Quantum Kolmogorov Complexity

André Berthiaume
School of CTI, 243 South Wabash Avenue, DePaul University
Chicago IL, 60604-2301, United States of America
E-mail: berthiaume@cs.depaul.edu

and

Wim van Dam
Computer Science Division, Soda Hall, University of California
Berkeley CA, 94720-1776, United States of America
E-mail: vandam@cs.berkeley.edu

and

Sophie Laplante
L.R.I., Université Paris-Sud
Bâtiment 490, 91405 Orsay, France
E-mail: Sophie.Laplante@lri.fr

In this paper we give a definition for quantum Kolmogorov complexity. In the classical setting, the Kolmogorov complexity of a string is the length of the shortest program that can produce this string as its output. It is a measure of the amount of innate randomness (or information) contained in the string.

We define the quantum Kolmogorov complexity of a qubit string as the length of the shortest *quantum* input to a universal quantum Turing machine that produces the initial qubit string with high fidelity. The definition of Paul Vitányi [25] measures the amount of classical information, whereas we consider the amount of *quantum* information in a qubit string. We argue that our definition is a natural and accurate representation of the amount of quantum information contained in a quantum state. Recently, Peter Gács [8] has also proposed two measures of 'quantum algorithmic entropy', which are based on the existence of a universal semi-density matrix. The latter definitions are related to Vitányi's and the one presented in this article, respectively.

1. INTRODUCTION

In classical computations, the Kolmogorov-Solomonoff-Chaitin (Kolmogorov, for short) complexity of a finite string is a measure of its randomness.[4, 14, 22] The Kolmogorov complexity of x is the length of the shortest program that produces x as its output. It can be seen as a lower bound on the optimal compression that x can undergo, and it is closely related to Shannon's information theory.[5, 21] Kolmogorov complexity has been shown to have a windfall of applications in fields as diverse as learning theory, complexity theory, combinatorics and graph theory, analysis of algorithms, to name just a few.

With the advent of quantum computation, it is natural to ask how to define the Kolmogorov complexity of qubit strings. The goal of this paper is to argue that the definition presented here is a natural and robust measure of the amount of quantum information contained in a qubit string, and that it has several appealing properties.

Finding such a robust definition for quantum Kolmogorov complexity has been of interest for many years (see for example the 1996 article [23]). More recently, Paul Vitányi [25] has also proposed a definition for quantum algorithmic complexity. Our definition differs fundamentally from Vitányi's for his definition measures the amount of *classical* information necessary to approximate the quantum state. Also Peter Gács [8] has discussed two definitions 'quantum algorithmic entropy', both of which are based on the notion of a universal semi-density matrix. One of Gács' definitions is close to ours, while the other is close to Vitányi's.

1.1. What is a good definition?

A good definition of quantum Kolmogorov complexity should meet the following fundamental criteria. These are intended to insure that it gives an accurate representation of the information content of a quantum string.

- It should be robust, that is, invariant under the choice of the underlying quantum Turing machine.
- It should bear a strong relationship with quantum information theory.
- For classical strings, it should be closely related to traditional Kolmogorov complexity complexity.

However, quantum Kolmogorov complexity should not be expected to always behave the way classical Kolmogorov complexity does. The reader may want to bear in mind quantum mechanical phenomena such as the no-cloning theorem[28], whose consequences we will discuss in Section 8.

A first attempt at defining quantum Kolmogorov complexity of a qubit string X is to consider the length of the shortest quantum program that produces X as its output. There are many questions that arise from this 'definition'.

Bits or qubits? The first question to consider is whether we want to measure the amount of algorithmic information of a string in bits, or in qubits. Note that the set of bit strings (programs) is countable, whereas for qubit strings this set is uncountable. Hence, any definition that measures in bits would have to overcome this apparent contradiction. Paul Vitányi [25] considers classical descriptions of qubit strings, whereas we consider qubit descriptions.

Exact or inexact? What does 'produce' mean? Is a minimal program required to produce the string X exactly, or only up to some fidelity? In the latter case, is the fidelity a constant? Otherwise, how is it parameterized? (For exact simulation, we can only hope to simulate a subclass of all the possible quantum Turing machines, say by restricting the set of their possible amplitudes. What would be a reasonable choice in such a scenario?) In this article we will use an approximation scheme.

What model of computation? The size of quantum circuits is not an appropriate measure since large circuits may be very simple to describe. The Turing machine model is the appropriate one to consider.

What is meant by 'quantum program?' A program is the input for a universal quantum Turing machine. If we want to count the size of a description in qubits, then we must allow for 'programs' to be arbitrary qubit strings. (These can be viewed as programs whose code may include some auxiliary 'hard-coded' qubit strings.)

One-time description or multiple generation? In the classical setting, the program that prints the string x can be run as many times as desired. Because of the no-cloning theorem of quantum physics however, we can no longer assume this property for our quantum descriptions. In general this will be due to the fact that it is not possible to recover the program without losing its output. This is closely related to another reason not to choose the multiple generation option. The complex-valued parameters α and β of a qubit $|q\rangle = \alpha|0\rangle + \beta|1\rangle$ can contain an unbounded amount of information. If we were able to reproduce q over and over again, then we would have to conclude that the single qubit q contains an unlimited amount of information. This contradicts the fact that the quantum mechanical system of q can only contain one bit of information.[10] For the above two reasons, we will not require a 'reusability' condition.

1.2. Organization

The paper will be organized as follows. In the Sections 2, 3, and 4 we give basic notation, definitions, prior work and some well-known theorems and lemmas that will be used in the paper. Our definition of quantum Kolmogorov complexity is given in Section 5. The invariance theorem for this definition is then proven in Section 6. Section 7 compares the properties of quantum and classical Kolmogorov complexity, including incompressibility and subadditivity. We give some typical quantum mechanical results on the complexity of copies in Section 8. Section 9 discusses the relationship with quantum information theory. Quantum Kolmogorov complexity gives us a way to express the amount of correlation in a bipartite system, which we briefly discuss in Section 10. We conclude with a discussion of possible extensions and future work.

2. PRELIMINARIES

In this section we fix our notation and definitions, and mention some results that will be used to prove the results in this paper.

2.1. Notation

We use x, y, $\ldots$ to denote finite, classical Boolean strings. When we write $|x\rangle$, we mean the quantum state vector in the standard basis that corresponds to the classical string x. In general we use $\phi, \psi, \dots$ to denote pure quantum states, while mixed states are represented by the letters ρ, σ, et cetera. We also use uppercase letters X, Y, ... for (mixed) quantum states that are strings of qubits. The terms quantum state, qubit string, and quantum register are used interchangeably (sometimes to emphasize the purpose of the quantum state at hand). Lower-case letters i, j, k, l, m, n denote integer indices (typically string lengths).

For classical strings over the alphabet $\{0,1\}$, $\ell(x)$ denotes the length of the string. For finite sets $A, |A|$ denotes the cardinality of the set. Concatenation of x, y is written as the juxtaposition xy, and the n-fold concatenation of x is written $x^{\otimes n}$.

For Hilbert spaces, we write $\mathcal{H}_d$ for the d-dimensional Hilbert space and $\mathcal{H}^{\otimes n}$ for the n-fold tensor product space $\mathcal{H} \otimes \cdots \otimes \mathcal{H}$. Similarly, we write $U^{\otimes n}$ for the n-fold fold tensor of the operation U, and $\psi^{\otimes n}$ for n copies of the state ψ. A pure quantum state ψ represented as a vector in such a Hilbert space is denoted by the ket $|\psi\rangle$. The *fidelity* between two pure states ψ and ϕ is the absolute value of the inner product of the two vectors: $|\langle\psi|\phi\rangle|$ (although some authors use the square of this value).

