A Hierarchy of Information Quantities for the Finite Block Length Analysis of Quantum Tasks

Marco Tomamichel[1] and Masahito Hayashi[1,2]

[1] *Centre of Quantum Technologies, National University of Singapore, Singapore.*[*]
[2] *Graduate School of Mathematics, Nagoya University, Nagoya, Japan.*[†]

We consider two fundamental tasks in quantum information theory, data compression with quantum side information as well as randomness extraction against quantum side information. We characterize these tasks for general sources using so-called one-shot entropies. These characterizations — in contrast to earlier results — enable us to derive tight second order asymptotics for these tasks in the i.i.d. limit. More generally, our derivation establishes a hierarchy of information quantities that can be used to investigate information theoretic tasks in the quantum domain: The one-shot entropies most accurately describe an operational quantity, yet they tend to be difficult to calculate for large systems. We show that they asymptotically agree (up to logarithmic terms) with entropies related to the quantum and classical information spectrum, which are easier to calculate in the i.i.d. limit. Our technique also naturally yields bounds on operational quantities for finite block lengths.

full version: `arXiv: 1208.1478`

INTRODUCTION

The characterization of information theoretic tasks that are repeated only once (the *one-shot* setting) or a finite number of times (the *finite block length* setting) has recently generated great interest in classical information theory [7, 11]. In particular, these studies investigate the asymptotic performance of information theoretic tasks in the second order, i.e. they determine precisely the term that scales proportional to $\sqrt{n}$ when we consider n independent and identically distributed (*i.i.d.*) uses of a source or channel. Among the tasks that have been studied are noiseless source coding [6, 9], Slepian-Wolf coding [1, 14], random number generation when the source distribution is known [6], the classical statistical evaluation used for parameter estimation in quantum cryptography [5], and channel coding [7, 10, 11].

Concurrently, progress has been made towards characterizing tasks utilizing quantum resources in the same setting. Two different techniques have been proposed to achieve this: The information spectrum method [3, 4] and one-shot entropies [12]. Combining these two approaches, we are able to derive the first second order expansion of an operational quantity utilizing quantum resources.

OVERVIEW OF RESULTS

Given a source that emits a random variable X and quantum side information B about X, we study the following two *operational quantities* in the one-shot setting.

- The maximal number of random and secret bits, ε-close to uniform and independent of B, that can be extracted from X is denoted $\ell^\varepsilon(X|B)$. This task was first investigated by Renner and König [13] in the quantum setting and has various applications in cryptography.
- The minimal length in bits of an encoding M of X, such that X can be recovered up to an error ε from B and M is denoted $m^\varepsilon(X|B)$. This is noiseless source compression with side information and has been investigated by Devetak and Winter [2] in the quantum setting.

Second Order Expansion of Operational Quantities: Our first contribution is to show that both the direct and converse bounds on the operational quantities converge to the same expression in the

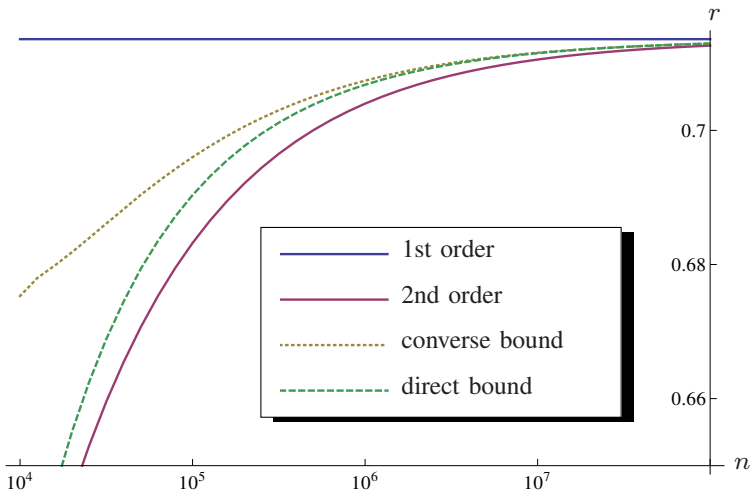

FIG. 1. The plot shows direct and converse bounds on $r = \frac{1}{n}\ell^\varepsilon(X^n|B^n)$ for $n \in [10^4, 10^8]$. Note that for increasing n these bounds first converge to each other, then to the second order asymptotics of the task, and finally to the Shannon rate.

second order. In particular, we find the following asymptotic expansion for n i.i.d. copies of the source and any non-trivial error parameter $0 < \varepsilon < 1$:

$$\begin{aligned} \ell^\varepsilon(X^n|B^n) &= nH(X|B) + \sqrt{n}\, s(X|B)\Phi^{-1}(\varepsilon^2) + O(\log n), \\ m^\varepsilon(X^n|B^n) &= nH(X|B) - \sqrt{n}\, s(X|B)\Phi^{-1}(\varepsilon) + O(\log n), \end{aligned} \tag{1}$$

where $H(X|B)$ is the conditional von Neumann entropy of the source, Φ is the cumulative normal distribution function, and $s(X|B)$ may be interpreted as a standard deviation of this entropy.

