Simulation of N-Phonon Entanglement and Tomography in a Bulk Acoustic Resonator

Semester Thesis

Michael Alexander Eichenberger

Supervision
Yu Yang

Prof. Dr. Yiwen Chu

HyQu - Hybrid Quantum Systems
Laboratory for Solid State Physics
ETH Zürich

August 30, 2022

.

Abstract

Hybrid quantum systems consisting of a superconducting qubit and a high-overtone bulk acoustic resonator are a promising platform for quantum information processing. Due to the spectral properties of the acoustic resonator, a single superconducting qubit can couple to a multitude of phonon modes. By encoding qubit states in these modes, a compact and controllable multi-qubit system is achieved. As a proof of principle of quantum information encoding and control, we want to generate and tomographically measure maximally entangled states between multiple phonon modes. This semester thesis tackles the techniques used to generate and measure multi-phonon entangled states. Specifically, pulse sequences for the generation of arbitrary W-states and state reconstruction via quantum tomography were developed and tested in simulations.

Contents

1 Overview

This document is intended to be a short and concise summary of the work conducted during the semester thesis. It is written with the intention that people in the *Hybrid Quantum systems* group can reconstruct my work. It is designed as a report much more than a thesis.

All of the data obtained during the semester project, as well as the written code, can be found on the *group drive* under *10 - Personal Folders/Michael/Phonon Entanglement* and */Control Gate*.

Project Overview: The main goal of my semester project was to design and test a sequence to generate and measure multi-phonon entanglement in the ħBAR device. For this, pulse-level simulations with the QIP processor framework from QuTip were carried out. The simulation environment can be found on gitlab: hyqu_simulator.

I managed to design and test (simulation) sequences that generate arbitrary sized (maximally entangled) W-states and measure the qubit states via multi-qubit tomography (quantum state tomography for the qubit-like subspaces of the phonon modes). The new sequences were then included in the hyqu_simulator framework.

As a next step, I tried to assess the robustness of the new protocols by testing them for various device parameter sets. Additionally, I optimized the qubit π-pulse sequence and the SWAP-gate sequence to maximize the fidelity of the entanglement protocols. The carried out simulations show that the main limiting factor are the coherence times of the current device. For high-coherence settings ($T_1 > 100\mu s$), properties such as the envelopes of the control pulses and the coupling to non-resonant phonon modes become relevant.

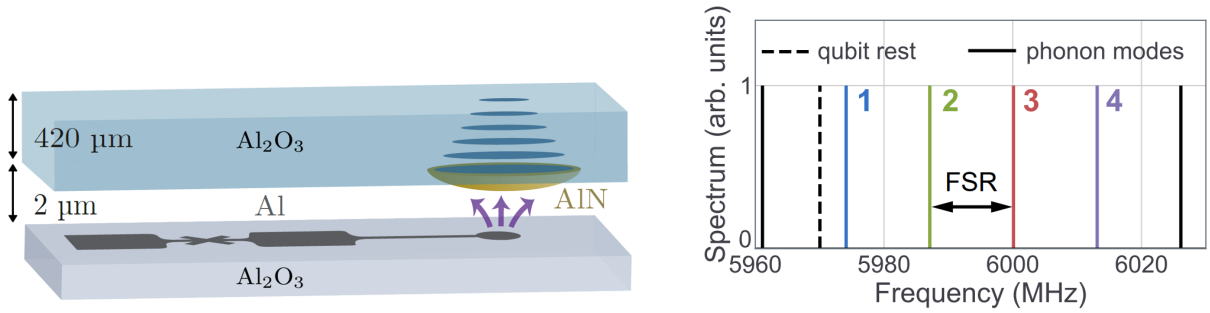
Lastly, after analyzing and optimizing the entanglement sequences, I started looking into the realization of an entangling gate between two phonon modes. A controlled phonon-phonon gate such as the CNOT- or CPHASE-gate is necessary to realize arbitrary quantum circuits on the ħBAR platform.

2 The System

We model the qubit - HBAR system by considering the following hamiltonian

$$\hat{H}/\hbar = \omega^q(t)\hat{a}^\dagger\hat{a} + \frac{1}{2}\alpha\hat{a}^\dagger\hat{a}^\dagger\hat{a}\hat{a} + \sum_{k=1}^{n}\omega_k^p\hat{b}_k^\dagger\hat{b}_k + \sum_{k=1}^{n}g_k\left(\hat{b}_k + \hat{b}_k^\dagger\right)\left(\hat{a} + \hat{a}^\dagger\right) + \Omega\cos\left(\omega^q(t)t + \delta(t)\right)\left(\hat{a} + \hat{a}^\dagger\right) \tag{1}$$

where ω^q is the tunable qubit-frequency, δ a tunable phase for the qubit drive, $\hat{a}$ the qubit annihilation operator, $\hat{b}_k$ the annihilation operator and ω_k^p the frequency of the k-th phonon mode, α the anharmonicity of the (transmon) qubit, g_k the coupling between the qubit



(a) (b)

Figure 1: (a) Sketch of the $\hbar$BAR device consisting of a transmon qubit and a high overtone bulk acoustic resonator (HBAR). The coupling between qubit and HBAR is realized via the piezoelectric effect [1]. (b) shows the phonon modes considered for the simulations: Four successive Laguerrre-Gaussian (0,0) modes. *qubit rest* indicates the default frequency of the qubit when not tuned for an interaction process.

and the k-th phonon mode and lastly Ω the strength of the external drive (applied to the qubit). Additionally to this hamiltonian, the decay- and decoherence times for both qubit (T_1^q, T_2^q) and phonon modes $(T_1^{p,k}, T_2^{p,k})$ are considered. Figure 1 shows a sketch of the actual device and the phonon modes considered.

2.1 Simulation Parameters

For the simulations, a realistic parameter set obtained from the actual device were used. This parameter set is denoted *hero sample parameters*. The name *hero sample* originates from this specific device sample having superior properties compared to the other samples (it is *our current hero*). The following table shows all the relevant parameter values:

$$\begin{aligned} \omega^q_{\text{rest}}/2\pi &= 5.97004 \text{ GHz} & T_1^q &= 10\ \mu\text{s} \\ \omega^p_1/2\pi &= 5.974115 \text{ GHz} & T_2^q &= 15\ \mu\text{s} \\ \text{FSR} &= 13 \text{ MHz} & T^p_{1,k} &= 77\ \mu\text{s} \\ g_k/2\pi &= 260 \text{ kHz} & T^p_{2,k} &= 104\ \mu\text{s} \end{aligned}$$

Figure 2: hero sample parameters. This parameter set was used for simulations if not stated differently.

In the carried out simulations two terms are often used: T_1 *limited scaling* and *hero sample scaling*. I used these terms to describe the process where the system parameters are swept, but the relation between them is held constant. Specifically, the two terms mean

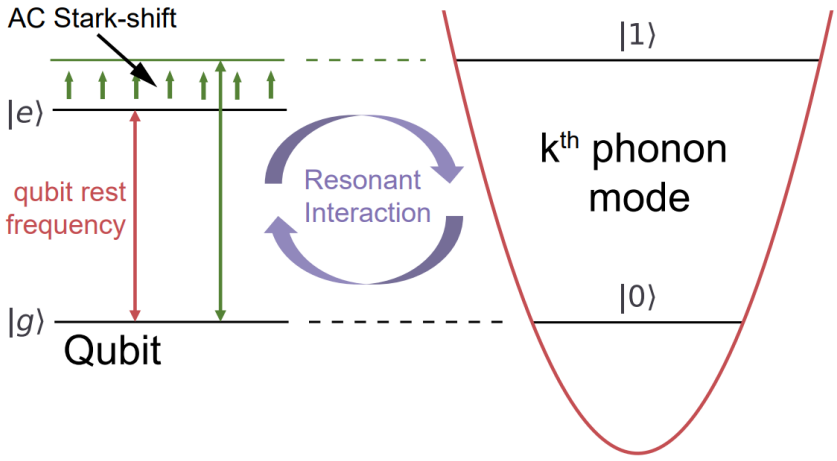


Figure 3: SWAP interaction between qubit and the k-th phonon mode.

$$\begin{array}{lccc} T_1 \textit{ limited scaling} & T_1^p = 7.7T_1^q & T_2^q \approx 2T_1^q & T_2^p \approx 2T_1^p \\ \textit{hero sample scaling} & T_1^p = 7.7T_1^q & T_2^q = 1.5T_1^q & T_2^p = 10.4T_1^q \end{array}$$

Note that for the simulation, the T_2 times must be chosen strictly smaller than $2T_1$, since else the QIP processor runs into zero division. For the simulations I chose $T_2 = 2{\cdot}0.9999T_1$.

2.2 iSWAP Gate and Generalization

Assume that the qubit frequency ω_q is tuned on resonance with the k-th phonon mode. If the phonon FSR is sufficiently large, we can neglect the coupling between the qubit and off-resonant phonons and the hamiltonian reduces to the simple beamsplitter interaction

$$H \approx g_k(a^\dagger b_k + a b_k^\dagger). \tag{2}$$

We can thus just consider the system composed of the qubit and the k-th phonon mode. Figure 3 shows a sketch of the qubit-phonon interaction.

If we evolve the system under this hamiltonian, we achieve the **generalized SWAP gate**:

$$U_{gswap} = e^{-i\hat{H}t} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(g_k t) & -i\sin(g_k t) & 0 \\ 0 & -i\sin(g_k t) & \cos(g_k t) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{3}$$

This unitary essentially describes Rabi-oscillations between the qubit and the k-th phonon mode. By setting the interaction times t to the values $t_{\text{iSWAP}} = \pi/(2g_k)$ and $t_{\sqrt{\text{iSWAP}}} = \pi/(4g_k)$ we can realize the iSWAP and the $\sqrt{\text{iSWAP}}$ gate:

$$\text{iSWAP} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \sqrt{\text{iSWAP}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -i & 0 \\ 0 & -i & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4}$$

The generalized SWAP gate U_{gswap} is the main component of both the entanglement generation sequence and the phonon tomography.

3 Phonon Entanglement

For pure-state quantum mechanics a state on a bipartite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ is considered *entangled* if it is not separable, where *separability* is defined as [2]

$$|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B \text{ is separable} \iff |\psi\rangle = |\phi_A\rangle \otimes |\phi_B\rangle =: |\Phi\rangle. \tag{5}$$

Thus a pure state is separable if it can be represented as a single product state. For mixed states, the condition for separability generalizes to

$$\rho \text{ is separable} \iff \rho = \sum_i p_i |\Phi_i\rangle\langle\Phi_i| \qquad \sum_i p_i = 1, \tag{6}$$

where ρ is the density matrix of the quantum state and $\forall i : |\Phi_i\rangle$ are product states. This means that if we take mixed states into account, an entangled state cannot be represented as a statistical mixture of product states. This definition still holds for multipartite systems, if we group the full system partition into two groups

$$\exists S : \quad \mathbf{H} = \bigotimes_i \mathcal{H}_i = \left(\bigotimes_{m \in S} \mathcal{H}_m \right) \otimes \left(\bigotimes_{n \notin S} \mathcal{H}_n \right) \equiv \mathcal{H}_A \otimes \mathcal{H}_B, \tag{7}$$

where $S \subset \{1, \ldots, N\}$ is a subset of system parts. Hence, for the multipartite system entanglement means that the state is entangled in respect to at least one bipartition of the full system. Maximal multipartite entanglement then would imply that for all subsets S there is residual bipartite entanglement. Unfortunately, this definition is rather elusive and the categorization of entanglement in multipartite quantum systems is much richer than just the embedding of bipartite entanglement [3].

In order to asses the entanglement of a system, we now want to be able to do two things: (1) Detect if the state is indeed entangled [4] and (2) quantify the entanglement [2]. For multipartite systems, designing quantities that serve one of these tasks is highly non-trivial and an active field of research [5–7]. Recently, a very general entanglement criterion for multipartite systems with discrete and continuous variables has been constructed using

entanglement witnesses [8, 9]. However, this construction comes with the cost of increased computational complexity.

For the entangled phonon-states analyzed in this work, I used the following construction to quantify the entanglement.

3.1 N-Negativity: Entanglement Measure for Phonon-States

We want to create a measure of entanglement which is applicable to an N-partite system. For this, we use the well-established negativity $\mathcal{N}$ [10] defined for bipartite systems as

$$\mathcal{N}(\rho, A) = \frac{1}{2}\left(||\rho^{T_A}||_1 - 1\right), \tag{8}$$

where $||\cdot||_1$ is the Schatten-1 norm and ρ^{T_A} is the partial transpose in respect to subsystem A. The negativity $\mathcal{N}$ vanishes for separable states and does not increase under local operations and classical communication (LOCC), thus it is an entanglement monotone (also for mixed states) [10]. This can now be extended to the multipartite-space by calculating $\mathcal{N}$ for all possible bipartitions of the full space and averaging over the respective results. Note that since the negativity is proper a measure of entanglement, the average of all bipartitions is still a measure of entanglement [2]. For a composite system $\mathbf{H} = \bigotimes_{i=1}^{N} \mathcal{H}_i$ the N-partite negativity, from now on called *N-negativity*, can thus be calculated as

$$\mathcal{N}_N(\rho) = \frac{1}{2^N - 2} \sum_{\emptyset \neq S \subsetneq \{1,\ldots,N\}} \mathcal{N}\left(\rho, \bigotimes_{i \in S} \mathcal{H}_i\right), \tag{9}$$

where S runs over all possible subsets of $\{1, \ldots, N\}$ except of the empty set $\emptyset$ and the full set $\{1, \ldots, N\}$. In theory, only subsets with $\lfloor N/2 \rfloor$ elements need to be considered (because for $\mathcal{H} = A \otimes B : \mathcal{N}(\rho, A) = \mathcal{N}(\rho, B)$). For simplicity in the calculation all terms are considered, since they average out and thus do not change the obtained value.

3.2 The W-state

By general swap gate $U_{\text{gswap}}(t)$ between the qubit and different phonon modes $1, \ldots N$ an N-dimensional W-state can be generated. An ideal W-state has the form

$$|W\rangle_N = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} |(2^k)_2\rangle \tag{10}$$

,

where the number in the kets indicates a bit-string (binary representation of the number), e.g. for $|W\rangle_3$:

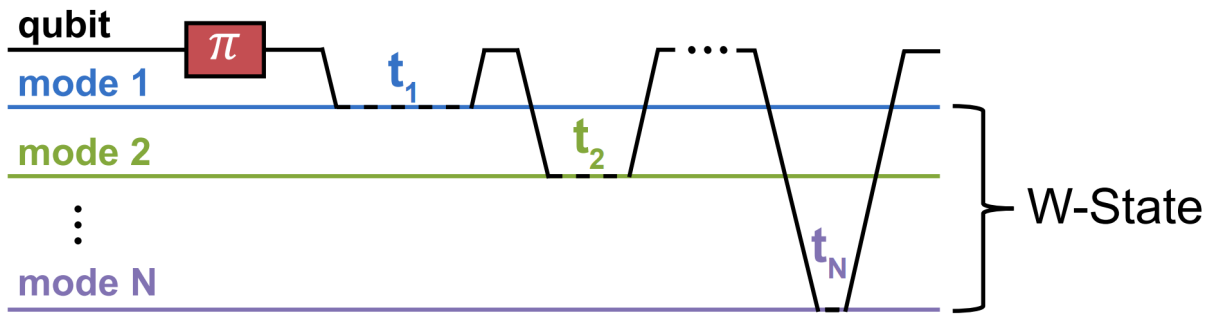


Figure 4: Generation sequence for the N-Phonon W-state. Note that with the last interaction, all the excitation from the qubit is deposited in the phonon modes, creating a W-state among the phonon modes.

$$|W\rangle_3 = \frac{1}{\sqrt{3}}(|001\rangle + |010\rangle + |100\rangle) = \frac{1}{3}(|(2^0)_2\rangle + |(2^1)_2\rangle + |(2^2)_2\rangle). \tag{11}$$

The achievable state in a real experiment has additional relative phases, because all the phonon modes rotate at a different frequency.

$$|W\rangle_N^{\text{real}} = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{\phi_k} |(2^k)_2\rangle \tag{12}$$

3.3 Entanglement Generation Sequence

The W-state is generated by first applying a π-Pulse to the qubit and then swapping the excitation to the N phonon modes. The goal here is to choose the interaction time t_k for the k-th phonon mode such that exactly the a fraction of $1/\sqrt{N}$ (amplitude) is swapped to the phonon mode. This condition can be calculated for each step, leading to the following expression for the k-th interaction time

$$t_k = \frac{1}{g_k} \arcsin\left(\frac{1}{\sqrt{N+1-k}}\right) \qquad k \in \{1, \ldots, N\}. \tag{13}$$

The total sequence thus contains of a qubit π-pulse and N qubit-phonon interactions with times t_k. Figure 4 shows a sketch of the sequence.

