The Operational Meaning of Min- and Max-Entropy

Robert König, Renato Renner, and Christian Schaffner

***Abstract*—In this paper, we show that the conditional min-entropy $H_{\min}(A|B)$ of a bipartite state ρ_{AB} is directly related to the maximum achievable overlap with a maximally entangled state if only local actions on the B-part of ρ_{AB} are allowed. In the special case where A is classical, this overlap corresponds to the probability of guessing A given B. In a similar vein, we connect the conditional max-entropy $H_{\max}(A|B)$ to the maximum fidelity of ρ_{AB} with a product state that is completely mixed on A. In the case where A is classical, this corresponds to the security of A when used as a secret key in the presence of an adversary holding B. Because min- and max-entropies are known to characterize information-processing tasks such as randomness extraction and state merging, our results establish a direct connection between these tasks and basic operational problems. For example, they imply that the (logarithm of the) probability of guessing A given B is a lower bound on the number of uniform secret bits that can be extracted from A relative to an adversary holding B.**

***Index Terms*—Entropy measures, max-entropy, min-entropy, operational interpretations, quantum information theory, quantum hypothesis testing, singlet fraction, single-shot information theory.**

I. Introduction

A CENTRAL goal of information theory is the (quantitative) analysis of processes involving the acquisition, transmission, and storage of information. For example, given a (noisy) communication channel, one may ask at which rate data can be transmitted reliably (this is the *channel capacity*). Or, given a source emitting signals, one may be interested in the amount of space needed to store the information in such a way that the signal can be recovered later (this is the *compression rate*). In the following, we call such quantities *operational* because they are defined by an actual information-processing task.

Traditionally, most operational quantities are defined *asymptotically* under the assumption that a certain process is repeated many times *independently*.[1] Consider, for example, the problem of data compression. For a random variable X and for $\varepsilon \geq 0$, let $\ell^{\varepsilon}_{\mathsf{compr}}(X)$ be the minimum length (measured in terms of bits) of an encoding $\mathrm{enc}(X)$ such that X can be recovered from $\mathrm{enc}(X)$ except with an error probability of at most ε. The *compression rate* of a source emitting a sequence of mutually independent pieces of data $X_1, \ldots, X_n$, each distributed according to P_X, is then defined by

$$r_{\mathsf{compr}}(P_X) := \lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{\ell^{\varepsilon}_{\mathsf{compr}}(X_1 \ldots X_n)}{n}. \tag{1}$$

It may be one of the most remarkable features of information theory that a huge variety of operational quantities can be expressed in terms of a few simple entropy measures. In fact, in the asymptotic case where a process is repeated many times independently, the (almost) only relevant entropy measure is the *Shannon entropy* (or its quantum-mechanical generalization, the *von Neumann entropy*). For example, the compression rate (1) of a source emitting data distributed according to P_X is equal to the Shannon entropy S of a random variable X with distribution P_X, i.e.,

$$r_{\mathsf{compr}}(P_X) = S(X). \tag{2}$$

This equality is also known as the *source-coding theorem* [1]. Another well-known example is the channel capacity. According to the *noisy-channel coding theorem* [1], the maximum rate at which information can be transmitted over a noisy communication channel is equal to a difference between two Shannon entropies [see (13)].

The situation is different in the nonasymptotic case or when the independence assumption is dropped. Here, the Shannon/von Neumann entropies no longer give a correct characterization of operational quantities.[2] Therefore, one has to replace them by more general entropy measures. In the past few years, several such generalizations have been developed, notably the *spectral entropy rates* [2], as well as *(smooth) min- and max-entropies* [3]. While both notions completely overcome the need for independence or Markovian assumptions, spectral entropy rates are (as suggested by their name) still restricted to asymptotic considerations. In contrast, smooth min- and max-entropies are fully general.[3] In particular, no repetition of random processes is required. That is, one may consider situations where a source only emits one single piece of information or where a channel is only used once.

The aim of this paper is to propose new operational interpretations of these nonasymptotic entropy measures. Our

Manuscript received August 15, 2008; revised May 04, 2009. Current version published August 19, 2009. The work of R. König was supported by the National Science Foundation (NSF) under Grants PHY-0456720 and PHY-0803371. The work of C. Schaffner was supported by the European Union fifth framework project QAP IST 015848 and the NWO VICI project 2004–2009. The work of R. Renner was supported by the Swiss National Science Foundation under Grant 200021-119868.

R. König is with the Institute for Quantum Information, California Institute of Technology, Pasadena, CA 91125 USA.

R. Renner is with the Institute for Theoretical Physics, ETH Zurich, Zurich 8093, Switzerland.

C. Schaffner is with the Centrum Wiskunde & Informatica (CWI), Amsterdam 1090, The Netherlands.

Communicated by P. Hayden, Associate Editor for Quantum Information Theory.

Digital Object Identifier 10.1109/TIT.2009.2025545

[1]The independence assumption is sometimes replaced by the less restrictive requirement that the process is Markovian.

[2]For example, the minimum compression length $\ell^{\varepsilon}_{\mathsf{compr}}(X)$ defined above can deviate arbitrarily from $S(X)$. This is readily verified by the following example: Let X be defined as the random variable which takes the value 0 with probability $\frac{1}{2}$, and with probability $\frac{1}{2}$ is equal to a uniformly distributed bit string of length n. Then, $S(X) \approx \frac{n}{2}$ while $\ell^{\varepsilon}_{\mathsf{compr}}(X) \approx n$ for any sufficiently small ε.

[3]The spectral entropy rates can be seen as asymptotic limits of smooth min-/max-entropies [4].

main findings are motivated and described in the following sections, which are organized as follows. In Section I-A, we review the notion of min-/max-entropies, our central object of interest. These entropy measures are the basis for the definition of *smooth* min-/max-entropies, which can be seen as generalizations of Shannon/von Neumann entropy, as indicated above. Their properties are discussed later in Section I-A. After this preparation, we will turn to connections between (smooth) min/max-entropies and operational quantities, starting with some important examples in Section I-B. We then summarize the new operational interpretations derived in this work as well as their implications in Section I-C.

A. (Smooth) Min-/Max-Entropy: Basic Definitions

1) Min-/Max-Entropy: We start with the definition of *conditional min-entropy*. This quantity (and the closely related conditional max-entropy) is the main object of study of this paper. In what follows, id_A denotes the identity on system A.

Definition 1: Let $\rho = \rho_{AB}$ be a bipartite density operator. The min-entropy of A conditioned on B is defined by

$$H_{\min}(A|B)_\rho := -\inf_{\sigma_B} D_\infty(\rho_{AB}||\mathsf{id}_A \otimes \sigma_B) \tag{3}$$

where the infimum ranges over all normalized density operators σ_B on subsystem B and where[4]

$$D_\infty(\tau||\tau') := \inf\{\lambda \in \mathbb{R} : \tau \leq 2^\lambda \tau'\}. \tag{4}$$

It is interesting to note that the Shannon/von Neumann entropy could be defined in a similar way. Namely, if we replace $H_{\min}$ by the von Neumann entropy S and D_∞ by the relative entropy[5] D in (3), we find

$$S(A|B)_\rho = -\inf_{\sigma_B} D(\rho_{AB}||\mathsf{id}_A \otimes \sigma_B).$$

This equality is readily verified using the fact that $D(\tau||\tau')$ is nonnegative for any normalized τ, τ' and equals zero if $\tau = \tau'$.

For a tripartite pure state $\rho = \rho_{ABC}$, the von Neumann entropy satisfies the equality[6]

$$S(A|B)_\rho = -S(A|C)_\rho. \tag{5}$$

The same is no longer true for the min-entropy. However, it turns out that the entropy obtained by replacing the system B by the "purifying system" C often appears in expressions characterizing operational quantities. This motivates the following definition.

Definition 2: Let $\rho = \rho_{AB}$ be a bipartite density operator. The max-entropy of A conditioned on B is defined by

$$H_{\max}(A|B)_\rho := -H_{\min}(A|C)_\rho \tag{6}$$

where the min-entropy on the right-hand side is evaluated for a purification ρ_{ABC} of ρ_{AB}.[7]

This is well defined because all purifications of ρ_{AB} are related by unitaries on C, and the quantity $H_{\min}(A|C)_\rho$ is invariant under such unitaries.

We point out that $H_{\min}$ and $H_{\max}$ could have been defined alternatively by starting from an expression for $H_{\max}$ and subsequent definition of $H_{\min}$ by purification [i.e., (6)]. In this sense, both quantities are equally fundamental.

If the state ρ is clear from the context, we will omit the subscript in $H_{\min}(A|B)_\rho$ and $H_{\max}(A|B)_\rho$. Also, in the special case where the system B is trivial (i.e., one-dimensional), we omit the conditioning and simply write $H_{\min}(A)$ and $H_{\max}(A)$. Note that the above definitions also apply to classical probability distributions P_X which can always be written as quantum states $\rho_X = \sum_x P_X(x)|x\rangle\langle x|$ for some orthonormal basis $\{|x\rangle\}_x$.

To get some more intuition for these definitions, it may help to compute their value for certain special states. One extreme case are product states $\rho_{AB} = \rho_A \otimes \rho_B$, for which one readily verifies that the min-entropy only depends on the maximum eigenvalue $\|\rho_A\|_\infty$ of ρ_A, i.e., $H_{\min}(A|B)_\rho = -\log_2 \|\rho_A\|_\infty$. Note that this corresponds to the Rényi entropy of order infinity of the density operator ρ_A. Similarly, we get $H_{\max}(A|B)_\rho = 2\log_2 \mathsf{tr}\sqrt{\rho_A}$, which is the Rényi entropy of order $\frac{1}{2}$ of ρ_A [see (24)]. Another extreme case is where ρ_{AB} is a pure state. Here, one finds $H_{\min}(A|B)_\rho = -\log_2(\mathsf{tr}\sqrt{\rho_A})^2$ and $H_{\max}(A|B)_\rho = \log_2 \|\rho_A\|_\infty$.

