ABSTRACT

Title of Dissertation: SCALING QUANTUM COMPUTERS WITH LONG CHAINS OF TRAPPED IONS

Laird Nicholas Egan
Doctor of Philosophy, 2021

Dissertation Directed by: Professor Christopher Monroe
Department of Physics
Joint Quantum Institute

Quantum computers promise to solve models of important physical processes, optimize complex cost functions, and challenge cryptography in ways that are intractable using current computers. In order to achieve these promises, quantum computers must both increase in size and decrease error rates.

To increase the system size, we report on the design, construction, and operation of an integrated trapped ion quantum computer consisting of a chain of 15 $^{171}Yb^{+}$ ions with all-to-all connectivity and high-fidelity gate operations. In the process, we identify a physical mechanism that adversely affects gate fidelity in long ion chains. Residual heating of the ions from noisy electric fields creates decoherence due to the weak confinement of the ions transverse to a focused addressing laser. We demonstrate this effect in chains of up to 25 ions and present a model that accurately describes the observed decoherence. To mitigate this noise source, we first propose a new sympathetic cooling scheme to periodically re-cool the ions throughout a quantum circuit, and then demonstrate its capability in a proof-of-concept

experiment.

One path to suppress error rates in quantum computers is through quantum error correction schemes that combine multiple physical qubits into logical qubits that robustly store information within an entangled state. These extra degrees of freedom enable the detection and correction of errors. Fault-tolerant circuits contain the spread of errors while operating the logical qubit and are essential for realizing error suppression in practice. We demonstrate fault-tolerant preparation, measurement, rotation, and stabilizer measurement of a distance-3 Bacon-Shor logical qubit in our quantum computer. The result is an encoded logical qubit with error rates lower than the error of the entangling operations required to operate it.

Scaling Quantum Computers with Long Chains of Trapped Ions

by

Laird Nicholas Egan

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2021

Advisory Committee:
Professor Christopher Monroe, Chair/Advisor
Professor Norbert Linke
Professor Alicia Kollár
Professor Vladimir Manucharyan
Professor Christopher Jarzynski

Acknowledgments

I will start by acknowledging that I was not a typical physics graduate student, since my previous 5 years were spent in industry working for a defense contractor. It is not lightly that one makes the decision to leave a well-paying career to pursue cutting-edge academic research with its long hours and often thankless work. I am very thankful to my co-workers Dave Coar, Manny Stockman, and Brian Holloway, who helped me to kick the idea around my head, set expectations on the experience, and have remained in touch throughout my entire journey.

On the other side of the fence, I am very thankful to my advisor, Chris Monroe, who took a bit of a risk on my non-traditional resume, unsure of how I would perform in an academic setting. He still welcomed me with open arms even after I told him that I wanted to graduate in 4 years and then immediately return to industry. Throughout the entire experience, he has always been supportive of my ambitions and an amazing resource for physics and career advice.

One of my goals when starting this journey, was that I wanted to be surrounded by the best and the brightest in the field. I found it in Marko Cetina, who led the experiment with wisdom, insight, and tenacity from when I arrived to the day I started writing this thesis. I also had the opportunity to work alongside several other extremely talented and creative post-docs in Kristi Beck, Michael Goldman, and Crystal Noel. I am especially grateful for Kristi, who took me under her wing when I first joined and put up with a lot of stupid questions while I learned the ropes. I also had the pleasure of spending many days in the lab with Drew Risinger,

Daiwei Zhu, Debopriyo Biswas, and Bahaa Harraz. Physics is a team sport and I could not have done it without your skills and good humor. I would also like to thank Jonathan Mizrahi, Kai Hudek, and Jason Amini, who designed the system before I arrived and encapsulated the proverb of planting trees in whose shade they shall never sit. Finally, I need to thank all of the remaining members of Chris's lab, of which there are too many to enumerate, but who I shared many great conversations and friendships with over lunch or the espresso machine.

Perhaps the most rewarding experience throughout my journey was the amazing collaboration I had with Ken Brown's group on our joint quantum error-correction results. Far be it from the head in the clouds stereotype of theorists, Ken was always willing to jump into the nitty-gritty of our machine and to adjust his ideas to meet us half-way, often resulting in spectacular data. I also had the pleasure of working closely with one of Ken's students, Dripto Debroy, who I shared the pain of writing multiple papers with and made calculating IARPA benchmarks an enjoyable experience. The rest of the theory team at Duke, including Michael Newman, Muyuan Li, James Leung, and the experimental team of Jungsang Kim, including Stephen Crain and Ye Wang were critical resources and fruitful partners on the IARPA LogiQ program.

My family has been an amazing support structure throughout this journey. I cannot articulate how grateful I am to my parents Dave and Tina, for instilling an intellectual curiosity in me and for enthusiastically following every development in my adventure. My brother, Garth, paved the way for doctoral degrees at UMD and was a great source of commiseration. My sister, Tessa, is an unbounded source of

inspiration and joy. I was very fortunate to pursue my degree near my in-laws, Jim, Amy, and Grace who provided many hot meals while I worked late nights and were always eager to hear another explanation of quantum computing.

Saving the best for last, I could not have done this without my incredible wife Liz. She was fully on-board with the concept that we may have to move across the country and find her a new job, just so I could take a pay-cut and work long hours. Her support has never wavered from that moment. Thank you for walking next to me on this fantastic journey we call life.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AC	Alternating Current (can also describe laser electric fields)
ADC	Analog-to-Digital Converter
AFM	Anti-Ferromagnetic
AM	Amplitude Modulation
AOM	Acousto-Optic Modulator
AWG	Arbitrary Waveform Generator
BSB	Blue Sideband
CMOS	Complementary Metal-Oxide-Semiconductor
CNOT	Controlled-Not
CW	Continuous Wave
DAC	Digital-to-Analog Converter
DC	Direct Current
DDS	Direct Digital Synthesis
DoF	Degree of Freedom
ECDL	External Cavity Diode Laser
EOM	Electro-Optic Modulator
FM	Frequency Modulation or Ferromagnetic (context dependant)
FPGA	Field-programmable Gate Array
FT	Fault-Tolerant
GHZ	Greenberger–Horne–Zeilinger
HF	Hyperfine
IR	Infrared
LO	Local Oscillator
MEMS	Microelectromechanical System
MS	Mølmer-Sørensen
MW	Microwave
NA	Numerical Aperture
nFT	non-Fault-Tolerant
NLTL	Non-linear Transmission Line
PDH	Pound-Drever-Hall
PM	Phase Modulation
PMT	Photo-Multiplier Tube
QCCD	Quantum Charge-Coupled Device
QHO	Quantum Harmonic Oscillator
RF	Radio Frequency
RFSoC	Radio Frequency System-on-Chip
RSB	Red Sideband
RMS	Root Mean Sum
RWA	Rotating Wave Approximations

SBC	Sideband Cooling
SBR	Saturable Bragg Reflector
SFG	Sum Frequency Generation
SHG	Second Harmonic Generation
SNR	Signal-to-Noise Ratio
SPAM	State Preparation and Measurement
SVD	Singular Value Decomposition
UHV	Ultra High Vacuum
UV	Ultra-Violet

Chapter 1: Introduction

1.1: Quantum Computers

Quantum computers offer a completely different computing paradigm than classical computers. Harnessing the fundamental properties of quantum mechanics, quantum computers promise to solve certain classes of problems that are impossible or intractable with modern classical computers. In some sense, quantum computers are unavoidable, as the Moore's law scaling of classical transistors eventually envisions that bits are reduced to the size of atoms, where quantum mechanics becomes the dominating principle. One of the first envisioned applications of quantum computers is to simulate and solve important physical processes that operate under the laws of quantum mechanics, including molecular structure and chemical reaction mechanisms [1, 2, 3, 4]. The key idea is that because nature is fundamentally quantum, a quantum computer will be more efficient at modeling these systems.

In the 90's, Peter Shor developed an algorithm for factoring large numbers using a quantum computer that is exponentially faster than its classical counterpart [5]. The computational difficulty of factoring large numbers is a key pillar of modern cryptography and thus threatened by the development of quantum comput-

ers. This algorithm expanded the idea that quantum computers may be useful in a more general sense for solving broad numerical problems. While Shor's algorithm is still many years from practical implementation, there is hope in the near-term that quantum computers will be able to solve other interesting numerical problems, such as optimizing complex cost functions, at speeds that exceed the ability of classical computers [6].

The advantages of quantum computers can be summarized by two main principles - quantum *superposition* and quantum *entanglement*. Classical computers operate on *bits*, which are like a switches that only exist in two possible states - either 0 (off) or 1 (on). Quantum bits, or *qubits*, can also take on definite states $|0\rangle$ or $|1\rangle$, but can also exist in a quantum superposition of these two states. A general quantum state $|\psi\rangle$ is described by,

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

where α and β are two complex numbers required to define the qubit state. For two qubits, four complex numbers are required, and for N qubits, 2^N numbers are required to fully describe the system. This exponential scaling is essential to the power of quantum computers. For example, to simulate a 100-qubit quantum computer using a classical computer would take at least 100,000 Yotta-bits (10^{24}) of classical data and possibly many more since in general α, β are continuous!

However, superpositions are fragile and are destroyed when the qubit is measured, returning $|0\rangle$ with probability $|\alpha|^2$ or $|1\rangle$ with probability $|\beta|^2$. Thus, even

though all possible 2^N quantum states can be simultaneously created and operated on, the measurement only supplies an N-bit output string, sampled randomly from total quantum state. This appears to be a bottleneck in quantum computers, however it can be circumvented by applying the second principle of quantum entanglement, or more generally quantum interference.

Quantum algorithms rely on the ability to interfere different quantum states together so that the undesired output states are deconstructively suppressed and the desired output states are constructively enhanced. If the final superposition of states is much less than 2^N, then the result can be efficiently extracted. As an extreme example, the Bernstein–Vazirani algorithm only requires one query of the quantum computer in order to obtain the output of the quantum algorithm [7]. Much of the art of quantum algorithm design is figuring out how to encode problems into quantum operations that create the desired interference patterns.

Quantum interference often results in the creation of entangled states. Entangled states have no classical analogue and led Einstein to conclude the quantum mechanical description of reality was incomplete [8], although so far experiments have not supported this position (see Bell's inequality [9]). A prototypical entangled state is written as,

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right)$$

If we only measure the first qubit and we obtain $|0\rangle$, then we know the second qubit must also be in $|0\rangle$, even though we have not measured it. Stranger still, if we apply a phase to the first qubit $|1\rangle \rightarrow e^{i\phi}|1\rangle$, then this phase is shared equally amongst

the two qubits, even if they are separated far apart. Quantum entanglement is a key computational resource available only to quantum computers and is a critical component of quantum algorithms.

Clearly, quantum computers do not simply offer a simple speed up over classical computers. Rather they open up an entirely new computational space that we are just beginning to explore. While there are already compelling reasons to build a quantum computer, we hope that some of the best applications are yet to come. The best way to develop new uses for quantum computers is to actually build them and put them in the hands of users. To cite a relevant historical comparison - no one could have imagined ubiquity of modern classical computers when the ENIAC was first being built to calculate artillery firing tables.

1.2: Trapped Ion Quantum Computers

There are several competing qubit technologies that all show potential to eventually create large-scale quantum computers. These technologies include trapped ions, superconducting circuits [10], neutral atoms [11], NV centers [12] and photonic systems [13]. In the future, other qubit technologies may also enter the race, such as topological qubits [14]. Spurred by recent private and commercial investment, ion traps and superconducting circuits have emerged as two of the most mature technologies.

Despite their successes, the two platforms have very different challenges moving forward. Superconducting qubits suffer from qubit inhomogeneity due to imperfect

fabrication control. On the other hand, atomic qubits are defined by fundamental constants of nature and are perfectly replicable. Superconducting qubits are fixed in place and hardwired with control lines and thus fundamentally stable, whereas the "wires" in trapped ion systems are focused laser beams, making the wire connectivity more flexible but also susceptible to misalignment. The most gainful path forward in scaling up superconducting system is improving the qubit itself. For trapped ion systems, the qubit is nearly perfect as is, with the primarily challenge being the "classical" control associated with positioning and transport of ions relative to some fixed number of laser beam "wires."

Trapped ions have long since achieved all of Divicenzo's criteria that were laid out as requirements for the implementation of a quantum computer [15]. The remaining challenge is to scale up the performance of the system to solve interesting and useful applications. The primary long-term vision for large scale ion trap quantum computers is to have a chip with multiple separate ion chains of length N. Entangling operations may be performed within a single chain through the shared motional bus, but then the chain may be broken up and ions may by shuttled around the trap and merged into a different chains as the quantum algorithm progresses. This is known as the Quantum CCD architecture [16] and is likely to support 100-1000 total qubits. This architecture is also likely to require additional "sympathetic cooling" ions that can recool the data ions to quench any excess motion generated from shuttling operations [17]. Ultimately several QCCD chips in separate vacuum chambers will have to be photonically connected [18] in order to achieve systems with 10,000-100,000 qubits.

A significant open question in ion trap architecture is how large to make N. There is an obvious benefit to $N \leq 2$, in that the laser beams no longer need to be focused to address individual qubits. Large beams can be directed at "gate" zones within the chip, and one or two qubits can be shuttled into this zone to perform single- or two-qubit gates. However there is also significant overhead associated with maintaining all-to-all connectivity in these systems, as significant shuttling overhead is required to execute every time-slice in the circuit. For example, in a recent demonstration [19] of an $N = 2$ system running a quantum volume (QV = 64) circuit consisting of 54 entangling operations, 97% of the total length of execution time (130 ms) was spent re-cooling and shuttling ions.

For $N > 2$, ions now need to be individually addressed, which comes at a design complexity cost, but can also significantly reduce the overhead shuttling costs since the on-demand "random-access" connectivity of the system can be much higher. The same QV = 64 circuit would take only ≈ 12.5 ms in an $N = 6$ system [20] with no shuttling required. A useful analogy is size of the bit register in classical CPUs. While it is possible to implement any classical computation on a 1-bit processor (essentially a Turing Machine), we know that this is inefficient and will have data bottlenecks. There is a good reason why CPU manufacturers have been pushing to increase the size of the CPU register (now 64-bit), which allows the processor to randomly access more data and increase the efficiency of the computation. The central theme of this thesis is exploring the implications of making N as large as practically possible in our quantum processor.

1.3: Thesis Outline

In Chapter 2, we outline the basic techniques required to trap, cool, prepare and readout ions that are common to any $^{171}Yb^{+}$ system. We also describe the spacing and motional mode structure in long chains. In Chapter 3, we describe how the ions are controlled as qubits. We start with a simple model and work our way up to describe the Raman frequency comb used in this system. The theory behind single-qubit and two-qubits gate operations is explained, with a focus on how high-fidelity can be achieved in practice. In Chapter 4, we describe the hardware and engineering approach used to build our ion trap quantum computer. In Chapter 5, we describe how the machine is operated in practice, including how to perform the required calibrations. Additionally, we provide various metrics that benchmark the performance of the machine. In Chapter 6, we discuss the feasibility of further scaling up the chain length. We identify a detrimental physical mechanism that affects long chains and then develop a strategy to mitigate it. Finally, in Chapter 7, we demonstrate a quantum error correction code as a means of scaling trapped ion quantum computers in the long-term. We conclude in Chapter 8 with an outlook towards the future of trapped ion quantum computing.

Chapter 2: Trapped Ion Basics

Trapped ions are an excellent platform for quantum information processing due to their naturally long coherence times [21] and high fidelity state preparation and measurement [22]. Since each ion is a fundamental building block of nature, the qubits are perfectly reproducible, which is a large advantage when ultimately scaling to many 1000s of qubits. In this chapter we will discuss how to contain and trap atomic ions and outline the basic control required over the atomic level structure to perform core functions such as cooling, state preparation and measurement.

2.1: Ion Traps

The first step to working with ions is to trap them in a localized area so that they can be measured, cooled, and controlled. In general, there are two different types of traps that can be used to confine ion in three-dimensional wells - RF Paul traps and Penning traps. RF Paul traps we first developed by Wolfgang Paul in 1935 [23] and for which he shared the Nobel Prize in Physics in 1989 [24]. A linear Paul trap uses a static confining electric field in one dimension, and a dynamic RF electric field in the other two dimensions to trap an ion. As long as certain stability parameters are met (related to the mass/charge of the particle, the time scale of the

dynamic field, and the trap voltages and geometry), then the dynamic field creates an effective confining field in those two dimensions.

The other type of trap is Penning trap, which was first conceptualized by Frans Penning [25] and built by Hans Dehmelt, for which he shared the Nobel Prize in Physics in 1989 with Paul [26]. In contrast, these traps do not use dynamic fields, which may make them preferable for precision measurement experiments. Here, a static homogeneous magnetic field provides confinement of charged particles in two dimensions (via the Lorentz force) and an inhomogeneous static electric field provides confinement in the third. In this work, we use a Linear RF Paul Trap, since they readily support one-dimensional ion chains that can be individually addressed with lasers for quantum computations.

2.1.1: Linear RF Paul Trap

An ion is trapped if it experiences a confining force along all three Cartesian axes. One can write a general electric potential with quadrupole terms as,

$$\Phi(x, y, z) = \frac{U}{2}\left(\alpha x^2 + \beta y^2 + \gamma z^2\right) \tag{2.1}$$

Applying Gauss's Law leads to the constraint,

$$\nabla^2\Phi(x, y, z) = \alpha + \beta + \gamma = 0 \tag{2.2}$$

It is easy to see that at least one of $\{\alpha, \beta, \gamma\}$ must be < 0, or in other words, deconfining. This is the basis of Earnshaw's theorem, which states static potentials alone cannot confine point charges, resulting from the fact that the electric field at a point in free-space must be divergenceless. Now, consider adding an oscillating RF quadrupole potential in addition to the the static potential, described generally by

$$\Phi(x, y, z) = \frac{U}{2}\left(\alpha x^2 + \beta y^2 + \gamma z^2\right) + \frac{\widetilde{U}}{2}\cos(\Omega_{RF} t)\left(\alpha' x^2 + \beta' y^2 + \gamma' z^2\right) \tag{2.3}$$

Defining $\hat{x}$ to be along the axis of the trap, then Gauss's law can be satisfied by setting,

$$\begin{aligned} -(\beta + \gamma) = \alpha > 0 \\ \beta' = -\gamma', \alpha' = 0 \end{aligned} \tag{2.4}$$

This describes a static confining potential along $\hat{x}$ (the axial direction) and a dynamic inverting quadrupole potential in the $\hat{y}\hat{z}$-plane (the radial directions). To see how the dynamic field can generate confinement, consider just the dynamic potential along the $\hat{z}$ direction. The electric field is given at a location z is given by:

$$E_z(t) = V z \cos(\Omega_{RF} t) \tag{2.5}$$

If the oscillating field is fast enough so that the ion only moves a small distance over the period of one RF cycle, then the position of the ion can be decomposed into an equilibrium value z_0 and a part that harmonically oscillates about that point with

amplitude $A = eVz_0/m\Omega_{RF}^2$,

$$z(t) = z_0 + A\cos(\Omega_{RF}t) = z_0 + \frac{eVz_0}{m\Omega_{RF}^2}\cos(\Omega_{RF}t) \tag{2.6}$$

The electric force felt by this ion is,

$$\begin{aligned} F_z(t) &= -eE_z(t) \\ &= -eV\cos(\Omega_{RF}t)\, z(t) \\ &= -eV\cos(\Omega_{RF}t)\left(z_0 + \frac{eVz_0}{m\Omega_{RF}^2}\cos(\Omega_{RF}t)\right) \\ &= -eVz_0\cos(\Omega_{RF}t) - \frac{e^2V^2z_0}{m\Omega_{RF}^2}\cos(\Omega_{RF}t)^2 \end{aligned} \tag{2.7}$$

Averaging over one RF period,

$$\overline{F_z} = -\frac{e^2V^2}{2m\Omega_{RF}^2}z_0 \tag{2.8}$$

Thus integrating over the "fast-motion" of the ion reveals a restoring force relative to the "slow motion" of the ion (z_0). This arises from a ponderomotive force that, in general, causes charged particles to seek areas of minimum field amplitude in an inhomogeneous oscillating electric field. The slow-motion is also called the "secular motion" of the ion, which is also well approximated by harmonic motion

with ω called the secular frequency,

$$\overline{F_z} = -kz = -\omega_{sec}^2 mz$$
$$\omega_{sec} = \frac{eV}{\sqrt{2}m\Omega_{RF}} \tag{2.9}$$

We note that the motion of the ion along $\hat{y}$ can be treated in a similar manner. Importantly, these two directions do not couple to each other and are completely separable. For a more careful analysis of the RF + DC confinement in the radial dimensions, Equation 2.3 can be used to write equations of motion that, with proper substitutions, resemble equations of the form $\frac{d^2x}{d\xi^2} + [a - 2q\cos(2\xi)]\, x = 0$, which are well known class of functions called the Mathieu functions. Here a and q represent unitless parameters that determine the stability of the confinement (which was assumed in Eq. 2.6). While detailed analysis can be found in other resources [27], here we will simply present the solution to the Mathieu equation,

$$z(t) = [z_0 + \widetilde{z}\cos(\omega_z t)]\left[1 - \frac{q_z}{2}\cos(\Omega_{RF} t)\right] \tag{2.10}$$

Where z_0 indicates the ions equilibrium position offset from the RF null, $\widetilde{z}$ is the amplitude of the ions secular motion and $q_z = \frac{2eV}{mR^2\Omega^2}$ is a dimensionless Mathieu parameter approximately equal to $q_z \approx 0.25$ in this work. Relative to the simple treatment above, the Mathieu solution contains a term that describes motion at the bare RF frequency applied to the trap electrodes, known as the "micromotion" of the ion. When the ion at at the RF null ($z_0 = 0$) and $q_z \ll 1$, the motion of the ion

is well approximated by harmonic motion. However, if there are stray electric fields that displace the ion from the RF null ($z_0 \neq 0$), then the micromotion becomes magnified by the displacement and can quickly dominate the secular motion. This can cause Doppler shifts that reduce the effectiveness of cooling and state detection, as well as reducing the the appropriateness of the harmonic approximation. Since the harmonic approximation is crucial to the operation of our two-qubit entangling gates, one should always seek to minimize excess micromotion by applying compensation fields until $z_0 \approx 0$ (similarly along the other two axes). There are several techniques to measure micromotion [28].

2.1.2: Micro-fabricated Surface Traps

In practice, linear RF Paul traps can be realized rather simply by situating four conducting rods along the corners of a rectangle (situated in the yz-plane), with RF connected to one diagonal pair and the other pair shorted to ground at the relevant RF frequency (typically through a capacitor). Two end cap electrodes with a gap in between are situated along the x-axis and provide the static confinement field. This particular design is called a "four-rod" trap and is schematically shown in Fig. 2.1(a). Four-rod traps provide limited optical access but the design can be modified to use thin "blade" electrodes that provide higher numerical aperture [20]. The DC pair of blades can be segmented to provide coarse shaping of the static axial potential. Both four-rod and blade traps are examples of "macro-fabricated" traps. These traps have dimensions of a few inches, are typically assembled by hand, and may have non-repeatable trap characteristics that require calibration.

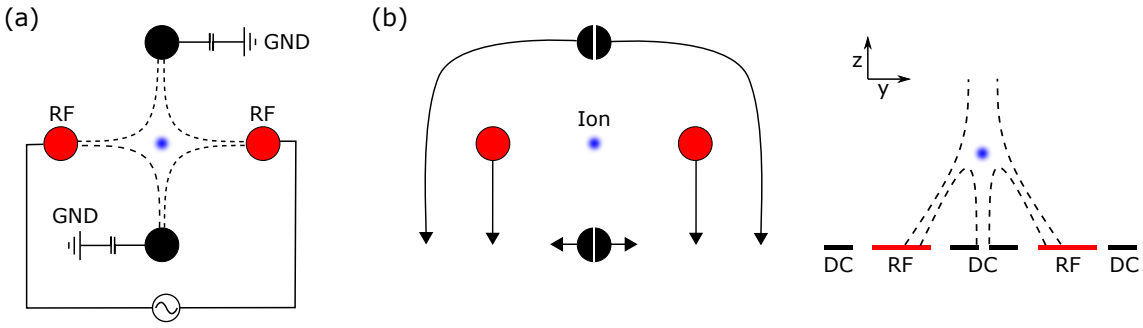

Figure 2.1: (a) A macro-fabricated four-rod trap places the RF confining voltage (red) on opposing diagonals of a square. The ground rods (black) are shorted to ground at RF frequency, but may hold a DC voltage. (b) The four-rod trap can be deformed into a micro-fabricated surface geometry and maintain a potential that traps ions (blue dot). Due to the symmetry breaking of the deformation, the surface geometry introduces higher-order multipole moments that can cause ion height dependent trap characteristics.

The push for scalable quantum computers has highlighted the need for precise traps with repeatable properties and many control electrodes required for shuttling ions between multiple axial trapping wells. Micro-fabrication techniques can achieve precise dimensional machining of electrodes at $\sim 1\ \mu$m resolution, at the cost of being limited to a mostly two-dimensional surface. A deformation of the four-rod trap to the a surface trap topology is shown in Fig. 2.1(b) and demonstrates how a trapping quadrupole-like RF field can still be realized above a plane. Although this work uses a micro-fabricated surface trap, micro-fabrication does not necessarily imply surface geometry. Using cantilevered MEMS-type features [29] or stacking of multiple 2D wafers [30] the geometry can be made to more closely resemble a blade-trap, while still realizing the precision benefits of micro-machining.

As of this thesis, there are two primary competing manufacturing processes for micro-fabricated traps: silicon CMOS and laser-etched fused silica. CMOS fabrication leverages the capabilities developed from the chip manufacturing industry to

build ion trap features out of silicon substrates. It is extremely precise and repeatable, but is limited to a small selection of available materials. In particular, while silicon is ideal for eventually integrating photonic waveguides [31] or peripheral electric elements (e.g., DACs [32], photodetectors [33]), it is not an ideal material from the ion's perspective. Silicon is a great photo-detector and so stray light from laser beams will create electrical carriers that may excite motion in the ion. Care must be taken to electrically shield any semiconducting material from line-of-sight of the ion.

The other manufacturing process uses focused femto-second lasers to "activate" areas of the substrate that are later etched away at a higher rate than the non-illuminated areas. After evaporating gold onto the substrate at an angle, the etched channels then define the electrode geometry. Compared to CMOS processes, which can involve hundreds of steps, the process here is much simpler, which typically means higher process yield and faster turnaround time for prototyping. Fused silica is also a preferable substrate since it is an insulator, and will not exhibit the photo-detector effect of silicon. On the other hand, UV light can ionize defects within the glass that leads to a slow charging/discharging of background electric field that can decompensate the ion from the RF null. We make a strong distinction between the "slow charging" from insulators and the "prompt charge response" from semiconducting photodiodes; the former can be calibrated for while the latter is quite damning for a trapped ion quantum computer. In this work, we use a CMOS micro-fabricated surface trap manufactured by Sandia National Laboratories [34].

The main drawback of micro-fabricated traps compared to their macro pre-

decessors are the voltage limitations that can be placed on the electrodes. While macro-traps can readily support up to $\sim$ 1 kV of RF voltage, it is risky to apply $>$ 300 V to a CMOS surface trap. This is a result of the relatively small separation ($\sim \mu$m) between RF and GND planes, with dielectric breakdown occurring within the insulating oxide layer that separates them. The lower RF voltage reduces the trap depth, which makes ion loss events more common due to residual background gas collisions. Additionally, it places a cap on the achievable secular frequency (Eq. 2.9), especially in higher mass ions, which can limit the speed of two-qubit gates.

Another consideration in surface traps is that the geometry breaks the rotational symmetry of the system, as can be seen in Fig. 2.1(b), and requires one to consider higher order multi-poles in the trapping potential. In simple surface trap models, the contribution from the hexapole term is roughly equal to the contribution from the quadrupole term [27]. The dominant effect of this in our system is a shift of the secular frequency that depends on the height of the ion from the trap surface (i.e., $\omega_z \rightarrow \omega_z(z)$). Intuitively, this is because the RF field lines are more concentrated below the ion than above it. Thus, for stability of radial motional mode frequencies in surface traps, it is critical to routinely check and calibrate the ion height, for example, by minimizing the micromotion along the vertical direction.

2.1.3: Ion Crystals and Normal Modes

In section 2.1.1, we described how the motion of one ion decomposes into three sets of independent motion along the axes of the trap. If $\omega_x \ll \omega_r$, then the

axial potential can support many ions, aligned along x. Assuming the ions are cold, they will arrange into a crystal that minimizes the total potential energy of the system. The condition for N ions to form a linear chain, rather than a zig-zag type arrangement is [35],

$$\frac{\omega_r}{\omega_x} > 0.77 \frac{N}{\sqrt{\log N}} \tag{2.11}$$

Assuming this condition is met and considering just the axial direction for now, the potential energy of the chain is given by a combination of the static axial confining potential $\Phi(x)$ and the inter-ion Coulomb repulsion.

$$\begin{aligned} U(\vec{x}) &= U_\Phi + U_{Coulomb} \\ &= \sum_i^N \Phi(x_i) + \frac{e^2}{4\pi\epsilon_0} \sum_{i<j} \frac{1}{|x_i - x_j|} \end{aligned} \tag{2.12}$$

The position of each ion can be solved for numerically by minimizing the potential energy,

$$\vec{x}^* = \underset{\vec{x}}{\operatorname{argmin}}\, U(\vec{x}) \tag{2.13}$$

It is useful to define a unit system for working with potentials. Native distance and energy units normalized to MHz can be defined as,

$$d_0 \equiv \left(\frac{e^2}{4\pi\epsilon_0 m (2\pi \times 1 \text{ MHz})} \right)^{1/3} \tag{2.14}$$

$$E_0 \equiv \frac{e^2}{4\pi\epsilon_0 d_0} \tag{2.15}$$

Then we can write a dimensionless potential energy $V = U/E_0$, using a di-

mensionless coordinate $u_i = x_i/l_0$ and expand the static potential in a polynomial basis as,

$$V(\vec{u}) = \sum_i \left(\sum_n X_n \frac{{u_i}^n}{n!} \right) + \frac{1}{2} \sum_{i \neq j} \frac{1}{|u_i - u_j|} \tag{2.16}$$

Where the coefficients X_n are unitless parameters that describe the relative strength of each term in the Taylor expansion of the static potential. The benefit to parameterizing the potential in this manner is that there is a straightforward conversion to real units. In real units, the static potential has the same form given by

$$U_\Phi(\vec{x}) = \sum_i \left(\sum_n c_n \frac{{x_i}^n}{n!} \right) \tag{2.17}$$

With conversion between real and dimensionless parameters,

$$c_n = \frac{E_0}{{d_0}^n} X_n \tag{2.18}$$

With this choice of units, there is also a very intuitive understanding of the X_2 parameter. For one ion in the trap with no other static fields applied, $X_2 = A$ will correspond to an axial mode frequency of $\sqrt{A}$ MHz. So for example $X_2 = 0.04$ will give a 200 kHz axial mode frequency. For $^{171}\mathrm{Yb}^+$, the conversion factors can be

summarized as follows.

$$\begin{aligned} c_1 &= 192\ e\mathrm{V/m} \times X_1 \\ c_2 &= 70.0\ e\mathrm{V/mm}^2 \times X_2 \\ c_3 &= 25.5\ \mathrm{k}e\mathrm{V/mm}^3 \times X_3 \\ c_4 &= 9.32\ \mathrm{M}e\mathrm{V/mm}^4 \times X_4 \end{aligned} \tag{2.19}$$

For the remainder of this thesis we will refer to the dimensionless parameters X_n. In our system, the Raman beams and detection optics are equispaced by 4.43μm at the ions. In order to align the ions to the imaging optics, the ions must be equispaced as well. Define the beam centers as $\vec{b}$ and the ion positions as $\vec{x}$. The goal is then to numerically optimize the trap parameters to minimize the standard deviation of the alignment error $\sigma_{\mathcal{E}} = \|\vec{b} - \vec{x}\|/\sqrt{N}$ for the N ions. Perfect alignment ($\sigma_{\mathcal{E}} = 0$) can only be achieved in an infinite square well, which unfortunately requires infinite terms in the Taylor expansion.

A quadratic potential only is not sufficient to achieve alignment across the chain, as shown in Fig. 2.2(a). Reasonable alignment ($\sigma_{\mathcal{E}} = 160$ nm) can be achieved with the addition of a quartic term as shown in Fig. 2.2(b). If we relax the requirement of alignment on the external edge ions, a further improvement ($\sigma_{\mathcal{E}} = 52$ nm) can be achieved on the internal ions, as shown in Fig. 2.2(c). In this case, the edge ions are generally not usable as qubits as they are offset by 760 nm from the Raman beams and detection optics. This results in a trade off between beam alignment and qubit number that will ultimately depend on system design (e.g. individual Raman

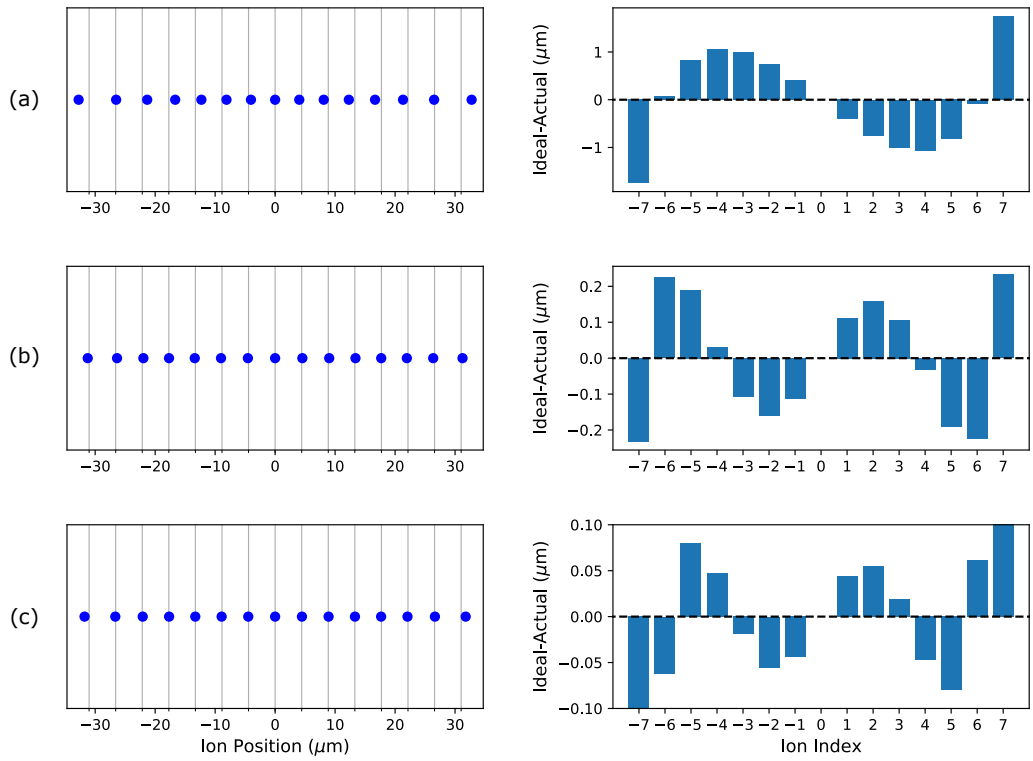

Figure 2.2: Optimizing 15 ions for alignment on imaging optics equispaced at 4.43 μm. The left column indicates the ion positions $\vec{x}$ (blue dots) along the x direction relative to the beam centers $\vec{b}$ (grey lines). The right column shows the alignment error $\vec{\mathcal{E}} = \vec{b} - \vec{x}$. (a) A quadratic potential only, $(X_2, X_4) = (0.029, 0)$, optimizing across the entire chain. (b) A quadratic and quartic potential, $(X_2, X_4) = (-0.005, 0.0023)$, optimizing across the entire chain. (c) A quadratic and quartic potential, $(X_2, X_4) = (0.00188, 0.00177)$, ignoring the end ions during optimization.

beam width) and the desired use case of the system. For high-fidelity operations, as in the case of this thesis, we bias towards beam alignment. In Section 5.1.3 we describe a useful secondary purpose for these edge ions.

Now that we have found the equilibrium positions of the ions in the chain, we now consider their harmonic motion about these points. The ion chain can be treated classically as a system of coupled harmonic oscillators. The Lagrangian of

this system is given by,

$$\begin{aligned}\mathcal{L} &= T - U \\ &= \sum_{i}^{N} \frac{m_i \dot{q}_{i,\alpha}^2}{2} - \sum_{i}^{N} U(q_{i,\alpha})\end{aligned} \tag{2.20}$$

Where $q_{i,\alpha}$ are generalized coordinates with i indexing the ions and $\alpha = x, y, z$ indexing the Cartesian coordinates. Unlike the potential presented in Eq. 2.12 which just considered the axial direction, we now must consider the all three Cartesian directions to describe the motion, including the pseudo-potential created by the RF. Assuming an upright quadrupole, the potential looks like,

$$\begin{aligned}U(\vec{r}) &= U_\Phi(\vec{x}) + U_{RF}(\vec{y}, \vec{z}) + U_{Coulomb}(\vec{r}) \\ &= \sum_{i}^{N} \Phi(x_i) + \sum_{i}^{N} \frac{m_i \omega_{sec}^2}{2}(y_i^2 + z_i^2) + \frac{e^2}{4\pi\epsilon_0} \sum_{i<j} \frac{1}{|\vec{r}_i - \vec{r}_j|}\end{aligned} \tag{2.21}$$

With $U_\Phi(\vec{x})$ left arbitrary, but may be a polynomial expansion such as Eq. 2.17. We can Taylor expand the potential energy U about the equilibrium points $q_i^{(0)}$ with $q_i = q_i^{(0)} + \widetilde{q}_i$,

$$\begin{aligned}U(q_i) &= U\left(q_i^{(0)}\right) + \widetilde{q}_i \left(\frac{\partial U}{\partial q_i}\right)_{q_i^{(0)}} + \sum_j \frac{\widetilde{q}_i \widetilde{q}_j}{2} \left(\frac{\partial^2 U}{\partial q_i \partial q_j}\right)_{q_i^{(0)}, q_j^{(0)}} \\ &= \frac{\widetilde{q}_i}{2} \sum_j \widetilde{q}_j U_{i,j}\end{aligned} \tag{2.22}$$

The first term in the expansion is a scalar and can be set to zero. The second term in the expansion is also zero as a consequence of static equilibrium (equivalent to minimizing the potential energy in Eq. 2.13). Thus the third term, which is just

the Hessian matrix ($\mathbf{U} = U_{i,j}$) of the potential, contains all the information about the motion of the ions about equilibrium. Assuming that the equations of motion are oscillatory of the form $\widetilde{q_i}(t) = b_i \exp^{-i\omega t}$ then the expanded Lagrangian equation of motion can be written as,

$$\sum_j U_{i,j} b_i = \sum_j \omega^2 T_{i,j} b_i \tag{2.23}$$

Where $T_{i,j}$ is just a diagonal matrix with entries corresponding to the mass of each ion in the chain, which we now label as $\mathbf{M}$. For a single-species ion chain, this reduces to just a constant, although here we explicitly allow for the possibility of chain of mixed-species (note for multiple species, ω_{sec} and therefore U_{RF} also depend on mass). Then identifying $\lambda = \omega^2$, Eq. 2.23 is just an eigenvalue problem, which can be solved numerically,

$$\mathbf{M}^{-1}\mathbf{U}\vec{b_k} = \lambda_k \vec{b_k} \tag{2.24}$$

The Hessian matrix $\mathbf{U}$ is a $3N \times 3N$ matrix, corresponding to each ion in each of the three Cartesian directions. Thus there will be $3N$ normal modes of motion in the chain, $\vec{b_k}$, which are unit vectors describing the participation of each ion in that mode, each oscillating at frequency $w_k = \sqrt{\lambda_k}$. The motional modes form an independent basis set to describe the motion of the chain and are therefore orthogonal. Typically we group the modes into three sets - the axial modes and the two transverse (synonymously called radial) set of modes. Note that the direction of the transverse modes may not necessarily align with the Cartesian coordinates,

depending on the trap geometry or applied static quadrupoles (e.g. yz quadrupole, Q_{yz}). The directions of the transverse modes defines the "principal axes" of the trap, which can be determined by evaluating the curvature of the non-Coulomb potential ($U_0 = U_\Phi + U_{RF}$) at $\vec{r} = \vec{0}$ using a similar eigenvalue method. The principal axes are then $\vec{b}_p$ where $\mathbf{H}$ denotes the Hessian operator.

$$\mathbf{H}(U_0)|_{x,y,z=0}\ \vec{b}_p = \lambda_p\ \vec{b}_p \tag{2.25}$$

An example of the mode spectrum ω_k, using U_0 for the system in described this thesis (see Section 4.2.1) is shown in Fig. 2.3. Note that for gates operating on the transverse modes (as in this thesis), it is important to keep the highest energy axial mode well separated from the lowest energy radial mode, so as to avoid exciting axial motion. Overlap of the radial modes, while not ideal, is mitigated by aligning the $\Delta\vec{k}$ of the Raman lasers to be aligned along the principal axis of the set of radial modes used for gates. As a side note, a consequence of requiring equispaced ions is that the radial modes are also roughly equispaced, which can simplify gate design (see Section 3.3.4).

An example of the radial mode participation vectors $\vec{b}_k$ for the mode spectrum given in Fig. 2.3 is shown in Fig. 2.4. The leftmost column describes the in-phase motional mode, synonymously called the lowest spatial-frequency mode, corresponding to the highest energy radial mode. The intuitive picture for this relation is that in-phase motion in the radial direction minimizes the distance between ions and thus maximizes the Coulomb interaction. For the axial modes, it is the opposite;

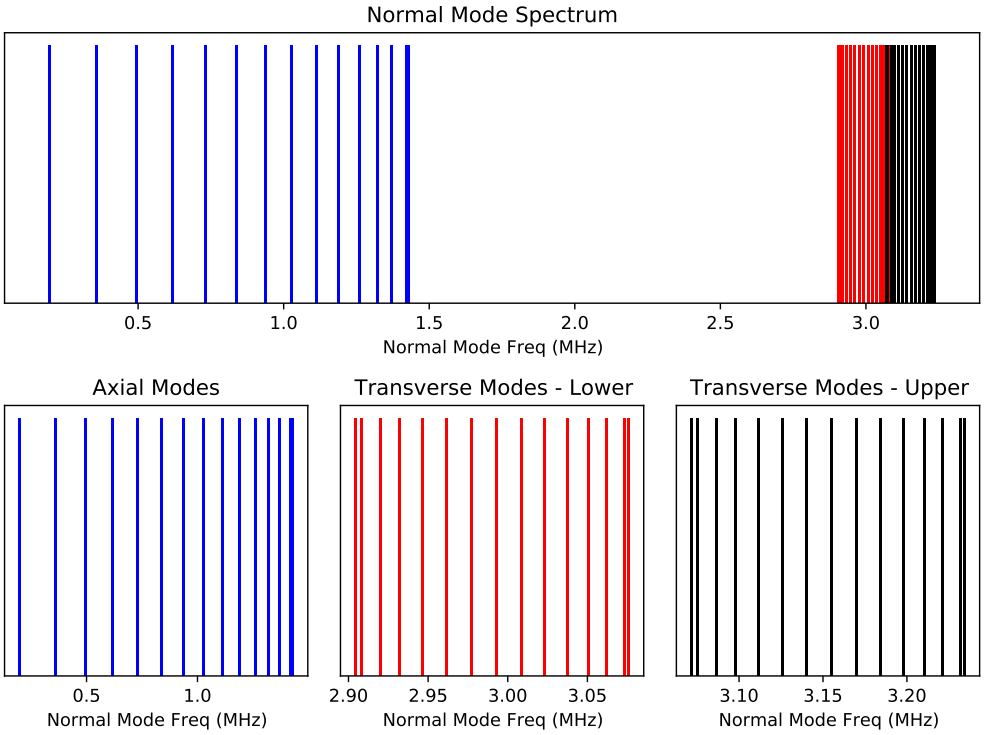

Figure 2.3: (Top) Full spectrum for all the normal modes in a 15-ion chain using the potential described in Section 4.2.1. (Bottom) A detailed view of each individual set of normal modes along each of the three principal axes of the trap.

in-phase motion in the axial direction maximizes the distance between ions and thus minimizes the Coulomb interaction, making it the lowest energy axial mode. Although historically this mode been called center-of-mass (COM) mode, we avoid that terminology in this thesis, because it does exhibit a spatial "bending" feature. Throughout this thesis, we consistently label the in-phase mode with mode index 1 for both axial and radial sets.

The rightmost column describes the out-of-phase motional mode, synonymously called the highest spatial-frequency mode, corresponding to the lowest energy radial mode (and historically called the "zig-zag" mode). The out-of-phase motion in the radial (axial) direction maximizes (minimizes) the distance between

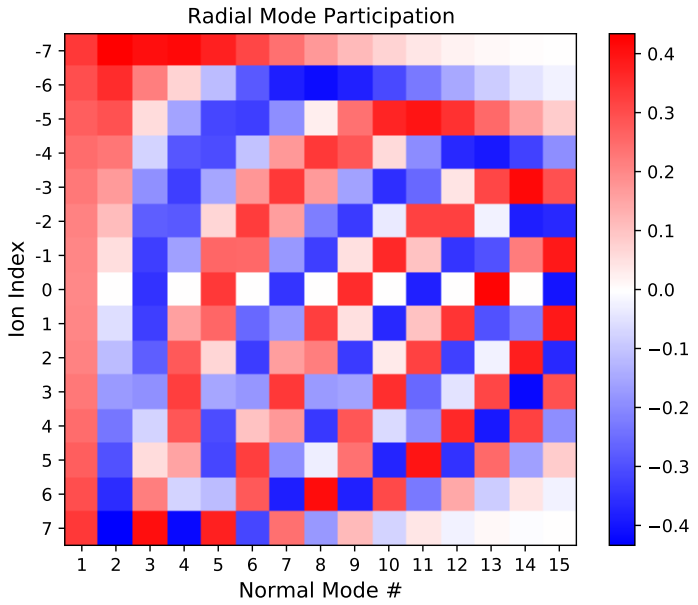

Figure 2.4: Lower radial mode participation matrix for 15-ion chain corresponding to the red spectrum in Figure 2.3. The left(right)-most column corresponds to the in(out-of)-phase motional mode with the highest(lowest) energy.

ions, thus minimizing (maximizing) the Coulomb interaction and making it the lowest (highest) energy mode amongst the mode set.

2.2: The $^{171}Yb^{+}$ Ion

There are a finite selection of elements to use for trapped ion quantum information processing. The most rigid requirement is that the ion have a hydrogen-like atomic level structure, most commonly found in ions with only a single outer electron (e.g. alkali earth metals), although some dual outer electron species also exhibit hydrogen-like spectra (e.g., Al^{+}). The complexity of atomic levels grows rapidly with increasing number of outer electrons and the amount of lasers required to control the atom quickly becomes impractical. The list of practical ion choices is thus typically

limited to: Be^+, Mg^+, Ca^+, Sr^+, Ba^+, Zn^+, Cd^+, Yb^+, and Hg^+. Among these choices, the commercial availability of laser wavelengths on the primary Doppler cooling line and/or the qubit transition should be considered for scaling systems into robust large-scale quantum computers.

There are also several different strategies to encode a perfect two-level system (that of an ideal qubit) into realistic multi-level atom. Zeeman qubits are encoded into ground state $S_{1/2}$ Zeeman levels in isotopes with typically zero nuclear spin. The qubits are inherently magnetic field sensitive, but also easily controlled with RF frequencies ($\sim$ MHz) and benefit from simple atomic level structure (no hyperfine levels). Additionally, there are no auxiliary states in the ground state other than the qubit states, making leakage impossible.

Hyperfine qubits in isotopes with non-zero nuclear spin ($I \neq 0$) are encoded into the ground state $S_{1/2}(F = I \pm 1/2, m_F = 0)$ levels typically split by frequencies on the GHz scale. Since $\Delta m_F = 0$, these qubits are called "clock-qubits" and are first-order magnetic field insensitive. Thus hyperfine qubits have many of the benefits of Zeeman qubits with longer practical coherence times. Typically an isotope nuclear spin of $I = 1/2$ is ideal since it minimizes the number of auxiliary states, although higher nuclear spins can be used with the cost of precise laser polarization control for state preparation and measurement. Optical addressing of hyperfine qubits typically requires two laser beams in a Raman configuration.

Finally, optical qubits make use of a low-lying D states to encode the qubit states, connected by an electrical quadrupole transition, typically $S_{1/2} \leftrightarrow D_{5/2}$. While the lifetimes of of the meta-stable D states are quite long ($\sim s$) compared

to coherent operations, they are not ∼infinite as in the case of Zeeman/HF qubits, which may pose a fundamental long-term limitation for ion trap quantum computers. Also, the local oscillator in this case is an optical frequency and must be stabilized to ∼Hz linewidths to maintain reasonable qubit coherence, a non-trivial task. On the other hand, qubit operations only require one laser beam, which can make the optical design simpler.

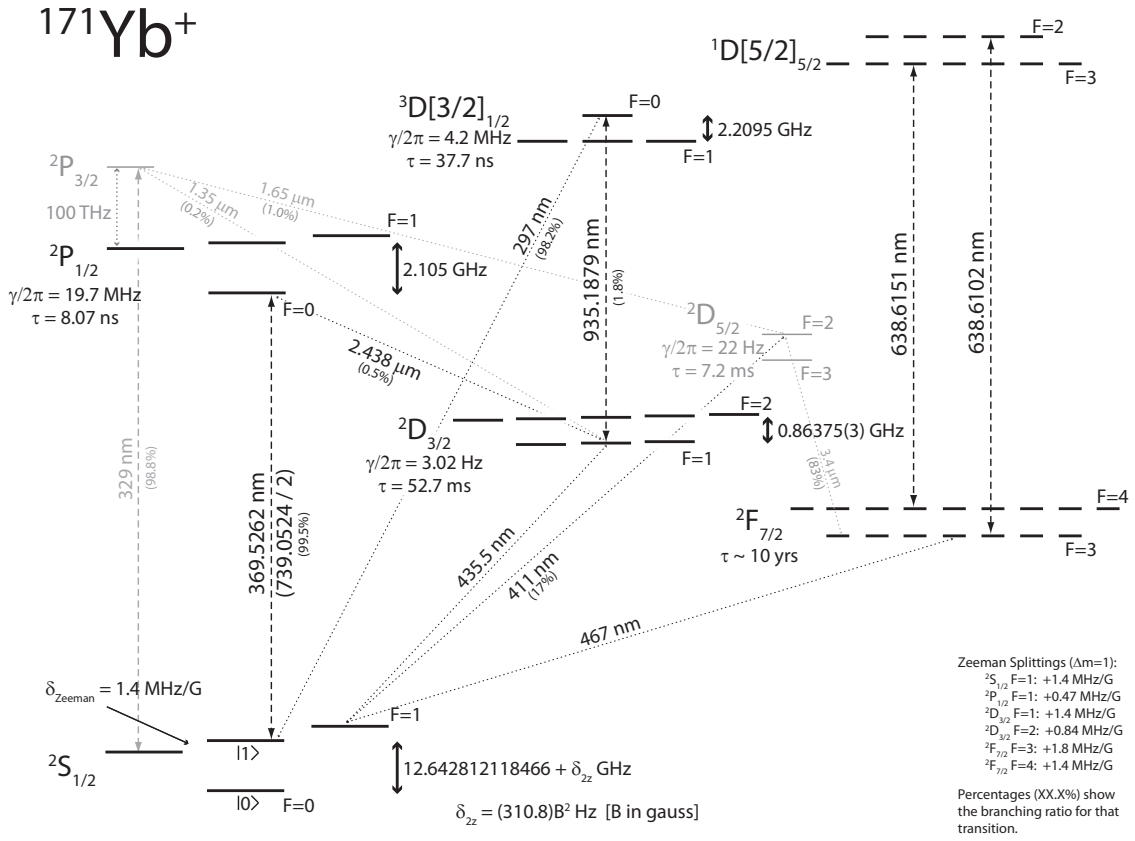

Figure 2.5: Relevant atomic energy levels for operating the $^{171}Yb^{+}$ ion as a qubit.

For hyperfine qubits, $^{171}Yb^{+}$ is a great choice because it is a stable isotope with nuclear spin $I = 1/2$. The primary Doppler cooling line $^2S_{1/2} \leftrightarrow ^2P_{1/2}$ is at 369.5 nm, which is UV but also within the range of where laser diodes are still

commercially available[1] [36] and compatible with optical fibers over short distances. For qubit control, the Raman wavelength that minimizes two-photon Stark shifts is at 355nm, where there are several commercially available high-power lasers[2].

The atomic level structure of $^{171}Yb^{+}$ is shown in Figure 2.5. The qubit is defined on the $^{2}S_{1/2}$ ground state, on the $\Delta m_F = 0$ hyperfine splitting as follows,

$$|0\rangle = |F = 0, m_F = 0\rangle$$
$$|1\rangle = |F = 1, m_F = 0\rangle \tag{2.26}$$

2.2.1: Ionization and Loading

The first step to trapping an ion is to ionize a neutral atom near the center of the RF null. Neutral atoms reside in solid fragments of an isotopically enriched (95%) $^{171}Yb^{+}$ source placed inside a stainless steel tube that is closed at one end. The tube is resistively heated by passing current (2.06 A) through a tungsten coiled filament welded to the closed end, causing the sample to heat up and atoms to evaporate. Because the oven is in a UHV environment, where the mean free path is $\sim$ km, the evaporated atoms become collimated by the aspect ratio of the tube and form a well directed flux beam. The atomic flux beam is oriented at the bottom of the chip trap, passing through a load slot milled through the chip, and arriving at the RF null located above the surface of the chip.

The neutrals are then photo-ionized in a two-stage process. First, the electron

[1] Nichia NDU1113E
[2] Paladin Compact 355-4000

in the neutral atom is excited from $^1S_0 \rightarrow {}^1P_1$ by a 398.9 nm laser. From this excited state, the continuum limit for ionizing the the electron is $\approx$ 394 nm. Several readily available wavelengths are suitable for this ionization process including the 369.5 nm Doppler cooling and 355 nm Raman lasers required to operate the $^{171}Yb^+$ ion. Typically, there is limited available laser power at 369.5 nm ($\sim$ mW), which can limit the ionization rate. In comparison, there is excess beam power ($\sim$ W) available from the Raman 355 nm laser, in particular from picking off the 0^{th} AOM order from the beam modulators, which is otherwise beam-dumped. This has the added engineering complexity associated with re-routing that beam with limited-space constraints, but with benefit of fast ionization rates.

In this system, we elect to use a third option, a free-running high-power (120 mW) laser diode[3] operating at 393 nm and modulated by a fast shutter system [37]. This ionization light is spatially multiplexed via polarization (beam cube) with the 399 nm neutral excitation beam into a fiber that is routed to a beam delivery box mounted on the vacuum chamber. A reflective collimator[4] at the fiber output collimates the beams independent of the wavelength. An achromatic lens[5] focuses the beams to a waist of 17 μm at the RF null. With 10 mW of 393 nm power and 5 μW of 399 nm power, we achieve an average ion loading rate of 0.5 s/load from a hot oven.

The neutral excitation and ionizing beams are overlapped with the RF null, oriented parallel to the chip surface and perpendicular to the velocity vector of the

[3]Nichia NDU4316
[4]Thorlabs RC08FC-F01
[5]Edmund Optics #65-980

atoms. This is important because the neutral atoms are thermally distributed with a velocity peak above room temperature and would see a broad Doppler shifted $^1S_0 \rightarrow {}^1P_1$ line for a neutral excitation beam oriented parallel to the flux. By orienting the beam perpendicular to the flux, we see the Doppler-free linewidth, which provides isotope selectivity.

2.2.2: Doppler Cooling

Now that a $^{171}Yb^+$ ion exists within the RF trapping region, it sees the confining psuedo-potential. However, the ion is quite hot because it maintains the thermal energy required to evaporate it out of the solid state. This energy must be removed from the ion in the form of photons in order for the ion to become well localized within the trap. To see how this is accomplished, consider a laser detuned by Δ from an electronic transition in an atom. Then the scattering rate of photons is given by [38],

$$\gamma_s = \frac{s\gamma/2}{1 + s + 4\Delta^2/\gamma^2} \tag{2.27}$$

where $s = I/I_s$ is the saturation parameter with saturation intensity given by $I_s = \pi hc/3\lambda^3\tau$ for the resonant wavelength of the transition λ and the lifetime of the excited state τ. γ is the angular linewidth of the transition $\gamma = 1/\tau$. The first order Doppler-shift seen by the ion is then given by $\delta_D = -\vec{k} \cdot \vec{v}$ where $\vec{v}$ is the velocity vector of the ion and $\vec{k}$ is the k-vector of the laser. Then the effective detuning seen by the ion is given by $\Delta_D = \Delta + \delta_D$. When the ion is moving in the same (opposite) direction as the photons, then the frequency of the photon is

red(blue)-shifted and the effective detuning decreases (increases). If the laser is detuned red of the transition, $\Delta < 0$, then the scattering rate will increase when the ion is moving in the opposite direction of the photons. Each absorption of a photon changes the momentum of the ion by $\hbar k$ in the direction of $\hat{k}$, whereas the emission of a photon is isotropic and averages to zero momentum transfer over several scattering events. Therefore, the momentum of an ion moving in a direction opposite to the k-vector of the laser is preferentially removed, resulting in a loss of kinetic energy and cooling of the ions.

