

Ion Trap Quantum Gates, Decoherence and Error Correction

P.L. Knight, M. Muraao, M.B. Plenio, and V. Vedral

Optics Section, The Blackett Laboratory, Imperial College, London, SW7 2BZ, UK

(September 7, 1998)

Abstract

We analyze the linear ion trap proposal for realizing a quantum computer. We calculate the limits on Shor's factorization algorithm due to spontaneous emission of the ions. A further contribution to decoherence which originates from instrumental fluctuations and imperfections is described. Fault-tolerant error correction is then applied to derive limits on the efficiency of ion trap computing.

I. INTRODUCTION

Since Shor's discovery [1,2] of an algorithm that allows the factorization of a large number by a quantum computer in polynomial time instead of an exponential time as in classical computing, interest in the practical realization of a quantum computer has greatly increased [3]. Recent advances in the preparation and manipulation of single ions, the use of coherent pulsed excitation in NMR, as well as the engineering of pre-selected cavity light fields have made quantum optics that field of physics which promises the first experimental realization of a quantum computer. Several proposals (reviewed in [4]) for experimental implementation have been made which rely on nuclear spins, quantum dots [5], cavity QED [6] and on ions in linear traps [7].

The realization of a quantum computer in a linear trap in particular was regarded as very promising, as it was thought that decoherence could be suppressed sufficiently to preserve the

large-scale superpositions necessary for significant quantum computation. Indeed, a single quantum gate in such an ion trap was realized by Monroe et al [8] and recently two ions were cooled down to the ground state of the motional degrees of freedom [9]. Nevertheless, the decoherence rate in these experiments would seem too high to allow the realization of extended quantum networks. This experiment was not solely limited by fundamental processes but rather by technical difficulties and one aim of future experiments is to reduce these technical problems to come closer to the fundamental limits, such that at least small networks could be realized.

However, there remains the question whether overcoming technical problems will be sufficient to realize practically useful computations such as factorization of large numbers on a quantum computer in a linear ion trap. To this end, it is necessary to investigate more thoroughly those limitations that are due to processes of more a fundamental nature than for example laser linewidths, pulse lengths etc.. An analysis of this problem was carried out in ref. [10] where the impact of spontaneous emission on the problem of factorization of a large number by a quantum computer was considered. It was shown that for realistic laser powers, factorization is limited to small numbers (at least if no sophisticated error correction methods are implemented). The limits that were obtained still depended there on the power of the applied lasers and left open the possibility of improving the maximal bitsize L_{max} of the largest factorizable number by using higher power lasers. Extremely high intensities, however, will lead to ionization or, more importantly, to a breakdown of the two-level-approximation [11,12]. Here we discuss intensity independent limits that can be derived, taking into account the breakdown of the two-level-approximation. We assume that the qubits are stored in the Zeeman sublevels of an ion which are then manipulated using detuned Raman transitions and discuss the limits on L_{max} on the basis of experimental parameters for several ions. We conclude that factorization without the use of efficient error correction methods is limited to almost trivial numbers if spontaneous emission is present. One can extend the analysis to take into account quantum error correction. Quantum error correction relies on the assumption that amplitude or phase errors retain population within

the two-dimensional space of the qubit. Here, however, we are concerned with 'leakage' errors out of the qubit space. However, the breakdown of the two-level approximation seriously affects the effectiveness of the quantum error correction. This shows that with present technology large scale factorization, or other demanding applications, seem out of reach.

In the second part we address the problem of so called accuracy thresholds in quantum computation [13,14]. In [13,14] it is noted that arbitrarily complicated (long) quantum computations can be performed, once the error rate of a quantum gate can be pushed below a certain threshold. We will discuss whether the required thresholds presented in [13,14] can be achieved or if spontaneous emission rules out this possibility (not to mention other error sources). We present a simple calculation to understand the order of magnitude of these thresholds and then calculate the spontaneous emission rate in a single quantum gate.

II. DECOHERENCE LIMITS TO QUANTUM FACTORING

A. The linear ion-trap model of quantum computation

Before we come to the derivation of limits imposed on quantum computation by spontaneous emission, we briefly describe the ion trap implementation of the quantum computer [7] on which we base our considerations. Several ions of mass M are trapped in a linear ion trap such that they are lined up and are well separated, ie the next neighbour distance is many wavelengths of the lasers used to manipulate the ions. This is necessary to be able to address ions separately and can lead to further restrictions on quantum computation in this model [15,16]. The motional degrees of freedom of the ions, and especially their collective center-of-mass (COM) motion with frequency ν , are cooled to the ground state. In order to be able to implement two-bit gates in this scheme we use the COM mode as a bus which allows us to create entanglement between different ions. This is achieved with an interaction that creates a phonon in the COM mode when we deexcite an ion, and annihilates a phonon

in the COM mode when an ion is excited. This interaction is generated by a laser which is detuned from the transition frequency by $\Delta = -\nu$. In that case the Hamiltonian in the Lamb-Dicke limit (i.e. tight confinement of the laser-cooled ion) , the RWA and a suitable interaction picture is given by [7]

$$H = \frac{\eta}{\sqrt{5L}} \frac{\Omega_{01}}{2} \left[|1\rangle\langle 0|a + |0\rangle\langle 1|a^\dagger \right] , \quad (1)$$

where $\eta = \frac{2\pi}{\lambda} \sqrt{\frac{\hbar}{2M\nu}}$ is the Lamb-Dicke parameter, Ω_{01} is the Rabi frequency on the $0 \leftrightarrow 1$ transition, where levels 0 and 1 represent the corresponding logical values of the qubit. The denominator $\sqrt{5L}$ originates from the fact that we are considering the COM mode, which has an effective mass $5LM$ because all $5L$ ions are oscillating in the trap potential. The trap has to contain $5L$ ions, as this is the number of ions that is required to implement Shor's algorithm to factorize an L bit number [1,17] (actually the correct number is $5L + 2$ but for simplicity we drop the 2). It is further assumed that coupling to other vibrational levels can be neglected. This assumption is a reasonable first approximation, as the closest lying vibrational mode has frequency $\sqrt{3}\nu$ independent of the number of ions in the trap. Nevertheless, this approximation breaks down when the Rabi frequency becomes too large, ie. if $\Omega_{01} \approx \nu$ [15,16]. For our considerations of the influence of spontaneous decay we neglect this effect although it may well become important in longer calculations. Using the Hamiltonian eq. (1) it can be shown that it is possible to construct a CNOT gate with only four π -rotations, and to construct the more involved Toffoli gate [18] with six π -rotations [7]. In the following we will define an elementary time step in terms of the computation time that is required to perform a CNOT gate. All other performance times can be re-expressed in units of that of the CNOT gate. The implementation of Shor's algorithm also requires a number of one bit operations, ie. operations that leave the COM mode unaffected during their performance. However, these operations can be performed much faster than two-bit gates, because $\frac{\eta}{\sqrt{5L}} \ll 1$ and also because fewer individual laser pulses are required. Therefore their contribution to the calculation time and to decoherence is small and will be neglected in the following.

