Towards Quantum Information Processing with Atomic and Solid State Systems
Signe Seidelin

To cite this version:

HAL Id: tel-00579308
https://tel.archives-ouvertes.fr/tel-00579308
Submitted on 23 Mar 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Towards Quantum Information Processing with Atomic and Solid State Systems

HABILITATION À DIRIGER DES RECHERCHES

Soutenue publiquement le 8 février 2011
par
Signe Louise SEIDELIN

Composition du jury :

Thomas COUDREAU        Rapporteur
Juan Ariel LEVENSON     Rapporteur
Jean-Michel RAIMOND     Rapporteur
Alain ASPECT            Examineur
Benoit BOULANGER        Examineur
Jean-Michel GÉRARD      Examineur

Travail de recherche effectué à

Institut Néel, équipe mixte CNRS/CEA/UJF ‘Nano-Physique et Semi-Conducteurs’, Grenoble, France

National Institute of Standards and Technology, Time and Frequency Division, Boulder, CO, USA.
Acknowledgment

Boulder, Colorado, USA

The first part of this work describes the experiments performed during my time in Dave Wineland’s group in Boulder. I wish to thank Dave with all my heart for this wonderful and exciting time. Dave has impressed me very deeply, and I truly see him as a role model of an ideal group leader.

I am more than grateful to Didi Leibfried. Without Didi, I would probably still be short-circuiting surface-electrode ion traps to this day! I thank Didi not only for his guidance with the experiments, but guidance with life in general!

I also want to thank John Chiavarini for a very enjoyable time together. John taught me all the cleanroom tricks, and patiently watched me break one wafer after another (at a certain point, he suggested that “we” should rather used fused silica instead of the expensive sapphire wafers, I wonder why?). I was indeed very sad when John left the group to go and start his position in Los Alamos. Now John is a scientist at the MIT Lincoln Laboratory, and I wish him all the best luck there.

Brad Blakestad also worked on ion traps, and we went through quite a few “electroplating-nightmares” together. And yes Brad, I confirm, I share your contempt for this horrible activity (see acknowledgments in ref. [1]!). But besides the electroplating, I enjoyed a lot working with Brad. Jason Amini joined us in these nightmares when he arrived in the group, and added fearlessly a few more to the equation. Just to give a few examples: “how to drill a few microns wide hole in a 500 microns thick wafer” or even worse “how to microfab 1 nF capacitors without shorts on quartz wafers”. I think that the term “to scale up” more applies to the number of problems than the number of qubits! I admire Jason’s courage and his countless Sundays spent in the cleanroom .. and his kindness towards helpless kittens, but that’s a whole different story. He also saved me once from being eaten by a Grizzly bear in Banff National park in Canada, what I am of course very grateful for. Ryan Epstein also helped me a lot with the trap project, and deserves a special thank you, both for this, but also for his good and refreshing sense of humor! Joe Britton obviously deserves my gratitude as well. It has been great to work with him. I thank him for countless inspiring discussions (including comments on this manuscript), a few crazy tandem bike rides, in addition to putting me on the 6 am (3 times a week!) competition swimming team at CU. I actually managed to make some progress: I went from being the slowest to the second slowest swimmer on the team in less than a year. And thanks to advice from Marie Delaney, I avoided to
go sports gliding with Joe, which probably saved our friendship!

Marie by the way helped me in many other ways. As a compatriot at NIST, she took me cross-country skiing, she showed me the best Boulder coffees shops, and she even taught me some basic Danish. Which turned out to be very useful, as Janus H. Wesenberg, another Dane, soon after joined the group. Thanks to Marie, I could impress him with an almost fluent Danish! Janus deserves many thanks for his help both in the lab and also with some quite advanced ion trap simulations. To stay among Danes, I also thank Anders Brusch, who arrived at the end of my stay, especially for sharing his secrets concerning how to handle “French Administration”.

Jonathan Home joined the group one year before I left, and became not just my office mate but also my very reliable hiking partner. Together, we have admired Long’s Peak from every possible angle, and I still miss the taste of “Jonathan’s Tuna-surprise”.

I could go on forever, but to make a long story short, every single group member contributed to making my stay at NIST wonderful. Thus I thank Jim Bergquist, John Bollinger, Manny Knill, Wayne Itano, Till Rosenband, Rainer Reichle, Roee Ozeri, Chris Langer, John Jost, David Hume, Nobu Shiga, Christian Ospelkaus, Windell Oskay, Piet Schmidt, Jeroen Koelemeij, James Chou, Aaron P. VanDevender, Luca Lorini, Kenton Brown and last but not least Tobias Schaetz (who made me laugh many times!).

Grenoble, France

The second part of this work covers my research done at Institut Néel in Grenoble while being an assistant professor at Université Joseph Fourier. When I arrived in Grenoble, the Néel institute was being directed by Alain Fontaine, and I thank him for welcoming me at his institute, now three years ago. Today, I wish to express my gratitude towards him for making me feel valuable to the institute.

I also wish to show my appreciation to Alain Schuhl as the person being in charge of the “habilitation à diriger des recherches”-program. His good advice concerning the outline of the manuscript helped with the writing and definitely made it appear much more coherent. At Néel, we are now so lucky as to have Alain Schuhl being the director of the institute.

I would also like to thank Henri Mariette as the leader of the research team “Nano-Physique et Semi-Conducteurs”. I am indeed very grateful for the understanding, patience and kindness he has shown me.

Both Joël Cibert as the former director of my department, the “Département Nano” and Hervé Courtois, the current director of this department, deserve my gratitude for their support, and for helping me to determine my future research directions.
Furthermore I thank Jean-Philippe Poizat and Alexia Auffèves for receiving me as a part of their team and for helpful discussions. Gilles Nogues’s arrival in this team in 2009 was extremely rewarding. Gilles has helped me on many occasions, including reading and commenting on this manuscript. I hope very much to be able to keep sharing new ideas and visions with him.

As we all know, the Ph.D. students are the true heroes in this game. This also applies to Mathieu Munsch, with whom I have had the great pleasure to work. We have gone through quite a few ups and downs together, what is common in experimental physics, but Mathieu’s optimism always conquered. Inah Yeo replaced Mathieu when he graduated in 2009, and she has truly impressed me with her enthusiasm and very hard work in the lab.

Many more people deserve my gratitude for making the past three years an inspiring and instructive time: Le Si Dang, Maxime Richard, Jacek Kasprzak, David Ferrand, Lucien Bescombes and Claire Le Gall just to mention a few.

I am equally grateful for efficient assistance from the administrative staff at Néel, not to forget Patrick Belmain who saved my computer (and me!) from recurrent crashes.

Also the members of our joint team “on-the-other-side-of-the-fence” deserve a special thank you. I thank every single one of them, but in particular Jean-Michel Gérard, Julien Claudon, Joël Bleuse and of course Régis André, whose humor and grace have warmed my heart on many occasions.

I would also like to use the occasion to thank my colleagues “on-the-other-side-of-town”, that is to say, my teaching colleagues. First and foremost, Gilbert Vincent. It would take several pages to list all that Gilbert has done for me, so I don’t even try. I am very sad that Gilbert will retire this summer. Now it’s up to me to stay fit enough to follow him on his mountain bike, if I want to still be able to spend time with him. This will not be an easy task! I also want to thank Christian Hoffmann, Julien Douady, Christophe Durand, Martina Flöser, Hélène Béa, Justine Laurent, Oriane Mollet, Bénédicte Hingant, Hermann Sellier, Guillaume Méjean, Estelle Moraux, Antoine Delon and Thierry Klein, and all the others with whom I have had the pleasure to engage in this noble activity.

The first few months of my research in Grenoble were spent at “Spectro” (now LiPhy), and I have kept a strong attachment with people there, both through teaching but also through ... eating! Along these lines, I have spent many enjoyable moments with Bahram Houchmandzadeh (trying to avoid using the word “ski”- if not, I will have to pay the coffee for everybody according to the rules of a sordid game that I am still trying to understand!), Irina Mihalcescu, Clément Nizak and many others.

Still at LiPhy, I also want to thank Hugues Guillem de Chatellus, Irène Wang, Théodora Scheul, Jean-Claude Vial, Jacques Dériouard and Thierry Dombre for their
I thank my mother-in-law for showing a genuine interest in my work!

I wish to express my gratitude towards the members in my examination committee. Thomas Coudreau, Ariel Levenson et Jean-Michel Raimond kindly accepted to review and report on the manuscript. I am equally very grateful to Benoit Boulanger for chairing the committee in addition to his help and good advice the everyday life. I genuinely thank Alain Aspect and Jean-Michel Gérard who also honored me by participating in this adventure. I thank them all for a very inspiring discussion at the end of my defense, with several questions that still make me ponder. In exchange, together they gave me valuable advice concerning my future research projects.

This event marks for me the end of an important chapter in my life, and the beginning of a new and very exciting one. I am more than delighted to be able to share this new chapter with Olivier Arcizet. I could not have found a better research partner. Not only his astonishing talents as a physicist but also as a human being have already made a deep impression on me.

To finish, I wish to thank my parents for being so supportive and understanding at all times. I wish I could make the sun shine more often in Denmark, and make the dark winter nights shorter, to express my gratitude ...

My last words are addressed to the person to whom I owe it all. Without him, nothing would have been possible. Yann, merci pour tout.
## Contents

### Introduction

1 Atomic Systems - Quantum Information Processing

1 Basics Techniques for Trapped Atomic Ions

1.1 The Ion Trapping Environment

1.1.1 Paul Traps

1.1.2 Trap Surroundings and Setup

1.2 Ion Manipulation with Lasers

1.2.1 Ion Species and Atomic Structure

1.2.2 Loading and Doppler Cooling

1.2.3 State Detection

1.2.4 Raman Transitions and Ground State Cooling

2 Quantum Computing with Trapped Atomic Ions

2.1 Quantum Gates for Ions

2.1.1 Fidelity and the Error Threshold

2.1.2 Benchmarking of Quantum Operations

2.1.3 Sources of decoherence

2.2 Entanglement as a Resource for Quantum Computing

2.2.1 A Six-Ion Schrödinger Cat State

2.2.2 Purification of Entanglement

2.3 Perspectives: The Next Milestone

3 Chip–Style Ion Microtraps for Quantum Computing

3.1 Towards a Scalable Microtrap

3.2 A Surface-Electrode Design

3.3 Fabrication Process and Practical Challenges

3.4 Testing and Characterizing the Surface-Electrode Trap

3.4.1 Trap Oscillation Frequencies

3.4.2 Heating Rates

3.5 State of the Art for Ion Micro Traps

3.5.1 Pushing the Chip Design Further

3.6 Perspectives: Microtraps in Progress
II Solid State Systems - Toward a Quantum Interface 55

4 Unity Photon Absorption Efficiency of a Single Quantum Dot 57
   4.1 Cavity Quantum Electrodynamics ................................. 58
      4.1.1 The Purcell Effect .............................................. 59
   4.2 A One-Dimensional System ......................................... 59
   4.3 Quantum Dots as Artificial Atoms .................................. 60
   4.4 Micro Cavities for Quantum Dots ................................... 62

5 Measurements of the Purcell Effect in a Semiconductor Microcavity 65
   5.1 Measurement of the Purcell Factor ................................. 65
   5.2 Micropillar Samples and Apparatus ................................. 66
      5.2.1 Photoluminescence Setup ...................................... 67
      5.2.2 Sample Characteristics ........................................ 68
      5.2.3 Time-resolved Measurements ................................... 69
      5.2.4 Modeling the System for Continuous-Wave Measurements .... 72
      5.2.5 Continuous-wave Measurements ................................. 74
   5.3 Discussion ............................................................... 78

Conclusion and Perspectives 79

A Compilation of Publications Relevant to Work Presented 83

Bibliography 193
Introduction and Outline

Einstein was a giant. His head was in the clouds, but his feet were on the ground. Those of us who are not so tall have to choose!


Richard Feynman was first to point out that a classical computer is not powerful enough to describe the quantum nature of a complex physical system. Instead, in a paper published in 1982, he suggested that one quantum system could ‘simulate’ another quantum system, and thereby he evoked for the first time the idea of a quantum computer. During the 1990s, it became clear that the imagined quantum computer also could be useful for other applications, for instance efficient factorization of prime numbers, as proposed by P. Shor, or certain search algorithms, for instance the one proposed by L. K. Grover.

In an attempt to clarify which physical systems hold promise as a quantum computer, D. P. DiVincenzo specified a number of criteria that the system must satisfy:

1. The quantum system must be ‘scalable’, that is, it must allow one to increase the number of qubits to meet the required level without loss of control.

2. There must be an ability to initialize all qubits into a known initial state, such as $|000...000\rangle$.

3. The coherence times of the qubits must be long compared to the gate operation times.

4. A universal set of quantum gates must exist, which will allow any quantum algorithm to be realized by successive applications of these gates. This universal set must include single-qubit operations that change the state of a single qubit, plus at least one two-qubit gate operation which can entangle the states of any two selected qubits.

5. There must be a method of reading out the state of any qubit.

---

This is in essence how DiVincenzo defined ‘scalability’. However, scalability today means different things to different people. If working with quantum algorithms, scalability might instead correspond to a sub-exponential growth of time required to solve a given problem, or it might be related to the increase of resources required to implement a sufficient number of error-correction codes etc.
In 1995, Ignacio Cirac and Peter Zoller suggested that an ensemble of cold confined atomic ions is an interesting system for quantum computing [6], and very quickly afterwards, criteria 2 to 5 were proven satisfied experimentally within this system (see for instance references [7,8]). These proof-of-principle experiments, although being convincing, were based on very few ions, and the ion trap designs used were not compatible with the scaling addressed in criterion 1. However, in contrast to some of the proposed systems such as nuclear magnetic resonance based quantum computing [9], no fundamental roadblock towards scalability had been identified. It is said that DiVincenzo did not include the scaling criteria in a first version of his list, but once he did, he put it as number one, emphasizing its importance. And indeed, today, a major part of the effort concerning quantum computing with trapped ions is directed towards scaling up the system. This also holds true for the work described in part I of this report. It is structured as follows:

Chapter 1 is a short introduction to well-established basic ion trapping techniques.

Chapter 2 demonstrates the capacity to create very large entangled state, and the application of a protocol for restoring lost entanglement.

Chapter 3 is a description of the progress in the development of potentially scalable ion trap.

Even though the progress discussed in part I is very encouraging, one of the yet unsolved issues with trapped ions is related to what has unofficially come to be known as the sixth DiVincenzo criterion, namely

6. Interconversion between stationary and flying qubits.

This criterion raises the question which concerns the distribution of quantum information. It is clear that photons are more suited for transporting quantum information over longer distances than ‘massive’ particle such as ions. The challenge then is to find a quantum system for which one can easily map quantum states between photons and material particles. Making a single ion absorb efficiently a single photon requires a perfect mode matching between the incident photon and the ion’s dipole radiation pattern. This is extremely difficult to achieve in free space, i.e. without a cavity surrounding the emitter. This is why, in part II of this work, we will explore a system in which an emitter is surrounded by a cavity. This is possible to do with trapped ions, but from a technological view point it might be easier to do by using solid state systems. In this work, we have used a quantum dot inserted into a semiconductor microcavity. These experiments are described in part II, which consists of two chapters organized as follows:

Chapter 4 is an introduction to some of the ideas from the field of cavity quantum electrodynamics, with accent on quantum dots in microcavities.

Chapter 5 describes a particular quantum dot-microcavity system, including several methods to probe the emitter-cavity interaction strength, which is quantified by the so-called Purcell Factor.

\(^2\text{We will define later what exactly we mean by ‘scalable’ ion traps.}\)
Part I of the work summarizes the results obtained during my three year long post-doc in Boulder, Colorado, USA, in the group of David Wineland at the ‘National Institute of Standards and Technology’ (NIST). My first year was devoted to the quantum information experiments. The success of these experiments is due to a shared effort between several members in the NIST group. The last two years were primarily devoted to designing and building a chip-style ion trap. During this time, I was working with two of the group’s Ph.D. students, Joe Britton and Brad Blakestad, who graduated in 2008 and 2010, respectively.

Part II describes the experiments performed from when I was appointed associate professor at the ‘Université Joseph Fourier’ (UJF) in Grenoble, and until today. This work is taking place at the ‘Institut Néel’, in the nano-physics department. The experiments described should in particular be credited to Mathieu Munsch, who completed his Ph.D. in our group in 2009. Since then, the experiments have been very successfully continued by our current Ph.D. student, Inah Yeo.

Even though part I and II both are about quantum information processing, each has its root in a different branch of physics. In essence, part I is based on atomic physics, whereas part II on solid state physics. I have therefore done my best to stay as simple in my writing as possible, in the hope that both experts and non-experts will benefit. The price to pay is lack of mathematical rigor and the risk of leaving out ‘details’ not necessarily considered as so by experts in the given field. I hope this price is worth paying.
Part I

Atomic Systems - Quantum Information Processing
CHAPTER 1

Basics Techniques for Trapped Atomic Ions

Science is facts; just as houses are made of stone, so is science made of facts; but a pile of stones is not a house, and a collection of facts is not necessarily science.


This chapter briefly outlines the main concepts related to ion trapping and manipulation, including how the ions are loaded into the trap, cooled and detected. It does not present new results, but provides the reader with the minimum background necessary to follow the discussion in chapter 2 and 3. Ion trapping techniques have been developed over a long period of time, making them quite sophisticated. With this in mind, it is clear that any brief description will necessarily be incomplete, and so will this one. A more complete review of basic ion trapping techniques can be found in the references [10,11].

1.1 The Ion Trapping Environment

In an ion trap, ions are confined by a coupling between external electric trapping fields and the atom’s net charge. This gives rise to a very deep trapping potential, typically around 1 eV. This is 40 times larger than \( k_B T \) at room temperature, which equals 0.025 eV. For neutral atoms, the depth of the trapping potential is much lower, typically around 1 meV\(^1\). The deep trapping potential associated with ion traps allows one to keep the ions trapped for hours, or, depending on trap geometry and surroundings, even days or months.

Earnshaw’s theorem states that charged particles cannot be confined in vacuum through the use of static electric fields alone because, according to Maxwell’s equations, the divergence of the electric field is zero. This restriction can be overcome through the combination of static electric fields with either static magnetic fields or oscillating electric fields. These two configurations are known as a Penning trap [12,13] and a radio frequency (RF) Paul trap [14], respectively. The Paul trap allows for a very high

---

\(^1\)This is true only for conservative trapping potentials, which are typically based on a coupling between an external field and the magnetic moment or optical dipole forces, in contrast to dissipative trap potentials, such as the magneto-optical trap (known as a MOT) in which the trap depth is deeper.
Figure 1.1: Illustration of a linear RF Paul trap. An oscillating potential is applied to the white RF electrodes which, together with DC electrodes, produce a quadrupole electric field pattern in the x-y plane. This provides radial confinement of the ions. Axial confinement is obtained by the segmented control electrodes: applying a higher DC potential to the ‘end-caps’ (green) relative to the middle fragment (gray) provides a potential minimum along the axis. If the axial confinement is weaker than radial confinement, which is usually the case, multiple ions will align in a chain along the z-axis. The distance between two rods in the x-y plane is typically of the order of 1 mm.

confinement of the ions, a desirable property for quantum information applications. So far, the Paul trap alone has been used for this purpose\(^2\). The ion trap described in chapter 3 is also based on the Paul trap design, and more generally, all future references are implicitly made to this particular type of trap.

1.1.1 Paul Traps

In the NIST quantum information processing (QIP) experiments we use what is called a ‘linear’ Paul trap, shown schematically in figure 1.1. A combination of static and oscillating fields create a quadrupole potential in the radial directions (x-y plane in figure), confining the ions between the rods. More precisely, an RF field is applied to two of the electrodes, while the other two electrodes are held at RF-ground. To obtain the axial confinement, static (DC) potentials are applied to the so-called control electrodes. The static (positive) potential applied to the middle section of the control electrodes is lower than the one applied to the end sections (also referred to as ‘end-caps’) in order to confine positively charged ions. When the axial confinement is weaker than the radial confinement, if several ions are present, they will align in a string, as illustrated in figure 1.1.

The motion of an ion in this field consists of two superimposed harmonically oscillating motions. One that is very fast (the micro motion) and one that is much slower.

---

\(^2\) Note that a theoretical proposal for using a Penning trap has been published [15], as well as experimental work in this direction [16][17]
1.1 The Ion Trapping Environment

The secular motion. Mathematically, the equation of motion can be analyzed in two different ways. Either as exact solutions of the Mathieu differential equations [10] or as solutions of a static effective potential called the ‘pseudopotential’ [13]. Whereas the Mathieu solutions provide insights on trap stability and high frequency motion, the pseudopotential approximation typically contains enough information for the analysis of trap designs, and has the advantage of being more simple to work with. For instance, the design of the ‘surface-electrode’ trap discussed in chapter 3 is based on pseudopotential simulations alone, which have proven sufficient for the purpose.

When we later discuss motional modes in the trapping potential for one or more ions, we implicitly refer to the secular motion as well. In the pseudopotential approximation, we can consider the trapping potential as harmonic in all three directions close to the potential minimum. For each direction $i = \{x, y, z\}$, we can assign a quantum number $n$ corresponding to the energy level in the trapping potential $E_n = \hbar \omega_i (n + 1/2)$. As mentioned, the two radial modes (in the $x,y$-plane in figure 1.1) have generally a higher frequency than the axial mode (in the $z$-plane). For the QIP experiments presented in this work, we always use the axial modes ($i = z$), and when nothing else is mentioned, $n$ is the quantum number corresponding to the given axial mode. If only one ion is present, there is only one axial mode, and in this case the ion oscillation frequency equals the trap frequency $\omega_z$. In general however, in one dimension, for $N$ ions, there will be $N$ motional modes, one of which is the center-of-mass mode. For a perfectly harmonic potential and identical ions, the center-of-mass frequency equals the single ion frequency, which in turn equals the effective trap frequency. All other modes are shifted in frequency, and if more than one ion is present, when we give a quantum number $n$, we will need to specify the associated axial mode.

1.1.2 Trap Surroundings and Setup

An ultra high vacuum environment ($\sim 1 \times 10^{-11}$ mbar) is required for doing quantum computing with trapped ions. At higher pressures, background collisions can cause ion heating and qubit decoherence. Collisions with background atoms or molecules may even lead to a chemical reaction, creating a different and undesired ion complex. For instance, $^9$Be$^+$ can react with H$_2$ forming a beryllium-hydride [10,19].

The ion trap vacuum system design that we use consists of a $\lambda/4$ coaxial RF resonator [20] surrounded by a quartz vacuum envelope with optical quality quartz windows for laser beam access and imaging. The boundary conditions for the RF cavity produce a voltage maximum at the tip of the resonator. The trap RF electrodes are attached to this tip. The resonator and glass envelope is attached to a standard stainless steel vacuum system, where the ultra high vacuum is maintained by use of an ion pump and occasional titanium sublimations. The setup is shown schematically in figure 1.2 including the imaging optics.

---

$^3$The micro motion, which is undesirable for quantum information purposes, is later minimized experimentally by fine tuning ‘by hand’ the voltages on the electrodes.

$^4$See reference [18] for a proposal of a scheme using the radial modes.
Chap 1 - Basics Techniques for Trapped Atomic Ions

1.2 Ion Manipulation with Lasers

Like neutral atoms, ions can be efficiently manipulated with laser beams. Alkali and alkali-like atoms have historically been a popular choice for neutral atom cooling- and trapping experiments. The single valence electron gives a simple and well-studied structure that is ideal for many experiments. To obtain a similar level scheme for ions, alkaline earth metals are often used. Once singly ionized they also contain a single valence electron, and therefore exhibit an atomic structure analogous to that of neutral alkali atoms. The main difference is that for the ion, the valence electron is generally more tightly bound due to the increased effective nuclear charge (the nuclear screening is diminished due to the absence of one electron). This tends to increase the atomic transition energies towards visible and UV wavelengths, whereas alkali-like atom transitions are typically closer to the infrared part of the spectrum.

Figure 1.2: Schematics of the trap surroundings. The trap is located at the tip of an RF resonator. Both resonator and trap is surrounded by a quartz envelope that is connected to a standard stainless steel vacuum system. The ions are imaged through an optical window, in this schematics by use of a CCD camera. An example of a string of four ions is shown to the right.
1.2 Ion Manipulation with Lasers

Figure 1.3: The internal energy level diagram for $^{9}$Be$^{+}$. Application of a magnetic field (13 G) both defines a quantization axis and breaks the degeneracy of the $m_F$ levels via the linear Zeeman effect. Hyperfine states $^{2}S_{1/2}|F = 2, m_F = -2⟩ ≡ |↓⟩$ and $^{2}S_{1/2}|F = 1, m_F = -1⟩ ≡ |↑⟩$ are selected to form the two-level qubit, with splitting $\omega_0/2\pi \approx 1.28$ GHz. The $^{2}P_{3/2}$ manifold is utilized during Doppler cooling, state detection, and state preparation (the corresponding transition is indicated by the dashed arrow). The $^{2}P_{1/2}$ manifold is utilized for ground state cooling, state preparation and quantum gates. Figure by Brad Blakestad [1].

1.2.1 Ion Species and Atomic Structure

Two different types of ions are used in this work, namely magnesium and beryllium ions. The QIP experiments presented in chapter 2 use $^{9}$Be$^{+}$ with a nuclear spin of 3/2 as the ion qubits. The level scheme of this ion is shown in figure 1.3. As indicated in the figure, two hyperfine states, $|F = 2, m_F = -2⟩$ and $|F = 1, m_F = -1⟩$ in the $^{2}S_{1/2}$ manifold, are chosen as the two-level qubit system. In the following, these states are denoted $|\downarrow⟩$ and $|\uparrow⟩$, respectively. The use of arrows is to make the analogy to a spin-$\frac{1}{2}$ system.

A strength of hyperfine qubits is their extremely long ‘spin’ relaxation time (equivalent to the $T_1$-time), which is on the order of $10^{15}$ s for $^{9}$Be$^{+}$ [10]. Although being separated by only 1.28 GHz, the qubit states can still be manipulated with optical transitions using Raman transitions, as we will discuss in section 1.2.4.

Some ion trapping groups use as qubits two levels which are separated not by an RF but by an optical wavelength, allowing one to make direct optical transitions between the two qubit states. An example is $^{40}$Ca$^{+}$ which is frequently used in the ion trapping
group lead by Rainer Blatt in Innsbruck. Here the qubit levels consist of an S and a metastable D state with a relaxation time of 1 s \[21\]. While it may be true that this is fairly long, it still can affect negatively the coherence time.

When we test the surface-electrode trap in chapter 3 we use magnesium ions. More precisely, during the initial stages of the testing, we use \( ^{24}\text{Mg}^+ \), which has the advantage of being simple to Doppler cool due to the absence of hyperfine structure. However, in order to measure accurately an ion heating rate, we will need to apply so-called sideband transitions, which in our case requires a hyperfine structure (see section 1.2.4). We therefore do these measurements with \( ^{25}\text{Mg}^+ \), possessing a hyperfine structure due to its nuclear spin of 5/2.

### 1.2.2 Loading and Doppler Cooling

Ions are loaded into a trap by ionizing neutral atoms passing near the trap center while cooling them with a Doppler laser beam. One easy way to ionize the atoms is by electron impact. To this end, a small wire is heated up which emits a sufficiently strong electron flux to ionize the atoms. This method is used for the experiments using \( ^{9}\text{Be}^+ \). However, using electron impact ionization significantly degrades the vacuum during load, and this combined with the shallow trap potential makes the method unsuccessful for loading ions into the surface-electrode trap (see chapter 3). Alternatively, we can use photo ionization which is more efficient and does not degrade the vacuum. For magnesium, this can be done using a two-photon transition at UV wavelengths \[22\].

For Doppler cooling, in contrast to neutral atoms, only a single laser beam is needed; trap strength far exceeds and therefore easily compensates the laser beam radiation pressure \[23\]. However, for motion perpendicular to the laser beam, the ion is potentially heated due to recoil. This can be avoided if the principal axes of the trapping potential are oriented such that they all have a non-zero projection onto the k-vector of the laser beam. In this configuration, all directions can be efficiently cooled with one single beam. In practice, this is achieved by orienting the laser beam adequately compared to the trapping axes.

For the QIP experiments using \( ^{9}\text{Be}^+ \), a transition (with a natural linewidth \( \Gamma_{\text{Be}} = 19.4 \text{ MHz} \)) between the \( ^2S_{1/2} \) and the \( ^2P_{3/2} \) manifold is utilized for Doppler cooling. The transition is driven by red-detuned 313 nm laser light. More accurately, we use a transition between \( ^2S_{1/2} \left| F = 2, m_F = 2 \right\rangle \) and \( ^2P_{3/2} \left| F = 3, m_F = 3 \right\rangle \). Using \( \sigma^+ \) polarized light, this becomes a cycling transition. To efficiently cool the ions both when they are very hot (i.e. when the Doppler shift due to atom motion is much larger than linewidth of the transition) but also at temperatures near the Doppler limit, we use two different detunings of the Doppler beam relative to this transition. One which is several hundreds of MHz detuned to cool very hot ions (during loading), one which is detuned by an amount equal to half of the natural linewidth of the transition, in order to interact with colder ions. As the cycling transition is not perfectly closed, a set of repumping beams are also needed, see for instance reference \[24\] for details on this.

For trap testing, when working with magnesium, Doppler cooling is achieved using 280 nm laser light. We use the cycling transition between the \( ^2S_{1/2} \) and the \( ^2P_{3/2} \) states. Also here we use two detunings, one very large (400 MHz), and one half of the natural
1.2 Ion Manipulation with Lasers

linewidth ($\Gamma_{\text{Mg}} = 40$ MHz).

The Doppler cooling can typically reduce the energy of the ion to a few motional quanta of the harmonic potential, depending on the trapping frequencies and the linewidth of the cooling transition. However, for quantum information purposes, one needs to get much closer to the ground state. This can be achieved only with sideband cooling, which is the subject of section 1.2.4.

1.2.3 State Detection

For both magnesium and beryllium, the Doppler transition is also used for detection. We use the Doppler beam which is closest to resonance, i.e. detuned approximately half a linewidth compared to the cycling transition. To simply confirm the presence of an ion, the fluorescence from the transition can be detected with a CCD camera. For the QIP experiments, it is necessary to be able to distinguish precisely between the two hyperfine states (figure 1.3) used for qubit levels. Due to the 1.28 GHz frequency difference between $|\downarrow\rangle$ and $|\uparrow\rangle$, we can selectively activate the cycling transition for one of the two state by choosing appropriately the corresponding laser detuning. Typically we set the laser on resonance for the $|\downarrow\rangle$ state, and consequently only this state fluoresces. As mentioned, selection rules and laser polarization make the transition closed, and the ion always decays back into the $|\downarrow\rangle$ state (reference [1] gives more details). By recording the fluorescence rate by the use of a photomultiplier tube, we can achieve an extremely precise state discrimination, better than 99.9 %.

1.2.4 Raman Transitions and Ground State Cooling

To do quantum information experiments, in addition to state detection, we need controlled transition between the two qubit levels, i.e. transition between $|\downarrow\rangle$ and $|\uparrow\rangle$. We call these transitions ‘single qubit rotations’ with the Bloch’s sphere for spin particles in mind.

Because of the small frequency separation between the levels, the transitions can not be done directly using optical wavelengths. Instead we use two-photon stimulated Raman transition using a virtual third level in the P state manifold. The frequency difference between the Raman beams corresponds to the hyperfine splitting, i.e. the energy difference between $|\downarrow\rangle$ and $|\uparrow\rangle$. This is illustrated by the gray lines in figure 1.4. The theory of this three-level, two-photon interaction is covered in [10,25]. Briefly, this complicated three level system can be simplified to a two-level analog when the two lasers are applied simultaneously, using an adiabatic elimination argument. The qualitative behavior is identical to that of a single-photon system, replacing the wave vector, the laser frequency and the phase as follows: $k \rightarrow k_1 - k_2$, $\omega_L \rightarrow \omega_{L1} - \omega_{L2}$, and $\phi \rightarrow \phi_1 - \phi_2$. Both here and in figure 1.4, the subscripts denote Raman beam 1 and 2, respectively. By applying the Raman beams continuously and setting $\hbar(\omega_{L2} - \omega_{L1})$ equal to be equal to the hyperfine splitting, we can observe Rabi oscillations between $|\downarrow\rangle$ and $|\uparrow\rangle$. Using well-defined durations of the Raman beam application time, we can place the entire population in the desired spin state. More precisely, by choosing the duration for the Raman beams to be half a Rabi period (a ‘pi-pulse’), we can transfer the full population...
The Raman transitions are also used for sideband cooling, by which the ion can be placed in the ground state of the trapping potential. To cool the motion of a single ion, or the center-of-mass motion for several ions, we do the following: instead of setting the frequency difference of the beams equal to the frequency of the transition $|\downarrow\rangle \leftrightarrow |\uparrow\rangle$ (which we call ‘the carrier’) we set it on the ‘red sideband’, which is detuned to the red by one quantum of the trapping frequency, $\omega_{\text{trap}}$ relative to the carrier. Hereby we can make transitions from $|\downarrow\rangle |n\rangle$ to $|\uparrow\rangle |n-1\rangle$, where the second bracket gives the motional quantum state for the trapping potential. This is illustrated by the red lines in figure 1.4. Repumping followed by spontaneous emission brings the ion back to the $|\downarrow\rangle$ state without changing the vibrational quantum number. By repeated applications of this transition, one can obtain $\langle n \rangle \leq 0.1$. One condition for the sideband cooling to be successful is that the recoil energy from the spontaneous emission ($\hbar\omega_{\text{recoil}}$) is smaller than the separation between the energy levels in the harmonic trapping potential (i.e. $\hbar\omega_{\text{trap}}$). This is quantified by the Lamb-Dicke parameter $\eta$ defined by $\eta^2 = \omega_{\text{recoil}}/\omega_{\text{trap}}$.
which needs to be smaller than one \([10]\), which requires a fairly strong ion confinement.

As mentioned earlier, if more ions are present, additional vibrational modes are present. These modes are shifted in frequency relative to the trap frequency, so to cool them, we must adjust the frequency difference between the Raman beams accordingly. Doing this, all vibrational modes can be cooled efficiently. In general, each time we start a QIP experiment, we cool all axial vibrational modes to the ground state.

The ion's average number of motional quanta (or its ‘temperature’) can be obtained by comparing the amplitude of the red (the cooling) and what we call the blue \((|↓⟩|n⟩ → |↑⟩|n + 1⟩)\) sideband (blue lines in figure 1.4). Roughly speaking, the closer the ion gets to the ground state, the weaker the red sideband becomes relative to the blue sideband. More precisely, the average number of quanta is given by \([26]\):

\[
\langle n \rangle = \frac{1}{A_B/A_R - 1} \tag{1.1}
\]

where \(A_B\) and \(A_R\) are the amplitudes of the blue and red sidebands, respectively. In practice, these amplitudes are obtained by scanning the laser frequencies across the sideband transitions using acousto-optical modulators.

One major use of the Raman beams are for the operating quantum gates. We will discuss how the quantum gates work in the next chapter, and we will use them to entangle trapped ions.
In section 1.2.4 we discussed how to make individual qubit rotations between the two spin states of an ion, by using Raman beams. Here, the two spin states are the hyperfine states which we denoted \( |\downarrow\rangle \) and \( |\uparrow\rangle \). Single qubit rotations associated with an entangling gate amounts to a ‘universal gate set’ in the sense of the DiVincenzo criteria \[5\] mentioned earlier. To complete our universal gate set, we therefore need an entangling gate between the spin states of two ions. The direct interaction between two ions’ spin states is negligible in our systems; however, an effective interaction can be achieved by coupling the spin states of the ions via their collective motion in the trap. This interaction is the corner stone for most quantum gates using trapped ions.

The original proposal to generate a universal two-qubit entangling gate in this manner was first suggested by I. Cirac and P. Zoller in 1995 \[6\], followed a few years later by a series of other schemes \[27\]–\[30\]. The Cirac-Zoller proposal turned out to be more difficult to implement than the later proposals, which were therefore demonstrated first, in particular in experiments performed at NIST \[31\]–\[32\]. The Cirac-Zoller proposal was later implemented in an Innsbruck experiment \[33\].

2.1 Quantum Gates for Ions

In the experiments that we will describe here, the main building block is the particular gate demonstrated in reference \[32\] which is known as a geometrical phase gate, and in the following, we will therefore focus on this gate alone. An excellent review on different ion gates can be found in \[24\].

One advantage of this gate is that, contrary to some of the first gates proposed, it does not require individual addressing of the ions. Indeed, the NIST setup used for the experiments presented here does not allow individual addressing of ions in a chain. To do so requires the laser beams to have a waist smaller than the separation between the ions, in our case often less than 5 \(\mu\)m. This is also true for the single qubit rotations; applied to a chain of ions, they will have a common action on all ions.
Figure 2.1: (a) By crossing two Raman beams in the trap center, we obtain a standing running wave. We adjust the spacing between the ions so that it matches an integer number of periods in the standing running wave. When this is done, each ion will be exposed to the same light intensity, for our particular spin states giving rise to a differential force $F_\downarrow = -2F_\uparrow$. (b) The Raman beams frequency difference is set equal to one quantum of the stretch mode of the two ion pair. This will excite increasingly the stretch mode during application of the beams. However, to instead excite and de-excite the mode, the frequency difference is shifted very slightly away from the stretch mode, as explained in the text.

A property of a quantum gate is its ‘conditioned action’: the action of a gate on one or more qubits depends on the initial state of these qubits. This conditioned action allows the gate to be used to create entanglement: in an entangled quantum state, the state of one qubit must depend on the states of the others. Consider for example a two-ion entangled state given by $\frac{1}{\sqrt{2}} \{\ket{\downarrow\uparrow} + \ket{\downarrow\uparrow}\}$ where for each ket the first arrow corresponds to ion 1 and the second arrow to ion 2. If a measurement is made and e.g. ion 1 is found in $\ket{\downarrow}$, we know that ion 2 must necessarily be in $\ket{\uparrow}$ and vice versa. This inter-ion spin dependence is created by the spin conditioned action of the gate.

In section 2.2.1 we will use the phase gate to produce a state of six entangled ions starting from a pure but unentangled state. In this section, for simplicity, we will discuss how the gate works for two trapped ions. The working principle of the gate remains the same when applied to a state containing a larger number of ions.

Both for two or more ions, we use as mentioned the axial motional modes for the gate. For a pair of two ion qubits in the harmonic trapping potential, along the axial direction, there will be two collective motional modes: a center-of-mass mode when the ions oscillate together in phase, and a ‘stretch’ mode, where they oscillate out of phase (also called a ‘breathing mode’). In order to excite the stretch mode, one needs ‘to pull the ions apart’, which in turn requires a force which direction is opposite on the two ions. We can create this differential force using the AC Stark shift on the ions from a standing running wave of laser light. The AC stark shift acts differently on different spin states, and it happens to be true that the force on the $\ket{\downarrow}$ state is opposite and two times larger than the force on the $\ket{\uparrow}$ state: $F_\downarrow = -2F_\uparrow$. This is illustrated in figure 2.1 (a). Note
that this relation between the forces is true only as long as the ions are exposed to the
same light intensity. To achieve this, we are careful to separate the ions by a distance
Corresponding to an integer number of wavelengths of the standing wave. If we then set
the frequency difference of the Raman beams equal to the frequency of the stretch mode
in order to conserve energy, the AC stark shift can excite the stretch mode whenever the
two ions are in different spin state, i.e. in the $|↓↑⟩$ or $|↑↓⟩$ state, see figure 2.1 (b).

If, on the other hand, the ions are in $|↓↓⟩$ or $|↑↑⟩$, no opposing force is created and
despite being subject to the same conditions as above, i.e. appropriate positioning and
laser frequencies, the stretch mode will not be excited for these two spin combinations.
Therefore, this scheme can potentially excite the stretch mode conditioned on the spin
states.

One could imagine that for the $|↓↓⟩$ and $|↑↑⟩$, the center-of-mass mode would be
excited due to a force being identical on the two ions. However, as the the frequency
of the Raman beams equals the stretch mode and not the center-of-mass mode, energy
conservation forbids that this mode becomes excited even though the force is susceptible
to do so.

Finally note that, by adjusting the spacing between ions and thereby the light field
appropriately, we can also prevent the gate from happening by nulling the differential
force, even though the combination of spin states and the frequency difference of the
Raman beams normally would result in an excitation. We will use this trick in the
purification experiment discussed in section 2.2.2.

Not only energy needs to be conserved but momentum as well. This is taken care
of by the geometry of the Raman beams, which are oriented so that the difference
of their wave vectors is along the axial direction of the trap. The recoil of the absorbed
quantum is therefore along this direction. In practical terms, this is achieved by crossing
the Raman beams in the trap center, with both beams at an angle of 45 degrees relative
to the trap axis.

In order to use this differential force for a phase gate, the frequency difference between
the Raman beams is not exactly one quantum of the stretch mode, but slightly higher.
Thereby the stretch mode is gradually excited and de-excited. If the time during which
the light drive force is applied is chosen correctly, one can make exactly one cycle of
excitation – de-excitation and return to the initial state. A good analogy is a swing: if
one pushes a swing slightly faster than the natural swing frequency and in the beginning
pushes it while it is moving away, the motion will gain in amplitude. But keeping the
same frequency, gradually one will start to push the swing while it is approaching, and
one will thereby little by little decrease the amplitude. For the ions, if the laser frequency
difference is $\delta$ higher than the stretch mode, the pulse time should be $\tau = 2\pi/\delta$ in order to
perform one cycle. This corresponds to making one closed circle in the generalized $x-p$
phase-space as illustrated in figure 2.2. Quantum mechanics states that the increased
energy due to the excitation of the stretch mode will add a phase to the wavefunction for
the two ions relative to the unperturbed state. The added phase which we will denote
$\Phi$ equals the area $A$ of the enclosed circle in phase space divided by $\hbar$. The truth table
of the gate is then as follows

$$
|↑↑⟩ \to |↑↑⟩, \quad |↓↓⟩ \to |↓↓⟩, \quad |↓↑⟩ \to e^{i\Phi} |↓↑⟩, \quad |↑↓⟩ \to e^{i\Phi} |↑↓⟩.
$$
Figure 2.2: The geometric phase gate from reference [32] represented in the generalized phase space. The Raman beams drive the stretch mode around a circle in phase space for $|\uparrow\downarrow\rangle$ and $|\downarrow\uparrow\rangle$. These states acquire a phase given by $\phi = A/\hbar$ where $A$ is the enclosed area of the circle. The states $|\uparrow\uparrow\rangle$ and $|\downarrow\downarrow\rangle$ are left untouched. Figure based on earlier version by Didi Leibfried [32].

Note that the spin states themselves are not changed, but acts more like a control bit, as the potential addition of the phase is conditioned on their state. The enclosed area and consequently $\Phi$ is controlled by the laser power. If chosen properly, the phase can be set equal to $\pi$ and thus realizing a $\pi$-phase gate.

This quantum phase gate is a main ingredient in the NIST quantum information experiments. Inserted into a sequence of single qubit rotations, it allows one to create maximally entangled states. An example is a Bell state given by $\frac{1}{\sqrt{2}}(|\downarrow\downarrow\rangle - i|\uparrow\uparrow\rangle)$, which was obtained in reference [32]. In this experiment, the ions are initialized in the vibrational ground state and the $|\downarrow\downarrow\rangle$ spin state by the techniques discussed in chapter 1. A $\pi/2$-pulse, common to both ions, then transforms the states into an equal superposition of all two-ion spin combinations $\frac{1}{2}(|\downarrow\downarrow\rangle + |\uparrow\uparrow\rangle + |\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)$. The phase gate with $\Phi = \pi/2$ is then applied, adding the phases to the $|\downarrow\uparrow\rangle$ and $|\uparrow\downarrow\rangle$ states alone, and the state becomes $\frac{1}{2}(|\downarrow\downarrow\rangle + |\uparrow\uparrow\rangle + e^{i\pi/2}|\uparrow\downarrow\rangle + e^{i\pi/2}|\downarrow\uparrow\rangle)$. Finally, a $3\pi/2$ pulse transforms this state into the desired Bell state. The relative phase between the two components in the entangled state, for this state $-i$, is determined by the relative phase between the two Raman beams. This is due to the fact that the phase of the standing wave pattern, which is responsible for the ion excitation, is determined by relative phase between the Raman beams. How the mentioned operations lead to the claimed transformations resulting in the Bell state $\frac{1}{\sqrt{2}}(|\downarrow\downarrow\rangle - i|\uparrow\uparrow\rangle)$ requires more details than what is given here. One approach consists in using the rotation matrices in the Bloch-sphere formalism applied to each intermediate state in order to visualize the unfolding of the sequence.

The geometrical phase gate can be generalized to work for any number of ion qubits, and will among others be used in section 2.2.1 to create a six-ion Schrödinger-cat state and in section 2.2.2 to perform entanglement purification. For a six-ion Schrödinger-cat
state, the dimension of the Bloch-sphere rotation matrices becomes \(2^6 \times 2^6 = 64 \times 64\) and pulse durations, phases and ion positions are obtained from a numerical approach. This is also true for the purification experiment.

### 2.1 Fidelity and the Error Threshold

As discussed in section 1.2.3 based on resonant fluorescence, we can determine very precisely the populations in \(|\downarrow\rangle\) and \(|\uparrow\rangle\) states. However, to determine the degree of entanglement of a given state, populations are not sufficient. For instance, a statistical mixture of an equal number of \(|\downarrow\rangle\) and \(|\uparrow\rangle\) components generate the same fluorescence as the Bell state \(\frac{1}{\sqrt{2}}(|\downarrow\downarrow\rangle - i |\uparrow\uparrow\rangle)\), so we need another way to quantify the entanglement.

One solution consists in defining the **fidelity** with respect to a desired state \(|\psi_{\text{perfect}}\rangle\) as

\[
F = \left| \langle \psi_{\text{perfect}} | \rho | \psi_{\text{perfect}} \rangle \right|^2, \tag{2.1}
\]

where \(\rho\) is the density matrix for the actual state. In the special case of a pure state, the fidelity can be written as the square of the overlap between the wavefunction of the perfect state and the obtained state \(|\psi_{\text{real}}\rangle\): 

\[
F = \left| \langle \psi_{\text{perfect}} | \psi_{\text{real}} \rangle \right|^2.
\]

For the Bell state, \(\frac{1}{\sqrt{2}}(|\downarrow\downarrow\rangle - i |\uparrow\uparrow\rangle)\), the fidelity given by equation 2.1 can be written as

\[
F_{\text{Bell}} = 1/2 \left( P_{\uparrow\uparrow} + P_{\downarrow\downarrow} \right) + |\rho_{\uparrow\uparrow,\downarrow\downarrow}|,
\]

where \(P_{\uparrow\uparrow}\) and \(P_{\downarrow\downarrow}\) are the probabilities of being in \(|\uparrow\uparrow\rangle\) and \(|\downarrow\downarrow\rangle\) respectively, measured directly by fluorescence, and \(\rho_{\uparrow\uparrow,\downarrow\downarrow}\) is the off-diagonal element of the density matrix, which corresponds to the coherence between \(|\uparrow\uparrow\rangle\) and \(|\downarrow\downarrow\rangle\).

The challenge is then to measure this off-diagonal element. This can be achieved by adding a second gate (a ‘decoding’ gate) in the end of the pulse sequence. The action of the decoding gate can roughly be thought of as rotating the density matrix of the entangled state, such that the off-diagonal elements become diagonal, i.e. the coherences turn into populations, which can therefore be measured directly via the fluorescence rate.

More precisely, the decoding gate substitutes the value of the diagonal element with a quantity that depends on the off-diagonal element:

\[
|\rho_{\downarrow\downarrow,\downarrow\downarrow}|_{\text{new}} = 1/4 \left\{ 1 - 2|\rho_{\uparrow\uparrow,\downarrow\downarrow}|_{\text{old}} \cos(2\theta) \right\}
\]

where \(\theta\) is the phase of the decoding gate (i.e. the phase difference between the Raman beams used for decoding). As this element is diagonal, it can as mentioned be obtained from the fluorescence rate. In practice, we record the fluorescence as a function of \(\theta\) and extract \(|\rho_{\uparrow\uparrow,\downarrow\downarrow}|\) from the amplitude of the oscillating signal. Note that also the decoding sequence contains potential errors, and consequently, the amplitude does not necessarily give the exact value, but a lower bound for the coherence of the state. The fidelity obtained is therefore potentially better than what we measure, but never worse. In the experiment from reference [32], the fidelity of this state was determined to be \(F_{\text{Bell}} = 0.97 \pm 0.02\).

The imperfect fidelity of a state arises not necessarily from the gate or decoding itself but can also come from initial state preparation or single qubit rotations. To distinguish the contributions, it is common to define an error probability per gate operation (or gate
fidelity) which takes into account only the errors arising from the gate. Even though this error rate can be very low for quantum gates with trapped ions, if long sequences are applied, errors will accumulate and become significant. Research concerning error correction protocols therefore plays a prominent role in quantum computing, and receives indeed major attention at the moment. Proof-of-principle quantum error correction experiments based on the use of ancilla qubits were implemented with trapped ions at NIST in 2004 [34].

Based on the work of E. Knill [35] and others, physicists often talk about a fault-tolerance ‘error threshold’. This threshold corresponds to the maximum number of errors one can tolerate per gate operation in order to make reliable computations. Depending on the physical system and the quantum error correction schemes used, people typically quote between $10^{-2}$ and $10^{-4}$ [35,36].

### 2.1.2 Benchmarking of Quantum Operations

To measure an error rate associated with a gate, the fidelity of the obtained state relative to the state which would result from a perfect gate, is measured. One difficulty is that a gate, when inserted into a string of gates, might behave differently, and therefore have a different error rate, depending on the whole sequence. Another difficulty relies in the fact that that the error rates to be measured are very low (ideally below $10^{-4}$). Randomization has been suggested as an effective tool for characterizing quantum noise [37]. In a NIST experiment, we benchmarked a one-qubit $\pi/2$-pulse using randomized benchmarking. The method involves applying a sequence of gates of varying length to a standard initial state. Each sequence is followed by a measurement which determines the fidelity of the obtained state. The average error per operation is obtained from the increase in the probability of having an error in the final state as a function of the sequence length. In this experiment, the error probability for the $\pi/2$-pulse was determined to be $0.00482(17)$ [38]. This is still on the high side of what is considered the fault-tolerance error threshold, but improvements should be conceivable as discussed in the following section.

### 2.1.3 Sources of decoherence

One of the major sources of errors in the quantum gates used for ions is decoherence linked to spontaneous scattering of photons. As discussed in section 1.2.4, we use stimulated Raman transitions via a virtual state near the P state manifold to change the spin states or perform quantum gates. Depending on the detuning between the virtual level and the closest P state, some part of the stimulated scattering into the desired state is replaced by spontaneous scattering [39]. This spontaneous scattering can potentially give rise to decoherence. However, the spontaneous emission rate can be decreased by increasing the detuning. This will decrease the gate speed (i.e. the Rabi flopping rate decreases), but the effect can be counteracted by increasing the laser intensity, depending on available power at the given wavelength [40].

\[\text{Note that contrary to the special case of the observations in reference [39], not only inelastic but also elastic spontaneous scattering can lead to decoherence, see reference [41].}\]
In addition to decoherence caused directly by application of the gates, the quantum states can decohere even when ‘left alone’. For instance, fluctuations in the surrounding magnetic fields can be a source of decoherence. It has been shown in an earlier NIST experiment that this can be overcome by using ‘magnetic-field insensitive qubits’, i.e. two qubit states being separated by an energy which, to first order, is independent of magnetic fields [42]. Another approach successfully implemented at NIST consists in encoding the qubits into the so-called ‘decoherence-free subspace’, spanned by superpositions of simple qubit states, engineered to make them immune to decoherence [7]. Finally, as we will discuss in section 3.4.2, ion heating can also lead to decoherence.

2.2 Entanglement as a Resource for Quantum Computing

The first experiments using entangled states, starting in the 1970, were carried out to prove the non-locality associated with quantum mechanics [43], and culminated in the early 1980s with the experiments by A. Aspect and coworkers. His experiments, using entangled pairs of photons, proved in a very convincing way non-locality by violating the Bell inequalities by 9 standard deviations [44,45].

One major problem with using photons alone for quantum computing is that they do not interact directly with each other. To perform a quantum gate, a conditional action is necessary, which requires some kind of interaction or strong non-linearity. One scheme using photons for quantum computing has been proposed [46], where the non-linear components are beam splitters and single photon detectors. This scheme has the drawback of being probabilistic, and it requires photo detectors of extreme performance in order to be below the fault-tolerance threshold.

Another use of photons for quantum information processing is to associate them with another physical system with which they can interact efficiently. For instance, we can couple them to a cavity containing an emitter, which leads to a strong non-linearity, the topic of part II of this work.

In contrast to photons, ions interact directly and strongly due to their Coulomb repulsion. As we saw in section 2.1, collective motional modes can be excited conditioned on the spin state of the individual ions. Quantum gates based on this interaction can therefore, as we saw, efficiently produce entangled states of ions.

Indeed, for a quantum computer to beat its classical counterpart, quantum states consisting of simple superpositions are insufficient – true gain comes from the use of the entangled states. Particularly powerful states are the so-called cat states, named after the Schrödinger cat [47]. These states consist of a superposition of two ‘maximally different’ quantum states. More generally, any state of the form

$$\frac{1}{\sqrt{2}} \left( |000...0\rangle + |111...1\rangle \right),$$

where 0 and 1 could be for instance the $|\downarrow\rangle$ and the $|\uparrow\rangle$ state of the qubit. For three qubits, the cat state is written as $\frac{1}{\sqrt{2}} \left( |\downarrow\downarrow\downarrow\rangle + |\uparrow\uparrow\uparrow\rangle \right)$, known as a GHZ state, named after D. Greenberger, M. Horne and A. Zeilinger, who showed that this state is particularly powerful for proving the non-locality in quantum mechanics [48]. Cat states are also very interesting for quantum information processing, especially in fault-tolerant quantum computing.
computing and for quantum communication, and some important error correction codes also rely on the use of these states. Finally, the use of cat states have been shown to increase sensitivity in interferometry using trapped ions, and similar experiments using Bose-Einstein condensates have been proposed.

Even though the high degree of entanglement contained in cat states makes them so attractive, it also makes them extremely sensitive to decoherence. If the spin of just one single qubit in the state becomes known, the spins of all qubits become instantly known, as they must necessarily be identical to that of the first qubit. All entanglement is therefore lost. This involuntary ‘readout’ of a qubit can be done by the environment via uncontrolled interactions, or by the loss of the physical particle representing the qubit.

A different class of entangled states is known as the W-states, discussed and compared to the cat states in reference. For \( N \) qubits, this state take the form

\[
\frac{1}{\sqrt{N}} \left( |0...001⟩ + |0...010⟩ + ... + |100...0⟩ \right).
\]

Contrary to the cat states, the W-states are robust to decoherence, as the measurement of one qubit only reveals a partial information about the full state. Therefore, even after read-out or loss of a qubit, some entanglement generally remains. Due to their complex structure, the detection of W-states is more complicated than for cat states. It typically requires quantum tomography, which is a reconstruction of all elements in the density matrix by repeated rotations and measurements (see for instance reference). This is far more time consuming than for the cat state, which only requires one decoding pulse and standard fluorescence measurements.

In 2005, while we created a six-ion cat state at NIST, an eight-ion W-state was created at the same time in the Innsbruck group. The two results were published side by side in the same issue of Nature. Together, these states represented at that time the current state-of-the art for the largest entangled states in any qubit system and have received massive interest from both the media and scientists. The six-ion state will be the subject of the next section.

### 2.2.1 A Six-Ion Schrödinger Cat State

Due to their sensitivity, being able to handle six-ion cat states requires a very high control over the system. Only due to an impressive amount of pre-existing technical know-how, which has been developed and refined over the course of many years in the NIST group, is this control possible. We will in the following give the main lines of how we create and analyze the cat-states, and the experimental results, in particular the fidelity measurements.

We create the entangled states containing from two up to six \(^{9}\text{Be}^+\) ions trapped in a multizone, linear Paul trap. We use axial center-of-mass (COM) frequencies in the range between \( \omega_{\text{COM}} / (2\pi) = 2.9 \) and 3.4 MHz. As in chapter 1, the qubit states are defined by two hyperfine states denoted \( |↓⟩ \) and \( |↑⟩ \). To begin the experiment, all axial motional modes are cooled to the ground state, and all qubits are brought to the \( |↓⟩ \) state by optical pumping, using the techniques described in 1.2.3.
To create the entangled state, we use the geometrical phase gate discussed in section 2.1, but due to the complexity of stretch modes for a six-ion chain, we now act on the more simple COM mode. The principle is the same except that the frequency difference between the Raman beams and the ion spacing are adjusted to fit the COM mode. The phase gate is inserted between two $\pi/2$ rotations, which are common to all qubits. The N-ion cat state obtained is ideally

$$|N \text{ Cat}\rangle = \frac{1}{\sqrt{2}}(|\downarrow, N\rangle + |\uparrow, N\rangle),$$  \hspace{1cm} (2.2)

where $|\downarrow, N\rangle = \prod_{i=1}^{N} |\downarrow\rangle$ and $|\uparrow, N\rangle = \prod_{i=1}^{N} |\uparrow\rangle$.

To determine the degree of entanglement of the obtained cat state, we need the fidelity of the state (section 2.1.1). For the cat state, we can write this as

$$F_{N \text{ Cat}} = \frac{1}{2} (P_{\downarrow N} + P_{\uparrow N}) + |\rho_{\downarrow N, \uparrow N}|,$$  \hspace{1cm} (2.3)

where $\rho_{\downarrow N, \uparrow N}$ is the coherence between $|\downarrow, N\rangle$ and $|\uparrow, N\rangle$. As discussed in section 2.1.1, a lower bound on this quantity is extracted from the amplitude of the fluorescence signal as a function of the phase $\theta$ of the decoding gate.

After decoding, in the ideal case (i.e. no errors during encoding and decoding), the state is transformed into

$$|N \text{ Cat}\rangle_{\text{decoded}} = -i \sin \left( \frac{N}{2} \theta \right) |\downarrow, N\rangle + i^{N+1} \cos \left( \frac{N}{2} \theta \right) |\uparrow, N\rangle.$$  \hspace{1cm} (2.4)

The amplitudes of the $|\downarrow, N\rangle$ and $|\uparrow, N\rangle$ states which can be measured by fluorescence now contain information about the coherences. The fluorescence rate is directly proportional to the probability of being in the $|\downarrow\rangle$ state, which has a component that oscillates as $\cos(N\theta)$. As in section 2.1.1, the amplitude of the oscillating signal is smaller than or equal to twice the value of $|\rho_{\downarrow N, \uparrow N}|$.

The fluorescence signal for four-, five- and six–ion cat states is shown in figure 2.3. First we note that that these signals confirm the fact that the fluorescence should oscillate as $\cos(N\theta)$ with $N$ the number of ions. This confirms that the states do indeed contain four-, five- and six–particle entanglement.

To be more quantitative, from the amplitudes, we obtain a lower bound on the fidelities for the three states: $F_{4 \text{ cat}} \geq 0.76(1)$, $F_{5 \text{ cat}} \geq 0.60(2)$ and $F_{6 \text{ cat}} \geq 0.509(4)$.

For a two-particle state, a fidelity above 0.5 in general proves entanglement. However, it has been shown that the fidelity is not enough to characterize entanglement for any state containing more than two qubits [58]. One approach to confirm the entanglement is based on the so-called entanglement witness operator [59]. If the witness operator has a negative expectation value, the state is definitely $N$-particle entangled.

The entanglement witness operator is particularly simple for a cat state (derived from reference [31]):

$$\langle W \rangle = 1 - 2F_{N \text{ Cat}}.$$
Figure 2.3: Fluorescence for cat states containing 4, 5 and 6 ions (top to bottom) as a function of the phase of the decoding gate, normalized to the equivalent single ion oscillation. The amplitude of the signal obtained by a sinusoidal fit gives a lower bound for the coherence.
2.2 Entanglement as a Resource for Quantum Computing

Therefore, due to the simple relationship between witness operator and fidelity, a measure of the latter is therefore sufficient to claim entanglement in the special case of cat-states: whenever the fidelity is above 0.5 the witness operator is negative, confirming genuine 4-, 5- and 6-ion entanglement in our experiment.

2.2.2 Purification of Entanglement

Most QIP protocols including quantum communications and teleportation require the qubits of an entangled state to be separated in space and each qubit transported to a different location. Often this transport results in decoherence and loss of entanglement. It has been proposed that in order to regain the lost entanglement, one can apply a ‘purification’ protocol [60]. By using quantum operations separately at each location, combined with classical communication between the locations, pairs of not-so-entangled states can be ‘distilled’ into pairs with a higher degree of entanglement. In other words, the entanglement is concentrated into fewer pairs.

