Non-Locality Without Counterfactual Reasoning

Stefan Wolf[1]

[1] *Faculty of Informatics, Università della Svizzera italiana, Via G. Buffi 13, 6900 Lugano, Switzerland*

Non-local correlations are usually understood through the outcomes of alternative measurements (on two or more parts of a system) that cannot altogether actually be carried out in an experiment. Indeed, a joint input/output — *e.g.*, measurement-setting/outcome — behavior is non-local if and only if the outputs for *all* possible inputs cannot coexist consistently. It has been argued that this counterfactual view is how Bell's inequalities and their violations are to be seen. We propose an alternative perspective which refrains from setting into relation the results of mutually exclusive measurements, but that is based solely on data actually available. Our approach uses *algorithmic complexity* instead of probability, implies non-locality to have similar consequences as in the probabilistic view, and is conceptually simpler yet at the same time more general than the latter.

Introduction.— In an eponymous text, *Asher Peres* (1978) [12] states that "unperformed experiments have no results." He argues first that it is not only illegitimate to speculate about unperformed experiments, but also that refraining from it frees physics of epistemological difficulties such as the EPR paradox [4]. Indeed, already *Ernst Specker* (1961) [17] had pointed out the counterfactual nature of the reasoning against the possibility of embedding quantum theory into propositional logic: "This is related to the scholastic speculations on the *infuturabili*, *i.e.*, the question whether divine omniscience extends to *what would have happened if something had happened that did not happen*." More recently, Zukowski and Brukner [20] suggested that non-locality is to be understood in terms of such *infuturabili*, called there "counterfactual definiteness." We propose an alternative line of argumentation which avoids relating outcomes of measurements that cannot be mutually carried out. Our reasoning employs *complexity* instead of probability: First, non-locality can be *defined* in such a setting, and second, it goes along with *similar consequences* as in the probabilistic view. The new approach is more general but uses, at the same time, fewer assumptions on the conceptual level.

Non-Locality with Counterfactual Reasoning.— Non-local correlations [4] are a fascinating feature of quantum theory. Conceptually challenging is the difficulty [2], [19] to explain their origin *causally*, *i.e.*, according to *Reichenbach's principle*, stating that a correlation between two space-time events can stem from a *common cause* (in the common past) or from a *direct influence* from one to the other [16]. More specifically, the difficulty manifests itself when *alternatives* are taken into account: The argument leading up to a Bell inequality relates outcomes of different measurements, only *one* of which can actually be realized. Does this mean that if we drop the assumption of *counterfactual definiteness* [20] (*i.e.*, of a consistent coexistence of the consequences of mutually exclusive starting points), then the paradox disappears? We argue the answer to be *negative*: Even in the *factuals-only view*, the joint properties — in terms of mutual compressibility — of the involved pieces of information are remarkable since certain *consequences* of non-local correlations, as they are known from the probability calculus, persist. An example is the complexity (instead of randomness) carried over from the inputs to the outputs if no-signaling holds.

In the probabilistic regime, a *Popescu-Rohrlich (PR) box* [14] gives rise to a mechanism of the following kind. Let A and B be the respective input bits to the box and X and Y the (unbiased) output bits with

$$X \oplus Y = A \cdot B \,. \tag{1}$$

This system is *no-signaling*, *i.e.*, the joint input-output behavior is useless for message transmission. According to Fine [8], the non-locality of the system (*i.e.*, conditional distribution) $P_{XY|AB}$ — meaning that it cannot be written as a convex combination of products $P^r_{X|A} \cdot P^r_{Y|B}$ —, is equivalent to the fact that there exists no "roof distribution" $P'_{X_0X_1Y_0Y_1}$ the marginal $P'_{X_iY_j}$ of which equals $P_{XY|A=i,B=j}$ for all $(i,j) \in \{0,1\}^2$: Non-locality is the impossibility of the outputs to *alternative inputs* to consistently coexist.

Such reasoning assumes and concludes that certain pieces of *classical*[1] *information exist or do not exist*. In this way of speaking, Fine's theorem [8] reads: *"The outputs cannot exist before the inputs do."* Let us make this qualitative statement more precise. We assume a perfect PR box, *i.e.*, a system always satisfying (1). Note that (1) alone does not uniquely determine $P_{XY|AB}$ since the marginal of X, for instance, is not fixed. If, however, we additionally require *no-signaling*, then the marginals, such as $P_{X|A=0}$ or $P_{Y|B=0}$, must be *perfectly unbiased* under the assumption that all four (X,Y)-combinations, *i.e.*, $(0,0),(0,1),(1,0)$, and $(1,1)$, are possible. To see this, assume $P_{X|A=0,B=0}(0) > 1/2$. By (1) we can conclude the same for Y: $P_{Y|A=0,B=0}(0) > 1/2$.

