Optimality of meta-converse for channel simulation

Mario Berta
Institute for Quantum Information,
RWTH Aachen University.
Aachen, Germany

Omar Fawzi
Univ Lyon, Inria,
ENS Lyon, UCBL, LIP.
Lyon, France

Aadil Oufkir
Institute for Quantum Information,
RWTH Aachen University.
Aachen, Germany
Email: oufkir@physik.rwth-aachen.de

***Abstract*—We study the effect of shared non-signaling correlations for the problem of simulating a channel using noiseless communication in the one-shot setting. For classical channels, we show how to round any non-signaling-assisted simulation strategy — which exactly corresponds to the meta-converse for channel simulation — to a strategy that only uses shared randomness. For quantum channels, we round any non-signaling-assisted simulation strategy to a strategy that only uses shared entanglement. As our main result, we prove a guarantee on the ratio of success probabilities of at least $(1-\frac{1}{\mathbf{e}})$, for both the classical and the quantum setting. We further show this ratio to be optimal. It can be improved to $(1-\frac{1}{t})$ using $O(\ln\ln(t))$ additional bits (qubits) of communication.**

I. Introduction

Channel simulation is the task of simulating a noisy channel using a noiseless channel (e.g., [1]–[6]). As such, it is the reverse task of channel coding [7]. As for channel coding, in the i.i.d. setting, the channel's capacity characterizes the minimal communication rate needed to simulate this channel when shared-randomness is allowed [2], [8], [9].

In the work [10] on classical channel coding, a simple and efficient algorithm is designed to return a code achieving a $(1-\frac{1}{\mathbf{e}})$-approximation of the maximum success probability that can be attained for a noisy classical channel. This algorithm is based on a relaxation of the problem where the sender and the receiver share non-signaling correlations — which exactly corresponds [11] to the well-known PPV meta-converse [12], [13].[1] The approximation ratio $(1-\frac{1}{\mathbf{e}})$ is tight and it is shown that it is NP-hard to achieve a strictly better ratio [10]. Moreover, some approximation algorithms are proposed for computing the maximum success probability of classical-quantum (CQ) channels [14], (deterministic) broadcast channels [15] and multiple access channels [16]. Finally, for quantum channel coding, finding optimal approximation algorithms for the success probability remains completely open.

In this work, we consider the task of simulating classical channels with shared-randomness and quantum channels with shared-entanglement in the one-shot setting. In the algorithmic point of view [10]–[12], [14]–[16], we have a complete description of the channel we want to simulate and we aim to find an optimal encoding-decoding scheme in order to maximize the success probability for a fixed communication size. Since we are interested in designing efficient algorithms to obtain near optimal codes for simulating classical channels, we consider the meta-converse for channel simulation which has been shown to lead to the correct high order refinements [5] (e.g., optimal bounds for coding in the finite block-length setting up to the second-order). Here, we show that the meta-converse is equally useful in the one-shot setting:

- Following [5], [17]–[22], we consider the meta-converse for channel simulation obtained by allowing non-signaling correlations between the sender and the receiver. When allowing non-signaling correlations, the optimal success probability for the simulation becomes a linear program (LP).
- This efficient LP meta-converse provides an upper bound on the success probability, as non-signaling correlations include shared-randomness. Our main result is to show that this bound gives an $(1-\frac{1}{\mathbf{e}})$-approximation of the maximum success probability and that this approximation ratio is tight. For this, we round the solution of the non-signaling program to a shared-randomness strategy for simulating a classical channel with the same communication size. After reducing the non-signaling program using symmetry, we use the rejection sampling technique [23] to simulate an optimal non-signaling channel using an optimal probability distribution that appears in the non-signaling program. Moreover, we exhibit channels for which the approximation ratio is arbitrarily close to $(1-\frac{1}{\mathbf{e}})$.
- If an additional $\ln(t)$-amount of communication is allowed, then we obtain a better approximation ratio $(1-\frac{1}{t})$. In particular, this allows us to prove the fact that the non-signaling simulation capacity of a classical channel is exactly the same as with shared-randomness [9].
- We generalize the previous approximation results to the quantum setting, where we compare shared-entanglement with non-signaling strategies using quantum rejection sampling [24].

Notations.

- For a positive integer $N\in\mathbb{N}$, we denote by $[N]$ the set of integers between 1 and N.
- The total variation (TV) distance between two probability distributions p and q on $[n]$ is

$$\|p-q\|_{\mathrm{TV}}=\frac{1}{2}\sum_{i=1}^{n}|p_i-q_i|.$$

[1]These results are stated in terms of the code size M for a fixed error probability ε. In this article, we are interested in maximizing the success probability $1-\varepsilon$ for a fixed given code size M.

- The one norm of a matrix M is defined as $\|M\|_1 = \mathrm{Tr}\,[|M|]$ where $|M| = \sqrt{MM^\dagger}$.
- The trace norm between two quantum states ρ and σ is defined as $\|\rho - \sigma\|_{\mathrm{tr}} = \frac{1}{2}\|\rho - \sigma\|_1$.
- The diamond distance between two quantum channels $\mathcal{N}$ and $\mathcal{M}$ is defined as

$$\|\mathcal{N} - \mathcal{M}\|_\diamond = \sup_{\sigma \text{ state}} \|(\mathrm{id} \otimes \mathcal{N})(\sigma) - (\mathrm{id} \otimes \mathcal{M})(\sigma)\|_{\mathrm{tr}}.$$

- For a quantum channel $\mathcal{N} : A \to B$, we define the Choi matrix $J_{\mathcal{N}} = \sum_{i,j=1}^{|A|} |i\rangle \langle j|_{A'} \otimes \mathcal{N}_{A\to B}(|i\rangle \langle j|_A) = (\mathrm{id} \otimes \mathcal{N}_{A\to B})(|w\rangle\langle w|_{A'A})$ where $|w\rangle = \sum_{i=1}^{|A|} |i\rangle_{A'} |i\rangle_A$.
- $A \succcurlyeq B$ stands for $A - B$ positive semi-definite.
- $\mathcal{N} \succcurlyeq_{\mathrm{CP}} \mathcal{M}$ stands for $\mathcal{N} - \mathcal{M}$ completely positive.

