Natural Computing (2007) 6:133–150
DOI 10.1007/s11047-006-9017-5

Entropy and algorithmic complexity in quantum information theory

FABIO BENATTI
Department of Theoretical Physics, University of Trieste, Strada Costiera, 11, 34014, Trieste, Italy
(E-mail: benatti@ts.infn.it)

Abstract. A theorem of Brudno says that the entropy production of classical ergodic information sources equals the algorithmic complexity per symbol of almost every sequence emitted by such sources. The recent advances in the theory and technology of quantum information raise the question whether a same relation may hold for ergodic quantum sources. In this paper, we discuss a quantum generalization of Brudno's result which connects the von Neumann entropy rate and a recently proposed quantum algorithmic complexity.

Key words: algorithmic complexity, classical and quantum information

1. Introduction

Information is physical for it is encoded into, transmitted by and manipulated through physical actions on physical carriers; increasing chip miniaturization is leading to a point where it will be necessary and also profitable to treat them as quantum entities. Quantum mechanics must then be used to describe the behavior of the information carriers and their dynamics: in a fashionable jargon, this would call for moving from *bits* to *qubits* and from classical to quantum information (Nielsen and Chuang, 2000).

A classical bit, a stochastic variable capable of values 0 and 1 with certain probabilities p_0 and p_1, may physically be described by a *classical spin* that may only be found either up or down along a vertical direction, all other directions in space being precluded to it. The up and down states may be denoted in the usual bra/ket notation as $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

A quantum bit or qubit can instead be considered as a quantum spin 1/2 system pointing in any direction in space; all possible superpositions of $\mid 0\rangle$, $\mid 1\rangle$, which are vectors in the Hilbert space $\mathbb{C}^2$, are then available as physical states: $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$, $|\alpha|^2 + |\beta|^2 = 1$.

In the following, we shall compare the relations between the rate at which entropy is produced by classical, respectively, quantum, ergodic sources, and the algorithmic complexity of the emitted strings of bits, respectively, qubits.

Classical algorithmic complexity theory as initiated by Kolmogorov (1965, 1968), Chaitin (1966) and Solomonoff (1964) aimed at giving firm mathematical ground to the intuitive notion of randomness. The idea is that random strings cannot have short descriptions by means of binary programs that run by Universal Turing Machines (UTMs) halt with these strings as outputs. Such an approach is on the one hand equivalent to Martin-Löf's which is based on the notion of *typicalness* (Uspenskii et al., 1990), and is on the other hand intimately connected with the notion of entropy. The latter relation is best exemplified in the case of infinitely long strings: by taking the ratio of the complexity with respect to the number of bits, one gets a *complexity per symbol* which a theorem of Brudno (1983) shows to be equal to the *entropy per symbol* of almost all sequences emitted by ergodic sources.

The fast development of quantum information and computation, with the formalization of the concept of Universal Quantum Turing Machines (UQTMs), quite naturally brought with itself the need of extending the notion of algorithmic complexity to the quantum setting. Within such a broader context, the ultimate goal is again a mathematical theory of the randomness of quantum objects. There are two possible algorithmic descriptions of qubit strings: either by means of bit-programs or of qubit-programs.

In the following, we consider a qubit-based *quantum algorithmic complexity*, constructed in terms of quantum descriptions of quantum objects. Based on Benatti et al. (2006), in Section 5 we show that Brudno's result essentially also holds in the quantum setting. In order to do that, classical algorithmic complexity theory is reviewed in Section 2, the notions of quantum ergodic sources and their von Neumann entropy (rate) are introduced in Section 3, while Quantum Turing Machines (QTM) and *quantum algorithmic complexity* are addressed in Section 4.

2. Classical information and algorithmic complexity

Both classical information and classical algorithmic complexity theory deal with sources emitting strings of bits $\mathbf{i}^{(n)} = i_1 i_2 \ldots\ i_n \in \{0,\ 1\}^n =: \Omega_2{}^{(n)}$ with probabilities $p(\mathbf{i}^{(n)})$ constituting stationary stochastic processes. Stationarity means that the probability of a string as $\mathbf{i}^{(n)}$ to be emitted does not depend on when the first bit i_1 is emitted. The family of probability distributions $\pi^{(n)} = \{p(\mathbf{i}^{(n)})\}$ over $\Omega_2{}^{(n)}$ gives rise to a probability measure π over all bit sequences in Ω_2. We shall denote by $(\Omega_2,\ \pi)$ any such stationary classical source and by $(\Omega_2{}^{(n)},\ \pi^{(n)})$ the statistical ensembles of strings of length n. In the following we shall focus upon stationary processes which are also ergodic: this means that the stationary probability distribution π cannot be decomposed into a convex combination of other stationary probability distributions.

A helpful physical picture associated with $(\Omega_2,\ \pi)$ is that of a classical ferromagnet consisting of infinitely many spins located at the integer sites along the x axis and pointing either up or down along the z axis. The measure π can then be interpreted as a state of the ferromagnet which gives probabilities to all local spin-configurations corresponding to the strings in $\Omega_2{}^{(n)}$, the other spins being either all up or all down.

Each statistical ensemble $(\Omega_2{}^{(n)},\ \pi^{(n)})$ of local configurations has Shannon entropy $H(\pi^{(n)}) = -\sum_{\mathbf{i}(n)} p(\mathbf{i}^{(n)}),\ \log_2 p(\mathbf{i}^{(n)})$, and the state of the infinite classical source (ferromagnet) is characterized by an entropy rate, or entropy density, or entropy production (Billingsley, 1965; Cover and Thomas, 1991)

$$s(\pi) := \lim_{n\to\infty} \frac{1}{n} H(\pi^{(n)}). \qquad (1)$$

The entropy rate gives a measure of the stochasticity of the source on the average, namely with respect to the given state π. Instead, according to Kolmogorov (1965), the purpose of algorithmic complexity theory is to assign a measure of irregularity to single strings or sequences and not to statistical ensembles of them. Kolmogorov's definition of the complexity of single objects was in terms of the difficulty of their reproducibility by algorithms run by universal computers as Turing machines (see Li and Vitanyi, 1997; Calude, 2002).

