Algebra of Nonlocal Boxes and the Collapse of Communication Complexity

Pierre Botteron, Anne Broadbent, Reda Chhaibi, Ion Nechita, Clément Pellegrini

Algebra of Nonlocal Boxes and the Collapse of Communication Complexity

Pierre Botteron [1], Anne Broadbent [2,3], Reda Chhaibi [1,4], Ion Nechita [5], and Clément Pellegrini[1]

[1]Institut de Mathématiques de Toulouse, Université de Toulouse, France.
[2]Nexus for Quantum Technologies, University of Ottawa, Canada.
[3]Department of Mathematics and Statistics, University of Ottawa, Canada.
[4]Institut Universitaire de France (IUF), France.
[5]Laboratoire de Physique Théorique, Université de Toulouse, CNRS, France.

Communication complexity quantifies how difficult it is for two distant computers to evaluate a function $f(X, Y)$, where the strings X and Y are distributed to the first and second computer respectively, under the constraint of exchanging as few bits as possible. Surprisingly, some nonlocal boxes, which are resources shared by the two computers, are so powerful that they allow to *collapse* communication complexity, in the sense that any Boolean function f can be correctly estimated with the exchange of only one bit of communication. The Popescu-Rohrlich (PR) box is an example of such a collapsing resource, but a comprehensive description of the set of collapsing nonlocal boxes remains elusive.

In this work, we carry out an algebraic study of the structure of wirings connecting nonlocal boxes, thus defining the notion of the "product of boxes" $\mathtt{P} \boxtimes \mathtt{Q}$, and we show related associativity and commutativity results. This gives rise to the notion of the "orbit of a box", unveiling surprising geometrical properties about the alignment and parallelism of distilled boxes. The power of this new framework is that it allows us to prove previously-reported numerical observations concerning the best way to wire consecutive boxes, and to numerically and analytically recover recently-identified noisy PR boxes that collapse communication complexity for different types of noise models.

Nonlocal boxes (NLBs) were introduced by Popescu and Rohrlich in 1994 as a theoretical generalization of quantum correlations [41]. When Alice and Bob share a pair of entangled states $|\Psi\rangle$, each of them can choose to measure their state in a certain basis depending on some "instructions" $x, y \in \{0, \dots, p\}$, and then each of them can encode their outcomes in respectively $a, b \in \{0, \dots, q\}$. Similarly, a two-party NLB is a "black box" shared between Alice and Bob, with some inputs x, y and some outputs a, b, and with the rule that Alice has access only to the left part and Bob only to the right part, see Figure 1. This way, we only study the statistics produced by the "hidden state" inside of the box and not the physical theory describing that state, which is the reason why nonlocal boxes are said to be *device-independent*. In this work, we consider one of the simplest scenarios, the CHSH scenario named after Clauser, Horne, Shimony, and Holt [16], where $p = q = 1$ and with two parties Alice and Bob (for more general scenarios, see [4, 5, 14, 20, 44]).

Generally, nonlocal boxes are non-signalling, meaning that they respect the relativistic constraint of no faster-than-light communication between parties (although there is a recent interest for partially-signalling scenarios [44]). These *non-signalling boxes* form a set $\mathcal{NS}$ defined by the

Pierre Botteron: pierre.botteron@math.univ-toulouse.fr
Anne Broadbent: abroadbe@uottawa.ca
Reda Chhaibi: reda.chhaibi@math.univ-toulouse.fr
Ion Nechita: ion.nechita@univ-tlse3.fr
Clément Pellegrini: clement.pellegrini@math.univ-toulouse.fr

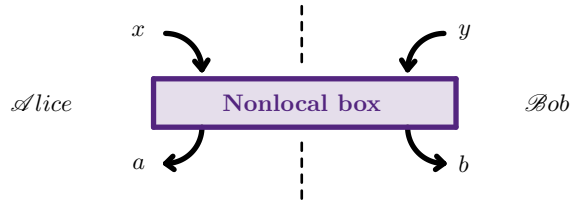


Figure 1: Representation of a nonlocal box.

following non-negativity and normalization equations, which ensure $\mathtt{P}$ is a well-defined conditional probability distribution:

$$\forall a,b,x,y \in \{0,1\}, \qquad \mathtt{P}(a,b\,|\,x,y) \geq 0 \qquad \text{and} \qquad \sum_{a,b} \mathtt{P}(a,b\,|\,x,y) = 1\,, \tag{1}$$

together with the conditions that the marginals of each party are independent of the other party's question:

$$\forall a,x \in \{0,1\}, \qquad \sum_{b\in\{0,1\}} \mathtt{P}(a,b\,|\,x,0) = \sum_{b\in\{0,1\}} \mathtt{P}(a,b\,|\,x,1) =: \mathtt{P}(a\,|\,x)\,, \tag{2}$$

$$\forall b,y \in \{0,1\}, \qquad \sum_{a\in\{0,1\}} \mathtt{P}(a,b\,|\,0,y) = \sum_{a\in\{0,1\}} \mathtt{P}(a,b\,|\,1,y) =: \mathtt{P}(b\,|\,y)\,. \tag{3}$$

The physical interpretation of Equations (2) and (3) is that Alice and Bob are space-like separated, so it would take more time for a light ray to move from Bob to Alice than the time needed for Alice to do her protocol and to receive her output of the box, and therefore Alice's marginal does not depend on Bob's input.

The best-known example of a non-signalling box is the $\mathtt{PR}$ box, named after Popescu and Rohrlich [41]. This box is designed to perfectly win at the $\mathtt{CHSH}$-game [16], *i.e.* the box produces outputs a,b such that $a \oplus b = xy$ with probability 1, where the symbol "$\oplus$" stands for the sum modulo 2.[1] More precisely, given an input pair (x,y), there are two possibilities for the outputs: either $(a=0, b=xy)$ or $(a=1, b=xy\oplus 1)$, each being output with probability $1/2$ by the $\mathtt{PR}$ box. In fact, it is possible to show that the $\mathtt{PR}$ box thus defined is the only box of $\mathcal{NS}$ that perfectly wins at the $\mathtt{CHSH}$-game. Let us give two other examples of boxes: (i) the *fully mixed* box $\mathtt{I}$, which outputs purely random bits a and b; and (ii) the deterministic boxes $\mathtt{P}_0, \mathtt{P}_1$ that always output $(0,0),(1,1)$ independently of the inputs. All these boxes are in $\mathcal{NS}$ and can be written as conditional probability distributions:

$$\begin{aligned} \mathtt{PR}\big(a,b\,|\,x,y\big) &:= \tfrac{1}{2}\mathbb{1}_{a\oplus b=xy}\,, & \mathtt{I}\big(a,b\,|\,x,y\big) &:= \tfrac{1}{4}\,, \\ \mathtt{P}_0\big(a,b\,|\,x,y\big) &:= \mathbb{1}_{a=b=0}\,, & \mathtt{P}_1\big(a,b\,|\,x,y\big) &:= \mathbb{1}_{a=b=1}\,, \end{aligned} \tag{4}$$

where the symbol "$\mathbb{1}$" stands for the indicator function. Two other important classes of correlations are the *local* set $\mathcal{L}$ and the *quantum* (tensor) set $\mathcal{Q}$, with strict inclusions $\mathcal{L} \subsetneq \mathcal{Q} \subsetneq \mathcal{NS}$, defined as:

$$\mathcal{L} = \left\{ \mathtt{P}(a,b\,|\,x,y) = \int_\lambda \mathbb{P}_A(a\,|\,x,\lambda)\,\mathbb{P}_B(b\,|\,y,\lambda)\,\mu(\lambda) \;:\; \mathbb{P}_A,\mathbb{P}_B,\mu \text{ are probability distributions} \right\},$$

$$\mathcal{Q} = \left\{ \mathtt{P}(a,b\,|\,x,y) = \langle\Psi|\Big(E_a^x \otimes F_b^y\Big)|\Psi\rangle \;:\; \begin{array}{l} \{E_a^x\}_a, \{F_b^y\}_b \text{ are POVMs over Hilbert spaces } \mathcal{H}_A, \mathcal{H}_B\,, \\ |\Psi\rangle \text{ is a vector of } \mathcal{H}_A \otimes \mathcal{H}_B \text{ of norm 1} \end{array} \right\}.$$

For more details on correlation sets and their separation, see [29].

Interestingly, some nonlocal boxes collapse what is called *communication complexity* (CC), a notion introduced by Yao in [48] and reviewed in [30, 43] that quantifies the difficulty of performing

[1] From a computing perspective, the $\mathtt{PR}$ box turns a distributed AND (xy) into a distributed XOR ($a\oplus b$), where "distributed" means that the computation requires both parties (Alice knows x and a, and Bob knows y and b).

a distributed computation. Say we want to evaluate a Boolean function $f : \{0,1\}^n \times \{0,1\}^m \to \{0,1\}$ using two distant computers, where the first computer receives as input $X \in \{0,1\}^n$, and the other computer receives as input $Y \in \{0,1\}^m$. The CC of f is then defined as the minimal number of bits that the computers need to communicate in order for the first computer to output the value $f(X,Y)$. For instance, when $n = m = 2$, $X = (x_1, x_2)$, $Y = (y_1, y_2)$, the CC of $f_1 := x_1 \cdot (y_1 \oplus y_2)$ equals 1, using the communication bit $y_1 \oplus y_2$, whereas it is possible to show that the CC of $f_2 := x_1 \cdot y_1 \oplus x_2 \cdot y_2$ equals 2, using communication bits y_1 and y_2; therefore f_2 is more complex than f_1 in the sense of CC. Yao also introduced in [48] a probabilistic version of CC, in which the computers can access shared randomness, and where for all X, Y the first computer has to output the correct value $f(X,Y)$ with probability at least $p > 1/2$ (p being independent of X and Y). Now, if a nonlocal box $\mathtt{P}$ is used in the protocol to compute the value $f(X,Y)$, we say that the box $\mathtt{P}$ *collapses* communication complexity if there exists a fixed $p > 1/2$ for which *any* Boolean function f, with arbitrary input size, can be correctly computed with only *one* bit of communication and probability p. In this definition, an arbitrary number of copies of the box $\mathtt{P}$ can be used in the protocol. Such a collapse is strongly believed to be unachievable in Nature since it would imply the absurdity that a single bit of communication is sufficient to distantly estimate any value of any Boolean function f [6, 10, 13, 18]. For more details on the link between nonlocal boxes and communication complexity, see [7].

Open question. Among the four examples of boxes listed in Equation (4), only the PR box collapses communication complexity [18], meaning that this box is very "powerful". We also know that some noisy versions of the PR box collapse CC for different types of noise [8, 10, 11, 13, 23]. On the other hand, we know that quantum correlations do *not* collapse communication complexity [17], and neither does a slightly wider set named "almost quantum correlations" [36]. To this day, the question is still open whether the remaining non-signalling boxes are collapsing, meaning that there is still a gap to be filled. We refer to [8, Fig. 2] for a figure that summarizes the situation.

Recent results. Two recent papers make progress on the question above. In [11] Brito, Moreno, Rai, and Chaves study correlation distillation in *quantum voids* [42], which are subsets of a face of $\mathcal{NS}$ where all nonlocal points are non-quantum. They prove strong distillation properties in 1- and 2- dimensional quantum voids and deduce that these regions collapse communication complexity, which partially answers the open question. More recently, in [23], Eftaxias, Weilenmann, and Colbeck propose a *sequential algorithm* to find a suitable sequence of wirings to collapse communication complexity, where a *wiring* is defined as a connection between boxes that allows the creation of a new box out of copies of a box (see Section 1). This allows us to numerically determine a collapsing region of nonlocal boxes, which is again a partial answer to the open question.

Our Results. We provide a new mathematical framework and algorithms in working towards addressing the open question. The ideas are based in part on the M.Sc. thesis of one of the authors [7].

(i) We introduce a new framework that we call the *algebra of boxes*. After recalling the definition of a wiring W, we use it to introduce a *product of boxes* $\mathtt{P} \boxtimes_{\mathsf{W}} \mathtt{Q}$. This leads to a natural embedding of the non-signalling set $\mathcal{NS}$ in an algebra, which we call an *algebra of boxes* and for which we characterize associativity and commutativity (see Proposition 7). This gives an algebraic perspective on protocols for correlation distillation—for instance, the non-associativity of the algebra of boxes tells us that the order in which the boxes are wired matters.

(ii) This framework gives rise to the fascinating notion of what we call the *orbit of a box*. The orbit of $\mathtt{P} \in \mathcal{NS}$ is roughly the set of all possible boxes that can be produced by wiring arbitrarily many copies of $\mathtt{P}$. This allows interesting visualizations of the hidden structure of boxes (see Figure 7), and surprisingly we observe that these orbits satisfy strong alignment and parallelism properties as shown in Theorem 10 and Corollary 11. Moreover, we derive the expression of the highest CHSH-valued box of the (tilted) orbit in Theorem 13, which explains the numerical observation reported in [23, Supplementary Material, II], and for which we derive an insightful linear-time algorithm that is exponentially more efficient compared to the

naive exponential-time computation of the entire orbit. In addition, we recover in Theorem 16 a similar result as in [23] stating that those methods lead to finding collapsing boxes via the recursive application of the multiplication $\cdot \boxtimes \mathtt{P}$ on the right.

(iii) We provide algorithms in our GitHub page [9] for the following task: given a box $\mathtt{P}$ that we want to show is collapsing, find an appropriate wiring W such that the orbit contains a collapsing box. The idea is to repeat several times in parallel a variant of the Gradient Descent Algorithm in order to find the most appropriate wiring W. These algorithms allow us to recover in Figure 10 similar new collapsing areas as in [23, Figure 3].

(iv) In Theorems 23 and 29, we show that our framework also allows us to recover some analytical results with a new proof based on the algebra of boxes: some triangles in the boundary of $\mathcal{NS}$ are collapsing [11]. To that end, algorithms of (iii) above were performed in order to identify a convenient wiring. Moreover, in Corollary 27, we recover a result from [42] with a new proof, based on communication complexity, showing that the triangle joining $\mathtt{PR}, \mathtt{P}_0, \mathtt{P}_1$ is a "quantum void".

Further Comparison with Recent Results. A contribution of our work is in providing a new algebraic framework for a unified perspective on three recent results [11, 23, 42]. Compared to [23], our work concurrently and independently[2] derives the existence of new collapsing boxes using right multiplication (Theorem 16) and we report a similar numerical result (Figure 10). However, as detailed in Remark 19, we note that the methods complement each other: as the authors of [23], we implement an algorithm finding a non-constant sequence of wirings, but we also implement an algorithm finding a constant sequence of wirings for analytical purposes. Moreover, as mentioned above, our Theorem 13 analytically proves a numerical observation reported in [23, Supplementary Material, II]. As for [11], we recover some of their results in Theorems 23 and 29, and although our proof reproduces previously-known analytical collapsing areas, we view our contribution as a new approach based on the algebra of boxes. Regarding [42], we recover one of their result in Corollary 27 with a new proof, based on communication complexity, showing that the triangle joining $\mathtt{PR}, \mathtt{P}_0, \mathtt{P}_1$ is a "quantum void". In summary, our new algebraic viewpoint unifies the results of [23] (see Subsection 4.1), [11] (see Subsection 4.2), and [42] (see Subsection 4.2).

Structure. This work is divided into four sections, a conclusion, and some appendices:

- Section 1: given a wiring W, we introduce the product of nonlocal boxes $\mathtt{P} \boxtimes_{\mathsf{W}} \mathtt{Q}$ and we study the new framework of *algebra of boxes*.
- Section 2: we define the notion of the *orbit* of a box $\mathtt{P}$, which consists of all boxes produced using copies of $\mathtt{P}$ and the product $\boxtimes$ previously defined, and we investigate its surprising geometric structure.
- Section 3: we present algorithms for the following task: given a box $\mathtt{P}$ that we want to show is collapsing, find an appropriate wiring W such that the orbit contains a collapsing box. These algorithms are based on Gradient Descents methods and are entirely accessible via our GitHub page [9].
- Section 4: we find collapsing boxes in two different ways: (i) numerically, using the algorithms of the previous section, and (ii) analytically, using the algebra of boxes and the orbit of a box.
- Conclusion: we conclude with some discussion, open problems, and avenues for future work.
- Appendices: we complete this work with supplementary figures and proofs.

1 Algebra of Boxes

The set of non-signalling boxes $\mathcal{NS}$ is the compact convex subset of the vector space $\mathcal{B} = \{\mathtt{P} : \{0,1\}^4 \to \mathbb{R}\}$ satisfying Equations (1), (2), (3). In this section, we propose to endow the vector

[2] Our results were first reported in the M.Sc. thesis of one of the authors [7], which appeared in the same month (June 2022) as the arXiv version of [23].

space $\mathcal{B}$ with a multiplication $\boxtimes_{\mathsf{W}}$, so that $\mathcal{B}$ becomes an algebra that we call *algebra of boxes* and that we denote $\mathcal{B}_{\mathsf{W}}$. To that end, we recall the notion of *wiring* W (deterministic then mixed), which, for the sake of simplicity, we define for only *two* boxes being connected — see Remark 6 for more generality. Then we provide some typical examples of wirings from the literature, and we finally introduce the algebra of boxes and characterize its associativity and commutativity.

1.1 Intuition Behind Wirings

Given two non-signalling boxes P and Q, it is possible to build a new box by *wiring* them together. This notion of wiring has found a great interest in the last two decades, especially with the following two goals: (i) attempting nonlocality distillation, *i.e.* we want to build a box that is "strongly nonlocal" starting from some boxes that are "weakly nonlocal" [6, 11, 13, 19, 21, 23, 24, 25, 27, 35]; (ii) finding sets that are closed under wirings, because it is argued that a consistent physical theory should, in principle, be closed under natural simple operations as wirings [1, 6, 31, 36, 37].

As one might guess, a wiring simply connects some outputs to some inputs under some rules, and it applies some pre- and post- processing operations to the carried bits. An example of wiring is presented in Figure 2 (a), where the wiring indeed connects some outputs to some inputs, but is counter-intuitive at first, since Alice and Bob do not use their share of the boxes in the same order: while Alice uses P then Q, Bob uses Q then P. This independence on the choice of the box order for each player generalizes quantum mechanics, in the sense that if Alice and Bob were sharing two entangled pairs instead of two nonlocal boxes, Alice would be able to measure her first particle and then the second one, while Bob would be able to do the converse, and they would still receive the outputs "instantaneously". Now, as in the quantum case, Alice receives an answer from the box P instantaneously even if Bob has not yet inputted a bit in his side of P, and she can use the output a_1 as a parametrization for the input x_2 of the box Q; similarly for Bob. This "instantaneous-answer" property of a box is typical of non-signalling correlations, as modelled by Equations (2) and (3) saying that Alice's marginal is independent of Bob's input, and vice-versa. Note that a wiring cannot link Alice's side to Bob's side, nor the opposite, since otherwise it could create a signalling box: there would be communication between parties.

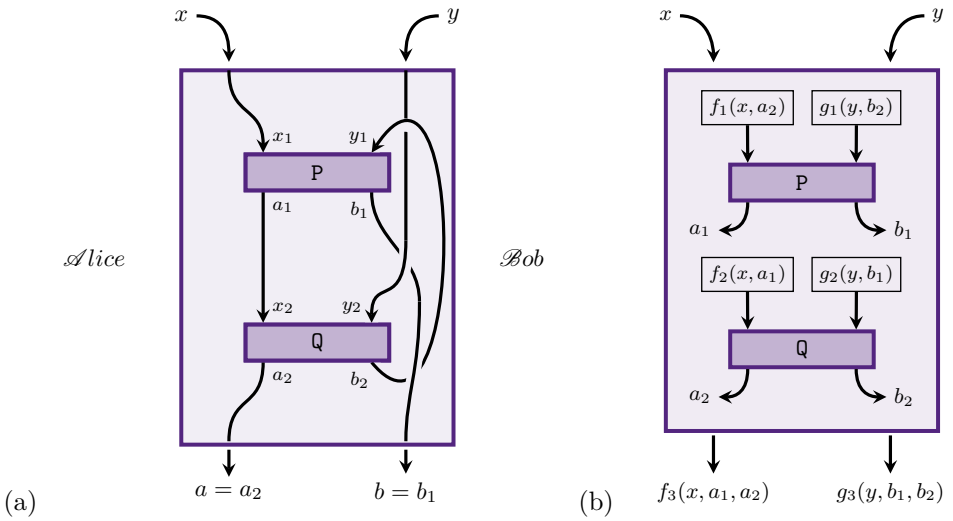


Figure 2: (a) Example of wiring between two boxes P and Q. (b) General wiring between two boxes P and Q.

1.2 Deterministic Wirings

Two boxes P and Q can be wired as in Figure 2 (b), using functions f_i and g_j depending on the global entries x and y and on the outputs a_k and b_ℓ of the boxes. Nevertheless, to be a valid wiring, the inputs on Alice's side must be in a valid order: the input x_2 of Q can depend on the output a_1 of P only if the input x_1 of P does not depend on the output a_2 of Q; the same should also hold on

Bob's side. In other words, the functions $f_1(x, a_2)$ and $f_2(x, a_1)$ cannot both depend on a_2 and a_1 respectively for the same value of x, and similarly for $g_1(y, b_2)$ and $g_2(y, b_1)$. These conditions are formalized in (5) and (6) of the following definition:

Definition 1 (Deterministic wiring). *A* deterministic wiring W *between two boxes* $\mathtt{P}, \mathtt{Q} \in \mathcal{NS}$ *consists in six Boolean functions* $f_1, f_2, g_1, g_2 : \{0,1\}^2 \to \{0,1\}$ *and* $f_3, g_3 : \{0,1\}^3 \to \{0,1\}$ *satisfying the* non-cyclicity conditions*:*

$$\forall x, \qquad \big(f_1(x,0) - f_1(x,1)\big)\big(f_2(x,0) - f_2(x,1)\big) \;=\; 0\,, \tag{5}$$

$$\forall y, \qquad \big(g_1(y,0) - g_1(y,1)\big)\big(g_2(y,0) - g_2(y,1)\big) \;=\; 0\,. \tag{6}$$

Given a wiring W and two boxes $\mathtt{P}, \mathtt{Q} \in \mathcal{NS}$, we obtain a new box that we denote $\mathtt{P} \boxtimes_{\mathsf{W}} \mathtt{Q}$. Formally, this new box is defined as the following conditional probability distribution:

$$\mathtt{P} \underset{\mathsf{W}}{\boxtimes} \mathtt{Q}(a,b\,|\,x,y) := \sum_{a_1,a_2,b_1,b_2} \mathtt{P}\Big(a_1,\, b_1 \,|\, f_1(x,a_2),\, g_1(y,b_2)\Big) \times\, \mathtt{Q}\Big(a_2,\, b_2 \,|\, f_2(x,a_1),\, g_2(y,b_1)\Big) \\ \times\; \mathbb{1}_{a=f_3(x,a_1,a_2)} \;\times\; \mathbb{1}_{b=g_3(y,b_1,b_2)}\,, \tag{7}$$

where the symbol $\mathbb{1}$ stands for the indicator function, taking value 1 if the indexed condition is satisfied, and 0 otherwise. It is important to specify the condition $\mathtt{P}, \mathtt{Q} \in \mathcal{NS}$, since in that case $\mathtt{P} \boxtimes_{\mathsf{W}} \mathtt{Q}$ is indeed a conditional probability distribution as shown in the proof of Fact 3; otherwise, if one requires only the condition on $\mathtt{P}, \mathtt{Q}$ to be conditional probability distributions (not necessarily lying in $\mathcal{NS}$), then it might happen that the product $\mathtt{P} \boxtimes_{\mathsf{W}} \mathtt{Q}$ is not a well-defined probability distribution: consider for example $\mathtt{P} = \mathtt{Q} = \mathbb{1}_{a=y}\mathbb{1}_{b=x}$ and the deterministic wiring $\mathsf{W} = (f_1 = x, f_2 = a_1, g_1 = b_2, g_2 = y, f_3 = 0, g_3 = 0)$.

Definition 2 (Closed under wirings). *A set* $X \subseteq \mathcal{NS}$ *is said to be* closed under wirings *if for all boxes* $\mathtt{P}, \mathtt{Q}$ *in* X *and all wirings* W*, the new box* $\mathtt{P} \boxtimes_{\mathsf{W}} \mathtt{Q}$ *is in* X *as well.*[3]

For the sake of completeness, we recall the fact that the non-signalling polytope $\mathcal{NS}$ is an example of a set that is closed under wirings.

