# Deliverable D3.1

# New Bell inequalities resistant to noise and imperfection

## Due date: Month 12 (April 30, 2014)

# Classical and quantum partition bound and detector inefficiency *

Sophie Laplante
LIAFA, Université Paris Diderot Paris 7

Virginie Lerays
LIAFA, Université Paris Diderot Paris 7

Jérémie Roland
ULB, QuIC, Ecole Polytechnique de Bruxelles

**Abstract**

We study randomized and quantum efficiency lower bounds in communication complexity. These arise from the study of zero-communication protocols in which players are allowed to abort. Our scenario is inspired by the physics setup of Bell experiments, where two players share a predefined entangled state but are not allowed to communicate. Each is given a measurement as input, which they perform on their share of the system. The outcomes of the measurements should follow a distribution predicted by quantum mechanics; however, in practice, the detectors may fail to produce an output in some of the runs. The efficiency of the experiment is the probability that the experiment succeeds (neither of the detectors fails).

When the players share a quantum state, this gives rise to a new bound on quantum communication complexity (**eff**$^*$) that subsumes the factorization norm. When players share randomness instead of a quantum state, the efficiency bound (**eff**), coincides with the partition bound of Jain and Klauck. This is one of the strongest lower bounds known for randomized communication complexity, which subsumes all the known combinatorial and algebraic methods including the rectangle (corruption) bound, the factorization norm, and discrepancy.

The lower bound is formulated as a convex optimization problem. In practice, the dual form is more feasible to use, and we show that it amounts to constructing an explicit Bell inequality (for **eff**) or Tsirelson inequality (for **eff**$^*$). We give an example of a quantum distribution where the violation can be exponentially bigger than the previously studied class of normalized Bell inequalities.

For one-way communication, we show that the quantum one-way partition bound is tight for classical communication with shared entanglement up to arbitrarily small error.

## 1 Introduction

How are Bell tests related to communication complexity? At a high level, both involve two distant players, Alice and Bob, who receive inputs $x, y$, respectively, and produce outputs $a, b$ according to some distribution $p(a, b|x, y)$. The goal of a Bell test is to show that a given distribution $p(a, b|x, y)$ (typically arising from performing measurements on a shared entangled state) cannot result from a so-called local hidden variable model, which we will call here *local protocol* (or zero-communication protocol) for simplicity. A local protocol is a protocol where Alice and Bob may use shared randomness only but no communication (nor shared entanglement). In practice they may also abort. The interesting quantity for us is the *efficiency*, that is, the probability that the players do not abort. A lower efficiency makes it easier to reproduce the distribution using a local protocol, and a meaningful measure is therefore the maximum efficiency such that a local protocol for $p(a, b|x, y)$ exists. In the communication complexity model, Alice and Bob have inputs $x$ and $y$ respectively and must minimize the communication between them in order to solve a distributed task (or equivalently

*Extended version of the previously published article: Sophie Laplante, Virginie Lerays and Jérémie Roland, "Classical and quantum partition bound and detector efficiency", Proceedings of the 39th International Colloquium on Automata, Languages and Programming, 2012

output $a, b$ according to distribution $p(a, b|x, y)$). Both are measures of how far a given distribution is from the set of local (zero-communication) distributions.

Massar and Buhrman *et al.* [Mas02, BHMR03] described how a communication protocol gives rise to a local protocol where the players can abort: if there is a $c$-bit communication protocol where Alice and Bob output $a, b$ with distribution $p(a, b|x, y)$ when Alice's input is $x$ and Bob's input is $y$, then there is a local protocol that outputs according to $\mathbf{p}$ (conditioned on the run not being aborted) whose probability of not aborting is $2^{-c}$. Both players use the shared randomness to guess a transcript, and if they disagree with the transcript, they abort. Otherwise they output according to the transcript.

In this paper we prove a much stronger relation between communication protocols and the notion of efficiency in Bell tests, and provide applications both in communication complexity and in Bell inequality violations for quantum distributions. More precisely, we show that **the efficiency bound** (i.e., the maximum efficiency of any local protocol that simulates $\mathbf{p}$) is in fact **equal to the partition bound** in classical communication complexity [JK10]. The partition bound is important because it is one of the 'strongest' known classical communication lower bound techniques. Moreover we obtain a **strong new bound for quantum communication complexity**, which is at least as strong as all previously known lower bound techniques for quantum communication complexity. We show that the one-way version of the quantum efficiency bound is tight.

We show that the efficiency bound is equivalent to finding **Bell inequalities that are resistant to the detection loophole**, exhibiting an unexpected connection between these notions. This enables us to exhibit a quantum distribution arising from measurements on an $n$-dimensional shared quantum state, but which provides exponential Bell violations.

## 1.1 Communication complexity and the partition bound

Recently, Jain and Klauck [JK10] proposed a new lower bound on randomized communication complexity which subsumes two families of methods: the algebraic methods, including the nuclear norm and factorization norm, and combinatorial methods, including discrepancy and the rectangle or corruption bound. The algebraic methods and discrepancy give lower bounds on quantum communication complexity, whereas the rectangle bound can show polynomial lower bounds on randomized communication complexity for problems known to have logarithmic quantum protocols.

A longstanding open problem is whether there are total functions for which there is an exponential gap between classical and quantum communication complexities. Many partial results have been given [NS96, BCWdW01, BYJK08, GKK$^+$08], most recently [KR11]. These strong randomized lower bounds all use the distributional model, in which the randomness of the protocol is replaced by randomness in the choice of inputs, which are sampled according to some hard distribution. The equivalence of the randomized and distributional models, due to Yao's minmax theorem [Yao83], comes from strong duality of linear programming. This technique appears to be inherently non-applicable to quantum communication complexity (see for instance [dGdW02] which considers a similar question in the setting of query complexity), and the rectangle bound, as a result, was understood to be a inapplicable to quantum lower bounds.

Contrary to previous combinatorial type lower bounds, the partition bound is proven directly for randomized protocols, without first going to the distributional model. Although the partition bound re-introduces linear programming duality, the dual variables can no longer be interpreted as a (hard) distribution on the inputs. By the same token, it is harder to get intuition on how to obtain concrete lower bounds for explicit functions.

## 1.2 Bell experiments

Quantum nonlocality gives us a different viewpoint from which to consider lower bounds for communication complexity. A fundamental question of quantum mechanics is 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. 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 write $\mathbf{p}$ for the distribution, and $p(a, b|x, y)$ for the individual probabilities.) A Bell test [Bel64] 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 $B(\mathbf{p})$ is chosen together with a threshold $\tau$ so that any local classical distribution $\mathbf{p}'$ verifies $B(\mathbf{p}') \leq \tau$, but the chosen distribution $\mathbf{p}$ violates this inequality: $B(\mathbf{p}) > \tau$.

Although there have been numerous experiments that have validated the predictions of quantum mechanics, none so far has been totally "loophole-free". A loophole can be introduced, for instance, when the state preparation and the measurements are imperfect, or when the detectors are partially inefficient so that no measurement is registered in some runs of the experiment, or if the entangled particles are so close that communication may have taken place in the course of a run of the experiment. In such cases, there are classical explanations for the results of the experiment. For instance, if the detectors were somehow coordinating their behavior, they may choose to discard a run, and though the conditional probability (conditioned on the run not having been discarded) may look quantum, the unconditional probability may very well be classical. This is called the detection loophole. When an experiment aborts with probability at most $1 - \eta$, we say that the efficiency is $\eta$. (Here we assume that individual runs are independent of one another.) To close the detection loophole, the efficiency has to be high enough so that the classical explanations are ruled out. Gisin and Gisin show for example that the EPR correlations can be reproduced classically with 75% detector efficiency [GG99]. However, in practice, whenever the detectors can be placed far apart enough to prevent communication from taking place (typically in optics setups), the efficiency is extremely small (on the order of 10%), which is far too small to close the detection loophole.

What can Bell tests tell us about communication complexity? Both are measures of how far a distribution is from the set of local distributions (those requiring no communication), and one would expect that if a Bell test shows a large violation for a distribution, simulating this distribution should require a lot of communication, and vice versa. Degorre *et al.* showed that the factorization norm amounted to finding large Bell inequality violations for a particular class of Bell inequalities [DKLR11]. Here, we introduce a new class of Bell inequalities whose violation corresponds to the partition bound.

### 1.3 Summary of results

If we assume there is a $c$-bit classical communication protocol where Alice and Bob output $a, b$ with distribution $p(a, b|x, y)$ when Alice's input is $x$ and Bob's input is $y$, then there is a protocol without communication that outputs according to $\mathbf{p}$ (conditioned on the run not being discarded) that uses shared randomness and whose efficiency is $2^{-c}$: both players guess a transcript, and if they disagree with the transcript, they abort. Otherwise they follow the protocol using the transcript. As others have observed [Mas02, BHMR03], one can immediately derive a lower bound: let $\eta$ be the maximum efficiency of a protocol without communication that successfully simulates $\mathbf{p}$ with shared randomness. We define $\mathbf{eff}(\mathbf{p}) = 1/\eta$, and $\log(\mathbf{eff}(\mathbf{p}))$ is a lower bound on the communication complexity of simulating $\mathbf{p}$. This gives a surprisingly strong bound. We show that it coincides with the partition bound (in the special case of computing functions).

When we turn to the dual formulation, we get a natural physical interpretation, that of Bell inequalities. To prove a lower bound amounts to finding a good Bell inequality and proving a large violation. This is similar to finding a hard distribution and proving a lower bound in the distributional model of communication; but it is much stronger since the Bell functional is not required to have positive coefficients that sum to one.

Our approach leads naturally to a new "quantum partition bound" which is a strong lower bound on quantum communication complexity. Let $\mathbf{eff}^*(\mathbf{p}) = 1/\eta^*$, where $\eta^*$ is the maximum efficiency of a protocol without communication that successfully simulates $\mathbf{p}$ with shared entanglement. In the one-way setting, we show that the quantum partition bound is tight.

Allowing for runs to be discarded with some probability has been studied in different models of computation such as post-selection, and zero-error (Las Vegas) randomized computation. (Jain and Klauck [JK10] in fact introduce a Las Vegas partition bound for zero-error protocols.) This is a stronger requirement than allowing a probability of error since the errors must be flagged. Lee and Shraibman give a proof of

the factorization norm ($\gamma_2$) lower bound on (quantum) communication complexity based on the best bias one can achieve with no communication [LS09a, Theorem 60] (attributed to Buhrman; see also Degorre *et al.* [DKLR11]). In light of our formulation of the (quantum) partition bound, it is an easy consequence that the (quantum) partition bound is an upper bound on $\gamma_2$ (see e.g. [LS09b] for definitions of the factorization norm $\gamma_2$ and the related nuclear norm $\nu$, as well as [DKLR11] for their extensions to the communication complexity of distributions), making it the strongest known bound on quantum communication complexity to date.