> The complexity of a bit string is the minimum length of a program for a Turing machine (TM) that produces the string.

The idea behind such a definition is easily captured by the following example: if one wants a computer to reproduce a list of N 1's, one has to issue a command of the form **PRINT** 1 N **TIMES**. Apart from the printing instructions, the increase of the number of bits in such a program is for large N determined by log N, where, from now onward, the logarithm is to be taken in base 2.

On the other hand, if one wants a computer to reproduce the N outcomes of a fair coin tossing, very rapidly the string grows patternless and the only thing to do is to tell the computer the digit to write one after the other. The length of such a program increases with N for large N.

On the whole, structureless strings offer no catch for writing down short programs that fed into a computer would get the given strings as outputs. The intuitive notion of random strings is thus mathematically characterized by the fact that, for large n, the shortest programs that reproduce them cannot do better than literal transcription.

More in detail, the algorithmic complexity $K(\mathbf{i}^{(n)})$ of a string $\mathbf{i}^{(n)} \in \Omega_2{}^{(n)}$ is the length (counted in the number of bits) of the shortest program p that run on a UTM $\mathfrak{U}$ yields the string as output, i.e. $\mathfrak{U}(p) = \mathbf{i}^{(n)}$. For infinite sequences $\mathbf{i} \in \Omega_2$, in analogy with the entropy rate, one defines the *complexity rate* as

$$k(\mathbf{i}) := \lim_n \frac{1}{n} K(\mathbf{i}^{(n)}), \tag{2}$$

where $\mathbf{i}^{(n)}$ is the string consisting of the first n bits of $\mathbf{i}$ (Alekseev and Yakobson, 1981).

A (classical) Turing Machine (TM) consists of a head moving along an infinite tape divided into cells carrying the symbols $\Sigma := \{0,1,\#\}$, # staying for blank, which can read, erase and replace them and move one step to the left or the right. The head is also equipped with a finite set of internal control states $Q = \{q_1, q_2, \ldots, q_f\}$. What the head does when it is positioned on a cell of the tape is determined by a program, that is by a list of instructions telling the head, on reading $s_i \in \Sigma$, when its internal state is q_i, which new symbol $s_f \in \Sigma$ to leave in the cell before moving a step $p \in \{L,R\}$ to the left or right and which new control state $q_f \in Q$ to assume.

The program is thus a list of instructions which specify which 2-ples $(q,\ s) \in Q \times \Sigma$ have to be associated with which 3-ples $(q,\ s,\ p) \in Q \times \Sigma \times \{L,\ R\}$; as such it can be represented by a map

$$\delta : (Q \times \Sigma) \times (Q \times \Sigma \times \{L, R\}) \mapsto \{0, 1\} \qquad (3)$$

where the value 1 means that the first two arguments determine the second three, whereas the value 0 means that the whole argument does not appear in the list of instructions. It turns out that there exist UTMs capable of acting as any of the TMs in the list.

In line of principle, the complexity of a given string may heavily depend on the machine on which a program is run which reproduces that string. However, the difference in the complexity of a given string upon changing UTM is bounded by a constant independent of the string; it follows that the complexity rate $k(\mathbf{i})$ in (2) is UTM-independent.

Intuitively, one expects a connection between the randomness of single strings and the average randomness of ensembles of strings. This is rigorously established for classical ergodic sources $(\Omega_2\ \pi)$, by a theorem of Brudno (1983) stating that the entropy rate equals the algorithmic complexity per symbol of almost all emitted bit strings:

$$k(\mathbf{i}) = s(\pi) \quad \pi - \text{a.e.} \qquad (4)$$

3. Quantum sources

Quantum sources can be thought as black boxes emitting strings of qubits instead of bits. The simplest quantum source is a black box emitting, at each stroke of time k, uncorrelated qubits in states $|\psi_{i_k}\rangle \in \mathbb{C}^2$, $i_k = \{1, 2, \ldots, \ell\}$, with probabilities $p(i_k)$.

Any such source is the quantum version of a classical Bernoulli source: at each click of the clock the statistical ensemble of emitted qubits is described by the mixed state or density matrix consisting of the mixture with weights $p(i)$ of the projectors $|\psi_i\rangle\langle\psi_i|$ onto the (normalized, but not necessarily orthogonal) states $|\psi_i\rangle$: $\rho := \sum_{i \in I} p(i) |\psi_i\rangle\langle\psi_i|$. The Bernoulli-like independence of the subsequent emissions is mathematically characterized by the fact that, after n clicks of the clock, the statistics of the emitted qubits is described by a density matrix which is the tensor product of ρ:

$$\rho^{\otimes n} := \sum_{\mathbf{i}^{(\mathbf{n})} \in I^n} p(\mathbf{i}^{(n)}) |\psi_{i_1} \otimes \psi_{i_2} \otimes \cdots \psi_{i_n}\rangle\langle\psi_{i_1} \otimes \psi_{i_2} \otimes \cdots \psi_{i_n}|$$

Physically speaking, such a quantum source as above has the structure of a quantum spin chain (quantum ferromagnet) consisting of uncorrelated qubits at each of its sites, the density matrices $\rho^{(n)}$ over n sites describing local states. However, typical ergodic states of quantum spin-chains have richer structures that could be used as quantum sources: the local states $\rho^{(n)} \neq \rho^{\otimes n}$, not anymore tensor products, would describe emitted n-qubit strings which are generic correlated density matrices.

We now briefly formalize the notion of quantum spin chains (Alicki and Fannes, 2001).