B. Derivation and discussion

One can store qubits in an ion trap in at least two ways. One way is within the two states of a metastable optical transition. This method certainly has the disadvantage that at practically all times the quantum computer will have a mean excitation of about 0.5 per qubit. Therefore, even if a qubit does not take part in a quantum gate operation, the qubit will decohere due to spontaneous emissions from the excited state of the qubit. This leads to a quite strong impact of spontaneous emission on the quantum computation and the metastable two-level system could well be a less favourable option [10–12]. To circumvent this problem, one could store the qubits in Zeeman sublevels of a stable ground state. In that case stored qubits would seem not to suffer decoherence due to spontaneous emission. However, how do we manipulate the qubit? It is (unfortunately) not practical to drive the qubit with a microwave field directly because the very long wavelength of this radiation does not allow us to address individual ions, as they cannot be separated far enough in a linear trap. Therefore we have to use a different method to manipulate the qubit. This method uses the presence of another level 2 which is coupled to the two qubit levels 0 and 1 by two lasers of Rabi frequency Ω_{02} and Ω_{12} . Each laser is strongly detuned from its transition, and its Rabi frequency is much smaller than the detuning. However, it is assumed that both detunings are equal, ie. the two photon detuning vanishes, while the one photon detuning Δ_2 is large. The advantage of this method is that due to the strong detuning Δ_2 the population in level 2 is small, so that spontaneous emissions from that level are rare. A Hamiltonian that implements an interaction where the qubit is excited while a phonon in the COM mode is deexcited and vice versa, is given by

$$H = -\hbar\Delta_2|2\rangle\langle 2| + \frac{\hbar\Omega_{02}}{2}[|2\rangle\langle 0| + |0\rangle\langle 2|] + \frac{\hbar\eta\Omega_{12}}{2\sqrt{5}L}\left[|2\rangle\langle 1|a + |1\rangle\langle 2|a^\dagger\right], \quad (2)$$

where again the Lamb-Dicke limit, the RWA and a suitable interaction picture is used. To make sure that the Raman transitions have the effect as full Rabi oscillations between levels 0 and 1 the condition

$$\Omega_{02} = \frac{\eta\Omega_{12}}{\sqrt{5L}} \quad (3)$$

has to be satisfied which we will assume in the following. The condition eq. (3) can be obtained from the solution of the time evolution

$$|\langle 0|\psi(t)\rangle|^2 = \left| \frac{\Omega_{02}^2 e^{-i(\Omega_{02}^2 + \eta^2\Omega_{12}^2/5L)t/4\Delta_2} + \eta^2\Omega_{12}^2/5L}{\Omega_{02}^2 + \Omega_{12}^2} \right|^2 \quad (4)$$

$$|\langle 1|\psi(t)\rangle|^2 = 1 - |\langle 0|\psi(t)\rangle|^2 \quad (5)$$

$$|\langle 2|\psi(t)\rangle|^2 = \frac{\Omega_{02}^2}{2\Delta_2^2} (1 + \cos \Delta_2 t) \quad (6)$$

which is valid for $\Gamma_{22}t \ll 1$ and for an initial state $|\psi(t)\rangle = |0\rangle$. To obtain $|\langle 0|\psi(t)\rangle|^2 = 0$ for some t eq. (3) has to be satisfied. In the Raman pulse implementation of qubits a CNOT gate can be implemented with four π pulses. One easily obtains the effective Rabi frequency of the qubit transition which is given by

$$\Omega_{eff} = \frac{\Omega_{02}^2}{2\Delta_2} \quad (7)$$

When transitions between level 0 and 2, and between level 1 and 2, are imperfect due to fluctuation of the driving lasers' intensity, or if the trap potential fluctuates, these fluctuations would lead decoherence which gives an exponential decay of Rabi oscillation observed in the experiment [19]. Here we describe briefly the decoherence of the ion trap system which originates by instrumental imperfections, using a master equation and the coupling of the system weakly to a quantum bosonic reservoir at temperature T [20]. The system-reservoir coupling H_{sr} describing an imperfect transition is given by

$$H_{sr} = \hbar \left(|2\rangle \langle 1| a + |1\rangle \langle 2| a^\dagger \right) \sum_l g'_l (B_l^\dagger + B_l), \quad (8)$$

and H'_{sr} describes fluctuations of the trap potential and is given by

$$H'_{sr} = \hbar a^\dagger a \sum_l g'_l (B_l^\dagger + B_l), \quad (9)$$

where ω_l is the l th reservoir frequency, and $B_l^\dagger$ and B_l are the creation and annihilation operators of the reservoir bosons.

Then the master equation for the reduced system operator in the interaction picture $\rho^I(t)$ due to the system-reservoir coupling is obtained using a “time convolution-less formalism” [21] and the rotating wave approximation on the master equation

$$\dot{\rho}^I(t) = \frac{1}{i\hbar} [H_e^I, \rho^I(t)] + \Gamma \rho^I(t) \quad (10)$$

with the damping term $\Gamma \rho^I(t)$ given by [20]

$$\begin{aligned} \Gamma \rho^I(t) = & \sum_l g_l^2 \int_0^t dt' \left\{ \left(\langle B_l^\dagger(t') B_l \rangle_B + \langle B_l(t') B_l^\dagger \rangle_B \right) \right. \\ & \times \left([C_s(-t') \rho^I(t), C_s^\dagger] + [C_s^\dagger(-t') \rho^I(t), C_s] \right) \\ & + \left(\langle B_l^\dagger(-t') B_l \rangle_B + \langle B_l(-t') B_l^\dagger \rangle_B \right) \\ & \left. \times \left([C_s, \rho^I(t) C_s^\dagger(-t')] + [C_s^\dagger, \rho^I(t) C_s(-t')] \right) \right\} \end{aligned} \quad (11)$$

where C_s represents the system operators (either $|1\rangle\langle 2|a^\dagger$ or $a^\dagger a$ for the imperfect transition case and the trap potential fluctuation cases), which couple to the reservoir. The time evolution of the system operators are determined by the effective Jaynes-Cummings interaction H_e , which is

$$H_e = \Omega_{eff} \left(|1\rangle\langle 0|a + |0\rangle\langle 1|a^\dagger \right). \quad (12)$$

