&1\end{pmatrix}=\begin{pmatrix}1/2&0\\0&1/2\end{pmatrix} \tag{7}$$

which is the "completely mixed state."

Note that the same density matrix will be obtained if **A** prepares an equal mixture of left-polarized and right-polarized photons. In fact, any equal mixture of two orthogonal pure states will yield the same density matrix.

Any source of quantum samples (that is, any imaginary **A** who secretly and randomly prepares quantum states according to some probability distribution) is called an *ensemble*. This can be viewed as the quantum counterpart of a stochastic variable. A density matrix completely describes **B**'s knowledge of the sample. Two different ensembles with the same density matrix are indistinguishable as far as **B** is concerned; when this is the case, there exists no measurement that can allow **B** a decision between the ensembles with probability of success better than chance.

The fact that a density matrix describes **B**'s *a priori* knowledge implies that additional classical information can change that density matrix. This is so, even when no measurement is performed and the quantum system remains untouched. Two typical cases of this are: 1) when **A** reveals to **B** information about the the outcome of her coin toss, or 2) when **A** and **B** share quantum entanglement (for example, Einstein–Podolsky–Rosen, or EPR, particles), and **A** sends the results of some measurements she performs on her system to **B**. Observe that, consequently, a density matrix is subjective in the sense that it depends on what **B** knows.

Example (Continued):

1) Suppose that, after **A** has sent an equal mixture of $|\mathsf{H}\rangle$ and $|\mathsf{V}\rangle$, she reveals to **B** that for that particular sample she prepared $|\mathsf{V}\rangle$. Then **B**'s density matrix changes, as far as he is concerned, from
$$\begin{pmatrix}1/2&0\\0&1/2\end{pmatrix}\quad\text{to}\quad\begin{pmatrix}0&0\\0&1\end{pmatrix}. \tag{8}$$
2) An identical change happens in the following situation: **A** prepares two EPR-correlated photons in a combined pure state
$$|\Psi^-\rangle=\frac{1}{\sqrt{2}}\left(|\mathsf{H}\rangle|\mathsf{V}\rangle-|\mathsf{V}\rangle|\mathsf{H}\rangle\right) \tag{9}$$
known as the singlet state. Following that, she sends one of the photons to **B**. As far as **B** is concerned, his photon's polarization will be described by the completely mixed state. On the other hand, if **A** and **B** measure both photons with respect to the same polarization (vertical, eliptical, etc.), we can predict from the overall state that their measurement outcomes will be anticorrelated. So if, upon making a measurement, **A** finds that her particle is horizontally polarized (i.e., $|\mathsf{H}\rangle$) and she tells this to **B**, then **B**'s density matrix will change according to (8).

As an aside, it is worthwhile to note that physicists sometimes disagree about whether the density matrix should be regarded as *the* state of a system or not. This, to some extent, can depend upon one's interpretation of quantum mechanics. Consider, for instance, the situation where **B** has not yet received the additional classical information to be sent by **A**. What is the state of his system? A pragmatist might answer that the state is simply described by **B**'s density matrix. Whereas a realist might argue that the state is really something different, namely, one of the pure states that go together to form that density matrix: **B** is merely ignorant of the "actual" state. For discussion of this topic we refer the reader to [7] and [11]. Here we leave this deep question unanswered and adhere to the pragmatic approach, which, in any case, is more relevant from an information-theoretical point of view.

[2] In general, we shall be fairly lax about the designations "matrix" and "operator," interchanging the two rather freely. This should cause no trouble as long as one keeps in mind that all operators discussed in this paper are linear.

Now let us describe how to compute the probability of a certain measurement result from the density matrix. Mathematically speaking, a density matrix ρ can be regarded as an object to which we can apply another operator E_x to obtain a probability. In particular, taking the trace of the product of the two matrices yields the probability that the measurement result is x given that the state was ρ, i.e., $\mathrm{Prob}\,[\text{result} = x|\text{state} = \rho] = \mathrm{Tr}\,(\rho E_x)$. Here the x serves as a label, connecting the operator E_x and the outcome x, but otherwise has no specific physical meaning. (This formula may help the reader understand the designation "density operator": it is required in order to obtain a probability density function for the possible measurement outcomes.)

Most generally, a quantum-mechanical *measurement* is described formally by a collection (ordered set) of operators, one for each outcome of the measurement.

Definition 2. (See [10]): Let $\mathcal{E} = \langle E_1, \cdots, E_m \rangle$ be a collection (ordered set) of operators such that 1) all the E_x are positive semi-definite operators, and 2) $\Sigma_x\, E_x = \mathrm{Id}$, where Id is the identity operator. Such a collection specifies a *Positive Operator-Valued Measure* (POVM) and corresponds to the most general type of measurement that can be performed on a quantum system.

Applying a POVM to a system whose state is described by a density matrix ρ results in a probability distribution according to

$$\mathrm{Prob}\,[\text{result} = x\,|\,\text{state} = \rho] = \mathrm{Tr}\,(\rho E_x) \tag{10}$$

where x ranges from 1 to m.

As an alternative for the designation POVM, one sometimes sees the term "Probability Operator Measure" used in the literature. It is a postulate of quantum mechanics that any physically realizable measurement can be described by a POVM. Moreover, for every POVM, there is *in principle* a physical procedure with which to carry out the associated measurement. Therefore, we can denote the set of all possible measurements, or equivalently the set of all POVM's, as $\mathcal{M}$.

Warning: It should be noted that the scheme of measurements defined here is the most general that can be contemplated within quantum mechanics. This is a convention that has gained wide usage within the physics community only relatively recently (within the last 15 years or so). Indeed, almost all older textbooks on quantum mechanics describe a more restrictive notion of measurement. In the usual approach, as developed by von Neumann, measurements are taken to be in one-to-one correspondence with the set of all Hermitian operators on the given Hilbert space. The eigenvalues of these operators correspond to the possible measurement results. The framework of POVM's described above can be fit within the older von Neumann picture if one is willing to take into account a more detailed picture of the measurement process, including all ancillary devices used along the way. The ultimate equivalence of these two pictures is captured by a formal result known as Neumark's Theorem [10].

A Projection Valued Measurement (PVM)—another name for the von Neumann measurements just described—is a special case of a POVM: it is given by adding the requirement that $E_x E_y = \delta(x,y) E_x$ (with $\delta(x,y) = 1$ if $x = y$ and 0 otherwise—i.e., the Kronecker-delta). With this requirement, the operators E_x are necessarily projection operators, and so can be thought of as the eigenprojectors of an Hermitian operator. One consequence of this is that the number of outcomes in a PVM can never exceed the dimensionality of the Hilbert space. General POVM's need not be restricted in this way at all; moreover, the E_x need not even commute.

Example: Measuring whether a photon is polarized according to angle α or to $\alpha + \pi/2$ is done by the POVM

$$\left\{ \begin{pmatrix} c^2 & cs \\ cs & s^2 \end{pmatrix}, \begin{pmatrix} s^2 & -cs \\ -cs & c^2 \end{pmatrix} \right\} \tag{11}$$

where $c = \cos\alpha$ and $s = \sin\alpha$. This is a PVM. When applied to a photon known to be in state $|\mathsf{H}\rangle$, for instance, this results in the probability distribution $\langle c^2, s^2 \rangle$, using (10).

An example of a POVM which is not a PVM is the symmetric three-outcome "trine" POVM: let $\gamma = \cos(\pi/3)$ and $\sigma = \sin(\pi/3)$

$$\left\{ \frac{2}{3}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \frac{2}{3}\begin{pmatrix} \gamma^2 & \gamma\sigma \\ \gamma\sigma & \sigma^2 \end{pmatrix}, \frac{2}{3}\begin{pmatrix} \gamma^2 & -\gamma\sigma \\ -\gamma\sigma & \sigma^2 \end{pmatrix} \right\} \tag{12}$$

which simplifies to

$$\left\{ \begin{pmatrix} \frac{2}{3} & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{6} & \frac{1}{6}\sqrt{3} \\ \frac{1}{6}\sqrt{3} & \frac{1}{2} \end{pmatrix}, \begin{pmatrix} \frac{1}{6} & -\frac{1}{6}\sqrt{3} \\ -\frac{1}{6}\sqrt{3} & \frac{1}{2} \end{pmatrix} \right\}. \tag{13}$$

Applying this POVM to the state $|\mathsf{V}\rangle$ results in the probability distribution $\langle 0, \frac{1}{2}, \frac{1}{2} \rangle$, again according to (10).