Given the subset $\mathbb{N}$ of positive integers along the real line, each site $i \in \mathbb{N}$ accommodates the algebra $(\mathcal{A})_i = M_2(\mathbb{C})$ of 2×2 complex matrices. To the spin configuration in the finite subset $[1, n] \subset \mathbb{N}$ it is associated the $2^n \times 2^n$ matrix algebra $\mathcal{A}_{[1,n]} := \bigotimes_{i \in [1,n]} \mathcal{A}_i$, that is the algebra of $2^n \times 2^n$ complex matrices acting on the Hilbert space $\mathbb{C}^{2^n} = \underbrace{\mathbb{C}^2 \otimes \mathbb{C}^2 \cdots \otimes \mathbb{C}^2}_{n\,\text{times}}$. Each such algebra can be embedded as a *local* sub-algebra into an infinite tensor product $\mathcal{A}^\infty$ by mapping $a \in \mathcal{A}_{[1,n]}$ into $a \otimes \bigotimes_{k=n+1}^{+\infty} (\mathbf{1})_k$. In such a context, $\mathcal{A}_{[1,n]}$ will be the sub-algebras corresponding to qubit strings of length n, while the quasi-local C^* algebra $\mathcal{A}^\infty$ will describe one-sided qubit sequences.

Stationary classical sources are essentially described by the statistics of emitted bit strings of any given length n independently of at which stroke of time the first of their bits has been emitted. Of course, independently of stationarity, given the probability distribution $\pi^{(n+1)}$ of strings of length $n + 1$, one gets that of the subset of strings of length n, $\pi^{(n)}$, by summing over the $n + 1$-th bit.

Analogously, if the local algebras $\mathcal{A}_{[1,n]}$ are equipped with density matrices $\rho^{(n)}$ such that $\rho^{(n)} = \mathrm{Tr}_1 \rho^{(n+1)} = \mathrm{Tr}_{n+1} \rho^{(n+1)}$, where Tr_j denotes the trace-operation over the degrees of freedom of the j-th spin, then the sequence of local states $\rho^{(n)}$ defines a stationary state Ψ on $\mathcal{A}^\infty$ such that, for all $a \in \mathcal{A}_{[1,n]}$, $\Psi(a) = \mathrm{Tr}_{[1,n]}\left(\rho^{(n)} a\right)$, where the trace is to be performed with respect to any orthonormal basis of the the Hilbert space $\mathbb{C}^{2^n}$ on which the local density matrices $\rho^{(n)}$ act. As much as classical sources are described by 2-uples (Ω_2, π), quantum sources will be described by 2-ples $(\mathcal{A}^\infty, \Psi)$.

Continuing with classical--quantum analogies, to the entropy rate of stationary classical sources there corresponds the von Neumann entropy rate of stationary quantum sources. Given the local density matrices $\rho^{(n)}$, their degree of mixedness is well measured by the von Neumann entropy $S(\rho^{(n)}) : -\mathrm{Tr}_{[1,n]}(\rho^{(n)} \log \rho^{(n)})$. Then, the following limit, the quantum entropy rate, exists (Alicki and Fannes, 2001)

$$s(\Psi) := \lim_{n\to\infty} \frac{1}{n} S(\rho^{(n)}). \tag{5}$$

Exactly as in the classical case, those stationary states on $\mathcal{A}^\infty$ which cannot be written as convex combinations of other stationary states are called ergodic. In the following, we shall consider ergodic quantum sources $(\mathcal{A}^\infty, \Psi)$. Ergodicity of Ψ has an important consequence concerning the minimum dimension of the Ψ-typical subspaces. These are finite dimensional subspaces associated with orthogonal projections $q \in \mathcal{A}_{[1,n]}$ supporting Ψ almost entirely: $\mathrm{Tr}(\rho^{(n)}q) \geq 1 - \varepsilon$, with $0 < \varepsilon \ll 1$.

Proposition 3.1 *If* $(\mathcal{A}^\infty, \Psi)$ *is an ergodic quantum source with entropy rate* s *and* $\beta_{\epsilon,n}(\Psi) := \min\{\log \mathrm{Tr}_n(q) |\ q \in \mathcal{A}_{[1,n]}$ *projector*, $\mathrm{Tr}(\rho^{(n)}q) \geq 1 - \epsilon\}$, *then* $\lim_{n\to\infty} \frac{1}{n} \beta_{\varepsilon,n}(\Psi) = s$ *for every* $0 < \varepsilon < 1$.

In other words, if a sequence of projections $p_n \in \mathcal{A}_{[1,n]}$ is such that $\lim_n 1/n \log \mathrm{Tr}(p_n) < s$, then $\lim_n \mathrm{Tr}(\rho^{(n)} p_n) = 0$. This is a quantum *Asymptotic Equipartition Property* (Bjelaković et al., 2004): if a projection $q \in \mathcal{A}_{[1,n]}$ almost supports Ψ, its associated subspace has dimension $\simeq 2^{n\ s}$.

4. QTMs and quantum algorithmic complexity

Algorithmic complexity measures the degree of randomness of a single object. It is defined as the minimal description length of the object, relative to a certain "machine", classically a UTM. In order to properly introduce a quantum counterpart to it, the setting is as follows: the objects to describe will be density operators in the local subalgebras $\mathcal{A}_{[1,n]}$ corresponding to n-qubit strings; the algorithms will themselves be quantum, i.e. density operators on a suitable Hilbert space; the reference machines will be QTMs as defined by Bernstein and Vazirani (1997), in particular universal QTMs.

4.1. *Quantum turing machines*

The detailed construction of UQTMs can be found in (Adleman et al., 1997; Bernstein and Vazirani, 1997): these machines work analogously to classical TMs, that is they consist of a read/write head, a set of internal control states and a tape. However, the transition functions among the machine's configurations (the programs or quantum algorithms, compare (3)) are given in terms of probability amplitudes, implying the possibility of linear superpositions of the machine's configurations:

$$\delta : Q \times \Sigma \times Q \times \Sigma \times \{L, R\} \to \tilde{\mathbb{C}} \cap \{z \in \mathbb{C} : |z| \leq 1\}, \tag{6}$$

where $\tilde{\mathbb{C}}$ denotes the set of all $z \in \mathbb{C}$ that can be computed by a (classical deterministic) algorithm up to accuracy 2^{-n} in polynomial time in n, such that the resulting global time evolution is unitary.