The following gives a summary of our results. Full definitions and statements are given in the main text. Let $\mathrm{prt}(\mathbf{p})$ be the partition bound for a distribution $\mathbf{p}$ (defined in Section 3.1). $R_0(\mathbf{p})$ denotes the communication complexity of simulating $\mathbf{p}$ exactly using shared randomness and classical communication, and $Q_0^*(\mathbf{p})$ denotes the communication complexity of simulating $\mathbf{p}$ exactly using shared entanglement and quantum communication. One-way communication, where only Alice sends a message to Bob, is denoted by the superscript $\rightarrow$. In the simultaneous messages model, each player sends a message to the referee, who does not know the inputs of either player, and has to produce the output. This is denoted by the superscript $\|$. Shared entanglement is indicated by the superscript $*$. For any distribution $\mathbf{p}$,

- Theorem 4: $\mathrm{prt}(\mathbf{p}) = \mathbf{eff}(\mathbf{p})$,
- Theorem 5: $Q_0^*(\mathbf{p}) \geq \frac{1}{2}\log(\mathbf{eff}^*(\mathbf{p}))$,
- Theorem 7: $\gamma_2(\mathbf{p}) \leq 2\ \mathbf{eff}^*(\mathbf{p})$ and $\nu(\mathbf{p}) \leq 2\ \mathbf{eff}(\mathbf{p})$ (for nonsignaling $\mathbf{p}$),
- Theorem 14: $R_\epsilon^{*,\|}(\mathbf{p}) \leq O(\mathbf{eff}^*(\mathbf{p}))$ and $R_\epsilon^*(\mathbf{p}) \leq O(\sqrt{\mathbf{eff}^*(\mathbf{p})})$.

In the case of one-way communication, the upper bounds are much tighter. The one-sided efficiency measure, which we denote $\mathbf{eff}^{\rightarrow}$ is given in Definition 5.

- Theorem 6 : $Q_0^{*,\rightarrow}(\mathbf{p}) \geq \frac{1}{2}\log(\mathbf{eff}^{*,\rightarrow}(\mathbf{p}))$ and
- Theorem 15: $Q_\epsilon^{*,\rightarrow}(\mathbf{p}) \leq \frac{1}{2}\log(\mathbf{eff}^{*,\rightarrow}(\mathbf{p})) + O(1)$.

We can use smoothing to handle $\epsilon$ error, and we can formulate the bounds to allow both $\epsilon$ error and $\eta$ efficiency. (In the latter case, for boolean functions, this is equivalent to relaxing the exactness constraints in the linear programs.) For simplicity we have omitted these details in this summary.

We prove strong Bell violations using these new techniques for two problems studied in [BRSdW11]. The first is based on the Hidden Matching problem, from which we derive a distribution that can be simulated with zero communication and an $n$-dimensional shared quantum state. For this distribution, we prove an exponential Bell inequality violation for one-way efficiency resistant Bell inequalities, where one player is allowed to abort. In contrast, Junge et al [JPPG+10, JP11] have shown a linear upper bound on normalized Bell inequalities, as a function of the dimension of the shared state. Therefore, our lower bound exhibits an exponential gap between the usual normalized Bell inequalities and the new (one-way) efficiency resistant Bell inequalities.

## 1.4 Related work

Massar exhibits a Bell inequality that is more robust against detector inefficiency based on the distributed Deutsch Josza game [Mas02]. The Bell inequality is derived from the lower bound on communication complexity for this promise problem [BCW98, BCT99]. Massar shows an upper bound of $\mathbf{eff}(\mathbf{p})$ on expected communication complexity of simulating $\mathbf{p}$. He also states, but does not claim to prove, that a lower bound can be obtained as the logarithm of the efficiency. Buhrman *et al.* [BHMR03, BHMR06] show how to get Bell inequalities with better resistance to detector inefficiency by considering multipartite scenarios where players share GHZ type entangled states. Their technique is based on the rectangle bound and they derive a general tradeoff between monochromatic rectangle size, efficiency, and communication. They show a general lower bound on multiparty communication complexity which is exactly as we describe above.

Buhrman *et al.* [BRSdW11] show gaps between quantum and classical winning probability for games where the players are each given inputs and attempt, without communication, to produce outputs that satisfy some predicate. In the classical case they use shared randomness and in the quantum case, they use shared entanglement. Winning probabilities are linear so these translate to large Bell inequality violations.

Vértesi *et al.* show that there is a distribution with boolean outputs $\mathbf{p} \in \mathcal{Q}$, based on partially entangled states, such that (in our language) $\mathbf{eff}^{\rightarrow}(\mathbf{p}) = \Omega(2^n)$ [VPB10]. Therefore, our results imply that $R_0^{\rightarrow}(\mathbf{p}) = n$. Since the states are nearly separable, however $R_\epsilon^{\rightarrow}(\mathbf{p}) = 0$ for large enough $\epsilon$.

Lower bounds for communication complexity of simulating distributions were first studied in a systematic way by Degorre *et al.* [DKLR11]. These bounds are shown to be closely related to the nuclear norm and factorization norm [LS09b], and the dual expressions are interpreted as Bell inequality violations.

Following up on the results in this paper, Kerenidis et al [KLL+12] used the notion of efficiency of zero-communication protocols to show that the information cost is bounded below by a relaxation of the partition bound which is larger than the smooth rectangle and $\gamma_2$ bounds. Jain and Yao [JY12] followed up with a strong direct product theorem for the communication complexity of all functions for which an optimal lower bound can be shown using the smooth rectangle bound. Using a similar notion of zero-communication protocols, Gavinski and Lovett [GL13] showed that the log rank conjecture is equivalent to an upper bound which is polylogarithmic in the rank on the zero-communication cost. The notion of zero-communication cost that they use is the non-constant efficiency (Definition 10 and Lemma 19.)

# 2 Preliminaries

## 2.1 Classical partition bound

The partition bound of Jain and Klauck [JK10] is given as a linear program, following the approach introduced by Lovász [Lov90] and studied in more depth by Karchmer *et al.* [KKN95]. It differs from the rectangle and other combinatorial bounds in that it is formulated directly on the randomized protocol, as opposed to first applying Yao's minmax theorem to reduce to a deterministic protocol with distributional inputs. From a $c$-bit, $\epsilon$-correct randomized protocol one can infer a distribution over rectangle partitions of size at most $2^c$, where each rectangle is assigned an output value $z$. Set weights $w_{R,z}$ to be the probability that rectangle $R$ occurs with label $z$ (the same rectangle may occur in different partitions, with different labels and different probabilities). This is a feasible solution to the following linear program.

**Definition 1** (Partition bound [JK10])**.** *Let $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be any partial function whose domain we write $f^{-1}$. Then $prt_\epsilon(f)$ is defined to be the optimal value of the linear program, where $R$ ranges over the rectangles from $\mathcal{X} \times \mathcal{Y}$ and $z$ ranges over the set $\mathcal{Z}$:*

$$\begin{aligned} \mathrm{prt}_\epsilon(f) = \min_{w_{R,z} \geq 0} \quad & \sum_{R,z} w_{R,z} & \\ \textit{subject to} \quad & \sum_{R:(x,y)\in R} w_{R,f(x,y)} \geq 1 - \epsilon & \forall x, y \in f^{-1} \\ & \sum_{z} \sum_{R:(x,y)\in R} w_{R,z} = 1 & \forall x, y \in \mathcal{X} \times \mathcal{Y} \end{aligned}$$

The feasible solution sketched above verifies all the constraints and the objective value is at most $2^c$, so $R_\epsilon(f) \geq \log(\mathrm{prt}_\epsilon(f))$. The partition bound subsumes almost all previously known techniques [JK10], in particular the factorization norm [LS09b], smooth rectangle [JK10] and rectangle or corruption bound [Yao83], and discrepancy [CG85, BNS89]. Bounds not known to be subsumed by the partition bound include the approximate rank [Kra96, BdW01] and the information complexity [CSWY01].

## 2.2 Local and quantum distributions

Given a distribution $\mathbf{p}$, how much communication is required if Alice is given $x \in \mathcal{X}$, Bob is given $y \in \mathcal{Y}$, and their goal is to output $a, b \in \mathcal{A} \times \mathcal{B}$ with probability $p(a, b|x, y)$?

Some classes of distributions are of interest and have been widely studied in quantum information theory since the seminal paper of Bell [Bel64]. The local deterministic distributions, denoted $\ell \in \mathcal{L}_{\text{det}}$, 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 exchanging any communication. The local distributions $\mathcal{L}$ are any distribution over the local deterministic strategies. Mathematically this corresponds to taking convex combinations of the local deterministic distributions, and operationally to zero-communication protocols with shared randomness.

We focus our attention in this paper on local strategies that are allowed to abort the protocol with some probability. When they abort, they output the symbol $\perp$. We will use the notation $\mathcal{L}^{\perp}_{\text{det}}$ and $\mathcal{L}^{\perp}$ to denote these strategies, where $\perp$ is added to the possible outputs for both players, and $\perp \notin \mathcal{A} \cup \mathcal{B}$. Therefore, when $\ell \in \mathcal{L}^{\perp}_{\text{det}}$ or $\mathcal{L}^{\perp}$, $l(a, b|x, y)$ is *not* conditioned on $a, b \neq \perp$ since $\perp$ is a legal output for such distributions.

The quantum distributions, denoted $\mathbf{q} \in \mathcal{Q}$, are the ones that result from applying measurements to each part of a shared entangled bipartite state. Each player outputs the measurement outcome. 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 a sampling problem as follows. Consider a boolean function $f : \mathcal{X} \times \mathcal{Y} \to \{0, 1\}$ whose communication complexity we wish to study (non-boolean 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 $\mathbf{p}_f$. 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 $f$. We are usually interested in distributions requiring nontrivial communication complexity, and lie well beyond these sets.

## 2.3 Communication complexity measures

We use the following notation for communication complexity of distributions. $R_\epsilon(\mathbf{p})$ is the minimum amount of communication necessary to reproduce the distribution $\mathbf{p}$ in the worst case, up to $\epsilon$ in total variation distance for all $x, y$. We write $|\mathbf{p} - \mathbf{p}'|_1 \leq \epsilon$ to mean that for any $x, y$, $\sum_{a,b} |p(a, b|x, y) - p'(a, b|x, y)| \leq \epsilon$.

$R^{\eta}_0(\mathbf{p})$ is the amount of communication needed to reproduce $\mathbf{p}$ exactly with a protocol which may abort with probability at most $1 - \eta$ for any input $x, y$ (the probability that it aborts may depend on $x, y$). The probability produced by the protocol is conditioned on the event that neither player aborts. When the player aborts it outputs $\perp$.

For quantum communication, we use $Q$ to denote quantum communication, and we use the superscript $*$ to denote the presence of shared entanglement. We use superscripts $\to$ for one-way communication (i.e, when only Alice can send a message to Bob), and $\|$ for simultaneous messages (i.e., when Alice and Bob cannot communicate with each other, but are only allowed to send a message to a third party who should produce the final output of the protocol). The usual relation $Q^{\eta,*}_\epsilon(\mathbf{p}) \leq R^{\eta}_\epsilon(\mathbf{p})$ holds for any $\epsilon, \eta, \mathbf{p}$. Moreover, since one can always output at random instead of aborting, which introduces at most $1 - \eta$ error for each $x, y$, we have the following relation between $R_\epsilon(\mathbf{p})$ and $R^{\eta}_\epsilon(\mathbf{p})$.

