Open Systems & Information Dynamics
Vol. 16, No. 1 (2009) 1–28
© World Scientific Publishing Company

World Scientific
www.worldscientific.com

Quantum Algorithmic Complexities and Entropy

Fabio Benatti

Dipartimento di Fisica Teorica, Università di Trieste, 34014 Trieste, Italy

and

Istituto Nazionale di Fisica Nucleare, Sezione di Trieste, 34014 Trieste, Italy

(Received: January 5, 2009)

Abstract. We review the basics of classical algorithmic complexity theory and two of its quantum extensions that have been prompted by the foreseeable existence of quantum computing devices. In particular, we will examine the relations between these extensions and the von Neumann entropy rate of generic quantum information sources of ergodic type.

1. Introduction

Among the many consequences of the development of information theory and computer science was the realization that information is physical [51, 9, 36] and thus that the limits to information processing tasks are ultimately set by the physical laws on which these processes are based. For instance, the present models of computation rely on physical phenomena that are modeled by deterministic and/or stochastic classical processes, while *quantum computation theory* [22, 40, 57, 45, 24] envisages the possibility of a new model of computation based on quantum mechanical laws. Such a new revolutionary perspective finds its motivations on one hand in the high pace at which chip miniaturization proceeds [22], and, on the other hand, in the suggestion [36, 41] that computing devices based on quantum physics might provide a more efficient description of quantum systems than classical (probabilistic) computers and, above all, from the discovery of quantum algorithms with more efficient performances with respect to what is classically achievable [57].

The most basic information processing device is abstractly modeled by a *Turing Machine* (TM) (see [67]); it consists of a bi-infinite tape $\boldsymbol{T}$ subdivided into cells containing either a blank symbol # or a symbol σ from a given alphabet $\widetilde{\Sigma}(=\{0,1\}$ for the sake of simplicity); a read/write head $\boldsymbol{H}$ moving along the tape which, when positioned on the i-th cell, reads the symbol $\sigma_i \in \Sigma$, leaves it unchanged or changes it into $\sigma'_i \in \Sigma$ and then proceeds to either the cell $i+1$ to the right (R) or to the cell $i-1$ to the left (L); a central processing unit $\boldsymbol{C}$ (CPU) capable of a finite number of control states q_i, which is updated at each computational step.

The list of possible moves defines a program for the TM; formally, it amounts to a *transition function* $\delta : Q \times \Sigma \mapsto Q \times \Sigma \times \{L, R\}$ which connects a pair (q, σ) consisting of an internal state q and symbol read σ to a new internal state q', a new symbol σ' and an order to the head to move to the next cell to the left, L, or to the right, R: $\delta(q, \sigma) = (q', \sigma', d)$, $d \in \{L, R\}$. A TM is probabilistic if (q, σ) is connected with more triplets (q', σ', d), each transition occurring with a certain probability $p_{q,\sigma}(q', \sigma', d)$; if $p_{q,\sigma}(q', \sigma', d) = 1$ for a unique (q', σ', d) and $= 0$ for all other triplets, then one is using a deterministic TM.

As a consequence, any TM can be identified by the set of rules defining the transition function δ and the transition probabilities (approximated by rational numbers). Each set of rules, that is any TM, corresponds to a certain task, a *computation*, to be performed on an input binary string $\boldsymbol{i}^{(n)}$ of length n belonging to the set Ω_2^* of strings of any finite length. Any computation can be assumed to start with the CPU control state in a chosen ready state q_r, the head positioned on a chosen 0-th cell and the input written on a finite number of cells extending from the 0-th one to its left, while all other cells to the left and to the right contain blank symbols. The computation then proceeds through a sequence of steps dictated by the transition function δ, each of them corresponding to a certain configuration of the TM that performs it. In order to determine when a computation terminates, a special state q_f among the control states is selected such that when the control unit is in this state, then the output is read off as the binary string which starts from the position of the head and continues to the right until the last symbol $\sigma_i \neq \#$.

Since programs consist of finite sets of rules, they — and thus the corresponding TMs — can be listed and identified by natural numbers. One can then construct a TM $\mathfrak{U}$ which, by being given the number n_A (written as a binary string) of a TM $\mathfrak{A}$ and an input $\boldsymbol{i}^{(n)} \in \Omega^*$, reads n_A and works as $\mathfrak{A}$ on $\boldsymbol{i}^{(n)}$. Such a TM $\mathfrak{U}$, which is able to simulate any other TM $\mathfrak{A}$, is called a *universal Turing machine* (UTM) and was used by Kolmogorov to constructively attack the issue of characterizing randomness, one of those intuitive notions which is most elusive from the mathematical point of view.

Consider a string $\boldsymbol{i}^{(n)} \in \Omega_2^*$ whose symbols are emitted independently with probabilities $p_{0,1} = 1/2$ (by a Bernoulli source, like tossing a fair coin); both the string $\boldsymbol{i}^{(n)}$ consisting of n 0s and a string $\boldsymbol{j}^{(n)}$ made of 0s and 1s distributed without any evident pattern occur with probability 2^{-n}. However, because of its regularity, $\boldsymbol{i}^{(n)}$ would be called nonrandom and, vice versa, because of the absence of regular structures, $\boldsymbol{j}^{(n)}$ would be called random [33, 67]. Presence and absence of patterns provide a useful clue to defining which strings or sequences are random and which are not; in fact, patterns in a string allow one to compress its description by means of a program run by a UTM $\mathfrak{U}$.

Any binary string $\boldsymbol{i}^{(n)} = i_1 i_2 \cdots i_n \in \Omega_2^*$ can always be reproduced by processing the program

$$\textbf{PRINT} \quad i_1 i_2 \cdots i_n \,,$$

which specifies the bits to print, one after the other. This program amounts to the literal transcription of the target string. Clearly, it is convenient to seek more clever ways to describe $\boldsymbol{i}^{(n)}$, that is shorter programs. In doing so, one is much helped by the presence of patterns; if $i_j = 0$ for all j, the following simple program could be used:

$$\textbf{PRINT} \quad \mathbf{0} \quad n \quad \textbf{TIMES} \,.$$

For large n, the length of such program goes as $\log_2 n$, that is as the number of bits necessary to specify the length of the string $\ell(\boldsymbol{i}^{(n)}) = n$. On the other hand, if $\boldsymbol{i}^{(n)}$ shows no pattern, there is no shorter effective description than literal transcription. In this case, the length of the effective description grows as n and not as $\log_2 n$.

It was a natural proposal then [49, 50, 27, 63, 64] to quantify the randomness content of a binary string by the length of the shortest program that, processed by a UTM, outputs the target. Any program p such that $\mathfrak{U}[p] = \boldsymbol{i}^{(n)}$ is an *effective description* of the string $\boldsymbol{i}^{(n)}$ and the length $\ell(p^*)$ of the shortest program p^* such that $\mathfrak{U}[p^*] = \boldsymbol{i}^{(n)}$ is called the *algorithmic complexity* of $\boldsymbol{i}^{(n)}$ [71, 65, 38, 39].

This idea and its further developments [29 − 31] have led to many interesting applications in mathematics and physics ranging from number theory to thermodynamics [67, 25, 59, 62]. Among these, an intriguing result was the connection between the complexity per symbol of symbolic trajectories of ergodic dynamical systems and their dynamical entropy production; this result is known as *Brudno's theorem* [23, 70, 3, 46]. The simplest instances of classical dynamical systems are information sources that emit symbols, one after the other: their dynamics is the shift along the emitted sequences and their structure entirely relies upon their state, namely upon the statistics with which these sequences occur. For them, Brudno's theorem asserts that almost all sequences of ergodic classical information sources are characterized by algorithmic complexity per symbol which equals the entropy per symbol of the source. Notice that entropy (and thus the entropy rate) quantifies the randomness of a statistical ensemble with a given probability distribution (the emitted strings of any length and their probabilities); instead, algorithmic complexity characterizes the randomness of individual objects (individual strings).

The idea of computing devices that process information on the atomic scale and thus work according to quantum mechanics prompted the extension of the notions of TMs and of UTMs to those of *quantum Turing machines*

(QTMs) [34] and to *universal* QTMs (UQTMs) [14]: roughly speaking, the latter ones are computing devices that work as classical TMs and UTMs, the only difference being that their configurations behave as vector states of a suitable Hilbert space. Namely, given any set of possible configurations, their linear superpositions are also possible configurations. In comparison with probabilistic TMs, possible configurations of a QTM are connected not by transition probabilities, rather by transition amplitudes [40] so that constructive and destructive interferences are possible.

Once the existence of UQTMs is foreseen, a very natural theoretical step is to try to formulate a quantum algorithmic complexity theory whose ultimate goal should be a theory of randomness of individual quantum states or qubit strings [38]. Qubit strings can be identified with pure or mixed states corresponding to local configurations of quantum spin chains; these are natural models of quantum information sources and are characterized by their von Neumann entropy rates [61, 16]. In the following, we shall briefly review the basics of classical algorithmic complexity and two proposals of quantum algorithmic complexity. Both of them are based on the use of UQTMs for the description of qubit strings; the first one [15] (referred to as quantum qubit complexity) considers descriptions of qubit strings by quantum programs, that is by other qubit strings. The second one [66], referred to as quantum bit complexity) restricts the descriptions of qubit strings to classical binary strings again being processed by UQTMs. We shall review a Brudno-like relation between quantum qubit complexity rate and von Neumann entropy rate [5] and study in more detail the possible existence of a similar relation for the quantum bit complexity.

2. Classical Algorithmic Complexity

We shall restrict attention to binary strings $\boldsymbol{i}^{(n)} = i_1 i_2 \cdots i_n \in \Omega_2^*$, $i_j = 0, 1$, of any length. With regard to the randomness of single binary strings, based on the motivations sketched above, one is led to the following

DEFINITION 1 The Kolmogorov complexity [33, 67] or plain algorithmic complexity of $\boldsymbol{i}^{(n)} \in \Omega_2^{(n)}$ is the length of the shortest binary program p such that $\mathfrak{U}[p] = \boldsymbol{i}^{(n)}$:

$$C_{\mathfrak{U}}[\boldsymbol{i}^{(n)}] \;:=\; \min\Big\{\ell(p)\;:\;\mathfrak{U}(p) = \boldsymbol{i}^{(n)}\Big\}\,.$$

One may deem the above definition not to be entirely objective as the dependence on the chosen UTM $\mathfrak{U}$ remains; however, plain algorithmic complexities of the same string $\boldsymbol{i}^{(n)}$ with respect to two different UTMs $\mathfrak{U}_{1,2}$ differ by a constant which does not depend on the string. Indeed, $\mathfrak{U}_1$ can simulate $\mathfrak{U}_2$ and vice versa; given $\boldsymbol{i}^{(n)}$, let p_1^* be such that $C_{\mathfrak{U}_1}(\boldsymbol{i}^{(n)}) = \ell(p_1^*)$ and let P_{12}

be the program, of length $\ell(P_{12}) = L_{12}$, which allows $\mathfrak{U}_2$ to simulate $\mathfrak{U}_1$. In order to make $\mathfrak{U}_2$ simulate $\mathfrak{U}_1$ on the input p_1^*, the programs P_{12} and p_1^* must be put together so that $\mathfrak{U}_1$ knows where the simulation instructions end and the string to be processed begins. This is achieved by concatenating P_{12} and p_1^* as $q = p_1^* \beta(P_{12})$, where

$$\boldsymbol{i}^{(n)} = i_1 i_2 \cdots i_n \ \mapsto \ \beta(\boldsymbol{i}^{(n)}) = i_1 i_1 i_2 i_2 \cdots i_n i_n 01$$

is the encoding of a string obtained by repeating each of its bits twice and marking the end with the pair of different bits 01: for this encoding one needs $\ell(\beta(P_{12})) = 2\,(L_{12} + 1)$ bits. In this way $\mathfrak{U}_2$ will first read $\beta(P_{12})$ being thus able to simulate $\mathfrak{U}_1$ on the subsequent portion p_1^* of the program q. Therefore, from Definition 1, it follows that

$$C_{\mathfrak{U}_2}(\boldsymbol{i}^{(n)}) \ \leq \ \ell(q) + A \ \leq \ \ell(p^*) + 2(L_{12} + 1) + A \ \leq \ C_{\mathfrak{U}_1}(\boldsymbol{i}^{(n)}) + A_{12}$$

and $C_{\mathfrak{U}_1}(\boldsymbol{i}^{(n)}) \leq C_{\mathfrak{U}_2}(\boldsymbol{i}^{(n)}) + A_{21}$, by changing $\mathfrak{U}_1$ into $\mathfrak{U}_2$; therefore,

$$\left| C_{\mathfrak{U}_1}(\boldsymbol{i}^{(n)}) - C_{\mathfrak{U}_2}(\boldsymbol{i}^{(n)}) \right| \ \leq \ A\,,$$

where A is a suitable constant which does not depend on the input $\boldsymbol{i}^{(n)}$.