Physically speaking, a QTM consists of a doubly-infinite tape of cells indexed by $\mathbb{Z}$ and a single read/write tape head that moves along the tape. A *state* of the QTM is a quantum state on a suitable Hilbert space $\mathcal{H}_{\mathrm{QTM}}$ spanned by the (classical) *configurations* of the QTM consisting of those (finitely many) tape cells containing non-blank symbols from Σ, the control state $q \in Q$ and the head position $k \in \mathbb{Z}$, denoting the index of the cell that the head is reading.

The QTM will thus conveniently be described by the Hilbert space $\mathcal{H}_{\mathrm{QTM}} := \mathbb{C}^Q \otimes \mathcal{H}_{\mathrm{in}} \otimes \mathcal{H}_{\mathrm{out}} \otimes \ell^2(\mathbb{Z})$, where $\mathbb{C}^Q$ is the Hilbert space pertaining to the q states of the control, $\ell^2(\mathbb{Z})$ is the Hilbert space relative to the *head position* **H**, classically an integer in $\mathbb{Z}$, here a state on $\ell^2(\mathbb{Z})$, while $\mathcal{H}_{\mathrm{in,out}}$ are isomorphic Hilbert spaces describing the tape further subdivided into an input, **I**, and output, **O**, track. At any finite time, only a finite set $T := \left\{(x_i)_{i\in\mathbb{Z}} \in \{0, 1, \#\} | x_i \neq \# \text{ for } \textit{finitely} \text{ many } i \in \mathbb{Z}\right\}$ of cells on any track can contain non-blank symbols, thus $\mathcal{H}_{\mathrm{in}} = \mathcal{H}_{\mathrm{out}} := \ell^2(T)$.

As for classical TMs, we only allow inputs consisting entirely of 0's and 1's, that is, they do not contain any blank symbol #, but they can have arbitrary length. Thus, we will choose the input being a state on the *Fock space* $\mathcal{H}_{\mathrm{Fock}} := \bigoplus_{k=0}^{\infty} \mathcal{H}_k$, where $\mathcal{H}_k$ is the linear space spanned by basis-vectors of the form $|\mathbf{i}^{(k)}\rangle, \mathbf{i}^{(k)}$ being a string in $\{0,1\}^k$.

Let $\mathcal{H}_{\leq n} := \bigoplus_{k=0}^{n} \mathcal{H}_k$ denote the subspaces of dimension $2^{n+1} - 1$. We define the *length* $\ell(\sigma) \in \mathbb{N}_0 \cup \{\infty\}$ of a qubit string $\sigma \in \mathcal{T}_1^+(\mathcal{H}_{\mathrm{Fock}})(\mathcal{T}_1^+(\mathcal{H})$ denoting the space of density matrices on $\mathcal{H})$ by

$$\ell(\sigma) := \min\{n \in \mathbb{N}_0 | \sigma \in \mathcal{T}_1^+(\mathcal{H}_{\leq n})\}. \tag{7}$$

To use these qubit strings $\sigma \in \mathcal{T}_1^+(\mathcal{H}_{\text{Fock}})$ as inputs (and outputs) for a QTM, they are first written onto the input track interval $[0, \ell(\sigma) - 1]$ by means of an embedding map e:$\{0,1\}^* \rightarrow$ T. Running a QTM M on an input $\sigma \in \mathcal{T}_1^+(\mathcal{H}_{\text{Fock}})$ means that the machine will start at time step $t = 0$ in the initial state $M^0(\sigma) := \underbrace{|q_o\rangle\langle q_o|}_{\mathbf{C}} \otimes \underbrace{e(\sigma)}_{\mathbf{I}} \otimes \underbrace{|\#\rangle\langle\#|}_{\mathbf{O}} \otimes \underbrace{|0\rangle\langle 0|}_{\mathbf{H}}$, i.e. the input σ is written onto the input track cells $[0, \ell(\sigma) - 1]$ (surrounded by blank symbols # on both sides), the output track is in the vector state $|\#\rangle$, i.e. it contains only blank symbols, the head position starts at position 0, and the control is in the initial state q_0.

Let $M^t(\sigma) \in \mathcal{T}_1^+(\mathcal{H}_{\text{QTM}})$ be the quantum state at time $t \in \mathbb{N}_0$ of the machine M as a whole, given the state $\sigma \in \mathcal{T}_1^+(\mathcal{H}_{\text{Fock}})$ as input. The discrete time steps are given by a unitary time evolution operator V_M, such that $M^t(\sigma) = (V_M)^t \ M^0(\sigma) \ ({V_M}^*)^t$, where V_M is derived from the transition function δ, given in the definition of the QTM, see Bernstein and Vazirani (1997).

The states of the different subsystems of M are given by the partial trace over all the other parts of the machine. For example, the state of the output track of M at time t is given by $M^t_{\mathbf{O}}(\sigma) = \text{Tr}_{\mathbf{C,I,H}}(M^t(\sigma))$. We still have to specify what we mean for the QTM to halt and to output a certain qubit string and we should do it in a way that make the action of the QTM correspond to a so-called *quantum operation*, that is to a completely positive and trace-preserving map on $\mathcal{T}_1^+(\mathcal{H}_{\text{Fock}})$. There is an obstacle to this program, namely that a given QTM may have different halting times for different inputs. This obstacle can be overcome by means of the halting and output conventions in Bernstein and Vazirani (1997) as summarized in the following definition and proposition.

Definition 4.1 (Halting and Output Conventions) A QTM M halts at time $t \in \mathbb{N}$ on input σ, if $\langle q_f | M^t_{\mathbf{C}}(\sigma) | q_f \rangle = 1$ and $\langle q_f | M^{t'}_{\mathbf{C}}(\sigma) | q_f \rangle = 0$ for every $t' < t$, where $q_f \in Q$ is the special state of the control (specified in the definition of M) signalling the halting of the computation.

M has proper output, if the output track at halting time t contains a proper qubit string ρ, starting in cell 0 and ending in cell $\ell(\rho) - 1$, such that all the other cells are blank, that is $e^{-1}({M_{\mathbf{O}}}^t(\sigma))$ is defined in $\mathcal{T}_1^+(\mathcal{H}_{\text{Fock}})$, where the embedding e has been defined after Eq. (7).