[1] *Classicality* of information is, as the limit of macroscopicity [9] when the representing system's size tends to infinite, an idealized notion implying that it can be measured without any disturbance, and that the outcome is always the same. It makes sense to say that a *classical bit U exists* (*i.e.*, takes a definite value) or does not exist.

From no-signaling we get $P_{X|A=0,B=1}(0) > 1/2$. Using symmetry, and no-signaling again, we obtain both $P_{X|A=1,B=1}(0) > 1/2$ and $P_{Y|A=1,B=1}(0) > 1/2$. This contradicts (1) since two bits which are both biased towards 0 cannot differ with certainty. Therefore, our original assumption was wrong: The outputs *must* be perfectly unbiased. Altogether, this means that X *as well as* Y *cannot exist*[2] *before* $f(A,B)$ *does for some nontrivial* $f : \{0,1\}^2 \to \{0,1\}$. The paradoxical aspect of non-locality — at least if a causal structure is in place — is the fact that *fresh pieces of information come to existence in a spacelike-separated way that are perfectly correlated.*

Non-Locality Without Counterfactual Reasoning.— We propose an understanding of non-locality without counterfactual definiteness. We use an asymptotic *Kolmogorov-complexity* calculus for binary strings to replace the probability calculus.

A recent article [13] suggests the use of Kolmogorov complexity in the context of non-local correlations with the objective of avoiding probabilities, but not of resigning from counterfactual arguments; in fact, the outcomes of alternative measurements *are* assumed to coexist in [13]; their argument builds on the relation between these alternative output data.

An asymptotic Kolmogorov calculus.— Let $\mathcal{U}$ be a fixed universal Turing machine (TM).[3] For a finite or infinite string s, the *Kolmogorov complexity* [10], [11] $K(s) = K_{\mathcal{U}}(s)$ is the length of the shortest program for $\mathcal{U}$ such that the machine outputs s. Note that $K(s)$ can be infinite if s is.

Let $a = (a_1, a_2, \ldots)$ be an infinite string. Then $a_{[n]} := (a_1, \ldots, a_n, 0, \ldots)$. We study the asymptotic behavior of, *e.g.*, $K(a_{[n]}) : \mathbf{N} \to \mathbf{N}$. For this function, we simply write $K(a)$, similarly $K(a \,|\, b)$ for $K(a_{[n]} \,|\, b_{[n]})$, the latter being the length of the shortest program outputting $a_{[n]}$ on input $b_{[n]}$. We write

$$K(a) \approx n :\Longleftrightarrow \lim_{n\to\infty} \big(K(a_{[n]})/n\big) = 1 \ .$$

We call a string a with this property *incompressible*. We also use $K(a_{[n]}) = \Theta(n)$, as well as

$$K(a) \approx 0 :\Longleftrightarrow \lim_{n\to\infty} \big(K(a_{[n]})/n\big) = 0 \Longleftrightarrow K(a_{[n]}) = o(n).$$

Note that *computable* strings a satisfy $K(a) \approx 0$.

Generally, for functions $f(n)$ and $g(n) \not\approx 0$, we write $f \approx g$ if $f/g \to 1$. *Independence of* a *and* b is then

$$K(a \,|\, b) \approx K(a)$$

or, equivalently, $K(a,b) \approx K(a) + K(b)$. If we introduce $I_K(x;y) := K(x) - K(x \,|\, y) \approx K(y) - K(y \,|\, x)$, independence of a and b is $I_K(a,b) \approx 0$.

In the same spirit, we can define *conditional independence*: We say that *a and b are independent given c* if

$$K(a,b \,|\, c) \approx K(a \,|\, c) + K(b \,|\, c)$$

or, equivalently, $K(a \,|\, b,c) \approx K(a \,|\, c)$, or $I_K(a;b \,|\, c) := K(a \,|\, c) - K(a \,|\, bc) \approx 0$.

