A Property of the Intrinsic Mutual Information

Matthias Christandl [*] Renato Renner [†] Stefan Wolf [‡]

Abstract

In the setting where two parties knowing random variables X and Y, respectively, want to generate a secret key by communication accessible to an adversary who additionally knows a finite random variable Z, the so-called *intrinsic information* between X and Y given Z, $I(X;Y\downarrow Z)$, proved useful for determining the number of extractable secret key bits. Given a tripartite probability distribution P_{XYZ}, this information measure is, however, hard to compute in general since a minimization has to be made over all possible discrete-output channels $P_{\overline{Z}|Z}$ the adversary could use for processing her information. We strongly simplify this by showing that it can, without loss of generality, be assumed that the output alphabets of these channels equal their input alphabet; this implies in particular that there exists an optimal channel which achieves the minimum, since the set of such channels is compact. The proofs of our results combine techniques from point-set topology, measure theory, and convex geometry.

1 Introduction

In the context of unconditionally secure key agreement, the following measure for conditional mutual information, called the *intrinsic information*, between two discrete random variables X and Y, given a third variable Z, was defined as follows in [4].

Definition 1. **[4]** The *intrinsic conditional mutual information* (*intrinsic information* for short) between X and Y given Z is defined as

$$I(X;Y\downarrow Z) := \inf_{\overline{Z}} \left(I(X;Y|\overline{Z})\right) , \tag{1}$$

where the infimum is taken over all discrete random variables $\overline{Z}$ such that $XY \to Z \to \overline{Z}$ is a Markov chain.

The minimization in (1) includes, in other words, all discrete conditional probability distributions, or discrete channels, $P_{\overline{Z}|Z}$.

The intrinsic information is useful in a context where two parties, being connected by a public channel, and having access to (repeated realizations of) random variables X and Y, respectively, want to generate a key being secret even if a possible adversary has some additional knowledge specified by Z. In fact, it was shown [4] that $I(X;Y\downarrow Z)$ is an upper bound on the rate $S = S(X;Y||Z)$ at which such a key can be extracted. Later results allowed for clarifying the role of $I(X;Y\downarrow Z)$ in secret-key agreement: In [5] it is shown that another information measure, the *reduced intrinsic information of X and Y, given Z*, can be defined as a variation of $I(X;Y\downarrow Z)$ and yields an even better (tight?) upper bound on S. The result of the present paper, namely a simplification of the representation of $I(X;Y\downarrow Z)$ in terms of P_{XYZ}, immediately carries over to the reduced intrinsic

[*]ETH Zürich, CH-8092 Zürich, Switzerland. Email: matthias.christandl@qubit.org.

[†]Computer Science Department, ETH Zürich, CH-8092 Zürich, Switzerland. Email: renner@inf.ethz.ch.

[‡]Département d'Informatique et R.O., Université de Montréal, Montréal, QC, Canada H3C 3J7. Email: wolf@iro.umontreal.ca. Supported by Canada's NSERC.

information measure. Another recent (unpublished) result states that $I(X;Y\downarrow Z)$ is a lower bound on the rate at which secret-key bits are required for distributing pieces of information X and Y by public communication, leaving a possible wire-tapper with no more information than Z.

The main disadvantage of the intrinsic information measure is obvious: Since it is the infimum of the conditional mutual information $I(X;Y|\overline{Z})$, taken over the set of all possible discrete conditional probability distributions $P_{\overline{Z}|Z}$, it is a priori not easy to compute — by the same reason as, for instance, the capacity of a noisy communication channel can be hard to determine. In particular, for proving that $I(X;Y\downarrow Z)>0$ holds it is not enough to show that $I(X;Y|\overline{Z})$ is strictly positive for all Markov chains $XY\to Z\to\overline{Z}$ [2]: The minimum might not be taken by any channel since the space of discrete channels (more precisely, the set of channels with a fixed input alphabet and discrete yet unbounded output alphabet) is not a compact set. In this paper, we take a step towards understanding $I(X;Y\downarrow Z)$ better: We prove that the minimum *is indeed taken* by a specific channel $P_{\overline{Z}|Z}$ and, moreover, that we can assume without loss of generality that *the alphabet of $\overline{Z}$ equals (or is contained in) the alphabet of Z*. In other words, in order to reach the minimum, it is not necessary to extend the range of Z when computing $\overline{Z}$.