II. Simulation of classical channels

Given a classical channel $W_{Y|X}$ of input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$ and an integer M, our goal is to simulate the channel W using a classical communication of at most M distinct messages and with an error probability as small as possible. More formally, we can describe a size M ε-simulation code for the channel $W_{Y|X}$ by a triple $(\{\mathcal{E}_s\}_{s\in\mathcal{S}}, \{\mathcal{D}_s\}_{s\in\mathcal{S}}, \mathcal{S})$ such that the synthesized channel

$$\widetilde{W}_{Y|X}(y|x) = \sum_{s\in\mathcal{S}} p_S(s) \sum_{i=1}^{M} \mathcal{E}_s(i|x)\mathcal{D}_s(y|i)$$

is ε close to the actual channel $W_{Y|X}$ in the worst case total variation distance. Here, it is allowed to use a shared random variable S on an arbitrary discrete set $\mathcal{S}$. This choice is motivated by the fact that strategies, assisted by shared-randomness (SR), are proven to achieve optimal channel simulation capacity [5], whereas strategies without assistance require a simulation rate strictly greater than the capacity [2], [9]. Allowing shared-randomness of arbitrary size makes the problem of computing the success probability computationally even harder. In the following, we describe a possible approximation algorithm given by the meta-converse.

SR strategies are special case of the so called non-signaling strategies. In this latter, the joint encoder-decoder map $N_{IY|XJ}$ satisfies

$$\begin{aligned} N_{IY|XJ}(i, y|x, j) &\geq 0 && \forall i, y, x, j, \\ \sum_{i,y} N_{IY|XJ}(i, y|x, j) &= 1 && \forall x, j, \\ \sum_{i} N_{IY|XJ}(i, y|x, j) &= N_{Y|J}(y|j) && \forall x, \\ \sum_{y} N_{IY|XJ}(i, y|x, j) &= N_{I|X}(i|x) && \forall j. \end{aligned}$$

We denote the set of non-signaling maps by $\mathcal{NS}(IY|XJ)$. Non-signaling strategies prove to be useful for simplifying the computation of the maximal success probability which is obtained by solving the following program

$$\begin{aligned} \mathrm{Success}^{\mathrm{NS}}(W, M) = \max_{N}\ & 1 - \sup_{x\in\mathcal{X}} \left\|\widetilde{W}(\cdot|x) - W(\cdot|x)\right\|_{\mathrm{TV}} \\ \text{s.t.}\ & N \in \mathcal{NS}(IY|XJ), \\ & \widetilde{W}(y|x) = \sum_{i=1}^{M} N(i, y|x, i). \end{aligned} \tag{1}$$

So, when we relax the constraints of the encoder-decoder to be non-signaling, we obtain a linear program. After a symmetry-based reduction, the program (1) becomes [5], [21], [25] (see [26] for a proof)

$$\begin{aligned} \mathrm{Success}^{\mathrm{NS}}(W, M) = \max_{\widetilde{W}, \zeta}\ & 1 - \sup_{x\in\mathcal{X}} \left\|\widetilde{W}(\cdot|x) - W(\cdot|x)\right\|_{\mathrm{TV}} \\ \text{s.t.}\ & \sum_y \widetilde{W}(y|x) = 1 \quad \forall x, \\ & \widetilde{W}(y|x) \geq 0 \quad \forall x, y, \\ & \widetilde{W}(y|x) \leq \zeta(y) \quad \forall x, y, \\ & \sum_y \zeta(y) = M. \end{aligned} \tag{2}$$

Observe that a strategy assisted by shared-randomness $N(iy|xj) = \sum_{s\in\mathcal{S}} p_S(s) \sum_{i=1}^{M} \mathcal{E}_s(i|x)\mathcal{D}_s(y|j)$ is non-signaling, so the non-signaling (NS) simulation success probability is greater than the shared-randomness (SR) simulation success probability

$$\mathrm{Success}^{\mathrm{NS}}(W, M) \geq \mathrm{Success}^{\mathrm{SR}}(W, M).$$

These values might be different in general. For the channel [17]

$$W : \binom{M^2}{M} \to \{1, 2, \ldots, M^2\},\ W(y|x) = \frac{1}{M}\,\mathbf{1}\{y \in x\}$$

we show that (see [26] for a proof)

$$\frac{\mathrm{Success}^{\mathrm{SR}}(W, M)}{\mathrm{Success}^{\mathrm{NS}}(W, M)} \underset{M\to\infty}{\longrightarrow} 1 - \frac{1}{\mathbf{e}}.$$

Since non-signaling strategies are computationally accessible and shared-randomness strategies are of interest, a natural question arises:

How large is the gap between the success probabilities of these strategies?

In the following, we round the solution of the non-signaling program (2) to a strategy that requires only shared-randomness. This allows us to prove a meta inequality between the channel using non-signaling resources and the constructed channel using only shared randomness.