There are two advantages to using the formalism of POVM's over that of PVM's. First, it provides a compact formalism for describing measurements that the PVM formalism has to stretch to obtain—by considering ancillary systems, extra time evolutions, etc., in the measurement process. Secondly, and most importantly, there are some situations that call for all these extra steps to obtain an optimal measurement. A simple example is that of having to distinguish between three possible states for a system with a two-dimensional Hilbert space: the optimal POVM will generally have three outcomes, whereas a direct von Neumann measurement on the system can only have two.

IV. Measures of Distinguishability

We have just seen that a measurement (a POVM) applied to a density matrix results in a probability distribution. Suppose now we have two density matrices defined over the same Hilbert space. Then we find ourselves back in the (classical) situation described in the previous section: comparing two probability distributions over the same outcome space $\mathcal{X}$. In particular, let ρ_0 and ρ_1 be two density matrices, and let $\mathcal{E} = \{E_1, \cdots, E_m\}$ denote a POVM. Let $p_0(\mathcal{E})$ denote the probability distribution obtained by performing the POVM $\mathcal{E}$ on a system in state ρ_0 according to (10); let $p_1(\mathcal{E})$ be defined likewise. Then we have

		$x=1$	$x=2$	$\cdots$	$x=m$
$p_0(\mathcal{E})$	π_0	$\mathrm{Tr}\,(\rho_0 E_1)$	$\mathrm{Tr}\,(\rho_0 E_2)$	$\cdots$	$\mathrm{Tr}\,(\rho_0 E_m)$
$p_1(\mathcal{E})$	π_1	$\mathrm{Tr}\,(\rho_1 E_1)$	$\mathrm{Tr}\,(\rho_1 E_2)$	$\cdots$	$\mathrm{Tr}\,(\rho_1 E_m)$

As before, π_0 and π_1 denote the *a priori* probabilities and are assumed to be equal to $\frac{1}{2}$.

This section discusses four notions of distinguishability for probability distributions and—by way of the connection above—also density matrices. The unique feature in the quantum case is given by the observer's freedom to choose the measurement. Since, of course, one would like to choose the quantum measurement to be as useful as possible, one should optimize each distinguishability measure over all measurements: the values singled out by this process give rise to what we call the quantum distinguishability measures.

The reader should note that being able to distinguish between probability distributions—that is, between alternative statistical hypotheses—is already an important and well-studied problem with a vast literature. It goes under the name of statistical classification, discrimination, or feature evaluation, and has had applications as far-flung as speech recognition and radar detection. For a general overview, consult [12]. The problem studied here is a special case of the general one, in the sense that we want to distinguish between two (and only two) discrete probability distributions with equal *a priori* probabilities.

In the following subsections each classical measure of distinguishability is discussed first, followed by a discussion of its quantum counterpart.

A. Probability of Error

Consider the following experimental situation where **B** is asked to distinguish between two stochastic variables. **A** provides him with one sample, x, with equal probability to have been secretly chosen from either X_0 or X_1. **B**'s task is to guess which of the two stochastic variables the sample came from, X_0 or X_1. Clearly, the average probability that **B** makes the right guess serves as a measure of distinguishability between the two probability distributions $p_0(x)$ and $p_1(x)$.

It is well known that **B**'s optimal strategy is to look at the *a posteriori* probabilities: given the sample x, his best choice is the t for which $r_t(x)$ is maximal (see the representation at the end of Section II). This strategy is known as *Bayes' strategy*. So the average probability of successfully identifying the distribution equals

$$\sum_{x\in X} p(x)\max\{r_0(x), r_1(x)\} = \frac{1}{2}\sum_{x\in X}\max\{p_0(x), p_1(x)\}. \tag{14}$$

Conversely, we can also express the probability that **B** fails.

Definition 3: The *probability of error between two probability distributions* is defined by

$$\mathsf{PE}(p_0, p_1) \stackrel{\text{def}}{=} \frac{1}{2}\sum_{x\in X}\min\{p_0(x), p_1(x)\}. \tag{15}$$

Two identical distributions have $\mathsf{PE} = \frac{1}{2}$, and two orthogonal distributions have $\mathsf{PE} = 0$.

Warning: PE is not a distance function: for example, when two distributions are close to one another, PE is *not* close to 0, but close to $\frac{1}{2}$.

In the quantum-mechanical case, the experimental setup is almost identical. **A** has two ensembles, one according to ρ_0, the other according to ρ_1. She provides **B** with a quantum sample chosen from one of the two ensembles with equal probability. Following a measurement, **B** must again guess from which ensemble the sample was drawn: the one under ρ_0 or the one under ρ_1.

For any fixed measurement, the Bayesian strategy of guessing the density operator with the largest posterior probability is the optimal thing to do. However, now **B** should as well make use of his extra degree of freedom: he can choose the measurement he applies to his sample. He should choose the measurement that minimizes his probability of error. So we define

Definition 4: The *probability of error between two density matrices* ρ_0 and ρ_1 is defined by

$$\mathsf{PE}(\rho_0, \rho_1) \stackrel{\text{def}}{=} \min_{\mathcal{E}\in\mathcal{M}} \mathsf{PE}(p_0(\mathcal{E}), p_1(\mathcal{E})) \tag{16}$$

where the POVM $\mathcal{E}$ ranges over the set of all possible measurements $\mathcal{M}$.

(More carefully, one should use "infimum" in this definition. However—since in all the optimization problems we shall consider here, the optima actually can be obtained—there is no need for the extra rigor.)

The question of finding an explicit formula for the optimal POVM in this definition was first studied by Helstrom [13, pp. 106–108]. He shows that the POVM $\mathcal{E}^*$ that *minimizes* $\mathsf{PE}(p_0(\mathcal{E}), p_1(\mathcal{E}))$ is actually a PVM. Knowing the optimal POVM, the probability of error can be expressed explicitly. The expression he gives is,

$$\mathsf{PE}(\rho_0, \rho_1) = \frac{1}{2} + \frac{1}{2}\sum_{\lambda_j\le 0}\lambda_j \tag{17}$$

where the λ_j denote the eigenvalues of the matrix $\Gamma = \rho_0 - \rho_1$.

This expression can be cleaned up a little in the following way. Consider the function $f(x) = \frac{1}{2}(x - |x|)$. It vanishes when $x \ge 0$ and is the identity function otherwise. Thus with its use, we can expand the summation in (17) to be over all the eigenvalues of Γ

$$\mathsf{PE}(\rho_0, \rho_1) = \frac{1}{2} + \frac{1}{2}\sum_{j=1}^{N} f(\lambda_j) \tag{18}$$

$$= \frac{1}{2} + \frac{1}{4}\operatorname{Tr}\Gamma - \frac{1}{4}\sum_{j=1}^{N}|\lambda_j| \tag{19}$$

$$= \frac{1}{2} - \frac{1}{4}\operatorname{Tr}|\Gamma|. \tag{20}$$

Hence we have the following proposition.

Proposition 1: Given two arbitrary density matrices ρ_0 and ρ_1, the probability of error equals

$$\mathsf{PE}(\rho_0, \rho_1) = \frac{1}{2} - \frac{1}{4}\sum_{j=1}^{N}|\lambda_j| = \frac{1}{2} - \frac{1}{4}\operatorname{Tr}|\rho_0 - \rho_1| \tag{21}$$

where the λ_j are the eigenvalues of $\rho_0 - \rho_1$.

$\mathsf{PE}(\rho_0, \rho_1)$ is, therefore, just a simple function of the distance between ρ_0 and ρ_1, when measured as the trace norm

of their difference. (An alternative derivation of this can be found in [14].)