A consequence is that the following is true for all random variables X, Y, and Z (where the range $\mathcal{Z}$ of Z is finite): If there exists a Markov chain $XY\to Z\to\overline{Z}$ such that $I(X;Y|\overline{Z})=0$ holds, then there exists in particular a Markov chain $XY\to Z\to\overline{Z}_{\rm fin}$, where $\overline{Z}_{\rm fin}$ is now a *finite* random variable with $\overline{\mathcal{Z}}_{\rm fin}=\mathcal{Z}$, such that $I(X;Y|\overline{Z}_{\rm fin})=0$ holds.

2 Simplifying the Representation of the Intrinsic Information — The Minimum is Taken and the Range Need Not Be Extended

Let us state the main result of this paper.

Theorem 1. *If the range $\mathcal{Z}$ of Z is finite, then there exists a finite random variable $\overline{Z}$, having the same range $\mathcal{Z}$, such that $XY\to Z\to\overline{Z}$ is a Markov chain and*

$$I(X;Y\downarrow Z)=I(X;Y|\overline{Z})\ .$$

Corollary 2. *If the range $\mathcal{Z}$ of Z is finite, then*

$$I(X;Y\downarrow Z)=\min_{\overline{Z}} I(X;Y|\overline{Z})$$

where the minimum is taken over all random variables $\overline{Z}$ with range $\mathcal{Z}$ such that $XY\to Z\to\overline{Z}$ is a Markov chain.

Corollary 3. *If the range $\mathcal{Z}$ of Z is finite, then the following statements are equivalent:*

1. *There exists a discrete random variable $\overline{Z}$ such that $XY\to Z\to\overline{Z}$ is a Markov chain, and X and Y are independent conditioned on $\overline{Z}$.*
2. *There exists a finite random variable $\overline{Z}$ with range $\mathcal{Z}$ such that $XY\to Z\to\overline{Z}$ is a Markov chain, and X and Y are independent conditioned on $\overline{Z}$.*
3. *$I(X;Y\downarrow Z)=0$.*

Corollary 2 is a direct consequence of Theorem 1, stating that the infimum over discrete channels from Z to $\overline{Z}$ in the definition of the intrinsic information can be replaced by a minimum over channels having an output alphabet of size $|\mathcal{Z}|$.

Corollary 3 follows from the simple fact that, if (and only if) for some random variable $\overline{Z}$, the conditional mutual information $I(X;Y|\overline{Z})$ is zero, then, conditioned on $\overline{Z}$, the two random variables X and Y are independent.

The proof of Theorem 1 is based on Carathéodory's Fundamental Theorem [3] (Lemma 4), which can be used to derive a general statement about the expectation of functions of random variables defined on an affine space (Lemma 5). This then leads to a property of the expectation of functions defined on probability distributions (Lemma 6), from which, finally, Theorem 1 follows.

Let us introduce some notation to be used for the proof of Theorem 1. The expectation of a function $f = f(\,\cdot\,)$ of some random variable U is denoted by $E_U[f(U)] = E[f(U)]$. Similarly, for a function $g(\,\cdot\,,\,\cdot\,)$ of two arguments, the expectation $E_U[g(U,\,\cdot\,)]$ is a function (of one argument), mapping its argument x to $E_U[g(U,x)]$. It turns out to be convenient to express the mutual information $I(U;V)$ between two random variables U and V as a function of their distribution $P_{UV} = P_{UV}(\,\cdot\,,\,\cdot\,)$, i.e., $I(P_{UV}(\,\cdot\,,\,\cdot\,)) := I(U;V)$. The conditional mutual information $I(U;V|W)$ can then be written as $E_W[I(P_{UV|W}(\,\cdot\,,\,\cdot\,,W))]$.

Lemma 4. *Let D be a discrete random variable with range $\mathcal{D}$, being a subset of an n-dimensional affine space $\mathcal{W}$ over $\mathbb{R}$. Then, there exists a random variable F with range $\mathcal{F}$ of size $|\mathcal{F}| \leq n+1$ such that $\mathcal{F} \subseteq \mathcal{D}$ and $E[F] = E[D]$.*

Proof. The expectation values $E[D]$ and $E[F]$, being elements of $\mathcal{W}$, can be regarded as convex combinations of elements of $\mathcal{D}$ and $\mathcal{F}$, respectively. The statement is thus a direct consequence of Carathéodory's Fundamental Theorem [3], saying that each point in the convex hull of a set S in $\mathbb{R}^n$ is a convex combination of $n+1$ or fewer points of S. □

