

Probing quantum-classical boundary with compression software

Hou Shun Poh ^{*}, Marcin Markiewicz ^{† ‡}, Paweł Kurzyński ^{* §}, Alessandro Cerè ^{* ¶}, Dagomir Kaszlikowski ^{* ¶}, and Christian Kurtsiefer ^{* ¶}

^{*}Center for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543, [†]Faculty of Physics, University of Warsaw, ul. Pasteura 5, PL-02-093 Warszawa, Poland, [‡]Institute of Theoretical Physics and Astrophysics, University of Gdansk, ul. Wita Stwosza 57, PL-80-952, Gdansk, Poland, [§]Faculty of Physics, Adam Mickiewicz University, Umultowska 85, 61-614 Poznań, Poland, and [¶]Department of Physics, National University of Singapore, 2 Science Drive 3, Singapore 117542

Submitted to Proceedings of the National Academy of Sciences of the United States of America

We experimentally demonstrate that it is impossible to simulate quantum bipartite correlations with a deterministic universal Turing machine. Our approach is based on the Normalized Information Distance (NID) that allows the comparison of two pieces of data without detailed knowledge about their origin. Using NID, we derive an inequality for output of two local deterministic universal Turing machines with correlated inputs. This inequality is violated by correlations generated by a maximally entangled polarization state of two photons. The violation is shown using a freely available lossless compression program. The presented technique may allow to complement the common statistical interpretation of quantum physics by an algorithmic one.

Kolmogorov Complexity | Quantum correlations | Description of reality | Turing Machine

Abbreviations: NID, Normalized Information Distance; NCD, Normalized Compression Distance

Introduction

The idea that physical processes can be considered as computations done on some universal machines traces back to Turing and von Neumann [1], and the growth of the computational power allowed for further development of these concepts. This resulted in a completely new approach to science in which the complexity of observed phenomena is closely related to the complexity of computational resources needed to simulate them [2]. In addition, there are physical phenomena that simply cannot be traced with analytical tools, which further motivated a computational approach to physics [3]. Moreover, the idea of quantum computation [4] lead to a discovery of a few problems that seem not efficiently traceable on classical computers but efficiently on a quantum version [5, 6].

Classical physics can be simulated on universal Turing machines, or other computationally equivalent models [7]. On the other hand, efficient simulation of quantum systems requires a replacement of deterministic universal Turing machines with quantum computers whose states are non-classically correlated. Such machines can even simulate any local quantum system efficiently [8, 9]. Can we experimentally distinguish between these two descriptions of the universe using a logically self-contained computational approach?

In this paper, we show that there are processes which cannot be simulated on local classical machines at all, independently of the available classical resources. We first introduce the notion of Kolmogorov complexity, a measure of the classical complexity of a phenomena, and later apply it to derive a bound on classical descriptions [10]. Next, we use the fact that Kolmogorov complexity can be approximated by compression algorithms [11].

We compress experimental data obtained from polarisation measurements on entangled photon pairs and show the violation of a classical bound.

Let’s consider the description of a machine, whether classical or quantum, that outputs a string x made of 0’s and 1’s. In the case of a Turing machine U , we can always write a program Λ that generates x . The simplest such program is obviously ‘PRINT x ’. However, this is not optimal: in many cases the program can be much shorter than the string itself.

This brings us to the concept of Kolmogorov complexity $K(x)$, the minimal length $l(\Lambda)$ of all programs Λ that reproduce a specific output x . If $K(x)$ is of the order of the length of the output $l(x)$ then our algorithmic description of x is inefficient, and x is called algorithmically random [12]. In most cases $K(x)$ cannot be computed [13]. To circumvent this issue, we can estimate $K(x)$ with some efficient lossless compression algorithm $C(x)$ [11].

We now extend this picture by considering two Turing machines U_A (Alice) and U_B (Bob), which are spatially separated. If these machines cannot communicate, they generate two output strings that are independent, although the programs fed into the machines can be correlated. Moreover, the input programs are classical bit strings so the correlations between them must be classical.