B. Kolmogorov Distance

Among (computational) cryptographers, another measure of distinguishability between probability distributions is used fairly often: the standard notions of exponential and computational indistinguishability [15], [16], [2] are based on it.

Definition 5: The *Kolmogorov distance between two probability distributions* is defined by

$$\mathsf{K}(p_0,p_1) \stackrel{\text{def}}{=} \frac{1}{2}\sum_{x\in X}|p_0(x)-p_1(x)|. \tag{22}$$

Two identical distributions have $\mathsf{K}=0$, and two orthogonal distributions have $\mathsf{K}=1$.

In some references the factor of $\frac{1}{2}$ plays no role, and the "Kolmogorov distance" is defined without it. Here we have included it because we wanted K to take values between 0 and 1.

Probability of error and Kolmogorov distance are closely related.

Proposition 2:

$$\mathsf{PE}(p_0,p_1)=\frac{1}{2}-\frac{1}{2}\mathsf{K}(p_0,p_1). \tag{23}$$

This is not very difficult to prove. The most important step is to split the sum over $\mathcal{X}$ into two disjoint subsums, one for which $p_0(x)<p_1(x)$, and one for which $p_0(x)\geq p_1(x)$. See [17].

In the quantum case, we must again optimize over all possible measurements. But here this means that we have to find the POVM that maximizes the Kolmogorov distance.

Definition 6: The *Kolmogorov distance between two density matrices* ρ_0 and ρ_1 is defined by

$$\mathsf{K}(\rho_0,\rho_1) \stackrel{\text{def}}{=} \max_{\mathcal{E}\in\mathcal{M}} \mathsf{K}(p_0(\mathcal{E}),p_1(\mathcal{E})) \tag{24}$$

where the POVM $\mathcal{E}$ ranges over the set of all possible measurements $\mathcal{M}$.

The relation between probability of error and Kolmogorov distance (23) shows that the two measurements that optimize PE and K are identical: $\mathcal{E}^*$ minimizes the function $\mathsf{PE}(p_0(\mathcal{E}),p_1(\mathcal{E}))$ if and only if it also maximizes $\mathsf{K}(p_0(\mathcal{E}),p_1(\mathcal{E}))$. See also [14, Appendix]. Combining (21) and (23) we get

Proposition 3: The Kolmogorov distance between two density matrices ρ_0 and ρ_1 equals

$$\mathsf{K}(\rho_0,\rho_1)=\frac{1}{2}\sum_{j=1}^{N}|\lambda_j|=\frac{1}{2}\operatorname{Tr}|\rho_0-\rho_1| \tag{25}$$

where the λ_j are the eigenvalues of $\rho_0-\rho_1$.

In the special case that ρ_0 and ρ_1 diagonalize in the same basis we are essentially back to the classical case. The two probability distributions can be read off from the diagonals, and (25) trivially reduces to (22).

Observe that $\operatorname{Tr}|\rho_0-\rho_1|$ is simply the *trace-norm distance* on operators [18], [19]. Hence K has the additional property of satisfying a triangle inequality. The trace-norm distance appears to be of unique significance within the class of all operator-norms because of its connection to probability of error.

C. Bhattacharyya Coefficient

Another distinguishability measure that has met widespread use—mostly because it is sometimes easier to evaluate than the others—is the Bhattacharyya coefficient. See [20], [21], and [17].

Definition 7: The *Bhattacharyya coefficient between two probability distributions* p_0 and p_1 is defined by

$$\mathsf{B}(p_0,p_1)=\sum_{x\in X}\sqrt{p_0(x)p_1(x)}. \tag{26}$$

Two identical distributions have $\mathsf{B}=1$, and two orthogonal distributions have $\mathsf{B}=0$.

Warning: B is also not a distance function: for instance, when two distributions are close to one another, B is *not* close to 0. It can, however, be easily related to a distance function by taking its arccosine.

The Bhattacharyya coefficient's greatest appeal is its simplicity: it is a sort of overlap measure between the two distributions. When their overlap is zero, they are completely distinguishable; when their overlap is one, the distributions are identical and hence indistinguishable. Moreover, the Bhattcharyya coefficient can be thought of geometrically as an inner product between p_0 and p_1, interpreted as vectors in an m-dimensional vector space. However, it does not appear to bear a simple relation to the probability of error in any type of statistical inference problem.

In the quantum case, we define a distinguishability measure by minimizing over all possible measurements.

Definition 8: The *Bhattacharyya coefficient between two density matrices* ρ_0 and ρ_1 is defined by

$$\mathsf{B}(\rho_0,\rho_1) \stackrel{\text{def}}{=} \min_{\mathcal{E}\in\mathcal{M}} \mathsf{B}(p_0(\mathcal{E}),p_1(\mathcal{E})) \tag{27}$$

where the POVM $\mathcal{E}$ ranges over the set of all possible measurements $\mathcal{M}$.

The following proposition provides a closed-form expression for this distinguishability measure.

Proposition 4. (Fuchs and Caves [22]): The quantum Bhattacharyya coefficient can be expressed as

$$\mathsf{B}(\rho_0,\rho_1)=\operatorname{Tr}\left(\sqrt{\sqrt{\rho_0}\rho_1\sqrt{\rho_0}}\right) \tag{28}$$

where the square root of a matrix ρ denotes any positive semidefinite matrix σ such that $\sigma^2=\rho$.

When ρ_0 and ρ_1 diagonalize in the same basis we are back to the classical case, and (28) reduces to (26), because the equality $\sqrt{\rho_0}\rho_1\sqrt{\rho_0}=\rho_0\rho_1$ now holds.

Surprisingly, it turns out that B is equivalent to another *nonmeasurement oriented* notion of distinguishability. Suppose $|\psi_0\rangle$ and $|\psi_1\rangle$ are pure states. When we think these two state vectors geometrically, a natural notion of distinguishability is the angle between $|\psi_0\rangle$ and $|\psi_1\rangle$, or any simple function of this angle like the inner product or overlap. In particular, we can define $\text{overlap}\,(|\psi_0\rangle,|\psi_1\rangle) := |\langle\psi_0|\psi_1\rangle|$ as a measure of distinguishability. The question is: what to do for mixed states?

The answer was given by Uhlmann [23].[3] If ρ_0 is the density matrix of a mixed state in the Hilbert space $\mathcal{H}_1$, then we can always extend the Hilbert space such that ρ_0 becomes a pure state in the combined Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$. More precisely, we can always find an extension $\mathcal{H}_2$ of $\mathcal{H}_1$ and a pure state $|\psi_0\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$, such that $\text{Tr}_2(|\psi_0\rangle\langle\psi_0|) = \rho_0$. Here the operator Tr_2 means to perform a partial-trace operation over the ancillary Hilbert space $\mathcal{H}_2$. When this condition holds, $|\psi_0\rangle$ is said to be a *purification* of ρ_0. Similarly, if $|\psi_1\rangle$ is the purification of ρ_1, we are back to a situation with two pure states, and we can apply the formula above, leading to the following generalized definition.

Definition 9: The *(generalized) overlap* between two density matrices is defined by

$$\text{overlap}\,(\rho_0,\rho_1) \stackrel{\text{def}}{=} \max|\langle\varphi_0|\varphi_1\rangle| \tag{29}$$

where the maximum is taken over all purifications $|\varphi_0\rangle$ and $|\varphi_1\rangle$ of ρ_0 and ρ_1, respectively.