Lemma 5. *Let $\mathcal{W}$ be an n-dimensional affine space over $\mathbb{R}$, f a real-valued function on $\mathcal{W}$, and V a finite random variable taking values in $\mathcal{W}$ (i.e., its range $\mathcal{V}$ is a finite subset of $\mathcal{W}$).*

Then, there exists a random variable U with range $\mathcal{U} \subseteq \mathcal{V}$ of size at most $n+1$, satisfying

$$E[U] = E[V] \quad \textit{and} \quad E[f(U)] \leq E[f(V)]\ . \tag{2}$$

Proof. Let $\overline{U}$ be a random variable with minimal range $\overline{\mathcal{U}} \subseteq \mathcal{V}$ (i.e., $|\overline{\mathcal{U}}|$ is minimal) such that (2) is satisfied for $U = \overline{U}$, and assume by contradiction that $|\overline{\mathcal{U}}| > n+1$. Since (2) trivially holds for $U = V$, $|\overline{\mathcal{U}}|$ is bounded by $|\mathcal{V}|$, i.e., $|\overline{\mathcal{U}}| \leq |\mathcal{V}| < \infty$. The minimal value $|\overline{\mathcal{U}}|$ is thus well-defined.

According to Lemma 4, there is a random variable $\overline{U}'$ with range $\overline{\mathcal{U}}'$ such that $E[\overline{U}'] = E[\overline{U}]$, $\overline{\mathcal{U}}' \subseteq \overline{\mathcal{U}}$, and $|\overline{\mathcal{U}}'| \leq n+1$. Let the set $\overline{\mathcal{U}}'$ be restricted to values with strictly positive probability, i.e., for all $v \in \overline{\mathcal{U}}'$, $P_{\overline{U}'}(v) > 0$, and define $\alpha(v) := P_{\overline{U}}(v)/P_{\overline{U}'}(v)$ for all $v \in \overline{\mathcal{U}}'$.

Let v_0 be an element of $\overline{\mathcal{U}}'$ such that $\alpha(v)$ is minimal for $v = v_0$, i.e., $\alpha(v) \geq \alpha(v_0)$ for all $v \in \overline{\mathcal{U}}'$, and let $\alpha_0 := \alpha(v_0)$. Note that

$$\sum_{v \in \overline{\mathcal{U}}'} \bigl(P_{\overline{U}'}(v) - P_{\overline{U}}(v)\bigr) = 1 - \mathrm{Prob}\,[\overline{U} \in \overline{\mathcal{U}}'] > 0\ ,$$

where the last inequality follows from the fact that, by assumption, the range of $\overline{U}$ is minimal and $|\overline{\mathcal{U}}| > n+1 \geq |\overline{\mathcal{U}}'|$, i.e., $\mathrm{Prob}\,[\overline{U} \notin \overline{\mathcal{U}}'] > 0$. There is thus at least one element $v \in \overline{\mathcal{U}}'$ such that $P_{\overline{U}}(v) < P_{\overline{U}'}(v)$ holds. Hence we have $\alpha(v) < 1$ and, consequently, $\alpha_0 < 1$.

Let $\overline{U}''$ be a new random variable on the set $\overline{\mathcal{U}}$ with

$$P_{\overline{U}''}(v) := \frac{1}{1-\alpha_0}\bigl(P_{\overline{U}}(v) - \alpha_0 \cdot P_{\overline{U}'}(v)\bigr) \tag{3}$$

for all $v \in \overline{\mathcal{U}}$. The random variable $\overline{U}''$ is well-defined since, as an immediate consequence of the definition of $\alpha(v)$ and $\alpha_0 \leq \alpha(v)$, the probabilities $P_{\overline{U}''}(v)$ are non-negative (for all $v \in \overline{\mathcal{U}}$), and, additionally, sum up to one:

$$\sum_{v \in \overline{\mathcal{U}}} P_{\overline{U}''}(v) = \frac{1}{1-\alpha_0}\Bigl(\sum_{v \in \overline{\mathcal{U}}} P_{\overline{U}}(v) - \alpha_0 \sum_{v \in \overline{\mathcal{U}}} P_{\overline{U}'}(v)\Bigr) = 1\ .$$

