A Quantum to Classical Phase Transition in Noisy Quantum Computers

Dorit Aharonov*

The fundamental problem of the transition from quantum to classical physics is usually explained by decoherence, and viewed as a gradual process. The study of entanglement, or quantum correlations, in noisy quantum computers implies that in some cases the transition from quantum to classical is actually a phase transition. We define the notion of entanglement length in d-dimensional noisy quantum computers, and show that a phase transition in entanglement occurs at a critical noise rate, where the entanglement length transforms from infinite to finite. Above the critical noise rate, macroscopic classical behavior is expected, whereas below the critical noise rate, subsystems which are macroscopically distant one from another can be entangled.

The macroscopic classical behavior in the super-critical phase is shown to hold not only for quantum computers, but for any quantum system composed of macroscopically many finite state particles, with local interactions and local decoherence, subjected to some additional conditions. This phenomenon provides a possible explanation to the emergence of classical behavior in such systems. A simple formula for an upper bound on the entanglement length of any such system in the super-critical phase is given, which in principle can be tested experimentally.

I. INTRODUCTION

Quantum computation is a fascinating subject which manifests the peculiarities of quantum mechanics and uses them in order to achieve an advantage in terms of computational power over classical computers. Shor's algorithm [25] is the most astonishing known example for such an advantage in computational power: It enables one to factor an integer in polynomial time using a quantum computer, whereas the best classical algorithm for this task is sub-exponential. The advantage of quantum algorithms such as Shor's algorithm over classical algorithms suggests, though not proves, that the computational complexity class of quantum computation is not polynomially equivalent to that of classical, even randomized, Turing machines.

In real life, quantum computers will be subjected to noise. We will make here the assumption that the noise is probabilistic and local, meaning that each particle, each time step, suffers a certain faulty event with probability η, which is referred to as the noise rate. Quantum computing is now known to maintain its full computational power even in the presence of such local noise, as long as the noise rate is weaker than a certain threshold η_0 [3,18,19]. On the other hand, it is known [2] that when the noise in the system is stronger than a much higher threshold, η_1, the quantum computation can be simulated efficiently by a classical Turing machine. Trying to put the two results together, we learn that there are two regimes of noise in the quantum computer, in which the computational power of the system is qualitatively different. For weak noise, the computational power is fully quantum, whereas for strong noise it becomes classical. This raises the following question: What is the physical difference between the two noise regimes, which reflects itself in the difference in computational power, and how does the transition between the two different physical behaviors occur?

It turns out that an answer to these questions can be given in terms of the behavior of entanglement, or quantum correlations in the system. Perhaps the best way to explain the notion of entanglement is by saying what entanglement is not: We say a state in the Hilbert space of a composite system $A \otimes B$ is non entangled, if two persons, Alice and Bob, each having access to one side of the system, could construct the overall state by applying local quantum operations on their side, and exchanging classical information, by, say, talking on the phone. Any state in the composite system $A \otimes B$ that cannot be constructed in this way is said to be *entangled*. Here we will be interested not only in whether or not states are entangled, but rather in the amount of entanglement in quantum states. Several possible definitions for the amount of bipartite entanglement have been suggested: The entanglement of formation [7], the asymptotic entanglement of formation [7,15] and the

*Computer Science division, UC Berkeley, E-mail: doria@cs.berkeley.edu

asymptotic entanglement of distillation [7,24]. In this paper we will use the entanglement of formation, but the results hold for any measure of entanglement which satisfies certain continuity requirements.

To study the behavior of entanglement in noisy quantum computers, we define the notion of *entanglement length*. Roughly speaking, the entanglement length is the rate of decay of the entanglement between two disjoint sets of particles, as a function of the distance between the two sets. This is the analogous quantity to correlation length in statistical physics, except that here we will be interested in correlations between two subsets of the system, rather than in two-point correlations. We study the behavior of the entanglement length in the noisy quantum computer as a function of the noise rate. We find that there exists a noise rate, η_1, which depends on the geometry of the system, such that the entanglement length is finite for $\eta_1 < \eta \leq 1$. This means that the entanglement between two sets of particles decays exponentially with the distance between the two sets for this range of noise rates. On the other hand, the entanglement length is shown to be infinite in the range $0 \leq \eta < \eta_0$, where η_0 is the threshold for fault tolerance. This means that the entanglement between two sets of particles can be large regardless of the distance between the two sets. These two facts show the existence of a phase transition in entanglement at a non-trivial noise rate $\eta_0 \leq \eta_c \leq \eta_1$. The system in the sub-critical regime behaves quantumly even on the macroscopic scale- two sets of particles, within macroscopic distance, can share a lot of entanglement, so there is long range entanglement in the system. In the super-critical phase, where entanglement decays exponentially with the distance, the system behaves classically on the macroscopic scale, because two macroscopic subsets within macroscopic distance are practically non entangled.

The results can be generalized to quantum systems which are not necessarily quantum computers. We show that macroscopic classical behavior is expected above the critical noise rate for any macroscopic quantum system with local interactions, local decoherence of a certain type, and some additional restrictions to be discussed later. This shows that the phase transition in entanglement is expected in any such quantum system which exhibits long range entanglement in the absence of noise. Moreover, our results can, in principle, be verified experimentally in any such system, as long as the density matrices of subsystems can be measured accurately enough. The entanglement length can then be numerically computed (this is a difficult computational task, but possible for small subsystems) and compared to the finite entanglement length which is predicted by our analysis.

The emergence of classical macroscopic behavior in large quantum systems has been an intriguing area of research for the last several decades. Perhaps the most common and acceptable explanation so far is by decoherence, i.e. interactions with the environment which cause the system to lose its quantum features. See, for example, [26,10] and references therein. This explanation, however, predicts a gradual transition from quantum to classical behavior. A very interesting implication of the results presented in this paper is that they suggest an alternative way to explain the transition from quantum macroscopic behavior to classical macroscopic behavior in certain physical systems, which is qualitatively different from the standard gradual explanation (but see in this context [17]). The origin for the abrupt phase transition from quantum to classical predicted by our results is that in our model, we combine the decoherence process with the assumption that noise, as well as interactions, are local, where the behavior we are interested in is global.

The first part of the proof involves showing that the entanglement decays exponentially with the distance when η is larger than a certain threshold. To do this, we use a method due to Aharonov and Ben-Or [2] to present the density matrix by mixtures of clustered states, and then we study the behavior of the sizes of the clusters, evolving in time, using a mapping of the problem to percolation. Known results from percolation theory [14] imply an exponentially decaying bound on the probability for distant sets of particles to be in the same cluster, and this implies exponentially small entanglement between the two sets. For the second part of the proof, i.e. in order to show that the entanglement length is infinite for weak noise, we use the threshold result for fault tolerant quantum computation [3,18,19], which enables one to create long range entangled states in the noisy quantum computer.

We start by defining the notion of entanglement, and entanglement length, and then proceed to prove the strong noise case and the weak noise case, which together imply the existence of a phase transition at a critical point. We conclude with generalizations to various quantum systems, and with several open questions regarding possible implications to the transition from quantum to classical.

II. ENTANGLEMENT

The notion of *entanglement* is associated with a state of a quantum system composed of two systems, A and B. The term entanglement refers to the quantum correlations between A and B, in a state which lives in $A \otimes B$. Remarkably, two parts of a composite quantum system can exhibit very strong correlations which cannot be explained classically, unless one drops a very important axiom in physics, namely locality. The remarkable phenomena which can be exhibited due to entanglement between two quantum systems were first discovered by Einstein, Podolsky and Rosen [11], more than 60 years ago, and manifested in Bell's inequality more than 30 years ago [4,5]. However, the elusive phenomenon of entanglement is still far from being understood.

In this paper we will be interested in the *amount* of entanglement in quantum systems; we therefore need a good measure of entanglement. One very important requirement on such a measure is that the entanglement in any state cannot increase by classical communication and local quantum operations on the A and B sides separately; this is the whole essence of the term entanglement. We will denote such a process involving local operations and classical communication by LOCC. A desired property of a measure of entanglement is additivity, meaning that the entanglement in $\rho \otimes \sigma$ should be equal to the sum of the entanglement in ρ and the entanglement in σ. This is required if entanglement is to be viewed as a physical resource.