It can be demonstrated that [22]

$$\text{overlap}\,(\rho_0,\rho_1) = \mathtt{B}\,(\rho_0,\rho_1). \tag{30}$$

Despite the rather Baroque appearance $\mathtt{B}\,(\rho_0,\rho_1)$ takes in (28), it is endowed with several very nice properties. For instance, $\mathtt{B}\,(\rho_0,\rho_1)$ is multiplicative over tensor products

$$\mathtt{B}\,(\rho_0\otimes\rho_1,\rho_2\otimes\rho_3) = \mathtt{B}\,(\rho_0,\rho_2)\,\mathtt{B}\,(\rho_1,\rho_3). \tag{31}$$

B's square is concave over one of its arguments; i.e., if $0 \le \mu_0,\mu_1 \le 1, \mu_0+\mu_1 = 1$ then

$$(\mathtt{B}\,(\rho,\mu_0\rho_0+\mu_1\rho_1))^2 \ge \mu_0(\mathtt{B}\,(\rho,\rho_0))^2+\mu_1(\mathtt{B}\,(\rho,\rho_1))^2. \tag{32}$$

Moreover, B itself is doubly concave[4]

$$\mathtt{B}(\mu_0\rho_0+\mu_1\rho_1,\mu_0\rho_2+\mu_1\rho_3) \ge \mu_0\mathtt{B}(\rho_0,\rho_2)+\mu_1\mathtt{B}(\rho_1,\rho_3). \tag{33}$$

D. Shannon Distinguishability

Now we come to the last, and maybe most important, notion of distinguishability. Mutual information, as defined by Shannon [25], can be used as a distinguishability measure between probability distributions [26], [12]. We assume that the reader is familiar with the (Shannon) entropy function H, the argument of which can be either a stochastical variable or a probability distribution. $H_2(p) = -p\log p-(1-p)\log(1-p)$ is the entropy of the distribution $\langle p, 1-p\rangle$.

[3] A nice review of this theorem in terms of finite-dimensional Hilbert space methods can be found in [24].

[4] The authors thank C. M. Caves for pointing this out to them.

Consider the following elementary example. Suppose we have two boxes, each containing colored balls. Let $t \in \mathcal{T} = \{0,1\}$ denote the identity of the boxes; and let us think of T as a stochastic variable. Then $\text{Prob}\,[T=t]$ is just the *a priori* probability π_t of Section II. Recall that in our case $\pi_0 = \pi_1 = \frac{1}{2}$, so $H(T) = 1$. Let X denote the stochastic variable corresponding to the color of a ball upon being drawn from a box, taking into account that the identity of the box is itself a stochastic variable. Recall that $\text{Prob}\,[X=x]$ was written as $p(x)$.

Consider the same experiment as in Section IV-A, in which **A** picks a ball from one of the two boxes and gives it to B. One can ask now: How much information does X (the color of a picked ball) convey about T (the identity of the box it came from)?

Information is defined as the reduction of uncertainty, where uncertainty is quantified using the Shannon entropy. Consider two quantities: 1) the average uncertainty of B about T before he was handed a sample (or ball), denoted H; and 2) his average uncertainty about T after he was handed a sample, denoted $H(T|X)$. This difference expresses the amount of information gained through the experiment, and can also be used as a measure of distinguishability between two distributions. Thus we obtain

$$\begin{aligned}
&\text{average information}\\
&\quad = H(T) - H(T|X) && (34)\\
&\quad = H_2\left(\tfrac{1}{2}\right) - \sum_{x\in\mathcal{X}} p(x)\,H(T|X=x) && (35)\\
&\quad = 1 - \sum_{x\in\mathcal{X}} p(x)\,H(r_0(x)) && (36)\\
&\quad = I(T;X). && (37)
\end{aligned}$$

This leads to the following definition.

Definition 10: The *Shannon distinguishability between two probability distributions* p_0 and p_1 is defined by

$$\mathtt{SD}\,(p_0,p_1) \stackrel{\text{def}}{=} I(T;X). \tag{38}$$

In the same fashion as all the other distinguishability measures, the Shannon distinguishability can be applied to the quantum case. We must find the measurement that optimizes it when tabulated for probability distributions obtained by applying a quantum measurement.

Definition 11: The *Shannon distinguishability between two density matrices* ρ_0 and ρ_1 is defined as

$$\mathtt{SD}\,(\rho_0,\rho_1) \stackrel{\text{def}}{=} \max_{\mathcal{E}\in\mathcal{M}} \mathtt{SD}\,(p_0(\mathcal{E}),p_1(\mathcal{E})) \tag{39}$$

where the POVM $\mathcal{E}$ ranges over the set of all possible measurements $\mathcal{M}$.

There is an unfortunate problem for this measure of distinguishabity: calculating the value $\mathtt{SD},(\rho_0,\rho_1)$ is generally a difficult problem. Apart from a few special cases, no explicit formula for SD solely in terms of ρ_0 and ρ_1 is known. Even stronger than that: no such formula can exist in the general case [27]. This follows from the fact that optimizing the Shannon

TABLE I
OVERVIEW OF THE FOUR DISTINGUISHABILITY MEASURES DISCUSSED IN THIS PAPER

	classical definition	when $p_0=p_1$	when $p_0 \perp p_1$	optimality criterion	quantum expression
PE	$\frac{1}{2}\sum \min\{p_0(x), p_1(x)\}$	1/2	0	min	$\frac{1}{2} - \frac{1}{4}\mathrm{Tr}\lvert\rho_0 - \rho_1\rvert$
K	$\frac{1}{2}\sum \lvert p_0(x) - p_1(x)\rvert$	0	1	max	$\frac{1}{2}\mathrm{Tr}\lvert\rho_0 - \rho_1\rvert$
B	$\sum \sqrt{p_0(x)p_1(x)}$	1	0	min	$\mathrm{Tr}\sqrt{\rho_0^{1/2}\rho_1\rho_0^{1/2}}$
SD	*via* $I(X;T)$	0	1	max	no simple form

distinguishability requires the solution of a transcendental equation. (See also [10] and [28] for a discussion of other aspects of SD.)

E. Overview

The material presented in the previous four subsections is summarized in Table I.

V. INEQUALITIES

We have seen already (23) that probability of error and Kolmogorov distance are related through the equality:

$$\mathsf{PE}(p_0,p_1) = \frac{1}{2} - \frac{1}{2}\mathsf{K}(p_0,p_1). \tag{40}$$

The other pairs of distinguishability measures are related through inequalities, some of which can be found in the literature [20], [29], [21], [17], [12].

Proposition 5: Let p_0 and p_1 be probability distributions. The following relations hold:

$$1-\mathsf{B}(p_0,p_1) \le \mathsf{K}(p_0,p_1) \le \sqrt{1-\mathsf{B}(p_0,p_1)^2} \tag{41}$$

$$1-H_2(\mathsf{PE}(p_0,p_1)) \le \mathsf{SD}(p_0,p_1) \le 1-2\,\mathsf{PE}(p_0,p_1) \tag{42}$$

$$1-\mathsf{B}(p_0,p_1) \le \mathsf{SD}(p_0,p_1) \le 1-H_2\left(\frac{1}{2}-\frac{1}{2}\sqrt{1-\mathsf{B}(p_0,p_1)}\right). \tag{43}$$

Before giving the proof of this proposition, we state its quantum equivalent. This is the main result of the paper.

Theorem 1: Proposition 5 can be generalized to the quantum scenario: one can substitute PE, K, B, and SD and use density matrices ρ_0 and ρ_1 as operands. Alternatively, using the quantum expressions, (41)–(43) can be expressed in the following, equivalent form:

$$1-\mathsf{B}(\rho_0,\rho_1) \le \frac{1}{2}\mathrm{Tr}\lvert\rho_0-\rho_1\rvert \le \sqrt{1-\mathsf{B}(\rho_0,\rho_1)^2} \tag{44}$$

$$1-H_2\left(\frac{1}{2}-\frac{1}{4}\mathrm{Tr}\lvert\rho_0-\rho_1\rvert\right) \le \mathsf{SD}(\rho_0,\rho_1) \le \frac{1}{2}\mathrm{Tr}\lvert\rho_0-\rho_1\rvert \tag{45}$$

$$1-\mathsf{B}(\rho_0,\rho_1) \le \mathsf{SD}(\rho_0,\rho_1) \le 1-H_2\left(\frac{1}{2}-\frac{1}{2}\sqrt{1-\mathsf{B}(\rho_0,\rho_1)}\right). \tag{46}$$

The importance of this theorem is that, while the quantum Shannon distinguishability is impossible to calculate in a closed form, the inequalities provide a useful way to bound it. We will use these bounds in an application in the next section.

