Robust Bell inequalities from communication complexity

Sophie Laplante[1], Mathieu Laurière[2], Alexandre Nolin[1], Jérémie Roland[3], and Gabriel Senno[4]

[1]IRIF, Université Paris-Diderot, Paris, France

[2]ORFE, Princeton University, Princeton, NJ 08544, USA

[3]Université Libre de Bruxelles, Brussels, Belgium

[4]ICFO-Institut de Ciencies Fotoniques, The Barcelona Institute of Science and Technology, 08860 Castelldefels (Barcelona), Spain

June 5, 2018

Sophie Laplante: laplante@irif.fr

Mathieu Laurière: lauriere@princeton.edu

Alexandre Nolin: nolin@irif.fr

Jérémie Roland: jroland@ulb.ac.be

Gabriel Senno: gabriel.senno@icfo.eu,

Abstract

The question of how large Bell inequality violations can be, for quantum distributions, has been the object of much work in the past several years. We say that a Bell inequality is normalized if its absolute value does not exceed 1 for any classical (*i.e.* local) distribution. Upper and (almost) tight lower bounds have been given for the quantum violation of these Bell inequalities in terms of number of outputs of the distribution, number of inputs, and the dimension of the shared quantum states. In this work, we revisit normalized Bell inequalities together with another family: inefficiency-resistant Bell inequalities. To be inefficiency-resistant, the Bell value must not exceed 1 for any local distribution, including those that can abort. This makes the Bell inequality resistant to the detection loophole, while a normalized Bell inequality is resistant to general local noise. Both these families of Bell inequalities are closely related to communication complexity lower bounds. We show how to derive large violations from any gap between classical and quantum communication complexity, provided the lower bound on classical communication is proven using these lower bound techniques. This leads to inefficiency-resistant violations that can be exponential in the size of the inputs. Finally, we study resistance to noise and inefficiency for these Bell inequalities.

1 Introduction

The question of achieving large Bell violations has been studied since Bell's seminal paper in 1964 [6]. In one line of investigation, proposals have been made to exhibit families of distributions which admit unbounded violations [33, 39, 40, 43]. In another, various measures of nonlocality have been studied, such as the amount of communication necessary and sufficient to simulate quantum distributions classically [7–9, 14, 37, 38, 44, 49, 50], or the resistance to detection inefficiencies and noise. More recently, focus has turned to giving upper and lower bounds on violations achievable, in terms of various parameters: number of players, number of inputs, number of outputs, dimension of the quantum state, and amount of entanglement [14, 22, 23].

Up until quite recently, violations were studied in the case of specific distributions (measuring Bell states), or families of distributions. Buhrman *et al.* [12] gave a construction that could be applied to several problems which had efficient quantum protocols (in terms of communication) and for which one could show a trade-off between communication and error in the classical setting. This still required an *ad hoc* analysis of communication problems. Recently Buhrman *et al.* [13] proposed the first general construction of quantum states along with Bell inequalities from any communication problem. The quantum states violate the Bell inequalities when there is a sufficiently large gap between quantum and classical communication complexity (a super-quadratic gap is necessary, unless a quantum protocol without local memory exists).

Table 1 summarizes the best known upper and lower bounds on quantum violations achievable with normalized Bell inequalities.

Parameter	Upper bound	Ad hoc lower bounds	Best possible lower bound from [13]
Number of inputs N	$2^c \leq N$ [14, 23, 34]	$\frac{\sqrt{N}}{\log(N)}$ [22]	$\frac{\sqrt{c}}{q} \leq \log(N)$
Number of outputs K	$O(K)$ [22]	$\Omega\left(\frac{K}{(\log(K))^2}\right)$ [12]	$\leq \log(K)$
Dimension d	$O(d)$ [23]	$\Omega\left(\frac{d}{(\log(d))^2}\right)$ [12]	$\leq \log\log(d)$

Table 1: Bounds on quantum violations of bipartite normalized Bell inequalities, in terms of the dimension d of the local Hilbert space, the number of settings (or inputs) N and the number of outcomes (or outputs) K per party. In the fourth column, we compare ad hoc results to the recent constructions of [13] (Theorem 2) which gives a lower bound of $\frac{\sqrt{c}}{q}$, where c (resp. q) stands for the classical (resp. quantum) communication complexity of simulating a distribution. We give upper bounds on their construction in terms of the parameters d, N, K.

1.1 Our results

We revisit the question of achieving large Bell violations by exploiting known connections with communication complexity. Strong lower bounds in communication complexity, equivalent to the partition bound, amount to finding *inefficiency-resistant Bell inequalities* [31]. These are Bell functionals that are bounded above by 1 on all local distributions *that can abort*, i.e. local distributions with an additional abort outcome ⊥ for each party. In an experimental Bell test, such a constraint would provide resistance to the detection

loophole, by associating an event where one of the detectors does not click to an abort outcome.

First, we study the resistance of normalized Bell inequalities to inefficiency. We show that, up to a constant factor in the value of the violation, any normalized Bell inequality can be made resistant to inefficiency while maintaining the normalization property (**Theorem 1**).

Second, we show how to derive large Bell violations from any communication problem for which the partition bound is bounded below and the quantum communication complexity is bounded above. The problems studied in communication complexity are far beyond the quantum set, but we show how to easily derive a quantum distribution from a quantum protocol. The Bell value we obtain is 2^{c-2q}, where c is the partition lower bound on the classical communication complexity of the problem considered, and q is an upper bound on its quantum communication complexity (**Theorem 3 and Corollary 1**). The quantum distribution has one extra output per player compared to the original distribution and uses the same amount of entanglement as the quantum protocol plus as many EPR pairs as needed to teleport the quantum communication in the protocol. Next, we show that the magnitude of the violations is unaltered in the presence of white noise, i.e. when the ideal quantum state giving rise to the correlations gets added a maximally mixed stated with some probability δ (**Theorem 4**).

Finally, we provide tools to build Bell inequalities from communication lower bounds in the literature. Lower bounds used in practice to separate classical from quantum communication complexity are usually achieved using corruption bounds and its variants. In **Theorem 6**, we give an explicit construction which translates these bounds into a suitable Bell functional. Table 2 summarizes the new results or the improvements that we obtain in this work.

Problem	Normalized Bell violations [13]	Inefficiency-resistant Bell violations (this work)
VSP [28, 45]	$\Omega\left(\sqrt[6]{n}/\sqrt{\log n}\right)$ $d = 2^{\Theta(n \log n)}, K = 2^{\Theta(n)}$	$2^{\Omega(\sqrt[3]{n}) - O(\log n)}$ $d = 2^{O(\log n)}, K = 3$
DISJ [1, 46, 47]	N/A	$2^{\Omega(n) - O(\sqrt{n})}$ $d = 2^{O(\sqrt{n})}, K = 3$
TRIBES [11, 21]	N/A	$2^{\Omega(n) - O(\sqrt{n} \log^2 n)}$ $d = 2^{O(\sqrt{n} \log^2 n)}, K = 3$
ORT [11, 48]	N/A	$2^{\Omega(n) - O(\sqrt{n} \log n)}$ $d = 2^{O(\sqrt{n} \log n)}, K = 3$

Table 2: Comparison of the Bell violations obtained by the general construction of Buhrman *et al.* [13] for normalized Bell violations (second column) and this work, for inefficiency-resistant Bell violations (see Propositions 4, 5, 6, and 7). The parameter n is the size of the input (typically, $N = 2^n$.) See Section 6.2 for the definitions of the communication problems appearing in this table. The construction of Buhrman *et al.* only yields a violation when the gap between classical and quantum complexities is more than quadratic. In the case where the gap is too small to prove a violation, we indicate this with "N/A".

1.2 Related work

The study of the maximum violation of Bell inequalities began with Tsirelson [51], who showed that for two-outcome correlation Bell inequalities, the maximum violation is bounded above by Grothendieck's constant. Tsirelson also raised the question of whether one can have unbounded violations of Bell inequalities. More precisely, he asked whether there exist families of Bell inequalities for which the amount of the violation grows arbitrarily large.

The first answer to this question came from Mermin [39], who gave a family of Bell inequalities for which a violation exponential in the number of parties is achieved. In the years that followed, several new constructions appeared for number of parties and number of inputs [3, 33, 35, 40, 43].

The study of upper bounds on violations of normalized Bell inequalities resumed in [14], where an upper bound of $O(K^2)$ (with K the number of outputs per player) and of $2^c \leq N$ (with c the communication complexity and N the number of inputs per player) were proven. In [23] the authors proved a bound of $O(d)$ in terms of the dimension d of the local Hilbert space, and in [22], the bound in terms of the number of outputs was improved to $O(K)$. In [22], Bell inequalities are constructed for which a near optimal, but probabilistic, violation of order $\Omega(\sqrt{m}/\log m)$, with $N = K = d = m$, is proven. In [12], the same violation, although requiring $N = 2^m$ inputs, is achieved for a family of Bell inequalities and quantum distributions built using the quantum advantage in one-way communication complexity for the Hidden Matching problem (with $K = d = m$). In the same paper, a violation of order $\Omega(m/(\log m)^2)$, with $K = d = m$ and $N = 2^m/m$ is achieved with the Khot-Vishnoi game; later, Junge et al. proved in [25] that N can be reduced to $O(m^8)$. Recently, an asymmetric version of this same game was introduced to allow one of the parties to only make dichotomic measurements, with a smaller (although almost optimal for this scenario) violation $\Omega(\sqrt{m}/(\log m)^2)$ [42].

For *inefficiency-resistant* Bell inequalities, the bounds in [22] do not apply. In fact, Laplante *et al.* proved in [31] a violation exponential in the dimension and the number of outputs for this type of Bell functionals, achieved by a quantum distribution built, as in [12], from the Hidden Matching communication complexity problem.

The connection exhibited in [12] between Bell violations and communication complexity is generalized by Buhrman *et al.* in [13] where a fully general construction is given to go from a quantum communication protocol for a function f to a Bell inequality and a quantum distribution violating it when the gap between classical and quantum communication complexity is larger than quadratic. The downside to this construction is that the quantum distribution has a double exponential (in the communication) number of outputs and the protocol to implement it uses an additional double exponential amount of entanglement. Also, this result does not apply for quantum advantages in a zero-error setting.

2 Preliminaries

2.1 Quantum nonlocality

Local, quantum, and nonsignaling distributions have been widely studied in quantum information theory since the seminal paper of Bell [6]. In an experimental setting, two players share an entangled state and each player is given a measurement to perform. The outcomes of the measurements are predicted by quantum mechanics and follow some probability distribution $p(a, b|x, y)$, where a is the outcome of Alice's measurement x, and

b is the outcome of Bob's measurement y.

We consider bipartite distribution families of the form $\mathbf{p} = (p(\cdot, \cdot|x, y))_{(x,y) \in \mathcal{X} \times \mathcal{Y}}$ with inputs $(x, y) \in \mathcal{X} \times \mathcal{Y}$ determining a probability distribution $p(\cdot, \cdot|x, y)$ over the outcomes $(a, b) \in \mathcal{A} \times \mathcal{B}$, with the usual positivity and normalization constraints. The set of probability distribution families is denoted by $\mathcal{P}$. For simplicity, we call simply "distributions" such probability distribution families. The expression "Alice's marginal" refers to her marginal output distribution, that is $\sum_b p(\cdot, b|x, y)$ (and similarly for Bob).

The *local deterministic distributions*, denoted $\mathcal{L}_{\det}$, are the ones where Alice outputs according to a deterministic strategy, i.e., a (deterministic) function of x, and Bob independently outputs as a function of y, without communicating. The *local distributions* $\mathcal{L}$ are obtained by taking distributions over the local deterministic strategies. Operationally, this corresponds to protocols with shared randomness and no communication. Geometrically, $\mathcal{L}$ is the convex hull of $\mathcal{L}_{\det}$.

A *Bell test* [6] consists of estimating all the probabilities $p(a, b|x, y)$ and computing a *Bell functional*, or linear function, on these values. The Bell functional B is chosen together with a threshold τ so that any local classical distribution $\boldsymbol{\ell}$ satisfies the *Bell inequality* $B(\boldsymbol{\ell}) \leq \tau$, but the chosen distribution $\mathbf{p}$ exhibits a *Bell violation*: $B(\mathbf{p}) > \tau$. By normalizing B, we can assume without loss of generality that $\boldsymbol{\ell}$ satisfies $B(\boldsymbol{\ell}) \leq 1$ for any $\boldsymbol{\ell} \in \mathcal{L}$, and $B(\mathbf{p}) > 1$.

In this paper, we will also consider strategies that are allowed to abort the protocol with some probability. When they abort, they output the symbol $\perp$ ($\perp$ denotes a new symbol which is not in $\mathcal{A} \cup \mathcal{B}$). We will use the notation $\mathcal{L}_{\det}^{\perp}$ and $\mathcal{L}^{\perp}$ to denote local strategies that can abort, where $\perp$ is added to the possible outputs for both players. When $\ell \in \mathcal{L}_{\det}^{\perp}$ or $\mathcal{L}^{\perp}$, $\ell(a, b|x, y)$ is *not* conditioned on $a, b \neq \perp$ since $\perp$ is a valid output for such distributions.

The *quantum distributions*, denoted $\mathcal{Q}$, are the ones that result from applying measurements x, y to their part of a shared bipartite quantum state. Each player outputs his or her measurement outcome (a for Alice and b for Bob). In communication complexity terms, these are zero-communication protocols with shared entanglement. If the players are allowed to abort, then the corresponding set of distributions is denoted $\mathcal{Q}^{\perp}$.

Boolean (and other) functions can be cast as sampling problems. Consider a boolean function $f : \mathcal{X} \times \mathcal{Y} \to \{0, 1\}$ (nonboolean functions and relations can be handled similarly). First, we split the output so that if $f(x, y) = 0$, Alice and Bob are required to output the same bit, and if $f(x, y) = 1$, they output different bits. Let us further require Alice's marginal distribution to be uniform, likewise for Bob, so that the distribution is well defined. Call the resulting distribution $\mathbf{p}_f$, that is, for any $a, b \in \{0, 1\}$ and $(x, y) \in \mathcal{X} \times \mathcal{Y}$, we have $p_f(a, b|x, y) = 1/2$ if $a \oplus b = f(x, y)$, and $p_f(a, b|x, y) = 0$ otherwise, $\oplus$ being the 1-bit XOR.

If $\mathbf{p}_f$ were local, f could be computed with one bit of communication using shared randomness: Alice sends her output to Bob, and Bob XORs it with his output. If $\mathbf{p}_f$ were quantum, there would be a 1-bit protocol with shared entanglement for f. Interesting distributions have nontrivial communication complexity and, therefore, they usually lie well beyond these sets.