Proposition 1. *Let $M' \in \mathbb{N}$. Let $\widetilde{W}^{\mathrm{NS}}$ be a feasible solution of the program (2). There exists $\widetilde{W}^{\mathrm{SR}}$, a strategy of size M' assisted by shared-randomness, that satisfies*

$$\widetilde{W}^{\mathrm{SR}}(y|x) \geq \left[1 - \left(1 - \frac{1}{M}\right)^{M'}\right] \widetilde{W}^{\mathrm{NS}}(y|x) \quad \forall x, y.$$

To construct the strategy $\widetilde{W}^{\mathrm{SR}}$ with shared-randomness, we use a standard tool from statistics [27] that has been applied previously to channel simulation [5], [22], [28] namely the *rejection sampling technique* [23].

Proof of Prop. 1. Let $(\widetilde{W}^{\mathrm{NS}}, \zeta)$ be a feasible solution of the program (2).

a) Shared randomness: Since $\sum_y \zeta(y) = M$ and for all y, $\zeta(y) \geq 0$, $\{\frac{\zeta(y)}{M}\}_y$ is a probability distribution. Let $\mathbf{Y_1}, \dots, \mathbf{Y_{M'}}$ be i.i.d. samples from $p_{\mathrm{initial}} = \{\frac{\zeta(y)}{M}\}_y$. This is the shared-randomness.

b) Encoding: For an input x, we run the rejection sampling algorithm [5] with M' steps for $p_{\mathrm{initial}} = \{\frac{\zeta(y)}{M}\}_y$ and $p_{\mathrm{target}} = \widetilde{W}^{\mathrm{NS}}(\cdot|x)$ and obtain $\widetilde{\mathbf{Y}} = \mathbf{Y_i}$ where $\mathbf{i}$ is the first index (or M' if it does not exist) such that

$$\mathbf{U_i} \leq \frac{1}{M} \cdot \frac{p_{\mathrm{target}}(\mathbf{Y_i})}{p_{\mathrm{initial}}(\mathbf{Y_i})} = \frac{\widetilde{W}^{\mathrm{NS}}(\mathbf{Y_i}|x)}{\zeta(\mathbf{Y_i})},$$

where $\mathbf{U_i} \sim \mathrm{Unif}([0,1])$, then encode $\mathcal{E}_s(m|x) = \delta_{m=\mathbf{i}}$.

c) Decoding: Decode as $\mathcal{D}_s(y|m) = \delta_{y=\mathbf{Y_m}}$. For finite M', we follow [5] and show, with a simple calculation, that the distribution of $\widetilde{\mathbf{Y}}$ satisfies for $y \in \mathcal{Y}$ that

$$p_{\widetilde{\mathbf{Y}}}(y) = \left(1 - \frac{1}{M}\right)^{M'-1} \left(p_{\mathrm{initial}}(y) - \frac{1}{M} \cdot p_{\mathrm{target}}(y)\right) + \left[1 - \left(1 - \frac{1}{M}\right)^{M'}\right] p_{\mathrm{target}}(y).$$

Indeed, let us denote $\mathbf{E}$ the binary random variable such that $\mathbf{E} = 1$ if and only if there is an $\mathbf{i} \leq M'$ such that $\mathbf{U_i} \leq \frac{1}{M} \cdot \frac{p_{\mathrm{target}}(\mathbf{Y_i})}{p_{\mathrm{initial}}(\mathbf{Y_i})}$. We have by letting $\lambda = \frac{1}{M}$ that

$$\begin{aligned} &p_{\widetilde{\mathbf{Y}}\mathbf{E}}(y, 0) \\ &= \sum_{y_1, \dots, y_{M'-1}} \prod_{i=1}^{M'-1} p_{\mathrm{initial}}(y_i) \left(1 - \lambda \cdot \frac{p_{\mathrm{target}}(y_i)}{p_{\mathrm{initial}}(y_i)}\right) \\ &\qquad \cdot p_{\mathrm{initial}}(y) \left(1 - \lambda \cdot \frac{p_{\mathrm{target}}(y)}{p_{\mathrm{initial}}(y)}\right) \\ &= \prod_{i=1}^{M'-1} \sum_{y_i} \left(p_{\mathrm{initial}}(y_i) - \lambda \cdot p_{\mathrm{target}}(y_i)\right) \\ &\qquad \cdot \left(p_{\mathrm{initial}}(y) - \lambda \cdot p_{\mathrm{target}}(y)\right) \\ &= (1-\lambda)^{M'-1} \left(p_{\mathrm{initial}}(y) - \lambda \cdot p_{\mathrm{target}}(y)\right). \end{aligned}$$

Similarly, we get

$$\begin{aligned} &p_{\widetilde{\mathbf{Y}}\mathbf{E}}(y, 1) \\ &= \sum_{j=1}^{M'} \sum_{y_1, \dots, y_{j-1}} \prod_{i=1}^{j-1} p_{\mathrm{initial}}(y_i) \left(1 - \lambda \cdot \frac{p_{\mathrm{target}}(y_i)}{p_{\mathrm{initial}}(y_i)}\right) \\ &\qquad \cdot p_{\mathrm{initial}}(y) \left(\lambda \cdot \frac{p_{\mathrm{target}}(y)}{p_{\mathrm{initial}}(y)}\right) \\ &= \sum_{j=1}^{M'} (1-\lambda)^{j-1} \cdot \lambda \cdot p_{\mathrm{target}}(y) \\ &= \left(1 - (1-\lambda)^{M'}\right) p_{\mathrm{target}}(y). \end{aligned}$$

Hence, we find

$$\begin{aligned} p_{\widetilde{\mathbf{Y}}}(y) &= p_{\widetilde{\mathbf{Y}}\mathbf{E}}(y, 0) + p_{\widetilde{\mathbf{Y}}\mathbf{E}}(y, 1) \\ &= \left(1 - \frac{1}{M}\right)^{M'-1} \left(p_{\mathrm{initial}}(y) - \frac{1}{M} \cdot p_{\mathrm{target}}(y)\right) \\ &\quad + \left[1 - \left(1 - \frac{1}{M}\right)^{M'}\right] p_{\mathrm{target}}(y). \end{aligned}$$