The master equation (10) can be solved by expanding all system operators in terms of the dressed states, which are eigenstates of the Jaynes-Cummings interaction Hamiltonian under certain conditions. We take the continuum limit of the reservoir modes and assume the system-reservoir coupling is weak. We also require the time scale of the reservoir variables to be much shorter than the system variables so we can take the Markovian limit. If we assume an initial condition of a product state $|1\rangle_a\langle 1| \otimes \sum_n p_n |n\rangle_m\langle n|$, the population of the level 1 is given by

$$P_1(t) = \frac{1}{2} \left\{ 1 + \sum_n p_n \cos(B_n t) e^{-A_n t} \right\}. \quad (13)$$

The damping rate A_n is

$$A_n = (n+1) \kappa(n) \{ \hat{n}(n) + 1/2 \} \equiv A_n^{in}, \quad (14)$$

or

$$A_n = \frac{1}{2} \kappa(n) \{ \hat{n}(n) + 1/2 \} \equiv A_n^{vi}, \quad (15)$$

for the imperfect laser transition case (14), and for the fluctuation of the vibrational potential case (15), where $\bar{n}(n)$ is the mean reservoir boson number given by

$$\hat{n}(n) = \left(e^{2\hbar\Omega_{eff}\sqrt{n+1}/k_B T} - 1 \right)^{-1}, \quad (16)$$

and $\kappa(n)$ is assumed to be given by

$$\kappa(n) \approx \left(2\hbar\Omega_{eff}\sqrt{n+1} \right)^d. \quad (17)$$

The coherent part B_n is given by

$$B_n = \sqrt{4\Omega_{eff}^2(n+1) - A_n^2}. \quad (18)$$

The results for the decoherence rates A_n^{in} (14) and A_n^{vi} (15) show that decoherence originates in the coupling to reservoir bosons at frequency of $2\Omega_{eff}\sqrt{n+1}$. The effective Rabi frequency $\Omega_{eff}/2\pi$ in the experiment [19] is around 100 kHz, so reservoir bosons with frequency 100 kHz may be responsible for decoherence. These reservoir bosons have much lower frequency than those responsible for spontaneous emission of atomic states, which is of order of GHz, and also population decay of motional states, which is of order 10 MHz. This low frequency nature of the reservoir boson suggests that the reservoir can be regarded as having a *high temperature* nature whereas in the optical frequency regime, the corresponding reservoir is often approximated at zero temperature. This high temperature limit represents classical noise where the reservoir operators commute. Introducing normalised values, $\tilde{A}_n^{in} = A_n^{in}/\Omega_{eff}$, $\tilde{A}_n^{vi} = A_n^{vi}/\Omega_{eff}$, $\tilde{\gamma}_0 = \gamma_0/\Omega_{eff}$, $\tilde{\kappa}(n) = \kappa(n)/\Omega_{eff}$, the normalised decoherence rates are

$$\tilde{A}_n^{in} = \tilde{A}_0^{in} (n+1)^{(d+1)/2}, \quad (19)$$

$$\tilde{A}_n^{vi} = \tilde{A}_0^{vi} (n+1)^{(d-1)/2}. \quad (20)$$

To get an exponent of 0.7 for the factor $(n + 1)$ as suggested by the experiment [19], we need $d \approx 0.4$ for the imperfect dipole transition case and $d \approx 2.4$ for the case of fluctuations of the vibrational potential. This then replicates the observed decay [19] given by eq. (13) with the decay factors A_n seen in the experiments. Candidates for the origin of the dephasing are discussed in [20] and include laser phase fluctuation in the Raman excitation, as well as fluctuations in the trap potential.

Let us now return to the question of spontaneous emission decoherence. The performance of a CNOT gate requires time for 4π -pulse operations. Thus from the definition of the effective Rabi frequency (eq. 7), we find for the elementary time-step,

$$\tau_{el} = \frac{8\pi\Delta_2}{\Omega_{02}^2} . \quad (21)$$

The probability for a spontaneous emission from level 2 has to be small in the entire network responsible for the factorization (which involves ϵL^3 gate operations, where ϵ is of order 200 – 1000). Thus we require

$$2\Gamma_{22}\rho_{22}\tau_{el}\epsilon L^3 = p_{em}^{(2)} \ll 1 . \quad (22)$$

This condition immediately yields

$$L = \left(\frac{\Delta_2 p_{em}^{(2)}}{8\pi\epsilon\Gamma_{22}} \right)^{1/3} . \quad (23)$$

which is already intensity independent. From eq. (23) one could conclude that one would be able to factorize gigantic numbers by making the ratio Δ_2/Γ_{22} sufficiently large, eg. $\Delta_2 = 10^{13}s^{-1}$, $\Gamma_{22} = 1s^{-1}$ and $\epsilon = 216$ would lead to $L = 1225$. But this reasoning is invalid, because we have neglected the fact that the two-level approximation breaks down because in the ion other levels, which we model by a level 3, also couple to levels 0 and 1 via the Raman pulses. This additional coupling is very important because it potentially involves levels that may have $\Gamma_{33} \cong 10^8 s^{-1}$. In addition the presence of level 3 changes the effective Rabi frequency in the qubit and may therefore lead to additional errors. Finally there is a limit to the size of Δ_2 from the nature of the qubit quantum numbers, as usually states 0

and 1 are hyperfine levels within a single manifold, which require a nuclear spin flip which is precluded for very large Δ_2 .