Fact 3. [1] *$\mathcal{NS}$ is closed under deterministic wirings.*

Proof. We need to show that the box $\mathtt{P} \boxtimes_{\mathsf{W}} \mathtt{Q}$ given in (7) is a well-defined conditional probability distribution that satisfies the non-signalling conditions (2) and (3). First, by non-negativity of $\mathtt{P}$ and $\mathtt{Q}$, the new box $\mathtt{P} \boxtimes_{\mathsf{W}} \mathtt{Q}$ is non-negative as well. Now, fix $x, y \in \{0,1\}$. The non-cyclicity conditions (5) and (6) tell us that f_1 or f_2 is constant in the second variable, and similarly for g_1 and g_2. Without loss of generality, up to changing the roles of both f_1, g_1 with respectively f_2, g_2, we only need to consider the following two non-exclusive cases:

- Case 1: the functions $f_1(x, a_2)$ and $g_1(y, b_2)$ are constant in the second variable, and we denote them $f_1(x)$ and $g_1(y)$;
- Case 2: the functions $f_2(x, a_1)$ and $g_1(y, b_2)$ are constant in the second variable, and we denote them $f_2(x)$ and $g_1(y)$.

In Case 1, we see that coefficients sum to one by normalization of $\mathtt{P}$ and $\mathtt{Q}$:

$$\sum_{a,b} \mathtt{P} \underset{\mathsf{W}}{\boxtimes} \mathtt{Q}(a,b\,|\,x,y) \;=\; \sum_{a_1,b_1} \mathtt{P}\big(a_1,\, b_1 \,|\, f_1(x),\, g_1(y)\big) \times \Bigg(\underbrace{\sum_{a_2,b_2} \mathtt{Q}\Big(a_2,\, b_2 \,|\, f_2(x,a_1),\, g_2(y,b_1)\Big)}_{=1} \Bigg) \;=\; 1\,.$$

[3] Notice that there exists as well a more general definition involving k boxes and m parties, but our manuscript is restricted to the simpler case of $k = m = 2$, which is the reason why we give a simpler definition here. In the general framework, many sets are known to be closed under wirings [1, 6, 31, 36, 37].

In Case 2, using the non-signalling conditions on $\mathtt{P}$ and $\mathtt{Q}$, we see that coefficients sum to one again:

$$\sum_{a,b} \mathtt{P} \underset{\mathsf{W}}{\boxtimes} \mathtt{Q}(a,b\,|\,x,y) \;=\; \sum_{a_2,b_1} \underbrace{\left(\sum_{a_1} \mathtt{P}\big(a_1,\, b_1 | f_1(x,a_2),\, g_1(y)\big)\right)}_{=\mathtt{P}(b_1\,|g_1(y))} \underbrace{\left(\sum_{b_2} \mathtt{Q}\Big(a_2,\, b_2 | f_2(x),\, g_2(y,b_1)\Big)\right)}_{=\mathtt{Q}(a_2\,|f_2(x))}$$
$$= \left(\sum_{b_1} \mathtt{P}\big(b_1\,|g_1(y)\big)\right)\left(\sum_{a_2} \mathtt{Q}\big(a_2\,|f_2(x)\big)\right) \;=\; 1\,.$$

Hence $\mathtt{P}\boxtimes_{\mathsf{W}}\mathtt{Q}$ is a conditional probability distribution. It only remains to check that $\mathtt{P}\boxtimes_{\mathsf{W}}\mathtt{Q}$ satisfies the non-signalling conditions (2) and (3). Fix $x, a \in \{0,1\}$. In Case 1, we have for all $y \in \{0,1\}$:

$$\sum_{b} \mathtt{P} \underset{\mathsf{W}}{\boxtimes} \mathtt{Q}(a,b\,|\,x,y) \;=\; \sum_{a_1,a_2} \mathtt{P}\big(a_1\,|\,f_1(x)\big)\,\mathtt{Q}\big(a_2\,|\,f_2(x,a_1)\big)\,\mathbb{1}_{a=f_3(x,a_1,a_2)} \;=:\; \mathtt{P}\boxtimes_W \mathtt{Q}(a\,|\,x)\,,$$

i.e. the result does not depend on y, which means that the marginal in b is well-defined. This is similar in Case 2, changing $f_1(x)$ into $f_1(x,a_2)$ and $f_2(x,a_1)$ into $f_2(x)$. Hence the first non-signalling condition (2) is satisfied, and the other one (3) follows in a similar way. □

1.3 Mixed Wirings

Using local randomness, one can generalize the class of deterministic wirings to the one of *mixed wirings*. The difference is that the functions f_i and g_j take values in $[0,1]$ instead of $\{0,1\}$. For instance, if $f_1(x,a_1) = p \in [0,1]$ for some fixed bits x and a_1, it means that Alice uses a Bernoulli distribution $\mathcal{B}(p)$ to input the bit 1 with probability p, or the bit 0 with probability $1-p$. In other words, we have 32 Bernoulli variables $(B_1,\dots,B_{32})$, whose parameters are stored in $\mathsf{W} = \big(f_1(0,0), f_1(0,1),\dots,g_3(1,1,1)\big) \in \mathbb{R}^{32}$, and the box product $\mathtt{P}\boxtimes\mathtt{Q}$ becomes the expected value of the deterministic wirings:

$$\mathtt{P} \underset{\mathsf{W}}{\boxtimes} \mathtt{Q} \;=\; \mathbb{E}\Big[\mathtt{P} \underset{\{B_i\}}{\boxtimes} \mathtt{Q}\Big]\,. \tag{8}$$

Note that this generalization of wirings does not change the definition of nonlocal boxes: the inputs and outputs of a box are still classical bits, not any real number between 0 and 1. In order to ensure a well-defined local order for both Alice and Bob, we will need to add a dependence relation between the variables B_i, namely the non-cyclicity condition, as for the deterministic wirings:

Definition 4 (Mixed wiring)**.** *A* mixed wiring W *between two boxes* $\mathtt{P},\mathtt{Q} \in \mathcal{NS}$ *consists in six functions* $f_1, f_2, g_1, g_2 : \{0,1\}^2 \to [0,1]$ *and* $f_3, g_3 : \{0,1\}^3 \to [0,1]$ *satisfying the* non-cyclicity conditions (5) *and* (6)*. Mixed wirings form a set that we denote* $\mathcal{W}$*.*

The set of mixed wirings W is not convex because non-cyclicity conditions (5) and (6) are non-affine equalities. For instance, consider the wirings W, W' with all coefficients 0 except the one corresponding to respectively $f_1(0,0) = 1$, $f_2'(0,0) = 1$; each of these wirings satisfies the non-cyclicity conditions (5) and (6), but the mean $\mathsf{W}'' = (\mathsf{W}+\mathsf{W}')/2$ does not:

$$\big(f_1''(0,0) - f_1''(0,1)\big)\big(f_2''(0,0) - f_2''(0,1)\big) \;=\; \big(1/2-0\big)\big(1/2-0\big) \;\neq\; 0\,,$$

hence the non-convexity of $\mathcal{W}$. The expression of $\mathtt{P}\boxtimes_{\mathsf{W}}\mathtt{Q}$ is the same as before, with the convention that $\mathtt{P}(a,b\,|\,\alpha,\beta)$ with $\alpha,\beta \in [0,1]$ means $(1-\alpha)(1-\beta)\mathtt{P}(ab\,|\,00) + (1-\alpha)\beta\mathtt{P}(ab\,|\,01) + \alpha(1-\beta)\mathtt{P}(ab\,|\,10) + \alpha\beta\mathtt{P}(ab\,|\,11)$, which gives:

$$
\begin{aligned}
&\mathtt{P} \boxtimes_{\mathsf{W}} \mathtt{Q}(a,b\,|\,x,y) \\
&= \sum_{a_1,a_2,b_1,b_2\in\{0,1\}} \Big[\mathtt{P}\big(a_1,\,b_1\,|\,0,\,0\big)(1-f_1(x,a_2))\,(1-g_1(y,b_2)) + \mathtt{P}\big(a_1,\,b_1\,|\,0,\,1\big)(1-f_1(x,a_2))\,g_1(y,b_2) \\
&\qquad + \mathtt{P}\big(a_1,\,b_1\,|\,1,\,0\big)\,f_1(x,a_2)\,(1-g_1(y,b_2)) + \mathtt{P}\big(a_1,\,b_1\,|\,1,\,1\big)\,f_1(x,a_2)\,g_1(y,b_2)\Big] \\
&\quad \times \Big[\mathtt{Q}\big(a_2,\,b_2\,|\,0,\,0\big)(1-f_2(x,a_1))\,(1-g_2(y,b_1)) + \mathtt{Q}\big(a_2,\,b_2\,|\,0,\,1\big)(1-f_2(x,a_1))\,g_2(y,b_1) \\
&\qquad + \mathtt{Q}\big(a_2,\,b_2\,|\,1,\,0\big)\,f_2(x,a_1)\,(1-g_2(y,b_1)) + \mathtt{Q}\big(a_2,\,b_2\,|\,1,\,1\big)\,f_2(x,a_1)\,g_2(y,b_1)\Big] \\
&\quad \times \Big[(1-f_3(x,a_1,a_2))\,\mathbb{1}_{a=0} + f_3(x,a_1,a_2)\,\mathbb{1}_{a=1}\Big] \times \Big[(1-g_3(y,b_1,b_2))\,\mathbb{1}_{b=0} + g_3(y,b_1,b_2)\,\mathbb{1}_{b=1}\Big]. \qquad (9)
\end{aligned}
$$

Using the probabilistic point of view of Equation (8), we see that a mixed wiring is a convex combination of deterministic wirings (the realization of Bernoulli variables being either 0 or 1). Hence, by linearity of the expectation $\mathbb{E}$, we deduce from Fact 3 that the set $\mathcal{NS}$ is closed under mixed wirings:

Fact 5. *$\mathcal{NS}$ is closed under mixed wirings.* □

Remark 6. Note that the formalism presented here is deliberately not the most general one, since this simpler version is enough to state our results in the next sections. For a more general framework, see [6]. For instance, here we require a deterministic local box order: knowing x, Alice perfectly knows which box she will use first, and similarly for Bob knowing y, but a more general mixed wiring would consist in setting a probability distribution on the different permutations of Alice's boxes and another one on Bob's boxes. In addition, here we only defined *wirings of depth* 2, but it is possible to have more complex wirings using k nonlocal boxes, thus obtaining a *wirings of depth k*, which play a significant role in the context of trivial communication complexity [23].

1.4 Typical Examples of Wirings

We now review typical wirings that are studied in the literature. See Figure 3 for an illustration of these wirings. Note that all of these wirings are *deterministic* wirings.

Example 0. The trivial wiring W_{triv} is defined as the wiring that does "nothing", in the sense that it outputs exactly the global inputs: $(a,b) = (x,y)$. Similarly, the linear wiring W_{lin} simply connects the output of a box to the input of the box immediately below.

Example 1. In [25], Forster *et al.* introduce a wiring $\mathsf{W}_{\oplus}$ in order to distill nonlocality. It consists in setting boxes in parallel and in taking the sum mod 2 of the outputs.

Example 2. In [13], Brunner and Skrzypczyk enhance the wiring from Example 1 in order to obtain a better distillation protocol of nonlocality. Their wiring W_{BS} is *adaptive*, in the sense that boxes are no longer in parallel: the second box's inputs a_2, b_2 are not simply equal to the previous box's outputs a_1, b_1, but they equal the the latter multiplied by the general inputs x, y. Their new protocol is so powerful that it allows to arbitrarily reduce the noise of any *correlated box* (defined as convex combinations of PR and SR) so that the PR box is almost perfectly simulated. As a consequence, communication complexity collapses; see the next section for more details.

Example 3. In [1], Allcock *et al.* study two variants of the previous wirings. First, their "distillation wiring" W_{dist} is similar to the one in Example 2: it is also adaptive and it also distills all correlated boxes. Second, their "AND wiring" $\mathsf{W}_{\wedge}$ resembles the one in Example 1: boxes are set in parallel, but we take the product of the outputs instead of the sum.

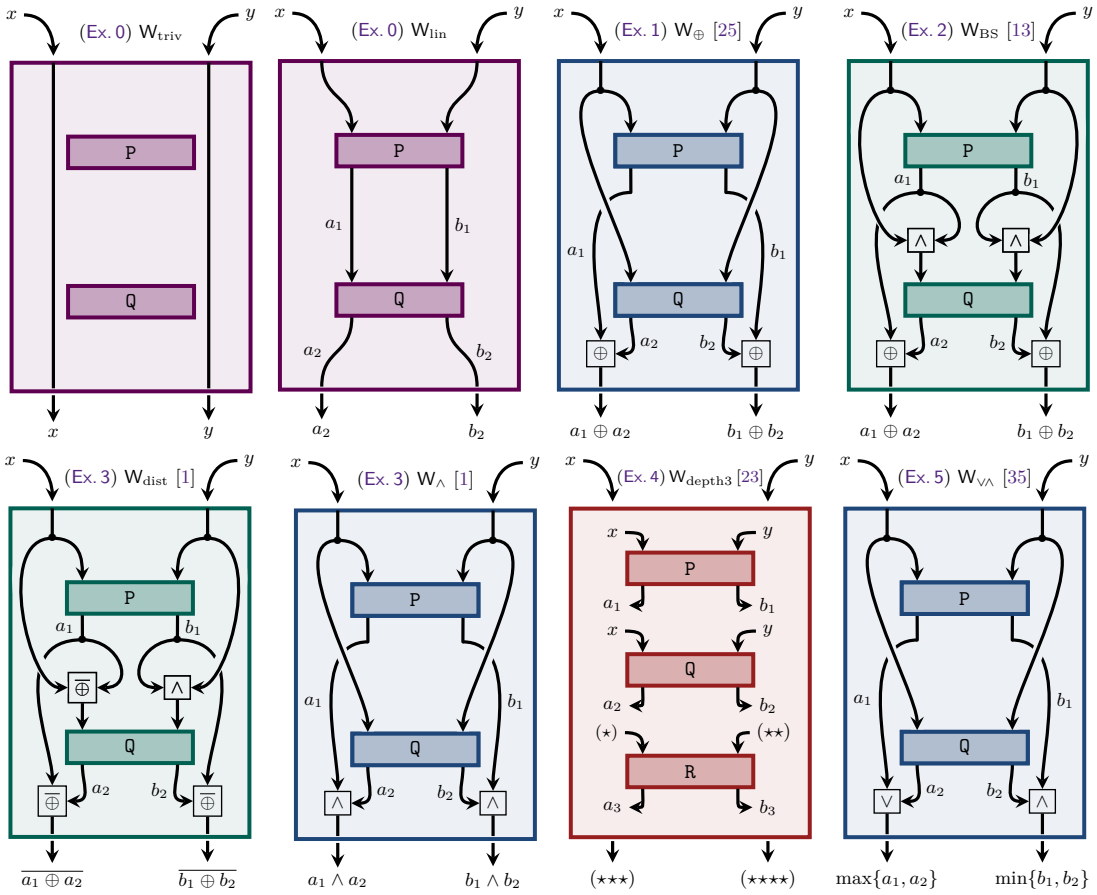


Figure 3: Typical examples of wirings. Wirings with the same color have similar internal structures. The overline bar is the NOT gate: $\overline{x} = x \oplus 1$. The symbol $(\star)$ stands for $xa_2 \vee x\overline{a_1} \vee \overline{xa_2}a_1$, and $(\star\star)$ for $yb_2 \vee y\overline{b_1}$, and $(\star\star\star)$ for $a_3a_2 \vee a_3\overline{a_1} \vee \overline{a_3a_2}a_1$, and $(\star\star\star\star)$ for $b_3b_2 \vee b_3\overline{b_1} \vee \overline{b_3b_2}b_1$.

Example 4. In [27], Høyer and Rashid study the depth-k generalizations of the wirings from Examples 1 and 2 and 3: they wire k boxes instead of only two. They also give an example of a depth-3 protocol that extends the known region of distillable boxes. In [23] this idea is improved upon, by constructing genuine depth-3 protocols, such as $\mathsf{W}_{\text{depth3}}$ drawn above. These protocols are strictly better than depth-2 protocols in terms of the collapse of CC, in the sense that they prove the existence of nonlocal boxes that are shown to collapse CC using this wiring but that cannot be distilled using any depth-2 wiring. To do so, they use their algorithm to single out the area of nonlocal boxes that are distillable by means of depth-2 wirings, and they show that their example $\mathsf{W}_{\text{depth3}}$ is out of this region.[4] Note that in our work, the study is limited to depth-2 wirings.

Example 5. More recently, in [35], Naik *et al.* defined the "OR-AND wiring" $\mathsf{W}_{\vee\wedge}$ in order to distill the nonlocality of quantum correlations. That wiring is a mix of the ones in Examples 1 and 3: it consists in setting boxes in parallel and in taking the maximum (the "OR") of Alice's outputs and the minimum (the "AND") of Bob's outputs.

1.5 Algebra of Boxes Induced by a Wiring

Let $\mathcal{B}$ be the vector space of all the functions $\{0,1\}^4 \to \mathbb{R}$, and consider a mixed wiring W. As defined in Equation (9), the operation $\boxtimes_{\mathsf{W}}$ is bilinear, so the vector space $\mathcal{B}$ equipped with

[4] We thank Mirjam Weilenmann for feedback and for pointing out that [12] states that the wirings $\mathsf{W}_{\oplus}$ and W_{dist} are sufficient to characterize the distillable region of the slice PR–SR–I.

the product $\boxtimes_{\mathsf{W}}$ is actually an algebra, which we denote by $\mathcal{B}_{\mathsf{W}}$ for that specific wiring $\mathsf{W} \in \mathcal{W}$. Its dimension is $\dim \mathcal{B}_{\mathsf{W}} = 2^4 = 16$. Note that the (affine) dimension of the non-signalling polytope $\mathcal{NS} \subseteq \mathcal{B}_{\mathsf{W}}$ is $\dim \mathcal{NS} = 16 - 8 = 8$ because there are 8 dependent variables in the affine conditions (1), (2), (3); see [3] for a more general expression.

Multiplication Table. In order to better understand the behavior of the box product $\boxtimes$, it is interesting to compute the product of some basic boxes: for instance the boxes $\mathtt{PR}$, $\mathtt{P}_0$, $\mathtt{P}_1$, $\mathtt{I}$ defined in Equation (4). In Figure 4, we present the multiplication table for the wiring W_{BS} from [13]. By bilinearity of the box multiplication, this table shows that the convex hull $\mathrm{Conv}\{\mathtt{PR}, \mathtt{P}_0, \mathtt{P}_1\}$ is stable under $\boxtimes$. On the contrary, observe that the convex hull $\mathrm{Conv}\{\mathtt{PR}, \mathtt{P}_0, \mathtt{P}_1, \mathtt{I}\}$ is not stable under $\boxtimes$: the product $\mathtt{I} \boxtimes \mathtt{PR}$ gives $\mathtt{Q}_1 := \frac{1}{4}\,\mathtt{PR} - \frac{1}{8}\big(\mathtt{P}_0 + \mathtt{P}_1\big) + \mathtt{I}$ which is out of the convex hull (nevertheless the affine hull $\mathrm{Aff}\{\mathtt{PR}, \mathtt{P}_0, \mathtt{P}_1, \mathtt{I}\}$ is stable under $\boxtimes$). Notice that we show in Proposition 22 that actually $\mathrm{Conv}\{\mathtt{PR}, \mathtt{P}_0, \mathtt{P}_1\} = \mathcal{NS} \cap \mathrm{Aff}\{\mathtt{PR}, \mathtt{P}_0, \mathtt{P}_1\}$. From this table, one may postulate that $\mathtt{P}_0$ is a right identity in the sense that $\mathtt{P} \boxtimes \mathtt{P}_0 = \mathtt{P}$ for all $\mathtt{P}$ in $\mathcal{NS}$, and it is indeed true as a simple consequence of formula (7). One may similarly verify that $\mathtt{I}$ is a right fixed point, in the sense that $\mathtt{P} \boxtimes \mathtt{I} = \mathtt{I}$ for all $\mathtt{P}$ in $\mathcal{NS}$, as it is possible to guess from the table. See all the multiplication tables of the typical depth-2 wirings in Appendix C.

P \ Q	$\mathtt{PR}$	$\mathtt{P}_0$	$\mathtt{P}_1$	$\mathtt{I}$
$\mathtt{PR}$	$\mathtt{PR}$	$\mathtt{PR}$	$\mathtt{PR}$	$\mathtt{I}$
$\mathtt{P}_0$	$\frac{1}{2}\big(\mathtt{P}_0 + \mathtt{P}_1\big)$	$\mathtt{P}_0$	$\mathtt{P}_1$	$\mathtt{I}$
$\mathtt{P}_1$	$\mathtt{PR}$	$\mathtt{P}_1$	$\mathtt{P}_0$	$\mathtt{I}$
$\mathtt{I}$	$\mathtt{Q}_1$	$\mathtt{I}$	$\mathtt{I}$	$\mathtt{I}$

Figure 4: Multiplication table of the operation $\boxtimes_{\mathsf{W}_{\mathrm{BS}}}$ induced by the wiring from [13]. Each cell displays the result of $\mathtt{P} \boxtimes \mathtt{Q}$. The box $\mathtt{Q}_1$ at the bottom left is $\mathtt{Q}_1 := \frac{1}{4}\,\mathtt{PR} - \frac{1}{8}\big(\mathtt{P}_0 + \mathtt{P}_1\big) + \mathtt{I}$. Further multiplication tables are available in Appendix C.

Non-Commutativity and Non-Associativity. A direct consequence of the multiplication table in Figure 4 is that the algebra $\mathcal{B}_{\mathsf{W}_{\mathrm{BS}}}$ induced by the wiring W_{BS} is non-commutative ($\mathtt{P}_0 \boxtimes \mathtt{PR} \neq \mathtt{PR} \boxtimes \mathtt{P}_0$) and non-associative ($(\mathtt{P}_0 \boxtimes \mathtt{P}_1) \boxtimes \mathtt{PR} \neq \mathtt{P}_0 \boxtimes (\mathtt{P}_1 \boxtimes \mathtt{PR})$). This non-associativity is at the root of interesting remarks, see drawings of the orbit of a box in the next section, Figure 7. Similarly, the algebra induced by the wiring $\mathsf{W}_{\mathrm{dist}}$ is both non-commutative and non-associative, but on the contrary, the algebras induced by $\mathsf{W} \in \{\mathsf{W}_{\mathrm{triv}}, \mathsf{W}_{\oplus}, \mathsf{W}_{\wedge}, \mathsf{W}_{\vee\wedge}\}$ are both commutative and associative. One may wonder if there exist induced algebras that are associative but not commutative, and the converse. To that end, here is a characterization of commutativity and associativity in a simple case where boxes are set in parallel and with the same input functions:

Proposition 7 (Characterization of commutativity and associativity)**.** *Assume* W *is a wiring such that* $f_1 = f_2 = f(x)$ *and* $g_1 = g_2 = g(y)$*. Then:*

(i) $\mathcal{B}_{\mathsf{W}}$ *is commutative if and only if the functions* $f_3(x, a_1, a_2)$ *and* $g_3(y, b_1, b_2)$ *are "symmetric" in the last two variables, in the sense that* $f_3(x, a_1, a_2) = f_3(x, a_2, a_1)$ *for all* x, a_1, a_2*, and similarly for* g_3*.*

If in addition $f(x) = x$ *and* $g(y) = y$*:*

(ii) $\mathcal{B}_{\mathsf{W}}$ *is associative if and only if the functions* $f_3(x, a_1, a_2)$ *and* $g_3(y, b_1, b_2)$ *are "associative" in the last two variables, in the sense that* $f_3(x, a_1, f_3(x, a_2, a_3)) = f_3(x, f_3(x, a_1, a_2), a_3)$ *for all* x, a_1, a_2, a_3*, and similarly for* g_3*.*

Proof. (i) First, from the expression (7), see that for all bits a, b, x, y and any boxes $\mathtt{P}, \mathtt{Q}$ in $\mathcal{B}_{\mathsf{W}}$, we

have:

$$\begin{aligned}
&\mathtt{P} \underset{\mathsf{W}}{\boxtimes} \mathtt{Q}(a,b\,|\,x,y) - \mathtt{Q} \underset{\mathsf{W}}{\boxtimes} \mathtt{P}(a,b\,|\,x,y) \\
&\quad = \sum_{a_1,a_2,b_1,b_2} \mathtt{P}\big(a_1,b_1\,|\,f(x),g(y)\big) \times \mathtt{Q}\big(a_2,b_2\,|\,f(x),g(y)\big) \times \Big[\mathbb{1}_{a=f_3(x,a_1,a_2)}\mathbb{1}_{b=g_3(y,b_1,b_2)} \\
&\qquad\qquad - \mathbb{1}_{a=f_3(x,a_2,a_1)}\mathbb{1}_{b=g_3(y,b_2,b_1)}\Big] .
\end{aligned}$$

Hence, if f_3 and g_3 are both symmetric in the last two variables, then the difference is null and the algebra is commutative. Conversely, suppose that $\mathcal{B}_{\mathsf{W}}$ is commutative, so that the left-hand side is null. Taking probability distributions $\mathtt{P}$ and $\mathtt{Q}$ that are always positive (such as $\mathtt{I}$), we have that the difference in the right-hand side has to be null for all $x, y, a, b, a_1, a_2, b_1, b_2$. Fix x, a_1, a_2 and consider $a := f_3(x, a_1, a_2)$, and similarly fix y, b_1, b_2 and consider $b := g_3(y, b_1, b_2)$. We obtain $1 - \mathbb{1}_{a=f_3(x,a_2,a_1)}\mathbb{1}_{b=g_3(y,b_2,b_1)} = 0$, which means that both indicator functions are equal to 1, and therefore both subscript equalities hold. Hence, this being true for any fixed x, a_1, a_2 and y, b_1, b_2, we obtain that f_3 and g_3 are symmetric as wanted.

