How Long Can a Quantum Memory Withstand Depolarizing Noise?

Fernando Pastawski,[1] Alastair Kay,[1,2] Norbert Schuch,[1] and Ignacio Cirac[1]
[1]*Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Str. 1, D-85748 Garching, Germany*
[2]*Centre for Quantum Computation, Centre for Mathematical Sciences, University of Cambridge, Cambridge CB3 0WA, United Kingdom*
(Received 8 May 2009; published 20 August 2009)

We investigate the possibilities and limitations of passive Hamiltonian protection of a quantum memory against depolarizing noise. Without protection, the lifetime of a single qubit is independent of N, the number of qubits composing the memory. In the presence of a protecting Hamiltonian, the lifetime increases at most logarithmically with N. We construct an explicit time-independent Hamiltonian which saturates this bound, exploiting the noise itself to achieve the protection.

DOI: 10.1103/PhysRevLett.103.080501 PACS numbers: 03.67.Pp, 03.67.Ac

A cornerstone for the majority of applications in quantum information processing is the ability to reliably store quantum information, protecting it from the adversarial effects of the environment. Quantum Error Correcting Codes (QECC) achieve this task by using a redundant encoding and regular measurements which allow for the detection, and subsequent correction, of errors [1–4]. An alternative approach uses so-called protecting Hamiltonians [5,6], which permanently act on the quantum memory and immunize it against small perturbations. Its most attractive feature is that, in contrast to QECC, it does not require any periodic action on the quantum memory, just encoding and decoding operations at the times of storing and retrieving the information. Whereas this approach may tolerate certain types of noise [7,8], it is not clear if it is suitable in the presence of depolarizing noise, something which QECC can deal with.

In this Letter, we give a complete answer to this question. More specifically, we consider the situation where a logical qubit is encoded in a set of N qubits and allowed to evolve in the presence of depolarizing noise and a protecting Hamiltonian. The goal is to find the strategy delivering the longest lifetime, τ, after which we can apply a decoding operation and reliably retrieve the original state of the qubit. By adapting ideas taken from [9], it is established that the lifetime cannot exceed $O(\log N)$. An analysis of the case in which no protecting Hamiltonian is used presents markedly different behavior depending on whether we intend to store classical or quantum information. Finally, we construct a static protecting Hamiltonian that saturates the upper bound $\tau \sim O(\log N)$. To this end, we first show how to achieve this bound using a time-dependent Hamiltonian protection which emulates QECC. We then introduce a clock gadget which exploits the noise to measure time—similar to radiocarbon dating—thus allowing us to simulate the previous time-dependent protection without explicit reference to time.

We consider a system of N qubits, each of which is independently subject to depolarizing noise at a rate r. The total state evolves as

$$\dot{\rho}(t) = -i[H(t), \rho(t)] - r\left[N\rho(t) - \sum_{n=1}^{N} \mathrm{tr}_n(\rho(t)) \otimes \frac{\mathbb{1}_n}{2}\right]. \tag{1}$$

Furthermore, we shall allow for an arbitrary encoding of the initial state as well as a final decoding procedure to recover the information.

Protection limitations.—Using purely Hamiltonian protection, a survival time of $\tau \sim O(\log N)$ is the maximum achievable. Intuitively, this is due to the fact that the depolarizing noise adds entropy to the system at a constant rate, while any reversible operation (i.e., Hamiltonian or unitary evolution) will never be able to remove this entropy from the system. Rather, in the best case, it can concentrate all the entropy in a subsystem, keeping the remaining part as clean as possible. This entropic argument was first presented in [9], where the authors investigated the power of reversible computation (both classical and quantum) subject to noise in the absence of fresh ancillas. To this end, they considered the information content $I(\rho) = N - S(\rho)$ of the system, with N the number of qubits and $S(\rho) = -\mathrm{tr}(\rho \log_2 \rho)$. The information content upper bounds the number of classical bits extractable from ρ, and thus ultimately also the number of qubits stored in ρ. While the original statement about the decrease of $I(\rho)$ is for discrete-time evolution, it can be straightforwardly generalized to the continuous time setting of Eq. (1), where it states that

$$\frac{dI(\rho)}{dt} \leq -rI(\rho),$$

which yields that the information of the system is smaller than $\frac{1}{2}$ after a time $\frac{\ln(2N)}{r}$.