Finally, a distribution is *nonsignaling* if for each player, its marginal output distributions, given by $p_A(a|x, y) = \sum_b p(a, b|x, y)$, for Alice, and $p_B(b|x, y) = \sum_a p(a, b|x, y)$, for Bob, do not depend on the other player's input. When this is the case, we write the marginals as $p_A(a|x)$ and $p_B(b|y)$. Operationally, this means that each player cannot influence the statistics of what the other player observes with his own choice of input. We note with $\mathcal{C}$ the set of nonsignaling distributions, also referred to as the *causal* set, and we

note $\mathcal{C}^\perp$ when we allow aborting. The well-known inclusion relations between these sets are $\mathcal{L} \subset \mathcal{Q} \subset \mathcal{C} \subset \mathcal{P}$.

For any Boolean function f, the distribution $\mathbf{p}_f$ is nonsignaling since the marginals are uniform. A fundamental question of quantum mechanics has been to establish experimentally whether nature is truly *nonlocal*, as predicted by quantum mechanics, or whether there is a purely classical (i.e., *local*) explanation to the phenomena that have been predicted by quantum theory and observed in the lab.

2.2 Measures of nonlocality

We have described nonlocality as a yes/no property, but some distributions are somehow more nonlocal than others. To have a robust measure of nonlocality, it should fulfill some common sense properties: for a fixed distribution, the measure should be bounded; it should also be convex, since sampling from the convex combination of two distributions can be done by first picking randomly one of the two distributions using shared randomness, and then sampling from that distribution. We also expect such a measure of nonlocality to have various equivalent formulations. Several measures have been proposed and studied: resistance to noise [2, 24, 26, 43], resistance to inefficiency [31, 35, 36], amount of communication necessary to reproduce them [9, 14, 37, 44, 49, 50], information-theoretic measures [10, 16, 17], etc.

In the form studied in this paper, normalized Bell inequalities were first studied in [14], where they appeared as the dual of the linear program for a well-studied lower bound on communication complexity, known as the nuclear norm ν [34] (the definition is given in Section 2.3). There are many equivalent formulations of this bound. For distributions arising from boolean functions, it has the mathematical properties of a norm, and it is related to winning probabilities of XOR games. It can also be viewed as a gauge, that is, a quantity measuring by how much the local set must be expanded in order to contain the distribution considered. For more general nonsignaling distributions, besides having a geometrical interpretation in terms of affine combinations of local distributions, it has also been shown to be equivalent to the amount of local noise that can be tolerated before the distribution becomes local [23].

A subsequent paper [31] studied equivalent formulations of the partition bound, one of the strongest lower bounds in communication complexity [20]. This bound also also has several formulations: the primal formulation can be viewed as resistance to detector inefficiency, and the dual formulation is given in terms of inefficiency-resistant Bell inequality violations.

In this paper, we show how to deduce large violations on quantum distributions from large violations on nonsignaling distributions, provided there are efficient quantum communication protocols for the latter.

2.3 Communication complexity and lower bounds

In classical communication complexity (introduced by [52]), two players each have a share of the input, and wish to compute a function on the full input. Communication complexity measures the number of bits they need to exchange to solve this problem in the worst case, over all inputs of a given size n. In this paper we consider a generalization of this model, where instead of computing a function, they each produce an output, say a and b, which should follow, for each (x, y), some prescribed distribution $p(a, b|x, y)$ (which depends on their inputs x, y). We assume that the order in which the players speak does not depend on the inputs. This is without loss of generality at a cost of a factor of 2 in the communication.

We use the following notation for communication complexity of distributions. $R_\epsilon(\mathbf{p})$ is the minimum number of bits exchanged in the worst case between players having access to shared randomness in order to output with the distribution $\mathbf{p}$ up to ϵ in total variation distance for all x, y. We call total variation distance between distributions the distance denoted by $|.|_1$, and defined as $|\mathbf{p} - \mathbf{p}'|_1 = \max_{x,y} \sum_{a,b} |p(a,b|x,y) - p'(a,b|x,y)|$. We use Q_ϵ to denote quantum communication complexity (see [54]) and Q_ϵ^* when, in addition to the communication of quantum states, the players are allowed to share entanglement.

To give upper bounds on communication complexity it suffices to give a protocol and analyze its complexity. Proving lower bounds is often a more difficult task, and many techniques have been developed to achieve this. The methods we describe here are complexity measures which can be applied to any function. To prove a lower bound on communication, it suffices to give a lower bound on one of these complexity measures, which are bounded above by communication complexity for any function. We describe here most of the complexity measures relevant to this work.

The nuclear norm ν, given here in its dual formulation and extended to nonsignaling distributions, is expressed by the following linear program [14, 34]. (There is a quantum analogue, γ_2, which is not needed in this work. We refer the interested reader to the definition for distributions in [14]).

Definition 1 ([14, 34]). *The nuclear norm ν of a nonsignaling distribution $\mathbf{p} \in \mathcal{C}$ is given by*

$$\nu(\mathbf{p}) = \max_B \qquad\qquad B(\mathbf{p})$$
$$subject\ to \qquad\qquad |B(\boldsymbol{\ell})| \leq 1 \quad \forall \boldsymbol{\ell} \in \mathcal{L}_{det}.$$

With error ϵ, $\nu_\epsilon(\mathbf{p}) = \min_{\mathbf{p}' \in \mathcal{C}: |\mathbf{p}' - \mathbf{p}|_1 \leq \epsilon} \nu(\mathbf{p}')$. We call any Bell functional that satisfies the constraint in the above linear program normalized Bell functional.

In this definition and in the rest of the paper, unless otherwise specified (in particular in Lemma 1), B ranges over vectors of real coefficients $B_{a,b,x,y}$ and $B(\mathbf{p})$ denotes $\sum_{a,b,x,y} B_{a,b,x,y} p(a,b|x,y)$, where a, b ranges over the nonaborting outputs and x, y ranges over the inputs. So even when B and $\mathbf{p}$ have coefficients on the abort events, we do not count them. Table 1 summarizes the known upper and lower bounds on ν for various parameters. The (log of the) nuclear norm is a lower bound on classical communication complexity.

Proposition 1 ([14, 34]). *For any nonsignaling distribution $\mathbf{p} \in \mathcal{C}$, $R_\epsilon(\mathbf{p}) + 1 \geq \log(\nu_\epsilon(\mathbf{p}))$, and for any Boolean function f, $R_\epsilon(f) \geq \log(\nu_\epsilon(\mathbf{p}_f))$.*

As lower bounds on communication complexity of Boolean functions go, ν is one of the weaker bounds, equivalent to the smooth discrepancy [20], and no larger than the approximate nonnegative rank and the smooth rectangle bounds [29]. More significantly for this work, up to small multiplicative constants, for Boolean functions, (the log of) ν is a lower bound on quantum communication, so it is useless to establish gaps between classical and quantum communication complexity. (This limitation, with the upper bound in terms of the number of outputs on normalized Bell violations, is a consequence of Grothendieck's theorem [18].)

The classical and quantum efficiency bounds, given here in their dual formulations, are expressed by the following two convex optimization programs. The classical bound is a generalization to distributions of the partition bound of communication complexity [20, 31]. This bound is one of the strongest lower bounds known, and can be exponentially

larger than ν (an example is the Vector in Subspace problem [28]). It is always at least as large as the relaxed partition bound which is in turn always at least as large as the smooth rectangle bound [20, 27]. Its weaker variants have been used to show exponential gaps between classical and quantum communication complexity. The definition we give here is a stronger formulation than the one given in [31]. We show they are equivalent in Appendix E.

Definition 2 ([31]). *The ϵ-error efficiency bound of a distribution $\mathbf{p} \in \mathcal{P}$ is given by*

$$\mathbf{eff}_\epsilon(\mathbf{p}) = \max_{B, \beta} \qquad \beta$$

$$subject\ to \qquad B(\mathbf{p}') \geq \beta \qquad \forall \mathbf{p}' \in \mathcal{P} \text{ s.t. } |\mathbf{p}' - \mathbf{p}|_1 \leq \epsilon,$$

$$B(\boldsymbol{\ell}) \leq 1 \qquad \forall \boldsymbol{\ell} \in \mathcal{L}_{det}^\perp.$$

We call any Bell functional that satisfies the second constraint in the above program inefficiency-resistant Bell functional. The ϵ-error quantum efficiency bound of a $\mathbf{p} \in \mathcal{P}$ is

$$\mathbf{eff}_\epsilon^*(\mathbf{p}) = \max_{B, \beta} \qquad \beta$$

$$subject\ to \qquad B(\mathbf{p}') \geq \beta \qquad \forall \mathbf{p}' \in \mathcal{P} \text{ s.t. } |\mathbf{p}' - \mathbf{p}|_1 \leq \epsilon,$$

$$B(\mathbf{q}) \leq 1 \qquad \forall \mathbf{q} \in \mathcal{Q}^\perp.$$

We denote $\mathbf{eff} = \mathbf{eff}_0$ and $\mathbf{eff}^* = \mathbf{eff}_0^*$ the 0-error bounds.

Although it may not be straightforward from the above definition due to the presence of absolute values, the program for the classical efficiency bound is linear, as a consequence of $\mathcal{L}_{det}^\perp$ being a polytope (See Appendix E and Ref. [31]).

For any given distribution $\mathbf{p}$, its classical communication complexity is bounded from below by the (log of the) efficiency bound. For randomized communication complexity with error ϵ, the bound is $\log(\mathbf{eff}_\epsilon)$ and for quantum communication complexity, the bound is $\log(\mathbf{eff}_\epsilon^*)$. Note that for any $\mathbf{p} \in \mathcal{Q}$, the quantum communication complexity is 0 and the $\mathbf{eff}^*$ bound is 1. For any function f, the efficiency bound $\mathbf{eff}_\epsilon(\mathbf{p}_f)$ is equivalent to the partition bound [20, 31].

Proposition 2 ([31]). *For any $\mathbf{p} \in \mathcal{P}$ and any $0 \leq \epsilon < 1/2$, $\mathrm{R}_\epsilon(\mathbf{p}) \geq \log(\mathbf{eff}_\epsilon(\mathbf{p}))$ and $Q_\epsilon^*(\mathbf{p}) \geq \frac{1}{2} \log(\mathbf{eff}_\epsilon^*(\mathbf{p}))$. For any $\mathbf{p} \in \mathcal{C}$ and any $0 \leq \epsilon \leq 1$, $\nu_\epsilon(\mathbf{p}) \leq 2\mathbf{eff}_\epsilon(\mathbf{p})$.*

Theorem 3 in Section 4 below involves upper bounds on the quantum efficiency bound. To give an upper bound on the quantum efficiency bound of a distribution $\mathbf{p}$, it is more convenient to use the primal formulation, and upper bounds can be given by exhibiting a local (or quantum) distribution with abort which satisfies the following two properties: the probability of aborting should be the same on all inputs x, y, and conditioned on not aborting, the outputs of the protocol should reproduce the distribution $\mathbf{p}$. The efficiency bound is inverse proportional to the probability of not aborting, so the goal is to abort as little as possible.

Proposition 3 ([31]). *For any distribution $\mathbf{p} \in \mathcal{P}$, $\mathbf{eff}^*(\mathbf{p}) = 1/\eta^*$, with η^* the optimal value of the following optimization problem (nonlinear, because $\mathcal{Q}^\perp$ is not a polytope).*

$$\max_{\mathbf{q} \in \mathcal{Q}^\perp, \zeta} \qquad \zeta$$

$$subject\ to \quad q(a,b|x,y) = \zeta p(a,b|x,y) \quad \forall x,y,a,b \in \mathcal{X} \times \mathcal{Y} \times \mathcal{A} \times \mathcal{B}$$

Moreover, for any $0 \leq \epsilon \leq 1$, $\mathbf{eff}_\epsilon^*(\mathbf{p}) = \min_{\mathbf{p}' \in \mathcal{P}: |\mathbf{p}' - \mathbf{p}|_1 \leq \epsilon} \mathbf{eff}^*(\mathbf{p}')$.

3 Properties of Bell inequalities

Syntactically, there are two differences between the normalized Bell functionals (Definition 1) and the inefficiency-resistant ones (Definition 2). The first difference is that the normalization constraint is relaxed: for inefficiency-resistant functionals, the Bell value for local distributions is, unlike the case for normalized Bell functionals, not bounded in absolute value; it is only bounded from above. Since this is a maximization problem, this relaxation allows for larger violations. This difference alone would not lead to a satisfactory measure of nonlocality, since one could obtain unbounded violations by shifting and dilating the Bell functional. The second difference prevents this. The upper bound is required to hold for a larger set of local distributions, those that can abort. This is a much stronger condition. Notice that a local distribution can selectively abort on configurations that would otherwise tend to keep the Bell value small, making it harder to satisfy the constraint.

In this section, we show that normalized Bell violations can be modified to be resistant to local distributions that abort, while preserving the violation on any nonsignaling distribution, up to a multiplicative factor of 3 (plus a term independent of the input and output sizes). This means that we can add the stronger constraint of resistance to local distributions that abort to Definition 1, incurring a loss of essentially a factor of 3, and the only remaining difference between the resulting linear programs is the relaxation of the lower bound (dropping the absolute value) for local distributions that abort.

Theorem 1. *Let B be a normalized Bell functional on $\mathcal{A} \times \mathcal{B} \times \mathcal{X} \times \mathcal{Y}$ and $\mathbf{p} \in \mathcal{C}$ a nonsignaling distribution such that $B(\mathbf{p}) \geq 1$. Then there exists a normalized Bell functional B^* on $(\mathcal{A} \cup \{\bot\}) \times (\mathcal{B} \cup \{\bot\}) \times \mathcal{X} \times \mathcal{Y}$ with 0 coefficients on the $\bot$ outputs such that : $\forall \mathbf{p} \in \mathcal{C}$, $B^*(\mathbf{p}) \geq \frac{1}{3}B(\mathbf{p}) - \frac{2}{3}$, and $\forall \ell \in \mathcal{L}_{\det}^{\bot}$, $|B^*(\ell)| \leq 1$.*

The formal proof of Theorem 1 is deferred to Appendix A, and we will only give its high-level structure in this part of the paper. First, we show (see Observation 1) how to rescale a normalized Bell functional so that it saturates its normalization constraint. Then, Definition 3 adds weights to abort events to make the Bell functional resistant to inefficiency. Finally, Lemma 1 removes the weights on the abort events of a Bell functional while keeping it bounded on the local set with abort, without dramatically changing the values it takes on the nonsignaling set. Our techniques are similar to the ones used in [36].

