Quantum Information Processing for Quantum Simulations

Traitement quantique de l'information pour les simulations quantiques

par

Camille Le Calonnec

Thèse présentée au département de physique

en vue de l'obtention du grade de docteur ès sciences (Ph.D.)

FACULTÉ des SCIENCES

UNIVERSITÉ de SHERBROOKE

&

ÉCOLE DOCTORALE des SCIENCES CHIMIQUE

CESQ – ISIS (UMR 7006)

UNIVERSITÉ de STRASBOURG

Sherbrooke, Québec, Canada, 12 avril 2024

Le Vendredi 12 avril 2024

le jury a accepté la thèse de Mlle Camille Le Calonnec dans sa version finale.

Membres du jury

Professeur Alexandre Blais
Directeur de recherche
Département de physique
Université de Sherbrooke

Professeur Guido Pupillo
Co-directeur de recherche
Laboratory of Quantum Physics
ISIS - CESQ

Professeur David Sénéchal
Président rapporteur
Département de physique
Université de Sherbrooke

Professeure Éva Dupont-Ferrier
Membre interne
Département de physique
Université de Sherbrooke

Professeure Sophia Economou
Membre externe
Department of Physics
Virginia Tech

Professeur Shannon Whitlock
Membre externe
Laboratory of Exotic Quantum Matter
ISIS - CESQ

To: My Family ♡

Remerciements

En tout premier je voudrais remercier Alexandre de m'avoir donné la chance de pouvoir vivre cette expérience de doctorat dans son groupe. Merci de m'avoir guidé tout au long, de m'avoir encouragé et de m'avoir donné toutes les opportunités possibles qu'un doctorant pouvait espérer avoir. Bien qu'Alexandre soit une personnalité dans le domaine des circuits supraconducteurs et qu'il soit très occupé, il est toujours présent pour ses étudiants, que ce soit pour discuter science ou autre. Je n'aurais pu imaginer une meilleure ambiance dans un groupe de recherche.

Je voudrais aussi remercier Guido de m'avoir donné la possibilité de faire mon doctorat en cotutelle avec la France. Cette cotutelle m'a permis de rencontrer des personnes incroyables et de participer à des évènements qui ont eu une influence importante dans ma vie. Je voudrais remercier les personnes du CESQ avec qui j'ai passé de bons moments durant mon séjour en France: Sven, Vineesha, Johannes, Guido, Guillermo, Tanul, Manuel, Tatjana, Francesco, David, Anna, Tom, Shannon. Special thank you to Sven, which has been a really fun friend!!

I would like to thank the people I've been working with during my thesis, namely:

- Alexandru for supervising me in the first years of my PhD, for sharing his expertise with me and for his life advices and support through the years. I also had the chance to work with Cat and Agustin, two scientifically impressive and inspiring characters.

- the people involved in the two-qubit gate project at Princeton: Pranav, Andrei, Alex, Sara and Charlie. A special thanks to Sara and Charlie, my main contacts in the last years, for teaching me a lot of things about the experiments and for their kindness and generosity.

- Maxime et Alexandre Choquette pour les précieuses discussions que nous avons eues au cours des années qui finalement ont abouti à adapt-QOCA. Merci Maxime d'avoir partagé des morceaux de codes avec moi, de m'avoir aidé à interpréter mes résultats et par-dessus tout, merci de m'avoir aidé avec Qiskit! Merci aussi à Alexandre Foley de l'AlgoLab pour notre courte (mais agréable) collaboration sur adapt-QOCA appliqué au modèle de Fermi-Hubbard.

Je voudrais remercier toutes les personnes qui sont passées par le groupe d'Alexandre pendant mon séjour à Sherbrooke car vous avez rendu la vie à Sherbrooke plus agréable.

Que ce soit MF avec sa bonne humeur, Alexander avec ses blagues toujours bien placées, Élie juste en étant Élie, Manuel for our Algorithm discussions, (big) Ben for being a book of knowledge and for being a great pool and bulb drawer, (small) Ben for always having my back at volleyball (cause everybody knows how bad I am), Cristobal(+Tania+Baby cause they're almost part of the group) for your infinite kindness, Ross for all your funny Gifs, Alex(+Sam) for being a great gossip listener, Johnathan for being the chillest and easy to leave roommate ever during covid, Choquette for our discussions on QOCA, Ronan for given us the feeling he had always been part of the family, Anthony for being such a great intern, Joachim for bringing a French vibe to the group (or maybe I should thank you for that) and Cat for being a badass model. Remerciements très très spéciaux à Othmane et Lautaro (♡) pour être restés avec moi très tard au bureau les soirs et les weekends pendant la rédaction (compliquée) de ma thèse. Merci d'avoir cuisiné pour moi et de m'avoir apporté des bonbons pour me remonter le moral! Merci aussi à notre coordonnatrice Catherine, qui facilite notre vie de tous les jours et qui a toujours eu une oreille attentive.

For me a PhD is much more than doing research, it's also about developing new skills, meeting people and personal growth. Alongside my PhD, I pursued this amazing program called QSciTech which aims at developing our skill in entrepreneurship (in quantum). I would like to thank the people responsible for the program, namely Yves Bérubé-Lauzière, Dominique Parisé and Vanessa Sandoval Romero, and all my teachers. I've learned so much from all the classes and I had a great time. I am very grateful I could be part of it!

I would also like to thank all the people from the student initiatives I've been part of during my PhD, namely all the people from QMAT, StrasQC (Mauricio, Manuel, Denis, Jean-Gab), the Quartier Quantique (Karl, Choquette, Agustin, Élie, Ioanna, Patrick) and DiPhUS.

I'm also sincerely grateful to Shannon Whitlock for giving me the opportunity to be in the organizing committee of the First European Spring School in Quantum Science and Technology co-organized by students of the Université de Strasbourg and the University of Heidelberg. This experience has contributed significantly to my personal growth and I got to know people who have had an important impact on my life. Thanks to all the other organizers: Johannes, Alexandra, Robert, Karen, Jean-Gab, Marie, Pierre. Special thank you to Johannes without whom the experience wouldn't have been the same. This was an incredible experience!

Je voudrais aussi remercier ma meilleure amie Sandra, présente dans les bons comme dans les mauvais moments, qui a été d'un soutien indéfectible durant cette thèse. Elle a toujours eu les mots justes pour me remonter le moral, et depuis toujours, nos discussions ont le pouvoir de me motiver quelque soit le sujet. Pour tout cela je lui suis sincèrement reconnaissante. Je voudrais aussi particulièrement remercier Martin et Baptiste pour leur amitié et leur soutien depuis que je les connais. Merci aussi à Alex, Bavneet, Ankita, Gurleen, Thomas and co.

Finalement, je voudrais remercier ma famille pour leur soutien et leur amour inconditionnel! Sans elle, j'en suis convaincue, je n'aurais pas pu terminer ma thèse! Merci à ma mère et à ma soeur d'avoir été présente lorsque j'avais besoin, et mon père pour m'avoir

porté jusque là... ♡ . And thank you Mark for your support and for believing in me. If you hadn't pushed me onto that plane, I most likely wouldn't have made it to the end!

Résumé

La recherche en informatique quantique a pris de l'ampleur dans les deux dernières décennies, notamment grâce aux financements massifs des gouvernements et de l'industrie dans les technologies quantiques. L'ordinateur quantique promet de résoudre des problèmes qui jusqu'à présent restaient insurmontables à cause du temps déraisonnable qu'il faudrait pour les résoudre sur les ordinateurs classiques les plus puissants. L'idée de l'ordinateur quantique consiste à utiliser des propriétés quantiques de la matière, telles que l'intrication et la superposition quantique des états, pour effectuer des calculs d'une nouvelle manière. L'informatique quantique requiert un nouveau type d'ordinateur où les bits d'information classique, généralement représentés par la présence ou non de courant à l'entrée d'un transistor, sont remplacés par des bits quantiques aussi appelés qubits, et qui sont représentés par l'état d'un système quantique. De plus, il est nécessaire de développer de nouveaux algorithmes capables d'être implémentés sur un ordinateur quantique et capables de tirer parti de l'avantage quantique. Cette thèse porte sur ces deux composantes complémentaires : la construction de portes logiques, un bloc fondamental de tout ordinateur, et la création de meilleurs algorithmes pour résoudre des problèmes pertinents sur ces machines.

Parmi les plateformes les plus avancées pour l'ordinateur quantique, on trouve les circuits supraconducteurs, qui sont la plateforme sur laquelle ont misé Google, IBM, Alice&Bob, Nord Quantique et bien d'autres compagnies. Malgré les avancées des dernières décennies, de nombreux défis restent à être relevés pour que les machines existantes soient en mesure de résoudre des problèmes pertinents et surpassent les ordinateurs classiques. L'un des défis auquel cette thèse s'attaque est la réalisation de portes logiques à deux qubits rapides et à haute fidélité. La rapidité est cruciale pour pouvoir réaliser le plus d'opérations possible avant que les qubits ne perdent leur cohérence quantique, et une haute fidélité est requise pour minimiser les erreurs de calcul et permettre d'implémenter de la correction d'erreurs quantiques. De plus, certains qubits supraconducteurs sont connus pour être affectés par une interaction indésirable, l'interaction ZZ qui, même si relativement faible, affecte la fidélité des portes logiques quantiques, empêchant d'atteindre des fidélités supérieures à 99.9%

nécessaires à la correction d'erreurs.

En parallèle avec l'amélioration des composantes de l'ordinateur quantique, il est important de développer des applications pour ce nouveau type de machine. Parmi les approches étudiées, les algorithmes quantiques variationnels (VQA) sont ceux qui sont les plus susceptibles d'être déployés sur les ordinateurs quantiques de petites tailles existants. Ce sont des algorithmes hybrides qui nécessitent à la fois de la puissance de calcul classique et quantique. Ces algorithmes semblent prometteurs pour résoudre de nombreux problèmes concrets d'intérêt industriel, comme obtenir les propriétés de molécules ou trouver les solutions de problèmes d'optimisation combinatoire. En effet, ces problèmes peuvent être formulés sous forme d'un Hamiltonien et résolus en déterminant son état fondamental. Dans ce but, les VQAs utilisent un circuit quantique paramétré appelé *ansatz*, dont les paramètres sont optimisés pour s'approcher de l'état fondamental. Un premier défi consiste à trouver un ansatz qui permette, pour un problème donné, de converger le plus efficacement possible vers la solution. Un deuxième défi est de trouver un circuit fait de portes logiques facilement implémentables sur les plateformes physiques actuelles.

Les travaux effectués dans cette thèse apportent des contributions à ces deux défis. Dans la première partie de la thèse, nous introduisons le dispositif que nous avons conçu et qui permet de réaliser des portes d'intrication paramétrique entre deux qubits supraconducteurs. Cette approche a permis à nos collaborateurs du groupe d'Andrew Houck de l'Université de Princeton de réaliser expérimentalement une opération de type $\sqrt{i\mathrm{SWAP}}$ en 15 ns avec une fidélité de 98.8%, et ce dans un dispositif où l'interaction ZZ est supprimée. Dans la seconde partie de la thèse, nous avons développé un ansatz pour les VQAs permettant de résoudre le modèle de Fermi Hubbard. Ce nouvel ansatz utilise les avantages de deux ansätze préexistants, et permet d'obtenir l'énergie et l'état fondamental du FHM avec une précision de plusieurs ordres de grandeur supérieure aux ansätze standards, et ce en réduisant significativement le nombre de portes d'intrication CNOT qui sont généralement imparfaites et longues à implémenter.

Summary

Quantum computing has gained momentum over the last two decades, thanks in particular to massive government and industry funding for quantum technologies. The quantum computer promises to solve problems that are currently untractable because of the unreasonable time it would take to solve them on the most powerful classical computers. The idea of the quantum computer is to use quantum properties of matter, such as entanglement and quantum superposition of states, to perform calculations in a new way. Quantum computing requires a new type of computer in which classical information bits, generally represented by the presence or absence of current at the input of a transistor, are replaced by quantum bits, also known as qubits, which are represented by the state of a quantum system. In addition, it is necessary to develop new algorithms that can be implemented on a quantum computer and that are able to benefit of the quantum advantage. This thesis focuses on these two complementary components: the construction of logic gates, a fundamental building block of any computer, and the design of algorithms for solving relevant problems on these machines.

Among the most advanced platforms for quantum computing are superconducting circuits, which are the platform on which Google, IBM, Alice&Bob, Nord Quantique and many other companies have decided to bet on. Despite the advances of recent decades, many challenges remain to be overcome before existing machines are able to solve relevant problems and outperform classical computers. One of the challenges this thesis tackles is the realization of fast and high-fidelity two-qubit gates. Speed is crucial to be able to perform as many operations as possible before the qubits lose their quantum coherence, and high fidelity is required to minimize computational errors and enable quantum error correction to be implemented. In addition, some superconducting qubits are known to be affected by an undesirable interaction, the ZZ interaction, which, although relatively weak, affects the fidelity of quantum logic gates, making it challenging to achieve fidelities above 99.9% required for quantum error correction.

Alongside the improvement of quantum computer components, it is important to develop

applications for this new type of machine. Among the approaches explored, variational quantum algorithms (VQA) are the family of algorithms most likely to be applied on existing small-scale quantum computers. These are hybrid algorithms that require both classical and quantum computing resources. VQAs seem promising for solving many concrete problems of industrial interest, such as obtaining the properties of molecules or finding solutions to combinatorial optimization problems. Indeed, these problems can be formulated in terms of a Hamiltonian and solved by determining its ground state. To this end, VQAs use a parameterized quantum circuit called *ansatz*, whose parameters are optimized to approximate the ground state. A first challenge is to find an ansatz which, for a given problem, converges as efficiently as possible to the solution. A second challenge is to find a circuit made of logic gates that can be easily implemented on current physical platforms.

The work carried out in this thesis tackles both of these challenges. In the first part of the thesis, we introduce a device we have designed that enables the implementation of parametric entanglement gates between two superconducting qubits. This approach has enabled our collaborators in Andrew Houck's group at Princeton University to experimentally realize a $\sqrt{i\text{SWAP}}$ operation in 15 ns with a fidelity of 98.8%, and this in a device where the ZZ interaction is suppressed. In the second part of the thesis, we developed an ansatz for VQAs to solve the Fermi Hubbard Model. This new ansatz uses the advantages of two pre-existing ansätze, and makes it possible to obtain the energy and ground state of the FHM with an accuracy several orders of magnitude better than standard ansätze, while significantly reducing the number of CNOT gates, which are generally imperfect and time-consuming to implement.

Abbreviations

Adapt-VQE Adaptive Derivative-Assembled Problem-Tailored ansatz Variational Quantum Eigensolver

BCH Baker-Campbell-Hausdor expansion

COBYLA Constrained Optimization BY Linear Approximations

C-Phase Controlled-Phase gate

CPTP Complete Positive Trace Preserving

CR Cross-Resonance gate

FHM Fermi-Hubbard Model

FQA Feedback-based Quantum Algorithm

GFQ Generalized Flux Qubit

HEA Hardware Efficient Ansatz

JJ Josephson junction

JW Jordan-Wigner

KNO Kerr Non-linear Oscillator

NISQ Parametrized Quantum Circuit

PQC Quantum Optimal Control inspired Ansatz

QAOA Quantum Approximate Optimization Algorithm

QOCA Quantum Optimal Control inspired Ansatz

QUBO Quadratic Unconstrained Binary Optimization

RWA Rotating-Wave Approximation

SPSA Simultaneous Perturbation Stochastic Approximation

SW Schrieffer-Wolff

SWPT Schrieffer-Wolff Perturbation Theory

SNAIL Superconducting Nonlinear Asymmetric Inductive eLement

TLS Two-Level System

UCC Unitary Coupled Cluster

VA Variational Algorithm

VHA Variational Hamiltonian Ansatz
VQA Variational Quantum Algorithm

Contents

List of Tables

List of Figures

Part I

Parametric two-qubit gate

Chapter 1

Quantum computing with superconducting circuits

In this first chapter, we introduce the basic concepts of superconducting circuits that will be useful for the following chapters. In Sect. 1.1, we explain how an electrical circuit can be formulated in the form of a Lagrangian and subsequently expressed as a Hamiltonian. The procedure that takes us from a circuit to a quantized Hamiltonian is known as circuit quantization. We provide an overview of the basic concept using the approach of Devoret [1, 2]. In Sect. 1.2, we describe three commonly used superconducting qubits and which will be useful in the next chapters: the fixed-frequency transmon, the flux-tunable transmon and the generalized flux qubit. Finally, in Sect. 1.3 we provide an overview of the different strategies that can be used to implement two-qubit gates between superconducting qubits.

I hope that this chapter will provide valuable context on how our work relates to the current state-of-the-art research.

1.1 Circuit quantization

In this section, we present the approach of Devoret to quantize an electrical circuit [1, 2]. The electrical circuits that we consider in the following are two-pole elements connected at nodes in simple networks. The most common elements are capacitances, inductances and Josephson junctions. The Josephson junction (JJ) is ubiquitous in quantum computing

with superconducting circuit because it is the best nonlinear element that is also non-dissipative at sub-Kelvin temperatures. This nonlinearity plays an important role for designing superconducting qubits.

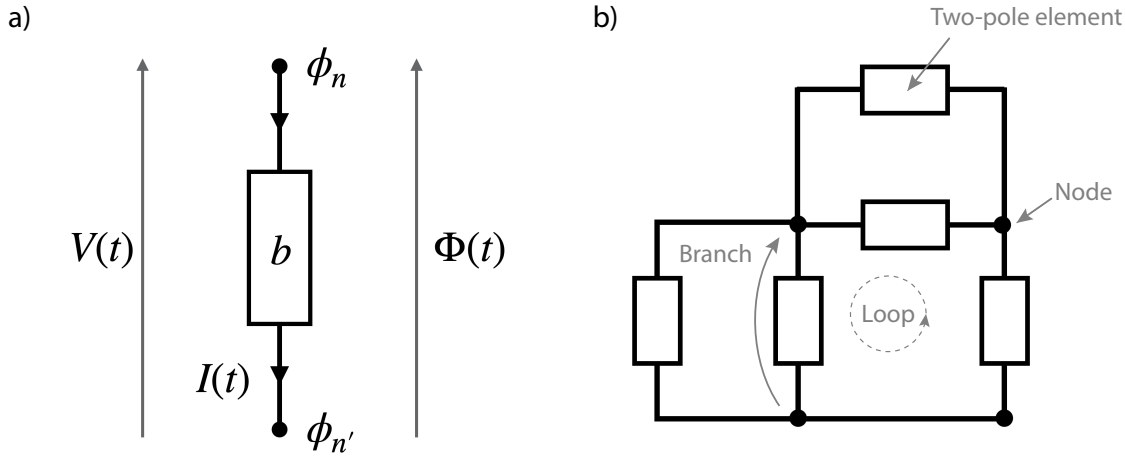


Figure 1.1 a) Diagram of a two-pole element on branch b, with a current I flowing through it and a voltage V across the branch. The branch flux Φ is defined between the flux nodes ϕ_n and $\phi_{n'}$. b) Diagram of an electrical circuit made of two-pole elements. The nodes are indicated by dots and the branches are the path between two consecutive nodes.

The branch charge and branch flux of any two-pole element are described by the voltage $V(t)$ and the current $I(t)$ that flows through it:

$$\Phi(t) = \int_{-\infty}^{t} dt' V(t') \tag{1.1}$$

$$Q(t) = \int_{-\infty}^{t} dt' I(t'), \tag{1.2}$$

where the branch flux can also be written as $\Phi = \phi_n - \phi_{n'}$, with ϕ_i the flux node at the nodes $i \in \{n, n'\}$, see Fig. 1.1. The energy absorbed by the element is given by:

$$E(t) = \int_{-\infty}^{t} dt' I(t') V(t'). \tag{1.3}$$

Using basic equations relating the voltage and the current, such as $V = L\frac{dI}{dt}$ and $I = \frac{dQ}{dt} = C\frac{dV}{dt}$ for a linear inductor (resp. capacitor) with inductance L (resp. capacitance C), one can determine the voltage and current flowing through a circuit element, and the energy stored in it, as a function of the branch flux and branch charge. In the following, we express those quantities in terms of the branch flux (see Table 1.1), note, however, that they could

be expressed in terms of the branch charge instead without loss of generality. Using this convention, one can identify the capacitive energies as the kinetic energies and the inductive energies as the potential energies.

Two-pole element	Current	Voltage	Energy	Symbol
Linear capacitor	$I_C = C\ddot{\Phi}$	$V_C = \dot{\Phi}$	$\frac{C}{2}\dot{\Phi}^2$	
Linear inductor	$I_L = \frac{1}{L}\Phi$	$V_L = \dot{\Phi}$	$\frac{1}{2L}\Phi^2$	
Josephson Junction	$I_J = \frac{2e}{\hbar}E_J \sin(\frac{2e}{\hbar}\Phi)$	$V_J = \dot{\Phi}$	$-E_J \cos\left(\frac{2e}{\hbar}\Phi\right)$	

Table 1.1 Common superconducting elements. For each circuit element, we define the current and the voltage flowing through the dipole, the energy stored in it and its symbol. For the linear capacitor, C is the capacitance, for the linear inductor, L is the inductance, and for the Josephson junction E_J is the Josephson energy.

Due to the constrains imposed by Kirchhoff's laws, the number of degrees of freedom of the circuit is smaller than the number of branches in the circuit. To get rid of the redundancy of variables when deriving the Hamiltonian of a superconducting circuit, we use the approach of Vool and Devoret [2], known as the *method of nodes*. Note that other methods exist to derive the Hamiltonian for a circuit which can better describe certain device circuits [3, 4].

The first step of the method of nodes is to define a *spanning tree* $\mathcal{T}$ for the circuit. This step consists in defining a reference node which is set to the ground, and identifying a set of branches that connect every node of the circuit to the ground through branches with a capacitance. The tree $\mathcal{T}$ should not form any loop. Also, note that the choice of spanning tree is not unique. To satisfy Kirchhoff's laws, the branch fluxes are defined as:

$$\begin{aligned} \Phi_{b\in\mathcal{T}} &= \phi_n - \phi_{n'} \\ \Phi_{b\in\overline{\mathcal{T}}} &= \phi_n - \phi_{n'} + \Phi^b_{\text{ext}} \end{aligned} \tag{1.4}$$

where $\overline{\mathcal{T}}$ is the complement of $\mathcal{T}$ and, ϕ_n and $\phi_{n'}$ are the flux nodes of the branch b. If the branch b is not in the spanning tree, but belongs to a loop, then the external static flux Φ^b_{ext} enclosed by the loop, is added to the flux branch.

The Lagrangian is obtained by subtracting the potential energy from the kinetic energy:

$$\mathcal{L} = E_{kin} - E_{pot} = \sum_{b\in B_C} \frac{C_b}{2}\dot{\Phi}_b^2 - \sum_{b'\in B_L} \frac{\Phi_{b'}^2}{2L_{b'}} + \sum_{b''\in B_J} E_{Jb''} \cos\left(2\pi\frac{\Phi_{b''}}{\Phi_0}\right), \tag{1.5}$$

where B_c, B_L and B_J are the branches that have respectively a capacitance, an inductor or a Josephson junction. We introduced the quantum flux $\Phi_0 = \frac{h}{2e}$. The flux branches Φ_b, $\Phi_{b'}$ and $\Phi_{b''}$ in Eq. (1.5) are written as Eq. (1.4) depending on whether the branches are in $\mathcal{T}$ or $\overline{\mathcal{T}}$. The Lagrangian is thus written in terms of the generalized coordinates $\{\boldsymbol{\phi}, \boldsymbol{\phi}_{\text{ext}}, \dot{\boldsymbol{\phi}}, \dot{\boldsymbol{\phi}}_{\text{ext}}\}$, where $\boldsymbol{\phi}$ and $\boldsymbol{\phi}_{\text{ext}}$ are vector notations of the flux nodes and of the external flux. If one does the exercise and applies Euler-Lagrange equations to the Lagrangian, one will find the equations of motion of the circuit, which are equivalent to Kirchhoff's laws.

To derive the circuit Hamiltonian from the Lagrangian in Eq. (1.5), we first define the conjugate variable of the flux nodes which here corresponds to the charge nodes:

$$q_n = \frac{\partial \mathcal{L}}{\partial \dot{\phi}_n}. \tag{1.6}$$

The Hamiltonian finally reads:

$$H = \sum_n q_n \dot{\phi}_n - \mathcal{L}. \tag{1.7}$$

From the Hamiltonian description of the circuit, it is then straightforward to quantize the Hamiltonian. One simply needs to replace ϕ_n, q_n and H by the operators $\hat{\phi}_n$, $\hat{q}_n$, $\hat{H}$ respectively, and replace the Poisson brackets by the commutator relation $[\hat{\phi}_n, \hat{q}_m] = i\hbar\delta_{nm}$,

In the laboratory, to observe quantum behavior in those systems, dissipation has to be much smaller than the separation between two levels of the Hamiltonian, which is not the case for dissipative circuits operated at room temperature. The circuits that we are considering are made of superconductors operated at milliKelvin temperatures. These superconducting circuits are non-dissipative, making it possible to measure quantum properties.

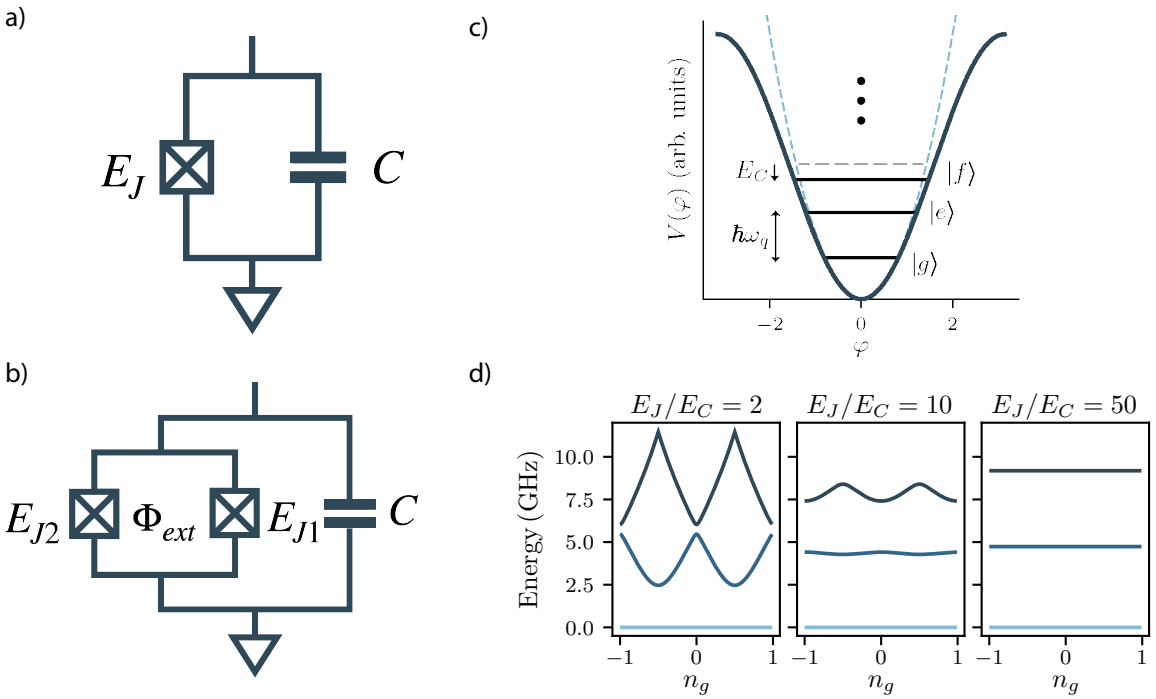


Figure 1.2 a) Diagram of the lumped element circuit of a fixed-frequency transmon, b) Diagram of the lumped element circuit of a flux-tunable transmon (bottom) threaded by a magnetic flux Φ_{ext}. c) One well of the cosine potential of the transmon qubit (full line), $V(\varphi) = -E_J \cos(\varphi)$, compared to the quadratic potential of the LC oscillator (dashed lines), $V(\varphi) = -\varphi^2$. The eigenstates of the transmon are denoted $\{|g\rangle, |e\rangle, |f\rangle, \cdots\}$, and ω_q is the qubit frequency. d) Illustrates the charge dispersion in the first energy levels of the transmon Hamiltonian Eq. (1.8) versus the charge bias n_g. The three plots are for different $E_\mathrm{J}/E_\mathrm{C}$ ratios, for a fixed plasma frequency $\omega_p/2\pi = 5\,\mathrm{GHz}$. c) and d) are adapted from Ref. [5].

1.2 Existing Superconducting qubits and couplers

1.2.1 The transmon qubit

One of the most commonly used superconducting qubits is the transmon qubit [5, 6]. This qubit is made of a capacitance in parallel with a Josephson junction, see Fig. 1.2 a). Using the procedure mentioned in the previous section one can find the transmon Hamiltonian:

$$\begin{aligned} \hat{H} &= \frac{(\hat{q} - q_g)^2}{2C} - E_\mathrm{J} \cos\left(2\pi \frac{\hat{\Phi}}{\Phi_0}\right) \\ &= 4E_\mathrm{C}(\hat{n} - n_g)^2 - E_\mathrm{J} \cos\left(\hat{\varphi}\right) \end{aligned} \tag{1.8}$$

where $\hat{q} = 2e\hat{n}$ is the charge operator, which is related to the operator representing the number of Cooper pairs on the superconducting island, $\hat{n}$. The charging energy is defined as $E_{\mathrm{C}} = \frac{e^2}{2C}$, with $C = C_J + C_S$ the sum of the junction's capacitance (C_J) and of the shunt capacitance (C_S). The parameter n_g represents a classical charge bias on the superconducting island of the transmon that is induced by its capacitive coupling to the environment or to a voltage source. Since $\hat{n}$ is discrete, the phase operator $\hat{\varphi} = 2\pi\hat{\Phi}/\Phi_0$ is compact and its eigenvalues are defined on $[0, 2\pi[$. The commutation relation is $[\hat{\varphi}, \hat{n}] = i$.

The cosine potential of a transmon qubit of frequency ω_q is plotted in Fig. 1.2 c) (solid line) along with the quadratic potential of an LC oscillator of same frequency (dashed line). While the quantized energy levels of an LC oscillator are equally spaced, the transmon levels are anharmonically separated. The *anharmonicity* is defined as the amount by which the second transmon energy level deviates from the one of a harmonic oscillator. We will show in a moment that the anharmonicity of the transmon is, to first approximation, equal to $-E_{\mathrm{C}}$. The number of levels confined in the transmon potential depends on the height of the well ($2E_{\mathrm{J}}$), and can be estimated by counting the number of harmonic levels that could fit in the well, *i.e.* $\sim 2E_{\mathrm{J}}/\sqrt{8E_{\mathrm{C}}E_{\mathrm{J}}}$. The transmon potential generally contains about 8 energy levels. Compared to the harmonic oscillator, the ground state and the first excited state of the transmon can be isolated and used as a qubit. In the absence of anharmonicity or in the case of small anharmonicity, the control of the qubit would cause unwanted transitions to higher energy levels.

Looking at the spectrum of the electrical circuit, in the left plot of Fig. 1.2 d) one can see that its energy levels depend on the offset charge n_g. Because of its coupling to the environment, the device of Fig. 1.2 a) can thus be sensitive to charge noise. We refer to the variations of the energy levels with n_g, $\max[E(n_g)] - \min[E(n_g)]$, as the *charge dispersion*. The choice of E_{J} and E_{C} is of critical importance to control the transmon's sensitivity to charge noise, as the charge dispersion decreases exponentially with the square root of $E_{\mathrm{J}}/E_{\mathrm{C}}$ [6]. Fig. 1.2 d) shows the transmon spectrum as a function of the offset charge for different ratios $E_{\mathrm{J}}/E_{\mathrm{C}}$. When the charging energy is comparable to the Josephson energy, *i.e.* $E_{\mathrm{J}} \sim E_{\mathrm{C}}$, variations in n_g significantly change the spectrum of the device. Such variations in the qubit frequency are to be avoided because they lead to dephasing. To mitigate this issue, transmons are designed in the regime where $E_{\mathrm{J}}/E_{\mathrm{C}} \gg 1$, commonly known as the *transmon regime*. In this regime, see right panel in Fig. 1.2 d), the first transmon levels are to a very good approximation insensitive to charge fluctuations with a charge dispersion almost equal to zero. Note, however, that having a large $E_{\mathrm{J}}/E_{\mathrm{C}}$ ratio comes at the cost of reducing the transmon anharmonicity. The ratio $E_{\mathrm{J}}/E_{\mathrm{C}}$ should therefore remain below 100 to keep some anharmonicity in the spectrum level. A small anharmonicity makes it

complicated to isolate the two lowest energy levels used as a qubit, since driving the qubit leads to unwanted transitions to higher energy levels.

We now write the charge number and phase operators in terms of the creation and annihilation operators:

$$\begin{aligned} \hat{\varphi} &= \varphi_{\mathrm{ZPF}}(\hat{b} + \hat{b}^\dagger), \\ \hat{n} &= n_{\mathrm{ZPF}}(\hat{b} - \hat{b}^\dagger), \end{aligned} \tag{1.9}$$

where we set the characteristic magnitude of the zero-point fluctuations of the charge number and the phase to:

$$\begin{aligned} \varphi_{\mathrm{ZPF}} &= \left(\frac{2E_{\mathrm{C}}}{E_{\mathrm{J}}}\right)^{\frac{1}{4}}, \\ n_{\mathrm{ZPF}} &= -\frac{i}{2}\left(\frac{E_{\mathrm{J}}}{2E_{\mathrm{C}}}\right)^{\frac{1}{4}}. \end{aligned} \tag{1.10}$$

In the transmon regime ($E_{\mathrm{J}}/E_{\mathrm{C}} \gg 1$), φ_{ZPF} is small and one can thus expand the cosine and approximate the Hamiltonian to fourth order in $\hat{\varphi}$:

$$\hat{H} \approx 4E_{\mathrm{C}}\hat{n}^2 + E_{\mathrm{J}}\frac{\hat{\varphi}^2}{2} - E_{\mathrm{J}}\frac{\hat{\varphi}^4}{4!}. \tag{1.11}$$

Notice that we dropped the offset charge term in Eq. (1.11) since the transmon is charge insensitive in this regime. The two first terms in Eq. (1.11) represent a harmonic oscillator and the last term appears as a non-linearity. Replacing $\hat{\varphi}$ and $\hat{n}$ by the creation and annihilation operators, the Hamiltonian can be rewritten as follows:

$$\hat{H} = \sqrt{8E_{\mathrm{J}}E_{\mathrm{C}}}\hat{b}^\dagger\hat{b} - \frac{E_{\mathrm{C}}}{12}\left(\hat{b}^\dagger + \hat{b}\right)^4. \tag{1.12}$$

It is reasonable to apply the rotating-wave approximation (RWA) and remove the terms with unequal number of $\hat{b}^\dagger$ and $\hat{b}$, since those terms will be oscillating fast in the frame rotating at the plasma frequency $\omega_p = \sqrt{8E_{\mathrm{J}}E_{\mathrm{C}}}$ compared to the prefactor E_{C}. The transmon Hamiltonian can thus be approximated by a Kerr non-linear oscillator (KNO):

$$\hat{H} = \hbar\omega_q\hat{b}^\dagger\hat{b} + \frac{\alpha}{2}\hat{b}^{\dagger 2}\hat{b}^2, \tag{1.13}$$

with a qubit frequency $\omega_q = \sqrt{8E_{\mathrm{J}}E_{\mathrm{C}}}/\hbar - E_{\mathrm{C}}/\hbar$ and an anharmonicity $\alpha = -E_{\mathrm{C}}$. The ground-to-first excited state transition has a frequency of ω_q and the first-to-second excited state transition is detuned from ω_q by the anharmonicity $-E_{\mathrm{C}}$, which is negative for a transmon. This small nonlinearity in the spectrum makes it possible to control the ground

and first excited states without populating the higher excited states. Therefore the two first levels of the transmon can be used as a qubit. The transmon is a multi-level system with a spectrum that is anharmonically spaced, it is for this reason often said to behave like an artificial atom. Notice that the higher transmon levels could be used to form a *qudit* instead of a qubit [7] [8] [9].

Transmon qubits have relaxation times T_1 of the order of tens to hundred of microseconds, with state-of-the-art $T_1 s$ around 300 to 500 μs [10–12]. The dephasing time T_2 is typically of the same order or magnitude than the T_1.

1.2.2 The flux-tunable transmon

The transmon can be rendered flux-tunable by integrating two JJs shunted in parallel instead of a single JJ, see Fig. 1.2 b). The two JJs form a SQUID loop. Following the approach introduced in Sect. 1.1, one finds the following Hamiltonian:

$$\hat{H} = 4E_{\mathrm{C}}\hat{n}^2 - E_{\mathrm{J1}}\cos\left(\hat{\varphi}\right) - E_{\mathrm{J2}}\cos\left(\hat{\varphi} - \varphi_{\mathrm{ext}}\right), \tag{1.14}$$

where E_{Ji} is the Josephson energy of junctions i in the SQUID loop and E_{C} represents the charging energy, encompassing the capacitances of the junctions as well as the shunt capacitance. The operators $\hat{n}$ and $\hat{\varphi}$ represent the charge number and phase across the circuit branch. The parameter Φ_{ext} is the external flux threaded through the SQUID loop, and $\varphi_{\mathrm{ext}} = 2\pi\Phi_{\mathrm{ext}}/\Phi_0$. By the means of trigonometric manipulation, the Hamiltonian can be rewritten with a single cosine potential as [6]:

$$\hat{H} = 4E_{\mathrm{C}}\hat{n}^2 - E_{\mathrm{J}}(\varphi_{\mathrm{ext}})\cos\left(\hat{\varphi} - \varphi_0\right) \tag{1.15}$$

with

$$E_{\mathrm{J}}(\varphi_{\mathrm{ext}}) = (E_{\mathrm{J1}} + E_{\mathrm{J2}})\cos\left(\frac{\varphi_{\mathrm{ext}}}{2}\right)\sqrt{1 + d^2\tan^2\left(\frac{\varphi_{\mathrm{ext}}}{2}\right)}, \tag{1.16}$$

where $d = (E_{\mathrm{J2}} - E_{\mathrm{J1}})/(E_{\mathrm{J1}} + E_{\mathrm{J2}})$ is the *junction asymmetry* and

$$\tan(\varphi_0) = \frac{E_{\mathrm{J2}}\sin(\varphi_{\mathrm{ext}})}{E_{\mathrm{J1}} + E_{\mathrm{J2}}\cos(\varphi_{\mathrm{ext}})}\tan\left(\frac{\varphi_{\mathrm{ext}}}{2}\right). \tag{1.17}$$

Note that φ_0 differs from the definition in Ref. [6], where $\tan(\varphi_0) = d\tan\left(\frac{\varphi_{\mathrm{ext}}}{2}\right)$. This difference arises from the fact the method of nodes is used here to derive the Hamiltonian while Ref. [6] uses the method of branches. Nevertheless, the Hamiltonians are equivalent up to a gauge transformation. The phase φ_0 can anyway be eliminated by a shift of variables

in the case where the external flux Φ_{ext} is constant.

Because of the external flux the transmon now has an effective Josephson energy which depends on φ_{ext}. This renders the transmon frequency tunable with,

$$\hbar\omega(\varphi_{\text{ext}}) = \sqrt{8E_{\text{C}}|E_{\text{J}}(\varphi_{\text{ext}})|} - E_{\text{C}} \tag{1.18}$$

and the anharmonicity remains fixed and negative as for the fixed-frequency transmon. The flux-tunable transmon is widely used as a qubit or a coupler [13–20]. One advantage of using the flux-tunable transmon as a qubit is that it allows for more flexibility in the device parameters, as current nanofabrication errors can typically shift the capacitances by $1-5\%$ and the Josephson energies by $\sim 5-20\%$ [21–23]. Flux-tunable transmons are also frequently used as couplers, where their frequency is tuned such as to turn on or off interactions between the qubits [14, 24]. The tunability of the qubits or the coupler is also used to implement gates, more details on how two-qubit gates are performed will be given in Sect. 1.3. It is important to note, however, that this flexibility in the flux-tunable transmon comes at the expense of increased sensitivity to flux noise. This flux noise ultimately leads to a decrease in the qubit phase coherence compared to fixed-frequency transmons. An additional drawback of flux-tunable qubits is the need for calibrating the device parameters, which can be a challenging task, especially for chips with a large number of qubits [25, 26].

1.2.3 The Generalized Flux Qubit

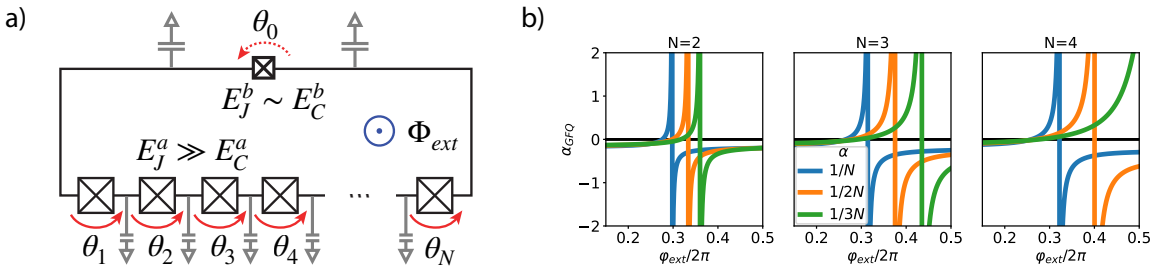


Figure 1.3 a) Diagram of the lumped element circuit of a generalized flux qubit, adapted from Ref. [27]. b) Anharmonicity of the GFQ α_{GFQ} in the KNO approximation versus flux. Each plot corresponds to a different N and each line to a different anisotropy ratio $\alpha = 1/N$, $1/2N$ and $1/3N$.

The generalized flux qubit (GFQ) is a variation of the flux qubit, where instead of a simple SQUID, one of the branches of the loop has an array of N Josephson junctions, as shown in Fig. 1.3 [28–33]. This circuit is also known as a Superconducting Nonlinear Asymmetric

Inductive eLement (SNAIL) developed by Frattini *et al.* [29]. The circuit branch with a single junction is called the *black-sheep junction*, and we will refer to the branch with multiple junctions as the junction array.

In principle, this circuit contains $N+1$ modes, defined by the flux branches θ_i. Under some assumptions it is possible to approximate the array of junction by a single mode, $\phi = \sum_{m=1}^{N} \theta_m$, usually called the *superinductance mode* [27]. The Hamiltonian of the GFQ derived from circuit quantization can then be expressed as follows:

$$\hat{H}(t) = 4E_{\mathrm{C}}\hat{n}^2 - \alpha E_{\mathrm{J}} \cos\left[\hat{\varphi} + \varphi_{\mathrm{ext}}\right] - \beta N E_{\mathrm{J}} \cos\left[\frac{\hat{\varphi}}{N}\right], \tag{1.19}$$

with $\hat{\varphi} = \phi/\phi_0$ where $\phi_0 = \frac{\hbar}{2e}$, and $\hat{n}$ respectively the superconducting phase difference and the Cooper pair number operators. The parameter E_{C} is the charging energy of the superinductance mode, which can be written in terms of the charging energies of the black sheep junction E_{C}^b, and of the array E_{C}^a as $E_{\mathrm{C}} = \frac{NE_{\mathrm{C}}^a E_{\mathrm{C}}^b}{NE_{\mathrm{C}}^a + E_{\mathrm{C}}^b}$. We assume that all the Josephson junctions are identical in the array and that each individual Josephson energy is E_{J}. The parameter β corresponds to a renormalization of the superinductance due to disorder in the junction array and to finite zero-point fluctuations. This quantity can be approximated as $\beta \sim \exp\left(-\frac{N-1}{4N}\sqrt{8E_{\mathrm{C}}^a/E_{\mathrm{J}}^a}\right)$, see Ref. [34]. The parameter α is simply the ratio between the Josephson energy of the black sheep junction and the Josephson energy of the junction array E_{J}. It parametrizes the anisotropy between the two branches of the loop. Finally, the loop can be threaded by an external flux $\varphi_{\mathrm{ext}} = \frac{\phi_{ext}}{\phi_0}$.

The Hamiltonian in Eq. (1.19) is obtained by assuming that [29]:

- All the Josephson junctions are identical in the N-junction array, with a Josephson energy E_J and a capacitance C_J.
- The external flux φ_{ext} is constant.
- The capacitance to the ground of each junction between the islands can be neglected. This is valid if the capacitance to the ground is much smaller than C_J/N^2.
- The single-mode approximation is valid if the plasma frequency of a junction, $\hbar\Omega_p = \sqrt{8E_J E_{C_J}}$, is higher frequency than the dynamics that is being studied. This is usually the case because the devices are designed with $\Omega_p/2\pi \equiv [20-40]$ GHz, which is to be compared with dynamics occurring in the GHz range.
- The phase slip rate in the array, which is proportional to $e^{-\sqrt{E_J/E_{C_J}}}$, is small compared to the inverse of the relevant timescale. This is usually the case because of the shunting

capacitor.

As for the transmon, the cosine potentials in Eq. (1.19) can be Taylor expanded and then expressed in terms of creation and annihilation operators. For $E_{\mathrm{J}}/E_{\mathrm{C}}$ large and $\alpha < 1/N$, the generalized flux qubit can also be expressed as a KNO with frequency and anharmonicity:

$$\omega_{GFQ}(\varphi_{\mathrm{ext}}) = \sqrt{8E_{\mathrm{J}}E_{\mathrm{C}}\left[\alpha\cos\left(\varphi_{\mathrm{ext}}\right)+\frac{\beta}{2N}\cos\left(\frac{\varphi_{\mathrm{ext}}}{N}\right)\right]} \tag{1.20}$$

$$\frac{\alpha_{GFQ}}{2} = -\frac{1}{2}E_{\mathrm{C}}\left[\frac{\alpha\cos\left(\varphi_{\mathrm{ext}}\right)+\frac{\beta}{N^3}\cos\left(\frac{\varphi_{\mathrm{ext}}}{N}\right)}{\alpha\cos\left(\varphi_{\mathrm{ext}}\right)+\frac{\beta}{2N}\cos\left(\frac{\varphi_{\mathrm{ext}}}{N}\right)}\right]. \tag{1.21}$$

In the limit where $\beta \to 0$ one recovers the transmon. For the GFQ, the anharmonicity can be made positive when $\cos(\varphi_{\mathrm{ext}})$ is negative if the number N and the anisotropy ratio α between the Josephson energy of the two branches is appropriately chosen. In Fig. 1.3 b), we show the anharmoticity of the GFQ in the KNO appromixation versus φ_{ext}, for different N and α. We see that it is easier to obtain positive anharmonicities for larger value of N and for $\alpha < 1/N$. Note that for $\alpha < 1/N$ (resp. $\alpha > 1/N$), the potential of the GQF in Eq. (1.19) features a single (multiple) well versus the superconducting phase φ.

It has been shown for qubits with opposite anharmonicities that it is possible to suppress unwanted ZZ interactions between the qubits [35–37]. The ZZ interaction is generated by a term in $\sigma_{Z_1}\sigma_{Z_2}$ in the Hamiltonian which shifts the qubit frequencies when both qubits are in the first excited state. While this ZZ term can be desired to generate CZ gates, it can also be unwanted. Spurious ZZ interaction is one of the main factors limiting current gate fidelities and a great deal of efforts have been made to address this issue [35–39]. In Chapter 3, we show that by using a GFQ as a coupler between transmon qubits, it is possible to cancel this ZZ interaction.

1.3 Two-qubit gates

There exist multiple approaches of generating two-qubit gates for superconducting qubits [5]. In this section, we briefly describe three ways of implementing gates in superconducting qubit architectures, depending on the presence of a coupler and on the tunability of the qubits.

1.3.1 Flux-tuned gates

The general idea of flux-tuned gates is to keep the qubits far detuned from any resonance when the gate is off and to tune specific energy levels into resonance to implement a certain gate. Here we consider one of the simplest architectures to implement a gate, that is, one or two flux-tunable qubits which are directly coupled through a capacitance. This system is described by a beam-splitter Hamiltonian:

$$\begin{aligned}\hat{H} =&\hbar\omega_1\hat{b}_1^\dagger\hat{b}_1 + \frac{\alpha_1}{2}\hat{b}_1^{\dagger 2}\hat{b}_1^2 + \hbar\omega_2\hat{b}_2^\dagger\hat{b}_2 + \frac{\alpha_2}{2}\hat{b}_2^{\dagger 2}\hat{b}_2^2 \\ &+ \hbar J(\hat{b}_1^\dagger\hat{b}_2 + \hat{b}_1\hat{b}_2^\dagger),\end{aligned} \tag{1.22}$$

where J depends on the Josephson and charging energies of the qubits, and on the coupling capacitance. This Hamiltonian describes the coherent exchange of an excitation between the two qubits.

By tuning the qubits at some specific frequencies, different gates can be activated using this Hamiltonian. For instance, by tuning the qubits on resonance, $\omega_1 = \omega_2$, it is possible to generate an iSWAP gate. This becomes evident by treating the qubits as two-level systems and moving to the rotating frame at the qubit frequencies:

$$\hat{H}_{\text{eff}} = \hbar J(\hat{\sigma}_{+1}\hat{\sigma}_{-2} + \hat{\sigma}_{-1}\hat{\sigma}_{+2}). \tag{1.23}$$

An iSWAP gate can thus be implemented with this Hamiltonian by letting the system evolve for a time $t_{i\text{SWAP}} = \pi/(2J)$.

Still using the Hamiltonian in Eq. (1.22), it is also possible to implement a controlled-phase gate (C-Phase). This can be achieved by taking advantage of the higher energy levels of the qubits [5, 40]. In fact, when $J \neq 0$, the energy level $\hbar\omega_{11}$ associates to the state with both qubits in the first excited state is no longer equal to the sum of the qubit energies $\hbar(\omega_{01} + \omega_{10})$. The amount by which the energies differ is denoted $\zeta = \omega_{11} - \omega_{01} - \omega_{10}$ and is generally referred to as the ZZ interaction. In the spectrum, this shift translates in an avoided crossing between the computational level $|1_1 1_2\rangle$ and the second excited state of one of the qubits $|0_1 2_2\rangle$, when tuning ω_1. Tuning the qubits in and out of this avoided crossing leads to the following unitary operation: $diag(1, e^{\phi_{01}}, e^{\phi_{10}}, e^{\phi_{11}})$, where $\phi_{ab} = \int dt\omega_{ab}$ is the dynamical phase accumulated during the flux excursion by the state $|a_1, b_2\rangle$ with $a\,b \in \{0, 1\}$. Up to single qubit phase gates, this operation is equivalent to a C-Phase gate, $\text{CPHASE}(\phi) = diag(1, 1, 1, e^{i\phi})$, where $\phi = \phi_{11} - \phi_{01} - \phi_{10} = \int dt\zeta(t)$. For $\phi = \pi$, this operation leads to a controlled-Z gate (CZ).

One of the main issues of this type of gate is that activating a gate can lead to frequency collisions when tuning the qubit frequency. These frequency collisions can be with other qubits, resonators, and couplers, or with to two-level defects that are almost ubiquitous in superconducting chips [41].

Above we showed how two-qubit gates can be performed by tuning the qubit frequencies at some resonance frequencies. The same principle can be applied using a resonator bus or a flux-tunable coupler instead of tunable qubits to turn on and off interactions [14, 24, 42, 43]. Using this approach, some of the best iSWAP-like gates have been implemented in about 12 ns with $\sim 99.5\%$ fidelity [24] and CZ gate in 30 ns with 99.07% fidelity [42]. Typical iSWAP times are around a few to tens of nanoseconds and CZ gate are generally in the order of tens of nanoseconds.

1.3.2 All-microwave gates

Previously we have shown how gates can be activated by changing a control parameter of the qubits and/or a coupler, like a flux line. Here, the idea is to turn on and off interactions using a microwave drive on the qubits and/or the coupler. These types of gates can be implemented with flux-tunable and fixed-frequency qubits and/or couplers.

As an example, we consider the cross-resonance (CR) gate [44, 45]. The CR gate is implemented between two detuned qubits that are either capacitively coupled or connected via a resonator bus. The main idea is to drive a qubit at the frequency of the other, thereby activating the CR term in $\hat{\sigma}_{\mathrm{z}1}\hat{\sigma}_{\mathrm{x}2}$. This CR term can ultimately be used to realize a CNOT gate.

Considering two transmons approximated as KNOs (see Sect. 1.2), the driven Hamiltonian can be expressed as follows:

$$\begin{aligned}\hat{H} =&\hbar\omega_1\hat{b}_1^\dagger\hat{b}_1 + \frac{\alpha_1}{2}\hat{b}_1^{\dagger 2}\hat{b}_1^2 + \hbar\omega_2\hat{b}_2^\dagger\hat{b}_2 + \frac{\alpha_2}{2}\hat{b}_2^{\dagger 2}\hat{b}_2^2 + \hbar J(\hat{b}_1^\dagger\hat{b}_2 + \hat{b}_1\hat{b}_2^\dagger)\\ &+ \hbar\epsilon(t)\left[\hat{b}_1^\dagger e^{-i\omega_d t} + \hat{b}_1 e^{i\omega_d t}\right],\end{aligned} \tag{1.24}$$

where ω_i and α_i are respectively the frequency and anharmonicity of qubit $i \in \{1, 2\}$, and J describes the coupling amplitude between the qubits. The term on the second line of Eq. (1.24) represents a drive of frequency ω_d on qubit 1 with a time-dependent envelop $\epsilon(t)$. Moving to the rotating frame at the qubit frequencies and applying a Schrieffer-Wolff transformation (see Appendix B) to second order in J to diagonalize the undriven part of

the Hamiltonian, one arrives at the following Hamiltonian [5]:

$$\begin{aligned}\hat{H} =& \frac{\hbar\delta_1}{2}\hat{\sigma}_{z1} + \frac{\hbar\delta_2}{2}\hat{\sigma}_{z2} + \chi_{12}\hat{\sigma}_{z1}\hat{\sigma}_{z2} \\ &+ \hbar\epsilon(t)\left[\hat{\sigma}_{x1} - J'\hat{\sigma}_{x2} - \frac{E_{C1}}{\Delta_{12}}J'\hat{\sigma}_{z1}\hat{\sigma}_{x2}\right]\end{aligned} \tag{1.25}$$

where the transmons are truncated to their two first energy levels at the end of the full calculation, and where:

$$\frac{\delta_1}{2} = \omega_1 + \frac{J^2}{\Delta_{12}} + \chi_{12} - \omega_d \qquad \chi_{12} = \frac{J^2}{\Delta_{12} + E_{C2}} + \frac{J^2}{\Delta_{12} + E_{C1}} \tag{1.26}$$

$$\frac{\delta_2}{2} = \omega_2 - \frac{J^2}{\Delta_{12}} + \chi_{12} - \omega_d \qquad J' = \frac{J}{\Delta_{12} - E_{C1}}. \tag{1.27}$$

In Eq. (1.25), the CR term $\hat{\sigma}_{z1}\hat{\sigma}_{x2}$ appears with a prefactor that is relatively small. One therefore wants the detuning between the qubits to be small compared to the anharmonicity E_{C1}, but large enough compared to J. On the other hand, a small qubit-qubit detuning is not ideal on multiqubit-device, as frequency crowding might become an issue. Driving qubit 1 gives rise to an effective drive term in $J'\epsilon(t)$ on qubit two, however, protocols exists to eliminate this term [46, 47]. Note that the term χ_{12}, referred to as the ZZ cross-Kerr term, only appears when taking into account the higher transmon levels. This ZZ interaction term is one of the main factors limiting current gate fidelities [35–39]. It has to be made as small as possible, which, on the other hand, compromises the gate speed. Note, however, that mitigation strategies exist to cancel χ_{12}, for instance it can be dynamically suppressed by using a microwave drive [39] or by using qubits and/or couplers with opposite anharmonicities [35, 36, 48].

Using this scheme, CR gates are generated in about $200 - 400\,\text{ns}$ for relatively low gate fidelities $(94 - 96\%)$ [46] compared to flux-tuned gates. However, an advantage of all-microwave gates is that they can be performed on fixed-frequency qubits, reducing the qubit sensitivity to noise and ultimately increases their lifetimes. Moreover, fixed-frequency qubits simplify the chip design, it namely reduces the electronics, as the same drive lines can be used for single- and two-qubit gates.

1.3.3 Parametric gates

The main idea of parametric gates is to activate off-resonant interactions by modulating a parameter of the device at a specific frequency. The device parameters that are typically

modulated are the qubits or coupler frequencies (if tunable), or a coupling parameter. By appropriately choosing the frequency of the modulation, it is possible to generate transitions between energy levels which otherwise would be far off resonant and weakly coupled. To understand how a parametric gate is performed, we illustrate it with an example.

Let's consider a fixed-frequency qubit capacitively coupled to a frequency-tunable qubit:

$$\hat{H} = \frac{\hbar\omega_1(t)}{2}\hat{\sigma}_{z1} + \frac{\hbar\omega_2}{2}\hat{\sigma}_{z2} + J\hat{\sigma}_{x1}\hat{\sigma}_{x2}, \tag{1.28}$$

where we suppose $\omega_1(t) = \omega_1 + \epsilon\sin(\omega_d t)$. Moving to the rotating frame at the frequency of the qubits with the unitary:

$$\hat{U} = e^{-\frac{i\hbar}{2}\int_0^t dt'\omega_1(t')\hat{\sigma}_{z1}} e^{-\frac{i\hbar}{2}\omega_2\hat{\sigma}_{z2}}, \tag{1.29}$$

and using the Jacobi-Anger expansion $e^{iz\cos(\theta)} = \sum_{n=-\infty}^{\infty} i^n J_n(z)e^{in\theta}$, where J_n is the nth-order Bessel function, the effective Hamiltonian can now be expressed as follows:

$$\hat{H}' = J\sum_{n=-\infty}^{\infty} i^n J_n\left(\frac{\epsilon}{\omega_d}\right)\left[e^{i(\omega_1-\omega_2-n\omega_d)t}\hat{\sigma}_{+1}\hat{\sigma}_{-2} + e^{i(\omega_1+\omega_2-n\omega_d)t}\hat{\sigma}_{+1}\hat{\sigma}_{+2} + \text{h.c.}\right]. \tag{1.30}$$

We see that the modulation of the qubit frequency in Eq. (1.28) induces sidebands in Eq. (1.30), making it possible to activate interactions between the qubits by putting into resonance $n\omega_d = (\omega_1 \pm \omega_2)$. When the modulation of the qubit frequency is off ($\epsilon = 0$), only the first order Bessel remains in Eq. (1.30), and if the qubits are far detuned from each other compared to J, then the qubit interaction is weak. Note also that the sidebands have an effective coupling $JJ_n\left(\frac{\epsilon}{\omega_d}\right)$ which decreases with n.

In the example above, modulating the qubit frequency at $\omega_d = \omega_1 - \omega_2$ generates a $\sqrt{i\text{SWAP}}$ gate. For more complex Hamiltonians, additional interaction terms, other than $\hat{\sigma}_{+1}\hat{\sigma}_{-2}$ and $\hat{\sigma}_{+1}\hat{\sigma}_{+2}$, will appear in the effective Hamiltonian, making it possible to activate other gates by modulation the tunable parameter at the appropriate frequencies [19, 49–51]. Conditional phase gates (CZ) can typically be performed in about 100-400 ns with fidelities below 98%, and iSWAP gates in 40-250 ns and fidelities $\sim 99\%$ and below [19, 51]. Ref. [51] reports a 44 ns iSWAP gate and a 124 ns CZ gate with an average process fidelity of 99.3% and 97.9% respectively, by performing a parametric-resonance gate.

Chapter 2

Floquet theory to extract gate properties

One of the most direct approaches to numerically determine the time of a gate involves solving for the system dynamics. For instance, to determine the time of an iSWAP gate between two qubits, one would analyze the qubit populations and identify the moment at which the qubits have exchanged states. When optimizing the device parameters, for instance, to maximize the gate speed, solving for the system dynamics for different parameters is a resource-intensive process. Moreover, as the drive amplitude increases, the frequency at which the gate should be driven to achieve a high-fidelity gate, changes due to ac-Stark shifts [52] and Bloch-Siegert shifts [53]. Consequently, the gate dynamics has to be repeated for different drive frequencies to identify the one that results in a high-fidelity gate. Based on the need to efficiently explore device parameters, we have developed a method based on Floquet theory that enables us to extract gate rates significantly faster than standard numerical methods. This method takes advantage of the fact that gates we consider are activated by periodically driving the system.

In the first part of this chapter, we introduce the basic concepts of Floquet theory. In the second part, we explain how one can obtain interaction amplitudes using only the Floquet spectrum.

2.1 Floquet theory

In solid-state physics, Bloch's theorem is used to describe the behavior of electrons in periodic lattices [54]. More precisely the theorem states that the solutions to the Schrödinger equation of a space periodic Hamiltonian can be written as a product of a plane wave and a periodic function. Floquet's theorem is the analog of Bloch's theorem, with the difference that the Hamiltonian is now time-periodic instead of space-periodic [55, 56]. The theorem is formulated as follows:

Theorem 1 ***(Floquet's theorem)***

Let $H(t)$ be a continuous and T-periodic Hamiltonian in the Hilbert space $\mathcal{H}$, and let $|\Psi(t)\rangle$ be a solution to the Schrödinger equation:

$$i\hbar\frac{d}{dt}|\Psi(t)\rangle = H(t)|\Psi(t)\rangle. \tag{2.1}$$

There exists a complete set of solutions to Eq. (2.1), $\{\,|\Psi_\alpha(t)\rangle\,|\,\alpha \in [1,\cdots,dim(\mathcal{H})]\}$, that can be written as:

$$|\Psi_\alpha(t)\rangle = e^{-\frac{i}{\hbar}\epsilon_\alpha t}|\Phi_\alpha(t)\rangle, \tag{2.2}$$

where $|\Phi_\alpha(t)\rangle$ are T-periodic functions called the Floquet modes, and ϵ_α are the Floquet quasi-energies which are time-independent. The solutions $|\Psi_\alpha(t)\rangle$ are called Floquet states.

A certain number of properties and implications can be derived from this theorem [56–58]. We enumerate some of the properties here:

Properties:

1. The Floquet modes inherit the completeness of the Floquet states, for a given time t, since $\sum_\alpha |\Psi_\alpha(t)\rangle\langle\Psi_\alpha(t)| = \sum_\alpha |\Phi_\alpha(t)\rangle\langle\Phi_\alpha(t)| = I$, and the modes are orthogonal:

$$\langle\Phi_\alpha(t)|\Phi_\beta(t)\rangle = \delta_{\alpha\beta} \qquad \forall t \quad \text{for } \alpha,\beta \in [1,\cdots,dim(\mathcal{H})]. \tag{2.3}$$

2. The Floquet modes and quasi-energies are not unique, there exists an equivalence class of modes and quasi-energies that lead to the same Floquet state.

For any $k \in \mathbb{Z}$:

$$\begin{cases} |\Phi_{\alpha,k}(t)\rangle = e^{ik\omega t} |\Phi_\alpha(t)\rangle \\ \epsilon_{\alpha,k} = \epsilon_\alpha + k\omega \end{cases} \tag{2.4}$$

with $\omega = \frac{T}{2\pi}$, the Floquet mode and quasi-energy $|\Phi_{\alpha,k}(t)\rangle$ and $\epsilon_{\alpha,k}$ lead to the same Floquet state $|\Psi_\alpha(t)\rangle$. Because the quasi-energies are defined modulo ω, the Floquet spectrum is folded in what we can call the *first Briouillin zone* in analogy to Bloch's theorem.

3. Inserting Eq. (2.2) in Eq. (2.1), one finds that the Floquet modes and quasi-energies are solution to the eigenvalue problem of the Floquet Hamiltonian $H_F = \left[H(t) - i\hbar\frac{d}{dt}\right]$:

$$H_F |\Phi_\alpha(t)\rangle = \left[H(t) - i\hbar\frac{d}{dt}\right] |\Phi_\alpha(t)\rangle = \epsilon_\alpha |\Phi_\alpha(t)\rangle . \tag{2.5}$$

H_F exists in the tensor product space of $\mathcal{H}$ and of the square-integrable T-periodic functions space $\mathcal{H}_\mathcal{T}$, which we call $\mathcal{H}_\mathcal{F} = \mathcal{H} \otimes \mathcal{H}_\mathcal{T}$. We define the inner product in the composite space as:

$$\langle\langle \Phi_{\alpha,n} | \Phi_{\beta,m} \rangle\rangle \quad = \quad \frac{1}{T}\int_0^T dt \langle \Phi_{\alpha,n}(t) | \Phi_{\beta,m}(t) \rangle = \delta_{\alpha\beta}\delta_{nm}. \tag{2.6}$$

4. If two Floquet Hamiltonians $\left(\hat{H}'(t) - i\partial_t\right)$ and $\left(\hat{H}(t) - i\partial_t\right)$ are unitarily equivalent, *i.e.*:

$$\hat{H}'(t) - i\partial_t = e^{-\hat{S}(t)}\left[\hat{H}(t) - i\partial_t\right] e^{\hat{S}(t)}, \tag{2.7}$$

with $e^{-\hat{S}(t)}$ a time-dependent unitary transformation, then their quasi-energies are identical.

The Floquet modes and quasi-energies are obtained by diagonalizing the evolution operator $U(t)$ at the first period of the drive, since:

$$U(T)|\Psi_\alpha(0)\rangle = |\Psi_\alpha(T)\rangle \qquad \Leftrightarrow \qquad U(T)|\Phi_\alpha(0)\rangle = e^{-\frac{i}{\hbar}\epsilon_\alpha T}|\Phi_\alpha(0)\rangle. \tag{2.8}$$

The quasi-energies are $\epsilon_\alpha = i\frac{\hbar\omega}{2\pi}\log(\text{eig}_\alpha) \pmod{\omega}$, where eig_α is the eigenvalue of $U(T)$ associated to $|\Phi_\alpha(0)\rangle$. The Floquet modes at time $t \in [0, T]$ can then be obtained by evolving $|\Phi_\alpha(0)\rangle$ up to a time t and multiplying it by $e^{\frac{i}{\hbar}\epsilon_\alpha T}$. The fact that the Floquet energies are solutions to the eigenvalue problem of the Floquet Hamiltonian in Eq. (2.5) means that they hold information about the time dependence of the Hamiltonian. In fact, the quasi-energies contain information such as ac-stark shifts and Bloch-Siegert shifts. We show in the next

section how one can take advantage of this knowledge to extract interaction amplitudes.

An analogy can be made between the solutions of a time-independent Hamiltonian and the solutions obtained from a time-periodic Hamiltonian. Let H be a time-independent Hamiltonian, and $\{|\phi_n\rangle, E_n\}$ the solutions to the time-independent Schrödinger equation. Any solution to the time-dependent Schrödinger equation can be written in the form:

$$|\Psi(t)\rangle = \sum_n c_n e^{-\frac{i}{\hbar}E_n t}|\phi_n\rangle \tag{2.9}$$

with $c_n \in \mathbb{C}$. Similarly, any solution of Eq. (2.1) can be written as:

$$|\Psi(t)\rangle = \sum_\alpha c_\alpha e^{-\frac{i}{\hbar}\epsilon_\alpha t}|\Phi_n(t)\rangle, \tag{2.10}$$

where the eigenenergies are replaced by the quasi-energies and the static eigenstates by the time-periodic Floquet modes. In the limit where the Hamiltonian is time independent, Eq. (2.10) converges to Eq. (2.9).

In Property 2, we see that for any k, $\epsilon_{\alpha,k}$ is associated to the same physical Floquet state $|\Psi_\alpha(t)\rangle$, and should therefore be observable. These Floquet levels will be experimentally observed provided that a matrix element exists for the transition of interest $\Delta_{\alpha\beta k} = \epsilon_\alpha - \epsilon_\beta + k\omega_d$, where $k \in \mathbb{Z}$ can be interpreted as the number of photons added or removed by the drive. In the framework of Floquet theory, the matrix elements are defined as:

$$X_{\alpha\beta k} = \frac{1}{T}\int_0^T dt e^{-ik\omega_d t}\langle\Phi_\beta(t)|X_{\alpha\beta k}|\Phi_\alpha(t)\rangle. \tag{2.11}$$

Now that we have introduced some basic concepts of Floquet theory, we explain how one can obtain interaction amplitudes using only the Floquet quasi-energies.

2.2 Extracting interaction amplitudes

As we will see in Chapter 3, approximate analytical descriptions of superconducting devices are often insufficient to reach quantitative agreements with experiments, and numerical simulations become unavoidable. Moreover, exploring the parameter space searching for device parameters that lead to fast and high-fidelity gates can becomes a resource intensive process. For these reasons, we have developed a method based on the concepts of the previous section, which allows us to extract gates rates, and more generally interaction amplitudes, significantly faster than standard numerical method.

To explain the intuition behind this approach, we first introduce an illustrative example: a two-level system (TLS) coupled to a quadrature of a quantized field. This problem is time-independent, however, as we will see later, it can be related to a periodically driven TLS. The objective of this example is to provide insights on how the Floquet quasi-energies are related to the dressed spectrum.

A two-level system coupled to a field:

Such a system can be modeled by the Rabi Hamiltonian [59]:

$$H = \frac{\omega_q}{2}\sigma_z + \omega_d \hat{a}^\dagger \hat{a} + g\left(\hat{a} + \hat{a}^\dagger\right)\sigma_x, \tag{2.12}$$

where the two first terms respectively represent the Hamiltonian of a qubit of frequency ω_q and a field of frequency ω_d. The last term on the right corresponds to the qubit-field coupling of amplitude g. We denote the eigenstates of the qubit $\{|g\rangle, |e\rangle\}$, and the field eigenstates $\{\,|N\rangle$ for $N \in \mathbb{N}\}$. If we consider that the qubit and the field are uncoupled (*i.e.* $g = 0$), then the eigenstates of the Hamiltonian are $|q, N\rangle \equiv |q\rangle \otimes |N\rangle$ for $q \in \{e, g\}$ and $N \in \mathbb{N}$. In the case where the field is on resonance with the qubit, the unperturbed spectrum exhibits degeneracies between the bare levels $|g, N+1\rangle$ and $|e, N\rangle$ for any N, see Fig. 2.2 a). In Fig. 2.1 b), this degeneracy translates in an exact crossing of the eigenstates (dashed lines) at the resonance. Turning on the coupling, the eigenstates $|e, N\rangle$ and $|g, N+1\rangle$ interact and hybridize, leading to the dressed states $|g(N)\rangle$ and $|e(N)\rangle$ which are linear combinations of the bare states, see Fig. 2.2 a). The hybridization of the states leads to level repulsion which removes the degeneracy at resonance and translates in an avoided crossing of the dressed states (solid lines) in Fig. 2.1 b). An exact expression can be found for the energy separation Ω between the two levels in the case where the Rabi Hamiltonian is treated in the RWA. The Rabi Hamiltonian in the RWA is nothing else than the Jaynes-Cummings Hamiltonian,

$$H = \frac{\omega_q}{2}\sigma_z + \omega_d \hat{a}^\dagger \hat{a} + g\left(\hat{a}\hat{\sigma}_+ + \hat{a}^\dagger \hat{\sigma}_-\right), \tag{2.13}$$

which is exactly solvable, and for which the energy separation of the avoided crossing reads [5, 59]:

$$\Omega = \sqrt{\Omega_1^2 + \Delta^2}, \tag{2.14}$$

where $\Omega_1 = 2g\sqrt{N}$ is the *Rabi frequency*, which we will also refer to as the interaction amplitude, and $\Delta = \omega_d - \omega_q$ is the qubit-field detuning. When the field frequency is on resonance with the qubit, $\omega_d = \omega_q$, the interaction leads to a perfect swap of excitation between the two energy levels in a time that is inversely proportional to $\Omega_1/2$. In the case

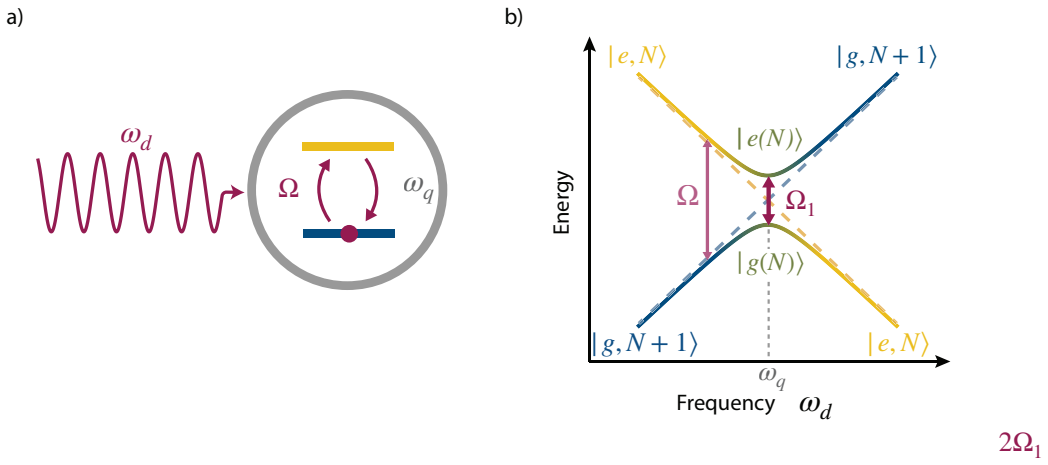


Figure 2.1 a) Illustration on a two-level system (gray) of interacting with a field (red). The excitation exchange happens with a rate Ω. The blue (yellow) level corresponds to the qubit ground (excited) state. b) Energy diagram of the qubit spectrum dressed by the drive as a function of the field frequency. The minimum distance in the avoided crossing is equal to $2\Omega_1$. The dashed lines correspond to the undriven spectrum while the solid lines are the driven spectrum. As the field and the qubit become off-resonant, the energy levels can be identified to the states at the extremities.

where the coupling becomes large compared to the qubit and coupler frequencies, Eq. (2.14) does not hold anymore for the Rabi Hamiltonian and the minimum distance in the avoided crossing is shifted from ω_q due to an important dressing of the energy levels, known as the Bloch-Siegert shift. This shift is of the order of $g^2/(\omega_q + \omega_d)$.

