

ANALOG COMPUTING WITH OPTICAL AND MICROWAVE BOSONIC SYSTEMS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Alen Senanian

August 2024

© 2024 Alen Senanian

ALL RIGHTS RESERVED

ANALOG COMPUTING WITH OPTICAL AND MICROWAVE BOSONIC SYSTEMS

Alen Senanian, Ph.D.

Cornell University 2024

Moore's law continues to push the boundaries of capabilities with today's digital electronics, albeit with a much slower rate than decades prior. The end of Dennard scaling has similarly made digital electronics difficult to continue scaling with energy efficiency. The breakdown of these two observations first made in the early days of computing have led to important consequences in today's computing needs: Large-scale compute needed by artificial intelligence systems now require warehouses full of computers, and embedded smart-sensors for edge computing are limited by the energy efficiency of digital electronics. Analog computers have emerged as a platform for performing sensing and machine learning tasks owing to their energy efficiency, the ability to interface directly with the analog world, and the robustness of certain tasks like machine learning to hardware imperfections. In this thesis, we present two experiments that demonstrate two novel applications of analog computing with physical systems. In the first experiment, we construct a highly-multimode frequency domain fiber laser that is capable of simulating physics in two- and three-dimensional large-scale lattices. We leverage the programmability and scale of our simulator to study exotic condensed matter phenomena, such as time-reversal symmetry-breaking, non-Hermitian physics, and dynamics in non-Euclidean geometries. In this work, we simulate lattices with up to 100,000 sites – orders of magnitude greater than previously achieved in photonic simulators. In the second experiment, we describe and perform a proof-of-principle demonstration of a new form of appli-

cation for quantum devices. In between the fields of quantum sensing and quantum computation, we perform microwave signal processing on ultra-low power signals, and propose a route towards achieving a quantum computational-sensing advantage: a quantum advantage in performing a computational task on analog signals that are inaccessible to any classical receiver. Our results provide the first step towards achieving such an advantage.

BIOGRAPHICAL SKETCH

Alen Senanian grew up in Glendale, California and graduated from Glendale High school in 2009. He then attended Pasadena City College studying electrical engineering after working as an electrician for a few years. Deeply inspired by his first foray into quantum mechanics, especially puzzled by the measurement problem, he changed his major to physics before transferring to University of California Santa Cruz in 2014. Alen received a B.S. in Physics in 2016. Alen began his research interests in University of California Santa Cruz under the mentorship of Onuttom Narayan, studying non-ergodic dynamics in classical coupled oscillators. Moving to Ithaca, NY to study physics at Cornell University, Alen joined Peter McMahon's group to build experimental setups, including quantum systems.

This document is dedicated to Sara Senanian.

ACKNOWLEDGEMENTS

First, thanks to Onuttom Narayan and Josh Deutsch, whose mentorship in my undergraduate years was instrumental in cultivating my passion for physics.

Thanks to Chris Wilson, Kripa Raman, Florian Theuss, Saba Karimeddiny, Sam Bright-Thonney, David Hathcock, Bozo Vareskic, Maciej Olszewski for the endless entertainment during my time at Cornell. Thanks to Maya Martirossyan for the many dinners, annual easter brunches, and providing a link to my Armenian roots while stranded in upstate New York. Thanks to Bill and Carrol Wilson for the thanksgiving dinners and for teaching me how to sail on Seneca lake. Thanks to Jerair Barmaksezian, Mery Bayamdzhyan, Vatche Bahudian, Marian Kourieh for the multi-decade friendship that you have kept up despite being on the other side of the country.

Thanks to Robert Page, our machine shop manager at LASSP, for agreeing to machine last-minute parts several times. I owe you a few beers.

Thanks to the McMahon group. In particular, thanks to Sridhar Prabhu for spending Christmas holiday helping me finish the QRC paper in time, and for countless other things. Thanks to Vladimir Kremenetski for his passion and curiosuty for physics and math, and for always being ready to discuss abstract concepts. Thanks to Shi-Yuan Ma, Federico Presutti, Maxwell Anderson, and the rest of the McMahon team for the thoughtful discussions and entertainment in lab.

Thanks to Tiyanu Wang for the entertainment at late hours in BD03, and for all the career advice. Thanks to Tatsuhiko Onodera for the constant reminder that, despite the demands of the academic life, physics is really fun. Thanks to Logan Wright for teaching me to be a responsible scientist, and that no idea is really too big.

Thanks to Valla Fatemi for all the technical experience I have gained working with Valla. Thanks to Valla for being in the weeds in the experiments with me, always ready to answer any question at any time of day.

Thanks to Peter McMahon for running a really special group. To be a part of Peter's group is an opportunity to find creative solutions to really big problems, with a high bar for quality research, and it's an opportunity to learn everything there is to be a scientist and running a lab. I feel very lucky to have been a part of this group.

Thanks to my family for the constant support, and for emphasizing education. Thanks to my parents for the big gamble to immigrate to the United States well into their 50s. Thanks to my sister Sara Senanian who, after accepting an offer to study at UC Berkeley for mechanical engineering just three years after immigrating to the United States, provided a constant stream of inspiration to myself and many of my cousins.

TABLE OF CONTENTS

Biographical Sketch	iii
Dedication	iv
Acknowledgements	v
Table of Contents	vii
1 Introduction	1
1.1 Analog computing	1
1.1.1 Analog computing with microwave and visible optics	4
1.2 Outline of this thesis	6
2 Background	8
2.1 Optical ring resonators	8
2.2 Superconducting microwave circuits for quantum computing	12
3 Programmable large-scale simulation of bosonic transport in optical synthetic frequency lattices	16
3.1 Introduction	16
3.2 Results	20
3.2.1 Measurements of band structure	22
3.2.2 Preparing arbitrary input states	23
3.2.3 Simulating a lattice with more than 100k lattice sites	25
3.3 Discussion	28
3.4 Experimental setup	30
3.4.1 Setup characterization	32
3.4.2 Real-space-occupation measurements	36
3.4.3 Band-structure measurements	39
3.4.4 Input-state preparation	44
3.5 Supplementary results	47
4 Microwave signal processing using an analog quantum reservoir computer	53
4.1 Introduction	53
4.1.1 Experimental setup and protocols	57
4.2 Results	60
4.2.1 Classification of time-independent signals	60
4.2.2 Classification of radio-frequency (RF) communication modulation schemes	62
4.2.3 Classification of filtered noise	64
4.3 Discussion	68
4.4 Experimental setup	70
4.5 System Hamiltonian & Reservoir description	73
4.5.1 Hamiltonian description	73
4.5.2 Reservoir description for time-independent signals	73

4.5.3	Reservoir description for slow varying time-dependent signals	79
4.6	Quantum reservoir characterization	82
4.6.1	CNOD	82
4.6.2	Reservoir unitary characterization	84
4.6.3	Qubit & parity measurements	87
4.6.4	Tuning T_2 via resonator-induced dephasing	88
4.7	Machine learning with the quantum reservoir	90
4.7.1	Output feature encoding	90
4.7.2	Training the linear layer	94
4.8	Supplementary information machine learning tasks	97
4.8.1	Classification of Radio-Frequency signals	97
4.8.2	Classification of noisy signals	98
4.9	Simulation of the quantum reservoir	100
4.9.1	Introduction	100
4.9.2	The advantage of continuous-time continuous-variable QRCs over discrete-time qubit-based QRCs	102
4.9.3	Comparison to other reservoirs	104
4.9.4	Multi-qubit reservoirs	106
4.10	Theoretical analysis of the expressivity of our QRC for time- independent signals	109
4.11	Leaky Echo State Networks (LESN)	112
4.11.1	Background	112
4.11.2	Digital reservoir comparison	113
5	Outlook	116
5.1	Toward the construction of a large-scale frequency domain neural net- work accelerator	116
5.2	Toward an advantage in sensing room-temperature signals using mi- crowave quantum oscillators	119
Bibliography		125

CHAPTER 1

INTRODUCTION

1.1 Analog computing

For much of computing history, computing has taken its form in physical, dynamical systems. In ancient through medieval times, physical systems were built to perform computations, whether to perform abstract numeric calculations or directly predict trajectories of physical systems, e.g. astrological bodies, by analogy. Examples of these early analog computers can be found across civilizations such as the Sumerian abacus, the south-pointing chariot from 5th-century BCE China [37], or the from 3rd-century BCE Greece [36]. One of the most successful analog computers today, the clock, is a dynamical system that serves as a highly-abstracted analog of our solar system, calculating the trajectories of the Earth's rotation and orbit around the sun.

Analog computers continued to develop in sophisticated mechanical, hydrodynamic, and other platforms until the invention of the transistor after the second world war, ushering in the era of digital computing. The incredible success of the transistor, likely the most impactful piece of technology ever created, enabled exponential improvement of computing power year over year, displacing analog computers and monopolizing the way we compute. This exponential growth in computing power, famously predicted in 1965 by Gordon Moore and subsequently known as *Moore's Law*, is owed to the doubling of transistors that can fit on a single chip every two years. A related scaling, known as *Dennard scaling*, was the observation by Robert Dennard that the power density was constant for a chip area even as you increase the number of transistors. The combination of Moore's law and Dennard scaling led to decades of sustained exponential growth in computing performance, leading to new

classes of computers, applications, and industries.

