

Suboptimality of Parity for Distilling Correlations with Nontrivial Marginals

Syed Affan Aslam,^{*} Areej Ilyas,[†] and Jibran Rashid[‡]

*School of Mathematics and Computer Science,
Institute of Business Administration, Karachi, Pakistan.*

(Dated: November 26, 2025)

We prove that the PARITY protocol is optimal for a general class of non-adaptive distillation protocols of all n player nonlocal boxes (NLBs) based on XOR games. The conditional distributions generated by these NLBs are assumed to have trivial local marginals. We also show that already for $n = 2$, PARITY is no longer optimal if the local marginals are non-trivial. The OR protocol is shown to perform better and in the process also slightly extend the known correlations that collapse communication complexity. This emphasizes again the need to understand the local properties of nonlocal systems in order to obtain a better characterization of the global behavior. We conclude by showing an equivalence between adaptive distillation protocols that use identical NLBs and PARITY protocol using nonidentical NLBs.

I. INTRODUCTION

Nonlocality serves as a benchmark for genuine quantum behavior. The research agenda aimed at developing our understanding of nonlocal correlations accessible to physical theories benefits by analyzing correlations produced by nonlocal boxes (NLBs) [1, 2]. These hypothetical models allow investigation of underlying probabilistic frameworks without necessarily identifying complete corresponding physical models. Consequently, they may yield a simpler mathematical model for specific resources and related analysis for information processing tasks[3–5].

We concern ourselves here with concentration of weak multipartite nonlocal correlations through local protocols. Analogous to entanglement distillation, nonlocality distillation protocols allow us to combine multiple copies of a weak correlation source to obtain a stronger one. Developing a better understanding of their properties allows us to rule out theories that allow correlations that do not satisfy *reasonable* information theoretic principles.

Høyer and Rashid [6] prove optimality of PARITY for distillation for the bipartite case. We generalize the claim by proving that the PARITY protocol is the optimal non-adaptive nonlocality distillation protocol for all n player XOR games. A key limitation in the proof is the assumption that the shared NLBs have trivial local marginals. We show that the OR protocol performs better nonlocality distillation for specific error models with non-trivial marginals. Naik et al. [7] have identified the OR-AND protocol for a different error model with similar results. We identify parameters for which the OR protocol performs better than any known two copy distillation protocol, non-adaptive or adaptive.

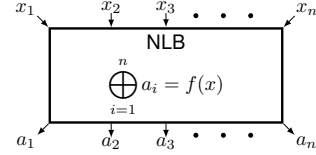


FIG. 1. A multipartite NLB N for XOR games is shared between n players. On input bit string x the box outputs bit string a such that the parity of a equals $f(x)$ with probability $p(a|x)$.

II. FRAMEWORK

We consider n player XOR games where players $P_1, P_2, \dots, P_n$ receive binary inputs $x_1, x_2, \dots, x_n$ and output bits $a_1, a_2, \dots, a_n$ respectively. So, the combined length n input x and output a bit strings are elements of $\{0, 1\}^n$. The winning condition is satisfied if parity of the output bits equals a boolean function f evaluated on input x , i.e.,

$$\bigoplus_{i=1}^n a_i = f(x).$$

Each player P_i decides a strategy based on common knowledge of the fixed function f and their input x_i . Given that there is no communication allowed between players after receiving the inputs, the no-signalling conditions limit the probability distributions that players can generate. An NLB N is a resource shared between players that generates conditional probability distributions $p(a|x)$ that satisfy the no-signalling conditions. One way to quantify the strength of the correlation produced by the NLB N is the value V given by

$$V(N) = \sum_{f(x)=\bigoplus_i a_i} p(a|x) - \sum_{f(x) \neq \bigoplus_i a_i} p(a|x). \quad (1)$$

We parametrize the distributions produced by noisy NLBs on input x by δ_x , such that

$$p(a|x) = \begin{cases} \frac{1}{2^n}(1 + \delta_x) & \text{if } \bigoplus_{i=1}^n a_i = 0 \text{ and} \\ \frac{1}{2^n}(1 - \delta_x) & \text{if } \bigoplus_{i=1}^n a_i = 1. \end{cases} \quad (2)$$

^{*} syedaffanaslam@gmail.com

[†] areej.ilyas@khi.iba.edu.pk

[‡] jrashid@iba.edu.pk

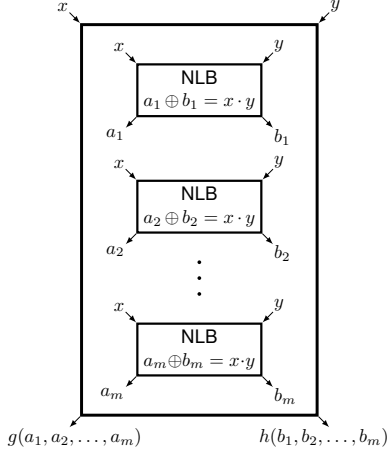


FIG. 2. A non-adaptive distillation protocol for two players using m bipartite NLBs. The PARITY protocol fixes g and h to be the parity function. All boxes receive the same input as the two players.