A natural way to construct a measure of entanglement is to ask whether there is an elementary unit of entanglement, so that any state can be constructed with LOCC given sufficiently many such entanglement units. It turns out that there exists exactly such a unit: the Bell state, $\frac{1}{\sqrt{2}}(|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle)$. It was shown [6,7] that any bipartite quantum state can be generated by Alice and Bob using only LOCC operations given that sufficiently many Bell states are a priori shared between A and B. One can try to use the number of elementary units required to construct a state as a good measure of the entanglement in this state. It is reasonable to take the asymptotic limit of such a process, and to define the entanglement in a state as the following limit. Let ρ be our quantum (pure or mixed) state, and let k_n be the number of Bell states required to generate a state, ρ_n, and let ρ_n approach the state $\rho^{\otimes n}$ as n tends to infinity. The asymptotic entanglement of formation [7,15] of the state ρ is defined to be the infimum over all such processes, of k_n/n, as n tends to infinity. Let us denote this measure by E_f^∞. This measure can be easily shown not to increase by LOCC, and is also clearly partially additive, i.e. $E_f^\infty(\rho \otimes \rho) = E_f^\infty(\rho) + E_f^\infty(\rho)$.

An equally natural definition would be the converse one, called the asymptotic entanglement of distillation, in which one is interested in generating as many Bell states as possible by applying LOCC on many copies of the original state ρ. The asymptotic limit of the ratio between the number of Bell states generated in this way, and n, was defined in [7] to be the asymptotic entanglement of distillation. A more rigorous definition was given by Rains [24]. Let us denote this measure by E_d^∞.

Fortunately, for pure states these two measures coincide, and have a very beautiful form. As was shown in [6], they are exactly the von-Neumann entropy of the reduced density matrix on one part of the system. If $\rho = |\phi\rangle\langle\phi|$,

$$E(A:B, |\phi\rangle\langle\phi|) = S(tr_A|\phi\rangle\langle\phi|) = S(tr_B|\phi\rangle\langle\phi|). \tag{2.1}$$

The entropy of entanglement possesses both additivity and monotonicity under LOCC, and also behaves nicely in many other ways.

The situation for mixed states, however, is much more interesting. It turns out that though the asymptotic distillable entanglement and the asymptotic entanglement of formation coincide on pure states, there are very interesting differences between them when mixed states are considered. Clearly, the asymptotic entanglement of distillation is not larger than the asymptotic entanglement of formation [7]. The question of whether there exist states in which E_f^∞ is strictly larger than E_d^∞ is still open. This irreversible process, in which the entanglement which was inserted into the state cannot be distilled, is called *bound entanglement* [16] and is now being extensively studied.

The asymptotic entanglement of formation is believed to be equal to the following quantity, called the *entanglement of formation* [7], and denoted by E_f. $E_f(\rho)$ is the least expected entropy of entanglement of any ensemble of pure states realizing ρ, or more formally:

$$E_f(A:B,\rho) = \min_{\sum_i w_i|\alpha_i\rangle\langle\alpha_i|=\rho} \sum_i w_i E(A:B, |\alpha_i\rangle\langle\alpha_i|), \tag{2.2}$$

with $w_i \geq 0$. The question of whether E_f is equal or not to E_f^∞ depends upon whether E_f is additive. It is believed that indeed it is the case that E_f is additive, but this is not known.

Let us survey what is known about the above three entanglement measures, in terms of convexity and continuity. Entanglement of formation is trivially convex. Asymptotic entanglement of formation can also be shown to be convex. Here is a sketch of the proof that $E_f^\infty(p\rho_0 + (1-p)\rho_1) \leq pE_f^\infty(\rho_0) + (1-p)E_f^\infty(\rho_1)$. Alice and Bob want to approximate the matrix $(p\rho_0 + (1-p)\rho_1)^{\otimes n}$. This matrix is equal to the mixture $\sum_{a_1,...a_n} p^{n_0}(1-p)^{n_1} \bigotimes_i \rho_{a_i}$, where a_i runs over the set $\{0,1\}$ and $n_0(n_1)$ is the number of $0'$s ($1'$s) in the string $a_1,...a_n$. It can be easily shown that all but exponentially small weight or probability is concentrated on strings of the form $a_1,...a_n$ in which n_0 and n_1 are within $n^{3/4}$ of their expected values, pn and $(1-p)n$, respectively. Hence, Alice and Bob can approximate the overall matrix be deciding classically on one such random string $a_1,...a_n$, according to the appropriate probability distribution, and neglecting the cases where the number of 0's and $1'$s are too far from the expected. Having decided on a string $a_1...a_n$, they can generate the matrix $\bigotimes_i \rho_{a_i}$ using the protocols for ρ_0 and ρ_1 separately. Since the number of $0'$s and $1'$s is close to the expected, and using the definition of E_f^∞ for ρ_0 and ρ_1, this requires at most $(pn + n^{3/4})E_f^\infty(\rho_0) + ((1-p)n + n^{3/4})E_f^\infty(\rho_1)$ plus a small correction. When dividing by n, the contribution of the extra $n^{3/4}$ terms tends to zero and we get the desired result. Currently it is not known whether asymptotic entanglement of distillation is convex or not.

As for continuity, the situation is even less clear. To discuss continuity, we use the trace metric on density matrices, which is defined as

$$\|\rho - \sigma\| = tr|\rho - \sigma| \tag{2.3}$$

where $|A| = \sqrt{A^\dagger A}$. It is known that the above three entanglement measures are continuous, in the sense that if a sequence of density matrices σ_n converges to a density matrix ρ in the trace metric, then the entanglement in σ_n converges to the entanglement in ρ:

$$\lim_{n\longrightarrow\infty} \sigma_n = \rho \quad \Longrightarrow \quad \lim_{n\longrightarrow\infty} E(\sigma_n) = E(\rho). \tag{2.4}$$

However, we will be interested in how different the entanglement of two very close density matrices can be. Entanglement of formation was recently shown [22] to have very strong continuity properties, in the following sense: Given two density matrices of a bipartite Hilbert space of dimension $d \times d'$, which are within ϵ distance one from another, in the trace metric, the entanglement of formation of the two matrices is at most ϵ times some linear function in $\log(d)$ and $\log(d')$, plus a term independent of d and d' which goes to 0 as ϵ goes to 0:

$$\begin{aligned} &|E_f(A:B,\ \rho) - E_f(A:B,\ \sigma)| \leq \\ &9|\rho-\sigma|\log(\max\{d,d'\}) - |\rho-\sigma|\log(|\rho-\sigma|)). \end{aligned} \tag{2.5}$$

This strong continuity implies that when two density matrices of n finite state particles are polynomially close one to another (in the number of particles), the entanglement of formation between them is also polynomially small. It is not yet known whether the asymptotic measures of entanglement possess these nice continuity properties, or not.

In this paper we work with the entanglement of formation, E_f, which is known to be both convex and strongly continuous. We remark that the phenomena presented in this paper depend very weakly on the exact properties of the measure of entanglement which is being used. The results in this paper hold, with straight forward modifications, for any measure of entanglement E which is continuous in a sufficiently strong sense, meaning that two density matrices which are ϵ apart have entanglements not different by more than some polynomial in the number of particles and in ϵ.

III. THE MODEL OF THE QUANTUM SYSTEM

We are interested in quantum systems composed of n two-state particles, embedded on a d-dimensional lattice. Such quantum particles are usually called *qubits* in the context of quantum computers. The Hilbert space of n such particles is the tensor product of n two dimensional complex vector space, $\mathcal{C}^2$, where the basis of

$\mathcal{C}^2$ is standardly taken to be $|0\rangle$ and $|1\rangle$. The system is initialized with a certain state (usually a tensor product state, but not necessarily) and evolves in time via interactions between the particles. Time is discretized into time steps and all interactions are assumed to be instantaneous and occur at integer times. In this model, particles interact only with their nearest neighbors on the lattice. An important assumption is that one particle cannot interact with more than one other particle at any given time. For simplicity, we will assume that the particles interact alternately with particles to each of their sides. For one dimension, i.e. an array of particles, this means that a particle interacts with a particle to its left and to its right alternately. The interaction graph can be easily viewed in a $d+1$ dimensional scheme, which for $d=1$ looks as follows:

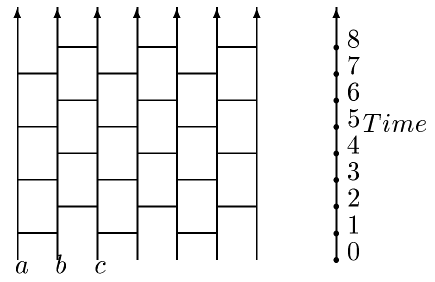

FIG. 1. The vertical axis corresponds to time, and the horizontal axis corresponds to space. Horizontal edges connect two interacting particles. Particles interact alternately with particles to their left or to their right.

Two particles can interact via an arbitrary interaction, and we do not assume anything about the nature or strength of each interaction. After the interactions are turned off, and before the next step of interactions is turned on, we apply the noise step. The noise is assumed here to be local and stochastic, meaning that each particle with a certain probability η undergoes an arbitrary fault process, associated with the noise operator $\mathcal{E}$ operating on the density matrix of one particle.