The technological growth propelled by Moore’s Law and Dennard scaling have in recent years begun to slow or end. Dennard scaling ended around 2006 [15], and has since left the industry to focus largely on multi-core processors to achieve energy efficiency. Moore’s law in turn has begun to slow down, effectively making the number of transistors double every 20 years, as opposed to every 2 years as it did at the end of the twentieth century [51]. Despite this technological slowing, the appetite for compute has skyrocketed in recent decades due to the rise in popularity of large AI models. The amount of computation required to train large models has been doubling every 3.4 months since 2012 [101], far outpacing the growth of processor performance even in its period of fastest growth at the end of the twentieth century. In turn, those tasked with training these models have resorted to hosting these models in large warehouses containing 100,000 servers acting as one single gigantic computer. In an entirely separate vein, the demand for small-scale but power-efficient computing has also grown as new applications in edge computing have emerged. In edge computing, small scale computations are performed on edge devices that can be equipped with sensors, and can therefore processes data directly and make real-time decisions before sending it to a larger cluster. Example trends in edge computing include smart cameras [16], in-sensor computing [137], and machine-enabled smart sensors [5]. Crucially, these technologies relies on highly energy-efficient compute. Whether the application concerns large-scale machine learning or low-power edge computing, the lack of a sustained Moore’s law and Dennard scaling have brought significant challenges to present-day digital computing.

In recent years, analog computing has been undergoing an renaissance, in part due to the challenges stated above faced by digital computing, but also due to a ma-

turity in emergent technologies outside the semiconductor industry such as telecommunications optics and other platforms which are less suited for digital computation. Additionally, quantum computation, a special kind of analog computing that promises enormous strides in both computing and sensing, has seen great progress in the past decade. Here, we describe quantum computers as analog since, even in gate-based paradigms, gates and states are parameterized by continuous-valued numbers. While digital computation will likely remain the only way to perform fault-tolerant (classical) computation in the foreseeable future, analog computers have gathered interest in the last decade as special-purpose devices that are highly adept at solving specific computational tasks – tasks that do not require perfect, error-free calculations. AI models, for example, do not require fault-tolerant computation, with some recent successful models such as diffusion-denoising probabilistic models [52] explicitly relying on built-in noise. Other example tasks that are robust to noise and hardware imperfections include simulations of differential equations and other applications of scientific computing, combinatorial optimization [91], and cryptography [97]. The computations in these tasks can have some noise while still being useful, making them good candidates for analog computers. What makes analog computers potentially better suited for the above tasks over their digital counterparts is due to three primary features inherent to analog computing. First, analog computers can readily make use of multiple degrees of freedom to perform massively parallel computations. Second, given the real world is analog and not digital, analog computers can directly interface with real-world data and process it in real-time, similar to today’s digital edge devices, but without the need to perform the analog-to-digital conversion. Finally, to solve a particular problem, the analog computer can be constructed to model the problem directly, such as the evolution of a given dynamical system, without the use of extra resources. Finally, platform-specific benefits such as very high energy

efficiency can make analog computers uniquely positioned to solve certain tasks.

The focus of this thesis is the construction of two analog computers built on two very different technologies, and showcase them on solving two highly-specialized, but very different problems. The first system is an multi-mode optical system that leverages the high bandwidths inherent to optics to perform large-scale analog simulations of lattice systems. The second system we construct is a small-scale but highly sensitive quantum mechanical system that performs machine learning tasks on ultra-low power microwave signals. While both of the systems studied in this thesis rely on different technologies, the studies on each of these systems offer a proof-of-principle demonstration of using an analog computer to solve a particular problem leveraging benefits unique to that technology.

1.1.1 Analog computing with microwave and visible optics

The recent interest in building optical analog computers follows several decades of an optical-communications boom, as well as thrusts in optical technology from advancing consumer electronics, such as cameras in smartphones and tablets. This commercial technological growth has been complemented by both academic developments in photonic integrated circuits [34] and exotic photonic materials [14]. These developments have enabled researchers to begin to address the limits faced by digital computers in modeling power-hungry neural networks, but also analog simulators and combinatorial optimizers.

Optical analog computers come with particular benefits that make it an attractive platform over other platforms for specific tasks; here, we discuss some of these benefits. In particular, extra large bandwidths, low loss, and the minimal cross-talk are

primary features of optical systems that we will discuss and that the work in this thesis leverages. For a more complete discussion on the benefits of optics, especially as it pertains to performing machine learning using analog optical systems, see Ref. [87].

In an extremely crude sense, visible optics¹ is similar to microwave optics, except the waves have a much shorter wavelength. Of course, this deceptively small difference leads to enormous differences in not just practical technological aspects, but fundamental ones too. Nonetheless, this is a physical truth with important practical consequences. For example, linear spatial structures capable of hosting light, e.g. resonators, are typically orders of magnitude smaller than those for microwaves, enabling compact and scalable structures in photonic integrated circuits. Another important consequence of the much shorter wavelength, in particular for quantum technologies, is that the energies required to excite photons at visible wavelengths is much larger than the thermal energy at room temperatures. This important fact enables quantum technologies at room temperatures, whereas microwave quantum devices can only operate in environments at a few millikelvin, which would otherwise be bombarded by thermal photons.

Both microwave and photonic technologies operate in a bandwidth on the order of the frequencies of their respective band. This gives optical platforms bandwidths many orders of magnitude larger than what is achievable in digital microwave electronics, which are typically limited to a few GHz. Even with analog microwaves, this bandwidth increases by only another factor of 5-10. One can leverage the large bandwidth at visible frequencies to construct systems with a number of degrees of freedom in the millions, something that is completely out of reach in the microwave regime. With global operations acting across optical bandwidths, this enables parallel

¹We will call visible optics as “optics” or “photonic” for simplicity. When microwave optics are discussed, the word microwave will always precede the word optics.

operations with extremely high throughput.

Due to the Coulombic interaction energies, nonlinear dynamics can be readily constructed in electronics, both analog and digital. Nonlinear materials at optical wavelengths on the other hand are very challenging to produce. On the other hand, the difficulty in light at optical frequencies interacting with matter comes with other benefits. Namely, light can travel in vacuum or other dielectric materials with remarkably low loss. Additionally, there is very little potential for cross-talk between different traveling waves. These benefits makes optics a very attractive platform for analog computing. For quantum applications, the lack of strong non-linearities has been a challenge for optical frequencies, despite the low loss available to optical approaches. As a result, quantum computing in the photonic domain has chiefly relied on single-photon detectors as the source of non-Gaussianity for compiled quantum circuits, and this has led to the development of measurement-based quantum computation (MBQC). Microwave quantum circuits on the other hand have benefit from the strong nonlinearity imbued by Josephson junctions, and rely on superconductivity to keep losses in the physical circuits to a minimum.

1.2 Outline of this thesis

Chapter 2 provides an introduction to the background tools needed in our study of analog simulators in classical visible optics and quantum microwave optics. We first derive the coupled mode equations for a ring resonator with an embedded amplifier and phase modulator from Maxwell's equations. We then introduce superconducting qubits as a platform for microwave quantum optics and briefly discuss the physics of transmons, the most popular implementation of a superconducting qubit.

Chapter 3 discusses our work in the construction of a large-scale photonic simulator in the frequency domain. In this system, linear bosonic lattice systems are simulated by encoding lattice sites in the longitudinal modes of a very long fiber ring cavity. Using a nonlinear element, these modes are then coupled together to host a variety of programmable lattices in one-, two-, and three-dimensions. We study and measure the band structure in each of the lattices we instantiate and study the dynamics of on these lattices for various excitations. Leveraging the programmability and especially the long-range connectivity, we instantiate systems with broken time-reversal symmetry, non-Hermitian transport, and non-Euclidian geometry in lattices with 10,000-100,000 lattices sites.

Chapter 4 discusses our work in a proof-of-concept demonstration of a new application between the fields of quantum sensing and quantum computing. Using a simple qubit-oscillator system, we use a reservoir computing framework to perform non-trivial signal-processing on ultra-low power signals. We evaluate the possibility of a new form of quantum advantage, a computational-sensing advantage, where we imagine a device capable of performing a computational task on signals that are outside the detection ability of any classical receiver. Our experiments in this chapter provide a proof-of-principle demonstration towards building such a device.

CHAPTER 2

BACKGROUND

This chapter provides a basic introduction to the physics of optical ring resonators and superconducting qubit circuits that much of the rest of the dissertation is based upon. We will first introduce ring resonators, optical amplifiers, and electro-optical modulation within the context of how each contribute to the large bandwidths we leverage in constructing an ultra-large photonic simulator. Next, we will introduce the physics of superconducting qubits, particularly, that of the transmon qubit. We will discuss how transmons can form a qubit subspace, then discuss the physics of a combined oscillator-qubit system.

2.1 Optical ring resonators

Light is a time-dependent electromagnetic field with wavelengths that span from sub-nanometers to kilometers and larger. Here we introduce the physics of electromagnetic fields in optical resonators. Electromagnetic fields, composed of an electric field \mathbf{E} and a magnetic field \mathbf{B} , are described by Maxwell's equations:

$$\begin{aligned}\nabla \cdot \mathbf{E} &= \frac{\rho}{\epsilon_0} \\ \nabla \cdot \mathbf{B} &= 0 \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\ \nabla \times \mathbf{B} &= \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t},\end{aligned}$$

where ρ is a charge density, \mathbf{J} is the current density, $\epsilon_0 = 8.85 \times 10^{-12}$ F/m is the electrical permittivity in vacuum and $\mu_0 = 4\pi \times 10^{-7}$ H/m is the magnetic permeability in vacuum. In the bulk of a dielectric medium at optical frequencies, there are no free

charges or currents, i.e. $\rho = 0$ and $\mathbf{J} = 0$, and Maxwell's equations can be reduced to Maxwell's wave equation:

$$\nabla^2 \mathbf{E} = \frac{1}{\epsilon_0 c^2} \frac{\partial^2 \mathbf{P}}{\partial t^2} + \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2}, \quad (2.1)$$

where $c = 1/\sqrt{\epsilon_0 \mu_0}$ is the speed of light, and \mathbf{P} is the polarization field describing the response of electrons in a material to an externally applied electric field:

$$\mathbf{P} = \epsilon_0 (\chi^{(1)} \mathbf{E} + \chi^{(2)} \mathbf{E}^2 + \dots). \quad (2.2)$$

The χ 's are tensors that describe the susceptibility of the material. The first term in the sum is the linear response of a material. The second and third term are the most commonly studied responses, and underlies much of non-linear visible optics. In microwave, where non-linearities can become extreme, higher order terms in the above sum can become important.