Sequence Time T_{gen}

By summing up the necessary swap gate times t_k for the entire W-state generation sequence we can find the total process time as

$$T_{gen} = T_\pi + \sum_{k=1}^{N} t_k = \frac{4}{\Omega} + \sum_{k=1}^{N} \frac{1}{g_k} \arcsin\left(\frac{1}{\sqrt{N+1-k}}\right) \tag{14}$$

.

We can find an upper bound for the generation time (see Appendix A)

$$T_{ges} < (0.5 + 1.2\sqrt{N-1})\ \mu s \tag{15}$$

where N is the number of entangled phonon modes. Therefore the entire process scales with $\mathcal{O}(\sqrt{N})$. Figure 5 shows the sequence times including the derived upper bounds.

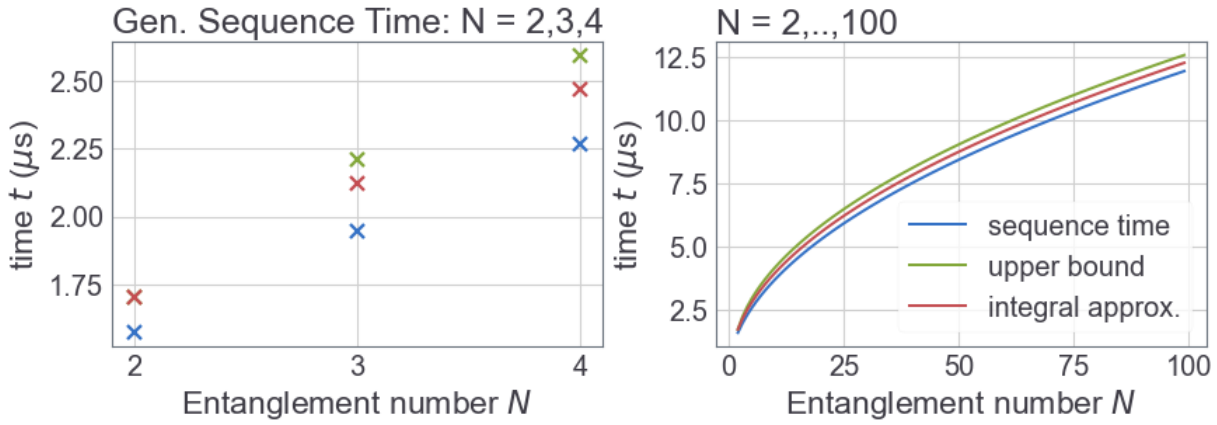


Figure 5: Generation sequence time as a function of the entanglement number N. *Integral approx.* indicates a second approximation method for the generation sequence time which yields an even better estimate (see Appendix A for details).

4 Phonon Tomography

We use quantum state tomography to reconstruct the density matrix of the multi-phonon system. Since our code space only includes qubit-like encodings in the $|0\rangle$ and $|1\rangle$ Fock states of the harmonic oscillator, we want to reconstruct the density matrix only of that (multi-qubit-like) subspace

$$\mathcal{H} = \bigotimes_{k=1}^{N} \mathcal{C}^2 \cong \mathcal{C}^{2^N}. \tag{16}$$

The desired density matrix $\hat{\rho}$ then is a positive, trace one hermitian operator on this Hilbert space

$$\hat{\rho} \in \text{End}\left(\bigotimes_{k=1}^{N} \mathcal{C}^2\right), \qquad \hat{\rho}^\dagger = \hat{\rho}, \qquad \text{Tr}[\hat{\rho}] = 1, \qquad \hat{\rho} \geq 0. \tag{17}$$

The **main idea** now is to use Linear Algebra to write $\hat{\rho}$ as the linear combination of a basis of $\{A \in \text{End}(\mathcal{H}) | A^\dagger = A\}$ we can measure in the experiment. For our purposes, the basis of choice is the tensor-product pauli basis, consisting of all possible tensor-product combinations of the Pauli operators $\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z$ plus the identity matrix $\mathbb{I}$. The individual basis vectors are then given by

$$\text{k-th basis vector:} \qquad \hat{\Sigma}_k = \bigotimes_{l_k} \hat{\sigma}_{l_k} \qquad \hat{\sigma}_{l_k} \in \{\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z, \mathbb{I}\}. \tag{18}$$

Here l_k iterates over all possible combinations of the Paulis plus identity. The set $\{\hat{\Sigma}_k\}_k$ is then a basis of the full space, since the Paulis plus identity are a basis of the hermitian single-qubit operators. In total there are 4^N basis vectors for given system size N. We now note that the chosen basis is *orthonormal* for the trace scalar product on the full space. For $\hat{A} = \bigotimes_{k=1}^{N} \hat{a}_k$ and $\hat{B} = \bigotimes_{k=1}^{N} \hat{b}_k$ it is defined as

$$\langle A, B \rangle = \prod_{k=1}^{N} \langle \hat{a}_k, \hat{b}_k \rangle = \frac{1}{2^N} \prod_{k=1}^{N} \text{Tr}\left[\hat{a}_k^\dagger \hat{b}_k\right] \tag{19}$$

Therefore we can describe the desired density matrix $\hat{\rho}$ fully by the coefficients r_k of the expansion in the Pauli basis

$$\hat{\rho} = \frac{1}{2^N} \sum_k r_k \hat{\Sigma}_k \qquad r_k = \text{Tr}\left[\hat{\Sigma}_k \hat{\rho}\right], \tag{20}$$

where the relation for the coefficients r_k follows from $\{\hat{\Sigma}_k\}_k$ being an ONB. We thus have reduced the problem to finding the coefficients r_k for the expansion. These coefficients are the *expectation value* of the hermitian operator $\hat{\Sigma}_k$

$$r_k = \text{Tr}\left[\hat{\Sigma}_k \hat{\rho}\right] = \langle \hat{\Sigma}_k \rangle. \tag{21}$$

Unfortunately, since the quantities directly observable in the experiment are the $\hat{\sigma}_z$ projectors $|g\rangle\langle g|$ and $|e\rangle\langle e|$, we still need to relate r_k to these projectors. This can be done by noticing two things:

1. The projectors of the $\hat{\sigma}_x$ and $\hat{\sigma}_y$ operators are also accessible via measurement of $|g\rangle\langle g|$ and $|e\rangle\langle e|$, by rotating the measurement frame first with a $\pi/2$-pulse in the x- or y-direction. We find

$$\begin{aligned}\mathrm{Tr}\left[|+i\rangle\langle+i|\hat{\rho}\right] &= \mathrm{Tr}\left[R_x^\dagger\left(-\frac{\pi}{2}\right)|g\rangle\langle g|R_x\left(-\frac{\pi}{2}\right)\hat{\rho}\right]\\ \mathrm{Tr}\left[|-i\rangle\langle-i|\hat{\rho}\right] &= \mathrm{Tr}\left[R_x^\dagger\left(-\frac{\pi}{2}\right)|e\rangle\langle e|R_x\left(-\frac{\pi}{2}\right)\hat{\rho}\right]\end{aligned}$$

$$\begin{aligned}\mathrm{Tr}\left[|+\rangle\langle+|\hat{\rho}\right] &= \mathrm{Tr}\left[R_y^\dagger\left(\frac{\pi}{2}\right)|g\rangle\langle g|R_y\left(\frac{\pi}{2}\right)\hat{\rho}\right]\\ \mathrm{Tr}\left[|-\rangle\langle-|\hat{\rho}\right] &= \mathrm{Tr}\left[R_y^\dagger\left(\frac{\pi}{2}\right)|e\rangle\langle e|R_y\left(\frac{\pi}{2}\right)\hat{\rho}\right]\end{aligned}$$

This means, that by applying a $R_y(\pi/2)$ pulse to the system before readout, we can measure in the x basis and by applying $R_x(-\pi/2)$ pulse we can measure in the y-basis.

2. The Pauli operators can be decomposed into projectors and thus the measurements of the projectors is related to the full Pauli. We have

$$\begin{aligned}\hat{\sigma}_z &= |g\rangle\langle g| - |e\rangle\langle e|\\ \hat{\sigma}_x &= |+\rangle\langle+| - |-\rangle\langle-|\\ \hat{\sigma}_y &= |+i\rangle\langle+i| - |-i\rangle\langle-i|\\ \mathbb{I} &= |g\rangle\langle g| + |e\rangle\langle e|\end{aligned}$$

And thus for a given operator $\hat{\Sigma}_k$ we can decompose it further into tensor products of projectors. We make the notation more convenient by writing the projector of the $+1$ eigenvalue state as $|0_{x,y,z,\mathbb{I}}\rangle\langle 0_{x,y,z,\mathbb{I}}|$. Similarly, for the -1 eigenvalue state $|1_{x,y,z,\mathbb{I}}\rangle\langle 1_{x,y,z,\mathbb{I}}|$. E.g. in this notation the state $|-i\rangle\langle-i|$ would become $|1_y\rangle\langle 1_y|$. In this notation the decomposition of $\hat{\Sigma}_k$ takes the following form

$$\hat{\Sigma}_k = \sum_{x\in\{0_{l_k},1_{l_k}\}^N} (-1)^{\oplus x}|x\rangle\langle x|. \tag{22}$$

Here $x = x_1x_2\ldots x_N$ is a bit-string (e.g. $|x\rangle = |0_z1_x0_x\ldots1_z\rangle$) of the 0 and 1 states and $\oplus$ is defined as

$$\oplus := \#\{x_s|x_s \in \{1_x, 1_y, 1_z\}\} \tag{23}$$

With these two ingredients we can now write the coefficients r_k as the sum of projector measurements obtainable in the experiment:

$$\begin{aligned} r_k &= \mathrm{Tr}\left[\hat{\Sigma}_k\hat{\rho}\right] \\ &= \mathrm{Tr}\left[\sum_{x\in\{0,1\}^N}(-1)^{\oplus x}|x\rangle\langle x|\hat{\rho}\right] \\ &= \sum_{x\in\{0_{l_k},1_{l_k}\}^N}(-1)^{\oplus x}\mathrm{Tr}\left[|x\rangle\langle x|\hat{\rho}\right], \end{aligned}$$

where on the third line we used the linearity of the trace. The expectation value $\mathrm{Tr}\left[|x\rangle\langle x|\hat{\rho}\right]$ can be obtained via conditional probabilities. We have

$$\mathrm{Tr}\left[|x\rangle\langle x|\hat{\rho}\right] = \langle x\rangle = \mathrm{P}\left((\text{phonon 1 is in state } x_1)\cap\cdots\cap(\text{phonon N is in state } x_N)\right) \quad (24)$$

Using the notation

$$\mathrm{P}\left(x_1,\ldots,x_N\right) \equiv \mathrm{P}\left((\text{phonon 1 is in state } x_1)\cap\cdots\cap(\text{phonon N is in state } x_N)\right) \quad (25)$$

we apply Bayes' law for conditional probabilities and obtain

$$\begin{aligned} \mathrm{Tr}\left[|x\rangle\langle x|\hat{\rho}\right] &= \mathrm{P}\left(x_1,\ldots,x_N\right) \\ &= \mathrm{P}\left(x_N|x_1,\ldots,x_{N-1}\right)\mathrm{P}\left(x_1,\ldots,x_{N-1}\right) \\ &= \ldots \\ &= \mathrm{P}\left(x_N|x_1,\ldots,x_{N-1}\right)\mathrm{P}\left(x_{N-1}|x_1,\ldots,x_{N-2}\right)\ldots\mathrm{P}\left(x_1\right) \\ &= \prod_{k=1}^{N}\mathrm{P}\left(x_k|x_1,\ldots,x_{k-1}\right) \end{aligned}$$

These conditional probabilities can now be directly obtained from the experiment by sequentially measuring the phonon states and taking the statistical average

$$\mathrm{P}\left(x_k|x_1,\ldots,x_{k-1}\right) \approx \langle\langle x_k\rangle\rangle \text{ and } x_1,...,x_{k-1} \text{ measured before} \quad (26)$$

where we assume that the statistical average $\langle\langle.\rangle\rangle$ is a good estimator of the population average $\langle.\rangle$. This method is called *Direct Inversion Tomography* [11] and has its quirks (see Section 8.4).

Summary

We can do direct inversion tomography and reconstruct the state density matrix $\hat{\rho}$ by expanding it in a Pauli-Operator basis and then expanding the Pauli basis into projectors. The expectation values for the projectors then translate to conditional probabilities obtainable from the experiment by calculating the statistical average of σ_z projector measurements. σ_x and σ_y projectors can be measured by rotating the measurement frame via additional $\pi/2$-pulses. The full density matrix takes the form

$$\hat{\rho} = \frac{1}{2^N} \sum_k \left(\sum_{x \in \{0_{l_k}, 1_{l_k}\}^N} (-1)^{\oplus x} \prod_{m=1}^{N} \mathrm{P}\left(x_m | x_1, \ldots, x_{m-1}\right) \right) \hat{\Sigma}_k \tag{27}$$

4.1 Tomography Sequence

The tomography sequence consists of sequential measurements of the N phonon mode. For each phonon-mode, the following procedure is carried out:

1. swap in Phonon Mode k to qubit (iSWAP)
2. apply $\pi/2$-pulse to qubit (determine measurement direction). The x-direction can be chosen arbitrarily from the xy-plane at the beginning of the tomography sequence by setting the parameter *tomography phase* ϕ ($\phi = 0 : \sigma_x$ pulse, $\phi = \pi/2 : \sigma_y$ pulse)
3. measure qubit state
4. wait time (t_wait)
5. cool qubit down with Mode k
   ...
6. swap in Phonon Mode k+1

The sequence for the k-th phonon mode is also illustrated in Figure 6. By repeating this sequence many times for each measurement axis combination we can obtain the needed measurement statistics. From these, we can calculate the conditional probabilities of the state in post-processing, which gives sufficient information to calculate the full state density matrix.

Sequency Time T_{tom}

The tomography sequence scales linear in the entanglment number N: $T_{tom} \in \mathcal{O}(N)$. The specific times needed for the tomography sequence are shown in Fig. 7.

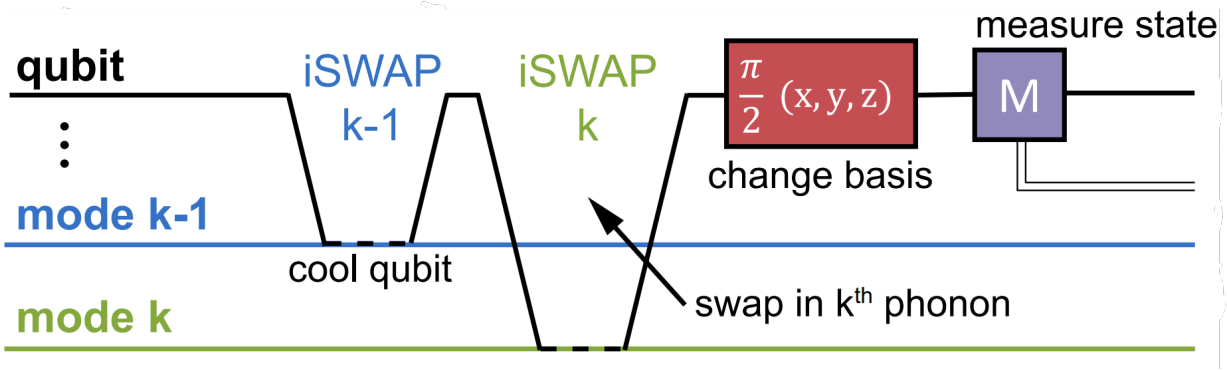


Figure 6: Tomography sequence for the k-th phonon mode in the N-phonon tomography. The depcited sequence must be repeated for all $1, \ldots, N$ phonon modes.