(ii) From (7) again, we have for all bits a, b, x, y and any boxes $\mathtt{P}, \mathtt{Q}, \mathtt{R}$ in $\mathcal{B}_{\mathsf{W}}$:

$$\begin{aligned}
&\mathtt{P} \underset{\mathsf{W}}{\boxtimes} (\mathtt{Q} \underset{\mathsf{W}}{\boxtimes} \mathtt{R})(a,b\,|\,x,y) - (\mathtt{P} \underset{\mathsf{W}}{\boxtimes} \mathtt{Q}) \underset{\mathsf{W}}{\boxtimes} \mathtt{R}(a,b\,|\,x,y) \\
&= \sum_{a_1,a_2,a_3,b_1,b_2,b_3} \mathtt{P}\big(a_1,b_1\,|\,x,y\big) \times \mathtt{Q}\big(a_2,b_2\,|\,x,y\big) \times \mathtt{R}\big(a_3,b_3\,|\,x,y\big) \times \Big[\mathbb{1}_{a=f_3(x,a_1,f_3(x,a_2,a_3))}\mathbb{1}_{b=g_3(y,b_1,g_3(y,b_2,b_3))} \\
&\qquad\qquad - \mathbb{1}_{a=f_3(x,f_3(x,a_1,a_2),a_3)}\mathbb{1}_{b=g_3(y,g_3(y,b_1,b_2),b_3)}\Big] .
\end{aligned}$$

A similar proof with double implication as in (i) applies, hence the associativity criterion follows. □

Now, it is easier to build an associative non-commutative induced algebra $\mathcal{B}_{\mathsf{W}'}$. Consider the wiring W' given by $f_1(x, a_2) = f_2(x, a_1) = x$, and $g_1(y, b_2) = g_2(y, b_1) = y$, and $f_3(x, a_1, a_2) := a_1$, and $g_3(y, b_1, b_2) := b_1$. This wiring satisfies the condition (ii) of the proposition and does not satisfy the condition (i), hence it is as wanted. Conversely, with similar arguments, a commutative non-associative algebra $\mathcal{B}_{\mathsf{W}''}$ is induced by the wiring W'' defined by the same f_1, f_2, g_1, g_2 and $f_3(x, a_1, a_2) := a_1a_2 \oplus 1$ and $g_3(y, b_1, b_2) := b_1b_2 \oplus 1$. Therefore, we obtain the table in Figure 5.

	Associativity	Non-associativity
Commutativity	$\mathsf{W}_{\text{triv}}, \mathsf{W}_{\oplus}, \mathsf{W}_{\wedge}, \mathsf{W}_{\vee\wedge}$	W''
Non-commutativity	W'	$\mathsf{W}_{\text{BS}}, \mathsf{W}_{\text{dist}}$

Figure 5: Associativity and commutativity of the induced algebra $\mathcal{B}_{\mathsf{W}}$, depending on the wiring W displayed in the table cell.

2 Orbit of a Box

In this section, we study the set of all boxes that can be generated given many copies of a starting box $\mathtt{P}$ and a wiring W. After introducing the *orbit* of a box, we provide some consequences to communication complexity. Subsequently, we study a particular example, W_{BS}, with which we find collapsing boxes in Subsection 4.2, and then we give some general remarks about other orbits. Finally, we conclude this section by giving the technical proof of the theorem stating that the "best" parenthesization is the multiplication on the right.

2.1 Definition

Given multiple copies of a non-signalling box $\mathtt{P} \in \mathcal{NS}$ and of a (mixed) wiring W, Alice and Bob can produce many other boxes, *e.g.* $(\mathtt{P} \boxtimes_{\mathsf{W}} \mathtt{P}) \boxtimes_{\mathsf{W}} \mathtt{P}$ or $\mathtt{P} \boxtimes_{\mathsf{W}} (\mathtt{P} \boxtimes_{\mathsf{W}} \mathtt{P})$. All of these new boxes are

again non-signalling because $\mathcal{NS}$ is closed under wirings, see Lemma 5. We call *orbit* of the box $\mathtt{P}$ (induced by the wiring W) the set of all of these possible new boxes:

$$\begin{aligned}\mathrm{Orbit}_{\mathsf{W}}(\mathtt{P}) &:= \Big\{\text{boxes } \mathtt{Q} \in \mathcal{NS} \text{ that can be produced by using finitely many times the box } \mathtt{P} \text{ and the wiring } \mathsf{W}\Big\} \\ &= \bigcup_{k\geq 1} \mathrm{Orbit}_{\mathsf{W}}^{(k)}(\mathtt{P}) \subseteq \mathcal{NS}\,,\end{aligned}$$

where $\mathrm{Orbit}^{(k)}(\mathtt{P})$ is called the *orbit of depth* k of $\mathtt{P}$ (or simply k-*orbit*), defined as:

$$\mathrm{Orbit}_{\mathsf{W}}^{(k)}(\mathtt{P}) := \Big\{\text{all possible products with } k \text{ times the term } \mathtt{P}, \text{ using the multiplication } \boxtimes_{\mathsf{W}}\Big\}\,.$$

When the context is clear, we overload the notation and write Orbit and $\mathrm{Orbit}^{(k)}$ respectively. In general, these k-orbits are not singletons for $k \geq 3$ since the algebra $\mathcal{B}_{\mathsf{W}}$ induced by W is not necessarily associative and commutative (see Figure 5). Actually, up to multiplicity, the cardinal $\#\,\mathrm{Orbit}^{(k)}$ is exactly the number of parenthesizations with k terms, which is the Catalan number $\frac{1}{k}\binom{2k-2}{k-1}$, which grows exponentially fast. Here are the 3- and 4- orbits:

$$\begin{aligned}\mathrm{Orbit}^{(3)}(\mathtt{P}) &= \Big\{(\mathtt{P}\boxtimes\mathtt{P})\boxtimes\mathtt{P},\ \mathtt{P}\boxtimes(\mathtt{P}\boxtimes\mathtt{P})\Big\}\,,\\ \mathrm{Orbit}^{(4)}(\mathtt{P}) &= \Big\{\big((\mathtt{P}\boxtimes\mathtt{P})\boxtimes\mathtt{P}\big)\boxtimes\mathtt{P},\ \big(\mathtt{P}\boxtimes(\mathtt{P}\boxtimes\mathtt{P})\big)\boxtimes\mathtt{P},\ (\mathtt{P}\boxtimes\mathtt{P})\boxtimes(\mathtt{P}\boxtimes\mathtt{P}),\ \mathtt{P}\boxtimes\big((\mathtt{P}\boxtimes\mathtt{P})\boxtimes\mathtt{P}\big),\ \mathtt{P}\boxtimes\big(\mathtt{P}\boxtimes(\mathtt{P}\boxtimes\mathtt{P})\big)\Big\}\,.\end{aligned}$$

Note that a k-orbit ($k \geq 2$) can be inductively computed using orbits with lower depth:

$$\mathrm{Orbit}^{(k)} = \bigcup_{1\leq\ell\leq k-1} \mathrm{Orbit}^{(\ell)} \boxtimes \mathrm{Orbit}^{(k-\ell)}\,,$$

which is the same recurrence relation as that of Catalan numbers.

2.2 Consequences to Communication Complexity

Assume Alice and Bob are given infinitely many copies of a nonlocal box $\mathtt{P}$, and assume they want to distantly compute (in finite time) the value of a Boolean function $f(X,Y)$, where $X,Y \in \{0,1\}^n$ are strings that are known by Alice and Bob respectively. Among all the possible protocols they can try to do in order to succeed, they can wire their copies of $\mathtt{P}$ in order to produce a "better" box. For example, starting from a noisy box $\mathtt{P}$, Alice and Bob can try to produce a box that is closer to the "perfect box" $\mathtt{PR}$ which satisfies $a \oplus b = xy$ without noise. Such a protocol is called a *distillation protocol* [13]. We call *collapsing box* a nonlocal box that collapses CC.

Find Collapsing Boxes Using the Orbit. Imagine Alice and Bob are able to produce a collapsing box $\mathtt{Q}$ after applying wirings to copies of a starting box $\mathtt{P}$. Then they can use that new box $\mathtt{Q}$ to distantly compute the value $f(X,Y)$, which means that they have a protocol to collapse communication complexity and therefore that $\mathtt{P}$ is collapsing. This point of view is particularly interesting since it implies that it is sufficient to find a single collapsing box in the union $\bigcup_{\mathsf{W}} \mathrm{Orbit}_{\mathsf{W}}(\mathtt{P})$ to deduce that $\mathtt{P}$ is collapsing as well. See an illustration in Figure 6 (a).

Find Collapsing Boxes Using a Cone. Once we find a collapsing box $\mathtt{P}$, we can deduce many other collapsing boxes: there is a convex cone taking origin at $\mathtt{P}$ that is collapsing as well. More precisely, given a box $\mathtt{P}$, denote $\mathcal{C}_{\mathtt{P}}$ the convex cone of boxes $\mathtt{R}$ for which there exists a local correlation $\mathtt{L} \in \mathcal{L}$ such that $\mathtt{P} = \lambda\,\mathtt{R} + (1-\lambda)\,\mathtt{L}$, with $\lambda \in [0,1]$. We claim that if $\mathtt{P}$ is collapsing, then any $\mathtt{R} \in \mathcal{C}_{\mathtt{P}}$ is collapsing as well. Indeed, assume Alice and Bob are given copies of a box $\mathtt{R}$. Then, they can use shared randomness to produce the wanted box $\mathtt{L}$ and the wanted convex coefficient λ, so that they can generate the box $\mathtt{P}$ with the relation $\mathtt{P} = \lambda\,\mathtt{R} + (1-\lambda)\,\mathtt{L}$. Now, as $\mathtt{P}$ is collapsing, they have a protocol that collapses communication complexity, hence $\mathtt{R}$ is collapsing as well. See an illustration in Figure 6 (b). In the study of collapsing boxes, notice that it is standard to assume that shared randomness is a "free" resource; for instance Brassard *et. al.* made that choice in [10] in their collapsing protocol.

Combining arguments from these last two paragraphs and by the fact that Alice and Bob can make a convex combination of boxes using shared randomness, we deduce a sufficient criterion for a box to collapse communication complexity:

Proposition 8 (Collapsing orbit). *Let* $\mathtt{P}$ *be a box in* $\mathcal{NS}$*. If there exists a box* $\mathtt{Q} \in \mathrm{Conv}\left(\mathcal{L} \cup \bigcup_{\mathsf{W}} \mathrm{Orbit}_{\mathsf{W}}(\mathtt{P})\right)$ *that collapses communication complexity, then* $\mathtt{P}$ *is collapsing as well. See Figure 6* (c).

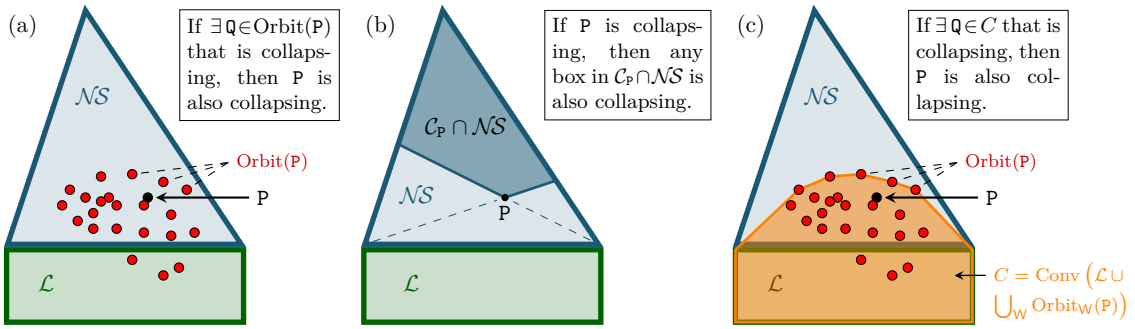


Figure 6: Orbits that collapse communication complexity.

2.3 Case Study: Orbit of W_{BS}

In this subsection, we focus our attention on the wiring W_{BS} inspired by Brunner and Skrzypczyk [13]. Denote $\boxtimes$ the corresponding box multiplication. Define the *shared randomness* box as $\mathtt{SR} := (\mathtt{P}_0 + \mathtt{P}_1)/2$; it is designed to output a couple (a, b) such that $a = b$ uniformly and independently of the inputs. From the multiplication table in Figure 4, one can see that the 2-dimensional affine space $\mathcal{A} := \mathrm{Aff}\{\mathtt{PR}, \mathtt{SR}, \mathtt{I}\}$ is stable under $\boxtimes$. As a consequence, the orbit $\mathrm{Orbit}(\mathtt{P})$ of any box $\mathtt{P}$ in $\mathcal{A}$ is itself included in $\mathcal{A}$, and as $\mathcal{A}$ is two-dimensional, it is particularly easy to draw the orbit of a box in that case. We represent an orbit in Figure 7.

By definition of the affine space $\mathcal{A}$, any box $\mathtt{A} \in \mathcal{A}$ can be uniquely written as $\mathtt{A} = c_1(\mathtt{A})\,\mathtt{PR} + c_2(\mathtt{A})\,\mathtt{SR} + c_3(\mathtt{A})\,\mathtt{I}$ for some real coefficients $c_i(\mathtt{A})$ that sum to 1, called *convex coordinates* of $\mathtt{A}$ in the affine basis $\{\mathtt{PR}, \mathtt{SR}, \mathtt{I}\}$. An interesting aspect of considering convex coordinates is that it gives a simple characterization of the parallelism property of lines:

$$\forall \mathtt{A}, \mathtt{B} \in \mathcal{A}, \qquad \mathrm{Aff}\{\mathtt{A}, \mathtt{B}\} \,||\, \mathrm{Aff}\{\mathtt{PR}, \mathtt{SR}\} \iff c_3(\mathtt{A}) = c_3(\mathtt{B})\,. \quad [5] \tag{10}$$

Moreover, in our case, we have an additional interesting property of the third convex coordinate:

Lemma 9. *The function* $1 - c_3(\cdot)$ *is multiplicative:*

$$\forall \mathtt{A}, \mathtt{B} \in \mathcal{A}, \qquad 1 - c_3(\mathtt{A} \boxtimes \mathtt{B}) = \big(1 - c_3(\mathtt{A})\big)\big(1 - c_3(\mathtt{B})\big)\,.$$

[5] Indeed, for $\mathtt{A} \neq \mathtt{B} \in \mathcal{A}$ whose convex coefficients are respectively a_1, a_2, a_3 and b_1, b_2, b_3, saying that the line $\mathrm{Aff}\{\mathtt{A}, \mathtt{B}\}$ is parallel to the line $\mathrm{Aff}\{\mathtt{PR}, \mathtt{SR}\}$ is equivalent to knowing that there exists a scalar $\lambda \in \mathbb{R}^*$ such that $\mathtt{A} - \mathtt{B} = \lambda\,(\mathtt{PR} - \mathtt{SR})$, *i.e.* there exists $\lambda \in \mathbb{R}^*$ such that $a_1 - b_1 = \lambda$ and $a_2 - b_2 = -\lambda$ and $a_3 - b_3 = 0$, *i.e.* we have two equations: $a_1 + a_2 = b_1 + b_2$ and $a_3 = b_3$. Finally, using the normalization condition $\sum_i a_i = \sum_j b_j = 1$, we see that these two equations are equivalent to simply imposing $a_3 = b_3$, as claimed.

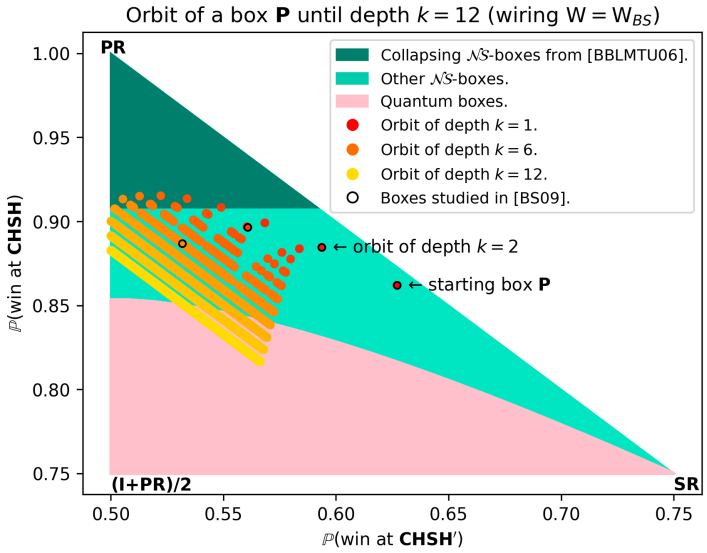


Figure 7: Example of a box orbit, drawn for depth up to $k = 12$. The quantum area $\mathcal{Q}$ (in pink) is drawn using formulas from [33]. Dark green represents the collapsing area that was found by Brassard *et. al.* in [10], which consists of all the boxes with CHSH-value higher than $\frac{3+\sqrt{6}}{6} \approx 0.91$. The orbit is drawn in yellow and orange dots — observe that it intersects the collapsing area in dark green, so Proposition 8 tells us that the starting box P is collapsing. The black circles represent the boxes that were studied in [13], doing "pairwise" multiplications: P, $\mathtt{P} \boxtimes \mathtt{P}$, $(\mathtt{P} \boxtimes \mathtt{P}) \boxtimes (\mathtt{P} \boxtimes \mathtt{P})$, *etc*... Each iteration is the wiring of two copies of the previous iteration, it gives a subset of our orbit. As displayed in the drawing and detailed in the proof of Proposition 16, our method allows us to find a larger set of boxes P that are collapsing.

Proof. The multiplication table induced by the wiring W_{BS} [13] is:

P \ Q	PR	SR	I
PR	PR	PR	I
SR	$\frac{1}{2}$PR + $\frac{1}{2}$SR	SR	I
I	$\frac{1}{4}$PR − $\frac{1}{4}$SR + I	I	I

, (11)

where each cell displays the result of $\mathtt{P} \boxtimes \mathtt{Q}$. For $\mathtt{A}, \mathtt{B} \in \mathcal{A}$ whose coefficients c_i are denoted a_1, a_2, a_3 and b_1, b_2, b_3 for the sake of readability, we use the bilinearity of the product $\boxtimes$ and we get:

$$\begin{aligned} \mathtt{A} \boxtimes \mathtt{B} = &\left[a_1b_1 + a_1b_2 + \tfrac{1}{2}a_2b_1 + \tfrac{1}{4}a_3b_1\right] \mathtt{PR} + \left[\tfrac{1}{2}a_2b_1 + a_2b_2 - \tfrac{1}{4}a_3b_1\right] \mathtt{SR} \\ &+ \left[a_1b_3 + a_2b_3 + a_3b_1 + a_3b_2 + a_3b_3\right] \mathtt{I}\,. \end{aligned}$$

Hence, using the normalization property of coefficients $\sum_i a_i = \sum_j b_j = 1$, the third coefficient simplifies as $b_3 + a_3(1 - b_3)$, which is equal to $1 - (1 - a_3)(1 - b_3)$ as wanted. □

Now, interestingly, we observe that the points of a given k-orbit are all aligned, and we even know the equation of the line:

Theorem 10 (Alignment). *For any $k \geq 1$ and $\mathtt{P} \in \mathcal{A}$, the points of* $\mathrm{Orbit}^{(k)}(\mathtt{P})$ *are all aligned on a line $\mathfrak{L}_k$ whose expression in convex coordinates is given by:*

$$\mathfrak{L}_k := \left\{\mathtt{A} \in \mathcal{A} \,:\, c_3(\mathtt{A}) = 1 - \big(1 - c_3(\mathtt{P})\big)^k\right\}.$$

Proof. We prove by induction on $k \geq 1$ that $\mathrm{Orbit}^{(k)} \subseteq \mathfrak{L}_k$. For $k = 1$, the 1-orbit contains only one element, namely P, which obviously satisfies $c_3(\mathtt{P}) = 1 - \big(1 - c_3(\mathtt{P})\big)$, so P indeed belongs to $\mathfrak{L}_1$.

Now, assume the result holds *until* some integer $k \geq 1$, and let $\mathtt{Q} \in \mathrm{Orbit}^{(k)}$. By definition, the box $\mathtt{Q}$ decomposes as $\mathtt{Q} = \mathtt{Q}_1 \boxtimes \mathtt{Q}_2$, for some $\mathtt{Q}_1 \in \mathrm{Orbit}^{(\ell)}$ and $\mathtt{Q}_2 \in \mathrm{Orbit}^{(k-\ell)}$ for some $1 \leq \ell \leq k-1$. By the induction hypothesis, we know that $c_3(\mathtt{Q}_1) = 1-\big(1-c_3(\mathtt{P})\big)^{\ell}$ and $c_3(\mathtt{Q}_2) = 1-\big(1-c_3(\mathtt{P})\big)^{k-\ell}$. Then using Lemma 9, we obtain:

$$\begin{aligned} c_3(\mathtt{Q}) &= 1 - \big(1 - c_3(\mathtt{Q}_1)\big)\big(1 - c_3(\mathtt{Q}_2)\big) \\ &= 1 - \big(1 - c_3(\mathtt{P})\big)^{\ell}\big(1 - c_3(\mathtt{P})\big)^{k-\ell} \\ &= 1 - \big(1 - c_3(\mathtt{P})\big)^{k}, \end{aligned}$$

which means that $\mathtt{Q}$ belongs to the line $\mathfrak{L}_k$. □

As a consequence, we see that all the points of the k-orbit have the same third convex coefficient, so using the equivalence given in Equation (10), we obtain:

Corollary 11 (Parallelism). *The supporting line $\mathfrak{L}_k$ of all the orbits $\mathrm{Orbit}^{(k)}$ are parallel to the diagonal line $\mathfrak{L}_D := \mathrm{Aff}\{\mathtt{PR}, \mathtt{SR}\}$:*

$$\forall k \geq 1, \qquad \mathrm{Orbit}^{(k)} \,||\, \mathfrak{L}_D \,.$$

In particular, all the orbits are parallel to each other:

$$\forall k, \ell \geq 1, \quad \mathrm{Orbit}^{(k)} \,||\, \mathrm{Orbit}^{(\ell)} \,. \qquad \square$$

Moreover, looking closely at the sequence of coefficients $1 - \big(1 - c_3(\mathtt{P})\big)^{k}$ and noticing that the diagonal line $\mathfrak{L}_D$ is defined by the equation $c_3(\mathtt{A}) = 0$, we see that:

Corollary 12 (Orbits move to the left). *Assume $\mathtt{P} \notin \mathfrak{L}_D$. Then the orbits are more and more distant from the diagonal line as k grows. Moreover, the sequence of lines $(\mathfrak{L}_k)_k$ tends to the line $\mathfrak{L}_\infty$ defined by the equation $c_3(\mathtt{A}) = 1$, which is exactly the line passing through $\mathtt{I}$ and parallel to the diagonal $\mathfrak{L}_D$.* □

It takes a lot of computational time to draw k-orbits of a box $\mathtt{P}$ as k grows, since it requires to compute $\frac{1}{k}\binom{2k-2}{k-1}$ elements (Catalan number), which grows exponentially. However, our goal is not to compute the whole orbit, but simply to determine whether or not the orbit intersects the known collapsing area (dark green). To that end, one may notice that it is enough to compute the "highest" box of each k-orbit in the y-coordinate (see Figure 7) and to check whether those "highest" boxes intersect the collapsing area (dark green area). This is the purpose of the following proposition, which displays a simple expression of the "highest" box of each k-orbit, and which allows much faster tests of a box $\mathtt{P}$ being collapsing or not without computing all the points of the orbit. We prove this result only in a subset of the orbit, that we call *tilted orbit*, which is easier to manipulate in inductions, and which is defined by $\widetilde{\mathrm{Orbit}}^{(1)}(\mathtt{P}) := \{\mathtt{P}\}$ and for $k \geq 2$:

$$\widetilde{\mathrm{Orbit}}^{(k)} := \Big(\mathtt{P} \boxtimes \widetilde{\mathrm{Orbit}}^{(k-1)}\Big) \cup \Big(\widetilde{\mathrm{Orbit}}^{(k-1)} \boxtimes \mathtt{P}\Big) = \bigcup_{\ell \in \{1,k-1\}} \widetilde{\mathrm{Orbit}}^{(\ell)} \boxtimes \widetilde{\mathrm{Orbit}}^{(k-\ell)} \subseteq \mathrm{Orbit}^{(k)} \,.$$

Note that the cardinality of that set is $\#\widetilde{\mathrm{Orbit}}^{(k+1)} = 2^k$, up to multiplicity. We call $\mathtt{CHSH}$-value the y-coordinate, indicating how "high" is a box:

$$\mathtt{CHSH}(\mathtt{P}) := \mathbb{P}\big(\text{win at } \mathtt{CHSH}\big) = \frac{1}{4}\sum_{a \oplus b = xy} \mathtt{P}(a, b \,|\, x, y) \,.$$

We say that a tilted orbit *distills the* $\mathtt{CHSH}$*-value* if it contains a box $\mathtt{Q}$ such that $\mathtt{CHSH}(\mathtt{Q}) \geq \mathtt{CHSH}(\mathtt{P})$. In the following theorem, we present the expression of the best parenthesization in terms of $\mathtt{CHSH}$-value, which explains the numerical observation reported in [23, Supplementary Material, II]:

Theorem 13 (Highest box). *Let* $\mathtt{P} \in \widetilde{\mathcal{A}}$ *be a box, and let* $k \geq 2$ *an integer such that the tilted* $(k-1)$*-orbit distills the* CHSH*-value. Then the highest* CHSH*-value of* $\widetilde{\mathrm{Orbit}}^{(k)}(\mathtt{P})$ *is achieved at a box whose expression is the product of* k *times* P on the right*:*

$$\mathtt{P}^{\boxtimes k} := \Big(\big((\mathtt{P} \boxtimes \mathtt{P}) \boxtimes \mathtt{P}\big) \cdots\Big) \boxtimes \mathtt{P} \in \underset{\mathtt{Q} \in \widetilde{\mathrm{Orbit}}^{(k)}(\mathtt{P})}{\mathrm{argmax}} \mathtt{CHSH}(\mathtt{Q}) \,.$$

Proof. See Subsection 2.5. □

Remark 14. In the Ph.D. thesis of Giorgos Eftaxias [22], the author presents three types of architectures of wirings in Subsection 5.5.1: the exponential architecture, that we call here pairwise multiplication, used in [13]; the linear architecture, that is the same as the one used in the theorem above; and the Fibonacci architecture. They present some subsets of $\mathcal{NS}$ for which the linear architecture seems to be the best one among the three (Remark 1), and some other subsets of $\mathcal{NS}$ for which it is the Fibonacci architecture (Subsection 5.F.2).

