Probing the quantum-classical boundary with compression software

Hou Shun Poh,[1] Marcin Markiewicz,[2,3] Paweł Kurzyński,[1,4]
Alessandro Cerè,[1] Dagomir Kaszlikowski,[1,5] and Christian Kurtsiefer[1,5]

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

We experimentally demonstrate an impossibility to reproduce quantum bipartite correlations with a deterministic universal Turing machine. We use 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 two-photon state. The violation is shown using a freely available lossless compressor. The presented technique may allow to complement the common statistical interpretation of quantum physics by an algorithmic one.

PACS numbers: 03.67.-a, 03.65.Ta, 42.50.Dv, 89.20.Ff

Introduction. The idea that physical processes can be considered as computations done on some universal machines traces back to Turing and von Neumann [1]. 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]. There are physical phenomena that cannot be traced with analytical tools, which further motivated a computational approach to physics [3]. Moreover, the idea of quantum computation [4] led to a discovery of a few problems efficiently traceable on quantum computers but not on classical ones [5, 6].

The question arises if the complexity of the output of a system can be used as a signature of its non-classicality. In this paper we show that there are processes which cannot be reproduced on local universal Turing machines (UTM) at all, independently of the available classical resources, following a similar approach by Fuchs [7]. We first revisit the concept of Kolmogorov complexity, a measure of the classical complexity of a phenomenon, and later apply it to derive a bound on classical descriptions [8]. Next, we use the fact that Kolmogorov complexity can be approximated by compression algorithms [9]. We then compress experimental data obtained from polarization measurements on entangled photon pairs and show a violation of the classical bound.

Kolmogorov complexity. Consider the description of a machine, whether classical or quantum, that outputs a string x of 0's and 1's. In the case of a UTM, 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 comparable to the length of the output $l(x)$ then our algorithmic description of x is inefficient, and x is called algorithmically random [10]. In most cases $K(x)$ is uncomputable [11]. To circumvent this issue, we can estimate $K(x)$ with some efficient lossless compression $C(x)$ [9].

Bipartite systems. We now extend this picture to two spatially separated UTM's U_A (Alice) and U_B (Bob). 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 strings using the *Normalized Information Distance* (NID) [8]. This distance compares two data sets without detailed knowledge about their origin. In practice, we evaluate an approximation to the NID, the *Normalized Compression Distance* (NCD) [9], using a lossless compression software, in our case the LZMA Utilities, based on the Lempel-Ziv-Markov chain algorithm [12].

We consider a model experiment, similar to the Bell test [13]: a source emits pairs of photons traveling to two separate polarization analyzers M_A (Alice) and M_B (Bob). Each analyzer has two outputs labeled 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 analyzers' outputs are bit strings (see figure 1).

The output x of each analyzer can be described as the output of a UTM, fed with the settings a_j or b_k, and a program Λ, which contains the information about generating the correct output for every detection event and for every setting. For a string of finite length $l(x) = N$, Λ has to describe the 4^N possible events. The length of the shortest Λ is the Kolmogorov complexity of the generated

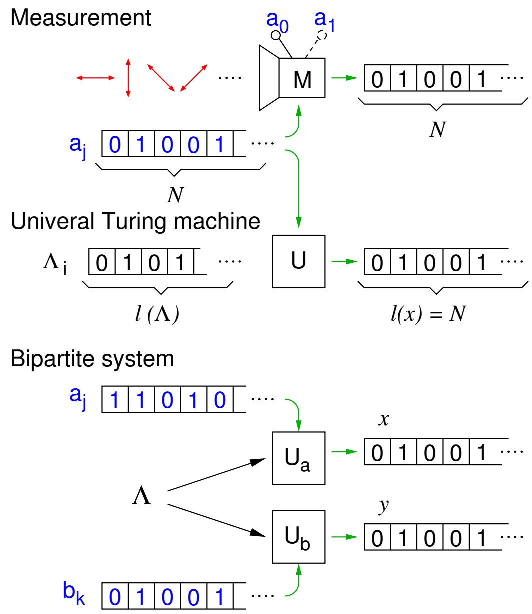

FIG. 1. Measurement: N particles enter a measuring device characterized by two polarizer settings a_0 and a_1 generating N-outcome bit strings. A Universal Turing machine (UTM) fed with a program Λ_i and information about the settings a_0 or a_1 can reproduce the string of length N. The bottom part shows a model to reproduce correlated strings x and y generated from measurements on a bipartite system with local UTMs and a common program Λ.

string. Next, we describe the output of the experiment as the output of two local non-communicating UTMs U_A and U_B. We feed Λ to both of them and obtain two output strings, x and y, both of length N. The program has to describe the behavior of all $2N$ events for all possible settings a_j and b_k, hence 16^N possibilities.

