Non-Local Box Complexity and Secure Function Evaluation

Marc Kaplan* Iordanis Kerenidis* Sophie Laplante* Jérémie Roland†

Abstract

A non-local box is an abstract device into which Alice and Bob input bits x and y respectively and receive outputs a and b, where a, b are uniformly distributed and $a \oplus b = x \wedge y$. Such boxes have been central to the study of quantum or generalized non-locality, as well as the simulation of non-signaling distributions. In this paper, we start by studying how many non-local boxes Alice and Bob need in order to compute a Boolean function f. We provide tight upper and lower bounds in terms of the communication complexity of the function both in the deterministic and randomized case. We show that non-local box complexity has interesting applications to classical cryptography, in particular to secure function evaluation, and study the question posed by Beimel and Malkin [BM04] of how many Oblivious Transfer calls Alice and Bob need in order to securely compute a function f. We show that this question is related to the non-local box complexity of the function and conclude by greatly improving their bounds. Finally, another consequence of our results is that traceless two-outcome measurements on maximally entangled states can be simulated with 3 non-local boxes, while no finite bound was previously known.

1 Introduction

Communication complexity. Communication complexity is a central model of computation, which was first defined by Yao in 1979 [Yao79]. It has found applications in many areas of theoretical computer science including Boolean circuit complexity, time-space tradeoffs, data structures, automata, formula size, etc. In this model Alice and Bob receive inputs x and y respectively and are allowed to communicate in order to compute a function $f(x, y)$. The goal is to find the minimum amount of communication needed for this task. In different variants of the model, we allow Alice and Bob to err with some probability, and to share common resources in an attempt to enable them to solve their task more efficiently.

One such resource is shared randomness. When Alice and Bob are not allowed any errors, shared randomness does not reduce the communication complexity. On the other hand, when they are allowed to err, a common random string can reduce the amount of communication needed. However, Newman's result tells us that shared randomness can be replaced by private randomness at an additional cost logarithmic in the input size [New91].

Another very powerful shared resource is entanglement. Using teleportation, Alice and Bob can transmit quantum messages by using their entanglement and only classical communication. This model has been proven to be very powerful, in some cases exponentially more efficient than the classical one. Another way to understand the power of entanglement is by looking at the CHSH game [CHSH69], where Alice and Bob receive uniformly random bits x and y respectively and their goal is to output bits a and b resp. such that $a \oplus b = x \wedge y$ without communicating. It is easy to conclude that even if Alice and Bob share randomness, their optimal strategy will be successful with probability 0.75 over the inputs. However, if they share entanglement, then there is a strategy that succeeds with probability approximately 0.85. This game proves that quantum entanglement can enable two parties to create correlations that are impossible to create with classical means.

Even though the setting of the previous game is not exactly the same as the model of communication complexity, we can easily transform one to the other. From now on, in our communication complexity model, instead of requiring Bob to output the value of the function $f(x, y)$, we require Alice and Bob to output two bits a and b respectively, such that $a \oplus b = f(x, y)$. We call this "computing f in parity". It is easy to see that the two models are equivalent up to one bit of communication.

*LRI - Université Paris-Sud
†NEC Laboratories America

Non-local boxes. As we said, entanglement enables Alice and Bob to succeed in the CHSH game with probability 0.85. But what if they shared some resource that would enable them to win the game with probability 1? Starting from such considerations, Popescu and Rohrlich [PR94, RP96, PR97] defined the notion of a non-local box. A non-local box is an abstract device shared by Alice and Bob. By one use of a non-local box, we mean that Alice inputs x, Bob inputs y, Alice gets an output a and Bob gets b where $a \oplus b = x \wedge y$, and the marginal distributions on a and b are uniform. The name non-local box is due to the property that one use of a non-local box creates correlations between two bits that are maximally non-local (allowing to win the CHSH game with probability one), but still does not allow to communicate, since taken separately, each bit is just an unbiased random coin. As such, a non-local box may be considered as a unit of non-locality. We note here an important property of a non-local box, namely that, similar to entanglement, one player can enter an input and receive an output even before the second player has entered an input.

The importance of the notion of a non-local box has become increasingly evident in the last years. Non-local boxes were first introduced to study (quantum or generalized) non-locality. In particular, it was shown than one of the most studied versions of the EPR experiment, where Alice and Bob perform projective measurements on a maximally entangled qubit pair, may be simulated using only one use of a non-local box [CGMP05]. More generally, it was shown that any non-signaling distribution over Boolean outputs may be exactly simulated with some finite number of non-local boxes (for finite input size) [BP05, JM05]. This was later generalized to any non-signaling distribution, except that the simulation may not always be performed exactly for non-Boolean outputs [FW09]. These results rely on the fact that the set of non-signaling distributions is a polytope, so it suffices to simulate the extremal vertices to be able simulate the whole set. In the context of non-locality, another application of non-local boxes is the study of pseudo-telepathy games [BM05].

It is easy to see that one use of a non-local box can be simulated with one bit of communication and shared randomness: Alice outputs a uniform bit r and sends x to Bob, who outputs $r \oplus x \cdot y$. However, the converse cannot possibly hold, since a non-local box cannot be used for communication.

The first question is what happens if we use non-local boxes as shared resource in the communication complexity model. Van Dam showed that for any Boolean function $f : \{0,1\}^n \times \{0,1\}^n \rightarrow \{0,1\}$, Alice and Bob can use 2^n non-local boxes and no communication at all and at the end output bits a and b such that $a \oplus b = f(x,y)$ [vD05]. In other words, if non-local boxes were physically implementable, then all functions would have trivial communication complexity. His results were strengthened by Brassard *et al.* who showed that even if a non-ideal non-local box existed, one that solves the CHSH game with probability 0.91, then still all functions would have trivial communication complexity [BBL$^+$06]. Note that in these results, the number of non-local boxes needed may be exponential in the input size and do not take into account any properties of the function and more precisely its communication complexity without non-local boxes. It also follows from the work of [BP05, BBL$^+$06] that for any Boolean function f, if f has a circuit with fan-in 2 (i.e. where every gate has two input bits) of size s, then there is a deterministic non-local box protocol of complexity $O(s)$, where the bits of the input of f are split arbitrarily among the players. This implies that exhibiting an explicit function for which the deterministic non-local box complexity is superlinear, would translate into a superlinear circuit lower bound for this function. This is a notoriously difficult problem, and while a simple counting argument shows that a random function requires exponential size circuits, the best lower bound to date for an explicit function is linear [LR01, IM02].

Secure function evaluation. Non-local boxes have also been studied in relation to cryptographic primitives such as Oblivious Transfer or Bit Commitment. Wolf and Wullschleger [WW05] showed that Oblivious Transfer is equivalent to a *timed* version of a non-local box (up to a factor of 2). To maintain the non-signaling property of the non-local box, one can define timed non-local box as having a predefined time limit, and if any of the players has not entered an input by this time, then some fixed input, say 0, is used instead. Subsequently, Buhrman *et al.* [BCU$^+$07] showed how to construct Bit Commitment and Oblivious Transfer by using non-local boxes that do not need to be timed but have to be trusted.

In this paper, we are interested in secure function evaluation, which is one of the most fundamental cryptographic tasks. In this model, Alice and Bob want to evaluate some function of their inputs in a way that does not leak any more information than what follows from the output of the function. It is known that there exists functions that cannot be evaluated securely in the information-theoretic setting ([BOGW88, CCD88, CK91, Kus92]). However, all functions can be computed securely in the information theoretic setting if the players have access to a black box that performs Oblivious Transfer or some other complete function, e.g. the AND function ([GV88, Kil88]).

There has been a lot of work trying to identify, in various settings, which functions can be easily evaluated in a secure way, *i.e.*, without any invocation of the black box, and which are hard to evaluate securely, *i.e.*, require at least one invocation of the black box ([CK91, Kus92, BMM99, Kil91, KKMO00, Kil00]). Moreover, Beaver [Bea96] showed that there exists a hierarchy of different degrees of hardness for the information-theoretic setting. In other words, for all k, there are functions that can be securely evaluated with k invocations of the AND box but cannot be computed with $k-1$ uses of the black box.

Beimel and Malkin [BM04] proposed a quantitative approach to secure function evaluation by studying how many calls to an Oblivious Transfer or other complete black box one needs in order to securely compute a given function f in the honest-but-curious model. For a Boolean function $f : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ and deterministic protocols, they provide a combinatorial characterization of the minimal number of AND calls required, which however does not lead to an efficient algorithm to determine how many ANDs are actually required. They also show that $2^{|\mathcal{X}|}$ ANDs are sufficient for any function. In the randomized case, they provide lower bounds depending on the truth-table of the function which can be at most of the order of n. They also state that "it would be very interesting to try and explore tighter connections with the communication complexity of the functions".

Finally, Naor and Nissim [NN01] have given some connections between the communication complexity of a function f and the communication complexity for securely computing f. These results, translated into the Beimel-Malkin and our model, only show that the number of ANDs is at most exponential in the communication complexity.

Summary of results. In this paper, we provide more evidence on the importance of non-local boxes by showing how they relate to different models of communication complexity as well as how they can be used as a tool to quantitatively study secure function evaluation.

First, we study how many non-local boxes are needed in order to distributively compute a Boolean function f. We define four different variants denoted $NL, NL_\varepsilon, NL^{||}, NL^{||}_\varepsilon$, where the first two are the deterministic and randomized non-local box complexity and the latter two are the deterministic and randomized complexity where the non-local boxes are used in a restricted manner: the non-local boxes are all used in parallel, that is, the input to any non-local box does not depend on the outcome of any other. We provide lower and upper bounds for all these models in terms of communication models and show that in many cases our bounds are tight.

For the deterministic parallel non-local box complexity, we show that $NL^{||}(f)$ is equal to the rank of the function f over $\mathbb{GF}_2$. This also implies that it is equivalent to the communication complexity of the function in the following model: Alice and Bob send to a referee one message each and the referee outputs the Inner Product of the two vectors mod 2. Moreover we show that $NL^{||}(f)$ is always greater than the deterministic communication complexity $D(f)$ and less than $2^{D(f)}$. These bounds are optimal as can be seen by looking at the functions of Inner Product and Disjointness.

In the randomized parallel case, we define a notion of approximate rank over $\mathbb{GF}_2$ which is equal to $NL^{||}_\varepsilon(f)$, under the assumption that the output of the protocol is the XOR of the outcomes of the non-local boxes. The notion of approximate rank over $\mathbb{R}$ has been used for communication complexity [Bd01] and gives upper and lower bounds in the randomized model.

For the deterministic non-local box complexity $NL(f)$, we show that it is at least the communication complexity $D(f)$ and, of course, smaller than $NL^{||}(f)$, which is again a tight bound. In the randomized case, we prove that it is bounded above by the communication complexity $R^{||,MAJ}(f)$ in the following model: Alice and Bob send to a referee one message each and the referee outputs 1 if for the majority of indices, the two messages are equal. This is a natural model of communication complexity that has appeared repeatedly, for example in the simulation of quantum protocols by classical ones and in various upper bounds on simultaneous messages [KNR99, Gro97, SZ08, LS08]. This model is also bounded above in terms of γ_2^∞, a quantity which has been used for upper and lower bounds on communication complexity [LS08].

In another application of our work, using the recent result of Regev and Toner [RT09], we show that traceless two-outcome measurements on maximally entangled states can be simulated with 3 non-local boxes. Previously, no finite bound was known for this case. In order to do this we need to extend our results from Boolean functions to any distribution.

Then, we look at the consequences of our results in the area of secure function evaluation. The main question we study is how many calls to a secure primitive one needs to make in order to securely evaluate a function f. Specifically, in the honest-but-curious model, we exactly characterize the number of secure AND boxes we need in order to evaluate f by the one-way communication complexity of f. Our proof will be reminiscent of our proofs for the non-local box complexity. In the malicious model, we upper bound the number of Oblivious Transfer boxes needed by the non-local box complexity of f, when the non-local boxes are used in order. This implies strong upper bounds in terms of the communication complexity as well as γ_2^∞. For the lower bounds, we show that the communication complexity of f remains a lower bound for OT-*optimal* protocols that securely evaluate f.

Our results show that non-local boxes, introduced for the study of quantum correlations or more general non-locality, can provide a novel way of looking at questions about classical communication complexity, secure function evaluation and complexity theory.

2 Preliminaries

2.1 Communication Complexity

Let $f : \mathcal{X} \times \mathcal{Y} \to \{0, 1\}$ be a bipartite Boolean function. We consider the following setting: Alice gets an input $x \in \mathcal{X}$ and Bob an input $y \in \mathcal{Y}$. The communication complexity of f is the number of bits that Alice and Bob have to send one to the other in order to compute the function. In the standard model, one player has to output the value of the function. Here, we consider the following variation. We say that Alice and Bob compute $f(x, y)$ in parity if after executing a communication protocol, Alice outputs a bit a and Bob outputs a bit b such that $a \oplus b = f(x, y)$, where $\oplus$ denotes both the logical XOR and the addition modulo 2. This model differs from the standard one by at most 1 bit.

We use the following notions of communication complexity. In probabilistic models, we assume that the players have a common source of randomness.

- $D(f)$ and $R_\varepsilon(f)$: deterministic and ε-bounded error communication complexity of $f(x, y)$ in parity.
- $D^\rightarrow(f)$ and $R_\varepsilon^\rightarrow(f)$: one-way deterministic and bounded-error communication complexity of $f(x, y)$ in parity.
- $D^{\|}(f)$ and $R_\varepsilon^{\|}(f)$: deterministic and bounded-error communication complexities in the model of simultaneous messages, where Alice and Bob each send a message to the referee and the referee outputs the value of the function $f(x, y)$.