Conjecture 15 (Dyck paths). *We conjecture that the same result actually holds without the tilde,* i.e. *the right multiplication* $\mathtt{P}^{\boxtimes k}$ *gives the highest* CHSH*-value of* $\mathrm{Orbit}^{(k)}(\mathtt{P})$*, as observed numerically. An idea of the proof could be to use Dyck paths. Each time we open/close a parenthesis, the path goes up/down respectively, which produces a certain Dyck path. The statement to be proved is that each time we convert a* $\vee$ *into a* $\wedge$*, the* CHSH*-value is non-decreasing. Then, we would have that the best Dyck path is necessarily the one that always goes up first and then always goes down, which corresponds to the multiplication of boxes on the right.*

As previously mentioned, the next theorem is concurrent and independent of the work of [23]:

Theorem 16 (New collapsing boxes). *The techniques described in* Subsection 2.2 *allow the discovery of new collapsing boxes. See new collapsing areas in* Figure 10*.*

Proof. See Figure 7 for an intuition of the proof. Take the starting box P with coordinates $(0.627, 0.862)$ in the affine plane $\mathcal{A} = \mathrm{Aff}\{\mathtt{PR}, \mathtt{SR}, \mathtt{I}\}$, where the coordinate system is given by the CHSH′- and CHSH- values of P. On the one hand, the tilted orbit of P intersects the collapsing area that was found in [10] (in dark green), since for instance $\mathtt{CHSH}(\mathtt{P}^{\boxtimes 5}) \approx 0.913 > 0.908 \approx \frac{3+\sqrt{6}}{6}$, so P is collapsing by Proposition 8. On the other hand, this box P does not lie in any of the previously-known collapsing areas from [8, 10, 11, 13, 18] (to the best of our knowledge, these five references are the only previous results showing a collapse of communication complexity, in addition to [23] which concurrently and independently found a similar result to ours as mentioned before). Indeed, it is not in the collapsing areas from [10, 18] since $\mathtt{CHSH}(\mathtt{P}) = 0.862 < 0.908$, nor is it in the collapsing area from [8] since $A + B \approx 14.13 < 16$ (using the authors' notation). The box P neither is in any of the collapsing regions found in [11] since it does not belong to the boundary $\partial\mathcal{NS}$ of the non-signalling set. The last area to check is the one from [13], which was numerically found. From a box P, they define a sequence of boxes using "pairwise" multiplications: $\mathtt{Q}_1 := \mathtt{P}$ and $\mathtt{Q}_{n+1} := \mathtt{Q}_n \boxtimes \mathtt{Q}_n$ for $n \geq 1$, and they check whether or not there exists an integer n such that $\mathtt{CHSH}(\mathtt{Q}_n) > \frac{3+\sqrt{6}}{6}$. But, for our starting box P, none of the $\mathtt{Q}_n$ satisfy this inequality: indeed, for $1 \leq n \leq 5$, it possible to check it by hand, for $n = 6$ we have $\mathtt{Q}_6 \in \mathcal{L}$, and for $n \geq 7$ we also have $\mathtt{Q}_n \in \mathcal{L}$ since $\mathcal{L}$ is closed under wirings [1]. Hence our example P is a new collapsing box. □

2.4 Some Other Orbits

In the previous subsection, we studied specifically the orbit of the wiring W_{BS} in the slice of $\mathcal{NS}$ passing through the boxes PR, SR, I. Here we comment on some examples of other orbits in three different ways: (i) it is possible to study the same wiring W_{BS} but in different slices of $\mathcal{NS}$; (ii) it is possible to study another wiring than W_{BS} but to keep the same slice as in the previous subsection; (iii) it is possible to change both the wiring and the slice. See Appendix A for many drawings.

(i) We keep the wiring W_{BS} and we consider the slice of $\mathcal{NS}$ passing through $\mathtt{PR}$, $\mathtt{P}_0$, $\mathtt{P}_1$. Notice that we prove in Proposition 22 that this slice is actually precisely the convex combination of the three points. We draw two examples of such an orbit in Figure 8, with two different starting boxes. We observe that both of them seem to recover the alignment and parallelism properties that we showed in Theorem 10 and its corollary. Again, these lines seem all parallel to what we called previously the diagonal line $\mathfrak{L}_D$, which is defined as the line passing through $\mathtt{PR}$ and $\mathtt{SR} = \frac{1}{2}(\mathtt{P}_0 + \mathtt{P}_1)$. Notice that we show in Theorem 23 that all the boxes of this triangle are actually collapsing, except the ones in the segment $\mathrm{Conv}\{\mathtt{P}_0, \mathtt{P}_1\}$, drawn in pink.

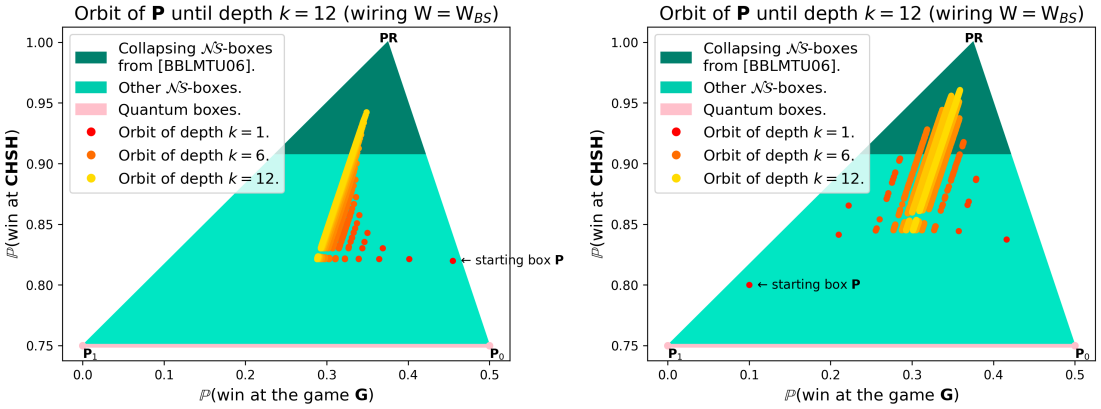


Figure 8: Orbit of W_{BS} in a different slice than in Subsection 2.3: here we consider the slice of $\mathcal{NS}$ passing through $\mathtt{PR}$, $\mathtt{P}_0$, and $\mathtt{P}_1$. We represent the orbit with two different starting boxes. Each orbit is drawn with depth going until $k = 12$. The game $\mathtt{G}$ is defined by the winning rule $a = 0$ and $b = y$. Notice that we give a proof based on CC that this triangle $\mathrm{Conv}\{\mathtt{PR}, \mathtt{P}_0, \mathtt{P}_1\}$ is a quantum void in Corollary 27, which is why the only quantum boxes in this triangle are actually local boxes.

(ii) Among the "typical" wirings defined in Subsection 1.4, the only ones that stabilize the plane $\mathrm{Aff}\{\mathtt{PR}, \mathtt{SR}, \mathtt{I}\}$ are $\mathsf{W}_{\oplus}$ and W_{BS}, see Appendix C. This is why, for these two wirings, we can conveniently draw the orbits in a plane. The orbit of $\mathsf{W}_{\oplus}$ is drawn in Appendix A (a). We observe that each k-orbit contains only one element, which is not surprising since we know from Figure 5 that its induced algebra is associative, meaning that the choice of parenthesization does not lead to a different result. In the same appendix, we also draw the orbit for three other wirings. Surprisingly, we observe that the three new orbits look the same as the orbit of W_{BS}.

(iii) We add the slice $\mathtt{PR}$, $\mathtt{P}_0$, $\mathtt{P}_1$ for each wiring of Appendix A. We observe that the alignment and parallelism properties seem to still hold in those cases, as before. Moreover, we see that the example (d) distills the $\mathtt{CHSH}$-value better than the other examples in that slice.

2.5 Proof of Theorem 13

Recall that $\mathtt{SR} := (\mathtt{P}_0 + \mathtt{P}_1)/2$ is the *shared randomness* box. Given a non-signalling box $\mathtt{P} \in \mathcal{NS}$, its $\mathtt{CHSH}$- and $\mathtt{CHSH}'$- values are defined as follows:

$$\mathtt{CHSH}(\mathtt{P}) := \frac{1}{4} \sum_{a \oplus b = xy} \mathtt{P}(a, b \,|\, x, y)\,, \qquad \text{and} \qquad \mathtt{CHSH}'(\mathtt{P}) := \frac{1}{4} \sum_{\substack{a \,\oplus\, b \\ = (x \,\oplus\, 1)(y \,\oplus\, 1)}} \mathtt{P}(a, b \,|\, x, y)\,.$$

For example, we have $\mathtt{CHSH}(\mathtt{PR}) = 1$ and $\mathtt{CHSH}(\mathtt{SR}) = \frac{3}{4}$ and $\mathtt{CHSH}(\mathtt{I}) = \frac{1}{2}$. Denote $\mathcal{A}$ the affine space $\mathcal{A} := \mathrm{Aff}\{\mathtt{PR}, \mathtt{SR}, \mathtt{I}\}$, and denote $\widetilde{\mathcal{A}} \subseteq \mathcal{A}$ the set of boxes $\mathtt{P}$ in the convex hull $\mathrm{Conv}\{\mathtt{PR}, \mathtt{SR}, \mathtt{I}\}$ whose $\mathtt{CHSH}$-value is $\geq 3/4$. We will prove our results in $\widetilde{\mathcal{A}}$; by the symmetry of the problem, similar results also hold in other areas, such as $2\mathtt{I} - \widetilde{\mathcal{A}}$ the symmetric of $\widetilde{\mathcal{A}}$ by $\mathtt{I}$.

Lemma 17 (Multiplying by $\mathtt{P}$ preserves the $\mathtt{CHSH}$-value order)**.** *Let $\mathtt{P} \in \widetilde{\mathcal{A}}$, and let $\mathtt{Q} \neq \mathtt{R} \in \mathcal{A}$ such that the line $\mathrm{Aff}\{\mathtt{Q}, \mathtt{R}\}$ is parallel to the diagonal line $\mathfrak{L}_D := \mathrm{Aff}\{\mathtt{PR}, \mathtt{SR}\}$. We have:*

$$\mathtt{CHSH}(\mathtt{Q}) \geq \mathtt{CHSH}(\mathtt{R}) \quad \Longrightarrow \quad \begin{cases} \mathtt{CHSH}(\mathtt{Q} \boxtimes \mathtt{P}) \geq \mathtt{CHSH}(\mathtt{R} \boxtimes \mathtt{P})\,, \\ \mathtt{CHSH}(\mathtt{P} \boxtimes \mathtt{Q}) \geq \mathtt{CHSH}(\mathtt{P} \boxtimes \mathtt{R})\,. \end{cases}$$

Proof. As the box $\mathtt{P}$ lies in $\widetilde{\mathcal{A}}$, it is of the form $\mathtt{P} = p_1\mathtt{PR} + p_2\mathtt{SR} + (1-p_1-p_2)\mathtt{I}$ for some coefficients $p_1, p_2 \geq 0$ such that $p_1 + p_2 \leq 1$. Rewrite it as $\mathtt{P} = \binom{p_1}{p_2}$, and similarly denote $\mathtt{Q} = \binom{q_1}{q_2}$ and $\mathtt{R} = \binom{r_1}{r_2}$ for some coefficients $q_i, r_j \in \mathbb{R}$. By the parallelism assumption, vectors $\mathtt{Q} - \mathtt{R}$ and $\mathtt{PR} - \mathtt{SR}$ have to be colinear, *i.e.* there must exist some $\lambda \in \mathbb{R}^*$ such that $\mathtt{Q} - \mathtt{R} = \lambda(\mathtt{PR} - \mathtt{SR}) = \lambda\binom{1}{-1}$, so we may rewrite the second coefficient of $\mathtt{R}$ as $r_2 = q_1 + q_2 - r_1$. With this notation, we can use the linearity of the function $\mathtt{CHSH}(\cdot)$ to see that condition $\mathtt{CHSH}(\mathtt{Q}) \geq \mathtt{CHSH}(\mathtt{R})$ simplifies to $(q_1 - r_1) \geq 0$:

$$\begin{aligned}\mathtt{CHSH}(\mathtt{Q}) - \mathtt{CHSH}(\mathtt{R}) &= (q_1 - r_1) \times \mathtt{CHSH}(\mathtt{PR}) + (q_2 - r_2) \times \mathtt{CHSH}(\mathtt{SR}) + \Big((1-q_1-q_2) - (1-r_1-r_2)\Big) \times \mathtt{CHSH}(\mathtt{I}) \\ &= (q_1 - r_1) \times 1 - (q_1 - r_1) \times \tfrac{3}{4} + 0 \times \tfrac{1}{2} \\ &= \tfrac{1}{4}(q_1 - r_1)\,.\end{aligned}$$

Now, using the multiplication table from Figure 4 and bilinearity of $\boxtimes$, we may compute the following expressions:

$$\begin{aligned}\mathtt{CHSH}(\mathtt{Q} \boxtimes \mathtt{P}) - \mathtt{CHSH}(\mathtt{R} \boxtimes \mathtt{P}) &= \tfrac{1}{8}(p_1 + 2p_2)(q_1 - r_1) \geq 0\,, \\ \mathtt{CHSH}(\mathtt{P} \boxtimes \mathtt{Q}) - \mathtt{CHSH}(\mathtt{P} \boxtimes \mathtt{R}) &= \tfrac{1}{16}(1 - p_1 + p_2)(q_1 - r_1) \geq 0\,,\end{aligned}$$

which gives the desired result. □

Lemma 18 (Right multiplication gives better CHSH-value)**.** *For any* $\mathtt{P} \in \widetilde{\mathcal{A}}$ *and* $\mathtt{Q} \in \widetilde{\mathrm{Orbit}}(\mathtt{P})$*, we have:*

$$\mathtt{CHSH}(\mathtt{Q}) \geq \mathtt{CHSH}(\mathtt{P}) \quad \Longrightarrow \quad \mathtt{CHSH}\Big(\mathtt{Q} \boxtimes \mathtt{P}\Big) \geq \mathtt{CHSH}\Big(\mathtt{P} \boxtimes \mathtt{Q}\Big).$$

Proof. Use the coordinate system (x, y) given by the $\mathtt{CHSH}'$- and $\mathtt{CHSH}$- values respectively in order to write $\mathtt{P}$ and $\mathtt{Q}$ as taking coordinates $(x_\mathtt{P}, y_\mathtt{P})$ and $(x_\mathtt{Q}, y_\mathtt{Q})$. For instance we have $\mathtt{PR} : (\frac{1}{2}, 1)$ and $\mathtt{SR} : (\frac{3}{4}, \frac{3}{4})$ and $\mathtt{I} : (\frac{1}{2}, \frac{1}{2})$. Use the multiplication table from Equation (11) and apply the bilinearity of $\boxtimes$ in order to obtain the following expression:

$$\mathtt{CHSH}\Big(\mathtt{Q} \boxtimes \mathtt{P}\Big) - \mathtt{CHSH}\Big(\mathtt{P} \boxtimes \mathtt{Q}\Big) = \tfrac{1}{8}(12x_\mathtt{P}y_\mathtt{Q} - 12y_\mathtt{P}x_\mathtt{Q} - 7x_\mathtt{P} + 7y_\mathtt{P} + 7x_\mathtt{Q} - 7y_\mathtt{Q}) =: f_\mathtt{P}(x_\mathtt{Q}, y_\mathtt{Q})\,.$$

For any fixed $\mathtt{P} \in \widetilde{\mathcal{A}}$, we want to show that $f_\mathtt{P}(x_\mathtt{Q}, y_\mathtt{Q}) \geq 0$. By construction, we know that $\mathtt{P} \in \mathfrak{L}_1$ and $\mathtt{Q} \in \mathfrak{L}_k$ for some $k \geq 1$, so by Corollary 12 we have $x_\mathtt{Q} + y_\mathtt{Q} \leq x_\mathtt{P} + y_\mathtt{P}$, which we may rewrite as $x_\mathtt{Q} \leq x_\mathtt{P} + y_\mathtt{P} - y_\mathtt{Q}$. As $\mathtt{P}$ lies in $\widetilde{\mathcal{A}}$, we have $y_\mathtt{P} \geq \frac{3}{4}$, so the first partial derivative is non-positive: $\frac{\partial}{\partial x_\mathtt{Q}} f_\mathtt{P}(x_\mathtt{Q}, y_\mathtt{Q}) = \frac{1}{8}(7 - 12y_\mathtt{P}) \leq -1/4 \leq 0$, which means that the function $f_\mathtt{P}(\cdot, y_\mathtt{Q})$ is decreasing over $\mathbb{R}$ for any fixed $y_\mathtt{Q}$. It yields the following inequalities:

$$f_\mathtt{P}(x_\mathtt{Q}, y_\mathtt{Q}) \;\geq\; f_\mathtt{P}(x_\mathtt{P} + y_\mathtt{P} - y_\mathtt{Q}, y_\mathtt{Q}) \;=\; \tfrac{3}{2}\big(y_\mathtt{Q} - y_\mathtt{P}\big)\big(x_\mathtt{P} + y_\mathtt{P} - \tfrac{7}{6}\big) \;\geq\; 0\,,$$

since both factors are non-negative: the first one is non-negative using the hypothesis $\mathtt{CHSH}(\mathtt{Q}) \geq \mathtt{CHSH}(\mathtt{P})$, and the second one is non-negative using $x_\mathtt{P} \geq 1/2$ and $y_\mathtt{P} \geq 3/4$ since $\mathtt{P} \in \widetilde{\mathcal{A}}$. Hence $f_\mathtt{P}$ is non-negative and we obtain the wanted result. □

Recall that the set $\widetilde{\mathrm{Orbit}}^{(k)}(\mathtt{P})$ is called the tilted k-orbit of the box $\mathtt{P}$ and contains some boxes $\mathtt{Q}$ that are generated by applying a wiring to copies of $\mathtt{P}$. We say that this tilted k-orbit *distills the* $\mathtt{CHSH}$*-value* if it contains a box $\mathtt{Q}$ such that $\mathtt{CHSH}(\mathtt{Q}) \geq \mathtt{CHSH}(\mathtt{P})$. In that distilling scenario, we can compute the expression of a box achieving the best $\mathtt{CHSH}$-value:

Statement (Theorem 13)**.** *Let* $\mathtt{P} \in \widetilde{\mathcal{A}}$ *be a box, and let* $k \geq 2$ *be an integer such that the tilted* $(k-1)$*-orbit distills the* $\mathtt{CHSH}$*-value. Then the highest* $\mathtt{CHSH}$*-value of* $\widetilde{\mathrm{Orbit}}^{(k)}(\mathtt{P})$ *is achieved at a box whose expression is the product of* k *times* $\mathtt{P}$ on the right*:*

$$\mathtt{P}^{\boxtimes k} := \Big(\big((\mathtt{P} \boxtimes \mathtt{P}) \boxtimes \mathtt{P}\big) \cdots\Big) \boxtimes \mathtt{P} \;\in\; \operatorname*{argmax}_{\mathtt{Q} \in \widetilde{\mathrm{Orbit}}^{(k)}(\mathtt{P})} \mathtt{CHSH}(\mathtt{Q})\,.$$

Proof. We prove the result by induction on $k \geq 2$. It is obviously true for $k = 2$ since $\widetilde{\text{Orbit}}^{(2)}(\mathtt{P})$ only contains $\mathtt{P} \boxtimes \mathtt{P}$. Now, fix $k \geq 2$ and assume $\mathtt{CHSH}(\mathtt{P}^{\boxtimes k}) \geq \mathtt{CHSH}(\mathtt{Q})$ for any $\mathtt{Q}$ in the tilted k-orbit (induction hypothesis). Assume as well that $\mathtt{CHSH}(\mathtt{P}^{\boxtimes k}) \geq \mathtt{CHSH}(\mathtt{P})$ (distillation hypothesis). We want to show that:

$$\mathtt{CHSH}(\mathtt{P}^{\boxtimes k+1}) \geq \mathtt{CHSH}(\mathtt{Q} \boxtimes \mathtt{P}) \qquad \text{and} \qquad \mathtt{CHSH}(\mathtt{P}^{\boxtimes k+1}) \geq \mathtt{CHSH}(\mathtt{P} \boxtimes \mathtt{Q}) \,,$$

for all $\mathtt{Q}$ in the tilted k-orbit. The first inequality follows from Lemma 17 using the relation $\mathtt{P}^{\boxtimes k+1} = \mathtt{P}^{\boxtimes k} \boxtimes \mathtt{P}$ and the induction assumption. For the other inequality, start from $\mathtt{CHSH}(\mathtt{P}^{\boxtimes k+1}) = \mathtt{CHSH}(\mathtt{P}^{\boxtimes k} \boxtimes \mathtt{P})$ and apply Lemma 18 in order to get $\geq \mathtt{CHSH}(\mathtt{P} \boxtimes \mathtt{P}^{\boxtimes k})$. Then conclude using Lemma 17 and the induction hypothesis in order to obtain $\geq \mathtt{CHSH}(\mathtt{P} \boxtimes \mathtt{Q})$ for any $\mathtt{Q}$ in the tilted k-orbit. □

3 Numerical Optimization on the Set of Wirings

We saw in the previous section that, given a non-signalling box $\mathtt{P}$, there may exist a wiring W that sufficiently distills the box $\mathtt{P}$ in order to collapse communication complexity. The question we address in this section is the following: if the box $\mathtt{P}$ is fixed, how to find a wiring W good enough to collapse communication complexity (when it is possible)? The difficulty is that, for each input $x, y \in \{0, 1\}$, there are 82 possible deterministic wirings [45], leading to a total number of $82^4 \approx 10^8$ possible deterministic wirings. So a naive discrete optimization over deterministic wirings seems inefficient. To that end, we present two optimization algorithms: (i) an algorithm that tests many different combinations of wirings and that is suitable for numerical simulations, and (ii) another one that finds a "uniform" collapsing wiring W in a whole region of boxes, which is appropriate for deriving an analytical proof (see the next section). This section might be skipped at first reading as it is more technical. See our GitHub page for the details of the algorithms [9].

