## Channel Simulation: Tight meta converse for error and strong converse exponents

Mario Berta, Michael X. Cao, Hao-Chung Cheng, Omar Fawzi, Aadil Oufkir, and Yongsheng Yao

**Introduction** – One of the most fundamental questions in Shannon theory is to ask at what optimal asymptotic rate one noisy channel $\mathcal{V}$ can simulate another noisy channel $\mathcal{W}$ — termed *channel interconversion* [1]. For the target channel $\mathcal{W} = \mathcal{I}$, this has famously been answered by Shannon's noisy channel coding theorem [2] and refined by its finite-blocklength [3, 4] and higher-order [3, 5–7] characterizations. The opposite task $\mathcal{V} = \mathcal{I}$ corresponds to channel simulation, which has been studied in different variants, e.g., in [8–11]. In contrast to channel coding, finite-blocklength and higher-order characterizations for channel simulation have only recently been understood better [12–14] and are a first step towards the ultimate goal of a refined analysis of channel interconversion beyond the first-order optimal asymptotic rate. In this submission, we expand on an algorithmic ideas that were previously explored for channel coding [15]. These methods give a simple and efficient algorithm that generates a code achieving at least $(1 - \frac{1}{e})$ of the maximum success probability that can ever be attained under the same message-size constraint. The algorithm is based on a natural linear programming relaxation of the problem and can be understood as allowing the sender and receiver to share no-signaling correlations [16] — which exactly corresponds to the well-known *PPV meta-converse* for channel coding [3, 4].

This motivates us to investigate no-signaling correlations, and more broadly the meta-converse for the task of channel simulation. Particularly, in this work, we consider the problem of simulating classical channels with shared-randomness assistance, as well as the problem of simulating classical-quantum (CQ) channels with shared-entanglement assistance. In what might be called the algorithmic point of view on Shannon theory [3, 15–19] — and similar as in aforementioned results on channel coding — one has a complete description of the channel to simulate and the goal is then to find good encoding-decoding schemes in order to maximize the success probability within a given message size. More specifically, we aim to design efficient approximation algorithms to obtain near-optimal codes for channel simulation. We then investigate the optimality of these approximation algorithms in terms of asymptotic expansions for the error exponent and the strong converse exponent.

### Main findings

- We propose an efficient algorithm in generating an optimal no-signaling protocol for channel simulation, and rounding it to a near-optimal protocol with solely shared-randomness or entanglement-assistance. This then provides novel lower bounds on the optimal success probability for channel simulation under message-size constraints. Our rounding bounds are quantitatively much stronger than their coding counterparts [15, 17] and this allows for information-theoretic applications.
- We apply our rounding result to determine the exact error exponent as well as the strong converse exponent of shared randomness-assisted classical channel simulation in the worst case total-variation distance. We find that these exponents can be written as simple optimizations over the channel's Rényi-capacities. Our rounding results also give the entanglement-assisted classical-quantum channel simulation error exponent and strong converse exponent in worst-case purified distance. These exponents are tightly termed using the sandwiched Rényi divergences.
- Strikingly, and in stark contrast to channel coding, there are no critical rates for both classical and classical-quantum channel simulation. This then allows for a tight characterization for arbitrary rates below and above the simulation capacity!

**Meta-converse for channel simulation under no-signaling assistance** – The study of channel simulation focuses on the trade-off between the approximation performance and the communication cost of the simulation. We measure the performance by the "success probability", which quantifies the distance between the synthesized and the target channel. In the one-shot setting, our goal is to characterize the maximal success probability for any given target channel $\mathcal{W}$ with the message size of the communication capped by some integer $M$, *i.e.*, we want to quantify

$$\mathrm{Success}^{\mathrm{A}}(\mathcal{W}, M) = \sup_{\widetilde{\mathcal{W}}} \left\{ 1 - \mathrm{dist}(\mathcal{W}, \widetilde{\mathcal{W}}) \,\middle|\, \begin{array}{l} \widetilde{\mathcal{W}} \text{ is the synthesized channel of some A-assisted pro-} \\ \text{tocol using communication of message size } M. \end{array} \right\} \quad (1)$$