For any Hamiltonian, if the energy separation can be determined at the minimum distance in the avoided crossing then the interaction amplitude can be found. This is the main idea on which our method is based on, as it is adapted to a time-periodic system.

A driven two-level system:

An analogy can be made between the Rabi Hamiltonian in Eq. (2.12) and a driven qubit in the case where the field can be treated classically. Indeed, by moving the Rabi Hamiltonian to the rotating frame at ω_d, then displacing the field mode as $\hat{a} \to \hat{a} + \alpha$ with $\alpha \in \mathbb{C}$, and dropping quantum fluctuations ($\hat{a}$) in favor of the classical part (α), one obtains the Hamiltonian of a driven qubit:

$$H(t) = \frac{\omega_q}{2}\sigma_z + A\cos(\omega_d t)\sigma_x, \tag{2.15}$$

with A the amplitude of the drive. The spectrum of the driven system is dressed by the drive and as in Fig. 2.1 exhibits an avoided crossing as a function of the drive frequency. Even

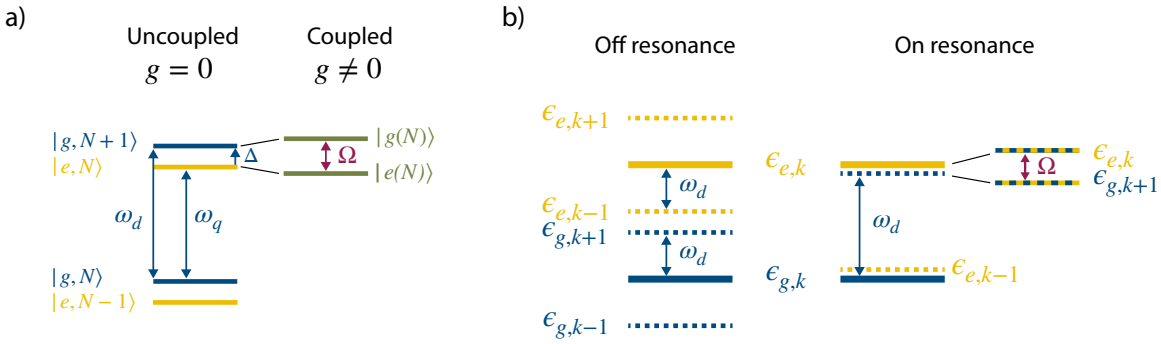


Figure 2.2 a) Diagram of the spectrum of a qubit coupled to a field. On the left the levels correspond to the uncoupled spectrum ($g = 0$) and on the right coupled spectrum ($g \neq 0$) with the dressed states $|g(N)\rangle$ and $|e(N)\rangle$. The letters g and e refer to the qubit ground state and excited states, and N to the number of excitations. Δ is the qubit-field detuning and Ω is the Rabi splitting. b) Diagram of the Floquet spectrum of a driven qubit. The energy levels in dashed lines correspond to Floquet energies modulo ω_d. If the drive couples the ground (g) and excited (e) states, then the quasi-energies $\epsilon_{e,k}$ and $\epsilon_{g,k+i}$ will repel each other when the drive is on resonance.

though the field mode ($\hat{a}^\dagger$ and $\hat{a}$) does not appear anymore in Eq. (2.15), some of its features can be found in the quasi-energies. In particular, in the quasi-energy $\epsilon_{\alpha,k} = \epsilon_\alpha + k\omega_d$, the integer k can be thought of as the number of excitations originating from the drive, similar to the number of field excitations N in the previous example.

As mentioned in Sect. 2.1, the quasi-energies account for drive effects and therefore effectively reproduce the dressed spectrum. By determining the Floquet energies at the minimum distance in the avoided crossing, we can extract the interaction amplitude, between the levels in question and ultimately determine the gate time. In the following chapters, we will use the term *gate rate* to denote the interaction amplitude. As seen previously, finding the quasi-energies only involves the numerical computation of the evolution operator of Eq. (2.15) for a single period of the drive. In Chapter 3, we consider gates using GHz drives, which typically corresponds to drive periods of $T \sim 1\,\text{ns}$. The drive period is thus around 1 to 3 orders of magnitude shorter than typical gate times. Extracting the gate rates with this Floquet method is thus significantly more efficient than standard numerical methods based on evolving the gate dynamics for times much longer than T. This method can naturally be extended to any time-periodic Hamiltonian and proves to be useful as long as the gate time is shorter than the period.

Also, more generally, other interactions such as cross-Kerr interactions can be formulated as linear combinations of the quasi-energies. For instance, in the undriven case, the ZZ

cross-Kerr interaction between two qubits is written as:

$$ZZ = E_{11} - E_{01} - E_{10} + E_{00} \tag{2.16}$$

where E_{ij} is the energy of the computational state $|ij\rangle$, with $i, j \in \{0, 1\}$. This relation can be extended to the driven case using the quasi-energies:

$$ZZ = \epsilon_{11} - \epsilon_{01} - \epsilon_{10} + \epsilon_{00}, \tag{2.17}$$

to extract the dynamical ZZ that is introduced by the drive.

Note that the Floquet spectrum can also be used to find the optimal drive frequency to implement a gate, as the frequency might be altered due to Bloch-Siegert shifts in the case of a strong drive. This is done by plotting the avoided crossing as a function of the drive frequency to find ω_d at which the avoided crossing is minimal and the interaction therefore maximal.

In the following chapter, we apply this Floquet method to numerically determine the gate time and the dynamical ZZ of a superconducting device, and to extract the shift in the drive frequency as the case of strong drives.

Chapter 3

Fast and high-fidelity parametric two-qubit gate design

The work presented in this chapter of the thesis was done in collaboration with the group of Andrew Houck at Princeton University and with members of the Blais group at the Université de Sherbrooke. In Sherbrooke, Alexandru Petrescu, Catherine Leroux and Agustin Di Paolo contributed in designing and developing the device model presented below. Alexandru Petrescu is one of the main contributors who developed the time-dependent Schrieffer-Wolff perturbation theory to analytically describe the designed device. At Princeton University, Pranav Mundada, Andrei Vrajitoarea, Sara Sussman and Charles Guinn, contributed to this chapter with useful discussions related to the device design.

The architecture we explore is illustrated in Fig. 3.1. It consists in two fixed-frequency transmon qubits labeled a and b (green) and a coupler labeled c (blue). The coupler is made of an array of Josephson junctions in parallel with a single JJ, which are shunted with a capacitance. The loop formed by the JJs can be threaded by a magnetic flux, and as we will see later, we will activate gates by modulating the flux through the loop. The qubits and the coupler are capacitively connected to each other. The Hamiltonian of this circuit is derived in Appendix A by using the standard approach [1, 2] and reads:

$$\hat{H}(t) = \hat{H}_a + \hat{H}_b + \hat{H}_c(t) + \hat{H}_g, \tag{3.1}$$

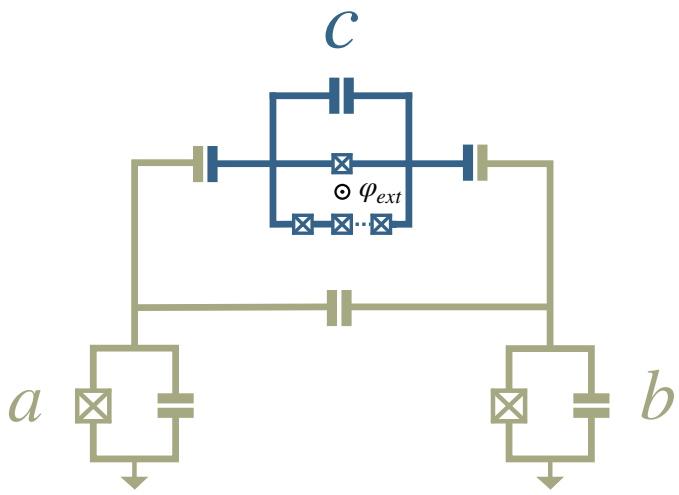


Figure 3.1 Device design. The qubits a and b (green) are fixed-frequency transmons capacitively coupled to a SNAIL used as a coupler (blue). The coupler c has a superconducting loop, made of N JJs in parallel with a black sheep junction, that is threaded by a magnetic flux φ_{ext}.

where the transmons and the coupler are described by,

$$\begin{aligned} \hat{H}_j &= 4E_{\mathrm{C}j}\hat{\boldsymbol{n}}_j^2 - E_{\mathrm{J}j}\cos(\hat{\boldsymbol{\varphi}}_j), \quad \text{for } j = a, b, \\ \hat{H}_c(t) &= 4E_{\mathrm{C}c}\hat{\boldsymbol{n}}_c^2 - \alpha E_{\mathrm{J}c}\cos\left[\hat{\boldsymbol{\varphi}}_c + \mu_\alpha\varphi_{\mathrm{ext}}(t)\right] - \beta N E_{\mathrm{J}c}\cos\left[\frac{\hat{\boldsymbol{\varphi}}_c}{N} + \mu_\beta\varphi_{\mathrm{ext}}(t)\right]. \end{aligned} \tag{3.2}$$

and where the coupling between the qubits and the coupler is,

$$\hat{H}_g = 4E_{\mathrm{C}ab}\hat{\boldsymbol{n}}_a\hat{\boldsymbol{n}}_b + 4E_{\mathrm{C}bc}\hat{\boldsymbol{n}}_b\hat{\boldsymbol{n}}_c + 4E_{\mathrm{C}ac}\hat{\boldsymbol{n}}_a\hat{\boldsymbol{n}}_c. \tag{3.3}$$

The operators $\hat{\boldsymbol{\varphi}}_j$ and $\hat{\boldsymbol{n}}_j$ are canonically conjugate and represent the superconducting phase difference, resp. the Cooper pair number for the bare mode of the transmons for the indices $j = a, b$ and of the coupler for $j = c$. They satisfy the commutation relations $[\hat{\boldsymbol{\varphi}}_j, \hat{\boldsymbol{n}}_k] = i\delta_{jk}$, where we set $\hbar = 1$. The parameters $E_{\mathrm{C}j}$ are the charging energies, which depend on the device capacitances (see Appendix A). The Josephson energies are denoted $E_{\mathrm{J}a}$, $E_{\mathrm{J}b}$ for the transmon modes, whereas $\beta E_{\mathrm{J}c}$ is the Josephson energy of one of the junctions in the N-junction array of the coupler. The parameter β corresponds to a renormalization of the superinductance due to disorder in the junction array and to finite zero-point fluctuations, as mentioned in the subsection about the generalized flux qubit in Sect. 1.2 [34]. The parameter α is a factor parametrizing the anisotropy between the two branches of the coupler, and accounts for a renormalization of the small junction energy due to hybridization with the

modes in the junction array. The two branches of the coupler involving Josephson junctions form a loop that can be threaded by an external flux, which we take to be:

$$\varphi_{\text{ext}}(t) = \overline{\varphi}_{\text{ext}} + \delta\varphi \sin(\omega_d t) \tag{3.4}$$

where $\overline{\varphi}_{\text{ext}}$ is a dc flux, ω_d is the frequency of the modulation and $\delta\varphi$ is its amplitude, which is assumed to be small compared to the flux quantum. As shown in Sect. 1.3, two-qubit gates can be activated by varying $\overline{\varphi}_{\text{ext}}$, or by parametrically modulating φ_{ext}. In the context of our work and in the following subsections, we only focus on parametric gates.

To provide intuition about the device properties, we first approximate the qubits and the coupler as Kerr nonlinear oscillators in a model which we refer to as the 3KNO model. This simple toy model captures the essential features of the Hamiltonian, and will give us a first intuition of the device properties. As part of the analytical description of the device, we introduce a new method based on a time-dependent Schrieffer-Wolff transformation to derive an effective Hamiltonian. This method allows us to identify the contribution of each normal mode (dressed by the drive) to the different effective interaction constants [48]. After studying the toy model, we apply the time-dependent Schrieffer-Wolff method to the full device model and extract the relevant interaction terms.

3.1 Toy model: three Kerr non-linear oscillators

3.1.1 Analytics

A) The model

As seen in Sect. 1.2, the transmon qubit can be approximated as a Kerr nonlinear oscillator (KNO) when operated in the transmon regime ($E_J \gg E_C$):

$$\hat{H}_j = \boldsymbol{\omega}_j \hat{\boldsymbol{j}}^\dagger \hat{\boldsymbol{j}} + \frac{\boldsymbol{\alpha}_j}{2} \hat{\boldsymbol{j}}^{\dagger 2} \hat{\boldsymbol{j}}^2 \quad \text{for} \quad j \in \{a, b\}, \tag{3.5}$$

where $\boldsymbol{\omega}_j = \sqrt{8E_{\text{J},}E_{\text{C}j}} - E_{\text{C}j}$ is the frequency of qubit $j \in \{a, b\}$ and its anharmonicity is $\boldsymbol{\alpha}_j = -E_{\text{C}j} < 0$. Similarly, we approximate the SNAIL coupler as a Kerr non-linear oscillator with a tunable frequency,

$$\hat{H}_c(t) = \boldsymbol{\omega}_c(t) \hat{\boldsymbol{c}}^\dagger \hat{\boldsymbol{c}} + \frac{\boldsymbol{\alpha}_c}{2} \hat{\boldsymbol{c}}^{\dagger 2} \hat{\boldsymbol{c}}^2. \tag{3.6}$$

Because of the flux tunability of the coupler, its parameters can be varied in time, we therefore model its frequency as:

$$\boldsymbol{\omega}_c(t) = \boldsymbol{\omega}_c + \delta \sin(\omega_d t), \tag{3.7}$$

where $\boldsymbol{\omega}_c$ can be tuned with a dc flux and, where ω_d and δ parametrize the frequency and the amplitude at which the coupler frequency is modulated in time. We will refer to the modulated term in Eq. (3.6) as the drive and will call ω_d and δ the drive frequency and amplitude. Because the coupler potential in Eq. (3.2) is a cosine, a periodic modulation of the flux will generate multiple harmonics of the coupler frequency $\boldsymbol{\omega}_c(t)$. Here we have neglected higher harmonics of $\boldsymbol{\omega}_c(t)$ and assume the coupler frequency to be of the form above. Also, for the sake of simplicity we neglect the time dependence of the coupler anharmonicity in the toy model but we will take it into account in the full device model in Sect. 3.2. Importantly, in contrast to the transmons, the anharmonicity of the coupler α_c can be positive.

Now, dropping all photon number nonconserving terms in the coupling,

$$\hat{H}_g = -\boldsymbol{g}_{ac}\hat{\boldsymbol{a}}^\dagger\hat{\boldsymbol{c}} - \boldsymbol{g}_{bc}\hat{\boldsymbol{b}}^\dagger\hat{\boldsymbol{c}} - \boldsymbol{g}_{ab}\hat{\boldsymbol{a}}^\dagger\hat{\boldsymbol{b}} + \text{ h.c.}, \tag{3.8}$$

the full Hamiltonian of this toy model finally reads:

$$\begin{aligned}\hat{H}_{\text{3KNO}} =& \boldsymbol{\omega}_a\hat{\boldsymbol{a}}^\dagger\hat{\boldsymbol{a}} + \frac{\boldsymbol{\alpha}_a}{2}\hat{\boldsymbol{a}}^{\dagger 2}\hat{\boldsymbol{a}}^2 \;+\; \boldsymbol{\omega}_b\hat{\boldsymbol{b}}^\dagger\hat{\boldsymbol{b}} + \frac{\boldsymbol{\alpha}_b}{2}\hat{\boldsymbol{b}}^{\dagger 2}\hat{\boldsymbol{b}}^2 \;+\; \boldsymbol{\omega}_c(t)\hat{\boldsymbol{c}}^\dagger\hat{\boldsymbol{c}} + \frac{\boldsymbol{\alpha}_c}{2}\hat{\boldsymbol{c}}^{\dagger 2}\hat{\boldsymbol{c}}^2 \\ &- \left(\boldsymbol{g}_{ac}\hat{\boldsymbol{a}}^\dagger\hat{\boldsymbol{c}} + \boldsymbol{g}_{bc}\hat{\boldsymbol{b}}^\dagger\hat{\boldsymbol{c}} + \boldsymbol{g}_{ab}\hat{\boldsymbol{a}}^\dagger\hat{\boldsymbol{b}} \;+\; \text{h.c.}\right).\end{aligned} \tag{3.9}$$

The modes in bold used until now are the bare qubit and coupler modes. The bare frequencies, anharmonicities and couplings are also indicated in bold.

B) Black-box quantization approach

To correctly take into account for the couplings between the circuit modes in our model, we use a black-box quantization approach [60]. The idea consists in treating the anharmonicities pertubatively and diagonalizing the rest of the Hamiltonian which includes the coupling terms, such as to re-express the Hamiltonian in the dressed basis. We thus decompose the Hamiltonian in two parts $\hat{H} = \hat{H}^{(0)} + \lambda\hat{H}^{(1)}$ with:

$$\begin{aligned}\hat{H}^{(0)} &= \boldsymbol{\omega}_a\hat{\boldsymbol{a}}^\dagger\hat{\boldsymbol{a}} + \boldsymbol{\omega}_b\hat{\boldsymbol{b}}^\dagger\hat{\boldsymbol{b}} + \boldsymbol{\omega}_c\hat{\boldsymbol{c}}^\dagger\hat{\boldsymbol{c}} - \left(\boldsymbol{g}_{ac}\hat{\boldsymbol{a}}^\dagger\hat{\boldsymbol{c}} + \boldsymbol{g}_{bc}\hat{\boldsymbol{b}}^\dagger\hat{\boldsymbol{c}} + \boldsymbol{g}_{ab}\hat{\boldsymbol{a}}^\dagger\hat{\boldsymbol{b}} + \text{H.c.}\right) \\ \lambda\hat{H}^{(1)} &= \frac{\boldsymbol{\alpha}_a}{2}\hat{\boldsymbol{a}}^{\dagger 2}\hat{\boldsymbol{a}}^2 + \frac{\boldsymbol{\alpha}_b}{2}\hat{\boldsymbol{b}}^{\dagger 2}\hat{\boldsymbol{b}}^2 + \frac{\boldsymbol{\alpha}_c}{2}\hat{\boldsymbol{c}}^{\dagger 2}\hat{\boldsymbol{c}}^2 + \delta\sin(\omega_d t)\hat{\boldsymbol{c}}^\dagger\hat{\boldsymbol{c}},\end{aligned} \tag{3.10}$$

where λ is a dimensionless parameter that we will set to 1 later, but will be useful to

keep track of the order of the perturbation theory. To find the normal modes dressed by the couplings, we diagonalize $\hat{H}^{(0)}$ to eliminate the coupling terms $\boldsymbol{g}_{ij}$. This leads to $\hat{H}^{(0)} = \omega_a \hat{a}^\dagger \hat{a} + \omega_b \hat{b}^\dagger \hat{b} + \omega_c \hat{c}^\dagger \hat{c}$, where ω_j and $\hat{j}$ for $j \in \{a, b, c\}$, are the dressed frequencies and modes. The bare and dressed modes can be expressed in relation to another as $\hat{\boldsymbol{\alpha}} = \sum_{\beta=a,b,c} u_{\alpha\beta} \hat{\beta}$, where the hybridization coefficients $u_{\alpha\beta}$, are either obtained numerically or by using symbolic computation tools like the Sneg package in Mathematica [61]. Even for a 3-mode problem, the analytical expressions for $u_{\alpha\beta}$ are rather complicated functions of $\boldsymbol{\omega}_j$ and $\boldsymbol{g}_{ij}$ and are therefore not reproduced here. Expressing the perturbation $\lambda H^{(1)}$ in terms of the normal modes, we have:

$$\begin{aligned} \lambda \hat{H}^{(1)} = \sum_{j=a,b,c} \frac{\alpha_j}{2} \left(u_{ja}\hat{a} + u_{jb}\hat{b} + u_{jc}\hat{c} \right)^{\dagger 2} \left(u_{ja}\hat{a} + u_{jb}\hat{b} + u_{jc}\hat{c} \right)^2 \\ + \delta \sin(\omega_d t) \left(u_{ca}\hat{a} + u_{cb}\hat{b} + u_{cc}\hat{c} \right)^\dagger \left(u_{ca}\hat{a} + u_{cb}\hat{b} + u_{jc}\hat{c} \right). \end{aligned} \tag{3.11}$$

This approach is valid if the anharmonicities and the drive amplitude can be treated pertubatively compared to the qubit-coupler couplings.

C) Time-dependent Schrieffer-Wolff transformation

In the previous subsection, we diagonalized the unperturbed Hamiltonian and expressed the perturbed Hamiltonian in the dressed-mode basis. As can be seen in Eq. (3.11), in this basis the perturbation now possesses counter-rotating terms. We now remove those fast oscillating terms by transforming the perturbed Hamiltonian. An effective Hamiltonian can be derived by performing a perturbative expansion of the unitary transformation, generally known as Schrieffer-Wolff perturbation theory (SWPT), to successively correct for the fast-rotating terms. More details can be found about the SWPT in Appendix B. In the context of Eq. (3.11), the perturbation is time-dependent, we will therefore perform a time-dependent SW transformation as explained below.

Keeping the partition of the Hamiltonian chosen in Eq. (3.10), the next step now consists in moving the Floquet Hamiltonian to the interaction picture with respect to the unperturbed Hamiltonian:

$$\begin{aligned} \lambda \hat{H}_I^{(1)}(t) - i\partial_t &= \hat{U}_0(t) \left[\hat{H} - i\partial_t \right] \hat{U}_0^\dagger(t) \\ &= \hat{U}_0(t) \lambda \hat{H}^{(1)}(t) \hat{U}_0^\dagger(t) - i\partial_t. \end{aligned} \tag{3.12}$$

where $\hat{U}_0(t) = e^{-i\int_0^t dt' \hat{H}^{(0)}}$. In order to get an effective Hamiltonian while getting rid of the fast-oscillating terms in Eq. (3.12), we apply a Schrieffer-Wolff (SW) transformation on the

interaction-picture Floquet Hamiltonian:

$$\hat{H}_{I,eff} - i\partial_t = e^{-i\hat{S}(t)} \left[\lambda \hat{H}_I^{(1)}(t) - i\partial_t\right] e^{i\hat{S}(t)}. \tag{3.13}$$

The exponentials are Taylor expanded and reordered with the Baker-Campbell-Hausdor (BCH) expansion. The generator $\hat{S}(t)$ can be written in powers of λ and solved iteratively to perform the RWA, see Appendix B for more details. Note that according to Property 4 of Chapter 2, the Floquet Hamiltonians $\left(\hat{H}_{I,eff} - i\partial_t\right)$ and $\left(\lambda \hat{H}_I^{(1)}(t) - i\partial_t\right)$ are unitarily equivalent and therefore have the same quasi-energies. In a certain sense, the perturbative expansion of $e^{-\hat{S}(t)}$ that we perform in Eq. (3.13) can be seen as an iterative approach to finding the Floquet spectrum. This procedure leads to:

$$\hat{H}_{I,eff} = \lambda \overline{\hat{H}}_I^{(1)} + \lambda^2 \bar{\hat{H}}_I^{(2)} + \cdots, \tag{3.14}$$

where $\overline{\hat{H}}_I^{(1)}$ is the time-independent part of $\hat{H}_I^{(1)}$,

$$\overline{\hat{H}}_I^{(1)} = \lim_{T\to\infty} \frac{1}{T} \int_0^T dt \hat{H}_I^{(1)}(t) \tag{3.15}$$

and which together with the oscillatory part $\widetilde{\hat{H}}_I^{(1)}$, gives:

$$\lambda \hat{H}_I^{(1)}(t) = \lambda \overline{\hat{H}}_I^{(1)} + \lambda \widetilde{\hat{H}}_I^{(1)}(t). \tag{3.16}$$

The generator $\hat{S}(t)$ at order one is chosen such as to cancel $\widetilde{\hat{H}}_I^{(1)}$. This procedure can be repeated iteratively order by order. The second order RWA term in the effective Hamiltonian then reads:

$$\lambda^2 \overline{\hat{H}}_I^{(2)} = \frac{1}{i} \overline{\left[\overline{\hat{H}}_I^{(1)}, \int_0^t \lambda \widetilde{\hat{H}}_I^{(1)}(t')dt'\right]} + \frac{1}{2i} \overline{\left[\lambda \widetilde{\hat{H}}_I^{(1)}(t), \int_0^t \lambda \widetilde{\hat{H}}_I^{(1)}(t')dt'\right]}. \tag{3.17}$$

D) Effective Hamiltonian

As an example we consider a drive modulation $\omega_d = \omega_b - \omega_a$. This modulation frequency will bridge the energy difference between the two qubits, allowing the iSWAP interaction $-i\left(\hat{a}^\dagger \hat{b} - \hat{b}^\dagger \hat{a}\right)$ to be resonant. We now apply the time-dependent SW transformation

described above on the 3KNO Hamiltonian and obtain:

$$\begin{aligned}\lambda\overline{\hat{H}}_I^{(1)} &= -iJ_{ab}^{(1)}\left(\hat{a}^\dagger\hat{b}-\hat{b}^\dagger\hat{a}\right)\\ &+\frac{\alpha_a^{(1)}}{2}\hat{a}^{\dagger 2}\hat{a}^2+\frac{\alpha_b^{(1)}}{2}\hat{b}^{\dagger 2}\hat{b}^2+\frac{\alpha_c^{(1)}}{2}\hat{c}^{\dagger 2}\hat{c}^2\\ &+\chi_{ab}^{(1)}\hat{a}^\dagger\hat{a}\hat{b}^\dagger\hat{b}+\chi_{bc}^{(1)}\hat{b}^\dagger\hat{b}\hat{c}^\dagger\hat{c}+\chi_{ca}^{(1)}\hat{c}^\dagger\hat{c}\hat{a}^\dagger\hat{a},\end{aligned} \tag{3.18}$$

where

$$J_{ab}^{(1)} = u_{ca}u_{cb}\frac{\delta}{2} \tag{3.19}$$

$$\alpha_j^{(1)} = \sum_{i=a,b,c} u_{ij}^4\alpha_i \quad \text{for } j=[a,b,c] \tag{3.20}$$

$$\chi_{jk}^{(1)} = \sum_{i=a,b,c} 2u_{ij}^2u_{ik}^2\alpha_i \quad \text{for } j,k=[a,b,c] \text{ and } j\neq k. \tag{3.21}$$

The first term of Eq. (3.18) corresponds to the iSWAP interaction of amplitude $J_{ab}^{(1)}$, which is linear in the drive amplitude δ, meaning that the interaction is off when undriven. The second line of Eq. (3.18) corresponds to the mode anharmonicities (the self-Kerr) and $\chi_{jk}^{(1)}$ the cross-Kerr interactions between modes j and k. The latter correspond to state-dependent frequency shift of mode j due to the presence of an excitation in mode k. The cross-Kerr interaction between the qubits, χ_{ab}, is usually referred to as the ZZ interaction.

In order to perform a fast and high fidelity iSWAP gate one wants to respectively maximize $J_{ab}^{(1)}$ while minimizing $\chi_{jk}^{(1)}$. The presence of the unwanted ZZ interaction results in a shift in the qubit frequencies, which prevents the flawless implementation of the iSWAP gate and ultimately diminishes the gate fidelity. If we first take a look at the gate rate in Eq. (3.19), we see that increasing the drive amplitude increases $J_{ab}^{(1)}$, thereby decreasing the gate time. However, the drive amplitude cannot be infinitely increased to speedup the gate, as the drive can then no longer be treated perturbatively and Eq. (3.19) breaks down. Now considering the cross-Kerr interactions in Eq. (3.21), we see that they are independent on the drive amplitude to first order in the RWA. Moreover they are written as linear combinations of qubits and coupler anharmonicities, with positive prefactors. This implies that the only way to get rid of the ZZ interactions is to have qubits and a coupler with opposite anharmonicities. Since transmons have by nature negative anharmonicities, a coupler with positive anharmonicity is necessary to cancel the cross-Kerr. However, a positive anharmonicity coupler alone is not a sufficient condition, as it also needs to compensate the hybridization coefficients u_{ij} that appear in the sum.

We now derive second order corrections to the effective Hamiltonian. Going to second

order in perturbation theory, we find that the iSWAP rate gets no additional contribution $\left(J_{ab}^{(2)}=0\right)$, while the cross-Kerr interaction between the qubit is modified by:

$$\begin{aligned}\chi_{ab}^{(2)} =& 2\frac{\left(\sum_{j=a,b,c} u_{aj}u_{bj}u_{cj}^2\alpha_j\right)^2}{\omega_a+\omega_b-2\omega_c}+\delta\frac{u_{ac}u_{bc}\sum_{j=a,b,c}u_{aj}u_{bj}\left(u_{aj}^2-u_{bj}^2\right)}{\omega_a-\omega_b}\\ &+4\frac{\left(\sum_{j=a,b,c} u_{aj}u_{bj}^2u_{cj}\alpha_j\right)^2}{\omega_a-\omega_c}+4\frac{\left(\sum_{j=a,b,c} u_{aj}^2u_{bj}u_{cj}\alpha_j\right)^2}{\omega_b-\omega_c}\\ &-2\frac{\left(\sum_{j=a,b,c} u_{aj}^3u_{bj}\alpha_j\right)^2}{\omega_a-\omega_b}+2\frac{\left(\sum_{j=a,b,c} u_{aj}u_{bj}^3\alpha_j\right)^2}{\omega_a-\omega_b}.\end{aligned} \tag{3.22}$$

The dominant terms are on the first line. As will be seen later, the coupler is generally placed between the two qubit frequencies and since in our case the qubits are far detuned in frequency and the coupler is close to the qubit frequencies, the denominator $(\omega_a+\omega_b-2\omega_c)$ will be small. Notice also that the second term of the first line is the lowest order contribution to what we call the *dynamical* cross-Kerr interaction, which is induced by the drive. Indeed, $\chi_{ab}^{(1)}$ does not inherit any feature from the drive while $\chi_{ab}^{(2)}$ shows a linear dependence with the drive amplitude δ. The drive induces a change in the ZZ interaction which is only captured to second order in perturbation theory.

E) Improving the effective Hamiltonian

The effective Hamiltonian derived above is already a good approximation and gives us some intuition of how the quantities of interest (J_{ab} and χ_{ab}) depend on the system and drive parameters. As mentioned earlier, one could have chosen other partitions for the (un)perturbed Hamiltonian in Eq. (3.10). For instance, using a partition inspired from the corrections obtained from the previous perturbation theory, we can get better agreements with exact numerics. More precisely, we choose a new partition where $\hat{H}^{(0)}$ now also contains the terms from the perturbation $\hat{H}^{(1)}$ of the previous partition that are diagonal in Fock space (Eq. (3.11)). The unperturbed Hamiltonian in the new partition reads:

$$\begin{aligned}\hat{H}^{(0)} =& \omega_a\hat{a}^\dagger\hat{a}+\omega_b\hat{b}^\dagger\hat{b}+\omega_c\hat{c}^\dagger\hat{c}\\ &+\delta\sin(\omega_d t)\left[u_{ca}^2\hat{a}^\dagger\hat{a}+u_{cb}^2\hat{b}^\dagger\hat{b}+u_{cc}^2\hat{c}^\dagger\hat{c}\right]\\ &+\frac{\alpha_a^{(0)}}{2}\hat{a}^{\dagger 2}\hat{a}^2+\frac{\alpha_b^{(0)}}{2}\hat{b}^{\dagger 2}\hat{b}^2+\frac{\alpha_c^{(0)}}{2}\hat{c}^{\dagger 2}\hat{c}^2\\ &+\chi_{ab}^{(0)}\hat{a}^\dagger\hat{a}\hat{b}^\dagger\hat{b}+\chi_{ac}^{(0)}\hat{a}^\dagger\hat{a}\hat{c}^\dagger\hat{c}+\chi_{bc}^{(0)}\hat{b}^\dagger\hat{b}\hat{c}^\dagger\hat{c}\end{aligned} \tag{3.23}$$

where $\alpha_j^{(0)}$ and $\chi_{ij}^{(0)}$ are the anharmonicities and cross-Kerr interactions obtained in Eq. (3.19)-

Eq. (3.21) that we relabeled to keep track of the order in the perturbation theory. The perturbed Hamiltonian now includes all the terms that appear in the expansion of Eq. (3.11) that are not diagonal in Fock space. Following the same procedure as before but with the new partition, we obtain:

$$J_{ab}^{(1)} = \frac{\delta}{2} u_{ca} u_{cb} \left[J_0\left(\frac{\delta u_{ca}^2}{\omega_d}\right) J_0\left(\frac{\delta u_{cb}^2}{\omega_d}\right) + 3 J_1\left(\frac{\delta u_{ca}^2}{\omega_d}\right) J_1\left(\frac{\delta u_{cb}^2}{\omega_d}\right) \right], \tag{3.24}$$

$$\chi_{ab}^{(1)} = 0 \tag{3.25}$$

$$J_{ab}^{(2)} = \frac{i\delta^2}{2} u_{ca} u_{cb} u_{cc}^2 \left(\frac{1}{\omega_a - \omega_c} - \frac{1}{\omega_b - \omega_c}\right) J_1\left(\frac{\delta u_{cc}^2}{\omega_d}\right) \prod_{j=a,b,c} J_0\left(\frac{\delta u_{cj}^2}{\omega_d}\right) + \dots \tag{3.26}$$

$\chi_{ab}^{(2)}$ contains too many terms to be written here, but can easily be calculated using symbolic computation. For clarity reasons, we only wrote a few terms of $J_{ab}^{(2)}$. The Bessel functions appear from a Jacobi-Anger expansion of the drive term in $\hat{H}^{(0)}$, when moving to the interaction picture. In fact, when moving to the interaction picture with $\hat{U}_0(t) = e^{-i\int_0^t dt' \hat{H}^{(0)}(t')}$, we expand the time-dependent exponentials using the Jacobi-Anger identity:

$$e^{i\hat{O}\frac{\delta}{\omega_d}\cos\omega_d t} = J_0\left(\frac{\hat{O}\delta}{\omega_d}\right) + 2\sum_{n=1}^{\infty} i^n J_n\left(\frac{\hat{O}\delta}{\omega_d}\right)\cos n\omega_d t, \tag{3.27}$$

for any operator $\hat{O}$, and where $J_n(z)$ is the nth Bessel function of the first kind. Since $\omega_d \delta \ll 1$, we truncate the expansion after the first Bessel function. We see that Eq. (3.24) reproduces similar results to the one obtained in Eq. (3.19), up to linear terms in δ. The cross-Kerr interaction between the qubits gets corrections to $\chi_{ab}^{(0)}$ to second order only.

F) Comparison to exact numerics

To evaluate the accuracy of the time-dependent perturbation theory described above, we compare it at different orders to numerical results. In Fig. 3.2, we show the gate rate and, the static and dynamical ZZ interaction versus the coupler frequency, for a given set of device parameters which can be found in the caption. In Fig. 3.2 a), the ZZ interaction is plotted versus the coupler frequency, in the absence of a drive. The blue dots are obtained numerically from the exact diagonalization of the Hamiltonian using 5 levels per qubit and coupler, while the full lines correspond to the analytical formulae of Eq. (3.21) and Eq. (3.22), considering first- and second-order perturbation theory. We find that to second order, the analytics and the numerics are already in very good agreement. The resonances at 4 and 5.5 GHz correspond to the qubit frequencies, and the resonance that appears approximately at $\boldsymbol{\omega}_c \approx (\boldsymbol{\omega}_a + \boldsymbol{\omega}_b)/2 = 2\pi \times 4.75\,\text{GHz}$, corresponds to a virtual two-photon excitation of

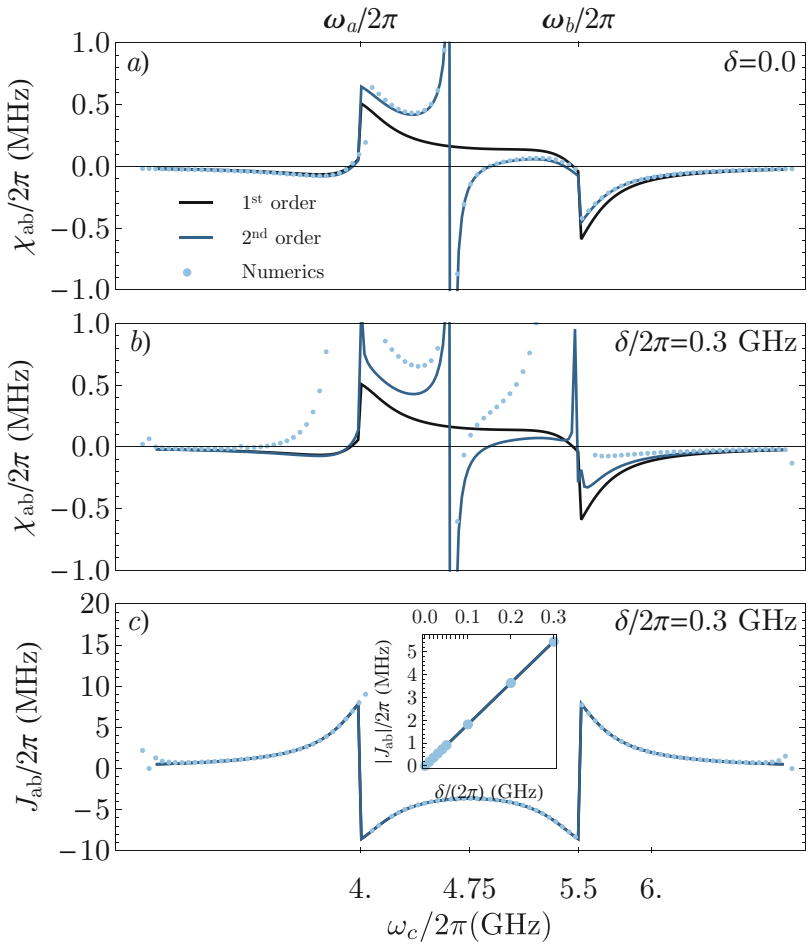


Figure 3.2 a) Static cross-Kerr interaction $\chi_{ab}(\omega_c)$, from first (black) and second-order RWA (blue), and from the full diagonalization of Eq. (2.5) (light blue points) for: $\boldsymbol{\omega}_a/2\pi = 4.0$, $\boldsymbol{\omega}_b/2\pi = 5.5$, $\boldsymbol{\alpha}_a/2\pi = -0.3$, $\boldsymbol{\alpha}_b/2\pi = -0.2$, $\boldsymbol{\alpha}_c/2\pi = 0.25$, $\boldsymbol{g}_{ab}/2\pi = 0.12$, $\boldsymbol{g}_{bc}/2\pi = -0.12$, and $\boldsymbol{g}_{ac}/2\pi = 0$, all in GHz. b) Analogue of a) for the *dynamical* cross-Kerr interaction at $\delta/2\pi = 0.3$ GHz. c) Same as b) for the gate interaction rate $J_{ab}(\omega_c)$. Inset: $J_{ab}(\delta)$ at $\boldsymbol{\omega}_c/2\pi = 4.25$ GHz. From Ref. [48].

the coupler mode.

Fig. 3.2 b) and c) depict the ZZ cross-Kerr interaction and the gate rate as a function of the coupler frequency, but now at $\delta \neq 0$. For the numerical results, we now use the Floquet numerics described in Chapter 2. First we find that the gate rate is in excellent agreement with exact Floquet numerics and this for any reasonable drive strength that can be treated perturbatively. The inset of Fig. 3.2 c) illustrates that for a given coupler frequency, the gate rate follows the expected linear behavior for increasing drive amplitudes. Unlike for J_{ab}, the

second-order perturbation theory is not sufficient to capture the effects of the drive on the ZZ interaction, even for moderate drive strength. Although we expect higher-order perturbation theory to replicate the numerical results, we did not go deeper into the analytics due to the substantial memory Mathematica requires.

Because the qubits are far off resonant the iSWAP rate is zero when the drive is off, however, as seen in Fig. 3.2 a) the ZZ interaction is always on. As we will see in Sect. 3.1.2, if the device parameters are appropriately chosen, one can find a coupler frequency for which the ZZ interaction vanishes during idle time and even during the gate. With the set of device parameters used in Fig. 3.2, we see that the cross-Kerr interaction can be canceled by tuning the coupler at a specific point between the qubit frequencies.

3.1.2 Numerical results

In this next section, we numerically explore the toy model without doing the RWA. As before, we will focus on the interaction rate of the iSWAP gate and the ZZ cross-Kerr interaction. In particular, we study the behavior of these two quantities as a function of the device parameters: qubit/coupler frequencies $(\omega_a, \omega_b, \omega_c)$, anharmonicities $(\alpha_a, \alpha_b, \alpha_c)$ and couplings (g_a, g_b, g_c). In the upcoming, we will perform a parameter exploration starting from the following typical device parameters:

$$\begin{aligned} &\omega_a/2\pi = 4\,\text{GHz} && \alpha_a/2\pi = -0.2\,\text{GHz} && g_a/2\pi = 0.05\,\text{GHz} \\ &\omega_b/2\pi = 5.5\,\text{GHz} && \alpha_b/2\pi = -0.2\,\text{GHz} && g_b/2\pi = -0.05\,\text{GHz} \\ &\omega_c/2\pi = [2.5, 7]\,\text{GHz} && \alpha_c/2\pi = 0.1\,\text{GHz} \end{aligned} \tag{3.28}$$

by varying one or two parameters at a time.

As a first step, we focus on the static properties of the device obtained from exact diagonalization of the Hamiltonian. Fig. 3.3 shows the ZZ interaction as the coupler frequency is varied a) for multiple coupler anharmonicities, spanning from negative to positive values, and b) for multiple coupling strength. We identify three main resonances which appear at the qubit frequencies $\omega_c/2\pi = 4$ and $5.5\,\text{GHz}$, and close to $\omega_c \approx (\omega_a + \omega_b)/2 = 2\pi \times 4.75\,\text{GHz}$. The resonance appearing between the two qubit frequencies results from an avoided crossing between the second coupler level with the state $|1_a 1_b 0_c\rangle$ where $|i_a i_b i_c\rangle$ is a dressed state of the full Hamiltonian that has maximum overlap with the bare qubits and coupler Fock state $|i\rangle_a |i\rangle_b |i\rangle_c$. States near resonances are strongly hybridized, making them challenging to distinguish. As a result, we will focus our attention on the regions slightly away from these resonances.

As expected from the results of Sect. 3.1.1, we see in Fig. 3.3 a) that for negative coupler anharmonicities there is no coupler frequency at which the ZZ interaction is zero. On the other hand, positive anharmonicities can result in multiple coupler frequencies where the ZZ interaction is zero. The zero-ZZ points appear when the coupler is close to the qubit frequencies. Such operation points are generally found when the coupler is between the qubit frequencies, however, it is not excluded to find such operation points when the coupler is placed above or below the qubit frequencies.

Fig. 3.3 b) shows again χ_{ab} as a function of the coupler frequency but now for different values of the coupling g_a. Varying instead the coupling g_b leads to qualitatively the same conclusions. As expect, the stronger the coupling g_a, the larger the cross-Kerr interaction becomes in absolute value. Interestingly, the zero-ZZ point remains relatively unchanged when varying the coupling strength as long as it is smaller or comparable to the coupler anharmonicity. This is true for most sets of parameters. Otherwise, the zero-ZZ point can get shifted or can eventually disappear. Moreover, increasing the detuning between the qubits, the effective coupling between the qubits decreases, leading to a qualitatively similar behavior as observed in Fig. 3.3 b).

As seen in Fig. 3.3, for realistic device parameters, the toy model typically leads to cross-Kerr interactions ranging from a few to tens of kHz. For transmon qubits, this is about an order of magnitude smaller than the typical dephasing time. In other words, any ZZ cross-Kerr interaction below 10 kHz would not be experimentally measurable. That being said, we will see below that using a model of the device based on its full Hamiltonian, the ZZ interaction is, in fact, on the order of a few MHz. It is therefore crucial to find operation points where this interaction can be turned off.

We now turn our attention to the driven case. In this context, we numerically obtain the results using the Floquet method described in Chapter 2. In Fig. 3.4 a) we plot in solid (dashed) lines the dynamical (static) ZZ cross-Kerr interaction as a function of the coupler frequency, for a drive amplitude $\delta/2\pi = 0.1\,\mathrm{GHz}$ ($\delta/2\pi = 0\,\mathrm{GHz}$). In Fig. 3.4 b) we plot the iSWAP gate rate J_{ab} as a function of the coupler frequency for a drive amplitude $\delta/2\pi = 0.1\,\mathrm{GHz}$. Each colored line in Fig. 3.4 corresponds to a different coupler anharmonicity. A first observation that can be made is that the gate rate J_{ab}, remains relatively unchanged when the coupler anharmonicity is varied. This is to be expected from Eq. (3.19) since the hybridization coefficients only depend on the qubits and coupler frequencies and the couplings. The gate rate is roughly an order of magnitude larger than the ZZ interaction. The iSWAP rate is largest close to the qubit-coupler resonances, which suggests that the two-qubit gate should ideally be operated when the coupler is between

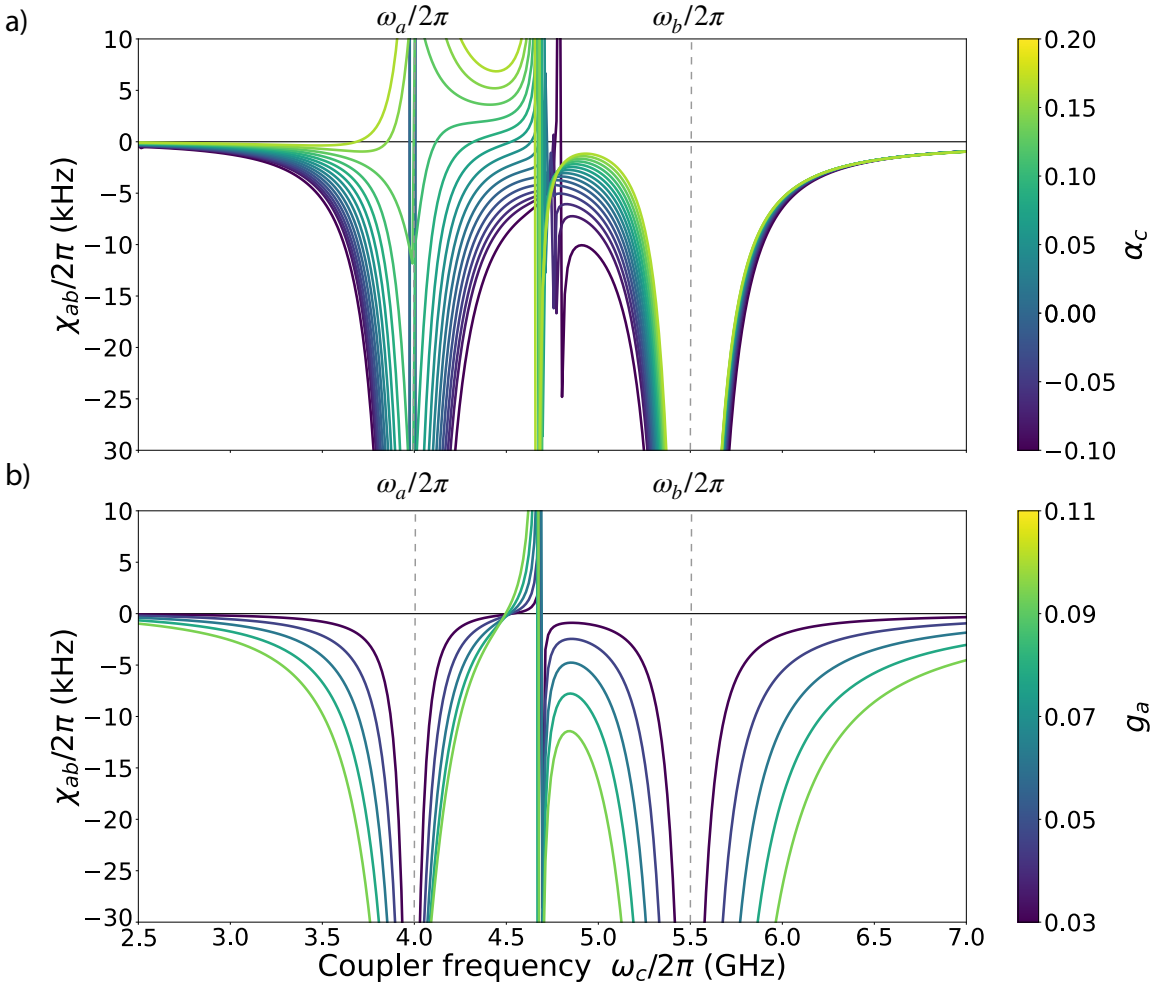


Figure 3.3 Toy model: static Cross-Kerr interaction. Cross-Kerr interaction as a function of the coupler frequency in the undriven case: a) for multiple coupler anharmonicities ranging from $\alpha_c/2\pi = [-0.1, 0.2]$ GHz, and b) for multiple coupling strength ranging from $g_c/2\pi = [0.03, 0.11]$ GHz. The other device parameters are the ones indicated in Eq. (3.28). The horizontal black line at 0 is a guide to the eye and the vertical dashed gray lines indicate the qubit frequencies.

the qubit frequencies or slightly above or below them. Additional resonances appear away from the qubit frequencies, especially at large drive amplitude. These are due to avoided crossings with higher energy levels.

Turning our attention to the cross-Kerr interaction, in Fig. 3.4 a) we compare the static ZZ in the absence of a drive (dashed lines) and the dynamical ZZ in the presence of a drive (solid lines). We note a small variation of the cross-Kerr interaction between the static and the dynamical ZZ. This variation is further accentuated as the drive amplitude increases. Ideally, we want the ZZ interaction to be zero during the gate and idle time. As an example,

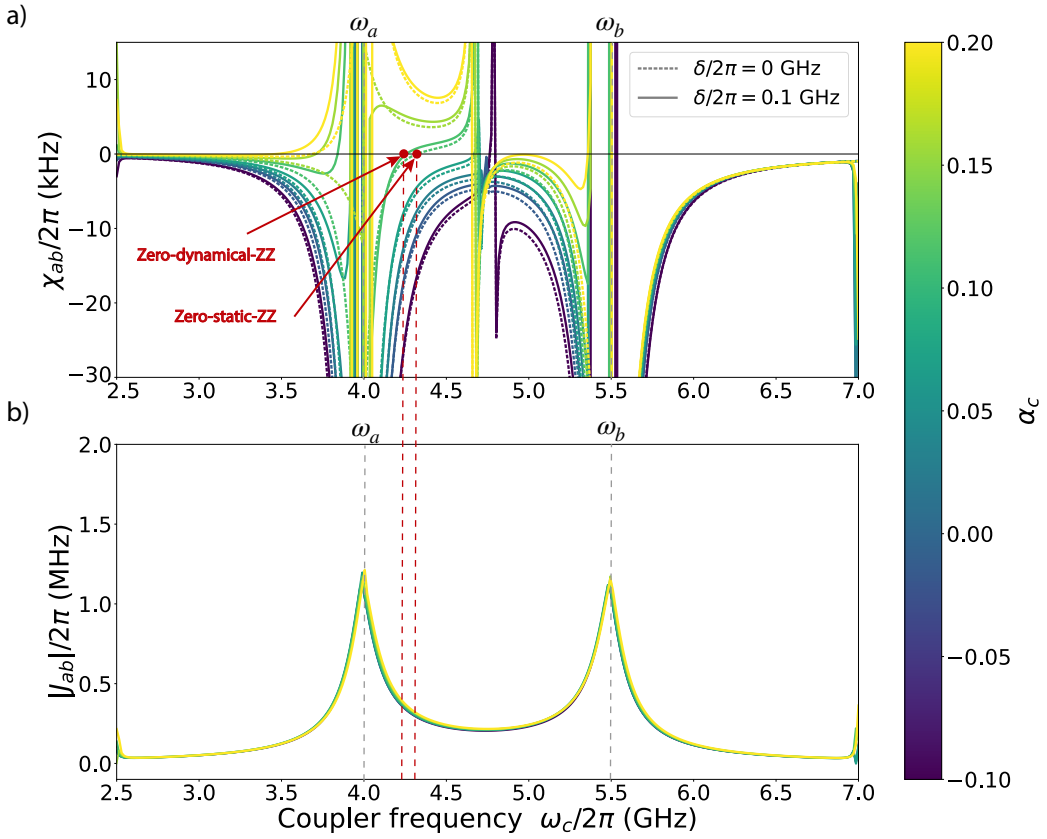


Figure 3.4 Dynamical device properties. a) Cross-Kerr interaction and, b) iSWAP interaction rate as a function of the coupler frequency for multiple coupler anharmonicities. The full lines correspond to the driven case with amplitude $\delta/2\pi = 0.1$ GHz, while the dashed lines correspond to the static case. For $\alpha_c/2\pi = 0.1\,\text{GHz}$, we indicate with red dots the points where the static and dynamical ZZ are zero for a given coupler frequency. The other device parameters are the ones indicated in Eq. (3.28). The horizontal black line at 0 is a guide to the eye and the vertical dashed gray lines indicate the qubit frequencies.

we indicate with red dots, in Fig. 3.4, the points where the static and dynamical ZZ are zero, for $\alpha_c/2\pi = 0.1\,\text{GHz}$. This observation implies that it is possible to find operation points where both the static and dynamical ZZ interactions are zero by tuning the coupler at the zero-static-ZZ point during idle time, and by shifting the coupler frequency at the zero-dynamical-ZZ point when operating the iSWAP gate. In the toy model, the dynamical ZZ can sometimes be canceled at a sweet spot with respect to the coupler frequency.

The iSWAP rate can be enhanced by increasing the drive amplitude or the qubit-coupler coupling rates. In Fig. 3.5 we show how the gate rate and the dynamical ZZ vary as the drive amplitude is increased. As expected from the results of Sect. 3.2.1 the gate rate increases

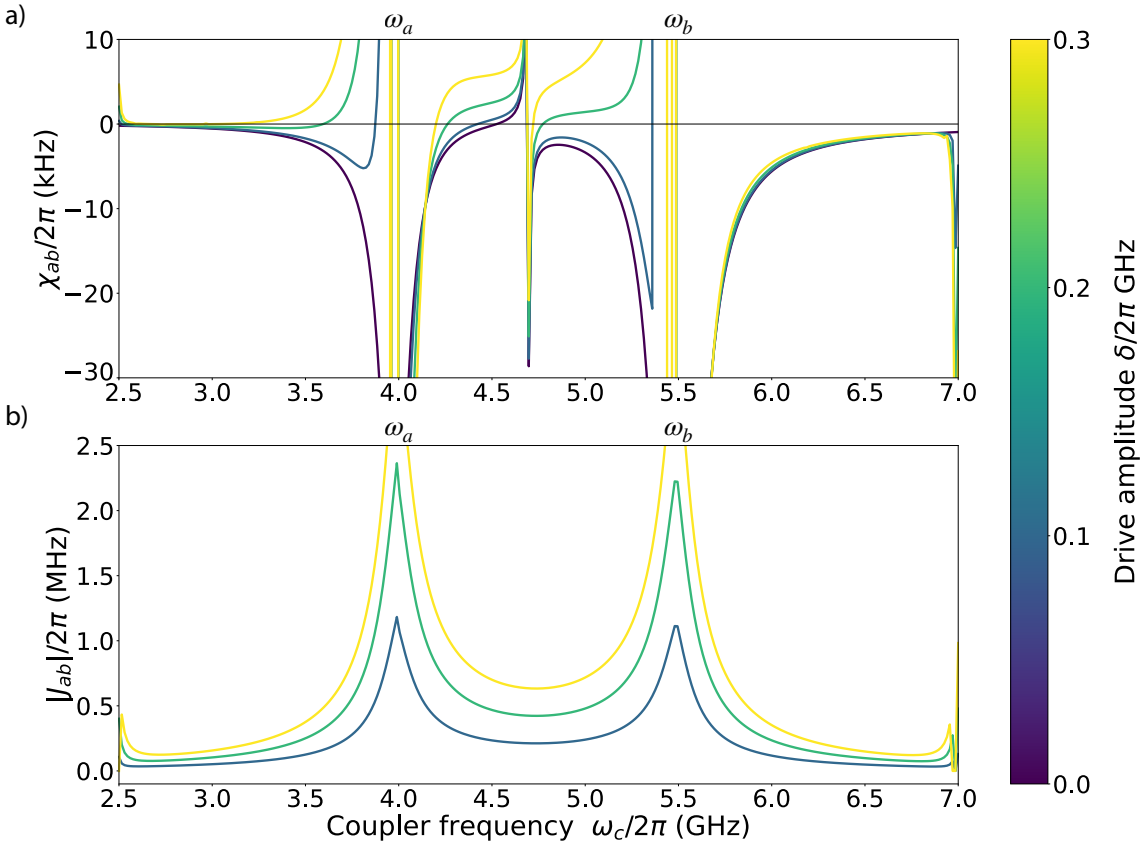


Figure 3.5 Dynamical device properties. a) Cross-Kerr interaction as a function of the coupler frequency for multiple drive amplitudes. b) iSWAP interaction rate as a function of the coupler frequency for multiple drive amplitudes. The results are plotted for $\delta/2\pi = [0, 0.1, 0.2, 0.3]$ GHz. For different drive amplitudes, we indicate with red dots the points where the dynamical ZZ are zero for a given coupler frequency. The device parameters are the ones indicated in Eq. (3.28). The horizontal black line at 0 a guide to the eye and the vertical dashed gray lines indicate the qubit frequencies.

almost linearly with δ. Introducing a drive also raises the ZZ cross-Kerr, but a zero-ZZ point can generally still be found. With the particular parameter configuration that we have chosen here, the coupler can be operated between the two qubits while keeping the ZZ cross-Kerr interaction to zero.

3.2 Full device

3.2.1 Analytics

In the previous section, to gain intuition on the workings of the coupler of Fig. 3.1, we have studied a toy model of the circuit. However, this model is significantly simplified and does not capture the quantitative features of the full device. For this reason, in this section, we derive the effective gate Hamiltonian of the full device and from it, extract the interaction terms. In contrast to the toy model, we go a step further and include the effects of the parametric drive on the bare modes. As a first step, we express the circuit Hamiltonian in terms of the creation and annihilation operators of the bare circuit modes. We then follow the approach used above for the toy model to derive the effective Hamiltonian and deduce the interaction rates. Finally, we compare our perturbation theory to the numerical analysis of the full model.

In the following, we focus on the parametric activation of an iSWAP interaction by modulating the flux through the coupler at the qubit-qubit detuning frequency.

A) Bare-mode Hamiltonian

Starting with the full-circuit Hamiltonian in Eq. (3.2), we express the bare charge number and phase operators in terms of the creation and annihilation operators:

$$\begin{aligned} \hat{\boldsymbol{\varphi}}_j &= \sqrt{\frac{\eta_j}{2}}(\hat{\boldsymbol{j}} + \hat{\boldsymbol{j}}^\dagger), \\ \hat{\boldsymbol{n}}_j &= -i\sqrt{\frac{1}{2\eta_j}}(\hat{\boldsymbol{j}} - \hat{\boldsymbol{j}}^\dagger), \end{aligned} \tag{3.29}$$

where $\hat{\boldsymbol{j}} \in \{\hat{\boldsymbol{a}}, \hat{\boldsymbol{b}}, \hat{\boldsymbol{c}}\}$ are the bare qubit or coupler modes, and η_j are chosen such that the quadratic terms, $\hat{\boldsymbol{j}}^2$ and $\hat{\boldsymbol{j}}^{\dagger 2}$, vanish in the time-averaged qubit or coupler Hamiltonian. In other words, we insert the modes of Eq. (3.29) in the qubit and coupler Hamiltonians, and after time-averaging these Hamiltonians, we eliminate the quadratic modes by appropriately choosing η_j. Mathematically speaking, this is equivalent to solving the following transcendental equation:

$$\mathcal{F}(\eta_j)\eta_j^2 = 8E_{\mathrm{C}j}/E_{\mathrm{J}j}, \tag{3.30}$$

for $j = a, b, c$, and with:

$$\begin{aligned}\mathcal{F}(\eta_{a(b)}) &\equiv e^{-\frac{\eta_{a(b)}}{4}}, \\ \mathcal{F}(\eta_c) &\equiv \alpha e^{-\frac{\eta_c}{4}} J_0(\mu_\alpha \delta\varphi) \cos(\mu_\alpha \overline{\varphi}_{\text{ext}}) \\ &\quad + \frac{\beta}{N} e^{-\frac{\eta_c}{4N^2}} J_0(\mu_\beta \delta\varphi) \cos(\mu_\beta \overline{\varphi}_{\text{ext}}),\end{aligned} \tag{3.31}$$

which are solved numerically. Note that in the transmon regime, $\mathcal{F}(\eta_{a(b)}) \sim 1$ and we recover Eq. (1.10) of the transmon qubit. Furthermore, $\mathcal{F}(\eta_c)$ depends on the drive amplitude, such that the coupler modes will capture drive effects like the ac-Stark shift already at low order in perturbation theory. Compared to the toy model, here, we go a step further and included the effects of the parametric drive on the bare modes.

Now that the bare modes are defined, we normal order expand the cosines and sines in Eq. (3.1) using:

$$\begin{aligned}\cos\hat{\boldsymbol{\varphi}}_j &= e^{-\frac{\eta_j}{4}} \sum_{\substack{m,n\geq 0 \\ m+n=\text{ even}}} \frac{\left(-\frac{\eta_j}{2}\right)^{\frac{m+n}{2}} \hat{\boldsymbol{j}}^{\dagger m} \hat{\boldsymbol{j}}^n}{m!n!}, \\ \sin\hat{\boldsymbol{\varphi}}_j &= e^{-\frac{\eta_j}{4}} \sqrt{\frac{\eta_j}{2}} \sum_{\substack{m,n\geq 0 \\ m+n=\text{ odd}}} \frac{\left(-\frac{\eta_j}{2}\right)^{\frac{m+n-1}{2}} \hat{\boldsymbol{j}}^{\dagger m} \hat{\boldsymbol{j}}^n}{m!n!}.\end{aligned} \tag{3.32}$$

Note that the sines appear in the coupler Hamiltonian when expanding the cosine potentials of sums. Furthermore, for the coupler Hamiltonian, we expand the trigonometric functions of φ_{ext} in Jacobi-Anger series over the harmonics of the frequency of the drive. The specific choice of η_j we made finally allows us to rewrite the transmon Hamiltonians up to fourth order as:

$$\begin{aligned}\hat{H}_a &= \boldsymbol{\omega}_a \hat{\boldsymbol{a}}^\dagger \hat{\boldsymbol{a}} + \frac{\boldsymbol{\alpha}_a}{2} \hat{\boldsymbol{a}}^{\dagger 2} \hat{\boldsymbol{a}}^2 \\ &\quad + \frac{\boldsymbol{\alpha}_a}{12} \left(\hat{\boldsymbol{a}}^4 + \hat{\boldsymbol{a}}^{\dagger 4}\right) + \frac{\boldsymbol{\alpha}_a}{3} \left(\hat{\boldsymbol{a}}^\dagger \hat{\boldsymbol{a}}^3 + \hat{\boldsymbol{a}}^{\dagger 3} \hat{\boldsymbol{a}}\right) + \cdots\end{aligned} \tag{3.33}$$

with analogous equations for modes $\hat{\boldsymbol{b}}$, and where the mode frequency and anharmonicity are defined as:

$$\begin{aligned}\boldsymbol{\omega}_{a(b)} &= \frac{4E_{\text{C}a(b)}}{\eta_{a(b)}} + \frac{1}{2}\mathcal{F}(\eta_{a(b)})\eta_{a(b)}E_{\text{J}a(b)}, \\ \boldsymbol{\alpha}_{a(b)} &= -E_{\text{C}a(b)}.\end{aligned} \tag{3.34}$$

In the first line of Eq. (3.33), we recover the KNO used for the toy model, and the second

line contains the contribution of quartic counter-rotating terms, which are number-non conserving.

In contrast to the qubit Hamiltonians, the coupler is time-dependent, we thus choose η_c such that the quadratic terms $\hat{\boldsymbol{c}}^2$ and $\hat{\boldsymbol{c}}^{\dagger 2}$, vanish in the time-averaged Hamiltonian. By "time-averaged" Hamiltonian, we mean $\overline{\hat{H}}_c = \lim_{T\to\infty} \frac{1}{T}\int_0^T dt \hat{H}_c(t)$. Now, expressing the coupler Hamiltonian as $\hat{H}_c(t) = \overline{\hat{H}}_c(t) + \widetilde{\hat{H}}_c(t)$, in terms of the creation and annihilation operators we obtain:

$$\begin{aligned}\overline{\hat{H}}_c =& \boldsymbol{\omega}_c \hat{\boldsymbol{c}}^\dagger \hat{\boldsymbol{c}} + \frac{\boldsymbol{\alpha}_c}{2}\hat{\boldsymbol{c}}^{\dagger 2}\hat{\boldsymbol{c}}^2 \\ &+ \frac{\boldsymbol{\alpha}_c}{12}\left(\hat{\boldsymbol{c}}^4 + \hat{\boldsymbol{c}}^{\dagger 4}\right) + \frac{\boldsymbol{\alpha}_c}{3}\left(\hat{\boldsymbol{c}}^\dagger \hat{\boldsymbol{c}}^3 + \hat{\boldsymbol{c}}^{\dagger 3}\hat{\boldsymbol{c}}\right) \\ &+ \boldsymbol{g}_{c,3}\left(\hat{\boldsymbol{c}}^3 + \hat{\boldsymbol{c}}^{\dagger 3} + 3\hat{\boldsymbol{c}}^\dagger \hat{\boldsymbol{c}}^2 + 3\hat{\boldsymbol{c}}^{\dagger 2}\hat{\boldsymbol{c}}\right) \\ &+ \boldsymbol{g}_{c,1}\left(\hat{\boldsymbol{c}} + \hat{\boldsymbol{c}}^\dagger\right) + \cdots\end{aligned} \tag{3.35}$$

where the coupler mode frequency and anharmonicity are defined as in Eq. (3.34) by replacing $a \to c$, and the cubic and monomial counter-rotating terms:

$$\begin{aligned}\boldsymbol{g}_{c,3} =& -\alpha\epsilon e^{-\frac{\eta_c}{4}}\eta_c^{3/2} J_0(\mu_\alpha \delta\varphi)\sin(\mu_\alpha \overline{\varphi}_{\text{ext}})\, E_{\text{Jc}}/(12\sqrt{2}) \\ &- \frac{\beta}{N^2} e^{-\frac{\eta_c}{4N^2}}\eta_c^{3/2} J_0(\mu_\beta \delta\varphi)\sin(\mu_\beta \overline{\varphi}_{\text{ext}})\, E_{\text{Jc}}/(12\sqrt{2}), \\ \boldsymbol{g}_{c,1} =& \alpha\epsilon e^{-\frac{\eta_c}{4}}\eta_c^{1/2} J_0(\mu_\alpha \delta\varphi)\sin(\mu_\alpha \overline{\varphi}_{\text{ext}})\, E_{\text{Jc}}/\sqrt{2} \\ &+ \beta e^{-\frac{\eta_c}{4N^2}}\eta_c^{1/2} J_0(\mu_\beta \delta\varphi)\sin(\mu_\beta \overline{\varphi}_{\text{ext}})\, E_{\text{Jc}}/\sqrt{2}.\end{aligned} \tag{3.36}$$

The time-dependent part of the Hamiltonian $\widetilde{\hat{H}}_c(t)$, contains many terms, making it impractical to list them all here. These expressions can be found in appendix D of Ref. [48] which we included in Appendix C of this thesis. A major difference between Eq. (3.33) and Eq. (3.35) is the presence of monomials ($\hat{\boldsymbol{c}}$ and $\hat{\boldsymbol{c}}^\dagger$) and cubic terms, which appear because the coupler Hamiltonian breaks the parity symmetry due to the presence of an external flux.

Now that the bare modes have been defined, the coupling Hamiltonian in Eq. (3.3) can be rewritten as:

$$\hat{H}_g = -\frac{2E_{Cab}}{\sqrt{\eta_a\eta_b}}(\hat{\boldsymbol{a}} - \hat{\boldsymbol{a}}^\dagger)(\hat{\boldsymbol{b}} - \hat{\boldsymbol{b}}^\dagger) - \frac{2E_{Cac}}{\sqrt{\eta_a\eta_c}}(\hat{\boldsymbol{a}} - \hat{\boldsymbol{a}}^\dagger)(\hat{\boldsymbol{c}} - \hat{\boldsymbol{c}}^\dagger) - \frac{2E_{Cbc}}{\sqrt{\eta_b\eta_c}}(\hat{\boldsymbol{b}} - \hat{\boldsymbol{b}}^\dagger)(\hat{\boldsymbol{c}} - \hat{\boldsymbol{c}}^\dagger). \tag{3.37}$$

Notice that due to the presence of η_c, the couplings account for the effects of the drive.

Finally, the full-circuit Hamiltonian that we numerically simulate to compare the an-

alytical results obtained in the following subsections is the sum of Eq. (3.33) for the two qubits, Eq. (3.35), Eq. (3.37) and the time-dependent part of the coupler Hamiltonian found appendix D of Ref. [48] which can be found in Appendix C of this thesis.

B) Displace the Hamiltonian

We now want to derive the effective gate Hamiltonian from the circuit Hamiltonian obtained above. Before performing the time-dependent SW transformation to derive the effective model, we displace the Hamiltonian in Eq. (3.1) to remove the contributions that are linear in $\hat{\boldsymbol{\varphi}}_j$ and $\hat{\boldsymbol{n}}_j$. Intuitively, this procedure corresponds to performing the Taylor expansion around the classical minimum of the potential energy. We consider a unitary displacement of Eq. (3.1) according to which $\hat{\boldsymbol{\varphi}}_j \to \hat{\boldsymbol{\varphi}}_j + \varphi_j(t)$ and $\hat{\boldsymbol{n}}_j \to \hat{\boldsymbol{n}}_j + n_j(t)$, with $j = a, b, c$. The linear terms of the Hamiltonian can be eliminated by choosing $\varphi_j(t)$ and $n_j(t)$ such that they obey the classical equations of motion associated with the full circuit Hamiltonian of Eq. (3.1). Writing down the Heisenberg equations of motion,

$$\frac{d\hat{\boldsymbol{\varphi}}_j(t)}{dt} = i[\hat{H}, \hat{\boldsymbol{\varphi}}_j] \quad \text{and} \quad \frac{d\hat{\boldsymbol{n}}_j(t)}{dt} = i[\hat{H}, \hat{\boldsymbol{n}}_j] \qquad \text{for } j = a, b, c \tag{3.38}$$

and then changing the Heisenberg-picture operators to classical variables $\hat{\boldsymbol{\varphi}}_j(t) \to \varphi_j(t)$ and $\hat{\boldsymbol{n}}_j(t) \to n_j(t)$, we find the displacements by solving the following differential equations:

$$\begin{aligned}
\dot{\varphi}_a &= 8E_{\mathrm{C}a} n_a + 4E_{\mathrm{C}ac} n_c + 4E_{\mathrm{C}ab} n_b, & \dot{n}_a &= -E_{\mathrm{J}a} \sin(\varphi_a), \\
\dot{\varphi}_b &= 4E_{\mathrm{C}ab} n_a + 8E_{\mathrm{C}b} n_b + 4E_{\mathrm{C}bc} n_c, & \dot{n}_b &= -E_{\mathrm{J}b} \sin(\varphi_b), \\
\dot{\varphi}_c &= 4E_{\mathrm{C}ac} n_a + 4E_{\mathrm{C}bc} n_b + 8E_{\mathrm{C}c} n_c, & \dot{n}_c &= -\alpha E_{\mathrm{J}c} \sin\left[\varphi_c + \mu_\alpha \varphi_{\mathrm{ext}}(t)\right] \\
& & & \quad - \beta E_{\mathrm{J}c} \sin\left[\frac{\varphi_c}{N} + \mu_\beta \varphi_{\mathrm{ext}}(t)\right].
\end{aligned} \tag{3.39}$$

These equations can be solved approximately by considering a trial form, $\varphi_j = \xi_j + \zeta_j \sin(\omega_d t)$, and equating the coefficients of the zeroth and first harmonics of the drive frequency ω_d. More details are given in appendix D of the paper Ref. [48] which can be found in Appendix C. Ultimately, the displaced variables can be seen as the fluctuations around a classical trajectory. Note that we expect the displacements for the transmons to be much smaller than for the coupler ($\zeta_{a,b} \ll \zeta_c$ and $\xi_{a,b} \ll \xi_c$).