4 Exponential violations from communication bounds

Recently, Buhrman *et al.* gave a general construction to derive normalized Bell inequalities from any sufficiently large gap between classical and quantum communication complexity.

Theorem 2 ([13]). *For any function f for which there is a quantum protocol using q qubits of communication but no prior shared entanglement, there exists a quantum distribution $\mathbf{q} \in \mathcal{Q}$ and a normalized Bell functional B such that $B(\mathbf{q}) \geq \frac{\sqrt{R_{1/3}(f)}}{6\sqrt{30}q}(1 - 2^{-q})^{2q}$.*

Their construction is quite involved, requiring protocols to be memoryless, which they show how to achieve in general, and uses multiport teleportation to construct a quantum distribution. The Bell inequality they construct expresses a correctness constraint.

In this section, we show how to obtain large inefficiency-resistant Bell violations for quantum distributions from gaps between quantum communication complexity and the efficiency bound for classical communication complexity. We first prove the stronger of

two statements, which gives violations of $\frac{\mathbf{eff}_\epsilon(\mathbf{p})}{\mathbf{eff}^\star_{\epsilon'}(\mathbf{p})}$. For any problem for which a classical lower bound c is given using the efficiency or partition bounds or any weaker method (including the rectangle bound and its variants), and any upper bound q on quantum communication complexity, it implies a violation of 2^{c-2q}.

Theorem 3. *For any distribution $\mathbf{p} \in \mathcal{P}$ and any $0 \le \epsilon' \le \epsilon \le 1$, if (B, β) is a feasible solution to the dual of $\mathbf{eff}_\epsilon(\mathbf{p})$ and $(\zeta, \mathbf{q})$ is a feasible solution to the primal for $\mathbf{eff}^\star_{\epsilon'}(\mathbf{p})$, then there is a quantum distribution $\overline{\mathbf{q}} \in \mathcal{Q}$ such that $B(\overline{\mathbf{q}}) \ge \zeta\beta$ and $B(\boldsymbol{\ell}) \le 1, \forall \boldsymbol{\ell} \in \mathcal{L}^\perp_{det}$, and in particular, if both are optimal solutions, then $B(\overline{\mathbf{q}}) \ge \frac{\mathbf{eff}_\epsilon(\mathbf{p})}{\mathbf{eff}^\star_{\epsilon'}(\mathbf{p})}$. The distribution $\overline{\mathbf{q}}$ has one additional output per player compared to the distribution $\mathbf{p}$.*

Proof. Let (B, β) be a feasible solution to the dual of $\mathbf{eff}_\epsilon(\mathbf{p})$, $\mathbf{p}'$ be such that $\mathbf{eff}^\star_{\epsilon'}(\mathbf{p}) = \mathbf{eff}^\star(\mathbf{p}')$ with $|\mathbf{p}' - \mathbf{p}|_1 \le \epsilon'$, and $(\zeta, \mathbf{q})$ be a feasible solution to the primal for $\mathbf{eff}^\star(\mathbf{p}')$. From the constraints, we have $\mathbf{q} \in \mathcal{Q}^\perp$, $q(a, b|x, y) = \zeta p'(a, b|x, y)$ for all $(a, b, x, y) \in \mathcal{A} \times \mathcal{B} \times \mathcal{X} \times \mathcal{Y}$, $B(\boldsymbol{\ell}) \le 1$ for all $\boldsymbol{\ell} \in \mathcal{L}^\perp_{det}$, and $B(\mathbf{p}'') \ge \beta$ for all $\mathbf{p}''$ s.t. $|\mathbf{p}'' - \mathbf{p}|_1 \le \epsilon$. Then $B(\mathbf{q}) = \zeta B(\mathbf{p}') \ge \zeta\beta$. However, $\mathbf{q} \in \mathcal{Q}^\perp$ but technically we want a distribution in $\mathcal{Q}$ (not one that aborts). So we add a new (valid) output 'A' to the set of outputs of each player, and they should output 'A' instead of aborting whenever $\mathbf{q}$ aborts. The resulting distribution, say $\overline{\mathbf{q}} \in \mathcal{Q}$ (with additional outcomes 'A' on both sides), is such that $B(\overline{\mathbf{q}}) = B(\mathbf{q})$ (since the Bell functional B does not have any weight on $\perp$ or on 'A'). $\square$

Theorem 2 and Theorem 3 are both general constructions, but there are a few significant differences. Firstly, Theorem 3 requires a lower bound on the partition bound in the numerator, whereas Theorem 2 only requires a lower bound on communication complexity (which could be exponentially larger). Secondly, Theorem 2 requires a quantum communication protocol in the denominator, whereas our theorem only requires an upper bound on the quantum efficiency bound (which could be exponentially smaller). Thirdly, our bound is exponentially larger than Buhrman *et al.*'s for most problems considered here, and applies to subquadratic gaps, but their bounds are of the more restricted class of normalized Bell inequalities. These differences imply, in principle, different applicability regimes for the two constructions, which we depict in Figure 1.

Theorem 3 gives an explicit Bell functional provided an explicit solution to the efficiency (partition) bound is given and the quantum distribution is obtained from a solution to the primal of $\mathbf{eff}^\star$ (Proposition 3). Recall that a solution to the primal of $\mathbf{eff}^\star$ is provided by a quantum zero-communication protocol that can abort, which conditioned on not aborting, outputs following $\mathbf{p}$. We can also start from a quantum protocol, as we show below. From the quantum protocol, we derive a quantum distribution using standard techniques.

Corollary 1. *For any distribution $\mathbf{p} \in \mathcal{P}$ and any $0 \le \epsilon' \le \epsilon \le 1$ such that $R_\epsilon(\mathbf{p}) \ge \log(\mathbf{eff}_\epsilon(\mathbf{p})) \ge c$ and $Q^*_{\epsilon'}(\mathbf{p}) \le q$, there exists an explicit inefficiency-resistant B derived from the efficiency lower bound, and an explicit quantum distribution $\overline{\mathbf{q}} \in \mathcal{Q}$ derived from the quantum protocol such that $B(\overline{\mathbf{q}}) \ge 2^{c-2q}$.*

Proof. Let (B, β) be an optimal solution to $\mathbf{eff}_\epsilon(\mathbf{p})$ and let c be such that $\mathbf{eff}_\epsilon(\mathbf{p}) = \beta \ge 2^c$. By optimality of B, we have $B(\mathbf{p}') \ge 2^c$ for any $\mathbf{p}'$ such that $|\mathbf{p}' - \mathbf{p}|_1 \le \epsilon$. Since $Q^*_{\epsilon'}(\mathbf{p}) \le q$, there exists a q-qubit quantum protocol (possibly using preshared entanglement) for some distribution $\mathbf{p}'$ with $|\mathbf{p}' - \mathbf{p}|_1 \le \epsilon' \le \epsilon$. Then, we can use teleportation to obtain a $2q$ classical bit, entanglement-assisted protocol for $\mathbf{p}'$. We can simulate it without communication by picking a shared $2q$-bit random string and running the protocol but without sending any messages. If the measurements do not match the string, output a
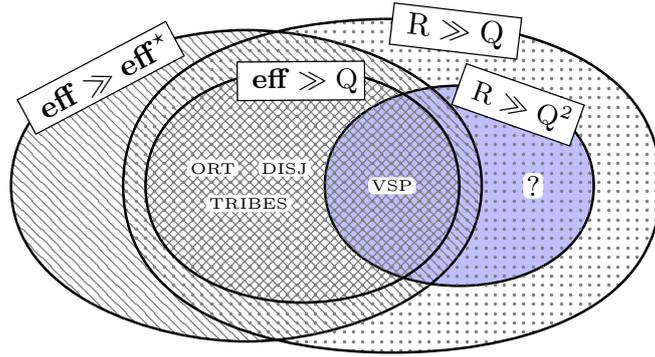
Figure 1: An illustration of the different applicability regions of Theorem 2 and Theorem 3. There are problems in the literature for which both constructions apply, e.g. VSP, and problems for which only our construction applies, e.g. DISJ, ORT, and TRIBES (see Table 2). Whether there exists any problem in the region in which $R \gg Q^2$ but $\mathbf{eff} \approx \mathbf{eff}^*$ (indicated by a question mark in the picture), for which Theorem 2 applies but our Theorem 3 does not, is, to our current knowledge, an open question.

new symbol 'A' (not in the output set of the quantum protocol and different from $\bot$). We obtain a quantum distribution $\overline{\mathbf{q}}$ such that $B(\overline{\mathbf{q}}) = B(\mathbf{p}')/2^{2q} \geq 2^{c-2q}$. $\qquad\square$

Most often, communication lower bounds are not given as efficiency or partition bounds, but rather using variants of the corruption bound. We show in Section 6.1 how to map a corruption bound to explicit Bell coefficients.

5 Noise-resistant violations from communication bounds

Normalized Bell inequalities are naturally resistant to any constant level of local noise in the sense that the value of the Bell functional can only decrease by constant additive and multiplicative terms. Indeed, if the observed distribution is $\tilde{\mathbf{p}} = (1-\varepsilon)\mathbf{p} + \varepsilon\boldsymbol{\ell}$ for some $\boldsymbol{\ell} \in \mathcal{L}$, then $B(\tilde{\mathbf{p}}) \geq (1-\varepsilon)B(\mathbf{p}) - \varepsilon$ since $\mid B(\boldsymbol{\ell}) \mid \leq 1$. In inefficiency-resistant Bell inequalities, relaxing the absolute value leads to the possibility that $B(\boldsymbol{\ell})$ has a large negative value for some local $\boldsymbol{\ell}$. (Indeed, such large negative values are inherent to large gaps between ν and $\mathbf{eff}$.) If this distribution were used as adversarial noise, the observed distribution, $(1-\varepsilon)\mathbf{p} + \varepsilon\boldsymbol{\ell}$, could have a Bell value much smaller than $B(\mathbf{p})$. This makes inefficiency-resistant Bell inequalities susceptible to adversarial local noise.

In Theorem 4 below, whose proof we defer to Appendix B, we show that the quantum violations arising from our construction in Corollary 1 are resistant to the addition of white noise. More specifically, we consider the standard noise model (see e.g., [15, 26, 36]) in which the (ideal) quantum state $|\psi\rangle$ gets added a maximally mixed state with some probability δ,

$$\rho := (1-\delta)|\psi\rangle\langle\psi| + \delta\frac{\mathbb{I}}{d}.$$

We show that even in this noisy scenario, the inefficiency-resistant Bell functional B and the quantum distribution $\mathbf{q}$ resulting from the construction in Corollary 1 are such that $B(\mathbf{q}) \geq 2^{c-2q}$. Intuitively, resistance to this kind of noise follows from: 1) B arising from a lower bound on bounded-error communication complexity and 2) the teleportation measurements needed to construct $\mathbf{q}$ still giving uniform outcomes when we replace a maximally entangled with a maximally mixed state.

Theorem 4. *For any* $\mathbf{p} \in \mathcal{P}$ *and any* $0 \leq \epsilon' \leq \epsilon \leq 1$ *such that* $\log(\mathbf{eff}_\epsilon(\mathbf{p})) \geq c$ *and* $Q^*_{\epsilon'}(\mathbf{p}) \leq q$, *there exists an explicit inefficiency-resistant* $\overline{B}$, *POVMs* $\{E_{a|x}\}_{a \in \mathcal{A} \cup \{S\}}$ *over a Hilbert space* $\mathcal{H}_\mathcal{A}$ *and* $\{E_{b|y}\}_{b \in \mathcal{B} \cup \{S\}}$ *over a Hilbert space* $\mathcal{H}_\mathcal{B}$ *and a quantum state* $|\psi\rangle \in \mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B}$, *such that for any* $0 \leq \delta \leq \epsilon - \epsilon'$, *we have that*

$$\overline{B}(\mathbf{q}_\delta) \geq 2^{c - 2q},$$

with

$$q_\delta(a, b | x, y) = \mathrm{tr}\left[E_{a|x} \otimes E_{b|y} \left((1 - \delta)|\psi\rangle\langle\psi| + \delta \frac{\mathbb{I}}{\dim(\mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B})} \right) \right].$$

Notice that $\mathbf{q}$ *has one additional output per player (labelled S) compared to* $\mathbf{p}$.

6 Explicit constructions

6.1 From corruption bound to Bell inequality violation

We now explain how to construct an explicit Bell inequality violation from the corruption bound. The corruption bound, introduced by Yao in [53], is a very useful lower bound technique. It has been used for instance in [46] to get a tight $\Omega(n)$ lower bound on the randomized communication complexity of Disjointness (whereas the approximate rank, for example, can only show a lower bound of $\Theta(\sqrt{n})$). Let us recall that a rectangle R of $\mathcal{X} \times \mathcal{Y}$ is a subset of that set of the form $R_A \times R_B$, where $R_A \subseteq \mathcal{X}$ and $R_B \subseteq \mathcal{Y}$.

Theorem 5 (Corruption bound [4, 30, 53]). *Let* f *be a (possibly partial) Boolean function on* $\mathcal{X} \times \mathcal{Y}$. *Given* $\gamma, \delta \in (0, 1)$, *suppose that there is a distribution* μ *on* $\mathcal{X} \times \mathcal{Y}$ *such that for every rectangle* $R \subseteq \mathcal{X} \times \mathcal{Y}$

$$\mu(R \cap f^{-1}(1)) > \gamma \mu(R \cap f^{-1}(0)) - \delta$$

Then, for every $\epsilon \in (0, 1)$, $2^{R_\epsilon(f)} \geq \frac{1}{\delta}\left(\mu(f^{-1}(0)) - \frac{\epsilon}{\gamma} \right)$.

See, e.g., Lemma 3.5 in [5] for a rigorous treatment. For several problems, such a μ is already known. In Theorem 6 below, whose proof we defer to Appendix C, we show how to construct a Bell inequality violation from this type of bound.