where A = SR (shared-randomness) in the classical setting and A = EA (entanglement-assisted) in the CQ setting, and 'dist' is some distance function between channels that itself can be expressed as a linear program (LP) or a semidefinite program (SDP). The latter includes the worst-case total-variation distance / trace distance (which corresponds to the diamond norm) or the worst-case purified distance. These two distances have operational interpretations and have been long studied in (quantum) channel simulations (*e.g.*, see [13, 20, 21]). However, direct computation of such maximal success probability is generally intractable, as it involves optimizing over unbounded encoding and decoding schemes. A more tractable approximation is to allow the sender and receiver to instead share no-signaling correlations [16, 20]; in which case the success probability can be determined by solving an LP (for classical channels) [13, 22] or an SDP (for CQ-channels) [20, 23] as

$$\mathrm{Success}^{\mathrm{NS}}(\mathcal{W}, M) = \sup_{\widetilde{\mathcal{W}}, V} \left\{1 - \mathrm{dist}(\mathcal{W}, \widetilde{\mathcal{W}}) \;\middle|\; \widetilde{\mathcal{W}}_x \preccurlyeq V \;\forall x \in \mathcal{X}, \quad \mathrm{Tr}\,[V] = M\right\} \tag{2}$$

where $\widetilde{\mathcal{W}}$ is the synthesized channel, $\mathcal{X}$ is the input set, and $V$ is some Hermitian operator (or a function when $\mathcal{W}$ is classical) on the output system.

Our first main result is a rounding procedure that converts any no-signaling (NS) solution of (2) into an SR- (when $\mathcal{W}$ is classical) or an EA-strategy (when $\mathcal{W}$ is CQ). We then establish a meta-inequality that bounds the rounded SR/EA-synthesized channel in the (worst-case) Loewner order relative to the original NS-synthesized channel.

**Proposition 1.** *Consider two messages sizes $M$ and $M'$. Suppose $\widetilde{\mathcal{W}}^{\mathrm{NS}}$ is a solution of the program (2) w.r.t. target channel $\mathcal{W}$ and message size $M$. There must exist a strategy $\widetilde{\mathcal{W}}^{\mathrm{A}}$ w.r.t. the message size $M'$, where $\mathrm{A} = \mathrm{SR}$ in the classical setting and $\mathrm{A} = \mathrm{EA}$ in the CQ setting, such that*

$$\widetilde{\mathcal{W}}_x^{\mathrm{A}} \succcurlyeq \left(1 - \left(1 - \frac{1}{M}\right)^{M'}\right) \cdot \widetilde{\mathcal{W}}_x^{\mathrm{NS}} \quad \forall x \in \mathcal{X}. \tag{3}$$

The proof of this proposition relies on a fine analysis of the rejection sampling technique [13, 24]. This meta-inequality allows us to derive the following tight inequalities between the optimal NS- and SR/EA- success probabilities under the (worst-case) trace and purified distances.

**Theorem 2.** *Let $M, M' \in \mathbb{N}$ and $\mathcal{W}$ be a classical or a CQ channel. We have*

$$1 \geq \frac{\mathrm{Success}^{\mathrm{A}}(\mathcal{W}, M')}{\mathrm{Success}^{\mathrm{NS}}(\mathcal{W}, M)} \geq 1 - \left(1 - \frac{1}{M}\right)^{M'}, \quad 1 \geq \frac{\mathrm{Success}^{\mathrm{A}}(\mathcal{W}, M \cdot \log(t))}{\mathrm{Success}^{\mathrm{NS}}(\mathcal{W}, M)} \geq 1 - \frac{1}{t} \quad \forall t > 1 \tag{4}$$

*where $\mathrm{A} = \mathrm{SR}$ in the classical setting and $\mathrm{A} = \mathrm{EA}$ in the CQ setting.*

In particular, letting $M' = M$, we obtain the same approximation factor $(1 - \frac{1}{\mathrm{e}})$ as in [15]. Additionally, using (4), a 0.999 approximation of the success probability can be achieved using only 2 bits of additional communication. Furthermore, we exhibit explicit channels for which the approximation ratio in (4) is asymptotically tight in the classical setting, thereby demonstrating the tightness of the meta-converse in this context.