Let $M : \mathcal{T}_1^+(\mathcal{H}_{\text{Fock}}) \longrightarrow \mathcal{T}_1^+(\mathcal{H}_{\text{Fock}})$ be the partial map defined as follows: if M halts at time $t \in \mathbb{N}$ on input σ and has proper output, then $M(\sigma)\!: = e^{-1}({M_{\mathbf{O}}}^t(\sigma))$. The latter will be called the output of M on input σ, otherwise $M(\sigma)$ is undefined.

Lemma 4.2 (QTMs are Quantum Operations) (Benatti et al., 2006). *For every QTM M and every $c \in \mathbb{N}$, there is a quantum operation $\mathcal{M}_c : \mathcal{T}_1^+(\mathcal{H}_{\leq c}) \to \mathcal{T}_1^+(\mathcal{H}_{\text{out}})$, such that*

$$M(\sigma) = e^{-1}(\mathcal{M}_c(\sigma))$$

for every $\sigma \in \mathcal{T}_1^+(\mathcal{H}_{\leq c})$ for which the action $M(\sigma)$ of the QTM is defined.

4.2. *Quantum algorithmic complexity*

Given the theoretical possibility of universal computing machines working in agreement with the quantum rules, it was a natural step to extend the problem of algorithmic descriptions to the quantum case.

Contrary to the classical case, where different formulations are equivalent, several inequivalent possibilities are available in the quantum setting. In the following, we shall use the definitions in (Berthiaume et al. 2001) which, roughly speaking, say that the algorithmic complexity of a qubit string ρ is the logarithm in base 2 of the dimension of the smallest Hilbert space (spanned by computational basis vectors) containing a quantum state that, once fed into a UQTM, makes the UQTM compute the output ρ and halt.

In general, quantum states cannot be perfectly distinguished; it thus makes sense to allow some tolerance in the accuracy of the machine's output, the accuracy of the reproduction being measured by the following distance between density matrices:

$$D(\rho,\sigma) := \frac{1}{2}\mathrm{Tr}|\rho - \sigma|. \tag{8}$$

The typical case we want to study is the (approximate) reproduction of a density matrix $\rho \in \mathcal{T}_1^+(\mathcal{A}^{(n)})$ by a QTM M. This means that there is a "quantum program" $\sigma \in \mathcal{T}_1^+(\mathcal{H}_{\text{Fock}})$, such that $M(\sigma) \approx \rho$, in the sense that $D(M(\sigma),\rho) \ll 1$.

We are particularly interested in the case that the program σ is shorter than ρ itself, i.e. that its length $\ell(\bullet)$ in (7) satisfy $\ell(\sigma) < \ell(\rho)$. On the whole, the minimum possible length $\ell(\sigma)$ for ρ will be defined as the *quantum algorithmic complexity* of ρ.

Definition 4.3 (Quantum Algorithmic Complexity) Let M be a QTM and $\rho \in \mathcal{T}_1^+(\mathcal{H}_{\text{Fock}})$ a qubit string. An approximation-scheme quantum complexity $QC_M^{\searrow 0}$ is defined by the minimal length $\ell(\sigma)$ of any density operator $\sigma \in \mathcal{T}_1^+(\mathcal{H}_{\text{Fock}})$, such that when given M as input together with any integer k, the output $M(k, \sigma)$ has trace distance from ρ smaller than $1/k$:

$$QC_M^{\searrow 0}(\rho) := \min\left\{\ell(\sigma) \,\middle|\, (\rho, M(k,\sigma)) \leq \frac{1}{k} \text{ for every } k \in \mathbb{N}\right\} \qquad (9)$$

. Some points are worth stressing:

- The *same* qubit program σ is accompanied by a classical specification of an integer k, which tells the program to what accuracy the computation of the output state must be accomplished.
- A noiseless transmission channel (implementing the identity transformation) between the input and output tracks can always be realized: this corresponds to classical literal transcription, so that automatically $QC_M^{\searrow 0}(\rho) \leq \ell(\rho) + c_M$ for some constant c_M. Of course, the key point in classical as well as quantum algorithmic complexity is that there are sometimes much shorter qubit programs than literal transcription.
- For the accuracy specification $\frac{1}{k}$ is not important; we can choose any computable function that tends to zero for $k \to \infty$, and we will always get an equivalent definition (in the sense of being equal up to some constant).
- In Bernstein and Vazirani (1997), it is proved that there is a universal QTM (UQTM) $\mathfrak{U}$ that can simulate with arbitrary accuracy every other machine M. Like in the classical case, one can thus fix $\mathfrak{U}$ for the rest of the paper and consider $QC^{\searrow 0}(\rho) := QC_{\mathfrak{U}}^{\searrow 0}(\rho)$.

5. Quantum Brudno's theorem

It turns out that the rates of the complexity $QC^{\searrow 0}$ of the typical pure states of qubit strings generated by an ergodic quantum source $(\mathcal{A}^\infty, \Psi)$ are asymptotically equal to the entropy rate $s(\Psi)$ of the source. A precise formulation of this result is the content of the following theorem. It can be seen as a quantum extension of Brudno's theorem as a convergence in probability statement, while the original formulation of Brudno's result is an almost sure statement.

Theorem 5.1 (Quantum Brudno Theorem) (Benatti et al., 2006) *Let* $(\mathcal{A}^{\infty}, \Psi)$ *be an ergodic quantum source with entropy rate s. For every* $\delta > 0$, *there exists a sequence of* Ψ*-typical projectors* $q_n(\delta) \in \mathcal{A}^{(n)}$, $n \in \mathbb{N}$, *i.e.* $\lim_{n\to\infty} \mathrm{Tr}\left(\rho^{(n)} q_n(\delta)\right) = 1$, *such that for every one-dimensional projector* $q \le q_n(\delta)$ *and n large enough* $\frac{1}{n} QC^{\searrow 0}(q) \in (s - \delta, s + \delta)$.

As for the classical Brudno's theorem (4), the proof of this theorem consists in proving appropriate lower and upper bounds.