The value V of the NLB N is then given by

$$\begin{aligned} V &= \sum_{a, x \in \{0,1\}^n} (-1)^{[f(x) = \oplus_{i=1}^n a_i]} p(a|x) \\ &= \sum_{x \in \{0,1\}^n} (-1)^{f(x)} \delta_x, \end{aligned} \quad (3)$$

where the value on input x is δ_x if $f(x) = 0$ and $-\delta_x$ otherwise. Note that the local marginals for each player P_i are trivial, i.e.,

$$p(a_i = 0|x_i) - p(a_i = 1|x_i) = 0.$$

Given m identical NLBs the goal of NLB distillation is to obtain an NLB that achieves as high a value as possible with no communication between the players. Høyer and Rashid [6] defined a protocol as *non-adaptive* (Figure 2) if each of the m NLBs takes as input, the original input bits of the players. We consider the same class of protocols here, but note that this is not the most general form of non-adaptive protocols, since the players could locally determine different inputs to the boxes. In our case, the protocol is completely determined by the boolean functions players use to fix their output bits. The PARITY protocol requires all players to output the parity of the bits received from the m NLBs. Our claim about NLBs of the form in Equation 2 is stated in Theorem 1.

Theorem 1 *The PARITY protocol is optimal for the class of non-adaptive distillation protocols that provide the same input to all NLBs with trivial marginals.*

We also prove Theorem 2 in Section IV that shows that the optimal protocol can change if the local marginals are non-trivial.

Theorem 2 *The OR protocol can perform better non-locality distillation than the PARITY protocol for NLBs with non-trivial marginals.*

The proof of optimality for PARITY has free parameters for the expected value for each input while requiring the local marginals to be trivial. The proof for Theorem 2 is obtained by characterizing the eight dimensional NLBs for the two player CHSH inequality [8] in terms of the expected values for each input and the four local marginals. It is open whether the proof of Theorem 1 can be generalized to allow for classification of optimal protocols in terms of local marginals.

III. PARITY IS OPTIMAL FOR XOR GAMES

We prove Theorem 1 by first calculating the value achieved by the parity protocol over m NLBs in Lemma 1.

Lemma 1 *The PARITY protocol over m NLBs of the form in Equation 2, attains the value*

$$V = \sum_{x \in \{0,1\}^n} (-1)^{f(x)} \delta_x^m.$$

Lemma 1 follows from the observation that on input bit string x the PARITY protocol using m NLBs, attains the value δ_j^m when $f(\hat{x}) = 0$ and $-\delta_j^m$ otherwise. This is a generalization of Theorem 2 from [9].

We complete the proof of Theorem 1 by proving that the value of all non-adaptive distillation protocols for trivial marginal NLBs is bounded by the value of the PARITY protocol.

Lemma 2 *The value of a non-adaptive distillation protocol using at most m NLBs with the same input, parametrized by Equation 2, is upper bounded by the value*

$$\max_{1 \leq k \leq m} \left| \sum_{x \in \{0,1\}^n} (-1)^{f(x)} \delta_x^k \right|.$$

Proof of Lemma 2 Let $\hat{a}_i = (a_i^1, a_i^2, \dots, a_i^m)$ be the length m tuple of the output bit string obtained by player i on input bit x_i . The combined outputs from the m NLBs received by the n players is given by $\hat{a} = (\hat{a}_1, \hat{a}_2, \dots, \hat{a}_n)$. Consider the situation where for the input $x = (x_1, x_2, \dots, x_n)$, all the n players output $\hat{0} = (0, 0, \dots, 0)$ at the end of the protocol. Note that for a fixed input x , we can drop the subscript from δ . Let each $\hat{A}_i$ in $\hat{A} = (\hat{A}_1, \hat{A}_2, \dots, \hat{A}_n)$ represent the set of strings in $\{0,1\}^m$ such that player i 's final output at the end of protocol is 0. Given that the n players input x into the m NLBs, the probability that they obtain output $\hat{a}$

is given by

$$\begin{aligned}
p(\hat{a}|x) &= \prod_{i=1}^m \left(\frac{1-\delta}{2^n} + \frac{\delta}{2^{n-1}} [\oplus_{j=1}^n a_j^i = 0] \right) \\
&= \left(\frac{1-\delta}{2^n} \right)^m \prod_{i=1}^m \left(1 + \frac{2\delta}{1-\delta} [\oplus_{j=1}^n a_j^i = 0] \right) \\
&= \left(\frac{1-\delta}{2^n} \right)^m \left(\frac{1+\delta}{1-\delta} \right)^{m-|\oplus_{i=1}^n \hat{a}_i|} \\
&= \frac{1}{2^{mn}} (1-\delta)^{|\oplus_{i=1}^n \hat{a}_i|} (1+\delta)^{m-|\oplus_{i=1}^n \hat{a}_i|}.
\end{aligned}$$

The probability $Q(\hat{A})$ to obtain output $\hat{0}$ can be obtained by summing over all output bit strings $\hat{a}_j \in \hat{A}_j$, i.e., $Q(\hat{A}(\delta))$ equals

$$\sum_{j=1}^n \sum_{\hat{a}_j \in \hat{A}_j} \frac{1}{2^{mn}} (1-\delta)^{|\oplus_{j=1}^n \hat{a}_j|} (1+\delta)^{m-|\oplus_{j=1}^n \hat{a}_j|}.$$