Since the ion is oscillating within the RF pseudo-potential, each half-period of motion will be blue-shifted, providing several opportunities for cooling. Unlike a free, untrapped atom, which requires counter-propagating sets of laser beams in all three Cartesian dimensions to cool (6 beams total), trapped ions only require a single Doppler cooling beam. The caveat is that the $\vec{k}$ of Doppler cooling beam must have projections along all the principal axes of the trap in order to fully cool the ion. In this system, the Doppler cooling beam is oriented at 45° relative to the trap axis ($\hat{x}$) and parallel to the chip surface (xy-plane). The radial axes of the trap are then tilted at 45° relative to the chip surface with a DC yz-quadrupole field (Q_{yz}). In this configuration, the Doppler cooling beam has projections onto the axial and both radial sets of motional modes.

The optimal detuning value of the Doppler cooling beam Δ will depend on the temperature of the ion. Since the Doppler cooling mechanism depends on the slope of the scattering rate curve, the maximum cooling efficiency will be achieved at the maximum slope. The location of this maximum slope will depend on both the ion

temperature due to the Doppler broadening and the applied laser power due to power broadening. In $^{171}Yb^+$ the primary cooling line is on $^2S_{1/2}(F=1) \rightarrow {}^2P_{1/2}(F=0)$ with linewidth $\gamma/2\pi = 19.7$ MHz. In all of our experiments with chains, we start with 1 ms of Doppler cooling at $\Delta = 2\pi \times -42$ MHz detuning and then proceed with 1 ms of second-stage Doppler cooling at $\Delta = 2\pi \times -8$ MHz with a weaker beam near one saturation intensity. The Doppler limit is given by $T_D = \hbar\gamma/2k_B$ which for $^{171}Yb^+$ is ≈ 480 μK. Although the average momentum from spontaneous emission averages to zero, $\langle p \rangle = 0$, the average squared momentum does not, $\langle p^2 \rangle \neq 0$, representing a heating process. The Doppler limit is the equilibrium where the heating rate equals the cooling rate.

As stated above, the primary Doppler cooling line used in $^{171}Yb^+$ is $^2S_{1/2}(F = 1) \rightarrow {}^2P_{1/2}(F = 0)$, although several factors make this process deviate from an ideal two-level system. The first is the existence of Zeeman levels in the $^2S_{1/2}(F = 1)$ state. From the excited $^2P_{1/2}(F = 0)$ level, the ion will probabilistically relax into any one of the Zeeman levels. Thus a mixture of linear $\hat{\pi}$ and circular $\hat{\sigma}_\pm$ polarizations is required to ensure the ion continues to scatter photons. Additionally, a magnetic bias field is required in destabilize coherent dark states that can arise from superpositions of angular momentum states that interfere to create net zero coupling to the $^2P_{1/2}(F = 0)$ level [39]. A magnetic field of $\mathbf{B} \approx 5$ G is applied perpendicular to the surface of the trap and the k-vector of the Doppler beam, defining an atomic quantization axis and lifting the degeneracy of the Zeeman states. The ideal polarization of light, which equalizes the Rabi frequency for all the Zeeman states, is linear and oriented at $\approx 55°$ relative to $\mathbf{B}$ [39].

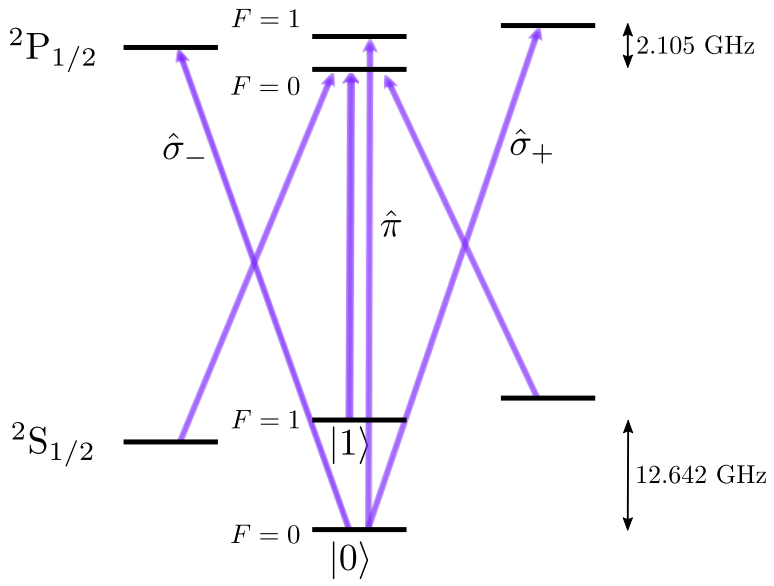

Figure 2.6: Schematic depicting the relevant atomic energy levels that must be addressed in order to efficiently cool $^{171}Yb^+$. Not shown is the $^2D_{3/2}$ manifold, which also must be actively depopulated.

The second deviation from the ideal level system is the existence of hyperfine structure in $^{171}Yb^+$. Driving the transition $^2S_{1/2}(F = 1) \rightarrow {}^2P_{1/2}(F = 0)$ can occasionally off-resonantly excite $^2P_{1/2}(F = 1)$, since the hyperfine splitting in the excited state is only 2.105 GHz. From here the ion can decay into $^2S_{1/2}(F = 0)$ which would cause cooling to cease. In order to bring the ion back into the cooling cycle, we also need to drive $^2S_{1/2}(F = 0) \rightarrow {}^2P_{1/2}(F = 1)$ as shown in Figure 2.6. This is accomplished by creating 14.748 GHz sidebands on the main cooling light with a resonant free-space EOM[6].

Finally, as shown in Figure 2.5, there is 0.5% probability that the ion decays into the $^2D_{3/2}$ manifold from the excited $^2P_{1/2}(F = 1)$ state. This D-state is metastable with $\tau = 52.7$ ms and removes the ion from the cooling cycle after

[6]Qubig, PM - Yb+_14.7

only ≈ 200 scattering events. An IR laser at 935 nm depopulates this state to the bracket $^3\mathrm{D}[3/2]_{1/2}$ state, which quickly decays back into the main cycling path. As above, the Zeeman and hyperfine structure also exists in the $^2\mathrm{D}_{3/2}$ manifold, which means the 935 nm laser requires a mixture of $\hat{\pi}$ and circular $\hat{\sigma}_\pm$ polarization, and sidebands at 3.072 GHz created by a resonant EOM in order ensure the ion cannot get stuck. In the end, all the auxiliary frequencies and lasers required for Doppler cooling $^{171}\mathrm{Yb}^+$ are only meant to ensure that the ion continues to scatter photons on the main transition so that the ion can cool efficiently.

We also note that if the ion collides with a background gas particle while it is in the excited $^2\mathrm{P}_{1/2}$ state, it can inelastically scatter into the long-lived ($\tau \sim 10$ yrs) $^2\mathrm{F}_{7/2}$ state. Empirically, we observe that application of high-power 355-nm Raman light can recover the ion from this dark state, likely through a nonlinear multi-photon process. A 638-nm laser can also be used to repump via the $^1\mathrm{D}[5/2]_{5/2}$ state.

2.2.3: State Preparation

Now that there is a cold, well-localized ion in the trap, we can use it to store quantum information. In order to run circuits, the ion must start in a well defined quantum state. We initialize the quantum state into $|0\rangle$ by means of optical pumping. We wish to clear out the $^2\mathrm{S}_{1/2}(F = 1)$ state, primarily by exciting to the $^2\mathrm{P}_{1/2}(F = 1)$ state, which has a high branching ratio into $|0\rangle$. Over several scattering events the probability of remaining in the $^2\mathrm{S}_{1/2}(F = 1)$ rapidly decays.

In this system, we accomplish this by adding 2.105 GHz sidebands to the same 369.5 nm light used for the main line of Doppler cooling (with $\Delta = 0$ and being

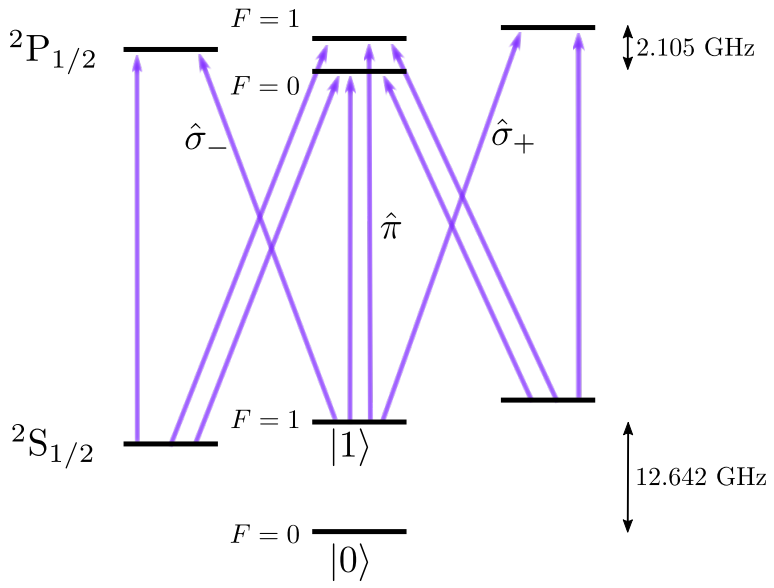

Figure 2.7: Schematic depicting the relevant atomic energy levels that must be addressed in order to efficiently pump $^{171}Yb^+$ into the $|0\rangle$ state. Not shown is the $^2D_{3/2}$ manifold, which also must be actively depopulated.

certain to turn off the 14.748 GHz sidebands). As shown in Figure 2.7, the $^2S_{1/2}(F = 1) \rightarrow {}^2P_{1/2}(F = 0)$ transition does not contribute to the pumping, but it also does not degrade fidelity significantly because it is 14.7 GHz off-resonant from the undesired $^2S_{1/2}(F = 0) \rightarrow {}^2P_{1/2}(F = 1)$ transition. As with Doppler cooling, this process requires 935 nm light with the proper polarization and frequency sideband in order to depopulate the $^2D_{3/2}$ manifold. We use a 100 μs "slow pump" procedure with a low saturation intensity to further reduce the risk of off-resonant excitation. After optical pumping, we measure 99.78% fidelity for the $|0\rangle$ state, with only 0.02% error due to the optical pumping procedure and the rest due to detection errors (see Section 5.3.1).

2.2.4: State Detection

At the conclusion of a quantum circuit, we must measure the state of all the individual ions in the chain to extract the results of the algorithm. The detection scheme in $^{171}Yb^{+}$ is based on state-dependent fluorescence. As shown in Figure 2.8, we illuminate the ion with light that is resonant with $^{2}S_{1/2}(F = 1) \rightarrow {}^{2}P_{1/2}(F = 0)$. Ions in $|0\rangle$ are off-resonant by 14.7 GHz from this light and will scatter very few photons, making it "dark". On the other hand, ions in $|1\rangle$ are resonant and will scatter many photons, making it "bright". Since the decay from $^{2}P_{1/2}(F = 0) \rightarrow {}^{2}2_{1/2}(F = 0)$ is forbidden, the ion will cycle in this transition, emitting many photons. Even if the ion decays to the $^{2}D_{3/2}$ manifold, it will likely get excited to $^{3}D[3/2]_{1/2}(F = 0)$ by 935 nm, which is forbidden to decay to $|0\rangle$, and therefore it will return to the detection cycle without an error.

Even though a single photon is enough to collapse the wavefunction of the qubit, the bright/dark states are long-lived over several hundred scattering events. In our typical detection window of 100 μs, a bright ion will scatter around 1000 photons, which makes it rather easy to gather sufficient SNR to discriminate between dark and bright states. Figure 2.9 shows a typical photon histogram for a state prepared in $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. When the ion is bright, we collect an average of ≈ 10 photons corresponding to $\approx 1\%$ total detection efficiency (see Section 4.5). The number of photons detected is Poisson distributed about this mean value. To distinguish bright from dark, we use a simple threshold discriminator set at 1.5 photons. On a single shot, if the the number of photons collected is below (above) this threshold

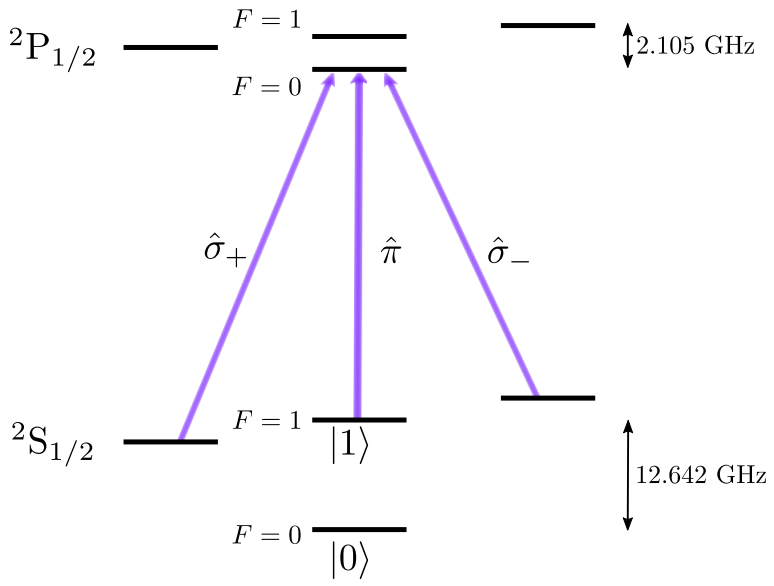

Figure 2.8: Schematic depicting the relevant atomic energy levels that must be addressed in order to perform state detection. Not shown is the $^2D_{3/2}$ manifold, which also must be actively depopulated.

then we declare the state $|0\rangle$ ($|1\rangle$), as shown in Figure 2.9. With this technique we achieve an average detection fidelity of 99.54% (see Section 5.3.1), which is close to the maximum theoretical detection fidelity of 99.85% at our detection efficiency [40].

The main detection error is caused by leakage from the cycling transition by 2.105 GHz off resonant excitation to $^2P_{1/2}(F = 1)$, which can cause a bright ion to go dark. For this reason, it is important to drive detection efficiency as high as possible to maximize the probability of detecting > 1 photon before the ion is pumped dark. Additionally, the detection window should be minimized to the point where there is no excess separation between the $|0\rangle$ and $|1\rangle$ state histograms. This reduces the error due to background dark counts (which would cause a dark ion to look bright), detection crosstalk (from neighboring bright ions), and dark-to-bright off-resonant excitation. Different discrimination schemes above and beyond a simple threshold,

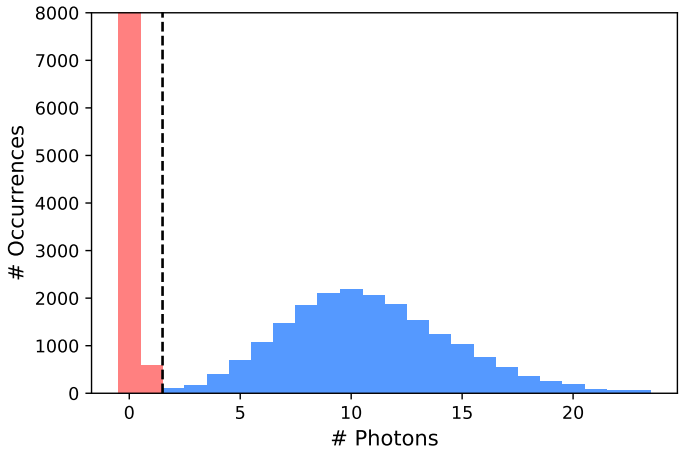

Figure 2.9: Typical photon histogram for a state prepared in $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, measured over a 100 μs detection window, and accumulated over many experimental repetitions. We use a threshold discriminator at 1.5 photons (dashed black line) to distinguish between the bright $|1\rangle$ state (blue) and the dark $|0\rangle$ state (red).

such as time-resolved photon counting [41] and adaptive Bayesian estimation [42] may improve detection fidelity. Additionally, the detection efficiency can readily be improved by increasing the quantum efficiency of the photon detector in the UV range, for example with a superconducting nanowire single-photon detector [43, 44]. Finally, other atomic transitions can supplement the state-dependant fluorescence scheme, such as shelving to the $^2F_{7/2}$ state [45], to decrease the off-resonant excitation errors.

2.2.5: The $^{171}Yb^+$ Hyperfine Structure

As described in Equation 2.26, the qubit is defined on the ground state hyperfine "clock" transition separated by $\omega_0/2\pi = 12,642,812,118.5 + \delta_2$ Hz, where $\delta_2 = (310.8)\mathbf{B}^2$ Hz is the second-order Zeeman shift for a magnetic field $\mathbf{B}$ in Gauss [40, 46]. This transition is called a clock state because $\Delta m_F = 0$, which

means it is first-order insensitive to magnetic field noise, which greatly increases its coherence time. There is still a quadratic sensitivity to magnetic field noise, although the effect is much smaller as long as $\mathbf{B}$ is kept small. Recall that some non-zero $\mathbf{B}$ is required to define the atomic quantization axis and prevent population trapping in dark states.

The total atomic system in $^{171}\mathrm{Yb}^{+}$ consists of two spin-1/2 particles, the neutron and the electron. The $F = 0$ state is a singlet state and the $F = 1$ state is a triplet state. Then in terms of the two spins in the system,

$$\begin{aligned} |0\rangle &= \frac{1}{\sqrt{2}}(|\downarrow_e\uparrow_n\rangle - |\uparrow_e\downarrow_n\rangle) \\ |1\rangle &= \frac{1}{\sqrt{2}}(|\downarrow_e\uparrow_n\rangle + |\uparrow_e\downarrow_n\rangle) \end{aligned} \tag{2.28}$$

Thus in order to flip the qubit state $|0\rangle \rightarrow |1\rangle$, there must be a phase shift of the electron spin $|\downarrow_e\rangle \rightarrow |\downarrow_e\rangle, |\uparrow_e\rangle \rightarrow -|\uparrow_e\rangle$. For this reason, the lifetime of the $|1\rangle$ is practically infinite, since even a magnetic dipole transition on the electron will not change the qubit state, due to the symmetry of the $m_F = 0$ singlet and triplet states. To create a phase shift in the electron spin we must apply a B-field that is parallel to the DC bias field $\mathbf{B}$ and has a frequency that is resonant with ω_0. In our system, a microwave horn antenna is used to couple the resonant MW frequency into free-space with linear polarization and the electric field perpendicular to $\mathbf{B}$. As a side note, the Raman interactions described in Section 3.1.2 create a fictitious B-field that is also oriented along $\mathbf{B}$ to generate this phase shift.

With microwave control of the qubit, we can characterize the natural coherence

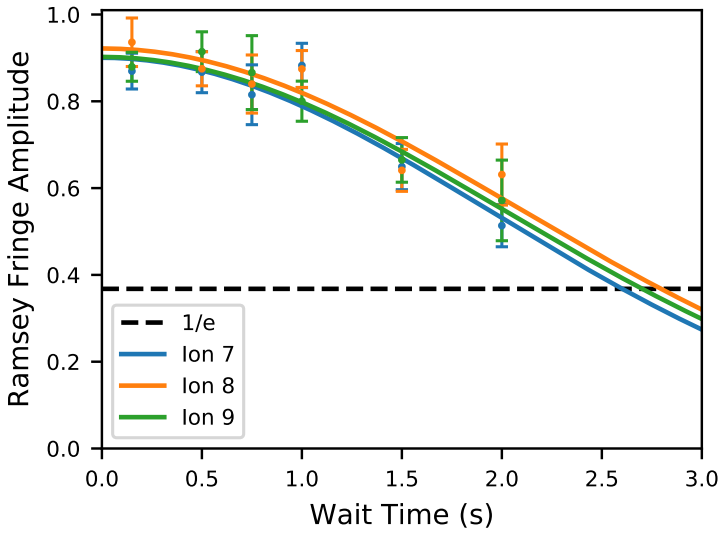

Figure 2.10: Microwave T_2 for three physical qubits in a 15-ion chain. The contrast is fit to a Gaussian decay $Ae^{-(\tau_{\text{wait}}/T_2)^2}$, with average fit value $T_2 = 2.84(16)$ s.

time of the qubit by preparing $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, waiting, and then applying a $\pi/2$ pulse of varying phase that attempts to return the qubit into a pure $|0/1\rangle$ state, called a Ramsey fringe. The extent to which the qubit does not perfectly return to $|0/1\rangle$ quantifies the decoherence in the system and is called the Ramsey fringe amplitude. Additionally, during the wait time, we suppress static magnetic field inhomogeneity using a dynamical decoupling technique that applies π-pulses with alternating 90° phase offsets, commonly known as an $(XY)^N$ pulse sequence [47], to periodically refocus the qubit spin. The resulting data is shown in Figure 2.10. We observe that the resulting decay is better fit to a Gaussian ($Ae^{-(\tau_{\text{wait}}/T_2)^2}$), compared to an exponential decay, with average $T_2 = 2.84(16)$ s. The coherence time of this echo experiment is limited by control noise rather than the qubit itself, as $T_2 > 1$ hour has been achieved in $^{171}\text{Yb}^+$ in other experiments [21]. In particular, residual magnetic field noise and/or drift in the LO used to generate the MW tone will result in phase

noise and cause decoherence. In the future magnetic field noise can be decreased by operating our qubit in a lower bias-field or by using magnetic shielding.

Chapter 3: From Ions To Qubits

In order to compute with qubits, one must be able to implement a universal gate set. Quantum gates on qubits are represented mathematically in a simple form using matrices. The challenge then is to control a physical Hamiltonian so that the desired unitary matrix is actually implemented on the ion. In the following sections, we describe the Hamiltonian for a trapped ion system and then depict how it can be controlled to implement digital single- and two-qubit gates that form a universal gate set.

3.1: Coherent Raman Control

Lasers are the primary tool for controlling trapped ions. To build intuition on these interactions, we first present a simple model of single CW laser beam and a qubit encoded on a dipole allowed transition. Then, we introduce modifications to this simple model to allow for two-photon Raman transitions on hyperfine qubits. Finally, we describe the pulsed Raman frequency comb that is used in this system for coherent control of the qubits.

3.1.1: Simple Trapped Ion Model

A trapped ion consists of two near-perfect quantum systems: a spin qubit rotating at its resonant frequency (ω_0) and quantum harmonic oscillator defined by the frequency of ion motion in a harmonic trap potential (ω_m).

$$\begin{aligned} H_{spin} &= \frac{1}{2}\hbar\omega_0\hat{\sigma_z} \\ H_{motion} &= \hbar\omega_m(a^\dagger a + \frac{1}{2}) \\ H_0 &= H_{spin} + H_{motion} \end{aligned} \tag{3.1}$$

Controlling the ion corresponds to controlling both of these degrees of freedom. For now, imagine that ω_0 is an optical frequency such that the trapped ion can be addressed with near-resonant laser light. For the simple case of a single laser beam addressing a dipole transition, the perturbing Hamiltonian can be written as:

$$H_1 = -\hat{d} \cdot \mathbf{E} \tag{3.2}$$

where $\mathbf{E}$ is the applied electric field at the ion location $\hat{x}$ and $\hat{d}$ is the dipole operator. This gives the total Hamiltonian,

$$H = H_0 + H_1 \tag{3.3}$$

Assume a traveling wave $\mathbf{E} = E_0\hat{\epsilon}\cos(\vec{k}\hat{x} + \omega t + \phi)$ and using $E_0(\hat{d} \cdot \hat{\epsilon}) = \Omega(\hat{\sigma}_+ + \hat{\sigma}_-)$, where Ω is the resonant Rabi frequency defined by the frequency with

which population is exchanged between $|\downarrow\rangle \leftrightarrow |\uparrow\rangle$. The spin lowering and raising operators $\hat{\sigma}_+/\hat{\sigma}_-$ are defined by $\hat{\sigma}_+ = \hat{\sigma}_x + i\hat{\sigma}_y = |\uparrow\rangle\langle\downarrow|$ and $\hat{\sigma}_- = \hat{\sigma}_x - i\hat{\sigma}_y = |\downarrow\rangle\langle\uparrow|$. Transforming to the interaction frame, $H_I = U^\dagger H U + i\frac{\partial U^\dagger}{\partial t}U$ characterized by the unitary $U = e^{-iH_0 t/\hbar}$, and using the RWA approximation ($\omega, \omega_0 \gg \Omega$) to ignore the fast rotating terms at $e^{\pm i(\omega+\omega_0)t}$, we can write,

$$H_I = \frac{\hbar\Omega}{2}\left(\hat{\sigma}_+ e^{i(\vec{k}\hat{x}-\mu t-\phi)} + h.c.\right) \tag{3.4}$$

Where $\mu = \omega - \omega_0$ parameterizes the detuning of the laser beam frequency from the resonance. In this simple case, we are only considering a single motional mode (one of three for a single ion). In practice, this is accomplished by aligning the $\vec{k}$ vector of the laser beam to be parallel with one of the principle trap axes (in this example, $\hat{k} \parallel \hat{x}$). Since the mode decomposition forms a independent basis set that describes the motion of an ion, the modes do not couple to each other and we can ignore the other two modes that are not addressed. The position of the ion can be expanded into two parts, the mean (stationary) position of the ion and the small harmonic oscillations about that mean, $\vec{k}\hat{x} = kx_{eq} + \eta(\hat{a}^\dagger + \hat{a})$ with the Lamb-Dicke parameter $\eta = kx_0$, $x_0 = \sqrt{\hbar/2m\omega_m}$ representing the zero-point spread of the ground state wave function, and $\hat{a}^\dagger/\hat{a}$ representing the phonon creation/annihilation operators. With this expansion, the interaction Hamiltonian becomes,

$$H_I = \frac{\hbar\Omega}{2}\hat{\sigma}_+ \exp[i\eta(\hat{a}e^{-i\omega_m t} + \hat{a}^\dagger e^{i\omega_m t})]e^{-i(\mu t-\phi)} + h.c. \tag{3.5}$$

Expanding the exponent $\exp[\eta x] \approx 1 + \eta x + \frac{\eta^2}{2}x^2 \ldots$ up to first order of the Lamb-Dicke parameter, we can write:

$$\begin{aligned} H_I &\approx \frac{\hbar\Omega}{2}[(\hat{\sigma}_+ e^{-i(\mu t-\phi)} + \hat{\sigma}_- e^{i(\mu t-\phi)})(1 + i\eta(\hat{a}e^{-i\omega_m t} + \hat{a}^\dagger e^{i\omega_m t}))] \\ &= \frac{\hbar\Omega}{2}[(\hat{\sigma}_+ e^{-i(\mu t-\phi)} + \hat{\sigma}_- e^{i(\mu t-\phi)}) + i\eta(\hat{\sigma}_+ e^{-i(\mu t-\phi)} + \hat{\sigma}_- e^{i(\mu t-\phi)})(\hat{a}e^{-i\omega_m t} + \hat{a}^\dagger e^{i\omega_m t})] \end{aligned} \tag{3.6}$$

There are three particular cases of interest we can look at with $\mu = 0, \pm\omega_m$. In each case, we again perform the RWA approximation and ignore terms that are oscillating on the order of ω_m.

Case 1: $\mu = 0$

Here the Hamiltonian reduces to

$$H_{carrier} = \frac{\hbar\Omega_{nn}}{2}(\hat{\sigma}_+ e^{i\phi} + \hat{\sigma}_- e^{-i\phi}) \tag{3.7}$$

This operation describes a spin transition only corresponding to $|\downarrow\rangle|n\rangle \leftrightarrow |\uparrow\rangle|n\rangle$. This is called the carrier transition.

Case 2: $\mu = -\omega_m$

Here the Hamiltonian reduces to

$$H_{RSB} = \frac{\hbar\Omega_{n-1,n}}{2}(\hat{a}\hat{\sigma}_+ e^{i\phi} + \hat{a}^\dagger\hat{\sigma}_- e^{-i\phi}) \tag{3.8}$$

This operation corresponds to flipping the spin and removing one quanta of motional energy, i.e., $|\downarrow\rangle|n\rangle \leftrightarrow |\uparrow\rangle|n-1\rangle$. This is called the red sideband tran-

sition.

Case 3: $\mu = \omega_m$

Here the Hamiltonian reduces to

$$H_{BSB} = \frac{\hbar\Omega_{n+1,n}}{2}(\hat{a}^\dagger\hat{\sigma}_+ e^{i\phi} + \hat{a}\hat{\sigma}_- e^{-i\phi}) \tag{3.9}$$

This operation corresponds to flipping the spin and adding one quanta of motional energy, i.e., $|\downarrow\rangle|n\rangle \leftrightarrow |\uparrow\rangle|n+1\rangle$. This is called the blue sideband transition.

We point out that in all cases, there is a modified Rabi frequency given by

$$\begin{aligned} \Omega_{m,n} &= \Omega_0|\langle m|e^{i\eta(\hat{a}^\dagger+\hat{a})}|n\rangle| \\ &= \Omega_0 D_{m,n} \end{aligned} \tag{3.10}$$

Where $D_{m,n}$ is the Debye-Waller factor. Since $D_{m,n} = D_{n,m}$, assume without loss of generality that $n \geq m$, then this factor is given by:

$$D_{m,n} = e^{-\eta^2/2}\eta^{n-m}\sqrt{\frac{m!}{n!}}L_m^{n-m}(\eta^2) \tag{3.11}$$

Where L_m^{n-m} is the generalized Laguerre polynomial. In the Lamb-Dicke regime, where $\eta^2(2n+1) \ll 1$, these Debye-Waller factors for the modification of the Rabi frequency in the carrier, red sideband, and blue sideband transitions are shown in

Fig. 3.1 and can be summarized as follows.

$$\begin{aligned} \Omega_{carrier} &\simeq \Omega_0[1-(n+1/2)\eta^2] \\ \Omega_{RSB} &\simeq \Omega_0\eta\sqrt{n} \\ \Omega_{BSB} &\simeq \Omega_0\eta\sqrt{n+1} \end{aligned} \tag{3.12}$$

By assumption of the Lamb-Dicke regime, the correction to the carrier frequency is small, whereas the Rabi frequency of the red/blue sideband transition is reduced by the Lamb-Dicke parameter and depends more strongly on the motional quantum state, a fact that we can exploit to perform thermometry on the ion. Another feature to note is that when the ion is in its motional ground state ($n = 0$) then $\Omega_{RSB} = 0$ whereas $\Omega_{BSB} = \Omega_0\eta$. Intuitively, there is no more motional energy to remove in the ground state. We can exploit this asymmetry to perform sub-Doppler cooling of ions that relies on the existence of a dark state (see Section 5.1.2).

3.1.2: Raman Transitions

Despite the simple model presented above (one-ion, single motional mode, single CW laser beam, dipole transition), it can readily be extended to more complex situations. Encoding a qubit in a dipole-allowed transition would be a poor choice due to the typical lifetimes of the excited state ($\sim$ ns). Thus the electric dipole assumption is the first to be discarded. To optically address a hyperfine transition, as in $^{171}\mathrm{Yb}^+$, we use two laser beams with frequency ω_1, ω_2, each detuned from possibly multiple excited (dipole-allowed) states by Δ_i, with the frequency difference

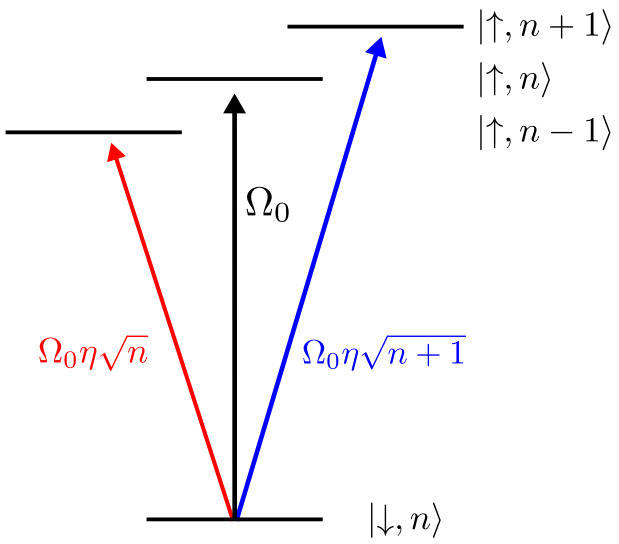

Figure 3.1: Simple model of a trapped ion showing the three basic operations we can achieve with a laser: a carrier transition (black), and red sideband transition (red) and a blue sideband transition (blue).

between the two beams tuned to the hyperfine qubit resonance frequency ω_0 plus some detuning μ, so that $\omega_1 - \omega_2 = \omega_0 + \mu$. When the resonance condition is satisfied (i.e., when the beat-note between the two lasers are tuned to hyperfine qubit frequency), the hyperfine states become coupled via a two-photon stimulated Raman transition. When the detuning from the excited state ($\sim \Delta_i$) is large enough, the excited state probabilities become negligible ($\sim 1/\Delta_i$). Since the time scale for excited state dynamics goes like Δ_i, compared to the typically much slower qubit state dynamics, which go like ($\sim \Omega$), the excited state can be adiabatically eliminated to create a coupling that is formally equivalent to the process described in Section 3.1.1. In our experiment, Δ_i is 33 and 67 THz, whereas Ω is ~ 500 kHz.

The electric fields of the two Raman beams, indexed by $j = 1, 2$, can be

described by:

$$\mathbf{E}_j = E_j \hat{\epsilon}_j \cos(\vec{k}_j \hat{x} + \omega_j t + \phi_j) \tag{3.13}$$

Making the following identifications:

$$\begin{aligned} \omega &\leftrightarrow \omega_1 - \omega_2 \\ \vec{k} &\leftrightarrow \Delta\vec{k} = \vec{k}_1 - \vec{k}_2 \\ \phi &\leftrightarrow \Delta\phi = \phi_1 - \phi_2 \end{aligned} \tag{3.14}$$

The carrier interaction Hamiltonian can be written almost identically to Eq. 3.7:

$$H_{carrier} = \frac{\hbar\Omega_0}{2}(\hat{\sigma}_+ e^{i\Delta\phi} + \hat{\sigma}_- e^{-i\Delta\phi}) + \delta^{(2)}\hat{\sigma}_z \tag{3.15}$$

We note that Raman transitions couple to motion along the direction defined by $\Delta\vec{k}$. If one wishes to maximize $\Delta\vec{k}$ (for example, to speed up two-qubit gates), the Raman beams are typically chosen to counter-propagate so that $|\Delta\vec{k}| = 2|\vec{k}|$ Additionally, there is a new term on the right hand side of the Hamiltonian that corresponds to an AC Stark shift (also know as the Autler-Townes effect) that shifts the qubit levels due to off-resonant coupling of the Raman beams to the excited states. The parameter $\delta^{(2)} = \delta^{(2)}_{\uparrow} - \delta^{(2)}_{\downarrow}$ characterizes the magnitude of the differential two-photon Stark shift between the two qubit states. The two-photon Stark shift

for each qubit state is defined as [48]:

$$\delta^{(2)}_{\uparrow/\downarrow} = \sum_j \left[\frac{E_j^2}{4\hbar^2} \sum_i \frac{|\langle \uparrow / \downarrow | \hat{d} \cdot \hat{\epsilon}_j | i \rangle|^2}{\Delta_{i,j}} \right] \tag{3.16}$$

So far, in the interaction Hamiltonian, we have rolled up the details of the laser polarization and the dipole coupling matrix into an effective Rabi frequency Ω. With Raman transitions, more consideration must be taken into the beam polarizations and it is worth explicitly writing down the expected Rabi frequency as [48]:

$$\Omega_0 = \frac{E_1 E_2}{4\hbar^2} \sum_i \frac{\langle \uparrow | \hat{d} \cdot \hat{\epsilon}_2 | i \rangle \langle i | \hat{d} \cdot \hat{\epsilon}_1 | \downarrow \rangle}{\Delta_i} \tag{3.17}$$

For the case of Raman transitions between hyperfine states of the $S_{1/2}$ manifold, where each Raman beam is tuned between the two excited $P_{1/2}$ and $P_{3/2}$ manifolds split by fine-structure (ω_{FS}) and detuned from the $P_{1/2}$ states by $\Delta_1 \approx \Delta_2 = \Delta$, the Rabi frequency is given by [49]:

$$\Omega_0 = P_{F,m_F \to F',m'_F} \frac{g_1 g_2}{3} \left(\frac{1}{\Delta} + \frac{1}{\Delta - \omega_{FS}} \right) \tag{3.18}$$

Where g_1, g_2 are the resonant one-photon Rabi frequencies of the each beam evaluated on the dipole element $\hat{d}_0$ of the cycling ${}^2S_{1/2}(F = 1, m_F = 1) \leftrightarrow {}^2P_{3/2}(F = 2, m_F = 2)$ transition [49]:

$$g_j = \frac{E_j}{2\hbar} |\hat{d}_0 \cdot \hat{\sigma}_+| \tag{3.19}$$

The prefactor in Eq. 3.18, $P_{F,m_F \to F',m'_F}$, contains all the information about the po-

larization of the Raman beams relevant to driving an $|F, m_F\rangle \rightarrow |F', m'_F\rangle$ transition in the $S_{1/2}$ ground state manifold. In general, the polarization of a laser beam can be expressed as a linear combination, $\hat{\epsilon}_i = e^i_- \hat{\sigma}_- + e^i_\pi \hat{\sigma}_\pi + e^i_+ \hat{\sigma}_+$, of the laser's left circular ($\hat{\sigma}_-$), linear ($\hat{\sigma}_\pi$), and right circular ($\hat{\sigma}_+$) polarization components with $|e^i_-|^2 + |e^i_\pi|^2 + |e^i_+|^2 = 1$. The polarization prefactor can be succinctly described by $P = i(\hat{\epsilon}_2^* \times \hat{\epsilon}_1) \cdot \hat{\mathbf{B}}$, where $\hat{\mathbf{B}}$ defines the quantization axis of the ion. The polarization prefactor can be enumerated for all the possible hyperfine transition as follows [50]:

$$\begin{aligned} P_{0,0\rightarrow 1,0} &= (e^0_- e^{1*}_- - e^0_+ e^{1*}_+) \\ P_{0,0\rightarrow 1,-1} &= -(e^0_- e^{1*}_\pi + e^0_\pi e^{1*}_+) \\ P_{0,0\rightarrow 1,1} &= (e^0_+ e^{1*}_\pi + e^0_\pi e^{1*}_-) \\ P_{1,0\rightarrow 1,-1} &= (e^0_- e^{1*}_\pi + e^0_\pi e^{1*}_+) \\ P_{1,0\rightarrow 1,1} &= (e^0_+ e^{1*}_\pi + e^0_\pi e^{1*}_-) \end{aligned} \tag{3.20}$$

For typical operations, we wish to maximize the qubit drive and minimize the coupling to the Zeeman levels to avoid qubit leakage errors. From the equations above, this condition can be enforced by ensuring that there is no π-light in the Raman beams ($\epsilon^0_\pi, \epsilon^1_\pi = 0$). Thus the polarization prefactor in most set-ups can be reduced to $P = e^0_- e^{1*}_- - e^0_+ e^{1*}_+$. This condition can be achieved by using linearly polarized light in both Raman beams that are perpendicular to each other and also to the magnetic field (colloquially called the "lin-perp-lin" configuration).

Continuing on in this specific case of hyperfine transitions with Raman beams

tuned between the $P_{1/2}$ and $P_{3/2}$ manifolds, we can write down the two-photon stark shifts for each laser beam j:

$$\begin{aligned}
\delta^{(2)}_{\downarrow,j} &= \frac{g_j^2}{3}\left(\frac{1}{\Delta} - \frac{1}{\Delta - \omega_{FS}}\right) \\
\delta^{(2)}_{\uparrow,j} &= \frac{g_j^2}{3}\left(\frac{1}{\Delta + \omega_0} - \frac{1}{\Delta - \omega_{FS} - \omega_0}\right) \\
\delta^{(2)}_{j} &= \delta^{(2)}_{\uparrow,j} - \delta^{(2)}_{\downarrow,j} \\
&\approx \frac{-g_j^2\omega_0}{3}\left(\frac{1}{\Delta^2} + \frac{1}{(\Delta - \omega_{FS})^2}\right)
\end{aligned} \tag{3.21}$$

Where the final approximation is due to $\omega_0 \ll \Delta, \omega_{FS}$. It is clear from Eq. 3.18 and 3.21 that for Raman beams tuned between the $P_{1/2}$ and $P_{3/2}$ manifolds, the couplings add constructively for the Rabi frequency, and subtract destructively for the 2-photon Stark shifts. In fact, there is a minimum in the 2-photon Stark shift relative to the Rabi frequency, $\delta^{(2)}/\Omega_0 \approx 2\times10^{-4}$ at $\Delta^* = \omega_{FS}/3$, which corresponds to a Raman wavelength of 355 nm [51]. This wavelength is also readily generated as the third harmonic of an Nd:YAG or Nd:YVO_4 laser, which means there are many commercially-available high-power sources, making this wavelength selection near ideal for $^{171}Yb^+$.

3.1.3: Raman Frequency Combs

In the previous section, we detailed how the phase of the Raman qubit drive, ϕ, depends on the phase difference between the two Raman beams, $\Delta\phi = \phi_1 - \phi_2$. In order to coherently control qubits over long gate sequences, this phase must be very stable. One way to generate two phase-locked laser beams is to split a

single laser beam with a beam-splitter and direct them to the ion along two arms of an interferometer. In this case, noise on $\Delta\phi$ will primarily result from optical path length fluctuations caused by acoustic vibrations on fast time-scales ($\sim$ kHz), air turbulence on medium time-scales ($\sim$ Hz) and thermal drifts on slow time-scales ($\sim$ mHz). Of course, in order to drive hyperfine transitions, the two beams must also have frequency components that are separated by the qubit frequency, ω_0. Generating this tone in a split-CW setup is challenging because AOMs do not operate at the required frequency ($\sim$ GHz) and EOMs have very low conversion efficiency, particularly in UV wavelengths. An alternative to splitting one CW laser beam, is to use two separate CW laser LOs running at a frequency difference near the hyperfine splitting. In this case, reliably locking the optical phase of the two beams at $\omega_0 = 12.6$ GHz to Hz level stability can be quite challenging and the environmental shielding requirements are more stringent because the two optical paths are now completely separate, compared to just the paths after the beam splitter (i.e. the interferometer is bigger).

A different solution is to use a pulsed Raman laser generated by a mode-locked optical frequency comb. Inside the laser cavity, a laser gain medium (in our case Nd:YVO_4) sets the gain bandwidth, $\Delta\nu_g$, that specifies the frequency range over which light is emitted from the medium. The laser cavity itself also supports discrete longitudinal laser modes that are spaced in frequency $\Delta\nu_c = \frac{c}{2L}$ set by the length of the cavity L. If $\Delta\nu_c \ll \Delta\nu_g$, then many laser modes are simultaneously supported within the cavity. So far the output of this laser is still CW (and multimode) and requires one more ingredient to generate pulses - a mode-locking element. The mode

locking element in our case is a passive Saturable Bragg Reflector (SBR) on one end of the cavity that becomes more reflective at higher laser intensities. Thus, the SBR will preferentially amplify high-intensity spikes that ultimately leads to all laser modes being phase locked to their maximum at the location of the SBR at a fixed point in time. As time evolves, this high-intensity spike travels spatially back and forth in the cavity and forms the pulse of our pulsed laser. This is all to highlight the point that mode-locked lasers can be thought of as an ensemble of phase-coherent CW laser beams. If any (or multiple) pairs of the CW beams in the ensemble are spaced by ω_0, then one can drive qubit transitions. As an added benefit, the large peak power of the pulses at the fundamental 1064 nm wavelength leads to high conversion efficiencies to UV third-harmonic 355 nm wavelength through nonlinear SHG and SFG processes.

Fortunately, the simple model for Raman transitions given in 3.1.2 can readily be extended to pulsed lasers. Instead of a constant electric field amplitude, we now consider one that is periodically modulated. The laser's time-varying electric field amplitude at a fixed point in space is given by [52]:

$$E_j(t) = E_j \sum_{n=1}^{N} p(t - n/f_{\text{rep}}) \tag{3.22}$$

where $p(t)$ is the pulse shape and f_{rep} is the repetition rate of the laser ($f_{\text{rep}} = \Delta\nu_c$). We assume for now that the Raman transition is being driven by two lasers with independent carrier frequencies and polarizations, but the difference between the two carrier frequencies is less than the common repetition rate. This can be accomplished

by splitting the output of a pulsed laser and passing the two beams through two AOMs with a frequency difference $f_A = (\omega_2 - \omega_1)/2\pi$ and two sets of polarization optics. In this setup, the phase of the Raman drive is also just set the by the phase difference of the two RF tones applied to the AOMs, $\Delta\phi = \phi_1 - \phi_2$.

In the regime where each laser pulse is weak and many pulses are required to drive the ion from $|\downarrow\rangle$ to $|\uparrow\rangle$, it is useful to analyze the effect of Raman driving in the frequency domain [53]. If the pulse width τ is much less than the repetition period $1/f_{\text{rep}}$, then the Fourier transform of the electric field amplitude is given by [52]:

$$\tilde{E}_j(\omega) = \sum_k \tilde{E}_{j,k}(\omega) = E_j \sum_k \delta(\omega - \omega_k)\ \tilde{p}(\omega - \omega_j) \tag{3.23}$$

where $\delta(\omega)$ is a sharply peaked function of width $\sim f_{\text{rep}}/N$. This describes an optical frequency comb whose teeth (indexed by k) are centered on the carrier frequency ω_j, spaced by the repetition rate $2\pi f_{\text{rep}}$, and modulated by the envelope function $\tilde{p}(\omega)$. This result is identical to our previous picture of many equally spaced CW laser modes supported simultaneously in the laser cavity and modulated by the laser gain bandwidth. For a mode-locked laser, the pulse shape is given by $p(t) = \sqrt{\pi/2}\ \text{sech}(\pi t/\tau)$ [52, 54], which results in an envelope function $\tilde{p}(\omega) = \tau\sqrt{\pi/2}\ \text{sech}(\omega\tau/2)$ and an RMS electric field amplitude:

$$\bar{E}_j = \left[\frac{f_{\text{rep}}}{N}\int_{-\infty}^{\infty} E_j\,(t)\,\mathrm{d}t\right]^{1/2} = \sqrt{f_{\text{rep}}\tau}\ E_j \tag{3.24}$$

Spectrally, the time-averaged electric field amplitude corresponding to the kth tooth

of the jth frequency comb is therefore

$$E_{j,k} = f_{\text{rep}} \int_{-\infty}^{\infty} |\tilde{E}_{j,k}(\omega)| \, \mathrm{d}\omega = E_j \; f_{\text{rep}} \; \tilde{p}(\omega_k - \omega_j) \tag{3.25}$$

This enables us to define the single-photon Rabi frequency $g_{j,k}$ due to the kth comb tooth, which is obtained by substituting $E_{j,k}$ for E_j in Eq. 3.19.

As long as the pulse width τ is short enough so that the frequency comb's bandwidth is larger than the qubit splitting ω_0, then pairs of teeth from each comb can be brought into resonance in order to drive the Raman transition. The resonance condition is therefore

$$|f_A + n f_{\text{rep}}| = \omega_0 / 2\pi \tag{3.26}$$

where n is an integer. Inserting $g_{j,k}$ into Eq. 3.18, summing over all resonant comb pairs, and assuming $f_{\text{rep}}\tau \ll 1$, we find the Raman Rabi frequency [52],

$$\begin{aligned} \Omega_n &= \sum_k P \; \frac{g_{1,k} \, g_{2,k+n}}{3} \left(\frac{1}{\Delta} + \frac{1}{\Delta - \omega_{FS}} \right) \\ &\approx \Omega_0 \; h_n \end{aligned} \tag{3.27}$$

due to this resonant comb tooth, where we have defined the comb-independent Rabi frequency

$$\Omega_0 = P \; \bar{E}_1 \bar{E}_2 \; \frac{|\mathbf{d}_0 \cdot \hat{\sigma}_+|^2}{12 \; \hbar^2} \left(\frac{1}{\Delta} + \frac{1}{\Delta - \omega_{FS}} \right) \tag{3.28}$$

and the envelope function,

$$h_k = k\pi f_{\text{rep}}\tau \ \text{csch}\left(k\pi f_{\text{rep}}\tau\right) \tag{3.29}$$

which has the property $h_0 = 1$.

Equations 3.28 and 3.29 cleanly divide the factors that determine the Raman Rabi frequency: the envelope function h_k is a simple function containing only the frequency comb's temporal characteristics—its pulse length and repetition rate—while the comb-independent Rabi frequency Ω_0 contains the parameters that describe any laser's interaction with a set of dipole-allowed atomic transitions—the transitions' dipole moments and the laser's polarization, intensity, and single-photon detunings. In this sense, using a pulsed laser is very similar to the model presented in Section 3.1.2 with a few caveats. Note that the Rabi frequency in Eq. 3.28 is proportional to the time-averaged electric field of the two Raman beams, which is a quantity that is convenient to measure experimentally, but somewhat obscures the fact that the pulses must be overlapped in time in order to drive the qubit transition. In practice this is easily accomplished by including a motorized delay stage on one arm of interferometer. The second caveat is that the frequency combs contain other frequency components that are off-resonant from the qubit transition and can cause four-photon Stark shifts, as discussed in the following section.

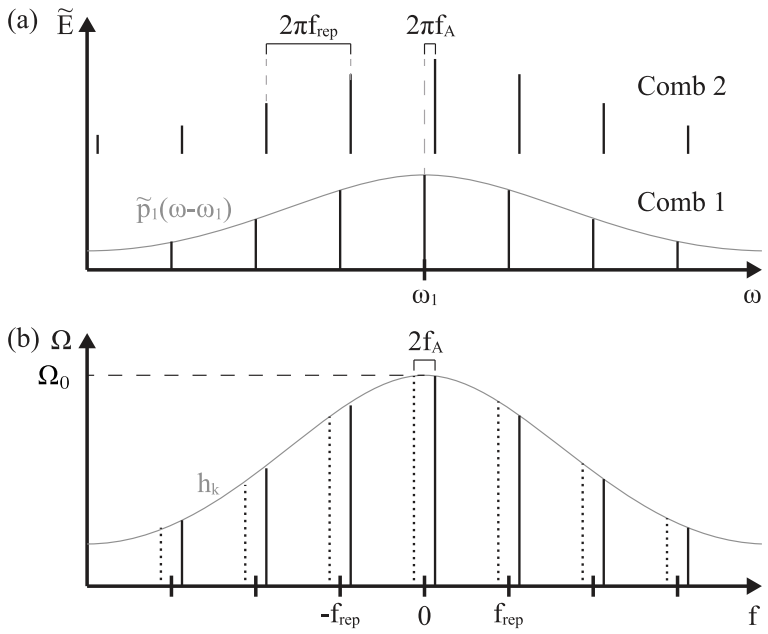

Figure 3.2: Schematic representations of the (a) optical frequency combs and (b) beatnote comb. The solid (dotted) teeth in the beatnote comb represent co(counter)-rotating terms. The curvatures of the envelope functions have been exaggerated relative to the comb spacings.

3.1.4: Four-Photon Stark Shifts

An intuitive picture of pulsed Raman manipulation is depicted in Fig 3.2. The two optical combs, are both centered on the optical carrier frequency ω_1 and are offset by the AOM difference frequency $2\pi f_A$. The beatnotes between the two optical combs create an RF beatnote comb, which is centered on zero and extends to frequencies of order $\sim 1/\tau$. The beatnote comb is naturally expressed in units of Ω_0, and its envelope is h_k.

There are two sets of teeth in the beatnote comb: the co-rotating teeth correspond, intuitively, to photon absorption from comb 2 and emission into comb 1 and appear at $kf_{\text{rep}} + f_A$, while the counter-rotating teeth correspond to absorp-

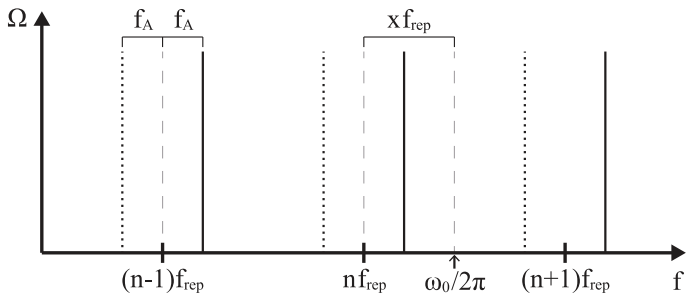

Figure 3.3: Schematic diagram of the near-resonant beatnote comb teeth, illustrating the relationship between x and f_A. As in Fig. 3.2, the solid (dotted) teeth in the beatnote comb represent co-(counter-)rotating terms.

tion from comb 1 and emission into comb 2 and appear at $kf_{\text{rep}} - f_A$. In order to drive the Raman transition from $|\downarrow\rangle$ to $|\uparrow\rangle$, f_{rep} and f_A must be set to place one beatnote comb tooth at $\omega_0/2\pi$, which is equivalent to the resonance condition given in Eq. 3.26. The other beatnote comb teeth, which are off-resonant and do not drive the transition, give rise to a four-photon Stark shift. These four-photon Stark shifts scale as $\sim \Omega^2/\Delta$ and can be $\times 100$ larger [50] than the two-photon Stark shifts discussed in Section 3.1.2. These shifts are somewhat unique to using the optical frequency comb approach. We can neglect Stark shifts due to the Zeeman states because the Raman Rabi frequency between $|\downarrow\rangle$ and those states is zero if there is no π component to either Raman laser's polarization (Eq. 3.20), as is the case a "lin-perp-lin" configuration.

A change in the laser parameter f_{rep} has the effect of shifting the frequencies of the beatnote comb teeth, thereby changing the resulting Stark shifts. We parameterize the variable f_{rep} with the dimensionless parameter,

$$x \equiv \frac{\omega_0/2\pi \bmod f_{\text{rep}}}{f_{\text{rep}}} \tag{3.30}$$

which ranges from 0 to 1. The resonance condition in Eq. 3.26 becomes simply,

$$f_A = (n \pm x)\, f_{\text{rep}} \tag{3.31}$$

for driving the qubit transition, which is illustrated in Fig. 3.3. Although n can be any integer, typically $|f_A| < f_{\text{rep}}$ because of the finite bandwidths of the AOMs.

The Stark shift on the qubit transition frequency due to the kth comb tooth is [50],

$$\delta_k^{(4)} = -\frac{|\Omega_k|^2}{2\Delta_k} \tag{3.32}$$

where Ω_k is given by Eq. 3.27 and Δ_k is the detuning of the kth tooth away from the qubit transition. Summing over all co- and counter-rotating comb teeth, we find a total Stark shift

$$\delta^{(4)} = \sum_k \delta_k^{(4)} = \delta_0^{(4)} \sum_k h_k^2 \left(\frac{1}{k - n - x + f_A/f_{\text{rep}}} + \frac{1}{k - n - x - f_A/f_{\text{rep}}} \right), \tag{3.33}$$

where we define the comb-independent Stark shift

$$\delta_0^{(4)} = -\frac{|\Omega_0|^2}{4\pi f_{\text{rep}}} \tag{3.34}$$

We note that Eq. 3.33 holds whether one comb tooth is resonant with the qubit transition or not, but we implicitly omit the resonant term, if there is one, from the sum. Although $\delta_0^{(4)}$ does depend on f_{rep}, the integer n, which specifies the pair

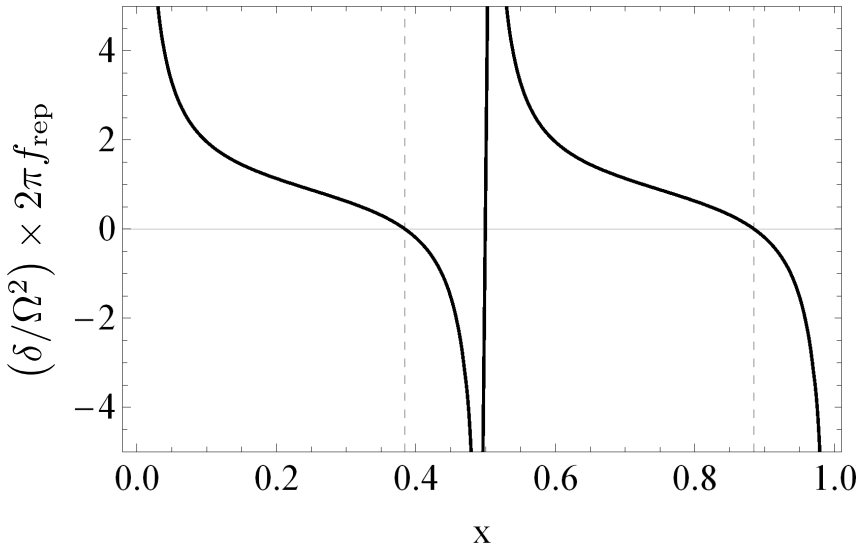

Figure 3.4: The 4-photon Stark shift versus f_{rep} (parameterized by x). The vertical lines indicate where the plot crosses zero.

of comb teeth immediately below resonance, is typically 107 in this experiment. Therefore, varying x from 0 to 1 changes f_{rep} by less than 1%, so δ_0 can effectively be treated as a constant prefactor.