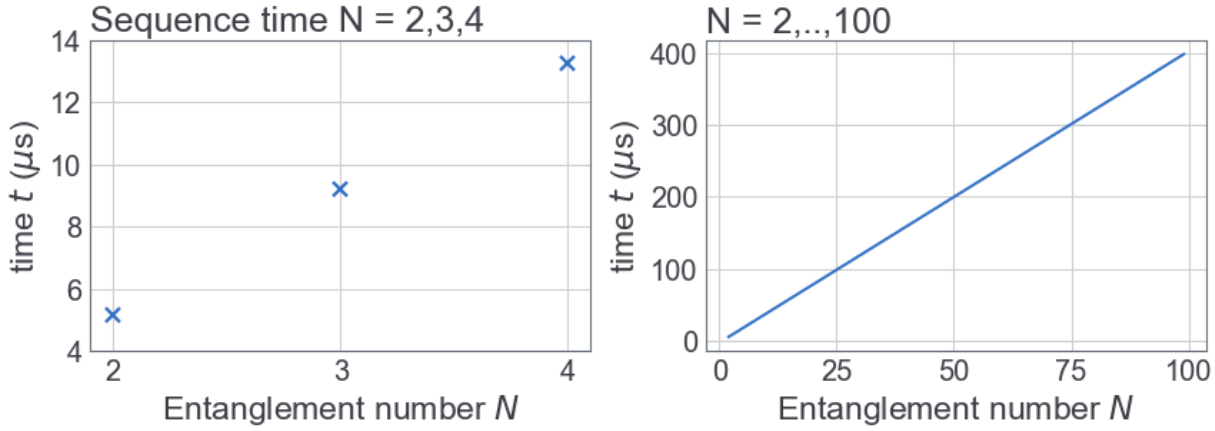


Figure 7: Tomography sequence time as a function of the entanglement number N. Here a wait time of $t_{\text{wait}} = 2\,\mu\text{s}$ was chosen.

5 Optimization

In this section the optimization of both the SWAP gate and the π-pulse are treated.

5.1 Optimizing the Swap Gate

The goal here is to analyze the SWAP gate characteristics in detail and understand in depth which system properties influence the SWAP gate fidelity. For this, various simulations were carried out. Figure 8 for example shows a sweep of the swap interaction time (swap ratio) and detuning paramters and the corresponding achieved fidelity. From this we can see that the ideal SWAP parameters seem to not be at the expected value of ratio=1 and detuning=0. Figure 9 shows the coherence time scaling of the SWAP fidelity and the optimal parameters. Here, for every coherence time parameter set the SWAP gate sweep

is performed. We see that the optimal parameters depend on the mode that is swapped. Especially mode 1 deviates from the expected ratio=1, detuning=0 behavior. Table 1 lists the obtained optimal parameters.

	mode 1	mode 2	mode 3	mode 4
max. fidelity	0.9389	0.9232	0.9231	0.9231
swap ratio	0.96410	0.99487	1.00513	1.00513
detuning kHz	56.41	15.38	5.13	5.13

Table 1: Optimized SWAP parameters (ratio and detuning) for the hero sample paramters.

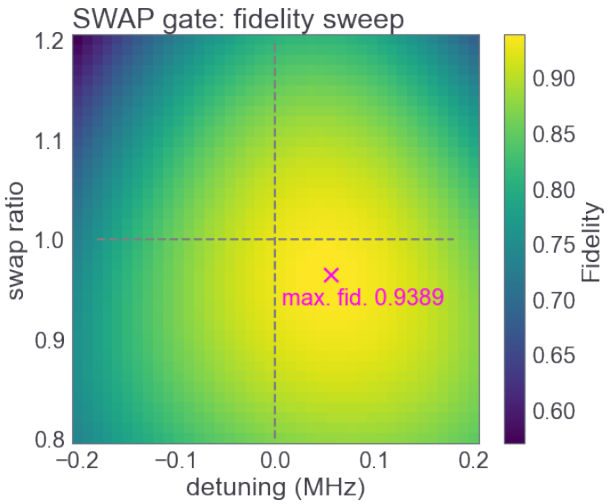


Figure 8: SWAP gate parameter ratio and detuning sweep for first phonon mode (mode 1) and 4 modes present in simulation. The sweep resolution, which was 0.01 for the swap ratio and 0.01 MHz for the detuning

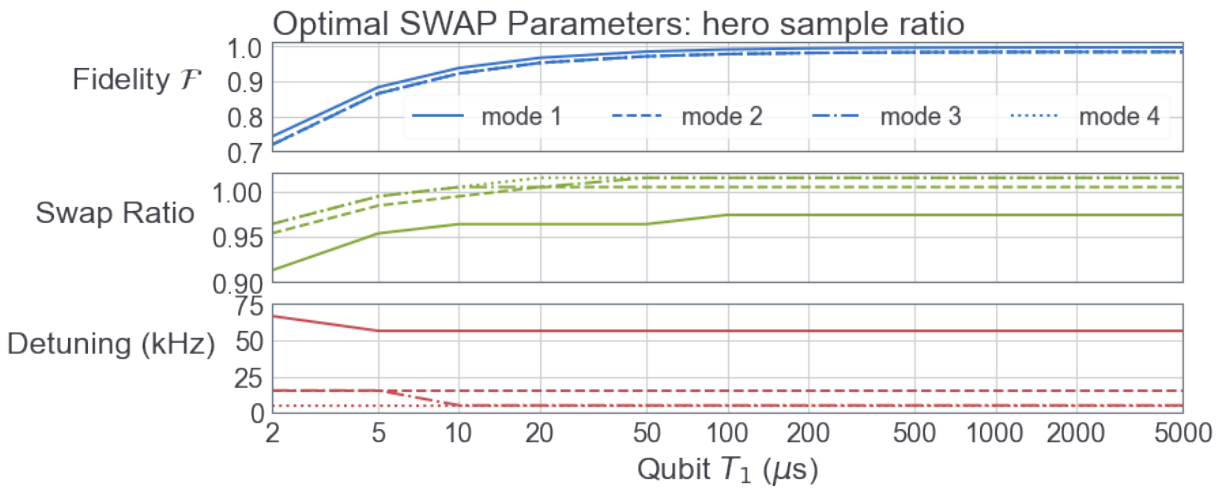


Figure 9: SWAP gate parameter optimization as a function of coherence time scaling (hero sample ratio).

5.2 Compare mode number influence

In a next step, we want to understand how the presence of phonon modes in the simulation impacts the optimized parameters. For this mode 1 is swapped while a different number of (higher) phonon modes is present in the simulations. The optimized paramters for each of the settings are shown in Figure 10. We can see that increasing the number of phonon modes in the simulation does not impact the optimal parameters significantly, thus a different mechanism leads to the shift from ratio=1, detuning=0.

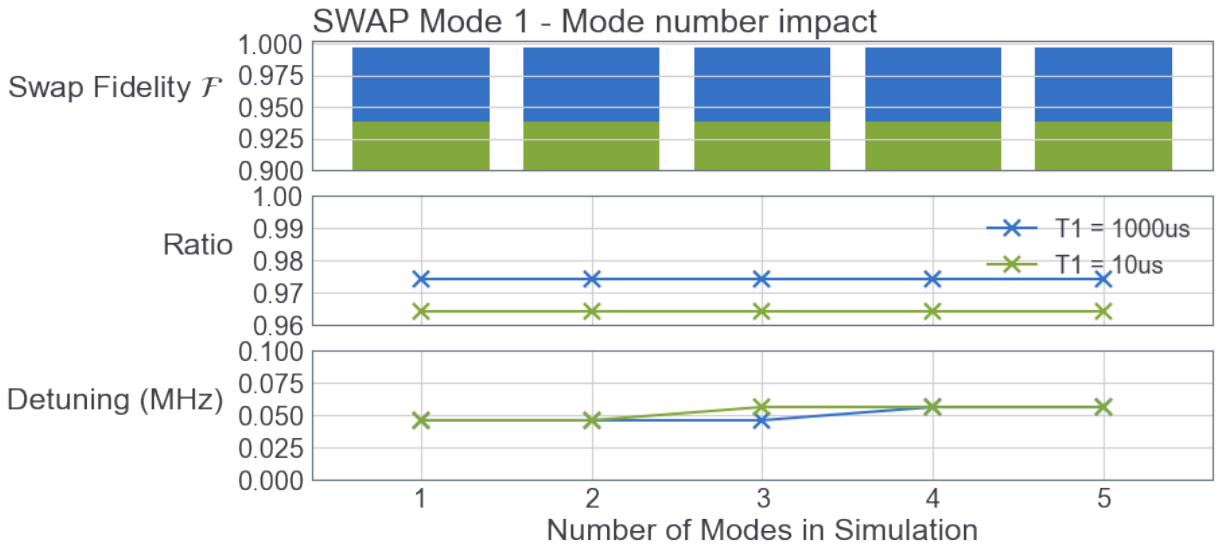
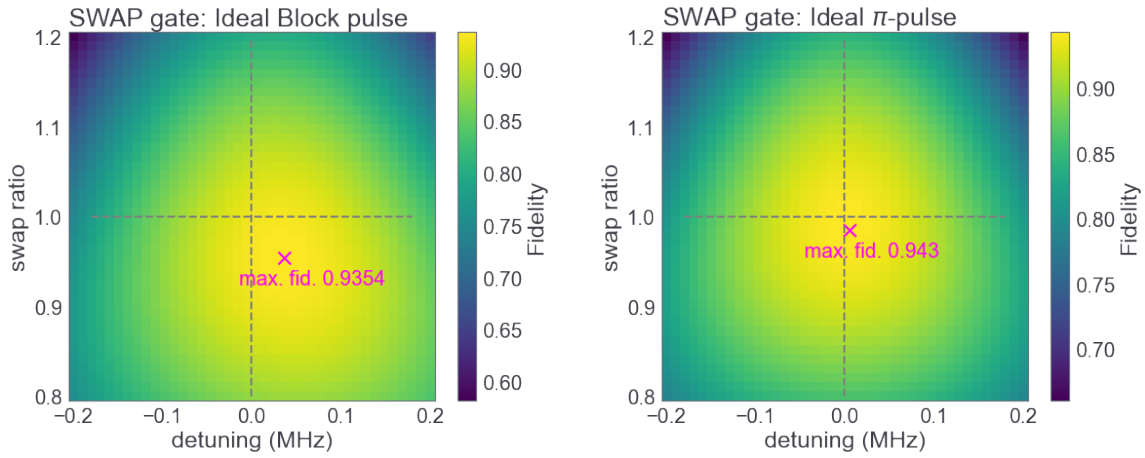


Figure 10: Impact of the mode number to optimal SWAP gates parameters (hero sample paramters). Here, phonon mode 1 swapped. Note that the obtained optimal value is limited by the sweep resolution, which was 0.01 for the swap ratio and 0.01 MHz for the detuning.

Origin of mode 1 detuning

From Figure 10 we can see that even if only one phonon mode is present in the simulation, the ideal SWAP fidelity is still not achieved for zero detuning, but for a higher detuning value. I wanted to investigate the origin of this shift. Figure 11 shows the SWAP parameter sweep for both an ideal SWAP pulse shape (no-gaussian tails, i.e. instant on/off behavior) and an ideal π-pulse (qubit initialized in excited state). We see that for the ideal π-pulse, the optimum is very close to the expected zero detuning settings. Thus we conclude that the SWAP parameter optimum is dependent on the preceding π-pulse. This can be understood as the SWAP optimum taking into account the already present error of the π-pulse. Thus, for further optimization the π-pulse should be optimized. It is not yet fully understood what the origin of the π-pulse infidelity is. I carried out a simulation without any phonon modes present and without decoherence, which yields a π-pulse fidelity of 1 (up to numerical error in the 10^{-6} range). This rules out problems with the simulation itself (at least up to the interaction with phonon modes). With only one phonon mode present (still without decoherence), the fidelity decreases to 0.990 and exhibits some dependence on the tail-length of the gaussian pulse (it increases with longer tails up to 0.992). In my opinion this suggests that the issue is frequency related, since it increases for more tail "cut-off". An explanation could be that the π-pulse on the qubit drives some interaction with the phonon, leading to a decrease in the π-pulse fidelity.



(a) (b)

Figure 11: SWAP parameter sweep for (a) SWAP-pulse with ideal block-shape (b) ideal π-pulse.

5.3 Optimizing the π-Pulse

The current pulse-shape used for the qubit π-pulse is a gaussian with cut tails at $\#\sigma$ (i.e. $\#\sigma$ selects up to what standard deviation the gaussian pulse is carried out), defined by the following equation:

$$c(t, \#\sigma) = \frac{\Omega}{A}\frac{\pi}{2} \exp\left(-\frac{(t-\#\sigma)^2\Omega^2}{2(\#\sigma)^2}\right)\cos(\phi) \tag{28}$$

where Ω is the drive-frequency (= qubit-frequency on resonance), A is a tunable amplitude factor and ϕ defines the an angle in the x-y-plane (allows for arbitrary rotation around the bloch-sphere) and $\#\sigma$ defines at what sigma value the tails of the gaussian are cut off. The pulse duration is fixed to $\#\sigma/\Omega$.

Optimize Amplitude

We can now calculate the amplitude factor A to realize an ideal π-pulse. For this we consider that the drive-hamiltonian (for $\phi = 0$) is given as (for simplicity we don't write the dependence on the additional parameter $\#\sigma$)

$$H(t) = c(t)(a + a^\dagger) \tag{29}$$

This leads to a unitary evolution (under H) of

$$U(t, t_0) = \exp\left[i\left(\int_{t_0}^{t} c(\tilde{t})d\tilde{t}\right)(a + a^\dagger)\right] \tag{30}$$

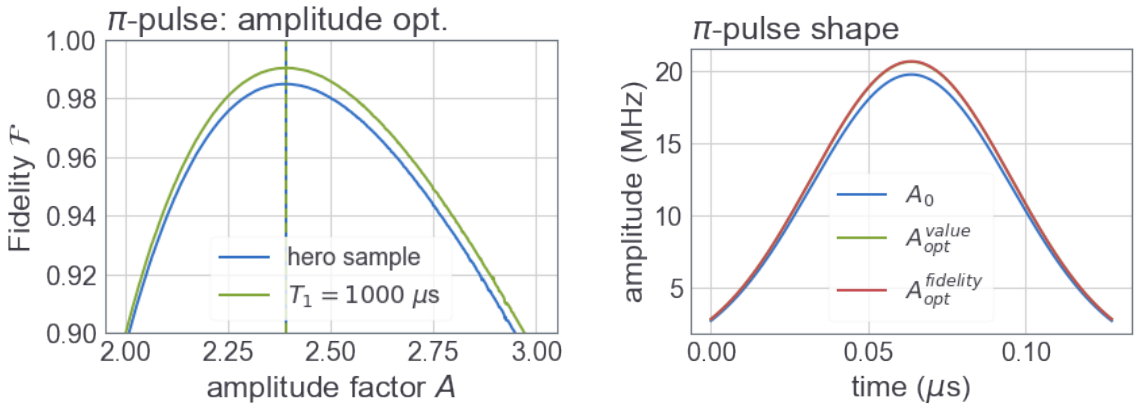
Thus, for an ideal π-pulse, we require

$$\int_{t_0}^{t} c(\tilde{t})d\tilde{t} \stackrel{!}{=} \frac{\pi}{2} \tag{31}$$

Thus in the case of the gaussian pulse envelope

$$\int_{0}^{\#\sigma/\Omega} \frac{\Omega}{A} \exp\left(-\frac{(t - \#\sigma)^2\Omega^2}{2(\#\sigma)^2}\right) dt \stackrel{!}{=} 1 \tag{32}$$

This equation can be numerically solved (I used *scipy.optimize.minimize* with the *Nelder-Mead* solver and a tolerance of 10^{-12}). For $\#\sigma = 4$ (this is the standard value chosen in the simulations) an amplitude factor of $A = 2.393$ was obtained. This lead to a small increase of the simulated π-pulse fidelity from 0.9801 to 0.9848. Note that directly optimizing the simulated fidelity leads to a very similar value of $A = 2.389$ with fidelity 0.9849. Figure 12 shows the initial and optimized π-pulse shapes. Note that these small increases in fidelity can most likely not be measured in the actual experiment. Rather, if the simulator is sufficiently accurate and precise, they allow to study the physical effects limiting the achievable fidelity. Thus, these optimizations mainly treat the question where a perfect system could get us, much more than trying to achieve fidelity increases in the actual experiment.