Entanglement purification experiments have already been done with photons [61–65], but in these experiments the success rate was low, and the purified pairs were necessarily destroyed at the end of the sequence, and thereby wasted for further use.

In 2006, we performed a purification experiment with trapped ions at NIST. In this experiment, by using two pairs of ions with a non-perfect entanglement, one pair of a higher degree of entanglement was obtained. The success rates of the protocol was between 35 and 65 %, much higher than the former experiments with photons, and the entanglement-boosted pair are not destroyed by the sequence, making it available for subsequent use. We will in the following review this experiment.

The main lines of our protocol follows the proposal from reference [60] but with a few modifications for practical purposes. We start by confining four $^{9}$Be$^{+}$ ions in a linear Paul trap. The qubit states are the same as in the cat experiment described above. As before, we begin by ground state cooling of all axial motional modes and optically pumping all ions to the $|\downarrow\rangle$ state. We then make two pairs of separately entangled ions using the geometrical phase gate on the center-of-mass mode. This step is depicted schematically in figure 2.4 (a). Each pair is a state of the form $\cos(\epsilon)|\downarrow\downarrow\rangle + i\sin(\epsilon)|\uparrow\uparrow\rangle$. The full state can then be written as a tensor product of the two independently entangled pairs

$$|\Psi\rangle = \{\cos(\epsilon)|\downarrow\downarrow\rangle_{12} + i\sin(\epsilon)|\uparrow\uparrow\rangle_{12}\} \otimes \{\cos(\epsilon)|\downarrow\downarrow\rangle_{34} + i\sin(\epsilon)|\uparrow\uparrow\rangle_{34}\},$$

(2.5)

where the subscripts refer to the ion annotation from figure 2.4 (a). Experimentally, the entanglement of both pairs can be created simultaneously with one pulse sequence applied to all ions. This is done by adjusting the spacing between the ions as explained in section 2.1. We adjust the distances between the ions in a way such that both pair 1–2 and 3–4 are subject to the phase gate but such that no entanglement between ions from different pairs exist.

For $\epsilon = \pi/4$ these are Bell states: $|\Psi_{\text{Bell}}\rangle = \frac{1}{\sqrt{2}}\{|\downarrow\downarrow\rangle + i|\uparrow\uparrow\rangle\}$. In the protocol, we need non-perfect states to purify, so we create on purpose ‘non-ideal’ states, i.e. states for which $\epsilon \neq \pi/4$. The goal is then to increase the entanglement of the state consisting of ion 3 and 4, at the expense of the entanglement in pair 1-2.
Figure 2.4: Schematics of the purification protocol. (a) Two pairs of separately entangled ions are created using one phase gate by adjusting the distances between the ions. (b) Ions are entangled gate ‘across’ the pairs. Ion 1 is entangled with ion 3 and ion 2 with ion 4. This is again achieved by adjusting the distance between the ions. (c) A fluorescence measurement is made on pair 1–2. (d) If the outcome of this measurement shows that one of the ions is fluorescing and the other not, the protocol has succeeded and the pair 3–4 has gained in entanglement.

In a second step, figure 2.4(b), we use a ‘crossed’ phase gate entangling each component of the first entangled pair to its counterpart in the second pair. This is also achieved by adapting the distance between the ions: we shift the ions such that the distance between ion 1 and 3 (and between ion 2 and 4) makes them subject to the phase gate. The original protocol suggests that qubit 2 and 3 are interchanged [60], but this is technically difficult in the trap we use, this is why we use instead the ‘crossed’ phase gate. This gate is the actual purification step. This time, contrary to when we prepared the initial states, we do an ‘ideal’ gate (that is, we do not introduce imperfections deliberately). We obtain (ideally) the state

\[ |\Psi\rangle = \frac{1}{2} \left\{ |\downarrow\uparrow\rangle_{12} |\downarrow\uparrow\rangle_{34} + |\uparrow\downarrow\rangle_{12} |\uparrow\downarrow\rangle_{34} + \cos(2\epsilon) (|\downarrow\downarrow\rangle_{12} |\downarrow\downarrow\rangle_{34} + |\uparrow\uparrow\rangle_{12} |\uparrow\uparrow\rangle_{34}) + \sin(2\epsilon) (|\downarrow\uparrow\rangle_{12} |\downarrow\uparrow\rangle_{34} + |\uparrow\downarrow\rangle_{12} |\uparrow\downarrow\rangle_{34}) \right\} \]

where the appearance of \(\epsilon\), which is potentially different from \(\pi/4\), comes from the initial state defined in equation 2.5. In order to gain entanglement for the state corresponding to qubit 3 and 4, upon measurement, the qubits 1 and 2 must be found in either \(|\uparrow\downarrow\rangle\) or \(|\downarrow\uparrow\rangle\). More precisely, if we find the first pair in the state \(|\uparrow\downarrow\rangle_{12}\), then ion 3 and 4 are ideally projected into the state

\[ |\Psi_{34 \text{ purif}}\rangle = \frac{1}{\sqrt{1 + \sin^2(2\epsilon)}} \times (|\uparrow\downarrow\rangle_{34} + \sin(2\epsilon) |\downarrow\uparrow\rangle_{34}) , \]  

(2.6)

whereas if we find the first pair in the state \(|\downarrow\uparrow\rangle_{12}\)

\[ |\Psi_{34 \text{ purif}}\rangle = \frac{1}{\sqrt{1 + \sin^2(2\epsilon)}} \times (|\downarrow\uparrow\rangle_{34} + \sin(2\epsilon) |\uparrow\downarrow\rangle_{34}) . \]  

(2.7)
In both cases, the fidelity of the (ideal) purified state of ion 3 and 4 relative to the maximally entangled state $|\Psi_{\text{max}}\rangle = \frac{1}{\sqrt{2}} \{|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle\}$ can be written as

$$F_{34 \text{ purif}} = |\langle \Psi_{34 \text{ purif}} | \Psi_{\text{max}} \rangle|^2 = \frac{4 \cos^4(\epsilon - \pi/4)}{3 + \cos[4(\epsilon - \pi/4)]},$$  \hspace{1cm} (2.8)

To appreciate the gain in entanglement, this expression should be compared to the fidelity of the initial state of ion 3 and 4 with respect to a perfect Bell state. This quantity is given by

$$F_{34 \text{ initial}} = |\langle \Psi_{34 \text{ initial}} | \Psi_{\text{Bell}} \rangle|^2 = \frac{1}{2}(\cos(\epsilon) + \sin(\epsilon))^2 = \cos^2(\epsilon - \pi/4),$$  \hspace{1cm} (2.9)

which is smaller than $F_{34 \text{ purif}}$ for $0 \leq \epsilon \leq \pi/2$, except for $\epsilon = \pi/4$ where they are equal.\footnote{This is to be expected as the initial state is already maximally entangled.}

On the other hand, if ion 1 and 2 are found to be in $|\downarrow\downarrow\rangle$ or $|\uparrow\uparrow\rangle$ state, ion 3 and 4 will also be in either $|\downarrow\downarrow\rangle$ or $|\uparrow\uparrow\rangle$, respectively, a state with zero entanglement. Therefore, to know if the protocol has succeeded, in a last step, figure 2.4 (c), we do a fluorescence measurement on ion 1 and 2. This measurement allows us to identify the failures, that is, the cases where we obtain $|\uparrow\uparrow\rangle_{12}$ or $|\downarrow\downarrow\rangle_{12}$. In the opposite case, as illustrated in...
figure 2.4 (d), we know that ion 3 and 4 have gained entanglement if one (and only one) of ion 1 and 2 fluoresces.

The fidelity of the purified pair is then measured. The pair consists of a statistical mixture of the two possible states given in equation 2.6 and 2.7. However, what we need to know is the fidelity relative to $|\Psi_{\text{max}}\rangle = \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)$, which is the same for each component of the mixture. Therefore, the fact that the final state is a statistical mixture, does in this case not require special care.

The data from the experiments are shown in figure 2.5. The fidelities are measured in the same way as in the cat experiment. The purified fidelity is measured after applying the purification protocol, whereas the unpurified fidelity is obtained from an identical reference experiment in which the purification step is omitted. The solid straight line represents the case where the fidelity is unchanged. For unpurified fidelities in the interval between 0.5 and 0.7, the data lie above this line, indicating a net gain in fidelity. The inset shows the expected gain obtained based on equations 2.8 and 2.9. According to this curve, we should expect a gain for any pair with an unpurified fidelity above 0.5. The reason why we do not observe the predicted gain for the points representing unpurified fidelities above 0.75 is due to experimental imperfections. Data corresponding to unpurified fidelities below 0.5 should not be considered, as these represent states with no initial entanglement.

Even though this gain in entanglement is relatively small, as the purified pair is intact and thus available for further use, one can imagine repeating a purification protocol several times. By each time starting with the pair from the previous round, the degree of entanglement can be gradually increased to reach a higher value by what could be called ‘entanglement pumping’. We have not yet implemented this scheme, but it would be interesting to do so.

2.3 Perspectives: The Next Milestone

The results of this chapter illustrate that trapped ions are a very interesting choice for a quantum computing system, and that trapped ions currently set the record for largest entangled states for any qubit system.

Even though an impressive number of quantum protocols have been implemented with trapped ions over the last decade [8], and maybe more than in any other qubit system, they have so far included only a very limited number of ion qubits. The first DiVincenzo criterion states that any quantum computing system should be scalable. For ions, contrary to some of the other proposed systems, there is no fundamental reason for not being so. No fundamental reason – but indeed a few ‘technical’ challenges. Among others, to scale up the system, we need to add qubits to the system, and not in any arbitrary way. In order to keep the ion’s motional modes under control, we cannot just keep adding ions to a given trap. We need a trap architecture allowing for a large number of ions, but which limits the number of motional modes. And even more

---

4The off-diagonal matrix elements corresponding to the states in equation 2.6 and 2.7 are identical, and even though the populations for $|\downarrow\uparrow\rangle$ are $|\uparrow\downarrow\rangle$ different, only the sum of them, which is the same in the case of the two components, enters the expression for the fidelity.
demanding, the trap design must allow for a fabrication procedure which itself is also scalable. Potentially scalable ion traps is the subject of the next chapter.
Chip–Style Ion Microtraps for Quantum Computing

A great pleasure in life is doing what people say you cannot do.

Walter Bagehot (1826-77) English economist, political journalist and critic.

Over the past decades, ion traps have been used not only for quantum computing purposes, but in a variety of physics experiments. The use of ion traps originates from mass spectroscopy [14], and has later been used for high precision atomic [66] and molecular spectroscopy [67, 68] and for extremely high accuracy atomic clocks [69, 70]. In this chapter, we will focus on the use of ion traps for quantum information science, which is a field of growing interest in the physics community [6, 8, 10, 71, 72].

To meet the strict requirements for a quantum computer based on trapped ions, the design of the traps needs very special care. This chapter is mainly devoted to the design and demonstration of an ion trap holding characteristics which seem promising for quantum computing applications [73]. In particular, its novel geometry makes it potentially scalable, and the associated ion heating rate is lower than predictions based on other traps. This first design is indeed a proof-of-principle, but we have also developed a more complex and sophisticated version, which has recently been successfully tested [74]. While the emphasis of this chapter will be on the first model, the last part of this chapter will briefly discuss this ‘second-generation’ trap.

3.1 Towards a Scalable Microtrap

The first requirement for an ion trap is that it does not affect neither the internal state of the ion (such as the spin state) nor the external state (such as motional state) in any uncontrolled, excessive way. This has already been achieved with different types of millimetric scale traps, an example being the trap made from gold coated alumina shown in figure 3.1. A second main concern is that the trap also needs to be ‘scalable’ in the sense that the number of trapped ions should in principle be unlimited, or, for practical purposes, large enough to implement useful algorithms. As discussed in section 1.1.1, the number of motional modes increases with the total number of ions in the harmonic potential. This is why we cannot just add more and more ions; the mode spectrum becomes too dense to resolve. One solution is to separate the trap into different trapping
zones [72]. The different trapping zones would then be interconnected via ‘junctions’. To make a network of parallel trapping zones, these junctions could for instance be ‘X’ shaped.

Though of high performance, earlier trap designs do not seem compatible with an extension towards a large number of trapping zones, and thereby they do not seem to satisfy the scaling criterion with respect to the number of qubits. Earlier designs are typically based on two-layer structures, and their fabrication is based on machining and assembling individual parts, as shown in the right hand side of figure 3.1. Not only the assembling, but also the resolution for laser machining of alumina is insufficient for extending the design to more complex architectures. As we will see, the two layer structure also makes it more challenging to make electrical contact to the different electrodes, as these are less accessible.

Microfabrication on the other hand holds the promise of interconnected trapping zones that would allow the desired scaling to reach the desired level [8,10]. Smaller traps also permit larger electric field gradients at the ions which seems to be needed for ion separation with low heating and for controlled transport through trap junctions [1,75]. More generally, adopting microfabrication, we can also benefit from many of the tools used by the semiconductor industry to ‘scale up’.

### 3.2 A Surface-Electrode Design

To avoid any assembling of individually fabricated parts altogether and use microfabrication techniques alone, a ‘surface-electrode geometry’ seems like a good approach [78]. In this particular geometry, the electrodes are positioned to lie on a single plane with the ion suspended above the plane. A schematics of a surface-electrode design composed of five electrodes is shown in figure 3.2 (b). This is quite different from the usual
design shown in figure 3.2 (a) where the ions are surrounded by the electrodes. In the surface-electrode geometry, the potential minimum is situated above the electrodes while the minimum in the standard trap geometry is situated in between the electrodes, see figure 3.2 (c,d).

Neutral cold atoms have been trapped using ‘surface’ traps (commonly referred to as atom chips) for more than a decade [79], but until this work, ions had not been trapped using this geometry. Simulations showed that it is indeed possible to create a potential minimum above a set of surface electrodes in an ion trap [78], though the trapping potential depth is smaller (see section 1.1). For instance, in our first surface-electrode trap [73] the trap depth is at best 200 meV, approximately a fifth of the depth of earlier traps used in the group.

The electrode design ended up being different from the originally proposed symmetric ‘five-electrode’ design shown in figure 3.2 (b): in a symmetric configuration, the two radial trapping axes are perpendicular and parallel to the trap surface, respectively, as illustrated in figure 3.3 (b). This means that if the Doppler cooling beam is parallel to the surface it will not be able to cool the perpendicular radial motion (see section 1.2.2). On the other hand, in a ‘four wire’ design, figure 3.3 (a), the radial trapping axes are tilted by 45 degrees, allowing a Doppler beam parallel to the surface to cool both radial directions. Alternately, we could have chosen to bring the Doppler laser beam in at a non-zero angle relative to the surface but the beam would then scatter on the electrodes. This scattered light could interfere with the fluorescence detection of the ion. Moreover, when the laser light strikes the exposed dielectrics, these seem to charge up. With this in mind, we chose the four wire design for the surface-electrode trap.

In figure 3.4 (a) we show an image of the central zone of the trap. In contrast to both the original linear Paul trap and the ‘five wire’ design, only one of the DC electrodes is segmented as the resulting axial confinement is sufficient. The advantage of this partial segmentation is that all electrode connections can be made from the sides, such that no ‘through-wafer’ connections are necessary. Moreover, the middle segment of the right hand DC electrode is split into two parts: this allows one to minimize micro motion by adjusting those two voltages independently. Finally, the gaps between the electrodes are never perpendicular to the central axis of the trap but always angled. This is done to minimize the risk of forming short circuits when loading, during which a flow of magnesium atoms are entering the trap center perpendicular to the electrodes, and therefore might fill the gaps with magnesium.

A pseudopotential simulation of the trapping potential is shown in figure 3.4 (b). The plotted potential is a cross section in the trap center, corresponding to the position between the arrows indicated in (a). The cross indicates the direction of the trapping axes and the position of the potential minimum, which is about 40 \( \mu \text{m} \) above the surface, corresponding to the expected position of the ion in the trap.

\[1\] In order to null micro motion, the minimum of the pseudopotential needs to be in exactly the same position as the DC minimum. This can be obtained by adjusting the two parts of the middle electrodes separately.
Figure 3.2: Figure comparing two different trap geometries. In (a) a schematics of the standard ion trap design discussed in section 1.1, and in (b) the surface-electrode design discussed in this chapter. Below each is a false color contour plot of the corresponding radial pseudopotential obtained from simulations (low potential: black, high potential: white).
### 3.3 Fabrication Process and Practical Challenges

Several practical aspects need consideration when building an ion trap.

- To begin, RF losses in the trap’s insulating substrate (the wafer) can degrade the quality factor of the resonator in which the trap is mounted, and can cause ohmic heating of the trap itself [80]. This can be avoided by use of low RF-loss insulator substrates such as quartz or alumina. Contrary to silicon which is routinely used in microfabrication, these are substrates for which the fabrication recipes are less developed, making their implementation less straightforward.

- Another concern is to make the neutral atom flux reach the trapping region without depositing material on insulating substrates, which might short circuit adjacent trap electrodes. One step to avoid this consists in making angled gaps between electrodes in the trap center as mentioned above. This is unfortunately not enough. But by making electrode structures that are tall relative to the width of the gaps provides additional shielding of the substrate. Taller structures are also desirable because they shield the ion from the substrate and vice versa; this avoids that the ion charges up the substrate which in return modifies the trapping potential.
The trap also needs to include low pass RF filters. The RF electrodes have a small capacitive coupling to each control electrode (typically $< 0.1 \text{ pF}$) which result in unwanted RF fields at the ion’s position. This RF potential is set to ground by a low pass RC filter (typically $R = 1 \text{ k}\Omega$ and $C = 1 \text{ nF}$) on each control electrode.

Finally, the trap needs to be operated in a ultra-high vacuum environment. This requires that the trap materials can resist a bake-out, typically around 220 degrees celsius over more than a week. Some metals tends to form unwanted alloys at their interfaces if heated, and some metals cease to adhere to the substrate. Finally, to maintain ultra-high vacuum, we need materials of low vapor pressure. These constraints all limit the choice of materials.

The challenge is then to build tall electrodes structures (several micrometers) using only materials fulling the above requirements. We found gold on polished fused quartz (or silica) to be a suitable choice.

The main steps of the fabrication recipe we developed for this purpose are shown schematically in figure 3.5. To make the electrode structures, we use photo lithography combined with metal evaporation and electro-deposition (electroplating). More precisely, outside the central trapping region, structures do not necessarily have to be thick, and we can therefore make them relatively easily by evaporation. For the central regions, we use electro deposition which allows us to build taller structures.

The steps are the following: a 0.030 $\mu$m titanium adhesion layer and a 0.100 $\mu$m copper seed layer are first evaporatively deposited onto the substrate. This deposition
### 3.4 Testing and Characterizing the Surface-Electrode Trap

We use $^{24}\text{Mg}^+$ for the initial tests of the trap. Due to the shallow well depth of the trapping potential, we use photoionization instead of electron impact for ionizing the atoms (see section 1.2.2). The Doppler cooling beam is applied continuously, while the photoionization beam needs to be applied for only a few seconds to create ions in the trap (once the neutral magnesium flux is established). Ions are detected with a CCD camera by observing fluorescence from the Doppler transition along a direction perpendicular to the laser axis.

---

**Table: Fabrication Steps for the Surface-Electrode Trap**

<table>
<thead>
<tr>
<th>Structure</th>
<th>Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electroplating seed layer</td>
<td>Quartz wafer, Ti/Cu, Ti/Au, Electroplated Au with thick resist pattern, Etch Cu seed layer and Ti adhesion layer, Evaporated Cu (Ti adhesion layer)</td>
</tr>
<tr>
<td>Pads &amp; resistors</td>
<td>Evaporated Au (Ti adhesion layer)</td>
</tr>
<tr>
<td>Trap structure</td>
<td></td>
</tr>
</tbody>
</table>

---

The subsequent etching of the seed layer tends to damage the narrow structures of the resistors by undercutting it, so we prefer to make them directly on the substrate.
to the trap surface, as in the view of figure 3.7. If more than ten ions are loaded into the trap the Coulomb repulsion becomes stronger than the axial trapping potential and a zig-zag shaped string will form (see for example [81]).

### 3.4.1 Trap Oscillation Frequencies

We then measure the oscillation frequencies in every direction of the principal axes for a single ion in the trap. This is done by ‘parametric heating’ [20]. To do so, we apply an oscillating field to a chosen DC electrode. When the frequency of the applied field is equal to one of the motional frequencies, the ion will heat, and we observe as a result of the heating a change in fluorescence rate. Depending on which mode we want to excite, we act on different electrodes. In the following we will refer to the electrode number designation of figure 3.6. To excite the axial mode, we apply the oscillating field to electrode 2, while both transverse modes can be excited using electrode 1. The trap oscillation frequencies depend on the operating voltages (both control and RF) applied to the electrodes. The optimal choice for voltages is determined using numerical solvers (boundary element method, see for instance reference [82]) subject to the constraint that the RF pseudopotential minimum overlaps the null points of the electric field from the static potential, in order to minimize RF micro motion (see for example [83]). For the experiments described here, the static potentials on each control electrode, expressed as a fraction of the potential $V_5$ on electrode 5 are $V_1 = 0.320$, $V_2 = 0.718$, $V_3 = 0.738$, and $V_4 = -0.898$. Table 3.1 summarizes our measurements for different values of $V_5$ and $V_{RF}$, and compares with simulations.
3.4 Testing and Characterizing the Surface-Electrode Trap

Figure 3.7: A false color image of one, two, three, six, and 12 ions confined in the surface-electrode trap ($V_{RF} = 103.5 \text{ V}$ and $V_5 = 2.00 \text{ V}$). The length scale is determined from a separate image of the electrodes whose dimensions are known. The horizontal white bars indicate the separation distance between the ions as predicted from the measured axial oscillation frequency. The ratio between transverse and axial oscillation frequencies makes it energetically favorable for the 12 ion string to break into a zig-zag shape.

### 3.4.2 Heating Rates

Using Doppler followed by Raman sideband cooling we can prepare a trapped ion’s harmonic motion in the ground state with $>99\%$ probability $^{10,84,85}$. This is the initial step in any QIP sequence. But the cooling procedure induces decoherence so we have to stop cooling the ions during the quantum operations. However, due to time dependent stray electric fields in the environment, the ions do not remain in the ground state but gradually gain motional quanta, and this heating can perturb the outcome of quantum operations performed with the ions. One source of heating is ambient electric fields that have a frequency component close to a secular frequency. We expect such fields to arise from the Johnson noise on the electrodes, but often the heating rates observed experimentally in a wide range of ion traps are several orders of magnitude larger than what the Johnson noise can account for $^{10,26,86}$. Instead, so-called ‘patch fields’ have been evoked to explain this excessive heating. These are patches on the electrodes of fluctuating potentials with sizes smaller than the ion-electrode spacing $^{26,87}$. Recent experiments $^{86,88}$ indicate that the heating is thermally activated, which is also consistent with the idea of patch fields $^{87}$.

Instead of announcing heating rates, one typically defines a spectral density of electric field fluctuations, $S_E$, measured in $(\text{V/m})^2/\text{Hz}$, which is independent of trapping frequencies. Figure 3.8 gives $S_E$ as a function of minimum ion-electrode distance $R$, obtained from ion heating measurement in a number of ions traps. The dependence of $S_E$ on the $R$ follows a $R^{-\alpha}$ scaling, where $\alpha \approx 3.5$ $^{26,86}$. Johnson noise (thermal heating) on the other hand scales as $R^{-2}$ $^{10,26}$, a second reason for discarding this as the main cause of ion heating. Heating due to patch-fields, on the other hand, is predicted to scale $R^{-4}$ $^{26,87}$.

For the purpose of ion quantum gates, microtraps are advantageous because their
speed increases as $R$ decreases\footnote{This is true for a fixed value of the voltages applied to the electrodes. In principle, one could also increase the gate speed by increasing the oscillation frequencies via the applied fields, but at best up to the voltage break-down limit.} \cite{10, 32, 72}. But the highly unfavorable dependence on motional heating makes it necessary to keep $R$ relatively large, at least until new understanding allows us to build microtraps going way beyond this scaling law. The goal of the following is not to understand the origin of the heating, but merely to measure the heating rate in the surface-electrode trap. We will use two types of techniques: one that is very simple and based on Doppler cooling alone, and one that is more elaborate, based on sideband cooling. The advantage of the latter is that it allows us to confirm the heating rate measurement very close to the ground state. This is necessary to do, as the heating might not be linear over the full range of motional states. The value for the heating rate near the ground state is particularly important, as the quantum gates are performed in this regime.

### Doppler Recooling Experiments

Due to the absence of nuclear spin in $^{24}\text{Mg}^+$, this ion does not have a hyperfine structure. This makes heating rate measurements via sideband cooling unfeasible (section \ref{sec:sideband}). Instead, we measure the heating rate by doing a ‘Doppler recooling’ experiment: by measuring the dynamics of the fluorescence of a hot ion during recooling, the initial ion temperature can be obtained. Our model which describes the recooling is presented in reference \cite{89}, which also gives the details of this approach. By letting the ion heat up over a fixed time interval and subsequently measuring its temperature, we establish a heating rate of $620 \pm 50$ quanta per second for an axial trapping frequency of 4.02 MHz and an ion-electrode distance of 40 μm \cite{90}. An example of a recooling fluorescence curve after a heat time (delay) of 25 seconds is shown in the inset of figure \ref{fig:3.9}. From

\begin{table}[h]
\centering
\begin{tabular}{c|c|ccc|c}
$V_5$ (V) & $V_{RF}$ (V) & $f_\parallel$ (MHz) & $f_{\perp 1}$ (MHz) & $f_{\perp 2}$ (MHz) & $U_T$ (meV) \\
\hline
5.00 & 103.5 & 2.83 & 15.78 & 17.13 & (Meas.) \\
2.00 & 103.5 & 1.84 & 15.87 & 16.93 & (Meas.) \\
5.00 & 46.2 & 2.77 & 5.28 & 8.29 & (Meas.) \\
\end{tabular}
\caption{Oscillation frequencies from measurements (Meas.) and simulated values (Sim.) for three different potential configurations using the scaling as explained in the text. The axial frequency is denoted $f_\parallel$, while $f_{\perp 1}$ and $f_{\perp 2}$ are the frequencies of the two transverse modes whose axes are indicated by the cross in figure \ref{fig:3.4}. The trap depth is denoted $U_T$. The uncertainties in the experimental values for the frequencies are approximately 0.10 MHz.}
\end{table}
3.4 Testing and Characterizing the Surface-Electrode Trap

Figure 3.8: A log-log plot of the spectral density of electric-field fluctuations $S_E$ inferred from observed ion motional heating rates. Data points show heating measurements in room temperature ion traps observed in different ion species by several research groups. A $1/d^4$ plot seems to follow approximately the observed data, ruling out thermal heating (Johnson noise), which scales as $1/d^2$, as a plausible explanation. Two outliers are identified: our surface-electrode trap, and a semiconductor-material ion trap from the group in Michigan, see section 3.6.

this curve, the ion’s average number of motional quanta $\langle n \rangle$ is deduced. The main plot gives $\langle n \rangle$ as a function of delay time based on 5 different recooling curves. The heating rate is extracted from the slope of the linear fit. The result of our measurement is added on the plot in figure 3.8.

Sideband Transitions

The heating rate measured by recooling is lower than what is to be expected from the empiric ion-electrode distance scaling law in figure 3.8. To confirm this encouraging result close to the ground state, we replaced the $^{24}\text{Mg}^+$ with $^{25}\text{Mg}^+$, allowing us to perform sideband transitions. The principle of the method consists in recording the red and blue sideband for each delay time in order to obtain $\langle n \rangle$, as outlined in section 1.2.4. Details of this specific measurements are given in reference [90]. The data is shown in figure 3.10. Note that the delay times are much shorter than in the recooling experiment, and consequently, we stay much closer to the ground state. In this experiment, we obtain a heating rate of $690 \pm 60$ quanta per second also using an axial trapping frequency of 4.02 MHz [90]. This confirms the results obtained using recooling, and places the heating rate well below the value expected according to the ion-electrode distance scaling law.
3.4 State of the Art for Ion Micro Traps

Most proposals for scalable ion traps involve multiple trapping zones, each containing only a few number of ions each [72]. As discussed earlier, if too many ions are held in the same trapping potential, the motional mode spectrum becomes so dense that individual addressing of each mode becomes extremely difficult. Instead, a few ions in parallel trap arrays seems like the best choice for adding qubits to the system. This approach
3.5 State of the Art for Ion Micro Traps

Figure 3.11: The second generation surface-electrode trap. This design features 2 linear loading zones (left and right side) and a ring structure consisting of 150 transport/storage/probing zones interconnected by six Y-junctions. The zoom shows a simulation of the radial components of the pseudopotential across a Y-junction. Trap schematics by Jason Amini and potential calculation by Janus H. Wesenberg.

relies on the ability to transport ions among the different zones in order to make them interact with each other. The transport should be done without excessive heating of the ions. In earlier NIST experiments, ions have been transported in straight lines in linear zones by ramping the voltages on different DC electrodes. This has been done with very little heating, but the real challenge turns out to be the transport through junctions connecting different trapping zones, in particular if this involves transport around corners. The transport through an X-junction with minimal heating has recently been achieved in trap specially designed for this purpose [1]. The low eating rate was due to a very careful consideration for the trap design and the voltage sequence used on the electrodes. However, this particular trap design was based on a three-layer geometry, which is difficult to realize by means of microfabrication techniques. Ion transport is a subject of major importance for all future trap designs; a detailed discussion on this subject can be found in reference [75].

3.5.1 Pushing the Chip Design Further

The surface-electrode trap design presented above contains one single zone and does not allow for controlled ion transport. The challenge is then to combine the advantages of the surface geometry and low heating rates with the possibility of parallel arrays of trapping zones and controlled ion transport comparable to the results in reference [1].

With these goals in mind, we have designed and built a second-generation surface-electrode trap [74]. The fabrication methods are very similar to the first surface-electrode trap, and it also consists of gold on fused quartz. The trap contains several distinct zones which are interconnected using ‘Y-junctions’. A schematics of the trap is shown figure 3.11 including a zoom on the Y-junction, together with the result of a pseudopotential simulation. The choice of Y-junctions is based on the simulation, which indicate that for the surface geometry, transport through a Y-shaped junction allows for less heating than other junction geometries [91]. Other improvements for this trap design
concern the out-lead wire technology. Whereas in the first generation surface-electrode trap those are gap-welded to the electrodes, they are here part of the microfabrication. More precisely, the metal layers are separated by an insulating oxide layer in which holes, also called vias, are etched. The inner walls of the vias are coated with metal, resulting in electrical connection between the electrodes and the backside of the chip. The electrical connections can then be made from below the chip, in our case achieved by using a pin grid array carrier [74]. Finally, the trap contains a special loading zone with a through-wafer loading slot. Thereby the atom flux enters from the backside, avoiding any contamination of trap electrodes or risk of short circuiting electrodes by deposit on the substrate.

The testing of this trap is still work in progress. So far, $^{24}\text{Mg}^+$ ions have been loaded and successfully transported between different zones of this trap. The heating rate associated with the transport remains to be measured.

3.6 Perspectives: Microtraps in Progress

Other groups worldwide are participating actively in the development of ion microtraps. A 3-layer trap containing a T-junction has been build in the ion trapping group Michigan, but unfortunately the heating associated with the transport over the junction was extremely high (comparable to the trapping potential height) [92]. The same group has made a 2-layer ion trap based on semiconductor materials, but the heating rate turned out to be high [93] (for comparison this heating rate is added to the graph in figure 3.8). The same group has also been investigating surface-electrode geometries.

Short after our surface-electrode trap paper was published, a group at MIT also demonstrated surface-electrode trapping [94,95]. The same group has shown that heating is suppressed in a cryogenic surface-electrode trap [88]. This is not necessarily incompatible with the explanation based on patch fields, as the effect of surface defaults and locally trapped charges might diminish at lower temperatures [87].

Inspired by the results of the NIST group, Sandia National Laboratories have started to work with ion traps [96] and have taken up the surface-electrode design short after our proof-of-principle-demonstration of this geometry.

Also in France, namely in the Paris ion trapping group at the university ‘Diderot’, surface-electrode geometry ion traps are being investigated [97]. Their future measurements of associated heating rates will be of interest to the community.

It is wonderful to see that ion trapping is a field in great expansion, and many groups are currently working on trap designs and fabrication methods. The above examples are just a few out of many. It would be hard to make an exhaustive list, and it would not remain up-to-date for very long anyway. This collective effort to make scalable traps is an important step in the quest for quantum computing with trapped ions.
Part II

Solid State Systems - Toward a Quantum Interface
Part I of this work hopefully managed to convince the reader that quantum computing with trapped ions actually does work for a limited number of qubits, and that the system is potentially scalable. One thing that we swept under the carpet was the question about how to distribute the information, not just between qubits during the quantum calculations, but over long distances. For instance, it is quite clear that we will never be able to transport an ion qubit over several kilometers without loss of the information encoded into the quantum state.

An obvious solution consists in using photons for transporting information. This is currently done in quantum cryptography. Photons can propagate unperturbed over very long distances in free space or in optical fibers. The reason why photons alone might not be good candidates for quantum information processing, is their lack of a mutual interaction necessary for realizing an efficient quantum gate. Moreover, photons are not adequate for storing information.

It therefore seems like any realistic quantum information processing system must be a hybrid system containing one kind of qubits for distributing the quantum information (i.e. photons) and another kind of qubits for executing the calculations and storing the information. These qubits could for instance be trapped ions, or quantum dots, or any other ‘material’ particle or system adequate for the purpose.

The main difficulty in realizing this hybrid system is to develop a reliable ‘quantum interface’ between the two types of qubits. The transfer of the quantum state needs to take place on the single qubit level without any loss of information whatsoever. The necessity for such a quantum interface has been expressed in terms of the DiVincenzo Criteria for a quantum computer. As mentioned in the introduction, it has indeed been suggested that a sixth criterion should indeed be added, namely ‘interconversion between stationary and flying qubits’, in order to recognize this additional requirement.

To accomplish this mapping, one could imagine that a single atom should be able
to absorb a single photon in free space. Based on the resonant free space absorption cross section for this process, given by $\sigma = \frac{3\lambda^2}{2\pi}$, one might expect a high absorption probability as long as the light beam is focused tightly enough. However, this cross section is only valid for perfect mode matching, i.e. when the incident beam profile corresponds to the dipole emission pattern of the atom [98]. This makes it very delicate to rely on a free space interaction in order to achieve an efficient quantum interface (see reference [99] for a potential exception to this statement for trapped ions).

A possible solution to this problem is to make use of a cavity. As we shall see, the cavity can greatly enhance the light-emitter interaction, and one can reach a regime where the absorption probability of a single photon gets very close to unity. When this happens, the desired mapping can be achieved in a reliable way.

The goal of this chapter is to discuss the physical properties of an emitter coupled to a cavity, with focus on the particular system we have chosen to study, namely a quantum dot inside a semiconductor micropillar cavity.

### 4.1 Cavity Quantum Electrodynamics

In an article published in 1946, Edward Mills Purcell is the first to mention that the presence of an electromagnetic resonator (a cavity) should have a strong effect on the radiative properties of an embedded emitter [100]. Two year later, in the same spirit, H.B. Casimir and D. Polder show that the interaction between a neutral atom and a conducting plane leads to a shift in the atomic energy [101]. In the beginning of the 1980s, D. Kleppner at MIT proposed that it should be possible to observe an inhibition of spontaneous emission by the use of a cavity [102]. In 1985, together with his group, D. Kleppner succeeded in doing the proposed experiment, using Rydberg atoms [103]. In the meantime, also using Rydberg atoms, Serge Haroche and coworkers in Paris performed experiments which convincingly witnessed the Purcell effect [104].

Drawing on these first experiments, the possibility of engineering the radiative characteristics of an emitter using a cavity has opened up a whole new field in physics, under the name ‘Cavity Quantum Electrodynamics’ or Cavity QED (see for instance reference [105] for an introduction). Whereas the above mentioned experiments used unconfined atoms, many experiments now combine the cavity with atom trapping (using for instance optical far-detuned dipole traps [106]). Pioneers of this field include E. Hinds (London), D. Heinzen (Austin, Texas), H. Kimble (Caltech, California), H. Walther and G. Rempe (both at Garching, Germany), see reference [107] for a review.

It is common to define different regimes according to the coupling strength between emitter and cavity. Along these lines, *strong coupling* corresponds to the regime where the coupling between the cavity and emitter, usually denoted $g$, is much larger than cavity’s coupling to the environment, denoted $\kappa$, and the spontaneous emission rate $\gamma$. Several experiments have been conducted in this regime using neutral atoms, beginning with the experiments in 1983 in the group of Serge Haroche [108].

The opposite case, which we refer to as *weak coupling*, corresponds to $g \ll \kappa, \gamma$, and finally, we call the intermediate regime the *Purcell regime*. This regime corresponds to the coupling governing the experiments described here.

Our goal in part II of this report is to use the Cavity QED via the Purcell effects to
enhance the absorption of a quantum dot embedded in a semiconductor microcavity, as a step in the quest of realizing a reliable quantum interface.

4.1.1 The Purcell Effect

The influence of the cavity is quantified by the so-called Purcell factor \( F \), which can be expressed as a function of the emitter lifetime on- and off resonance with the cavity. If \( \Delta \) denotes the frequency difference (the detuning) between the emitter and the cavity, the Purcell effect can be written as

\[
F = \frac{\tau(\Delta = \infty)}{\tau(\Delta = 0)} - 1,
\]

(4.1)

For an emitter perfectly coupled to the cavity, the Purcell-factor only depends on the cavity parameters and takes on the value denoted \( F_P \) which is given by

\[
F_P = \frac{3}{4\pi^2} \left(\frac{\lambda}{n}\right)^3 \frac{Q}{V}.
\]

(4.2)

where \( \lambda \) is the wavelength for the given transition, \( n \) the refractive index of the cavity, \( Q \) is the cavity quality factor (\( Q = \lambda/\Delta\lambda \)), and \( V \) the mode volume. When we say perfectly coupled we mean that the detuning between the emitter and the cavity is zero, that the emitter is located at the cavity field anti-node, that its linewidth is much smaller than that of the cavity, and that the polarization is the same as that of the cavity mode. The value of \( F_P \) therefore corresponds to the maximum value of the ‘observed’ Purcell factor given in equation 4.1. From equation 4.2, we see that it is important to have a high quality factor and a small mode volume in order to achieve a significant enhancement of the light-emitter coupling. Semiconductor micropillar cavities and photonic crystal cavities allow for very small mode volumes (\( V \approx (\lambda/n)^3 \)) combined with relatively high quality factors (\( Q \approx 10^5 \)). We will discuss the micropillar cavity in more details in section 4.4.

Note that both equation 4.1 and equation 4.2 give the Purcell factor at resonance. It is possible to define an ‘enhancement factor’ which is a function of detuning, by multiplying the resonant Purcell factor with the empty cavity lineshape \( L(\Delta) = 1/(1 + \Delta^2/\kappa_0^2) \) (a Lorentzian of width \( \kappa_0 \)). The enhancement factor corresponding to the the two expressions is then given by \( FL(\Delta) \) and \( F_P L(\Delta) \), respectively.

4.2 A One-Dimensional System

If the emitter’s coupling with the cavity is sufficiently strong, the emission into the cavity mode will be largely dominating over emission into the remainder of the \( 4\pi \) solid angle, and the system exhibits a ‘one-dimensional’ behavior. This idea has been proposed in the group of H. J. Kimble in 1995 [110, 111]. The effect was first demonstrated experimentally in 1999 using Rydberg atoms in a high \( Q \) microwave cavity [112], and in 2007 using a quantum dot inserted into a photonic crystal [113]. Very recently it has also been achieved with a system of a quantum dot combined with a special type of a micropillar cavity [114].
This one-dimensional behavior can be used to reach a regime where the absorption probability of a single photon can get very close to unity for a given emitter. This is exactly what is needed for a quantum interface: the presence of the photon is mapped onto the excited state of the emitter. It is also possible to map the polarization of the photon onto the excited state; as some transitions can only be excited using a particular polarization, we can excite the emitter conditioned on the photon polarization.

It is possible to push this system further and observe a ‘giant non-linearity’. This has been discussed with focus on the case of a quantum dot in a micropillar in reference [115]. The idea is illustrated in figure 4.1. An emitter in the ground state will reflect a weak field (less than one photon per emitter lifetime), whereas an emitter in the excited state will transmit the field. The reflection is due to destructive interference effects between the incident and re-emitted field. This system has even been compared to a CNOT gate: the photon transmission is conditioned upon the state of the emitter, which therefore plays the role of the control bit. We will not push the analogy further here, but just retain the possibility of interacting with an emitter on the single photon level.

### 4.3 Quantum Dots as Artificial Atoms

A quantum dot is an interesting candidate for the choice of emitter. Quantum dots consist of nano-scale crystals made from a certain class of semi-conductor materials. These materials are composed of chemical elements typically from the periodic groups II-VI or III-V. The quantum dot consists of a low band gap material surrounded by a high band gap material, confining the electrons to the low band gap material in three spatial dimensions. The spectrum of the electrons confined to the low band gap material therefore consists of discrete levels. A schematics of the energy levels associated with a quantum dot is given in figure 4.2. Because of the obtained electron confinement together with the level structure, quantum dots are often referred to as artificial atoms.

The discrete energy levels can be divided into two groups: one that corresponds to the valence band, and one that corresponds to the conduction band. In the ground state,
4.3 Quantum Dots as Artificial Atoms

Figure 4.2: Schematics of a quantum dot. Left hand side illustrates the idea of non-resonant excitation followed by photoluminescence. Right hand side depicts our particular quantum dot, namely InAs embedded in GaAs, and indicate the notion of an exciton. The bandgaps at 4 K are $E_{g_{\text{InAs}}} = 0.415$ eV and $E_{g_{\text{GaAs}}} = 1.51$ eV, respectively. Figure by Mathieu Munsch [109].

the levels in the valence band are occupied by electrons whereas the conduction band is empty. When an electron is excited to a level in the conduction band, leaving a ‘hole’ in the valence band behind, we say we have created an exciton. When two electrons are excited, we obtain a biexciton, which energy is shifted relative to that of the exciton, due to the altered Coulomb interactions [116].

Electron-hole pairs can be created by non-resonant pumping, that is, by using a laser with an energy higher than the band gap of the surrounding material, as illustrated in figure 4.2. However, in the experiments described in chapter 5, we typically pump the quantum dot in the wetting layer in order to minimize absorption by the Bragg mirrors. Then, via non-radiative processes, the electron relaxes into the lowest energy levels of the conduction band, and the exciton is obtained. It is also possible to pump the system resonantly [117–119] or quasi-resonantly (phonon-assisted excitation) but we will use only non-resonant pumping here. The exciton decays by emitting a photon which can then be detected. This is referred to as photoluminescence, and is a standard way of characterizing quantum dots.

In our experiments, we work with self-assembled quantum dots, obtained by molecular beam epitaxy (MBE). More precisely, we use InAs inserted into GaAs (from the III-V group) as indicated in figure 4.2. The wetting layer corresponds to an energy slightly below the GaAs bandgap (1.51 eV) which corresponds to a wavelength slightly above 817 nm. The photons emitted are in the near infra-red part of the spectrum (900 nm).

---

1 The wetting layer is the initial layer of atoms (in our case InAs) epitaxially grown on a surface, from which the self-assembled quantum dots are created. This layer is energetically situated a little below the large band gap.
Figure 4.3: Scanning Electron Microscopy (SEM) image of a micropillar cavity, obtained at ‘Laboratoire de Photonique et de Nanostructures’ (LPN). The quantum dot is located in the plane between two Bragg mirrors, consisting of layers of semiconductors. The arrows denoted ‘in’ and ‘out’ indicate the direction of the laser beam used for probing the system and the emission from the cavity.

4.4 Micro Cavities for Quantum Dots

Several types of microcavities are technologically compatible with quantum dots. Among the most successful, one can mention microdiscs [120], micropillars [120] or photonic crystals [121] which, in addition to a small mode volume, can have a very high quality factor.

In the measurements presented in chapter 5, we use the micropillar structures. The pillars are obtained by etching a cylindrical structure into a planar cavity based on Bragg mirrors. Bragg mirrors consist of multiple layers of alternating materials with varying refractive index, confining the modes longitudinally. The planar cavity using Bragg mirrors was first introduced in the 1980s [122]. The pillar shape has for our purpose the advantage of a highly directional emission pattern, which allows for an efficient mode matching of the incoming beam (figure 4.3). The resonant cavity modes for a micropillar can be determined from the model described in reference [123]. As indicated in figure 4.3, it is possible to insert a quantum dot into the micropillar. The quantum dot can be brought on resonance with the cavity by applying an electric field to the sample [114,124] or, as done in our experiments, by varying its temperature.
4.4 Micro Cavities for Quantum Dots

Micropillar cavities containing quantum dots have been successfully used by Jean-Michel Gerard and coworkers for making highly efficient single-photon sources [125]. The efficiency of a single-photon source depends on the probability for a photon to be emitted into the cavity mode, which can be very high for this type of cavity. As a high probability of emission into the mode also means a high emitter absorption probability, the efficiency of single photon sources suggests that this system also is appropriate for mapping photons onto quantum dots.

In the following chapter, we will characterize a particular micropillar-quantum dot system with focus on different types of measurements of the Purcell factor, i.e. the figure of merit for the cavity-emitter interaction.
Measurements of the Purcell Effect in a Semiconductor Microcavity

No effect that requires more than 10 percent accuracy in measurement is worth investigating.

Walther Nernst (1864-1941), physicist and chemist.

In this chapter we will present our experimental setup aiming at realizing a ‘quantum interface’. As described in chapter 4, this interface requires a strong light-matter interaction, which can be achieved by the use of a cavity. Even though the strong coupling regime has been reached in several experiments worldwide (see section 4.1) many applications can be carried out in the intermediate coupling regime, i.e. the Purcell regime. This is also the case for the quantum interface: it has been shown that this coupling is sufficient to reach a unity absorption probability on the single photon level [115].

In this intermediate regime, the signature of the coupling is a reduced lifetime of the emitter and a funneling of the spontaneous emission into the cavity mode. The Purcell factor is a figure of merit for quantifying these effects. As a first step in realizing the quantum interface, it is therefore important to develop reliable methods to measure accurately the Purcell factor in solid state systems. Indeed, the first precise measurements of the Purcell effect in atomic systems go back to the 1980s [104], but the environment in solid state systems adds complexity to the experiment, making the interpretation of the measurements less straightforward. We will in this chapter present different methods for measuring the Purcell factor in solid state systems, based on a data analysis which includes appropriately the complex environment.

5.1 Measurement of the Purcell Factor

Two main types of measurements are possible. The first method is in principle the most simple one. It consists in comparing the lifetime of a quantum dot at and far from resonance with the cavity mode. The lifetimes are measured using a time-resolved setup [126]. Inserting the obtained lifetimes into equation 4.1, the Purcell factor can be derived. We will apply this method in section 5.2.3.
The time-resolved method only works as long as the quantum dot lifetime is longer than the time resolution of the detector, or more generally, longer than any other time scales involved, such as the exciton creation time (capture and relaxation of electrons and holes inside the quantum dot). For very large Purcell factors, the lifetime of the quantum dot is significantly shortened, and might consequently be too short to be directly measured.

Instead, the Purcell effect can be estimated from measurements under continuous-wave (CW) excitation [120]. This can be understood as follows: when the frequency of the quantum dot approaches the cavity resonance frequency, the pump rate required to saturate the emission of the quantum dot is higher due to the shortening of the exciton lifetime. This also results in an increase of the spontaneous emission rate. Both the increase in pump rate required to reach saturation or the emission rate at saturation can be measured in order to determine the Purcell factor. We will do this in section 5.2.5.

In addition to these dynamical effects, the Purcell effect also produces a preferential funneling of the spontaneous emission of the quantum dot into the cavity mode. As we detect only the photons emitted into the cavity mode, this effect increases the photon collection efficiency. This redirection of the light into the cavity mode is therefore partially responsible for the observed increase in the spontaneous emission rate. The strength of this geometric effect is also related to the Purcell effect, and therefore provide yet another method for measuring the Purcell factor. This will also be done in section 5.2.5.

5.2 Micropillar Samples and Apparatus

Our samples are fabricated at ‘Laboratoire de Photonique et de Nanostructures’ (LPN) in Marcoussis near Paris. A sample consists of a layer of InAs/GaAs self-assembled QDs, grown by molecular beam epitaxy, located at the center of a λ-microcavity. The cavity mirrors consist of two planar Bragg mirrors, made of alternating layers of Al\(_{0.1}\)Ga\(_{0.9}\)As and Al\(_{0.95}\)Ga\(_{0.05}\)As. The top mirror contains 28 pairs of these layers, while the bottom layer contains 32. The quality factor of the planar cavity (i.e. before etching into the cylindrical pillar shape) is 14000. The planar cavity is then etched to form a micropillar with a diameter of a few micrometers (2.3 \(\mu\)m for our specific sample). The density of the quantum dots in the plane of the cavity is approximately \(2.5 \times 10^9\) QDs/cm\(^2\).

The etching of the Bragg mirrors into a micropillar can deteriorate the quality factor of the cavity. To measure the final micropillar quality factor, we perform a photoluminescence measurement at high pump laser power, such that the ensemble of quantum dots acts as a spectrally broad light source, which is used for probing the cavity [125][127]. From this measurement, we extract a quality factor of our specific 2.3 \(\mu\)m diameter sample mentioned above of \(Q = \nu/\Delta \nu = 7500\), see reference [109] for details on this measurement. We will in the following use this value for \(Q\) when we calculate the ‘bare’ cavity linewidth \(\kappa_0 = \nu/Q\). Finally, using equation 4.2 the maximum value for the Purcell Factor becomes \(F_P = 18.6\).
5.2 Micropillar Samples and Apparatus

Figure 5.1: The setup used for optical studies of the micropillar-quantum dot system. To excite the sample, depending on the type of experiment, we use either a pulsed (a mode-locked Titanium-Sapphire) or a continuous-wave laser (a laser diode). For detection, we use a monochromator followed by an avalanche photodiode or a CCD camera. Figure by Mathieu Munsch [109].

5.2.1 Photoluminescence Setup

A schematics of the full setup is shown in figure 5.1. Our sample is located in a cryostat held at 4 K. For the CW measurements, the quantum dots are excited using a standard laser diode emitting at 825 nm, while for the time-resolved measurements, we use a pulsed Ti:Sa laser centered at 825 nm (80 MHz repetition rate and 1 ps pulse width). In both cases, this corresponds to a non-resonant excitation (see section 4.3). Within our pump power range, for InAs/GaAs quantum dots, the capture and relaxation time of the charge carriers is less than 50 ps [126]. For our typical quantum dot-cavity coupling, this is always faster than the radiative decay and consequently we do not take it into account in the data analysis. The emitted light is recollected after passing through a spectrometer (0.03 nm resolution). The spectrometer has two output channels: one channel leads to a CCD camera used for the CW measurements, the other to an avalanche photodiode (APD) for time-resolved measurements. The overall time resolution is 80 ps.
Figure 5.2: The full spectrum recorded at 4 K showing the inhomogeneous line, with a zoom on the section of interest including two isolated quantum dots ($X_a$ and $X_b$) and their respective biexcitons ($XX_a$ and $XX_b$), and the cavity mode ($C$). Our Purcell factor measurements are concerned with $X_a$ alone. The data in red is recorded at low pump power, while the data in gray is for high pump power. For all wavelength below 922 nm, the intensity is multiplied by a factor of three.

### 5.2.2 Sample Characteristics

To give an overview of the photoluminescence pattern arising from the particular micropillar that we will study more closely in the following, we show in figure 5.2 the entire inhomogeneous line, which includes hundreds of quantum dots, centered around 895 nm. The inset in figure 5.2 shows a zoom on the region of interest. The micropillar has been etched such that the cavity resonance (indicated by C in the inset) is located on the low energy wing of this inhomogeneous line, where the quantum dot density is spectrally very low. This allows us to optically isolate the single quantum dot (denoted $X_a$) that we want to scan through the cavity resonance. We also note that its corresponding biexciton ($XX_a$) is blue shifted by about 1 nm. This is much larger than the cavity linewidth, and for a given temperature, we can therefore keep the biexciton off resonance with the cavity, while scanning the exciton through resonance. For this specific micropillar, the quantum dot $X_a$ is at resonance with the cavity at 19.5 K. At this temperature, a second quantum dot ($X_b$) appears about 3 cavity linewidths away (with its biexciton $XX_b$ even further away), and is therefore also minimally affected by the cavity. All other quantum dots are much further detuned.

In fig. 5.3 we show the temperature dependence of the cavity resonance wavelength, as well as the quantum dot emission wavelengths for $X_a$ and $X_b$. The cavity frequency
varies linearly with temperature due to the temperature-dependent refractive index (and a smaller contribution due to thermal expansion/contraction of the cavity). The quantum dot exciton energies follow the expected quadratic temperature dependence of the semiconductor bandgap \[128,130\]. Due to this difference in temperature dependence, we can vary the quantum dot-cavity detuning \[131,135\] and thereby put the quantum dot on or off resonance, or scan it across the profile of the cavity. In our experiments, we will focus only on the quantum dot denoted \(X_a\) and its biexciton \(XX_a\) and we will in the following omit the subscript \(a\).

5.2.3 Time-resolved Measurements

To determine the lifetime of the quantum dot, we use our time-resolved setup. We use a narrow window centered on our particular quantum dot to eliminate light from other quantum dots. We start by placing the quantum dot off resonance, and record the photoluminescence decay rate. We use a model which includes an exciton and a biexciton decay. In fact, even though we only detect the light corresponding to the exciton wavelength, we observe a component of the decay curve with the biexciton lifetime. This is because the exciton is partially populated \textit{via} decaying biexcitons. Solving the rate equations for the corresponding three-level system, we obtain a dual
time constant exponential decay of the exciton level, with time constants $\tau_x$ and $\tau_{xx}$. The fit gives the following values

$$\tau_x(\Delta \sim \infty) = 0.80 \pm 0.05 \text{ ns} \quad \text{and} \quad \tau_{xx}(\Delta \sim \infty) = 0.40 \pm 0.02 \text{ ns}.$$  \hspace{1cm} (5.1)

The value for $\tau_{xx}$ is confirmed by an independent experiment in which the window is centered on the biexciton line, and the fluorescence signal fitted to a mono-exponentially decaying curve.

The obtained biexciton lifetime $\tau_{xx}$ can be used as a fixed parameter when we fit the data for the resonant case. As mentioned, due to their mutual energy difference, even when the exciton is scanned through resonance, the biexciton remains sufficiently detuned not to be influenced by the Purcell effect and its lifetime thus remains constant. In all of what follows, we therefore assume $\tau_{xx}$ to take on the value given in equation 5.1.

Despite the fact that we can use $\tau_{xx}$ as a fixed parameter, the resonant case is less straightforward to analyze. Based on our observations, we realize that even far detuned quantum dots tend to emit light into the cavity mode. Recent experiments worldwide confirm this observation\cite{135,139}. Theoretical studies have attributed this phenomenon to different causes, including power-broadening, phonon relaxation \cite{140}, multi-exciton complexes \cite{137}, pure dephasing \cite{141} and charges in the neighborhood of the quantum dot \cite{142}.

At resonance, where our particular quantum dot is spectrally overlapping with the
cavity, the spectral window centered on the quantum dot will necessarily include a part of this contribution as well. We have plotted the decay curve for three different pump powers in figure 5.4. Based on this figure, we observe that the additional emission from detuned emitters increases at higher powers. More precisely, for high pump powers, we observe a quasi mono-exponential decay, due to the fact that the (unwanted) collective contribution from the detuned emitters dominates over the emission of our exciton. As shown by the lowest lying curve in shown in figure 5.4, when decreasing the pump power, we recover the bi-exponential decay, which reflects the Purcell shortened exciton lifetime, and the unmodified biexciton lifetime. However, even in this power range, the off-resonant emission is still present.

Therefore, to obtain the correct lifetime, we need to include these detuned emitters in our model. Within our pumping ranges, this can be done by adding a two-level system to our model, which lifetime corresponds to the ‘average’ lifetime of the emitters contributing to the signal. This lifetime can be measured in an independent experiment, in which all quantum dots are far detuned. We obtain \(0.8 \pm 0.05\) ns. Using this value, together with the biexciton lifetime obtained above as fixed parameters, we can now fit the decay curves at resonance, with the Purcell enhanced exciton lifetime as the only free parameter (see reference [109] for the full expressions used for fitting), and we obtain \(\tau_x(\Delta \sim 0) = 0.20 \pm 0.01\) ns. The excellent agreement between data and the highly constrained fit in figure 5.4 validates our model.

Finally, using the obtained resonant lifetime together with the off-resonant lifetime given in equation 5.1, we obtain from equation 4.1 a Purcell factor of \(F = 3.0 \pm 0.5\). Based on the above discussion, we remark that the low power condition is a necessary but not sufficient criterion for measuring the correct lifetime. Indeed, even for powers well below saturation, only the complete model fit the data correctly. In figure 5.5 we have plotted the exciton lifetime obtained by measurements equivalent to those in figure 5.4 for different values of the detuning together with a Lorentzian fit. As expected, the

Figure 5.5: Exciton lifetimes measured at low power as a function of quantum dot-cavity detuning. The solid curve is a Lorentzian fit to the data.
exciton lifetime as a function of detuning mimics the profile of the cavity mode.

5.2.4 Modeling the System for Continuous-Wave Measurements

To measure the Purcell factor using CW excitation as we will do in the next section, we need to develop a model describing the steady state populations of the different quantum dot levels inspired by the work in reference [144]. We use a three-level system, which includes a ground state, an exciton and a biexciton as shown in figure 5.6. The quantum dot is pumped non-resonantly by a rate $r$ and decays via a radiative process, either into the cavity or into ‘leaky’ modes.

The rate of photons emitted by the exciton into the cavity is equal to $p_x(\Delta, r)\Gamma(\Delta)$, where $p_x(\Delta, r)$ is the population of the exciton state, and $\Delta$ is the emitter-cavity detuning. The exciton decay rate into leaky modes is $p_x(\Delta, r)\gamma$ while for the biexciton, the corresponding rate is $p_{xx}(\Delta, r)\gamma_{xx}$. We assume $\gamma$ and $\gamma_{xx}$ to be independent of the detuning and identical to that of the bulk material.

At low pump rates, only the exciton state is populated. When the pump rate is
Figure 5.7: The photoluminescence of the quantum dot (illustrated as a triangle) arriving at the detector can be separated into two channels: one part emitted into loss channels \( \gamma \) but redirected to the detector with a probability \( \chi_{\text{leak}} \) and the part emitted into the cavity \( \Gamma(\Delta) \), and detected with a probability \( \chi_{\text{cav}} \). We have shown that \( \chi_{\text{leak}} \ll \chi_{\text{cav}} \) \[143\] and we therefore only consider the direct channel in the following.

increased above a certain level (which we shall call the saturation pump rate) the biexciton starts to become populated too. When this happens, the light intensity emitted at the exciton frequency starts to drop, as photons are partially emitted at the biexciton frequency. Therefore, if we detect only the fluorescence from the decay of the exciton state, we observe a maximum in the light intensity at the saturation pump rate, and we will call this intensity the ‘saturation intensity’.

Within this model, we can then write \( \Gamma(\Delta) = \gamma F L(\Delta) \), where \( F \) is the Purcell factor at zero detuning given by equation \[4.1\] and \( L(\Delta) = 1/(1+\Delta^2/\kappa_0^2) \) is the empty cavity line shape. In this model, when pumping with a rate \( r \), the normalized excitonic population, obtained by solving the rate equations, is then given by

\[
p_x(\Delta, r) = \frac{1}{1 + \frac{r}{\gamma_{21}} + \frac{\gamma + \Gamma(\Delta)}{r}}. \tag{5.2}
\]

The role of the cavity is not only to enhance the cycling rate for the exciton, but also to efficiently funnel the emitted photons into the cavity mode. Provided that the emission pattern of the cavity is directional, which is the case for micropillars, the coupling with a properly positioned detector can be very efficient, whereas the coupling of leaky modes with the detector is very limited. These geometrical efficiencies are indicated as \( \chi_{\text{cav}} \) and \( \chi_{\text{leak}} \) in figure \[5.7\]. We have shown that \( \chi_{\text{cav}}/\chi_{\text{leak}} \approx 15 \) \[109,143\] and we will in the following analysis neglect the photons arriving at the detector via leaky modes. Within this approximation, we detect an intensity

\[
I(\Delta, r) \propto \Gamma(\Delta)p_x(\Delta, r). \tag{5.3}
\]

The three-level model will be used in the data analysis for the CW measurements in the next section.
5.2.5 Continuous-wave Measurements

To determine the Purcell factor using CW excitation, we record the photoluminescence intensity as a function of the pump power. We repeat this measurement for several cavity-quantum dot detunings. Typical data sets are depicted in figure 5.8. As outlined in section 5.1, these data sets can be analyzed using several methods in order to determine the Purcell factor. Details concerning each method can be found in the references [109, 143]. Basically, we can use either the saturation pump rate measurement (i.e. the pump rate required to saturate the quantum dot), the saturation intensity (i.e. the intensity emitted by the quantum dot at saturation) or the effect of funneling, which amounts to doing a comparison between the intensity emitted by the quantum dot at low and high, but constant, pump power. We will discuss these three methods in the following.