We expand the inner-most product in its 2^m terms and express each entry as an evaluation of $\oplus_{j=1}^n \hat{a}_j$ on one of the 2^m characters χ_z . Here $\chi_z(\oplus_{j=1}^n \hat{a}_j) = (-1)^{z \cdot (\oplus_{j=1}^n \hat{a}_j)}$ is a character for the group $\mathbb{Z}_2^m$. This gives

$$\begin{aligned}
Q(\hat{A}(\delta)) &= \frac{1}{2^{nm}} \sum_{j=1}^n \sum_{\hat{a}_j \in \hat{A}_j} \sum_{z \in \{0,1\}^m} \chi_z(\oplus_{j=1}^n \hat{a}_j) \delta^{|z|} \\
&= \sum_{z \in \{0,1\}^m} \delta^{|z|} \prod_{j=1}^n \left(\sum_{\hat{a}_j \in \hat{A}_j} \frac{1}{2^m} \chi_z(\hat{a}_j) \right) \\
&= \sum_{z \in \{0,1\}^m} \delta^{|z|} \prod_{j=1}^n \sum_{s_j} \frac{1}{2^m} \chi_z(s_j) [s_j \in \hat{A}_j].
\end{aligned}$$

We now introduce a set of functions $f_1, f_2, \dots, f_n$ that take the value +1 if $s_j \in \hat{A}_j$ and -1 otherwise, to express $Q(\hat{A}(\delta))$ as

$$\begin{aligned}
&\sum_{z \in \{0,1\}^m} \delta^{|z|} \left(\prod_{j=1}^n \sum_{s_j} \frac{1}{2^m} \chi_z(s_j) \left(\frac{f_j(s_j) + 1}{2} \right) \right) \\
&= \sum_{z \in \{0,1\}^m} \frac{\delta^{|z|}}{2^n} \prod_{j=1}^n \left(\hat{f}_z^j + [z = 0] \right),
\end{aligned}$$

where $\hat{f}_z^j = \sum_{s_j} \frac{1}{2^m} \chi_z(s_j) f_j(s_j)$ is the Fourier coefficient. Using a similar process we can compute $R(\hat{A}(\delta))$, the probability for the n players to output even parity. Let k be a bit string in $\{0,1\}^n$ with even parity. We define the i^{th} element of $\hat{A}^k$ to be the set $\hat{A}_i$ if $k_i = 0$ and its complement $\bar{\hat{A}}_i$ if $k_i = 1$, i.e., the set containing bit strings on which player i 's final output is 1. For example, if k is equal to the all ones length n bit string $k = \hat{1} = (1, 1, \dots, 1)$, then $\hat{A}^k$ corresponds to the set of strings on which the final output of all n players is 1.

Here we equate $Q(\hat{A}^{k=\hat{0}})$ to $Q(\hat{A})$. For a fixed k , the corresponding output probability $Q(\hat{A}^k(\delta))$ is given by,

$$\begin{aligned}
&\sum_{z \in \{0,1\}^m} \delta^{|z|} \left(\prod_{j=1}^n \sum_{s_j} \frac{1}{2^m} \chi_z(s_j) \left(\frac{(-1)^{k_j} f_j(s_j) + 1}{2} \right) \right) \\
&= \sum_{z \in \{0,1\}^m} \frac{\delta^{|z|}}{2^n} \prod_{j=1}^n \left((-1)^{k_j} \hat{f}_z^j + [z = 0] \right).
\end{aligned}$$

The probability $R(\hat{A}(\delta))$ to output even parity bit string is obtained by summing over all Q 's,

$$R(\hat{A}(\delta)) = \sum_k Q(\hat{A}^k(\delta)).$$

Expanding the summation, the middle terms cancel out leaving the expression

$$R(\hat{A}(\delta)) = \frac{1}{2} \left(1 + \sum_{z \in \{0,1\}^m} \delta^{|z|} \prod_j \hat{f}_z^j \right).$$

The expected value for a fixed δ can then be expressed as

$$R(\hat{A}(\delta)) - (1 - R(\hat{A}(\delta))) = 2R(\hat{A}(\delta)) - 1.$$

Finally, we can determine the bound on the value V by summing over the biases for each input string x to obtain

$$\begin{aligned}
V &= \sum_{x \in \{0,1\}^n} (-1)^{f(x)} (2R(\delta_x) - 1) \\
&= \sum_{z \in \{0,1\}^m} \prod_{j=1}^n \hat{f}_z^j \left(\sum_x (-1)^{f(x)} \delta_x^{|z|} \right) \\
&\leq \sum_{z \in \{0,1\}^m} \prod_{j=1}^n |\hat{f}_z^j| \left(\left| \sum_x (-1)^{f(x)} \delta_x^{|z|} \right| \right) \\
&\leq \max_{1 \leq k \leq m} \left| \sum_x (-1)^{f(x)} \delta_x^k \right| \sum_{z \in \{0,1\}^m} \prod_{j=1}^n |\hat{f}_z^j| \\
&\leq \max_{1 \leq k \leq m} \left| \sum_x (-1)^{f(x)} \delta_x^k \right|,
\end{aligned}$$

where the last inequality follows from the $\hat{f}_z^j$ being normalized functions. $\square$

Combining Lemmas 1 and 2 proves Theorem 1. Next we restrict our attention to two player NLBs for the CHSH inequality, but generalize the error model to allow for non-trivial marginals.

