Foundations of Physics, Vol. 35, No. 12, December 2005 (© 2005)
DOI: 10.1007/s10701-005-8660-5

Information Theories with Adversaries, Intrinsic Information, and Entanglement*

Karol Horodecki,[1] Michał Horodecki,[2] Pawel Horodecki,[3] and Jonathan Oppenheim[4]

Received July 6, 2005

There are aspects of privacy theory that are analogous to quantum theory. In particular one can define distillable key and key cost in parallel to distillable entanglement and entanglement cost. We present here classical privacy theory as a particular case of information theory with adversaries, where similar general laws hold as in entanglement theory. We place the result of Renner and Wolf—that intrinsic information is lower bound for key cost—into this general formalism. Then we show that the question of whether intrinsic information is equal to key cost is equivalent to the question of whether Alice and Bob can create a distribution product with Eve using I_M bits of secret key. We also propose a natural analogue of relative entropy of entanglement in privacy theory and show that it is equal to the intrinsic information. We also provide a formula analogous to the entanglement of formation for classical distributions.

KEY WORDS: Classical privacy; quantum entanglement; intrinsic information; key cost; relative entropy distance.

1. INTRODUCTION

There is a deep connection between quantum entanglement and classical privacy (see e.g. Refs. 1, 2). In entanglement theory, one of the basic questions is how many singlets Alice and Bob can draw from quantum state

[*] It is our pleasure to dedicate this paper to Asher Peres on the occasion of his seventieth birthday.

[1] Department of Mathematics, Physics and Computer Science, University of Gdańsk, Poland.
[2] Institute of Theoretical Physics and Astrophysics, University of Gdańsk, 80-952 Gdańsk, Poland; e-mail: michalh@iftia.univ.gda.pl
[3] Faculty of Applied Physics and Mathematics, Technical University of Gdańsk, 80-952 Gdańsk, Poland.
[4] Dept. of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, UK.

ϱ_{AB} by use of *local operations and classical communication*, while in privacy theory one asks how many shared random secret bits (private key) can Alice and Bob obtain from probability distribution p_{ABE} shared with an adversary Eve, by use of *local operations and public communication*. The protocols of distilling key were used to build protocols of distillation of singlets.[3] The teleportation[4] via singlets can be thought as analogue of performing one-time pad by using secret key.[2] Recently more and more connections between entanglement and privacy have been obtained (see e.g. Refs. 5,6–9). In particular, the ideas from entanglement have been transferred into privacy theory.[6,9] The notion of key cost,[10] i.e. number of bits of private key, needed to create a probability distribution p_{ABE} was introduced.[1] Renner and Wolf[6] showed that there is irreversibility between distillable key and key cost, similar to that in entanglement theory. In particular they proved that a function called the intrinsic information I_{intr} is a lower bound for key cost.

In this paper we develop connections between entanglement and classical privacy theory. The basic question is whether we can have a result equivalent to Ref. 11. There, the following was shown. Consider the entanglement of formation of a quantum state with density matrix ρ_{AB} held between two parties (Alice and Bob) and defined by

$$E_F = \inf \sum_i p_i S(Tr_A |\psi_i\rangle\langle\psi_i|), \tag{1}$$

where the *infimum* is taken over all decompositions of

$$\rho_{AB} = \sum_i p_i |\psi_i\rangle\langle\psi_i|, \tag{2}$$

and S is the Von Neumann entropy and Tr_A is partial trace over Alice's system. Now consider the regularization of this function, i.e.

$$E_c = \lim_{n\to\infty} \frac{E_F(\rho_{AB}^{\otimes n})}{n}. \tag{3}$$

Then, it was shown that this regularized entanglement of formation is equal to the entanglement cost of creating the state out of singlets. The result in entanglement theory was based on two ingredients: (1) regularized E_F [12] is lower bound for the entanglement cost, (2) pure states can be transformed reversibly into each other. Result (1) comes essentially from a general relation between so called *monotones* and *transitions rates*. Ingredient (2) was obtained [13] by constructing protocols that transform reversibly pure states into singlets (then trivially, going through singlet, any pure state can be reversibly transformed into any other one). It implies that E_F is also an upper bound of entanglement cost.