Remark 19 (Comparison with [11, 23]). We now compare and contrast our methods with two recent works that also study optimization over wirings:

(i) In [11], the authors suggest reducing the 82^4 possible deterministic wirings for Alice and Bob to only 3152 by simply considering the ones that preserve the $\mathtt{PR}$ box, *i.e.* wirings W such that $\mathtt{PR} \boxtimes_{\mathsf{W}} \mathtt{PR} = \mathtt{PR}$, and then doing a discrete optimization over that smaller set. This smaller set encompasses for instance the wirings W_{BS}, W_{dist} but discards $\mathsf{W}_{\oplus}$, $\mathsf{W}_{\wedge}$, $\mathsf{W}_{\vee\wedge}$ (see definitions in Subsection 1.4); see also [23, Supplementary Material, I] which mentions that even some optimal wirings are discarded. This technique allows them to analytically prove that many new areas of boxes are collapsing.

(ii) In [23], the authors use a mix of exhaustive search and linear programming. For each of Bob's 82^2 extremal half-wirings, they apply linear programming to optimize Alice's half-wiring, and then they select the best pair of half-wirings. This allows them to numerically find optimal wirings for any pair of boxes, which leads them to discover new collapsing boxes.

(iii) In our work, we use an efficient variant of the Gradient Descent algorithm, based on Line Search methods, frequent resets, and parallel descents. A limitation in the method from [11] could come from the fact that many wirings are discarded; this is why we choose to take our feasible set to be the entire set of mixed wirings $\mathcal{W} \subseteq [0, 1]^{32}$. In [23, Supplementary Material, II], the authors implement a sequence of different optimal wirings, which we do similarly in what we call later Task A, but we also implement a uniform version of it in Task B, which allows us to find a single optimal wiring for a whole region of boxes (instead of a sequence of wirings) and then to prove by hand the collapse of communication complexity for those boxes. In this manner, we recover both the numerical results of [23] (see details in Subsection 4.1) and the analytical results of [11] (see details in Subsection 4.2).

3.1 Goals of the Algorithms

Task A. To prove that a box $\mathtt{P}$ is collapsing, a particular case of Proposition 8 says that it is enough to find a finite sequence of wirings $(\mathsf{W}_1, \dots, \mathsf{W}_N)$ such that the following box is collapsing:

$$\mathtt{P}_{N+1} := \Big(\big((\mathtt{P} \underset{\mathsf{W}_1}{\boxtimes} \mathtt{P}) \underset{\mathsf{W}_2}{\boxtimes} \mathtt{P} \big) \underset{\mathsf{W}_3}{\boxtimes} \dots \Big) \underset{\mathsf{W}_N}{\boxtimes} \mathtt{P} \,.$$

Note that we need to specify the parenthetization because the different products $\boxtimes_{\mathsf{W}_i}$ are potentially non-associative. Among the numerous possibilities, we choose the parenthesization on the left because it is easy to implement and because it is the best one when the wiring is W_{BS}, see Theorem 13. This algorithm will consist in an iterative construction of the sequence $(\mathsf{W}_i)_i$: first, find a wiring W_1 such that the $\mathtt{CHSH}$-value of the box $\mathtt{P}_2 := \mathtt{P} \boxtimes_{\mathsf{W}_1} \mathtt{P}$ is the highest possible, then find W_2 such that the $\mathtt{CHSH}$-value of the box $\mathtt{P}_3 := \mathtt{P}_2 \boxtimes_{\mathsf{W}_2} \mathtt{P}$ is the highest possible, so on and so forth until the N-th iteration. If the $\mathtt{CHSH}$-value of the box $\mathtt{P}_{N+1}$ is above the threshold $\frac{3+\sqrt{6}}{6} \approx 0.91\%$, we know that communication complexity collapses [10], so the starting box $\mathtt{P}$ is collapsing as well. Otherwise, we cannot conclude whether $\mathtt{P}$ is collapsing or not.

Task B. The goal of this algorithm is essentially the same as the first one, but we add a strong constraint: we want all the W_i to be the same wiring W:

$$\Big(\big((\mathtt{P} \underset{\mathsf{W}}{\boxtimes} \mathtt{P}) \underset{\mathsf{W}}{\boxtimes} \mathtt{P} \big) \underset{\mathsf{W}}{\boxtimes} \dots \Big) \underset{\mathsf{W}}{\boxtimes} \mathtt{P} \;=:\; \mathtt{P}^{\boxtimes_{\mathsf{W}} N+1} \,.$$

In that sense, this is a "uniform" version of the first algorithm. The interest of this algorithm is that it helps to give analytical proofs (see Section 4): if the value $\mathtt{CHSH}(\mathtt{P}^{\boxtimes_{\mathsf{W}} N})$ is above the threshold $\frac{3+\sqrt{6}}{6} \approx 0.91\%$ for some N, then by continuity of $\boxtimes_{\mathsf{W}}$, there is an open neighborhood around $\mathtt{P}$ such that for any $\mathtt{Q}$ close enough to $\mathtt{P}$ we also have that $\mathtt{CHSH}(\mathtt{Q}^{\boxtimes_{\mathsf{W}} N})$ is above the threshold, and therefore the whole neighborhood of $\mathtt{P}$ is collapsing. This technique will help to discover wide collapsing areas and to provide analytical proofs by hand.

3.2 Toy Example ($N = 1$)

In this subsection, we treat the case when there is only one product $\boxtimes_{\mathsf{W}}$ between two boxes $\mathtt{Q}, \mathtt{P} \in \mathcal{NS}$. We detail the maximization algorithm we use: a Projected Gradient Descent. The optimization problem consists in finding W^* as follows:

$$\mathsf{W}^* \;=\; \underset{\mathsf{W} \in \mathcal{W}}{\operatorname{argmax}} \;\; \Phi(\mathsf{W}) \,. \tag{12}$$

where the objective function is $\Phi(\mathsf{W}) := \mathtt{CHSH}(\mathtt{Q} \boxtimes_{\mathsf{W}} \mathtt{P})$ for some fixed non-signalling boxes $\mathtt{Q}, \mathtt{P}$, and where $\mathcal{W}$ is the set of mixed wirings introduced in Definition 4, which we recall below.

The Constraint $\mathsf{W} \in \mathcal{W}$. Recall that a *mixed wiring* W between two boxes $\mathtt{Q}, \mathtt{P} \in \mathcal{NS}$ is the data of six functions $f_1, f_2, g_1, g_2 : \{0,1\}^2 \to [0,1]$ and $f_3, g_3 : \{0,1\}^3 \to [0,1]$ satisfying the following *non-cyclicity conditions*:

$$\forall x, \qquad \big(f_1(x,0) - f_1(x,1)\big)\big(f_2(x,0) - f_2(x,1)\big) \;=\; 0 \,, \tag{13}$$

$$\forall y, \qquad \big(g_1(y,0) - g_1(y,1)\big)\big(g_2(y,0) - g_2(y,1)\big) \;=\; 0 \,. \tag{14}$$

Recall that the corresponding diagram can be found in Figure 2 (b), and that mixed wirings form a set that we denote $\mathcal{W}$. In our algorithms, we view W a real vector with $4 \times 2^2 + 2 \times 2^3 = 32$ variables. This vector stores each value of each function:

$$\mathsf{W} = [f_1(0,0) \quad f_1(0,1) \quad f_1(1,0) \quad f_1(1,1) \quad g_1(0,0) \quad \dots]^\top \in \mathbb{R}^{32} \,. \tag{15}$$

To satisfy the normalization constraint that the f_i, g_j take value in $[0,1]$, and the non-cyclicity conditions (13) and (14) (which are non-linear conditions), we implement a projection function $\mathtt{proj} : \mathbb{R}^{32} \to \mathbb{R}^{32}$ in Algorithm 1. Notice that our real code is written in a vectorized fashion

Algorithm 1: Projection function `proj` on the feasible set $\mathcal{W}$. Vectorized version in our GitHub page [9].

Data: $\mathsf{W} = \begin{bmatrix} w_1 & \dots & w_{32} \end{bmatrix} = \begin{bmatrix} f_1(0,0) & f_1(0,1) & f_1(1,0) & f_1(1,1) & g_1(0,0) & \dots & g_3(1,1,1) \end{bmatrix} \in \mathbb{R}^{32}$.
Result: $\texttt{proj}(\mathsf{W}) \in \mathcal{W}$.

$\mathsf{W} \leftarrow \begin{bmatrix} \max\{w_1, 0\} & \dots & \max\{w_{32}, 0\} \end{bmatrix}$;
$\mathsf{W} \leftarrow \begin{bmatrix} \min\{w_1, 1\} & \dots & \min\{w_{32}, 1\} \end{bmatrix}$;
for $x \in \{0,1\}$ **do**
  if $|f_1(x,0) - f_1(x,1)| \leq |f_2(x,0) - f_2(x,1)|$ **then** $f_1(x,0), f_1(x,1) \leftarrow \big(f_1(x,0) + f_1(x,1)\big)/2$;
  else $f_2(x,0), f_2(x,1) \leftarrow \big(f_2(x,0) + f_2(x,1)\big)/2$;
end
for $y \in \{0,1\}$ **do**
  if $|g_1(y,0) - g_1(y,1)| \leq |g_2(y,0) - g_2(y,1)|$ **then** $g_1(y,0), g_1(y,1) \leftarrow \big(g_1(y,0) + g_1(y,1)\big)/2$;
  else $g_2(y,0), g_2(y,1) \leftarrow \big(g_2(y,0) + g_2(y,1)\big)/2$;
end
return W

and is difficult to read as such, so we only present the idea here. Moreover, we use the package PyTorch [39] for automatic differentiation.

The Objective Function $\Phi(\mathsf{W}) := \texttt{CHSH}(\texttt{Q} \boxtimes_{\mathsf{W}} \texttt{P})$. In our algorithms, we view a box $\texttt{P}$ as a $2\times2\times2\times2$-tensor: $\texttt{P}[a,b,x,y] := \texttt{P}(a,b\,|\,x,y)$ with $a,b,x,y \in \{0,1\}$, whose entries are float numbers between 0 and 1. Two things need to be computed separately: $\texttt{Q} \boxtimes_{\mathsf{W}} \texttt{P}$ and then $\texttt{CHSH}(\cdot)$. On the one hand, the product $\texttt{P} \boxtimes_{\mathsf{W}} \texttt{P}$ is computed using Equation (9), which we vectorized in our algorithm using five types of operations: tensor transposition $\top$, tensor sum $+$, tensor product $\otimes$, contraction of tensors $\cdot$, and entry-wise multiplication $*$; see details in the pdf document of our GitHub page [9]. On the other hand, the function $\texttt{CHSH}(\cdot)$ is a linear function that computes the `CHSH`-value of a box, implemented with a dot product as follows:

$$\texttt{CHSH}(\texttt{R}) := \frac{1}{4} \sum_{a \oplus b = xy} \texttt{R}(a,b\,|\,x,y) \;=\; \langle \texttt{R}, \texttt{T} \rangle \,,$$

where $\texttt{T}$ is the $2 \times 2 \times 2 \times 2$-tensor defined as $\texttt{T}[a,b,x,y] = 1/4$ if $a \oplus b = xy$, and $= 0$ otherwise.

3.2.1 Naive Gradient Descent

To gain insight into the complexity of the optimization problem, we begin by studying a basic algorithm, the Projected Gradient Descent, with a small learning rate ($\alpha \ll 1$) and a lot of iterations ($K \gg 1$). We will obtain a histogram of the frequency of the different results we obtain, see Figure 9 (a).

Projected Gradient Descent. We implement a "projected" version of the Gradient Descent algorithm in order to satisfy the constraint $\mathsf{W} \in \mathcal{W}$ at each step. It simply means that each iteration is projected on the feasible set:

$$\mathsf{W}^{k+1} = \texttt{proj}\big(\mathsf{W}^k + \alpha\, \nabla\Phi(\mathsf{W}^k)\big) \,,$$

where $\alpha \in \mathbb{R}$ is the learning rate. Our implementation can be found in Algorithm 2. We compute the gradient of the objective function using the automatic differentiation Python package `torch.autograd` that provides us with the commands `backward` and `grad`. As we do not have a good intuition of what could be a good wiring W in $\mathcal{W}$ to start with given a fixed box $\texttt{P}$, we take a random initialization: W^0 is uniformly generated in the hypercube $[0,1]^{32}$. As such, the vector W^0 is not necessarily a well-defined mixed wiring since it does not necessarily satisfy the non-cyclicity conditions (13) and (14), but this problem is fixed after one iteration in the Projected Gradient

Algorithm 2: Projected Gradient Descent. More details on our GitHub page [9].

Data: $\Phi : \mathbb{R}^{32} \to \mathbb{R}$ objective function, $\alpha \in \mathbb{R}$ learning rate, $K \in \mathbb{N}$ number of iterations, $\varepsilon > 0$ tolerance.

Result: $\mathsf{W}_{\text{out}} \approx \text{argmax}_{\mathsf{W}}\, \Phi(\mathsf{W}) \in \mathcal{W} \subseteq \mathbb{R}^{32}$.

$\mathsf{W} \sim \mathcal{U}([0,1]^{32})$;
for $k \in \{0, \dots, K-1\}$ **do**
  $\mathsf{W}_{\text{old}} \leftarrow \mathsf{W}$;
  $\mathsf{W} \leftarrow \texttt{proj}\big(\mathsf{W} + \alpha\, \nabla\Phi(\mathsf{W})\big)$ using Algorithm 1 ;
  if $||\mathsf{W} - \mathsf{W}_{\text{old}}||_\infty < \varepsilon$ **then** break ;
end
return $\mathsf{W}_{\text{out}} := \mathsf{W} \in \mathbb{R}^{32}$.

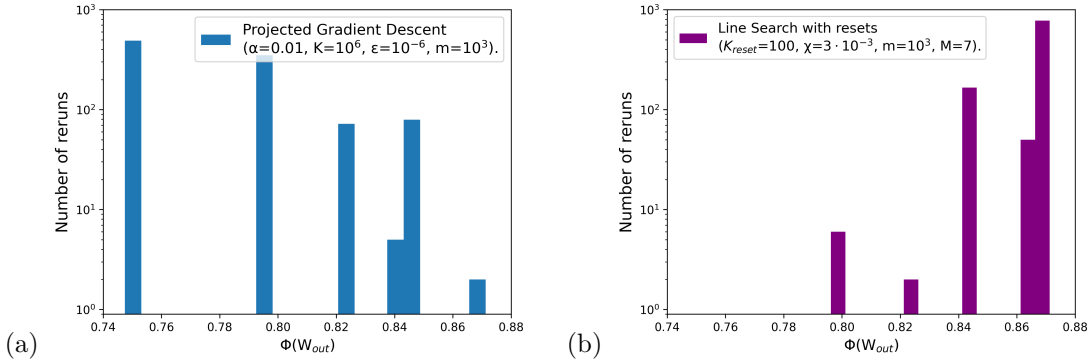


Figure 9: Histogram of the evaluations of the objective function Φ applied at the output W_{out} of (a) Algorithm 2 and (b) Algorithm 3. As expected, we observe that the latter is more efficient than the former in maximizing Φ, for equivalent computation duration.

Descent algorithm since the wiring is then projected. Otherwise, one can also directly apply $\texttt{proj}$ to W^0. The notation $\mathsf{W} \sim \mathcal{U}(X)$ means that we uniformly generate W in the set X.

Estimating the Proportion of "Good" Outputs. We use Algorithm 2 with a learning rate $\alpha = 0.01$, a number of iterations of $K = 10^6$, a tolerance of $\varepsilon = 10^{-6}$, and we obtain the histogram presented in Figure 9 (a). Recall that the objective function is $\Phi(\mathsf{W}) := \texttt{CHSH}(\texttt{Q} \boxtimes_{\mathsf{W}} \texttt{P})$; this histogram is drawn with $\texttt{Q} = \texttt{P} = p\texttt{PR} + q\texttt{SR} + (1-p-q)\texttt{I}$, where $p = 0.39$ and $q = 0.6$. The number of reruns is $m = 10^3$, done simultaneously in parallel, which is faster than doing m descents one after another[6]. We observe that the results concentrate on certain discrete values. These values correspond to different attractive points in different basins of attraction (recall that the initial W is taken uniformly at random in $[0,1]^{32}$). As we want to maximize Φ, we are interested in the highest concentrated value ≈ 0.87. In that example, we observe that the proportion of starting wirings such that W_{out} is only $\chi \approx {}^2/_{10^3} = 0.2\%$ using this basic Gradient Descent algorithm. This information tells us that the function Φ is difficult to maximize, which is why we present a more efficient algorithm in the following subsection.

3.2.2 More Efficient Algorithm: Line Search with Resets

In this subsubsection, we present a variant of the Gradient Descent algorithm called Line Search, which we enhance with frequent resets of bad outcomes. See [38] for a standard reference book in numerical optimization. The idea of this algorithm is, instead of always keeping the same α, to

[6] In order to do m gradient descents in parallel efficiently, we "parallelize" Algorithm 2: instead of viewing W as a 32-vector, we view it as a $(32 \times m)$-matrix, where each column represents a different wiring. Comparing this method with a naive FOR loop, we observe a factor of 100 in the speed gain. Notice that when the descent is done, one can post-select the best wiring among the m columns of W_{out}. See our GitHub page [9].

estimate the best coefficient α_k at each step of the descent:

$$\begin{cases} \alpha_k = \operatorname{argmax}_{\alpha\in\mathbb{R}} \Phi\big(\mathsf{W}^k + \alpha\nabla\Phi(\mathsf{W}^k)\big)\,, \\ \mathsf{W}^{k+1} = \mathtt{proj}\big(\mathsf{W}^k + \alpha_k\,\nabla\Phi(\mathsf{W}^k)\big)\,. \end{cases}$$

As we observed in the previous subsection, the proportion χ of "good" starting wirings is very weak, which is why we apply frequent resets: we do $m = 10^3$ descents in parallel but only $K_{\text{reset}} = 100$ steps, then we keep only the best $m \cdot \chi$ wirings and we reset all the others to a new random initialization. Then we repeat that procedure but we reset fewer wirings (say, at the j-th repetition, keep for instance the best $j \cdot m \cdot \chi$ wirings), and we repeat this procedure $1/\chi$ times. In the end, most of the wirings should be in the good basin of attraction, so we can apply one final run of line search, with many more steps so that it converges to the attractor. See Algorithm 3, and we obtain results of Figure 9 (b).

Algorithm 3: Line Search with Resets. More details on our GitHub page [9].

Data: $\Phi : \mathbb{R}^{32} \to \mathbb{R}$ objective function, $K_{\text{reset}} \in \mathbb{N}$ number of iterations before reset, $\chi \in [0,1]$ proportion of "good" random wirings, $m \in \mathbb{N}$ number of descents in parallel, $M \in \mathbb{N}$ number of iterations to compute the best learning rate α^*.
Result: $\mathsf{W}_{\text{out}} \in \mathbb{R}^{32\times m}$, where each column is $\approx \operatorname{argmax}_{\mathsf{W}} \Phi(\mathsf{W})$.

$\mathsf{W} = (32 \times m$ null matrix), whose columns are denoted W_i ;
for $j \in \{0, \dots, \lfloor 1/\chi \rfloor - 1\}$ **do**
  $\mathsf{W} \leftarrow$ among the m columns of W, keep the $j \cdot m \cdot \chi$ ones giving the highest values for the objective function Φ, and reset all the other columns randomly with $\mathcal{U}([0,1]^{32})$;
  if $j = \lfloor 1/\chi \rfloor - 1$ **then** $K_{\text{reset}} \leftarrow 10 \cdot K_{\text{reset}}$;
  for $k \in \{0, \dots, K_{\text{reset}} - 1\}$ **do**
    for $i \in \{0, \dots, m-1\}$ **do** $\alpha_i^* \leftarrow \operatorname{argmax}_{\alpha>0} \Phi\big(\mathsf{W}_i + \alpha\nabla\Phi(\mathsf{W}_i)\big)$ using M iterations ;
    $\mathsf{W} \leftarrow \Big[\mathtt{proj}\big(\mathsf{W}_i + \alpha_i^*\,\nabla\Phi(\mathsf{W}_i)\big)\Big]_i$ using Algorithm 1 ;
  end
end
return $\mathsf{W}_{\text{out}} = \mathsf{W} \in \mathbb{R}^{32\times m}$.

3.3 Task A

Algorithm A is presented in Algorithm 4; it simply consists in applying the toy case $\mathtt{Q} \boxtimes \mathtt{P}$ from the previous subsection recursively N times. We want to find a sequence of wirings $\mathsf{W}_1, ..., \mathsf{W}_N$ such that the CHSH-value is above the following threshold:

$$\mathtt{CHSH}\Bigg(\underbrace{\Big(\big((\mathtt{P} \underset{\mathsf{W}_1}{\boxtimes} \mathtt{P}) \underset{\mathsf{W}_2}{\boxtimes} \mathtt{P}\big) \underset{\mathsf{W}_3}{\boxtimes} \dots\Big) \underset{\mathsf{W}_N}{\boxtimes} \mathtt{P}}_{=:\,\mathtt{P}_{N+1}}\Bigg) > \frac{3+\sqrt{6}}{6}\,.$$

A consequence is that the box P collapses communication complexity [10]. Notice that for some boxes $\mathtt{P} \in \mathcal{NS}$, it might not be possible to find such a sequence of wirings because it is impossible to distill them by any means. This algorithm is used in Subsection 4.1 in order to plot the new regions of collapsing nonlocal boxes.

Remark 20. Going further, once we find $(\mathsf{W}_1^*, ..., \mathsf{W}_N^*)$, it is possible to do a "backward" process: for all $i \in \{1, ..., N\}$, fix W_j^* for $j \neq i$, optimize the function $\mathsf{W}_i \mapsto \mathtt{CHSH}(\mathtt{P}_{N+1})$ and update W_i^*. It is also possible to use neural network methods to optimize all the W_i "at the same time".

Algorithm 4: Task A. More details on our GitHub page [9].

Data: $\mathtt{P} \in [0,1]^{2+2+2+2}$ box, $N \in \mathbb{N}$ number of box products, $(K_{\text{reset}}, \chi, m, M)$ parameters for Algorithm 3.
Result: $[\mathsf{W}_1^*, \dots, \mathsf{W}_n^*] \in \mathcal{W}^n$ for some $n \leq N$.

$\mathtt{P}_1 \leftarrow \mathtt{P}$;
for $n \in \{1, \dots, N\}$ **do**
  $\mathsf{W}_n^* \in [0,1]^{32} \leftarrow \operatorname{argmax}_{\mathsf{W}} \mathtt{CHSH}(\mathtt{P}_n \boxtimes_{\mathsf{W}} \mathtt{P})$ by picking the best column among the m columns of the output $\mathsf{W}_{\text{out}} \in \mathbb{R}^{32\times m}$ of Algorithm 3 ;
  $\mathtt{P}_{n+1} \leftarrow \mathtt{P}_n \boxtimes_{\mathsf{W}_n^*} \mathtt{P}$;
  if $\mathtt{CHSH}(\mathtt{P}_{n+1}) > \frac{3+\sqrt{6}}{6}$ **then return** $[\mathsf{W}_1^*, \dots, \mathsf{W}_n^*]$;
end
return *"Nothing found."*

3.4 Task B

Task B is a "uniform" version of task A, in the sense that we want all the W_i's to be the same. In order words, we want to find a wiring W and an integer N such that:

$$\mathtt{CHSH}\Bigg(\underbrace{\Big(\big((\mathtt{P} \underset{\mathsf{W}}{\boxtimes} \mathtt{P}) \underset{\mathsf{W}}{\boxtimes} \mathtt{P}\big) \underset{\mathsf{W}}{\boxtimes} \dots\Big) \underset{\mathsf{W}}{\boxtimes} \mathtt{P}}_{=:\, \mathtt{P}^{\boxtimes_{\mathsf{W}}(N+1)}}\Bigg) > \frac{3+\sqrt{6}}{6}\,.$$

This algorithm is used in the proof of Theorem 29 in order to find appropriate collapsing wirings for the analytical proof.