To achieve a more restrictive, but again intensity independent, upper bound for L than the one given in eq. (23) we now take into account the presence of one more level 3 which acquires population during the execution of the quantum gate. One can give a complete analysis of the full four level system; however, this is very tedious, although not complicated in principle. We assume that the transition to level 2 are quadrupole-allowed while the transition to an extraneous level 3 is dipole allowed. In this article we will restrict ourselves to the limiting case that

$$\frac{\Omega_{03}^2}{\Delta_3} \ll \frac{\Omega_{02}^2}{\Delta_2} . \quad (24)$$

ie. that level 2 determines the dynamics and that the other level 3 contributes a small correction. The Rabi frequency on a transition is not independent of the decay constant on that same transition. In fact we find [10]

$$\frac{\Omega_{01}^2}{\Gamma_{11}} = \frac{6\pi c^3 \epsilon_0}{\hbar \omega_{01}^3} E^2 , \quad (25)$$

which can be derived along similar lines to the usual derivation of the relation between the Einstein A and B coefficient. Inserting this into eq. (24) yields

$$L = \left(\frac{\Delta_2 p_{em}^{(3)}}{8\pi \epsilon \Gamma_{22}} \right)^{1/3} \ll \left(\frac{\Delta_3 p_{em}^{(2)}}{8\pi \epsilon \Gamma_{33}} \right)^{1/3} , \quad (26)$$

where Δ_3 and Γ_{33} is more or less determined because the lasers are extremely far detuned from the $i \leftrightarrow 3$ transitions.

To calculate the influence of level 3 on the decoherence, we first have to find the average population in that level. For this we note that during the implementation of a CNOT a full 4π -rotation will be performed. That means that on average the population in both levels 0 and 1 are equal. Therefore the much slower coupling to level 3 sees an averaged population in both levels with no average coherence. We can therefore average over two contributions to ρ_{33} which originate from the initial states $|0\rangle$ and $|1\rangle$ and obtain

$$\rho_{33} = \frac{1}{2} \frac{\Omega_{13}^2 + \eta'^2 \Omega_{03}^2 / 5L}{4\Delta_3^2} , \quad (27)$$

where η' is the Lamb-Dicke parameter for the $0 \leftrightarrow 3$ transition. Eq. (27) contains an additional factor $1/2$ compared to eq. (6). This is due to the fact that level 3 is strongly decaying and that therefore $\Gamma_{33}t \gg 1$. Note also that the roles of the lasers on the $i \leftrightarrow 3$ transitions has been reversed compared to their role on the $i \leftrightarrow 2$ transitions; the phonon is now destroyed by an excitation from level 0 and not from level 1 anymore. This is due to the fact that $i \leftrightarrow 3$ transitions are dipole allowed (E1) while the $i \leftrightarrow 2$ transition is quadrupole allowed (E2). Again an emission from level 3 should not occur during the whole calculation which means

$$\begin{aligned} p_{em}^{(3)} &= 2\Gamma_{33}\rho_{33}\tau_{el}\epsilon L^3 \\ &= \frac{16\pi^2\epsilon^2\Gamma_{22}\Gamma_{33}(\Omega_{13}^2 + \eta'^2\Omega_{03}^2/5L)}{\Delta_3^2\Omega_{02}^2 p_{em}^{(2)}} L^6 . \end{aligned} \quad (28)$$

Now we have to find out which term in the brackets of eq. (28) is dominant. Because we need $\Omega_{02} = \eta\Omega_{12}/\sqrt{5L}$, see eq. (3), and because we are in the Lamb-Dicke limit, ie. $\eta/\sqrt{5L} \ll 1$, the laser on the $0 \leftrightarrow 2$ transition has to be weaker than that one on the $1 \leftrightarrow 2$ transition. But again $\eta'/\sqrt{5L} \ll 1$ and therefore $\Omega_{13} \ll \eta'\Omega_{03}/\sqrt{5L}$. This implies

$$\begin{aligned} p_{em}^{(3)} &= \frac{16\pi^2\epsilon^2}{\Delta_3^2 p_{em}^{(2)}} \frac{\Gamma_{22}\Gamma_{33}\Omega_{13}^2}{\Omega_{02}^2} L^6 \\ &= \frac{16\pi^2\epsilon^2\Gamma_{22}\Gamma_{33}}{\Delta_3^2 p_{em}^{(2)}} \frac{\Gamma_{22 \rightarrow 00}}{\Omega_{02}^2} \frac{\Omega_{13}^2}{\Gamma_{33 \rightarrow 11}} \frac{\Gamma_{33 \rightarrow 11}}{\Gamma_{22 \rightarrow 00}} L^6 \end{aligned} \quad (29)$$

where we have inserted the decay coefficients $\Gamma_{ii \rightarrow jj}$ from level i to level j to be able to apply eq. (25) which then yields

$$p_{em}^{(3)} = \frac{16\pi^2\epsilon^2\Gamma_{22}\Gamma_{33}}{\Delta_3^2 p_{em}^{(2)}} \left(\frac{E_{12}}{E_{02}} \right)^2 \left(\frac{\omega_{02}}{\omega_{31}} \right)^3 \frac{\Gamma_{33 \rightarrow 11}}{\Gamma_{22 \rightarrow 00}} L^6 , \quad (30)$$

where E_{i2} is the electric field of the laser on the $i \leftrightarrow 2$ transition. We define the constant

$$\alpha = \frac{\Gamma_{33}}{\Gamma_{22}} \frac{\Gamma_{22 \rightarrow 00}}{\Gamma_{33 \rightarrow 11}} \left(\frac{E_{02}}{E_{12}} \right)^2 \quad (31)$$

where the $\Gamma_{ii \rightarrow jj}$ are the decay coefficients from level i to level j and E_{i2} are the laser field strengths, and obtain

$$L = \left\{ \frac{\Delta_3^2 p_{em}^{(2)} p_{em}^{(3)} \alpha}{16\pi^2 \epsilon^2 \Gamma_{33}^2} \left(\frac{\omega_{13}}{\omega_{02}} \right)^3 \right\}^{1/6}. \quad (32)$$

where the $p_{em}^{(i)}$ are the probabilities to suffer a spontaneous emission from level i . Note that $E_{02} \ll E_{12}$ as we are in the Lamb-Dicke regime and the phonon is excited on the $1 \leftrightarrow 2$ transition. The ratio of the decay coefficients in α will be of the order of unity. Therefore α is much smaller than 1. In fact, because of $\Omega_{02} = \eta \Omega_{12} / \sqrt{5L}$, it is a good approximation to assume $\alpha = \beta \eta^2 / 5L$ which then yields

$$L = \left\{ \frac{\Delta_3^2 p_{em}^{(2)} p_{em}^{(3)} \eta^2 \beta}{80\pi^2 \epsilon^2 \Gamma_{33}^2} \left(\frac{\omega_{13}}{\omega_{02}} \right)^3 \right\}^{1/7}. \quad (33)$$