In privacy theory, the question translates to: is the intrinsic information $I_{\rm intr}$ of the probability distribution p_{ABE} equal to the cost of producing such distribution by Alice and Bob out of ideal private key? (regularization is not needed because $I_{\rm intr}$ is additive). That $I_{\rm intr}$ is a lower bound for key cost was proven in Ref. 6.

In this paper we show that the following analogue of result (2) will give the converse (i.e. that $I_{\rm intr}$ is *upper* bound for key cost). Namely, one should show that Alice and Bob can transform reversibly any distribution *product with Eve* into any other distribution product with Eve. We point out that one ingredient of the proof of this result is already known: one knows that out of distribution product with Eve, Alice and Bob can produce an amount of private key equal to the mutual information of the key. We then conclude that to prove or disprove the equality between $I_{\rm intr}$ and key cost, one should show that to prepare a distribution product with Eve, it is enough to share an amount of key equal to the mutual information of the distribution (or prove that this is impossible).

Note added: Very recently, a forgotten paper of Wyner[14] has been discovered through the archeological work of Andreas Winter. There, a single letter formula for the *common information* was given which can be strictly larger than the mutual information. Shortly after it's discovery, it was abandoned, the thinking presumably being that distillable common randomness was a more natural definition of common information. Wyner's function between two distributions X and Y is

$$C(X:Y) \equiv \inf_W I(X,Y:W), \tag{4}$$

where the infimum is taken over W in the Markov chain $X \to W \to Y$. This can thus be interpreted as the key cost of creating a distribution product with Eve. This result allows us to complete many of the unanswered questions originally posed here. Throughout our paper, one can replace our function $J(X,Y)$ with that of Wyner. This implies an explicit formula for the key cost of distributions X, Y, Z with the variable Z belonging to the adversary (Proposition 1, with a lower bound following straight-forwardly from monotonicity of the function). This function can be strictly greater than the intrinsic information.

We also discuss the result of Renner and Wolf giving a lower bound for key cost and put it into a general formalism which is a natural modification of the one in entanglement theory. Namely, we show that similarly as in the case of entanglement cost, one can use here a general relation between monotones and transition rates. The suitable modification of the scheme from entanglement theory was provided in Ref. 15. However, the authors didn't consider conditions that are essential in asymptotic rate transitions: regularization and asymptotic continuity. Here we fill this gap. Moreover we generalize the scheme so that it become independent of privacy theory. What we obtain is *information theory with adversaries*. We restrict ourselves to the case of two adversaries. Each of them have their own class of operations. What is resource for one of them is not for other one and vice versa.

Finally, seeking for other potential advantages from the analogy between privacy and entanglement, we define the counterpart of relative entropy of entanglement.(16) We then show that unlike in entanglement theory, it is equal to the intrinsic information.

2. INFORMATION THEORIES: THE ROLE OF MONOTONES

To begin with, let us note that what we are dealing with are various *information theories*. Essentially, a specific information theory is determined by a class of operations. The most important notion is *transition rate* between states optimized over the class of operations. Other important notions are *monotones*: functions that cannot increase under the class of operations. An example is entanglement theory, where monotones are entanglement measures, transition rates are, in particular, entanglement of formation and entanglement of distillation (number of singlets that one can obtain from a given state). Of course, another example is also the "mother theory" i.e. Shannon's one, where the central theorem tells us about the optimal simulation of a noiseless channel by a noisy one. The role of states is played by channels; the transition—by *simulating* one channel with another one; the operations are local ones (coding and decoding). In cryptography there are two classes of operations: one is the class of operations of trusted parties Alice and Bob, and the second by their adversary Eve. Still one can design a similar formalism, following Ref. 15. Such theories we will call *information theories with adversary*.