Idea of the Algorithm. First, find a wiring $\mathsf{W}_1^* = \operatorname{argmax}_{\mathsf{W}} \mathtt{CHSH}(\mathtt{P} \boxtimes_{\mathsf{W}} \mathtt{P})$ with a Gradient Descent algorithm, and then evaluate the powers of $\mathtt{P}$ with that wiring W_1^* until $N+1$, *i.e.* compute $\mathtt{CHSH}\big(\mathtt{P}^{\boxtimes_{\mathsf{W}_1^*} n}\big)$ for $n = 1, \dots, N+1$. If one of those evaluations is greater than the threshold $(3+\sqrt{6})/6$ from [10], then we can stop the algorithm, it means that the wiring W_1^* achieves the goal. Otherwise, compute $\mathsf{W}_2^* = \operatorname{argmax}_{\mathsf{W}} \mathtt{CHSH}\big(\mathtt{P}^{\boxtimes_{\mathsf{W}} 3}\big)$ and repeat the same evaluation process of the powers of $\mathtt{P}$ as in the previous step. Proceed like this until computing W_M^*, where $M \in \mathbb{N}$ is some hyper-parameter. Typically, we take $M \leq N$ because it is a lot faster to evaluate the N-th power of $\mathtt{P}$ than to optimize the N-th power of $\mathtt{P}$. See the details in Algorithm 5.

Algorithm 5: Task B. More details on our GitHub page [9].

Data: $\mathtt{P} \in [0,1]^{2+2+2+2}$ box, $N \in \mathbb{N}$ maximal tested box power, $L \in \mathbb{N}$ maximal optimized box power, $(K_{\text{reset}}, \chi, m, M)$ parameters for Algorithm 3.
Result: $\mathsf{W}^* \in \mathcal{W}$.

for $\ell \in \{1, \dots, L\}$ **do**
  $\mathsf{W}_\ell^* \leftarrow \operatorname{argmax}_{\mathsf{W}} \mathtt{CHSH}\big(\mathtt{P}^{\boxtimes_{\mathsf{W}} \ell+1}\big)$ using Algorithm 3 ;
  for $n \in \{1, \dots, N+1\}$ **do**
    if $\mathtt{CHSH}\big(\mathtt{P}^{\boxtimes_{\mathsf{W}_\ell^*} n}\big) > \frac{3+\sqrt{6}}{6}$ **then return** W_ℓ^* ;
  end
end
return *"Nothing found."*

4 New Collapsing Boxes

In this section, we present collapsing boxes found in two different ways. (i) First with a numerical approach, using the algorithms (Section 3). (ii) Then with an analytical approach, using the algebra of boxes (Section 1) and the orbit of a box (Section 2).

4.1 Numerical Results

Using Algorithm 4 that addresses Task A, we obtain many collapsing boxes. Some samples are drawn in Figure 10 on some slices of the non-signalling set $\mathcal{NS}$, but note that this algorithm also applies more generally to any desired slice. As previously mentioned, this work is concurrent and independent of the work of [23]. In the drawings, some boxes are denoted $\mathtt{P}_{\mathrm{L}}$ and $\mathtt{P}_{\mathrm{NL}}$, let us recall their definition here. The local set $\mathcal{L}$ and the non-signalling set $\mathcal{NS}$ are polytopes, *i.e.* the convex hull of a finite number of extremal points. The first set $\mathcal{L}$ admits exactly 16 extremal points, called *local extreme points* and denoted $\mathtt{P}_{\mathrm{L}}^{\mu,\nu,\sigma,\tau}$, where $\mu,\nu,\sigma,\tau \in \{0,1\}$. These 16 points are as well extremal points of $\mathcal{NS}$, together with 8 additional extremal points, called *non-local extreme points* and denoted $\mathtt{P}_{\mathrm{NL}}^{\mu,\nu,\sigma}$. They are defined as follows [2, 3]:

$$\begin{aligned}
&\bullet \text{ Local extremal points:} \quad \mathtt{P}_{\mathrm{L}}^{\mu,\nu,\sigma,\tau}(a,b\,|\,x,y) \;:=\; \begin{cases} 1 & \text{if } a=\mu\,x\oplus\nu \text{ and } b=\sigma\,y\oplus\tau, \\ 0 & \text{otherwise,} \end{cases} \\
&\bullet \text{ Nonlocal extremal points:} \quad \mathtt{P}_{\mathrm{NL}}^{\mu,\nu,\sigma}(a,b\,|\,x,y) \;:=\; \begin{cases} 1/2 & \text{if } a\oplus b=xy\oplus\mu\,x\oplus\nu\,y\oplus\sigma, \\ 0 & \text{otherwise.} \end{cases}
\end{aligned} \tag{16}$$

Note that $\mathtt{PR}=\mathtt{P}_{\mathrm{NL}}^{000}$ and $\mathtt{P}_0=\mathtt{P}_{\mathrm{L}}^{0000}$ and $\mathtt{P}_1=\mathtt{P}_{\mathrm{L}}^{0101}$ (we remove the commas in the superscripts for the sake of simplicity of notations).

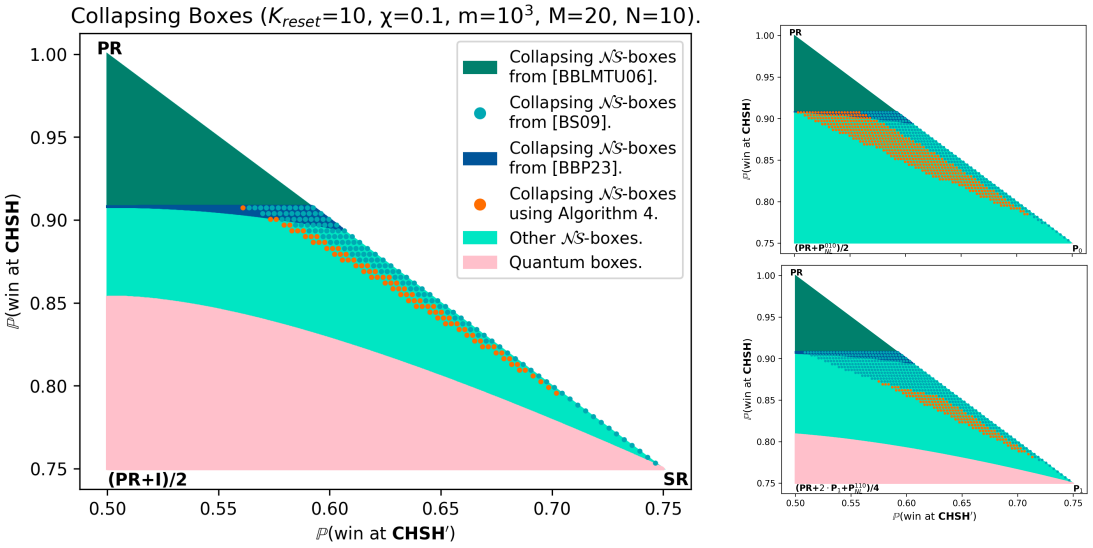


Figure 10: In orange are drawn the collapsing boxes outputted by Algorithm 4. Each drawing represents a different slice of $\mathcal{NS}$; the extreme points of the triangles indicate which slice is drawn and the definition of the boxes $\mathtt{P}_{\mathrm{NL}}$ can be found in Equation (16). The three graphs have the same color legend, displayed at the center, and they are all configured with the same algorithm parameters $(K_{\mathrm{reset}},\chi,m,M,N)$, detailed at the top. We adopt the following convention: (i) boxes that are numerically determined are drawn with dots, (ii) boxes that are analytically determined are drawn in plain regions (there exist explicit equations describing those regions). Notice that the left drawing is similar to [23, Figure 3], which was found using a different algorithm as detailed in Remark 19. The quantum set $\mathcal{Q}$ (in pink) is drawn using formulas from [33]. References: [BBLMTU06]=[10], [BS09]=[13], [BBP23]=[8].

Observe in Figure 10 that, depending on the chosen slice, the collapsing area does not always have the same "shape" nor the same "area". Moreover, notice in the graphs that there seems to exist a collapsing area in the neighborhood below the diagonal segments joining $\mathtt{PR}$ and respectively $\mathtt{SR}$, $\mathtt{P}_0$, $\mathtt{P}_1$. This is actually true. Indeed, we analytically show below in Theorem 23 that those three segments are collapsing, and we also know that the box product $\mathtt{P}\boxtimes_{\mathsf{W}}\mathtt{Q}$ is continuous in $\mathtt{P}$ and $\mathtt{Q}$ for any W (it is even bilinear, recall the expression in (9)), so distillation protocols are continuous and in some sense the orbits are also "continuous", hence there exists an open neighborhood below these diagonal segments that collapses communication complexity.

Remark 21 (Continuous extension of a finite collapsing set). The algorithm only provides us with

finitely many collapsing boxes, but we can still deduce a continuous "extension" of that collapsing set. Indeed, as explained in Subsection 2.2, if we know that a box $\mathtt{P} \in \mathcal{NS}$ collapses communication complexity, then we also know that the cone $\mathcal{C}_{\mathtt{P}}$ is collapsing, where $\mathcal{C}_{\mathtt{P}}$ denotes a certain cone taking origin at $\mathtt{P}$ represented in Figure 6 (b). As a consequence, if we list the collapsing boxes $\{\mathtt{P}_k\}_{1\leq k\leq K}$ obtained by numerical means, we can deduce what follows:

$$\text{The union of cones } \bigcup_{k=1}^{K} \mathcal{C}_{\mathtt{P}_k} \text{ is a collapsing set.}$$

4.2 Analytical Results

4.2.1 The Triangle PR, $\mathtt{P}_0$, $\mathtt{P}_1$ is Collapsing

In this subsection, we extend the result [13] from Brunner and Skrzypczyk, who showed that any box in the segment joining the boxes PR and SR $:= (\mathtt{P}_0 + \mathtt{P}_1)/2$ is collapsing (except the box SR, which is classical). In the following theorem, we recover a result from [11] stating that any box in the triangle joining the boxes $\mathtt{PR}, \mathtt{P}_0, \mathtt{P}_1$ is collapsing (except the boxes in the segment joining $\mathtt{P}_0$ and $\mathtt{P}_1$, which are classical), with a new proof, based on the algebra of boxes. Recall that PR is the non-signalling box that outputs (a, b) such that $a \oplus b = xy$ when (x, y) is inputted, and $\mathtt{P}_0$ and $\mathtt{P}_1$ are respectively the deterministic boxes that output $(0, 0)$ and $(1, 1)$ independently of the inputs. Recall also that the *convex hull* of a set $\{\mathtt{Q}_1, \dots, \mathtt{Q}_N\}$ is the set of all possible convex combinations of these $\mathtt{Q}_i$:

$$\mathrm{Conv}\{\mathtt{Q}_1, \dots, \mathtt{Q}_N\} := \left\{ \sum_{i=1}^{N} q_i \, \mathtt{Q}_i \text{ such that } q_i \geq 0 \text{ and } \textstyle\sum_i q_i = 1 \right\},$$

and the *affine hull* of $\{\mathtt{Q}_1, \dots, \mathtt{Q}_N\}$ has the same definition but without the non-negativity constraint:

$$\mathrm{Aff}\{\mathtt{Q}_1, \dots, \mathtt{Q}_N\} := \left\{ \sum_{i=1}^{N} q_i \, \mathtt{Q}_i \text{ such that } q_i \in \mathbb{R} \text{ and } \textstyle\sum_i q_i = 1 \right\} \supseteq \mathrm{Conv}\{\mathtt{Q}_1, \dots, \mathtt{Q}_N\}.$$

Proposition 22. *The triangle* $C := \mathrm{Conv}\{\mathtt{PR}, \mathtt{P}_0, \mathtt{P}_1\}$ *is a face on the boundary of* $\mathcal{NS}$.

Proof. First, we prove the equality between the sets C and $A := \mathcal{NS} \cap \mathrm{Aff}\{\mathtt{PR}, \mathtt{P}_0, \mathtt{P}_1\}$, meaning that the convex hull C is actually a slice of $\mathcal{NS}$. The first inclusion $C \subseteq A$ is trivial because the three points $\mathtt{PR}, \mathtt{P}_0, \mathtt{P}_1$ are in $\mathcal{NS}$ and because $\mathcal{NS}$ is a polytope so it is stable under taking convex combination. Conversely, recall that by definition $\mathtt{CHSH}(\mathtt{P}) := \frac{1}{4} \sum_{a\oplus b=xy} \mathtt{P}(a, b \,|\, x, y)$, and $\mathtt{CHSH}'(\cdot)$ and $\mathtt{CHSH}''(\cdot)$ are defined similarly but with respective summand conditions $a \oplus b = (x \oplus 1) \cdot (y \oplus 1)$ and $a \oplus b = x \cdot (y \oplus 1)$. As these three functions are linear, they preserve alignment and convexity. It means that if a box $\mathtt{P}$ is of the form $\mathtt{P} = \sum_i q_i \, \mathtt{Q}_i$ for some reals q_i and boxes $\mathtt{Q}_i$, then $\mathtt{CHSH}(\mathtt{P}) = \sum_i q_i \, \mathtt{CHSH}(\mathtt{Q}_i)$, and similarly with $\mathtt{CHSH}'$ and $\mathtt{CHSH}''$. We apply the preservation of alignment in Figure 11 representing the 24 extremal points of $\mathcal{NS}$ [3], and we obtain the following inclusions:

$$\text{(a)} \qquad A \subseteq \mathrm{Conv}\{\mathtt{PR}, \mathtt{P}_0, \mathtt{P}_1, \mathtt{P}_L^{1011}, \mathtt{P}_L^{1110}, \mathtt{P}_{NL}^{111}\};$$
$$\text{(b)} \qquad A \subseteq \mathrm{Conv}\{\mathtt{PR}, \mathtt{P}_0, \mathtt{P}_1, \mathtt{P}_L^{1000}, \mathtt{P}_L^{1101}, \mathtt{P}_{NL}^{100}\}.$$

Now, taking the intersection, we obtain $A \subseteq C$, which yields the wanted equality $A = C$, and C is indeed a slice of $\mathcal{NS}$.

It remains to show that the slice C is included in the boundary $\partial\mathcal{NS}$, so that it is indeed a face. Assume that there is a point $\mathtt{P}$ in C of the form $\mathtt{P} = q\mathtt{Q}_1 + (1-q)\mathtt{Q}_2$ with $q > 0$ and $\mathtt{Q}_1, \mathtt{Q}_2 \in \mathcal{NS}$. Applying the convexity preservation property of $\mathtt{CHSH}(\cdot)$, $\mathtt{CHSH}'(\cdot)$, $\mathtt{CHSH}''(\cdot)$ in Figure 11, we obtain the following two necessary conditions:

$$\text{(a)} \qquad \mathtt{Q}_1, \mathtt{Q}_2 \in \mathrm{Conv}\{\mathtt{PR}, \mathtt{P}_0, \mathtt{P}_1, \mathtt{P}_L^{1011}, \mathtt{P}_L^{1110}, \mathtt{P}_{NL}^{111}\};$$
$$\text{(b)} \qquad \mathtt{Q}_1, \mathtt{Q}_2 \in \mathrm{Conv}\{\mathtt{PR}, \mathtt{P}_0, \mathtt{P}_1, \mathtt{P}_L^{1000}, \mathtt{P}_L^{1101}, \mathtt{P}_{NL}^{100}\}.$$

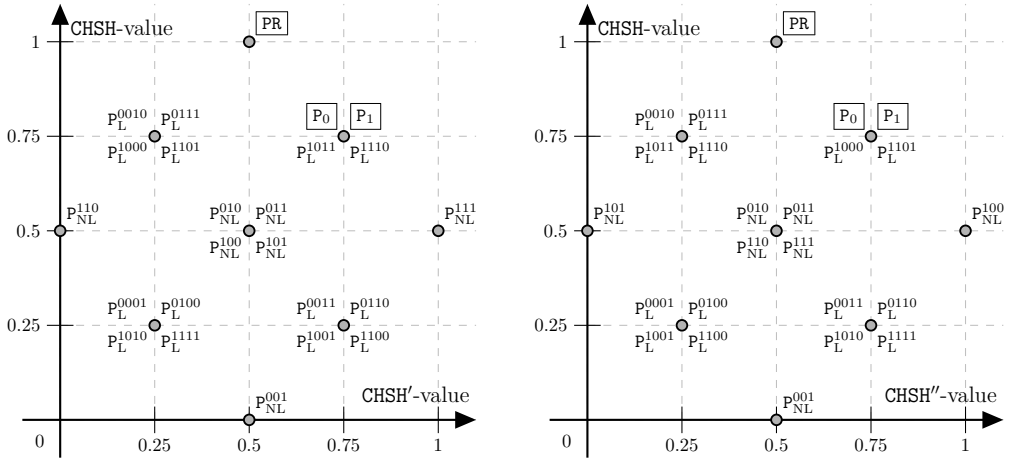


Figure 11: Computation of the CHSH-, CHSH′- and CHSH″- values of the 24 extremal points of $\mathcal{NS}$ [3], where by definition $\mathtt{CHSH}(\mathtt{P}) := \frac{1}{4}\sum_{a\oplus b=xy} \mathtt{P}(a,b\,|\,x,y)$, and CHSH′ and CHSH″ are defined similarly but with respective summand conditions $a\oplus b = (x\oplus 1)\cdot(y\oplus 1)$ and $a\oplus b = x\cdot(y\oplus 1)$.

Then, taking the intersection, we get $\mathtt{Q}_1, \mathtt{Q}_2 \in C$ and therefore $C \subseteq \partial\mathcal{NS}$ as wanted. □

Theorem 23. *The face C of $\mathcal{NS}$ is collapsing, except in the segment* $\mathrm{Conv}\{\mathtt{P}_0, \mathtt{P}_1\}$.

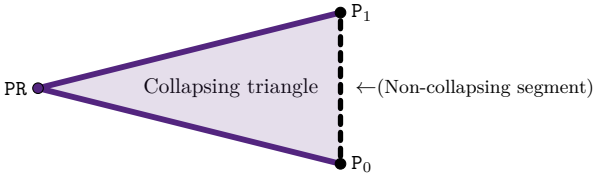


Figure 12: Illustration of Theorem 23.

Proof. Denote $\mathtt{T} := \mathrm{Conv}\{\mathtt{PR}, \mathtt{P}_0, \mathtt{P}_1\} \setminus \mathrm{Conv}\{\mathtt{P}_0, \mathtt{P}_1\}$, and consider a convex combination of the form $\mathtt{P}_{\alpha,\beta} := \alpha\mathtt{PR} + \beta\mathtt{P}_0 + (1-\alpha-\beta)\mathtt{P}_1 \in \mathtt{T}$ with $\alpha, \beta \geq 0$ and $\alpha \neq 0$ and $\alpha+\beta \leq 1$. Similarly fix other convex coefficients $(\alpha_0, \beta_0) =: u_0$ with the same conditions. We want to build a sequence $(u_k)_k = \big((\alpha_k, \beta_k)\big)_k$ such that $(\mathtt{P}_{u_k})_k$ tends to the PR box. Denote $\boxtimes$ the box product induced by the wiring W_{BS} (see definition in Subsection 1.4). By bilinearity of $\boxtimes$ and using the multiplication table in Figure 4, computations lead to $\mathtt{P}_{\alpha,\beta} \boxtimes \mathtt{P}_{\alpha_0,\beta_0} = \mathtt{P}_{\tilde{\alpha},\tilde{\beta}}$ where:

$$\begin{bmatrix}\tilde{\alpha}\\ \tilde{\beta}\end{bmatrix} = A\begin{bmatrix}\alpha\\ \beta\end{bmatrix} + b\,, \qquad A := \begin{bmatrix}1-\alpha_0 & -\alpha_0\\ -1+\alpha_0+\beta_0 & -1+\frac{3}{2}\alpha_0+2\beta_0\end{bmatrix}, \qquad b := \begin{bmatrix}\alpha_0\\ 1-\alpha_0-\beta_0\end{bmatrix}.$$

From this remark, we define the following sequence:

$$u_0 := (\alpha_0, \beta_0)\,, \qquad u_{k+1} = A\,u_k + b\,.$$

We easily identify that $\ell := (1,0)$ is a fixed point of $x \mapsto A\,x + b$, so it yields:

$$u_{k+1} - \ell = \big(A\,u_k + b\big) - \big(A\ell + b\big) = A\,(u_k - \ell) = A^{k+1}\,(u_0 - \ell)\,,$$

where the last equality follows from an induction on k. But the matrix A admits exactly two distinct[7] eigenvalues $\lambda_1 = 1 - a/2$ and $\lambda_2 = -1 + a + 2b$. So A is diagonalizable, and its power

[7]The eigenvalues λ_1 and λ_2 are distinct because otherwise we would have $2 = \frac{3}{2}(a+b) + \frac{b}{2}$, which is achieved only if both $a+b=1$ and $b=1$, which is equivalent to $a=0$ and $b=1$ and which contradicts the assumption $a \neq 0$.

$A^k = P\begin{bmatrix}\lambda_1^k & 0\\ 0 & \lambda_2^k\end{bmatrix}P^{-1}$ tends to the null matrix because $|\lambda_1|, |\lambda_2| < 1$, where P is an invertible matrix. Hence, from the above equation, the sequence $(u_k)_k$ tends to ℓ, and by continuity we have that the sequence of boxes $(\mathtt{P}_{u_k})_k \subseteq \mathbb{R}^{16}$ converges to $\mathtt{P}_\ell = \mathtt{PR}$. But Brassard *et al.* showed that there is an open neighbor around PR that collapses communication complexity [10]. Therefore, we know that the sequence $(\mathtt{P}_{u_k})_k$ reaches this collapsing neighbor for some k large enough, and using Proposition 8 we conclude that any starting box $\mathtt{P}_{u_0} \in \mathtt{T}$ is indeed collapsing. □

Remark 24 (Why is $\mathrm{Conv}\{\mathtt{P}_0, \mathtt{P}_1\}$ non-collapsing?). It is not surprising that the boundary segment $\mathrm{Conv}\{\mathtt{P}_0, \mathtt{P}_1\}$ of T is not in the collapsing area because this segment is included in the local set $\mathcal{L}$, which is itself included in the quantum set $\mathcal{Q}$, for which it is known that communication complexity does not collapse [17].

Remark 25 (Left multiplication does not give the same result). In the proof, we defined our sequence of boxes $(\mathtt{P}_{u_k})_k$ based on right multiplication. One could instead try to do the left multiplication: $\mathtt{P}_{u_{k+1}} = \mathtt{P}_{u_0} \boxtimes \mathtt{P}_{u_k}$. In that case, similar computations lead to:

$$u_{k+1} = A'u_k + b', \qquad A' := \begin{bmatrix} 1-\alpha_0-\beta_0 & 0 \\ -1+\alpha_0+\frac{3}{2}\beta_0 & -1+\alpha_0+2\beta_0 \end{bmatrix}, \qquad b' := \begin{bmatrix} \alpha_0 \\ 1-\alpha_0-\beta_0 \end{bmatrix}.$$

The map $x \mapsto A'x + b'$ admits a unique fixed point $\ell' := \left(\frac{\alpha_0}{\alpha_0+\beta_0}, \frac{\beta_0}{2(\alpha_0+\beta_0)}\right)$ (no division by 0 since $\alpha_0 > 0$). The matrix A' is already in the triangular form, its eigenvalues are $\lambda'_1 := 1-\alpha_0-\beta_0$ and $\lambda'_2 := -1+\alpha_0+2\beta_0$, and again $|\lambda'_1|, |\lambda'_2| < 1$ so $\mathtt{P}_{\alpha_k,\beta_k} \to \mathtt{P}_{\ell'}$. But in that case $\mathtt{P}_{\ell'}$ is not the PR box, so we cannot apply Ref. [10] to build a collapsing protocol from any starting box.

Remark 26 (Pairwise multiplication gives the same result). It is also possible to try the pairwise multiplication: $\mathtt{P}_{u_{k+1}} = \mathtt{P}_{u_k} \boxtimes \mathtt{P}_{u_k}$, which is the way the authors of [13] originally proved that the segment $\mathrm{Conv}\{\mathtt{PR}, \mathtt{SR}\}\backslash\{\mathtt{SR}\}$ is collapsing. But this pairwise multiplication does not behave as well as with the right multiplication, iterations $u_k = (\alpha_k, \beta_k)$ are non-affine here:

$$u_{k+1} = \begin{bmatrix} -1 & -1 & 0 \\ 1 & 5/2 & 2 \end{bmatrix} \begin{bmatrix} \alpha_k^2 \\ \alpha_k\beta_k \\ \beta_k^2 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ -2 & -2 \end{bmatrix} u_k + \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Nevertheless, the result still holds using this pairwise multiplication, but the proof we found is much more technical, see Appendix B.