In a ring resonator, the waves propagate in one dimension within a linear medium with periodic boundary conditions such that $E(x, t) = E(x + L, t)$ for a resonator of length L . One can rewrite Maxwell's equation for an electric field with a fixed polarization along z traveling in a ring resonator with periodic boundary conditions:

$$\frac{2\pi c}{nL} \frac{\partial E_z}{\partial \phi} = -\frac{\partial E_z}{\partial t} - E_z \frac{1}{n} \frac{\partial n}{\partial t} \quad (2.3)$$

where ϕ is the coordinate along the resonator with $\phi = \phi + 2\pi$, and $n = \sqrt{1 + \chi^{(1)}}$ is the refractive index of the material, which in general can depend on ϕ and t . Expanding the electric field within the resonator in terms of the resonator modes, we obtain

$$E_z(\phi, t) = \sum_m C_m a_m(t) e^{-im\phi} + c.c. \quad (2.4)$$

The prefactors C_m are fixed so that $|a_m|^2$ represents the photon number flux for mode m . Substituting this into Maxwell's equations, for a homogenous and time-independent material, i.e. $n(\phi, t) = n$, and making use of the orthogonality of $e^{-im\phi}$,

we arrive at the dynamical equation for the mode amplitudes

$$\dot{a}_m(t) = i\omega_m a_m(t). \quad (2.5)$$

The eigenvalues $\omega_m = m \times 2\pi c/nL$ are the frequencies of the m th mode in the resonator, and form an equally-spaced comb of frequencies.

Equation 2.5 describes the time-evolution of modes in a ring cavity in the absence of loss. Loss is typically mitigated with an optical amplifier, which typically introduce nonlinear effects. The nonlinear gain and loss in a ring resonator are typically modeled as

$$\dot{a}_m(t) = \left(\frac{g}{1 + u/E_s} - \ell \right) a_m(t) \quad (2.6)$$

where g is the linear gain, ℓ is the loss, and $u = \sum_m |a_m|^2$ is the power in the cavity. Equation 2.6 is related to Eq. 2.5 with the addition of the gain and loss terms, in addition to the rotating frame transformation $a_m \rightarrow a_m e^{i\omega_m t}$. The non-trivial steady state solution to Eq. 2.6 is given by $\sum_m |a_m(t)|^2 = E_s(\frac{g}{\ell} - 1)$. This solution describes a stable manifold on a hypersphere in the space of amplitudes centered at $\{a_m(t)\} = 0$ with radius $\sqrt{E_s(g/\ell - 1)}$. This regime describes a laser that is above threshold, with intracavity power u proportional to the gain g . The trivial solution with $\{a_m(t)\} = 0$ describes a laser that is below threshold, with the transition occurring at $g = \ell$.

To summarize the discussion so far: with an optical amplifier residing inside a cavity, we can sustain a set of equally spaced modes. These modes are separated in frequency by an amount inversely proportional to the resonator length, and will generally span a bandwidth supported by the amplifier. Optical amplifiers can generally support bandwidths of 10-100s of THz. A fiber loop cavity that is a couple hundred meters long will have mode spacing on the order of a few MHz. By adding an optical amplifier within such a fiber loop enables the support of potentially millions of frequency modes that could encode and process information. The discussion so far

however, has only touched on the ability to encode or store information.

To manipulate information in a fiber loop cavity, there are a few avenues. Here, we will only cover the use of an electro-optic modulator to manipulate the light stored in the longitudinal modes in a fiber cavity. The electro-optic effect is the nonlinear effect in which a material's refractive index will change as a response to an applied electric field. An electro-optic modulator (EOM) is constructed from a material with a strong second-order nonlinearity, typically Lithium Niobate (LiNbO_3). A plane wave passing through an EOM that is being driven by an AC voltage will pick up an additional phase, i.e. $a(t) \rightarrow a(t)e^{i\phi(t)}$, where $\phi(t)$ is proportional to the applied RF bias.

We can model an EOM residing in the cavity by treating the refractive index in a resonator as time-independent and homogeneous, except for a section of length w , i.e. $n(\phi, t) = n_0 + n_1(t)\Theta(\phi/\phi_w)$, where Θ is the Heaviside function and $\phi_w = 2\pi \times 2w/L$. Substituting this into Eq. 2.3 and neglecting the last term, since the RF modulation speeds $\partial n/\partial t$ will be much slower than the dynamics at optical frequencies, we obtain

$$\begin{aligned}\dot{a}_m(t) &= i \frac{c}{L} \sum_l \left[\sqrt{lm} \int_0^{2\pi} \frac{e^{i(l-m)\phi}}{n_0 + n_1(t)\Theta(\phi/\phi_w)} d\phi \right] a_l(t) \\ &\approx i \frac{c}{L} \sum_l \left[\sqrt{lm} \int_0^{2\pi} \frac{e^{i(l-m)\phi}}{n_0} \left(1 - \frac{n_1(t)}{n_0} \Theta(\phi/\phi_w) \right) d\phi \right] a_l(t) \\ &= i\omega_m a_m(t) - i \frac{2\pi w}{n_0^2 L^2} n_1(t) \sum_l \sqrt{lm} \operatorname{sinc} \left(\frac{(l-m)w}{2\pi L} \right) a_l(t)\end{aligned}\quad (2.7)$$

The first term is the same term in Eq. 2.5 describing the oscillation of the mode field in the cavity, and the second term describes the coupling between the modes, with coupling strength determined by the magnitude of the AC refractive index $n_1(t)$. By choosing a modulation with a specific frequency, one can selectively couple modes that are spaced apart by the modulation frequency. To see this, one can substitute $n_1(t) = \cos(\Omega t)$ with $\Omega = \omega_n - \omega_{n+1}$. Then, by going in the rotating frame and dropping

fast oscillating terms, the only terms that will remain are nearest-neighbor modes. A $\cos(k\Omega t)$ modulation tone will couple modes that k apart. A secondary way to see this is to observe that the action of adding a phase to a propagating wave can be described in the basis of frequency modes as

$$U = F e^{iN} F^\dagger \quad (2.8)$$

where N is a diagonal matrix encoding the modulation of the refractive index $n(t)$ such that $N_{ii} \propto n(t = i)$ and $F_{jk} = (e^{-2\pi i/M})^{jk}$ is the discrete Fourier transform over M modes. The equation above sets constraints on the time evolution operator U , namely that it is a unitary Toeplitz operator, that is, a unitary matrix with constant diagonals. Such an operator describes a translationally invariant coupling.

2.2 Superconducting microwave circuits for quantum computing

Here, we give an introduction to the simplest type of superconducting microwave qubit, the transmon qubit [71]. The transmon is composed by a Josephson junction shunted by a capacitance. The corresponding Hamiltonian can be written in the charge basis $\{|n\rangle\}$, where n refers to the number of cooper pairs on the superconducting island and can take any integer value:

$$H = 4E_C(n - n_g)^2 - \frac{E_J}{2} \sum_n |n\rangle\langle n+1| + \text{H.c..} \quad (2.9)$$

The first term is the charging energy of the capacitor, where $E_C = e^2/2C$ is the single electron charging energy of the capacitor with capacitance C , and n_g is an offset charge set by a local voltage biasing line. The second term is the energy associated with Cooper pairs tunneling through the Josephson junction, with a characteristic Josephson energy $E_J = \phi_0^2/2L_J$, where ϕ_0 is the reduced flux quantum and L_J is

the inductance of the junction. The second term in the Hamiltonian of Eq. 2.9 is like a nearest-neighbor tight-binding model, where instead of lattice sites, the term describes the energy required to change the offset charge. For a transmon, $E_J \gg E_C$, and this term dominates. We can therefore diagonalize this term in the flux basis ϕ by using the identity $|n\rangle = \frac{1}{2\pi} \int_0^{2\pi} d\phi e^{-in\phi} |\phi\rangle$:

$$H = 4E_C(n - n_g)^2 - E_J \cos(\phi). \quad (2.10)$$

The success of the transmon is owed largely to its inherent exponential suppression to charge noise, or fluctuations in n_g . To quantize the transmon, we first note that typically in the transmon limit $E_J/E_C > 50$, which puts the first few excitations deep in the cosine well. We can therefore expand the cosine potential to fourth-order to obtain

$$H = 4E_C(n - n_g)^2 - E_J \left(1 - \frac{\phi^2}{2!} + \frac{\phi^4}{4!} - \dots \right). \quad (2.11)$$

We can quantize the above Hamiltonian by introducing the creation and annihilation operators, defined as

$$\begin{aligned} n &= i \left(\frac{E_J}{8E_C} \right)^{1/4} \frac{a - a^\dagger}{\sqrt{2}} \\ \phi &= \left(\frac{E_C}{8E_J} \right)^{1/4} \frac{a + a^\dagger}{\sqrt{2}} \end{aligned}$$

These operators obey the commutation relation $[a, a^\dagger] = 1$. Substituting in these operators into Eq. 2.11 and normal ordering, we obtain

$$H/\hbar = \omega_a a^\dagger a - \alpha a^\dagger a^\dagger a a \quad (2.12)$$

where $\hbar\omega_a = \sqrt{8E_J E_C} - E_C$ is the frequency of the transmon in its ground state, and α is the anharmonicity, or self-Kerr, or the transmon. The first term is that of a harmonic oscillator, and the second term is an anharmonic correction. This anharmonicity is what enables one to isolate a qubit subspace within the full transmon manifold.