We slightly abuse notation by sometimes letting the state symbols $\phi, \rho, \dots$ also stand for the corresponding density matrices. Hence, a pure state ϕ as a Hilbert space vector is denoted by $|\phi\rangle$, whereas its density matrix $|\phi\rangle\langle\phi|$ can also be denoted by ϕ.

A density matrix can always be decomposed as a mixture of pure, orthogonal states: $\rho = \sum_i p_i |\phi_i\rangle\langle\phi_i|$, with $p_1, p_2, \dots$ a probability distribution over the mutually orthogonal states $\phi_1, \phi_2, \dots$ The matrix ρ represents a pure state if and only if $\rho^2 = \rho$, in which case we can also say $\sqrt{\rho} = \rho$. The square root of a general mixed state is thus described by

$$\sqrt{\rho} \;=\; \sqrt{\sum_i p_i|\phi_i\rangle\langle\phi_i|} \;=\; \sum_i \sqrt{p_i}|\phi_i\rangle\langle\phi_i|.$$

We use the above rule for the generalization of the fidelity to mixed states. The fidelity between two density matrices ρ and σ is defined by

$$\text{Fidelity}(\rho,\sigma) \;=\; \operatorname{tr}\left(\sqrt{\sqrt{\rho}\cdot\sigma\cdot\sqrt{\rho}}\right). \tag{1}$$

For pure states ϕ and ψ, the above definition coincides again with the familiar $|\langle\phi|\psi\rangle|$. If $\text{Fidelity}(\rho,\sigma) = 1$, then $\rho = \sigma$, and vice versa.

An ensemble $\mathcal{E}$ is a specific distribution $p_1, p_2, \dots$ over a set of (mixed) states $\rho_1, \rho_2, \dots$. We denote this by $\mathcal{E} = \{(\rho_i, p_i)\}$. The average state of such an ensemble $\mathcal{E}$ is $\rho = \sum_i p_i \rho_i$. It is important to realize that an average state can correspond to several different ensembles. When an ensemble is used to produce a sequence of states ρ_i according to the probabilities p_i, we speak of a *source* $\mathcal{E}$.

The length of a quantum state is denoted by $\ell(X)$, by which we mean the smallest l for which X sits in the 2^l-dimensional Hilbert space (in the standard basis).

A transformation $\mathbf{S}$ on the space of density matrices is allowed by the laws of quantum mechanics if and only if it is a completely positive, trace preserving mapping.

2.2. Classical Kolmogorov complexity

The Kolmogorov complexity of a string, in the classical setting, is the length of the shortest program that prints this string on an empty input.[15]

Formally, this is stated first relative to a partial computable function, which as we know can be computed by a Turing machine.

DEFINITION 2.1. (Classical Kolmogorov complexity) Fix a Turing machine T that computes the partial computable function Φ. For any pair of strings $x, y \in \{0,1\}^*$, the Kolmogorov complexity C of x relative to y (with respect to Φ) is defined as

$$C_\Phi(x|y) \;=\; \text{Min}\{\ell(p) : \Phi(p,y) = x\}.$$

When y is the empty string, we simply write $C_\Phi(x)$. Also the notation $C_T(x|y)$ is used.

The key theorem on which rests the robustness of Kolmogorov complexity is the *invariance theorem.* This result states that the length of shortest programs does not depend by more than an additive constant on the underlying Turing machine. In the classical case, this theorem is proven with the existence of a universal Turing machine. This machine has two inputs: a finite description of the original Turing machine, and the program that this Turing machine executes to output the string. More formally, the invariance theorem in the classical case can be stated as follows.

THEOREM 2.1. *(Classical invariance theorem) There is a universal partial computable function Φ_0 such that for any partial computable Φ there is a constant c_Φ with*

$$C_{\Phi_0}(x|y) \;\leq\; C_\Phi(x|y) + c,$$

for all strings x and y.

Giving an invariance theorem in the quantum mechanical case will be key to showing that our definition of quantum Kolmogorov complexity is robust.

Since for any string x of length n, $C(x) \leq n + O(1)$, a string which has complexity at least n is called *incompressible.* The existence of incompressible strings is a crucial fact of Kolmogorov complexity.

PROPOSITION 2.1. *(Classical incompressibility) For every string length n, there is a string x of length n such that $C(x) \geq n$.*

The proof that there exists incompressible strings is a simple application of the pigeonhole principle. By comparing the number of strings of length n (2^n) and the number of programs of length smaller than n ($2^n - 1$ in total), one must conclude that there is at least one string of length n which is not the output of any of the program of length $< n$.

2.3. Entropy of classical sources

The Shannon entropy of a random source that emits symbols from an alphabet is a measure of the amount of randomness in the source.[5, 21]

DEFINITION 2.2. (Shannon entropy) Let A be a random source $\{(x_i, p_i)\}$, which emits letter x_i (independently) with probability p_i. The Shannon entropy H of A is $H(A) = -\sum_i p_i \log p_i$.

In the classical setting, Kolmogorov complexity and Shannon entropy are closely related, as we describe now. This is an important property of Kolmogorov complexity, and one would expect a similarly strong relationship to hold between quantum Kolmogorov complexity and quantum entropy.

Shannon's noiseless coding theorem states that the entropy corresponds to the average number of bits required on average to encode sequences of character emitted by a random source.

PROPOSITION 2.2. *(Shannon's noiseless coding theorem [21]) Consider a classical channel A that is used to transmit letters taken from an ensemble $\{(x_i, p_i)\}$, where the x_i are the letters and p_i the corresponding probabilities. Then, the Shannon entropy $H(A)$ gives the following bounds.*

1. *For any $\epsilon, \delta > 0$, there is an n such that there is an encoding that on n letters encodes on average the letters with $H(A) + \delta$ bits, for which the probability of successfully decoding is at least $1 - \epsilon$.*
2. *For any ϵ, δ, there is an n such that for any δ', there is an $\epsilon' < \epsilon$ such that if the channel encodes n letters with less than $n(H(A) - \delta')$ bits, then the probability of success is no bigger than $2^{-n(\delta'-\delta)} + \epsilon'$.*

In the classical case, the Kolmogorov complexity of a string is bounded by the entropy of a source 'likely to have emitted this string'. A brief summary of the argument is included here. (Details can be found in [15, page 180].)

Let x be a (long) binary string. It can be broken down into m blocks of length k, where each block is thought of as a character in an alphabet of size 2^k. Define the frequency f_i of a character c_i to be the number of times it appears as a block in x, and let A_m represent the source $\{(c_i, f_i/m)\}$. To reconstruct x, it suffices to provide the frequency of each character ($\sum_i \log f_i$ bits) and then specify x among the strings that share this frequency pattern. With some manipulations, the following can be shown.

PROPOSITION 2.3. *(Correspondence between Shannon entropy and Kolmogorov complexity) For any string x and A_m a corresponding source defined in the discussion above*

$$C(x) \;<\; m(H(A_m) + \gamma),$$

where γ vanishes to zero as m goes to infinity.

2.4. Quantum information theory

We have seen that in the classical setting, Kolmogorov complexity is very closely related to Shannon entropy. In this section we describe the quantum, or Von

Neumann, entropy, related measures, and important properties which will be used in the proofs of our results.

DEFINITION 2.3. (Von Neumann entropy) The Von Neumann entropy of a mixed state ρ is defined as $S(\rho) = \mathrm{tr}(-\rho \log \rho)$. If we decompose ρ into its mutually orthogonal eigenstates ϕ_i with eigenvalues p_i, we see that

$$S(\rho) \;=\; S\left(\sum_i p_i|\phi_i\rangle\langle\phi_i|\right) \;=\; -\sum_i p_i \log p_i,$$

the latter being the Shannon entropy of the probability distribution $p_1, p_2, \ldots$

The entropy of finite systems is robust against small changes. This *insensitivity* of S over the space of finite dimensional density matrices ρ is expressed by the following lemma.