In table 1 we present some values for L calculated from eq. (33) for three candidate ions Ca^+ , Hg^+ , Ba^+ . We assume that $p_{em}^{(2)} = p_{em}^{(3)} = 1$ (an optimistic choice) and observe that the resulting limits for the numbers that can be factorized on a quantum computer are very small, especially if $\eta = 0.01$. Nevertheless one may think that transitions in Ba^+ are promising; however, this optimism is somewhat misplaced because a very stable transition is also very slow. We can calculate the factorization time using Ba^+ and assuming $\eta = 1$ and find

$$T = 0.013L^3 \quad (34)$$

which gives for $L = 10$ a value of $T = 13s$, during which no decohering event of any kind is allowed to happen [11]. It will be difficult to isolate the quantum computer from the environment for such a long time. This shows very clearly that the solution to the problem of decoherence cannot be to employ extremely stable transitions for qubits but to find methods that enable the quantum computer to cope with a certain level of decoherence, ie we need a form of quantum error correction. In principle it is indeed possible to implement quantum error correction in quantum computers but it remains open whether these methods will in fact improve the prospects of quantum computation. In the next section we will address this problem and derive limitations to quantum factoring including the use of quantum error correction methods. An analysis in the same spirit can be carried out if we include quantum

error correction into our considerations. Before we start with the discussion of the two-level system case, we first define two constants, q and c , describing the necessary overheads in the number of qubits and the number of additional quantum gate operations.

- The number of required qubits rises by a factor of q as all qubits will have to be encoded and will be kept encoded during the calculation. Repeated decoding and encoding would raise the number of operations substantially and lead to unnecessary errors.
- The number of operations required to implement one quantum logical operation on encoded qubits rises by a factor of c because we have to perform more than one quantum gate to implement one quantum logic operation on the encoded qubits. In addition we also have to perform quantum error corrections periodically.

The value of the factors q and c depends strongly on the quantum error correction code that is used. If one wants to protect a single qubit against one general error, then it is known that $q = 5$ is the minimal possible value [13]. The number c for this code is uncertain in general for the most optimal codes but is at least $c = 5$ for a CNOT gate [22] and probably much more for other gates such as a Toffoli gate. One then obtains for L , given leakage errors out of the qubit which conventional error correction techniques fail to handle,

$$L = \left\{ \frac{\beta \eta^2 p_{em}^{(3)} \Delta_3^2}{80 \pi^2 c^2 q \epsilon \Gamma_{33}^{out} \Gamma_{33 \rightarrow 11}} \left(\frac{\omega_{13}}{\omega_{02}} \right)^3 \left(\frac{p_{fail}}{\epsilon} \right)^{\frac{1}{2}} \right\}^{\frac{2}{11}}. \quad (35)$$

Here Γ_{33}^{out} is the decay coefficient for a decay that leads out of the qubit levels, and p_{fail} is the probability that the whole computation fails due to second order errors. In table 2.2 we give values for the same ions as in table 1. Although the values in table 2.2 look quite promising, it should be noted that the computation time for large values of L is extremely long and can reach values of the order of days. During all of this time the computer would have to be isolated from the environment and all other sources of errors would have to be negligible during that time.

The results of this section have been derived under the assumption that leakage errors, i.e. transitions to levels other than the qubit levels, cannot be detected. Of course it is not impossible in principle to detect leakage errors. One could for example use the qubit as the control bit of a quantum computation that flips a target qubit if our control bit is populated and leaves the target bit unaffected if the control bit does not have any population [24]. Subsequently one would have to detect whether the target bit has been flipped or not. This could in principle be done using quantum jump techniques [25–28] to detect whether the target bit is in state 0 or state 1. However, there are a number of problems with this approach. Quantum jump techniques for the detection of qubit states require the scattering of at least $10^3 - 10^4$ photons (assuming today’s counter efficiencies) before our detection probability can come close to unity. This requires a long time, much larger than $10^{-4}s$, which would lengthen the quantum computation even further. In addition the scattering of 10^4 photons off a transition can lead to an excitation of phonons in the ion trap which can affect the subsequent calculation. For a Lamb-Dicke parameter $\eta = 0.1$ and 10 ions in the trap this would lead, on average, to the excitation of around 10 phonons when we check every qubit for leakage errors. Finally, the detection of a leakage error does not tell us to which level in the ion the population has leaked to. Therefore repumping seems to be unrealistic and instead the ion has to be replaced. This requires the storage of additional reserve ions which further increases the number of ions in the trap thereby slowing down the computer even more. In summary this indicates that leakage errors limit the performance of quantum error correction codes when applied in real atoms.

III. ACCURACY THRESHOLDS TO QUANTUM COMPUTATION

A. Quantum Computer

An input to a quantum computer is a string of quantum bits– *qubits*. A quantum computer is viewed as consisting of two main parts: *quantum gates* and *quantum wires* [29].

By *basic* quantum gates we mean any set of quantum gates which can perform any desired quantum computation. A universal quantum gate is the one whose combination can be used to simulate any other quantum gate [29–31]. A quantum wire is used as a representation of that part of computation of any qubit where the evolution is a simple identity operation (i.e. no gate operates on the qubit), as well as the time qubit spends during the gate operation.

B. Errors in Computation

There are two main types of errors occurring during a quantum computation process. The first type are errors due to the interaction of qubits with the external environment which cause spontaneous emission and decoherence type of errors. These errors are easily represented as unitary evolutions of the joint qubit+environment system, after which the environment is traced out of the picture [32]. This results in transforming an initially pure state of the qubit into a statistical mixture of qubits. The most general type of interaction of this kind can be represented by using (P)ositive (O)perator (V)alued (M)easure type operators acting on the qubit only:

$$\rho \longrightarrow \sum A_i \rho A_i^\dagger \quad (36)$$

where $\sum A_i^\dagger A_i = I$, i.e. the POVM is complete. This is an incoherent type of error.

The second type of error is an error in quantum gates where the qubit is “over-rotated” by a certain small amount. This is a coherent type of error which is simply represented by a unitary transformation on the qubit:

$$\rho \longrightarrow V_i \rho V_i^\dagger \quad (37)$$

where the “true” transformation should have been $\rho \longrightarrow U_i \rho U_i^\dagger$. The difference between U and V signifies the size of the error. This type of error can be represented as a special case of POVM operators in eq. (36), but the error it produces is different in nature to an incoherent error. We will discuss in the next section how to deal with this (assumption 4.).