Remark 1 Because of the above argument and of the fact that one is interested in very long strings, the additive constant can be in most cases safely neglected and the explicit dependence of $C_{\mathfrak{U}}$ on the UTM $\mathfrak{U}$ dropped. Also, one has to notice that algorithmic complexity deals with shortness of descriptions and not with the amount of time and memory resources that are needed to perform certain computational tasks.

The latter problem is rather the concern of *computational complexity theory*, while in algorithmic complexity theory one has in line of principle no limitations of time and memory (this is somewhat inherent in the assumed infinity of the tape of TMs).

LEMMA 1 *The plain algorithmic complexity is upper bounded as follows*

$$C(\boldsymbol{i}^{(n)}) \ \leq \ A + \ell(\boldsymbol{i}^{(n)}) \ = \ A + n\,, \tag{1}$$

where A is a constant which does not depend on $\boldsymbol{i}^{(n)}$.

Proof. As seen in the introduction, one can always choose as the effective description the program which tells $\mathfrak{U}$ to print the bits of $\boldsymbol{i}^{(n)}$ one after the other; the constant A takes care of a fixed amount of bits that specify the printing instructions. □

The above upper bound is in general rough; indeed, there are strings with patterns that may be used to compress their descriptions. However, when n increases, (1) turns out not too bad.

LEMMA 2 *The number of strings $\boldsymbol{i}^{(n)} \in \Omega^{(n)}$ with plain algorithmic complexity strictly smaller than $c > 0$ is bounded by*[a]

$$\#\Big\{\boldsymbol{i}^{(n)} \in \Omega^{(n)} \,:\, C_{\mathfrak{U}}(\boldsymbol{i}^{(n)}) < c\Big\} \;\leq\; 2^c - 1\,. \tag{2}$$

Proof. The number of binary programs with length smaller than c equals the number of binary strings with $\lfloor c \rfloor - 1$ digits at the most, whence

$$\#\{p \,:\, \ell(p) < c\} \;=\; \sum_{j=1}^{\lfloor c \rfloor - 1} 2^j \;=\; 2^{\lfloor c \rfloor} - 1 \;\leq\; 2^c - 1\,.$$

□

The above result can be recast in the following way: for $0 < \alpha < n$, the fraction of binary strings of length n with algorithmic complexity larger than $n - \alpha$ is larger than

$$\frac{2^n - 2^{n-\alpha}}{2^n} \;=\; 1 - \frac{1}{2^\alpha}$$

which, by choosing $1 \ll \alpha \ll n$, can be made close to 1.

Given any UTM $\mathfrak{U}$, it is suggestive to associate to binary strings $\boldsymbol{i}^{(n)}$ *algorithmic probabilities* defined by

$$P_{\mathfrak{U}}(\boldsymbol{i}^{(n)}) \;:=\; \sum_{p\,:\,\mathfrak{U}[p] = \boldsymbol{i}^{(n)}} 2^{-\ell(p)}\,. \tag{3}$$

That is, the programs that processed by $\mathfrak{U}$ output $\boldsymbol{i}^{(n)}$ are given weights equal to the probabilities that their binary strings occur by tossing a fair coin and these weights are summed. Clearly, the quantity $P_{\mathfrak{U}}(\boldsymbol{i}^{(n)})$ cannot converge because if p is such that $\mathfrak{U}[p] = \boldsymbol{i}^{(n)}$, all programs of the form pq obtained by concatenating p with any other binary string $q \in \Omega_2^*$ would also output $\boldsymbol{i}^{(n)}$. The program p is called a *prefix* for the program pq and a program p is said to have the *prefix property* if it is not the prefix of any other halting program; the divergence of the above sum is avoided if one restricts to UTMs that accept only programs that have the prefix property: these UTMs are called *prefix* UTMs.

A prefix TM can be realized [28] as a TM with a control unit, two tapes and two read-write heads. The first tape, the program tape, is entirely occupied by the program which is written as a binary string between two blank symbols marking its beginning and its end; the program is read by a head that can only read, halt and move right. The second tape, the work tape, is, as in the case of an ordinary TM, two-way infinite and the head on it can read, write 0,1, leave a blank #, halt or move both right and left.

[a] If c is not an integer, c is to be understood as $\lfloor c \rfloor$, the largest integer not exceeding c: $\lfloor c \rfloor \leq c < \lfloor c \rfloor + 1$.

The computation starts with the head on the program tape scanning the first blank symbol, the other head on the 0-th cell of the work tape, only finitely many of its cells possibly carrying nonblank symbols, and with the control unit in its initial ready state q_r. Then, in agreement with the symbols read by the two heads and the control unit internal state, the head on the working tape erases and writes or does nothing and then moves left, right or stays, the head on the program tape either moves right or stays, while the control unit updates its internal state.

The computation terminates if the reading head on the program tape reaches the end of the program, in which case, the output is what is written on the work tape to the right of the cell being scanned by the head until only cells with blank symbols are found. The program halts if and only if the head on the program tape reaches the end of the tape.

By restricting to prefix UTMs, one introduces a refined version of algorithmic complexity [28].

DEFINITION 2 The prefix algorithmic complexity of $\boldsymbol{i}^{(n)} \in \Omega_2^{(n)}$ is the length of the shortest program p such that $\mathfrak{U}[p] = \boldsymbol{i}^{(n)}$, where $\mathfrak{U}$ is any chosen reference prefix UTM:

$$\mathrm{K}(\boldsymbol{i}^{(n)}) \;=\; \min\Big\{\ell(p) \,:\, \mathfrak{U}[p] = \boldsymbol{i}^{(n)}\,,\ \mathfrak{U} \text{ a prefix UTM}\Big\}\,.$$

Since the set of programs with the prefix property is smaller than the set of all programs, then $\mathrm{C}(\boldsymbol{i}^{(n)}) \leq \mathrm{K}(\boldsymbol{i}^{(n)})$. On the other hand, if p is such that $\mathrm{C}(\boldsymbol{i}^{(n)}) = \ell(p)$, its self-delimiting encoding $p^* := \beta(\ell(p))p$ discussed after Definition 1 has the prefix property; therefore, it follows that

$$\mathrm{K}(\boldsymbol{i}^{(n)}) \;\leq\; \ell(p^*) \;\leq\; \mathrm{C}(\boldsymbol{i}^{(n)}) + 2\log\ell(p) + C\,.$$

Remark 2 If $\mathfrak{U}$ is a prefix UTM, not only the quantity $P_{\mathfrak{U}}(\boldsymbol{i})$ in (3) exists, but so does the *Chaitin magic number* $\Omega = \sum_{\boldsymbol{i}\in\Omega^*} P_{\mathfrak{U}}(\boldsymbol{i})$ [29, 30, 67]. Thus, $P_{\mathfrak{U}}(\boldsymbol{i})/\Omega$ represents the probability that $\boldsymbol{i}$ is the output of $\mathfrak{U}$ running a binary program p of length $\ell(p)$ randomly chosen according to the Bernoulli uniform probability distribution that assigns probability $2^{-\ell(p)}$ to anyone of them. Since short programs have higher probabilities, random strings have smaller algorithmic probabilities than regular ones. Because changing prefix UTM introduces corrections that are independent of the chosen string $\boldsymbol{i}^{(n)}$, the quantity $P_{\mathfrak{U}}(\boldsymbol{i}^{(n)})$ plays the role of a *universal semi-measure* and is the key quantity in a whole theory that goes under the name of *algorithmic probability*. Its relations to algorithmic complexity, such as the intriguing equality

$$\mathrm{K}(\boldsymbol{i}^{(n)}) \;=\; -\log_2 P_{\mathfrak{U}}(\boldsymbol{i}^{(n)}) + A\,,$$

where A is a constant independent of $\boldsymbol{i}^{(n)}$, are thoroughly reviewed in [67].

As already stressed in the introduction, algorithmic complexity theory was born to put the notion of randomness of individual objects, bit strings for concreteness, on a solid mathematical ground. However, instead of single bit strings $\boldsymbol{i}^{(n)} \in \Omega_2^*$, one may consider the statistical ensemble of all binary strings of length n, $\Omega_2^{(n)}$, and the probabilities $p(\boldsymbol{i}^{(n)})$ with which they occur, for instance being emitted by a classical information source. In such a case, given the probability distribution $\pi^{(n)} = \{p(\boldsymbol{i}^{(n)})\}_{\boldsymbol{i}^{(n)} \in \Omega_2^{(n)}}$, the Shannon entropy

$$H(\pi^{(n)}) \ := \ -\sum_{\boldsymbol{i}^{(n)} \in \Omega^{(n)}} p(\boldsymbol{i}^{(n)}) \log_2 p(\boldsymbol{i}^{(n)}) \tag{4}$$

is a good indicator of the randomness of the information source or of the average randomness of the strings of length n it emits.

Of course, classical information sources emit strings of any length and, usually, they can be assumed to be stationary, that is the probability of a string $\boldsymbol{i}^{(n)} = i_1 i_2 \cdots i_n$ being emitted does not depend on which use of the source i_1 has occurred. This condition is mathematically rephrased as follows:

$$\sum_{i_1=0}^{1} p(i_1 i_2 \cdots i_n) \ = \ p(i_2 i_3 \cdots i_n)\,. \tag{5}$$

Notice that, for nonstationary sources the probability of a string depends on when it starts being emitted; thus, since i_2 is the second symbol in the probabilities appearing in the sum, the probability on the right hand side which has i_2 as first symbol need not in general result from the summation. A second compatibility relation must also generically hold: it states that the sum of the probabilities of occurrence of strings of length $n+1$ with the same prefix of length n must be the probability of the prefix,

$$\sum_{i_{n+1}=0}^{1} p(i_1 i_2 \cdots i_n i_{n+1}) \ = \ p(i_1 i_2 \cdots i_n)\,. \tag{6}$$

A fundamental result of probability theory [18] states that a family of probability distributions $\pi^{(n)}$ satisfying the above two conditions gives rise to a translation-invariant probability measure π on the set Ω_2 of binary sequences equipped with the σ-algebra generated by *cylinder sets*, that is by intersections and unions of all possible sets of sequences with fixed prefixes $\boldsymbol{i}^{(n)} = i_1 i_2 \cdots i_n \in \Omega_2^*$.