We make here another assumption, and restrict the noise to be one of the following two processes. The first process, namely independent stochastic collapses, is a process in which at each time step, each particle is measured with independent probability η, in a fixed but arbitrary basis. Alternatively, we can use the depolarization model, in which at each time step, each particle, with independent probability η, is replaced by a particle in a completely mixed state. In the rest of the paper, we will assume that the noise model is independent stochastic collapses, but all results can be easily stated using stochastic depolarization.

This model of noisy quantum circuits is somewhat weaker than the general model of noisy quantum circuits usually discussed in the literature, due to our restriction that only nearest neighbor particles are allowed to interact. Nevertheless, it was shown in [3,13] that, like general quantum circuits, such d-dimensional noisy quantum circuits are capable of performing fault tolerant quantum computation, as long as the noise rate η is smaller than a certain threshold, η_0. The threshold, however, is worse than the threshold without the nearest neighbor restriction, by one or two orders of magnitude, depending on the dimension. The fact that such systems are capable of performing quantum computation will prove to be important for the result of the phase transition in these systems.

IV. ENTANGLEMENT LENGTH

For quantum systems which are embedded on a lattice, the notion of distance between sets of particles is well defined. In this case, one can define the *entanglement length* in the system. We would like to define an analogous quantity to the standard correlation length from statistical physics. In this case, one says that the correlation length in a physical system is ξ if the correlation between the outcomes of a certain observable O measured at sites a and b, decays exponentially with the distance between them, $d(a,b)$, where the distance is scaled in the decay factor by ξ:

$$< O_a O_b > \propto e^{-\frac{d(a,b)}{\xi}}. \tag{4.1}$$

More precisely, ξ is defined to be the following quantity:

$$\xi^{-1} = \lim_{d(a,b) \longmapsto \infty} \left(\frac{-log < O_a O_b >}{d(a,b)} \right), \tag{4.2}$$

where the system which is considered is usually assumed to be spatially homogeneous so the choice of a, b does not matter.

In analogy with correlation length, we could have defined the entanglement length in the quantum system to be μ if the entanglement between two particles, a and b, decays exponentially with the distance between them, $d(a, b)$, where the distance is scaled in the decay factor by μ:

$$E(a : b) \propto e^{-\frac{d(a,b)}{\mu}}. \tag{4.3}$$

However, there are a few problems with this definition, which will force us to modify it slightly. The first modification is necessary due to the fact that entanglement is a non-local quantity. It might well be that the system contains a lot of entanglement, but small subsets of the system are completely unentangled. For example, in fault tolerant quantum computers, the entanglement is bound to be shared by large sets of qubits, and in order to see entanglement it is necessary to probe large subsets of the system. We will therefore be interested not in two point correlations, but in entanglement between two sets A and B of arbitrary sizes.

Another problem is the following. In systems which are homogeneous in space and time, one can easily take the limit of the size of the system to infinity, and therefore the asymptotic behavior in equation (4.2) is well defined. However, we are interested in fault tolerant quantum computers, which are not homogeneous in space, nor in time. Roughly speaking, we will say that the entanglement length in the system is μ if the entanglement between any two sets A and B is *bounded* by a function which decays exponentially with the distance between the sets, where the decay factor is scaled by μ. The fact that we are interested in a bound, and not in the exact behavior of any pair of sets, allows for non homogeneity in space. To allow for non homogeneity in time, we will consider the average entanglement between A and B over time. This corresponds to the following behavior:

$$< E(A : B) >_t \propto poly(|A|, |B|) e^{-\frac{d(A,B)}{\mu}}. \tag{4.4}$$

where $|A|$ is the number of particles in A and similarly for B. We allow the additional polynomial factor due to the fact that for sets which are not too large compared to the distance between them, the exponential decay dominates the polynomial in the sizes of A and B, and what we will see is merely an exponential decaying behavior. We claim that it is not reasonable to consider two sets of particles which are very large, and to study the behavior of the entanglement they share as a function of the distance between them, in the range where that distance is extremely small compared to the sizes of the sets.

The characterization of μ by equation (4.4) is very helpful to keep in mind, and in fact this is the definition of entanglement length which we will be using in this paper. However the above definition is not quite rigorous, and for the sake of completeness we next show how to make it rigorous by giving it a similar form to that of equation (4.2). In order to do that, we first need to make the notion of a quantum system more precise. In the non-homogeneous case, it is not clear what the notion of an infinite system means. We therefore define a quantum (infinite) system to be a sequence of quantum systems, $\{Q_n\}_{n=1}^{\infty}$, where Q_n consists of n particles, and we think of n as growing to ∞. For a given n, Q_n is a finite system in space, which evolves in time from $t = 0$ to $t = T_n$. Since each Q_n is finite in space, in order to take a limit similar to that of equation (4.2), we need to consider a sequence of pairs of sets, A and B which belong to larger and larger systems. We thus add a subscript n to the subsets A_n and B_n, indicating that they belong to the quantum system Q_n. We would now like to translate the fact that we are interested in sets which are not too large compared with their distance to a precise restriction on the sequences of sets $\{A_n\}$,$\{B_n\}$. The weakest condition which we can impose, to avoid pathologic cases, is that $\lim_{n \longmapsto \infty} |A_n| \cdot |B_n| / exp(d(A_n, B_n)) = 0$, meaning that the sizes of A_n and B_n are not growing exponentially or faster than exponentially with the distance between them. Finally, we want to take care of the fact that we are interested in the largest entanglement length which can be observed in the system. This corresponds to taking the infimum over all such sequences of A_n and B_n. All this translates to the following definition:

Definition 1 *The entanglement length* μ *of a quantum system* $\{Q_n\}_{n=1}^{\infty}$ *is defined by:*

$$\mu^{-1} = \inf \liminf_{n \longmapsto \infty} \left(\frac{-log < E(A_n : B_n) >_t}{d(A_n, B_n)} \right)$$

where the infimum is taken over all pairs of sequences, $\{A_n\}_{n=1}^{\infty}, \{B_n\}_{n=1}^{\infty}$*, such that* A_n *and* B_n *are disjoint sets in* Q_n*,* $\lim_{n \longmapsto \infty} d(A_n, B_n) = \infty$ *and* $\lim_{n \longmapsto \infty} |A_n| \cdot |B_n| / exp(d(A_n, B_n)) = 0$.

Note that if one plugs into definition 1 the exponential behavior of equation (4.4), the contribution of the polynomial factor in equation (4.4) tends to zero due to the requirements on A_n and B_n, and the correct μ pops out. Though definition 1 might seem complicated, calculating the above infimum turns out to be very simple in all our applications.

The notion of a family of quantum systems might seem rather unrestricted at this point. A natural family of circuits to consider is the family of quantum circuits associated with a certain computational problem. The nth circuit Q_n would be the quantum circuit solving the problem for input strings of n bits, and the entanglement is averaged over the computational time.

V. CLUSTERED DENSITY MATRICES

We now proceed to study the entanglement length in d-dimensional noisy quantum circuits, in the strong noise regime. In this case, we will try to bound the entanglement in the system from above. Now a very useful observation is in place. After a particle was hit by the noise process, it is no longer entangled with the rest of the system. In other words, in both noise processes which we consider, the density matrix after applying the noise process can be written in the form

$$\mathcal{E} \circ \rho = (1 - \eta)\rho + \eta \sum_i p_i \rho_i^Q \otimes \rho_i^q \tag{5.1}$$

where the index q refers to the noisy particle, and Q refers to the rest of the system. For example, for the stochastic noise process, in which the last two state particle is measured in the basis $\{|0\rangle, |1\rangle\}$, the resulting density matrix would be of the form:

$$\mathcal{E} \circ \rho = (1 - \eta)\rho + \eta \sum_{i=0}^{1} Pr(i) \rho_i^Q \otimes |i\rangle\langle i| \tag{5.2}$$

where ρ_i^Q is the density matrix of all but the last particle, under the condition that the last particle is measured to be in the state $|i\rangle$. We use this observation as follows. We will aim to present the density matrix in such a way that lack of entanglement translates to tensor product structure. In other words, we will present a density matrix as a mixture of tensor product states, as follows:

$$\begin{aligned} \rho(t) &= \sum_i w_i \rho_i(t), \\ \rho_i(t) &= \rho_i^1(t) \otimes \cdots \otimes \rho_i^{m_i}(t) \end{aligned} \tag{5.3}$$

where $\rho_i^j(t)$ is a density matrix which describes a set of particles A_i^j, and for each i the sets A_i^j are a partition of the system. These sets of supposedly entangled particles are called clusters. It should be understood here that given a density matrix, there is no single way to present it as a mixture of clustered states. However, we will define the representation according to the dynamics of the process which generates ρ, so that our representation will be well defined.