LEMMA 2.1. *(Insensitivity of Von Neumann entropy; see Section II.A in [26]) If a sequence of $d \times d$ dimensional density matrices $\rho_1, \rho_2, \ldots$, has $\lim_{k\to\infty} \rho_k = \rho$, then also $\lim_{k\to\infty} S(\rho_k) = S(\rho)$.*

Proof. The convergence of $\rho_1, \rho_2, \ldots$ to ρ is understood to use some kind of norm for the density matrices that is continuous in the matrix entries $\langle i|\rho|j\rangle$. (The operator norm $|\rho| = \mathrm{tr}(\rho\rho^*)$, for example.) The entropy $S(\rho)$ is a continuous function of the finite set of eigenvalues of ρ. These eigenvalues are also continuous in the entries of ρ. ∎

A source $\mathcal{E} = \{(\rho_i, p_i)\}$ has an associated Von Neumann entropy $S(\rho)$ of the average state $\rho = \sum_i p_i \rho_i$. Schumacher's noiseless coding theorem [20] shows how to obtain an encoding with average letter-length $S(\rho)$ for a source of pure states, where the fidelity of the encoding goes to 1 as the number of letters emitted by the source goes to infinity. (A survey can be found in Preskill's lecture notes [19, page 190], Nielsen's thesis [17, Chapter 7], or the recent book by Nielsen and Chuang [18].)

We will use a slightly stronger result, which gives a universal compression scheme, that is, one that does not depend on the source itself, but only on its entropy. This result is due to Jozsa *et al.* [12], building upon the work of Jozsa and Schumacher [13].

THEOREM 2.2. *(Universal quantum compression [12, 13]) For any $\epsilon, \delta > 0$, there is a block size $n = n(\epsilon, \delta)$ such that for every entropy bound S, there is an encoding scheme for blocks of n states that works for any pure state source $\mathcal{E} = \{(\phi_i, p_i)\}$ with the following properties. Let $\rho = \sum_i p_i|\phi_i\rangle\langle\phi_i|$ be the average state, with all $|\phi_i\rangle \in \mathcal{H}_d$, and ρ has entropy $S(\rho) \le S$, then for the encoding and decoding scheme it holds that:*

1. *Each $|\phi_i\rangle$ can be encoded by a code word σ_i that has length $\ell(\sigma_i) \le S + \delta + \frac{1}{n}(d^2 \log(n+1))$.*
2. *For each i, the fidelity between ϕ_i and the decoding of σ_i is at least $1 - \epsilon$.*

We continue the section by defining the 'χ quantity' for ensembles.

DEFINITION 2.4. (Holevo's chi quantity [10]) For an ensemble $\mathcal{E} = \{(\rho_i, p_i)\}$, with $\rho = \sum_i p_i\rho_i$, Holevo's chi quantity equals

$$\chi(\mathcal{E}) \;=\; S(\rho) - \sum_i p_i S(\rho_i).$$

Note that the χ quantity depends not only on ρ, but also on the specific pairs (ρ_i, p_i).

The following monotonicity property of Lindblad and Uhlmann will be very useful later in the paper.

THEOREM 2.3. *(Lindblad-Uhlmann monotonicity [16, 24]) Let $\mathcal{E} = \{(\rho_i, p_i)\}$ be an ensemble, and $\mathbf{S}$ a completely positive, trace preserving mapping. For every such $\mathcal{E}$ and $\mathbf{S}$, it holds that: $\chi(\mathbf{S}(\mathcal{E})) \leq \chi(\mathcal{E})$, where $\mathbf{S}(\mathcal{E})$ is the transformed ensemble $\{(\mathbf{S}(\rho_i), p_i)\}$.*

Further background on these measures of quantum information and their properties can be found in [19, Chapter 5] and [26]. Another good source is Nielsen's thesis [17].

3. SYMMETRIC SUBSPACES

We use the symmetric subspace of the Hilbert space to prove some of our results on copies of quantum states. Let $\mathcal{H}_d$ be a Hilbert space of dimension d with the basis states labeled $|1\rangle, \ldots, |d\rangle$. The *symmetric subspace* $\mathcal{SYM}(\mathcal{H}_d, m)$ of the m-fold tensor product space $\mathcal{H}_d^{\otimes m}$ contains the states that are invariant under permutation of its m parts. As a consequence, it is a subspace spanned by as many basis vectors as there are multisets of size m of $\{1, \ldots, d\}$. If $A = \{i_1, \ldots, i_m\}$ is such a multiset of $\{1, \ldots, d\}$, then $|A\rangle$ is the normalized superposition of all the different permutations of $i_1, \ldots, i_m$. The set of the different vectors $|A\rangle$ (ranging over the multisets A) is an orthogonal basis of the symmetric subspace $\mathcal{SYM}(\mathcal{H}_d, m)$. This shows that the dimension of the symmetric subspace is $\binom{m+d-1}{d-1}$, because choosing such a multiset is equivalent to splitting a sequence of m zeroes into d (possibly empty) intervals. (If j_i is the size of the of ith interval, then this number also represents that the element $i \in \{1, \ldots, d\}$ appears j_i times in the multiset. The number of ways of splitting a sequence of size m into d intervals is $\binom{m+d-1}{d-1}$.)

The symmetric subspace $\mathcal{SYM}(\mathcal{H}_d, m)$ is the smallest subspace of $\mathcal{H}_d^{\otimes m}$ that contains all the pure states of the form $|\phi\rangle^{\otimes m}$ for all $|\phi\rangle \in \mathcal{H}_d$.

As an example, consider the symmetric subspace $\mathcal{SYM}(\mathcal{H}_2, 3)$. For every qubit $\alpha|0\rangle + \beta|1\rangle$, we can indeed express any three-fold copy in the four dimensions of $\mathcal{SYM}(\mathcal{H}_2, 3)$:

$$\begin{aligned}(\alpha|0\rangle + \beta|1\rangle)^{\otimes 3} \;&=\; \alpha^3|000\rangle \;+\; \alpha^2\beta(|001\rangle + |010\rangle + |100\rangle) \;+\\ &\quad\; \alpha\beta^2(|011\rangle + |101\rangle + |110\rangle) \;+\; \beta^3|111\rangle\\ &=\; \alpha^3|\{0,0,0\}\rangle \;+\; \alpha^2\beta\sqrt{3}|\{0,0,1\}\rangle \;+\\ &\quad\; \alpha\beta^2\sqrt{3}|\{0,1,1\}\rangle \;+\; \beta^3|\{1,1,1\}\rangle.\end{aligned}$$

We thus reach the important conclusion that there exists a unitary transformation from the 3 qubits of the symmetric subspace $\mathcal{SYM}(\mathcal{H}_2, 3)$ to the two qubits of the space spanned by the vectors $|\{0,0,0\}\rangle$, $|\{0,0,1\}\rangle$, $|\{0,1,1\}\rangle$ and $|\{1,1,1\}\rangle$. The generalization of this compression result for all values d and m is presented in Section 8. For more information on the symmetric subspace and its properties, see the paper by Barenco *et al.* [1].

4. ACCUMULATION OF ERRORS

The following lemma is used to bound the error introduced when composing two inexact quantum procedures.

LEMMA 4.1. *(Fidelity of composition) Let ρ_1, ρ_2 and ρ_3 be three density matrices.*

$$\textit{If } \mathrm{Fidelity}(\rho_1, \rho_2) \geq 1 - \delta_1 \quad \textit{and} \quad \mathrm{Fidelity}(\rho_2, \rho_3) \geq 1 - \delta_2,$$

then $\mathrm{Fidelity}(\rho_1, \rho_3) \geq 1 - 2\delta_1 - 2\delta_2$.