By going from strings to sequences, Shannon entropy generally diverges and one then considers the entropy rate of the source

$$h(\pi) \ := \ \lim_{n\to+\infty} \frac{1}{n} H(\pi^{(n)})\,. \tag{7}$$

Because of the stationarity of the probability measure π and the subadditivity of the Shannon entropy [18, 32, 60, 68], the above limit exists and, in general, $h(\pi) \leq H(\pi^{(1)}) \leq 1$.

The number $h(\pi)$ is rather interesting: if the source is ergodic, that is if π cannot be written as a convex decomposition of other stationary probability measures on Ω_2, then, by the Shannon-Mc Millan-Breiman theorem [18, 60], $h(\pi)$ represents the maximal compression rate of the source. Namely, for sufficiently large n, the statistics of the source allows one to encode its strings of length n, which are 2^n, into a subset containing roughly $2^{nh(\pi)} \leq 2^n$ strings, which still carries almost exactly the information delivered by the source [33]. Indeed, the sender can send the receiver those strings which occur with probability almost $2^{-nh(\pi)}$, which are less than those possibly emitted by the source, and encode all the other ones by a same fixed string; the probability that the receiver makes an error in wrongly identifying the latter strings becomes negligibly small with increasing n. Moreover, the compression rate given by $h(\pi)$ is maximal in the sense that, if one tries to compress more, that is by encoding the strings emitted by the source into $2^{n\alpha}$ strings with $\alpha < h(\pi)$, then the probability of an error at the receiver's end of the transmission channel tends to 1.

The other reason why the entropy rate of an ergodic source is an interesting parameter comes from its relation to the algorithmic complexity rate of its sequences. As well as for the Shannon entropy, in the limit of longer and longer strings, one considers the algorithmic complexity per symbol [46]; namely, given a sequence $\boldsymbol{i} \in \Omega_2$, one looks for its prefixes $\boldsymbol{i}^{(n)} = i_1 i_2 \cdots i_n$ of increasing length and defines $\mathrm{c}(\boldsymbol{i}) := \limsup_{n\to\infty} \frac{1}{n}\mathrm{C}(\boldsymbol{i}^{(n)})$. The following result holds [23, 70].

THEOREM 1 *Let (Ω_2, π) denote an ergodic binary source with entropy rate $h(\pi)$. Then,*

$$\mathrm{c}(\boldsymbol{i}) \;=\; \lim_{n\to\infty} \frac{1}{n}\mathrm{C}(\boldsymbol{i}^{(n)}) \;=\; h(\pi)\,, \tag{8}$$

for π-almost all $\boldsymbol{i} \in \Omega_2$.

Remark 3 The scope of Brudno's theorem is wider than in the above theorem; indeed, it applies to generic stationary discrete-time dynamical systems. By partitioning the phase-space of such systems into finitely many nonintersecting atoms (a so-called coarse-graining), one can associate to their trajectories sequences consisting of the labels of the atoms visited by the phase-point at each successive tick of time. The time-invariant states of such systems provide a family of probability measures for the coarse-grained trajectories of any length n that fulfil conditions (5) and (6). In this way, each coarse-graining yields a symbolic model, whereby the dynamics amounts to the shift along sequences of symbols distributed according to a definite translation-invariant probability measure [3]; these symbolic models have each an entropy rate and the maximal one is known as *dynamical entropy* or *Kolmogorv-Sinai entropy* of the dynamical system [32]. In the same way, one associates to the coarse-grained trajectories their complexity per symbol; then, Brudno's theorem in its general form asserts that the maximal complexity rate of almost all trajectories of ergodic classical dynamical systems equals their dynamical entropy.

3. Quantum Algorithmic Complexities

The quantum algorithmic complexity which we are going to examine in the following is based on the hypothesis that in a more or less distant future quantum computers will effectively supplant the present model of computation [40, 57, 45]) and will be used to effectively describe quantum states.

As in the classical realm, one is mainly interested in qubit strings of increasing length; therefore, the most convenient mathematical framework for accommodating states of discrete quantum systems of increasing dimension is provided by quantum spin chains.

Spin chains of commutative type, that is infinite lattices carrying at each site spins which are capable of only two states, up ($i = 0$) and down ($i = 1$) along the z axis, say, already naturally appear when dealing with the set of binary strings Ω_2^*. In fact, $\Omega_2^* \ni \boldsymbol{i}^{(n)} = i_1 i_2 \cdots i_n$ can be thought of as a local configuration (from site 1 to site n) of a doubly infinite commutative spin chain where the spins $1/2$ at all other sites outside the given interval are either all pointing up or down. Local probability distributions $\pi^{(n)} = \{p(\boldsymbol{i}^{(n)})\}_{\boldsymbol{i}^{(n)} \in \Omega_2^{(n)}}$ on the statistical ensembles of strings of length n correspond to diagonal density matrices

$$\rho^{(n)} := \sum_{\boldsymbol{i}^{(n)} \in \Omega_2^{(n)}} p(\boldsymbol{i}^{(n)}) \, | \, \boldsymbol{i}^{(n)} \, \rangle \langle \, \boldsymbol{i}^{(n)} \, | \, ,$$

where the strings $\boldsymbol{i}^{(n)}$ are naturally associated with the elements of the so-called orthogonal *computational basis* in a 2^n-dimensional Hilbert space. Notice also that the Shannon entropy $H(\pi^{(n)})$ of the probability distribution $\pi^{(n)}$ (see (4)) coincides with the von Neumann entropy

$$S(\rho^{(n)}) := -\operatorname{Tr}(\rho^{(n)} \log_2 \rho^{(n)}) \, . \tag{9}$$

A quantum spin chain $\mathcal{A}$ is a one-dimensional lattice each of whose sites i carries a same full matrix algebra $(\boldsymbol{A})_i = M_d(\mathbb{C})$ describing a d-level system; if $d = 2$, $\mathcal{A}$ describes a qubit spin chain [20, 4]. More precisely, local spin configurations located within finite integer intervals $[p, q]$, $p \leq q$, are described by *local algebras* that are finite tensor products of single site matrix algebras,

$\boldsymbol{A}_{[p,q]} := \otimes_{i=p}^{q}(\boldsymbol{A})_i$. Tensor products $\boldsymbol{A}_{[p,q]}$ are then turned into infinite tensor product by tensorizing them with the infinite tensor products $1_{\{-\infty,p-1]}$ and $1_{[q+1,+\infty\}}$ of as many identity matrices $1 \in M_d(\mathbb{C})$ as the sites outside the intervals $[p,q]$; in such a way one can consider the so-called $*$-algebra $\mathcal{A}^*$ containing all products and linear combinations of elements of different $\boldsymbol{A}_{[p,q]}$. Finally, $\mathcal{A}$ is the so-called *quasi-local* C^* algebra obtained by norm-closure of $\mathcal{A}^*$ with respect to the norm that reduces to the usual matrix norm when restricted to the local algebras $\boldsymbol{A}_{[p,q]}$ [20, 21].

As much as in the commutative case, one is usually interested in stationary quantum sources; they correspond to quantum spin chains $\mathcal{A}$ equipped with a global state ω, that is with a normalized, positive functional (expectation) on $\mathcal{A}$ that is shift-invariant. Namely, $\mathcal{A}$ is endowed with a natural isomorphism σ such that

$$\sigma\left(1_{\{-\infty,-1]}\otimes(A^0)_0\otimes 1_{[1,+\infty\}}\right) \;=\; 1_{\{-\infty,0]}\otimes(A^0)_1\otimes 1_{[2,+\infty\}}$$

and the statistics of the source is fixed by expectation ω such that $\omega(1)=1$ and $\omega\circ\sigma=\omega$. Furthermore, the restrictions of ω to the local algebras $\mathcal{A}^{(n)} := \boldsymbol{A}_{[0,n-1]}$ are density matrices $\rho^{(n)} \in \mathcal{A}^{(n)}$ acting on the Hilbert spaces $\mathbb{H}^{(n)} := (\mathbb{C}^d)^{\otimes n}$:

$$\begin{aligned}\omega\left(1_{\{-\infty,-1]}\otimes(A^0)_0\otimes(A^1)_1\otimes\cdots(A^q)_q\otimes 1_{[q+1,+\infty\}}\right)\\ = \mathrm{Tr}\left(\rho^{(n)}\,A^0\otimes A^1\otimes\cdots A^q\right).\end{aligned}$$

The set $\{\rho^{(n)}\}_{n\geq 0}$ satisfies the quantum versions of conditions (5) and (6):

$$\mathrm{Tr}_0(\rho^{(n)}) \;=\; \rho^{(n-1)} \;=\; \mathrm{Tr}_n(\rho^{(n)})\,, \tag{10}$$

where Tr_k denotes the partial trace taken with respect to any orthonormal basis in the Hilbert space corresponding to the k-th lattice site.

Vice versa, given a one-parameter family of density matrices $\{\rho^{(n)}\}_{n\geq 0}$, $\rho^{(n)} \in \mathcal{A}^{(n)}$, satisfying conditions (10), it defines a shift-invariant state ω on the quasi-local C^* algebra $\mathcal{A}$ such that $\omega\lceil\mathcal{A}^{(n)} = \rho^{(n)}$.

The average randomness of local spin configurations $\rho^{(n)}$ is measured by the von Neumann entropy (9) and that of the global state ω by the *von Neumann entropy rate*

$$s(\omega) \;:=\; \lim_{n\to+\infty}\frac{1}{n}S(\rho^{(n)})\,. \tag{11}$$

The limit exists because ω is assumed to be translation-invariant and because the von Neumann entropy is subadditive [69, 58].

One can clearly see that classical information sources correspond to quantum spin chains whose sites carry diagonal matrix algebras instead of full noncommutative matrix algebras. At the same time, one notices that there is a larger variety of qubit strings than of bit strings. Let us fix $d = 2$ and choose in each single qubit Hilbert space $\mathbb{C}^2$ a *computational basis* $|0\rangle, |1\rangle$. In order to be as general as possible, superpositions of qubit states of different lengths k are allowed: they correspond to vectors in the Fock-like Hilbert space $\mathbb{H}_F := \bigoplus_{k=0}^{\infty} \mathbb{H}_k$. More generally, qubit strings will be represented by density matrices $\rho \in \mathbb{B}_1^+(\mathbb{H}_F)$ acting on $\mathbb{H}_F$.

EXAMPLE 1 Any bit string $\boldsymbol{i} \in \{0,1\}^*$ identifies a computational basis vector in $\mathbb{H}_F$: the empty string λ corresponds to the vacuum $|\Omega_F\rangle$, the 1-qubit subspace $\mathbb{H}_1$ is spanned by $|0\rangle$, $|1\rangle$, while the k-qubit subspace $\mathbb{H}_k$ is generated by the vectors corresponding to the bit strings of length k, $\boldsymbol{i}^{(k)} \in \Omega_2^{(k)}$, namely by $|\boldsymbol{i}^{(k)}\rangle = |i_1 i_2 \cdots i_k\rangle$, $i_j = 0, 1$. Generic qubit strings amount to density matrices in $\mathbb{B}_1^+(\mathbb{H}_{\leq n})$ acting on the Hilbert space $\mathbb{H}_{\leq n} := \bigoplus_{k=0}^{n} \mathbb{H}_k$, of dimension $\sum_{k=0}^{n} 2^k = 2^{n+1} - 1$.