**Lemma 1.** *For any $\epsilon, \eta$ and any distribution $\mathbf{p}$, we have $R_{\epsilon+(1-\eta)}(\mathbf{p}) \leq R^{\eta}_\epsilon(\mathbf{p})$.*

For all the models of randomized communication, we assume shared randomness between the players. Except in the case of simultaneous messages, this is the same as private randomness up to a logarithmic additive term [New91]. (Ref. [DKLR11] sketches a proof of how to adapt Newman's theorem to the case of simulating distributions.)

# 3 Partition bound and detector inefficiency

## 3.1 The partition bound for distributions

We extend the partition bound to the more general setting of simulating a distribution $p(a,b|x,y)$ instead of computing a function. Protocols with shared randomness and communication also lead to a distribution over rectangle partitions; however, since each player ouputs a value, the label associated with each rectangle is a local deterministic distribution, denoted by $\ell$. The following definition applies to protocols that use communication and allow the players to abort a run with some fixed probability $1-\eta$. The partition bound corresponds to the case $\eta = 1$ and the Las Vegas partition bound [JK10] is closely related to the case $\eta = 1/2$.

**Definition 2.** *For any distribution $\mathbf{p} = p(a,b|x,y)$, over inputs $x \in \mathcal{X}, y \in \mathcal{Y}$ and outputs $a \in \mathcal{A}, b \in \mathcal{B}$, define $\mathrm{prt}^\eta(\mathbf{p})$ to be the optimal value of the following linear program. The variables of the program are $\eta_{x,y}$ and $w_{R,\ell}$, where $R$ ranges over the rectangles from $\mathcal{X}\times\mathcal{Y}$ and $\ell$ ranges over the local deterministic distributions with inputs in $R$ and with outputs in $\mathcal{A}\times\mathcal{B}$.*

$$\begin{aligned}
\mathrm{prt}^\eta(\mathbf{p}) = \min_{w_{R,\ell\geq 0},\eta_{x,y}} \quad & \sum_{R,\ell\in\mathcal{L}_{\mathrm{det}}} w_{R,\ell} & \\
\textit{subject to} \quad & \sum_{R,\ell\in\mathcal{L}_{\mathrm{det}}:x,y\in R} w_{R,\ell}\cdot l(a,b|x,y) = p(a,b|x,y)\cdot\eta_{x,y} & \forall x,y,a,b \in \mathcal{X}\times\mathcal{Y}\times\mathcal{A}\times\mathcal{B} \\
& \eta \leq \eta_{x,y} \leq 1 & \forall x,y \in \mathcal{X}\times\mathcal{Y}.
\end{aligned}$$

*When $\eta = 1$ we write $\mathrm{prt}(\mathbf{p}) = \mathrm{prt}^1(\mathbf{p})$, and the linear program simplifies:*

$$\begin{aligned}
\mathrm{prt}(\mathbf{p}) = \min_{w_{R,\ell}\geq 0} \quad & \sum_{R,\ell} w_{R,\ell} & \\
\textit{subject to} \quad & \sum_{R,\ell:x,y\in R} w_{R,\ell}\cdot l(a,b|x,y) = p(a,b|x,y) & \forall x,y,a,b \in \mathcal{X}\times\mathcal{Y}\times\mathcal{A}\times\mathcal{B}
\end{aligned}$$

*For randomized communication with error, $\mathrm{prt}^\eta_\epsilon(\mathbf{p}) = \min_{|p'-p|_1\leq\epsilon}\mathrm{prt}^\eta(\mathbf{p}')$.*

**Theorem 2.** *For any distribution $\mathbf{p}$, $R^\eta_\epsilon(\mathbf{p}) \geq \log(\mathrm{prt}^\eta_\epsilon(\mathbf{p}))$.*

We have included a direct proof of the theorem in Appendix A, which for $\eta = 1$ is essentially the same as the original partition bound, sketched above, where output values $z$ are replaced with local deterministic strategies $\ell$. We will now turn to an alternative, arguably simpler, proof by introducing the efficiency bound.

## 3.2 The efficiency bound

For any distribution $\mathbf{p}$, $\mathbf{eff}(\mathbf{p})$ is the inverse of the maximum efficiency sufficient to simulate it classically with shared randomness, without communication.

**Definition 3.** *For any distribution $\mathbf{p}$ with inputs $\mathcal{X}\times\mathcal{Y}$ and outputs in $\mathcal{A}\times\mathcal{B}$, $\mathbf{eff}(\mathbf{p}) = 1/\zeta_{\mathrm{opt}}$, where $\zeta_{\mathrm{opt}}$ is the optimal value of the following linear program. The variables are $\zeta$ and $q_\ell$, where $\ell$ ranges over local deterministic protocols with inputs taken from $\mathcal{X}\times\mathcal{Y}$ and outputs in $\mathcal{A}\cup\{\perp\}\times\mathcal{B}\cup\{\perp\}$.*

$$\begin{aligned}
\zeta_{\mathrm{opt}} = \max_{\zeta,q_\ell\geq 0} \quad & \zeta & \\
\textit{subject to} \quad & \sum_{\ell\in\mathcal{L}^\perp_{\mathrm{det}}} q_\ell l(a,b|x,y) = \zeta p(a,b|x,y) & \forall x,y,a,b \in \mathcal{X}\times\mathcal{Y}\times\mathcal{A}\times\mathcal{B} \\
& \sum_{\ell\in\mathcal{L}^\perp_{\mathrm{det}}} q_\ell = 1 &
\end{aligned}$$

*For randomized communication with error, define $\mathbf{eff}_\epsilon(\mathbf{p}) = \min_{|p'-p|_1\leq\epsilon}\mathbf{eff}(\mathbf{p}')$.*

The first constraint says that the local distribution, conditioned on both outputs differing from $\perp$, equals $\mathbf{p}$, and the second is a normalization constraint. Note that the efficiency $\zeta$ is the same for every input $x, y$. This is surprisingly important and the relaxation $\zeta_{x,y} \geq \zeta$ does not appear to coincide with the partition bound. Other more realistic variants (for the Bell setting), such as players aborting independently of one another, could be considered as well. (We note that this would not result in a linear program.)

**Theorem 3.** *[Mas02, BHMR03]* $R_\epsilon(\mathbf{p}) \geq \log \mathbf{eff}_\epsilon(\mathbf{p})$.

*Proof.* Let $P$ be a randomized communication protocol for a distribution $\mathbf{p}'$ with $|\mathbf{p} - \mathbf{p}'|_1 \leq \epsilon$, using $t$ bits of communication. $P$ is a convex combination of deterministic protocols: we denote by $\Lambda$ the source of public randomness and by $P_\lambda$ the deterministic protocol corresponding to $\lambda \in \Lambda$. We assume that the total number of bits exchanged is independent of the execution of the protocol, introducing dummy bits at the end of the protocol if necessary. Let $q_l$ be the following distribution over local deterministic protocols $\ell$: using shared randomness Alice and Bob first pick some random value $\lambda$ according to $\Lambda$ and then they pick a transcript $T \in \{0,1\}^t$. If $T$ is consistent with $P_\lambda$, Alice outputs according to $P_\lambda$, otherwise she outputs $\perp$; similarly for Bob. We claim that for each $\lambda \in \Lambda$ only one transcript is valid for Alice and Bob simultaneously, so the probability that neither player outputs $\perp$ is exactly $2^{-t}$. Indeed, define $A_x = \{T : \exists y', T = T_{x,y'}\}$ and $B_y = \{T : \exists x', T = T_{x',y}\}$. Then $A_x \cap B_y = \{T_{x,y}\}$ because $T$ in $A_x \cap B_y$ means that there exist $x'$ and $y'$ such that $T = T_{x',y} = T_{x,y'}$. By the rectangle property of the transcripts of $P_\lambda$ it must be the case that $T = T_{x,y}$. Furthermore, if we condition on not aborting, this protocol does exactly the same thing as $P$. This distribution therefore satisfies the constraints of $\mathbf{eff}(\mathbf{p}')$ with $\zeta = 2^{-t}$. □

**Theorem 4.** *For any distribution* $\mathbf{p}$, $\mathbf{eff}(\mathbf{p}) = \mathrm{prt}(\mathbf{p})$.

*Proof.* In the partition bound, a pair $(\ell, R)$, where $\ell$ is a local distribution with outputs in $\mathcal{A} \times \mathcal{B}$ and $R$ is a rectangle, determines a local distribution $\ell_R$ with outputs in $(\mathcal{A} \cup \{\perp\}) \times (\mathcal{B} \cup \{\perp\})$, where Alice outputs as in $\ell$ if $x \in R$, and outputs $\perp$ otherwise (similarly for Bob). Let $(a_0, b_0) \in \mathcal{A} \times \mathcal{B}$ be an arbitrary pair of outputs. In the efficiency bound, a distribution $\ell \in \mathcal{L}^\perp_{\mathrm{det}}$ defines both a rectangle being the set of inputs where neither Alice nor Bob abort, and a local distribution $\ell' \in \mathcal{L}_{\mathrm{det}}$ where Alice outputs as $\ell$ if the output is different from $\perp$ and $a_0$ otherwise (similarly for Bob with $b_0$). We can transform the linear program for $\mathrm{prt}(\mathbf{p})$ into the linear program for $\mathbf{eff}(\mathbf{p})$ by making the change of variables: $\zeta = \left(\sum_{R,\ell} w_{R,\ell}\right)^{-1}$ and $q_{\ell_R} = \zeta\, w_{R,\ell}$. □

We define $\mathbf{eff}^\eta(\mathbf{p})$ which is equal to $\mathrm{prt}^\eta(\mathbf{p})$. The details are given in Appendix B.

### 3.3 Lower bound for quantum communication complexity

By replacing local distributions by quantum distributions we get a strong new lower bound on quantum communication that subsumes the factorization norm. Insomuch as the partition bound is an extension of the rectangle bound, this quantum analogue of the partition bound can be thought of as a quantum extension of the rectangle bound, circumventing Yao's minmax theorem.

**Definition 4.** *For any distribution* $\mathbf{p}$ *with inputs* $\mathcal{X} \times \mathcal{Y}$ *and outputs* $\mathcal{A} \times \mathcal{B}$, $\mathbf{eff}^*(\mathbf{p}) = 1/\eta^*$, *with* $\eta^*$ *the optimal value of the following (non-linear) program.*

$$\begin{array}{lll} \max\limits_{\zeta, \mathbf{q} \in \mathcal{Q}^\perp} & \zeta & \\ \textit{subject to} & q(a,b|x,y) = \zeta p(a,b|x,y) & \forall x,y,a,b \in \mathcal{X}\times\mathcal{Y}\times\mathcal{A}\times\mathcal{B} \end{array}$$

*As before, we let* $\mathbf{eff}^*_\epsilon(\mathbf{p}) = \min_{|p'-p|_1 \leq \epsilon} \mathbf{eff}^*(\mathbf{p}')$.