Theorem 6. *Let* f *be a (possibly partial) Boolean function on* $\mathcal{X} \times \mathcal{Y}$, *where* $\mathcal{X}, \mathcal{Y} \subseteq \{0, 1\}^n$. *Fix* $z \in \{0, 1\}$. *Let* μ *be an input distribution, and* $(U_i)_{i \in I}$ *(resp.* $(V_j)_{j \in J}$*) be a family of pairwise nonoverlapping subsets of* $f^{-1}(\bar{z})$ *(resp. of* $f^{-1}(z)$*). Assume that there exists* $g : \mathbb{N} \to (0, +\infty)$ *such that, for any rectangle* $R \subseteq \mathcal{X} \times \mathcal{Y}$

$$\sum_{i \in I} u_i \mu(R \cap U_i) \geq \sum_{j \in J} v_j \mu(R \cap V_j) - g(n). \tag{1}$$

Then, the Bell functional B *given by the following coefficients: for all* $a, b, x, y \in \{0, 1\} \times \{0, 1\} \times \mathcal{X} \times \mathcal{Y}$,

$$B_{a,b,x,y} = \begin{cases} 1/2(-u_i g(n)^{-1} \mu(x, y)) & \text{if } (x, y) \in U_i \text{ and } a \oplus b = z, \\ 1/2(v_j g(n)^{-1} \mu(x, y)) & \text{if } (x, y) \in V_j \text{ and } a \oplus b = z, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

satisfies

$$B(\boldsymbol{\ell}) \leq 1, \quad \forall \boldsymbol{\ell} \in \mathcal{L}_{det}^\perp, \tag{3}$$

$$B(\mathbf{p}_f) = \frac{1}{2 \cdot g(n)} \sum_j v_j \mu(V_j) \tag{4}$$

and for any $\mathbf{p}' \in \mathcal{P}$ such that $|\mathbf{p}' - \mathbf{p}_f|_1 \leq \epsilon$:

$$B(\mathbf{p}') \geq \frac{1}{2 \cdot g(n)} \left[\sum_j v_j \mu(V_j) - \epsilon \left(\sum_j |v_j| \mu(V_j) + \sum_i |u_i| \mu(U_i) \right) \right]. \qquad (5)$$

For many other problems in the literature, such as Vector in Subspace and Tribes, stronger variants of the corruption bound are needed to obtain good lower bounds. These stronger variants have been shown to be no stronger than the partition bound (more specifically, the relaxed partition bound) [27]. The generalization in Theorem 6 of the hypothesis of Theorem 5, which the reader might have noticed, allow us to construct explicit Bell functionals also for these problems.

6.2   Some specific examples

Using Corollary 1 and the construction to go from a corruption bound (or its variants) to a Bell inequality (Theorem 6), we give explicit Bell inequalities and violations for several problems studied in the literature. Since our techniques also apply to small gaps, we include problems for which the gap between classical and quantum communication complexity is polynomial.

Vector in Subspace   In the Vector in Subspace Problem $\mathrm{VSP}_{0,n}$, Alice is given an $n/2$ dimensional subspace of an n dimensional space over $\mathbb{R}$, and Bob is given a vector. This is a partial function, and the promise is that either Bob's vector lies in the subspace, in which case the function evaluates to 1, or it lies in the orthogonal subspace, in which case the function evaluates to 0. Note that the input set of $\mathrm{VSP}_{0,n}$ is continuous, but it can be discretized by rounding, which leads to the problem $\widetilde{\mathrm{VSP}}_{\theta,n}$ (see [28] for details). Klartag and Regev [28] show that the VSP can be solved with an $O(\log n)$ quantum protocol, but the randomized communication complexity of this problem is $\Omega(n^{1/3})$. As shown in [27], this is also a lower bound on the relaxed partition bound. Hence Corollary 1 yields the following.

Proposition 4. *There exists a Bell inequality B and a quantum distribution $\overline{\mathbf{q}}_{VSP} \in \mathcal{Q}$ such that $B(\overline{\mathbf{q}}_{VSP}) \in 2^{\Omega(n^{1/3}) - O(\log n)}$ and for all $\ell \in \mathcal{L}_{det}^{\perp}$, $B(\ell) \leq 1$.*

Note that the result of [28] (Lemma 4.3) is not of the form needed to apply Theorem 6. It is yet possible to obtain an explicit Bell functional following the proof of Lemma 5.1 in [27].

Disjointness   In the Disjointness problem, the players receive two sets and have to determine whether they are disjoint or not. More formally, the Disjointness predicate is defined over $\mathcal{X} = \mathcal{Y} = \mathcal{P}([n])$ by $\mathrm{DISJ}_n(x,y) = 1$ iff x and y are disjoint. It is also convenient to see this predicate as defined over length n inputs, where $\mathrm{DISJ}_n(x,y) = 1$ for $x,y \in \{0,1\}^n$ if and only if $|\{i : x_i = 1 = y_i\}| = 0$. The communication complexity for DISJ_n is $\Omega(n)$ using a corruption bound [46] and there is a quantum protocol using $O(\sqrt{n})$ communication [1]. Combining these results with ours, we obtain the following.

Proposition 5. *There is a quantum distribution $\overline{\mathbf{q}}_{\mathrm{DISJ}} \in \mathcal{Q}$ and an explicit Bell inequality B satisfying: $B(\overline{\mathbf{q}}_{\mathrm{DISJ}}) = 2^{\Omega(n) - O(\sqrt{n})}$, and for all $\ell \in \mathcal{L}_{det}^{\perp}$, $B(\ell) \leq 1$.*

The proof is deferred to the Appendix (see Section D.1).

Tribes. Let $r \geq 2$, $n = (2r + 1)^2$. Let $\mathrm{TRIBES}_n : \{0, 1\}^n \times \{0, 1\}^n \to \{0, 1\}$ be defined as: $\mathrm{TRIBES}_n(x, y) := \bigwedge_{i=1}^{\sqrt{n}} \left(\bigvee_{j=1}^{\sqrt{n}} \left(x_{(i-1)\sqrt{n}+j} \wedge y_{(i-1)\sqrt{n}+j} \right) \right)$. The Tribes function has an $\Omega(n)$ classical lower bound [19] using the smooth rectangle bound and a $O(\sqrt{n}(\log n)^2)$ quantum protocol [11]. Combining these results with ours, we obtain the following.

Proposition 6. *There is a quantum distribution $\overline{\mathbf{q}}_{\mathrm{TRIBES}} \in \mathcal{Q}$ and an explicit Bell inequality B satisfying: $B(\overline{\mathbf{q}}_{\mathrm{TRIBES}}) = 2^{\Omega(n) - O(\sqrt{n}(\log n)^2)}$, and for all $\boldsymbol{\ell} \in \mathcal{L}_{det}^{\perp}$, $B(\boldsymbol{\ell}) \leq 1$.*

The proof is deferred to the Appendix (see Section D.2).

Gap Orthogonality. The Gap Orthogonality (ORT) problem was introduced by Sherstov as an intermediate step to prove a lower bound for the Gap Hamming Distance (GHD) problem [48]. We derive an explicit Bell inequality for ORT from Sherstov's lower bound of $\Omega(n)$, shown in [27] to be a relaxed partition bound. (Applying Corollary 1 also gives a (nonexplicit) violation for GHD.) The quantum upper bound is $O(\sqrt{n} \log n)$ by the general result of [11]. In the ORT problem, the players receive vectors and need to tell whether they are nearly orthogonal or far from orthogonal. More formally, we consider the input space $\{-1, +1\}^n$ (to stick to the usual notations for this problem), and we denote $\langle \cdot, \cdot \rangle$ the scalar product on $\{-1, +1\}^n$. Let $\mathrm{ORT}_n : \{-1, +1\}^n \times \{-1, +1\}^n \to \{-1, +1\}$ be the partial function defined as in [48] by: $\mathrm{ORT}_n(x, y) = -1$ if $|\langle x, y \rangle| \leq \sqrt{n}$, and $\mathrm{ORT}_n(x, y) = +1$ if $|\langle x, y \rangle| \geq 2\sqrt{n}$. Combining the results mentioned above with ours, we obtain the following.

Proposition 7. *There is a quantum distribution $\overline{\mathbf{q}}_{\mathrm{ORT}} \in \mathcal{Q}$ and an explicit Bell inequality B satisfying: $B(\overline{\mathbf{q}}_{\mathrm{ORT}}) = 2^{\Omega(n) - O(\sqrt{n} \log n)}$, and for all $\boldsymbol{\ell} \in \mathcal{L}_{det}^{\perp}$, $B(\boldsymbol{\ell}) \leq 1$.*

The proof is deferred to the Appendix (see Section D.3).

7 Discussion

We have given three main results. First, we showed that normalized Bell inequalities can be modified to be bounded in absolute value on the larger set of local distributions that can abort without significantly changing the value of the violations achievable with nonsignaling distributions. Then, we showed how to derive large inefficiency-resistant Bell violations from any gap between the partition bound and the quantum communication complexity of some given distribution $\mathbf{p}$. The distributions $\mathbf{q}$ achieving the large violations are relatively simple (only 3 outputs for boolean distributions $\mathbf{p}$) and are resistant to the addition of white noise to the ideal entangled state giving rise to them. Finally, we showed how to construct explicit Bell inequalities when the separation between classical and quantum communication complexity is proven via the corruption bound.

From a practical standpoint, the specific Bell violations we have studied are probably not feasible to implement, because the parameters needed are still impractical or the quantum states are infeasible to implement. However, our results suggest that we could consider functions with small gaps in communication complexity, in order to find practical Bell inequalities that are robust against uniform noise and detector inefficiency. As an example, let us show how we could design a Bell test from a gap in communication complexity using our techniques. Consider a boolean function f with a communication complexity lower bound $R_\epsilon(\mathbf{p}_f) \geq \log(\mathbf{eff}_\epsilon(\mathbf{p}_f)) \geq c(n)$, and accepting a quantum protocol using at most $q(n)$ qubits of communication with error at most $\epsilon' < \epsilon$,

where n is the size of the inputs. The construction from Theorem 4 then leads to a Bell inequality $\overline{B}$, a quantum state $|\psi\rangle \in \mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B}$ and POVMs $\{E_{a|x}\}$ and $\{E_{b|y}\}$ over $\mathcal{H}_\mathcal{A}$ and $\mathcal{H}_\mathcal{B}$ respectively, such that, for any $\delta \leq \epsilon - \epsilon'$, $\overline{B}(\mathbf{q}_\delta) \geq 2^{c(n)-2q(n)}$ for $\mathbf{q}_\delta(a,b|x,y) = \mathrm{tr}\left[E_{a|x} \otimes E_{b|y}\left((1-\delta)|\psi\rangle\langle\psi| + \delta\frac{\mathbb{I}}{\dim(\mathcal{H}_\mathcal{A}\otimes\mathcal{H}_\mathcal{B})}\right)\right]$. More precisely, $|\psi\rangle$ is obtained by preparing $q(n)$ EPR pairs shared between Alice and Bob, and the POVMs $\{E_{a|x}\}$ and $\{E_{b|y}\}$ arise from the players performing the local unitaries described by the quantum protocol together with a total of $q(n)$ local Bell measurements (each qubit communicated by Alice to Bob in the quantum communication protocol for $\mathbf{p}_f$ will turn into a Bell measurement on Alice's side, and vice versa on Bob's side) and a final measurement on both sides to obtain outcomes (a,b) in $\mathbf{p}_f$. Practically, these measurements will not always be successful due to the imperfect efficiency of the detectors. Suppose that the efficiency for each Bell measurement is η_0, and that the overall efficiency for the final measurements is at least $\eta_0^{2q(n)}$ (these final measurements involve $2q(n)$ qubits, which justifies such an expected scaling). Therefore, factoring in this efficiency and assuming that the noise affecting the experiment is $\delta < \epsilon - \epsilon'$, the observed Bell value will be $\tilde{B}(\mathbf{q}') \geq 2^{c(n)-2q(n)}\eta_0^{2q(n)}$, hence we still observe a Bell violation (a value $\tilde{B}(\mathbf{q}') > 1$) as long as n is chosen high enough (or the detector efficiencies are large enough) to have $\eta_0 > 2^{-\left(\frac{c(n)}{2q(n)}-1\right)}$.

The preceding analysis, although accounting for detectors whose efficiency decreases exponentially with the number of EPR measurements $q(n)$, assumes that the (uniform) δ-noise affects the quantum state both globally and independently of n. This is a quite standard noise model, used in various studies of Bell inequalities such as [15, 26, 36]. In a practical scenario, one might nevertheless want to consider an arguably more realistic noise model where the error increases with the dimension of the global quantum state (this would happen for example in the case of noisy teleportation, where each individual EPR pair gets added a maximally mixed state). As is, our construction is not resistant to such a noise model as it is based on a classical lower bound for *bounded-error* communication complexity, which can therefore only tolerate bounded noise. One way around this issue would be to turn the bounded-error classical lower bound into a bound valid for low success probabilities using standard amplification techniques (see, e.g. [30]). Another possible solution would be to reduce the noise in the quantum protocol by using quantum error correction (see, e.g. [41]). We leave these ideas for future work.

Lastly, we comment on upper bounds for the violation of inefficiency-resistant Bell inequalities. First, since (the log of) the efficiency bound is a lower bound on communication complexity, these violations are bounded above by the number of inputs per side. Next, for dimension d and number of outcomes K, we obtain the upper bound $\mathbf{eff}_\epsilon(\mathbf{q}) \leq 2^{O((\frac{Kd}{\epsilon})^2 \log^2(K))}$ for quantum distributions, by combining known bounds. Indeed, we know that $R_\epsilon(\mathbf{p}) \leq O((\frac{K\nu(\mathbf{p})}{\epsilon})^2 \log^2(K))$ for any $\mathbf{p} \in \mathcal{C}$ (see [14]). Combining this with the bounds $\mathbf{eff}_\epsilon(\mathbf{p}) \leq 2^{R_\epsilon(\mathbf{p})}$ (Proposition 2), and $\nu(\mathbf{q}) \leq O(d)$ for any $\mathbf{q} \in \mathcal{Q}$ (see [23]), gives the desired upper bound. Hence unbounded violations are possible for $K = 3$ outputs per side.

Acknowledgments

References

[1] S. Aaronson and A. Ambainis. Quantum search of spatial regions. *Theory of Computing*, 1:47–79, 2005. DOI: 10.4086/toc.2005.v001a004.

[2] A. Acín, T. Durt, N. Gisin, and J. I. Latorre. Quantum nonlocality in two three-level systems. *Physical Review A*, 65:052325, 2002. DOI: 10.1103/PhysRevA.65.052325.

[3] M. Ardehali. Bell inequalities with a magnitude of violation that grows exponentially with the number of particles. *Physical Review A*, 46:5375–5378, 1992. DOI: 10.1103/PhysRevA.46.5375.

[4] L. Babai, P. Frankl, and J. Simon. Complexity classes in communication complexity theory. In *Proc. 27th FOCS*, pages 337–347. IEEE, 1986. DOI: 10.1109/SFCS.1986.15.