C. Fault-tolerant Computation

The idea of fault tolerant quantum computation [22,33] is to encode the qubits in such a way that the encoding does not introduce more errors than previously present. If the error stays at the same level we then keep performing the error correction until the error has decreased in magnitude [23,33,34]. The present state of the art requires 5 – 10 qubits to encode a single qubit against a single error. It is the iterative application “in depth” of the encoding that will enable us to reduce error to an arbitrarily small level providing it is below a certain level to start with. In other words we will be encoding the encoding bits. Before giving the main argument in detail we list the assumptions that went into our calculations.

D. Assumptions

1. Qubits err independently. While this is usually true for incoherent errors it is not true for coherent ones. As an example take any conditional gate (a Controlled-Not for example). If the error affects the control qubit it is always transferred to the target bit. This problem is already taken care of by the way we perform fault tolerant quantum computation [33], and therefore we disregard it completely while maintaining the original assumption.
2. There are three basic incoherent errors which occur at the same rate (isotropic decay is assumed for simplicity). In the Born-Markov approximation the master equation for a single qubit is of the Lindblad form:

$$\dot{\rho} = \sum_i A_i^\dagger \rho A_i - \frac{1}{2} A_i^\dagger A_i \rho - \frac{1}{2} \rho A_i^\dagger A_i + i[H, \rho] \quad (38)$$

Three basic single qubit errors are represented by substituting the Pauli spin matrices instead of A 's in the above. Let us see what happens when we substitute σ_3 as the only entry for A , and neglect the unitary part of the evolution. Then the diagonal elements of ρ will not change since σ_3 commutes with the qubits Hamiltonian. The off-diagonals (assumed real for convenience) do change, however, according to:

$$\rho_{01}(t) = \rho_{01}(t=0)e^{-\gamma t} \quad (39)$$

where γ is the decay rate and t is time. Therefore what is initially a coherent superposition of the qubits two states will evolve in time into an incoherent mixture. This is going to be the basis of incoherent errors in our calculation. We assume that the probability that there is no error in n qubits after time t is

$$p_{ne} = e^{-n\Gamma t} \quad (40)$$

where $\Gamma = 3\gamma$ is the decay rate for all the errors together. The probability that there is *at least one error* is therefore $p_e = 1 - p_{ne}$. For convenience we introduce $\eta = \Gamma t$ henceforth.

3. Incoherent and coherent errors are independent. This, and the next, are the most insecure of our assumptions. To assume otherwise would complicate the argument enormously, and we avoid the complication for the sake of clarity of our argument and calculation.
4. Errors are treated “classically”. This partly follows from the previous assumption. We assume that the coherent error per gate is ϵ (no error is $1 - \epsilon$) in a completely classical fashion; this implies that either the qubit was correctly rotated with probability $1 - \epsilon$ or it was over-rotated with the probability ϵ . These two possibilities exist in a classical combination in our approximation rather than in a coherent superposition which really is the case. Moreover, we take the combination of coherent and incoherent errors to be a simple addition. This can be justified on the grounds that an error can always be defined as a distance (some proper metric) between what should be the outcome of a certain operation and what in fact is the actual outcome. Since any metric satisfies the triangle inequality this implies that the maximum rate of error after two steps would be the sum of the errors from each individual step.

5. Basic gates are all two-qubit gates. It is true that almost all two-bit gates are universal. Shor [1], however, also used a Toffoli (three-bit)-gate for his fault-tolerant computation, but this does not change our argument.
6. The number of gates needed to encode and fault tolerantly correct basic gate are of the order $\mathcal{O}(l^2)$, where l is the overall number of qubits [22].

E. Calculation

Our heuristic argument for stable quantum computation runs as follows: we require that the probability of error after the fault-tolerantly encoded basic gate is of higher order (i.e. the error is smaller) than the probability of error after the unencoded gate (that is the whole point of encoding and fault-tolerant error correction!). From this we derive the bound on the size of allowed errors in the wires and in the gates. We argue that when we encode the encoding bits again we reduce the error further and that by doing this we can reduce the error arbitrarily for an arbitrarily long computation. Therefore given certain initial limits on the error rate in the gates and wires we can stabilize any computation to a desirably small error rate, given an unlimited amount of time.

The probability of having any of the three basic errors in the first as well as in the second wire is η -giving the overall first order wire error of 2η . The error in the gate is ϵ . We assume that the overall error of the whole basic gate is $\leq 2\eta + \epsilon$.

Suppose that the basic gate is now encoded fault tolerantly against a single error of any kind, using l qubits. Then the overall second order error is at the end of the gate:

$$\eta^*(\eta, \epsilon, l) = \left(1 - \frac{l(l-1)}{2}l^4\eta^2\right)l^2\epsilon + \frac{l(l-1)}{2}l^4\eta^2(1 - l^2\epsilon) \quad , \quad (41)$$

i.e. it is equal to having error in the wires (this time it is second order) and not in the gates plus having error in the gates and not in the wires. The term $l(l-1)/2$ comes from choosing two out of l gates to err and the factor l^4 comes from the fact that there are l^2 gates so that

the error is transformed according to $\eta \rightarrow l^2\eta$ and that the error is of second order. What we require is that the fault tolerant error correction reduces the error. Hence:

$$(1 - \frac{l(l-1)}{2}l^4\eta^2)l^2\epsilon + \frac{l(l-1)}{2}l^4\eta^2(1-\epsilon) \leq 2\eta + \epsilon . \quad (42)$$

Since the LHS is $> \eta$ we can simplify the above condition without a greater loss in generality to:

$$(1 - \frac{l(l-1)}{2}l^4\eta^2)l^2\epsilon + \frac{l(l-1)}{2}l^4\eta^2(1-\epsilon) \leq \eta , \quad (43)$$

The solutions to the equation derived from the above are:

$$\eta_{\pm} = \frac{1 \pm \sqrt{1 - 2(l^8\epsilon - 2l^{10}\epsilon^2)}}{(l^6 - 2l^8\epsilon)} , \quad (44)$$

We require that $\eta \in \mathcal{R}$ (and that $0 \leq \eta \leq 1/2$) so that we have the following two regimes of error:

1. $0 < \eta < \eta_+$ and $e \leq e_-$
2. $0 < \eta < \eta_-$ and $e \geq e_+$

where

$$\epsilon_{\pm} = \frac{1}{2l^2}(1 \pm \sqrt{1 - 2l^{-6}}) \quad (45)$$

We now imagine that the output of the first encoded basic gate is fed into the next one (or part of the output into one next basic gate: and the rest into another next basic gate, it does not really change the conclusion). It is evident that if condition 1 holds, further encoding can only decrease the error. The residual error that we did not take into account is $\sim l^3(l^2\eta)^3 = l^9\eta^3$ (i.e. the second order error is not corrected by our encoding). In the extreme (worst) case when $\epsilon = \epsilon_- \sim l^{-8}$ we get $\eta \sim l^{-6}$, which means that the residual uncorrected error is $\sim l^{-9}$. Obviously, this error can accumulate over time if the computation is sufficiently long. In general, however the residual error after n in depth encodings is $l^{-\mathcal{O}(n)}$ which can be made arbitrarily small using sufficiently large n . Unfortunately, in this case

the number of elementary gates needed for each new encoding would square, i.e. the time of computation ($\sim$ the number of elementary gates) would increase exponentially. So, whereas in theory any computation can be performed with arbitrary accuracy given an initial $\eta \sim l^{-6}$ per gate, the time of computation will grow as $1/(\text{residual error})$.