This is due to the fact that transitions in an nonlinear oscillator have unique frequencies and therefore can individually be addressed. For a transmon, the anharmonicity typically resides between 150-300 MHz.

Next, we describe the physics of a transmon coupled to an oscillator. The resultant interaction between the two is what will enable us to readout the state of the transmon qubit. A linear coupling of the transmon to the oscillator will result in the two modes hybridizing into a set of eigenmodes. Two important consequences of this hybridization is that (1) the linear oscillator mode will inherent an anharmonicity from the junction, and that (2) the combined system will have a nonlinear interaction.

Adding an additional resonator to the Hamiltonian in Eq. 2.11 with linear coupling to the transmon, we obtain

$$H/\hbar = \omega_t t^\dagger t + \omega_r r^\dagger r + g(t^\dagger r + \text{H.c.}) - \frac{E_J}{4!} \phi^4 \quad (2.13)$$

where we have renamed the transmon creation and annihilation operators from a to t, r and r^\dagger is the creation and annihilation operators for the resonator mode with resonance frequency ω_r , and g is the strength of the linear coupling. In the above equation, we have kept the nonlinearity in the flux basis. Diagonalizing out the third term to remove the linear coupling introduces a new basis \tilde{t} and \tilde{r} . The flux operator in this basis is then $\phi = \sum_n \phi_n (a_n + a_n^\dagger)$, where $\{a_n\} = \{\tilde{t}, \tilde{r}\}$, and the ϕ_n 's are determined by the requirement that the linear part of the Hamiltonian is diagonal in the operators. Substituting in ϕ , we obtain

$$H/\hbar = \omega_t \tilde{t}^\dagger \tilde{t} + \omega_r \tilde{r}^\dagger \tilde{r} - \frac{E_J}{4!} \left(\sum_n \phi_n (a_n + a_n^\dagger) \right)^4 \quad (2.14)$$

By expanding the last term, normal ordering, and removing fast oscillating terms, we obtain the so-called dispersive Hamiltonian

$$H/\hbar = \omega_t \tilde{t}^\dagger \tilde{t} + \omega_r \tilde{r}^\dagger \tilde{r} - \frac{\chi_{tt}}{2} \tilde{t}^\dagger \tilde{t}^\dagger \tilde{t} \tilde{t} - \frac{\chi_{rr}}{2} \tilde{r}^\dagger \tilde{r}^\dagger \tilde{r} \tilde{r} - \chi_{tr} \tilde{t}^\dagger \tilde{t} \tilde{r}^\dagger \tilde{r} \quad (2.15)$$

Here, the second and third term are the renormalized anharmonicity and the inherited anharmonicity of the transmon and resonator respectively, and the final term is the cross-Kerr interaction between them. This last term describes a state-dependent frequency shift of the resonator depending on the state of the transmon, or a frequency shift of the transmon depending on the state of the resonator. As mentioned previously, the dispersive Hamiltonian is what enables readout of the transmon qubit state. By measuring the frequency of a dedicated "readout resonator", one measures and projects the state of the qubit.

CHAPTER 3

PROGRAMMABLE LARGE-SCALE SIMULATION OF BOSONIC TRANSPORT IN OPTICAL SYNTHETIC FREQUENCY LATTICES

This chapter is based on the following publication:

- Senanian, A., Wright, L.G., Wade, P.F., Doyle, H.K. and McMahon, P.L., 2023. Programmable large-scale simulation of bosonic transport in optical synthetic frequency lattices. *Nature Physics*, 19(9), pp.1333-1339.

3.1 Introduction

Simulations have long been used to understand emergent phenomena in complex many-body systems. Special-purpose analog simulators trade off the generality of digital implementations for either scalability or access to regimes challenging for digital computers. In this regard, photonic analog simulators [74, 79, 48, 92, 122, 121] complement developments in platforms like superconducting circuits [58, 67] and ultracold atoms [100] by enabling, in principle, extremely large-scale simulations. Photonic simulation has a long history and has led to the discovery of a variety of phenomena challenging to realize in conventional condensed-matter systems, such as topological phase transitions [94, 82, 27, 81, 76] and non-Hermitian exceptional points [31, 53, 126], which in turn has led to new photonic devices with applications far beyond basic physical science [6, 54].

Although telecommunication technologies routinely utilize the high-bandwidth inherent to optics, harnessing the frequency parallelism of light for large-scale analog simulation has largely remained unexplored. One promising approach is to im-

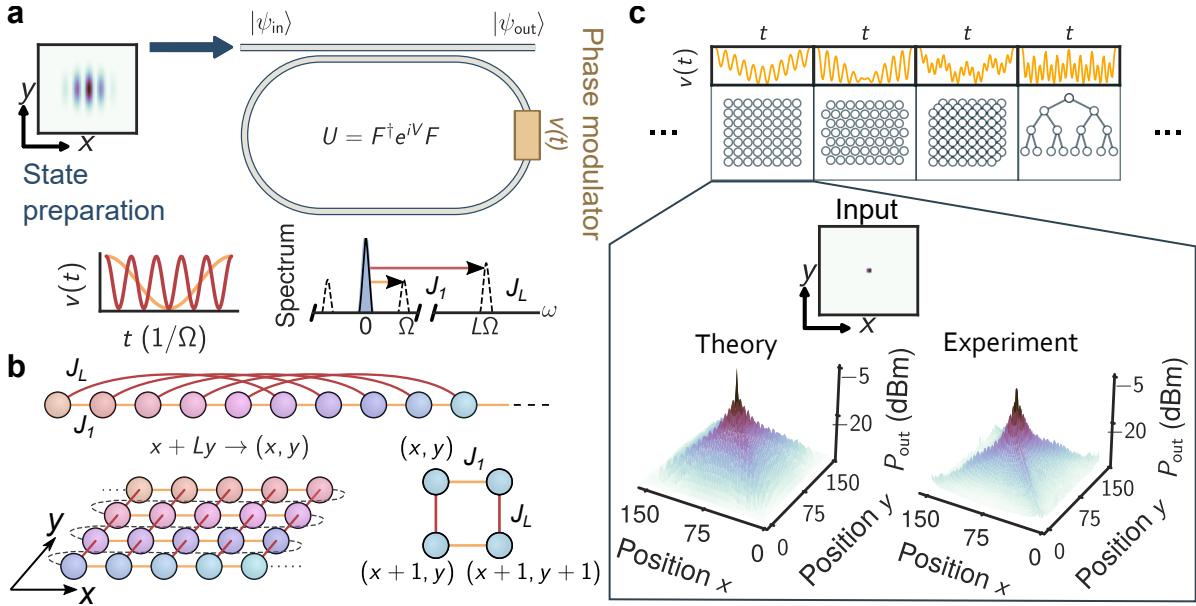


Figure 3.1: Simulations of large-scale bosonic transport with programmable photonic simulator. (a) Dynamic modulation of a fiber ring resonator couples frequency components of the intra-cavity field in each round-trip, represented in the basis of the frequency modes as $U = F^\dagger e^{iV} F$. Here, $V \propto \text{diag}(\vec{v})$ is the diagonal voltage operator defining the modulation signal $\vec{v} = (v(t_1), v(t_2), \dots, v(1/\Omega))$, and F is the discrete Fourier transform. The components of the voltage modulation define the coupling (bottom). By modulating at multiples of the mode spacing (Ω), we only couple long lived modes of the cavity, allowing for injected signals to propagate in frequency for many multiples of the round-trip time. (b) Engineered long-range coupling maps the one-dimensional spectrum to $L \times L$ two-dimensional lattice with twisted boundary conditions. As L grows large, the lattice approaches a smooth 2D plane. (c) A set of voltage signals defining lattices in 2 and 3 dimensions (top; see main text and Figure 3.12 for details), and the response to a single frequency drive (single-site) for a twisted 2D square lattice with over 20,000 lattice sites compared with tight-binding simulations of a 2D square lattice (bottom).

plement synthetic frequency dimensions [113, 95, 28, 56, 121, 29], in which optical frequency modes are mapped to lattice sites to perform bosonic analog simulations. Simulators using synthetic frequency dimensions have been shown to be versatile, implementing synthetic electric and magnetic fields [10, 75, 33], non-Hermitian coupling [120, 121], and nonlinear interactions [33, 118]. So far, however, these demonstrations have been confined to small lattice sizes, with limited programmability and restricted initial conditions.

Here, we demonstrate a frequency-mode-based platform that can simulate transport of arbitrary excitations in planar and non-planar optical lattices with up to 100,000 sites – orders of magnitude greater than achieved previously in photonic simulators with programmable geometry [132]. By pursuing a dense spectrum with MHz mode spacing, we leverage developments in both optical frequency combs and high-frequency optoelectronics to manipulate and probe a large number of optical frequency modes in a ring cavity. Additionally, these technologies enable a wide-bandwidth measurement scheme and arbitrary encoding of both amplitude and phase of input states, allowing fine resolution and control of bosonic transport. We leverage the scale and programmability of our simulator with three key demonstrations. First, our simulator can read out momentum-space features with high resolution, enabling us to measure band structures of various multi-dimensional lattices. Second, we observe both the real- and momentum-space signatures of time-reversal-symmetry breaking due to an effective gauge field in a two-dimension triangular lattice, an important step towards the realization of the photonic quantum valley Hall effect [47, 83, 62]. Third, we observe hierarchical transport across several orders of magnitude of length scales within a non-planar tree-like network, a toy model for quantum gravity and p -adic AdS/CFT correspondence [46, 9].