5.1. *Lower bound*

For classical TMs, there are no more than $2^{c+1} - 1$ different programs of length $\ell \le c$. This can be used as a "counting argument" for proving the lower bound of Brudno's Theorem in the classical case (Chaitin, 1966). A similar statement holds for QTMs as it can be showed by elaborating on an argument due to (Berthiaume et al., 2001) which states that there cannot be more than $2^{\ell+1} - 1$ mutually orthogonal one-dimensional projectors p with quantum complexity $QC^{\searrow 0}(p) \le \ell$.

Proposition 5.2 (Quantum Counting Argument) *Let* $0 < \delta < 1/e$, $c \in \mathbb{N}$ *a natural number with* $c \ge \frac{1}{\delta}\left(4 + 2\log\frac{1}{\delta}\right)$, *h a linear subspace of an arbitrary Hilbert space* $\mathcal{H}'$, *and* $\mathcal{E} : \mathcal{T}_1^+(\mathcal{H}_{\mathrm{Fock}}) \to \mathcal{T}_1^+(\mathcal{H}')$ *a quantum operation. Let* ${N^{\delta}}_c$ *be a maximum cardinality subset of orthonormal vectors from the set*

$$A_c^{\delta}(\mathcal{E}, h) := \{\phi \in h : \|\phi\| = 1, \exists \sigma_{\phi} \in \mathcal{T}_1^+(\mathcal{H}_{\le c})\, s.t.\, D(\mathcal{E}(\sigma_{\phi}), |\phi\rangle\langle\phi|) \le \delta\},$$

of all normalized vectors in h which are reproduced within δ by the operation $\mathcal{E}$ *on some input of length* $\le$ *c. Then it holds that* $\log|N_c^{\delta}| < c + 1 + \frac{2+\delta}{1-2\delta}\delta c$.

The result of above puts a limit to the number of orthogonal vectors that can be δ-approximated by a same quantum operation on qubit strings of a fixed length. When the quantum operation $\mathcal{E}$ corresponds to a quantum algorithm computed by a QTM as in Section 4, Proposition 5.2 can then be used to show the following.

Corollary 5.3 (Lower Bound for $\frac{1}{n} QC^{\searrow 0}$) (Benatti et al., 2006) *Let* $(\mathcal{A}^{\infty}, \Psi)$ *be an ergodic quantum source with entropy rate s. Let* $(p_n)_{n\in\mathbb{N}}$

with $p_n \in \mathcal{A}^{(n)}$ be an arbitrary sequence of Ψ-typical projectors. Then, for every $0 < \delta < 1/e$, there is a sequence of Ψ-typical projectors $q_n(\delta) \leq p_n$ such that for n large enough $\frac{1}{n}QC^{\searrow 0}(q) > s - \delta$ is satisfied for every one-dimensional projector $q \leq q_n(\delta)$.

The proof is based on Proposition 3.1 and hinges upon the fact that the maximal number of orthogonal vectors in the typical subspace associated with p_n which violate the lower bound in the corollary cannot increase as $2^{n\ s}$, because of Proposition 5.2. The argument following Proposition 3.1 says that, when n is large enough, the projection $P_{\mathcal{O}_n}$ onto the subspace $\mathcal{O}_n$ generated by them is such that $\mathrm{Tr}\left(\rho^{(n)} P_{\mathcal{O}_n}\right) \leq \epsilon$. It is thus the orthogonal complement $\mathcal{O}_n^{\perp}$ which becomes typical; also, all its vectors must fulfil the lower bound, otherwise the maximality of the set spanning $\mathcal{O}_n$ would be violated.

5.2. *Upper bound*

According to the previous result, for large m, the quantum complexity rate $\frac{1}{m}QC^{\searrow 0} > s - \delta$ with high probability. We further show

Proposition 5.4 (Upper Bound) (Benatti et al., 2006) *Let $(\mathcal{A}^{\infty}, \Psi)$ be an ergodic quantum source with entropy rate s. Then, for every $0 < \delta < 1/e$, there is a sequence of Ψ-typical projectors $q_m(\delta) \in \mathcal{A}_{[1,m]}$ such that for every one-dimensional projector $q \leq q_m(\delta)$ and m large enough $\frac{1}{m}QC^{\searrow 0}(q) < s + \delta$.*

The upper bound is proved by explicitly providing a short quantum algorithm (with program length increasing like $m(s + \delta)$) that computes q within arbitrary accuracy. This can be done by means of *quantum universal typical subspaces* as constructed in (Kaltchenko and Yang, 2003), the essential point being that in this way one needs no *a priori* information about the state Ψ on $\mathcal{A}^{\infty}$.

Proposition 5.5 (Universal Typical Subspaces) *Let $s > \ 0$ and $\varepsilon > 0$. There exists a sequence of projectors $Q_{s,\varepsilon}^{(n)} \in \mathcal{A}_{[1,n]}$, $n \in \mathbb{N}$, such that for n large enough $\mathrm{Tr}(Q^{(n)}{}_{s,\varepsilon}) \leq 2^{n(s+\varepsilon)}$ and for every ergodic quantum state Ψ on $\mathcal{A}^{\infty}$ with entropy rate $s(\Psi) \leq \ s$ it holds that $\lim_{n\to\infty}$ $\mathrm{Tr}\left(\rho^{(n)} Q_{s,\varepsilon}^{(n)}\right) = 1$.*

Let $0 < \varepsilon < \delta/2$ be such that $r := s + \varepsilon$ is rational, and let $q_m := Q_{s,\varepsilon}{}^{(m)}$ be the universal projector sequence of the previous theorem. Such a sequence is *independent* of the choice of the ergodic state Ψ, as long as its entropy rate $s(\Psi) \leq s$.

Since $\mathrm{Tr}(Q^{(n)}_{s,\varepsilon}) \leq 2^{n(s+\varepsilon)}$, for m large enough, there exists some unitary transformation U^* that transforms every one-dimensional projector $q \leq q_m$ into a qubit string $\tilde{q} := U^* q U$ of length $\ell(\tilde{q}) = \lceil mr \rceil$ projecting onto a vector in $\mathcal{H}_{\lceil mr \rceil}$.