1) Saturation pump measurements

We use the saturation pump power $P_{\text{sat}}$ as a function of detuning. This method has already been used for micropillars [120], microdiscs [120] and photonic crystals [121]. The expression can be derived from equation 5.3 by determining the pump rate corresponding...
5.2 Micropillar Samples and Apparatus

Figure 5.9: Normalized saturation intensity for the single quantum dot as a function of detuning. The solid line is a fit to the data corresponding to equation 5.5.

<table>
<thead>
<tr>
<th>$\Delta / \kappa_0$</th>
<th>0.0</th>
<th>0.2</th>
<th>0.4</th>
<th>0.6</th>
<th>0.8</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_{\text{sat}}$</td>
<td>0.0</td>
<td>0.2</td>
<td>0.4</td>
<td>0.6</td>
<td>0.8</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Figure 5.9: Normalized saturation intensity for the single quantum dot as a function of detuning. The solid line is a fit to the data corresponding to equation 5.5.

to the maximum intensity, $r_{\text{sat}}$, which is proportional to $P_{\text{sat}}$. We obtain

$$P_{\text{sat}}(\Delta) \propto \sqrt{1 + F L(\Delta)}.$$  \hspace{1cm} (5.4)

where $F$ is the Purcell factor and $L(\Delta) = 1/(1 + \Delta^2/\kappa_0^2)$ is a Lorentzian of width $\kappa$ the empty cavity line shape. The data are fitted with this expression to obtain the Purcell factor $F = 3.7 \pm 1.0$. The relatively large error is due to the uncertainty of $P_{\text{sat}}$, estimated from the individual fit of the different curves in figure 5.8.

2) Saturation intensity measurements

We here use the photoluminescence intensity at saturation, denoted $I_{\text{sat}}$, as a function of detuning. The expression can again be obtained from equation 5.3

$$I_{\text{sat}}(\Delta) \propto \frac{F L(\Delta)}{1 + \sqrt{2 + 2F L(\Delta)}}.$$ \hspace{1cm} (5.5)

The data, which includes the points from figure 5.8 plus additional data points obtained in the same manner, is shown in figure 5.9. The width of this curve is comparable to the width $\kappa_0$ of the cavity (approximately 30 GHz), much larger than the intrinsic quantum dot linewidth (a few GHz). This is to be expected, as the data is recorded by scanning the temperature, and thereby the spectral position of the quantum dot relative to the cavity, making the quantum dot act as a probe for the cavity profile. From the fit based on equation 5.5, also shown in figure 5.9, we extract a Purcell-factor of $F = 2.4 \pm 1.2$.

Here, the intrinsic uncertainty in the saturation intensity measurement is quite small, but the error on the Purcell factor is amplified by the fitting procedure, resulting in the stated error. Moreover, this error does not include the uncertainty on the measurement of the cavity linewidth $\kappa_0$. The Purcell factor obtained with this method is particularly sensitive to the value of $\kappa_0$. Therefore, if the uncertainty on $\kappa_0$ is large, this method should be avoided.
3) Redirection of the photoluminescence intensity

The last method to be discussed is the less intuitive one. It is based on the fact that the Purcell effect does not only increase the spontaneous emission rate (i.e., it decreases the lifetime) of an emitter, but it has a ‘geometrical’ effect as well. More precisely, photons are funneled more efficiently into the cavity mode. As we detect only the photons emitted into the cavity mode ($\chi_{\text{cav}}/\chi_{\text{leak}} \gg 1$), even without a change in the emission rate, the Purcell effect shows up as an increase in the number of detected photons. In order to separate the two effects, that is, the increase in the spontaneous emission rate and the redirection of light into the cavity, we need to consider two different regimes according to the excitation power:

1. Above saturation, we observe both the effect of the geometrical redirection of the emission pattern and shortening of the quantum dot lifetime due to the coupling with the cavity. As a consequence, the light intensity follows the $\mathcal{L}(\Delta)$ profile of the cavity. More precisely, in the regime well above saturation equation (5.2) can be written as $p_x(\Delta, r) \approx \gamma_{xx}/r$, and we get

$$I(\Delta) \propto \gamma F \mathcal{L}(\Delta) \frac{\gamma_{xx}}{r}. \quad (5.6)$$

The width of the intensity curve is equal to the width of the cavity (only its amplitude depends on $F$).

2. Well below saturation of the quantum dot, we still observe the geometrical redirection, but not the increase in the spontaneous emission rate, which always remains

Figure 5.10: Measurements of the quantum dot intensity at fixed pump power (30 $\mu W$ and 300 $\mu W$, respectively). The two curves can be thought of as the intersection of the curves in fig. 5.8 with vertical lines centered at $P=30 \mu W$ and $P=300 \mu W$ (with several more similar curves added). It is important to notice that the vertical scale is different for the two curves; if they had been plotted using the same vertical scale, they would have had their wings superposed.
5.2 Micropillar Samples and Apparatus

equal to the pump rate. This is because the spontaneous emission rate can never be higher than the pump rate: each photon absorbed can give rise to only one photon spontaneously emitted. As the spontaneous lifetime is shorter than the pump rate in the non-saturated regime, the two will be equal.

The increase in the number of detected photons, when bringing the quantum dot on resonance with the cavity, is therefore due to the redirection of photons alone. Well below saturation we can neglect \(1 + r/\gamma_x\) in equation 5.2 and using this approximation in equation 5.3, we find that

\[
I(\Delta) \propto \frac{rF \mathcal{L}(\Delta)}{1 + F \mathcal{L}(\Delta)}, \tag{5.7}
\]

Inserting the expression for the cavity \(\mathcal{L}(\Delta)\) this can be written as

\[
I(\Delta) \propto \frac{rF}{1 + F} \left\{ 1 + \frac{\Delta^2}{\left(1 + F / \kappa \sqrt{1 + F} \right)^2} \right\}, \tag{5.8}
\]

This function is broader than the Lorentzian profile of the cavity mode, \(\mathcal{L}(\Delta)\), by a factor \(\sqrt{1 + F}\).

This broadening relative to the curve given in equation 5.6 can be understood intuitively based on the following. As the detuning \(\Delta\) approaches zero, the re-direction which is responsible for the increase in detected light intensity becomes more and more prominent. However, as more and more of the light has been re-directed into the cavity, the increase in detected light ‘saturates’ as we cannot observe the ‘dynamical’ component of the Purcell effect, due to pump power limitations. We can now also intuitively see why the width of this curve is related to the Purcell factor: the larger \(F\), the ‘earlier’ (i.e. for larger \(\Delta\)) the majority of the light has been re-directed into the cavity, making the gain in the detected light slow down. This is seen as a saturation in profile near \(\Delta = 0\), i.e. the curve ‘flattens out’.

By taking ratio of the widths for high (equation 5.6) and low (equation 5.8) pump rate measurements, we can extract a value for the Purcell factor. We use again data corresponding to intensity measurement as a function of detuning, but now for a fixed pump power: one case in which the fixed pump power is very low (30 \(\mu W\)) and one for which it is high (300 \(\mu W\)). These two data sets can be thought of as the intersections between a vertical line in the data represented in figure 5.8. When these data are plotted as a function of detuning, we obtain the two curves shown in figure 5.10. From the ratio of the two widths, we extract a Purcell-factor of \(F = 3.2 \pm 0.9\), where the stated uncertainty arises from the intensity measurements, which is the dominant source of error in this case.

\(^3\)As seen in figure 5.8 we do not reach powers high enough to make the approximation in equation 5.6 valid (at higher powers, the signal to noise ratio becomes too low). In the data analysis, we use instead the exact expression for the intensity. We have chosen to write equation 5.6 to simplify the discussion.
5.3 Discussion

The agreement between time-resolved measurements and CW measurement suggests that both methods yield reliable values for the Purcell factor. While the time-resolved measurements are more precise, they are limited to emitters with Purcell-enhanced lifetimes longer than the time resolution of the detector. For high Purcell factors, this can be a limiting factor. Continuous-wave measurements do not have this limitation, and the setup is very simple, as no time-resolved components are required.

The Purcell factor for the particular sample studied in this chapter is of the order of 3, which is 6 times lower than the theoretical value (see equation 4.2), which assumes a perfect spatial and spectral coupling between cavity and emitter. As we can control the spectral position very precisely, our lower value certainly reflects the fact that the spatial coupling of the quantum dot is not perfect, i.e. the quantum dot is not located in the center of the micropillar cavity.

Although being relatively small, this value actually corresponds to the Purcell factor for the to-date most efficient single-photon sources based on micropillar cavities [125]. The equivalence between the emission and absorption probability makes us confident that the system is an adequate candidate for the planned purpose, namely for mapping photons onto a material particle in order to form a quantum interface. A quantitative analysis of the system’s predicted capacity to exhibit this non-linear behavior, including experimental considerations such as signal-to-noise issues, has also been given in [109]. This analysis confirms our intuition, but even more important, preliminary experimental results starts to show the signature of the absorption of a single quantum dot, a very encouraging result for us. These recent results are obtained primarily by Inah Yeo, our Ph.D. student succeeding Mathieu Munsch.
Conclusion and Perspectives

Physicists like to think that all you have to do is say, these are the conditions, now what happens next?


Scaling down to scale up

The goal of part I of this work was to show the reader that cold confined ions represent an interesting choice as the elementary bits for a quantum computer. This system has already been proven compatible with the following DiVincenzo criteria [5]: initialization, long coherence times, a universal set of quantum gates, and a reliable method of read-out for the qubits. The recent progress in ion trap designs indicate that trap architecture should not be a limiting factor for reaching scalable quantum information processing with confined atomic ions.

To be able to scale up the system, defined in this work as the capacity to posses and control an increased number of qubits, we have adapted the ion trap designs accordingly. To keep on increasing the number of qubits without making the traps too bulky, it is necessary to ‘scale down’ the size of the current traps. Microfabrication offers indeed the benefit of being able to produce smaller traps.

However, the trapped ion quantum computer is not just the trap. The vacuum system, several voluminous lasers and space consuming laser optics must accompany the ion trap to make the quantum information processing feasible, and the full equipment in general occupies a large laboratory. Figure 5.11 shows a journalist’s reaction to the surface-ion trap discussed in chapter 3. The expression ‘a large quantum computer’ is unfortunately very close to the reality. Although well-meant by the journalist, it is thus a sharp reminder to us that there is still a lot of work to be done before such a quantum computer can ‘easily be mass produced’! Consequently, in addition to the ion traps, vacuum system, lasers and laser optics and mirrors must also be made ‘scalable’, or to use a more appropriate term, be miniaturized. Scientist and engineers from the MEMS community are collaborating with the ion trap community towards this goal. For instance, J. Kim and coworkers at Duke University, USA, are working towards developing suitable MEMS mirrors [145]. Scientists at Sandia National Laboratories, USA, are working on MEMS-based arrays of ion micro-traps [146].

4Microelectromechanical systems.
New Ion Trap May Lead To Large Quantum Computers

Physicists at the National Institute of Standards and Technology (NIST) have designed and built a novel electromagnetic trap for ions that could be easily mass produced to potentially make quantum computers large enough for practical use. The new trap, described in the June 30 issue of Physical Review Letters, may help scientists surmount what is currently the most significant barrier to building a working quantum computer—scaling up components and processes that have been successfully demonstrated individually.

Figure 5.11: From Science Daily concerning the surface-electrode ion trap discussed in chapter 3 (http://www.sciencedaily.com/releases/2006/07/060707020823.htm).

The ‘Sixth’ DiVincenzo Criterion Revisited

Based on the above, it seems that trapped ions might provide a valuable ground for quantum computation, and definitely worthy of further investigation. However, as discussed earlier, the sixth DiVincenzo criterion, namely ‘interconversion between stationary and flying qubits’ has not yet been proved satisfied for trapped ions.

Part II of this work suggested to increase the photon-emitter interaction via the Purcell effect by placing the emitter inside a cavity. We used a micropillar cavity combined with a quantum dot to explore these ideas. Preliminary results suggest that even for a relatively weak Purcell effect, it should be possible to map the state of a photon onto the state of the quantum dot. This is indeed very promising, all the more so because a new technology for fabricating micropillars has very recently been developed, which leads to higher Purcell factors. More precisely, at the ‘Laboratoire de Photonique et de Nanostructures’ (LPN) in Marcoussis, Pascale Senellart and coworkers can now position a quantum dot within a pillar microcavity with a 50 nm accuracy [147], whereas for our samples, quantum dots were arbitrarily distributed in the plane of the cavity. Moreover, in situ photolithography allows the LPN team to achieve, in addition to this spatial mode matching, a spectral matching between the frequency of the quantum dot and cavity emission. Both lead to an improved emitter-cavity interaction, and in turn to a stronger Purcell effect.

With this in mind, it should be possible to map a single photons onto a single quantum dot in a reliable way. However, one point still needs to be addressed: the exciton might not be suited as a ‘quantum bit’. Its lifetime being extremely short (nanoseconds) it is at least not a good candidate for storing quantum information. Instead of using the exciton, it is possible to use a charged quantum dot, and then use the electron spin. By applying a magnetic field, the degeneracy of the two spin states is lifted, and these states can therefore form the quantum bit. The lifetime of the single electron’s
spin state (i.e. the time it takes for a non-equilibrium spin configuration to relax to the equilibrium state, also referred to as the $T_1$-time) is much longer than the exciton lifetime. The lifetime does therefore not limit the dephasing time (the $T_2$-time), which can then exceed 100 ns \[148\]. This is fairly long compared to the time required to flip the spin of the electron, which is typically of the order of 10 to 100 ps. However, the ratio between $T_2$- and the spin-flip time still remains much smaller than if using the hyperfine state of a cold, confined ion.

**The Ultimate Choice of a Quantum Computing System**

I believe that it is still too early to commit to one single system for a quantum computer, or even to establish the components of a hybrid system. I would not feel comfortable having to choose either cold confined ions or quantum dots, or any other system or hybrid system. Ions are excellent for high-fidelity quantum gates, whereas quantum dots clearly lack behind in this aspect. On the contrary, from a technological view point it might be simpler to associate a quantum dot with a microcavity, even though ion-cavity systems, since the first proposal in 2002 \[149\], also are getting a steadily increasing attention worldwide (see for instance references \[150\]–\[153\] and references therein).

Neutral atoms have also been inserted very successfully into cavities, but they lack the advantage of the straightforward trapping and manipulation of ions, which exist due to their charge. Finally, in another variant of the quantum dot system definitely worth considering, the micropillar cavity is replaced by a photonic crystal. This type of system would then greatly benefit from the recent development of integrated circuits for photonics.

Even though the future should reveal some kind of fundamental limitation bringing one of the systems at a deadlock, or worse, that we ‘have just forgotten something important’, I am sure that this research at no-account would be wasted. The acquired knowledge can be transposed and adapted to a different system, and the resulting technological progress will always contain more universal components equally adequate for a different purpose.

The best way to choose the ideal system is probably to push each one of them to their maximum performance and see how far they will reach. That is why, in my opinion, it is good that many different systems are being investigated in parallel by different groups worldwide. It is also probable that not only one but several systems will prove successful, or that a given system is ideal for one particular purpose, but not for all. For instance, it might well be that one system is adequate for quantum simulations, and one for factorizing prime numbers. I guess that only the future can give us the answer.

**Beyond Quantum Information Processing**

In this report, I have tried to present a coherent story under the common title ‘quantum information processing’. Doing this, I have in a certain way reduced the content to a quest for a quantum computer. However, it is my hope that future research in this direction not only will bring us closer to the realization of such a device, but also might
lead to a deeper understanding of the quantum world. Despite the fact that quantum mechanics has been thoroughly tested, and no ‘errors’ in the theory have been discovered, further insight might still be gained. Even though we are forced to accept that even Nature cannot predict the outcome of a measurement for certain quantum states, it might for instance be possible to obtain further knowledge concerning the mechanism of the ‘collapse’ of the wavefunction upon measurement. This is an interesting question, also related to decoherence, which can be thought of as the result of continuous measurements on the state performed by the environment. Decoherence is a subject of major fundamental interest, as it dictates the limit between the classical and quantum world.

If the single steps on the paths towards a quantum computer do not in themselves lead to a deeper understanding of the quantum world, the quantum computer as an instrument for research certainly will. If we can control and exploit a sufficiently large number of qubits, quantum simulations of physical systems approaching the classical limit will open the door to territory which so far has remained unexplored, theoretically as well as experimentally.

More generally, I hope that the overwhelming technological progress over the last few decades will lead to experiments which are susceptible of addressing fundamental questions. Sure, it would be great to have a quantum computer, but would it not be even greater to find out more about Nature?
APPENDIX A

COMPILATION OF PUBLICATIONS
RELEVANT TO WORK PRESENTED

Towards scalable ion traps for quantum information processing
J. Amini, H. Uys, J. H. Wesenberg, S. Seidelin, J. Britton, J. J. Bollinger, D. Leibfried,
C. Ospelkaus, A. P. Van Devender and D. J. Wineland
NEW JOURNAL OF PHYSICS 12, 033031 (2010)

Continuous-wave versus time-resolved measurements of Purcell factors of quantum dots
in semiconductor microcavities
M. Munsch, A. Mosset, A. Auffèves, S. Seidelin, J. P. Poizat, J.-M. Gérard, A. Lemaître,
I. Sagnes, and P. Senellart
PHYSICAL REVIEW B 80, 115312 (2009)

Randomized benchmarking of quantum gates
E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R.
Ozeri, S. Seidelin, and D. J. Wineland
PHYSICAL REVIEW A 77, 012307 (2008).

Fluorescence during Doppler cooling of a single trapped atom
J. H. Wesenberg, R. J. Epstein, D. Leibfried, R. B. Blakestad, J. Britton, J. P. Home,
W. M. Itano, J. D. Jost, E. Knill, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland
PHYSICAL REVIEW A 76, 053416 (2007).

Simplified motional heating rate measurements of trapped ions
R. J. Epstein, S. Seidelin, D. Leibfried, J. H. Wesenberg, J. J. Bollinger, J. M. Amini,
R. B. Blakestad, J. Britton, J. P. Home, W. M. Itano, J. D. Jost, E. Knill, C. Langer,
R. Ozeri, N. Shiga, and D. J. Wineland
PHYSICAL REVIEW A 76, 033411 (2007).
Errors in trapped-ion quantum gates due to spontaneous photon scattering
R. Ozeri, W. M. Itano, R. B. Blakestad, J. Britton, J. Chiaverini, J. D. Jost, C. Langer,
D. Leibfried, R. Reichle, S. Seidelin, J. H. Wesenberg, and D. J. Wineland
PHYSICAL REVIEW A 75, 042329 (2007).

Experimental purification of two-atom entanglement
Ozeri, S. Seidelin, D. J. Wineland

Transport dynamics of single ions in segmented microstructured Paul trap arrays
R. Reichle, D. Leibfried, R. Blakestad, J. Britton, J. Jost, E. Knill, C. Langer, R. Ozeri,
S. Seidelin, and D. J. Wineland
FORTSCHRIT DER PHYSIK 54, 666 (2006).

Microfabricated surface-electrode ion trap for scalable quantum information processing
S. Seidelin, J. Chiaverini, R. Reichle, J. J. Bollinger, D. Leibfried, J. Britton, J. H.
Wesenberg, R. B. Blakestad, R. J. Epstein, D. B. Hume, W. M. Itano, J. D. Jost, C.
Langer, R. Ozeri, N. Shiga, and D. J.Wineland

Creation of a six-atom “Schrödinger cat” state
D. Leibfried, E. Knill, S. Seidelin, J. Britton, R. B. Blakestad, J. Chiaverini, D.B. Hume,
W.M. Itano, J.D. Jost, C. Langer, R. Ozeri, R. Reichle and D. J. Wineland
NATURE 438, 639 (2005).
Toward scalable ion traps for quantum information processing

J M Amini1,5, H Uys2, J H Wesenberg3, S Seidelin4, J Britton, J J Bollinger, D Leibfried, C Ospelkaus, A P VanDevender and D J Wineland

National Institute of Standards and Technology, Time and Frequency Division, 325 Broadway, Boulder, CO 80305, USA
E-mail: Jason.Amini@gtri.gatech.edu

Received 20 August 2009
Published 16 March 2010
Online at http://www.njp.org/
doi:10.1088/1367-2630/12/3/033031

Abstract. In this paper, we report the design, fabrication and preliminary testing of a 150 zone ion trap array built in a ‘surface-electrode’ geometry micro-fabricated on a single substrate. We demonstrate the transport of atomic ions between the legs of a ‘Y’-type junction and measure the in-situ heating rates for the ions. The trap design demonstrates the use of a basic component design library that can be quickly assembled to form structures optimized for a particular experiment.

Online supplementary data available from stacks.iop.org/NJP/12/033031/mmedia

1 Present address: Georgia Tech Quantum Institute, GTRI/STL, CRB Building, 400 10th Street NW, Atlanta, GA 30318, USA.
2 Present address: Council for Scientific and Industrial Research, Meiring Naudé Road, Brummeria, Pretoria, South Africa.
3 Present address: Centre for Quantum Technologies, National University of Singapore, 117543 Singapore.
4 Present address: Institut Neel-CNRS, BP 166, 25, rue des Martyrs, 38042 Grenoble Cedex 9, France.
5 Author to whom any correspondence should be addressed.
1. Introduction

The basic components of a quantum information processor using trapped ions have been demonstrated in a number of experiments [1–3]. To perform complex algorithms that are not tractable with classical computers, these components need to be integrated and scaled to larger numbers of quantum bits (qubits). Both integration and scaling can be achieved by making trap arrays with many zones. In one possible scheme, information is shared between zones by physically transporting the ions between trapping zones that have various specialized functions such as detection, storage and logic gates [4–6]. We report here on the design, fabrication and preliminary testing of a large array built in a ‘surface-electrode’ geometry [7, 8] and report the first transport of atomic ions through a surface-electrode trap junction. Transport of ions through a junction has been demonstrated previously in multi-layer (three-dimensional) trap electrode geometries [9, 10]. The surface-electrode trap described here is composed of 150 zones and six ‘Y’ type junctions and is in principle scalable to an arbitrarily large number of zones. It demonstrates the use of a basic component design library that can be quickly assembled to form structures designed for a particular experiment or, in the future, a particular algorithm. Microfabricated on a single substrate, the traps are amenable to rapid mass fabrication.

Ion trap design, fabrication and characterization for quantum information experiments can be a difficult and time-consuming process. With a design library and fabrication techniques as demonstrated in this paper, we illustrate the use of pre-designed and pre-characterized, modular components that can be assembled into trap designs for specific experiments. The design library includes components for Y junctions, loading, transport and generic ‘experiment’ regions. We combined six of the Y junctions into a hexagonal ring that includes two loading components and two experiment regions, as shown in figure 1. Fabrication of five of the unmounted traps on a single 76 mm diameter wafer took one week.

The ion trap described in this paper and shown in figures 1(a) and (b) is a planar version (surface-electrode trap) of a radio-frequency (rf) Paul trap [7, 8] that confines ions by combining a ponderomotive potential generated by an rf electric field and static electric potentials, as shown in figure 2. In the configuration shown in figure 1, the ponderomotive potential (also called the pseudopotential) does not form a fully three-dimensional trapping potential. Instead, it forms a
confining tube and static or slowly varying control potentials applied to segmented electrodes confine ions along the axis of the tube(s). By changing the control potentials slowly with respect to the rf period, the ions can be smoothly transported along the tube. For added versatility, the trap includes junctions that combine multiple ponderomotive tubes and allow the ions to switch between multiple paths.

2. Fabrication

The trap fabrication is based on gold-on-quartz structures reported in [7, 8], where a quartz wafer (which has low rf loss) is coated with a patterned gold conducting layer. The monolithic...
Figure 2. Example rf ponderomotive potential contour. For the experimental conditions used when testing the trap (51 V peak rf at 90.7 MHz and trapping $^{24}\text{Mg}^+$), this contour corresponds to 1.3 meV. By changing the potentials on the control electrodes, the ion is moved along the two-dimensional confining tube formed by ponderomotive potential. Junctions join multiple tubes and the control potentials move the ion between the junction’s legs.

construction of surface-electrode traps provides a basis for scalable ion trap structures. However, as the trap complexity increases, distributing the potentials to the control electrodes becomes increasingly difficult. The design of the trap described here could not be realized using a single conducting layer because control electrodes between the rf ‘rails’ form isolated islands surrounded by other control electrodes. To counter this problem, we extended the single conducting layer fabrication used in [8] to multiple conducting layers, as shown in figure 3(a). The top conducting layer forms a nearly continuous conducting plane that shields the ion from the potentials on the second (lower) conducting layer (see [11]).

The fabrication process begins with a 380 $\mu$m thick amorphous quartz wafer. The first conducting layer is 300 nm of evaporated Au with 20 nm Ti adhesion layers deposited on both sides of the Au. These layers are lithographically patterned and etched (wet etched for Au and plasma etched for Ti) to form the interconnects (figure 3). A 1 $\mu$m layer of SiO$_2$ is then deposited by chemical-vapor deposition, forming the insulation between the two layers. The Ti layer deposited on top of the Au acts as an adhesion layer for this oxide. The top conducting layer is 1 $\mu$m of evaporated Au with a 20 nm Ti adhesion layer, patterned using the liftoff technique [12]. Holes etched through the oxide, here denoted as vias, allow electrical connections to be deposited between metal layers. The vias are plasma etched with a process that results in sloped side walls by laterally etching the photoresist at the same rate as the oxide is etched. Because the liftoff technique, used for the top layer of Au, requires a directional flux of Au during the deposition, vias with a vertical side wall would not necessarily have the walls coated with Au, resulting in electrical discontinuities between the two conducting layers.

The trap was loaded by passing a neutral flux of $^{24}\text{Mg}$ through slots in the wafer. To form these slots, tapered channels through the quartz wafer were mechanically drilled from the back side of the wafer to within 30 $\mu$m from the top front surface at locations under the two loading
zones. The remaining 30 µm membrane was then milled using a focused ion beam to form a 12 µm by 20 µm slot in the surface connecting the drilled channel (figure 3(b)).

3. Trap geometry

The trap design, shown in figure 1(c), incorporates the use of a library of patterns that can be connected together to form more complex structures. The core of the trap design is the six Y junctions that are assembled to form a hexagonal ring. Two loading regions at either end
feed ions into the hexagonal ring. Inserted into two legs of the hexagon are components that can combine and separate pairs of ions for entangling and distributing the ions. Except for the loading regions, the outward legs of the hexagon are terminated in this design. However, these legs could be extended in future designs to integrate more hexagonal rings or other components such as memory storage regions.

Junctions between multiple ponderomotive tubes have been demonstrated in larger, multi-layer traps [9, 10], but these designs are not as convenient for scaling as the surface-electrode traps, in part because of the difficulty of alignment and assembly. An ideal junction would produce an rf ponderomotive zero at the axis of the ponderomotive confining tube along all three legs, merging in the junction center. Since such an exact ponderomotive zero is not possible for all points along the confining ponderomotive tube of a Y junction [13], we numerically optimized the shape of the rf rails in the junction (see section 3.2) to minimize the magnitude of the ponderomotive potential along a continuous path to the junction center. An iterative algorithm began with an initial shape (figure 4(a)) to generate the design shown in figure 4(c). Other choices for optimization criteria generate alternative junction geometries that might have benefits over the design that is used in this trap. For example, as shown in [10], efficient transport through a junction is possible even if the junction has large deviations from the ideal of an exact ponderomotive zero.

The control electrodes for the junction are narrow near the junction center for increased spatial control and some electrodes are connected on the surface to allow connections from other parts of the trap to pass under the junction. The final junction design was added to the component library for assembly into the hexagonal ring.

The limited number of available vacuum-feedthrough electrical connections available in our test apparatus required that single potentials be shared among multiple control electrodes. The loading zone and three junctions on the left half of the trap share 12 potentials distributed by the left-side control bus (figure 1(b)). The right-side loading zone and junctions have a similar control bus. The two experimental zones are independent of these buses and independent of each other. In this way, one set of control lines can be used to load and transport an ion while the experimental regions are configured to independently hold and manipulate other ions. A total of 48 separately controllable potentials were applied to the chip and distributed to the 150 control electrodes.

3.1. The design library

The trap geometry shown in figure 1(c) was assembled from a library of component designs. Each component has a specific function such as loading or transporting ions. A sample from this library is shown in figure 5. By developing and testing a component library, other trap designs can be assembled from the same library and use precalculated transport waveforms (see section 3.3). This components library includes connections to the control electrodes using the lower conducting layer. Standardizing the connections in this way greatly simplifies the design process since the connections tend to be very dense around complex components such as the junctions.

3.2. Electrode geometry optimization

The optimization procedure for the junction shape begins with a Y junction that has straight-edged rf electrodes, as shown in figure 4(a). The two rf rails in each arm of the Y junction
Figure 4. Optimization of junction rf electrodes (shown in red). (a) Initial shape used for optimization and (b) the ponderomotive potential $\Phi_{pp}$ at the center of the ponderomotive tube. The horizontal axes of (b) and (d) represent the position along the ponderomotive tube starting at the junction center ($x = 0$) and proceeding outward along the tube to the right. The ponderomotive potential is evaluated for the experimental conditions (51 V peak rf at 90.7 MHz and trapping $^{24}\text{Mg}^+$). (c) Optimized junction shape and (d) its ponderomotive potential, which is 300 times smaller than the unoptimized case shown in (b).

have different widths, 40 and 60 $\mu$m, respectively, in order to rotate the rf quadrupole axis. This rotation simplifies the control potentials needed to allow a single laser beam, which is parallel to the electrode surface, to Doppler cool a trapped ion [16]. Close to the junction, the minimum value of the rf ponderomotive potential is not zero at every point along the center of the ponderomotive tube, but instead forms ponderomotive ‘barriers’, as shown in figures 4(b) and (d). The optimization objective was to find appropriate electrode shapes that minimized the absolute height of these barriers. The inner and outer edges of the initial electrodes are broken up into 16 line elements of roughly equal lengths. The locations of the points at which these line elements connect (shown as dots in figures 4(a) and (c)) serve as optimization parameters. The Nelder–Mead simplex optimization algorithm [17] was then used to systematically move these locations around in order to minimize the height of the barriers while maintaining the
Figure 5. Example components from the design library. The components are assembled to form larger trapping structures such as the design detailed in this paper. Additional components provide standardized electrical connections on the lower conducting layer for the control electrodes. In the experimental regions, one zone on either side was elongated to provide storage zones. A narrow central electrode and additional electrodes outside the rf rails provide an electrical potential wedge for separating groups of ions as in [14, 15]. The shape of the connection pads was chosen to provide wide-angle access to various electrodes surrounding the pads.

The shape of the junction allows for threefold rotational symmetry of the junction. The algorithm uses the Biot–Savart-like law for two-dimensional electrostatic potentials derived in [18], which provides an analytical solution for the electric field contributions of any straight-line edge of a two-dimensional region at fixed potential and surrounded by a ground plane. Working with two-dimensional electrodes having straight-line edges therefore allows for rapid addition of the contributions due to the different edges. We also assume that the potential is specified everywhere on a two-dimensional plane, i.e. the gapless plane approximation [19–21]. Figure 4(c) illustrates the optimized electrode shapes and the corresponding ponderomotive barriers are shown in figure 4(d), achieving a suppression factor of greater than 300 over the initial ponderomotive barrier height in figure 4(b).

The shape of the rf rails was modified slightly to meet the constraints of our fabrication equipment including mitering of acute angles. The rf rails were then truncated at 270 µm from the junction center and encapsulated with control electrodes into a component for inclusion in the component library as shown in figure 6. The perimeter of the rf rails is listed in table 1. The ponderomotive barriers were calculated with the use of both the Biot–Savart-type integral in the gapless approximation and the use of a boundary element calculation. The resulting ponderomotive barriers are shown in figure 7(a) and an ion height above the surface is given in figure 7(b). As expected the ponderomotive trapping frequencies become weaker close to the center of the junction, as shown in figure 7(c). The boundary element calculation includes the effect of the gaps between the electrodes and the effect of the lower conducting layer as it passes under the gaps.

The control electrodes were designed once the rf electrodes shapes were determined. For adequate control along the ponderomotive tubes in the straight transport sections, the control electrodes were divided into 60 µm long segments (roughly 1.5 times the ion/surface distance). Near the junction, the control electrodes were designed with narrower electrodes to increase the spatial control of the potentials. The junction design also allows connections from the inside of the hexagonal ring to the outside to pass under the control electrodes.
3.3. Transport waveforms

We apply time-dependent potentials (waveforms) to the control electrodes to transport the ions along the ponderomotive tube. The process of generating these waveforms begins with a calculation to determine a set of points, $x_{\text{center}}$, that trace out the center of the ponderomotive tube and are separated by $1 \mu m$ intervals along the region the ion is to be transported. The goal then is to find, for each of the points in the set $x_{\text{center}}$, the appropriate control electrode potentials that trap the ion in a harmonic potential well centered at that point. The potentials are constrained to not displace the ion transversely to the tube, where it will experience rf micromotion [16]. Smoothly switching between the potentials that give rise to wells centered at successive points along the ponderomotive tube will then form a waveform that transports the ion along the tube.

The process of calculating these potentials is divided into two steps for each of the $1 \mu m$ spaced locations, say the point $x_{\text{center}}^{(i)}$. In the first step, we use constrained linear programming optimization to ensure that the control electrode potentials do not create an electric field at $x_{\text{center}}^{(i)}$. Given $n$ participating control electrodes, let $\mathbf{A}$ represent the $3 \times n$ matrix containing in the $j$th column the contributions to the electric field component at $x_{\text{center}}^{(i)}$ in the $x$, $y$ and $z$ directions, respectively, if the $j$th electrode is held at $1$ V and all others at $0$ V. The null space of $\mathbf{A}$ then gives a set of $n - 3$ vectors $\mathbf{v}$, each satisfying the condition $\mathbf{A}\mathbf{v} = 0$, i.e. the entries in each vector $\mathbf{v}$ give a combination of potentials that may be applied to the $n$ electrodes without producing an electric field at $x_{\text{center}}^{(i)}$. The desired potential well may be constructed from a linear combination
Table 1. Coordinates along the perimeter of one of the junction’s three rf rails after the optimized shape was modified for inclusion in the design library (see the text and figure 6(a)). The 5 \( \mu \text{m} \) gap between the rf and control electrodes was centered on this perimeter, cutting equally into both sets of electrodes as shown in figure 6(b). The other two rails can be obtained from these coordinates by rotations of 120\(^\circ\) and 240\(^\circ\) about the origin.

<table>
<thead>
<tr>
<th>Vertex</th>
<th>( x (\mu\text{m}) )</th>
<th>( y (\mu\text{m}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>270.0</td>
<td>80.0</td>
</tr>
<tr>
<td>2</td>
<td>172.0</td>
<td>80.0</td>
</tr>
<tr>
<td>3</td>
<td>134.5</td>
<td>74.7</td>
</tr>
<tr>
<td>4</td>
<td>80.1</td>
<td>155.9</td>
</tr>
<tr>
<td>5</td>
<td>7.8</td>
<td>148.9</td>
</tr>
<tr>
<td>6</td>
<td>-39.7</td>
<td>188.8</td>
</tr>
<tr>
<td>7</td>
<td>-84.3</td>
<td>266.0</td>
</tr>
<tr>
<td>8</td>
<td>-118.9</td>
<td>246.0</td>
</tr>
<tr>
<td>9</td>
<td>-91.4</td>
<td>198.4</td>
</tr>
<tr>
<td>10</td>
<td>-42.7</td>
<td>124.6</td>
</tr>
<tr>
<td>11</td>
<td>-13.0</td>
<td>90.7</td>
</tr>
<tr>
<td>12</td>
<td>1.8</td>
<td>70.0</td>
</tr>
<tr>
<td>13</td>
<td>2.0</td>
<td>19.7</td>
</tr>
<tr>
<td>14</td>
<td>1.2</td>
<td>8.5</td>
</tr>
<tr>
<td>15</td>
<td>4.0</td>
<td>3.4</td>
</tr>
<tr>
<td>16</td>
<td>14.9</td>
<td>7.1</td>
</tr>
<tr>
<td>17</td>
<td>31.4</td>
<td>52.3</td>
</tr>
<tr>
<td>18</td>
<td>38.1</td>
<td>53.6</td>
</tr>
<tr>
<td>19</td>
<td>58.8</td>
<td>29.7</td>
</tr>
<tr>
<td>20</td>
<td>189.7</td>
<td>21.3</td>
</tr>
<tr>
<td>21</td>
<td>220.4</td>
<td>20.0</td>
</tr>
<tr>
<td>22</td>
<td>270.0</td>
<td>20.0</td>
</tr>
</tbody>
</table>

of these vectors while maintaining the zero field condition. We also constrain the potentials to be in the range \( \pm 5 \text{ V} \), which is within the range our control electronics can generate.

In the second step, the potential well of the desired trapping frequency, and centered on \( \mathbf{x}^{(i)}_{\text{center}} \), is now generated from a linear combination of the above solutions. We use a Nelder–Mead simplex optimization algorithm [17], which varies the contribution of each null space vector in such a linear combination. In each iteration the potential due to the current linear combination of null space vectors is calculated at ten points along the tube, extending over \( \pm 50 \mu\text{m} \) from the well center. The algorithm minimizes the root-mean-square difference between the calculated potential and a target harmonic potential. This method of fitting to the \( \pm 50 \mu\text{m} \) interval ensures that the well has sufficient depth along the rf tube. During the calculation we ignore the effect of the applied potentials on the trapping in the radial direction. However, choosing the control electrode target potential to have a trapping frequency significantly smaller than the local radial ponderomotive trapping frequencies ensures that the potential does not become anti-trapping in the radial direction. We post-check that this is indeed the case, and that the trap has sufficient depth in all directions.
Figure 7. Trapping characteristics of the final junction design (junction center located at $x = 0$). (a) Ponderomotive barriers calculated (solid line) using the Biot–Savart-like boundary integrals in the gapless approximation and (dashed line) using a boundary element code to calculate the potentials including $5 \mu m$ gaps between the electrodes and the lower layer interconnects (see the text). The boundaries in the calculations are defined by the coordinates given in table 1. The ponderomotive potential is evaluated for the experimental conditions (51 V peak rf at 90.7 MHz and trapping $^{24}\text{Mg}^+\text{)}$). The ponderomotive barrier rising to the right of $x = 300 \mu m$ does not affect transport in the vicinity of the junction. It is caused by the adjacent loading zone, which has been included in this calculation. (b) Ion height as a function of position. (c) Ponderomotive radial trapping frequencies as a function of position.

Ideally the potentials calculated by the above method would vary smoothly as the ion is transported along the ponderomotive tube, but we find that the algorithm intermittently produced sharp, unwanted jumps in the potentials that form wells centered on successive $x_{\text{center}}^{(i)}$. These jumps most likely occur because all $n = 14$ independent control electrodes were included in our optimization for precise specification of the potential at many points, therefore allowing many different, equally viable optimized solutions. To remove these jumps, the potentials are post-smoothed by replacing the potentials associated with a well centered on $x_{\text{center}}^{(i)}$ by the average of the potentials associated with the two nearest neighboring wells, i.e. those wells centered on $x_{\text{center}}^{(i-1)}$ and $x_{\text{center}}^{(i+1)}$. This averaging can be repeated several times, first for all $x_{\text{center}}^{(i)}$ where $i$ is odd and then for all $x_{\text{center}}^{(i')}$ where $i'$ is even, and so on. After ten levels of smoothing the zero-field condition at each well center is only weakly violated. We estimate this violation to be no greater than that caused by uncertainties in the trap fabrication, which limits the precision with which we can specify the location of the tube center to within 1 $\mu m$.

4. Results

To test the basic features of the trap, $^{24}\text{Mg}^+$ ions were loaded by photoionizing (285 nm) a neutral flux of $^{24}\text{Mg}$ created in a resistively heated oven. To prevent contamination and possible shorting of the trap electrodes by the neutral flux, the oven was located behind the chip (as viewed in figures 1(b) and (c)). Some of the flux then enters the loading regions through two $12 \mu m$ by $20 \mu m$ slots (see figure 3(b)). From 1 to 10 ions were typically loaded, depending on the loading duration, neutral flux, and photoionization intensity. The ponderomotive potential is significantly perturbed in the vicinity of the loading slots, primarily because the dielectric conditions.
quartz distorts the rf electric field. After loading, the ions were located 20 µm to the side of the slot (along the axis) due to this perturbation.

The trap was enclosed in a vacuum chamber that had several glass viewports for laser access and for imaging the ions. The viewport used for the imaging was 6.5 cm in diameter and located 2.5 cm above the trap surface. We suspect that photoelectrons emitted from the trap surface by the photoionization beam and the Mg$^+$ Doppler cooling beams (280 nm) were charging the glass in the viewport. This apparent charging prevented the trap from holding ions without applying strong compensation fields using the control electrodes. The timescale for the charging was minutes, while discharging required several hours. To mitigate this problem, we installed a grounded gold screen (42 µm spacing with 83% light transmission) 5 mm above the trap, parallel to the surface. At this distance, the screen does not significantly affect the trap potentials but it greatly reduced the apparent charging effects [22].

4.1. Transport

Ions were successfully loaded in both loading regions of the trap and transported to and through the first adjacent Y junction, as shown in figure 8. We transported individual or groups of ions 850 µm from the loading zone to the leg of the junction in 2 ms (figure 8(a)). Potentials applied to the control electrodes then transfer the ion to either of the other two legs. Figure 8(b) shows a trapped ion in three locations around the leftmost junction of figure 1. We were able to transport ions between any pair of the junction’s three legs and in any order. We also transported ions to the center of the junction directly from the loading zone, as shown in figure 8(a).

In these preliminary experiments, we were unable to use the calculated waveforms for adiabatic transport between the junction’s legs because of unexpected barriers to the ion transport around 30 µm from the junction center. Instead, this part of the transport used a waveform that initially generated a symmetric set of three wells, one on each leg. During transport, the control potentials were adjusted briefly to form a single well on one leg with a barrier on all three legs that prevented the ion from exiting the vicinity of the junction. The ion then Doppler cools into the single well. Finally, the original symmetric three-well potential was restored with the ion in the new position.

Transport between the legs in this initial test was thus ballistic and required that a laser cooling beam overlap the final location of the transport in each leg. The overlap of these cooling beams covered much of the junction and the cooling was applied continuously during transport. This is in contrast to the ion transport through a junction demonstrated in both [9, 10], where the transport was not assisted by laser cooling. Stray fields, varying over the course of a day, also made the transport erratic. With careful adjustment of the waveforms, an ion could be transported around the legs of the junction dozens of times before being lost. Transport loss is expected to be high due to the large ion heating from this type of transport. Transport between two of the junctions seemed technically feasible but practically difficult because of the need for continual adjustment of the waveform to compensate for time varying charging fields. We did transport ions just past the midpoint between two of the junctions. However, much of the work with this trap has been focused on determining the source of the barrier and the stray fields, which is still not understood.

For a further demonstration of transport through this junction, please see the online supplementary material for this paper, available at stacks.iop.org/NJP/12/033031/mmedia. The included video shows a consecutive sequence of six transports clockwise around the junction followed by six transports counterclockwise.

Figure 8. (a) Diagram of a Y junction with an inset showing an ion trapped in the junction center. (b) Transport through the leftmost Y junction shown in figure 1. Each photo (false color) shows an ion in one of three locations around the junction. All ions were transported to the junction from a load zone attached to the junction’s left leg. To illustrate the electrode locations in the photos, a laser beam was directed at the quartz substrate so that the scattered light illuminates the gaps between the electrodes.

4.2. The heating rate

One characteristic parameter of an ion trap is the in situ heating rate of the ions due to fluctuating electric fields at the ion location. Typically, miniature rf Paul traps have observed heating rate orders of magnitude greater than expected from simple Johnson noise on the control electrodes [16, 23, 24]. The source of these noisy electric fields is still not understood, but they can be approximately characterized according to ion–electrode distance. The heating rate for a single ion at the location shown in figure 1(c) was observed to be \( \dot{n} = 87(11) \times 10^3 \) phonon s\(^{-1}\) at a 3.5 MHz axial frequency and a 38 \( \mu \)m ion-to-surface distance, using the recooling method described in [24]. The corresponding electric field noise spectral density seen by the ion is \( S_E(\omega) \approx \dot{n}(4m\hbar\omega/e^2) \approx 1.3(2) \times 10^{-9} \) V\(^2\) m\(^{-2}\) Hz\(^{-1}\) [24–26], where \( m \) and \( e \) are the ion mass and charge, respectively. This places this trap in the average range of heating rates for its size scale.

We also characterized a separate linear surface-electrode trap made with 1 \( \mu \)m thick electroplated gold. This trap demonstrated transport over 20 zones (greater than 1 mm travel) and, as with the hexagonal trap, used multiple conducting layers for interconnects. The linear
trap was designed with a subset of the component library used for the trap of figure 1. The heating rate in this trap gave $S_E(\omega) < 6 \times 10^{-11} \text{V}^2 \text{m}^{-2} \text{Hz}^{-1}$ as determined by the axial heating of a $^{24}\text{Mg}^+$ ion at 4.5 MHz and 40 $\mu$m ion–electrode distance using the recooling method. This bound is more than an order of magnitude lower than the evaporated gold trap.

The relatively low heating measured for the electroplated-gold surface-electrode trap described in [8] as compared to other traps of similar dimensions has elicited conjecture that electroplated gold might have low heating rates in general. The results from the linear trap reported here further support this conjecture, but given the wide variation in heating rates measured between traps of similar materials, we cannot draw a definitive conclusion from these measurements.

5. Conclusions

The remaining issues that need to be investigated with this type of trap include the sources of stray electric fields and apparent discrepancies between the observed ion positions and those predicted from potential calculations. The design components are not limited to the gold-on-quartz construction and could be realized with other choices of materials and fabrication processes. Different strategies for electrode geometry optimization should also be explored. We expect that as the component library is refined and expanded, the design and construction of ion traps for future quantum information processing experiments will become considerably faster and more reliable.

Acknowledgments

We thank M G Blain, J J Hudgens, A Pimentel and J Gallegos of Sandia National Laboratories for the focused ion beam milling of the loading slots and wirebonding. We acknowledge the funding by IARPA, ONR and the NIST quantum information program. HU acknowledges support from the South African Council for Scientific and Industrial Research (CSIR). This paper is a contribution of NIST and is not subject to US copyright.

© US Government

References


[9] Hensinger W K, Olmschenk S, Stick D, Hucul D, Yeo M, Acton M, Deslauriers L, Monroe C and
Rabchuk J 2006 T-junction ion trap array for two-dimensional ion shuttling storage and manipulation Appl. Phys. Lett. 88 034101

High-fidelity transport of trapped-ion qubits through an X-junction trap array Phys. Rev. Lett. 102 153002


2nd edn (Englewood-Cliffs, NJ: Prentice-Hall)


2009 Entangled mechanical oscillators Nature 459 683


investigation of planar ion traps Phys. Rev. A 73 032307

cooling and low heating of trapped Cd+ 111 ions Phys. Rev. A 70 043408


in cryogenic surface-electrode ion traps Phys. Rev. Lett. 100 013001

Continuous-wave versus time-resolved measurements of Purcell factors for quantum dots in semiconductor microcavities

M. Munsch, A. Mosset, A. Auffèves, S. Seidelin, and J. P. Poizat
CEA/CNRS/UJF Joint Team “Nanophysics and Semiconductors,” Institut Néel–CNRS, BP 166, 25 rue des Martyrs, 38042 Grenoble Cedex 9, France

J.-M. Gérard
CEA/CNRS/UJF Joint Team “Nanophysics and Semiconductors,” CEA/INAC/SP2M, 17 rue des Martyrs, 38054 Grenoble, France

A. Lemaitre, I. Sagnes, and P. Senellart
Laboratoire de Photonique et de Nanostructures, LPN/CNRS, Route de Nozay, 91460 Marcoussis, France

Received 3 June 2009; revised manuscript received 3 August 2009; published 10 September 2009

The light-emission rate of a single quantum dot can be drastically enhanced by embedding it in a resonant semiconductor microcavity. This phenomenon is known as the Purcell effect and the coupling strength between emitter and cavity can be quantified by the Purcell factor. The most natural way for probing the Purcell effect is a time-resolved measurement. However, this approach is not always the most convenient one and alternative approaches based on a continuous-wave measurement are often more appropriate. Various signatures of the Purcell effect can indeed be observed using continuous-wave measurements (increase in the pump rate needed to saturate the quantum dot emission, enhancement of its emission rate at saturation, and change in its radiation pattern), signatures which are encountered when a quantum dot is put on resonance with the cavity mode. All these observations potentially allow one to estimate the Purcell factor. In this paper, we carry out these different types of measurements for a single quantum dot in a pillar microcavity and we compare their reliability. We include in the data analysis the presence of independent, nonresonant emitters in the microcavity environment, which are responsible for a part of the observed fluorescence.

DOI: 10.1103/PhysRevB.80.115312 PACS number(s): 42.50.Pq, 78.67.Hc, 78.90.+t

I. INTRODUCTION

Coupling an emitter to a cavity strongly modifies its radiative properties, giving rise to the observation of cavity quantum electrodynamics (CQED) effects, which can be exploited in the field of quantum information and fundamental tests of quantum mechanics. A variety of systems allows one to implement different CQED schemes, ranging from Rydberg atoms and alkaline atoms in optical cavities to superconducting devices, as well as for semiconducting quantum dots (QDs) (for an early review, see Ref. 5) coupled to optical solid-state cavities. Thanks to impressive recent progress in nanoscale fabrication techniques, vacuum Rabi splitting, giant optical nonlinearities at the single-photon level, and vacuum Rabi oscillation in the temporal domain have been demonstrated for single InAs/GaAs QDs coupled to microcavities. Success in sophisticated CQED experiments requires first of all an efficient enhancement of the spontaneous emission (SE) of an emitter coupled to a resonant single mode cavity. The dynamical role of the cavity is quantified by the so-called Purcell factor $F_P$, namely, the ratio between the emitter’s SE rate with and without the cavity. For an emitter perfectly coupled to the cavity the Purcell factor only depends on the cavity parameters and takes on the value denoted $F_P$ which is given by

$$F_P = \frac{3}{4\pi^2} \left( \frac{\lambda}{V} \right)^3 \left( \frac{\lambda}{n} \right)^3$$

where $Q$ is the cavity quality factor, $V$ the cavity volume, $\lambda$ the wavelength for the given transition, and $n$ the refractive index.

The Purcell effect using QDs as emitters has first been observed when coupled to pillar type microcavities in the late 1990s. Moreover, when its radiation pattern is directive, the cavity efficiently funnels the spontaneously emitted photons in a single direction of space. This geometrical property allows one to implement efficient sources of single photons or even single, indistinguishable photons. A high Purcell factor also enhances the visibility of CQED-based signals such as QD-induced reflection. Beyond its seminal role, the Purcell factor appears thus as an important parameter which measures the ability of a QD-cavity system to show CQED effects and has therefore become a figure of merit for quantifying these effects. It is obviously important to develop reliable methods to measure accurately this figure of merit.

Two types of measurements are possible. The first one is the most intuitive and simply consists in comparing the lifetime of a QD at and far from resonance with the cavity mode, using a time-resolved setup. This is feasible only as long as the resonant QD lifetime is longer than the time resolution of the detector or more generally longer than any other time scales involved, such as the exciton creation time (capture and relaxation of electron and holes inside the QD). For a large Purcell factor, this might be a limiting condition. Instead, the Purcell effect can be estimated from measurements under continuous-wave (CW) excitation. When approaching QD-cavity resonance, the pump rate required to saturate the emission of the QD is higher due to the shortening of the exciton lifetime. The Purcell effect also produces a preferential funneling of the QD SE into the cavity mode and...
thus increases the photon collection efficiency in the output cavity channel. Measuring either the saturation pump rate or the photoluminescence (PL) intensity as a function of detuning enables thereby one to measure the Purcell factor.

This paper first aims at evaluating the consistency of these different methods and to compare their accuracy. Moreover, both methods suffer from the same problem, related to the fact that the cavity is illuminated by many other sources in addition to the particular QD being studied. Even far detuned QDs can efficiently emit photons at the cavity frequency. This feature has been observed by several groups worldwide, leading to theoretical effort to understand this phenomenon. All the models involve the decoherence-induced broadening of the QDs combined with cavity filtering and enhancement. Even though one can easily isolate the contribution of the single QD when it is far detuned from the cavity mode, this becomes much more difficult near resonance when other sources emitting via the cavity have to be taken into account. With this aim, we have developed a model which includes these contributions and therefore enables us to fit the experimental data and to derive a correct value of the Purcell factor.

II. Sample Characteristics and Setup

To fabricate the samples, a layer of InAs self-assembled QDs is grown by molecular-beam epitaxy and located at the center of a \( \lambda \)-GaAs microcavity surrounded by two planar Bragg mirrors, consisting of alternating layers of \( \text{Al}_{0.1}\text{Ga}_{0.9}\text{As} \) and \( \text{Al}_{0.95}\text{Ga}_{0.05}\text{As} \). The top (bottom) mirror contains 28 (32) pairs of these layers. The quality factor of the planar cavity is 14 000. In a subsequent step, the planar cavity is etched in order to form a micropillar containing the QDs. The specific micropillar discussed in the following has a diameter of 2.3 \( \mu \text{m} \) and the density of the quantum dots is approximately \( 2.5 \times 10^{10} \text{QDs/cm}^2 \).

The etching of the Bragg mirrors into a micropillar can deteriorate the quality factor of the cavity. To measure the micropillar quality factor, we perform a photoluminescence measurement at high power such that the ensemble of QDs act as a spectrally broad light source, which is used for probing the cavity. From this measurement, we extract a quality factor of our specific 2.3-\( \mu \text{m} \)-diameter sample mentioned above of \( Q=\lambda/\Delta\lambda =7500 \). This value agrees (to within 10\%) with reflectivity measurement using white light. We will, in the following section, use the corresponding bare cavity linewidth \( \kappa_0=\lambda/Q \) (in nanometers). Using Eq. (1) together with the measured value for the quality factor, we obtain \( F_p =18.6 \).

Our sample is located in a cryostat held at 4 K. For the continuous-wave measurements, the QDs are excited using a standard laser diode emitting at 820 nm while for the time-resolved measurements, we use a pulsed Ti:Sa laser centered at 825 nm (80 MHz repetition rate and 1 ps pulse width). In both cases, this corresponds to an off-resonant excitation in the GaAs barrier. In our pump power range, for InAs QDs, the capture and relaxation of the charge carriers is less than 50 ps. For our given QD-cavity coupling, this is much faster than the radiative decay and consequently we do not take it into account in the data analysis.

![FIG. 1.](image)

**FIG. 1.** (Color online) The full spectrum recorded at 4 K showing the inhomogeneous line, which presents a zoom on the section of interest including two isolated quantum dots (\( X_a \) and \( X_b \)) and their respective biexcitons (\( XX_a \) and \( XX_b \)), and the cavity mode (C).

The emitted light is recollected after passing a spectrometer (1.5 m focal and 0.03 nm resolution). The spectrometer has two output channels: one channel leads to a charge coupled device camera (for the CW measurements), the other to an avalanche photodiode with a 40 ps time resolution which, combined with a 5 ps resolution for the data-acquisition card and 65 ps resolution due to the spectrometer, gives us an overall resolution of 80 ps.

In Fig. 1 we give an overview of the different lines observed in a typical photoluminescence experiment for our particular micropillar to be studied in the following. Centered around 895 nm, we observe what is usually referred to as the inhomogeneous line, composed of hundreds of QDs. The micropillar has been processed such that the cavity resonance is located on the low-energy wing of this inhomogeneous line, where the QD density is very low, allowing us to optically isolate one single QD (denoted \( X_a \)) to be studied and, in particular, scanned through cavity resonance. We also note that its corresponding biexciton (\( XX_a \)) is blueshifted by about 1 nm, an amount which is larger than the cavity linewidth. For a given temperature, we can therefore make the biexciton off-resonance with the cavity while having the exciton centered at resonance. For this specific micropillar, this happens at 19.5 K. In this case, a second QD (\( X_b \)) appears about three cavity linewidths away (with its biexciton \( XX_b \) even further away) and is therefore also minimally affected by the cavity. All other QDs are much further detuned. Throughout the paper, we will assume that the radiative efficiency of the emitter is unity, which is a very good approximation for self-assembled QDs at low temperature.

In Fig. 2 we show the temperature dependence of the cavity resonance frequency as well as the two relevant QD emission wavelengths. The cavity frequency varies due to a temperature-dependent refractive index while the QD excitation energy follows the expected temperature dependence of the GaAs band gap. Due to this difference in temperature dependence, we can vary the QD-cavity detuning.

III. Continuous-Wave Measurements

Even though the Purcell effect is a dynamical phenomenon, it can be measured without a time-resolved setup. This
can be understood as follows. As the emitter’s lifetime decreases near resonance due to the Purcell effect, it becomes harder to saturate the optical transition. This can be quantified by measuring the increase in the pump rate required to saturate the emitter (see Sec. III A) or by measuring the actual cycling rate in a PL measurement at saturation (Sec. III B). So by comparing the on- and off-resonant saturation pump rate or PL intensity, the Purcell factor can be measured. More recently it has been demonstrated that one can also extract the Purcell factor due to the change in the fraction of SE that is funneled into the cavity mode. This is done by measuring the SE rate as a function of detuning for fixed pump power, as will be done in Sec. III C.

An illustration of the principle is given in Fig. 3(a). A QD is embedded in a cavity whose fundamental mode is nearly resonant with the excitonic $X_a$ transition (Fig. 1). We denote $\Delta$ the detuning between the excitonic transition and the cavity mode. The QD is nonresonantly pumped with a rate $r$ and decays by emitting photons either in the cavity mode or in other leaky modes with a rate which we suppose to be independent of the detuning $\Delta$ and identical to that of the bulk material (which is a reasonable approximation for QDs in micropillar cavities). As suggested by the PL spectra shown in part II, the QD should be modeled by a three-level system which includes the biexciton [Fig. 3(b)]. In the following we will concentrate solely on $X_a$, so for simplicity we will omit the subscript $a$. We denote $\gamma$ and $\gamma_{X_a}$ the coupling of the exciton ($X$) and biexciton (XX) transitions with the leaky modes. In addition to the leaky modes, the $X$ transition is coupled to the cavity mode with a rate $\Gamma(\Delta) = \gamma F \mathcal{L}(\Delta)$ where $F$ is the effective Purcell factor experienced by the QD, taking into account that it is not perfectly coupled to the cavity (in contrast to $F_P$ given in Eq. (1), which is only an upper bound for $F$). Moreover, $\mathcal{L}(\Delta) = 1/(1 + \Delta^2/\kappa_0^2)$ is a Lorentzian of width $\kappa_0$ corresponding to the empty cavity line shape. When pumping with a rate $r$, the average excitonic population is then given by

$$p_X(\Delta, r) = \frac{1}{1 + \frac{r}{\gamma X_a} + \frac{\gamma + \Gamma(\Delta)}{r}}.$$  

As mentioned in the first part of this paper, the role of the cavity is not only to enhance the cycling rate for the exciton ($X$) but also to efficiently funnel the emitted photons into the cavity mode. Provided that the emission pattern of the cavity is directional, which is the case for micropillars, the coupling with a conveniently positioned detector can be very efficient, whereas the coupling between leaky modes and detector remains poor. These geometrical efficiencies are, respectively, denoted $X_{cav}$ and $X_{leak}$ [see Fig. 3(a) and Ref. 24]. The PL intensity from our single QD collected by the detector can thus be written in the following way:

$$I_{X, \text{det}}(\Delta, r) = I_X^{\text{leak}}(\Delta, r) + I_X^{\text{cav}}(\Delta, r),$$  

where

$$I_X^{\text{leak}}(\Delta, r) = X_{leak} \gamma p_X(\Delta, r)$$  

is the PL intensity emitted through the leaky modes and

$$I_X^{\text{cav}}(\Delta, r) = X_{cav} \Gamma(\Delta) p_X(\Delta, r)$$  

the detected PL intensity emitted spatially into the cavity mode. Please note that the notation $cav$ applies to geometrical considerations but not to the emission frequency (this PL contribution is indeed emitted at the QD frequency). In our experiment, to separate $I_{X, \text{det}}$ from the PL intensity from all other light sources, we use of the spectrometer to select a window centered on our selected QD (see the inset in Fig. 1) and we then fit the line shape corresponding to the single QD with a Lorentzian function. When the QD-cavity detuning is large, it is easy to separate the QD line shape from the cavity but as the detuning decreases, they will partially overlap with each other. When this happens, to avoid that a part of the cavity peak erroneously is included in the single QD line shape, we also do a Lorentzian fit on the cavity profile, which we then subtract from the QD line shape. Note that in doing this, we also involuntarily omit from $I_{X, \text{det}}$ the part of the QD PL which is emitted at the cavity frequency but this part constitutes a small fraction of the total signal.