For the model of simultaneous messages, we also consider some natural restrictions on how the referee computes the output from the messages he receives from the players. We assume the messages sent are of the same length. Suppose the referee receives bits $\mathbf{a} = (a_1, \ldots, a_t)$ from Alice, and $\mathbf{b} = (b_1, \ldots, b_t)$ from Bob. If the referee always computes a predefined function $g(\mathbf{a}, \mathbf{b})$, then we write $D^{\|,g}(f)$ or $R_\varepsilon^{\|,g}(f)$ to be the length of the message sent by the players (not the sum of these lengths, as is done in the standard model). In this paper, we will consider two functions, the inner product modulo 2, $IP_2(\mathbf{a}, \mathbf{b}) = \bigoplus_i (a_i \cdot b_i)$ (where $\cdot$ denotes the multiplication over $\mathbb{GF}_2$, which corresponds to the logical AND) and the majority function, $MAJ(\mathbf{a}, \mathbf{b}) = MAJ(a_1 \oplus b_1, \ldots, a_t \oplus b_t)$.

2.2 Non-local box Complexity

Definition 1 (Non-local box). *A non-local box is a device shared by two parties, which on one side takes Boolean input x and immediately produces Boolean output a, and on the other side takes Boolean input y and immediately produces Boolean output b, according to the following distribution:* $\mathbf{p}_{NL}(a, b|x, y) = \begin{cases} \frac{1}{2} & \textit{if } a \oplus b = x \cdot y \\ 0 & \textit{otherwise.} \end{cases}$

Let us stress the importance of timing in this definition. Indeed, Alice should receive her output a from the box as soon as she has entered her input x, no matter if Bob has already entered his input or not (and vice-versa). This is possible because the input-output distribution is non-signaling, that is, the marginal distribution of Alice's input a does not depend on Bob's input y, since $p(a|x, y) = 1/2$ for any a, x, y. In other words, from Alice's point of view, a is just an unbiased random bit. The reason for this definition is to mimic an EPR experiment, where Alice obtains her measurement outcome as soon as she performs her measurement, independently of whether Bob has performed his measurement or not. One could potentially consider *timed* versions of non-local boxes as well. For example, imagine that the outputs of the box are produced only after both players enter their inputs. In this case, the box becomes signaling, since a player can transmit a bit by either entering an input or not into the box. As we have said, there are also versions of non-local boxes that are both timed and non-signaling, for example a box that has a predefined time limit, and if any of the players has not entered an input by that time, then some fixed input, say 0, is used instead. In this paper, we only consider non-local boxes that are also non-signaling.

We study a model akin to communication complexity, where Alice and Bob use non-local boxes instead of communication. In a non-local box protocol, Alice and Bob wish to compute some function $f : \mathcal{X} \times \mathcal{Y} \to \{0, 1\}$. Alice gets an input $x \in \mathcal{X}$, Bob gets an input $y \in \mathcal{Y}$, and they have to compute $f(x, y)$ in parity. Recall that it means that at the end of the protocol, Alice outputs $a \in \{0, 1\}$ and Bob $b \in \{0, 1\}$, such that $a \oplus b = f(x, y)$. For a protocol P, we will write $P(x, y) = (a, b)$. In the course of the protocol, Alice and Bob are allowed shared randomness and may use non-local boxes, but they may not communicate. Bob is not allowed to see Alice's inputs to the non-local boxes, nor does he see the outcome on Alice's side, and likewise for Alice.

Definition 2. *For any function $f : \mathcal{X} \times \mathcal{Y} \to \{0, 1\}$, $NL(f)$ is the smallest t such that there is a protocol that computes f in parity exactly, using t non-local boxes.*

We will label the non-local boxes with labels from 1 to t. (Recall that in general, Alice and Bob are not required to use the t non-local boxes in the same order.) We relax the exactness condition and allow the protocol's outcome to be incorrect with constant probability ε.

Definition 3. *For any function $f : \mathcal{X} \times \mathcal{Y} \to \{0, 1\}$, $NL_\varepsilon(f)$ is the smallest t such that there is a protocol P using t non-local boxes, with $\Pr[P(x, y) = (a, b)$ with $a \oplus b = f(x, y)] \geq 1 - \varepsilon$.*

We will also study two variants of the general model, where the non-local boxes are used in a restricted manner. First, we assume that the non-local boxes are used in parallel, that is, the input to any non-local box does not depend on the outcome of any other. In this model, we denote the complexity $NL^{||}$ in the exact case, and $NL_\varepsilon^{||}$ in the ε error case.

Second, we define the model where both players use the non-local boxes in the same order, that is, the non-local boxes are labeled from 1 to t and Alice's input to the non-local box with label i does not depend on the outputs from the non-local boxes labeled from $i + 1$ to t (similarly for Bob). Note that in the most general case, Alice and Bob may use their t non-local boxes with labels 1 through t in whichever order they want. For example, Alice may use the non-local box with label 3 first, then use the output in order to compute the input for the non-local box with label 1, while Bob might use the non-local box with label 1 first and so forth. The complexity in this model is denoted NL^{ord} in the exact case, and $NL_\varepsilon^{\mathrm{ord}}$ in the ε error case. It is clear that this model is more powerful than the parallel model but less powerful than the general non-local box complexity. In fact, we will only use this last variant when we talk about secure function evaluation. Note also that in all these models, the non-local boxes are still non-signaling and Alice and Bob receive the outputs of the non-local boxes immediately after they enter their inputs.

Finally, we consider a restriction where the players always output the same predefined function g of the outputs of the non-local boxes. Let $(a_1, b_1), \dots, (a_t, b_t)$ be the outcomes of the t non-local boxes in some particular run of a protocol. Of particular interest are protocols where Alice outputs $a = a_1 \oplus \cdots \oplus a_t$ and Bob outputs $b = b_1 \oplus \cdots \oplus b_t$. The function g is used in a superscript to denote the complexity of a function f in this model, NL^g in the deterministic case, and NL_ε^g in the ε error case, and in particular, $NL^{||,\oplus}$ and $NL_\varepsilon^{||,\oplus}$ when the non-local boxes are in parallel and $g = \oplus$.

2.3 Secure Function Evaluation

We will consider the following cryptographic primitives.

Definition 4 (Oblivious transfer)**.** *A 2-1 Oblivious Transfer (OT) is a device which on input bits p_0, p_1 for Alice and q for Bob, outputs bit b to Bob, such that $b = p_q$.*

Definition 5 (Secure AND)**.** *A secure AND is a device which on input bits p for Alice and q for Bob, outputs bit a to Alice, such that $a = p \cdot q$.*

While at first view, these definitions seem similar to the definition of the non-local box, note that the timing properties are different: for the cryptographic primitives, the outputs are produced only after all the inputs have been entered into the device. It is precisely this subtlety that has led to confusion when trying to use non-local boxes to implement cryptographic primitives, in particular for bit commitment, when timing is particularly important, since a cheating Alice could wait until the reveal phase before committing her bit into the non-local box, without Bob ever realizing it [BCU$^+$07]. However, we will see that this is not an issue for our results on secure computation.

Let $f : \mathcal{X} \times \mathcal{Y} \to \{0, 1\}$ be a bipartite Boolean function. We study the number of cryptographic primitives required to securely compute f. In all the models we consider, we require perfect privacy.

In the honest-but-curious model, perfect privacy means that an honest-but-curious player, i.e. a player who follows the protocol properly while keeping a record of his internal coins and the messages from the other player, does not learn *more than required* about the other player's input. *Not more than required* means, for models where the function must be computed in parity, that the players should learn nothing about the other's input, while for models where one of the player should output the function, it means that this player should learn nothing more than what he can infer from his input and the value of the function, while the other player should learn nothing. In other words, anything that can be obtained by an honest-but-curious player after participating in the protocol, can be obtained from the input and output available to that player. Formally, a protocol is private (or secure) if there is an efficient way to simulate the view of each honest-but-curious player, where the view is defined as the player's input, his internal coins and the messages that he receives during the protocol.

In the malicious model, privacy must still hold even if the player does not follow the protocol. Formally, one defines privacy of a protocol by saying that any real execution of the protocol, for any malicious player, just emulates an ideal protocol, where the players interact with a third trusted party. We refrain from getting into the details here, since we only argue privacy for an extremely simple case of protocols, where there is no communication and the players use non-signaling non-local boxes. Such protocols are trivially secure, since no information is exchanged between the two players (neither by communication nor by the non-local boxes).

Let us note that AND may not be used as a primitive in the malicious model, so we will consider the OT primitive instead. Moreover, in this model, it is known that perfect privacy [DM99] cannot be achieved without randomness. Therefore, in this setting we do not consider the deterministic model. Our bounds in the randomized malicious model also hold for the weaker honest-but-curious model.

- $AND(f)$: number of secure AND gates required to securely compute $f(x, y)$ (*not* in parity) in the deterministic, honest-but-curious model. We note that we can allow free two-way communication without in fact changing the complexity [BM04].
- $OT_\varepsilon(f)$: number of 2-1 Oblivious Transfer calls required to compute $f(x, y)$ in parity with perfect privacy and ε error over the players' private coins, assisted with (free) two-way communication, in the malicious model.

2.4 Complexity Measures

We will compare non-local box complexity to traditional models of communication complexity and prove upper and lower bounds for this new model. Some of these bounds are in terms of factorization norms [LS08] and related measures.

Definition 1. *Let M be a real matrix. The γ_2 norm of M is $\gamma_2(M) = \min_{X^TY=M} col(X)col(Y)$, where $col(N)$ is the largest Euclidian norm of a column of N.*

We consider two representations of a Boolean function $f : X \times Y \to \{0, 1\}$ by a matrix. The first one is the communication matrix *i.e.*, the matrix M_f whose entries are $M_f[x, y] = f(x, y)$. The other representation is the sign matrix S_f whose entry $S_f[x, y]$ is 1 is $f(x, y) = 0$ and -1 otherwise. One representation can be deduced from the other by a linear transformation.

Although we won't directly use the sign matrix representation outside of this section, it is important for the factorization norm lower bound. Indeed, it is known that for any Boolean function f, $2 \log(\gamma_2(S_f))$ is a lower bound on the deterministic communication complexity of f [LS08]. In the rest of the paper, we will write $\gamma_2(f)$ for $\gamma_2(S_f)$.

We now introduce a "smoothed" version of this measure. Intuitively, $\gamma_2^\alpha(A)$ is the minimum value of $\gamma_2(M)$ for matrices inside a ball of a fixed radius centered on A. This version of γ_2 naturally comes out when one wants to lower bound the bounded error communication complexity in the randomized and quantum case.

Definition 2. *Let M be a sign matrix and $\alpha \geq 1$. $\gamma_2^\alpha(M) = \min\{\gamma_2(N) : \forall i,j\ 1 \leq M[i,j]N[i,j] \leq \alpha\}$. In particular, $\gamma_2^\infty(M)$ is the minimum γ_2 norm over all matrices N such that $1 \leq M[i,j]N[i,j]$.*

The measures γ_2^α and γ_2^∞ give upper and lower bounds for bounded-error communication complexity [LS08]: $2\log(\gamma_2^\alpha(f)/\alpha) \leq R_\varepsilon(f)$ and $R_\epsilon^{||,MAJ}(f) \leq O((\gamma_2^\infty(f))^2)$ (implicit in [LS08]), where $\alpha = \frac{1}{1-2\varepsilon}$.

The discrepancy of a sign matrix M over inputs $X \times Y$ with respect to distribution μ over the inputs is $Disc_\mu(M) = \max_R \sum_{(x,y)\in R} \mu(x,y)M[x,y]$, where R ranges over all sets of the form $R = S \times T$. It is known that $\gamma_2^\infty(f) = \Theta(\frac{1}{Disc(f)})$, and for any α, $\gamma_2^\infty(f) \leq \gamma_2^\alpha(f)$ [LS08].

Finally, for a Boolean function, the L_1 norm is defined as the sum of the absolute values of its Fourier coefficients.

Definition 3. *Let $f : \{-1,1\}^{2n} \to \{-1,1\}$, and denote by α_S the Fourier coefficients of f, that is $f(x) = \sum_{S\subseteq\{0,1\}^{2n}} \alpha_S \chi_S(x)$ where $\chi_S(x) = \prod_{i\in S} x_i$. The L_1 norm of f is defined by $L_1(f) = \sum_S |\alpha_S|$.*

We can think of the $2n$ bits of input of the function as equally split between Alice and Bob. Grolmusz uses this notion to upper bound the randomized communication complexity by proving that $R_\epsilon(f) \leq O(L_1^2(f))$ [Gro97].

3 Deterministic non-local box complexity

3.1 Characterization of $NL^{||,\oplus}$ in terms of rank

We start by studying a restricted model of non-local box complexity, where the non-local boxes are used in parallel and at the end of the protocol, Alice and Bob output the parity of the outputs of their non-local boxes respectively. We will show that the complexity of f in this model is equal to the rank of the communication matrix of f over $\mathbb{GF}_2$. It is known that this rank is equal to the minimum m, such that $f(x,y)$ can be written as $f(x,y) = \bigoplus_{i=1}^m a_i(x)\cdot b_i(y)$ (see also [Bd01]).

This restricted variant of non-local box complexity is exactly the one that appears in van Dam's work [vD05], where he shows that any Boolean function f can be computed by such a protocol of complexity 2^n. Moreover, we prove that the restriction that the players output the XOR of the outcomes of the non-local boxes is without loss of generality.

Theorem 1. $NL^{||,\oplus}(f) = \mathrm{rank}_{\mathbb{GF}_2}(M_f) = D^{||,IP_2}(f)$.