Moreover, the second constraint of the program (2) reads

$$\forall y: \ p_{\mathrm{initial}}(y) = \frac{\zeta(y)}{M} \geq \frac{1}{M} \widetilde{W}^{\mathrm{NS}}(y|x) = \frac{1}{M} \cdot p_{\mathrm{target}}(y).$$

Therefore, we get

$$\begin{aligned} p_{\widetilde{\mathbf{Y}}}(y) &= \left(1 - \frac{1}{M}\right)^{M'-1} \left(p_{\mathrm{initial}}(y) - \frac{1}{M} \cdot p_{\mathrm{target}}(y)\right) \\ &\quad + \left[1 - \left(1 - \frac{1}{M}\right)^{M'}\right] p_{\mathrm{target}}(y) \\ &\geq \left[1 - \left(1 - \frac{1}{M}\right)^{M'}\right] p_{\mathrm{target}}(y). \end{aligned}$$

By observing that $p_{\widetilde{\mathbf{Y}}}(y) = \widetilde{W}^{\mathrm{SR}}(y|x)$, where $\widetilde{W}^{\mathrm{SR}}(y|x) = \sum_{s \in \mathcal{S}} p_S(s) \sum_{m=1}^{M'} \mathcal{E}_s(m|x) \mathcal{D}_s(y|m)$, we deduce the required inequality. □

As a direct corollary of Prop. 1, we can control the gap between the success probabilities of strategies assisted by shared-randomness and non-signaling boxes. Although we focus on the worst-case total variation distance, it is noteworthy that Prop. 1 also permits to control the gap between the success probabilities of SR and NS strategies under average-case total variation, as well as under average and worst-case Bhattacharyya distance.

Corollary 1.1. *Let $M, M' \geq 1$ and W be a channel. We have*

$$1 \geq \frac{\mathrm{Success}^{\mathrm{SR}}(W, M')}{\mathrm{Success}^{\mathrm{NS}}(W, M)} \geq \left[1 - \left(1 - \frac{1}{M}\right)^{M'}\right].$$

In particular, when $M' = M$, the gap between $\mathrm{Success}^{\mathrm{SR}}(W, M)$ and $\mathrm{Success}^{\mathrm{NS}}(W, M)$ is at most $1 - \frac{1}{\mathbf{e}}$. Furthermore, by choosing $M' = \ln(t)M$, we can show that the gap is at most $1 - \frac{1}{t}$ and thus approaches 1 as $t \to \infty$. This implies the fact that the non-signaling assistance does not help to reduce the asymptotic simulation capacity of a classical channel [5], [9] (see [26] for a proof)

$$C^{\mathrm{SR}}(W) = C^{\mathrm{EA}}(W) = C^{\mathrm{NS}}(W).$$

Finally, as shown in [26], the bounds of Cor. 1.1 are tight. These simulation results are analogous to the known rounding results for channel coding [10]. However, the techniques needed are different. In fact, the simulation meta-converse is related to the smoothed max-divergence [21], while the coding PPV meta-converse can be phrased in terms of the hypothesis testing divergence [11].

Proof of Cor. 1.1. Observe that for two probability distributions p and q, $1-\|p-q\|_{\mathrm{TV}}=\sum_i \min(p_i,q_i)$. Let $x\in\mathcal{X}$, by Prop. 1 we have

$$\begin{aligned}1 &- \left\|\widetilde{W}^{\mathrm{SR}}(\cdot|x)-W(\cdot|x)\right\|_{\mathrm{TV}}\\ &= \sum_y \min\left(\widetilde{W}^{\mathrm{SR}}(y|x),\ W(y|x)\right)\\ &\geq \sum_y\left[1-\left(1-\frac{1}{M}\right)^{M'}\right]\min\left(\widetilde{W}^{\mathrm{NS}}(y|x),\ W(y|x)\right)\\ &= \left[1-\left(1-\frac{1}{M}\right)^{M'}\right]\left(1-\left\|\widetilde{W}^{\mathrm{NS}}(\cdot|x)-W(\cdot|x)\right\|_{\mathrm{TV}}\right),\end{aligned}$$

where we used $\left(1-(1-1/M)^{M'}\right)\leq 1$. Choosing the optimal feasible solution $\widetilde{W}^{\mathrm{NS}}$ of the program (2) and taking the minimum on x on both sides give the desired inequality. □

III. Simulation of quantum channels

We are able to provide approximation algorithms in the quantum setting as well. Note that, in contrast, the corresponding question for quantum channel coding remains open. In this section, we generalize the rounding results of Section II to the simulation of quantum channels with quantum communication and with entanglement-assisted strategies (EA). EA strategies are more natural in the quantum setting and are strictly more powerful for quantum channels than shared-randomness strategies [9].