An example of typical experimental data is pictured in Fig. 4, where the PL intensities for different detunings $\Delta$ are plotted. As we generally measure the pump power denoted $P$ and not the pump rate $r$, we have chosen to plot the data as a function of the former (and we do the same in the graphs to

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{fig3.png}
\caption{(Color online) (a) The PL of the QD (illustrated as a triangle) arriving at the detector can be separated into two channels: one part emitted into loss channels (\(\gamma\)) and directed to the detector with a probability \(X_{\text{leak}}\) and the part emitted into the cavity \(\Gamma(\Delta)\) and detected with a probability \(X_{\text{cav}}\). (b) Three-level scheme including the exciton [X] and biexciton [XX]. The notations are defined in the text.}
\end{figure}
follow). This also means that $P_{\text{sat}}$ is the pump power corresponding to the pump rate $r_{\text{sat}}$.

For each detuning, the maximal intensity $I_{\text{sat},\text{det}}(\Delta, r_{\text{sat}}(\Delta))$ is reached when the $X$ transition is saturated, where $r_{\text{sat}}(\Delta)$ denote the pump rate required to saturate the transition (saturation pump rate). Note that the highest values of $I_{\text{sat},\text{det}}(\Delta, r_{\text{sat}})$ and corresponding $r_{\text{sat}}$ are reached at resonance, which is coherent with the enhancement of the $X$ transition rate induced by the cavity. In the following, we will analyze the curves presented in Fig. 4 (and further equivalent curves not added to the graph for clarity), in four different ways (Secs. III A and III D).

A. Saturation pump rate measurements

In the first method the Purcell factor is extracted from the saturating pumping rate intensity as a function of detuning $r_{\text{sat}}(\Delta)$, corresponding to black filled circles in Fig. 4. This method has been proposed as a substitute for the time-resolved measurements and has been widely used for micropillars,\textsuperscript{20} microdiscs,\textsuperscript{20} and photonic crystals.\textsuperscript{28} The analytic expressions can be found by determining the pump rate corresponding to the maximum intensity of Eq. (3). We obtain

$$r_{\text{sat}}(\Delta) \propto \sqrt{1 + F \mathcal{L}(\Delta)}. \quad (6)$$

In Fig. 5(a) we have plotted the data and the fit according to Eq. (6) where we have imposed the bare cavity linewidth based on independent measurements. From the fit, we extract a Purcell factor of

$$F = 3.7 \pm 1.0, \quad (7)$$

which the relatively large error is due to the uncertainty of $r_{\text{sat}}$. The slope of the baseline in Fig. 5(a) is due to the increase in temperature for increased detuning. As mentioned in Sec. II, we use an optical excitation obtained through the pumping of the GaAs barrier material. The diffusion length of the electrons and holes increases with temperature so that

\begin{align*}
\frac{\Delta}{\kappa} & = 0, \\
\frac{\Delta}{\kappa} & = -0.14, \\
\frac{\Delta}{\kappa} & = -0.32, \\
\frac{\Delta}{\kappa} & = -0.56, \\
\frac{\Delta}{\kappa} & = -0.35, \\
\frac{\Delta}{\kappa} & = -2.72.
\end{align*}

B. Saturation PL intensity measurements

Another similar approach again based on the black filled circles in Fig. 4 has been used in recent papers.\textsuperscript{29,30} This method corresponds to exploiting directly the maximum intensity of Eq. (3) given by

$$I_{\text{sat},\text{det}}(\Delta, r_{\text{sat}}) \propto \frac{F \mathcal{L}(\Delta)}{1 + \sqrt{1 + 2 F \mathcal{L}(\Delta)}}, \quad (8)$$

where we have made the assumption that $\chi_{\text{cav}} \gg \chi_{\text{cav}}$ (see Sec. III D), which is valid for micropillars, but not necessarily for photonic crystals.

In Fig. 6 we have plotted the data and the fit (the maximum normalized to one) according to Eq. (8), again with the bare cavity linewidth fixed. From the fit, we extract a Purcell factor of

$$X \approx 3.8.$$
In this case, the intrinsic uncertainty in the PL measurement is quite small but is amplified by the fitting procedure, resulting in the stated error.

C. PL intensity with fixed pump rate

In the two previous sections, we have used the data corresponding to the saturation pump rate and intensity. Instead, we can also use the emitted PL intensity, not at saturation, but for a fixed pump rate. This amounts to using the PL intensity corresponding to the intersection of the curves in Fig. 4 with a straight vertical cut. In particular, we have plotted in Fig. 7 the PL intensity for powers below and above saturation. The fit corresponds again to Eq. (3) but this time with the pump rate fixed (\( r = 30 \) and 300 \( \mu W \) for the two curves, respectively). From both curves we have subtracted a global offset corresponding to the PL intensity \( I_{\text{X, det}} \) at \( \Delta = \infty \).

Below saturation the change in the light intensity \( I_{\text{X, det}} \) as the QD is scanned across the cavity resonance is due to the geometrical redirection of the emission alone (a modification in the emission pattern). What we detect is a projection of a fraction of the micropillar emission pattern onto the microscope aperture. More precisely, for low powers (well below saturation) \( p_\Delta (\Delta, r) = \frac{\gamma_\Delta}{\gamma_\Delta + r} \), and we obtain

\[
F_{\text{X}}(\Delta) \approx \frac{F \mathcal{L}(\Delta)}{1 + F \mathcal{L}(\Delta)} = \beta(\Delta),
\]

where we have defined the function \( \beta(\Delta) \) which can be interpreted as the fraction of the emission pattern overlapping with the cavity mode. This function is broader than the Lorentzian profile of the cavity mode by a factor \( \sqrt{F+1} \).

Above saturation, the geometrical redirection of the emission pattern is still present but the light intensity \( F_{\text{X}} \approx \mathcal{L}(\Delta) \) follows now the \( \mathcal{L}(\Delta) \) profile of the cavity owing to the additional effect of the larger emission rate of the quantum dot caused by the shortening of its lifetime. More precisely, in the regime well above saturation we have \( p_\Delta (\Delta, r) = \gamma_\Delta/r \) and we get

\[
F_{\text{X}}(\Delta, r) \propto F \mathcal{L}(\Delta).
\]

From the ratio of the two widths, we extract a Purcell factor of

\[
F = 3.2 \pm 0.9,
\]

where the stated uncertainty arises from the intensity measurements, which is the dominant source of error in this case.

As mentioned in Sec. II, we can assume that the charge relaxation is much faster than the radiative decay. Due to the nonlinear power dependence of charge relaxation, the validity of this assumption has so far been required throughout the analysis. The method here presented, however, is based on measurements using a constant pump power and is therefore robust against such nonlinearities. This method therefore remains valid even when the stated assumption no longer holds true.

D. PL intensity ratio at low and high pump rate

This method also consists in comparing the light emitted by the single QD for different detunings but only requires four of the measurements used above (below and above saturation at resonance and far from resonance). Here we do not subtract the offset due to \( \chi_{\text{leak}} \) as done above, which has the advantage that it allows us to quantify \( \chi_{\text{cav}}/\chi_{\text{leak}} \). We define as \( \epsilon(\Delta, r) \) the following ratio:

\[
\epsilon(\Delta, r) = \frac{I_{\text{X, det}}(0, r)}{I_{\text{X, det}}(\Delta, r)} = \frac{p_\Delta(0, r)}{p_\Delta(\Delta, r)} \times \frac{\chi_{\text{leak}} + \chi_{\text{cav}} F}{\chi_{\text{leak}} + \chi_{\text{cav}} F} \mathcal{L}(\Delta),
\]

where the parameter \( \alpha(\Delta) \) depends on the cavity funneling properties.

For pump rates below the pump rate required to saturate (where \( p_\Delta(\Delta, r) = \gamma_\Delta/\gamma_0 \))

\[
\epsilon_{\text{below}}(\Delta) = \alpha(\Delta) \times \frac{1 + F \mathcal{L}(\Delta)}{1 + F},
\]

whereas above the saturation pump rate (again using that \( p_\Delta(\Delta, r) = \gamma_\Delta/r \))

\[
\epsilon_{\text{above}}(\Delta) = \alpha(\Delta).
\]

Taking the ratio between \( \epsilon_{\text{below}} \) and \( \epsilon_{\text{above}} \), \( \alpha(\Delta) \) cancels and with an independent measurement of \( \mathcal{L}(\Delta) \) (see Sec. II), we obtain a Purcell factor of

\[
F = 2.5 \pm 0.5,
\]

where the error arises from the uncertainty on the intensity measurements. From the separate value of \( \epsilon_{\text{above}} \) (or \( \epsilon_{\text{below}} \)) we get

\[
\frac{\chi_{\text{cav}}}{\chi_{\text{leak}}} \sim 15 \pm 4.5,
\]

confirming that the cavity is much better coupled to the detector than the leaky modes. This ratio depends on the radia-
FIG. 8. (Color online) Lifetime measurements at different pump powers of the quantum dot while (a) far detuned from the cavity and (b) close to resonance. In (a) the solid line corresponds to $P = 3P_{\text{sat}}$ and dotted line to $P = P_{\text{sat}}$. In (b) we have $P = P_{\text{sat}}$ (solid line) and $P = P_{\text{sat}}/10$ (dotted line) and $P = P_{\text{sat}}/30$ (dashed line).

Note that we have only included the presence of exciton and biexciton in all given formulas. We have, however, repeated the above analysis, allowing for all orders of exciton levels, without any significant change in final results within the range of used pump powers.

IV. TIME-RESOLVED MEASUREMENTS

As a way to confirm our continuous-wave measurements of the Purcell factor, we have performed a detailed study of the lifetime as a function of the detuning, using time-resolved spectroscopy. This technique has been used extensively for many different systems since it was the first method to be used. In fact, the Purcell factor can be written as

$$F = \frac{\tau(\Delta = 0)}{\tau(\Delta = \infty)} - 1,$$

where $\tau$ is the lifetime of the QD and $\Delta$ again is the detuning. Opposite Eq. (1), this definition also applies to an emitter that is not perfectly coupled to the cavity (within the approximation where $\gamma_{\text{leak}} = \gamma_{\text{sat}}$, the latter denoting the SE of the QD into the unprocessed, or bulk, material).

In Fig. 8 we show the measured lifetime of our quantum dot for different pump powers. In (a) the QD is detuned from the cavity resonance while in (b) it is at resonance. In the first case, (a), we show data corresponding to two different powers. When $P = P_{\text{sat}}$ (dotted line) the QD exhibits the typical monoexponential decay (also the case for any power below $P_{\text{sat}}$). When $P > P_{\text{sat}}$ (solid line), the effect of the biexciton can be observed as a rounding off of the curve at short time, which corresponds to the delay in the recombination of the exciton. The data fit very well with a model including three levels (a ground state, the exciton, and biexciton states) and we extract the exciton and biexciton lifetimes, which are the same for the two different powers,

$$\tau_X = 0.80 \pm 0.05 \text{ ns and } \tau_{XX} = 0.40 \pm 0.02 \text{ ns.}$$

As the biexciton is not influenced by the Purcell effect (for the detunings used in this experiment), the obtained value can be used as a fixed parameter when we then fit the data for the resonant case. Note that all our fits have been convoluted with the experimental system’s response time (80 ps time resolution). On the contrary, the resonant case (b) shows a power dependency that cannot be explained by our simple three-level model used above. We clearly observe in Fig. 8(b) a change from a quasimonoexponential decay to a biexponential decay, when lowering the pump rate. We exclude a prominent role of dark excitons since a monoexponential behavior is observed in the nonresonant case (a). In addition, the fact that the second lifetime of the exponential decay is fast (less than 1 ns) also tends to eliminate this hypothesis. We believe that this behavior is due to detuned emitters, which contribute to the collected intensity via the cavity emission. Recent experiments7,10,16,21 show that QDs could emit photons in the cavity mode even at rather high detunings (several times the cavity linewidth). In contrast to CW measurements where we could separate the emission of our QD from the one of the cavity using appropriate Lorentzian fits, in the present case we do not have access to the full spectra and therefore cannot use the same technique. Instead we must select a frequency window around the QD line, for which we integrate all PL. This makes us unable to filter out the cavity component which overlaps in frequency with the chosen window [when close to resonance, as in Fig. 8(b)]. As a result we measure two different times: the shorter one is the lifetime of our single QD (undergoing Purcell effect), whereas the longer one corresponds to the lifetime of other detuned emitters. The higher the pump power, the more dominant is the signal due to the contribution of the detuned emitters. Therefore, at high powers, the light from other emitters tends to make the signal invisible for our single QD. This is illustrated in Fig. 9, where we have shown the spectra corresponding to three different pump powers, ranging from high (a) to low (c) but for a fixed detuning. The fraction of light emitted via the cavity clearly dominates at high powers but decreases when lowering the pump power.

This is why, for high powers, only one lifetime can be observed [upper curves in Fig. 8(b)] and this lifetime is obviously no longer the QD radiative lifetime but corresponds to the light emitted via the cavity. Only for lower pump power, the true lifetime also becomes visible (lower curve) as seen by the biexponential decay. We therefore need to include these additional emitters that we can model (within
our pumping range) with a two-level system whose lifetime corresponds to an average lifetime, which can be measured in an independent experiment in which all QDs are far detuned. We obtain $0.8 \pm 0.05$ ns.

The exciton lifetime is the only free parameter in our fits (the biexciton lifetime is a fixed parameter). The excellent agreement between data and fit seems to validate our model. We find for the nonresonant and resonant cases, $\tau(\Delta=-0.2\kappa_0) = 0.80 \pm 0.05$ ns and $\tau(\Delta=0) = 0.2 \pm 0.01$ ns, which give a Purcell factor of

$$F = 3.0 \pm 0.5,$$

where the stated uncertainty arises from the exponential fit. Based on the above discussion, we remark that the low power condition is a necessary but not sufficient criterion for measuring the correct lifetime. Indeed, although all shown powers in Fig. 8 are below saturation, only the complete model gives the right lifetime. In Fig. 10 we have plotted the decay time of 0.8 ns and 0.5 ns, which give a Purcell factor of $3.0 \pm 0.5$, where the stated uncertainty arises from the exponential fit.

Based on the above discussion, we remark that the low power condition is a necessary but not sufficient criterion for measuring the correct lifetime. Indeed, although all shown powers in Fig. 8 are below saturation, only the complete model gives the right lifetime. In Fig. 10 we have plotted the decay time of 0.8 ns and 0.5 ns, which give a Purcell factor of $3.0 \pm 0.5$, where the stated uncertainty arises from the exponential fit.

**V. FINAL DISCUSSION**

We have presented several ways to measure the Purcell factor, which is an important figure of merit in CQED. All our CW measurements agree with each other, within the experimental uncertainty, for a Purcell factor of $3.0 \pm 0.4$. We emphasize that in our evaluation of the errors, we have not taken into account the stated 10% uncertainty for the bare cavity linewidth (see Sec. II). A simple PL measurement of the cavity linewidth has a negligible uncertainty but when probing the cavity by reflectivity measurements, this value turns out to be about 10% different. We also point out that the value measured by reflectivity is systematically higher than the one measured in PL. We will here revisit the obtained results for the Purcell factor in order to see how a 10% deviation on the quality factor would affect the values. While the first method (based on saturation pump power, in Sec. III A) does not depend on this parameter, all the other CW methods here presented do. In particular, the second technique, which uses the saturation intensity (Sec. III B), drastically depends on this parameter. In our case, an uncertainty of 10% on the quality factor would make the measurement based on this method useless. Though we still can fit the data with a correct shape, the obtained Purcell factor is absurd and exceeds the theoretical value. Finally, concerning the third method (Sec. III C), the modification of the Purcell factor induced by the 10% change in the quality factor amounts to 20%, which is slightly below the stated error due to the imprecision on the measurement. Therefore, this error is not significantly increased when allowing the given deviation on the quality factor. The time-resolved measurements also agree within the error bars with the CW measurements. The fact that we clearly do not observe a single exponential decay at resonance confirms the hypothesis that other light sources contribute to the light emitted into the cavity channel. In particular, for the time-resolved measurement, if not including this light in our model, the lifetime appears to be pump power dependent, even when we pump way below saturation which is clearly nonphysical. We thus underline that the commonly adopted criterion that the time-resolved spectroscopy of an exciton has to be made below saturation might not be sufficient. If additional emitters are present in the environment of the considered QD, it might be adequate to include their presence in the data analysis.

In conclusion, the agreement of the time-resolved measurements with the CW measurements suggests that both methods are reliable. The dramatic influence of the cavity linewidth uncertainty on the Purcell-factor error bars might be a reason for preferring $Q$-independent measurements such as time-resolved spectroscopy. On the other hand, the time-resolved measurements suffer from a lower signal-to-noise ratio and for some systems (photonic crystals, in particular, where the radiation pattern is less favorable), this becomes a
limiting factor, making the CW measurements more desirable. In that case, based on above considerations, we advise to use the method based on the saturation intensity with precaution, unless a very precise measurement of the cavity quality factor is available. If this is not the case, the other CW methods here presented seem more robust against an uncertainty on this parameter.

*signe.seidelin@grenoble.cnrs.fr

12. This requires that the emitter-cavity detuning is zero, that the emitter is located at the field antinode, and that it is quasimonochromatic (linewidth much smaller than the cavity linewidth), and that the polarization is the same as that of the cavity mode.

**ACKNOWLEDGMENTS**

We thank B. Gayral and M. Richard for fruitful discussions and M. Rigault for his help and enthusiasm in the initial stages of the measurements. The authors acknowledge financial support from the IST-FET European project QAP (Contract No. 15848) and by the ANR project QSWITCH.
Randomized benchmarking of quantum gates

E. Knill,* D. Leibfried, R. Reichle,† J. Britton, R. B. Blakestad, J. D. Jost, C. Langer,‡ R. Ozeri,§ and D. J. Wineland
National Institute of Standards and Technology, MC 891, 325 Broadway, Boulder, Colorado 80305, USA

(AReceived 26 July 2007; published 8 January 2008)

A key requirement for scalable quantum computing is that elementary quantum gates can be implemented with sufficiently low error. One method for determining the error behavior of a gate implementation is to perform process tomography. However, standard process tomography is limited by errors in state preparation, measurement and one-qubit gates. It suffers from inefficient scaling with number of qubits and does not detect adverse error-compounding when gates are composed in long sequences. An additional problem is due to the fact that desirable error probabilities for scalable quantum computing are of the order of 0.0001 or lower. Experimentally proving such low errors is challenging. We describe a randomized benchmarking method that yields estimates of the computationally relevant errors without relying on accurate state preparation and measurement. Since it involves long sequences of randomly chosen gates, it also verifies that error behavior is stable when used in long computations. We implemented randomized benchmarking on trapped atomic ion qubits, establishing a one-qubit error probability per randomized π/2 pulse of 0.00482(17) in a particular experiment. We expect this error probability to be readily improved with straightforward technical modifications.

DOI: 10.1103/PhysRevA.77.012307 PACS number(s): 03.67.Lx, 03.67.Pp, 37.10.De, 37.10.Gh

I. INTRODUCTION

In principle, quantum computing can be used to solve computational problems having no known efficient classical solutions, such as factoring and quantum physics simulations, and to significantly speed up unstructured searches and Monte Carlo simulations [1–4]. In order to realize these advantages of quantum computing, we need to coherently control large numbers of qubits for many computational steps. The smallest useful instances of the above-mentioned algorithmic applications require hundreds of qubits and many millions of steps. A quantum computing technology that realistically can be used to implement sufficiently large quantum computations is said to be “scalable.” Current quantum computing technologies that promise to be scalable have demonstrated preparation of nontrivial quantum states of up to eight qubits [5] but it is not yet possible to apply more than a few sequential two-qubit gates without excessive loss of coherence. Although there have been experiments to determine the behavior of isolated gates applied to prepared initial states [5–15], there have been no experiments to determine the noise affecting gates in a general computational context.

An important challenge of quantum computing experiments is to physically realize gates that have low error whenever and wherever they are applied. Studies of fault-tolerant quantum computing suggest that in order to avoid excessive resource overheads, the probability of error per unitary gate should be well below 10−5 [16–18]. The current consensus is that it is a good idea to aim for error probabilities below 10−7. What experiments can be used to verify such low error probabilities? One approach is to use process tomography to establish the complete behavior of a quantum gate. This requires that the one-qubit “analysis” gates employed in the tomography have lower error than the bound to be established on the gate under investigation. If this requirement is met, process tomography gives much useful information about the behavior of the gate, but fails to establish that the gate will work equally well in every context where it may be required. In particular, because process tomography requires preparing known initial states and one-qubit analysis gates, there is no obvious way to determine how well a gate works when used as the 4th gate of a computation. To do so would require characterizing all the previous gates and, in some independent fashion, the analysis gates. Process tomography can also be very time consuming as its complexity scales exponentially with the number of qubits.

We propose a randomized benchmarking method to determine the error probability per gate in computational contexts. Randomization has been suggested as a tool for characterizing features of quantum noise in Ref. [19]. The authors propose implementing random unitary operators \( U \) followed by their inverses \( U^{-1} \). Under the assumption that the noise model can be represented by a quantum operation acting independently between the implementations of \( U \) and \( U^{-1} \), the effect of the randomization is to depolarize the noise. The average fidelity of the process applied to a pure initial state is the same as the average over pure states of the fidelity of the noise operation. (The latter average is known as the average fidelity and is closely related to the entanglement fidelity of an operation [20].) They also show that the average fidelity can be obtained with few random experiments. They then consider self-inverting sequences of random unitary operations of arbitrary length. Assuming that the noise can be represented by quantum operations that do not depend on the choice of unitaries, the fidelity decay of the sequence is

[Knill@boulder.nist.gov]

*Present address: University of Ulm, Ulm, Germany.
†Present address: Lockheed Martin, Huntsville, Alabama.
‡Present address: Weizmann Institute of Science, Rehovot, Israel.
§Present address: Institut Néel, Grenoble, France.
shown to represent the strength of the noise. Our randomized benchmarking procedure simplifies this procedure by restricting the unitaries to Clifford gates and by not requiring that the sequence is strictly self-inverting. An alternative approach to verifying that sequences of gates realize the desired quantum computation is given in Ref. [21]. In this approach, successively larger parts of quantum networks are verified by making measurements involving their action on entangled states. This “self-testing” strategy is very powerful and provably works under minimal assumptions on gate noise. It is theoretically efficient but requires significantly more resources and multisystem control than randomized benchmarking.

Our randomized benchmarking method involves applying random sequences of gates of varying lengths to a standard initial state. Each sequence ends with a randomized measurement that determines whether the correct final state was obtained. The average computationally relevant error per gate is obtained from the increase in error probability of the final measurements as a function of sequence length. The random gates are taken from the Clifford group [22], which is generated by \( \pi/2 \) rotations of the form \( e^{i\sigma x \pi/4} \) with \( \sigma \) a product of Pauli operators acting on different qubits. The restriction to the Clifford group ensures that the measurements can be of one-qubit Pauli operators that yield at least one deterministic one-bit answer in the absence of errors. The restriction is justified by the fact that typical fault-tolerant architectures (those based on stabilizer codes) are most sensitive to errors in elementary Clifford gates such as the controlled NOT. Provided the errors in these gates are tolerated, other gates needed for universality are readily implemented by purified state preparation [16,23]. Note that the results of Ref. [19] hold if the unitaries are restricted to the Clifford group, because the Clifford group already has the property that noise is depolarized. We believe that randomized benchmarking yields computationally relevant errors even when the noise is induced by, and depends on, the gates, as is the case in practice. The relevant assumptions are given in Sec. V.

Randomized benchmarking as discussed and implemented here gives an overall average fidelity for the noise in gates. To obtain more specific information, the technique needs to be refined. In Ref. [24], randomization by error-free one-qubit unitaries is used to obtain more detailed information about noise acting on a multiqubit system. Randomized benchmarking can be adapted to use similar strategies.

II. RANDOMIZED BENCHMARK OF ONE QUBIT

For one qubit, our randomized benchmarking procedure consists of a large number of experiments, where each experiment consists of a pulse sequence that requires preparing an initial quantum state \( \rho \), applying an alternating sequence of either major axis \( \pi \) pulses or identity operators (“Pauli randomization”) and \( \pi/2 \) pulses (“computational gates”), and performing a final measurement \( M \). The pulse sequence between state preparation and measurement begins and ends with \( \pi \) pulses. For one qubit, the initial state is \( |0\rangle \). Because the major axis \( \pi \) and \( \pi/2 \) rotations are in the Clifford group, the state is always an eigenstate of a Pauli operator during the pulse sequence. The Pauli randomization applies unitary operators (“Pauli pulses”) that are (ideally) of the form \( e^{i\pi \sigma b \pi/2} \), where the sign \( \pm \) and \( b=x,y,z \) are chosen uniformly at random and we define \( \sigma_0 \) to be the identity operator. For ideal pulses, the choice of sign determines only a global phase. However, in an implementation, the choice of sign can determine a physical setting that may affect the error behavior. The computational gates are \( \pi/2 \) pulses of the form \( e^{i\pi \sigma_i n_i} \), with \( n_i \). The sign and \( n_i \) are chosen uniformly at random, except for the last \( \pi/2 \) pulse, where \( n_i \) is chosen so that the final state is an eigenstate of \( \sigma_i \). The computational gates generate the Clifford group for one qubit. Their choice is motivated by the fact that they are experimentally implementable as simple pulses. The final measurement is a von Neumann measurement of \( \sigma_\mu \).

The length \( l \) of a randomized pulse sequence is its number of \( \pi/2 \) pulses. The \( \pi/2 \) pulses are considered to be the ones that advance a computation. The \( \pi \) pulses serve only to randomize the errors. One can view their effect as being no more than a change of the Pauli frame. The Pauli frame consists of a Pauli operator that needs to be applied to obtain the intended computational state in the standard basis [16]. We call the \( \pi/2 \) and Pauli pulse combinations randomized computational gates. In principle, we can determine a pulse error rate by performing \( N \) experiments for each length \( l = 1, \ldots, L \) to estimate the average probability \( p_l \) of the incorrect measurement outcome (or “error probability”) for sequences of length \( l \). The relationship between \( l \) and \( p_l \) can be used to obtain an average probability of error per pulse. Suppose that all errors are independent and depolarizing. Let the depolarization probability of an operation \( A \) be \( \tilde{d} \) and consider a specific pulse sequence consisting of operations \( A_0, A_1 A_2, \ldots, A_{2l-1} A_{2l} \), where \( A_0 \) is the state preparation, \( A_1 \), \( A_2 \) and the following pairs are the randomized computational gates, and \( A_{2l-1} \) the measurement. For the measurement, we can assume that the error immediately precedes a perfect measurement. The state after \( A_l \) is a known eigenstate of a Pauli operator or completely depolarized. Depolarization of the state is equivalent to applying a random Pauli or identity operator, each with probability 1/4. The probability of the state’s not having been depolarized is \( \Pi_{l=0}^{l=L} (1 - \tilde{d}^l) \). In particular, we can express \( p_l = \mathbb{E}[1 - (1 - d)^l] / 2 \), where the function \( \mathbb{E}(\cdots) \) gives the expectation over the random choices of the \( A_i \). The factor of 1/2 in the expression for \( p_l \) arises because depolarization results in the correct state 1/2 of the time. The choices of the \( A_i \) are independent except for the last \( \pi/2 \) pulse. Assume that the depolarization probability of the last \( \pi/2 \) pulse does not depend on the previous pulses. We can then write \( p_l = \mathbb{E} \left[ 1 - (1 - d)^{l} \right] / 2 \), where \( d \) is the average depolarization probability of a random combination of one \( \pi/2 \) and one Pauli pulse (a randomized computational gate) and \( d_l \) combines
the depolarization probabilities of the preparation, initial Pauli pulse, and measurement. Thus $p_l$ decays exponentially to 1/2, and the decay constant yields $d$.

A commonly used metric to describe the deviation of an implemented gate from the intended gate is the average fidelity $F_\alpha$, which is defined as the uniform average over pure input states of the fidelity of the output state with respect to the intended output state. We are interested in the average computationally relevant error per step consisting of a randomized computational gate ("average error" for short). This is given by the expectation over gates of $1 - F_\alpha$ and relates to the depolarization parameter $d$ of the previous paragraph by $1 - F_\alpha = d/2$. In our implementation of the randomized computational gates, the $\pi$ pulses around the $z$ axis are implemented by changes in rotating frame and do not involve actively applying a pulse. Therefore, on average, the angular distance of the randomized gate’s action is $\pi$. As a result, $(1 - d/2)$ represents the average fidelity of pulses with action $\pi$.

Although estimates of $p_l$ are sufficient to obtain the average error for a randomized computational gate, it is useful to consider the error behavior of specific randomized computations and even fixed instances of the randomized sequences. For this purpose, the sequences are generated by first producing $N_\sigma$ random sequences consisting of $L$ random computational gates, where the gates are chosen independently without considering the final state. These sequences are considered to be a sample of typical computations. Each sequence is then truncated at different lengths. For each length, a $\pi/2$ pulse is appended to ensure that the final state is an eigenstate of $\sigma_z$. The sign of this final pulse is random. The resulting sequences are randomized by inserting the random Pauli pulses. We can then perform experiments to determine the probability of incorrect measurement outcomes for each such sequence and for each truncated computation after randomization by Pauli pulses. To be specific, the procedure is implemented as follows.

**Randomized benchmarking for one qubit.** This obtains measurement statistics for $N_\xi N_\gamma N_\eta N_\zeta$ experiments, where $N_\xi$ is the number of different computational gates, $N_l$ is the number of lengths to which the sequences are truncated, $N_\eta$ is the number of Pauli randomizations for each gate sequence, and $N_\zeta$ is the number of experiments for each specific sequence.

1. Pick a set of lengths $l_1 < l_2 < \ldots < l_{N_l}$. The goal is to determine the probability of error of randomized computations of each length.
2. Do the following for each $j = 1, \ldots, N_G$.
   1. Choose a random sequence $G = \{G_1, \ldots\}$ of $l_N$ computational gates.
   2. For each $k = 1, \ldots, N_l$ do the following.
      1. Determine the final state $p_j$ obtained by applying $G_{l_j} \ldots G_1$ to $|0\rangle$, assuming no error.
      2. Randomly pick a final computational gate $R$ among the two $\pi/2$ pulses that result in an eigenstate of $\sigma_z$ when applied to $p_j$. Record which eigenstate is obtained.

(2b3) Do the following for each $m = 1, \ldots, N_P$.

(2b3a) Choose a random sequence $T = \{P_1, \ldots\}$ of $l_k + 2$ Pauli pulses.

(2b3b) Experimentally implement the pulse sequence that applies $P_{l_k+2} R P_{l_k+1} G_{l_k} \ldots G_1 P_1$ to $|0\rangle$ and measures $\sigma_\gamma$, repeating the experiment $N_\zeta$ times.

(2b3c) From the experimental data and the expected outcome of the experiments in the absence of errors [from step (2b2) and the chosen Pauli pulses], obtain an estimate $p_{j/l_k/m}$ of the probability of error. Record the uncertainty of this estimate.

The probabilities of error $p_l$ are obtained from the $p_{j/l_k/m}$ by averaging $p_l = \frac{\sum_{j=1}^{N_G} \sum_{m=1}^{N_P} p_{j/l_k/m}}{N_\eta N_\zeta}$. We also obtain the probabilities of error for each computational gate sequence $p_{l_k/m} = \frac{\Sigma_{m=1}^{N_P} p_{j/l_k/m}}{N_\eta}$. If the errors are independent and depolarizing, the $p_{j/l_k/m}$ and the $p_{l_k/m}$ should not differ significantly from the $p_{l_k}$. (Significance is determined by the statistical error in the measurement of the $p_{j/l_k/m}$ by the $N_\eta$ experimental repetitions and the method for inferring expectations of $\sigma_z$ from the actual measurements performed.) However, if the errors are systematic in the sense that each implemented pulse differs from the ideal pulse by a pulse-dependent unitary operator, this can be observed in the distribution of the $p_{l_k/m}$ over $m$. In this case, the final state of each implemented pulse sequence is pure. The deviation of these pure states from the expected states is distributed over the Bloch sphere as $m$ and $j$ are varied. For example, consider the case where $p_l$ is close to 1/2. If the errors are systematic, the $p_{j/l_k/m}$ are distributed as the probability amplitude of $|1\rangle$ for a random pure state. In particular, we are likely to find many instances of $j$ and $m$ where $p_{j/l_k/m}$ is close to 0 or 1, that is, differs significantly from 1/2. In contrast, if the error is depolarizing, the $p_{j/l_k/m}$ are all close to 1/2 independent of $j$ and $m$.

III. TRAPPED-ION-QUBIT IMPLEMENTATION

We determined the computational relevance error probabilities for computational gates on one qubit in an ion trap. The qubit was represented by two ground-state hyperfine levels of a $^9$Be$^+$ ion trapped in a linear radio-frequency Paul trap briefly described in Ref. [25]. It is the same trap that has been used in a several quantum information processing experiments [26–30]. The two qubit states are $|\downarrow\rangle$ $(F = 2, m_F = -2)$ and $|\uparrow\rangle$ $(F = 1, m_F = -1)$, where for our purposes, we identify $|\downarrow\rangle$ with $|0\rangle$ and $|\uparrow\rangle$ with $|1\rangle$. The state $|\downarrow\rangle$ is prepared by optical pumping, after laser cooling the motional states of the ion. We can distinguish between $|\downarrow\rangle$ and $|\uparrow\rangle$ by means of state-dependent laser fluorescence. Computational gates and Pauli pulses involving $x$- or $y$-axis rotations were implemented by means of two-photon stimulated Raman transitions. To ensure that the pulses were not sensitive to the remaining excitations of the motional degrees of freedom, we used copropagating Raman beams. It was therefore not necessary to cool to the motional ground state and only Doppler cooling was used. Pulses involving $z$-axis rotations were
FIG. 1. Fidelity as a function of the number of steps for each randomized sequence. The fidelity (given by 1 minus the probability of error) is plotted on a logarithmic scale. The fidelity for the final state is measured for each randomized sequence. There are 32 points for each number of steps, corresponding to $N_p=8$ randomizations of each of $N_C=4$ different computational sequences. We used $N_p=8160$ for these experiments. Different symbols are used for the data for each computational sequence. The standard error of each point is between 0.001 (near fidelities of 1) and 0.006 (for the smaller fidelities). The scatter greatly exceeds the standard error, suggesting that coherent errors contribute significantly to the loss of fidelity.

implemented by programmed phase changes of one of the Raman beams. This changes the phase of the rotating reference frame and is equivalent to the desired $z$-axis rotation. The $z$-axis rotations were accompanied by a delay equivalent to the corresponding $x$ and $y$ pulses.

The Raman beams were switched on and off and shifted in phase and frequency as necessary by means of acousto-optic modulators controlled by a field-programmable gate array (FPGA). The pulse sequences were written in a special-purpose pulse-programming language and precompiled onto the FPGA. The version of the FPGA in use for the experiments was limited to about 100 computational pulses. The longest sequence in our experiments consisted of 96 computational gates. Our initial implementations clearly showed the effects of systematic errors in the distribution of the error probabilities of individual sequences. This proved to be a useful diagnostic and we were able to correct these systematics to some extent. One of the largest contributions to systematic errors was due to Stark shifts. To correct for these shifts, we calibrated them and adjusted phases in the pulse sequences.

IV. EXPERIMENTAL RESULTS

We generated $N_C=4$ random computational sequences and truncated them to the $N_f=17$ lengths $\{2,3,4,5,6,8,10,12,16,20,24,32,40,48,64,80,96\}$. Each truncated sequence was Pauli randomized $N_p=8$ times. Each final pulse sequence was applied to an ion a total of 8160 times in four groups that were interleaved with the other experiments in a randomized order. Pulse durations, qubit-resonant frequencies and Stark shifts were recalibrated automatically at regular intervals. The number of experiments per pulse sequence was sufficient to obtain the probability of incorrect measurement outcome with a statistical error small compared to the variation due to randomization and systematic errors. Figure 1 plots the fidelity (one minus the probability of incorrect measurement outcome) of each of the $4 \times 17 \times 8 = 544$ final pulse sequences against the length of the corresponding computational sequence. As explained in the figure caption, the variation in fidelity for each length shows that nondepolarizing errors contribute significantly to error. Figure 2 plots the average fidelity over the eight Pauli randomizations of each computational sequence truncated to the different lengths. Pauli randomization removes coherent errors, significantly reducing the variation in fidelities for different computational sequences. The remaining variation could be due to the small sample of eight Pauli randomizations used to obtain the average. The empirical average probability of error per randomized computational gate can be obtained by fitting the exponential decay and was found to be 0.00482(17). The fit was consistent with a simple exponential decay, which suggests that these gates behave similarly in all computational contexts. The error bars represent standard deviation as determined by nonparametric bootstrapping [31]. In what follows, if the fits are good, error bars are determined from nonlinear least-squares fits. In the cases where we can obtain a useful estimate of an error per randomized computational gate but the fits are poor, we used nonparametric bootstrapping.

For our experimental setting, it is possible to perform experiments to quantify the different types of errors as a consistency check. The results of these experiments are in the Appendix and are consistent with the randomized benchmarking data.

V. THEORETICAL CONSIDERATIONS

The average error per randomized computational gate is obtained by fitting an exponential. For general error models,
it is possible that the initial behavior of the measured error probabilities does not represent the average error of interest, and it is the eventual decay behavior that is of interest. In this case, randomized benchmarking determines an asymptotic average error probability (AAEP) per randomized computational gate. It is desirable to relate the empirical AAEP to the average error probability (AEP) of a single randomized computational gate. As discussed above, the AAEP agrees with the AEP if the error of all operations is depolarizing and independent of the gates. It can be seen that for depolarizing errors, this relationship holds even if the error depends on the gates. In general, one can consider error models with the following properties.

**Memoryless errors.** The errors of each gate are described by a quantum operation. In particular, the “environment” for errors in one gate is independent of that in another.

**Independent errors.** For gates acting in parallel on disjoint qubits, each gate’s errors are described by a quantum operation acting on only that gate’s qubits.

**Stationary errors.** The errors depend only on the gate, not on where and when in the process the error occurs.

**Subsystem preserving errors.** The errors cause no leakage out of the subsystem defining the qubits.

Although the AAEF need not be identical to the AEP, we conjecture that there are useful bounds relating the two error probabilities. In particular, if the AAEF is zero then there is a fixed logical frame in which the AEP is zero. Trivially, if the AEP is zero, then the AAEF is zero.

Randomized benchmarking involves both Pauli randomization and computational gate randomization. The expected effect of Pauli randomization is to ensure that, to first order, errors consist of random (but not necessarily uniformly random) Pauli operators. Computational gate randomization ensures that we average errors over the Clifford group. If, as in our experimental implementation, the computational gates generate only the Clifford group, it takes a few steps for the effect to be close to averaging over the Clifford group. This process is expected to have the effect of making all errors equally visible to our measurement, even though the measurement is fixed in the logical basis and the last step of the randomized computation is picked so that the answer is deterministic in the absence of errors.

**VI. BENCHMARKING MULTIPLE QUBITS**

Scalable quantum computing requires not only having access to many qubits, but also the ability to apply many low-error quantum gates to these qubits. The error behavior of gates should not become worse as the computation proceeds. Randomized benchmarking can verify the ability to apply many multiqubit gates consistently.

Randomized benchmarking can be applied to two or more qubits by expanding the set of computational gates to include multiqubit gates. The initial state is $|0 \cdots 0\rangle$. Pauli randomization is performed as before and is expected to convert the error model to probabilistic Pauli errors to first order. Because the size of the Clifford group for two or more qubits is large, one cannot expect to effect a random Clifford group element at each step. Instead, one has to rely on rapid mixing of random products of generators of the Clifford group to achieve (approximate) multiqubit depolarization. The number of computational steps that is required for approximate depolarization depends on the computational gate set. An example of a useful gate set consists of controlled NOTs (alternatively, controlled sign flips) combined with major-axis $\pi/2$ pulses on individual qubits. By including sufficiently many one-qubit variants of each gate, one can ensure that each step’s computational gates are randomized in the product of the one-qubit Clifford groups. This helps by having the effect of equalizing the probability of Pauli product errors of the same weight (see Ref. [24]).

The one-qubit randomized benchmark has a last step that ensures a deterministic answer for the measurement. For $n > 1$ qubits, one cannot expect deterministic answers for each qubit’s measurement, as this may require too complex a Clifford transformation. Instead, one can choose a random Pauli product that stabilizes the last state and apply a random product of one-qubit $\pi/2$ pulses with the property that this Pauli product is turned into a product of $\sigma_z$ operators. If there is no error, measuring $\sigma_z$ for each qubit and then computing the appropriate parity of the measurement outcomes gives a known deterministic answer. With error, the probability of obtaining the wrong parity can be thought of as a one-qubit error probability $p$ for the sequence. If the error is completely depolarizing on all qubits, with depolarization probability $d$, then $p=d/2$, just as for one qubit. One expects that for sufficiently long sequences, $p$ increases exponentially toward 1/2 so that the asymptotic average error probability per randomized computational gate can be extracted as for one qubit.

![FIG. 3. Measurement of phase decoherence with refocusing. We measured the probability of $|1\rangle$ as a function of time for the standard refocused decoherence measurement. The pulse sequence consisted of a $\pi/2$ pulse at phase 0 followed by a delay of $T/2$, a $\pi$ pulse at phase $\pi$, another delay of $T/2$ and a final $\pi/2$ pulse at phase $\pi$. The straight line shows the fit for exponential decay on the interval from 1 to 200 $\mu$s. Its extrapolation to larger times is shown dashed. The deviation from an exponential decay at larger times can be attributed to slow phase drifts that are no longer refocused by the single $\pi$ pulse in the pulse sequence. From the fit, the contribution of unrefocusable phase decoherence to the error probability per step is 0.0037(1). The standard deviation of the plotted points ranges from 0.002 for values near 1 to 0.008 for the smallest values, similar to the apparent scatter of the plotted points.](012307-5)
ACKNOWLEDGMENTS

We thank K. Coakley, S. Glancy, and R. Epstein for their comments on the manuscript. Our work was supported by the Disruptive Technology Office (DTO) under Contract No. 712868 and by the National Institute of Standards and Technology (NIST). It is a contribution of NIST, an agency of the U.S. government, and is not subject to U.S. copyright.

APPENDIX: DIRECT ERROR CHARACTERIZATIONS

We performed experiments to directly quantify the different types of errors in our pulses. These experiments characterize only the initial error (the error of the first gate) and serve as a consistency check for the randomized benchmarking data.

Known sources of errors include (a) phase errors due to fluctuating magnetic fields and changes in path length between the two Raman beams (they are merged on a polarizer), (b) amplitude errors due to changes in beam position at the ion and intensity fluctuations not being sinusoidal. Since the contribution to the probability of error is derived from the short-time behavior, the effect of the different models on the inferred probability of error per step is small. For the fit shown, the inferred contribution to the probability of error per step is 0.0090(7), larger than the error per step derived from Fig. 2. This is likely due to the fact that in the randomized sequences, the centering of the explicit $\pi$ pulses in their intervals reduces this contribution by refocusing.

FIG. 4. Measurement of phase decoherence without refocusing. The randomized benchmark does not systematically refocus changes in frequency. To estimate the contribution to error from decoherence including refocusable decoherence, we performed the experiment of Fig. 3 without the refocusing pulse. This is essentially an on-resonance Ramsey experiment. It was not experimentally possible to eliminate the oscillatory shape of the curve by calibrating the frequency indicating that the oscillation was not simply caused by detuning from the resonant frequency. However, the shape is similar to what one would expect from a roughly periodic change in frequency that is not synchronized with the experiment. Such changes could come from magnetic field fluctuations and phase noise due to air currents in the paths of the two Raman beams. To estimate the contribution to the probability of error per step, we fitted an exponentially decaying cos$(t)$ curve to the points with time coordinates less than 220 $\mu$s. The extrapolation of the fitted curve (dashed) clearly deviates from the data. Note that for sinusoidal phase noise, the curve should be related to a decaying Bessel function. Fits to such a function also deviate from the experimental data, consistent with the phase fluctuations not being sinusoidal. Since the contribution to the probability of error is derived from the short-time behavior, the effect of the different models on the inferred probability of error per step is small. For the fit shown, the inferred contribution to the probability of error per step is 0.0090(7), larger than the error per step derived from Fig. 2. This is likely due to the fact that in the randomized sequences, the centering of the explicit $\pi$ pulses in their intervals reduces this contribution by refocusing.

FIG. 5. Contribution of spontaneous emission to phase decoherence. To experimentally determine the contribution of spontaneous emission to decoherence, we applied the two Raman beams separately for half the time of each arm of the refocused decoherence measurement and compared the resulting data to that of Fig. 3 [32]. The points shown here were obtained by dividing the probabilities measured by the corresponding probabilities of Fig. 3, interpolating between the nearest points to match the time coordinates. The straight line shows the fitted exponential decay. The fit was weighted and used linear approximation to determine standard deviations of the points. The standard deviations used range from 0.003 to 0.015, which is substantially less than the apparent scatter of the plotted points. The inferred contribution to the error probability per step is 0.00038(3). This contribution can be estimated theoretically [32], which for the relevant configuration gives a value of approximately 0.0003.

FIG. 6. Rabi flopping experiment. To determine the contribution to the probability of error per step due to pulse area error and associated decoherence, we performed a Rabi flopping experiment. We fitted the points to a decaying cosine curve with a possible phase offset and both linear and quadratic decay. Again, we restricted the fit to an initial segment of the data (black curve). The extrapolation (dashed curve) shows significant deviations. The random uncertainty in the points ranges from 0.002 to 0.007, less than the symbol size of the plotted points. The apparent scatter in the points near the end of the curve is likely due to slow fluctuations in pulse amplitude. The contribution to the probability of error per step as detected in this experiment is 0.006(3) if the calibration were based on this experiment. Automatically calibrated pulse times fluctuated by around 0.02 $\mu$s. For pulse times differing by this amount, the contribution to the error per step is 0.007(3).

We thank K. Coakley, S. Glancy, and R. Epstein for their comments on the manuscript. Our work was supported by the Disruptive Technology Office (DTO) under Contract No. 712868 and by the National Institute of Standards and Technology (NIST). It is a contribution of NIST, an agency of the U.S. government, and is not subject to U.S. copyright.
from the upper levels required for the stimulated Raman transition.

Phase decoherence can be measured by observing the decay of signal in a Ramsey spectrometry experiment of the qubit with or without refocusing [32]. Figure 3 shows the probability of observing $|1\rangle$ at the end of a refocused Ramsey experiment as a function of the delay between the first and last $\pi/2$ pulse. By fitting the initial part of the curve to an exponential decay, one can infer the contribution of unrefocussable phase error to each step of the Pauli randomized sequences. We obtained an estimate of 0.0037 (1) for this contribution. Figure 4 shows the probability of observing $|1\rangle$ in a similar experiment but with the refocusing pulse omitted. This is an on-resonance Ramsey experiment. The fit suggests a contribution of 0.0090 (7) for the error per step. This is larger than the inferred error from the randomized experiments, which can be explained by the refocusing effects of the Pauli randomization. See the caption of Fig. 4 for a discussion of fitting issues. We note that our benchmarking experiments, as well as the error characterizations in this section, were performed without line triggering the experiments, thereby making them sensitive to phase shifts caused by 60 Hz magnetic field fluctuations. Greatly improved decoherence times are typically obtained if such triggering is used.

The contribution of spontaneous emission to phase decoherence can be determined by a refocused Ramsey experiment where the two Raman beams are on separately half the time during the intervals between the pulses [32]. To determine the desired contribution, the probability of $|1\rangle$ as a function of time is compared to the data shown in Fig. 3. The results of the comparisons are in Fig. 5. The inferred contribution to the error probability per step is 0.00038 (3), well below the contribution of the other sources of error.

The effect of amplitude fluctuations can be estimated from the loss of visibility of a Rabi flopping experiment. The data are shown in Fig. 6. Modeling the Rabi flopping curve is nontrivial and the fits are not very good. Nevertheless, we can estimate a contribution to the error probability per step from the behavior of the curve during the first few oscillations. This gives a contribution of 0.006 (3), consistent with the probability of error per step obtained in the randomized experiments. Note that the contribution measured here also includes errors due to phase fluctuations during the computation pulses.

Fluorescence during Doppler cooling of a single trapped atom


National Institute of Standards and Technology, Boulder, Colorado 80305, USA

(Received 9 July 2007; published 26 November 2007)

We investigate the temporal dynamics of Doppler cooling of an initially hot single trapped atom in the weak-binding regime using a semiclassical approach. We develop an analytical model for the simplest case of a single vibrational mode for a harmonic trap, and show how this model allows us to estimate the initial energy of the trapped particle by observing the time-dependent fluorescence during the cooling process. The experimental implementation of this temperature measurement provides a way to measure atom heating rates by observing the temperature rise in the absence of cooling. This method is technically relatively simple compared to conventional sideband detection methods, and the two methods are in reasonable agreement. We also discuss the effects of rf micromotion, relevant for a trapped atomic ion, and the effect of coupling between the vibrational modes on the cooling dynamics.

DOI: 10.1103/PhysRevA.76.053416 PACS number(s): 32.80.Lg, 32.80.Pj, 42.50.Vk

1. INTRODUCTION

Laser cooling of trapped neutral atoms and atomic ions is a well-established technique: for example, cooling to the motional ground state [1–3] and motional state tomography [4] are routinely performed with resolved motional-sideband excitation techniques. Sideband techniques require the natural linewidth $\Gamma$ of the cooling transition to be small compared to the vibrational frequency of the trapped particle, in order to allow the motional sidebands to be resolved. Many experiments are, however, conducted in the “weak-binding regime,” where $\Gamma$ is larger than the oscillation frequency. Here, the cooling process is essentially the same as Doppler cooling of free atoms, because the spontaneous decay process is short compared to the atom’s oscillation period [5]. Even in experimental setups that implement sideband techniques, an initial stage of such “Doppler cooling” is often employed. The first examinations of Doppler cooling of trapped ions [5–11] did not take into account the effects of micromotion due to the trapping rf field. After cooling and heating effects related to micromotion were observed, these effects were explained theoretically [12–14] by including the effects of micromotion.

Here, we consider Doppler cooling of a single trapped atom or ion. While most previous work has focused on the final stages of cooling, our focus will be on the temporal dynamics of the cooling, particularly in the “hot regime” where the Doppler shift due to atom motion is comparable to or much larger than $\Gamma$. For the one-dimensional (1D) case we find that the cooling rate can be calculated analytically in the weak-binding regime without assuming the atom to be in the Lamb-Dicke regime. For a trapped ion, when we take rf micromotion into consideration, stable, highly excited states emerge when only one mode is considered [15]. When all three vibrational modes of the ion are considered, we find that couplings between the modes tend to break the stability of such points, allowing cooling to reach the Doppler limit.

A practical application of our results is to estimate the initial motional energy of an atom or ion from observations of the time dependence of the fluorescence during the cooling process. As mentioned above, sideband spectroscopy is the conventional technique for characterizing motional states, and it has been used to characterize the heating rate of ions in the absence of cooling [1,2,16–20]. However, it is more complicated to implement experimentally than Doppler cooling, requiring more laser beams. Currently, considerable effort is being devoted to understanding the heating observed in ion traps [16–20]. This heating is anomalous in that it is typically much higher than would be expected from thermal electronic noise (Johnson noise) that results from resistance associated with the trap electrodes. Because of its pervasive nature and detrimental effects to quantum state manipulation, it is important to understand and eliminate its cause. A less complicated technique for measuring temperature could simplify this work.

We note that cooling of a single ion is significantly different than cooling a large number of ions in the same trap where the degrees of freedom are strongly mixed. However, the cooling and heating of a single ion is interesting because heating can only arise from external fields, not from the effects of ion-ion collisions that can lead to heating for two or more ions [21].

This paper is structured as follows. In Sec. II we present a semiclassical model of the Doppler cooling process for a bound atom confined in one dimension in the weak-binding regime in a static harmonic well (no micromotion). In Secs. III and IV we analyze the fluorescence vs time predicted by the model. Here, we consider single cooling trajectories and average over these with a given distribution of initial motional energies. We derive expressions useful for estimating initial temperature from fluorescence observations in these sections. Section V discusses how to minimize the total measurement time required to estimate the mean initial energy.

---

*janus.wesenberg@nist.gov
†Present address: Areté Associates, Longmont, CO 80501, USA.
‡Present address: Lockheed Martin, Huntsville, AL, USA.
§Present address: Weizmann Institute of Science, Rehovot, Israel.
°Present address: Institut Néel, Grenoble, France.
In Sec. VI we consider the effects of other motional modes with and without taking into account any rf micromotion experienced by such modes. The primary result of this section is most relevant for ions confined in miniature linear Paul traps [22,23]. In these traps, ions are bound along the axis by a static harmonic well. In the transverse direction, ions are bound by a ponderomotive pseudopotential where the ions’ motion must also include the effects of rf micromotion. Typically, these traps are operated under conditions where the ions’ oscillation frequencies in the transverse direction are much higher than along the axis. Moreover, it is typically observed that the anomalous heating as a function of oscillation frequency $\omega$ drops off as $1/\omega$ or faster. Therefore, to a good approximation the ion heating is dominated by that along the axis of the trap. In this case the results of Secs. II–IV will be valid to a good approximation. These conclusions are supported by a more complete solution of the problem described in Sec. VI C. Section VII suggests possible modifications to the basic experimental protocol that might provide improved sensitivity of the temperature measurements. Section VIII concludes the paper.

II. MODEL

We consider a semiclassical model of Doppler cooling of a single weakly trapped atom [5,11]. We will initially consider only a single mode of motion, taken to be along the $z$ direction. We assume a harmonic potential with oscillation frequency $\omega$. In Sec. VI we consider a more detailed model that includes three dimensions and micromotion relevant for trapped ions.

The atom is Doppler cooled by a single laser beam of angular frequency $\nu_{\text{las}}$ and wave vector $k$, detuned by $\Delta = \nu_{\text{las}} - \nu_{\text{g}}$ from the resonance frequency $\nu_{\text{g}}$ of a two-level, or “cycling,” transition between two internal states $|g\rangle$ and $|e\rangle$, of the atom. We write the coupling Hamiltonian as

$$H^{(c)} = \hbar \Omega_{\text{Rabi}} (|e\rangle \langle g| + |g\rangle \langle e|) \cos(k \cdot x - \nu_{\text{las}}t),$$

where $x$ is the atom position, $2\pi\hbar$ is Planck’s constant, and $\Omega_{\text{Rabi}}$ is the resonant Rabi frequency.

We assume the atom is weakly bound in the $z$ direction, that is, $\omega_z$ is much smaller than the excited state decay rate $\Gamma$. The atom’s level populations are then approximately in steady state with respect to the instantaneous effective detuning $\Delta_{\text{eff}} = \Delta + \Delta_0$, including the Doppler shift $\Delta_0 = -v_z k_z$, where $v_z$ and $k_z$ are the $z$ components of the velocity and wave vector. The excited state population is then [24]

$$\rho_{e}(v_z) = \frac{s/2}{1 + s + (2\Delta_{\text{eff}}/\Gamma)^2}, \quad (2)$$

Here $s$ is the saturation parameter, proportional to the cooling beam intensity, $s = 2(\Omega_{\text{Rabi}}/\Gamma)^2$.

The excited state population is associated with the photon scattering rate $dN/dt$ by the relation $dN/dt = \Gamma \rho_{e}(v_z)$. While the momentum kicks associated with photon emission are assumed to average to zero over many absorption-emission cycles, the absorbed photons will impart a velocity-dependent momentum transfer due to the scattering that can be described by a velocity-dependent force

$$F_z(v_z) = m \frac{dv_z}{dt} = \hbar k_z \Gamma \rho_{e}(v_z), \quad (3)$$

where $m$ is the atom’s mass. This velocity-dependent force will in general change the motional energy $E$ of the atom. If the relative change in energy over a motional cycle is small, we can average the effect of $F_z$ over the oscillatory motion to find the evolution of $E$:

$$\frac{dE}{dt} = \langle v_z F_z(v_z) \rangle, \quad (4)$$

where the average is over one motional cycle. The average energy change per scattering event is $dE/dN = h\nu_{\text{g}} v_z = -\hbar \Delta_0$. In addition to $F_z$, the atom will experience a stochastic force due to photon recoil that, assuming isotropic emission, will cause heating at a rate [11]

$$\left( \frac{dE}{dt} \right)_{\text{recoil}} = \frac{4(\hbar k_z)^2}{3m} d\frac{N}{dt}, \quad (5)$$

where $(\hbar k_z)^2/2m$ is the recoil energy associated with the scattering. We will mostly ignore the effects of recoil heating in what follows, since it will be important only near the cooling limit.

III. ANALYSIS

We will now analyze the time dependence of the atom fluorescence during the Doppler cooling process, as predicted by the model introduced above.

To simplify the algebra, we will scale energies by $\hbar$ times half the power-broadened linewidth, and time by the resonant scattering rate:

$$e = E/E_0, \quad E_0 = \frac{\hbar \Gamma}{2} \sqrt{1 + s}, \quad (6a)$$

$$\delta = \hbar \Delta/E_0, \quad (6b)$$

$$r = \frac{(\hbar k_z)^2}{2m} / E_0, \quad (6c)$$

$$\tau = t/t_0, \quad t_0 = \left( \frac{\Gamma}{\sqrt{2}/(1 + s)} \right)^{-1}. \quad (6d)$$

As an example of typical values, we consider a trapped $^{25}\text{Mg}^+$ ion, where the $^{25}\text{Mg}^+ : F_{23/2} F_3/2$ cooling transition at 279.6 nm has a natural linewidth of $\Gamma = 2\pi \times 41.4$ MHz. At a detuning of $\Delta = -2\pi \times 20$ MHz with $s = 0.9$ and $k_z/k = 0.71$, we find that $E_0/k = 1.4$ mK, where $k$ is Boltzmann’s constant and $t_0 = 16$ ns. The detuning and recoil parameters are $\delta = -0.70$ and $r = 0.0018$.

For a given motional energy $e$, the instantaneous Doppler shift $\delta_0 = \hbar \Delta_0/E_0$ has a maximal value of $\delta_0 = 2\sqrt{s}$. In the scaled units, the average change in energy per scattering event is $-\delta_0$, so $\delta_0$ is also equal to the maximal change in energy per scattering event. As explained below, the energy at which the maximal Doppler shift is equal to the power-
broadened linewidth, $e=1/r$, is of interest during the cooling process. For reference we note that this energy corresponds to

$$
\frac{1}{r}E_0 = \frac{\hbar^2 \Gamma^2}{4 \hbar^2 k_r^2}.
$$

(7)

For the typical experimental parameters considered above, $E_0/r$ is equal to $k_B \times 0.7$ K or $3700 \hbar \omega_c$ for $\omega_c=2\pi \times 4.0$ MHz.