Monotones in information processing. In Ref. 17 following(18,19) a general paradigm was formulated concerning asymptotic rates of transferring states into other states by means of restricted classes of manipulations. The main notion is asymptotic rate of transition. Given a class of operations, one can ask, at what rate it is possible to transform state ρ into σ, given large n independent copies of ρ,

$$\rho^{\otimes n} \to \sigma^{\otimes m}. \tag{5}$$

The above means that acting on $\rho^{\otimes n}$ by allowed operations, we get some state σ_m that is close to the desired state $\sigma^{\otimes m}$ for large n. The optimal rate $R(\rho \to \sigma)$ is defined as infimum of m/n, where we take the limit of large n.

Such rates are limited by the amount of *resources* contained in the source and target states. By *resource* one means any quantity that cannot be increased by the allowed operations. Clearly, one cannot have such a rate that would increase a resource, because in definition of rate, only allowed operations can be used, which by definition of resource cannot increase it. Mathematically, a resource is described by a monotone, i.e. a function of state that cannot increase upon acting by allowed operations

$$M(\Lambda(\rho)) \leqslant M(\rho) \tag{6}$$

for any allowed operation Λ and state ρ. The statement, that transition rates are limited by amount of resources present in the states, was made more precise for asymptotic transition rates in Ref. 17 (see also Ref. 20). We have

Theorem 1. (*Central inequality of information theories*).

For any function M, which satisfies (1) asymptotic continuity (2) monotonicity, one obtains

$$R(\rho \to \sigma) \leqslant \frac{M^\infty(\rho)}{M^\infty(\sigma)} \tag{7}$$

for $M^\infty(\sigma) \neq 0$, where $M^\infty(\rho) = \limsup \frac{1}{n} M(\rho^{\otimes n})$. Asymptotic continuity means here that if $||\gamma_n - \tilde{\gamma}_n|| \to 0$ then $|M(\gamma_n) - M(\tilde{\gamma}_n)|/\log \dim \mathcal{H}_n \to 0$, where states $\gamma_n, \tilde{\gamma}_n$ act on Hilbert space $\mathcal{H}_n$.

The above statement says no more, but that resource M cannot be increased by transitions. The initial amount of the resource per input copy is $M^\infty(\rho)$ and the final amount is just $R(\rho \to \sigma) M^\infty(\sigma)$.

The above formalism originates from entanglement theory, where the class is local operations and classical communication. As argued in Ref. 17 the above result does not rely on any feature of entanglement. It can be applied to any situation. In particular it worked in situation, where the resource was local information,[21] or just information.[22]

3. INFORMATION THEORIES WITH ADVERSARIES

The paradigm also applies to classical theories, by taking the class of operations that can be performed classically, and restricting input states to those diagonal in a fixed basis. An example is the classical theory of privacy manipulations.

There is however a slight difference between classical privacy theory and the general scheme mentioned in the previous section. Namely, in the latter there is no malicious adversary. The needed modification was done by Ref. 15. We will now shortly recall that paradigm, modifying it to be suitable in the asymptotic case (i.e. we put regularization and asymptotic continuity into the game). Also, we extract an abstract feature of the scheme, to make it completely general, rather than referring only to privacy. Simply, we consider any situation where there are adversaries, so that their operations are assumed to be directed against each other.

As noted, in standard theory, there is one allowable class of operations, and one wants to do ones best by using it. The state cannot change, if we will not apply an operation. In classical privacy theory, we have at least two actors that play against each other. We will provide a description from the point of view of one of them—call him X . His adversary is traditionally called Eve. The latter has her own allowed class of operations, and *can change state* even though X does not apply *any* operation. Thus, in particular we have to redefine the notion of state transition. But first let us ask, what is resource for *X* in such situation? Of course, we again have that

(i) Resource cannot *increase* under X operations
However it is not the end. Second postulate is needed:

(ii) Resource cannot *decrease* under operation of adversary Eve

It is easy to see that the last postulate is reasonable: it assumes that an adversary is as malicious as can be, and always acts optimally. So, if she has any possibility to decrease X 's abilities, she will do this. Thus something, that can be destroyed by her, cannot be treated as a resource by trusted parties. Consequently in the scenario with an adversary, we have a double postulate of monotonicity. The *monotones* will now be called any functions of state, that satisfy (i) and (ii).