As shown in Bernstein and Vazirani (1997), a UQTM can implement every classical algorithm, and it can apply every unitary transformation U (when given an algorithm for the computation of U) on its tapes within any desired accuracy. We can thus feed $\tilde{q}$, plus some classical instructions including a subprogram for the computation of U, as input into the UQTM $\mathfrak{U}$. This UQTM starts by computing a classical description of the transformation U, and subsequently applies U to $\tilde{q}$, recovering the original projector $q = U\tilde{q}U^*$ on the output tape. Since $U = U(q_m)$ depends on Ψ only through its entropy rate $s(\Psi)$, the subprogram that computes U does not have to be supplied with additional information on Ψ and will thus have fixed length independent of m.

The qubit program that reproduces q thus amounts to a quantum decompression algorithm $\mathfrak{A}$, which is, formally, a mapping (r is rational)

$$\mathfrak{A} : \mathbb{N} \times \mathbb{N} \times \mathbb{Q} \times \mathcal{H}_{Fock} \to \mathcal{H}_{Fock}, \qquad (k, m, r, \tilde{q}) \mapsto q = \mathfrak{A}(k, m, r, \tilde{q}),$$

where k is the integer appearing in the Definition 4.3 of the quantum algorithmic complexity, m is the index of the universal projector, $r = s + \varepsilon$ is almost the entropy rate and $\tilde{q}$ is the rotated projection to be reproduced as output of $\mathfrak{A}$.

The parameters $(m, r, \tilde{q})$ are encoded into a single qubit string σ, while, according to the definition of $QC^{\searrow 0}$, the parameter k is not a part of σ, but is given as a second parameter. For instance, m can be encoded by giving $\lfloor \log m \rfloor$ 1's, followed by one 0, followed by the $\lfloor \log m \rfloor + 1$ binary digits of m. Let $|M\rangle\langle M|$ denote the corresponding projector in the computational basis. The parameter r can be encoded in any way, since it does not depend on m. The only constraint is that the description must be self-delimiting, i.e. it must be clear and decidable at what position the description for r starts and ends. The descriptions will also be given by a computational basis vector (or rather the corresponding projector) $|R\rangle\langle R|$.

The descriptions are then stuck together, and the input $\sigma(\tilde{q})$ is given by $\sigma(\tilde{q}) := |M\rangle||R\rangle\langle R| \otimes \tilde{q}$. Choosing m large enough such that

$\log\mathrm{Tr}(Q^{(m)}{}_{s,\varepsilon}) \leq m(s+\varepsilon)$, it follows that $\ell(\sigma(\tilde{q})) = 2\lfloor\log m\rfloor + 2 + c + \lceil mr\rceil$, where $c \in \mathbb{N}$ is some constant which depends on r, but not on m.

Now, $\sigma(\tilde{q})$ can be fed into the reference UQTM $\mathfrak{U}$ together with a description of the algorithm $\mathfrak{A}$ of fixed length c' which depends on r, but not on m. This will give a qubit string $\sigma_{\mathfrak{U}}(\tilde{q})$ of length

$$\begin{aligned}\ell(\sigma_{\mathfrak{U}}(\tilde{q})) &= 2\lfloor\log m\rfloor + 2 + c + \lceil mr\rceil + c'\\ &\leq 2\log m + m\left(s + \frac{1}{2}\delta\right) + c'',\end{aligned}$$

where c'' is again a constant which depends on r, but not on m.

Also, the matrix U which rotates (decompresses) a compressed (short) qubit string $\tilde{q}$ back into the typical subspace can be constructed in a way that it can achieve the required accuracy $D\big(\mathfrak{U}(\sigma_u(\tilde{q}), k), q\big) < \frac{1}{k}$, for every $k \in \mathbb{N}$. It thus follows that

$$\frac{1}{m}QC^{\searrow 0}(q) \leq 2\frac{\log m}{m} + s + \frac{1}{2}\delta + \frac{c''}{m},$$

whence, if m is large enough, the upper bound.

The proof of the quantum version of Brudno's theorem is as follows: let $\tilde{q}_m(\delta)$ be the Ψ-typical projector sequence given in Proposition 5.4, i.e. the complexity $\frac{1}{m}QC^{\searrow 0}$ of every one-dimensional projector $q \leq \tilde{q}_m(\delta)$ is upper bounded by $s+\delta$. Due to Corollary 5.3, there exists another sequence of Ψ-typical projectors $p_m(\delta) \leq \tilde{q}_m(\delta)$ such that, additionally, $\frac{1}{m}QC^{\searrow 0}(q) > s - \delta$ holds.

6. Conclusions

In this paper we have reviewed an extension of Brudno's theorem to the quantum setting, though in a slightly weaker form which is due to the absence of a natural concatenation of qubits. The quantum Brudno's relation proved in this paper is not a pointwise relation as in the classical case, rather a kind of convergence in probability which connects the *quantum complexity per qubit* with the von Neumann entropy rate of quantum ergodic sources. Possible strengthening of this relation following the strategy which permits the formulation of a quantum Breiman theorem starting from the quantum Shannon–McMillan noiseless coding theorem (Bjelaković et al., 2003) will be the matter of future investigations.

The ultimate goal behind any quantum extension of classical algorithmic complexity is a better comprehension of the concept of *randomness* in quantum mechanics, as much as algorithmic complexity sheds light on classical randomness. Also, Brudno's theorem links it to the chaoticity of the left-shift along (binary) sequences, thus, via phase–space coarse graining, to the chaoticity of classical dynamics embodied by the dynamical entropy of Kolmogorov and Sinai (KS entropy) (Alekseev and Yakobson, 1981; Frigg, 2004).

Passing to the quantum realm, the qualities of any proposal of *quantum algorithmic complexity* will therefore have to be tested against questions like *how random is a given qubit string? how random is a given quantum dynamics?* However, with respect to the classical setting, where there exists essentially one algorithmic complexity and one entropy rate (dynamical entropy), in the quantum case there are manifold possible extensions of both with the property that they coincide with their classical counterparts when used in commutative (classical) contexts, departing from each other otherwise.