The Kolmogorov complexity $K(x,y)$ of two bit strings 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.

A distance measure between x and y called *Normalized Information Distance* (NID) was introduced in [8]:

$$\mathrm{NID}(x,y) \,=\, \frac{K(x,y) \,-\, \min\{K(x),K(y)\}}{\max\{K(x),K(y)\}}\,. \qquad (1)$$

The NID is a metric and thus obeys the triangle inequality

$$\mathrm{NID}(x,y) \,+\, \mathrm{NID}(y,z) \,\geq\, \mathrm{NID}(x,z)\,. \qquad (2)$$

It holds up to a correction of order $\log(l(x))$, which can be neglected for sufficiently long strings [8].

We consider bit strings x_{a_j} and y_{b_k} generated by Alice and Bob with fixed settings a_j and b_k. Inequality (2) transforms to

$$\mathrm{NID}(x_{a_0},y_{b_0}) + \mathrm{NID}(y_{b_0},y_{b_1}) \geq \mathrm{NID}(x_{a_0},y_{b_1})\,. \qquad (3)$$

However, $\mathrm{NID}(y_{b_0},y_{b_1})$ 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 $\mathrm{NID}(x_{a_1},y_{b_0}) + \mathrm{NID}(x_{a_1},y_{b_1}) \geq \mathrm{NID}(y_{b_0},y_{b_1})$, and combine it with (3) to get:

$$\mathrm{NID}(x_{a_0},y_{b_0}) + \mathrm{NID}(x_{a_1},y_{b_0}) + \mathrm{NID}(x_{a_1},y_{b_1}) \geq \mathrm{NID}(x_{a_0},y_{b_1})\,. \qquad (4)$$

We introduce a parameter S' quantifying the degree of violation of Eq. (4):

$$\begin{aligned} S' &= \mathrm{NID}(x_{a_0},y_{b_1}) - \mathrm{NID}(x_{a_0},y_{b_0}) \\ &\quad - \mathrm{NID}(x_{a_1},y_{b_0}) \,-\, \mathrm{NID}(x_{a_1},y_{b_1}) \leq 0 \end{aligned} \qquad (5)$$

To test this inequality, we have to address the following problem. We can set up a source to generate entangled photon pairs in an arbitrary state but 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 stated as: the same preparation results in the same complexity of observed phenomena.

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

The statistical approach uses an ensemble of all possible N-bit strings and looks for their average Kolmogorov complexity. The ensemble average is the Shannon entropy $H(X)$ [11] and thus

$$\langle \mathrm{NID}(x,y)\rangle = \frac{H(x,y) - \min\{H(x),H(y)\}}{\max\{H(x),H(y)\}}\,. \qquad (6)$$

Inequality (4) becomes an entropic Bell inequality by Braunstein and Caves [14] 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 such that $\vec{a_0}\cdot\vec{b_1} = \cos 3\theta$, $\vec{a_0}\cdot\vec{b_0} = \vec{a_1}\cdot\vec{b_0} = \vec{a_1}\cdot\vec{b_1} = \cos\theta$, inequality (4) is violated for an appropriate range of θ. An expected quantum value of S' as a function of θ is shown in Fig. 2a. The maximal violation is $S' = 0.24$ for $\theta = 8.6°$. This statistical approach requires the assumption that output of the systems are *independent identically distributed* (i.i.d.).