Quantum Turing machines act on and construct superpositions of vector qubit strings and, more generally, convex combinations of projections onto vector qubit strings of different lengths. Moreover, like their classical counterparts, QTMs comprise different parts as a read/write head $\boldsymbol{H}$, a control unit $\boldsymbol{C}$ and one or more tapes $\boldsymbol{T}$, all of them capable of being in states that are either Hilbert space vectors in $\mathbb{H}_{\boldsymbol{H}}$, $\mathbb{H}_{\boldsymbol{C}}$ and $\mathbb{H}_{\boldsymbol{T}}$, or density matrices acting on them. Therefore, the QTM configurations too are generically described by density matrices acting on appropriate Hilbert spaces. Notice that mixed states are quite typical in such context for they naturally appear when one is interested in the state of the read/write head, say, and therefore traces over the Hilbert spaces corresponding to the other QTM components.

The notion of QTM as a computing device working according to quantum mechanics was first proposed by Deutsch [34]. A full and detailed analysis can be found in [14] and [1] and further developments in connection with the notion of universality in [56]. In the following, we shall assume the existence of such machines and define their action on input qubit strings $\sigma \in \mathbb{B}_1^+(\mathbb{H}_F)$ as completely positive linear maps $\mathbb{U} : \mathbb{B}_1^+(\mathbb{H}_F) \to \mathbb{B}_1^+(\mathbb{H}_F)$. These maps emerge from the fact that QTMs act unitarily on their inputs, but, when they halt, the output is obtained by tracing over the Hilbert space $\mathbb{H}_{\boldsymbol{H}} \otimes \mathbb{H}_{\boldsymbol{C}}$ of the head and internal control unit. The combination of partial tracing and unitary transformations corresponding to the action of $\mathfrak{U}$ on input density matrices amounts to the action of a completely positive map $\mathbb{U}$ [4].

We have seen in Sect. 1 that the definition of algorithmic complexity rests on a solid ground because the length of the shortest effective description of a bit string is essentially independent of the computer that computes it once this is chosen from the class of universal Turing machines. Clearly, any definition of quantum complexity based on using QTMs will also need the existence of universal QTMs in order to be essentially machine-independent.

In [56] some problems relative to the proposal in [14] are considered and finally resolved through an operative definitions of UQTM which is fully consistent from the point of view of quantum algorithmic complexity.

Remark 4 Unlike in the classical situation where there are countably many bit strings, there are uncountably many qubit strings that can be arbitrarily close to one another. In order to quantify how close two qubit strings $\rho, \sigma \in \mathbb{B}_1^+(\mathbb{H}_F)$ actually are, the *trace distance* is a commonly used tool in quantum information [57]:

$$D(\rho, \sigma) \;:=\; \mathrm{Tr}\left(\sqrt{(\rho-\sigma)^\dagger(\rho-\sigma)}\right) . \tag{12}$$

It amounts to the sum of the absolute values of the eigenvalues of the Hermitean, generically nonpositive definite, matrix $\rho - \sigma$.

The same approximation problem arises when one wants to make a QTM act according to a unitary transformation described by a unitary matrix U. In general, a UQTM is only able to apply a unitary transformation U on some segment of its tape within an accuracy of δ, if it is supplied with a complex matrix $\tilde{U}$ as input such that

$$\|U - \tilde{U}\| \;\leq\; \frac{\delta}{2(10\sqrt{d})^d} \,,$$

d being the size of the matrix. The machine cannot apply U exactly; in fact, it only knows an approximation $\tilde{U}$. It also cannot apply $\tilde{U}$ directly, for $\tilde{U}$ is only approximately unitary, and the machine can only work unitarily. Instead, it will effectively apply another unitary transformation V which is close to $\tilde{U}$ and thus close to U, such that $\|V - U\| < \delta$. Let $|\psi\rangle := U|\psi_0\rangle$ be the output that one wants to have from $\mathfrak{U}$ and let $|\phi\rangle := V|\psi_0\rangle$ be the approximation that is really computed by the machine. Then, both the norm and trace distances are small: $\||\phi\rangle - |\psi\rangle\| < \delta$, $D(|\,\phi\,\rangle\langle\,\phi\,|, |\,\psi\,\rangle\langle\,\psi\,|) < \delta$ [5].

In the following, we shall consider two recent proposals of algorithmic complexity of quantum states; both stem from the idea that complexity should be associated with difficulty of description. They can roughly be summarized as follows:

- *quantum qubit complexity* (QC_q): quantum states are described by means of other quantum states that are processed by UQTMs [15];
- *quantum bit complexity* (QC_c): quantum states are described by classical programs run by UQTMs [66].

Remark 5 There are other proposals of quantum algorithmic complexities, not based on the use of UQTMs. Indeed, as in [53, 54, 55], one may choose to relate the complexity of qubit strings to the complexity of the (classical) description of the quantum circuits that construct a given quantum state. Or, as in [37], one may extend the notion of universal semi-measure (see (3) and Remark 2) and define a *quantum universal semi-density matrix.*

3.1. Quantum qubit complexity

The quantum extension of algorithmic complexity put forward in [15] closely follows the classical steps by seeking the shortest possible quantum descriptions of a target qubit string in terms of other density matrices involving the smallest possible number of qubits.

In the commutative setting, the length of a bit string is simply the number of bits it consists of; in the quantum setting, the number of qubits involved fixes the Hilbert space dimension. Therefore, the following definition naturally extends the notion of the program length.

DEFINITION 3 The length of a qubit string $\rho \in \mathbb{B}_1^+(\mathbb{H}_F)$ is

$$\ell(\rho) \;:=\; \min\{n \in \mathbb{N}_0 \,:\, \rho \in \mathbb{B}_1^+(\mathbb{H}_{\leq n})\}\,, \tag{13}$$

where $\ell(\rho) = \infty$ if this set is empty.

Because of the uncountable number of quantum states in a neighborhood of the target qubit string (see the discussion at the beginning of Remark 4), one seeks not exact, but only approximate reproductions whose accuracy can be taken care of by the trace distance (12). The following ones are then two natural possible extension of Definition 1.

DEFINITION 4 Let $\mathfrak{U}$ be a QTM and $\rho \in \mathbb{B}_1^+(\mathbb{H}_F)$ a qubit string. For every $\delta \geq 0$, the *finite accuracy quantum complexity* $\mathrm{QC}^{\delta}_{\mathfrak{U}}(\rho)$ is defined as the minimal length $\ell(\sigma)$ of quantum programs $\sigma \in \mathbb{B}_1^+(\mathbb{H}_F)$ such that the corresponding output $\mathfrak{U}[\sigma]$ has a trace distance from ρ smaller than δ,

$$\mathrm{QC}^{\delta}_{\mathfrak{U}}(\rho) \;:=\; \min\Big\{\ell(\sigma) \,:\, D\left(\rho, \mathfrak{U}[\sigma]\right) \leq \delta\Big\}\,. \tag{14}$$

Similarly, an *approximation scheme quantum complexity* $\mathrm{QC}_{\mathfrak{U}}$ is defined as the minimal length $\ell(\sigma)$ of any density operator $\sigma \in \mathbb{B}_1^+(\mathbb{H}_F)$, such that when processed by $\mathfrak{U}$ together with any integer k, the output $\mathfrak{U}[k,\sigma]$ has trace distance from ρ smaller than $1/k$, for all k:

$$\mathrm{QC}_{\mathfrak{U}}(\rho) \;:=\; \min\Big\{\ell(\sigma) \,:\, D\left(\rho\,, \mathfrak{U}[k,\sigma]\right) \leq \frac{1}{k} \text{ for every } k \in \mathbb{N}\Big\}\,. \tag{15}$$

The results in [56] allow one to prove the independence (up to an additive constant) of the above definitions from the chosen QTM $\mathfrak{U}$ if this is universal.

Accordingly, one can thus fix an arbitrary UQTM and, like in the classical case, drop reference to it and set

$$\mathrm{QC}_q\,(\rho) \;:=\; QC_{\mathfrak{U}}(\rho)\,, \qquad \mathrm{QC}^{\delta}_q\,(\rho) \;:=\; \mathrm{QC}^{\delta}_{\mathfrak{U}}(\rho)\,. \tag{16}$$

Remark 6 Definition 4 is essentially equivalent to that in [15], the only technical difference being the use of the trace distance rather than the fidelity.

Remark 7 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. Notice that in (15) the minimal length has to be sought among σ such that any of them yields an approximation of ρ within $1/k$ for all k: this is an effective procedure.

Remark 8 Let $k \in \mathbb{N}$ and $\sigma \in \mathbb{B}_1^+(\mathbb{H}_F)$. Let $\beta(k)$ denote the string that consists of at most $\lfloor \log_2 k \rfloor$ bits of the binary expansion of k, each repeated twice and ending with 01. Let $|\beta(k)\rangle\langle\beta(k)|$ be the corresponding projector in the computational basis. The pair (k,σ) can then be encoded into a single qubit string $\mathcal{C}(k,\sigma) := |\,\beta(k)\,\rangle\langle\,\beta(k)\,| \otimes \sigma$. Note that

$$\ell(\mathcal{C}(k,\sigma)) \;=\; 2\lfloor \log k \rfloor + 2 + \ell(\sigma)\,. \tag{17}$$

Remark 9 The exact choice of the accuracy $1/k$ is not important; choosing any computable function that tends to zero for $k \to \infty$ will yield an equivalent definition (in the sense of being equal up to some constant). The same is true for the choice of the encoding $\mathcal{C}$: as long as k and σ can both be computably decoded from $\mathcal{C}(k,\sigma)$ and as long as there is no way to extract additional information on the desired output ρ from the k-description part of $\mathcal{C}(k,\sigma)$, the results will be equivalent up to a suitable constant.

EXAMPLE 2 Every UQTM $\mathfrak{U}$ can implement the identity transformation; namely, one can always perform a literal transcription, so that automatically $\mathrm{QC}^{\delta}_{\mathfrak{U}}(\rho) \leq \ell(\rho) + c_U$ for some constant c_U. Of course, the key point in classical as well as in quantum algorithmic complexity is that there sometimes exist much shorter qubit programs than just literal transcription.

EXAMPLE 3 The finite accuracy and approximation scheme QC_{q} are related to each other by the following inequality: for every QTM $\mathfrak{U}$ and every $k \in \mathcal{N}$, $\mathrm{QC}^{1/k}_q\,(\rho) \leq \mathrm{QC}_q\,(\rho) + 2\lfloor \log k \rfloor + 2$, for all $\rho \in \mathbb{B}_1^+(\mathbb{H}_F)$. Indeed, if $\mathrm{QC}_{\mathfrak{U}}(\rho) = \ell$, there is $\sigma \in \mathbb{B}_1^+(\mathbb{H}_F)$ with $\ell(\sigma) = \ell$, such that $D(\mathbb{U}[k,\sigma],\rho) \leq 1/k$ for every $k \in \mathbb{N}$. Then $\sigma' := \mathcal{C}(k,\sigma)$, where $\mathcal{C}$ is the encoding in Remark 8, is such that $D\,(\mathbb{U}[\sigma'],\rho) \leq 1/k$ and

$$\mathrm{QC}^{1/k}_q\,(\rho) \;\leq\; \ell(\sigma') \;\leq\; 2\lfloor \log k \rfloor + 2 + \ell \;=\; 2\lfloor \log k \rfloor + 2 + \mathrm{QC}_q\,(\rho)\;,$$

where the second equality follows from (17).