**Theorem 5.** $Q^*_\epsilon(\mathbf{p}) \geq \frac{1}{2} \log \mathbf{eff}^*_\epsilon(\mathbf{p})$.

*Proof.* Let $Q$ be a $t$ qubit communication protocol for $\mathbf{p}'$ with $|\mathbf{p}' - \mathbf{p}|_1 \leq \epsilon$. We use teleportation and classical communication to send every qubit, hence obtaining an entanglement-assisted protocol using at most $2t$ bits of classical communication. We introduce dummy bits so that the number of bits exchanged is exactly $2t$, independently of the execution of the protocol. Then proceed as before: guess the a uniform classical transcript using the shared randomness. Then Alice and Bob will simulate the previous entangled-assisted communication protocol (performing the measurements) checking that the communication in the transcript is consistent with their execution of the protocol. If it is the case, they output according to the protocol, and they abort otherwise. If we fix the outputs of the measurements then the protocol is deterministic so, as before there is exactly one transcript which is valid simultaneously for $x$ and $y$, and the efficiency is $2^{-2t}$. Conditioning on not aborting, this protocol outputs exactly with the same distribution as before. The result is a protocol using zero communication and entanglement with efficiency $2^{-2t}$, satisfying the constraints. □

Since the local distributions form a subset of the quantum distributions, $\mathbf{eff}^*(\mathbf{p}) \leq \mathbf{eff}(\mathbf{p})$ for any $\mathbf{p}$.

### 3.4 One-way efficiency bound

Because of its rectangle-based formulation, the partition and efficiency bounds can easily be tailored to the case of one-way communication protocols. In the case of the partition bound, we consider only rectangles of the form $X \times Y$ with $Y = \mathcal{Y}$. In the case of the efficiency bound, this amounts to only letting Alice abort the protocol. The set of local (resp. quantum) distributions where only Alice can abort is denoted $\mathcal{L}_{\text{det}}^{\perp A}$ (resp. $\mathcal{Q}^{\perp A}$).

**Definition 5.** *Define* $\mathbf{eff}^{\rightarrow}$ *and* $\mathbf{eff}^{*,\rightarrow}$ *as*

$$
\begin{aligned}
(\mathbf{eff}^{\rightarrow}(\mathbf{p}))^{-1} = \max_{\zeta, q_\ell \geq 0} \quad & \zeta \\
\textit{subject to} \quad & \sum_{\ell \in \mathcal{L}_{\text{det}}^{\perp A}} q_\ell l(a,b|x,y) = \zeta p(a,b|x,y) \qquad \forall a \in \mathcal{A}, b \in B, x, y \in \mathcal{X} \times \mathcal{Y} \\
& \sum_{\ell \in \mathcal{L}_{\text{det}}^{\perp A}} q_\ell = 1 \\
(\mathbf{eff}^{*,\rightarrow}(\mathbf{p}))^{-1} = \max_{\zeta, \mathbf{q} \in \mathcal{Q}^{\perp A}} \quad & \zeta \\
\textit{subject to} \quad & q(a,b|x,y) = \zeta p(a,b|x,y) \qquad \forall a \in \mathcal{A}, b \in B, x, y \in \mathcal{X} \times \mathcal{Y}.
\end{aligned}
$$

**Theorem 6.** $R_0^{\rightarrow}(\mathbf{p}) \geq \log \mathbf{eff}^{\rightarrow}(\mathbf{p})$ *and* $Q_0^{*,\rightarrow}(\mathbf{p}) \geq \frac{1}{2} \log \mathbf{eff}^{*,\rightarrow}(\mathbf{p})$.

The proof is similar to the two-way case.

### 3.5 Efficiency is larger than $\gamma_2$

Jain and Klauck show that the partition bound is an upper bound on $\gamma_2$ for boolean functions (in fact they show that the weaker smooth rectangle bound is an upper bound on $\gamma_2$) [JK10]. The lower bounds $\nu$ and $\gamma_2$ were extended to nonsignaling distributions by Degorre *et al.* [DKLR11]. Nonsignaling distributions, including quantum distributions, have marginal distributions independent of the other player's input.

**Definition 6** (Non-signaling distributions)**.** *A distribution* $\mathbf{p}$ *is nonsignaling if* $\forall a, x, y, y', \sum_b p(a,b|x,y) = \sum_b p(a,b|x,y')$*, and* $\forall b, x, x', y, \sum_a p(a,b|x,y) = \sum_a p(a,b|x',y)$.

**Definition 7** ([DKLR11])**.** *For any nonsignaling distribution* $\mathbf{p}$*,*

1. $\nu(\mathbf{p}) = \min\{\sum_i |\, q_i \,|\colon \exists \mathbf{p}_i \in \mathcal{L}, q_i \in \mathbb{R}, \mathbf{p} = \sum_i q_i \mathbf{p}_i\}$,

2. $\gamma_2(\mathbf{p}) = \min\{\sum_i |q_i| : \exists \mathbf{p}_i \in \mathcal{Q}, q_i \in \mathbb{R}, \mathbf{p} = \sum_i q_i \mathbf{p}_i\}$,

Recall from Section 2.2 the definition of the distribution $\mathbf{p}_f$ for any boolean function $f$. It was shown that for any Boolean function $f$, the factorization norm $\gamma_2(f) = \Theta(\gamma_2(\mathbf{p}_f))$, and similarly for the nuclear norm, $\nu(f) = \Theta(\gamma_2(\mathbf{p}_f))$ [DKLR11].

**Theorem 7.** *For any nonsignaling* $\mathbf{p}$, $\nu(\mathbf{p}) \leq 2\ \mathbf{eff}(\mathbf{p})$ *and* $\gamma_2(\mathbf{p}) \leq 2\ \mathbf{eff}^*(\mathbf{p})$.

*Proof.* We sketch the proof for $\gamma_2$ vs $\mathbf{eff}^*$. The proof for $\nu$ vs. $\mathbf{eff}$ is similar.

Let $\zeta, \mathbf{q}$ be an optimal solution for $\mathbf{eff}^*(\mathbf{p})$. Then $q(a,b|x,y) = \zeta p(a,b|x,y)$, $\forall x,y,a,b \in \mathcal{X}\times\mathcal{Y}\times\mathcal{A}\times\mathcal{B}$, where $\mathbf{q}$ outputs $\perp$ with probability $1-\zeta$ for every $x, y$. Define $\tilde{\mathbf{q}} \in \mathcal{Q}$ as the distribution where the players output according to $\mathbf{q}$ unless their outcome is $\perp$, in which case they output independently from the other player, uniformly at random from $\mathcal{A}$ or $\mathcal{B}$. Let $\mathbf{r}$ be the distribution $\tilde{\mathbf{q}}$ conditioned on one of the players having output $\perp$ when they ran $\mathbf{q}$. We can write $\tilde{q}(a,b|x,y) = q(a,b|x,y) + (1-\zeta)r(a,b|x,y)$. Notice that $\mathbf{r} \in \mathcal{Q}$. Therefore, on $\mathcal{A}\times\mathcal{B}$, $\mathbf{p} = \frac{1}{\zeta}\mathbf{q} = \frac{1}{\zeta}\tilde{\mathbf{q}} - \frac{1-\zeta}{\zeta}\mathbf{r}$ . This is an affine combination of quantum distributions (that do not output $\perp$) so $\gamma_2(\mathbf{p}) \leq |\frac{1}{\zeta}| + |-\frac{1-\zeta}{\zeta}| = 2\mathbf{eff}^*(\mathbf{p}) - 1$. □

For Boolean functions, the gap between $\nu$ and $\gamma_2$ is known to be at most a multiplicative constant (by Grothendieck's inequality). However, there is no immediate way to conclude similarly for $\mathbf{eff}$ vs. $\mathbf{eff}^*$. Since these are stronger bounds, determining the largest possible gap between these measures could lead to further evidence towards the existence, or not, of exponential gaps between quantum and classical communication complexity for total boolean functions.

# 4 Detector resistant Bell inequalities

## 4.1 Dual formulation of the efficiency bound

Proving lower bounds in the standard formulation of $\mathbf{eff}$ is difficult since it is formulated as the multiplicative inverse of a maximisation problem, which translates to a universal quantifier on all the variables of the optimization problem. To prove concrete lower bounds, we will use the dual formulation, expressed as a maximisation problem, where it suffices to check a feasible solution against all of the constraints.

**Lemma 8** (Dual formulation of the efficiency bounds)**.** *For any distribution* $\mathbf{p}$*,*

$$\begin{aligned}
\mathbf{eff}(\mathbf{p}) = \max_B \quad & B(\mathbf{p}) & \\
\textit{subject to} \quad & B(\ell) \leq 1 & \forall \ell \in \mathcal{L}_{\text{det}}^{\perp} \\
\mathbf{eff}^*(\mathbf{p}) = \max_B \quad & B(\mathbf{p}) & \\
\textit{subject to} \quad & B(\mathbf{q}) \leq 1 & \forall \mathbf{q} \in \mathcal{Q}^{\perp}
\end{aligned}$$

*where* $B$ *ranges over all real linear functionals, with coefficients* $B_{abxy}$ *and for any distribution* $\mathbf{p}$*,* $B(\mathbf{p}) = \sum_{(a,b,x,y)\in\mathcal{A}\times\mathcal{B}\times\mathcal{X}\times\mathcal{Y}} B_{abxy} p(a,b|x,y)$*. (There are no coefficients for the abort outcomes.)*

The first part uses linear programming duality and the second can be shown using Lagrange multipliers.

These expressions are clearly reminiscent of Bell inequalities, except for the introduction into the constraints of local strategies that may abort. Let us compare the classes of Bell inequalities above, with those stemming from the lower bound $\nu$, introduced in [DKLR11] and studied in Junge *et al.* and Buhrman *et al.* [JPPG+10, BRSdW11]. The dual of $\nu$ can be formulated as follows:

**Proposition 9** ([DKLR11])**.**

$$\begin{aligned}
\nu(\mathbf{p}) = \max_B \quad & B(\mathbf{p}) & \\
\textit{subject to} \quad & |B(\ell)| \leq 1 & \forall \ell \in \mathcal{L}_{\text{det}}
\end{aligned}$$

We will call the family of Bell inequalities satisfying the constraints of the dual formulation of $\nu$ *normalized Bell inequalities*, and those satisfying the constraints of the dual expression for $\mathbf{eff}$, *inefficiency resistant Bell inequalities*. Theorems 7 and 3 tell us that $\frac{1}{2}\nu(\mathbf{p}) \leq \mathbf{eff}(\mathbf{p}) \leq 2^{R(\mathbf{p})}$. This is not immediately apparent by comparing the two dual expressions, since there are two differences between the two families of Bell inequalities. One is a relaxation of the normalized Bell inequalities by removing the absolute value in the constraints. This increases the value of $\mathbf{eff}$, and one might worry that this will lead to unbounded violations. (Of course this is ruled out by the upper bound on $\mathbf{eff}$.) The second difference goes in the other direction: the constraint on inefficiency resistant inequalities is required to hold for all local strategies including those that may abort. Notice that to get a value of $\mathbf{eff}$ that is larger than $\nu$, there have to be local strategies that abort whose Bell value is (far) less than -1, otherwise one is just exhibiting a normalized Bell inequality with added constraints.

Concretely, how does one go about finding a feasible solution to the dual? Consider a distribution $\mathbf{p}$ for which we would like to find a lower bound. We construct a Bell inequality $B(\mathbf{p}) = \sum_{a,b,x,y} B_{abxy} p(a,b|x,y)$ so that $B(\mathbf{p})$ is large, and $B(\ell)$ is small for every $\ell \in \mathcal{L}^{\perp}$. The goal is to balance the coefficients $B_{abxy}$ so that they correlate well with the distribution $\mathbf{p}$ and badly with local strategies. For $B(\ell)$ to be small for local strategies, we apply a small weight or even a penalty (negative weight) when the local strategy is incorrect. For $B(\mathbf{p})$ to be large, we assign a positive coefficient when the outcome is correct, or if it should occur with high probability. Weights can be zero when the input is not contributing to the hardness of the problem.