C) Driven black-box quantization approach

Just as for the toy model, we define the normal modes by partitioning the Hamiltonian into a perturbed and unperturbed part, $\hat{H} = \hat{H}^{(0)} + \lambda\hat{H}^{(1)}$, where we define:

$$\begin{aligned}\hat{H}^{(0)} &= \boldsymbol{\omega}_a\hat{\boldsymbol{a}}^\dagger\hat{\boldsymbol{a}} + \boldsymbol{\omega}_b\hat{\boldsymbol{b}}^\dagger\hat{\boldsymbol{b}} + \boldsymbol{\omega}_c\hat{\boldsymbol{c}}^\dagger\hat{\boldsymbol{c}} + \hat{H}_g \\ &\equiv \omega_a\hat{a}^\dagger\hat{a} + \omega_b\hat{b}^\dagger\hat{b} + \omega_c\hat{c}^\dagger\hat{c}.\end{aligned} \tag{3.40}$$

The second line is written in terms of the dressed modes and frequencies (not in bold). The bare and normal modes are related to each other by:

$$\hat{\boldsymbol{\varphi}}_\alpha = \sum_{\beta=a,b,c} \frac{u_{\alpha\beta}}{\sqrt{2}}(\hat{\beta} + \hat{\beta}^\dagger) \quad , \quad \hat{\boldsymbol{n}}_\alpha = \sum_{\beta=a,b,c} \frac{v_{\alpha\beta}}{i\sqrt{2}}(\hat{\beta} - \hat{\beta}^\dagger), \tag{3.41}$$

where the hybridization coefficients $u_{\alpha\beta}$ and $v_{\alpha\beta}$ are defined in appendix E of Ref. [48] which can be found in Appendix C. The perturbed term $\lambda\hat{H}^{(1)}$ is made of the remaining terms.

D) Effective Hamiltonian

Following the same method as for the toy model, we move to the interaction picture defined by the Hamiltonian $\hat{H}^{(0)}$ of Eq. (3.40). For a modulation frequency $\omega_d = \omega_a - \omega_b$, we obtain the effective Hamiltonian of the full device after a time-dependent Schrieffer-Wolff transformation at first order in the RWA, the effective Hamiltonian of the full device:

$$\hat{H}_{\text{eff}} = \lambda\overline{\hat{H}}_I^{(1)} = J_{ab}^{(1)}(-i\hat{a}^\dagger\hat{b} + \text{H.c.}) \tag{3.42}$$

$$+ \frac{\alpha_a^{(1)}}{2}\hat{a}^{\dagger 2}\hat{a}^2 + \frac{\alpha_b^{(1)}}{2}\hat{b}^{\dagger 2}\hat{b}^2 + \frac{\alpha_c^{(1)}}{2}\hat{c}^{\dagger 2}\hat{c}^2 \tag{3.43}$$

$$+ \chi_{ab}^{(1)}\hat{a}^\dagger\hat{a}\hat{b}^\dagger\hat{b} + \chi_{bc}^{(1)}\hat{b}^\dagger\hat{b}\hat{c}^\dagger\hat{c} + \chi_{ca}^{(1)}\hat{c}^\dagger\hat{c}\hat{a}^\dagger\hat{a} \tag{3.44}$$

$$+ J_{ab;a}^{(1)}(-i\hat{a}^\dagger\hat{a}\hat{a}^\dagger\hat{b} + \text{H.c.}) \tag{3.45}$$

$$+ J_{ab;b}^{(1)}(-i\hat{b}^\dagger\hat{b}\hat{a}^\dagger\hat{b} + \text{H.c.}) \tag{3.46}$$

$$+ J_{ab;c}^{(1)}(-i\hat{c}^\dagger\hat{c}\hat{a}^\dagger\hat{b} + \text{H.c.}) \tag{3.47}$$

$$+ K_{ab}^{(1)}(\hat{a}^{\dagger 2}\hat{b}^2 + \text{H.c.}). \tag{3.48}$$

The interaction terms that appear in the above Hamiltonian are:

$$
\begin{aligned}
J_{ab}^{(1)} &= -\frac{u_{aa}u_{ab}}{2}J_1(\zeta_a)\sin(\xi_a)E_{\mathrm{J}a} - \frac{u_{ba}u_{bb}}{2}J_1(\zeta_b)\sin(\xi_b)E_{\mathrm{J}b} \\
&\quad -\frac{u_{ca}u_{cb}}{2}\alpha J_1(\zeta_c+\mu_\alpha\delta\varphi)\sin(\xi_c+\mu_\alpha\overline{\varphi}_{\mathrm{ext}})E_{\mathrm{J}c}^{(\alpha)} \\
&\quad -\frac{u_{ca}u_{cb}}{2}\frac{\beta}{N}J_1\left(\frac{\zeta_c}{N}+\mu_\beta\delta\varphi\right)\sin\left(\frac{\xi_c}{N}+\mu_\beta\overline{\varphi}_{\mathrm{ext}}\right)E_{\mathrm{J}c}^{(\beta)},
\end{aligned} \tag{3.49}
$$

$$
\alpha_j^{(1)} = -\frac{1}{8}\sum_{i=a,b,c} u_{ij}^4 E'_{\mathrm{J},i}, \tag{3.50}
$$

$$
\chi_{jk}^{(1)} = -\frac{1}{4}\sum_{i=a,b,c} u_{ij}^2 u_{ik}^2 E'_{\mathrm{J},i}, \tag{3.51}
$$

$$
J_{ab;j}^{(1)} = -\frac{u_{cj}^2}{4}J_{ab}^{(1)}, \tag{3.52}
$$

$$
\begin{aligned}
K_{ab}^{(1)} &= -\frac{u_{aa}^2u_{ab}^2}{16}J_2(\zeta_a)\cos(\xi_a)E_{\mathrm{J}a} - \frac{u_{ba}^2u_{bb}^2}{16}J_2(\zeta_b)\cos(\xi_b)E_{\mathrm{J}b} \\
&\quad -\frac{u_{ca}^2u_{cb}^2}{16}\alpha J_2(\zeta_c+\mu_\alpha\delta\varphi)\cos(\xi_c+\mu_\alpha\overline{\varphi}_{\mathrm{ext}})E_{\mathrm{J}c}^{(\alpha)} \\
&\quad -\frac{u_{ca}^2u_{cb}^2}{16}\frac{\beta}{N^3}J_2\left(\frac{\zeta_c}{N}+\mu_\beta\delta\varphi\right)\cos\left(\frac{\xi_c}{N}+\mu_\beta\overline{\varphi}_{\mathrm{ext}}\right)E_{\mathrm{J}c}^{(\beta)},
\end{aligned} \tag{3.53}
$$

for $j,k\in\{a,b,c\}$, and with the effective Josephson energies:

$$
\begin{aligned}
E'_{\mathrm{J}a} &\equiv e^{-\frac{u_{aa}^2}{4}-\frac{u_{ab}^2}{4}-\frac{u_{ac}^2}{4}}J_0(\zeta_a)\cos(\xi_a)E_{\mathrm{J}a}, \\
E'_{\mathrm{J}b} &\equiv e^{-\frac{u_{ba}^2}{4}-\frac{u_{bb}^2}{4}-\frac{u_{bc}^2}{4}}J_0(\zeta_b)\cos(\xi_b)E_{\mathrm{J}b}, \\
E'_{\mathrm{J}c} &\equiv \alpha J_0(\zeta_c+\mu_\alpha\delta\varphi)\cos(\xi_c+\mu_\alpha\overline{\varphi}_{\mathrm{ext}})E_{\mathrm{J}c}^{(\alpha)} \\
&\quad + \frac{\beta}{N^3}J_0\left(\frac{\zeta_c}{N}+\mu_\beta\delta\varphi\right)\cos\left(\frac{\xi_c}{N}+\mu_\beta\overline{\varphi}_{\mathrm{ext}}\right)E_{\mathrm{J}c}^{(\beta)}, \\
E_{\mathrm{J}c}^{(\alpha)} &\equiv e^{-\frac{u_{ca}^2}{4}-\frac{u_{cb}^2}{4}-\frac{u_{cc}^2}{4}}E_{\mathrm{J}c}, \\
E_{\mathrm{J}c}^{(\beta)} &\equiv e^{-\frac{u_{ca}^2}{4N^2}-\frac{u_{cb}^2}{4N^2}-\frac{u_{cc}^2}{4N^2}}E_{\mathrm{J}c}.
\end{aligned} \tag{3.54}
$$

We notice that Eq. (3.42)–(3.44) also appear in the effective model of the 3KNO. The full model, however, presents additional interactions corresponding to photon-number-conditioned iSWAP terms, Eq. (3.45)–(3.47), and a coherent two-photon exchange Eq. (3.48). In Eq. (3.48), the occurrence of second order Bessel functions indicates that $K_{ab}^{(1)}$ is generated by the second harmonic of the drive. Furthermore, considering that the terms are also fourth order in the hybridization coefficients, the contribution from the two-photon exchange term is minor in comparison to the resonant interaction.

All the interaction terms depend on the drive, either directly with the presence of the Bessel function of $\delta\varphi$ and ζ_j, or indirectly through the hybridization coefficients. It is important to remember that in contrast to the toy model, the drive amplitude now appears in the hybridization coefficients, as we take into account the effect of the drive when rewriting the Hamiltonian in terms of the dressed modes. Moreover, the drive frequency implicitly depends on ω_d since the dressed qubit frequencies ω_a, ω_b do. In Chapter 2, we introduced a numerical method to find the adjusted parametric drive frequency.

As for the toy model, we performed the perturbation theory up to second order. Because of the complexity of the full-device model, we do not improve the starting point of the perturbation theory to better partition the (un)perturbed Hamiltonians, as we did for the toy model. Nevertheless, we further corrected the effective Hamiltonian to second order in perturbation theory, using Eq. (3.17). Because the parity-breaking terms in $\hat{\varphi}^3$ significantly dress the coupler frequency, the qubit-coupler resonances are shifted. We chose to take into account the corrections to the coupler poles at second order in perturbation theory, by reparametrizing the external flux as $\overline{\varphi}_{\text{ext}} \to \overline{\varphi}'_{\text{ext}}(\overline{\varphi}_{\text{ext}})$ such that $\omega_c(\overline{\varphi}'_{\text{ext}}) = \omega_c^{(2)}(\overline{\varphi}_{\text{ext}})$. In other terms, we redefine the dressed coupler modes in a self-consistent way to better capture the qubit-coupler resonances. We will see below that these corrections are validated by the numerics.

E) Comparison to exact numerics

We compare the analytical results obtained from the second-order perturbation theory, with the exact Floquet numerics described in Chapter 2. The Hamiltonian used for the numerics is the one derived at the end of Sect. 3.2 B). In Fig. 3.6, we study the behavior of the iSWAP rate (J_{ab}) and the ZZ cross-Kerr interaction (χ_{ab}) versus the dc flux $\overline{\varphi}_{\text{ext}}$, for a realistic set of parameters defined in the figure caption. For J_{ab}, we observe excellent agreement between the numerical results and the second-order perturbative results. Remarkably, the agreement persists even at large drive amplitudes. The two main resonances located at $\overline{\varphi}_{\text{ext}}/2\pi = 0.29$ and 0.37 correspond to the qubit-coupler resonances. Note that the coupler frequency lies above (below) the qubits frequencies for $\overline{\varphi}_{\text{ext}}/2\pi < 0.29$ (> 0.37). The additional poles appearing when $\omega_c > \omega_{a/b}$ and $\omega_c < \omega_{a/b}$ are expected to be captured at higher order within the perturbation theory. For the given device parameters and a relatively strong drive, the gate rate can reach approximately $J_{ab} \sim 20$ MHz, corresponding to a ~ 25 ns $\sqrt{i\text{SWAP}}$ gate which is relatively fast for a two-qubit gate. Shifting the focus to the ZZ interaction, we see that the analytical predictions only match well the numerical results in the absence of drive. Similarly to the toy model, the effect of the dynamical ZZ is expected to emerge at fourth order in perturbation theory. Within this parameter configuration, the

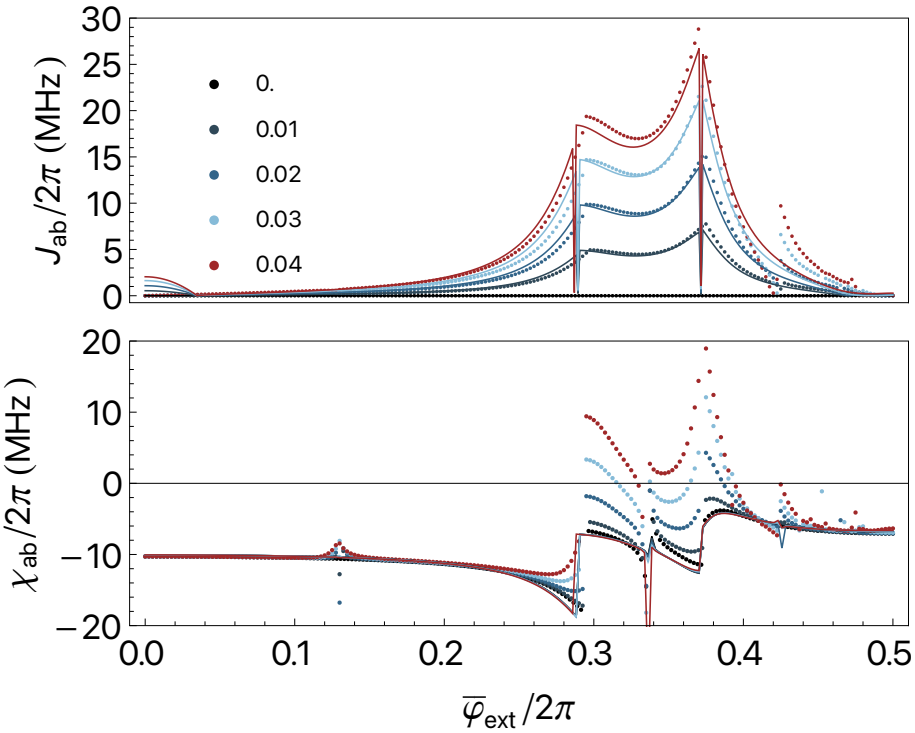


Figure 3.6 Interaction rates of the full circuit as a function of the external dc flux $\overline{\varphi}_{\text{ext}}$. The dots represent exact numerical simulations obtained from the Floquet method described in Chapter 2, with Hilbert space dimension 10 per mode. The lines correspond to the analytical second-order RWA calculations. Color (see legend) encodes parametric drive amplitude $\delta\varphi/2\pi$. Parameter choices: $C_a = 134.205$ fF, $C_b = 134.218$ fF, $C_c = 75.987$ fF, $C_{ac} = 11.11$ fF, $C_{bc} = 11.22$ fF, $C_{ab} = 0$, $E_{\text{J}a}/2\pi = 37$ GHz, $E_{\text{J}b}/2\pi = 27$ GHz, $E_{\text{J}c}/2\pi = 50$ GHz, $\alpha = 0.258$, $\beta = 1$, and $N = 3$, $\mu_\alpha = 5/6$ and $\mu_\beta = -1/18$. We attribute large discontinuities in the numerical curves to state tracking errors near avoided crossings. From Ref. [48].

ZZ cross-Kerr interaction does not exhibit a dc flux point for which $\chi_{ab} = 0$ in the static case. However, multiple zero ZZ points emerge with sufficiently large drive amplitudes.

As mentioned, the Hamiltonian used for the numerical results is the one derived at the end of Sect. 3.2 B). Note, however, that the spectrum of this latter Hamiltonian differs from that of the full cosine Hamiltonian in Eq. (3.1) by tens to hundreds of MHz. This discrepancy specially affects the cross-Kerr interaction by a few to tens of MHz. In Sect. 3.2.2 and Chapter 4, we will consequently use the full Hamiltonian in Eq. (3.1) to study more accurately the device properties and to compare our model to the experimental data.

F) Other parametric gates

According to the frequency of the parametric flux modulation, multiple interactions can

Gate	Bosonic operator	Drive frequency	Dominant unwanted interaction
iSWAP/beam-splitter	$-i\hat{a}^\dagger\hat{b} + i\hat{b}^\dagger\hat{a}$	$\omega_a - \omega_b$	$\hat{a}^\dagger\hat{a}\hat{b}^\dagger\hat{b}$
Two-mode squeezing	$-i\hat{a}^\dagger\hat{b}^\dagger + i\hat{b}\hat{a}$	$\omega_a + \omega_b$	$\hat{a}^\dagger\hat{a}\hat{b}^\dagger\hat{b}$
CZ/Ising-ZZ	$\hat{a}^\dagger\hat{a}\hat{b}^\dagger\hat{b}$	no drive	
CNOT	$-i(\hat{a} - \hat{a}^\dagger)\hat{b}^\dagger\hat{b}$	ω_a	$-i(\hat{a} - \hat{a}^\dagger)\hat{a}^\dagger\hat{a}$

Table 3.1 List of the most accessible gate Hamiltonians realizable with a parametric drive in the analyzed architecture. Adapted from Ref. [48].

be activated. This device is not limited to the iSWAP gate. For instance, by modulating the coupler flux at $\omega_d = \omega_a + \omega_b$ the dominant activated interaction is a two-mode squeezing term $\hat{a}^\dagger\hat{b}^\dagger + \hat{a}\hat{b}$. For this modulation, the derivation of the effective Hamiltonian with the method previously described, leads to the same Hamiltonian as the one in Eqs. (3.42)–(3.48), with the difference that $\hat{a}^\dagger\hat{b}$ is replaced by $\hat{a}^\dagger\hat{b}^\dagger$, in Eq. (3.42) and Eqs. (3.45)–(3.48). The interaction constants remain identical to Eqs. (3.49)–(3.54), with the difference that the values of the hybridization coefficients differ due to their dependence on the modulation frequency.

A CNOT gate can also be parametrically activated by modulating the flux at $\omega_d = \omega_a$, where qubit b is the control qubit and qubit a is the target. Applying the above method with this modulation choice of frequency, we find the following effective Hamiltonian:

$$\begin{aligned}\lambda\overline{\hat{H}}_I^{(1)} = &-i\Omega_{a;b}(\hat{a} - \hat{a}^\dagger)\hat{b}^\dagger\hat{b} - i\Omega_{a;c}(\hat{a} - \hat{a}^\dagger)\hat{c}^\dagger\hat{c} \\ &- i\Omega_a(\hat{a} - \hat{a}^\dagger) - i\Omega_{a;a}(\hat{a}^\dagger\hat{a}\hat{a} - \hat{a}^\dagger\hat{a}^\dagger\hat{a}),\end{aligned} \tag{3.55}$$

where the rates $\Omega_{i;j}$ depend on the device parameters. The first term of the first line of Eq. (3.55) corresponds to the cross-resonance interaction used to perform the CNOT gate. The second term is conditioned on the state of the coupler and it is generally neglected because the coupler is in the ground state with $\langle\hat{c}^\dagger\hat{c}\rangle \approx 0$. The other terms of the second line contain single qubit operations which can typically be corrected for by using additional single qubit rotations.

Table 3.1 summarizes some of the gates which can be implemented in this architecture. For instance, even without flux modulation, but simply by tuning the dc flux within a region where the ZZ cross-Kerr interaction is large, it is possible to generate a CZ gate. Gates not listed in this table include the possibility of activating more advanced native gates using multiple drive tones.

3.2.2 Numerical results

In Sect. 3.2.1 we have seen that it becomes difficult to gain an intuition about the dependence of the gate rate and the cross-Kerr interaction on the parameters of the device. Moreover in order to be in quantitative agreement with experiments, performing numerical simulations of the full device Hamiltonian in Eq. (3.1) becomes necessary.

As a first step, we study the static (undriven) device properties as a function of its parameters. We focus on a set of qubit and coupling parameters, and vary the coupler parameters α, $E_{\mathrm{C}c}$, $E_{\mathrm{J}c}$, N around the following typical device parameters:

$$\begin{aligned} E_{\mathrm{C}a}/2\pi &= 0.2\,\mathrm{GHz} & E_{\mathrm{J}a}/2\pi &= 18\,\mathrm{GHz} & & \\ E_{\mathrm{C}b}/2\pi &= 0.2\,\mathrm{GHz} & E_{\mathrm{J}b}/2\pi &= 34\,\mathrm{GHz} & & \\ E_{\mathrm{C}ac}/2\pi &= -0.2\,\mathrm{GHz} & E_{\mathrm{C}bc}/2\pi &= 0.2\,\mathrm{GHz} & E_{\mathrm{C}ab}/2\pi &= 0. \end{aligned} \tag{3.56}$$

In Fig. 3.7 a), we plot the smallest value of the absolute value of the ZZ interaction that is obtained in the range $\overline{\varphi}_{\mathrm{ext}}/2\pi = [0, 0.5]$ as a function of $E_{\mathrm{J}c}$ (x-axis) and $E_{\mathrm{C}c}$ (y-axis). Every plot corresponds to a different number of Josephson junctions in the array $N = 2, 3, 4$ and 5 (left to right columns), and a ratio between the Josephson energies of the branches of the coupler loop $\alpha = 1/2N$ (first line), $1/(N+1)$ (third line) and the intermediate value $\frac{1}{2}(\frac{1}{2N} + \frac{1}{N+1})$ (second line). Note that we want $\alpha < 1/N$ such that the coupler potential exhibits a single well. All the plots use the same color scale, with values ranging from 0 to 1 MHz, and any point in the maps that has a minimum ZZ above 1 MHz is plotted in white.

A first observation is that for any choice of N between 2 and 5, it is possible to find multiple coupler parameters that allow a zero-ZZ point. In fact, in almost all the cases considered here, there exist large regions in parameter space where ZZ can be canceled. However, even though multiple choices of N feature zero-ZZ regions, it is easier to find the latter for larger values of N. In all the plots, we also identify regions where ZZ is non-zero due to an avoided crossing of higher energy states. We will get back to these features below.

These large zero-ZZ regions are particularly promising as it indicates that a deviation in the coupler parameters, which are typically $\sim 1 - 10\%$ for $E_{\mathrm{C}c}$ and $\sim 5 - 10\%$ for $E_{\mathrm{J}c}$, will not prevent the device to have a zero-ZZ point, assuming that the qubit parameters are fixed. We have further studied the impact of deviations of all the device parameters on the existence of a zero-ZZ point. We numerically explored 10^6 devices with $N = 4$, assuming a Gaussian distribution of all the parameters around the mean values of Eq. (3.56), $E_{Jc}/2\pi = 90\,\mathrm{GHz}$ and $E_{Js}/2\pi = 15\,\mathrm{GHz}$ for the black sheep junction with a standard deviation of 10%. We

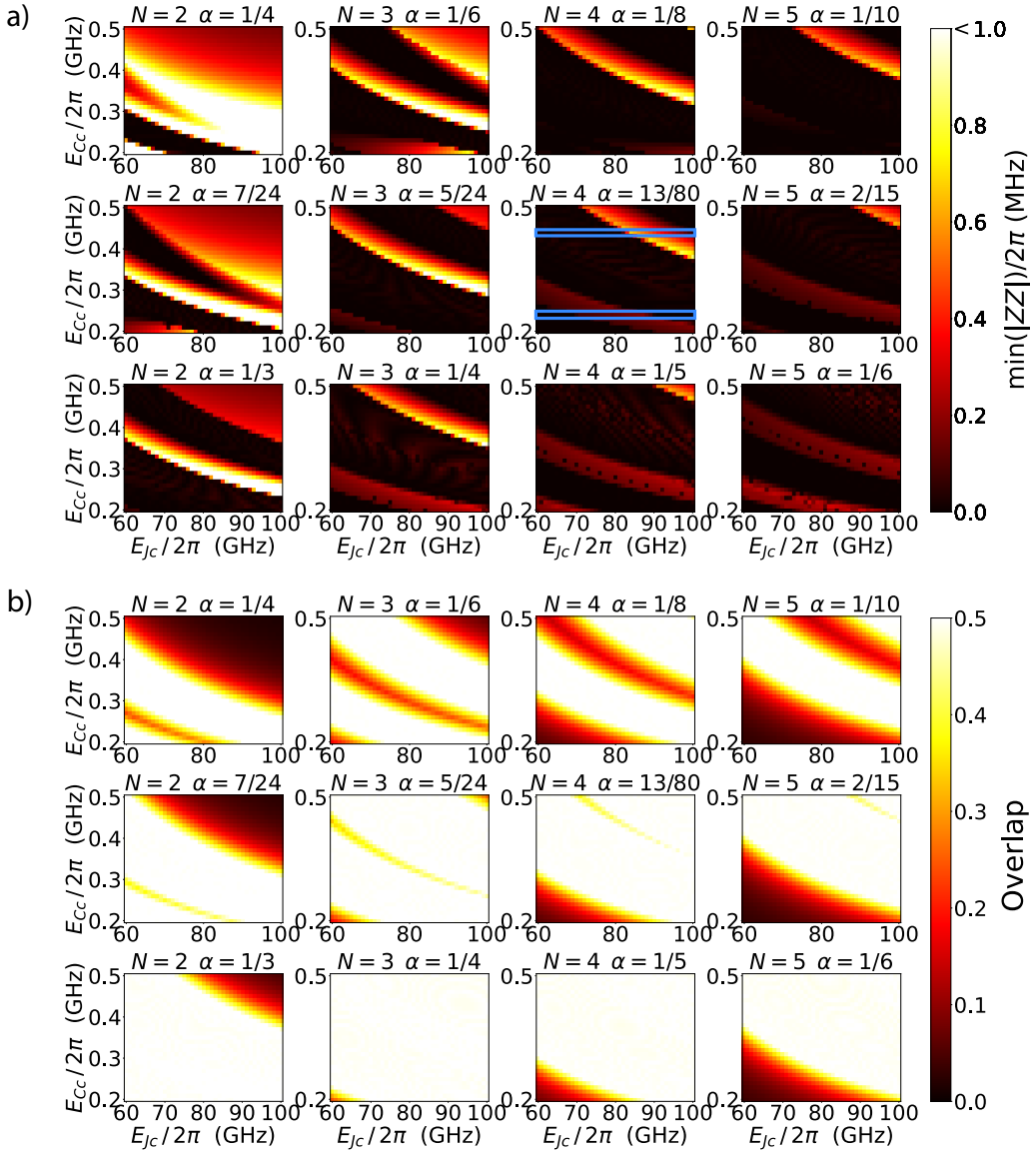


Figure 3.7 a) Plot of the minimum ZZ and b) Plot of the overlap between the dressed qubit state and the bare coupler state at minimum ZZ, for coupler parameters along $E_{\mathrm{Cc}}/2\pi = [0.2, 0.5]$ GHz and $E_{\mathrm{Jc}}/2\pi = [60, 100]$ GHz. Each plot corresponds to a number of JJs in the array $N = 2, 3, 4$ or 5 with a ratio α between the Josephson energies of the branches of the coupler loop. The plots in each column correspond to a specific N. The plots in the top row correspond to $\alpha = 1/2N$, the last row to $\alpha = 1/(N+1)$ and the row in middle to $\alpha = \frac{1}{2}(\frac{1}{2N} + \frac{1}{N+1})$. All the plots are at the same scale and any value of $|ZZ| > 1$ MHz is plotted in white. The qubit and coupling parameters used in this figure are the ones indicated in Eq. (3.56). The two blue boxes in the subfigure representing $N = 4$ with $\alpha = 13/80$ in a) are the regions for which we explore the spectra and the ZZ interaction versus flux in Fig. 3.8.

found that over 10^6 parameter sets, 50% have zero-ZZ operating points.

In addition to having zero-ZZ points, we also require fast gates with low leakage. For instance, if the coupler parameters are chosen such that the coupler frequency is very far from the qubits then it is rather easy to set ZZ to zero, however, the gate is then typically slower. To better understand the different regions of the plots of Fig. 3.7 and develop some intuition, we plot the spectra and the static ZZ versus flux for multiple device parameters. In Fig. 3.8, we consider devices with 4 JJs in the coupler array and $\alpha = 13/80$. The top figures represent the ZZ interaction χ_{ab} versus flux, for a fixed value $E_{Cc}/2\pi = 0.25$ in a) and 0.45 GHz in b). The colored lines correspond to a different $E_{Jc}/2\pi$ ranging from 60 to 100 GHz. As a guide to the eye, the regions explored in Fig. 3.8 correspond to the blue boxes in Fig. 3.7 a) for $N = 4$ and $\alpha = 13/80$. The bottom plots of Fig. 3.8 represent the energy levels versus flux for a fixed E_{Cc} and multiple E_{Jc}. The first excited state of the qubit a is indicated in red, while that of qubit b in orange. The first (second) excited states of the coupler are plotted in multiple colors, each corresponding to a with different E_{Jc}, in solid (dashed transparent) lines. As expected, the coupler frequency increases with E_{Cc} and E_{Jc}. Depending on the device parameters, the coupler can either cross both qubit frequencies, cross a single qubit or none of them.

For $E_{Cc}/2\pi = 0.25$ GHz, the coupler frequency is mainly below the qubits and the second level of the coupler can in some cases cross the higher frequency qubits, something which should be avoided to prevent qubit leakage to coupler states. In the upper plot of Fig. 3.8 b), we observe the presence of resonances in the ZZ interaction that correspond to avoided crossings between qubits ($|1_a 0_b 0_c\rangle$ or $|0_a 1_b 0_c\rangle$) and the coupler ($|0_a 0_b 1_c\rangle$), or between the level $|1_a 1_b 0_c\rangle$ and a state involving the first or second excited state of the coupler ($|q_a q_b 1_c\rangle$ or $|q_a q_b 2_c\rangle$). Since the higher energy spectrum is very crowded, this region is not shown. From Fig. 3.7 a) for $N = 4$ and $\alpha = 13/80$, and in the region where $E_{Cc}/2\pi = 0.25$ GHz, we expect a zero-ZZ point for the Josephson energies $E_{Jc}/2\pi = 60$ and 100 GHz. In Fig. 3.8, we do observe zero-ZZ points for $E_{Jc}/2\pi = 60$, 90 and 100 GHz, while for $E_{Jc}/2\pi = 70$ and 80 GHz we see that the presence of a resonance leads to a finite ZZ interaction. Note that the existing zero-ZZ points are located where the coupler is below the qubit frequencies.

Now looking at Fig. 3.8 b), for $E_{Cc}/2\pi = 0.45$ GHz, we see that an avoided crossing for $E_{Jc}/2\pi = 80$ and 90 GHz prevents the appearance of a zero-ZZ point. This avoided crossing is responsible for the large feature in the upper right corner in Fig. 3.7 a) for $N = 4$ and $\alpha = 13/80$. Interestingly, we find that for $E_{Cc}/2\pi = 0.45$ GHz and $E_{Jc}/2\pi = 100$ GHz, a zero-ZZ point can be found close to a sweet spot with respect to the flux (*i.e.* $\partial\chi_{ab}/\partial\bar{\varphi}_{ext} = 0$), meaning that the ZZ interaction is first order insensitive to flux in the coupler flux line. Note

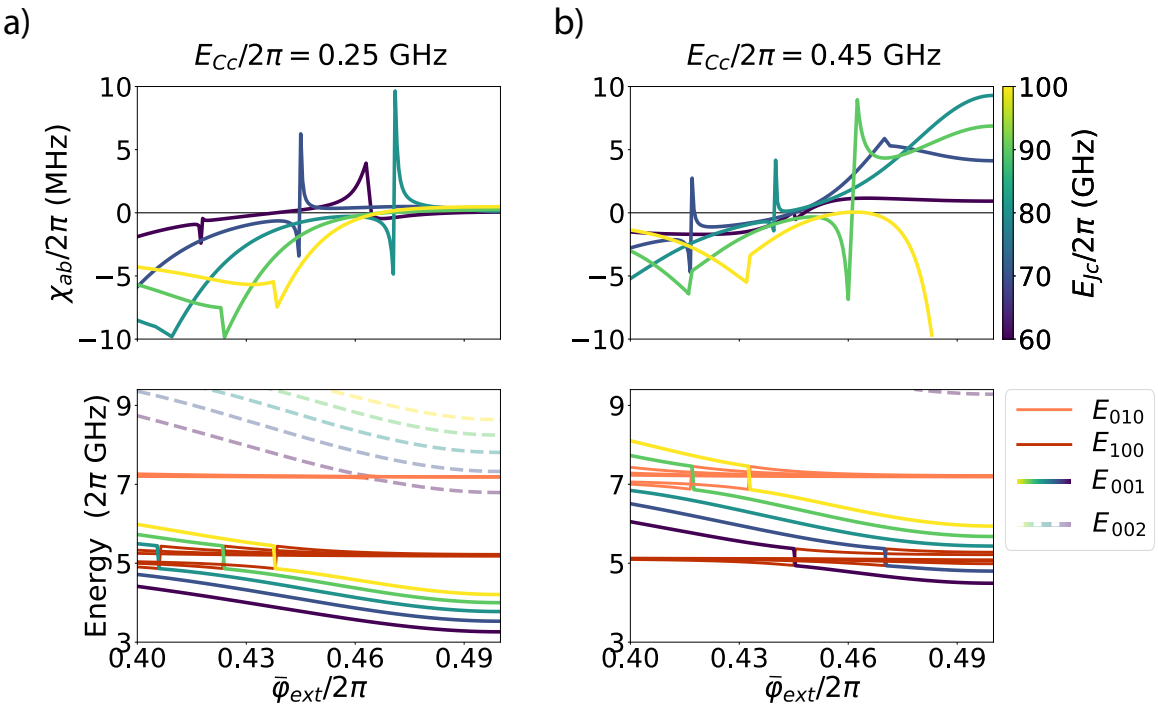


Figure 3.8 Spectra (bottom) and ZZ cross-Kerr interaction (top) in the flux range $\overline{\varphi}_{\text{ext}}/2\pi = [0.4, 0.5]$ for a) $E_{Cc}/2\pi = 0.25\,\text{GHz}$ (left) and b) $E_{Cc}/2\pi = 0.45\,\text{GHz}$ (right), for multiple values of E_{Jc}. Each lines in the top figures represent the ZZ interaction for different values of $E_{Jc}/(2\pi)$ ranging from 60 to 100 GHz. The horizontal black line at 0 is a guide to the eye. The bottom figures represent the device spectra, where the first excited states of the qubits are indicated in fixed red and orange colors, while the first (second) excited states of the coupler are plotted in multiple colors each associated to a device with different E_{Jc}, in solid (dashed transparent) lines. We denote E_{abc} the energy level of a state $|a_a b_b, c_c\rangle$. The qubit and coupling parameters for these figures are the ones indicated in Eq. (3.56), and the fixed coupler parameters are $N = 4$ with $\alpha = 13/80$. This figure explores the spectra and the ZZ interaction for the parameter regimes indicated by the blue boxes in Fig. 3.7.

also that for $E_{Cc}/2\pi = 0.45\,\text{GHz}$ the existing zero-ZZ points are located where the coupler is between the qubits. This is interesting as it allows the coupler to be simultaneously close in frequency to both qubits, maximizing the qubit-coupler interaction and, consequently, the effective qubit-qubit interaction.

By studying more extensively the spectra for the other plots of Fig. 3.7 a), we notice that the zero-ZZ regions are often located where the coupler is close in frequency to one of the qubits. Strong qubit-coupler hybridization can be responsible for shorter qubit lifetimes, since the coupler decay would impact the qubit due to qubit-coupler entanglement. In Fig. 3.7 b), we characterize this hybridization by plotting for the same parameters as Fig. 3.7

a) the overlap between the dressed qubit states and the bare coupler state at the flux value that minimizes the absolute value of the ZZ interaction. Fig. 3.7 b) is obtained for the same parameter regimes as in a). The larger the overlap is, the stronger the qubits and the coupler are hybridized. We see that in almost all the regions where zero-ZZ points exist, the qubits and the coupler are strongly hybridized. This is due to the fact the zero-ZZ points appear when the coupler approaches the qubit frequency. For $N = 2$ and 3, no regions can be found where ZZ is canceled while keeping the qubits and the coupler weakly hybridized. On the other hand, for $N = 4$ and 5 it is possible to both have zero-ZZ points and a small overlap between the dressed qubit and the bare coupler. This is for instance the case with $N = 5$ and $\alpha = 1/10$ in the lower left corner. However, these regions generally correspond to situations where the coupler is relatively far below the qubit frequencies, which is expected to reduce the speed of the gate. Moreover, as mentioned previously, the strong hybridization that exists between the qubits and the coupler at the zero-ZZ point can become an issue for the qubit coherence time. Indeed, the Purcell-type decay of the qubit due to the coupler is roughly given by the square of the qubit-coupler overlap times the coupler decay rate. Given the large overlaps that are observed, it it thus important for the coupler to have a large quality factor.

While the above results were obtained for the undriven system, we now turn out attention to the properties of the device in the presence of a drive. Namely, we now characterize the gate rate. In Fig. 3.9, we plot in a) to d), for a given choice of N and α, the minimum ZZ points, the overlap of the dressed qubits with the bare coupler state, and the iSWAP rate J_{ab} at the minimum ZZ point when the coupler is modulated with an amplitude $\delta\varphi/2\pi = 0.01$. Fig. 3.9 a) to d) correspond to different $N = 1$, 2, 3 and 4, and we chose α according to the observation of zero-ZZ regions in Fig. 3.7 and from the study of the spectra in those regions. Since simulating the gate rates take a significant amount of time, for each N and α, we only evaluate the gate rate in the regions surround by the blue box in the plots of ZZ and of the overlaps in Fig. 3.10. The gate rates J_{ab} are plotted on the same scale and we do not evaluate J_{ab} when the minimum ZZ point is larger than 0.1 MHz, those points are indicated in white. The maximum gate rates obtained for $N = 2$, 3, 4 and 5 are respectively 35.7, 15.4, 27.6 and 14.4 MHz, corresponding to $\sqrt{i\text{SWAP}}$ gates of around 14 ns, 32 ns, 18 ns and 35 ns. These gates are fast compared to typical parametric two-qubit gate times as shown in Sect. 1.3.3. For instance, cross-resonance gates are typically performed in a few hundreds of nanoseconds and CZ gates in $40 - 250$ ns. Note that Ref. [51] has achieved an iSWAP gate in 44 ns by performing a parametric-resonance gate. Faster two-qubit gates have been implemented using flux-tunable qubits or couplers. Note that in our case the drive amplitude could be increased, further reducing the gate times. The regions where the gate rate is the largest appear when the coupler is located between the qubit frequencies. When the

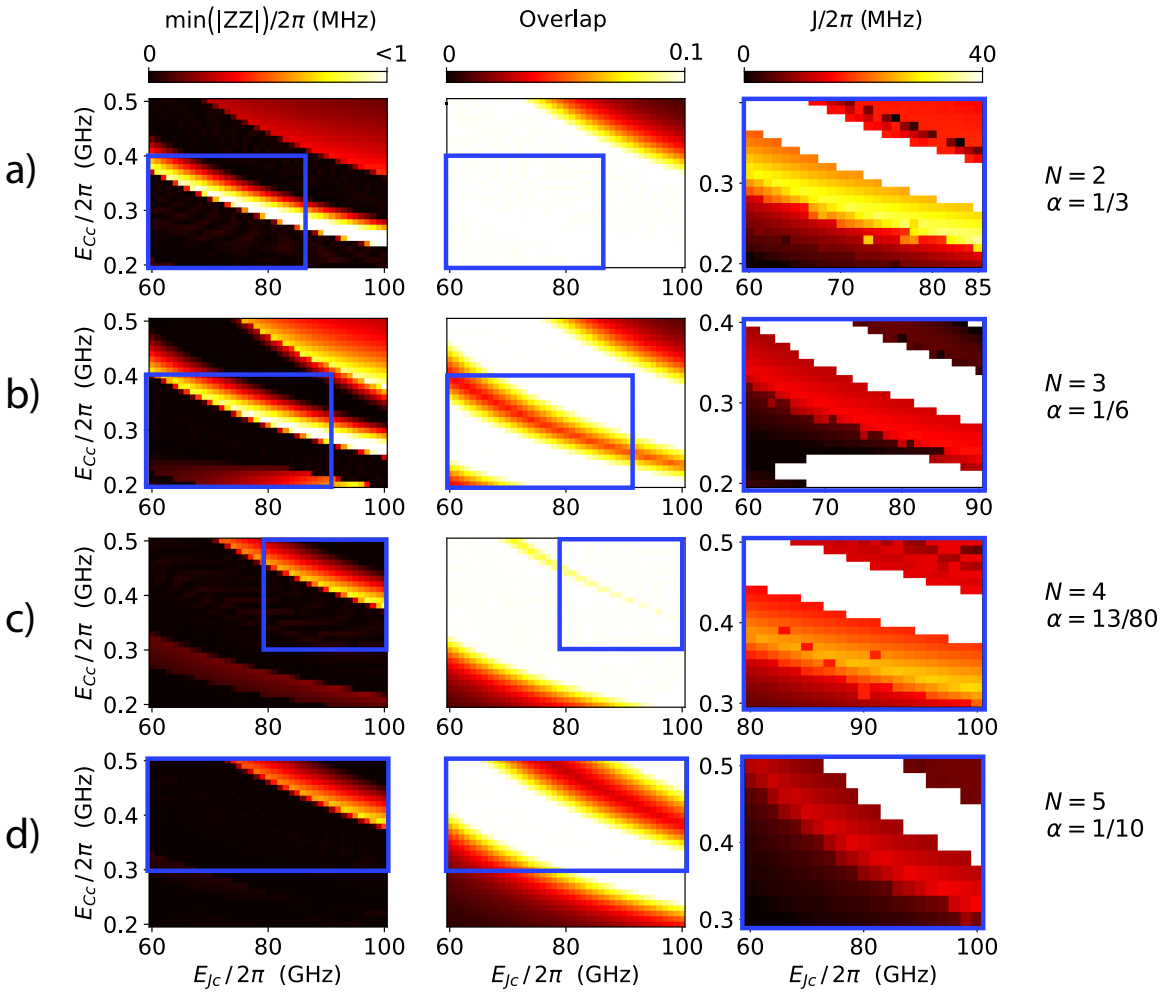


Figure 3.9 Colormap of ZZ, the overlap and J_{ab} versus E_{Cc} and E_{Jc} for multiple (N, α): a) $N = 2\&\alpha = 1/3$, b) $N = 3\&\alpha = 1/6$, c) $N = 4\&\alpha = 13/80$ and d) $N = 5\&\alpha = 1/10$. All the colormaps are at the same scale and any value of $|ZZ| > 1\,\text{MHz}$ is plotted in white. For each N and α, the gate rate is simulated in the regions surround by the blue box in the colormaps of ZZ and of the overlaps. We omit to simulate J_{ab} where the minimum ZZ point is above 0.1 MHz, those points are indicated in white. The qubit and coupling parameters used in this figure are the ones indicated in Eq. (3.56).

coupler is simultaneously close in frequency to both qubits, it maximizes the qubit-coupler interaction and, consequently, the effective qubit-qubit interaction. As mentioned earlier, the decrease in the gate rate in the lower left (upper right) corner of the plots can be explained by the fact that the coupler frequency decreases (increases) further away below (above) the qubits frequencies.

We now focus on a specific set of device parameters to study how the gate rate and the

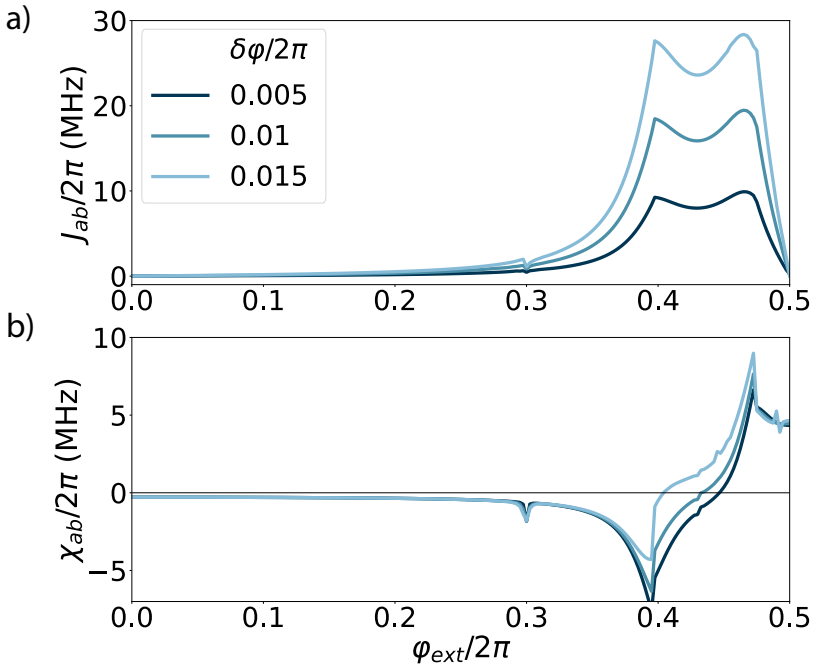


Figure 3.10 a) iSWAP gate J_{ab} rate and b) dynamical ZZ interaction χ_{ab} versus flux. Each colored line corresponds to a different drive amplitude $\delta\varphi/2\pi$ ranging from 0.005 to 0.015. The two main resonances at $\varphi_{ext}/2\pi = 0.4$ and 0.4725 correspond to the avoided crossing between the qubits and the coupler. The device parameters used in this figure are the ones indicated in Eq. (3.56).

dynamical ZZ vary as the drive amplitude increases. For this, we use the first parameter fit of a device that was experimentally fabricated by our collaborators from the Houck Lab at Princeton University:

$$\begin{aligned}
E_{\mathrm{C}a}/2\pi &= 0.205\,\mathrm{GHz} & E_{\mathrm{J}a}/2\pi &= 19.65\,\mathrm{GHz} & E_{\mathrm{C}ac}/2\pi &= -0.175\,\mathrm{GHz} \\
E_{\mathrm{C}b}/2\pi &= 0.205\,\mathrm{GHz} & E_{\mathrm{J}b}/2\pi &= 34.4\,\mathrm{GHz} & E_{\mathrm{C}bc}/2\pi &= 0.185\,\mathrm{GHz} \\
E_{\mathrm{C}c}/2\pi &= 0.4\,\mathrm{GHz} & E_{\mathrm{J}c}/2\pi &= 91.2\,\mathrm{GHz} & E_{\mathrm{C}ab}/2\pi &= 0\,\mathrm{GHz} \\
N &= 4 & \alpha &= 0.164\,. & &
\end{aligned} \tag{3.57}$$

Note that the qubit and coupling parameters are relatively close to the ones used in Eq. (3.56). In Fig. 3.10 a), we show the iSWAP gate rate J_{ab} versus flux and in b) the dynamical ZZ cross-Kerr interaction. We see that the gate rate is the largest around $\varphi_{ext}/2\pi = 0.35$ and 0.5, where the coupler is the closest to the qubits. Each line corresponds to a different drive amplitude $\delta\varphi$. The static zero-ZZ point is located between the qubits at $\varphi_{ext}/2\pi = 0.45$. The gate rate increases linearly with $\delta\varphi$ and reached 25 MHz corresponding to a 20 ns $\sqrt{i\mathrm{SWAP}}$ gate. This gate time is comparable to what is achieved with flux-tuned two-qubit

gates, but is particularly fast for a parametric two-qubit gate. The gate time also has to be compared to the qubit lifetime. Considering that fast gates can be implemented with this device, the impact of Purcell decay caused by the coupler may not be as significant if it remains within reasonable limits.

Chapter 4

Theoretical support to experiment

The work presented in this chapter of the thesis was done in collaboration with the group of Andrew Houck at Princeton University and with members of the Sherbrooke group. At Princeton University, Pranav Mundada, Andrei Vrajitoarea, Sara Sussman and Charles Guinn fabricated and measured the devices which will be presented below. Alex Place also contributed in fabricating the devices. In Sherbrooke, Alexandru Petrescu, Catherine Leroux and Agustin Di Paolo contributed in analyzing the experiments presented below.

In this section, we present and analyze the experimental results obtained together with our collaborators from the Houck Lab at Princeton University. We consider two generations of the experiment, corresponding to two devices which we refer to as device 1 and device 2 below. We first show in Sect. 4.1 a theoretical fit of the spectra of these two devices and of the static ZZ interaction versus flux, after which we reproduce in Sect. 4.2 Autler-Townes spectroscopy experiments using the Floquet quasi-energy spectrum to calibrate the drive amplitude. Finally, in Sect. 4.3 we characterize $\sqrt{i\text{SWAP}}$ gates implemented on the two devices. More details about the device fabrication and experiments can be found in the thesis of Sara Sussman in Ref. [62].

4.1 Parameter fit

As mentioned previously, the circuit parameters achieved experimentally deviate from the targeted parameters. The device is characterized using electromagnetic simulations to determine the circuit capacitances, and witness JJs are fabricated next to the device to determine the Josephson energies in the circuit. The parameter fit is then fine-tuned by fitting the device spectrum obtained from spectroscopy experiments. Following this approach, we find that the parameters of device 1 are:

$$\begin{aligned}
&E_{Ca}/2\pi = 0.205\,\text{GHz} && E_{Ja}/2\pi = 19.65\,\text{GHz} && E_{Cac}/2\pi = -0.18\,\text{GHz} \\
&E_{Cb}/2\pi = 0.205\,\text{GHz} && E_{Jb}/2\pi = 34\,\text{GHz} && E_{Cbc}/2\pi = 0.18\,\text{GHz} \\
&E_{Cc}/2\pi = 0.4\,\text{GHz} && E_{Jc}/2\pi = 91.2\,\text{GHz} && E_{Cab}/2\pi = 0\,\text{GHz} \\
&N = 4 && \alpha = 0.164\,.
\end{aligned} \tag{4.1}$$

In the experimental setup, only the qubits have readout resonators, the coupler energy levels are measured by strongly driving through the qubit readout lines [62]. In Fig. 4.1, we plot the device spectrum and the ZZ interaction measured experimentally versus flux along with a fit (dashed lines) obtained from numerical diagonalization of the Hamiltonian of Eq. (3.1) using the above parameter values. The fit is in good agreement with the spectrum and the ZZ interaction over a large flux range. A zero-ZZ point can be found at $\bar{\varphi}_{\text{ext}}/2\pi = 0.45$ which appears when the coupler is between the qubit frequencies. We remind from Sect. 3.2.2 that the gate rate is generally maximized when the zero-ZZ point is located where the coupler is between the two qubits. In Fig. 4.1 b), we observe two resonances in the ZZ interaction around $\bar{\varphi}_{\text{ext}}/2\pi = 0.43$ and 0.47, which are respectively due to the crossing of levels $|110\rangle$ and $|002\rangle$, and the crossing of the first excited state of the qubit a with the coupler. Despite a small shift in the location of the second resonance, the two resonances are well captured by the numerical simulations.

Following the above mentioned approach to fit the parameters of device 2, we find:

$$\begin{aligned}
&E_{Ca}/2\pi = 0.205\,\text{GHz} && E_{Ja}/2\pi = 13.65\,\text{GHz} && E_{Cac}/2\pi = 0.085\,\text{GHz} \\
&E_{Cb}/2\pi = 0.208\,\text{GHz} && E_{Jb}/2\pi = 26.55\,\text{GHz} && E_{Cbc}/2\pi = -0.08\,\text{GHz} \\
&E_{Cc}/2\pi = 0.445\,\text{GHz} && E_{Jc}/2\pi = 67.5\,\text{GHz} && E_{Cab}/2\pi = 0.01\,\text{GHz} \\
&N = 4 && \alpha = 0.1888\,.
\end{aligned} \tag{4.2}$$

The Josephson energies of the qubits are smaller in this second device to reduce their frequencies and detune them further from the readout resonators. The qubit-coupler

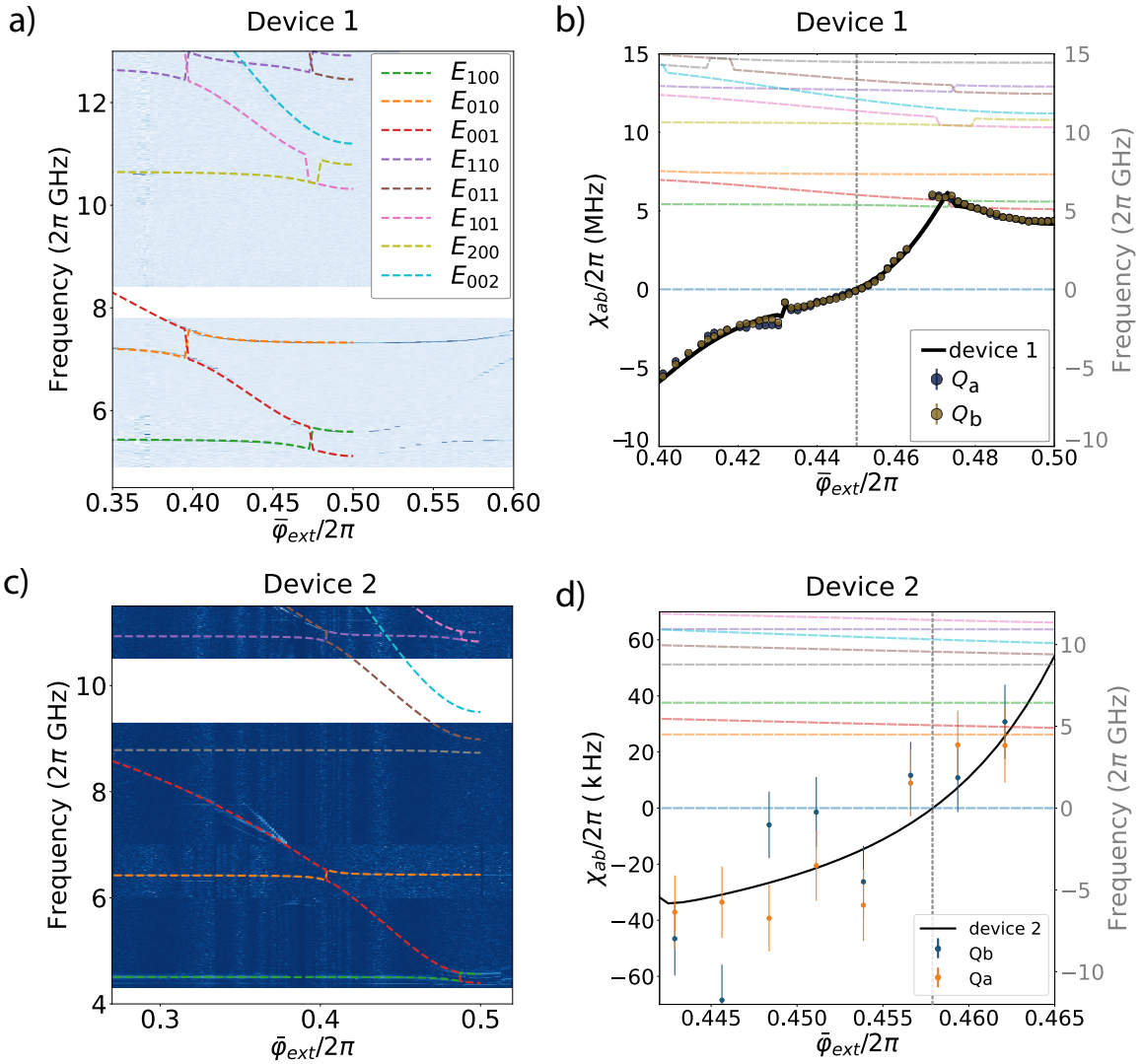


Figure 4.1 a) Spectrum of device 1 obtained from spectroscopy with an overlaid fit versus flux. The colored dashed lines correspond to the fitted spectrum using the parameters in Eq. (4.1). The energy levels are labeled as E_{abc}. b) ZZ cross-Kerr (χ_{ab}) interaction of device 1 versus flux. The blue and gold dots are the ZZ interaction measured through qubit a and b respectively. The black line is the fit obtained of device 1 and the dashed colored lines are the same energy levels as in a) and are provided as a visual aid to understand the resonances in the ZZ curve. The dashed vertical line indicated the zero-ZZ point. c) Same as a) but for device 2. d) Same as b) but for device 2.

couplings E_{Cac} and E_{Cbc} are lower in this device compared to device 1, but this is due to a fabrication error. The spectrum and the ZZ interaction are plotted in Fig. 4.1 c) and d) along with a fit. Note that for device 2 the ZZ interaction was measured in a smaller flux range around the zero-ZZ point compared to device 1. This might give the wrong impression

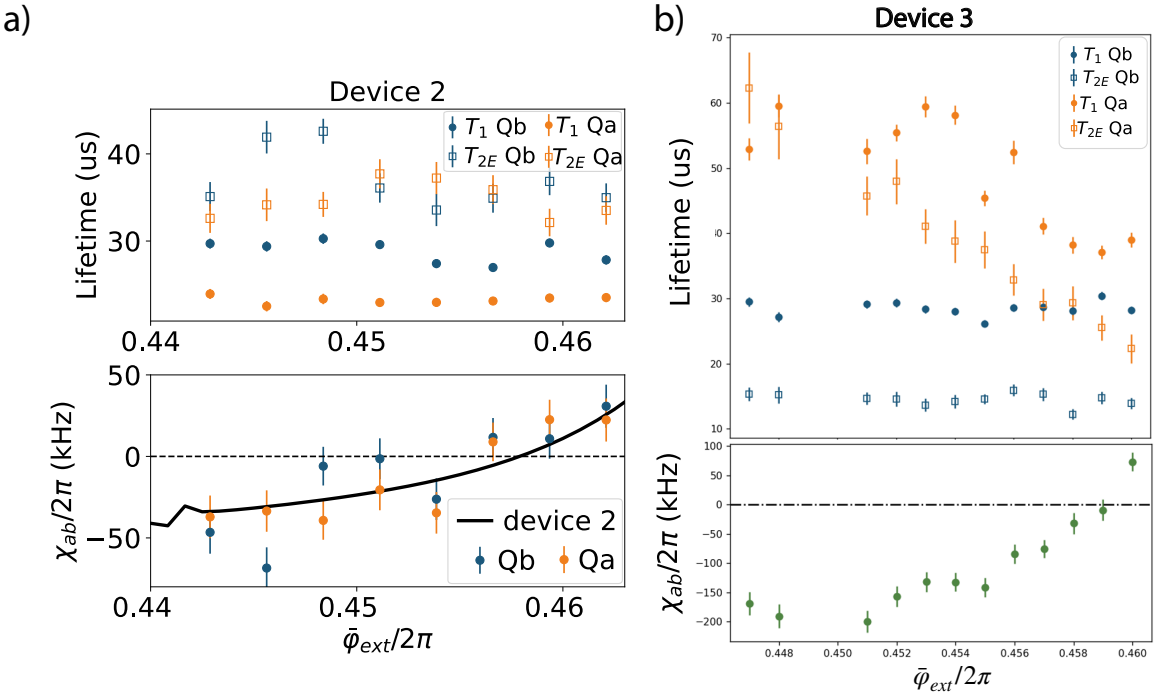


Figure 4.2 Qubit lifetime and ZZ cross-Kerr interaction (χ_{ab}) as a function of flux $\phi_{ext} = \bar{\varphi}_{ext}/2\pi$. a) For device 2 presented in Eq. (4.2). b) For a third device for which the parameters were not fitted. In both a) and b), qubit a (b) are the low (high) frequency qubits. Figures from Ref. [63].

that the ZZ interaction is overall smaller than for device 2 since the scale is now in kHz. For device 2 the zero-ZZ point is located at $\bar{\varphi}_{\text{ext}}/2\pi = 0.457$, where the coupler frequency is here again between the qubits.

The qubit lifetimes for device 1 measured at $\bar{\varphi}_{\text{ext}}/2\pi = 0$ are $T_1^{(a)} \sim 80\,\mu$s for qubit a and $T_1^{(b)} \sim 8\,\mu$s for qubit b. The dephasing times of the qubits are $T_2^{(a)} = 20\,\mu$s and $T_2^{(b)} = 4\,\mu$s. The poor lifetime of qubit b is attributed to the readout resonator which is closer to the qubit than targeted, leading to a fast qubit decay via the Purcell effect. Because the coupler was not coupled to a readout resonator, its lifetime was not measured. As long as the coupler and the qubits states are not too strongly hybridized, the qubits should not be affected by the lifetime of the coupler. Note, however, that at the flux point where the ZZ interaction is zero, we suspect qubit a to be strongly hybridized with the coupler since they are near an avoided crossing (*e.g.* see Fig. 4.1 b) at $\bar{\varphi}_{\text{ext}}/2\pi = 0.45$). The qubit-coupler entanglement could therefore have an impact on $T_1^{(a)}$ and $T_2^{(a)}$ at the zero-ZZ point. For device 1 the qubit lifetimes has, however, not been measured at this point. It has nevertheless been measured over larger flux ranges for other devices than device 1 and 2.

The upper plot of Fig. 4.2 a) depicts the qubit lifetimes close to a zero-ZZ point for device 2. The plot below represents the ZZ interaction in the same flux range together with a fit (black line) obtained with the parameters in Eq. (4.2). The qubit T_1s are $\sim 23\,\mu$s and $30\,\mu$s, and T_2s $\sim 35\,\mu$s, and remain relatively unchanged with flux. These qubit T_1s and T_2s are relatively low for typical fixed-frequency transmon qubits. We suspect these short qubit lifetimes to be due to fabrication errors rather than to the strong qubit-coupler hybridization because the T_1s and T_2s remain constant with flux around the zero-ZZ point [62]. Fig. 4.2 b) also shows the qubit lifetime and the ZZ interaction in a flux range close to the zero-ZZ point for a third device for which the parameters were not fitted but are expected to be similar to those of device 2. Here, we see in Fig. 4.2 b) that the T_1 and T_2 of qubit a are cut by half as we get closer to the zero-ZZ flux point. This is to be expected if qubit a and the coupler are strongly hybridized.

4.2 Autler-Townes spectroscopy

Autler-Townes spectroscopy is used in the present work to convert the drive power (in dBm) to flux quanta (in units of Φ_0). With this conversion factor in hand, it is possible to ensure that the drive power corresponds to a small fraction of the quantum flux. The Autler-Townes spectroscopy corresponds to two-tone spectroscopy [59], where a first drive of frequency ω_d is used to activate an interaction, while a second drive of frequency ω_p, which we will refer to as the *probe*, is used to probe the spectrum dressed by the first drive. This experiment provides a spectroscopy plot which shows how an energy level is affected by a drive ω_d due to ac-Stark and Bloch-Siegert shifts. Ultimately, it reproduces the avoided crossing of Fig. 2.1 shown in the Floquet section.

Here, the first drive is activated by modulating the flux through the coupler at the zero-ZZ point $\bar{\varphi}_{\text{ext}}/2\pi = 0.45$ and the dressed spectrum is probed by measuring the qubits as in Sect. 4.1. Fig. 4.3 shows the results of the Autler-Townes spectroscopy of device 1 around the coupler and qubit frequencies. The Floquet quasi-energy associated to the qubit and coupler states are overlaid (red dots) on the experimental data. We see that the quasi-energies are in excellent agreement with the experimental data, for $\delta\varphi/2\pi = 0.01$. For qubit a, see Fig. 4.3 a), the avoided crossing around $\omega_d/2\pi = 0.5\,$GHz corresponds to an interaction with the coupler. This drive frequency $\omega_d/2\pi = 0.5\,$GHz corresponds to the detuning between qubit a and the coupler at the zero-ZZ point. The avoided crossing around $\omega_d/2\pi = 1.9\,$GHz corresponds to the qubit-qubit interaction as it appears at the qubit-qubit detuning at the zero-ZZ point. As explained in Sect. 2.2, the minimum distance in the

avoided crossing along ω_d corresponds to the interaction amplitude between two states. The gate time can thus be obtained from the Autler-Townes experiment.

Similarly, in Fig. 4.3 b), we observe two avoided crossing in the qubit b spectroscopy at $\omega_d/2\pi = 1.9$ and 1.4 GHz, respectively corresponding to the qubit-qubit and qubit b-coupler detunings. For the coupler, in Fig. 4.3 c), the experimental data is qualitatively reproduced by the Floquet energies, however, quantitatively there is an overall shift of $\sim 10\,\text{MHz}$ of the quasi-energies along ω_p. Because the coupler does not have its own readout line, it is indirectly probed through the drive line of qubit a (which is closer in frequency to the coupler than qubit b) using a stronger drive amplitude. We attribute the frequency shift with the experimental data to the resulting strong ac-Stark. Nevertheless, the quasi-energies reproduce correctly the main avoided crossings at $\omega_d/2\pi \approx 0.55\,\text{GHz}$ and $1.44\,\text{GHz}$. Multiple lines appear in the coupler spectrum which we were not able to explain, but which we also attribute to the strong probe drive.

Experimentally calibrating the power of the drive, *i.e.* making the correspondence between the drive power applied experimentally and the drive amplitude $\delta\varphi$ of Eq. (3.1), is not an easy task. We use the quasi-energies to fit the avoided crossings and determine what fraction of a flux quantum, $\delta\varphi$, the drive power corresponds to. The stronger the drive amplitude is, the larger the avoided crossings are. In the particular case of Fig. 4.3, we reproduced the experimental data with a drive amplitude $\delta\varphi/2\pi = 0.01$, which we now know is associated to drive power applied experimentally.

4.3 Gate characterization

4.3.1 Chevron pattern

In the present work, we focus on the implementation of a $\sqrt{i\text{SWAP}}$ gate. While an iSWAP gate exchanges the states of the qubits and introduces a relative phase between the states $|01\rangle$ and $|10\rangle$, as:

$$\begin{aligned} |10\rangle &\rightarrow i|01\rangle \quad , \quad |00\rangle \rightarrow |00\rangle \\ |01\rangle &\rightarrow i|10\rangle \quad , \quad |11\rangle \rightarrow |11\rangle \end{aligned} \tag{4.3}$$

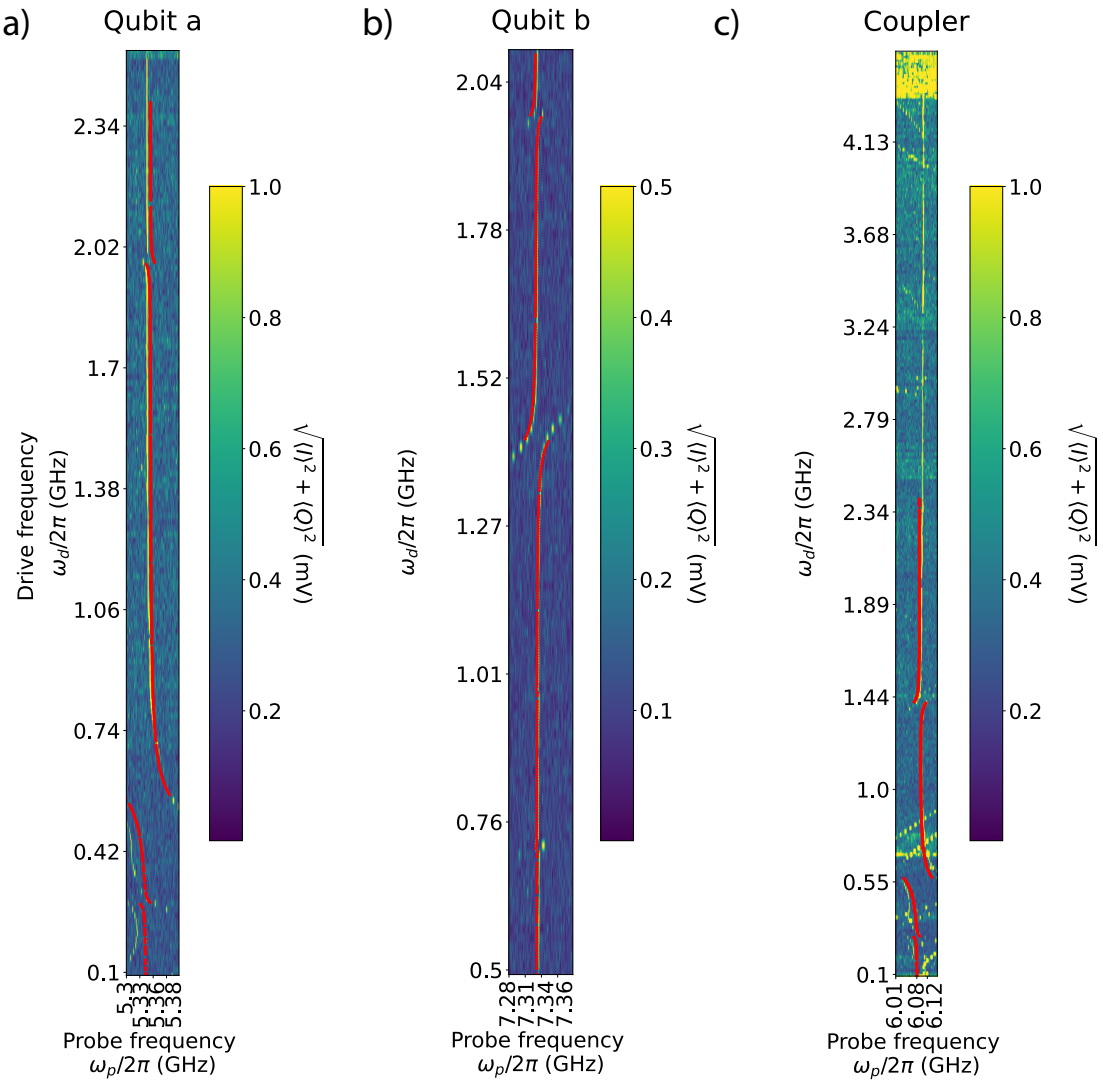


Figure 4.3 Autler-Townes spectrum close to the coupler and qubit frequencies, with an overlaid fit (red dots) of the Floquet quasi-energies for a drive amplitude $\delta\varphi/2\pi = 0.01$. For the coupler plot, there is an overall shift of the quasi-energies of about 10 MHz along ω_p.

the $\sqrt{i\text{SWAP}}$ is realized in half the time of an iSWAP gate and maximally entangles the qubits states

$$\begin{aligned} |10\rangle &\rightarrow (|10\rangle + i|01\rangle)/\sqrt{2} \\ |01\rangle &\rightarrow (i|10\rangle + |01\rangle)/\sqrt{2} \end{aligned} \tag{4.4}$$

while keeping $|00\rangle$ and $|11\rangle$ unchanged.

As discussed in Chapter 3, the $\sqrt{i\text{SWAP}}$ gate is generated by modulating the flux through

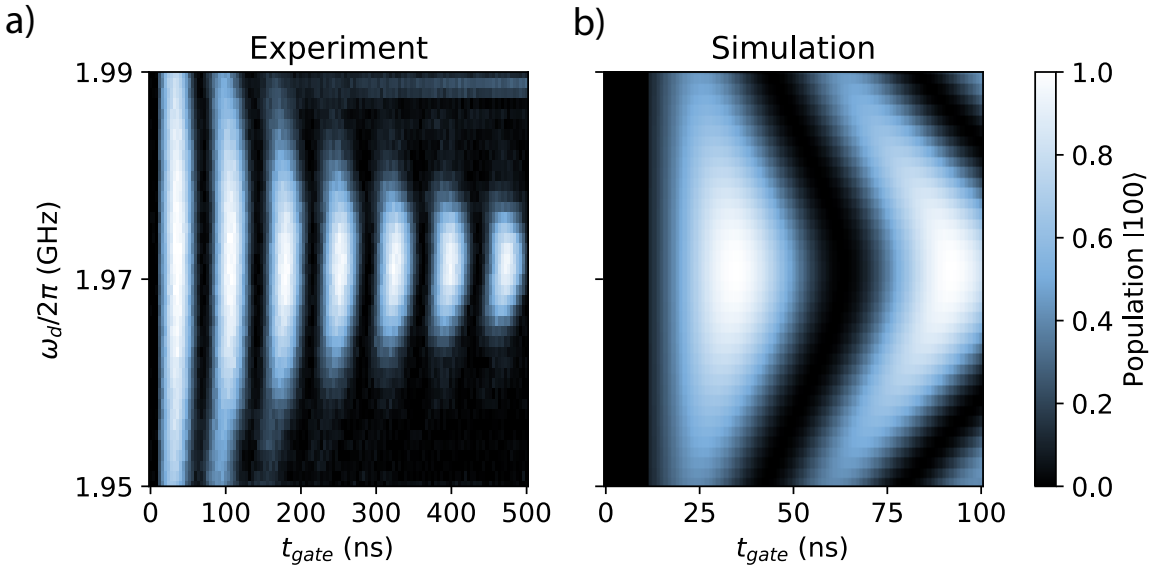


Figure 4.4 Plot of the population in qubit a versus the pulse time for multiple drive frequencies ω_d for device 1. a) Experimental data b) Simulated data using the device parameters in Eq. (4.1) and $\delta\varphi/2\pi = 0.01$.

the coupler at $\omega_d = \omega_b - \omega_a$. As soon as the interaction is turned on, the qubits exchange excitations leading to a swap of the population between the qubits. Because of the ac-Stark and Bloch-Siegert shifts, the optimal drive frequency that leads to a perfect exchange of the qubit populations is shifted from the undriven qubit detuning. Experimentally, the optimal drive frequency is obtained by measuring the qubit population in time for multiple ω_d around $\omega_b - \omega_a$ and identifying the frequency that leads to a maximum population exchange.