We determine the complexity of the generated strings using the *Normalized Information Distance* (NID) [10]. This distance allows for a comparison of two data sets without detailed knowledge about their origin. In practice, we evaluate an approximation to the NID, the *Normalized Compression*

Significance

Composite quantum systems show correlations that can not be explained by any composition of models for individual systems. An often cited indication for this is the violation of a Bell inequality, which is inherently based on a statistical argument. Inspired by the idea that the universe may be simulated by universal computers like Turing machines, we replace statistical tests with one that evaluates the complexity of experimental measurement results, and preset an algorithmic equivalent of a Bell inequality. We evaluate the complexity of a physical system of entangled photon pairs by assessing the compressibility of measurement data, and find correlations beyond any “classical” model. This technique can complement the common statistical interpretation of quantum physics with an algorithmic one.

Reserved for Publication Footnotes

Distance (NCD) [11], using a lossless compression software, in our case the LZMA Utilities, based on the Lempel-Ziv-Markov chain algorithm [14].

Simulation by deterministic universal Turing machines

We consider a model experiment, similar to the one used for testing the Bell inequalities [15]: a source emits pairs of photons that travel to two separate polarization analyzers M_A (Alice) and M_B (Bob). Each analyzer has two outputs associated with bit values 0 and 1, and can be set along directions a_0 or a_1 for M_A , and b_0 or b_1 for M_B . The record of the outputs from each analyzer forms a bit string (see figure 1).

The output x of each individual analyzer can be described as the output of a Turing machine U , fed with the settings a_j or b_k , and a program Λ . The program will contain the information for generating the correct output for every detection event and for every setting.

If we consider a string of finite length $l(x) = N$, Λ will have to describe the 4^N possible events. The length of the shortest Λ is equal to the Kolmogorov complexity of the generated string.

Next, we consider the simulation of the experiment with two local non-communicating machines U_A and U_B (see figure 2). We feed a program Λ to both of them and obtain two output strings, x and y , both of length N . In this case, the program has to describe the behavior of all $2N$ events for all possible settings a_j and b_k , hence 16^N possible events.

Normalized Information Distance. The Kolmogorov complexity of two bit strings $K(x, y)$ is the length of the shortest program generating them simultaneously. $K(x, y)$ can be shorter than $K(x) + K(y)$ if x and y are correlated - the more correlated they are, the simpler it is to compute one string knowing the other. This idea was further carried out by Cilibrasi and Vitanyi [11] who constructed a distance measure between x and y called *Normalized Information Distance* (NID),

$$\text{NID}(x, y) = \frac{K(x, y) - \min\{K(x), K(y)\}}{\max\{K(x), K(y)\}}. \quad [1]$$

The NID obeys all required properties of a metric, in particular, the triangle inequality

$$\text{NID}(x, y) + \text{NID}(y, z) \geq \text{NID}(x, z). \quad [2]$$

The above inequality holds up to a correction of order $\log(l(x))$, which can be neglected for sufficiently long strings [11].

Information Inequality. We consider the bit strings x_{a_j} and y_{b_k} generated by Alice and Bob with fixed setting a_j and b_k . Equation 2 then transforms into

$$\text{NID}(x_{a_0}, y_{b_0}) + \text{NID}(y_{b_0}, y_{b_1}) \geq \text{NID}(x_{a_0}, y_{b_1}). \quad [3]$$

However, $\text{NID}(y_{b_0}, y_{b_1})$ cannot be determined experimentally because the strings y_{b_0} and y_{b_1} come from measurements of incompatible observables. We therefore use the triangle inequality

$$\text{NID}(x_{a_1}, y_{b_0}) + \text{NID}(x_{a_1}, y_{b_1}) \geq \text{NID}(y_{b_0}, y_{b_1}), \quad [4]$$

and combine it with inequality (3) to obtain a quadrangle inequality:

$$\text{NID}(x_{a_0}, y_{b_0}) + \text{NID}(x_{a_1}, y_{b_0}) + \text{NID}(x_{a_1}, y_{b_1}) \geq \text{NID}(x_{a_0}, y_{b_1}). \quad [5]$$

Similar to various tests of Bell inequalities, we introduce a scalar quantity S' that quantifies the degree of violation of equation 5:

$$\begin{aligned} S' &= \text{NID}(x_{a_0}, y_{b_1}) - \text{NID}(x_{a_0}, y_{b_0}) \\ &\quad - \text{NID}(x_{a_1}, y_{b_0}) - \text{NID}(x_{a_1}, y_{b_1}) \leq 0 \end{aligned} \quad [6]$$

In order to experimentally test this inequality, we have to address the following problem. We can set up a source to generate entangled photon pairs in a state of our choosing, but we cannot control the nature of the measurement. For every experimental run i with the same preparation the resulting string x_{i,a_j} can be different. Consequently, the corresponding program Λ_i is different for every experimental run.