2) Smooth Min-/Max-Entropy: The *smooth* min/max-entropy of a state ρ is defined by the corresponding (nonsmooth) min/max-entropy for an "optimal" state ρ' in a ε-neighborhood of ρ, where ε is called *smoothness parameter*.

Definition 3: Let $\rho = \rho_{AB}$ be a bipartite density operator and let $\varepsilon \geq 0$. The ε-smooth min- and max-entropy of A conditioned on B are given by

$$H_{\min}^\varepsilon(A|B)_\rho := \sup_{\rho'} H_{\min}(A|B)_{\rho'}$$
$$H_{\max}^\varepsilon(A|B)_\rho := \inf_{\rho'} H_{\max}(A|B)_{\rho'}$$

where the supremum ranges over all density operators $\rho' = \rho'_{AB}$ which are ε-close to ρ.[8]

[4]For commuting density operators τ and τ', the quantity $D_\infty(\tau||\tau')$ corresponds to the (classical) relative Rényi entropy of order ∞. In general, the relative Rényi entropy of order α of two probability distributions P and Q is defined as $D_\alpha(P,Q) := \frac{1}{\alpha-1}\log_2 \sum_x P_X(x)^\alpha Q(x)^{1-\alpha}$, and D_∞ is obtained in the limit $\alpha \to \infty$.

[5]Note that the relative entropy (also knows as Kullback–Leibler divergence) $D(\tau||\tau') := \mathsf{tr}(\tau(\log_2 \tau - \log_2 \tau'))$ is also defined for unnormalized operators τ'.

[6]Note that, by definition, $S(A|B) = S(AB) - S(B)$ and $S(A|C) = S(AC) - S(C)$. The equality then follows from the fact that, by the Schmidt decomposition, $S(AB) = S(C)$ and $S(B) = S(AC)$.

[7]In the existing literature, $H_{\max}$ and $H_{\max}^\varepsilon$ are sometimes defined in a different manner (closely related to the Rényi entropy of order 0). It can be shown, however, that the smooth variants of these definitions only deviate by an additive term which is logarithmic in the smoothness parameter (see [5]).

[8]In the classical case, smooth entropies are usually defined with respect to the trace distance $\delta_{\mathsf{tr}}(\rho,\sigma) = \frac{1}{2}\|\rho - \sigma\|_1$. Quantum mechanically, distance measures based on the fidelity $F(\rho,\sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1$ are more suitable because they are invariant under purifications. Candidates are the *Bures distance* $\|\rho - \sigma\|_B = \sqrt{2 - 2F(\rho,\sigma)}$ and the *angle* $\|\rho - \sigma\|_A = \arccos F(\rho,\sigma)$. The corresponding definitions are essentially equivalent because of the inequalities $1 - F(\rho,\sigma) \leq \delta_{\mathsf{tr}}(\rho,\sigma) \leq \sqrt{1 - F(\rho,\sigma)^2}$.

3) Basic Properties: It follows directly from the definitions that the same kind of duality between min- and max-entropy holds between the corresponding smooth versions, namely

$$H_{\max}^{\varepsilon}(A|B)_{\rho} = -H_{\min}^{\varepsilon}(A|C)_{\rho} \tag{7}$$

for a purification ρ_{ABC} of ρ_{AB}.

As already indicated, smooth min-/max-entropies can be seen as generalizations of the Shannon/von Neumann entropy S. More precisely, the latter can be written in terms of the former [6], [5], i.e.,

$$S(A|B)_{\rho} = \lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}H_{\min}^{\varepsilon}(A^n|B^n)_{\rho^{\otimes n}} \tag{8}$$

$$S(A|B)_{\rho} = \lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}H_{\max}^{\varepsilon}(A^n|B^n)_{\rho^{\otimes n}}. \tag{9}$$

Note that the two statements are trivially equivalent because of (5) and (7).

Given these asymptotic relations, it is not surprising that smooth min-/max-entropies share various properties with the Shannon/von Neumann entropy. For example, they are strongly subadditive, i.e.,

$$H_{\min}^{\varepsilon}(A|B) \geq H_{\min}^{\varepsilon}(A|BC) \tag{10}$$

and likewise for $H_{\max}^{\varepsilon}$. In fact, inequality (10) can be seen as a generalization of the strong subadditivity of the von Neumann entropy $S(A|B) \geq S(A|BC)$, which can be recovered by virtue of identity (8), i.e., for any ρ_{ABC}

$$\begin{aligned} S(A|B)_{\rho} &\overset{(8)}{=} \lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}H_{\min}^{\varepsilon}(A^n|B^n)_{\rho^{\otimes n}} \\ &\overset{(10)}{\geq} \lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}H_{\min}^{\varepsilon}(A^n|B^nC^n)_{\rho^{\otimes n}} \\ &\overset{(8)}{=} S(A|BC)_{\rho}. \end{aligned}$$

Interestingly, despite its generality, inequality (10) is easy to prove, as we will see at the end of Section I-C.

B. Operational Quantities in Terms of Smooth Min-/Max-Entropy

The main reason for considering (smooth) min-/max-entropies is that they are well suited for the characterization of operational quantities in the most general case. Recall that expressions for operational quantities involving the Shannon/von Neumann entropy, e.g., (2), are typically only valid asymptotically, under the assumption that certain resources can be used many times independently. Interestingly, the structure of such expressions essentially remains the same if one drops these assumptions, except that smooth entropies take the place of Shannon/von Neumann entropy. The purpose of this section is to illustrate this phenomenon with a few examples.

1) Data Compression: We start with the example of data compression, which has already been introduced above. For a random variable X and $\varepsilon \geq 0$, let again $\ell_{\mathsf{compr}}^{\varepsilon}(X)$ be the minimum length of an encoding from which the value of X can be recovered correctly with probability at least $1-\varepsilon$. It can then be shown that $\ell_{\mathsf{compr}}^{\varepsilon}(X)$ is essentially equal to the smooth max-entropy of X [3]. More precisely, we have

$$\ell_{\mathsf{compr}}^{\varepsilon}(X) = H_{\max}^{\varepsilon'}(X) + O(\log 1/\varepsilon) \tag{11}$$

for some $\varepsilon' \in [\frac{1}{2}\varepsilon, 2\varepsilon]$. The O-notation indicates that equality holds up to an additive term of the order $\log 1/\varepsilon$.[9] In typical applications, this logarithmic term is much smaller than the other quantities occurring in the expression. In particular, the term is independent of the size of the resource (in our case, the random variable X), and thus becomes irrelevant in the asymptotic limit of large resources.

We stress that (11) is valid for a *single* realization of the random variable X, and thus strictly generalizes Shannon's source coding theorem described at the beginning of this section. Identity (2) can be recovered as an asymptotic limit of (11), for X consisting of many independent and identically distributed (i.i.d.) pieces $X_1, \ldots, X_n$, i.e.,

$$\begin{aligned} r_{\mathsf{compr}}(P_X) &\overset{(1)}{=} \lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{\ell_{\mathsf{compr}}^{\varepsilon}(X_1\ldots X_n)}{n} \\ &\overset{(11)}{=} \lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}H_{\max}^{\varepsilon}(X_1\ldots X_n) \\ &\overset{(9)}{=} S(X). \end{aligned}$$

2) Channel Coding: As a second example, we consider the noisy-channel coding problem. For any $\varepsilon \geq 0$, let $\ell_{\mathsf{transm}}^{\varepsilon}(X \to Y)$ be the maximum number of bits that can be transmitted in *one use* of a classical noisy channel $X \to Y$ (specified by a conditional probability distribution $P_{Y|X}$) with maximum error probability ε. As shown in [7], this quantity is given by

$$\ell_{\mathsf{transm}}^{\varepsilon}(X\to Y) = \max_{P_X}\left(H_{\min}^{\varepsilon'}(X) - H_{\max}^{\varepsilon'}(X|Y)\right) + O(\log 1/\varepsilon) \tag{12}$$

for some $\varepsilon' \in [\frac{1}{2}\varepsilon, 2\varepsilon]$.

Similarly to the example of source coding, we may consider the special case where the channel allows many mutually independent transmissions. The figure of merit then is the *channel capacity*

$$r_{\mathsf{transm}}(P_{Y|X}) := \lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{\ell_{\mathsf{transm}}^{\varepsilon}(X^n \to Y^n)}{n}$$

that is, the maximum rate at which information can be transmitted by n uses of the channel $X \to Y$, in the limit of large n. Using the nonasymptotic statement (12) together with (8) and (9), we find

$$\begin{aligned} r_{\mathsf{transm}}(P_{Y|X}) &= \max_{P_X}(S(X) - S(X|Y)) \\ &= \max_{P_X} I(X:Y). \end{aligned} \tag{13}$$

This is Shannon's well-known noisy-channel coding theorem.

3) Privacy Amplification: Let X be a classical random variable and let B be (possibly quantum-mechanical) side information. The goal of *randomness extraction* is to compute a bit

[9]Note that the smooth entropies are monotonic functions of ε. Equality (11) is thus just a way to state that the operational quantity $\ell_{\mathsf{compr}}^{\varepsilon}(X)$ lies in the interval $[H_{\max}^{2\varepsilon}(X), H_{\max}^{\varepsilon/2}(X)]$, up to some additive constant of the order $\log 1/\varepsilon$.

string $f(X)$ which is uniform and independent of the side information B. Randomness extraction is crucial for a number of applications, particularly in the context of cryptography, where it is also called *privacy amplification* [8]. For example, in a key-agreement scheme, one may want to turn a (only partially secure) raw key X into a fully secure key $f(X)$. Security of $f(X)$ is then akin to uniformity relative to side information B held by a potential adversary.