Having modified monotonicity, let us now consider the definition of rate.[6,15] In the scenario with adversaries, a definition of rate transition as layed out in the introduction does not make sense: X cannot obtain any state he wishes, because there is an adversary, who can change the state. Thus the very notion of transition does not make sense as it is. Therefore

X should not aim to obtain some fixed target state, but rather, to obtain state that is in a sense *no worse* than the required state. We can say that state ρ is no worse than ρ', if there exists operation Λ of adversary Eve, such that $\Lambda(\rho') = \rho$. The state ρ is now indeed no worse than ρ', as the adversary can only increase the resources.

We can now define asymptotic rate of transition $R(\rho \to \sigma)$ as follows: Take n copies of ρ. X acts on them, and gets some state σ'_m, which is no worse than σ''_m, which in turn is asymptotically close to $\sigma^{\otimes m}$. The rate is again given by supremum of m/n over such protocols in the limit of large n.

Having defined both rates and monotones, we can now prove the following

Theorem 2. (*Central inequality of information theories with adversaries*). The formula (7) holds for information theories with adversary.

Proof. Let us start with n copies of ρ and check what happens to a chosen monotone M during any protocol that realizes transition between ρ and σ. First due to monotonicity (i) we have

$$M(\rho^{\otimes n}) \geqslant M(\sigma'_m). \tag{8}$$

Now, because the obtained state σ'_m is "not worse" than σ''_m, i.e. $\sigma''_m = \Lambda_E(\sigma'_m)$ for some Eve operation Λ_E, we can use monotonicity (ii) we get

$$M(\sigma'_m) \geqslant M(\sigma''_m). \tag{9}$$

From these we get

$$\frac{1}{n} M(\rho^{\otimes n}) \geqslant \frac{M(\sigma''_m)}{m} \frac{m}{n}. \tag{10}$$

Due to asymptotic continuity we have

$$\frac{M(\sigma''_m)}{m} \approx \frac{M(\sigma^{\otimes m})}{m}. \tag{11}$$

Since in optimal protocol m/n tends to rate we get

$$M^{\infty}(\rho) \geqslant R(\rho \to \sigma) M^{\infty}(\sigma). \tag{12}$$

This ends the proof. □

Remark. In classical theory, the regularization will usually not matter: the quantities are most often additive. Two examples are: maximal mutual information over a channel, and intrinsic information.

Corollary 1. For additive monotones, satisfying $M(\sigma) = 1$ we have $R(\rho \to \sigma) \leqslant M(\rho)$ and for those satisfying $M(\rho) = 1$ we have $R(\rho \to \sigma) \geqslant M(\sigma)$.

It should be noted here, that in this abstract approach, roles of Eve and X are symmetric. So far we considered transition rates and monotones from the point of view of X. We could reverse the reasoning, and obtain the same theorem for Eve, i.e. her optimal transition rates are bounded by her monotones.

4. CLASSICAL PRIVACY THEORY

An example of theory with adversaries is classical privacy theory. There the role of X is played by Alice and Bob and the role of Eve is just played by the eavesdropper Eve. (Let us recall here, that since the theory with adversaries is completely symmetric between adversaries, so we could reverse the problem, and analyze resources of Eve. We do not consider this problem here). The operations of Alice and Bob are local operations and public communication. Local operations transform the distribution as usual. The communication acts on the distribution in such a way that apart from obvious consequence which is modification of Alice and Bob distribution, any bit of communication is copied and the copy is added to Eve 's system. Though one might think that this is Eve's operations, we treat it as an element of Alice and Bob operations, because it is automatically associated with any of Alice and Bob's public communication.