Proof. We say that a bi-partite, pure state ϕ^{AB} is the 'purification' of the (mixed) state ρ if we obtain ρ by tracing out the B part of ϕ^{AB}: $\rho = \mathrm{tr}_B(\phi^{AB})$. (For more on this definition, see [7].) The lemma now follows from the fact that the fidelity between two (mixed) states ρ_1 and ρ_2 equals the maximum 'pure state fidelity' $|\langle\phi_1|\phi_2\rangle|$, with ϕ_i the purifications of ρ_i. ∎

This lemma is especially powerful in combination with the basic result that the fidelity between two states cannot decrease under a quantum mechanical transformation. It enables us to prove the following result that bounds the error of two consecutive operations.

LEMMA 4.2. *(Fidelity after two transformations) If U_1 and U_2 are two quantum-mechanical transformations and ρ_1, ρ_2, ρ_3 are density matrices such that*

$$\mathrm{Fidelity}(\rho_2, U_1(\rho_1)) \geq 1 - \delta_1 \quad \text{and} \quad \mathrm{Fidelity}(\rho_3, U_2(\rho_2)) \geq 1 - \delta_2, \tag{2}$$

then, for the combined transformation U_2U_1,

$$\mathrm{Fidelity}(\rho_3, U_2 \cdot U_1(\rho_1)) \geq 1 - 2\delta_1 - 2\delta_2. \tag{3}$$

Proof. ¿From $\mathrm{Fidelity}(\rho_2, U_1(\rho_1)) \geq 1 - \delta_1$, and the nondecreasing property of the fidelity it follows that $\mathrm{Fidelity}(U_2(\rho_2), U_2 \cdot U_1(\rho_1)) \geq 1 - \delta_1$. Lemma 4.1 concludes the proof. ∎

In order to give bounds on the complexity of several copies of a state, as we do in Section 8, we also need the following bound on the total error in the n-fold tensor product of the approximation of a given state.

LEMMA 4.3. *(Fidelity of copies) Let $\rho_1^{\otimes n}$ and $\rho_2^{\otimes n}$ be the n-fold copies of the mixed states ρ_1 and ρ_2, then* $\mathrm{Fidelity}(\rho_1^{\otimes n}, \rho_2^{\otimes n}) = (\mathrm{Fidelity}(\rho_1, \rho_2))^n$. *Hence, if* $\mathrm{Fidelity}(\rho_1, \rho_2) \geq 1 - \delta$, *then* $\mathrm{Fidelity}(\rho_1^{\otimes n}, \rho_2^{\otimes n}) \geq 1 - n\delta$.

Proof. This follows directly from the Fidelity definition of Equation 1. ■

5. QUANTUM KOLMOGOROV COMPLEXITY

We define the *quantum Kolmogorov complexity QC* of a string of qubits X, relative to a quantum Turing machine M, as the length of the shortest qubit string that, when given as input to M, produces on the output register the qubit string X. (Note that we only allow M that have computable transition amplitudes. See the articles [2, 6], and particularly Definition 3.2.2 in [2], for a further description of this computational model.)

5.1. Input/Output conventions

First we will specify in more detail what is meant by the 'input' and 'output' of a quantum computation.

We consider quantum Turing machines with two heads on two one-way infinite tapes: one input/work tape, and one output tape. We allow both tapes to be changed because we want to be able to move the input qubits to the output tape.

For a QTM M with a single input, when we say M starts with input Y, we mean that M starts with the quantum state $|Y\$00\cdots\rangle$ on its input tape, and $|00\cdots\rangle$ on the output tape. The $ symbol is a special endmarker (or blank) symbol.

Note that testing for the end of the input can be done without disturbing the input, since we assume that the '$' state is orthogonal to the '0' and '1' states. (This is analogous to the classical case, where where Turing machine inputs are encoded in a three-letter alphabet; nevertheless we consider the actual input to be encoded only over the characters 0 and 1.) A string is a proper input if the endmarker symbol appears only once and is not in superposition with any other position of the tape. We dismiss any non-proper inputs.

For a QTM with multiple inputs, we assume that there is a convention for encoding the multiple inputs so that they can be individually recovered. For example, when we write $M(Y_1, Y_2)$, we may assume that the input tape is initialized to $|1^{\ell(Y_1)}0Y_1Y_2\$00\cdots\rangle$: the sequence of ones $1^{\ell(Y_1)}$ is unambigiously indicated by the leftmost zero in the string, and with the thus obtained value $\ell(Y_1)$ we can seperate Y_1 and Y_2 from the remainder of the sequence. Likewise, for multiple outputs, if we write $M(Y_1, Y_2) = (X_1, X_2)$, we mean that X_1 and X_2 must be encoded according to a prearranged convention so that X_1 and X_2 can be recovered individually from the output tape. (Note that we do not define prefix-free complexity in this paper. The programs themselves need not be prefix-free.)

We let $M^T(X)$ denote the contents of the output tape after T steps of computation. We consider only QTMs which do not modify their output tape after they have 'halted'. (Because of reversibility, they may modify the input tape after reaching the halting state.) The output $M(X)$ is the content of the output tape at any time after M has stopped changing its output tape.

5.2. Definitions

For some fidelity function $f : \mathbb{N} \to [0, 1]$ we will now define the corresponding quantum Kolmogorov complexity.

DEFINITION 5.1. (Quantum Kolmogorov complexity with fidelity f) For any quantum Turing machine M and qubit string X, the f-approximation quantum Kolmogorov complexity, denoted $QC^f_M(X)$, is the length of the smallest qubit string P such that for any fidelity parameter k we have $\text{Fidelity}(X, M(P, 1^k)) \geq f(k)$.

Note that we require that the same string P be used for all approximation parameters k. This way the program cannot depend on a particular value of k. (Otherwise k itself might contain information about the string we want to describe.)

Note also that we allow the string X, the program P, and the output $M(P, 1^k)$ to be mixed states for the following reasons. There is no reason why the approximation $M(P, 1^k)$ of a pure state X has to be pure as well. By allowing mixed states we avoid this problem, and, as a bonus, get also a definition for the complexity of mixed states. Because the fidelity and the time evolution of M is properly defined for mixtures this causes no serious problems. (Clearly, the program P_ρ that simply moves ρ from the input to the output tape will have to be mixed as well, which explains the necessity of mixed input strings.)

If f is the constant function 1, we thus have the following definition.

DEFINITION 5.2. (Quantum Kolmogorov complexity with perfect fidelity) The perfect fidelity quantum Kolmogorov complexity is $QC^1_M(X)$.

The problem with this definition is that it is not known whether an invariance theorem can be given for this perfect-fidelity Kolmogorov complexity. This is because the invariance theorems that are known for quantum computers deal with *approximating* procedures rather than with exact simulations. We therefore prove an invariance theorem for a weaker, limiting version, where the output of M must have high fidelity with respect to the target string X: $\text{Fidelity}(X, M(P, 1^k)) \approx 1$.

DEFINITION 5.3. (Quantum Kolmogorov complexity with bounded fidelity) For any constant $\epsilon < 1$, $QC^\epsilon_M(X)$ is the constant-fidelity quantum Kolmogorov complexity.

Again there are problems with this definition. First, it may be the case that some strings are very easy to describe up to a given constant, but inherently very hard to describe for a smaller error. Second, it may be the case that some strings are easier to describe up to a given constant on one machine, but not on another machine. For these two reasons, this definition does not appear to be robust.