Let $\overline{\mathcal{U}}'' \subseteq \overline{\mathcal{U}}$ be the minimal set such that $P_{\overline{U}''}(v)$ is strictly positive for all $v \in \overline{\mathcal{U}}''$. By definition, v_0 is contained in $\overline{\mathcal{U}}' \subset \overline{\mathcal{U}}$. On the other hand, since

$$P_{\overline{U}''}(v_0) = \frac{1}{1-\alpha_0}\big(P_{\overline{U}}(v_0) - \underbrace{\alpha_0 \cdot P_{\overline{U}'}(v_0)}_{P_{\overline{U}}(v_0)}\big) = 0 \ ,$$

v_0 is not contained in $\overline{\mathcal{U}}'' \subseteq \overline{\mathcal{U}}$. The set $\overline{\mathcal{U}}''$ is thus strictly smaller than the set $\overline{\mathcal{U}}$, i.e., $|\overline{\mathcal{U}}''| < |\overline{\mathcal{U}}|$.

Rewriting equation (3) as $P_{\overline{U}}(v) = (1-\alpha_0) \cdot P_{\overline{U}''}(v) + \alpha_0 \cdot P_{\overline{U}'}(v)$, the expectation of $g(\overline{U})$, for any function g defined on $\overline{\mathcal{U}}$, can be expressed as

$$E[g(\overline{U})] = (1-\alpha_0) \cdot E[g(\overline{U}'')] + \alpha_0 \cdot E[g(\overline{U}')] \ . \tag{4}$$

Letting g be the identity function, and recalling that $E[\overline{U}'] = E[\overline{U}]$ and $\alpha_0 < 1$, we directly obtain

$$E[\overline{U}''] = E[\overline{U}'] = E[\overline{U}] \ .$$

Furthermore, for $g = f$, it follows directly from (4) that $E[f(\overline{U}')]$ and $E[f(\overline{U}'')]$ can not both be larger than $E[f(\overline{U})] \leq E[f(V)]$, i.e., (2) holds for either $U = \overline{U}'$ or $U = \overline{U}''$ (or both). Since $|\overline{\mathcal{U}}'| \leq n+1 < |\overline{\mathcal{U}}|$ and $|\overline{\mathcal{U}}''| < |\overline{\mathcal{U}}|$, this is a contradiction to the assumption that $\overline{U}$ has minimal range. □

Lemma 6. *Let $(Z^i, V^i)_{i\in\mathbb{N}}$ be a sequence of pairs of discrete random variables where the range of Z^i is a finite set $\mathcal{Z}$ (for all $i \in \mathbb{N}$).*

Then, for any real-valued continuous function f defined on the set of all probability distributions over $\mathcal{Z}$, $\mathcal{P}(\mathcal{Z})$, there is a pair $(\overline{Z}, \overline{V})$ of finite random variables, both having range $\mathcal{Z}$, satisfying

$$P_{\overline{Z}} = \lim_{i\to\infty} P_{Z^i} \tag{5}$$

$$E_{\overline{V}}[f(P_{\overline{Z}|\overline{V}}(\,\cdot\,,\overline{V}))] \leq \lim_{i\to\infty} E_{V^i}[f(P_{Z^i|V^i}(\,\cdot\,,V^i))] \tag{6}$$

(if these limits are defined).

Proof. The probability distribution of Z^i conditioned on V^i (for $i \in \mathbb{N}$) can be considered as a random variable $X^i := P_{Z^i|V^i}(\,\cdot\,,V^i)$ with range $\mathcal{P}(\mathcal{Z})$. The expectation value on the right-hand side of (6) can then be written as

$$E[f(X^i)] = E_{V^i}[f(P_{Z^i|V^i}(\,\cdot\,,V^i))] \ , \tag{7}$$

and the expectation of X^i satisfies

$$E[X^i] = E_{V^i}[P_{Z^i|V^i}(\,\cdot\,,V^i)] = \sum_v P_{V^i}(v) P_{Z^i|V^i}(\,\cdot\,,v) = P_{Z^i}(\,\cdot\,) \ . \tag{8}$$