Transitions between distributions product with Eve. In entanglement theory, it is proven that there is asymptotic reversibility for pure states.[13] Any pure state can be reversibly transformed into the two qubit singlet-like state,

$$|\psi_+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \tag{13}$$

with rate S_A (entropy of either of the subsystems). The potential analogue of pure states in privacy theory are probability distributions product against the $AB : E$ cut. We call them *private distributions*. We do not know whether transitions between private distributions can be made reversible. What is at present known is that for any private distribution of random variables X, Y shared by Alice and Bob, they can transform it reversibly into $I(X : Y)$ bits of perfect key. Here $I(X, Y)$ is Shannon mutual information

$$I(X, Y) = H(X) + H(Y) - H(X, Y) \equiv H(X) - H(Y|X), \tag{14}$$

where $H(X)$ is Shannon entropy of probability distribution of random variable X, and $H(X, Y)$ is entropy of joint probability distribution; $H(Y|X) = H(Y) - H(X, Y)$ is called conditional entropy.

To get key, Alice performs part of the so called Slepian–Wolf compression,[23] sending the information about X (while holding a copy of it X') to Bob by use of $H(X|Y)$ bits, so that Bob has now X on his side. The mutual information between Alice and Bob is now $H(X)$ (X and X' are perfectly correlated), while the mutual information between Alice and Eve is no more than $H(X|Y)$. Thus by Refs. 24, 25 the can distill $I(X, Y)$ bits of key.

The converse would be possible, if for example we had a scheme of visible compression of information carried by probability distributions that achieve the lower bound given by the Holevo quantity obtained in Ref. 26. However until now it is not known whether such a scheme exists.

To see how it would work, suppose that we have the following result on compression of information from sources with mixed signal states. Consider a source at Alice's site emitting with probability $p(x)$ a distribution $q_x(y)$. The task is to reproduce asymptotically faithfully $q_x(y)$ at Bob's site, using a minimal amount of bits. One can imagine the following

Conjecture 1. The minimal number of bits is equal to $I_{\text{compr}} = I(X, Y)$ where X and Y are distributed according to $p(x, y) = p(x)q_x(y)$.

More precisely, the conjecture states that in the asymptotic regime of long sequences, when $R = I_{\text{compr}} + \epsilon$ bits are used with $I_{\text{compr}} = I(X, Y)$, the obtained distributions $\tilde{q}_x(y)$ would satisfy high fidelity criterion

$$\sum_x p(x) \sum_y \sqrt{q_x(y)}\sqrt{\tilde{q}_x(y)} \to 1. \tag{15}$$

This problem was developed in Ref. 27. There are good reasons to believe that this conjecture is false, and we hope to address this in a future work. However, basing on the conjecture one can easily provide the protocol for formation of (X,Y) with arbitrary distribution $p(x,y)$. Namely, Alice picks at random x with probability $p(x)$ and by use of $I(X,Y)$ bits reproduces at Bob's side $q_x(y)$ which she chooses to be $p(y|x)$. From (15) it then follows that

$$\sum_{xy}\sqrt{p(x,y)}\sqrt{\tilde{p}(x,y)}=\sum_{x}p(x)\sum_{y}\sqrt{q(y|x)}\sqrt{\tilde{q}(y|x}\to 1. \tag{16}$$

We should mentioned here that one can achieve the bound of conjecture if additional resource is allowed: shared random bits.[28] However from our point of view this is useless, because the additional shared randomness would be correlated with Eve, and the produced final distribution would not be product with Eve.

4.1. Intrinsic Information as Analogue of Entanglement of Formation

Entanglement of formation E_F of a state ϱ_{AB} can be defined as follows. Consider purification ψ_{ABC} of the state ϱ_{AB}. Let Charlie perform such measurement on his system, that for any outcome i, Alice and Bob share some pure state ψ_i^{AB}. When he tells them honestly the outcome of his measurement, they will share on average $\sum_i p_i E(\psi_i)$ of pure entanglement. Entanglement of formation is the minimal average entanglement over all Charlie measurements. Thus entanglement of formation is pure entanglement available for Alice and Bob, when Charlie is adversary, who nevertheless honestly applies some rules.[29] When Alice and Bob cannot trust Charlie, the available pure entanglement is distillable entanglement, which is then obviously not greater than E_F.