The asymptotic expansion in (1) should be read as follows. First, independently of the allowed error ε, the operational quantities approach the von Neumann entropy for large n. This is the first order expansion and has been known previously [2, 12]. Our second order expansion now shows how fast the quantities convergence as a function of ε and the standard deviation of the source. For a small[1] error parameter ε, this means that $\ell^\varepsilon(X^n|B^n)$ is smaller than $nH(X|B)$ by a term scaling with $\sqrt{n}$ while $m^\varepsilon(X^n|B^n)$ is larger than the Shannon limit by a term scaling with $\sqrt{n}$. This is in accordance with our intuition that these tasks are harder to fulfill for small n and ε.

Finite Block Length Analysis: Our analysis naturally yields both direct and converse bounds on the operational quantities for finite n, which can be evaluated numerically. We give an example of such a finite block length analysis in Fig. 1. For this purpose, we consider the state that results when transmitting either $|0\rangle$ or $|1\rangle$ through the complementary channel of a Pauli channel with a phase error $p = 0.05$ that is independent of the bit flip error. The resulting state is

$$c\rho_{XB} = \frac{1}{2}|0\rangle\langle 0| \otimes |\phi^0\rangle\langle\phi^0| + \frac{1}{2}|1\rangle\langle 1| \otimes |\phi^1\rangle\langle\phi^1|, \quad \text{where} \quad |\phi^x\rangle = \sqrt{p}\,|0\rangle + (-1)^x\sqrt{1-p}\,|1\rangle\,.$$

We are interested in how much randomness can be extracted from X if we require $\varepsilon = 10^{-6}$.

Converse bounds for finite n are very relevant because they allow us to investigate how close a given protocol is to the maximal achievable rate. From Fig. 1, for example, we can deduce that

Hierarchy of Information Quantities: Furthermore, we establish a hierarchy of information quantities (cf. Figure 2) that can be used to analyze quantum information tasks beyond the examples

[1] Note that $\Phi^{-1}(\varepsilon)$ is negative for small ε and changes sign when ε exceeds $1/2$.

Class	Role	Quantities
Class 1	Describing the optimal performance of the protocol. The calculation is very difficult, even for small systems.	$m^\varepsilon(X\|B)_\rho$ $l^\varepsilon(X\|B)_\rho$,
Class 2	One-shot bound for general sources. The calculation is possible for small systems, using an SDP.	$H_h^{\varepsilon\pm\eta}(A\|B)_\rho$, $H_{\min}^{\varepsilon\pm\eta}(A\|B)_\rho$, $D_h^{\varepsilon\pm\eta}(\rho\|\sigma)$, $D_{\max}^{\varepsilon\pm\eta}(\rho\|\sigma)$
Class 3	Quantum version of information spectrum.	$D_s^{\varepsilon\pm\delta}(\rho\|\sigma)$, $D_s^{\varepsilon\pm\delta}(\mathcal{E}_\sigma(\rho)\|\sigma)$
Class 4	Classical information spectrum. The calculation is approximately possible for i.i.d. sources.	$D_s^{\varepsilon\pm\delta}(P_{\rho,\sigma}\|Q_{\rho,\sigma})$
Class 5	Second order asymptotics. The calculation is easy for arbitrarily large n.	$nD(\rho\|\sigma)+\sqrt{n}\,s(\rho\|\sigma)\Phi^{-1}(\varepsilon)$ $nH(X\|B)+\sqrt{n}\,s(X\|B)\Phi^{-1}(\varepsilon)$

Classes	Difference	Method
Class 1 $\to$ Class 2	$O(\log n)$ with $\eta \propto \frac{1}{\sqrt{n}}$	Random coding and monotonicity.
Class 2 $\to$ Class 4	$O(\log n)$ with $\delta \propto \frac{1}{\sqrt{n}}$	Relations between entropies.
Class 4 $\to$ Class 5	$O(1)$	Berry-Esseen Theorem.

FIG. 2. Hierarchy of information quantities. We consider the tasks for a constant $\varepsilon \in (0,1)$ and consider n i.i.d. repetitions of tasks for large n. Note that Class 3 and Class 4 are unified if ρ and σ commute.

discussed above. The hierarchy is partly inspired by recent results in hypothesis testing and constitutes the main technical contribution of this paper.

The operational quantities (Class 1) are highly dependent on the problem considered. For example, in the case of randomness extraction, this quantity depends on the precise security requirement we impose on the extracted random variable. In a first step, we thus bound them in terms of one-shot entropies (Class 2) as in (**??**). These quantities in general also depend on the exact specification of the considered problem. They can be formulated as semi-definite optimization problems[2] (SDPs) and can be calculated for small examples. For large block lengths (e.g., $n \gg 10$), however, these optimization problems quickly become intractable as their complexity scales exponentially in n.