While it is possible to drive transitions using only one Raman beam that contains either one comb ("resonant operation" with $x = 0/1$) or two combs ("phase insensitive operation" since $\Delta k = 0$), both cases require elliptical or circular polarization. This complicates the picture a bit, since a single comb can interfere with itself to create *intra*-beam Stark Shifts. In this experiment, with the polarizations set to be linear in both beams, Raman transitions are driven with a single comb in both beams ("phase-sensitive operation") and Stark shifts arise only from *inter*-beam (i.e., cross-beam) interference.

In Fig. 3.4, we compare the ratio of the Stark shift given in Eq. 3.33 to the Rabi frequency given by $\Omega = \Omega_0\ h_n$, over a variety of f_{rep} values corresponding

to $x \in [0, 1]$. Interestingly, there are particular values of f_{rep} at which the four-photon Stark Shift is completely nulled, which reduces pulsed Raman operation to the simplicity of CW Raman operation. Since f_{rep} depends on the cavity length inside the laser, this is typically set during manufacturing and cannot be specified by the customer at purchase. It is clear that the four-photon Stark shift can change dramatically with f_{rep}, leading to "good" or "bad" lasers resulting from probabilistic f_{rep} values sampled over the manufacturing distribution. In order to take control over this parameter, we modify the laser to place the output coupler of the laser cavity on a movable stage, thus allowing us to tune the length of the cavity (Section 4.4.1) and null the four-photon Stark shift (Section 5.2.1).

3.2: Digital Single Qubit Gates

In driving single qubit interactions, coupling to any sort of motion is undesirable, as this can cause leakage out of the qubit subspace by entangling spin with motion. Therefore, we use the carrier transition Hamiltonian given in Eq. 3.7, whose operator has the form $(\hat{\sigma}_+ e^{i\phi} + \hat{\sigma}_- e^{-i\phi})$. For $|0\rangle = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ and $|1\rangle = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$, the spin lowering and raising operators are given by,

$$\hat{\sigma}_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{3.35}$$

$$\hat{\sigma}_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{3.36}$$

Thus it is easy to see that when $\phi = 0$ and $\phi = -\pi/2$, one can easily create

the digital Pauli $\hat{X}$ and $\hat{Y}$ gates.

$$(\hat{\sigma}_+ e^{i0} + \hat{\sigma}_- e^{-i0}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \hat{X} \tag{3.37}$$

$$(\hat{\sigma}_+ e^{i(-\pi/2)} + \hat{\sigma}_- e^{-i(-\pi/2)} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \hat{Y} \tag{3.38}$$

This is for the specific case in which $\theta = \Omega t = \pi$, commonly called a "π-pulse". More generally, θ and ϕ are continuous parameters controlled by the duration of the carrier pulse and the phase difference between the two Raman beams. In the more general case, this Hamiltonian creates a unitary of the form,

$$\hat{R}(\theta, \phi) = \begin{pmatrix} \cos(\frac{\theta}{2}) & -i\sin(\frac{\theta}{2})e^{-i\phi} \\ -i\sin(\frac{\theta}{2})e^{i\phi} & \cos(\frac{\theta}{2}) \end{pmatrix} \tag{3.39}$$

This unitary describes a rotation in the Bloch sphere of arbitrary angle about an arbitrary axis in the xy-plane and is native to the trapped ion system. Commonly used are rotation axes aligned with the x and y axis, so it is worthwhile to explicitly define $\hat{R}_x(\theta)$ and $\hat{R}_y(\theta)$ as,

$$\hat{R}_x(\theta) = \begin{pmatrix} \cos(\frac{\theta}{2}) & -i\sin(\frac{\theta}{2}) \\ -i\sin(\frac{\theta}{2}) & \cos(\frac{\theta}{2}) \end{pmatrix} \tag{3.40}$$

$$\hat{R}_y(\theta) = \begin{pmatrix} \cos(\frac{\theta}{2}) & -\sin(\frac{\theta}{2}) \\ \sin(\frac{\theta}{2}) & \cos(\frac{\theta}{2}) \end{pmatrix} \tag{3.41}$$

The single qubit gates $\hat{R}_x(\theta)$ and $\hat{R}_y(\theta)$ are universal by themselves, but is

also useful to define a rotation about the z-axis,

$$\hat{R}_z(\theta) = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix} \tag{3.42}$$

This gate can be implemented physically and natively in the ion trap system by detuning far from the carrier/motional modes and using Stark shifts from the Raman beams. On the other hand, this operation can also be implemented virtually by simply advancing the phase of any subsequent RF pulses by θ for whichever ion acquired the $\hat{R}_z(\theta)$ gate. This approach requires one to keep track of all the ion phases in the chain and be able to shift them independently, which in this system is done automatically at the waveform compiling step and then applied to the individual addressing beams. Virtual gates are strongly preferred, as they do not consume any wall-clock time in the circuit and their accuracy is only limited by the phase resolution/bit depth of the RF source.

Single qubit gates are fast ($\sim$ 1 μs) relative to their two-qubit counterparts ($\sim$ 225 μs), which means we can perform extra rotations as a means to mitigate errors with minimal time cost relative to the total circuit. Such sequences are known as composite pulse sequences in which a single imperfect pulse with some error ϵ is replaced with a sequence of pulses, which reduces the error to $O(\epsilon^n)$, where n can be chosen arbitrarily at the cost of extra rotation angle [55]. There are many such sequences designed to suppress against amplitude, frequency, or phase errors. On the time scales of single qubit gates, amplitude errors are the dominant source of noise in this system due to either slow alignment drifts of the tightly focused

individual beams or fast motion of the ions (Sec. 6.1).

The two main pulses used in ion trap quantum computers are the BB1 pulse [56] and the SK1 pulse [55]. Both suppress amplitude errors on arbitrary input states, however the SK1 pulse also suppresses small amplitude errors on neighboring idle qubits due to Rabi crosstalk (Sec. 5.3.2). For a first-order correction, the SK1 pulse consists of three pulses [55],

$$\begin{aligned} \mathrm{SK1}(\theta,\phi)|\psi\rangle &= R_2(2\pi,\phi_2)R_1(2\pi,\phi_1)R(\theta,\phi)|\psi\rangle \\ \phi_1 &= \phi - \arccos\left(\frac{\theta}{4\pi}\right) \\ \phi_2 &= \phi + \arccos\left(\frac{\theta}{4\pi}\right) \end{aligned} \tag{3.43}$$

Where R is the desired pulse (plus error) and R_1/R_2 are the correction pulses, with the one caveat that if $\theta = 0$, then the correction pulses should obviously not be performed. In this work we use SK1 pulses and benchmark their performance on a 15-ion chain in Section 5.3.3. Additionally, we apply a shaped Gaussian envelope to the entire composite SK1 pulse, rather than use square pulses, to avoid excitation of axial motion due to prompt charge response. For a shaped SK1 pulse, the total integrated area under the Gaussian curve should be $A = \theta + 4\pi$, with the appropriately timed phase flips delimiting the different pulses in the sequence.

3.3: Digital Two Qubit Gates

The prototypical entangling gate is the CNOT gate, which flips a target qubit, depending on the state of a control qubit. If the control qubit is in a superposition

state, then this operation readily generates entanglement. While the CNOT is easy to understand and model from a theoretical standpoint, it is generally hard to implement physically due its asymmetric nature (i.e., there is a distinction between the control qubit and the target qubit). The Ciraq-Zoller gate [57] uses the asymmetry of the motional ground state, where there is no RSB transition available for $|0\rangle$, to implement the asymmetry of the CNOT gate. Therefore, the gate requires high-fidelity ground state cooling and is not robust to heating of the modes that may occur throughout a long circuit sequence.

One could imagine that symmetrical gates need not require ground state cooling and are more natural to implement physically in an ion trap system over a range of possible motional states. The CNOT gate can be symmetrized with single qubit gates only, which makes it equivalent from the perspective of a universal gate set. In particular, if the state vector is written as $\{c_{00}, c_{01}, c_{10}, c_{11}\}$ in the z-basis, then the CNOT gate is written as,

$$\text{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{3.44}$$

Then with the following transformation, we arrive at a symmetrical matrix representing a ZZ entangling operation, which applies a phase shift to the coefficients of the anti-symmetric spin states in the two-qubit state vector.

$$ZZ = [R_z(\pi/2) \otimes R_z(\pi/2)\, H] \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} [I \otimes H] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.45}$$

An XX interaction, which is implemented by the Mølmer-Sørensen interaction and described in the following section, looks identical to ZZ when the state vector is represented in the x-basis. To see how an XX interaction transforms a state in the z-basis, we can rotate the x-axis up to z, perform the ZZ interaction, and then return the x-axis back to its origin.

$$[R_y(\pi/2) \otimes R_y(\pi/2)] \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} [R_y(-\pi/2) \otimes R_y(-\pi/2)] \tag{3.46}$$

$$= \frac{e^{i\pi/4}}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 & -i \\ 0 & 1 & -i & 0 \\ 0 & -i & 1 & 0 \\ -i & 0 & 0 & 1 \end{pmatrix} \tag{3.47}$$

This operation corresponds to a collective spin flip, $|00\rangle \rightarrow \frac{e^{i\pi/4}}{\sqrt{2}} (|00\rangle - i|11\rangle) \rightarrow |11\rangle$. This highlights the duality that a phase-shift in the x-basis corresponds to a spin flip in the z-basis. Additionally, we point out that the XX interaction can be wrapped in single-qubit gates to create an effective ZZ interaction, which can be used to operate gates in the "phase-insensitive" configuration that is immune interferometric optical phase noise [19, 58].

3.3.1: Mølmer-Sørensen Interaction

In this section we describe the Mølmer-Sørensen interaction and show how it implements an XX entangling operation mediated by the shared motional bus of the ion chain. One obvious benefit to this approach is that because the motional modes are global with respect to the chain, the entangling operations have all-to-all

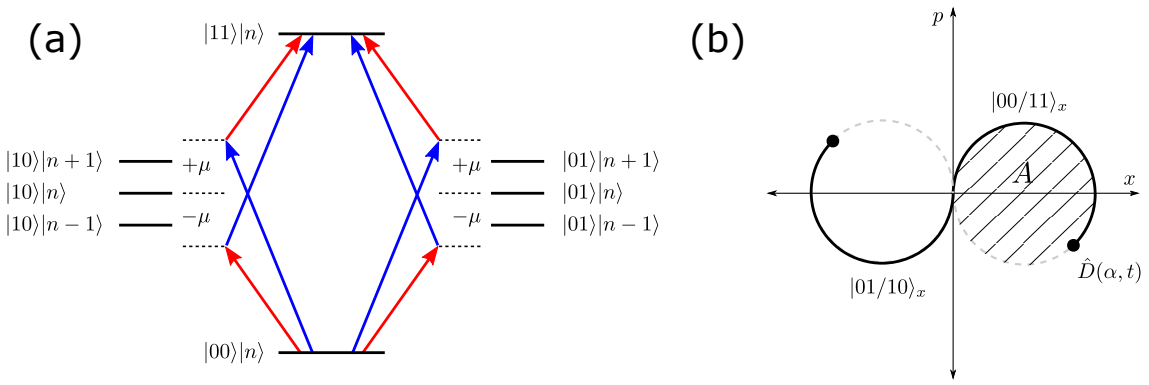

Figure 3.5: The Mølmer-Sørensen interaction consists of applying symmetrical red/blue tones near the motional modes. (a) Weak-field coupling of the MS interaction can be thought of a resonance condition between virtual excited modes. The dependence on the motional state n on the red/blue sideband factors out when both paths are considered. (b) General-field coupling creates a spin-dependent force that splits the even/odd parity spin states in phase-space. When motion is closed at the end of the interaction, the spin states acquire a geometric phase shift proportional to the total area (A) enclosed in phase-space.

connectivity. A downside to all-to-all connectivity is that parallel entangling gates are more difficult to implement, since the MS interaction is pair-wise over all the illuminated ions, $\sum_{i,j} X_i X_j$. For example if one wishes to implement gates X_1X_2 and X_3X_4 in parallel, the cross terms $\{X_1X_3, X_1X_4, X_2X_3, X_2X_4\}$ must be nulled to avoid driving unwanted interactions. While this can be accomplished through laser pulse-shaping, demonstrations so far have shown a decrease in fidelity relative to serial operations [59, 60].

Another complication with coupling to the motional bus to generate entanglement is that the Rabi frequency of the RSB/BSB depends on the motional state (Sec. 3.1.1). Mølmer and Sørensen are credited with first devising a scheme that is insensitive to the motional state [61, 62]. The scheme involves applying a bichromatic tone, corresponding to a simultaneous blue and red sideband, symmetrically

detuned from the carrier transition by $\pm\mu$ and detuned from the motional modes by $\pm\delta = |\mu| - \omega_k$. In the weak-coupling regime, where $n+1/n-1$ motional states are negligibly excited ($\eta\Omega \ll \delta$), the MS interaction uses resonance conditions within virtual excited motional states to create collective spin flips. The fact that this is insensitive to the motional state can be observed by examining Fig. 3.5(a) and considering the second-order perturbation [63],

$$\widetilde{\Omega} = 2 \sum_{m=n\pm1} \frac{\langle 11, n|H_I|m\rangle\langle m|H_I|00, n\rangle}{\delta} = \frac{(\eta\Omega)^2}{\delta} \tag{3.48}$$

Traversing the $(n-1)$ or $(n+1)$ path, we pick up a factor of (n) or $(n+1)$ from the Debye-Waller factors, respectively. However the δ from each path has opposite sign, so when adding these two paths together, the dependence on n is canceled in the effective Rabi frequency. The result is the same for the odd-parity states $\{|01, n\rangle, |10, n\rangle\}$. The weak-field regime is often used in quantum simulation experiments to model Ising chains with a tunable power-law coupling parameter [64]. In the general-field regime, the MS interaction excites motion via a spin-dependent force, which strongly entangles spin with motion. As shown in Fig. 3.5(b), the even parity states are driven in one direction in the QHO phase-space and the odd parity states in the other direction. As long as the motion is closed at the end of the interaction, then the two parity spin states acquire a relative geometric phase proportional to the total area enclosed, with no residual entanglement to the motion.

To understand the MS interaction in the general-field regime, we start from the interaction Hamiltonian given in Equation 3.4 and consider applying two tones

(blue/red) with symmetric detuning $\pm\mu$ from the carrier transition [65],

$$\begin{aligned}
H_I &= \frac{\hbar\Omega}{2}\hat{\sigma}_+\left(e^{-i(\Delta k\hat{x}+\mu t+\Delta\phi_b)}+e^{i(\Delta k\hat{x}-\mu t+\Delta\phi_r)}\right)+h.c. \\
&= \frac{\hbar\Omega}{2}\hat{\sigma}_+\left(e^{-i(\Delta k\hat{x}+\mu t+\phi^{(m)}+\phi^{(s)})}+e^{i(\Delta k\hat{x}-\mu t-\phi^{(m)}+\phi^{(s)})}\right)+h.c. \\
&= \hbar\Omega\left(\cos\mu t+\phi^{(m)}-\Delta k\hat{x}\right)\left(e^{i\phi^{(s)}}\hat{\sigma_-}+e^{-i\phi^{(s)}}\hat{\sigma_+}\right) \\
&= \hbar\Omega\left(\cos\mu t+\phi^{(m)}-\Delta k\hat{x}\right)\left(\hat{\sigma_x}\cos\phi^{(s)}-\hat{\sigma_y}\sin\phi^{(s)}\right) \\
&= \hbar\Omega\hat{\sigma_n}\left(\cos\mu t+\phi^{(m)}-\Delta k\hat{x}\right)
\end{aligned} \tag{3.49}$$

Where we have defined the spin and motion phase as $\phi^{(s)} = (\Delta\phi_b + \Delta\phi_r)/2$ and $\phi^{(m)} = (\Delta\phi_b - \Delta\phi_r)/2$ and defined $\hat{\sigma_n} = \left(\hat{\sigma_x}\cos\phi^{(s)} - \hat{\sigma_y}\sin\phi^{(s)}\right)$. For now we will assume $\phi^{(s)} = 0$ corresponding to an XX interaction, although this can easily be adjusted to YY by setting $\phi^{(s)} = -\pi/2$ (or along any other axis in the equator). Expanding the position of ion j as in Eq. 3.5, but now considering all the normal motional modes of a chain, with the Lamb-Dicke parameter η_k, the mode participation vector b_j^k, the mode frequency ω_k, and the annihilation/creation operators $\hat{a}_k/\hat{a}_k^\dagger$ all indexed by motional mode k,

$$H_I = \hbar\Omega_j\hat{\sigma}_j^x\left[\cos\mu t+\phi_j^{(m)}-\sum_k\eta_k b_j^k(\hat{a}_k e^{-i\omega_k t}+\hat{a}_k^\dagger e^{i\omega_k t})\right] \tag{3.50}$$

As in Section 3.1.1, we can expand this expression according to the power of

η_k to give [65],

$$H_I = \hbar\Omega_j\hat{\sigma}_j^x \Bigg[\cos(\mu t + \phi_j^{(m)}) + \sin(\mu t + \phi_j^{(m)}) \sum_k \eta_k b_j^k (\hat{a}_k e^{-i\omega_k t} + \hat{a}_k^\dagger e^{i\omega_k t}) - \frac{1}{2}\cos(\mu t + \phi_j^{(m)}) \sum_k \sum_l \eta_k \eta_l b_j^k b_j^l (\hat{a}_k e^{-i\omega_k t} + \hat{a}_k^\dagger e^{i\omega_k t})(\hat{a}_l e^{-i\omega_l t} + \hat{a}_l^\dagger e^{i\omega_l t})\Bigg] \tag{3.51}$$

The first term is a single-qubit interaction, which is far off-resonant and can be ignored in the case of $\mu \approx \omega_k \approx 3$ MHz as in the situation of two-qubit gates on the radial modes. When this interaction is applied to multiple ions, we simply sum H_I over the ion indices j. The unitary that describes this interaction can be obtained by using the Magnus expansion, which terminates exactly after the first two terms [66],

$$\begin{aligned} U(\tau) &= \exp\left[-i\int_0^\tau dt H_I(t) - \frac{1}{2}\int_0^\tau dt_1 \int_0^{t_1} dt_2 [H_I(t_1), H_I(t_2)]\right] \\ &= \exp\left[i\sum_j \phi_j(\tau)\hat{\sigma}_j^x + i\sum_{i<j} \Theta_{ij}(\tau)\hat{\sigma}_i^x\hat{\sigma}_j^x\right] \end{aligned} \tag{3.52}$$

The first summation corresponds to spin-phonon coupling with the following definitions [65],

$$\phi_j(\tau) = -i\sum_k \left(\alpha_j^k(\tau)\hat{a}_k^\dagger - \alpha_j^{k*}(\tau)\hat{a}_k\right) \tag{3.53}$$

$$\alpha_j^k(\tau) = -\frac{i}{\hbar}\eta_k b_j^k \int_0^\tau \chi_j(t) e^{i\omega_k t} dt \tag{3.54}$$

$$\chi_j(t) = \hbar\Omega_j(t)\sin\left(\mu(t)t + \phi_j^{(m)}\right) \tag{3.55}$$

Note that $\phi_j(\tau)$ has exactly the form of the displacement operator $\hat{D}(\alpha) = \exp\left[\alpha\hat{a}^\dagger - \alpha^*\hat{a}\right]$, which describes a force that depends on the spin ($\hat{\sigma}_j^x$), as introduced earlier. In order to close motion and disentangle spin from motion at the end of the interaction, we must enforce $\alpha_j^k(\tau) = 0$ for each motional mode k and each ion j. In short chains this may be achieved by simply tuning the length of the interaction (τ), although in longer chains with more motional modes this may not be possible. Thus, we have explicitly allowed the Rabi frequency $\Omega_j(t)$ and the detuning $\mu(t)$ to be time-dependent in the definition of $\chi_j(t)$, corresponding to the other two control knobs in the system.

The second summation is the spin-spin interaction that is of interest for entangling gates and corresponds to the phase-space area enclosed for each motional mode [65],

$$\Theta_{ij}(\tau) = \frac{2}{\hbar^2}\sum_k \eta_k^2 b_i^k b_j^k \int_0^\tau dt_1 \int_0^{t_1} dt_2 \; \chi_i(t_1)\chi_j(t_2)\sin\left[\omega_k(t_1 - t_2)\right] \tag{3.56}$$

When $\Theta_{ij}(\tau) = \pi/4$ at the end of the operation, we obtain the matrix that looks like Eq. 3.45 when the state vector is represented in the x-basis. The fidelity of this MS operation (averaged over a thermal distribution of motional states with

temperature T) relative to the ideal XX unitary matrix is given by [65],

$$\overline{F} = \frac{1}{10}\left[4 + 2(\Gamma_i + \Gamma_j)|\sin(2\Theta_{ij})| + \Gamma_+ + \Gamma_-\right] \tag{3.57}$$

$$\Gamma_{i(j)} = \exp\left[-2\sum_k |\alpha^k_{i(j)}|^2 \coth\left(\frac{\hbar\omega_k}{2k_BT}\right)\right] \tag{3.58}$$

$$\Gamma_\pm = \exp\left[-2\sum_k |\alpha^k_i \pm \alpha^k_j|^2 \coth\left(\frac{\hbar\omega_k}{2k_BT}\right)\right] \tag{3.59}$$

Thus when $\alpha^k_j(\tau) = 0$ (for all ions j) and $\Theta_{ij} = \pm\pi/4$, then $\overline{F} = 1$, corresponding to the two conditions, (1) spin is disentangled from motion at the end of the gate and (2) the total area enclosed in phase-space creates the differential phase-shift required for a maximally-entangling operation like that described in Eq. 3.45. These constraints are general for both serial and parallel gate operations. In the case of parallel gates, the first condition remains unchanged, however the second condition must be slightly modified so that the intended two-qubit interactions have $\Theta_{ij} = \pm\pi/4$, while the unwanted two-qubit interactions have $\Theta_{ij} = 0$.

We also note that the sign of Θ_{ij} is allowed to be either positive or negative, as both can be readily compiled into a CNOT gate through slightly different single-qubit gate operations [67] and the unconstrained sign allows more flexibility in the gate design. Finally, we note that some quantum algorithms may make use of less than fully entangling operations, $|\Theta_{ij}| < \pi/4$. In AM gate schemes, this can be achieved by simply scaling the solution for $|\Theta_{ij}| = \pi/4$, although it may save time-cost to compute a second set of solutions for a smaller angle (e.g., $|\Theta_{ij}| = \pi/8$) as small-angle gates can in general run faster. In the following section, we describe

pulse-shaping techniques to satisfy these constraints under serial gate operations between just two ions.

3.3.2: AM Laser Pulse Shaping

The three control knobs available within a gate waveform to achieve high gate fidelity are the total gate duration (τ_{gate}), the laser power at the ions $\Omega(t)$ (i.e., amplitude modulation), and the gate detuning $\mu(t)$ (i.e., frequency/phase modulation). The gate duration τ_{gate} does not have enough fine control for long chains with many motional modes, and should always be minimized within the bounds of achievable laser power to minimize errors due to heating and other decoherence sources. There have been many demonstrations of AM gates [20, 68], FM gates [69, 70], PM gates [71], multi-tone gates [72, 73], and combinations thereof (e.g., AM-PM [74], AM-FM [75, 76]). For the work presented in this thesis, we use AM gates, which have been successful in predecessor experiments [77]. We present a slight modification of this pulse-shaping scheme to use linear interpolation to smooth out impulse jumps between different amplitude segments to avoid excitation of axial motion due to prompt charge response (Section 2.1.2). In Section 5.3.4 we characterize AM-interpolated gates on a chain of 15 ions and demonstrate high-fidelity operations that appear to be limited by noise sources other than the gate solution itself. While FM/PM gates may be preferable to use in the future, we expect the primary benefit to be slightly faster gates. FM gates make better utilization of the available laser power, as AM gates are necessarily de-rated from the maximum Rabi frequency in order to achieve control.

For discrete AM gates, assume the laser sequence is divided into n_{seg} equal length segments over the total gate length τ_{gate}, each with a constant amplitude and the same fixed detuning μ. Then $\mathbf{\Omega} = (\Omega_1, \Omega_2, \ldots, \Omega_{n_{seg}})$ is a column vector describing the Rabi frequency of each segment, which we wish to optimize. The start and stop times of the n-th segment is given by $\tau_{1,n} = (n-1)\tau_{gate}/n_{seg}$ and $\tau_{2,n} = n\tau_{gate}/n_{seg}$. Because Ω_n is constant, it can be pulled out of the integral in Eq. 3.54 and the spin-motion term is then given by [65],

$$\alpha_j^k = \mathbf{A}_j^k \mathbf{\Omega} \tag{3.60}$$

$$\mathbf{A}_j^k(n) = -i\eta_k b_j^k \int_{\tau_{1,n}}^{\tau_{2,n}} \sin(\mu t) e^{i\omega_k t} dt \tag{3.61}$$

Where $\mathbf{A}_j^k(n)$ is a row vector whose n-th component describes the net displacement caused by the n-th waveform segment on ion j in the motional mode k. Similarly for the spin-spin term, $\Omega_p\Omega_q$ is constant and can be pulled out of the double integral in Eq. 3.56, which can now be expressed as a matrix given by [65],

$$\Theta_{ij} = \mathbf{\Omega}^T \boldsymbol{\gamma} \mathbf{\Omega} \tag{3.62}$$

$$\boldsymbol{\gamma}(p,q) = \begin{cases} \sum_k \eta_k^2 b_i^k b_j^k \int_{\tau_{1,p}}^{\tau_{2,p}} dt_1 \int_{\tau_{1,q}}^{\tau_{2,q}} dt_2 \sin(\mu t_1)\sin(\mu t_2)\sin\left[\omega_k(t_1-t_2)\right] & (p>q) \\ 2\sum_k \eta_k^2 b_i^k b_j^k \int_{\tau_{1,p}}^{\tau_{2,p}} dt_1 \int_{\tau_{1,p}}^{t_1} dt_2 \sin(\mu t_1)\sin(\mu t_2)\sin\left[\omega_k(t_1-t_2)\right] & (p=q) \\ \boldsymbol{\gamma}(q,p) & (p<q) \end{cases} \tag{3.63}$$

The requirement that motion is closed at the end of the gate (i.e., $\alpha_j^k = 0$ for

all motional modes $1 \leq k \leq N$ and for each ion indexed by $j = 0, 1$) can be simply summarized as follows,

$$\begin{pmatrix} \mathbf{A}_0^1 \\ \mathbf{A}_0^2 \\ \vdots \\ \mathbf{A}_1^N \end{pmatrix} \mathbf{\Omega} = \vec{0} \tag{3.64}$$

Equation 3.64 is a homogeneous system of $2N$ linear equations with n_{seg} unknowns. If $n_{seg} = 2N$, then there is a unique solution for $\mathbf{\Omega}$ that can be calculated exactly. Since Eq. 3.64 is linear with respect to $\mathbf{\Omega}$, if $\mathbf{\Omega}_0$ is a solution then $a\mathbf{\Omega}_0$ is also a solution with arbitrary scale factor a. Thus as long as motion is closed and $\Theta_{ij} \neq 0$, then the solution can be scaled to achieve the desired entangling gate angle (within the bounds of available laser power). If $n_{seg} > 2N$, then the system is under-determined and in general has infinitely many solutions. In this case one may wish to optimize within the null-space to find solutions that minimize Rabi frequency or some metric for robustness. For long ion chains where N may be large, it is a more scalable solution to over-determine the problem, $n_{seg} < 2N$, which in general does not have an exact solution. However, good solutions can still be found through an optimization procedure. The obvious quantity to optimize is the fidelity given in Eq. 3.57, however this is nonlinear in $\mathbf{\Omega}$. In the limit of high-fidelity (small α), then the exponentials in the fidelity equation can be expanded in terms of α to

produces an approximate expression that is quadratic in $\mathbf{\Omega}$ (similar to Θ_{ij}) [65],

$$\begin{aligned} F &\approx 1 - \frac{4}{5}\sum_k (|\alpha_i^k|^2 + |\alpha_j^k|^2) \coth \frac{\hbar\omega_k}{2k_BT} \\ &= 1 - \frac{4}{5}\mathbf{\Omega}^T \left[\sum_k \left(\mathbf{A}_i^{k\dagger}\mathbf{A}_i^k + \mathbf{A}_j^{k\dagger}\mathbf{A}_j^k\right) \coth \frac{\hbar\omega_k}{2k_BT}\right] \mathbf{\Omega} \\ &= 1 - \frac{4}{5}\mathbf{\Omega}^T\mathbf{M}\mathbf{\Omega} \end{aligned} \tag{3.65}$$

The matrix $\mathbf{M}$ is Hermitian, which means that $\mathbf{\Omega}^T\mathbf{M}\mathbf{\Omega} = \mathbf{\Omega}^T\mathrm{Re}[\mathbf{M}]\mathbf{\Omega}$. With this definition, we want to minimize $\mathbf{\Omega}^T\mathbf{M}\mathbf{\Omega}$ subject to the constraint $\mathbf{\Omega}^T\boldsymbol{\gamma}\mathbf{\Omega} = \pm\pi/4$. The method of Lagrange multipliers can be used to minimize the scalar function $\mathcal{L}(\mathbf{\Omega}, \lambda) = \mathbf{\Omega}^T\mathbf{M}\mathbf{\Omega} - \lambda(\mathbf{\Omega}^T\boldsymbol{\gamma}\mathbf{\Omega} \mp \pi/4)$ as a function of $\mathbf{\Omega}$ and multiplier λ. This function is designed to have stationary points that satisfy $\frac{\partial\mathcal{L}}{\partial\mathbf{\Omega}} = 0$ and $\frac{\partial\mathcal{L}}{\partial\lambda} = 0$, leading to equations,

$$\frac{\partial\mathcal{L}}{\partial\mathbf{\Omega}} = 0 \Rightarrow \mathbf{M}\mathbf{\Omega} = \lambda\boldsymbol{\gamma}\mathbf{\Omega} \tag{3.66}$$

$$\frac{\partial\mathcal{L}}{\partial\lambda} = 0 \Rightarrow \mathbf{\Omega}^T\boldsymbol{\gamma}\mathbf{\Omega} = \pm\pi/4 \tag{3.67}$$

Equation 3.66 is just a generalized eigenvalue problem. After calculating the eigenvectors $\mathbf{\Omega}$, we scale each so that $\mathbf{\Omega}^T\boldsymbol{\gamma}\mathbf{\Omega} = \pm\pi/4$, and then calculate the exact fidelity (not approximate) as given in Eq. 3.57. The solution with the highest fidelity is the optimal solution.

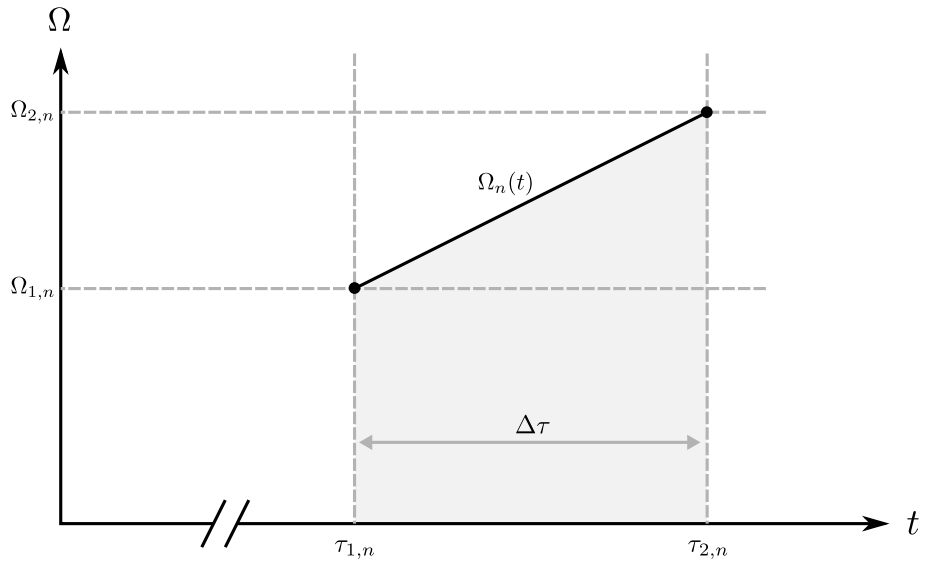

Figure 3.6: Linear parameterization of the n-th segment of an AM XX gate waveform. The are two free parameters corresponding the Rabi frequency at the start and end of the segment.

3.3.3: Linearly Interpolated Gate Solutions

We now expand upon the framework from the previous section to generate pulse solutions where the Rabi frequency of each segment varies in a linear fashion over the length of the segment and is continuous between segments. We now parameterize each segment n with two Rabi frequencies corresponding to the starting amplitude $\Omega_{1,n}$ at $\tau_{1,n} = (n-1)\tau_{gate}/n_{seg}$ and the end amplitude $\Omega_{2,n}$ at $\tau_{2,n} = n\tau_{gate}/n_{seg}$, with $\Delta\tau = \tau_{2,n} - \tau_{1,n}$ constant for all n. This parameterization is depicted in Fig. 3.6. The Rabi frequency at any point during segment n is then given by,

$$\begin{aligned}\Omega_n(t) &= \frac{\Omega_{2,n} - \Omega_{1,n}}{\Delta\tau}\, t + \frac{\tau_{2,n}\Omega_{1,n} - \tau_{1,n}\Omega_{1,n}}{\Delta\tau} \\ &= \frac{1}{\Delta\tau}\left[(\tau_{2,n} - t)\,\Omega_{1,n} + (t - \tau_{1,n})\,\Omega_{2,n}\right]\end{aligned} \tag{3.68}$$

There are now $2n_{seg}$ free parameters, with $\mathbf{\Omega} = \left(\Omega_{1,1}, \Omega_{2,1} \ldots \Omega_{1,n_{seg}}\Omega_{2,n_{seg}}\right)$.

When Ω was constant, we could pull it out of the spin-motion integral in Eq. 3.54. However now there is a linear factor of t, so we should expect to see terms like $\int t\sin(\mu t)e^{i\omega t}dt$. With this in mind, we can now re-define the spin-motion matrix given in Eq. 3.60 as follows, with $1 \leq n \leq n_{seg}$,

$$\alpha_j^k = \mathbf{A}_j^k \mathbf{\Omega} \tag{3.69}$$

$$\mathbf{A}_j^k(2n-1) = -\frac{i\eta_k b_j^k}{\Delta\tau}\left[\tau_{2,n}\int_{\tau_{1,n}}^{\tau_{2,n}} \sin(\mu t)e^{i\omega_k t}dt - \int_{\tau_{1,n}}^{\tau_{2,n}} t\sin(\mu t)e^{i\omega_k t}dt\right] \tag{3.70}$$

$$\mathbf{A}_j^k(2n) = -\frac{i\eta_k b_j^k}{\Delta\tau}\left[\int_{\tau_{1,n}}^{\tau_{2,n}} t\sin(\mu t)e^{i\omega_k t}dt - \tau_{1,n}\int_{\tau_{1,n}}^{\tau_{2,n}} \sin(\mu t)e^{i\omega_k t}dt\right] \tag{3.71}$$

The spin-spin matrix is more complicated. Once it is multiplied out, it includes terms that look like $\Omega_p\Omega_q$ for different segments p and q, which start to involve many cross terms once $\Omega(t)$ is allowed to vary linearly. Now using $t_{p/q}$ to denote the independent time parameter of each segment, we can represent all the cross-terms with a 2×2 matrix called $\mathbf{B}_{pq}$,

$$\begin{aligned}\Omega_p(t_p)\Omega_q(t_q) &= \frac{1}{\Delta\tau^2}\left[(\tau_{2,p}-t_p)\,\Omega_{1,p} + (t_p-\tau_{1,p})\,\Omega_{2,p}\right]\left[(\tau_{2,q}-t_q)\,\Omega_{1,q} + (t_q-\tau_{1,q})\,\Omega_{2,q}\right] \\ &= \begin{pmatrix}\Omega_{1,q} & \Omega_{2,q}\end{pmatrix}\mathbf{B}_{pq}\begin{pmatrix}\Omega_{1,p}\\ \Omega_{2,p}\end{pmatrix}\end{aligned} \tag{3.72}$$

with,

$$\mathbf{B}_{pq} = \frac{1}{\Delta\tau^2}\begin{pmatrix}\tau_{2,p}\tau_{2,q} - \tau_{2,q}t_p - \tau_{2,p}t_q + t_pt_q & -\tau_{1,p}\tau_{2,q} + \tau_{2,q}t_p + \tau_{1,p}t_q - t_pt_q \\ -\tau_{2,p}\tau_{1,q} + \tau_{1,q}t_p + \tau_{2,p}t_q - t_pt_q & \tau_{1,p}\tau_{1,q} - \tau_{1,q}t_p - \tau_{1,p}t_q + t_pt_q\end{pmatrix} \tag{3.73}$$

The analogous matrix for the diagonal case ($\Omega_p\Omega_p$), is given by $\mathbf{D}_{pp}$,

$$\Omega_p(t_1)\Omega_p(t_2) = \begin{pmatrix}\Omega_1 & \Omega_2\end{pmatrix}\mathbf{D}_{pp}\begin{pmatrix}\Omega_1\\ \Omega_2\end{pmatrix} \tag{3.74}$$

$$\mathbf{D}_{pp} = \frac{1}{\Delta\tau^2}\begin{pmatrix}\tau_{2,p}^2 - \tau_{2,p}t_1 - \tau_{2,p}t_2 + t_1t_2 & -\tau_{1,p}\tau_{2,p} + \tau_{1,p}t_2 + \tau_{2,p}t_1 - t_1t_2\\ -\tau_{1,p}\tau_{2,p} + \tau_{1,p}t_1 + \tau_{2,p}t_2 - t_1t_2 & \tau_{2,p}^2 - \tau_{1,p}t_1 - \tau_{1,p}t_2 + t_1t_2\end{pmatrix} \tag{3.75}$$

Again, when $\Omega_p\Omega_q$ was constant, it could be pulled out of the spin-spin double integral in Eq. 3.56. However now there is both a linear factor of $t_{p/q}$ but also a quadratic factor t_pt_q, so we should expect to see these terms inside the integrals. Define,

$$\Lambda_0 = \frac{1}{\Delta\tau^2}\sum_k \eta_k^2 b_i^k b_j^k \int_{\tau_{1,p}}^{\tau_{2,p}} dt_p \int_{\tau_{1,q}}^{\tau_{2,q}} \sin(\mu t_p)\sin(\mu t_q)\sin[\omega_k(t_p - t_q)]dt_q \tag{3.76}$$

$$\Lambda_p = \frac{1}{\Delta\tau^2}\sum_k \eta_k^2 b_i^k b_j^k \int_{\tau_{1,p}}^{\tau_{2,p}} dt_p \int_{\tau_{1,q}}^{\tau_{2,q}} t_p\sin(\mu t_p)\sin(\mu t_q)\sin[\omega_k(t_p - t_q)]dt_q \tag{3.77}$$

$$\Lambda_q = \frac{1}{\Delta\tau^2}\sum_k \eta_k^2 b_i^k b_j^k \int_{\tau_{1,p}}^{\tau_{2,p}} dt_p \int_{\tau_{1,q}}^{\tau_{2,q}} t_q\sin(\mu t_p)\sin(\mu t_q)\sin[\omega_k(t_p - t_q)]dt_q \tag{3.78}$$

$$\Lambda_{pq} = \frac{1}{\Delta\tau^2}\sum_k \eta_k^2 b_i^k b_j^k \int_{\tau_{1,p}}^{\tau_{2,p}} dt_p \int_{\tau_{1,q}}^{\tau_{2,q}} t_pt_q\sin(\mu t_p)\sin(\mu t_q)\sin[\omega_k(t_p - t_q)]dt_q \tag{3.79}$$

Since the matrix $\mathbf{B}_{pq}$ contains all the information about the cross terms that will show up in the spin-spin matrix (Eq. 3.60), we substitute $t_{p/q} \to \Lambda_{p/q}$ in $\mathbf{B}_{pq}$ to arrive at $\widetilde{\mathbf{B}}_{pq}$, which contains all the spin-spin information between segment p and

q,

$$\widetilde{\mathbf{B}}_{pq} = \begin{pmatrix} \tau_{2,p}\tau_{2,q}\Lambda_0 - \tau_{2,q}\Lambda_p - \tau_{2,p}\Lambda_q + \Lambda_{pq} & -\tau_{1,p}\tau_{2,q}\Lambda_0 + \tau_{2,q}\Lambda_p + \tau_{1,p}\Lambda_q - \Lambda_{pq} \\ -\tau_{2,p}\tau_{1,q}\Lambda_0 + \tau_{1,q}\Lambda_p + \tau_{2,p}\Lambda_q - \Lambda_{pq} & \tau_{1,p}\tau_{1,q}\Lambda_0 - \tau_{1,q}\Lambda_p - \tau_{1,p}\Lambda_q + \Lambda_{pq} \end{pmatrix} \tag{3.80}$$

Similar integrals can be defined in the diagonal case and substituted into $\mathbf{D}$ to form $\widetilde{\mathbf{D}}$,

$$\Xi_0 = \frac{1}{\Delta\tau^2}\sum_k \eta_k^2 b_i^k b_j^k \int_{\tau_{1,p}}^{\tau_{2,p}} dt_1 \int_{\tau_{1,p}}^{t_1} \sin(\mu t_1)\sin(\mu t_2)\sin[\omega_k(t_1-t_2)]dt_2 \tag{3.81}$$

$$\Xi_1 = \frac{1}{\Delta\tau^2}\sum_k \eta_k^2 b_i^k b_j^k \int_{\tau_{1,p}}^{\tau_{2,p}} dt_1 \int_{\tau_{1,p}}^{t_1} t_1\sin(\mu t_1)\sin(\mu t_2)\sin[\omega_k(t_1-t_2)]dt_2 \tag{3.82}$$

$$\Xi_2 = \frac{1}{\Delta\tau^2}\sum_k \eta_k^2 b_i^k b_j^k \int_{\tau_{1,p}}^{\tau_{2,p}} dt_1 \int_{\tau_{1,p}}^{t_1} t_2\sin(\mu t_1)\sin(\mu t_2)\sin[\omega_k(t_1-t_2)]dt_2 \tag{3.83}$$

$$\Xi_{12} = \frac{1}{\Delta\tau^2}\sum_k \eta_k^2 b_i^k b_j^k \int_{\tau_{1,p}}^{\tau_{2,p}} dt_1 \int_{\tau_{1,p}}^{t_1} t_1 t_2\sin(\mu t_1)\sin(\mu t_2)\sin[\omega_k(t_1-t_2)]dt_2 \tag{3.84}$$

$$\widetilde{\mathbf{D}}_{pp} = \begin{pmatrix} \tau_{2,p}^2\Xi_0 - \tau_{2,p}\Xi_1 - \tau_{2,p}\Xi_2 + \Xi_{12} & -\tau_{1,p}\tau_{2,p}\Xi_0 + \tau_{1,p}\Xi_2 + \tau_{2,p}\Xi_1 - \Xi_{12} \\ -\tau_{1,p}\tau_{2,p}\Xi_0 + \tau_{1,p}\Xi_1 + \tau_{2,p}\Xi_2 - \Xi_{12} & \tau_{2,p}^2\Xi_0 - \tau_{1,p}\Xi_1 - \tau_{1,p}\Xi_2 + \Xi_{12} \end{pmatrix} \tag{3.85}$$

Now the spin-spin matrix has the condensed form for $1 \le p, q \le n_{seg}$,

$$\Theta_{ij} = \mathbf{\Omega}^T\boldsymbol{\gamma}\mathbf{\Omega} \tag{3.86}$$

$$\boldsymbol{\gamma}(p,q) = \begin{pmatrix} \gamma_{2p-1,2q-1} & \gamma_{2p-1,2q} \\ \gamma_{2p,2q-1} & \gamma_{2p,2q} \end{pmatrix} = \begin{cases} \widetilde{\mathbf{B}}_{pq} & (p > q) \\ 2\widetilde{\mathbf{D}}_{pp} & (p = q) \\ \widetilde{\mathbf{B}}_{qp} & (p < q) \end{cases} \tag{3.87}$$

We have now reformulated the problem to have the same form as in the previous section - minimize $\mathbf{\Omega}^T\mathbf{M}\mathbf{\Omega}$ (with M constructed from $\mathbf{A}_j^k$ given in Eq. 3.70 and 3.71) subject to the constraint $\mathbf{\Omega}^T\boldsymbol{\gamma}\mathbf{\Omega} = \pm\pi/4$ (with $\boldsymbol{\gamma}$ given in Eq. 3.87). However, the motivation for linearly-interpolated gates was to smooth out impulse jumps in the laser power that could potentially excite axial motion through prompt charge response in the chip trap. So far, we have not accomplished this because at each segment boundary, there are now two free-parameters and there is no guarantee that they are equal. Thus, we wish to enforce that $\Omega_{2,n} = \Omega_{1,n+1}$ for all $1 \leq n < n_{seg}$. Additionally, we wish the waveform to smoothly turn on/off at the start/end of the waveform, $\Omega_{1,1} = \Omega_{2,n_{seg}} = 0$. We can succinctly represent all these constraints as,

$$
\begin{aligned}
0 &= (\Omega_{1,1} - 0)^2 + (\Omega_{2,1} - \Omega_{1,2})^2 + \cdots + (\Omega_{2,n_{seg}-1} - \Omega_{1,n_{seg}})^2 + (\Omega_{2,n_{seg}} - 0)^2 \\
&= \Omega_{1,1}^2 + \Omega_{2,1}^2 + \Omega_{1,2}^2 - 2\Omega_{2,1}\Omega_{1,2} + \cdots + \Omega_{2,n_{seg}-1}^2 + \Omega_{1,n_{seg}}^2 - 2\Omega_{2,n_{seg}-1}\Omega_{1,n_{seg}} + \Omega_{2,n_{seg}}^2 \\
&= \mathbf{\Omega}^T \begin{pmatrix}
1 & 0 & 0 & 0 & & & & \\
0 & 1 & -1 & 0 & \cdots & \cdots & \cdots & \cdots \\
0 & -1 & 1 & 0 & & & & \\
0 & 0 & 0 & \ddots & & & & \\
\vdots & & & & \ddots & 0 & 0 & 0 \\
\vdots & & & & 0 & 1 & -1 & 0 \\
\vdots & & & & 0 & -1 & 1 & 0 \\
\vdots & & & & 0 & 0 & 0 & 1
\end{pmatrix} \mathbf{\Omega} \\
&= \mathbf{\Omega}^T\mathbf{C}\mathbf{\Omega}
\end{aligned}
\tag{3.88}
$$

Where C is called the "continuity matrix" defined below. This matrix has dimension

$2n_{seg} \times 2n_{seg}$ and rank $n_{seg} + 1$. Therefore it has a null space of size $n_{seg} - 1$.

$$\mathbf{C}(i,j) = \begin{cases} 1, & i = j \\ -1, & (i = 2n \;\&\; j = 2n+1) \parallel (i = 2n+1 \;\&\; j = 2n) \text{ for any } 1 \leq n < n_{seg} \\ 0, & \text{else} \end{cases}$$

We wish to operate the Lagrange multiplier optimization in a regime where the continuity condition is guaranteed. To accomplish this, we find the vector null space of $\mathbf{C}$ and then project the fidelity matrix $\mathbf{M}$ and spin-spin matrix $\boldsymbol{\gamma}$ into this constrained space where continuity is guaranteed. The vector null space of $\mathbf{C}$ can be found using SVD, or equivalently because $\mathbf{C}$ is positive semi-definite, finding the eigenvalues of $\mathbf{C}$ and picking out those eigenvectors with eigenvalue $\lambda = 0$. Defining the column null space of $\mathbf{C}$ as $\mathbf{V}_C$ with dimension $2n_{seg} \times n_{seg} - 1$, then the constrained fidelity and spin-spin matrices are given by,

$$\mathbf{M}_C = \mathbf{V}_C^T \mathbf{M} \mathbf{V}_C \tag{3.89}$$

$$\boldsymbol{\gamma}_C = \mathbf{V}_C^T \boldsymbol{\gamma} \mathbf{V}_C \tag{3.90}$$

Now we can proceed to solving the Lagrange multiplier optimization. Note that the null space of $\mathbf{C}$ is size $n_{seg} - 1$, which is equivalent to the number degrees of freedom in the system. This means that the condition for an over-determined system is now $n_{seg} < 2N + 1$. As before, we wish to minimize the scalar function $\mathcal{L}_C(\boldsymbol{\Omega}_C, \lambda) = \boldsymbol{\Omega}_C^T \mathbf{M}_C \boldsymbol{\Omega}_C - \lambda(\boldsymbol{\Omega}_C^T \boldsymbol{\gamma}_C \boldsymbol{\Omega}_C \mp \pi/4)$ with the corresponding stationary

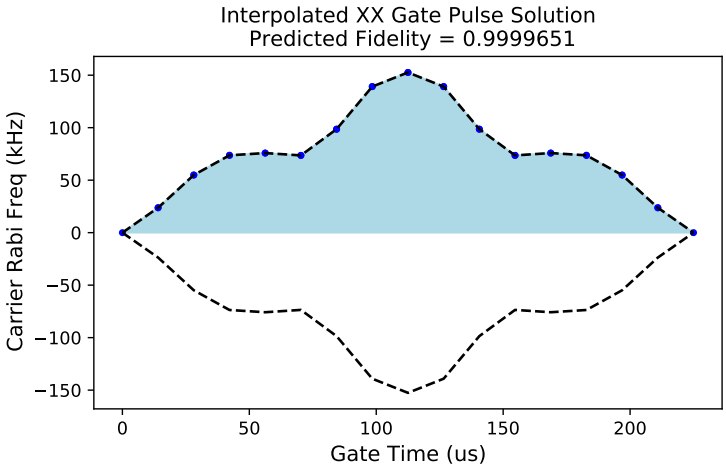

Figure 3.7: Example gate solution on ions $(-6, 5)$ in a 15-ion chain. The y-axis corresponds to the Rabi frequency of *each* red/blue sideband when brought into resonance with the carrier. Ω is allowed to go negative, although in this case it does not as evidenced by the filled blue volume about the x-axis.

points,

$$\mathbf{M}_C\mathbf{\Omega}_C = \lambda\boldsymbol{\gamma}_C\mathbf{\Omega}_C \tag{3.91}$$

$$\mathbf{\Omega}_C^T\boldsymbol{\gamma}_C\mathbf{\Omega}_C = \pm\pi/4 \tag{3.92}$$

After solving the generalized eigenvalue problem, the full $\mathbf{\Omega}$ vector can be recovered by $\mathbf{\Omega} = \mathbf{V}_C\mathbf{\Omega}_C$. Similar to before, we scale each so that $\mathbf{\Omega}^T\boldsymbol{\gamma}\mathbf{\Omega} = \pm\pi/4$, and then pick the solution with the highest exact fidelity as the optimal solution.

The result of this optimization procedure is shown in Figure 3.7, with the gate solution on ions $(-6, -5)$ relative to the center ion in a chain of 15 ions. The mode frequencies and the mode participation matrix used for this solution are the same as show in Fig. 2.4 and 2.3. For this maximally entangling gate, $n_{seg} = 16$ and $\tau_{gate} = 225$ μs. The predicted fidelity for this gate, in the presence of no other noise

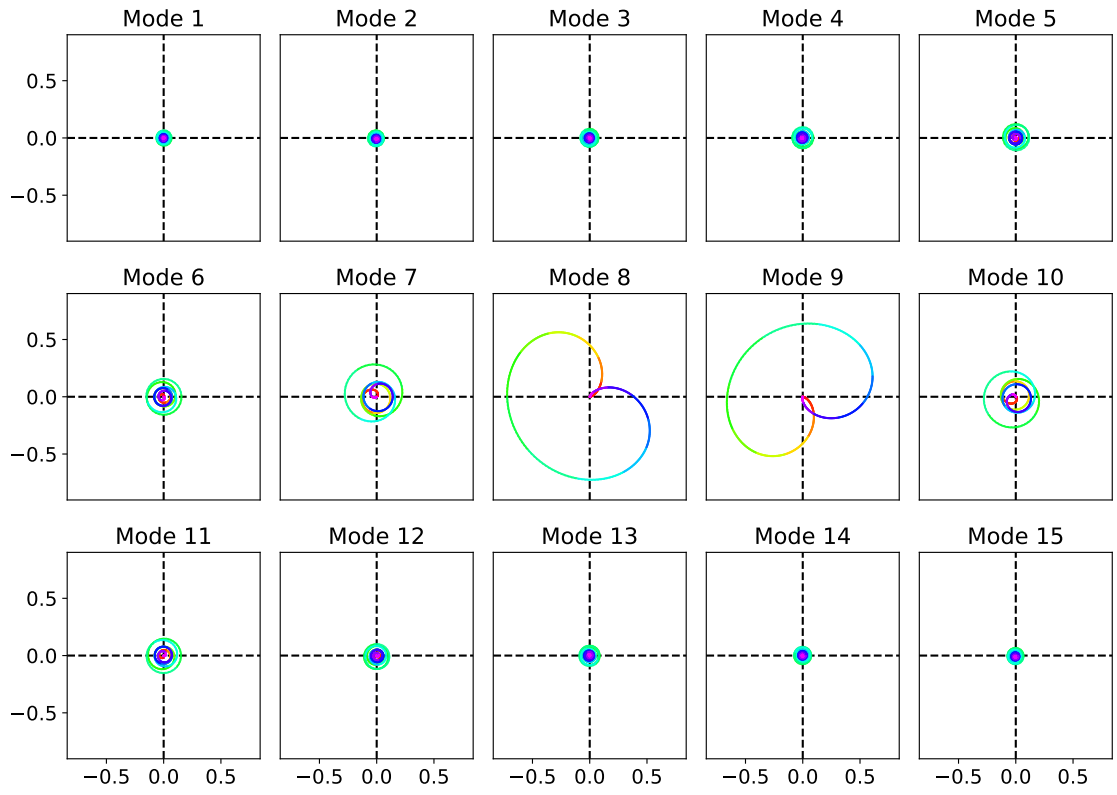

Figure 3.8: Phase space trajectories for the gate solution shown in Fig. 3.7. Mode 1 corresponds to the lowest spatial frequency mode. The trajectories shown are ion independent in the sense that they are not yet scaled by the mode participation factors. Red corresponds to $t = 0$.

sources is $> 99.99\%$.

The phase space trajectories for each mode under the influence of this gate solution is shown in Fig. 3.8. The curves correspond exactly to Eq. 3.54, where the factor of b_j^k has been omitted to make the trajectories ion independent. Mode 8 and 9 are most strongly excited because the gate frequency sits in between them. Thus, failure to close these two modes will result in the largest error. As the modes become more detuned from the drive frequency, their phase space performs small oscillations about the origin, approaching the weak-field limit described in Sec. 3.3.1. In practice, the far away modes can sometimes be completely ignored [75].

3.3.4: Optimizing Gates for Robustness

So far we have assumed that the gate detuning relative to the carrier μ, has been fixed at some arbitrary value. While it is possible to find a relatively high fidelity solution for all μ, some choices of μ require prohibitively high laser power. This depends on mode-participation of the ions on modes that are strongly driven and enclose the most phase space. For example, an XX gate that includes ion 0 (the center ion in an odd-number chain) that is detuned close to an odd-spatial mode will have zero participation in that mode $b_0^{2n} = 0$ (where in-phase index is $n = 1$). Thus on average, it will be detuned further from the active modes and require more power to accumulate the phase required. Therefore, the choice of μ should take into consideration the joint mode participation vectors of each ion in the gate.