The maximum number of uniform and independent bits that can be extracted from X is directly given by the smooth min-entropy of X. More precisely, let $\ell^{\varepsilon}_{\mathsf{extr}}(X|B)$ be the maximum length of a bit string $f(X)$ that can be computed from X such that $f(X)$ is ε-close to a string Z which is perfectly uniform and independent of the side information B.[10] One can show that [6], [9]

$$\ell^{\varepsilon}_{\mathsf{extr}}(X|B) = H^{\varepsilon'}_{\min}(X|B) + O(\log 1/\varepsilon)$$

where $\varepsilon' \in [\frac{1}{2}\varepsilon, 2\varepsilon]$. In the special case where B is independent of X, this corresponds to the *leftover hash lemma* [10], [8]. For later reference, we also note that

$$\ell^{\varepsilon}_{\mathsf{extr}}(X|B) \geq H_{\min}(X|B) + O(\log 1/\varepsilon) \tag{14}$$

which holds because $H^{\varepsilon}_{\min}(X|B)$ is monotonically increasing in ε and equals $H_{\min}(X|B)$ for $\varepsilon = 0$.

4) Decoupling: The previous result can be extended to a fully quantum-mechanical setting as follows. Let A and B be two quantum systems. The goal is to find a maximum subsystem A' of A such that the state on A' is completely mixed and decoupled from B (conditioned on a suitable measurement on the remaining part of A). Let $\ell^{\varepsilon}_{\mathsf{decpl}}(A|B)$ be the maximum size of A' (measured in qubits) such that this is possible up to a distance ε.[11] One then finds [11]–[13][12]

$$\ell^{\varepsilon}_{\mathsf{decpl}}(A|B) = H^{\varepsilon'}_{\min}(A|B) + O(\log 1/\varepsilon) \tag{15}$$

with $\varepsilon' \in [\frac{1}{2}\varepsilon, 2\varepsilon]$.

5) State Merging: In the same manner as privacy amplification generalizes to decoupling in the fully quantum case, data compression (and its relatives such as coding with side information) extends to a fully quantum setting; this is referred to as state merging. The setting is described by a tripartite pure state $|\Psi_{ABC}\rangle$. The aim is to redistribute the A-part to the system B by local operations and classical communications (LOCC) between A and B. Depending on the (reduced) state $\rho = \rho_{AB}$, this either consumes or generates bipartite entanglement. Let $(-)\ell^{\varepsilon}_{\mathsf{merg}}(A|B)_{\rho}$ be the minimal (maximal) number of ebits of entanglement required (generated) by this process [the distinction between consumed/generated entanglement is reflected by the sign of the quantity $\ell^{\varepsilon}_{\mathsf{merg}}(A|B)_{\rho}$], such that the outcome is ε-close to the desired output.[13] One then finds

$$\ell^{\varepsilon}_{\mathsf{merg}}(A|B)_{\rho} = H^{\varepsilon'}_{\max}(A|B)_{\rho} + O(\log 1/\varepsilon) \tag{16}$$

where again $\varepsilon' \in [\frac{1}{2}\varepsilon, 2\varepsilon]$ (see [13] for details). In fact, the $\geq$-part of this statement is a direct consequence of decoupling result above [12], the arguments in [11] (cf., also Section III-B), and the definition of $H_{\max}(A|B)_{\rho}$.

C. Contribution: Min-/Max-Entropies as Operational Quantities

In this paper, we show that min-/max-entropies have direct[14] operational interpretations. We begin by presenting the corresponding results for the special case where we condition classical information X on a (possibly) quantum system B. The fully general case is discussed in Section I-C2.

1) Uncertainty About Classical Information: Consider an agent with access to a (classical or quantum) system B whose state ρ^x_B depends on a classical random variable X. This situation can be described by a classical-quantum state

$$\rho = \rho_{XB} := \sum_x P_X(x)|x\rangle\langle x| \otimes \rho^x_B \tag{17}$$

with $\{|x\rangle\}_x$ a family of mutually orthogonal vectors representing the (classical) values of X.

a) Min-Entropy of Classical Information is Guessing Probability: Let $p_{\mathsf{guess}}(X|B)$ be the probability that the agent correctly guesses X when using an optimal strategy; that is, $p_{\mathsf{guess}}(X|B) = \sum_x P_X(x)\mathsf{tr}(E_x\rho^x_B)$, where the optimal measurement strategy is described by the positive operator-valued measure (POVM) $\{E_x\}_x$ on B that maximizes this expression. Note that conditions for the optimality of a POVM $\{E_x\}_x$ in this hypothesis testing problem were found by Holevo [14] and independently by Yuen, Kennedy, and Lax [15]. These works also use semidefinite programming duality in a similar fashion as in this paper. Here, we are interested in the optimal value of this optimization problem. We show that (cf., Theorem 1)

$$p_{\mathsf{guess}}(X|B) = 2^{-H_{\min}(X|B)_{\rho}} \tag{18}$$

where the entropy is evaluated for the state ρ_{XB} given by (17).

If no side information B is available or, more generally, if the state of B is independent of X, we have $2^{-H_{\min}(X|B)} = \|\rho_X\|_{\infty} = \max_x P_X(x)$ as noted in Section I-A. Identity (18) then reduces to the trivial fact that the maximum probability of correctly guessing X without prior information is equal to $\max_x P_X(x)$.

Note that previously, only the upper bound [16]

$$p_{\mathsf{guess}}(X|B) \leq 2^{-H_{\min}(X|B)_{\rho}}$$

[10]See paragraph on max-entropy of classical information in Section I-C for more details.

[11]This distance is quantitatively expressed by the *decoupling accuracy*; see below.

[12]This is based on a tightened version [12] of a bound obtained in [11], which shows that projecting onto a random subspace of dimension $\dim A'$ achieves decoupling. More precisely, it can be shown [12] that the decoupling accuracy of the residual state is, on average over the measurement outcome, exponentially small in the difference $H_{\min}(A|B) - \log \dim A'$.

[13]Closeness is measured in terms of the distance of the output state $\rho'_{ABB_1B_2C}$ of the protocol to the state $|\Phi_{AB}\rangle^{\otimes \ell} \otimes |\Psi_{B_1B_2C}\rangle$, where $|\Phi_{AB}\rangle$ is an ebit between A and B, ℓ is the number of ebits generated, and $|\Psi_{B_1B_2C}\rangle$ is identical to $|\Psi_{ABC}\rangle$ when identifying the subsystems B_1 and A as well as B_2 and B.

[14]The term *direct* refers to the fact that no smoothing is required, in contrast to the examples of Section I-B.

and the lower bound [17]

$$2^{-H_2(X|B)_\rho} \leq p_{\mathsf{guess}}(X|B)$$

were known, where the left-hand side is the average guessing probability when the square root measurement [18] is used, that is

$$H_2(X|B)_\rho = -\log \mathrm{tr}\left(\left((\mathrm{id}_X \otimes \rho_B^{-1/2})\rho_{XB}\right)^2\right).$$

b) Max-Entropy of Classical Information is Security of Key: The secrecy of X when used as a key in the presence of an adversary with access to system B is conventionally measured in terms of the distance of the state ρ_{XB} [cf., (17)] to a product state of the form $\tau_X \otimes \rho_B$, where τ_X is the completely mixed state (corresponding to the uniform distribution on X) and where ρ_B is the reduced state on subsystem B. This models an *ideal* situation where the key is perfectly uniform and independent of the adversary's system. If the trace distance is used, then this distance is directly related to the *distinguishing advantage* between the real and the ideal system.

One may relax the above and may only require that the desired state is of the form $\tau_X \otimes \sigma_B$, for some *arbitrary* density operator σ_B. When using the trace distance, this relaxed definition is equivalent to the above up to a factor of 2. Also, since the trace distance and the fidelity are essentially equivalent, we can use the fidelity. We then get the following measure for the secrecy of X relative to B:

$$\begin{aligned} p_{\mathsf{secr}}(X|B)_\rho &:= |X| \max_{\sigma_B} F(\rho_{XB}, \tau_X \otimes \sigma_B)^2 \\ &= \max_{\sigma_B} \left(\sum_x \sqrt{P_X(x)} F(\rho_B^x, \sigma_B)\right)^2 \end{aligned}$$

where $|X|$ is the alphabet size of X (we include this factor here for convenience). We show that (cf., Theorem 3)

$$p_{\mathsf{secr}}(X|B)_\rho = 2^{H_{\max}(X|B)_\rho}. \tag{19}$$

If no side information B is available or, more generally, if B is independent of X, we obtain $2^{H_{\max}(X|B)_\rho} = (\sum_x \sqrt{P_X(x)})^2$ (cf., Section I-A). Identity (19) then simply expresses the fact that the secrecy of X in this case is quantified by the distance of P_X to the uniform distribution (where distance is measured in terms of the fidelity).

