Quantum Kolmogorov Complexity

André Berthiaume
School of CTI
DePaul University, Chicago
berthiaume@cs.depaul.edu

Wim van Dam
C.W.I. Amsterdam
Centre for Quantum Computation
University of Oxford
wimvdam@qubit.org

Sophie Laplante
L.R.I.
Université Paris Sud
Sophie.Laplante@lri.fr

Abstract

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 Vitányi [20] measures the amount of classical information, whereas we consider the amount of* quantum *information in a qubit string. We argue that our definition is natural and is an accurate representation of the amount of quantum information contained in a quantum state.*

1 Introduction

In classical computations, the Kolmogorov-Solomonoff-Chaitin (Kolmogorov, for short) complexity of a finite string is a measure of its randomness.[3, 11, 18] The Kolmogorov complexity of x is the length of the shortest program which 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 information theory.[4, 17]

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 what is a good definition for the Kolmogorov complexity of quantum strings. Our goal in this paper is to argue that our definition is a natural and robust measure the amount of quantum information contained in a quantum string, which has several appealing properties.

Recently, Paul Vitányi [20] has also proposed a definition for quantum algorithmic complexity. Our definition differs significantly from Vitányi's: the definition he proposes is a measure of the amount of *classical* information necessary to approximate the quantum state.

The paper will be organized as follows: In Section 3, we give basic notation, definitions, prior work and some theorems that will be used in proofs in the paper. In Section 4 we give our definition of quantum Kolmogorov complexity. In Section 5 we prove the invariance theorem. Section 6 compares the properties of quantum and classical Kolmogorov complexity, including incompressibility, subadditivity, and the complexity of copies. Section 7 discusses the relationship with quantum information theory. We conclude with a discussion of possible extensions and future work.

2 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.
- It should be closely related to classical complexity on classical strings.

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 phenomena such as the no-cloning theorem, whose consequences we will discuss later in the paper.[23]

0-7695-0674-7/00 $10.00 © 2000 IEEE

2.1 Critical issues

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 bit strings (programs) are countable, whereas qubit strings are uncountable, so any definition that measures in bits would have to overcome this apparent contradiction. Paul Vitányi [20] 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 the Turing machines, say by restricting the set of possible amplitudes. What would be a reasonable choice?) We will use an approximation scheme.

What model of computation? 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 for a quantum Turing machine is its input, and if we want to count program length in qubits, 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 cannot assume that the shortest program can be run several times to produce several copies of the same string. This may be due to the fact that it is not possible to recover the program without losing its output. There is also a second reason not to choose the multiple generation option. The complex-valued parameters α and β of a qubit $|q\rangle = \alpha|0\rangle + \beta 1$ can contain an unbounded amount of information. If we would be 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.[8] For the above two reason, we will not require a 'reusability' condition.

3 Preliminaries

We start with some notation, definitions, and results that will be used to prove the results in this paper.

3.1 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 $\phi,\psi,\ldots$ to denote quantum pure states. Mixed states are represented by the letters ρ,σ etc. We also use uppercase letters $X,Y,\ldots$ 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 or string lengths.

For classical strings over the alphabet $\Sigma = \{0,1\}$, $\ell(x)$ 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^n.

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

We slightly abuse notation by sometimes letting the state symbols $\phi,\rho,\ldots$ 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 $p_1,p_2,\ldots$ a probability distribution over the mutually orthogonal states $\phi_1,\phi_2,\ldots$. 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 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

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

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 specific distribution $p_1, p_2, \ldots$ over a set of (mixed) states $\rho_1, \rho_2, \ldots$. We denote this by $\mathcal{E} = \{(\rho_i, p_i)\}$. The average state of such an ensemble $\mathcal{E}$ is $\rho = \sum_i p_i \rho_i$. An average state corresponds 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 \$ on the space of density matrices is allowed by the laws of quantum mechanics if and if only it is a completely positive, trace preserving mapping.