Having established an upper bound for the scaling of τ with N, let us analyze whether this bound can be reached under different circumstances. We start with the simplest case where we use no Hamiltonian protection (i.e., $H = 0$) and show that τ is independent of N; that is, no quantum memory effect can be achieved. For that, we note that the effect of Eq. (1) on each qubit may be expressed in terms of a depolarizing channel

$$\mathcal{E}_t(\rho) = \lambda(t)\rho + [1 - \lambda(t)]\frac{\mathbb{1}}{2}$$

where $\lambda(t) = e^{-rt}$. For $t \geq t_{\rm cl}$, where $\lambda(t_{\rm cl}) = \frac{1}{3}$, the resulting channel is entanglement breaking [10]. This remains true if one incorporates encoding and decoding steps. It is simple to prove that for entanglement breaking qubit channels, the average fidelity [11] is upper-bounded by $2/3$. Thus, the lifetime τ is smaller than $t_{\rm cl} = \frac{\ln 3}{r}$, which is independent of N.

The previous argument does not apply to classical information, for which an optimal storage time that is logarithmic in N may be achieved. The classical version of Eq. (1), taking $H(t) \equiv 0$, is a system of N classical bits subject to bit flipping noise (each bit is flipped at a rate $r/2$). In this case, encoding in a repetition code, and decoding via majority voting, yields an asymptotically optimal information survival time $O(\log N)$. Using optimal estimation [12] and this classical protocol in the encoding phase, the bound $2/3$ may be asymptotically reached for the quantum case.

Time-dependent protection.—We will now use the ideas of QECC to build a simple circuit based model that reaches the upper bound on the protection time. This model assumes that unitary operations can be performed instantaneously, which is equivalent to having a time-dependent protecting Hamiltonian with unbounded strength; we will show how to remove both requirements later on. Instead of using a repetition code, we encode the qubit to be protected in an l level concatenated QECC [2–4] (i.e., l levels of the QECC nested into each other), which requires $N = d^l$ qubits, where d is the number of qubits used by the code. Each level of the QECC can provide protection for a constant time $t_{\rm prot} < t_{\rm cl}$, and thus, after $t_{\rm prot}$, one layer of decoding needs to be executed. Each decoding consists of a unitary $U_{\rm dec}$ on each d-tuple of qubits in the current encoding level; after the decoding, only one of each of the d qubits is used further (Fig. 1). The total time that such a concatenated QECC can protect a qubit is given by $t_{\rm prot} l = t_{\rm prot}\log_d N \sim O(\log N)$, as in the classical case.

Time-independent protection.—In the following, we show that the same $O(\log N)$ protection time which we can achieve using a time-dependent protection circuit can also be obtained from a time-independent protecting

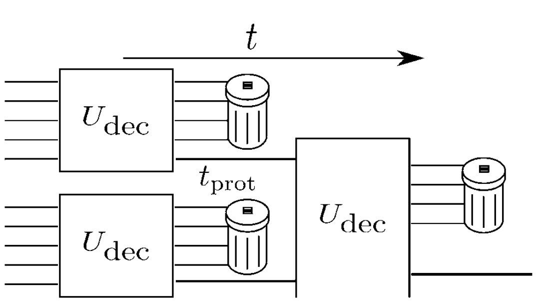

FIG. 1. Decoding a nested QECC. The "discarded" qubits carry most of the entropy and are not used further.