It is reasonable to assume that for every two experimental runs i and i' the complexity of the generated strings remains the same: $K(x_{i,a_j}) = K(x_{i',a_j})$ and $K(x_{i,a_j}, y_{i,b_k}) = K(x_{i',a_j}, y_{i',b_k})$. Without these assumptions the same physical preparation of the experiment has different consequences and thus the notion of preparation loses its meaning. More generally, the predictive power of science can be expressed by saying that the same preparation results in the same complexity of observed phenomena.

Estimation of Kolmogorov complexity

In general the Kolmogorov complexity cannot be evaluated, but it can be estimated. One can adapt two conceptually different approaches.

Statistical Approach. This approach takes into account the ensemble of all possible N -bit strings and asks about their average Kolmogorov complexity. It can be shown that this average equals the Shannon entropy $H(X)$ of the ensemble [13], and thus

$$\langle \text{NID}(x, y) \rangle = \frac{H(x, y) - \min\{H(x), H(y)\}}{\max\{H(x), H(y)\}}. \quad [7]$$

Inequality (5) becomes a type of entropic Bell inequality introduced by Braunstein and Caves [16] if local entropies are maximal, i.e., $H(x) = H(y) = N$. They showed that for a maximally entangled polarization state of two photons, and polarizer angles obeying the constraints $\vec{a}_0 \cdot \vec{b}_1 = \cos 3\theta$, $\vec{a}_0 \cdot \vec{b}_0 = \vec{a}_1 \cdot \vec{b}_0 = \vec{a}_1 \cdot \vec{b}_1 = \cos \theta$, inequality (5) is violated for an appropriate range of θ . Calculating the entropy $H(x, y)$ using the probability distributions predicted by quantum mechanics, it is possible to obtain the expected value of S' as a function of θ (Fig. 4a). The maximal violation of this inequality is $S' = 0.24$, with a separation of $\theta = 8.6^\circ$.

Algorithmic approach. It is possible to avoid a statistical description of our experiment following the ideas pioneered in [11]. There, it was shown that the Kolmogorov complexity can be well approximated by the application of compression algorithms. This approximation introduces the new distance called *Normalized Compression Distance* (NCD)

$$\text{NCD}(x, y) = \frac{C(x, y) - \min\{C(x), C(y)\}}{\max\{C(x), C(y)\}}, \quad [8]$$

where $C(x)$ is the length of the compressed string x , and $C(x, y)$ is the length of the compressed concatenated strings x, y . Replacing NID with NCD in equation 6 leads to a new inequality:

$$\begin{aligned} S' \rightarrow S &= \text{NCD}(x_{a_0}, y_{b_1}) - \text{NCD}(x_{a_0}, y_{b_0}) \\ &\quad - \text{NCD}(x_{a_1}, y_{b_0}) - \text{NCD}(x_{a_1}, y_{b_1}) \leq 0. \end{aligned} \quad [9]$$

This expression can be tested experimentally because the NCD distance measure is operationally defined.

Choice of compressor

Before moving to the experiment, we need to ensure the suitability of the compression software we use to evaluate the NCD. For this, we numerically simulate the outcome of an experiment, based on a distribution of results predicted by quantum physics. Among the packages we tested, we found that the LZMA Utility [14] approaches the Shannon limit [17] most closely.

The simulation also allows us to verify the angle that maximizes the violation 9. The results of the simulation are presented in figure 4. More details on the generation of the simulated data and the choice of the compressor are provided in the Appendix.

