A Property of the Intrinsic Mutual Information

Matthias Christandl
Centre for Quantum Computation
DAMTP, University of Cambridge
Cambridge CB3 0WA
United Kingdom
e-mail: matthias.christandl@qubit.org

Renato Renner
Computer Science Department
ETH Zürich
CH-8092 Zürich
Switzerland
e-mail: renner@inf.ethz.ch

Stefan Wolf
Département d’Informatique et R.O.
Université de Montréal
Montréal, QC
Canada H3C 3J7
e-mail: wolf@iro.umontreal.ca

Abstract — The so-called intrinsic mutual information is an important measure in the context of information-theoretic secret-key agreement. We prove a property of this information measure which, in particular, strongly simplifies its computation. More generally, our result is useful for analyzing the correlation of two random variables conditioned on a third one.

I. Definitions and Motivation

The intrinsic (mutual) information [2] between two discrete random variables X and Y, given a third random variable Z, is defined as

\[ I(X; Y | Z) := \inf_Z I(X; Y | Z), \]

where the infimum is taken over all discrete random variables Z such that XY \to Z \to Z is a Markov chain. This minimization includes, in other words, all discrete conditional probability distributions, or discrete channels, \( P_{Z|XY} \).

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 possesses some knowledge, specified by Z. In fact, it was shown [2] that \( I(X; Y | Z) \) is an upper bound on the rate \( S = \mathbb{S}(X; Y | Z) \) at which such a key can be extracted. Another recent result [3] states that \( I(X; Y | Z) \) 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.

Since the intrinsic information is defined by an infimum ranging over the set of all possible discrete conditional probability distributions \( P_{Z|XY} \), it is a priori not easy to compute. In particular, to prove that \( I(X; Y | Z) > 0 \) holds, it is not enough to show that \( I(X; Y | Z) \) is strictly positive for all Markov chains XY \to Z \to Z: The minimum might not be attained by any particular channel since the space of discrete channels is not a compact set. Our result is a step towards the better understanding of \( I(X; Y | Z) \): We prove that the minimum is indeed taken by a specific channel \( P_{Z|XY} \), and, moreover, that this minimum can be reached for a channel whose output alphabet is not larger than the alphabet of Z.

As a consequence, the following is true for all random variables X, Y, and Z (where the range Z of Z is finite): If there exists a Markov chain XY \to Z \to Z such that \( I(X; Y | Z) = 0 \) holds, then there exists a Markov chain XY \to Z \to Zfin, where \( Z_{\text{fin}} \) is now a finite random variable with range \( Z_{\text{fin}} = Z \), such that \( I(X; Y | Z_{\text{fin}}) = 0 \) holds.

II. Main Results and Conclusions

Theorem. If the range Z of Z is finite, then there exists a finite random variable Z, having the same range Z, such that XY \to Z \to Z is a Markov chain and

\[ I(X; Y | Z) = I(X; Y | Z). \]

The infimum over discrete channels from Z to Z in the definition of the intrinsic information can thus be replaced by a minimum over channels with output alphabet Z.

Corollary 1. If the range Z of Z is finite, then

\[ I(X; Y | Z) = \min_Z I(X; Y | Z), \]

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

In particular, this result simplifies the task of proving that the intrinsic information of a given triple of random variables is non-vanishing [1]. If and only if \( I(X; Y | Z) \), the mutual information of random variables X and Y with respect to Z, vanishes, then X and Y are independent conditioned on Z. This immediately proves the following corollary.

Corollary 2. If the range Z of Z is finite, then the following statements are equivalent:

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

References


1Supported by the Swiss National Science Foundation (SNF).

2Supported by Canada’s NSERC.