3.2 Classical Kolmogorov complexity

The Kolmogorov complexity of a string, in the classical setting, is the length of the shortest program which prints this string on an empty input.[12]

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

Definition 1 *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 of x relative to y (with respect to Φ) is defined as*

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

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 theorem 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 1 *There is a universal partial computable function Φ_0 such that for any partial computable Φ and pair of strings x, y,*

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

where c is a constant depending only on Φ.

Giving an invariance theorem will be key to showing that 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 1 *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 $< n$.

3.3 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.[4, 17]

Definition 2 *Let A be a random source that 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 to encode sequences of character emitted by a random source.

Proposition 2 ***Shannon's noiseless coding [17]:*** *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 their corresponding probabilities. Then*

1. *for any ϵ, δ, 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 $P_{\text{success}} \geq 1 - \epsilon$;*
2. *for any ϵ, δ, there is an n such that for any δ', there is an ϵ' such that if the channel encodes n letters, each letter with less than $H(A) - \delta'$ bits per letter, then the probability of success $P_{\text{success}} \leq 2^{-n(\delta'-\delta)} + \epsilon'$.*

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 [12, 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 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, it can be shown that

Proposition 3

$$C(x) \quad < \quad m(H(A) + \gamma),$$

where A is the source defined in the discussion above, and γ vanishes as m goes to infinity.

3.4 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 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, we see that*

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

where $H(p)$ is the Shannon entropy of the probability distribution $p_1, p_2, \ldots$

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 [16] 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 [15, page 190] or in Nielsen's thesis [14, Chapter 7].)

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. [9], building upon the work of Jozsa and Schumacher [10].

Theorem 2 ***Universal quantum compression (see [10, 9]):*** *Consider pure state sources $\mathcal{E} = \{(\phi_i, p_i)\}$. For any ϵ, δ, there is an $n = n(\epsilon, \delta)$ such that for any entropy bound S, there is an encoding scheme that works for any source of Von Neumann entropy at most S that has 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 $S(\rho) \leq S$, then*

1. *Each $|\phi_i\rangle$ can be encoded by a code word σ_i, which has length $\leq S + \delta + \frac{1}{n}(d^2 \log(n+1))$.*
2. *For each i, $\mathrm{Fidelity}(\phi_i, \sigma_i) \geq 1 - \epsilon$.*

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

Definition 4 ***Holevo's chi quantity [8]:*** *For an ensemble $\mathcal{E} = \{(\rho_i, p_i)\}$, with $\rho = \sum_i p_i \rho_i$, Holevo's chi quantity equals*

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

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

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

Theorem 3 ***Lindblad-Uhlmann monotonicity [13, 19]:*** *Let $\mathcal{E} = \{(\rho_i, p_i)\}$ be an ensemble, and $\$$ a completely positive, trace preserving mapping. For every such $\mathcal{E}$ and $\$$, it holds that: $\chi(\$(\mathcal{E})) \leq \chi(\mathcal{E})$, where $\$(\mathcal{E})$ is the transformed ensemble $\{(\$(\rho_i), p_i)\}$.*

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

Lemma 1 ***Insensitivity of Von Neumann entropy (see Section II.A in [21]):*** *If a sequence $\rho_1, \rho_2, \ldots$, 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. □

Further background on these measures of quantum information and their properties can be found in [15, Chapter 5]. Another good source is Nielsen's thesis [14].

3.5 Symmetric spaces

We use the symmetric subspace of the Hilbert space to show 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 $\mathrm{Sym}(\mathcal{H}_D^m)$ of the m-fold tensor product space $\mathcal{H}_D^m$ is a subspace spanned by as many basis vectors as there are multisets of size m of $\{1, \ldots, D\}$. Let $A = \{i_1, \ldots, i_m\}$ be such a multiset of $\{1, \ldots, D\}$. Then, $|s_A\rangle$ is the normalized superposition of all the different permutations of

$i_1, \ldots, i_m$. The set of the different vectors $|s_A\rangle$ (ranging over the multisets A) is an orthogonal basis of the symmetric subspace $\mathrm{Sym}(\mathcal{H}_D^m)$. Hence the dimension of the symmetric subspace is $\binom{m+D-1}{D-1}$. (This is because choosing a multiset is the same thing as splitting m consecutive elements into D (possibly empty) intervals, where the size of ith interval represents the number of times the ith element appears in the multiset. The number of ways of splitting an interval of size m into D intervals is $\binom{m+D-1}{D-1}$.)