Experiment

In our experiment (see figure 3), the output of a grating-stabilized laser diode (LD, central wavelength 405 nm) passes through a single mode optical fiber (SMF) for spatial mode filtering, and is focused to a beam waist of $80\text{ }\mu\text{m}$ into a 2 mm thick BBO crystal. In this crystal (cut for type-II phase-matching), photon pairs are generated via spontaneous parametric down-conversion (SPDC) in a slightly non-collinear configuration. A half-wave plate ($\lambda/2$) and a pair of compensation crystals (CC) take care of the temporal and transversal walk-off [18].

Two spatial modes (labeled A and B) of down-converted light, defined by the SMFs for 810 nm, are matched to the pump mode to optimize the collection [19]. In type-II SPDC, each down-converted pair consists of an ordinary and extraordinarily polarized photon, corresponding to horizontal (H) and vertical (V) in our setup. A pair of polarization controllers (PC) ensures that the SMFs do not affect the polarization of the collected photons. To arrive at an approximate singlet Bell state, the phase ϕ between the two decay possibilities in the polarization state $|\psi\rangle = \frac{1}{\sqrt{2}}(|H\rangle_A|V\rangle_B + e^{i\phi}|V\rangle_A|H\rangle_B)$, is adjusted to $\phi = \pi$ by tilting the CC.

In the polarization analyzers (figure 3), the photons from SPDC are projected onto arbitrary linear polarization by $\lambda/2$ plates, set to half of the analyzing angles $\theta_{A(B)}$, and polarization beam splitter (PBS) in each analyzer. Photons are detected by avalanche photo diodes (APDs), and corresponding detection events from the same pair identified by a coincidence unit (CU) if they arrive within $\approx \pm 3\text{ ns}$ of each other.

The quality of polarization entanglement is tested by probing the polarization correlations in a basis complementary to the intrinsic HV basis of the crystal. With interference filters (IF) of 5 nm bandwidth (FWHM) centered at 810 nm, in the 45° linear polarization basis we observe a visibility $V_{45} = 99.9 \pm 0.1\%$. The visibility in the natural H/V basis of the type-II down-conversion process also reaches $V_{HV} = 99.9 \pm 0.1\%$. A separate test of a CHSH-type Bell inequality [20] leads to a value of $S = 2.826 \pm 0.0015$. This indicates a relatively high quality of polarization entanglement; the uncertainties in the visibilities are obtained from propagated Poissonian counting statistics.

Measurement and Data Post-processing. We record two-fold coincidences of detection events between detectors at A and B. For each PBS, the transmitted output is associated with 0 and the reflected one with 1. The resulting binary strings x from A, and y from B are written into two individual binary files. From these, we calculate the NCD using equation 8. This procedure is repeated for each of the four settings (a_0, b_0) , (a_1, b_0) , (a_1, b_1) , and (a_0, b_1) in order to obtain the value for S .

To remove the bias due to differences in the detection efficiency of the APDs in the experiment, we also measure for each setting the associated orthogonal ones (see Appendix for details).

Results

The inequality is experimentally tested by evaluating S in equation 9 for a range of θ ; the obtained values [points (c), (d) in figure 4] are consistently lower than the trace (a) calculated via entropy using equation 7, and than a simulation with the same compressor (b). This is because the LZMA Utility is not working exactly at the Shannon limit, and also due to imperfect state generation and detection.

For $\theta = 8.6^\circ$ we collected results from a large number of photon pairs. Although we set out in this work to avoid a statistical argument in the interpretation of measurement results, we do resort to statistical techniques to assess the confidence in an experimental finding of a violation of inequality equation 9. To estimate an uncertainty of the experimentally obtained values for S , this large data set was subdivided into files with length greater than 10^5 bits. The results from all the subdivided files are then averaged to obtain the final result of $S(\theta = 8.6^\circ) = 0.0494 \pm 0.0076$, with the latter indicating a relatively small standard deviation over these different subsets.

Discussion

There is a trend to look at physical systems and processes as programs run on a computer made of the constituents of our universe. We could show that this is not possible if one uses a computation paradigm of a local deterministic Turing machine. Although this has been already extensively researched in quantum information theory, we present a complementary algorithmic approach for an explicit, experimentally testable example. This algorithmic approach is complementary to the orthodox Bell inequality approach to quantum nonlocality [15] that is statistical in its nature.