Since, for each $i \in \mathbb{N}$, the value of the random variable X^i is fully determined by the value of the discrete random variable V^i, X^i must be discrete as well. Therefore, for each $\varepsilon_i > 0$, there is a finite subset $\mathcal{U}^i$ of the range of X^i such that the random variable U^i, defined by $P_{U^i}(v) := P_{X^i|v\in\mathcal{U}^i}(v)$ (for $v \in \mathcal{U}^i$), satisfies

$$|E[f(U^i)] - E[f(X^i)]| \leq \varepsilon_i \tag{9}$$

and, for all $z \in \mathcal{Z}$,

$$|E[U^i](z) - E[X^i](z)| \leq \varepsilon_i \ . \tag{10}$$

The range $\mathcal{U}^i$ of U^i, consisting of elements of $\mathcal{P}(\mathcal{Z})$, is a finite subset of the $(n-1)$-dimensional affine space $\{v = (v_1, \ldots, v_n) \in \mathbb{R} : \sum_{i=1}^n v_i = 1\}$ where $n := |\mathcal{Z}|$. We can thus apply Lemma 5 which implies that there exists a random variable T^i with range $\mathcal{T}^i \subseteq \mathcal{U}^i$ of size (at most) n satisfying

$$\begin{aligned} E[f(T^i)] &\leq E[f(U^i)] && (11) \\ \text{and} \qquad E[T^i] &= E[U^i] \ . && (12) \end{aligned}$$

Let $\varepsilon_i := 1/i$. Then, combining (11) with (7) and (9), and, similarly, combining (12) with (8) and (10), leads to

$$E[f(T^i)] \leq E_{V^i}[f(P_{Z^i|V^i}(\,\cdot\,, V^i))] + 1/i \tag{13}$$

and, for all $z \in \mathcal{Z}$,

$$|E[T^i](z) - P_{Z^i}(z)| \leq 1/i \tag{14}$$

(for all $i \in \mathbb{N}$).

Let $v^{(1)}, \ldots, v^{(n)}$ be the n elements of the range $\mathcal{T}^i \subseteq \mathcal{P}(\mathcal{Z})$ and define the function

$$\begin{array}{rccc} p^i : & \mathcal{Z} \times \{1, \ldots, n\} & \longrightarrow & \mathbb{R}^+ \\ & (z, k) & \longmapsto & P_{T^i}(v^{(k)}) \cdot v^{(k)}(z) \ . \end{array}$$

Then, for all $z \in \mathcal{Z}$,

$$\sum_{k=1}^n p^i(z,k) = \sum_{v \in \mathcal{T}^i} P_{T^i}(v) \cdot v(z) = E[T^i](z) \ . \tag{15}$$

Since the range of T^i is a subset of $\mathcal{P}(\mathcal{Z})$, it follows that $\sum_{z \in \mathcal{Z}} \sum_{k=1}^n p^i(z,k) = 1$, which implies that the functions p^i (for $i \in \mathbb{N}$) are bounded by the interval $[0, 1]$. Because the set of bounded functions defined on a finite set is compact, the sequence $(p^i)_{i \in \mathbb{N}}$ has a convergent subsequence $(p^{i_j})_{j \in \mathbb{N}}$, i.e., the limes $\overline{p} := \lim_{j \to \infty} p^{i_j}$ exists.

From (15) and (14), we obtain, for all $z \in \mathcal{Z}$,

$$\sum_{k=1}^n \overline{p}(z,k) = \lim_{j \to \infty} \sum_{k=1}^n p^{i_j}(z,k) = \lim_{j \to \infty} E[T^{i_j}](z) = \lim_{i \to \infty} P_{Z^i}(z) \ . \tag{16}$$

Since $\overline{p}$ is thus a probability distribution, we can define two random variables $\overline{Z}$ and $\overline{V}$ with range $\mathcal{Z} = \{z^{(1)}, \ldots, z^{(n)}\}$, such that

$$P_{\overline{Z}\overline{V}}(z, z^{(k)}) = \overline{p}(z,k)$$

(for all $(z,k) \in \mathcal{Z} \times \{1, \ldots, n\}$). Equation (5) is then a direct consequence of (16). Furthermore, from the continuity of f, we get

$$E_{\overline{V}}[f(P_{\overline{Z}|\overline{V}}(\,\cdot\,, \overline{V}))] = \lim_{j \to \infty} E[f(T^{i_j})] \tag{17}$$

which, together with (13), implies (6) and thus concludes the proof. □

Proof (of Theorem 1). The main idea of the proof is to express the mutual conditional information $I(X;Y|W)$ (for some random variable W) in terms of an expectation value of a function σ of conditional probabilities, such that Lemma 6 can be applied.