The class of Hamiltonians that our system is able to simulate is

$$H = \sum_{i < j} J_{i-j} a_i^\dagger a_j + \text{H.c.} \quad (3.1)$$

Hamiltonians in this class describe non-interacting bosons on translationally invariant lattices. a_l^\dagger and a_l are, respectively, the bosonic creation and annihilation operators for the l th lattice site. The lattice geometries are defined by the complex tunneling rates $J_k = |J_k|e^{i\phi_k}$, which encode the translationally invariant coupling of sites a distance k apart with amplitude $|J_k|$ and phase ϕ_k . The main goal of the photonic

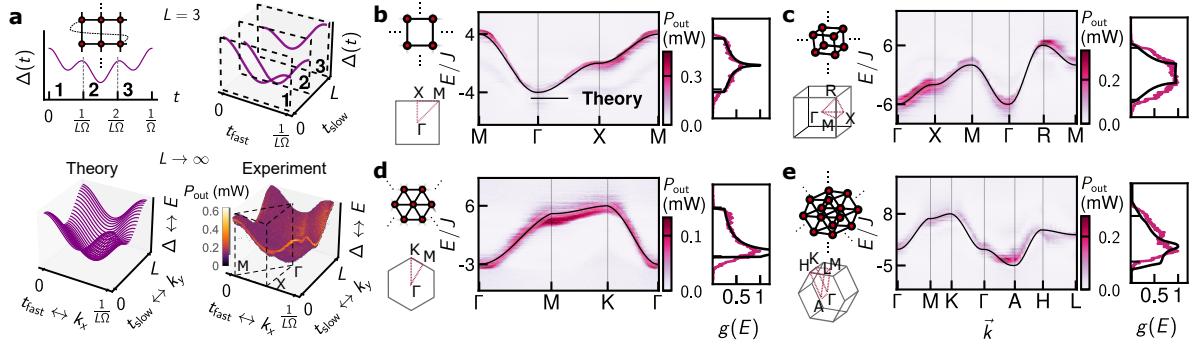


Figure 3.2: Optical band-structure measurements of 2D and 3D lattices. (a) (Top) The band structure for a twisted 2D square lattice is measured from the time-domain response of the cavity to scanning single-frequency injection as a function of the detuning Δ [28], here demonstrated using a pedagogical example with linear lattice size $L = 3$. This time-trace output is sliced up into chunks of length L , allowing the reconstruction of a full 2D band structure measured in a single shot (see Figure 3.14). (Bottom) Reconstructed band structure for a 2D square lattice with large L , comparing theoretical (left) with experimental results (right) for $L = 100$. As L grows large, the effect of the twisted boundary condition in the band structure becomes negligible, and the measured band structure approximates that of a regular 2D square lattice. (b) Data of the full band structure plotted along slices that connect special points of the Brillouin zone, compared with theoretical results for a true 2D square lattice (black). These points, highlighted in the bottom-left, denote locations in momentum space with high symmetry. The density of states $g(E)$ is directly measured by summing the time-domain response (right). Band structures and density of states for (c) 2D triangular lattice ($L = 100$), (d) 3D square lattice ($L = 28$), and (e) 3D hexagonal lattice ($L = 28$).

simulator we present in this work is to be able study the transport of a variety of single-particle excitations in any Hamiltonian in the class defined by Eq. (1), where the complex parameters J_k can be programmed arbitrarily – this allows us to study a diversity of different lattices, including ones that are multidimensional.

3.2 Results

In the synthetic-frequency-dimensions approach [132] that we adopt, the lattice operators a_l, a_l^\dagger are associated with the l th frequency mode of an optical cavity, spaced apart by Ω (the free spectral range of the cavity). The tunneling rates $(J_k)_{k=1,2,3,\dots}$ are physically realized using a phase modulator within the optical cavity (Fig. 3.1a) – intuitively, the modulator creates optical sidebands at the frequencies contained in the modulation signal $v(t)$ and, by setting these sidebands at harmonics of Ω , these sidebands cause coupling between cavity modes. Additionally, by setting the amplitudes and phases of frequency components in $v(t)$ appropriately, different lists of tunneling amplitudes ($|J_k|$) and relative phases (ϕ_k) can be programmed (Fig. 3.1a), which in turn realize different lattice geometries (Fig. 3.1c). Our simulator includes a gain/loss balance term in the Hamiltonian, but this is kept close to zero. Additionally, phase modulation at Ω produces higher-order sidebands, but these can be suppressed to negligible amplitudes by appropriately choosing the modulation voltage (see Methods).

Although the Hamiltonian in Eq. (1) nominally describes a 1D lattice, we can implement effective higher-dimensional lattices by suitably programming the couplings (J_k) to reflect the local geometry of a target higher-dimensional lattice. For example, an effective $L \times L$ 2D square lattice can be realized in a non-local 1D lattice by coupling nearest-neighbors and L th nearest-neighbors, i.e., $(J_k)_{k=1,2,3,\dots} = (J_1, 0, \dots, 0, J_L, 0, \dots)$ (Fig. 3.1b). This produces a 2D lattice with a twisted boundary condition [113, 135], and only approximates a true 2D lattice once L is made to be very large. Here, any local excitation with finite lifetime will become insensitive to the boundary. Thus, by pursuing a large number of modes, we can realize effective lattices in higher dimensions that approximate the true physics. The vanishing

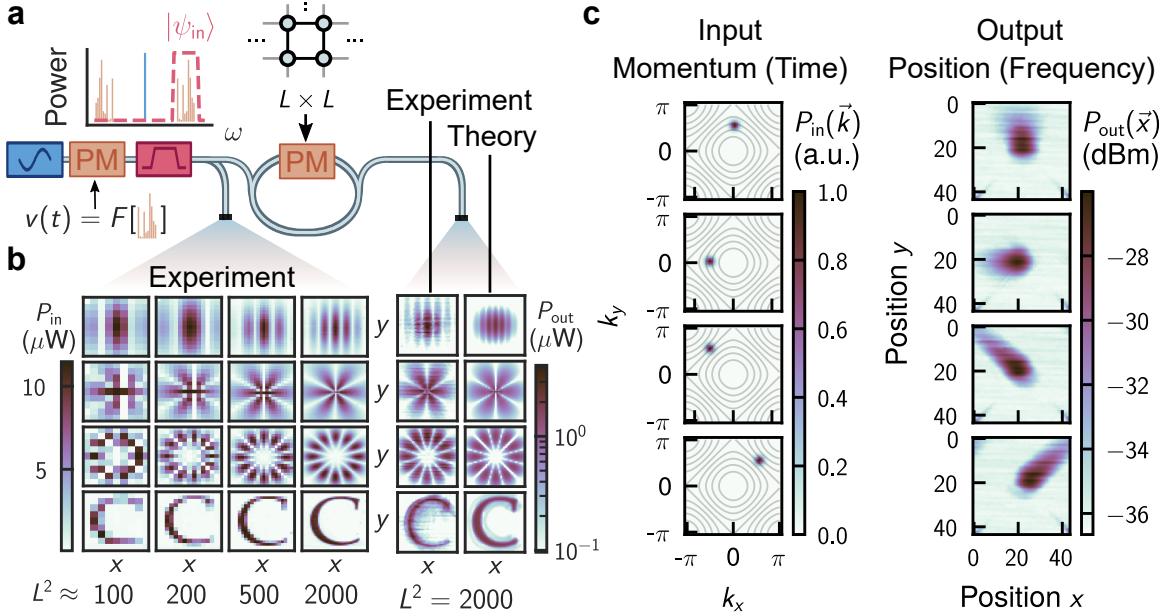


Figure 3.3: Input state preparation. (a) Scheme for preparing arbitrary input states: a single frequency tone is modulated with a signal encoding both amplitude and phase of a given state, producing symmetric sidebands (orange spectrum). The initial tone and the unwanted sideband is then rejected with a bandpass filter (red envelope), leaving only the positive sidebands which are sent into the cavity. (b) Experimental measurements of input states for increasing number of modes programmed in the input signal for: a standing wavepacket, an angular wave enveloped with a Gaussian centered at zero, higher angular state enveloped with an offset radial Gaussian, and the Cornell University logo. The steady state outputs of these states for a 2D $L \times L$ square lattice are shown to the right, along with comparison with theory. (c) We excite momentum eigenstates of a 2D square lattice with momenta in various directions enveloped with a Gaussian. Left shows the representation of the input state in momentum space $\vec{k} = (k_x, k_y)$, right shows the experimental steady state in position space $\vec{x} = (x, y)$. Here, local momentum eigenstates are continuously excited at the center, and propagate with a well-defined momentum before decaying.

effects of the twisted boundary condition can be seen in the steady-state response to a single-site excitation in the comparison shown in Fig. 3.1c between our experiment and simulations of a true 2D tight-binding lattice with hard boundary conditions.

3.2.1 Measurements of band structure

The frequency-multiplexed platform has a convenient encoding of reciprocal space for lattice systems. In the mapping from lattice sites to frequency modes, time maps to momentum [28]. Therefore, since the Fourier components of the modulation signal define the connectivity, the modulation signal in the time domain defines the band structure. For a 1D lattice, this correspondence is exact: the modulation signal $v(t) = -V_0 \cos(\Omega t)$ couples nearest-neighbor modes, implementing a 1D tight-binding chain with band structure $E(k) = -J \cos(ka)$. Here the lattice spacing a is identified with the mode spacing Ω , and momentum k with time t . More generally, the action of phase modulation on the frequency modes can be expressed as a unitary operator $U = F^\dagger e^{iV} F$, where F is the discrete Fourier transform, and $V \propto \text{diag}(\vec{v})$ is a diagonal matrix whose values are proportional to the voltage signal $\vec{v} = (v(t_1), v(t_2), \dots, v(1/\Omega))$. In our simulator, the operator U implements the time evolution defined by the Hamiltonian in Eq. (1). Thus, the modulation signal $v(t)$ defines the time evolution in a diagonal basis, and therefore encodes energy eigenvalues of the lattice, i.e., the band structure. As a consequence, this permits us to encode arbitrary lattices that have with a single-band band structure. Additionally, it provides us direct access to momentum-space lattice measurements [28].