Additionally, there are some choices of μ that are more robust to gate detuning errors. While the RF tone is generally very stable in frequency, the modes may shift relative to this frequency due to drifts in the RF power or DC quadrupoles, which will look identical to a mode detuning error. These shifts have been characterized in this system to be < 500 Hz over several hours (Sec. 5.2.2). The approach to find robust solutions is a numerical one - scan over the entire mode spectrum, at each point find the optimal solution, and then evaluate the average fidelity subject to a ± 1 kHz gate detuning error. The gate with the best average fidelity is the optimally robust solution. The result of this this procedure is shown in Figure 3.9, where the black dashed line indicates the choice of μ that results in the optimally robust solution (corresponding to the minimum in the middle plot). This gate solution is

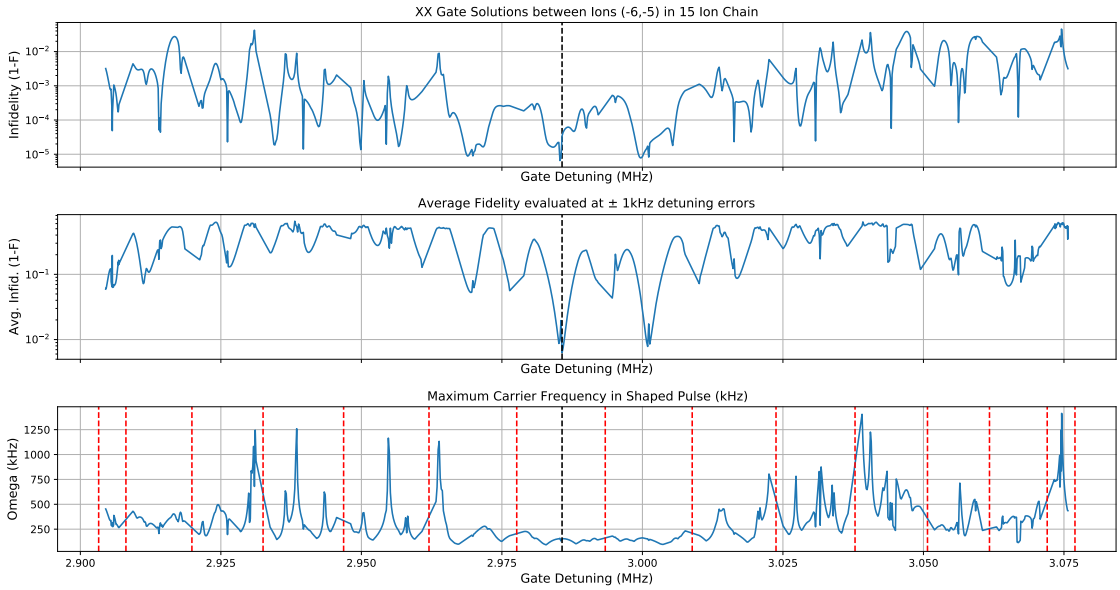

Figure 3.9: Example gate detuning scan for an XX gate on ions $(-6, -5)$ in a 15-ion chain. The optimally robust detuning corresponding to the solution shown in Fig. 3.7 is indicated by the black dashed line in all the plots. (Top) Minimum predicted infidelity of the optimal solution computed at different gate detunings across the mode spectrum. (Middle) Average infidelity of the optimal solution in the plot above, subjected to a ± 1 kHz gate error. (Bottom) Maximum carrier Rabi frequency of the optimal solution for each gate detuning. Locations of the motional modes are indicated by red dashed lines.

the one depicted in Fig. 3.7. This procedure is repeated across all $\binom{N}{2}$ possible gate pairs to generate the full set. Due to reflection symmetry of the chain, gate (i, j) has the same solution as $(-i, -j)$, for an ion indexing scheme where the center ion is 0 in an odd-length chain. So in practice only $\binom{N}{2}/2$ solutions need to be computed.

Empirically we observe that the optimally robust solutions tend to fall nearly in the middle of two modes. Intuitively, for a mode spectrum where all the modes are equally spaced by δ, a gate centered between two of these modes will be detuned from the mode spectrum by $\Delta_i \in \{\ldots, 3\delta/2, \delta/2, -\delta/2, -3\delta/2 \ldots\}$. A square bichromatic pulse will drive phase space loops with period $\tau_i = 1/\Delta_i$. Thus all the modes will

close at $\tau = 2N/\delta$ with integer N. In reality the modes are not perfectly equispaced and we use amplitude as a control knob and not τ, but the underlying symmetry hints at why the chosen detunings may be particularly robust.

Additionally, one may also set a threshold for laser power and discard any solutions that exceed this to guarantee practical operations. Further considerations for the choice of μ are to avoid the lowest frequency spatial mode, because it will be the mode predominantly heated by spatially uniform anomalous electric field noise. Additionally, we observe through numerical analysis that location of the highest spatial frequency mode is highly sensitive to voltage miscalibrations and should be avoided. Since the middle 5 modes in a 15-ion chain are stable and do not heat, they are used in this system as the primary workhorse modes for two qubit gates.

To verify that the example gate solution is robust to a wide-range of gate detuning errors, we perform a more detailed analysis on the optimally robust gate solution. Define the gate detuning error by ϵ_μ and the fidelity of the gate subject to the error given by Eq. 3.57 with the substitution $\mu \to \mu + 2\pi\epsilon_\mu$ in Eq. 3.55. The result of scanning ϵ_μ over the range $[-2.5, 2.5]$ kHz is shown in Figure 3.10. This particular solution achieves 10^{-4} error over 300 Hz of bandwidth, 10^{-3} error over 1.025 kHz of bandwidth, 10^{-2} error over 2.300 kHz of bandwidth, and 10^{-1} error over 4.475 kHz bandwidth. This level of robustness is sufficient in our system due to the relative stability of the modes (Sec. 5.2.2) and the fact that other noise sources (other than motional closure) seemingly dominate the gate fidelity (Sec. 5.3.4).

In the future, there are straight forward ways to enforce robustness in the computation of $\mathbf{\Omega}$ rather than to rely on a brute force numerical search. For one

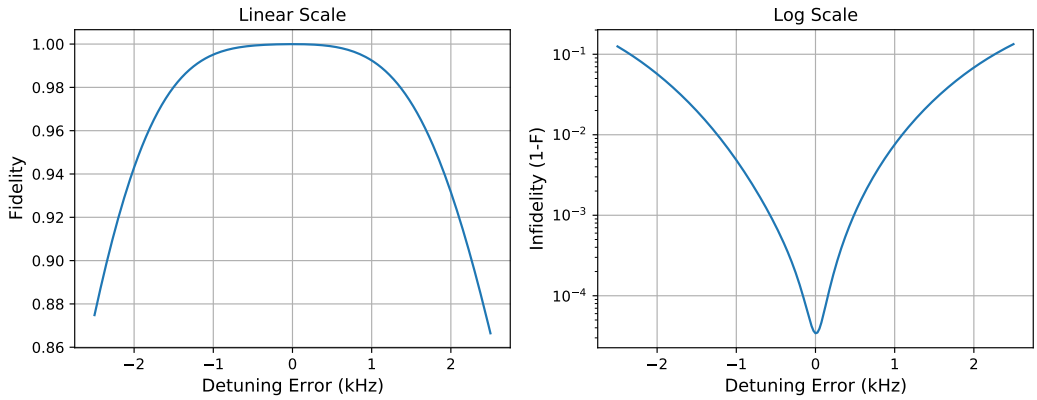

Figure 3.10: Predicted gate detuning error vs. gate fidelity/error for the gate shown in Fig. 3.7 on a 15-ion chain. (Left) Linear scale gate fidelity. (Right) Log scale gate error

approach, in addition to enforcing $\alpha_j^k = 0$, one can also enforce the higher order moments to also be zero, $\frac{\partial \alpha}{\partial \mu} = 0$, $\frac{\partial^2 \alpha}{\partial \mu^2} = 0$, etc, in the Lagrange optimization formulation. Including only up to the 2nd moment has been shown to double the bandwidth of 10^{-3} error relative to what is shown here [70].

Chapter 4: Integrated Ion Trap System

In this chapter we will describe the hardware components of the experimental system. A systems engineering approach was used to design this experiment, starting from requirements dictated by the physics of the ions. A trapped ion quantum computer consists of several complex subsystems, working in harmony to achieve high fidelity operations. To the extent that is is possible, it is useful to conceptually divide the system into modular subsystems that inherit requirements from the ions. To scale a system successfully, there are also a few key design principles that must be considered. The primary challenge when scaling up a system size is to maintain the fidelity demonstrated on smaller systems. As the system grows in size, the complexity of the control grows as well. The failure or under-performance of a single component in the system can turn the quantum computer into a quantum random number generator. Thus, robustness and reliability at the component level are two key design principles required for scaling.

As the number of qubits in the system grows, so do the required calibrations required to run it at high fidelity. Passive stability of the system, whether it be the optical addressing system or the background electric field seen by the ion, is critical to reducing the calibration overhead. Passive stability is often achieved

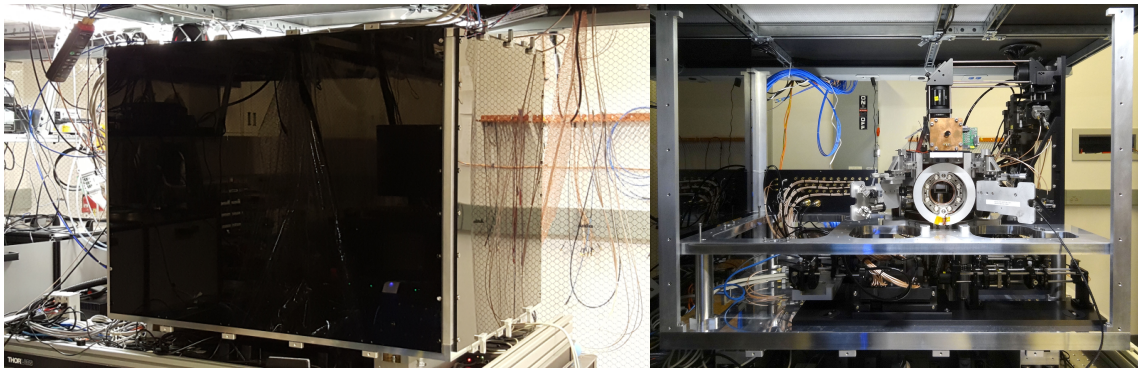

Figure 4.1: A picture of the integrated ion trap quantum computer. (Left) The sealed box provides environmental isolation and stability of the enclosed optical systems. (Right) Inside the box, the vacuum chamber sits on a raised plate. The Raman and imaging optics primarily reside underneath, on the bottom plate.

by miniaturizing key subsystems, bolting down unused degrees of freedom in the system, and maintaining stable environmental parameters such as temperature and humidity. Even with a passively stable design, calibration values are still expected to drift and the number of calibration routines quickly exceeds the attention span of a single human. Thus automation is another key factor that allows a large system to run at high fidelity for extended period of time. This includes autoloading ions after a loss event, tracking beam alignment, or periodically calibrating two-qubit gate amplitudes. To achieve automation, the relevant degrees of freedom need to actuated (by motor, piezo, voltage, etc.) and commanded by a centralized control system that maintains awareness about the state of the machine.

In Figure 4.1, we show the integrated ion trap quantum computer used in this thesis. The vacuum chamber, imaging system, and Raman system all reside in a sealed black box, which exists to shield the sensitive components from environmental fluctuations and provide passive stability. All of the CW lasers reside on an adjacent

optical table and are routed into the box via optical fibers. Auxiliary electronics, such as the control FPGA, power supplies, and frequency sources, are mounted in a nearby server rack or sit on a hanging rack above the black box.

4.1: Vacuum Chamber Subsystem

In general, qubits need to be well isolated from the environment, otherwise they will exchange thermal energy with the environmental bath and quickly decohere. In trapped ion systems collisions with background gas particles can eject ions from the trap or scramble the output of a quantum computation. Thus, the chip trap and oven sources must sit in a UHV environment with low residual background pressure. Internal chamber parts must be meticulously cleaned with solvents, assembled, sealed in the chamber and baked under vacuum in order to achieve the lowest possible pressures. In this system we achieve a pressure of 7.11×10^{-11} Torr, measured by directly observing hopping rates of a single ion in a double well potential.

The internal stack up of the vacuum system is shown in Figure 4.2. The baseplate shown at the bottom of the picture is mounted internally to the top 6" Conflat feed-through flange of the chamber (the chip trap sits upside down in our system). The feed-through flange includes a 100-pin micro D-sub for the DC electrodes, and 8 electrical feed-throughs for oven, RF, and ground connections. A small hole exists in the center of the flange with a tapered adaptor that ends in a 2.75" Conflat viewport to provide optical access through the slot. The HOA chip is

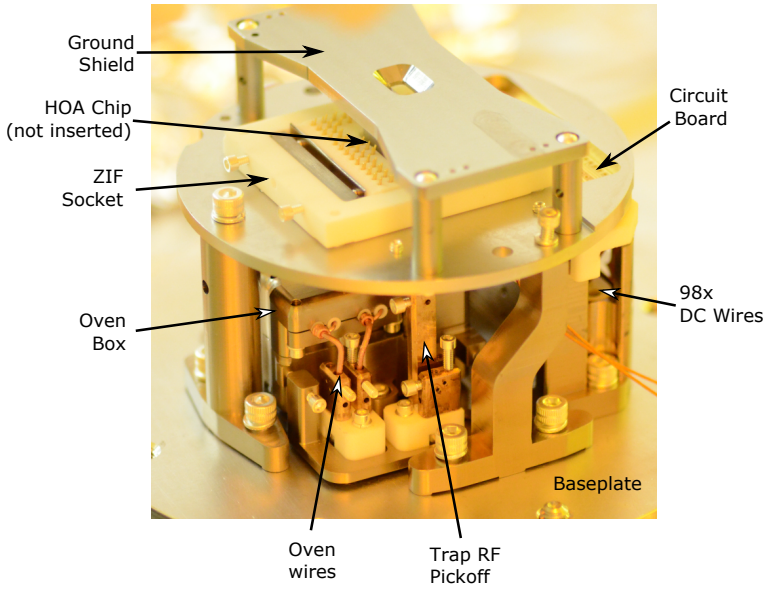

Figure 4.2: Picture of internal vacuum stack-up. Copper rods that provide trap RF connections are in the rear of the picture and are not visible.

inserted in a ZIF-socket that sits on stand-offs from the baseplate. A circuit board underneath the ZIF-socket routes electrical signals from the chip trap to the proper wires from the feed-through flange. A metal ground shield is installed above the chip to shield stray electric fields from the nearby viewport. Underneath the ZIF-socket is the oven box, which contains the thermal ovens aimed at various parts of the trap. This system has a total of 5 different ovens. Two isotopically enriched Yb thermal ovens are aimed at either the load slot or the quantum region of the trap. A thermal Barium oven is aimed at a separate load slot. Finally two ablation targets, one Yb and one Ba are aimed at the remaining two load slots.

The external structure of the vacuum chamber system is shown in Figure 4.3. The entire chamber is raised to a secondary level to allow for Raman and imaging optics on the bottom level. At the top and bottom of the chamber are 6" Conflat openings (along $\hat{z}$). The top opening is sealed by the feed-through flange and the

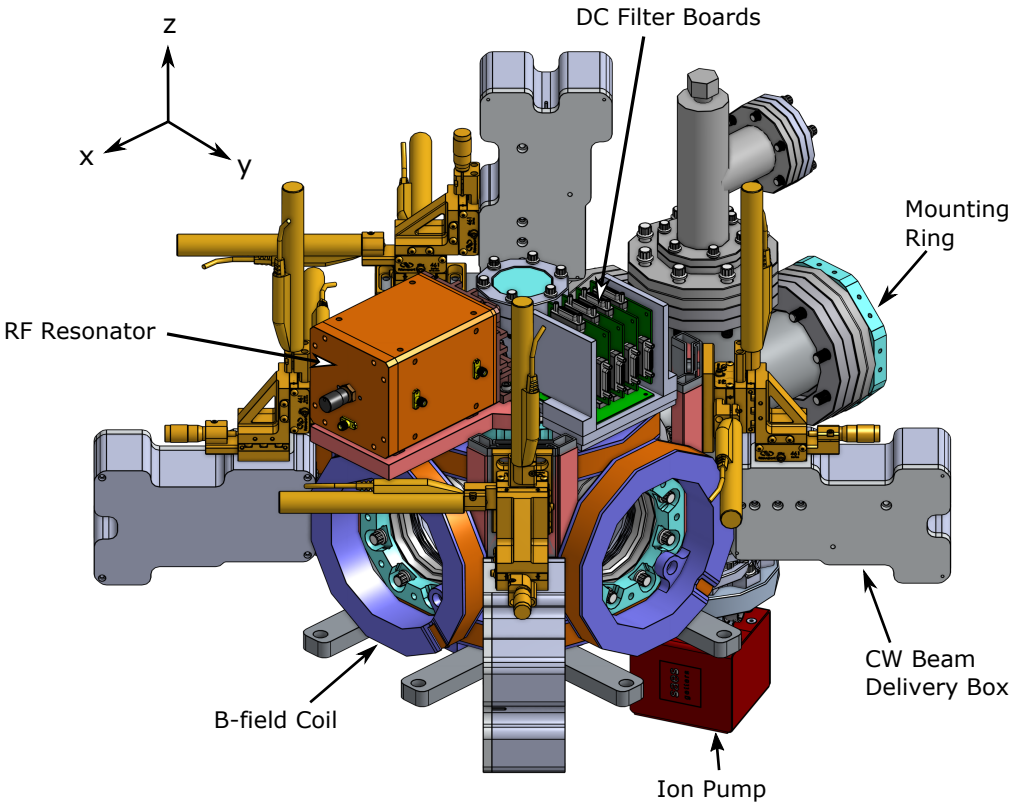

Figure 4.3: Model of external vacuum chamber system. The entire chamber sits on a raised platform, allowing optical access from underneath.

bottom opening is sealed by a re-entrant viewport that provides a short working distance to the ions for the imaging system and individual Raman beams. Along the main $\hat{x}, \hat{y}$ axes in the horizontal plane are four 4.5" Conflat openings. Three of these are sealed with viewports, and the fourth leads to a cross that supports the ion pump/NEG, vacuum gauge, and another viewport that provides clear optical access. Oriented at 45° relative to the main horizontal axes are four 1.33" Conflat openings that are sealed with viewports for CW beam delivery.

Each of the 4.5" and 6" flanges are sealed along with a captured steel mounting ring that provides tapped mounting points. The RF resonator and DC filter boards are mounted directly on the chamber, via the mounting ring, near their respective

feed-through connections to ensure solid ground connection and shielding of the output signals. The output of each is well shielded to minimize the risk of external noise entering the system. Also attached to these mounting rings are Helmholtz coils along the three orthogonal directions that allow us to specify an arbitrary magnetic field vector. In typical operation of this system, a magnetic field of ≈ 5 G is applied along the $\hat{z}$ axis. Finally associated with each 1.33" viewport opening is a CW beam delivery box. Each box includes a fiber collimator, polarization optics, and a focusing lens. Additionally a pick-off lens and small photodiode mounted in the box provide power monitoring and tweaking up of fiber-coupling. The output beam on the box can be aligned onto the ions via two motorized linear actuators (CONEX-TRB12CC) transverse to the direction of beam propagation. By mounting the CW beam delivery optics directly to the vacuum chamber, which is in turn referenced to the ion position via the internal stack-up, misalignment drifts are heavily mitigated.

4.2: Trap Subsystem

The trap subsystem includes the ion trap itself, but also RF resonator used to generate the trapping voltage and the DC voltage source to provide static confining potentials.

4.2.1: Sandia HOA 2.1.1

The micro-fabricated chip trap in our system is an HOA-2.1.1 chip trap manufactured by Sandia National Laboratories [34]. The trap geometry consists of a

$\approx$ 3 mm linear slot is the middle of the trap, which we call the "quantum region" and is where we build up long chains and perform quantum circuits. As shown in the top of Figure 4.4, the surface of the trap is raised 1.81 mm above the surface of the chip carrier to allow optical access for focused beams. Similarly, the trap is narrowed in the quantum region with a signature bow-tie shape to allow for a maximum NA of 0.11 for beams parallel to the chip surface. For beams perpendicular to the chip, passing through the 60 μm wide slot in the quantum region, the available NA is 0.25.

On each end of the quantum region, there is a Y-junction that the splits into two arms each. Each of the four arms has a slot milled through the substrate called a "load slot". There are 94 independent DC control electrodes that allow for shuttling and transport of ions, as well as fine control of the axial potential of the chain. Each DC electrode has an integrated capacitor of 1.05 nF, that shorts to ground at RF frequency to avoid parasitic coupling with the RF. As shown in the bottom of Figure 4.4, in order to build up a long chain, a single ion is loaded in the load slot, and then shuttled and merged into the existing chain in the quantum region. When the desired number of ions is reached, this process halts. While there is a thermal oven aimed at the quantum region, it is not used to avoid contamination of the surface near the ion chain and also because the loading process is less controlled, making it more difficult to load deterministic chain lengths.

We apply 220 V to the RF electrodes at a frequency of $\Omega_{RF} = 2\pi \times 36.06$ MHz, resulting in a radial secular frequency of $\omega_{sec} = 2\pi \times 3.155$ MHz. The RF null is located at a height of $z = 68$ μm above the surface of the nearest metal electrode

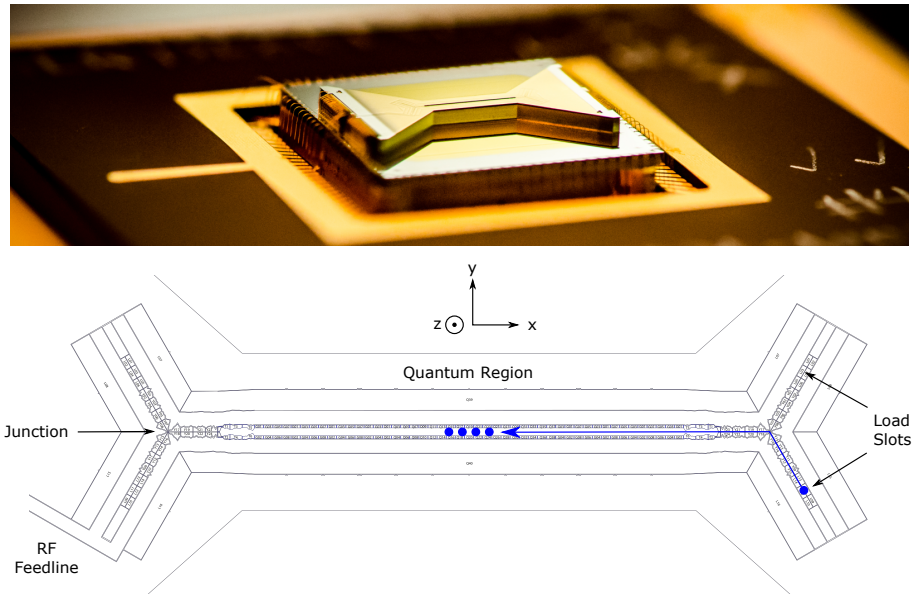

Figure 4.4: (Top) A picture of the the HOA-2.1.1 chip trap. The central quantum region is used to perform circuits on long chains, and is raised and tapered in order to allow optical beam access. (Bottom) Each ion is individually loaded in a load slot, and then shuttled and merged with an existing chain in the quantum region to deterministically load a given chain length.

with a characteristic trap distance of $R = 140$ μm. The potential that we use to model the potential of the trap is given in unitless spatial coordinate defined by d_0 (Eq. 2.14) by,

$$V(x,y,z) = \frac{X_1}{1!}(x) + \frac{X_2}{2!}(x^2 - y^2/2 - z^2/2) + \frac{X_3}{3!}(x^3) + \frac{X_4}{4!}(x^4 - 3x^2y^2 - 3x^2z^2) + \frac{X_{bend}}{3!}(x^2 z) - Q_{yz}(yz) - Q_{xz}(xz) + f_{sec}^2(y^2/2 + z^2/2)(1 - 3z/h) \tag{4.1}$$

Where X_i are unitless constants defined in Equation, 2.17. X_{bend} is a term that is used to describe a bending of the chain in the vertical dimension. This decompensation is a result of the lack of segmented outer electrodes in the design and empirically scales with X_4 as $X_{bend} \sim 13.5X_4$. Q_{ij} are quadrupoles where $Q_{yz} = 0.5$ is used

to tilt the principle axes of the trap at 45°. Finally, h is the location of the RF null and the scaling term is used to describe the asymmetry of radial confinement in the vertical direction. In each case, the unitless constants scale a particularly defined voltage primitive that is applied to the DC trap electrodes designed to apply the desired potential term without decompensating the ion from the RF null. In typical operation of a 15-ion chain, the in-phase mode along each trap axis is $(\nu_x, \nu_{y+z}, \nu_{y-z}) = (0.193, 3.077, 3.234)$ MHz.

4.2.2: RF Resonator

Creating 220 V to supply to the RF line in the trap is a non-trivial task. Also, the secular motion of the ion, f_{sec}, is directly proportional to this voltage, which means that the radial motional mode locations and therefore the two-qubit gate fidelity also depends on the stability of the RF voltage applied to the trap. Active stabilization of the voltage over short time scales is required to stabilize the modes over the time-scale of a gate and decrease motional decoherence due to fast phase noise on the motional modes. The pick-off for fast stabilization is before the vacuum chamber, so thermal drifts in the vacuum chamber and trap itself can still cause drifts. Thus, passive stability of this voltage over long time scales is required to combat motional mode calibration overhead.

The trap voltage is created by a voltage step-up transformer. A shielded helical quarter-wave resonator [78], consisting of a copper coil, serves as the inductor in a resonant LC-circuit in series with the capacitance of the trap itself ($C \approx 12$ pF). When attached to the trap, the resonator has a quality factor $Q = 140$. As shown

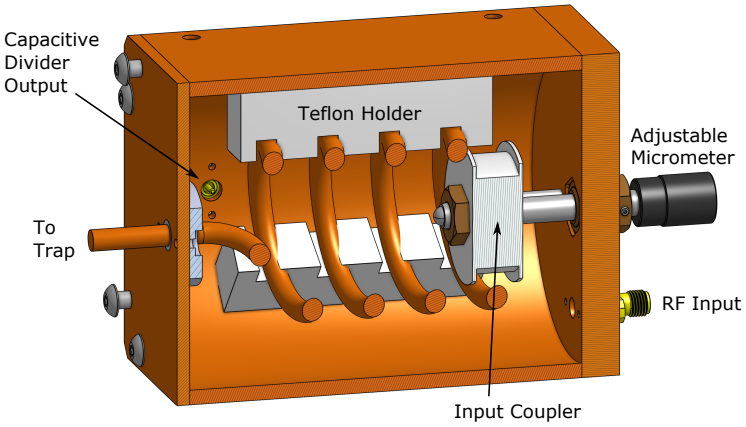

Figure 4.5: Model of the RF resonator used to general high-voltage trap RF. The resonant frequency when connected to the trap is designed at 36 MHz.

in Figure 4.5, an RF frequency source is fed via SMA into wound input coupling antenna. The antenna can be translated via an adjustable micrometer to impedance match the input signal to 50Ω. The signal couples into the helical copper coil that is mechanically stabilized by Teflon holders inside the cavity. Although the inside of the cavity has a cylindrical bore, the outside is square to ensure rigid mechanical mounting to the chamber.

At the other end of the coil, a low-temperature coefficient (Vishay QUAD HIFREQ) capacitive divider (36:1) picks off the amplified RF signal. This pickoff is fed into a rectifier and compared against a precision reference voltage (LM399H), with the difference signal serving as the input into a digital PID controller [79]. The PID controller outputs the feedback signal into the IF port of a mixer that sits between the frequency source (LO port of mixer) and the resonator (RF port of mixer). Since chip traps can easily breakdown if too much voltage is applied, one should be sure that even when the mixer is full saturated, the applied voltage from

the resonator is still well below the maximum voltage rating of the trap.

4.2.3: DC Voltage Controller

In order to transport ions around the chip, we must be able to independently control all 98 DC electrodes on the HOA and implement well timed sequences that execute shuttling operations. A custom built circuit board houses 25 DAC integrated circuit chips (16-bit, 4 channels, DAC8734) that are controlled by an FPGA (Opal Kelly XEM6010-LX150). Voltage solutions consist of several "lines", where each line specifies one output state for all 98 electrodes. A shuttling graph defines nodes (corresponding to a start/end line) and edges (all the lines between the start/end line). Voltage solutions and shuttling graphs are uploaded to the FPGA on-board memory from the main control computer via USB. For "coherent" shuttling operations that need to happen at precise times within a quantum circuit, the latency in USB communication is unacceptable. The FPGA also has interfaces for one serial SPI line and hardware trigger. Within a quantum circuit, the main control FPGA can specify a shuttling edge via the SPI line, that then executes in a deterministic amount of time at the start of the trigger (also from the main control FPGA).

The voltage lines from the controller box must also be heavily filtered before they reach the trap, as any noise that is near the motional frequencies of the ion will cause it to heat. A 5^{th}-order low-pass RC filter is applied to each voltage line immediately before the lines enter the vacuum feed-through. The corner frequency of the filters is designed at 1 kHz, which is a conservative design choice that limits the speed of shuttling operations to ~ 1 ms, but heavily mitigates risk of external

lab noise impacting the ion.

4.3: CW Laser Subsystem

The CW laser system provides all of the ionization, cooling, pumping, and detection required to trap and maintain ions and to use them as qubits. Each laser system is a self-contained module that sits on a nearby optical table and is routed to the beam delivery boxes mounted on the vacuum chamber by optical fiber.

The 935-nm, 399-nm, and 393-nm modules are quite simple. The 935-nm light is generated by an ECDL with a cateye configuration (AOSense). A pickoff optic couples the light into a fiber and is routed to a multichannel wavemeter (High-Finesse) for frequency measurement and stabilization. The remaining light passes through a free-space resonant 3.07 GHz EOM and polarization optics before being fiber-coupled and delivered to the vacuum chamber. The 935-nm is not delivered to a beam delivery box, rather $\approx$ 12 mW of light is focused along the axis of the ion chain to a waist of 650 μm to ensure uniform illumination of all the ions. After passing through the chamber, the beam is retro-reflected and directed over the Yb thermal oven load slot to ensure repumping during loading.

The 399-nm light is generated from an AR coated diode (Nichia NDUA116T) inserted in a commercial ECDL body (Moglabs) with a Littrow configuration. As with 935-nm, the light is first picked off and directed to another port on the wavemeter for frequency locking. The remaining 399-nm path includes a fast shutter [37] for binary amplitude control and polarization optics. Then it is frequency multiplexed

with 369.5-nm and 393-nm light, fiber-coupled and directed to the load slot that contains the Yb thermal oven. The 393-nm light is generated from a free-running laser diode (Nichia NDU4316) with no external cavity. The frequency is not stabilized since all the light is higher energy than the ionization threshold for neutral Yb. Similar to the 399-nm path, a fast shutter provides binary amplitude control before the light is multiplexed into the load slot fiber.

Due to the variety of functions it must perform (cooling, pumping, detection), the 369.5-nm module is a bit more complex. The 369.5-nm light is generated from a laser diode (Nichia NDU113E) inserted in a commercial ECDL body (Moglabs) with a Littrow configuration. The frequency is stabilized by PDH locking [80, 81] the laser to a transfer cavity. The length of the cavity is locked to a 780-nm laser that is in turn locked to the ^{87}Rb D2 ($F = 2 \rightarrow F' = 3$) line via modulation transfer spectroscopy on a Rb vapor cell [82]. With this frequency stabilization scheme, the 369.5-nm light is referenced directly to an atomic transition and thus protects against long term frequency drifts.

The 369.5-nm light is broken into two paths, the pump/detect path and the cooling path. The cooling path includes an always-on 14.748 GHz free-space EOM for repumping sidebands. The pump/detect path includes a switchable 2.105 GHz EOM that is turned on when performing optical pumping. Each path includes an unfolded double-pass AOM consisting of two separate AOMs (IntraAction ASM-2101LA8) that provide amplitude and frequency control. The effective double-pass set-up reduces leakage light that could decohere the qubit. These paths are recombined on a beam-splitter, with one output port going to the load slot fiber and the

other going to the quantum region fiber. The beam in the quantum region features a cylindrical lens that focuses the 369.5-nm light to a beam waist of $w_z \approx 20$ μm and $w_x \approx 90$ μm at the ion plane to ensure equal illumination across a chain.

4.4: Raman Subsystem

The Raman system provides the individual addressing of the ions and communicates gate waveforms to the qubits. A Raman operation requires two lasers beams. In this system we use a "global beam" that illuminates all the ions, and then 32 "individual beams" that provide the individual addressability. The global beam will illuminate idle qubits when operations are being performed on active qubits, but the only effect of this is a two-photon stark shift on the idle qubits (~ 50 Hz) which can be calibrated out (and assuming single-beam 4-photon stark shifts have been nulled through polarization control).

Both Raman beams are generated from the output of a 4W 355-nm laser (Coherent Paladin Compact 355-4000). A motorized $\lambda/2$ waveplate, a high-power beam cube, and a beam dump control the total power entering the system. The beams are then split with a 1:3.5 (individual:global) power ratio at another beam cube into the individual path and the global path. These two beam paths are depicted schematically in Figure 4.6.

The global beam path includes a motorized delay stage, which is used to temporally overlap pulses of the two Raman beams. A wideband, single channel, AOM (L3Harris) is used to modulate the beam. Due to the wide bandwidth, it

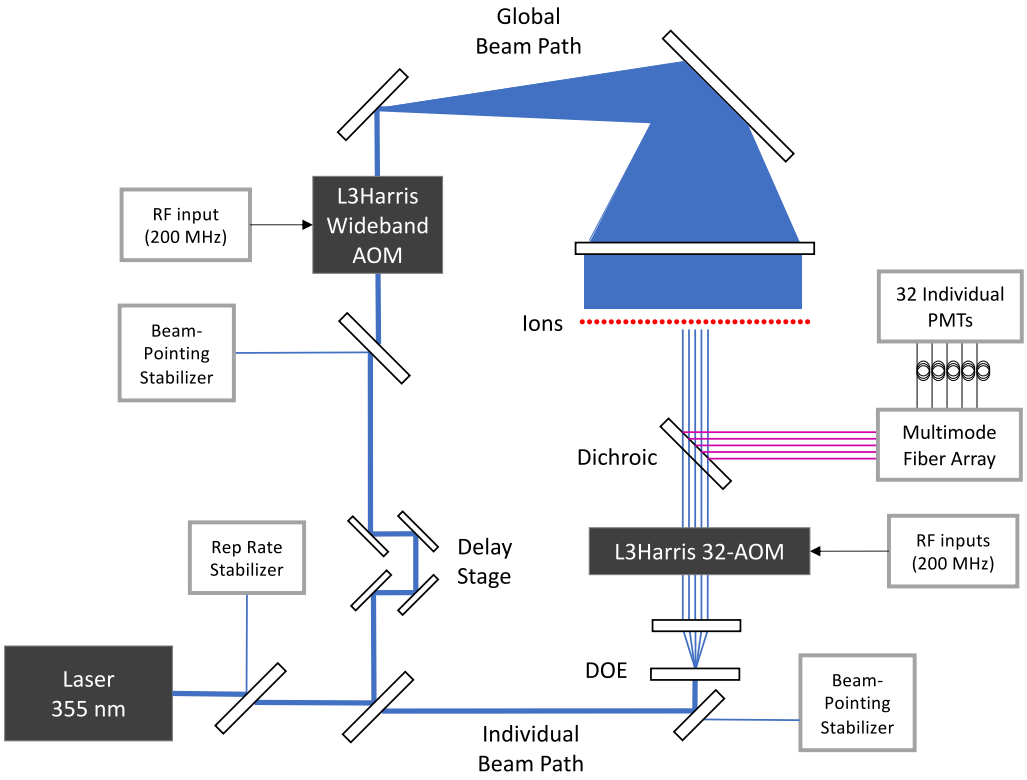

Figure 4.6: Schematic showing the basic components of the Raman subsystem. The global beam addresses the entire ion chain whereas the individual beams are controlled with a 32-channel AOM to implement gates on selected ions.

is typically preferred to put the red/blue tones of a two-qubit gate waveform on the global beam. The global beam path is designed to be *telecentric*, which means that even though the red and blue beams are deflected at different angles out of the AOM, they will overlap spatially at the ion plane (albeit with slightly different angles of incidence). The relatively low NA of the global beam path also mitigates this angular separation at the ions. After the AOM, the global beam is directed parallel to the chip surface along $\hat{y}$, shown in Figure 4.7. The polarization of the global beam $\hat{\epsilon}_g$ is set to linear along the $\hat{x}$ axis and is focused to an elliptical $1/e^2$ waist of $w_z = 10$ μm and $w_x = 166$ μm at the ion plane. There is ≈ 750 mW in the global beam at normal operating power.

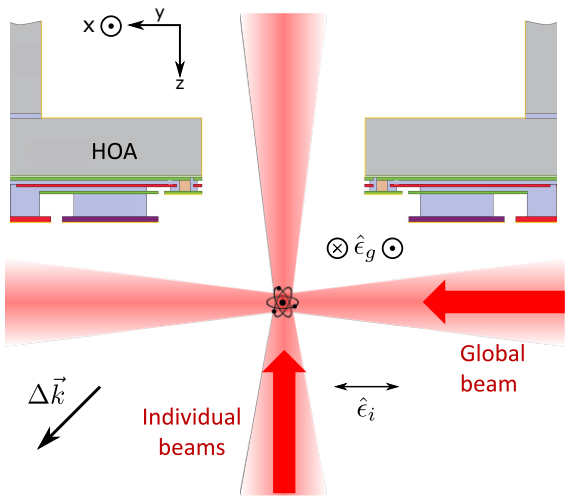

Figure 4.7: Schematic of the Raman geometry. The ion chain is oriented along $\hat{x}$ and the net $\Delta\vec{k}$ of the Raman beams is along $\hat{y}+\hat{z}$, which couples to one set of the radial modes tilted at 45° via a static Q_{yz} quadrupole.

The individual beam path includes a diffractive optical element that splits a single beam into 33 equal power beams. The beams are then directed to a 32-channel AOM (L3Harris) with nominal channel spacing of 450 μm and a center drive frequency of 200 MHz. By driving specific channels on the AOM, we can selectively turn on/off individual beams at the output. The output beams are then spatially overlapped with the imaging system using the reflection of a dichroic mirror. A custom 0.63 NA lens (Photon Gear) situated in the re-entrant window at the bottom of the chamber focuses each individual beam along the $\hat{z}$ axis, through the slot in the HOA (Figure 4.7). The polarization of the individual beams $\hat{\epsilon}_i$ is set to linear along the $\hat{y}$ axis. Each beam is focused to a circular $1/e^2$ radius of $w = 870$ nm and spaced by 4.43 μm at the ion plane. There is ≈ 1 mW in each beam at normal operating power.

In both paths, there are angle and position beam sensors that are used to

stabilize beam pointing. The net $\Delta\vec{k}$ of the Raman beams is along $\hat{y} + \hat{z}$. We typically operate the machine with a Rabi frequency of 500 kHz. At full power we have achieved a Rabi frequency of 2 MHz, although we purposely de-rate from this value to leave laser power headroom as burns in the system develop. For two-qubit gates there is a nice separation of responsibilities between the global and individual beam. As discussed above, the global path is better suited for communicating frequency content to the ions, including either static red/blue tones or FM modulation waveforms. Then the individual beams are responsible for communicating phase and amplitude information. Putting phase information in the individual beams is a requirement for using software R_z gates. Amplitude information could be put on either beam, but due to the 2-photon stark shift of the global beam on idle qubits, this effect is easier to calibrate out when the global beam intensity is held constant.

4.4.1: Rep Rate Control and Stabilization

To control the four-photon Stark shifts, we need control over the repetition rate of the laser (see Section 3.1.4). The repetition rate is defined by the length of the cavity and is set during the manufacturing process of the laser. Using a cleanroom to avoid contamination of the high power laser cavity, we open up the Paladin laser and install a custom output coupler on a translatable stage (SmarAct SLC-1730-S-HV) to dynamically tune the length of the cavity. The custom output coupler is shown in Figure 4.8. To achieve realignment of the optical cavity, a fiber based autocollimator system [83] is used along with the three degrees of freedom (tip, tilt, rotation) available on the output coupler mount (SmarAct STT-25.4-HV,

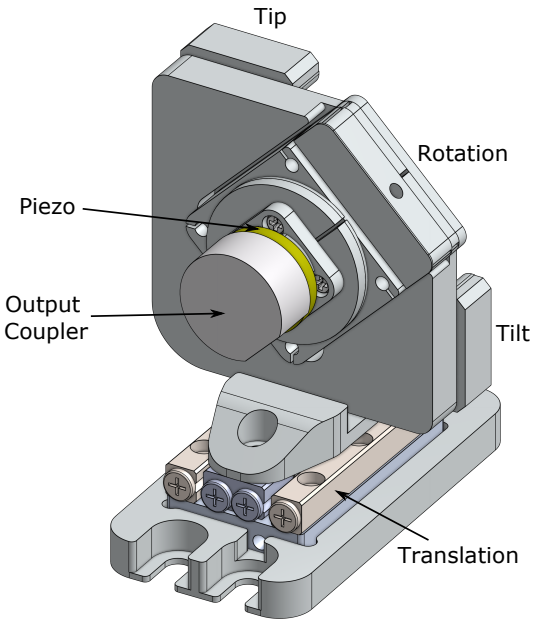

Figure 4.8: Modified Paladin output coupler that enables tuning of the repetition rate of the laser.

SR-1908-HV). A piezo ring sandwiched in between the output coupler and the mount is used to stabilize the length of the cavity.

One flaw in this particular design is that the bandwidth of the piezo is limited to $\sim$ 1 kHz due to mechanical resonances in the mirror mount. This bandwidth is insufficient since acoustic noise can range up to 20 kHz in frequency. Even with the piezo feedback enabled, we measure the standard deviation of phase noise to be 52° over 10 s, largely at discrete acoustic frequencies. To further reduce this phase jitter, we feed-forward the residual phase noise onto the the RF signals going into the AOM, as has been demonstrated previously [84]. With this, the standard deviation of phase noise is reduced to 1.1° over 10 s. This solution is practical since our RF control system only uses three independent RF channels, each requiring

a feed-forward circuit. In future systems, where an RF channel is dedicated to each AOM channel, it may be preferable to suppress the noise at the source. With improved design, piezo mirrors in cavities can suppress noise up to 180 kHz [85].

4.5: Imaging Subsystem

The imaging system is used for detecting the state of the qubits at the end of a quantum circuit. In this system the Raman system and the imaging system are tightly coupled because they both they are combined into the same objective lens via a dichroic mirror. This has some advantages in that once the paths are overlapped, only the fine alignment of the shared objective lens needs to be tracked. However, in future systems it may make sense to decouple these two systems, which will mean imaging and Raman alignment will need to be calibrated independently.

Starting at the ions, light is collected through the bottom re-entrant viewport by a custom 0.63 NA objective lens (PhotonGear) with an object working distance of 16 mm. The objective images the ions into an intermediate imaging plane approximately 200 mm from the back of the lens with a measured magnification of $M = 9.2$. The objective is mounted on a six degree-of-freedom hexapod stage (PhysikInstrumente H811) that provides rough alignment to the ions (~ 1 μm at ion plane). A bounce mirror underneath the motorized stage reroutes the imaging light to the bottom plate, where a piezo-controlled mirror provides fine alignment ($\sim$ 10nm at ion plane). The imaging light is transmitted through the dichoric mirror that couples to the Raman system and passes through another set of lens that provides an

Figure 4.9: Multi-mode fiber array and PMT imaging system. (Left) The input face of the multi-mode fiber array under backside illumination. (Middle) On the output side, the fiber bundle is broken out out in 32 individual optical fiber cords. (Right) Each fiber output is connected to an individual narrow-band filter and PMT module.

addition magnification of $M = 3$ for a total magnification of $M = 28$. Finally the light is imaged onto either a camera (Andor Zyla) or a custom multi-mode fiber array (FTO), chosen by a motorized flip mirror.

While the camera is useful for initial alignment and debugging purposes, it is not suitable for fast high fidelity state detection due to its low frame rate, high background dark counts and spatial/electrical crosstalk. The fiber array consists of 32 multi-mode fibers, each with 0.22 NA and 100 μm core/125 μm cladding diameter, and spaced by a 125 μm pitch. With a diffraction limited ion spot size of $\lambda/2\text{NA} \approx 300$ nm and a total magnification of 28, the ion image at the fiber array will have a spot size of 8.2 μm with a pitch of 125 μm (the magnification of the imaging system is set by the multi-mode fiber pitch relative to the ion spacing). The NA of the fiber also helps to filter out background scattered light that may be incident on the fiber array. We characterize the spatial crosstalk of this imaging system in Section 5.3.1.

Previous experiments have used an integrated multi-channel PMT array, where

the dominant form of crosstalk was electrical [77]. The fiber array mitigates this by breaking out the fiber bundle into individual fibers on the other end, as shown in Figure 4.9. Each individual fiber is then connected to an individual fiber-adapted PMT module (Hamamatsu H10682-210), which is shielded and physically separated from the other sensors. With this design we see no evidence of electrical crosstalk. Additionally a narrow-band 369.5-nm filter (Semrock FF01-370/6-25) sits in between the fiber bundle output and the PMT module to further reduce background light. Each photon that is detected on the PMT sends out a digital voltage spike to a coaxial cable, which are then counted by the main control system.

We estimate that $\eta_C = 7.4\%$ of the total light is collected from the ion, given by,

$$\eta_C = [1 - \cos(\arcsin(\mathrm{NA}))]\ \eta_{\mathrm{GS}}\ \eta_{\mathrm{MM}} \tag{4.2}$$

Where NA= 0.63 is the numerical aperture of the initial objective lens, $\eta_{GS} = 88\%$ is clipping due to misalignment of the in-vacuum ground shield, and $\eta_{MM} = 75\%$ is an estimated coupling efficiency of the light into the multi-mode fiber. Then the total estimated detection efficiency is $\eta_D = 1.1\%$ of the total light from the ion, given by,

$$\eta_D = \eta_C\ \eta_F\ \eta_{QE} \tag{4.3}$$

Where $\eta_F = 90\%$ is transmission of the narrow-band filter and $\eta_{QE} = 17\%$ is the quantum efficiency of the detector.

4.6: Control Subsystem

There are many different pieces of hardware that must be commanded and controlled with precise timing at various points within an experiment in order for the system to function. On the slow scale, ion lifetimes, laser powers, and voltage calibrations must be tracked and managed. On the medium scale, CW lasers with correct frequency and sideband structure need to be switched on/off in order to perform cooling, preparation and detection operations. And on the fast scale, frequency and phases of RF sources must be rapidly switched to execute quantum circuits.

In this system, we draw a line between "coherent" sequences (that is, where the phase of the qubit matters), and "incoherent" sequences. A main control FPGA executes all of the incoherent processes such as cooling, state preparation, and photon counting. Then, at the start of a quantum circuit, the system is handed off to the coherent RF subsystem via a trigger. The coherent RF subsystem defines the circuit and contains all the gate waveforms with correct frequency/phase/amplitude, all routed to the correct subset of ions. When the coherent sequence is finished, the system is handed back over to the incoherent main control subsystem, whose immediate task is to perform state detection and count photons, before continuing on to the next task.

4.6.1: Main Control Subsystem

The main control system consists of an FPGA (Xilinx KC705) mounted to a custom breakout board. The breakout board provides 32 digital output lines (SN64BCT25244DWR, line driver), 24 digital input lines, 8 DAC channels (DAC8568), 8 ADC channels (AD7607), and control over 10 DDS chips (AD9912). The DDS chips are a single-tone programmable frequency source, suitable for running CW laser AOMs or simple quantum experiments (e.g. Rabi, Ramsey, frequency scans, etc.), however their capability falls short of what is required to run complex quantum circuits. In particular, because they are single-tone, two-channels would have to combined to create a bichromatic waveform required for two-qubit gates. The four-photon stark shift during two-qubits gates depends critically on the amplitude balance between these two tones, so when the two tones come from different sources, their relative amplitude must be meticulously tracked and calibrated. Additionally, the amplitude control on our particular chip (AD9912) is too slow to implement AM two-qubit gates. Even with other DDS chips with faster amplitude control (AD9910), a more dedicated system with native dual-tone support is preferred.

The main FPGA is programmed via the ARTIQ ecosystem, which features the ability to take high-level programming syntax (Python) running on the host computer and compile it onto the FPGA hardware. This approach has the benefit of the nanosecond timing resolution afforded by the FPGA, while maintaining a Python codebase that is readily accessible without knowledge of low-level FPGA programming languages (e.g. Verilog, VHDL). While FPGAs excel with simple out-

put sequences that are tightly timed, they struggle with more complex mathematical operations involving floating point arithmetic. The FPGA is able to hand off these tasks to the host computer via a remote procedure call (RPC) and wait for the response before continuing on with its timing sequence.

The ARTIQ ecosystem also features a networked approach to hardware device drivers that is able to support equipment hooked up to various computers across the lab. This makes the system scalable, because wires need not be routed across the lab to a centralized control computer. Rather, the centralized control computer interacts with auxiliary computers across the lab through the LAN and a client/server model. The ARTIQ ecosystem also features a GUI for user interaction, an experiment scheduling system that is critical to automation, and databases that store experimental data and calibration information. With this structure, the main control system is able to maintain awareness about the state of the machine and automatically schedule experiments and calibrations.

4.6.2: Coherent RF Control Subsystem

The most complex gate waveform in this system is the two-qubit gate, which consists of a bichromatic tone with fast modulation ($\sim$ 10 μs) of the amplitude (or phase/frequency for other gate designs). Additionally, these waveforms must be reconfigurably routed to any 2 of the 32 channels of the multichannel AOM in order to implement all-to-all connectivity of the ion chain. An AWG is a natural choice as an RF source, because they are designed to implement arbitrarily complex waveforms. However, AWGs are rather expensive and it may be cost prohibitive to

allocate a dedicated AWG to each of the 32 AOM channels. Additionally, most the time spent by a qubit in a quantum algorithm is idle time where its is not actively being operated on. AWGs sample waveform data written in memory, and so this idle time must still be written and uploaded to the AWG across all the channels, potentially forming a data bottleneck.

The solution in this system is to use a single AWG card with four output channels (Keysight M3202A). One channel is routed to a oscilloscope for monitoring and debugging purposes. One channel is used to drive the global beam. The remaining two channels drive the individual beams and are routed to a custom built 5×16 switch network. Each of the 5 inputs are split into 16 and connected to the input channels of 16 SP5T RF switchs (EVAL-ADRF5250). The 16 outputs of all switchs are connected the middle channels of the 32-channel AOM. A single switch can point an output channel to any one of the 5 input channels (or ground) specified by a 3-bit digital register. At a given time, the state of the switch network is fully defined by 3×16 bits. As an example, the switch network state for a single two qubit gate between ions i and j will point the AOM channel corresponding to ion i at the first AWG channel and the AOM channel corresponding to ion j at the second AWG channel, with all the other output channels connected to ground. It is important that there is an individual AWG channel for each ion, since software R_Z gates may require different phases on the two ions. The switch network state is controlled by 2 digital pulser cards (PXI-6541) that are programmed in conjunction with the AWG. The AWG triggers the cards to switch states on the backplane of a common chassis, as it steps through gate waveforms. The other 3 inputs to the

switch network are DDS channels for simple experiments that wish to bypass the complex AWG/switch network compilation procedure

While the current system only supports 16 AOM channels, it is possible to extend it to support the full 32 channels of the individual AOM. However, the system has limitations even at its current size. In particular, complicated circuits can take ∼minutes to compile, relative to the ∼seconds required to gather shot noise statistics. Thus, the system is not particularly suitable for algorithms that require ∼thousands of different circuits to generate a result, such as QAOA [86] or GST [87].

In the near-future, we plan to upgrade the system using RFSoC technology (Xilinx Zynq UltraScale+), which will make it possible to give each AOM channel a dedicated RF line. More importantly, using a reduced waveform instruction set, it is feasible to stream waveforms directly to the cards, completely eliminating compilation overhead. The system is available as an evaluation board (ZCU111) or from third party manufacturers (Pentek 5950). The latter option also provides the ability to phase synchronize channels across multiple cards (Pentek 5903), which may be required for active crosstalk cancellation or a phase-insensitive gate configuration.

Chapter 5: System Operation, Calibration, and Characterization

5.1: System Operation

5.1.1: Autoloading Ions

The average lifetime of a 15-ion chain in our system is ∼30–45 min. In order to reliably run circuits over extended periods of time, the system must detect ion loss events and automatically reload a chain with the correct number of ions. To detect loss events, photon counts are monitored during the second-stage cooling process of each experimental shot. If a loss event is detected, the system will automatically attempt to reload a chain. Recall that each ion is individually loaded in the load slot and then shuttled/merged into the existing chain in the quantum region. The flowchart for this process is depicted in Fig. 5.1(a).

There are no detection optics above the load slot, which makes this process probabilistic (i.e., we do not know whether we have loaded an ion in the load slot until we try to merge it into the quantum region). After warming up the neutral Yb oven, each load attempt begins by clearing the loading region ("eject") by applying

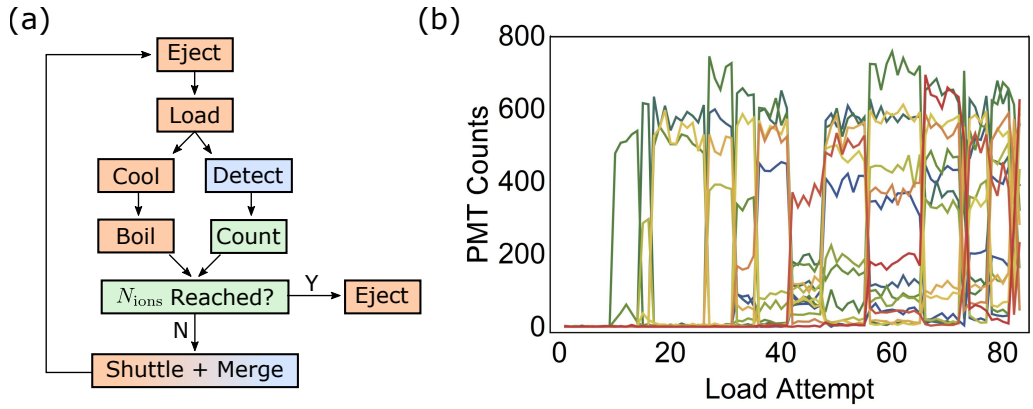

Figure 5.1: Autoloading chains of ions. Each ion is individually loaded in the load slot, and then shuttled and merged with the existing chain in the quantum region until the desired chain length is achieved. (a) Flowchart for autoloading chains. Orange boxes occur in the load slot. Blue boxes occur in the quantum region. Green boxes are calculations that are processed on the host computer. (b) Typical example of the detection counts in the quantum region versus the load attempt. Chain load events are detected by looking for transitions in the detection counts.

high DC voltages to the outer electrodes of the load slot. After returning to a trapping solution, we attempt to load by simultaneously switching on the 399/393-nm ionization beams. Then we switch the load beams off and allow the ion to cool (if an ion was actually trapped). Simultaneously, we collect photons from the existing chain in the quantum region. If we attempt to merge a hot ion into a cold chain it can melt the crystal and eject ions from the trap. To throw out ions that have not sufficiently cooled during the cooling stage, we lower the trapping potential so that hot ions are "boiled" out of the load slot. Simultaneously, we calculate on the host computer whether we think an ion was loaded on the previous load attempt and correspondingly update the estimated length of the chain in the quantum region. If we have reached the desired chain length, the loaded ion is ejected from the load slot, otherwise we shuttle it to the quantum region and merge it with the existing chain.

We use a fixed axial potential for building up chains in the quantum region, which means the ions are not necessarily aligned to imaging optics. On each load attempt we collect photons in the quantum region and look for transitions in the PMT pattern to detect whether an ion was successfully added to the chain. In the "detect" block of the autoloading process, we collect a vector of PMT counts $\vec{c}_i$ from the quantum region corresponding to load attempt i. The pattern of $\vec{c}_i$ as ions are added to the chain is shown in Fig. 5.1(b). To detect transitions, we define the metric m,

$$m = \frac{||\vec{c}_{i-1} \cdot \vec{c}_i||^2}{1 + ||\vec{c}_{i-1}|| \cdot ||\vec{c}_i||} \tag{5.1}$$

Where $||\cdot||$ indicates the vector-norm. If $\vec{c}_{i-1} \approx \vec{c}_i$, then $m \approx 1$, whereas if $\vec{c}_{i-1} \perp \vec{c}_i$, then $m = 0$. By setting a threshold τ, we say that an ion has been added to the chain when $m < \tau$ and increment the estimated chain length by one. For first ion, this metric is noisy, thus we add an additional constraint that $||\vec{c}_i||^2 > 20^2$ photons.