In Fig. 4.4, we plot, for device 1, the population of qubit a as a function of time and along the drive frequency, when initializing the state in $|0_a 1_b 0_c\rangle$ at $t = 0$. Fig. 4.4 a) and b) respectively correspond to the experimental data and the simulated data. Note that both figures are plotted on different time scales. Each point is obtained by using a flat-top Gaussian pulse, with a ramp up/down time of $N_\sigma\sigma$ where $\sigma = 3\,\text{ns}$ is the standard deviation of the Gaussian and $N_\sigma = 2$, and a flat top time of t_f. The gate duration is thus $t_{\text{gate}} = 2N_\sigma\sigma + t_f$. The pattern observed in Fig. 4.4 is referred to as a chevron. For a given ω_d we see that the qubit population oscillates in time. When the drive is off-resonant the oscillations gradually disappear and the qubit returns to its ground state. Conversely, on resonance, the oscillations repeat until the qubit decays. A full population swap occurs in $t_{\text{gate}} = 32\,\text{ns}$ around $\omega_d/2\pi = 1.97\,\text{GHz}$, for both the simulation and the experiment. This corresponds to a 16 ns $\sqrt{i\text{SWAP}}$ gate, which makes it a fast two-qubit gate compared to state-of-the-art parametric gates.

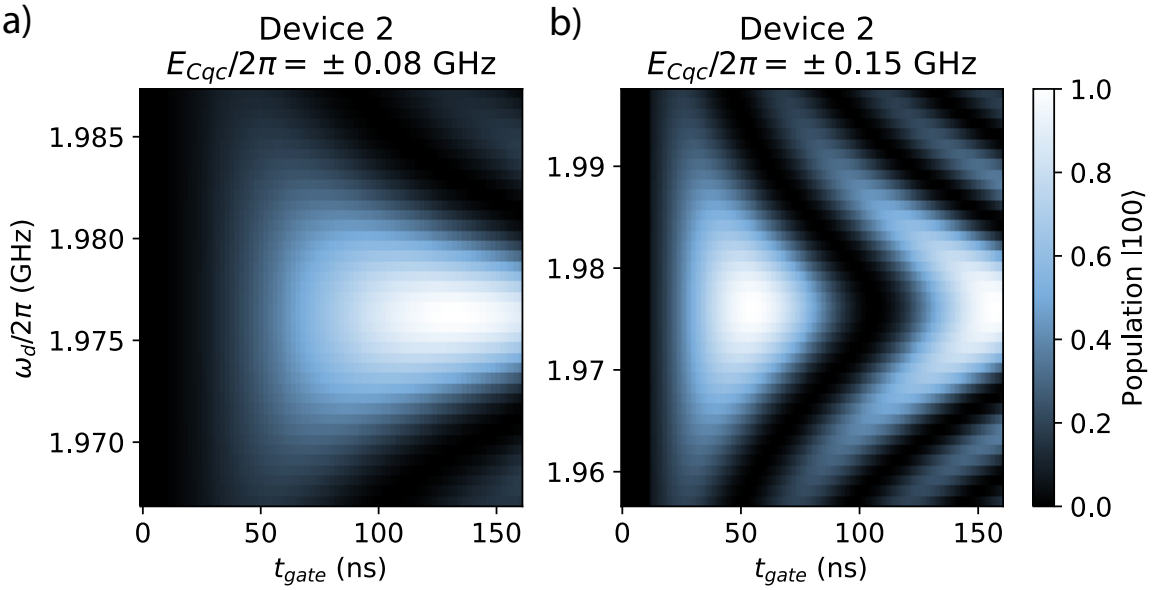


Figure 4.5 Simulation of the population in qubit a versus the pulse time t_{gate} for multiple drive frequencies ω_d for device 2 with $\delta\varphi/2\pi = 0.01$. a) Simulated data using the device parameters in Eq. (4.2) b) Simulated data using the device parameters in Eq. (4.2) but with larger couplings $E_{Cbc}/2\pi = -E_{Cac}/2\pi = 0.15\,\text{GHz}$.

For device 2, the time evolution of the qubit population has been measured for a few drive frequencies only, therefore we do not have a complete figure of the chevron pattern. However, the population at the optimal drive frequency showed slow $\sqrt{i\text{SWAP}}$ gates of $\sim 190\,\text{ns}$ for an uncalibrated drive power. In Fig. 4.5 a), we simulate the chevron for device 2 using the same drive amplitude as for device 1. A full population exchange occurs for $t_{\text{gate}} = 132\,\text{ns}$, which is to be compared to the $t_{\text{gate}} = 32\,\text{ns}$ obtained with the device 1 using the same drive amplitude. The reason for the gate being slow, is the low qubit-coupler couplings E_{Cac} and E_{Cbc}, which are half the size as those in device 1. Following this observation, Fig. 4.5 b) shows the expected results for a device that has the same parameters as device 2 but with larger couplings $E_{Cbc}/2\pi = -E_{Cac}/2\pi = 0.15\,\text{GHz}$. We see that the oscillations are more rapid, with a $\sqrt{i\text{SWAP}}$ obtained in 33 ns.

4.3.2 Gate fidelity and leakage

Before studying the leakage and the fidelity of the gate for device 1 and 2, we introduce and define some concepts used to characterize gates [64, 65].

In the device that we consider, the Hilbert space $\mathcal{H}$ spanned by the qubits and the coupler is larger than the qubit subspace in which the ideal dynamics occurs. To study the

leakage to the non-computational subspace, we split the Hilbert space into two subspaces χ_1 and χ_2. We define χ_1 the 4-dimentional qubit subspace made of the states $|q_a q_b 0_c\rangle$ with q_a, $q_b \in \{0, 1\}$, and χ_2 the $(d-4)$-dimensional leakage subspace comprising all the other states, where $d = \dim(\mathcal{H})$. The full Hilbert space is written as a direct sum of the latter two subspaces, $\mathcal{H} = \chi_1 \oplus \chi_2$. The state leakage of a density matrix $\rho \in \mathcal{D}(\mathcal{H})$ in the set of density matrices of the full Hilbert space is defined as:

$$L(\rho) = \mathrm{Tr}\left[\mathbb{I}_2 \rho\right] = 1 - \mathrm{Tr}\left[\mathbb{I}_1 \rho\right] \tag{4.5}$$

where $\mathbb{I}_1$ and $\mathbb{I}_2$ are respectively the projectors onto the subspaces χ_1 and χ_2.

Any quantum operation on ρ, including the leakage errors, can be described by a complete positive trace preserving (CPTP) map [64]. We call $\mathcal{E}$ the CPTP map describing a leakage error introduced by an imperfect gate. Note that a leakage error can be divided in two categories, leakage and seepage errors, which transfer population to and from the leakage subspace respectively. In the following work, we focus on the leakage and define the *leakage rate* L_1 as the average leakage of $\mathcal{E}$ [65]:

$$L_1(\mathcal{E}) = \int d\psi_1 L\left(\mathcal{E}\left(|\psi_1\rangle\langle\psi_1|\right)\right) = L\left(\mathcal{E}\left(\frac{\mathbb{I}_1}{d_1}\right)\right) \tag{4.6}$$

where $d_1 = \dim(\chi_1) = 4$, and the integral is taken over the Haar measure of all states in χ_1.

While characterizing leakage informs us about losses to the leakage subspace, it does not inform us about the performance of the gate within the computational subspace. To address this, we introduce a commonly used metric, the average gate fidelity [65]:

$$\begin{aligned} \overline{F}(\mathcal{E}) &= \int d\psi_1 \, \langle\psi_1| \mathcal{E}\left(|\psi_1\rangle\langle\psi_1|\right) |\psi_1\rangle \\ &= \frac{d_1 F_{\mathrm{pro}}(\mathcal{E}) + 1 - L_1(\mathcal{E})}{d_1 + 1} \end{aligned} \tag{4.7}$$

where the integral is taken over the states of χ_1 and $F_{\mathrm{pro}}(\mathcal{E})$ is called the process fidelity of $\mathcal{E}$ with the identity map on χ_1, defined as [64]:

$$F_{\mathrm{pro}}(\mathcal{E}) = \frac{1}{d_1^2} \mathrm{Tr}\left[\left(\mathbb{I}_1 \otimes \mathbb{I}_1\right) \mathcal{S}_{\mathcal{E}}\right] \tag{4.8}$$

where $\mathcal{S}_{\mathcal{E}}$ is the superoperator representation of $\mathcal{E}$.

In this section, we compare the gate operation including leakage U_{gate} with the unitary of a perfect gate operation U_{target}. The map of the leakage error is $\mathcal{E} = \mathcal{U}_{\mathrm{target}}^{\dagger} \mathcal{U}_{\mathrm{gate}}$, where $\mathcal{U}_{\mathrm{target}}$ is the CPTP map associated to the targeted gate and $\mathcal{U}_{\mathrm{gate}}$ the map represents the

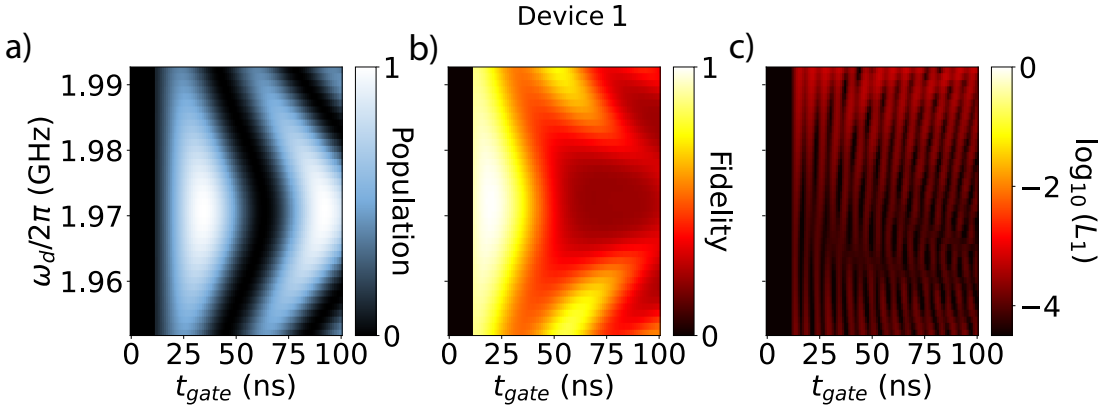


Figure 4.6 Simulation of a) the population in qubit a, b) the average gate fidelity and c) the leakage L_1 versus the pulse time t_{gate} and as a function of the drive frequencies ω_d for device 1. The time t_{gate} corresponds to the duration of the full gate pulse. The data is obtained using the device parameters in Eq. (4.1).

imperfect gate.

In Fig. 4.6 a), b) and c), we respectively show the population in qubit a, the average gate fidelity of a $\sqrt{i\text{SWAP}}$ gate and the leakage L_1 as a function of time and of the drive frequencies ω_d for device 1. Note that panel a) is identical to Fig. 4.4 b), where we observe a full population swap occurs in $t_{\text{gate}} = 32\,\text{ns}$ around $\omega_d/2\pi = 1.97\,\text{GHz}$. The simulated average gate fidelity is maximized at $t_{\sqrt{i\text{SWAP}}} = 20\,\text{ns}$, including the ramp times, with $\overline{F} = 99.6\%$ and a leakage $L_1 = 1 \times 10^{-4}$. The fidelity can be improved to $\overline{F} = 99.8\%$ by correcting for single-qubit rotations.

Overall the leakage is relatively low with a maximum value $L_1 = 6 \times 10^{-4}$. We see a fringe structure in the leakage which is due to leakage into the coupler. A population exchange still occurs between the qubits when the coupler is in its higher excited energy levels, *i.e.* between $|1_a 0_b 1_c\rangle \leftrightarrow |0_a 1_b 1_c\rangle$ and $|1_a 0_b 2_c\rangle \leftrightarrow |0_a 1_b 2_c\rangle$. The frequency of these transitions are respectively $\omega_d/2\pi = 1.991$ and $1.992\,\text{GHz}$ when the drive if off, but we suspect the drive to dress the energy levels, thereby shifting the frequency for a perfect swap of excitations between those levels. A wider simulation of the chevron for $\omega_d/2\pi > 1.99\,\text{GHz}$ shows a chevron pattern in the leakage plot.

More generally, the generated gate operation can be compared to an fSim gate represented

by the unitary matrix [66]:

$$\mathrm{fSim}(\theta,\xi,\chi,\gamma,\varphi) = \begin{matrix} \\ |00\rangle \\ |10\rangle \\ |01\rangle \\ |11\rangle \end{matrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & e^{-i\gamma-i\xi} & -ie^{-i\gamma+i\chi} & 0 \\ 0 & -ie^{-i\gamma-i\chi} & e^{-i\gamma+i\xi} & 0 \\ 0 & 0 & 0 & e^{-2i\gamma-i\varphi} \end{pmatrix}, \tag{4.9}$$

where the parameters $(\theta,\xi,\chi,\gamma,\varphi)$ respectively parametrize an iSWAP gate, (ξ,χ,γ) single-qubit gate angles and φ a CZ gate. By optimizing all the parameters of the fSim to maximize the fidelity with the imperfect U_{gate} we have obtained, it is possible to uncover its iSWAP, CZ and single-qubit gate components. For instance, the optimized parameters (ξ,χ,γ) of the fSim can then be used to correct for single-qubit rotations, which can further improve the gate fidelity. Here, the advantage of comparing our gate to the fSim is that it captures the dominant interactions of our Hamiltonian while having fewer free parameters to optimize over than an arbitrary two-qubit unitary. The fSim gate with the highest fidelity to U_{gate} may not necessarily be the intended target gate, but it can still be a maximally entangling. Optimizing the parameters of the fSim gate in the context of the numerical simulation of the gate implemented with device 1, we now find a fidelity of 99.99(8)%.

For device 1, an fSim gate close to $\sqrt{i\mathrm{SWAP}}$ has been experimentally achieved in 15 ns with a fidelity of 98.8% using a combination of two-qubit cross-entropy benchmarking with single-qubit randomized benchmarking as is done in Ref. [66]. Given the coherence times measured in Sect. 4.1, a fidelity of 99.8% would be expected for a 15 ns gate. Assuming that the optimized fSim gate is the closest resembling the implemented gate, the fSim fidelity suggests the existence of leakage. Leakage has been explored by benchmarking the purity of the two-qubit gate and it has been found to be lower than what is expected from noise, in agreement with the difference found between the coherence-limited fidelity and the measure fidelity. Moreover, the purity decay for the single-qubit gates observed is similar to the two-qubit gate purity decay, even though the two-qubit gate is about 4 times faster than single-qubit gates. In device 1, the gate fidelity is therefore limited by leakage. This could be explained by a decrease of the qubit coherence times induced by the presence of a strong drive, or by the population of the coupler at the end of the pulse. Compared to typical parametric two-qubit gates, our gate is one of the fastest that exist, and this for a fidelity comparable to some of the best parametric gates.

In Fig. 4.7 a), b) and c) we show the same plots as in Fig. 4.6 but for simulations of device 2. As shown in the previous section, for this device a full population exchange occurs

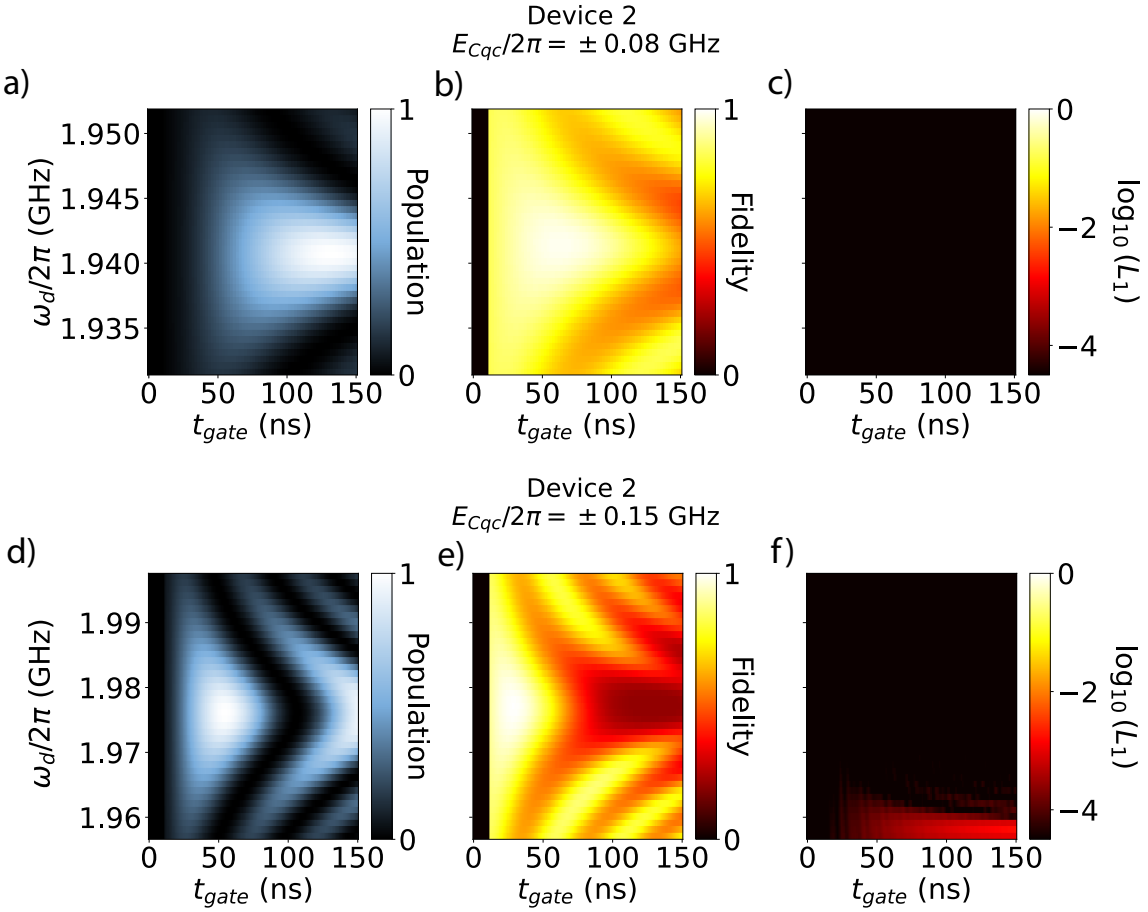


Figure 4.7 Simulation of a) the population in qubit a, b) the average gate fidelity and c) the leakage L_1 versus the pulse time t_{gate} and as a function of the drive frequencies ω_d for device 2. The data is obtained using the device parameters in Eq. (4.2). d), e) and f) are the same as a), b) and c) respectively, but for device 2 with larger couplings $E_{Cbc}/2\pi = -E_{Cac}/2\pi = 0.15\,\text{GHz}$.

for $t_{\text{gate}} = 132\,\text{ns}$. The simulated average gate fidelity is maximized at $t_{\sqrt{i\text{SWAP}}} = 64\,\text{ns}$, including the ramp times, with $\overline{F} = 98.4\%$ and a leakage $L_1 = 3 \times 10^{-6}$. It is unclear why the fidelity is lower compared to device 1, since dissipation is not accounted for in these simulations, and the L_1 leakage is very low. However, the closest fSim gate gives a fidelity of 99.99(9)%. In terms of gate time and gate fidelity, device 2 is not as favorable as device 1.

In Fig. 4.7 d), e) and f) we show the same plots as in a), b) and c) for the same device but with larger couplings $E_{Cbc}/2\pi = -E_{Cac}/2\pi = 0.15\,\text{GHz}$. Here, the simulated average gate fidelity is maximized at $t_{\sqrt{i\text{SWAP}}} = 29\,\text{ns}$, including the ramp times, with $\overline{F} = 99.6\%$ and a leakage $L_1 = 2 \times 10^{-5}$, similar to device 1. We note the appearance of leakage around $\omega_d/2\pi = 1.96\,\text{GHz}$ which could be due to a transition between the levels $|2_a 1_b 0_c\rangle \leftrightarrow |1_a 2_b 1_c\rangle$.

The closest fSim gate gives a fidelity of 99.99(8)%. It is here again unclear why the fidelity is high compared to device 2 with lower coupling.

Conclusion

In Part I of this thesis, we have introduced a new superconducting circuit design that allows for the cancellation of the unwanted ZZ interaction between the qubits and enables the execution of parametric gates in less than 50 ns. Our design is made of far-detuned fixed-frequency transmons capacitively coupled through a generalized flux qubit used as a coupler, see Fig. 3.1. The ingredient which makes it possible to cancel the unwanted ZZ interaction between the qubits is the positive anharmonicity of the coupler. In Chapter 3, we have shown that it is possible to cancel the ZZ interaction in the case where the qubits and the coupler have opposite anharmonicities. To the best of our knowledge, qubits with opposite anharmonicities have been used to cancel the ZZ interaction, but not couplers with opposite anharmonicities to the qubits. We chose the qubits to be far detuned such as to strongly suppress qubit-qubit interactions during idle time. However, this also comes at the cost of reducing gate interactions. To overcome this, we activate gates by parametrically modulating the flux through the coupler loop. This modulation provides the necessary energy to bridge the gap between energy levels which otherwise would be far off-resonant.

We developed a theoretical model to describe the superconducting device in Fig. 3.1. Together with Alexandru Petrescu, we developed a time-dependent Schrieffer-Wolff perturbation theory which to second order in the perturbation already reproduces well numerical simulations in the case of moderate to strong drive amplitudes. Despite the qualitative agreements of the perturbation theory with the full circuit Hamiltonian in Eq. (3.1), full numerical simulations are in order to quantitatively predict experimental results. Also, one of my personnal contributions was the development of a numerical method based on Floquet theory which allows us to determine the interaction amplitudes been two energy levels when activating a gate, significantly faster than standard methods based on simulating the gate dynamics.

We studied the behavior of relevant quantities like the ZZ interaction, the gate rate or the qubit-coupler hybridization, with the device parameters. The simulation shown in

Chapter 3 and Chapter 4 were conducted by myself. In our exploration for finding favorable device parameters, we have shown that for any number of JJs between 2 and 5 in the array of the coupler can exhibit a zero-ZZ operating point provided that the charging energy E_{Cc}, E_{Jc} and the anisotropy ratio α between Josephson energies of coupler branches are appropriately chosen. However, we have seen that it is easier to find a device with a zero-ZZ point for $N = 4$ and $N = 5$ and $\alpha \sim 1/2N$. Note that fabricated devices always deviate from the targeted parameters with a Gaussian distribution of all the parameters with a 10% standard deviation. We showed that starting with a set of target parameter that exhibits a zero-ZZ point, out of 10^6 device 50% of the fabricated device will still have a zero-ZZ operating point. When a zero-ZZ point exists, it is always located between $\bar{\varphi}_{\text{ext}}/2\pi = [0.3, 0.5]$, where the coupler anharmonicity is positive, and the coupler frequency is generally close to one of the qubits or between the two qubits. Situations where the zero-ZZ point is located where the coupler is between the qubit frequencies are particularly interesting as they allow the coupler to be simultaneously close in frequency to both qubits, thereby maximizing the qubit-coupler interaction and, consequently, the effective qubit-qubit interaction. We show that it is possible to reach $\sqrt{i\text{SWAP}}$ gate times in 10 to 40 ns using a moderate drive amplitude. An issue though of the coupler being close to the qubit frequencies is that they can strongly hybridize with the coupler, ultimately leading to a reduction of the qubit lifetime due to Purcell-type decay through the coupler. In the case of strong qubit-coupler hybridization, it is thus important for the coupler to have a large quality factor. This might be one of the main limiting factors of this architecture.

Our experimental collaborators have achieved a 15 ns fSim gate close to $\sqrt{i\text{SWAP}}$ with a fidelity of 98.8%, making it a state-of-the-art parametric two-qubit gate. They have shown that the gate fidelity is limited due to leakage. Despite the leakage limitation, this gate is one of the fastest that exists for a fidelity comparable to some of the best parametric gates. One of the major challenges, though, is the fabrication of those devices. First because parameter deviations can significantly change the device properties, and second because fabrication errors affect the qubit lifetimes. The devices that were fabricated typically had qubit lifetimes about $T_1, T_2 \sim 20 - 40\,\mu$s at the zero-ZZ point. A way to increase the qubit lifetime would be to further decrease the qubit frequencies, as the quality factor, which depends on the material used for fabricating the qubit is roughly given by the qubit frequency time its relaxation rate. In the present work, we have mainly focused on qubits detuned by about 2 GHz, but such large detuning might not be necessary, this could be further explored. Moreover, given that the qubit lifetimes are comparable to flux-tunable qubit, we could also consider building a similar device as in Fig. 3.1 but with flux-tunable transmon qubits instead. This would also alleviate some of the fabrication issues. In this work, we only focused on iSWAP like gates, however, as shown in Chapter 3 other interactions are available

on this device. It would be interesting to explore other gates. Using two drive modulations of the coupler flux can further diversify the implementable gates.

Part II

Variational Quantum Algorithm

Before relevant problems can be solved on quantum computers, numerous challenges need to be overcomed both on the hardware and software sides. In Part I of this thesis, we introduced a superconducting device aimed at improving the speed and fidelity of two-qubit gates, which shows our efforts in improving quantum hardware. In Part II, our focus shifts to one of the software aspects of quantum computing: quantum algorithms. Quantum algorithms have garnered interest for their possible applications in cryptography, quantum chemistry, combinatorial optimization, machine learning, quantum materials, finance, *etc.*, where classical algorithms struggle to provide solutions. In fact, certain problems are currently very resource intensive and require excessively long simulation times, thus the need for exploring alternatives beyond classical methods. A wide range of classically intractable problems can be formulated in terms of Hamiltonians, with their resolutions relying on the determination of its ground state. For instance, numerous problems relevant for industry applications can be mapped to the graph coloring problem in combinatorial optimization. The Quadratic Unconstrained Binary Optimization (QUBO) technique [67] which can be used to solve the graph coloring, Max-Cut, and scheduling management problems, is unfortunately known to be an NP-hard problem. Remarkably, it turns out that the objective function formulated by the QUBO problem can be expressed as an Ising model [68], which ultimately can be written as a Hamiltonian [69]. Finding the ground state of the problem Hamiltonian on a quantum processor thus amounts to solving the problem at hand.

Variational algorithms (VA) are already extensively used in physics and chemistry [70], for instance, to find the ground states of some molecule or to study the time evolution of a quantum system. The idea behind VAs is to design a parametrized variational form, also called an *ansatz*, describing the wave function of the system. The parameters of the variational form are then optimized by minimizing a cost function until convergence. The convergence of VAs in terms of speed and precision highly depends on the structure of the ansatz and its ability to explore the Hilbert space, its design is thus of major importance. The advantage of VAs compared to exact solutions obtained through Schrödinger's equation is that VAs only explore a relevant subspace of the Hilbert space, if the ansatz is appropriately chosen. This approach partially addresses the challenge of solving Schrödinger's equation, which becomes exponentially complex on classical computers as the system size increases. VAs have shown to be effective to solve some problems, however, as soon as the optimal solution is highly entangled, the variational forms struggle to efficiently approximate the wave function. Quantum algorithms have been introduced as a new strategy to tackle these issues. In fact, it has been shown that the spectrum and the eigenstates of a Hamiltonian can be obtained using the phase estimation algorithm introduced by Kitaev in the 90s [71]. Based on this work, a certain number of algorithms have been developed, some showing for example that molecular energies can be computed with a polynomial scaling [72, 73]. Unfortunately,

despite the progress of the last decades to reduce the resources needed to implement those quantum algorithms, in terms of the number of qubits and gates, they remain unrealistic to implement on today's noisy intermediate-scale quantum (NISQ) devices.

In this context variational quantum algorithms (VQA) were developed to get around these challenges. Inspired from their classical analog, VQAs may be the family of algorithms most susceptible to show quantum advantage on NISQ devices [74]. Similarly to VAs, the idea consists in designing a parametrized quantum circuit, also called an ansatz (or ansätze in the plural), that prepares a quantum state from some initial state. The circuit parameters are then optimized such that the final state converges towards the solution of the problem being addressed. VQAs are hybrid quantum-classical algorithms, where the quantum state is prepared on a quantum processor but the parameter optimization is performed classically until convergence, as shown in Fig. 5.0. Similar to VAs, the convergence of the algorithm highly depends on the structure of the ansatz, meaning the choice and the order of the operations in the ansatz, and is thus of critical importance. In the last decades, different types of ansätze were proposed in the literature to address various problems, each having its own set of strengths and weaknesses. In this second part of this thesis, we will focus on a new type of ansatz which we will apply to the Fermi-Hubbard problem. This ansatz is a hybridization of a pre-existing ansatz, the Quantum Optimal Control inspired Ansatz (QOCA) [75], and a protocol for designing ansätze, the Adaptive Derivative-Assembled Problem-Tailored ansatz Variational Quantum Eigensolver (adapt-VQE) [76], which we will describe in the following chapters.

In Chapter 5, we will first introduce the building blocks and the basic concepts of VQAs and will present the Fermi-Hubbard model (FHM). Chapter 6 will describe our newly developed ansatz, *adapt-QOCA*, as well as the two algorithms on which it is based on, *i.e.* adapt-VQE and QOCA. Finally, in Chapter 7, we benchmark our ansatz on the FHM. We will study how adapt-QOCA performs in terms of energy, fidelity and two-qubit gate counts compared to QOCA and adapt-VQE. We also explain the ingredients that allows adapt-QOCA to outperform other ansätze.

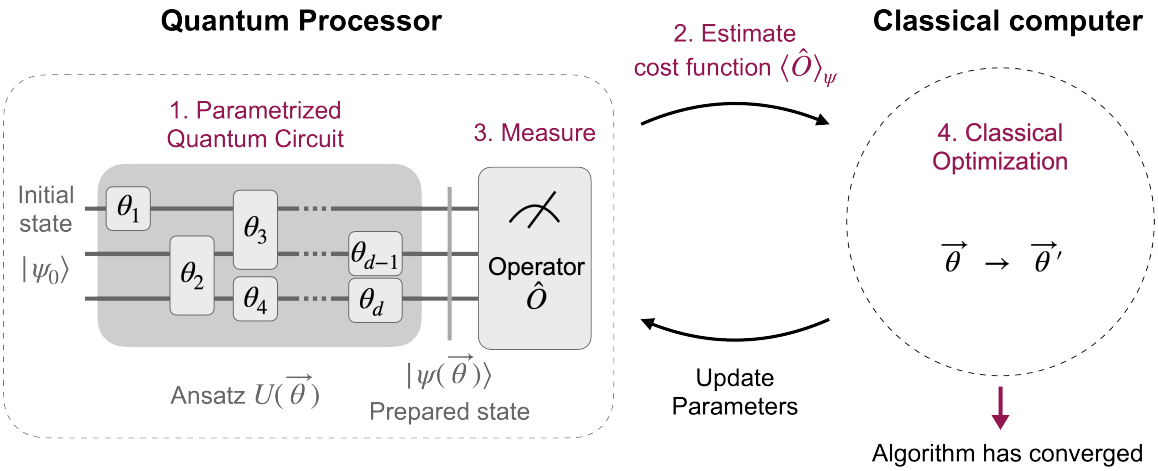


Figure 5.0 Variational Quantum Algorithm. Illustration of the main blocks of a VQA: the PQC or ansatz, the measurement and estimation of a cost function represented by an operator $\hat{O}$, and the classical optimization. A state $|\psi(\theta)\rangle$ is prepared by applying the ansatz $U(\vec{\theta})$ on an initial state $|\psi_0$. A cost function is estimated by measuring some operator $\hat{O}$ on the prepared state. The parameters of the PQC $\vec{\theta} = (\theta_1, \cdots, \theta_d)$ are then classically optimized and updated until the algorithm has converged.

Chapter 5

Introduction

In this section we describe the main building blocks of VQAs [74], see Fig. 5.0:

1. The variational form or parametrized quantum circuit (Sect. 5.1),
2. The cost function to be minimized (Sect. 5.2),
3. The measurement of the cost function (Sect. 5.3),
4. The classical optimization (Sect. 5.4).

5.1 Parametrized Quantum Circuits

The central part of a VQA is the *Parametrized Quantum Circuits* (PQC), which is also referred to as *variational form* or *ansatz* (or ansätze in the plural). It is described by a unitary operation $U(\vec{\theta})$ that depends on a certain number of parameters $\vec{\theta} = (\theta_1, \theta_2, \cdots)$. The PQC takes as an input an initial state $|\psi_0\rangle$ and outputs a state that depends on the circuit parameters, $|\psi(\theta)\rangle = U(\theta)|\psi_0\rangle$. Ultimately we want to find the parameters that will lead to the solution of the problem that is being solved. In VQAs, these parameters are found by optimizing a cost function using a classical subroutine.

The initial state $|\psi_0\rangle$ can be any state within the Hilbert space, but ideally, it is chosen to be close to the optimal solution, to facilitate the convergence of the VQA. For instance, in quantum chemistry one would choose the Hartree-Fock state. In some cases no educated guess for the initial state is known or is easy to prepare. In those cases, a reference state such as $|0\rangle^{\otimes n}$, or a random state can be chosen.

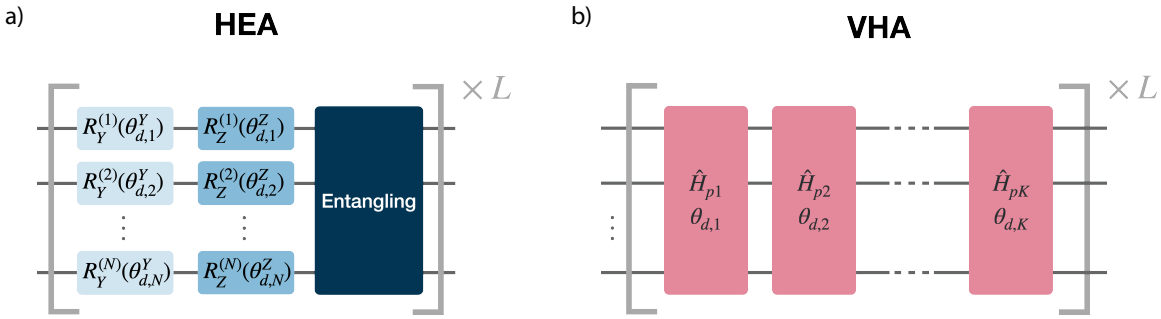


Figure 5.1 Standard ansätze. Illustration of a single layer of two well-known ansätze: a) the hardware efficient ansatz (HEA) and b) the variational Hamiltonian ansatz (VHA). a) The ansatz is made of single-qubit rotations $R_\sigma^{(n)}(\theta_{d,n}^\sigma)$ of angle $\theta_{d,n}^\sigma$ around the axis associated to the Pauli matrix $\sigma \in \{Y, Z\}$ on qubit n. At the end of a layer of HEA an unparametrized entangling gate is applied to all the qubits. The main block of HEA is repeated L times. b) The ansatz is decomposed in a set of parametrized operations $H_{\mathrm{p}k}$ with variational parameters $\{\theta_{d,k}\}$ and where the terms $H_{\mathrm{p}k}$ come from a decomposition of the problem Hamiltonian $H_{\mathrm{p}} = \sum_{k=1}^{K} c_k H_{\mathrm{p}k}$ with $\{c_k \forall k \in \{1, 2, \cdots, K\}\}$ a set of coefficients. The main block of VHA is repeated L times.

The performance of a VQA, in terms of speed and precision of convergence, highly depends on the form of the ansatz. However, its optimal structure is not known beforehand. Just as the choice of initial state can be informed by an educated guess, the PQC can be designed by leveraging the knowledge we possess about the problem we aim to solve. In this sense, PQCs can be divided in two broad categories: hardware-efficient ansätze and problem-inspired ansätze. The first is designed such as to adapt to hardware constraints, while the latter is designed taking into account the physics of the problem, generally requiring long and highly connected circuits (*i.e.* circuits with multiple entangling gates and requiring a high qubit connectivity). We now take a quick overview of two standard ansätze to illustrate these two categories, and provide some context on how our work relates to the existing ansätze.

Hardware Efficient Ansatz:

The hardware efficient ansatz (HEA) consists in a sequence of parametrized single-qubit rotations on each qubit, followed by an unparametrized entangling gate on all the qubits [77] (see Fig. 5.1 a)):

$$U_{\mathrm{HEA}}(\theta) = \prod_{d=1}^{L} \left[U_{\mathrm{ent}} \prod_{n=1}^{N} R_Z^{(n)}(\theta_{d,n}^Z) R_Y^{(n)}(\theta_{d,n}^Y) \right]. \tag{5.1}$$

Here, $R_\sigma^{(n)}(\theta) = e^{-i\theta\sigma/2}$ denotes a single-qubit rotation of angle θ around the axis associated to the Pauli matrix $\sigma \in \{Y, Z\}$ on qubit n. Note that the operators in the product in Eq. (5.2) are ordered by decreasing d, *i.e.* $\prod_{d=1}^{L} U_d = U_L \cdots U_1$. The unitary U_{ent} represents an entangling gate on the N qubit. In practice, the specific entangling gates that are used depend on the physical platform on which the VQA is run: one simply chooses those gates that can be realized with high fidelity. This choice explains the name of the ansatz. The block in brackets in Eq. (5.1) is sequentially applied L times, with L the circuit *depth*.

Since the HEA ansatz is problem-agnostic and intended to address a wide range of problems, it has to be able to cover a large part of the Hilbert space by varying the circuit parameters. The ability of an ansatz to generate states from the Hilbert space is generally characterized by its expressibility [78], a metric that compares the distribution of states obtained from sampling the parameters of a PQC to the ensemble of Haar-random states. As appealing as highly expressive ansätze may sound to explore vast Hilbert spaces, this feature has been shown to lead to issues that may hinder efficient optimization, such as the presence of barren plateaus [79], *i.e.* where the optimization landscape is flat. We will provide a brief discussion of this matter in Sect. 5.2. Unfortunately, it has been shown that HEA exhibits barren plateaus because of its high expressibility [79, 80].

Variational Hamiltonian Ansatz:

The variational Hamiltonian Ansatz (VHA) implements a parametrized version of the evolution of the problem Hamiltonian H_{p} that we are trying to solve [81]. The evolution operator $U = e^{-iH_{\text{p}}t}$ is generally approximated using the Suzuki-Trotter expansion [82]:

$$\begin{aligned} e^{-iH_p t} &= \lim_{n\to\infty} \left[\prod_k e^{-ic_k H_{\text{p}k} \frac{t}{n}} \right]^n \\ &= \left[\prod_k e^{-ic_k H_{\text{p}k} \frac{t}{n}} \right]^n + \mathcal{O}\left(\frac{t^2}{n}\right), \end{aligned} \tag{5.2}$$

where we decompose $H_{\text{p}} = \sum_k c_k H_{\text{p}k}$, with $\{H_{\text{p}k}\}$ a set of operators which do not commute with each other, and $\{c_k\}$ a set of coefficients. The decomposition is ideally chosen such that $e^{-ic_k H_{\text{p}k}}$ can be easily implemented with single- and two-qubit gates. Following this intuition, the VHA is designed by replacing the fixed coefficients c_k in Eq. (5.2) by variational parameters. The variational form of VHA thus writes:

$$U_{\text{VHA}}(\theta) = \prod_{d=1}^{L} \left[\prod_k e^{-i\theta_{d,k} H_{\text{p}k}} \right] \tag{5.3}$$

where L layers of the block in brackets are sequentially applied, and $\theta = \{\theta_{d,k}\}$ are the variational parameters (see Fig. 5.1 b)). We note that the Suzuki-Trotter expansion does not provide any prescription on the order in which different $H_{\mathrm{p}k}$ should appear in the decomposition of Eq. (5.3). In practice, this choice can lead to widely different convergence properties of the VQA.

Note that there exist many more ansätze. For instance, a well-known problem-inspired ansatz for chemistry is the Unitary Coupled Cluster (UCC). It is inspired from classical computation to describe the electronic structure of molecules in quantum chemistry [83]. This ansatz aims to capture the electron correlation effects that are missing in the Hartree-Fock approximation, and which are necessary to predict and accurately describe certain chemical properties. Another well-known ansatz for solving combinatorial optimization problems such as the Max-Cut problem is the Quantum Approximate Optimization Algorithm (QAOA) [84]. There also exist different methods for designing an ansatz instead of having a fixed-structure ansatz. For instance, the Adaptive Derivative-Assembled Problem-Tailored ansatz Variational Quantum Eigensolver (adapt-VQE) [76] introduced a strategy to dynamically build an ansatz based on the operations that have the biggest impact on the energy of the prepared state. We will describe this algorithm in more detail in Sect. 6.2. A review of existing PQCs can be found in Ref. [74].

5.2 The cost function

The *cost function*, also called *objective function*, is another central component of VQAs, as it is the function that will be optimized. In the context of VQAs, the cost function is formulated in terms of an operator, and is evaluated by measuring its expectation value. The cost function is very frequently the Hamiltonian of a given problem, for which we want to minimize to find the ground state energy. It can, however, happen that one does not want to optimize the energy of a Hamiltonian but rather wants to reach a target state, in other words, optimize the state fidelity. For instance, the excited state algorithms [85] and other algorithms in quantum machine learning [86] use the fidelity as part of the objective function. Other objective functions such as the Gibbs cost function [87] or the conditional value-at-risk function [88] have also shown to do at least as well for solving some combinatorial optimization problems. Regardless of the quantity being measured, in the end, it boils down to expressing an objective function $C(\theta)$ as the expectation value of an operator $\hat{O}$ on the prepared state:

$$C(\theta) = \mathrm{Tr}\left[\hat{O}U(\theta)\hat{\rho}_0 U^\dagger(\theta)\right] = \langle\hat{O}\rangle_\psi \tag{5.4}$$

where $U(\theta)$ is the unitary of the parametrized circuit, $\hat{\rho}_0 = |\psi_0\rangle\langle\psi_0|$ is the initial input state in the n-qubit Hilbert space, and $|\psi\rangle = U(\theta)|\psi_0\rangle$ is the output state of the PQC. If the cost function is the state fidelity, then $\hat{O}$ is the projector of the target state. When minimizing the energy, the cost function $\hat{O}$ is rather the problem Hamiltonian. The objective function may have to be decomposed in a linear combination of operators that can actually be measured experimentally.

VQAs are confronted by many challenges that need to be overcome, especially regarding the optimization task. For example, it has been shown that classically training VQAs is, under certain assumptions, a NP-hard problem and that the optimization will generally stay trapped in local minima [89]. Another important limitation is the possible existence of *barren plateaus* [79, 90]. A cost function is said to exhibit barren plateaus if, when training its parameter over the optimization landscape, the gradients of the cost function with respect to the variational parameters are exponentially suppressed (on average) with the number of qubits. Therefore, if a cost function exhibits barren plateaus it will be complicated to run an optimization to converge for more than a few qubits. Barren plateaus thus specially affect gradient-based optimization methods, but it has also been shown that even gradient-free methods can be sensitive to barren plateaus [91, 92]. Note that the cost function does not only depend on $\hat{O}$, but also depends on the prepared state since $C(\theta) = \langle\psi(\theta)|\hat{O}|\psi(\theta)\rangle$, its optimization is thus influenced by the structure of the PQC. In particular, it has been shown that the higher the expressibiliy of a circuit is, the most likely it will exhibit barren plateaus [90, 93]. A lot of efforts are currently made to alleviate these challenges [90, 94–96].

5.3 The measurement

After a quantum state $|\psi(\theta)\rangle$ has been prepared with a PQC, the cost function has to be evaluated. As mentioned in Sect. 5.2, the cost function is estimated by measuring, on the prepared state, the expectation value of the objective function described by an operator $\hat{O}$, as seen in Eq. (5.4). Ideally, one would directly measure $\hat{O}$ in its eigenbasis, however quantum computers are generally constrained to measure qubits in the Z-basis only. To measure $\hat{O}$, one can in principle perform a basis change by applying a unitary transformation (involving multiple qubits) on the state at the end of the PQC. However, this is usually a complicated and expensive operation. Moreover, $\hat{O}$ does not always have to be an observable, and in such a situation its measurement is more complicated. To simplify the discussion, in what follows we will focus on observables, and more specifically on the case where $\hat{O}$ is a Hamiltonian. Fortunately, there exists different methods that allow us to evaluate the

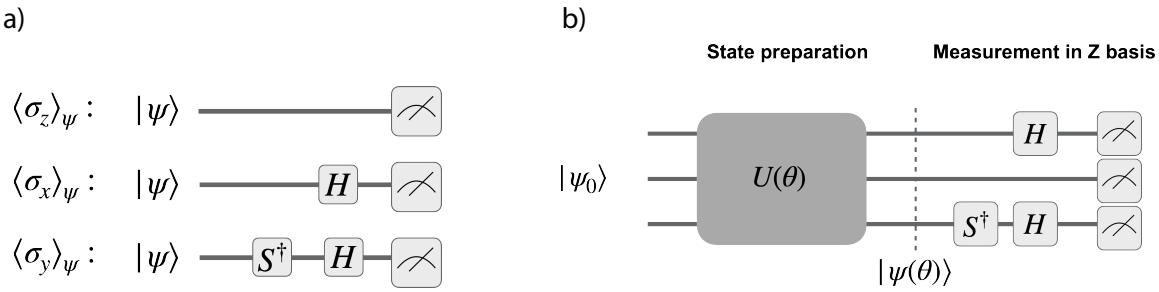


Figure 5.2 Measurement of Pauli matrices. a) Illustrates the gates that need to be applied on a state $|\psi\rangle$ to measure the Pauli matrices σ_x, σ_y and σ_z, in the Z basis. b) Illustrates an example to measure the Pauli string $\sigma_x \otimes \sigma_z \otimes \sigma_y$, on a state $|\psi(\theta)\rangle$ prepared by a PQC $U(\theta)$. H and S are respectively the Hadamard and the phase gates.

expectation value by breaking down the cost function observable into a sum of operators that can be measured more easily. In what follows we will explain one of the most commonly used measurement procedures which consists in decomposing $\hat{O}$ into Pauli strings.

Measurement of Pauli strings:

One of the simplest ways to estimate expectation values on current devices is to decompose the operator $\hat{O}$ associated to the cost function into a sum of Pauli strings, consisting in a tensor product of Pauli matrices, and then measure the individual Pauli strings. In fact, any operator can be uniquely decomposed into Pauli strings as $\hat{O} = \sum_{k=1}^{M} c_k \hat{P}_k$, with $c_k \in \mathbb{C}$ and M the number of Pauli strings $\hat{P}_k$ in the expansion of $\hat{O}$. The k-th Pauli string in the expansion is written as $\hat{P}_k = \otimes_{i=1}^{n} \hat{\sigma}_k^{(i)}$, with $\hat{\sigma}_k^{(i)} \in \{\hat{I}, \hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z\}$ the Pauli matrix that is applied on qubit i. The expectation value of $\hat{O}$ is then simply a linear combination of the expectation values of the individual Pauli strings, $\langle\hat{O}\rangle_\psi = \sum_{k=1}^{M} c_k \langle\hat{P}_k\rangle_\psi$.

As mentioned earlier, Pauli strings can easily be evaluated on NISQ devices, as the Pauli Z basis is their natural measurement basis. For a single qubit, measuring the expectation value of the X and Y Pauli matrices implies moving from the Z to the X/Y basis, which can easily be done using the identity:

$$\begin{aligned} \hat{\sigma}_x &= H\hat{\sigma}_z H, \\ \hat{\sigma}_y &= SH\hat{\sigma}_z HS^\dagger, \end{aligned} \tag{5.5}$$

with $H = (\hat{\sigma}_x + \hat{\sigma}_y)/\sqrt{2}$ the Hadamard gate and $S = \sqrt{\hat{\sigma}_z}$ the phase gate. These transformations are equivalent to applying H (resp. $HS^\dagger$) on the state to be measured, see Fig. 5.2 a). The expectation value of a Pauli operator is then estimated by repeating the measurements a

certain number of times and extracting the probability p of measuring the qubit in state $|0\rangle$. The measurement probability is related to the expectation value by $\langle\hat{\sigma}_z\rangle_\psi = 2p - 1$. For a n-qubit Pauli string $\hat{P}$, its expectation value is then simply $\langle\hat{P}\rangle_\psi = \prod_{k\in K}(2p(k) - 1)$, with K the ensemble of qubits on which Pauli operations (other than the identity) are applied, and $p(k)$ is the probability of measuring qubit k in state $|0\rangle$.

In principle, the number of Pauli strings to measure in $\hat{O}$ could scale as 4^n with n the number of qubits, *i.e.* the maximum number of Pauli strings. For physical problems, or by taking advantage of symmetries of the problem, the scaling is generally not as dramatic. For instance, for chemistry problems, the number of Pauli strings to measure in $\hat{O}$, typically scales as n^4 with the number of qubits. This scaling is due to the fact that the Hamiltonian is generally only written with two-body interaction terms. Nevertheless, it can quickly become prohibitively large. The measurement overhead has to be considered seriously, as it can become a limiting factor for quantum algorithms. Previous work by Wecker *et al.* [81] has shown, for example, that $\sim 10^{13}$ measurements would be required for each energy evaluation of the Fe_2S_2 molecule with a milliHartree precision. Approaches exist to reduce the measurement overhead, for example by simultaneously measuring Pauli strings that commute by measuring them in a common basis [97, 98]. Note though that moving to this shared basis comes at the cost of adding a unitary (involving multiple qubits) at the end of the PQC. Also, finding and optimally assembling the terms that commute, and that can hence be measured simultaneously, is equivalent to the minimum clique cover problem. This task is thus inherently NP-hard. Fortunately, good heuristics exist for assembling the operators [99–101].

There exist other alternatives that are more efficient than measuring Pauli strings. For instance, an approach based on a Cartan subalgebra has shown one of the largest reductions in the required number of measurements for chemistry problems [102]. Note, however, that most of the existing methods that enable the reduction of the number of measurements generally rely on knowledge of the underlying problem. For an overview of other methods, we refer the reader to Ref. [74].

Measurement precision:

Intuitively, one would like to minimize the number of terms in the decomposition of $\hat{O} = \sum_k h_k\hat{O}_k$ such that fewer operators have to be measured. However another factor worth considering comes into play: the measurement precision. In principle any partition of $\hat{O}$, even if non-optimal, will lead to the same final result. However, the number of measurements to reach a certain precision of the cost function can change significantly. The precision is

defined by the variance of the estimator of the cost function:

$$\epsilon^2 = \mathrm{Var}\left[\widehat{C(\theta)}\right] = \mathrm{Var}\left[\widehat{\langle \hat{O} \rangle_\psi}\right] \tag{5.6}$$

where $\widehat{\langle \hat{O} \rangle_\psi}$ denotes the estimator of $\langle \hat{O} \rangle_\psi$. In statistics, the *estimator* of an expectation value is an approximation of the actual expectation value for a finite number of samples. The estimator converges to the expectation value in the limit of an infinite number of measurements. Just as for the expectation value, one can write the estimator as a sum of estimators $\widehat{\langle \hat{O} \rangle_\psi} = \sum_k h_k \widehat{\langle \hat{O}_k \rangle_\psi}$. Using this expression, the variance of the measured cost function is:

$$\mathrm{Var}\left[\widehat{\langle \hat{O} \rangle_\psi}\right] = \sum_k h_k^2 \mathrm{Var}\left[\widehat{\langle \hat{O}_k \rangle_\psi}\right] + \sum_{j \neq k} h_i h_j \mathrm{Cov}\left[\widehat{\langle \hat{O}_j \rangle_\psi}, \widehat{\langle \hat{O}_k \rangle_\psi}\right] \tag{5.7}$$

where $\mathrm{Cov}[A, B]$ denotes the covariance of two random variables A and B, in our case the estimators of the expectation values of $\langle \hat{O}_j \rangle_\psi$ and $\langle \hat{O}_k \rangle_\psi$. If the terms $\hat{O}_k$ are sampled by independent measurements of the state $|\psi\rangle$, then the second term on the right side of Eq. (5.7) vanishes because the estimators are then uncorrelated. On the other hand, if the terms $\hat{O}_k$ are estimated using the same measurements, then the covariances remain and can ultimately increase or decrease the variance of $\widehat{\langle \hat{O} \rangle_\psi}$. A change in the variance affects the required number of measurements to reach a certain precision. In fact, if we consider that the terms $\hat{O}_k$ are normalized such that the eigenvalues are smaller than one, and if we assume that each term in $\hat{O}$ is measured the same number of times, then the maximum number of total measurements m required to reach a precision of ϵ is related by [81]:

$$m \leq \frac{M}{\epsilon^2} \sum_{k=1}^{M} h_j^2. \tag{5.8}$$

This shows that finding the best partitioning of $\hat{O}$ to minimize the number of measurements is far from trivial.

5.4 Classical optimization

VQAs are hybrid quantum-classical algorithms, which means that one part of the algorithm is performed on a quantum processor and another part on a classical computer. Subsections Sect. 5.1 to Sect. 5.3 have focused on the parts that are performed on a quantum processor. We have seen how parametrized circuits are designed, and have shown how the cost function can be defined and evaluated via measurements. Now remains the question of how to update the parameters of the PQC. There exists many methods to do this, and it generally involves classical resources which we briefly cover in this subsection.

The task of the classical optimizer consists in updating the circuit parameters to minimize a cost function which generally is a complicated non-linear function that can live in a very large parameter space. To achieve quantum advantage in the NISQ era, it is essential to minimize the number of evaluations of the cost function (done on a quantum processor). In fact, remember that the cost function can only be estimated up to a given precision that depends on the number of qubit measurements and on the operator being measured. This means that a single iteration of the classical optimization requires to prepare and measure the quantum state several times, which can take a long time. Moreover, the cost function evaluation can be affected by errors in the preparation of the state due to noise, qubit coherence times, *etc.* It is therefore crucial to choose an optimization method that minimizes the function evaluations and that is resilient to noise. The different optimizers can typically be divided in two categories: gradient-based and gradient-free optimization.

Among the gradient-based approaches, the simplest way of updating the parameter vector at a step $(n+1)$-th of the optimization is using the rule $\boldsymbol{\theta}^{(n+1)} = \boldsymbol{\theta}^{(n)} - \eta \nabla_{\boldsymbol{\theta}} C(\theta)$, where η is the learning rate of the optimizer and $C(\theta)$ the objective function to be minimized. The gradient with respect to each parameter can then be obtained, for example, using the finite difference method: $\partial_{\theta_i} C(\boldsymbol{\theta}) = (C(\boldsymbol{\theta} + h\boldsymbol{e_i}) - C(\boldsymbol{\theta} - h\boldsymbol{e_i}))/2h$, with h a small coefficient and $\boldsymbol{e_i}$ the unit vector with 1 at the i-th element such that the parameter θ_i is varied by h. This method approximates best the gradient when h tends to zero, which can only be done at the cost of increasing the number of measurements. There exist many variations of this method, and other gradient-based methods, like stochastic gradient descent, Hessian-aided gradient descent, quantum natural gradient descent, quantum imaginary time evolution among others [74]. Methods also exist where the gradient is evaluated directly on the quantum processor by calculating the analytical gradient [103–106].

Among the gradient-free approaches, the first experimental implementation of a VQA used the Nelder-Mead method [107]. More advantageous methods have been used since,

that are more resilient to statistical noise like the *simultaneous perturbation stochastic approximation* (SPSA) algorithm which require fewer evaluations of the cost function regardless of the dimensions of the parameter space [77, 108]. Optimization using reinforcement learning and neural networks have also been explored [109–112]. Surrogate model-based optimization, often used when a function evaluation is costly, have in the last years shown to be promising for VQAs with limited quantum resources [113]. For a more complete review of optimizers commonly used for PQC optimization see Ref. [74]. Generally speaking, the efficiency of classical optimization routines highly depends on the problem that is being solved, on the objective function, on the initial parameters, on the form of the PQC, on state preparation errors *etc.* Work has been done to compare different approaches in specific contexts [99, 104, 114].

Note that optimization-free quantum algorithms are also being developed to address the problem of classical optimization in VQAs. Those algorithms update the parameters of the PQC layer-by-layer using a feedback loop conditioned on the qubit measurement of the previous layer. The parameter update follows a feedback law generally inspired from quantum control theory and which under some assumptions guarantees to the algorithm to asymptotically converge to the solution of the problem that is being solved. Feedback-based quantum algorithm also referred to as FQAs have for example been used to solve combinatorial problem [115–117] and finding the ground state of the Fermi-Hubbard model [118]. A caveat is that those methods currently require very deep quantum circuits in order to reach high accuracy.

In our work, we are interested in improving the performance of the QOCA ansatz, which will be described in Sect. 6.3, compared to existing ansätze for solving the Fermi-Hubbard model. We did not insist on finding the best possible optimizer, and rather chose to compare the ansätze for similar numerical conditions. In the following sections, the optimizer we use to test the performance of different ansätze is the *Constrained Optimization BY Linear Approximations* COBYLA. It is a derivative-free optimization method that approximates the objective function and constraints using a linear model to explore the parameter space for the optimal solution. The search is done in a trust region and the distance by which the search can move from the current parameters in each iteration is set at the beginning of the optimization but adapts dynamically based on the performance of previous iterations.

5.5 Fermi-Hubbard model

5.5.1 The model

The Fermi-Hubbard model (FHM) is widely used in condensed matter physics to study strongly correlated materials. It is one of the simplest model that exists to describe interacting electrons on a lattice, it can for example be used to describe high-critical temperature superconductors. The FHM is of major interest to explain and predict important properties of matter such as conductor-isolator phase transitions. Because of its importance in physics, we use this model to benchmark the performance of our ansatz compared to other standard VQAs.

The FHM describes the interaction of fermions on an S-site lattice, where each site can be occupied by at most two fermions, provided that they satisfy Pauli's exclusion principle. In this model, the Coulomb repulsion between electrons that are not on the same sites are neglected. Translated mathematically, the Fermi-Hubbard Hamiltonian takes the form:

$$\hat{H}_{\text{FHM}} = \underbrace{-t \sum_{\substack{\langle i,j \rangle \\ \sigma=\{\uparrow,\downarrow\}}} \hat{a}^\dagger_{i\sigma}\hat{a}_{j\sigma}}_{\equiv T} + \underbrace{U\sum_{i=1}^{S}\hat{n}_{i\uparrow}\hat{n}_{i\downarrow} - \mu \sum_{\substack{i=\{1\cdots S\} \\ \sigma=\{\uparrow,\downarrow\}}} \hat{n}_{i\sigma}}_{\equiv V}, \tag{5.9}$$

where i, j are the indices of the lattice sites, and σ represents the spin degree of freedom of the electrons. The operator $\hat{a}^{(\dagger)}_{i\sigma}$ is the fermionic annihilation (creation) operator removing (creating) an electron with spin σ on site j, while the occupation operator of the spin orbital $i\sigma$ is defined as $\hat{n}_{i\sigma} = \hat{a}^\dagger_{i\sigma}\hat{a}_{i\sigma}$.

The first term of the Hamiltonian, T, typically referred to as the kinetic term, corresponds to the hopping of an electron (more generally a fermion) between sites i and j at a rate $-t$. An electron with spin σ is removed from site j and created on site j, conserving the spin and the total number of particles. Notice also that the summation is done over nearest-neighbor sites $\langle i, j\rangle$ only, limiting hopping to neighboring sites.

The second term corresponds to the on-site Coulomb repulsion when two electrons (with opposite spins) are on the same site. This term hinders double occupancy with an energy penalty U. The last term is proportional to the chemical potential and sets the number of particles. The two last terms are generally referred to as the potential terms, V.

Unfortunately, the FHM is only exactly solvable for a 1D chain [119]. For higher-

dimensional lattices, this model has been extensively studied by numerical means on classical computers [120]. However, solving the FHM becomes exponentially complex as sites are added to the model. The largest lattices that can be solved exactly via exact diagonalization contain around 24 sites [121]. Naturally, the complexity of solving the FHM also depends on the parameter regime. For example, the kinetic and potential terms are both individually solvable as they are diagonal in the momentum and position space, respectively, and their ground states are made of delocalized plane waves, resp. localized wave functions. This makes the intermediate coupling regime, where $U/t \sim 1$, a particularly interesting case to study as the two processes compete. In this context, neither of the two parts of the Hamiltonian can be treated perturbatively, as they now contribute equally to the Hamiltonian, making it a particularly complicated regime to solve.

5.5.2 Mapping $\hat{H}_{\mathsf{FHM}}$ to a qubit Hamiltonian

The Fermi-Hubbard Hamiltonian presented in Eq. (5.9) is formulated using fermionic operators. To be able to solve the FHM on a quantum computer, the Hamiltonian first has to be reformulated in terms of the qubit operators, usually the Pauli matrices. The two main ingredients for mapping the problem Hamiltonian to a qubit Hamiltonian consist in:

1. mapping the quantum states,
2. rebuilding the fermionic algebra in terms of the Pauli operators.

Multiple methods exist for mapping a Hamiltonian such as the Jordan-Wigner (JW) mapping [122], the parity mapping [123] and the Bravyi-Kitaev mapping [124]. The JW mapping is one of the simplest and intuitive transformation one can think of to perform this task. Through this mapping, the occupation of a site is directly mapped on two qubits as $(0, \uparrow, \downarrow, \uparrow\downarrow) \rightarrow (|00\rangle, |01\rangle, |10\rangle, |11\rangle)$. An S-site lattice is thus mapped onto $2S$ qubits, where we arrange the qubits as $|q_1 \cdots q_{2S}\rangle = |p_1^{\uparrow} \cdots p_S^{\uparrow}; p_1^{\downarrow} \cdots p_S^{\downarrow}\rangle$ with p_j^{σ} representing a particle of spin $\sigma \in \{\uparrow, \downarrow\}$ on site $j \in \{1, \cdots, S\}$. The fermionic operators are mapped to Pauli operators as:

$$\begin{aligned} \hat{a}_{j\sigma} &\rightarrow \left(\hat{X}_{j\sigma} + i\hat{Y}_{j\sigma}\right) \bigotimes_{k=1}^{j-1} \hat{Z}_{k\sigma} \\ \hat{a}_{j\sigma}^{\dagger} &\rightarrow \left(\hat{X}_{j\sigma} - i\hat{Y}_{j\sigma}\right) \bigotimes_{k=1}^{j-1} \hat{Z}_{k\sigma} \end{aligned} \tag{5.10}$$

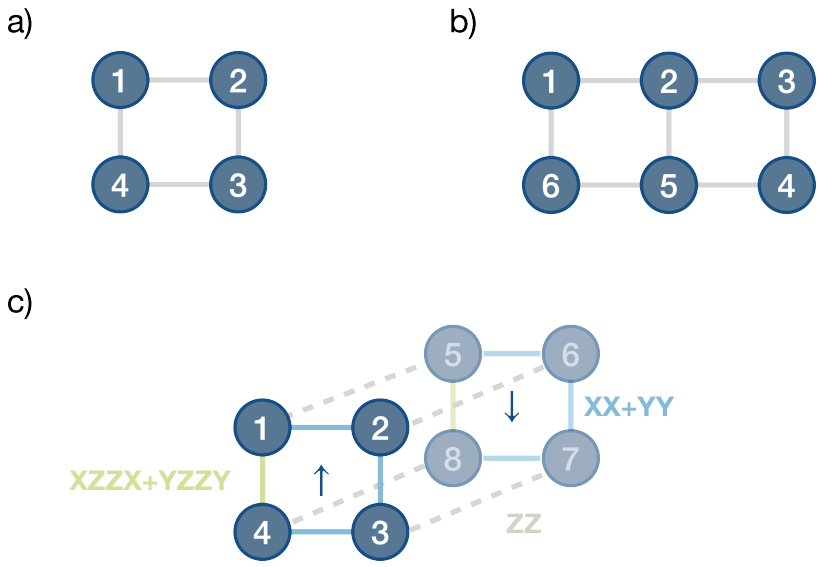


Figure 5.3 FHM lattice sites. Illustration of a: a) 2×2 lattice and b) 2×3 lattice, with labeling of the sites. c) Diagram of the qubits representing a 2×2 lattice. The 4 first (last) qubits represent spin up (down) particles. The lines connecting the qubits represent the qubit interaction terms appearing in the Fermi-Hubbard Hamiltonian: ZZ (dashed gray), $XX + YY$ (blue), $X_1Z_2Z_3X_4 + Y_1Z_2Z_3Y_4$ and $X_5Z_6Z_7X_8 + Y_5Z_6Z_7Y_8$ (green).

for all $\sigma \in \{\uparrow, \downarrow\}$ and $j \in \{1, 2, \cdots, S\}$, and with $\hat{X}$, $\hat{Y}$, $\hat{Z}$ the Pauli matrices. Note from Sect. 5.5.2 that as the number of lattice sites increases, the number of qubit operations in the Pauli strings that are not the identity increases linearly, ultimately leading to more two-qubit gates to simulate them. There exist more efficient ways of mapping the problem to a qubit Hamiltonian, as for instance the Bravyi-Kitaev mapping which involves fewer qubits to simulate a fermionic operator than the JW transformations.

In the following chapters, we benchmark different VQAs on a 2×2 and a 2×3 lattice, in the intermediate coupling regime with $t = 1$, $U = 4$ and at half-filling, *i.e.* $\mu = U/2$. Organizing the lattice sites as in Fig. 5.3, the qubit Hamiltonian takes the form:

$$\hat{H}^{qb}_{\text{FHM}} = -\frac{t}{2} \sum_{\substack{\langle i,j \rangle, j<i \\ \sigma=\{\uparrow\downarrow\}}} \left(\hat{X}_{i\sigma}\hat{X}_{j\sigma} + \hat{Y}_{i\sigma}\hat{Y}_{j\sigma} \right) \bigotimes_{k=i+1}^{j-1} \hat{Z}_{k\sigma} + \frac{U}{4} \sum_{i=1}^{S} \hat{Z}_{i\uparrow}\hat{Z}_{i\downarrow}. \tag{5.11}$$

Note that the single $\hat{Z}$s that appear from the onsite interaction and the chemical potential cancel each other at half-filling.

Chapter 6

Adaptive approaches for QOCA

The work presented in this chapter greatly benefitted from discussions with Maxime Dion from the Algo Lab of the Institut Quantique at the Université de Sherbrooke and from Alexandre Choquette-Poitevin presently at IBM Quantum.

6.1 Quantum Optimal Control inspired Ansatz

The *Quantum Optimal Control inspired Ansatz* (QOCA) has shown promising results for the Fermi-Hubbard Model (FHM) and the water molecule [75]. The QOCA ansatz is an adapted form of the VHA, exploiting its problem-inspired nature to confine the variational search within a specific symmetry sector of the Hilbert space. Conversely, QOCA introduces additional terms to the ansatz that break the problem's symmetry, aiming to discover potential shortcuts within the Hilbert space.

As seen in Sect. 5.1, the VHA implements a parametrized version of the trotterized time evolution of a given problem Hamiltonian, H_{problem}. The QOCA ansatz draws inspiration from quantum optimal control by constructing the ansatz not only from the problem Hamiltonian but from an extended Hamiltonian that incorporates a set of drive terms:

$$\hat{H}_{\text{QOCA}}(t) = \hat{H}_{\text{problem}} + \sum_k c_k(t)\hat{H}_{\text{d}k}, \tag{6.1}$$

with $\{\hat{H}_{\text{d}k}\}$ a set of drive terms and $c_k(t)$ the time-dependent amplitudes of the control drives. The drive terms are chosen such that $[H_{\text{problem}}, \hat{H}_{\text{d}k}] \neq 0\ \forall k$. The drive terms here

play the role of the control Hamiltonian in quantum optimal control. The state-preparation unitary of QOCA is then simply the trotterization of Eq. (6.1), which reads:

$$\hat{U}_{\mathrm{QOCA}}\left(\vec{\theta},\vec{\delta}\right)=\prod_{l=1}^{L}\left[\prod_{j}e^{i\theta_{j,l}\hat{H}_{j}}\prod_{k}e^{i\delta_{k,l}\hat{H}_{\mathrm{d}k}}\right], \tag{6.2}$$

where the problem Hamiltonian is decomposed as $\hat{H}_{\mathrm{problem}}=\sum_j \hat{H}_j$, and $\theta_{j,l}$ is the parameter of the term $\hat{H}_j$ in the l-th layer. Similarly, $\delta_{k,l}$ is the parameter of the drive term $\hat{H}_{\mathrm{d}k}$ in the l-th layer, which represents the amplitude of the control drives. This unitary implements L blocks of the QOCA ansatz, where L is the circuit depth. One such block is represented at the top of Fig. 6.1 a).

The presence of symmetry-breaking terms in the drive part allows the VQA to search for the ground state by temporarily exiting the intended symmetry sector of the problem Hamiltonian. This ability to venture outside the symmetry sector has the advantage that the optimizer can find shortcuts in the Hilbert space, thereby accelerating the convergence process. Moreover, QOCA is for this reason also expected to be resilient to symmetry-breaking errors. In fact, for the VHA ansatz, if such errors occur, the inability of the ansatz to return to the symmetry sector, prevents the algorithm from finding the ground state.

In the paper introducing QOCA [75], this ansatz is used to solve the FHM for an $(n\times m)$-lattice (with $S=n\times m$ sites), using linear drives for fermions:

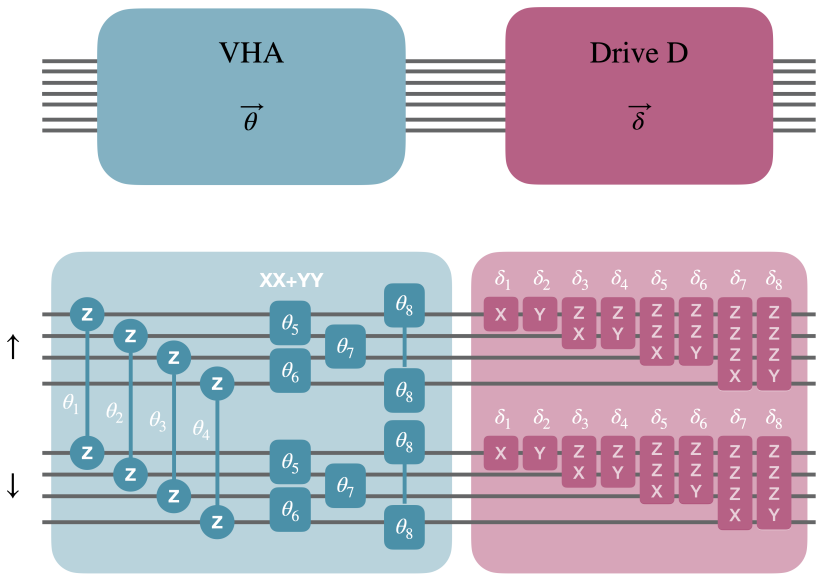
$$\hat{H}_{\mathrm{d1}}=\sum_{j=1}^{S}(\hat{a}_j^{\dagger}+\hat{a}_j)\quad\text{and}\quad\hat{H}_{\mathrm{d2}}=\sum_{j=1}^{S}i(\hat{a}_j^{\dagger}-\hat{a}_j), \tag{6.3}$$

which are separately applied on both spins. This drive has no physical meaning but is chosen because it does not conserve the number of particles, and thus also does not commute with H_{problem}. By applying the Jordan-Wigner transformation to map the fermionic linear drive to a qubit Hamiltonian, one obtains the following drive terms:

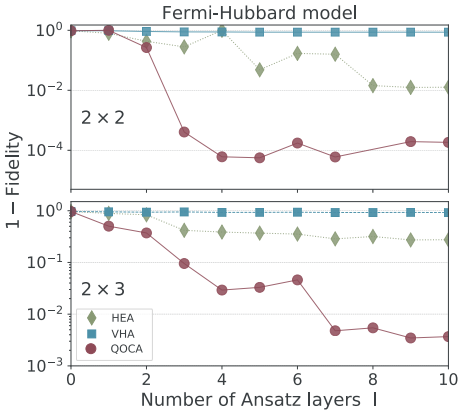
$$\hat{H}_{\mathrm{d1}}=\sum_{j=1}^{S}\hat{X}^{(j)}\bigotimes_{k<j}\hat{Z}^{(k)}\qquad,\qquad\hat{H}_{\mathrm{d2}}=\sum_{j=1}^{S}\hat{Y}^{(j)}\bigotimes_{k<j}\hat{Z}^{(k)}, \tag{6.4}$$

with $\hat{A}^{(j)}\in\{\hat{X},\hat{Y},\hat{Z}\}$ the Pauli matrices applied on qubit j. In other words, the drive terms are $\hat{X}$ and $\hat{Y}$ applied on each qubit j with a chain of $\hat{Z}$s on the qubits with $k<j$.

a)



b)



c)

2×2 (8 qubits)

Ansatz	Max Fid.	L	n_θ/L	n_{CX}/L
HEA	0.9876	9	16	7
VHA	0.1343	8	8	56
QOCA	0.999	9	12	40

2×3 (12 qubits)

Ansatz	Max Fid.	L	n_θ/L	n_{CX}/L
HEA	0.7276	10	24	11
VHA	0.0804	10	13	116
QOCA	0.9965	9	25	172

Figure 6.1 QOCA ansatz a) One layer of the QOCA ansatz represented by a parametrized quantum circuit. The ansatz is divided in two parts: one block of VHA (blue) parametrized by $\vec{\theta}$ and one block comprising the drive terms (red) parametrized by $\vec{\delta}$. The circuit below is the QOCA ansatz for a 2×2 FHM. It contains a block of VHA involving ZZ and $XX + YY$ Pauli operations on the given qubits, which respectively correspond to the potential and kinetic part of the Hamiltonian in Eq. (5.11). The second block (red) comprises the drive terms involving single- and multi-qubit Pauli operations of Eq. (6.4). The 4 top (bottom) qubits represent the 4 sites with spin up (down). b) State infidelities with respect to the target state as a function of the number of layers L of the HEA, the VHA and the QOCA ansatz. Top (resp. bottom) panel is for a 2×2 (resp. 2×3) lattice. The initial state is $|+\rangle^{\otimes N}$ in both cases. c) Maximum fidelity with respect to the ground state of the FHM, attained for L ansatz layers, each requiring a number n_θ/L of variational parameters and n_{CX}/L CNOTs per layer, assuming an all-to-all connectivity. b) and c) are adapted from Ref. [75].