(a) (b)

Figure 12: (a) π-pulse amplitude factor sweep for two different parameter sets (hero sample ratio). (b) π-pulse shape before and after amplitude optimization.

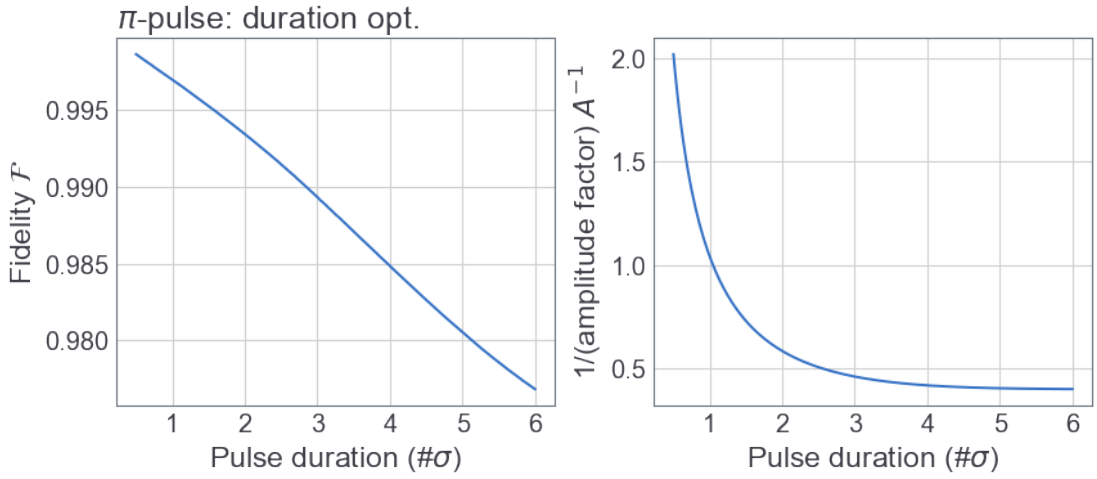


Figure 13: π-pulse duration ($\#\sigma$) sweep for hero sample paramters. (a) π-pulse fidelity with (b) corresponding amplitude factor A.

Optimize cut-off $\#\sigma$

In a second step I tried to optimize the duration of the gaussian pulse by trying different cut-off values $\#\sigma$ for the gaussian. Here, for every chosen cut-off $\#\sigma$, the amplitude A is numerically optimized (Nelder-Mead minimization). Figure 13 shows the found relations between cut-off $\#\sigma$ and optimal amplitude factor A. Figure 14 shows the optimized pulse shapes for some of the tested cut-offs. We can see that for short pulse durations, the optimal pulse converges towards a dirac-delta pulse, which is not achievable in the real experiment. Consequently, for the time being the old value of $\#\sigma = 4$ was kept for simulations.

5.4 Coherence-time scaling

The SWAP gate was benchmarked by simulating the gate fidelity for various parameter sets, shown in Fig. 15 and 16. The abbreviations have the following meaning

- **Hero** hero sample ratio
- **T1lim** T_1-limited ratio
- **T1q** T_1^q scales (other coherence times fixed at hero sample values)
- **T1/T2q** only T_1^q, T_2^q scales (other coherence times fixed at hero sample values)
- **1M** only phonon mode 1 present in simulation
- **OptPi** optimized π-pulse (amplitude value)
- **IdPi** ideal π-pulse (state initialized in ideal qubit e state)
- **OptD** SWAP parameters optimized for every T_1^q value respectively (see Fig. 9)

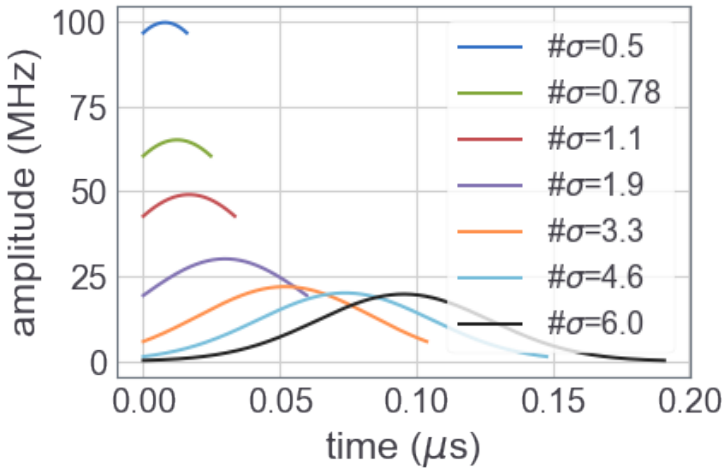


Figure 14: Optimal π-pulse shapes for different durations ($\#\sigma$) (hero sample paramters). Note that for short durations the ideal pulse converges to a dirac-delta dstribution.

- **OptS** SWAP parameters optimized with single value set (average for $T_1^q > 10\,\mu s$ of the obtained optimal parameters from Fig. 9)

From the simulations we find that the biggest limiting factor of the gate fidelity of the current device is the coherence times. Already a T_1^q increase to 20 µs would improve the fidelity by over 2%. Also, we can see that through the optimization of both *pi*-pulse and SWAP parameters a fidelity increase of 2% could be achieved. This is not a lot, but for the tomography sequence with $2N$ swaps it does makes a difference. From Fig. 16 we can also see that for long coherence times the π-pulse seems to be the main limiting factor.

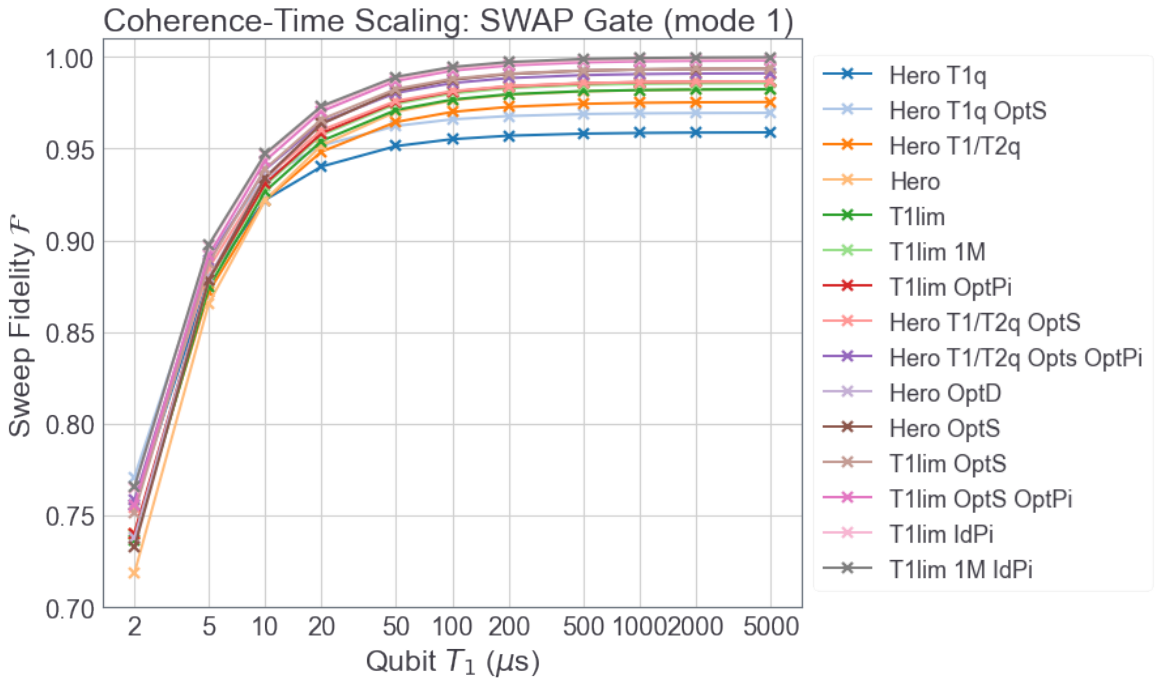


Figure 15: Coherence-time scaling of the SWAP gate. See list in 5.4 for an explanation of the abbreviations. In these simulation 4 phonon modes of dimension 2 were present.

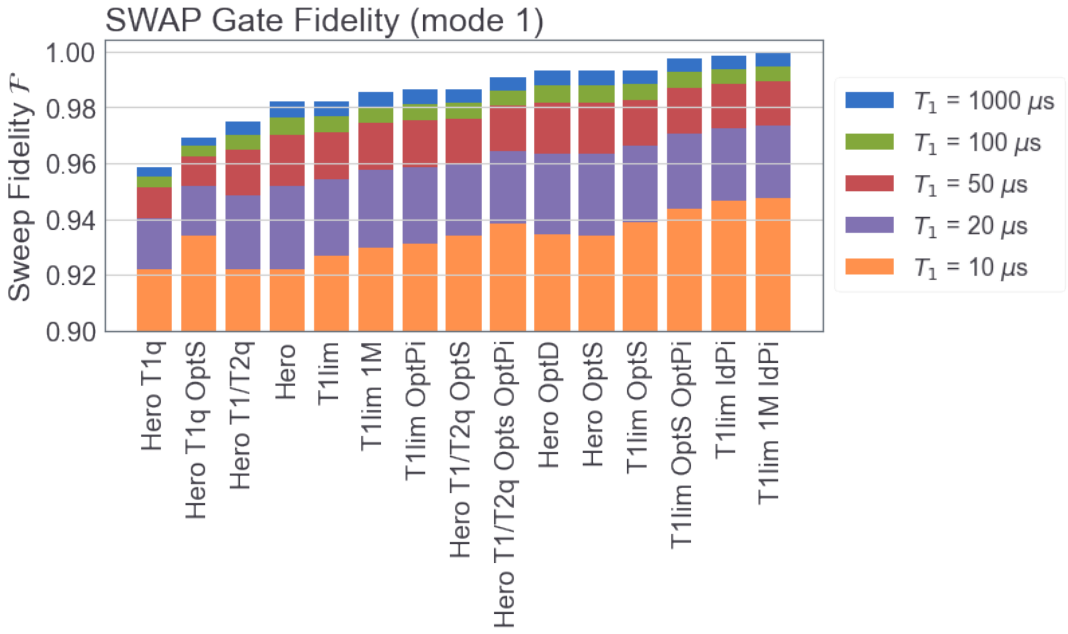


Figure 16: Coherence-time scaling of the SWAP gate for a subset of the simulated qubit decay times T_1^q. See list in 5.4 for an explanation of the abbreviations. In these simulation 4 phonon modes of dimension 2 were present.

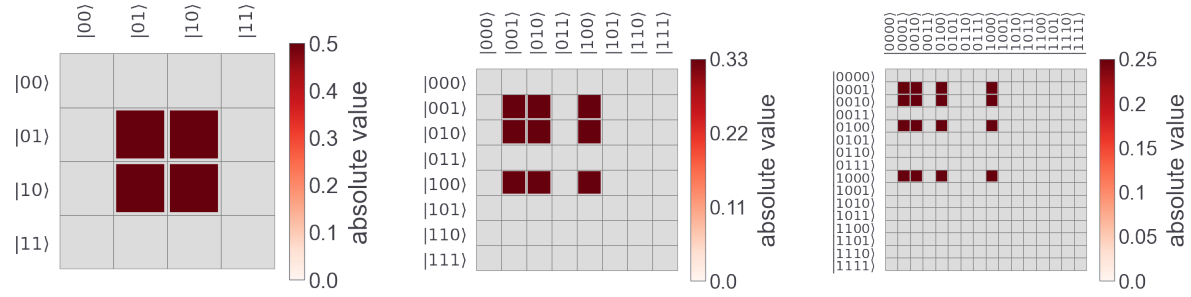
6 Simulation Results

This section contains the carried out simulations on entangled state generation and tomography. If not specified differently, the used simulation parameters are as listed in Table 6. Note that for this section, the π-pulse amplitude is optimized if not mentioned, as opposed to the SWAP parameters which are only optimized if specifically indicated.

number of phonon modes	4
phonon mode Hilbert space dimension	2
tomography phase ϕ	0
tomography wait time t_{wait}	2 µs
coherence times	hero sample

Table 2: Default simulation parameters. These values were used if not stated differently. The *tomography phase* ϕ indicates an additional rotation angle for the tomography coordinate system ($x' = \phi, y' = \phi + \pi/2$).

In order to get a feeling for the depicted Hinton-diagrams, the ideal 2,3 and 4 phonon W-states are shown in Figure 17. For non-ideal states, an additional plot is need to depict the complex phase of the density matrix entries.

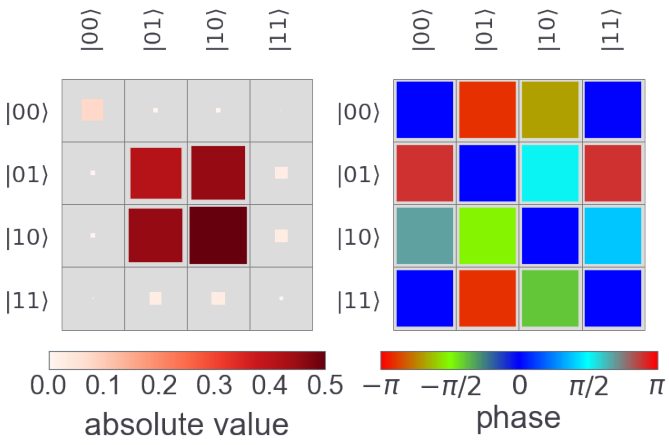


(a) N=2, $\mathcal{N}_N = 0.5$ (b) N=3, $\mathcal{N}_N = 0.471$ (c) N=4, $\mathcal{N}_N = 0.462$

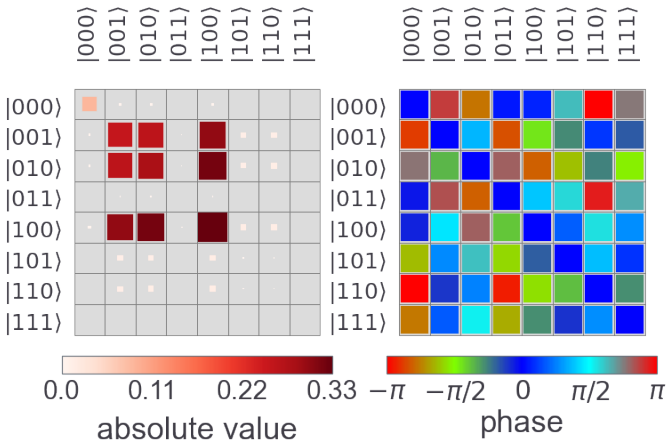
Figure 17: Hinton diagrams of ideal W-states for N=2,3 and 4 phonon modes.