5.1.2: Ground State Cooling

At the start of a circuit, the ions should be cooled as close to the motional ground state as possible. The Doppler limit is still several quanta of energy, and therefore a sub-Doppler cooling technique must be used in order to achieve ground state cooling. Sub-Doppler techniques typically rely on pumping the ion into a dark state, so that they no longer scatter photons, which is a heating process. Raman sideband cooling [88] consists of applying a π-pulse on the RSB transition, followed by optical pumping, $|\downarrow, n\rangle \rightarrow |\uparrow, n-1\rangle \rightarrow |\downarrow, n-1\rangle$ so that the net effect is to remove

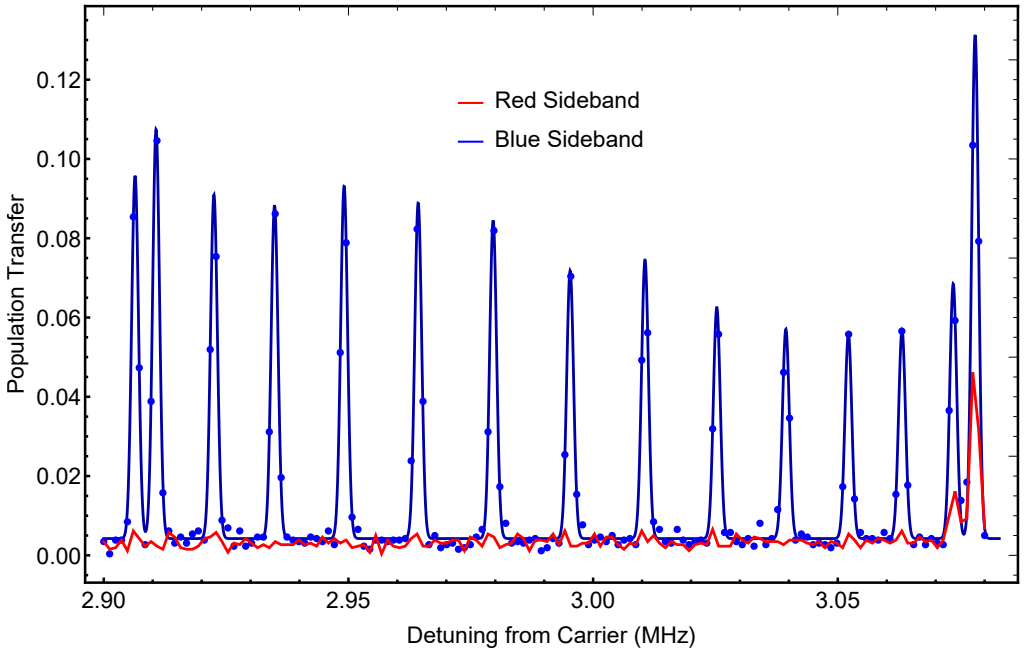

Figure 5.2: Blue/Red sideband frequency scan across the radial mode spectrum following SBC on a 15-ion chain. The population transfer is averaged across all the ions. The asymmetry between the red and blue sideband indicates that the motional modes are near the ground state.

one quanta of energy. Repeating this process eventually results in $|\downarrow, n = 0\rangle$, which is a dark state under application of either a RSB pulse or optical pumping. SBC is immediately followed by optical pumping and then the quantum circuit.

The result of SBC on a chain of 15-ions is shown in Figure 5.2, where we scan the radial gate modes on both the red and blue sideband. The higher-spatial frequency modes do not heat from anomalous electric field noise, and are therefore very close to the motional ground state $\overline{n} < 0.05$ (the RSB transfer does not exceed the noise floor of the measurement). The two lowest spatial frequency modes do couple to noisy electric fields and therefore heat during the process, resulting in $\overline{n} \approx 0.5$ for the highest energy in-phase mode.

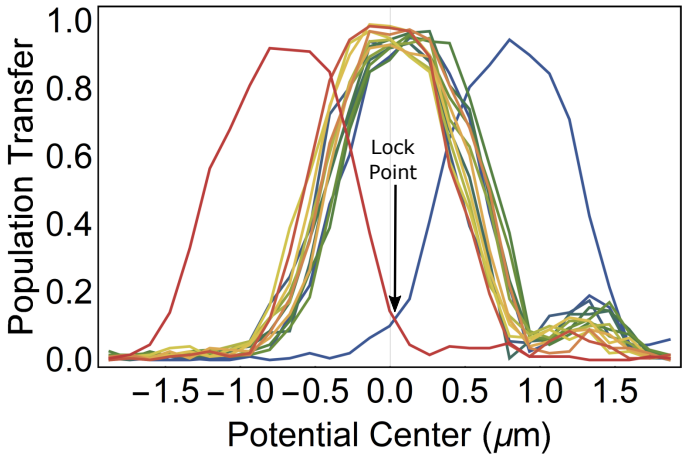

Figure 5.3: The population transfer as chain of 15 ions are scanned across the individual Raman beam by translating the axial potential. The middle 13 ions are well aligned to the Raman beams, but the 2 edge ions are not (see Section 2.1.3). We can use their crossing point as a signal to lock the Raman beams to the remaining 13 ions. The +7(-7) edge ion corresponds to the red (blue) curve.

5.1.3: Ion-based Beam Pointing Lock

It is critical to maintain precise overlap between the ions and the Raman beams to stabilize gate amplitudes. The chain can move relative to the beams due to stray background electric field (X_1) or the beams can move relative to the chain due to thermal drifts or air turbulence. In this sense, the ions themselves are the best sensors for any misalignment. In Figure 5.3, we show the population transferred from a fixed Raman pulse as a chain of 15 ions is translated across the individual beams. The middle 13 beams all achieve good alignment at $x = 0$ μm, which means they are equispaced by the design of the potential (Section 2.1.3). The two edge ions deviate from equispacing, again by design, however they have a crossing at $x = 0$ μm that can be used to stabilize the individual beam pointing.

On each experimental shot, before main circuit preparation, we perform a mini-

calibration experiment. The experiment involves Doppler cooling, optical pumping, and then a single Raman pulse on the two edge ions tuned to achieve roughly half transfer at $x = 0$ μm. State detection is performed on the ions and the states are fed directly into an integrator feedback loop $\mathcal{I}$ that controls a piezo mirror that steers the beam positions along $\hat{x}$. For example, if the +7 (-7) ion measures in the $|0\rangle$ state then the integrator is incremented $\mathcal{I}$ $-= g$ ($\mathcal{I}$ $+= g$), where g corresponds to the gain of the feedback loop. The feedback output is scaled so that $\mathcal{I} = 0$ corresponds to the middle of the piezo voltage range with nominal pointing at $x = 0$ μm. In this notation, $\mathcal{I} < 0$ ($\mathcal{I} > 0$) corresponds to steering the beams in the $-\hat{x}$ ($+\hat{x}$) direction. The lock bandwidth is limited by the experimental duty cycle, which depends on the length of the circuit that follows the pointing calibration experiment, but in this system is $\approx$ 200 Hz. Since we do not wish to inject quantum shot noise into the pointing lock, the locking bandwidth is limited to $\sim$ 10 Hz.

5.2: System Calibration

5.2.1: Nulling Stark Shifts via Laser Repetition Rate

In Section 3.1.4, we showed that the cross-beam four-photon Stark shift caused by off-resonant comb teeth can by theoretically set to zero by adjusting the repetition rate of the laser. Here "cross-beam" refers to the off-resonant *inter*-comb teeth that scales like the product of the intensities of the individual and global beam ($\sim I_g I_i$). This is in comparison to the *intra*-beam four-photon stark shifts due to residual polarization circularity in either the global or individual beam alone ($\sim I_g^2, I_i^2$).

Additionally, there are the standard two-photon Stark shifts (Eq. 3.21) that scale like the intensity of a single beam ($\sim I_g, I_i$). Thus the total Stark shift during a carrier operation can be described as,

$$\delta = \alpha_{\mathrm{i}}^{(2)} I_{\mathrm{i}} + \alpha_{\mathrm{g}}^{(2)} I_{\mathrm{g}} + \alpha_{\mathrm{i}}^{(4)} I_{\mathrm{i}}^2 + \alpha_{\mathrm{g}}^{(4)} I_{\mathrm{g}}^2 + \alpha_{\mathrm{x}}^{(4)} I_{\mathrm{i}} I_{\mathrm{g}} \tag{5.2}$$

Where α is a constant that contains all the atomic physics of the shift. The cross-beam parameter $\alpha_{\mathrm{x}}^{(4)}$ depends on f_{rep} and is the free-parameter we wish to tune. We assume $\alpha_{\mathrm{i}}^{(4)}$ and $\alpha_{\mathrm{g}}^{(4)}$ can be set to zero using polarizers and waveplates. It might be tempting to tune f_{rep} so that the total Stark shift is zero ($\delta = 0$), however static shifts can in principle be calibrated out in circuits using RZ gates. More detrimental to the fidelity of circuits is fast Stark shift noise or slow drifts that occur at the timescales faster than can be calibrated out. The strategy taken in this experiment is to null the dependence on the individual beam intensity, I_i, with the assumption that it will be the dominant intensity fluctuations in the system due to the small beam size. The goal is then to tune f_{rep} so that,

$$\alpha_{\mathrm{i}}^{(2)} I_{\mathrm{i}} = -\alpha_{\mathrm{x}}^{(4)} I_{\mathrm{i}} I_{\mathrm{g}} \tag{5.3}$$

Which would eliminate the dependence of δ on I_i. The quantity $\alpha_{\mathrm{i}}^{(2)} I_{\mathrm{i}}$ can easily be measured by turning on the individual beam only during one leg of a MW Ramsey echo interferometer. In our experiment, $\alpha_{\mathrm{i}}^{(2)} I_{\mathrm{i}} = -76$ Hz at the individual intensity used to drive single-qubit gates. The cross-beam stark-shift $\alpha_{\mathrm{x}}^{(4)} I_{\mathrm{i}} I_{\mathrm{g}}$ is more difficult

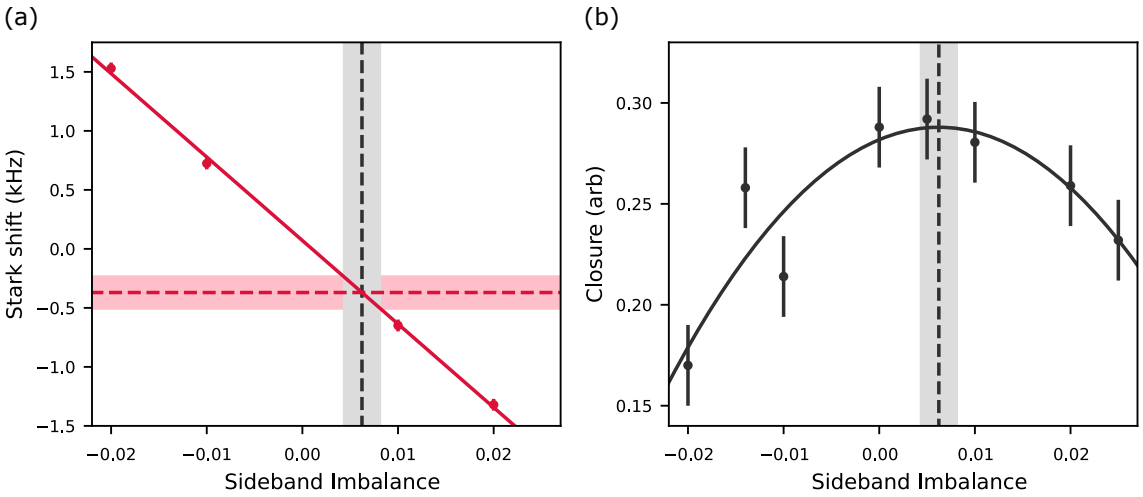

Figure 5.4: (a) Changing the sideband imbalance of a bichromatic tone creates a linear shift in the Stark shift measured during a MW Ramsey experiment. (b) To balance the two tones, we find where the optimal motional closure occurs from the experiment described in Fig. 5.5. This value indicates at which sideband imbalance the two tones become balanced (grey bar), which allows us to deduce the stark shift due to the off-resonant comb teeth (red bar in (a)).

to measure because the interaction necessarily drives qubit rotations. One approach to measure the shift due to off-resonant comb teeth *only*, is to split the resonant comb-tooth with two tones symmetrically detuned ($\Delta = \pm \sim 2$ MHz) from the carrier transition (similar to a MS interaction). If the tones are significantly detuned from the carrier and motional modes then no spin flip occurs. Additionally if the two tones are balanced in amplitude, then the Stark shifts from each tone cancels exactly. Since the splitting of the off-resonant comb teeth is much less than f_{rep}, this does not significantly alter their contribution to the four-photon Stark shift. Thus, placing a *balanced* bichromatic tone in one leg of a MW Ramsey echo interferometer measures the stark shift due to the off-resonant comb teeth only.

The challenge then reduces to ensuring that the red and blue tones are truly balanced in amplitude. There is no guarantee of this experimentally due to slopes

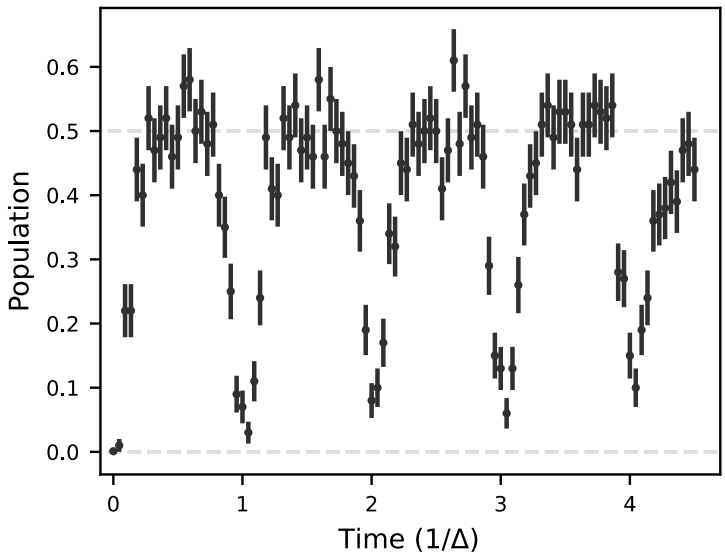

Figure 5.5: Motional closure experiment. An MS interaction is tuned Δ away from a single motional mode, which entangles spin with motion and causes the spin to randomize. At periods of $1/\Delta$, the motional excitation closes (i.e. disentangles from spin) and the spin returns to the $|0\rangle$ state.

in AOM efficiency, optical path differences, and other frequency-dependent effects in the RF chain. Scanning the amplitude imbalance between the two tones creates a linear change in the Stark shift (Fig. 5.4(a)) and somewhere along that line lies the point where the tones are truly balanced at the ion. We define the sideband imbalance (κ) by $I_b/I_r = \frac{1+\kappa}{1-\kappa}$ and find the true κ value by looking at the optimal motional closure, since the MS Hamiltonian only factorizes properly (disentangles spin from motion) when these two tones are balanced. With one ion, detuning close to a motional mode by Δ excites motion that returns to the initial state (i.e., the motion closes) with period $1/\Delta$. This experiment is shown in Fig 5.5. Sitting on the fourth dip for sensitivity and measuring the absolute value of the amplitude of the dip over a range of κ values yields the quadratic curve shown in Fig. 5.4(b). The peak of the quadratic curve gives the location where the two tones are balanced.

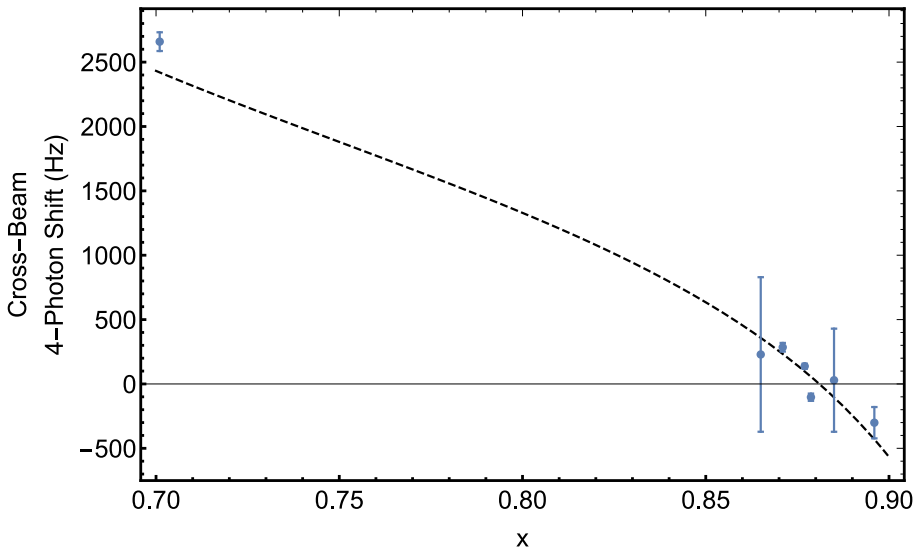

Figure 5.6: Nulling the 4-photon stark shift by tuning the repetition rate of the laser. Experimentally we find the value of f_{rep} that nulls the sensitivity to individual beam intensity at $x = 0.877$.

Equipped with this knowledge, we can return to Fig. 5.4(a) and finally deduce the cross-beam stark shift.

Repeating the procedure above at varying values of f_{rep} (or equivalently, varying values of x defined in Eq. 3.30) allows us to iterate until the optimality condition in Eq. 5.3 is met. The result of this process is shown in Fig. 5.6. To within the experimental uncertainty, we find the optimal value at $x = 0.877$, with the cross beam stark shift measured as $\alpha_{\text{x}}^{(4)} I_{\text{i}} I_{\text{g}} = 120(56)$ Hz. The precision in this measurement is limited by the uncertainty in the quadratic fit of motional closure versus sideband imbalance (Fig. 5.4(b)) since the cross-beam stark shift has a rather large slope of 700 Hz per 0.01 units of κ. The solid line corresponds to the four-photon Stark shift model given in Eq. 3.34 with a one-parameter fit to the pulse duration, $\tau = 11.7$ ps.

5.2.2: Voltage Calibrations

The locations of the radial modes and therefore the gate fidelity depends strongly on the DC potential applied to the chain. These voltages must be calibrated to ensure consistent high-fidelity operation of the machine. In Equation 4.1, we provide the model used to describe the HOA potential. In particular there are four terms used to describe a Taylor expansion of the axial potential in linear, quadratic, cubic and quartic terms with corresponding unitless multipliers X_1, X_2, X_3 and X_4. In Section 2.1.3, we describe how these multipliers relate to physical values. For our normal 15-ion chain operation, we desire odd terms to be zero (X_1,X_3=0) and the even terms to achieve optimal equispacing of the ions ($X_2, X_4 = 0.00188, 0.00177$).

Calibrating X_1 and X_2 is accomplished on a single ion. First, we apply a strong axial confinement of $X_2 = 0.4$ (axial frequency 630 kHz) to find the center of the well x_0. Then we relax the axial confinement to $X_2 = 0.01$ (axial frequency 100 kHz) and look to see if the ion shifted in position relative to x_0, indicating a stray linear field. Then we adjust X_1 until the ion is centered again at x_0. Next, we scan the axial sidebands to look for the motional peak at 100 kHz. Even though our Raman has no $\Delta\vec{k}$ along the axial direction, by offsetting the center of the beam relative to the ion, we can couple to axial motion (see Section 6.1.2). A long, low-power pulse should be used to avoid Stark shifts from the nearby carrier transition. Finally, we add an offset to X_2 until the mode is placed exactly at 100 kHz.

Calibrating X_3 and X_4 is accomplished on a chain of 15 ions, where the applied potential now has a calibrated X_1 and X_2 value. The location of the center ion

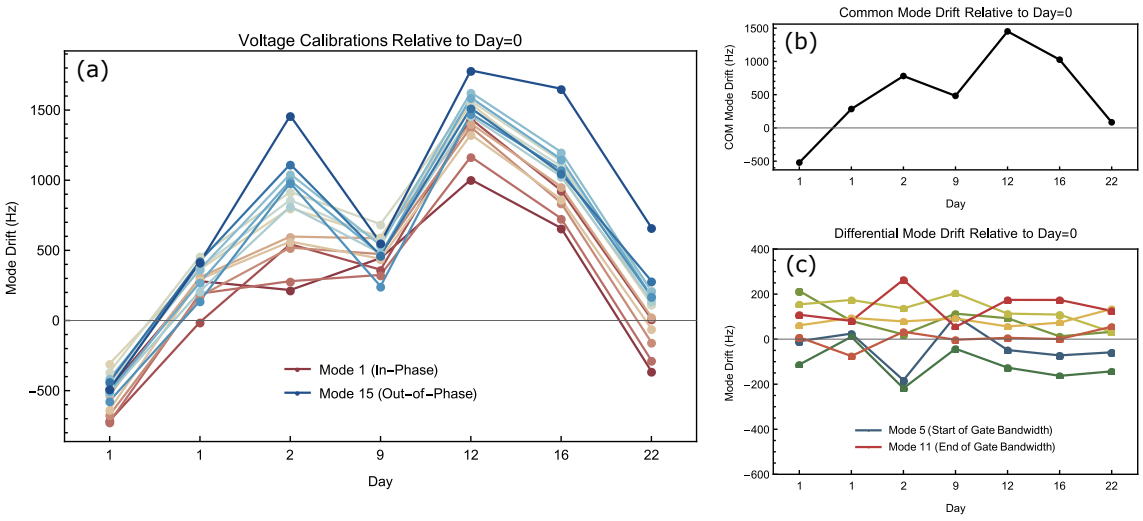

Figure 5.7: The consistency of the mode calibrations over the course of approximately a month.

relative to the previous value x_0, gives the signal required for tweaking X_3. We note that these calibrations are sequential and X_3 should not be calibrated until X_1 has been properly set. We adjust X_3 until the chain is again centered at x_0. For X_4, we again look at the frequency of the axial modes, but this time we use one of the high-spatial frequency modes, which are particularly sensitive to shifts in the quartic potential. We adjust X_4 until the mode overlaps with our reference value. The reference value is generated from a numerical mode solver that models the potential described in Eq. 4.1. The absolute frequency of this mode is less important (assuming the ions are still well spaced) than the stability and repeatability of this mode.

The consistency of this voltage calibration procedure is shown over the timescale of approximately a month in Fig. 5.7(a). On a given day, the measured mode spectrum $\vec{\omega}_i$ is compared against a reference mode spectrum $\vec{\omega}_0$ measured on day zero, with the mode drift given by $\vec{\omega}_0 - \vec{\omega}_i$. In general, we see that the mode error relative

to the reference spectrum is primarily a fixed constant offset across the spectrum. The common mode drift is defined as the difference in the mean of the two spectra, $\overline{\vec{\omega}_0} - \overline{\vec{\omega}_i}$, and is shown in Fig. 5.7(b). This value drifts over the range of ± 1 kHz, but is easily calibrated out by adjusting the DC quadrupole field Q_{yz} or by offsetting the gate detuning of all the gates.

The differential mode drift is the error that is not explained by a simple offset in the mode spectrum. This error is defined by $(\vec{\omega}_0 - \overline{\vec{\omega}_0}) - (\vec{\omega}_i - \overline{\vec{\omega}_i})$ and is shown in Fig. 5.7(c). This error is not simple to calibrate out, however the magnitude of this error is roughly ± 200 Hz and is contained within the 10^{-4} detuning error band of the optimized gate solutions. Therefore, by compensating only the common mode drift, we have high confidence that the gates will maintain good motional closure following the voltage calibration. However, we note that the effect of the differential mode drift may also change the mode participation factors, and therefore the XX gate angle, which should be re-calibrated following a voltage calibration. Empirically we observe that the voltage calibration remains valid over ~ 4 hours.

5.2.3: Gate Calibrations

Gates must have both the correct amplitude and phase. The one-qubit gate is fundamentally more robust due to the use of composite pulse sequences that suppress amplitude fluctuations. Assuming the Stark shift dependence on individual beam intensity has been nulled (see Section 5.2.1), then the single-qubit gates is also more robust from a phase perspective. The two-qubit gate does not benefit from composite pulse sequences, as the extra rotations significantly decrease fidelity. Additionally,

as shown in Fig. 5.4(a), a few percent amplitude imbalance in the red/blue tones of the two-qubit gate waveform can cause $\sim$ kHz Stark shifts. This effect is not present in single-qubit gates because they operate on the carrier transition. Thus, the primary operation that needs to be continually calibrated in this system is the two-qubit gate.

Assuming the static potential has been calibrated, the radial mode frequencies are well within the gate detuning error that can be accommodated by our optimized, robust gates. Therefore, the detuning of the gate relative to the modes does not have to be calibrated in order to achieve high-fidelity operation. To calibrate the gate angle (amplitude) for a given gate, we apply $[XX(\pi/4)]^{2N+1}$, $N = 0, 1, 2\ldots$ pulses to a state initialized to $|00\rangle$ and look at $P(|00\rangle)$ and $P(|11\rangle)$. For a perfect $\pi/4$ angle these probabilities should be equal, $P(|00\rangle) = P(|11\rangle) = 0.5$. We adjust a global multiplier to the entire AM gate solution until this is achieved. Increasing N will increase the sensitivity of this calibration due to coherent build up of the error, but may also exhibit time-dependant effects that can systematically miscalibrate the gate angle. We use $N = 0$ with several thousand repetitions to reduce the effects of shot noise.

As shown in Fig. 5.4(a), the imbalance of the red/blue tones in the gate waveform can have an outsized effect on the Stark shift seen by the qubits participating in the gate. We can measure this Stark shift on a given gate by initializing the qubits in $|{+}{+}\rangle$ and then applying $[XX(\pi/4)XX(-\pi/4)]^N$, $N = 1, 2, 3\ldots$ Finally, we apply $R_y(\pi/2)$ pulses to the two qubits. If there has been no Stark shift, then the qubits should transfer to the $|11\rangle$ state. Applying this experiment while sweep-

ing the relative amplitude of the red/blue tones in the waveform yields sinusoidal curves for each qubit. When these qubits are spatially separated in the chain, they may each see a different red/blue tone imbalance. Thus we set the calibrated tone imbalance to be the mean of where each qubit has maximum transfer to $|1\rangle$. The residual error, which we call the differential Stark shift, is taken care of by automatically applying differential software $R_z(\pm\theta)$ gates paired with every application of the XX gate on that ion pair.

5.3: System Characterization

5.3.1: State Preparation And Measurement Error

We characterize the state preparation and measurement (SPAM) errors on a single ion, initially prepared in the $|0\rangle$ state using optical pumping. We can apply a following SK1 [55] π-rotation to prepare $|1\rangle$. To measure the qubit state, 369 nm light that is resonant with the $|^2S_{1/2}, F = 1\rangle \leftrightarrow |^2P_{1/2}, F = 0\rangle$ transition is directed onto the ion and the scattered photons are coupled into a multi-mode fiber bundle and detected on an array of pig-tailed PMTs. We determine that the ion is bright (dark) when we detect > 1 (≤ 1) photons within a 100 μs window. The SPAM error for $|0\rangle$ (the dark state) is 0.22(2)%, and can be broken down into dark-to-bright excitation (off resonant 14.75 GHz $|^2S_{1/2}, F = 0\rangle \rightarrow |^2P_{1/2}, F = 1\rangle$), incomplete pumping (preparation error), and background dark counts. When the ion is prepared in $|1\rangle$ (the bright state), we measure a SPAM error of 0.71(4)%, which can be broken down into bright-to-dark excitation (off resonant 2.105 GHz $|^2S_{1/2}, F =$

$1\rangle \rightarrow |^2P_{1/2}, F = 1\rangle$), threshold error (overlap of bright/dark histograms), state preparation error (single qubit gate error). The average single-qubit SPAM error is then 0.46(2)%. Table 5.1 describes the SPAM error budget, derived either from separate measurements or by fitting Poisson curves to the histogram of photon count event frequency.

0-state and 1-state error	**Error**
SPAM error on dark ion $\equiv \|0\rangle$	0.22%
Dark to bright pumping	0.13%
Preparation error - incomplete pumping	0.02%
Background dark counts (measured with no ion qubit)	0.07%
SPAM error on bright ion $\equiv \|1\rangle$	0.71%
Bright to dark pumping	0.55%
Thresholding error	0.12%
Preparation error (1-qubit randomized benchmarking)	0.03%

Table 5.1: State preparation and measurement error budget for a single ion in our system.

Detection crosstalk is a source of correlated error. We can characterize this by spatially scanning a 3-ion chain over 3 PMTs (by applying a linear axial field) and simultaneously applying detection light. The result of this experiment is shown in Figure. 5.8. At 0 V/mm, we can see that all three ions are aligned to the PMTs, as expected. Shifting the ions off to the side so that only the final ion is aligned to to PMT3, we can read the nearest neighbor crosstalk as 0.9% on PMT2, and the next-nearest neighbor crosstalk as 0.2% on PMT1. Assuming a Poisson distribution with $\lambda = 9$ this corresponds to a 0.31% false alarm rate on the nearest neighbor and 0.02% false alarm rate on next-nearest neighbor. Note that a this error only occurs when the target is bright and the neighbor is dark. For bright-bright or dark-dark combinations there is no error, so in practice the average crosstalk detection error

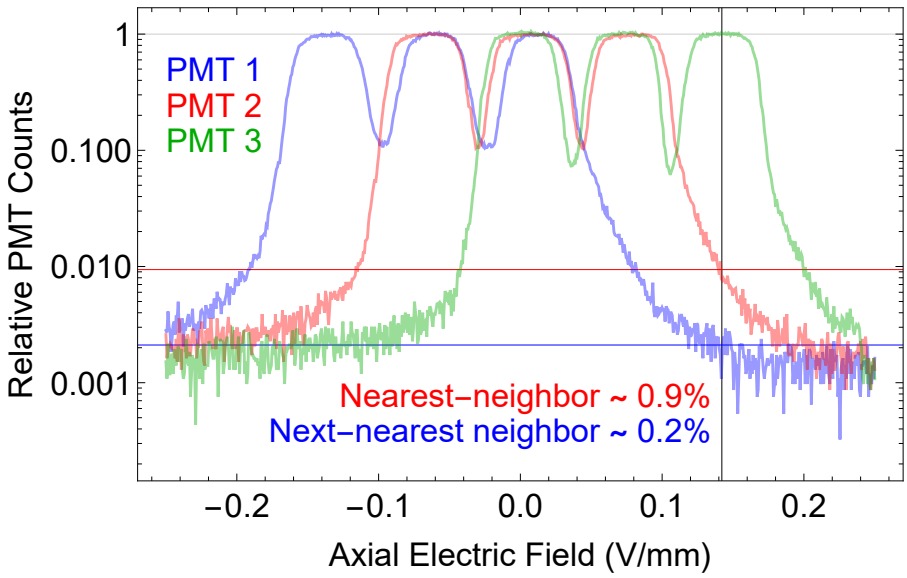

Figure 5.8: Log-scale plot of detection crosstalk in our system. Scattered photons from resonant detection light are measured on three channels of a multi-mode fiber bundle + PMT array. The three ions are scanned spatially across the detectors using a linear axial field.

is reduced by a factor of 2.

5.3.2: Rabi Crosstalk

Crosstalk errors are caused by unwanted Rabi drive spilling over onto idle qubits during active addressing of the target qubits. In one sense, crosstalk errors are a favorable type of error since they are unitary, which means they can in principle be reversed (e.g. through echoing protocols). On the other hand, crosstalk errors are correlated, which means they can be difficult to model and may break error-correction schemes that assume simple stochastic noise models. On a chain, Rabi crosstalk can be characterized by addressing one ion with a square carrier pulse and measuring flopping on the idle qubits. In post-processing, we normalize the Rabi frequency of the idle qubits by the Rabi frequency of the addressed ion. The results

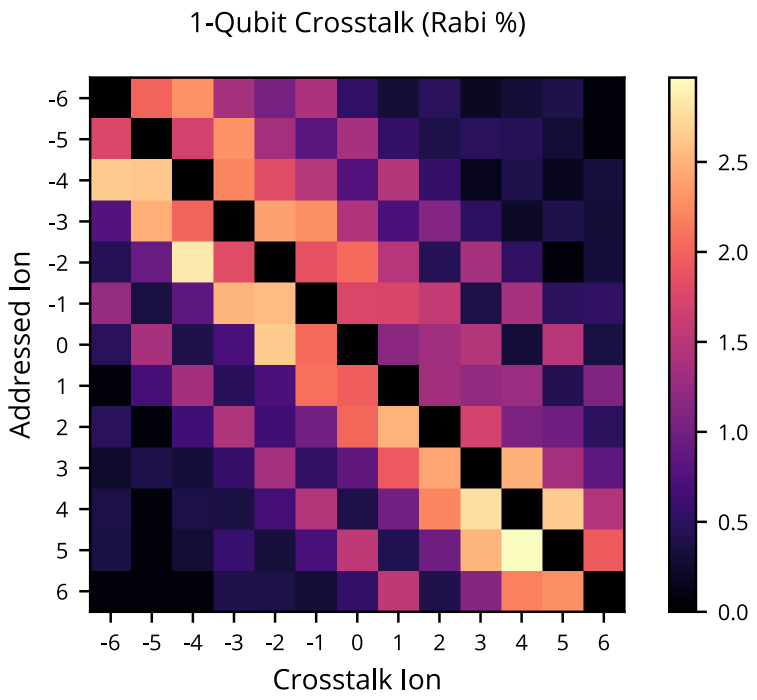

Figure 5.9: Rabi crosstalk measured in a chain of 15 ions by driving the addressed ion with a square carrier pulse and observing flopping on the idle qubits. Crosstalk Rabi frequency is normalized to a percentage of the addressed ion Rabi frequency.

of this experiment are shown in Figure 5.9.

As shown in Figure 5.10, there are three expected sources of crosstalk in our system - optical, electrical, and acoustical. For a perfect individual Gaussian beam with a waist of 0.85 μm, the expected electric field amplitude one ion spacing away (4.43 μm) from the beam center is many order of magnitudes below what is observed. Thus we would expect optical crosstalk to be dominated by beam aberrations and to primarily impact the nearest neighbors. There can also be electrical crosstalk due to capacitive pickup in the RF feed lines in the AOM cell itself that can drive main-lobe acoustic waves on unwanted channels. The design of the circuit boards in the AOM has odd numbered channels entering from one side, and even numbered channels entering from the other side. Thus we expect any electrical crosstalk to be

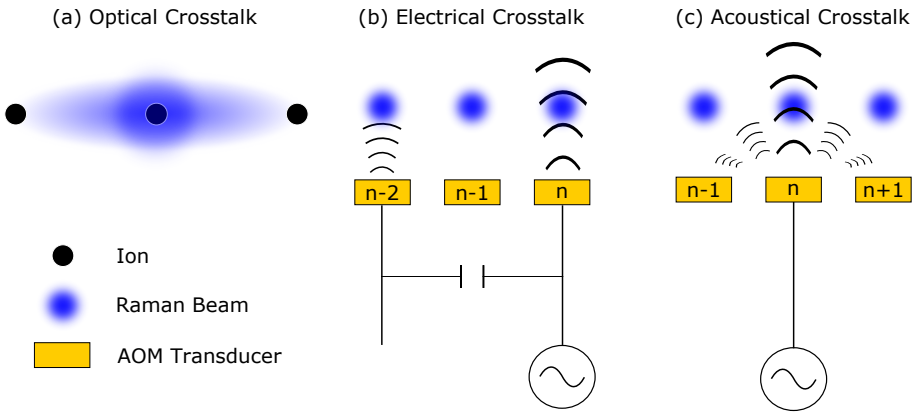

Figure 5.10: There are three different sources of Rabi crosstalk in our system. (a) Optical crosstalk due to beam aberrations causing electric field spill-over onto nearest neighbor ions. (b) Electrical crosstalk from RF capacitive pickup that drives main-lobe acoustic waves on next-nearest neighbor channels. (c) Acoustic crosstalk from diffractive side-lobes of the driven sound wave. Evidence suggests this is the dominant source of Rabi crosstalk in our system.

primarily next-nearest neighbor. Finally, there can be acoustic crosstalk within the AOM crystal from side-lobes caused by diffraction of the acoustic column. Due to interference effects of acoustic waves, we expect these effects to be highly dependent on the spatial location of the beams and/or ions and to exhibit fringe-like features. The dominant feature in our crosstalk matrix is nearest neighbor (± 1) and next-nearest neighbor (± 2) crosstalk in the range of 2-3%. However there is also a fringe at ± 5 with $\sim 1.5\%$ crosstalk which is difficult to explain with optical or electrical crosstalk.

To investigate this further, we drive the center AOM channel in a chain of 15, while rastering the ion chain through the beams using DC voltages on the chip trap. By observing the neighboring ions in this process, we can map out the spatial pattern of the crosstalk. The result of this experiment is shown in 5.11 at two different RF powers into the 32-ch AOM, corresponding to a 13× reduction in RF

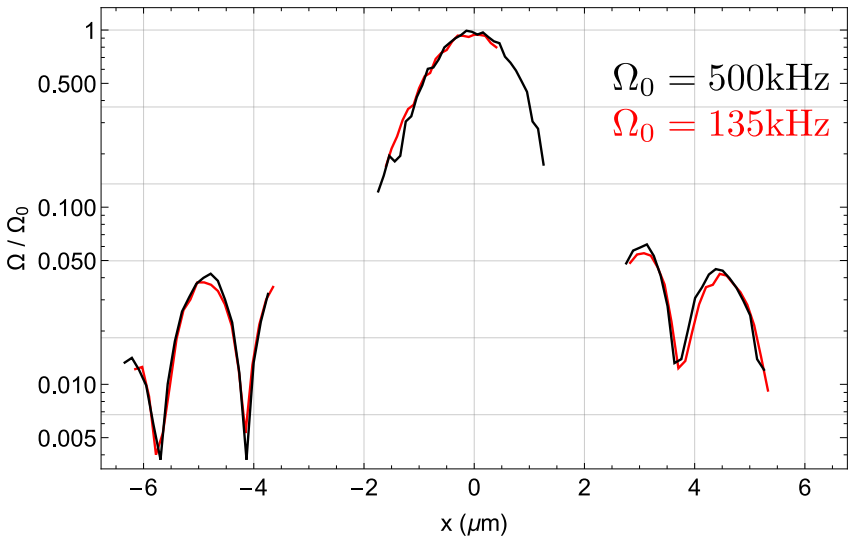

Figure 5.11: The spatial pattern of crosstalk onto the nearest-neighbor ions exhibits fringe patterns consistent with acoustic crosstalk. In this experiment, the center individual beam is driven while rastering an ion chain through the beams using DC voltages.

power between the black and red curves. There is apparently no dependence of crosstalk with RF or optical power. Notably, there is a distinctive fringe pattern that is consistent with the beams on the crosstalk ions being activated by weak acoustic waves - either through acoustic or electrical crosstalk, although nearest-neighbor here suggests acoustical. Through discussion with the AOM manufacturer, we identified an AOM-beam alignment issue likely responsible some amount of excess crosstalk. In particular, the input beams are not perfectly orthogonal to input crystal face. With perfect alignment, we expect the optimal crosstalk to drop to 1-2% on mostly nearest neighbor, which is in line with other work on similar hardware [77].

Practically, the impact of crosstalk in single-qubit gates can be suppressed quadratically using SK1 pulses, however two-qubit gates will still be sensitive. Assuming that the crosstalk originates from acoustic waves, we can attempt to drive

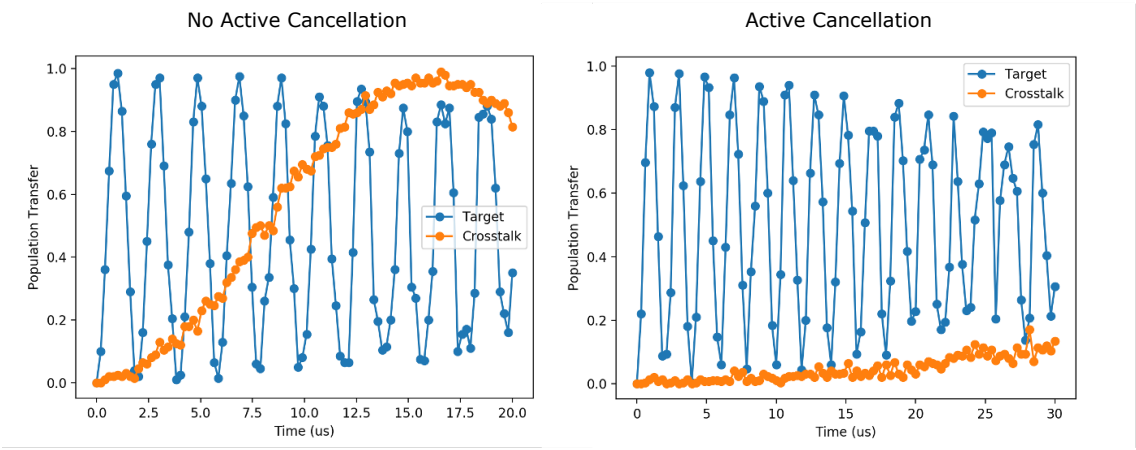

Figure 5.12: (left) Excitation on ion +1 (orange) due to beam applied to ion 0 (blue) without any cancellation. (right) Same but with a weak tone of correct phase/amplitude applied to the AOM channel corresponding to ion +1 (orange). We achieve 10× supression of the relative crosstalk Rabi frequency using active cancellation.

the idle crosstalk channel with a weak tone that cancels the acoustic wave at the location of the idle beam. The results of this experiment are shown in 5.12, demonstrating a 10× suppression in the relative crosstalk Rabi frequency. We note that both the phase and amplitude of the cancellation tone must be calibrated to null the crosstalk. In general, we find that the cancellation tone is not simply 180° out of phase with the drive tone, which is consistent with acoustical crosstalk and path length differences. Additionally, we point out that unlike electrical crosstalk, which can theoretically be canceled perfectly, acoustical crosstalk can never be completely nulled due to the difference in wavevectors between the cancellation wave and the crosstalk wave. The scalability of this cancellation scheme may ultimately be limited by the stability of the cancellation tone phases, which may drift on thermal time-scales.

5.3.3: Single Qubit Gate Performance

Measuring the fidelity of *one* single-qubit gate will likely always be limited by detection errors. One proposed method to overcome this limitation is to observe the scaling of error with long sequences of single-qubit gates. In particular, randomized benchmarking protocols use sequences of random Clifford gates to generate estimates of the computationally relevant errors without relying on accurate state preparation and measurement [89]. We use a sequence of up to 20 random Clifford gates, which are then decomposed into our native rotation gates and implemented using SK1 composite pulses (we are limited to only 20 Clifford gates by the details of our AWG system). Each random sequence is followed by its inverse in order to, in principle, echo out the gates completely and return the qubit to the initial $|0\rangle$ state. The degree to which the qubit does not return to the initial state quantifies the infidelity of the circuit.

The measured occupation of the $|0\rangle$ ground state as a function of the number of the applied Clifford gates is shown in Figure 5.13. This benchmarking procedure is performed on a single ion, as well as on an individual qubit in a chain of 15 ions, so as to detect any adverse affects arising from an increase in the system size. The fitted slope of the occupation of the $|0\rangle$ state as a function of the number of the applied Clifford gates indicates a per-Clifford error of $3.4(8) \times 10^{-4}$ on the 15-ion chain. We note that on average, each Clifford operation is composed of 1.875 native gates [90], corresponding to an error of $1.8(3) \times 10^{-4}$ per native gate. The offset in the fit is consistent with SPAM errors.

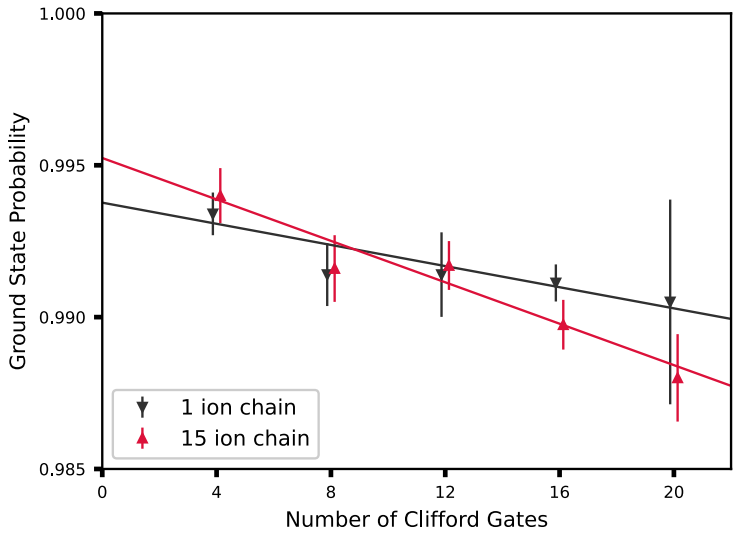

Figure 5.13: The probability to measure a single ion in the ground state after a variable number of Clifford gates in a randomized benchmarking sequence. The slope of the line indicates the per-Clifford fidelity, while the y-intercept indicates the SPAM error. Fit function for 1 ion chain is $0.9938(8) - N * 1.7(7) \times 10^{-4}$, and for 15 ion chain $0.995(1) - N * 3.4(8) \times 10^{-4}$. Error bars shown are the standard error of the mean.

5.3.4: Two Qubit Gate Performance

In general, to measure the fidelity of a multi-qubit state relative to some ideal state, one is required to gather information about the full density matrix, which scales poorly as the number of qubits grows. However, the density matrix of a Bell state created from a single XX gate is sparse and so we can take a short cut. Starting in $|00\rangle$ and applying a single XX, the two-qubit density matrix of this state will be,

$$\rho_{ideal} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \rho_{00} & 0 & 0 & \rho_{03} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \rho_{03}^{*} & 0 & 0 & \rho_{33} \end{pmatrix} \tag{5.4}$$

So then we only need to infer three values in order to calculate the fidelity of the state, ρ_{00}, ρ_{33}, and ρ_{03}. The error of the gate is then determined by observing the

deviation from these ideal values. Measuring this state, we can directly extract ρ_{00} and ρ_{33} since the probability to measure $|00\rangle/|11\rangle$ is just ρ_{00}/ρ_{33} (i.e., $P(|00\rangle) = \rho_{00}$ and $P(|11\rangle) = \rho_{33}$). Then we only need to infer the value of the off-diagonal term.

Consider the effect of a $\hat{R}(\pi/2, \pi/4)$ rotation on each qubit in the entangled state, with the rotation matrix given by $\hat{R}_A = \hat{R}(\pi/2, \pi/4) \otimes \hat{R}(\pi/2, \pi/4)$. Then the transformed density matrix is given by $\tilde{\rho}_{ideal} = \hat{R}_A \cdot \rho_{ideal} \cdot \hat{R}_A^\dagger$. Define the average parity Π of a state as the sum of the even parity populations minus the sum of odd parity populations, $\Pi = [P(|00\rangle) + P(|11\rangle)] - [P(|01\rangle) + P(|10\rangle)]$. When the qubits are perfectly correlated, $\Pi = 1$, and when the qubits are perfectly anti-correlated, $\Pi = -1$. Applying this observable to the density matrix $\tilde{\rho}_{ideal}$, yields $\Pi = -2i\rho_{03}$, which gives us a direct measure of the off-diagonal term in the density matrix. So when $\rho_{03} = i/2$, as in the ideal state, then $\Pi = 1$. While, we have only considered the ideal state, it can be shown that the fidelity of any mixed state relative to this ideal state can measured with the same approach [91]. Thus the fidelity of the Bell state created from a single XX gate is given by,

$$\mathcal{F} = \frac{1}{2}\left(P(|00\rangle) + P(|11\rangle) + \Pi\right) \tag{5.5}$$

In the example above, we knew ahead of time which phase to apply to the "analysis pulses" in order to maximize Π, but in practice we need to scan the phase $\hat{R}_A(\phi) = \hat{R}(\pi/2, \phi) \otimes \hat{R}(\pi/2, \phi)$. This results in a "parity fringe" curve like that shown in Figure 5.14, where Π is taken to be the amplitude of the sine curve. This exact calculation of fidelity only requires two experiments - one without analysis

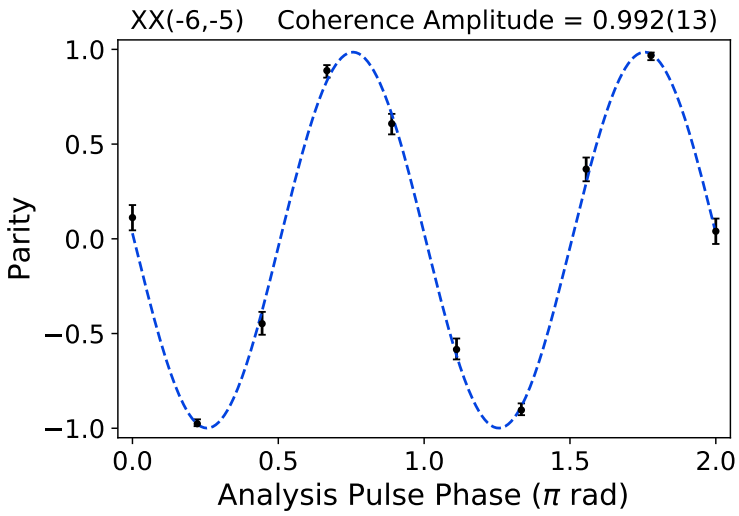

Figure 5.14: The parity fringe after a single XX gate on ions $(-6, -5)$ in a 15-ion chain. The parity fringe amplitude (0.992), is one of two numbers required to calculate the fidelity of the gate.

pulses to measure the populations, and one with analysis pulses to measure the coherence.

Measuring the gate fidelity of a single two-qubit gate also suffers from SPAM limitations. Fortunately, the fidelity measurement presented above works for any odd number of XX gates (an even number of XX gates generates no entanglement). We can take a similar approach to the single qubit gate characterization by repeating successive XX gates and observe the decay in fidelity versus number of gates. However, simply repeating XX gates can build up a coherent error that adds quadratically and could over-estimate the error per gate. For example, we anticipate the dominant error in the $XX(\pi/4)$ gate will be an over or under-rotation error by a small angle ϵ resulting in $XX(\pi/4 + \epsilon)$. A sequence of N successive applications of the gate will then will result in an accumulated angle error of $N\epsilon$. We take this situation to be the "worst-case" scenario.

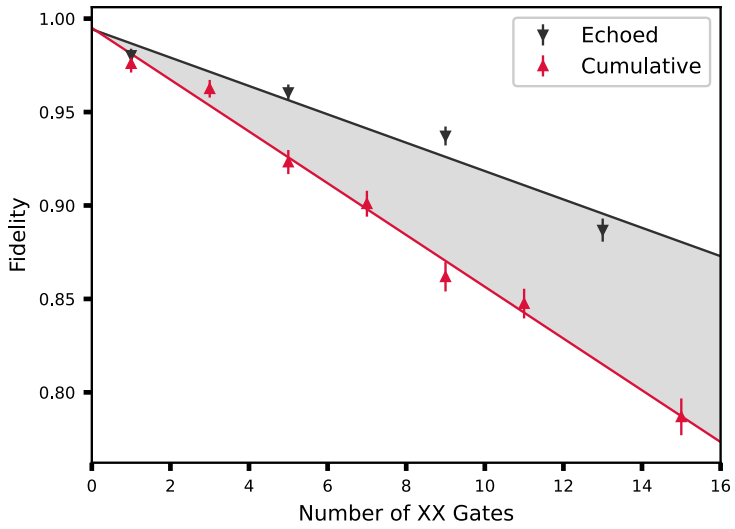

Figure 5.15: Fidelity estimate for a gate on ions $(-6, -5)$ in a 15-ion chain. The cumulative sequence (red) $XX(\pi/4)XX(\pi/4)...$ coherently builds errors, and the echoed sequence (black) $XX(\pi/4)XX(-\pi/4)...$ coherently cancels errors The true gate fidelity lies somewhere in the shaded region. The slope of the two lines determines the error per gate, or the estimated fidelity to be in the range of $98.5-99.3\%$.

On the other hand, if the phase of the gate is flipped by π with each successive application, $XX(\pi/4+\epsilon)XX(-\pi/4-\epsilon)...$, then this over/under rotation error is suppressed to the extent that it is stable between applications. We also note this echo sequence will also suppress other forms of coherent errors, such as gate crosstalk. We take the echoed gate sequence to be the "best-case" scenario. Within a circuit, we expect the true fidelity of a single XX gate to fall between these two extremes. As we increase the number of XX gates in the sequence, the slope of the fidelity gives the error per gate. In Figure 5.15, we show the results of these two sequences for a gate between ions $(-6, -5)$ in the chain of 15 ions. The estimated fidelity for this gate is then bounded within $98.5 - 99.3\%$.

Examining the cumulative XX gate experiment can yield insight into the error

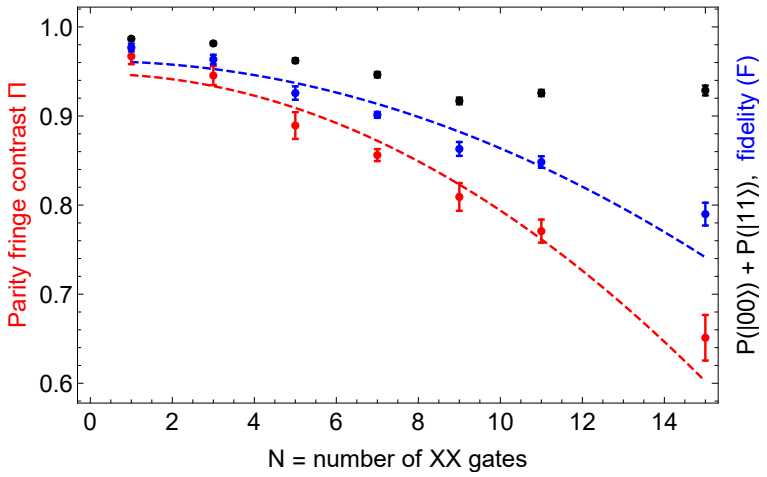

Figure 5.16: Error analysis for repeated XX gates on ions (-6,-5) in a 15-ion chain. The black points indicate $P(|00\rangle) + P(|11\rangle)$, which is a good metric for motional gate closure. The red points indicate the parity fringe coherence, which is a good metric for amplitude noise. The points are fit to $\Pi = 1 - (0.053 - (0.055N)^2/2)$. The blue points indicate the calculated fidelity from Eq. 5.5 with a fit to $F = 1 - (0.038 - (0.044N)^2/2)$

mechanisms that limit gate performance. As discussed in Section 3.3.1, there are two conditions to achieve a high-fidelity MS interaction - motional closure (i.e., phase space is closed and motion is disentangled from spin) and geometric phase enclosed (i.e., that gate is maximally entangling with $\chi = \pi/4$). Failure to close motion at the end of the gate results in residual entanglement of spin and motion. This can indicate an error in gate design or motional mode calibration error. Since the MS-interaction is a two-body interaction, it acts on the even parity subspace $|00/11\rangle$, assuming the qubits start in $|00\rangle$. Thus, leakage into the odd parity subspace $|01/10\rangle$ indicates residual entanglement with motion. Therefore the metric, $1 - P(|00\rangle) - P(|11\rangle)$, is useful for quantifying motional closure. This metric is shown after N two-qubit gates in Figure 5.16.

On the other hand, the off-diagonal term, quantified by Π, provides informa-

tion about the geometric phase enclosed by the gate. A decrease in Π indicates either a systematic miscalibration of the gate amplitude, or worse, amplitude noise that results in decoherence between $|00\rangle$ and $|11\rangle$. This metric is also sensitive to phase noise during the gate, although we anticipate amplitude noise to be the dominating contributor. The parity fringe coherence as a function of N is shown in Figure 5.16. The gate fidelity, which depends on these two numbers, is also shown in Figure 5.16. Clearly, the gate fidelity is limited by the parity fringe coherence, rather than the even parity subspace error. This indicates that amplitude noise appears to be the limiting factor, as opposed to the gate design. Amplitude noise can either appear from pointing noise of the focused individual Raman beams, or from movement of the ions with respect to fixed beams (see Section 6.1). While there are certainly other mechanisms that cause errors in two-qubit gates (spontaneous emission, motional coherence, stark shift noise, Rabi frequency imbalance, etc.), the effect of these errors are predicted to be significantly lower than the decrease that we observe. A more detailed discussion of the XX error budget can be found in other works [65, 92, 93].

In the future, a more standardized approach to fidelity estimation such as randomized benchmarking or gate set tomography [87, 89] would be preferred. However, these characterization protocols typically require thousands of relatively deep circuits to generate an accurate estimate, which are expensive to run in this system due to the particular implementation of the RF control system. The above characterization is particular for a low power gate. As the required power in a gate increases, we empirically observe a corresponding decrease in fidelity. We attribute

this to cross-talk, which we verify by looking at leakage into the odd parity subspace conditioned on a spin flip in an idle qubit.