Formally, a size-M entanglement-assisted ε-simulation code for $\mathcal{W}_{A\to B}$ is a triple $(\mathcal{E}_{A_iE_A\to A_o},\mathcal{D}_{E_BB_i\to B_o},\sigma_{E_AE_B})$ such that $|A_o|=|B_i|=M$, $A_i=A$, $B_o=B$ and the synthesized channel

$$\widetilde{\mathcal{W}}_{A\to B}(\rho)=\mathcal{D}_{E_BB_i\to B}\,\mathrm{id}^q_{A_o\to B_i}\,\mathcal{E}_{AE_A\to A_o}(\rho_A\otimes\sigma_{E_AE_B})$$

is ε close to the channel $\mathcal{W}_{A\to B}$ in the diamond distance, where $\mathrm{id}_q=\mathrm{id}^q_{A_o\to B_i}$ denotes the quantum identity channel of dimension M. For a fixed size M, maximizing the success probability $1-\varepsilon$ of EA strategies is computationally hard. In order to approximate the success probability, we use the meta-converse obtained by considering non-signaling strategies Π whose Choi matrices satisfy [21]

$$\begin{aligned}&J_\Pi\succcurlyeq 0,\quad \mathrm{Tr}_{A_oB_o}[J_\Pi]=\mathbb{I}_{A_iB_i} && \text{(CP), (TP)},\\ &\mathrm{Tr}_{A_o}[J_\Pi]=\frac{\mathbb{I}_{A_i}}{|A_i|}\otimes\mathrm{Tr}_{A_oA_i}[J_\Pi] && (A\not\to B),\\ &\mathrm{Tr}_{B_o}[J_\Pi]=\frac{\mathbb{I}_{B_i}}{|B_i|}\otimes\mathrm{Tr}_{B_iB_o}[J_\Pi] && (B\not\to A).\end{aligned}$$

Using [21, Cor. 2] or [25], the non-signaling program can be written as the following semi-definite program (SDP)

$$\begin{aligned}\mathrm{Success}^{\mathrm{NS}}(\mathcal{W},M)=\max_{\widetilde{\mathcal{W}}_{A\to B},V_B}\ & 1-\left\|\mathcal{W}_{A\to B}-\widetilde{\mathcal{W}}_{A\to B}\right\|_\diamond\\ \text{s.t.}\quad & \widetilde{\mathcal{W}}_{A\to B}\quad\text{quantum channel},\\ & J_{\widetilde{\mathcal{W}}}\preccurlyeq \mathbb{I}_{A'}\otimes V_B,\\ & \mathrm{Tr}[V_B]=M^2.\end{aligned}\tag{3}$$

Notably, when only classical instead of quantum communication is allowed, with the classical identity channel

$$\mathrm{id}_c(\rho)=\mathrm{id}^c_{A_o\to B_i}(\rho)=\sum_{x=1}^N\langle x|\,\rho\,|x\rangle\,|x\rangle\!\langle x|,$$

the non-signaling program is similar to (3) except that M^2 is replaced by N (see [26] for a proof). Since EA strategies are non-signaling, the success probability of NS strategies is at least as its EA counterparts

$$\mathrm{Success}^{\mathrm{NS}}(\mathcal{W},M)\geq\mathrm{Success}^{\mathrm{EA}}(\mathcal{W},M).$$

Now, we can inquire about the extent to which these quantities can differ. Importantly, the answer to this question is similar to Cor. 1.1 although the context and the proof strategies are slightly different. We round a feasible solution of the program (3) to an EA strategy. Then we compare the distance between the resulting channel of this strategy to the feasible non-signaling channel. We prove a meta-inequality between the EA and the NS channels that can be of independent interest.

Proposition 2. *Let $M,M'\in\mathbb{N}$. Let $\widetilde{\mathcal{W}}^{\mathrm{NS}}$ be a feasible solution of the program* (3)*. There exists $\widetilde{\mathcal{W}}^{\mathrm{EA}}$, an EA strategy of size M', that satisfies*

$$\widetilde{\mathcal{W}}^{\mathrm{EA}}\succcurlyeq_{\mathrm{CP}}\left[1-\left(1-\frac{1}{M^2}\right)^{M'^2}\right]\widetilde{\mathcal{W}}^{\mathrm{NS}}.$$

The proof is based on a quantum analog of rejection sampling technique similar to the convex split technique (e.g., [24], [28]–[30]).

Proof of Prop. 2. Let $(\widetilde{\mathcal{W}}^{\mathrm{NS}}_{A\to B},V_B)$ be a feasible solution of the program (3). A bipartite state $|\phi\rangle\!\langle\phi|_{A'A}$ can be written as

$$|\phi\rangle_{A'A}=(O_\phi\otimes\mathbb{I}_A)\,|w\rangle_{A'A}\,,\text{ where }|w\rangle_{A'A}=\sum_{i=1}^{|A|}|i\rangle_{A'}\,|i\rangle_A\,.$$

The state $|\phi\rangle\!\langle\phi|_{A'A}$ is pure so $\langle\phi|\phi\rangle=1$ implying

$$\mathrm{Tr}\left[O_\phi O_\phi^\dagger\right]=\langle w|\,(O_\phi^\dagger O_\phi\otimes\mathbb{I}_A)\,|w\rangle=\langle\phi|\phi\rangle=1.$$

The constraints of the program (3) imply

$$\begin{aligned}0&\preccurlyeq\mathrm{id}\otimes\widetilde{\mathcal{W}}^{\mathrm{NS}}_{A\to B}(|\phi\rangle\!\langle\phi|_{A'A})\\ &=\mathrm{id}\otimes\widetilde{\mathcal{W}}^{\mathrm{NS}}_{A\to B}((O_\phi\otimes\mathbb{I})\cdot|w\rangle\!\langle w|\cdot(O_\phi^\dagger\otimes\mathbb{I}))\\ &=(O_\phi\otimes\mathbb{I})\cdot J_{\widetilde{\mathcal{W}}^{\mathrm{NS}}}\cdot(O_\phi^\dagger\otimes\mathbb{I})\\ &\preccurlyeq(O_\phi\otimes\mathbb{I})\cdot(\mathbb{I}_{A'}\otimes V_B)\cdot(O_\phi^\dagger\otimes\mathbb{I})\\ &=(O_\phi O_\phi^\dagger\otimes V_B).\end{aligned}$$