We now briefly discuss a quantum version of Theorem 1 which connects the von Neumann entropy rate $s(\omega)$ (see (11)) of an ergodic quantum spin chain $(\mathcal{A},\omega)$ to the qubit complexity $\mathrm{QC}_q(\rho)$ and $\mathrm{QC}^\delta_q(\rho)$ of qubit strings that are pure states $\rho = |\psi\rangle\langle\psi|$ of the chain. Like for classical information sources, ergodic quantum spin chains are spin chains endowed with shift-invariant states which cannot be written as convex combinations of other shift invariant states. In such context, the key notion is that of *typical subspaces*.

DEFINITION 5 For a quantum spin chain $(\mathcal{A},\omega)$, the projectors $p_n \in \mathcal{A}^{(n)}$ such that $\omega(p_n) = \mathrm{Tr}(\rho^{(n)}\, p_n) \geq 1-\varepsilon$ are termed *typical projectors* and the subspaces of $(\mathbb{C}^d)^{\otimes n}$ onto which they project are called *typical subspaces*.

The following result, a *quantum equipartition theorem*, relates the spectrum of local states to the dimension of high probability subspaces [16, 42, 43].

PROPOSITION 1 *Let $(\mathcal{A},\omega)$ be an ergodic quantum source with entropy rate $s(\omega)$; further, for any $\varepsilon > 0$, let*

$$\beta_{\varepsilon,n}(\omega) \;:=\; \min\Big\{\log_2 \mathrm{Tr}(q)\;:\;\mathcal{A}^{(n)} \ni q = q^\dagger = q^2\,,\ \mathrm{Tr}(\rho^{(n)}\, q) \geq 1-\varepsilon\Big\}\,.$$

Then, for every $0<\varepsilon<1$, $\displaystyle\lim_{n\to\infty}\frac{1}{n}\beta_{\varepsilon,n}(\omega) = s(\omega)$.

Typical projections $p_n \in \mathcal{A}^{(n)}$ are such that they carry almost all the probability of ω, $\omega(p_n) = \mathrm{Tr}(\rho^{(n)}\, p_n) \geq 1-\varepsilon$, while the dimension of the corresponding subspaces, $\mathrm{Tr}(p_n)$, increases as $2^{n\, s(\omega)}$. Notice that in the above proposition $0<\varepsilon<1$ is otherwise arbitrary; therefore, a convenient way to interpret it is as follows: if a sequence of projectors $p_n \in \mathcal{A}^{(n)}$ is such that

$$\limsup_{n\to\infty}\frac{1}{n}\log_2 \mathrm{Tr}(p_n) \;<\; s(\omega)\,,$$

they cannot become eventually typical so that $\lim_{n\to+\infty}\mathrm{Tr}(\rho^{(n)}p_n) = 0$.

The quantum version of Brudno's theorem proved in [5] essentially states that there are sequences of typical subspaces of $(\mathbb{C}^2)^{\otimes n}$, such that the complexity rates $\frac{1}{n}\mathrm{QC}_q(q)$ and $\frac{1}{n}\mathrm{QC}^\delta_q(q)$ of any one-dimensional projector q onto a state belonging to these subspaces can be made arbitrarily close to the entropy rate by choosing n large enough.

THEOREM 2 *Let $(\mathcal{A},\omega)$ be an ergodic quantum source with entropy rate $s(\omega)$. For every $\delta > 0$, there exists a sequence of typical projectors $q_n(\delta) \in \mathcal{A}^{(n)}$, $n \in \mathbb{N}$, i.e. $\lim_{n\to\infty}\mathrm{Tr}(\rho^{(n)}q_n(\delta)) = 1$, such that for every one-dimensional projector $q \leq q_n(\delta)$ and n large enough*

$$\frac{1}{n}\mathrm{QC}_q(q) \in (s(\omega)-\delta, s(\omega)+\delta)\ , \tag{18}$$

$$\frac{1}{n}\mathrm{QC}^\delta_q(q) \in (s(\omega)-\delta(2+\delta)s(\omega)\,,\ s(\omega)+\delta)\ . \tag{19}$$

Moreover, $s(\omega)$ is optimal in the sense that every sequence of projectors $q_n \in \mathcal{A}^{(n)}$, $n \in \mathbb{N}$, that for large n may be represented as a sum of mutually orthogonal one-dimensional projectors that all violate the lower bounds in (18) and (19) for some $\delta > 0$, has an asymptotically vanishing expectation value with respect to ω.

The proof of the above result consists of two steps; in the first one, the lower bounds to the quantum qubit complexity rates are established by a quantum version of the counting argument in Lemma 2. An argument due to [15] states that there cannot be more than $2^{\ell+1} - 1$ mutually orthogonal one-dimensional projectors p with quantum complexity $\mathrm{QC}_q(p) \leq \ell$. The proof is based on the Holevo's χ-quantity [57] and the result can be used to provide an explicit upper bound on the maximal number of orthogonal one-dimensional projectors that can be approximated within trace distance δ by the action of completely positive maps $\mathbb{U}$ on density matrices σ of length $\ell(\sigma) \leq c$.

LEMMA 3 *Let $0 < \delta < 1/e$, $c \in \mathbb{N}$ such that $c \geq \frac{1}{\delta}\left(4 + 2\log\frac{1}{\delta}\right)$, $\mathcal{K}$ a linear subspace of an arbitrary Hilbert space $\mathbb{K}$, and $\mathbb{U} : \mathbb{B}_1^+(\mathbb{H}_F) \to \mathbb{B}_1^+(\mathbb{K})$ a completely positive trace-preserving map. Let N_c^δ be a maximum cardinality subset of orthonormal vectors from the set $A_c^\delta(\mathbb{U}, \mathcal{K})$ of all normalized vectors in $\mathcal{K}$ which are reproduced within δ by $\mathbb{U}$ acting on some input of length $\leq c$:*

$$A_c^\delta(\mathbb{U}, \mathcal{K}) \ := \ \Big\{ |\phi\rangle \in \mathcal{K} \ : \ \exists\, \sigma_\phi \in \mathbb{B}_1^+(\mathbb{H}_{\leq c}),\ D\left(\mathbb{U}[\sigma_\phi], |\,\phi\,\rangle\langle\,\phi\,|\right) \leq \delta \Big\}.$$

Then, $\log_2 |N_c^\delta| < c + 1 + \dfrac{2+\delta}{1-2\delta}\delta c.$

In the second step, for any projection onto a state vector $|\Psi\rangle$ belonging to subspaces which are *universally typical* [44] and for a fixed accuracy, one constructs a concrete qubit string σ of length roughly $n\, s(\omega)$ that run by the UQTM $\mathfrak{U}$ gives an output $\mathfrak{U}[\sigma]$ such that $D(\mathfrak{U}[\sigma], |\,\Psi\,\rangle\langle\,\Psi\,|) \leq \varepsilon$. Since σ need not be an optimal qubit program, its length gives an upper bound to the quantum qubit complexity of $|\Psi\rangle$ which increases like $n\, s(\omega)$.

For achieving the lower bound it is of fundamental importance that typical subspaces be not only typical but also universal; namely, that their typicality depends not on the precise state ω on the spin chain, but only on whether the entropy rate of the state is smaller than a certain fixed amount. Indeed, one has

PROPOSITION 2 *Let $s > 0$ and $\varepsilon > 0$. There exists a sequence of projectors $Q_{s,\varepsilon}^{(n)} \in \mathcal{A}_n$, $n \in \mathcal{N}$, such that for n large enough*

$$\mathrm{Tr}(Q_{s,\varepsilon}^{(n)}) \ \leq \ 2^{n(s+\varepsilon)} \tag{20}$$

and for every ergodic quantum state ω on $\mathcal{A}_{\mathbb{Z}}$ with entropy rate $s(\omega) \leq s$ it holds

$$\lim_{n\to\infty} \omega^{(n)}(Q^{(n)}_{s,\varepsilon}) \;=\; \mathrm{Tr}(\rho^{(n)} Q^{(n)}_{s,\varepsilon}) \;=\; 1\,. \tag{21}$$

The existence of universal typical subspaces guarantees that in order to describe their vectors one need not describe the restriction to them of the global state ω, but only a suitable bound s as close as one likes to its entropy rate $s(\omega)$. Roughly speaking, vectors $|\Psi\rangle$ in the universal typical subspace are not in general written as qubit strings so their lengths are not $\simeq n\,s$; however, they can always be unitarily transformed into qubit strings $|\Phi\rangle = U|\Psi\rangle$ of roughly that length; when $|\Phi\rangle$ is presented to a UQTM $\mathfrak{U}$ together with a suitable classical description of an approximated version V of the unitary matrix U, $\mathfrak{U}$ acting on the qubit program containing $|\,\Phi\,\rangle\langle\,\Phi\,|$ and a qubit rendering of the classical (bit) description of V will output a (close) approximation of $|\,\Psi\,\rangle\langle\,\Psi\,|$.

Remark 10 Unlike in Theorem 1 where the result holds almost everywhere, its quantum generalization given above essentially holds in probability. The major obstruction to a stronger quantum version comes form the difficulty of extending to *qubit* strings what is natural for *bit* strings, namely their concatenation [17].

EXAMPLE 4 Consider a quantum spin chain $(\mathcal{A},\omega)$ of Bernoulli type with a state ω which is the tensor product of tracial states $\rho = 1/2$ for each qubit; this quantum source is mixing, thus ergodic and its entropy rate is $s(\omega) = -\,\mathrm{Tr}\,\rho\log_2\rho = 1$. Then, the quantum version of Brudno's theorem states that there exists a sequence of subspaces $\mathbb{K}_n \subseteq \mathbb{H}_F$ of high probability, such that for any $\varepsilon > 0$, by taking n sufficiently large,

$$1-\varepsilon \;\leq\; \frac{1}{n}\mathrm{QC}_q\left(|\,\Psi\,\rangle\langle\,\Psi\,|\right) \;\leq\; 1+\varepsilon\,,$$

for all pure qubit states $\Psi \in \mathbb{K}_n$.

3.2. Quantum bit complexity

The approach to quantum algorithmic complexity proposed in [66] is based on descriptions of qubit strings $|\Psi\rangle \in \mathbb{H}_n$ by means of self-delimiting classical programs $p \in \Omega_2^*$ instead of generic qubit strings. These classical programs are presented to a fixed UQTM $\mathfrak{U}$ as computational basis vectors $|p\rangle$ which, after being processed by $\mathfrak{U}$, output normalized vectors $|\mathfrak{U}[p]\rangle \in \mathbb{H}_n$. The accuracy of the reproduction of the target $|\Psi\rangle$ by the output $|\mathfrak{U}[p]\rangle$ is taken care of by the logarithm of the modulus of the scalar product $\langle\Psi|\mathfrak{U}[p]\rangle$.

DEFINITION 6 ([66, 37]) The quantum bit complexity $\mathrm{QC}_c(\Psi)$ of n-qubit vector states $|\Psi\rangle \in \mathbb{H}_n$ is given by

$$\mathrm{QC}_c(\Psi) \;:=\; \min\Big\{\ell(p) \;+\; \lceil -\log_2 |\langle\Psi|\mathfrak{U}[p]\rangle|^2\rceil\Big\}\,,$$

where $p \in \Omega_2^*$ is any self-delimiting binary program.

The logarithm is a *penalty for bad approximations*: indeed, it diverges if $|\mathfrak{U}[p]\rangle$ is orthogonal Ψ, while it vanishes when $|\mathfrak{U}[p]\rangle = |\Psi\rangle$. Therefore, the quantum bit complexity as defined by Vitanyi is a tradeoff between the length of the classical description and the allowed errors in the reproduction of the target qubit string.