In the second case, however, $\eta_+ \leq 0$ and $\eta_- \geq 1/2$. Thus in the best case, when $\epsilon = \epsilon_+$, η tends asymptotically to $1/2$. This would imply, for example, that a qubit in an initially pure state, e.g. $|0\rangle$, would eventually evolve into a completely random mixture at the output, of the form $1/2(|0\rangle\langle 0| + |1\rangle\langle 1|)$. Therefore under condition 2 the error rate is too big to allow any useful computation. However, under condition 1 the stabilization of error is clearly accomplished.

We next interpret our result in the light of the ion-trap realization of quantum computation.

IV. THE ERROR RATE IN ONE QUANTUM GATE

In the previous section we have seen how to estimate the accuracy threshold for quantum computation with a simple argument. In explicit constructions of fault-tolerant networks one can calculate the error threshold more precisely and it is found that the incoherent error rate per quantum gate should not be higher than around $10^{-5} - 10^{-6}$ [13]. In this last section we will now see whether accuracies of that order can be achieved in a linear ion trap realisation of the quantum computer using as qubits Zeeman sublevels in the ions given in tables 1 and 2. We emphasise that we take into account only the spontaneous emission of the ions and assume all the other errors disappear.

We calculate the probability to suffer at least one spontaneous emission during the implementation of N quantum gates. This probability has to be smaller than unity. Again we represent the qubit by two Zeeman sublevels and use Raman pulses to transfer population between the two states. For the time required to perform N quantum gates we find from the Hamiltonian eq. (2)

$$T = N \frac{8\pi\Delta_2}{\Omega_{02}^2} . \quad (46)$$

From that we obtain the probability for a spontaneous emission from level 2 as

$$p_2 = \frac{8\Gamma_{22}N}{\Delta_2} \quad (47)$$

Now again we have to take into account the fact that the two-level approximation can break down. This leads to an additional independent source of spontaneous emission via extraneous levels. One finally obtains in a completely analogous way to the derivation in Section 2 for the probability to have a spontaneous emission from an extraneous level

$$p_3 = \frac{80\Gamma_{33}^2\pi^2N^2L}{\Delta_{13}^2\beta\eta^2} \left(\frac{\omega_{12}}{\omega_{13}}\right)^3 . \quad (48)$$

The total probability of a spontaneous emission is $p_{tot} = p_2 + p_3$ and therefore the error rate per quantum gate is

$$r = \frac{p_{tot}}{N} = \sqrt{\frac{320L}{\beta}} \frac{\pi\Gamma_{33}}{\Delta_{13}\eta} \left(\frac{\omega_{12}}{\omega_{13}}\right)^{3/2} . \quad (49)$$

We can use the data for the ions given in table 1 and 2 in eq. (49). If we assume $\eta = 1, \beta = 1$ as well as in [13] $L = 7$ and an optimistic $N = 10^6$ we see that even for Barium the probability for at least one emission is almost one. This means that unless the encoding procedures given in [13,14] can be improved substantially the accuracy threshold for quantum computation may not be achievable. Work to this end is underway.

V. CONCLUSIONS

In this article we have studied the impact of decoherence and spontaneous emission on the practical applicability of quantum computation in linear ion traps and especially the possibility of using a quantum computer to factorize large numbers. We conclude that with present technology such a factorization will not be possible, even if we employ sophisticated methods of quantum error correction. We have shown that the numbers that can be factorized will be restricted to almost trivial sizes. We then investigated the minimal error

rate per quantum gate and compared it to recently established accuracy thresholds that would, in principle, allow arbitrarily complicated quantum computations. We find that the presently known thresholds may not be easily achievable because of spontaneous emission alone. Other sources of error would lead to even stronger limitations. We conclude that new physical ideas are therefore necessary if the goal of practically useful quantum computation is to be reached.

VI. ACKNOWLEDGEMENTS

We acknowledge discussions with D.F.V. James and R.J. Hughes and thank them for providing us with the experimental data for the ions. This work was supported by a European Community Network, the UK Engineering and Physical Sciences Research Council, by a Feodor-Lynen grant of the Alexander von Humboldt Foundation, by the Knight Trust, and by the Japan Society for the Promotion of Science.

REFERENCES

- [1] Shor, P. W. (1994): in *Proceedings of the 35th Annual Symposium on the Foundations of Computer Science, Los Alamitos*, CA (IEEE Computer Society Press, New York), p. 124.
- [2] Ekert, A. and Josza, R., *Rev. Mod. Phys.* **68**, 733 - 753 (1996).
- [3] A. Steane, *Rep. Prog. Phys.***61**, 117-173 (1998); R.J. Hughes et al, : *The Los Alamos Trapped Ion Quantum Computer Experiment*, submitted to *Fortschritte der Physik*, quant-ph/9708050; Vedral, V. and Plenio, M.B., *Basics of Quantum Computation*, to appear in *Prog. Quant. Elect.* (1998).
- [4] Barenco, A., *Contemp. Phys.* **37**, 375-39 (1996).
- [5] Barenco, A., Deutsch, D., Ekert, A. and Josza, R., *Phys. Rev. Lett.* **74**, 4083-4086 (1995).
- [6] Sleator, T. and Weinfurter, H., *Phys. Rev. Lett* **74**, 4087-4090 (1995).
- [7] Cirac, J.I. and Zoller, P., *Phys. Rev. Lett.* **74**, 4091-4094 (1995).
- [8] Monroe, C., Meekhof, D.M., King, B.E., Itano, W.M. and Wineland, D.J., *Phys. Rev. Lett* **75**, 4714-4717 (1995).
- [9] B.E. King, C.S. Wood, C.J. Myatt, Q.A. Turchette, D. Leibfried, W.M. Itano, C. Monroe, and D.J. Wineland, e-print quant-ph/9803023.
- [10] Plenio, M.B., and Knight, P.L., *Phys. Rev. A* **53**, 2986-2990 (1996).
- [11] Plenio, M.B., and Knight, P.L., *Proc. Roy. Soc. A* **453**, 2017-2041 (1997).
- [12] Plenio, M.B., and Knight, P.L., in *New Developments on Fundamental Problems in Quantum Physics*, edited by M. Ferrero and A. van der Merwe (Kluwer, Dordrecht), p. 311 (1997).