An immediate consequence of Theorem 2 is that SR (EA in the CQ setting) has the same error and success probability exponents as those obtained with NS-strategies. This leads us to a *novel method for deriving error and strong converse exponents* for channel simulation: one derives the NS exponents first, then apply the rounding bounds obtained via approximation algorithms (*e.g.*, Theorem 2). We believe this *new bottom up approach to Shannon theory* can provide new insights into exponent analysis for other information tasks as well.

As demonstrated in the following examples, the choice of distance measure significantly impacts the channel simulation error and strong converse exponents. In the following, we analyze the exponents with respect to the purified distance in the CQ setting. For the total-variation distance, however, we limit our analysis to the classical setting.

**Application 1: Exponents in purified distance** – Error and strong converse exponents of channel simulation are prominent examples showing the importance of our approximation results. It is known [21] that with a communication rate $r$ above the channel capacity $C(\mathcal{W})$, the optimal simulation error decays exponentially as the number of channel uses tends to infinity, *i.e.*, $\varepsilon(\mathcal{W}^{\otimes n}, e^{nr}) := 1 - \mathrm{Success}(\mathcal{W}^{\otimes n}, e^{nr}) = \exp(-nE(r) + o(n))$, and the decay rate $E(r)$ is known as the error exponent. Similarly, we can show that with a communication rate $r$ below the channel capacity $C(\mathcal{W})$, the optimal simulation error converges to 1 exponentially fast as the number of channel uses grows, *i.e.*,

$1-\varepsilon(\mathcal{W}^{\otimes n}, e^{nr}) = \mathrm{Success}(\mathcal{W}^{\otimes n}, e^{nr}) = \exp(-nS(r)+o(n))$, and we call the convergence rate $S(r)$ the strong converse exponent.

Previously, channel simulation error exponents are only known below a certain critical rate threshold [21]. Given that known bounds on channel coding error exponents also exhibit a critical rate [25–29], one might expect a similar critical rate for channel simulation. However, we show that the exponents for channel simulation can be characterized *for all rates*, revealing a key distinction between channel coding and its reverse task, channel simulation.

**Theorem 3.** *For any classical or CQ channel $\mathcal{W}$, and $r \geq 0$, we have the error exponent and the strong converse exponent for simulating $\mathcal{W}$ with A-assisted communication of rate $r$ as*

$$\lim_{n\to\infty} -\frac{1}{n}\log \varepsilon^{\mathrm{A}}(\mathcal{W}^{\otimes n}, e^{nr}) = \frac{1}{2}\sup_{\alpha\geq 0}\ \alpha\left(r - \widetilde{I}_{\alpha+1}(\mathcal{W})\right), \tag{5}$$

$$\lim_{n\to\infty} -\frac{1}{n}\log \mathrm{Success}^{\mathrm{A}}(\mathcal{W}^{\otimes n}, e^{nr}) = \sup_{\frac{1}{2}\leq\alpha\leq 1}\ \frac{1-\alpha}{\alpha}\left(\widetilde{I}_{\alpha}(\mathcal{W}) - r\right), \tag{6}$$

*respectively, where* $\mathrm{A} \in \{\mathrm{SR}, \mathrm{NS}\}$ *in the classical setting,* $\mathrm{A} \in \{\mathrm{EA}, \mathrm{NS}\}$ *in the CQ setting, and* $\widetilde{I}_{\alpha}(\mathcal{W})$ *is the channel sandwiched mutual information of order $\alpha$.*

For the error exponents, we improve the achievability result of [21], which involves an optimization similar to (5), but with $\alpha \in [0,1]$. Our proof proceeds by showing (5) for no-signaling strategies then deducing the SR/EA-error exponents using Theorem 2.