For a 2×2 and a 2×3 lattice, Ref. [75] variationally optimizes the energy and reports the infidelity of the final state compared to the exact ground state, as shown in Fig. 6.1 b). The simulations were conducted for the Hamiltonian in Eq. (5.11), with $\mu = U/2$ and $U/t = 4$, and with the state initialized in $|+\rangle^{\otimes n}$ which is already in the symmetry sector of the ground state of the Hamiltonian. The QOCA ansatz is compared to the VHA and the HEA ansatz in terms of fidelity, circuit depth, number of parameters and number of CNOTs. The table in Fig. 6.1 c) shows that while HEA improves the state fidelity up to 10^{-2} with 9 circuit layers, VHA, on the other hand, fails to improve at all the infidelity even though the ansatz possesses all the necessary terms to explore the symmetry sector of the problem Hamiltonian. In contrast, QOCA does significantly better than both VHA and HEA in terms of fidelity, and this for shallow circuits, *i.e.* for a small number of layers than HEA. In terms of the number of variational parameters per layer, QOCA is comparable to HEA. However, one of the major drawbacks of QOCA, especially for NISQ devices, is that it requires many CNOT gates in each layer.

The QOCA ansatz shows better convergence in terms of speed and accuracy than standard approaches, but a certain number of questions remain. Are all the drive terms necessary or can the ansatz be made shallower? Is there a better way of ordering the operators in the ansatz? Also, the results obtained in Ref. [75] by optimizing the parameters of all the layers at once but, do better optimization strategies exist? In our work we have investigated those questions. In particular, we draw inspiration from an adaptive method to build ansätze, called *adapt-VQE* [76], which we briefly describe in the following section.

6.2 Adapt-VQE

The *Adaptive Derivative-Assembled Problem-Tailored ansatz Variational Quantum Eigensolver* (adapt-VQE) was introduced in Ref. [76] as a method to dynamically design an ansatz. It was initially used to solve chemistry problems and has shown to outperform the well-known UCCSD ansatz in terms of accuracy, of the number of variational parameters and of the number of two-qubit gates. The general idea of Adapt-VQE is to build and optimize an ansatz as follows (see Fig. 6.2):

Step 0. Start with an empty circuit corresponding to the unitary $\hat{U} = I$. Choose a set of operators $\mathcal{O} = \left\{\hat{O}_k \,|\, \hat{O}_k = -\hat{O}_k^\dagger\right\}$, called the *operator pool*, which will be used to iteratively build the ansatz. The initial state that is applied on the circuit is noted $|\psi_0\rangle$, and should ideally be an educated guess, such as the Hartree-Fock state for chemistry problems.

Step 1. **I.** For each $\hat{O}_k \in \mathcal{O}$, evaluate the energy gradient of the state $|\psi(\theta_1)\rangle = e^{\theta_1 \hat{O}_k}|\psi_0\rangle$, with respect to the parameter θ_1, if the operator $\hat{O}_k$ were to be added to the ansatz. The energy gradient is the derivative of the expectation value of the problem Hamiltonian $\hat{H}_\mathrm{p}$, which can be estimated by measuring the commutator of $\hat{H}_\mathrm{p}$ with $\hat{O}_k$ on the state $|\psi_0\rangle$ since:

$$\left.\frac{\partial E}{\partial \theta_1}\right|_{\theta_1=0} = \left.\frac{\partial}{\partial \theta_1}\langle\psi_0|e^{-\theta_1\hat{O}_k}\hat{H}_\mathrm{p}e^{\theta_1\hat{O}_k}|\psi_0\rangle\right|_{\theta_1=0} = \langle\psi_0|[\hat{H}_\mathrm{p},\hat{O}_k]|\psi_0\rangle. \tag{6.5}$$

The gradient is generally estimated at $\theta_1 = 0$. Finally, the operator in the pool which maximizes the norm of the energy gradient in Eq. (6.5), $\hat{A}_1 = \max_{\hat{O}_k\in\mathcal{O}}\left|\frac{\partial}{\partial\theta_1}\langle H_p\rangle\right|$, is then added to the PQC, leading to the unitary $\hat{U}(\theta_1) = e^{\theta_1\hat{A}_1}$. In principle, all the commutators $[\hat{H}_\mathrm{p},\hat{O}_k]$ could be measured in parallel on different quantum processors.

II. The parameter θ_1 is then classically optimized to minimize the energy $\langle\hat{H}_\mathrm{p}\rangle$ on the prepared state $\hat{U}(\theta_1)|\psi_0\rangle$ using a VQE subroutine. The parameter θ_1 is generally initialized to zero but note that it does not have to be the case. We refer to the optimized parameter as $\theta_1^{opt} = \min_{\theta_1}\langle H_p\rangle$ and define $|\psi_1\rangle$ the state after this first optimization.

Step 2. **I.** For each $\hat{O}_k \in \mathcal{O}$, estimate the energy gradient of the state $|\psi(\theta_2)\rangle = e^{\theta_2\hat{O}_k}|\psi_1\rangle$, with respect to the parameter θ_2. This is equivalent to measuring $\langle\psi_1|[\hat{H}_\mathrm{p},\hat{O}_k]|\psi_1\rangle$. The pool operator which maximizes the energy gradient norm, $\hat{A}_2 = \max_{\hat{O}_k\in\mathcal{O}}\left|\frac{\partial}{\partial\theta_2}\langle H_p\rangle\right|$, is then added to the circuit, leading to the updated ansatz $\hat{U}(\theta_1,\theta_2) = e^{\theta_2\hat{A}_2}e^{\theta_1\hat{A}_1}$.

II. The parameters (θ_1, θ_2) are then classically optimized to minimize the energy $\langle \hat{H}_\mathrm{p} \rangle$ on the prepared state $\hat{U}(\theta_1, \theta_2)|\psi_0\rangle$ using a VQE subroutine. The parameters are generally initialized to $(\theta_1, \theta_2) = (\theta_1^{opt}, 0)$. We refer to the optimized parameters as $(\theta_1^{opt}, \theta_2^{opt}) = \min_{\theta_1,\theta_2} \langle H_p \rangle$, where θ_1 is also reoptimized, and define $|\psi_2\rangle$ the state after this second step. Note that θ_1^{opt} is redefined after the optimization, therefore it does not correspond to the initialization value.

Step d. The PQC is adaptively built by repeating this process up to a given layer which depends on the chosen convergence criterion. The convergence criterion can, for instance, be a threshold on the energy or on the values of the measured commutators. The adapt-VQE procedure can be written recurrently at step d as:

I. For each $\hat{O}_k \in \mathcal{O}$, estimate the energy gradient of the state $|\psi(\theta_d)\rangle = e^{\theta_d \hat{O}_k}|\psi_{d-1}\rangle$, with respect to the parameter θ_d, where $|\psi_{d-1}\rangle$ is the optimized state found at the previous step. This is equivalent to measuring $\langle \psi_{d-1}|[\hat{H}_\mathrm{p}, \hat{O}_k]|\psi_{d-1}\rangle$. The operator in the pool which maximizes the energy gradient norm on $|\psi(\theta_d)$ is then added to the circuit, leading to $\hat{U}(\vec{\theta}) = e^{\theta_d \hat{A}_d} e^{\theta_{d-1}\hat{A}_{d-1}} \cdots e^{\theta_1 \hat{A}_1}$, with $\vec{\theta}$ the parameters written in vector form.

II. The parameters $\vec{\theta}$ are then classically optimized to minimize the energy using a VQE subroutine. The newly added parameter is generally initialized to zero while the other parameters are set to the values found from the optimization at step $(d-1)$. Note that different optimization strategies of the parameters are possible.

Dynamically growing the ansatz through iterative selection of the pool operator that has the largest impact on the energy offers the advantage of generating what appears to be a nearly optimal ansatz, with a minimum number of operators and parameters. Nevertheless, this advantage comes at the cost of increasing the number of measurements. In fact, compared to fixed-structure ansätze like the HEA and VHA, adapt-VQE requires evaluating the commutators of the problem Hamiltonian with each pool operator. Since this task has to be performed at each iteration, it ultimately results in a substantial increase of the measurement overhead. Efforts are made to reduce the number of measurements [125, 126]. Another advantage of adapt-VQE is that it is expected to be more resilient to barren plateaus and less susceptible to local minima [127]. Several variations of adapt-VQE have been developed since its invention to further improve its performance in solving specific problems, ranging from chemistry to combinatorial optimization [125, 126, 128–130].

Originally, the operator pool consisted of fermionic operators present in chemistry Hamiltonians [76], however, alternative pools have been utilized since. For instance, Ref. [128]

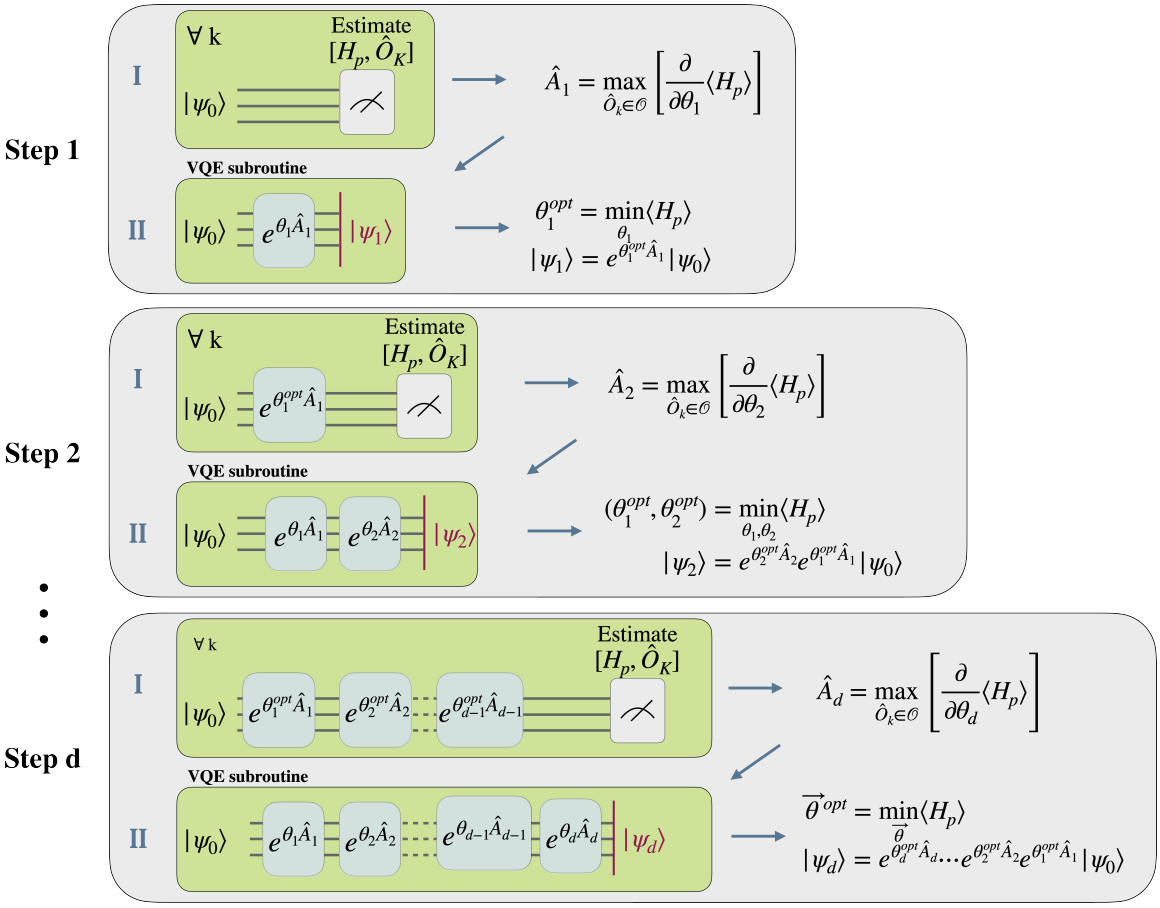


Figure 6.2 Procedure for implementing adapt-VQE Each step is decomposed into two parts I and II. In part I, all the commutators $[H_p, \hat{O}_k]$ with $\hat{O}_k \in \mathcal{O}$ are evaluated on a prepared state. The pool operator that has the largest variation in the state's energy is then added to the circuit in part II. All the circuit parameters are then optimized in part II using a VQE subroutine. **Step 1**, all the commutators are evaluated on the initial state $|\psi_0\rangle$. The operator that maximizes the expectation value of the commutator, $\hat{A}_1$, is added to the circuit. The parameter θ_1 is optimized, leading to the state $|\psi_1\rangle$. **Step 2**, all the commutators are evaluated on $|\psi_1\rangle$. The pool operator that maximizes the expectation value of the commutator, $\hat{A}_2$, is added to the circuit with a parametrization θ_2. All the parameters are then optimized, leading to the state $|\psi_2\rangle$. **Step d** The PQC is iteratively grown and optimized up to a certain layer d.

suggested a more hardware efficient operator pool made of the individual Pauli strings that appear from the fermionic operators mapping. As this pool constitutes a decomposition of the fermionic operator pool, it encompasses a significantly larger number of terms. However, it was shown that the number of operators within the pool could be further reduced to exhibit linear scaling with the number of qubits. Ref. [128] demonstrated that although the parameter count required for convergence increased, this particular pool substantially

reduced both the circuit depth and the two-qubit gates count.

Another strategy to further reduce the circuit depth consists in adding multiple operators in a single adapt-VQE iteration [130]. In fact, it can happen that multiple pool operators acting on separate qubits exhibit large energy gradients. This strategy has been shown to lead to shallower circuits while keeping a similar number of two-qubit gates for chemistry problems.

Adapt-VQE has also been implemented alongside the *Quantum Approximate Optimization Algorithm* (QAOA) to address combinatorial problems [84, 129]. Similarly to the QOCA, the ansatz in QAOA consists of alternating layers of two unitaries generated by a cost Hamiltonian and a mixer Hamiltonian. The cost Hamiltonian is specific to the optimization problem in hand, while different choices are possible for the mixer Hamiltonians. In Ref. [129], the mixer part of the QAOA ansatz is constructed adaptively built by adding operators from a pool, employing a strategy similar to the adapt-VQE approach. In this particular case, the operator pools used comprise single- and/or two-qubit gates, making it more hardware-efficient. This combination of adapt-VQE and QAOA, named adapt-QAOA, has shown to converge faster than QAOA, while reducing the required number of CNOT gates and parameters.

6.3 Adaptive approaches for QOCA

One of the motivations that led us to introduce an adaptive approach for QOCA is related to the structure of the drive part. More precisely, we wanted to understand if all the drive terms were necessary for QOCA to converge well or if it could be made shallower, and understand if a better way of ordering the operators in the ansatz existed.

6.3.1 Adapt-drive QOCA

To investigate the question related to the structure of the drive part of QOCA, we combine QOCA with an adaptive method similar to adapt-VQE. As described in Sect. 6.1, a layer of QOCA is made of a block of VHA and a block composed of drive terms. In adapt-drive QOCA, we instead iteratively build the drive part by adding one operator at a time, from a predefined pool, as it is done in adapt-VQE. More precisely, adapt-drive QOCA follows the subsequent step:

1. Predefine a pool of operators $\mathcal{O}$ which will be used to build the ansatz.
2. Prepare an initial state $|\psi_0\rangle$, and start with an empty circuit.
3. Add a block of the VHA to the empty PQC.
4. Perform a VQE subroutine to optimize all the parameters in the PQC.
5. a) For each $\hat{O}_k \in \mathcal{O}$, estimate the expectation value of the commutator of the problem Hamiltonian with $\hat{O}_k$ on the state prepared by the PQC. The pool operator maximizing the expectation value of the commutator is then added to the PQC with its own parameter.

   b) Perform a VQE subroutine to optimize all the parameters in the ansatz. As an initial guess, we use the parameters found from the preceding optimization and assign a zero value to the newly added parameter.
6. Repeat step 5 to iteratively grow the PQC until the chosen convergence condition is attained. The condition can for instance be a threshold on the measured energy or on the values of the commutators.
7. Go back to step 3 until a second chosen convergence criterion is not yet reached. The condition can for instance be a threshold on the measured energy or the number of operators added to the circuit.

Just as QOCA, this algorithm can in principle be utilized for any cost minimization problem. In the following, we benchmark this method on the Fermi-Hubbard model, allowing us to compare to known results for QOCA [75].

6.3.2 Adapt-QOCA

Let's now consider a second strategy to adaptively build an ansatz, which we call *adapt-QOCA*. The idea here is to adaptively grow the ansatz by adding to the PQC the pool operator that has the biggest impact on the energy, just as in adapt-VQE. However, a key distinction lies in the fact that when the commutator of the problem with all the pool operators falls below a certain threshold, a layer of VHA is added to the circuit. This strategy was inspired from the observation that adapt-VQE[1] failed to find the ground state of the FHM and by noticing that adding a layer of VHA in adapt-drive QOCA generally kickstarted the convergence of the state.

[1] Using a pool containing all the terms present in QOCA.

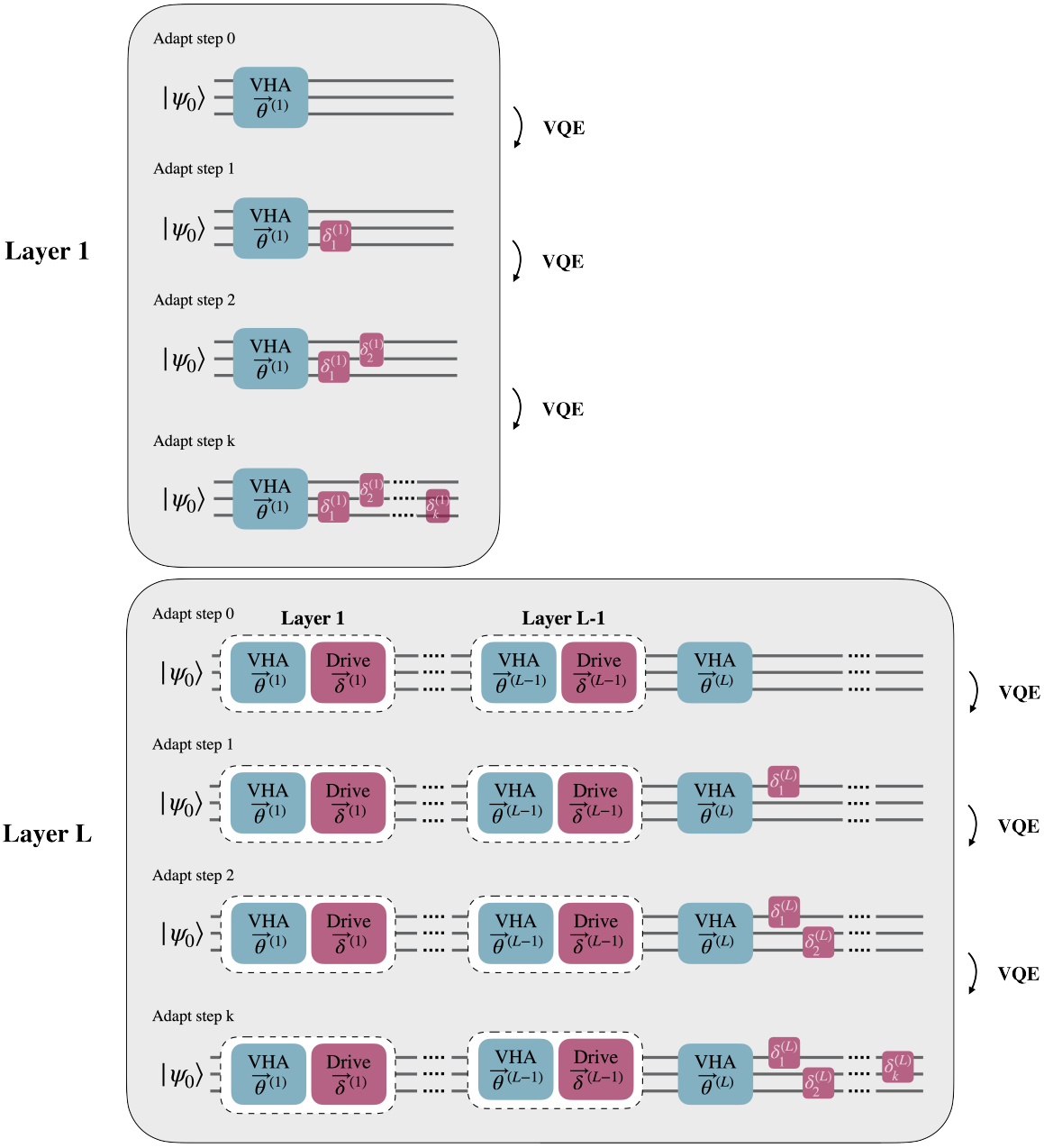


Figure 6.3 Procedure for implementing adapt-drive QOCA Each layer of adapt-QOCA is iteratively built and optimized. In the first layer, a first block of VHA (blue) is added to the circuit with parameters $\vec{\theta}^{(1)}$, where the upper index indicates the layer of adapt-drive QOCA. The circuit is then optimized via a VQE subroutine. The ansatz is then iteratively grown using the adapt-VQE approach up to a step k defined by the user. Each of the added operators has a parameter $\delta_j^{(1)}$ for $j \in \{0, \cdots, k\}$, where the upper index indicates the adapt-drive QOCA layer and the lower index the step in the adapt-VQE procedure. All the circuit parameters are optimized via a VQE subroutine before adding a new operator to the PQC. An L-layer adapt-drive QOCA circuit is built layer by layer repeating the procedure in layer 1.

Chapter 7

Benchmark adapt-QOCA

The work presented in this chapter greatly benefitted from discussions with Maxime Dion from the Algo Lab of the Institut Quantique at the Université de Sherbrooke and from Alexandre Choquette-Poitevin presently at IBM Quantum.

In this chapter, we benchmark the ansätze presented in Sect. 6.3 on the Fermi-Hubbard Model (FHM) for a 2×2 and a 2×3 lattice, in the intermediate coupling regime with $t = 1$, $U = 4$ and at half-filling, *i.e.* $\mu = U/2$, see Eq. (5.11).

In Sect. 7.1 we define various possible pools and establish the context in which the simulations were done, including optimization conditions, initial state *etc.* Sect. 7.2 subsequently investigates the convergence of the states prepared by adapt-drive QOCA and adapt-QOCA, comparing their energy and fidelity to the original QOCA. Finally, in Sect. 7.3, we examine the pool operators selected by both approaches and study their performance in terms of circuit depth, before summarizing the results in Sect. 7.3.2.

7.1 Setting up the simulations

7.1.1 Operator pools

Various options of pools can be considered:

1. *Drive pool*: This pool is made of the drive terms of the original QOCA ansatz as described in Eq. (6.4). For clarity, we only show the pools for the 2×2 FHM:

$$\begin{aligned}\mathcal{O}_{\text{drive}} = &\{(X_1 + X_5), (Z_1X_2 + Z_5X_6), (Z_1Z_2X_3 + Z_5Z_6X_7), (Z_1Z_2Z_3X_4 + Z_5Z_6Z_7X_8)\} \\ &\cup \{(Y_1 + Y_5), (Z_1Y_2 + Z_5Y_6), (Z_1Z_2Y_3 + Z_5Z_6Y_7), (Z_1Z_2Z_3Y_4 + Z_5Z_6Z_7Y_8)\}\,.\end{aligned} \tag{7.1}$$

   The qubits are ordered as $q_1q_2q_3q_4q_5q_6q_7q_8 = p_1^\uparrow p_2^\uparrow p_3^\uparrow p_4^\uparrow p_1^\downarrow p_2^\downarrow p_3^\downarrow p_4^\downarrow$ with p_j^σ representing a particle of spin $\sigma \in \{\uparrow, \downarrow\}$ on site $j \in \{1, 2, 3, 4\}$, see the ordering in Fig. 6.1. The qubits corresponding to a spin up (down) particle are represented in blue (gray).

2. *VHA pool*: This pool is made of the terms present in VHA, and therefore also present in QOCA. For the FHM, these terms are obtained from Eq. (5.11) and are represented in Fig. 5.3. For a 2×2 lattice, the VHA pool is:

$$\begin{aligned}\mathcal{O}_{\text{VHA}} = &\{Z_1Z_5, Z_2Z_6, Z_3Z_7, Z_4Z_8\} \\ &\cup \{(X_1X_2 + Y_1Y_2) + (X_5X_6 + Y_5Y_6)\} \\ &\cup \{(X_3X_4 + Y_3Y_4) + (X_7X_8 + Y_7Y_8)\} \\ &\cup \{(X_2X_3 + Y_2Y_3) + (X_6X_7 + Y_6Y_7)\} \\ &\cup \{(X_1Z_2Z_3X_4 + Y_1Z_2Z_3Y_4) + (X_5Z_6Z_7X_8 + Y_5Z_6Z_7Y_8)\}\,.\end{aligned} \tag{7.2}$$

3. *QOCA pool*: Contains all the terms present in QOCA: $\mathcal{O}_{\text{QOCA}} = \mathcal{O}_{\text{VHA}} \cup \mathcal{O}_{\text{drive}}$. For a 2×2 lattice, the QOCA pool possesses 16 operators with 8 terms in $\mathcal{O}_{\text{VHA}}$ and 8 terms in $\mathcal{O}_{\text{drive}}$.

Notice that the above pool operators respect the spin symmetry by simultaneously performing the same operations on qubits representing fermions with opposite spins. For example, when an operation X is applied to qubit 1 (site 1 spin $\uparrow$) it is also applied on qubit 5 (site 1 spin $\downarrow$), and they will both share the same parameter in the PQC. As suggested in Ref. [131], alternative drive Hamiltonians could be used in QOCA. In adapt-QOCA and adapt-drive QOCA, they would lead to different drive pools.

In the context of adapt-drive QOCA, the drive part is built adaptively. A choice has therefore to be made for the number of operators N_{ops} the drive part of a single layer of adapt-drive QOCA should contain. To guide this choice, it is worth noting that the drive part in the original QOCA ansatz comprises 8 (resp. 13) terms for the 2×2 (resp. 2×3) FHM. We explore the consequence of using different values of N_{ops}, ranging from 1 to 10 (up to 12 for the 2×3 FHM). Furthermore, we have established a criterion: if, during the adaptive growth of the drive part (as described in steps 5 to 6 of the procedure in Sect. 6.3.1) the commutators fall below a given threshold, we do not add any additional operators to the drive part. Instead, we proceed to the next layer of adapt-drive QOCA by adding a VHA block. In other words, an adapt-drive QOCA layer will possess at most N_{ops} operators in the drive part.

7.1.2 Thresholds

In the context of adapt-drive QOCA, we need to establish two convergence criteria at step 6 and 7 of the procedure outlined in Sect. 6.3. At step 6.b., the criteria for adding an additional operator to the drive part are as follows: if N_{ops} operators have been added to the drive part, or if the largest expectation value of the commutators of the problem Hamiltonian with the pool operators fall below 10^{-6}, then the algorithm moves to step 8. In the case of step 7, we consider that convergence has been reached when 5 layers (resp. 10 layers) of adapt-QOCA have been added to the PQC for the FHM 2×2 (resp. 2×3).

For adapt-QOCA, we consider that if the largest expectation value of the commutators of the problem Hamiltonian with the pool operators falls below 10^{-6}, then a block of VHA is added to the PQC. The algorithm stops when N_{ops} operators have been added to the PQC.

7.1.3 Initial state

Following Ref. [75]We initialize the quantum state in $|+\rangle^{\otimes n}$ with $n = 8$ or 12 for a 2×2 and 2×3 lattice, respectively. This state is half-filled, ensuring that it starts in the correct particle-number symmetry sector of the Hilbert space, with the average site occupation $\langle \hat{N} \rangle = S$, where $\hat{N} = \sum_{i=1,\sigma=\{\uparrow,\downarrow\}}^{S} \hat{n}_{i\sigma}$. We know from Fig. 7.1 b) that even though this initial state lies in the right symmetry sector VHA and HEA converge poorly, while QOCA exhibits better convergence. Better initial guesses are possible, but QOCA and its variations are designed to converge to the ground state even when starting from an initial state that has minimal overlaps with the target state.

7.1.4 Circuit simulation and optimization

The circuit simulations were performed using the opensource Qiskit package [132]. The circuits are built by transpiling the exponential of the Pauli operators in the ansatz into Hadamard, CNOTs, R_Z rotation, S and $S^\dagger$ phase gates using built-in Qiskit functions. A house made code is used to build the ansätze and perform the VQE subroutines.

The results presented for adapt-drive QOCA and adapt-QOCA in the following sections, are obtained using the classical optimizer COBYLA. We set the tolerance of the final accuracy at $tol = 10^{-9}$ and allow COBYLA to evaluate the cost function at most $maxiter = 10^6$ times for the 2×2 FHM, and 10^5 times for a 2×3 lattice. Note that we have reduced the maximum number of function evaluations for the latter case as simulations are very long for PQCs with many parameters. Based on a study of the energy landscape with respect to the variation of ansatz parameters, we estimated reasonable values for the distance in parameter space by which the search can move from the current parameters. This distance is denoted *rhobeg* in Qiskit. We obtained the best results for values between $rhobeg = \pi/1.05$ and $\pi/0.95$. All the other optimizer options were set to the default values.

The results obtained in Sect. 7.2 for the QOCA ansatz are obtained using COBYLA with $tol = 10^{-9}$, $maxiter = 10^6$ and $rhobeg = \pi/8$ (resp. $\pi/0.95$) for the FHM 2×2 (resp. 2×3). For an L-layer QOCA ansatz, we initialize all the parameters at zero and subsequently optimize them all at once as it is done in Ref. [75]. Note that using these optimization parameters the results are improved compared to Ref. [75]. It is worth mentioning that we also experimented with a layer-wise optimization, but it did not yield as favorable results.

Note that COBYLA is only efficient for a moderate number of variational parameters, around hundreds. To simulate larger lattices, other optimizers could be more advantageous.

7.2 Fermi-Hubbard model

In this section, we examine how the energy and fidelity of the state prepared by the adapt-drive QOCA and adapt-QOCA, approach the ground state when adding operators to the PQC. In particular, we will focus on the FHM.

7.2.1 FHM 2×2

Fig. 7.1 a) shows the error in the energy compared to the exact energy for the 2×2 FHM, as a function of the number of operators/parameters in the circuit, while Fig. 7.1 b) shows the state infidelity compared to the ground state. As a reference, the results obtained for the original QOCA are represented by the gray line, with dots at multiples of 16 on the x-axis corresponding to the number of parameters in a single QOCA layer. Note that the energies and the infidelities achieved here surpass those reported in Ref. [75] as we have further explored the optimization parameters. The results of adapt-drive QOCA with the drive pool $\mathcal{O}_{\text{drive}}$ are represented in solid colored lines for different values of N_{ops}. Each dot corresponds to the results of a VQE subroutine, *i.e.* each time a pool operator or a block of VHA is added to the circuit. For a given line, the large gaps that appear between dots (along the x-axis) occur when a layer of VHA is added to the circuit, which involves adding 8 parameters at once. Additionally, as previously mentioned in the adapt-drive QOCA protocol (step 6), it is important to note that there are instances where, for a given layer in adapt-QOCA and for a given N_{ops}, the values of the commutators $\left([H_{\text{p}}, \hat{O}_k]\, \forall \hat{O}_k \in \mathcal{O}_{\text{drive}}\right)$ fall below the chosen threshold before the number of operators added to the drive part reaches N_{ops}. This is visible, for example in Fig. 7.1 a) for $N_{\text{ops}} = 8$ at 30 on the x-axis, where only 3 operators are added to the drive part. Note also that the results are identical for $N_{\text{ops}} > 8$ as no more than 8 operators are added in the drive part. A first observation is that regardless of the number of operators added to the drive part using the adapt-VQE approach, adapt-drive QOCA always does significantly better than QOCA in terms of accuracy and number of parameters. In fact, the error in the energy and the infidelity already fall below 10^{-4} for the same number of operators than in a 2-layer QOCA ansatz. In terms of highest accuracy, $N_{\text{ops}} = 1$ does better than the others. However, if we consider that a precision of 10^{-7} in the energy is sufficient, then $N_{\text{ops}} = 2$, 7 or 8 do better in terms of the number of operators in the ansatz. In short, it is difficult to make a definitive judgment on which N_{ops} performs best because the results vary a lot depending on the parameters of the optimizer. Note, however, that regardless of the optimizer parameters that we tested, adapt-drive QOCA still systematically did better than QOCA. Ultimately, these results show that all the drive operators in QOCA are not necessary and that a better ordering exists. In Sect. 7.3, we will study in more detail the occurrence of the operators chosen by adapt-drive QOCA, and will compare the convergence of the algorithm in terms of the circuit depth and number of CNOTs in the ansatz.

Knowing that adapt-drive QOCA does better than QOCA, it raises the question whether directly using adapt-VQE could do even better than adapt-QOCA. To answer this question,

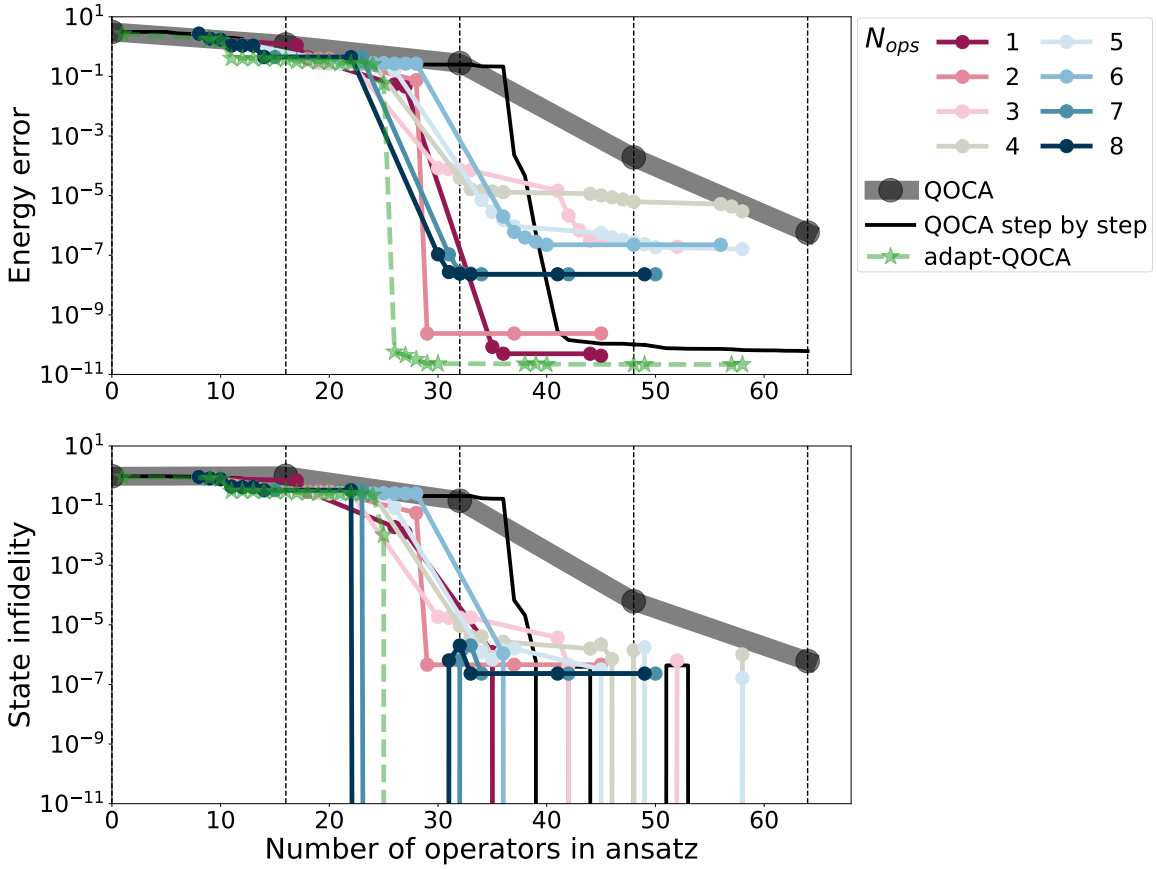


Figure 7.1 Ansätze performance for the 2x2 FHM. a) Error in the energy compared to the exact energy (in units of the hopping term t) and b) State infidelity compared to the ground state, as a function of the number of operators in the ansatz. As a reference, the vertical dashed black lines correspond to the number of operators in layers 1 to 4 of the QOCA ansatz. The gray dotted line corresponds to the results obtained using the QOCA ansatz with simultaneous optimization of all the parameters, while the solid black line is the result of iterative parameter optimization. The colored dotted lines represent the results obtained with adapt-drive QOCA with the drive pool, for multiple N_{ops} ranging from 1 to 8. Results from adapt-QOCA are depicted as a green dashed line with star markers.

we have implemented adapt-VQE using the QOCA pool $\mathcal{O}_{\text{QOCA}}$. In our exploration of optimization parameters, and even by relaxing the convergence criteria, adapt-VQE builds an ansatz that is unable to reduce the energy error. In fact, at every iteration the commutator values never exceed 10^{-6}, except for the first operator chosen, and adapt-VQE remained stuck in a local minimum from the start. As a result, and based on the observation that adapt-drive QOCA sometimes drastically improves in accuracy after a block of VHA is added to the PQC, we have tested another variation of adapt-drive QOCA which we refer to as adapt-QOCA (see Sect. 6.3.2). The idea is to add a layer of VHA to the circuit

when the values of the commutators fall below a chosen threshold, here 10^{-6}. The results of adapt-QOCA are plotted as a green dashed line with star markers. We notice that adapt-VQE adds a block of VHA at the beginning, providing a first small kick that proves to be necessary for convergence. For the FHM 2×2, adapt-QOCA performs better in terms of accuracy than adapt-drive QOCA. The advantage of adapt-drive QOCA over adapt-QOCA is that it requires fewer evaluations of the commutators and therefore fewer measurements. For the FHM 2×2, and as we will see with the FHM 2×3, adapt-QOCA converges better than QOCA and seems to be a lower limit to adapt-drive QOCA.

The parameter optimization strategies differ between QOCA and the adapt-VQE-type approaches. In fact, QOCA optimizes all its parameters at once, whereas in the adaptive approaches, parameters are optimized sequentially each time a new operator is added to the circuit. To verify that the performance disparities between QOCA and both adapt-drive QOCA and adapt-QOCA do not solely stem from their optimization strategies, we compare the results in the case where the parameters of QOCA are first all set to zero, and are then iteratively optimized one after another. We refer to this strategy as "QOCA step by step", which is depicted in Fig. 7.1 in a solid black line. We see that the optimization strategy plays an important role in the convergence of the ansatz. In fact, the step by step optimization shows significantly better results than the original optimization strategy. When prioritizing the highest accuracy, QOCA step by step outperforms adapt-drive QOCA for most choices of N_{ops}. However, if we consider that a precision of 10^{-5} in the energy is sufficient, then adapt-drive QOCA and adapt-QOCA still exhibit better or comparable performance in terms of the number of operators.

Other choices of pools:

We also tested adapt-drive QOCA using the QOCA pool, $\mathcal{O}_{\text{QOCA}}$, to give it more flexibility and ensure fairness in comparison to adapt-QOCA. We have depicted the results in Fig. 7.2 for N_{ops} ranging from 1 to 5. For $N_{\text{ops}} > 5$ the results for the QOCA pool become identical. As one could expect, the QOCA pool generally does better than the drive pool as it allows for more flexibility in the choice of operators. In Sect. 7.3, we will study the occurrence of the operators in the circuit grown by adapt-drive QOCA and adapt-QOCA. This will provide insights into which operators play an important role for the convergence of the prepared state. Building on our observation in Sect. 7.3.2 that only few drive terms were select from $\mathcal{O}_{\text{QOCA}}$, we also tested adapt-drive QOCA and adapt-QOCA with the VHA pool $\mathcal{O}_{\text{VHA}}$. However, it became evident quite rapidly that this pool did not lead to any improvement in the energy error for the latter ansätze, highlighting the critical role played by the drive terms in quick starting the convergence process.

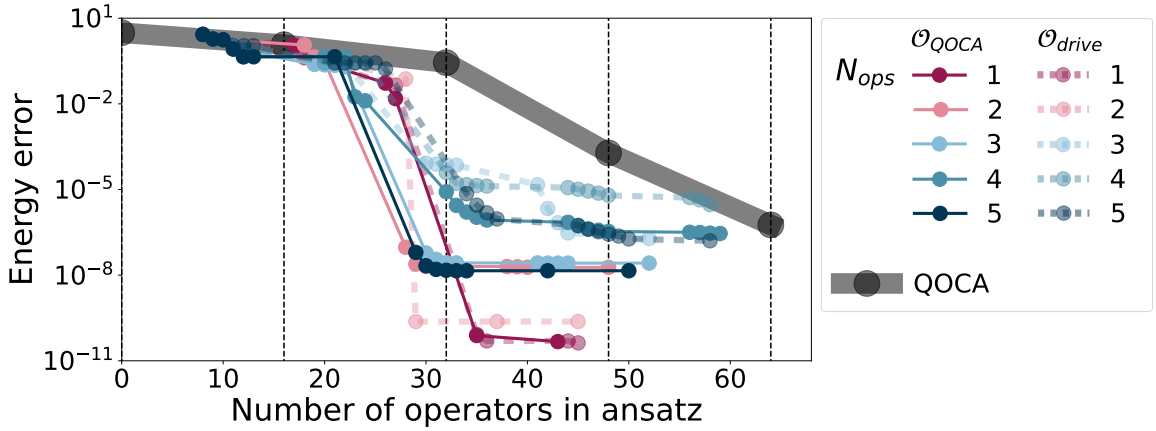


Figure 7.2 Adapt-drive QOCA with different pools. Error in the energy compared to the exact energy (in units of the hopping term t) as a function of the number of operators in the ansatz. As a reference, the vertical dashed black lines correspond to the number of operators in layers 1 to 4 of the QOCA ansatz. The gray dotted line corresponds to the results obtained using the QOCA ansatz with simultaneous optimization of all the parameters. The solid colored lines represent the results obtained with adapt-QOCA with the pool $\mathcal{O}_{\text{drive}}$, and in transparent dashed line with the pool $\mathcal{O}_{\text{QOCA}}$, for multiple N_{ops} ranging from 1 to 5.

7.2.2 FHM 2×3

We now illustrate the above results on a larger lattice, the 2×3 FHM. In Fig. 7.3, we present the same quantities as in Fig. 7.1, with the distinction that the QOCA pool is used for adapt-drive QOCA instead of the drive pool. This choice is motivated by the fact that the QOCA pool generally yielded better results than the drive pool for the FHM 2×2. It is important to note that we have, nonetheless, tested adapt-drive QOCA with the drive pool and have drawn similar conclusions as for the FHM 2×2. Furthermore, for the sake of clarity we plot the results for odd values of N_{ops} only, but note that even values of N_{ops} exhibit the same trend.

Similar to the findings for the FHM 2×2, we observe that adapt-drive QOCA and adapt-QOCA outperform QOCA in terms of energy and fidelity. Additionally, the energy and infidelity of the state prepared by adapt-QOCA once again serves as a lower bound for adapt-drive QOCA. We also tested adapt-VQE with the QOCA pool and, as previously, found improvement in the energy error and the state fidelity.

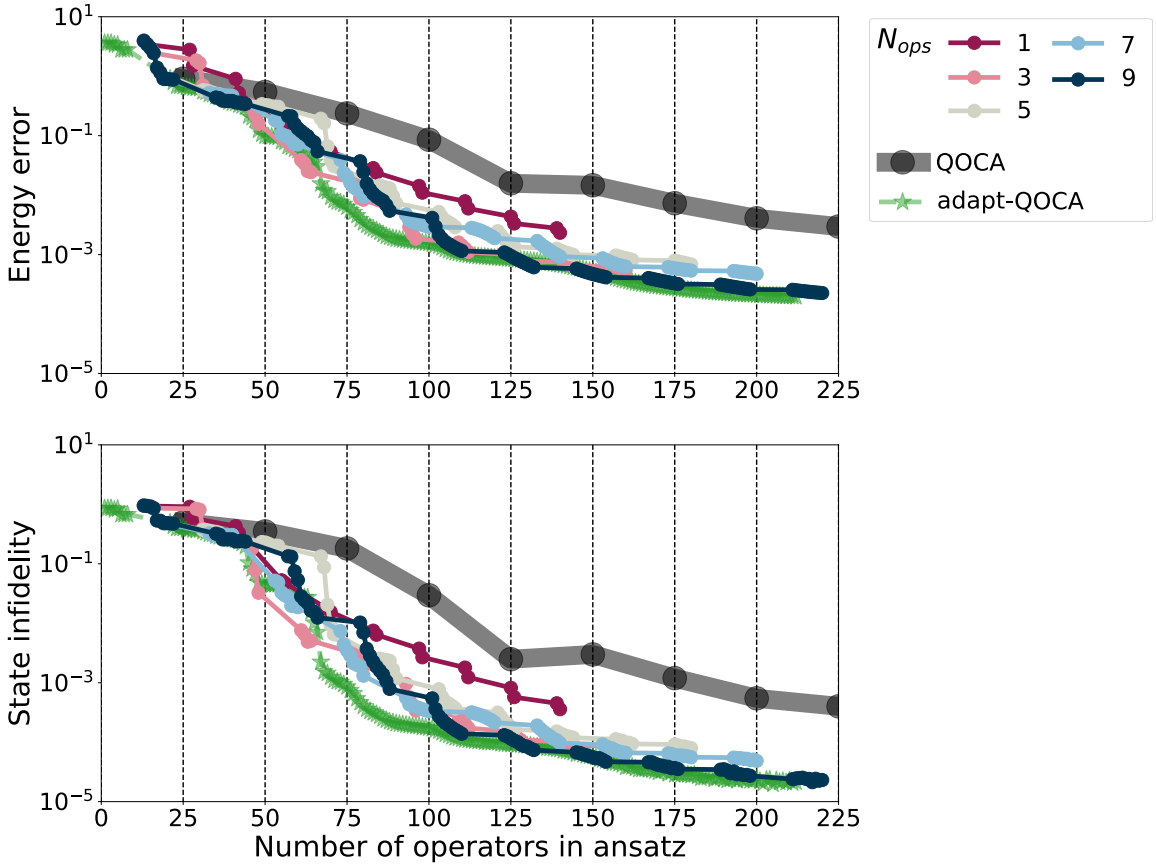


Figure 7.3 Ansätze performance for the 2×3 FHM. Same figure as in Fig. 7.1 for the 2×3 FHM, and with the difference that adapt-drive QOCA was grown with the QOCA pool, and only for odd values of N_{ops} ranging from 1 to 9.

7.3 Chosen operators and gate count

7.3.1 Number of CNOTs

As demonstrated in the preceding section, both adapt-QOCA and adapt-drive QOCA achieve higher accuracy while requiring fewer operators compared to QOCA. Nevertheless, having fewer operators does not necessarily mean fewer CNOTs in the ansatz or smaller circuit depth. In fact, if the operators selected from the pool require more two-qubit gates, particularly the operators involving many qubits, these ansätze can require more CNOTs. In Fig. 7.4, we show how the total number of CNOTs increases as the energy error improves. The CNOT gate count is obtained by first transpiling the ansätze from the exponentiation of the Pauli operations to Hadamard, CNOTs, R_Z rotations and, S and $S^\dagger$ phase gates using built-in Qiskit functions, assuming all-to-all qubit connectivity. Note that Qiskit automatically

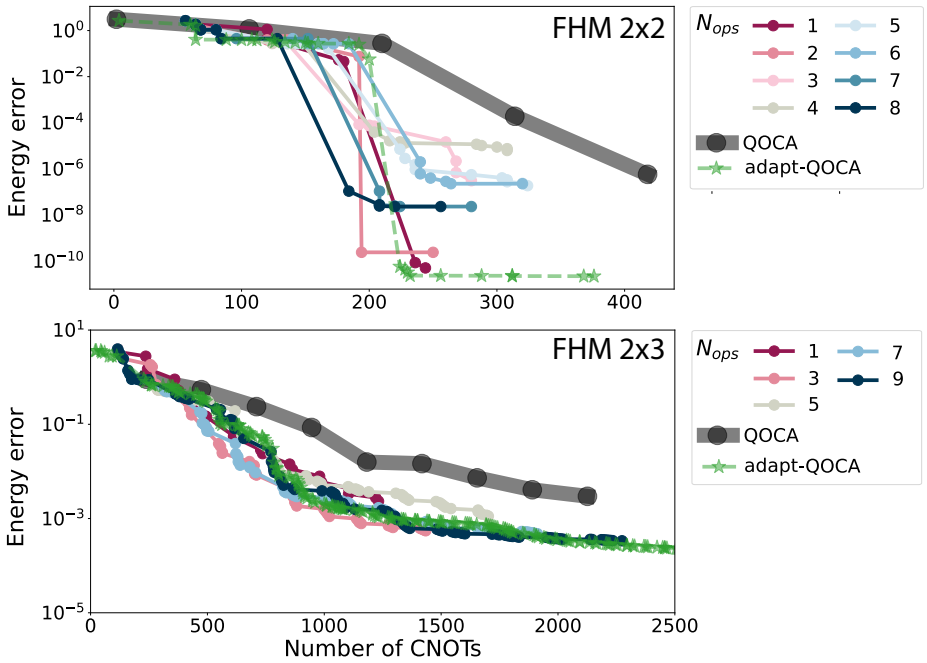


Figure 7.4 Gate count and circuit depth. Error in the energy compared to the exact energy (in units of the hopping term t) as a function of the number of CNOTS in the ansatz for: a) FHM 2×2, b) FHM 2×3. The gray dotted line corresponds to the results obtained using the QOCA ansatz with simultaneous optimization of all the parameters. The solid colored lines represent the results obtained with adapt-drive QOCA with the pool $\mathcal{O}_{\text{drive}}$ for multiple N_{ops} ranging from 1 to 8. Results from adapt-QOCA are depicted as a green dashed line with star markers.

simplifies the circuit by performing back-to-back gate cancellation when possible. For both the FHM 2×2 and 2×3, we see that adapt-QOCA and adapt-drive QOCA require substantially fewer CNOT gates than QOCA to reach a same precision, adapt-QOCA and adapt-drive QOCA are comparable. Nevertheless, we observe that there exist values of N_{ops} for which adapt-drive QOCA reaches the same energy error with fewer two-qubit gates than adapt-QOCA. This is particularly visible for the FHM 2×2 (resp. 2×3) with $N_{\text{ops}} = 2$ and 8 (resp. $N_{\text{ops}} = 3$).

7.3.2 Chosen operators

Our investigation now delves into the occurrence of the operators within the circuits grown by adapt-drive QOCA and adapt-QOCA. This study aims to provide insights into which operators contribute to the convergence of the prepared state and to understand the reasons that cause the adaptive approaches to outperform QOCA. In Fig. 7.5, the bar charts

represent, for the FHM 2×2, the number of times the operators in $\mathcal{O}_{\text{QOCA}}$ appear in the ansatz built by adapt-QOCA (left) and in adapt-drive QOCA for $N_{ops} = 7$ (right). For clarity, the labels over the bar charts are associated to the operators in the legend. Going from the left to the right, the 8 first operators, are the operators present in $\mathcal{O}_{\text{VHA}}$ and are ordered in the way they would appear in VHA. The 8 following operators are the drive operators in $\mathcal{O}_{\text{drive}}$. The tables below the bar charts show the order in which the operators are added to the PQCs following the arrow that represents the operator ordering in the ansatz. Each block of VHA that is added can be visually seen in the table as a series of light blue cells appearing in a diagonal shape. Fig. 7.6 shows the same figure for the FHM 2×3, and with the difference that $N_{ops} = 9$ for adapt-drive QOCA. Note that $\mathcal{O}_{\text{VHA}}$ now contains 13 operators (13 first from the left to right in the table) and 17 terms in the drive pool $\mathcal{O}_{\text{drive}}$.

As a first observation, we see that for both adapt-drive QOCA and adapt-QOCA some pool operators are never selected. This is the case for the FHM 2×2 and 2×3, and can be expected to extend to larger lattices. The operators in question are:

$$\left\{ \sum_{\sigma=\{\uparrow,\downarrow\}} \bigotimes_{j=1}^{i-1} Z_{j\sigma} Y_{i\sigma} \mid \forall \text{ odd } i \in \{1, 2, \cdots, S\} \right\} \tag{7.3}$$

and

$$\left\{ \sum_{\sigma=\{\uparrow,\downarrow\}} \bigotimes_{j=1}^{i-1} Z_{j\sigma} X_{i\sigma} \mid \forall \text{ even } i \in \{1, 2, \cdots, S\} \right\} \tag{7.4}$$

with $A_{j\sigma} \in \{X, Y, Z\}$ the Pauli matrix applied on the qubit representing a particle on site j and spin σ. Note that even though the operators $ZZXIII$ and $ZZZZXI$ were not selected in the particular case of Fig. 7.6, *i.e.* with $N_{ops} = 9$ for adapt-drive QOCA, these operators appear for other choices of N_{ops}.

We observe that very few drive operators are added to the ansätze. For the FHM 2×2, the drive terms represent 5 to 25% of the circuit operators, while for the FHM 2×3 they only represent only 5 to 10%. In the particular case of Fig. 7.6, the drive terms only appear at the beginning of the ansatz, but note that for some choices of N_{ops} they can appear later. Even though few drive terms are selected, those appear to be necessary for the convergence of the algorithm. In fact, as we tested adapt-QOCA and adapt-drive QOCA with the VHA pool, we noticed that the algorithm did not convergence at all. The drive terms seem to give a kick for the ansatz to be able to start converging.

Turning our attention to adapt-QOCA, we observe that in the early phase of the PQC, only drive operators have non-zero commutators and are therefore added to the circuit before

all the commutators fall below 10^{-6}. A block of VHA is thus added rather soon in the circuit for both FHM lattices. This layer of VHA is of major importance, as without it, the procedure would reduce to performing the original adapt-VQE using the QOCA pool which, as mentioned earlier, fails at improving the energy and state fidelity. Constructing the ansatz solely by selecting the next best operator appears to have its limitations in this context. Adding an appropriately chosen combination of operators instead of a single operator at a time could maybe help to see further away and help better organize the operators in the ansatz. Note that we do not have a good criterion for what an appropriate combination would be, even though VHA seemed to be sufficient to outperform the original adapt-VQE and adapt-QOCA in this context. Finally, we observe that adding a single layer of VHA at the beginning is sufficient to induce the converge. Note that in Fig. 7.5, the following layers of VHA are introduced after adapt-QOCA has already reached an energy error below 10^{-10}.

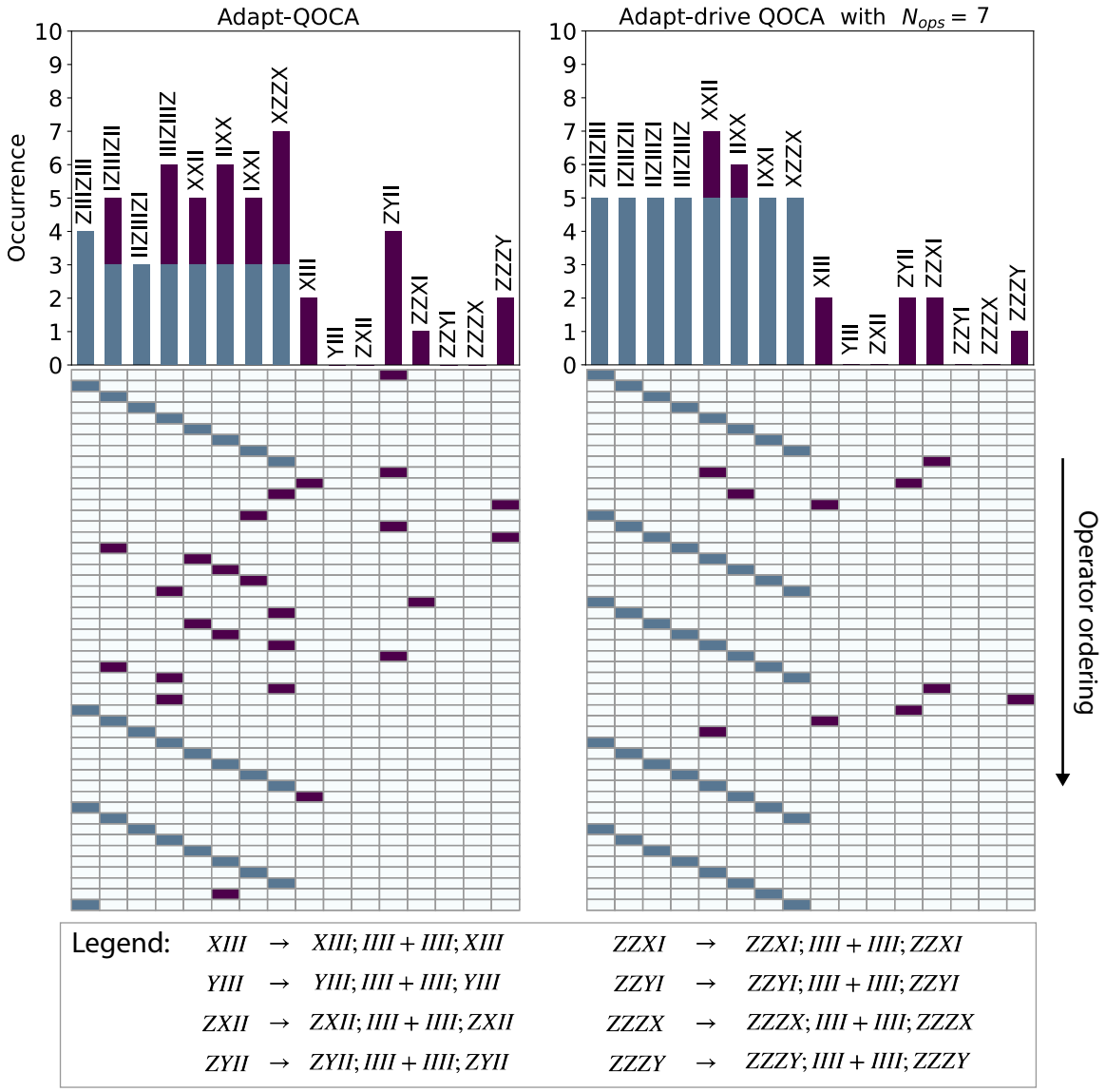


Figure 7.5 Operator occurrence FHM 2×2. The bar charts represent the number of times the operators in $\mathcal{O}_{\text{QOCA}}$ appear in the circuit built by adapt-QOCA (left) and in adapt-drive QOCA for $N_{ops} = 7$ (right). The labels on the bar chart are defined in the legend. The 8 first operators (from the left to the right) are the operators of $\mathcal{O}_{\text{VHA}}$ and the 8 following are the operators in $\mathcal{O}_{\text{drive}}$. The tables below the bar charts show the order in which the operators are added to the PQC along the *operator ordering* arrow. When a block of VHA is added to the ansatz the 8 first operators on the left are added one after another (in blue). The number of operators added from a VHA block is also indicated in blue in the bar chart.

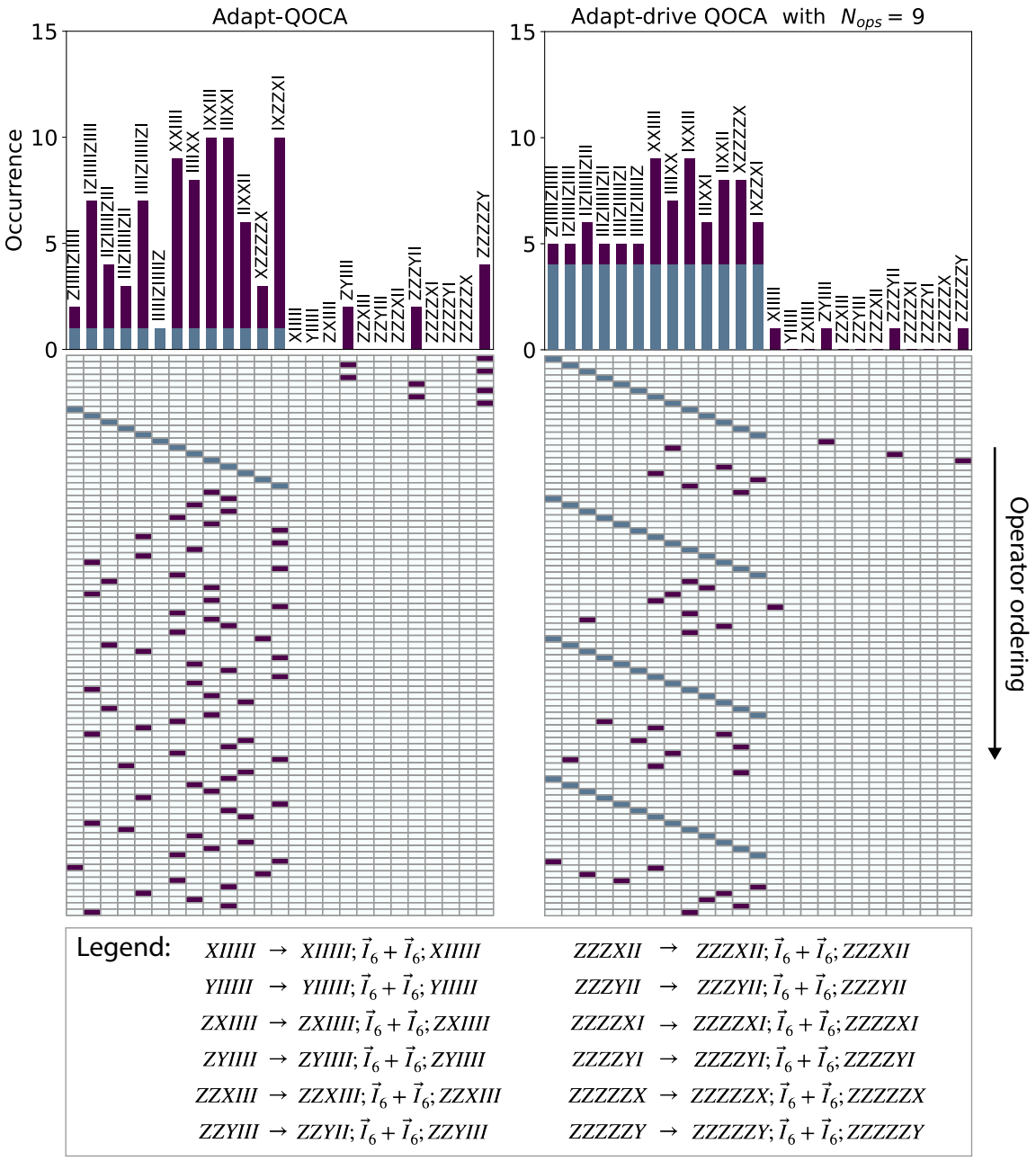


Figure 7.6 Operator occurrence FHM 2×3. The bar charts represent the number of times the operators in $\mathcal{O}_{\text{QOCA}}$ appear in the circuit built by adapt-QOCA (left) and in adapt-drive QOCA for $N_{ops} = 9$ (right). The labels on the bar chart are defined in the legend, where $\vec{I}_6$ is an abbreviation for $IIIIII$. The 13 first operators (from the left to the right) are the operators of $\mathcal{O}_{\text{VHA}}$ and the 13 following are the operators in $\mathcal{O}_{\text{drive}}$. The tables below the bar charts show the order in which the operators are added to the PQC along the *operator ordering* arrow. When a block of VHA is added to the ansatz the 13 first operators on the left are added one after another (in blue). The number of operators added from a VHA block is also indicated in blue in the bar chart.

Conclusion

During my thesis, I developed two new approches to build ansätze, namely adapt-drive QOCA and adapt-QOCA, which are combinations of the QOCA ansatz with an adapt-VQE approach. Originally, I designed adapt-drive QOCA with the idea to better design the drive part of the QOCA ansatz while keeping the VHA part fixed. The drive part is adaptively designed by choosing the operators from a pool which have the largest impact on the energy gradient of the state prepared by the ansatz. The outcomes presented in Sect. 7.2 and Sect. 7.3 showed that the ansätze grown by adapt-drive QOCA enable the convergence to the ground state with higher accuracy than QOCA. Furthermore this is achieved using fewer operators and parameters, which I have shown also led to fewer CNOTs and shallower circuits. These statements are true for any choices of N_{ops}. It is nevertheless important to note that some values of N_{ops} enable a better convergence than others. However, it is complicated to make a definitive judgment on which N_{ops} performs best as the results vary substantially depending on the parameters of the optimizer. Based on the results of adapt-drive QOCA and knowing that adapt-VQE failed for the FHM when preparing the initial state in the $|+\rangle^{\otimes n}$ state, I came to develop adapt-QOCA. Adapt-QOCA is comparable in performance if not better than adapt-QOCA for most N_{ops}. In fact, for the FHM 2×2 and 2×3, it seemed to effectively serve as a lower bound on the energy error and infidelity for adapt-drive QOCA. An additional advantage of adapt-QOCA is that it does not require to choose a value for N_{ops}. Nevertheless, a disadvantage of adapt-QOCA over adapt-drive QOCA is that it necessitates more measurements to evaluate the commutators.

I have tested adapt-drive QOCA and adapt-QOCA using three different pools: $\mathcal{O}_{\text{drive}}$, $\mathcal{O}_{\text{VHA}}$ and $\mathcal{O}_{\text{QOCA}}$. The VHA pool showed to be insufficient to enable any improvement in the energy error for the latter ansätze, highlighting the critical role played by the drive terms in quick starting the convergence process. On the other hand, the drive pool showed a significant improvement over QOCA. As one could expect, the QOCA pool generally did better than the drive pool as it allows for more flexibility in the operator choice. From the examination of the operators selected from the pool we noticed that some drive operators

were never chosen. Building on our observations for the FHM 2×2 and 2×3, we expect that the unselected drive operators could be extended to larger lattices as described in Eq. (7.3) and Eq. (7.4). Furthermore we notice that using the QOCA pool; rather few drive operators are selected, only 5 to 25% of the time for a 2×3 lattice, and 2 to 10% for a 2×3 lattice. These few drive terms are nevertheless necessary for the algorithms to convergence to the ground state.

Appendices

appendix A

Circuit quantization of the full device

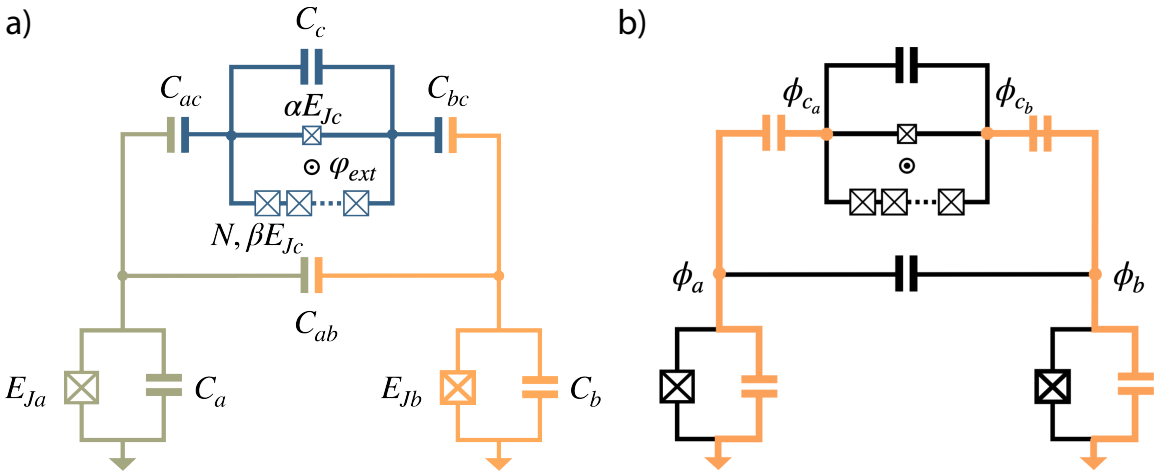


Figure A.1 Device design. Compared to Fig. 3.1, a) indicates the circuit parameters. Qubit a is shown in green, qubit b in orange and the coupler in blue. b) Shows in orange the spanning tree used here to quantize the circuit. The circuit parameters are indicated in black.

To quantize the circuit illustrated in Fig. A.1, we follow the approach described in Sect. 1.1 using the method of nodes [2]. As a first step, we introduce the circuit modes ϕ_a, ϕ_b, ϕ_{c_a} and ϕ_{c_b}, see Fig. 1.3 a) and define a spanning tree which is shown in orange in Fig. 1.3 b). As described in Sect. 1.1, we can express the kinetic and potential components of the

Lagrangian as follows:

$$\begin{aligned} E_{\text{kin}} &= \sum_{b\in B_C} \frac{C_b}{2}\dot{\Phi}_b^2 \\ &= \frac{1}{2}C_a\phi_a^2 + \frac{1}{2}C_b\phi_b^2 + \frac{1}{2}C_{ab}(\phi_a-\phi_b)^2 \\ &+ \frac{1}{2}C_c\left(\phi_{c_a}-\phi_{c_b}\right)^2 + \frac{1}{2}C_{ac}\left(\phi_a-\phi_{c_a}\right)^2 + \frac{1}{2}C_{bc}\left(\phi_b-\phi_{c_b}\right)^2 \end{aligned} \tag{A.1}$$

$$\begin{aligned} E_{\text{pot}} &= \sum_{b'\in B_L} \frac{\Phi_{b'}^2}{2L_{b'}} + \sum_{b''\in B_J} E_{Jb''}\cos\left(2\pi\frac{\Phi_{b''}}{\Phi_0}\right) \\ &= -E_{\text{J}a}\cos\left(\frac{\phi_a}{\phi_0}\right) - E_{\text{J}b}\cos\left(\frac{\phi_b}{\phi_0}\right) \\ &- \alpha\epsilon E_{\text{J}q}\cos\left(\frac{\left(\phi_{c_a}-\phi_{c_b}\right)+\phi_{ext}}{\phi_0}\right) - N\beta E_{\text{J}q}\cos\left(\frac{\left(\phi_{c_a}-\phi_{c_b}\right)}{N\phi_0}\right). \end{aligned} \tag{A.2}$$

The Lagrangian finally reads:

$$\begin{aligned} \mathcal{L} = E_{\text{kin}} - E_{\text{pot}} &= \frac{1}{2}\dot{\boldsymbol{\phi}}^\dagger \boldsymbol{C}\dot{\boldsymbol{\phi}} + E_{\text{J}a}\cos\left(\varphi_a\right) + E_{\text{J}b}\cos\left(\varphi_b\right) \\ &+ \alpha\epsilon E_{\text{J}c}\cos\left(\left(\varphi_{q_1}-\varphi_{q_2}\right)+\varphi_{\text{ext}}\right) - N\beta E_{\text{J}c}\cos\left(\frac{\left(\varphi_{q_1}-\varphi_{q_2}\right)}{N}\right), \end{aligned} \tag{A.3}$$

where we have introduced the vector notation for the flux nodes, $\boldsymbol{\phi}^t = (\phi_a, \phi_b, \phi_{c_a}, \phi_{c_b})$ and define $\varphi_n = \phi_n/\phi_0$ with $n \in \{a, b, c_a, c_b\}$, and $\Phi_0 = h/2e = 2\pi\phi_0$. The capacitance matrix reads:

$$\boldsymbol{C} = \begin{array}{c} \\ \phi_a \\ \phi_b \\ \phi_{c_a} \\ \phi_{c_b} \end{array} \begin{array}{c} \begin{array}{cccc} \phi_a & \phi_b & \phi_{c_a} & \phi_{c_b} \end{array} \\ \begin{pmatrix} C_a + C_{ac} + C_{ab} & -C_{ab} & -C_{ac} & 0 \\ -C_{ab} & C_b + C_{bc} + C_{ab} & 0 & -C_{bc} \\ -C_{ac} & 0 & C_c + C_{ac} & -C_c \\ 0 & -C_{bc} & -C_c & C_c + C_{bc} \end{pmatrix} \end{array}. \tag{A.4}$$

We now perform a basis change such that $\phi_+ = \phi_{c_a} + \phi_{c_n}$ and $\phi_- = \phi_{c_a} - \varphi_{c_n}$, resulting in:

$$\begin{aligned} \mathcal{L} &= \frac{1}{2}\dot{\boldsymbol{\phi}}'^\dagger \boldsymbol{C}'\dot{\boldsymbol{\phi}}' + E_{\text{J}a}\cos\left(\varphi_a\right) + E_{\text{J}b}\cos\left(\varphi_b\right) \\ &+ \alpha\epsilon E_{\text{J}c}\cos\left(\varphi_- + \varphi_{\text{ext}}\right) - N\beta E_{\text{J}c}\cos\left(\frac{\varphi_-}{N}\right), \end{aligned} \tag{A.5}$$

with $\boldsymbol{\phi'}^t = (\phi_a, \phi_b, \phi_+, \phi_-)$ and

$$\boldsymbol{C'} = \begin{array}{c} \\ \phi_a \\ \phi_b \\ \phi_{c_+} \\ \phi_{c_-} \end{array} \begin{array}{c} \begin{array}{cccc} \phi_a & \phi_b & \phi_{c_+} & \phi_{c_-} \end{array} \\ \begin{pmatrix} C_a + C_{ac} + C_{ab} & -C_{ab} & -\frac{1}{2}C_{ac} & 0 \\ -C_{ab} & C_2 + C_{bc} + C_{ab} & -\frac{1}{2}C_{bc} & 0 \\ -\frac{1}{2}C_{ac} & \frac{1}{2}C_{bc} & C_q + \frac{1}{4}C_{ac} + \frac{1}{4}C_{bc} & \frac{1}{4}(C_{ac} - C_{bc}) \\ -\frac{1}{2}C_{ac} & -\frac{1}{2}C_{bc} & \frac{1}{4}(C_{ac} - C_{bc}) & \frac{1}{4}(C_{ac} + C_{bc}) \end{pmatrix} \end{array}. \tag{A.6}$$

We now proceed with a Legendre transform to obtain the charge nodes, which we denote $\boldsymbol{q} = \frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{\phi}}'} = \boldsymbol{C'}\dot{\boldsymbol{\phi}}'$ in vector notation. Remarquably, the charge node q_+ is a constant of motion, because $\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}_+}\right) = \dot{q}_+ = 0$. This variable can be seen as a net charge distribution on the qubits and coupler pads, and can therefore be ignored. Accordingly, the Hamiltonian takes the form $H = \frac{1}{2}\boldsymbol{q}^\dagger \tilde{\boldsymbol{C}}^{-1}\boldsymbol{q} + E_{\text{pot}}$ where $\tilde{\boldsymbol{C}}$ is a 3×3 matrix corresponding to the matrix presented in Eq. (A.6), with the exclusion of the third row and column.