For harmonic oscillations, the instantaneous Doppler shift $\delta_D$ is distributed according to the probability density

$$
P_D(\delta_D; \delta_0) = \int_0^{2\pi} \delta_{\text{Dirac}}[\delta_D - \delta_M \sin(d)] \frac{d\phi}{2\pi}
$$

$$
= \begin{cases} 
\frac{1}{\pi \sqrt{\delta_M^2 - \delta_D}} & \text{if } |\delta_D| < \delta_M, \\
0 & \text{otherwise}, 
\end{cases}
$$

(8)

where $\delta_{\text{Dirac}}$ is the Dirac $\delta$ function. Examples of this probability distribution for two different values of $\delta_M$ are shown in the insets of Fig. 1. Since the average energy change per scattering event is $-\delta_D$, and the instantaneous scattering rate is $1/(1+\delta_D^2)$, where $\delta_M=\delta+\delta_D$, the rate of change of $e$ averaged over the secular oscillations, given by Eq. (4), takes the form

$$
\frac{de}{d\tau} = \int -\delta_D P_D(\delta_M; \delta_D) \frac{1}{1+(\delta+\delta_D)^2} d\delta_D.
$$

(9)

We can evaluate the integral as detailed in Appendix A, to find that

$$
\frac{de}{d\tau} = \frac{1}{2\sqrt{er}} [\text{Re}(Z) + \delta \text{Im}(Z)],
$$

(10a)

$$
\approx \frac{\delta}{2\sqrt{er}}, \quad e \gg (1 + \delta^2)/r,
$$

(10b)

where $Z=\text{Z}(\delta, \delta_M)=\sqrt{1-(\delta+\delta_M^2)^2}/4er$. The asymptotic approximation (10b) corresponds to approximating $P_D(\delta_M; \delta_D)$ by $P_D(\delta_D; 0)$, which is reasonable in the hot regime, where the peaks of $P_D$ have small overlap with the Lorentzian line profile. From Eqs. (10a) and (10b), we can evaluate $e$ as a function of time; an example is shown in the lower part of Fig. 1.

The scattering rate averaged over the motion is analogous to Eqs. (10a) and (10b) and is given by

$$
\frac{dN}{d\tau} = \int P_D(\delta_M; \delta_D) \frac{1}{1+(\delta+\delta_D)^2} d\delta_D = \frac{1}{2\sqrt{er}} \text{Im}(Z),
$$

(11)

as illustrated in the upper part of Fig. 1. In the limit of $e \gg (1 + \delta^2)/r$, we find $dN/d\tau = 1/(2\sqrt{er})$, so that according to Eq. (10b) we have in this limit $de/dN = \delta$. In physical units, this corresponds to each photon on average extracting an energy of $\hbar \Delta$. This can be understood by noting that, in the limit of $e \gg (1 + \delta^2)/r$, $P_D(\delta_M; \delta_D)$ is to a good approxi-

FIG. 1. Scattering rate (top) and energy (bottom) as a function of time during the Doppler cooling of a single atom from an initial energy $E_0$ [in scaled units, see Eq. (6a)] of $10_r/\tau$, as given by Eqs. (11), (10a), and (10b), at a detuning of half the power-broadened linewidth $\delta=1$; see Eq. (6b)]. The dotted curve in the lower plot shows the energy as a function of time predicted by the asymptotic approximation of Eq. (12). Insets in the upper plot show (at two different times) the two components appearing in the integral defining $dN/d\tau$ in Eq. (11), the probability density of the effective detuning $P_D(\delta_M; \delta_D)$ (solid) and the Lorentzian line profile, $L(\delta_M)=1/(1+\delta_M^2)$ (dashed), as functions of $\delta_M=\delta+\delta_D$ at $e=8/r$ and $e=e_c$ [Eq. (14)]. Scattering events where the atom is moving toward the laser so that $\delta_D>0$, corresponding to the rightmost peak of $P_D$, result in cooling, and vice versa. The energies of maximal cooling and scattering rates, $e_c$ and $e_c$, are given by Eqs. (13) and (14).

mation uniform over the Lorentzian line profile (see, for example, the upper left inset of Fig. 1), and so the value of $\delta_M=\delta+\delta_D$ averaged over the scattering events will be nearly zero. Since each scattering event extracts an energy of $-\delta_D$, the average cooling per scattering event should indeed be $\delta$.

The time dependence of $e$ is formally found by integrating $de/d\tau$ as given by Eqs. (10a) and (10b). For the asymptotic approximation (10b) we find

$$
e(\tau) = \left(\frac{e_0^3}{2} + \frac{3\delta^2}{4\sqrt{r}}\right)^{2/3}, \quad e \gg (1 + \delta^2)/r,
$$

(12)

where $e_0$ is the energy at $\tau=0$, as plotted in the lower part of Fig. 1. For the exact expression derived from Eq. (10a), we must resort to numerical methods to find $e(t)$. Nevertheless, we do find analytically that the cooling rate is maximal for $e$ related to $\delta$ by

$$
e = e_c = \frac{1 + \delta^2}{2r} \cos \left[ \frac{1}{3} \arccos \left( \frac{1 - \delta^2}{1 + \delta^2} \right) \right],
$$

(13)

which quantifies our previous observation that $1/r$ is a typical energy scale of the cooling process.
The behavior of $dN/d\tau$ is qualitatively different for $\delta$ being smaller or larger than a critical detuning $\delta_c = -1/\sqrt{3}$. For $\delta<\delta_c$, $dN/d\tau$ has a maximum at an energy $e = \bar{e}$, where

$$e = \bar{e} = \frac{1}{4\tau} \left( \delta - \sqrt{3} \right) \left( \delta + 1/\sqrt{3} \right). \tag{14}$$

For the example parameters listed below Eqs. (6a)–(6d), $\delta_c$ corresponds to a detuning of $\delta_c E_0/\hbar = -2\pi \times 16.5$ MHz. The maximal scattering rate is reached when one of the peaks of the Doppler distribution (8) is in resonance with the cooling transition, as illustrated in the upper right-hand inset in Fig. 1. In the regime where a maximum exists, the maximal scattering rate is found to exceed the steady state scattering rate by a factor of

$$\left. \frac{dN}{d\tau} \right|_{e = \bar{e}} = \frac{\sqrt{3/\delta} \left( 1 + \delta^2 \right)}{4} \left. \frac{dN}{d\tau} \right|_{e = 0}. \tag{15}$$

Closer to resonance, i.e., when $\delta_c < \delta < 0$, no maximum occurs, as illustrated in Fig. 2.

We emphasize that the only approximations made above are the weak-binding approximation and the omission of recoil heating. In particular, the trapped particle is not assumed to be in the Lamb-Dicke regime. For the weak-binding regime, $\delta_{\text{eff}} > 1$ implies that the motion is outside the Lamb-Dicke regime. To find the cooling rate predicted by Eqs. (10a) and (10b) in the Lamb-Dicke limit, we note that to first order in $e$ we find $\delta e/d\tau \approx 4\delta e/2(1 + \delta^2)$. This corresponds to $e$ decreasing exponentially with $\tau$. Except for the omission of recoil heating, the value of the decay time agrees with previous work that assumed the atom was in the Lamb-Dicke regime [5,25].

In the above analysis, we have ignored recoil heating [given by Eq. (5)]. In the limit of $e \gg \bar{e}$, the ratio of heating to cooling is seen to be $4\tau/3(\delta_0^2)$, which is a small fraction for typical experimental parameters. For $e < \bar{e}$, the cooling is less efficient and the contribution from recoil becomes more significant, leading to a nonzero steady-state energy.

Nevertheless, ignoring recoil heating is reasonable when observing fluorescence scattering, since the scattering rate has almost reached its steady state value when the effect of recoil becomes important. Therefore we have omitted recoil heating in this analysis to make $\delta$ the only free parameter and simplify the discussion. Recoil can be included in calculations by combining Eqs. (5), (10a), and (11).

IV. THERMAL AVERAGING

An application of the analysis presented above is to estimate the initial motional energy of a trapped atom from the time-dependent fluorescence observed during the cooling process. Using this method, we can estimate the average rate of heating experienced by a trapped atom in the absence of cooling by first allowing the atom to heat up without cooling for a certain period and then observing the time dependence of the fluorescence as the atom is recooled. As discussed in Sec. III, when $e \gg (1 + \delta^2)/\tau$, the average cooling per scattering event is $\delta$. The approximate total number of photons scattered during the cooling of an atom with initial motional energy $e_{\text{initial}}$ can consequently be approximated by $\bar{e} e_{\text{initial}}/\delta$.

For the example parameters given in Sec. III, this corresponds to $\approx 3200$ photons for $e_{\text{initial}} = 4/\tau$, corresponding to $k_B T = 2.8$ K. With typical photon detection efficiencies of less than $10^{-3}$, very few photons are registered in a single experiment. We must therefore repeat many experimental cycles consisting of a heating period and a cooling period.

We now consider the form of the fluorescence signal when averaged over many such experimental cycles. Here, we will assume that the heating is stochastic and take the distribution $P_\theta(e)$ of the motional energies at the beginning of each cooling period to be the Maxwell-Boltzmann distribution with mean energy $\bar{e}$,

$$P_\theta(e) = \frac{1}{\bar{e}} e^{-(e-\bar{e})/\bar{e}}. \tag{16}$$

However, the results below hold for any form of $P_\theta(e)$.

The thermally averaged scattering rate is conveniently written in terms of the propagator $\Xi$ of $e$: Let $\Xi(e_0, \tau)$ denote the energy at time $\tau$ of an atom with initial energy $e(\tau = 0) = e_0$. We can then write the thermally averaged scattering rate at time $\tau$ as

$$\left\langle \frac{dN}{d\tau} \right\rangle_\bar{e} = \int_0^\infty P_\theta(e') \frac{dN}{d\tau} \bigg|_{e = \Xi(e', \tau)} de'. \tag{17}$$

This can be efficiently computed numerically by noting that $\Xi(e(\tau_1), \tau_2) = \Xi(e, \tau_1 + \tau_2)$, as detailed in Appendix B. Figure 3 shows the thermally averaged scattering rate for a few different parameters.

The fluorescence predicted by Eq. (17) has been found to be in agreement with experimentally observed fluorescence. We show one experimental data set for comparison in Fig. 4; the experiments are more fully described in Ref. [20]. Furthermore, the resulting estimated heating rates have been found to agree reasonably well with results obtained using the Raman sideband technique [17,20]. This agreement may at first seem surprising, given that the two methods probe
very different energy scales and that the Raman sideband technique is not sensitive to collisional heating. For the measurements based on the Raman sideband technique the ion was allowed to heat for only a few milliseconds, thereby gaining a few motional quanta. In contrast, the measurement results presented in Fig. 4 are based on 25 s heating periods, allowing the ion to gain many motional quanta. However, the results should agree if, as is currently believed, the heating rate is dominated by the effects of stochastic field fluctuations originating at the electrodes [16]. In this case, we would not expect any thermalization effects before the ion temperature is comparable to that of the electrodes. We would also not expect the effective temperature of the electrodes to be below the ambient temperature (∼300 K).

V. OPTIMAL EXPERIMENTAL PARAMETERS

We now examine how the total measurement time required to reach a given accuracy in the heating rate estimate depends on the choice of experimental parameters. As the recoil parameter $r$ will be fixed by the choice of atom, we consider only the choice of optimal values for initial ion energy $\bar{e}$, cooling laser detuning $\delta$, and saturation parameter $s$.

The experimental signal during a single recooling cycle will consist of an initial deviation from the steady state scattering rate. For a given initial motional energy, both the size and duration of this deviation increase with decreasing detuning, as illustrated by Fig. 2. For a given experimental setup, the optimal detuning is decided as a compromise between recooling signal and ability to recool atoms that have been highly excited by, e.g., collisions.

For a given value of $\bar{e}$, the experimental signal, in terms of the number of photons scattered before steady state is reached, does not depend on the laser beam intensity. Since $\bar{e}$ is the average initial energy relative to $E_0$, which is proportional to the power-broadened linewidth, a lower laser beam intensity will give a larger signal for a given heating period. This suggests using the smallest feasible laser intensity, requiring a compromise with respect to robust cooling and detector dark counts. From this standpoint we want to keep the saturation parameter below, but probably close to, 1.

For a given detuning and laser intensity, an additional choice of the length of the heating period in each experimental cycle has to be made: Should we perform a relatively low number of cycles with long heating periods or more cycles with shorter heating periods? To answer this question, we estimate the total measurement time $T_{\text{tot}}$ required to reach a certain relative accuracy in the estimate of the heating rate. We assume a constant heating rate and assume that the total time is dominated by the heating periods, so that $T_{\text{tot}}$ is proportional to the average initial energy $\bar{E} \approx \bar{e} \sqrt{1 + s}$, and to the number of runs.

We consider a setup where the observed fluorescence is collected in sequential time bins that are short compared to the total time required for the cooling process. In the limit where the distribution of the integrated number of counts $n_i$ in time bin $i$ is described by a normal distribution with vari-

![Fig. 3](image-url)  
**FIG. 3.** (Color online) Thermally averaged scattering rate vs time, for $\delta = -\sqrt{3}$ (top) and $-1/\sqrt{3} = \delta_1$ (bottom). In both figures, $P_0$ is assumed to be a thermal distribution [Eq. (16)] with $\bar{e}$ equal to one, two, and four times $e_c$ for the solid, dashed, and dash-dotted lines, respectively. Note that, for $\delta < \delta_1 = -1/\sqrt{3}$, the initial fluorescence is larger than the steady state fluorescence for low values of $\bar{e}$; this is attributed to the local maximum in the fluorescence vs time as illustrated in Fig. 2.

![Fig. 4](image-url)  
**FIG. 4.** Experimentally observed fluorescence during Doppler cooling of a single $^{23}\text{Mg}^+$ ion [20] compared with the fluorescence predicted by the simple 1D model. Data points (triangles) indicate the observed scattering rates, obtained by integrating over many experiments. In each experiment, the time-resolved fluorescence is recorded during ion recooling, after the ion has been allowed to heat up for a period of 25 s. The experimental parameters are those given after Eqs. (6a)–(6d). Error bars are based on counting statistics. The solid curve is the scattering rate predicted by Eq. (17), assuming the motional energy of the ion after the heating period to be given by the Maxwell-Boltzmann distribution (16) with $\bar{e} = 5.1/\pi$, corresponding to a temperature of 3.9 K. Since $\bar{e}$ is the only free parameter of Eq. (17), the estimated value was extracted by a single-parameter fit, and agrees reasonably well with an independent temperature estimate of 3.4±0.3 K extrapolated from heating rates measured in the same trap by use of the Raman sideband technique [20].
ance $\sigma_\varnothing$, we can estimate the uncertainty on the maximum-likelihood estimate of $\bar{\varnothing}$ for a given data set by [26]

$$1/\sigma(\bar{\varnothing})^2 = \sum_i \left( \frac{\partial N_i}{\partial \bar{\varnothing}} \right)^2 / \sigma_i^2.$$  

(18)

It follows from Eqs. (10a), (11), and (17), that in the 1D case the cooling dynamics can be rewritten in a form independent of $r$ by reparametrizing in terms of $N_r, \bar{\varnothing}, r$, and $\varpi$. We will denote the reparametrized scattering rate by

$$\tilde{R}_\varnothing(\bar{\varnothing}, \varpi) = \left( \frac{\partial N_r}{\partial \varpi} \right)_{\bar{\varnothing} \varpi}.$$  

(19)

Since the relative uncertainty on the heating rate estimate is equal to $\sigma(\bar{\varnothing})/\bar{\varnothing}$ and $\sigma_i = \sqrt{N_i}$, Eqs. (18) and (19) allow us to estimate the time required to obtain a given relative uncertainty on the heating rate:

$$\frac{T_{\text{tot}}}{\sqrt{1 + s}} \propto \left[ \tilde{\varnothing} \int_0^\infty \left( \frac{\partial \tilde{R}_\varnothing(\bar{\varnothing}, \varpi)}{\partial \bar{\varnothing}} \right)^2 \frac{dq}{\tilde{R}_\varnothing(\bar{\varnothing}, \varpi)} \right]^{-1}.$$  

(20)

Note that the right-hand side depends only on $\delta$ and $\tilde{\varnothing}$. Figure 5 shows $T_{\text{tot}}/\sqrt{1 + s}$ calculated for different detunings. The figure confirms that a low detuning is indeed favorable, and also shows that, for a given detuning, $T_{\text{tot}}$ decreases with increasing $\tilde{\varnothing}$. This is not surprising, given that the time to extract a certain amount of energy increases with atom temperature, as illustrated by Eq. 1. From Fig. 5 and the definition of $\varnothing_c$ [Eq. (13)], we can see that the heating period should be chosen long enough to get a significant signal, i.e., $\tilde{\varnothing} > \varnothing_c$, but the optimal heating period must be decided based on other experimental parameters such as trap depth and background gas collision rate.

VI. COOLING IN THREE DIMENSIONS

So far, we have considered only cooling in one dimension. In this section we will consider the effect of the vibrational modes in other directions on the cooling process. Our goal is to gain a qualitative understanding of the effects of the transverse modes on the cooling dynamics of the $\varnothing$ mode, with the intent of establishing to what extent the simple 1D model presented above is a reasonable approximation.

A. 3D cooling of neutral atoms

For a neutral atom, the confinement transverse to $\varnothing$ is not associated with micromotion, as it is for ions in a linear Paul trap, and the 1D weak-binding model extends immediately to three dimensions. Following Sec. III, let $\varepsilon_i, i = \{x, y, z\}$, denote the motional energy in mode $i$, $\delta_{\varnothing i} = -\hbar k_i \nu / E_0$ the Doppler shift, and $\delta_{\varnothing i}^0 = 2\sqrt{\varepsilon_i r_i}$ the maximum Doppler shift. Although the cooling of all modes is formally identical in the absence of micromotion, we will discuss the cooling dynamics with a focus on the $\varnothing$ mode.

In experiments it is typically easy to make the frequencies of the three modes incommensurate, which we will assume here. In that case, we can write the rate of change of $\varepsilon_c$ as

$$\frac{d\varnothing_c}{dt} = \frac{-\delta_{\varnothing i}^0}{1 + (\delta + \sum_j \delta_{\varnothing j}^0)^2} \int P_D(\delta_{\varnothing i}^0, \delta_{\varnothing j}^0) d\delta_{\varnothing i} d\delta_{\varnothing j}.$$  

(21)

where $R_\varnothing$ is the effective line profile experienced by the $\varnothing$ mode, obtained by convolving the Lorentzian line profile with the distribution $P_D(\delta_{\varnothing i}^0, \delta_{\varnothing j}^0)$ of the combined Doppler shift $\delta_{\varnothing} = \delta_{\varnothing i}^0 + \delta_{\varnothing j}^0$ due to the $x$ and $y$ “spectator” modes,

$$P_D(\delta_{\varnothing i}^0, \delta_{\varnothing j}^0) = \int P_D(\delta_{\varnothing i}^0; u) P_D(\delta_{\varnothing j}^0; \delta_{\varnothing i}^0 - u) du.$$  

(22)

For the 3D case, $R_\varnothing$ replaces the Lorentzian factor $1/[1 + (\delta + \delta_{\varnothing}^0)^2]$ of Eq. (9). As illustrated in Figs. 6 and 7(a), $P_D(\varnothing; \delta_{\varnothing}^0)$ is peaked (diverges) at $\delta_{\varnothing}^0 = \pm \delta_{\varnothing M}^0$. If $|\delta_{\varnothing M}^0 - \delta_{\varnothing i}^0| \approx 2$, the peaks are separated by more than the width of the Lorentzian profile, and $R_\varnothing$ will be double peaked, as illustrated by the solid line in Fig. 7(b). It follows from Eq. (21) that the cooling rate in the limit of small $\delta_{\varnothing M}^0$ is proportional to the slope of $R_\varnothing$ at $\delta$, and that the rate of change of $\varnothing_c$ is positive if the slope is negative. If $-|\delta_{\varnothing M}^0 - \delta_{\varnothing i}^0| < \delta < 0$, this will result in heating of the $\varnothing$ mode, at least as long as $\delta_{\varnothing i}^0 + \delta_{\varnothing j}^0$ are both inside the peaks of $R_\varnothing$, that is, when $\delta_{\varnothing i}^0 < |\delta_{\varnothing M}^0 - \delta_{\varnothing i}^0| + \delta$, as illustrated by the solid curve in Fig. 7(c) for $(\delta_{\varnothing i}^0, \delta_{\varnothing j}^0) = (0, 4)$ and $(1, 3)$ [13,15]. The figure also shows that this thermalization or energy equilibration effect is not present if $|\delta_{\varnothing M}^0 - \delta_{\varnothing i}^0| \approx 1$, as $R_\varnothing$ is not double peaked in this case. Mathematically, $d\varnothing_c/dt, R_\varnothing$, and $P_D(\varnothing; \delta_{\varnothing i}^0)$ are all conveniently expressed as convolution integrals of functions with known Fourier transforms.

The dashed lines in Fig. 8 show the cooling rates predicted by Eq. (21) for the case of the $x$ spectator mode excited, for various energies of the two modes. When $\delta_{\varnothing M}^0$ is large compared to $\delta_{\varnothing i}^0$, we see from the right-hand side of Fig. 8(b) that the $\varnothing$ cooling rate is almost unaffected by the $x$ spectator mode. This can be understood by noting that, in the
limit where \( P_{\beta} (\delta_{M}^{(1)}; \delta_{0}^{(1)}) \) is uniform over the values of \( \delta_{0}^{(1)} \) where \( R_{\beta}(\delta_{0}^{(1)}) \) is nonzero, the symmetry of \( R_{\beta} \) implies that the average energy change per scattering event is \( \bar{\delta} \) as also discussed in Sec. III. Since \( R_{\beta}(\bar{\delta}_{\text{eff}}) = 0 \) for \( \bar{\delta}_{\text{eff}} > 1 + \delta_{M}^{(1)} + \delta_{M}^{(2)} \), this implies that the temperature of the spectator modes will not affect the cooling rate in this limit. At lower values of \( \delta_{M}^{(2)} \), we generally see a decrease in the cooling rate in a gradual approach to the thermalization regime discussed above.

The consequences of thermalization and equilibration process are complex, when considering the full 3D cooling problem. Consider for instance the case where only one mode is initially hot. According to the discussion above, this will result in heating of the two remaining modes, until the fastest-heating mode has reached a value of \( \delta_{M} \) similar to that of the initially hot mode. After this thermalization, the modes will be cooled simultaneously at a cooling rate significantly lower than the cooling rate for a single hot mode.

At this point, it is worth reconsidering the validity of our omission of recoil heating. The recoil heating rate as given by Eq. (5) is seen to have a maximum value of \( 4r/3 \) at the resonant scattering rate. It is clear from Fig. 8 that for typical values of \( r \) on the order of \( 10^{-3} \), recoil is insignificant at high energies.

B. 3D cooling of ions including the effects of micromotion

For an ion in a linear Paul trap, if we take the \( z \) direction to be the axis, confinement in the transverse \( x \) and \( y \) directions is provided by the ponderomotive potential of a rf quadrupole field. The full 3D cooling problem including micromotion on the transverse modes is very complex even in

\[
R_{\beta}(\delta_{\text{eff}}; \beta) = \sum_{\text{ even } n} J_{n}(\beta) \frac{1}{1 + (\delta_{\text{eff}} - n\tilde{\Omega})^{2}}. \tag{23}
\]

where \( \delta_{\text{eff}} = \bar{\delta} + \sum_{j} \delta_{0}^{(j)} \) is the effective detuning, \( \beta = |\mathbf{a} \cdot \mathbf{k}| \) is the micromotion modulation index, \( \tilde{\Omega} = \hbar \Omega / E_{0} \) is the scaled rf frequency, and \( J_{n} \) is the \( n \)th Bessel function.
In contrast to the situation in Ref. [13], we are considering a case where \( \beta \) changes during the secular motion. Since \( \beta \) and \( \delta_{\mathrm{ext}} \) depend on \( \vec{x} \) and \( \vec{x} \), respectively, we parametrize the secular motion by the instantaneous phases \( \phi_i \), where \( \vec{x}_i = x_i^{(0)} \cos[\phi_i(t)] \), and where \( \vec{x}_i^{(0)} \) is slowly varying and \( \phi_i = \omega_i t \). Choosing the \( x \) and \( y \) axes so that the rf field is proportional to \((\vec{x} - \vec{y}) \cos(\Omega t)\), we find that \( \delta_{\Omega}^{(0)} = \delta_{\Omega}^{(z)} \sin(\phi_i) \). In the limit where the transverse confinement is modified only weakly by static potentials [22,23], so that \( \omega_i = \omega_z \), we find in the pseudopotential approximation that \( \beta = \sqrt{2} \delta_{M}^{(z)} \cos(\phi_i) - \delta_{M}^{(z)} \cos(\phi_i)/\Omega \), which we note to be independent of the secular frequencies. In this case we have

\[
\frac{dE_i}{d\tau} = - \int \delta R_{\mu}(\delta_{\mathrm{ext}}(\phi),\beta(\phi_i, \phi_j)) \frac{d\phi_i}{\Omega}, \tag{24}
\]

where the integral is over \( [0,2\pi) \) in all dimensions. Note that, since the modulation index depends only on the transverse components of the motion, the effect of excited transverse modes on the cooling of the \( z \) mode can still be described in terms of an effective line profile similar to \( R_i \) in Eq. (21).

For the cooling of the transverse modes, the effects of micromotion on the cooling rates are pronounced, as illustrated by the solid curves in Fig. 8(a). A very clear qualitative difference from the cooling rate in the micromotion-free case is that, at sufficiently high rf frequencies (\( \Omega_{\text{eff}} > 4.4 \) for \( \delta = -1 \)), stable points for the transverse mode energies develop even when the remaining modes are cold. This effect has been discussed in Ref. [15] and is attributed to the heating peak of the Doppler distribution becoming resonant with a micromotion sideband, as described by Eq. (23). This might be related to the bistable behavior reported in some single-ion experiments [21,27,28]. The stability breaks down when thermalization is taken into consideration. Consider, for instance, the stable point indicated in Fig. 8(a) for \( \delta_{M}^{(z)} = 7.6 \). Here, it is clear from the figure that, when the \( z \) mode has heated to \( \delta_{M}^{(z)} > 2 \), cooling of the \( x \) mode will commence.

When \( \delta_{M}/\Omega \leq \sqrt{2} \) for the transverse modes, we find that only the \( J_0 \) term of Eq. (23) contributes significantly, and the argument of Sec. VI A that the cooling rate for the \( z \) mode is not affected by excited transverse modes when \( \delta_{M}^{(z)} > 1 + \delta_{M}^{(z)} + \delta_{M}^{(z)} \) also applies here, as illustrated by Fig. 8(b).

Finally, another effect with respect to micromotion is that the presence of uncontrolled stray fields can result in the ion experiencing micromotion even at the ion equilibrium position. At temperatures where \( \delta_{M} < \Omega \), the first-order effect according to Eq. (23) of this will be a reduction of the central spectral component by a factor of \( J_0(\beta) \); see, for example, Ref. [29]. We note that this effect can be compensated by using an effective saturation parameter based on the steady state fluorescence observed in the trap.

It is clear from the results in this section that we cannot ignore the transverse modes if their associated maximal Doppler shifts are comparable to that of the \( z \) mode. If, however, we assume the transverse modes are cold enough to avoid the heating effects described in Figs. 7 and 8, we have seen
modes. This would, however, reduce the efficiency of cooling of the transverse modes [11].

C. Departures from the weak-binding, low-saturation limit

In most experimental situations, we will not strictly satisfy the requirements of low saturation or weak binding. In particular, for the trap referenced in Fig. 4 the secular frequencies of the transverse modes are approximately equal to half the 41.4 MHz linewidth of the Doppler cooling transition, making the weak-binding assumption only approximate. Also, the illustrated data were obtained at a saturation parameter of 0.9, outside the validity region of the line-profile model that accounts for rf micromotion (24). To validate our claim that the fluorescence signal predicted by the 1D model is a good approximation if the heating rate is assumed to be a strongly decreasing function of $\omega_0$, we performed a numerical Monte Carlo simulation of the fluorescence, based on integrating the optical Bloch equations through a large number of cooling trajectories. For each trajectory, we propagate the density matrix $\rho$ through a large number of cooling trajectories. For each trajectory, we propagate the density matrix $\rho$ through the trap potential at $2t\omega_0$, which leads to amplification of the signal change due to a given amount of additional heating, as illustrated in Fig. 10.

Alternatively, parametric amplification [30–32] could be employed after the heating cycle to modify the thermal distribution. Parametric amplification can be implemented by modulating the $z$ trap potential at $2\omega_0$, and leads to amplification of one quadrature of the motion while damping the other quadrature. For a low value of $\overline{E}$, parametric amplification would increase the fraction of experiments in which the atom is in the slow-cooling regime at the beginning of the cooling process, thus increasing the signal for a given heating period, as illustrated in Fig. 11.

The performance of either of the modified protocols outlined in this section will be determined by the uncertainty of the implementation parameters.

VII. MODIFIED EXPERIMENTAL PROTOCOLS

We consider two modifications to the experimental protocol to reduce the total measurement time. Both are motivated by the fact that the size of the signal from a given amount of heating increases with increased initial energy.

One approach would be to coherently add a known amount of energy to the $z$ mode at the start of the heating period. If the added energy is enough to bring the atom into the slow-cooling regime, this will increase the signal change due to a given amount of additional heating, as illustrated in Fig. 10.

In conclusion, we have shown that the motional energy of a trapped atom or ion can be estimated from the temporal changes in fluorescence observed when Doppler cooling is applied. Specifically, the initial energy can be estimated by fitting Eq. (17), where the mean initial motional energy is the only free variable, to the observed fluorescence. Our analysis assumes the oscillation frequency of the atoms is much smaller than the linewidth of the optical transition used for different assumptions for the frequency dependence of the heating to the data set presented in Fig. 4. We find that, if we assume the transverse modes are not heated, we obtain a temperature estimate of 3.9 K, in agreement with the result of fitting the 1D model to the data, as illustrated by Fig. 4. The $z$ mode temperature of 3.7 K estimated from the $\omega^{-1.4}$ model found in Ref. [20] is close to, and slightly smaller than, this value, and agrees with the temperature estimate of $3.4\pm0.3$ K based on extrapolating heating rates measured with the Raman sideband technique for the same trap configuration. This particular form of the frequency dependence of the heating rate was observed for the same trap when the Raman sideband technique [20] was used, and similar frequency dependencies have been observed in other geometries [16,18]. If we instead assume an $\omega^{-1}$ dependence of the heating, the results change only slightly.

Our main conclusions from the simulation results are that the primary effect of the presence of weakly heated spectator modes will be to slow down cooling due to thermalization. If $\omega^{-1.4}$ heating of the transverse modes is assumed, the 1D model will somewhat overestimate the motional temperature of the axial mode.
Doppler cooling and the motional energy at the start of the cooling is thermal.

Compared to Raman sideband transition methods for heating rate measurements, this method is simpler to implement experimentally but requires longer measurement duration for traps with low heating rates. On the other hand, for high heating rates, where sideband cooling is inefficient, this may be the method of choice. We have shown that, in the typical situation, where the time for heating dominates, the total measurement time decreases with decreasing laser intensity, decreasing laser detuning, and increased heating period duration. We have compared the trade-off between these parameters [Fig. 5 and Eq. (20)]. Finally, we indicate (Sec. VII) that the total measurement time can be reduced by adding additional energy to more quickly bring the ion into the low-fluorescence regime.

By comparison with various models of three-dimensional Doppler cooling, we have established that under typical experimental conditions the effects of the high-frequency modes are small, and that they will lead to temperature estimates that are somewhat higher than the actual temperature of the low-frequency mode.

ACKNOWLEDGMENTS

This work was supported by NIST and the Disruptive Technology Office (DTO) under Contract No. 712868. J.H.W. acknowledges support from The Danish Research Agency. R.J.E. acknowledges support from the National Research Council. S.S. acknowledges support from the Carlsberg Foundation. J.P.H. acknowledges support from the Lindemann Foundation. We thank J. J. Bollinger and C. Ospelkaus for comments on the manuscript.

APPENDIX A: INTEGRALS

The integrals appearing in Eqs. (9), (11), and (21) are all convolution integrals of elements with analytical Fourier transforms and can thus be easily evaluated in Fourier space. For the 1D integrals, the inverse Fourier transform can also be performed analytically. Here we present a more direct approach to evaluating the 1D integrals.

For $a, b \in \mathbb{R}$ we define $Z(a, b)$ as

$$Z = \int_0^{2\pi} \frac{d\phi}{\sin(\phi) - z} = -\frac{1}{z} \sqrt{\frac{z^2}{z^2 - 1}},$$

(A1)

where $z=(a+i)/b$. Noting that

$$\frac{1}{x-z} = b \left[ \frac{1}{1 + (a-bx)^2} \right] \left[ (bx-a) + i \right],$$

(A2)

we find that, according to Eq. (A1),

$$\int_0^{2\pi} \frac{b \sin(\phi)}{1 + [a-b \sin(\phi)]^2} \frac{d\phi}{2\pi} = \frac{1}{b} [\text{Re}(Z) + a \text{Im}(Z)],$$

and

$$\int_0^{2\pi} \frac{1}{1 + [a - b \sin(\phi)]^2} \frac{d\phi}{2\pi} = \frac{1}{b} \text{Im}(Z).$$

Taking the branch cut discontinuity for $\sqrt{\cdot}$ to be along the negative real axis, we have for $b > 0$ that $\sqrt{(-iz)^2} = -iz$, so that

$$Z(a, b) = \frac{ib}{\sqrt{b^2 - (a+i)^2}}, \quad b > 0.$$  

(A3)

APPENDIX B: NUMERICAL CALCULATION OF THE AVERAGED SCATTERING RATE

In this section we present an efficient numerical method for evaluating the averaged scattering rate given by Eq. (17).
Introducing $e_n = \Xi(e_0, n\Delta\tau)$, for $n = 0, 1, \ldots$, we note that
$\Xi(e_n, n\Delta\tau) = e_{m+n}$. The values of $e_n$ are the energies along a
single cooling trajectory. If the scattering rate can be considered
constant on time scales of $\Delta\tau$, we find that the thermally averaged scattering rate, as given
by Eq. (17), averaged over the same intervals can be approximated by

$$
\frac{dN}{d\tau}[\Xi(e_0, \tau)] = R_n, \quad \tau \in [(n - 1)\Delta\tau, n\Delta\tau], \quad (B1)
$$

\[ \begin{align*}
\bar{R}_n &= \sum_{n=0}^{\infty} R_{n+m} \int_{e_{m+1}}^{e_m} P_0(e')de'. \\
&= 2e^{-e_{m+1}} \sinh \left( \frac{\Delta e}{2} \right).
\end{align*} \]

Since the values of the $R_n$ are independent of $P_0$, $\bar{R}_n$ is easily
calculated for different $P_0$ by list convolution.

For the Maxwell-Boltzmann distribution, a numerically stable form of the weight factors appearing in (B2) is

$$
\int_{e_{m+1}}^{e_m} e^{-e'/\beta}de' = 2e^{-e_{m+1}} \sinh \left( \frac{\Delta e}{2} \beta \right).
$$
Simplified motional heating rate measurements of trapped ions

National Institute of Standards and Technology, Boulder, Colorado 80305, USA

Received 10 July 2007; published 19 September 2007

We have measured motional heating rates of trapped atomic ions, a factor that can influence multi-ion quantum logic gate fidelities. Two simplified techniques were developed for this purpose: one relies on Raman sideband detection implemented with a single laser source, while the second is even simpler and is based on time-resolved fluorescence detection during Doppler recooling. We applied these methods to determine heating rates in a microfabricated surface-electrode trap made of gold on fused quartz, which traps ions 40 μm above its surface. Heating rates obtained from the two techniques were found to be in reasonable agreement. In addition, the trap gives rise to a heating rate of 300±30 s⁻¹ for a motional frequency of 5.25 MHz, substantially below the trend observed in other traps.

DOI: 10.1103/PhysRevA.76.033411

PACS number(s): 32.80.Lg, 03.67.Lx, 32.80.Pj, 42.50.Vk

I. INTRODUCTION

Control of the quantum states of trapped ions has progressed significantly over the past few years. Many of the necessary requirements for quantum-information processing have been demonstrated in separate experiments, such as high-fidelity state preparation, readout, single- and two-qubit gates, and long-lived single-qubit coherence (see, e.g., Ref. [1]). One of the limitations thus far in scaling to larger numbers of ions has been the lack of a suitable trap architecture. A critical benchmark for trap design is the heating rate of an ion’s motional degrees of freedom due to electric field noise from the trap electrodes. As current quantum gates rely on the coupled motion of two or more ions, noise in the motion can degrade gate fidelities [2]. To facilitate the determination of these heating rates, we have developed two measurement methods that have reduced hardware complexity compared to that of more traditional methods [3–16]. Here, we report details of these two methods.

The ion trap used for this study is a monolithic design made of gold on fused quartz, where all trap electrodes reside in a single plane [17,18]. This “surface-electrode” geometry has the potential to greatly simplify the trap fabrication process and electrode wiring topology, thereby enabling the creation of large multiplexed trap arrays. Heating rates were previously measured in a nearly identical trap by recooling measurements. The axial trap frequency was found to be in reasonable agreement between the values obtained with the two techniques. In addition, heating rates were measured at several trap frequencies and the electric field noise was found to have approximately 1/2 character in this particular trap. Finally, although the trap discussed here was fabricated by the same process as that of Ref. [18], the heating rates are found to be somewhat lower, possibly due to cleaner electrode surfaces [20].

II. DOPPLER RECOOLING

The Doppler recooling method is based on the observation that the near-resonance fluorescence rate from an ion is influenced by its motional temperature due to the Doppler effect. By monitoring the fluorescence as a function of time during Doppler cooling of an initially hot ion, one can determine the initial temperature of the ion averaged over many experimental runs. In the experiments discussed here, an ion is first cooled close to the Doppler limit. Then it is allowed to heat up for a variable amount of time (the delay time) by turning off the Doppler cooling laser beam. The laser is subsequently turned back on and the fluorescence is monitored as a function of time until the ion’s fluorescence rate reaches its steady-state value. By fitting a theoretical model [19] to the data, the ion’s temperature at the end of the delay time can be extracted. The model is a one-dimensional semiclassical description of Doppler cooling in the “weak-binding” limit, where the ion’s motional frequency is much smaller than the linewidth of the Doppler cooling transition (see Ref. [19] for details). An attractive feature of this technique is its relative simplicity. It requires only one low-power red-detuned laser beam and no magnetic fields [18], in contrast to the Raman sideband technique discussed in Sec. III.

Figure 1 displays the average number of axial vibrational quanta, ⟨n⟩, for various delay times obtained from Doppler recooling measurements. The axial trap frequency was ω/2π=4.02 MHz. Here the axial direction is the direction of weakest binding in the trap and is controlled primarily by static potentials [18]. Each value of ⟨n⟩ was obtained by
fitting the model of Ref. [19] to a Doppler recooling trace, as exemplified by the inset data. The model yields the thermal recooling time for a delay time of 25 s with fit (solid line) to the model of Ref. [19].

To obtain these data, magnesium ions were created in the trap through a two-photon photoionization process using 1–10 mW of 285 nm laser excitation [22]. This ionization method was found to significantly reduce the required Mg oven temperature (and the concomitant pressure rise) compared to electron impact ionization. All experiments were carried out with $^{25}\text{Mg}^+$ ions in a magnetic field of $B=10$ G for consistency with the Raman measurements discussed below. In addition, the Doppler beam (“blue Doppler”) saturation parameter was 0.9 throughout. Due to the hyperfine structure of $^{25}\text{Mg}^+$, a second laser beam (“red Doppler”) was used to repump out of the $F=2$ ground-state manifold. This additional beam was not necessary in the measurements of Ref. [18] performed at $B=0$ with $^{24}\text{Mg}^+$, an isotope without hyperfine structure. These beams have the same intensities, polarizations, and detunings as in the Raman experiments described below (see Figs. 2 and 3).

III. SIMPLIFIED RAMAN SIDEBAND DETECTION

Our Raman sideband detection apparatus has been simplified compared to more commonly used schemes [3–16]. Instead of relying on three or more lasers, Fig. 2 depicts how the two Raman beams and two Doppler cooling beams were derived from a single 280-nm source: a frequency-quadrupled fiber laser. The frequency-doubled output of the laser at 560 nm was frequency-locked to an iodine vapor absorption line. The 560 nm light was then doubled to produce 280 nm light. Multiple acousto-optic modulators (AOMs) were used as frequency shifters and on-off switches, controlling the 280-nm beams that we call red Raman (RR), blue Raman (BR), red Doppler (RD), and blue Doppler (BD). The AOMs (boxes) are labeled by the frequency shift they impart to the deflected beams; 2$X$ indicates double-pass configuration. The frequency shift is noted for each beam relative to the $^2\Sigma_{1/2}^+(3, 3\rightarrow 3, 4)$+$^2P_{3/2}^-, (4, 4)$ cycling transition of $^{25}\text{Mg}^+$. Inset: geometry of beams, trap axis, and magnetic field $B$, including the photoionization (PI) beam.

FIG. 1. Average number of axial vibrational quanta, $\langle n \rangle$, as a function of delay time obtained by the Doppler recooling method with an axial trap frequency of 4.02 MHz. The fit (solid line) gives a heating rate of $d\langle n \rangle/dt=620\pm50$ s$^{-1}$. Inset: ion fluorescence versus recooling time for a delay time of 25 s with fit (solid line) to the model of Ref. [19].

FIG. 2. Schematic of the laser beams. All beams are derived from one 280-nm source by use of acousto-optic modulators (AOMs) as frequency shifters and switches; they are labeled red Raman (RR), blue Raman (BR), red Doppler (RD), and blue Doppler (BD). The AOMs (boxes) are labeled by the frequency shift they impart to the deflected beams; 2$X$ indicates double-pass configuration. The frequency shift is noted for each beam relative to the $^2\Sigma_{1/2}^+(3, 3\rightarrow 3)+^2P_{3/2}^-, (4, 4)$ cycling transition of $^{25}\text{Mg}^+$. Inset: geometry of beams, trap axis, and magnetic field $B$, including the photoionization (PI) beam.

FIG. 3. Ion fluorescence as a function of Raman frequency detuning relative to the Raman carrier ($\Delta n=0$) transition (not shown). Fits (solid lines) to the Raman sideband amplitudes yield $\langle n \rangle=0.34\pm0.08$, the average number of axial vibrational quanta after Raman cooling with an axial trap frequency of 5.25 MHz. Inset: partial level diagram of $^{25}\text{Mg}^+$ showing the laser beams and their polarizations relative to the quantizing magnetic field. Relevant $|F, m_F\rangle$ levels are indicated. For $^{25}\text{Mg}^+$ in a magnetic field of 10 G, the $^2\Sigma_{3/2}^+(3, 3\rightarrow 3)$ and $^2\Sigma_{1/2}^+(2, 2\rightarrow 2)$ levels are split by 1.81 GHz. The $^2\Sigma_{1/2}^-\leftrightarrow ^2P_{3/2}^-$ transition wavelength is 280 nm.
(BD). Referring to the level diagram of $^{25}$Mg$^+$ in the inset of Fig. 3, a scheme relying on multiple lasers would typically employ two $\sigma^-$-polarized laser beams (derived from the same laser) to drive the $^2S_{1/2} \leftrightarrow ^2P_{3/2}$ transition (not shown) for optical pumping of the $F=2$ and $F=3$ ground states to the $|F=3,m_F=\pm 3\rangle$ state. A second laser would be needed to drive the $^2S_{1/2}, |3, -3\rangle \leftrightarrow ^2P_{3/2}, |4, -4\rangle$ cycling transition for state detection. In addition, a third laser would be used to generate two beams for far-off-resonant Raman transitions [4]. By contrast, we only required a single laser by using the $^2S_{1/2} \leftrightarrow ^2P_{3/2}$ transition for Doppler cooling and state preparation and detection (with possibly reduced state preparation fidelity) and by accepting relatively low Raman detunings.

The double-passed AOMs (Fig. 2) generated two beams for Raman transitions with an adjustable frequency difference near 1810 MHz and detunings from the $^2P_{3/2}, |4, -2\rangle$ state of approximately 900 MHz, as shown in the inset of Fig. 3. These relatively small Raman detunings, given the optical transition linewidth of 41.4 MHz, led to significantly reduced coherence of Raman transitions through incoherent photon scattering [12]. In particular, the Rabi flopping decay time for the red sideband $|3, -3\rangle \rightarrow |2, -2\rangle |n=1\rangle$ Raman transition [5] was approximately one Rabi oscillation period after sideband cooling to $\langle n \rangle \approx 1$. Despite these compromises, we were able to achieve reasonable sideband detection contrast [23] and cool the axial mode to $\langle n \rangle \approx 0.34 \pm 0.08$ quanta for an axial trap frequency of 5.25 MHz (Fig. 3), which was sufficient for heating measurements.

In a typical experimental sequence, we first performed Doppler cooling and $|3, -3\rangle$ state preparation with BD and RD for 300 $\mu$s followed by BD for 20 $\mu$s and then RD for 20 $\mu$s. Next we applied 25–30 cycles of resolved sideband cooling [4] on the ion’s axial mode of motion. One cycle of sideband cooling consisted of driving the red Raman sideband transition $|3, -3\rangle |n\rangle \rightarrow |2, -2\rangle |n-1\rangle$ using RR and BR, where the pulse length was adjusted to be an approximate $\pi$ pulse for $n=1$. The cycle was completed by repumping to the $|3, -3\rangle$ state with RD for 8 $\mu$s and then BD for 0.3 $\mu$s. These pulses typically enabled cooling of the ion to $\langle n \rangle \approx 1$ axial vibrational quanta.

In order to measure heating rates, all beams were turned off for a specified delay period (usually 0–5 ms) to let the ion heat up after sideband cooling. Then a Raman analysis pulse was applied with variable BR frequency detuning; the pulse time was chosen such that an approximate $\pi$ pulse was effective when resonant with the blue sideband transition $|3, -3\rangle |n=0\rangle \rightarrow |2, -2\rangle |n=1\rangle$. Finally, ion fluorescence was detected with BD applied for 50 $\mu$s. The sequence was repeated for different BR detunings to sweep out the sidebands, as in Fig. 3, where each data point was typically an average of several hundred experiments (1400 for the data in Fig. 3). The Rabi flopping $\pi$ time on the red sideband was 1–5 $\mu$s, depending on the axial trap frequency and laser beam intensities. Typical laser powers were 1 mW (BD), 10 $\mu$W (RD), 90 $\mu$W (BR), and 40 $\mu$W (RR); beam waists were estimated to be 15–30 $\mu$m.

Ideally, if $\langle n \rangle=0$, the analysis pulse leaves the ion in the fluorescing $|3, -3\rangle$ state when resonant with the red sideband transition, which is then forbidden. However, when resonant with the blue sideband transition, which is allowed, the analysis pulse promotes the ion to the nonfluorescing $|2, -2\rangle$ state. When $\langle n \rangle \neq 0$, the red sideband is not forbidden, but the ratio of the red and blue sidebands can be used to extract $\langle n \rangle$ (see below and [3,4]).

**IV. TECHNIQUE COMPARISON AND SURFACE TRAP RESULTS**

Figure 4 shows values of $\langle n \rangle$ extracted from Raman sideband measurements at multiple delay times for the same 4.02-MHz axial trap frequency as Fig. 1. Each value of $\langle n \rangle$ is obtained from Gaussian fits to the sidebands as in Fig. 3. Assuming $n$ has a thermal distribution, then $\langle n \rangle = R \left\langle n^2 \right\rangle^2 - \langle n \rangle$, where $R$ is the ratio of the red and blue sideband amplitudes [4,8]. A weighted linear fit to $\langle n \rangle$ versus delay time gives a heating rate; the average value obtained from this and other similar data runs is $d\langle n \rangle/dt = 690 \pm 60$ s$^{-1}$. This compares reasonably well with the value obtained from Doppler recolcing (see Sec. II). Likewise, at a different trap frequency of 2.86 MHz, we find heating rates of $1470 \pm 150$ s$^{-1}$ and $1260 \pm 130$ s$^{-1}$ for Raman and Doppler recolcing techniques, respectively.

Figure 5 displays heating rates, for a range of axial trap frequencies, measured with the Raman sideband technique. A weighted power-law fit yields $d\langle n \rangle/dt \propto \omega^{-2.4 \pm 0.4}$. From the heating rates we can calculate the electric field noise spectral density $S_E(\omega) = d\langle n \rangle/dt \propto \omega^{-2.4 \pm 0.4}$, where $m$ is the ion mass and $e$ is the electron charge, respectively. The explicit factor of $\omega$ in this equation and the measured frequency dependence of $d\langle n \rangle/dt$, we find $S_E \propto \omega^{-1.4 \pm 0.4}$ for our surface-electrode trap. A similar frequency dependence has been observed in ion traps of different geometries [8,11,14] and may give some insight into the heating mechanism.

In Fig. 6, we put these heating results in perspective by plotting values of $S_E(\omega)$ and $a_0 S_E(\omega)$ versus $d$, the distance between the ion and the nearest electrode, for several different ion traps reported in the literature (as similarly done in Ref. [11]). The surface trap studied here has $d=40$ $\mu$m. For
comparison, all the traps plotted have approximately room-temperature electrodes. It has recently been found, however, that cooling the electrodes can significantly reduce the heating rates [14,16]. While the fundamental heating mechanism is not understood, the predominant explanation is that the electrodes are covered in patches of varying potential that fluctuate with an unknown frequency dependence. If we assume that these fluctuating patch potentials have a size $\ll d$, then $S_E$ should scale as $d^{-4}$ (indicated by the gray shaded bands) [8]. A similar dependence on $d$ was observed in Ref. [14], where $d$ could be varied in the same trap (Fig. 6, open circles).

Concerning the frequency spectrum of the noise, Fig. 6 shows several cases where the values of $\omega S_E$ for a given trap are bunched together, indicating that $S_E \propto 1/\omega$ is a better assumption than $S_E$ being independent of $\omega$. In most traps shown, however, $S_E$ actually depends more strongly on $\omega$. It is unclear whether this is intrinsic to the traps or due to external noise sources.

As can be seen, the values of $S_E$ and $\omega S_E$ for the NIST surface trap are over an order of magnitude lower than what might be expected from the trend. The significant scatter in the data highlights the importance of other parameters that have yet to be fully quantified, such as the microscopic properties of the electrodes (purity, roughness, crystallinity, etc.). For example, there is some evidence that electrode contamination (due to the ion loading process) influences the electric field noise [8]. In our apparatus, the loading geometry is such that the electrode surfaces become coated with a small amount of Mg during each loading attempt. While we have not measured a systematic change in the heating rate due to loading in this trap, we cannot rule out the influence of surface contamination.

V. CONCLUSION

According to the results presented here, the simple Doppler recooling technique is a reasonably accurate tool for trap characterization. It has several advantages, including simplicity and relatively small resource requirements (a single low-power laser, no magnetic fields, etc.), which results in lower cost and setup time. The primary disadvantage is that delay times can be inconveniently long for low heating rates. For example, delay times of approximately 1 min and averaging durations of several hours were needed for a heating rate of 300 s$^{-1}$ with our experimental parameters. This would be particularly troublesome if the uncooled ion lifetime in the trap (set by background gas collisions) were comparable to the delay time. However, weighing these factors, the Doppler recooling technique may still prove useful for rapid characterization of new ion traps. Means to potentially reduce the measurement time for this method are outlined in Ref. [19].

Accepting somewhat increased complexity, a simplified Raman sideband detection apparatus is shown to be suitable for heating rate measurements in the single-quantum regime. Using relatively small Raman beam detunings (900 MHz) and the same $2S_1/2 \leftrightarrow 2P_{3/2}$ transition for Doppler cooling, state preparation, and detection enables a single laser to supply all necessary beams. Despite reduced sideband cooling efficiency and Rabi coherence, cooling of a single mode to $n \approx 1$ is achieved with significantly fewer resources than more common Raman sideband detection experiments [3–16].

Finally, heating rates for the new surface-electrode trap geometry [18] appear to be manageable for large-scale
quantum-information processing. Compared to other ion traps (Fig. 6), the rates measured here are significantly lower than one might expect for the electrode-ion separation in this trap. A future surface-electrode trap design—having multiple zones with differing values of $d$ and a loading scheme that does not contaminate the electrodes—would be useful to further characterize the scaling of the electric field noise.

ACKNOWLEDGMENTS

This work was supported by the Disruptive Technology Office (DTO) under Contract No. 712868, by the DOD Multidisciplinary University Research Initiative (MURI) program administered by the Office of Naval Research, and by NIST. S.S., J.H.W., and J.P.H. acknowledge financial support from the Carlsberg Foundation, the Danish Research Agency, and the Lindemann Trust, respectively.

[21] The measured heating rates varied from run to run by somewhere more than expected from the estimated uncertainties of a given run for unknown reasons. Quoted heating rates are averages of data taken on different runs.
[23] The optical windows were found to have spatially inhomogeneous birefringence leading to reductions in both optical polarization purity and atomic state preparation fidelity.
Errors in trapped-ion quantum gates due to spontaneous photon scattering

NIST Boulder, Time and Frequency Division, Boulder, Colorado 80305, USA.
(Received 6 November 2006; published 27 April 2007)

We analyze the error in trapped-ion, hyperfine qubit, quantum gates due to spontaneous scattering of photons from the gate laser beams. We investigate single-qubit rotations that are based on stimulated Raman transitions and two-qubit entangling phase gates that are based on spin-dependent optical dipole forces. This error is compared between different ion species currently being investigated as possible quantum-information carriers. For both gate types we show that with attainable laser powers the scattering error can be reduced to below current estimates of the fault-tolerance error threshold.

I. INTRODUCTION

Quantum bits, or qubits, that are encoded into internal states of trapped ions are an interesting system for quantum-information processing (QIP) studies [1,2]. Internal states of trapped ions can be well isolated from the environment, and very long coherence times are possible [2–5]. The internal states of several ion qubits can be deterministically entangled, and quantum gates can be carried out between two ion qubits [6–12].

Among different choices of internal states, qubits that are encoded into pairs of ground-state hyperfine or Zeeman states benefit from negligible spontaneous decay rates [2]. The small energy separation between the two states of such qubits (typically in the radio-frequency or microwave domain) allows for phase coherence between a local oscillator and a qubit superposition state over relatively long times [2,4,13,14].

The quantum gates that are performed on hyperfine ion qubits typically use laser beams. Since light couples only very weakly to the electron spin, spin manipulations rely instead on the spin-orbit coupling of levels that are typically excited nonresonantly through allowed electric-dipole transitions. Spin manipulations therefore require a finite amplitude in the excited electronic state, and spontaneous scattering of photons from the laser beams during the gate is inevitable.

Fault-tolerant quantum computation demands that the error in a single gate be below a certain threshold. Current estimates of the fault-tolerance error threshold range between $10^{-2}$ and $10^{-4}$ [15–17]. These estimates rely on specific noise models and error-correction protocols and should be considered as guidelines only. However, the general view is that for fault tolerance to be practical, the error probability in quantum gates should be at least as small as $10^{-4}$. It is, therefore, worth exploring the limitations to the fidelity of quantum gates performed on trapped ions with laser light using this level of error as a guideline [18].

In his 1975 paper [19], Mollow showed that the effect of a quantum coherent field on an atom is equivalent to that of a classical field plus a quantum vacuum field. The error due to the interaction with light can be categorized into two parts. The first is the error due to noise in classical laser parameters, such as intensity or phase [2,20]. The second part originates from the quantum nature of the electromagnetic field and is due to vacuum fluctuations—i.e., the spontaneous scattering of photons [21,22].

Ion-qubit levels and transitions

Most ion species considered for QIP studies have a single valence electron, with a $^5S_{1/2}$ electronic ground state and $^2P_{1/2}$ and $^2P_{3/2}$ electronic excited states. Some of the ions also have $D$ levels with lower energy than those of the excited state $P$ levels. Ions with a nonzero nuclear spin also have hyperfine structure in all of these levels. A small magnetic field is typically applied to remove the degeneracy between different Zeeman levels. Here we consider qubits that are encoded into a pair of hyperfine levels of the $^5S_{1/2}$ manifold. Figure 1 illustrates a typical energy level structure.

To allow for a straightforward comparison between different ion species, we investigate qubits that are based on clock transitions—i.e., a transition between the $|F=I−1/2,m_F=0⟩=|↑⟩$ and $|F=I+1/2,m_F=0⟩=|↓⟩$ hyperfine levels in the $S_{1/2}$ manifold of ions with a half odd-integer nuclear spin $I$, at a small magnetic field. For this transition, the total Raman and Rayleigh photon scattering rates, as well as the Rabi frequency, are independent of $I$, and the comparison between different ion species depends only on other atomic constants. Superpositions encoded into these states are also more resilient against magnetic field noise [3,5,11]. Even though our quantitative results apply only to this configuration, for other choices of hyperfine or Zeeman qubit states (including those with $I=0$) the results will not change significantly.

Gates are assumed to be driven by pairs of Raman beams detuned by $Δ$ from the transition between the $S_{1/2}$ and the $P_{1/2}$ levels (see Fig. 1). We further assume that the Raman beams are linearly polarized and Raman transitions are
driven by both $\sigma_+$ photon pairs and $\sigma_-$ photon pairs. The two beams in a Raman pair are designated as red Raman and blue Raman beams that are detuned by $\Delta$ from the transition to the $P_{1/2}$ level, represented by the two straight arrows. The angular frequency difference between the two beams equals the angular frequency separation between the qubit levels $\omega_0 - \omega_f = \omega_0$. Some ion species have $D$ levels with energies below the $P$ manifold. Wavy arrows illustrate examples of Raman scattering events.

The Rabi frequency between the two clock states is [23]

$$\Omega_R = \frac{g_{bf} \delta_f}{3} (b_r - b_o r_r) \frac{\omega_f}{\Delta} \Delta (\Delta - \omega_0),$$

(1)

where $g_{bf} = E_{bf} \langle P_{3/2}, F=I+1/2, m_f=F | \hat{d} \cdot \hat{S}_e | P_{1/2}, F=I \rangle / |E_{bf}|$ is the peak electric field of the $b$ or $r$ beam at the position of the ion, respectively, and $\hat{d} \cdot \hat{S}_e$ is the right circular component of the electric-dipole operator. The right and left circular polarization components of the $b$ and $r$ beams are $b_{\sigma_+}$ and $r_{\sigma_+}$, respectively. The $P_{1/2}$ and $P_{3/2}$ excited levels are separated by an angular frequency $\omega_f$.

The total spontaneous photon scattering rate from these beams is given by [23]

$$\Gamma_{\text{total}} = \frac{\gamma}{3} \left[ g_{b/}^2 (b_r^2 + b_o^2) + g_{r/}^2 (r_r^2 + r_o^2) \right] \left[ \frac{1}{\Delta} + \frac{2}{(\Delta - \omega_f)^2} \right].$$

(2)

Here $\gamma$ is the natural linewidth of the $P_{1/2}$ and $P_{3/2}$ levels [27]. We now assume linearly polarized Raman beams with $b_+ = b_0 = r_+ = -r_0 = 1/\sqrt{2}$ [23]. We further assume $g_b = g_r = g$. The time for a $\pi$ rotation is

$$\tau_\pi = \frac{\pi}{2 \Omega_R}. $$

(3)

Combining Eqs. (1)–(3) the probability to scatter a photon during $\tau_\pi$ is given by

$$P_{\text{total}} = \left( \frac{\pi \gamma}{\omega_f} \right) \frac{2 \Delta^2 + (\Delta - \omega_f)^2}{|\Delta (\Delta - \omega_f)|}. $$

(4)

The dashed line in Fig. 2 shows $P_{\text{total}} \text{ vs } \Delta$, where the laser detuning is expressed in units of the excited-state fine-structure splitting and the scattering probability is given in units of $\gamma/\omega_f$. The total scattering probability has a global minimum of

$$P_{\text{min}} = \frac{2 \sqrt{2} \pi \gamma}{\omega_f},$$

(5)

when the laser detuning is between the two fine-structure manifolds $[\Delta = (\sqrt{2} - 1) \omega_f]$. The asymptotic value of $P_{\text{total}}$ for large positive or negative detuning,

$$P_{\text{asym}} = \frac{3 \pi \gamma}{\omega_f},$$

(6)

is only slightly larger than the global minimum.
Previous studies estimated decoherence by assuming that any photon scattering will immediately decohere a hyperfine superposition \[23\]. Under this assumption the lowest possible gate error equals \(P_{\text{min}}\) and ions with a small \(\gamma/\omega_{\text{rf}}\) ratio benefit from a lower gate error minimum.

Two kinds of off-resonance photon scattering occur in the presence of multiple ground states: inelastic Raman scattering which transfers population between ground states and Rayleigh elastic scattering which does not change ground-state populations. Since no energy or angular momentum is exchanged between the photon and ion internal degrees of freedom, no information about the qubit state is carried away by a Rayleigh-scattered photon. Rayleigh scattering, therefore, does not necessarily lead to decoherence \[23–25\]. This was experimentally shown in \[25\] where, when of equal rate from both qubit levels, off-resonance Rayleigh scattering of photons did not affect the coherence of a hyperfine superposition. Decoherence in the presence of light was shown to be dominated by Raman scattering. The guideline used in \[23\] is therefore overly pessimistic.