EXAMPLE 5 For generic qubit strings $\Psi \in \mathbb{H}_n$, the following universal upper bound holds [66]: if $\Psi \in \mathbb{H}_n$ is normalized

$$\mathrm{QC}_c(\Psi) \;\leq\; 2n + C\,,$$

where C is a constant independent of Ψ.

Indeed, by expanding

$$|\Psi^{(n)}\rangle \;=\; \sum_{\boldsymbol{i}^{(n)}\in\Omega_2^{(n)}} c(\boldsymbol{i}^{(n)})\,|\boldsymbol{i}^{(n)}\rangle\,,$$

with respect to the computational basis, there must be at least one $\boldsymbol{i}_*^{(n)}$ such that $|c(\boldsymbol{i}_*^{(n)})|^2 \geq 2^{-n}$. Let $p \in \Omega_2^*$ be a self-delimiting program such that $|\mathfrak{U}[p]\rangle = |\boldsymbol{i}_*^{(n)}\rangle$. Then,

$$\mathrm{QC}_c(\Psi) \;\leq\; \ell(p) + \lceil -\log_2 |\langle\mathfrak{U}[p]|\Psi\rangle|^2\rceil \;\leq\; 2n + C\,.$$

EXAMPLE 6 Suppose that for an orthonormal basis $\mathcal{B} := \{|b_i\rangle\}_{i=1}^{2^n}$ in $\mathbb{H}_n$ there exists a self-delimiting program $q_{\mathcal{B}}$, of classical prefix-complexity $\mathrm{K}(\mathcal{B}) = \ell(q_{\mathcal{B}})$, which computes it exactly and let $p_i \in \Omega_2^*$ such that $|\mathfrak{U}[p_i]\rangle = |b_i\rangle$. Let us fix $|b_i\rangle \in \mathcal{B}$; if p_i is any program such that $|\mathfrak{U}[p_i]\rangle = |b_i\rangle$ then no penalty for a bad approximation is to be payed. However, the shortest among such program, say p_*, need not achieve the minimum in the previous definition; thus,

$$\mathrm{QC}_c(b_i) \;\leq\; \ell(p_*)\,, \tag{22}$$

while, in general, $\mathrm{QC}_c(b_i) = \ell(q_*) + \lceil -\log_2 |\langle\mathfrak{U}[q_*]|b_i\rangle|^2\rceil$ for some bit string $q_* \in \Omega_2^*$. Let $\mathfrak{U}$ process binary programs in *dovetailed fashion* [67, 33]; that is the programs are listed lexicographically as binary strings and $\mathfrak{U}$ starts executing the first step of program 1, then the second step of program 1 and the first step of program 2, and so on. It follows that at the k-th step of the process $\mathfrak{U}$ will have executed $k - j + 1$ steps of the j-th program with

$1 \leq j \leq k$. Then, q_* can be used to construct the vector $|\mathfrak{U}[q_*]\rangle \in \mathbb{H}_n$ whose coefficients $\langle b_j|\mathfrak{U}[q_*]\rangle$ in the expansion with respect to the ONB $\mathcal{B}$ provide probabilities $|\langle b_j|\mathfrak{U}[q_*]\rangle|^2$ that can be used to construct a Shannon-Fano-Elias code-word $q(i)$ for $|b_i\rangle$ [33]. Therefore, $q_{\mathcal{B}}$, q_* and $q(i)$ can be used to construct a self-delimiting program $q = q_{\mathcal{B}} q_* q(i)$ such that $\mathfrak{U}$ does the following:

- it constructs the directly computable basis $\mathcal{B}$ and the vector $|\mathfrak{U}[q_*]\rangle$;
- it computes the Shannon-Fano-Elias code for $\mathcal{B}$ with respect to $|\mathfrak{U}[q_*]\rangle$;
- it outputs the vector with code-word $q(i)$.

Since $|\mathfrak{U}[q]\rangle = |b_i\rangle$, one estimates

$$\begin{aligned} \ell(p_*) \ \leq \ \ell(q) \ &\leq \ \ell(q_*) + \ell(q(i)) + \mathrm{K}(\mathcal{B}) + C \\ &= \ \mathrm{QC}_c\,(b_i) + \mathrm{K}(\mathcal{B}) + C\,, \end{aligned} \tag{23}$$

whence, from (22), one gets that, up to an additive constant,

$$\mathrm{QC}_c\,(b_i) \ = \ \min\{\ell(p) \,:\, |\mathfrak{U}[p]\rangle = |b_i\rangle\}\,.$$

We can now argue that quantum bit complexity and prefix complexity agree on classical bit strings. Indeed, let $\mathcal{B} := \{|\boldsymbol{i}^{(n)}\rangle\}_{\boldsymbol{i}^{(n)}} \in \Omega_2^{(n)}$ be the computational basis; then, the shortest program that tells $\mathfrak{U}$ how to generate it is such that $\ell(q_{\mathcal{B}}) = O(1)$. Therefore, for all $\boldsymbol{i} \in \Omega_2^*$, $\mathrm{QC}_c\,(|\boldsymbol{i}\rangle) = \mathrm{K}(\boldsymbol{i})$ up to an additive constant.

In [37], a lower bound to the quantum bit complexity of a subset of $\Psi \in \mathbb{H}_n$ is obtained. For the benefit of the reader, we provide the proof of this result by adapting it to the needs of the next section. For any $\Psi \in \mathbb{H}_n$ and $\alpha \geq 0$, let us define the subsets

$$\Omega_2^* \ \supseteq \ \Pi_\alpha(\Psi) \ := \ \left\{p \in \Omega_2^* \,:\, -\log_2|\langle\mathfrak{U}[p]|\Psi\rangle|^2 < \alpha\right\}.$$

Set $\mathrm{QC}_\alpha(\Psi) := \min\limits_{p\in\Pi_\alpha(\Psi)} \ell(p)$; then, $\alpha \geq \beta \Longrightarrow \Pi_\beta(\Psi) \subseteq \Pi_\alpha(\Psi)$, whence

$$\alpha \ \geq \ \beta \ \Longrightarrow \ \mathrm{QC}_\beta(\Psi) \ \geq \ \mathrm{QC}_\alpha(\Psi) \ \geq \ \mathrm{QC}_\infty(\Psi)\,,$$

where $\mathrm{QC}_\infty(\Psi)$ is the length of the shortest classical program p whose only constraint is that $|\mathfrak{U}[p]\rangle$ must not be orthogonal to $|\Psi\rangle$, $|\langle\mathfrak{U}[p]|\Psi\rangle| > 0$. Therefore, if $\mathrm{QC}_c\,(\Psi)$ is attained at q, that is if

$$\mathrm{QC}_c\,(\Psi) \ = \ \ell(q) + \underbrace{\lceil -\log_2|\langle\mathfrak{U}[q]|\Psi\rangle|^2\rceil}_{\beta}\,,$$

then, $\ell(q) \geq QC_\beta(\Psi) \geq QC_\alpha(\Psi)$ for all $\alpha \geq \beta$ and

$$\mathrm{QC}_c\,(\Psi) \ \geq \ \mathrm{QC}_\infty(\Psi) + \beta\,. \tag{24}$$

The following Lemma shows that there are vectors $\Psi \in \mathbb{H}_n$ for which $\mathrm{QC}_\infty(\Psi)$ cannot be small.

LEMMA 4 *Let $\mathbb{K}(d) \subseteq \mathbb{H}_n$ be a d-dimensional subspace; for all $0 \leq a \leq \log_2 d$ there exists $\mathbb{K}_a \subseteq \mathbb{K}(d)$ of dimension $d_a \geq d - 2^a$ such that $\mathrm{QC}_\infty(\Psi) \geq a$ for all $\Psi \in \mathbb{K}_a$.*

Proof. From Lemma 2, there are less than 2^a programs $p \in \Omega_2^*$ with $\ell(p) < a$; it follows that the subspace $\mathbb{H}(a)$ linearly spanned by the corresponding vectors $|\mathfrak{U}[p]\rangle$ has dimension smaller than 2^a. Let $\mathbb{K}(d) \subseteq \mathbb{H}_n$ be any subspace of dimension $d \geq 2^a$ and choose $\mathbb{K}_a \subseteq \mathbb{K}(d)$ orthogonal to $\mathbb{H}(a)$ and thus of dimension $d_a < d - 2^a$. Now, $|\Psi\rangle \in \mathbb{K}_a$ satisfies $\mathrm{QC}_\infty(\Psi) \geq a$, unless there is a program p with $\ell(p) < a$ such that $\langle \mathfrak{U}[p] | \Psi \rangle \neq 0$; this is impossible since, by construction, $|\Psi\rangle$ is orthogonal to the linear span of $|\mathfrak{U}[p]\rangle$ with $\ell(p) < a$. □

When the number of qubits increases, we have seen that the rate of the quantum qubit complexity QC_q can be controlled by means of high probability subspaces. Instead, in the case of the bit quantum complexity QC_c, one has to argue in terms of volumes of vectors. More concretely, one estimates the *relative mass* of unit vectors $|\Psi\rangle \in \mathbb{K}_a$ that satisfy $\mathrm{QC}_\alpha(\Psi) < r$.

Consider Ψ as a point $\boldsymbol{u} \in \mathbb{R}^{2d_a}$ on the unit sphere S_{2d_a} whose coordinates are the real and imaginary parts of the Fourier coefficients of the expansion of $|\Psi\rangle$ with respect to a chosen orthonormal basis in the subspace $\mathbb{K}_a$. Let

$$S_{2d_a}(\theta) \ = \ \int_0^\theta \mathrm{d}\phi \, A_{2d_a - 1}(\sin\phi) \, ,$$

denote the area of the sector of S_{2d_a} consisting of unit vectors $\boldsymbol{u} \in \mathbb{R}^{2d_a}$ which have scalar product $1 \geq \boldsymbol{u} \cdot \boldsymbol{e} \geq \cos\theta$ with respect to a fixed vector $\boldsymbol{e}$. In the above expression

$$A_n(t) \ = \ t^{n-1} \frac{2\,\pi^{n/2}}{\Gamma(n/2)} \, ,$$

with $\Gamma(z)$ the Euler Gamma function, is the area of the unit sphere S_n in $\mathbb{R}^n$ of radius t (notice that area of the unit sphere and area of the sector of angle θ are related by $A_n(1) = S_n(\pi)$).