Finally, we write the quantized Hamiltonian as follows:

$$\begin{aligned} \hat{H} =& 4E_{\text{C}a}\hat{n}_a^2 - E_{\text{J}a}\cos(\hat{\varphi}_a) + 4E_{\text{C}b}\hat{n}_b^2 - E_{\text{J}b}\cos(\hat{\varphi}_b) \\ & 4E_{\text{C}c}\hat{n}_c^2 - \alpha\epsilon E_{\text{J}c}\cos\left(\hat{\varphi}_c + \varphi_{\text{ext}}\right) - N\beta E_{\text{J}c}\cos\left(\frac{\hat{\varphi}_c}{N}\right) \\ & + 4E_{\text{C}ab}\hat{n}_a\hat{n}_b + 4E_{\text{C}bc}\hat{n}_b\hat{n}_c + 4E_{\text{C}ac}\hat{n}_a\hat{n}_c \end{aligned} \tag{A.7}$$

where $\boldsymbol{n} = \frac{1}{2e}\boldsymbol{q}$ and the coupler mode $\hat{q}_{c_-}$ ($\hat{\phi}_{c_-}$) is renamed $\hat{q}_c$ ($\hat{\phi}_c$). The charging energies $E_{\text{C}j}$ for $j \in \{a, b, c\}$ and the couplings $E_{\text{C}ij}$ for $i, j \in \{a, b, c\}$ and $i \neq j$, are defined by the elements of the inverse of the capacitance matrix:

$$\begin{aligned} E_{\text{C}j} &= [\tilde{\boldsymbol{C}}^{-1}]_{jj} \\ E_{\text{C}ij} &= [\tilde{\boldsymbol{C}}^{-1}]_{ij}. \end{aligned} \tag{A.8}$$

Note that the function that relates the capacitances to the charging energies and couplings is not bijective because a mode has been removed by going from C' to $\tilde{C}$. It can sometimes be complicated to recover the capacitances from a set of charging energies and couplings.

appendix B

Appendix: Schrieffer-Wolff transformation

The Schrieffer–Wolff (SW) transformation is a technique used to derive an effective Hamiltonian by perturbatively diagonalizing a Hamiltonian to finite order in the interaction [133].

B.1 Time-independent Schrieffer-Wolff transformation

Let's now consider a time-independent Hamiltonian $H = H_0 + \lambda H_1$, which can be written in terms of an unperturbed diagonal Hamiltonian H_0 retaining most of the physics and λH_1 a small off-diagonal perturbation. The parameter λ is introduced here to keep track of the order in the perturbation theory, and will be set to one at the end of the calculation.

The SW transformation is a unitary transformation on H, commonly written:

$$H_{\text{eff}} = e^{-S} H e^{S} \tag{B.1}$$

with S the generator of the transformation, satisfying $S^\dagger = -S$, and chosen such that H_{eff} is diagonal to first order in the perturbation. Since λH_1 is small, S is also expected to be small, it can therefore be assumed that the generator can be expanded as a series in λ:

$S = \lambda S^{(1)} + \mathcal{O}(\lambda^2)$. The effective Hamiltonian H_{eff} can now be expanded in powers of λ using the Baker-Campbell-Haussdorf (BCH) formula:

$$\begin{aligned} H_{\text{eff}} &= H + [S, H] + \frac{1}{2}[S, [S, H]] + \cdots \\ &= H_0 + \lambda \left(H_1 + [S^{(1)}, H_0] \right) + \lambda^2 \left([S^{(1)}, H_1] + \frac{1}{2} \left[S^{(1)}, [S^{(1)}, H_0] \right] \right) + \mathcal{O}(\lambda^3). \end{aligned} \tag{B.2}$$

To remove the first-order term in the equation above, the generator has to be chosen such that $H_1 + [S^{(1)}, H_0] = 0$. Finally, the SW transformation yields:

$$H_{\text{eff}} = H_0 + \frac{1}{2}[S^{(1)}, H_1] + \mathcal{O}(\lambda^3). \tag{B.3}$$

Note that in principle the expansion of Eq. (B.1) in λ can be truncated at any order to obtain the effective Hamiltonian.

B.2 Time-dependent Schrieffer-Wolff perturbation theory

Let's consider a time-dependent Hamiltonian $\hat{H}(t) = \hat{H}^{(0)}(t) + \lambda \hat{H}^{(1)}(t)$, where $\hat{H}^{(0)}(t)$ is the exactly solvable unperturbed part of the Hamiltonian and $\lambda \hat{H}^{(1)}(t)$ is a small perturbation. The Floquet Hamiltonian is defined as:

$$H_F = \hat{H}(t) - i\partial_t. \tag{B.4}$$

We apply a time-dependent SW transformation on the Floquet Hamiltonian:

$$\hat{H}_{\text{eff}}(t) - i\partial_t = e^{-\hat{S}(t)} \left[\hat{H}(t) - i\partial_t \right] e^{\hat{S}(t)} \tag{B.5}$$

where $\hat{S}(t)$ is the time-dependent generator, which we again assume can be expanded as a series in λ: $\hat{S}(t) = \lambda \hat{S}^{(1)}(t) + \lambda^2 \hat{S}^{(2)}(t) + \cdots$. The effective Hamiltonian H_{eff}, can then be expanded in powers of λ by using the BCH expansion:

$$\begin{aligned} \hat{H}_{\text{eff}}(t) - i\partial_t &= e^{-\hat{S}(t)} \left(\hat{H}(t) - i\partial_t \right) e^{\hat{S}(t)} \\ &= \hat{H} - i\dot{\hat{S}} + \left[\hat{H}, \hat{S} \right] - \frac{i}{2} \left[\dot{\hat{S}}, \hat{S} \right] - i\partial_t + \cdots \\ &= \sum_{k=1}^{\infty} \lambda^k \left[\hat{H}^{(k)}(t) - i\dot{\hat{S}}^{(k)} \right] - i\partial_t \end{aligned} \tag{B.6}$$

where $\lambda^k \hat{H}^{(k)}(t) = \lambda^k \overline{\hat{H}}^{(k)} + \lambda^k \widetilde{\hat{H}}^{(k)}(t)$ can be written with a time-averaged part $\overline{\hat{H}}^{(k)}$ and an oscillatory part $\widetilde{\hat{H}}^{(k)}(t)$. The effective Floquet Hamiltonian can recursively be corrected for, order by order, by iteratively finding $\lambda^k \hat{S}^{(k)}$ and $\lambda^k \hat{H}^{(k)}$. The kth-order term in the generator $\lambda^k \hat{S}^{(k)}$ is chosen such that $\lambda^k \hat{H}^{(k)}$ is time independent at order k and less. In other words, the generator has to satisfy:

$$\lambda^k \hat{S}^{(k)} = -i \int_0^t dt' \lambda^k \widetilde{\hat{H}}^{(k)}(t') \quad \forall k. \tag{B.7}$$

In the context of Sect. 3.2.1, the effective Hamiltonian at second order is shown in Eq. (3.17).

appendix C

Publication: Accurate Methods for the Analysis of Strong-Drive Effects in Parametric Gates

Accurate Methods for the Analysis of Strong-Drive Effects in Parametric Gates

Alexandru Petrescu,[1,*] Camille Le Calonnec,[1,2,†] Catherine Leroux,[1] Agustin Di Paolo,[1,‡] Pranav Mundada,[3] Sara Sussman,[4] Andrei Vrajitoarea,[3] Andrew A. Houck,[3] and Alexandre Blais[1,5]

[1] *Institut Quantique and Département de Physique, Université de Sherbrooke, Sherbrooke, Québec J1K 2R1, Canada*

[2] *CESQ and ISIS (UMR 7006), Université de Strasbourg and CNRS, Strasbourg 67000, France*

[3] *Department of Electrical Engineering, Princeton University, Princeton, New Jersey 08544, USA*

[4] *Department of Physics, Princeton University, Princeton, New Jersey 08544, USA*

[5] *Canadian Institute for Advanced Research, Toronto, Ontario M5G 1M1, Canada*

(Received 14 October 2022; accepted 25 January 2023; published 3 April 2023)

The ability to perform fast, high-fidelity entangling gates is a requirement for a viable quantum processor. In practice, achieving fast gates often comes with the penalty of strong-drive effects that are not captured by the rotating-wave approximation. These effects can be analyzed in simulations of the gate protocol, but those are computationally costly and often hide the physics at play. Here, we show how to efficiently extract gate parameters by directly solving a Floquet eigenproblem using exact numerics and a perturbative analytical approach. As an example application of this toolkit, we study the space of parametric gates generated between two fixed-frequency transmon qubits connected by a parametrically driven coupler. Our analytical treatment, based on time-dependent Schrieffer-Wolff perturbation theory, yields closed-form expressions for gate frequencies and spurious interactions, and is valid for strong drives. From these calculations, we identify optimal regimes of operation for different types of gates including *i*SWAP, controlled-Z, and CNOT. These analytical results are supplemented by numerical Floquet computations from which we directly extract drive-dependent gate parameters. This approach has a considerable computational advantage over full simulations of time evolutions. More generally, our combined analytical and numerical strategy allows us to characterize two-qubit gates involving parametrically driven interactions, and can be applied to gate optimization and cross-talk mitigation such as the cancelation of unwanted *ZZ* interactions in multiqubit architectures.

DOI: 10.1103/PhysRevApplied.19.044003

I. INTRODUCTION

With considerable advances in state preparation, gate operation, measurement fidelity, and coherence time, superconducting qubits have become one of the leading platforms for quantum information processing [1–3]. Systems consisting of up to a few dozen qubits have been recently deployed by a number of research groups [4–6]. As these architectures are scaled up, a challenge is to engineer two-qubit interactions to realize gates that are fast enough compared to the decoherence times of the qubits, while at the same time obtaining operation fidelities that are sufficiently high to satisfy a threshold for quantum error correction [7,8]. To realize fast and high-fidelity two-qubit gates, precise modeling of the dynamics of small multiqubit systems is necessary, but becomes computationally difficult as the number of degrees of freedom increases. Moreover, to achieve fast gates, drives that are strong in the sense of the rotating-wave approximation (RWA) are necessary, in which case beyond-RWA corrections become relevant.

A dominant source of infidelity in gate operation consists of cross-Kerr interactions, or the *ZZ* terms in Pauli-matrix notation. These terms are either static due to the connectivity of qubits, or dynamically generated by control drives. In the case of many two- and single-qubit gates, *ZZ* terms produce spurious entanglement that cannot be mitigated by local single-qubit operations. There are active experimental efforts to reduce the effect of *ZZ* interactions [9–14]. Moreover, the presence of nonlocal *ZZ* interactions, and of higher-order cross-Kerr terms, can indicate the onset of quantum chaotic behavior in systems of many coupled qubits [15].

*alexandru.petrescu@minesparis.psl.eu; Current affiliation: LPENS, Département de physique, École normale supérieure, Centre Automatique et Systèmes (CAS), MINES ParisTech, Université PSL, Sorbonne Université, CNRS, Inria, 75005 Paris, France.

†camille.le.calonnec@usherbrooke.ca

‡Current affiliation: Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

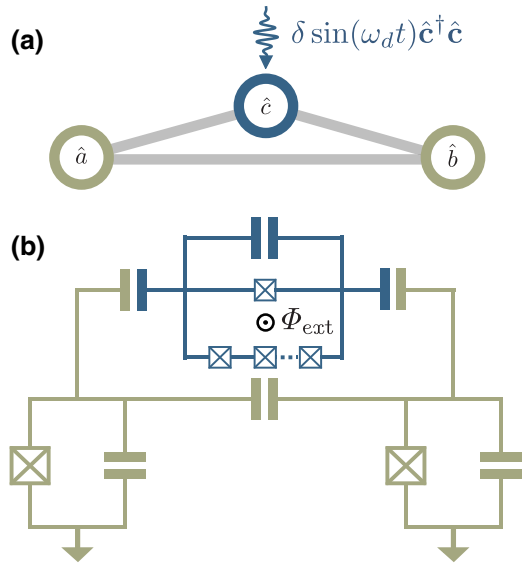


FIG. 1. (a) Graph representation of the model. The three bare modes have mutual capacitive couplings (light gray edges); mode c is parametrically driven. (b) Superconducting circuit implementation: modes a and b are transmon qubits; the coupler mode c is implemented as a generalized capacitively shunted flux qubit (see text).

In this paper, we present a computationally efficient set of analytical and numerical tools to characterize and tailor gate Hamiltonians. As an example application of these tools, we consider flux-tunable parametric coupler architectures [16,17] schematically illustrated in Fig. 1. We develop two complementary approaches, both of which start from a treatment of the Floquet Hamiltonian, which can capture non-RWA effects exactly [18–20]. Our analytical approach starts from the quantization of the *driven* superconducting circuit. More specifically, while we adopt a normal-mode picture such as in black-box quantization [21] or energy-participation-ratio approaches [22], the mode frequencies and impedances, and as a result self- and cross-Kerr interactions, depend on the strength of the drive. This dependence is accounted for in an expansion over the harmonics of the drive. From this, we obtain analytical expressions of ac-Stark shifted transition frequencies and interaction strengths. Of note, as compared to previous work, normal modes are defined here by taking into account drive-induced corrections [23] to the Josephson potential [24]. Due to its similarity to black-box quantization, this analytical technique can be easily generalized to circuits containing multiple qubits and couplers.

To obtain corrections to the effective interaction strengths, our approach relies on a time-dependent Schrieffer-Wolff perturbation theory [25–27], which consists of a hierarchy of unitary transformations applied to the time-dependent Floquet Hamiltonian [26,28]. We make the explicit choice to work in the transmon limit of small anharmonicity [29], expressed in terms of the small dimensionless parameter $\sqrt{8E_C/E_J}$, whereas drive effects are included in a series expansion over the harmonics of the drive frequency and then integrated into the exact treatment of the normal-mode Hamiltonian. This approach allows us to identify the contribution of each driven normal mode to the different effective interaction constants.

Our formalism is equally applicable to strong anharmonicities, where one has to formulate the Hamiltonian in the energy eigenbasis. The cross-resonance gate [30,31] has been modeled [27] with such methods, with the notable difference that drive effects were included in the perturbative expansion, something which requires the calculation of higher-order corrections as the drive strength is increased. In contrast, here we show that by effectively performing a series resummation over drive-amplitude contributions, we can model effects such as gate-rate saturation with drive power that are frequently observed (see e.g. Refs. [31,32]) without the need to evaluate high-order terms in perturbation theory.

On the other hand, with our numerical approach, we show how gate parameters and, more precisely, the data from a two-tone spectroscopy experiment, can be extracted from a solution of the Floquet eigenproblem [18,19]. This is efficient by comparison to the simulation of Hamiltonian dynamics over the full duration of the gate protocol: Floquet methods rely on integrating the dynamics over one period of the parametric drive, on the order of 1 ns, which is typically 3 orders of magnitude shorter than the gate duration. By construction, the parameters extracted from this approach account for renormalization by the drives. We are then able to benchmark the convergence of the analytical approach by direct comparison to the numerical result. In the context of superconducting circuit architectures, Floquet numerical methods have also been used to model instabilities in transmon qubits under strong drives [23,33], to obtain corrections beyond linear-response theory for the bilinear interaction between two cavities mediated by a driven ancilla [34], to model a strongly driven controlled-phase gate between transmon qubits [35], or to enhance the coherence of fluxonium qubits [36,37].

The remainder of this paper is structured as follows. In Sec. II we introduce the circuit model, as well as a pedagogical toy model from which all qualitative features of the full theory can be extracted, and illustrate how to obtain the different gate Hamiltonians. In Sec. III we introduce the basic concepts for second-order RWA, based on a Schrieffer-Wolff transformation of the Floquet Hamiltonian. Section IV captures in more detail the complexity of the problem with an analysis of the three-mode theory derived from the full-circuit Hamiltonian. In Sec. V, we describe in detail a method to extract effective gate Hamiltonians from a Floquet analysis. In Sec. V B, we compare all previous approaches using simulations based on the

numerical integration of the Schrödinger equation. Finally, we summarize in Sec. VI.

II. MODEL HAMILTONIAN

As a concrete example of our approach, we consider a model for a parametric coupler consisting of three nonlinear bosonic modes interacting capacitively [16], see Fig. 1(a). The qubit modes $\hat{a}$ and $\hat{b}$ are assumed to be far-detuned, making the beam splitter (or *i*SWAP) qubit-qubit interactions negligible in the rotating-wave approximation. Those modes are capacitively coupled to a third mode, the coupler $\hat{c}$. The latter can be parametrically modulated in order to activate interactions between the two qubit modes, for example a *i*SWAP-type gate on which we mostly focus here.

A. Superconducting circuit

A possible realization of this three-mode system is shown in Fig. 1(b) and consists of two fixed-frequency transmon qubits interacting via a capacitively shunted flux qubit whose two branches contain one and N Josephson junctions, respectively [38–40]. In a single-mode approximation, this generalized flux qubit plays the role of coupler mode and the parametric drive is realized by modulating the reduced external flux $\varphi_{\text{ext}} = 2\pi\,\Phi_{\text{ext}}/\Phi_0$, with Φ_{ext} the flux threading the coupler loop and Φ_0 the flux quantum. For certain values of the static external flux, the coupler has a positive anharmonicity, which is helpful in obtaining gates with a vanishing ZZ interaction [10,14,41,42]. We stress that we use this specific circuit implementation for illustration purposes only, and that the methods presented here apply beyond the weakly anharmonic regime.

Quantizing the circuit of Fig. 1(c) using the standard approach [43,44] yields the Hamiltonian (see Appendix B for a detailed derivation)

$$\hat{H}(t) = \hat{H}_a + \hat{H}_b + \hat{H}_c(t) + \hat{H}_g, \tag{1}$$

where the transmons and the coupler are described by

$$\begin{aligned}
\hat{H}_j &= 4E_{Cj}\hat{\mathbf{n}}_j^2 - E_{Jj}\cos(\hat{\boldsymbol{\varphi}}_j),\ j = a, b,\\
\hat{H}_c(t) &= 4E_{Cc}\hat{\mathbf{n}}_c^2 - \alpha E_{Jc}\cos\left[\hat{\boldsymbol{\varphi}}_c + \mu_\alpha\varphi_{\text{ext}}(t)\right]\\
&\quad - \beta N E_{Jc}\cos\left[\frac{\hat{\boldsymbol{\varphi}}_c}{N} + \mu_\beta\varphi_{\text{ext}}(t)\right].
\end{aligned} \tag{2}$$

These expressions use pairs of canonically conjugate superconducting phase difference and Cooper-pair number for the bare modes, $\left[\hat{\boldsymbol{\varphi}}_j, \hat{\mathbf{n}}_k\right] = i\delta_{jk}$ for the mode indices $j, k = a, b$, or c, and we set $\hbar = 1$. The Josephson energies are denoted E_{Ja}, E_{Jb} for the transmon modes, whereas βE_{Jc} is the Josephson energy of one of the N-array junctions in the coupler, and α is a factor parametrizing the anisotropy between the two branches. The parameter β is a renormalization of the superinductance due to disorder in the junction array and finite zero-point fluctuations (see Appendix B). Moreover, the parameter α accounts for a renormalization of the small junction energy due to hybridization with the modes in the junction array. Furthermore, E_{Ca}, E_{Cb}, and E_{Cc} are charging energies. In the transmon regime, E_{Ja}/E_{Ca} and $E_{Jb}/E_{Cb} \gtrsim 50$ [29].

The coupler loop is threaded by an external flux φ_{ext}, which can be modulated in time with a modulation amplitude $\delta\varphi$, taken to be small compared to the flux quantum

$$\varphi_{\text{ext}}(t) = \overline{\varphi}_{\text{ext}} + \delta\varphi\sin(\omega_d t). \tag{3}$$

As discussed by You *et al.* [45], quantization of the coupler loop under time-dependent flux imposes that the external flux be included in both branches of the potential energy in $\hat{H}_c(t)$, with weighting factors $\mu_{\alpha,\beta}$ determined by the capacitive energies of the two branches (see Appendix B for a detailed derivation). This subtlety is significant, as the details of the flux modulation determine the parametric interactions between the two qubit modes.

Finally, the three bare modes interact through linear terms induced by the capacitive coupling

$$\hat{H}_g = 4E_{Cab}\hat{\mathbf{n}}_a\hat{\mathbf{n}}_b + 4E_{Cbc}\hat{\mathbf{n}}_b\hat{\mathbf{n}}_c + 4E_{Cca}\hat{\mathbf{n}}_c\hat{\mathbf{n}}_a. \tag{4}$$

The introduction of normal modes will eliminate this linear coupling Hamiltonian.

B. Toy model for circuit Hamiltonian

In this subsection, we introduce a simple model, which captures the essential qualitative features of the Hamiltonian of Eq. (1). Our toy model consists of three linearly coupled Kerr-nonlinear oscillators and has the form given in Eq. (1) now with

$$\begin{aligned}
\hat{H}_a &= \boldsymbol{\omega}_a\hat{\mathbf{a}}^\dagger\hat{\mathbf{a}} + \frac{\boldsymbol{\alpha}_a}{2}\hat{\mathbf{a}}^{\dagger 2}\hat{\mathbf{a}}^2,\\
\hat{H}_b &= \boldsymbol{\omega}_b\hat{\mathbf{b}}^\dagger\hat{\mathbf{b}} + \frac{\boldsymbol{\alpha}_b}{2}\hat{\mathbf{b}}^{\dagger 2}\hat{\mathbf{b}}^2,\\
\hat{H}_c(t) &= \boldsymbol{\omega}_c(t)\hat{\mathbf{c}}^\dagger\hat{\mathbf{c}} + \frac{\boldsymbol{\alpha}_c}{2}\hat{\mathbf{c}}^{\dagger 2}\hat{\mathbf{c}}^2,\\
\hat{H}_g &= -\mathbf{g}_{ab}\hat{\mathbf{a}}^\dagger\hat{\mathbf{b}} - \mathbf{g}_{bc}\hat{\mathbf{b}}^\dagger\hat{\mathbf{c}} - \mathbf{g}_{ca}\hat{\mathbf{c}}^\dagger\hat{\mathbf{a}} + \text{H.c.}
\end{aligned} \tag{5}$$

Comparing to the full-circuit model, note that $\boldsymbol{\omega}_{a(b)} \approx \sqrt{8E_{Ja(b)}E_{Ca(b)}} - E_{Ca(b)}$ whereas the anharmonicities of the transmon qubits are negative and amount to $\boldsymbol{\alpha}_{a(b)} \approx -E_{Ca(b)}$. In an experimental implementation, the parameters defining the coupler—the anharmonicity $\boldsymbol{\alpha}_c$ and the frequency $\boldsymbol{\omega}_c(t)$—can be varied by applying a time-dependent external flux to activate a chosen gate.

The parametric drive resulting from the flux modulation of Eq. (3) is modeled by a modulation of the coupler

frequency at a frequency ω_d

$$\boldsymbol{\omega}_c(t) = \boldsymbol{\omega}_c + \delta \sin(\omega_d t). \tag{6}$$

In a more detailed analysis of the coupler (see Sec. IV), we take into account the time dependence of the anharmonicity $\boldsymbol{\alpha}_c$, but we choose to neglect it in this toy model.

Note that we reduce the complexity of the problem in a few ways: we truncate the Josephson expansion to include only quartic terms. All photon-number nonconserving terms are dropped. Higher harmonics of the drive of Eq. (6) are neglected, and we do not consider the ac-Stark shifts of the various coupling constants. All of these contributions are taken into account in the analysis of the full circuit Hamiltonian in Sec. IV. Thus the toy model is significantly simpler than the full circuit theory, but nonetheless still contains the necessary ingredients that allow us to illustrate the general method introduced in this paper.

III. PERTURBATIVE EXPANSION

In this section, we introduce a perturbative expansion to obtain successive corrections to the effective Hamiltonian in the rotating-wave approximation. To simplify the discussion, we focus on the toy model and come back to the full circuit Hamiltonian in the next section. Our approach relies on a sequence of unitary transformations amounting to a time-dependent Schrieffer-Wolff treatment of the Floquet Hamiltonian in the normal-mode representation, an approach used before to derive corrections to the lifetime of driven transmon qubits [26,28]. Time-dependent extensions of Schrieffer-Wolff transformations have been shown to be necessary to capture effects of drives in the dispersive regime of circuit QED [25], with quantitative agreement with experiment in the analysis of the cross-resonance gate [27]. A notable difference from prior work on microwave-activated two-qubit gates is that, in performing a normal-mode transformation, we are able to obtain good agreement with exact numerics already at second order in perturbation theory. For example, the calculation in Ref. [27] relies on an expansion in capacitive couplings and drive power, which would require us, in the setup presented here, to go to higher order (fourth) in the calculation to obtain results comparable to the normal-mode approach.

A. Formalism

As usual, our starting point is a decomposition of the system Hamiltonian into an unperturbed, exactly solvable part, and a perturbation:

$$\hat{H} = \hat{H}^{(0)}(t) + \lambda \hat{H}^{(1)}(t). \tag{7}$$

Here, we introduce the dimensionless power-counting parameter λ to keep track of the order in perturbation theory, to be set at the end of the calculation to unity, $\lambda \to 1$. Now we move to the interaction picture with respect to $\hat{H}^{(0)}$. Letting $\hat{U}_0(t) = \mathcal{T} e^{-i\int_0^t dt' \hat{H}^{(0)}(t')}$, where $\mathcal{T}$ is the time-ordering operator, we find for the interaction-picture Floquet Hamiltonian

$$\begin{aligned}\lambda \hat{H}_I^{(1)}(t) - i\partial_t &= \hat{U}_0^\dagger(t)\left[\hat{H}^{(0)} + \lambda \hat{H}^{(1)}(t) - i\partial_t\right]\hat{U}_0(t) \\ &= \hat{U}_0^\dagger(t)\, \lambda \hat{H}^{(1)}(t)\, \hat{U}_0(t) - i\partial_t.\end{aligned} \tag{8}$$

In the above we assume only that the unperturbed time-evolution operator $\hat{U}_0(t)$ is known. Equation (8) can be seen as a unitary transformation between two Floquet Hamiltonians [19]. Thus, the Floquet quasienergies corresponding to $\lambda \hat{H}_I^{(1)}(t) - i\partial_t$ must be identical to those of $\hat{H}^{(0)} + \lambda \hat{H}^{(1)}(t) - i\partial_t$, while the eigenstates are related by $\hat{U}_0(t)$.

In an iterative Schrieffer-Wolff approach, we treat the operator $\lambda \hat{H}^{(1)}(t)$ as a small perturbation from which we derive corrections to the known Floquet quasienergies of $\hat{H}^{(0)}$ [26,28]. To this end, we consider a unitary transformation on the interaction-picture Floquet Hamiltonian $\hat{H}_I(t) - i\partial_t \equiv \lambda \hat{H}_I^{(1)}(t) - i\partial_t$, and the corresponding Baker-Campbell-Hausdorff (BCH) expansion in powers of the generator of this unitary, that is

$$\begin{aligned}\hat{H}_{I,\text{eff}} - i\partial_t &\equiv e^{-\hat{G}_I(t)}[\hat{H}_I(t) - i\partial_t]e^{\hat{G}_I(t)} \\ &= \hat{H}_I - i\dot{\hat{G}}_I + [\hat{H}_I, \hat{G}_I] - \frac{i}{2}[\dot{\hat{G}}_I, \hat{G}_I] - i\partial_t + \cdots\end{aligned} \tag{9}$$

This equation defines the effective Hamiltonian, whose spectrum is equal (up to a desired precision in λ) to that of the original driven theory. The generator $\hat{G}_I(t)$ can be solved for iteratively in powers of λ (see Appendix A), which allows us to perform the rotating-wave approximation order by order

$$\hat{H}_{I,\text{eff}} = \lambda \overline{\hat{H}_I}^{(1)} + \lambda^2 \overline{\hat{H}_I}^{(2)} + \ldots \tag{10}$$

where the terms on the right-hand side are defined below.

To obtain a lowest-order term of the effective Hamiltonian, $\lambda \overline{\hat{H}_I}^{(1)}$, we separate the interaction picture Hamiltonian into oscillatory and nonoscillatory terms with the notation

$$\lambda \hat{H}_I^{(1)}(t) \equiv \lambda \overline{\hat{H}_I}^{(1)} + \lambda \widetilde{\hat{H}_I}^{(1)}(t), \tag{11}$$

where we define the constant part of a time-dependent operator $\hat{O}(t)$ by [46]

$$\overline{\hat{O}} \equiv \lim_{T\to\infty} \frac{1}{T}\int_0^T dt \hat{O}(t), \tag{12}$$

whereas the oscillatory part of the operator is

$$\widetilde{\hat{O}}(t) \equiv \hat{O}(t) - \overline{\hat{O}}. \tag{13}$$

Since the time-averaging operation removes all terms that are oscillatory in time, $\overline{\hat{H}}_I^{(1)}$ is the first-order RWA Hamiltonian [46], whereas $\widetilde{\hat{H}}_I^{(1)}$ is canceled by an appropriate choice of the corresponding term at order λ in the generator.

One can iterate this procedure at every order, collecting terms that are oscillatory and then canceling them. The second-order RWA Hamiltonian (for a derivation, see Appendix A) reads

$$\lambda^2 \overline{\hat{H}}_I^{(2)} = \frac{1}{i}\overline{\left[\overline{\hat{H}}_I^{(1)}, \int_0^t \lambda \widetilde{\hat{H}}_I^{(1)}(t')dt'\right]} + \frac{1}{2i}\overline{\left[\lambda\widetilde{\hat{H}}_I^{(1)}(t), \int_0^t \lambda \widetilde{\hat{H}}_I^{(1)}(t')dt'\right]}. \tag{14}$$

This form becomes analogous to the second term in the Magnus expansion [46,47] when the perturbation has a vanishing mean, i.e., $\overline{\hat{H}}_I^{(1)} = 0$.

B. Black-box quantization approach to the toy model

Expressing the toy-model Hamiltonian as the sum of static quadratic terms, $\hat{H}^{(0)}$, and of time-dependent and Kerr terms, $\hat{H}^{(1)}(t)$, the first step in deriving parametrically activated interactions between the transmon modes is to diagonalize the former, which we write as

$$\hat{H}^{(0)} = \begin{pmatrix} \hat{\mathbf{a}}^\dagger & \hat{\mathbf{b}}^\dagger & \hat{\mathbf{c}}^\dagger \end{pmatrix} \begin{pmatrix} \boldsymbol{\omega}_a & \mathbf{g}_{ab} & \mathbf{g}_{ca} \\ \mathbf{g}_{ab} & \boldsymbol{\omega}_b & \mathbf{g}_{bc} \\ \mathbf{g}_{ca} & \mathbf{g}_{bc} & \boldsymbol{\omega}_c \end{pmatrix} \begin{pmatrix} \hat{\mathbf{a}} \\ \hat{\mathbf{b}} \\ \hat{\mathbf{c}} \end{pmatrix}. \tag{15}$$

This diagonalization is achieved with an orthonormal transformation $\hat{\boldsymbol{\alpha}} = \sum_{\beta=a,b,c} u_{\alpha\beta}\hat{\beta}$, for $\alpha = a, b, c$, and which is chosen such that $\hat{H}^{(0)}$ takes the form

$$\hat{H}^{(0)} = \omega_a \hat{a}^\dagger \hat{a} + \omega_b \hat{b}^\dagger \hat{b} + \omega_c \hat{c}^\dagger \hat{c}, \tag{16}$$

where $\hat{a}$, $\hat{b}$, and $\hat{c}$ are the normal modes and $\omega_{a,b,c}$ the corresponding mode frequencies. The $u_{\alpha\beta}$ are hybridization coefficients encoding the connectivity of the three modes through the capacitive couplings $\hat{H}_g$ entering in $\hat{H}^{(0)}$. In this normal-mode basis, the remainder of the Hamiltonian reads

$$\begin{aligned}\lambda\hat{H}^{(1)}(t) &= \sum_{j=a,b,c} \frac{\boldsymbol{\alpha}_j}{2}(u_{ja}\hat{a} + u_{jb}\hat{b} + u_{jc}\hat{c})^{\dagger 2}(u_{ja}\hat{a} + u_{jb}\hat{b} + u_{jc}\hat{c})^2 \\ &+ \delta \sin(\omega_d t)(u_{ca}\hat{a} + u_{cb}\hat{b} + u_{cc}\hat{c})^\dagger(u_{ca}\hat{a} + u_{cb}\hat{b} + u_{cc}\hat{c}).\end{aligned} \tag{17}$$

The expression above illustrates that coupling between the normal modes arises from the nonlinearity and the parametric drive.

Our choice of unperturbed Hamiltonian $\hat{H}^{(0)}$ and perturbation $\lambda\hat{H}^{(1)}$ in Eqs. (16) and (17), respectively, is guided by black-box quantization [21]: the unperturbed Hamiltonian is linear and diagonal in the normal-mode basis, whereas the perturbation consists of Kerr-nonlinear terms, on the one hand, and quadratics appearing from the normal-mode expansion of the parametric drive, on the other hand. As we show below, while better choices are possible (see Sec. III C), this choice leads to a simple and intuitive form for the effective Hamiltonian.

As an example of the many common types of interactions that can be activated by a parametric drive [3], an *i*SWAP interaction between the transmon modes arises if we set the modulation to be at the frequency difference between the two transmon modes

$$\omega_d = \omega_b - \omega_a. \tag{18}$$

This choice yields the first-order RWA Hamiltonian

$$\begin{aligned}\lambda\overline{\hat{H}}_I^{(1)} &= J_{ab}^{(1)}\left(-i\hat{a}^\dagger\hat{b} + \text{H.c.}\right) \\ &+ \frac{\alpha_a^{(1)}}{2}\hat{a}^{\dagger 2}\hat{a}^2 + \frac{\alpha_b^{(1)}}{2}\hat{b}^{\dagger 2}\hat{b}^2 + \frac{\alpha_c^{(1)}}{2}\hat{c}^{\dagger 2}\hat{c}^2 \\ &+ \chi_{ab}^{(1)}\hat{a}^\dagger\hat{a}\hat{b}^\dagger\hat{b} + \chi_{bc}^{(1)}\hat{b}^\dagger\hat{b}\hat{c}^\dagger\hat{c} + \chi_{ca}^{(1)}\hat{c}^\dagger\hat{c}\hat{a}^\dagger\hat{a}.\end{aligned} \tag{19}$$

The first row of this equation contains the *i*SWAP interaction of amplitude $J_{ab}^{(1)}$. The second row contains the mode anharmonicities, and the third row contains cross-Kerr interactions, the first of which is the *ZZ* term.

The coupling constants in the above effective Hamiltonian result from the normal-mode transformation of the quadratic part of the toy model and take the form

$$\begin{aligned} J_{ab}^{(1)} &= u_{ca}u_{cb}\frac{\delta}{2}, \quad \alpha_j^{(1)} = \sum_{i=a,b,c} u_{ij}^4 \boldsymbol{\alpha}_i, \\ \chi_{jk}^{(1)} &= \sum_{i=a,b,c} 2u_{ij}^2 u_{ik}^2 \boldsymbol{\alpha}_i, \end{aligned} \tag{20}$$

for all $j, k = a, b, c$, and $j \neq k$. In practice, one wants to maximize $J_{ab}^{(1)}$ to obtain a fast gate, while minimizing

the cross-Kerr interactions $\chi_{jk}^{(1)}$ to avoid the accumulation of coherent errors. Cross-Kerr interactions are a source of infidelity for a *i*SWAP-type gate, as well as in other gate implementations [10–14,48]. In the first-order RWA Hamiltonian, to cancel the cross-Kerr interaction between the two transmons, we use a coupler with a positive anharmonicity [42] $\boldsymbol{\alpha}_c > 0$, together with $\boldsymbol{\alpha}_a, \boldsymbol{\alpha}_b < 0$, which is distinct from using qubits of opposite anharmonicities [10,41], or couplers with negative anharmonicity [14,49]. Equation (20) forms the basis for the optimization of the gate parameters. Before pursuing this further, we first derive useful corrections to the gate Hamiltonian from the oscillatory part of the Hamiltonian, $\widetilde{\hat{H}}_I^{(1)}(t)$. Finally, note that the first-order term $\chi_{jk}^{(1)}$ is only a *static*, i.e., δ-independent, cross-Kerr interaction.

At second order in perturbation theory, there is no correction to the *i*SWAP gate frequency $J_{ab}^{(2)} = 0$. In the regimes of interest, where the coupler frequency is close enough to the qubit frequencies for the interaction between the coupler and the qubits to be non-negligible, the dominant contribution to the second-order RWA correction to the cross-Kerr interaction $\chi_{ab}^{(2)}$ is

$$\chi_{ab}^{(2)} \approx 2\frac{\left(\sum_{j=a,b,c} u_{aj}\, u_{bj}\, u_{cj}^2\, \boldsymbol{\alpha}_j\right)^2}{\omega_a + \omega_b - 2\omega_c} + \delta\frac{u_{ac}u_{bc}\left[u_{aa}^3 u_{ba}\boldsymbol{\alpha}_a - u_{bb}^3 u_{ab}\boldsymbol{\alpha}_b\right]}{\omega_a - \omega_b}. \tag{21}$$

The full expression for $\chi_{ab}^{(2)}$ can be found in Appendix C. Inspecting the hybridization coefficients $u_{\alpha\beta}$ and the denominators in Eq. (21), we deduce that the second-order correction to the *static* cross-Kerr interaction, corresponding to the first term in Eq. (21), arises from a virtual two-photon excitation of the coupler (generated by the commutator $[\hat{a}\hat{b}\hat{c}^{\dagger 2}, \hat{a}^\dagger\hat{b}^\dagger\hat{c}^2]$). This correction would not be present in a two-level approximation [16]. On the other hand, the second term in Eq. (21) is the lowest-order contribution to the *dynamical* cross-Kerr interaction.

C. Improving the starting point of the perturbation theory

As mentioned in the previous subsection, other choices for $\hat{H}^{(0)}$ and $\lambda\hat{H}^{(1)}$ are possible, which give better accuracy in comparisons with exact Floquet numerics. In this subsection, we take the unperturbed Hamiltonian $\hat{H}^{(0)}(t)$ to consist of the Fock-space diagonal part of $\hat{H}(t)$, namely

$$\begin{aligned}\hat{H}^{(0)}(t) &= \omega_a\hat{a}^\dagger\hat{a} + \omega_b\hat{b}^\dagger\hat{b} + \omega_c\hat{c}^\dagger\hat{c} \\ &\quad + \delta\sin(\omega_d t)\left[u_{ca}^2\hat{a}^\dagger\hat{a} + u_{cb}^2\hat{b}^\dagger\hat{b} + u_{cc}^2\hat{c}^\dagger\hat{c}\right] \\ &\quad + \frac{\alpha_a^{(0)}}{2}\hat{a}^{\dagger 2}\hat{a}^2 + \frac{\alpha_b^{(0)}}{2}\hat{b}^{\dagger 2}\hat{b}^2 + \frac{\alpha_c^{(0)}}{2}\hat{c}^{\dagger 2}\hat{c}^2 \\ &\quad + \chi_{ab}^{(0)}\hat{a}^\dagger\hat{a}\hat{b}^\dagger\hat{b} + \chi_{ac}^{(0)}\hat{a}^\dagger\hat{a}\hat{c}^\dagger\hat{c} + \chi_{bc}^{(0)}\hat{b}^\dagger\hat{b}\hat{c}^\dagger\hat{c},\end{aligned} \tag{22}$$

where the quartic couplings in the last two rows are exactly those defined in Eq. (20), with the superscript now changed from 1 to 0 to reflect their presence in the unperturbed Hamiltonian.

We expect this starting point, Eq. (22), to lead to more precise results, because of two reasons. Firstly, the perturbation $\hat{H}^{(1)}$ is now off diagonal in Fock space. The effects of the anharmonicities of the modes are now included at the level of $\hat{H}^{(0)}$, and, in particular, we expect a dressing of contributions corresponding to two-photon excitations, such as Eq. (21). Secondly, due to the second row of Eq. (22), we can derive the effect of harmonics of the drive frequency through a Fourier expansion of the time-evolution operator $\hat{U}_0(t)$, defined below.

Following the steps of the previous subsections, we evaluate the time-evolution operator with respect to the unperturbed Hamiltonian $\hat{H}^{(0)}(t)$, that is

$$\hat{U}_0(t) = e^{-i\int_0^t dt'\hat{H}^{(0)}(t')}, \tag{23}$$

with $[\hat{H}^{(0)}(t), \hat{H}^{(0)}(t')] = 0$. The time dependence in the exponent is handled via the Jacobi-Anger identity [50]

$$e^{i\hat{O}\frac{\delta}{\omega_d}\cos\omega_d t} = J_0\left(\frac{\hat{O}\delta}{\omega_d}\right) + 2\sum_{n=1}^{\infty} i^n J_n\left(\frac{\hat{O}\delta}{\omega_d}\right)\cos n\omega_d t, \tag{24}$$

for any operator $\hat{O}$, where $J_n(z)$ is the nth Bessel function of the first kind [50]. This expansion allows us to keep track of all harmonics of the drive. In practice, since the modulation amplitude is small, $\delta/\omega_d \ll 1$, only a few terms will be necessary.

We obtain to first order (and truncating after the first Bessel function)

$$J_{ab}^{(1)} = \frac{\delta}{2}u_{ca}u_{cb}\left[J_0\left(\frac{\delta u_{ca}^2}{\omega_d}\right)J_0\left(\frac{\delta u_{cb}^2}{\omega_d}\right) + 3J_1\left(\frac{\delta u_{ca}^2}{\omega_d}\right)J_1\left(\frac{\delta u_{cb}^2}{\omega_d}\right)\right], \tag{25}$$

which agrees, up to linear terms in δ, with the expression in Eq. (20). On the other hand, the cross-Kerr interaction at this order is vanishing $\chi_{ab}^{(1)} = 0$.

To second order in perturbation theory, the dominant contributions to $J_{ab}^{(2)}$ are

$$J_{ab}^{(2)} = \frac{i\delta^2}{2} u_{ca} u_{cb} u_{cc}^2 \left(\frac{1}{\omega_a - \omega_c} - \frac{1}{\omega_b - \omega_c} \right) \times J_1\left(\frac{\delta u_{cc}^2}{\omega_d} \right) \prod_{j=a,b,c} J_0\left(\frac{\delta u_{cj}^2}{\omega_d} \right) + \ldots . \quad (26)$$

We do not reproduce here the full form containing 20 terms. The second-order contribution $\chi_{ab}^{(2)}$ is nonvanishing, and contains approximately 450 terms in expanded form. Despite the complexity of these full expressions, they are easy to derive and manipulate with symbolic computation tools [51]. Focusing on the static cross-Kerr interaction, i.e., in the $\delta \to 0$ limit, the dominant correction resulting from the above changes amounts to our previous Eq. (21), but replacing the denominator of the first term of that expression by a form that faithfully includes the contribution from the anharmonicities, as expected in the case of a virtual two-photon excitation of the coupler mode. That is, approximately

$$\omega_a + \omega_b - 2\omega_c \to \omega_a + \omega_b - (2\omega_c + u_{cc}^4 \boldsymbol{\alpha}_c). \quad (27)$$

The full expression for the corrected static cross-Kerr interaction can be found in Appendix C.

In Fig. 2 we compare analytical results to numerical results obtained from exact diagonalization (at $\delta = 0$), or a solution of the Floquet eigenspectrum at $\delta \neq 0$ (see Sec. V). We find that the agreement between numerics and analytics is excellent for the gate interaction strength, as well as for the static $\delta = 0$ cross-Kerr interaction. However, we find that second-order perturbation theory is insufficient to reproduce the effects of the drive on the cross-Kerr interaction, even for modest drive amplitudes. We expect that higher-order perturbation theory should correctly capture the drive-amplitude dependence of the anharmonicities, but these contributions have been inaccessible in our study due to the large memory demands of the computer algebra manipulations.

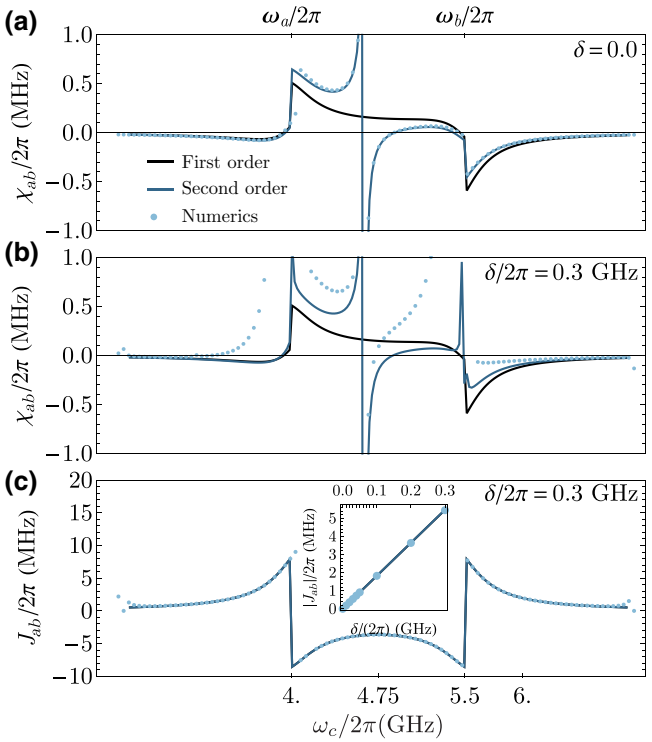


FIG. 2. (a) Static cross-Kerr interaction $\chi_{ab}(\boldsymbol{\omega}_c)$, from first- (black) and second-order RWA (blue), and from the full diagonalization of Sec. V (light blue points) for $\boldsymbol{\omega}_a/2\pi = 4.0$, $\boldsymbol{\omega}_b/2\pi = 5.5$, $\boldsymbol{\alpha}_a/2\pi = -0.3$, $\boldsymbol{\alpha}_b/2\pi = -0.2$, $\boldsymbol{\alpha}_c/2\pi = 0.25$, $\mathbf{g}_{ab}/2\pi = 0.12$, $\mathbf{g}_{bc}/2\pi = -0.12$, all in GHz, and $\mathbf{g}_{ca}/2\pi = 0$. (b) Analogue of (a) for *dynamical* cross-Kerr interaction at $\delta/2\pi = 0.3$ GHz. (c) Same as (b) for the gate interaction rate $J_{ab}(\boldsymbol{\omega}_c)$. Inset: $J_{ab}(\delta)$ at $\boldsymbol{\omega}_c/2\pi = 4.25$ GHz.

IV. FULL-CIRCUIT HAMILTONIAN

Building on the previous results, we now turn to deriving an effective Hamiltonian for the full-circuit Hamiltonian of Eqs. (1) and (2). The full-circuit model goes beyond the toy model in that it systematically includes the effects of the parametric drive on all of the coupling constants. Although the simplicity of the toy model is useful in developing an intuitive understanding of the effect of parametric drives on the system, the full-circuit model can lead to more accurate comparisons with experimental data.

The full-circuit theory is constructed with the following steps: we first introduce creation and annihilation operators for the bare circuit modes starting from the first-order RWA driven circuit Hamiltonian in Sec. IV A. Because the drive is taken into account at that level, the frequencies and zero-point fluctuations of these bare modes will be explicitly corrected by the drive. In Sec. IV B, we perform a normal-mode transformation amounting to a *driven* black-box quantization approach. We then show in Sec. IV C how a variety of quantum gates can be addressed by appropriate choices of the parametric drive frequency. Lastly, we find corrections to the desired gate Hamiltonian using a time-dependent Schrieffer-Wolff perturbation theory in Sec. IV C.

A. Bare-mode Hamiltonian

To define the bare modes, we begin with the full-circuit model Hamiltonian of Eqs. (1) and (2). We normal order expand the Josephson cosine potentials in this Hamiltonian over a set of creation and annihilation operators, which we

define as follows:

$$\hat{\boldsymbol{\varphi}}_a = \sqrt{\frac{\eta_a}{2}}(\hat{\mathbf{a}} + \hat{\mathbf{a}}^\dagger),$$
$$\hat{\mathbf{n}}_a = -i\sqrt{\frac{1}{2\eta_a}}(\hat{\mathbf{a}} - \hat{\mathbf{a}}^\dagger), \tag{28}$$

with analogous equations for modes $\hat{\mathbf{b}}$ and $\hat{\mathbf{c}}$. The coefficients $\eta_{a,b,c}$ are chosen such that terms proportional to $\hat{\mathbf{a}}^2, \hat{\mathbf{b}}^2$, and $\hat{\mathbf{c}}^2$ vanish in the time-averaged Hamiltonian. This amounts to three transcendental equations:

$$\mathcal{F}(\eta_j)\eta_j^2 = 8E_{Cj}/E_{Jj}, \tag{29}$$

for $j = a, b, c$, where we define the form factors

$$\mathcal{F}(\eta_{a(b)}) \equiv e^{-\eta_{a(b)}/4},$$
$$\mathcal{F}(\eta_c) \equiv \alpha e^{-\eta_c/4} J_0(\mu_\alpha \delta\varphi)\cos\left(\mu_\alpha \overline{\varphi}_{\text{ext}}\right) + \frac{\beta}{N} e^{-\eta_c/4N^2} J_0(\mu_\beta \delta\varphi)\cos\left(\mu_\beta \overline{\varphi}_{\text{ext}}\right). \tag{30}$$

Note that in the transmon limit $\mathcal{F}(\eta_{a(b)}) \approx 1$ we recover the usual expression $\eta_{a(b)} \approx \sqrt{8E_{Ca/b}/E_{Ja/b}}$. Secondly, the parameter η_c depends on the parametric drive amplitude $\delta\varphi$, which indicates that the mode c impedance is drive dependent. This has consequences for the precision of the calculation of coupling constants dressed by the parametric drives. In particular, it allows us to capture the ac-Stark shift of the coupler mode at the lowest order in perturbation theory. In what follows, sine and cosine functions of the phase are normal order expanded according to Eq. (D7) of Appendix D. In turn, trigonometric functions of the flux modulation are expanded in Jacobi-Anger series over the harmonics of the frequency of the drive.

Using the above definitions, the transmon Hamiltonian $\hat{H}_a$ takes the familiar form

$$\hat{H}_a = \boldsymbol{\omega}_a \hat{\mathbf{a}}^\dagger \hat{\mathbf{a}} - E_{Ja}\left(\cos\hat{\boldsymbol{\varphi}}_a + e^{-\eta_a/4}\frac{\hat{\boldsymbol{\varphi}}_a^2}{2}\right). \tag{31}$$

The second term on the right-hand side contains the nonlinear part of the Josephson potential, i.e., the inductive part is subtracted. Up to quartic order, $\hat{H}_a$ takes the form

$$\hat{H}_a = \boldsymbol{\omega}_a \hat{\mathbf{a}}^\dagger \hat{\mathbf{a}} + \frac{\boldsymbol{\alpha}_a}{2}\hat{\mathbf{a}}^{\dagger 2}\hat{\mathbf{a}}^2 + \frac{\boldsymbol{\alpha}_a}{12}\left(\hat{\mathbf{a}}^4 + \hat{\mathbf{a}}^{\dagger 4}\right) + \frac{\boldsymbol{\alpha}_a}{3}\left(\hat{\mathbf{a}}^\dagger \hat{\mathbf{a}}^3 + \hat{\mathbf{a}}^{\dagger 3}\hat{\mathbf{a}}\right) + \cdots \tag{32}$$

The first row of this expression is a Kerr-oscillator Hamiltonian as in the toy model of Sec. II, whereas the second row contains corrections from quartic counter-rotating terms. Here, we have introduced the mode frequency and anharmonicity, which take the forms [29]

$$\boldsymbol{\omega}_a = \frac{4E_{Ca}}{\eta_a} + \frac{1}{2}\mathcal{F}(\eta_a)\eta_a E_{Ja} \approx \sqrt{8E_{Ca}E_{Ja}} - E_{Ca},$$
$$\boldsymbol{\alpha}_a = -E_{Ca}. \tag{33}$$

Note that for the approximate equality in the first row we use a Taylor expansion of Eq. (29) for η_a. The equations for mode $\hat{\mathbf{b}}$ are identical from the above with a change of subscripts and operators $a \to b$.

The coupler Hamiltonian differs from that of the transmon modes in two fundamental ways: it breaks parity symmetry due to the external flux, and it is time dependent. Following Eqs. (13) and (12), we write this time-dependent Hamiltonian as

$$\hat{H}_c(t) = \overline{\hat{H}}_c(t) + \widetilde{\hat{H}}_c(t). \tag{34}$$

The creation and annihilation operators of the coupler mode can then be defined by extracting the quadratic part of the time-averaged coupler Hamiltonian. Using Eq. (28) where $a \to c$ together with

$$\overline{\cos\left[\hat{\boldsymbol{\varphi}}_c + \mu_\alpha \varphi_{\text{ext}}(t)\right]} = \cos(\mu_\alpha \overline{\varphi}_{\text{ext}}) J_0(\mu_\alpha \delta\varphi)\cos(\hat{\boldsymbol{\varphi}}_c) \tag{35}$$

and a similar relation for the second branch of the coupler [see Eq. (2)], we find in analogy to Eq. (31) for the Hamiltonian of the transmon mode

$$\begin{aligned}\overline{\hat{H}}_c = {} & \boldsymbol{\omega}_c \hat{\mathbf{c}}^\dagger \hat{\mathbf{c}} \\ & - \alpha E_{Jc} J_0(\mu_\alpha \delta\varphi)\cos(\mu_\alpha \overline{\varphi}_{\text{ext}})\left(\cos\hat{\boldsymbol{\varphi}}_c + e^{-\eta_c/4}\frac{\hat{\boldsymbol{\varphi}}_c^2}{2}\right) \\ & - \beta N E_{Jc} J_0(\mu_\beta \delta\varphi)\cos(\mu_\beta \overline{\varphi}_{\text{ext}}) \\ & \times \left(\cos\frac{\hat{\boldsymbol{\varphi}}_c}{N} + e^{-\eta_c/4N^2}\frac{\hat{\boldsymbol{\varphi}}_c^2}{2N^2}\right) \\ & + \alpha E_{Jc} J_0(\mu_\alpha \delta\varphi)\sin(\mu_\alpha \overline{\varphi}_{\text{ext}})\sin\hat{\boldsymbol{\varphi}}_c \\ & + \beta N E_{Jc} J_0(\mu_\beta \delta\varphi)\sin(\mu_\beta \overline{\varphi}_{\text{ext}})\sin\frac{\hat{\boldsymbol{\varphi}}_c}{N}.\end{aligned} \tag{36}$$

Crucially, in this first-order rotating-wave approximation of the parametric drive, the Josephson energy is renormalized by the factor $J_0(\mu_{\alpha,\beta}\delta\varphi)$, see also Ref. [23]. We interpret this as an effective reduction of the Josephson potential barrier, and consequently an increase of phase fluctuations, in the presence of drives. Moreover, the presence of the nonzero external flux results in the parity breaking sine terms in Eq. (36).

The second term of $\hat{H}_c(t)$ in Eq. (34), the oscillatory part, take the form

$$\widetilde{\hat{H}}_c(t) = -\alpha E_{Jc}\cos\left[\widetilde{\hat{\boldsymbol{\varphi}}_c + \mu_\alpha \varphi_{\rm ext}(t)}\right] - \beta N E_{Jc}\cos\left[\widetilde{\frac{\hat{\boldsymbol{\varphi}}_c}{N} + \mu_\beta \varphi_{\rm ext}(t)}\right], \quad (37)$$

which can be expanded in a Jacobi-Anger series in harmonics oscillating at the frequency $n\omega_d$, where n is an integer.

As above, the next step is to expand the coupler Hamiltonian up to quartic terms in the creation and annihilation operators. In contrast to the transmon Hamiltonian of Eq. (32), parity breaking leads to the appearance of monomials of odd order. The nonoscillatory part is

$$\begin{aligned}\overline{\hat{H}}_c = {}& \boldsymbol{\omega}_c \hat{\mathbf{c}}^\dagger \hat{\mathbf{c}} + \frac{\boldsymbol{\alpha}_c}{2}\hat{\mathbf{c}}^{\dagger 2}\hat{\mathbf{c}}^2 \\ & + \frac{\boldsymbol{\alpha}_c}{12}\left(\hat{\mathbf{c}}^4 + \hat{\mathbf{c}}^{\dagger 4}\right) + \frac{\boldsymbol{\alpha}_c}{3}\left(\hat{\mathbf{c}}^\dagger \hat{\mathbf{c}}^3 + \hat{\mathbf{c}}^{\dagger 3}\hat{\mathbf{c}}\right) \\ & + \mathbf{g}_{c,3}\left(\hat{\mathbf{c}}^3 + \hat{\mathbf{c}}^{\dagger 3} + 3\hat{\mathbf{c}}^\dagger \hat{\mathbf{c}}^2 + 3\hat{\mathbf{c}}^{\dagger 2}\hat{\mathbf{c}}\right) \\ & + \mathbf{g}_{c,1}\left(\hat{\mathbf{c}} + \hat{\mathbf{c}}^\dagger\right) + \cdots . \end{aligned} \quad (38)$$

The first row of the above expression takes the form of the coupler Hamiltonian in the approximation of the toy model of Sec. II, while the remaining rows contain number-nonconserving terms up to quartic order. Here, the parametric drive-dependent mode frequency and anharmonicity read

$$\begin{aligned}\boldsymbol{\omega}_c &= \frac{4E_{Cc}}{\eta_c} + \frac{1}{2}\mathcal{F}(\eta_c)\eta_c E_{Jc}, \\ \boldsymbol{\alpha}_c &= -E_{Cc}, \end{aligned} \quad (39)$$

while the prefactors of the counter-rotating terms are

$$\begin{aligned}\mathbf{g}_{c,3} = {}& -\alpha\epsilon e^{-\eta_c/4}\eta_c^{3/2}J_0(\mu_\alpha\delta\varphi)\sin\left(\mu_\alpha\overline{\varphi}_{\rm ext}\right)E_{Jc}/(12\sqrt{2}) \\ & - \frac{\beta}{N^2}e^{-\eta_c/4N^2}\eta_c^{3/2}J_0(\mu_\beta\delta\varphi)\sin \\ & \times\left(\mu_\beta\overline{\varphi}_{\rm ext}\right)E_{Jc}/(12\sqrt{2}), \\ \mathbf{g}_{c,1} = {}& \alpha\epsilon e^{-\eta_c/4}\eta_c^{1/2}J_0(\mu_\alpha\delta\varphi)\sin\left(\mu_\alpha\overline{\varphi}_{\rm ext}\right)E_{Jc}/\sqrt{2} \\ & + \beta e^{-\eta_c/4N^2}\eta_c^{1/2}J_0(\mu_\beta\delta\varphi)\sin\left(\mu_\beta\overline{\varphi}_{\rm ext}\right)E_{Jc}/\sqrt{2}. \end{aligned} \quad (40)$$

The contribution from the oscillatory part $\widetilde{\hat{H}}_c(t)$ is too lengthy to be reproduced here, and is given up to the second harmonic of the parametric modulation frequency ω_d in Table II of Appendix D.

Finally, the last term of the full circuit Hamiltonian to consider is the linear interaction $\hat{H}_g$ induced by the capacitive coupling. Using Eq. (28), this Hamiltonian takes the form

$$\hat{H}_g = -\frac{2E_{Cab}}{\sqrt{\eta_a\eta_b}}(\hat{\mathbf{a}} - \hat{\mathbf{a}}^\dagger)(\hat{\mathbf{b}} - \hat{\mathbf{b}}^\dagger) + \ldots \quad (41)$$

where the ellipsis represents two more terms corresponding to the cyclic permutations of the mode indices.

The Hamiltonian specified by Eq. (32) and its equivalent for the b transmon mode, together with Eq. (38), the terms summarized in Table II of Appendix D, and Eq. (41), form the basis of the full-circuit numerical simulation performed in Sec. V.

B. Driven black-box quantization approach for parametrically activated interactions

We now follow the procedure developed with the toy model in Sec. III B to obtain effective gate Hamiltonians under parametric modulations. To do so, we first displace the Hamiltonian of Eq. (1) by approximate solutions to the corresponding classical equations of motion, with the aim of removing all contributions linear in the coordinates $\hat{\boldsymbol{\varphi}}_j, \hat{\mathbf{n}}_j$. As such, this avoids keeping the linear terms as part of the perturbative expansion. This procedure is detailed in Appendix D and amounts to making the following replacement in Eq. (1):

$$\hat{\boldsymbol{\varphi}}_j \to \hat{\boldsymbol{\varphi}}_j + \xi_j + \zeta_j \sin(\omega_d t). \quad (42)$$

The parameters ζ_j, ξ_j are found numerically. The procedure above can best be understood as a change of frame in which all coordinates $\hat{\boldsymbol{\varphi}}_j$, $\hat{\mathbf{n}}_j$ represent quantum fluctuations about a known classical trajectory. In particular, should the amplitude responses ζ_j be neglected, then ξ_j would be the amount by which phase variables need to be displaced such that the subsequent Taylor expansion is performed around the classical minimum of the potential energy.

Next, we collect under $\hat{H}^{(0)}$ the time-independent quadratic terms, in a procedure analogous to the one above. We then eliminate the linear coupling $\hat{H}_g$ of Eq. (41) from $\hat{H}^{(0)}$ through a normal-mode transformation

$$\begin{aligned}\hat{H}^{(0)} &= \boldsymbol{\omega}_a\hat{\mathbf{a}}^\dagger\hat{\mathbf{a}} + \boldsymbol{\omega}_b\hat{\mathbf{b}}^\dagger\hat{\mathbf{b}} + \boldsymbol{\omega}_c\hat{\mathbf{c}}^\dagger\hat{\mathbf{c}} + \hat{H}_g \\ &\equiv \omega_a\hat{a}^\dagger\hat{a} + \omega_b\hat{b}^\dagger\hat{b} + \omega_c\hat{c}^\dagger\hat{c}. \end{aligned} \quad (43)$$

The linear transformation is determined by a set of 18 hybridization coefficients that relate bare mode coordinates to normal-mode coordinates (see Appendix E for the

procedure to compute these coefficients)

$$\hat{\boldsymbol{\varphi}}_\alpha = \sum_{\beta=a,b,c} \frac{u_{\alpha\beta}}{\sqrt{2}}(\hat{\beta}+\hat{\beta}^\dagger),$$
$$\hat{\mathbf{n}}_\alpha = \sum_{\beta=a,b,c} \frac{v_{\alpha\beta}}{i\sqrt{2}}(\hat{\beta}-\hat{\beta}^\dagger), \tag{44}$$

for $\alpha = a, b, c$. We stress that the hybridization coefficients $u_{\alpha\beta}$, $v_{\alpha\beta}$ depend on the amplitude of the parametric drive. As a result, drive effects such as the ac-Stark shift of the nonlinear oscillators are accounted for already at the level of the normal-mode decomposition of the circuit Hamiltonian.

In analogy with our treatment in Sec. II of the toy model, we take $\hat{H}^{(0)}$ to be the unperturbed Hamiltonian with respect to which the interaction picture is defined. Primarily in order to keep the expressions more concise, we opt to neglect the corrections analyzed in Sec. III C. With $\hat{H}^{(0)}$ the unperturbed Hamiltonian, the remaining interaction terms are $\lambda\hat{H}^{(1)}(t) \equiv \hat{H} - \hat{H}^{(0)}$. Expressing these in the interaction picture with respect to $\hat{U}_0 = e^{-i\hat{H}^{(0)}t}$ as in Eq. (8), we find

$$\lambda\hat{H}_I^{(1)}(t) = \hat{U}_0^\dagger(t)\left[\hat{H}(t) - \hat{H}^{(0)}\right]\hat{U}_0(t), \tag{45}$$

which, as before, is decomposed into oscillatory and nonoscillatory parts.

As a first example, to realize a beam-splitter interaction, the first-order RWA Hamiltonian is obtained in the form of Eq. (19) for a modulation frequency that satisfies

$$\omega_d = \omega_b - \omega_a. \tag{46}$$

Of note, as already mentioned, the right-hand side of the above definition depends implicitly on the drive frequency ω_d, since it is defined in terms of ac-Stark shifted normal-mode frequencies. In Sec. V we present a numerical procedure to obtain the parametric drive frequency.

With this choice of modulation frequency, the effective Hamiltonian takes the form

$$\begin{aligned}
\lambda\overline{\hat{H}}_I^{(1)} &= J_{ab}^{(1)}(-i\hat{a}^\dagger\hat{b} + \text{H.c.}) \\
&+ \frac{\alpha_a^{(1)}}{2}\hat{a}^{\dagger 2}\hat{a}^2 + \frac{\alpha_b^{(1)}}{2}\hat{b}^{\dagger 2}\hat{b}^2 + \frac{\alpha_c^{(1)}}{2}\hat{c}^{\dagger 2}\hat{c}^2 \\
&+ \chi_{ab}^{(1)}\hat{a}^\dagger\hat{a}\hat{b}^\dagger\hat{b} + \chi_{bc}^{(1)}\hat{b}^\dagger\hat{b}\hat{c}^\dagger\hat{c} + \chi_{ca}^{(1)}\hat{c}^\dagger\hat{c}\hat{a}^\dagger\hat{a} \\
&+ J_{ab;a}^{(1)}(-i\hat{a}^\dagger\hat{a}\hat{a}^\dagger\hat{b} + \text{H.c.}) \\
&+ J_{ab;b}^{(1)}(-i\hat{b}^\dagger\hat{b}\hat{a}^\dagger\hat{b} + \text{H.c.}) \\
&+ J_{ab;c}^{(1)}(-i\hat{c}^\dagger\hat{c}\hat{a}^\dagger\hat{b} + \text{H.c.}) \\
&+ K_{ab}^{(1)}(\hat{a}^{\dagger 2}\hat{b}^2 + \text{H.c.}).
\end{aligned} \tag{47}$$

In contrast to the effective gate Hamiltonian Eq. (19) obtained for the toy model, there are additional terms in the last four rows, namely photon-number-conditioned beam-splitter terms and a photon-pair beam-splitter term. The couplings appearing in the above Hamiltonian are

$$\begin{aligned}
J_{ab}^{(1)} &= -\frac{u_{aa}u_{ab}}{2}J_1(\zeta_a)\sin(\xi_a)E_{Ja} \\
&- \frac{u_{ba}u_{bb}}{2}J_1(\zeta_b)\sin(\xi_b)E_{Jb} \\
&- \frac{u_{ca}u_{cb}}{2}\alpha J_1(\zeta_c + \mu_\alpha\delta\varphi)\sin(\xi_c + \mu_\alpha\overline{\varphi}_{\text{ext}})E_{Jc}^{(\alpha)} \\
&- \frac{u_{ca}u_{cb}}{2}\frac{\beta}{N}J_1\left(\frac{\zeta_c}{N} + \mu_\beta\delta\varphi\right)\sin \\
&\times\left(\frac{\xi_c}{N} + \mu_\beta\overline{\varphi}_{\text{ext}}\right)E_{Jc}^{(\beta)}, \\
\alpha_j^{(1)} &= -\frac{1}{8}\sum_{i=a,b,c} u_{ij}^4 E'_{J,i}, \\
\chi_{jk}^{(1)} &= -\frac{1}{4}\sum_{i=a,b,c} u_{ij}^2 u_{ik}^2 E'_{J,i}, \\
J_{ab;j}^{(1)} &= -\frac{u_{cj}^2}{4}J_{ab}^{(1)}, \text{ for } j = a, b, c, \\
K_{ab}^{(1)} &= -\frac{u_{aa}^2 u_{ab}^2}{16}J_2(\zeta_a)\cos(\xi_a)E_{Ja} \\
&- \frac{u_{ba}^2 u_{bb}^2}{16}J_2(\zeta_b)\cos(\xi_b)E_{Jb} \\
&- \frac{u_{ca}^2 u_{cb}^2}{16}\alpha J_2(\zeta_c + \mu_\alpha\delta\varphi)\cos(\xi_c + \mu_\alpha\overline{\varphi}_{\text{ext}})E_{Jc}^{(\alpha)} \\
&- \frac{u_{ca}^2 u_{cb}^2}{16}\frac{\beta}{N^3}J_2\left(\frac{\zeta_c}{N} + \mu_\beta\delta\varphi\right)\cos \\
&\times\left(\frac{\xi_c}{N} + \mu_\beta\overline{\varphi}_{\text{ext}}\right)E_{Jc}^{(\beta)},
\end{aligned} \tag{48}$$

where we use

$$\begin{aligned}
E'_{Ja} &\equiv e^{-u_{aa}^2/4 - u_{ab}^2/4 - u_{ac}^2/4}J_0(\zeta_a)\cos(\xi_a)E_{Ja}, \\
E'_{Jb} &\equiv e^{-u_{ba}^2/4 - u_{bb}^2/4 - u_{bc}^2/4}J_0(\zeta_b)\cos(\xi_b)E_{Jb}, \\
E'_{Jc} &\equiv \alpha J_0(\zeta_c + \mu_\alpha\delta\varphi)\cos(\xi_c + \mu_\alpha\overline{\varphi}_{\text{ext}})E_{Jc}^{(\alpha)} \\
&+ \frac{\beta}{N^3}J_0\left(\frac{\zeta_c}{N} + \mu_\beta\delta\varphi\right)\cos\left(\frac{\xi_c}{N} + \mu_\beta\overline{\varphi}_{\text{ext}}\right)E_{Jc}^{(\beta)}, \\
E_{Jc}^{(\alpha)} &\equiv e^{-u_{ca}^2/4 - u_{cb}^2/4 - u_{cc}^2/4}E_{Jc}, \\
E_{Jc}^{(\beta)} &\equiv e^{-u_{ca}^2/4N^2 - u_{cb}^2/4N^2 - u_{cc}^2/4N^2}E_{Jc}.
\end{aligned} \tag{49}$$

The above expressions depend on the drive both explicitly, through the arguments of the Bessel functions, and implicitly, through the hybridization coefficients u_{jk}. While

the amplitude responses on the two transmon qubits are expected to be small, i.e., $\zeta_{a,b} \ll \zeta_c$ and $\xi_{a,b} \ll \xi_c$, these contributions show that, through hybridization, all Josephson elements contribute to the resonant parametric interaction. The first three lines of Eq. (48) are similar in form to those obtained for the toy model in Eq. (20). Of the two additional classes of terms possible in the full-circuit model at this order in perturbation theory, the photon-pair beam-splitter term, in the last row of Eq. (47), is generated by the second harmonic of the drive. However, since this term is fourth order in the hybridization coefficients, it can only become comparable to the beam-splitter interaction at vanishing external flux $\overline{\varphi}_{\rm ext} \approx 0$, or if $\delta\varphi$ is set to cancel $J_{ab}^{(1)}$.