[5] P. Beame, T. Pitassi, N. Segerlind, and A. Wigderson. A strong direct product theorem for corruption and the multiparty communication complexity of disjointness. *Computational Complexity*, 15(4):391–432, 2006. DOI: 10.1007/s00037-007-0220-2.

[6] J. S. Bell. On the Einstein Podolsky Rosen paradox. *Physics*, 1:195, 1964. DOI: 10.1103/PhysicsPhysiqueFizika.1.195.

[7] C. Branciard and N. Gisin. Quantifying the nonlocality of Greenberger-Horne-Zeilinger quantum correlations by a bounded communication simulation protocol. *Physical Review Letters*, 107(2):020401, 2011. DOI: 10.1103/PhysRevLett.107.020401.

[8] J. B. Brask and R. Chaves. Bell scenarios with communication. *Journal of Physics A: Mathematical and Theoretical*, 50(9):094001, 2017. DOI: 10.1088/1751-8121/aa5840.

[9] G. Brassard, R. Cleve, and A. Tapp. Cost of exactly simulating quantum entanglement with classical communication. *Physical Review Letters*, 83(9):1874, 1999. DOI: 10.1103/PhysRevLett.83.1874.

[10] N. Brunner, D. Cavalcanti, A. Salles, and P. Skrzypczyk. Bound nonlocality and activation. *Physical Review Letters*, 106:020402, 2011. DOI: 10.1103/PhysRevLett.106.020402.

[11] H. Buhrman, R. Cleve, and A. Wigderson. Quantum vs classical communication and computation. In *Proc. 30th STOC*, pages 63–68, 1998. DOI: 10.1145/276698.276713.

[12] H. Buhrman, O. Regev, G. Scarpa, and R. de Wolf. Near-optimal and explicit Bell inequality violations. *Theory of Computing*, 8(1):623–645, 2012. DOI: 10.4086/toc.2012.v008a027.

[13] H. Buhrman, Ł. Czekaj, A. Grudka, Mi. Horodecki, P. Horodecki, M. Markiewicz, F. Speelman, and S. Strelchuk. Quantum communication complexity advantage implies violation of a Bell inequality. *Proceedings of the National Academy of Sciences*, 113(12):3191–3196, 2016. DOI: 10.1073/pnas.1507647113.

[14] J. Degorre, M. Kaplan, S. Laplante, and J. Roland. The communication complexity of non-signaling distributions. *Quantum Information & Computation*, 11(7-8):649–676, 2011. URL http://dl.acm.org/citation.cfm?id=2230916.2230924.

[15] T. Durt, D. Kaszlikowski, and M. Żukowski. Violations of local realism with quantum systems described by N-dimensional Hilbert spaces up to $N = 16$. *Physical Review A*, 64(2):024101, 2001. DOI: 10.1103/PhysRevA.64.024101.

[16] M. Forster, S. Winkler, and S. Wolf. Distilling nonlocality. *Physical Review Letters*, 102:120401, 2009. DOI: 10.1103/PhysRevLett.102.120401.

[17] R. Gallego, L. E. Würflinger, A. Acín, and M. Navascués. Operational framework for nonlocality. *Physical Review Letters*, 109:070401, 2012. DOI: 10.1103/PhysRevLett.109.070401.

[18] A. Grothendieck. Résumé de la théorie métrique des produits tensoriels topologiques. *Boletim da Sociedade de Matemática de São Paulo*, 8:1–79, 1953. URL https://www.revistas.usp.br/resenhasimeusp/article/view/74836.

[19] P. Harsha and R. Jain. A strong direct product theorem for the tribes function via the smooth-rectangle bound. In *Proc. 33rd FSTTCS*, volume 24, pages 141–152, 2013. ISBN 978-3-939897-64-4. DOI: 10.4230/LIPIcs.FSTTCS.2013.141.

[20] R. Jain and H. Klauck. The partition bound for classical communication complexity and query complexity. In *Proc. 25th CCC*, pages 247–258, 2010. DOI: 10.1109/CCC.2010.31.

[21] T. S. Jayram, R. Kumar, and D. Sivakumar. Two applications of information complexity. In *Proc. 35th STOC*, pages 673–682, 2003. ISBN 1-58113-674-9. DOI: 10.1145/780542.780640.

[22] M. Junge and C. Palazuelos. Large violation of Bell inequalities with low entanglement. *Communications in Mathematical Physics*, 306(3):695–746, 2011. DOI: 10.1007/s00220-011-1296-8.

[23] M. Junge, C. Palazuelos, D. Pérez-García, I. Villanueva, and M. M. Wolf. Unbounded violations of bipartite Bell inequalities via operator space theory. *Communications in Mathematical Physics*, 300(3):715–739, 2010. DOI: 10.1007/s00220-010-1125-5.

[24] M. Junge, C. Palazuelos, D. Pérez-García, I. Villanueva, and M. M. Wolf. Operator space theory: A natural framework for Bell inequalities. *Physical Review Letters*, 104: 170405, 2010. DOI: 10.1103/PhysRevLett.104.170405.

[25] M. Junge, T. Oikhberg, and C. Palazuelos. Reducing the number of questions in nonlocal games. *Journal of Mathematical Physics*, 57(10):102203, 2016. DOI: 10.1063/1.4965831.

[26] D. Kaszlikowski, P. Gnaciński, M. Żukowski, W. Miklaszewski, and A. Zeilinger. Violations of local realism by two entangled N-dimensional systems are stronger than for two qubits. *Physical Review Letters*, 85:4418–4421, 2000. DOI: 10.1103/PhysRevLett.85.4418.

[27] I. Kerenidis, S. Laplante, V. Lerays, J. Roland, and D. Xiao. Lower bounds on information complexity via zero-communication protocols and applications. *SIAM Journal on Computing*, 44(5):1550–1572, 2015. DOI: 10.1137/130928273.

[28] B. Klartag and O. Regev. Quantum one-way communication can be exponentially stronger than classical communication. In *Proc. 43rd STOC*, pages 31–40, 2011. ISBN 978-1-4503-0691-1. DOI: 10.1145/1993636.1993642.

[29] G. Kol, S. Moran, A. Shpilka, and A. Yehudayoff. Approximate nonnegative rank is equivalent to the smooth rectangle bound. In *Automata, Languages, and Programming*, pages 701–712. Springer, 2014. DOI: 10.1007/978-3-662-43948-7_58.

[30] E. Kushilevitz and N. Nisan. *Communication complexity.* Cambridge University Press, 1997. ISBN 978-0-521-56067-2. DOI: 10.1017/CBO9780511574948.

[31] S. Laplante, V. Lerays, and J. Roland. Classical and quantum partition bound and detector inefficiency. In *Proc. 39th ICALP*, pages 617–628, 2012. DOI: 10.1007/978-3-642-31594-7_52.

[32] S. Laplante, M. Laurière, A. Nolin, J. Roland, and G. Senno. Robust Bell Inequalities from Communication Complexity. In Anne Broadbent, editor, *11th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2016)*, volume 61 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 5:1–

5:24, Dagstuhl, Germany, 2016. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. ISBN 978-3-95977-019-4. DOI: 10.4230/LIPIcs.TQC.2016.5. URL http://drops.dagstuhl.de/opus/volltexte/2016/6686.

[33] W. Laskowski, T. Paterek, M. Żukowski, and Č Brukner. Tight multipartite Bell's inequalities involving many measurement settings. *Physical Review Letters*, 93(20): 200401, 2004. DOI: 10.1103/PhysRevLett.93.200401.

[34] N. Linial and A. Shraibman. Lower bounds in communication complexity based on factorization norms. *Random Structures & Algorithms*, 34(3):368–394, 2009. ISSN 1098-2418. DOI: 10.1002/rsa.20232.

[35] S. Massar. Nonlocality, closing the detection loophole, and communication complexity. *Physical Review A*, 65:032121, 2002. DOI: 10.1103/PhysRevA.65.032121.

[36] S. Massar, S. Pironio, J. Roland, and B. Gisin. Bell inequalities resistant to detector inefficiency. *Physical Review A*, 66:052112, 2002. DOI: 10.1103/PhysRevA.66.052112.

[37] T. Maudlin. Bell's inequality, information transmission, and prism models. In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, pages 404–417. JSTOR, 1992. DOI: 10.1086/psaprocbienmeetp.1992.1.192771.

[38] K. Maxwell and E. Chitambar. Bell inequalities with communication assistance. *Physical Review A*, 89(4):042108, 2014. DOI: 10.1103/PhysRevA.89.042108.

[39] D. N. Mermin. Extreme quantum entanglement in a superposition of macroscopically distinct states. *Physical Review Letters*, 65(15):1838, 1990. DOI: 10.1103/PhysRevLett.65.1838.

[40] K. Nagata, W. Laskowski, and T. Paterek. Bell inequality with an arbitrary number of settings and its applications. *Physical Review A*, 74(6):062109, 2006. DOI: 10.1103/PhysRevA.74.062109.

[41] M. A Nielsen and I. Chuang. *Quantum computation and quantum information.* AAPT, 2002. DOI: 10.1017/CBO9780511976667.

[42] C. Palazuelos and Z. Yin. Large bipartite Bell violations with dichotomic measurements. *Physical Review A*, 92:052313, 2015. DOI: 10.1103/PhysRevA.92.052313.

[43] D. Pérez-García, M. M. Wolf, C. Palazuelos, I. Villanueva, and M. Junge. Unbounded violation of tripartite Bell inequalities. *Communications in Mathematical Physics*, 279 (2):455–486, 2008. DOI: 10.1007/s00220-008-0418-4.

[44] S. Pironio. Violations of Bell inequalities as lower bounds on the communication cost of nonlocal correlations. *Physical Review A*, 68(6):062102, 2003. DOI: 10.1103/PhysRevA.68.062102.

[45] R. Raz. Exponential separation of quantum and classical communication complexity. In *Proc. 31st STOC*, pages 358–367, 1999. ISBN 1-58113-067-8. DOI: 10.1145/301250.301343.

[46] A. A. Razborov. On the distributional complexity of disjointness. *Theoretical Computer Science*, 106(2):385 – 390, 1992. ISSN 0304-3975. DOI: 10.1016/0304-3975(92)90260-M.

[47] A. A. Razborov. Quantum communication complexity of symmetric predicates. *Izvestiya: Mathematics*, 67(1):145, 2003. DOI: 10.1070/IM2003v067n01ABEH000422.

[48] A. A. Sherstov. The communication complexity of gap Hamming distance. *Theory of Computing*, 8(1):197–208, 2012. DOI: 10.4086/toc.2012.v008a008.

[49] M. Steiner. Towards quantifying non-local information transfer: finite-bit non-locality. *Physics Letters A*, 270(5):239–244, 2000. DOI: 10.1016/S0375-9601(00)00315-7.

[50] B. F. Toner and D. Bacon. Communication cost of simulating Bell correlations. *Physical Review Letters*, 91(18):187904, 2003. DOI: 10.1103/PhysRevLett.91.187904.

[51] B. S. Tsirel'son. Quantum analogues of the Bell inequalities. The case of two spatially separated domains. *Journal of Soviet Mathematics*, 36(4):557–570, 1987. DOI: 10.1007/BF01663472.

[52] A. C. C. Yao. Some complexity questions related to distributed computing. In *Proc. 11th STOC*, pages 209–213, 1979. DOI: 10.1145/800135.804414.

[53] A. C. C. Yao. Lower bounds by probabilistic arguments. In *Proc. 24th FOCS*, pages 420–428. IEEE, 1983. DOI: 10.1109/SFCS.1983.30.

[54] A. C. C. Yao. Quantum circuit complexity. In *Proc. 34th FOCS*, pages 352–361. IEEE, 1993. DOI: 10.1109/SFCS.1993.366852.

A Proof of Theorem 1

Observation 1. *Let B be a nonconstant normalized Bell functional and $\mathbf{p} \in \mathcal{C}$ such that $B(\mathbf{p}) \geq 1$. Consider $\boldsymbol{\ell}^- \in \mathcal{L}_{\det}^\perp$ such that $B(\boldsymbol{\ell}^-) = m = \min\{B(\boldsymbol{\ell})|\boldsymbol{\ell} \in \mathcal{L}_{\det}^\perp\}$ and $\boldsymbol{\ell}^+ \in \mathcal{L}_{\det}^\perp$ such that $B(\boldsymbol{\ell}^+) = M = \max\{B(\boldsymbol{\ell})|\boldsymbol{\ell} \in \mathcal{L}_{\det}^\perp\}$. We have $m < M$ because B is nonconstant. The Bell functional $\tilde{B}$ defined by $\tilde{B}(\cdot) = \frac{1}{M-m}(2B(\cdot) - M - m)$, is such that $\tilde{B}(\boldsymbol{\ell}^+) = 1$, $\tilde{B}(\boldsymbol{\ell}^-) = -1$, $|\tilde{B}(\boldsymbol{\ell})| \leq 1$ for all $\boldsymbol{\ell} \in \mathcal{L}_{\det}^\perp$, and $\tilde{B}(\mathbf{p}) \geq B(\mathbf{p})$ since B is normalized.*

Definition 3 below is the first step of the construction. It takes two marginal distributions $\mathbf{m}_A$ and $\mathbf{m}_B$, and a normalized Bell functional B, and constructs a Bell functional $B_{\mathbf{m}_A,\mathbf{m}_B}^\perp$ whose value over every distribution $\mathbf{p} \in \mathcal{C}^\perp$ coincides with the value of B over the distribution $\mathbf{p}' \in \mathcal{C}$ obtained from $\mathbf{p}$ by replacing the abort events with samples from $\mathbf{m}_A$ and $\mathbf{m}_B$.

Definition 3. *For all two families of distributions, $\mathbf{m}_A = (m_A(\cdot|x))_{x \in \mathcal{X}}$ over outcomes in $\mathcal{A}$ for Alice and $\mathbf{m}_B = (m_B(\cdot|y))_{y \in \mathcal{Y}}$ over outcomes in $\mathcal{B}$ for Bob, and any normalized Bell functional B with coefficients only on nonaborting events, we define the Bell functional $B_{\mathbf{m}_A,\mathbf{m}_B}^\perp$ on $(\mathcal{A} \cup \{\perp\}) \times (\mathcal{B} \cup \{\perp\}) \times \mathcal{X} \times \mathcal{Y}$ by*

$$(B_{\mathbf{m}_A,\mathbf{m}_B}^\perp)_{a,b,x,y} = B_{a,b,x,y} + \chi_{\{\perp\}}(a) \sum_{a' \neq \perp} m_A(a'|x) B_{a',b,x,y}$$

$$+ \chi_{\{\perp\}}(b) \sum_{b' \neq \perp} m_B(b'|y) B_{a,b',x,y} + \chi_{\{\perp\}}(a)\chi_{\{\perp\}}(b) \sum_{a',b' \neq \perp} m_A(a'|x)m_B(b'|y) B_{a',b',x,y}$$

where $\chi_{\mathcal{S}}$ is the indicator function for set $\mathcal{S}$ taking value 1 on $\mathcal{S}$ and 0 everywhere else.