The sector area can be bounded from above as follows:

$$S_{2d_a}(\theta) \ = \ \frac{2\pi^{d_a - 1/2}}{\Gamma(d_a - 1/2)} \int_0^\theta \mathrm{d}\phi \, \sin^{2d_a - 1}\phi \ \leq \ \frac{2\pi^{d_a - 1/2}}{\Gamma(d_a - 1/2)} \sin^{2(d_a - 1)}\theta \, . \quad (25)$$

Let $\widetilde{S}_{2d_a}(\theta)$ denote the sector area $S_{2d_a}(\theta)$ normalized to that of the unit sphere, $A_{2d_a}(1)$. Then

$$\widetilde{S}_{2d_a}(\theta) \ := \ \frac{S_{2d_a}(\theta)}{A_{2d_a}(1)} \ \leq \ \frac{2\pi^{d_a - 1/2}}{\Gamma(d_a - 1/2)} \frac{\Gamma(d_a)}{2\pi^{d_a}} \sin^{2(d_a - 1)}\theta \quad (26)$$

$$< \ d_a \, \mathrm{e}^{-(d_a - 1)(\theta - \pi/2)^2} \, , \quad (27)$$

where the last inequality comes from expanding $f(\theta) := \log\sin\theta$ around $\pi/2$,

$$f(\theta) \;=\; -\frac{1}{2}(\theta-\pi/2)^2+\frac{1}{6}f^{'''}(\overline{\theta})\,(\theta-\pi/2)^3 \;\leq\; -\frac{1}{2}(\theta-\pi/2)^2,$$

with $\theta \leq \overline{\theta} \leq \pi/2$, and from the fact that $f^{'''}(\overline{\theta}) \geq 0$. We now use (27) to estimate the relative volume, $F_a(p,\alpha)$, of the subset

$$\mathcal{F}_a(p,\alpha) \;:=\; \Big\{|\psi\rangle \in \mathbb{K}_a \;:\; -\log_2|\langle\mathfrak{U}[p]|\psi\rangle|^2 < \alpha\Big\}$$

consisting of vectors with penalty smaller than α with respect to a given output $|\mathfrak{U}[p]\rangle$. Since $2^{-\alpha/2} < |\langle\mathfrak{U}[p]|\psi\rangle| = \cos\theta = \sin(\pi/2-\theta) \leq \pi/2-\theta$,

$$F_a(p,\alpha) \;<\; d_a\,\mathrm{e}^{-(d_a-1)2^{-\alpha}}\,.$$

From this inequality we further deduce

LEMMA 5 *The set*

$$\mathcal{F}^r_a(\alpha) \;:=\; \Big\{|\psi\rangle \in \mathbb{K}_a \;:\; \mathrm{QC}_\alpha(\psi) < r\Big\}$$

has relative volume $F^r_a(\alpha)$ such that $F^r_a(\alpha) < d_a\,2^r\mathrm{e}^{-(d_a-1)2^{-\alpha}}$.

Proof. If $|\psi\rangle \in \mathbb{K}_a$ is such that $\mathrm{QC}_\alpha(\psi) < r$, then $-\log_2|\langle\psi|\mathfrak{U}[p]\rangle|^2 < \alpha$ for at least one program p with $\ell(p) < r$; the result then follows since there are $\leq 2^r$ such programs. □

The complement $\mathcal{G}^r_a(\alpha)$ of the set $\mathcal{F}^r_a(\alpha)$ consists of $|\psi\rangle \in \mathbb{K}_a$ such that either

$$-\log_2|\langle\psi|\mathfrak{U}[p]\rangle|^2 \;\geq\; \alpha \qquad \text{or} \tag{28}$$

$$-\log_2|\langle\psi|\mathfrak{U}[p]\rangle||^2 \;<\; \alpha \qquad \text{and} \qquad \ell(p) \geq r\,. \tag{29}$$

In other words, from (24) and Lemma 4, it turns out that $\mathcal{G}^r_a(\alpha)$ consists of $|\psi\rangle \in \mathbb{K}_a$ such that

$$\mathrm{QC}_c\,(\psi) \;\;\geq\;\; \mathrm{QC}_\infty(\psi)+\alpha \;\geq\; a+\alpha \quad \text{or} \tag{30}$$

$$\mathrm{QC}_c\,(\psi) \;\;\geq\;\; r-\log_2|\langle\psi|\mathfrak{U}[p]\rangle||^2 \;\geq\; r\,. \tag{31}$$

Notice that the relative volume $G^r_a(\alpha)$ of $\mathcal{G}^r_a(\alpha)$ is large, $G^r_a(\alpha) \geq 1-\varepsilon$, if the relative volume of $\mathcal{F}^r_a(\alpha)$ is small, $F^r_a(\alpha) \leq \varepsilon$.

PROPOSITION 3 *For any $\varepsilon \geq 0$ and $n \in \mathbb{N}$ large enough, there exists a subspace $\mathbb{K}_{n-1} \subset \mathbb{H}_n$ of dimension $\geq 2^{n-1}$ containing a subset $\mathcal{G}_{n-1}$ of relative volume $G_{n-1} \geq 1-\varepsilon$ such that, for all $|\Psi\rangle \in \mathcal{G}_{n-1}$,*

$$2-\varepsilon \;\leq\; \frac{\mathrm{QC}_c\,(\Psi)}{n} \;\leq\; 2+\varepsilon\,.$$

Proof. Choose $\mathbb{H}(d) = \mathbb{H}_n$ and $a = n - 1$ in Lemma 4; then, there exists a subspace $\mathbb{K}_{n-1} \subset \mathbb{H}_n$ of dimension $d_{n-1} \geq 2^n - 2^{n-1} = 2^{n-1}$ such that $\mathrm{QC}_\infty(\Psi) \geq n - 1$ for all $\Psi \in \mathbb{K}_{n-1}$. Setting $r = 2n$ and $\alpha = n - 1 - 2\log_2 n$ in Lemma 5, one gets

$$F^{2n}_{n-1}(n - 1 - 2\log_2 n) \; < \; \mathrm{e}^{-(1-2^{-n+1})n^2 + (3n-1)\log 2}.$$

As for n sufficiently large the $\Psi \in \mathbb{K}_{n-1}$ violating $\mathrm{QC}_{n-1-2\log_2 n}(\Psi) < 2n$ have relative volume $\geq 1 - \varepsilon$, the result follows from the lower bound in Example 4.1 and the upper bounds (30) and (31). □

3.2.1. QC_c *for ergodic qubit sources*

If the n-qubit states Ψ are emitted by a quantum source $(\mathcal{A}, \omega)$, the result of Proposition 3 is independent of the state of the corresponding quantum spin chain and thus of the overall statistics with which they occur.

Thus one could say that, for very large n, a fraction of mass almost 1 of n-qubit vector states belonging to a subspace of dimension not less than 2^{n-1} has quantum bit complexity per symbol close to 2 that is twice the quantum qubit complexity per symbol of all pure states in high probability subspaces of a Bernoulli quantum source (see Example 3). However, in the latter case the fact that the complexity rate $\simeq 1$ follows from the specific structure of the state ω on the quantum spin chain $\mathcal{A}$.

We shall now assume that the n-qubit vectors Ψ are qubit strings emitted by an ergodic qubit source $(\mathcal{A}, \omega)$ with entropy rate $s(\omega) \leq 1$; more specifically, we shall identify them with projectors $|\,\Psi\,\rangle\langle\,\Psi\,|$ belonging to the local subalgebras $\mathcal{A}^{(n)} = M_2(\mathbb{C})^{\otimes n}$.

We can thus exploit the existence of sequences of universal typical projections $Q^{(n)}_{s,\delta}$ as in the description of the proof of Theorem 2.

The first step is a generalization of Example 5 that upper bounds the quantum bit complexity taking into account the entropy of the source.

LEMMA 6 *Let $(\mathcal{A}, \omega)$ be an ergodic qubit source. For all $\varepsilon > 0$, there exists a sequence of typical projections Q_n such that for all $\Psi \in \widetilde{\mathbb{K}}_n := Q_n\mathbb{H}_n$ one has*

$$\mathrm{QC}_c(\Psi) \; \leq \; 2n(s(\omega) + \varepsilon) + C\,,$$

where C is a constant independent of Ψ.

Proof. Consider the sequence of typical projectors $Q_n := Q^{(n)}_{s,\delta}$ and the corresponding typical subspaces $\mathbb{K}_n$ of dimension $q_n = \mathrm{Tr}(Q_n)$ with chosen orthonormal bases $\{\Phi_j\}_{j=1}^{q_n}$. For all normalized $\Psi \in \mathbb{K}_n$, there must be an index k such that $|\langle\Psi|\Phi_k\rangle|^2 \geq q_n^{-1}$. Further, let U be the unitary operator that rotate the basis vectors Φ_j into the computational basis such that $|\Phi_k\rangle = U|\boldsymbol{k}^{(n)}\rangle$, with $\boldsymbol{k}^{(n)}$ a digit string of length $\lceil \log_2 q_n \rceil$.

In order to prove the upper bounds in Theorem 2, we use the fact that, for any $\eta > 0$, a unitary rotation U can be approximated by another unitary V such that $\|U - V\| \leq \eta$ which can be implemented by a universal QTM $\mathfrak{U}$. Moreover, the description of V requires a classical program consisting of an amount of bits which is independent of ω and of its restriction to the subalgebra $\mathcal{A}^{(n)} \ni |\,\Psi\,\rangle\langle\,\Psi\,|$.

Let p be the program which provides the UQTM $\mathfrak{U}$ with the specification of $\boldsymbol{k}^{(n)}$ and V such that $|\mathfrak{U}[p]\rangle = V|\boldsymbol{k}^{(n)}\rangle$; then, using (20),

$$\ell(p) \ = \ \log_2 q_n \ + \ C_1 \ \leq \ n(s(\omega) + \varepsilon) + C_1 \,,$$

with C_1 independent of $\boldsymbol{k}^{(n)}$ and V. Since

$$|\,\langle\Psi|\mathfrak{U}[p]\rangle - \langle\Psi|\Phi_k\rangle\,| \ = \ |\langle\,\Psi\,|\,(V - U)\,|\boldsymbol{k}^{(n)}\,\rangle| \ \leq \ \eta \,,$$

it turns out that, for an appropriate choice of η,

$$|\langle\Psi|\mathfrak{U}[p]\rangle|^2 \ \geq \ q_n^{-1}\,(1 - \eta\sqrt{q_n})^2 \ \geq \ q_n^{-1}\,C_2 \,,$$

with C_2 a constant independent of q_n, whence the result follows from

$$-\log_2|\langle\Psi^{(n)}|\mathfrak{U}[p]\rangle|^2 \ \leq \ n(s(\omega) + \varepsilon) + C_3 \,.$$

□

The second step consists in a generalization of Lemma 4 which shows that the typical subspaces $\mathbb{K}_n$ contain subspaces $\widetilde{\mathbb{K}}_n$ which are themselves typical and such that, for any $|\Psi\rangle \in \widetilde{\mathbb{K}}_n$, $\mathrm{QC}_\infty(\Psi)$ diverges with $n \to +\infty$.

LEMMA 7 *Let $(\mathcal{A}, \omega)$ be an ergodic qubit source. For all $\varepsilon > 0$, there exists a sequence of universal typical projections P_n such that*

$$\mathrm{QC}_\infty(\Psi) \ \geq \ n(s(\omega) - 2\varepsilon) \,,$$

for all $\Psi \in \widetilde{\mathbb{K}}_n = P_n\mathbb{K}_n$.

Proof. Consider universal typical projections Q_n and the corresponding subspaces $\mathbb{K}_n$ of the previous Lemma. The existence of subspaces $\widetilde{\mathbb{K}}_n \subseteq \mathbb{K}_n$ with the stated properties is proved as in the proof of Lemma 4 by choosing $\mathbb{H}(d) = \mathbb{K}_n$ and $a = n(s(\omega) - 2\varepsilon)$. The subspaces $\widetilde{\mathbb{K}}_n$ are orthogonal complements within $\mathbb{K}_n$ of subspaces $\mathbb{H}(n(s(\omega) - 2\varepsilon))$ of dimension $\leq 2^{n(s(\omega)-2\varepsilon)}$. Let R_n project onto subspaces $\mathbb{H}(n(s(\omega) - 2\varepsilon)) := R_n\mathbb{K}_n$; then Proposition 1 (see in particular the discussion following it) gives $\lim_{n\to\infty}\omega(R_n) = 0$. Thus, the typicality of the projections P_n follows from the typicality of Q_n since $P_n = Q_n - R_n$. □

We can now conclude with a result where the state of the quantum spin chain plays a role in extending Proposition 3.