Let $\mathcal{P}(\mathcal{Z})$ be the space of all probability distributions on the range $\mathcal{Z}$ of Z. For any random variable W (with range $\mathcal{W}$) such that $XY \to Z \to W$ is a Markov chain, $I(X;Y|W)$ can be rewritten as

$$\begin{aligned} I(X;Y|W) &= \sum_{w \in \mathcal{W}} P_W(w) \cdot I(P_{XY|W}(\,\cdot\,, \cdot\,, w)) \\ &= \sum_{w \in \mathcal{W}} P_W(w) \cdot \sigma(P_{Z|W}(\,\cdot\,, w)) = E_W[\sigma(P_{Z|W}(\,\cdot\,, W))] \ . \end{aligned} \tag{18}$$

where σ is the function

$$\begin{array}{rccl} \sigma: & \mathcal{P}(\mathcal{Z}) & \longrightarrow & \mathbb{R} \\ & P & \longmapsto & I\left(\sum_{z\in\mathcal{Z}} P_{XY|Z}(\,\cdot\,,\,\cdot\,,z)\cdot P(z)\right)\ . \end{array}$$

According to the definition of the intrinsic information, there is a sequence $(W^i)_{i\in\mathbb{N}}$ of discrete random variables satisfying

$$\lim_{i\to\infty} I(X;Y|W^i) = I(X;Y\downarrow Z) \tag{19}$$

such that $XY \to Z \to W^i$ is a Markov chain. Consequently, using (18) and (19), we obtain

$$I(X;Y\downarrow Z) = \lim_{i\to\infty} E_{W^i}[\sigma(P_{Z|W^i}(\,\cdot\,,W^i))]\ . \tag{20}$$

Since σ is obviously continuous, we can directly apply Lemma 6 for $(Z^i,V^i) := (Z,W^i)$, leading to a random variable $\overline{V}$ with range $\mathcal{Z}$, such that $XY \to Z \to \overline{V}$ is a Markov chain and

$$E_{\overline{V}}[\sigma(P_{Z|\overline{V}}(\,\cdot\,,\overline{V}))] \le \lim_{i\to\infty} E_{W^i}[\sigma(P_{Z|W^i}(\,\cdot\,,W^i))]\ .$$

Thus, using (18) (for $W=\overline{V}$) and (20), we finally have $I(X;Y|\overline{V}) \le I(X;Y\downarrow Z)$. Setting $\overline{Z} := \overline{V}$, Theorem 1 follows from the fact that the intrinsic information $I(X;Y\downarrow Z)$, being defined as an infimum, cannot be larger than $I(X;Y|\overline{V})$. □

3 Concluding Remarks

We have shown that the intrinsic information measure, which is of importance in the context of unconditionally secure key agreement, is easier to compute, to understand, and to manipulate than previously believed: The *infimum* in the definition of $I(X;Y\downarrow Z)$, taken over the set of arbitrary discrete-output channels, can be replaced by a *minimum*, taken over all channels whose output alphabet equals the input alphabet.

Using this result, it is a much easier task to prove for instance that the intrinsic information of a distribution is non-vanishing. Such proofs were used in [2] and [1] in order to analyze tripartite probability distributions arising when so-called bound entangled quantum states are measured; these distributions are believed to satisfy $S(X;Y||Z) = 0$ whereas $I(X;Y\downarrow Z) > 0$. It is a fundamental problem to prove the existence of such "bound information" which cannot be used for the generation of a secret key by any protocol [2],[1],[5].

References

[1] N. Gisin, R. Renner, and S. Wolf, Linking classical and quantum key agreement: is there a classical analog to bound entanglement?, in *Algorithmica*, vol. 34, pp. 389–412, 2002.

[2] N. Gisin and S. Wolf, Linking classical and quantum key agreement: is there "bound information"?, in *Proceedings of CRYPTO 2000*, Lecture Notes in Computer Science, vol. 1880, pp. 482–500, Springer-Verlag, 2000.

[3] P. M. Gruber and J. M. Wills (Eds.), *Handbook of Convex Geometry*, Vol. A, Elsevier Science Publishers, Amsterdam, 1993.

[4] U. Maurer and S. Wolf, Unconditionally secure key agreement and the intrinsic conditional information, *IEEE Transactions on Information Theory*, Vol. 45, No. 2, pp. 499–514, 1999.

[5] R. Renner, J. Skripsky, and S. Wolf, A new measure for conditional mutual information and its properties, manuscript, 2002.