Observation 2. *Let $f_{\mathbf{m}_A,\mathbf{m}_B} : \mathcal{C}^\perp \to \mathcal{C}$ be the function that replaces abort events on Alice's (resp. Bob's) side by a sample from $\mathbf{m}_A$ (resp. $\mathbf{m}_B$) (note that $f_{\mathbf{m}_A,\mathbf{m}_B}$ preserves locality). Then, for every $\mathbf{m}_A$, $\mathbf{m}_B$ and B as in Definition 3, the Bell functional $B_{\mathbf{m}_A,\mathbf{m}_B}^\perp$ satisfies that $B_{\mathbf{m}_A,\mathbf{m}_B}^\perp(\mathbf{p}) = B(f_{\mathbf{m}_A,\mathbf{m}_B}(\mathbf{p}))$, $\forall \mathbf{p} \in \mathcal{C}^\perp$, so $B_{\mathbf{m}_A,\mathbf{m}_B}^\perp(\mathbf{p}) = B(\mathbf{p})$, for all $\mathbf{p} \in \mathcal{C}$, and $|B_{\mathbf{m}_A,\mathbf{m}_B}^\perp(\boldsymbol{\ell})| \leq 1$, for all $\boldsymbol{\ell} \in \mathcal{L}^\perp$.*

Next, in Lemma 1 below, we do without the abort coefficients in the Bell functionals $B_{\mathbf{m}_A,\mathbf{m}_B}^\perp$.

Lemma 1. *Let B' be a normalized Bell functional on $\mathcal{A}^\perp \times \mathcal{B}^\perp \times \mathcal{X} \times \mathcal{Y}$ (possibly with nonzero weights on $\perp$). Then the Bell functional B'' on the same set defined by*

$$B''_{a,b,x,y} = B'_{a,b,x,y} - B'_{a,\perp,x,y} - B'_{\perp,b,x,y} + B'_{\perp,\perp,x,y}, \tag{6}$$

for all $(a,b,x,y) \in (\mathcal{A} \cup \{\perp\}) \times (\mathcal{B} \cup \{\perp\}) \times \mathcal{X} \times \mathcal{Y}$
satisfies :

1. *If $a = \bot$ or $b = \bot$ then $B''_{a,b,x,y} = 0$*

2. *for all $\mathbf{p} \in \mathcal{C}$,*

$$B''(\mathbf{p}) = B'(\mathbf{p}) - B'(\mathbf{p}_{A,\bot}) - B'(\mathbf{p}_{\bot,B}) + B'(\mathbf{p}_{\bot,\bot}), \qquad (7)$$

where $\mathbf{p}_{A,\bot} \in \mathcal{L}^{\bot}$ (resp. $\mathbf{p}_{\bot,B} \in \mathcal{L}^{\bot}$) is the local distribution obtained from $\mathbf{p}$ if Bob (resp. Alice) replaces all of his (resp. her) outputs by $\bot$, and $\mathbf{p}_{\bot,\bot} \in \mathcal{L}^{\bot}$ is the local distribution where both Alice and Bob always output $\bot$. In Item 2 above, for all $\mathbf{p}'$,

$$B'(\mathbf{p}') = \sum_{(a,b) \in \mathcal{A}^{\bot} \times \mathcal{B}^{\bot}} \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} B'_{a,b,x,y} \mathbf{p}'(a,b|x,y)$$

where the first sum is also over the abort events.

Proof. Item 1 follows from (6). We prove Item 2. For $\mathbf{p} \in \mathcal{C}^{\bot}$ with marginals $\mathbf{p}_A$ and $\mathbf{p}_B$, we have: for all $y \in Y$, $p_A(a|x) = \sum_{b \in \mathcal{B}^{\bot}} p(a,b|x,y)$, and for all $x \in X$, $p_B(b|y) = \sum_{a \in \mathcal{A}^{\bot}} p(a,b|x,y)$. For the remainder of this proof, summations involving a (resp. b) are over $a \in \mathcal{A}^{\bot}$ (resp. $b \in \mathcal{B}^{\bot}$).

By definition, $p_{A,\bot}(a,b|x,y) = p_A(a|x)\chi_{\{\bot\}}(b)$, $p_{\bot,B}(a,b|x,y) = \chi_{\{\bot\}}(a)p_B(b|y)$, and $p_{\bot,\bot}(a,b|x,y) = \chi_{\{\bot\}}(a)\chi_{\{\bot\}}(b)$. We have:

$$
\begin{aligned}
B''(\mathbf{p}) &= \sum_{a,b,x,y} \left[B'_{a,b,x,y} - B'_{a,\bot,x,y} - B'_{\bot,b,x,y} + B'_{\bot,\bot,x,y} \right] p(a,b|x,y) \\
&= \sum_{a,b,x,y} B'_{a,b,x,y} p(a,b|x,y) - \sum_{a,x,y} B'_{a,\bot,x,y} \sum_b p(a,b|x,y) \\
&\quad - \sum_{b,x,y} B'_{\bot,b,x,y} \sum_a p(a,b|x,y) + \sum_{x,y} B'_{\bot,\bot,x,y} \sum_{a,b} p(a,b|x,y) \\
&= B'(\mathbf{p}) - \sum_{a,x,y} B'_{a,\bot,x,y} p_A(a|x) - \sum_{b,x,y} B'_{\bot,b,x,y} p_B(b|y) + \sum_{x,y} B'_{\bot,\bot,x,y} \\
&= B'(\mathbf{p}) - B'(\mathbf{p}_{A,\bot}) - B'(\mathbf{p}_{\bot,B}) + B'(\mathbf{p}_{\bot,\bot}). \qquad \square
\end{aligned}
$$

We are now ready to prove Theorem 1.

Proof of Theorem 1. If B is constant, since it is normalized by assumption, we have $B \equiv 1$. Thus, we can simply take B^* defined by: for all $(x,y) \in \mathcal{X} \times \mathcal{Y}$, $B^*_{a,b,x,y} = B_{a,b,x,y}$ if $(a,b) \in \mathcal{A} \times \mathcal{B}$, and $B^*_{a,b,x,y} = 0$ otherwise.

Now, let us assume that B is not constant and let $\boldsymbol{\ell}^-, \boldsymbol{\ell}^+ \in \mathcal{L}_{\det}$, and $\tilde{B}$ constructed from B as in Observation 1 satisfying $\tilde{B}(\boldsymbol{\ell}^-) = -1$ and $\tilde{B}(\boldsymbol{\ell}^+) = 1$. Since $\boldsymbol{\ell}^-$ and $\boldsymbol{\ell}^+$ are deterministic distributions, we have: $\boldsymbol{\ell}^- = \boldsymbol{\ell}_A^- \otimes \boldsymbol{\ell}_B^-$ and $\boldsymbol{\ell}^+ = \boldsymbol{\ell}_A^+ \otimes \boldsymbol{\ell}_B^+$, for some marginals $\boldsymbol{\ell}_A^-, \boldsymbol{\ell}_B^-, \boldsymbol{\ell}_A^+$, and $\boldsymbol{\ell}_B^+$. We consider the replacing Bell functional $B^{\bot}_{\boldsymbol{\ell}_A^-, \boldsymbol{\ell}_B^-}$ (resp. $B^{\bot}_{\boldsymbol{\ell}_A^+, \boldsymbol{\ell}_B^+}$) from Definition 3 constructed from $(\tilde{B}, \boldsymbol{\ell}_A^-, \boldsymbol{\ell}_B^-)$ (resp. from $(\tilde{B}, \boldsymbol{\ell}_A^+, \boldsymbol{\ell}_B^+)$). Taking $B' = \frac{1}{2}(B^{\bot}_{\boldsymbol{\ell}_A^-, \boldsymbol{\ell}_B^-} + B^{\bot}_{\boldsymbol{\ell}_A^+, \boldsymbol{\ell}_B^+})$, we have $|B'(\boldsymbol{\ell})| \leq 1$, for all $\boldsymbol{\ell} \in \mathcal{L}^{\bot}$, and therefore we can apply Lemma 1 to get B'' from B'. Since $B'(\mathbf{p}_{\bot,\bot}) = \frac{1}{2}(B^{\bot}_{\boldsymbol{\ell}_A^-, \boldsymbol{\ell}_B^-}(\mathbf{p}_{\bot,\bot}) + B^{\bot}_{\boldsymbol{\ell}_A^+, \boldsymbol{\ell}_B^+}(\mathbf{p}_{\bot,\bot})) = \frac{1}{2}(\tilde{B}(\boldsymbol{\ell}^-) + \tilde{B}(\boldsymbol{\ell}^+)) = 0$, by (7) we have for all $\mathbf{p} \in \mathcal{C}^{\bot}$, $B''(\mathbf{p}) = B'(\mathbf{p}) - B'(\mathbf{p}_{A,\bot}) - B'(\mathbf{p}_{\bot,B})$. Hence, denoting $B^* = \frac{1}{3}B''$, B^* satisfies all the required properties since $|B'(\boldsymbol{\ell})| \leq 1$ for all $\boldsymbol{\ell} \in \mathcal{L}^{\bot}$ and therefore we have for all $\mathbf{p} \in \mathcal{C}$, $B^*(\mathbf{p}) \geq \frac{1}{3}B'(\mathbf{p}) - \frac{1}{3}|B'(\mathbf{p}_{A,\bot})| - \frac{1}{3}|B'(\mathbf{p}_{\bot,B})| \geq \frac{1}{3}B'(\mathbf{p}) - \frac{2}{3}$, and for all $\boldsymbol{\ell} \in \mathcal{L}^{\bot}$, $|B^*(\boldsymbol{\ell})| \leq \frac{1}{3}|B'(\boldsymbol{\ell})| + \frac{1}{3}|B'(\boldsymbol{\ell}_{A,\bot})| + \frac{1}{3}|B'(\boldsymbol{\ell}_{\bot,B})| \leq 1$. $\square$

B   Proof of Theorem 4

Proof. Let (B, β) be an optimal solution to $\mathbf{eff}_\epsilon(\mathbf{p})$ and let c be such that $\mathbf{eff}_\epsilon(\mathbf{p}) = \beta \geq 2^c$. Recall that, by optimality of B, we have

$$B(\mathbf{p}') \geq 2^c \text{ for any } \mathbf{p}' \text{ such that } |\mathbf{p}' - \mathbf{p}|_1 \leq \epsilon. \tag{8}$$

As in the proof of Corollary 1, we go from a q-qubit quantum protocol for a distribution $\mathbf{p}'$ with $|\mathbf{p}' - \mathbf{p}|_1 \leq \epsilon'$ to a $2q$-bit entanglement-assisted protocol for $\mathbf{p}'$ by replacing quantum communication with teleportation. Let $\mathcal{H}_{\mathcal{A}'}$ (resp. $\mathcal{H}_{\mathcal{B}'}$) be Alice's (resp. Bob's) local Hilbert spaces in this protocol and, without loss of generality, let the initial state of the protocol be

$$|\psi\rangle := |\beta_{00}\rangle^{\otimes q}|\phi\rangle$$

with $|\beta_{00}\rangle := 1/\sqrt{2}(|00\rangle + |11\rangle)$ and some fixed $|\phi\rangle \in \mathcal{H}_{\mathcal{A}'} \otimes \mathcal{H}_{\mathcal{B}'}$. Next, we go to a zero-communication protocol Π by instructing the players not to communicate the teleportation measurements' outcomes (and hence not to perform the correcting unitaries) and to output a new symbol $S \notin \mathcal{A} \cup \mathcal{B} \cup \{\perp\}$ whenever they get a teleportation measurement's outcome different from 00 (the outcome corresponding to the case in which no correcting unitary is required); notice that this happens with probability $1 - 2^{-2q}$. The local unitaries and local measurements performed by the players during the execution of Π effectively induce POVMs $\{E_{a|x}\}_{a \in \mathcal{A} \cup \{S\}}$ and $\{E_{b|y}\}_{b \in \mathcal{B} \cup \{S\}}$ over $\mathcal{H}_{\mathcal{A}} := \mathcal{H}_{\mathcal{A}'} \otimes \mathbb{C}^{2^q}$ and $\mathcal{H}_{\mathcal{B}} := \mathbb{C}^{2^q} \otimes \mathcal{H}_{\mathcal{B}'}$ respectively and a corresponding quantum distribution $\mathbf{q}$ such that the probability of the players outputting (a, b) on inputs (x, y) is given by $q(a, b|x, y) = \text{tr}\left[(E_{a|x} \otimes E_{b|y})|\psi\rangle\langle\psi|\right]$. Notice that

$$q(a \neq S, b \neq S|x, y) = \text{tr}\left[(\mathbb{I} - E_{S|x}) \otimes (\mathbb{I} - E_{S|y})|\psi\rangle\langle\psi|\right] = 2^{-2q}, \tag{9}$$

which together with the fact that we started from a quantum protocol for $\mathbf{p}'$, implies that $\mathbf{q}$ is of the form

$$q(a, b|x, y) = \frac{1}{2^{2q}}p'(a, b|x, y) + (1 - \frac{1}{2^{2q}})r_1(a, b|x, y),$$

with $\mathbf{r_1}$ supported on events in which at least one of the players output S. Therefore, letting $\overline{B}$ be the Bell functional such that $\overline{B}_{a,b,x,y} = B_{a,b,x,y}$ if $a \neq S$ and $b \neq S$, and $\overline{B}_{a,b,x,y} = 0$ otherwise, we have that

$$\overline{B}(\mathbf{q}) = \frac{1}{2^{2q}}B(\mathbf{p}') \geq 2^{c-2q}.$$

Next, suppose that we run protocol Π with the initial state

$$\rho := (1 - \delta)|\psi\rangle\langle\psi| + \delta\left(\frac{\mathbb{I}}{2^{2q}} \otimes \frac{\mathbb{I}}{\dim(\mathcal{H}_{\mathcal{A}}')\dim(\mathcal{H}_{\mathcal{B}}')}\right),$$

and let $\mathbf{q}_\delta$ be the resulting quantum distribution. Notice that the $2q$ teleportation measurements will still give uniform outcomes when performed with $\mathbb{I}/4$ instead of $|\beta_{00}\rangle$, i.e. (9) still holds when we replace $|\psi\rangle$ with ρ, and hence we have that

$$
\begin{aligned}
q_\delta(a, b|x, y) &= \text{tr}\left[(E_{a|x} \otimes E_{b|y})\rho\right] \\
&= (1 - \delta)\left[\frac{1}{2^{2q}}p'(a, b|x, y) + (1 - \frac{1}{2^{2q}})r_1(a, b|x, y)\right] + \\
&\qquad \delta\left[\frac{1}{2^{2q}}n(a, b|x, y) + (1 - \frac{1}{2^{2q}})r_2(a, b|x, y)\right] \\
&= \frac{1}{2^{2q}}\left[(1 - \delta)p'(a, b|x, y) + \delta n(a, b|x, y)\right] + (1 - \frac{1}{2^{2q}})\left[(1 - \delta)r_1(a, b|x, y) + \delta r_2(a, b|x, y)\right]
\end{aligned}
$$

for distributions $\mathbf{r_2}$ supported on events with at least one of the players outputting S and $\mathbf{n}$ on events in which none of the players output S.