An equivalent definition of the symmetric subspace is that it is the smallest subspace that contains all the states of the form $|\phi\rangle^m$, for all $|\phi\rangle \in \mathcal{H}_D$. (For more on the symmetric subspace and its properties, see the paper by Barenco et al. [1].)

3.6 Accumulation of errors

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

Lemma 2 ***Fidelity of composition:*** *If* $\mathrm{Fidelity}(\rho, \rho') \geq 1-\delta_1$ *and* $\mathrm{Fidelity}(\rho', \rho'') \geq 1-\delta_2$, *then* $\mathrm{Fidelity}(\rho, \rho'') \geq 1 - 2\delta_1 - 2\delta_2$.

Proof: This follows from the fact that the fidelity between two mixed states ρ and σ equals the maximum 'pure state fidelity' $|\langle\phi|\psi\rangle|$, where ϕ and ψ are 'purifications' of ρ and σ. (See [6] for more details on this.) □

In order to give bounds on the complexity of several copies of a state, as we do in Section 6.3, we need the following bound on the total error in the n-fold tensor product of the approximation of a given state.

Lemma 3 *Let* ρ^n *and* σ^n *be the* n*-fold copies of the mixed states* ρ *and* σ, *then* $\mathrm{Fidelity}(\rho^n, \sigma^n) = (\mathrm{Fidelity}(\rho, \sigma))^n$.

Proof: This follows directly from the definition $\mathrm{Fidelity}(\rho, \sigma) = \mathrm{tr}\left(\sqrt{\sqrt{\rho} \cdot \sigma \cdot \sqrt{\rho}}\right)$. □

4 Quantum Kolmogorov Complexity

We define the *quantum Kolmogorov complexity* QC of a string of qubits, relative to a quantum Turing machine M, as the length of the shortest qubit string which when given as input to M, produces on its output register the qubit string. (Note that we only allow M that have computable transition amplitudes. See the articles [2, 5], and particularly Definition 3.2.2 in [2], for a further description of this computational model.)

4.1 Input/Output Conventions

We give some precisions about what is meant by 'input' and 'output'.

We consider quantum Turing machines with two heads on two one-way infinite tapes. We allow the input tape to be changed. This is required: for example, the contents of the input may have to be moved 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 also assume that there is a convention for encoding the multiple inputs so that they can be individually recovered. For example, when we write $M(P, Y)$, we may assume that the input tape is initialized to $|1^{\ell(P)}PY\$00\cdots\rangle$. We only count the length of X and Y for the length of the input. Likewise, for multiple outputs, if we write $M(P, Y) = (X_1, X_2)$, 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.

4.2 Definitions

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

Definition 5 ***Quantum Kolmogorov complexity with fidelity f:*** *For any quantum Turing machine* M *and qubit string* X, *the* f*-approximation quantum Kolmogorov complexity, denoted* $QC_M^f(X)$, *is the length of the smallest qubit string* P *such that for any fidelity parameter* k *we have* $\mathrm{Fidelity}(X, M(P, 1^k)) \geq f(k)$.

Note that we require that the same string P be used for all approximation parameters k.

We will say that program P M-computes X with fidelity $f(k)$ if $\forall k$, $\mathrm{Fidelity}(M(P, 1^k), X) \geq f(k)$.

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

Definition 6 ***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 the ideal Kolmogorov complexity. This is because the invariance theorems that are known for quantum computers deal with *approximating* procedures. 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: $\text{Fidelity}(X, M(P)) \approx 1$.