2) Uncertainty About Quantum Information: We now discuss the fully general case, where we have an arbitrary bipartite state $\rho = \rho_{AB}$. The min-/max-entropies carry the following operational interpretations.

a) Min-Entropy is Maximum Achievable Singlet Fraction: Define the maximally entangled state

$$|\Phi_{AB}\rangle := \frac{1}{\sqrt{d}} \sum_x |x_A\rangle |x_B\rangle$$

where $\{|x_A\rangle\}_{x=1}^d$ is an orthonormal basis of subsystem A (of dimension d) and $\{|x_B\rangle\}_{x=1}^d$ is a family of mutually orthogonal vectors on subsystem B (we assume that $\dim A \leq \dim B$). We define the "quantum correlation" $q_{\mathsf{corr}}(A|B)_\rho$ as the maximum overlap with the singlet[15] state $|\Phi_{AB}\rangle$ that can be achieved by local quantum operations $\mathcal{E}$ (trace-preserving completely positive maps) on subsystem B, that is

$$\begin{aligned} &q_{\mathsf{corr}}(A|B)_\rho \\ &\quad := d \max_{\mathcal{E}} F\left((\mathrm{id}_A \otimes \mathcal{E})(\rho_{AB}), |\Phi_{AB}\rangle\langle\Phi_{AB}|\right)^2. \end{aligned} \tag{20}$$

We show that (cf., Theorem 2)

$$q_{\mathsf{corr}}(A|B)_\rho = 2^{-H_{\min}(A|B)_\rho}. \tag{21}$$

Note that in the case where the information is classical, i.e., if $\rho = \rho_{XB}$ is of the form (17), we have

$$q_{\mathsf{corr}}(X|B)_\rho = \max_{\mathcal{E}} \sum_x P_X(x) \langle x|\mathcal{E}(\rho_B^x)|x\rangle.$$

The operation $\mathcal{E}$ can be interpreted as a guessing strategy, so that $\langle x|\mathcal{E}(\rho_B^x)|x\rangle$ becomes the probability of correctly guessing X if $X = x$. We thus recover the maximum guessing probability p_{guess} as a special case, i.e.,

$$q_{\mathsf{corr}}(X|B)_\rho = p_{\mathsf{guess}}(X|B).$$

b) Max-Entropy is Decoupling Accuracy: The *decoupling accuracy* is a parameter that can be seen as the quantum analog of the error probability in classical coding theorems and is also called *quantum error* in [11] and [19]; it measures the quality of decoupling as follows. It is defined as the distance of ρ_{AB} to the product state $\tau_A \otimes \sigma_B$, where τ_A is the completely mixed state on A and σ_B is an arbitrary density operator. In a cryptographic setting, it quantifies how random A appears from the point of view of an adversary with access to B. As above for classical A, we define a fidelity-based version of this quantity as

$$q_{\mathsf{decpl}}(A|B)_\rho := d_A \max_{\sigma_B} F(\rho_{AB}, \tau_A \otimes \sigma_B)^2 \tag{22}$$

where d_A is the dimension of A and τ_A is the completely mixed state on A. We show that (cf., Theorem 3)

$$q_{\mathsf{decpl}}(A|B)_\rho = 2^{H_{\max}(A|B)_\rho}. \tag{23}$$

It is immediately obvious that this generalizes the security parameter for a classical key X, i.e., for $\rho = \rho_{XB}$ of the form (17), we have

$$q_{\mathsf{decpl}}(X|B)_\rho = p_{\mathsf{secr}}(X|B)_\rho.$$

3) Implications: A main implication of our results is that they establish a connection between seemingly different operational quantities. For example, because the number $\ell_{\mathsf{extr}}^\varepsilon(X|B)$ of uniform bits that can be extracted from X with respect to side information B is lower bounded by $H_{\min}(X|B)$ [see (14)], we find that

$$\ell_{\mathsf{extr}}^\varepsilon(X|B) \geq -\log_2 p_{\mathsf{guess}}(X|B) + O(\log 1/\varepsilon).$$

[15]In the literature, the expression "singlet" often refers to the maximally entangled two-qubit state $\frac{1}{\sqrt{2}}(|0\rangle|1\rangle - |1\rangle|0\rangle)$. Here we use the expressions "singlet" and "singlet fraction" more generally for any maximally entangled state $|\Phi_{AB}\rangle$. This is justified because definition (20) gives the same value independent of the choice of the maximally entangled state $|\Phi_{AB}\rangle$.

In other words, the negative logarithm of the guessing probability of X tells us how many uniform bits we can extract from X (relative to some system B). This connection between randomness extraction and guessing entropy may be useful for applications, e.g., in cryptography. Here, the derivation of lower bounds on the amount of extractable randomness is usually a central part of the security analysis (see [6] and [20]–[22]).

Our results can also be used to prove additivity properties of the min-/max-entropies. One of them is additivity of the min-/max-entropies for independent systems. Let $\rho_{AA'BB'} = \rho_{AB} \otimes \rho_{A'B'}$. Then, by the definition of q_{corr}

$$q_{\mathsf{corr}}(AA'|BB') \geq q_{\mathsf{corr}}(A|B) \cdot q_{\mathsf{corr}}(A'|B').$$

By virtue of (21), this is equivalent to

$$H_{\min}(AA'|BB') \leq H_{\min}(A|B) + H_{\min}(A'|B').$$

Note that the opposite inequality follows immediately from the definition of $H_{\min}$ and the additivity of D_∞. We thus have

$$H_{\min}(AA'|BB')_\rho = H_{\min}(A|B) + H_{\min}(A'|B')$$

and, equivalently [by the definition (6)]

$$H_{\max}(AA'|BB')_\rho = H_{\max}(A|B) + H_{\max}(A'|B').$$

A second example is the strong subadditivity of conditional min-entropy (10). Here, it suffices to notice that every trace-preserving completely positive map $\mathcal{E}$ acting on B can also be understood as acting on registers B and C, hence

$$q_{\mathsf{corr}}(A|B)_\rho \leq q_{\mathsf{corr}}(A|BC)_\rho$$

for every quantum state ρ_{ABC}. By (21), this is equivalent to

$$H_{\min}(A|B)_\rho \geq H_{\min}(A|BC)_\rho.$$

The extension to smooth min-entropy (10) is straightforward (see [6, Lemma 3.2.7]).

Our results also simplify the calculation of the min-/max-entropies. As an example, let us calculate the entropy $H_{\max}(A|B)_\rho$ for a state of the form $\rho_{AB} = \rho_A \otimes \rho_B$. By (23), it suffices to determine the quantity $q_{\mathsf{decpl}}(A|B)_\rho$, which is given by

$$q_{\mathsf{decpl}}(A|B)_\rho = \max_{\sigma_B} d_A F(\rho_A \otimes \rho_B, \tau_A \otimes \sigma_B)$$

where τ_A is the completely mixed state on the d_A-dimensional Hilbert space A. Using the multiplicativity of the fidelity, we find

$$\begin{aligned} q_{\mathsf{decpl}}(A|B)_\rho &= d_A F(\rho_A, \tau_A) \max_{\sigma_B} F(\rho_B, \sigma_B) \\ &= d_A F(\rho_A, \tau_A) \\ &= \|\sqrt{\rho_A}\|_1^2. \end{aligned}$$

We thus obtain

$$H_{\max}(A|B)_\rho = 2 \log \mathsf{tr} \sqrt{\rho_A} \tag{24}$$

for any $\rho = \rho_{AB}$ of the form $\rho_A \otimes \rho_B$. This corresponds to the Rényi entropy of order $\frac{1}{2}$, which is hence the natural counterpart to the min-entropy (Rényi entropy of order ∞). As noted in [3], the Rényi entropy of order α, for any $\alpha < 1$, is—up to small additive terms of the order $\log \frac{1}{\varepsilon}$—determined by a smoothed version of $H_0(\rho_A) := \log_2 \mathsf{rank}(\rho_A)$. The max-entropy $H_{\max}$ of a density operator can thus be interpreted as a measure for its rank.

4) Outline of the Remainder of This Paper: In Section II, we discuss some mathematical preliminaries, in particular, semidefinite programming, which plays a crucial role in our arguments. Our main results are then stated and proved in Section III.

II. Some Technical Preliminaries

A. Semidefinite Programming

Our central tool will be the duality between certain pairs of semidefinite programs. It will be convenient to use a fairly general formulation of this duality; a derivation of the results summarized in this section can be found, e.g., in [23, Section 6]. The presentation here follows this reference, but specializes certain statements to the situation of interest for simplicity. We start by introducing a few definitions.

A subset $K \subset \mathcal{V}$ of a vector space $\mathcal{V}$ is called a *convex cone* if $0 \in K$ and $\mu v + \nu w \in K$ for all nonnegative $\mu, \nu \geq 0$ and $v, w \in K$. A convex cone K gives rise to a partial order relation $\leq_K$ on $\mathcal{V}$, defined by $v \leq_K w$ if and only if $w - v \in K$. If $\mathcal{V}$ is a Euclidean space with inner product $\langle \cdot, \cdot \rangle$, then the *dual cone* $K^* \subset \mathcal{V}$ of K is defined by $K^* = \{v \in \mathcal{V} | \langle v, w \rangle \geq 0 \text{ for all } w \in K\}$. The *interior* $\mathsf{int} K \subset K$ is the subset of points $w \in K$ for which there exists an open ball centered around w and contained in K.