Hamiltonian. The basic idea of our construction is to implement the time-dependent Hamiltonian presented before in a time-independent way. To this end, a clock is built which serves as control. The time-independent version performs the decoding gates conditioned on the time estimate provided by the clock. In order to obtain a clock from (1) with a time-independent H, we will make use of the noise acting on the system: we add a number, K, of "clock qubits" which we initialize to $|1\rangle^{\otimes K}$ and let the depolarizing noise act on them. The behavior of the clock qubits is thus purely classical; they act as K classical bits initialized to 1 which are being flipped at a rate $r/2$. Thus, the polarization k, defined by the number of "1" bits minus the number of "0" bits has an average expected value of $\bar{k}(t) = Ke^{-rt}$ at time t. Conversely, this provides the time estimate

$$\tilde{t}(k) = \min\left(\frac{\ln(K/k)}{r}, t_{\rm max}\right). \tag{2}$$

Here, $t_{\rm max}$ is the maximum time for which we expect that estimate to be reliable, which depends on K and the precision of the estimate cf. (5) below. Particular realizations of this random process of bit flips can be described by a polarization trajectory $k(t)$. Good trajectories are defined to be those such that

$$|k(t) - \bar{k}(t)| < K^{1/2+\epsilon} \tag{3}$$

for all $0 \leq t \leq t_{\rm max}$. For appropriate parameters $t_{\rm max}$ and $0 < \epsilon < \frac{1}{2}$, the following theorem states that almost all trajectories are good and can provide accurate time estimates.

Theorem (Depolarizing clock).—For $K \geq 16$, good trajectories have a probability

$$P[k(t)\,\text{good traj.}] \geq 1 - K\frac{rt_{\rm max} + \exp[-3K^{2\epsilon}/8]}{\exp[K^{2\epsilon}/8]}. \tag{4}$$

Furthermore, for any good trajectory $k(t)$, the time estimate $\tilde{t}$ returned by the clock will differ from the real time t by at most

$$\frac{\delta}{2} := \frac{1}{rK^{1/2-\epsilon}}e^{rt_{\rm max}} \geq |\tilde{t}[k(t)] - t|. \tag{5}$$

(For fixed δ, this implies that $t_{\rm max}$ will scale logarithmically with the number of qubits.)

Note that the theorem does not simply state that any time evolution will be outside (3) for an exponentially small amount of time (which is easier to prove), but that there is only an exponentially small number of cases in which (3) is violated at all. Although the former statement would in principle suffice to use the clock in our construction, the stronger version of the theorem makes the application of the clock, and, in particular, the error analysis, more transparent and will hopefully lead to further applications of the clock gadget.

Proof.—To prove the theorem, note that each of the bits undergoes an independent exponential decay so that the total polarization is the sum of K identical independent random variables. We can thus use Hoeffding's inequality

[13] to bound the probability of finding a polarization far from the expected average value $\bar{k}(t)$,

$$\Pr[|k(t) - \bar{k}(t)| \geq K^{1/2+\epsilon}] \leq 2e^{-(K^{2\epsilon}/2)}. \qquad (6)$$

This already implies that most of the trajectories violate (3) for no more than an exponentially small amount of time. To see why (6) implies that most trajectories are good trajectories, we bound the average number of times a trajectory leaves the region (3) of good trajectories. Since a nongood trajectory must leave (3) at least once, it is also an upper bound on the probability of nongood trajectories. Hence, it suffices to consider the average rate $R(t)$ at which processes leave (3), and integrate over t to obtain a bound on the probability of trajectories which are not good.

The rate at which a process leaves the set of good trajectories has two sources, as illustrated in Fig. 2: First, the system can undergo a spin flip, thus leaving the region defined by (3) vertically (rate R_v), and second, it can leave it horizontally if the time t passes the maximum time allowed by (3) for the current value $k(t)$ of the polarization (rate R_h). A vertical leave can occur only if $|k(t) - \bar{k}(t)| \geq K^{1/2+\epsilon} - 2 \geq K^{1/2+\epsilon}/2$, provided $K \geq 16$ (a spin flip changes $k(t)$ by ± 2). Equation (6) thus gives an average bound

$$R_v(t) \leq Kre^{-K^{2\epsilon}/8}.$$

A horizontal leave can only occur at discrete times extremizing (3),

$$t \in \mathcal{T} = \{t: \bar{k}(t) + K^{1/2+\epsilon} \in \mathbb{N}\},$$

and the probability of a trajectory fulfilling $k(t) = \bar{k}(t) + K^{1/2+\epsilon}$ may again be bounded using (6), such that

$$R_h(t) \leq 2e^{-K^{2\epsilon}/2} \sum_{\tau \in \mathcal{T}} \delta(t - \tau).$$