Finally, leting $\mathbf{p}'' := (1 - \delta)\mathbf{p}' + \delta\mathbf{n}$, and noticing that

$$
\begin{aligned}
|\mathbf{p} - \mathbf{p}''|_1 &= |(1 - \delta)\mathbf{p}' + \delta\mathbf{n}|_1 \\
&\le \epsilon' + \delta|\mathbf{p}' - \mathbf{n}|_1 \\
&\le \epsilon' + \delta \le \epsilon,
\end{aligned}
$$

with the last inequality following from $\delta \le \epsilon - \epsilon'$, we get from (8) that,

$$
\overline{B}(\mathbf{q}_\delta) = \frac{1}{2^{2q}}B(\mathbf{p}'') \ge 2^{c-2q}.
$$

$\square$

C Proof of Theorem 6

Proof. Let us first set $B_{z,x,y} = B_{a,b,x,y}$ for all $a \oplus b = z$. Let $\boldsymbol{\ell} \in \mathcal{L}_{det}^\perp$. Then, we have:

$$
B(\boldsymbol{\ell}) = \sum_{(x,y)\in R} B_{z,x,y} + \sum_{(x,y)\in S} B_{z,x,y}
$$

where R and S are the two rectangles where $\boldsymbol{\ell}$ outputs z. Let us take a rectangle R. Then :

$$
\sum_{(x,y)\in R} B_{z,x,y} = \frac{1}{2 \cdot g(n)}\left(\sum_j v_j\mu(V_j \cap R) - \sum_i u_i\mu(U_i \cap R)\right) \le 1/2
$$

with the inequality following from (1). This proves (3).

Let us now compute $B(\mathbf{p}_f)$. By linearity of B and the definition of its coefficients, we have:

$$
\begin{aligned}
B(\mathbf{p}_f) &= \sum_{a,b,x,y} B_{a,b,x,y}\mathbf{p}_f(a,b|x,y) \\
&= \frac{1}{2}\sum_{(x,y)\in f^{-1}(z),a,b} B_{a,b,x,y}\chi_{\{z\}}(a \oplus b) + \frac{1}{2}\sum_{(x,y)\in f^{-1}(\bar{z}),a,b} B_{a,b,x,y}\chi_{\{\bar{z}\}}(a \oplus b) \\
&= 1/2\sum_j \sum_{(x,y)\in V_j} v_j g(n)^{-1}\mu(x,y) \\
&= \frac{1}{2 \cdot g(n)}\sum_j v_j\mu(V_j)
\end{aligned}
$$

(for the third equality we used the fact that $B_{a,b,x,y} = 0$ when $a \oplus b = \bar{z}$). This proves (4).

Moreover, for any family of additive error terms $\Delta(a,b|x,y) \in [-1,1]$ such that

$$
\sum_{a,b} |\Delta(a,b|x,y)| \le \epsilon \qquad\qquad \forall x,y \in \mathcal{X} \times \mathcal{Y},
$$

denoted collectively as $\boldsymbol{\Delta}$, we have

$$
\begin{aligned}
|B(\boldsymbol{\Delta})| &= \left| \sum_{a,b,x,y} B_{a,b,x,y} \Delta(a,b|x,y) \right| \\
&= \frac{1}{2 \cdot g(n)} \left| \sum_{a,b \,:\, a \oplus b = z} \left[\sum_i \sum_{(x,y) \in U_i} (-u_i) \mu(x,y) \Delta(a,b|x,y) + \sum_j \sum_{(x,y) \in V_j} v_j \mu(x,y) \Delta(a,b|x,y) \right] \right| \\
&\leq \frac{1}{2 \cdot g(n)} \left[\sum_i \sum_{(x,y) \in U_i} |u_i| \mu(x,y) \left(\sum_{a,b} |\Delta(a,b|x,y)| \right) + \sum_j \sum_{(x,y) \in V_j} |v_j| \mu(x,y) \left(\sum_{a,b} |\Delta(a,b|x,y)| \right) \right] \\
&\leq \frac{\epsilon}{2 \cdot g(n)} \left[\sum_i |u_i| \mu(U_i) + \sum_j |v_j| \mu(V_j) \right]
\end{aligned}
$$

From this calculation and (4), we obtain, for $\mathbf{p}' = \mathbf{p}_f + \boldsymbol{\Delta}$:

$$
B(\mathbf{p}') = B(\mathbf{p}_f) + B(\boldsymbol{\Delta}) \geq \frac{1}{2 \cdot g(n)} \left[\sum_j v_j \mu(V_j) - \epsilon \left(\sum_j |v_j| \mu(V_j) + \sum_i |u_i| \mu(U_i) \right) \right],
$$

which proves (5).

$\square$

D   Explicit examples

Let us formulate a special case of Theorem 6 that will be useful in the examples. Here there is just one subset in $f^{-1}(0)$ and one in $f^{-1}(1)$.

Corollary 2. *Let f be a (possibly partial) Boolean function on $\mathcal{X} \times \mathcal{Y}$, where $\mathcal{X}, \mathcal{Y} \subseteq \{0,1\}^n$. Given $\gamma \in (0,1)$ and $g : \mathbb{N} \to (0,1)$, suppose that there is a distribution μ on $\mathcal{X} \times \mathcal{Y}$ such that: for any rectangle $R \subseteq \mathcal{X} \times \mathcal{Y}$,*

$$
\mu(R \cap f^{-1}(1)) > \gamma \mu(R \cap f^{-1}(0)) - g(n). \tag{10}
$$

Then μ satisfies (1) with $z = 0$, $i = j = 1$, $U_1 = f^{-1}(1)$, $V_1 = f^{-1}(0)$, $u_1 = 1$, $v_1 = \gamma$. Let B be defined by (2), that is: for all $a,b,x,y \in \{0,1\} \times \{0,1\} \times \mathcal{X} \times \mathcal{Y}$,

$$
B_{a,b,x,y} = \begin{cases} -\frac{1}{2 \cdot g(n)} \mu(x,y) & \text{if } f(x,y) = 1 \text{ and } a \oplus b = 0 \\ \frac{\gamma}{2 \cdot g(n)} \mu(x,y) & \text{if } f(x,y) = 0 \text{ and } a \oplus b = 0 \\ 0 & \text{otherwise.} \end{cases}
$$

Then, B satisfies

$$
\begin{aligned}
B(\boldsymbol{\ell}) &\leq 1, \quad \forall \boldsymbol{\ell} \in \mathcal{L}_{det}^{\perp}, \\
B(\mathbf{p}_f) &= \frac{\gamma}{2 \cdot g(n)} \mu(f^{-1}(0))
\end{aligned}
$$

and for any $\mathbf{p}' \in \mathcal{P}$ such that $|\mathbf{p}' - \mathbf{p}_f|_1 \leq \epsilon$:

$$
B(\mathbf{p}') \geq \frac{1}{2 \cdot g(n)} \left[\gamma \mu(f^{-1}(0)) - \epsilon(\gamma \mu(f^{-1}(0)) + \mu(f^{-1}(1))) \right].
$$

D.1 Disjointness

In [46], Razborov proved the following.

Lemma 2 ([46])**.** *There exist two distributions μ_0 and μ_1 with $\mathrm{supp}(\mu_0) \subseteq \mathrm{DISJ}_n^{-1}(1)$ and $\mathrm{supp}(\mu_1) \subseteq \mathrm{DISJ}_n^{-1}(0)$, such that: for any rectangle R in the input space,*

$$\mu_1(R) \geq \Omega(\mu_0(R)) - 2^{\Omega(n)}.$$

Following his proof, one can check that we actually have:

$$\mu_1(R) \geq \frac{1}{45}\mu_0(R) - 2^{-\epsilon n + \log_2(2/9)}.$$

So, letting $\mu := (\mu_0 + \mu_1)/2$,

$$\mu(R \cap f^{-1}(0)) \geq \frac{1}{45}\mu(R \cap f^{-1}(1)) - 2^{-\epsilon n + \log_2(4/9)}. \tag{11}$$

Remark 1. *Actually,* $\mathrm{supp}(\mu_1) = A_1 := \{(x,y) \ : \ |x| = |y| = m, |x \cap y| = 1\} \subseteq \mathrm{DISJ}_n^{-1}(0)$.

Note that by this construction, $\mu(f^{-1}(0)) = \mu(f^{-1}(1)) = 1/2$. Combining (11) with Corollary 2 (with $g(n) = 2^{-\epsilon n + \log_2(4/9)}$), we obtain:

Corollary 3. *There exists a Bell inequality B satisfying:* $\forall \ell \in \mathcal{L}_{det}^{\perp}$, $B(\ell) \leq 1$,

$$B(\mathbf{p}_{\mathrm{DISJ}_n}) = \frac{1}{90}2^{\epsilon n - \log_2(4/9)},$$

and for any distribution $\mathbf{p}' \in \mathcal{P}$ such that $|\mathbf{p}' - \mathbf{p}_{\mathrm{DISJ}_n}|_1 \leq \varepsilon$,

$$B(\mathbf{p}') \geq 2^{\epsilon n - \log_2(4/9)}\frac{1 - 46\epsilon}{90}.$$

More precisely, Theorem 6 gives an explicit construction of such a Bell inequality: we can define B as:

$$B_{a,b,x,y} = \begin{cases} -2^{\epsilon n - \log_2(4/9)}\mu(x,y) & \text{if } \mathrm{DISJ}_n(x,y) = 0 \text{ and } a \oplus b = 1 \\ \frac{1}{45}2^{\epsilon n - \log_2(4/9)}\mu(x,y) & \text{if } \mathrm{DISJ}_n(x,y) = 1 \text{ and } a \oplus b = 1 \\ 0 & \text{otherwise.} \end{cases}$$

To obtain Proposition 5, we use Corollary 1 together with the fact that $Q_{\varepsilon'}(\mathrm{DISJ}_n) = O(\sqrt{n})$.

D.2 Tribes

Let $n = (2r+1)^2$ with $r \geq 2$ and let $\mathrm{TRIBES}_n : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ be defined as:

$$\mathrm{TRIBES}_n(x,y) := \bigwedge_{i=1}^{\sqrt{n}} \left(\bigvee_{j=1}^{\sqrt{n}} (x_{(i-1)\sqrt{n}+j} \text{ and } y_{(i-1)\sqrt{n}+j}) \right).$$

In [19][Sec. 3] the following is proven:

Lemma 3. *There exists a probability distribution μ on $\{0,1\}^n \times \{0,1\}^n$ for which there exist numbers $\alpha, \lambda, \gamma, \delta > 0$ such that for sufficiently large n and for any rectangle R in the input space:*

$$\gamma\mu(U_1 \cap R) \geq \alpha\mu(V_1 \cap R) - \lambda\mu(V_2 \cap R) - 2^{-\delta n/2+1}$$

where $U_1 = \text{TRIBES}_n^{-1}(0)$, $\{V_1, V_2\}$ forms a partition of $\text{TRIBES}_n^{-1}(1)$ and $\mu(U_1) = 1 - 7\beta^2/16$, $\mu(V_1) = 6\beta^2/16$, $\mu(V_2) = \beta^2/16$ with $\beta = \frac{r+2}{r+1}$.

In [19], the coefficients are $\alpha = 0.99, \lambda = \frac{16}{3(0.99)^2}$ and $\gamma = \frac{16}{(0.99)^2}$ (the authors say these values have not been optimized).

Combining this result with our Theorem 6 (taking $z = 1, i = 1, j = 2, U_1, V_1, V_2$ as in Lemma 3, $u_1 = \gamma, v_1 = \alpha, v_2 = -\lambda$, and $g(n) = 2^{-\delta n/2+1}$), we obtain:

Corollary 4. *There exists a Bell inequality satisfying: $\forall \boldsymbol{\ell} \in \mathcal{L}_{det}^{\perp}$, $B(\boldsymbol{\ell}) \leq 1$,*

$$B(\mathbf{p}_{\text{TRIBES}_n}) = 2^{\delta n/2-1}\frac{\beta^2}{16}(6\alpha - \lambda),$$

and for any distribution $\mathbf{p}' \in \mathcal{P}$ such that $|\mathbf{p}' - \mathbf{p}_{\text{TRIBES}_n}|_1 \leq \varepsilon$,

$$B(\mathbf{p}') \geq 2^{\delta n/2-1}\left[\frac{\beta^2}{16}(6\alpha - \lambda) - \epsilon(\gamma(1 - 7\beta^2/16) + \lambda\beta^2/16 + \alpha 6\beta^2/16)\right].$$

More precisely, Theorem 6 provides a Bell inequality B yielding this bound, defined as:

$$B_{a,b,x,y} = \begin{cases} -\gamma 2^{\delta n/2-1}\mu(x,y) & \text{if } (x,y) \in U_1 \text{ and } a \oplus b = 1 \\ \alpha 2^{\delta n/2-1}\mu(x,y) & \text{if } (x,y) \in V_1 \text{ and } a \oplus b = 1 \\ -\lambda 2^{\delta n/2-1}\mu(x,y) & \text{if } (x,y) \in V_2 \text{ and } a \oplus b = 1 \\ 0 & \text{otherwise.} \end{cases}$$

To obtain Proposition 6, we use Corollary 1 together with the fact that $Q_{\varepsilon'}(\text{TRIBES}_n) = O(\sqrt{n}(\log n)^2)$.