Thus, we relate the one-shot entropies to the quantum information spectrum (Class 3) and then the classical information spectrum of the corresponding Nussbaum-Szkoła distributions (Class 4), which can often be approximated even for large n. Finally, the classical information spectrum allows us to evaluate the second order asymptotics precisely (Class 5).

Alternatively, one can see the one-shot entropies as providing us a "microscopic" analysis of the performance of a task for general sources, whereas the information spectrum and their asymptotic expansion quantities give a "macroscopic" view that can be approximately calculated for sources with sufficient structure.

We show that the following quantities are equivalent in an appropriate sense:

$$D_{\max}^{\sqrt{1-\varepsilon}}(\rho|\sigma) \approx D_h^\varepsilon(\rho|\sigma) \approx D_s^\varepsilon(\rho\|\sigma) \approx D_s^\varepsilon(\mathcal{E}_\sigma(\rho)\|\sigma) \approx D_s^\varepsilon(P_{\rho,\sigma}\|Q_{\rho,\sigma}).$$

First, note that the smoothing parameter is varied by a constant term in some relations. More importantly, the equivalence only holds up to additive terms $\log\theta(\sigma)$, where $\theta(\sigma)$ is the logarithm of the ratio between the largest and smallest eigenvalue of σ.

In the i.i.d. setting, it is evident that $\theta(\sigma^n)$ grows at most linearly in n and this additive term thus grows at most as $O(\log n)$. Hence, our results imply that the smoothing parameter for all

[2] The min-entropy can be formulated as an SDP [8] and an extension of this to the smooth min-entropy is possible. The SDP for hypothesis testing is discussed in Section **??**.

these quantities can be chosen as $\varepsilon \pm \mathrm{poly}(n)^{-1}$ without incurring a penalty that grows faster than $O(\log n)$.

---

* cqtmarco@nus.edu.sg
† masahito@math.nagoya-u.ac.jp

[1] D. Baron, M. A. Khojastepour, and R. G. Baraniuk. Redundancy Rates of Slepian-Wolf Coding. In *Proc. 42nd Annual Allerton Conf. on Comm., Control, and Computing*, 2004.

[2] I. Devetak and A. Winter. Classical Data Compression with Quantum Side Information. *Phys. Rev. A*, 68(4), Oct. 2003. DOI: 10.1103/PhysRevA.68.042301.

[3] T. S. Han. *Information-Spectrum Methods in Information Theory*. Springer, 2002.

[4] T. S. Han and S. Verdu. Approximation Theory of Output Statistics. In *Proc. IEEE ISIT*, pages 153–153. IEEE, 1993. DOI: 10.1109/ISIT.1993.748468.

[5] M. Hayashi. *Quantum Information — An Introduction*. Springer, 2006.

[6] M. Hayashi. Second-Order Asymptotics in Fixed-Length Source Coding and Intrinsic Randomness. *IEEE Trans. on Inf. Theory*, 54(10):4619–4637, Oct. 2008. DOI: 10.1109/TIT.2008.928985.

[7] M. Hayashi. Information Spectrum Approach to Second-Order Coding Rate in Channel Coding. *IEEE Trans. on Inf. Theory*, 55(11):4947–4966, Nov. 2009. DOI: 10.1109/TIT.2009.2030478.

[8] R. König, R. Renner, and C. Schaffner. The Operational Meaning of Min- and Max-Entropy. *IEEE Trans. on Inf. Theory*, 55(9):4337–4347, Sept. 2009. DOI: 10.1109/TIT.2009.2025545.

[9] I. Kontoyiannis. Second-Order Noiseless Source Coding Theorems. *IEEE Trans. on Inf. Theory*, 43(4):1339–1341, July 1997. DOI: 10.1109/18.605604.

[10] Y. Polyanskiy. *Channel Coding: Non-Asymptotic Fundamental Limits*. PhD thesis, Princeton University, Nov. 2010.

[11] Y. Polyanskiy, H. V. Poor, and S. Verdu. Channel Coding Rate in the Finite Blocklength Regime. *IEEE Trans. on Inf. Theory*, 56(5):2307–2359, May 2010. DOI: 10.1109/TIT.2010.2043769.

[12] R. Renner. *Security of Quantum Key Distribution*. PhD thesis, ETH Zurich, Dec. 2005. arXiv: quant-ph/0512258.

[13] R. Renner and R. König. Universally Composable Privacy Amplification Against Quantum Adversaries. In *Proc. TCC*, volume 3378 of *LNCS*, pages 407–425, Cambridge, USA, 2005. DOI: 10.1007/978-3-540-30576-7_22.

[14] V. Y. F. Tan and O. Kosut. The Dispersion of Slepian-Wolf Coding. In *Proc. IEEE ISIT*, 2012.