Our goal would be to find a way to represent the matrix as a mixture of clustered states with as small clusters as possible. The intuition is that we want to give an upper bound on the amount of entanglement in the system. When all the clusters are of size one, there is no entanglement in the system. We will see later, that this can be generalized to say that small localized clusters imply no entanglement between distant sets. We will thus try to keep the clusters as small as possible. The way we do this is as follows. In a quantum computer, the initial state is a basic state, which is a pure state in which all particles are in tensor product with one another:

$$\rho = \rho(1) \otimes \rho(2) \otimes \cdots \otimes \rho(n). \tag{5.4}$$

Thus, for $t = 0$, all clusters are of size 1. Given any clustered states description of the density matrix at time t, we can obtain a clustered state description for the matrix at time $t+1$ as follows. From each participant in the mixture, $\rho_i(t)$, we obtain $\rho_i(t+1)$ which will be a mixture of clustered states. $\rho(t+1)$ will then be a mixture of all $\rho_i(t+1)$. To obtain $\rho_i(t+1)$ from $\rho_i(t)$, we first apply the interaction step, and then apply the noise step.

To apply the interaction step of time t, we apply for each interaction at that time step the unitary transformation corresponding to the interaction, on the appropriate pair of particles. If the two particles are in one cluster in $\rho_i(t)$, then we simply apply the appropriate unitary matrix, corresponding to the interaction, on the density matrix describing this cluster, and there is no need to change the clusters. However, if the two particles are from two different clusters, we can no longer keep the two clusters in tensor product, because in general they will be entangled after the interaction. Therefore, we first join the two clusters together, by taking the tensor product of the density matrices describing the two clusters, and then apply the appropriate unitary matrix on the new big cluster. The resulting state after all interactions of time step t were applied is therefore, in general, a clustered state with larger clusters than the state $\rho_i(t)$.

We then apply the noise step on the resulting clustered state. Recall that a measurement detaches a particle from its cluster, and thus after a measurement the particle is a cluster of its own. To apply the noise process, we transform the state to a mixture of states, which are the results of all possible combinations of which particles where measured, with the appropriate probabilities. Clusters in the state can only shrink due to this process.

We would now like to understand the typical size of clusters in this representation of the density matrix. Before we do this in a more formal way, let us gain some intuition. If the system were noise free, very soon all the clusters would become one giant cluster of n particles. What makes the situation more interesting are the stochastic collapses, which separate a measured particle from its cluster, thereby decreasing the size of the cluster by one, and creating another cluster of size one. One can view the noisy quantum evolution in time as a struggle between two forces: The interactions, which tend to join clusters and entangle the different parts of the system, and the stochastic collapses, which tend to detach particles from their clusters, thereby destroying this entanglement constantly. A crucial point here is that the two competing forces are matched in power, since they both operate on order of n particles each time step. We thus expect a critical error rate, η_c, at which the two forces are equal, and at which some transition between the dominance of the entangling interaction process transforms to the dominance of the disentangling noise process. We now go on to see this phenomenon more rigorously, using a map to a percolation process.

VI. THE PERCOLATION PROCESS

It turns out that the dynamics of the clusters in the above description are intimately connected with a percolation process on the *quantum circuit* itself. The percolation process on the graph is defined as follows: For each time step, each vertical edge, along the ith wire, between time t and $t+1$, is erased with independent probability η. In the cluster picture, this corresponds to the collapse of the ith particle between time steps t and $t+1$, which disentangles the particle's past and future. Thus an event in the probability space of the noise process, i.e. a specific choice of which particles collapsed during the process, is mapped to an event in the percolation process, in which the corresponding vertical edges are erased. Since events in the stochastic noise process correspond to members in the mixed density matrix, we have a map between clustered states arising from our cluster dynamics, and realizations of the percolation process. This map preserves the probability measure.

We now claim that clusters in the clustered state correspond to connected components in the percolation process:

Lemma 1 Correspondence Lemma: *Two particles a and b are in the same cluster at time t, in one* *realization of the noise process in the cluster model, if and only if (a,t) and (b,t) are connected by a path which uses only edges up to time t in the corresponding realization of the percolation model.*

Proof: To prove this combinatorial lemma we use induction on t. For the base of the induction, $t = 0$, the correspondence is true by definition. Let us now assume that the lemma is correct for t, and prove it for $t+1$. To apply the induction step, the following observation comes in handy. Each path that connects $(a, t+1)$ and $(b, t+1)$ in the percolation process, is actually a concatenation of alternating paths, occurring either after time t or at times up to t. We denote the points at which the different concatenated paths connect one to another by $(x_1,t_1), ..., (x_{2k}, t_{2k})$. It is easy to see that there is always an even number of such points, and that $t_1 = t_2 = ... = t_{2k} = t$. Let us call the particles $x_1, ..., x_{2k}$ the connection particles. We shall also denote $a = x_0$, $b = x_{2k+1}$. A schematic example for the one dimensional quantum circuit case is shown in figure 2.

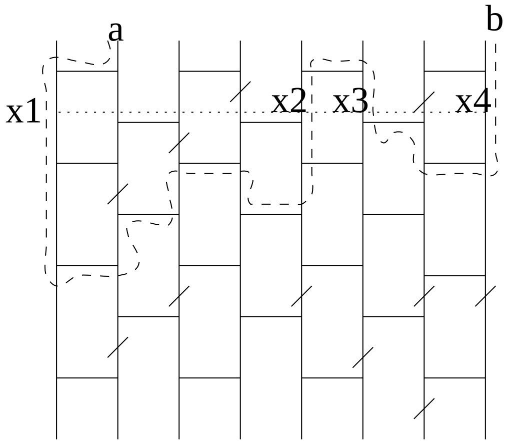

FIG. 2. A path connecting the two particles, a and b, at time $t+1$ can be represented as a concatenation of paths which are restricted alternately to the time intervals $[0,t]$ and $[t, t+1]$. The path $(a, t+1) \longmapsto (b, t+1)$ in the figure is a concatenation of the paths $(a, t+1) \longmapsto (x_1, t) \longmapsto (x_2, t) \longmapsto (x_3, t) \longmapsto (x_4, t) \longmapsto (b, t+1)$.

Let us now prove the first direction: Let a and b be two particles connected at time $t+1$ in the percolation model. We want to show that a and b are in the same cluster at time $t+1$. To see this, we show that all particles $a = x_0, x_1, .., x_{2k}, x_{2k+1} = b$ are in the same cluster at time $t+1$. For the pairs of particles x_{2i+1} and x_{2i+2}, i.e. pairs in which the first particle has odd index, this is true since they are connected by a path confined to time steps t or earlier, so by the induction assumption, they are in the same cluster at time t. Moreover, none of the connection particles collapsed between time steps t and $t+1$, due to the fact that they connect between a path before time t and a path after time t. Therefore x_{2i+1} and x_{2i+2} are in the same cluster also at time $t+1$. (In the schematic example, this shows that particles x_1 and x_2 are in the same cluster at time $t+1$, and similarly x_3 and x_4.) Now by definition of the connection particle, there is a path after time t connecting x_{2i} and x_{2i+1}, which means that x_{2i} and x_{2i+1} interact at time $t+1$. At the edges of the chain, i.e. for $i = 0$ or $i = k$, it might be that x_{2i} and x_{2i+1} are the same particle, and therefore they are trivially connected. (In the example, this corresponds to the interaction at time $t+1$ between a and x_1, and between x_2 and x_3, and to the fact that $b = x_4$.) By the definition of the clusters' evolution in time, the fact that particles interact imply that their clusters are joined, and therefore x_{2i} and x_{2i+1} are in the same cluster at time $t+1$. Combining this with the fact that x_{2i+1} and x_{2i+2} are in the same cluster at $t+1$, this implies that all the particles $a = x_0, x_1, .., x_{2k}, x_{2k+1} = b$ are in the same cluster at time $t+1$, which completes one direction of the induction step.