The dual of the one-way efficiency bound can also be interpreted as Bell inequality violations.

**Lemma 10** (Dual formulation for one-way efficiency)**.** *For any* $\mathbf{p}$*,*

$$\begin{aligned}
\mathbf{eff}^{\rightarrow}(\mathbf{p}) = \max_{B_{abxy}} \quad & B(\mathbf{p}) & \\
\textit{subject to} \quad & B(\ell) \leq 1 & \forall \ell \in \mathcal{L}^{\perp A}_{\text{det}} \\
\mathbf{eff}^{*,\rightarrow}(\mathbf{p}) = \max_{B} \quad & B(\mathbf{p}) & \\
\textit{subject to} \quad & B(\mathbf{q}) \leq 1 & \forall \mathbf{q} \in \mathcal{Q}^{\perp A}
\end{aligned}$$

Notice that the constraint need only be verified for those local strategies where Alice is allowed to abort.

## 4.2 Bell violation for the Hidden Matching distribution

We show how to apply the efficiency bound to derive an exponential efficiency-resistant Bell violation for the Hidden Matching problem [BYJK08, GKK+08, BRSdW11]. The Hidden Matching problem can be formulated as a game that can be won with probability 1 by a quantum protocol with zero communication and an $O(n)$-dimensional shared quantum state. Buhrman *et al.* show a normalized Bell violation of $\Omega(\frac{\sqrt{n}}{\log(n)})$ for this game. Our exponential violation should also be compared to the results of Junge *et al.* [JPPG+10] who show that if a dostribution $\mathbf{p}$ can be simulated with a quantum protocol with an $n$ dimensional shared quantum state, then $\nu(p) \leq n$. As we have discussed above, we are dealing here with a different family of Bell inequalities. More precisely, we establish that there can be an exponential gap between $\nu$ and $\mathbf{eff}^{\rightarrow}$.

We apply the partition bound method on the Hidden Matching probability distribution that we define here. The Hidden Matching distribution is based on the Hidden Matching problem of [BYJK08] adapted to the setting of games by Buhrman *et al.* [BRSdW11]. We use many of the ideas and techniques from the latter to give the efficiency bound, but some added tricks are needed to derive the exponentially larger Bell violations.

**Definition 8** (Hidden Matching distribution)**.** *Alice receives* $x \in \{0,1\}^n$ *and Bob receives a matching* $M$ *over vertices* $\{1,\ldots,n\}$*. Alice has to output* $a \in \{0,1\}^{\log(n)}$ *and Bob has to output* $d \in \{0,1\}$ *and* $(i,j) \in M$ *according to the following distribution, which we call the* Hidden Matching distribution*:*

$$\text{HM}(a,d,i,j|x,M) = \begin{cases} \frac{2}{n^2} & \text{if } \langle a, i \oplus j \rangle \oplus d = x_i \oplus x_j \\ 0 & \text{otherwise.} \end{cases}$$

**Theorem 11** ([BRSdW11])**.** $\mathrm{HM} \in \mathcal{Q}$*, that is,* $Q_0^*(\mathrm{HM}) = 0$*, and* $\mathbf{eff}^*(\mathrm{HM}) = 1$*. The zero-communication quantum protocol for* HM *uses an* $n$*-dimensional shared quantum state.*

**Theorem 12.** *There exists a constant* $\mathcal{C} > 0$ *such that for any* $0 \leq \epsilon < \frac{1}{2}$*,* $\mathbf{eff}_\epsilon^{\rightarrow}(\mathrm{HM}) \geq \frac{2^{\frac{\sqrt{n-1}}{2\mathcal{C}}}}{n}(\frac{1}{2} - \epsilon)$*.*

From the one-way version of Lemma 1, we obtain as a corollary:

**Corollary 13.** *There exists a constant* $\mathcal{C} > 0$ *such that for any* $0 < \eta \leq 1$ *and* $0 \leq \epsilon < \eta - \frac{1}{2}$*,* $R_\epsilon^{\eta,\rightarrow}(\mathrm{HM}) \geq \frac{\sqrt{n-1}}{2\mathcal{C}} - \log(n) + \log(\eta - \epsilon - \frac{1}{2})$*.*

We give a bound on $\mathbf{eff}_\epsilon^{\rightarrow}(\mathrm{HM}) = \min_{\{p':|p'-\mathrm{HM}|_1 \leq \epsilon\}}(\mathbf{eff}^{\rightarrow}(\mathbf{p}'))$. To do that, we need to find a Bell functional which is bounded from above for any local deterministic strategy where Alice may abort and which is large for any distribution close to the Hidden Matching distribution. The main idea is to define a Bell functional which translates the bias of the winning probability of the strategy putting negative weights when the strategy fails and positive ones when it succeeds. Then we show that for any local deterministic strategy where Alice may abort, this bias is small, whereas it is big for any distribution close to the Hidden Matching distribution. Details of the proof are given in Appendix C.

# 5 Upper bounds for one- and two-way communication

The efficiency bound subsumes most known lower bound techniques for randomized communication complexity. How close is it to being tight? An upper bound on randomized communication is proven by Massar [Mas02]. We give a similar bound for entanglement assisted communication complexity in terms of $\mathbf{eff}^*$. Our bounds are stated for zero-error communication complexity where the players may abort with some probability $1 - \eta$. The weaker statement with $\epsilon$ error can be derived using Lemma 1.

**Theorem 14.** *For any distribution* $\mathbf{p}$ *with outputs in* $\mathcal{A}, \mathcal{B}$*,*

1. *[Mas02]* $R^{\eta,\|}(\mathbf{p}) \leq \log(\frac{1}{1-\eta})\mathbf{eff}(\mathbf{p}) \log(\#(\mathcal{A} \times \mathcal{B}))$
2. $R^{*,\eta,\|}(\mathbf{p}) \leq \log(\frac{1}{1-\eta})\mathbf{eff}^*(\mathbf{p}) \log(\#(\mathcal{A} \times \mathcal{B}))$
3. $R^{*,\eta}(\mathbf{p}) \leq O\left(\sqrt{\log(\frac{1}{1-\eta})\mathbf{eff}^*(\mathbf{p})}\right)$

*Proof.* For the first item, let $P$ be a zero-error, zero-communication protocol with shared randomness for $\mathbf{p}$ which has efficiency $\zeta = \frac{1}{\mathbf{eff}(\mathbf{p})}$. Alice and Bob run the protocol $N = \lceil \log(\frac{1}{1-\eta})\frac{1}{\zeta} \rceil$ times and send their outcome to the referee in each run. If the referee finds a valid run (where neither player aborts), he produces the corresponding outputs; otherwise he aborts. Since each run has a probability $\zeta$ of producing a valid run, the probability that the referee aborts is $(1-\zeta)^N \leq e^{-\zeta N} \leq 1 - \eta$.

For the second item, the proof is the same but the players share entanglement to run the protocol with shared entanglement and efficiency $\frac{1}{\mathbf{eff}^*(\mathbf{p})}$.

If multiple rounds of communication are allowed, then a quadratic speedup is possible in the quantum case by using a protocol for disjointness [BCW98, HdW02, AA05] on the input $u, v$ of length $N$, where $u_i$ is 0 if Alice aborts in the $i$th run and 1 otherwise, similarly for $v$ with Bob. □

For one-way communication complexity, the quantum partition bound is tight, up to arbitrarily small inefficiency. We give the results for quantum communication since the rectangle bound is already known to be tight for randomized communication complexity [JKN08].

**Theorem 15.** *For any distribution* $\mathbf{p}$ *and efficiency* $\eta < 1$*,* $Q_0^{*,\eta,\rightarrow}(\mathbf{p}) \leq \frac{1}{2}\log(\mathbf{eff}^{*,\rightarrow}(\mathbf{p})) + \log\log(1/(1-\eta))$*.*

*Proof.* Let $(\zeta, \mathbf{q})$ be an optimal solution for $\mathbf{eff}^{*,\rightarrow}(\mathbf{p})$. For any $x, y$, if we sample $a, b$ according to $\mathbf{q}$, $\Pr_{\mathbf{q}}[a \neq \perp | x] = \zeta$ and $\Pr_{\mathbf{q}}[a, b | x, y] = \zeta p(a, b|x, y)$ for all $a, b \neq \perp$ and all $x, y$. Let Alice and Bob simulate this quantum distribution $N = \lceil \log(\frac{1}{1-\eta})\frac{1}{\zeta} \rceil$ times, keeping a record of the outputs $(a_i, b_i)$ for $i \in [N]$. Since this distribution is quantum, this requires no communication (only shared entanglement). Alice then

communicates an index $i \in [N]$ such that $a_i \neq \perp$, if such an index exists, or just a random index if $a_i = \perp$ for all $i \in [N]$. Alice and Bob output $(a_i, b_i)$ corresponding to this index.

The correctness of the protocol follows from the fact that $\Pr_{\mathbf{q}}[a_i = \perp(\forall i)] = (1-\zeta)^N \leq e^{-\zeta N} \leq 1 - \eta$. The protocol then requires $\log N = -\log \zeta + \log\log(\frac{1}{1-\eta})$ bits of classical communication. Using superdense coding, this can be replaced by $\frac{1}{2}\log N$ qubits of quantum communication. □

Finally, we show that $R_0^{\eta,\rightarrow}$ depends on $\eta$ by at most an additive constant. The same is also true in the quantum model.

**Lemma 16.** *For any distribution* $\mathbf{p}$ *and efficiencies* $0 < \eta \leq \eta' < 1$, $R_0^{\eta,\rightarrow}(\mathbf{p}) \leq R_0^{\eta',\rightarrow}(\mathbf{p}) \leq R_0^{\eta,\rightarrow}(\mathbf{p}) - \log\eta + \log\log(1/(1-\eta'))$.

*Proof (sketch).* The proof is as above, except that we start from a protocol for $\mathbf{p}$ with efficiency $\eta$ instead of a quantum distribution. Note that Alice only needs to send the communication corresponding to the original protocol for the successful attempt. □

# 6 Conclusion and open problems

There are many questions to explore. In experimental setups, in particular with optics, one is faced with the very real problem that in most runs of an experiment, no outcome is recorded. The frequency with which apparatus don't yield an outcome is called detector inefficiency. Can we find explicit Bell inequalities for quantum distributions that are very resistant to detector inefficiency? For experimental purposes, it is also important for the distribution to be feasible to implement. One way to achieve this could be to prove stronger bounds for the inequalities based on the GHZ paradox given by Buhrman *et al.* [BHMR06]. Their analysis is based on a tradeoff derived from the rectangle bound. It may be possible to give sharper bounds with our techniques. Another is to consider asymmetric Bell inequalities and dimension witnesses [BPA+08, VPB10]. Here, Alice prepares a state and Bob makes a measurement. The goal is to have a Bell inequality demonstrating that Alice's system has to be large. The dimension is exponential in the size of Alice's message to Bob, so proving a lower bound on one-way communication complexity gives a lower bound on the dimension. In order to close the detection loophole, one can also consider more realistic models of inefficiency, where the failure to produce a measurement outcome is the result of either the entangled state not being produced, or the detector of each player failing independently. This could be exploited by defining a stronger version of the partition/efficiency bound that also takes into account the probabilities of events where only one of the players produces a valid outcome. While such a variation of the efficiency bound is meaningful for Bell tests, we have not considered it here as it might not be a lower bound on communication complexity.