To extend the momentum-to-time analogy to 2D and 3D, we require large enough number of modes to eliminate the finite-size effects from the twisted boundary condition. For the above example of an effective 2D lattice, slices of the band structure along the slow axis (i.e., the axis corresponding to transport along nearest-neighbors) suffer from an asymmetry near the boundaries of the Brillouin zone (Fig. 3.2a), prominent for small L . This is due to the twisted boundary conditions, which makes the two directions no longer independent, since L hops along Ω will reach the same posi-

tion as a single hop along $L\Omega$. Concretely, the asymmetry can be seen by comparing the two-tone signal we use to generate a 2D lattice, $v(t) = -2V_0 \cos(\Omega t) - 2V_0 \cos(L\Omega t)$, and a true 2D tight-binding lattice with nearest-neighbor hopping, which has a band structure $E(\vec{k}) = -2J \cos(k_x a) - 2J \cos(k_y a)$. The latter has two independent reciprocal lattice vectors, k_x and k_y . For $L \gg 1$ however, we can rely on a separation of timescales and treat $\Omega' = L\Omega$ as an effective independent degree of freedom. This approach can be extended to higher-dimensional lattices, e.g. for a 3D square lattice, Ω , $L\Omega$, and $L^2\Omega$ are the independent degrees of freedom.

Fig. 3.2a outlines how we extend the methods introduced in Ref. [28] to measure the band structure of 2D lattice in a single-shot, then slice up the measured band structure in periods of $T_{\text{fast}} = 1/L\Omega$ to reconstruct the 2D full band structure. See Figures 3.14 and 3.15 in Section 3.4.3 for full details on this reconstruction. As $L \rightarrow \infty$, the band structure of our effective 2D square lattice approaches that of a regular 2D square lattice, as seen when comparing Figs. 2a and b. Slices through high symmetry points of the full band structure are shown in Fig. 3.2c-f for a 2D square, 2D triangular, a 3D simple cubic, and a 3D hexagonal lattice, along with the respective density of states for each. Theoretical curves for ordinary tight-binding lattices are shown in black.

3.2.2 Preparing arbitrary input states

High-bandwidth telecommunications optoelectronics enable the study of transport in our platform for arbitrary input states. Our scheme is enabled by 12-GHz electro-optic modulation, summarized schematically in Fig. 3.3a. This technique allows us to specify the amplitude and phases of input excitations for up to around 4000 lat-

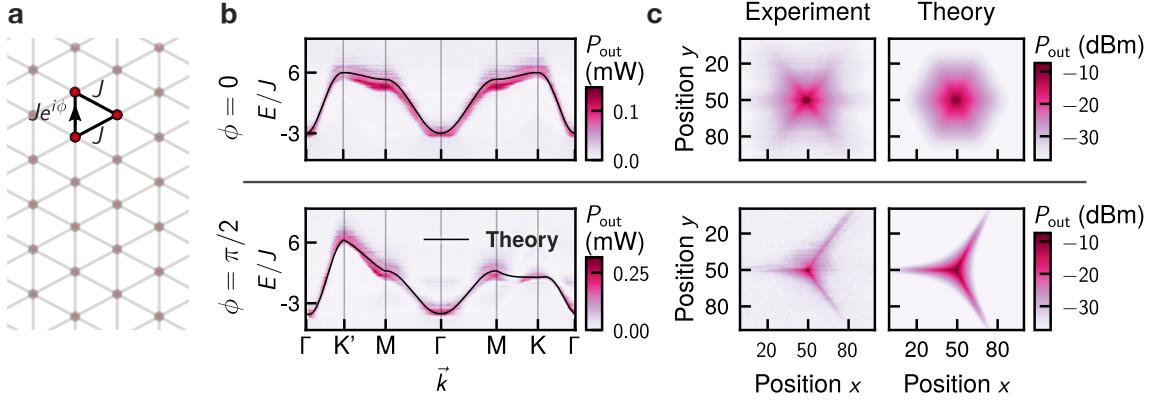


Figure 3.4: Time-reversal-symmetry breaking in a 2D triangular lattice due to an effective gauge field. (a) Complex hopping terms induces a nonzero local magnetic flux within a plaquette of a triangular lattice. Here, a relative phase is added to one of three directions incident on a given lattice site. **(b)** The introduction of the magnetic field breaks time reversal symmetry, as can be seen in the asymmetry of the K and K' points in the band structure after performing Pierelis substitution (bottom). **(c)** Measured steady state spectral response due to a single-site injection under the influence of the synthetic local magnetic field. The presence of the synthetic field leads to a departure from 6-fold symmetry to 3-fold symmetry in the transport. Experimental data (left) is compared with simulations (right).

tice sites, limited primarily by a bandpass filter (see Methods). Figure 3.3b shows experimental measurements of various input states, including standing wavepacket eigenstates, angular wavepackets and a Cornell ‘C’. The right column displays their respective steady-state response. Here, we are continuously exciting local states and observing their steady-state dynamics in the presence of loss. The full control over both amplitude and phase enable us to excite states with net momentum. Fig. 3.3c shows the steady state response of momentum eigenstates of a 2D square lattice enveloped with a Gaussian for a discrete set of nonzero input momenta. Here, we can directly observe locally excited momentum states propagate in different directions. The momentum of each input state is labeled by its respective momentum distribution within the Brillouin zone, shown in the left column of Fig. 3.3c.

By programming the phases and detunings of the coupling Hamiltonian

(Eq. (3.1)), we implemented synthetic magnetic and electric fields respectively (see Fig. 3.23 in Section 3.5 for measurements for synthetic electric fields) [75, 30, 78, 19, 10, 134, 106, 105, 88, 133, 102], as well as non-Hermitian models (see Figs. 3.24 and 3.25 in Section 3.5 for realizations of the Hatano-Nelson model in one and two dimensions). Figure 3.4 shows the effect of a synthetic gauge field applied to a 2D triangular lattice, giving rise to a global zero magnetic field, but nonzero local magnetic field. Adding a relative phase along nearest neighbor hoppings results in an accumulated phase of either $e^{i\phi}$, or $e^{-i\phi}$, indicating a local nonzero magnetic flux going around each plaquette. The sign of the accumulated phase alternates between neighboring plaquettes, making the total magnetic flux through the lattice vanish globally. Shown in Fig. 3.4b, the addition of this field breaks time-reversal symmetry, which for the triangular and honeycomb lattices, maps the K to K' points [47]. This results in a reduction of a 6-fold symmetry to a 3-fold symmetry in the transport of injected light, where propagation of light is prohibited in certain directions, as shown in the heterodyne measurements of the steady-state density (Fig. 3.4). This time-reversal-symmetry breaking with local non-zero fields is one key ingredient in observations of the quantum-valley Hall effect seen in honeycomb lattices [47, 83, 62].

3.2.3 Simulating a lattice with more than 100k lattice sites

In addition to lattices found in traditional condensed-matter systems, our photonic simulator is capable of simulating systems not realizable in crystalline materials. Systems with non-planar connectivity are particularly interesting given their realization in solid state systems are impractical, yet contain rich physics. Periwal et al. [100] recently experimentally demonstrated a simulation of a graph with exotic long-range connectivity given by