In privacy theory, consider distribution (X,Y,Z). Similarly, Eve performs any operation on her variable $Z\to Z'$ and then reveals the value to Alice and Bob. Given revealed value z', Alice and Bob share distribution $p(x,y|z')$ product with Eve, so that on average they share $I(X,Y|Z')\equiv\sum_z' p(z')I(X,Y|Z'=z')$ bits of key, assuming Eve revealed the values honestly. Intrinsic information $I_{\rm intr}$ is defined as infimum of $I(X,Y|Z')$ over Eve's operations. It is obviously greater, than distillable key (as it is key drawn with some help of Eve, while distillable key is drawn without any help of Eve).

We will now reproduce the proof of Ref. 6 that intrinsic information of distribution (X,Y,Z) is a lower bound for amount of key which is needed to create such distribution.

To this end let us recall how it was proven that E_F^∞ is entanglement cost E_c.[11] That E_F^∞ is lower bound for E_c follows from inequality (7).

I.e. it follows from the fact E_F is an asymptotically continuous monotone,[3,30] that for singlet has value $\log d$. The converse was shown by construction of a protocol of creation of state. The protocol is suggested by the original definition of E_F: Alice and Bob pick at random i, produce state ψ_i from optimal decomposition, and forget i.

Let us now consider intrinsic information. We will argue that the result of Ref. 6 on the lower bound follows from Corrolary 2. Let us check that $I_{\rm intr}$ satisfies needed properties: (1) monotonicity of $I_{\rm intr}$ over local actions (Alice, Bob and Eve) follows from the very definition of $I_{\rm intr}$ (cf. Ref. 15); monotonicity under public communication was proven in Ref. 6; 2) asymptotic continuity of $I_{\rm intr}$ is established by the following inequality

$$|I_{\rm intr}(p) - I_{\rm intr}(q)| \leqslant 2H(\epsilon, 1-\epsilon) + \epsilon \min(\log \mathcal{X}, \log \mathcal{Y}) \tag{17}$$

essentially proven in Ref. 6; here $\mathcal{X}, \mathcal{Y}$ are ranges of X, Y and $||p-q|| = \epsilon$ (a proof can be also obtained by use of Refs. 31, 32) (3) additivity was applied without proof in Ref. 6 (the proof can be easily obtained from the proof of additivity of "squashed entanglement" Ref. 32); (4) for 1 bit of key, $I_{\rm intr} = 1$, as for private distribution product with Eve, $I_{\rm intr}$ is equal to mutual information. Thus by Corrolary 1 we get that intrinsic information is lower bound for key cost.

The protocol of formation, is also similar to the quantum one. Consider the optimal channel of Eve. Roughly speaking, Alice picks at random $\tilde{z}$, announce it to Bob (so that Eve also get to know it). Then the task is to produce $p(x, y|\tilde{z})$ which is a private (product) distribution. If now the cost of producing it requires $I(X : Y|\tilde{Z})$ secret bits, then on average they Alice and Bob need just $I_{\rm intr}$ of private bits. Since however we do not know at present, how many bits of key is needed to produce a private distribution, we can only get an upper bound in terms of this unknown function which we denote by $J(X, Y)$.

Proposition 1. Cost of producing a distribution (X, Y, Z) out of perfect key is bounded from above by $J(X, Y \downarrow Z)$, given by

$$J(X, Y \downarrow Z) = inf_\Lambda J(X, Y|\Lambda(Z)), \tag{18}$$

where $J(X, Y|Z) = \sum_z p(z) J(X, Y|Z = z)$ with $J(X, Y)$ being cost of producing private distribution (X, Y); infimum is taken over channels Λ.

To prove the proposition, let us take the optimal channel Λ, and call the produced random variable $\tilde{Z}$. Alice will pick a sequence of $\tilde{z}$'s of length n. The set of all sequences can be divided into strongly typical sequences,[(23)] that have $f(\tilde{z})$ occurrences of each $\tilde{z}$ (where $np_{\tilde{z}}(1-\delta) \leqslant f(\tilde{z}) \leqslant np_{\tilde{z}}(1+\delta)$) and all other sequences are negligible, since they occur with ϵ probability. (δ and ϵ can be made arbitrarily small by choosing large n). If the selected sequence is not strongly typical, Alice and Bob abort. If it is strongly typical, Alice and Bob will prepare $np_{\tilde{z}}(1-\delta)$ copies of distribution $p(x,y|\tilde{z})$ for each $\tilde{z}$. Thus the total distribution will be n copies of $p(x,y,z)$ and the amount of key that has to be used is $np_{\tilde{z}}J(X,Y|\tilde{Z})$, which gives $J(X,Y|\tilde{Z})$ per copy. Note, that we implicitly assumed that $\tilde{z}$ has fixed range (independent of n). However it also works if it is not the case. This ends the proof of the proposition.