IV. DISTILLING NLBS WITH NON-TRIVIAL MARGINALS

The statement of the *local quantum measurement* principle [10] is a proven fact rather than an assumption. It

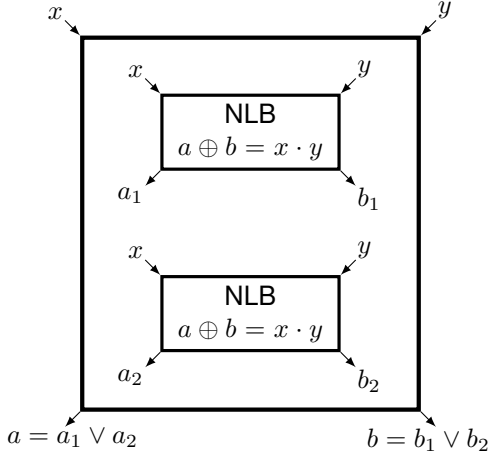


FIG. 3. The non-adaptive OR protocol for two NLBs.

states that in a two player setting if the local systems of Alice and Bob can be described by a quantum system, then their nonlocal correlations also admit a quantum description. The assumption however, lies within the statement. If we assume that the local distributions of Alice and Bob can always be produced by measurements on a quantum system, then according to the conclusion of Barnum et al. [10], their joint distribution also admits a

quantum description. The mystery lies in what exactly it entails for the local distributions to be quantum.

Here we study the influence of local distributions on nonlocality distillation. We generalize the NLB definition to include non-trivial local marginals and then consider them as a resource for nonlocality distillation. For binary input and output, any bipartite no-signalling distribution $p_{ab|xy}$ can be decomposed as

$$p_{ab|xy} = \frac{1}{4} \left(1 + (-1)^a E_x + (-1)^b E_y + (-1)^{a \oplus b} E_{xy} \right), \quad (4)$$

where

$$E_{xy} = p_{00|xy} + p_{11|xy} - p_{01|xy} - p_{10|xy},$$

$$E_x = p_{a=0|x} - p_{a=1|x},$$

$$E_y = p_{b=0|y} - p_{b=1|y}.$$

Let the expected value for each input xy be given by

$$\delta_1 = E_{xy=00}, \delta_2 = E_{xy=01}, \delta_3 = E_{xy=10} \text{ and } \epsilon = E_{xy=11}.$$

Similarly, let the four local marginals be represented by

$$\alpha = E_{x=0}, \beta = E_{x=1}, \gamma = E_{y=0} \text{ and } \omega = E_{y=1}.$$

A local marginal is said to be trivial if it is equal to zero and non-trivial otherwise. The general description of an NLB N with non-trivial marginals is given by

$$N = \frac{1}{4} \begin{pmatrix} 1+\alpha+\gamma+\delta_1 & 1+\alpha-\gamma-\delta_1 & 1+\gamma-\alpha-\delta_1 & 1-\alpha-\gamma+\delta_1 \\ 1+\alpha+\omega+\delta_2 & 1+\alpha-\omega-\delta_2 & 1+\omega-\alpha-\delta_2 & 1-\alpha-\omega+\delta_2 \\ 1+\beta+\gamma+\delta_3 & 1+\beta-\gamma-\delta_3 & 1+\gamma-\beta-\delta_3 & 1-\beta-\gamma+\delta_3 \\ 1+\beta+\omega+\epsilon & 1+\beta-\omega-\epsilon & 1+\omega-\beta-\epsilon & 1-\beta-\omega+\epsilon \end{pmatrix}. \quad (5)$$

Note that the value $V = \delta_1 + \delta_2 + \delta_3 - \epsilon$ attained by the NLB N in Equation 5 is the same as an NLB with trivial marginals. This is a consequence of the fact that the CHSH inequality does not depend on the local marginals. At the same time, we may not choose the values of the local marginals arbitrarily. The value attained by an NLB restricts the values of the marginals in order to ensure no-signalling constraints are not violated. For example, $V = 4$ limits the only valid local marginals for a no-signalling NLB to zero.

Given that non-trivial marginals seem to require more resources to simulate, we consider the possibility that NLBs with non-trivial marginals may have a higher distillation value than those with trivial marginals. Even though the local marginals do not contribute to the value V of a single NLB N , we show in Table I that via non-locality distillation protocols an NLB with non-trivial marginals can get distilled to a strictly higher value than the identical NLB with trivial marginals.

This example illustrates that the nonlocal properties of

the correlations of an NLB also depend on properties local to Alice and Bob. Distillation provides an operational framework to distinguish NLBs based on the strength of their correlations. Figure 3 gives the OR protocol, which we now examine in terms of its distillation power for non-trivial marginals.