Any process that can be simulated on a local universal Turing machine can be encoded as a program that is fed into it. For every such a program there exists its shortest description called Kolmogorov complexity, which in most of the cases can only be approximated using compression software. Moreover, such a description must obey distance properties as shown in [10, 11]. By testing equation 9, we showed that this is unattainable in the specific case of polarization-entangled photon pairs. Therefore, there exist physical processes that cannot be simulated on local universal Turing machines.

There are two fundamentally different notions of complexity in computer science. On one hand, computational complexity, mainly researched on in quantum information science, studies how much resources are needed to solve a computational problem. These studies focus on complexity classes such as P, NP [21], and its main concern is, given an input program, how efficiently it can be computed. On the other hand, algorithmic complexity deals with a problem of what the most efficient encoding of an input program is. This complementary problem to computation complexity has not yet received enough attention in quantum information science, and it would require a further work on quantum version of Kolmogorov complexity [22].

We would like to stress that our analysis of the experimental data is purely and consistently algorithmic. We do not resort to statistical methods that are alien to the concept of computation. If this approach can be extended to all quantum experiments, it would allow us to bypass the commonly used statistical interpretation of quantum theory.

ACKNOWLEDGMENTS. We acknowledge the support of this work by the National Research Foundation & Ministry of Education in Singapore, partly through the Academic Research Fund MOE2012-T3-1-009. P.K. and D.K. are also supported by the Foundational Questions Institute (FQXi). A.C. also thanks Andrea Baronchelli for the hints on the use of compression software.

Appendix: Symmetrization of detector efficiencies

The experimental setup (figure 3) uses four APDs: D_{HA} , D_{VA} (Alice), and D_{HB} , D_{VB} (Bob) to register photon pair events in the two spatial modes. By denoting events at D_H and D_V as 1 and 0, the four possible output patterns are 00, 01, 10, and 11, where the least and most significant bit corresponds to the Alice and Bob mode, respectively. Due to differences in the losses in the transmitted and reflected port of the PBS, efficiencies in coupling light into the APDs, and the quantum efficiencies of APDs, the detection efficiencies for the four output combinations are different. The resulting effective pair efficiencies are then given by the product of the contributing detection efficiencies η_{VB} , η_{HB} , η_{VA} , and η_{HA} .

This asymmetry will skew the statistics of the measurement results. We symmetrize the effective pair efficiencies for each (θ_A, θ_B) measuring also the following settings: $(\theta_A + 45^\circ, \theta_B)$, $(\theta_A, \theta_B + 45^\circ)$, and $(\theta_A + 45^\circ, \theta_B + 45^\circ)$. This procedure swaps the output ports of the PBS at which each polarization state is detected. The resulting outcomes are then interleaved, providing a uniform detection probability for the four possible outcomes. The effective pair detection efficiency for all four combinations is then $(\eta_{VB}\eta_{VA} + \eta_{VB}\eta_{HA} + \eta_{HB}\eta_{VA} + \eta_{HB}\eta_{HA})/4$.

Appendix: Choice of compressor

In order to evaluate the NCDs of the binary strings, we need to choose a compression algorithm that performs close to the Shannon limit. This is necessary to ensure that it does not introduce any unintended artifacts that lead to an overestimation of the violation. Preferably we want to work in the regime where the obtained NCDs always underestimate the violation. For this purpose, we characterized four compression algorithms implemented by freely available compression programs: *lzma* [14], *bzip2* [23], *gzip* [24], and *lzw* [25]. To eliminate the overhead associated with the compression of ASCII text files, we save data in a binary format.

For this characterization and a simulation of the experiment, we need to generate a “random” string of bits (1, 0)

or pairs of bits (00, 01, 10, and 11) of various length with various probability distributions. We generate these strings using the *MATLAB* [26] function *randsample()* that uses the pseudo random number generator *mt19937ar* with a long period of $2^{19937} - 1$. It is based on the Mersenne Twister [27], with zigurat [28] as the algorithm that generates the required probability distribution. The complexity of this (deterministic) source of pseudorandom numbers should be high enough to not be captured as algorithmic.