Finally, let us mention, that from the above reasoning it clearly follows that intrinsic information has the interpretation of cost of producing given distribution from private distributions. Because we do not know if there is reversibility among private distributions themselves, the question of whether I_{intr} is cost key remains open.

5. RELATIVE ENTROPY DISTANCE IS EQUAL TO INTRINSIC INFORMATION

Here we will consider the direct analogue of quantum relative entropy of entanglement.[16,33] This quantity will be T of Ref. 15 optimized over Eve's channels. We will show it is equal to intrinsic information.

The counterpart of separable states are distributions that have zero intrinsic information. They are such distributions q where there exist channel Γ_E such that

$$\Gamma_E(\tilde{q}) = q, \tag{19}$$

where

$$q(x,y,z) = q(x|z)q(y|z)q(z). \tag{20}$$

That is after Eve applies channel Γ_E, given outcome z on her side, Alice and Bob share product distribution $q(x|z)q(y|z)$. We define the *relative entropy of privacy* by

$$K_R(p) = \inf_{\Lambda_E,\tilde{q}} S(\Lambda_E(p)|\tilde{q}), \tag{21}$$

where the infimum is taken over all Eve's channels and distributions $\tilde{q}$ satisfying (19); $S(\rho|\sigma) = \mathrm{Tr}\rho\log\rho - \mathrm{Tr}\rho\log\sigma$.

We will now simplify the definition a bit. Due to monotonicity of relative entropy, we have

$$K_R(p) \geqslant \inf_{\Lambda_E,\tilde{q},\Gamma_E} S(\Gamma_E\Lambda_E(p)|\Gamma_E(\tilde{q})) \tag{22}$$

for such Γ that $\Gamma(\tilde{q})$ is of the form (20). Then we have

$$\inf_{\Lambda_E,\tilde{q},\Gamma_E} S(\Gamma_E\Lambda_E(p)|\Gamma_E(\tilde{q})) \geqslant \inf_{\Lambda_E,q,\Gamma_E} S(\Gamma_E\Lambda_E(p)|q), \tag{23}$$

where Γ_E is now not restricted and q is of the form (20). Thus we get $K_R(p) \geqslant \inf_{\Lambda_E,q} S(\Lambda_E(p)|q)$. Also the converse inequality holds: the infimum in definition (21) of K_R is taken over greater set. Thus

$$K_R(p) = \inf_{\Lambda_E,q} S(\Lambda_E(p)|q), \tag{24}$$

where the infimum is taken only over probability distributions of the form (20).

Let us now evaluate $\inf_q S(p|q)$ where p is any fixed distribution, while q is of the form (20). It turns out that the optimal q is such that

$$q(z) = p(z); \quad q(x|z) = p(x|z); \quad q(y|z) = p(y|z), \tag{25}$$

where $p(z)$ is Eve's marginal distribution of p and $p(x|z)$ is the conditional distribution calculated from marginal distribution p^{AE} of p, similarly for $p(y|z)$. Then direct calculation gives

$$\inf_q S(p|q) = I(X : Y|Z). \tag{26}$$

Thus relative entropy distance from q satisfying (20) is equal to conditional mutual information. Then K_R is equal to intrinsic information.

ACKNOWLEDGMENTS

We thank Claude Crepeau, Matthias Christandl, Norbert Lütkenhaus and Andreas Winter on helpful discussions. This work is supported by the Polish Ministry of Scientific Research and Information Technology under the (solicited) grant no. PBZ-MIN-008/P03/2003 and by EU grants RESQ (IST-2001-37559), QUPRODIS (IST-2001-38877) and PROSECCO (IST-2001-39227). JO additionally acknowledges the support of a grant from the Cambridge-MIT Institute.