A family of lower bound techniques still not subsumed by the efficiency bound are the information theoretic bounds such as information complexity [CSWY01]. It was recently shown that information complexity is an upper bound on discrepancy [BW11], and this upper bound was subsequently extended to a relaxation of the partition bound [KLL+12]. This *relaxed* partition bound also subsumes most algebraic and combinatorial lower bound techniques, with the notable exception of the partition bound itself, and we would therefore like to see connections one way or the other between information complexity and the partition bound.

Finally, the quantum partition bound is of particular interest. It is hard to apply since it is not linear, and it amounts to finding a Tsirelson inequality, a harder task to be sure than finding a good Bell inequality, that can nevertheless be approached via semidefinite programming relaxations [NPA08, DLTW08]. On the other hand, it is a very strong bound and one can hope to get a better upper bound on quantum communication complexity. Finding tight bounds complexity would be an important step to proving the existence, or not, of exponential gaps for total boolean functions.

# 7 Acknowledgements

We wish to particularly thank Ronald de Wolf, Raghav Kulkarni and Iordanis Kerenidis for many fruitful discussions. Research funded in part by the EU grants QCS, QAlgo, ANR Jeune Chercheur CRYQ, ANR Blanc QRAC and EU ANR Chist-ERA DIQIP. J.R. acknowledges support from the action Mandats de Retour of the Politique Scientifique Fédérale Belge, and the Belgian ARC project COPHYMA.

# References

[AA05] S. Aaronson and A. Ambainis. Quantum search of spatial regions. *Theory of Computing*, 1:47–79, 2005. arXiv:quant-ph/0303041, doi:10.4086/toc.2005.v001a004.

[BCT99] G. Brassard, R. Cleve, and A. Tapp. Cost of exactly simulating quantum entanglement with classical communication. *Phys. Rev. Lett.*, 83:1874–1877, 1999. arXiv:quant-ph/9901035, doi:10.1103/PhysRevLett.83.1874.

[BCW98] H. Buhrman, R. Cleve, and A. Wigderson. Quantum vs classical communication and computation. In *Proc. 30th STOC*, pages 63–68, 1998. arXiv:quant-ph/9802040, doi:10.1145/276698.276713.

[BCWdW01] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf. Quantum fingerprinting. *Phys. Rev. Lett.*, 87(16):167902, 2001. doi:10.1103/PhysRevLett.87.167902.

[BdW01] Harry Buhrman and Ronald de Wolf. Communication complexity lower bounds by polynomials. In *Annual Conference on Structure in Complexity Theory*, 2001. URL: citeseer.ist.psu.edu/buhrman99communication.html.

[Bel64] J. S. Bell. On the Einstein Podolsky Rosen paradox. *Physics*, 1:195, 1964.

[BHMR03] H. Buhrman, P. Høyer, S. Massar, and H. Röhrig. Combinatorics and quantum nonlocality. *Phys. Rev. Lett.*, 91:048301, 2003. arXiv:quant-ph/0209052, doi:10.1103/PhysRevLett.91.047903.

[BHMR06] H. Buhrman, P. Høyer, S. Massar, and H. Röhrig. Multipartite nonlocal quantum correlations resistant to imperfections. *Phys. Rev. A*, 73:012321, 2006. doi:10.1103/PhysRevA.73.012321.

[BNS89] L. Babai, N. Nisan, and M. Szegedy. Multiparty protocols and logspace-hard pseudorandom sequences. In *Proc. 21st STOC*, pages 1–11, 1989. doi:10.1145/73007.73008.

[BPA+08] N. Brunner, S. Pironio, A. Acín, N. Gisin, A. Méthot, and V. Scarani. Testing the dimension of Hilbert spaces. *Phys. Rev. Lett.*, 100:210503, 2008. arXiv:0802.0760, doi:10.1103/PhysRevLett.100.210503.

[BRSdW11] H. Buhrman, O. Regev, G. Scarpa, and R. de Wolf. Near-optimal and explicit Bell inequality violations. In *Proc. 26th CCC*, pages 157–166, 2011. arXiv:1012.5043, doi:10.1109/CCC.2011.30.

[BW11] M. Braverman and O. Weinstein. A discrepancy lower bound for information complexity. Technical Report 12-164, ECCC, 2011. arXiv:1112.2000.

[BYJK08] Z. Bar-Yossef, T.S. Jayram, and I. Kerenidis. Exponential separation of quantum and classical one-way communication complexity. *SIAM J. Comput.*, 38(1):366–384, 2008. doi:10.1145/1007352.1007379.

[CG85] B. Chor and O. Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. In *Proc. 26th FOCS*, pages 429–442, 1985. doi:10.1109/SFCS.1985.62.

[CSWY01] A. Chakrabarti, Y. Shi, A. Wirth, and A. Yao. Informational complexity and the direct sum problem for simultaneous message complexity. In *Proc. 42nd FOCS*, pages 270–278, 2001. doi:10.1109/SFCS.2001.959901.

[dGdW02] M. de Graaf and R. de Wolf. On quantum versions of the Yao principle. In *Proc. 19th STACS*, pages 347–358, 2002. doi:10.1007/3-540-45841-7_28.

[DKLR11] J. Degorre, M. Kaplan, S. Laplante, and J. Roland. The communication complexity of non-signaling distributions. *Quantum Information and Computation*, 11(7 –8):649–676, 2011. arXiv:0804.4859.

[DLTW08] A. C. Doherty, Y.-C. Liang, B. Toner, and S. Wehner. The quantum moment problem and bounds on entangled multi-prover games. In *Proc. 23rd CCC*, pages 199–210, 2008. arXiv:0803.4373, doi:10.1109/CCC.2008.26.

[dW08] R. de Wolf. A brief introduction to Fourier analysis on the boolean cube. *Theory of Computing Library–Graduate Surveys*, 1:1–20, 2008. doi:10.4086/toc.gs.2008.001.

[GG99] B. Gisin and N. Gisin. A local hidden variable model of quantum correlation exploiting the detection loophole. *Phys. Lett. A*, 260:323–327, 1999. arXiv:quant-ph/9905018, doi:10.1016/S0375-9601(99)00519-8.

[GKK+08] D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. de Wolf. Exponential separation for one-way quantum communication complexity, with applications to cryptography. *SIAM J. Comput.*, 38(5):1695–1708, 2008. doi:10.1145/1250790.1250866.

[GL13] D. Gavinsky and S. Lovett. En route to the log-rank conjecture: new reductions and equivalent formulations. Technical Report TR13-080, ECCC, 2013. URL: http://eccc.hpi-web.de/report/2013/080/.

[HdW02] P. Høyer and R. de Wolf. Improved quantum communication complexity bounds for disjointness and equality. In *Proc. 19th STACS*, pages 299–310, 2002. doi:10.1007/3-540-45841-7_24.

[JK10] R. Jain and H. Klauck. The partition bound for classical complexity and query complexity. In *Proc. 25th CCC*, pages 247–258, 2010. arXiv:0910.4266, doi:10.1109/CCC.2010.31.

[JKN08] R. Jain, H. Klauck, and A. Nayak. Direct product theorems for communication complexity via subdistribution bounds. In *Proc. 40th STOC*, pages 599–608, 2008. doi:10.1145/1374376.1374462.

[JP11] Marius Junge and Carlos Palazuelos. Large Violation of Bell Inequalities with Low Entanglement. *Communications in Mathematical Physics*, 306(3):695–746, 2011. arXiv:1007.3043, doi:10.1007/s00220-011-1296-8.

[JPPG+10] Marius Junge, Carlos Palazuelos, D. Pérez-García, Ignacio Villanueva, and Michael M. Wolf. Operator Space Theory: A Natural Framework for Bell Inequalities. *Phys. Rev. Lett.*, 104(17):170405, 2010. arXiv:0912.1941, doi:10.1103/PhysRevLett.104.170405.

[JY12] R. Jain and P. Yao. A strong direct product theorem in terms of the smooth rectangle bound. Technical report, 2012. URL: http://arxiv.org/abs/1209.0263, arXiv:1209.0263.

[KKN95] M. Karchmer, E. Kushilevitz, and N. Nisan. Fractional covers and communication complexity. *SIAM J. Discrete Math.*, 8(1):76–92, 1995. doi:10.1109/SCT.1992.215401.

[KLL+12] I. Kerenidis, S. Laplante, V. Lerays, J. Roland, and D. Xiao. Lower bounds on information complexity via zero-communication protocols and applications. Technical Report 12-038, ECCC, 2012. arXiv:1204.1505.

[KR11] B. Klartag and O. Regev. Quantum one-way communication can be exponentially stronger than classical communication. In *Proc. 43rd STOC*, pages 31–40, 2011. arXiv:1009.3640, doi:10.1145/1993636.1993642.

[Kra96] M. Krause. Geometric arguments yield better bounds for threshold circuits and distributed computing. *Theoretical Computer Science*, 156:99–117, 1996.

[Lov90] L. Lovász. *Communication Complexity: a Survey, in: Paths, Flows, and VLSI Layout.* Springer, B.H. Korte edition, 1990.

[LS09a] T. Lee and A. Shraibman. Lower bounds in communication complexity. *Foundations and Trends in Theoretical Computer Science*, 3(4):263–399, 2009. doi:10.1561/0400000040.

[LS09b] N. Linial and A. Shraibman. Lower bounds in communication complexity based on factorization norms. *Random Structures and Algorithms*, 34(3):368–394, 2009. doi:10.1002/rsa.20232.

[Mas02] S. Massar. Non locality, closing the detection loophole and communication complexity. *Phys. Rev. A*, 65:032121, 2002. arXiv:quant-ph/0109008, doi:10.1103/PhysRevA.65.032121.

[New91] I. Newman. Private vs. common random bits in communication complexity. *Information Processing Letters*, 39(2):61–71, 1991. doi:10.1016/0020-0190(91)90157-D.

[NPA08] M. Navascués, S. Pironio, and A. Acín. A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. *New Journal of Physics*, 10(7):073013, 2008. arXiv:0803.4290, doi:10.1088/1367-2630/10/7/073013.

[NS96] I. Newman and M. Szegedy. Public vs. private coin flips in one round communication games. In *Proc. 28th STOC*, pages 561–570, 1996. doi:10.1145/237814.238004.

[VPB10] T. Vértesi, S. Pironio, and N. Brunner. Closing the detection loophole in Bell experiments using qudits. *Phys. Rev. Lett.*, 104:060401, 2010. arXiv:0909.3171, doi:10.1103/PhysRevLett.104.060401.

[Yao83] A. C. Yao. Lower bounds by probabilistic arguments. In *Proc. 24th FOCS*, pages 420–428, 1983. doi:10.1109/SFCS.1983.30.