A stronger notion of approximability is the existence of an approximation *scheme.* (See, for example, the book by Garey and Johnson [9, Chapter 6] for more on approximation algorithms and approximation schemes.) For constant-approximability, different algorithms (with different sizes) can exist for different constants. In an approximation scheme, a single program takes as auxiliary input an approximation parameter k, and produces an output that approximates the value we want within the approximation parameter. This is the model we wish to adopt for quantum Kolmogorov complexity.

DEFINITION 5.4. (Quantum Kolmogorov complexity with fidelity converging to 1) The complexity $QC^{\uparrow 1}_M(X)$ is equal to $QC^f_M(X)$, where $f(k) = 1 - \frac{1}{k}$.

We choose to encode the fidelity parameter in unary, and the convergence function to be $f(k) = 1 - \frac{1}{k}$ so that the model remains robust when polynomial time bounds are added. We discuss this further in Section 6.

We may also define $QC^{\uparrow 1}_M(X|Y)$, the complexity of producing X when Y is given as an auxiliary input, in the usual way.

6. INVARIANCE

To show that our definition is robust we must show that the complexity of a qubit string does not depend on the underlying quantum Turing machine.

We use the following result, proved in the paper of Bernstein and Vazirani [2]. To be precise, we use the notation $\ulcorner M \urcorner$ to denote the classical description of the quantum Turing machine M. (Recall that we only consider quantum Turing machines whose amplitudes can be computed to arbitrary precision with a finite classical description.)

THEOREM 6.1. *(Universal quantum Turing machine [2]) There exists a universal quantum Turing machine U with a finite classical description such that the following holds. For any quantum Turing machine M (which has a finite classical description), for any pure state X, for any approximation parameter k, and any number of time steps T, we have* $\text{Fidelity}(U(\ulcorner M \urcorner, X, 1^k, T), M^T(X)) \geq 1 - \frac{1}{k}$. *Recall that M^T is the contents of the output tape of M after T time steps.*

THEOREM 6.2. *(Quantum invariance theorem) There is a universal quantum Turing machine U such that for any quantum Turing machine M and qubit string X,*

$$QC^{\uparrow 1}_U(X) \;\leq\; QC^{\uparrow 1}_M(X) + c_M,$$

where c_M is a constant depending only on M.

Proof. The proof of Theorem 6.2 follows from the existence of a universal quantum Turing machine, as proven by Bernstein and Vazirani [2]. Let U be this UTM. The constant c_M represents the size of the finite description that U requires to calculate the transition amplitudes of the machine M. Let P be the state that witnesses that $QC^{\uparrow 1}_M(X) = \ell(P)$, and hence $\text{Fidelity}(X, M(P, 1^k)) \geq 1 - \frac{1}{k}$ for every k.

With the description corresponding to c_M, U can simulate with arbitrary accuracy the behavior of M. Specifically, U can simulate machine M on input $(P, 1^{4k})$ with a fidelity of $1 - \frac{1}{4k}$. Therefore, by Lemma 4.1, $\text{Fidelity}(X, U(M, P, 1^{4k})) \geq 1 - \frac{1}{k}$. ■

The same holds true for the conditional complexity, that is, there exists a UTM U such that for all quantum machines M and quantum strings X, Y we have $QC^{\uparrow 1}_U(X|Y) \leq QC^{\uparrow 1}_M(X|Y) + c_M$.

Henceforth, we will fix a universal quantum Turing machine U and simply write $QC(X)$ instead of $QC^{\uparrow 1}_U(X)$. Likewise we write $QC(X|Y)$ instead of $QC^{\uparrow 1}_U(X|Y)$.

We also abuse notation and write M instead of $\ulcorner M \urcorner$ to represent the code of the quantum Turing machine M used as an input to the universal Turing machine.

The simplest application of the invariance theorem is the following proposition.

PROPOSITION 6.1. *There exists a constant c, such that for any qubit string X, $QC(X) \leq \ell(X) + c$. The value of c depends only on our choice of the underlying universal Turing machine.*

Proof. Consider the quantum Turing machine M that moves its input to the output tape, yielding $QC_M(X) = \ell(X)$. The proposition follows by invariance. ∎

We may also define time-bounded QC is the usual way, that is, fix $T : \mathbb{N} \to \mathbb{N}$ a fully-time-computable function. Then $QC^T(X|Y)$ is the length of the shortest program which on input $Y, 1^k$, produces X on its output tape after $T(\ell(X) + \ell(Y))$ computation steps. The Bernstein and Vazirani simulation entails a polynomial time blowup (polynomial in the length of the input and the length of the fidelity parameter encoded in unary), so there is a polynomial time blowup in the corresponding invariance theorem.

7. PROPERTIES OF QUANTUM KOLMOGOROV COMPLEXITY

In this section we compare classical and quantum Kolmogorov complexity by examining several properties of both. We find that many of the properties of the classical complexity, or natural analogues thereof, also hold for the quantum complexity. A notable exception is the complexity of m-fold copies of arbitrary qubit strings, which we will describe in Section 8.

7.1. Correspondence for classical strings

We would like to show that for classical states, classical and quantum Kolmogorov complexity coincide, up to a constant additive term.

PROPOSITION 7.1. *There is a constant c, such that for every finite, classical string x, it holds that $QC(x) \leq C(x) + c$.*

(The constant depends only on the underlying universal Turing machine.)

Proof. This is clear: the universal quantum computer can also simulate any classical Turing machine. ∎

The converse is also true, as shown by Peter Gács [8].

PROPOSITION 7.2. *(See [8] for the proof.) There is a constant c, such that for every finite, classical string x, it holds that $C(x) \leq QC(x) + c$.*

7.2. Quantum incompressibility

In this section, we show that there exist quantum-incompressible strings. Our main theorem is a very general form of the incompressibility theorem with some useful special cases as corollaries.

Assume we want to consider the minimal-length programs that describe a set of quantum states. In general, these may be pure or mixed states. We will use the following notation throughout the proof. The mixed states $\rho_1, \dots, \rho_M$ be are the target strings (those we want to produce as output). Their minimal-length programs will be $\sigma_1, \dots, \sigma_M$, respectively. The central idea is that if the states ρ_i are sufficiently different, then the programs σ_i must be different as well. We turn this into a quantitative statement with the use of the insensitive chi quantity in combination with the monotonicity of quantum mechanics.

Earlier, Michał Horodecki used a similar technique to prove a closely related result [11], which shows that the Holevo quantity is a lower bound for the optimal compression rate for ensembles of mixed states.

THEOREM 7.1. *For any set of strings $\rho_1, \dots, \rho_M$ such that $\forall i, QC(\rho_i) \leq l$, this l is bounded from below by*

$$l \ \geq \ S(\rho) - \tfrac{1}{M}\sum_i S(\rho_i),$$

where ρ is the 'average' density matrix $\rho = \frac{1}{M}\sum_i \rho_i$.

(Stated slightly differently, this says that there is an i such that $QC(\rho_i) \geq S(\rho) - \frac{1}{M}\sum_i S(\rho_i)$.)

Proof. Take $\rho_1, \dots, \rho_M$ and their minimal programs $\sigma_1, \dots, \sigma_M$ (and hence $QC(\rho_i) = \ell(\sigma_i)$). Let $\mathbf{S}^k$ be the completely positive, trace preserving map corresponding to the universal QTM U with fidelity parameter k. With this, we define the following three uniform ensembles:

- the ensemble $\mathcal{E} = \{(\rho_i, \frac{1}{M})\}$ of the original strings,
- $\mathcal{E}_\sigma$ the ensemble of programs $\{(\sigma_i, \frac{1}{M})\}$, and
- the ensemble of the k-approximations $\tilde{\mathcal{E}}^k = \mathbf{S}^k(\mathcal{E}_\sigma) = \{(\tilde{\rho}_i^k, \frac{1}{M})\}$, with $\tilde{\rho}_i^k = \mathbf{S}^k(\sigma_i)$.