Since $\mathrm{Tr}[V_B]=M^2$ and $\mathrm{Tr}\left[O_\phi O_\phi^\dagger\right]=1$, the matrix $O_\phi O_\phi^\dagger\otimes\frac{V_B}{M^2}$ is a quantum state satisfying for all unit vectors $|\phi\rangle$ that

$$O_\phi O_\phi^\dagger\otimes\frac{V_B}{M^2}\succcurlyeq\left(\frac{1}{M^2}\right)\cdot\mathrm{id}\otimes\widetilde{\mathcal{W}}^{\mathrm{NS}}_{A\to B}(|\phi\rangle\!\langle\phi|_{A'A}).$$

Consider the following scheme inspired by [28].

a) Shared entanglement: Let $|V\rangle_{E_AE_B}$ be a purification of $\frac{V_B}{M^2}$ ($|E_A| = |E_B| = |B|$). $|V\rangle^{\otimes M'^2}_{E_AE_B}$ is a part of the shared-entanglement between Alice and Bob.

b) Encoding: For an input state $|\phi\rangle\langle\phi|$, we have

$$O_\phi O_\phi^\dagger \otimes \frac{V_B}{M^2} \succcurlyeq \left(\frac{1}{M^2}\right) \cdot \mathrm{id} \otimes \widetilde{\mathcal{W}}^{\mathrm{NS}}_{A\to B}(|\phi\rangle\langle\phi|_{A'A})$$

so we can write for some state $\zeta_{A'B}$ that

$$O_\phi O_\phi^\dagger \otimes \frac{V_B}{M^2} = (1-\lambda)\,\sigma_{A'B} + \lambda\zeta_{A'B},$$

where $\lambda = 1 - \frac{1}{M^2}$ and $\sigma = \mathrm{id} \otimes \widetilde{\mathcal{W}}^{\mathrm{NS}}_{A\to B}(|\phi\rangle\langle\phi|_{A'A})$. Consider a purification of $O_\phi O_\phi^\dagger \otimes \frac{V_B}{M^2}$ as

$$|OV\rangle = \sqrt{1-\lambda}\,|0\rangle_S\,|\sigma\rangle_{E_AF_AE_BF_B} + \sqrt{\lambda}\,|1\rangle_S\,|\zeta\rangle_{E_AF_AE_BF_B},$$

where $|F_A| = |F_B| = |A'| = |A|$. Since

$$\begin{aligned}\mathrm{Tr}_{SE_AF_AA'}\left[|OV\rangle\langle OV|\right] &= (1-\lambda)\,\sigma_B + \lambda\zeta_B\\ &= \frac{V_B}{M^2} = \mathrm{Tr}_{E_A}\left[|V\rangle\langle V|_{E_AE_B}\right],\end{aligned}$$

by Uhlmann's theorem [31], there is an isometry $\mathcal{I}$ such that,

$$\mathcal{I}_{E_A\to SE_AF_AA'} \cdot \left(|V\rangle^{\otimes M'^2}_{E_AE_B}\right) = |OV\rangle^{\otimes M'^2}_{SE_AF_AE_BF_B}.$$

Alice can then apply the isometry $\mathcal{I}_{E_A\to SE_AF_AA'}$ on her part of the shared entangled state $|V\rangle^{\otimes M'^2}_{E_AE_B}$. The resulting state is,

$$\begin{aligned}&|OV\rangle^{\otimes M'^2}\\ &= \sum_{x\in\{0,1\}^{M'^2}} \sqrt{1-\lambda}^{|\bar{x}|}\sqrt{\lambda}^{|x|}\,|x\rangle_{S^{M'^2}} \otimes |\sigma\rangle^{\otimes|\bar{x}|} \otimes |\zeta\rangle^{\otimes|x|},\end{aligned}$$

where $|x| = M'^2 - |\bar{x}| = \sum_i x_i$. Finally, Alice measures the system S and observes $\mathbf{x} = x \in \{0,1\}^{M'^2}$ with probability $(1-\lambda)^{|\bar{x}|}(\lambda)^{|x|}$. The probability that $\mathbf{x}$ contains '0' is

$$\mathbb{P}\left[0 \in \mathbf{x}\right] = 1 - \mathbb{P}\left[\mathbf{x} = 1\cdots1\right] = 1 - \lambda^{M'^2}.$$

Let $\mathbf{I}$ be the index of the first '0' in $\mathbf{x}$ if $\mathbf{x} \neq 1\cdots1$ and M'^2 otherwise. Alice sends $\mathbf{I} \in [M'^2]$ to Bob using super dense coding [32] and the quantum identity channel of dimension M'.

c) Decoding: Bob returns the $\mathbf{I}$'s (post-measurement) copy of E_BF_B. This state is denoted $\mathrm{id} \otimes \widetilde{\mathcal{W}}^{\mathrm{EA}}_{\mathbf{I}}(|\phi\rangle\langle\phi|)$ with

$$\begin{aligned}&\mathrm{id} \otimes \widetilde{\mathcal{W}}^{\mathrm{EA}}(|\phi\rangle\langle\phi|)\\ &= \mathbb{E}_{\mathbf{I}}\left[\mathrm{id} \otimes \widetilde{\mathcal{W}}^{\mathrm{EA}}_{\mathbf{I}}(|\phi\rangle\langle\phi|)\right]\\ &= \sum_{x\neq1\cdots1} \mathbb{P}\left[\mathbf{x} = x\right]\sigma_{E_BF_B} + \mathbb{P}\left[\mathbf{x} = 1\cdots1\right]\zeta_{E_BF_B}\\ &= \left(1-\lambda^{M'^2}\right)\mathrm{id} \otimes \widetilde{\mathcal{W}}^{\mathrm{NS}}_{A\to B}(|\phi\rangle\langle\phi|_{A'A}) + \lambda^{M'^2}\zeta_{A'B}\\ &\succcurlyeq \left[1 - \left(1 - \frac{1}{M^2}\right)^{M'^2}\right]\mathrm{id} \otimes \widetilde{\mathcal{W}}^{\mathrm{NS}}_{A\to B}(|\phi\rangle\langle\phi|_{A'A}).\end{aligned}$$