Uncomputability from PR boxes and incompressible inputs.— Let now (a,b,x,y) be infinite binary strings with

$$x_i \oplus y_i = a_i \cdot b_i \quad \text{for all} \quad i = 1,2,\ldots \ . \tag{2}$$

Although the intuition is that the strings stand for the inputs and outputs of a PR box, no dynamic meaning is attached to them (or to the "box," for that matter) since there is *no free choice of an input and no generation of an output in function of this input*; all we have are four fixed strings satisfying the PR condition. Nothing prevents us from defining this static situation to be *no-signaling*:

$$K(x \,|\, a) \approx K(x \,|\, ab) \quad \text{and} \quad K(y \,|\, b) \approx K(y \,|\, ab) \ . \tag{3}$$

Recall the mechanism enabled by the maximal non-locality of the PR box: *If the inputs are not entirely fixed, then the outputs must be completely unbiased, as long as the system is no-signaling.* We can now draw a statement of similar flavor but entirely within *actual* data (see Figure 1).

Theorem 1. *If the input pair to a PR box is incompressible and no-signaling holds, then the outputs are uncomputable.*

Proof. Let $(a,b,x,y) \in (\{0,1\}^{\mathbf{N}})^4$ with (2), no-signaling (3), and $K(a,b) \approx 2n$, *i.e.*, the input pair is incompressible. We conclude $K(a \cdot b \,|\, b) \approx n/2$. Note first that $b_i = 0$ implies $a_i \cdot b_i = 0$, and second that any further compression of $a \cdot b$, given b, would lead to "structure in (a,b)," *i.e.*, a possibility of describing or programming (a,b) in shorter than $2n$. Observe now

$$K(x \,|\, b) + K(y \,|\, b) \gtrsim K(a \cdot b \,|\, b) \tag{4}$$

since x and y together determine $a \cdot b$. Now, (4) implies

$$K(y \,|\, b) \gtrsim K(a \cdot b \,|\, b) - K(x \,|\, b) \gtrsim n/2 - K(x) \ . \tag{5}$$

On the other hand,

$$K(y \,|\, ab) \approx K(x \,|\, ab) \le K(x) \ . \tag{6}$$

Now, no-signaling (3), together with (5) and (6), implies $n/2 - K(x) \lesssim K(x)$, hence, $K(x) \gtrsim n/4 = \Theta(n)$. The string x must be *uncomputable*. □

Theorem 1 raises a number of questions: Does a similar result hold for the *conditional* complexities $K(x \,|\, a)$

[2] Actually, there cannot even exist a classical value *arbitrarily weakly* correlated with either X or Y.

[3] We can assume a *fixed* machine here since the introduced asymptotic notions are independent of this choice.

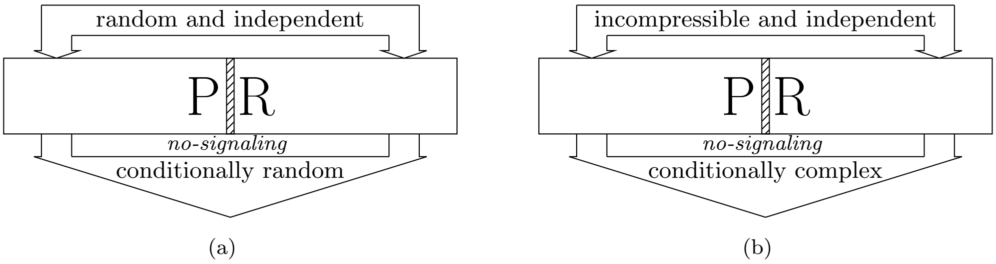

Figure 1. The traditional (a) vs. the new (b) view: Non-locality *à la* Popescu/Rohrlich (PR) plus no-signaling leads to the output inheriting *randomness* (a) or *complexity* (b), respectively, from the input.

and $K(y\,|\,b)$, and from *quantum* non-local correlations? Can we give a *general definition* of non-locality and does a similar result as the above hold with respect to *any* non-local correlation? In the remainder, we address these questions.

Conditional uncomputability of a PR box' outputs.—

Theorem 2. *Under the conditions of Theorem 1, we have* $K(x\,|\,a) = \Theta(n)$ *and* $K(y\,|\,b) = \Theta(n)$.