In this paper we reexamine the errors due to spontaneous photon scattering on single-qubit gates (rotations of the equivalent spin-1/2 vector on the Bloch sphere) and two-qubit (entanglement) gates, where the qubits are based on ground-state hyperfine levels and manipulated with stimulated Raman transitions. We compare between different ion species and examine different Raman laser parameters. In Sec. II we analyze the contribution to the gate error due to spontaneous Raman scattering. Following a Raman scattering event the ion qubit is projected into one of its ground states and spin coherence is lost. This error was also addressed in \[26\]. In Sec. III we examine the error due to Rayleigh scattering. The Rayleigh scattering error results primarily from the photon’s recoil momentum kick. We show that both types of gate errors can be reduced to small values, while keeping the gate speed constant, with the use of higher laser intensity.

II. RAMAN SCATTERING ERROR

In a Raman photon scattering event energy and angular momentum are exchanged between the scattered photon and the ion’s internal degrees of freedom. The polarization and frequency of the scattered photon (with respect to those of the laser) become entangled with the ions’ internal state. Therefore, after tracing out the photon degrees of freedom, the ions’ spin coherence is lost. In other words, Raman scattering serves as a measurement of the ion-qubit internal state. Following a spontaneous Raman scattering event the ion qubit is projected into one of the ground states in the \(S_{1/2}\) manifold. For ions with low-lying \(D\) levels, Raman scattering events can also transfer the ion from the qubit levels into one of the \(D\) levels.

A. Single-qubit gate

For a single-qubit gate we choose to look at the fidelity of a \(\pi\) rotation around the \(x\) axis of the Bloch sphere, represented by the Pauli operator \(\hat{\sigma}_x\). This gate is assumed to be driven by a copropagating Raman beam pair, where the frequency difference between the beams equals the frequency separation between the two qubit states, \(\omega_{\text{rf}}\). It is straightforward to generalize this case to other rotations.

As a measure of the error in the rotation, we use the fidelity of the final state (characterized by density matrix \(\hat{\rho}_{\text{final}}\)) produced by the erroneous gate as defined by

\[
F = \langle \Psi | \hat{\rho}_{\text{final}} | \Psi \rangle,
\]

where \(|\Psi\rangle\) is the ideal final state. Given an initial state \(|\Psi_{\text{init}}\rangle\), \(|\Psi\rangle\) can be written as

\[
|\Psi\rangle = \hat{\sigma}_x |\Psi_{\text{init}}\rangle.
\]

In the presence of off-resonance Raman scattering, the density matrix that describes the state of the qubit after the gate has the form

\[
\hat{\rho}_{\text{final}} = (1 - P_{\text{Raman}}) |\Psi\rangle \langle \Psi | + \hat{\rho}_e.
\]

The erroneous part of the density matrix, \(\hat{\rho}_e = \sum |i\rangle \langle i|\), is composed of projectors into different levels \(|i\rangle\) of lower energy. Here \(P_{\text{Raman}} = \sum |i\rangle \langle i|\) is the probability for a spontaneous Raman scattering event to occur during the gate.

Note that some Raman scattering events keep the ion within the qubit manifold. Using Eq. (7) the contribution of \(\hat{\rho}_e\) to the fidelity is positive and not strictly zero. For simplicity, we neglect this contribution and put a lower bound on the gate fidelity:

\[
F \geq 1 - P_{\text{Raman}}.
\]

In what follows we assume this expression to be an equality. The error in the gate due to spontaneous Raman photon scattering is hence given by the Raman scattering probability

\[
e = 1 - F = P_{\text{Raman}}.
\]

We first examine the error due to Raman scattering back into the \(S_{1/2}\) manifold \(\epsilon_s\). The Raman scattering rate back into the \(S_{1/2}\) manifold is calculated to be \[25\]

\[
\Gamma_{\text{Raman}} = \frac{2 \gamma}{9} \left[ g_0^2 (b_0^2 + b_0^2) + g_0^2 (r_0^2 + r_0^2) \right] \frac{\omega_{\text{rf}}}{\Delta (\Delta - \omega_{\text{rf}})}.
\]

For the same laser parameters as above, the probability to scatter a Raman photon during the gate is

\[
P_{\text{Raman}} = \frac{2 \pi \gamma}{3} \frac{\omega_{\text{rf}}}{|\Delta (\Delta - \omega_{\text{rf}})|} = \epsilon_s.
\]

The solid line in Fig. 2 shows \(P_{\text{Raman}}\) vs \(\Delta\). The Raman scattering probability decays quadratically with \(\Delta\) for \(\Delta \gg \omega_{\text{rf}}\). Qualitatively, this is because Raman scattering involves a rotation of the electron spin. Electron spin rotations are achieved through spin-orbit coupling in the excited state. This coupling has opposite-sign contributions from the two fine-structure levels. Therefore as we detune far compared to the fine-structure splitting, those two contributions nearly cancel \[24\].

Using Eq. (1) we can write
The ratio $\gamma/g^2$ can be expressed in terms of atomic constants and the peak electric field amplitude $E$ at the position of the ion [27]:

$$\gamma/g^2 = \frac{4h\omega_{1/2}}{\pi\epsilon_0 c^2 E^2}.$$  (15)

Here $\omega_{1/2}$ is the frequency of the transition between the $S_{1/2}$ and $P_{3/2}$ levels ($D_2$ line), $\epsilon_0$ is the vacuum permittivity, and $c$ is the speed of light. Assuming Gaussian laser beams, at the center of the beam,

$$E^2 = \frac{4P}{\pi w_0^2 \epsilon_0}.$$  (16)

Here $P$ is the power in each of the Raman beams and $w_0$ is the beam waist at the position of the ion. The probability to scatter a single Raman photon can be written as

$$P_{\text{Raman}} = \frac{2\pi \gamma |\Omega_R|}{g^2}.$$  (14)

This result is essentially the same as Eq. (5) of [26]. We can now rearrange this expression to put an upper bound on the required power for a desired gate speed and error:

$$P = \frac{2\pi w_0 |\Omega_R|}{\lambda_{3/2} \epsilon_3} h \omega_{1/2} \Omega_R,$$  (17)

where $\lambda_{3/2} = c/\omega_{1/2}$.

Assume that the ratio of the beam waist to the transition wavelength is constant for different ion species. In this case, the power needed to obtain a given Rabi frequency and to keep the error below a given value would scale linearly with the optical transition frequency. A more realistic assumption might be that the Raman beam waist is not diffraction limited and is determined by other experimental considerations, such as the interior distance in the trap or beam pointing fluctuations. In this case, assuming that $w_0$ is constant, the required power would scale as the optical transition frequency cubed. Either way, ion species with optical transitions of longer wavelength are better suited in the sense that less power is required for the same gate speed and error requirements. In addition, high laser power is typically more readily available at longer wavelengths. Finally, we note that the error is independent of the fine-structure splitting as long as we have sufficient power to drive the transition. The transition wavelengths of different ions are listed in Table I.

Figure 3 shows the laser power needed per Raman beam for a given error due to spontaneous Raman scattering into the $S_{1/2}$ manifold. Here we assume $\Omega_R/2\pi = 0.25$ MHz ($\tau_R = 1$ $\mu$s) and $w_0 = 20$ $\mu$m. Different lines correspond to the different ion species listed on the figure legend. Table II lists the error in a single-ion-qubit gate due to Raman scattering.

![Fig. 3](image-url)
back into the $S_{1/2}$ manifold for the same parameters as in Fig. 3 and assuming 10 mW in each of the Raman beams. The power $P_0$ needed in each of the gate beams for $\epsilon_S = 10^{-4}$ is also listed in the table. As can be seen, ions with shorter transition wavelength require a larger detuning and accordingly higher laser power to maintain a low gate error. For most ions, a few milliwatts of laser power is enough to reduce the gate error to below $10^{-4}$.

Equation (13) can be solved to give the required detuning values for a given $\epsilon_S$. The number of such detuning values comes from the number of crossings of a horizontal line, set at the desired error level, with the solid curve in Fig. 2. When $\epsilon_S$ is higher than the minimum $P_{\text{Raman}}$ inside the fine-structure manifold—i.e., $\epsilon_S > 8\pi\gamma/3\omega_l$ (see Fig. 2)—then four different detuning values yield the same $\epsilon_S$, two outside and two inside the fine-structure manifold. When $\epsilon_S$ is lower than this value, only two detuning values, both of which are outside the fine-structure manifold, yield $\epsilon_S$. Those detuning values $\Delta_0$ that are below the $S_{1/2} \rightarrow P_{1/2}$ transition ($\Delta < 0$) and correspond to $\epsilon_S = 10^{-4}$ are listed in Table II for different ions. For most ions $\Delta_0$ is in the few hundred gigahertz range and its magnitude is much smaller than $\omega_l$.

We now consider the error for various ions caused by Raman scattering into low-lying $D$ levels $\epsilon_D$. As transitions between levels in the $S$ and $D$ manifolds do not necessarily involve electron spin rotations, the error suppression discussed preceding Eq. (13) will not occur. Instead, the Raman scattering rate into the $D$ levels will be given by the total scattering rate times a fixed branching ratio $f$. When driven resonantly the $P_{1/2}$ and $P_{3/2}$ levels decay to the $D$ manifold with different (but often similar) branching ratios. Here we assume that for a detuning large compared to the fine-structure splitting, the branching ratio is essentially independent of the laser detuning and is given by the average of the two resonant branching ratios [46,47]. Table I lists $f$, for various ion species, obtained from [48–52]. The error due to Raman scattering into $D$ levels is given by

$$\epsilon_D = f P_{\text{total}}.$$

Using Eq. (13) we can write the ratio of the errors due to Raman scattering into the different manifolds:

$$\epsilon_D/\epsilon_S = 3f/2 \left( 2\Delta^2 + (\Delta - \omega_l)^2 \right) / \omega_l^2.$$

(20)

For $|\Delta| < \omega_l$, the Raman scattering error is dominated by scattering back into the $S_{1/2}$ levels. For ions with $1/f \gg 1$ the two errors become comparable at a detuning $\Delta = \sqrt{2}\omega_l/3\sqrt{f}$. When the detuning becomes large compared to the fine-structure splitting, scattering into low-lying $D$ levels dominates. For most ions considered here, $|\Delta_0| < \omega_l$ and (perhaps with the exception of $^{137}$Ba$^+$) $\epsilon_S$ is the more dominant source of error. The ratios $\epsilon_D/\epsilon_S$ when $\epsilon_S = 10^{-4}$ (i.e., $\Delta = \Delta_0$) are given in Table II for different ions.

Due to the asymptotic value of the total scattering rate in the $|\Delta| > \omega_l$ limit [Eq. (6)], $\epsilon_D$ has an asymptotic value which gives a lower bound to the Raman scattering error:

$$\epsilon_{D_{\text{as}}} = 3\pi\gamma f / \omega_l.$$

(21)

Table II lists $\epsilon_{D_{\text{as}}}$ for various ions. For all ion species considered this value is below the assumed estimates for the fault tolerance threshold.

### B. Two-qubit gate

A universal quantum gate set is complete with the addition of two-qubit entangling gates. During the last few years there have been several proposals and realizations of two ion-qubit gates [1,8–12,53–55]. Here we focus on gates that use spin-dependent forces in order to imprint a geometric phase on certain collective spin states [8,9,11,12,53–55]. We examine only gates that are implemented with a continuous nonresonant pulse rather than those using multiple short pulses [56]. Again, to compare different ion species we examine ion qubits that are encoded into hyperfine clock states. When the laser detuning is large compared to the hyperfine splitting, the differential light force between clock levels is negligible [4,57]. However, a phase gate can be applied between spin states in the rotated basis (superpositions of clock states that lie on the equatorial plane of the Bloch sphere).
In this scheme the ions traverse a trajectory in phase space that is conditioned on their mutual wave vectors lying parallel to the trap axis. One beam of each pair is parallel to the magnetic field which sets the quantization direction. The beams’ polarizations in each pair are assumed to be linear, perpendicular to each other and to the magnetic field. Wavy arrows illustrate examples of photon scattering directions.

[53,54]. In this scheme the ions traverse a trajectory in phase space that is conditioned on their mutual spin state (in the rotated basis [11]). The phase the ions acquire is proportional to the total area encircled in phase space. This geometric phase gate was demonstrated in [8] and was realized on clock states in [11].

This form of phase gate is implemented with two different Raman fields that are slightly off resonance with upper and lower motional sidebands of the spin-flip transition. For simplicity, we assume here that the gate is driven by two independent pairs of Raman beams—i.e., a total of four beams. Most experimental implementations of this phase gate thus far have used a three-beam geometry [8,11]. It is straightforward to generalize the treatment below to the three-beam case.

Typical conditions for the gate are such that the angular frequency difference between the beams is \((\omega_0+\omega_{trap}-\delta)\) in one Raman pair and \((\omega_0-\omega_{trap}+\delta)\) in the other. Here \(\omega_{trap}\) is the angular frequency of the normal mode that the gate excites and \(\delta\) is the Raman field detuning from that motional sideband. Under these conditions the ions will traverse \(K\) full circles in phase space for a gate duration of \(\tau_{gate}=2\pi K/\delta\). Typically \(\delta\) is chosen to be much smaller than \(\omega_{trap}\) to avoid coupling to the pure spin flip (“carrier”) transition or the motional “spectator” mode and \(K\) is usually chosen to be 1 to minimize the gate time.

Figure 4 depicts the assumed geometry of the laser beams. The two beams comprising each Raman pair intersect at right angles at the position of the ions, such that the difference in their wave vectors is parallel to the trap axis. With this choice, the Raman fields couple, to a very high degree, only to the motion along the trap axis. The beam polarizations are assumed to be linear, perpendicular to each other and to the magnetic field axis. The beams’ relative frequencies can be arranged such that the final state is insensitive to the optical phase at the ions’ position [57]. Generalizing this treatment to other Raman beam geometries is straightforward.

As in the single-qubit case, Raman scattering will project one of the ion qubits into one of its states below the \(P\) manifold. In the appropriate basis the ideal gate operation is represented by [9,11]

\[
\hat{U} = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & e^{i\phi} & 0 & 0 \\
0 & 0 & e^{i\phi} & 0 \\
0 & 0 & 0 & 1
\end{pmatrix},
\]

where typically \(\phi=\pi/2\). The output state is \(|\Psi\rangle = \hat{U}|\Psi_{in}\rangle\).

The density matrix following the erroneous gate will be of the form

\[
\hat{\rho}_{\text{final}} = (1 - 2P_{\text{Raman gate}})|\Psi\rangle\langle\Psi| + \hat{\rho}_e,
\]

where \(P_{\text{Raman gate}}\) is the probability that one of the ions scattered a Raman photon during the gate [58,59]. For a gate consisting of \(K\) circles, the Raman detuning \(\delta\) is chosen such that the gate time is given by [53]

\[
\tau_{\text{gate}} = \frac{\pi}{2(\Omega_k|\eta|)} \sqrt{K} = \tau_0 \sqrt{K}. \tag{24}
\]

Here \(\eta=\Delta k z_0\) is the Lamb-Dicke parameter, where \(\Delta k = \sqrt{2k_1}\) is the wave vector difference between the two beams that drive the gate (for the particular geometry of Fig. 4) and \(k_1\) is the laser beam wave vector magnitude. The root mean square of the spatial spread of the ground-state wave function of one ion for the normal mode that the gate excites is

\[
z_0 = \sqrt{\hbar/4M\omega_{trap}}, \tag{25}
\]

where \(M\) is the mass of an individual ion. The single-ion carrier Rabi frequency \(\Omega_k\) is given in Eq. (1).

Using similar considerations to those used in the single-ion gate and, as before, neglecting any (positive) contribution of \(\hat{\rho}_e\) to the fidelity, we can set an upper bound for the power needed for a certain error in the gate due to Raman photon scattering into \(S_{1/2}\) levels:

\[
P = \frac{2\pi}{3\epsilon_x} \left( \frac{2\pi w_0}{\hbar \omega_{trap}} \right)^2 \hbar \omega_{trap} |\Omega_k| \frac{4\sqrt{K}}{\eta}. \tag{26}
\]

This required gate power is \(4\sqrt{K}/\eta\) times larger than that needed for the same error in the single-ion-qubit \(\pi\) rotation given in Eq. (18). The factor of \(4\sqrt{K}/\eta\) is due to the longer two-ion-qubit gate duration compared to single-qubit rotations, and the factor of 4 is due to the presence of two ions and the pair of required Raman fields.

When comparing different ion species we can fix different parameters, depending on experimental constraints or requirements. For example, here we choose as fixed parameters the beam waist \(w_0\) (for the reasons given above), the gate time \(\tau_{\text{gate}}\) (assuming a certain computation speed is desired), and the mode frequency \(\omega_{trap}\) (which sets the time scale for various gates). With these choices, heavier ions pay the price of smaller \(\eta\) and therefore higher-power requirements per given gate time and error. A different approach would be to choose \(\eta\) fixed, in which case heavier ions will need a lower
For a fixed gate time a lower $\omega_{\text{trap}}$ leads to stronger, off-resonant, coupling of the Raman fields to the carrier or even larger due to this coupling. The Lamb-Dicke parameter $\eta$ for the above trap frequency is also listed for different ions.

### Table III

<table>
<thead>
<tr>
<th>Ion</th>
<th>$10^4 \varepsilon_s$</th>
<th>$P_0$ (mW)</th>
<th>$\Delta_0/2\pi$ (THz)</th>
<th>$\varepsilon_{\text{scat}}/\varepsilon_s$</th>
<th>$10^4 \varepsilon_{\text{total}}$</th>
<th>$\eta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^9\text{Be}^+$</td>
<td>3.6</td>
<td>36</td>
<td>$-1.20$</td>
<td>N.A.</td>
<td>N.A.</td>
<td>0.194</td>
</tr>
<tr>
<td>$^{25}\text{Mg}^+$</td>
<td>11.1</td>
<td>111</td>
<td>$-7.28$</td>
<td>N.A.</td>
<td>N.A.</td>
<td>0.130</td>
</tr>
<tr>
<td>$^{43}\text{Ca}^+$</td>
<td>13.6</td>
<td>136</td>
<td>$-10.42$</td>
<td>1.01</td>
<td>1.06</td>
<td>0.071</td>
</tr>
<tr>
<td>$^{67}\text{Zn}^+$</td>
<td>41.1</td>
<td>411</td>
<td>$-24.96$</td>
<td>N.A.</td>
<td>N.A.</td>
<td>0.11</td>
</tr>
<tr>
<td>$^{87}\text{Sr}^+$</td>
<td>26.5</td>
<td>265</td>
<td>$-20.34$</td>
<td>0.52</td>
<td>0.50</td>
<td>0.048</td>
</tr>
<tr>
<td>$^{111}\text{Cd}^+$</td>
<td>64.3</td>
<td>643</td>
<td>$-35.44$</td>
<td>N.A.</td>
<td>N.A.</td>
<td>0.081</td>
</tr>
<tr>
<td>$^{137}\text{Ba}^+$</td>
<td>37.4</td>
<td>374</td>
<td>$-30.67$</td>
<td>1.65</td>
<td>1.46</td>
<td>0.034</td>
</tr>
<tr>
<td>$^{171}\text{Yb}^+$</td>
<td>57.5</td>
<td>575</td>
<td>$-32.89$</td>
<td>0.01</td>
<td>0.007</td>
<td>0.038</td>
</tr>
<tr>
<td>$^{199}\text{Hg}^+$</td>
<td>149.7</td>
<td>1497</td>
<td>$-49.52$</td>
<td>0.003</td>
<td>0.001</td>
<td>0.078</td>
</tr>
</tbody>
</table>

For the choice where $w_0$, $\tau_{\text{gate}}$, and $\omega_{\text{gate}}$ are fixed, we can write

$$P = \frac{8\pi^2}{3\varepsilon_s} \frac{K^3}{\tau_{\text{gate}}^3} w_0^2 \omega_{3/2} M \omega_{\text{trap}}$$

or, equivalently,

$$\varepsilon_s = \frac{8\pi^2}{3P} \frac{K^3}{\tau_{\text{gate}}^3} w_0^2 \omega_{3/2} M \omega_{\text{trap}}$$

Table III lists $\varepsilon_{D_{\text{scat}}}$ for the different ion species.

![Figure 5](image-url)
III. RAYLEIGH SCATTERING ERROR

Since Rayleigh photon scattering is elastic, no energy or angular momentum is transferred between the photons’ and ions’ internal degrees of freedom. Therefore, these degrees of freedom remain uncorrelated. Rayleigh scattering does not necessarily lead to direct spin decoherence.

In situations where Rayleigh scattering rates from the two-ion-qubit states are different, Rayleigh scattering of photons will eventually measure the qubit state and lead to decoherence. In fact, the most common ion-qubit detection method relies on state-selective Rayleigh scattering of photons on a cycling transition. In most Raman gates, however, the laser is detuned from resonance by much more than the qubit levels’ energy separation, \( \Delta \gg \omega_0 \), and the laser polarization is typically linear to suppress differential Stark shifts. Under these conditions Rayleigh scattering rates from the two qubit levels are almost identical and the error due to the rate difference is negligible. For the clock transition qubit states considered here, Rayleigh scattering rates are almost identical, regardless of the laser polarization (for more details on this error see Sec. III B).

The main effect of Rayleigh scattering is the momentum recoil it imparts to the ion qubit. For a single-qubit gate, the Raman beams are usually arranged in a copropagating geometry. In this configuration, since the Lamb-Dicke parameter is very small \( \eta = (\omega_0 / \omega_{rad} ) k \), the effect of ion motion is negligible on the gate operation. Therefore, Rayleigh scattering has a negligible effect on single-qubit gates.

In the two-qubit phase gate, a mode of motion is excited that is entangled with the two-ion collective spin state. In this case, recoil from photon scattering perturbs the ion’s motion through phase space and contributes to the gate error. The ion-qubit trajectory is distorted in two ways. The larger distortion arises from the direct recoil momentum displacement. A second, much smaller, distortion arises from the contribution of the recoil to the appearance of nonlinearities in the gate operation due to deviations from the Lamb-Dicke regime (see Sec. III B).

In Sec. III A we calculate the error in a two-qubit gate due to direct recoil phase-space displacement. In Sec. III B we elaborate on the two other sources of error mentioned above: namely, errors due to uneven Rayleigh scattering rates and errors due to deviations from the Lamb-Dicke regime.

A. Rayleigh scattering recoil error

In elastic Rayleigh scattering, energy and momentum are not exchanged between the ions’ and photons’ internal degrees of freedom. However, momentum and energy are exchanged between the photons’ and ions’ external degrees of freedom. The scattered photon direction will be different from that of the laser beam, causing the ion to recoil. This recoil acts on the ion qubit as a phase-space momentum displacement, distorting the ion’s trajectory through phase space and causing an error in the entangling-gate phase. Note that the momentum imparted to the ion (and therefore the deviation from the desired gate phase) and the momentum that is carried by the scattered photon (namely, its scattering direction) are correlated. From this point of view the gate infidel-

ity again arises due to the entanglement between the scattered photons’ and ions’ (this time external) degrees of freedom (for more details on this point of view see the Appendix).

In the Lamb-Dicke regime (for a thermal state \( \eta^{3/2} \approx 1 \), where \( \eta \) is the average mode population), the gate operation can be approximated as a series of finite displacements in phase space [9]:

\[
\hat{U} = \prod_{k=1}^{N} \hat{D}(\Delta \alpha_k).
\]

Displacements through phase space, \( \Delta \alpha_k \), are conditioned on the joint spin state of the two ions and depend on the gate parameters. For certain gate parameters, the displacement is zero for the two parallel spin states \( |\uparrow\uparrow\rangle \) and \( |\downarrow\downarrow\rangle \) and nonzero with opposite sign for the \( |\uparrow\downarrow\rangle \) and \( |\downarrow\uparrow\rangle \) states, where \( \uparrow \) and \( \downarrow \) hereafter refer to the rotated basis rather than the clock levels. Using the commutation relations between phase-space displacements, we can write the gate operation as a single displacement times an overall phase [9]:

\[
\hat{U} = \hat{D} \left( \sum_{k=1}^{N} \Delta \alpha_k \right) \exp \left[ i \frac{\operatorname{Im}(\alpha \beta)}{2} \hat{D}(\alpha + \beta) \right],
\]

where \( N \) is the total number of infinitesimal displacements. When the net displacement is zero—i.e., \( \Sigma \Delta \alpha_k = 0 \)—the two-ion motion returns to its initial state and spin and motion are disentangled at the end of the gate. The gate operation on the affected spin states can be written as

\[
\hat{U} = \hat{I} e^{i \phi}.
\]

The phase acquired,

\[
\phi = \text{Im} \left( \sum_{j=2}^{N} \Delta \alpha_j \sum_{l=1}^{j-1} \Delta \alpha_l \right),
\]

is the same for the \( |\uparrow\downarrow\rangle \) and \( |\downarrow\uparrow\rangle \) states and proportional to the enclosed phase-space area. Figure 6(a) shows an example for the trajectory traversed in phase space during an ideal gate in a reference frame rotating at the mode frequency \( \omega_{rad} \).

In the presence of laser light there is a finite probability \( P_{\text{Rayleigh}} \) that a Rayleigh photon will be scattered by one of the ions during the gate [58]. The probability for a Rayleigh scattering event to occur during a single-qubit gate is given by the difference between Eqs. (4) and (13):

\[
P_{\text{Rayleigh}} = \frac{\pi \gamma 3 \Delta^2 - 2 \Delta \omega_0 + \omega_0^2/3}{|\Delta(\Delta - \omega_0)|}.
\]

Since in the limit of \( |\Delta| \gg \omega_0 \) all scattering events are Rayleigh scattering, \( P_{\text{Rayleigh}} \) and \( P_{\text{total}} \) have the same asymptotic value \( P_{\text{ext}} \).

Using Eq. (36) and the factor for the extra required power in the two-qubit gate,
The erroneous gate can be again written as a sum of displacements, those due to the gate drive and recoil, which occurs at some random time during the gate. For a particular two-ion spin state,

\[ \hat{U}_e = \prod_{i=0}^{N} \hat{D}(\Delta \alpha_i) \hat{D}(\beta) \prod_{i=0}^{M} \hat{D}(\Delta \alpha_i). \] (41)

By use of the commutation relation (32), the gate is written as

\[ \hat{U}_e = \hat{D}(\beta) e^{i(\Delta \phi + \Delta \phi)}, \] (42)

where the phase error \( \Delta \phi \) is determined to be

\[ \Delta \phi = \text{Im} \left[ \beta \sum_{i=1}^{N} \Delta \alpha_i^* \sum_{k=0}^{M} \Delta \alpha_k \beta^* \right]. \] (43)

Writing displacements as a function of the gate time

\[ \Delta \alpha_k = \alpha(t_k) = \frac{\partial \alpha(t)}{\partial t} dt, \] (44)

we can write the phase error as a function of the gate displacement before and after the scattering time \( t_{\text{scat}} \):

\[ \Delta \phi = \text{Im} \left[ \beta \int_{t_{\text{scat}}}^{t_{\text{gate}}} \frac{\partial \alpha(t)}{\partial t} dt + \beta^* \int_{t_{\text{scat}}}^{t_{\text{gate}}} -\frac{\partial \alpha(t)}{\partial t} dt \right]. \] (45)

Since \( \alpha(t_{\text{gate}}) = 0 \), we get

\[ \Delta \phi = \text{Im} \left[ \beta \alpha(t_{\text{scat}}) - \beta^* \alpha(t_{\text{scat}}) \right]. \] (46)

Since the accumulated displacement at the moment the photon was scattered \( \alpha(t_{\text{scat}}) \) is of equal magnitude but opposite sign for the \( |\uparrow\downarrow\rangle \) and \( |\downarrow\uparrow\rangle \) states, we get

\[ \Delta \phi_{\uparrow\downarrow} = -\Delta \phi_{\downarrow\uparrow} = \Delta \phi. \] (47)

The erroneous gate can be therefore represented by the operator

\[ \hat{U}_e = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & e^{i(\Delta \phi + \Delta \phi)} & 0 & 0 \\ 0 & 0 & e^{i(\Delta \phi - \Delta \phi)} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \] (48)

The final-state fidelity following the erroneous gate depends on \( |\Psi_{\text{init}}\rangle \). For a general initial state,

\[ |\Psi_{\text{init}}\rangle = \mu |\uparrow\uparrow\rangle + \sigma |\uparrow\downarrow\rangle + \gamma |\downarrow\uparrow\rangle + \delta |\downarrow\downarrow\rangle, \] (49)

and using Eq. (7), we can write for the gate fidelity

\[ F = \| \mu \|^2 + e^{i\Delta \phi} |\sigma|^2 + e^{-i\Delta \phi} |\gamma|^2 + |\delta|^2. \] (50)

The most relevant fidelity for fault tolerance considerations is that averaged over all possible initial states. The fidelity we calculate here is that due to the worst-case input state: that is, the input state that minimizes Eq. (50). The worst-case fidelity is clearly smaller than the average fidelity and
therefore gives a conservative estimate to the gate error [62]. The worst-case fidelity for this kind of error was calculated in [63]. For $\Delta \phi / \phi < 1$ it is

$$F = \cos^2 \Delta \phi.$$ (51)

This minimal fidelity is the result of an input state with $|\chi|=|\gamma|=1/2$.

Photon scattering occurs in only a small fraction $P_{\text{Rayleigh gate}}$ of the gates. Averaging over all gates performed, we get

$$F_{\text{gate}} = 1 - P_{\text{Rayleigh gate}}(1 - \langle \cos^2 \Delta \phi \rangle) = 1 - \epsilon_R.$$ (52)

where the angular brackets correspond to an average over all possible $\Delta \phi$'s—i.e., due to different values of $\beta$ and different scattering times.

To perform the above average we need to explicitly write the different displacements. In the rotating frame the cumulative gate displacement after a time $t$ can be written as [9,53]

$$\alpha(t) = \frac{i}{2\sqrt{K}}[e^{-i\hat{\beta}} - 1]e^{i\Phi_L},$$ (53)

where $\Phi_L$ is the gate phase, determined by the optical phase difference between the gate beams at the ion’s position. In the rotating frame, the recoil displacement can be written as

$$\beta = |\beta|e^{i\omega_{\text{Ray}}t}.$$ (54)

Substituting Eqs. (54) and (53) into Eq. (46), we can write the averaged term in the fidelity:

$$\langle \cos^2 \Delta \phi \rangle = \frac{\delta}{2\pi} \int_{|\beta|=\delta} |\beta|_{\text{max}} S|\beta| \cos \left( \frac{|\beta|}{\sqrt{K}} \cos(\omega_{\text{Ray}}t + \delta t - \Phi_L) - \Phi_L \right) d\beta.$$ (55)

Here we assume that the probability of scattering at different time intervals during the gate is uniform. The probability distribution for different recoil displacement magnitude $|\beta|$ is given by $S(|\beta|)$, which is determined by the geometry of the Raman beams with respect to the trap axis and the probability distribution of photon scattering directions.

With the laser beam configuration assumed in Fig. 4, recoil due to photon absorption can be imparted only at 45° to the trap axis, whereas recoil due to photon emission can be imparted in any direction. The maximum allowed displacement following an absorption-emission cycle is therefore

$$|\beta|_{\text{max}} = k_L Z_0(1 + \sqrt{2})/2 = \frac{\eta}{2}(1 + \sqrt{2}).$$ (56)

For proper gate operation we require $\eta << 1$; thus we expand Eq. (55) in powers of $|\beta|$, including terms to order $|\beta|^2$. Further, since typically $\delta \omega_{\text{Ray}} \ll 1$, we neglect terms proportional to $\delta \omega_{\text{Ray}}$. With these approximations the gate fidelity is independent of $\Phi_L$ and is given by

$$F_{\text{gate}} = 1 - P_{\text{Rayleigh gate}}(1 - \langle \cos^2 \Delta \phi \rangle) = 1 - \epsilon_R.$$ (52)

where the angular brackets correspond to an average over all possible $\Delta \phi$'s—i.e., due to different values of $\beta$ and different scattering times.

To perform the above average we need to explicitly write the different displacements. In the rotating frame the cumulative gate displacement after a time $t$ can be written as [9,53]

$$\alpha(t) = \frac{i}{2\sqrt{K}}[e^{-i\hat{\beta}} - 1]e^{i\Phi_L},$$ (53)

where $\Phi_L$ is the gate phase, determined by the optical phase difference between the gate beams at the ion’s position. In the rotating frame, the recoil displacement can be written as

$$\beta = |\beta|e^{i\omega_{\text{Ray}}t}.$$ (54)

Substituting Eqs. (54) and (53) into Eq. (46), we can write the averaged term in the fidelity:

$$\langle \cos^2 \Delta \phi \rangle = \frac{\delta}{2\pi} \int_{|\beta|=\delta} |\beta|_{\text{max}} S|\beta| \cos \left( \frac{|\beta|}{\sqrt{K}} \cos(\omega_{\text{Ray}}t + \delta t - \Phi_L) - \Phi_L \right) d\beta.$$ (55)

Here we assume that the probability of scattering at different time intervals during the gate is uniform. The probability distribution for different recoil displacement magnitude $|\beta|$ is given by $S(|\beta|)$, which is determined by the geometry of the Raman beams with respect to the trap axis and the probability distribution of photon scattering directions.

With the laser beam configuration assumed in Fig. 4, recoil due to photon absorption can be imparted only at 45° to the trap axis, whereas recoil due to photon emission can be imparted in any direction. The maximum allowed displacement following an absorption-emission cycle is therefore

$$|\beta|_{\text{max}} = k_L Z_0(1 + \sqrt{2})/2 = \frac{\eta}{2}(1 + \sqrt{2}).$$ (56)

For proper gate operation we require $\eta << 1$; thus we expand Eq. (55) in powers of $|\beta|$, including terms to order $|\beta|^2$. Further, since typically $\delta \omega_{\text{Ray}} \ll 1$, we neglect terms proportional to $\delta \omega_{\text{Ray}}$. With these approximations the gate fidelity is independent of $\Phi_L$ and is given by

$$F_{\text{gate}} = 1 - P_{\text{Rayleigh gate}}(1 - \langle \cos^2 \Delta \phi \rangle) = 1 - \epsilon_R.$$ (52)

where the angular brackets correspond to an average over all possible $\Delta \phi$'s—i.e., due to different values of $\beta$ and different scattering times.

To perform the above average we need to explicitly write the different displacements. In the rotating frame the cumulative gate displacement after a time $t$ can be written as [9,53]

$$\alpha(t) = \frac{i}{2\sqrt{K}}[e^{-i\hat{\beta}} - 1]e^{i\Phi_L},$$ (53)

where $\Phi_L$ is the gate phase, determined by the optical phase difference between the gate beams at the ion’s position. In the rotating frame, the recoil displacement can be written as

$$\beta = |\beta|e^{i\omega_{\text{Ray}}t}.$$ (54)

Substituting Eqs. (54) and (53) into Eq. (46), we can write the averaged term in the fidelity:

$$\langle \cos^2 \Delta \phi \rangle = \frac{\delta}{2\pi} \int_{|\beta|=\delta} |\beta|_{\text{max}} S|\beta| \cos \left( \frac{|\beta|}{\sqrt{K}} \cos(\omega_{\text{Ray}}t + \delta t - \Phi_L) - \Phi_L \right) d\beta.$$ (55)

Here we assume that the probability of scattering at different time intervals during the gate is uniform. The probability distribution for different recoil displacement magnitude $|\beta|$ is given by $S(|\beta|)$, which is determined by the geometry of the Raman beams with respect to the trap axis and the probability distribution of photon scattering directions.

With the laser beam configuration assumed in Fig. 4, recoil due to photon absorption can be imparted only at 45° to the trap axis, whereas recoil due to photon emission can be imparted in any direction. The maximum allowed displacement following an absorption-emission cycle is therefore

$$|\beta|_{\text{max}} = k_L Z_0(1 + \sqrt{2})/2 = \frac{\eta}{2}(1 + \sqrt{2}).$$ (56)

For proper gate operation we require $\eta << 1$; thus we expand Eq. (55) in powers of $|\beta|$, including terms to order $|\beta|^2$. Further, since typically $\delta \omega_{\text{Ray}} \ll 1$, we neglect terms proportional to $\delta \omega_{\text{Ray}}$. With these approximations the gate fidelity is independent of $\Phi_L$ and is given by

$$F_{\text{gate}} = 1 - P_{\text{Rayleigh gate}}(1 - \langle \cos^2 \Delta \phi \rangle) = 1 - \epsilon_R.$$ (52)

where the angular brackets correspond to an average over all possible $\Delta \phi$'s—i.e., due to different values of $\beta$ and different scattering times.

To perform the above average we need to explicitly write the different displacements. In the rotating frame the cumulative gate displacement after a time $t$ can be written as [9,53]
leig photon will be scattered during the gate is

\[ P_{\text{Rayleigh gate}} = \frac{4\sqrt{K}P_r}{\eta}. \quad (63) \]

The recoil error in a two-qubit gate is therefore asymptotically bound by

\[ \epsilon_{\text{Rayleigh}} = \frac{5\pi \eta \gamma}{2 \sqrt{K} a_t} \approx \frac{5\pi^3 \gamma}{a_t K} \sqrt{\frac{h}{2M\omega_{\text{trap}}K^2}}. \quad (64) \]

Table IV lists \( \epsilon_{\text{Rayleigh}} \) for different ion species for \( \omega_{\text{trap}}/2\pi = 5 \text{ MHz} \) and a single-circle \((K=1)\) gate. With the exception of \( ^9\text{Be}^+ \), the error due to photon recoil in a two-ion-qubit gate is below \( 10^{-4} \). For this error heavier ions benefit due to their smaller recoil.

It is possible to reduce \( \epsilon_{\text{Rayleigh}} \) by choosing a smaller \( \eta/\sqrt{K} \) (by increasing the trap frequency and/or performing multiple-circle gates). Correspondingly more laser power, proportional to \( K/\eta^2 \), is then required in order not to increase the Raman scattering error \( \epsilon_{\text{R}} \) or reduce the gate speed.

### B. Other Rayleigh scattering errors

In Sec. III A we calculated the error due to the recoil imparted to the ion qubit during Rayleigh photon scattering. As noted, Rayleigh scattering of photons adds two more contributions to the gate error.

The first contribution \( \epsilon_{\text{R}} \) is due to a difference in the Rayleigh scattering rates from the two ground-state levels. Assume that the Rayleigh scattering rate is \( \Gamma_{\text{Rayleigh}} \) from one qubit level and \( \Gamma_{\text{Rayleigh}} + d\Gamma \) from the other. A measurement of the qubit level will be conclusive once the difference in the number of photons that are scattered is larger than the standard deviation of the number of photons scattered from each level. Since the number of photons that are scattered follows a Poisson distribution, a measurement will occur after photons have been scattered for a time \( t \) such that

\[ d\Gamma t > \sqrt{\Gamma_{\text{Rayleigh}}}. \quad (65) \]

The error rate is, therefore, given by \( d\Gamma t / \Gamma_{\text{Rayleigh}} \) and the error during a gate by

\[ \epsilon_{\text{R}} = (d\Gamma t / \Gamma_{\text{Rayleigh}})^2 P_{\text{Rayleigh}}. \quad (66) \]

For the above gate parameters, the difference in scattering rates is due solely to the difference in detuning between the two qubit levels. For \( \Delta \gg \omega_0 \), which is typically required to reduce other scattering errors to low levels, this error can be approximated by \( \epsilon_{\text{R}} = (\omega_0 / \Delta)^2 P_{\text{Rayleigh}} \). The contribution of \( \epsilon_{\text{R}} \) to the total error at a realistic laser detuning is very small.

We performed an accurate calculation of the difference in the scattering rates between the two clock levels and verified that indeed, for all ion species and laser detunings discussed above, \( \epsilon_{\text{R}} \) is negligible.

A second source of Rayleigh scattering error is through the contribution of the recoil momentum displacement to nonlinearities in the gate evolution. The gate error due to nonlinearities was calculated in [61] and in the present context is proportional to \( P_{\text{Rayleigh gate}} \eta^2 \text{Var}(n) \), where \( \text{Var}(n) \) is the variance of the motional-mode distribution due to recoil. The recoil momentum displacement magnitude is of order \( \eta \). Therefore, starting from the ground state and following a single scattering event, \( \text{Var}(n) = \eta \). The gate error due to this effect will be proportional to \( P_{\text{Rayleigh gate}} \eta^2 \eta^2 \), and significantly smaller than other scattering errors discussed above.

### IV. DISCUSSION

We have calculated the errors due to photon scattering in single-pulse single-qubit gates and two-qubit phase gates implemented with stimulated Raman transitions. These errors present a fundamental limit to the gate fidelity in trapped-ion QIP experiments that use these kinds of gates and should be a significant factor when choosing a specific ion as a quantum information carrier for fault-tolerant quantum computing schemes.

Three main errors occur from spontaneously scattering photons during a gate. Two errors are due to Raman scattering either back into the \( S_{1/2} \) manifold \( (\epsilon_{\text{D}}) \) or to low-lying \( D \) levels \( (\epsilon_{\text{D}}) \). The third error is due to the Rayleigh scattering recoil during a two-qubit gate \( (\epsilon_{\text{R}}) \). For most ions currently considered for QIP experiments, the dominant error for realistic laser parameters is \( \epsilon_{\text{D}} \). This error can be typically reduced to below current estimates for the fault tolerance threshold with the use of relatively high (but probably attainable) laser power. This makes the availability of high-power laser sources at the relevant wavelength important. This error is also reduced for ions with a relatively large \( S \rightarrow P \) transition wavelength.

Among those three errors, only \( \epsilon_{\text{D}} \) cannot be reduced below a certain value by the use of higher laser intensity. Therefore, it may eventually be advantageous to choose ions that do not have low-lying \( D \) levels. However, for most ions considered here, \( \epsilon_{\text{D}} \) is still small at realistic laser parameters compared to the other Raman error \( \epsilon_{\text{R}} \).

The masses of the different ions play an interesting role in the scattering error. Since it is harder to transfer momentum to heavier ions, they suffer from a larger \( \epsilon_{\text{R}} \) in two-qubit gates (or alternatively from the need for more laser power for a given gate speed and error level). For the same reason, lighter ions suffer from a larger Rayleigh recoil error \( \epsilon_{\text{R}} \). Examining Eqs. (28) and (64), the Raman error scales as \( \omega_{\text{trap}}^2 M \omega_{\text{trap}} \) and the Rayleigh error scales as \( 1/\omega_0 \omega_{\text{trap}} \) (neglecting differences in wavelength and natural linewidths). Since both \( \omega_0 \) and \( \omega_{\text{trap}} \) are generally larger for heavier ions, the Raman error is larger and Rayleigh error is smaller for heavier ions by more than is indicated by the kinetic argument above. Since for most ions \( \epsilon_{\text{R}} \) is the dominant error, lighter ions seem to currently have a lower overall error due to photon scattering. Also, because both errors are a function of \( \omega_{\text{trap}} \), lighter ions will reach the same error level and power requirements with a higher \( \omega_{\text{trap}} \) when compared to heavier ions. This allows for faster gate operation. For a given trap geometry and applied potentials, the axial trap frequency scales as \( 1/\sqrt{M} \) and the radial trap frequencies scale as \( 1/M \). For the ion crystal to remain along the axial trap direction the radial frequencies have to be larger than the axial frequencies; therefore, the limiting frequency is the ra-
dial frequency. In this case, $M_{\omega_{\text{trap}}}$ is independent of the ions’ mass. Heavier ions under these conditions will have a lower $\omega_{\text{trap}}$, leading to slower gate operation.

In conclusion, to minimize the effect of scattering on the fidelity of trapped-ion-qubit gates, one needs to strike a balance between the desirable characteristics of long wavelength, light mass, the availability of high-power laser sources and, if possible, the lack of low-lying $D$ levels when choosing a specific ion as a quantum-information carrier.

Finally, we remind the reader that here we have focused on hyperfine ion qubits and gates that rely on off-resonant Raman transitions applied in a continuous pulse. Other kinds of trapped-ion gates or qubits could have different limitations on the gate fidelity due to spontaneous photon scattering.

ACKNOWLEDGMENTS

We thank E. Knill for many insightful discussions and comments and J. Home, J. Amini, and A. Sørensen for helpful comments. This work is supported by DTO and NIST.

APPENDIX: DEPHASING vs ENTANGLEMENT WITH DIFFERENT PHOTON MODES

In Sec. III we calculated the error due to Rayleigh photon scattering without considering the scattered photon degrees of freedom. Rather, we considered dephasing due to the random phase that is generated by the scattered photon’s recoil.

As noted, we can also view the gate error caused by Rayleigh scattering of photons as arising from the entanglement between the photon and the ion-qubit external degrees of freedom. For example, a photon that is scattered during the gate, while the two spin wave packets are displaced from each other, can be collected by an ideal imaging system. Since the two-ion collective spin is entangled with their position, the position at which the photon is detected at the image plane can give “which-way” information, thereby measuring the ions’ spin, collapsing the entanglement, and causing an error in the gate. Experiments investigating atomic decoherence due to this effect were performed in neutral-atom interferometers [64,65]. This effect was also calculated in [66] for a single ion in a superposition of two different motional coherent states.

The general equivalence of these two points of view (dephasing versus entangling with the environment) was explained in [67]. Here we show this equivalence for the trapped ion case where the ion is undergoing an (ideally) spin-dependent closed-loop displacement. For simplicity we examine the case of a single ion; it is, however, straightforward to generalize the following proof to the case of two ions.

We first calculate the final-state fidelity by examining entanglement with the photon modes. As in the two-qubit gate, a single ion is prepared in an equal superposition of spin states and is cooled to the motional ground state:

$$|\Psi\rangle = \left( \frac{1}{\sqrt{2}} |\uparrow\rangle + \frac{1}{\sqrt{2}} |\downarrow\rangle \right) \otimes |0\rangle_M \otimes |0\rangle_k. \quad (A1)$$

Here kets with subscript $M$ represent motional states and kets with the subscript $k$ the different photon modes, which are initially empty (neglecting the laser mode). The ion is subsequently driven by an oscillating force which is detuned from its motional resonance. The direction of force is opposite for the two different spin states. Ideally the two spin states would traverse opposite circular trajectories in phase space (in a frame rotating at the motional mode frequency).

At the end of the gate drive, both parts of the superposition (ideally) return to the ground state of motion and acquire the same geometric phase $\phi_{\text{gate}}$ which is proportional to the phase-space area encircled. The state at the end of the ideal gate drive is

$$|\Psi_{\text{ideal}}\rangle = e^{i\phi_{\text{gate}}} \left( \frac{1}{\sqrt{2}} |\uparrow\rangle + \frac{1}{\sqrt{2}} |\downarrow\rangle \right) \otimes |0\rangle_M \otimes |0\rangle_k. \quad (A2)$$

Now assume the ion scattered a photon during the gate drive. Immediately before the scattering event, the gate evolution produces

$$|\Psi\rangle = \left( \frac{1}{\sqrt{2}} e^{i\phi} |\uparrow\rangle |\alpha\rangle + \frac{1}{\sqrt{2}} e^{-i\phi} |\downarrow\rangle |\alpha\rangle \right) \otimes |0\rangle_k. \quad (A3)$$

where $\pm \alpha$ are the conditional phase-space displacement for the two spin states (we have dropped the subscript $M$) and $\phi$ is the geometric phase accumulated by the time of scattering. Immediately after scattering, a single photon is created in mode $k'$ and the ion correspondingly recoils:

$$|\Psi\rangle = \frac{1}{4\pi} \int \delta(E_k') \left( \frac{1}{\sqrt{2}} e^{i\phi} e^{i q \cdot \hat{r}} |\uparrow\rangle |\alpha\rangle + \frac{1}{\sqrt{2}} e^{-i\phi} e^{i q \cdot \hat{r}} |\downarrow\rangle |\alpha\rangle \right) \otimes |1\rangle_{k'} \otimes |0\rangle_k d\vec{k}'. \quad (A4)$$

Here $q' = \vec{k}' - \vec{k}$ is the wave vector difference between the scattered photon and the laser beam. The delta function $\delta(E_k')$ enforces energy conservation. The electromagnetic field is now represented by a superposition of states in which a single photon was scattered into a certain mode while all other modes are empty. Each part of this superposition is correlated with the corresponding momentum recoil operator acting on the trapped ion. Recoil into directions other than the trap axis will give rise to motion which is common to both parts of the superposition and can therefore be traced over. We therefore neglect recoil into the dimensions other than the trap axis direction and approximate $e^{i q \cdot \hat{r}} \approx e^{i q_z \hat{z}}$, where $q_z' = q_z$. By use of Eq. (40) this momentum recoil can be written as a phase-space displacement $e^{i q_z' \hat{z}} = D(\beta')$, where $\beta' = iq_z' z_0$ and $z_0$ is the root mean square of the ground-state spatial spread. Using Eq. (32) we add the recoil to the gate displacement and the state after scattering can be written as

$$|\Psi\rangle = \frac{1}{4\pi} \int \delta(E_k') \left( \frac{1}{\sqrt{2}} e^{i\phi} e^{i \text{Im}(\beta' a^*) |\uparrow\rangle |\alpha + \beta'\rangle} + \frac{1}{\sqrt{2}} e^{-i\phi} e^{i \text{Im}(\beta' a^*) |\downarrow\rangle |\alpha + \beta'\rangle} \right) \otimes |k'\rangle d\vec{k}'. \quad (A5)$$

where
\[ |k\rangle = |1\rangle_k \otimes |0\rangle_k. \]
Since the total gate displacement is equal to zero, the remaining part of the gate displacement is \( \mp \alpha \) (depending on the spin state). Using Eq. (32) again, we can write the state of the system after the gate has completed:
\[
|\Psi\rangle = \frac{1}{4\pi} \int \delta(E_{k'}) \left( \frac{1}{\sqrt{2}} e^{i\phi_{gate} - i\alpha |\beta|^* |\beta'|} |\uparrow\rangle \langle \beta'\rangle + \frac{1}{\sqrt{2}} e^{i\phi_{gate} + i\alpha |\beta|^* |\beta'|} |\downarrow\rangle \langle \beta'\rangle \right) \otimes |k\rangle \otimes |k\rangle \otimes |k\rangle.
\]
We can now define the gate phase error \( \Delta \phi' = \text{Im}(\alpha |\beta|^* - \alpha |\beta'|) \) and write the corresponding density operator
\[
\hat{\rho} = \frac{1}{2(4\pi)^2} \int \delta(E_{k'}) \delta(E_{k''}) \left( e^{-i\Delta \phi'} |\uparrow\rangle \langle \uparrow| + e^{-i\Delta \phi'} |\downarrow\rangle \langle \downarrow| + e^{i\Delta \phi'} e^{-i\Delta \phi''} |\downarrow\rangle \langle \downarrow| \right) \otimes |\beta\rangle \langle \beta| \otimes |k\rangle \otimes |k\rangle \otimes |k\rangle.
\]
Since we have no information about the mode into which the photon was scattered, we reduce the above density matrix with a trace over the photon modes, using \( \text{tr}(|k\rangle \langle k'|) = 4\pi |\delta(k-k')| \)
\[
\hat{\rho} = \frac{1}{8\pi} \int \delta(E_{k'}) \left( \frac{1}{2} |\uparrow\rangle \langle \uparrow| + \frac{1}{2} |\downarrow\rangle \langle \downarrow| + \frac{1}{2} e^{i\Delta \phi'} |\uparrow\rangle \langle \downarrow| + \frac{1}{2} e^{-i\Delta \phi'} |\downarrow\rangle \langle \uparrow| \right) \otimes |\beta\rangle \langle \beta| \otimes |k\rangle \otimes |k\rangle \otimes |k\rangle.
\]
\[
\hat{\rho} = \frac{1}{8\pi} \int \delta(E_{k'}) \left( \frac{1}{2} |\uparrow\rangle \langle \uparrow| + \frac{1}{2} |\downarrow\rangle \langle \downarrow| + \frac{1}{2} e^{i\Delta \phi'} |\uparrow\rangle \langle \downarrow| + \frac{1}{2} e^{-i\Delta \phi'} |\downarrow\rangle \langle \uparrow| \right) \otimes |\beta\rangle \langle \beta| \otimes |k\rangle \otimes |k\rangle \otimes |k\rangle.
\]
Note that the coherences of the density matrix in Eq. (A8) are given by an average of \( e^{i\Delta \phi'} \). A large spread in \( \Delta \phi' \) will "wash out" coherence and leave an incoherent statistical mixture. Tracing over the motional degrees of freedom we are now ready to evaluate the fidelity with respect to the ideal gate output state [Eq. (A2)],
\[
F = \frac{1}{2} + \frac{1}{2} \langle \cos(2\Delta \phi') \rangle = \langle \cos^2 \Delta \phi' \rangle,
\]
where averaging is performed with respect to all of the photon modes:
\[
\langle \cos^2 \Delta \phi' \rangle = \frac{1}{4\pi} \int \delta(E_{k'}) \cos^2 \Delta \phi' \, d\kappa'.
\]
We now turn to calculate the fidelity looking only at dephasing as in Sec. III A—i.e., loss of fidelity due to the photon scattering random phase. The gate output state is then
\[
|\Psi_{\text{ideal}}\rangle = e^{i\phi_{\text{gate}}} \left( \frac{1}{\sqrt{2}} |\uparrow\rangle + \frac{e^{i\phi}}{\sqrt{2}} |\downarrow\rangle \right).
\]
where \( \Delta \phi \) is the random scattering phase [Eq. (46)]. The fidelity of this state with respect to the ideal gate output state is identical to that calculated considering entanglement with the photon modes in Eq. (A9).
In situations where the detuning is not large compared to the wavelength of light. Therefore, scattering events of the two ions are not correlated.

Since the off-resonant coupling to the carrier scales as \(1/\eta\), even with our choice of constant \(\omega_{\text{trap}}\), lighter ions suffer less from errors due to this coupling. Furthermore, lighter ions require lower rf trap potentials for the same \(\omega_{\text{trap}}\).

For realistic parameters \(P_{\text{test}} \ll 1\); therefore, we neglect the possibility that more than one photon is scattered during the gate.

We assume that the distance between the ions is large compared to the wavelength of light. Therefore, scattering events of the two ions are not correlated.

For a small phase error, the error averaged over all possible input states is a factor of 2 lower than the worst-case input error.

For realistic parameters \(P_{\text{test}} \ll 1\); therefore, we neglect the possibility that more than one photon is scattered during the gate.

For a small phase error, the error averaged over all possible input states is a factor of 2 lower than the worst-case input error.

http://physics.nist.gov/PhysRefData/ASD/
Entanglement is a necessary resource for quantum applications—entanglement established between quantum systems at different locations enables private communication and quantum teleportation, and facilitates quantum information processing. Distributed entanglement is established by preparing an entangled pair of quantum particles in one location, and transporting one member of the pair to another location. However, decoherence during transport reduces the quality (fidelity) of the entanglement. A protocol to achieve entanglement 'purification' has been proposed to improve the fidelity after transport. This protocol uses separate quantum operations at each location and classical communication to distil high-fidelity entangled pairs from lower-fidelity pairs. Proof-of-principle experiments distilling entangled photon pairs have been carried out. However, these experiments obtained distilled pairs with a low probability of success and required destruction of the entangled pairs, rendering them unavailable for further processing. Here we report efficient and non-destructive entanglement purification with atomic quantum bits. Two noisy entangled pairs were created and distilled into one higher-fidelity pair available for further use. Success probabilities were above 35 per cent. The many applications of entanglement purification make it one of the most important techniques in quantum information processing.

Recent efforts to realize practical quantum information processing devices based on various physical systems have led to impressive new developments. Much-anticipated applications of quantum information processing devices are quantum communication and quantum computing. Large-scale implementations require distributing each quantum bit (qubit) of an entangled pair of qubits to separate locations. In quantum communication, entangled pairs are the fundamental resource for quantum teleportation, entanglement-based quantum cryptography and other protocols. They also underlie several promising schemes for quantum computing, which are needed for fast coupling of distant qubits and purification succeeds if the two measurement outcomes are different.

To restore fidelity lost during transport of entangled qubits, we can use entanglement purification to distil a smaller number of high-fidelity entangled qubits. The simplest instance of the entanglement purification protocol of ref. 4 distills one entangled pair of qubits from two imperfectly entangled ones. Here the qubits at one location are compared by means of local quantum operations. The results of the comparison are shared through classical communication between locations, and if they are consistent, a higher-fidelity pair of entangled qubits is obtained. Assuming no error in the comparison and sufficiently entangled initial pairs, the process can be iterated in multiple ‘rounds’ to obtain arbitrarily high-fidelity entangled pairs. Previous experiments, using photons, included demonstrations of entanglement purification and entanglement concentration, which required specific input states. However, in these experiments, entangled pairs were obtained by post-selection with a low success probability of $<10^{-3}$ per try. As two pairs are needed for an experiment, purification success probabilities were less than $10^{-6}$ per try. In addition, successful comparisons required destruction of entanglement, so the purified entangled pairs were not available for further use.

Here we report experiments that faithfully implemented the purification protocol proposed in ref. 4. An important feature of this protocol is that it works for all input states with sufficient fidelity with respect to the desired ideal entangled state. We tested the protocol on a family of input states that are approximately pure. Although other methods such as entanglement concentration can achieve better fidelity for our input states, this is at the cost of lower success probability, and our goal was to demonstrate a purification

---

**Experimental purification of two-atom entanglement**


Entanglement is a necessary resource for quantum applications—entanglement established between quantum systems at different locations enables private communication and quantum teleportation, and facilitates quantum information processing. Distributed entanglement is established by preparing an entangled pair of quantum particles in one location, and transporting one member of the pair to another location. However, decoherence during transport reduces the quality (fidelity) of the entanglement. A protocol to achieve entanglement 'purification' has been proposed to improve the fidelity after transport. This protocol uses separate quantum operations at each location and classical communication to distil high-fidelity entangled pairs from lower-fidelity pairs. Proof-of-principle experiments distilling entangled photon pairs have been carried out. However, these experiments obtained distilled pairs with a low probability of success and required destruction of the entangled pairs, rendering them unavailable for further processing. Here we report efficient and non-destructive entanglement purification with atomic quantum bits. Two noisy entangled pairs were created and distilled into one higher-fidelity pair available for further use. Success probabilities were above 35 per cent. The many applications of entanglement purification make it one of the most important techniques in quantum information processing.

Recent efforts to realize practical quantum information processing devices based on various physical systems have led to impressive new developments. Much-anticipated applications of quantum information processing devices are quantum communication and quantum computing. Large-scale implementations require distributing each quantum bit (qubit) of an entangled pair of qubits to separate locations. In quantum communication, entangled pairs are the fundamental resource for quantum teleportation, entanglement-based quantum cryptography and other protocols. They also underlie several promising schemes for quantum computing, which are needed for fast coupling of distant qubits and purification succeeds if the two measurement outcomes are different.

To restore fidelity lost during transport of entangled qubits, we can use entanglement purification to distil a smaller number of high-fidelity entangled qubits. The simplest instance of the entanglement purification protocol of ref. 4 distills one entangled pair of qubits from two imperfectly entangled ones. Here the qubits at one location are compared by means of local quantum operations. The results of the comparison are shared through classical communication between locations, and if they are consistent, a higher-fidelity pair of entangled qubits is obtained. Assuming no error in the comparison and sufficiently entangled initial pairs, the process can be iterated in multiple 'rounds' to obtain arbitrarily high-fidelity entangled pairs. Previous experiments, using photons, included demonstrations of entanglement purification and entanglement concentration, which required specific input states. However, in these experiments, entangled pairs were obtained by post-selection with a low success probability of $<10^{-3}$ per try. As two pairs are needed for an experiment, purification success probabilities were less than $10^{-6}$ per try. In addition, successful comparisons required destruction of entanglement, so the purified entangled pairs were not available for further use.

Here we report experiments that faithfully implemented the purification protocol proposed in ref. 4. An important feature of this protocol is that it works for all input states with sufficient fidelity with respect to the desired ideal entangled state. We tested the protocol on a family of input states that are approximately pure. Although other methods such as entanglement concentration can achieve better fidelity for our input states, this is at the cost of lower success probability, and our goal was to demonstrate a purification

---

**Figure 1 | Network diagrams for purification.** a. The original protocol. The qubits of both pairs are connected by CNOT gates across the pairs, after which the two qubits of pair 1 are measured. The measurement outcomes are compared by classical communication and purification succeeds if they are the same. b. Experimental implementation in this work. The ion qubits of both entangled pairs are (1) rotated by $R(n/2, n/4)$ (see Methods), (2) connected by two-ion $z = n/4$-phase gates across the pairs (marked with Z), (3) rotated by $R(n, 3n/4)$, (4) connected by two-ion $z = n/4$-phase gates again and (5) finally rotated by $R(n/2, n/4)$. The ions of pair 1 are then measured, and purification succeeds if the two measurement outcomes are different. c-f, Gates applied to ions and positions of the ions. The potential well zones are shown schematically under the ions. c. With all ions trapped in one potential well, entangled pair 1 (black) and pair 2 (grey) are created by a phase gate as described in Methods. d. Corresponding ions in each pair are connected by phase gates (as shown in b). e. The pairs are separated into zones $z2$ and $z3$, and $p1$ is measured by state-dependent fluorescence. f. After moving the ions, the fidelity of $p2$ is determined in $z2$ by state tomography.