In [26], it is shown that the state $\left(1-\lambda^{M'^2}\right)\sigma_{E_BF_B} + \lambda^{M'^2}\zeta_{E_BF_B}$ can be written as $\mathrm{id} \otimes \widetilde{\mathcal{W}}^{\mathrm{EA}}(|\phi\rangle\langle\phi|)$ for a quantum channel $\widetilde{\mathcal{W}}^{\mathrm{EA}}$. □

As a consequence of Prop. 2, we can control the gap between the EA and NS success probabilities under the diamond distance. We note that a similar statement is implied by Prop. 2 for fidelity-based distances as well.

Corollary 2.1. *Let $M, M' \geq 1$ and $\mathcal{W}$ be a quantum channel. We have*

$$1 \geq \frac{\mathrm{Success}^{\mathrm{EA}}(\mathcal{W}, M')}{\mathrm{Success}^{\mathrm{NS}}(\mathcal{W}, M)} \geq \left[1 - \left(1 - \frac{1}{M^2}\right)^{M'^2}\right].$$

Similar to the classical case, when $M' = M$, the gap between $\mathrm{Success}^{\mathrm{EA}}(\mathcal{W}, M)$ and $\mathrm{Success}^{\mathrm{NS}}(\mathcal{W}, M)$ is at most $1 - \frac{1}{\mathbf{e}}$. Furthermore, by choosing $M' = \ln(t)M$, we can show that the gap is at most $1 - \frac{1}{t}$ and approaches 1 as $t \to \infty$. This proves the fact that the non-signaling correlations do not help to reduce the asymptotic entanglement-assisted simulation capacity of a quantum channel [17], [21] (see [26] for a proof) $Q^{\mathrm{EA}}(\mathcal{W}) = Q^{\mathrm{NS}}(\mathcal{W})$.

Proof of Cor. 2.1. Let $\alpha = 1 - \left(1 - \frac{1}{M^2}\right)^{M'^2}$ and $\phi = |\phi\rangle\langle\phi|_{A'A}$ be a pure state. Prop. 2 implies the existence of the state ζ such that

$$\widetilde{\mathcal{W}}^{\mathrm{EA}}_{A\to B}(\phi_{A'A}) = \alpha\widetilde{\mathcal{W}}^{\mathrm{NS}}_{A\to B}(\phi_{A'A}) + (1-\alpha)\zeta.$$

So, we have by the triangle inequality that

$$\begin{aligned}&1 - \left\|\widetilde{\mathcal{W}}^{\mathrm{EA}}_{A\to B}(\phi_{A'A}) - \mathcal{W}_{A\to B}(\phi_{A'A})\right\|_{\mathrm{tr}}\\ &= 1 - \left\|\alpha\widetilde{\mathcal{W}}^{\mathrm{NS}}_{A\to B}(\phi_{A'A}) + (1-\alpha)\zeta - \mathcal{W}_{A\to B}(\phi_{A'A})\right\|_{\mathrm{tr}}\\ &\geq 1 - \alpha\left\|\widetilde{\mathcal{W}}^{\mathrm{NS}}_{A\to B}(\phi_{A'A}) - \mathcal{W}_{A\to B}(\phi_{A'A})\right\|_{\mathrm{tr}}\\ &\quad - (1-\alpha)\left\|\zeta - \mathcal{W}_{A\to B}(\phi_{A'A})\right\|_{\mathrm{tr}}\\ &\geq 1 - \alpha\left\|\widetilde{\mathcal{W}}^{\mathrm{NS}}_{A\to B}(\phi_{A'A}) - \mathcal{W}_{A\to B}(\phi_{A'A})\right\|_{\mathrm{tr}} - (1-\alpha)\\ &= \alpha\left(1 - \left\|\widetilde{\mathcal{W}}^{\mathrm{NS}}_{A\to B}(\phi_{A'A}) - \mathcal{W}_{A\to B}(\phi_{A'A})\right\|_{\mathrm{tr}}\right).\end{aligned}$$

Choosing the optimal feasible solution $\widetilde{\mathcal{W}}^{\mathrm{NS}}$ of the program (3) and taking the infimum over $\phi_{A'A}$ yield the Corollary. □

IV. Conclusion

In this paper, we studied approximation algorithms of simulating classical and quantum channels. We proved rounding results relating the success probabilities of SR/EA and NS strategies. In particular, we showed that NS assistance is not proving any further advantage for simulation capacities. Yet, NS programs are more tractable for approximating the success probability. It would be interesting to generalize these results to the simulation of classical-quantum (CQ), quantum-classical (QC) and/or network channels, as well as investigating computational hardness results of achieving an improved approximation ratio of $(1 - \frac{1}{\mathbf{e}} + \varepsilon)$ for some $\varepsilon > 0$.

Acknowledgment

MB and AO acknowledge funding by the European Research Council (ERC Grant Agreement No. 948139). OF acknowledges support from the European Research Council (ERC Grant AlgoQIP, Agreement No. 851716).

REFERENCES

[1] Y. Steinberg and S. Verdu, "Channel simulation and coding with side information," *IEEE Transactions on Information Theory*, vol. 40, no. 3, pp. 634–646, May 1994.