REFERENCES

1. N. Gisin and S. Wolf, http://arXiv.org/abs/quant-ph/0005042.
2. D. Collins and S. Popescu, *Phys. Rev. A* **65**, 032321 (2002).
3. C. H. Bennett, D. P. DiVincenzo, J. Smolin, and W. K. Wootters, *Phys. Rev. A* **54**, 3824 (1997).
4. C. H. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres, and W. K. Wootters, *Phys. Rev. Lett.* **70**, 1895 (1983).
5. A. Acin, L. Massanes, and N. Gisin, *Phys. Rev. Lett.* **91**, 167901 (2003).
6. R. Renner and S. Wolf, *Advances in Cryptology – EUROCRYPT '03, Lecture Notes in Computer Science* (Springer, 2003).
7. I. Devetak and A. Winter, "Distillation of secret key and entanglement from quantum states", *http://arXiv.org/abs/quant-ph/0306078*.
8. A. Acin, I. Cirac, and L. Massanes, *http://arXiv.org/abs/quant-ph/0311064*.
9. K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, *Phys. Rev. Lett.* **94**, 160502 (2005).
10. The term *information of formation* [6] is sometimes used, while we take the liberty of using the term *key cost* both for additional clarity and to distinguish it from other usages of this term [34].
11. P. Hayden, M. Horodecki, and B. Terhal, *J. Phys. A* **34**, 6891 (2001).
12. Regularization of the function f is given by $f^{\infty}(\rho) = \lim_n f(\rho^{\otimes n})/n$ where ρ is quantum or classical state.
13. C. H. Bennett, H. Bernstein, S. Popescu, and B. Schumacher, *Phys. Rev. A* **53**, 2046 (1996).
14. A. D. Wyner, *IEEE Trans. Inf. Theory* **IT-21**, 163 (1975).
15. N. J. Cerf, S. Massar, and S. Schneider, *Phys. Rev. A* **66**, 042309 (2002).
16. V. Vedral and M. B. Plenio, *Phys. Rev. A* **57**, 1619 (1998).
17. M. Horodecki, *Quantum Inform. Comp.* **1**, 3 (2001).
18. G. Vidal, *J. Mod. Opt.* **47**, 355 (2000).
19. M. Horodecki, P. Horodecki, and R. Horodecki, *Phys. Rev. Lett.* **84**, 2014 (2000).
20. M. Donald, M. Horodecki, and O. Rudolph, *J. Math. Phys.* **43**, 4252 (2002).
21. M. Horodecki, K. Horodecki, P. Horodecki, R. Horodecki, J. Oppenheim, A. Sen(De), and U. Sen, *Phys. Rev. Lett.* **90**, 100402 (2003).
22. M. Horodecki, P. Horodecki, and J. Oppenheim, *Phys. Rev. A* **67**, 062104 (2003).
23. T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, 1991).
24. U. Maurer and S. Wolf, *Lecture Notes Comput. Sci.* **1807**, 351 (2000).
25. I. Csiszar and J. Korner, *IEEE Trans. Inform. Theory* **24**, 339 (1978).
26. M. Horodecki, *Phys. Rev. A* **57**, 3364 (1998).
27. A. Winter, *http://arXiv.org/abs/quant-ph/0208131*.
28. W. Dur, G. Vidal, and I. Cirac, *Phys. Rev. A* **64**, 022308 (2001).
29. Similar scenario was independently considered by M. Christandl and R. Renner (in preparation).
30. M. A. Nielsen, *Phys. Rev. A* **61**, 064301 (2000).
31. R. Alicki and M. Fannes, *J. Phys. A* **37**,(2003).
32. M. Christandl and A. Winter, *http://arXiv.org/abs/quant-ph/0308088*.
33. V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, *Phys. Rev. Lett* **78**, 2275 (1997).
34. J. Oppenheim, M. Horodecki, and R. Horodecki, *Phys. Rev. Lett.* **90**, 010404 (2003).