By the monotonicity of Theorem 2.3 we know that for every k, $\chi(\tilde{\mathcal{E}}^k) \leq \chi(\mathcal{E}_\sigma)$. The chi quantity of the ensemble $\mathcal{E}_\sigma$ is upper bounded by the maximum size of its strings: $\chi(\mathcal{E}_\sigma) \leq \max_i\{\ell(\sigma_i)\} \leq l$. Thus the only thing that remains to be proven is that $\chi(\tilde{\mathcal{E}}^k)$, for sufficiently big k, is 'close' to $\chi(\mathcal{E})$. This will be done by using the insensitivity of the Von Neumann entropy.

By definition, for all i, $\lim_{k\to\infty} \mathrm{Fidelity}(\rho_i, \tilde{\rho}_i^k) = 1$, and hence $\lim_{k\to\infty} \tilde{\rho}_i^k = \rho_i$. Because the ensembles $\mathcal{E}$ and $\tilde{\mathcal{E}}^k$ have only a finite number (M) of states, we can use Lemma 2.1, to obtain $\lim_{k\to\infty} \chi(\tilde{\mathcal{E}}^k) = \chi(\mathcal{E})$. This shows that for any $\delta > 0$, there exists a k such that $\chi(\mathcal{E}) - \delta \leq \chi(\tilde{\mathcal{E}}^k)$. With the above inequalities we can therefore conclude that $\chi(\mathcal{E}) - \delta \leq l$ holds for arbitrary small $\delta > 0$, and hence that $l \geq \chi(\mathcal{E})$. ■

The following four corollaries are straightforward with the above theorem.

COROLLARY 7.1. *For every length n, there is an incompressible classical string of length n.*

Proof. Apply Theorem 7.1 to the set of classical strings of n bits: $\rho_x = |x\rangle\langle x|$ for all $x \in \{0,1\}^n$. All ρ_x are pure states with zero Von Neumann entropy, hence the lower bound on l reads $l \geq S(\rho)$. The average state $\rho = 2^{-n}\sum_x |x\rangle\langle x|$ is the total mixture $2^{-n}I$ with entropy $S(\rho)=n$, hence indeed $l \geq n$. ∎

COROLLARY 7.2. *For any set of orthogonal pure states $|\phi_1\rangle,\ldots,|\phi_M\rangle$, the smallest l such that for all i, $QC(\phi_i) \leq l$ is at least $\log M$. (Stated differently, there is an i such that $QC(\phi_i) \geq \log M$.)*

Proof. All the pure states have zero entropy $S(\phi_i)=0$, hence by Theorem 7.1: $l \geq S(\rho)$. Because all ϕ_is are mutually orthogonal, this Von Neumann entropy $S(\rho)$ of the average state $\rho = \frac{1}{M}\sum_i |\phi_i\rangle\langle\phi_i|$ equals $\log M$. ∎

COROLLARY 7.3. *For every length n, at least $2^n - 2^{n-c}+1$ mutually orthogonal qubit strings of length n have complexity at least $n-c$.*

COROLLARY 7.4. *For any set of pure states $|\phi_1\rangle,\ldots,|\phi_M\rangle$, the smallest l such that for all i, $QC(\phi_i) \leq l$ is at least $S(\rho)$, where $\rho = \frac{1}{M}\sum_i |\phi_i\rangle\langle\phi_i|$.*

8. THE COMPLEXITY OF COPIES

It is trivial to copy a classical bit string x to the m-fold state $x^{\otimes m}$. As long as we know the integer m, the complexity of $x^{\otimes m}$ is no bigger than that of the single copy x, or in Kolmogorov complexity terms: $C(x^{\otimes m}|m) \leq C(x)+O(1)$. This no longer holds in the case of quantum information, as it is in general not possible to copy an unknown quantum state.[28] Typically for a quantum state X, the complexity $QC(X^{\otimes m}|m)$ will grow as m gets bigger. This should not surprise us because a large number m of copies enables us to estimate the amplitudes of X more accurately than a single copy would. Hence, we can 'extract' more information from $X^{\otimes m}$ if we have more copies of X. An obvious upper bound on the quantum Kolmogorov complexity of $X^{\otimes m}$ is $QC(X^{\otimes m}|m) \leq m\cdot QC(X)$. The two main theorems of this section tell us that, despite the 'no cloning' phenomenon of quantum mechanics, it is possible to compress copies of pure states. This result is established with the help of the theory of symmetric subspaces. We start with the general upper bound.

THEOREM 8.1. *There exists a constant c, such that for an arbitrary pure state X and integer m it holds that*

$$QC(X^{\otimes m}|m) \;\leq\; \log\binom{m+2^{QC(X)}-1}{2^{QC(X)}-1}+c, \tag{4}$$

and hence $QC(X^{\otimes m}) \leq \log\binom{m+2^{QC(X)}-1}{2^{QC(X)}-1}+O(\log m)$.

Proof. First we sketch the proof, omitting the effect of the approximation. Consider a pure qubit string X whose minimal-length program is P_X. To produce m copies of X, it suffices to produce m copies of P_X and execute these m programs.

We can always assume that this P_X is a pure state, because for a mixture of programs, any of the pure programs in the mixtures will produce X as well. Let l be the length $QC(X)$ of P_X; we denote the 2^l-dimensional Hilbert space by $\mathcal{H}$. Consider $\mathcal{H}^{\otimes m}$, the m-fold tensor product of $\mathcal{H}$. The symmetric subspace $\mathcal{SYM}(\mathcal{H},m)$ is d-dimensional, where $d = \binom{m+2^l-1}{2^l-1}$. The sequence $P_X^{\otimes m}$ sits in this symmetric subspace, and can therefore be encoded exactly using $\log d + O(\log m)$ qubits, where the m term is used to describe the rotation from the d-dimensional space to the m copies in $\mathcal{H}^{\otimes m}$. Hence, given m, the quantum Kolmogorov complexity of $X^{\otimes m}$ is bounded from above by $\log d + O(1)$ qubits.

For the full proof, we will need to take into account the effect of the imperfect fidelities and prove that we can reach a fidelity not smaller than $1 - \frac{1}{k}$.

The first part of the computation consists of the mapping from the d dimensions to the symmetric subspace $\mathcal{SYM}(\mathcal{H},m)$. This is the transformation

$$|i\rangle \;\rightarrow\; |A_i\rangle, \tag{5}$$

for $1 \leq i \leq d$ which labels all the multisets $A_i \subseteq \{1,\ldots,2^l\}$ of size m. We approximate this unitary transformation with enough accuracy such that the output has fidelity $\geq 1 - \frac{1}{4k}$ with the perfect state $P_X^{\otimes m}$.

Next, we execute the programs P_X with a fidelity parameter of $4km$. Hence the joint, m-fold evolution $U_2^{\otimes m}$ establishes $\text{Fidelity}(X^{\otimes m}, U_2^{\otimes m}(P_X^{\otimes m})) \geq 1 - \frac{1}{4k}$ (*cf.* Lemma 4.3).

We finish the proof by employing Lemma 4.2, which tells us that the overall fidelity-error of the above two transformations cannot be bigger than $\frac{1}{k}$. ■

This upper bound is also very close to being tight for some X, as we show in the next theorem.

THEOREM 8.2. *(Incompressibility for copies of quantum states) For every m and n, there is an n-qubit state X such that*

$$QC(X^{\otimes m}) \geq \log \binom{m+2^n-1}{2^n-1}$$

.