Definition 7 ***Quantum Kolmogorov complexity with bounded fidelity:*** *For any constant* $\epsilon < 1$, $QC^\epsilon_M(X)$ *is the constant-fidelity quantum Kolmogorov complexity.*

There are two 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 [7, 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 8 ***Quantum Kolmogorov complexity with fidelity converging to 1:*** *The* $QC^{\uparrow 1}_M(X)$ *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 5.

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.

5 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 $\overline{M}$ 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 4 ***Universal quantum Turing machine (see [2]):*** *There exists a universal quantum Turing machine* U *that has 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, $\text{Fidelity}(U(\overline{M}, X, 1^k, T), M^T(X)) \geq 1 - \frac{1}{k}$. *Recall that* M^T *is the contents of the output tape of* M *after* T *time steps.*

Theorem 5 *There is a universal quantum Turing machine* U *such that for any quantum Turing machine* M *and qubit strings* X,

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

where c_M *is a constant depending only on* M.

Proof: The proof follows from the existence of a universal quantum Turing machine, as proven by Bernstein and Vazirani [2]. Let U be this UTM as mentioned above. 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 witness 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 2, $\text{Fidelity}(X, U(M, P, 1^{4k})) \geq 1 - \frac{1}{k}$. □

The same holds true for the conditional complexity, that is, $\exists U \forall M, X, Y$, $QC^{\uparrow 1}_U(X|Y) \leq QC^{\uparrow 1}_M(X|Y) + c_M$.

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 $\overline{M}$ to represent the code of the quantum Turing machine M used as an input to the universal Turing machine.

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.

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

Proposition 4 *For any qubit string X, $QC(X) \leq \ell(X) + c$, where c is a constant depending 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. □

6 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.

6.1 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 5 *For any finite, classical string x, $QC(x) \leq C(x) + O(1)$.*

(The constant hidden by the big-O notation depends only on the underlying universal Turing machine.)

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

We leave as a tantalizing open question whether the converse is also true, that is:

Open Problem 1 *Is there a constant c such that for every finite, classical string x, $C(x) \leq QC(x) + c$?*

6.2 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. We state 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 $\rho_1, \ldots, \rho_M$ be are the target strings (those we want to produce as output). Their minimal-length programs will be $\sigma_1, \ldots, \sigma_M$, 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.

Theorem 6 *For any set of strings $\rho_1, \ldots, \rho_M$ 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 $\rho_1, \ldots, \rho_M$ and their minimal programs $\sigma_1, \ldots, \sigma_M$ (and hence $QC(\rho_i) = \ell(\sigma_i)$). Let $\$^k$ 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 $\tilde{\mathcal{E}}^k = \$^k(\mathcal{E}_\sigma) = \{(\tilde{\rho}_i^k, \frac{1}{M})\}$, with $\tilde{\rho}_i^k = \$^k(\sigma_i)$.

By the monotonicity of Theorem 3 we know that for every k, $\chi(\tilde{\mathcal{E}}^k) \leq \chi(\mathcal{E}_\sigma)$. The chi factor 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, $\lim_{k\to\infty} \text{Fidelity}(\rho_i, \tilde{\rho}_i^k) = 1$, 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 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 1 *For every length n, there is an incompressible classical string of length n.*

Proof: Apply Theorem 6 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 $l \geq n$. □

0-7695-0674-7/00 $10.00 © 2000 IEEE

Corollary 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 6: $l \geq S(\rho)$. 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 3 *For every length* n, *at least* $2^n - 2^{n-c}+1$ *qubit strings of length* n *have complexity at least* $n - c$.

Corollary 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* $\rho = \frac{1}{M}\sum_i |\phi_i\rangle\langle\phi_i|$.

6.3 The complexity of copies

A case where quantum Kolmogorov complexity behaves differently from classical Kolmogorov complexity is that, in general, the relation $C(x^m) \leq C(x) + O(\log m)$ does not hold, as we show below. We give an upper and a lower bound for the Kolmogorov complexity of X^m.