The first part of this characterization involves establishing the minimum string length required for the compression algorithms to perform consistently. We start by generating binary strings, x , with equal probability of 1’s and 0’s, i.e. random strings, of varying length. For each x , we evaluate the compression overhead Q as

$$Q = \frac{C(x) - H(x)}{l(x)}. \quad [10]$$

For a good compressor, we expect Q to be close to 0. From figure 5, it can be seen that for all the compressors, Q starts to converge after about 10^5 bits, setting the minimum string length required for the compressors to work consistently. The *lzw* compressor fails this test, converging to a Q of 0.37 for long string, while *bzip2*, *gzip*, and *lzma* give a Q below 10^{-1} .

In the second part of this characterization, test the compressors with strings with a known amount of correlation. We generate a random string x of length 10^7 using the same technique already described. We then generate a second string y of equal length and with probability p of being correlated to x . For $p = 0$ the two strings are equal, i.e. perfectly correlated. For $p = 0.5$ they are uncorrelated.

The two strings x and y are then combined to form the string xy : to avoid artifacts due to the limited data block size of the compression algorithms, the elements of x and y are interleaved. We then compress xy and evaluate the compression overhead Q as a function of p . The results for different compressors are shown in figure 6. Although there are ranges of p where *bzip2* and *gzip* perform better than *lzma*, the latter shows a more uniform performance over the entire interval of p . It is reasonable to assume that the use of *lzma* should reduce the possibility of artifacts in the estimation of the NCD also for the data obtained from the experiment.

1. Coveney PV, Highfield R (1996) *Frontiers of Complexity: The Search for Order in a Chaotic World*. (Fawcett Columbine).
2. Wolfram S (1985) Undecidability and intractability in theoretical physics. *Phys. Rev. Lett.* 54(8):735–738.
3. Moore C (1990) Unpredictability and undecidability in dynamical systems. *Phys. Rev. Lett.* 64(20):2354–2357.
4. Feynman RP (1982) Simulating physics with computers. *Int. J. Theor. Phys.* 21(6-7):467–488.
5. Shor PW (1994) Algorithms for quantum computation: discrete logarithms and factoring. (IEEE, Santa Fe, NM), pp. 124–134.
6. Grover LK (1996) A fast quantum mechanical algorithm for database search. p. 212.
7. Wolfram S (2001) *New Kind of Science?* (Wolfram Media, Inc.).
8. Deutsch D (1985) Quantum theory, the church-turing principle and the universal quantum computer. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 400(1818):97–117.
9. Lloyd S (1996) Universal Quantum Simulators. *Science* 273(5278):1073–1078.
10. Li M, Chen X, Li X, Ma B, Vitányi PMB (2004) The similarity metric. *Information Theory, IEEE Transactions on* 50(12):3250–3264.
11. Cilibrasi R, Vitányi PMB (2005) Clustering by compression. *Information Theory, IEEE Transactions on* 51(4):1523–1545.
12. Li M, Vitányi PMB (2009) *An Introduction to Kolmogorov Complexity and Its Applications*, Texts in Computer Science. (Springer New York).
13. Cover TM, Thomas JA (2006) *Elements of Information Theory*. (Wiley-Interscience), 2nd edition.
14. Pavlov I (year?) (<http://www.7-zip.org/sdk.html>).
15. Bell JS (1964) On the einstein-podolsky-rosen paradox. *Physics* 1(3):195–200.
16. Braunstein SL, Caves CM (1988) Information-theoretic bell inequalities. *Phys. Rev. Lett.* 61(6):662–665.
17. Shannon CE (1948) A Mathematical Theory of Communication. *Bell System Technical Journal* 27(3):379–423.
18. Kwiat PG et al. (1995) New high-intensity source of polarization-entangled photon pairs. *Phys. Rev. Lett.* 75(24):4337–4341.
19. Kurtsiefer C, Oberparleiter M, Weinfurter H (2001) High-efficiency entangled photon pair collection in type-II parametric fluorescence. *Phys. Rev. A* 64(2):023802.
20. Clauser JF, Horne M, Shimony A, Holt R (1969) Proposed Experiment to Test Local Hidden-Variable Theories. *Phys. Rev. Lett.* 23(15):880–884.
21. Papadimitriou CH (1994) *Computational Complexity*, Theoretical computer science. (Addison-Wesley).
22. Mueller M (year?) Ph.D. thesis (Technical University of Berlin, Berlin).
23. Seward J (year?) (<http://www.bzip.org/>).
24. Gailly JL, Adler M (year?) (<http://www.gzip.org/>).
25. Welch T (1984) A technique for high-performance data compression. *Computer* 17(6):8–19.
26. (year?) (*MATLAB R2010a*, The MathWorks, Inc., Natick, Massachusetts, United States.).
27. Matsumoto M, Nishimura T (1998) Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 8(1):3–30.