Proof. Note first

$$K(x\,|\,a) \approx 0 \Leftrightarrow K(x\,|\,ab) \approx K(y\,|\,ab) \approx 0 \Leftrightarrow K(y\,|\,b) \approx 0,$$

i.e., the two quantities of interest are negligible simultaneously. In order to show that they are both $\Theta(n)$, we assume $K(x\,|\,a) \approx 0 \approx K(y\,|\,b)$ instead. Then, there exist programs P_n and Q_n, both of length $o(n)$, computing functions f_n and g_n with

$$f_n(a_{[n]}) \oplus g_n(b_{[n]}) = a_{[n]} \cdot b_{[n]} \ . \tag{7}$$

For fixed (families of) functions f_n and g_n, asymptotically how many $(a_{[n]}, b_{[n]})$ can there be at most that satisfy (7)? This question boils down to a parallel-repetition analysis of the PR game. A result by Raz [15] implies that this number is of order $(2-\Theta(1))^{2n}$. Therefore, the two programs P_n and Q_n — together with the index of length $(1-\Theta(1))2n$ to single out the correct pair $(a_{[n]}, b_{[n]})$ within the candidates' list of length $(2-\Theta(1))^{2n}$ — leads to a program of length $o(n) + (1-\Theta(1))2n$ with output $(a_{[n]}, b_{[n]})$, in conflict with (a,b)'s incompressibility. □

Conditional uncomputability from quantum correlations: Chained Bell inequality and magic-square game.— Unfortunately, the quantum-physically achievable approximations to the PR box do not seem to allow for a similar reasoning immediately. We can, however, use the violation of the *chained Bell inequality* (see, *e.g.*, [3], [7]) realizable in the laboratory [18].

To the chained Bell inequality belongs the following idealized system: Let $A, B \in \{1, \ldots, m\}$ be the inputs. We assume the "promise" that B is congruent to A or to $A+1$ modulo m. Given this, the outputs $X, Y \in \{0,1\}$ must satisfy

$$X \oplus Y = \chi_{A=m,B=1} \ , \tag{8}$$

where $\chi_{A=m,B=1}$ is the characteristic function of the event $\{A = m, B = 1\}$. Barrett, Hardy, and Kent [3] showed that if A and B are random, then X and Y must be almost perfectly unbiased if the system is no-signaling. More precisely, they were able to show such a statement from the gap between the error probability of the best quantum $(\Theta(1/m^2))$ as opposed to classical $(\Theta(1/m))$ strategy for winning the game.

Theorem 3. *Let* $(a,b,x,y) \in (\{1,\ldots,m\}^n)^2 \times (\{0,1\}^n)^2$ *be respecting the promise and such that*

$$K(a,b) \approx (\log m + 1) \cdot n \ ,$$

i.e., (a,b) *is incompressible conditioned on the promise; the system is no-signaling* (3)*; the fraction of quadruples* (a_i, b_i, x_i, y_i), $i = 1, \ldots, n$, *with* (8) *is* $1 - \Theta(1/m^2)$. *Then* $K(x) = \Theta(n)$. *In particular,* x *is uncomputable.*

Proof. Since $K(a,b)$ is maximal, we have[4]

$$K(\chi_{a=m,b=1}\,|\,b) = h(\Theta(1/m))n \ . \tag{9}$$

Note first that this is the maximal complexity of a string in which the fraction of 1's is of order $\Theta(1/m)$, since the number of such strings is

$$\log \binom{n}{\Theta(1/m)n} \approx h(\Theta(1/m))n \ .$$

Second, if the complexity were any lower, this would lead to the possibility of compressing (a,b), which is excluded. Now, we have

$$K(x\,|\,b) + K(y\,|\,b) + h(\Theta(1/m^2))n \gtrsim K(\chi_{a=m,b=1}\,|\,b)$$

[4] h is the binary entropy $h(p) = -p\log p - (1-p)\log(1-p)$. Usually, p is a probability, but h is invoked here merely as an approximation for binomial coefficients.

since one possibility for generating the string $\chi_{a=m,b=1}$ from position 1 to n is to generate $x_{[n]}$ and $y_{[n]}$, as well as the string indicating the positions where (8) is violated; the complexity of the latter is at most

$$\log \binom{n}{\Theta(1/m^2)n} \approx h(\Theta(1/m^2))n \ .$$

Together with (9), we get

$$K(y \,|\, b) \geq \Theta(1/m)n - K(x) \qquad (10)$$

if m is sufficiently large. On the other hand,

$$\begin{aligned} K(y \,|\, ab) &\lesssim K(x \,|\, ab) + h(\Theta(1/m^2))n \\ &\leq K(x) + h(\Theta(1/m^2))n \ . \end{aligned} \qquad (11)$$