In general, ion chains feature all-to-all two-qubit gate connectivity; however, some gates require more optical power than others to achieve maximal entanglement. These differences in power requirements can be understood by examining the mode participation symmetries in the chain. For example, ion 0, the center ion, requires high power in nearly all of its gates because it only participates in the even spatial modes (i.e., $b_{0,2n} = 0$, $n = 1, 2, ..., 7$ where $b_{i,1}$ is mode-participation factor of the highest-frequency in-phase radial mode for ion i). So on average, for a fixed gate frequency, the modes that drive entanglement are further detuned from the gate. We note that this is unique to our choice of amplitude modulated (AM) gates with a fixed frequency; phase/frequency-modulated (PM/FM) gates or multi-tone gates may have different chain symmetry considerations.

In 5.17, we present the power requirements for the gates in our system. Each gate is optimized according to Section 3.3.4. Once the optimal gate detuning is fixed for each gate, we calculate the root-mean-square (RMS) Rabi frequency (Ω_{rms}) of the AM waveform for *each* red/blue sideband when brought into resonance with the carrier transition. In our system, we use equal Rabi frequencies to drive both ions i, j in the gate ($\Omega_{i,rms} = \Omega_{j,rms}$), although this need not be case. The Lamb-Dicke factor ($\eta \approx 0.08$) converts carrier Rabi frequency to sideband frequency and this factor is normalized by the gate duration ($\tau_{gate} = 225\ \mu$s).

Using Fig. 5.17 as a cost matrix, we optimize the ion $\rightarrow$ qubit mapping so that the required gates for a given circuit minimize the total cost. In general, we

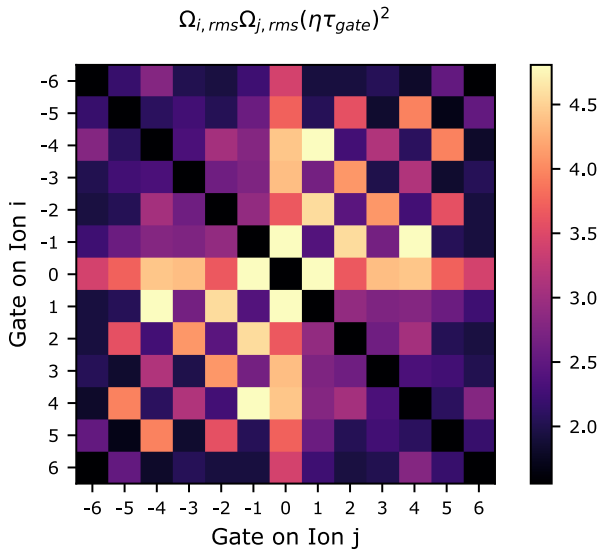

Figure 5.17: For an XX gate on ions i, j, the (RMS) Rabi frequency (Ω_{rms}) of the AM waveform for *each* red/blue sideband when brought into resonance with the carrier transition is normalized by the Lamb-Dicke factor ($\eta \approx 0.08$) and the gate duration ($\tau_{gate} = 225\mu$s). Gate power is used as a proxy for crosstalk, and the ion→qubit mapping is chosen to minimize this cost matrix.

observe that each half of the chain has strong coupling to itself, and the two halves of the chain couple well to each other as long as symmetry of the chain is obeyed (e.g., gates where the ions are with both odd or both even integer offsets from the center of the chain couple well, but mixed even and odd integer offsets do not).

5.3.5: Anomalous Heating Rates

The ions are observed to heat up while left along in the dark. This mechanism is called anomalous heating and is typically attributed to surface physics on the trap electrodes that causes electric field noise at the ions. This heating can cause detrimental effects within a single XX gate or over the course of a circuit. It is important that this heating rate be driven as low as possible through careful design and cleaning of the trap.

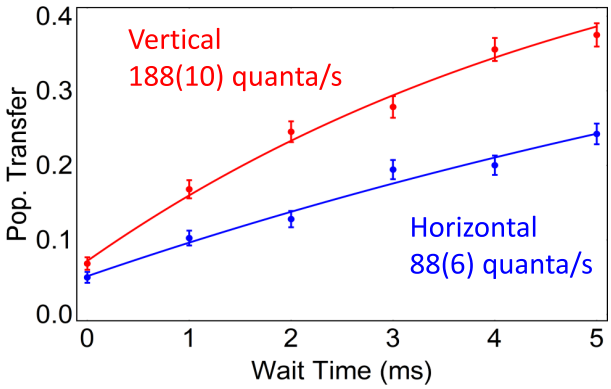

Figure 5.18: Heating rate measurement for the HOA chip trap. Population transferred from a fixed RSB pulse on a single ion with radial modes oriented parallel (horizontal) and perpendicular (vertical) to the chip trap surface. Scanning the wait time before the RSB pulse allows us to deduce the average phonon number as a function of time (i.e., the heating rate).

To characterize the anomalous heating rates in the HOA chip trap, we trap a single ion and orient the principal axes of the trap parallel and perpendicular to the trap surface. We first characterize the carrier Rabi frequency Ω_0. With the known Lambe-Dicke parameter η, the average phonon population can be deduced by applying a fixed duration RSB pulse on the radial mode and observing the population transferred. The result of this experiment is shown in Figure 5.18. We observe a heating rate to 188(10) quanta/s for the perpendicular (vertical) radial mode and 88(6) quanta/s for the parallel (horizontal) radial mode. This factor of two difference is consistent with anomalous heating caused by fluctuating patch potentials on the trap electrodes [94].

Chapter 6: Scaling to Long Ion Chains

In the previous chapter, we reported on performance benchmarks for a 15 ion chain. There are obvious benefits to trying to increase this chain length further to increase the computational power of the quantum computer while maintaining simultaneous all-to-all connectivity. While ion trap architectures with multiple chains in a QCCD-style system [16, 19] will ultimately be required, especially to perform mid-circuit measurement, there is significant overhead to the shuttling operations required to achieve effective all-to-all connectivity across the multiple chains. Thus, there are advantages to trying to increase the base chain length as long as possible, even if that system contains multiple chains [95]. The natural question becomes "what ultimately limits the length of the chain?" There are obvious engineering and control challenges to increasing the length of the chain, including voltage calibrations, two-qubit gate design, and RF control, however none of these appear fundamental. Following published work [96], in Section 6.1 we identify a challenge at the fundamental physics level that poses a threat to the scalability of long chains, and then in Section 6.2 we propose a mitigating solution and provide proof-of-concept demonstration of its capability.

6.1: Axial Heating

The fidelity of entangling quantum gates between trapped ion qubits relies on the control of normal modes of motion. It is therefore common to use high-frequency radial modes to mediate the entanglement [97], since they are more easily laser-cooled to the ground state and are less susceptible to heating from electric field noise [94, 98]. However, heating of spectator modes can also degrade quantum gate fidelity [99]. When addressing individual ions in long chains, the motion of the ions in the weakly-confined axial direction in particular can spoil the coupling of the ions to tightly-focused individual-addressing laser beams.

Consider the effect of axial ($\hat{x}$) motion of a chain of trapped ions, each of mass m, with axial normal mode frequencies ω_k. Each ion is addressed by an array of focused laser beams that drive Rabi oscillations between two qubit states ($|0\rangle$ and $|1\rangle$). The beams are directed perpendicular to $\hat{x}$ so that, throughout their axial motion, the ions experience fixed phases of the Rabi drive. The instantaneous qubit Rabi frequency Ω_i of the ion i is proportional to the electric field amplitude of the laser beam at the position of this ion [20, 100]. This is true for both direct optical qubit transitions [100] or two-beam optical Raman transitions where one of the beams has a uniform intensity profile, as used here. For small axial deviations of the i-th ion about its equilibrium position $x = x_0$, the Rabi frequency can be Taylor expanded to second order as,

$$\Omega_i(x) = \Omega_{i,0} + \Omega_i'(x_0)(x - x_0) + \frac{1}{2}\Omega_i''(x_0)(x - x_0)^2 \tag{6.1}$$

with $\Omega_0 = \Omega(x_0)$. The linear term $\Omega_i'(x_0)$ in Eq. 6.1 is zero if the beam center is aligned with x_0. However if the ion is offset from the center of the beam, either due to misalignment or non-perfect equispacing of the ions (see Section 2.1.3), then this term results in a spin-dependent axial force proportional to the gradient of the Rabi frequency along the axial direction. We will discuss this effect in Section 6.1.2, but for now we will assume the beam is centered on x_0 (i.e., $\Omega_i'(x_0) = 0$). When $\Omega_i \ll \omega_k$ (as in the case of two qubit gates), the ion will experience a time-averaged Rabi frequency

$$\begin{aligned} \overline{\Omega}_i &= \Omega_{i,0} + \Omega_{i,0}'' \frac{A^2}{2} \int_0^{2\pi} \frac{1}{2\pi} \cos^2\phi \; d\phi \\ &= \Omega_{i,0} + \frac{1}{2}\Omega_{i,0}'' \sum_{k=1}^{N} b_{ik}^2 \frac{E_k}{m\omega_k^2} \end{aligned} \tag{6.2}$$

where $\Omega_{i,0}$ is the Rabi frequency at the equilibrium position of ion i, $\Omega_{i,0}''$ is its curvature along the $\hat{x}$-axis, and A is the amplitude of motion given by $E = m\omega^2 A^2/2$ for a single motional mode. For many motional modes, we need to include the sum in Eq. 6.2 representing the mean-squared displacement of the i-th ion from its equilibrium, with b_{ik} the participation of this ion in axial mode k. $E_k = \hbar\omega_k(n_k+1/2)$ is the energy of excitation of mode k, with n_k the number of motional quanta in the same mode. Here, E_k is assumed to be constant during a Rabi oscillation or single gate operation.

Since cooling and heating are incoherent processes, assume that the energies E_k follow a thermal Boltzmann distribution at temperatures T_k, and write $k_B T_k = \hbar\omega_k\bar{n}_k$ with $\bar{n}_k \gg 1$ the average axial vibrational occupancy number of mode k.

The spread in the axial mode energy results in a spread in the average value of the Rabi frequency during the pulse, which results in decoherence. Unlike the ion position, which, for an ion in a thermal state of motion, is Gaussian-distributed, in the same state, the ion's energy follows an exponential distribution. This results in much stronger relative fluctuations in $\overline{\Omega}$ compared to a process whose average Rabi frequency depends linearly on the ion's position.

Starting in $|0\rangle$ and driving the carrier transition on the ith qubit, the probability to find the state in $|1\rangle$ can be evaluated by performing the thermal average, denoted by $\langle\cdot\rangle$, over all the modes k as

$$p_{|1\rangle}(t) = \left\langle \sin^2\left(\frac{\overline{\Omega}_i t}{2}\right)\right\rangle = \frac{1}{2} - \frac{1}{2}\mathrm{Re}\, e^{i\Omega_{i,0}t}\prod_{k=1}^{N}\sum_{n_k=0}^{\infty} p_k(n_k)\exp\left(ib_{ik}^2\Omega_{i,0}''t\frac{E_k}{2m\omega_k}\right), \tag{6.3}$$

where the occupation numbers $p_k(n_k) = e^{\hbar n_k\omega_k/(k_BT_k)}/(1-e^{-\hbar\omega_k/(k_BT_k)})$ correspond to Boltzmann distributions at the mode temperatures T_k. Performing the sums over n_k we then obtain

$$p_{|1\rangle}(t) = \frac{1}{2} - \frac{1}{2}\mathrm{Re}\, e^{i\Omega_{i,0}t}\prod_{k=1}^{N}\frac{\left(1-e^{-\hbar\omega_k/k_BT}\right)\exp\left(\frac{i\Omega_i''t}{2}\frac{\hbar b_{ik}^2}{2m\omega_k}\right)}{1-\exp\left\{-\frac{\hbar\omega_k}{k_BT}\left(1+i\Omega_{i,0}\theta_{ik}t\right)\right\}}, \tag{6.4}$$

where the decay parameters θ_{ik} are defined below in Equation 6.7.

Since the mean occupation numbers $\bar{n}_k \approx k_BT_k/(\hbar\omega_k)$ of the modes that contribute to the gate error are much larger than one. Expanding Eq. 6.4 to leading

order in $\hbar\omega_k/k_B T_k \sim 1/n_k$, we obtain

$$p_{|1\rangle}(t) = \frac{1}{2} - \frac{1}{2}\mathrm{Re}\, e^{i\Omega_{i,0}t} \prod_{k=1}^{N} \frac{1}{1 + i\Omega_{i,0}\theta_{ik}t} \exp\left(\frac{i\Omega_i'' t}{2} \frac{\hbar b_{ik}^2}{2m\omega_k}\right). \tag{6.5}$$

The last phase term describes the change in the average Rabi frequency due to the zero-point motion of the ions, with no additional decoherence. In this regime, it can be neglected, yielding,

$$p_{|1\rangle}(t) = \frac{1 - C\cos(\Omega_{i,0}t + \phi)}{2}, \tag{6.6}$$

The Rabi oscillations exhibit a phase advance of $\phi = \sum_k \arctan(\Omega_{i,0}\theta_{ik}t)$ and a loss in contrast by a factor of $C = \prod_k (1 + \theta_{ik}^2\Omega_{i,0}^2 t^2)^{-1/2}$. Here, the decay parameter for ion i due to mode k is defined as

$$\theta_{ik} = -\frac{k_B T_k b_{ik}^2}{2m\omega_k^2} \frac{\Omega_{i,0}''}{\Omega_{i,0}} = -b_{ik}^2 \xi_k^2 \frac{\Omega_{i,0}''}{\Omega_{i,0}} \bar{n}_k, \tag{6.7}$$

and characterizes the decoherence per Rabi cycle of ion i due to the motion in mode k. Here, $\xi_k = \sqrt{\hbar/(2m\omega_k)}$ is the zero-point spatial spread of the ion motion.

To probe the decoherence caused by the thermal axial motion of the qubit, we confine a single ion with varying static confinement strength along the axial $\hat{x}$ direction. Following Raman sideband cooling [101] of the radial motion of the ion to an average occupation number $\bar{n}$ of less than 0.15 quanta, we optically pump the ion into $|0\rangle$. After a 5-ms delay, we drive carrier Rabi oscillations of the ion, obtaining the data shown in Fig. 6.1. We observe that, when the axial frequency of the ion

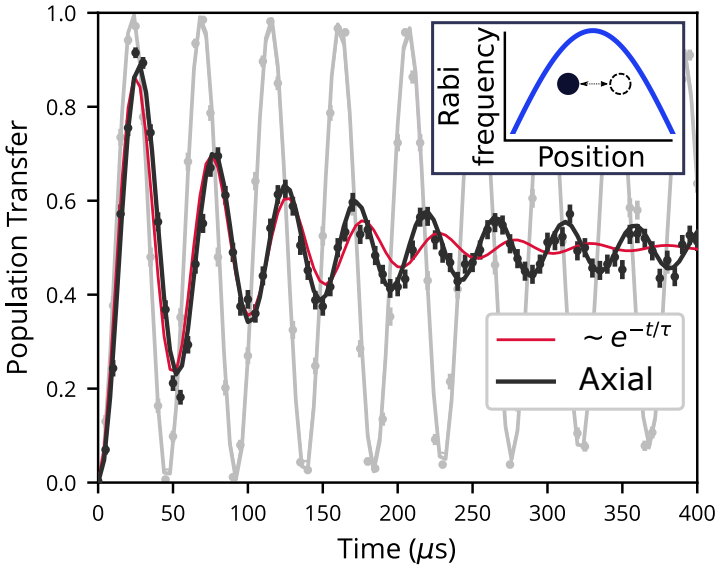

Figure 6.1: Rabi oscillations on a single ion with axial frequency 140 kHz (black) and 710 kHz (gray). The black and gray lines correspond to fits to Eq. 6.6. The red line corresponds to a fit to a two-state model with phase damping. Inset: illustration of the axial motion of the ion in the tightly-focused Raman beam.

is decreased from 700 kHz to 140 kHz, the Rabi oscillations exhibit a sharp decay. The observed oscillations do not exhibit an exponential decay with a constant phase shift, as would be expected from pure phase damping, but instead agree with the model from Eqs. 6.6-6.7.

We investigate the spatial dependence of the decoherence by applying static trap voltages to move the ion along $\hat{x}$. First, we map out the spatial profile of our tightly-focused Raman beam by tightly confining the ion and then driving it with a carrier Raman pulse of duration $\tau = 0.5$ μs. We use the fraction of the ion population transferred to $|1\rangle$ to determine the Rabi angle $\Omega\tau$, as shown in Fig. 6.2(a). A Gaussian fit to the obtained data ($\Omega\tau \sim e^{-x^2/w^2}$) yields the $1/e^2$ intensity-radius of the tightly focused beam $w = 870(25)$-nm. On one shoulder of the tightly focused beam, we observe a deviation from the Gaussian shape, which we ascribe to the

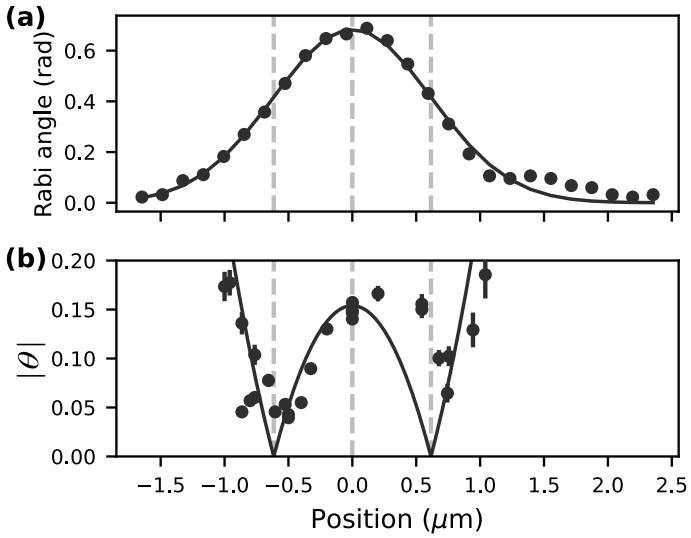

Figure 6.2: (a) The angle of the Rabi oscillation of a single ion depending on the position of the ion in a trap with tight axial confinement ($\omega_0 = 2\pi\times630$ kHz). (b) The absolute value of the decay parameter obtained by fit of the Rabi oscillation to the model from Eq. 6.6 for a weakly axially confined ion ($\omega_0 = 2\pi\times140$ kHz) depending on the position of the ion. The solid line indicates a prediction based on the Gaussian fit from (a) and an axial ion temperature corresponding to $\bar{n} = 280$ quanta.

mode profile of the laser.

Next, we relax the single ion's axial confinement and perform a carrier Rabi oscillation experiment, as in Fig. 6.1, at each set position of the ion. We fit the resulting data to the model from Eq. 6.6 to extract the decay parameter θ, with the results shown in Fig. 6.2(b). We compare the obtained $\theta(x)$ data to the prediction $\theta = 2(\xi_1/w)^2\,(1-2x^2/w^2)\,\bar{n}_1$, which assumes a Gaussian shape of the tightly-focused beam, with the ion's axial average thermal vibrational number of $\bar{n}_1 = 280$ at the expected Doppler cooling temperature.

We observe good qualitative agreement, with the laser mode shape likely responsible for the discrepancy at positive $\hat{x}$ values. The shape in Fig. 6.2(b) is in striking contrast to decoherence caused by beam fluctuations, which increases on

the sides of the focused beam. Rather the measured decoherence parameter reaches its minimum values near the inflection points of the Gaussian curve. The measurements from Figs. 6.1 and 6.2 demonstrate that both the spatial and the temporal behavior of the observed Rabi oscillations of an ion under weak axial confinement are well described by the model of Eqs. 6.6-6.7.

We now consider the effect of axial motion on a chain of ions. After sufficient time t_{w} following laser cooling, the temperature T_k of the axial mode k will be dominated by the work done by noisy background electric fields [94, 98]. If this field is uniform in space, its work on mode k will be proportional to $(\sum_i b_{i,k})^2$. Since the decoherence due to the decay parameter scales as $1/\omega_k^2$, we expect the heating of the lowest-frequency "in-phase" axial mode ($m = 1$), to strongly dominate the gate error budget (recall for radial modes the in-phase mode has the largest ω_k, whereas the opposite is true for axial modes). In this case, the effects of the axial motion on the i-th ion are captured by the single decay parameter

$$\theta_i \equiv \theta_{i1} = b_{i1}^2 \left(\sum_j b_{j1} \right)^2 \theta(t_{\mathrm{w}}), \tag{6.8}$$

where $\theta(t_{\mathrm{w}})$ is the decay parameter of a single ion in a trap with axial frequency equal to ω_1, following Eq. 6.7. Note that for harmonic axial confinement where the in-phase mode is the center-of-mass mode ($b_{i1} = N^{-1/2}$), the decay parameter θ_i becomes equal to that for a single ion.

In the experiment, we use a combination of quadratic and quartic axial potentials (see Section 2.1.3) to prepare near-equispaced chains of 15 (25) ions with 4.4 μm

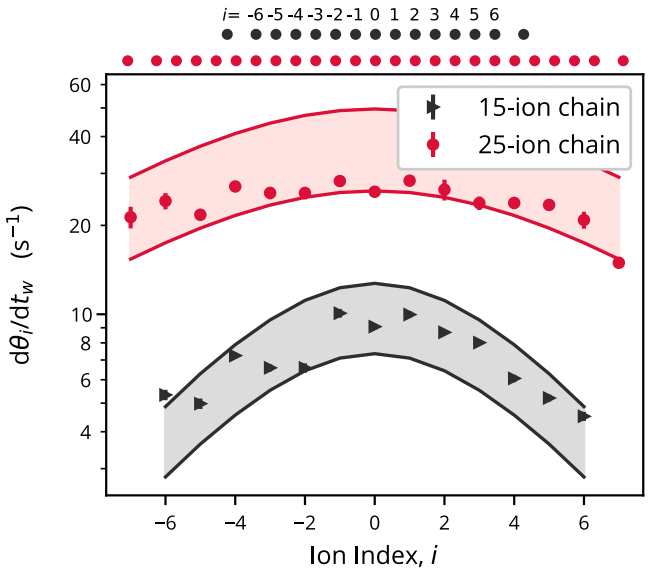

Figure 6.3: The rate of change of the decay parameter θ_i as function of the ion index (i) in a chain of near-equispaced 15 ions (black) and 25 ions (red). The error bars are statistical from fits of $\theta_i(t_w)$ to a linear increase with the wait time t_w. The solid lines correspond to predictions based on power-law spectral density of electric field noise with exponent $\alpha = [0.8, 1]$ and the independently measured heating rate of one ion at 3 MHz of $\dot{\bar{n}}_r = 88(6)$ quanta/s (see Section 5.3.5).

ion spacing, obtaining 193 (123) kHz as the lowest axial mode frequency. Following sideband cooling of the radial modes of the ion chain, we drive simultaneous Raman Rabi oscillations on the middle 13 (15) ions in the chain after a variable wait time t_{w} following the Doppler and sideband cooling. We fit the oscillations of each ion to the model from Eqs. 6.6-6.7 to determine the rate of change of its decay parameter, and we show the results of these measurements in Fig. 6.3. The observed variation of the decay parameters across the chains follows the factor b_{i1}^2 from Eq. 6.8, with the Rabi oscillations of the middle ions exhibiting increased decay due their higher participation in the lowest-frequency axial mode. This effect is colloquially called the "soft belly" of the chain and is not observed with a quadratic confinement term only (in which case b_{i1} is flat).

Electric-field noise in ion traps is empirically observed to follow a power-law with frequency $\omega^{-\alpha}$, with exponent α between 0 and 2 [94]. To check for consistency of our observations with this behavior, we use sideband spectroscopy of a single ion to measure the heating rate of $\dot{\bar{n}}_r = 88(6)$ quanta/s for a 3-MHz radial mode parallel to the trap surface (Section 5.3.5). In Fig. 6.3, we show the predictions based on Eqs. 6.7 and 6.8 and $\alpha = 1$. We also show predictions for $\alpha = 0.8$, which we deduced independently from $d\theta/dt_{\mathrm{w}}$ for a single ion as a function of the axial trap frequency. We observe good quantitative agreement with our data, suggesting that electric-field noise in our system is consistent with previously observed values.

The Rabi frequency $\Omega_{i,j}$ of the entangling dynamics between the $|00\rangle$ and $|11\rangle$ states of ions i and j during two-qubit entangling gates is proportional to the product of the single-qubit Rabi frequencies on the two addressed ions. If the two ions are centered on the maxima of their respective individual-addressing beams, the joint decay parameter corresponding to axial mode k is $\theta_{ik} + \theta_{jk}$. The fidelity of the obtained two-qubit state is bounded from above by $F_{ij} = \mathrm{Tr}\left(\hat{\rho}\left(t\right)\hat{\rho}_{\chi}\right) = \left\langle \cos^2\left(\bar{\Omega}_{i,j}t - \chi\right)\right\rangle$, where χ is the desired two-qubit gate angle, and $\bar{\Omega}_{i,j}$ is the mean value of $\Omega_{i,j}$ during the ions' axial motion. Performing the thermal average as in Eq. 6.6, we find the state fidelity bound after N_g successive fully-entangling gates ($\chi = N_g\pi/4$) is

$$F_{ij} = \frac{1}{2} + \frac{1}{2}\prod_{m}\frac{1}{\sqrt{1 + \left(N_g\pi/2\right)^2\left(\theta_{ik} + \theta_{jk}\right)^2}}. \tag{6.9}$$

To check this prediction, we apply linear amplitude modulated gates (see Section 3.3.3) that address ions $(-6, -5)$ in chains of 15 (25) ions that perform one

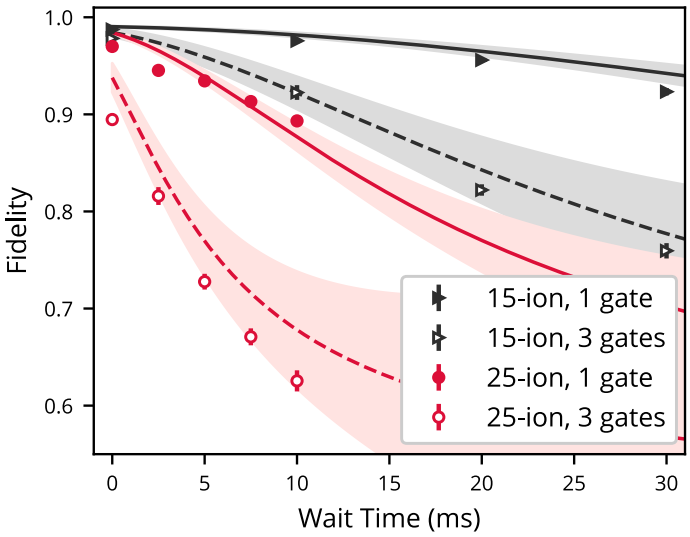

Figure 6.4: The fidelity of one (three) fully entangling gate(s) on two ions (indices i=−6 and j=−5 as in Fig. 6.3) in a chain of 15 and 25 ions as a function of the waiting time before the gate(s). The error bars denote 1σ uncertainty of the weighted average of several measurements of the fidelity. The shaded areas show the predictions based on Eq. 6.9, adjusted down by the error in state preparation and measurement in our ion chains (0.9%, Section 5.3.1). The shaded area in the theory predictions reflect the uncertainty in our determination of the decay parameters θ_i and θ_j.

or three successive 225 μs (500 μs)-long entangling gates between the target ions, after a variable wait time t_{w} following Doppler and sideband cooling. Successive gates are expected to magnify the decoherence effect. After applying the entangling gate(s), we measure the $|00\rangle, |11\rangle$ subspace population $p_{00} + p_{11}$. Separately, we apply additional $\pi/2$ pulses with variable phase to both ions and extract the parity fringe contrast C, to witness entanglement [91]. We independently determine θ_i and θ_j in our chains as a function of t_w by repeating the measurements from Fig. 6.3 and fitting the resulting decay parameters to a linear increase with the wait time t_w.

We compute the gate fidelities from our measurements as $F = (p_{00} + p_{11} + C)/2$ [91] and show these results in Fig. 6.4. With 15 ions, we achieve a baseline two qubit gate fidelity of 98.5% and with 25 ions we achieve a gate fidelity of 97.5% on

the two selected ions. We compare these results to the predictions based on Eq. 6.9, with only the lowest ($k = 1$) axial mode contributing. We observe good agreement between our measurements and the model of Eq. 6.9. In particular, in chains of 15 (25) ions, after t_w=10 (2.5) ms, corresponding to the time it takes to complete 40 (5) sequential individual entangling gates, our model explains most of the observed loss in gate fidelity.

In ion chains, the mode frequency of in-phase axial mode scales roughly inversely proportional to the ion number $\omega_1 \sim 1/N$ (see Section 6.1.1). Assuming electric field noise with exponent α and using Eqs. 6.7-6.8, we obtain $d\theta_i/dt_w \sim N^{2+\alpha}$ in such chains, implying that the entangling gate error scales as $t_w^2 N^{4+2\alpha}$. For $\alpha \approx 1$ as seen in many ion trap experiments [94], this results in a scaling of gate errors proportional to N^6, which is strongly unfavorable to the scalability of long chains without a mitigating strategy. In this data, the decrease in fidelity of a single gate after $t_{\mathrm{w}} = 10$ ms is only 1% in a 15-ion chain, compared to 7.5% on a 25-chain. High fidelity operations on chains than longer 15 ions could be achieved by significantly lowering the base heating rate, but will ultimately succumb to the same scaling limitation, assuming the other engineering challenges are met. For example, if $\dot{\bar{n}}$ were reduced to 10 qaunta/s at 3 MHz (relative to the 88 quanta/s demonstrated here), it would be one of the lowest heating rates ever reported to date in a surface trap [102]. Yet the chain length could only increase by $\sim \times 2$ (31 ions) and maintain similar performance to the 15-ion chain shown here.

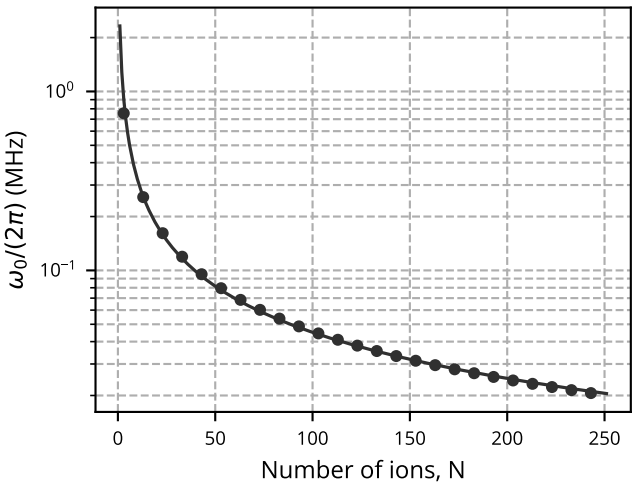

Figure 6.5: The lowest axial frequency ω_1 of a chain of $^{171}Yb^+$ ions as a function of the number of ions N in the potential given by Eq. 6.10 with $d = 4.4$ μm. The solid line corresponds to the fitted $N^{-0.856}$ power law.

6.1.1: Lowest Axial Frequency in Equispaced Ion Chains

In the limit of many ions ($N \to \infty$), an equispaced ion chain can be approximated by a continuous charge distribution with linear charge density e/d, where d is the ion spacing. The Coulomb potential of this charge is exactly countered by the applied trap potential

$$V(x) = \frac{e^2}{4\pi\epsilon_0 d} \ln \frac{(N/2)^2}{(N/2)^2 - (x/d)^2}, \tag{6.10}$$

which holds the ions in place. To model near-equispaced chains, we use $V(x)$ as the trap potential and numerically find the equilibrium ion positions $x_{i,0}$. We find that, when using the potential from Eq. 6.10, for all $N < 250$, the deviation of the equilibrium ion positions from that of an equal-spaced chain is at most 0.02 d.

For ion positions x_i near the equilibrium, the total energy can be written as

$\frac{e^2}{2d^3}\sum_{i,j=1}^{N}(x_i - x_{i,0})Q_{i,j}(x_j - x_{j,0})$, where

$$\begin{aligned} Q_{i,i} &= \frac{2(N/2)^2 + 2(x_{i,0}/d)^2}{((N/2)^2 - (x_{i,0}/d)^2)^2} + \sum_{j\neq i}\frac{2d^3}{|x_{i,0} - x_{j,0}|^3} \\ Q_{i\neq j} &= -\frac{2d^3}{|x_{i,0} - x_{j,0}|^3}. \end{aligned} \tag{6.11}$$

The axial mode frequencies are then found as $\omega_k = \omega_u\sqrt{\lambda_k}$, where λ_k are the eigenvalues of the matrix $Q_{i,j}$ and $\omega_u = \sqrt{e^2/(4\pi\epsilon_0 m d^3)}$ is the unit frequency. The calculated frequency ω_1 of the lowest-frequency axial mode is plotted as a function of the ion number, N, in Fig. 6.5. For chains of up to 250 ions, we observe good agreement with the $N^{-0.86}$ scaling that was predicted for chains in harmonic traps by Ref. [103].

6.1.2: Spin-Dependent Axial Coupling

In Section 6.1, we assumed that the beam was aligned to the equilibrium ion position x_0, and therefore $\Omega'(x_0) = 0$ in the Taylor expansion of Ω. If the beam is misaligned there can be undesired coupling to the axial motion [104]. Assume a beam with waist w is misaligned by δ_x to the ion position $\hat{x}$. Then,

$$\begin{aligned} \hat{H} &= \frac{\Omega_0}{2}e^{-(\hat{x}+\delta_x)^2/w^2}\hat{\sigma}_x \\ &= \frac{\Omega_0}{2}e^{(-\hat{x}^2 - 2\hat{x}\delta_x - \delta_x^2)/w^2}\hat{\sigma}_x \\ &\approx \frac{\Omega_0}{2}e^{-\delta_x^2/w^2}\hat{\sigma}_x\left(1 - \frac{2\delta_x\hat{x}}{w^2}\right) \\ &= \frac{\Omega_1}{2}\hat{\sigma}_x\left(1 - \frac{2\delta_x\hat{x}}{w^2}\right) \end{aligned} \tag{6.12}$$

Where $\Omega_1 = \Omega_0 e^{-\delta_x^2/w^2}$ is the derated Rabi frequency due to the misalignment. Writing the position operator as $\hat{x} = \xi_0(\hat{a} + \hat{a}^\dagger)$ with $\xi_0 = \sqrt{\hbar/2m\omega_1}$ and ω_1 the frequency of the in-phase axial mode, then,

$$\begin{aligned}\hat{H} &\approx \frac{\Omega_1}{2}\hat{\sigma}_x\left(1 - \frac{2\delta_x\xi_0}{w^2}(\hat{a} + \hat{a}^\dagger)\right) \\ &= \frac{\Omega_1}{2}\hat{\sigma}_x\left(1 - \frac{2\sqrt{2}\delta_x\xi_0}{w^2}\hat{X}\right) \\ &= \frac{\Omega_1}{2}\hat{\sigma}_x - \hat{\sigma}_x\hat{X}F\omega_1\end{aligned} \tag{6.13}$$

Where $\hat{X} = (\hat{a}+\hat{a}^\dagger)/\sqrt{2}$ and $F = \sqrt{2}\Omega_1\delta_x x_0/w^2\omega_1$ is the magnitude of the spin-dependent force on the in-phase axial mode. For a constant carrier Rabi frequency, this force will create the spin-dependent displacement operator,

$$\hat{D} = \exp\left[\hat{\sigma}_x\hat{a}^\dagger\alpha - \hat{\sigma}_x\hat{a}\alpha^*\right] \tag{6.14}$$

with $\alpha = F(e^{i\omega_1 t} - 1)/\sqrt{2}$. If the Rabi frequency is not constant, as in a pulse with a Gaussian amplitude envelope, the ion will follow a state-dependent trajectory in phase space. If the phase space is not closed at the end of the operation, then there will be residual spin-motion entanglement that looks like decoherence. However, if the force is ramped slowly enough relative to $1/\omega_1$, then the net α will be small and the ion will make excursions that approximately close. Unfortunately, sudden phase flips throughout a smooth amplitude profile, as in the case of shaped SK1 pulses, will cause a sudden change of force that may cause motion not to close. Pulse-shaping techniques like those described in Section 3.3.2, taking into account

the axial in-phase mode, would mitigate this effect.

Additionally, if single-qubit gates are run in parallel, with both ions misaligned to their respective beams, then this spin-dependent force will create entanglement between the two ion spins, just as in the case of the MS interaction. To mitigate this risk in our system, single-qubit gates are run sequentially.

6.2: Sympathetic Cooling

To address the challenge posed by axial heating, we propose to intersperse coherent operations with periodic sympathetic cooling [105] of axial modes via coolant ions that are distributed throughout the ion chain [19, 99, 106]. Since the gate error is proportional to t_w^2, in the limit of frequent cooling to the same initial axial temperature T_0, we expect sympathetic cooling to strongly suppress of the effects of axial heating on gate fidelity.

The choice of a sympathetic cooling ion will lend itself to different cooling schemes. There have been demonstrations of using $^{138}Ba^+$ ions as a sympathetic coolant [19] and the system described in this thesis contains Barium ovens. The D1 line in Barium is sufficiently detuned (493 nm) from the D1 line in Ytterbium (369 nm) that Doppler cooling can be performed on the sympathetic cooling Ba ions without decohering the data qubit stored in Yb. Since recoil from axial Doppler cooling will heat the radial modes, typically this is followed by resolved sideband cooling on the radial modes before performing further operations on the data qubits [19].

The mass ratio between $^{138}Ba^+$and $^{171}Yb^+$ is $r = 171/138 = 1.24$, which also

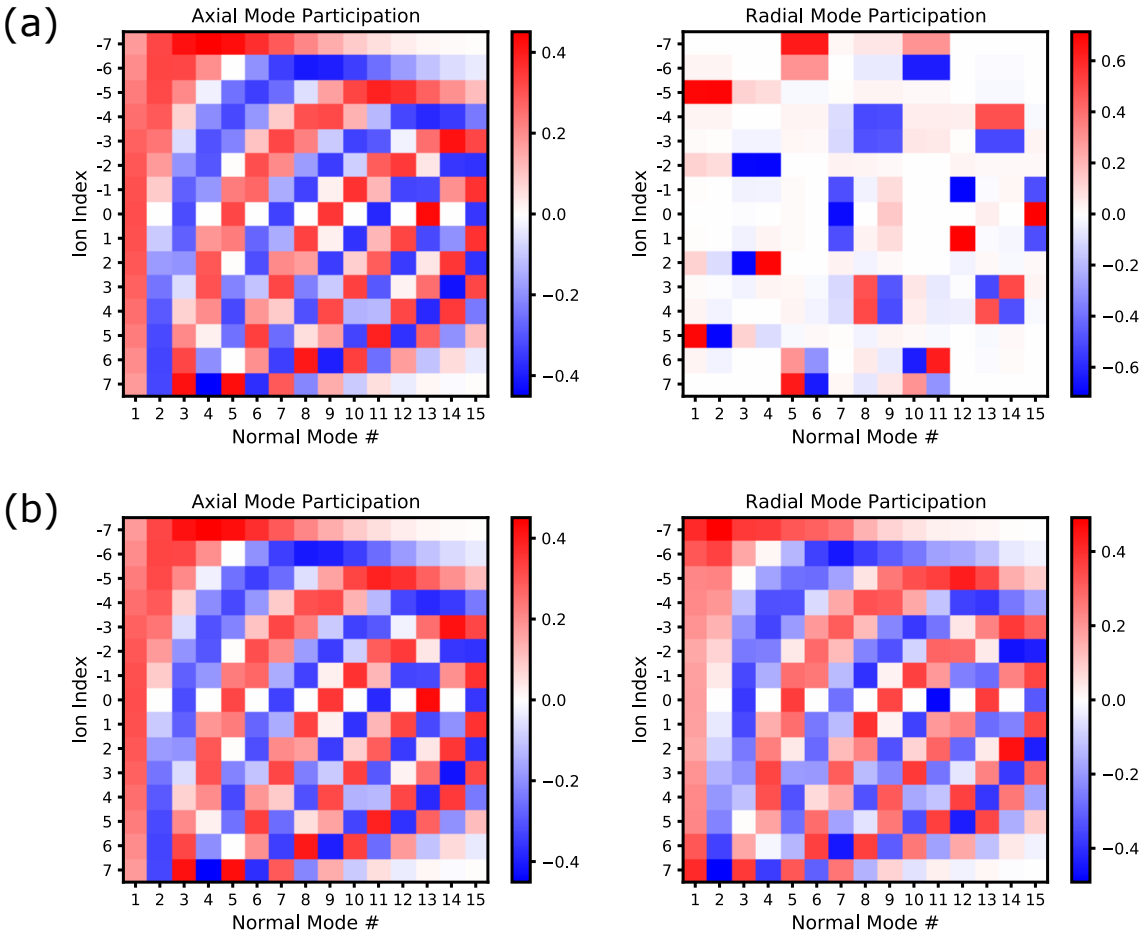

Figure 6.6: Choice of $^{138}Ba^{+}$vs $^{172}Yb^{+}$ for sympathetic cooling in a 15-ion chain. (a) $^{171}Yb^{+} \times 11$,$^{138}\,Ba^{+} \times 4$ with Barium ions located at ion index $i = \{-5, -2, 2, 5\}$. The axial mode participation matrix is unaffected by the mass difference, whereas the radial mode participation matrix is strongly decoupled, breaking the all-to-all connectivity of the chain. (b) $^{171}Yb^{+} \times 11$,$^{172}\,Yb^{+} \times 4$ with Yb172 ions located at ion index $i = \{-5, -2, 2, 5\}$. Both the axial and radial mode participation matrices are largely unaffected due to the near unity mass ratio. Compare to the radial mode participation matrix in a pure chain of $^{171}Yb^{+} \times 15$ in the same potential as Figure 2.4.

scales the radial secular frequency by the same amount. This difference in radial secular frequency causes radial modes on long chains to decouple into Ba-only modes and Yb-only modes. Not only does this make it difficult to cool the radial modes of Yb using Ba, but it also spoils the all-to-all connectivity of the chain. Recall that to efficiently entangle two ions via the MS interaction, they both must have high participation in *at least* one motional mode. As shown in Fig. 6.6(a), the Yb-only radial modes form groups of $3-4$ ions that couple well within the grouping, but not to other groupings. This effect is not observed in the axial mode spectrum because the axial potential is dominated by the Coulomb interaction, which depends on charge, rather than the RF psuedopotential, which depends on mass (and charge). While only one arrangement of the ions is shown here, this decoupling is common to all arrangements. Thus large mass ratios in a sympathetic cooling ion are not conducive to operating gates on the radial modes [107]. Using an even isotope of Yb, in particular $^{172}Yb^{+}$, as a sympathetic coolant leaves the radial mode spectrum almost unchanged Fig. 6.6(b).

With an even-isotope Yb, Doppler cooling on the D1 line would readily cause decoherence on the $^{171}Yb^{+}$ data qubits. Thus, our solution is to apply the narrow-line sideband cooling scheme from Ref. [108] (demonstrated on $^{171}Yb^{+}$, but applied to $^{172}Yb^{+}$ here). The ground $^{2}S_{1/2}$ manifold is coupled to the $^{2}D_{3/2}$ manifold via a quadrupole transition transition using a 435-nm laser beam. To obtain maximal coupling, this beam should be incident at 45° relative to the magnetic field and be linearly polarized in the plane defined by the magnetic field and its direction of propagation [109]. Two 435-nm tones separated by the appropriate Zeeman splitting

allow simultaneous addressing of both $\Delta m = 0$ transitions. Following the scheme from [108], the two tones are detuned to the red relative to the respective carrier transitions so as to cool the motional modes of the ion chain. State reset is provided by a 935-nm beam that couples the $^2D_{3/2}$ manifold to the $^3D[3/2]_{1/2}$ states, which decay to the ground state via emission of a 297-nm photon. We note that this cooling scheme would work with any even-isotope Yb, although we will focus on $^{172}Yb^+$ because it minimizes the mass ratio and is the only other isotope available in the 171-isotopically enriched source in this system ($\approx 2.5\%^{172}Yb^+$).

Crosstalk to $^{171}Yb^+$ in this cooling scheme can arise from off-resonant absorption of the 435-nm light, as well as from absorption of the 297-nm photons that are scattered by the $^{172}Yb^+$ ions. Both crosstalk processes depend on the isotope shifts of the relevant transitions. The shifts of the 935-nm transition are available in [110]. The shifts of the 297-nm transition for even Yb isotopes were measured by [111], while the shifts of the 435-nm transition for even isotopes were measured by [112]. A King plot of the modified 435-nm (297-nm) isotope shifts as a function of the modified 935-nm isotope shifts allow us to determine the shift of the centroid of the $^{171}Yb^+$ 435-nm (297-nm) line relative to $^{172}Yb^+$ as $\Delta_{435} = 2\pi \times 1.4$ GHz ($\Delta_{297} = -2\pi \times 1.0$ GHz). Using these shifts, we determine the detunings of the 435-nm, 935-nm, and 297-nm light that is involved in cooling $^{172}Yb^+$ from the transitions in $^{171}Yb^+$, and show these in Fig. 6.7.

Decoherence of qubits during the cooling of $^{172}Yb^+$ can arise via off-resonant scattering of 435-nm photons by the $^{171}Yb^+$ ions. The rate Γ_{435} for this process is bounded from above by $\Omega_{435}^2/(4(\Delta_{435}^{1\to 2})^2\tau_D)$, where Ω_{435} is the resonant 435-nm

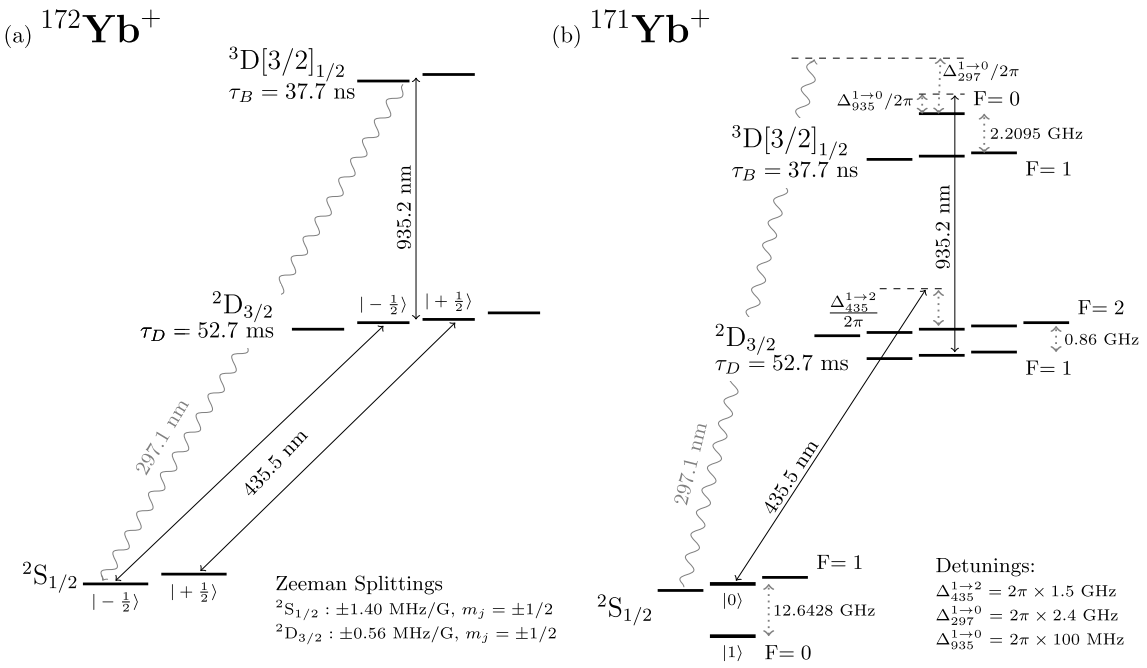

Figure 6.7: Diagram of the states and transitions involved in (a) the cooling scheme of ^{172}Yb$^+$ and (b) the crosstalk to the ^{171}Yb$^+$ qubit ions. Lifetime of the ^{2}D$_{3/2}$ state is cited from [113], and that of the ^{3}D[3/2]$_{1/2}$ state is from [114]. The hyperfine splittings of the ^{3}D[3/2]$_{1/2}$ and ^{2}D$_{3/2}$ states are from [40].

Rabi frequency and $\tau_D = 52.7$ ms is the lifetime of the ^{2}D$_{3/2}$ state [113]. In order to remain in the sideband regime and obtain low temperatures, Ω_{435} needs to be smaller than the motional frequency of the cooled modes. Using the radial mode frequency of 3 MHz, we find that $\Gamma_{435} < 2 \times 10^{-5}$ s^{-1}.

Further decoherence is induced by the scattering, by ^{171}Yb$^+$, of 297-nm photons via the off-resonant, two-step, $F = 1 \to F = 1 \to F = 0$ transition that is driven by the combination of the 435-nm and the 935-nm light used to cool ^{172}Yb$^+$. We model this process using three-level optical Bloch equations. Assuming 435-nm and 935-nm Rabi frequencies of $2\pi \times 3$ MHz, we find that the rate of scattering of 297-nm photons by a ^{171}Yb$^+$ ion via this process, and the resulting decoherence rate, are both smaller than 2×10^{-5} s^{-1}. Stronger crosstalk arises due to absorption, by

^{171}Yb$^+$, of the 297-nm photons that are scattered by ^{172}Yb$^+$. Given the near-unity branching ratio of the 297-nm transition, a single ^{172}Yb$^+$ ion at unity saturation will scatter 297-nm photons at the rate $\Gamma_{sc} = \Gamma/4$, where $\Gamma = 1/\tau_B$, with $\tau_B = 37.7$ ns the lifetime of the ^{3}D$[3/2]_{1/2}$ state [114]. These scattered photons will be absorbed by a ^{171}Yb$^+$ ion at distance d at the maximal rate

$$\Gamma_{297}^{\max} = 6\pi \left(\frac{\lambda}{2\pi}\right)^2 \frac{\Gamma/4}{4\pi d^2} \left(\frac{\Gamma/2}{\Delta_{1\to 0}^{297}}\right)^2, \tag{6.15}$$

where $\lambda = 297$ nm. Considering a single ^{171}Yb$^+$ ion in an infinite chain with spacing d and fraction r of ^{172}Yb$^+$ ions, yields the expression for the scattering rate R, with $\Delta = \Delta_{297}^{\max}$. Setting $d = 4\ \mu$m and $r = 0.5$ then yields $R < 2{\times}10^{-3}\ \mathrm{s}^{-1}$, corresponding to 2×10^{-3} error per qubit after 2000 XX gates in a 25-ion chain. Since only the first several axial modes need to be cooled, fewer coolant ions and a lower cooling duty cycle would likely suffice. Moreover, sideband-cooling is a dark-state cooling scheme, and some of the entropy-removing spontaneous scatterings are elastic. Therefore, the true error rate is likely at least an order of magnitude smaller.

The modification of existing CW laser systems is straightforward to support ^{172}Yb$^+$. For loading ^{172}Yb$^+$, the neutral excitation line is detuned by $\Delta_{399}^{1/2\to 3/2} = 2\pi \times -304$ MHz relative to the ^{1}P$_1$ ($F = 3/2$) line in neutral ^{171}Yb [115, 116]. We added a double pass AOM setup running at $\sim$ 150MHz to shift the frequency during loading. A shutter blocking the 0-order after the AOM, combined with an RF switch for the AOM tone, allows us to switch from the $(0, 0)$-order for ^{171}Yb to the $(-1, -1)$-order for ^{172}Yb, with good efficiency and near perfect purity.

Despite the frequency purity and the fact that 399-nm beam is perpendicular to the oven flux, when $I_{399} \sim I_{sat}$ and resonant with the ^{172}Yb line, we observe loading of ^{171}Yb at a rate an order of magnitude more than expected. We attribute this to oven flux that is bouncing off the sidewalls of the load slot and acquiring a Doppler shift relative to the neutral excitation beam. Isotopic purity is important for loading chains with a deterministic mixture ratio of 171/172. We empirically observed that this could be achieved by turning off $^{171}Yb^+$ cooling light in the load slot while attempting to load $^{172}Yb^+$ (while keeping $^{171}Yb^+$ cooling light on in the quantum region to cool chains that may be building up there). We find that even if a $^{171}Yb^+$ ion is trapped in the load slot, it does not make it through the junction of the HOA trap, unless it has been cooled.

For Doppler cooling of $^{172}Yb^+$, the D1 line is detuned by $\Delta_{369}^{1\to0} = 2\pi \times 3.7$ GHz relative to the $^2S_{1/2}(F = 1) \to^2 P_{1/2}(F = 0)$ line in $^{171}Yb^+$ [110, 117]. The approach here was to build a second local oscillator using a laser diode ECDL and beat-note lock it to the $^{171}Yb^+$ 369-nm laser (which is in turn frequency referenced to a neutral Rb transition). Finally, the overall cooling power of this scheme has a non-monotonic dependence on the 935-nm power [108], which can broaden the 435-nm transition and decrease the cooling efficiency. To give flexibility in optimizing the 935-nm power relative to the 435-nm power, we installed a single-sideband optical serrodyne frequency shifter [118] using an NLTL and an in-fiber EOM. With this we can tune the power in the frequency tone at +100 MHz for $^{172}Yb^+$ by $0 - 95\%$ relative to nominal power at the $^{171}Yb^+$ frequency.

Loading a two ion chain with exactly one $^{171}Yb^+$ and one $^{172}Yb^+$ allows us

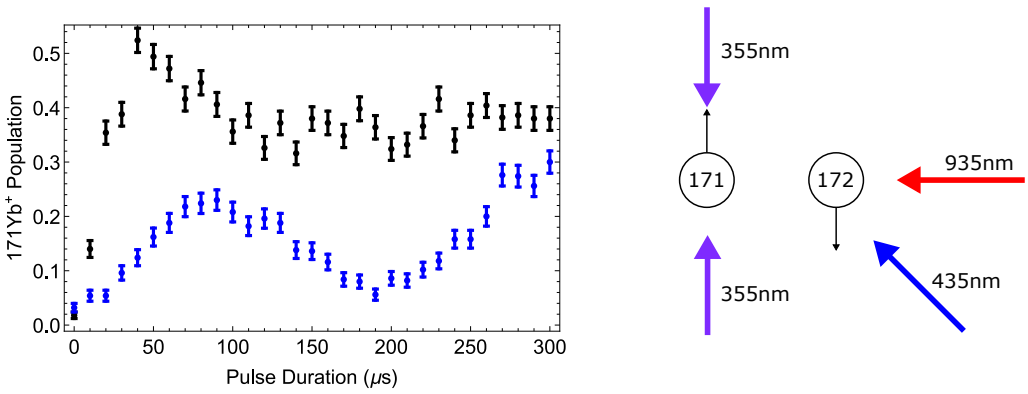

Figure 6.8: Demonstration of sympathetic cooling on a chain of two ions. (Left) red sideband Rabi flop on the out-of-phase "tilt" mode using the standard Raman configuration on the $^{171}Yb^{+}$ ion. The black curve is taken immediately following Doppler cooling and optical pumping corresponding to $\bar{n} = 2.87$ quanta. The blue curve is after an additional 1 ms of sympathetic cooling on the $^{172}Yb^{+}$ ion corresponding to $\bar{n} = 0.42$ quanta. (Right) The laser geometry and ion configuration for this experiment.

to demonstrate this cooling scheme in a proof-of-principle experiment. To probe temperature we can look at a BSB/RSB flop on the lower energy radial "tilt" mode using Raman spectroscopy on the $^{171}Yb^{+}$ ion. Immediately following Doppler cooling and optical pumping, we measure $\bar{n} = 2.87$ quanta on the radial tilt mode. Then, two tones are applied to the 435-nm AOM at -2.976 MHz relative to the $\Delta m = 0$ transitions corresponding the RSB of the tilt mode. Simultaneously, the 935-nm light is shifted to be resonant with the $^{172}Yb^{+}$ $^{2}D_{3/2} \rightarrow^{3} D[3/2]_{1/2}$ transition to pump the ion back into the $^{2}S1/2$ manifold. After 1 ms of sympathetic cooling, we again perform a BSB/RSB flop and measure $\bar{n} = 0.42$ quanta, demonstrating that cooling is achieved. The RSB flop before cooling (black) and after cooling (blue) is shown in Figure 6.8. Qualitatively, the RSB is much weaker after cooling. The laser geometry is also shown on the right of Figure 6.8, with the 435-nm light oriented at 45° relative to the chain.