Proof. Fix m and n and let $\mathcal{H}$ be the 2^n-dimensional Hilbert space. Consider the (continuous) ensemble of all m-fold tensor product states $X^{\otimes m}$: $\mathcal{E} = \{(X^{\otimes m}, \mu)\}$, where $\mu^{-1} = \int_{X\in\mathcal{H}} dX$ is the appropriate normalization factor. The corresponding average state is calculated by the integral $\rho = \mu \int_{X\in\mathcal{H}} X^{\otimes m} dX$. This mixture is the totally mixed state in the symmetric subspace $\mathcal{SYM}(\mathcal{H},m)$ (see Section 3 in [27]), and hence has entropy $S(\rho) = \log\binom{m+2^n-1}{2^n-1}$. Because all $X^{\otimes m}$ are pure states, we can use Corollary 7.4 to prove the existence of a X for which $QC(X^{\otimes m}) \geq \log\binom{m+2^n-1}{2^n-1}$. ■

The results of this section can be viewed as a refinement of the no-cloning theorem, in the following sense. The quantity $QC(X^{\otimes m}|m)$, for any state X, gives a measure of how clonable that particular state is. Theorem 8.2 tells us that there exist strings that are 'maximally non-clonable'.

8.1. Subadditivity

Consider the following subadditivity property of classical Kolmogorov complexity.

PROPOSITION 8.1. *For any x and y, $C(x,y) \leq C(x) + C(y|x) + O(\log(C(x)))$.*

In the classical case, we can produce x, and then produce y from x, and print out the combination of x and y. In the quantum case, producing Y from X may destroy X. In particular, with $X = Y$, the immediate quantum analogue of Proposition 8.1 would contradict the $m = 2$ case of Theorem 8.2.

A natural quantum extension of this result is as follows.

PROPOSITION 8.2. *For any pair of quantum strings X, Y, we have $QC(X,Y) \leq QC(X,X) + QC(Y|X) + O(\log(QC(X)))$.*

9. QUANTUM INFORMATION THEORY

In this section we establish a relationship between quantum compression theory and the bounded-fidelity version of quantum Kolmogorov complexity.

One would like to give a direct analogue of Proposition 2.3. We prove below a slightly weaker statement, for bounded-fidelity complexity.

We believe that the direct analogue of Proposition 2.3 may not hold for quantum Kolmogorov complexity. The argument can be summarized as follows. In the classical case, given a string x, we can define a source A such that x is in the so-called 'typical subspace' of A. This allows us to give a short, exact description of x. In the quantum case, we may also define a quantum source likely to have emitted a given qubit string X (in an appropriate tensor space). However, we do not get that X is *in* the typical subspace of this source, only that it is *close* to the typical subspace. How close it can be guaranteed to be depends on the length of X. Therefore, for a fixed string length n, we may not be able to get an encoding of arbitrary high fidelity.

THEOREM 9.1. *Let U be the universal quantum Turing machine from [2]. Then for any ϵ, δ there is an n such that for any d-dimensional $\mathcal{H}$, and any qubit string $X = |\phi_1\rangle \otimes \cdots \otimes |\phi_n\rangle \in \mathcal{H}^{\otimes n}$,*

$$QC_U^{\epsilon}(X) \;\leq\; n(S(\rho) + \delta + \tfrac{1}{n}(d^2 \log(n+1))),$$

where $\rho = \frac{1}{n}\sum_i |\phi_i\rangle\langle\phi_i|$.

Proof. Fix ϵ, δ. Apply Theorem 2.2 with $\epsilon' = \frac{\epsilon}{4}, \delta' = \delta$, and let $n = n(\epsilon', \delta')$ be the value from the theorem. Let $|\phi_1\rangle \otimes \cdots \otimes |\phi_n\rangle \in \mathcal{H}^{\otimes n}$ be the string for whose quantum Kolmogorov complexity we want to give an upper bound. By Theorem 2.2(1), we get that the length of the encoding is what was given in the statement of the theorem. By simulating the decoding algorithm to a precision of $\frac{\epsilon}{4}$, together with Theorem 2.2(2), and Lemma 4.1, we have that the fidelity of the encoding is at least $1-\epsilon$. That completes the proof. ■

10. THE COMPLEXITY OF CORRELATIONS

In this section we will use quantum Kolmogorov complexity to quantify the complexity of the correlation between two systems. For a bipartite state ρ_{AB} we denote this quantity by $KCor(\rho_{AB})$, which is defined as follows.

DEFINITION 10.1. (Quantum Kolmogorov complexity of correlations) Consider a bipartite state ρ_{AB} of $n+m$ qubits where n qubits are on A's side and B has the remaining m qubits. The quantum Kolmogorov complexity $KCor$ of the correlation between A and B is defined by

$$KCor(\rho_{AB}) \;=\; QC(\rho_{AB}|\rho_A,\rho_B),$$

where $\rho_A = \mathrm{tr}_B(\rho_{AB})$ and $\rho_B = \mathrm{tr}_A(\rho_{AB})$.

Because the complexity $KCor(\rho_{AB})$ can never be bigger than $QC(\rho_{AB})$, the following general upper bound holds.

PROPOSITION 10.1. *There exists a constant c such that for every bipartite, $n+m$-qubit state ρ_{AB} we have*

$$KCor(\rho_{AB}) \;\leq\; n+m+c. \tag{6}$$

Proof. Apply Proposition 6.1 to the relation $KCor(\rho_{AB}) \leq QC(\rho_{AB})$. ∎

The gap between the correlation complexity $KCor$ and the Kolmogorov complexity can be made arbitrarily big as is shown by the next lemma.

LEMMA 10.1. *There exists a constant c such that for any combination of lengths n and m, there is an $n+m$-qubit string ρ_{AB} with maximum Kolmogorov complexity $QC(\rho_{AB}) \geq n+m$, combined with a constant lower bound on the complexity of the correlation $KCor(\rho_{AB}) \leq c$.*

Proof. Consider the set of classical strings of length $n+m$. Clearly, these states can be expressed as tensor products $X_{AB} = X_A \otimes X_B$, where X_A (X_B) are n (m) bit strings. By the program of size c that moves the inputs X_A and X_B to the output tape (thus producing X_{AB}) we obtain $KCor(X_{AB}) = QC(X_{AB}|X_A,X_B) \leq c$. On the other hand, by Corollary 7.1, at least one of these strings X_{AB} also has to obey $QC(X_{AB}) \geq n+m$. ∎

The central idea behind the definition of $KCor$ is that we consider the complexity of the correlations 'high' when the partial states ρ_A and ρ_B do not contain much information about the total configuration ρ_{AB}. In this sense it is possible that all the complexity of a state is contained in its correlations. The following lemma expresses this result.

LEMMA 10.2. *For every length n, there exists a bipartite, $n+n$-qubit state ρ_{AB} with maximum correlation complexity $KCor\rho_{AB} \geq 2n$.*

Proof. First we consider the $n = 1$ case of two distributed qubits. Take the four Bell states $|\phi^1_{AB}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, $|\phi^2_{AB}\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle)$, $|\phi^3_{AB}\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$, and $|\phi^4_{AB}\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$. As these states are mutually orthogonal, we can use the uniform source $\mathcal{E} = \{(\phi^i_{AB}, \frac{1}{4})\}$ to encode two bits of information.[3] It is also straightforward to see that all the partially traced out states are identical to the same totally mixed qubit: $\phi^i_A = \phi^i_B = \frac{1}{2}(|0\rangle\langle + || 1\rangle\langle 1|) = \frac{1}{2}I$ for all i. Hence, for one of the ϕ's we must have $KCor(\phi^i_{AB}) = QC(\phi^i_{AB}|\frac{1}{4}I \otimes I) \geq 2$.