As in the case of the toy model, going to second order in perturbation theory using Eq. (14) we find corrections to the coupling constants derived above to first order. Noting that parity-breaking terms significantly dress the coupler $0 \to 1$ transition frequency, we absorb this renormalization into a reparametrization of the external flux $\overline{\varphi}_{\rm ext} \to \overline{\varphi}'_{\rm ext}(\overline{\varphi}_{\rm ext})$ such that $\omega_c(\overline{\varphi}'_{\rm ext}) = \omega_c^{(2)}(\overline{\varphi}_{\rm ext})$, i.e., we absorb the corrections to the coupler pole at second order in perturbation theory into a redefinition of the coupler normal mode, in a self-consistent approach that can be further validated with exact numerics.

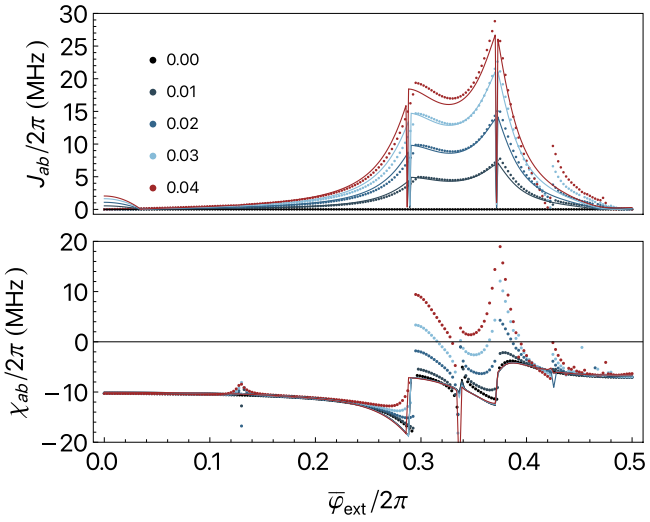


FIG. 3. Coupling constants in the effective Hamiltonian for the full circuit as a function of external dc flux $\overline{\varphi}_{\rm ext}$. Dots (lines) represent Floquet two-tone spectroscopy data with Hilbert-space dimension 10 per mode (second-order RWA calculations). Color (see legend) encodes parametric drive amplitude $\delta\varphi/2\pi$. Parameter choices: $C_a = 134.205$ fF, $C_b = 134.218$ fF, $C_c = 75.987$ fF, $C_{ac} = 11.11$ fF, $C_{bc} = 11.22$ fF, $C_{ab} = 0$, $E_{Ja}/2\pi = 37$ GHz, $E_{Jb}/2\pi = 27$ GHz, $E_{Jc}/2\pi = 50$ GHz, $\alpha = 0.258$, $\beta = 1$, and $N = 3$, $\mu_\alpha = 5/6$ and $\mu_\beta = -1/18$. We attribute large discontinuities in the numerical curves to state tracking errors near avoided crossings (see Sec. V).

In Fig. 3 we show a comparison between exact Floquet numerics (see Sec. V) and second-order perturbation theory for the full-circuit model. We find that the analytics reproduce with good accuracy the numerical results for the gate interaction rate J_{ab} in the region where the coupler $0 \to 1$ frequency lies between the two transmons: $\omega_a < \omega_c < \omega_a$. There are poles in the numerical gate rate J_{ab} for $\omega_c < \omega_a$ or for $\omega_b < \omega_c$ that we expect to capture only at third order in perturbation theory. The numerical cross-Kerr interaction, as in the case of the toy model, only agrees well with analytics in the static case $\delta\varphi = 0$. Focusing our attention on the curves obtained from Floquet numerics, we see that with a typical set of parameter gate rates as large as $J_{ab}/2\pi \sim 20$ MHz (equivalent to a 25 ns $\sqrt{i\rm SWAP}$ gate) can be achieved while maintaining a vanishing dynamical cross-Kerr interaction. The tools presented in this paper feed into a larger scale optimization of the circuit parameters, which forms the subject of a future study.

C. Other parametric gates

The space of parametric gates is not limited to beam-splitter-type, or red sideband, terms. Indeed, different interactions can be activated by appropriate choices of the frequency of the parametric drive [16,52–58]. For example, if instead the modulation frequency targets the blue sideband,

$$\omega_d = \omega_a + \omega_b, \tag{50}$$

then the resulting interaction is a two-mode squeezing term. The effective gate Hamiltonian is formally the same as Eq. (47) with the simple modification

$$\hat{a}^\dagger \hat{b} \to \hat{a}^\dagger \hat{b}^\dagger, \tag{51}$$

in the first line and in the last four lines of Eq. (47). The coupling constants remain formally as in Eq. (48). Note, though, that quantitatively the rates will differ, since the classical responses and the hybridization coefficients are dependent on drive frequency. As opposed to the beam-splitter interaction, we expect [59] nonadiabatic effects at the larger modulation frequency Eq. (50), which will require higher orders in perturbation theory beyond the scope of this work.

It is also possible to obtain a CNOT interaction induced by a parametric drive at $\omega_d = \omega_a$, which makes the a transmon mode into the target mode of a cross-resonance protocol [30,31]. Following the same procedure as in the preceding subsection, with this choice of modulation frequency we arrive at the effective gate Hamiltonian

$$\begin{aligned}\lambda\overline{\hat{H}}_I^{(1)} &= -i\Omega_{a;b}(\hat{a} - \hat{a}^\dagger)\hat{b}^\dagger\hat{b} - i\Omega_{a;c}(\hat{a} - \hat{a}^\dagger)\hat{c}^\dagger\hat{c} \\ &\quad - i\Omega_a(\hat{a} - \hat{a}^\dagger) - i\Omega_{a;a}(\hat{a}^\dagger\hat{a}\hat{a} - \hat{a}^\dagger\hat{a}^\dagger\hat{a}).\end{aligned} \tag{52}$$

The first term of above expression generates the cross-resonance gate, while the second term is a coupler-state conditional drive on mode a, which is negligible for $\langle \hat{c}^\dagger \hat{c} \rangle \approx 0$. On the other hand, the second row contains local operations on qubit a.

The coupling constants in Eq. (52) take the form

$$\Omega_a = u_{ca} E''_{J,c}, \Omega_{a;a} = u_{ca}^3 E'''_{J,c}/2,$$
$$\Omega_{a;b} = u_{ca} u_{cb}^2 E'''_{J,c}, \Omega_{a;c} = u_{ca} u_{cc}^2 E'''_{J,c}, \tag{53}$$

where we define

$$\begin{aligned} E''_{J,c} &= \frac{\alpha}{\sqrt{2}} J_1 \left(\zeta_c + \mu_\alpha \delta\varphi\right) \cos\left(\xi_c + \mu_\alpha \overline{\varphi}_{\text{ext}}\right) E^{(\alpha)}_{J,c} \\ &\quad + \frac{\beta}{\sqrt{2}} J_1 \left(\frac{\zeta_c}{N} + \mu_\beta \delta\varphi\right) \cos\left(\frac{\xi_c}{N} + \mu_\beta \overline{\varphi}_{\text{ext}}\right) E^{(\beta)}_{J,c}, \\ E'''_{J,c} &= -\frac{\alpha}{\sqrt{2}} J_1 \left(\zeta_c + \mu_\alpha \delta\varphi\right) \cos\left(\xi_c + \mu_\alpha \overline{\varphi}_{\text{ext}}\right) E^{(\alpha)}_{J,c} \\ &\quad - \frac{\beta}{\sqrt{2}N^2} J_1 \left(\frac{\zeta_c}{N} + \mu_\beta \delta\varphi\right) \cos\left(\frac{\xi_c}{N} + \mu_\beta \overline{\varphi}_{\text{ext}}\right) E^{(\beta)}_{J,c}. \end{aligned} \tag{54}$$

For brevity, in the expressions above we drop the smaller contributions proportional to $J_1(\zeta_{a,b})$. While in the standard cross-resonance gate protocol the gate is activated by a microwave tone on one of the qubits [30,31], here it is the coupler mode c that is parametrically driven. This protocol to achieve a CNOT gate is advantageous if the coupler mode is much more strongly coupled to the transmon modes a and b than their direct capacitive coupling. In the standard cross-resonance protocol [31,32], the CNOT gate rate $\Omega_{a;b}$ saturates as a function of the amplitude of the parametric drive; in this model saturation could be in part due to the Bessel function J_1. Table I summarizes the different interactions that can be obtained for different choices of modulation frequencies.

V. FLOQUET NUMERICS

In this section we use exact numerical Floquet methods to extract the effective gate Hamiltonian from quasienergy spectra. Floquet theory validates the results obtained using perturbation theory in Secs. III and IV C. On the other hand, this numerically exact method is applicable beyond the regime of validity of perturbation theory. In this section, we first briefly introduce the method and the notation in Sec. V A and, as an example application, return to our toy model to extract the cross-Kerr interaction χ_{ab} and the $\sqrt{i\text{SWAP}}$ gate amplitude J_{ab}. Using these results, we show how to adjust the system parameters such as to cancel the dynamical cross-Kerr interaction during an $\sqrt{i\text{SWAP}}$ gate. Then, in Sec. V B, we apply the method to the full-circuit Hamiltonian. In particular, we perform a numerical experiment analogous to two-tone spectroscopy for the parametrically driven circuit. For completeness, an introduction to Floquet theory is presented in Appendix F.

A. Effective Hamiltonian from Floquet spectra

Our analysis starts from the observation that the effective Hamiltonian is unitarily equivalent to the Floquet Hamiltonian according to Eqs. (8) and (9), and therefore their quasienergy spectra (see Appendix F) are identical. In the laboratory frame, we can write

$$\hat{H}_{\text{eff}} - i\partial_t = e^{-\hat{G}(t)} \left[\hat{H}(t) - i\partial_t\right] e^{\hat{G}(t)}. \tag{55}$$

The perturbative expansion for $e^{-\hat{G}(t)}$, and consequently that for $\hat{H}_{\text{eff}}$, is therefore an iterative approach to finding the Floquet spectrum.

In this section we compute the Floquet spectrum exactly and show how the parameters of the effective Hamiltonian can be extracted from it. ac-Stark shifted normal-mode frequencies, self- and cross-Kerr interactions, and gate amplitudes are formulated as linear combinations of appropriately identified eigenvalues of the Floquet Hamiltonian. For illustration, in this subsection we confine our attention to the Floquet analysis of the toy model of Eq. (5).

To identify states in the Floquet quasienergy spectrum, we find eigenvectors that have a maximum overlap with a set of known, unperturbed states. We let the state $|i_a i_b i_c\rangle$ be the eigenstate of the time-independent Schrödinger equation for the undriven Hamiltonian, that has maximum overlap with the Fock state $|i_a\rangle |i_b\rangle |i_c\rangle$, and denote its eigenenergy by $E_{i_a i_b i_c}$. Finally, we define $|i_a i_b i_c\rangle_F$ as the Floquet eigenmode having maximum overlap with $|i_a i_b i_c\rangle$, and we denote its quasienergy with $\epsilon_{i_a i_b i_c}$. In what follows, we label kets by three integers as above, in the order $a - b - c$.

TABLE I. List of the most accessible gate Hamiltonians realizable with a parametric drive in the analyzed architecture.

Gate	Bosonic operator	Drive frequency	Dominant unwanted interaction	Equation
iSWAP and beam splitter	$-i\hat{a}^\dagger \hat{b} + i\hat{b}^\dagger \hat{a}$	$\omega_a - \omega_b$	$\hat{a}^\dagger \hat{a} \hat{b}^\dagger \hat{b}$	Eq. (47)
Two-mode squeezing	$-i\hat{a}^\dagger \hat{b}^\dagger + i\hat{b}\hat{a}$	$\omega_a + \omega_b$	$\hat{a}^\dagger \hat{a} \hat{b}^\dagger \hat{b}$	Eqs. (47) and (51)
CZ and Ising ZZ	$\hat{a}^\dagger \hat{a} \hat{b}^\dagger \hat{b}$	No drive		Eq. (47)
CNOT	$-i(\hat{a} - \hat{a}^\dagger)\hat{b}^\dagger \hat{b}$	ω_a	$-i(\hat{a} - \hat{a}^\dagger)\hat{a}^\dagger \hat{a}$	Eq. (52)
CSWAP	$-i\hat{c}^\dagger \hat{c}(\hat{a}^\dagger \hat{b} - \hat{b}^\dagger \hat{a})$	$\omega_a - \omega_b$	$-i\hat{a}^\dagger \hat{b} + i\hat{b}^\dagger \hat{a}$	Eq. (47)

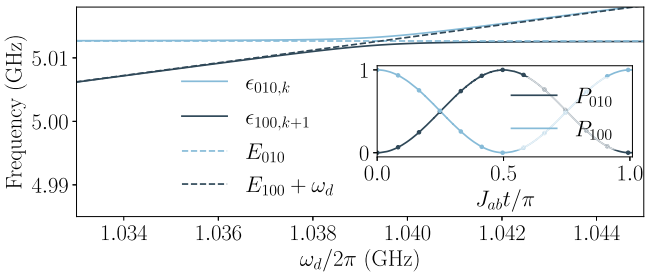


FIG. 4. Quasienergies of the Floquet modes with maximum overlap with eigenstates $|0_a 1_b 0_c\rangle$ and $|1_a 0_b 0_c\rangle$, for the toy model. The light and dark blue dashed lines correspond to the eigenenergies of the uncoupled system. The inset shows the population of the Floquet states, $P_{n_a n_b n_c}(t) = |_F\langle n_a n_b n_c|\psi(t)\rangle|^2$, compared to the state populations of a two-level system (dots), driven resonantly with Rabi rate J_{ab}, where J_{ab} is the gate amplitude obtained from the avoided crossing in the Floquet spectrum.

With these definitions, the gate amplitude J_{ab} has a natural interpretation in the Floquet formalism. As shown above, the $\sqrt{i\text{SWAP}}$ interaction arises in the toy model if

$$\omega_d = \omega_b - \omega_a \equiv E_{100} - E_{010}. \tag{56}$$

Since the parametric drive enters via a term proportional to $\hat{\mathbf{c}}^\dagger\hat{\mathbf{c}}$, which couples the undriven eigenstates in the two-state manifold $\{|100\rangle, |010\rangle\}$, there is an avoided crossing between the Floquet modes $|100; k+1\rangle_F$ and $|010; k\rangle_F$, as shown in Fig. 4. Because the gate operation is analogous to Rabi oscillations in the two-state manifold $\{|100\rangle, |010\rangle\}$, the size of the avoided crossing is twice the effective gate amplitude, $2J_{ab}$. For example, if an excitation is originally prepared in the transmon b, then population dynamics would obey $P_{010}(t) \equiv |_F\langle 010|\psi(t)\rangle|^2 = \sin^2(J_{ab}t)$ and $P_{100} = 1 - P_{010}$, in full agreement with exact numerics (inset of Fig. 4). Away from the avoided crossing, the difference between the dressed states and the undriven states corresponds to the ac-Stark shift of the transmon normal modes due to the off-resonant drive.

Note that, in practice, the two-state manifold $\{|100\rangle, |010\rangle\}$ is coupled by the drive to other levels. The resonant drive frequency ω_d is then slightly shifted from Eq. (56) due to the ac-Stark effect induced by these additional couplings, and the exact value can be determined numerically by minimizing the size of the anticrossing.

The dynamical cross-Kerr interaction χ_{ab} is written in terms of a Walsh transform [15] of the quasienergies

$$\chi_{ab}(\delta) = \epsilon_{110} - \epsilon_{100} - \epsilon_{010} + \epsilon_{000}, \tag{57}$$

and reduces to the static cross-Kerr when the parametric drive is turned off:

$$\chi_{ab}(0) = E_{110} - E_{100} - E_{010} + E_{000}. \tag{58}$$

Along with J_{ab} and χ_{ab}, any ac-Stark-shifted quantity pertaining to the effective Hamiltonian can, in principle, be obtained by taking appropriate linear combinations of the quasienergies in the Floquet spectrum.

Since the Floquet quasienergy spectrum can be obtained from the propagator $\hat{U}(2\pi/\omega_d, 0)$ over one period of the drive (Appendix F), the Floquet method is numerically efficient as compared to the simulation of the dynamics over the complete gate time. The period of the drive is on the order of 1 ns, which is between 2 and 3 orders of magnitude shorter than the gate times studied here. Due to its relatively small computational footprint, the Floquet method allows us to efficiently search for optimal gate parameters, e.g., a maximal J_{ab} with a minimal residual cross-Kerr interaction, χ_{ab}. As an example, in Fig. 5 we study the behavior of J_{ab} and χ_{ab} as a function of the bare coupler frequency $\boldsymbol{\omega}_c$ for different choices of drive amplitude, δ, and bare coupler anharmonicity, $\boldsymbol{\alpha}_c$.

From these studies we can, for example, find parameters for which the cross-Kerr interaction χ_{ab} vanishes. As

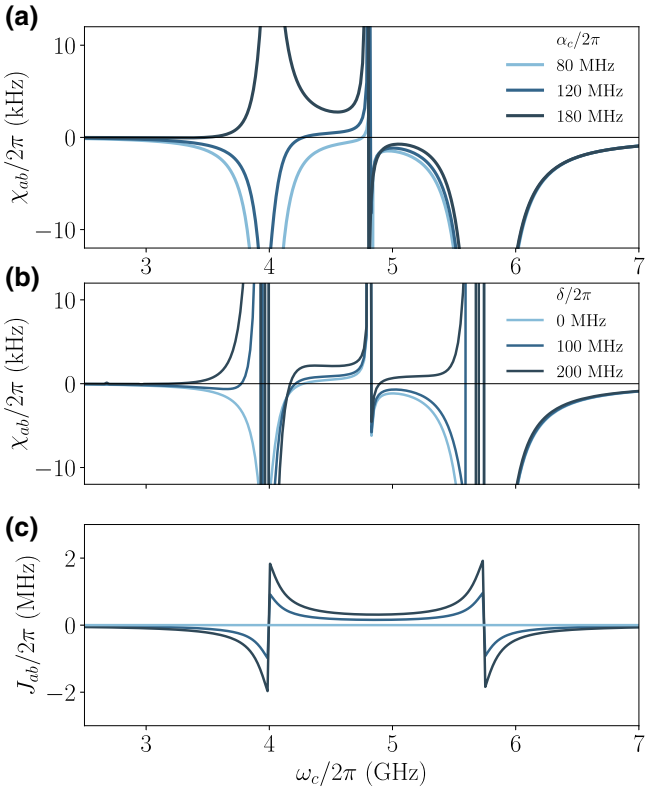


FIG. 5. (a) Static χ_{ab} interaction at $\delta = 0$ for the toy model versus the coupler frequency for different values of the coupler anharmonicities, with remaining parameters $\boldsymbol{\omega}_a/2\pi = 4.0$, $\boldsymbol{\omega}_b/2\pi = 5.75$, $\boldsymbol{\alpha}_a/2\pi = \boldsymbol{\alpha}_b/2\pi = -0.2$, and $\mathbf{g}_{ac}/2\pi = -\mathbf{g}_{bc}/2\pi = 0.05$ GHz. (b) Dynamical χ_{ab} versus bare coupler frequency for the parameters above and $\boldsymbol{\alpha}_c/2\pi = 0.12$ GHz, for different values of the drive amplitude. (c) Gate amplitude J_{ab} for the parameters in (b). The Hilbert-space dimension for each mode is 5.

already mentioned, this is helpful to obtain high-fidelity two-qubit gates and relies on choosing a positive coupler anharmonicity, $\boldsymbol{\alpha}_c$. In Fig. 5 we find that, while varying $\boldsymbol{\alpha}_c$ does not affect J_{ab} to lowest order in perturbation theory, it has a considerable impact on χ_{ab}. Indeed, Fig. 5(a) shows the static χ_{ab} for multiple values of $\boldsymbol{\alpha}_c$, and illustrates that it is possible to fine tune $\boldsymbol{\alpha}_c$ to cancel χ_{ab}. We observe empirically that, in the RWA, whenever the bare anharmonicities obey $\boldsymbol{\alpha}_a^{-1} + \boldsymbol{\alpha}_b^{-1} + \boldsymbol{\alpha}_c^{-1} = 0$, one can find $\boldsymbol{\omega}_c$ for which $\chi_{ab} = 0$ at a sweet spot, where $\partial\chi_{ab}/\partial\delta \approx 0$.

For the dynamical cross-Kerr interaction χ_{ab} one observes complex variations with δ. The main resonances appear for $\boldsymbol{\omega}_c \approx \boldsymbol{\omega}_a$, $\boldsymbol{\omega}_b$, and $(\boldsymbol{\omega}_a + \boldsymbol{\omega}_b)/2$ but the slopes and the sign of χ_{ab} change and additional resonances appear away from the qubit frequencies, especially when the drive amplitude and coupling strengths $g_{ab,ac}$ are sufficiently large. As illustrated in Fig. 5(b), by tuning the drive amplitude it is possible to find a bare coupler frequency, $\boldsymbol{\omega}_c$, for which the effective $\chi_{ab}(\delta) = 0$. On the other hand, as seen in Fig. 5(c), the gate rate increases with δ without qualitative changes of its dependence on $\boldsymbol{\omega}_c$. Therefore, as the gate is turned on or off by varying δ, one can adjust the bare coupler frequency $\boldsymbol{\omega}_c$ to maintain the instantaneous $\chi_{ab}(\delta) = 0$. This defines a cross-Kerr-free curve in the parameter space $(\boldsymbol{\omega}_c, \delta)$ connecting the "off" point $\delta = 0, \chi_{ab}(0) = 0, J_{ab}(0) = 0$ to the "on" point $\delta \neq 0, \chi_{ab}(\delta) = 0, J_{ab}(\delta) \neq 0$. We study this in detail on the realistic full-circuit model in Sec. V B.

B. Full-circuit simulation

In this section, we apply the Floquet numerical method to the full-circuit Hamiltonian of Sec. IV A. We study the dependence of the coupling constants in the effective gate Hamiltonian versus dc flux and as a function of the drive amplitude.

Figure 6 shows the analogues of the plots in Fig. 5 for the gate amplitude $J_{ab}(\overline{\varphi}_{\text{ext}})$ and of the cross-Kerr $\chi_{ab}(\overline{\varphi}_{\text{ext}})$ now for the full-circuit Hamiltonian. State tracking is performed as described in the previous subsection. However, in the vicinity of avoided crossings, it is impossible to identify with certainty the states generated by the relatively large capacitive couplings considered here. We therefore introduce exclusion regions where state tracking is unreliable. Even though the tracking is expected to be complicated by the presence of counter-rotating terms coupling states with different photon numbers in the full device Hamiltonian of Sec. IV A, we find that this is not a significant source of tracking error, as compared to errors due to large hybridization.

In Fig. 6(b), we represent χ_{ab} versus the dc flux $\overline{\varphi}_{\text{ext}}$ for different values of the flux-drive amplitude. Unlike the toy model, χ_{ab} does not go to zero away from the qubit-coupler resonances. This is because the qubit-coupler detuning saturates as a function of $\overline{\varphi}_{\text{ext}}$, as opposed to the toy model

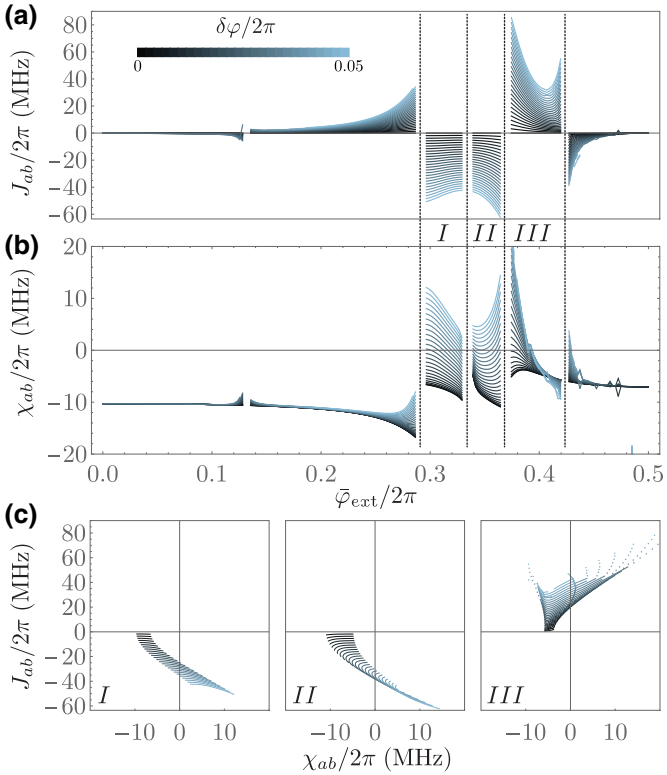


FIG. 6. Same parameter choices as Fig. 3. Gate amplitude J_{ab} (a) and cross-Kerr χ_{ab} (b) for different parameteric drive amplitudes $\delta\varphi$ (encoded in curve color) and as a function of the static flux $\overline{\varphi}_{\text{ext}}$ from Floquet simulations. Data has been excluded where deficient state tracking in the vicinity of avoided crossings led to unphysical discontinuities in the quantities. (c) for regions I, II, and III identified in (a) and (b), we eliminate the common parameter $\overline{\varphi}_{\text{ext}}$ and plot $J_{ab}(\chi_{ab})$. This allows us to identify those regimes in which the $\sqrt{i\text{SWAP}}$ gate interaction can be turned on, while maintaining a vanishing dynamical χ_{ab}.

where the detuning could be increased arbitrarily. In the undriven case (black), we see that for this set of device parameters there does not exist a flux value for which χ_{ab} vanishes. However, increasing the drive amplitude allows for an active cancelation of the dynamical χ_{ab} at some flux value. The corresponding behavior of J_{ab} is shown in Fig. 6(a). In Fig. 6(c), we synthesize the numerical results into three favorable regions of operation for the parametric gate, denoted I, II, and III, respectively [see (a)]. For these regions, we eliminate the external flux and plot directly the gate amplitude J_{ab} against the dynamical cross-Kerr interaction χ_{ab}. This allows us to determine regimes of optimal $\sqrt{i\text{SWAP}}$ gate operation. We conclude that gate amplitudes as high as 40 MHz, corresponding to a gate time of 12.5 ns, can be achieved with vanishing cross-Kerr interaction for these parameter choices.

For both J_{ab} and χ_{ab} there exist peaks away from the qubit-coupler resonances, situated at $\bar{\varphi}_{\text{ext}}/2\pi \approx 0.13, 0.42$. These correspond to avoided crossings appearing in the

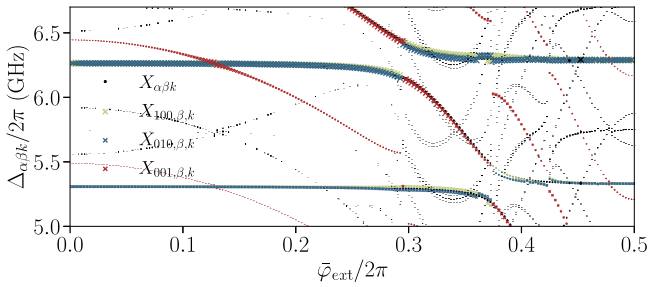


FIG. 7. Two-tone spectroscopy data from Floquet numerics. Each point corresponds to a possible transition and its size is weighed by the matrix element of the charge operator of bare qubit a, $\hat{X} = \hat{\mathbf{n}}_a$. Parameters chosen as in Fig. 3 with $\delta\varphi/2\pi = 0.0$ (black dots) and 0.03 (crosses). The subscripts α, β sweep over the subset of Floquet modes $\{|000\rangle_F, |100\rangle_F, |010\rangle_F, |001\rangle_F\}$, whereas the drive photon number k takes integer values between -15 and 15.

driven Floquet spectrum, corresponding to the hybridization of Floquet levels involving distinct numbers of drive photons. For example, the Floquet level $|100, k\rangle_F$ can couple to the Floquet level $|001, k-1\rangle_F$. This can be seen by unfolding the Floquet spectrum in spectroscopy simulations (see Fig. 7).

To exemplify the full extent of the Floquet analysis, we generate two-tone spectroscopy data from our simulations according to Eqs. (F2) and (F3) in Appendix F, by focusing on the experimentally relevant situation where the parametric drive is on, while the (second) probe tone acts on the bare charge operator $\hat{\mathbf{n}}_a$. In Fig. 7 we represent the numerically computed spectrum close to the two-qubit transition frequencies. The size of each point is proportional to the absolute value of the corresponding matrix element. The black dots correspond to transition frequencies in the undriven spectrum. As expected, the dot sizes are larger for the transitions involving the probed qubit a. The large avoided crossings around $\bar{\varphi}_{\rm ext}/2\pi \approx \{0.32, 0.38\}$ result from the capacitive couplings between the coupler and the qubits. Secondary avoided crossings appear between the coupler mode and the transmons near $\bar{\varphi}_{ext}/2\pi \approx \{0.13, 0.42\}$ in the driven spectrum, and are responsible for the secondary poles mentioned in the discussion of the coupling constants of the effective Hamiltonian, Fig. 6. Furthermore, as we detail in Appendix G, counter-rotating terms induce corrections when attempting an accurate comparison with spectroscopic data from experiments.

VI. CONCLUSION

In summary, we present two complementary methods for the analysis of parametrically activated two-qubit gates, one based on analytical time-dependent Schrieffer-Wolff perturbation theory, and one based on numerical Floquet methods. Although we mostly focus on coupler-mediated parametric $\sqrt{i\text{SWAP}}$ gates, a larger collection of gates can be generated in the same model Hamiltonian. The methods presented here allow one to efficiently evaluate the terms present in the effective gate Hamiltonian.

For the $\sqrt{i\text{SWAP}}$ interaction, we show that with experimentally accessible parameters, a gate frequency of approximately 40 MHz corresponding to a gate time as short as 12.5 ns can be obtained with vanishing dynamical cross-Kerr interaction. This fast gate is achieved by working with large capacitive couplings between the qubits and the coupler, while canceling the cross-Kerr interactions by setting the coupler anharmonicity to positive values, and choosing the right modulation amplitude. Optimization of realistic device parameters based on close agreements between the Floquet simulations and the experimental data will be published elsewhere [60].

We argue that the analytical method introduced here and which is based on a drive-dependent normal-mode expansion is a computationally efficient strategy to organize the perturbation theory as compared to an energy eigenbasis calculation, for it allows the parameters of the effective Hamiltonian to be obtained at lower orders in perturbation theory. Moreover, this strategy is suitable in the regime of comparatively large linear couplings, where the dispersive approximation breaks down. Nonetheless, we show that higher orders in analytical perturbation theory are needed for full agreement with exact numerical results, especially for higher-order interactions, such as the dynamical cross-Kerr. Generating higher-order contributions efficiently using computer algebra techniques is the subject of future studies. On the other hand, this work indicates that Floquet numerical methods, as compared to full time-dynamics simulations, is a numerically efficient and exact method for minute optimization studies of parametric gates.

ACKNOWLEDGMENTS

We thank Joachim Cohen, Charles Guinn, Moein Malekakhlagh, and Baptiste Royer for useful discussions. We are grateful to Ross Shillito for help with optimizing numerical simulations. This work is undertaken thanks to funding from NSERC, the Canada First Research Excellence Fund, and the U.S. Army Research Office Grant No. W911NF-18-1-0411.

APPENDIX A: TIME-DEPENDENT SCHRIEFFER-WOLFF TRANSFORMATION

To obtain equations for $\hat{G}_I(t)$, we assume that the generator can be expanded as a series in λ, that is

$$\hat{G}_I(t) = \lambda \hat{G}_I^{(1)}(t) + \lambda^2 \hat{G}_I^{(2)}(t) + \cdots, \tag{A1}$$

and collect powers of λ in the BCH expansion of Eq. (9)

$$e^{-\hat{G}_I}(\hat{H}_I - i\partial_t)e^{\hat{G}_I} = \lambda\hat{H}_I^{(1)} - i\lambda\dot{\hat{G}}_I^{(1)} + [\lambda\hat{H}_I^{(1)}, \lambda\hat{G}_I^{(1)}] - \frac{i}{2}[\lambda\dot{\hat{G}}_I^{(1)}, \lambda\hat{G}_I^{(1)}] - i\lambda^2\dot{\hat{G}}_I^{(2)} - i\partial_t + O(\lambda^3). \tag{A2}$$

The above expansion can be expressed compactly

$$e^{-\hat{G}_I}(\hat{H}_I - i\partial_t)e^{\hat{G}_I} = \sum_{k=1}^{\infty}\lambda^k\left[\hat{H}_I^{(k)}(t) - i\dot{\hat{G}}_I^{(k)}\right] - i\partial_t. \tag{A3}$$

Provided a prescription for $\lambda^k\hat{G}_I^{(k)}(t)$, we have a recursive way of determining higher-order corrections to the interaction Hamiltonian: knowledge of $\lambda\hat{H}_I^{(1)}(t)$ allows one to determine $\lambda^2\hat{H}_I^{(2)}$, then $\lambda^3\hat{H}_I^{(3)}$ etc.

The kth-order term in the generator, $\lambda^k\hat{G}_I^{(k)}(t)$, is determined by the condition that the Hamiltonian be free of oscillatory terms of order λ^k or less. This condition can be formulated explicitly if we write, as in Eq. (11),

$$\lambda^k\hat{H}_I^{(k)}(t) = \lambda^k\overline{\hat{H}}_I^{(k)} + \lambda^k\widetilde{\hat{H}}_I^{(k)}(t). \tag{A4}$$

Then oscillatory terms $\lambda^k\widetilde{\hat{H}}_I^{(k)}$ are canceled for every k if

$$\lambda^k\hat{G}_I^{(k)}(t) = \frac{1}{i}\int_0^t\lambda^k\widetilde{\hat{H}}_I^{(k)}(t). \tag{A5}$$

Note that, in the above expression, we impose the boundary condition $\hat{G}_I^{(k)}(0) = 0$ by specifying the lower limit of the integration. Noting that Eq. (A5) implies

$$\lambda^k\widetilde{\hat{G}}_I^{(k)}(t) = \frac{1}{i}\int^t\lambda^k\widetilde{\hat{H}}_I^{(k)}(t),$$
$$\lambda^k\overline{\hat{G}}_I^{(k)}(t) = \frac{1}{i}\left[\int^t\lambda^k\widetilde{\hat{H}}_I^{(k)}(t)\right]_{t=0}. \tag{A6}$$

The dc part of the generator, $\lambda^k\overline{\hat{G}}_I^{(k)}$, is nonvanishing here as a result of the boundary condition in Eq. (A5), as opposed to the zero time-average property of kick operators, to which the generator studied here is related [47].

With the above formalism in place, we are now ready to compute perturbative corrections. From Eq. (9) we identify the λ^2 correction to the interaction-picture Hamiltonian

$$\lambda^2\hat{H}_I^{(2)}(t) = [\lambda\hat{H}_I^{(1)}, \lambda\hat{G}_I^{(1)}] - \frac{i}{2}[\lambda\dot{\hat{G}}_I^{(1)}, \lambda\hat{G}_I^{(1)}]. \tag{A7}$$

Going ahead and solving the RWA condition in Eq. (A5) at order λ^1, we find the order-λ^2 RWA Hamiltonian

$$\lambda^2\overline{\hat{H}}_I^{(2)} = \frac{1}{i}\overline{\left[\overline{\hat{H}}_I^{(1)}, \int_0^t\lambda\widetilde{\hat{H}}_I^{(1)}(t')dt'\right]} + \frac{1}{2i}\overline{\left[\lambda\widetilde{\hat{H}}_I^{(1)}(t), \int_0^t\lambda\widetilde{\hat{H}}_I^{(1)}(t')dt'\right]}. \tag{A8}$$

This procedure can be iterated to higher orders, with increasing complexity due to the proliferation of terms from nested commutators in the BCH expansion.

APPENDIX B: CIRCUIT QUANTIZATION

In this Appendix we derive the model Hamiltonian of Eq. (1) from the circuit Lagrangian corresponding to Fig. 8. Assuming the individual modes of the junction array have small impedance, guaranteed by sufficiently large Josephson energy, the junction array can be described by an effective one-dimensional Lagrangian where the total phase difference across the array is spread evenly through the junctions. The effective one-dimensional Lagrangian associated with the bare coupler mode is

$$\mathcal{L}_c = \sum_{k=\alpha,\beta}\frac{C_k}{2}\dot{\boldsymbol{\phi}}_k^2 + \alpha E_{Jc}\cos\left[\boldsymbol{\varphi}_\alpha\right] + \beta N E_{Jc}\cos\left[\frac{\boldsymbol{\varphi}_\beta}{N}\right], \tag{B1}$$

where $\boldsymbol{\phi}_\alpha$ is the branch flux across the small junction and the shunt capacitor with total capacitance C_α, $\boldsymbol{\phi}_\beta$ is the branch flux across the junction array with effective capacitance C_β, and $\boldsymbol{\varphi}_k = 2\pi\boldsymbol{\phi}_k/\Phi_0$ are the associated reduced phase variables, and Φ_0 is the superconducting flux quantum. The phases $\boldsymbol{\varphi}_\alpha$ and $\boldsymbol{\varphi}_\beta$ are constrained by the fluxoid quantization, $\boldsymbol{\varphi}_\alpha + \boldsymbol{\varphi}_\beta = \varphi_{\text{ext}}$. We define the alternative coordinates

$$\boldsymbol{\varphi}_\alpha = \boldsymbol{\varphi}_c + \mu_\alpha\varphi_{\text{ext}},$$
$$\boldsymbol{\varphi}_\beta = -\boldsymbol{\varphi}_c - N\mu_\beta\varphi_{\text{ext}}, \tag{B2}$$

with $\mu_\alpha - N\mu_\beta = 1$, such that the capacitive energy in the Lagrangian is now purely quadratic in $\dot{\boldsymbol{\phi}}_c$. We thus require $C_\alpha\mu_\alpha + C_\beta N\mu_\beta = 0$. We obtain

$$\mu_\alpha = \frac{C_\beta}{C_\alpha + C_\alpha},$$
$$\mu_\beta = -\frac{1}{N}\frac{C_\alpha}{C_\alpha + C_\beta}. \tag{B3}$$

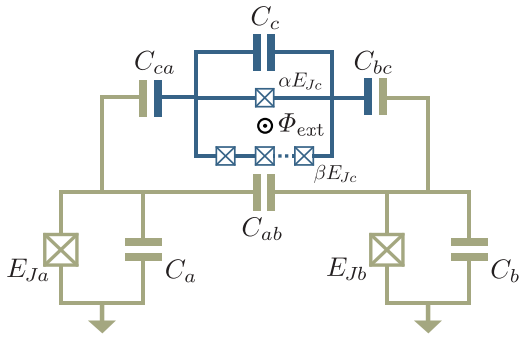


FIG. 8. Circuit schematic and notations used in the derivation of the circuit Lagrangian in Appendix B. The coupler consists of two branches of total capacitances C_α and C_β (not indicated in the figure). The α branch consists of a single Josephson junction, while the 'β' branch contains N junctions in series. The bare coupler and transmon modes are connected capacitively through coupling capacitances $C_{ab,bc,ca}$.

Up to time-dependent scalar terms, we obtain the form

$$\mathcal{L}_c = \frac{C_c}{2}\dot{\boldsymbol{\phi}}_c^2 + \alpha E_{Jc}\cos\left[\boldsymbol{\varphi}_c + \mu_\alpha\varphi_{\text{ext}}\right] + \beta N E_{Jc}\cos\left[\frac{\boldsymbol{\varphi}_c}{N} + \mu_\beta\varphi_{\text{ext}}\right]. \tag{B4}$$

Moreover, for the two bare transmon modes $j = a, b$ the Lagrangian reads

$$\mathcal{L}_j = \frac{C_j}{2}\dot{\boldsymbol{\phi}}_j^2 + E_{Jj}\cos\boldsymbol{\varphi}_j. \tag{B5}$$

The total Lagrangian of the system then takes the form

$$\mathcal{L} = \mathcal{L}_a + \mathcal{L}_b + \mathcal{L}_c + \mathcal{L}_g, \tag{B6}$$

where we introduce the capacitive coupling between the three bare modes

$$\mathcal{L}_g = \frac{C_{ab}}{2}\dot{\boldsymbol{\phi}}_a\dot{\boldsymbol{\phi}}_b + \frac{C_{bc}}{2}\dot{\boldsymbol{\phi}}_b\dot{\boldsymbol{\phi}}_c + \frac{C_{ca}}{2}\dot{\boldsymbol{\phi}}_c\dot{\boldsymbol{\phi}}_a. \tag{B7}$$

APPENDIX C: PERTURBATION THEORY FOR THE TOY MODEL

In this section, we reproduce expressions for the cross-Kerr interaction obtained to second-order in perturbation theory for the toy model. The full expression of the second-order RWA correction to the cross-Kerr interaction in Sec. III B reads

$$\begin{aligned}\chi^{(2)}_{ab,\text{Sec. III B}} = &\frac{4\left(\sum_{j=a,b,c}u_{aj}^2u_{bj}u_{cj}\boldsymbol{\alpha}_j\right)^2}{\omega_b-\omega_c} + \frac{4\left(\sum_{j=a,b,c}u_{aj}u_{bj}^2u_{cj}\boldsymbol{\alpha}_j\right)^2}{\omega_a-\omega_c} + \frac{2\left(\sum_{j=a,b,c}u_{aj}u_{bj}u_{cj}^2\boldsymbol{\alpha}_j\right)^2}{\omega_a+\omega_b-2\omega_c} \\ &- \frac{2\left(\sum_{j=a,b,c}u_{aj}^3u_{bj}\boldsymbol{\alpha}_j\right)^2}{\omega_a-\omega_b} + \frac{2\left(\sum_{j=a,b,c}u_{aj}u_{bj}^3\boldsymbol{\alpha}_j\right)^2}{\omega_a-\omega_b} + \frac{u_{ac}u_{bc}\sum_{j=a,b,c}u_{aj}u_{bj}\left(u_{aj}^2-u_{bj}^2\right)}{\omega_a-\omega_b}\delta.\end{aligned} \tag{C1}$$

The second-order correction to the static cross-Kerr interaction as calculated in Sec. III C is

$$\begin{aligned}\chi^{(2)}_{ab,\text{Sec. III C}} = &\frac{4\left(\sum_{j=a,b,c}u_{aj}^2u_{bj}u_{cj}\boldsymbol{\alpha}_j\right)^2}{\omega_b-\omega_c+\sum_{j=a,b,c}2u_{aj}^2\left(u_{bj}^2-u_{cj}^2\right)\boldsymbol{\alpha}_j} + \frac{4\left(\sum_{j=a,b,c}u_{aj}u_{bj}^2u_{cj}\boldsymbol{\alpha}_j\right)^2}{\omega_a-\omega_c+\sum_{j=a,b,c}2u_{bj}^2\left(u_{aj}^2-u_{cj}^2\right)\boldsymbol{\alpha}_j} \\ &+ \frac{2\left(\sum_{j=a,b,c}u_{aj}u_{bj}u_{cj}^2\boldsymbol{\alpha}_j\right)^2}{\omega_a+\omega_b-2\omega_c+\left(2u_{aj}^2u_{bj}^2-u_{cj}^4\right)\boldsymbol{\alpha}_j} - \frac{2\left(\sum_{j=a,b,c}u_{aj}^3u_{bj}\boldsymbol{\alpha}_j\right)^2}{\omega_a-\omega_b+\sum_{j=a,b,c}\left(u_{aj}^4-2u_{aj}^2u_{bj}^2\right)\boldsymbol{\alpha}_j} \\ &+ \frac{2\left(\sum_{j=a,b,c}u_{aj}u_{bj}^3\boldsymbol{\alpha}_j\right)^2}{\omega_a-\omega_b+\sum_{j=a,b,c}\left(2u_{aj}^2u_{bj}^2-u_{bj}^4\right)\boldsymbol{\alpha}_a}.\end{aligned} \tag{C2}$$

The expression for the dynamical cross-Kerr interaction, $\chi^{(2)}_{ab,\text{Sec. III C}}$ at $\delta \neq 0$, is available from the formalism, but it is too lengthy to be reproduced here. In the main text, an evaluation of this expression is used in making direct comparisons to exact numerics.

APPENDIX D: DETAILS FOR FULL-CIRCUIT HAMILTONIAN

In this Appendix, we record a number of results used in Sec. IV, in particular solutions to classical equations of motion in Appendix D 1, the formulae used for normal-ordered expansions in Appendix D 2, and the time-dependent terms in the coupler Hamiltonian in Appendix D 3.

1. Classical equations of motion

We consider a time-dependent unitary displacement of Eq. (1) according to which

$$\hat{\boldsymbol{\varphi}}_j \rightarrow \hat{\boldsymbol{\varphi}}_j + \varphi_j(t),\ \hat{\mathbf{n}}_j \rightarrow \hat{\mathbf{n}}_j + n_j(t), \tag{D1}$$

with $j = a, b, c$. Requiring that the Taylor expansion of the displaced Hamiltonian do not contain any linear terms amounts to having $\varphi_j(t), n_j(t)$ obey the classical equations of motion associated with the Hamiltonian of Eq. (1). We obtain these by writing down the six Heisenberg equations of motion $d\hat{\boldsymbol{\varphi}}_j/dt = i[\hat{H}, \hat{\boldsymbol{\varphi}}_j]$, $d\hat{\mathbf{n}}_j/dt = i[\hat{H}, \hat{\mathbf{n}}_j]$ for $j = a, b, c$, then passing Heisenberg-picture operators to classical variables $\hat{\boldsymbol{\varphi}}_j(t) \rightarrow \varphi_j(t), \hat{\mathbf{n}}_j(t) \rightarrow n_j(t)$.

$$\begin{aligned}
\dot{\varphi}_a &= 8E_{Ca}n_a + 4E_{Cca}n_c + 4E_{Cab}n_b,\\
\dot{\varphi}_b &= 4E_{Cab}n_a + 8E_{Cb}n_b + 4E_{Cbc}n_c,\\
\dot{\varphi}_c &= 4E_{Cca}n_a + 4E_{Cbc}n_b + 8E_{Cc}n_c,\\
\dot{n}_a &= -E_{Ja}\sin(\varphi_a),\\
\dot{n}_b &= -E_{Jb}\sin(\varphi_b),\\
\dot{n}_c &= -\alpha E_{Jc}\sin\left[\varphi_c + \mu_\alpha\varphi_{\text{ext}}(t)\right]\\
&\quad - \beta E_{Jc}\sin\left[\frac{\varphi_c}{N} + \mu_\beta\varphi_{\text{ext}}(t)\right].
\end{aligned} \tag{D2}$$

Differentiating the first three equations we eliminate the charge coordinates to obtain a set of second-order equations for the phase coordinates

$$\begin{aligned}
&\ddot{\varphi}_a + \omega_{pa}^2\sin\varphi_a + 4E_{Cab}E_{Jb}\sin\varphi_b + 4E_{Cca}E_{Jc}\left\{\alpha\sin\left[\varphi_c + \mu_\alpha\varphi_{\text{ext}}(t)\right] + \beta\sin\left[\frac{\varphi_c}{N} + \mu_\beta\varphi_{\text{ext}}(t)\right]\right\} = 0,\\
&\ddot{\varphi}_b + \omega_{pb}^2\sin\varphi_b + 4E_{Cab}E_{Ja}\sin\varphi_a + 4E_{Cbc}E_{Jc}\left\{\alpha\sin\left[\varphi_c + \mu_\alpha\varphi_{\text{ext}}(t)\right] + \beta\sin\left[\frac{\varphi_c}{N} + \mu_\beta\varphi_{\text{ext}}(t)\right]\right\} = 0,\\
&\ddot{\varphi}_c + \alpha\omega_{pc}^2\sin\left[\varphi_c + \mu_\alpha\varphi_{\text{ext}}(t)\right] + \beta\omega_{pc}^2\sin\left[\frac{\varphi_c}{N} + \mu_\beta\varphi_{\text{ext}}(t)\right] + 4E_{Cca}E_{Ja}\sin\varphi_a + 4E_{Cbc}E_{Jb}\sin\varphi_b = 0,
\end{aligned} \tag{D3}$$

where we define three plasma frequencies $\omega_{pj} = \sqrt{8E_{Cj}E_{Jj}}$ for $j = a, b, c$. These equations can be solved approximately by considering a trial form

$$\varphi_j = \zeta_j\sin(\omega_d t) + \xi_j, \tag{D4}$$

and equating coefficients of the zeroth and first harmonics of the drive frequency ω_d. This leads to six coupled transcendental equations

$$\begin{aligned}
&\omega_{pa}^2\sin(\xi_a)J_0(\zeta_a) + 4E_{Cab}E_{Jb}\sin(\xi_b)J_0(\zeta_b)\\
&\quad + 4E_{Cca}E_{Jc}\left[\alpha\sin(\xi_c + \mu_\alpha\overline{\varphi}_{\text{ext}})J_0(\zeta_c + \mu_\alpha\delta\varphi) + \beta\sin\left(\frac{\xi_c}{N} + \mu_\beta\overline{\varphi}_{\text{ext}}\right)J_0\left(\frac{\zeta_c}{N} + \mu_\beta\delta\varphi\right)\right] = 0,\\
&\omega_{pb}^2\sin(\xi_b)J_0(\zeta_b) + 4E_{Cab}E_{Ja}\sin(\xi_a)J_0(\zeta_a)\\
&\quad + 4E_{Cbc}E_{Jc}\left[\alpha\sin\left(\xi_c + \mu_\alpha\overline{\varphi}_{\text{ext}}\right)J_0\left(\zeta_c + \mu_\alpha\delta\varphi\right) + \beta\sin\left(\frac{\xi_c}{N} + \mu_\beta\overline{\varphi}_{\text{ext}}\right)J_0\left(\frac{\zeta_c}{N} + \mu_\beta\delta\varphi\right)\right] = 0,\\
&\alpha\omega_{pc}^2\sin(\xi_c + \mu_\alpha\overline{\varphi}_{\text{ext}})J_0(\zeta_c + \mu_\alpha\delta\varphi) + \beta\omega_{pc}^2\sin\left(\frac{\xi_c}{N} + \mu_\beta\overline{\varphi}_{\text{ext}}\right)J_0\left(\frac{\zeta_c}{N} + \mu_\alpha\delta\varphi\right)\\
&\quad + 4E_{Cca}E_{Ja}\sin(\xi_a)J_0(\zeta_a) + 4E_{Cbc}E_{Jb}\sin(\xi_b)J_0(\zeta_b) = 0,\\
&\quad - \omega_d^2\zeta_a + 2\omega_{pa}^2\cos(\xi_a)J_1(\zeta_a) = 0,\\
&\quad - \omega_d^2\zeta_b + 2\omega_{pb}^2\cos(\xi_b)J_1(\zeta_b) = 0,\\
&\quad - \omega_d^2\zeta_c + 2\alpha\omega_{pc}^2\cos(\xi_c + \mu_\alpha\overline{\varphi}_{\text{ext}})J_1(\zeta_c + \mu_\alpha\delta\varphi) + 2\beta\omega_{pc}^2\cos\left(\frac{\xi_c}{N} + \mu_\beta\overline{\varphi}_{\text{ext}}\right)J_1\left(\frac{\zeta_c}{N} + \mu_\beta\delta\varphi\right) = 0.
\end{aligned} \tag{D5}$$

The equations above are solved numerically by searching for the root closest to the response of the decoupled system ($E_{Cab} = E_{Cca} = E_{Cbc} = 0$) to a static external field ($\delta\varphi = 0$), i.e., zero amplitude response $\zeta_j = 0$, in addition to $\xi_{a,b} = 0$ and ξ_c the minimizer of the static potential of the coupler defined by the current conservation condition

$$\alpha \sin(\xi_c + \mu_\alpha \overline{\varphi}_{\rm ext}) + \beta \sin\left(\frac{\xi_c}{N} + \mu_\beta \overline{\varphi}_{\rm ext}\right) = 0. \quad \text{(D6)}$$

We finally make the assumption that displacing the quadratures by the classical solutions obtained above will remove from the Hamiltonian, to a good approximation, the terms that are linear in the quadratures $\hat{\boldsymbol{\varphi}}_j, \hat{\mathbf{n}}_j$ for $j = a, b, c$.

2. Normal-ordered expansions of trigonometric functions. Jacobi-Anger expansions

Sine and cosine are expanded in normal order using the following two expressions [61] [recall that $\hat{\varphi}_a = \sqrt{\eta_a/2}(\hat{a} + \hat{a}^\dagger)$]:

$$\cos\hat{\varphi}_a = e^{-\eta_a/4} \sum_{\substack{m,n\geq 0 \\ m+n=\text{ even}}} \frac{\left(-\frac{\eta_a}{2}\right)^{m+n/2} \hat{a}^{\dagger m}\hat{a}^n}{m!n!},$$
$$\sin\hat{\varphi}_a = e^{-\eta_a/4}\sqrt{\frac{\eta_a}{2}} \sum_{\substack{m,n\geq 0 \\ m+n=\text{ odd}}} \frac{\left(-\frac{\eta_a}{2}\right)^{m+n-1/2} \hat{a}^{\dagger m}\hat{a}^n}{m!n!}, \quad \text{(D7)}$$

with analogous expressions for the operators $\hat{b}$ and $\hat{c}$.

3. Time-dependent terms in the coupler Hamiltonian

Terms corresponding to the Jacobi-Anger expansion up to the second harmonic of the drive in the bare coupler Hamiltonian $\widetilde{H}_c(t)$ in Sec. IV are listed in Table II. The operator monomial at the beginning of each row is to be multiplied by the sum of the two following columns, and then results from all rows are to be summed. The coefficients of the missing monomials $\hat{\mathbf{c}}, \hat{\mathbf{c}}^2, \hat{\mathbf{c}}^3, \hat{\mathbf{c}}^\dagger\hat{\mathbf{c}}^2, \hat{\mathbf{c}}^4, \hat{\mathbf{c}}^\dagger\hat{\mathbf{c}}^3$ are obtained by Hermitian conjugation.

APPENDIX E: NORMAL-MODE TRANSFORMATION

In Sec. IV, we make use of a normal-mode transformation that eliminates the off-diagonal capacitive coupling terms from the time-independent quadratic Hamiltonian. In this section we provide the steps to obtain the normal-mode coefficients.

Consider the quadratic form (repeated indices are summed over):

$$\hat{H} = A_{\alpha\beta}\hat{\bar{n}}_\alpha\hat{\bar{n}}_\beta + B_{\alpha\beta}\hat{\bar{\varphi}}_\alpha\hat{\bar{\varphi}}_\beta. \quad \text{(E1)}$$

We make a simplification by assuming that there are no off-diagonal inductive terms, $B_{\alpha\beta} \propto \delta_{\alpha\beta}$, which is valid for the circuit studied here. The diagonalization involves three steps:

Step 1. Rescale the variables so that the diagonal part of the Hamiltonian, the inductive part, contains terms with the *same* inductive energy. For this, let us define the square root of the product of the inductive energies $B = \left(\prod_\alpha B_{\alpha\alpha}\right)^{1/2}$ and the dimensionless coefficients

TABLE II. Time-dependent terms, up to quartics, in the bare coupler Hamiltonian.

Monomial	$J_1(\delta)$	$J_2(\delta)$
$\hat{\mathbf{c}}^\dagger$	$\sqrt{2}\alpha\epsilon e^{-\eta_c/4}\sqrt{\eta_c}E_{Jc}J_1\left(\delta\mu_\alpha\right)\sin\left(t\omega_d\right)\cos\left(\mu_\alpha\overline{\varphi}_{\rm ext}\right) +$ $\sqrt{2}\beta N e^{-\eta_c/4N^2}\sqrt{\eta_c/N^2}E_{Jc}J_1\left(\delta\mu_\beta\right)\sin\left(t\omega_d\right)\cos\left(\mu_\beta\overline{\varphi}_{\rm ext}\right)$	$\sqrt{2}\alpha\epsilon e^{-\eta_c/4}\sqrt{\eta_c}E_{Jc}J_2\left(\delta\mu_\alpha\right)\cos\left(2t\omega_d\right)\sin\left(\mu_\alpha\overline{\varphi}_{\rm ext}\right) +$ $\sqrt{2}\beta N e^{-\eta_c/4N^2}\sqrt{\eta_c/N^2}E_{Jc}J_2\left(\delta\mu_\beta\right)\cos\left(2t\omega_d\right)\sin\left(\mu_\beta\overline{\varphi}_{\rm ext}\right)$
$\hat{\mathbf{c}}^\dagger\hat{\mathbf{c}}^\dagger$	$-\frac{1}{2}\alpha\epsilon e^{-\eta_c/4}\eta_c E_{Jc}J_1\left(\delta\mu_\alpha\right)\sin\left(t\omega_d\right)\sin\left(\mu_\alpha\overline{\varphi}_{\rm ext}\right)$ $-\beta\eta_c e^{-\frac{\eta_c}{4N^2}}E_{Jc}J_1\left(\delta\mu_\beta\right)\sin\left(t\omega_d\right)\sin\left(\mu_\beta\overline{\varphi}_{\rm ext}\right)/2N$	$\frac{1}{2}\alpha\epsilon e^{-\eta_c/4}\eta_c E_{Jc}J_2\left(\delta\mu_\alpha\right)\cos\left(2t\omega_d\right)\cos\left(\mu_\alpha\overline{\varphi}_{\rm ext}\right)$ $+\beta\eta_c e^{-\frac{\eta_c}{4N^2}}E_{Jc}J_2\left(\delta\mu_\beta\right)\cos\left(2t\omega_d\right)\cos\left(\mu_\beta\overline{\varphi}_{\rm ext}\right)/2N$
$\hat{\mathbf{c}}^\dagger\hat{\mathbf{c}}$	$-\alpha\epsilon e^{-\eta_c/4}\eta_c E_{Jc}J_1\left(\delta\mu_\alpha\right)\sin\left(t\omega_d\right)\sin\left(\mu_\alpha\overline{\varphi}_{\rm ext}\right)$ $-\beta\eta_c e^{-\frac{\eta_c}{4N^2}}E_{Jc}J_1\left(\delta\mu_\beta\right)\sin\left(t\omega_d\right)\sin\left(\mu_\beta\overline{\varphi}_{\rm ext}\right)/N$	$+\alpha\epsilon e^{-\eta_c/4}\eta_c E_{Jc}J_2\left(\delta\mu_\alpha\right)\cos\left(2t\omega_d\right)\cos\left(\mu_\alpha\overline{\varphi}_{\rm ext}\right)$ $+\beta\eta_c e^{-\frac{\eta_c}{4N^2}}E_{Jc}J_2\left(\delta\mu_\beta\right)\cos\left(2t\omega_d\right)\cos\left(\mu_\beta\overline{\varphi}_{\rm ext}\right)/N$
$\hat{\mathbf{c}}^\dagger\hat{\mathbf{c}}^\dagger\hat{\mathbf{c}}^\dagger$	$-\alpha\epsilon e^{-\frac{\eta_c}{4}}\eta_c^{3/2}E_{Jc}J_1\left(\delta\mu_\alpha\right)\sin\left(t\omega_d\right)\cos\left(\mu_\alpha\overline{\varphi}_{\rm ext}\right)/6\sqrt{2}$ $-\beta\eta_c e^{-\frac{\eta_c}{4N^2}}\sqrt{\frac{\eta_c}{N^2}}E_{Jc}J_1\left(\delta\mu_\beta\right)\sin\left(t\omega_d\right)\cos\left(\mu_\beta\overline{\varphi}_{\rm ext}\right)/6\sqrt{2}N$	$-\alpha\epsilon e^{-\frac{\eta_c}{4}}\eta_c^{3/2}E_{Jc}J_2\left(\delta\mu_\alpha\right)\cos\left(2t\omega_d\right)\sin\left(\mu_\alpha\overline{\varphi}_{\rm ext}\right)/6\sqrt{2}$ $-\beta\eta_c e^{-\frac{\eta_c}{4N^2}}\sqrt{\frac{\eta_c}{N^2}}E_{Jc}J_2\left(\delta\mu_\beta\right)\cos\left(2t\omega_d\right)\sin\left(\mu_\beta\overline{\varphi}_{\rm ext}\right)/6\sqrt{2}N$
$\hat{\mathbf{c}}^\dagger\hat{\mathbf{c}}^\dagger\hat{\mathbf{c}}$	$-\alpha\epsilon e^{-\frac{\eta_c}{4}}\eta_c^{3/2}E_{Jc}J_1\left(\delta\mu_\alpha\right)\sin\left(t\omega_d\right)\cos\left(\mu_\alpha\overline{\varphi}_{\rm ext}\right)/2\sqrt{2}$ $-\beta\eta_c e^{-\frac{\eta_c}{4N^2}}\sqrt{\frac{\eta_c}{N^2}}E_{Jc}J_1\left(\delta\mu_\beta\right)\sin\left(t\omega_d\right)\cos\left(\mu_\beta\overline{\varphi}_{\rm ext}\right)/2\sqrt{2}N$	$-\alpha\epsilon e^{-\frac{\eta_c}{4}}\eta_c^{3/2}E_{Jc}J_2\left(\delta\mu_\alpha\right)\cos\left(2t\omega_d\right)\sin\left(\mu_\alpha\overline{\varphi}_{\rm ext}\right)/2\sqrt{2}$ $-\beta\eta_c e^{-\frac{\eta_c}{4N^2}}\sqrt{\frac{\eta_c}{N^2}}E_{Jc}J_2\left(\delta\mu_\beta\right)\cos\left(2t\omega_d\right)\sin\left(\mu_\beta\overline{\varphi}_{\rm ext}\right)/2\sqrt{2}N$
$\hat{\mathbf{c}}^\dagger\hat{\mathbf{c}}^\dagger\hat{\mathbf{c}}^\dagger\hat{\mathbf{c}}^\dagger$	$1/48\alpha\epsilon e^{-\eta_c/4}\eta_c^2 E_{Jc}J_1\left(\delta\mu_\alpha\right)\sin\left(t\omega_d\right)\sin\left(\mu_\alpha\overline{\varphi}_{\rm ext}\right)$ $+\beta\eta_c^2 e^{-\frac{\eta_c}{4N^2}}E_{Jc}J_1\left(\delta\mu_\beta\right)\sin\left(t\omega_d\right)\sin\left(\mu_\beta\overline{\varphi}_{\rm ext}\right)/48N^3$	$-1/48\alpha\epsilon e^{-\eta_c/4}\eta_c^2 E_{Jc}J_2\left(\delta\mu_\alpha\right)\cos\left(2t\omega_d\right)\cos\left(\mu_\alpha\overline{\varphi}_{\rm ext}\right)$ $-\beta\eta_c^2 e^{-\frac{\eta_c}{4N^2}}E_{Jc}J_2\left(\delta\mu_\beta\right)\cos\left(2t\omega_d\right)\cos\left(\mu_\beta\overline{\varphi}_{\rm ext}\right)/48N^3$
$\hat{\mathbf{c}}^\dagger\hat{\mathbf{c}}^\dagger\hat{\mathbf{c}}^\dagger\hat{\mathbf{c}}$	$1/12\alpha\epsilon e^{-\eta_c/4}\eta_c^2 E_{Jc}J_1\left(\delta\mu_\alpha\right)\sin\left(t\omega_d\right)\sin\left(\mu_\alpha\overline{\varphi}_{\rm ext}\right)$ $+\beta\eta_c^2 e^{-\frac{\eta_c}{4N^2}}E_{Jc}J_1\left(\delta\mu_\beta\right)\sin\left(t\omega_d\right)\sin\left(\mu_\beta\overline{\varphi}_{\rm ext}\right)/12N^3$	$-1/12\alpha\epsilon e^{-\eta_c/4}\eta_c^2 E_{Jc}J_2\left(\delta\mu_\alpha\right)\cos\left(2t\omega_d\right)\cos\left(\mu_\alpha\overline{\varphi}_{\rm ext}\right)$ $-\beta\eta_c^2 e^{-\frac{\eta_c}{4N^2}}E_{Jc}J_2\left(\delta\mu_\beta\right)\cos\left(2t\omega_d\right)\cos\left(\mu_\beta\overline{\varphi}_{\rm ext}\right)/12N^3$
$\hat{\mathbf{c}}^\dagger\hat{\mathbf{c}}^\dagger\hat{\mathbf{c}}\hat{\mathbf{c}}$	$+1/8\alpha\epsilon e^{-\eta_c/4}\eta_c^2 E_{Jc}J_1\left(\delta\mu_\alpha\right)\sin\left(t\omega_d\right)\sin\left(\mu_\alpha\overline{\varphi}_{\rm ext}\right)$ $+\beta\eta_c^2 e^{-\frac{\eta_c}{4N^2}}E_{Jc}J_1\left(\delta\mu_\beta\right)\sin\left(t\omega_d\right)\sin\left(\mu_\beta\overline{\varphi}_{\rm ext}\right)/8N^3$	$-1/8\alpha\epsilon e^{-\eta_c/4}\eta_c^2 E_{Jc}J_2\left(\delta\mu_\alpha\right)\cos\left(2t\omega_d\right)\cos\left(\mu_\alpha\overline{\varphi}_{\rm ext}\right)$ $-\beta\eta_c^2 e^{-\frac{\eta_c}{4N^2}}E_{Jc}J_2\left(\delta\mu_\beta\right)\cos\left(2t\omega_d\right)\cos\left(\mu_\beta\overline{\varphi}_{\rm ext}\right)/8N^3$

$f_\alpha = \sqrt{B/B_{\alpha\alpha}}$. Then we introduce alternative canonically conjugate coordinates:

$$\hat{\varphi}'_\alpha = f_\alpha^{-1}\hat{\tilde{\varphi}}_\alpha, \qquad \hat{n}'_\alpha = f_\alpha \hat{\tilde{n}}_\alpha. \tag{E2}$$

In terms of the alternative coordinates, and letting $A'_{\alpha\beta} = A_{\alpha\beta}/(f_\alpha f_\beta)$ (no implicit summation), we have

$$\hat{H} = A'_{\alpha\beta}\hat{n}'_\alpha \hat{n}'_\beta + B\delta_{\alpha\beta}\hat{\varphi}'_\alpha\hat{\varphi}'_\beta. \tag{E3}$$

Step 2. Diagonalize the capacitive coupling matrix A'. We assume here that this is possible and is achieved by an orthonormal matrix S, such that

$$A'_{\alpha\beta} = (S^T)_{\alpha\mu}D_{\mu\nu}S_{\nu\beta} = S_{\mu\alpha}D_{\mu\nu}S_{\nu\beta}, \tag{E4}$$

with $D_{\mu\nu}$ a diagonal matrix. Rewriting the above as $A'_{\alpha\beta}(S^T)_{\beta\gamma} = (S^T)_{\alpha\mu}D_{\mu\nu}S_{\nu\beta}(S^T)_{\beta\gamma} = (S^T)_{\alpha\mu}D_{\mu\gamma}$, or $\mathbf{A}' \cdot \mathbf{S}^T = \mathbf{S}^T \cdot \mathbf{D}$, then the matrix $\mathbf{S}$ contains the eigenvectors of $\mathbf{A}'$ on its *rows*. This diagonalization leads to

$$\hat{H} = S_{\mu\alpha}D_{\mu\nu}S_{\nu\beta}\hat{n}'_\alpha\hat{n}'_\beta + B\delta_{\alpha\beta}\hat{\varphi}'_\alpha\hat{\varphi}'_\beta. \tag{E5}$$

Inspecting the first term, we again define alternative coordinates

$$\hat{n}''_\mu = S_{\mu\alpha}\hat{n}'_\alpha, \qquad \hat{\varphi}''_\mu = S_{\mu\alpha}\hat{\varphi}'_\alpha. \tag{E6}$$

One can verify that the alternative double-primed coordinates are canonically conjugate because the transformation is orthonormal: $[\hat{n}''_\mu, \hat{\varphi}''_\nu] = S_{\mu\alpha}S_{\nu\beta}[\hat{n}'_\alpha, \hat{\varphi}'_\beta] = iS_{\mu\alpha}S_{\nu\beta}\delta_{\alpha\beta} = iS_{\mu\alpha}S_{\nu\alpha} = i\delta_{\mu\nu}$. With this, we obtain a diagonal form for the Hamiltonian

$$\hat{H} = \hat{n}''_\alpha D_{\alpha\beta}\hat{n}''_\beta + B\delta_{\alpha\beta}\hat{\varphi}''_\alpha\hat{\varphi}''_\beta, \tag{E7}$$

where in the second term we use the fact that the orthogonal transformation preserves the inner product.

Step 3. Finally, we need to undo the rescaling transformation of step 1. That is, introduce a third and last pair of canonically conjugate coordinates, the *normal-mode coordinates*

$$\hat{\varphi}_\alpha = f_\alpha\hat{\varphi}''_\alpha, \qquad \hat{n}_\alpha = f_\alpha^{-1}\hat{n}''_\alpha. \tag{E8}$$

At last the quadratic Hamiltonian reads

$$\begin{aligned}\hat{H} &= \hat{n}_\alpha f_\alpha f_\beta D_{\alpha\beta}\hat{n}_\beta + \frac{B\delta_{\alpha\beta}}{f_\alpha f_\beta}\hat{\varphi}_\alpha\hat{\varphi}_\beta \\ &= \hat{n}_\alpha f_\alpha f_\beta D_{\alpha\beta}\hat{n}_\beta + \hat{\varphi}_\alpha B_{\alpha\beta}\hat{\varphi}_\beta.\end{aligned} \tag{E9}$$

This is the final normal-mode Hamiltonian.

Hybridization coefficients. It is helpful to summarize the normal-mode transformation by skipping over the intermediate variables (primed, and double primed). For this we have to invert the definitions of the intermediate coordinates to obtain

$$\begin{aligned}\hat{\tilde{\varphi}}_\alpha &= \sum_\beta f_\alpha S_{\beta\alpha} f_\beta^{-1}\hat{\varphi}_\beta \equiv \sum_\beta U_{\alpha\beta}\hat{\varphi}_\beta, \\ \hat{\tilde{n}}_\alpha &= \sum_\beta f_\alpha^{-1} S_{\beta\alpha} f_\beta \hat{n}_\beta \equiv \sum_\beta V_{\alpha\beta}\hat{n}_\beta,\end{aligned} \tag{E10}$$

where we use $\hat{\varphi}'_\alpha = S_{\mu\alpha}\hat{\varphi}''_\mu$ and $\hat{n}'_\alpha = S_{\mu\alpha}\hat{n}''_\mu$. Note that $\mathbf{U} \cdot \mathbf{V}^T = 1$, i.e., the transformation from bare to normal modes is canonical.

Creation and annihilation operators. Lastly, we consider the creation and annihiliation operators. In order for squeezing terms to disappear in the Hamiltonian, we need

$$\begin{aligned}\hat{\tilde{\varphi}}_\alpha &= \sum_{\beta=a,b,c} \frac{u_{\alpha\beta}}{\sqrt{2}}(\hat{\beta} + \hat{\beta}^\dagger), \\ \hat{\tilde{n}}_\alpha &= \sum_{\beta=a,b,c} \frac{v_{\alpha\beta}}{i\sqrt{2}}(\hat{\beta} - \hat{\beta}^\dagger),\end{aligned} \tag{E11}$$

where

$$u_{\alpha\beta} = U_{\alpha\beta}\sqrt{\epsilon_\beta},\ v_{\alpha\beta} = \frac{V_{\alpha\beta}}{\sqrt{\epsilon_\beta}}. \tag{E12}$$

Finally, we obtain the hybridization coefficients entering Eq. (44) in the main text. The approach given in this Appendix generalizes to an arbitrary number of modes with off-diagonal coupling in either the capacitive matrix, or in the inductive matrix.

APPENDIX F: FLOQUET THEORY

This Appendix provides a practical summary of Floquet theory. The spectrum of a monochromatically driven system can be obtained from the Floquet formalism [20], according to which the time-dependent Schrödinger equation for a periodically driven Hamiltonian $\hat{H}(t) = \hat{H}(t + 2\pi/\omega_d)$ can be recast into a numerically solvable eigenproblem for the so-called Floquet Hamiltonian [19]

$$\left[\hat{H}(t) - i\partial_t\right]|\phi_\alpha(t)\rangle = \epsilon_\alpha\,|\phi_\alpha(t)\rangle\,. \tag{F1}$$

The eigenvalues are the *quasienergies* ϵ_α, and whose eigenvectors are the Floquet modes, which are periodic functions of time with $|\phi_\alpha(t)\rangle = |\phi_\alpha(t + 2\pi/\omega_d)\rangle$. In terms of these, the solution to the time-dependent Schrödinger is $|\psi_\alpha(t)\rangle = e^{-i\epsilon_\alpha t}\,|\phi_\alpha(t)\rangle$. Of note, the solutions to Eq. (F1) are only defined up to an integer multiple k of the drive frequency ω_d, for if $\{\epsilon_\alpha, |\phi_\alpha(t)\rangle\}$ is a solution, then so is $\{\epsilon_{\alpha k} \equiv \epsilon_\alpha + k\omega_d, |\phi_{\alpha k}(t)\rangle = e^{-i\omega_d t}\,|\phi_\alpha(t)\rangle\}$, which is a consequence of the periodicity of the Floquet modes.