Let us now prove the other direction of the induction step. We want to prove that there is a path connecting $(a, t+1)$ and $(b, t+1)$ in the percolation process, assuming that a and b are in the same cluster at time $t+1$. This cluster of time step $t+1$, which contains a and b, was generated by joining together several smaller clusters, which existed after the noise step of time step t. It is easy to see that there is a subset of those clusters, $C_{-1}, C_0, ..., C_{k-1}, C_k$, such that $a \in C_{-1}$ and $b \in C_k$, and such that each two subsequent clusters, C_i and C_{i+1}, were connected at time $t+1$ by a unitary gate. Note that each cluster C_i, except maybe C_{-1} and C_k, consists of at least two particles: one particle, denoted by x_{2i+1}, participates in an interaction with a particle from the preceding cluster in the chain, C_{i-1}, and the other particle, x_{2i+2}, participated in an interaction with a particle from the next cluster in the chain, C_{i+1}. To construct a path from a to b at time $t+1$, we first note that by the induction assumption, there is a path connecting the two particles in the same cluster, x_{2i+1} and x_{2i+2}, at time t. Moreover, these particles did not collapse between time step t and $t+1$, because if they did collapse, they would have belonged to a cluster consisting of one particle exactly. Hence there is a path connecting them at time $t+1$. A path between the second particle of C_i and the first particle of C_{i+1} exists because by definition they interact at time $t+1$. Paths from a to the first particle x_1, and from the last particle x_{2k} to b, exist either because they interact at time t, or they are simply the same particle. This enables us to construct a concatenated path from a to b at time $t+1$. $\Box$

We can now investigate the sizes of the connected components in the percolation model and then translate our findings to the cluster model. Such percolation processes [14] are known to exhibit a phase transition at a critical point, above which there is a connected component of linear size, but below which the typical connected components are of logarithmic size at the most. Moreover, the connected components in the sub-critical phase are localized, in the sense that the probability for two particles to be in the same connected component decays exponentially with the distance between these particles.

Let us first release the restriction to a percolation on a rectangle of size $n \times T$ corresponding to the quantum circuit, and consider the infinite lattice. We show the existence of a phase transition in connectivity for percolation on the infinite lattice, and from this we will be able to get information about the finite case. One more simplification which is useful is to notice that by contracting each edge which corresponds to an interaction to one point, we do not change the connectivity properties of the process, and the resulting percolation process is the usual model of bond percolation. For example, in the two dimensional lattice associated with the one dimensional quantum circuit, the interaction edges are exactly the horizontal edges in figure 1, and so after contracting each of these edges to one point, the resulting percolation process is the standard bond percolation on the square lattice, which is rotated by 45 degrees. The contraction therefore transforms the problem to standard bond percolation on translational invariant lattices, which is pretty well understood.

In bond percolation on translational invariant lattices one usually uses p as the probability for one edge to be present, so in our case $p = 1 - \eta$. One can define the critical p, p_c, to be the smallest probability in which the point 0 (an arbitrary fixed point in the lattice) belongs to an infinite connected component with positive probability. In translational invariant lattices, this is also equal to the smallest p in which the expected size of $0'$s connected component (the set of all points connected to the point 0) is infinite [21,1,8]

$$\begin{aligned} p_c &= inf\{p \mid Pr_p(|H(0)| = \infty) > 0\} = \\ &= inf\{p \mid E_p(|H(0)|) = \infty\}. \end{aligned} \tag{6.1}$$

where $H(0)$ is the connected component of 0.

A theorem by Hammersley [8,14], asserts that for translational invariant lattices of any finite dimension d, for $p < p_c$, the probability $\tau(x, y : p)$ for x to be connected to y, decays exponentially with the distance:

$$\tau(x, y : p) \leq \exp\left(-\frac{d(x,y)}{\xi_p}\right) \tag{6.2}$$

with $\xi_p < \infty$. Above p_c, the probability for the point 0 to be connected to infinity is larger than 0 by definition of p_c, so $\xi_p = \infty$ for $p > p_c$.

What can we infer from equation (6.2) regarding the finite case of percolation on the quantum circuit? Note that percolation on the circuit up to time T is simply a restriction of the infinite case to an $n \times T$ rectangle. Thus, the probability for two points in the rectangle to be connected by a path within the rectangle, is smaller than their probability to be connected by a path in the infinite lattice. Hence, an upper bound on the probability for connectedness which holds for the infinite lattice case necessarily holds also for the restricted finite case. Therefore equation (6.2) holds also for the percolation on the $n \times T$ quantum circuit. This means that below the critical p_c, the connected components in the percolation process in the quantum circuit are very small with very high probability.

We end this section with some facts about the value of p_c. In particular, we would like to know that p_c is bounded away from 0 and from 1, so that the two phases of weak and strong connectivity are non trivial. It is easy to calculate the critical probability p_c in the case of the infinite lattice corresponding to a one dimensional quantum circuit. This infinite lattice turns out to be simply the square lattice Z^2, after contracting the interaction edges. For percolation on Z^2 it is known [14] that the critical probability is exactly a half. Let us denote by $p_c(d+1)$ the critical p_c for bond percolation on the infinite lattice corresponding to a quantum circuit of dimension d. Hence:

$$p_c(1+1) = \frac{1}{2}. \tag{6.3}$$

For higher dimensions, we can bound $p_c(d+1)$ away from 0 and 1, but the bounds are not tight:

Lemma 2 $\frac{1}{3} \leq p_c(d+1) \leq \frac{1}{2^{1/d}}$.

Proof: The upper bound comes from the fact that the projection of the $d+1$ percolation with parameter p on a $1+1$ percolation gives percolation in $1+1$ with parameter p^d. This is true since after a particle interacts with a particle along one axis, it waits d time steps before it interacts again with a neighbor along the same axis. This gives the upper bound, since if the original process had exponentially decaying correlations, it cannot be that in the projected process we are above the phase transition where the point 0 is connected with constant probability to infinity. The lower bound is derived from a standard argument which reduces the problem to a branching process. A branching process is a process which starts with one node, and each nodes gives birth to k nodes with some probability distribution $p(k)$, independent of the other nodes. It is a standard result (See [12], for example) that when the expected number of descendents for each node is less than 1, the dynasty dies in finite time with probability 1. To construct the corresponding branching process, observe that the degree of the lattice, after contracting the horizontal edges, is exactly 4, regardless of the dimension d. Thus, there are four edges going out from the point 0. Starting with the point 0, we regard each of its neighbors for which the connecting edge is present in the percolation process as an ancestor of a dynasty. Each such node has three edges going out from it to three neighbors in the lattice, apart from the edge going back to 0. The descendents of the node are those of the neighbors for which the connecting edge is present in the percolation process. Each of these descendents has its own descendents, and so on. When we encounter a node which is already in the dynasty, we do not count it again; this way we have a tree. Clearly, in this way we can reach any node which is in the connected component of 0 in the percolation process. Thus if the branching process is finite, then the connected component of the point 0 is finite. It is easy to see that the branching process starting at each neighbor of the point 0 is finite with probability 1 if $p < \frac{1}{3}$. This is true since the expected number of descendents of each node is exactly the number of its neighbors which are not yet in the dynasty, times p. Since the degree of the graph is 4, and one of the neighbors is the node's ancestor, the expected number of descendents is at most $3p$, which is less than 1 for p less than $\frac{1}{3}$. This gives the desired result. $\Box$

VII. FINITE ENTANGLEMENT LENGTH

The analysis of the percolation process has taught us that when p is smaller than the critical point for percolation, p_c, the connected components in the system are small. We will now go back to the density matrices picture, and use this result to show that the entanglement length is finite when $\eta > 1 - p_c$. The intuition behind the proof is to use the correspondence lemma together with equation (6.2), to show that the clusters are small and localized for $\eta > 1 - p_c$. More precisely, for $\eta > 1 - p_c$, the density matrix of the quantum system can be approximated by a mixture of clustered states with localized clusters of logarithmic size. Thus, distant subsets of particles are with high probability contained in different clusters. This means that most of the weight of the density matrix is concentrated on states in which there is no entanglement between the two subsets. The weight of the states in which there is entanglement between the two sets decays exponentially with the distance between the sets. By continuity of entanglement, this implies that the entanglement between the two sets decays exponentially with the distance. The rate of the decay is the *entanglement length* of the system. Note that in such a situation the entanglement between the two sets becomes negligible already when the distance is of the order of $\log(n)$ particles. The above reasoning translates to the following theorem:

Theorem 1 *Consider a d dimensional quantum circuit with nearest neighbor interactions, subjected to local noise of the type of stochastic depolarization or stochastic collapses, with noise rate η. If the circuit is initialized* *with an unentangled state, i.e. a tensor product state, and if $\eta > 1 - p_c(d+1)$, then the entanglement of formation between any two sets of particles A and B at any time $t \geq 0$ decays exponentially with the distance between the two sets:*

$$E_f(A:B) \leq \min\{|A|,|B|\}|A| \cdot |B| e^{-\frac{d(A,B)}{\xi_{1-\eta}}}.$$

For a general initial state, a similar formula is true except for a correction term which decays exponentially with time:

$$E_f(A:B,\rho(t)) \leq \min\{|A|,|B|\} \Bigg(|A| \cdot |B| e^{-\frac{d(A,B)}{\xi_{1-\eta}}} +$$

$$+ n \min\{|A|,|B|\} e^{-\frac{t}{\xi_{1-\eta}}} \Bigg).$$

Proof: Let us start with the simple case, in which the initial state is a complete tensor product, i.e. all clusters are of size 1. By equation (6.2), the probability for two particles A and B to be connected decays exponentially in the distance between them. The correspondence lemma implies that the probability for two particles to be in the same cluster at time t is equal to the probability they are connected in the percolation model at time t. i.e. the probability for two particles from A and B to be in the same cluster is bounded above by $\exp(-d(A,B)/\xi_{1-\eta})$. Thus, the probability for any pair of particles from A and B to be in the same cluster is bounded above by $|A| \cdot |B| \exp(-d(A,B)/\xi_{1-\eta})$. The density matrix can thus be written as a mixture of one density matrix with weight smaller than $|A| \cdot |B| \exp(-d(A,B)/\xi_{1-\eta})$, and another density matrix which is a mixture of density matrices, where in all these matrices all the particles in A are in different clusters than all the particles in B. The reduced density matrix to A, B of the second matrix is thus separable, and contains no entanglement between A and B. By convexity of entanglement of formation, the entanglement in the entire density matrix is bounded above by the entanglement in the first density matrix, times the weight of this matrix. The entanglement of the first matrix is at most the number of particles in the system, and this gives the desired result. (For measures of entanglement which are not convex, but strongly continuous, one should replace the term $\min\{|A|,|B|\}$ by the appropriate polynomial from the continuity bound.)

We now proceed to the general initial state. We will give an upper bound for the case in which the initial state is one big cluster, and any other case is trivially implied by it. To do this, we have to understand where we have used the fact that the initial state is not entangled. This was used for the base of the induction in the correspondence lemma, where the fact that all clusters are of one particle corresponds to the fact that in the percolation graph, the initial connected components at time 0 are all of size 1. To adapt the situation to the case in which all particles are in one big cluster at time 0, we add a horizontal line of length n connecting all particles to one big connected component at time $t = 0$. The correspondence lemma then goes through. However, equation (6.2) no longer holds. To correct it, we add to it a term which corresponds to the probability for A to be connected to B by a path that goes through time $t = 0$, i.e. through the additional new line we have added to the graph. For such a path to exist, both A and B need to be connected to time 0. The probability for any one of the particles in A to be connected to any one of the n particles at time 0 is at most $n|A|$ times the probability for one particle at time t to be connected to one particle at time 0, which is at most $exp(-\frac{t}{\xi_{1-\eta}})$ by equation (6.2). The same argument applies for the connection from B to time 0, and this gives the desired result. $\Box$

This shows that the system cannot create entanglement between distant sets of particles: Roughly speaking, the typical range of entanglement is microscopic. This is true for any initial condition, where the relaxation time to the typical unentangled state is of the order of $\log(n)$ steps.

This result implies an upper bound on the entanglement length in the quantum system above the critical noise rate, and in particular shows that it is finite. This is done by taking the limit in the definition of entanglement length, which gives:

Corollary 1 *The entanglement length μ_η of a d dimensional quantum circuit with nearest neighbor interactions, subjected to local noise of the type of stochastic depolarization or stochastic collapses, with noise rate η, satisfies:*

$$\mu_\eta \leq \xi_{1-\eta}$$

where ξ_p is defined in equation (6.2). In particular μ_η is finite for $\eta > 1 - p_c(d+1)$.

This gives a bound on the entanglement length, in terms of the correlation length in classical bond percolation. The correlation length of a given lattice can be easily estimated by computer experiments, and analytical bounds are given in [8,14].

VIII. INFINITE ENTANGLEMENT LENGTH

We now want to concentrate on the other noise regime, and show that below the critical noise the entanglement length is infinite. One might naively think that this can be deduced from the fact that the density matrix is a mixture of clustered states with linear sized clusters. However, there is a difficulty in pursuing the connection between clusters and entanglement for this purpose, because of the following reason. The density matrix is actually a mixture of many clustered states. The mixture of two clustered states, with very large clusters, can be a density matrix in which the clusters are of size one. One example of such a case is a mixture of the two states, $\frac{1}{\sqrt{2}}(|0^n\rangle + |1^n\rangle)$ and $\frac{1}{\sqrt{2}}(|0^n\rangle - |1^n\rangle)$, the mixture of which is a non entangled state. Thus, the sizes of the clusters can be used for upper-bounds on entanglement, but it is not clear how to use them in order to show a lower bound on the entanglement in the system.

We therefore need to use different techniques for lower bounds on entanglement. We will use techniques from quantum computation. A quantum computer embedded on a lattice is a special case of the quantum systems we are discussing. The particles are quantum bits, and the interactions are fixed according to the algorithm. Therefore, corollary 1 shows that the entanglement length is finite above the critical noise rate also in fault tolerant quantum computers. For fault tolerant quantum computers we can also analyze the other side of the noise scale, and show that the entanglement length in the system is infinite if the noise rate is below a certain threshold. We will use the threshold result [3,18,19] for fault tolerant quantum computation, which shows that quantum computation can be made robust to noise, using quantum error correcting codes, as long as the noise is smaller than a certain threshold. In fact, here we need the slightly stronger version of the threshold result [3,13], which asserts that this can be done even when the quantum system is embedded on a d dimensional lattice. The threshold is then $\eta_0(d)$, which for $d = 1$ is estimated to be 10^{-7} [3]. In the fault tolerant range, two distant sets of qubits can be entangled, and remain entangled for a long time, with the amount of entanglement independent of the distance.

We now give an example of a quantum computer which exhibits entanglement among distant parts of the system when $\eta < \eta_0(d)$, but the entanglement length is finite for noise $\eta > 1 - p_c(d+1)$. The idea is that a fault tolerant computer can simulate any quantum state, including states which contain entanglement between sets of qubits which are far apart. Hence, we will construct a quantum algorithm in which there is entanglement between two parts of the system, and make it fault tolerant. This can be done in many ways, but here is a simple example, for $d = 1$. Divide the set of qubits to three sets, A, B, C. We will create entanglement among A and B, while leaving the qubits in the middle, C, in a basic state. This will be done by constructing the state:

$$\frac{1}{\sqrt{2}}\left(|0^m\rangle_A \otimes |0^n\rangle_C \otimes |0^q\rangle_B + |1^m\rangle_A \otimes |0^n\rangle_C \otimes |1^q\rangle_B\right) \tag{8.1}$$

on a fault tolerant quantum computer, and keeping this state for a long time, by applying error corrections. This state indeed contains entanglement between the two registers A and B, which are n sites apart. The algorithm which constructs such a state is very simple, and uses only two basic quantum gates: The Hadamard gate, which is a one qubit gate applying the following unitary transformation

$$\begin{aligned} |0\rangle &\longmapsto \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \\ |1\rangle &\longmapsto \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle), \end{aligned} \tag{8.2}$$

and the controlled NOT gate which is a two qubit gate applying the following unitary transformation (said to be applied *from* the first qubit to the second one):

$$|a\rangle \otimes |b\rangle \longmapsto |a\rangle \otimes |a \oplus b\rangle, \tag{8.3}$$

where $\oplus$ means addition mod 2. Using these gates, it is easy to create the state

$$\frac{1}{\sqrt{2}}(|0^{m+q}\rangle + |1^{m+q}\rangle) \tag{8.4}$$

on the first $m+q$ qubits, by applying a Hadamard gate on the first qubit and then controlled NOT gates from the first qubit to the second, from the second to the third, and so on. Then, we want to swap the $m+1, ..., m+q$ qubits to the right. To do this, we first swap the last qubit in B with qubits to its right until it gets to the last site. In the same way we bring the one before last qubit in B to the one before last site, and so on until all qubits in B are in the right most sites, which achieves the desired state with only nearest neighbor interactions.