The inequality (4) follows immediately by integrating $R_h(t) + R_v(t)$ from 0 to $t_{\max}$.

Assuming that $k(t)$ corresponds to a good trajectory, the accuracy of the time estimate (2) may be bounded by

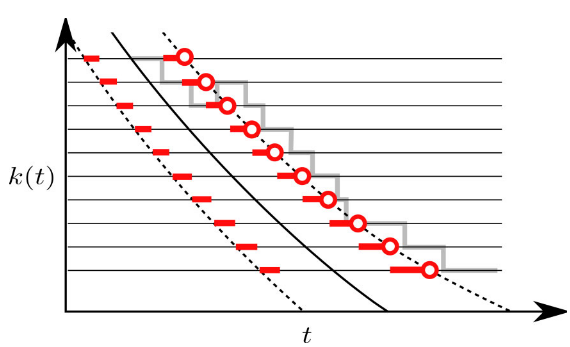

FIG. 2 (color online). A steplike trajectory in light gray illustrates the two ways of leaving region (3) of good trajectories (dashed lines): either a spin flip can take the polarization out of the thickly marked regions, or polarization may leave region (3) as time passes without a spin flip (hollow dots).

applying the mean value theorem to $\bar{k}$:

$$|\tilde{t}[k(t)] - t| = \frac{|\bar{k}\{\tilde{t}[k(t)]\} - \bar{k}(t)|}{|\bar{k}'(t_{\mathrm{interm}})|} \leq \frac{K^{\epsilon}}{r\sqrt{K}} e^{rt_{\max}}.$$

□

Clock dependent Hamiltonian.—Let us now show how the decoding circuit can be implemented using the clock gadget. The circuit under consideration consists of the decoding unitaries $U_{\mathrm{dec}}^{l,k}$ (decoding the k'th encoded qubit in level l, acting on d qubits each); after a time interval t_{prot} (the time one level of the code can protect the qubit sufficiently well), we perform all unitaries $U_{\mathrm{dec}}^{l,k}$ at the current level l—note that they act on distinct qubits and thus commute. Each of these unitaries can be realized by applying a d-qubit Hamiltonian $H_{\mathrm{dec}}^{l,k}$ for a time $t = t_{\mathrm{dec}}$. Thus, we have to switch on all the $H_{\mathrm{dec}}^{l,\cdot}$ for $t \in [t_l, t_l + t_{\mathrm{dec}}]$, where $t_l = lt_{\mathrm{prot}} + (l-1)t_{\mathrm{dec}}$.

In order to control the Hamiltonian from the noisy clock, we define clock times $k_{l,\mathrm{on}} = \lfloor \bar{k}(t_l) \rfloor$ and $k_{l,\mathrm{off}} = \lceil \bar{k}(t_l + t_{\mathrm{dec}}) \rceil$, and introduce a time-independent Hamiltonian which turns on the decoding Hamiltonian for level l between $k \in [k_{l,\mathrm{on}}, k_{l,\mathrm{off}}]$,

$$H = \sum_l (H_{\mathrm{dec}}^{l,1} + \cdots + H_{\mathrm{dec}}^{l,d^{L-l}}) \otimes \Pi_l. \qquad (7)$$

The left part of the tensor product acts on the N code qubits, the right part on the K clock (qu)bits, and

$$\Pi_l = \sum_{k=k_{l,\mathrm{on}}}^{k_{l,\mathrm{off}}} \sum_{w_x=(k+N)/2} |x\rangle\langle x|,$$

where x is an N-bit string with Hamming weight w_x. The initial state of the system is, as for the circuit construction, the product of the encoded qubit in an l-level concatenated code and the maximally polarized state $|1\rangle^{\otimes K}$ on the clock gadget.