Now, (3), (10), and (11) together imply $K(x) = \Theta(n)$. □

As above, the application of Raz' parallel-repetition result [15] allows for proving *conditional* uncomputability:

$$K(x \,|\, a) = \Theta(n) \ .$$

In order to see that, note first

$$K(y \,|\, ab) \lesssim K(x \,|\, ab) + h(\Theta(1/m^2))n \ .$$

Therefore, the assumption $K(x \,|\, a) \approx K(x \,|\, ab) \approx 0$ leads, just as above, to a program generating $(a_{[n]}, b_{[n]})$ of length

$$\begin{aligned} & h(\Theta(1/m^2))n + (1 - h(\Theta(1/m)))(\log m + 1) \\ & = (1 - \Theta(1))(\log m + 1) \not\approx (\log m + 1)n \end{aligned} \qquad (12)$$

for fixed, sufficiently large m, in conflict with the assumption that (a, b) be incompressible. In (12) we use that for fixed output-alphabet sizes, Raz' result bounds the success probability in the parallel repetition of the game by $2^{-c \cdot 1/m \cdot n}$ for some constant c, and that $1/m^2 < c/m$ holds for sufficiently large m.

For any non-local behavior characterizable by a condition that is always satisfiable with entanglement, but not *without* this resource — so called "pseudo-telepathy" games [5] —, a similar reasoning shows that incompressibility of the inputs leads to uncomputability of at least one of the two outputs, even given the corresponding input. We illustrate the argument with the example of the *magic-square game* [1]: Let $(a, b, x, y) \in (\{1, 2, 3\}^{\mathbf{N}})^2 \times (\{1, 2, 3, 4\}^{\mathbf{N}})^2$ be the quadruple of the inputs and outputs, respectively, and assume that the pair (a, b) is incompressible as well as $K(x \,|\, a) \approx 0 \approx K(y \,|\, b)$. Then there exist $o(n)$-length programs P_n, Q_n such that $x_n = P_n(a_{[n]})$ and $y_n = Q_n(b_{[n]})$. Again, Raz' parallel-repetition theorem [15] implies that the length of a program generating $(a_{[n]}, b_{[n]})$ is, including the employed sub-routines P_n and Q_n, of order $(1 - \Theta(1))\mathrm{len}(a_{[n]}, b_{[n]})$ — in violation of the incompressibility of (a, b).

General factual definition of (non-)locality.— We propose the following definition of when a no-signaling quadruple $(a, b, x, y) \in (\{0, 1\}^{\mathbf{N}})^4$ (where a, b are the inputs and x, y the outputs) is *local*. There exists $\lambda \in \{0, 1\}^{\mathbf{N}}$ such that

$$\begin{aligned} & K(a, b, \lambda) \approx K(a, b) + K(\lambda) \ , \\ & K(x \,|\, a\lambda) \approx 0 \ , \text{ and } K(y \,|\, b\lambda) \approx 0 \ . \end{aligned} \qquad (13)$$

Sufficient conditions for locality are $K(a, b) \approx 0$ or $K(x, y) \approx 0$ because of $\lambda := (x, y)$. At the other end of the scale, we expect that for any non-local "system," the fact that $K(a, b)$ is maximal implies x or y to be conditionally uncomputable, given a or b, respectively.

It is a natural question whether the given definition harmonizes with the probabilistic understanding. Indeed, the latter can be seen as a special case of the former: If the (fixed) strings are *typical sequences* of a stochastic process, our non-locality definition implies non-locality of the corresponding conditional distribution. The reason is that a hidden variable of the distribution immediately gives rise, through sampling, to a λ in the sense of (13). Note, however, that our formalism is more general since most strings *cannot* be seen as typical sequences of such a process.

Conclusion.— We propose a view on non-locality that does not rely on relating outcomes of measurements that cannot be actually carried out altogether. It is based on the notion of *complexity* rather than probability. In the argument, Kolmogorov complexity — defined with respect to Turing machines — can be replaced by other computing models with finite machine descriptions. While the resulting reasoning, leading to similar mechanisms enabled by non-locality, is more general than the probabilistic regime, a central assumption — the existence of results of unperformed measurements — can even be dropped. Our statements and reasoning are asymptotic and apply to *sufficiently long finite strings*.

In the derivation of Bell inequalities and, hence, the analysis of real-life experiments demonstrating non-locality, the assumption that the outcomes of all alternative measurements exist together is used to be implicitly made. Our line of reasoning suggests a more *direct* discussion, referring only to the data at hand. Note that whereas Kolmogorov complexity is uncomputable itself, upper bounds on the quantity can be obtained easily from any compression algorithm [6].