This algorithm by itself is not fault tolerant, and in the presence of any amount of noise, i.e. $\eta > 0$, the entanglement in the system will be lost immediately. However, we can make this algorithm fault tolerant by the methods in [3,18,19], as long as η is smaller than $\eta(d)$, the threshold for fault tolerance for d-dimensional quantum computers [3,13]. These results are too complicated to explain here in details. In a nutshell, fault tolerance is achieved by encoding the qubits using quantum error correcting codes, and computing the algorithm on the encoded states, while applying quantum error correction on the state frequently. Each qubit is replaced by polylog(n) qubits, encoding its state. The state $|0\rangle$ is encoded by the state $|S_0\rangle$ of polylog(n) qubits, and similarly $|1\rangle$ is encoded by the state $|S_1\rangle$. Let us denote by A', B' and C' the qubits encoding the original sets of qubits A, B, and C, respectively. If no fault occurs, at the end of the algorithm the state of the system will be in the state (8.1) encoded by the quantum error correcting code:

$$|S_0^m\rangle_{A'} \otimes |S_0^n\rangle_{C'} \otimes |S_0^q\rangle_{B'} + |S_1^m\rangle_{A'} \otimes |S_0^n\rangle_{C'} \otimes |S_1^q\rangle_{B'} \tag{8.5}$$

normalized by a factor of $\frac{1}{\sqrt{2}}$. The entanglement in this state will remain there for ever if errors do not occur. However, errors do occur. By fault tolerance, this means that at the end of the computation the density matrix is polynomially close to a density matrix ρ which can be corrected to the correct state (8.5) by noiseless quantum error corrections. Due to continuity of entanglement, it suffices to argue that such ρ contains a constant amount of entanglement. But this is true since we know that ρ can be corrected to the state (8.5) by local operations not involving interactions between A' and B'. Since entanglement cannot increase by local operations, the entanglement between A' and B' in ρ is at least as that in the state encoding (8.5), i.e. one entanglement unit. The distance between the actual density matrix and a correctable density matrix ρ is, by [3], at most the number of time steps t divided by a polynomial in n, because the error introduced each time step is at most 1 over a polynomial in n. This distance is smaller than some constant ϵ as long as the number of time steps is polynomial in the size of the system n. Thus, by strong continuity the entanglement between A' and B' will remain bounded from below by a constant for polynomially many time steps. The average entanglement over time is very close to one, since the time it takes to construct the state is much smaller than the polynomial time for which the entanglement remains in the system. This proves the existence of a non trivial sub-critical side of the phase transition:

Theorem 2 *The entanglement length in the d dimensional fault tolerant quantum computer defined above satisfies*

$$\mu_\eta = \infty$$

for η smaller than the threshold for fault tolerance in d dimensional quantum computers, i.e. $\eta < \eta_0(d)$. $\square$

Corollary 1 and theorem 2 together imply the existence of a phase transition in entanglement length in fault tolerant quantum computers, since they imply the existence of a critical point, at which the entanglement length transforms from infinite to finite.

IX. GENERALIZATIONS TO OTHER QUANTUM SYSTEMS

The model of noisy quantum circuits actually describes not only quantum systems designed to serve as computational devices, but a much broader class of physical systems as well. We first claim that any d dimensional quantum system, in which the particles are located in space with low enough density, and in which interactions occur only between particles which are not too far apart, can be modeled by a quantum circuit. This can be done by discretizing the medium to very small cells, such that each cell contains at most one particle. Time will be discretized to sufficiently small intervals such that a particle can only move to a neighbor cell in one time step. Then, the movement of particles can be modeled by an interaction between an occupied and an un-occupied cell, and since the density of particles is low, one particle never interacts with more than one other particle at the same time, so the notion of quantum gate is appropriate.

We argue that the noise model which is used in this paper, stochastic collapses, actually captures a much more general scenario. It turns out that stochastic collapses are equivalent to the following process of local decoherence, in which each particle interacts with its own independent thermal bath [23]. Consider a particle-bath interaction which causes the off diagonal elements of the particle's density matrix to decay with a certain rate, γ. This may be true when the density matrix is written in a basis which is not necessarily the computational basis. We assume the process is Markovian, so that the environment of the particle is renewed each time step. If the particle-bath interaction operates for time Δt, the (i,j) element of the density matrix transforms to

$$\rho_{i,j} \longmapsto \rho_{i,j} \exp(-\gamma \Delta t (1 - \delta_{i,j})). \tag{9.1}$$

Thus, the off diagonal elements of the density matrix of the particle decay by a certain factor each time step. If we set $\exp(-\gamma\Delta t) = 1 - \eta$, we get

$$\rho_{i,j} \longmapsto (1-\eta)\rho_{i,j} + \eta \rho_{i,j}\delta_{i,j}, \tag{9.2}$$

which is equivalent to a measurement with probability η. Thus stochastic collapses are a good model for such decoherence processes, as long as the decoherence during interactions can be neglected, for example in low density systems or when interactions between particles in the system are instantaneous. Similarly, the depolarization process can be presented as a gradual change of the density matrix of one particle.

The results generalize to non-instantaneous interactions, as long as the interactions last for less than a time step, so that there is a time interval in which decoherence takes place when the particle is not participating in any interaction. This would be obvious if there is no noise during the interactions. However, the results hold even if these interactions are both non-instantaneous and noisy. To see this, we argue in the following way. Let δ characterize the decoherence rate during the interaction. The proof of the upper bound on the entanglement length, for large η, holds regardless of the amount of noise in the interactions, δ, because the proof only uses the noise occurring between interactions. The proof of the lower bound on the entanglement length, for small η, goes through as long as $\eta + \delta$ is smaller than the threshold for fault tolerance. Hence, it is straight forward to see that the results hold for such non-instantaneous interactions as well.

One can relax several other assumptions which we have made. The simplest generalization is to particles of more than two states. Clearly, the results apply to particles with any finite number of states, not necessarily qubits, and different particles need not have the same number of possible states. Another assumption is the restriction of the particles to interact alternately. However, this exact form of the alternating interactions is not important, since any interaction graph of nearest neighbor interactions is a subgraph of the alternating graph, if time is scaled by a factor of $2d$.

The above arguments show that the analysis done in this paper regarding upper bounds on the entanglement length in the super-critical phase in quantum computers holds also for a wide variety of quantum systems with macroscopically many finite-state particles, with local instantaneous interactions and appropriate local decoherence processes. Therefore theorem 1 and corollary 1 can be generalized to these cases. In such quantum systems, our analysis provides an explanation to the emergence of macroscopic classical behavior above the critical noise rate, as will be discussed in the last section. If in addition such a system generates long range entanglement in the absence of weak or no noise, then it is proven to exhibit the same phenomenon of phase transition in entanglement length.

X. OPEN PROBLEMS AND QUANTUM-CLASSICAL TRANSITION

We have proved the existence of a phase transition in entanglement in noisy quantum computers, using a newly defined parameter, the entanglement length. However, we merely introduced in this paper a new phenomenon. The list of questions which remain open is extremely large, and varies on different physical fields, and there are many implications of the results which require further study.

A set of open questions regarding the phase transition comes from statistical physics. For example, what are the critical exponents related to this phase transition? What is the universality class of this phase transition? In fact, it is not clear that there is only one critical point here. In the case of the quantum computer, or other quantum systems, there might well be an intermediate regime of noise, in between the two thresholds, for which the entanglement behaves in a different way, i.e. its dependence on the distance is neither an exponential decay nor constant. The question of showing that there is only one critical point at which a transition occurs from exponential decay to independence on the distance remains open.

An important question is how general these results are, in particular those results concerning the weak noise regime. In this paper, we have been able to show the existence of a non trivial sub-critical phase in fault tolerant quantum computers. Are there more natural quantum systems, in particular systems which are homogeneous (or periodic) in space and in time, which have local interactions and local noise, which are able to maintain long range entanglement in the presence of weak or zero local noise for a long time? Does a random quantum system, i.e. in which random interactions are applied, exhibit long range quantum correlations? Such systems will provide more examples for quantum systems in which a phase transition in entanglement length occurs.

Indeed, it is worthwhile to ask which of our assumptions regarding the properties of the physical system are essential, and which are technical. Intuitively, the most important assumption is that of locality of both the noise and the interactions. This is true since locality is what makes the two competing forces, the interactions which tend to entangle the system and the noise which tends to disentangle it, comparable in power. It is the fact that two forces of approximately even power compete which gives rise to the existence of a critical point in which a phase transition occurs. It seems, however, that the assumption on the exact form of the noise process might not be so important. An open question is to characterize and understand the exact class of noise operators $\mathcal{E}$ for which the results hold. The question is what noise operators can be decomposed into a process in which with a certain probability or amplitude the particle becomes disentangled from the other particles in the system. It also seems that the discretization of the interactions is not essential. However it is not clear how to generalize the results to the case in which the particle interacts simultaneously with all its neighbors and the environment. It seems that a considerably different approach would be needed to deal with this case.