Theorem 7 $QC(X^m) \leq \log \binom{m+2^{QC(X)}-1}{2^{QC(X)}-1} + O(\log m) + O(\log QC(X))$.

Proof: First we sketch the proof, omitting the effect of the approximation. Consider any 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 make m runs of P_X.

Let l be the length of P_X; we call $\mathcal{H}$ the 2^l-dimensional Hilbert space. Consider $\mathcal{H}^m = \mathcal{H} \otimes \cdots \otimes \mathcal{H}$, the m-fold tensor product of $\mathcal{H}$. The symmetric subspace $\mathrm{Sym}(\mathcal{H}^m)$ is d-dimensional, where $d = \binom{m+2^l-1}{2^l-1}$. The state P_X^m sits in this symmetric subspace, and can therefore be encoded exactly using $\log d + O(\log m) + O(\log l)$ qubits, where the $O(\log m)$ and $O(\log l)$ terms are used to describe the rotation onto $\mathrm{Sym}(\mathcal{H}^m)$. Hence, the quantum Kolmogorov complexity of X^m is bounded from above by $\log d + O(\log m) + O(\log l)$ qubits.

For the full proof, we will need to take into account the effect of the imperfect fidelities of the different computations.

To achieve a fidelity of $1 - \frac{1}{k}$, we will compute m copies of the minimal program P_X to a fidelity of $1 - \frac{1}{4km}$. On each copy, we simulate the program with fidelity of $1 - \frac{1}{4km}$, and thus obtain the strings $\tilde{X}_i$ $(1 \leq i \leq m)$, each of which has (according to Lemma 2) fidelity $1 - \frac{1}{km}$ with the target string X. By Lemma 3 we get a total fidelity of at least $1 - \frac{1}{k}$.

We now proceed to the details of the proof. First we introduce some notation.

Assume that for some QTM M, $QC_M(X) \leq \ell(P_X) = l$, where P_X M-computes X (with fidelity $1 - \frac{1}{k}$ for any k.)

Let R be the rotation that takes qubit strings $X^m \in \mathrm{Sym}(\mathcal{H}^m)$ to qubit strings of length $\lceil \log(\dim(\mathrm{Sym}(\mathcal{H}^m))) \rceil$. More precisely, R is the rotation that takes the ith basis state of $\mathrm{Sym}(\mathcal{H}^m)$ to the ith classical basis state of the Hilbert space of dimension $2^{\lceil \log(\dim(\mathrm{Sym}(\mathcal{H}^m))) \rceil}$.

For any fidelity parameter δ, R^{-1} can be computed efficiently and to arbitrary precision. By that we mean that for any δ, there is a transformation R_δ^{-1} for which the following holds: Let $Z = R(X^m)$ for some $X \in \mathcal{H}$. If $\widetilde{X^m} = R_\delta^{-1}(Z)$, then for each i, the mixed state $\tilde{X}_i$ obtained from X by tracing out all components that do not correspond to the ith copy of X, is such that $\mathrm{Fidelity}(X, \tilde{X}_i) \geq 1 - \delta$.

We now define the program that witnesses the upper bound on $QC(X^m)$ claimed in the theorem.

Let M' be the quantum Turing machine that does the following on input $(Z, l, m, 1^k)$.

1. Computes $Z' = R_{1/4km}^{-1}(Z)$. (When Z is an m-proper input, which we specify below, then $Z' \approx Y^m$ for some $Y \in \mathcal{H}$.)
2. On each 'copy' $\tilde{Y}_i$ of Y, runs the QTM $M(\tilde{Y}_i, 1^{4km})$. (That is, $\tilde{Y}_i$ is the result of tracing out all but the positions of Z' that correspond to the ith block of l qubits.)

The input Z is an 'm-proper input' if for some Y, $Z = R(Y^m)$. (Note that Z is exactly $R(Y^m)$, not an approximation up to some fidelity.)