[2] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, "Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem," *IEEE transactions on Information Theory*, vol. 48, no. 10, pp. 2637–2655, 2002.

[3] M. Berta, "Quantum side information: Uncertainty relations, extractors, channel simulations," *arXiv preprint arXiv:1310.4581*, 2013.

[4] M. Sudan, H. Tyagi, and S. Watanabe, "Communication for generating correlation: A unifying survey," *IEEE Transactions on Information Theory*, vol. 66, no. 1, pp. 5–37, 2019.

[5] M. X. Cao, N. Ramakrishnan, M. Berta, and M. Tomamichel, "Channel simulation: Finite blocklengths and broadcast channels," *arXiv preprint arXiv:2212.11666*, 2022.

[6] L. Yu, V. Y. Tan *et al.*, "Common information, noise stability, and their extensions," *Foundations and Trends® in Communications and Information Theory*, vol. 19, no. 2, pp. 107–389, 2022.

[7] C. E. Shannon, "A mathematical theory of communication," *The Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.

[8] M. Berta, M. Christandl, and R. Renner, "The quantum reverse Shannon theorem based on one-shot information theory," *Communications in Mathematical Physics*, vol. 306, pp. 579–615, 2011.

[9] C. H. Bennett, I. Devetak, A. W. Harrow, P. W. Shor, and A. Winter, "The quantum reverse Shannon theorem and resource tradeoffs for simulating quantum channels," *IEEE Transactions on Information Theory*, vol. 60, no. 5, pp. 2926–2959, 2014.

[10] S. Barman and O. Fawzi, "Algorithmic aspects of optimal channel coding," *IEEE Transactions on Information Theory*, vol. 64, no. 2, pp. 1038–1045, 2017.

[11] W. Matthews, "A linear program for the finite block length converse of Polyanskiy–Poor–Verdú via nonsignaling codes," *IEEE Transactions on Information Theory*, vol. 58, no. 12, pp. 7036–7044, 2012.

[12] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.

[13] M. Hayashi, "Information spectrum approach to second-order coding rate in channel coding," *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 4947–4966, 2009.

[14] O. Fawzi, J. Seif, and D. Szilágyi, "Approximation algorithms for classical-quantum channel coding," in *2019 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2019, pp. 2569–2573.

[15] O. Fawzi and P. Fermé, "Broadcast channel coding: Algorithmic aspects and non-signaling assistance," *arXiv preprint arXiv:2310.05515*, 2023.

[16] ——, "Multiple-access channel coding with non-signaling correlations," *IEEE Transactions on Information Theory*, 2023.

[17] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter, "Zero-error channel capacity and simulation assisted by non-local correlations," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 5509–5523, 2011.

[18] D. Leung and W. Matthews, "On the power of PPT-preserving and non-signalling codes," *IEEE Transactions on Information Theory*, vol. 61, no. 8, pp. 4486–4499. [Online]. Available: https://doi.org/10.1109/TIT.2015.2439953

[19] R. Duan and A. Winter, "No-Signalling-Assisted Zero-Error Capacity of Quantum Channels and an Information Theoretic Interpretation of the Lovász Number," *IEEE Transactions on Information Theory*, vol. 62, no. 2, pp. 891–914, Dec. 2015.

[20] X. Wang, W. Xie, and R. Duan, "Semidefinite Programming Strong Converse Bounds for Classical Capacity," *IEEE Transactions on Information Theory*, vol. 64, no. 1, pp. 640–653, Aug. 2017.

[21] K. Fang, X. Wang, M. Tomamichel, and M. Berta, "Quantum channel simulation and the channel's smooth max-information," *IEEE Transactions on Information Theory*, vol. 66, no. 4, pp. 2129–2140, 2019.

[22] M. X. Cao, N. Ramakrishnan, M. Berta, and M. Tomamichel, "One-shot point-to-point channel simulation," in *2022 IEEE International Symposium on Information Theory (ISIT)*, 2022, pp. 796–801.

[23] P. Harsha, R. Jain, D. McAllester, and J. Radhakrishnan, "The communication complexity of correlation," *IEEE Transactions on Information Theory*, vol. 56, no. 1, pp. 438–449, 2010.

[24] A. Anshu, V. K. Devabathini, and R. Jain, "Quantum communication using coherent rejection sampling," *Physical review letters*, vol. 119, no. 12, p. 120506, 2017.

[25] R. Takagi, K. Wang, and M. Hayashi, "Application of the Resource Theory of Channels to Communication Scenarios," *Physical Review Letters*, vol. 124, no. 12, p. 120502, Mar. 2020.

[26] M. Berta, O. Fawzi, and A. Oufkir, "Optimality of meta-converse for channel simulation," *To appear*, 2024.

[27] V. Neumann, "Various techniques used in connection with random digits," *Notes by GE Forsythe*, pp. 36–38, 1951.

[28] M. X. Cao, R. Jain, and M. Tomamichel, "One-shot simulation of classical-quantum channels," *private communication*.

[29] H.-C. Cheng and L. Gao, "Tight one-shot analysis for convex splitting with applications in quantum information theory," *arXiv preprint arXiv:2304.12055*, 2023.

[30] H.-C. Cheng, L. Gao, and M. Berta, "Quantum broadcast channel simulation via multipartite convex splitting," *arXiv preprint arXiv:2304.12056*, 2023.

[31] A. Uhlmann, "The "transition probability" in the state space of a *-algebra," *Reports on Mathematical Physics*, vol. 9, no. 2, pp. 273–279, 1976.

[32] C. H. Bennett and S. J. Wiesner, "Communication via one- and two-particle operators on einstein-podolsky-rosen states," *Physical review letters*, vol. 69, no. 20, pp. 2881–2884, 1992.