In short, our main result is as follows. If the settings' description in a non-locality experiment is incompressible, then the outcomes must be uncomputable, even given the respective inputs: *If the experimenter is able to generate an incompressible string, then the measured photons must be able to come up with a non-computable behavior as well.* This gives an all-or-nothing flavor to the Church-Turing hypothesis, since "beyond-TM" computations either do not exist at all, or they occur even in individual photons. Tightened versions of our results give rise to a physical *incompressibility-amplification and -expansion mechanism*.

Acknowledgement. The author thanks Mateus Araújo, Ämin Baumeler, Harvey Brown, Caslav Brukner, Bora Dakic, Paul Erker, Jürg Fröhlich, Arne Hansen, Marcus Huber, Alberto Montina, Benno Salwey, Andreas Winter, and Magdalena Zych for enlightening discussions. The author is supported by the Swiss National Science Foundation (SNF), the National Centre of Competence in Research "Quantum Science and Technology" (QSIT), and the COST action on Fundamental Problems in Quantum Physics. We thank all members of the *Facoltà indipendente di Gandria* for their support.

[1] P. K. Aravind, Bell's theorem without inequalities and only two distant observers, *Foundations of Physics Letters*, Vol. 15, No. 4, pp. 397-405, 2002.

[2] T. J. Barnea, J.-D. Bancal, Y.-C. Liang, N. Gisin, Tripartite quantum state violating the hidden influence constraints, *Phys. Rev. A*, Vol. 88, No. 022123, 2013.

[3] J. Barrett, L. Hardy, and A. Kent, No-signalling and quantum key distribution, *Phys. Rev. Lett.*, Vol. 95, No. 010503, 2005.

[4] J. S. Bell, On the Einstein-Podolsky-Rosen paradox, *Physics*, Vol. 1, pp. 195–200, 1964.

[5] G. Brassard, A. Broadbent, and A. Tapp, Quantum pseudo-telepathy, quant-ph/0407221, 2004.

[6] R. Cilibrasi and P. Vitányi, Clustering by compression, *IEEE Transactions on Information Theory* Vol. 51, No. 4, 1523-1545, 2005.

[7] R. Colbeck and R. Renner, Free randomness can be amplified, *Nature Physics*, Vol. 8, pp. 450–454, 2012.

[8] A. Fine, Hidden variables, joint probability, and the Bell inequalities, *Phys. Rev. Lett.*, Vol. 48, p. 291, 1982.

[9] N. Gisin, Quantum measurement of spins and magnets, and the classical limit of PR-boxes, quant-ph/1407.8122, 2014.

[10] A. N. Kolmogorov, Three approaches to the quantitative definition of information, Problemy Peredachi Informatsii, Vol. 1, No. 1, pp. 3–11, 1965.

[11] M. Li and P. Vitányi, *An Introduction to Kolmogorov Complexity and Its Applications*, Springer-Verlag, 2008.

[12] A. Peres, Unperformed experiments have no results, *American Journal of Physics*, Vol. 46, No. 7, pp. 745–747, 1978.

[13] H. S. Poh, M. Markiewicz, P. Kurzyński, A. Cerè, D. Kaszlikowski, and C. Kurtsiefer, Probing quantum-classical boundary with compression software, quant-ph/1504.03126, 2015.

[14] S. Popescu and D. Rohrlich, Non-locality as an axiom, *Foundations of Physics*, Vol. 24, pp. 379, 1994.

[15] R. Raz, A parallel repetition theorem, *SIAM Journal on Computing*, Vol. 27, No. 3, pp. 763–803, 1998.

[16] H. Reichenbach, The principle of the common cause, in *The Direction of Time*, Chap. 19, pp. 157–167, California Press, Berkeley, 1956.

[17] E. Specker, Die Logik nicht gleichzeitig entscheidbarer Aussagen, *Dialectica*, Vol. 14, pp. 239–246, 1960.

[18] T. E. Stuart, J. A. Slater, R. Colbeck, R. Renner, and W. Tittel, An experimental test of all theories with predictive power beyond quantum theory, *Phys. Rev. Lett.*, Vol. 109, No. 020402, 2012.

[19] C. Wood and R. Spekkens, The lesson of causal discovery algorithms for quantum correlations: Causal explanations of Bell-inequality violations require fine-tuning, *New J. Phys.*, Vol. 17, No. 033002, 2015.

[20] M. Zukowski and C. Brukner, Quantum non-locality - It ain't necessarily so…, *J. Phys. A: Math. Theor.*, Vol. 47, No. 424009, 2014.