6.1 State Generation

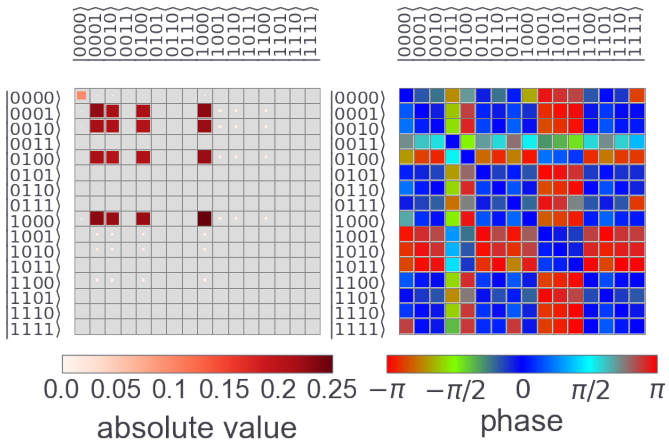
In this section the state generation is analyzed without the tomography. Thus, here we only look at the state preparation fidelity which cannot be assessed in the real experiment. Figure 18 shows the resulting density matrices after state generation for 2,3 and 4 phonon modes. We see that they are different from the ideal states by not having an equal distribution in the populated states. Also, there always is a finite population of the ground state due to decay. Lastly, the individual components of generated states pick up a relative phase due to the different frequency of the phonon modes. These phases are plotted on the right side.



(a) N=2, $\mathcal{F} = 0.914$, $\mathcal{N}_N = 0.416$



(b) N=3, $\mathcal{F} = 0.894$, $\mathcal{N}_N = 0.378$



(c) N=4, $F = 0.882$, $\mathcal{N}_N = 0.358$

Figure 18: Hinton diagrams of the generated W-states with optimized SWAP parameters for N=2,3 and 4 phonon modes. This is the state preparation **without** tomography.

Figure 19 shows the obtained fidelities for the state generation sequence. We see that increasing the Hilbert space dimension of the phonon modes has almost no impact to the fidelity, making simulations with only 2 phonon modes to be sufficient. We also find that with optimized SWAP parameters the generation fidelity can be increased by roughly 0.4%. Lastly, by reversing the sequence (populate highest phonon mode first) it is shown that ideally crossing phonon modes should be minimized as much as possible to get high fidelities.

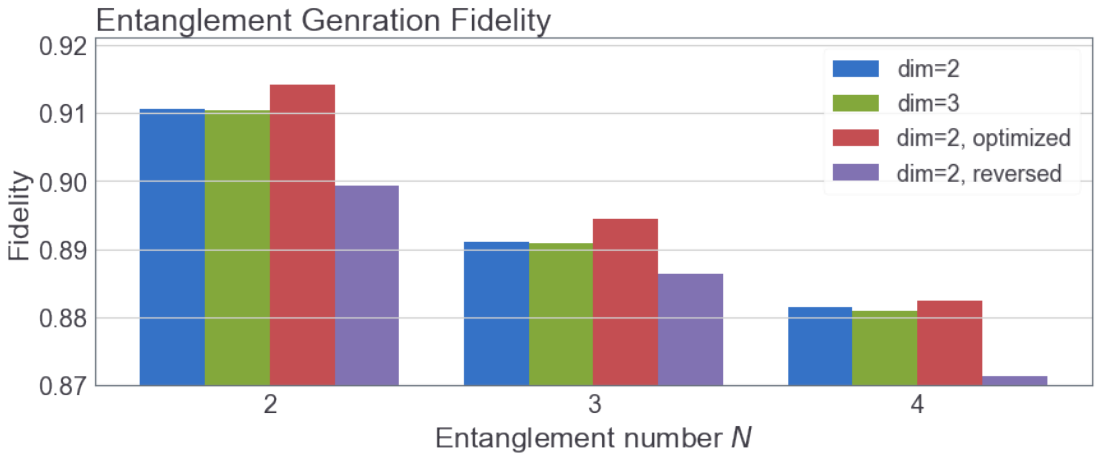


Figure 19: State generation fidelity for different settings. "dim" indicates the dimension of the phonon mode Hilbert space. We can see that the phonon mode dimension does not impact the generation fidelities considerably. This legitimizes only using dim=2 for the simulations. "reversed" describes the state generation, where the phonon mode with the highest frequency was populated first (and all others passed by the qubit as a consequence). We see that this leads to lower state preparation fidelity. Thus the crossing of non-target phonon modes should be avoided.

As a last analysis for the state preparation, Figure 20 shows the fidelity scaling with the system coherence time for both T_1-limited and hero sample ratio. We can see that for long coherence times the fidelities are independent of the specific scaling relation. In this region, the pulse shapes of the applied gates limit the achievable fidelity.

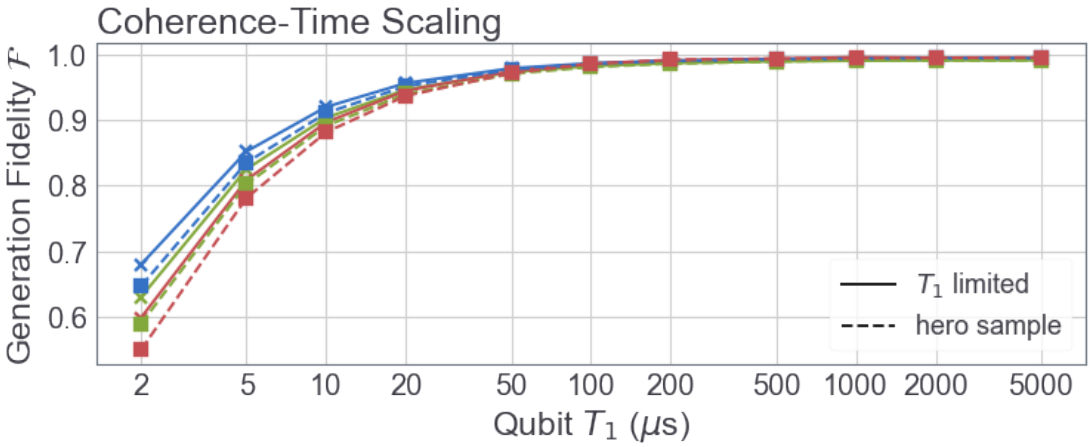
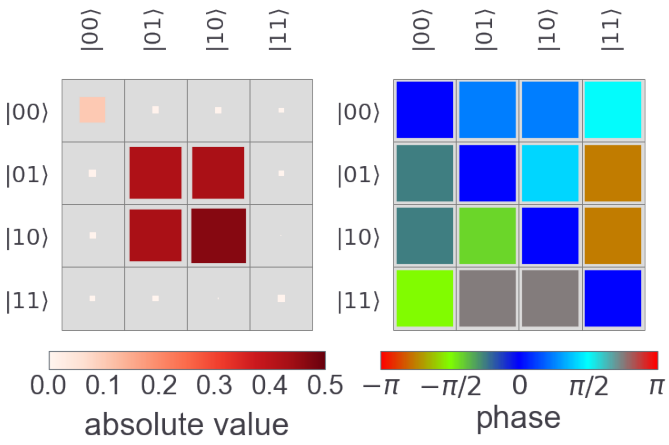


Figure 20: Coherence-time scaling of the entanglement generation sequence. Here we compare the T_1-limited ratio to the hero sample ratio for N=2,3 and 4 phonon modes. We can see that for long coherence times the two scalings converge, indicating that the specific scaling ratio does not impact the protocol fidelity for large enough coherence times. In this, the pulse shapes of the applied gates limit the achievable fidelity.

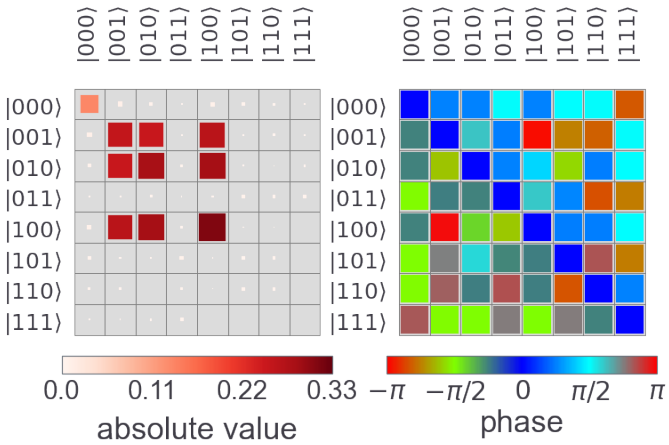
6.2 State Tomography

In this section the implemented state tomography sequence is tested with various parameters. Note that "optimized" for this section means that the SWAP parameters were optimized, but **not** for the state generation. Figure 21 shows the reconstructed density matrices for initialized ideal W-states. Figure 22 shows the reconstructed density matrix for the generated W-states. We can see that for the generated states and tomography a considerable amount of the population has already decayed back to the ground state.

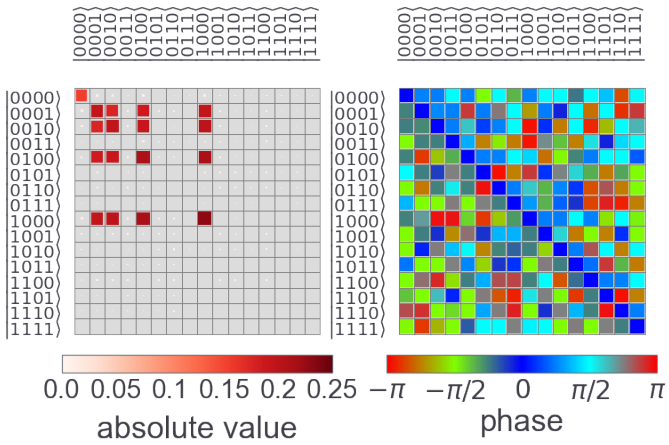
Figure 23 now shows the achievable tomography fidelity for both the ideal initial state and the generated state. We observe that for N=3 and 4 for the ideal initial state the optimization does not increase the achievable fidelity considerably (although it does also not decrease). My explanation for this behavior is that the hero sample coherence times the optimization does not influence the tomography greatly. The increase for the generation then mostly comes from the optimized π-pulse and **not** the optimized SWAP parameters (see Section 8.1 for more details).



(a) N=2, $\mathcal{F} = 0.874$, $\mathcal{N}_N = 0.376$

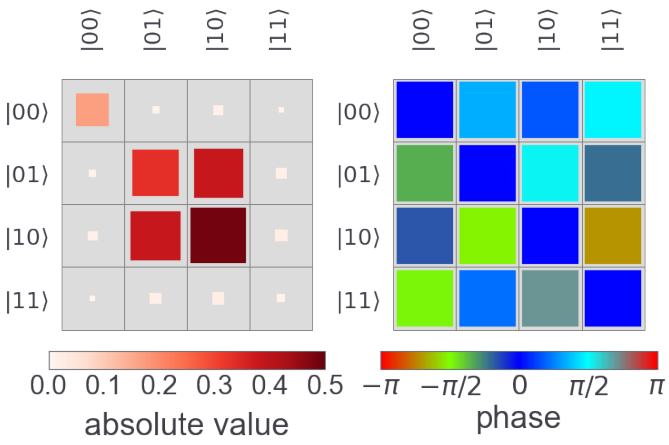


(b) N=3, $\mathcal{F} = 0.821$, $\mathcal{N}_N = 0.319$

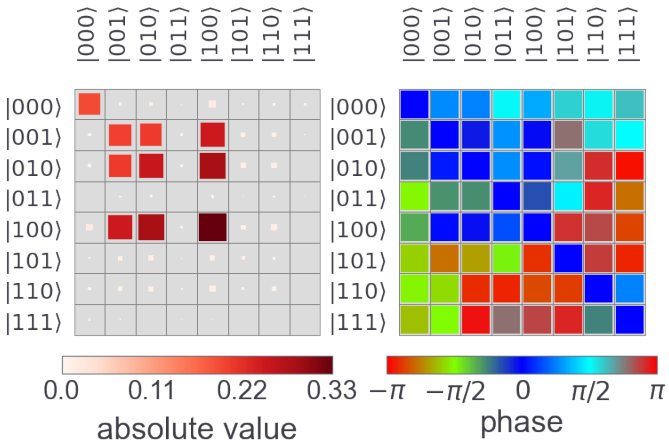


(c) N=4, $\mathcal{F} = 0.789$, $\mathcal{N}_N = 0.298$

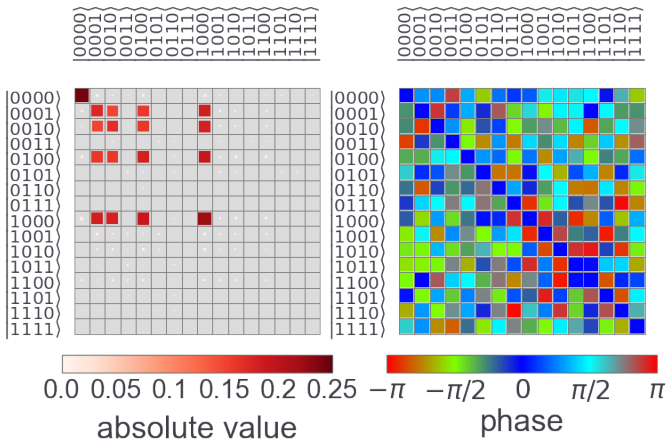
Figure 21: Hinton diagrams of the tomography results for ideal W-states with N=2,3 and 4 phonon modes. This is the tomography **without** state preparation. The SWAP parameters were optimized.



(a) N=2, $\mathcal{F} = 0.793$, $\mathcal{N}_N = 0.302$

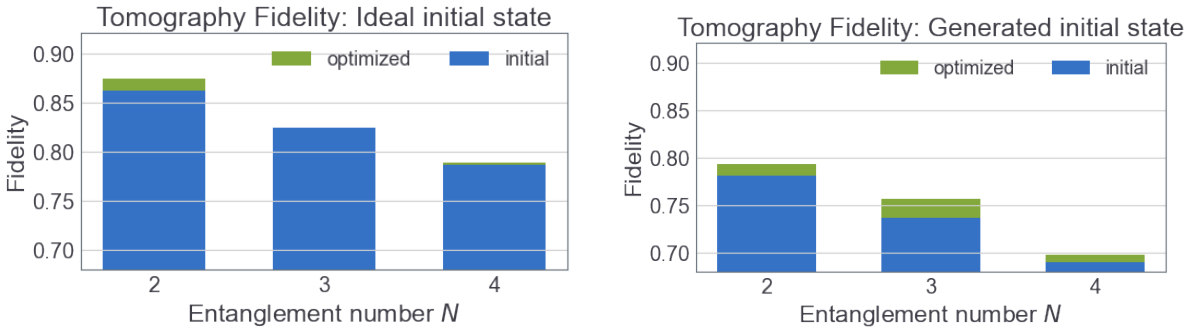


(b) N=3, $\mathcal{F} = 0.757$, $\mathcal{N}_N = 0.266$



(c) N=4, $\mathcal{F} = 0.698$, $\mathcal{N}_N = 0.224$

Figure 22: Hinton diagrams of the tomography results for the generated W-states with N=2,3 and 4 phonon modes. This is state preparation **and** tomography combined. The SWAP parameters were optimized.



(a) Ideal state

(b) Generated state

Figure 23: Achievable tomography fidelities for $N = 2,3$ and 4 phonon modes (hero sample parameters). We can see that the optimization of the SWAP gate leads to a fidelity increase in the 1-2% range. Note that here the initial state was **without** optimized π-pulse.

Next, figure 24 shows the scaling behavior of the tomography sequence with coherence times. We can see that even for very long coherence times the achievable fidelity is not approaching 1.0. This means that the fidelity is limited due to other effects such as the pulse shapes of the SWAP gates and the π-pulse. An additional reason could also be the loss of population to crossed phonon modes when swapping. Figure 25 shows the scaling of the N-Negativity $\mathcal{N}_N$ for the tomography and generation sequence.