PROPOSITION 4 *Let $(\mathcal{A},\omega)$ be an ergodic quantum source. For all $\varepsilon > 0$, there exists a sequence of typical projections $P_n \in \mathcal{A}^{(n)}$ each containing a subset $\mathcal{G}_n$ with relative volume $G_n \geq 1-\varepsilon$ such that, for all $\Psi \in \mathcal{G}_n$,*

$$2\,s(\omega)-\varepsilon \;\leq\; \frac{\mathrm{QC}_c\,(\Psi)}{n} \;\leq\; 2\,s(\omega)+\varepsilon\,.$$

Proof. We argue as in the proof of Proposition 3. Consider the typical subspaces $\widetilde{\mathbb{K}}_n$ of the previous Lemma and the upper bound on relative volumes in Lemma 5. Choose $d_a = \mathrm{Tr}(P_n)$, $a = n(s(\omega)-2\varepsilon)$, $r = 2n(s(\omega)-\varepsilon)$ and $\alpha = n(s(\omega)-\varepsilon) - 2\log_2 n$. Then,

$$F_a^r(\alpha) \;<\; \exp\Big(-n^2(1-2^{-n\varepsilon}-2^{-n(s(\omega)-\varepsilon)}) + \\ 3n(s(\omega)-\varepsilon)\log 2 + \log(1-2^{-n\varepsilon})\Big)\,.$$

Therefore, the set of $\Psi \in \widetilde{\mathbb{K}}_n$ such that $\mathrm{QC}_c\,(\Psi) \geq 2n(s(\omega)-3\varepsilon) - 2\log_2 n$ has the relative volume which tends to 1 with increasing n. □

In view of the above result, for $n \gg 1$, a large fraction of n-qubit vectors of high probability subspace with respect to ω have a rate of quantum bit complexity which becomes arbitrarily close to twice the entropy rate of the ergodic quantum binary source which emits them.

4. Conclusions

The birth and rapid development of quantum information and computation theory have prompted the extension of many classical concepts to the quantum realm. In particular, the notion of algorithmic complexity has been generalized in a number of ways in order to address the issue of randomness of quantum states.

Among the results of classical algorithmic complexity theory, an interesting one for its connections with the complexity of the time-evolution is the equality proved by Brudno between the complexity per symbol of (almost all) the trajectories of ergodic classical dynamical systems and their (dynamical) entropy rate.

In this paper, after a short review of the basic tenets of algorithmic complexity in a commutative setting, we have considered two extensions to quantum systems that are based on the reproduction of quantum states by means of universal quantum Turing machines that are provided with quantum descriptions of these states in one definition, which we have termed quantum qubit complexity, and with classical descriptions in the other definition, which we have referred to as bit quantum complexity.

We have briefly sketched how a Brudno-like relation exists between the qubit quantum complexity per symbol of quantum ergodic sources and their von Neumann entropy rate and found an analogous relation for the quantum bit complexity but with a scaling factor of 2 and a validity that, unlike in the former case, involves not only typical subspaces, but also volumes of vectors in these subspaces of relative mass almost 1.

Bibliography

[1] L. M. Adleman, J. Demarrais, M. A. Huang, SIAM J. Comput. **26**, 1524 (1997).

[2] G. Alber *et al.*, *Quantum Information: An Introduction to Basic Theoretical Concepts and Experiments*, ed., Springer Tracts in Mod. Phys. **173**, Springer-Verlag, Berlin, 2001.

[3] V. M. Alekseev, M. V Yakobson, Phys. Rep. **75**, 287 (1981).

[4] R. Alicki, M. Fannes, *Quantum Dynamical Systems*, Oxford University Press, Oxford, 2001.

[5] F. Benatti, T. Krüger, M. Müller *et al*, Commun. Math. Phys. **265**, 437 (2006).

[6] G. Benenti, G. Casati, G. Strini, *Principles of Quantum Computation and Information*, World Scientific, Singapore, 2004.

[7] I. Bengtsson, K. Życzkowski, *Geometry of Quantum States: an Introduction to Quantum Entanglement*, University Press, Cambridge, 2006.

[8] C. H. Bennett, Sci. Amer. **241**, 22 (1979).

[9] C. H. Bennett, Int. J. Th. Phys. **21**, 905 (1982).

[10] C. H. Bennett, Sci. Amer. **257**, 88 (1987).

[11] C. H. Bennett, D. P. Di Vincenzo, J. Smolin *et al.*, Phys. Rev. A **54**, 3824 (1996).

[12] C. H. Bennett, P. Gács, M. Li *et al.*, Proc. 25th ACM Symp. Theory of Computation, ACM Press, 1993.

[13] C. H. Bennett, R. Landauer, Sci. Amer. **253**, 38 (1985).

[14] E. Bernstein, U. Vazirani, SIAM Journal on Computing **26**, 1411 (1997).

[15] A. Berthiaume, W. Van Dam, S. Laplante, J. Comput. System Sci. **63**, 201 (2001).

[16] I. Bjelaković, T. Krüger, R. Siegmund-Schultze *et al.*, Invent. Math. **155**, 203 (2004).

[17] I. Bjelaković, T. Krüger, R. Siegmund-Schultze et al., *Chained Typical Subspaces — a Quantum Version of Breiman's Theorem*, arXiv: quant-ph70301177.

[18] P. Billingsley, *Ergodic Theory and Information*, J. Wiley, New York, 1965.

[19] D. Bouwmeester *et al.*, *The Physics of Quantum Information*, ed., Springer-Verlag, Berlin, 2000.

[20] O. Bratteli, D. W. Robinson, *Operator Algebra and Quantum Statistical Mechanics I*, Springer, New York, 1987.

[21] O. Bratteli, D. W. Robinson, *Operator Algebra and Quantum Statistical Mechanics II*, Springer, New York, 1981.

[22] S. L. Braunstein, *Quantum Computing*, ed., Wiley-VHC, Weinheim, 1999.

[23] A. A. Brudno, Trans. Moscow Math. Soc. **2**, 127 (1983).

[24] D. Bruss, G. Leuchs, *Lectures on Quantum Information*, Wiley-Vch GmbH & Co. KGaA, 2007.

[25] C. S. Calude, *Information and Randomness. An Algorithmic Perspective*, Springer, Berlin, 2002.

[26] N. J. Cerf, G. Leuchs, E. S. Polzik,*Quantum Information with Continuous Variables of Atmos and Light*, eds., Imperial College Press, World Scientific, 2007.

[27] G. J. Chaitin, J. Assoc. Comp. Mach. **13**, 547 (1966).

[28] G. J. Chaitin, J. of the ACM **22**, 329 (1975).

[29] G. J. Chaitin, *Information, Randomness and Incompleteness*, World Scientific, Singapore, New Jersey, Hong Kong, 1987.

[30] G. J. Chaitin, *The Limits of Mathematics*, Springer, Singapore, 1998.

[31] G. J. Chaitin, *Algorithmic Information Theory*, Cambridge University Press, Cabridge, 1988.

[32] I. P. Cornfeld, S. V. Fomin, Ya. G. Sinai, *Ergodic Theory*, Springer, New York, 1982.

[33] T. M. Cover, J. A. Thomas, *Elements of Information Theory*, Wiley Series in Telecommunications, John Wiley & Sons, New York, 1991.

[34] D. Deutsch, Proc. R. Soc. Lond. **A400**, 97 (1985).

[35] L. Diosi, *A Short Course in Quantum Information Theory: an Approach from Theoretical Physics*, Springer, Berlin, 2007.

[36] R. P. Feynman, *Feynman Lectures on Computation*, Addison-Wesley, 1995.

[37] P. Gács, J. Phys. A: Math. Gen. **34**, 6859 (2001).

[38] P. Gács, *Lecture Notes on Descriptional Complexity and Randomness. Technical Report*, Comput. Sci. Dept., Boston Univ., 1988.

[39] P. Grünwald, P. Vitanyi, *Shannon Information and Kolmogorov Complexity*, arXiv: cs/0410002.

[40] J. Gruska, *Quantum Computing*, Mc Graw Hill, London, 1999.

[41] A. J. P. Hey, *Feynman and Computation*, ed., Perseus Books, Reading MS, 1999.

[42] F. Hiai, D. Petz, Commun. Math. Phys. **143**, 99 (1991).

[43] F. Hiai, D. Petz, J. Functional Anal. **125**, 287 (1994).

[44] A. Kaltchenko, E. H. Yang, Quantum Information and Computation **3**, 359 (2003).

[45] P. Kaye, R. Laflamme, M. Mosca, *An Introduction to Quantum Computing*, Oxford University Press, Oxford UK, 2007.

[46] G. Keller, *Wahrsheinlichkeittheorie*, Lecture Notes, Universität Erlangen-Nurnberg, 2003.

[47] A. N. Kolmogorov, Dokl. Akad. Nauk SSSR **119**, 861 (1958).

[48] A. N. Kolmogorov, Dokl. Akad. Nauk SSSR **124**, 754 (1959).

[49] A. N. Kolmogorov, Problems of Information Transmission **1**, 4 (1965).

[50] A. N. Kolmogorov, IEEE Trans. Inform. Theory **14**, 662 (1968).

[51] R. Landauer, IBM J. Research **3**, 183 (1961).

[52] M. Lebellac, *A Short Introduction to Quantum Information and Quantum Computation*, Cambridge Univeristy Press, 2006.

[53] C. E. Mora, H. J. Briegel, *Algorithmic Complexity of Quantum States*, arXiv: quant-ph/0412172.

[54] C. E. Mora, H. J. Briegel, Phys. Rev. Lett. **95**, 200503 (2005).

[55] C. E. Mora, H. J. Briegel, B. Kraus, *Quantum Kolmogorov complexity and its applications*, arXiv: quant-ph/0610109.

[56] M. Müller, *Strongly Universal Quantum Turing Machines and Invariance of Kolmogorov Complexity*, arXiv: quant-ph/0605030.

[57] M. A. Nielsen, I. L. Chuang, *Quantum Information and Quantum Computation*, Cambridge University Press, Cambridge UK, 2000.

[58] M. Ohya, D. Petz, *Quantum Entropy and Its Use*, Springer, Berlin, Heidelberg, New York, 1993.

[59] J. Rissanen, *Information and Complexity in Statistical Modeling*, Springer Verlag, 2006.

[60] P. C. Shields, *The Ergodic Theory of Disrcete Sample Paths*, AMS, Providence ,1996.

[61] B. Schumacher, Phys. Rev. A **51**, 2738 (1995).

[62] B. Schumacher, *Entropy, Complexity and Computation*, http://physics.kenyon.edu/coolphys/thrmcmp/newcomp.htm.

[63] R. J. Solomonoff, Inform. Contr. **7**, 1 (1964).

[64] R. J. Solomonoff, Inform. Contr. **7**, 224 (1964).

[65] V. A. Uspenskii, A. L. Semenov, A. Kh. Shen, Russian Math. Surveys **45**, 121 (1990).

[66] P. Vitanyi, IEEE Trans. Inform. Theory **47/6**, 2464 (2001).

[67] M. Li, P. Vitanyi, *An Introduction to Kolmogorov Complexity and Its Applications*, 2nd ed, Springer,New York, Berlin, Heidelberg, 1997.

[68] P. Walters, *An Introduction to Ergodic Theory*, Graduate Texts in Mathematics **79**, Springer, New York, 1982.

[69] A. Wehrl, Rev. Mod. Phys. **50**, 221 (1978).

[70] H. White, Erg. Th. Dyn. Sys. **13**, 807 (1993).

[71] A. K. Zvonkin, L. A. Levin, Russian Mathematical Surveys **25**, 83 (1970).