28. Marsaglia G, Tsang WW (2000) The ziggurat method for generating random variables. *Journal of Statistical Software* 5(8):1–7.

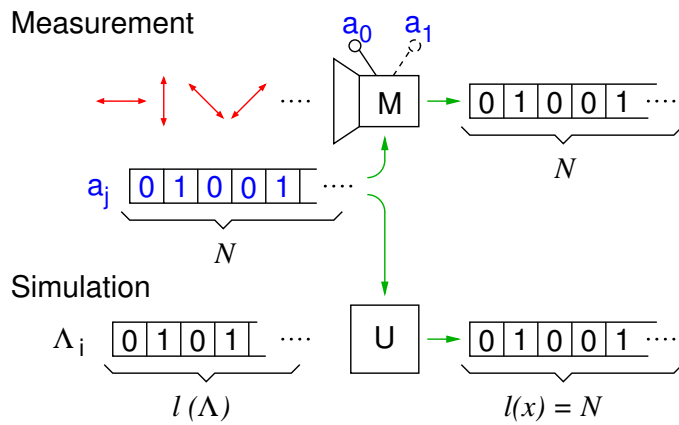


Fig. 1. Measurement: N particles enter a measuring device, characterized by two possible polarizer settings a_0 and a_1 , which generates a bit string of N measurement outcomes. Simulation: a universal Turing machine is fed with a program elements Λ_i and information about the chosen setting a_0 or a_1 . It delivers an output string of length N .

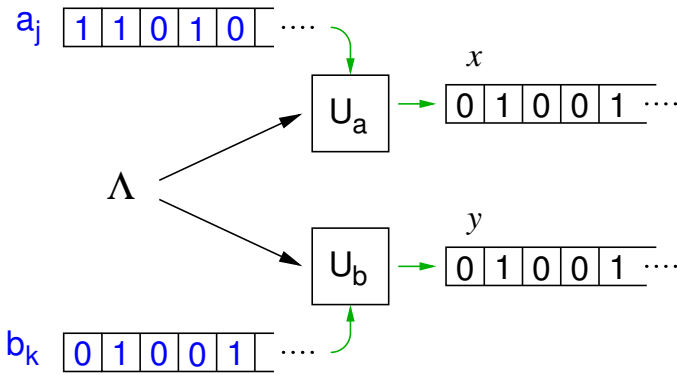


Fig. 2. Local classical machines simulating the generation of strings x and y by correlation measurements on an entangled state.

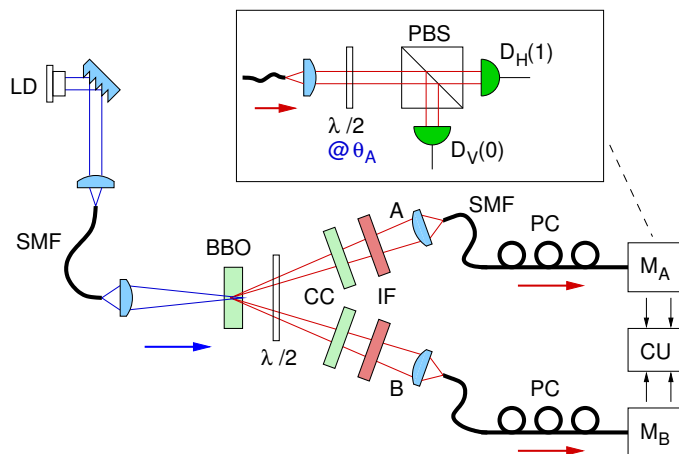


Fig. 3. Schematic of the experimental set-up. Polarization correlations of entangled-photon pairs are measured by the polarization analyzers M_A and M_B , each consisting of a half wave plate ($\lambda/2$) followed by a polarization beam splitter (PBS). All photons are detected by Avalanche photodetectors D_H and D_V , and registered in a coincidence unit (CU).