The strong converse exponents are completely new. While we follow the same procedure of first establishing the NS-exponents before the SR/EA-ones, we find that even the NS-strong converse exponents are challenging to derive. We rely on 'the auxiliary channel' technique introduced by Haroutunian [30]. To bound the smoothed max information, we carefully choose a smoothing channel depending on an auxiliary channel. This auxiliary channel enables us to use Chebyshev-like approximations and invoke a variational representation of the $\alpha$-mutual information.

**Application 2: Exponents in the total-variation distance** – We are able to establish channel simulation error and strong converse exponents in the classical setting in the total-variation distance.

**Theorem 4.** *For any classical channel $\mathcal{W}$ and $r \geq 0$, we have the error exponent and the strong converse exponent for simulating $\mathcal{W}$ with SR-assisted communication of rate $r$ as*

$$\lim_{n\to\infty} -\frac{1}{n}\log \varepsilon^{\mathrm{SR}}(\mathcal{W}^{\otimes n}, e^{nr}) = \sup_{\alpha\geq 0}\ \alpha\left(r - I_{\alpha+1}(\mathcal{W})\right), \tag{7}$$

$$\lim_{n\to\infty} -\frac{1}{n}\log \mathrm{Success}^{\mathrm{SR}}(\mathcal{W}^{\otimes n}, e^{nr}) = \sup_{0\leq\alpha\leq 1}\ (1-\alpha)\left(I_{\alpha}(\mathcal{W}) - r\right), \tag{8}$$

*respectively, where $I_{\alpha}(\mathcal{W})$ is the channel mutual information of order $\alpha$.*

To characterize the NS exponents, we employ the method of types (see, *e.g.*, [25]) along with the post-selection technique [31, 32]. The SR exponents then follows by applying Theorem 2. Notably, we achieve stronger non-asymptotic exponents with only *polynomial prefactors*. Interestingly, the channel simulation achievability bound for the strong converse exponent is similar to the well-known sphere packing bound — the channel coding converse error exponent [25, 30, 33]. Establishing a non-asymptotic bound with polynomial prefactors for the sphere packing bound was achieved using some non-standard large deviations techniques [34, 35]. Therefore, we believe that our proof strategy can be used to analyze the exponents in other information tasks as well.

**Conclusion** – In this work, we proposed an algorithm for no-signaling assisted channel simulation and showed its connection to the entanglement / shared-randomness assisted channel simulation. Such connection turned out to be tight in the large-deviation regime and helped us to derive the channel simulation error and strong converse exponents in both the worst-case purified distance and the total-variation distance. As a surprising result, the expressions for such exponents do not have a critical rate, and hold valid for all communication rates. It is not clear yet whether this conclusion persists in the fully quantum setting since our rounding is currently only good enough for classical-quantum channels and the vanilla convex-split technique does lead to a critical rate [21]. More generally, for the problem of channel interconversion, rounding is challenging as the problem lacks the full symmetry that permits the reduction of the no-signaling program for the success probability. As a starting point in this direction, we are currently working on direct interconversion schemes.