On the other hand, it is possible to avoid a statistical description of our experiment following Ref. [9] where it was shown that the Kolmogorov complexity can be

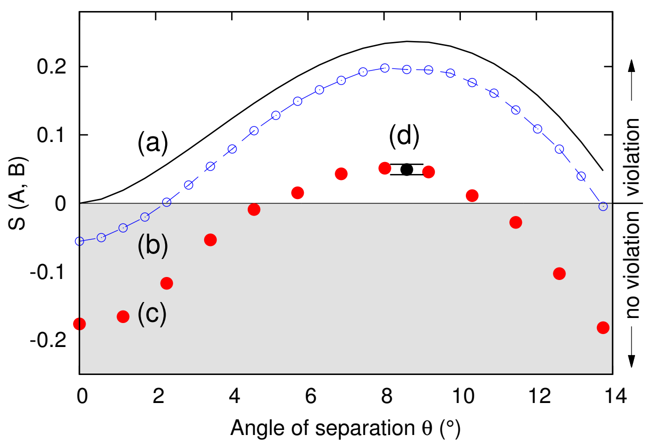

FIG. 2. Plots of S versus angle of separation θ. (a) Result obtained from Eq. (6), (b) result obtained from using the LZMA compressor on numerically generated data, (c) measurement of S in the experiment shown in figure 3, and (d) longer measurement at the optimal angle $\theta = 8.6°$.

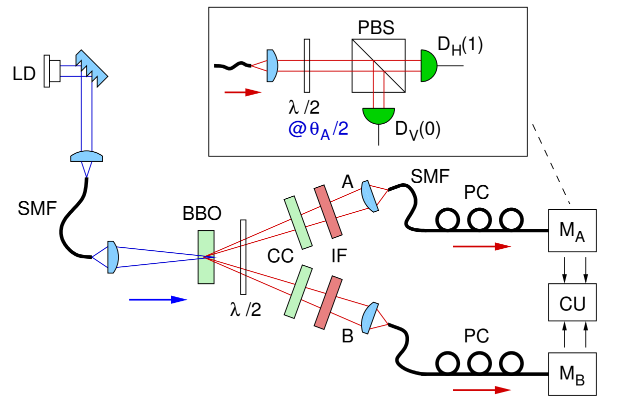

FIG. 3. Schematic of the experimental set-up. Polarization correlations of entangled-photon pairs are measured by the polarization analyzers M_A and M_B, each consisting of a half wave plate ($\lambda/2$) followed by a polarization beam splitter (PBS). All photons are detected by Avalanche photodetectors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

well approximated by compression algorithms. In [9] the *Normalized Compression Distance* (NCD) is introduced

$$\mathrm{NCD}(x,y) = \frac{C(x,y) - \min\{C(x),C(y)\}}{\max\{C(x),C(y)\}}\,, \qquad (7)$$

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 Eq. (5) leads to a new inequality:

$$\begin{aligned} S' \to S = \mathrm{NCD}(x_{a_0},y_{b_1}) - \mathrm{NCD}(x_{a_0},y_{b_0}) \\ - \mathrm{NCD}(x_{a_1},y_{b_0}) - \mathrm{NCD}(x_{a_1},y_{b_0}) \leq 0\,. \quad (8) \end{aligned}$$

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

Most compression algorithms use some prediction about the data composition. If it matches this prediction, the compression can be done efficiently. To conduct an 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 [12] approaches the Shannon limit [15] most closely (see the Supplementary materials).

In general our method can be used for data from any source by finding a suitable compression algorithm [9]. Thus, we are not limited to i.i.d. sources, as it is commonly assumed in standard statistical ensemble-based experiments, like, for instance, Bell-type tests.

The numerical simulation also verifies the angle that maximizes the violation of (8). The results of this simulation are presented in figure 2. More details on the generation of the simulated data and the choice of the compressor are provided in the Supplementary materials.

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 μm into a 2 mm thick BBO crystal cut for type-II phase-matching. There, 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 [16]. Two spatial modes (A, B) of down-converted light, defined by the SMFs for 810 nm, are matched to the pump mode to optimize the collection [17]. 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 $|\psi\rangle = 1/\sqrt{2}\left(|H\rangle_A|V\rangle_B + e^{i\phi}|V\rangle_A|H\rangle_B\right)$ is adjusted to $\phi = \pi$ by tilting the CC.