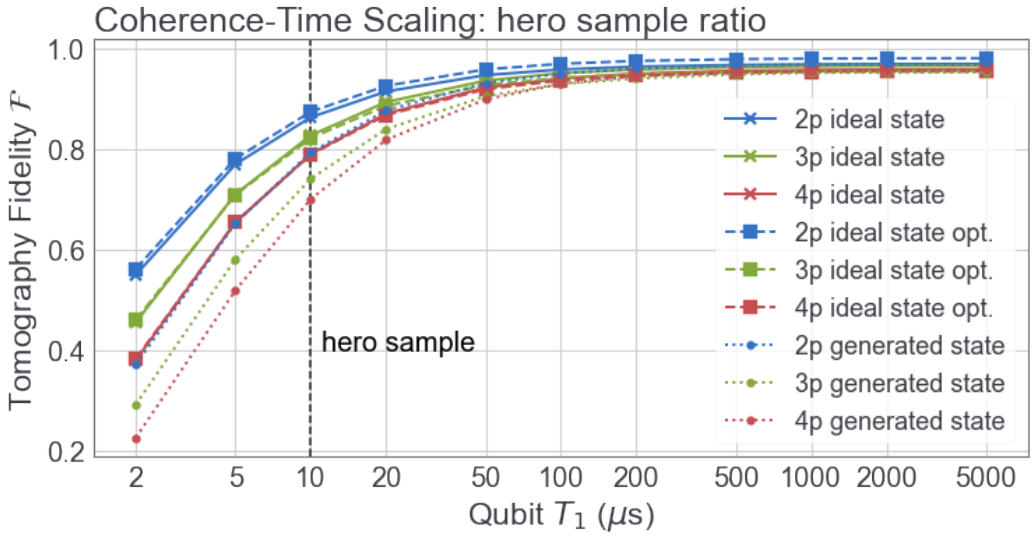


Figure 24: Coherence-time scaling of the fidelity for the tomography sequence. Here we compare the tomography of an ideal initialized state to tomography **and** state generation for the hero sample ratio. "opt" indicates that the SWAP parameters were optimized.

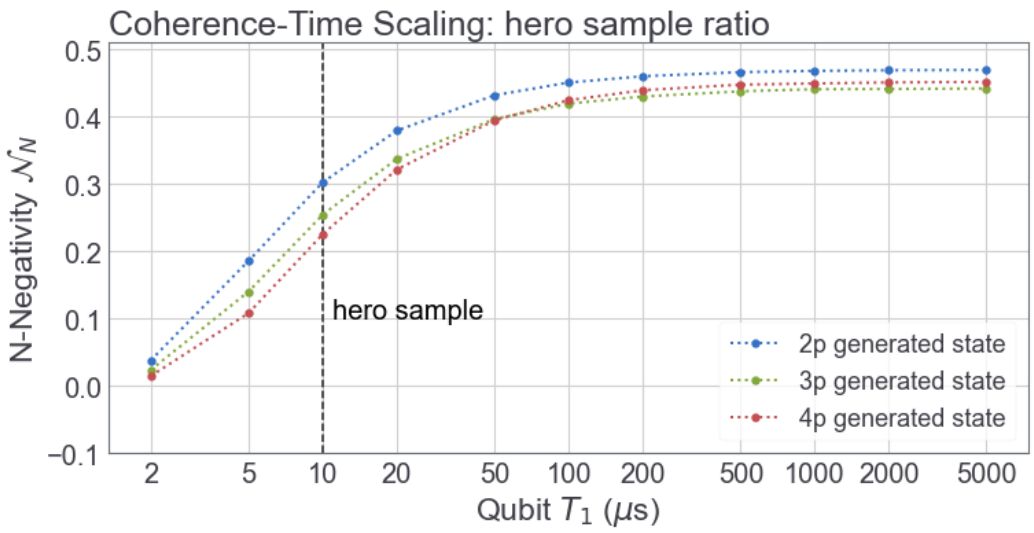


Figure 25: Coherence-time scaling of the N-Negativity $\mathcal{N}_N$ for the generation & tomography sequence.

Phonon Number and Dimension impact

Here I wanted to analyze what effect the phonon number in the simulation (higher frequency modes) and the phonon mode Hilbert space size has on the tomography fidelities. Figure 26 shows a Sweep over these two dimensions for the 2-phonon tomography. We see that the fidelity decrease for Hilbert space dimensions > 2 is approximately 0.2%.

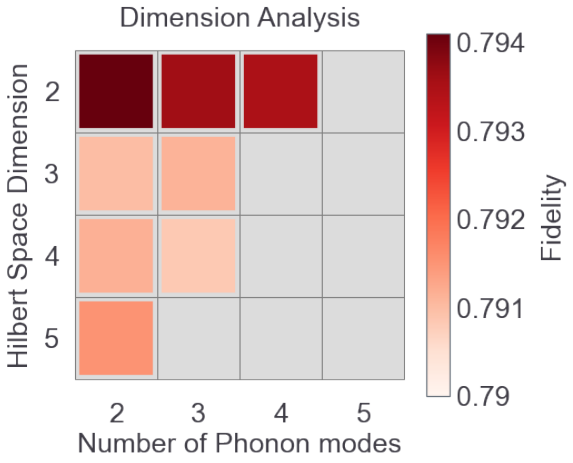


Figure 26: impact on the 2-phonon tomography fidelity (mode 1 and 2) for different phonon number present in simulation and different phonon mode Hilbert space size. We can see that the impact of the phonon mode number and Dimension on the fidelity is in the 0.1%, justifying only simulating with Hilbert space size 2 and 4 phonon modes. Note that for the grey boxes the simulation was not carried out.

Tomography Phase Sweep

Here, the tomography phase ϕ is swept in order to see what effect it has on the reconstructed state. This means that the phases set for the σ_x and σ_y rotation axes are rotated with an angle ϕ relative to the initial value $\phi_x = 0, \phi_y = \pi/2$. The sweep is carried out both for an initialized ideal W-state and for the generated state. We can see that for the ideal state with maximum fidelity the phase does not have an influence on the tomography fidelity. This changes for the generated state, where the fidelity has a local minimum and maximum. My intuitive explanation for this behavior is that for the error prone generated state the tomography phase can be used to "align" the infidelity induced by the tomography to the non-ideal generated state. With this, we can pick a measurement basis that mitigates the stacking of generation plus tomography infidelities.

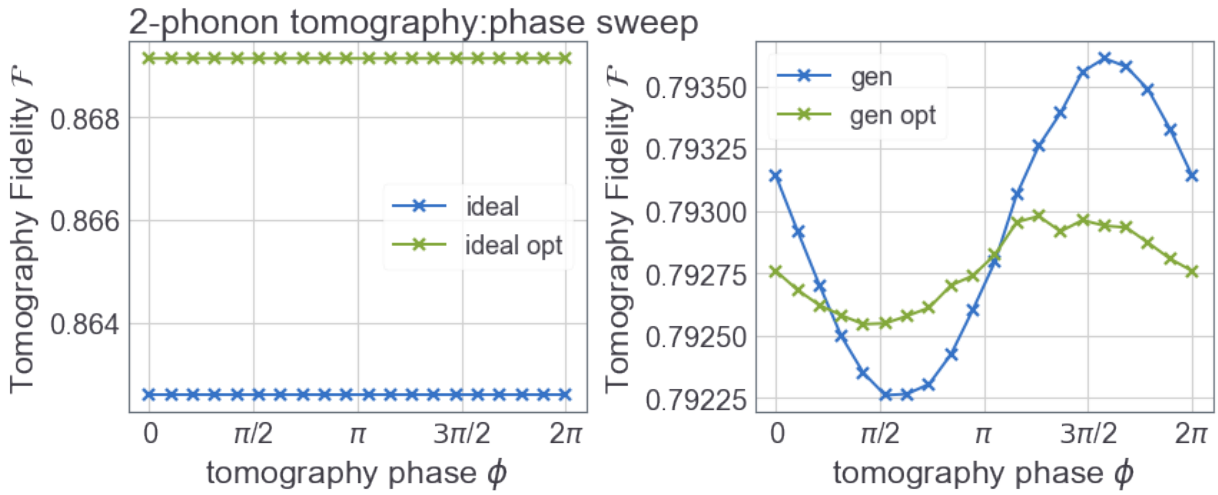


Figure 27: Tomography phase ϕ sweep for 2-phonon tomography. (a) ideal initialized W-state and (b) generated W-state. "opt" indicates that the simulation was carried out with optimized SWAP parameters. We see that the phase has no impact on the ideal state tomography (as theoretically expected) as opposed to the generated state, where we can see local maxima and minima of the fidelity. I explain this phase dependence by the non-ideal character of both generation and tomography: By choosing the phase correctly, we can "align" the infidelities of generation and tomography so that they add up most destructively. For the ideal generation we have a perfectly symmetric input state with respect to the tomography phase, therefore we cannot "align" the infidelity. Note that due to the periodicity of the phase the obtained fidelities are 2π-periodic.

Tomography wait time t_{wait} Sweep

In Figure 28 the tomography wait time t_{wait} is swept for 2-phonon tomography. Two different regimes are separated: (1) The scaling from very low times to 1 µs, where decay is not a dominant effect. (2) From 1 to 10 µs, where decay becomes important. We can see that for the longer wait times, the achievable fidelity decreases nearly linearly with the wait time, showing the impact of decay for long sequence times. In the short time regime, the obtained data is rather surprising: The fidelity fluctuates in the 1-2% range depending on the wait time. Since decay nor decoherence is a dominant factor in this regime, the effect must be depending on the different phase oscillation speeds of the phonon modes. Figure 29 shows a high resolution sweep for the wait times up to 0.02 µs. From this we can see that the fidelity fluctuates up to 1% in magnitude with frequencies in the tens of MHz range. Thus, in the real experiment the wait time should be tuned to acquire a local maximum.

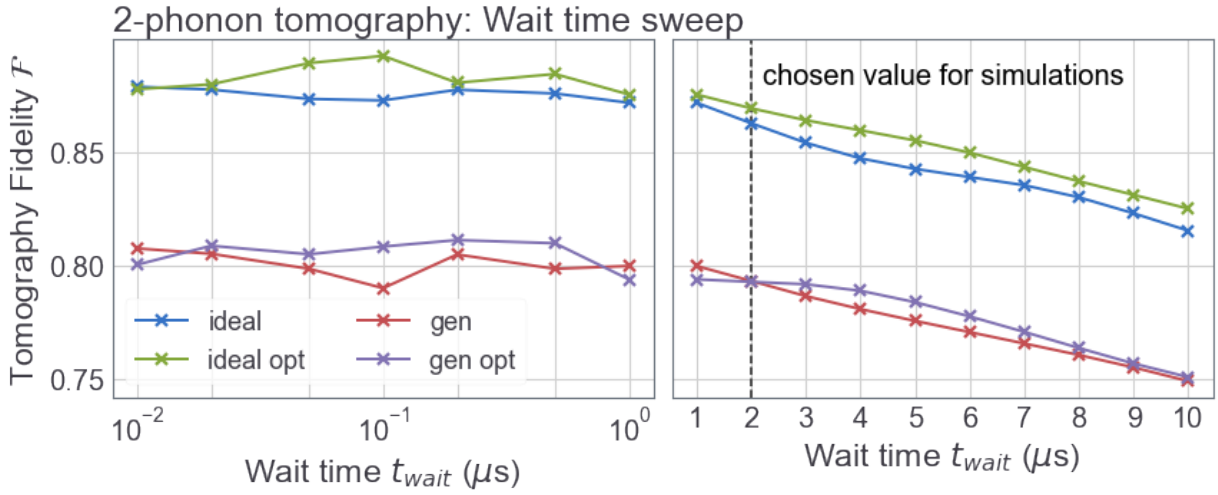


Figure 28: Tomography wait time t_{wait} sweep for 2-phonon tomography. (a) logarithmic wait time increase from 0.01 to 1 µs (b) linear increase from 1 to 10 µs. "ideal" indicates an initialized ideal W-state, "gen" the generated one. "opt" indicates that the simulation was carried out with optimized SWAP parameters. We can see that below 1 µs the achieved fidelity is fluctuating rather strongly. Since these times are much shorter than the coherence time this is not an effect connected to decay or decoherence, but could be connected to the different rotating speeds of the different phonon modes. Maybe specific phase relations during SWAP gates can increase or decrease the state reconstruction fidelity. For wait times $>$ 1 µs the fidelity seems to decrease linearly with the wait time. This effect is related to the decay and decoherence of the system.

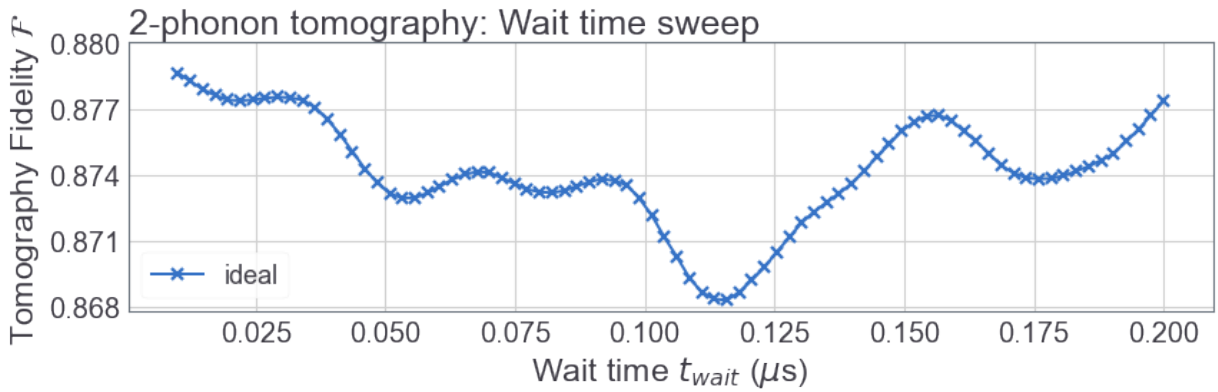


Figure 29: Tomography wait time t_{wait} sweep for 2-phonon tomography with an initialized ideal W-state and high temporal resolution. We can see that the achieved Fidelity does fluctuate up to 1% with a frequency in the range of tens of MHz.

6.3 Entanglement

After analyzing generation and tomography with the fidelity as the measurement metric, we also want to look at the entanglement in the created states. For this we calculated the proposed N-negativity $\mathcal{N}_N$ entanglement measure described in Section 3.1. Figure 30 shows the obtained N-negativity $\mathcal{N}_N$ for the optimized state generation and tomography.

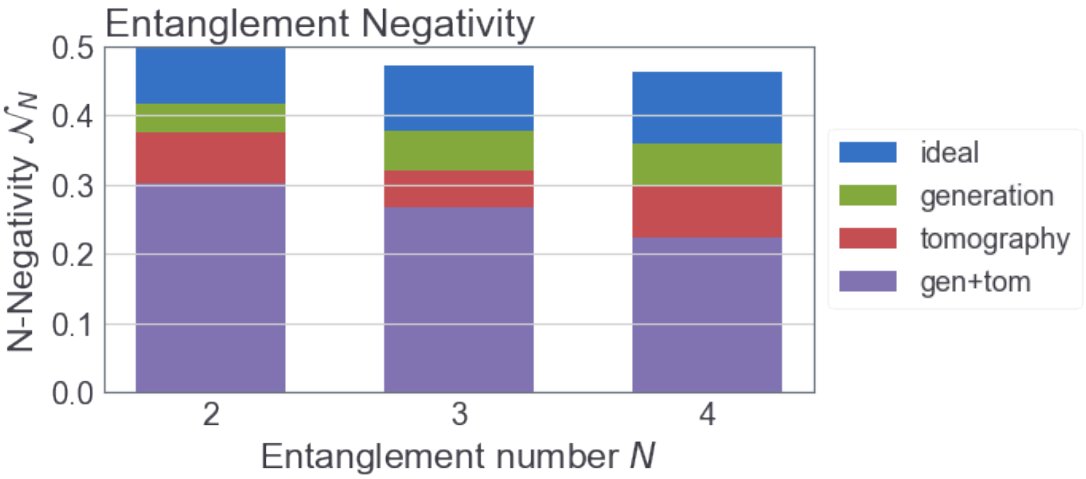


Figure 30: N-negativity $\mathcal{N}_N$ for generated W-state, tomography of the ideal W-state and tomography of the generated W-state for N=2,3 and 4. The SWAP parameters are optimized. In blue, the ideal entanglement number of the respective W-state is shown (calculated for the ideal W-state).