Let $\mathcal{V}_1$ and $\mathcal{V}_2$ be Euclidean spaces with inner products $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$, respectively. A linear map $E^* : \mathcal{V}_2 \to \mathcal{V}_1$ is called *dual of* or *adjoint to* a linear map $E : \mathcal{V}_1 \to \mathcal{V}_2$ if

$$\langle E v_1, v_2 \rangle_2 = \langle v_1, E^* v_2 \rangle_1, \qquad \text{for all } v_1 \in \mathcal{V}_1, v_2 \in \mathcal{V}_2.$$

For a given map E, the dual map E^* is necessarily unique if it exists. The two linear programming problems we are interested in are defined in terms of a pair of such maps. They are referred to as the *primal* and *dual* problem, and are specified by parameters $c \in \mathcal{V}_1$ and $b \in \mathcal{V}_2$. The programs are expressed by the following optimizations:

$$\begin{aligned} \gamma^{\mathsf{primal}} &= \inf_{\substack{v_1 \geq_{K_1} 0 \\ E v_1 \geq_{K_2} b}} \langle v_1, c \rangle_1 \\ \gamma^{\mathsf{dual}} &= \sup_{\substack{v_2 \geq_{K_2^*} 0 \\ E^* v_2 \leq_{K_1^*} c}} \langle b, v_2 \rangle_2. \end{aligned} \tag{25}$$

We will usually assume that the sets we optimize over are nonempty. (In the language of linear programming, there exists a *feasible plan* and a *dual feasible plan*.) The *weak duality theorem* states that $\gamma^{\mathsf{primal}} \geq \gamma^{\mathsf{dual}}$. We are particularly interested in conditions for equality. (This is referred to as a *zero duality gap*.) A simple criterion is *Slater's interiority condition*, which states the following.

Lemma 1: Suppose that there is an element $v \in \mathsf{int}K_1$ such that $Ev - b \in \mathsf{int}K_2$. Suppose further that the infimum in (25) is attained. Then, $\gamma^{\mathsf{primal}} = \gamma^{\mathsf{dual}}$.

B. Quantum Operations

Let $\mathcal{H}_A$ be a Hilbert space and let $\mathcal{L}(\mathcal{H}_A)$ be the set of linear maps $E : \mathcal{H}_A \to \mathcal{H}_A$. An element $E \in \mathcal{L}(\mathcal{H}_A)$ is called *nonnegative* (written $E \geq 0$) if $\langle\psi|E|\psi\rangle \geq 0$ for all $|\psi\rangle \in \mathcal{H}_A$. A *positive* element E (written $E > 0$) is defined in the same way with a strict inequality.

An *operation* is a linear map $\mathcal{E} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$. It is called *trace-preserving* if $\mathsf{tr}(\mathcal{E}(E)) = \mathsf{tr}(E)$ for all $E \in \mathcal{L}(\mathcal{H}_A)$. It is *unital* if it maps the identity on $\mathcal{H}_A$ to the identity on $\mathcal{H}_B$, i.e., if $\mathcal{E}(\mathsf{id}_A) = \mathsf{id}_B$. The map is called *positive* if $\mathcal{E}(E) \geq 0$ for all $E \geq 0$. It is *completely positive* (CP) if $\mathsf{id}_R \otimes \mathcal{E} : \mathcal{L}(\mathcal{H}_R \otimes \mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_R \otimes \mathcal{H}_B)$ is positive for any auxiliary space $\mathcal{H}_R$, where id_R is the *identity operation*. A *quantum operation* is a completely positive trace-preserving map (CPTP). The *adjoint map* of an operation $\mathcal{E} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$ is the unique map $\mathcal{E}^\dagger : \mathcal{L}(\mathcal{H}_B) \to \mathcal{L}(\mathcal{H}_A)$ satisfying

$$\mathsf{tr}(F_B\mathcal{E}(E_A)) = \mathsf{tr}(\mathcal{E}^\dagger(F_B)E_A)$$

for all $E_A \in \mathcal{L}(\mathcal{H}_A)$ and $F_B \in \mathcal{L}(\mathcal{H}_B)$. Note that $(\mathcal{E}^\dagger)^\dagger = \mathcal{E}, \mathsf{id}_A^\dagger = \mathsf{id}_A$ and $(\mathcal{E} \otimes \mathcal{F})^\dagger = \mathcal{E}^\dagger \otimes \mathcal{F}^\dagger$ for two maps $\mathcal{E}$ and $\mathcal{F}$. Two easily verified properties which follow directly from this definition are

$$\mathcal{E} \text{ is unital if and only if } \mathcal{E}^\dagger \text{ is trace-preserving} \tag{26}$$

and

$$\mathcal{E} \text{ is positive if and only if } \mathcal{E}^\dagger \text{ is positive.}$$

In particular, the last statement implies that

$$\mathcal{E} \text{ is completely positive (CP) if and only if } \mathcal{E}^\dagger \text{ is CP.} \tag{27}$$

Statements (26) and (27) can be summarized as follows. Let us define $\mathsf{CPTPM}(\mathcal{H}_A, \mathcal{H}_B)$ as the set of quantum operations $\mathcal{E} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$ and $\mathsf{CPUM}(\mathcal{H}_B, \mathcal{H}_A)$ as the set of completely positive unital maps $\mathcal{F} : \mathcal{L}(\mathcal{H}_B) \to \mathcal{L}(\mathcal{H}_A)$. We then have the following.

Lemma 2: The adjoint map

$$\dagger : \mathsf{CPTPM}(\mathcal{H}_A, \mathcal{H}_B) \to \mathsf{CPUM}(\mathcal{H}_B, \mathcal{H}_A)$$

is a bijection with inverse

$$\dagger : \mathsf{CPUM}(\mathcal{H}_B, \mathcal{H}_A) \to \mathsf{CPTPM}(\mathcal{H}_A, \mathcal{H}_B).$$

Let d_A be the dimension of $\mathcal{H}_A$ and let $\{|x\rangle_A\}_{x\in[d_A]}$ be an orthonormal basis of $\mathcal{H}_A$. (We will restrict our attention to finite-dimensional Hilbert spaces.) Let $\mathcal{H}_{A'} \cong \mathcal{H}_A$ be a Hilbert space with orthonormal basis $\{|x\rangle_{A'}\}_{x\in[d_A]}$. The *maximally entangled state* on $\mathcal{H}_A \otimes \mathcal{H}_{A'}$ is defined as

$$|\Phi_{AA'}\rangle = \frac{1}{\sqrt{d_A}} \sum_{|x\rangle\in[d_A]} |x\rangle_A \otimes |x\rangle_{A'}. \tag{28}$$

The *Choi–Jamiołkowski-map* J takes operations $\mathcal{E} : \mathcal{L}(\mathcal{H}_{A'}) \to \mathcal{L}(\mathcal{H}_B)$ to operators $J(\mathcal{E}) \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$. It is defined as

$$J(\mathcal{E}) = d_A(\mathsf{id}_A \otimes \mathcal{E})(|\Phi_{AA'}\rangle\langle\Phi_{AA'}|).$$

It has the following well-known properties. The equivalence of statements i) and ii) in the following lemma is an immediate consequence of Lemma 2.

Lemma 3 (Choi–Jamiołkowski isomorphism [24]): Let $\mathcal{H}_A \cong \mathcal{H}_{A'}$ and $\mathcal{H}_B$ be arbitrary Hilbert spaces. The map J bijectively maps:

i) the set $\mathsf{CPTPM}(\mathcal{H}_{A'}, \mathcal{H}_B)$ to the set of operators $F_{AB} \geq 0$ with $\mathsf{tr}_B F_{AB} = \mathsf{id}_A$;
ii) the set $\mathsf{CPUM}(\mathcal{H}_{A'}, \mathcal{H}_B)$ to the set of operators $E_{AB} \geq 0$ with $\mathsf{tr}_A E_{AB} = \mathsf{id}_B$.

Another concept we will need is the notion of *classicality*, which allows us to treat ensembles as quantum states. We will say that a Hermitian operator E_{AB} on a bipartite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ is *classical relative to an orthonormal basis* $\{|x_A\rangle\}_{x\in[d_A]}$ *of* $\mathcal{H}_A$ if it is a linear combination of operators of the form $|x\rangle\langle x| \otimes E_B$, where $x \in [d_A]$ and E_B is a Hermitian operator on $\mathcal{H}_B$.

III. Main Results and Their Derivation

We are now ready to prove our main statements. We first focus on the min-entropy in Section III-A. The interpretation of max-entropy will be derived in Section III-B.

A. Proof of the Operational Characterization of $H_{\min}$

With Lemma 1 from Section II-A, it is straightforward to prove the following statement. Note that we restrict our attention to finite-dimensional Hilbert spaces. Since the optimizations are now taken over compact sets, we can replace $\inf$ and $\sup$ by $\min$ and $\max$, respectively.

Lemma 4: Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be finite-dimensional Hilbert spaces, and let ρ_{AB} and σ_B be nonnegative operators on $\mathcal{H}_A \otimes \mathcal{H}_B$ and $\mathcal{H}_B$, respectively. Then

$$\min_{\substack{\sigma_B \geq 0 \\ \mathsf{id}_A \otimes \sigma_B \geq \rho_{AB}}} \mathsf{tr}(\sigma_B) = \max_{\substack{E_{AB} \geq 0 \\ \mathsf{tr}_A(E_{AB}) = \mathsf{id}_B}} \mathsf{tr}\left(\rho_{AB}E_{AB}\right). \tag{29}$$

In addition, if ρ_{AB} is classical on $\mathcal{H}_A$ relative to an orthonormal basis $\{|x\rangle\}_x$, then the maximization on the right-hand side of (29) can be further restricted to those operators E_{AB} which are classical on $\mathcal{H}_A$ relative to $\{|x\rangle\}_x$.