Proof of Proposition 5: We start by proving (41). To get the left-hand inequality, note

$$1-\mathsf{B}(p_0,p_1) = \frac{1}{2}\left(\sum_{x\in\mathcal{X}} p_0(x) + \sum_{x\in\mathcal{X}} p_1(x) - 2\sum_{x\in\mathcal{X}}\sqrt{p_0(x)p_1(x)}\right) \tag{47}$$

$$= \frac{1}{2}\sum_{x\in\mathcal{X}} \lvert\sqrt{p_0(x)} - \sqrt{p_1(x)}\rvert^2 \tag{48}$$

$$\le \frac{1}{2}\sum_{x\in\mathcal{X}} \lvert p_0(x) - p_1(x)\rvert \tag{49}$$

$$= \mathsf{K}(p_0,p_1). \tag{50}$$

The inequality in the penultimate step holds for each term individually. To get the right-hand inequality, we simply use the Schwarz inequality

$$\mathsf{K}(p_0,p_1)^2 = \frac{1}{4}\left(\sum_{x\in\mathcal{X}} \lvert p_0(x) - p_1(x)\rvert\right)^2 \tag{51}$$

$$= \frac{1}{4}\left(\sum_{x\in\mathcal{X}} \lvert\sqrt{p_0(x)} - \sqrt{p_1(x)}\rvert \cdot \lvert\sqrt{p_0(x)} + \sqrt{p_1(x)}\rvert\right)^2 \tag{52}$$

$$\le \frac{1}{4}\sum_{x\in\mathcal{X}} (\sqrt{p_0(x)} - \sqrt{p_1(x)})^2 \cdot \sum_{x\in\mathcal{X}} (\sqrt{p_0(x)} + \sqrt{p_1(x)})^2 \tag{53}$$

$$= \frac{1}{4}(2-2\,\mathsf{B}(p_0,p_1))(2+2\,\mathsf{B}(p_0,p_1)) \tag{54}$$

$$= 1-\mathsf{B}(p_0,p_1)^2. \tag{55}$$

In order to prove the left inequality of (42), we observe that this is a special case of the Fano inequality (see, for instance, [30])

$$H(T|X) \le H_2(\mathsf{PE}(p_0,p_1)) + \mathsf{PE}(p_0,p_1)\log(\#T-1) \tag{56}$$

where $\#\mathcal{T} = 2$ is the cardinality of the set $\mathcal{T}$.

For the right-hand inequality of (42) we expand $I(X;T)$ as $H(T) - H(T|X)$ to obtain an inequality between SD and PE. (See also [29].) Recall the definitions of $r_t(x)$ and $p(x)$, observing that $r_1(x) = 1 - r_0(x)$ and that $2\min\{r, 1-r\} \le H_2(r)$ for all r between 0 and 1 (see Fig. 1). Hence, we obtain

$$\mathsf{SD}(p_0,p_1) = 1 - \sum_{x\in\mathcal{X}} p(x)H_2(r_0(x)) \tag{57}$$

$$\le 1 - \sum_{x\in\mathcal{X}} p(x)\cdot 2\min\{r_0(x), 1-r_0(x)\} \tag{58}$$

$$\le 1 - 2\mathsf{PE}(p_0,p_1). \tag{59}$$

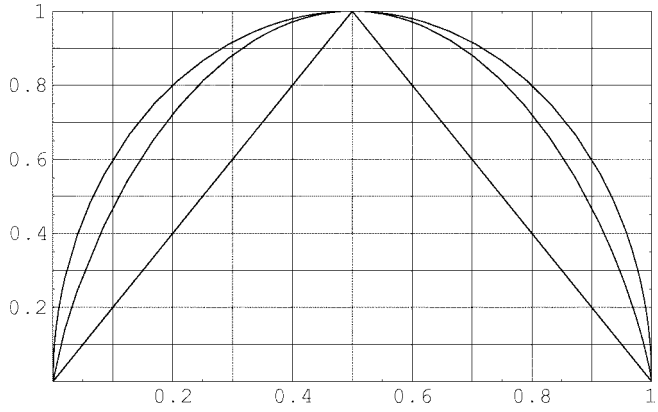

Fig. 1. $2\min\{x, 1-x\} \leq \mathsf{H}_2(x) \leq 2(x(1-x))^{1/2}$. (Formally this is proven by looking at the first and second derivatives.)

The left-hand inequality of (43) is obtained in a similar way. Using the fact that $H_2(r) \leq 2\sqrt{r(1-r)}$ (see Fig. 1), we get

$$\mathsf{SD}(p_0, p_1) = 1 - \sum_{x\in\mathcal{X}} p(x) H_2(r_0(x)) \tag{60}$$

$$\geq 1 - \sum_{x\in\mathcal{X}} p(x) \cdot 2\sqrt{r_0(x)(1-r_0(x))} \tag{61}$$

$$= 1 - \sum_{x\in\mathcal{X}} \sqrt{p_0(x)p_1(x)} \tag{62}$$

$$= 1 - \mathsf{B}(p_0, p_1). \tag{63}$$

For the right-hand side of (43) we define the function

$$g(r) := \frac{1}{2} - \frac{1}{2}\sqrt{1-r^2}. \tag{64}$$

If we let $h(r) = 2\sqrt{r(1-r)}$ for $0 \leq r \leq 1$, then we get that $h(r) = h(1-r)$, that $gh(r) = \min\{r, 1-r\}$, and that $hg(r) = r$. So

$$H_2(r_0(x)) = H_2(\min\{r_0(x), 1-r_0(x)\}) = H_2(gh(r_0(x)))$$

and

$$\mathsf{SD}(p_0, p_1) = 1 - \sum_{x\in\mathcal{X}} p(x) H_2(r_0(x)) \tag{65}$$

$$= 1 - \sum_{x\in\mathcal{X}} p(x) H_2(g(h(r_0(x)))) \tag{66}$$

$$\leq 1 - H_2\left(g\left(\sum_{x\in\mathcal{X}} p(x) h(r_0(x))\right)\right) \tag{67}$$

$$= 1 - H_2(g(\mathsf{B}(p_0, p_1))) \tag{68}$$

$$= 1 - H_2\left(\frac{1}{2} - \frac{1}{2}\sqrt{1 - (\mathsf{B}(p_0, p_1))^2}\right). \tag{69}$$

Here we used Jensen's inequality on the (composite) function $H_2(g(r))$; this function is convex. This concludes the proof of Proposition 5. □

The main tool in proving the quantum versions of these inequalities is in noting that all the bounds are appropriately monotonic in their arguments.

Proof of Theorem 1: First we prove (44). We start with the first inequality. Let $\mathcal{E}_B^*$ denote a POVM that optimizes B and define $\mathcal{E}_K^*$ likewise

$$1 - \mathsf{B}(\rho_0, \rho_1) = 1 - \mathsf{B}(p_0(\mathcal{E}_B^*), p_1(\mathcal{E}_B^*)) \tag{70}$$

$$\leq \mathsf{K}(p_0(\mathcal{E}_B^*), p_1(\mathcal{E}_B^*)) \tag{71}$$

$$\leq \mathsf{K}(p_0(\mathcal{E}_K^*), p_1(\mathcal{E}_K^*)) \tag{72}$$

$$= \mathsf{K}(\rho_0, \rho_1). \tag{73}$$

The second inequality of (44) follows from

$$\mathsf{K}(\rho_0, \rho_1) = \mathsf{K}(p_0(\mathcal{E}_K^*), p_1(\mathcal{E}_K^*)) \tag{74}$$

$$\leq \sqrt{1 - \mathsf{B}(p_0(\mathcal{E}_K^*), p_1(\mathcal{E}_K^*))^2} \tag{75}$$

$$\leq \sqrt{1 - \mathsf{B}(p_0(\mathcal{E}_B^*), p_1(\mathcal{E}_B^*))^2} \tag{76}$$

$$= \sqrt{1 - \mathsf{B}(\rho_0, \rho_1)^2}. \tag{77}$$

Equations (45) and (46) are proven in an identical way. In particular, in (44) the functions on the extreme left, $f(x) = 1-x$, and on the extreme right, $f(x) = \sqrt{1-x^2}$, are both monotonically decreasing. In addition, B must be minimized whereas K must be maximized. The same is true for (45) and (46). □

VI. Exponential Indistinguishability

As already described in Section I, in the solution of various cryptographic tasks, one often actually designs a whole *family* of protocols. These are parameterized by a *security parameter* n: a number that might denote the length of some string, the number of rounds, or the number of photons transmitted, for instance. Typically the design of a good protocol requires that the probability of cheating for each participant vanishes exponentially fast, i.e., is of the order $O(\varepsilon^{-n})$, with ε between 0 and 1. As an example, one technique is to compare the protocol implementation (the family of protocols) with the *ideal protocol specification* and to prove that these two become exponentially indistinguishable [3].