Error analysis.—We now perform the error analysis for the protecting Hamiltonian (7). In order to protect the quantum information, we will require that the error probability per qubit in use is bounded by the same threshold p^* after each decoding step is completed (i.e., at $t = t_l + t_{\mathrm{dec}} + \frac{\delta}{2}$). We will restrict to the space of good trajectories, since we know from the clock theorem that this accounts for all but an exponentially small fraction, which can be incorporated into the final error probability.

We will choose K large enough to ensure that the error $\frac{\delta}{2} \geq |\tilde{t} - t|$ in the clock time satisfies $\delta \ll t_{\mathrm{prot}}, t_{\mathrm{dec}}$. In this way, we ensure that the decoding operations are performed in the right order [14] and with sufficient precision. We may thus account for the following error sources between $t_l + t_{\mathrm{dec}} + \delta/2$ and $t_{l+1} + t_{\mathrm{dec}} + \delta/2$: (i) Inherited errors from the previous rounds which could not be corrected for. By assumption, these errors are bounded by $p_{\mathrm{inher}} \leq p^*$. (ii) Errors from the depolarizing noise during the free evolution of the system. The system is sure to evolve freely for a time $t_{\text{prot}} - \delta$, i.e., the noise per qubit is bounded by $p_{\text{evol}} \leq 1 - \exp[-r(t_{\text{prot}} - \delta)] \leq r(t_{\text{prot}} - \delta)$. (iii) Errors during the decoding. These errors affect the decoded rather than the encoded system and stem from two sources: On the one hand, the time the Hamiltonian is active has an uncertainty $t_{\text{dec}} \pm \delta$, which gives an error in the implemented unitary of not more than $\exp[\delta \| H_{\text{dec}}^{k,l} \|] - 1$. On the other hand, depolarizing noise can act during the decoding for at most a time $t_{\text{dec}} + \delta$. In the worst case, noise on any of the code qubits during decoding will destroy the decoded qubit, giving an error bound $d\{1 - \exp[-r(t_{\text{dec}} + \delta)]\} \leq dr(t_{\text{dec}} + \delta)$. Thus, the error on the decoded qubit is

$$p_{\text{dec}} \leq \exp[\| H_{\text{dec}}^{k,l} \| \delta] - 1 + dr(t_{\text{dec}} + \delta).$$

Since the noise is Markovian (i.e., memoryless), the clock does not correlate its errors in time. In summary, the error after one round of decoding is at most $B(p_{\text{inher}} + p_{\text{evol}}) + p_{\text{dec}}$, which we require to be bounded by p^* again. Here, $B(p)$ is a property of the code, and returns the error probability of the decoded qubit, given a probability p of error on each of the original qubits; for example, for the 5-qubit perfect QECC [15], $B(p) \leq 10p^2$.

We will now show that it is possible to fulfill the required conditions by appropriately defining the control parameters. First, we choose $p^* \leq 1/40$ to have the QECC [15] work well below threshold. We may take $t_{\text{prot}} := \frac{p^*}{r}$ and $t_{\text{dec}} := \frac{p^*}{4dr}$. To minimize imprecision in the implemented unitaries, the decoding Hamiltonians are chosen to be of the minimal possible strength, $\| H_{\text{dec}}^{k,l} \| \leq \frac{2\pi}{t_{\text{dec}}}$. Finally, we take $\delta := \frac{p^* t_{\text{dec}}}{8\pi}$. Inserting the proposed values in the derived bounds, it is straightforward to show that $B(p_{\text{inher}} + p_{\text{evol}}) + p_{\text{dec}} < p^*$.