# A Proof of Theorem 2

We give the proof that for any distribution $\mathbf{p}$, $R^{\eta}_{\epsilon}(\mathbf{p}) \geq \log(\mathrm{prt}^{\eta}_{\epsilon}(\mathbf{p}))$.

*Proof of Theorem 2.* Let $\mathcal{P}$ be a protocol that simulates $\mathbf{p}$ with $\eta$ detector efficiency and $c$ bits of communication in the worst case, up to $\epsilon$ error in total variation distance. Let $\mathbf{p}'$ be the distribution produced by $\mathcal{P}$. We can think of $\mathcal{P}$ as a probability distribution over fully deterministic protocols $\{P_i\}$, where $P_i$ is chosen with probability $q(i)$. Each deterministic protocol $P_i$ further decomposes into $2^c$ rectangles $\{R_{i,j}\}$ and in each rectangle, the players apply a local strategy $\ell_{i,j}$ defined over inputs in $R$ and outputs in $\mathcal{A} \cup \{\perp\} \times \mathcal{B} \cup \{\perp\}$.

From this we construct a feasible solution to the linear program for $\mathrm{prt}^{\eta}(\mathbf{p}')$. For any rectangle $R \subseteq X \times Y$ and any local distribution $l$ defined over inputs $R$ and outputs in $\mathcal{A} \cup \{\perp\} \times \mathcal{B} \cup \{\perp\}$, we set

$$w_{R,\ell} = \sum_{i,j : R = R_{i,j} \text{ and } \ell = \ell_{i,j}} q(i).$$

Intuitively, $w_{R,\ell}$ is the probability of finding rectangle $R$ paired together with local strategy $\ell$ when choosing a deterministic protocol from $\mathcal{P}$. Each pair $R, l$ might appear in several of the deterministic protocols in $\mathcal{P}$, so we take the sum of the probabilities where this pair occurs.

First we claim that the objective function is $2^c$.

$$\begin{aligned}\sum_{R,\ell} w_{R,\ell} &= \sum_{R,\ell} \sum_{i,j:R=R_{i,j} \text{ and } \ell=\ell_{i,j}} q(i) \\ &= \sum_{i,j} \sum_{R_{i,j},\ell_{i,j}} q(i) \\ &= 2^c \sum_i q(i) \\ &= 2^c.\end{aligned}$$

Now, we claim that all the constraints are verified. For the first constraint, fix any $a, b, x, y \in \mathcal{A}\times\mathcal{B}\times\mathcal{X}\times\mathcal{Y}$. By assumption, $\mathcal{P}$ outputs according to $p'(a, b|x, y)$, conditioned on having output a value in $A \times B$. Let us explicitly calculate the (unconditional) probability that $\mathcal{P}$ outputs $a, b$ on input $x, y$. With probability $w_{R,\ell}$, $\mathcal{P}$ outputs according to the local strategy $l$ applied on a rectangle $R$ containing $x, y$. So the probability of outputting $a, b$ is $\sum_{R:x,y\in R,\ell} w_{R,\ell} \cdot l(a, b|x, y)$. The conditional probability is obtained by dividing by the probability of outputting some $a', b' \in \mathcal{A} \times \mathcal{B}$ on input $x, y$. This is precisely the quantity $\eta_{x,y}$.

The second constraint follows from the efficiency of $\mathcal{P}$. This completes the proof.

□

# B Efficiency bound for protocols with bounded efficiency

In order to prove lower bounds on simulating $\mathbf{p}$ with efficiency $\eta < 1$, we define the following generalization of $\mathbf{eff}(\mathbf{p})$.

**Definition 9.** *For any distribution $\mathbf{p}$ with inputs in $X \times Y$ and outputs $A \times B$, define $\mathbf{eff}^\eta(\mathbf{p}) = 1/\zeta_{\text{opt}}$, where $\zeta_{\text{opt}}$ is the optimal value of the following linear program. The variables are $\zeta, \zeta_{xy}$ and $q_\ell$, where $\ell$ ranges over all local deterministic protocols with inputs taken from $\mathcal{X} \times \mathcal{Y}$ and outputs in $\mathcal{A}\cup\{\perp\}\times\mathcal{B}\cup\{\perp\}$.*

$$\begin{aligned}\zeta_{\text{opt}} = \max_{\zeta,\zeta_{xy},q_\ell\geq 0} \quad & \zeta \\ \text{subject to} \quad & \sum_{\ell\in\mathcal{L}^\perp_{\text{det}}} q_\ell l(a, b|x, y) = \zeta_{xy} p(a, b|x, y) && \forall x, y, a, b \in \mathcal{X}\times\mathcal{Y}\times\mathcal{A}\times\mathcal{B} \\ & \sum_{\ell\in\mathcal{L}^\perp_{\text{det}}} q_\ell = 1 \\ & \eta\zeta \leq \zeta_{xy} \leq \zeta && \forall x, y \in \mathcal{X}\times\mathcal{Y}.\end{aligned}$$

*For randomized communication with error, we define $\mathbf{eff}^\eta_\epsilon(\mathbf{p}) = \min_{|p'-p|_1\leq\epsilon} \mathbf{eff}^\eta(\mathbf{p}')$.*

This provides a lower bound for $R^\eta_\epsilon(\mathbf{p})$, which is equivalent to the lower bound obtained from $\text{prt}^\eta_\epsilon(\mathbf{p})$ (we omit the proofs of these statements as they closely follow the lines of the special case $\eta = 1$).

**Lemma 17.** *For any distribution $\mathbf{p}$, we have $R^\eta_\epsilon(\mathbf{p}) \geq \log \mathbf{eff}^\eta_\epsilon(\mathbf{p})$.*

**Theorem 18.** *For any distribution $\mathbf{p}$, $\mathbf{eff}^\eta_\epsilon(\mathbf{p}) = \text{prt}^\eta_\epsilon(\mathbf{p})$.*

We can also study the maximum $\eta$ such that $\mathbf{p}$ can be simulated with efficiency $\eta_{xy} \geq \eta$ on input $x, y$, without any communication. We denote the inverse of this quantity by $\text{eff}^{\text{nc}}(\mathbf{p})$.

**Definition 10.** *For any distribution $\mathbf{p}$ with inputs in $X \times Y$ and outputs $A \times B$, define $\text{eff}^{\text{nc}}(\mathbf{p}) = 1/\eta$, where $\eta$ is the maximum $\eta$ such that $R^\eta(\mathbf{p}) = 0$.*

This quantity can be seen as a relaxation of $\mathbf{eff}(\mathbf{p})$, where we no longer require the inefficiency to be the same for all inputs. Indeed, it can be rewritten as follows.

**Lemma 19.** *For any distribution* $\mathbf{p}$*, we have* $\mathrm{eff}^{\mathrm{nc}}(\mathbf{p}) = 1/\zeta_{\mathrm{opt}}$*, where* $\zeta_{\mathrm{opt}}$ *is the optimal value of the following linear program.*

$$\begin{aligned}
\zeta_{\mathrm{opt}} = \max_{\zeta,\zeta_{xy},q_\ell \geq 0} \quad & \zeta \\
\text{subject to} \quad & \sum_{\ell \in \mathcal{L}_{\mathrm{det}}^{\perp}} q_\ell l(a,b|x,y) = \zeta_{xy} p(a,b|x,y) && \forall x,y,a,b \in \mathcal{X}\times\mathcal{Y}\times\mathcal{A}\times\mathcal{B} \\
& \sum_{\ell \in \mathcal{L}_{\mathrm{det}}^{\perp}} q_\ell = 1 \\
& \zeta \leq \zeta_{xy} && \forall x,y \in \mathcal{X}\times\mathcal{Y}.
\end{aligned}$$

By comparing the linear programs for the different quantities, we immediately obtain the following relations:

**Lemma 20.** *For any distribution* $\mathbf{p}$*, we have* $\eta \cdot \mathrm{eff}^{\mathrm{nc}}(\mathbf{p}) \leq \mathbf{eff}^{\eta}(\mathbf{p}) \leq \eta \cdot \mathbf{eff}(\mathbf{p})$.

# C Lower bound for a Hidden Matching distribution

We first recall an application of KKL inequality as explained in [dW08] which we use in the proof.

**Lemma 21.** *Let* $A$ *be a subset of* $\{0,1\}^n$*. Let* $S$ *be a subset of* $\{1 \dots n\}$*. We define* $\beta_S = \mathbb{E}_{x \in A}\left((-1)^{S \cdot x}\right)$ *where* $S \cdot x = \sum_{i \in S} x_i$*. Let* $\mathcal{S}_2$ *be the set of subsets of* $\{1 \dots n\}$ *of size 2. There exists an absolute constant* $\mathcal{C}$ *such that*

$$\sum_{S \in \mathcal{S}_2} \beta_S^2 \leq \mathcal{C} \log\left(\frac{2^n}{|A|}\right)^2.$$

We now prove the theorem.

*Proof of Theorem 12.* Let $\mathbf{p}'$ be such that $|\mathbf{p}' - \mathrm{HM}|_1 \leq \epsilon$. We lower bound $\mathbf{eff}^{\rightarrow}(\mathbf{p}')$ using the dual of $\mathbf{eff}^{\rightarrow}$ (Lemma 8).

$$\begin{aligned}
\mathbf{eff}^{\rightarrow}(\mathbf{p}') = \max_{B_{x,M,a,d,i,j}} \quad & \sum_{x,M,a,d,i,j} B_{x,M,a,d,i,j} \cdot p'(a,d,i,j|x,M) \\
\text{subject to} \quad & \sum_{x \in X(\ell),M,a,d,i,j} B_{x,M,a,d,i,j} \cdot l(a,d,i,j|x,M) \leq 1 && \forall \ell \in \mathcal{L}_{\mathrm{det}}^{\perp A},
\end{aligned}$$

where we let $X(\ell)$ be the set of inputs for which Alice does not abort when following the local deterministic strategy $\ell$.

We exhibit coefficients that satisfy the constraints and give us a good lower bound for the objective function for each $\mathbf{p}'$ close to HM.

To give an upper bound on the Bell value of any local deterministic strategy that may output $\perp$, we will use the fact that such a strategy leads to a partition of Alice's inputs, where she doesn't abort, into rectangles. We will show an upper bound on the bias of each rectangle using the analysis from [BRSdW11]. However, in their analysis, the Bell value of the local strategy depends on the size of the rectangle, which will result in a poor upper bound. We will need to consider two different cases. If the rectangle is small enough, then we obtain a sufficiently good upper bound as is. If the rectangle is too big, we will need to subtract from the coefficients some constant that we will call $\mu$. Notice that in $\mathbf{eff}$, the constraint is that

the Bell value of any local deterministic strategy is less than 1, but the absolute value is not bounded as in $\nu$. This is why we can subtract without violating the constraint. The overall weight of those $\mu$ will not significantly affect the Bell value of a distribution close to the Hidden Matching distribution so the objective value will remain large.