Proof. We start by showing that $NL^{||,\oplus}(f) \leq \mathrm{rank}_{\mathbb{GF}_2}(M_f)$. Let $\mathrm{rank}_{\mathbb{GF}_2}(M_f) = t$, *i.e.*, $f(x,y) = \bigoplus_{i\in[t]} p_i(x)\cdot q_i(y)$. Then we construct a protocol that uses t non-local boxes in parallel, where Alice and Bob output the parity of the outcomes of the non-local boxes and for every input (x,y) the output of the protocol is equal to $f(x,y)$. The inputs of Alice and Bob to the i-th non-local box are the bits $p_i(x)$ and $q_i(y), i \in [t]$ respectively and let a_i, b_i the outputs of the non-local box such that $a_i \oplus b_i = p_i(x)\cdot q_i(y)$. Alice and Bob output at the end of the protocol the value $(\bigoplus_{i\in[t]} a_i) \oplus (\bigoplus_{i\in[t]} b_i) = \bigoplus_{i\in[t]} p_i(x)\cdot q_i(y) = f(x,y)$.

Conversely, if there exists a protocol where Alice and Bob use t non-local boxes in parallel with inputs $p_i(x), q_i(y)$ and outputs a_i, b_i, their final output is $(\bigoplus_{i\in[t]} a_i) \oplus (\bigoplus_{i\in[t]} b_i)$ and it always equals $f(x,y)$, then we have $f(x,y) = (\bigoplus_{i\in[t]} a_i) \oplus (\bigoplus_{i\in[t]} b_i) = \bigoplus_{i\in[t]} p_i(x)\cdot q_i(y)$ and hence $\mathrm{rank}_{\mathbb{GF}_2}(M_f) \leq t$.

From this last argument, we get $D^{||,IP_2}(f) \leq NL^{||,\oplus}(f)$ since the players can send p_i and q_i to the referee who computes the inner product. For the converse, if the referee receives m_A, m_B from each player and computes their inner product mod 2, the players can instead input each bit of the message into a non-local box and output the parity of the outputs to obtain the same result. ☐

For the next corollary, we use the fact that $\log(2\mathrm{rank}_{\mathbb{F}}(M_f) - 1) \leq D(f) + 1$ for any field $\mathbb{F}$ (see [KN97]). (The plus one on the right side of the inequality appears because in our model where the value of the function is distributed among the players, the communication complexity can be one bit less than in the standard model.)

Corollary 1. $NL^{||,\oplus}(f) \leq 2^{D(f)}$.

On the other hand, it is easy to see that the one-way communication complexity is a lower bound on the non-local box complexity.

Lemma 1. $D^{\rightarrow}(f) \leq NL(f)$.

Proof. For any deterministic non-local box protocol of complexity t, Alice can send her t inputs to the non-local boxes to Bob, and since the protocol is always correct, in particular it is correct if both players assume that the output of Alice's non-local boxes are 0, Alice can output using this assumption. Bob can then compute his outputs of the non-local boxes and complete the simulation of the protocol. This shows that the one-way communication complexity is at most t. □

Notice that similarly to the traditional model, this implies an upper bound on the simultaneous messages model when computing in parity as well since for deterministic communication complexity, $D^{||}(f) \leq D^{\rightarrow}(f) + D^{\leftarrow}(f) + 2$. To see this, it suffices to see that Alice's message plus her output, together with Bob's message plus his output, determine a monochromatic rectangle in the communication matrix.

3.2 Removing the XOR restriction

In this section we show that both in the general and in the parallel model of deterministic non-local box complexity, we can assume without loss of generality that the players output the XOR of the outcomes of the non-local boxes.

Theorem 2. $NL(f) \leq NL^{\oplus}(f) \leq NL(f) + 2$.

Proof. Let P any protocol that uses t non-local boxes and at the end Alice outputs $A(x, \mathbf{a})$ and Bob outputs $B(y, \mathbf{b})$, where $\mathbf{a} = (a_1, \ldots, a_t)$ and $\mathbf{b} = (b_1, \ldots, b_t)$ are Alice and Bob's non-local box outputs. Instead of outputting these values they use another two non-local boxes with inputs $(A(x, \mathbf{a}) \oplus (\bigoplus_{i\in[t]} a_i), 1), (1, B(y, \mathbf{b}) \oplus (\bigoplus_{i\in[t]} b_i))$. Denote by $(a_{t+1}, b_{t+1}), (a_{t+2}, b_{t+2})$ the outputs of the non-local boxes. We have

$$a_{t+1} \oplus b_{t+1} = A(x, \mathbf{a}) \oplus \bigoplus_{i\in[t]} a_i \quad , \quad a_{t+2} \oplus b_{t+2} = B(y, \mathbf{b}) \oplus \bigoplus_{i\in[t]} b_i$$

Finally,

$$\begin{aligned} (\bigoplus_{i\in[t+2]} a_i) \oplus (\bigoplus_{i\in[t+2]} b_i) &= (\bigoplus_{i\in[t]} a_i) \oplus (\bigoplus_{i\in[t]} b_i) \oplus A(x, \mathbf{a}) \\ &\quad \oplus B(y, \mathbf{b}) \oplus (\bigoplus_{i\in[t]} a_i) \oplus (\bigoplus_{i\in[t]} b_i) \\ &= A(x, \mathbf{a}) \oplus B(y, \mathbf{b}). \end{aligned}$$

□

Unlike the general case, showing that in the parallel case we can assume that the players output the XOR of the outputs of the non-local boxes is not a trivial statement.

Theorem 3. $NL^{||}(f) \leq NL^{||,\oplus}(f) \leq NL^{||}(f) + 2$.

We proceed by providing two lemmas before proving our theorem.

Lemma 2. *Let a, b the outcomes of a non-local box and F, G, H arbitrary Boolean coefficients that do not depend on a, b. If for all a, $(F \cdot a) \oplus (G \cdot b) \oplus H = 0$, then $F = G$.*

Proof. Denote by p, q the inputs to the non-local box. By setting $a = 0$ and $a = 1$, we have $(G \cdot p \cdot q) \oplus H = 0$ and $F \oplus G \oplus (G \cdot p \cdot q) \oplus H = 0$. This implies $F = G$. □

We now fix some notation. Let $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ and P a protocol that computes f with zero error and uses t non-local boxes in parallel. Let $p_i(x)$, $q_i(y)$ the inputs to the i-th non-local box and a_i, b_i the corresponding outputs. We also note $\mathbf{a} = (a_1, \dots, a_t)$ and $\mathbf{b} = (b_1, \dots, b_t)$. Let $A(x, \mathbf{a}) = \bigoplus_{S \subseteq [t]} A_S(x) \cdot a_S$ and $B(y, \mathbf{b}) = \bigoplus_{S \subseteq [t]} B_S(y) \cdot b_S$ the final outputs of Alice and Bob, where A_S are polynomials in x, B_S polynomials in y, $a_S = \prod_{i \in S} a_i$ and $b_S = \prod_{i \in S} b_i$. Then, from the correctness of the protocol, we have

$$\forall (x, y, \mathbf{a}), f(x, y) = A(x, \mathbf{a}) \oplus B(y, \mathbf{b}).$$

We show that, without loss of generality, we may assume that the inputs to the non-local boxes satisfy some linear independence condition.

Definition 4. *A set of bipartite functions $\{f_i | i \in T\}$ is linearly independent if $\bigoplus_{i \in T} C_i \cdot f_i(x, y) = \alpha(x) \oplus \beta(y)$, for some $C_i \in \{0,1\}$ and functions α, β and for all x and y, implies $C_i = 0 \; \forall i \in T$ and $\alpha(x) = \beta(y)$ for all x and y.*

Lemma 3. *Let P be a protocol for f using t non-local boxes in parallel. Then there exists another protocol whose output is always equal to the one of P, uses $t' \leq t$ non-local boxes in parallel and the inputs to the non-local boxes are such that the set $\{p_i(x) \cdot q_i(y) | i \in [t']\}$ is linearly independent.*

Proof. Suppose that $\bigoplus_{i \in [t]} C_i \cdot p_i(x) \cdot q_i(y) = \alpha(x) \oplus \beta(y)$, with $C_k = 1$ for some $k \in [t]$. Then $p_k(x) \cdot q_k(y) = \alpha(x) \oplus \beta(y) \oplus \bigoplus_{i \in [t] \setminus \{k\}} C_i \cdot p_i(x) \cdot q_i(y)$. Since $p_k(x) \cdot q_k(y) = a_k \oplus b_k$, Alice and Bob do not need to use the k-th non-local box when implementing protocol P, it suffices for Alice to set $a_k = \alpha(x) \oplus \bigoplus_{i \in [t] \setminus \{k\}} C_i \cdot a_i$ and for Bob to set $b_k = \beta(y) \oplus \bigoplus_{i \in [t] \setminus \{k\}} C_i \cdot b_i$, which implies a new protocol with $t - 1$ non-local boxes. By repeating this procedure, they can build a protocol using $t' \leq t$ non-local boxes and such that the whole set $\{p_i(x) \cdot q_i(y) | i \in [t']\}$ is linearly independent. $\square$

Proof of Theorem 3. Since by definition $NL^{||}(f) \leq NL^{||,\oplus}(f)$, it suffices to show that $NL^{||,\oplus}(f) \leq NL^{||}(f) + 2$. Let $NL^{||}(f) = t$ and let P be a deterministic protocol for f, using t non-local boxes. Let $A(x, \mathbf{a}) = A_\emptyset(x) \oplus \bigoplus_{S \subseteq [t]} A_S(x) \cdot a_S$ and $B(y, \mathbf{b}) = B_\emptyset(y) \oplus \bigoplus_{S \subseteq [t]} B_S(y) \cdot b_S$ the outputs of Alice and Bob respectively, where the subsets S are non-empty. First, in order to simulate the two local terms $A_\emptyset(x)$ and $B_\emptyset(y)$, Alice and Bob can use two non-local boxes with inputs $(A_\emptyset(x), 1)$ and $(1, B_\emptyset(y))$. For the rest of the proof, all the subsets we consider are non-empty. We proceed by proving two claims about the outputs of the protocol.

Claim 1. *For all $(x, y, \mathbf{a})$ and for all $T \subseteq [t]$,* $\quad \bigoplus_{S : T \subseteq S} A_S(x) \cdot a_{S \setminus T} = \bigoplus_{S : T \subseteq S} B_S(y) \cdot b_{S \setminus T}$.

Proof. We prove this claim by induction on the size of T. By definition, the protocol satisfies for all $(x, y, \mathbf{a})$, $f(x, y) = A(x, \mathbf{a}) \oplus B(y, \mathbf{b})$. By factorizing the k-th non-local box, we get the following expression for every $(x, y, \mathbf{a})$:

$$f(x, y) = \bigoplus_{S : k \notin S} (A_S(x) \cdot a_S \oplus B_S(y) \cdot b_S) \oplus a_k \cdot (\bigoplus_{S : k \in S} A_S(x) \cdot a_{S \setminus \{k\}}) \oplus b_k \cdot (\bigoplus_{S : k \in S} B_S(y) \cdot b_{S \setminus \{k\}}).$$

We can now use Lemma 2, and have that

$$\bigoplus_{S : k \in S} A_S(x) \cdot a_{S \setminus \{k\}} = \bigoplus_{S : k \in S} B_S(y) \cdot b_{S \setminus \{k\}},$$

for all $(x, y, \mathbf{a})$ and for all $k \in [t]$. Hence, the claim is true for any subset T with $|T| = 1$. Suppose for the induction that it is true for any set of size $n \in [t-1]$ and consider any T such that $|T| = n$. Let $k \notin T$,

$$\begin{aligned} \bigoplus_{S : T \subseteq S} A_S(x) \cdot a_{S \setminus T} \quad &= \quad \bigoplus_{S : T \subseteq S} B_S(y) \cdot b_{S \setminus T}, \\ \bigoplus_{S : T \subseteq S, k \notin S} A_S(x) a_{S \setminus T} \oplus a_k (\bigoplus_{S : T \cup \{k\} \subseteq S} A_S(x) a_{S \setminus T \cup \{k\}}) \quad &= \quad \bigoplus_{S : T \subseteq S, k \notin S} B_S(y) b_{S \setminus T} \oplus b_k (\bigoplus_{S : T \cup \{k\} \subseteq S} B_S(y) b_{S \setminus T \cup \{k\}}). \end{aligned}$$

Applying Lemma 2 in the previous equation proves the claim for any $T \cup \{k\}$ and hence any set of size $n + 1$, which concludes the proof of Claim 1. $\square$

Claim 2. *For all (x, y) and for all $T \subseteq [t]$, we have:*

$$|T| > 1 \quad \Rightarrow \quad A_T(x) = B_T(y) = 0,$$
$$|T| = 1 \quad \Rightarrow \quad A_T(x) = B_T(y).$$

Proof. We prove this claim by downward induction on the size of T, starting with $|T| = t$, that is, $T = [t]$. We immediately obtain from Claim 1 that $A_{[t]}(x) = B_{[t]}(y)$. As a consequence, these do not depend on x or y, and we may define $C_{[t]} = A_{[t]}(x) = B_{[t]}(y)$. Moreover, we can define $A_S(x) = B_S(y) = 0$ for any $S \supsetneq [t]$.