For instance, the quantum algorithmic complexity presented in this paper (Berthiaume et al., 2001) is based on *quantum descriptions* (qubit strings as inputs of QTMs) of qubit strings, the notion put forward by Vitányi (2001) is instead based on *classical descriptions* (classical bit strings as inputs of QTMs) of qubit strings (see Svozil (1990) in relation to this approach, where quantum states are proposed as providers of classically random strings), while the algorithmic complexity in Mora and Briegel (2005) is essentially the classical description complexity of *quantum circuits* that serve to the concrete preparation of the qubit strings. Further, in Gács (2001) no reference is made to circuits or to UQTMs, rather the proposed quantum algorithmic complexity is constructed by extending the notion of *universal semimeasure* to that of *universal semidensity matrix*.

Classically, the approach based on constructive semimeasures is equivalent to the descriptional one; quantum mechanically, though relations can be established between the approaches of (Berthiaume et al., 2001; Vitányi, 2001; Gács, 2001), one does not expect classical-like equivalences among the various definitions. It is indeed a likely consequence of the very structure of quantum theory that a same classical notion may be extended in different inequivalent ways, all of them reflecting a specific aspect of that structure.

This fact is already apparent in the case of quantum dynamical entropies (compare for instance Alicki and Narnhofer (1995)) that extend the KS entropy where the CNT dynamical entropy of Connes, Narnhofer and Thirring coincides with the quantum entropy rate and the AF entropy developed by Alicki and Fannes does not, the former possibly vanishing and the other not for the same dynamical system. Indeed, each of the two extensions captures dynamical features precluded to the other. Analogously, it is therefore possible that there may exist different, equally suitable notions of “quantum randomness”, each one of them reflecting a different facet of a same multiform property.

The CNT entropy is stricly related to the quantum complexity used in this paper and thus describes that aspect of randomness which is related to compressibility, termed *chaoticness in* (Uspenskii et al., 1990). The AF entropy brings into play the unavoidable randomness due to quantum measurement processes that provide classical descriptions; it might then be related to Vitányi’s complexity and to noisy-channel quantum compressibility, whence to capacity. On the other hand, the algorithmic complexity of (Gács, 2001) points to a relation to a possible definition of *quantum randomness texts*, whence to an extension of the classical notion of *typicality* as related to Martin-Löf texts (Uspenskii et al., 1990). Finally, the circuit complexity, together with the previous one, seems best suited to attack the intriguing problem of quantum entanglement, which may perhaps represent a most radical departure from classical randomness.

References

Adleman LM, Demarrais J and Huang MA (1997) Quantum computability. SIAM Journal on Computing 26: 1524–1540

Alicki R and Fannes M (2001) Quantum Dynamical Systems. Oxford University Press

Alicki R and Narnhofer H (1995) Comparison of dynamical entropies for the noncommutative shifts. Letters in Mathematical Physics 33: 241–247

Alekseev VM and Yakobson MV (1981) Symbolic dynamics and hyperbolic dynamic systems. Physics Reports 75: 287

Benatti F, Krueger T, Mueller M, Siegmund-Schultze R and Szkoa A (2006) Entropy and Algorithmic Complexity in Quantum Information Theory: A Quantum Brudno’s Theorem. Commun. Math. Phys. 265: 437–461

Bernstein E and Vazirani U (1997) Quantum complexity theory. SIAM Journal on Computing 26: 1411–1473

Berthiaume A, Dam WVan and Laplante S (2001) Quantum Kolmogorov complexity. Journal of Computer and System Sciences 63: 201–221

Billingsley P (1965) Ergodic Theory and Information. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, New York

Bjelaković I, Krüger T, Siegmund-Schultze Ra and Szko a A (2003) Chained Typical Subspaces-a Quantum Version of Breiman's Theorem, quant-ph/0301177

Bjelaković I, Krüger T, Siegmund-Schultze Ra and Szkoa A (2004) The Shannon–McMillan theorem for ergodic quantum lattice systems. Inventiones Mathematicae 155: 203–222

Brudno AA (1983) Entropy and the complexity of the trajectories of a dynamical system Transactions of the Moscow Mathematical Society 2: 127–151

Calude C (2002) Information and Randomness. An Algorithmic Perspective, 2nd edn. Springer-Verlag, Berlin

Chaitin GJ (1966) On the length of programs for computing binary sequences Journal of the Association for Computing Machinery 13: 547–569

Cover TM and Thomas JA (1991) Elements of Information Theory, Wiley Series in Telecommunications. John Wiley & Sons, New York

Frigg R (2004) In what sense is the Kolmogorov–Sinai entropy a measure for chaotic behaviour? Bridging the gap between dynamical systems theory and communication theory British Journal for the Philosophy of Science 55: 411–434

Gács P (2001) Quantum algorithmic entropy Journal of Physics A: Mathematical and General 34: 6859–6880

Kaltchenko A and Yang EH (2003) Universal compression of ergodic quantum sources. Quantum Information and Computation 3(4): 359–375

Kolmogorov AN (1965) Three approaches to the quantitative definition on information Problems of Information Transmission 1: 4–7

Kolmogorov AN (1968) Logical basis for information theory and probability theory IEEE Transactions on Information Theory 14: 662–664

Li M and Vitanyi P (1997) An introduction to Kolmogorov complexity and its applications. Springer Verlag

Mora C and Briegel HJ (2005) Algorithmic complexity and entanglement of quantum states. Phys. Rev. Lett. 95:200503

Nielsen MA and Chuang IL (2000) .Quantum computation and quantum information. Cambridge University Press, Cambridge

Solomonoff RJ (1964) A formal theory of inductive inference. Information and Control 7, 1–22, 224–254

Svozil K (1990) The quantum coin toss-testing microphysical undecidability Physics Letters A 143: 433–437

Uspenskii VA, Semenov AL and Shen AKh (1990) Can an individual sequence of zeros and ones be random. Russian Math. Survey 45(1): 121–189

Vitányi P (2001) Quantum Kolmogorov complexity based on classical descriptions IEEE Transactions on Information Theory 47(6): 2464–2479