---

1National Institute of Standards and Technology, Boulder, Colorado 80305, USA. Present address: University of Ulm, 89069 Ulm, Germany.
atomic ion qubits that were then distilled. The protocol succeeded with probabilities between 35% and 65%, depending on the initial fidelity. We demonstrated a gain in fidelity for a range of initial fidelities. In principle, the purified pairs are available for further rounds of the purification protocol or for use in other algorithms.

Our experimental procedure follows the original proposal of ref. 4 that is shown in Fig. 1a. We first confined four \( ^{40}\text{Ar}^+ \) ions in one trapping zone of a linear multi-zone Paul trap\(^{19} \) where they formed a linear array along the weakest axis of the trapping potential (Fig. 1c), which was approximately harmonic in three dimensions. Qubits were implemented with two hyperfine ground states of each ion, \( |F=1, m_F=1\rangle \) and \( |F=2, m_F=-2\rangle \), designated respectively \( |\uparrow\rangle \) and \( |\downarrow\rangle \) for simplicity. We implemented qubit rotations \( R(0, \phi) \) (see Methods) by driving stimulated Raman transitions with two laser fields\(^{15} \). Multi-qubit operations were realized with geometric phase gates generalized from the two-qubit phase gate described in ref. 16 (see Methods).

To prepare the initial state of the four-qubit register, all four modes of motion of the linear array were cooled to the ground state\(^{17} \), and the internal qubit states were optically pumped to \( |\downarrow\rangle \). We then generated two pairs of qubits with varying degrees of entanglement, as depicted schematically in Fig. 1c. A phase gate embedded between rotations on all ions entangled pair 1 (p1, consisting of ions 1 and 2) and pair 2 (p2, consisting of ions 3 and 4) separately at the same time, while all four ions were held in the same trap potential well. Ideally, after this operation the state of the two pairs is:

\[
|\Phi\rangle_{p1} \otimes |\Phi\rangle_{p2} = \left( \cos(\varepsilon)|\downarrow\downarrow\rangle_{12} + i\sin(\varepsilon)|\uparrow\uparrow\rangle_{12} \right) \\
\otimes \left( \cos(\varepsilon)|\downarrow\downarrow\rangle_{34} + i\sin(\varepsilon)|\uparrow\uparrow\rangle_{34} \right)
\]

(1)

For angle \( \varepsilon = \pi/4 \), the state is a product of two maximally entangled Bell pairs \( |\Phi_{\text{Bell}}\rangle = \frac{1}{\sqrt{2}} \left( |\downarrow\downarrow\rangle + i|\uparrow\uparrow\rangle \right)_j \) with \( j = \{p1, p2\} \). Experimentally, the states \( |\Phi\rangle \) were obtained with a fidelity of approximately 0.75 for \( \varepsilon = \pi/4 \), limited by imperfections in the phase gate operation. Ideally, for general \( \varepsilon \), the fidelity of each of the created pairs with respect to this Bell pair is:

\[
F_{\text{pair}} = \left| \langle \Phi | \Phi_{\text{Bell}} \rangle_{j} \right|^2 = \cos^2(\varepsilon - \pi/4)
\]

(2)

The states \( |\Phi\rangle \) can be viewed as the ideal Bell pairs (\( \varepsilon = \pi/4 \)) perturbed by a coherent combination of phase (sign flip) errors for \( \varepsilon \neq \pi/4 \). To lowest order in the deviation of \( \varepsilon \) from \( \pi/4 \), the errors act on just one of the qubits and are detectable by the purification process. By applying the purification protocol to this family of states, we are able to determine how well the experimental implementation performs on this type of phase error. To determine the performance on mixtures of these states, we use weighted combinations of the data for different \( \varepsilon \).

We implemented a slight variation of the ref. 4 proposal shown in Fig. 1b. We used one purification phase gate connecting each member of the first entangled pair to its counterpart in the second pair (Fig. 1d). The first two qubits (p1) were then measured to determine whether purification succeeded (Fig. 1e). Execution of the phase gate ideally results in the state:

\[
|\Psi\rangle = \frac{1}{2} \left( |\uparrow\downarrow\rangle_{12}|\uparrow\downarrow\rangle_{34} + |\downarrow\uparrow\rangle_{12}|\downarrow\uparrow\rangle_{34} \\
+ \cos(2\varepsilon)(|\downarrow\downarrow\rangle_{12}|\downarrow\downarrow\rangle_{34} + |\uparrow\uparrow\rangle_{12}|\uparrow\uparrow\rangle_{34}) \\
+ \sin(2\varepsilon)(|\uparrow\downarrow\rangle_{12}|\downarrow\uparrow\rangle_{34} + |\downarrow\uparrow\rangle_{12}|\uparrow\downarrow\rangle_{34}) \right)
\]

(3)

The quantum correlations between the entangled pairs are similar to those produced in the original proposal. However, measurement on the first two qubits must show they are either \( |\uparrow\rangle \) or \( |\downarrow\rangle \) for the purification to succeed, and the resulting entangled state of qubits 3 and 4 is then rotated such that when \( \varepsilon = \pi/4 \) the state is \( |\Psi_{+}\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rangle_{12} + |\downarrow\rangle_{12}) \). For general \( \varepsilon \), the probability of measuring that ions 1 and 2 are different is \( P_{\text{diff}} = \frac{1}{4} (1 + \sin^2(2\varepsilon)) \), in which case the state of the second pair of qubits is projected to:

\[
|\Psi_{\text{pair}}\rangle = \frac{1}{\sqrt{1 + \sin^2(2\varepsilon)}} \left[ \frac{1}{2} \left( |\downarrow\downarrow\rangle_{12} + |\uparrow\downarrow\rangle_{12} \right) \\
+ \sin(2\varepsilon)(|\uparrow\downarrow\rangle_{12}|\downarrow\uparrow\rangle_{34} + |\downarrow\uparrow\rangle_{12}|\uparrow\downarrow\rangle_{34}) \right]
\]

(4)

Instead of \( P_{\text{diff}} = \frac{1}{4} (1 + \sin^2(2\varepsilon)) \), the state of the second pair of qubits is projected to:

\[
F_{\text{pair}} = \left| \langle \Psi_{\text{pair}} | \Psi_{+} \rangle \right|^2 = \frac{4 \cos^2(\varepsilon - \pi/4)}{3 + \cos(4\varepsilon - \pi/4)} \geq F_{\text{pair}} \quad (0 \leq \varepsilon \leq \pi/2)
\]

(5)

Near \( \varepsilon = \pi/4 \), the fidelity of the unpurified pairs decreases quadratically in \( \varepsilon - \pi/4 \), whereas the fidelity of the purified pair decreases quadratically for the measurements and determination of the fidelity of the purified pair, we separated groups of ions in the multi-zone architecture of the trap\(^{15} \). We chose an order of the ions that made the measurement on p1 and the subsequent tomography on p2 as simple as possible. If we were distilling remote entanglement, then both pairs p1 and p2 would have one qubit in each location. Results of the
All ions were transferred into a separation zone and separated into the two initial pairs within 380 μs. Pair p1 returned to the original interaction zone, where the state of the two ions was detected by state-dependent fluorescence (Fig. 1c). We used a detection period of 200 μs, during which we detected about 0.5 photons if both ions were in state |j⟩, and about 10 photons on average for each ion in state |g⟩. We declared the purification a success when the number of counts was in an interval determined by the property that the probability of falsely declaring success with both ions in state |j⟩ was less than 0.01 (0.05). We then moved pair p1 out of the interaction zone and the purified pair p2 into it (Fig. 1f) in about 470 μs. The state of p2 was analysed by a tomography procedure that determined all four Bell-state populations without distinguishing between the ions. This yielded an experimental value of the fidelity $F_{p2}$ (see Methods). The complete experimental sequence, including tomography, took about 1.54 ms.

For a comparison, each attempt at purification was followed by a reference experiment in which the entangled pairs were prepared in the same way but the purification phase gate was omitted. Instead we applied a common rotation $R(\pi/2, -\pi/4)$ to both pairs, transforming $\{|\psi\rangle, \Omega\rangle\}$ into $\{|\psi\rangle, \psi\rangle\}$. The fidelity $F_{p2}$ of pair p2 was then determined by the same tomography procedure. The fidelity of p1 was characterized in separate experiments and found to be equal to that of p2 within experimental uncertainties. Therefore $F_{p2}$ is the fidelity of a prepared Bell pair if purification was not used. The same noise processes act on the ions in the purification and the reference experiment after the time in the procedure where the purification gate would be applied. To determine the effectiveness of purification, $F_{p2}$ was compared to $F_{p2}$ (Fig. 2). The final fidelity of the purified pair reached a maximum of about 0.75. Nevertheless, we demonstrated an improvement $F_{p2} - F_{p2}$ for 0.5 $F_{p2} \approx 0.7$. Ideally, neither fidelity should drop below 0.5 for the family of states used here. The apparent improvements discernible in the figure for $F_{p2} < 0.5$ are explained by the common sources of errors equally affecting the purification and reference experiments after the time when the purification gate is implemented (or not). Depending on the initial fidelity, the probability of success varied between 35% and 65%.

The data shown in Fig. 2 are for the family of states given in equation (1) perturbed by experimental noise. From this data, one can infer the performance of the purification protocol on mixtures of such states. We chose a mixture that approximates a uniform distribution of angles $\epsilon$ in equation (1) between $\epsilon = 0$ and $\epsilon = \pi/4$. The unpurified fidelity for this mixture was $0.614 \pm 0.0015$, and the purified fidelity was $0.629 \pm 0.0015$. This implies a statistically significant improvement in fidelity of 0.015 $\pm 0.002$ for this particular mixture.

As the purification protocol took place with the two pairs in the same trapping zone and with operations simultaneously acting on all ions, it was necessary to verify that the implemented operations did not introduce unintended correlations. We confirmed experimentally that, subject to the limitations of the tomographic analysis, state preparation resulted in independent pairs of entangled ions and, if the purification gate was applied directly to the state $|\Psi\rangle = |\psi\rangle, |\psi\rangle$, the state of ions 1 and 3 was independent of the state of ions 2 and 4, within experimental error.

Thresholds of tolerable error rates for entanglement purification in quantum repeaters are of the order of a few per cent, much less than the imperfections present in our current experiment. To improve state preparation and purification in future implementations, better control of classical parameters such as magnetic field and laser intensity will be required. Decoherence due to spontaneous emission could be reduced by an appropriate choice of the laser-beam detuning. More advanced multi-zone traps and sympathetic cooling will make it possible to implement purification with the ions of each entangled pair transported to separated trap zones. With these improvements, purified Bell pairs with sufficient fidelity to violate Bell inequalities should be feasible. Because the two pairs of entangled states in equation (1) have identical angles $\epsilon$, our procedure tests the behaviour of the purification protocol in the case of collective phase errors. To determine the behaviour for other phase errors, one could use individual laser addressing of the ion or magnetic field gradients. In summary, we have demonstrated entanglement purification with relatively high success rates in a potentially scalable system. The protocol and success rates demonstrated, together with the availability of the purified pair, could enable ‘entanglement pumping’ by repetitive application of the purification protocol. Ideally, the fidelity of the remaining pair(s) can be ‘pumped’ arbitrarily close to 1, but in practice will never exceed a limit imposed by imperfections in the purification gates and the measurements used to produce the purified pairs. In addition to uses in quantum communication and large-scale quantum information processing, remote entangled atoms could be useful in more fundamental experiments, such as a loophole-free test of local hidden-variable theories. The multi-segmented trap architecture used here should allow the distribution of entangled particles to separate locations for exploring repetitive protocols in future experiments.

**METHODS**

**Gate operations.** A general single-qubit rotation $R(0, \phi)$ transforms the qubit states as $R(0, \phi)|j\rangle = \cos(\phi/2)|j\rangle - i \sin(\phi/2)|g\rangle$, and $R(0, \phi)|g\rangle = -i \sin(\phi/2)|j\rangle + \cos(\phi/2)|g\rangle$. These rotations were applied uniformly to all ions that resided in the trap zone z2 addressed by the Raman laser beams (Fig. 1c–f). Phase gates for entangling different combinations of pairs of ions in a string of four ions are a generalization of the phase gate described in ref. 16. For the arrangement of laser beams in the experiment, the phase of the dipole force repeated every 213 nm along the alignment direction of the ions. The equilibrium between mutual Coulomb repulsion and the confinement of the external trap potential determined the positions of the ions relative to the trap centre to be $s = -1.437, -0.454, 0.454, 1.437$, with $s = \sqrt{1/2}(\Delta_{sm\text{out}}/\Delta_{sm\text{in}})$, $\Delta_{sm\text{in}}$ the elementary charge, $\Delta_{sm\text{out}}$ the vacuum permittivity, $m$ the mass of the beryllium ion and $\Delta_{sm\text{out}}/\Delta_{sm\text{in}}$ the axial centre of mass (COM) frequency. For Be+, $s = 7.31 \mu$m ($\Delta_{s\text{out}} = 200 \mu$s). By choosing the strength of the external potential appropriately, we achieved a pattern of dipole forces that equally coupled ion 1 to 2, and ion 3 to 4, with negligible coupling of other ion pairs. Within a coupled pair, phases changed according to $|1\rangle, |1\rangle, |1\rangle, |1\rangle$ to $|1\rangle, |1\rangle, |1\rangle, |1\rangle$, where $s$ was determined by the time of the operation and the beam intensities.

To entangle two pairs of ions, we first applied a global rotation $R(\pi/2, 0)$ to all four ions, followed by the phase gate with $0 \leq s \leq \pi/4$ on the axial mode at $\omega = 2\pi \times 6.742$ MHz $\approx 2.41\Delta_{s\text{out}}$, which has normalized mode amplitudes $(\frac{1}{\sqrt{2}} - \frac{i}{\sqrt{2}})$ and $\frac{1}{\sqrt{2}}$, respectively. We then applied a refocussing $R(\pi, \pi)$ rotation, the same phase gate, and finally another $R(\pi/2, 0)$ rotation. Pair preparation took less than 100 μs. Calculations show that residual phases due to unwanted couplings only degrade the pair fidelity by 0.007. The observed loss in fidelity can be attributed to the following causes: during interaction, preparation pair $F$ will lose a maximum of 7% of probability of one of the ions undergoing an absorption–spontaneous emission cycle. Phase instabilities of the single qubit rotations caused by interferometric phase fluctuations between Raman beams contributed of the order of 10% over the 1.5-ms duration of the experiment. Further imperfections included laser beam pointing and intensity noise, leading to fluctuations in the ion–laser couplings (Rabi frequencies) of the order of 5%. We also estimated an error of 5% due to fluctuations in the trapping potential, caused by noisy potentials applied to the trap electrodes. The last three errors also included systematic drifts that contributed to the statistically significant scatter in the measured fidelities (Fig. 2).

The purification gate was similar to the first gate. Two phase gates $Z$ with angle $\pi/4$ were embedded into single qubit rotations in a sequence $R(\pi/2, \pi/4) - Z - R(\pi, \pi/4) - Z - R(\pi/2, \pi/4)$ applied to all ions. The phase gates $Z$ coupled ion 1 to 3 and 2 to 4 (see Fig. 1c) and were executed on the COM mode at a frequency of 4.07 MHz (duration 65 μs). Calculations show that unwanted couplings in this gate degrade the fidelity by 0.007. The actual loss of fidelity due to the purification procedure was similar to those discussed above. Additional loss of fidelity in the purified pair was caused by imperfect state discrimination of pair 1 (see main text). The photon count thresholds for inferring that exactly one ion was in state $|j\rangle$ were set so that the probability of error is at most 5%.

**State tomography.** During detection (duration 200 μs) we recorded between 0.15 $\leq N_{\text{avg}} \leq 0.8$ counts if all ions are projected into $|j\rangle$, and between $8 \leq N_{\text{avg}} \leq 12$ additional average counts for each ion in state $|g\rangle$. Count averages $N_{\text{avg}}$ and $L_{\text{avg}}$ were derived by fitting mixtures of poissonian distributions to count histograms for
the relevant detection periods and to reference histograms obtained by preparing all ions to be observed in the $|1\rangle$ or $|\bar{1}\rangle$ states. We used the maximum likelihood method for fitting the histograms, and parametric-bootstrap resampling for determining standard errors in inferred quantities. To obtain the fidelity of the second Bell pair (ions 3 and 4), we applied the tomographic rotation $T(\theta_a, \phi_b) = R(\pi/2, \phi_a) R(\pi/2, \phi_b)$ to both ions simultaneously, with $\phi_a = 0$, $\pi/4$, $\pi/2$ and $\phi_b = 0$, $\pi/4$, $\ldots$, $7\pi/4$. From the photon counts obtained after the tomographic rotations, we inferred the density matrix in the Bell basis before the rotations by maximum likelihood methods. The inferred entries included the Bell states populations but not the coherences between the singlet and triplet states. We used a version of the maximum likelihood algorithm described in ref. 22. In the reference experiment, we extended the tomographic inference by including the counts for the measurement of ions 1 and 2, which yielded the symmetrized density matrices on ions 3 and 4 conditional on the number of ions 1 and 2 in state $|\bar{1}\rangle$. We used this information to check that the state in the reference experiment was consistent with the first and second pair being independently entangled.

In separate experiments, we checked that the behaviour of the purification gate was consistent with correlating only ions 1 and 3, and (separately) ions 2 and 4. In both cases, we found the inferred density matrix to be consistent with the independence assumption, in the sense that they match the terms of an independent density matrix with probability of error less than 0.01.

Received 19 June; accepted 3 August 2006.

Acknowledgements This work was supported by the Disruptive Technology Office (DTO), by a DoD Multidisciplinary University Research Initiative (MURI) programme administered by the Office of Naval Research and by NIST. R.R. was supported by the Alexander von Humboldt Foundation. We thank S. Glancy and W. Itano for comments on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to D.L. (dl@boulder.nist.gov).
Transport dynamics of single ions in segmented microstructured Paul trap arrays

R. Reichle¹,²,*, D. Leibfried², R. B. Blakestad², J. Britton¹, J. D. Jost², E. Knill², C. Langer², R. Ozeri¹, S. Seidelin², and D. J. Wineland²

¹ University of Ulm, Albert-Einstein-Allee 11, 89069 Ulm, Germany
² National Institute for Standards and Technology, 325 Broadway, Boulder/CO 80305, USA

Received
Published online 4 August 2006

Key words Transport, quantum information processing, ion traps, segmented Paul trap
PACS 32.80.Pj, 03.67.Lx

It was recently proposed to use small groups of trapped ions as qubit carriers in miniaturized electrode arrays that comprise a large number of individual trapping zones, between which ions could be moved [1, 2]. This approach might be scalable for quantum information processing with a large numbers of qubits. Processing of quantum information is achieved by transporting ions to and from separate memory and qubit manipulation zones in between quantum logic operations. The transport of ion groups in this scheme plays a major role and requires precise experimental control and fast transport times. In this paper we introduce a theoretical framework to study ion transport in external potentials that might be created by typical miniaturized Paul trap electrode arrays. In particular we discuss the relationship between classical and quantum descriptions of the transport and study the energy transfer to the oscillatory motion during near-adiabatic transport. Based on our findings we suggest a numerical method to find electrode potentials as a function of time to optimize the local potential an ion experiences during transport. We demonstrate this method for one specific electrode geometry that should closely represent the situation encountered in realistic trap arrays.

© 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

1 Introduction

Quantum information processing is a rapidly evolving field of physical science. Its practical importance arises from the exponential speedup in computation of certain algorithmic tasks over classical computation [3]. Building an actual device that can process quantum information, however, is technologically difficult due to the need for qubits that can be processed and read out with high fidelities and the extreme sensitivity of the quantum mechanical states stored in these units against external uncontrolled perturbations. A promising technical approach as shown over the last decade, is to use strings of ions as physical qubits confined in linear electromagnetic Paul traps [1,4]. These strings are stored in a single trap and constitute a one dimensional crystallized structure whose vibrational modes can be laser cooled to their ground states. The strong mutual coupling of the ions by Coulomb forces in such a crystal has been proposed and utilized to create arbitrary superpositions of quantum states of the ionic internal states (1–4). In the last few years methods were developed that enable quantum state engineering with high precision and long coherence times [7–11]. The necessary criteria [12] for large-scale quantum computation have been demonstrated in the past years, and small algorithms have been implemented successfully [13–17]. However, as in other approaches aiming towards quantum computation, scaling to many qubits is challenging. Considerable overhead is required by quantum error correcting schemes that permit robust quantum computation and make large-scale implementations feasible. To scale up a linear string of many ions, a rapidly growing

* Corresponding author E-mail: rainer.reichle@uni-ulm.de

© 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
number of vibrational degrees of freedom needs to be controlled and cooled to the ground state for reliable processing. This is extremely difficult to realize. A more recent proposal [1, 2] has been made to circumvent this problem by using small arrays of a few qubits that are shuttled around in two-dimensional microstructures to process and store quantum states at various locations.

An initial systematic study showed that coherent transport of ions in linear trap arrays is possible with nearly no loss in contrast during the motion [18]. In this experiment an adiabatic transport of a qubit was performed over a distance of 1.2 \( \text{mm} \) in a time span of about 54 \( \mu \text{s} \) with negligible heating. Currently, there are strong efforts under way to demonstrate the possibility of building large-scale ion trap structures. For example, suggestions have been made to combine miniaturized ion chips directly with CMOS electronics to handle the resources required to control the many electric potentials [19]. Moreover, fast transport requires excellent experimental control of all these potentials.

A detailed scheme of how a viable architecture of an ion trap processor could look has been recently studied by Steane [20], fully incorporating quantum error correcting codes. The physical gate rate of this proposed 300 qubit processor unit was found to be limited by

\[
\tau_g = \frac{2}{\nu_{\text{COM}}} + \frac{10}{\nu_r} + \tau_{\text{cool}} + \tau_p \approx \text{a few \( \mu \text{s} \)},
\]

with the first two terms being an average time of the part of a typical gate that involve motion that is times for splitting \( \sim 2\nu_{\text{COM}}^{-1} \), recombining and moving \( \sim 10\nu_r^{-1} \) a small ion string, where \( \nu_{\text{COM}} \) and \( \nu_r \) are typical axial and radial trapping frequencies, respectively. The last two terms correspond to cooling after the transport has been done, and the time duration of conducting the actual phase gate, respectively. On the other hand, if large amounts of energy are transferred to the ions, longer cooling times might be needed. Inserting typical operating conditions shows that the first two terms make up a considerable part of the performance of the physical gate rate. In order to keep this part as small as possible we need designs for electrode structures enabling fast qubit transport.

In the following we present a theoretical framework that governs the transport dynamics of ions trapped in a time varying external potential. In Sect. 2 the equations of motion for the transport are derived. Sect. 3 discusses the general classical solution in terms of an Ermakov parametrization. This approach is useful to express the quantum approach presented in Sect. 5, which uses the Heisenberg picture following the approach of Kim et al. [21]. In Sect. 4 we point out some well-known properties of a quantum harmonic oscillator exposed to a transporting force for the simpler case when its frequency is kept constant. Sect. 5 presents the general quantum solutions and the interrelation between classical and quantum transport. Based on this framework we discuss in Sect. 6 a well-controlled regime for the transport and also include first order perturbations to the transport dynamics. In Sect. 7, we present numerical optimization routines to extract optimum switching of potentials for the transport and study miniaturization of electrode structures to estimate the required resources for a well-controlled transport. Finally, a simple electrode model is used to find a practical rule for the segmentation of ion traps revealing insight into the resources needed for large-scale layouts, that should be also applicable for more general trap arrays.

## 2 Classical equations of motion

A linear segmented Paul trap, e.g. as used in recent experiments [5, 14–16, 22, 24], consists typically of two alumina wafers with gold coated electrode surfaces of a few micrometer thickness. The slotted wafers provide electrical RF and DC fields for 3D confinement of ions. The arrangement for control electrodes is schematically sketched in Fig. 3 where only a single layer is shown. The confinement along the \( x \)-axis is achieved solely by electrostatic fields whereas the remaining two orthogonal radial directions correspond to a dynamical trapping by ponderomotive RF forces. In this article we limit ourselves to transport along a single dimension \( x \) from \( -b/2 \) to \( b/2 \). If we denote the coordinate of the ion in the laboratory frame by \( q \)
then we have from Newton’s equation of motion
\[ \ddot{q}(t) + \frac{Q}{m} \frac{\partial \phi(q,t)}{\partial q} = 0, \quad q(-t_0) = -\frac{b}{2}, \quad \dot{q}(-t_0) = 0, \quad (2) \]
with two initial conditions as the equations on the rhs; \( Q \) is the elementary charge and \( m \) the mass of the transported ion. We assume a time interval and location of the ion starting at \( -t_0 \) and \(-b/2\), and ending at \(+t_0\) and \(b/2\), respectively. In order to make use of coherent states of a harmonic oscillator (that do not spread in time) we are interested in designing the time-dependent electrical potential as
\[ \phi(q,t) = \phi_{\text{local}}(q - q_0(t)) + \varphi_0(t), \quad (3) \]
where \( \phi_{\text{local}}(q) \sim m\omega_0^2 q^2 / 2Q \) is purely quadratic with constant curvature in a sufficiently large range around the minimum, and \( \varphi_0(t) \) is a time-dependent offset with no influence on the dynamics. Here, we prescribe the dynamics by specifying a desired harmonic frequency \( \omega_0 \) and the temporal shift of the harmonic well by a transport function \( q_0(t) \). The residual, uncontrolled force caused by insufficient flexibility in creating the desired harmonic potential deteriorates the transport performance. Its effect can be described by the difference potential or residual acceleration, i.e.
\[ \phi_{\text{res}}(q,t) = \phi(q,t) - \varphi_0(t) = -\frac{m\omega_0^2(q - q_0(t))^2}{2Q} \quad \text{and} \quad a_{\text{res}}(q) = -\frac{Q}{m} \frac{\partial \phi_{\text{res}}(q)}{\partial q}, \quad (4) \]
respectively. Due to imperfect realization of the harmonic well \( \phi_{\text{res}}(q,t) \) adds fluctuating parts to the ideal harmonic potential as a function of position or time, critically depending on the electrode structure used. In Sect. 7 we will discuss a numerical scheme for approximating \( \phi(q,t) \) based on superpositions of individual electrode potentials in an optimal way.

We finally can write down the classical equation of motion
\[ \ddot{u} + \omega_0^2 u = -\dot{q}_0(t) + a_{\text{res}}(u + q_0(t)) \quad (5) \]
which we transformed into a frame moving with \( q_0(t) \) by \( u = q - q_0 \). The net acceleration on the rhs corresponds to an external force and displaces the ion from its equilibrium position \( u = 0 \) in this frame. Since we will treat only the first two perturbation terms we expand the final equation of motion around the minimum of the well and rearrange some terms to get
\[ \ddot{u} + \omega_0^2 \left( 1 - a'_{\text{res}}[q_0(t)] / \omega_0^2 \right) u - a''_{\text{res}}[q_0(t)] / 2 u^2 + \ldots = -\dot{q}_0(t) + a_{\text{res}}[q_0(t)] \quad (6) \]
where primes denote differentiation with respect to \( u \). For the following discussion we abbreviate \( \omega^2(t) = \omega_0^2 \left( 1 - a'_{\text{res}}[q_0(t)] / \omega_0^2 \right) \) and write \( f(t) = -\dot{q}_0(t) + a_{\text{res}}[q_0(t)] \) for the rhs of Eq. (6). For certain electrode structures, we can disregard terms involving the second and higher order derivatives of \( a_{\text{res}}(q) \) (cf. Sect. 7). We will make this assumption throughout the paper. In that case Eq. (6) simplifies to the equation of motion of a parametrically driven and forced harmonic oscillator with the Hamiltonian
\[ H(t) = \frac{p^2}{2m} + \frac{m\omega^2(t)}{2} u^2 - m f(t) u \quad (7) \]
and \( p = m\dot{u} \).

3 Classical dynamics of ion transport

To obtain a general classical solution with an arbitrary frequency modulation we first consider the formalism which is most often used in conjunction with time-dependent invariants within so called Lewis-Riesenfeld
methods [30]. These approaches have been shown to be successful in the quantization of time-dependent harmonic oscillators with many different kinds of time-dependencies. Here, we discuss the general classical solution using the Ermakov equation and its generalized phase equation for time-dependent frequencies. We then employ in Sect. 5 the approach of Kim et al. [21] to express the general quantum solution in terms of its classical solution.

3.1 Homogeneous solution
Neglecting higher order terms we find the homogeneous part of the solution of Eq. (6) by setting \( f(t) = 0 \), thus solving

\[
\ddot{u}_c + \omega^2(t) u_c = 0
\]

for an arbitrary time-dependent frequency \( \omega(t) \). For this, it is most convenient to make the ansatz

\[
u_1 = \rho(t)e^{i\mu(t)} \quad \nu_2 = \rho(t)e^{-i\mu(t)},
\]

introducing an amplitude function \( \rho(t) \) and a phase function \( \mu(t) \), both real. Inserting Eq. (9) into Eq. (8) and considering real and imaginary parts results in the two equations

\[
\ddot{\rho} - \rho \dot{\mu}^2 + \omega^2(t) \rho = 0, \quad 2\dot{\rho}\dot{\mu} + \rho \ddot{\mu} = 0.
\]

\( \rho \) is an integrating factor for the second equation on the right so that we can write

\[
\rho^2 \ddot{\mu} = 1,
\]

where we have chosen the integration constant as 1. The constant on the rhs of Eq. (11) has the SI units \( \text{m}^2 \text{rad/s} \) that should be taken into account at the end. If we substitute this back into the first equation of Eq. (10) we obtain the Ermakov equation for the amplitude function \( \rho(t) \)

\[
\ddot{\rho} + \omega^2(t) \rho = 1/\rho^3.
\]

For periods of constant frequency \( \omega = \omega_0 \) the general solution is \(^1\)

\[
\rho(t) = \pm \omega_0^{-1/2} \sqrt{\cosh \delta + \sinh \delta \sin(2\omega_0 t + \theta)},
\]

where \( \delta, \theta \) are constants of integration, their values depend on the past evolution [30]. The solution for the generalized phase is easily obtained once \( \rho \) is known. From Eq. (11) we have

\[
\mu(t) = \int_{-t_0}^{t} dt' \rho(t')^{-2}.
\]

The general homogeneous solution is then given by

\[
u_h(t) = \alpha_c \rho(t) \cos(\mu(t) + \varphi),
\]

with the classical amplitude \( \alpha_c \) and initial phase \( \varphi \) fixed by the initial conditions.

\(^1\) The general solution of this equation is easily obtained by first using \( \dot{\rho} \) as an integrating factor with the integration constant \( 2\omega \cosh \delta \). This equation is immediately transformed to a harmonic oscillator by \( x = \rho^2 - \cosh \delta/\omega \).
3.2 Green’s function and general solution

We use the general framework of Green’s functions to define a particular solution to the inhomogeneous case of Eq. (6), where we again terminate the expansion, i.e. $a_{res}^{(n)}[\eta] = 0$ for $n \geq 2$, to stay in a harmonic regime. Using the two independent homogeneous solutions of Eq. (9) we can determine the causal Green’s function

$$G(t, t') = \theta(t - t')\rho(t)\rho(t')\sin(\mu(t) - \mu(t')),$$

with $\theta(t - t')$ the Heaviside function. Employing $G(t, t')$, a particular solution is given by

$$u_p(t) = \int_{-\infty}^{t} dt' G(t, t')f(t') = \rho(t)\int_{-\infty}^{t} dt' \sin(\mu(t) - \mu(t'))\rho(t')f(t').$$

For later convenience we define the auxiliary function $\zeta(t) = i e^{-i\mu(t)} \int dt' e^{i\mu(t')} \rho(t')f(t')$ which will be useful for expressing the general quantum solution. In this notation we can abbreviate the particular solution by $u_p(t) = \rho(t)\{\zeta(t) + \zeta^*(t)\}/2$.

Thus, we obtained the general solution as the sum of the general homogeneous solution Eq. (15) and a particular solution

$$u_c(t) = u_h(t) + u_p(t) \equiv \rho(t)/2 \left\{ a_c e^{i(\mu(t) + \varphi)} + \zeta(t) \right\} + c.c.$$  

Higher derivatives, like velocity and acceleration, can easily be found from the general solution in Eq. (18) by using the Leibniz rule. For initial conditions where we start in the classical ground state $a_c = 0$ we define the quantity $\Xi(t) = u_p(t) + i\omega(t)u_p(t)$, assuming that the transport starts at $-t_0 > -\infty$, i.e. later than the infinite past, and demand that it takes a finite amount of time. With the help of the last definition the energy transferred to the oscillator at instants $t_1$ (where $f(t_1) \equiv 0$) is then given by

$$W(t_1) = m|\Xi(t_1)|^2/2,$$

with

$$\Xi(t_1) = \int_{-t_0}^{t_1} dt'\rho(t')\left\{ \rho(t_1)\dot{\mu}(t_1)\cos(\Delta\mu_1 t') + \left\{ \dot{\rho}(t_1) + i\omega(t_1)\rho(t_1) \right\}\sin(\Delta\mu_1 t') \right\}f(t'),$$

and $\Delta\mu_1 t' = \mu(t_1) - \mu(t')$. We will call $\Xi(t)$ the adiabatic suppression amplitude and its absolute square the adiabatic suppression factor. Thus, we have derived the classical energy transfer for arbitrary frequency evolutions and arbitrary external transport forces. To evaluate this expression one first must solve for the explicit time-dependence of $\rho$ and $\mu$ according to Eqs. (12), (14) by integrating the Ermakov equation and the phase equation, and finally compute the transferred energy at different times using Eqs. (19), (20).

3.3 Adiabatic limit

Since we are mainly interested in an adiabatic solution we can simplify the last expression by considering adiabatic expansions of the homogeneous solution for a parametrically driven harmonic oscillator. We introduce an adiabatic time scale $\mathcal{T}$ such that

$$\dot{\omega}/\omega = \mathcal{T}^{-1} \quad \text{for} \quad \mathcal{T} \gg \omega^{-1}.$$  

The general adiabatic expansion of the differential equations Eqs. (12), (14) is readily obtained [29]
\[
\rho(t) = \frac{1}{\sqrt{\omega(t)}} + \frac{1}{8} \frac{\dot{\omega}(t)}{\omega(t)^{7/2}} - \frac{3}{16} \frac{\dot{\omega}(t)^2}{\omega(t)^{9/2}} + \cdots 
\]
and
\[
\dot{\rho}(t) = \omega(t) - \frac{1}{4} \frac{\ddot{\omega}(t)}{\omega(t)^2} + \frac{3}{8} \frac{\dot{\omega}(t)^2}{\omega(t)^3} + \cdots 
\]

This procedure is equivalent to a perturbative approach on the first term in Eq. (12) [30]. We require that at instants \(-t_0, t_1\), i.e., at times when we measure the oscillator’s energy, the frequency has settled into a constant. Also for the following discussion we define that the oscillator’s initial frequency at \(-t_0\) is \(\omega(-t_0) = \omega_0\), so that \(\rho(-t_0) = 1/\sqrt{\omega_0}\), \(\dot{\rho}(-t_0) = 0\). Taking into account only the lowest order of the expansion in Eq. (22) the expression in Eq. (20) reduces to

\[
\Xi(t_1) = \sqrt{\omega(t_1)} \int_{-t_o}^{t_1} dt' f(t') \omega(t')^{-1/2} e^{i \Delta \mu_{1\nu}} \quad \text{with} \quad \Delta \mu_{1\nu} = \int_{t'}^{t_1} dt \omega(\tau) 
\]

providing the adiabatic energy transfer in the first order of frequency modulation.

4 Quantum and classical, dragged harmonic oscillators with constant frequency

Husimi [27] and Kerner [28] independently considered the forced quantum mechanical oscillator and found exact analytical expressions for their wavefunctions and propagators. We review some of their early ideas because they provide insight into the close relationship of the quantum and classical solution. In this paragraph we assume the frequency is independent of time. The corresponding Hamiltonian is given by Eq. (7) with \(\omega(t) = \omega_0\).

Following Husimi and Kerner, we can “uncouple” the classical oscillation by the transformation

\[
\Psi(u, t) = \phi(u', t) \exp(i m u_c u'/\hbar), \quad (24)
\]

with \(u' = u - u_c\) and \(u_c\) at first undefined. Inserting Eq. (24) into the time-dependent Schrödinger equation gives

\[
i \hbar \frac{\partial \phi}{\partial t} = \left( -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial u'^2} + \frac{1}{2} m \omega_0^2 u'^2 \right) \phi + m(\ddot{u}_c + \omega_0^2 u_c - f) u' \phi - (m/2)(\dot{u}_c^2 - \omega_0^2 u_c^2 + 2 f u_c) \phi. \quad (25)
\]

On the rhs we see that we can make the second term vanishing if we choose \(u_c\) to satisfy

\[
\ddot{u}_c + \omega_0^2 u_c = f = 0,
\]

i.e. if \(u_c\) satisfies the classical solution of Eq. (7). With this choice one can easily identify the classical action

\[
L(t) = (m/2)(\dot{u}_c^2 - \omega_0^2 u_c^2 + 2 f u_c)
\]

of a forced harmonic oscillator in the third term on the rhs of Eq. (25).

Furthermore, if we make the ansatz

\[
\phi(u', t) = \chi(u', t) \exp \left[ i \frac{\hbar}{\hbar} \int_{-\infty}^{t} dt' L(t') \right],
\]

we can absorb this term as a time-dependent phase into \(\phi\). The remaining part of the wavefunction, \(\chi\), then needs only to obey the usual harmonic oscillator wave equation in the frame defined by the classical trajectory with its internal coordinate \(u'\)

\[
i \hbar \frac{\partial \chi}{\partial t} = \left( -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial u'^2} + \frac{1}{2} m \omega_0^2 u'^2 \right) \chi. \quad (26)
\]
In this way one can achieve a separation of the forced harmonic oscillator from the unforced oscillator in a frame moving with the classical trajectory. The wavepacket does not become deformed by the homogeneously acting force. The quantum solution becomes displaced and only a phase is accumulated.

To determine further properties we can assume now a stationary state with energy $\epsilon_n$ for the solution of Eq. (26)

$$\chi_n(u', t) = u_n(u') \exp(-i\epsilon_n t) \quad \epsilon_n = (n + \frac{1}{2})\hbar\omega_0,$$

and evaluate transition probabilities at time $t$ for the oscillator to be in the number state $u_m$ if it was initially in the number state $u_n$

$$P_{mn}(t) = \left| \int_{-\infty}^{\infty} u_m(u - u_c(t))u_n(u)e^{im\omega_c(t)u/\hbar \, du} \right|^2.$$  

Husimi and Kerner showed that these transition moments can be evaluated analytically

$$P_{mn}(t) = (\mu/\nu)^2 \gamma^{\nu - \mu} e^{-\gamma} (2\nu - \mu)(\gamma))^2 \quad \text{with} \quad \gamma(t) = m/2\hbar\omega_0 |\dot{u}_c + i\omega u_c|^2$$  

by using generating functions for the Hermite polynomials [27,28]. In Eq. (27), $\nu$ is the greater while $\mu$ is the lesser of $m$ and $n$, respectively. $L^{\nu - \mu}_m$ denote the associated Laguerre polynomials, and its time-dependent argument $\gamma(t)$ describes the classical energy transfer in units of $\hbar \omega_0$. From Eq. (27) we see the classical character of the quantum solution: the transition probabilities are solely defined by the classical quantity $\gamma(t)$. Also, if we consider starting from the ground state $n = \mu = 0$ and using $L^0_0(\gamma) \equiv 1$ the probability distribution $P_{00}(t)$ becomes a poissonian, and thus we find the signature of a coherent state.

With this relation the expectation values for the mean energy and the dispersion of the energy distribution are then immediately obtained

$$\langle E_m \rangle_n \equiv \hbar\omega_0 \left( \sum_m m \, P_{mn} + 1/2 \right) = \hbar\omega_0 \left( n + \gamma + 1/2 \right) = \epsilon_n + \hbar\omega_0\gamma$$  

$$\langle (\Delta E_m)^2 \rangle_n \equiv \langle \hbar\omega_0 \rangle^2 \langle (m - \langle m \rangle)^2 \rangle = (\hbar\omega_0)^2(2n + 1)\gamma = 2\epsilon_n \hbar \omega_0 \gamma,$$

where $\epsilon_n$ is the initial energy before the force acts on the wavepacket. This is indicated in Eq. (28) by the subscripts on the lhs. Corresponding expressions for the classical solution

$$\langle E \rangle_{E_0} = E_0 + W \quad \langle (\Delta E)^2 \rangle_{E_0} = 2E_0W$$

are found if we average over the initial classical phase that are completely analogous to the quantum solutions.\(^2\) $E_0$ is the classical energy before the transport and $W \equiv \hbar\omega_0\gamma(t)$ the classical energy transfer. The mean energy and the energy spread increase linearly with the energy transfer in both solutions although the energy distributions of the classical and quantum solution are quite different [27]. Also, the zero point energy makes a difference between the classical and quantum description. If the system is initialized in its quantum ground state, transport can create a dispersion of the wavepacket due to $\epsilon_0 > 0$, while this is not the case for the classical ground state, i.e. if $E_0 = 0$.

5 The dragged quantum harmonic oscillator

Many methods have been developed to find exact quantum states of time-dependent oscillators. The generalized invariant method by Lewis and Riesenfeld [30] has been very successful in finding exact quantum

\(^2\) This result is easily derived by averaging the general classical solution in Eq. (18) given in Sect. 4 for a constant frequency over the phase interval $[0, 2\pi]$. 

www.fp-journal.org
motion in terms of wavefunctions and propagators. For the interpretation of time-dependent quantum systems and for showing its relationships to their classical solution, however, the Heisenberg picture is more appropriate since the Heisenberg operators for position and momentum obey similar equations of motion than the corresponding classical quantities. In this paragraph we aim to interpret the quantum solution using its classical analogue and therefore use the general approach of Kim et al. [21] that is based on the general invariant theory but acts in a Heisenberg picture, in contrast to the original approach.

The general invariant theory starts out by defining an invariant operator \( I(t) \) that satisfies the Heisenberg equation of motion. Ji et al. [23] used a Lie algebra approach to find the most general form of the solution with some integration constants \( c_i, \ i = 1, 2, 3 \), arbitrary defining the initial conditions (see discussion at the end of this paragraph). If we fix these parameters according to the conditions of Eq. (3.4) in [25] the generalized invariant is of the form

\[
I_T(t) = \frac{\omega_I}{2} \left( B^\dagger(t) B(t) + \frac{1}{2} \right),
\]

(30)

with \( \omega_I \) as a constant of motion, and the annihilation and creation operators are

\[
B(t) = \sqrt{\frac{m}{2}} \left\{ \left( \rho^{-1} - i \dot{\rho} \right) \hat{q}(t) - \zeta \right\} + i \frac{\rho}{\sqrt{2m}} \hat{p}(t)
\]

\[
B^\dagger(t) = \sqrt{\frac{m}{2}} \left\{ \left( \rho^{-1} + i \dot{\rho} \right) \hat{q}(t) - \zeta^* \right\} - i \frac{\rho}{\sqrt{2m}} \hat{p}(t).
\]

(31)

It holds that \([B(t), B^\dagger(t)] = 1\), where \( B(t), B^\dagger(t) \) are solely represented by the classical quantities \( \rho = \rho(t), \mu = \mu(t), \zeta = \zeta(t) \) as introduced in previous paragraphs. \( \hat{q}(t), \hat{p}(t) \) refer here to the Heisenberg operators for position and momentum and we have assumed in addition that \( f(-t_0) = 0 \).

From the Heisenberg equations of motion for \( B, \) i.e. \( dB(t)/dt = -i [B(t), \hat{H}(t)] \), one can obtain the simple time evolution for these annihilation and creation operators

\[
B(t) = e^{-i\mu(t)} B(-t_0) \quad B^\dagger(t) = e^{i\mu(t)} B^\dagger(-t_0),
\]

(32)

with \( \mu(t) \) the phase function. Their evolution in time is a simple time-dependent phase-shift mediated by the generalized classical phase referenced to the initial time \(-t_0\). This last property guarantees the time-independence of the invariant and the equivalence to the Hamiltonian (if \( f(-t_0) = 0 \) at the time \(-t_0\):

\[
I_T(t) = I_T(-t_0) = \hat{H}(-t_0).
\]

(33)

Following Kim et al. we can equate hermitian and anti-hermitian parts on both sides of Eqs. (32) by using the relations Eqs. (31) to determine the time-dependent Heisenberg operators for position and momentum

\[
\hat{q}(t) = \rho(t) \left\{ \hat{q}(-t_0) \sqrt{\omega_0} \cos \mu(t) + \frac{\hat{p}(-t_0)}{m \sqrt{\omega_0}} \sin \mu(t) \right\} + u_p(t)
\]

\[
\hat{p}(t) = \hat{q}(-t_0) \sqrt{\omega_0} \frac{\hat{p}(t) \cos \mu(t) - \rho(t)^{-1} \sin \mu(t)}{\rho(t) \cos \mu(t) - \rho(t)^{-1} \sin \mu(t)}
\]

\[
+ \frac{\hat{p}(-t_0)}{\sqrt{\omega_0}} \left\{ \frac{\rho(t)^{-1} \cos \mu(t) + \hat{p}(t) \sin \mu(t)}{\rho(t)^{-1} \cos \mu(t) + \hat{p}(t) \sin \mu(t)} \right\} + m \hat{u}_p(t),
\]

(35)

where \( \hat{q}(-t_0), \hat{p}(-t_0) \) denote position and momentum operator at time \(-t_0\), respectively. Similar as \( u_p(t) \), the classical velocity can be expressed as \( 2 \hat{u}_p(t) = \{ \hat{p} - i \rho^{-1} \}, \zeta + (\hat{p} + i \rho^{-1} \} \zeta^* \).

Our chosen initial conditions, Eq. (33), cast momentum and position operators into their standard form

\[
\hat{q}(-t_0) = \frac{1}{\sqrt{2m \omega_0}} \left\{ B + B^\dagger \right\} \quad \hat{p}(-t_0) = -i \sqrt{\frac{m \omega_0}{2}} \left\{ B - B^\dagger \right\}
\]

(36)
taking $\hbar = 1$. Kim et al. define a more general Fock state space based on number states of the invariant rather than on the Fock state space of the Hamiltonian and point out its importance and advantageous properties. However, due to our choice of the initial conditions these two state spaces are identical and their distinction is irrelevant for our discussion. We can define the Fock basis in the usual way by taking the operators at $-t_0$ according to
\begin{equation}
|n\rangle_B = |n, -t_0\rangle_B \quad \text{with} \quad |n, t\rangle_B = \frac{B^n(t)}{\sqrt{n!}} |0, t\rangle_B,
\end{equation}
where the vacuum state $|0, t\rangle_B$ is extracted from $B(t)|0\rangle_B = 0$. Furthermore, we introduce the time-independent coherent states in this Fock basis
\begin{equation}
|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle_B,
\end{equation}
with the complex amplitude $\alpha = |\alpha|e^{-i\varphi}$, because these states are the closest quantum equivalent to the classical solution and include the oscillator ground state for $\alpha = 0$. With these definitions the expectation values for the Heisenberg position and momentum operators from Eqs. (34,35) can be calculated using Eq. (36) and Eq. (38)
\begin{equation}
\langle \alpha | \hat{q}(t) | \alpha \rangle = \sqrt{\frac{2}{m}} \rho(t) |\alpha\rangle \cos[\mu(t) + \varphi] + u_p(t) \\
\langle \alpha | \hat{p}(t) | \alpha \rangle = \sqrt{2m} |\alpha\rangle \left\{ \dot{\rho}(t) \cos[\mu(t) + \varphi] - \rho^{-1}(t) \sin[\mu(t) + \varphi] \right\} + m \ddot{u}_p(t) \\
\equiv m \frac{d}{dt} \langle \alpha | \hat{q}(t) | \alpha \rangle.
\end{equation}
This way we retrieve exactly the same form for the mean values of position and momentum for the quantum solution as we obtained in Eq. (18) for the classical solution. If we disregard the zero point energy in $\langle \alpha | H(-t_0) | \alpha \rangle / \omega_0 = |\alpha|^2 + 1/2 \sim |\alpha|^2$ and set the matrix element equal to the potential energy at a classical turning point, we have $a_c \approx \sqrt{2} x_0 |\alpha|$ making the homogeneous solution of the classical and quantum formulations and hence the total solution identical. Here, $x_0 = \hbar / \omega_0 m$ is the extension of the ground state wave function of the harmonic oscillator. Alternatively, a full quantum description in the Schrödinger picture can be obtained by employing the time evolution operator that can be represented as a product of time-dependent displacement and squeezing operators [21,27].

Similarly we can compute the dispersions of $\hat{q}(t), \hat{p}(t)$ in the coherent state
\begin{equation}
\langle \alpha | (\Delta q(t))^2 | \alpha \rangle = \rho^2 / 2m \\
\langle \alpha | (\Delta p(t))^2 | \alpha \rangle = (\rho^{-2} + \dot{\rho}^2) m / 2.
\end{equation}
From the dispersion for the momentum we see that the wavepacket generally spreads solely due to the presence of the terms $\rho^2$ and $\dot{\rho}^2$. These matrix elements do not depend on the force because the force acts homogeneously in space and equally on the whole wavepacket. After periods of frequency modulations the dispersions in Eq. (40) are both time-dependent and exhibit oscillatory behaviour revealing a certain amount of squeezing [23]. For example, if we assume that after the transport we end up with a nonzero $\delta$ we can use the exact solution in Eq. (13) and evaluate the rhs of Eqs. (40). Then, the dispersions for $q(t), p(t)$ are proportional to
\begin{equation}
(cosh \delta \pm sinh \delta sin(2\omega_0 t + \theta)),
\end{equation}
distinguishable only by the $+$ and $-$ sign and constant prefactors, respectively. Therefore after the transport, the dispersions oscillate with twice the harmonic frequency and a relative phase shift of $\pi$. The strength of this squeezing oscillation is thus solely ruled by the classical quantity $\delta$. 

© 2006 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim www.fp-journal.org
Finally, our classical initial conditions $\rho(-t_0) = 1/\sqrt{\omega_0}$ together with the choice of the free parameters $c_1 = c_3 = \omega_0/m$, $c_2 = 0$ which we used to define the annihilation and creation operators in Eqs. (34), (35), provide the correct initial dispersions of the quantum formulation in this approach

$$\langle \alpha | (\Delta q(-t_0))^2 | \alpha \rangle = x_0^2/2 \quad \langle \alpha | (\Delta p(-t_0))^2 | \alpha \rangle = \hbar^2/2x_0^2.$$  

(42)

### 6 Transport dynamics in a well-controlled regime

In the following we consider an idealized situation for the transport, i.e. we assume that we could produce arbitrarily shaped external potentials in the experiment while locally maintaining parabolic potentials around $q_0$, i.e. $1 \gg |a_{\text{res}}(q_0)|/\omega_0^2$ and $|\dot{q}_0| \gg |a_{\text{res}}(q_0)|$ for all positions $q_0$ or times $q_0(t)$. Deviations from these ideal conditions due to constraints in realistic trap configurations will be evaluated in Sect. 7. In the ideal case we find from Eq. (23)

$$\Xi(t_1) = -e^{i\omega_0 t_1} \int_{-t_0}^{t_1} dt' e^{-i\omega_0 t'} \tilde{\Phi}_0(t').$$  

(43)

For $t_0, t_1 \to \infty$ we arrive at the well-known result that the transferred energy corresponds to the squared 

The two parts increase the amount of the transferred energy independently. For a real transport, where start and stop positions differ from each other, we need anti-symmetric parts in the transport function. A simple conclusion from this is that any symmetric part of the transport function can only increase the transferred energy while not contributing to the purpose of the transport, therefore we only need to consider anti-symmetric functions as candidates for transport, i.e. we take $q_0^2(t) \equiv 0$. By partially integrating Eq. (43) two times and using initial conditions for the start and stop position and velocities Eq. (2), we can also rewrite the integral in Eq. (43) as a direct functional of $q_0(t)$ that has a similar appearance but with an additional term. By symbols in this text with an extra tilde we denote quantities that are divided by the half span is fixed. Before we systematically study expression Eq. (43) we will consider two examples for which analytical solutions exist.

6.1 Two analytical examples

First we take a sine function for the transport function $q_0(t)$ as used in the experiments described in [18]

$$q_0(t) = \sin(t\pi/2t_0) \quad \text{for} \quad -t_0 < s < t_0.$$  

A graph of this function is given in Fig. 1(a). Inserting $\hat{q}_0(t)$ into Eq. (43) we find

$$\Xi(t_0) \equiv \Xi(x/\omega_0) = \omega_0 \frac{2\cos(x)}{1 - (2x/\pi)^2} \times \text{phase \ with} \quad x = \omega_0 t_0,$$  

(46)
where we converted the time variables to dimensionless units, \( x \equiv \omega_0 t_0 \), so that \( x/2\pi \) corresponds to the number of oscillation cycles. In these variables, \( |\Xi(t_0)|/\omega_0| \) is independent of the frequency and plotted in Fig. 1(b). The energy transfer is decaying overall, but shows some oscillations arising from the dependence on the energy transfer on the phase of the internal oscillation at \( t_0 \). From Eq. (45) we see that for an extreme nonadiabatic transport, i.e. \( x \approx 0 \), we have gained the full potential energy of \( m \omega_0^2 b^2/2 \). Depending on the exact transport duration we observe regular intervals where the energy drops to zero and no energy remains in the internal oscillator’s motion after the interval length \( 2t_0 \). This is due to the phase sensitivity of the transport. From Eq. (46) we have the proportionality \( |\Xi|^2 \propto (1 + \cos(2x)) \), so that we expect the first zero for \( x/2\pi = 1/4 \). However, for a transport in a harmonic well we need at least half an oscillation period for the ion to move to the other turning point, therefore we have instead \( x/2\pi = 3/4 \) which is seen in Fig. 1(b) as the first root of the adiabatic factor. The denominator in Eq. (46) cancels the first root. The adiabatic energy transfer corresponds to the envelope of this function and is given by \( 4/(1 - (2x/\pi)^2)^2 \). As we will see in the following the decay of the envelope can be sped up for different choices of the transport function.

Ideally, we want to have \( mb^2 \omega_0 |\Xi/\omega_0|^2/8\hbar < 1 \) in order to limit the maximum transferred energy to a few vibrational quanta. Let us consider some typical parameters for traps currently in use; we choose the axial frequency \( \omega = 2\pi \cdot 3 \text{ MHz} \), a typical average transport distance of about four traps (= control electrode widths), i.e. \( b = 400 \mu \text{m} \), and \( m = 9.01218 \text{ u} \) equal to the mass of a Beryllium ion. Then the adiabatic suppression factor should obey

\[
|\Xi/\omega_0|^2 < 2 \cdot 10^{-8}.
\]  

(47)

In Fig. 1(b) we have not plotted the whole range until this criterion is fulfilled. It is satisfied for about \( x/2\pi > 30 \). Thus, for the given case adiabatic transport happens on a rather long time scale, i.e. durations of \( 2 \cdot x/2\pi = 2t_0/T = 60 \) cycles. The transport in the experiment [18] which has used this transport function was performed over three times this distance requiring that \( |\Xi/\omega_0|^2 \) is lower by a factor of 9 more. The adiabatic envelope has decayed to this value at about \( x/2\pi = 52 \) yielding a transport duration of \( 2t_0/T \approx 104 \) cycles. Using a sine transport function the experimentally measured limit was around \( 2t_0/T \approx 157 \) oscillation cycles (where \( 2t_0 = 54 \mu \text{s} \) and \( \omega_0 = 2\pi 2.9 \text{ MHz} \)). This appears reasonable because the electrode array that was used in [18] was rather sparse, thus not allowing for full control and maintaining the conditions assumed in this paragraph properly. Also, the envelope in this region is quite flat; so within the uncertainties of the experiment, the experimentally observed limit is in reasonable agreement with our estimation.
We will look at an error function transport which turns out to be advantageous to the sine function in the second example

\[
\tilde{q}_0(s) = \frac{\text{Erf}(2s/t_p)}{\text{Erf}(2t_0/t_p)} \quad \text{for} \quad -t_0 < s < t_0,
\]

where we renormalized it to arrive at the times \( \pm t_0 \) at the start and end position. In addition we have introduced another time \( t_p \) which is nearly reciprocal to the slope of the transport function at the central point \( t = 0 \). Fig. 2(a) is a graph of this function for \( (t_p = 1, t_0 = 3/2) \) in arbitrary time units. Since we truncate the error function, we violate the second initial condition in Eq. (2) in a strict sense. However, we are interested only in settings where \( t_p/2t_0 < 1 \), so that this constraint for the velocity can be satisfied arbitrarily well. The adiabatic suppression factor can be evaluated analytically

\[
\frac{\Xi(x/\omega_0)/\omega_0}{2} \approx 2e^{-y^2/16} \text{Re} \{ \text{Erf} [2x/y + iy/4] / \text{Erf} [2x/y] \} \times \text{phase factor}
\]

\[
2x/y \to 2e^{-y^2/16} \times \text{phase factor},
\]

neglecting the part resulting from the finite initial and final velocities, and using the dimensionless variable \( y = \omega_0 t_p \). Fig. 2(b) illustrates the situation for \( y = 12 \) and \( x/2\pi \) in a range of time intervals the same as for the sine transport but also satisfying \( 2t_0/t_p > 1 \). It is clear that by using the error function the transport can be performed much faster than with a sine transport function, while still satisfying inequality Eq. (47). The full transport can now be performed in \( 2\pi/2\pi \approx 6 \) cycles with tolerable energy transfer. Interestingly, taking the limit for large \( 2x/y \) ratios in Eq. (49) removes the phase-sensitivity completely. However, we also want to note that the differences observed in these examples depend on experimental circumstances, e.g. for very short transport distances, the adiabatic suppression factor does not have to be small. In this case the differences between the adiabatic suppression factors is marginal in a qualitative sense. This can be seen in comparing Fig. 1(b) and Fig. 2(b) for cases when only about \( |\Xi/\omega_0|^2 < 10^{-3} \) is required, e.g. occurring for transport distances much less than an electrode width. On the other hand we find an interesting and advantageous distance scaling behaviour from Eq. (49): transporting longer distances does not require much longer time intervals. For example, Steane [20] estimated that within large-scale operation for the processing of a typical gate an average transport distance of \( \approx 22 \) traps is needed. By employing an error function transport we find that this is feasible with less than a quantum of transferred energy using the parameters \( (y = 13, x/2\pi = 4) \), i.e. in already about 8 oscillation cycles, only about a third more time than for a transport over 4 traps. The average velocity for such a transport is then considerably higher.
6.2 Near-optimum transport functions

In an attempt to optimize the transport function we can expand expression Eq. (49) up to the first order correction \[31\]

\[
\tilde{\Xi}(x/\omega_0)_{y} \rightarrow 2\omega_0 \left( e^{-b^2} - \frac{e^{-r^2}}{\sqrt{\pi}\sqrt{r^2 + b^2}} \cos(2rb + \arctan(b/r)) \right) / \text{Erf}[r] \times \text{phase factor},
\]

(50)

with \( r = 2x/y \equiv 2t_0/t_p, \ b = y/4 \). Because of \( 2rb = x \) the zeros of the suppression factor \( \Xi^2 \) are equally spaced as in the sinusoidal example if we disregard the phase in the cosine function, i.e. half periods \( \Delta(x/2\pi) = 1/2 \) since the particle can arrive from two different turning points at the end of the transport.

From this expansion it is clear that the ratio \( r \) basically determines the magnitude of the second term on the rhs of Eq. (50) and suppresses the phase sensitivity as it increases. If \( r \) is chosen large enough the energy transfer is dominated by \( b \). To find some conditions that are close to optimum we proceed in the following way: first, from the transport distance and achievable frequency we can evaluate the upper bound for the adiabatic suppression factor as in Eq. (47). Because we have to satisfy \( 2t_0 > t_p \) we then choose \( b \) large enough to suppress the first exponential factor to fulfill the given criterion. This procedure defines the asymptotic value of energy transfer for large \( x/2\pi \) as shown in Fig. 2(b). We then choose the interval length \( x/2\pi = rb/\pi \) by defining \( r \) such that we are just in the asymptotic range. The near phase-insensitivity can then be thought of as a result of the extremely slow start where the phase information in the limiting case \( r \rightarrow \infty \) in Eq. (49) gets totally lost.