Now let $n \geq 2$ and suppose that for any set S of size equal or larger to $n+1$, we have $A_S(x) = B_S(y) = 0$, and for any set S of size n, we have $A_S(x) = B_S(y)$. From Claim 1, we obtain that for all sets T of size $n-1$,

$$A_T(x) \oplus \bigoplus_{k \notin T} C_{T \cup \{k\}} \cdot a_k = B_T(y) \oplus \bigoplus_{k \notin T} C_{T \cup \{k\}} \cdot b_k,$$

where we have defined $C_S = A_S(x) = B_S(y)$ for all S of size n. Since $a_k \oplus b_k = p_k(x) \cdot q_k(y)$, we have $\bigoplus_{k \notin T} C_{T \cup \{k\}} \cdot p_k(x) \cdot q_k(y) = A_T(x) \oplus B_T(y)$, and by linear independence, we conclude that $C_{T \cup \{k\}} = 0$ and $A_T(x) = B_T(y)$. Using the same argument for any T of size $n-1$, we obtain that $A_S(x) = B_S(y) = 0$ for all S of size n, and $A_T(x) = B_T(y)$ for all T of size $n-1$, which concludes the proof of Claim 2. $\square$

This claim implies that the protocol P outputs the parity of two local terms plus the outcomes of the non-local boxes, and as a consequence $NL^{||,\oplus}(f) \leq t + 2$. $\square$

3.3 Optimality of our bounds

We show here that the bounds we proved in the previous section on the parallel and general non-local box complexity ($D(f) \leq NL^{||}(f) \leq 2^{D(f)}$ and $D(f) \leq NL(f) \leq NL^{||}(f)$ respectively) are optimal by giving examples of functions that saturate them. The first function we consider is the Inner Product function, $IP(x, y) = \oplus_i (x_i \wedge y_i)$, with $x, y \in \{0,1\}^n$. For this function we have that $D(IP) = NL(IP) = NL^{||}(IP) = n$.

The second function we consider is the Disjointness function, which is equal to $DISJ(x, y) = \vee_i (x_i \wedge y_i)$, with $x, y \in \{0,1\}^n$. It is well-known that for the communication matrix of the Disjointness function we have $\mathrm{rank}_{\mathbb{GF}_2}(M_{DISJ}) = 2^n$ and hence $NL^{||,\oplus}(DISJ) = 2^n$. On the other hand, we have $D(DISJ) = n$ and show that $NL(DISJ) = O(n)$. We describe below a simple protocol for the Disjointness function that follows from [BBL$^+$06] and was pointed out to us by Troy Lee and Falk Unger. The Disjointness function also provides an example of an exponential separation between deterministic parallel and general non-local box complexity.

Proposition 1. $NL(DISJ) \leq O(n)$.

Proof. On input $x = x_1 \cdots x_n, y = y_1 \cdots y_n$, Alice and Bob use n non-local boxes with inputs (x_i, y_i) and get outputs a_i, b_i with $a_i \oplus b_i = x_i \cdot y_i$. Then, they can use 2 non-local boxes in order to compute the OR of two such distributed bits since $(a_k \oplus b_k) \vee (a_\ell \oplus b_\ell) = (a_k \vee a_\ell) \oplus (a_k \vee b_\ell) \oplus (b_k \vee a_\ell) \oplus (b_k \vee b_\ell)$. The terms $(a_k \vee a_\ell)$ and $(b_k \vee b_\ell)$ can be locally computed by Alice and Bob respectively and hence they only need to use two non-local boxes with inputs $(\neg a_k, \neg b_\ell)$ and $(\neg a_\ell, \neg b_k)$ to compute the remaining terms. By combining n such distributed OR computations they compute $\vee_i (a_i \oplus b_i)$ and hence output the value of $DISJ(x, y)$ after using $3n$ non-local boxes. $\square$

4 Randomized non-local box complexity

In this section, we consider protocols that use shared randomness and have success probability at least $2/3$. We start by comparing the parallel non-local box complexity to communication complexity. Then we exactly characterize $NL_\varepsilon^{||,\oplus}$ in terms of the approximate rank (over $\mathbb{GF}_2$) of the communication matrix.

4.1 Upper and lower bounds for NL_ϵ

Theorem 4. $R_\epsilon^{\rightarrow}(f) \leq NL_\varepsilon(f) \leq NL_\varepsilon^{||,\oplus}(f) \leq 2^{R_\epsilon(f)}$.

Proof. For the first inequality, Alice sends all her inputs to the non-local boxes to Bob. They use the shared randomness to simulate the output of Alice's non-local boxes, which Alice can use to compute her output, and Bob uses to compute his outputs to the non-local boxes, and compute his output.

For the last inequality, let us fix a randomized communication protocol P for f using t bits of communication. We can write P as a distribution over deterministic protocols P_r each using at most t bits of communication, and computing some Boolean function f_r. By Corollary 1, $NL^{||,\oplus}(f_r) \leq 2^t$. Taking the same distribution over the non-local box protocols for f_r, we get $NL_\varepsilon^{||,\oplus}(f) \leq 2^t$ as claimed. □

Note that in fact any $NL_\epsilon^{||,g}$ protocol can be simulated in the simultaneous messages communication model, so in fact $R_\epsilon^{||}(f) \leq NL_\epsilon^{||,g}(f)$, for any g.

The approximate rank over the reals has been shown to be a useful complexity measure for communication complexity [Bd01]. For non-local box complexity, we now define the notion of approximate rank over $\mathbb{GF}_2$.

Definition 5. *Let $\mathcal{P}_t$ denote the convex hull of Boolean matrices with rank over $\mathbb{GF}_2$ at most t. Then for a $[0,1]$-valued*[1] *matrix A the approximate rank over $\mathbb{GF}_2$ is defined by $\varepsilon{-}\mathrm{rank}_{\mathbb{GF}_2}(A) = \min\{t : \exists A' \in \mathcal{P}_t \text{ with } \parallel A - A' \parallel_\infty \leq \varepsilon\}$.*

The next proposition gives an alternative definition of the approximate rank for Boolean matrices. This definition enables us to relate the approximate rank to the non-local box complexity.

Proposition 2. *$\varepsilon{-}\mathrm{rank}_{\mathbb{GF}_2}(M_f)$ is the minimum t, such that there exists a set of Boolean matrices $A_1, \ldots, A_R$, and a probability distribution over $[R]$ with the following properties:*

- *For every $r \in [R]$, $\mathrm{rank}_{\mathbb{GF}_2}(A_r) \leq t$,*
- *For every (x, y), $\mathrm{Prob}_r[M_f[x,y] = A_r[x,y]] \geq 1 - \varepsilon$.*

Proof. Suppose that $\varepsilon{-}\mathrm{rank}_{\mathbb{GF}_2}(M_f) = t$, and let $A \in \mathcal{P}_t$ such that $\parallel M_f - A \parallel_\infty \leq \varepsilon$. Denote by $A_1, \ldots, A_d$ the vertices of $\mathcal{P}_t$. By definition, $A = \sum_i \mu_i A_i$, with $\sum_i \mu_i = 1$ and $\forall i,\ \mu_i \geq 0$. For any (x, y), picking $A_i[x, y]$ with probability μ_i has expected value $E_\mu(A_i[x, y]) = A[x, y]$. It follows that $\mathrm{Prob}_\mu[M_f[x, y] \neq A_i[x, y]] = \mid M_f[x, y] - E_\mu(A_i[x, y]) \mid \leq \varepsilon$. Moreover, for any i, $\mathrm{rank}_{\mathbb{GF}_2}(A_i) \leq t$. This proves that the set $A_1, \ldots, A_d$ and μ have the desired properties. The proof goes conversely as well.

□

Theorem 5. *For any Boolean function f, $NL_\varepsilon^{||,\oplus}(f) = \varepsilon{-}\mathrm{rank}_{\mathbb{GF}_2}(M_f)$*

Proof. Fix a randomized protocol for f that uses t non-local boxes in parallel and is correct with probability at least $1 - \varepsilon$. We assume that the players share a random string r drawn according to some probability distribution from a set R. Since the protocol uses the boxes in parallel, their inputs only depend on the players' inputs x and y, and on the shared random string r. Let $p_i(x, r)$ and $q_i(x, r)$ denote Alice's and Bob's input to the i-th non-local box, and a_i, b_i the outputs.

Since Alice and Bob output the XOR of the outcomes of the non-local boxes, the final outcome of the protocol for a fixed r is some function $g_r(x, y) = \bigoplus_i a_i \oplus \bigoplus_i b_i = \bigoplus_i p_i(x, r) \oplus q_i(y, r)$. Let us the consider the matrices $A_r = M_{g_r}$. We know that $\mathrm{rank}_{\mathbb{GF}_2}(A_r) \leq t$ for every r. Moreover, the correctness of the protocol implies $\mathrm{Prob}_r[M_f(x, y) = A_r(x, y)] \geq 1 - \varepsilon$. This proves that $\varepsilon{-}\mathrm{rank}_{\mathbb{GF}_2}(M_f) \leq t$.

Conversely, suppose that $\varepsilon{-}\mathrm{rank}_{\mathbb{GF}_2}(M_f) = t$. Then fix a set of Boolean matrices $A_1, \ldots, A_R$ such that for every $r \in [R]$, $\mathrm{rank}_{\mathbb{GF}_2}(A_r) \leq t$ and $\mathrm{Prob}_r[M_f(x, y) = A_r(x, y)] \geq 1 - \varepsilon$. Consider the following protocol for f: Alice and Bob pick at random $r \in R$ and compute the function $g_r(x, y) = A_r(x, y)$. As $\mathrm{rank}_{\mathbb{GF}_2}(A_r) \leq t$, computing g_r requires at most t non-local boxes in parallel. It is straightforward to check that this protocol is correct with probability at least $1 - \varepsilon$. Hence, $NL_\varepsilon^{||,\oplus}(f) \leq t$. □

[1] While in this section we are only interested in Boolean matrices, we give the definition for the general case of [0, 1]-valued matrices as it will be useful in the next section.

In the randomized case, it is easy to get rid of the XOR restriction in general non-local box protocols, since the proof for the deterministic case still goes through. On the other hand, for the parallel case, this appears to be a surprisingly deep question, which remains open. The main obstacle appears to be related to the inherent randomness of the non-local boxes.

Next, we relate the general non-local box complexity to the following model of communication: Alice and Bob send to a referee one message each and the referee outputs 1 if for the majority of indices, the two messages are equal. We denote the communication complexity in this model by $R_\epsilon^{||,MAJ}(f)$. This is a natural model of communication complexity that has appeared repeatedly in the simulation of quantum protocols by classical ones, as well as various upper bounds on simultaneous messages [KNR99, Gro97, SZ08, LS08].

Theorem 6. $R_\varepsilon^{\rightarrow}(f) \leq NL_\varepsilon(f) \leq O(R_\epsilon^{||,MAJ}(f))$.

Proof. For the lower bound, Alice and Bob can use shared randomness to simulate the output of Alice's non-local boxes. Alice then computes her inputs to the non-local boxes, and sends them to Bob. From Alice's inputs and outputs to the non-local boxes, Bob may compute his inputs and outputs. The players may then compute their outputs to the protocol, which have the same probability distribution as the original protocol.

For the upper bound, fix a t-bit simultaneous protocol for f, where the referee receives two messages $\mathbf{a}$ and $\mathbf{b}$ of size t from Alice and Bob and outputs $MAJ(a_1 \oplus b_1, \ldots, a_t \oplus b_t)$. It is well-known, by using an addition circuit, that the majority of t bits can be computed by a circuit of size $O(t)$ with AND, NOT gates. Moreover, the distributed AND of two bits can be computed using two non-local boxes [BBL+06]. We conclude that the non-local box complexity of the distributed Majority is $O(t)$ and hence the theorem follows. □

Our theorem implies the following relation between non-local box complexity and factorization norms.

Corollary 2. $2\log(\gamma_2^\alpha(f)/\alpha) \leq NL_\varepsilon(f) \leq O((\gamma_2^\infty(f))^2)$, *where* $\alpha = \frac{1}{1-2\varepsilon}$.

Proof. It follows from our theorem and the inequalities $2\log(\gamma_2^\alpha(f)/\alpha) \leq R_\varepsilon(f)$ (see [LS08]) and $R_\epsilon^{||,MAJ}(f) \leq O((\gamma_2^\infty(f))^2)$ (also implicitly in [LS08]). □

It is known that $\gamma_2^\infty(f) = \Theta(\frac{1}{Disc(f)})$, and also that for any α, $\gamma_2^\infty(f) \leq \gamma_2^\alpha(f)$ [LS08]. Hence, since discrepancy gives a lower bound on the quantum communication complexity with entanglement $Q_\varepsilon^*(f)$ [Kre95], we get the following corollary.

Corollary 3. $NL_\epsilon(f) \leq O(2^{2Q_\varepsilon^*(f)})$.

Finally, we can relate the non-local box complexity of a function f, to the L_1 norm of the Fourier coefficients of f by using a result by Grolmusz. Grolmusz showed that for any Boolean function f, there exists a randomized public coin protocol that solves f with complexity $O(L_1^2(f))$. This protocol can be easily transformed into a simultaneous messages protocol where the referee outputs the distributed majority of the message bits. Hence,

Corollary 4. $NL_\epsilon(f) \leq O(L_1^2(f))$.

Let us make here a last remark about the proof of Theorem 6. We started from a Simultaneous Messages protocol where the referee outputs a Majority function and we constructed a non-local box protocol with complexity equal to the communication complexity. If we look at this protocol, we can see that Alice and Bob can use their non-local boxes in the same order. This will be useful when we relate non-local boxes to secure function evaluation.

Corollary 5. $R_\varepsilon^{\rightarrow}(f) \leq NL_\varepsilon(f) \leq NL_\varepsilon^{\mathrm{ord}}(f) \leq O(R_\epsilon^{||,MAJ}(f))$.

4.2 Optimality of our bounds and an efficient parallel protocol for Disjointness

In the deterministic case, we showed that our bounds are tight and also that the parallel and the general non-local box complexity can be exponentially different. Is the same true for the randomized case?

In fact, the Disjointness and Inner Product functions almost saturate our bound in terms of γ_2^∞ for the general randomized non-local box complexity. More precisely, for the Disjointness function, we have that $NL_\varepsilon(DISJ) = \Theta(n)$

(since $R_\varepsilon(DISJ) = \Omega(n)$ and $NL(DISJ) \leq O(n)$) and using discrepancy [KN97, Exercise 3.32], we have $(\gamma_2^\infty(DISJ))^2 = \Theta(n^2)$. On the other hand, for the Inner Product function we have $NL_\varepsilon(IP) = \Theta(n)$ but $(\gamma_2^\infty(IP))^2 = \Theta(2^n)$.