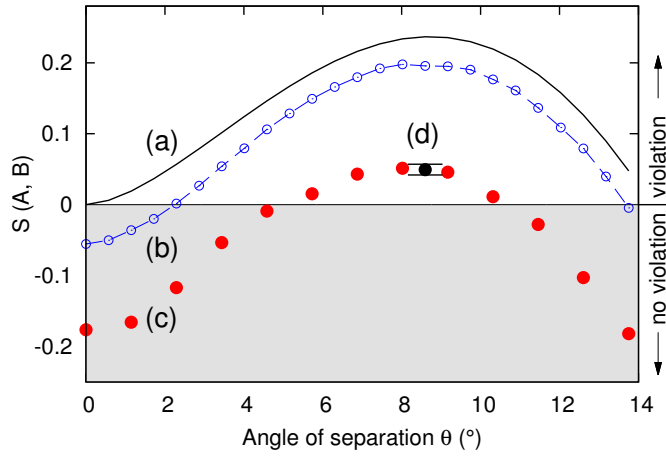


Fig. 4. Plots of S versus angle of separation θ . (a) Result obtained from equation 7 (b) result obtained from using the LZMA compressor on a simulated data ensemble, (c) measurement of S in the experiment shown in figure 3, and (d) longer measurement at the optimal angle $\theta = 8.6^\circ$.

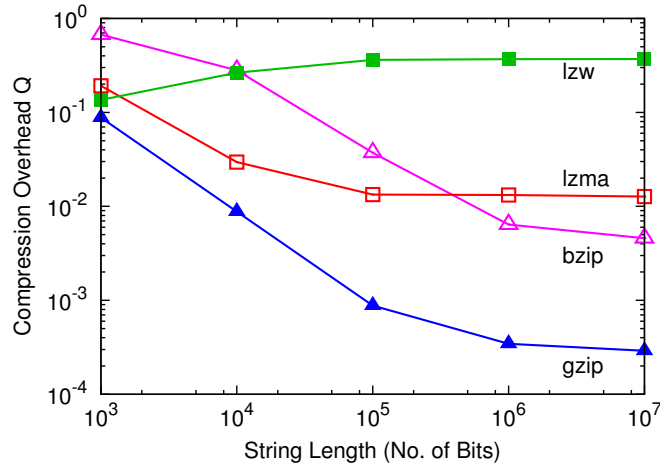


Fig. 5. Comparison of the compression overhead Q obtained using four different compression algorithms on pseudo-random strings of varying lengths. The expected value for an ideal compressor is 0. From this characterization we can exclude *lzw* as a useful compressor for our application.

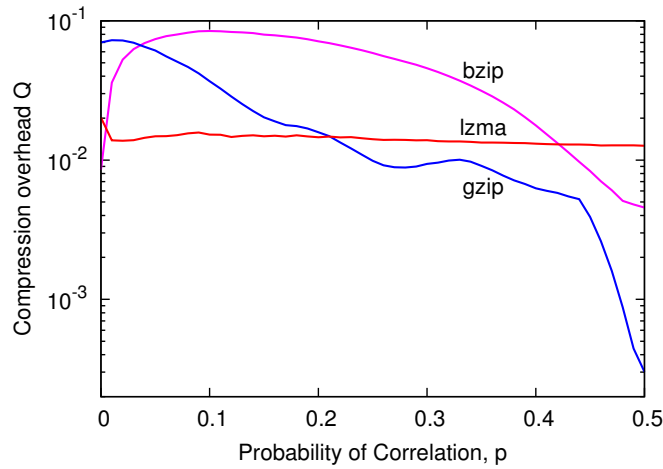


Fig. 6. Compression overhead Q for the string xy as a function of the probability of pairwise correlation p between the bits of the generating strings x and y for three different compressors: *bzip*, *gzip*, and *lzma*.