Proof: For a nonnegative operator E_{AB} with $\mathsf{tr}_A E_{AB} \leq \mathsf{id}_B$, we can define the operator

$$E'_{AB} = E_{AB} + \kappa_A \otimes (\mathsf{id}_B - \mathsf{tr}_A(E_{AB}))$$

where κ_A is an arbitrary normalized density operator on $\mathcal{H}_A$. We then have

$$\mathsf{tr}\left(\rho_{AB}E'_{AB}\right) \geq \mathsf{tr}(\rho_{AB}E_{AB})$$

with $E'_{AB} \geq 0$ and $\mathsf{tr}_A(E'_{AB}) = \mathsf{id}_B$. This shows that we can extend the maximization on the right-hand side of (29) to all

operators E_{AB} whose partial trace $\mathsf{tr}_A(E_{AB})$ is bounded by id_B (instead of being equal to id_B). The claim is therefore equivalent to

$$\min_{\substack{\sigma_B \geq 0 \\ \mathsf{id}_A \otimes \sigma_B \geq \rho_{AB}}} \mathsf{tr}(\sigma_B) = \max_{\substack{E_{AB} \geq 0 \\ \mathsf{tr}_A(E_{AB}) \leq \mathsf{id}_B}} \mathsf{tr}(\rho_{AB} E_{AB}). \tag{30}$$

To relate this to the general linear programming problem (25), we define $\mathcal{V}_1 = \mathsf{Herm}(\mathcal{H}_B)$ and $\mathcal{V}_2 = \mathsf{Herm}(\mathcal{H}_A \otimes \mathcal{H}_B)$ as the (real) vector spaces of Hermitian operators on $\mathcal{H}_B$ and $\mathcal{H}_A \otimes \mathcal{H}_B$, respectively, with standard Hilbert–Schmidt inner product. Furthermore, we define the convex cones K_1 and K_2 as the set of nonnegative operators in $\mathsf{Herm}(\mathcal{H}_B)$ and $\mathsf{Herm}(\mathcal{H}_A \otimes \mathcal{H}_B)$, respectively. We claim that these cones are self-dual, i.e., $K_1^* = K_1$ and $K_2^* = K_2$. This is easily seen from the spectral decomposition of a Hermitian operator. Finally, we define $E : \mathcal{V}_1 \to \mathcal{V}_2$ as the linear map $E(\theta_B) = E\theta_B := \mathsf{id}_A \otimes \theta_B$. It is easy to check that the adjoint $E^* : \mathcal{V}_2 \to \mathcal{V}_1$ is equal to the partial trace $\mathsf{tr}_A : \mathsf{Herm}(\mathcal{H}_A \otimes \mathcal{H}_B) \to \mathsf{Herm}(\mathcal{H}_B)$; indeed, for all $\theta_B \in \mathsf{Herm}(\mathcal{H}_B)$ and $F_{AB} \in \mathsf{Herm}(\mathcal{H}_A \otimes \mathcal{H}_B)$, we have

$$\begin{aligned} \langle E\theta_B, F_{AB} \rangle_2 &= \langle \mathsf{id}_A \otimes \theta_B, F_{AB} \rangle_2 \\ &= \mathsf{tr}((\mathsf{id}_A \otimes \theta_B) F_{AB}) \\ &= \mathsf{tr}(\theta_B \mathsf{tr}_A(F_{AB})) \\ &= \langle \theta_B, \mathsf{tr}_A(F_{AB}) \rangle_1. \end{aligned}$$

We also set $b = \rho_{AB}$ and $c = \mathsf{id}_B$. With these definitions, we conclude that the two optimization problems defined by (30) are a special instance of (25); the claim is equivalent to the statement that the duality gap vanishes. According to Lemma 1, it suffices to check Slater's interiority condition. For this purpose, we set $v = 2\lambda_{\max}(\rho_{AB}) \cdot \mathsf{id}_B$, where $\lambda_{\max}$ denotes the maximal eigenvalue. Clearly, v is in the interior of K_1. We also have

$$Ev - b = 2\lambda_{\max}(\rho_{AB}) \mathsf{id}_{AB} - \rho_{AB} > 0$$

hence $Ev - b \in \mathsf{int} K_2$; this proves the claim (30).

To prove the claim about the case where ρ_{AB} is classical relative to an orthonormal basis $\{|x\rangle\}_x$ of $\mathcal{H}_A$, we simply set $\mathcal{V}_2 = \mathrm{span}\{|x\rangle\langle x|\}_x \otimes \mathsf{Herm}(\mathcal{H}_B)$ equal to the set of Hermitian operators that are classical on $\mathcal{H}_A$. The remainder of the proof is identical to the general case. □

Observe that the left-hand side of (29) is equivalent to a minimization of the distance measure D_∞ from (4), i.e., we have

$$\begin{aligned} &\log \min_{\substack{\sigma_B \geq 0 \\ \mathsf{id}_A \otimes \sigma_B \geq \rho_{AB}}} \mathsf{tr}(\sigma_B) \\ &= \min_{\substack{\sigma_B \geq 0 \\ \mathsf{tr}(\sigma_B) = 1}} D_\infty(\rho_{AB} \| \mathsf{id}_A \otimes \sigma_B) \\ &= -H_{\min}(A|B)_\rho. \end{aligned} \tag{31}$$

Let us discuss the case where ρ_{XB} is classical on X. Lemma 4 allows us to show that the min-entropy $H_{\min}(X|B)_\rho$ is equivalent to the "guessing-entropy" of X given B.

Theorem 1: Let $\rho_{XB} = \sum_x p_x |x\rangle\langle x| \otimes \rho_B^x$ be classical on $\mathcal{H}_X$. Then

$$H_{\min}(X|B)_\rho = -\log p_{\mathsf{guess}}(X|B)_\rho$$

where $p_{\mathsf{guess}}(X|B)_\rho$ is the maximal probability of decoding X from B with a POVM $\{E_B^x\}_x$ on $\mathcal{H}_B$, i.e.,

$$p_{\mathsf{guess}}(X|B)_\rho := \max_{\{E_B^x\}_x} \sum_x p_x \mathsf{tr}(E_B^x \rho_B^x).$$

Proof: According to (31), it suffices to show that the right-hand side of (29) is equal to $p_{\mathsf{guess}}(X|B)_\rho$. But this is a direct consequence of the fact that every nonnegative operator E_{XB} with $\mathsf{tr}_X(E_{XB}) = \mathsf{id}_B$ which is classical on $\mathcal{H}_X$ has the form

$$E_{XB} = \sum_x |x\rangle\langle x| \otimes E_B^x$$

where the family $\{E_B^x\}_x$ is a POVM on $\mathcal{H}_B$. □

The Choi–Jamiołkowski isomorphism yields an operational interpretation of the min-entropy in the general case. We can express the min-entropy as the maximal achievable singlet fraction as follows.

Theorem 2: The min-entropy of a state ρ_{AB} on $\mathcal{H}_A \otimes \mathcal{H}_B$ can be expressed as

$$H_{\min}(A|B)_\rho = -\log q_{\mathsf{corr}}(A|B)_\rho \tag{32}$$

where $q_{\mathsf{corr}}(A|B)_\rho$ is the maximal achievable singlet fraction, i.e.,

$$q_{\mathsf{corr}}(A|B)_\rho := d_A \max_{\mathcal{F}} F\left((\mathsf{id}_A \otimes \mathcal{F})(\rho_{AB}), |\Phi_{AA'}\rangle\langle\Phi_{AA'}|\right)^2$$

with maximum taken over all quantum operations $\mathcal{F} : \mathcal{L}(\mathcal{H}_B) \to \mathcal{L}(\mathcal{H}_{A'}), \mathcal{H}_{A'} \cong \mathcal{H}_A$ and $|\Phi_{AA'}\rangle$ defined by (28).

Proof: Let us rewrite statement (32) as

$$\begin{aligned} &\min_{\substack{\sigma_B \geq 0 \\ \mathsf{tr}(\sigma_B) = 1}} D_\infty(\rho_{AB} \| \mathsf{id}_A \otimes \sigma_B) \\ &= \log\left(d_A \cdot \max_{\mathcal{F}} F\left((\mathsf{id}_A \otimes \mathcal{F})(\rho_{AB}), |\Phi_{AA'}\rangle\langle\Phi_{AA'}|\right)^2\right) \end{aligned} \tag{33}$$

where $|\Phi_{AA'}\rangle$ is the maximally entangled state. Let E_{AB} be a nonnegative operator on $\mathcal{H}_A \otimes \mathcal{H}_B$ with $\mathsf{tr}_A E_{AB} = \mathsf{id}_B$, and let $\mathcal{E} = J^{-1}(E_{AB}) \in \mathsf{CPUM}(\mathcal{H}_{A'}, \mathcal{H}_B)$ be the unital map corresponding to E_{AB} under the Choi–Jamiołkowsi isomorphism [cf., Lemma 3(ii)]. Let $\mathcal{F} = \mathcal{E}^\dagger : \mathcal{L}(\mathcal{H}_B) \to \mathcal{L}(\mathcal{H}_{A'})$ be the adjoint quantum operation (cf., Lemma 2). By definition of $\mathcal{E}$ and the adjoint $(\mathsf{id}_A \otimes \mathcal{E})^\dagger = \mathsf{id}_A \otimes \mathcal{E}^\dagger = \mathsf{id}_A \otimes \mathcal{F}$, we have

$$\begin{aligned} \mathsf{tr}(\rho_{AB} E_{AB}) &= d_A \mathsf{tr}\left(\rho_{AB}(\mathsf{id}_A \otimes \mathcal{E})(|\Phi_{AA'}\rangle\langle\Phi_{AA'}|)\right) \\ &= d_A \mathsf{tr}\left((\mathsf{id}_A \otimes \mathcal{E})^\dagger(\rho_{AB}) |\Phi_{AA'}\rangle\langle\Phi_{AA'}|\right) \\ &= d_A \mathsf{tr}\left((\mathsf{id}_A \otimes \mathcal{F})(\rho_{AB}) |\Phi_{AA'}\rangle\langle\Phi_{AA'}|\right). \end{aligned}$$