To simplify the analysis we begin by restricting $\delta_1 = \delta_2 = \delta_3 = 1$ and equal but not necessarily trivial local marginals, i.e., $\alpha = \beta = \gamma = \omega$. This results in the *correlated* NLB given by

$$N = \frac{1}{4} \begin{pmatrix} 2+2\alpha & 0 & 0 & 2-2\alpha \\ 2+2\alpha & 0 & 0 & 2-2\alpha \\ 2+2\alpha & 0 & 0 & 2-2\alpha \\ 1+2\alpha+\epsilon & 1-\epsilon & 1-\epsilon & 1-2\alpha+\epsilon \end{pmatrix}.$$

Proof of Theorem 2 Value V_{OR} of the OR protocol for

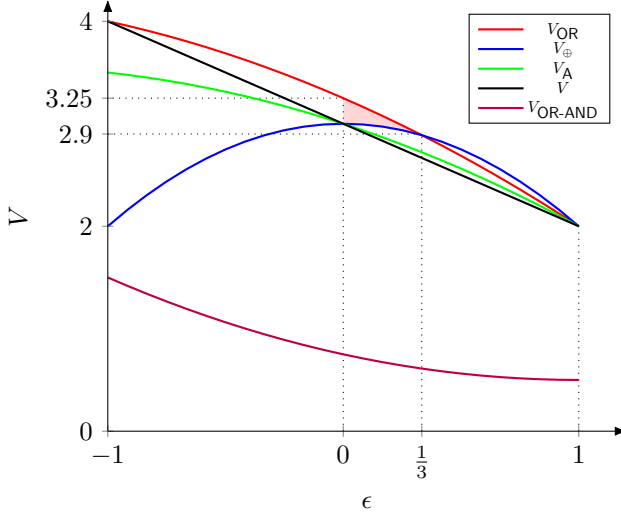


FIG. 4. The values for various two copy distillation protocols for $\alpha = \frac{1}{2}$. V_{OR} is the highest distilled value for $\epsilon \in [0, \frac{1}{3}]$. Note that even though the curve for V_{OR} remains at the top for $\epsilon < 0$, these values do not correspond to a valid NLB.

two copies of *correlated* NLBs is given by,

$$V_{\text{OR}} = (\alpha + \frac{1}{2})(3 - \epsilon) - \frac{\epsilon^2}{4} + \frac{5}{4} - 2\alpha$$

$$= \left(\left(\alpha + \frac{1}{2} \right) V \right) + \left(\frac{1}{4} V_{\oplus} \right) + \frac{1}{2} - 2\alpha.$$

It suffices to show that $V_{\text{OR}} > \max(V_{\oplus}, V)$, where $V = 3 - \epsilon$ is the value of a single *correlated* NLB and $V_{\oplus} = 3 - \epsilon^2$ is the value of the PARITY protocol on two copies of the NLB. Solving the inequality we obtain the requirement that for

$$\max(1 - 4\alpha, 2\alpha - 1) \leq \epsilon < \frac{4\alpha - 1}{3} \text{ and } \alpha \in \left[\frac{1}{4}, 1 \right],$$

OR protocol performs better nonlocality distillation than the PARITY protocol. $\square$

δ_1	δ_2	δ_3	ϵ	a	b	c	d	V	V_A
1.00	1.00	1.00	-0.70	0.01	0.01	0.01	0.01	3.70	3.8360
1.00	1.00	1.00	-0.70	0.0	0.0	0.0	0.0	3.70	3.8275
0.92	0.92	0.92	-0.22	0.01	0.01	0.01	0.01	2.98	2.9924
0.92	0.92	0.92	-0.22	0.0	0.0	0.0	0.0	2.98	2.9867
0.917	0.917	0.917	-0.22	0.01	0.01	0.01	0.01	2.971	2.97539
0.917	0.917	0.917	-0.22	0.0	0.0	0.0	0.0	2.971	2.96971

TABLE I. Column V gives the value attained by the NLB and V_A is the protocol of Allcock et al. [11] for 2 copies. Note that the distillation value is always higher for the NLB with non-trivial marginals.

Figure 4 shows that the OR protocol performs optimal nonlocality distillation for $\alpha = \frac{1}{2}$ and $\epsilon \in [0, \frac{1}{3}]$. It also

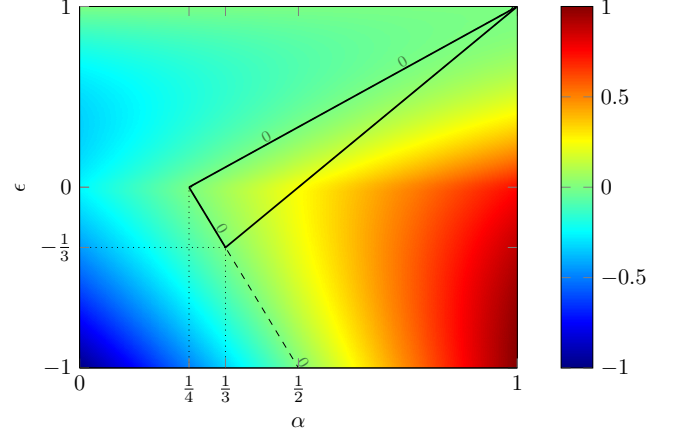


FIG. 5. The bounded triangle corresponds to the region where OR is optimal among the known protocols for distilling two copies of *correlated* NLBs.

outperforms that *adaptive* Allcock et al. [11] protocol. Figure 5 depicts the entire region in terms of ϵ and α where the OR protocol is the optimal known distillation protocol for two copies of the NLB. Table II summarizes the behavior of the OR protocol in the optimal distillation region. If we increase the parameter space and allow