Definition 12: Let $\{X_0\} = \langle X_0^{(1)}, X_0^{(2)}, X_0^{(3)}, \cdots\rangle$ denote a *family* of stochastic variables with corresponding distributions $\langle p_0^{(1)}, p_0^{(2)}, p_0^{(3)}, \cdots\rangle$. Let $\{X_1\}$ be defined similarly. Then $\{X_0\}$ and $\{X_1\}$ are *exponentially indistinguishable* if there exists an n_0 and an ε between 0 and 1 such that

$$\forall n \geq n_0 \colon \mathsf{K}(p_0^{(n)}, p_1^{(n)}) \leq \varepsilon^n. \tag{78}$$

Examples of exponentially indistinguishable stochastic-variable families can be constructed easily. For instance, let X_0^n be uniformly distributed over $\{0,1\}^n$, the set of strings of length n. That is, for each $x \in \{0,1\}^n$, we have $p_0^{(n)}(x) = 2^{-n}$. This defines the family of uniform distributions over $\{X_0\}$. Let $\{X_1\}$ be defined identically, *except* that $p_1^{(n)}(0^n) = 0$, while $p_1^{(n)}(1^n) = 2^{-n+1}$. So for $\{X_1\}$, 0^n, the word with all zeroes, has zero probability; while 1^n, the word with all ones, has double the probability it had in the uniform distribution. Clearly, the two families $\{X_0\}$ and $\{X_1\}$ are exponentially indistinguishable.

The reader should be aware that in *computational* cryptography more refined notions of distinguishability have been

defined [2]. For *polynomial indistinguishability*, it is only required that the families converge as fast as $1/n^k$, for any $k>0$. Though we will not argue it formally here, it is not hard to see that the proof of Lemma 1 generalizes to apply to polynomially indistinguishable families.

Yet another refinement is *computational indistinguishability*. For it, a sample is given to a Turing machine which outputs a 0 or 1, and we look at the Kolmogorov distance of the possible outputs. After maximizing over all Turing machines, we say the stochastic-variable families are computationally indistinguishable if the distance between them converges as $1/n^k$ for any $k>0$. Computational indistinguishability has turned out to be extremely powerful for defining notions as pseudorandom number generators [15] and zero knowledge protocols [2]. All these notions of protocol indistinguishability have in common that if a distinguisher is given a sample and restricted to polynomial-time calculations, then he will not be able to identify the source of the sample.

Above, in Definition 12 we have followed the computational-cryptographic tradition in defining exponential indistinguishability via the Kolmogorov distance. However, this choice is in no way crucial: the next lemma shows that we could have taken any of the four distinguishability measures. In other words, K, PE, B, and SD turn out to be equivalent when we require exponentially fast convergence.[5]

Lemma 1: Let $\{X_0\}$ and $\{X_1\}$ be two families of stochastic variables that are exponentially indistinguishable with respect to *one* of the distinguishability measures K, PE, B, SD. Then $\{X_0\}$ and $\{X_1\}$ are exponentially indistinguishable with respect to *each* of K, PE, B, SD.

Proof: The equivalence between exponential indistinguishability for PE and K follows from (23). The other equivalences follow from (41)–(43). For instance, the proof that exponential indistinguishability for K implies exponential indistinguishability for B goes as follows. Suppose

$$[\exists n_0,\varepsilon][\forall n\geq n_0]\colon \mathsf{K}(p_0^{(n)},p_1^{(n)})\leq\varepsilon^n. \tag{79}$$

Using the left-hand side of (41), it follows at once that $\mathsf{B}(p_0^{(n)},p_1^{(n)})\geq 1-\varepsilon^n$. It then follows from the fact that $\mathsf{B}(p_0^{(n)},p_1^{(n)})$ is bounded above by unity, that we obtain the desired exponential convergence.

For the reverse direction: if $1-\mathsf{B}(p_0^{(n)},p_1^{(n)})\leq\varepsilon^n$ then

$$1-(\mathsf{B}(p_0^{(n)},p_1^{(n)}))^2\leq\varepsilon^n(1+\mathsf{B}(p_0^{(n)},p_1^{(n)}))\leq 2\varepsilon^n \tag{80}$$

so $\mathsf{K}(p_0^{(n)},p_1^{(n)})\leq\sqrt{2}\sqrt{\varepsilon}^n$ using the right-hand side of (41). If we choose $\overline{\varepsilon}<\sqrt{\varepsilon}$ and $\overline{n}_0>n_0$ such that $\sqrt{2}<(\overline{\varepsilon}/\sqrt{\varepsilon})^{\overline{n}_0}$, then $\mathsf{K}(p_0^{(n)},p_1^{(n)})\leq\overline{\varepsilon}^n$ for $n\geq\overline{n}$.

The other implications are proven in a similar way. As far as expressions involving $H_2(x)$ are concerned, it is sufficient to recall (see Fig. 1) that

$$2\min\{x,1-x\}\leq H_2(x)\leq 2\sqrt{x(1-x)}. \tag{81}$$

This concludes the proof. □

[5]There is a small technicality here: indistinguishable distributions have PE $=\frac{1}{2}$ and B $=1$, so exponential indistinguishability means convergence to those values, instead of convergence to 0, as is the case with K and SD.

The obvious next step is to define exponential indistinguishability for density matrices, and to show that the choice of the distinguishability measure is immaterial.

Definition 13: Let $\{\rho_0^{(n)}\}=\langle\rho_0^{(1)},\rho_0^{(2)},\rho_0^{(3)},\cdots\rangle$ denote a *family* of density matrices defined over the Hilbert spaces $\mathcal{H}^{(1)},\mathcal{H}^{(2)},\mathcal{H}^{(3)}\cdots$, Let $\{\rho_1^{(n)}\}$ be defined similarly. Then $\{\rho_0^{(n)}\}$ and $\{\rho_1^{(n)}\}$ are *exponentially indistinguishable* if there exists an n_0 and an ε between 0 and 1 such that

$$\forall n\geq n_0\colon \mathsf{K}(\rho_0^{(n)},\rho_1^{(n)})\leq\varepsilon^n. \tag{82}$$

An example that makes use of this definition will be presented in the next section. However, first let us conclude with the quantum analog of Lemma 1.

Theorem 2: Let $\{\rho_0^{(n)}\}$ and $\{\rho_1^{(n)}\}$ be two families of density matrices which are exponentially indistinguishable with respect to *one* of the distinguishability measures K, PE, B, SD. Then $\{\rho_0^{(n)}\}$ and $\{\rho_1^{(n)}\}$ are exponentially indistinguishable with respect to *each* of K, PE, B, SD.

Proof: This follows now immediately from the proof of Lemma 1 using Theorem 1. □

VII. Applications

Let us now look at an application of the quantum-exponential indistinguishability idea. In particular, we look at the problem of the parity bit in quantum key distribution as studied in [5]. Let $|\psi_0\rangle=\binom{\cos\alpha}{\sin\alpha}$ and $|\psi_1\rangle=\binom{\cos\alpha}{-\sin\alpha}$, and let ρ_0 and ρ_1 be the corresponding density matrices. That is, the bits 0 and 1 that contribute to constructing a cryptographic key are encoded into a physical system—a photon, say—via ρ_0 and ρ_1. Likewise, the bit string $z=z_1z_2\cdots z_n$ is represented by n different photons, the ith photon being in state ρ_i. Thus the combined state for the string z is given by

$$\rho_z=\rho_{z_1}\otimes\rho_{z_2}\otimes\cdots\otimes\rho_{z_n} \tag{83}$$

where $\otimes$ stands for the tensor product.