D.3 Gap Orthogonality

Let f_n be the partial functions over $\{-1,+1\}^n \times \{-1,+1\}^n$ by $f_n(x,y) = \text{ORT}_{64n}(x^{64}, y^{64})$, that is:

$$f_n(x,y) = \begin{cases} -1 & \text{if } |\langle x,y\rangle| \leq \sqrt{n}/8 \\ +1 & \text{if } |\langle x,y\rangle| \geq \sqrt{n}/4. \end{cases}$$

In [48], Sherstov proves the following result.

Lemma 4 ([48]). *Let $\delta > 0$ be a sufficiently small constant and μ the uniform measure over $\{0,1\}^n \times \{0,1\}^n$. Then, $\mu(f_n^{-1}(+1)) = \Theta(1)$ and for all rectangle R in $\{0,1\}^n \times \{0,1\}^n$ such that $\mu(R) > 2^{-\delta n}$,*

$$\mu(R \cap f_n^{-1}(+1)) \geq \delta\mu(R \cap f_n^{-1}(-1)).$$

This implies that if we put uniform weight on inputs of ORT_{64n} of the form (x^{64}, y^{64}) and put 0 weight on the others, we get a distribution μ' satisfying the constraints of Corollary 2 for ORT_{64n} together with $\gamma = \delta$ from Lemma 4 and $g(64n) = 2^{\delta n}$.

To get a distribution satisfying the constraints of Corollary 2 on inputs of ORT_{64n+l} for all $0 \leq l \leq 63$ we extend μ' as follows:

$$
\tilde{\mu}(xu, yv) = \begin{cases} \mu'(x,y) & \text{if } u = +1^l, v = -1^l \text{ and } \left(\langle x,y \rangle < -\sqrt{64n} \text{ or } 0 \leq \langle x,y \rangle \leq \sqrt{64n} \right) \\ \mu'(x,y) & \text{if } u = +1^l, v = +1^l \text{ and } \left(-\sqrt{64n} \leq \langle x,y \rangle < 0 \text{ or } \langle x,y \rangle > \sqrt{64n} \right) \\ 0 & \text{otherwise} \end{cases}
$$

Using this distribution $\tilde{\mu}$ together with $\gamma = \delta$ from Lemma 4 and with $g(n) = 2^{-\delta n}$ we obtain, from Corollary 2, a Bell inequality violation for ORT_{64n+l} for all $0 \leq l \leq 63$:

Corollary 5. *There exists a Bell inequality B satisfying: $\forall \boldsymbol{\ell} \in \mathcal{L}_{det}^{\perp}$, $B(\boldsymbol{\ell}) \leq 1$,*

$$
B(\mathbf{p}_{\mathrm{ORT}_{64n+l}}) = 2^{\delta n} \delta \tilde{\mu}(\mathrm{ORT}_{64n+l}^{-1}(-1)),
$$

and for any distribution $\mathbf{p}' \in \mathcal{P}$ such that $|\mathbf{p}' - \mathbf{p}_{\mathrm{ORT}_{64n+l}}|_1 \leq \varepsilon$,

$$
B(\mathbf{p}') \geq 2^{\delta n} \left(\delta \tilde{\mu}(\mathrm{ORT}_{64n+l}^{-1}(-1)) - \epsilon [\delta \tilde{\mu}(\mathrm{ORT}_{64n+l}^{-1}(-1)) + \tilde{\mu}(\mathrm{ORT}_{64n+l}^{-1}(+1))] \right).
$$

More precisely, Theorem 6 gives an explicit construction of such a Bell inequality: we can define B as:

$$
B_{a,b,x,y} = \begin{cases} -2^{\delta n} \tilde{\mu}(x,y) & \text{if } (x,y) \in \mathrm{ORT}_{64n+l}^{-1}(+1) \text{ and } a \oplus b = -1 \\ \delta 2^{\delta n} \tilde{\mu}(x,y) & \text{if } (x,y) \in \mathrm{ORT}_{64n+l}^{-1}(-1) \text{ and } a \oplus b = -1 \\ 0 & \text{otherwise.} \end{cases}
$$

To obtain Proposition 7, we use Corollary 1 together with the fact that $Q_{\varepsilon'}(\mathrm{ORT}_n) = O(\sqrt{n} \log n)$.

E Equivalent formulations of the efficiency bounds

In [31], the zero-error efficiency bound was defined in its primal and dual forms as follows

Definition 4 ([31]). *The efficiency bound of a distribution $\mathbf{p} \in \mathcal{P}$ is given by*

$$
\begin{aligned}
\mathbf{eff}(\mathbf{p}) = \min_{\zeta, \mu_\ell \geq 0} \quad & \frac{1}{\zeta} \\
\text{subject to} \quad & \sum_{\ell \in \mathcal{L}_{det}^{\perp}} \mu_\ell \ell(a,b|x,y) = \zeta p(a,b|x,y) \quad \forall (a,b,x,y) \in \mathcal{A} \times \mathcal{B} \times \mathcal{X} \times \mathcal{Y} \\
& \sum_{\ell \in \mathcal{L}_{det}^{\perp}} \mu_\ell = 1 \\
= \max_{B} \quad & B(\mathbf{p}) \\
\text{subject to} \quad & B(\boldsymbol{\ell}) \leq 1 \quad \forall \boldsymbol{\ell} \in \mathcal{L}_{det}^{\perp}
\end{aligned}
$$

The ϵ-error efficiency bound was in turn defined as $\min_{\mathbf{p}' \in \mathcal{P} |\mathbf{p}' - \mathbf{p}|_1 \leq \epsilon} \mathbf{eff}(\mathbf{p}')$. In this appendix, we show that this is equivalent to the definition used in the present article (Definition 2). In the original definition, the Bell functional could depend on the particular $\mathbf{p}'$. We show that it is always possible to satisfy the constraint with the same Bell functional for all $\mathbf{p}'$ close to $\mathbf{p}$.

In order to prove this, we will need the following notions.

Definition 5. *A* distribution error $\boldsymbol{\Delta}$ *is a family of additive error terms* $\Delta(a,b|x,y) \in [-1,1]$ *for all* $(a,b,x,y) \in \mathcal{A} \times \mathcal{B} \times \mathcal{X} \times \mathcal{Y}$ *such that*

$$\sum_{a,b} \Delta(a,b|x,y) = 0 \qquad\qquad \forall (x,y) \in \mathcal{X} \times \mathcal{Y}.$$

For any $0 \le \epsilon \le 1$, *the set* Δ_ϵ *is the set of distribution errors* $\boldsymbol{\Delta}$ *such that*

$$\sum_{a,b} |\Delta(a,b|x,y)| \le \epsilon \qquad\qquad \forall (x,y) \in \mathcal{X} \times \mathcal{Y}.$$

This set is a polytope, so it admits a finite set of extremal points. We denote this set by Δ_ϵ^{ext}.

We will use the following properties of Δ_ε.

Fact 1. *For any distribution* $\mathbf{p} \in \mathcal{P}$, *we have*

$$\{\mathbf{p}' \in \mathcal{P} | \ |\mathbf{p}' - \mathbf{p}|_1 \le \epsilon\} \subseteq \{\mathbf{p} + \boldsymbol{\Delta} | \ \boldsymbol{\Delta} \in \Delta_\varepsilon\}$$

The reason why the set on the right-hand side might be larger is that $\mathbf{p} + \boldsymbol{\Delta}$ might not be a valid distribution. In order to ensure that this is the case, it is sufficient to impose that all obtained purposed probabilities are nonnegative, leading to the following property.

Fact 2. *For any distribution* $\mathbf{p} \in \mathcal{P}$, *we have*

$$\{\mathbf{p}' \in \mathcal{P} | \ |\mathbf{p}' - \mathbf{p}|_1 \le \epsilon\} = \{\mathbf{p} + \boldsymbol{\Delta} | \ \boldsymbol{\Delta} \in \Delta_\varepsilon \ \& \ p(a,b|x,y) + \Delta(a,b|x,y) \ge 0 \ \forall a,b,x,y\}$$

We are now ready to prove the following theorem.

Theorem 7. *Let* $\mathbf{p} \in \mathcal{P}$ *be a distribution,* $\mathbf{eff}_\varepsilon(\mathbf{p})$ *be defined as in Definition 2 and* $\mathbf{eff}(\mathbf{p})$ *be defined as in Definition 4. Then, we have*

$$\mathbf{eff}_\varepsilon(\mathbf{p}) = \min_{\mathbf{p}' \in \mathcal{P}: |\mathbf{p}' - \mathbf{p}|_1 \le \epsilon} \mathbf{eff}(\mathbf{p}').$$

Proof. Let $\overline{\mathbf{eff}}_\varepsilon(\mathbf{p}) = \min_{\mathbf{p}' \in \mathcal{P}: |\mathbf{p}' - \mathbf{p}|_1 \le \epsilon} \mathbf{eff}(\mathbf{p}')$. We first show that $\mathbf{eff}_\varepsilon(\mathbf{p}) \le \overline{\mathbf{eff}}_\varepsilon(\mathbf{p})$. Let (B, β) be an optimal feasible point for $\mathbf{eff}_\varepsilon(\mathbf{p})$, so that

$$\begin{aligned} \mathbf{eff}_\varepsilon(\mathbf{p}) &= \beta, \\ B(\mathbf{p}') &\ge \beta && \forall \mathbf{p}' \text{ s.t. } |\mathbf{p}' - \mathbf{p}|_1 \le \epsilon, \\ B(\boldsymbol{\ell}) &\le 1 && \forall \boldsymbol{\ell} \in \mathcal{L}_{det}^\perp. \end{aligned}$$

Therefore (B, β) is also a feasible point for $\mathbf{eff}(\mathbf{p}')$ for all $\mathbf{p}' \in \mathcal{P}$ such that $|\mathbf{p}' - \mathbf{p}|_1 \le \epsilon$, so that $\mathbf{eff}(\mathbf{p}') \ge \beta$ for all such $\mathbf{p}'$, and $\overline{\mathbf{eff}}_\varepsilon(\mathbf{p}) \ge \beta = \mathbf{eff}_\varepsilon(\mathbf{p})$.

It remains to show that $\mathbf{eff}_\varepsilon(\mathbf{p}) \ge \overline{\mathbf{eff}}_\varepsilon(\mathbf{p})$. In order to do so, we first use the primal

form of $\mathbf{eff}(\mathbf{p}')$ in Definition 4 to express $\overline{\mathbf{eff}}_\varepsilon(\mathbf{p})$ as follows

$$
\begin{aligned}
\overline{\mathbf{eff}}_\varepsilon(\mathbf{p}) = &\min_{\substack{\mathbf{p}'\in\mathcal{P}\\ \text{s.t } |\mathbf{p}'-\mathbf{p}|_1\leq\epsilon}} \quad \mathbf{eff}(\mathbf{p}')\\
= &\min_{\zeta,\mu_\ell\geq 0,\mathbf{p}'\in\mathcal{P}} \quad \frac{1}{\zeta}\\
&\text{subject to} \quad \sum_{\ell\in\mathcal{L}_{det}^{\perp}} \mu_\ell\ell(a,b|x,y) = \zeta p'(a,b|x,y) \quad \forall (a,b,x,y)\in\mathcal{A}\times\mathcal{B}\times\mathcal{X}\times\mathcal{Y}\\
&\qquad\qquad\quad \sum_{\ell\in\mathcal{L}_{det}^{\perp}} \mu_\ell = 1, \quad |\mathbf{p}'-\mathbf{p}|_1\leq\epsilon\\
= &\min_{\zeta,\mu_\ell\geq 0,\mathbf{\Delta}\in\Delta_\varepsilon} \quad \frac{1}{\zeta}\\
&\text{subject to} \quad \sum_{\ell\in\mathcal{L}_{det}^{\perp}} \mu_\ell\ell(a,b|x,y) =\\
&\qquad\qquad\quad \zeta[p(a,b|x,y) + \Delta(a,b|x,y)] \qquad \forall(a,b,x,y)\in\mathcal{A}\times\mathcal{B}\times\mathcal{X}\times\mathcal{Y}\\
&\qquad\qquad\quad \sum_{\ell\in\mathcal{L}_{det}^{\perp}} \mu_\ell = 1,
\end{aligned}
$$

where the last equality follows from Fact 2 and the fact that the first condition of the program imposes that $p(a,b|x,y) + \Delta(a,b|x,y)$ is nonnegative (since $\sum_\ell \mu_\ell\ell(a,b|x,y)$ is nonnegative). Since Δ_ε is a polytope, $\overline{\mathbf{eff}}_\varepsilon(\mathbf{p})$ can be expressed as the following linear program

$$
\begin{aligned}
\overline{\mathbf{eff}}_\varepsilon(\mathbf{p}) = &\min_{\zeta,\mu_\ell\geq 0,\nu_\Delta\geq 0} \quad \frac{1}{\zeta}\\
&\text{subject to} \quad \sum_{\ell\in\mathcal{L}_{det}^{\perp}} \mu_\ell\ell(a,b|x,y) = \zeta[p(a,b|x,y)+\\
&\qquad\qquad\quad \sum_{\mathbf{\Delta}\in\Delta_\varepsilon^{ext}} \nu_\Delta\Delta(a,b|x,y)] \qquad\qquad \forall(a,b,x,y)\in\mathcal{A}\times\mathcal{B}\times\mathcal{X}\times\mathcal{Y}\\
&\qquad\qquad\quad \sum_{\ell\in\mathcal{L}_{det}^{\perp}} \mu_\ell = 1, \qquad \sum_{\mathbf{\Delta}\in\Delta_\varepsilon^{ext}} \nu_\Delta = 1.
\end{aligned}
$$

Note that this can be written in standard LP form via the change of variables $\zeta' = 1/\zeta, \mu_\ell = \zeta w_\ell$. By LP duality, we then obtain

$$
\begin{aligned}
\overline{\mathbf{eff}}_\varepsilon(\mathbf{p}) = &\max_{B,\beta} \qquad\qquad\qquad\quad \beta\\
&\text{subject to} \qquad\qquad B(\mathbf{p}+\mathbf{\Delta}) \geq \beta \qquad \forall\mathbf{\Delta}\in\Delta_\varepsilon,\\
&\qquad\qquad\qquad\qquad\quad B(\boldsymbol{\ell}) \leq 1 \qquad\qquad \forall\boldsymbol{\ell}\in\mathcal{L}_{det}^{\perp}.
\end{aligned}
$$

Comparing this to the definition of $\mathbf{eff}_\varepsilon(\mathbf{p})$ (Definition 2) and together with Fact 1, we therefore have $\overline{\mathbf{eff}}_\varepsilon(\mathbf{p}) \leq \mathbf{eff}_\varepsilon(\mathbf{p})$.

□