Range for ϵ	α	δ	ϵ	V	$V_{\oplus}$	V_{OR}	V_A
$[-0.04, 0.013]$	0.26	1.00	0.01	2.99	2.9999	3.00	3.1112
$[-0.20, 0.066]$	0.30	1.00	0.01	2.99	2.9999	3.04	3.0914
$[-0.30, 0.133]$	0.35	1.00	0.01	2.99	2.9999	3.09	3.0667
$[-0.20, 0.200]$	0.40	1.00	0.01	2.99	2.9999	3.14	3.0419
$[-0.10, 0.266]$	0.45	1.00	0.01	2.99	2.9999	3.19	3.0172
$[0.00, 0.333]$	0.50	1.00	0.01	2.99	2.9999	3.24	2.9924

TABLE II. Comparison of the distilled values for *correlated* NLBs attained by various protocols.

for $\delta_1 = \delta_2 = \delta_3$, and two local marginals given by $\alpha = \gamma$ and $\beta = \omega$, the values attained for the resulting *symmetric* NLBs are listed in Appendix A. Table III provides a snapshot of how the two non-adaptive protocols behave within this region.

Finally, the known region for trivial communication complexity [12] is marginally increased by considering distillation of NLBs with non-trivial marginals. Table IV identifies parameter values for *symmetric* NLBs with a single non-trivial marginal. The OR protocol distills these NLBs beyond the required threshold of $4\sqrt{2/3}$.

V. PROTOCOLS WITH NONIDENTICAL BOXES

Consider whether given an adaptive protocol on n identical NLBs, we can construct a non-adaptive protocol on n nonidentical NLBs that simulates the adaptive protocol. Note that we are not asking for a trivial construc-

α	β	δ	ϵ	V	$V_{\oplus}$	V_{OR}
0.43	0.42	0.99	0.28	2.69	2.8619	2.850
0.43	0.43	0.99	0.28	2.69	2.8619	2.857
0.43	0.44	0.99	0.28	2.69	2.8619	2.864
0.44	0.43	0.99	0.28	2.69	2.8619	2.857
0.44	0.44	0.99	0.28	2.69	2.8619	2.864
0.44	0.45	0.99	0.28	2.69	2.8619	2.871
0.45	0.44	0.99	0.28	2.69	2.8619	2.864
0.45	0.45	0.99	0.28	2.69	2.8619	2.871
0.45	0.46	0.99	0.28	2.69	2.8619	2.878

TABLE III. Comparison of distilled values via non-adaptive protocols for *symmetric* NLBs.

α	δ	ϵ	V	$V_{\oplus}$	V_A	V_{OR}
0.42	0.99	-0.16	3.13	2.9744	3.101575	3.2683
0.41	0.99	-0.18	3.15	2.9676	3.121425	3.2735
0.40	0.99	-0.20	3.17	2.9600	3.141275	3.2781
0.39	0.99	-0.22	3.19	2.9516	3.161125	3.2821
0.38	0.99	-0.24	3.21	2.9424	3.180975	3.2855
0.37	0.99	-0.26	3.23	2.9324	3.200825	3.2883
0.36	0.99	-0.28	3.25	2.9216	3.220675	3.2905

TABLE IV. Choice of parameters that are now included in the region of trivial communication complexity via nonlocality distillation.

tion where we simply model the final output of the adaptive protocol as another box. We want to take the adaptiveness of a protocol using identical boxes and push it into the individual boxes. At the expense of having non-identical boxes we then obtain a non-adaptive protocol that produces the same output as the adaptive protocol. For example, the adaptive protocol of Brunner and Skrzypczyk [13] (Figure 6) maps two isotropic NLBs

$$N = \frac{1}{4} \begin{pmatrix} 1+\delta & 1-\delta & 1-\delta & 1+\delta \\ 1+\delta & 1-\delta & 1-\delta & 1+\delta \\ 1+\delta & 1-\delta & 1-\delta & 1+\delta \\ 1-\delta & 1+\delta & 1+\delta & 1-\delta \end{pmatrix},$$

to the NLB given by

$$\frac{1}{4} \begin{pmatrix} 1+\delta^2 & 1-\delta^2 & 1-\delta^2 & 1+\delta^2 \\ 1+\delta^2 & 1-\delta^2 & 1-\delta^2 & 1+\delta^2 \\ 1+\delta^2 & 1-\delta^2 & 1-\delta^2 & 1+\delta^2 \\ \frac{2-\delta^2-\delta}{2} & \frac{2+\delta^2+\delta}{2} & \frac{2+\delta^2+\delta}{2} & \frac{2-\delta^2-\delta}{2} \end{pmatrix}.$$

To construct the equivalent non-adaptive protocol, for input $xy \in \{00, 01, 10\}$ we can set the corresponding probability distributions for N_1 and N_2 equal to $(1+\delta, 1-\delta, 1-\delta, 1+\delta)$. For input $xy = 11$, assume that N_1 and N_2 have expected values δ_1 and δ_2 , respectively. Then, the probability that the non-adaptive protocol outputs 00 on input 11 is given by

$$p_{ab=00|xy=11} = \frac{1 - \delta_1 \delta_2}{4}.$$

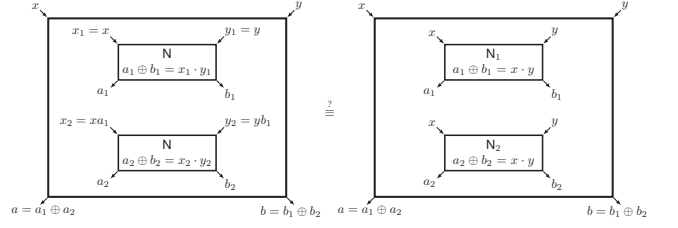


FIG. 6. Does there exist an equivalence between the adaptive protocol on the left using two identical NLBs and the non-adaptive protocol on the right using two nonidentical NLBs N_1 and N_2 .