In the polarization analyzers (figure 3), 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 in each analyzer. Photons are detected by avalanche photo diodes (APD), and corresponding detection events from the same pair identified by a coincidence unit if they arrive within $\approx \pm 3$ 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 $V_{45} = 99.9\pm0.1\%$. The visibility in the natural H/V basis of the type-II down-conversion process also reaches $V_{\rm HV} = 99.9\pm0.1\%$. A separate test of a CHSH-type Bell inequality [18] leads to a value of $S = 2.826 \pm 0.0015$. This indicates a high quality of polarization entanglement; the uncertainties in the visibilities are obtained from propagated Poissonian counting statistics.

We record 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 Eq. (7). 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 Supplementary materials for details).

The inequality is experimentally tested by evaluating S in Eq. (8) for a range of θ; the results [points (c), (d) in figure 2] are consistently lower than the trace (a) calculated via entropy using Eq. (6), 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 $\theta = 8.6°$ 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 Eq. (8). 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 these files are then averaged to obtain the final result of $S(\theta = 8.6°) = 0.0494 \pm 0.0076$, 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 show that this is not possible if one uses a computation paradigm of a local UTM. 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 [13] that is statistical in its nature.

The Kolmogorov complexity of the output of local UTM must obey distance properties as shown in [8, 9], and can be approximated by compression. The distance properties lead to inequality Eq. (8), which we find violated in the specific case of polarization-entangled photon pairs. Therefore, at least this physical processes can not be encoded as programs on local UTMs.

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. In addition, the algorithmic approach does not use the notion of an ensemble and the i.i.d. assumption. The compression treats the string of data as a single entity, and does not ignore correlations between subsequent string elements. Our approach allows us therefore to omit the notion of probability, at least for the case at hand. If it can be extended to other quantum experiments, it would offer an alternative with less assumptions to the commonly used statistical interpretation of quantum theory.

We have become aware of a recent work by Wolf [19] inspired by the ideas presented in this work, where this algorithmic approach is used to provide a different viewpoint on nonlocality that does not require counterfactual reasoning.

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.

[1] P. V. Coveney and R. Highfield, *Frontiers of Complexity: The Search for Order in a Chaotic World* (Fawcett Columbine, 1996).
[2] S. Wolfram, Phys. Rev. Lett. **54**, 735 (1985).
[3] C. Moore, Phys. Rev. Lett. **64**, 2354 (1990).
[4] R. P. Feynman, Int. J. Theor. Phys. **21**, 467 (1982).
[5] P. W. Shor, in *Proceedings of the 35th Annual Symposium on Foundations of Computer Science* (IEEE, Santa Fe, NM, 1994) pp. 124–134.
[6] L. K. Grover, in *Proceedings, 28th Annual ACM Symposium on the Theory of Computing* (1996) p. 212.
[7] C. A. Fuchs, in *Workshop on Squeezed States and Uncertainty Relations*, NASA Conference Publication 3135, edited by D. Han, Y. S. Kim, and W. W. Zachary (NASA, Washington, DC, 1992) pp. 83–85.
[8] M. Li, X. Chen, X. Li, B. Ma, and P. M. B. Vitányi, Information Theory, IEEE Transactions on **50**, 3250 (2004).
[9] R. Cilibrasi and P. M. B. Vitányi, Information Theory, IEEE Transactions on **51**, 1523 (2005).
[10] M. Li and P. M. B. Vitanyi, *An Introduction to Kolmogorov Complexity and Its Applications*, Texts in Computer Science (Springer New York, 2009).
[11] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. (Wiley-Interscience, 2006).
[12] I. Pavlov, http://www.7-zip.org/sdk.html.
[13] J. S. Bell, Physics **1**, 195 (1964).

[14] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **61**, 662 (1988).

[15] C. E. Shannon, Bell System Technical Journal **27**, 379 (1948).

[16] P. G. Kwiat, K. Mattle, H. Weinfurter, A. Zeilinger, A. V. Sergienko, and Y. Shih, Phys. Rev. Lett. **75**, 4337 (1995).

[17] C. Kurtsiefer, M. Oberparleiter, and H. Weinfurter, Phys. Rev. A **64**, 023802 (2001).

[18] J. F. Clauser, M. Horne, A. Shimony, and R. Holt, Phys. Rev. Lett. **23**, 880 (1969).

[19] S. Wolf, arXiv (2015), 1505.07037.