[1] M. Sudan, H. Tyagi, and S. Watanabe, IEEE Transactions on Information Theory **66**, 5 (2019).
[2] C. E. Shannon, The Bell system technical journal **27**, 379 (1948).
[3] Y. Polyanskiy, H. V. Poor, and S. Verdú, IEEE Transactions on Information Theory **56**, 2307 (2010).
[4] M. Hayashi, IEEE Transactions on Information Theory **55**, 4947 (2009).
[5] J. Wolfowitz, *Coding theorems of information theory*, Vol. 31 (Springer Science & Business Media, 2012).
[6] V. Strassen, in *Transactions of the Third Prague Conference on Information Theory, Statistical Decision Functions, Random Processes* (Czechoslovak Academy of Sciences, Prague, 1962) pp. 689–723.
[7] M. Tomamichel and V. Y. Tan, IEEE Transactions on Information Theory **59**, 7041 (2013).
[8] Y. Steinberg and S. Verdú, IEEE Transactions on Information Theory **40**, 634 (1994).
[9] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, IEEE transactions on Information Theory **48**, 2637 (2002).
[10] M. Berta, M. Christandl, and R. Renner, Communications in Mathematical Physics **306**, 579 (2011).
[11] C. H. Bennett, I. Devetak, A. W. Harrow, P. W. Shor, and A. Winter, IEEE Transactions on Information Theory **60**, 2926 (2014).
[12] N. Ramakrishnan, M. Tomamichel, and M. Berta, IEEE Transactions on Information Theory **69**, 5041 (2023).
[13] M. X. Cao, N. Ramakrishnan, M. Berta, and M. Tomamichel, IEEE Transactions on Information Theory **70**, 6780 (2024).
[14] M. X. Cao, R. Jain, and M. Tomamichel, in *2024 IEEE International Symposium on Information Theory (ISIT)* (IEEE, 2024) pp. 1421–1425.
[15] S. Barman and O. Fawzi, IEEE Transactions on Information Theory **64**, 1038 (2017).
[16] W. Matthews, IEEE Transactions on Information Theory **58**, 7036 (2012).
[17] O. Fawzi, J. Seif, and D. Szilágyi, in *2019 IEEE International Symposium on Information Theory (ISIT)* (IEEE, 2019) pp. 2569–2573.
[18] O. Fawzi and P. Fermé, IEEE Transactions on Information Theory **70**, 1693 (2024).
[19] O. Fawzi and P. Fermé, IEEE Transactions on Information Theory (2024).
[20] K. Fang, X. Wang, M. Tomamichel, and M. Berta, IEEE Transactions on Information Theory **66**, 2129 (2019).
[21] K. Li and Y. Yao, arXiv preprint arXiv:2112.04475 (2021).
[22] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter, IEEE Transactions on Information Theory **57**, 5509 (2011).
[23] M. Berta, O. Fawzi, and A. Oufkir, in *2024 IEEE International Symposium on Information Theory (ISIT)* (IEEE, 2024) pp. 1209–1214.
[24] V. Neumann, Notes by GE Forsythe , 36 (1951).
[25] I. Csiszár and J. Körner, *Information theory: coding theorems for discrete memoryless systems* (Cambridge University Press, Cambridge, England, UK, 2011).
[26] M. Dalai, IEEE Transactions on Information Theory **59**, 8027 (2013).
[27] H.-C. Cheng, L. Gao, and M. Berta, arXiv preprint arXiv:2304.12056 (2023).
[28] J. M. Renes, arXiv preprint arXiv:2407.11118 (2024).
[29] K. Li and D. Yang, arXiv preprint arXiv:2407.12403 (2024).
[30] E. A. Haroutunian, M. E. Haroutunian, A. N. Harutyunyan, *et al.*, Foundations and Trends® in Communications and Information Theory **4**, 97 (2008).
[31] M. Christandl, R. König, and R. Renner, Physical review letters **102**, 020504 (2009).
[32] M. Tomamichel and M. Hayashi, IEEE Transactions on Information Theory **64**, 1064 (2017).
[33] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp, Information and Control **10**, 65 (1967).
[34] Y. Altuğ and A. B. Wagner, IEEE Transactions on Information Theory **60**, 1592 (2014).
[35] H.-C. Cheng, M.-H. Hsieh, and M. Tomamichel, IEEE Transactions on Information Theory **65**, 2872 (2019).