If we run the above QTM M' on input $(R(P_X^m), l, m, 1^k)$ then the output of M' is $M'(R(P_x^m), l, m, 1^k) = \widetilde{X^m} = \tilde{X}_1 \cdots \tilde{X}_m$. (Recall that l is the length of P_X.)

It remains to show the following claims.

Claim 1 $\mathrm{Fidelity}(\widetilde{X^m}, X^m) \geq 1 - \frac{1}{k}$.

Claim 2 *The length of the program above for* M' *is* $\leq \log d_{l,m} + O(\log l) + O(\log m)$, *where* $d_{l,m} = \binom{m+2^l-1}{2^l-1}$.

Claim 2 follows immediately from the fact that the total length of the inputs $R(P_X^m), l, m$ is $\log d + O(\log l) + O(\log m)$.

We prove Claim 1. Since we chose a precision $\delta = \frac{1}{4km}$ in step 1, $\forall i$, $\mathrm{Fidelity}(P_X, \tilde{Y}_i) \geq 1 - \frac{1}{4km}$. Furthermore, since the computation at step 2 introduces at most an error of $\frac{1}{4km}$, $\forall i$, $\mathrm{Fidelity}(X, \tilde{X}_i) \geq 1 - \frac{1}{km}$ (by Lemma 2.) Therefore by Lemma 3, $\mathrm{Fidelity}(\widetilde{X^m}, X^m) \geq (1 - \frac{1}{km})^m \geq 1 - \frac{1}{k}$. This completes the proof of Claim 1.

Claim 1 and Claim 2 together give us that $QC_{M'}(X^m) \leq \log d_{l,m} + O(\log l) + O(\log m) \leq$

0-7695-0674-7/00 $10.00 © 2000 IEEE

$\log d_{n,m} + O(\log n) + O(\log m)$, where n is the length of X and an upper bound on its complexity. By invariance, we can conclude that $QC(X^m) \leq \log d_{n,m} + O(\log n) + O(\log m) + O(1)$, which proves the theorem. □

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

Theorem 8 *For every m and n, there is an n-qubit state X such that $QC(X^m) \geq \log \binom{m+2^n-1}{2^n-1}$.*

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^m: $\mathcal{E} = \{(X^m, \mu)\}$, where $\mu^{-1} = \int_{X \in \mathcal{H}} dX$ is the appropriate normalization factor. The corresponding average state is calculated by the integral $\rho = \mu \int_{X \in \mathcal{H}} X^m dX$. This mixture is the totally mixed state in the symmetric subspace $\mathrm{Sym}(\mathcal{H}^m)$ (see Section 3 in [22]), and hence has entropy $S(\rho) = \log \binom{m+2^n-1}{2^n-1}$. Because all X^m are pure states, we can use Corollary 4 to prove the existence of a X for which $QC(X^m) \geq \log \binom{m+2^n-1}{2^n-1}$. □

6.4 Subadditivity

Consider the following subadditivity property of classical Kolmogorov complexity.

Proposition 6 *For any x and y, $C(x,y) \leq C(x) + C(y|x) + O(1)$.*

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 6 would contradict Theorem 8 (for $m = 2$).

A natural quantum extension of this result is as follows.

Proposition 7 *For any X, Y, $QC(X,Y) \leq QC(X,X) + QC(Y|X) + O(1)$.*

7 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 3. However, we believe that such a statement does 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.

We now prove a slightly weaker statement, for bounded-fidelity complexity.

Theorem 9 *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 $X = |\phi_1\rangle \otimes \cdots \otimes |\phi_n\rangle \in \mathcal{H}^n$,*

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

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

Proof: Fix ϵ, δ. Apply Theorem 2 with $\epsilon' = \frac{\epsilon}{4}, \delta' = \delta$, and let $n = n(\epsilon', \delta')$ be the value from the theorem. Let $|\phi_1\rangle \otimes \cdots \otimes |\phi_n\rangle \in \mathcal{H}^n$ be the string for whose quantum Kolmogorov complexity we want to give an upper bound. By Theorem 2, item 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, item 2, and Lemma 2, we have that the fidelity of the encoding is at least $1 - \epsilon$. That completes the proof. □

8 Extensions and Future Work

We have argued that the QC of Definition 8 is a robust notion of Kolmogorov complexity for the quantum setting. Nevertheless, it would be interesting to see if an invariance theorem can be shown for the ideal quantum Kolmogorov complexity of Definition 6.