Perhaps the most interesting open problem raised by these results is to further understand their implications on our understanding of the transition from quantum to classical physics. By studying the entanglement in quantum computers we were able to shed some new light on the fundamental question of the transition from quantum to classical; Our results suggest that in some cases this transition is actually a phase transition. Moreover, the paper suggests that the emergence of classical macroscopic phenomena in large quantum systems can be attributed in the case of fault tolerant quantum computers and in other cases to the fact that the noise rate is larger than a certain critical point, so that the entanglement length is finite. The suggestion to explain the transition from quantum to classical macroscopic behavior as a phase transition in entanglement is fundamentally different from the standard point of view of gradual transition, usually explained by decoherence. What is the applicability of this explanation? This question is of course related to the understanding of the generality of the phase transition result. The relation of the phase transition to other abrupt quantum/classical phenomena, such as the threshold for violation of Bell inequalities [17], should also be understood. Another important question is whether there exists some classical or quasi-classical description of the behavior of a quantum system in its finite entanglement length phase. In essence, we expect such a description to exist due to the general intuition that entanglement is what captures the "quantumness" of the system. This intuition raises another related question which is whether the existence of a phase transition in entanglement induces other quantum phase transitions at the same critical point in the same system.

A very interesting problem is to come up with a better order parameter related to entanglement, rather than the entanglement length. There are many problems with the entanglement length as an order parameter. The most important one is that it might be that the system is very entangled, but the entanglement between any two distant subsets is zero. Such is for example a system in the state $\frac{1}{\sqrt{2}}(|0^n\rangle + |1^n\rangle)$, for which any subsystem is non-entangled. Entanglement in such very entangled quantum systems will not be detected when looking at sub-systems, and the entanglement length will therefore contain no information about the actual behavior of the entanglement in the system. Another motivation for this question is provided by [2], where the sizes of the clusters is analyzed relaxing the restriction to nearest neighbor interactions. The sizes of the clusters in this case indeed transform from logarithmic to linear at a critical noise rate. However, the notion of entanglement length cannot be defined in a system without geometry, so it is not clear how to define an order parameter which exhibits the phase transition in this case.

An important observation is that the model discussed here deals with quantum systems out of thermal equilibrium. The quantum systems we consider here can be in a steady state, but the density matrix is not in the Gibbs distribution of the eigenvectors of some Hamiltonian, and there is no temperature associated with the system. Thus the notion of zero noise rate does not coincide with that of zero temperature. In other words, long range entanglement when the noise is zero does not necessarily coincide with long range entanglement in the ground state of the Hamiltonian of the system. The fact that the quantum computer does not achieve thermal equilibrium despite the noise is explained by the fact that the system is constantly cooled by quantum error corrections. It is left as an open problem to further investigate possible thermal equilibrium phase transitions in entanglement, and the connection to the non-equilibrium phase transition presented here. The reader is referred to [20] and references therein for an introduction to non-equilibrium phase transitions.

We conclude with the intriguing open problem of experimental verification of these results. The entanglement length is an observable, and in principle it can be measured. Such an experiment might teach us a lot about quantum entanglement in large scales. Unfortunately, there is no existing physical realization of a quantum computer of more than several qubits, and hence the existence of a phase transition in entanglement length in fault tolerant quantum computer cannot be verified at this stage. However, perhaps it is possible to test the results regarding the upper bound on entanglement length in other quantum systems for which theorem 1 applies, as was discussed in the previous section. One way to go, is to find a quantum system in which it is possible to measure with high enough accuracy the joint density matrix of two subsets of the system. The entanglement between the two sub-systems can then be numerically approximated, using a (doable, but extremely difficult) minimization over equation (2.2). The entanglement can then be found as a function of the distance between the sets, from which the entanglement length can be deduced. It would be extremely interesting to design and perform such an experiment in order to measure the entanglement length in a concrete quantum system, and to compare it with the entanglement length predicted by the upper bound given in corollary 1.

XI. ACKNOWLEDGMENTS

I am most grateful to Michael Ben-Or and to Michael Nielsen. Discussions with them inspired this work. I am also in debt to Joseph Imri, Noam Nisan, David Ruelle, and Wojtek Zurek for interesting and stimulating discussions. I would like to thank Jennifer Chayes, Christian Borgs, Jeong Han Kim, David Aldous and Oded Schramm for useful comments about classical percolation. Thanks to Julia Kempe, Daniel Lidar and Michael Nielsen for useful comments and corrections on early drafts of this paper. This research was supported in part by University of California President's postdoctoral fellowship and NSF grant CCR-9800024.

---

[1] M. Aizenman and D.J. Barsky, Sharpness of the phase transition in percolation models, *Commun. Math. Phys* **108** pp. 489–526 (1987)

[2] D. Aharonov and M. Ben-Or, Polynomial Simulations of Decohered Quantum Computers, *37th Annual Symposium on Foundations of Computer Science* (FOCS) pp 46–55 (1996)

[3] D. Aharonov and M. Ben-Or, Fault-Tolerant Quantum Computation with Constant Error, *Proc. of the 29th Annual ACM Symposium on Theory of Computing* (STOC) (1997), Journal version: Fault-Tolerant Quantum Computation with Constant Error Rate, in press (1999)

[4] J.S. Bell, On the Einstein-Podolsky-Rosen paradox, *Physics* **1** pp. 195–200, (1964)

[5] J.S Bell, On the problem of hidden variables in quantum theory, Rev. Mod. Phys. **38** 447-52, (1966) *Speakable and unspeakable in quantum mechanics*, Cambridge University Press (1987)

[6] C. Bennett, H. Bernstein, S. Popescu, B. Schumacher, Concentrating Partial Entanglement by Local Operations, *Phys. Rev. A* **53**, 2046 (1996)

[7] C. Bennett, D.P. DiVincenzo, J. Smolin and W. Wootters, Mixed State Entanglement and Quantum Error Correction, *Phys.Rev. A* **54** pp. 3824–3851 (1996)

[8] J.T. Chayes, A. Puha and T. Sweet, Independent and dependent percolation, IAS/Park City Mathematics Series, Vol. 6, Probability Theory and Applications (E. Hsu and S. Varadhan, eds.), AMS (1999)

[9] D.P. DiVincenzo, T. Mor, P.W. Shor, J.A. Smolin, B.M. Terhal, Unextendible Product Bases, Uncompletable Product Bases and Bound Entanglement, quant-ph/9908070 (1999)

[10] D. Giulini, E. Joos, C. Keifer, J. Kupsch, I.-O. Stamatescu, and H. D. Zeh, Decoherence and the appearance of a classical world in quantum theory, Springer Verlag Berlin Heidelberg (1996)

[11] A. Einstein, N. Rosen and B. Podolsky, *Phys. Rev.* **47**, 777 (1935)

[12] Feller, An introduction to Probability Theory and its Applications, Vol I, Wiley, pp. 275–276 (1961)

[13] D. Gottesman, Fault-Tolerant Quantum Computation with Local Gates, *J. Modern Optics* **47**, pp. 333–345 (2000).

[14] G. Grimmett, Percolation, Springer-Verlag, New York (1989).

[15] M. Horodecki, P. Horodecki, R. Horodecki, Limits for entanglement measures *Phys. Rev. Lett.* **84** 2014 (2000)

[16] M. Horodecki, P. Horodecki and R. Horodecki, Mixed-state entanglement and distillation: is there a "bound" entanglement in nature? *Phys.Rev.Lett.* **80** pp. 5239-5242 (1998)

[17] L.A. Khalfin and B.S. Tsirelson, Quantum/Classical Correspondence in the light of Bell Inequalities, *Foundations of Physics* **22** 7, pp. 879–948 (1992)

[18] A. Yu. Kitaev, Quantum Computations: Algorithms and Error Corrections, *Russian Math. Surveys*, **52**:6, pp. 1191–1249 (1997)

[19] E. Knill, R. Laflamme and W. Zurek, Resilient quantum computation, *Science*, **279**, p.342 (1998)

[20] J. Marro and R. Dickman, Nonequilibrium Phase Transitions in Lattice Models, Cambridge University Press (1999)

[21] Menshikov M V, Molchanov S A and Sidorenko A F, Percolation theory and some applications, *Itogi Nauki i Techniki (Series of Probability Theory, Mathematical Statistics, Theoretical Cybernetics)* **24** 53-110, (1986)

[22] M.A. Nielsen, Continuity Bounds for Entanglement, quant-ph/9908086 (1999)

[23] G.M. Palma, K-A. Suominen and A.K. Ekert, Quantum computers and dissipation, *Proc. Roy. Soc. Lond. A* **452** pp. 567–584 (1996)

[24] E. Rains, A rigorous treatment of distillable entanglement, quant-ph/9809078 (1998)

[25] P.W. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, *SIAM J. Comp.*, **26**, No. 5, pp 1484–1509 (1997)

[26] Zurek W H, Decoherence and the transition from quantum to classical,*Physics Today* 44(10), pp. 36–44 (1991)