Scaling this up to a longer chain will require optimization the cooling parameters such as 935-nm power and the number/arrangement on $^{172}Yb^{+}$ sympathetic cooling ions in the chain. An added engineering challenge is to make sure that the configuration of ions within the chain is constant at the start of each quantum circuit. Even though the shift in mode frequencies may be within the range of what the gate can tolerate for detuning errors (see Section 3.3.4), in high-fidelity operations the mode-participation factors would require re-calibration of the gate amplitude if the ion configuration were to change. In the room-temperature system described here, collisions with residual background gas particles re-order the chain randomly on the scale of 4-5 minutes in a 15 ion chain. In order to put the chain back into the nominal configuration, a pairwise split-swap-merge protocol using the DC electrodes of the chip trap will have to be developed to implement a bubble-sort algorithm. With re-ordering implemented, the chain length will now be limited by the speed and efficiency of the cooling relative to the underlying heating rate.

Chapter 7: Fault-Tolerant Quantum Error-Correction

In addition to scaling up the system size, error rates also need to significantly decrease in order to reach compelling applications. For example, the qubit error rate required to factor a 2048-bit number using Shor's algorithm is estimated around 10^{-9} or smaller [119]. Quantum error correcting codes combine multiple physical qubits into *logical* qubits that robustly store information within an entangled state [120, 121, 122]. With quantum error-correction, the error rate requirement can be achieved on abstracted logical qubits, even if the error rate on the physical qubits is significantly higher [123, 124]. In this example, with error-correction the physical error rate required to factor a 2048-bit number is only 10^{-3}, at the cost of increasing the number of physical qubits to 20 million relative to 14 thousand logical qubits [119]. Fundamentally, error correction is a trade-off between physical qubit number and the physical error rate. Decreasing the error rate often requires battling fundamental physical processes, whereas increasing the qubit number is primarily an engineering challenge, making this trade-off favorable. Since the exact qubit overhead required to run quantum error-corrected algorithms depends on the underlying

physical error rate, the error-correction code architecture, and desired algorithmic error rate, there is obvious interest in driving the overhead as low as possible through further reductions in physical error rates or improved code architecture.

However, quantum error-correcting codes are not enough on their own. Fault-tolerant (FT) operations, which limit the ways in which errors can spread throughout the system, must also be used. A distance-3 error-correcting code is capable of correcting an arbitrary error on any one of the physical qubits that make up the logical qubit. However, interacting with logical qubits and performing computations with them involves circuits that may consist of several entangling operations. If not careful, a single entangling gate error can propagate a correlated error through the circuit and corrupt the logical qubit state beyond the error-correction capability of the code. Fault-tolerance is a statement being able to correct any single *circuit fault*, or in other words to ensure that errors in the circuit contain themselves to within the error-correction capabilities of the code. Without fault-tolerance, the logical error rate may be limited by faults at critical circuit locations that cascade into logical failures, negating the advantage of error-correction. So while error-correction is critical to scaling up quantum computers in theory, fault-tolerance is about manifesting that benefit in practice.

So far, FT state preparation, detection, and operations have been demonstrated using quantum error detecting codes with four data qubits [125, 126, 127, 128, 129]. These codes can identify when errors have occurred, but do not extract enough information to correct them. There have also been quantum demonstrations of classical repetition codes to correct quantum errors restricted along one axis

[130, 131, 132, 133, 134, 135]. In other work, qubits have been encoded into quantum error correcting codes that can correct all single qubit errors, but the encoding procedure was not fault-tolerant [136] and the system was not large enough to measure the error syndromes non-destructively using ancilla [137, 138]. Parallel work on bosonic codes has demonstrated encoded operations [139, 140], fault-tolerant detection, one-axis [141], and two-axis [142] error correction on encoded qubits. For both qubit codes and bosonic codes, fault-tolerant state preparation of a code capable of correcting all single-qubit errors has not been achieved to date.

In the subsequent sections, and following the work presented in [143], we demonstrate all of the primitives required for FT operation of an encoded qubit: FT preparation, FT measurement, FT logical gates, and FT stabilizer measurement. Our demonstration protects against any single circuit fault (along any axis and without postselection), realizing quadratic error suppression *in principle*. In practice, this requires local errors and high-fidelity operations on a large enouh system size to support the code. To investigate the error suppression properties of fault-tolerance in our code, we compare non-fault-tolerant (nFT) preparation, nFT logical gates, and nFT stabilizer measurement to their FT counterparts and compare the relative error rates under various decoding schemes.

7.1: Bacon-Shor 13 Subsystem Code

As shown in Figure 7.1, we choose to implement a [[9,1,3]] Bacon-Shor code [144, 145]. It is a distance-3 code, which means it is able to correct any single-qubit error.

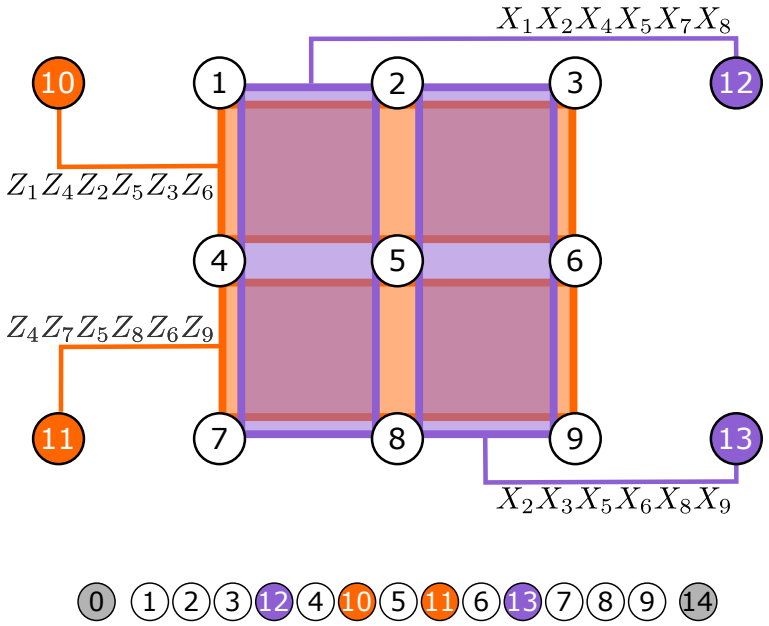

Figure 7.1: (Top) The Bacon-Shor code is a [[9,1,3]] subsystem code that encodes 9 data qubits into 1 logical qubit. Four weight-6 stabilizers are mapped to ancillary qubits 10, 11, 12, and 13, for measuring potential errors in the X (purple) and Z basis (orange). (Bottom) Due to the all-to-all connectivity in the system, the code can be embedded arbitrarily into the ion chain. An example mapping to a 15-ion chain is shown here, with the two edge ions unused as qubits.

There are two X-stabilizers (purple) and two Z-stabilizers (orange), each weight-6, that can be mapped to 4 ancillary qubits (10–13) to detect/correct errors in the code. The logical operators Pauli operators X_L (Z_L) act upon on the columns (rows) of the code. The stabilizers and logical operators are defined in Table 7.1. The code subspace is defined as the set of all states for which all the stabilizers have a $+1$ eigenvalue, $S_i|\psi\rangle_L = +1|\psi\rangle_L$. If any of the stabilizers measure a -1 eigenvalue, then an error has occurred and the quantum state is no longer in code subspace. By examining the structure of the stabilizer eigenvalues, we can decode where we suspect an error has occurred and attempt to apply a correction that will return the state to the code subspace.

This code is well-suited to near-term ion-trap quantum computing architec-

Stabilizers	Logical Operators
$S_1 = Z_1Z_4Z_2Z_5Z_8Z_8$	$X_L = X_1X_4X_7$
$S_2 = Z_4Z_7Z_5Z_8Z_6Z_9$	$Z_L = Z_1Z_2Z_3$
$S_3 = X_1X_2X_4X_5X_7X_8$	
$S_4 = X_2X_3X_5X_6X_8X_9$	

Table 7.1: Bacon-Shor stabilizers and logical operators.

tures for two reasons. First, Bacon-Shor codes can be prepared fault-tolerantly without intermediate measurement. Compared with the typical projective preparation of topological codes, unitary preparation requires fewer gates and less ancillary qubits. This allows us to demonstrate FT primitives with fewer resources and without intermediate measurements. Second, this code choice is a reasonable midpoint between the qubit efficiency of the 7-qubit Steane code and the robustness of the Surface-17 code [146]. Although the Bacon-Shor stabilizers are weight-6 and non-local, they can be fault-tolerantly measured using only one ancilla per stabilizer [147] and leverage the all-to-all connectivity in the device.

As a subsystem code, the Bacon-Shor code is a generalization of Shor's code that has additional degrees of freedom known as gauge qubits [121]. Of the 9 DoF present in the original data qubits, 4 are pinned down by fixing the eigenvalue of the code stabilizers. Of the 5 remaining DoF, 1 DoF is the logical qubit and the other 4 DoFs are the gauge qubits. For particular choices of gauge, the logical states are products of GHZ states:

$$\begin{aligned} |0/1\rangle_L \otimes |X\rangle_G &= \frac{1}{2\sqrt{2}}(|+++\rangle \pm |---\rangle)^{\otimes 3}, \\ |+/-\rangle_L \otimes |Z\rangle_G &= \frac{1}{2\sqrt{2}}(|000\rangle \pm |111\rangle)^{\otimes 3}, \end{aligned} \tag{7.1}$$

where $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$ and $|X/Z\rangle_G$ refer to different states of the gauge qubits. The gauge qubits are not protected to the same distance as the logical degree of freedom, but are useful for designing efficient FT protocols. One basis for the gauge qubits corresponds to fixing 4 constraints on the eigenvalues of the operators shown in the Table 7.2.

X-gauges	Z-gauges
X_1X_2	Z_1Z_4
X_4X_5	Z_2Z_5
X_7X_8	Z_3Z_6
X_2X_3	Z_4Z_7
X_5X_6	Z_5Z_8
X_8X_9	Z_6Z_9

Table 7.2: Gauge operators in the Bacon-Shor 13 code. Gauge choices $|X\rangle_G(|Z\rangle_G)$ correspond to all $X(Z)$-type gauge operators having eigenvalue $+1$.

It should be noted that this is not an independent set of operators because the stabilizers of the code, which are products of gauges, already have their eigenvalues fixed to $+1$. As such, if the X-gauges X_1X_2 and X_4X_5 both have eigenvalue $+1$ on a given logical state, then the eigenvalue of X_7X_8 will also be $+1$. We refer to a state in which all $X(Z)$-type gauge operators have eigenvalue $+1$ as the $|X\rangle_G(|Z\rangle_G)$ gauge. It should be noted that these gauge operators do not commute, so these two gauges are mutually exclusive. When decoding the Bacon-Shor code, we can only identify operators up to a product of gauges. Thus, logical primitives may also apply a gauge operator to the logical state. This leaves the logical qubit unaffected, but alters the state of the gauge qubits.

As an additional note of caution, logical Pauli operators on a subsystem code must decompose as a tensor product of operations on the logical and gauge degrees

of freedom. When a logical Pauli operator that acts non-trivially on the gauge subsystem is used to generate a continuous unitary operator, it will entangle the logical and gauge subsystems. As these gauge subsystems are less protected than the logical subsystem, the quantum information will also be less protected. Consequently, one must design continuous logical operators around logical Pauli operations that commute with the entire gauge group, ensuring that it acts trivially on the gauge subsystem.

Bacon-Shor codes support a wide range of FT operations, including state measurement, state preparation, gates, and stabilizer measurement. FT state measurement (in the X/Z basis) is performed by simply measuring all the data qubits (in the X/Z basis). From this information, one can recover relevant stabilizer outcomes as correlations among the single data-qubit outcomes.

FT state preparation in the X_L/Z_L basis is achieved by the definition of the product states. Fault-tolerance, as a design principle, ensures faults on physical operations do not propagate to uncorrectable multi-qubit failures in the circuit. As seen in Eq. 7.1, not all Bacon-Shor logical states require global entanglement. It is precisely this decomposition into decoupled GHZ states that allows Bacon-Shor to be prepared unitarily and fault-tolerantly. In the X_L/Z_L basis, the logical information is encoded redundantly into the relative phase of each state. While a single circuit fault may corrupt one of the three GHZ states, the information can be recovered from the other two.

Fault-tolerance in logical gates is often achieved via *transversal gates*, which are physical operations that act independently on each qubit in a code block.

Bacon-Shor codes have transversal constructions, when allowing permutations, for $\{CNOT_L,\ H_L,\ Y_L\left(\pi/2\right),\ X_L\}$ [148, 149, 150]. Here, $Y(\theta)$ indicates exponentiation of the Pauli-$\hat{Y}$ matrix, $e^{-i\theta\hat{Y}/2}$. FT non-Clifford logical gates, which are required for universality, can be achieved through magic state distillation [151].

Finally, measuring error syndromes requires interacting ancillae with multiple data qubits, which could cause damaging correlated errors. However, fault-tolerance is achieved by carefully ordering the interactions, so that correlated errors can be reduced to low-weight errors up to a benign transformation of the gauge subsystem [147, 152].

7.2: Fault-tolerant State Measurement and Decoding

A prototypical quantum memory experiment should involve three elements: (1) FT state preparation, (2) N rounds of FT syndrome extraction, and (3) FT state measurement. In the final round of the experiment, in order to recover the encoded measurement outcome, one must measure all of the data qubits individually. Afterwards, correlations among these single-qubit outcomes must be post-processed into the relevant stabilizer information. This post-processed information is then combined with any previously extracted syndromes, and then collectively decoded to produce a correction. When N is large, the decoder can become quite complex and may use maximum likelihood estimation or union-find algorithms [152], but in our case where $N = 0$, the decoder takes on a simple form.

Global measurement at the end of each circuit provides the state of all nine

data qubits. From this data, we can calculate the raw total parity, $Z_L = Z_1 Z_2 ... Z_8 Z_9$, and the eigenvalue of the two Z stabilizers, S_1 and S_2. The post-processed total parity, Z'_L from the different decoding schemes is then given by Table 7.3.

Protocol	**If**	**Then**	**Else**
Raw	True	$Z'_L = Z_L$	
Correction	$S_1 = -1 \parallel S_2 = -1$	$Z'_L = -Z_L$	$Z'_L = Z_L$
Detection	$S_1 = -1 \parallel S_2 = -1$	Discard data	$Z'_L = Z_L$

Table 7.3: Decoding protocols for the $N = 0$ case of logical encoding and then measurement.

Raw simply returns the total parity of all the data qubits, regardless of whether an error has occurred. Error *correction* attempts to apply a correction to the data qubits based on the stabilizer outcomes, and yields an expected quadratic suppression of uncorrelated errors (i.e. corrects any single error). Note that no data is discarded in this decoding scheme. Error *detection* is performed by post-selecting experimental shots that have all +1-eigenvalues of the stabilizers, and is equivalent to projecting the data into the code subspace. This will yield an expected cubic suppression of uncorrelated errors (i.e. detects any pair of errors). Importantly, the error-correction decoding step will occur even with large FT quantum computers, involves no post-selection, and does not differ in any way from quantum memory experiments where $N > 0$.

7.3: Fault-tolerant Encoding

We embed the 9 data qubits and 4 ancilla qubits of the Bacon-Shor-13 code in a single chain of 15 ions (Figure 7.1), with the two end ions left idle to obtain

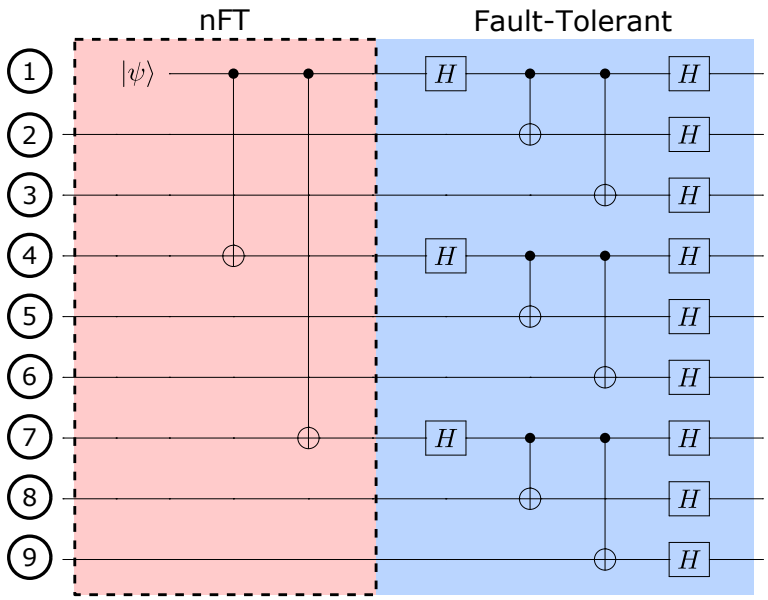

Figure 7.2: Encoding circuit for creating logical qubit states. The right subcircuit (blue) is used for FT preparation of Z-logical basis states. X-logical basis states can be created by omitting the final Hadamard gates. The left subcircuit (red, dashed) can be optionally prepended for nFT preparation of arbitrary logical states.

uniform spacing of the central 13 ions. The mapping of the code onto the chain is chosen to minimize two-qubit gate crosstalk (see Section 5.3.4).

The encoding circuit used to create logical states is shown in Figure 7.2. The right sub-circuit (blue) is FT because there are no entangling operations between independent GHZ states that would allow errors to propagate; however it is limited to preparation of only Z and X basis states. One may prepend an optional sub-circuit (red, dashed) that enables the encoding of arbitrary $|\psi\rangle_L$ states, controlled by a single physical qubit state $|\psi\rangle$. This circuit can produce global entanglement, and allows the possibility of early errors spreading between the separate GHZ states. As a consequence, this circuit loses the FT properties of the X and Z basis preparation circuits. To directly investigate the properties of fault-tolerance, we compare the encoding performance of the right FT sub-circuit to the full nFT circuit with $|\psi\rangle \in$

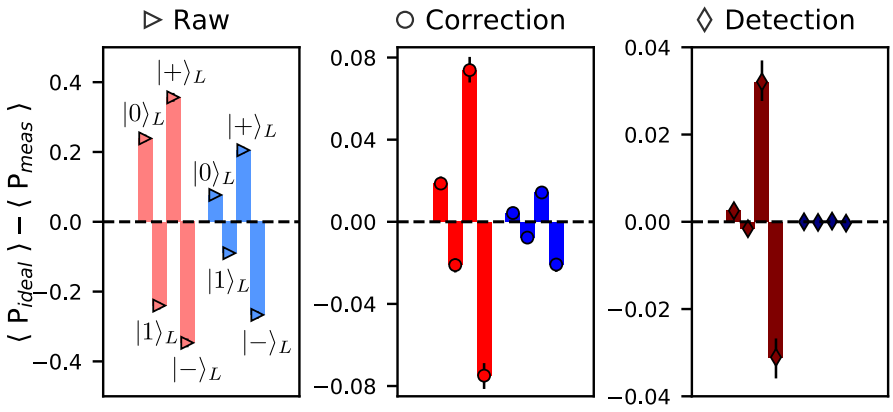

Figure 7.3: Errors for the key basis states of the encoded logical qubit. In both the FT (blue) and nFT (red) case, the measured expectation value of the parity ($P_{meas} = \langle Z/X \rangle_L$) is compared against the ideal parity of the logical state ($P_{ideal} = \pm 1$). We compare the two state preparation methods under the raw, correction, and detection decoding protocols described in Section 7.2. The data is from running the circuit in Figure 7.2

$\{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$.

After measuring the data qubits, the logical measurement outcome is determined by calculating the total parity of all the data qubits in the Z-basis, $Z_L = Z_1 Z_2 ... Z_8 Z_9$. From Eq. 7.1, the $|0\rangle_L$ state has even parity ($\langle Z \rangle_L = +1$) while $|1\rangle_L$ has odd parity ($\langle Z \rangle_L = -1$). Similarly, the $|+/-\rangle_L$ states have even/odd parity in the X_L basis; a $Y_L(-\pi/2)$ operation following the encoding circuit, maps $\langle X \rangle_L \rightarrow \langle Z \rangle_L$. The decoded parity compared to the ideal parity of each logical Z, X basis state is presented in Figure 7.3.

Using the FT circuit and performing error correction, we prepare $|0\rangle_L, |1\rangle_L, |+\rangle_L$, and $|-\rangle_L$ states with respective errors $0.21(4)\%, 0.39(5)\%, 0.71(7)\%$, and $1.04(9)\%$. We note that the average state preparation and measurement error for a single physical qubit in the Z basis is $0.46(2)\%$ (see Section 5.3.1) compared to $0.30(3)\%$ in the logical qubit. This is one context in which the logical qubit clearly outperforms our

physical qubit. For the nFT circuit the respective errors are 0.93(8)%, 1.05(9)%, 3.7(2)%, and 3.8(2)%. The error-detection experiment presents particularly strong evidence for fault-tolerance. We observe a remarkable gap in the failures between the nFT and FT protocols: averaged over the basis states, we see 2 failures of FT error-detection over 13,288 post-selected shots, compared with 197 failures over 12,105 post-selected shots when using nFT error-detection. This agrees with a local error model where we expect cubic suppression of FT error-detection, in stark contrast with nFT error-detection, which can fail due to a single circuit fault. The observed two orders-of-magnitude difference lends further evidence that these circuits, which are fault-tolerant in principle, are also fault-tolerant in practice.

The comparison of FT and nFT encoding circuits among the basis states is a bit superficial because there is no reason to choose a nFT circuit when a FT circuit exists. However, the nFT preparation circuit is useful for creating arbitrary quantum states, and in particular the magic states $|H_x\rangle_L = e^{-i\pi\hat{Y}/8}|0\rangle_L$ and $|H_y\rangle_L = e^{-i\pi\hat{X}/8}|0\rangle_L$, which can be distilled to implement FT non-Clifford gates [151, 153]. These states are depicted on the logical Bloch sphere along with their $\langle Z_L\rangle$ projection in Figure 7.4. After encoding, we perform $Y_L(\pi/2)$ rotations to bound the fidelity of these states.

To calculate the fidelity of our magic state preparation circuit for the state $|H_x\rangle_L$, we can compute the fidelity between a mixed state ρ, which represents the

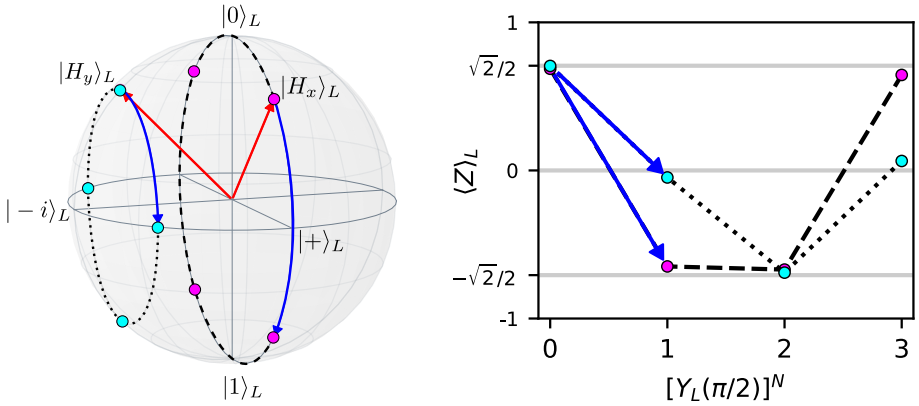

Figure 7.4: (Left) Magic states $|H_x\rangle_L$ (magenta) and $|H_y\rangle_L$ (cyan) are directly encoded using the full nFT circuit from Figure 7.2. Subsequent $Y_L(\pi/2)$ rotations (blue arrows) are used to bound the fidelity. (Right) Experimental $\langle Z\rangle_L$ values for the states depicted on the left.

experimentally prepared state, and the ideal pure state $|\psi\rangle$ as

$$\begin{aligned} F &= \langle\psi|\rho|\psi\rangle \\ &= \mathrm{Tr}[\rho|\psi\rangle\langle\psi|] \\ &= \frac{1}{2}\left(1 + \langle X\rangle_\rho\langle X\rangle_\psi + \langle Y\rangle_\rho\langle Y\rangle_\psi + \langle Z\rangle_\rho\langle Z\rangle_\psi\right) \\ &= \frac{1}{2}\left(1 + \langle X\rangle_\rho\frac{1}{\sqrt{2}} + \langle Z\rangle_\rho\frac{1}{\sqrt{2}}\right). \end{aligned} \tag{7.2}$$

The expectation values $\langle Z\rangle_\rho$ and $\langle X\rangle_\rho$ can be extracted by measuring logical Z operator, $\langle Z\rangle_L$ before and after a logical $Y_L(\pi/2)$ operation. This analysis leads to the following fidelities as shown in Table 7.4. Importantly, after error correction, the calculated $|H_x\rangle_L$ encoding fidelity is 97(1)%, which is above the distillation threshold of 92.4% [153].

The same procedure cannot be applied to the $|H_y\rangle_L$ state, as the [[9,1,3]] Bacon-Shor code does not allow for fault-tolerant measurement in the logical Y

Processing Technique	Fidelity
Raw	0.85(1)
Correction	0.97(1)
Detection	0.98(1)

Table 7.4: Fidelities for the $|H_x\rangle_L$ magic state preparation circuit under different processing techniques.

basis. Using the constraint

$$\langle X\rangle^2 + \langle Y\rangle^2 + \langle Z\rangle^2 \leq 1$$

we can only numerically bound the fidelity of the $|H_y\rangle_L$ state to the range $0.75 \leq F \leq 0.99$. However, we argue that the fidelities for preparing $|H_x\rangle_L$ and $|H_y\rangle_L$ should be very similar, as the preparation circuit only differs in the phase of a single qubit gate, a quantity which we control to $\approx 400\mu$rad limited by the AWG bit depth. Thus the $|H_y\rangle_L$ state fidelity should be very similar to values shown in Table 7.4.

7.4: Fault-tolerant Logical Gates

We implement a $Y_L(\theta)$ rotation on the encoded qubit, which can only be performed transversally for a discrete set of angles [154]. This is a consequence of the Eastin-Knill theorem, which states that continuous operations cannot be implemented transversally. In the case of Bacon-Shor, the smallest transversal $Y_L(\theta)$ rotation we can create is $Y_L(\pi/2)$, which is generated by applying a physical $Y(\pi/2)$ to each data qubit, followed by relabeling the data qubit indices in post-processing, as shown in Figure 7.5 (left, blue). We compare the performance of this FT rotation with a nFT circuit which implements $Y_L(\theta) = Y_1 Z_2 Z_3 X_4 X_7(\theta)$ (Figure 7.5, left,

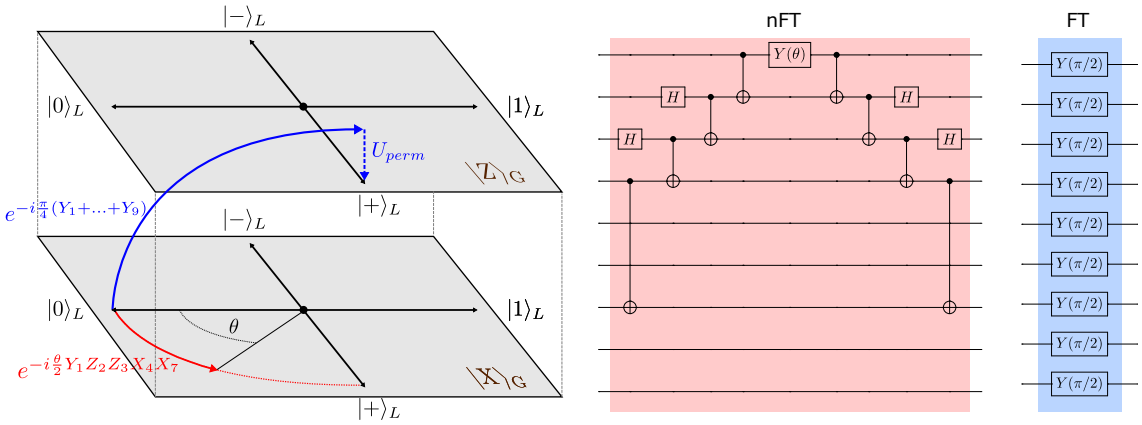

Figure 7.5: (Left) A schematic depicting different logical operations. A FT discrete logical rotation (blue) operating on $|0\rangle_L$ is a transversal operation, $Y_L(\pi/2) = Y(\pi/2)^{\otimes 9}$, that leaves the code subspace (gray planes) and returns via a permutation of qubit labeling (U_{perm}). A nFT continuous logical rotation (red) operating on $|0\rangle_L$ is a 5-qubit entangling operation, $Y_L(\theta) = Y_1Z_2Z_3X_4X_7(\theta)$, that rotates through the code subspace. At $\theta = \pi/2$, these gates are equivalent up to a gauge transformation. (Right) The circuit for the nFT gate capable of creating any state along the red curve in the schematic on the left. The circuit for the FT gate shown by the blue curve in the schematic on the left.

red). The rotation angle is controlled by a single qubit $Y(\theta)$ rotation on data qubit 1. In a perfect system, these rotations are equivalent for $\theta = N\pi/2$, $N \in \mathbb{Z}$ on the logical qubit, but differ in their operation on the gauge qubits. The circuits for each of these operations is shown in the right-hand side of Figure 7.5. The nFT gate (red) generates entanglement among the separate GHZ states, and so the failure of a single operation in the circuit can lead to the failure of the logical qubit. On the other hand, the FT circuit (blue) is clearly fault-tolerant by its transversal construction.

The results of these different gate operations on the logical qubit are shown in Figure 7.6 (left). In both cases, the expectation value of the logical Z operator is fit to a decaying sinusoid $\langle Z\rangle_L = A\cos(\theta)e^{-\Gamma\theta/\frac{\pi}{2}}$. The gate error per $\pi/2$ angle, corresponding to fit parameter Γ, is 0.3(1)% for the FT gate after error-correction.

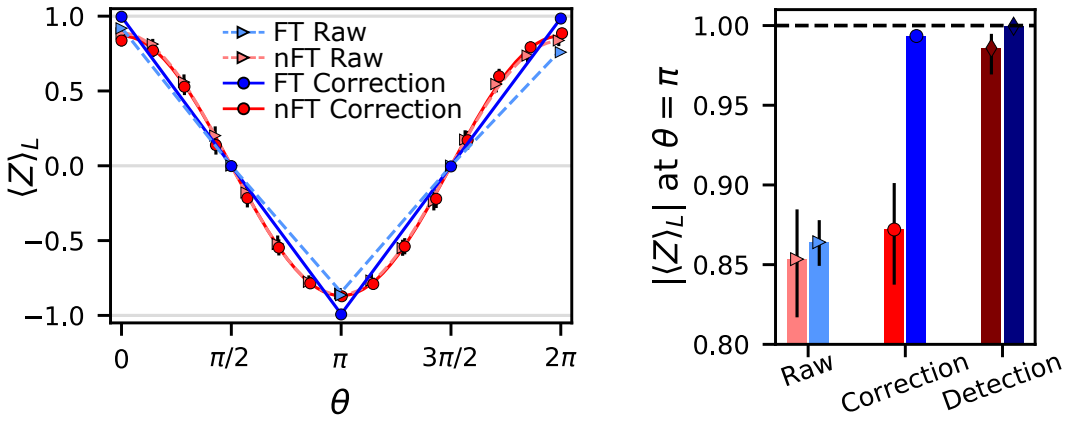

Figure 7.6: The data from running the circuits in Figure 7.5. (Left) Logical Rabi flop comparing FT (blue) and nFT (red) logical operations. (Right) Detailed view at $\theta = \pi$.

This error rate explains the additional error present for the $|+/-\rangle_L$ states in Figure 7.3, which require two additional $Y_L(\pi/2)$ gates for state preparation and measurement. The error at $\theta = \pi$, the maximum gate angle required with optimized circuit compilation, is shown in Figure 7.6 (right). The error for the FT gates and nFT continuous rotations is 0.33(18)% and 6.4(1.6)%, respectively, after error correction. Compared to the FT circuit, error correction on the nFT rotation provides minimal gains, indicative of a high proportion of weight-2 errors relative to weight-1 errors. In contrast, $\langle Z \rangle_L$ recovers quite significantly after error detection, indicating that there are still few weight-3 or higher errors in the system. This is a striking example of the value of fault-tolerance, which minimizes the impact of correlated weight-2 errors on the logical qubit.

7.5: Logical Qubit Memory

Performance of the logical qubit as a quantum memory can be characterized by measuring the coherence of $|+\rangle_L$ versus time, as in a typical Ramsey experiment.

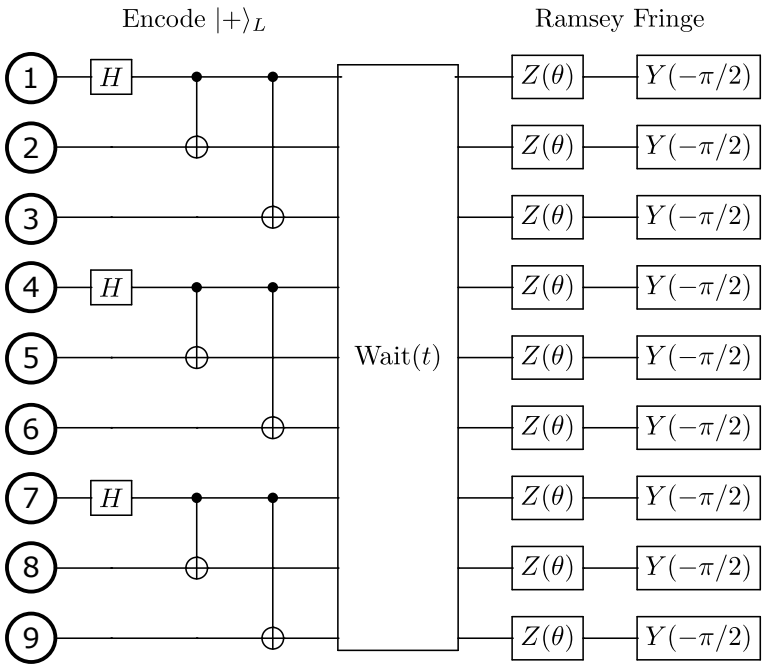

Figure 7.7: A logical Ramsey experiment to characterize the logical qubit memory coherence as a function of time.

The circuit for this experiment is shown in Figure 7.7. After encoding $|+\rangle_L$ and waiting a certain time, a varying $Z(\theta)$ gate is applied to every data qubit, followed by $Y_L(-\pi/2)$. A fit of $\langle X\rangle_L$ depending on θ to a Ramsey fringe yields the Ramsey amplitude for that wait time.

As shown in Equation 7.1, the logical $|+\rangle$ state we use is composed of three GHZ states $\frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$. Due to the structure of these states, if a $Z(\theta)$ gate is applied to each qubit, the three gates will coherently combine, and the end result will be the same as if a $Z(3\theta)$ gate had been applied on any single qubit. In this case, we only need to scan theta from $\theta = [-\pi/3, \pi/3]$ to reveal the full Ramsey fringe.

For the "raw" processing case, any Z error flips the logical output. As a result the cases where 1 or 3 errors occur lead to $|-\rangle_L$ states, while while cases with 0 or

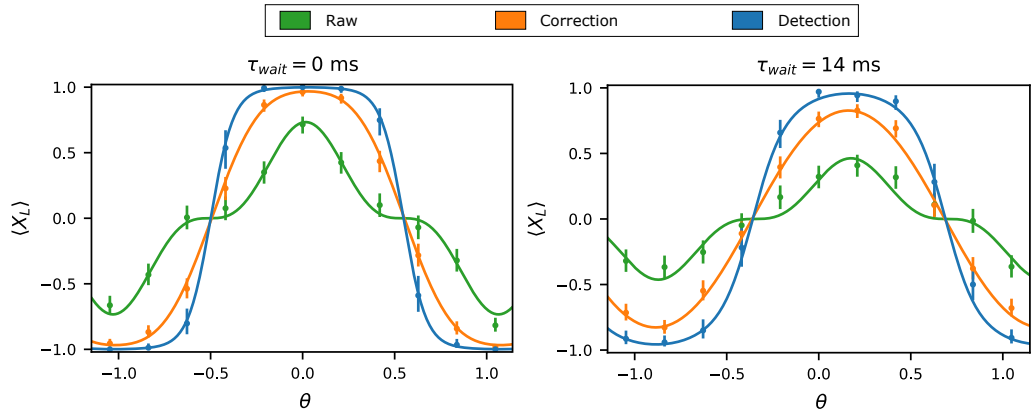

Figure 7.8: Examples of logical Ramsey fringe fitting at two different wait times, 0ms (left) and 14ms (right). The error model well matches the experimental data.

2 errors lead to $|+\rangle_L$. Consequently the expectation value of X_L can be thought of as the squared amplitude of cases which lead to $|+\rangle_L$, subtracted by the squared amplitude of cases which result in $|-\rangle_L$. This results in a curve

$$\begin{aligned}\langle X_L \rangle &= \cos(3\theta/2)^6 - 3\cos(3\theta/2)^4 \sin(3\theta/2)^2 + 3\cos(3\theta/2)^2 \sin(3\theta/2)^4 - \sin(3\theta/2)^6 \\ &= \cos(3\theta)^3.\end{aligned} \tag{7.3}$$

In the "corrected" processing case, the state can tolerate a single error without having its logical information corrupted. As a result error cases with 0 or 1 errors lead to $|+\rangle_L$, while 2 or 3 lead to $|-\rangle_L$. This results in the curve

$$\langle X_L \rangle = \cos(3\theta/2)^6 + 3\cos(3\theta/2)^4 \sin(3\theta/2)^2 - 3\cos(3\theta/2)^2 \sin(3\theta/2)^4 - \sin(3\theta/2)^6. \tag{7.4}$$

Lastly the 'detected' processing method is slightly more complex, as postselection means we must renormalize the expectation value. The case with 0 errors

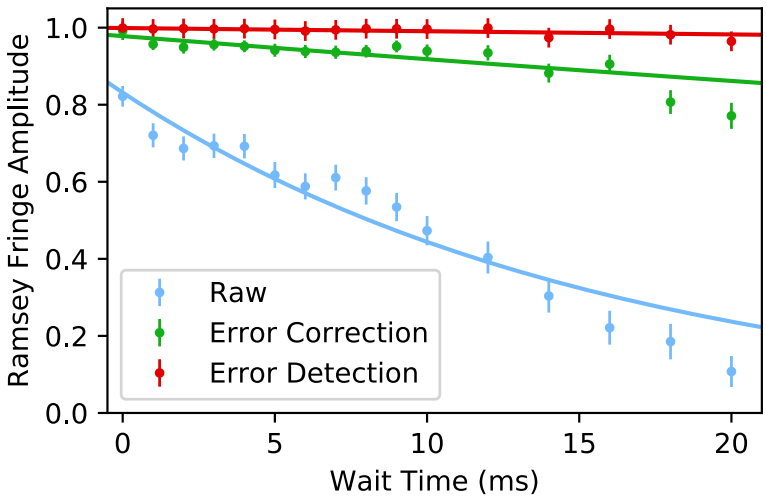

Figure 7.9: The logical qubit Ramsey fringe amplitude as a function of wait time, measuring the coherence of the $|+\rangle_L$ state.

leads to $|+\rangle_L$, while the case with 3 errors leads to $|-\rangle_L$. Cases with 1 or 2 errors must set off at least one stabilizer, and as a result those runs will be removed from the dataset. As a result the probabilities must be renormalized, leading to the curve

$$\langle X_L \rangle = \frac{\cos(3\theta/2)^6 - \sin(3\theta/2)^6}{\cos(3\theta/2)^6 + \sin(3\theta/2)^6}. \tag{7.5}$$

These models well describe the experimental data, as shown in Fig. 7.8(b). From the fits we can extract the amplitude and phase shift of each curve.

Repeating this fitting procedure at varying wait times leads to the coherence decay curves shown in Figure 7.9. Fitting the data in each decoder case to a decaying exponential Ae^{-t/T_2^*}, for the raw, error correction, and error detection decoding schemes, we measure a T_2^* of 16(2) ms, 160(30) ms, and 1200(400) ms.

To understand the phase flip errors in the logical qubit, we measure the T_2^* of a physical qubit in a chain of 15 ions. This is accomplished via a laser Raman Ramsey sequence on the center ion, $RY(\pi/2) - \tau_{\text{wait}} - RZ(\theta) - RY(-\pi/2)$, with

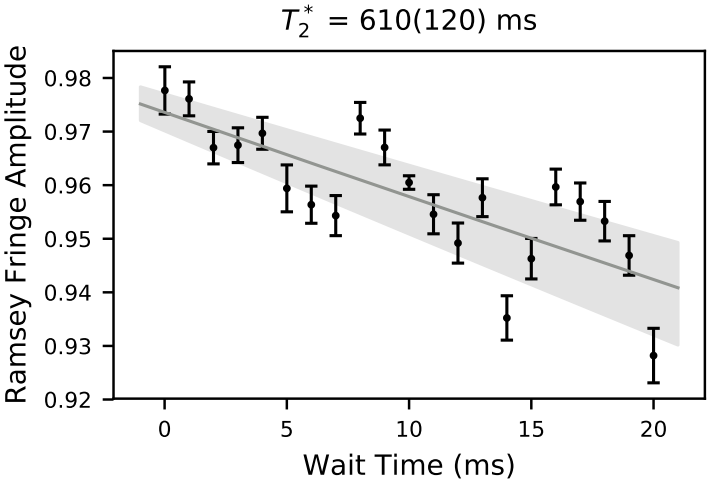

Figure 7.10: Ramsey fringe amplitude as a function of time for a single qubit is a chain of 15 ions. The shaded region indicates the 1σ uncertainty in the exponential decay least-squares fit.

no echoes. At each wait time τ_{wait}, the phase θ is swept, and the resulting data is fit to a sinusoid to extract the contrast. The Ramsey contrast is fit to a decaying exponential $Ae^{-\tau_{\text{wait}}/T_2^*}$ to extract T_2^*. The results of this experiment are shown in Figure 7.10. We find $T_2^* = 610(120)$ ms for a physical qubit in a chain of 15 ions. We attribute the physical qubit decoherence primarily to control noise, rather than to fundamental qubit decoherence. In particular, we note that there are features of revivals at ≈ 8 ms and 16 ms, corresponding to noise at ≈ 125 Hz. We assign this to mechanical fluctuations (e.g., fans) that shift the standing wave of the optical Raman beams relative to the ions. In particular, this error is coherent and global with respect to its effect across the chain. In the future, this effect can be mitigated by switching to a "phase-insensitive" configuration [155].

The GHZ states that make up the Bacon-Shor code ($|000\rangle + |111\rangle$) are three times as sensitive to phase noise as our physical qubit. To understand the implication

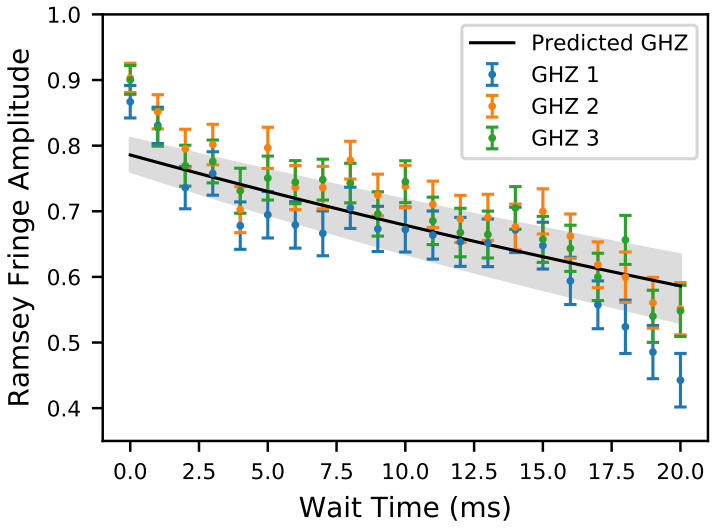

Figure 7.11: Ramsey fringes for each independent GHZ state are analyzed from the logical T_2^* experiment in Fig. 7.9. For the theoretical prediction, we apply the fitted dephasing noise in Fig. 7.10 to a numerical simulation of a three-qubit GHZ state. The shaded region indicates the 1σ uncertainty in the exponential decay least-squares fit in Fig. 7.10, propagated through the numerical simulation.

of the physical T_2^* on the logical T_2^*, we run numerical simulations to extrapolate the measured phase noise to a GHZ state. We assume that the Pauli-Z noise in the middle of the Ramsey sequence is Gaussian distributed with some width Δ_Z. Using the fit from Figure 7.10, we can numerically solve for the width of the noise spectrum Δ_Z. Once this value is found, we re-run the simulation with that noise spectrum on a three-qubit GHZ state to extract the predicted contrast. In Figure 7.11, we compare this predicted value to the three individual GHZ states measured in the logical qubit memory experiment. We conclude that almost all the dephasing in the logical qubit that we observe is explained by the observed T_2^*-decay in the physical qubit. We note that this is the same experimental data presented in Figure 7.9, just post-processed to analyze individual GHZ states rather than to perform error-correction.

7.5.1: Decoherence-Free Subspaces

The construction of the $|+\rangle_L$ state ($|000\rangle + |111\rangle$) in the Bacon-Shor code, leads to increased sensitivity to phase noise. Worse still, this error accelerates as the code size increases to distance-4 ($|0000\rangle + |1111\rangle$) and distance-5 ($|00000\rangle + |11111\rangle$), a feature that is unfavorable for the scalability of error-correction in our system. In other work, GHZ-like states that are constructed out of decoherence free subspaces have been shown to be robust to this type of noise [156, 157]. We define the ferromagnetic (anti-ferromagnetic) GHZ state of size n as,

$$\begin{aligned} |FM_n\rangle &= \frac{1}{\sqrt{2}}\left(|000\ldots\rangle + |111\ldots\rangle\right), \\ |AFM_n\rangle &= \frac{1}{\sqrt{2}}\left(|010\ldots\rangle + |101\ldots\rangle\right), \end{aligned} \tag{7.6}$$

In our system, where the phase noise is global and coherent (due to common beam path fluctuations or noise in the magnetic field), the even length AFM states $|AFM_{2n}\rangle$ will experience no decoherence, as each half of the GHZ state will acquire the same phase, resulting in a non-measurable global phase. The odd length AFM states $|AFM_{2n+1}\rangle$ will experience decoherence on the scale of a single physical qubit. To test this theory in our system, we create $|FM_n\rangle$ and $|AFM_n\rangle$ states for $n = 3, 4, 5$ and measure their coherence as function of time in Figure 7.12. For all n, the AFM GHZ states have significantly longer coherence than their FM counterparts. Notably, the large drop in contrast over the first 2 ms in the FM states is almost completely absent from the AFM states.

We can extend this increased coherence of AFM states, to the full Bacon-Shor

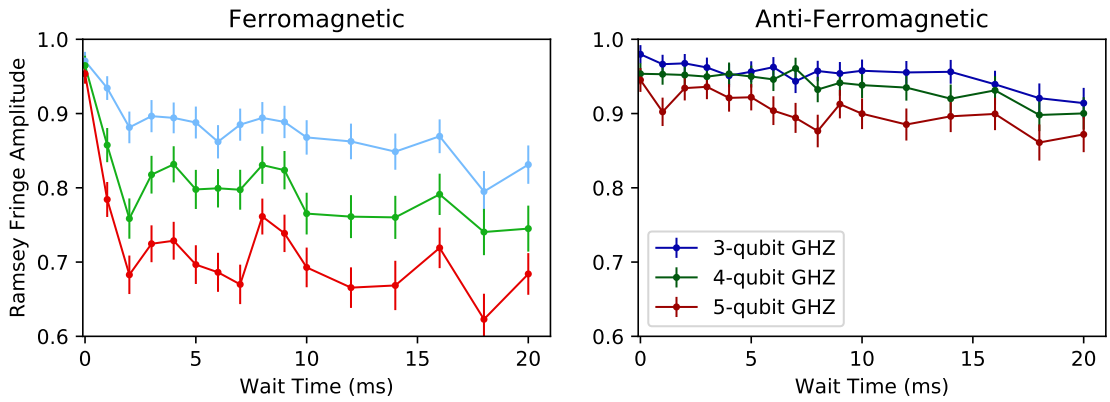

Figure 7.12: (Left) The Ramsey fringe amplitude as a function of wait time, for varying length ferromagnetic GHZ states $|000\ldots\rangle + |111\ldots\rangle$. (Right) The Ramsey fringe amplitude as a function of wait time, for varying length anti-ferromagnetic GHZ states $|010\ldots\rangle + |010\ldots\rangle$.

code using the gauge degrees of freedom. Define the $|\overline{Z}\rangle_G$ gauge as the state in which all the Z gauge operators (Table 7.2) have a -1 eigenvalue. Then the definition of the $|+\rangle_L$ state becomes,

$$|+\rangle_L \otimes |\overline{Z}\rangle_G = \frac{1}{2\sqrt{2}}(|010\rangle + |101\rangle)^{\otimes 3}, \tag{7.7}$$

Which we refer to as the AFM Bacon-Shor code. Repeating the logical qubit memory experiment, we show the coherence of the AFM logical qubit in Figure 7.13. As before, we fit the data in each decoder case to a decaying exponential Ae^{-t/T_2^*}. For the raw, error correction, and error detection decoding schemes, we measure a T_2^* of 47(6) ms, 460(60) ms, and 6600(1300) ms. For the case of error correction, the AFM Bacon-Shor code has a three times longer coherence time than that of FM code and is roughly consistent with the physical qubit coherence time. We stress that because we only modified the gauge degrees of freedom in the code, this does

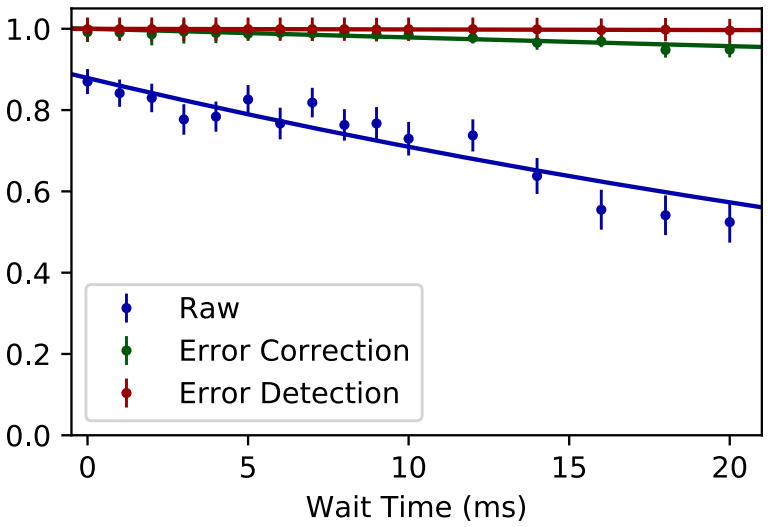

Figure 7.13: The logical qubit Ramsey fringe amplitude as a function of wait time, measuring the coherence of the AFM $|+\rangle_L$ state. In Figure 7.9, the $|+\rangle_L$ took the form of the product state $|000\rangle + |111\rangle$ (ferromagnetic). Here the $|+\rangle_L$ takes the form of the product state $|010\rangle + |101\rangle$ (anti-ferromagnetic).

not change any of the error-correction properties of the code or impact any of the other results we have previously demonstrated. In this case, manipulation of the gauge qubits significantly increases the performance of the logical qubit.

7.6: Fault-tolerant Stabilizer Measurements

In stabilizer measurements, fault-tolerance is achieved by a specific ordering of the interactions between the ancilla and the stabilizer block [147]. To test this theory, we insert a variable $Z(\theta)$ error on ancilla 12 during the measurement of a single stabilizer ($S_3 = X_1X_2X_4X_5X_7X_8$) and compare the impact of this error in a FT ordering and a nFT ordering. The circuits for these experiments are shown in Figure 7.14. In both cases, without correction, a Z error on the ancilla qubit will propagate to an X error on the data qubit and flip $\langle Z\rangle_L$.

The results of this experiment are shown in Figure 7.15. At the extreme case

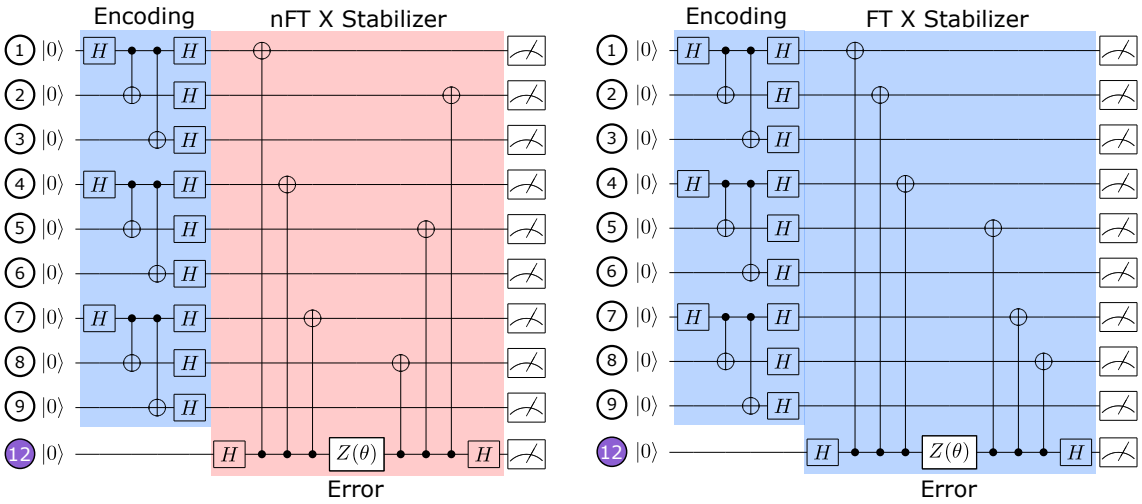

Figure 7.14: Non-fault-tolerant (red, right) and fault-tolerant (blue, right) stabilizer measurement orderings, performed on a FT-encoded $|0\rangle_L$ state. In both cases, a variable error $Z(\theta)$ is introduced on the ancilla qubit in the middle of the stabilizer measurement operation to test how the error propagates.

of $\theta = \pi$, the raw parity is nearly identical in the two cases, but after correction, the FT stabilizer measurement recovers the correct logical parity whereas the nFT stabilizer measurement induces a logical fault. This is because the FT gate ordering propagates a correlated error that decomposes as the product of (at worst) a single qubit fault and a benign transformation of the gauge subsystem. By comparison, the nFT gate ordering propagates a correlated error that corresponds to a logical operator and directly corrupts the logical subsystem.

At $\theta = 0$, (i.e., when no error is added) the error-corrected error rates for $|0\rangle_L$ after the nFT and FT stabilizer measurement are 0.76(22)% and 0.20(13)%, respectively, compared to a baseline encoding error of 0.23(13)%. To within statistical error, there is no distinction between performing the FT stabilizer measurement or not, providing strong evidence that this procedure does not corrupt the logical qubit state beyond the error-handling capabilities of the code. On the other hand, there *is*

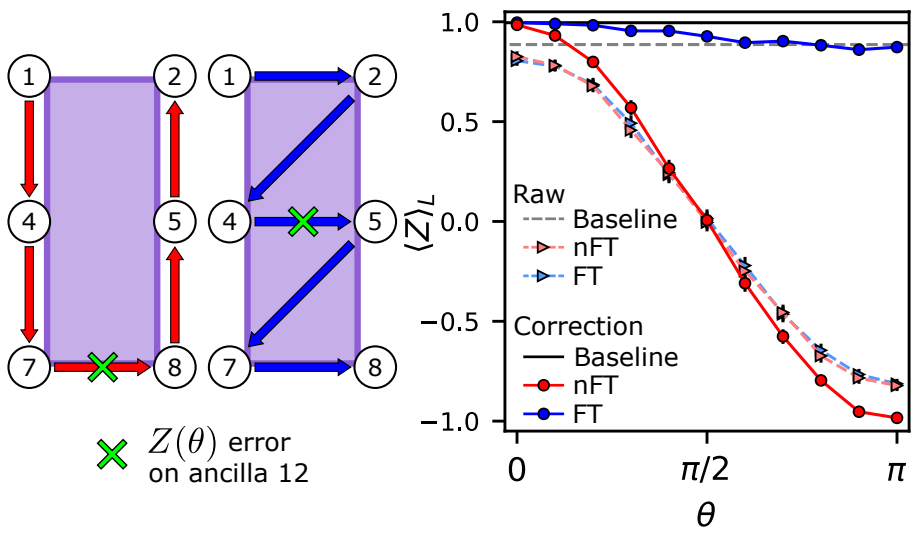

Figure 7.15: (Left) Schematic depicting where the artificial error occurs for the nFT (red) and FT (blue) ordering of the stabilizer measurement. (Right) The data from running the circuits in Figure 7.14. Expectation value of the logical Z operator after encoding $|0\rangle_L$ (Baseline, grey/black line), and then performing the nFT (red) or FT measurement (blue) of a single X-type stabilizer with a $Z(\theta)$ error inserted on the ancilla during measurement.

a statistically significant difference (p-value < 0.015) between the nFT and FT ordering. This again demonstrates the value of fault-tolerance in an apples-to-apples comparison: in two circuits of identical complexity, performing the circuit fault-tolerantly yields an average 4-times reduction in error. The fact that this reduction is not larger speaks to the precise phase control in our system.