- [13] E. Knill, R. Laflamme, and W. Zurek, *Accuracy threshold for Quantum Computation*, preprint quant-ph/9610011.
- [14] D. Aharonov and M. Ben-Or, *Fault-tolerant quantum computation with constant error*, preprint (1996).
- [15] Hughes, R.J., James, D.F.V., Knill, E.H., Laflamme, R. and Petschek, A.G., Phys. Rev. Lett. **77** 3240-3243 (1996).
- [16] James, D.F.V. and Hughes, R.J., *Global limits on quantum computation with cold trapped ions I: two level qubits*, preprint (1996).
- [17] Vedral, V., Barenco, A. and Ekert, A., Phys. Rev. A **54**, 147-153 (1996).
- [18] Barenco, A., Bennett, C.H., Cleve, R., DiVincenzo, D.P., Margolus, N., Shor, P., Sleator, T., Smolin, J.A. and Weinfurter, H., Phys. Rev. A **52**, 3457-3467 (1995).
- [19] Meekhof, D.M., Monroe, C., King, B.E., Itano, W.M., and Wineland, D.J., Phys. Rev. Lett. **76**, 1796-1799 (1996).
- [20] Murao, M., and Knight, P.L., *Decoherence in nonclassical motional states of a trapped ion*, quant-ph/9803041.
- [21] Shibata, F. and Arimitsu, T., J. Phys. Soc. Jpn **49**, 891-897 (1980).
- [22] Shor, P.W., *Fault-Tolerant Quantum Computation*, quant-ph/9605011.
- [23] Plenio, M.B., V. Vedral and Knight, P.L., Phys. Rev. A **55**, 3327-3332 (1997).
- [24] Preskill, J., private communication (1996).
- [25] Bergquist, J.C., Hulet, R.G., Itano, W.M. and Wineland, D.J., Phys. Rev. Lett. **57**, 1699 (1986).
- [26] Sauter, Th., Blatt, R., Neuhauser, W. and Toschek, P.E., Opt. Comm. **60**, 287-292 (1986).

- [27] Sauter, Th., Neuhauser, W., Blatt, R. and Toschek, P.E., Phys. Rev. Lett. **57**, 1696-1698 (1986).
- [28] Nagourney, W., Sandberg, J. and Dehmelt, H. Phys. Rev. Lett. **56**, 2797 (1986).
- [29] D. Deutsch, Proc. R. Soc. London Ser. A **425**, 73 (1989).
- [30] DiVincenzo, D.P., Phys. Rev. A **50**, 1015 (1995).
- [31] A. Barenco, Proc. R. Soc. London Ser. A **444**, 678 (1995).
- [32] P.L. Knight, M.B. Plenio, and V. Vedral, Phil. Trans. R. Soc. Lond. A **355**, 2381 (1997).
- [33] Shor, P.W., Phys. Rev. A **52**, R2493-2496 (1995).
- [34] DiVincenzo, D.P. and Shor, P.W., Phys. Rev. Lett. **77**, 3260- 3263 (1996).

FIGURES

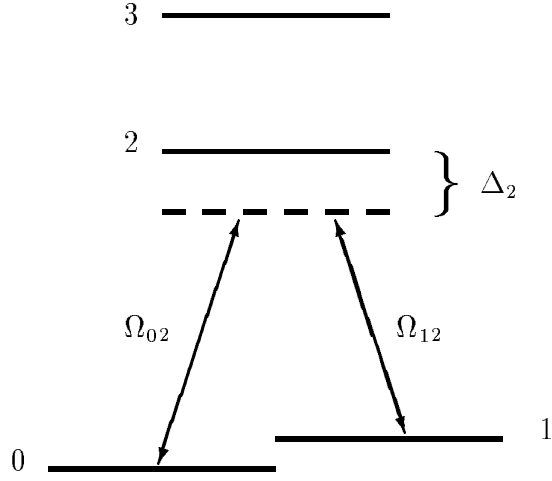


FIG. 1. The two lower levels 0 and 1 of the Λ system are Zeeman sublevels of a stable ground state and represent the logical values 0 and 1 of the qubit. They are coupled via Raman pulses that are strongly detuned from the intermediate level 2. The two-photon detuning is assumed to be zero while the one photon detuning Δ_2 is much larger than the associated Rabi frequencies on the $i \leftrightarrow 2$ transition Ω_{i2} . The lasers performing the Raman pulses may also couple the two lower states to an extraneous level 3 which may subsequently decay spontaneously.

TABLES

L(Lamb-Dicke parameter)	Ca ⁺	Hg ⁺	Ba ⁺
$L(\eta = 1)$	14.0	4.2	24.4
$L(\eta = 0.01)$	4.0	1.4	6.4

TABLE I. For several possible systems the upper limit on the bitsize L of the number N that can be factorized on a quantum computer is calculated according to eq. (33). A qubit is stored in two Zeeman sublevels of a stable ground state.

L(Lamb-Dicke parameter)	Ca ⁺	Hg ⁺	Ba ⁺
$L(\eta = 1)$	27.0	26	38
$L(\eta = 0.01)$	5.0	4.9	7.2

TABLE II. For several possible systems the upper limit on the bitsize L of the number N that can be factorized on a quantum computer is calculated according to eq. (35). A qubit is stored in two Zeeman sublevels of a stable ground state and a single error correcting code is used to protect the system. We have assumed $q = c = 5$.

Lamb-Dicke parameter	Ion	Ca ⁺	Hg ⁺	Ba ⁺
$\eta = 1$	r=	$2.03 \cdot 10^{-6}$	$9.26 \cdot 10^{-6}$	$0.44 \cdot 10^{-6}$

TABLE III. The error rate r per quantum gate if the qubit is represented by Zeeman sublevels. In [13] an accuracy threshold of roughly 10^{-6} was given. Spontaneous emission alone seems to rule out the possibility to achieve such a limit unless the procedures in [13,14] can be improved substantially.