7 Towards a two-phonon gate

7.1 Available Gates

In the current ARQC-System we can use the following types of gates for the system containing the qubit and phonon modes (let us just consider one mode for this discussion):

1. **Single-qubit-Gates** U^q_{single}. This family of gates allows for full control of the qubit state by making arbitrary rotations on the qubit Bloch-sphere. These gates are implemented via the external drive.
2. **Generalized Swap-Gate** $U_{gswap}(\alpha)$. This gate implements a rotation in the $\{|01\rangle, |10\rangle\}$ subspace for arbitrary rotation angle α.
3. **Qubit-Phase Gate** $U^q_{phase}(q)$. This gate changes the phase of states where the qubit is in the Fock 1 state. It can be realized by tuning the qubit with a stark shift such that a relative phase of q with respect to the phonon mode is acquired.
4. **Phonon-Phase Gate** $U^p_{phase}(p)$. Analogous to the qubit-phase gate, but affecting the states where the phonon is in the Fock 1 state.

In order to realize full **Universality** of the available gates, we need to additionally implement a qubit-phonon control gate like the CPHASE or CNOT [12]. Via additional SWAP gates this can then be extended to a phonon-phonon entangling gate.

7.2 Insufficiency of gates (2)-(3) for CPHASE

We can show that even in the two-qubit subspace $\mathcal{H} = (\mathcal{C}^2)^{\otimes 2}$ the operations (2)-(3) do not suffice to realize a CPHASE gate (see Appendix B). Therefore we need to include single qubit gates (1) in order to achieve the operation.

For the two-qubit subspace both the CPHASE and the CNOT gate can be decomposed into the generalized SWAP gate and single-qubit operations:

7.3 CPHASE Decomposition

This decomposition includes a "rotated" variation of the $\sqrt{iSWAP}$ gate. I think it should be possible to decompose this into a standard $\sqrt{iSWAP}$ and single-qubit rotations, but I did not investigate it further since the decomposition is not applicable for our system (see Section 7.5). The full decomposition reads

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2}+\frac{i}{2} & \frac{1}{2}-\frac{i}{2} & 0 \\ 0 & \frac{1}{2}-\frac{i}{2} & \frac{1}{2}+\frac{i}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2}+\frac{i}{2} & \frac{1}{2}-\frac{i}{2} & 0 \\ 0 & \frac{1}{2}-\frac{i}{2} & \frac{1}{2}+\frac{i}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & 0 & 0 & i \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -i \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

7.4 CNOT Decomposition

This composition was found in the QuTip source code [13] and achieves a CNOT gate with various single-qubit gates and two $\sqrt{\text{iSWAP}}$ gates. The total sequence is rather involved, therefore we first define the necessary single-qubit operations to write the decomposition in a more compact form:

$$R_x(\theta) = e^{i\frac{\theta}{2}X} = \cos\left(\frac{\theta}{2}\right) I + i \sin\left(\frac{\theta}{2}\right) \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \cos\frac{\theta}{2} & i\sin\frac{\theta}{2} \\ i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix}$$
$$R_y(\theta) = e^{i\frac{\theta}{2}Y} = \cos\left(\frac{\theta}{2}\right) I + i \sin\left(\frac{\theta}{2}\right) \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \begin{bmatrix} \cos\frac{\theta}{2} & \sin\frac{\theta}{2} \\ -\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix}$$
$$R_z(\theta) = e^{i\frac{\theta}{2}Z} = \cos\left(\frac{\theta}{2}\right) I + i \sin\left(\frac{\theta}{2}\right) \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} \cos\frac{\theta}{2} + i\sin\frac{\theta}{2} & 0 \\ 0 & \cos\frac{\theta}{2} - i\sin\frac{\theta}{2} \end{bmatrix}$$

We now write single-qubit operations on the qubit with the superscript q, e.g. $R_x^q(\theta)$ and on the phonon mode with superscript p, e.g. $R_y^p(\theta)$. Additionally, we introduce the unitary $U_{\text{phase}}^{\text{glob}}(\phi)$, which just adds a global phase. Then the full decomposition reads

$$R_y^q\left(-\frac{\pi}{2}\right) R_x^q\left(\frac{\pi}{2}\right) R_x^p\left(-\frac{\pi}{2}\right) \sqrt{\text{iSWAP}} R_x^q(\pi) \sqrt{\text{iSWAP}} R_y^q\left(\frac{\pi}{2}\right) U_{\text{phase}}^{\text{glob}}\left(\frac{\pi}{4}\right) R_z^q\left(\frac{\pi}{2}\right) U_{\text{phase}}^{\text{glob}}\left(\frac{3\pi}{2}\right) \tag{33}$$

Since we are usually not interested in global phases, the sequence can be shortened a bit. Up to global phase, the following protocol still realizes a CNOT gate

$$R_x^p\left(\frac{3\pi}{2}\right) R_y^q\left(\frac{3\pi}{2}\right) R_x^q\left(\frac{\pi}{2}\right) \sqrt{\text{iSWAP}} R_x^q(\pi) \sqrt{\text{iSWAP}} R_y^q\left(\frac{\pi}{2}\right) R_z^q(\pi) \tag{34}$$

7.5 The Bosonic Problem

Because of the bosonic nature of the phonon modes we need to check whether the usage of the two-qubit subspace $\mathcal{H} = (\mathcal{C}^2)^{\otimes 2}$ is appropriate for our problem. For this we include the next higher Fock level $|2\rangle$ of the phonon mode in the decompositions. Unfortunately, this shows that we cannot use the two above proposed sequences to realize a controlled gate. The problem lies in the usage of the SWAP-operation: If we start in the $|11\rangle$ state (first entry qubit, second entry phonon), a generalized SWAP gate always leads to information leakage into the $|02\rangle$ state. Thus our code space is not closed under the generalized swap and this operation **cannot** be used to realize a controlled qubit-phonon gate. Figure 31 shows the operator matrix of the iSWAP and $\sqrt{\text{iSWAP}}$ gate with non-vanishing $|02\rangle - |11\rangle$ elements.

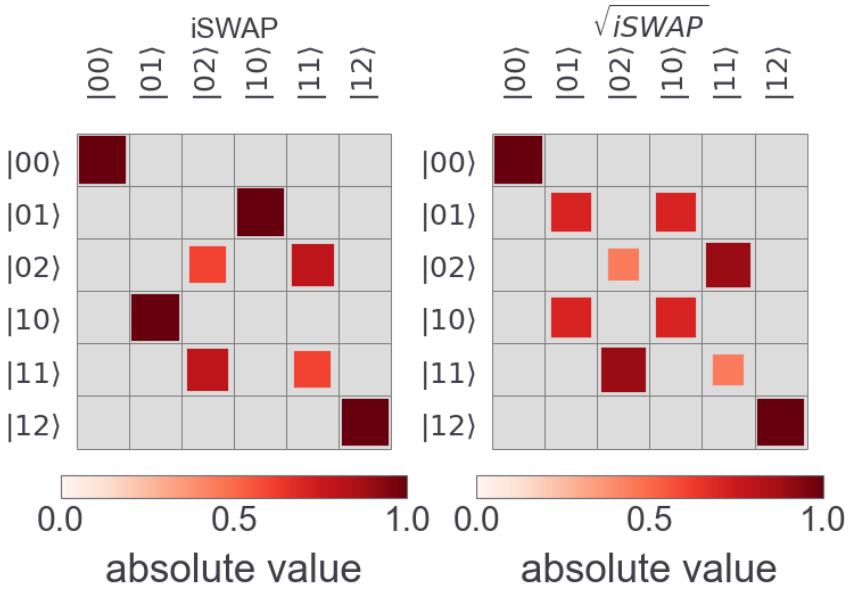


Figure 31: The SWAP-type gates have elements that lead to leakage out of the code-space. Here, the first entry of the states indicates the qubit population, the second entry the phonon population.

This can be made even more clear if the state $|11\rangle$ is evolved under the two gates. Figure 32 shows the occupation probability after the iSWAP and $\sqrt{\text{iSWAP}}$ gates.

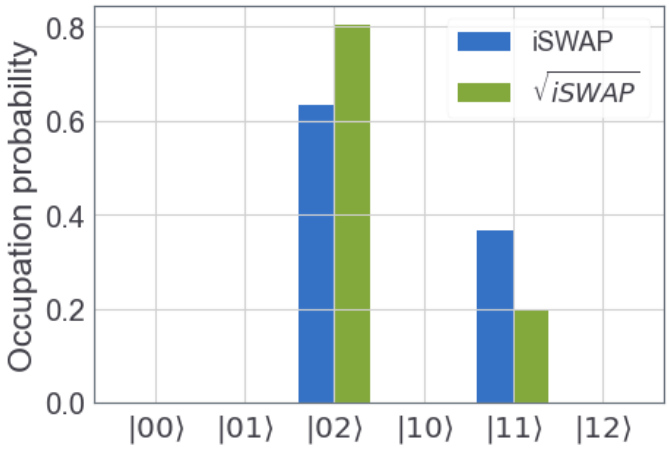


Figure 32: Occupation Probability of the system states after the iSWAP or $\sqrt{\text{iSWAP}}$ gate acted on the state $|11\rangle$.

Therefore, we have to consider other mechanisms to realize a two-qubit gate.

7.6 Two-phonon gate possibilities

Since a direct decomposition with the SWAP gates is not possible, we need to consider other approaches to realize a controlled gate between two phonon modes. The general idea here is to realize a controlled gate between the qubit and a phonon mode first and then modify it to a phonon-phonon gate via additional swapping. I have found the following methods to be candidates for this purpose:

1. **Dispersive Regime**. By tuning the qubit frequency far away from the phonon mode, a dispersive interaction between qubit and phonon mode can be achieved. By controlling this interaction, a CPHASE gate can be realized. We start with the dispersive hamiltonian (holds if $\Delta = |\omega^q - \omega^p| \gg g$) in the rotating frame of the qubit and the phonon:

$$H_\mathrm{d} = \frac{g^2}{\Delta} a^\dagger a b^\dagger b. \tag{35}$$

For the computational states between qubit and phonon we thus find

$$\begin{aligned} H_\mathrm{d}|g0\rangle &= 0 \\ H_\mathrm{d}|g1\rangle &= 0 \\ H_\mathrm{d}|e0\rangle &= 0 \\ H_\mathrm{d}|e1\rangle &= \frac{g^2}{\Delta}|e1\rangle \end{aligned}$$

Thus the time evolution under this hamiltonian is given by

$$U_{\mathrm{d}}(t) = \exp(-iH_{\mathrm{d}}t) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \exp\left(-i\frac{g^2}{\Delta}t\right) \end{bmatrix}, \tag{36}$$

By choosing the time t we can then achieve an arbitrary CPHASE operation between the qubit and the phonon. For this gate, the gate time is inversely proportional to the detuning. However, the detuning needs to be sufficiently large for the dispersive regime condition to hold. Thus it requires long coherence times for high fidelity.

2. **Sideband iSWAP** [14]. This gate is implemented by tuning the qubit on resonance with the $|e1\rangle - |f0\rangle$ transition, where $|f\rangle$ is the second excited state of the transmon and 0 and 1 indicate the fock state of a phonon mode. By performing two consecutive iSWAP gates, a CPHASE gate between qubit and the phonon mode is realized. In our system this implementation is currently not possible, since the tuning range via the stark-shift is not large enough.

3. **Sideband Modulation** [15]. For this gate the qubit frequency is parametrically modulated (e.g. with a time-dependent stark-shift). If the modulation frequency equals the frequency difference between the qubit and the phonon, then these two "virtually" couple. Modulating the qubit frequency at the frequency of the $|e1\rangle - |f0\rangle$ transition might allow the realization of a CPHASE gate. It can be shown that the achievable coupling strength between qubit and phonon is $\approx g/2$, where g is the initial coupling. Therefore the gate is most likely to be slower than the direct *Sideband iSWAP*.

4. **The Cross-Resonance Gate** [16]. This gate is implemented by applying a drive to the qubit with the resonance frequency of a phonon mode. The gate operation is then realized via the qubit-phonon interaction, effectively implementing a CNOT gate (up to single qubit rotations). Unfortunately, this gate has a maximum coupling strength of $g/4$, where g is the initial qubit-phonon coupling, making the sequence too long for current device parameters (hero sample). Also, it is not clear if the gate only applies quibt-qubit like systems and there is leakage out of the code space for our setting.

5. **Optimal Control** [17]. Optimal Control uses machine learning algorithms to find the most ideal composed gate sequence out of an available set of sub-gates. It is a heuristic approach and thus is not guaranteed to succeed.

6. **Phonon Blockade** [18]. Phonon blockade describes the process, that if one phonon is present in an oscillator, the excitation of a second one is prevented. If we could manage to introduce a phonon-blockade process into our system, then the code-space

cannot be left (= no leakage) and thus the above described decompositions are valid again. Phonon blockade could potentially be achieved by coupling the mechanical resonator quadratically to a 2-level system (e.g. another qubit) and using the energy shift of the dressed states.

8 Challenges & Improvements

There are some open ends I was not able to tackle during the time of the Semester Thesis:

8.1 SWAP gate optimization

As of now, the SWAP gate optimization was done for the full sequence π-pulse + SWAP. Thus the parameter optimization also accounted for the non-ideal π-pulse. But this π-pulse is not present in the tomography sequence, thus the ideal parameters for tomography could potentially differ from the found optimum. To solve this issue, the optimization should be done again for only the SWAP gate without the preceding π-pulse. However, it is questionable if all these optimizations are necessary in the first place, since the main limiting factor are still the coherence times. Without better coherence times these fine-tune optimizations will not be impactful.

8.2 Analytical Fidelity Estimation

During the semester project I tried to find an analytic expression for the fidelities of both the W-state generation and the phonon tomography. The idea was that in the limit of short gate times the fidelity for a single sub-gate can be expressed as

$$F(t) = 1 - e^{\Gamma_1 t}, \tag{37}$$

where Γ_1 is the decay time of the system. For high fidelities (or equivalently short times t), the total fidelity of the gate is thus approximately

$$F_{tot} = \prod_{i=1}^{N} F(t_i). \tag{38}$$

Now, The main component of this analysis is that for the SWAP gate, we can think of the excitation being swapped between qubit and phonon, thus the effective decay time is something in between the qubit and phonon decay time. For the full swap the occupation time is $1/2, 1/2$ and therefore the system decay time can be estimated as

$$\Gamma_1^{avg} = \frac{1}{2}\left(\frac{1}{T_1^{qubit}} + \frac{1}{T_1^{phonon}}\right). \tag{39}$$

I applied these calculations to the generation sequence and compared it to the obtained results from simulation, as shown in Figure 33. Although the obtained values somewhat match the simulation results, the analytical expression lacks a solid foundation. At this point it is still more of a heuristic approach. One way to improve it is to account for the fact that during the generation sequence, the phonon-qubit occupation is not 50/50, but more leaning towards the qubit. Thus the fidelities of the estimation would slightly decrease. This could be modelled by taking a weighted sum of the decay rates: $\Gamma_1^{avg} = \frac{w_1}{T_1^{qubit}} + \frac{w_2}{T_1^{phonon}}$, where w_1, w_2 are the weights.