Now let $Z_0^{(n)}$ denote all the strings of length n with even parity (i.e., with overall exclusive–or equal to 0) and $Z_1^{(n)}$ all strings of length n with odd parity. Then define

$$\rho_j^{(n)}=\frac{1}{2^{n-1}}\sum_{z\in Z_j^{(n)}}\rho_z \tag{84}$$

for $j=0,1$. In [5] these two density matrices are explicitly calculated in order to compute their Shannon distinguishability as a function of n and α. This is extremely important because the parity bit appears in the proof of security [1] of the BB84 key exchange protocol [31].

Here we compute the distinguishability between $\rho_0^{(n)}$ and $\rho_1^{(n)}$ in terms of Kolmogorov distance and Bhattacharyya coefficient. For the special case $n=2$ we also study the inequalities obtained in Theorem 5, as an illustration of how tight the bounds are. Observe that, at this point in time, the problem of the parity bit is one of the few nontrivial (i.e., multidimensional Hilbert-space) examples for which the Shannon distinguishability, Kolmogorov distance (and related

probability of error), and Bhattacharyya coeffecient can be computed. For the next few paragraphs the reader is advised to consult [5], or to take (85), (97), and (98) below as given.

First let us calculate the Kolmogorov distance $\mathsf{K}(\rho_0^{(n)}, \rho_1^{(n)})$ as a function of n and α. In [5] it is shown that

$$\Delta^{(n)} = \frac{1}{2}\left(\rho_0^{(n)} - \rho_1^{(n)}\right)$$

has nonzero entries only on the secondary diagonal. Moreover, it is not difficult to see that all these entries equal $c^n s^n$, where $c = \cos\alpha$, $s = \sin\alpha$. Therefore, $\Delta^{(n)}$ has 2^{n-1} eigenvalues equal to $-c^n s^n$, and 2^{n-1} eigenvalues equal to $+c^n s^n$, so

$$\mathsf{K}(\rho_0^{(n)}, \rho_1^{(n)}) = \sum |\lambda_j| = |2cs|^n = |\sin 2\alpha|^n = |S|^n \quad (85)$$

where $S = \sin 2\alpha$. Clearly, $\{\rho_0^{(n)}\}$ and $\{\rho_1^{(n)}\}$ are exponentially indistinguishable for all values of $\alpha \neq \pi/2$. (Note that in [5] exponential indistinguishability is proven only for the case that $\alpha \approx 0$.)

Computing the Bhattacharrya coefficient between $\rho_0^{(n)}$ and $\rho_1^{(n)}$ is a more elaborate calculation. In [5, eqs. (19) and (20)] it is shown that, with a minor change of basis, $\rho_0^{(n)}$ and $\rho_1^{(n)}$ can be taken to be block-diagonal with 2×2 blocks. These blocks are of the form

$$\sigma_0^{(n,k)} = \begin{pmatrix} c^{2(n-k)}s^{2k} & c^n s^n \\ c^n s^n & c^{2k}s^{2(n-k)} \end{pmatrix} \text{ for even parity} \quad (86)$$

and

$$\sigma_0^{(n,k)} = \begin{pmatrix} c^{2(n-k)}s^{2k} & -c^n s^n \\ -c^n s^n & c^{2k}s^{2(n-k)} \end{pmatrix} \text{ for odd parity} \quad (87)$$

where k ranges between 0 and n. For each $0 \leq k \leq \lfloor n/2 \rfloor$, the blocks $\sigma_p^{(n,k)}$ and $\sigma_p^{(n,n-k)}$ each make an appearance a total of $\frac{1}{2}\binom{n}{k}$ times.

With this as a starting point, let us develop a convenient notation. If σ is an $n \times n$ positive semidefinite matrix of the form

$$\begin{pmatrix} \sigma^u & \mathbf{0}_{pq} \\ \mathbf{0}_{qp} & \sigma^l \end{pmatrix} \quad (88)$$

where σ^u is a $p \times p$ matrix, σ^l is a $q \times q$ matrix, $\mathbf{0}_{pq}$ is a $p \times q$ matrix, and $n = p + q$, then we shall write this as $\sigma = \sigma^u \oplus \sigma^l$. In this fashion, we have

$$\rho_0^{(n)} = \bigoplus_{k=1}^{2^{n-1}} \rho_{(0,k)} \quad (89)$$

for the appropriate 2×2 matrices $\rho_{(0,k)}$. Similarly for $\rho_1^{(n)}$.

It is not difficult to see that the following three equalities hold:

$$\mathrm{Tr}\,(\sigma^u \oplus \sigma^l) = \mathrm{Tr}\,(\sigma^u) + \mathrm{Tr}\,(\sigma^l) \quad (90)$$

$$(\sigma_0^u \oplus \sigma_0^l)(\sigma_1^u \oplus \sigma_1^l) = (\sigma_0^u \sigma_1^u) \oplus (\sigma_0^u \sigma_1^u) \quad (91)$$

$$\sqrt{\sigma^u \oplus \sigma^l} = \sqrt{\sigma^u} \oplus \sqrt{\sigma^l}. \quad (92)$$

From this it follows that

$$\mathrm{Tr}\sqrt{\sqrt{\sigma_0}\sigma_1\sqrt{\sigma_0}} = \mathrm{Tr}\sqrt{\sqrt{\sigma_0^u}\sigma_1^u\sqrt{\sigma_0^u}} + \mathrm{Tr}\sqrt{\sqrt{\sigma_0^l}\sigma_1^l\sqrt{\sigma_0^l}} \quad (93)$$

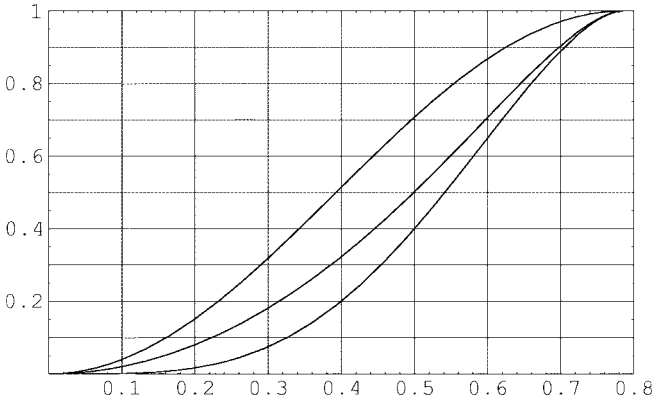

Fig. 2. Equation (45) for the parity bit with $n = 2$ and with $\alpha \in [0, \pi/4]$ on the horizontal axis.