The case of parallel non-local box complexity is more interesting. We can give a simple parallel protocol for the Disjointness function of complexity $O(n)$, hence showing that the exponential separation does not hold anymore. It is an open question whether or not parallel and general randomized non-local box complexity are polynomially related.

Proposition 3. $NL^{||}_{1/3}(DISJ) \leq O(n)$.

Proof. The idea is to reduce the Disjointness problem to a problem of calculating an Inner Product, which we know how to do with n parallel non-local boxes. In order to solve the general Disjointness problem with high probability, Alice and Bob proceed as follows: they look at a shared random string $r_1, \ldots, r_n$ and consider the strings $x \wedge r$ and $y \wedge r$ as inputs. In other words, they pick a random subset of their input bits, by picking each index with probability $1/2$. Then they perform an Inner Product calculation on their new inputs by using n non-local boxes in parallel. Let $a \oplus b = IP(x \wedge r, y \wedge r)$. It is easy to see that if $DISJ(x, y) = 0$, then $IP(x \wedge r, y \wedge r) = 0$ for all r. On the other hand, if $DISJ(x, y) = 1$, *i.e.*, if the intersection is non-empty, then $\mathrm{Prob}_r[IP(x \wedge r, y \wedge r) = 1] = 1/2$, since for a random subset, the probability that the size of the intersection on this subset is odd is exactly the same as the probability that the intersection is even. Hence, we have a one-sided error algorithm for Disjointness that is always correct when $DISJ(x, y) = 0$ and is correct with probability $1/2$ when $DISJ(x, y) = 1$.

We can get a two-sided error algorithm in the following way: Alice and Bob simulate the protocol above until they obtain the outputs a, b. Then, using their shared randomness, they output $a \oplus b$ with probability $1 - p$, and $0 \oplus 1$ or $1 \oplus 0$ with probability $p/2$. It is easy to see that when $DISJ(x, y) = 0$ then the success probability is $1 - p$ and when $DISJ(x, y) = 1$ the success probability is $p + (1 - p)/2 = (1 + p)/2$. Taking $p = 1/3$ makes the overall success probability of our algorithm $2/3$. □

5 Non-local boxes and measurement simulation

The question of the nature of non-local distributions arising from measurements of bipartite quantum states dates back to Maudlin [Mau92], who used communication complexity to quantify non-locality. Perhaps a more natural question is how many non-local boxes are required to simulate quantum distributions, since non-local boxes maintain the non-signaling property of these distributions. For binary measurements on maximally entangled qubit pairs it is known that 1 use of a non-local box suffices [CGMP05].

In this section we present another application of our results on non-local boxes. Using the recent breakthrough of Regev and Toner [RT09], who give a two-bit one-way protocol for simulating two-outcome measurements on entangled states for arbitrary dimensions, we show that this can be done with 3 non-local boxes. Previously, no finite upper bound was known for this problem.

Let $\mathbf{p}$ be a distribution over measurement outcomes $\mathcal{A} \times \mathcal{B}$, conditioned on measurements $\mathcal{X} \times \mathcal{Y}$. For measurements on quantum states, the distribution is non-signaling, that is, the marginal distributions do not depend on the other player's measurement:

$$\forall a \in \mathcal{A}, b \in \mathcal{B}, x \in \mathcal{X}, x' \in \mathcal{X}, y \in \mathcal{Y}, y' \in \mathcal{Y}, \quad p(a|x, y) = p(a|x, y') \text{ and } p(b|x, y) = p(b|x', y).$$

Therefore we write the marginals $p(a|x)$ and $p(b|y)$. In this paper we focus on distributions with uniform marginals over $\mathcal{A} = \mathcal{B} = \{0, 1\}$. These distributions are in bijection with $[0, 1]$-valued matrices.

Definition 6 (Correlation matrix)**.** *Let* $\mathbf{p}$ *be a distribution with uniform marginals over* $\mathcal{A} = \mathcal{B} = \{0, 1\}$*, conditioned on measurements* $\mathcal{X} \times \mathcal{Y}$*. The correlation matrix* $C_\mathbf{p} : \mathcal{X} \times \mathcal{Y} \to [0, 1]$ *of* $\mathbf{p}$ *is defined as* $C_\mathbf{p}(x, y) = \Pr[a \oplus b = 1 | x, y]$*, where* a, b *are distributed according to* $\mathbf{p}$*.*

It is not hard to prove that the set of $[0, 1]$-valued matrices is the convex hull of the set of Boolean matrices. This implies that the corresponding non-signaling distributions can be written as convex combinations of distributions of the following form. For any $f : \mathcal{X} \times \mathcal{Y} \to \{0, 1\}$, we define the associated distribution

$$\mathbf{p}_f(a, b|x, y) = \begin{cases} \frac{1}{2} & \text{if } f(x, y) = a \oplus b \\ 0 & \text{otherwise.} \end{cases}$$

In other words, the correlation matrix $C_{\mathbf{p}_f}(x,y)$ is Boolean and coincides with the communication matrix M_f. Observe that any protocol for f simulates the distribution $\mathbf{p}_f$, since we may assume without loss of generality that the outcomes are uniformly distributed (otherwise, Alice and Bob can flip their outcomes according to a shared random bit).

Just as for functions, we can define the communication and non-local box complexities of a distribution $\mathbf{p}$. When error ϵ is allowed, we require that the distribution $\mathbf{p}'$ simulated by the protocol be such that $\| C_{\mathbf{p}} - C_{\mathbf{p}'} \|_\infty \leq \epsilon$. Since binary distributions with uniform marginals may be represented as convex combinations of distributions arising from functions, we can generalize some results of the previous section to this case:

Theorem 7. *For any distribution* $\mathbf{p}$ *with uniform marginals over* $\mathcal{A} = \mathcal{B} = \{0,1\}$*, we have*

- $R_\epsilon^{\rightarrow}(\mathbf{p}) \leq NL_\varepsilon(\mathbf{p}) \leq NL_\varepsilon^{||,\oplus}(\mathbf{p}) \leq 2^{R_\epsilon(\mathbf{p})}$,
- $NL_\varepsilon^{||,\oplus}(\mathbf{p}) = \varepsilon\mathrm{-rank}_{\mathbb{GF}_2}(C_{\mathbf{p}})$.

Proof. For the first inequality, the non-local boxes are simulated by one-way communication the same way as they were for functions: the players use shared randomness to simulate Alice's output to the non-local boxes, then Alice computes her inputs to the non-local boxes according to those outputs, and sends them to Bob (see proof of Theorem 6 for details).

Let $t = R_\epsilon(\mathbf{p})$, and consider the corresponding randomized communication protocol. Let the shared randomness take value r with probability p_r. For any possible value r, the protocol will compute a Boolean function f_r such that $D(f_r) \leq t$, and $\| C_{\mathbf{p}} - \sum_r p_r M_{f_r} \|_\infty \leq \epsilon$. From Corollary 1, $NL^{||,\oplus}(\mathbf{p}) \leq 2^t$. Executing the non-local box protocol for f_r with probability p_r, we obtain a non-local box protocol simulating $\mathbf{p}$ with error at most ϵ, so that $NL_\varepsilon^{||,\oplus}(\mathbf{p}) \leq 2^t$.

Let $t = \varepsilon\mathrm{-rank}_{\mathbb{GF}_2}(C_{\mathbf{p}})$. By definition, there exist Boolean matrices A_r and a probability distribution p_r such that $\mathrm{rank}_{\mathbb{GF}_2}(A_r) \leq t$ and $\| C_{\mathbf{p}} - \sum_r p_r A_r \|_\infty \leq \epsilon$. Let f_r be the Boolean function described by communication matrix A_r. Theorem 1 implies that $NL^{||,\oplus}(f_r) \leq t$. Executing the non-local box protocol for f_r with probability p_r, we obtain a non-local box protocol simulating $\mathbf{p}$ with error at most ϵ, so that $NL_\varepsilon^{||,\oplus}(\mathbf{p}) \leq t$. The proof goes conversely as well. $\square$

The upper bound on the non-local box complexity in terms of the communication complexity may be slightly improved. While this is an insignificant improvement for most applications involving Boolean functions, this becomes relevant when considering low communication complexity distributions, as is the case for some quantum distributions.

Theorem 8. *For any distribution* $\mathbf{p}$ *with uniform marginals over* $\mathcal{A} = \mathcal{B} = \{0,1\}$*, we have* $NL_\varepsilon^{||}(\mathbf{p}) \leq 2^{R_\epsilon(\mathbf{p})} - 1$.

Proof. We build on an idea presented in [DLR07] to replace communication by non-local boxes. Let $t = R_\epsilon(\mathbf{p})$, and let us first consider the case of one-way communication protocols. Denote $m_A(x)$ the message sent by Alice to Bob, $A(x)$ is Alice's output, and $B(m,y)$ is Bob's output when he receives message m. Suppose without loss of generality that one message is exactly the all-zero string $\mathbf{0}$. We use one non-local box for each message except $\mathbf{0}$. In the non-local box for message m, Alice inputs 1 if $m_A(x) = m$ and 0 otherwise. Bob inputs $B(m,y) \oplus B(\mathbf{0},y)$. At the end, Alice outputs $\bigoplus a_i \oplus A(x)$ and Bob outputs $\bigoplus b_i \oplus B(\mathbf{0},y)$, where (a_i,b_i) is the output of the i-th non-local box. It is easy to check that the output is always $A(x) \oplus B(m_A(x),y)$.

In the two-way communication case, the proof is more involved. We proceed recursively, at each step removing the last bit of communication. We handle the different communication scenarios by doubling the number of protocols at each step. Throughout this proof, $T^{(k)}$ will be a k-bit transcript, and $A_i^{(k)}(T^{(k)},x)$ and $B_i^{(k)}(T^{(k)},y)$ will be the players' outputs for the 2^{t-k} different k-bit communication protocols indexed by i. The outputs from the different protocols may then be used as inputs to non-local boxes, effectively replacing communication by non-local boxes. More specifically, suppose we have a deterministic t bit communication protocol that computes $f(x,y)$ in parity:

$$f(x,y) = A_0^{(t)}(T^{(t)},x) \oplus B_0^{(t)}(T^{(t)},y).$$

We prove by downward induction on k, from $k = t$ to $k = 0$, that $f(x, y)$ may be written as

$$f(x,y) \quad = \quad A_0^{(k)}(T^{(k)},x) \oplus B_0^{(k)}(T^{(k)},y) \oplus \bigoplus_{i=1}^{2^{t-k}-1} A_i^{(k)}(T^{(k)},x) \cdot B_i^{(k)}(T^{(k)},y), \tag{1}$$

which shows that $f(x, y)$ can be computed with k bits of communication (to produce the outputs $A_i^{(k)}(T^{(k)}, x)$ and $B_i^{(k)}(T^{(k)}, y)$), followed by $2^{t-k} - 1$ non-local boxes in parallel.

It will then follow by induction that

$$f(x,y) \quad = \quad A_0^{(0)}(x) \oplus B_0^{(0)}(y) \oplus \bigoplus_{i=1}^{2^{t}-1} A_i^{(0)}(x) \cdot B_i^{(0)}(y),$$

so $f(x, y)$ can be computed with $2^t - 1$ non-local boxes in parallel.

Let us consider the k-bit protocols with outputs $A_i^{(k)}(T^{(k)}, x)$ and $B_i^{(k)}(T^{(k)}, y)$ from Eq. (1) and focus on the k-th bit of the transcript $T^{(k)}$. Since both players must agree, depending on the transcript so far $T^{(k-1)}$, whether this bit is communicated by Alice to Bob or vice-versa, we may define a Boolean function $d_k = d_k(T^{(k-1)})$, which gives the direction of this bit, say $d_k(T^{(k-1)})$ is 1 if the bit is communicated by Alice to Bob, and 0 otherwise. Let us now focus on the bit strings $T^{(k-1)}$ such that $d_k(T^{(k-1)}) = 1$. Since Alice may compute the next bit to be communicated from her input and the $k - 1$ first bits of the transcript, we may write it as $c_i^{(k)} = c_i^{(k)}(T^{(k-1)}, x)$, and her output as $A_i^{(k)}(T^{(k-1)}, x)$. As for Bob's output $B_i^{(k)}(T^{(k)}, y)$, we use a construction from [DLR05] to replace one bit of communication by a non-local box:

$$B_i^{(k)}(T^{(k)},y) = B_i^{(k)}(T^{(k-1)}0,y) \oplus c_i^{(k)} \cdot [B_i^{(k)}(T^{(k-1)}0,y) \oplus B_i^{(k)}(T^{(k-1)}1,y)].$$

Therefore, when $d_k(T^{(k-1)}) = 1$ we may write $f(x, y)$ as

$$\begin{aligned} f(x,y) \quad = \quad & A_0^{(k)}(T^{(k-1)},x) \oplus B_0^{(k)}(T^{(k-1)}0,y) \oplus c_0^{(k)}(T^{(k-1)},x) \cdot [B_0^{(k)}(T^{(k-1)}0,y) \oplus B_0^{(k)}(T^{(k-1)}1,y)] \\ & \oplus \bigoplus_{i=1}^{2^{t-k}-1} A_i^{(k)}(T^{(k-1)},x) \cdot B_i^{(k)}(T^{(k-1)}0,y) \\ & \oplus \bigoplus_{i=1}^{2^{t-k}-1} A_i^{(k)}(T^{(k-1)},x) \cdot c_i^{(k)} \cdot [B_i^{(k)}(T^{(k-1)}0,y) \oplus B_i^{(k)}(T^{(k-1)}1,y)]. \end{aligned}$$

For $0 \le i \le 2^{t-k} - 1$, we define the following output functions,

$$\begin{aligned} A_i^{(k-1)}(T^{(k-1)},x) &= A_i^{(k)}(T^{(k-1)},x) \\ B_i^{(k-1)}(T^{(k-1)},y) &= B_i^{(k)}(T^{(k-1)}0,y) \\ B_{i+2^{t-k}}^{(k-1)}(T^{(k-1)},y) &= B_i^{(k)}(T^{(k-1)}0,y) \oplus B_i^{(k)}(T^{(k-1)}1,y) \\ A_{2^{t-k}}^{(k-1)}(T^{(k-1)},x) &= c_0^{(k)}(T^{(k-1)},x), \\ A_{i+2^{t-k}}^{(k-1)}(T^{(k-1)},x) &= A_i^{(k)}(T^{(k-1)},x) \cdot c_i^{(k)} \qquad \text{(if } i \neq 0\text{)}, \end{aligned}$$

when $d_k(T^{(k-1)}) = 1$, and similar expressions, with A and B swapped, when $d_k(T^{(k-1)}) = 0$. Thus, we may write

$$f(x,y) \quad = \quad A_0^{(k-1)}(T^{(k-1)},x) \oplus B_0^{(k-1)}(T^{(k-1)},y) \oplus \bigoplus_{i=1}^{2^{t-k+1}-1} A_i^{(k-1)}(T^{(k-1)},x) \cdot B_i^{(k-1)}(T^{(k-1)},y).$$

□

The simpler proof in the case of one-way protocol can be used to derive an explicit protocol using 3 non-local boxes to simulate the correlations arising from 2-outcome measurements made on an entangled bipartite state. By Tsirelson's theorem [Tsi85], the problem of simulating these correlations reduces to the following problem.