This result easily generalizes to the $n+n$-qubit case if we take the n-fold tensor product of the above source. We can use the words of this $\mathcal{E}^{\otimes n}$ to encode $2n$ bits of information, while the partially traced out words all equal the totally mixed n qubit state $2^{-n}I$. This shows that for at least one of the words it must hold that its correlation complexity is not smaller than $2n$. ■

It would be incorrect to think that the complexity $KCor$ is 'yet another measure of entanglement'. It is true that tensor product states $X_A \otimes X_B$ have a low correlation complexity, but so have highly entangled states like $(\frac{1}{\sqrt{2}}|0_A 0_B\rangle + \frac{1}{\sqrt{2}}|1_A 1_B\rangle)^{\otimes n}$. Moreover, the definition also covers the complexity of purely classical correlations. Rather than quantifying entanglement, we expect the above definition to be useful in the context of 'communication complexity theory'. The last section of this article will explain this point further.

11. EXTENSIONS AND FUTURE WORK

We have argued that the QC of Definition 5.4 is a robust notion of Kolmogorov complexity for the quantum setting. It would be interesting to see if an invariance theorem can be shown for the ideal quantum Kolmogorov complexity of Definition 5.2. It would also be interesting to see if the invariance theorem (Theorem 7.1) can be improved in general.

Kolmogorov complexity in the classical setting is a good tool for showing lower bounds in computational complexity. For instance, one can show lower bounds in classical communication by using classical Kolmogorov complexity. A simple example is the following lower bound on the communication complexity of the equality function. We sketch the argument here. Assume that there is a protocol that decides whether two strings of length n are equal, in which t bits are exchanged. Consider an incompressible string x of length n, and simulate the protocol on input (x, x). Let T be the transcript of the communication on that input. Now we argue that the Kolmogorov complexity of the string can be bounded above by a function of t. To print x, we use the transcript and the protocol to find x as follows. Without loss of generality, assume that the second player always decides whether or not to accept the input. For every candidate z for x, simulate the protocol for the second player on input z, and use the transcript to obtain the communication that the second player would have received from the first player. Because the protocol is sound, the simulation will only accept if $z = x$. We output whenever a string is found that causes the protocol to accept. This program which prints x is of size (roughly) t, and therefore we have $n \leq C(x) \leq t$, from which we can conclude that the communication complexity of the equality function is at least n.

Could a similar argument be carried over to the quantum setting? If so, then by applying this framework to other problems in quantum complexity, quantum Kolmogorov complexity could become a powerful new tool in proving lower bounds.

The number of applications of classical Kolmogorov complexity is countless, and it is our hope that this definition will lead to a similar wide variety of applications in quantum complexity theory.

ACKNOWLEDGMENT

We would like to thank several people for interesting discussions or comments on this work: Jérémy Barbay, Harry Buhrman, Richard Cleve, David Deutsch, Peter Gács, Michał Horodecki, Frédéric Magniez, Miklos Santha, Paul Vitányi, John Watrous, and Ronald de Wolf.

This work has been supported by Wolfson College Oxford, Hewlett-Packard, European TMR Research Network ERP-4061PL95-1412, the Institute for Logic, Language and Computation in Amsterdam, an NSERC postdoctorate fellowship, the EU fifth framework project QAIP IST-1999-11234, and the TALENT grant S 62-552 of the Netherlands Organization for Scientific Research (NWO).

REFERENCES

1. A. Barenco, A. Berthiaume, D. Deutsch, A. Ekert, R. Jozsa, and C Macchiavello, "Stabilisation of Quantum Computations by Symmetrisation", *SIAM Journal on Computing,* Volume 26, No. 5, pp. 1541–1557 (1997)
2. E. Bernstein and U. Vazirani, "Quantum Complexity Theory", *SIAM Journal on Computing,* Volume 26, No. 5, pp. 1411–1473 (1997)
3. C. Bennett and S. Wiesner, "Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states", *Physical Review Letters,* Volume 69, pp. 2881–2884 (1992)
4. G. Chaitin, "On the length of programs for computing finite binary sequences", *Journal of the ACM,* Volume 13, No. 4, pp. 547–569 (1966)
5. T.M. Cover and J.A. Thomas, *Elements of Information Theory*, John Wiley & Sons, Wiley Series in Telecommunications (1991)
6. D. Deutsch, "Quantum theory, the Church-Turing principle and the universal quantum computer", *Proceedings of the Royal Society of London A,* Volume 400, pp. 97–117 (1985)
7. C.A. Fuchs and J. van de Graaf, "Cryptographic Distinguishability Measures for Quantum Mechanical States", *IEEE Transactions on Information Theory,* Volume 45, No. 4, pp. 1216–1227 (1999)
8. P. Gács, "Quantum Algorithmic Entropy", arXiv:quant-ph/0011046 (2000)
9. M.R. Garey and D.S. Johnson, *Computers and Intractibility, A Guide to the Theory of NP Completeness,* W.H. Freeman (1979)
10. A.S. Holevo, "Bounds for the Quantity of Information Transmitted by a Quantum Communication Channel", *Problemy Peredachi Informatsii,* Volume 9, No. 3, pp. 3–11 (1973); English Translation in *Problems in Information Transmission,* Volume 9, pp. 177–183 (1973)
11. M. Horodecki, "Limits for compression of quantum information carried by ensembles of mixed states", *Physical Review A,* Volume 57, No. 5, pp. 3364–3369 (1998)
12. R. Jozsa, M. Horodecki, P. Horodecki, R. Horodecki, "Universal Quantum Information Compression", *Physical Review Letters,* Volume 81, pp. 1714–1717 (1998)
13. R. Jozsa and B. Schumacher, "A New Proof of the Quantum Noiseless Coding Theorem", *Journal of Modern Optics,* Volume 41, pp. 2343–2349 (1994)
14. A.K. Kolmogorov, "Three approaches to the quantitative definition of information", Problems of Information Transmission, Volume 1, pp. 1–7 (1965)
15. M. Li and P. Vitányi, *An Introduction to Kolmogorov Complexity and its Applications,* Second Edition, Springer Verlag (1997)
16. G. Lindblad, "Completely Positive Maps and Entropy Inequalities", *Communications in Mathematical Physics,* Volume 40, pp. 147–151 (1975)
17. M.A. Nielsen, PhD thesis, University of New Mexico (1998); arXiv:quant-ph/0011036 (2000)

18. M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information,* Cambridge University Press (2000)

19. J. Preskill, "Quantum Computing" (1998); course notes available at URL: http://www.theory.caltech.edu/people/preskill/ph229/

20. B. Schumacher, "Quantum Coding", *Physical Review A,* Volume 51, No. 4, pp. 2738–2747 (1995)

21. C.E. Shannon and W. Weaver, *The mathematical theory of communication,* University of Illinois Press (1949)

22. R. Solomonoff, "A preliminary report on a general theory of inductive inference", technical report ZTB-138, Zator Company, Cambridge, Mas. (1960)

23. K. Svozil, "Quantum algorithmic information theory", *Journal of Universal Computer Science* Volume 2, pp. 311–346 (1996).

24. A. Uhlmann, "Relative Entropy and the Wigner-Yanase-Dyson-Lieb Concavity in an Interpolation Theory", *Reviews in Mathematical Physics,* Volume 54, pp. 21–32 (1977)

25. P. Vitányi, "Three Approaches to the Quantitative Definition of Information in an Individual Pure Quantum State", *Proceedings of the 15th Annual Conference on Computational Complexity,* pp. 263–270 (2000)

26. A. Wehrl, "General Properties of Entropy", *Reviews of Modern Physics,* Volume 50, No. 2, pp. 221–260 (1978)

27. R.F. Werner, "Optimal cloning of pure states", *Physical Review A,* Volume 58, pp. 1827–1832 (1998)

28. W.K. Wootters and W.H. Zurek, "A single quantum cannot be cloned", *Nature,* Volume 229, pp. 802–803 (1982)