Information about the monochromatically driven system can be obtained from the quasienergy spectra. For example, two-tone spectroscopy experiments where a weak tone is used to probe the spectra of the driven system can be modeled in the linear response regime [23]. In such experiments, probe-tone-induced transitions occur at frequency differences

$$\Delta_{\alpha\beta k} = \epsilon_\alpha - \epsilon_\beta + k\omega_d, \tag{F2}$$

provided that the operator corresponding to the probe tone, denoted generically as $\hat{X}$, has a nonzero matrix element between the corresponding Floquet modes. With the above notation, the corresponding matrix elements read

$$X_{\alpha\beta k} = \frac{1}{T}\int_0^T dt\, e^{-ik\omega_d t}\langle\phi_\beta(t)|\hat{X}|\phi_\alpha(t)\rangle, \tag{F3}$$

where $T = 2\pi/\omega_d$ is the period of the drive. This takes the form of a Fourier series coefficient $f_k = 1/T\int_0^T dt' e^{-ik(2\pi/T)t'} f(t')$ of the matrix element of the operator X between the two Floquet modes $\left|\phi_{\alpha,\beta}(t)\right\rangle$.

Numerically, the Floquet spectrum is efficiently obtained from the time-evolution operator over one period of the drive, which has a compact expression in terms of the Floquet modes [20]

$$\begin{aligned}\hat{U}(t+T,t) &= \mathcal{T} e^{-i\int_t^{t+T}\hat{H}(t')dt'} \\ &= \sum_\alpha e^{-i\epsilon_\alpha T}\,|\phi_\alpha(t)\rangle\,\langle\phi_\alpha(t)|\,,\end{aligned} \tag{F4}$$

where $\mathcal{T}$ is the time-ordering operator. According to the above expression, the Floquet modes at time $t = 0$, $|\phi_\alpha(0)\rangle$, are the eigenvectors of $U(T,0)$, whereas the quasienergies are obtained modulo an integer multiple of ω_d from the eigenvalues. The time-dependence over one period of the drive is obtained by propagating each mode $|\phi_\alpha(0)\rangle$ with the time-evolution operator $\hat{U}(t,0)$ in the interval $0 < t \leq T$.

To summarize, the steady-state dynamics can be obtained from the propagator $\hat{U}(t,0)$ over a single period of the drive, which makes the Floquet method an efficient alternative to numerical simulation of the dynamics over the complete gate time. Indeed, the period of the drive, on the order of 1 ns is between 2 to 3 orders of magnitude shorter than the typical gate times. In this work we obtain the quantities above by using the QuTip implementation of the Floquet formalism [62], to which we have contributed [63], amended by a numerically efficient evaluation of the time-evolution operator developed by Shillito *et al.* [64].

APPENDIX G: NON-RWA EFFECTS IN FLOQUET SIMULATIONS OF THE FULL DEVICE

In this Appendix we briefly discuss the role of counter-rotating terms in the Floquet simulations of the full device

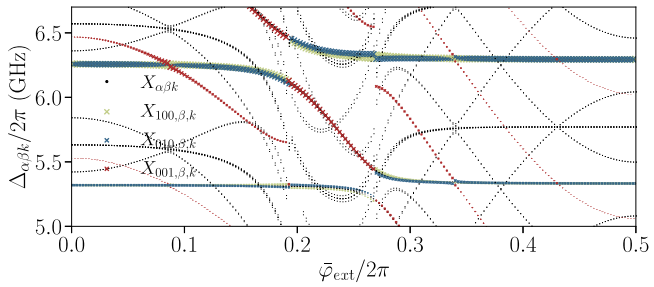


FIG. 9. Floquet eigenspectrum for a rotating-wave approximation in which all photon-number nonconserving terms are removed from the full-circuit Hamiltonian analyzed in Sec. V B. This figure is to be compared to the analogous result for the full Hamiltonian in Fig. 7.

Hamiltonian. Counter-rotating terms (among which the parity-breaking cubic terms play a significant role) in the coupler Hamiltonian induce a helpful correction to the coupler frequency, as can be seen by comparing Figs. 9 to 7. This indicates, among other things, that a mere approximation of the coupler Hamiltonian as a Kerr nonlinear oscillator, as done in the case of the toy model, would be insufficient for precise comparisons with experimental data. Moreover, the speedup obtained by using the Floquet method, together with the numerically efficient method for computing the time-evolution operator, enables us to study non-RWA effects efficiently as compared to full time dynamics.

[1] M. Kjaergaard, M. E. Schwartz, J. Braumüller, P. Krantz, J. I.-J. Wang, S. Gustavsson, and W. D. Oliver, Superconducting qubits: Current state of play, Annu. Rev. Condens. Matter Phys. **11**, 369 (2020).

[2] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, A quantum engineer's guide to superconducting qubits, Appl. Phys. Rev. **6**, 021318 (2019).

[3] A. Blais, A. L. Grimsmo, S. M. Girvin, and A. Wallraff, Circuit quantum electrodynamics, Rev. Mod. Phys. **93**, 025005 (2021).

[4] R. Versluis, S. Poletto, N. Khammassi, B. Tarasinski, N. Haider, D. J. Michalak, A. Bruno, K. Bertels, and L. DiCarlo, Scalable Quantum Circuit and Control for a Superconducting Surface Code, Phys. Rev. Appl. **8**, 034021 (2017).

[5] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, and D. A. Buell, *et al.*, Quantum supremacy using a programmable superconducting processor, Nature **574**, 505 (2019).

[6] A. D. Corcoles, A. Kandala, A. Javadi-Abhari, D. T. McClure, A. W. Cross, K. Temme, P. D. Nation, M. Steffen, and J. M. Gambetta, Challenges and opportunities of near-term quantum computing systems, Proc. IEEE **108**, 1338 (2020).

[7] D. Gottesman, Stabilizer codes and quantum error correction, (1997), arXiv preprint ArXiv:quant-ph/9705052.

[8] D. Aharonov and M. Ben-Or, Fault-tolerant quantum computation with constant error rate, SIAM J. Comput. **38**, 1207 (2008).

[9] P. Mundada, G. Zhang, T. Hazard, and A. Houck, Suppression of Qubit Crosstalk in a Tunable Coupling Superconducting Circuit, Phys. Rev. Appl. **12**, 054023 (2019).

[10] J. Ku, X. Xu, M. Brink, D. C. McKay, J. B. Hertzberg, M. H. Ansari, and B. L. T. Plourde, Suppression of Unwanted *zz* Interactions in a Hybrid Two-Qubit System, Phys. Rev. Lett. **125**, 200504 (2020).

[11] Y. Sung, L. Ding, J. Braumüller, A. Vepsäläinen, B. Kannan, M. Kjaergaard, A. Greene, G. O. Samach, C. McNally, D. Kim, A. Melville, B. M. Niedzielski, M. E. Schwartz, J. L. Yoder, T. P. Orlando, S. Gustavsson, and W. D. Oliver, Realization of High-Fidelity CZ and *zz*-free *i*SWAP Gates with a Tunable Coupler, Phys. Rev. X **11**, 021058 (2021).

[12] A. Noguchi, A. Osada, S. Masuda, S. Kono, K. Heya, S. P. Wolski, H. Takahashi, T. Sugiyama, D. Lachance-Quirion, and Y. Nakamura, Fast parametric two-qubit gates with suppressed residual interaction using the second-order nonlinearity of a cubic transmon, Phys. Rev. A **102**, 062408 (2020).

[13] A. Kandala, K. X. Wei, S. Srinivasan, E. Magesan, S. Carnevale, G. A. Keefe, D. Klaus, O. Dial, and D. C. McKay, Demonstration of a High-Fidelity CNOT Gate for Fixed-Frequency Transmons with Engineered *ZZ* Suppression, Phys. Rev. Lett. **127**, 130501 (2021).

[14] P. Zhao, D. Lan, P. Xu, G. Xue, M. Blank, X. Tan, H. Yu, and Y. Yu, Suppression of Static *ZZ* Interaction in an All-Transmon Quantum Processor, Phys. Rev. Appl. **16**, 024037 (2021).

[15] C. Berke, E. Varvelis, S. Trebst, A. Altland, and D. P. DiVincenzo, Transmon platform for quantum computing challenged by chaotic fluctuations, Nat. Commun. **13**, 2495 (2022).

[16] D. C. McKay, S. Filipp, A. Mezzacapo, E. Magesan, J. M. Chow, and J. M. Gambetta, Universal Gate for Fixed-Frequency Qubits via a Tunable Bus, Phys. Rev. Appl. **6**, 064007 (2016).

[17] M. C. Collodo, J. Herrmann, N. Lacroix, C. K. Andersen, A. Remm, S. Lazar, J.-C. Besse, T. Walter, A. Wallraff, and C. Eichler, Implementation of Conditional Phase Gates Based on Tunable *ZZ* Interactions, Phys. Rev. Lett. **125**, 240502 (2020).

[18] J. H. Shirley, Solution of the Schrödinger equation with a hamiltonian periodic in time, Phys. Rev. **138**, B979 (1965).

[19] H. Sambe, Steady states and quasienergies of a quantum-mechanical system in an oscillating field, Phys. Rev. A **7**, 2203 (1973).

[20] M. Grifoni and P. Hänggi, Driven quantum tunneling, Phys. Rep. **304**, 229 (1998).

[21] S. E. Nigg, H. Paik, B. Vlastakis, G. Kirchmair, S. Shankar, L. Frunzio, M. H. Devoret, R. J. Schoelkopf, and S. M. Girvin, Black-Box Superconducting Circuit Quantization, Phys. Rev. Lett. **108**, 240502 (2012).

[22] Z. K. Minev, Z. Leghtas, S. O. Mundhada, L. Christakis, I. M. Pop, and M. H. Devoret, Energy-participation quantization of Josephson circuits (2020), ArXiv:2010.00620.

[23] L. Verney, R. Lescanne, M. H. Devoret, Z. Leghtas, and M. Mirrahimi, Structural Instability of Driven Josephson Circuits Prevented by an Inductive Shunt, Phys. Rev. Appl. **11**, 024003 (2019).

[24] A. Petrescu, B. Royer, and A. Blais, *in preparation*.

[25] L. S. Theis and F. K. Wilhelm, Nonadiabatic corrections to fast dispersive multiqubit gates involvingzcontrol, Phys. Rev. A **95**, 022314 (2017).

[26] A. Petrescu, M. Malekakhlagh, and H. E. Türeci, Lifetime renormalization of driven weakly anharmonic superconducting qubits. II. The readout problem, Phys. Rev. B **101**, 134510 (2020).

[27] M. Malekakhlagh, E. Magesan, and D. C. McKay, First-principles analysis of cross-resonance gate operation, Phys. Rev. A **102**, 042605 (2020).

[28] M. Malekakhlagh, A. Petrescu, and H. E. Türeci, Lifetime renormalization of weakly anharmonic superconducting qubits. I. Role of number nonconserving terms, Phys. Rev. B **101**, 134509 (2020).

[29] J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Charge-insensitive qubit design derived from the Cooper pair box, Phys. Rev. A **76**, 042319 (2007).

[30] 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).

[31] J. M. Chow, A. D. Córcoles, J. M. Gambetta, C. Rigetti, B. R. Johnson, J. A. Smolin, J. R. Rozen, G. A. Keefe, M. B. Rothwell, M. B. Ketchen, and M. Steffen, Simple All-Microwave Entangling Gate for Fixed-Frequency Superconducting Qubits, Phys. Rev. Lett. **107**, 080502 (2011).

[32] S. Sheldon, E. Magesan, J. M. Chow, and J. M. Gambetta, Procedure for systematically tuning up cross-talk in the cross-resonance gate, Phys. Rev. A **93**, 060302 (2016).

[33] J. Cohen, A. Petrescu, R. Shillito, and A. Blais, Reminiscence of classical chaos in driven transmons, (2022), ArXiv:2207.09361.

[34] Y. Zhang, B. J. Lester, Y. Y. Gao, L. Jiang, R. J. Schoelkopf, and S. M. Girvin, Engineering bilinear mode coupling in circuit QED: Theory and experiment, Phys. Rev. A **99**, 012314 (2019).

[35] S. Krinner, P. Kurpiers, B. Royer, P. Magnard, I. Tsitsilin, J.-C. Besse, A. Remm, A. Blais, and A. Wallraff, Demonstration of an All-Microwave Controlled-Phase Gate Between Far-Detuned Qubits, Phys. Rev. Appl. **14**, 044039 (2020).

[36] P. S. Mundada, A. Gyenis, Z. Huang, J. Koch, and A. A. Houck, Floquet-Engineered Enhancement of Coherence Times in a Driven Fluxonium Qubit, Phys. Rev. Appl. **14**, 054033 (2020).

[37] Z. Huang, P. S. Mundada, A. Gyenis, D. I. Schuster, A. A. Houck, and J. Koch, Engineering Dynamical Sweet Spots to Protect Qubits from $1/f$ Noise, Phys. Rev. Appl. **15**, 034065 (2021).

[38] J. Q. You, X. Hu, S. Ashhab, and F. Nori, Low-decoherence flux qubit, Phys. Rev. B **75**, 140515 (2007).

[39] M. Steffen, F. Brito, D. DiVincenzo, M. Farinelli, G. Keefe, M. Ketchen, S. Kumar, F. Milliken, M. B. Rothwell, J.

Rozen, and R. H. Koch, Quantum information storage using tunable flux qubits, J. Phys.: Condens. Matter **22**, 053201 (2010).

[40] F. Yan, S. Gustavsson, A. Kamal, J. Birenbaum, A. P. Sears, D. Hover, T. J. Gudmundsen, D. Rosenberg, G. Samach, and S. Weber, *et al.*, The flux qubit revisited to enhance coherence and reproducibility, Nat. Commun. **7**, 12964 (2016).

[41] X. Xu and M. Ansari, *ZZ* Freedom in Two-Qubit Gates, Phys. Rev. Appl. **15**, 064074 (2021).

[42] P. Zhao, P. Xu, D. Lan, X. Tan, H. Yu, and Y. Yu, Switchable Next-Nearest-Neighbor Coupling for Controlled Two-Qubit Operations, Phys. Rev. Appl. **14**, 064016 (2020).

[43] S. Reynaud, E. Giacobino, J. Zinn-Justin, eds., *Fluctuations quantiques : Les Houches, Session LXIII, 27 juin-28 juillet 1995* (Elsevier Amsterdam, New York, https://www.worldcat.org/title/fluctuations-quantiques-les-houches-session-lxiii-27-juin-28-juillet-1995-quantum-fluctuations/oclc/36407840, 1997).

[44] U. Vool and M. Devoret, Introduction to quantum electromagnetic circuits, Int. J. Circuit Theory Appl. **45**, 897 (2017).

[45] X. You, J. A. Sauls, and J. Koch, Circuit quantization in the presence of time-dependent external flux, Phys. Rev. B **99**, 174512 (2019).

[46] M. Mirrahimi and P. Rouchon, *Dynamics and control of open quantum systems* (2021).

[47] N. Goldman and J. Dalibard, Periodically Driven Quantum Systems: Effective Hamiltonians and Engineered Gauge Fields, Phys. Rev. X **4**, 031027 (2014).

[48] S. Krinner, S. Lazar, A. Remm, C. Andersen, N. Lacroix, G. Norris, C. Hellings, M. Gabureac, C. Eichler, and A. Wallraff, Benchmarking Coherent Errors in Controlled-Phase Gates due to Spectator Qubits, Phys. Rev. Appl. **14**, 024042 (2020).

[49] E. A. Sete, A. Q. Chen, R. Manenti, S. Kulshreshtha, and S. Poletto, Floating Tunable Coupler for Scalable Quantum Computing Architectures, Phys. Rev. Appl. **15**, 064063 (2021).

[50] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables* Vol. 55 (Courier Corporation, New York, 1964).

[51] R. Žitko, Sneg – mathematica package for symbolic calculations with second-quantization-operator expressions, Comput. Phys. Commun. **182**, 2259 (2011).

[52] P. Bertet, C. J. P. M. Harmans, and J. E. Mooij, Parametric coupling for superconducting qubits, Phys. Rev. B **73**, 064512 (2006).

[53] A. O. Niskanen, Y. Nakamura, and J.-S. Tsai, Tunable coupling scheme for flux qubits at the optimal point, Phys. Rev. B **73**, 094506 (2006).

[54] A. Niskanen, K. Harrabi, F. Yoshihara, Y. Nakamura, S. Lloyd, and J. S. Tsai, Quantum coherent tunable coupling of superconducting qubits, Science **316**, 723 (2007).

[55] F. Beaudoin, M. P. da Silva, Z. Dutton, and A. Blais, First-order sidebands in circuit QED using qubit frequency modulation, Phys. Rev. A **86**, 022305 (2012).

[56] M. Reagor, C. B. Osborn, N. Tezak, A. Staley, G. Prawiroatmodjo, M. Scheer, N. Alidoust, E. A. Sete, N. Didier, and M. P. da Silva, *et al.*, Demonstration of universal parametric entangling gates on a multi-qubit lattice, Sci. Adv. **4**, eaao3603 (2018).

[57] S. A. Caldwell, N. Didier, C. A. Ryan, E. A. Sete, A. Hudson, P. Karalekas, R. Manenti, M. P. da Silva, R. Sinclair, and E. Acala, *et al.*, Parametrically Activated Entangling Gates using Transmon Qubits, Phys. Rev. Appl. **10**, 034050 (2018).

[58] N. Didier, E. A. Sete, M. P. da Silva, and C. Rigetti, Analytical modeling of parametrically modulated transmon qubits, Phys. Rev. A **97**, 022330 (2018).

[59] M. Roth, M. Ganzhorn, N. Moll, S. Filipp, G. Salis, and S. Schmidt, Analysis of a parametrically driven exchange-type gate and a two-photon excitation gate between superconducting qubits, Phys. Rev. A **96**, 062323 (2017).

[60] Pranav S. Mundada and Sara Sussman, *et al.*, *in preparation* (2021).

[61] D. Marcos, P. Rabl, E. Rico, and P. Zoller, Superconducting Circuits for Quantum Simulation of Dynamical Gauge Fields, Phys. Rev. Lett. **111**, 110504 (2013).

[62] J. Johansson, P. Nation, and F. Nori, Qutip 2: A python framework for the dynamics of open quantum systems, Comput. Phys. Commun. **184**, 1234 (2013).

[63] C. Le Calonnec, https://github.com/qutip/qutip/pull/1248 (2020).

[64] R. Shillito, J. A. Gross, A. Di Paolo, E. Genois, and A. Blais, Fast and differentiable simulation of driven quantum systems, Phys. Rev. Res. **3**, 033266 (2021).

Bibliography

[1] S. Reynaud, E. Giacobino, and Jean Zinn-Justin, editors. *Fluctuations quantiques : Les Houches, Session LXIII, 27 juin-28 juillet 1995*. Elsevier Amsterdam, New York, https://www.worldcat.org/title/fluctuations-quantiques-les-houches-session-lxiii-27-juin-28-juillet-1995-quantum-fluctuations/oclc/36407840, 1997. (Cited on pages 2 and 25.)

[2] Uri Vool and Michel Devoret. Introduction to quantum electromagnetic circuits. *International Journal of Circuit Theory and Applications*, 45(7):897–934, Jun 2017. ISSN 0098-9886. doi: 10.1002/cta.2359. URL http://dx.doi.org/10.1002/cta.2359. (Cited on pages 2, 4, 25, and 118.)

[3] Guido Burkard, Roger H. Koch, and David P. DiVincenzo. Multilevel quantum description of decoherence in superconducting qubits. *Phys. Rev. B*, 69:064503, Feb 2004. doi: 10.1103/PhysRevB.69.064503. URL https://link.aps.org/doi/10.1103/PhysRevB.69.064503. (Cited on page 4.)

[4] Jascha Ulrich and Fabian Hassler. Dual approach to circuit quantization using loop charges. *Phys. Rev. B*, 94:094505, Sep 2016. doi: 10.1103/PhysRevB.94.094505. URL https://link.aps.org/doi/10.1103/PhysRevB.94.094505. (Cited on page 4.)

[5] Alexandre Blais, Arne L. Grimsmo, S. M. Girvin, and Andreas Wallraff. Circuit quantum electrodynamics. *Rev. Mod. Phys.*, 93:025005, May 2021. doi: 10.1103/RevModPhys.93.025005. URL https://link.aps.org/doi/10.1103/RevModPhys.93.025005. (Cited on pages 6, 12, 13, 15, and 21.)

[6] Jens Koch, Terri M. Yu, Jay Gambetta, A. A. Houck, D. I. Schuster, J. Majer, Alexandre Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. Charge-insensitive qubit design derived from the cooper pair box. *Phys. Rev. A*, 76:042319, Oct 2007. doi: 10.1103/PhysRevA.76.042319. URL https://link.aps.org/doi/10.1103/PhysRevA.76.042319. (Cited on pages 6, 7, and 9.)

[7] D. Gottesman. Fault-tolerant quantum computation with higher-dimensional systems. *Chaos, Solitons & Fractals*, 10:1749–1758, 1999. ISSN 0960-0779. doi: 10.1016/S0960-0779(98)00218-5. URL https://www.sciencedirect.com/science/article/pii/S0960077998002185. (Cited on page 9.)

[8] Yuchen Wang, Zixuan Hu, Barry C. Sanders, and Sabre Kais. Qudits and high-dimensional quantum computing. *Frontiers in Physics*, 8, 2020. ISSN 2296-424X. doi: 10.3389/fphy.2020.589504. URL https://www.frontiersin.org/articles/10.3389/fphy.2020.589504. (Cited on page 9.)

[9] R. Bianchetti, S. Filipp, M. Baur, J. M. Fink, C. Lang, L. Steffen, M. Boissonneault, A. Blais, and A. Wallraff. Control and tomography of a three level superconducting artificial atom. *Phys. Rev. Lett.*, 105:223601, Nov 2010. doi: 10.1103/PhysRevLett.105.223601. URL https://link.aps.org/doi/10.1103/PhysRevLett.105.223601. (Cited on page 9.)

[10] R. Barends, J. Kelly, A. Megrant, D. Sank, E. Jeffrey, Y. Chen, Y. Yin, B. Chiaro, J. Mutus, C. Neill, P. O'Malley, P. Roushan, J. Wenner, T. C. White, A. N. Cleland, and John M. Martinis. Coherent josephson qubit suitable for scalable quantum integrated circuits. *Phys. Rev. Lett.*, 111:080502, Aug 2013. doi: 10.1103/PhysRevLett.111.080502. URL https://link.aps.org/doi/10.1103/PhysRevLett.111.080502. (Cited on page 9.)

[11] Alexander P. M. Place, Lila V. H. Rodgers, Pranav Mundada, Basil M. Smitham, Mattias Fitzpatrick, Zhaoqi Leng, Anjali Premkumar, Jacob Bryon, Andrei Vrajitoarea, Sara Sussman, Guangming Cheng, Trisha Madhavan, Harshvardhan K. Babla, Xuan Hoang Le, Youqi Gang, Berthold Jäck, András Gyenis, Nan Yao, Robert J. Cava, Nathalie P. de Leon, and Andrew A. Houck. New material platform for superconducting transmon qubits with coherence times exceeding 0.3 milliseconds. *Nature Communications*, 12(1):1779, 2021. ISSN 2041-1723. doi: 10.1038/s41467-021-22030-5. URL https://doi.org/10.1038/s41467-021-22030-5.

[12] M. Carroll, S. Rosenblatt, P. Jurcevic, I. Lauer, and A. Kandala. Dynamics of superconducting qubit relaxation times. *npj Quantum Information*, 8(1):132, 2022. ISSN 2056-6387. doi: 10.1038/s41534-022-00643-y. URL https://doi.org/10.1038/s41534-022-00643-y. (Cited on page 9.)

[13] David C. McKay, Stefan Filipp, Antonio Mezzacapo, Easwar Magesan, Jerry M. Chow, and Jay M. Gambetta. Universal gate for fixed-frequency qubits via a tunable bus. *Phys. Rev. Applied*, 6:064007, Dec 2016. doi: 10.1103/PhysRevApplied.6.064007. URL https://link.aps.org/doi/10.1103/PhysRevApplied.6.064007. (Cited on page 10.)

[14] Fei Yan, Philip Krantz, Youngkyu Sung, Morten Kjaergaard, Daniel L. Campbell, Terry P. Orlando, Simon Gustavsson, and William D. Oliver. Tunable coupling scheme for implementing high-fidelity two-qubit gates. *Phys. Rev. Appl.*, 10:054062, Nov 2018. doi: 10.1103/PhysRevApplied.10.054062. URL https://link.aps.org/doi/10.1103/PhysRevApplied.10.054062. (Cited on pages 10 and 14.)

[15] Eyob A. Sete, Angela Q. Chen, Riccardo Manenti, Shobhan Kulshreshtha, and Stefano Poletto. Floating tunable coupler for scalable quantum computing architectures. *Phys. Rev. Applied*, 15:064063, Jun 2021. doi: 10.1103/PhysRevApplied.15.064063. URL https://link.aps.org/doi/10.1103/PhysRevApplied.15.064063.

[16] J. Stehlik, D. M. Zajac, D. L. Underwood, T. Phung, J. Blair, S. Carnevale, D. Klaus, G. A. Keefe, A. Carniol, M. Kumph, Matthias Steffen, and O. E. Dial. Tunable coupling architecture for fixed-frequency transmon superconducting qubits. *Phys. Rev. Lett.*, 127:080505, Aug 2021. doi: 10.1103/PhysRevLett.127.080505. URL https://link.aps.org/doi/10.1103/PhysRevLett.127.080505.

[17] Youngkyu Sung, Leon Ding, Jochen Braumüller, Antti Vepsäläinen, Bharath Kannan, Morten Kjaergaard, Ami Greene, Gabriel O. Samach, Chris McNally, David Kim, Alexander Melville, Bethany M. Niedzielski, Mollie E. Schwartz, Jonilyn L. Yoder, Terry P. Orlando, Simon Gustavsson, and William D. Oliver. Realization of high-fidelity cz and zz-free iswap gates with a tunable coupler. *Phys. Rev. X*, 11:021058, Jun 2021. doi: 10.1103/PhysRevX.11.021058. URL https://link.aps.org/doi/10.1103/PhysRevX.11.021058.

[18] X. Li, T. Cai, H. Yan, Z. Wang, X. Pan, Y. Ma, W. Cai, J. Han, Z. Hua, X. Han, Y. Wu, H. Zhang, H. Wang, Yipu Song, Luming Duan, and Luyan Sun. Tunable coupler for realizing a controlled-phase gate with dynamically decoupled regime in a superconducting circuit. *Phys. Rev. Appl.*, 14:024070, Aug 2020. doi: 10.1103/PhysRevApplied.14.024070. URL https://link.aps.org/doi/10.1103/PhysRevApplied.14.024070.

[19] S. A. Caldwell, N. Didier, C. A. Ryan, E. A. Sete, A. Hudson, P. Karalekas, R. Manenti, M. P. da Silva, R. Sinclair, E. Acala, N. Alidoust, J. Angeles, A. Bestwick, M. Block, B. Bloom, A. Bradley, C. Bui, L. Capelluto, R. Chilcott, J. Cordova, G. Crossman, M. Curtis, S. Deshpande, T. El Bouayadi, D. Girshovich, S. Hong, K. Kuang, M. Lenihan, T. Manning, A. Marchenkov, J. Marshall, R. Maydra, Y. Mohan, W. O'Brien, C. Osborn, J. Otterbach, A. Papageorge, J.-P. Paquette, M. Pelstring, A. Polloreno, G. Prawiroatmodjo, V. Rawat, M. Reagor, R. Renzas, N. Rubin, D. Russell, M. Rust, D. Scarabelli, M. Scheer, M. Selvanayagam, R. Smith, A. Staley, M. Suska, N. Tezak, D. C. Thompson, T.-W. To, M. Vahidpour, N. Vodrahalli, T. Whyland, K. Yadav, W. Zeng, and C. Rigetti. Parametrically activated entangling gates using transmon qubits. *Phys. Rev. Applied*, 10:034050, Sep 2018. doi: 10.1103/PhysRevApplied.10.034050. URL https://link.aps.org/doi/10.1103/PhysRevApplied.10.034050. (Cited on page 16.)

[20] Nicolas Didier, Eyob A. Sete, Marcus P. da Silva, and Chad Rigetti. Analytical modeling of parametrically modulated transmon qubits. *Phys. Rev. A*, 97:022330, Feb 2018. doi: 10.1103/PhysRevA.97.022330. URL https://link.aps.org/doi/10.1103/PhysRevA.97.022330. (Cited on page 10.)

[21] Alexis Morvan, Larry Chen, Jeffrey M. Larson, David I. Santiago, and Irfan Siddiqi. Optimizing frequency allocation for fixed-frequency superconducting quantum processors. *Phys. Rev. Res.*, 4:023079, Apr 2022. doi: 10.1103/PhysRevResearch.4.023079. URL https://link.aps.org/doi/10.1103/PhysRevResearch.4.023079. (Cited on page 10.)

[22] J M Kreikebaum, K P O'Brien, A Morvan, and I Siddiqi. Improving wafer-scale josephson junction resistance variation in superconducting quantum coherent circuits.

Superconductor Science and Technology, 33(6):06LT02, apr 2020. doi: 10.1088/1361-6668/ab8617. URL https://dx.doi.org/10.1088/1361-6668/ab8617.

[23] N. Muthusubramanian, P. Duivestein, C. Zachariadis, M. Finkel, S. L. M. van der Meer, H. M. Veen, M. W. Beekman, T. Stavenga, A. Bruno, and L. DiCarlo. Wafer-scale uniformity of dolan-bridge and bridgeless manhattan-style josephson junctions for superconducting quantum processors. 2023. URL https://arxiv.org/abs/2304.09111. arXiv:2304.09111 [quant-ph]. (Cited on page 10.)

[24] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando G. S. L. Brandao, David A. Buell, Brian Burkett, Yu Chen, Zijun Chen, Ben Chiaro, Roberto Collins, William Courtney, Andrew Dunsworth, Edward Farhi, Brooks Foxen, Austin Fowler, Craig Gidney, Marissa Giustina, Rob Graff, Keith Guerin, Steve Habegger, Matthew P. Harrigan, Michael J. Hartmann, Alan Ho, Markus Hoffmann, Trent Huang, Travis S. Humble, Sergei V. Isakov, Evan Jeffrey, Zhang Jiang, Dvir Kafri, Kostyantyn Kechedzhi, Julian Kelly, Paul V. Klimov, Sergey Knysh, Alexander Korotkov, Fedor Kostritsa, David Landhuis, Mike Lindmark, Erik Lucero, Dmitry Lyakh, Salvatore Mandrà, Jarrod R. McClean, Matthew McEwen, Anthony Megrant, Xiao Mi, Kristel Michielsen, Masoud Mohseni, Josh Mutus, Ofer Naaman, Matthew Neeley, Charles Neill, Murphy Yuezhen Niu, Eric Ostby, Andre Petukhov, John C. Platt, Chris Quintana, Eleanor G. Rieffel, Pedram Roushan, Nicholas C. Rubin, Daniel Sank, Kevin J. Satzinger, Vadim Smelyanskiy, Kevin J. Sung, Matthew D. Trevithick, Amit Vainsencher, Benjamin Villalonga, Theodore White, Z. Jamie Yao, Ping Yeh, Adam Zalcman, Hartmut Neven, and John M. Martinis. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, 2019. doi: 10.1038/s41586-019-1666-5. URL https://doi.org/10.1038/s41586-019-1666-5. (Cited on pages 10 and 14.)

[25] Paul V. Klimov, Andreas Bengtsson, Chris Quintana, Alexandre Bourassa, Sabrina Hong, Andrew Dunsworth, Kevin J. Satzinger, William P. Livingston, Volodymyr Sivak, Murphy Y. Niu, Trond I. Andersen, Yaxing Zhang, Desmond Chik, Zijun Chen, Charles Neill, Catherine Erickson, Alejandro Grajales Dau, Anthony Megrant, Pedram Roushan, Alexander N. Korotkov, Julian Kelly, Vadim Smelyanskiy, Yu Chen, and Hartmut Neven. Optimizing quantum gates towards the scale of logical qubits. 2023. URL https://arxiv.org/abs/2308.02321. arXiv:2308.02321 [quant-ph]. (Cited on page 10.)

[26] Paul V. Klimov, Julian Kelly, John M. Martinis, and Hartmut Neven. The snake optimizer for learning quantum processor control parameters. 2020. URL https://arxiv.org/abs/2006.04594. arXiv:2006.04594 [quant-ph]. (Cited on page 10.)

[27] David G. Ferguson, A. A. Houck, and Jens Koch. Symmetries and collective excitations in large superconducting circuits. *Phys. Rev. X*, 3:011003, Jan 2013. doi: 10.1103/PhysRevX.3.011003. URL https://link.aps.org/doi/10.1103/PhysRevX.3.011003. (Cited on pages 10 and 11.)

[28] Fei Yan, Youngkyu Sung, Philip Krantz, Archana Kamal, David K. Kim, Jonilyn L. Yoder, Terry P. Orlando, Simon Gustavsson, and William D. Oliver. Engineering

framework for optimizing superconducting qubit designs, 2020. URL https://arxiv.org/abs/2006.04130. arXiv:2006.04130 [quant-ph]. (Cited on page 10.)

[29] N. E. Frattini, U. Vool, S. Shankar, A. Narla, K. M. Sliwa, and M. H. Devoret. 3-wave mixing Josephson dipole element. *Applied Physics Letters*, 110(22), 05 2017. ISSN 0003-6951. doi: 10.1063/1.4984142. URL https://doi.org/10.1063/1.4984142. 222603. (Cited on page 11.)

[30] T. P. Orlando, J. E. Mooij, Lin Tian, Caspar H. van der Wal, L. S. Levitov, Seth Lloyd, and J. J. Mazo. Superconducting persistent-current qubit. *Phys. Rev. B*, 60:15398–15413, Dec 1999. doi: 10.1103/PhysRevB.60.15398. URL https://link.aps.org/doi/10.1103/PhysRevB.60.15398.

[31] Matthias Steffen, Shwetank Kumar, David P. DiVincenzo, J. R. Rozen, George A. Keefe, Mary Beth Rothwell, and Mark B. Ketchen. High-coherence hybrid superconducting qubit. *Phys. Rev. Lett.*, 105:100502, Sep 2010. doi: 10.1103/PhysRevLett.105.100502. URL https://link.aps.org/doi/10.1103/PhysRevLett.105.100502.

[32] Fei Yan, Simon Gustavsson, Archana Kamal, Jeffrey Birenbaum, Adam P. Sears, David Hover, Ted J. Gudmundsen, Danna Rosenberg, Gabriel Samach, S. Weber, Jonilyn L. Yoder, Terry P. Orlando, John Clarke, Andrew J. Kerman, and William D. Oliver. The flux qubit revisited to enhance coherence and reproducibility. *Nature Communications*, 7(1):12964, 2016. ISSN 2041-1723. doi: 10.1038/ncomms12964. URL https://doi.org/10.1038/ncomms12964.

[33] Vladimir E. Manucharyan, Jens Koch, Leonid I. Glazman, and Michel H. Devoret. Fluxonium: Single cooper-pair circuit free of charge offsets. *Science*, 326(5949): 113–116, 2009. doi: 10.1126/science.1175552. URL https://www.science.org/doi/abs/10.1126/science.1175552. (Cited on page 10.)

[34] Giovanni Viola and Gianluigi Catelani. Collective modes in the fluxonium qubit. *Phys. Rev. B*, 92:224511, Dec 2015. doi: 10.1103/PhysRevB.92.224511. URL https://link.aps.org/doi/10.1103/PhysRevB.92.224511. (Cited on pages 11 and 26.)

[35] Peng Zhao, Dong Lan, Peng Xu, Guangming Xue, Mace Blank, Xinsheng Tan, Haifeng Yu, and Yang Yu. Suppression of static zz interaction in an all-transmon quantum processor. *Phys. Rev. Applied*, 16:024037, Aug 2021. doi: 10.1103/PhysRevApplied.16.024037. URL https://link.aps.org/doi/10.1103/PhysRevApplied.16.024037. (Cited on pages 12 and 15.)

[36] Peng Zhao, Peng Xu, Dong Lan, Ji Chu, Xinsheng Tan, Haifeng Yu, and Yang Yu. High-contrast zz interaction using superconducting qubits with opposite-sign anharmonicity. *Phys. Rev. Lett.*, 125:200503, Nov 2020. doi: 10.1103/PhysRevLett.125.200503. URL https://link.aps.org/doi/10.1103/PhysRevLett.125.200503. (Cited on page 15.)

[37] Jaseung Ku, Xuexin Xu, Markus Brink, David C. McKay, Jared B. Hertzberg, Mohammad H. Ansari, and B. L. T. Plourde. Suppression of unwanted zz interactions

in a hybrid two-qubit system. *Phys. Rev. Lett.*, 125:200504, Nov 2020. doi: 10.1103/PhysRevLett.125.200504. URL https://link.aps.org/doi/10.1103/PhysRevLett.125.200504. (Cited on page 12.)

[38] Pranav Mundada, Gengyan Zhang, Thomas Hazard, and Andrew Houck. Suppression of qubit crosstalk in a tunable coupling superconducting circuit. *Phys. Rev. Applied*, 12:054023, Nov 2019. doi: 10.1103/PhysRevApplied.12.054023. URL https://link.aps.org/doi/10.1103/PhysRevApplied.12.054023.

[39] Atsushi Noguchi, Alto Osada, Shumpei Masuda, Shingo Kono, Kentaro Heya, Samuel Piotr Wolski, Hiroki Takahashi, Takanori Sugiyama, Dany Lachance-Quirion, and Yasunobu Nakamura. Fast parametric two-qubit gates with suppressed residual interaction using the second-order nonlinearity of a cubic transmon. *Phys. Rev. A*, 102:062408, Dec 2020. doi: 10.1103/PhysRevA.102.062408. URL https://link.aps.org/doi/10.1103/PhysRevA.102.062408. (Cited on pages 12 and 15.)

[40] L. DiCarlo, J. M. Chow, J. M. Gambetta, Lev S. Bishop, B. R. Johnson, D. I. Schuster, J. Majer, A. Blais, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf. Demonstration of two-qubit algorithms with a superconducting quantum processor. *Nature*, 460(7252): 240–244, 07 2009. doi: 10.1038/nature08121. URL https://doi.org/10.1038/nature08121. (Cited on page 13.)

[41] Clemens Müller, Jared H Cole, and Jürgen Lisenfeld. Towards understanding two-level-systems in amorphous solids: insights from quantum circuits. *Reports on Progress in Physics*, 82(12):124501, oct 2019. doi: 10.1088/1361-6633/ab3a7e. URL https://dx.doi.org/10.1088/1361-6633/ab3a7e. (Cited on page 14.)

[42] Yu Chen, C. Neill, P. Roushan, N. Leung, M. Fang, R. Barends, J. Kelly, B. Campbell, Z. Chen, B. Chiaro, A. Dunsworth, E. Jeffrey, A. Megrant, J. Y. Mutus, P. J. J. O'Malley, C. M. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, Michael R. Geller, A. N. Cleland, and John M. Martinis. Qubit architecture with high coherence and fast tunable coupling. *Phys. Rev. Lett.*, 113:220502, Nov 2014. doi: 10.1103/PhysRevLett.113.220502. URL https://link.aps.org/doi/10.1103/PhysRevLett.113.220502. (Cited on page 14.)

[43] 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, 04 2014. ISSN 1476-4687. doi: 10.1038/nature13171. URL https://doi.org/10.1038/nature13171. (Cited on page 14.)

[44] Chad Rigetti and Michel Devoret. Fully microwave-tunable universal gates in superconducting qubits with linear couplings and fixed transition frequencies. *Phys. Rev. B*, 81:134507, Apr 2010. doi: 10.1103/PhysRevB.81.134507. URL https://link.aps.org/doi/10.1103/PhysRevB.81.134507. (Cited on page 14.)

[45] Jerry M. Chow, A. D. Córcoles, Jay M. Gambetta, Chad Rigetti, B. R. Johnson, John A. Smolin, J. R. Rozen, George A. Keefe, Mary B. Rothwell, Mark B. Ketchen, and M. Steffen. Simple all-microwave entangling gate for fixed-frequency superconducting qubits. *Phys. Rev. Lett.*, 107:080502, Aug 2011. doi: 10.1103/PhysRevLett.107.080502. URL https://link.aps.org/doi/10.1103/PhysRevLett.107.080502. (Cited on page 14.)

[46] A. D. Córcoles, Jay M. Gambetta, Jerry M. Chow, John A. Smolin, Matthew Ware, Joel Strand, B. L. T. Plourde, and M. Steffen. Process verification of two-qubit quantum gates by randomized benchmarking. *Phys. Rev. A*, 87:030301, Mar 2013. doi: 10.1103/PhysRevA.87.030301. URL https://link.aps.org/doi/10.1103/PhysRevA.87.030301. (Cited on page 15.)

[47] Sarah Sheldon, Easwar Magesan, Jerry M. Chow, and Jay M. Gambetta. Procedure for systematically tuning up cross-talk in the cross-resonance gate. *Physical Review A*, 93(6), Jun 2016. ISSN 2469-9934. doi: 10.1103/physreva.93.060302. URL http://dx.doi.org/10.1103/PhysRevA.93.060302. (Cited on page 15.)

[48] Alexandru Petrescu, Camille Le Calonnec, Catherine Leroux, Agustin Di Paolo, Pranav Mundada, Sara Sussman, Andrei Vrajitoarea, Andrew A. Houck, and Alexandre Blais. Accurate methods for the analysis of strong-drive effects in parametric gates. *Phys. Rev. Appl.*, 19:044003, Apr 2023. doi: 10.1103/PhysRevApplied.19.044003. URL https://link.aps.org/doi/10.1103/PhysRevApplied.19.044003. (Cited on pages 15, 27, 34, 42, 43, 44, 47, and 48.)

[49] Matthew Reagor, Christopher B. Osborn, Nikolas Tezak, Alexa Staley, Guenevere Prawiroatmodjo, Michael Scheer, Nasser Alidoust, Eyob A. Sete, Nicolas Didier, Marcus P. da Silva, Ezer Acala, Joel Angeles, Andrew Bestwick, Maxwell Block, Benjamin Bloom, Adam Bradley, Catvu Bui, Shane Caldwell, Lauren Capelluto, Rick Chilcott, Jeff Cordova, Genya Crossman, Michael Curtis, Saniya Deshpande, Tristan El Bouayadi, Daniel Girshovich, Sabrina Hong, Alex Hudson, Peter Karalekas, Kat Kuang, Michael Lenihan, Riccardo Manenti, Thomas Manning, Jayss Marshall, Yuvraj Mohan, William O'Brien, Johannes Otterbach, Alexander Papageorge, Jean-Philip Paquette, Michael Pelstring, Anthony Polloreno, Vijay Rawat, Colm A. Ryan, Russ Renzas, Nick Rubin, Damon Russel, Michael Rust, Diego Scarabelli, Michael Selvanayagam, Rodney Sinclair, Robert Smith, Mark Suska, Ting-Wai To, Mehrnoosh Vahidpour, Nagesh Vodrahalli, Tyler Whyland, Kamal Yadav, William Zeng, and Chad T. Rigetti. Demonstration of universal parametric entangling gates on a multi-qubit lattice. *Science Advances*, 4(2):eaao3603, 2018. ISSN 2375-2548. doi: 10.1126/sciadv.aao3603. URL https://www.science.org/doi/abs/10.1126/sciadv.aao3603. (Cited on page 16.)

[50] Yinqi Chen, Konstantin N. Nesterov, Hugh Churchill, Javad Shabani, Vladimir E. Manucharyan, and Maxim G. Vavilov. Voltage activated parametric entangling gates on gatemons. 2023. URL https://arxiv.org/abs/2304.08469. arXiv:2304.08469 [quant-ph].

[51] Eyob A. Sete, Nicolas Didier, Angela Q. Chen, Shobhan Kulshreshtha, Riccardo Manenti, and Stefano Poletto. Parametric-resonance entangling gates with a tunable coupler. *Phys. Rev. Appl.*, 16:024050, Aug 2021. doi: 10.1103/PhysRevApplied.16.024050. URL https://link.aps.org/doi/10.1103/PhysRevApplied.16.024050. (Cited on pages 16 and 53.)

[52] S. H. Autler and C. H. Townes. Stark effect in rapidly varying fields. *Phys. Rev.*, 100:703–722, Oct 1955. doi: 10.1103/PhysRev.100.703. URL https://link.aps.org/doi/10.1103/PhysRev.100.703. (Cited on page 17.)

[53] F. Bloch and A. Siegert. Magnetic resonance for nonrotating fields. *Phys. Rev.*, 57:522–527, Mar 1940. doi: 10.1103/PhysRev.57.522. URL https://link.aps.org/doi/10.1103/PhysRev.57.522. (Cited on page 17.)

[54] Felix Bloch. Über die quantenmechanik der elektronen in kristallgittern. *Zeitschrift für Physik*, 52(7):555–600, 07 1929. ISSN 0044-3328. doi: 10.1007/BF01339455. URL https://doi.org/10.1007/BF01339455. (Cited on page 18.)

[55] Gaston Floquet. Sur les équations différentielles linéaires à coefficients périodiques. *Annales scientifiques de l'École Normale Supérieure, Serie 2*, 12:47–88, 1883. doi: 10.24033/asens.220. URL http://www.numdam.org/articles/10.24033/asens.220/. (Cited on page 18.)

[56] Milena Grifoni and Peter Hänggi. Driven quantum tunneling. *Physics Reports*, 304(5):229 – 354, 1998. ISSN 0370-1573. doi: https://doi.org/10.1016/S0370-1573(98)00022-2. URL http://www.sciencedirect.com/science/article/pii/S0370157398000222. (Cited on page 18.)

[57] Hideo Sambe. Steady states and quasienergies of a quantum-mechanical system in an oscillating field. *Phys. Rev. A*, 7:2203–2213, Jun 1973. doi: 10.1103/PhysRevA.7.2203. URL https://link.aps.org/doi/10.1103/PhysRevA.7.2203.

[58] Jon H. Shirley. Solution of the schrödinger equation with a hamiltonian periodic in time. *Phys. Rev.*, 138:B979–B987, May 1965. doi: 10.1103/PhysRev.138.B979. URL https://link.aps.org/doi/10.1103/PhysRev.138.B979. (Cited on page 18.)

[59] editor Raymond Y. Chiao. *Amazing Light: A Volume Dedicated to Charles Hard Townes on His 80th Birthday*. Springer, New York, 1996. ISBN QC685 A43 1996. [1996] ©1996. (Cited on pages 21 and 61.)

[60] Simon E. Nigg, Hanhee Paik, Brian Vlastakis, Gerhard Kirchmair, S. Shankar, Luigi Frunzio, M. H. Devoret, R. J. Schoelkopf, and S. M. Girvin. Black-box superconducting circuit quantization. *Physical Review Letters*, 108(24), Jun 2012. ISSN 1079-7114. doi: 10.1103/physrevlett.108.240502. URL http://dx.doi.org/10.1103/PhysRevLett.108.240502. (Cited on page 28.)

[61] Rok Žitko. Sneg – mathematica package for symbolic calculations with second-quantization-operator expressions. *Computer Physics Communications*, 182(10):2259–2264, 2011. ISSN 0010-4655. doi: https://doi.org/10.1016/j.

cpc.2011.05.013. URL https://www.sciencedirect.com/science/article/pii/S0010465511001792. (Cited on page 29.)

[62] Sara Sussman. Quantum computing with an open source qubit controller, 2023. Ph.D. thesis, Princeton University. (Cited on pages 57, 58, and 61.)

[63] C. et al. Guinn. Progress on a tunable coupler architecture for parametric gates between far-detuned fixed-frequency transmon qubits: Part 2. *Bulletin of the American Physical Society*, 2023. (Cited on page 60.)

[64] Christopher J. Wood, Jacob D. Biamonte, and David G. Cory. Tensor networks and graphical calculus for open quantum systems. 2015. URL https://arxiv.org/abs/1111.6950. arXiv:1111.6950 [quant-ph]. (Cited on pages 65 and 66.)

[65] Christopher J. Wood and Jay M. Gambetta. Quantification and characterization of leakage errors. *Phys. Rev. A*, 97:032306, Mar 2018. doi: 10.1103/PhysRevA.97.032306. URL https://link.aps.org/doi/10.1103/PhysRevA.97.032306. (Cited on pages 65 and 66.)

[66] B. Foxen, C. Neill, A. Dunsworth, P. Roushan, B. Chiaro, A. Megrant, J. Kelly, Zijun Chen, K. Satzinger, R. Barends, F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, S. Boixo, D. Buell, B. Burkett, Yu Chen, R. Collins, E. Farhi, A. Fowler, C. Gidney, M. Giustina, R. Graff, M. Harrigan, T. Huang, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, P. Klimov, A. Korotkov, F. Kostritsa, D. Landhuis, E. Lucero, J. McClean, M. McEwen, X. Mi, M. Mohseni, J. Y. Mutus, O. Naaman, M. Neeley, M. Niu, A. Petukhov, C. Quintana, N. Rubin, D. Sank, V. Smelyanskiy, A. Vainsencher, T. C. White, Z. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis. Demonstrating a continuous set of two-qubit gates for near-term quantum algorithms. *Phys. Rev. Lett.*, 125:120504, Sep 2020. doi: 10.1103/PhysRevLett.125.120504. URL https://link.aps.org/doi/10.1103/PhysRevLett.125.120504. (Cited on page 68.)

[67] Young-Hyun Oh, Hamed Mohammadbagherpoor, Patrick Dreher, Anand Singh, Xianqing Yu, and Andy J. Rindos. Solving multi-coloring combinatorial optimization problems using hybrid quantum algorithms. 2019. URL https://arxiv.org/abs/1911.00595. arXiv:1911.00595 [quant-ph]. (Cited on page 75.)

[68] Fred Glover, Gary Kochenberger, and Yu Du. A tutorial on formulating and using qubo models. 2019. URL https://arxiv.org/abs/1811.11538. arXiv:1811.11538 [cs.DS]. (Cited on page 75.)

[69] Andrew Lucas. Ising formulations of many np problems. *Frontiers in Physics*, 2, 2014. ISSN 2296-424X. doi: 10.3389/fphy.2014.00005. URL https://www.frontiersin.org/articles/10.3389/fphy.2014.00005. (Cited on page 75.)

[70] Hans Feldmeier and Jürgen Schnack. Molecular dynamics for fermions. *Rev. Mod. Phys.*, 72:655–688, Jul 2000. doi: 10.1103/RevModPhys.72.655. URL https://link.aps.org/doi/10.1103/RevModPhys.72.655. (Cited on page 75.)

[71] Alexei Y. Kitaev, A. H. Shen, and Mikhail N. Vyalyi. Classical and quantum computation. In *Graduate Studies in Mathematics*, 2002. (Cited on page 75.)

[72] 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. doi: 10.1126/science.1113479. URL https://www.science.org/doi/abs/10.1126/science.1113479. (Cited on page 75.)

[73] Hefeng Wang, Sabre Kais, Alán Aspuru-Guzik, and Mark R. Hoffmann. Quantum algorithm for obtaining the energy spectrum of molecular systems. *Phys. Chem. Chem. Phys.*, 10:5388–5393, 2008. doi: 10.1039/B804804E. URL http://dx.doi.org/10.1039/B804804E. (Cited on page 75.)

[74] Kishor Bharti, Alba Cervera-Lierta, Thi Ha Kyaw, Tobias Haug, Sumner Alperin-Lea, Abhinav Anand, Matthias Degroote, Hermanni Heimonen, Jakob S. Kottmann, Tim Menke, Wai-Keong Mok, Sukin Sim, Leong-Chuan Kwek, and Alán Aspuru-Guzik. Noisy intermediate-scale quantum algorithms. *Rev. Mod. Phys.*, 94:015004, Feb 2022. doi: 10.1103/RevModPhys.94.015004. URL https://link.aps.org/doi/10.1103/RevModPhys.94.015004. (Cited on pages 76, 78, 81, 84, 86, and 87.)

[75] Alexandre Choquette, Agustin Di Paolo, Panagiotis Kl. Barkoutsos, David Sénéchal, Ivano Tavernelli, and Alexandre Blais. Quantum-optimal-control-inspired ansatz for variational quantum algorithms. *Phys. Rev. Res.*, 3:023092, May 2021. doi: 10.1103/PhysRevResearch.3.023092. URL https://link.aps.org/doi/10.1103/PhysRevResearch.3.023092. (Cited on pages 76, 91, 92, 93, 94, 99, 103, 104, and 105.)

[76] Harper R. Grimsley, Sophia E. Economou, Edwin Barnes, and Nicholas J. Mayhall. An adaptive variational algorithm for exact molecular simulations on a quantum computer. *Nature Communications*, 10(1):3007, 2019. ISSN 2041-1723. doi: 10.1038/s41467-019-10988-2. URL https://doi.org/10.1038/s41467-019-10988-2. (Cited on pages 76, 81, 94, 95, and 96.)

[77] Abhinav Kandala, Antonio Mezzacapo, Kristan Temme, Maika Takita, Markus Brink, Jerry M. Chow, and Jay M. Gambetta. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature*, 549(7671):242–246, 2017. ISSN 1476-4687. doi: 10.1038/nature23879. URL https://doi.org/10.1038/nature23879. (Cited on pages 79 and 87.)

[78] Sukin Sim, Peter D. Johnson, and Alán Aspuru-Guzik. Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms. *Advanced Quantum Technologies*, 2(12):1900070, 2019. doi: https://doi.org/10.1002/qute.201900070. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/qute.201900070. (Cited on page 80.)

[79] Jarrod R McClean, Sergio Boixo, Vadim N Smelyanskiy, Ryan Babbush, and Hartmut Neven. Barren plateaus in quantum neural network training landscapes. *Nature Communications*, 9(1):1–6, 2018. doi: 10.1038/s41467-018-07090-4. URL https://www.nature.com/articles/s41467-018-07090-4. (Cited on pages 80 and 82.)

[80] Martin Larocca, Piotr Czarnik, Kunal Sharma, Gopikrishnan Muraleedharan, Patrick J. Coles, and M. Cerezo. Diagnosing Barren Plateaus with Tools from Quantum Optimal Control. *Quantum*, 6:824, September 2022. ISSN 2521-327X. doi: 10.22331/q-2022-09-29-824. URL https://doi.org/10.22331/q-2022-09-29-824. (Cited on page 80.)

[81] Dave Wecker, Matthew B. Hastings, and Matthias Troyer. Progress towards practical quantum variational algorithms. *Phys. Rev. A*, 92:042303, Oct 2015. doi: 10.1103/PhysRevA.92.042303. URL https://link.aps.org/doi/10.1103/PhysRevA.92.042303. (Cited on pages 80, 84, and 85.)

[82] Masuo Suzuki. Generalized trotter's formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems. *Communications in Mathematical Physics*, 51(2):183 – 190, 1976. doi: 10.1007/BF01609348. URL https://link.springer.com/article/10.1007/BF01609348. (Cited on page 80.)

[83] Andrew G. Taube and Rodney J. Bartlett. New perspectives on unitary coupled-cluster theory. *International Journal of Quantum Chemistry*, 106(15):3393–3401, 2006. doi: https://doi.org/10.1002/qua.21198. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/qua.21198. (Cited on page 81.)

[84] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A quantum approximate optimization algorithm. 2014. URL https://arxiv.org/abs/1411.4028. arXiv:1411.4028 [quant-ph]. (Cited on pages 81 and 98.)

[85] Jakob S. Kottmann, Abhinav Anand, and Alán Aspuru-Guzik. A feasible approach for automatically differentiable unitary coupled-cluster on quantum computers. *Chem. Sci.*, 12:3497–3508, 2021. doi: 10.1039/D0SC06627C. URL http://dx.doi.org/10.1039/D0SC06627C. (Cited on page 81.)

[86] Marcello Benedetti, Delfina Garcia-Pintos, Oscar Perdomo, Vicente Leyton-Ortega, Yunseong Nam, and Alejandro Perdomo-Ortiz. A generative modeling approach for benchmarking and training shallow quantum circuits. *npj Quantum Information*, 5(1):45, May 2019. ISSN 2056-6387. doi: 10.1038/s41534-019-0157-8. URL https://doi.org/10.1038/s41534-019-0157-8. (Cited on page 81.)

[87] Li Li, Minjie Fan, Marc Coram, Patrick Riley, and Stefan Leichenauer. Quantum optimization with a novel gibbs objective function and ansatz architecture search. *Phys. Rev. Res.*, 2:023074, Apr 2020. doi: 10.1103/PhysRevResearch.2.023074. URL https://link.aps.org/doi/10.1103/PhysRevResearch.2.023074. (Cited on page 81.)

[88] Panagiotis Kl. Barkoutsos, Giacomo Nannicini, Anton Robert, Ivano Tavernelli, and Stefan Woerner. Improving Variational Quantum Optimization using CVaR. *Quantum*, 4:256, April 2020. ISSN 2521-327X. doi: 10.22331/q-2020-04-20-256. URL https://doi.org/10.22331/q-2020-04-20-256. (Cited on page 81.)

[89] Lennart Bittel and Martin Kliesch. Training variational quantum algorithms is np-hard. *Phys. Rev. Lett.*, 127:120502, Sep 2021. doi: 10.1103/PhysRevLett.127.120502. URL https://link.aps.org/doi/10.1103/PhysRevLett.127.120502. (Cited on page 82.)

[90] Marco Cerezo, Akira Sone, Tyler Volkoff, Lukasz Cincio, and Patrick J. Coles. Cost function dependent barren plateaus in shallow parametrized quantum circuits. *Nature Communications*, 12(1):1791, 2021. ISSN 2041-1723. doi: 10.1038/s41467-021-21728-w. URL https://doi.org/10.1038/s41467-021-21728-w. (Cited on page 82.)

[91] Andrew Arrasmith, M. Cerezo, Piotr Czarnik, Lukasz Cincio, and Patrick J. Coles. Effect of barren plateaus on gradient-free optimization. *Quantum*, 5:558, October 2021. ISSN 2521-327X. doi: 10.22331/q-2021-10-05-558. URL https://doi.org/10.22331/q-2021-10-05-558. (Cited on page 82.)

[92] M Cerezo and Patrick J Coles. Higher order derivatives of quantum neural networks with barren plateaus. *Quantum Science and Technology*, 6(3):035006, jun 2021. doi: 10.1088/2058-9565/abf51a. URL https://dx.doi.org/10.1088/2058-9565/abf51a. (Cited on page 82.)

[93] Zoë Holmes, Kunal Sharma, M. Cerezo, and Patrick J. Coles. Connecting ansatz expressibility to gradient magnitudes and barren plateaus. *PRX Quantum*, 3:010313, Jan 2022. doi: 10.1103/PRXQuantum.3.010313. URL https://link.aps.org/doi/10.1103/PRXQuantum.3.010313. (Cited on page 82.)

[94] M. Bilkis, M. Cerezo, Guillaume Verdon, Patrick J. Coles, and Lukasz Cincio. A semi-agnostic ansatz with variable structure for quantum machine learning. 2023. URL https://arxiv.org/abs/2103.06712. arXiv:2103.06712 [quant-ph]. (Cited on page 82.)

[95] A V Uvarov and J D Biamonte. On barren plateaus and cost function locality in variational quantum algorithms. *Journal of Physics A: Mathematical and Theoretical*, 54(24):245301, may 2021. doi: 10.1088/1751-8121/abfac7. URL https://dx.doi.org/10.1088/1751-8121/abfac7.

[96] Edward Grant, Leonard Wossnig, Mateusz Ostaszewski, and Marcello Benedetti. An initialization strategy for addressing barren plateaus in parametrized quantum circuits. *Quantum*, 3:214, December 2019. ISSN 2521-327X. doi: 10.22331/q-2019-12-09-214. URL https://doi.org/10.22331/q-2019-12-09-214. (Cited on page 82.)

[97] Ophelia Crawford, Barnaby van Straaten, Daochen Wang, Thomas Parks, Earl Campbell, and Stephen Brierley. Efficient quantum measurement of Pauli operators in the presence of finite sampling error. *Quantum*, 5:385, January 2021. ISSN 2521-327X. doi: 10.22331/q-2021-01-20-385. URL https://doi.org/10.22331/q-2021-01-20-385. (Cited on page 84.)

[98] Tzu-Ching Yen, Vladyslav Verteletskyi, and Artur F. Izmaylov. Measuring all compatible operators in one series of single-qubit measurements using unitary transformations. *Journal of Chemical Theory and Computation*, 16(4):2400–2409, 2020.

doi: 10.1021/acs.jctc.0c00008. URL https://doi.org/10.1021/acs.jctc.0c00008. PMID: 32150412. (Cited on page 84.)

[99] Jarrod R McClean, Jonathan Romero, Ryan Babbush, and Alán Aspuru-Guzik. The theory of variational hybrid quantum-classical algorithms. *New Journal of Physics*, 18 (2):023023, feb 2016. doi: 10.1088/1367-2630/18/2/023023. URL https://dx.doi.org/10.1088/1367-2630/18/2/023023. (Cited on pages 84 and 87.)

[100] Vladyslav Verteletskyi, Tzu-Ching Yen, and Artur F. Izmaylov. Measurement optimization in the variational quantum eigensolver using a minimum clique cover. *The Journal of Chemical Physics*, 152(12):124114, 03 2020. ISSN 0021-9606. doi: 10.1063/1.5141458. URL https://doi.org/10.1063/1.5141458.

[101] Xavier Bonet-Monroig, Ryan Babbush, and Thomas E. O'Brien. Nearly optimal measurement scheduling for partial tomography of quantum states. *Phys. Rev. X*, 10: 031064, Sep 2020. doi: 10.1103/PhysRevX.10.031064. URL https://link.aps.org/doi/10.1103/PhysRevX.10.031064. (Cited on page 84.)

[102] Tzu-Ching Yen and Artur F. Izmaylov. Cartan subalgebra approach to efficient measurements of quantum observables. *PRX Quantum*, 2:040320, Oct 2021. doi: 10.1103/PRXQuantum.2.040320. URL https://link.aps.org/doi/10.1103/PRXQuantum.2.040320. (Cited on page 84.)

[103] Gian Giacomo Guerreschi and Mikhail Smelyanskiy. Practical optimization for hybrid quantum-classical algorithms. 2017. URL https://arxiv.org/abs/1701.01450. arXiv:1701.01450 [quant-ph]. (Cited on page 86.)

[104] Jonathan Romero, Ryan Babbush, Jarrod R McClean, Cornelius Hempel, Peter J Love, and Alán Aspuru-Guzik. Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz. *Quantum Science and Technology*, 4(1): 014008, oct 2018. doi: 10.1088/2058-9565/aad3e4. URL https://dx.doi.org/10.1088/2058-9565/aad3e4. (Cited on page 87.)

[105] Kosuke Mitarai and Keisuke Fujii. Methodology for replacing indirect measurements with direct measurements. *Phys. Rev. Res.*, 1:013006, Aug 2019. doi: 10.1103/PhysRevResearch.1.013006. URL https://link.aps.org/doi/10.1103/PhysRevResearch.1.013006.

[106] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii. Quantum circuit learning. *Phys. Rev. A*, 98:032309, Sep 2018. doi: 10.1103/PhysRevA.98.032309. URL https://link.aps.org/doi/10.1103/PhysRevA.98.032309. (Cited on page 86.)

[107] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J. Love, Alán Aspuru-Guzik, and Jeremy L. O'Brien. A variational eigenvalue solver on a photonic quantum processor. *Nature Communications*, 5(1):4213, 2014. ISSN 2041-1723. doi: 10.1038/ncomms5213. URL https://doi.org/10.1038/ncomms5213. (Cited on page 86.)

[108] J.C. Spall. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Transactions on Automatic Control*, 37(3): 332–341, March 1992. doi: 10.1109/9.119632. URL https://ieeexplore.ieee.org/document/119632. (Cited on page 87.)

[109] Sami Khairy, Ruslan Shaydulin, Lukasz Cincio, Yuri Alexeev, and Prasanna Balaprakash. Reinforcement-learning-based variational quantum circuits optimization for combinatorial problems. 2019. URL https://arxiv.org/abs/1911.04574v1. arXiv:1911.04574 [cs.LG]. (Cited on page 87.)

[110] Guillaume Verdon, Michael Broughton, Jarrod R. McClean, Kevin J. Sung, Ryan Babbush, Zhang Jiang, Hartmut Neven, and Masoud Mohseni. Learning to learn with quantum neural networks via classical neural networks. 2019. URL https://arxiv.org/abs/1907.05415. arXiv:1907.05415 [quant-ph].

[111] Max Wilson, Sam Stromswold, Filip Wudarski, Stuart Hadfield, Norm M. Tubman, and Eleanor Rieffel. Optimizing quantum heuristics with meta-learning, 2019. URL https://arxiv.org/abs/1908.03185. arXiv:1908.03185 [quant-ph].

[112] Jiahao Yao, Paul Köttering, Hans Gundlach, Lin Lin, and Marin Bukov. Noise-robust end-to-end quantum control using deep autoregressive policy networks. 2020. URL https://arxiv.org/abs/2012.06701. arXiv:2012.06701 [quant-ph]. (Cited on page 87.)

[113] Stasja Stanisic, Jan Lukas Bosse, Filippo Maria Gambetta, Raul A. Santos, Wojciech Mruczkiewicz, Thomas E. O'Brien, Eric Ostby, and Ashley Montanaro. Observing ground-state properties of the fermi-hubbard model using a scalable algorithm on a quantum computer. *Nature Communications*, 13(1):5743, 2022. ISSN 2041-1723. doi: 10.1038/s41467-022-33335-4. URL https://doi.org/10.1038/s41467-022-33335-4. (Cited on page 87.)

[114] Xavier Bonet-Monroig, Hao Wang, Diederick Vermetten, Bruno Senjean, Charles Moussa, Thomas Bäck, Vedran Dunjko, and Thomas E. O'Brien. Performance comparison of optimization methods on variational quantum algorithms. *Phys. Rev. A*, 107:032407, Mar 2023. doi: 10.1103/PhysRevA.107.032407. URL https://link.aps.org/doi/10.1103/PhysRevA.107.032407. (Cited on page 87.)

[115] Alicia B. Magann, Kenneth M. Rudinger, Matthew D. Grace, and Mohan Sarovar. Feedback-based quantum optimization. *Phys. Rev. Lett.*, 129:250502, Dec 2022. doi: 10.1103/PhysRevLett.129.250502. URL https://link.aps.org/doi/10.1103/PhysRevLett.129.250502. (Cited on page 87.)

[116] Alicia B. Magann, Kenneth M. Rudinger, Matthew D. Grace, and Mohan Sarovar. Lyapunov-control-inspired strategies for quantum combinatorial optimization. *Phys. Rev. A*, 106:062414, Dec 2022. doi: 10.1103/PhysRevA.106.062414. URL https://link.aps.org/doi/10.1103/PhysRevA.106.062414.

[117] Alicia B. Magann, Sophia E. Economou, and Christian Arenz. Randomized adaptive quantum state preparation. *Phys. Rev. Res.*, 5:033227, Sep 2023. doi:

10.1103/PhysRevResearch.5.033227. URL https://link.aps.org/doi/10.1103/PhysRevResearch.5.033227. (Cited on page 87.)

[118] James B. Larsen, Matthew D. Grace, Andrew D. Baczewski, and Alicia B. Magann. Feedback-based quantum algorithm for ground state preparation of the fermi-hubbard model, 2023. URL https://arxiv.org/abs/2303.02917. arXiv:2303.02917 [quant-ph]. (Cited on page 87.)

[119] Elliott H. Lieb and F. Y. Wu. Absence of mott transition in an exact solution of the short-range, one-band model in one dimension. *Phys. Rev. Lett.*, 20:1445–1448, Jun 1968. doi: 10.1103/PhysRevLett.20.1445. URL https://link.aps.org/doi/10.1103/PhysRevLett.20.1445. (Cited on page 88.)

[120] Thomas Schäfer, Nils Wentzell, Fedor Šimkovic, Yuan-Yao He, Cornelia Hille, Marcel Klett, Christian J. Eckhardt, Behnam Arzhang, Viktor Harkov, François-Marie Le Régent, Alfred Kirsch, Yan Wang, Aaram J. Kim, Evgeny Kozik, Evgeny A. Stepanov, Anna Kauch, Sabine Andergassen, Philipp Hansmann, Daniel Rohe, Yuri M. Vilk, James P. F. LeBlanc, Shiwei Zhang, A.-M. S. Tremblay, Michel Ferrero, Olivier Parcollet, and Antoine Georges. Tracking the footprints of spin fluctuations: A multimethod, multimessenger study of the two-dimensional hubbard model. *Phys. Rev. X*, 11:011058, Mar 2021. doi: 10.1103/PhysRevX.11.011058. URL https://link.aps.org/doi/10.1103/PhysRevX.11.011058. (Cited on page 89.)

[121] Susumu Yamada, Toshiyuki Imamura, and Masahiko Machida. 16.447 tflops and 159-billion-dimensional exact-diagonalization for trapped fermion-hubbard model on the earth simulator. In *Proceedings of the 2005 ACM/IEEE Conference on Supercomputing*, SC '05, page 44, USA, 2005. IEEE Computer Society. ISBN 1595930612. doi: 10.1109/SC.2005.1. URL https://doi.org/10.1109/SC.2005.1. (Cited on page 89.)

[122] P. Jordan and E. Wigner. Uber das paulische aquivalenzverbot. *Z. Physik*, 47: 631–651, Sep 1928. doi: 10.1007/BF01331938. URL https://link.springer.com/article/10.1007/BF01331938. (Cited on page 89.)

[123] Jacob T. Seeley, Martin J. Richard, and Peter J. Love. The Bravyi-Kitaev transformation for quantum computation of electronic structure. *The Journal of Chemical Physics*, 137(22):224109, 12 2012. ISSN 0021-9606. doi: 10.1063/1.4768229. URL https://doi.org/10.1063/1.4768229. (Cited on page 89.)

[124] Sergey B. Bravyi and Alexei Yu. Kitaev. Fermionic quantum computation. *Annals of Physics*, 298(1):210–226, 2002. ISSN 0003-4916. doi: https://doi.org/10.1006/aphy.2002.6254. URL https://www.sciencedirect.com/science/article/pii/S0003491602962548. (Cited on page 89.)

[125] John S. Van Dyke, George S. Barron, Nicholas J. Mayhall, Edwin Barnes, and Sophia E. Economou. Scaling adaptive quantum simulation algorithms via operator pool tiling. 2022. URL https://arxiv.org/abs/2206.14215. arXiv:2304.09111 [quant-ph]. (Cited on page 96.)

[126] V. O. Shkolnikov, Nicholas J. Mayhall, Sophia E. Economou, and Edwin Barnes. Avoiding symmetry roadblocks and minimizing the measurement overhead of adaptive variational quantum eigensolvers. *Quantum*, 7:1040, June 2023. ISSN 2521-327X. doi: 10.22331/q-2023-06-12-1040. URL https://doi.org/10.22331/q-2023-06-12-1040. (Cited on page 96.)

[127] Harper R. Grimsley, George S. Barron, Edwin Barnes, Sophia E. Economou, and Nicholas J. Mayhall. Adaptive, problem-tailored variational quantum eigensolver mitigates rough parameter landscapes and barren plateaus. *npj Quantum Information*, 9(1):19, 03 2023. ISSN 2056-6387. doi: 10.1038/s41534-023-00681-0. URL https://doi.org/10.1038/s41534-023-00681-0. (Cited on page 96.)

[128] Ho Lun Tang, V.O. Shkolnikov, George S. Barron, Harper R. Grimsley, Nicholas J. Mayhall, Edwin Barnes, and Sophia E. Economou. Qubit-adapt-vqe: An adaptive algorithm for constructing hardware-efficient ansätze on a quantum processor. *PRX Quantum*, 2:020310, Apr 2021. doi: 10.1103/PRXQuantum.2.020310. URL https://link.aps.org/doi/10.1103/PRXQuantum.2.020310. (Cited on pages 96 and 97.)

[129] Linghua Zhu, Ho Lun Tang, George S. Barron, F. A. Calderon-Vargas, Nicholas J. Mayhall, Edwin Barnes, and Sophia E. Economou. An adaptive quantum approximate optimization algorithm for solving combinatorial problems on a quantum computer. 2022. URL https://arxiv.org/abs/2005.10258. arXiv:2005.10258 [quant-ph]. (Cited on page 98.)

[130] Panagiotis G. Anastasiou, Yanzhu Chen, Nicholas J. Mayhall, Edwin Barnes, and Sophia E. Economou. Tetris-adapt-vqe: An adaptive algorithm that yields shallower, denser circuit ansätze. 2022. URL https://arxiv.org/abs/2209.10562. arXiv:2209.10562 [quant-ph]. (Cited on pages 96 and 98.)

[131] Alexandre Choquette-Poitevin. Algorithmes variationnels quantiques pour la simulation de systèmes quantiques, 2020. Master's thesis, Université de Sherbrooke. (Cited on page 102.)

[132] Qiskit contributors. Qiskit: An open-source framework for quantum computing, 2023. URL https://zenodo.org/records/6027041. (Cited on page 104.)

[133] J. R. Schrieffer and P. A. Wolff. Relation between the anderson and kondo hamiltonians. *Phys. Rev.*, 149:491–492, Sep 1966. doi: 10.1103/PhysRev.149.491. URL https://link.aps.org/doi/10.1103/PhysRev.149.491. (Cited on page 121.)