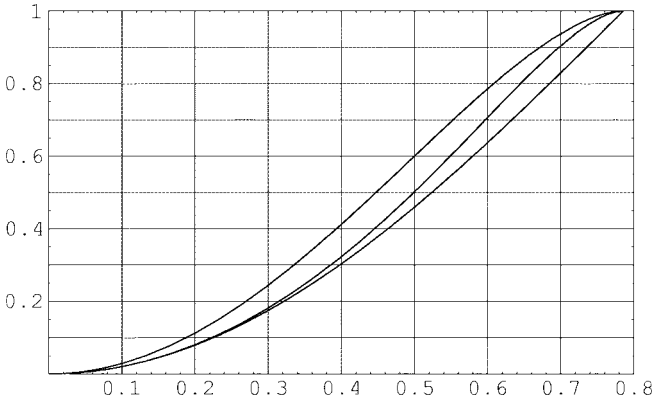

Fig. 3. Equation (46) for the parity bit with $n = 2$ and with $\alpha \in [0, \pi/4]$ on the horizontal axis.

which we can write in a shorthand notation as[6]

$$\mathsf{B}(\sigma_0, \sigma_1) = \mathsf{B}(\sigma_0^u, \sigma_1^u) + \mathsf{B}(\sigma_0^l, \sigma_1^l). \quad (94)$$

Thus we can evaluate $\mathsf{B}(\rho_0^{(n)}, \rho_1^{(n)})$ by evaluating each block individually and summing the results. In particular, we find that

$$\begin{aligned} \mathsf{B}(\sigma_0^{(n,k)}, \sigma_1^{(n,k)}) &= \mathsf{B}(\sigma_0^{(n,n-k)}, \sigma_1^{(n,n-k)}) \\ &= |c^{2(n-k)}s^{2k} - c^{2k}s^{2(n-k)}|. \end{aligned} \quad (95)$$

Summing up over all blocks of $\rho_i^{(n)}$ we get

$$\mathsf{B}(\rho_0^{(n)}, \rho_1^{(n)}) = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{k} |c^{2(n-k)}s^{2k} - c^{2k}s^{2(n-k)}|. \quad (96)$$

For the case $n = 2$ this expression reduces to

$$\mathsf{B}(\rho_0^{(2)}, \rho_1^{(2)}) = |c^4 - s^4| = |(c^2 - s^2)(c^2 + s^2)| = |C| \quad (97)$$

where $C = \cos 2\alpha$.

For the Shannon distinguishability in the special case $n = 2$, [5, eq. (43)] gives

$$\mathsf{SD}(\rho_0^{(2)}, \rho_1^{(2)}) = \frac{1}{2}(1 + C^2)H_2\left(1 - \frac{C^2}{1 + C^2}\right) + \frac{S^2}{2}. \quad (98)$$

We are now in a position to substitute (85), (97), and (98) into (44), (45), and (46). Observe that (44) holds automatically,

[6]Note that the expressions in this shorthand version are not proper Bhattacharyya coefficients: they are not normalized properly.

in fact with equality on the right-hand side. Equations (45) and (46) are illustrated in Figs. 2 and 3, respectively. The horizontal axis represents the angle α between $|\psi_i\rangle$ and $\binom{1}{0}$, meaning that for $\pi/4\,(\approx 0.785)$ the states $|\psi_0\rangle$ and $|\psi_1\rangle$ are orthogonal. The fact that the bounds based on the Bhattacharyya coefficient are fairly tight can be explained by the fact that the function $2\sqrt{x(1-x)}$ resembles $H_2(x)$ quite well.

ACKNOWLEDGMENT

We are grateful to T. Mor for helping us realize the usefulness of these results. Special thanks are also due M. Boyer and C. Crépeau.

REFERENCES

[1] E. Biham, M. Boyer, G. Brassard, J. van de Graaf, and T. Mor, "Security of quantum key distribution against all collective attacks," Tech. Rep. 9801022, LANL Quant-ph archives, 1998. [Online]. Available: http://xxx.lanl.gov/ps/quant-ph/9801022.

[2] S. Goldwasser, S. Micali, and C. Rackoff, "The knowledge complexity of interactive proof-systems," *SIAM. J. Comput.*, vol. 18, no. 1, pp. 186–208, 1989.

[3] D. Beaver, "Secure multiparty protocols and zero-knowledge proof systems tolerating a faulty minority," *J. Cryptol.*, vol. 4, no. 2, pp. 75–122, 1991.

[4] C. A. Fuchs, "Distinguishability and accessible information in quantum theory," Ph.D. dissertation, University of New Mexico, Albuquerque, 1995.

[5] C. H. Bennett, T. Mor, and J. Smolin, "The parity bit in quantum cryptography," *Phys. Rev. A*, vol. 54, no. 3, p. 2675, 1996. [Online]. Available: http://xxx.lanl.gov/ps/quant-ph/9604040.

[6] R. I. G. Hughes, *The Structure and Interpretation of Quantum Mechanics*. Boston, MA: Harvard Univ. Press, 1989.

[7] C. Isham, *Lectures on Quantum Theory*. London, U.K.: Imperial College Press, 1995.

[8] G. Brassard, C. Crépeau, R. Jozsa, and D. Langlois, "A quantum bit commitment scheme provably unbreakable by both parties," in *Proc. 34th IEEE Symp. Foundations of Computer Science*, 1993, pp. 362–371.

[9] Sudbery, *Quantum Mechanics and the Particles of Nature—An Outline for Mathematicians*. Cambridge, U.K.: Cambridge Univ. Press, 1986.

[10] A. Peres, *Quantum Theory: Concepts and Methods*. Dordrecht, The Netherlands: Kluwer, 1993.

[11] N. D. Mermin, "The Ithaca interpretation of quantum mechanics," Tech. Rep. 9609013, LANL Quant-ph Archives, 1996 [Online]. Available: http://xxx.lanl.gov/ps/quant-ph/9609013.

[12] M. Ben-Bassat, "Use of distance measures, information measures and error bounds in feature evaluation," in *Handbook of Statistics, Vol. 2—Classification, Pattern Recognition and Reduction of Dimensionality*, P. R. Krishnaiah and L. N. Kanal, Eds. Amsterdam, The Netherlands: North-Holland, 1982, pp. 773–791.

[13] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Mathematics in Science and Engineering), vol. 123. New York: Academic, 1976.

[14] C. A. Fuchs, "Information Gain vs. State Disturbance in Quantum Theory," *Fort. Phys.*, vol. 46, pp. 535–565, 1998.

[15] A. C.-C. Yao, "Protocols for secure computations," in *Proc. 23rd IEEE Symp. Foundations of Computer Science* (Chicago, IL, 1982), pp. 160–164.

[16] S. Goldwasser and S. Micali, "Probabilistic encryption," *J. Comput. Syst. Sci.*, vol. 28, no. 2, pp. 270–299, 1984.

[17] G. T. Toussaint, "Comments on 'The divergence and Bhattacharyya distance measures in signal selection'," *IEEE Trans. Commun. Technol.*, vol. COM-20, p. 485, 1972.

[18] M. Reed and B. Simon, *Methods of Modern Mathematical Physics—Part I: Functional Analysis*. New York: Academic, 1972.

[19] A. Orlowski, "Measures of distance between quantum states," in *Proc. 4th Workshop Physics and Computation—PhysComp '96* (New England Complex Systems Institute, Boston, MA, 1996), pp. 239–242.

[20] T. Kailath, "The divergence and Bhattacharyya distance measures in signal selection," *IEEE Trans. Commun. Technol.*, vol. COM-15, no. 1, pp. 52–60, 1967.

[21] G. T. Toussaint, "Some functional lower bounds on the expected divergence for multihypothesis pattern recognition, communication, and radar systems," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-1, pp. 384–385, 1971.

[22] C. A. Fuchs and C. M. Caves, "Mathematical techniques for quantum communication theory," *Open Syst. Inform. Dynamics*, vol. 3, no. 3, pp. 345–356, 1995.

[23] A. Uhlmann, "The 'transition probability' in the state space of a $*$-algebra," *Reps. Math. Phys.*, vol. 9, pp. 273–279, 1976.

[24] R. Jozsa, "Fidelity for mixed quantum states," *J. Modern Opt.*, vol. 41, no. 12, pp. 2315–2323, 1994.

[25] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423 and 623–656, 1948.

[26] D. V. Lindley, "On a measure of the information provided by an experiment," *Ann. Math. Statist.*, vol. 27, pp. 986–1005, 1956.

[27] C. A. Fuchs and C. M. Caves, "Ensemble-dependent bounds for accessible information in quantum mechanics," *Phys. Rev. Lett.*, vol. 73, no. 23, pp. 3047–3050, 1994.

[28] T. Mor, "Quantum memory in quantum cryptography," Ph.D. dissertation, Technion–Israel Inst. Technol., Haifa, Israel, 1997.

[29] M. E. Hellman and J. Raviv, "Probability of error, equivocation, and the Chernoff bound," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 368–372, July 1970.

[30] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley Series in Telecommunications). New York: Wiley, 1991.

[31] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," in *Proc. IEEE Int. Conf. Computers, Systems and Signal Processing* (Bangalore, India, 1984), pp. 175–179.