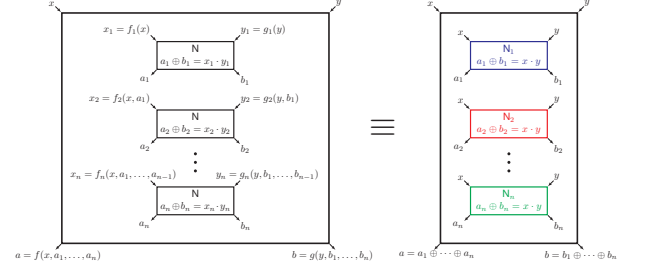


FIG. 7. General equivalence between adaptive protocols using identical NLBs and non-adaptive PARITY protocol using nonidentical NLBs.

Equating with the output probability of the adaptive protocol, we obtain

$$\frac{1 - \delta_1 \delta_2}{4} = \frac{2 - \delta^2 - \delta}{8},$$

$$\delta_1 \delta_2 = \frac{\delta(1 + \delta)}{2}.$$

Setting $\delta_1 = \delta$ and $\delta_2 = \frac{1+\delta}{2}$ gives the same probability distribution as the adaptive protocol and completes the equivalence construction. In summary, Alice and Bob can non-adaptively use two NLBs, $N_1 = N$ and N_2 given by

$$N_2 = \frac{1}{4} \begin{pmatrix} 1+\delta & 1-\delta & 1-\delta & 1+\delta \\ 1+\delta & 1-\delta & 1-\delta & 1+\delta \\ 1+\delta & 1-\delta & 1-\delta & 1+\delta \\ 1 - \frac{1+\delta}{2} & 1 + \frac{1+\delta}{2} & 1 + \frac{1+\delta}{2} & 1 - \frac{1+\delta}{2} \end{pmatrix},$$

to obtain the same output distribution as the adaptive protocol. Note that both the boxes N_1 and N_2 have a values less than or equal to the original NLB for all $\delta \geq 0$. We can generalize this construction to establish Theorem 3 (Figure 7).

Theorem 3 *The output distribution of an adaptive protocol using n identical isotropic NLBs can be simulated by a non-adaptive PARITY protocol over n nonidentical NLBs N_i .*

Proof of Theorem 3 The probability $p_{00|xy}$ that the adaptive protocol outputs $ab = 00$ on input xy is given

by a polynomial $q_{xy}(\delta)$ of degree at most n , having form

$$q_{xy}(\delta) = \frac{1}{4^n} \sum_{\substack{f(x,a_1,\dots,a_n)=a \\ g(y,b_1,\dots,b_n)=b}} \prod_{j=1}^n (1 + (-1)^{[a_j \oplus b_j \neq x_j \cdot y_j]} \delta).$$

The same probability for a non-adaptive protocol using n nonidentical NLBs N_i is given by

$$\frac{1}{4} \left(1 + \prod_{i=1}^n \delta_{i,xy}(\delta) \right),$$

where $\delta_{i,xy}(\delta)$ is the expected value of the i^{th} NLB on input xy . Equating the two, we obtain

$$\prod_{i=1}^n \delta_{i,xy}(\delta) = 4q_{xy}(\delta) - 1.$$

Setting each $\delta_{i,xy}(\delta) = c_{i,1}\delta + c_{i,0}$ with the restriction that $\delta_{i,xy}(\delta) \in [-1, 1]$ establishes the equivalence. $\square$

Unfortunately, the proof is not constructive, i.e. we do not obtain the expressions for each $\delta_{i,xy}(\delta)$. For example,

if we can bound the value of each N_i by the value of N we can prove non-distillability of isotropic NLBs in general.

VI. DISCUSSION

It remains to obtain a complete characterization of optimal non-adaptive distillation protocols in the presence of non-trivial marginals. One possibility is to use the recent algebraic characterization of distillation protocols by Botteron et al. [14, 15] for this task. Another generalization to consider is the combination of distillation protocols using non-identical NLBs, similar to the approach of Eftaxias, Weilenmann and Colbeck [16].

Farkas et al. [17] recently showed that nonlocality is not a sufficient condition for the security of device independent QKD protocols. At the same time, Woollorton, Brown and Colbeck [18] and Farkas [19] have shown that even arbitrarily small nonlocality can be used for DIQKD. This implies that the structure of the nonlocal correlation rather than just the nonlocal value, plays a crucial role in DIQKD protocols. A better understanding of nonlocality distillation in the presence of non-trivial marginals may be useful in identifying the nonlocality requirements for DIQKD.