- Alice receives a unit vector $\vec{x} \in \mathbb{R}^n$
- Bob receives a unit vector $\vec{y} \in \mathbb{R}^n$
- Alice outputs $A \in \{-1, 1\}$, Bob outputs $B \in \{-1, 1\}$ such that the correlation equals the inner product of the two vectors: $E[AB] = \vec{x} \cdot \vec{y}$.

Corollary 9. *There is a protocol for simulating traceless two-outcome measurements on maximally entangled states, using 3 non-local boxes in parallel.*

Proof. We sketch the protocol of Regev and Toner. Assume that the inputs to the problem are two unit vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$.

Alice and Bob share a random dimension 3 subspace of $\mathbb{R}^n$. Let G be the matrix of the projection onto this subspace. Alice and Bob start by applying a transformation $\vec{x'} = C(\vec{x}), \vec{y'} = C(\vec{y})$ (see [RT09] for details of this transformation C), then project their vectors on the random subspace, $\vec{x''} = G\vec{x'}, \vec{y''} = G\vec{y'}$. Let $\mathrm{sgn} : \mathbb{R} \mapsto \{-1, 1\}$ be the sign function, that is, $\mathrm{sgn}(x) = 1$ if $x \geq 0$ and $\mathrm{sgn}(x) = -1$ otherwise. Alice lets $\alpha_i = \mathrm{sgn}(x''_i)$ for $i = 0, 1, 2$ and lets $c_i = \alpha_0 \cdot \alpha_i$ for $i = 1, 2$. Alice outputs $A = \alpha_0$ and sends (c_1, c_2) to Bob. Bob outputs $B = \mathrm{sgn}(\vec{y''} \cdot \vec{z}_{c_1,c_2})$, where $\vec{z}_{c_1,c_2} = (1, c_1, c_2)$.

In the protocol with 3 non-local boxes, labeled $(1, -1), (-1, 1), (-1, -1)$, Alice inputs 1 into the box labeled (c_1, c_2) if $(c_1, c_2) \neq (1, 1)$, and 0 into the other boxes. Bob inputs $(1 - \mathrm{sgn}(\vec{y''} \cdot \vec{z}_{m_1,m_2}) \cdot (\mathrm{sgn}(\vec{y''} \cdot \vec{z}_{1,1})))/2$ into the box labeled (m_1, m_2). Let the outputs of the non-local box labeled m be (a_m, b_m). Then Alice outputs $A = \alpha_0 \cdot (-1)^{\bigoplus_m a_m}$ and Bob outputs $B = \mathrm{sgn}(\vec{y''} \cdot \vec{z}_{1,1}) \cdot (-1)^{\bigoplus_m b_m}$.

□

6 Secure Function Evaluation

6.1 Honest-but-curious model

In the honest-but-curious model, it is well known that OT and AND are equivalent (up to a factor of 2).

Claim 3. *One AND may be simulated by one OT. One OT may be simulated by two ANDs. These simulations preserve security in the honest-but-curious model.*

In high level, a non-local box creates an additive share of an AND, however, as we have pointed out before, one needs to be careful about the timing issues. In what follows, we take advantage of the similarity of a non-local box and the AND primitive in order to characterize exactly the number of AND calls necessary for the secure computation of a function f.

As a starting point, we consider the most basic model, namely deterministic secure computation with ANDs in the honest-but-curious model. Beimel and Malkin [BM04] have shown that $AND(f) \leq 2^{|\mathcal{X}|}$. We show that it is characterized by the one-way communication complexity of f. Note that as Beimel and Malkin, we only consider perfect privacy. Extending the results to almost-perfect privacy is left as an open question.

Theorem 10. $AND(f) = 2^{D^{\rightarrow}(f)}$.

Proof. $[AND(f) \leq 2^{D^{\rightarrow}(f)}]$. Let P be a one-way communication protocol for f using $t = D^{\rightarrow}(f)$ bits of communication, where, on input x, Alice sends a message $m(x) \in \{0, 1\}^t$ to Bob and outputs $A(x)$, while, on input y, Bob outputs $B(y, m(x))$. We now a build a secure protocol for f using 2^t secure ANDs. We label the AND gates by a t-bit string i. Let $m = m(x)$. For the AND gate labeled i, Alice inputs 1 iff $m = i$, while Bob inputs $B(y, i)$. Let a_i be the outputs of the AND gates (received by Alice). Note that $a_m = B(y, m)$, and $a_i = 0$ for all $i \neq m$. It then suffices for Alice to output $A(x) \oplus a_m$. The correctness of the protocol is immediate. The privacy for Alice is trivial since Bob does not receive the output of the ANDs, and as a consequence no information from Alice. The privacy for Bob follows from the fact that the only possibly non-zero output that Alice receives from the ANDs is $a_m = B(y, m)$, which she can deduce from $f(x, y)$ and her input x.

$[D^{\rightarrow}(f) \leq \log(AND(f))]$. Let P be a secure protocol for f using $t = AND(f)$ AND gates. Beimel and Malkin showed that in the deterministic case, we can assume without loss of generality that there is no communication between Alice and Bob. In the protocol P, Alice and Bob input p_i and q_i, respectively, in the AND gate labeled $i \in [t]$, and Alice receives the output $a_i = p_i \cdot q_i$. Since Bob does not receive any information, his inputs to the AND gates only depend on his input y, that is, $q_i = q_i(y)$. We show that the same holds for Alice.

Let $\mathbf{a} = (a_1, \cdots, a_t)$ be the vector of outputs from the AND gates. For fixed x, since the protocol is deterministic, and Alice should only learn whether $f(x, y)$ is 0 or 1, she should only receive two possible vectors, say $\mathbf{a}^0(x)$ when $f(x, y) = 0$ and $\mathbf{a}^1(x)$ otherwise. Note that if there exists some $x_0 \in \mathcal{X}$ such that $f(x_0, y)$ is constant for all $y \in \mathcal{Y}$, say $f(x_0, y) = 0$, Alice only receives one possible vector $\mathbf{a}^0(x_0)$ when $x = x_0$. In that case, we can fix $\mathbf{a}^1(x_0)$ arbitrarily to any vector different from $\mathbf{a}^0(x_0)$. Let $m = m(x)$ be the first index such that $a^0_m(x) \neq a^1_m(x)$. For any $i < m(x)$, $a^0_i(x) = a^1_i(x)$, hence Alice knows in advance what outputs she will receive from the first $m(x) - 1$ gates. Therefore, Alice does not need these outputs (since she may infer them by herself) and we may assume without loss of generality that she inputs $p_i(x) = 0$ in the first $m(x) - 1$ gates. For the AND gate number m, $a^0_m(x) \neq a^1_m(x)$, so it has to be the case that $p_m(x) = 1$ (otherwise a_m is always 0). From the output of that gate, Alice already knows the value $f(x, y)$ (depending on whether the output is $a^0_m(x)$ or $a^1_m(x)$), so she does not need the outputs of the last $t - m(x)$ AND gates, and we can assume without loss of generality that she just inputs $p_i(x) = 0$ for all $i > m(x)$.

To summarize, we can always assume that Alice inputs $p_i(x) = 0$ in all AND gates, except for some index $i = m(x)$ where she inputs 1. For this AND gate, the output will therefore coincide with Bob's input $q_m(y)$. From the definition of $a^0_m(x)$ and $a^1_m(x)$, we then have for this output $q_m(y) = a^0_m(x)$ iff $f(x, y) = 0$, that is, in turn, $f(x, y) = q_m(y) \oplus a^0_m(x)$. We are now ready to build a one-way protocol for f. It suffices for Alice to compute the index of the relevant AND gate $m = m(x)$ and to send it to Bob. Then, Bob sets his output to $B(y, m) = q_m(y)$, while Alice sets hers to $A(x) = a^0_m(x)$. $\square$

One can say that this shows that for most functions, randomization is necessary in order to construct efficient protocols even in the honest-but-curious model.

6.2 Malicious model

As we said, the AND primitive cannot be used in the malicious model: indeed, a dishonest Alice may input 1 in all ANDs, and she obtains Bob's input for free, which still allows her to compute the AND. Therefore, for a dishonest Alice, each AND is just equivalent to a bit of communication (Bob sends his input to Alice), and this does not allow for unconditional secure computation. For this reason, in the malicious model we consider the OT primitive. Moreover, it is known that in the malicious model, deterministic secure computation is impossible [DM99], so we consider the case where Alice and Bob may use private coins and the protocol can have ε error.

Due to their non-signaling property, protocols using non-local boxes only and no communication, such as those presented in the previous sections, are trivially secure even against malicious players. Indeed, the non-signaling property implies that the view of the protocol by a possibly dishonest player is always independent from the actions of the other player. We show that certain such protocols may be transformed into protocols using OTs, namely the protocols where Alice and Bob use their non-local boxes in the same order. At this point, we don't know if this type of protocols are strictly weaker than general non-local box protocols. Nevertheless, our upper bounds in terms of communication complexity hold for such protocols as well (Corollary 5) and hence they translate into upper bounds on $OT_\varepsilon(f)$.

Theorem 11. *For any* $\varepsilon \geq 0$, $OT_\varepsilon(f) \leq NL^{\mathrm{ord}}_\varepsilon(f)$.

Proof. Let us consider a protocol for f using t non-local boxes in order and no communication. Let us denote $(p_1, \ldots, p_t)$ and $(a_1, \ldots, a_t)$ the inputs and outputs of Alice's non-local boxes and $(q_1, \ldots, q_t)$ and $(b_1, \ldots, b_t)$ the inputs and outputs of Bob's non-local boxes. The fact that they use these non-local boxes in order implies that p_i (and q_i) can only depend on the inputs and outputs of the first $(i - 1)$ non-local boxes but not on the remaining ones. Note that still Alice and Bob get their outputs immediately when they enter their inputs.

We now replace each non-local box with an OT starting from the first one, keeping the distribution of the view of the protocol exactly the same. Alice and Bob know how to pick the inputs to the first non-local box p_1, q_1 since they only depend on their inputs (x, y) and the randomness. To replace this non-local box, Alice picks a random bit r_1 and inputs $\{r_1, r_1 \oplus p_1\}$ to the OT box; Bob inputs q_1 and hence, his output becomes $r_1 \oplus p_1 \cdot q_1$. Finally, Alice and Bob set the outputs of the simulated non-local box to $a_1 = r_1$ and $b_1 = r_1 \oplus p_1 \cdot q_1$. The simulation of the distribution of the outputs of the non-local box is perfect, since a_1, b_1 are unbiased random bits and $a_1 \oplus b_1 = p_1 \cdot q_1$.

Alice and Bob continue with the simulation of the remaining non-local boxes until the end (Alice using a fresh private random bit for each NLB). Note that for each non-local box, Alice and Bob can compute the inputs p_i, q_i from exactly the correct distribution, since they only depend on the previous $(i-1)$ non-local boxes which have been perfectly simulated. Hence, at the end, we obtain a protocol for f with the same success probability as the original one. Note that this construction works only when the non-local boxes are used in order.

It remains to prove that the new protocol with Oblivious Transfer boxes that we constructed is still secure. Privacy for Bob is immediate since he only interacts with Alice through the OTs (there is no additional communication), and Alice obtains no output from the OTs. Privacy for Alice follows from the fact that the only information that Bob receives from Alice during the protocol is the outputs of the OTs, and that these output bits are independent from each other and from Alice's input (since Alice uses independent private random bits to generate her OT inputs). □

From the above theorem we can conclude that all the upper bounds that we had for the $NL_\varepsilon^{\mathrm{ord}}$ complexity (see Corollaries 2-5) translate into upper bounds for $OT_\varepsilon(f)$.

The construction used to replace a NLB by a OT is due to Wolf and Wullschleger [WW05]. In this reference, this construction is used to prove that OT is equivalent to NLB, but note that this is strictly speaking incorrect due to the different timing properties of OT and NLB, as pointed out in [BCU$^+$07].

We now turn our attention to lower bounds. In this case, we first note that our results are not general, i.e. they only hold for a special type of protocols that we call 'OT-optimal' secure protocols. An 'OT-optimal' secure protocol is one where the inputs to the function remain private as usual, but we also require that for all the OT calls, there is always an input that remains perfectly private throughout the protocol.