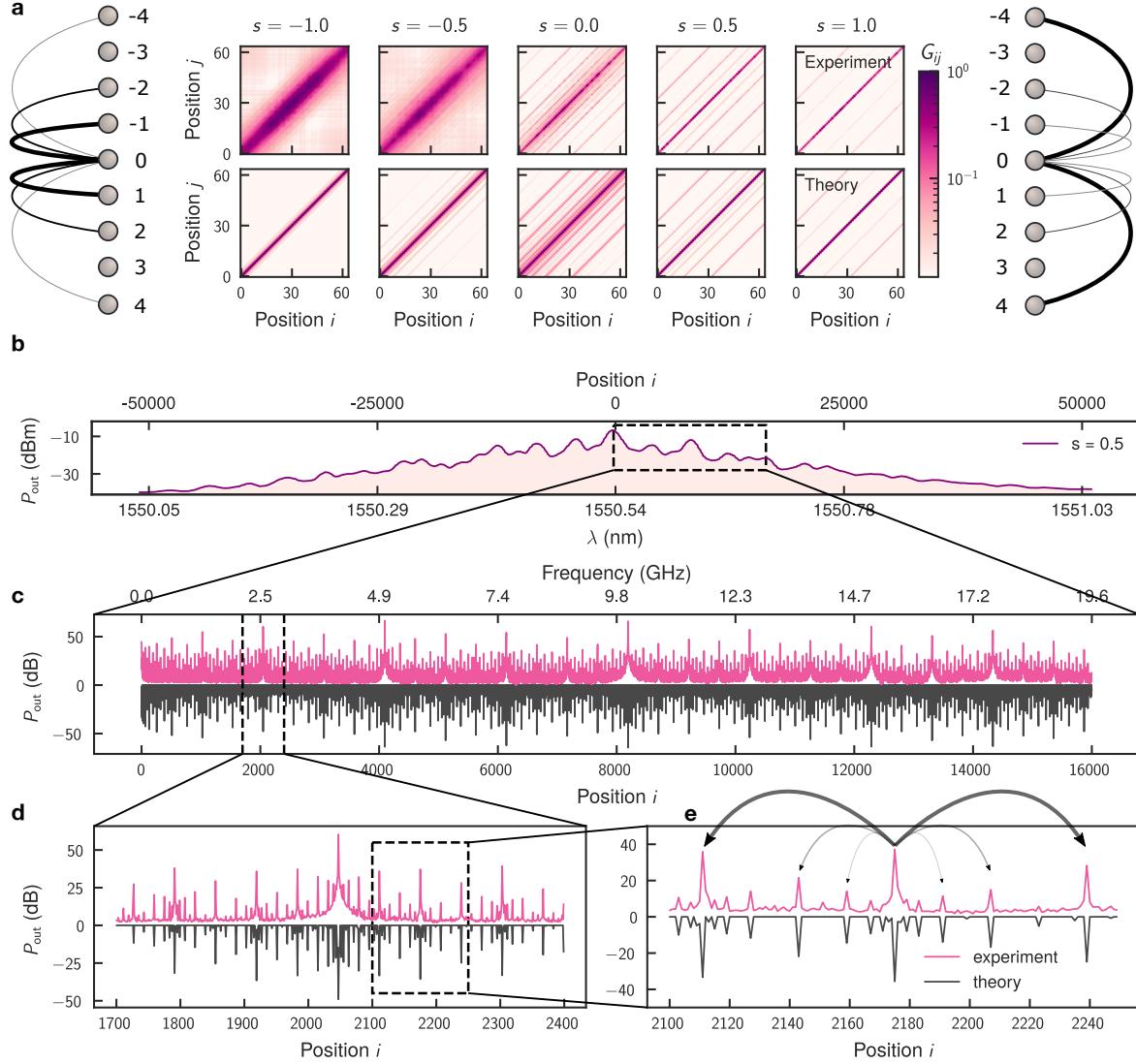


Figure 3.5: Simulations of bosonic transport in a tree-like geometry with a graph comprising over 100,000 sites. (a) Non-equilibrium correlation measurements for a 1D chain with non-local connectivity (see Eq. (3.2)). As s is tuned from -1 (left) to $+1$ (right), the correlations transition from locally decaying to tree-like [100]. The lattice cartoons on the left and right schematically show the coupling form for a single lattice site (position 0). (b) Optical spectrum measurement for response to a single site injection with $s = 0.5$. (c-e) RF spectrum measurement showing the lattice occupation with single-site resolution of the zoom-in in the full optical spectrum, comparing with simulations for windows of 20 GHz (c), 500 MHz (d), and 100 MHz (e).

$$J_{i-j} \propto \begin{cases} |i-j|^s & |i-j| = 2^n, n \in \{0, 1, 2, \dots\} \\ 0 & \text{otherwise.} \end{cases} \quad (3.2)$$

This describes a system that can be continuously changed, using the parameter s , from an Archimedean-geometry regime in which correlations between sites decay with distance $|i - j|$, to a non-Archimedean-geometry regime in which the correlations between sites have a tree-like structure. The hierarchical geometry of this tree-like system is a toy model for p -adic AdS/CFT correspondence [9] studied previously using atomic ensembles in an optical lattice with 16 sites [100], shown here for the first time in a photonic system. We experimentally show this transition in Fig. 3.5a in the measurements of correlations of the lattice as s is tuned (see Methods section for details on the correlation measurements and Fig. 3.22 in Section 3.5 for another example). Near the transition, at $s = 0.5$, the lattice exhibits both strong local and nonlocal connectivity, resulting in dense yet extremely large lattices, shown in both optical spectrum measurements in Fig. 3.5b and RF spectrum measurements in Fig. 3.5c.

While some graphs, such as the tree-like example depicted in Fig. 3.5, result in occupations that span 100,000 or more lattice sites, quantifying the absolute size of our simulator requires some nuance. On one hand, based on the dispersion and bandwidth of the elements inside the cavity, we believe the lattices we simulate span several THz, corresponding to millions of lattice sites. On the other hand, as in real material systems, local excitations in lattices that have only short- or medium-range connectivity will typically not be able to propagate to very distant sites before their amplitude decays below the noise floor of the detector. For example, in the experiments we performed with 2D lattices, the steady-state response was detectable at most $\sim 10^4$ lattice sites in the vicinity of the injected wavepacket (see Fig. 3.1c and

Fig. 3.21 in Section 3.5).

3.3 Discussion

The demonstrations performed in this paper cover only a small fraction of the bosonic physics that can be simulated with frequency-domain coupling of photonic modes. Simple modifications to the presented experimental setup, such as dispersion compensation, reducing total intra-cavity loss, and reducing the input power far below the gain-saturation power, should substantially increase the number of accessible lattice sites. Additionally, stabilizing the cavity phase with respect to the phase of the input state would remove decoherence effects limiting the propagation of our input excitations. With these upgrades, observing dynamics on lattices spanning millions of sites (or more) would become feasible. By adding multiple spatial modes [127] or bi-directional propagation [27], simulations of topological phenomena found in higher-dimensional gapped multi-band systems may be realized. With coupled cavities, defects and hard lattice edges [29] may be implemented; in our simulator, this would enable the study of propagating edge modes in high-dimensional lattices. By varying the intra-cavity phase modulation over multiple cavity periods, time-dependent lattices may be realized, which would allow the study of new non-equilibrium phases [8], and the implementation of very wide convolutional optical neural networks [35]. Similarly, the use of stroboscopic modulation, as opposed to continuous wave modulation, would enable timed measurements of transient dynamics, such as observing band evolutions [131] and time-resolved spectral measurements. In unmodulated cavities, Kerr nonlinearities give rise to locked combs defined by dissipative cavity solitons [70]. In the frequency domain, the Kerr effect produces highly non-local, four-mode interactions [118, 33]. Since Kerr interactions may be

programmed by modifying intra-cavity dispersion and spectral loss, and by introducing additional mode families [127], it should be possible to realize both new types of intricately tailored Kerr frequency combs, as well as simulations of the statistical mechanics of graphs with higher-order interactionse [8], which should allow the observation of emergent multi-stable states and abrupt synchronization (mode-locking) transitions.

Photonic simulators have over the past decade been established as robust platforms for exploring condensed-matter phenomena, including some that have been inaccessible in material systems. In this paper we have demonstrated a large-scale, programmable photonic simulator using synthetic frequency dimensions. We used our simulator to study several models with a variety of different geometries, including a lattice with tree-like connectivity that has, to our knowledge, not previously been realized outside of cold-atom experiments and that would be impractical to realize at scale in most simulator platforms without the use of synthetic dimensions due to the model's highly non-local interactions. Looking to the future, our simulator could be extended to even larger sizes through dispersion and loss engineering, modified to support the study of topological phenomena, and augmented with a nonlinearity that induced higher-order interactions between lattice sites. With extensions to fully utilize the many THz of bandwidth that is in principle available in optics, programmable synthetic-frequency-dimension photonic simulators may soon explore high-dimensional nonlinear physics, both near and far from equilibrium. At all scales, advances in this platform will benefit the development of tailored light sources and optical signal processors. However, it is in the THz-spanning ultra-large-scale regime that photonic simulators seem most compelling to us as analog simulators, with prospects to explore – and discover – entirely new and unexpected physical phenomena.

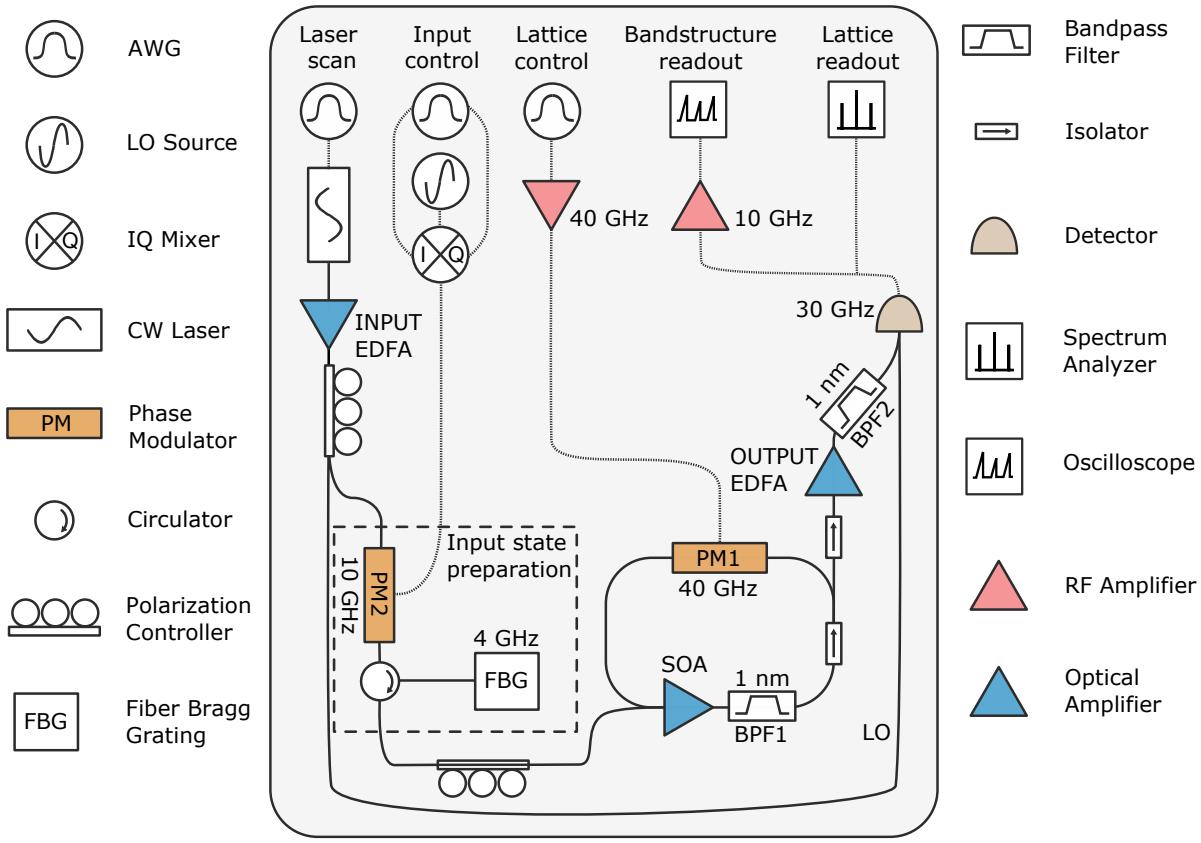


Figure 3.6: Schematic of the experimental setup.