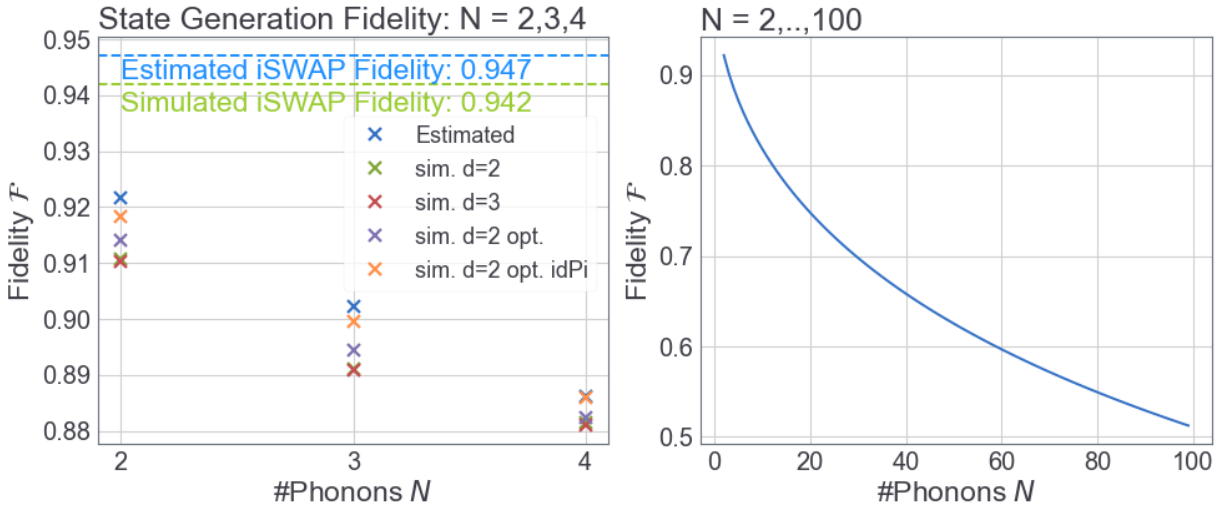


Figure 33: Analytical fidelity estimation for the W-state generation sequence. d here indicates the phonon Hilbert space dimension used for the simulation. "opt" indicates that the SWAP parameters are optimized. "idPi" indicates that an ideal π-pulse was used, i.e. the qubit initialized in the excited state.

8.3 Phonon-Decay

An interesting topic I was not able to look into enough is the decay properties of entangled systems. If more time would have been present, I would have solved the master equation analytically for a system containing two phonon modes. Of special interest is the impact of the initial state on the ground state population rate, i.e. how fast the total system state decays to the ground state. For W-states, there are published calculations [19] which indicate that the ground state decay rate does not increase with the entanglement size. This calculation would also give more insight into which weights to use for the analytical fidelity estimation from above.

8.4 Alternative Tomography Methods

The topic of quantum state tomography and density matrix reconstruction has become increasingly important for the field of quantum information processing. A variety of research has been conducted on this topic and different tomography methods were developed [11]. In this work, I used the method of *Direct Inversion*, which assumes that for some measurement operator A the sample mean $\langle\langle A\rangle\rangle$ is a good estimator for the population mean $\langle A\rangle$. Although this method has been used extensively in the field [20, 21], it also comes with some drawbacks: The Direct Inversion method does not limit the reconstructed density matrices to a subspace of physical results, specifically in the reconstruction process the condition $\mathrm{Tr}[\rho] = 1$ is not enforced, leading to possible unphysical results.

Note that in the simulations conducted in this semester thesis, this was not a problem since through the simulation the population mean $\langle A\rangle$ could directly be obtained, preventing this issue. However, for an actual experiment it could be beneficial to limit the reconstructed states to the space of physical density matrices by implementing a different tomography method based on *Likelihood-Estimation*, e.g. Maximum-Likelihood estimation (MLE) or Bayesian-Mean estimation (BME). Specifically MLE has been successfully used in quantum state tomography and has shown to be a robust alternative to Direct Inversion that ensures the trace condition [22].

8.5 Entanglement Measure

In order to make the certification of the entanglement of the measured states more solid, it could be beneficial to implement a scheme based on entanglement witnesses similar to the one used in [9]. In our setup, the implementation can directly be done in the discrete basis. Some possible witnesses for the three phonon W-state are already treated as an example in the article.

8.6 Two-Phonon Gate

Of course a big open task is to find out how to realize a two-mode gate between two phonon modes. A next step would be to simulate the proposed gates in 7.6 and see which of them could be realized in an actual experiment.

References

[1]Y. Chu, P. Kharel, W. H. Renninger, L. D. Burkhart, L. Frunzio, P. T. Rakich, and R. J. Schoelkopf, "Quantum acoustics with superconducting qubits", Science **358**, 199–202 (2017).

[2]S. Szalay, "Multipartite entanglement measures", Physical Review A **92**, 042329 (2015).

[3]W. Dür, G. Vidal, and J. I. Cirac, "Three qubits can be entangled in two inequivalent ways", Physical Review A **62**, 062314 (2000).

[4]N. Friis, G. Vitagliano, M. Malik, and M. Huber, "Entanglement certification from theory to experiment", Nature Reviews Physics **1**, 72–87 (2019).
[5]S. Haddadi and M. Bohloul, "A brief overview of bipartite and multipartite entanglement measures", International Journal of Theoretical Physics **57**, 3912–3916 (2018).
[6]M. Huber, F. Mintert, A. Gabriel, and B. C. Hiesmayr, "Detection of high-dimensional genuine multipartite entanglement of mixed states", Physical review letters **104**, 210501 (2010).
[7]F.-L. Wu, S.-Y. Liu, W.-L. Yang, S.-M. Fei, and H. Fan, "A Necessary and Sufficient Entanglement Criterion of N-qubit System Based on Correlation Tensor", arXiv preprint arXiv:2205.05323 (2022).
[8]J. Sperling and W. Vogel, "Multipartite entanglement witnesses", Physical review letters **111**, 110503 (2013).
[9]S. Gerke, J. Sperling, W. Vogel, Y. Cai, J. Roslund, N. Treps, and C. Fabre, "Full multipartite entanglement of frequency-comb Gaussian states", Physical review letters **114**, 050501 (2015).
[10]G. Vidal and R. F. Werner, "Computable measure of entanglement", Physical Review A **65**, 032314 (2002).
[11]R. Schmied, "Quantum state tomography of a single qubit: comparison of methods", Journal of Modern Optics **63**, 1744–1758 (2016).
[12]M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2010).
[13]QuTip developers and contributors, *CNOT sqrtiswap decomposition*, (2022) `https://qutip.org/docs/latest/modules/qutip/qip/circuit.html` (visited on 06/25/2022).
[14]R. Naik, N. Leung, S. Chakram, P. Groszkowski, Y. Lu, N. Earnest, D. McKay, J. Koch, and D. Schuster, "Random access quantum information processors using multimode circuit quantum electrodynamics", Nature communications **8**, 1–7 (2017).
[15]J. Strand, M. Ware, F. Beaudoin, T. Ohki, B. Johnson, A. Blais, and B. Plourde, "First-order sideband transitions with flux-driven asymmetric transmon qubits", Physical Review B **87**, 220505 (2013).
[16]C. Rigetti and M. Devoret, "Fully microwave-tunable universal gates in superconducting qubits with linear couplings and fixed transition frequencies", Phys. Rev. B **81**, 134507 (2010).
[17]B. Riaz, C. Shuang, and S. Qamar, "Optimal control methods for quantum gate preparation: a comparative study", Quantum Information Processing **18**, 1–26 (2019).
[18]H. Xie, C.-G. Liao, X. Shang, M.-Y. Ye, and X.-M. Lin, "Phonon blockade in a quadratically coupled optomechanical system", Phys. Rev. A **96**, 013861 (2017).
[19]A. R. Carvalho, F. Mintert, and A. Buchleitner, "Decoherence and multipartite entanglement", Physical review letters **93**, 230501 (2004).
[20]Y.-x. Liu, L. F. Wei, and F. Nori, "Tomographic measurements on superconducting qubit states", Phys. Rev. B **72**, 014547 (2005).
[21]J. M. Chow, *Quantum information processing with superconducting qubits* (Yale University, 2010).

[22]S. Filipp, P. Maurer, P. J. Leek, M. Baur, R. Bianchetti, J. M. Fink, M. Göppl, L. Steffen, J. M. Gambetta, A. Blais, and A. Wallraff, "Two-Qubit State Tomography Using a Joint Dispersive Readout", Phys. Rev. Lett. **102**, 200402 (2009).
[23]Y. J. Bagul and R. M. Dhaigude, "Simple efficient bounds for arcsine and arctangent functions", (2021).

A Entanglement Generation Time Estimation

The total sequence length of the state generation for N Phonon Entanglement is given by the π-pulse length plus the sum of the swap times t_k

$$T_{gen} = T_\pi + \sum_{k=1}^{N} t_k = \frac{4}{\Omega} + \sum_{k=1}^{N} \frac{1}{g_k} \arcsin\left(\frac{1}{\sqrt{N+1-k}}\right) \tag{40}$$

In order to find the dependence of the sequence time on the phonon number $T_{gen} \equiv T_{gen}(N)$ we use an upper bound for $\arcsin(x)$ on $x \in (0,1)$ [23]:

$$\frac{\arcsin x}{x} < \frac{1}{\sqrt{1-x^2}} \qquad x \in (0,1) \tag{41}$$

By plugging in $x = 1/\sqrt{N+1-k}$ we find

$$\begin{aligned}
\frac{1}{g_k} \arcsin\left(\frac{1}{\sqrt{N+1-k}}\right) &< \frac{1}{g_k} \frac{1}{\sqrt{N+1-k}} \frac{1}{\sqrt{1-\frac{1}{N+1-k}}} \\
&= \frac{1}{\sqrt{N+1-k}} \frac{\sqrt{N+1-k}}{\sqrt{N-k}} \\
&= \frac{1}{\sqrt{N-k}}
\end{aligned}$$

If we treat the $k = N$ term separately ($\arcsin(1) = \pi/2$) and assume that the coupling constants for all phonon modes are similar $\forall k : g_k \approx g$ we can write

$$\begin{aligned}
T_{gen} &< \frac{4}{\Omega} + \underbrace{\frac{\pi}{2g}}_{k=N \text{ Term}} + \frac{1}{g}\left(\frac{1}{\sqrt{1}} + ... + \frac{1}{\sqrt{N-1}}\right) \\
&= \frac{4}{\Omega} + \frac{1}{g}\left(\frac{\pi}{2} + \sum_{n=1}^{N-1} \frac{1}{\sqrt{n}}\right)
\end{aligned}$$

We can now use the relation $m \geq 2$: $2\sqrt{m} - 2 < \sum_{n=1}^{m} \frac{1}{\sqrt{n}} < 2\sqrt{m} - 1$ and plug this in for $m = N - 1$ to obtain

$$\begin{aligned} T_{ges} &< \frac{4}{\Omega} + \frac{1}{g}\left(\frac{\pi}{2} + 2\sqrt{N-1} - 1\right) \\ &= \left(\frac{4}{\Omega} + \frac{1}{g}\left(\frac{\pi}{2} - 1\right)\right) + \frac{2}{g}\sqrt{N-1} \\ &\approx (0.5 + 1.2\sqrt{N-1})\ \mu\text{s} \end{aligned}$$

Thus we have found that the generation sequence is $\mathcal{O}(\sqrt{N})$.

In a different approach we can also replace the sum with an integral $\sum_{k=1}^{N} \approx \int_1^N d\tilde{N}$ and find

$$\begin{aligned} T_{gen} = T_\pi + \sum_{k=1}^{N} t_k &= \frac{4}{\Omega} + \sum_{k=1}^{N} \frac{1}{g_k} \arcsin\left(\frac{1}{\sqrt{N+1-k}}\right) \\ &\approx \frac{4}{\Omega} + \int_1^N d\tilde{N} \frac{1}{g} \arcsin\left(\frac{1}{\sqrt{\tilde{N}}}\right) \\ &= \frac{4}{\Omega} + \frac{1}{g}\left(\sqrt{N-1} + N \arcsin\left(\frac{1}{\sqrt{N}}\right)\right) \end{aligned}$$

.

This again leads to the generation time being of order $\mathcal{O}(\sqrt{N})$, but is a better estimate to the total generation time T_{gen}.

B Insufficiency of using $U_{\text{gswap}}, U^q_{\text{phase}}$ and U^p_{phase} for a CPHASE gate

We can write the Unitaries of these operations in the qubit and phonon system with basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ where the first entry denotes the qubit:

$$U_{gswap}(\alpha) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\alpha) & -i\sin(\alpha) & 0 \\ 0 & -i\sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$U^{q}_{phase}(q) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{iq} & 0 \\ 0 & 0 & 0 & e^{iq} \end{bmatrix}$$

$$U^{p}_{phase}(p) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & e^{ip} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{ip} \end{bmatrix}$$

In order to have a **universal set of gates** for our system, we need to show that we can realize a Control-Phase Gate (CPhase-Gate), since all single-qubit gates are achievable using a SWAP and the qubit-control.

$$\text{CPhase} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{i\varphi} \end{bmatrix} \tag{42}$$

for some $\varphi \neq 0$. We will now show, that the three operations U_{gswap}, U^{q}_{phase} and U^{p}_{phase} are not sufficient to realize a CPhase-Gate.

Proof

Let us assume that a decomposition of the CPhase-gate into U_{gswap}, U^{q}_{phase} and U^{p}_{phase} exists. Write this decomposition as

$$\text{CPhase} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{i\varphi} \end{bmatrix} = \prod_{k} U_k \qquad U_k \in \{U_{gswap}, U^{q}_{phase}, U^{p}_{phase}\} \tag{43}$$

We now introduce indices m and n that count the occurrences of U^{q}_{phase} and U^{p}_{phase} respectively. For each occurrence, denote the corresponding phase parameter as q_m and p_n, such that for the m-th occurrence of U^{q}_{phase} the unitary is $U^{q}_{phase}(q_m)$.

We know notice that for arbitrary rotation angle α we have $\det(U_{gswap}(\alpha)) = 1$. For the Qubit- and Phonon-Phase Gate we have

$$\det(U^q_{phase}(q_m)) = e^{iq_m} e^{iq_m} = e^{2iq_m} \quad \det(U^p_{phase}(p_n)) = \cdots = e^{2ip_n} \tag{44}$$

Thus for the total decomposition

$$\begin{aligned}
\det(\text{CPhase}) = e^{i\varphi} &\overset{\text{decomp.}}{=} \det\left(\prod_k U_k\right) \\
&= \prod_k \det(U_k) \\
&= \prod_m \det(U^q_{phase}(q_m)) \prod_n \det(U^p_{phase}(p_n)) \\
&= \prod_m e^{2iq_m} \prod_n e^{2ip_n} \\
&= \exp\left(2i\left(\sum_m q_m + \sum_n q_n\right)\right)
\end{aligned}$$

From this follows a necessary condition for the sum of all q_m and p_n:

$$\varphi \overset{!}{=} 2\left(\sum_m q_m + \sum_n q_n\right) \mod 2\pi \tag{45}$$

.

Next, we observe that the $\{|11\rangle\}$ subspace is left unchanged by the U_{gswap} operation. Thus only the Qubit- and Phonon-Phase gate can contribute to the entry $e^{i\varphi}$ of the CPhase gate. This leads to a second condition

$$\begin{aligned}
e^{i\varphi} &= \prod_k (U_k)_{22} = \prod_m (U^q_{phase}(q_m))_{22} \prod_n (U^p_{phase}(p_n))_{22} \\
&= \prod_m e^{iq_m} \prod_n e^{ip_n} \\
&= \exp\left(i\left(\sum_m q_m + \sum_n p_n\right)\right)
\end{aligned}$$

This leads to another necessary condition for the sum of the q_m and p_n

$$\varphi \stackrel{!}{=} \left(\sum_m q_m + \sum_n q_n \right) \mod 2\pi \tag{46}$$

From these two condition it directly follows that $\varphi = 0 \mod 2\pi$. Therefore the only CPhase gate that can be achieved is the trivial Identity gate ($\varphi = 0$). $\square$