Observe that the operators $E_{AB} \geq 0$ with $\mathsf{tr}_A E_{AB} = \mathsf{id}_B$ are in one-to-one correspondence with quantum operations $\mathcal{F} \in \mathsf{CPTPM}(\mathcal{H}_B, \mathcal{H}_{A'})$ constructed in this fashion. The claim (33), therefore, follows from Lemma 4 and (31). □

Remark 1: The result of Theorem 2 can be extended to give an alternative expression for the maximal achievable fidelity with a nonmaximally entangled state $|\Psi_{AA'}\rangle = \sum_\lambda \sqrt{\lambda} |\lambda\rangle_A |\lambda\rangle_{A'} \in \mathcal{H}_A \otimes \mathcal{H}_{A'}$. We assume that $\mathcal{H}_{A'} \cong \mathcal{H}_A$ and that $|\Psi_{AA'}\rangle$ has

maximal Schmidt rank. Let $\tau_A = \mathrm{tr}_{A'}|\Phi_{AA'}\rangle\langle\Phi_{AA'}|$ be its reduced density operator. Then

$$\min_{\substack{\sigma_B \geq 0 \\ \mathrm{tr}(\sigma_B)=1}} D_\infty(\rho_{AB}\|\tau_A^{-1}\otimes\sigma_B) = \log \max_{\mathcal{F}} F\left((\mathsf{id}_A\otimes\mathcal{F})(\rho_{AB}), |\Psi_{AA'}\rangle\langle\Psi_{AA'}|\right)^2 \quad (34)$$

for any bipartite state ρ_{AB} on $\mathcal{H}_A\otimes\mathcal{H}_B$. Statement (34) follows by substituting $(\tau_A^{1/2}\otimes\mathsf{id}_B)\rho_{AB}(\tau_A^{1/2}\otimes\mathsf{id}_B)$ for ρ_{AB} in (33), using the fact that conjugating with an invertible matrix does not change operator inequalities, and $\sqrt{d_A}\cdot(\tau_A^{1/2}\otimes\mathsf{id}_B)|\Phi_{AA'}\rangle = |\Psi_{AA'}\rangle$.

B. Proof of the Operational Characterization of $H_{\max}$

To obtain the operational characterization of $H_{\max}$, we use Theorem 2. Recall the definition of the decoupling accuracy of a bipartite state $\rho = \rho_{AB}$, that is

$$q_{\mathsf{decpl}}(A|B)_\rho := d_A \max_{\sigma_B} F(\rho_{AB}, \tau_A\otimes\sigma_B)^2$$

where d_A is the dimension of $\mathcal{H}_A$ and τ_A is the completely mixed state on $\mathcal{H}_A$. We begin by showing the following lower bound on the decoupling accuracy.

Lemma 5: For all bipartite states ρ_{AB}, we have

$$2^{H_{\max}(A|B)_\rho} \leq q_{\mathsf{decpl}}(A|B)_\rho.$$

Proof: Let $\rho_{ABC} = |\varphi_{ABC}\rangle\langle\varphi_{ABC}|$ be a purification of ρ_{AB}, and let $\mathcal{F}: \mathcal{L}(\mathcal{H}_C)\to\mathcal{L}(\mathcal{H}_{A'})$ be a quantum operation that satisfies (cf., Theorem 2)

$$2^{-H_{\min}(A|C)} = d_A F((\mathsf{id}_A\otimes\mathcal{F})(\rho_{AC}), |\Phi_{AA'}\rangle)^2.$$

Let $\rho'_{AA'BR} = |\varphi'_{AA'BR}\rangle\langle\varphi'_{AA'BR}|$ be a purification of $\rho'_{ABA'} = (\mathsf{id}_{AB}\otimes\mathcal{F})(\rho_{ABC})$. We then have

$$2^{H_{\max}(A|B)} = 2^{-H_{\min}(A|C)} = d_A F(\rho'_{AA'}, |\Phi_{AA'}\rangle)^2. \quad (35)$$

However

$$F(\rho'_{AA'}, |\Phi_{AA'}\rangle) = F(|\varphi'_{AA'BR}\rangle, |\Phi_{AA'}\rangle\otimes|\theta_{BR}\rangle)$$

for some state $|\theta_{BR}\rangle$ on $\mathcal{H}_B\otimes\mathcal{H}_R$. By the monotonicity of the fidelity, we therefore get

$$F(\rho'_{AA'}, |\Phi_{AA'}\rangle) \leq F(\rho_{AB}, \tau_A\otimes\mathrm{tr}_R|\theta_{BR}\rangle\langle\theta_{BR}|) \leq \max_{\sigma_B} F(\rho_{AB}, \tau_A\otimes\sigma_B)$$

where we used the fact that $\rho'_{AB} = \rho_{AB}$. Inserting this into (35) gives the claim. □

The proof of the converse inequality closely follows a derivation in [19]. We include it here for completeness.

Lemma 6: For all bipartite states ρ_{AB}, we have

$$2^{H_{\max}(A|B)_\rho} \geq q_{\mathsf{decpl}}(A|B)_\rho.$$

Proof: We use the following fact, which is a consequence of the fact that all purifications of a fixed state are related by a unitary transformation on a (possibly extended) ancilla. If $|\phi_{ABCC'}\rangle$ has a reduced state of the form $\mathrm{tr}_{CC'}|\phi_{ABCC'}\rangle\langle\phi_{ABCC'}| = \tau_A\otimes\sigma_B$, where τ_A is the completely mixed state on $\mathcal{H}_A$, then there exists a unitary $U_{CC'}$ such that

$$(\mathsf{id}_{AB}\otimes U_{CC'})|\phi_{ABCC'}\rangle = |\Phi_{AC}\rangle|\theta_{BC'}\rangle \quad (36)$$

for some state $|\theta_{BC'}\rangle$ on $\mathcal{H}_B\otimes\mathcal{H}_{C'}$, where $|\Phi_{AC}\rangle$ denotes the fully entangled state on $\mathcal{H}_A\otimes\mathcal{H}_C$ (without loss of generality, we can assume that $d_A\leq d_C$).

Let σ_B be an arbitrary density matrix on $\mathcal{H}_B$. Let $\rho_{ABC} = |\psi_{ABC}\rangle\langle\psi_{ABC}|$ be a purification of ρ_{AB}, where we assume the dimension of $\mathcal{H}_C$ to be sufficiently large.

According to the definition of the fidelity, there exists a purification $|\phi_{ABCC'}\rangle$ of $\tau_A\otimes\sigma_B$ such that

$$F(\rho_{AB}, \tau_A\otimes\sigma_B) = F(|\psi_{ABC}\rangle|0_{C'}\rangle, |\phi_{ABCC'}\rangle).$$

Applying the unitary $U_{CC'}$ from (36) gives

$$F(\rho_{AB}, \tau_A\otimes\sigma_B) = F(|\psi'_{ABCC'}\rangle, |\Phi_{AC}\rangle|\theta_{BC'}\rangle)$$

where $|\psi'_{ABCC'}\rangle = (\mathsf{id}_{AB}\otimes U_{CC'})|\psi_{ABC}\rangle|0_{C'}\rangle$ because of the invariance of the fidelity under unitary operations. Using the monotonicity of the fidelity, we conclude that

$$F(\rho_{AB}, \tau_A\otimes\sigma_B) \leq F\left(\mathrm{tr}_{BC'}|\psi'_{ABCC'}\rangle\langle\psi'_{ABCC'}|, |\Phi_{AC}\rangle\right) = F((\mathsf{id}_A\otimes\mathcal{F})(\rho_{AC}), |\Phi_{AC}\rangle)$$

where $\mathcal{F}: \mathcal{L}(\mathcal{H}_C)\to\mathcal{L}(\mathcal{H}_C)$ is the quantum operation $\mathcal{F}(\rho) = \mathrm{tr}_{C'}(U_{CC'}(\rho\otimes|0\rangle\langle 0|_{C'})U_{CC'}^\dagger)$. Squaring both sides of the previous inequality, multiplying by d_A, taking the maximum over all quantum operations, and using Theorem 2 therefore gives

$$d_A F(\rho_{AB}, \tau_A\otimes\sigma_B)^2 \leq 2^{-H_{\min}(A|C)_\rho}.$$

Since σ_B was arbitrary, we can maximize the left-hand side over all σ_B. The claim then follows from the definitions of $q_{\mathsf{decpl}}(A|B)_\rho$ and $H_{\max}(A|B)_\rho$. □

In summary, we have shown the following result.

Theorem 3: Let ρ_{AB} be a state on $\mathcal{H}_A\otimes\mathcal{H}_B$, and let τ_A be the completely mixed state on $\mathcal{H}_A$. Then

$$H_{\max}(A|B)_\rho = \log q_{\mathsf{decpl}}(A|B)_\rho$$

where $q_{\mathsf{decpl}}(A|B)_\rho$ is the decoupling accuracy, defined by

$$q_{\mathsf{decpl}}(A|B)_\rho := d_A \max_{\sigma_B} F(\rho_{AB}, \tau_A\otimes\sigma_B)^2$$

with the maximum taken over all normalized states σ_B on $\mathcal{H}_B$.