-
- [1] L. Khalfin and B. Csirel'son, Quantum and quasi-classical analogs of Bell inequalities, Symposium on the Foundations of Modern Physics, 441 (1985).
 - [2] S. Popescu and D. Rohrlich, Quantum nonlocality as an axiom, Foundations of Physics **24**, 3 (1994).
 - [3] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and S. Roberts, Nonlocal correlations as an information theoretic resource, Physical Review A **71**, 022101 (2005), quant-ph/0404097.
 - [4] H. Buhrman, R. Cleve, S. Massar, and R. Wolf, Nonlocality and communication complexity, Reviews of Modern Physics **82**, 665 (2010), 0809.3173v4.
 - [5] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Reviews of Modern Physics **86**, 419 (2014), 1303.2849v3.
 - [6] P. Høyer and J. Rashid, Optimal protocols for nonlocality distillation, Physical Review A **82**, 042118 (2010), 1009.1668v1.
 - [7] S. Naik, G. Sidhardh, S. Sen, A. Roy, A. Rai, and M. Banik, Distilling nonlocality in quantum correlations, Physical Review Letters **130**, 220201 (2023).
 - [8] J. Clauser, M. Horne, A. Shimony, and R. Holt, Proposed experiment to test local hidden-variable theories, Physical Review Letters **23**, 880 (1969).
 - [9] M. Forster, Bounds for nonlocality distillation protocols, Physical Review A **83**, 062114 (2011), 0808.0651v4.
 - [10] H. Barnum, S. Beigi, S. Boixo, M. Elliott, and S. Wehner, Local quantum measurement and no-signaling imply quantum correlations, Physical Review Letters **104**, 140401 (2010), 0910.3952v2.
 - [11] J. Allcock, N. Brunner, N. Linden, S. Popescu, P. Skrzypczyk, and T. Vertési, Closed sets of nonlocal correlations, Physical Review A **80**, 062107 (2009), 0908.1496.
 - [12] G. Brassard, H. Buhrman, N. Linden, A. Méthot, A. Tapp, and F. Unger, A limit on nonlocality in any world in which communication complexity is trivial, Physical Review Letters **96**, 250401 (2006), quant-ph/0508042v1.
 - [13] N. Brunner and P. Skrzypczyk, Nonlocality distillation and post-quantum theories with trivial communication complexity, Physical Review Letters **102**, 160403 (2009), 0901.4070.
 - [14] P. Botteron, A. Broadbent, R. Chhaibi, I. Nechita, and C. Pellegrini, Algebra of nonlocal boxes and the collapse of communication complexity, Quantum **8**, 1402 (2024).
 - [15] P. Botteron, A. Broadbent, and M. Proulx, Extending the known region of nonlocal boxes that collapse communication complexity, Physical Review Letters **132**, 070201 (2024).
 - [16] G. Eftaxias, M. Weilenmann, and R. Colbeck, Advantages of multicopy nonlocality distillation and its application to minimizing communication complexity, Physical Review Letters **130**, 100201 (2023).
 - [17] M. Farkas, M. Balanzó-Juandó, K. Łukanowski, J. Kołodyński, and A. Acín, Bell nonlocality is not sufficient for the security of standard device-independent quantum key distribution protocols, Physical Review Letters **127**, 050503 (2021).
 - [18] L. Woollorton, P. Brown, and R. Colbeck, Device-independent quantum key distribution with arbitrarily small nonlocality, Physical Review Letters **132**, 210802 (2024).
 - [19] M. Farkas, Unbounded device-independent quantum key

rates from arbitrarily small nonlocality, Physical Review Letters **132**, 210803 (2024).

Appendix A: Symmetric NLBs with Two Non-Trivial Marginals

We fix $\delta_1 = \delta_2 = \delta_3$, $\alpha = \gamma$ and $\beta = \omega$ in Equation 5 to obtain *symmetric* NLBs given by

$$N = \frac{1}{4} \begin{pmatrix} 1+2\alpha+\delta & 1-\delta & 1-\delta & 1-2\alpha+\delta \\ 1+\alpha+\beta+\delta & 1+\alpha-\beta-\delta & 1+\beta-\alpha-\delta & 1-\alpha-\beta+\delta \\ 1+\alpha+\beta+\delta & 1+\beta-\alpha-\delta & 1+\alpha-\beta-\delta & 1-\alpha-\beta+\delta \\ 1+2\beta+\epsilon & 1-\epsilon & 1-\epsilon & 1-2\beta+\epsilon \end{pmatrix}. \quad (\text{A1})$$

The values attained by distillation protocols are

$$\begin{aligned} V_{\text{OR}} &= \frac{1}{4}(3\delta^2 - \epsilon^2) + \frac{1}{2}(3\delta - \epsilon) + (\beta + 2\delta - 2)\alpha + (\delta - \epsilon)\beta - \frac{1}{2}(\alpha^2 + \beta^2 - 1), \\ V_{\text{A}} &= \frac{1}{4}(11\delta^2 + 2\delta - 2\epsilon\delta - 2\epsilon - \epsilon^2 + (b - a + 2d)(\delta - \epsilon)) \quad \text{and} \\ V_{\text{OR-AND}} &= 2\alpha^2 + \frac{1}{4}\epsilon^2 - \frac{1}{2}\epsilon - \frac{3}{4}\delta^2 + \frac{3}{2}\delta - \frac{1}{2}. \end{aligned}$$