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.

9 Acknowledgements

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

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, and the EU fifth framework project QAIP IST-1999-11234.

References

[1] Adriano Barenco, André Berthiaume, David Deutsch, Artur Ekert, Richard Jozsa, and Chiara Macchiavello, "Stabilisation of Quantum Computations by Symmetrisation", *SIAM Journal on Computing*, Volume 26, No. 5, pp. 1541–1557 (1997)

[2] Ethan Bernstein and Umesh Vazirani, "Quantum Complexity Theory", *SIAM Journal on Computing*, Volume 26, No. 5, pp. 1411–1473 (1997)

[3] Gregory Chaitin, "On the length of programs for computing finite binary sequences", Journal of the ACM, Volume 13, No. 4, pp. 547–569 (1966)

[4] Thomas M. Cover and Joy A. Thomas, *Elements of Information Theory*, John Wiley & Sons, Wiley Series in Telecommunications (1991)

[5] David 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)

[6] Christopher A. Fuchs and Jeroen van de Graaf, "Cryptographic Distinguishability Measures for Quantum Mechanical States", *IEEE Transactions on Information Theory*, Volume 45, No. 4, pp. 1216–1227 (1999)

[7] Michael R. Garey and David S. Johnson, *Computers and Intractibility, A Guide to the Theory of NP Completeness*, W.H. Freeman (1979)

[8] Alexander 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)

[9] Richard Jozsa, Michał Horodecki, Paweł Horodecki, Ryszard Horodecki, "Universal Quantum Information Compression", *Physical Review Letters*, Volume 81, pp. 1714–1717 (1998)

[10] Richard Jozsa and Benjamin Schumacher, "A New Proof of the Quantum Noiseless Coding Theorem", *Journal of Modern Optics*, Volume 41, pp. 2343–2349 (1994)

[11] Andrei K. Kolmogorov, "Three approaches to the quantitative definition of information", Problems of Information Transmission, Volume 1, pp. 1–7 (1965)

[12] Ming Li and Paul Vitányi, *An Introduction to Kolmogorov Complexity and its Applications*, Second Edition, Springer Verlag (1997)

[13] Göran Lindblad, "Completely Positive Maps and Entropy Inequalities", *Communications in Mathematical Physics*, Volume 40, pp. 147–151 (1975)

[14] Michael Nielsen, PhD thesis, University of New Mexico, 1998

[15] John Preskill, "Quantum Computing" (1998); course notes available at URL: `http://www.theory.caltech.edu/people/preskill/ph229/`

[16] Benjamin Schumacher, "Quantum Coding", *Physical Review A*, Volume 51, No. 4, pp. 2738–2747 (1995)

[17] Claude E. Shannon and Warren Weaver, *The mathematical theory of communication*, University of Illinois Press (1949)

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

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

[20] Paul Vitányi, "Three Approaches to the Quantitative Definition of Information in an Individual Pure Quantum State", *Proceedings of the 15th Annual Conference on Computational Complexity* (2000) (these proceedings)

[21] Alfred Wehrl, "General Properties of Entropy", *Reviews of Modern Physics*, Volume 50, No. 2, pp. 221–260 (1978)

[22] Reinhard F. Werner, "Optimal cloning of pure states", *Physical Review A*, Volume 58, pp. 1827–1832 (1998)

[23] William K. Wootters and Wojceich H. Zurek, "A single quantum cannot be cloned", *Nature*, Volume 229, pp. 802–803 (1982)

0-7695-0674-7/00 $10.00 © 2000 IEEE