IV. Conclusion

In information theory, entropies are generally interpreted as *measures of uncertainty*. One method to make this interpretation more precise is to establish relations between entropy measures

TABLE I
OPERATIONAL INTERPRETATIONS OF (SMOOTH) MIN- AND MAX-ENTROPIES. THE APPROXIMATION ($\approx$) INDICATES THAT EQUALITY HOLDS UP TO AN ADDITIVE TERM OF ORDER $\log \frac{1}{\varepsilon}$ AND FOR AN APPROPRIATE CHOICE OF THE SMOOTHNESS PARAMETER ε'

goal (extremal state)	quality of a state ρ_{AB} (measured in terms of overlap)	amount of extremal states contained in a state ρ_{AB} (measured in # of qubits on A)
A fully entangled with B classical A fully determined by B	$\left.\begin{array}{l} -\log q_{\mathsf{corr}}(A\|B) \\ -\log p_{\mathsf{guess}}(A\|B) \end{array}\right\} = H_{\min}(A\|B)$	$-\ell^{\varepsilon}_{\mathsf{merg}}(A\|B) \overset{(16)}{\approx} -H^{\varepsilon'}_{\max}(A\|B)$
A fully mixed and indep. of B classical A uniform and indep. of B	$\left.\begin{array}{l} \log q_{\mathsf{decpl}}(A\|B) \\ \log p_{\mathsf{secr}}(A\|B) \end{array}\right\} = H_{\max}(A\|B)$	$\ell^{\varepsilon}_{\mathsf{decpl}}(A\|B) \overset{(15)}{\approx} H^{\varepsilon'}_{\min}(A\|B)$

and *operational quantities*, that is, quantities that characterize actual information-theoretic tasks.

Here, we consider a general scenario consisting of a (possibly quantum-mechanical) system A as well as an observer with (quantum or classical) *side information* B. The uncertainty of the observer about the state of system A then depends on the distribution of these states as well as the correlation between A and B.

There are two extreme situations, namely, when A is completely undetermined and when A is determined. Taking into account the side information B, these two situations are described as follows.

1) The state of A is fully correlated with (parts of) B.[16]
2) The state of A is uniformly distributed and independent of the side information B.

Note that in the first case, the requirement is merely that A is correlated with *parts* of B. This is because the side information B may consist of additional information that is unrelated to A.

For any given state ρ_{AB}, we may characterize the uncertainty of A given B by the distance to these extreme situations. If we take as a distance measure the *overlap* (i.e., the square of the *fidelity*), we retrieve the definitions of $q_{\mathsf{corr}}(A|B)$ and $q_{\mathsf{decpl}}(A|B)$ [see (20) and (22), respectively]. Our main results imply that these correspond to $H_{\min}(A|B)$ and $H_{\max}(A|B)$, respectively. We thus conclude that $H_{\min}(A|B)$ quantifies the closeness to a situation where A is determined by B, and, likewise, $H_{\max}(A|B)$ corresponds to the closeness to a situation where A is independent of B (see second column of Table I).

Given a bipartite state ρ_{AB}, we may also ask for the number of maximally entangled or completely independent qubits one can extract from A. Very roughly speaking, this is the idea underlying the definitions of $\ell^{\varepsilon}_{\mathsf{merg}}(A|B)$ and $\ell^{\varepsilon}_{\mathsf{decpl}}(A|B)$, respectively (see Section I-B for more details, in particular, the interpretation of negative quantities). Remarkably, these quantities are (approximately) given by the smooth entropies $H^{\varepsilon'}_{\max}(A|B)$ and $H^{\varepsilon'}_{\min}(A|B)$ (see last column of Table I).[17]

Despite these similarities between the (previously known) operational interpretations summarized in the last column of Table I and those given in the second column (the ones derived here), there are at least two fundamental differences. The first is that the new interpretations are exact and, in particular, valid without a smoothness parameter. In contrast, all previously established interpretations only hold up to additive terms of the order $\log \frac{1}{\varepsilon}$, where ε is a smoothness parameter (whose meaning is that of an error or failure probability). A second difference is that there does not seem to exist an obvious asymptotic counterpart for our identities. In particular, there are no analogous operational interpretations of the von Neumann entropy.

The results of this paper suggest that studying operationally defined quantities may be a viable approach to identifying relevant single-shot information measures in a multipartite setting. Of particular interest is the conditional mutual information, which has only recently been given an asymptotic interpretation [25].

[16]In the general case where A and B are quantum-mechanical systems, full correlation is akin to maximal entanglement.

[17]Note that compared to the discussion of the distance, the role of $\max$ and $\min$ is interchanged.

REFERENCES

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

[2] S. Verdú and T. S. Han, "A general formula for channel capacity," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1147–1157, Jul. 1994.

[3] R. Renner and S. Wolf, "Smooth Rényi entropy and applications," in *Proc. Int. Symp. Inf. Theory*, 2004, p. 233.

[4] N. Datta and R. Renner, "Smooth Rényi entropies and the quantum information spectrum," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2807–2815, Jun. 2009.

[5] M. Tomamichel, R. Colbeck, and R. Renner, "A fully quantum asymptotic equipartition property," 2008 [Online]. Available: arXiv:0811.1221

[6] R. Renner, "Security of quantum key distribution," Ph.D. dissertation, ETH Zurich, Zurich, Switzerland, 2005 [Online]. Available: arXiv.org:quant-ph/0512258

[7] R. Renner, S. Wolf, and J. Wullschleger, "The single-serving channel capacity," in *Proc. IEEE Int. Symp. Inf. Theory*, 2006, pp. 1424–1427.

[8] C. H. Bennett, G. Brassard, C. Crépeau, and U. Maurer, "Generalized privacy amplification," *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1915–1923, Nov. 1995.

[9] R. König and R. Renner, "Sampling of min-entropy relative to quantum knowledge," 2007 [Online]. Available: arXiv.org:0712.4291

[10] R. Impagliazzo, L. A. Levin, and M. Luby, "Pseudo-random generation from one-way functions," in *Proc. 21st Annu. ACM Symp. Theory Comput.*, New York, NY, 1989, pp. 12–24.

[11] M. Horodecki, J. Oppenheim, and A. Winter, "Quantum state merging and negative information," 2005 [Online]. Available: arXiv.org:quant-ph/0512247

[12] A. Winter and R. Renner, "Single-shot state merging," 2007, unpublished.

[13] M. Berta, "Single-shot quantum state merging," M.S. thesis, ETH Zurich, Zurich, Switzerland, 2008.

[14] A. S. Holevo, "Statistical decision theory for quantum systems," *J. Multivar. Anal.*, vol. 3, no. 4, pp. 337–394, 1973.
[15] H. P. Yuen, R. S. Kennedy, and M. Lax, "Optimum testing of multiple hypotheses in quantum detection theory," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 2, pp. 125–134, Mar. 1975.
[16] M. Christandl and R. Renner, "Guessing and min-entropy," 2006, unpublished.
[17] H. Buhrman, M. Christandl, P. Hayden, H.-K. Lo, and S. Wehner, "Security of quantum bit string commitment depends on the information measure," *Phys. Rev. Lett.*, vol. 97, no. 25, 2006, 250501.
[18] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, and W. K. Wootters, "Classical information capacity of a quantum channel," *Phys. Rev. A, Gen. Phys.*, vol. 54, no. 3, pp. 1869–1876, Sep. 1996.
[19] M. Horodecki, S. Lloyd, and A. Winter, "Quantum coding theorem from privacy and distinguishability," *Open Syst. Inf. Dyn.*, vol. 15, pp. 47–, 2008 [Online]. Available: arXiv.org:quant-ph/0702006
[20] I. B. Damgård, S. Fehr, R. Renner, L. Salvail, and C. Schaffner, "A tight high-order entropic quantum uncertainty relation with applications," in *Lecture Notes in Computer Science*. Berlin, Germany: Springer-Verlag, 2007, vol. 4622, pp. 360–378.
[21] S. Wehner, C. Schaffner, and B. M. Terhal, "Cryptography from noisy storage," *Phys. Rev. Lett.*, vol. 100, no. 22, 2008, 220502.
[22] C. Schaffner, B. Terhal, and S. Wehner, "Robust cryptography in the noisy-quantum-storage model," 2008 [Online]. Available: arxiv.org:0807.1333
[23] A. Barvinok, *A Course in Convexity*, ser. Graduate Studies in Mathematics. Providence, RI: AMS, 2002, 54.
[24] A. Jamiołkowski, "Linear transformations which preserve trace and positive semidefiniteness of operators," *Rev. Mod. Phys.*, vol. 3, pp. 275–278, 1972.
[25] I. Devetak and J. Yard, "The operational meaning of quantum conditional information," 2006 [Online]. Available: arXiv:quant-ph/0612050

Robert König received a diploma in theoretical physics from ETH Zurich, Switzerland, in 2003. He subsequently worked as a Research Assistant in the cryptography and information security group at ETH. In 2005, he moved to the Department of Applied Mathematics and Theoretical Physics of the University of Cambridge, U.K., where he received a Ph.D. in 2007.

Currently, he is a Postdoctoral Scholar at the Institute for Quantum Information, California Institute of Technology, Pasadena.

Renato Renner was born on December 11, 1974, in Lucerne, Switzerland. He received a degree in theoretical physics from ETH Zurich, Switzerland, in 2000 and the Ph.D. degree from the Computer Science Department, ETH Zurich, in 2005, working on a thesis in the area of quantum cryptography.

Between 2005 and 2007, he held an HP research fellowship in the Department for Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, U.K. Since 2007, he has been an Assistant Professor in Theoretical Physics at ETH Zurich. His research interests are ranging from quantum information science to foundations of quantum mechanics.

Prof. Renner is a member of the American Physical Society. He has been a member of the technical program committee of the 2008 IEEE International Symposium on Information Theory (ISIT).

Christian Schaffner received the diploma degree in mathematics from ETH Zurich, Switzerland, in 2003 and the Ph.D. degree in computer science from Århus University, Århus, Denmark, in 2007.

Currently, he holds a postdoctoral position at Centrum Wiskunde & Informatica (CWI), Amsterdam, The Netherlands. His research interests include quantum cryptography, cryptographic protocols, and (quantum) information theory.