The number of code qubits required is $N := d^l$, with $l := \lceil \frac{\tau}{t_{\text{prot}} + t_{\text{dec}}} \rceil$. The required logarithmic clock lifetime $t_{\max} = \tau$ and the precision δ are obtained by taking $\epsilon = 1/6$ and $K := (\frac{2e^{r\tau}}{r\delta})^3$, by virtue of Eq. (5) of the clock theorem. For any fixed r and p^*, this allows a lifetime $\tau \sim O[\log(N + K)]$.

Conclusions.—In this Letter, we have considered the ability of a Hamiltonian to protect quantum information from decoherence. While without a Hamiltonian, quantum information is destroyed in constant time, the presence of time-dependent control engenders protection for logarithmic time, which is optimal. As we have shown, the same level of protection can be attained with a time-independent Hamiltonian. The construction introduced a noise-driven clock which allows a time-dependent Hamiltonian to be implemented without explicit reference to time.

Since depolarizing noise is a limiting case of local noise models, it is expected that the time-independent Hamiltonian developed here can be tuned to give the same degree of protection against weaker local noise models, although these models may admit superior strategies. For instance, noise of certain forms (such as dephasing) allows for storage of ancillas, potentially yielding a linear survival time by error correcting without decoding. In the case of amplitude damping noise, the noise itself distills ancillas so that the circuit can implement a full fault-tolerant scheme, which gives an exponential survival time, assuming that one can redesign the clock gadget to also benefit from these properties.

Whether the same degree of protection can be obtained from a Hamiltonian which is local on a 2D or 3D lattice geometry remains an open question [16]. However, intuition suggests this might be impossible; the crucial point in reversibly protecting quantum information from depolarizing noise is to concentrate the entropy in one part of the system. Since the speed of information (and thus entropy) transport is constant due to the Lieb-Robinson bound [17], the rate at which entropy can be removed from a given volume is proportional to its surface area, while the entropy increase goes as the volume. It thus seems impossible to remove the entropy sufficiently quickly, although this argument is not fully rigorous, and the question warrants further investigation.

We acknowledge the DFG for support (projects MAP and FOR635), and the EU (QUEVADIS and SCALA). A. S. K. is also supported by Clare College, Cambridge.

[1] P. Shor, in *Proceedings of the 37th Annual Symposium on Foundations of Computer Science, 1996* (IEEE Computer Society Press, Burlington, Vermont, 1996), p. 56.
[2] D. Aharonov and M. Ben-Or, arXiv:quant-ph/9611025.
[3] D. Gottesman, Phys. Rev. A **57**, 127 (1998).
[4] D. Gottesman, arXiv:0904.2557.
[5] A. Y. Kitaev, Ann. Phys. (N.Y.) **303**, 2 (2003).
[6] D. Bacon, Phys. Rev. A **73**, 012340 (2006).
[7] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, J. Math. Phys. (N.Y.) **43**, 4452 (2002).
[8] R. Alicki, M. Horodecki, P. Horodecki, and R. Horodecki, arXiv:0811.0033.
[9] D. Aharonov, M. Ben-Or, R. Impagliazzo, and N. Nisan, arXiv:quant-ph/9611028.
[10] M. Horodecki, P. W. Shor, and M. B. Ruskai, Rev. Math. Phys. **15**, 629 (2003).
[11] M. A. Nielsen, Phys. Lett. A **303**, 249 (2002).
[12] S. Massar and S. Popescu, Phys. Rev. Lett. **74**, 1259 (1995).
[13] W. Hoeffding, J. Am. Stat. Assoc. **58**, 13 (1963).
[14] The noisy clock has the potential to run backwards in time within its accuracy.
[15] R. Laflamme, C. Miquel, J. P. Paz, and W. H. Zurek, Phys. Rev. Lett. **77**, 198 (1996).
[16] A first step is to incorporate the notion of boundedness. By controlling each decoding unitary in a given round from a different clock (which does not affect the scaling properties), a constant bound to the sum of Hamiltonian terms acting on any given finite subsystem can be shown.
[17] E. H. Lieb and D. W. Robinson, Commun. Math. Phys. **28**, 251 (1972).