Consider the following coefficients to the Bell functional.

$$B_{x,M,a,d,i,j} = \Phi'_{x,M,a,d,i,j} + \mu_{x,M},$$

where

$$\mu_{x,M} = -\frac{2^{\frac{\sqrt{n-1}}{2\mathcal{C}}}}{n2^{n+1}|\mathcal{M}_n|}$$
$$\Phi'_{x,M,a,d,i,j} = \frac{2^{\frac{\sqrt{n-1}}{2\mathcal{C}}}}{n2^n|\mathcal{M}_n|}\delta_{(i,j)\in M}\cdot(-1)^{\langle a,i\oplus j\rangle\oplus d\oplus x_i\oplus x_j}$$

where $\delta$ is the Kronecker function, and $\mathcal{M}_n$ is the set of matchings over edges $\{1,\dots n\}$.

**Verifying the constraints.** Let $\ell \in \mathcal{L}_{det}^{\perp}$ and $X = X(\ell)$, the set of inputs for which Alice does not abort when following the local deterministic strategy $\ell$. The strategy $\ell$ partitions the set $X$ into $\bigcup_a X_a$ where Alice outputs $a$, and $\mathcal{M}_n$ into $\bigcup_{d,i,j} R_{d,i,j}$ where Bob outputs $(d,i,j)$ because $\ell$ is local and deterministic.

First, we want to bound from above the value:

$$\sum_{x\in X,M,a,d,i,j} B_{x,M,a,d,i,j}\cdot l(a,d,i,j|x,M) = \sum_a \sum_{x\in X_a}\sum_{i,j,d}\sum_{M\in R_{i,j,d}} B_{x,M,a,d,i,j}.$$

We bound each term of the sum, for fixed $a$.

Let us first see what happens on small rectangles that is, when $X_a$ is small.

**Claim 1.** *If* $|X_a| \le 2^{n-\frac{\sqrt{n-1}}{2\mathcal{C}}}$ *then* $\sum_{x\in X_a}\sum_{i,j,d}\sum_{M\in R_{i,j,d}} B_{x,M,a,d,i,j} \le \frac{1}{n}$.

*Proof.* Since the $\mu_{x,M}$ are negative,

$$\begin{aligned}
\sum_{x\in X_a}\sum_{i,j,d}\sum_{M\in R_{i,j,d}} B_{x,M,a,d,i,j} &= \sum_{x\in X_a}\sum_{i,j,d}\sum_{M\in R_{i,j,d}} \mu_{x,M} + \sum_{x\in X_a}\sum_{i,j,d}\sum_{M\in R_{i,j,d}} \Phi'_{x,M,a,d,i,j} \\
&\le \sum_{x\in X_a}\sum_{i,j,d}\sum_{M\in R_{i,j,d}} \Phi'_{x,M,a,d,i,j} \\
&= \frac{2^{\frac{\sqrt{n-1}}{2\mathcal{C}}}}{n2^n}\sum_{x\in X_a, M\in\mathcal{M}_n}\left(\sum_{d,i,j|l(a,d,i,j|x,M)=1}\frac{(-1)^{x_i\oplus x_j\oplus d\oplus\langle a,i\oplus j\rangle}}{|\mathcal{M}_n|}\delta_{(i,j)\in M}\right) \\
&\le \frac{2^{\frac{\sqrt{n-1}}{2\mathcal{C}}}}{n2^n}|X_a|,
\end{aligned}$$

where we have used the fact that there is exactly one tuple $(d,i,j)$ such that $l(a,d,i,j|x,M)=1$, because $l$ is deterministic and Bob doesn't abort.
Since $|X_a| \le 2^{n-\frac{\sqrt{n-1}}{2\mathcal{C}}}$ then this sum is less than $\frac{1}{n}$. □

Now let us consider the case of the large rectangles.

**Claim 2.** *If* $|X_a| \ge 2^{n-\frac{\sqrt{n-1}}{2\mathcal{C}}}$ *then* $\sum_{x\in X_a}\sum_{i,j,d}\sum_{M\in R_{i,j,d}} B_{x,M,a,d,i,j} \le 0$.

*Proof.*

$$\begin{aligned}\sum_{x\in X_a}\sum_{i,j,d}\sum_{M\in R_{i,j,d}} B_{x,M,a,d,i,j} &= \sum_{x\in X_a,M}\mu_{x,M} + \sum_{x\in X_a}\sum_{i,j,d}\sum_{M\in R_{i,j,d}}\Phi'_{x,M,a,d,i,j}\\ &= -\frac{2^{\frac{\sqrt{n-1}}{2\mathcal{C}}}}{n2^{n+1}}|X_a| + \sum_{x\in X_a}\sum_{i,j,d}\sum_{M\in R_{i,j,d}}\Phi'_{x,M,a,d,i,j}.\end{aligned}$$

Let $\beta^a_{i,j} = \mathbb{E}_{x\in X_a}((-1)^{x_i\oplus x_j})$ and $q^a_{i,j} = \sum_d \sum_{M\in R_{d,i,j}} \frac{(-1)^{\langle a,i\oplus j\rangle\oplus d}}{|\mathcal{M}_n|}\delta_{(i,j)\in M}$. Then

$$\begin{aligned}\sum_{x\in X_a}\sum_{i,j,d}\sum_{M\in R_{i,j,d}}\Phi'_{x,M,a,d,i,j} &= \frac{2^{\frac{\sqrt{n-1}}{2\mathcal{C}}}}{n2^n}\sum_{i,j}\sum_{x\in X_a}\sum_d\sum_{M\in R_{d,i,j}}\frac{(-1)^{x_i\oplus x_j\oplus d\oplus\langle a,i\oplus j\rangle}}{|\mathcal{M}_n|}\delta_{(i,j)\in M}\\ &= \frac{2^{\frac{\sqrt{n-1}}{2\mathcal{C}}}}{n2^n}\sum_{i,j}|X_a|\beta^a_{i,j}\left(\sum_d\sum_{M\in R_{d,i,j}}\frac{(-1)^{\langle a,i\oplus j\rangle\oplus d}}{|\mathcal{M}_n|}\delta_{(i,j)\in M}\right)\\ &\le \frac{2^{\frac{\sqrt{n-1}}{2\mathcal{C}}}}{n2^n}|X_a|\sqrt{\sum_{i,j}|\beta^a_{i,j}|^2}\sqrt{\sum_{i,j}|q^a_{i,j}|^2}.\end{aligned}$$

The last line follows from the Cauchy-Schwarz inequality.

On one hand,

$$|q^a_{i,j}| \le \sum_d\sum_{M\in R_{d,i,j}}\frac{\delta_{(i,j)\in M}}{|\mathcal{M}_n|} = \mathrm{Prob}_{M\in\mathcal{M}_n}(l \text{ outputs}(i,j)\in M) \le \frac{1}{n-1},$$

and $\sum_{i,j}|q^a_{i,j}| \le 1$, so $\sqrt{\sum_{i,j}|q^a_{i,j}|^2} \le \frac{1}{\sqrt{n-1}}$. On the other hand, the KKL inequality gives us (with $A = X_a$):

$$\sqrt{\sum_{i,j}|\beta^a_{i,j}|^2} \le \mathcal{C}\times\log\left(\frac{2^n}{|X_a|}\right).$$

Hence,

$$\begin{aligned}\sum_{x\in X_a}\sum_{i,j,d}\sum_{M\in R_{i,j,d}}\Phi'_{x,M,a,d,i,j} &\le \frac{2^{\frac{\sqrt{n-1}}{2\mathcal{C}}}}{n2^n}|X_a|\mathcal{C}\log\left(\frac{2^n}{|X_a|}\right)\times\frac{1}{\sqrt{n-1}}\\ &\le \frac{2^{\frac{\sqrt{n-1}}{2\mathcal{C}}}}{n2^{n+1}}|X_a|,\end{aligned}$$

because $|X_a| \ge 2^{n-\frac{\sqrt{n-1}}{2\mathcal{C}}}$ implies that $\mathcal{C}\log\left(\frac{2^n}{|X_a|}\right)\frac{1}{\sqrt{n-1}} \le \frac{1}{2}$. □

From Claims 1 and 2, we obtain:

$$\sum_{x\in X,M,a,d,i,j} B_{x,M,a,d,i,j}\cdot l(a,d,i,j,|,x,M) \le \sum_{a||X_a|\le 2^{n-\frac{\sqrt{n-1}}{2\mathcal{C}}}}\frac{1}{n} \le 1$$

**Value of the objective function.** Let $\mathbf{p}'$ be a distribution such that $|\mathbf{p}'-\mathrm{HM}|_1 \leq \epsilon$. For any $x, M, a, d, i, j$, we define $\epsilon_{x,M,a,d,i,j} = |p'(a,d,i,j|x,M) - \mathrm{HM}(a,d,i,j|x,M)|$ and for any $x, M$, we have

$$\sum_{a,d,i,j} \epsilon_{x,M,a,d,i,j} \leq \epsilon.$$

We want to lower bound

$$\sum_{x,M,a,d,i,j} B_{x,M,a,d,i,j} \cdot p'(a,d,i,j,|x,M).$$

Recall that we have set $B_{x,M,a,d,i,j} = \Phi'_{x,M,a,d,i,j} + \mu_{x,M}$. We will consider the two terms separately. Since $\mathbf{p}'$ is a distribution,

$$\sum_{x,M,a,d,i,j} \mu_{x,M} \cdot p'(a,d,i,j|x,M) = \sum_{x,M} \mu_{x,M} = -\frac{2^{\frac{\sqrt{n-1}}{2\mathcal{C}}}}{n2^{n+1}} 2^n = -\frac{2^{\frac{\sqrt{n-1}}{2\mathcal{C}}}}{2n}$$

We also have

$$\begin{aligned}
&\sum_{x,M,a,d,i,j} \Phi'_{x,M,a,d,i,j} \cdot p'(a,d,i,j,|x,M) \\
&\geq \frac{2^{\frac{\sqrt{n-1}}{2\mathcal{C}}}}{n2^n|\mathcal{M}_n|} \sum_{x,M} \left( \sum_{a,d,i,j:x_i \oplus x_j = d \oplus \langle a, i \oplus j \rangle} \delta_{(i,j)\in M} (\mathrm{HM}(a,d,i,j|x,M) - \epsilon_{x,M,a,d,i,j}) \right. \\
&\qquad \left. + \sum_{a,d,i,j:x_i \oplus x_j \neq d \oplus \langle a, i \oplus j \rangle} \delta_{(i,j)\in M} (-\mathrm{HM}(a,d,i,j|x,M) - \epsilon_{x,M,a,d,i,j}) \right) \\
&= \frac{2^{\frac{\sqrt{n-1}}{2\mathcal{C}}}}{n} - \frac{2^{\frac{\sqrt{n-1}}{2\mathcal{C}}}}{n2^n|\mathcal{M}_n|} \sum_{x,M,a,d,i,j} \epsilon_{x,M,a,d,i,j} \delta_{(i,j)\in M} \\
&\geq \frac{2^{\frac{\sqrt{n-1}}{2\mathcal{C}}}}{n} - \frac{2^{\frac{\sqrt{n-1}}{2\mathcal{C}}}}{n} \epsilon.
\end{aligned}$$

Finally we get the value of the objective function

$$\mathbf{eff}^{\rightarrow}_{\epsilon}(\mathrm{HM}) \geq \sum_{x,M,a,d,i,j} B_{x,M,a,d,i,j} \cdot p'(a,d,i,j|x,M) \geq \frac{2^{\frac{\sqrt{n-1}}{2\mathcal{C}}}}{n} (\frac{1}{2} - \epsilon).$$

□