The intersection of the quantum set $\mathcal{Q}$ with the boundary $\partial\mathcal{NS}$ of the non-signalling set was recently studied in [15]. Moreover, the notion of the *quantum void* was introduced and studied in [11, 42], which consist in a subset of $\partial\mathcal{NS}$ for which all quantum correlations are actually local. A direct corollary of the previous theorem allows us to single out the quantum correlations of the face $C = \mathrm{Conv}\{\mathtt{PR}, \mathtt{P}_0, \mathtt{P}_1\}$: they are exactly the ones in the segment $\mathrm{Conv}\{\mathtt{P}_0, \mathtt{P}_1\}$. Indeed, on the one hand, it is known that quantum correlations do *not* collapse communication complexity [17], and on the other hand, local correlations are particular cases of quantum correlations, so we recover the following statement from [42] with a new proof, based on communication complexity:

Corollary 27. *The face $C = \mathrm{Conv}\{\mathtt{PR}, \mathtt{P}_0, \mathtt{P}_1\} \subseteq \partial\mathcal{NS}$ is a quantum void:*

$$\mathcal{Q} \cap C = \mathrm{Conv}\{\mathtt{P}_0, \mathtt{P}_1\}.$$ □

4.2.2 Other Collapsing Triangles

The fact that the triangle $\mathtt{T} := \mathrm{Conv}\{\mathtt{PR}, \mathtt{P}_0, \mathtt{P}_1\}\backslash\,\mathrm{Conv}\{\mathtt{P}_0, \mathtt{P}_1\}$ is collapsing induces many further collapsing triangles. In this subsection, we give some examples of such collapsing triangles, thus recovering some results of [11] with a new proof, based on the algebra of boxes.

Proposition 28. *Let* $\mathtt{P}, \mathtt{Q}, \mathtt{R}$ *be three boxes. If there exists a wiring* $\mathsf{W} \in \mathcal{W}$ *that induces the multiplication table below, then the triangle* $\mathrm{Conv}\{\mathtt{P}, \mathtt{Q}, \mathtt{R}\} \setminus \mathrm{Conv}\{\mathtt{Q}, \mathtt{R}\}$ *is collapsing.*

	$\mathtt{P}$	$\mathtt{Q}$	$\mathtt{R}$
$\mathtt{P}$	$\mathtt{P}$	$\mathtt{P}$	$\mathtt{P}$
$\mathtt{Q}$	$\frac{1}{2}\big(\mathtt{Q}+\mathtt{R}\big)$	$\mathtt{Q}$	$\mathtt{R}$
$\mathtt{R}$	$\mathtt{P}$	$\mathtt{R}$	$\mathtt{Q}$

Proof. For any $\alpha, \beta \in \mathbb{R}$, consider the convex combination $\mathtt{P}_{\alpha,\beta} := \alpha\mathtt{P} + \beta\mathtt{Q} + (1-\alpha-\beta)\mathtt{R}$. We have the equality $\mathtt{P}_{\alpha,\beta} \boxtimes \mathtt{P}_{\alpha_0,\beta_0} = \mathtt{P}_{\tilde{\alpha},\tilde{\beta}}$, with the same coefficients $(\tilde{\alpha}, \tilde{\beta})$ as in the proof of Theorem 23. Then, applying the same proof gives the desired result. □

Theorem 29. *All the triangles drawn in Figure 13 are collapsing.*

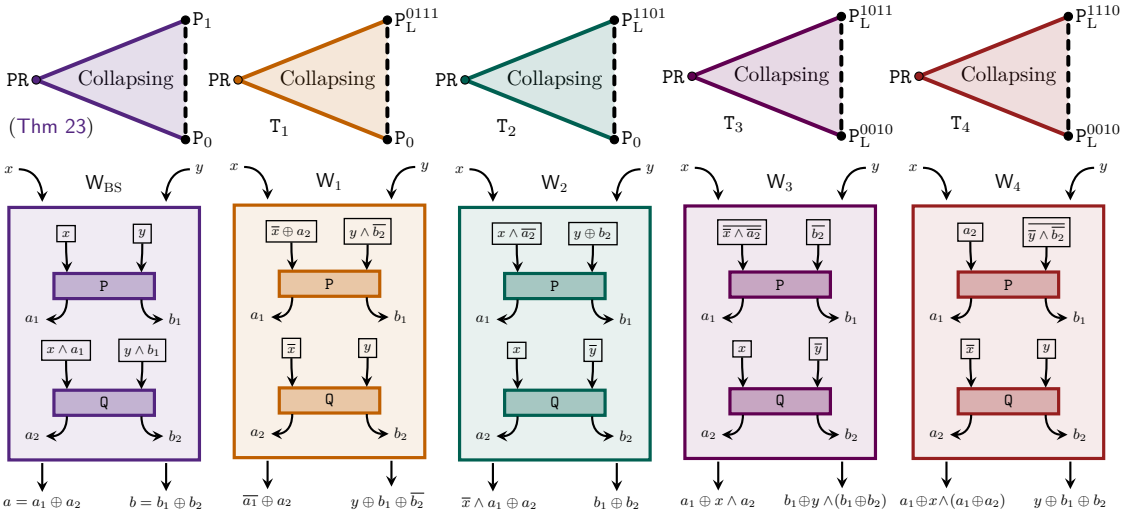


Figure 13: Examples of collapsing triangles, together with wirings that can be used in Proposition 28 to show a collapse of communication complexity. The definition of the boxes $\mathtt{P}_\mathrm{L}$ and $\mathtt{P}_\mathrm{NL}$ can be found in Equation (16).

Proof. The proof follows directly from Proposition 28 applied to what follows:

Triangle	$\mathtt{P}$	$\mathtt{Q}$	$\mathtt{R}$	Wiring $\mathsf{W} \in \mathcal{W} \subseteq \mathbb{R}^{32}$
$\mathtt{T}$	$\mathtt{PR}$	$\mathtt{P}_0$	$\mathtt{P}_1$	$\mathsf{W}_{\mathrm{BS}} = [0,0,1,1,0,0,1,1,0,0,0,1,0,0,0,1,0,1,1,0,0,1,1,0,0,1,1,0,0,1,1,0]$
$\mathtt{T}_1$	$\mathtt{PR}$	$\mathtt{P}_\mathrm{L}^{0111}$	$\mathtt{P}_0$	$\mathsf{W}_1 = [1,0,0,1,0,0,1,0,1,1,0,0,0,0,1,1,1,0,0,1,1,0,0,1,1,0,0,1,0,1,1,0]$
$\mathtt{T}_2$	$\mathtt{PR}$	$\mathtt{P}_0$	$\mathtt{P}_\mathrm{L}^{1101}$	$\mathsf{W}_2 = [0,0,1,0,0,1,1,0,0,0,1,1,1,1,0,0,0,1,1,0,0,0,1,1,0,1,1,0,0,1,1,0]$
$\mathtt{T}_3$	$\mathtt{PR}$	$\mathtt{P}_\mathrm{L}^{0010}$	$\mathtt{P}_\mathrm{L}^{1011}$	$\mathsf{W}_3 = [0,1,1,1,1,0,1,0,0,0,1,1,1,1,0,0,0,0,1,1,0,1,1,0,0,0,1,1,0,0,1,1]$
$\mathtt{T}_4$	$\mathtt{PR}$	$\mathtt{P}_\mathrm{L}^{1000}$	$\mathtt{P}_\mathrm{L}^{1110}$	$\mathsf{W}_4 = [0,1,0,1,0,1,1,1,1,1,0,0,0,0,1,1,0,0,1,1,0,0,1,1,0,1,1,0,1,0,0,1]$

where the notation $\mathsf{W} \in \mathbb{R}^{32}$ comes form Equation (15). See a representation of these wirings in Figure 13, bottom row. □

The wirings of Figure 13 are arbitrary examples of collapsing wirings that were obtained using Algorithm 5; many more wirings can be found in other triangles of $\mathcal{NS}$ using the same algorithm, which is accessible via our GitHub page [9]. Notice that these wirings are all different from the ones used in the proof of [11]. Now, an interesting problem would be to understand better the structure of the set $\mathcal{W}$ so that, given a triangle in $\mathcal{NS}$, we know how to construct a collapsing wiring W without using a search algorithm.

Conclusion

Our new algebraic perspective on nonlocal boxes allowed us to discover a surprising structure of what we called the orbit of a box, with some strong alignment and parallelism properties (see Figure 7). As a consequence, this deeper understanding of the algebraic structure of nonlocal boxes enabled us to recover in Section 4 some collapsing regions of $\mathcal{NS}$ that were recently also found in [11, 23]—for instance boxes in the triangle joining $\mathtt{PR}, \mathtt{P}_0, \mathtt{P}_1$ (except boxes in the segment $\mathtt{P}_0, \mathtt{P}_1$). The importance of our results is emphasized by considering the many known impossibility results [6, 34, 46]. Note also that according to our present intuition of Nature [6, 10, 13, 18], a direct consequence is that the collapsing boxes we presented are unlikely to exist in Nature.

We made advances towards answering the open question of determining which nonlocal boxes do indeed collapse communication complexity, but there is still a gap to be filled: for instance, as plotted in light blue in Figure 10, there are still regions of $\mathcal{NS}$ for which it is unknown whether there is a collapse of communication complexity.

Further work includes the study of the "square root" of a box: here we introduced and studied the properties of the product $\mathtt{P} \boxtimes \mathtt{P}$, but one might be interested in finding all the boxes $\mathtt{Q}$ such that $\mathtt{Q} \boxtimes_{\mathsf{W}} \mathtt{Q} = \mathtt{P}$ for some fixed wiring $\mathsf{W} \in \mathcal{W}$. If one knows that $\mathtt{P}$ is collapsing, then the wiring W yields a collapsing protocol, thus all the square roots $\mathtt{Q}$ are also collapsing.

Another interesting problem is the one mentioned at the end of Section 4. Given a triangle of boxes in $\mathcal{NS}$, we already provided in Section 3 and in our GitHub page [9] an algorithm that intends to find a wiring W that makes this triangle collapse communication complexity. Now, it would be interesting to address the question of how to construct such a collapsing wiring W without using a search algorithm, simply by knowing the target triangle of $\mathcal{NS}$. This can then help to understand better the structure of the set $\mathcal{W}$ and importantly to better understand protocols for correlation distillation.

In Subsection 4.2, we give examples of collapsing sets of dimension $d = 2$ (triangles). An open question from [11] consists in finding higher-dimensional sets of boxes that are all distillable to the $\mathtt{PR}$ box, which therefore collapses CC. For dimension $d = 3$, the authors provide explicit examples of distillable sets, but it is still unknown whether these sets are collapsing. For dimensions $d \geq 4$, they prove that distillable quantum voids are impossible due to the presence of isotropic boxes.

Another interesting problem would be to generalize and study the notion of "product of boxes" with more than two boxes. Indeed, our study is limited to wirings connecting two boxes, but it is possible to consider more general wirings, connecting n boxes, which are known to be strictly more powerful: for instance, some genuine depth-3 wirings have significance in the context of trivial communication complexity [23]. As a consequence, it may be that the multi-product of boxes $\mathsf{W}(\mathtt{P}_1, \ldots, \mathtt{P}_n)$ gives rise to similar structures of orbits as the one we found in Figure 7, which would be useful in the study of n-box distillation protocols.

In this work, we chose to study the principle of communication complexity. Although this principle alone cannot rule out the quantum set [36], a clever idea would be to combine it with other principles, such as *nonlocal computation* [32], *information causality* [28, 40], *macroscopic locality* [37], *local orthogonality* [26], *nonlocality swapping* [47], *many-box locality* [49], in working towards a comprehensive information-based description of Quantum Mechanics.

Acknowledgements

We thank Mirjam Weilenmann for the many helpful comments during the finalization process of the manuscript. We also thank Giorgos Eftaxias for comments on a preliminary version of this work. We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), [funding reference number ALLRP 580876 - 22]. This work was also supported by the MITACS grant FR113029, and by the ANR projects ESQuisses, grant number ANR-20-CE47-0014-01. P.B. was supported by the Institute for Quantum Technologies in Occitanie. A.B. acknowledges the support of the Air Force Office of Scientific Research under award number FA9550-20-1-0375. R.C. recognizes support from the ANR-3IA Artificial and Natural Intelligence Toulouse Institute. I.N. was supported by STARS, grant number ANR-20-CE40-0008, and by the PHC program *Star* (Applications of random matrix theory and abstract harmonic analysis to quantum information theory). C.P. was supported by the ANR projects Q-COAST ANR- 19-CE48-0003, "Quantum Trajectories" ANR-20-CE40-0024-01, and "Investissements d'Avenir" ANR-11-LABX-0040 of the French National Research Agency.

References

[1] Jonathan Allcock, Nicolas Brunner, Noah Linden, Sandu Popescu, Paul Skrzypczyk, and Tamás Vértesi. "Closed sets of nonlocal correlations". In: *Phys. Rev. A* 80 (6 Dec. 2009), p. 062107. DOI: 10.1103/PhysRevA.80.062107.

[2] Jonathan Allcock, Nicolas Brunner, Marcin Pawlowski, and Valerio Scarani. "Recovering part of the boundary between quantum and nonquantum correlations from information causality". In: *Phys. Rev. A* 80 (4 Oct. 2009), p. 040103. DOI: 10.1103/PhysRevA.80.040103.

[3] Jonathan Barrett, Noah Linden, Serge Massar, Stefano Pironio, Sandu Popescu, and David Roberts. "Nonlocal correlations as an information-theoretic resource". In: *Phys. Rev. A* 71 (2 Feb. 2005), p. 022101. DOI: 10.1103/PhysRevA.71.022101.

[4] Jonathan Barrett and Stefano Pironio. "Popescu-Rohrlich Correlations as a Unit of Nonlocality". In: *Phys. Rev. Lett.* 95 (14 Sept. 2005), p. 140401. DOI: 10.1103/PhysRevLett.95.140401.

[5] Mohammad Bavarian and Peter W. Shor. "Information Causality, Szemerédi-Trotter and Algebraic Variants of CHSH". In: *Proceedings of the 2015 Conference on Innovations in Theoretical Computer Science*. ITCS '15. Rehovot, Israel: Association for Computing Machinery, 2015, pp. 123–132. ISBN: 9781450333337. DOI: 10.1145/2688073.2688112.

[6] Salman Beigi and Amin Gohari. "Monotone Measures for Non-Local Correlations". In: *IEEE Transactions on Information Theory* 61.9 (2015), pp. 5185–5208. DOI: 10.1109/TIT.2015.2452253.

[7] Pierre Botteron. "Nonlocal Boxes and Communication Complexity". Under the supervision of Anne Broadbent, Ion Nechita and Clément Pellegrini. MA thesis. Université Paul Sabatier (Toulouse), June 2022. URL: https://pierre-botteron.github.io/Articles/2022-06-MSc-Thesis.pdf.

[8] Pierre Botteron, Anne Broadbent, and Marc-Olivier Proulx. "Extending the Known Region of Nonlocal Boxes that Collapse Communication Complexity". In: *Phys. Rev. Lett.* 132 (7 2024), p. 070201. DOI: 10.1103/PhysRevLett.132.070201.

[9] Pierre Botteron and Réda Chhaibi. *GitHub page*. [Accessed: November 2023]. URL: https://github.com/Pierre-Botteron/Algebra-of-Boxes-code.

[10] Gilles Brassard, Harry Buhrman, Noah Linden, André Allan Méthot, Alain Tapp, and Falk Unger. "Limit on Nonlocality in Any World in Which Communication Complexity Is Not Trivial". In: *Phys. Rev. Lett.* 96 (25 June 2006), p. 250401. DOI: 10.1103/PhysRevLett.96.250401.

[11] Samuraí Brito, George Moreno, Ashutosh Rai, and Rafael Chaves. "Nonlocality distillation and quantum voids". In: *Phys. Rev. A* 100 (1 July 2019), p. 012102. DOI: 10.1103/PhysRevA.100.012102.

[12] Nicolas Brunner, Daniel Cavalcanti, Alejo Salles, and Paul Skrzypczyk. "Bound Nonlocality and Activation". In: *Phys. Rev. Lett.* 106 (2 Jan. 2011), p. 020402. DOI: 10.1103/PhysRevLett.106.020402.

[13] Nicolas Brunner and Paul Skrzypczyk. "Nonlocality Distillation and Postquantum Theories with Trivial Communication Complexity". In: *Physical Review Letters* 102.16 (Apr. 2009). ISSN: 1079-7114. DOI: 10.1103/physrevlett.102.160403.

[14] Harry Buhrman and Serge Massar. "Causality and Tsirelson's bounds". In: *Phys. Rev. A* 72 (5 Nov. 2005), p. 052103. DOI: 10.1103/PhysRevA.72.052103.

[15] Kai-Siang Chen, Gelo Noel M. Tabia, Chellasamy Jebarathinam, Shiladitya Mal, Jun-Yi Wu, and Yeong-Cherng Liang. "Quantum correlations on the no-signaling boundary: self-testing and more". In: *Quantum* 7 (July 2023), p. 1054. ISSN: 2521-327X. DOI: 10.22331/q-2023-07-11-1054.

[16] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. "Proposed Experiment to Test Local Hidden-Variable Theories". In: *Phys. Rev. Lett.* 23 (15 Oct. 1969), pp. 880–884. DOI: 10.1103/PhysRevLett.23.880.

[17] Richard Cleve, Wim van Dam, Michael Nielsen, and Alain Tapp. *Quantum Entanglement and the Communication Complexity of the Inner Product Function*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999, pp. 61–74. ISBN: 978-3-540-49208-5. DOI: 10.1007/3-540-49208-9˙4.

[18] Wim van Dam. "Nonlocality & Communication Complexity". Ph.D. thesis. Departement of Physics: University of Oxford, 1999. URL: https://sites.cs.ucsb.edu/~vandam/oxford_thesis.pdf.

[19] Dejan D. Dukaric and Stefan Wolf. *A Limit on Non-Locality Distillation*. 2008. arXiv: 0808.3317 [quant-ph].

[20] Frédéric Dupuis, Nicolas Gisin, Avinatan Hasidim, André Allan Méthot, and Haran Pilpel. "No nonlocal box is universal". In: *Journal of Mathematical Physics* 48.8 (2007), p. 082107. DOI: 10.1063/1.2767538.

[21] Helen Ebbe and Stefan Wolf. "Multi-User Non-Locality Amplification". In: *IEEE Transactions on Information Theory* 60.2 (2014), pp. 1159–1167. DOI: 10.1109/TIT.2013.2292515.

[22] Giorgos Eftaxias. "Theory-independent topics towards quantum mechanics: ψ-ontology and nonlocality distillation". PhD thesis. University of Bristol (Quantum Engineering Centre for Doctoral Training), Mar. 2022. URL: https://research-information.bris.ac.uk/en/studentTheses/theory-independent-topics-towards-quantum-mechanics.

[23] Giorgos Eftaxias, Mirjam Weilenmann, and Roger Colbeck. "Advantages of Multicopy Nonlocality Distillation and Its Application to Minimizing Communication Complexity". In: *Phys. Rev. Lett.* 130 (10 Mar. 2023), p. 100201. DOI: 10.1103/PhysRevLett.130.100201.

[24] Giorgos Eftaxias, Mirjam Weilenmann, and Roger Colbeck. "Multisystem measurements in generalized probabilistic theories and their role in information processing". In: *Phys. Rev. A* 108 (6 2023), p. 062212. DOI: 10.1103/PhysRevA.108.062212.

[25] Manuel Forster, Severin Winkler, and Stefan Wolf. "Distilling Nonlocality". In: *Phys. Rev. Lett.* 102 (12 Mar. 2009), p. 120401. DOI: 10.1103/PhysRevLett.102.120401.

[26] Tobias Fritz, Ana Belén Sainz, Remigiusz Augusiak, Jonatan Bohr Brask, Rafael Chaves, Anthony Leverrier, and Antonio Acín. "Local orthogonality as a multipartite principle for quantum correlations". In: *Nature Communications* 4.1 (Aug. 2013), p. 2263. DOI: 10.1038/ncomms3263.

[27] Peter Hoyer and Jibran Rashid. "Optimal protocols for nonlocality distillation". In: *Phys. Rev. A* 82 (4 Oct. 2010), p. 042118. DOI: 10.1103/PhysRevA.82.042118.

[28] Prabhav Jain, Mariami Gachechiladze, and Nikolai Miklin. *Information causality as a tool for bounding the set of quantum correlations*. 2023. arXiv: 2308.02478 [quant-ph].

[29] Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen. *MIP*=RE*. 2022. arXiv: 2001.04383 [quant-ph].

[30] Eyal Kushilevitz and Noam Nisan. *Communication Complexity*. Cambridge University Press, 1996. DOI: 10.1017/CBO9780511574948.

[31] Ben Lang, Tamás Vértesi, and Miguel Navascués. "Closed sets of correlations: answers from the zoo". In: *Journal of Physics A: Mathematical and Theoretical* 47.42 (Oct. 2014), p. 424029. DOI: 10.1088/1751-8113/47/42/424029.

[32] Noah Linden, Sandu Popescu, Anthony J. Short, and Andreas Winter. "Quantum Nonlocality and Beyond: Limits from Nonlocal Computation". In: *Phys. Rev. Lett.* 99 (18 Oct. 2007), p. 180502. DOI: 10.1103/PhysRevLett.99.180502.

[33] Lluis Masanes. *Necessary and sufficient condition for quantum-generated correlations*. 2003. arXiv: quant-ph/0309137 [quant-ph].

[34] Ryuhei Mori. "Three-input majority function as the unique optimal function for the bias amplification using nonlocal boxes". In: *Phys. Rev. A* 94 (5 Nov. 2016), p. 052130. DOI: 10.1103/PhysRevA.94.052130.

[35] Sahil Gopalkrishna Naik, Govind Lal Sidhardh, Samrat Sen, Arup Roy, Ashutosh Rai, and Manik Banik. "Distilling Nonlocality in Quantum Correlations". In: *Phys. Rev. Lett.* 130 (22 June 2023), p. 220201. DOI: 10.1103/PhysRevLett.130.220201.

[36] Miguel Navascués, Yelena Guryanova, Matty J. Hoban, and Antonio Acín. "Almost quantum correlations". In: *Nature Communications* 6.1 (May 2015), p. 6288. DOI: 10.1038/ncomms7288.

[37] Miguel Navascués and Harald Wunderlich. "A glance beyond the quantum model". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 466.2115 (2009), pp. 881–890. DOI: 10.1098/rspa.2009.0453.

[38] Jorge Nocedal and Stephen J. Wright, eds. *Numerical Optimization*. en. Springer Series in Operations Research and Financial Engineering. New York: Springer-Verlag, 1999. ISBN: 9780387987934. DOI: 10.1007/b98874.

[39] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. "Automatic differentiation in PyTorch". In: *OpenReview* (2017). URL: https://openreview.net/forum?id=BJJsrmfCZ.

[40] Marcin Pawłowski, Tomasz Paterek, Dagomir Kaszlikowski, Valerio Scarani, Andreas Winter, and Marek Żukowski. "Information causality as a physical principle". In: *Nature* 461.7267 (Oct. 2009), pp. 1101–1104. DOI: 10.1038/nature08400.

[41] Sandu Popescu and Daniel Rohrlich. "Quantum nonlocality as an axiom". In: *Foundations of Physics* 24.3 (Mar. 1994), pp. 379–385. DOI: 10.1007/BF02058098.

[42] Ashutosh Rai, Cristhiano Duarte, Samuraí Brito, and Rafael Chaves. "Geometry of the quantum set on no-signaling faces". In: *Phys. Rev. A* 99 (3 Mar. 2019), p. 032106. DOI: 10.1103/PhysRevA.99.032106.

[43] Anup Rao and Amir Yehudayoff. *Communication Complexity: and Applications*. Cambridge University Press, 2020. DOI: 10.1017/9781108671644.

[44] Denis Rosset, Amin Baumeler, Jean-Daniel Bancal, Nicolas Gisin, Anthony Martin, Marc-Olivier Renou, and Elie Wolfe. *Algebraic and geometric properties of local transformations*. 2020. arXiv: 2004.09405 [quant-ph].

[45] Anthony J. Short, Sandu Popescu, and Nicolas Gisin. "Entanglement swapping for generalized nonlocal correlations". In: *Phys. Rev. A* 73 (1 Jan. 2006), p. 012101. DOI: 10.1103/PhysRevA.73.012101.

[46] Noah Shutty, Mary Wootters, and Patrick Hayden. "Tight Limits on Nonlocality from Nontrivial Communication Complexity; a.k.a. Reliable Computation with Asymmetric Gate Noise". In: *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*. 2020, pp. 206–217. DOI: 10.1109/FOCS46700.2020.00028.