6.3 High-frequency limit, adiabatic transport, and approximate trajectories

To better understand the behaviour discussed in previous sections, we present a few more general considerations. We write the time-dependence of the transport function according to \( q_0(t) = \vartheta_0(t/t_0) = \vartheta_0(\tau) \) so that \( \vartheta_0 \) only depends on the dimensionless variable \( \tau \) (for the error function example we also keep the ratio \( t_p/t_0 \) fixed). Making the substitution \( t' \rightarrow \tau \) in the integral in Eq. (43) we find

\[
\Xi(t_0)/\omega_0 = (\omega_0t_0)^{-1} \int_{-1}^{1} d\tau e^{-i\omega_0t_0\tau} \vartheta_0''(\tau) \times \text{phase factor}
\]

\[
\omega_0t_0 \rightarrow \infty \quad 2\sum_{n=0}^{N} \frac{\vartheta_0(n+2)(-1)^{n+2}}{(-\omega_0t_0)^{n+2}} \cos(\omega_0t_0 + n\pi/2) \times \text{phase factor}.
\]

The exponent in the integral relates the two time scales in \( \omega_0 \) and \( t_0 \). Using the method outlined in appendix A we expand it into the sum given in the second line in the limit \( \omega_0t_0 \rightarrow \infty \) assuming that \( \vartheta_0 \) is sufficiently smooth. In this expansion the derivatives at the start position (and end position due to anti-symmetry) define the energy transfer in the transport, and thus fully characterize the transport function for the transferred energy in the adiabatic limit. This provides us with a reason for the difference we observed above for the error function and sine examples. The second derivative for the sinusoidal transport is nonzero at \( \pm t_0 \) and much larger than in the case of the error function. In the latter all derivatives are damped by a gaussian while the ones for the sine transport alternate. Furthermore, we see that we can in general decrease the transferred energy for larger values of the product \( \omega_0t_0 \), i.e. by taking \( \omega_0 \) to infinity (high-frequency limit), we can lower the adiabatic suppression arbitrarily, on the other hand, slowing down the motion by increasing the length of the duration of the transport \( 2t_0 \), we move into the adiabatic regime. For infinitely slow motion we end up with zero transferred energy. These two limiting cases are formally equivalent because the energy transfer depends only on their relative time scale. We can perform the same expansion starting from Eq. (43) directly and use the relation \( u_x(t) = [\Xi(t) - \Xi^*(t)]/2i\omega_0 \) to find approximate trajectories valid in the same
limits
\[
\lim_{t \to 0} \sum_{n=0}^{\infty} \frac{1}{\omega_n^{n+2}} \left[ \cos \left( \frac{n\pi}{2} \right) q_0^{(n+2)}(t) - \cos \left( \frac{n\pi}{2} \right) q_0^{(n+2)}(-t_0) \right].
\]

### 7 Regularized trap-electrode waveforms, potential fluctuations and aspect-ratio rule

#### 7.1 Determination of waveforms

So far, we have said nothing about how to determine the waveforms applied to the electrodes. As soon as we have the waveforms at hand for a given model electrode configuration we can determine the magnitudes of perturbations. This is done in the next section. Here, we seek optimum solutions for a given electrode structure in order to keep the uncontrolled part of perturbations. This is done in the next section. Here, we seek optimum solutions for a given electrode configuration we can determine the magnitudes and dimensionless time-dependent amplitudes \( a_m(t) \) of the form

\[
\phi(q, t) = \sum_m a_m(t) \phi_m(q).
\]

To optimize waveforms for the time-dependent amplitudes for the transport problem we find a measure of the discrepancy by integrating over the residual non-matched part according to

\[
\min_{a_m, \phi_0} \int_{q_0(t) - \delta q}^{q_0(t) + \delta q} \left| \phi_{\text{res}}(q, t) \right|^2 dq \quad \forall \ t
\]

while \( \phi(q, t) \) from Eq. (52) enters here through Eq. (4). For any time \( t \) we want to find a set \( a_m, \phi_0 \) for which expression Eq. (53) is minimal. The integration is performed over an interval moving with the minimum of the parabolic potential well, i.e. \([q_0(t) - \delta q, q_0(t) + \delta q]\) and assuming a unity weight factor in the integrand. We do not consider in this range any lag of the ion due to acceleration and deceleration since for an adiabatic transport and experimental conditions the lag is much smaller compared to the optimization range. \( \phi_0(t) \) represents here another degree of freedom that does not perturb the dynamics but might allow one to more optimally choose the harmonic potential well by arbitrarily offsetting the desired parabolic potential for best fit. Condition Eq. (53) is readily converted into an linear system of equations by taking partial derivatives for the amplitudes \( a_m \) and \( \phi_0 \), and setting them all equal to zero. The minimization problem in Eq. (53) then reads

\[
\sum_{m=1}^{n_m} a_m(t) \int_{\mathcal{S}_q} dq \phi_m \phi_j + \phi_{0}(t) \int_{\mathcal{S}_\varphi_0} dq \phi_j = \eta \frac{m\omega^2}{2Q} \int_{\mathcal{S}_q} dq (q - q_0(t))^2 \phi_j \quad \forall \ j,
\]

where we dropped the explicit integral bounds and arguments of the potentials \( \phi_m(q) \) for the sake of simplicity. Bold symbols denote matrices, underlined symbols vectors.

The optimization problem can then be formulated in terms of the linear system

\[
\mathbf{S} + \mathbf{q} \equiv \left( \mathbf{S}_q \ \mathbf{S}_\varphi_0 \right) \cdot \left( \begin{array}{c} a \\ \varphi_0 \end{array} \right) = \eta \mathbf{K} \quad \forall \ t,
\]

with \( \eta = \frac{m\omega^2}{2Q} \). All quantities are functionals of \( q_0 \) and for a given transport function \( q_0(t) \), we need to solve the equations at every point in time. As a result we obtain the waveforms \( a_m(t) \equiv a_m[q_0(t)] \).
Typically we choose an optimization range of $2\delta q = 0.5W$ for electrodes of width $W$. This is usually much smaller than the mean distance between most of the contributing electrodes to the center of the parabolic well. Thus, due to the slow decay of the axial potentials, the curvatures of distant electrodes are similar, and their contribution differs locally only by a multiplication constant. This is particularly true for experimental situations where the high electrode density typically makes the system Eq. (55) nearly singular. A straightforward least-square method, such as

$$a_{t0} = \arg \min_{a_{t0}} \left\{ \| S_{t0} \cdot a_{t0} - \eta K \|^2 \right\},$$

is therefore not well suited for finding waveform amplitudes. For high electrode density, a tiny step $q_0 \rightarrow q_0 + \delta$ might change the individual electrode amplitudes exponentially fast. In these cases the matrix $S_{t0}$ in Eq. (55) filters out too much information from $a_{t0}$ to invert this system properly. In mathematical language these kind of problems belong to the family of discrete ill-posed problems that can be numerically solved using regularization approaches [35]. Here, the lost information is fed back in the minimization process via a Lagrangian multiplier concerning the smoothness of $a_m(t)$, or curvature etc. in amplitude space. If we apply a Tikhonov regularization to the given problem we have to solve

$$a_{t0,\nu} = \arg \min_{a_{t0}} \left\{ \| S_{t0} \cdot a_{t0} - \eta K \|^2 + \nu^2 \| L(a_{t0} - a_{t0}^0) \|^2 \right\}$$

in order to determine smooth time-dependent waveform amplitudes $a_m(t)$. In Eq. (57) the regularization parameter $\nu$ corresponds to a weight factor between the original least-square minimization and the additional side constraints, while $a_{t0}^0$ can be used to find solutions near a prescribed setting. The smoothing properties of this optimization originate from a common and simultaneous minimization of both terms. $L$ is a linear operator that can be used to feed back different kinds of information to the amplitudes. For the results given here we took for $L$ the unity operator, and also $a_{t0}^0 = 0$. Since we only want to limit the amplitudes $a_m$ to some appropriate experimental values and stabilize the solution, our interest is not to determine the overall minimum of Eq. (57) in a self-consistent way. For our convenience we choose $\nu$ manually to make the parameters compatible with available technology.

We can summarize the advantages of these methods to the current optimization problem:

1. The regularization method selects only nearby electrodes for creating a local parabolic potential, and disregards tiny linear contributions from distant electrodes which would require large amplitudes to effect small changes.
2. The choice of the regularization parameter limits the amplitudes $a_m$ to practical experimental values.
3. It is robust against changing the electrode density (here, governed by the widths $W$). This will be of importance in the next section.
4. It stabilizes the output waveforms and smoothes sharp features in the time-dependence of the amplitudes. Different constraints can be set via the $L$ operator, defining bounds or curvatures in the amplitude space.

For more detailed information we refer the reader to the mathematical literature [35]. A typical example of a parabolic potential created through superposition of an array of electrodes and the time-dependence of amplitudes is shown in Fig. 4 and further discussed in the next section.

### 7.2 Potential fluctuations and aspect-ratio rule

Based on a reasonable multi-electrode structure we want to estimate how well we can meet the requirements on transport potentials stated above, in particular, how stringently we can meet $|a_{m}(q_{0})|/\omega_{0}^{\omega} \ll 1$ and $|a_{res}(q_{0})| \ll |q_{0}|$. We employ the definition of waveforms and the method from the previous section for extracting the residual, uncontrolled potential $\phi_{res}$ from which we perturbatively derive the effect of
imperfections on the transport. As a simple model electrode structure for transport in single and multi-layer traps, we use the “railway track” electrode configuration sketched in Fig. 3(a) which might be a simple model for transport in single and multi-layer traps [19, 22, 32, 33]. The transport occurs along the long arrow where we assume the ion is held radially by RF fields and controlled axially by the electrical fields arising from the potentials of the “stripe” electrodes depicted in Fig. 3(a). We are mainly interested in the scaling behaviour as a guideline for general design rules. Waveforms that are actually used in experiments should be based on more accurate numerical potentials and generalized versions of Eq. (54) for all three dimensions. We are finally interested in the trade off between adding electrodes, by shrinking the electrode distances/widths along x, and the amount of control that is gained in that way.

We can model this arrangement as a sum over the potentials \( \phi_m(x) \) of several infinitely long (in the y-direction) stripe electrodes that are distributed along the x-axis

\[
\phi(x) = \sum_m a_m \phi_m(x) = \frac{1}{\pi} \sum_m a_m \arctan\left( \frac{W}{1 + (\hat{x} - mW)^2 - W^2/4} \right), \tag{58}
\]

where each \( \phi_m(x) \) is the exact solution of the Poisson equation for an infinitely long stripe at position \( m \cdot W \) that is embedded in a ground plane. For convenience we choose for the individual potentials in this basis set a potential on the electrodes of \( \phi(x_i) \) which is obtained from

\[
\omega^2(W, z_{ion}) = \frac{Q}{m_{ion}^2} \frac{\partial^2}{\partial \hat{x}^2} \phi_0(\hat{x}) \bigg|_{\hat{x}=0} = -\frac{2a_0}{\pi m} \frac{1}{z_{ion}^2} \frac{W}{(W/2)^2 + 3W^2/2 + 1}. \tag{59}
\]

The proportionality \( \omega \propto z_{ion}^{-1} \) is rather unusual and stems from the fact that we scale only a single dimension (along x). The last factor of the second equation is solely defined by the aspect ratio \( W \). We denote symbols with a hat as quantities normalized to the ion distance \( z_{ion} \) to the surface, e.g. the normalized electrode width \( \hat{W} = W/z_{ion} \). Fig. 3(b) shows the behaviour of \( \phi_0(\hat{x}) \) for various geometric aspect ratios \( \hat{W} \). We see that a plateau-like structure starts to form for \( \hat{W} \sim 2 \) and larger resulting in small field gradients along the transport direction in the center of each stripe electrode. The maximum frequency at the center of electrode \( m = 0 \) is obtained from

\[
\omega(t, \hat{
abla} \phi) \bigg|_{\hat{x}=0} = \frac{\omega_0}{\sqrt{1 + \frac{m^2}{n^2}}}. \tag{60}
\]

where each \( \phi_m(x) \) is the exact solution of the Poisson equation for an infinitely long stripe at position \( m \cdot W \) that is embedded in a ground plane. For convenience we choose for the individual potentials in this basis set a potential on the electrodes of \( \phi(x_i) \) which is obtained from

\[
\omega^2(W, z_{ion}) = \frac{Q}{m_{ion}^2} \frac{\partial^2}{\partial \hat{x}^2} \phi_0(\hat{x}) \bigg|_{\hat{x}=0} = -\frac{2a_0}{\pi m} \frac{1}{z_{ion}^2} \frac{W}{(W/2)^2 + 3W^2/2 + 1}. \tag{59}
\]

The proportionality \( \omega \propto z_{ion}^{-1} \) is rather unusual and stems from the fact that we scale only a single dimension (along x). The last factor of the second equation is solely defined by the aspect ratio \( W \). We denote symbols with a hat as quantities normalized to the ion distance \( z_{ion} \) to the surface, e.g. the normalized electrode width \( \hat{W} = W/z_{ion} \). Fig. 3(b) shows the behaviour of \( \phi_0(\hat{x}) \) for various geometric aspect ratios \( \hat{W} \). We see that a plateau-like structure starts to form for \( \hat{W} \sim 2 \) and larger resulting in small field gradients along the transport direction in the center of each stripe electrode. The maximum frequency at the center of electrode \( m = 0 \) is obtained from

\[
\omega(t, \hat{
abla} \phi) \bigg|_{\hat{x}=0} = \frac{\omega_0}{\sqrt{1 + \frac{m^2}{n^2}}}. \tag{60}
\]
While for smaller ratios the frequency fluctuations are in the percentage range (middle panel), the emulation of the uncontrolled acceleration $-\omega_0^2 a_{\text{res}}(\xi)/\omega_0^2$, and the middle and lower graphs the frequency modulation $\omega(t)$ for various aspect ratios. In both figures a dramatic change of the curves is observed around $W \sim 1$. While for smaller ratios the frequency fluctuations are in the percentage range (middle panel), the emulation of the potential for larger ratios is much worse due to the constraint $|a_m| \leq 2$ (lower panel). Frequencies drop by more than 90% already for the $W = 2$ calculation. Only in the strongly regularized scheme, did we find a direct correlation of the solution vector $a_m$ to the choice of the regularization parameter. In the weakly regularized scheme the amplitudes were limited by other lower bounds, and the waveform solution appeared similar over a large range of $\nu$, but exhibited a much noisier behaviour. This enhanced sensitivity is an indication that inclusion of more electrodes (smaller $W$) does not improve the quality of the solution anymore. The linear system becomes more singular and exhibits more rank-deficiency, i.e. rows and columns become more equal and their inclusion adds more redundancy. For the given parameters we observed that for $W = 0.5$ the transition from a regularized to a weakly regularized solution occurred. Therefore, our results indicate that $W = 0.5 - 1.0$ should be optimal for the configuration discussed here. For larger aspect- ratios we found that the coverage of curvatures of the individual potentials along the transport axis is not sufficient for the necessary amount of control.

The other constraint, i.e. $|\hat{a}_{\text{res}}(\xi)| \ll |\hat{\phi}_{\text{res}}|$, of a controlled transport force, has to be interpreted dynamically. Since the acceleration force depends on the time duration in which the transport is performed, this requirement can be violated for a slower transport. In Fig. 5) we show that in the initial phase the perturbations overwhelm the transporting acceleration for aspect ratios $W \geq 1.5$ or larger. Results for smaller aspect ratios are not given in this figure because the transport force by far dominates the excess force and lead to a fully controlled transport.

In general, fluctuations in the frequency and transport force affect an energy transfer according to Eq. (23). This introduces violations to the symmetry of the transporting force and leads to an enhanced energy transfer as seen from Eq. (44). We have not included higher order terms in our discussion, because we aim for experimental conditions to perform a transport in the well-controlled regime. However, they are inevitable...
Fig. 4 (a) Creation of a harmonic well by superposing potentials of the electrodes of the stripe configuration of Fig. 3(a). The fine lines represent individual electrode potentials that sum up to the total potential represented by the thicker line. The dashed-dotted line shows the ideal harmonic potential and the two vertical lines indicate the range of optimization at this given location. (b) Waveform amplitudes of electrodes 18 to 24 for a transport from trap 19 to 23 for $\hat{W} = 1$. Waveforms from more distant electrodes still contribute but are not shown for the sake of clarity. Amplitudes 21–24 are marked by symbols for better visualizing their traces. The transport is done according to an error function, while the abscissa represents time in units of oscillation periods $t/T$.

Fig. 5 First and second order perturbation to the transport (thick lines), and their dependence on various aspect ratios $\hat{W}$. The abscissa $t/T$ corresponds to the number of oscillation cycles with period $T = 2\pi/\omega_0$. The labels on the graphs annotate the value of $\hat{W}$. (Upper panel) Residual, uncontrolled acceleration $-\tilde{a}_{\text{res}}(q_0)/\omega_0^2$ (dimensionless) which should be compared to $\tilde{q}_0/\omega_0^2$ (dotted line). This is only given for the large aspect ratios $\hat{W} = 1.5, 2.0$. The graphs for smaller ratios would be close to zero on this scale. (Lower panel) Relative frequency modulation $\omega(q_0)/\omega_0$ during an error function transport for various aspect ratios. For explanations see text.

for longer transport distances and other types of motion, such as nonadiabatic transport, or splitting of ion groups where they might lead to large energy transfers.
8 Conclusions

In conclusion, we have analyzed the dynamics of single ion transport in microstructured linear Paul trap arrays. We have modeled the transport by a forced and parametrically excited harmonic oscillator and have presented a theoretical framework for its description. We have derived exact analytical expressions for the classical as well as quantum dynamics and reviewed their related properties. In particular we have expressed the Heisenberg operators by the approach of Kim et al. [21] through the dynamical quantities of the related classical solution. We have given explicit analytical expressions for the classical energy transfer involved in these transport phenomena and derived expressions for the lowest order deviations from ideal transport that will necessarily appear for unfavourable ion trap layouts. For current trap technology we have evaluated durations for a fast adiabatic transport and found that they depend strongly on the external force employed in the transport. According to these results, the adiabatic single ion transports of reference [18] could be sped up by more than an order of magnitude with negligible energy transfer to the motion. We determined appropriate transport waveforms and found that with an adiabatic transport over four electrode stripes of size roughly equal to the distance of the ion to the nearest electrode and frequencies in the range of $\approx 9 \text{ MHz}$ is feasible in about 6 oscillation cycles. Our results also indicate that a full control over the transport is available, where perturbations to a harmonic oscillator potential are negligible at all positions and times.

By directly relating deviations from these ideal potentials to the aspect ratio of the trap, we have found a practical design rule that should be valid for trap layouts more general than the one given here. The ratio of a control electrode width to the distance to the ion should be in the range $0.5 - 1$ for a well-controlled regime. This provides important insight into the amount of resources needed to realize large scale implementations of ion trap based quantum computers. Transport in a confining well of constant frequency might also enable continuous cooling processes during the transport. If eventually experiments allow one to maintain a well-controlled regime during the transport, performing quantum processing during transport is conceivable, possibly leading to appreciably shorter processing times.

Acknowledgements R.R. acknowledges support by the Alexander von Humboldt Foundation during the course of this work. Work also supported by DTO and NIST. We thank A. Steane, T. Rosenband and J. Wesenberg for helpful comments on the manuscript.

A Integral expansion

We employ a mathematical theorem proven within the formalism of h-transforms, see for example theorem 3.2 of [36]: If $g(\tau)$ has $N + 1$ continuous derivatives while $g^{(N+2)}$ is piecewise continuous on the real axis $[a, b]$ then

$$I(\lambda) = \int_a^b e^{-i\lambda \tau} g(\tau) d\tau \xrightarrow{\lambda \to \infty} \sum_{n=0}^{N} \frac{(-1)^n}{(-i\lambda)^{n+1}} \left[ g^{(n)}(b)e^{-i\lambda b} - g^{(n)}(a)e^{-i\lambda a} \right].$$

(60)

If we also require $\lim_{\tau \to a^+} g^{(n)}(\tau) = \lim_{\tau \to b^-} g^{(n)}(\tau) = 0$ for $n = 0, \ldots, N - 1$ it holds that $I(\lambda) = o(\lambda^{-N})$.

References

Microfabricated Surface-Electrode Ion Trap for Scalable Quantum Information Processing


Time and Frequency Division, NIST, Boulder, Colorado 80305, USA

(Received 9 February 2006; published 30 June 2006)

Individual laser-cooled $^{24}\text{Mg}^+$ ions are confined in a linear Paul trap with a novel geometry where gold electrodes are located in a single plane and the ions are trapped 40 $\mu$m above this plane. The relatively simple trap design and fabrication procedure are important for large-scale quantum information processing (QIP) using ions. Measured ion motional frequencies are compared to simulations. Measurements of ion recooling after cooling is temporarily suspended yield a heating rate of approximately 5 motional quanta per millisecond for a trap frequency of 2.83 MHz, sufficiently low to be useful for QIP.

Recent interest in coherent quantum state control and methods to realize practical quantum information processing (QIP) has led to impressive developments in quantum processing using several different physical systems [1]. Single quantum bit (qubit) rotations, two-qubit gates, and simple quantum algorithms have been implemented. However, perhaps the most significant challenge for any possible physical implementation of a quantum processor is to devise methods that scale to very large numbers of quantum information carriers.

The system of trapped ions is an interesting candidate for QIP because the basic requirements [2] have been demonstrated in separate experiments [1], and several schemes for scaling this system to large numbers of qubits have been proposed [1,3–7]. One approach is based on a network of interconnected processing and memory zones where ion qubits are selectively shuttled between zones [3,6]. Within this approach, miniature linear trap arrays [8–11] and a three layer $T$-junction trap [10] have been demonstrated. Since the speed of most multi-ion qubit gates is proportional to the ions’ motional frequencies and these frequencies are inversely proportional to the square of the trap dimensions, we would like to decrease the size of these dimensions. To do this robustly, microfabrication techniques are required. Three-dimensional traps have been demonstrated with boron-doped silicon [12] and monolithically fabricated gallium-arsenide electrodes [11]. A significant simplification in fabrication could be achieved if all trap electrodes reside on a single surface and the ions are trapped above this surface [13]. In this case, the trapping electric fields would be the analog of magnetic fields used in “chip” traps for neutral atoms (see [14] and references therein). Surface-electrode ion traps have the potential added benefit for scaling that microelectronics for electrode potential control can be fabricated below the plane of the electrodes [15].

Recently, macroscopic charged particles have been confined in a surface-electrode trap [16]. Storage of atomic ions, however, requires substantially different experimental parameters. In this Letter we report the first demonstration of stable confinement of atomic ions in a surface-electrode trap. The trap is constructed with standard and scalable microfabrication processes. We load $^{24}\text{Mg}^+$ into this trap, measure the motional frequencies of the ions, and find reasonable agreement with those determined from simulations. We also determine a motional heating rate of the ion(s) that is low enough to allow for high fidelity logic operations.

The standard linear radio-frequency (rf) Paul trap [17] consists of four parallel rods whose centers are located on the vertices of a square [Fig. 1(a)]. An rf potential is applied to two opposing rods with the other two (control electrode) rods held at rf ground. This configuration creates a nearly harmonic ponderomotive pseudopotential in the $\hat{x}$-$\hat{y}$ plane. Longitudinal confinement for a single trapping zone is obtained by segmenting the control electrodes along their length and applying appropriate static potentials to the different segments. Several variations on this design have been demonstrated [8–12], but it is very desirable to simplify their construction. A straightforward way to modify the 3D design of Fig. 1(a) is to place the four rods in a common plane, with alternating rf and control electrodes [13]; one version of this geometry is shown in Fig. 1(b). In this design, the rods are replaced with flat electrodes, as shown in Fig. 2.

![FIG. 1. (a) Standard linear rf Paul trap; (b) surface-electrode geometry where all electrodes reside in a single plane, with the ions trapped above this plane.](image-url)
We can fabricate this electrode structure by means of photolithography and metal deposition using evaporation and electrodeposition. For the substrate we use polished fused quartz, a material with low rf loss. A 0.030 µm titanium adhesion layer and a 0.100 µm copper seed layer are first evaporatively deposited onto the substrate. This deposition is uniform except for small areas for resistors where the quartz is left exposed. Resistors (~1 kΩ) and leads are fabricated through a lift-off process that entails patterning them with standard photolithography and evaporation of a 0.013 µm titanium adhesion layer followed by 0.300 µm of gold. Resistors are fabricated directly on the quartz substrate; leads are fabricated on top of the copper seed layer. The gold electrodes near the trapping region are electroplated onto the copper seed layer after a second photolithographic patterning step. Afterward, the exposed initial seed and adhesion layers are etched away to isolate electrodes and leads. The trap electrodes are plated to a thickness of ~6 µm so that the ratio of height to interelectrode spacing is relatively high (the interelectrode spacing is ~8 µm). This should reduce alteration of the trapping potential due to stray charges that may collect on the exposed insulator between electrodes.

We create ions in the trap by photoionizing thermally evaporated neutral magnesium atoms. The magnesium source is realized by resistively heating a stainless steel tube containing solid magnesium, which is sealed at both ends and has a small slit from which evaporated magnesium atoms emerge. With a planar electrode geometry, there is a risk of shorting electrodes to each other due to magnesium deposited onto the trap structure. To reduce this risk, in a last processing step we perform a controlled hydrofluoric acid (HF) etch of the central trap region. The HF etches away a small part of the titanium adhesion layer and the substrate, without affecting the electrodes. The result is an ~2 µm horizontal undercut of the electrodes to help prevent shorting due to deposition from the magnesium source. As a further precaution, we direct the magnesium flux nearly parallel to the surface and avoid as much as possible having the channels between electrodes be parallel to the flux (Fig. 2).

We use five independent control electrodes to provide sufficient degrees of freedom to be able to overlap the electric field null point of the static potential and the rf pseudopotential minimum. We create a low impedance path for the rf to ground on the control electrodes with capacitors (820 pF) that are surface mounted directly onto the chip (in the future, capacitors could be included as part of the fabrication process). Gold ribbons for applying the electrode potentials are gap welded to contact pads.

The trap structure is mounted in a copper tube that also serves as part of an rf transformer [18] and the entire structure is surrounded by a quartz envelope. The system is baked under vacuum prior to operation to reach a base pressure below 10^-8 Pa with the use of an ion getter pump combined with a titanium sublimation pump.

As we describe below, the trap well depth $U_\text{RF}$ for a surface-electrode trap is fairly shallow [13], not much above the mean kinetic energy of the neutral atoms before they are ionized. Nevertheless, we can load $^{24}\text{Mg}^+$ ions efficiently by resonant two-photon photoionization (PI) at 285 nm [19]. The PI laser, resonant with the $3s^3\text{S}_0 \rightarrow 3s3p^1\text{P}_1$ electric dipole transition in neutral magnesium, copropagates with a Doppler-cooling beam tuned approximately 400 MHz below the $3s^2\text{S}_{1/2} \rightarrow 3p^2\text{P}_{1/2}$ electric dipole transition in $^{24}\text{Mg}^+$ at 280 nm. The laser beams are parallel to the trap surface, and at an angle of approximately 45° with respect to the trap z axis as shown in Fig. 2(a).

Since the laser beam direction has significant overlap with all principal trap axes, cooling will be efficient in all directions [13]. During loading, both the Doppler-cooling and PI beams have 2 mW power and waists of ~40 µm. The atomic flux of magnesium intersects the laser beams at the trap [Fig. 2(a)]. The cooling beam is applied continuously, while the PI beam needs to be applied for only a few seconds to create ions in the trap. Ions are loaded with the first configuration of trap potentials of Table I, which has the largest trap depth. Ions are detected by observing $3p^2\text{P}_{1/2} \rightarrow 3s^2\text{S}_{1/2}$ fluorescence along a direction perpendicular to the trap surface with a CCD camera as in the view of Fig. 2(b). Despite the fact that the center of the laser beam is only 40 µm above the surface, which increases the risk of scatter from light striking the trap electrodes, the signal-to-background ratio for scattered light from the ions is greater than 100 when the Doppler-cooling laser is tuned for minimum temperature, approxi-
mately 20 MHz (one-half linewidth) below resonance with intensity slightly below saturation.

We measure the oscillation frequencies of an ion in the trap by applying an oscillating field to a control electrode and observing a change in fluorescence rate when the frequency of the applied field is equal to one of the motional frequencies, thereby heating the ion [18]. To excite the axial (transverse) mode, we apply the oscillating field to electrode 2 (1).

As an aid in initially determining the correct operating conditions, trapping potentials are determined using numerical solvers (boundary element method) subject to the constraint that the rf pseudopotential minimum overlaps the null points of the electric field from the static potential, to minimize rf micromotion (see, for example, [20]). For the experiments described here, the static potentials on each control electrode, expressed as a fraction of the potential $V_5$ on electrode 5 [Fig. 2(b)], are $V_1 = 0.320$, $V_2 = 0.718$, $V_3 = 0.738$, and $V_4 = -0.898$.

The peak potential amplitude $V_{\text{rf}}$ applied to the rf electrode (87 MHz) is difficult to measure directly. To determine it, we measure the three mode frequencies for three different experimental conditions shown in Table I and compare them to simulations. Inputs to the simulations are $V_5$, the measured electrode dimensions, and the ratio of the two rf powers applied to the trap (which can be accurately measured). A least squares fit, weighting all 9 frequencies equally, gives the values of $V_{\text{rf}}$. Since, under ideal circumstances, one set of conditions is sufficient to extract $V_{\text{rf}}$, an indication of the agreement between experiment and theory is given by the level of agreement between measured and predicted frequencies for all three cases (a few percent here). We believe the primary cause for the disagreement is due to the presence of additional stray static potentials on the electrodes. Finally, we numerically determine $U_T$ in the pseudopotential approximation [21,22] for the three cases of Table I. A transverse cross section [the $\hat{x}$-$\hat{y}$ plane of Fig. 1(b)] of the trapping potential for $V_5 = 5$ V and $V_{\text{rf}} = 103.2$ V near the central trap region is shown in Fig. 3.

In Fig. 4, we show groups of ions for the $V_{\text{rf}} = 103.2$ V and $V_5 = 2.00$ V configuration. The separation of the ions is related to the center-of-mass axial oscillation frequency [3], and the horizontal bars indicate the expected ion separations according to the measurement of this frequency. When the number of ions becomes large enough, the string breaks into a zigzag configuration (see, for example, [23]).

Because $^{24}\text{Mg}^+$ lacks hyperfine structure, we cannot easily determine ion motional heating rates near the quantum limit using Raman transitions [24] (without applying large magnetic fields) due to poor internal state discrimination. An approximate value can be determined as follows. After nulling rf micromotion [20], we adjust the cooling beam for minimum temperature (see above). The beam is then blocked for a period to allow the ion to heat up. Upon resuming cooling, the fluorescence is initially smaller than its steady state value due to Doppler broadening; as the ion recools, the Doppler broadening lessens and the fluorescence increases until it reaches its steady state value given by the Doppler-cooling limit. We examine this recooling data for conditions where $\{f_{\text{rf}}, f_{\perp 1}, f_{\perp 2}\} = \{2.83, 15.78, 17.13\}$ MHz (Table I). Since heating ($dE/dt$) is observed to be stronger at lower frequencies [25,26], we will assume that the axial direction is predominantly heated. We detect fluorescence in successive time bins of 50 $\mu$s duration and find that the initial averaged fluorescence decreases after blocking the cooling beam a few seconds. We compare these results to simulations where we assume the ion (after the beam-blocked period) under-

<table>
<thead>
<tr>
<th>$V_5$ (V)</th>
<th>$V_{\text{rf}}$ (V)</th>
<th>$f_1$ (MHz)</th>
<th>$f_{\perp 1}$ (MHz)</th>
<th>$f_{\perp 2}$ (MHz)</th>
<th>$U_T$ (meV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.00</td>
<td>103.2(5)</td>
<td>2.83</td>
<td>15.78</td>
<td>17.13</td>
<td>(Expt.)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.77</td>
<td>15.62</td>
<td>17.17</td>
<td>177(6)</td>
</tr>
<tr>
<td>2.00</td>
<td>103.2(5)</td>
<td>1.84</td>
<td>15.87</td>
<td>16.93</td>
<td>(Expt.)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.75</td>
<td>16.18</td>
<td>16.78</td>
<td>193(11)</td>
</tr>
<tr>
<td>5.00</td>
<td>46.1(2)</td>
<td>2.85</td>
<td>5.28</td>
<td>8.29</td>
<td>(Sim.)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.77</td>
<td>5.02</td>
<td>8.73</td>
<td>6(1)</td>
</tr>
</tbody>
</table>

FIG. 3. A transverse cross section of the simulated trapping potential (both rf pseudopotential and static potentials included) for potentials corresponding to $V_5 = 5.00$ V and $V_{\text{rf}} = 103.2$ V ($U_T = 177$ meV). The cross indicates the directions of the normal mode axes $n_{\perp 1}$ and $n_{\perp 2}$, and the expected position for an ion at the center of the trap in the $\hat{x}$-$\hat{y}$ plane. The separation between contour lines corresponds to 5 meV. The electrodes are depicted to scale in the lower part of the figure and labeled as in Fig. 2(b).
goes harmonic oscillations with an amplitude corresponding to the Maxwell-Boltzmann energy distribution. Cooling is assumed to proceed as described in [27], which then determines fluorescence versus time. Comparing experiments and simulations for blocking times between 1 and 10 s, we determine an axial heating rate of 4.8 ± 0.4 quanta per millisecond (statistical uncertainty). This heating rate should be sufficiently small to allow high fidelity QIP operations [28].

The heating determined this way includes the effects of ion or neutral collisions and should give an upper limit on the heating that affects QIP [29]. Also, the heating rate determined here agrees with that determined from the escape time from the shallow well (\( U_T \approx 6 \text{meV} \)). In [11], this method gave a significantly higher value than that determined using Raman transitions.

Johnson noise in the resistance of the \( RC \) filters on the control electrodes is a potential source of heating [3], but we theoretically estimate it to contribute only 1 quantum per second. Therefore the observed heating is apparently dominated by anomalous heating as observed in other experiments [25,26]. We note that in the presence of Doppler-cooling light, the lifetime of the ion is several hours (in the trap with \( U_T = 177 \text{meV} \)), presumably limited by chemical reactions [30].

An important next step is to replace \(^{24}\text{Mg}^+\) with, for example, \(^{25}\text{Mg}^+\) or \(^{9}\text{Be}^+\), to study heating at the quantum level [24,25]. A longer-term goal is to design and fabricate a 2D array of surface-electrode trapping zones [6].

We thank E. Langlois and J. Moreland for advice on the microfabrication processes and Y. Le Coq for helpful comments on the manuscript. This work was supported by the Advanced Research and Development Activity (ARDA) under Contract No. MOD-7171.05 and NIST. S.S. wishes to thank the Carlsberg Foundation for financial support. Likewise, J.H.W. thanks the Danish Research Agency.

*Electronic address: seidelin@boulder.nist.gov

1Present address: Los Alamos National Laboratory, Los Alamos, NM 87545, USA.

2Present address: Quantum Information Processing, University of Ulm, D-89069 Ulm, Germany.


[29] The duration of most QIP experiments (~1 ms) is small compared to the mean time between (coherence-destroying) collisions (typically >>1 s). Hence collisions affect only a negligible fraction of the QIP experiments.


FIG. 4. Images of 1, 2, 3, 6, and 12 ions confined in the surface-electrode trap (\( V_H = 103.2 \text{V} \) and \( V_V = 2.00 \text{V} \)). The length scale is determined from a separate image of the electrodes whose dimensions are known. The horizontal bars indicate the separation distance between the ions as predicted from the measured axial oscillation frequency. The ratio between transverse and axial oscillation frequencies makes it energetically favorable for the 12 ion string to break into a zigzag shape (see, for example, [23]).
Creation of a six-atom ‘Schrödinger cat’ state

D. Leibfried1, E. Knill1, S. Seidelin1, J. Britton1, R. B. Blakestad1, J. Chiaverini1†, D. B. Hume1, W. M. Itano1, J. D. Jost1, C. Langer1, R. Ozeri1, R. Reichle1 & D. J. Wineland1

Among the classes of highly entangled states of multiple quantum systems, the so-called ‘Schrödinger cat’ states are particularly useful. Cat states are equal superpositions of two maximally different quantum states. They are a fundamental resource in fault-tolerant quantum computing12 and quantum communication, where they can enable protocols such as open-destination teleportation1 and secret sharing2. They play a role in fundamental tests of quantum mechanics3 and enable improved signal-to-noise ratios in interferometry4. Cat states are very sensitive to decoherence, and as a result their preparation is challenging and can serve as a demonstration of good quantum control. Here we report the creation of cat states of up to six atomic qubits. Each qubit’s state space is defined by two hyperfine ground states of a beryllium ion; the cat state corresponds to an entangled equal superposition of all the atoms in one hyperfine state and all atoms in the other hyperfine state. In our experiments, the cat states are prepared in a three-step process, irrespective of the number of entangled atoms. Together with entangled states of a different class created in Innsbruck5, this work represents the current state-of-the-art for large entangled states in any qubit system.

One promising candidate system for scalable universal quantum information processing (QIP) consists of atomic ions that are confined in electromagnetic traps and manipulated with laser beams5. Most of the basic ingredients for QIP10 have been demonstrated separately in the last few years in this system. Furthermore, some simple algorithms that could serve as primitives for larger scale QIP, including quantum error correction, teleportation, and the semiclassical quantum Fourier transform, have been implemented in the atomic-ion system.

Before large-scale QIP with atomic ions can become a reality, several challenges must be met successfully. In addition to building larger trap arrays and improving the classical control systems, it is necessary to demonstrate the ability to reliably create and maintain highly entangled states of many ions. Among such states, the so-called cat states, named after Schrödinger’s cat1, are of particular interest. Cat states are equal superpositions of two maximally different states (see below) and play a distinguished role in the atomic-ion system.

Before large-scale QIP with atomic ions can become a reality, several challenges must be met successfully. In addition to building larger trap arrays and improving the classical control systems, it is necessary to demonstrate the ability to reliably create and maintain highly entangled states of many ions. Among such states, the so-called cat states, named after Schrödinger’s cat1, are of particular interest. Cat states are equal superpositions of two maximally different states (see below) and play a distinguished role in the atomic-ion system.

Before large-scale QIP with atomic ions can become a reality, several challenges must be met successfully. In addition to building larger trap arrays and improving the classical control systems, it is necessary to demonstrate the ability to reliably create and maintain highly entangled states of many ions. Among such states, the so-called cat states, named after Schrödinger’s cat1, are of particular interest. Cat states are equal superpositions of two maximally different states (see below) and play a distinguished role in the atomic-ion system.

Before large-scale QIP with atomic ions can become a reality, several challenges must be met successfully. In addition to building larger trap arrays and improving the classical control systems, it is necessary to demonstrate the ability to reliably create and maintain highly entangled states of many ions. Among such states, the so-called cat states, named after Schrödinger’s cat1, are of particular interest. Cat states are equal superpositions of two maximally different states (see below) and play a distinguished role in the atomic-ion system.

Before large-scale QIP with atomic ions can become a reality, several challenges must be met successfully. In addition to building larger trap arrays and improving the classical control systems, it is necessary to demonstrate the ability to reliably create and maintain highly entangled states of many ions. Among such states, the so-called cat states, named after Schrödinger’s cat1, are of particular interest. Cat states are equal superpositions of two maximally different states (see below) and play a distinguished role in the atomic-ion system.

Before large-scale QIP with atomic ions can become a reality, several challenges must be met successfully. In addition to building larger trap arrays and improving the classical control systems, it is necessary to demonstrate the ability to reliably create and maintain highly entangled states of many ions. Among such states, the so-called cat states, named after Schrödinger’s cat1, are of particular interest. Cat states are equal superpositions of two maximally different states (see below) and play a distinguished role in the atomic-ion system.

Before large-scale QIP with atomic ions can become a reality, several challenges must be met successfully. In addition to building larger trap arrays and improving the classical control systems, it is necessary to demonstrate the ability to reliably create and maintain highly entangled states of many ions. Among such states, the so-called cat states, named after Schrödinger’s cat1, are of particular interest. Cat states are equal superpositions of two maximally different states (see below) and play a distinguished role in the atomic-ion system.

Before large-scale QIP with atomic ions can become a reality, several challenges must be met successfully. In addition to building larger trap arrays and improving the classical control systems, it is necessary to demonstrate the ability to reliably create and maintain highly entangled states of many ions. Among such states, the so-called cat states, named after Schrödinger’s cat1, are of particular interest. Cat states are equal superpositions of two maximally different states (see below) and play a distinguished role in the atomic-ion system.

Before large-scale QIP with atomic ions can become a reality, several challenges must be met successfully. In addition to building larger trap arrays and improving the classical control systems, it is necessary to demonstrate the ability to reliably create and maintain highly entangled states of many ions. Among such states, the so-called cat states, named after Schrödinger’s cat1, are of particular interest. Cat states are equal superpositions of two maximally different states (see below) and play a distinguished role in the atomic-ion system.

Before large-scale QIP with atomic ions can become a reality, several challenges must be met successfully. In addition to building larger trap arrays and improving the classical control systems, it is necessary to demonstrate the ability to reliably create and maintain highly entangled states of many ions. Among such states, the so-called cat states, named after Schrödinger’s cat1, are of particular interest. Cat states are equal superpositions of two maximally different states (see below) and play a distinguished role in the atomic-ion system.

Before large-scale QIP with atomic ions can become a reality, several challenges must be met successfully. In addition to building larger trap arrays and improving the classical control systems, it is necessary to demonstrate the ability to reliably create and maintain highly entangled states of many ions. Among such states, the so-called cat states, named after Schrödinger’s cat1, are of particular interest. Cat states are equal superpositions of two maximally different states (see below) and play a distinguished role in the atomic-ion system.

Before large-scale QIP with atomic ions can become a reality, several challenges must be met successfully. In addition to building larger trap arrays and improving the classical control systems, it is necessary to demonstrate the ability to reliably create and maintain highly entangled states of many ions. Among such states, the so-called cat states, named after Schrödinger’s cat1, are of particular interest. Cat states are equal superpositions of two maximally different states (see below) and play a distinguished role in the atomic-ion system.

Before large-scale QIP with atomic ions can become a reality, several challenges must be met successfully. In addition to building larger trap arrays and improving the classical control systems, it is necessary to demonstrate the ability to reliably create and maintain highly entangled states of many ions. Among such states, the so-called cat states, named after Schrödinger’s cat1, are of particular interest. Cat states are equal superpositions of two maximally different states (see below) and play a distinguished role in the atomic-ion system.

Before large-scale QIP with atomic ions can become a reality, several challenges must be met successfully. In addition to building larger trap arrays and improving the classical control systems, it is necessary to demonstrate the ability to reliably create and maintain highly entangled states of many ions. Among such states, the so-called cat states, named after Schrödinger’s cat1, are of particular interest. Cat states are equal superpositions of two maximally different states (see below) and play a distinguished role in the atomic-ion system.

Before large-scale QIP with atomic ions can become a reality, several challenges must be met successfully. In addition to building larger trap arrays and improving the classical control systems, it is necessary to demonstrate the ability to reliably create and maintain highly entangled states of many ions. Among such states, the so-called cat states, named after Schrödinger’s cat1, are of particular interest. Cat states are equal superpositions of two maximally different states (see below) and play a distinguished role in the atomic-ion system.

Before large-scale QIP with atomic ions can become a reality, several challenges must be met successfully. In addition to building larger trap arrays and improving the classical control systems, it is necessary to demonstrate the ability to reliably create and maintain highly entangled states of many ions. Among such states, the so-called cat states, named after Schrödinger’s cat1, are of particular interest. Cat states are equal superpositions of two maximally different states (see below) and play a distinguished role in the atomic-ion system.

Before large-scale QIP with atomic ions can become a reality, several challenges must be met successfully. In addition to building larger trap arrays and improving the classical control systems, it is necessary to demonstrate the ability to reliably create and maintain highly entangled states of many ions. Among such states, the so-called cat states, named after Schrödinger’s cat1, are of particular interest. Cat states are equal superpositions of two maximally different states (see below) and play a distinguished role in the atomic-ion system.
state to another is known as ‘local operations and classical communication’ (LOCC)\(^\text{17}\). For \(N > 2\) there are \(N\)-particle entangled states that cannot be transformed into each other, even if \(|\psi\rangle\) need only be obtained from \(|\phi\rangle\) with non-zero probability of success\(^\text{18}\). This implies that such states belong to different entanglement classes and that it is not possible to compare the amount of entanglement by using LOCC transformations. One of the classes with genuine \(N\)-particle entanglement is characterized by the cat states.

Establishing genuine \(N\)-particle entanglement in the experiment requires making measurements that clearly distinguish the produced state from any incompletely entangled state. One approach is based on so-called entanglement witness operators\(^\text{19}\). If an entanglement witness operator has a negative expectation value for a state, then that state is definitely \(N\)-particle entangled. Thus, one way to experimentally determine the presence of entanglement is to measure the expectation value of a well-chosen witness and show that it is negative with sufficient statistical significance. \(N\)-particle entanglement of cat states can be proved with a particularly simple witness operator (derived from ref. 15) based on the projector onto the ideal state \(|N\text{ Cat}\rangle\): 

\[
W = 1 - 2|N\text{ Cat}\rangle\langle N\text{ Cat}| 
\]

The expectation value of this operator is directly related to the fidelity in equation (3):

\[
\langle W \rangle = 1 - 2F_{N\text{Cat}} \leq 1 - 4|G_{N|N}| 
\]

A nice feature of this entanglement measure is that it is negative, then copies of \(|\Psi_0\rangle\) can be purified by LOCC to nearly pure cat states by means of a simple and robust purification procedure\(^\text{19}\).

Another strategy for proving \(N\)-particle entanglement is based on a ‘depolarization’ method. Using LOCC operations, the original density matrix can be transformed into a standard, partially depolarized form in which \(\rho\) can be transformed into a standard, partially depolarized form: 

\[
\rho = \frac{1}{2} \sum_{i,j} G_{ij} |i\rangle \langle j| + \frac{1}{2} \sum_{i,j} W_{ij} |i\rangle \langle j| 
\]

where \(G_{ij}\) are the elements of the density matrix, \(W_{ij}\) are the elements of the depolarization matrix, and \(G_{ij} + W_{ij} = 1\). The depolarization matrix \(W_{ij}\) is a rotation around the quantization axis. We take \(G_{ij}\) as a sequence of \(N\) ions in state \(|\uparrow\rangle\) or \(|\downarrow\rangle\), not all states equal, and \(|\uparrow\rangle\) corresponds to \(|\uparrow\rangle\) with \(|\downarrow\rangle\) interchanged. The quantities \(G_{N|N}\) and \(P_j + P_j\) are not changed by the depolarization method. Consequently, we can obtain them directly from observations of the prepared state without actually implementing the depolarization. Inequality (7) can be satisfied even if \(F < 0.5\), but the purification process to obtain nearly pure cat states with LOCC from multiple copies may no longer be simple. For our experimental six-qubit cat state (below), since \(W\) was found to be only slightly negative, we used the depolarization method to conclusively establish that it was six-qubit entangled.

By extending methods used previously to create two- and three-qubit entangled states\(^\text{14}\), we entangled up to six ions in states that approximate \(N\)-qubit cat states. We confined \(^{133}\text{Cs}\) ions to the axis of a linear Paul trap with axial centre-of-mass (COM) frequencies between \(\omega_{\text{COM}}(2\pi) = 2.6\) MHz and \(\omega_{\text{COM}}(2\pi) = 3.4\) MHz and radial COM frequencies of approximately 8 MHz (ref. 22). All \(N\) axial motional modes of the ions were cooled to the ground state by extending the method of ref. 23. We prepared the internal state of each ion in the \(|F = 2, m_F = -2\rangle = |\uparrow\rangle\) hyperfine ground state by optical pumping, where \(F\) and \(m_F\) are the total angular momentum and the component of the angular momentum along the quantization axis. We take \(|F = 1, m_F = -1\rangle = |\downarrow\rangle\) as the other qubit state. The encoding operations \(U_N\) were realized using two-photon stimulated Raman transitions uniformly applied to all ions, incorporating a phase gate \(G_{\chi_N}\) that is an extension of the gate described in ref. 14. The phase gate was implemented by two laser beams that uniformly illuminate all ions with a relative detuning \(\omega_{\text{COM}} + \delta\) with \(\delta \ll \omega_{\text{COM}}\), exerting a state-dependent axial optical dipole force on the ions\(^\text{14}\). The spacing of the ions was chosen such that the force was proportional to \(|F_f\rangle\) (see Methods). If this dipole force is applied for a duration \(t_g = 2\pi/\delta\), the motion of the COM mode is excited and de-excited in such a way that each state on which a non-zero net force acts \((|F_f\rangle \neq 0)\) traverses a circle in phase space\(^\text{18}\) and acquires a phase given by the area circumscribed in phase space. This area is proportional to the square of the net force and therefore to \(|\langle F_f \rangle|\) for that state. In the experiment, the strength and detuning were adjusted to yield a phase of \(\frac{\pi}{2}\) on each component of the wavefunction, thus realizing the third (middle) operator in equation (2). For \(N = 5\) (and therefore \(\xi = 1\)) the left operator (up to trivial phases) was realized by an appropriate change of the final pulse: 

\[
\exp \left[ i\frac{\pi}{2}J_3 \right] \exp \left[ i\frac{\pi}{2}J_3 \right] \rightarrow \exp \left[ i\frac{\pi}{2}J_3 \right].
\]

After creating each cat state, we determined the populations in substates with equal numbers of \(|\uparrow\rangle\) components by observing state-dependent fluorescence (see Fig. 1 and Methods). The most important information on the quality of the states resides in the magnitude of coherence, \(G_{N|N}\). A lower bound on this quantity can be extracted by executing the interferometry algorithm described in ref. 7 on the states. Here \(|\Psi_0\rangle\) is ‘decoded’ by applying the operation \(U_{N\phi}\) to it. \(U_{N\phi}\) differs from \(U_N\) only by replacing the two \(J_3\) operations by the rotations \(|\phi\rangle = \cos(\phi) |\uparrow\rangle + \sin(\phi) |\downarrow\rangle\). Note that the effect of applying \(U_{N\phi}\) is equivalent to applying \(\exp\left[-iJ_3\phi U_N\right]\exp\left[iJ_3\phi\right]\). Since this is followed by measurement, the final \(\exp\left[-iJ_3\phi\right]\) has no observable effect. The first operator, \(\exp\left[iJ_3\phi\right]\), is a rotation around the z-axis of the Bloch sphere and uniquely ‘labels’ the coherence between \(|\uparrow, N\rangle\) and \(|\downarrow, N\rangle\) with a phase that evolves at \(N\) times \(\phi\) (ref. 24). The net effect is that \(U_{N\phi}\) transfers the coherence into a population difference. Ideally:

\[
|\Psi_0\rangle = U_{N\phi}|N\text{ Cat}\rangle = -i \sin \left( \frac{N}{2} \phi \right) |\uparrow, N\rangle + \cos \left( \frac{N}{2} \phi \right) |\downarrow, N\rangle
\]

Figure 1 Histogram and residuals of the \(|6\text{Cat}\rangle\). Upper plot, the experimental histogram is fitted to a sum of seven poissonian distributions with mean values corresponding to 0, 1, 2, …, 6 ions in state \(|\downarrow\rangle\). The fit yields the populations \(P_N\). Number of counts per 200 µs

© 2005 Nature Publishing Group

640
In the experiment we observed a fluorescence signal, which has a contribution that oscillates N-times sinusoidally as \( \phi \) is ramped through \( 2\pi \). This oscillation can only arise from the coherence between \( |1\rangle \) and \( |N\rangle \) between applications of \( U_N \) and \( U_{N-1} \). It has a maximal amplitude \( A_{\text{max}} \) determined by the difference in the fluorescence signal with all ions in state \( |1\rangle \) and the signal with all ions in state \( |N\rangle \), which is achieved if all the operations are implemented perfectly. In the imperfect case, the actual amplitude of the oscillation \( A \) yields a lower bound on \( |C_{\text{ion}}| \) via the inequality \( |C_{\text{ion}}| \geq \frac{A}{A_{\text{max}}} \). Because the imperfections in \( U_{N-1} \) match those of \( U_N \), up to the high accuracy with which pulse phases are controlled, this bound is overly pessimistic. However, it cannot be improved without introducing additional assumptions.

Figure 2 shows the fluorescence as a function of phase \( \phi \) for 4, 5 and 6 ions obtained by the method of the previous paragraph. As the number of ions increases, the coherence is more strongly affected by several sources of imperfection. Most importantly, decoherence due to spontaneous emission increases in proportion to the number of ions and the duration of the encoding operation \( U_N \) (for six ions we estimate an 18% probability of spontaneous emission per gate). In addition, the susceptibility to magnetic field noise, which washes out the fluorescence signal with all ions in state \( |j\rangle \), is ramped over a range of \( 2\pi \); up to the high accuracy with which pulse phases are controlled, this bound is overly pessimistic. However, it cannot be improved without introducing additional assumptions.

Figure 2 shows the fluorescence as a function of phase \( \phi \) for 4, 5 and 6 ions obtained by the method of the previous paragraph. As the number of ions increases, the coherence is more strongly affected by several sources of imperfection. Most importantly, decoherence due to spontaneous emission increases in proportion to the number of ions and the duration of the encoding operation \( U_N \) (for six ions we estimate an 18% probability of spontaneous emission per gate). In addition, the susceptibility to magnetic field noise, which washes out the fluorescence signal with all ions in state \( |j\rangle \), is ramped over a range of \( 2\pi \); up to the high accuracy with which pulse phases are controlled, this bound is overly pessimistic. However, it cannot be improved without introducing additional assumptions.

Figure 2 shows the fluorescence as a function of phase \( \phi \) for 4, 5 and 6 ions obtained by the method of the previous paragraph. As the number of ions increases, the coherence is more strongly affected by several sources of imperfection. Most importantly, decoherence due to spontaneous emission increases in proportion to the number of ions and the duration of the encoding operation \( U_N \) (for six ions we estimate an 18% probability of spontaneous emission per gate). In addition, the susceptibility to magnetic field noise, which washes out the fluorescence signal with all ions in state \( |j\rangle \), is ramped over a range of \( 2\pi \); up to the high accuracy with which pulse phases are controlled, this bound is overly pessimistic. However, it cannot be improved without introducing additional assumptions.

Figure 2 shows the fluorescence as a function of phase \( \phi \) for 4, 5 and 6 ions obtained by the method of the previous paragraph. As the number of ions increases, the coherence is more strongly affected by several sources of imperfection. Most importantly, decoherence due to spontaneous emission increases in proportion to the number of ions and the duration of the encoding operation \( U_N \) (for six ions we estimate an 18% probability of spontaneous emission per gate). In addition, the susceptibility to magnetic field noise, which washes out the fluorescence signal with all ions in state \( |j\rangle \), is ramped over a range of \( 2\pi \); up to the high accuracy with which pulse phases are controlled, this bound is overly pessimistic. However, it cannot be improved without introducing additional assumptions.

Figure 2 shows the fluorescence as a function of phase \( \phi \) for 4, 5 and 6 ions obtained by the method of the previous paragraph. As the number of ions increases, the coherence is more strongly affected by several sources of imperfection. Most importantly, decoherence due to spontaneous emission increases in proportion to the number of ions and the duration of the encoding operation \( U_N \) (for six ions we estimate an 18% probability of spontaneous emission per gate). In addition, the susceptibility to magnetic field noise, which washes out the fluorescence signal with all ions in state \( |j\rangle \), is ramped over a range of \( 2\pi \); up to the high accuracy with which pulse phases are controlled, this bound is overly pessimistic. However, it cannot be improved without introducing additional assumptions.

Figure 2 shows the fluorescence as a function of phase \( \phi \) for 4, 5 and 6 ions obtained by the method of the previous paragraph. As the number of ions increases, the coherence is more strongly affected by several sources of imperfection. Most importantly, decoherence due to spontaneous emission increases in proportion to the number of ions and the duration of the encoding operation \( U_N \) (for six ions we estimate an 18% probability of spontaneous emission per gate). In addition, the susceptibility to magnetic field noise, which washes out the fluorescence signal with all ions in state \( |j\rangle \), is ramped over a range of \( 2\pi \); up to the high accuracy with which pulse phases are controlled, this bound is overly pessimistic. However, it cannot be improved without introducing additional assumptions.

Figure 2 shows the fluorescence as a function of phase \( \phi \) for 4, 5 and 6 ions obtained by the method of the previous paragraph. As the number of ions increases, the coherence is more strongly affected by several sources of imperfection. Most importantly, decoherence due to spontaneous emission increases in proportion to the number of ions and the duration of the encoding operation \( U_N \) (for six ions we estimate an 18% probability of spontaneous emission per gate). In addition, the susceptibility to magnetic field noise, which washes out the fluorescence signal with all ions in state \( |j\rangle \), is ramped over a range of \( 2\pi \); up to the high accuracy with which pulse phases are controlled, this bound is overly pessimistic. However, it cannot be improved without introducing additional assumptions.
alignment direction of the ions. The inter-ion distance is determined by the force equilibrium between mutual Coulomb repulsion of the ions and the confining force of the external trap potential, causing inter-ion distances to be unequal for \( N > 3 \). Nevertheless, a spacing where the dipole force has approximately the same phase for each ion can be achieved. As an example, the positions of four ions relative to the trap centre are \( \pm 1.437, -0.454, 0.454, 1.437 \), where \( s = \sqrt{s^2/(4\pi e_0 m_0 w_{\text{COM}}^2)} \) is a universal scaling parameter, with \( e \) the elementary charge, \( e_0 \) the vacuum permeability, \( m_0 \) the mass of the beryllium ion and \( w_{\text{COM}} \) the axial COM frequency. We can take advantage of the fact that the two different distances have a ratio very close to an integer ratio, \( 1.437/0.454 \approx 3 \), to observe interference.

Determination of populations from state dependent fluorescence. During one detection period (duration 200 \( \mu s \)) we typically detect on average \( \bar{N}_0 = 0.5 \) counts if all ions are projected into | 1 \rangle, and about \( \bar{N}_1 = 10 \) additional counts for each ion in state | 2 \rangle. The parameters \( \bar{N}_0 \) and \( \bar{N}_1 \) were derived by fitting mixtures of poissonian distributions to reference photon-count histograms obtained by running many experiments for each of a small number of states including | 1 \rangle, | 2 \rangle, | 3 \rangle. We used the maximum likelihood method for fitting the histograms and parametric-bootstrap resampling for determining standard errors in inferred quantities25. Under the assumption that each ion fluoresces equally, the histograms should be well approximated by \( \bar{N}_0 \text{Pois}(\lambda) + \bar{N}_1 \text{Pois} (\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4) \), where \( \lambda \) is the number of experiments contributing to the histogram and \( \text{Pois}(\lambda) \) is a poissonian distribution with mean \( \lambda \). The parameters \( \lambda_0 \) and \( \lambda_1 \) determine the maximum possible amplitude of the phase oscillations in Fig. 2. This maximum amplitude is required for inferring a lower bound on \( |E|_{\text{DIST}} \). The probabilities \( P_0, \ldots , P_4 \) were obtained by a maximum likelihood estimation of their values based on experimental population histograms obtained by direct observation of the prepared cat states. Up to 39,900 experiments were used to acquire these histograms. The desired witness expectations were computed according to the fits. From the populations \( P_0 \), we determine the quantity \( 1/2(P_0 + P_N) = 1/2(P_0 + P_N) \), the first term in equation (4). From the remaining populations, we can find upper bounds on \( P_0 \). For example (\( N = 4 \)), \( P_{\text{Pois}} \leq P_0 \). Figure 1 shows the measured histogram for the six-ion cat state together with the residuals between data and the fitted distribution. The histograms and fits for four and five ions look similar. From \( P_j \), expectations were computed according to the fits. From the populations \( P_j \), we can find upper bounds on \( P_{\text{Pois}} \).

Received 27 July; accepted 20 September 2005.


Acknowledgements This work was supported by the US National Security Agency (NSA) and Advanced Research and Development Activity (ARDA), by the Department of Defence Multidisciplinary University Research Initiative (MURI) programme administered by the Office of Naval Research and by NIST. We thank S. Glancy and J. Bollinger for comments on the manuscript. This paper is a contribution by the National Institute of Standards and Technology and not subject to US copyright.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to D.L. (dill@boulder.nist.gov).


182 BIBLIOGRAPHY