Intuitively, since we try to minimize the number of OTs that we use, it should be the case that these OT calls are really necessary, in the sense that one of the two inputs should remain secure throughout the protocol. If for example both inputs are revealed at some point during the protocol, then one may not use this OT at all, resulting into a protocol with fewer OT calls. Even though intuitively our definition seems natural, at this point, we do not know whether this assumption can be done without loss of generality. Extending our results to general protocols remains an open question.

In high level, we need to impose the condition that the messages of the players are as usual independent of their input (otherwise the messages would reveal information about the inputs that need to remain private) but they are also independent of the inputs to the OT boxes (otherwise a malicious player could get information about the inputs to the OT box, hence showing that the protocol does not use the minimum number of OTs).

Let us fix some notation. Consider a protocol for the secure computation of a function f, using communication and OT boxes. A denotes the messages from Alice to Bob; B Bob's messages, S and T Alice and Bob's inputs to the OT boxes and O the outputs of the OT boxes. Note that only Bob receives these outputs. We assume that at every round i of the protocol, A_i is Alice's message, B_i is Bob's message and $S_i = (S_i^0, S_i^1), T_i, O_i$ are the inputs and the output of the i-th OT box. (In some rounds we may not have communication, or the communication can proceed in several rounds; these cases can be handled in the same way as in the proof below.) $A_{[i]}, B_{[i]}, S_{[i]}, T_{[i]}, O_{[i]}$ is the concatenation of the first i messages of Alice, messages of Bob, inputs and outputs of the i-th first OT boxes respectively.

Definition 6. *$\widehat{OT}_\varepsilon(f)$ is the number of 2-1 Oblivious Transfer calls required to compute $f(x, y)$ in parity with perfect privacy and ε error over the players' private coins, assisted with (free) two-way communication, in the malicious model, subject to the additional conditions that for each i,*

$$\begin{aligned}
\mathrm{Prob}[A_i|A_{[i-1]}, B_{[i-1]}, S_{[i]}, x, r] &= \mathrm{Prob}[A_i|A_{[i-1]}, B_{[i-1]}, r], \\
\mathrm{Prob}[B_i|A_{[i]}, B_{[i-1]}, T_{[i]}, O_{[i]}, y, r] &= \mathrm{Prob}[B_i|A_{[i]}, B_{[i-1]}, r],
\end{aligned}$$

where r is the shared random string used in the protocol.

Let us see exactly what our definition says and how it is related to our intuitive definition of 'optimal' protocols. Let us consider the first condition (a similar discussion holds for the second condition, by swapping Alice and Bob's roles). We claim that the distribution of A_i conditioned on $(A_{[i-1]}, B_{[i-1]}, r)$ should be independent of $(S_{[i]}, x)$. Imagine that it is not the case. Then there exist two different strings $(S_{[i]}, x)$ and $(S_{[i]}, x)'$, such that the distribution of A_i conditioned on $(A_{[i-1]}, B_{[i-1]}, r)$ is different depending on whether Alice's inputs are $(S_{[i]}, x)$ or $(S_{[i]}, x)'$. This is a contradiction to some strong notion of privacy. Bob, knowing $(A_{[i-1]}, B_{[i-1]}, r)$ and receiving A_i, will get some information about whether Alice's inputs are $(S_{[i]}, x)$ or $(S_{[i]}, x)'$. This means that first, if these two strings differ in x, then Bob learns information about the input, which cannot happen; and second, if they differ in one of Alice's inputs to some OT box, then the malicious Bob could get information about both Alice's inputs. A malicious Bob can do this by for example picking an OT call at random and input a random bit into the OT box. With non-zero probability, this would be exactly the box for which he can get information about one input bit and with half probability this bit will be different than the one he learned from the OT box. Hence, we believe that our definition captures exactly the notion of an 'OT-optimal' protocol where the inputs x, y as well as the inputs to the OT boxes must remain secure throughout the protocol.

Theorem 12. *For any Boolean function f, $\widehat{OT}_\varepsilon(f) = \Omega(R_\varepsilon(f))$*

Proof. We want to show that even in the randomized case, communication doesn't help a lot. In other words, we want to show that $O(t)$ bits of communication are sufficient, where t is the number of OT boxes the players use. Recall that we assume that the protocol is optimal in the following sense: privacy is preserved if both the inputs of the players and one of the inputs to the Oblivious Transfer boxes remain secure throughout the protocol. We start with a perfectly secure protocol that uses t OT boxes where Bob receives the outputs, and arbitrary two-way communication between Alice and Bob. We show how to obtain a protocol using $O(t)$ bits of communication (and no OT boxes) by proceeding in four steps.

1. First, we show that the optimality conditions imply that we can entirely suppress the communication from Alice to Bob. We defer the proof of this part to the end of the proof. After this first step, we have a perfectly secure protocol using t OT boxes where Bob receives the outputs, no communication from Alice to Bob and arbitrary communication from Bob to Alice.

2. The second step is to invert the t OT boxes, meaning that we simulate each OT box where Bob receives the output by an OT box where Alice receives the output. It is well-known that this is possible if we add one bit of communication from Alice to Bob for each OT box [WW06].

   For completeness we include the description of the simulation. Let (s_0, s_1) be Alice's inputs to the OT box and t be Bob's index. At the end, Bob receives s_t. Alice and Bob simulate this by an OT box, where Bob inputs $(r, r \oplus t)$, for a uniformly random bit r, and Alice inputs $s_0 \oplus s_1$. Let a denote Alice's output. After this, Alice sends the bit $a \oplus s_0$ to Bob and Bob outputs $y = a \oplus s_0 \oplus r$. First, we have $y = r \oplus ((s_0 \oplus s_1) \cdot t) \oplus s_0 \oplus r = s_t$. Also, Alice does not learn anything about t, since there is no communication from Bob to Alice and Bob gets only one bit from Alice that is either s_0 or s_1.

   This step requires to reintroduce t bits of communication from Alice to Bob. Note that the simulation is such that the new protocol is still optimal, that is, it still satisfies the conditions of Definition 6, where Alice and Bob's roles are swapped. Hence, we now have a perfectly secure protocol using t OT boxes where Alice receives the outputs, t bits of communication from Alice to Bob and arbitrary communication from Bob to Alice.

3. The third step is to suppress the communication from Bob to Alice. For this step, we can reuse the analysis of step 1, since after step 2, Alice is the one who receives the output of the OT boxes, i.e. the situation is exactly the same as in step 1, with the roles of Alice and Bob reversed. After this, we end up with a perfectly secure protocol using t OT boxes where Alice receives the outputs, t bits of communication from Alice to Bob, and no communication from Bob to Alice.

4. Finally, Alice and Bob can simulate these t OT boxes by communicating $2t$ bits and hence we end up with a communication protocol of complexity $3t$. This concludes the proof of the theorem.

We now show how to perform the first step of the proof. The goal is to have Alice and Bob use their shared randomness in order to pick Alice's messages without her sending any bit. On the other hand, Bob is going to send the same messages to Alice as before and they will also use the same OT boxes. Bob can fix his private randomness in the beginning of the protocol. Alice is going to start with a uniform distribution on her private randomness and during the protocol she will update this distribution in order to remain consistent with the protocol up to that point.

We now describe the original protocol in more detail:

- Alice and Bob pick their private randomness r_A and r_B uniformly at random.
- For every round i of the protocol
  - Alice and Bob use an OT box with inputs S_i and T_i respectively and Bob receives output O_i. Alice's input to the i-th OT box is a fixed function of $(A_{[i-1]}, B_{[i-1]}, S_{[i-1]}, x, r, r_A)$ and Bob's input is a fixed function of $(A_{[i-1]}, B_{[i-1]}, T_{[i-1]}, O_{[i-1]}, y, r, r_B)$.
  - Alice computes her message A_i as a function of $(A_{[i-1]}, B_{[i-1]}, S_{[i]}, x, r, r_A)$ and sends it to Bob.
  - Bob computes his message B_i as a function of $(A_{[i]}, B_{[i-1]}, T_{[i]}, O_{[i]}, y, r, r_B)$ and sends it to Alice.

We look at the distribution induced by this protocol $\mathrm{Prob}[r_A, r_B, A, B, S, T, O|x, y, r]$ and have

$$\begin{aligned}
&\mathrm{Prob}[r_A, r_B, A, B, S, T, O|x, y, r] \\
=\ &\mathrm{Prob}[r_A] \cdot \mathrm{Prob}[r_B] \cdot \prod_{i\in[t]} \mathrm{Prob}[S_i|A_{[i-1]}, B_{[i-1]}, S_{[i-1]}, x, r, r_A] \\
&\cdot \mathrm{Prob}[T_i|A_{[i-1]}, B_{[i-1]}, T_{[i-1]}, O_{[i-1]}, y, r, r_B] \cdot \mathrm{Prob}[O_i|S_i, T_i] \\
&\cdot \mathrm{Prob}[A_i|A_{[i-1]}, B_{[i-1]}, S_{[i]}, x, r, r_A] \cdot \mathrm{Prob}[B_i|A_{[i]}, B_{[i-1]}, T_{[i]}, O_{[i]}, y, r, r_B].
\end{aligned}$$

We now give a new protocol where Alice and Bob use their shared randomness to simulate Alice's messages. The distribution remains exactly the same as in the original protocol. In order for this to hold, Alice needs to "update" her private randomness to retain consistency. Here is the new protocol:

- Alice and Bob pick their private randomness r_A and r_B uniformly at random.
- For every round i of the protocol
  - Alice and Bob use an OT box with inputs S_i, T_i respectively and Bob receives output O_i.
    Alice picks her input S_i according to $\mathrm{Prob}[S_i|A_{[i-1]}, B_{[i-1]}, S_{[i-1]}, x, r]$. Bob's input is the same functions of $(A_{[i-1]}, B_{[i-1]}, T_{[i-1]}, O_{[i-1]}, y, r, r_B)$ as in the original protocol.
  - Alice and Bob simulate Alice's message by sampling from the distribution $\mathrm{Prob}[A_i|A_{[i-1]}, B_{[i-1]}, r]$, i.e. they pick a next message from all messages that are consistent in the original protocol with the shared randomness r and the transcript so far, averaged over the private randomness r_A for Alice. The key point is that due to the optimality condition in Definition 6, Bob also knows the distribution of Alice's message A_i when averaged over r_A, since it does not depend on $(S_{[i]}, x)$.
  - Bob computes his message B_i as the same function of $(A_{[i]}, B_{[i-1]}, T_{[i]}, O_{[i]}, y, r, r_B)$ as in the original protocol and sends it to Alice.
  - Alice "updates" her private randomness by picking r_A according to $\mathrm{Prob}[r_A|A_{[i]}, B_{[i]}, S_{[i]}, x, r]$.

We need to show that the distribution corresponding to the above protocol is exactly the same as in the original protocol and also that this a well-defined procedure. We have for the new protocol:

$$\begin{aligned}
&\mathrm{Prob}[r_A, r_B, A, B, S, T, O|x, y, r] = \mathrm{Prob}[r_A] \cdot \mathrm{Prob}[r_B] \\
&\quad \cdot \prod_{i\in[t]} \mathrm{Prob}[S_i|A_{[i-1]}, B_{[i-1]}, S_{[i-1]}, x, r] \cdot \mathrm{Prob}[T_i|A_{[i-1]}, B_{[i-1]}, T_{[i-1]}, O_{[i-1]}, y, r, r_B] \cdot \mathrm{Prob}[O_i|S_i, T_i] \\
&\quad \cdot \mathrm{Prob}[A_i|A_{[i-1]}, B_{[i-1]}, r] \cdot \mathrm{Prob}[B_i|A_{[i]}, B_{[i-1]}, T_{[i]}, O_{[i]}, y, r, r_B] \cdot \frac{\mathrm{Prob}[r_A|A_{[i]}, B_{[i]}, S_{[i]}, x, r]}{\mathrm{Prob}[r_A|A_{[i-1]}, B_{[i-1]}, S_{[i-1]}, x, r]}
\end{aligned}$$

Note that as it should be the distribution of r_A after the ℓ-th round is exactly

$$\mathrm{Prob}[r_A] \cdot \prod_{i \in [\ell]} \frac{\mathrm{Prob}[r_A|A_{[i]}, B_{[i]}, S_{[i]}, x, r]}{\mathrm{Prob}[r_A|A_{[i-1]}, B_{[i-1]}, S_{[i-1]}, x, r]} = \mathrm{Prob}[r_A|A_{[\ell]}, B_{[\ell]}, S_{[\ell]}, x, r]$$

We now show that the distributions which correspond to the two protocols are the same. It is easy to see that we need to prove the following fact

$$\begin{aligned}&\mathrm{Prob}[S_i|A_{[i-1]}, B_{[i-1]}, S_{[i-1]}, x, r, r_A] \cdot \mathrm{Prob}[A_i|A_{[i-1]}, B_{[i-1]}, S_{[i]}, x, r, r_A] = \\ &\quad \mathrm{Prob}[S_i|A_{[i-1]}, B_{[i-1]}, S_{[i-1]}, x, r] \cdot \mathrm{Prob}[A_i|A_{[i-1]}, B_{[i-1]}, r] \cdot \frac{\mathrm{Prob}[r_A|A_{[i]}, B_{[i]}, S_{[i]}, x, r]}{\mathrm{Prob}[r_A|A_{[i-1]}, B_{[i-1]}, S_{[i-1]}, x, r]}\end{aligned}$$

We have

$$\begin{aligned}&\mathrm{Prob}[S_i|A_{[i-1]}, B_{[i-1]}, S_{[i-1]}, x, r, r_A] \cdot \mathrm{Prob}[A_i|A_{[i-1]}, B_{[i-1]}, S_{[i]}, x, r, r_A] \\ &= \mathrm{Prob}[S_i, A_i|A_{[i-1]}, B_{[i-1]}, S_{[i-1]}, x, r, r_A] = \frac{\mathrm{Prob}[S_i, A_i, r_A|A_{[i-1]}, B_{[i-1]}, S_{[i-1]}, x, r]}{\mathrm{Prob}[r_A|A_{[i-1]}, B_{[i-1]}, S_{[i-1]}, x, r]} \\ &= \mathrm{Prob}[S_i, A_i|A_{[i-1]}, B_{[i-1]}, S_{[i-1]}, x, r] \cdot \frac{\mathrm{Prob}[r_A|A_{[i]}, B_{[i-1]}, S_{[i]}, x, r]}{\mathrm{Prob}[r_A|A_{[i-1]}, B_{[i-1]}, S_{[i-1]}, x, r]} \\ &= \mathrm{Prob}[S_i|A_{[i-1]}, B_{[i-1]}, S_{[i-1]}, x, r] \cdot \mathrm{Prob}[A_i|A_{[i-1]}, B_{[i-1]}, r] \cdot \frac{\mathrm{Prob}[r_A|A_{[i]}, B_{[i]}, S_{[i]}, x, r]}{\mathrm{Prob}[r_A|A_{[i-1]}, B_{[i-1]}, S_{[i-1]}, x, r]}\end{aligned}$$

For the last equation we used first that $\mathrm{Prob}[r_A|A_{[i]}, B_{[i-1]}, S_{[i]}, x, r] = \mathrm{Prob}[r_A|A_{[i]}, B_{[i]}, S_{[i]}, x, r]$ since B_i is independent of r_A (for fixed $A_{[i]}, S_{[i]}$) and more importantly that we have $\mathrm{Prob}[A_i|A_{[i-1]}, B_{[i-1]}, S_{[i]}, x, r] = \mathrm{Prob}[A_i|A_{[i-1]}, B_{[i-1]}, r]$. The last equality comes from the privacy of the protocol and the fact that it is optimal.