In this experiment, because the input and measured state was $|0\rangle_L$, the measurement is only sensitive to X-type errors. To check if the FT stabilizer is introducing Z-type errors into the logical qubit, we perform the same experiment (with no artificial error added, $\theta = 0$) on the $|+\rangle_L$ input state. $|+\rangle_L$ is measured in the $\langle X \rangle_L$ basis by a transversal $Y_L(-\pi/2)$ gate that maps $\langle X \rangle_L \rightarrow \langle Z \rangle_L$. We also check a Z stabilizers on both $|0/+\rangle_L$ input states. The results are shown in table 7.5. We note that from the logical T_2^* experiment (Section 7.5), we expect the error on the

$|+\rangle_L$ state to increase by 1.8% over the $\approx$ 1.5ms required to measure the stabilizer, which is not included in the baseline encoding error below.

Input State	Baseline Encoding (% Error)	FT Z Stabilizer (% Error)	FT X Stabilizer (% Error)
$\|0\rangle_L$	0.23(13)	0.41(10)	0.20(13)
$\|+\rangle_L$	0.45(11)	3.3(3)	2.1(2)

Table 7.5: Logical error rates for both Z and X FT stabilizers on different logical input states, after error-correction.

Now that we have verified the theory of stabilizer fault-tolerance on a single stabilizer measurement, we proceed to directly measuring the full set of stabilizers with four additional ancilla qubits. First, the state is fault-tolerantly encoded into the $|0\rangle_L$ state. Then, an artificial single-qubit error is applied to a particular data qubit. Finally, the full set of stabilizers, in sequential order X and then Z, are mapped to the ancilla qubits in a single shot. The circuit for this experiment is shown in Figure 7.16.

If no error has occurred, all four stabilizers commute with the logical qubit state and the ancilla qubits should remain in the $|0\rangle$ state. Conversely, if an error did occur on a data qubit, the stabilizers that do not commute with that error flip the state of the ancilla to $|1\rangle$. For example, a Pauli Y error on data qubit 1 anticommutes with both the X and Z stabilizers that measure it, resulting in a flip of ancilla qubits 10 and 12. The result of this experiment for all possible stabilizer bit strings is shown in Figure 7.17. Since the high and low values of the stabilizers are clearly distinguishable, this demonstrates the ability to, on average, simultaneously identify arbitrary single qubit errors along both X and Z axes using the stabilizer

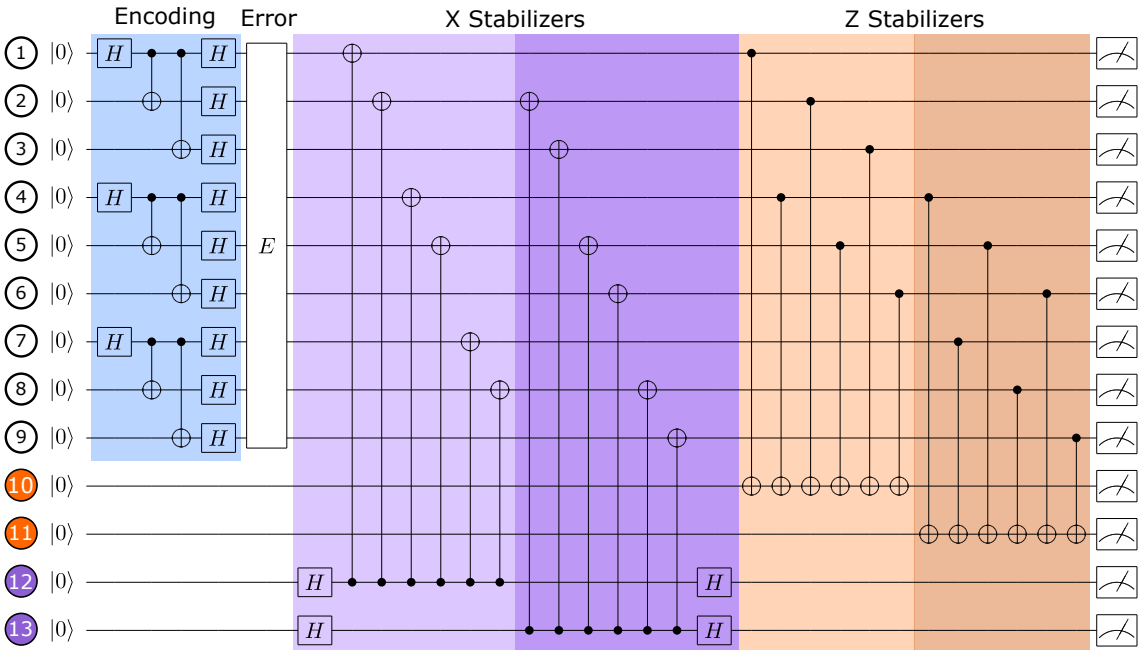

Figure 7.16: After encoding $|0\rangle_L$, different Pauli errors are purposely introduced on a selected data qubit in the code. To detect the error, each stabilizer eigenvalue is mapped onto the state of the corresponding ancilla qubit.

outcomes.

For a given artificial error, each stabilizer measurement yields a deterministic eigenvalue measurement (e.g., $\{+1, -1, -1, +1\}$) that is mapped to the ancilla qubit state (e.g., $\{0, 1, 1, 0\}$). Defining the error as the absolute value of difference between the expected ancilla bit string and the measured populations, and averaging across all the artificial errors, we obtain the following total error for each stabilizer measurement:

Stabilizer	**Total Error** (ϵ_{S_i})
$S_1 = Z_1Z_4Z_2Z_5Z_3Z_6$	0.244(3)
$S_2 = Z_4Z_7Z_5Z_8Z_6Z_9$	0.298(6)
$S_3 = X_1X_2X_4X_5X_7X_8$	0.179(3)
$S_4 = X_2X_3X_5X_6X_8X_9$	0.248(3)

Table 7.6: Total error rate for all the stabilizers in a full syndrome extraction experiment. This error includes encoding error, errors introduced from previous stabilizers and logical qubit T_2^* dephasing error.

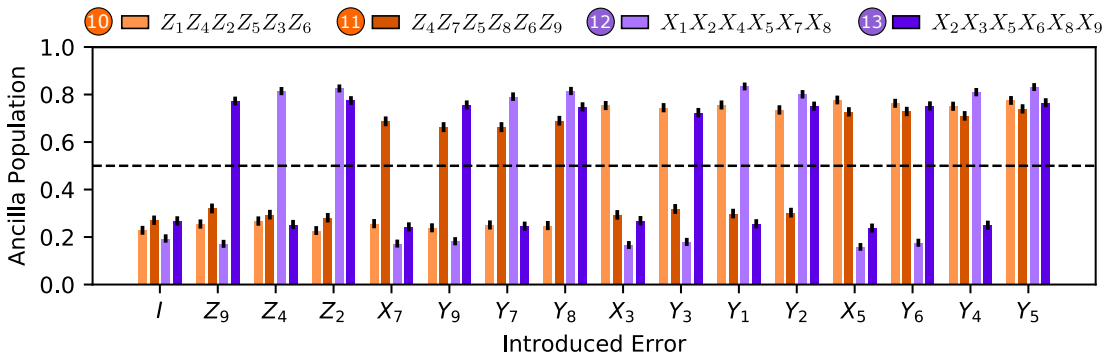

Figure 7.17: The data from running the circuit in Figure 7.16. The ideal ancilla population is 0/1 depending on whether an error *did not*/*did* anticommute with the stabilizer block. The colored bars correspond to the measured population of the different ancilla qubits.

In this experiment, the stabilizers are measured in the order S_3, S_4, S_1, S_2. The raw encoding of the $|0\rangle_L$ state has a base $\epsilon_{enc} = 0.038(2)$ error, which we assume is isotropic in the sense that all stabilizer measurements should see the error equally. Additionally, stabilizer measurements will detect errors introduced by itself or previous stabilizer measurements, which we assume to be isotropic as well. The per stabilizer error can be calculated by the differential error between successive stabilizer measurements. We calculate $\epsilon_Z = 0.064(7)$ per Z-stabilizer (avg. 98.9% gate fidelity) and $\epsilon_X = 0.069(5)$ per X-stabilizer (avg. 98.8% gate fidelity). Finally, we observe an error offset on the X-stabilizers relative to the Z-stabilizers of $\epsilon_{T_2^*} = 0.072(5)$, consistent with a Z-type error caused by the logical qubit dephasing (T_2^*) over the wall-clock time it takes to measure the X-stabilizers (≈ 3 ms). Then the total stabilizer measurement error for each ancilla qubit is well explained by the following error model:

$$\begin{aligned}
\epsilon_{S_1} &= \epsilon_{enc} + 2\epsilon_X + \epsilon_Z \\
\epsilon_{S_2} &= \epsilon_{enc} + 2\epsilon_X + 2\epsilon_Z \\
\epsilon_{S_3} &= \epsilon_{enc} + \epsilon_{T_2^*} + \epsilon_X \\
\epsilon_{S_4} &= \epsilon_{enc} + \epsilon_{T_2^*} + 2\epsilon_X
\end{aligned} \tag{7.8}$$

While the circuit shown in Fig. 7.16 is remarkably accurate given the complexity (30 two-qubit gates in total), the error rate is still too high to improve the logical qubit fidelity over even a single round of stabilizer readout and feedback. Comparing the stabilizer eigenvalues obtained from the ancilla qubits $S_{1,a}$ and $S_{2,a}$ (the data shown in Figure 7.17) to the stabilizer eigenvalues obtained from the data qubits at the end of the circuit $S_{1,d}$ and $S_{2,d}$ (the process described in Section 7.2), we observe a covariance of $\text{cov}(S_{1,a}, S_{1,d}) = 0.80$ and $\text{cov}(S_{2,a}, S_{2,d}) = 0.82$. A general rule of thumb is that the false-positive rate from the full syndrome extraction needs to be between 5-10% in order for the syndrome extraction to be beneficial. In other theoretical modeling work of the Bacon-Shor 13 code, the physical two-qubit gate needs to achieve $\approx 99.9\%$ fidelity in order to surpass the pseudo-threshold where the logical qubit outperforms the physical qubit [146].

Ultimately, the logical qubit state must be stabilized over many rounds of error-correction, which will involve mid-circuit detection of the ancilla qubits and conditional feedback. Intermediate measurement of a subset of qubits in the ion chain requires physically breaking the chain and shuttling idle qubits far away from the measured qubits in order to reduce decoherence due to resonantly scattered

369.5-nm photons during the detection process. After the measured qubits are reset, the separate chain segments must be re-merged into a single chain in order to resume with computations. This process is likely to heat the axial and/or radial modes, and so to maintain high-fidelity operations, the chain must be re-cooled. This can be accomplished using the sympathetic cooling scheme described in Section 6.2.

Chapter 8: Outlook

It is clear that ultimately trapped ion quantum computers will require sympathetic cooling. In the QCCD architecture, sympathetic cooling ions are required to re-cool multiple ions chains after they are shuffled and transported around a multi-zone ion trap. It also appears as though sympathetic cooling will be able to mitigate the dominant scaling limitation in a single long chain of ions, which is axial heating. While there are other scaling considerations (e.g., gate speed, computational complexity of entangling waveform, etc.), none scale as unfavorably (N^6) as the error due to axial heating. The picture that emerges is a maximum chain length where the trap-specific anomalous heating rate equals the sympathetic cooling rate. Practically, the cooling rate should be at least twice as fast as the heating rate, as gate operations still need to be performed.

Therefore, there is still strong motivation to reduce anomalous heating rates, as this will set the maximum chain length and/or minimize the requirements on sympathetic cooling (e.g., mixture ratio, cooling beam power, etc.). Further studies are also needed to characterize the sympathetic cooling rates in a chain and perform trade studies on the coolant mixture ratio and placement within a chain to maximize performance. One clear upgrade path is to operate the vacuum chamber at cryogenic

temperatures (~ 10 K), which has been shown to reduce anomalous heating [94] and will eliminate some of the complications with operating a room-temperature mixed-species ion chain, namely reordering of the chain due to background gas collisions.

There is also further work required to determine how two-qubit gate fidelities can reach $> 99.9\%$, which is roughly where quantum error-correction surpasses the threshold and can be used for further error suppression. While this fidelity has been achieved in systems with large beams that are limited to $N = 2$ chains [92, 158], it has not yet been realized with focused individual addressing beams, which support $N > 2$ chains. However, there appears to be no fundamental limitation to this approach. A recent demonstration using focused beams on a small chain has achieved 99.5% fidelity [93] and with further engineering it seems possible to achieve 99.9%.

With the integration of sympathetic cooling, reductions in anomalous heating rates, and further improvements to the two-qubit gate fidelity, it seems very possible to achieve high-fidelity operations on a chain of $N \approx 30$ ions. This is coincidentally the number of ions required to hold two logical qubits with the Bacon-Shor 13 code. Thus, one can imagine the QCCD architecture, but with logical qubits shuttled around the chip rather than physical qubits. Since quantum error-correction is expected to take up a significant portion of time in a fully fault-tolerant quantum algorithm, there is good reason to keep a single logical qubit well-localized with strong inter-connectivity, particularly for stabilizer readout. Of course, even with just two chains of $N \approx 30$ ions, we enter the regime where the quantum computer can no longer be simulated classically, and we hope to find exciting algorithms to

explore new physics or solve difficult numerical problems.

Appendix A: Error Bars and Fitting Quantum Data

In this appendix, we describe the fundamental statistics of quantum experiments, including calculation of error bars. Additionally, we describe how to fit data from quantum experiments that may consist of several individual quantum experiments. Finally we describe how to derive error bars from these fits.

A.1: The Binomial Distribution

The output of a quantum experiment on a single qubit is a Bernoulli Trial that succeeds (i.e., measures $|1\rangle$) with probability p. We wish to estimate $\hat{p}$ and obtain error bars on this estimate. To do this, we must repeat this Bernoulli Trial several times to obtain statistics. The probability of measuring k successes in a sequence of n Bernoulli trials is given by the Binomial Distribution, $k \sim B(n, p)$. Or in other words:

$$P(k|n,p) = \binom{n}{k} p^n (1-p)^{n-k} \tag{A.1}$$

One approach to estimating $\hat{p}$ is to use Maximum Likelihood Estimation (MLE) to maximize $P(p|n,k)$ over all possible $p \in [0,1]$. Thus:

$$
\begin{aligned}
\hat{p}_{MLE} &= \underset{p}{\text{argmax}}\, P(p|n,k) \\
&= \underset{p}{\text{argmax}}\left(\frac{P(k|n,p)P(p)}{P(k)}\right) && \text{From Bayes' Law} \\
&= \underset{p}{\text{argmax}}(P(k|n,p)) && \text{Assuming } p \sim U(0,1) \\
&= \underset{p}{\text{argmax}}(p^n(1-p)^{n-k}) && \text{Ignoring constants in } n,k \\
&= \underset{p}{\text{argmax}}(n\log p + (n-k)\log(1-p)) && \text{Log is monotonic increasing} \quad \text{(A.2)}
\end{aligned}
$$

Taking the derivative of the result of (A.2), setting it equal 0, and solving to p yields the familiar result:

$$\hat{p}_{MLE} = \frac{k}{n} \tag{A.3}$$

This fits with our naive intuition that we should just take the mean over the measurement output string (e.g., mean($[0,1,0,1,1,0,0,...]$)), but in this case it is also the MLE estimator!

A.2: Error Bars on $\hat{p}$

In general, we should expect that $\hat{p} \neq p$, since we can only ever estimate $\hat{p}$ through statistical means, which are inherently random. Thus we need to understand what the error bars on $\hat{p}$ are. It is important to note that the error bars are generally asymmetric, meaning there are different upper bounds than lower bounds. This asymmetry is magnified near the extreme of the interval $[0,1]$. One

easy way to see this is to imagine flipping a fair coin ($p = 0.5$) five times. There is a $(0.5)^5 = 0.03125$ probability that you see all heads. In this case, $\hat{p} = 1$ and the upper bound is 0 because probability cannot be larger than 1. One would expect that our lower bound includes the true value, $p = 0.5$.

There are many approximate solutions to find the error bounds on $\hat{p}$ including the Wilson score, Agresti-Coull, and Jeffreys interval. Here we will present the "exact" Clopper-Pearson solution that uses the underlying binomial probabiliy distribution. Note: "exact" is a bit of a misnomer, because the approximations work quite well, but here it means that we directly solve for the error bounds. The Clopper-Pearson interval contains all values of p not rejected by a given significance level α ($\alpha = 0.05$ for 95% confidence interval).

The lower limit (getting at least k successes) is given by:

$$\sum_{x=k}^{n} \binom{n}{x} {p_L}^x (1 - p_L)^{n-x} = \frac{\alpha}{2} \tag{A.4}$$

The upper limit (getting at most k successes) is given by:

$$\sum_{x=0}^{k} \binom{n}{x} {p_U}^x (1 - p_U)^{n-x} = \frac{\alpha}{2} \tag{A.5}$$

We wish to solve for p_L and p_U. The Beta Distribution has important links to the Binomial Distribution. In the Bayesian picture, it is known as the *conjugate*

prior. This means that prior on p has the same algebraic form as the posterior $P(k|p,n)$. Or, in other words, the Beta Distribution describes the probability distribution on p, which we have imperfect knowledge of, given our sampled data. This is an important property for iterative update schemes where you may want to adaptively sample from some Binomial Distribution and stop as soon as you hit a given confidence level. For now, we will just use the following identity:

$$\sum_{x=k}^{n} \binom{n}{x} p^x (1-p)^{n-x} = \int_0^p Beta(x; \alpha = k, \beta = n-k+1) dx \tag{A.6}$$

So our lower bound is just given by the CDF of the Beta Distribution, which we wish to invert to solve for p_L. The upper bound can be similarly solved for. The inverse of the CDF is called the percentage point function PPF and answers the question "what probability should I choose to contain x percent of the total population?". For example PPF(0.5) gives the median of any probability distribution.

We obtain the closed form solution for the bounds:

$$(p_L, p_U) = (Beta_{PPF}(\frac{\alpha}{2}; k, n-k+1), Beta_{PPF}(1-\frac{\alpha}{2}; k+1, n-k)) \tag{A.7}$$

The Beta PPF function is available in most software packages. In Python, one can use the `scipy.stats.beta.ppf(q,a,b)` function.

A.3: Fitting Functions on Binomial Data

The previous sections have only dealt with a single data point drawn from a Binomial Distribution with constant p, however in most quantum experiments we wish to fit a function to many such data points, each drawn from a different Binomial Distribution with a varying p_i value. The function we wish to fit usually only has a few degrees of freedom, so it not possible to maximize the MLE for each data point. Thus, we take the approach that we should maximize the total MLE for all of the data points.

We wish to fit a function $p_i = f(x_i|\theta)$ where θ describes the parameters of the model that we wish to optimize. The total likelihood of parameter choice θ is given by:

$$\mathcal{L}(\theta) = \prod_i P(f(x_i|\theta)|n_i, k_i) \tag{A.8}$$

Typically we like to use the log likelihood, since it converts products to sums and makes the optimization less susceptible to numerical rounding errors. Also, since most numerical optimization routines prefer minimization to maximization, we will use the negative log likelihood. Recalling from Section 1 that $P(p|n,k) \propto P(k|n,p) = B(k,n,p)$, we get:

$$\begin{aligned}\hat{\theta}_{MLE} &= \underset{\theta}{\text{argmax}}(\mathcal{L}(\theta)) \\ &= \underset{\theta}{\text{argmin}}(-\log \mathcal{L}(\theta)) \\ &= \underset{\theta}{\text{argmin}}(-\sum_i \log B(k_i, n_i, f(x_i|\theta)))\end{aligned}$$

This minmization can readily be solved using numerical optimizations. In Python, one can use `scipy.optimize.minimize(fun, x0)`.

A.4: Error Bars on $\hat{\theta}$

As before, we expect $\hat{\theta} \neq \theta$, and so we wish to obtain error bars on the fit parameters. For the MLE estimator (unbiased, efficient) the variance is given by the Cramer-Rao Bound:

$$\text{var}(\hat{\theta}) = \frac{1}{I(\hat{\theta})} \tag{A.9}$$

Where $I(\hat{\theta})$ is the Fisher information matrix. For a model with only a single parameter, the Fisher information is just the second derivative of the negative log-likelihood at the optimal point $\hat{\theta}$

$$I(\hat{\theta}) = \left.\frac{d^2(-\log \mathcal{L}(\theta))}{d\theta^2}\right|_{\hat{\theta}} \tag{A.10}$$

For models with more than one parameter, one should calculate the Hessian,

$H(\hat{\theta})$ and take the diagonals to get:

$$\sigma_{\hat{\theta}} = \frac{1}{\sqrt{\text{diag}(H(\hat{\theta}))}} \tag{A.11}$$

As a practical note, if using a quasi-Netwonian solver (e.g., L-BFGS-B) for the MLE maximization in Section 3, one should not trust the Hessian that is (often) returned as a result of the optimization. It is often a poor approximation and is conditional on the path that the optimizer took (i.e., it can depend on the initial guess of the parameters). It is better to independently approximate the Hessian at $\hat{\theta}$ using a finite differences method. In Python, `numdifftools.Hessian(f)` does exactly this.

Bibliography

[1] Richard P Feynman. Quantum mechanical computers. *Foundations of Physics*, 16(6):507–531, 1986.

[2] Daniel S Abrams and Seth Lloyd. Simulation of many-body fermi systems on a universal quantum computer. *Physical Review Letters*, 79(13):2586, 1997.

[3] Alán Aspuru-Guzik, Anthony D Dutoi, Peter J Love, and Martin Head-Gordon. Simulated quantum computation of molecular energies. *Science*, 309(5741):1704–1707, 2005.

[4] Markus Reiher, Nathan Wiebe, Krysta M Svore, Dave Wecker, and Matthias Troyer. Elucidating reaction mechanisms on quantum computers. *Proceedings of the National Academy of Sciences*, 114(29):7555–7560, 2017.

[5] Peter W Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Review*, 41(2):303–332, 1999.

[6] John Preskill. Quantum computing in the nisq era and beyond. *Quantum*, 2:79, 2018.

[7] Ethan Bernstein and Umesh Vazirani. Quantum complexity theory. *SIAM Journal on computing*, 26(5):1411–1473, 1997.

[8] Albert Einstein, Boris Podolsky, and Nathan Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical review*, 47(10):777, 1935.

[9] John S Bell. On the einstein podolsky rosen paradox. *Physics Physique Fizika*, 1(3):195, 1964.

[10] Michel H Devoret and Robert J Schoelkopf. Superconducting circuits for quantum information: an outlook. *Science*, 339(6124):1169–1174, 2013.

[11] Mark Saffman. Quantum computing with neutral atoms. *National Science Review*, 6(1):24–25, 2019.

[12] CE Bradley, J Randall, MH Abobeih, RC Berrevoets, MJ Degen, MA Bakker, M Markham, DJ Twitchen, and TH Taminiau. A ten-qubit solid-state spin register with quantum memory up to one minute. *Physical Review X*, 9(3):031045, 2019.

[13] Pieter Kok, William J Munro, Kae Nemoto, Timothy C Ralph, Jonathan P Dowling, and Gerard J Milburn. Linear optical quantum computing with photonic qubits. *Reviews of modern physics*, 79(1):135, 2007.

[14] Chetan Nayak, Steven H Simon, Ady Stern, Michael Freedman, and Sankar Das Sarma. Non-abelian anyons and topological quantum computation. *Reviews of Modern Physics*, 80(3):1083, 2008.

[15] David P DiVincenzo. The physical implementation of quantum computation. *Fortschritte der Physik: Progress of Physics*, 48(9-11):771–783, 2000.

[16] David Kielpinski, Chris Monroe, and David J Wineland. Architecture for a large-scale ion-trap quantum computer. *Nature*, 417(6890):709–711, 2002.

[17] Murray D Barrett, B DeMarco, T Schaetz, V Meyer, D Leibfried, J Britton, J Chiaverini, WM Itano, B Jelenković, JD Jost, et al. Sympathetic cooling of $^{9}Be^{+}$ and $^{24}Mg^{+}$ for quantum logic. *Physical Review A*, 68(4):042302, 2003.

[18] Kenneth R Brown, Jungsang Kim, and Christopher Monroe. Co-designing a scalable quantum computer with trapped atomic ions. *npj Quantum Information*, 2(1):1–10, 2016.

[19] Juan M Pino, Jennifer M Dreiling, Caroline Figgatt, John P Gaebler, Steven A Moses, CH Baldwin, M Foss-Feig, D Hayes, K Mayer, C Ryan-Anderson, et al. Demonstration of the trapped-ion quantum CCD computer architecture. *Nature*, 529:209–213, 2021.

[20] Shantanu Debnath, Norbert M Linke, Caroline Figgatt, Kevin A Landsman, Kevin Wright, and Christopher Monroe. Demonstration of a small programmable quantum computer with atomic qubits. *Nature*, 536(7614):63–66, 2016.

[21] Pengfei Wang, Chun-Yang Luan, Mu Qiao, Mark Um, Junhua Zhang, Ye Wang, Xiao Yuan, Mile Gu, Jingning Zhang, and Kihwan Kim. Single ion-qubit exceeding one hour coherence time. *Nature Communications*, 12:233, 2021.

[22] Justin E Christensen, David Hucul, Wesley C Campbell, and Eric R Hudson. High-fidelity manipulation of a qubit enabled by a manufactured nucleus. *npj Quantum Information*, 6(1):1–5, 2020.

[23] Wolfgang Paul and Helmut Steinwedel. Ein neues massenspektrometer ohne magnetfeld. *Zeitschrift für Naturforschung A*, 8(7):448–450, 1953.

[24] Wolfgang Paul. Electromagnetic traps for charged and neutral particles. *Reviews of modern physics*, 62(3):531, 1990.

[25] FM Penning. Ein neues manometer für niedrige gasdrucke, insbesondere zwischen l0- 3 und 10- 5 mm. *Physica*, 4(2):71–75, 1937.

[26] Hans Dehmelt. Experiments with an isolated subatomic particle at rest. *Reviews of modern physics*, 62(3):525, 1990.

[27] R Bradford Blakestad. *Transport of trapped-ion qubits within a scalable quantum processor*. PhD thesis, University of Colorado, Boulder, 2010.

[28] DJ Berkeland, JD Miller, James C Bergquist, Wayne M Itano, and David J Wineland. Minimization of ion micromotion in a paul trap. *Journal of applied physics*, 83(10):5025–5033, 1998.

[29] Dan Stick, WK Hensinger, Steven Olmschenk, MJ Madsen, Keith Schwab, and Chris Monroe. Ion trap in a semiconductor chip. *Nature Physics*, 2(1):36–39, 2006.

[30] Simon Ragg, Chiara Decaroli, Thomas Lutz, and Jonathan P Home. Segmented ion-trap fabrication using high precision stacked wafers. *Review of Scientific Instruments*, 90(10):103203, 2019.

[31] Robert J Niffenegger, Jules Stuart, Cheryl Sorace-Agaskar, Dave Kharas, Suraj Bramhavar, Colin D Bruzewicz, William Loh, Ryan T Maxson, Robert McConnell, David Reens, et al. Integrated multi-wavelength control of an ion qubit. *Nature*, 586(7830):538–542, 2020.

[32] J Stuart, R Panock, CD Bruzewicz, JA Sedlacek, R McConnell, IL Chuang, JM Sage, and J Chiaverini. Chip-integrated voltage sources for control of trapped ions. *Physical Review Applied*, 11(2):024010, 2019.

[33] Daniel H Slichter, Varun B Verma, D Leibfried, Richard P Mirin, Sae Woo Nam, and David J Wineland. UV-sensitive superconducting nanowire single photon detectors for integration in an ion trap. *Optics express*, 25(8):8705–8720, 2017.

[34] Peter Lukas Wilhelm Maunz. High optical access trap 2.0., 1 2016.

[35] Andrew M Steane. The ion trap quantum information processor. *arXiv preprint quant-ph/9608011*, 1996.

[36] Marko Cetina et al. *Hybrid approaches to quantum information using ions, atoms and photons*. PhD thesis, Massachusetts Institute of Technology, 2011.

[37] Grace H Zhang, Boris Braverman, Akio Kawasaki, and Vladan Vuletić. Note: Fast compact laser shutter using a direct current motor and three-dimensional printing. *Review of Scientific Instruments*, 86(12):126105, 2015.

[38] Harold J Metcalf and Peter van der Straten. Laser cooling and trapping of atoms. *JOSA B*, 20(5):887–908, 2003.

[39] DJ Berkeland and MG Boshier. Destabilization of dark states and optical spectroscopy in zeeman-degenerate atomic systems. *Physical Review A*, 65(3):033413, 2002.

[40] Steve Olmschenk, Kelly C Younge, David L Moehring, Dzmitry N Matsukevich, Peter Maunz, and Christopher Monroe. Manipulation and detection of a trapped Yb^+ hyperfine qubit. *Physical Review A*, 76(5):052314, 2007.

[41] Rachel Noek, Geert Vrijsen, Daniel Gaultney, Emily Mount, Taehyun Kim, Peter Maunz, and Jungsang Kim. High speed, high fidelity detection of an atomic hyperfine qubit. *Optics letters*, 38(22):4735–4738, 2013.

[42] AH Myerson, DJ Szwer, SC Webster, DTC Allcock, MJ Curtis, G Imreh, JA Sherman, DN Stacey, AM Steane, and DM Lucas. High-fidelity readout of trapped-ion qubits. *Physical Review Letters*, 100(20):200502, 2008.

[43] SL Todaro, VB Verma, KC McCormick, DTC Allcock, RP Mirin, DJ Wineland, SW Nam, AC Wilson, D Leibfried, and DH Slichter. State readout of a trapped ion qubit using a trap-integrated superconducting photon detector. *Physical Review Letters*, 126(1):010501, 2021.

[44] Stephen Crain, Clinton Cahall, Geert Vrijsen, Emma E Wollman, Matthew D Shaw, Varun B Verma, Sae Woo Nam, and Jungsang Kim. High-speed low-crosstalk detection of a $^{171}Yb^+$ qubit using superconducting nanowire single photon detectors. *Communications Physics*, 2(1):1–6, 2019.

[45] CL Edmunds, TR Tan, AR Milne, A Singh, MJ Biercuk, and C Hempel. Scalable hyperfine qubit state detection via electron shelving in the $^2D_{5/2}$ and $^2F_{7/2}$ manifolds in $^{171}Yb^+$. *arXiv preprint arXiv:2012.14606*, 2020.

[46] Peter TH Fisk, Matthew J Sellars, Malcolm A Lawn, and G Coles. Accurate measurement of the 12.6 GHz "clock" transition in trapped $^{171}Yb^+$ ions. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 44(2):344–354, 1997.

[47] Terry Gullion, David B Baker, and Mark S Conradi. New, compensated carr-purcell sequences. *Journal of Magnetic Resonance (1969)*, 89(3):479–484, 1990.

[48] David J Wineland, M Barrett, J Britton, J Chiaverini, B DeMarco, Wayne M Itano, B Jelenković, Christopher Langer, D Leibfried, V Meyer, et al. Quantum information processing with trapped ions. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 361(1808):1349–1361, 2003.

[49] R Ozeri, Wayne M Itano, RB Blakestad, J Britton, J Chiaverini, John D Jost, C Langer, D Leibfried, Rainer Reichle, Signe Seidelin, et al. Errors in trapped-ion quantum gates due to spontaneous photon scattering. *Physical Review A*, 75(4):042329, 2007.

[50] Aaron C Lee, Jacob Smith, Philip Richerme, Brian Neyenhuis, Paul W Hess, Jiehang Zhang, and Christopher Monroe. Engineering large stark shifts for control of individual clock state qubits. *Physical Review A*, 94(4):042308, 2016.

[51] WC Campbell, J Mizrahi, Q Quraishi, C Senko, D Hayes, D Hucul, DN Matsukevich, P Maunz, and C Monroe. Ultrafast gates for single atomic qubits. *Physical review letters*, 105(9):090502, 2010.

[52] David Hayes, Dzmitry N Matsukevich, Peter Maunz, David Hucul, Qudsia Quraishi, Steve Olmschenk, W Campbell, Jonathan Mizrahi, Crystal Senko, and Christopher Monroe. Entanglement of atomic qubits using an optical frequency comb. *Physical review letters*, 104(14):140501, 2010.

[53] J Mizrahi, B Neyenhuis, KG Johnson, WC Campbell, C Senko, D Hayes, and C Monroe. Quantum control of qubits and atomic motion using ultrafast laser pulses. *Applied Physics B*, 114(1):45–61, 2014.

[54] A.E. Siegman. *Lasers*. University Science Books, 1986.

[55] Kenneth R Brown, Aram W Harrow, and Isaac L Chuang. Arbitrarily accurate composite pulse sequences. *Physical Review A*, 70(5):052318, 2004.

[56] Stephen Wimperis. Broadband, narrowband, and passband composite pulses for use in advanced NMR experiments. *Journal of Magnetic Resonance, Series A*, 109(2):221–231, 1994.

[57] Juan I Cirac and Peter Zoller. Quantum computations with cold trapped ions. *Physical review letters*, 74(20):4091, 1995.

[58] Patricia J Lee, Kathy-Anne Brickman, Louis Deslauriers, Paul C Haljan, Lu-Ming Duan, and Christopher Monroe. Phase control of trapped ion quantum gates. *Journal of Optics B: Quantum and Semiclassical Optics*, 7(10):S371, 2005.

[59] Caroline Figgatt, Aaron Ostrander, Norbert M Linke, Kevin A Landsman, Daiwei Zhu, Dmitri Maslov, and Christopher Monroe. Parallel entangling operations on a universal ion-trap quantum computer. *Nature*, 572(7769):368–372, 2019.

[60] Nikodem Grzesiak, Reinhold Blümel, Kenneth Wright, Kristin M Beck, Neal C Pisenti, Ming Li, Vandiver Chaplin, Jason M Amini, Shantanu Debnath, Jwo-Sy Chen, et al. Efficient arbitrary simultaneously entangling gates on a trapped-ion quantum computer. *Nature communications*, 11(1):1–6, 2020.

[61] Klaus Mølmer and Anders Sørensen. Multiparticle Entanglement of Hot Trapped Ions. *Physical Review Letters*, 82(9):1835–1838, March 1999.

[62] Anders Sørensen and Klaus Mølmer. Quantum computation with ions in thermal motion. *Physical review letters*, 82(9):1971, 1999.

[63] Anders Sørensen and Klaus Mølmer. Entanglement and quantum computation with ions in thermal motion. *Physical Review A*, 62(2):022311, 2000.

[64] R Islam, EE Edwards, K Kim, S Korenblit, C Noh, H Carmichael, G-D Lin, L-M Duan, C-C Joseph Wang, JK Freericks, et al. Onset of a quantum phase transition with a trapped ion quantum simulator. *Nature communications*, 2(1):1–6, 2011.

[65] Yukai Wu, Sheng-Tao Wang, and L-M Duan. Noise analysis for high-fidelity quantum entangling gates in an anharmonic linear paul trap. *Physical Review A*, 97(6):062325, 2018.

[66] Shi-Liang Zhu and ZD Wang. Unconventional geometric quantum computation. *Physical review letters*, 91(18):187902, 2003.

[67] Dmitri Maslov. Basic circuit compilation techniques for an ion-trap quantum machine. *New Journal of Physics*, 19(2):023035, 2017.

[68] Taeyoung Choi, Shantanu Debnath, TA Manning, Caroline Figgatt, Z-X Gong, L-M Duan, and Christopher Monroe. Optimal quantum control of multimode couplings between trapped ion qubits for scalable entanglement. *Physical review letters*, 112(19):190502, 2014.

[69] Pak Hong Leung, Kevin A Landsman, Caroline Figgatt, Norbert M Linke, Christopher Monroe, and Kenneth R Brown. Robust 2-qubit gates in a linear ion crystal using a frequency-modulated driving force. *Physical review letters*, 120(2):020501, 2018.

[70] Reinhold Blumel, Nikodem Grzesiak, and Yunseong Nam. Power-optimal, stabilized entangling gate between trapped-ion qubits. *arXiv preprint arXiv:1905.09292*, 2019.

[71] Alistair R Milne, Claire L Edmunds, Cornelius Hempel, Federico Roy, Sandeep Mavadia, and Michael J Biercuk. Phase-modulated entangling gates robust to static and time-varying errors. *Physical Review Applied*, 13(2):024022, 2020.

[72] Jake Lishman and Florian Mintert. Trapped-ion entangling gates robust against qubit frequency errors. *Physical Review Research*, 2(3):033117, 2020.

[73] Yotam Shapira, Ravid Shaniv, Tom Manovitz, Nitzan Akerman, and Roee Ozeri. Robust entanglement gates for trapped-ion qubits. *Physical review letters*, 121(18):180502, 2018.

[74] Christopher DB Bentley, Harrison Ball, Michael J Biercuk, Andre RR Carvalho, Michael R Hush, and Harry J Slatyer. Numeric optimization for configurable, parallel, error-robust entangling gates in large ion registers. *Advanced Quantum Technologies*, 3(11):2000044, 2020.

[75] Kevin A Landsman, Yukai Wu, Pak Hong Leung, Daiwei Zhu, Norbert M Linke, Kenneth R Brown, Luming Duan, and C Monroe. Two-qubit entangling gates within arbitrarily long chains of trapped ions. *Physical Review A*, 100(2):022332, 2019.

[76] Pak Hong Leung and Kenneth R Brown. Entangling an arbitrary pair of qubits in a long ion crystal. *Physical Review A*, 98(3):032318, 2018.

[77] Shantanu Debnath. *A programmable five qubit quantum computer using trapped atomic ions*. PhD thesis, University of Maryland, College Park, 2016.

[78] JD Siverns, LR Simkins, S Weidt, and WK Hensinger. On the application of radio frequency voltages to ion traps via helical resonators. *Applied Physics B*, 107(4):921–934, 2012.

[79] David R Leibrandt and Jason Heidecker. An open source digital servo for amo physics experiments. *arXiv preprint arXiv:1508.06319*, 2015.

[80] RWP Drever, John L Hall, FV Kowalski, J_ Hough, GM Ford, AJ Munley, and H Ward. Laser phase and frequency stabilization using an optical resonator. *Applied Physics B*, 31(2):97–105, 1983.

[81] Eric D Black. An introduction to pound–drever–hall laser frequency stabilization. *American journal of physics*, 69(1):79–87, 2001.

[82] DJ McCarron, SA King, and SL Cornish. Modulation transfer spectroscopy in atomic rubidium. *Measurement science and technology*, 19(10):105601, 2008.

[83] IK Ilev. Fiber-optic autocollimation refractometer. *Optics communications*, 119(5-6):513–516, 1995.

[84] R Islam, WC Campbell, T Choi, SM Clark, CWS Conover, S Debnath, EE Edwards, B Fields, D Hayes, D Hucul, et al. Beat note stabilization of mode-locked lasers for quantum information processing. *Optics letters*, 39(11):3238–3241, 2014.

[85] Travis C Briles, Dylan C Yost, Arman Cingöz, Jun Ye, and Thomas R Schibli. Simple piezoelectric-actuated mirror with 180 khz servo bandwidth. *Optics express*, 18(10):9739–9746, 2010.

[86] Edward Farhi and Aram W Harrow. Quantum supremacy through the quantum approximate optimization algorithm. *arXiv preprint arXiv:1602.07674*, 2016.

[87] Robin Blume-Kohout, John King Gamble, Erik Nielsen, Jonathan Mizrahi, Jonathan D Sterk, and Peter Maunz. Robust, self-consistent, closed-form tomography of quantum logic gates on a trapped ion qubit. *arXiv preprint arXiv:1310.4492*, 2013.

[88] Dl Wineland and Hans Dehmelt. Proposed 10^{14} $\delta\nu/\nu$ laser fluorescence spectroscopy on Ti^{+} mono-ion oscillator III. In *Bulletin of the American Physical Society*, volume 20, pages 637–637, 1975.

[89] Emanuel Knill, Dietrich Leibfried, Rolf Reichle, Joe Britton, R Brad Blakestad, John D Jost, Chris Langer, Roee Ozeri, Signe Seidelin, and David J Wineland. Randomized benchmarking of quantum gates. *Physical Review A*, 77(1):012307, 2008.

[90] R Barends, J Kelly, A Megrant, A Veitia, D Sank, E Jeffrey, T C White, J Mutus, A G Fowler, B Campbell, Y Chen, Z Chen, B Chiaro, A Dunsworth, C Neill, P O'Malley, P Roushan, A Vainsencher, J Wenner, A N Korotkov, A N Cleland, and John M Martinis. Superconducting quantum circuits at the surface code threshold for fault tolerance. *Nature*, 508(7497):500–503, 2014.

[91] Cass A Sackett, David Kielpinski, Brian E King, Christopher Langer, Volker Meyer, Christopher J Myatt, M Rowe, QA Turchette, Wayne M Itano, David J Wineland, et al. Experimental entanglement of four particles. *Nature*, 404(6775):256–259, 2000.

[92] Christopher J Ballance. *High-fidelity quantum logic in Ca+*. PhD thesis, University of Oxford, 2014.

[93] Ye Wang, Stephen Crain, Chao Fang, Bichen Zhang, Shilin Huang, Qiyao Liang, Pak Hong Leung, Kenneth R Brown, and Jungsang Kim. High-fidelity two-qubit gates using a microelectromechanical-system-based beam steering system for individual qubit addressing. *Physical Review Letters*, 125(15):150505, 2020.

[94] M Brownnutt, M Kumph, P Rabl, and R Blatt. Ion-trap measurements of electric-field noise near surfaces. *Reviews of modern Physics*, 87(4):1419, 2015.

[95] Xin-Chuan Wu, Dripto M Debroy, Yongshan Ding, Jonathan M Baker, Yuri Alexeev, Kenneth R Brown, and Frederic T Chong. Tilt: Achieving higher fidelity on a trapped-ion linear-tape quantum computing architecture. *arXiv preprint arXiv:2010.15876*, 2020.

[96] M Cetina, LN Egan, CA Noel, ML Goldman, AR Risinger, D Zhu, D Biswas, and C Monroe. Quantum gates on individually-addressed atomic qubits subject to noisy transverse motion. *arXiv preprint arXiv:2007.06768*, 2020.

[97] Shi-Liang Zhu, Chris Monroe, and L-M Duan. Trapped ion quantum computation with transverse phonon modes. *Physical review letters*, 97(5):050505, 2006.

[98] Quentin A Turchette, BE King, D Leibfried, DM Meekhof, CJ Myatt, MA Rowe, CA Sackett, CS Wood, WM Itano, C Monroe, et al. Heating of trapped ions from the quantum ground state. *Physical Review A*, 61(6):063418, 2000.

[99] G-D Lin, S-L Zhu, Rajibul Islam, Kihwan Kim, M-S Chang, Simcha Korenblit, Christopher Monroe, and L-M Duan. Large-scale quantum computation in an anharmonic linear ion trap. *EPL (Europhysics Letters)*, 86(6):60004, 2009.

[100] Rainer Blatt and David Wineland. Entangled states of trapped atomic ions. *Nature*, 453(7198):1008–1015, 2008.

[101] Dietrich Leibfried, Rainer Blatt, Christopher Monroe, and David Wineland. Quantum dynamics of single trapped ions. *Reviews of Modern Physics*, 75(1):281, 2003.

[102] Dustin A Hite, Yves Colombe, Andrew C Wilson, Kenton R Brown, U Warring, R Jördens, John D Jost, KS McKay, DP Pappas, D Leibfried, et al. 100-fold reduction of electric-field noise in an ion trap cleaned with in situ argon-ion-beam bombardment. *Physical review letters*, 109(10):103001, 2012.

[103] JP Schiffer. Phase transitions in anisotropically confined ionic crystals. *Physical review letters*, 70(6):818, 1993.

[104] Adam D West, Randall Putnam, Wesley C Campbell, and Paul Hamilton. Tunable transverse spin–motion coupling for quantum information processing. *Quantum Science and Technology*, 6(2):024003, 2021.

[105] DJ Larson, James C Bergquist, John J Bollinger, Wayne M Itano, and David J Wineland. Sympathetic cooling of trapped ions: A laser-cooled two-species nonneutral ion plasma. *Physical review letters*, 57(1):70, 1986.

[106] Chin-wen Chou, DB Hume, JCJ Koelemeij, David J Wineland, and T Rosenband. Frequency comparison of two high-accuracy Al^+ optical clocks. *Physical review letters*, 104(7):070802, 2010.

[107] Ksenia Sosnova, Allison Carter, and Christopher Monroe. Character of motional modes for entanglement and sympathetic cooling of mixed-species trapped-ion chains. *Physical Review A*, 103(1):012610, 2021.

[108] Melissa Revelle, Craig William Hogle, Brandon Ruzic, Peter Lukas Wilhelm Maunz, Kevin Young, and Daniel Lobser. Demonstration of sideband cooling on the $^{171}Yb^+$ quadrupole transition. *Sandia National Laboratories Report No. SAND2019-0668C671715*, 1 2019.

[109] Christian Roos. *Controlling the quantum state of trapped ions*. PhD thesis, University of Innsbruck, 2 2000.

[110] James J McLoughlin, Altaf H Nizamani, James D Siverns, Robin C Sterling, Marcus D Hughes, Bjoern Lekitsch, Björn Stein, Seb Weidt, and Winfried K Hensinger. Versatile ytterbium ion trap experiment for operation of scalable ion-trap chips with motional heating and transition-frequency measurements. *Physical Review A*, 83(1):013406, 2011.

[111] Hendrik M Meyer, Matthias Steiner, Lothar Ratschbacher, Christoph Zipkes, and Michael Köhl. Laser spectroscopy and cooling of Yb^+ ions on a deep-uv transition. *Physical Review A*, 85(1):012502, 2012.

[112] Ian Counts, Joonseok Hur, Diana PL Aude Craik, Honggi Jeon, Calvin Leung, Julian C Berengut, Amy Geddes, Akio Kawasaki, Wonho Jhe, and Vladan Vuletić. Evidence for nonlinear isotope shift in Yb^+ search for new boson. *Physical Review Letters*, 125(12):123002, 2020.

[113] N Yu and L Maleki. Lifetime measurements of the $4f^{14}5d$ metastable states in single ytterbium ions. *Physical Review A*, 61(2):022507, 2000.

[114] RW Berends, EH Pinnington, B Guo, and Q Ji. Beam-laser lifetime measurements for four resonance levels of Yb II. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 26(20):L701, 1993.

[115] Michaela Kleinert, ME Gold Dahl, and Scott Bergeson. Measurement of the Yb I $^1S_0 - ^1P_1$ transition frequency at 399 nm using an optical frequency comb. *Physical Review A*, 94(5):052511, 2016.

[116] A Guttridge, SA Hopkins, SL Kemp, D Boddy, R Freytag, MPA Jones, MR Tarbutt, EA Hinds, and SL Cornish. Direct loading of a large Yb MOT on the transition. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 49(14):145006, 2016.

[117] Michael W Lee, Marie Claire Jarratt, Christian Marciniak, and Michael J Biercuk. Frequency stabilization of a 369 nm diode laser by nonlinear spectroscopy of ytterbium ions in a discharge. *Optics express*, 22(6):7210–7221, 2014.

[118] DMS Johnson, JM Hogan, S-W Chiow, and MA Kasevich. Broadband optical serrodyne frequency shifting. *Optics letters*, 35(5):745–747, 2010.

[119] Craig Gidney and Martin Ekerå. How to factor 2048 bit rsa integers in 8 hours using 20 million noisy qubits. Preprint at https://arxiv.org/abs/1905.09749, 2019.

[120] Daniel Eric Gottesman. *Stabilizer Codes and Quantum Error Correction*. PhD thesis, California Institute of Technology, 1997.

[121] Peter W Shor. Scheme for reducing decoherence in quantum computer memory. *Physical Review A*, 52(4):R2493, 1995.

[122] Emanuel Knill and Raymond Laflamme. Theory of quantum error-correcting codes. *Physical Review A*, 55(2):900, 1997.

[123] Dorit Aharonov and Michael Ben-Or. Fault-tolerant quantum computation with constant error rate. *SIAM Journal on Computing*, 2008.

[124] Emanuel Knill, Raymond Laflamme, and W Zurek. Threshold accuracy for quantum computation. Preprint at https://arxiv.org/abs/quant-ph/9610011, 1996.

[125] Antonio D Córcoles, Easwar Magesan, Srikanth J Srinivasan, Andrew W Cross, Matthias Steffen, Jay M Gambetta, and Jerry M Chow. Demonstration of a quantum error detection code using a square lattice of four superconducting qubits. *Nature Communications*, 6(1):1–10, 2015.

[126] Maika Takita, Andrew W Cross, AD Córcoles, Jerry M Chow, and Jay M Gambetta. Experimental demonstration of fault-tolerant state preparation with superconducting qubits. *Physical Review Letters*, 119(18):180501, 2017.

[127] Norbert M Linke, Mauricio Gutierrez, Kevin A Landsman, Caroline Figgatt, Shantanu Debnath, Kenneth R Brown, and Christopher Monroe. Fault-tolerant quantum error detection. *Science Advances*, 3(10):e1701074, 2017.

[128] Robin Harper and Steven T Flammia. Fault-tolerant logical gates in the ibm quantum experience. *Physical Review Letters*, 122(8):080504, 2019.

[129] Christian Kraglund Andersen, Ants Remm, Stefania Lazar, Sebastian Krinner, Nathan Lacroix, Graham J Norris, Mihai Gabureac, Christopher Eichler, and Andreas Wallraff. Repeated quantum error detection in a surface code. *Nature Physics*, pages 1–6, 2020.

[130] David G Cory, MD Price, W Maas, Emanuel Knill, Raymond Laflamme, Wojciech H Zurek, Timothy F Havel, and Shyamal S Somaroo. Experimental quantum error correction. *Physical Review Letters*, 81(10):2152, 1998.

[131] John Chiaverini, Dietrich Leibfried, Tobias Schaetz, Murray D Barrett, RB Blakestad, J Britton, Wayne M Itano, Juergen D Jost, Emanuel Knill, Christopher Langer, et al. Realization of quantum error correction. *Nature*, 432(7017):602–605, 2004.

[132] Philipp Schindler, Julio T Barreiro, Thomas Monz, Volckmar Nebendahl, Daniel Nigg, Michael Chwalla, Markus Hennrich, and Rainer Blatt. Experimental repetitive quantum error correction. *Science*, 332(6033):1059–1061, 2011.

[133] Matthew D Reed, Leonardo DiCarlo, Simon E Nigg, Luyan Sun, Luigi Frunzio, Steven M Girvin, and Robert J Schoelkopf. Realization of three-qubit quantum error correction with superconducting circuits. *Nature*, 482(7385):382–385, 2012.

[134] Diego Riste, Stefano Poletto, M-Z Huang, Alessandro Bruno, Visa Vesterinen, O-P Saira, and Leonardo DiCarlo. Detecting bit-flip errors in a logical qubit using stabilizer measurements. *Nature Communications*, 6(1):1–6, 2015.

[135] Julian Kelly, Rami Barends, Austin G Fowler, Anthony Megrant, Evan Jeffrey, Theodore C White, Daniel Sank, Josh Y Mutus, Brooks Campbell, Yu Chen, et al. State preservation by repetitive error detection in a superconducting quantum circuit. *Nature*, 519(7541):66–69, 2015.

[136] Ming Gong, Xiao Yuan, Shiyu Wang, Yulin Wu, Youwei Zhao, Chen Zha, Shaowei Li, Zhen Zhang, Qi Zhao, Yunchao Liu, et al. Experimental verification of five-qubit quantum error correction with superconducting qubits. Preprint at https://arxiv.org/abs/1907.04507, 2019.

[137] Daniel Nigg, Markus Mueller, Esteban A Martinez, Philipp Schindler, Markus Hennrich, Thomas Monz, Miguel A Martin-Delgado, and Rainer Blatt. Quantum computations on a topologically encoded qubit. *Science*, 345(6194):302–305, 2014.

[138] Yi-Han Luo, Ming-Cheng Chen, Manuel Erhard, Han-Sen Zhong, Dian Wu, Hao-Yang Tang, Qi Zhao, Xi-Lin Wang, Keisuke Fujii, Li Li, Nai-Le Liu, Kae Nemoto, William J. Munro, Chao-Yang Lu, Anton Zeilinger, and Jian-Wei Pan. Quantum teleportation of physical qubits into logical code-spaces. Preprint at https://arxiv.org/abs/2009.06242, 2020.

[139] Reinier W Heeres, Philip Reinhold, Nissim Ofek, Luigi Frunzio, Liang Jiang, Michel H Devoret, and Robert J Schoelkopf. Implementing a universal gate set on a logical qubit encoded in an oscillator. *Nature communications*, 8(1):1–7, 2017.

[140] Christa Flühmann, Thanh Long Nguyen, Matteo Marinelli, Vlad Negnevitsky, Karan Mehta, and Jonathan P Home. Encoding a qubit in a trapped-ion mechanical oscillator. *Nature*, 566(7745):513–517, 2019.

[141] Nissim Ofek, Andrei Petrenko, Reinier Heeres, Philip Reinhold, Zaki Leghtas, Brian Vlastakis, Yehan Liu, Luigi Frunzio, SM Girvin, Liang Jiang, et al. Extending the lifetime of a quantum bit with error correction in superconducting circuits. *Nature*, 536(7617):441–445, 2016.

[142] P Campagne-Ibarcq, A Eickbusch, S Touzard, E Zalys-Geller, NE Frattini, VV Sivak, P Reinhold, S Puri, S Shankar, RJ Schoelkopf, et al. Quantum error correction of a qubit encoded in grid states of an oscillator. *Nature*, 584(7821):368–372, 2020.

[143] Laird Egan, Dripto M Debroy, Crystal Noel, Andrew Risinger, Daiwei Zhu, Debopriyo Biswas, Michael Newman, Muyuan Li, Kenneth R Brown, Marko Cetina, et al. Fault-tolerant operation of a quantum error-correction code. *arXiv preprint arXiv:2009.11482*, 2020.

[144] Dave Bacon. Operator quantum error-correcting subsystems for self-correcting quantum memories. *Physical Review A*, 73(1):012340, 2006.

[145] Panos Aliferis and Andrew W Cross. Subsystem fault tolerance with the Bacon-Shor code. *Phys. Rev. Lett.*, 98(22):220502, 2007.

[146] Dripto M Debroy, Muyuan Li, Shilin Huang, and Kenneth R Brown. Logical performance of 9 qubit compass codes in ion traps with crosstalk errors. *Quantum Science and Technology*, 5(3):034002, 2020.

[147] Muyuan Li, Daniel Miller, and Kenneth R Brown. Direct measurement of Bacon-Shor code stabilizers. *Physical Review A*, 98(5):050301, 2018.

[148] P.W. Shor. Fault-tolerant quantum computation. *Proceedings of 37th Conference on Foundations of Computer Science*, 1996.

[149] Barbara M Terhal. Quantum error correction for quantum memories. *Reviews of Modern Physics*, 87(2):307, 2015.

[150] Eric Dennis, Alexei Kitaev, Andrew Landahl, and John Preskill. Topological quantum memory. *Journal of Mathematical Physics*, 43(9):4452–4505, 2002.

[151] Sergey Bravyi and Alexei Kitaev. Universal quantum computation with ideal clifford gates and noisy ancillas. *Physical Review A*, 71(2):022316, 2005.

[152] Muyuan Li, Daniel Miller, Michael Newman, Yukai Wu, and Kenneth R Brown. 2D compass codes. *Physical Review X*, 9(2):021041, 2019.

[153] Ben W Reichardt. Quantum universality from magic states distillation applied to CSS codes. *Quantum Information Processing*, 4(3):251–264, 2005.

[154] Bryan Eastin and Emanuel Knill. Restrictions on transversal encoded quantum gate sets. *Physical Review Letters*, 102(11):110502, 2009.

[155] IV Inlek, G Vittorini, D Hucul, C Crocker, and C Monroe. Quantum gates with phase stability over space and time. *Physical Review A*, 90(4):042316, 2014.

[156] Daniel A Lidar, Isaac L Chuang, and K Birgitta Whaley. Decoherence-free subspaces for quantum computation. *Physical Review Letters*, 81(12):2594, 1998.

[157] David Kielpinski, V Meyer, MA Rowe, Cass A Sackett, Wayne M Itano, C Monroe, and David J Wineland. A decoherence-free quantum memory using trapped ions. *Science*, 291(5506):1013–1015, 2001.

[158] John P Gaebler, Ting Rei Tan, Y Lin, Y Wan, R Bowler, Adam C Keith, S Glancy, K Coakley, E Knill, D Leibfried, et al. High-fidelity universal gate set for $^{9}\mathrm{be}^{+}$ ion qubits. *Physical review letters*, 117(6):060505, 2016.