[47] Paul Skrzypczyk, Nicolas Brunner, and Sandu Popescu. "Emergence of Quantum Correlations from Nonlocality Swapping". In: *Phys. Rev. Lett.* 102 (11 Mar. 2009), p. 110402. DOI: 10.1103/PhysRevLett.102.110402.

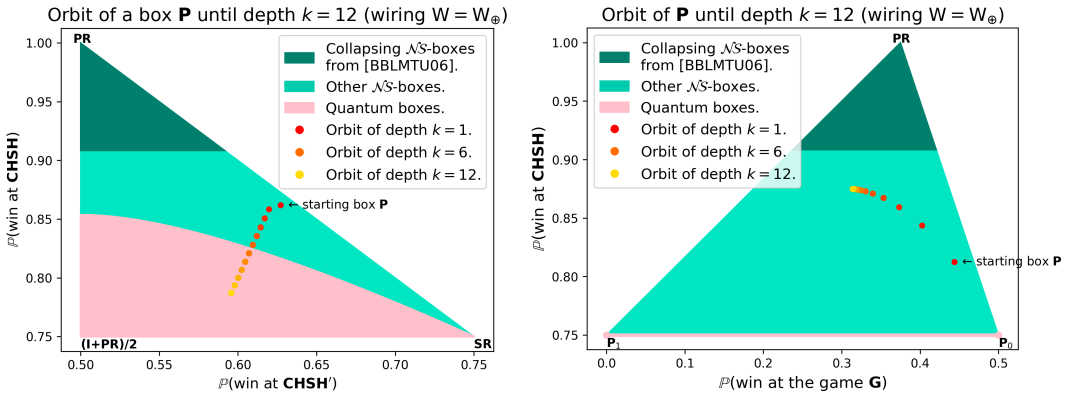
[48] Andrew Chi-Chih Yao. "Some Complexity Questions Related to Distributive Computing (Preliminary Report)". In: *Proceedings of the Eleventh Annual ACM Symposium on Theory of Computing*. STOC '79. Atlanta, Georgia, USA: Association for Computing Machinery, 1979, pp. 209–213. ISBN: 9781450374385. DOI: 10.1145/800135.804414.

[49] Yuqian Zhou, Yu Cai, Jean-Daniel Bancal, Fei Gao, and Valerio Scarani. "Many-box locality". In: *Phys. Rev. A* 96 (5 Nov. 2017), p. 052108. DOI: 10.1103/PhysRevA.96.052108.

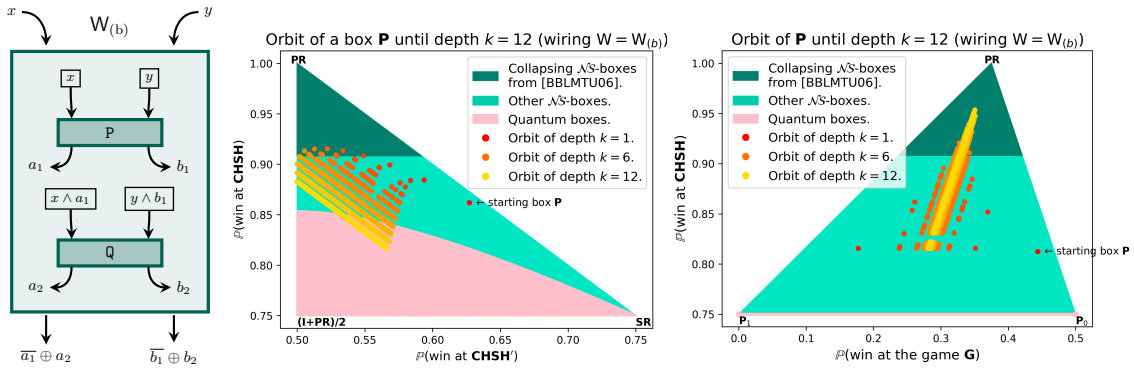
A Drawing of Some Orbits

We present below some examples of orbits to illustrate Subsection 2.4, using different wirings, each time in two different slices of $\mathcal{NS}$. Each orbit is drawn with depth going until $k = 12$. The game $\mathtt{G}$ is defined by the winning rule $a = 0$ and $b = y$.

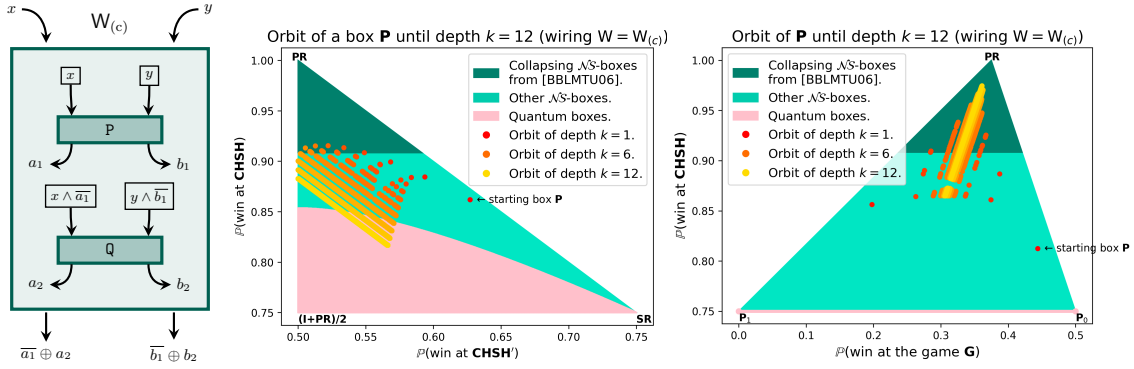
(a) For $\mathsf{W}_{\oplus}$ (see definition in Subsection 1.4):



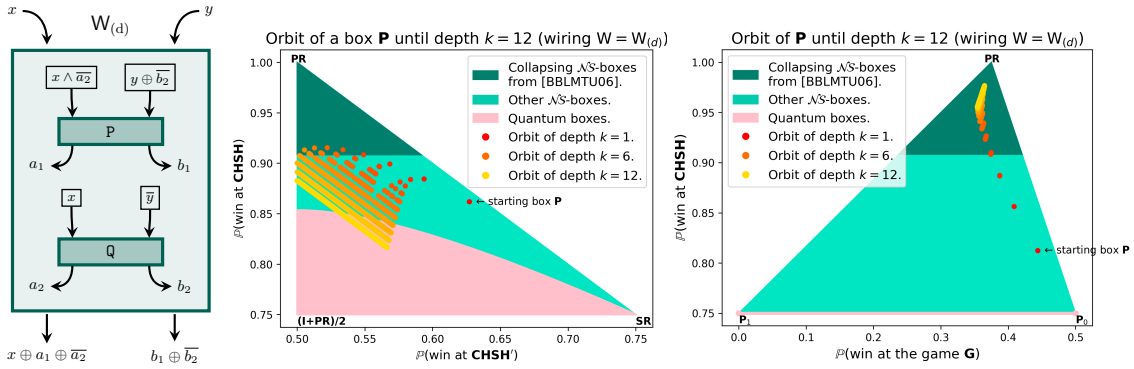
(b) For $\mathsf{W}_{(\mathrm{b})} := [0., 0., 1., 1., 0., 0., 1., 1., 0., 0., 0., 1., 0., 0., 0., 1., 1., 0., 0., 1., 1., 0., 0., 1., 1., 0., 0., 1., 1., 0., 0., 1.]$:



(c) For $\mathsf{W}_{(\mathrm{c})} := [0., 0., 1., 1., 0., 0., 1., 1., 0., 0., 1., 0., 0., 0., 1., 0., 1., 0., 0., 1., 1., 0., 0., 1., 1., 0., 0., 1., 1., 0., 0., 1.]$:



(d) For $\mathsf{W}_{(\mathrm{d})} := [0., 0., 1., 1., 0., 0., 1., 1., 0., 0., 1., 0., 0., 0., 1., 0., 0., 1., 1., 0., 0., 1., 1., 0., 0., 1., 1., 0., 0., 1., 1., 0.]$:



B An Alternative Proof of Theorem 23, using Pairwise Multiplication

Lemma 30 (Convergence sufficient conditions). *Let $\mathcal{K} \subseteq \mathbb{R}^2$ be a compact convex set, and let F, f_β, g_α be functions such that:*

$$F : \left\{ \begin{array}{ccc} \mathcal{K} & \longrightarrow & \mathcal{K} \\ (\alpha,\beta) & \longmapsto & \big(f_\beta(\alpha), g_\alpha(\beta)\big)\,. \end{array} \right.$$

where $f_\beta : I_\beta \to I_\beta$ and $g_\alpha : J_\alpha \to J_\alpha$, where $I_\beta := \mathcal{K} \cap \big(\mathbb{R} \times \{\beta\}\big)$ and $J_\alpha := \mathcal{K} \cap \big(\{\alpha\} \times \mathbb{R}\big)$ are viewed as intervals of $\mathbb{R}$. Suppose that:

(i) *The function F is continuous;*

(ii) *The f_β's are strictly increasing and uniformly bounded in the "left part" of the interior of $\mathcal{K}$:*

$$\exists \mu_*,\ \forall (\alpha,\beta) \in \mathrm{Int}(\mathcal{K}),\ \alpha < \mu_* \Longrightarrow \alpha < f_\beta(\alpha) < \mu_*\,; \tag{17}$$

(iii) *For all boundary element $(\alpha,\beta) \in \partial\mathcal{K}$ such that $\alpha < \mu_*$, either the resulting inequalities of (17) are satisfied, or $\alpha = f_\beta(\alpha)$ and the image $F(\alpha,\beta)$ lies in the interior $\mathrm{Int}(\mathcal{K})$ of $\mathcal{K}$.*

Then the sequence defined by $(\alpha_0,\beta_0) \in \mathcal{K}$, $\alpha_0 < \mu_*$, and $(\alpha_{k+1},\beta_{k+1}) = F(\alpha_k,\beta_k)$ $(k \geq 0)$, satisfies:

$$\alpha_k \quad \xrightarrow[k\to+\infty]{} \quad \mu_*\,.$$

Proof. A proof by induction based on (ii) and (iii) shows that the sequence $(\alpha_k)_k$ is increasing and bounded from above by μ_*, so it converges to some $\alpha_\infty \leq \mu_*$. Let us show that $\alpha_\infty = \mu_*$. As $\mathcal{K}$ is compact, Bolzano–Weierstrass' Theorem tells us that the sequence $(\beta_k)_k$ admits a converging subsequence $\big(\beta_{\varphi(k)}\big)_k$, whose limit is β_∞. Then, by continuity of F, we have:

$$\big(\alpha_{\varphi(k)+1}, \beta_{\varphi(k)+1}\big) = F\big(\alpha_{\varphi(k)}, \beta_{\varphi(k)}\big) \quad \xrightarrow[k\to+\infty]{} \quad F\big(\alpha_\infty,\beta_\infty\big) = \big(f_{\beta_\infty}(\alpha_\infty), g_{\alpha_\infty}(\beta_\infty)\big)\,.$$

Therefore the sequence $\big(\alpha_{\varphi(k)+1}\big)_k$ tends to $f_{\beta_\infty}(\alpha_\infty)$, but we know from before that it also tends to α_∞, so the limits coincide:

$$f_{\beta_\infty}(\alpha_\infty) = \alpha_\infty\,.$$

In other words, we showed that α_∞ is a fixed point of f_{β_∞}. Now, by contradiction if α_∞ were different from μ_*, *i.e.* if $\alpha_\infty < \mu_*$, then (ii) and (iii) would imply either the contradiction $\alpha_\infty < f_{\beta_\infty}(\alpha_\infty)$, or the inlcusion $\big(f_{\beta_\infty}(\alpha_\infty), g_{\alpha_\infty}(\beta_\infty)\big) \in \mathrm{Int}(\mathcal{K})$, but then (ii) would give $f_{\beta_\infty}(\alpha_\infty) < f_{\beta_\infty} \circ f_{\beta_\infty}(\alpha_\infty)$, which is again the contradiction $\alpha_\infty < f_{\beta_\infty}(\alpha_\infty)$. Hence $\alpha_\infty = \mu_*$ and we obtain the wanted result. □

Theorem 31 (Alternative proof of Theorem 23). *The set $T := \mathrm{Conv}\{\mathtt{PR}, \mathtt{P}_0, \mathtt{P}_1\} \setminus \mathrm{Conv}\{\mathtt{P}_0, \mathtt{P}_1\}$ is collapsing.*

Proof. Consider boxes of the form $\mathtt{P}_{\alpha,\beta} := \alpha\mathtt{PR} + \beta\mathtt{P}_0 + (1-\alpha-\beta)\mathtt{P}_1$, where $\alpha,\beta \geq 0$ and $\alpha \neq 0$ and $\alpha+\beta \leq 1$. Denote Δ the set of all such α's and β's, so that $T = \{\mathtt{P}_{\alpha,\beta} : (\alpha,\beta) \in \Delta\}$. The triangle Δ is not compact but it can be written as a union of compact convex sets:

$$\Delta = \bigcup_{n\geq 1} \mathcal{K}_n\,,$$

where $\mathcal{K}_n := \Delta \cap \{\alpha \geq 1/n\}$. Denote $\boxtimes$ the box product induced by the wiring W_{BS} inspired from [13]. Fix $(\alpha_0,\beta_0) \in \Delta$, and fix an integer $n \geq 1$ large enough so that $(\alpha_0,\beta_0) \in \mathcal{K}_n$. We want to build a protocol

that collapses communication complexity using the starting box $\mathtt{P}_{\alpha_0,\beta_0}$ as a resource. To do so, we will define a sequence of boxes $(\mathtt{P}_{\alpha_k,\beta_k})_k$ that are eventually collapsing for k large enough. First, see that the closure $\overline{T}$ of the triangle T is stable under $\boxtimes$, because $\overline{T}$ equals $\mathcal{NS} \cap \mathrm{Aff}\{\mathtt{PR}, \mathtt{P}_0, \mathtt{P}_1\}$ (see Proposition 22) and the two intersected sets are stable under $\boxtimes$ [8]. Moreover, by bilinearity of $\boxtimes$ and using the multiplication table in Figure 4, we obtain:

$$\mathtt{P}_{\alpha,\beta} \boxtimes \mathtt{P}_{\alpha,\beta} = \mathtt{P}_{\tilde{\alpha},\tilde{\beta}}\,, \tag{18}$$

with $\tilde{\alpha} := 2\alpha - \alpha^2 - \alpha\beta$ and $\tilde{\beta} := 1 - 2\alpha + \alpha^2 + \frac{5}{2}\alpha\beta - 2\beta + 2\beta^2$. We then consider the function $F : \mathcal{K}_n \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ such that $F(\alpha,\beta) := (\tilde{\alpha},\tilde{\beta})$ and we define the sequence $(\alpha_{k+1},\beta_{k+1}) := F(\alpha_k,\beta_k)$ for $k \geq 0$. Let us show that the conditions of Lemma 30 are satisfied. Due to the stability of $\overline{T}$ under $\boxtimes$, we have the inclusion $F(\mathcal{K}_n) \subseteq \overline{\Delta}$, and as

$$\tilde{\alpha} \;=\; 2\alpha - \;\alpha(\alpha+\beta) \geq \alpha \geq 1/n\,, \tag{19}$$

we obtain the inclusion $F(\mathcal{K}_n) \subseteq \mathcal{K}_n$, where we used $\alpha + \beta \leq 1$. (i) The function F is polynomial so continuous. (ii) Condition (17) is satisfied for $\mu_* := 1$: on the one hand f_β is strictly increasing due to (19) applied with $\alpha + \beta < 1$ in the interior $\mathrm{Int}(\mathcal{K}_n)$; on the other hand, by strict growth of f_β we have $f_\beta(\alpha) = \tilde{\alpha} < f_{\tilde{\beta}}\big(\tilde{\alpha}\big)$, which is ≤ 1 because $F(\mathcal{K}_n) \subseteq \mathcal{K}_n$. Hence we indeed have $\alpha < f_\beta(\alpha) < 1$ in the interior of $\mathcal{K}_n$. (iii) Let $(\alpha,\beta) \in \partial\mathcal{K}_n = \{\beta = 0\} \cup \{\alpha = 1/n\} \cup \{\alpha + \beta = 1\}$ such that $\alpha < 1$. If $\beta = 0$ or $\alpha = 1/n$, then inequalities of Condition (17) hold. Otherwise we have $\alpha + \beta = 1$ and in that case:

$$\begin{cases} \tilde{\alpha} &= \alpha \in \left(\frac{1}{n}, 1\right), \\ \tilde{\beta} &= \frac{1}{2}(1-\alpha)(1+\beta) > 0\,, \\ \tilde{\alpha} + \tilde{\beta} &= 1 - \frac{1}{2}(1-\alpha)(1-\beta) < 1\,, \end{cases}$$

which implies that $(\tilde{\alpha},\tilde{\beta}) \in \mathrm{Int}(\mathcal{K}_n)$. Finally, we may assume that $\alpha_0 < 1$ (otherwise $\alpha_0 = 1$ and then $\mathtt{P}_{\alpha_0,\beta_0} = \mathtt{PR}$, which is collapsing) and all the conditions of Lemma 30 are satisfied. The lemma tells that the sequence $(\alpha_k)_k$ tends to $\mu_* = 1$, so $(\mathtt{P}_{\alpha_k,\beta_k})_k \subseteq \mathbb{R}^{16}$ tends to $\mathtt{PR}$. But we know from [10] that there is a non-empty open set around $\mathtt{PR}$ that is collapsing. So there exists a finite k for which the box $\mathtt{P}_{\alpha_k,\beta_k}$ is collapsing and we obtain the wanted collapsing protocol. □

C Some Multiplication Tables

Find below the multiplication tables of $\boxtimes_{\mathsf{W}}$ for different wirings W. Each cell displays the result of $\mathtt{P} \boxtimes_{\mathsf{W}} \mathtt{Q}$, where $\mathtt{P}$ lies in the first column and $\mathtt{Q}$ lies in the first line. $\mathtt{P}_{10} := \mathtt{P}_{\mathrm{L}}^{0100}$. $\mathtt{P}_{01} := \mathtt{P}_{\mathrm{L}}^{0001}$. $\mathtt{Q}_1 := -\frac{1}{8}\big(\mathtt{P}_0 + \mathtt{P}_1\big) + \frac{1}{4}\,\mathtt{PR} + \mathtt{I}$. $\mathtt{Q}_2 := \frac{1}{4}\big(\mathtt{PR} + \mathtt{P}_{\mathrm{NL}}^{011}\big) + \frac{1}{2}\,\mathtt{I}$. $\mathtt{Q}_3 := \frac{1}{4}\,\mathtt{PR} + \frac{1}{2}\,\mathtt{I} + \frac{3}{8}\mathtt{P}_0 - \frac{1}{8}\mathtt{P}_1$. $\mathtt{Q}_4 := \frac{3}{8}\,\mathtt{P}_0 - \frac{1}{8}\,\mathtt{P}_1 + \frac{3}{4}\,\mathtt{I}$. $\mathtt{Q}_5 := \frac{1}{4}\,\mathtt{PR} + \frac{1}{8}(\mathtt{P}_0 + \mathtt{P}_1) + \frac{1}{2}\,\mathtt{P}_{10}$. $\mathtt{Q}_6 := \frac{3}{16}(\mathtt{P}_0 + \mathtt{P}_1) + \frac{1}{16}\,\mathtt{P}_{01} + \frac{9}{16}\,\mathtt{P}_{10}$. Find an algorithm to compute a multiplication table in our GitHub page [9]. Recall the definitions of $\mathtt{P}_{\mathrm{L}}$ and $\mathtt{P}_{\mathrm{NL}}$ in Equation (16).

$\mathsf{W}_{\mathrm{triv}}$	PR	$\mathtt{P}_0$	$\mathtt{P}_1$	I
PR	$\mathtt{P}_{\mathrm{L}}^{1010}$	$\mathtt{P}_{\mathrm{L}}^{1010}$	$\mathtt{P}_{\mathrm{L}}^{1010}$	$\mathtt{P}_{\mathrm{L}}^{1010}$
$\mathtt{P}_0$	$\mathtt{P}_{\mathrm{L}}^{1010}$	$\mathtt{P}_{\mathrm{L}}^{1010}$	$\mathtt{P}_{\mathrm{L}}^{1010}$	$\mathtt{P}_{\mathrm{L}}^{1010}$
$\mathtt{P}_1$	$\mathtt{P}_{\mathrm{L}}^{1010}$	$\mathtt{P}_{\mathrm{L}}^{1010}$	$\mathtt{P}_{\mathrm{L}}^{1010}$	$\mathtt{P}_{\mathrm{L}}^{1010}$
I	$\mathtt{P}_{\mathrm{L}}^{1010}$	$\mathtt{P}_{\mathrm{L}}^{1010}$	$\mathtt{P}_{\mathrm{L}}^{1010}$	$\mathtt{P}_{\mathrm{L}}^{1010}$

$\mathsf{W}_{\oplus}$ [25]	PR	$\mathtt{P}_0$	$\mathtt{P}_1$	I
PR	$\frac{\mathtt{P}_0+\mathtt{P}_1}{2}$	PR	PR	I
$\mathtt{P}_0$	PR	$\mathtt{P}_0$	$\mathtt{P}_1$	I
$\mathtt{P}_1$	PR	$\mathtt{P}_1$	$\mathtt{P}_0$	I
I	I	I	I	I

W_{BS} [13]	PR	$\mathtt{P}_0$	$\mathtt{P}_1$	I
PR	PR	PR	PR	I
$\mathtt{P}_0$	$\frac{\mathtt{P}_0+\mathtt{P}_1}{2}$	$\mathtt{P}_0$	$\mathtt{P}_1$	I
$\mathtt{P}_1$	PR	$\mathtt{P}_1$	$\mathtt{P}_0$	I
I	$\mathtt{Q}_1$	I	I	I

$\mathsf{W}_{\mathrm{dist}}$ [1]	PR	$\mathtt{P}_0$	$\mathtt{P}_1$	I
PR	PR	PR	PR	I
$\mathtt{P}_0$	$\frac{\mathtt{P}_0+\mathtt{P}_1}{2}$	$\mathtt{P}_1$	$\mathtt{P}_0$	I
$\mathtt{P}_1$	PR	$\mathtt{P}_0$	$\mathtt{P}_1$	I
I	$\mathtt{Q}_2$	I	I	I

$\mathsf{W}_{\wedge}$ [1]	PR	$\mathtt{P}_0$	$\mathtt{P}_1$	I
PR	$\frac{\mathtt{P}_0+\mathtt{P}_1}{2}$	$\mathtt{P}_0$	PR	$\mathtt{Q}_3$
$\mathtt{P}_0$	$\mathtt{P}_0$	$\mathtt{P}_0$	$\mathtt{P}_0$	$\mathtt{P}_0$
$\mathtt{P}_1$	PR	$\mathtt{P}_0$	$\mathtt{P}_1$	I
I	$\mathtt{Q}_3$	$\mathtt{P}_0$	I	$\mathtt{Q}_4$

$\mathsf{W}_{\vee\wedge}$ [35]	PR	$\mathtt{P}_0$	$\mathtt{P}_1$	I
PR	$\frac{\mathtt{PR}+\mathtt{P}_{10}}{2}$	$\frac{\mathtt{P}_0+\mathtt{P}_{10}}{2}$	$\frac{\mathtt{P}_1+\mathtt{P}_{10}}{2}$	$\mathtt{Q}_5$
$\mathtt{P}_0$	$\frac{\mathtt{P}_0+\mathtt{P}_{10}}{2}$	$\mathtt{P}_0$	$\mathtt{P}_{10}$	$\frac{\mathtt{P}_0+\mathtt{P}_{10}}{2}$
$\mathtt{P}_1$	$\frac{\mathtt{P}_1+\mathtt{P}_{10}}{2}$	$\mathtt{P}_{10}$	$\mathtt{P}_1$	$\frac{\mathtt{P}_1+\mathtt{P}_{10}}{2}$
I	$\mathtt{Q}_5$	$\frac{\mathtt{P}_0+\mathtt{P}_{10}}{2}$	$\frac{\mathtt{P}_1+\mathtt{P}_{10}}{2}$	$\mathtt{Q}_6$

[8] Indeed, (i) the set $\mathcal{NS}$ is stable under $\boxtimes$ because it is closed under wirings [1], (ii) the affine plane $\mathrm{Aff}\{\mathtt{PR}, \mathtt{P}_0, \mathtt{P}_1\}$ is stable under under $\boxtimes$ as a consequence of the multiplication table in Figure 4 and by bilinearity of $\boxtimes$.