Now, let us make sure that all these probabilities are non-zero. This follows from the privacy of the protocol. Let's say that after the $(i-1)$-th round Alice has been able to update her private randomness to a consistent r_A. In the next round, Alice picks input S_i for the OT box from the distribution of all inputs S_i that are consistent for some r_A and Alice and Bob pick a message A_i as Alice's next message. The distribution from which Alice and Bob picked A_i is from all messages for which for shared randomness r, there exists an input x and inputs $S_{[i]}$ such that there exists a string r_A so that A_i is consistent with $(x, r, r_A, A_{[i-1]}, S_{[i]})$. From privacy, if there exists an x and inputs $S_{[i]}$ for which the transcript $A_{[i]}$ is consistent for some r_A, then it has to be consistent for all inputs x and all inputs $S_{[i]}$ and for some other r_A. Otherwise, Bob will gain information about x or $S_{[i]}$. Hence, there will always be a choice of r_A which is consistent with the protocol. This finishes the first step, and consequently, the whole proof. □

7 Conclusion and open questions

We have shown various upper and lower bounds on non-local box complexity, and shown how the upper bounds could be translated into bounds for secure function evaluation. We have also shown how to simulate quantum correlations arising from binary measurements on bipartite entangled states using 3 non-local boxes. Note that combining these last two results also implies that such quantum correlations may also be simulated using 3 OT boxes. The advantage is that, while non-local boxes may not be actually realized (due to their violation of Tsirelson's bound), OT boxes may be implemented under computational assumptions. Note that such a simulation with OT boxes breaks the timing properties of the EPR experiment, but this is unavoidable when simulating quantum correlations using classical resources, due to the violation of Bell inequalities. Moreover, contrary to a simulation with communication such as in [RT09], using OT boxes preserves the cryptographic properties of the experiment, that is, Alice does not learn anything about Bob's measurement (and vice versa).

During our investigations, we have come across a series of interesting open questions.

- For randomized non-local box complexity in parallel, can we remove the XOR restriction, that is, is the case that $NL_\varepsilon^{||}(f) \approx NL_\varepsilon^{||,\oplus}(f) = \varepsilon{-}\mathrm{rank}_{\mathbb{GF}_2}(M_f)$? The proof for the deterministic does not carry over because of the inherent randomness of the non-local boxes, which could be used to save on the number of non-local boxes when some error probability is authorized.

- While the Disjointness function provides an example of exponential gap between parallel and general deterministic non-local box complexity, the gap disappears in the randomized model. Is it always the case that parallel and general randomized non-local box complexities are polynomially related?

- In general, non-local boxes could be used in a different order on Alice and Bob's side. Does this provide any advantage, that is, are there functions for which $NL_\varepsilon(f) < NL_\varepsilon^{\mathrm{ord}}(f)$?

- As for secure function evaluation, we proved that the communication complexity is a lower bound on OT complexity only under some optimality assumption. Can this assumption be made without loss of generality?

- Finally, another interesting question is whether we can prove an analogue of Theorem 12 for non-local boxes. Ideally, we would like to prove that for secure computation with ordered non-local boxes, communication does not help. Indeed, due to the reduction from ordered non-local box protocols to OT protocols and vice versa, this would imply that $NL_\varepsilon^{\mathrm{ord}}(f)$ is exactly $OT_\varepsilon(f)$, and not just an upper bound. This would of course provide even more motivation to study non-local box complexity in the context of secure function evaluation. Note that working with non-local boxes instead of OT boxes provides a few advantages. First, protocols using non-local boxes (and no communication) are necessarily secure, even in the malicious model. Second, contrary to OT protocols, such non-local box protocols do not require private randomness (except the inherent randomness of the non-local boxes) to ensure security in this model.

Acknowledgements

We would like to thank Troy Lee and Falk Unger for pointing out the deterministic protocol for Disjointness. We thank Ronald de Wolf for suggesting a randomized protocol for Disjointness with bias $1/\log(n)$, using Valiant-Vazirani. J. R. would like to thank Joe Kilian for useful comments.

References

[BBL+06] G. Brassard, H. Buhrman, N. Linden, A. A. Methot, A. Tapp, and F. Unger. Limit on nonlocality in any world in which communication complexity is not trivial. *Physical Review Letters*, 96(25):250401, 2006.

[BCU+07] H. Buhrman, M. Christandl, F. Unger, S. Wehner, and A. Winter. Implications of superstrong nonlocality for cryptography. *Proc. of the Royal Society A*, 462:1919–1932, 2007.

[Bd01] H. Buhrman and R. de Wolf. Communication complexity lower bounds by polynomials. In *Proc. of the 16th Annual IEEE Conference on Computational Complexity*, pages 120–130, 2001.

[Bea96] D. Beaver. Correlated pseudorandomness and the complexity of private computations. In *Proc. of the 28th Annual ACM Symposium on the Theory of Computing*, pages 479–488, 1996.

[BM04] A. Beimel and T. Malkin. A quantitative approach to reductions in secure computation. In Springer Berlin / Heidelberg, editor, *Proc. of the First Theory of Cryptography Conference*, volume 2951, pages 238–257, 2004.

[BM05] A. Broadbent and A. A. Méthot. On the power of non-local boxes. *Theoretical Computer Science*, 358(1):3–14, 2005.

[BMM99] A. Beimel, T. Malkin, and S. Micali. The all-or-nothing nature of two-party secure computation. In *Advances in Cryptology - CRYPTO '99*, pages 80–97, 1999.

[BOGW88] M. Ben-Or, S. Goldwasser, and A. Wigderson. Completeness theorems for noncryptographic faulttolerant distributed computations. In *Proc. of the 20th Annual ACM Symposium on the Theory of Computing*, pages 1–10, 1988.

[BP05] J. Barrett and S. Pironio. Popescu-Rohrlich correlations as a unit of nonlocality. *Physical Review Letters*, 95:140401, 2005.

[CCD88] D. Chaum, C. Crepeau, and I. Damgard. Multiparty unconditionally secure protocols. In *Proc. of the 20th Annual ACM Symposium on the Theory of Computing*, pages 11–19, 1988.

[CGMP05] N. J. Cerf, N. Gisin, S. Massar, and S. Popescu. Simulating Maximal Quantum Entanglement without Communication. *Physical Review Letters*, 94(22):220403, 2005.

[CHSH69] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt. Proposed Experiment to Test Local Hidden-Variable Theories. *Physical Review Letters*, 23:880–884, 1969.

[CK91] B. Chor and E. Kushilevitz. A zero-one law for boolean privacy. *SIAM J. on Discrete Mathematics*, 4(1):36–47, 1991.

[DLR05] J. Degorre, S. Laplante, and J. Roland. Simulating quantum correlations as a distributed sampling problem. *Physical Review A*, 72(062314), 2005.

[DLR07] J. Degorre, S. Laplante, and J. Roland. Classical simulation of traceless binary observables on any bipartite quantum state. *Physical Review A*, 75(012309), 2007.

[DM99] Y. Dodis and S. Micali. Lower bounds for oblivious transfer reductions. In *EUROCRYPT*, pages 42–55, 1999.

[FW09] M. Forster and S. Wolf. The universality of non-local boxes. In AIP Conference Proceedings, editor, *Proc. of the Ninth International Conference on Quantum Communication, Measurement and Computing*, volume 1110, pages 117–120, 2009.

[Gro97] V. Grolmusz. On the power of circuits with gates of low l_1 norms. *Theoretical Computer Science A*, 188:117–127, 1997.

[GV88] O. Goldreich and R. Vainish. How to solve any protocol problem - an efficiency improvement. In *Advances in Cryptology - CRYPTO '87*, pages 73–86, 1988.

[IM02] K. Iwama and H. Morizumi. An explicit lower bound of $5n - o(n)$ for boolean circuits. In *Proc. of the 27th International Symposium on Mathematical Foundations of Computer Science*, pages 353–364, 2002.

[JM05] N. S. Jones and L. Masanes. Interconversion of nonlocal correlations. *Physical Review A*, 72(5):052312, 2005.

[Kil88] J. Kilian. Basing cryptography on oblivious transfer. In *Proc. of the 20th Annual ACM Symposium on the Theory of Computing*, pages 20–31, 1988.

[Kil91] J. Kilian. A general completeness theorem for two-party games. In *Proc. of the 23th Annual ACM Symposium on the Theory of Computing*, pages 553–560, 1991.

[Kil00] J. Kilian. More general completeness theorems for two-party games. In *Proc. of the 32nd Annual ACM Symposium on the Theory of Computing*, pages 316–324, 2000.

[KKMO00] J. Kilian, E. Kushilevitz, S. Micali, and R. Ostrovsky. Reducibility and completeness in private computations. *SIAM J. Comput.*, 28(4):1189–1208, 2000.

[KN97] E. Kushilevitz and N. Nisan. *Communication complexity*. Cambridge University Press, New York, 1997.

[KNR99] I. Kremer, N. Nisan, and D. Ron. On randomized one-way communication complexity. *Computational Complexity*, 8(1):21–49, 1999.

[Kre95] I. Kremer. Quantum communication. Master's thesis, The Hebrew University of Jerusalem, 1995.

[Kus92] E. Kushilevitz. Privacy and communication complexity. *SIAM J. on Discrete Mathematics*, 5(2):273–284, 1992.

[LR01] O. Lachish and R. Raz. Explicit lower bound of 4.5n - o(n) for boolean circuits. In *Proc. of the 36th annual ACM symposium on Theory of computing*, pages 399–408, 2001.

[LS08] N. Linial and A. Shraibman. Lower bounds in communication complexity based on factorization norms. *Random Structures and Algorithms*, 2008.

[Mau92] T. Maudlin. Bell's inequality, information transmission, and prism models. In *Biennal Meeting of the Philosophy of Science Association*, pages 404–417, 1992.

[New91] I. Newman. Public vs. private coin flips in one round communication games. *Information Processing Letters*, 39(2):67–71, 1991.

[NN01] M. Naor and K. Nissim. Communication preserving protocols for secure function evaluation. In *Proc. of the 33rd Annual ACM Symposium on the Theory of Computing*, 2001.

[PR94] S. Popescu and D. Rohrlich. Causality and nonlocality as axioms for quantum mechanics. *Foundations of Physics*, pages 379–385, 1994.

[PR97] S. Popescu and D. Rohrlich. Action and passion at a distance. In M. Horne R. S. Cohen and J. Stachel, editors, *Potentiality, Entanglement, and Passion-at-a-Distance*, pages 197–206, Dordrecht, 1997. Kluwer Academic Publishers.

[RP96] D. Rohrlich and S. Popescu. Nonlocality as an axiom for quantum theory. In A. Mann and M. Revzen, editors, *The Dilemma of Einstein, Podolsky and Rosen – 60 Years Later*, volume 12 of *Annals of the Israel Physical Society*, page 152. Institute of Physics Publishing, 1996.

[RT09] O. Regev and B. Toner. Simulating quantum correlations with finite communication. *SIAM J. Comput.*, 39(4):1562–1580, 2009.

[SZ08] Y. Shi and Y. Zhu. Tensor norms and the classical communication complexity of bipartite quantum measurements. *SIAM J. Comput.*, 38(3):753–766, 2008.

[Tsi85] B. S. Tsirelson. Problems of the theory probability distributions ix. *Zapiski Math. Inst. Steklov (LOMI)*, 142:174–194, 1985. English translation in Quantum analogues of the Bell inequalities. The case of two spatially separated domains, J. Soviet Math. 36, 557–570 (1987).

[vD05] W. van Dam. Implausible Consequences of Superstrong Nonlocality. Technical Report quant-ph/0501159, arXiv e-Print archive, 2005.

[WW05] S. Wolf and J. Wullschleger. Oblivious transfer and quantum non-locality. In *Proc. of the International Symposium on Information Theory*, pages 1745–1748, 2005.

[WW06] S. Wolf and J. Wullschleger. Oblivious transfer is symmetric. In *EUROCRYPT*, pages 222–232, 2006.

[Yao79] A. C.-C. Yao. Some complexity questions related to distributive computing. In *Proc. of the 11th annual ACM Symposium on Theory of Computing*, pages 209–213, 1979.