3.4 Experimental setup

The experimental setup, schematically shown in Fig. 3.6, consists of a long fiber ring with a free spectral range (FSR) $\Omega = 1.226045$ MHz, modulated with a 40-GHz phase modulator (LN27S-FC from Thorlabs) to produce the lattices in the main text and below. The losses in the cavity are compensated by an semiconductor optical amplifier (SOA) from Thorlabs (SOA1117S). A 1 nm optical filter (OZ Optics BTF-11-11-1525/1565-9) is placed in the cavity to reduce ASE generated by the SOA, and an isolator (Thorlabs IO-H-1550APC) is placed to reduce parasitic resonances particularly at low cavity power.

In addition to the cavity, we implement arbitrary input states by a filtering mod-

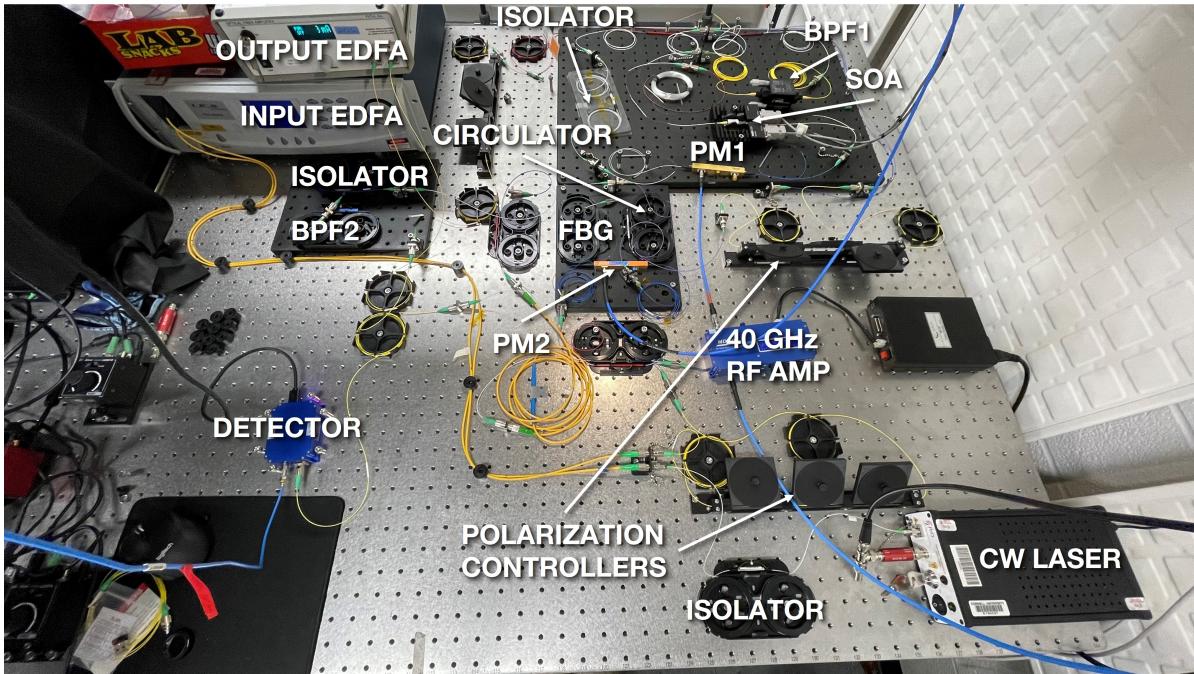


Figure 3.7: Photograph of the experimental setup.

ulations of an injection. The injection laser (RIO3335-3-00-1-BZ7), centered at 1550.57 nm, is first amplified before being split with a 50:50 coupler. One arm serves as the local oscillator (LO) for heterodyne detection, and the other arm is sent through a phase modulator (EOSpace) to prepare the input state. The polarization is adjusted so that it is aligned with the crystal axis of the phase modulator. After passing through the modulator, the input state is encoded in sidebands centered 12 GHz away from the injection. A fiber Bragg grating (FBG) mirror reflects a 4 GHz portion of the spectrum centered 12 GHz away from the LO, while the rest, containing the other sideband and the LO frequency, is dumped (see Fig. 3A). We place a circulator in between the phase modulator and the FBG to collect the reflected light. While this prescription results in 20 dB of insertion loss, we find 30 dB of isolation between the filtered sideband and the LO resulting in a clean single-side band modulation. The modulation of both the input state phase modulator and the cavity phase modulator are controlled by a single AWG (Keysight M8195A). The polarization of the prepared input state pro-

grammed by a dedicated AWG is set to align with the crystal axis of the intra-cavity EOM.

Finally, the output of the cavity is combined with the LO before being detected by a 30-GHz photodetector (Optilab PD-30-M-K-DC), resulting in a 12-GHz heterodyne detection of the output. One component of resultant RF signal is amplified before sent to an oscilloscope (Tektronix DSA72004) for band structure spectroscopy in the time-domain. The other component of the RF signal is sent to the spectrum analyzer (Tektronix RSA5126A) to perform direct readout of the lattice over 26 GHz. For band-structure measurements, the heterodyne arm is turned off. The spectrum measured in Fig. 5B was measured with an optical spectrum analyzer (Ando AQ6317).

3.4.1 Setup characterization

Here we outline characterizations of the cavity, particularly characterizations of the cavity gain and losses, coherence time, and the free spectral range (FSR).

The intracavity SOA compensates for roundtrip losses in the cavity. However, the SOA contributes significantly to noise by the unwanted production of amplified spontaneous emission (ASE). In order to reduce ASE, we placed a filter in the cavity, and minimized roundtrip losses to reduce the operating point of the SOA. The cavity losses were reduced to 5 dB, 4 dB of which originate from the insertion loss of the cavity EOM. In addition to limiting the noise, reducing the operating point of the SOA has the added benefit of reducing the contribution of the ASE to the gain saturation. The left panel of Fig. 3.8 shows the effect of the gain saturation on the dynamics of a nearest-neighbor optical lattice. For a large ratio of input power to saturating power, the transport of the optical power along the 1D chain is limited. From simulations, we

find that a ratio of $P_{\text{seed}}/P_{\text{sat}} < 1/100$ is enough to reach thousands of modes above the noise floor of our setup. The right side of Fig. 3.8 shows the experimental measured value of gain saturation at about 4 dBm, thus for all of the measurements in the main text, we had an input power of below -20 dBm going into the cavity.

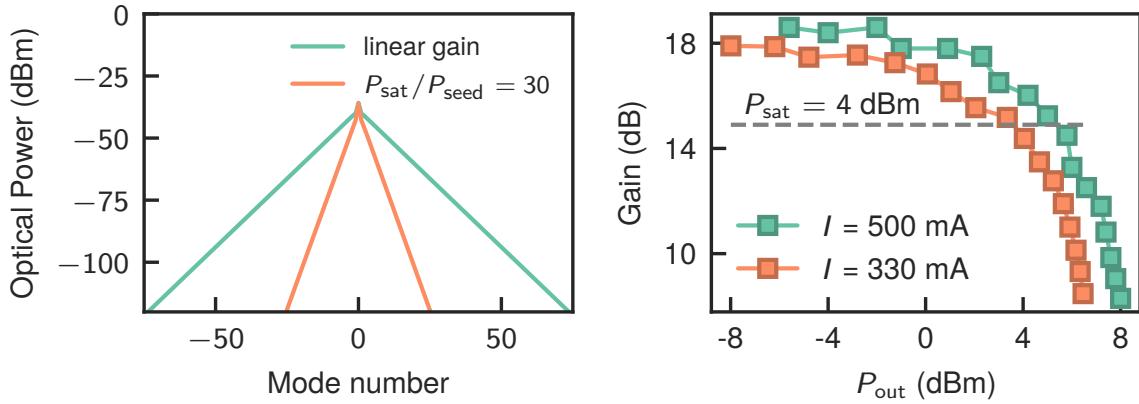


Figure 3.8: **Modeling and characterization of gain saturation.** Left: simulations of cavity dynamics for linear gain (green curve) and saturating gain (orange curve). Right: Experimental characterization of semiconductor amplifier at various levels of pump current.

Knowledge of the cavity FSR is needed to ~ 10 Hz to in order sustain the coherence over many roundtrips needed to instantiate large lattices. This was measured via different methods, differing in course-grained vs fine-grained characterization. At the first step, we measured the FSR by exciting the cavity far above the lasing threshold and measuring the mode excitations. This procedure gives us the FSR to within 10 kHz. Next, we placed the cavity just below the lasing threshold and performed spectral measurements due to a single-site injection, and maximized the transport observed over the modes as we varied the modulation frequency. Shown in Fig. 3.9 are the responses in a 1D chain as a function of modulation frequency at cavity threshold. To further increase the sensitivity of the measurement, the same procedure was applied at larger integer multiple harmonics. These are plotted in Fig. 3.9B.

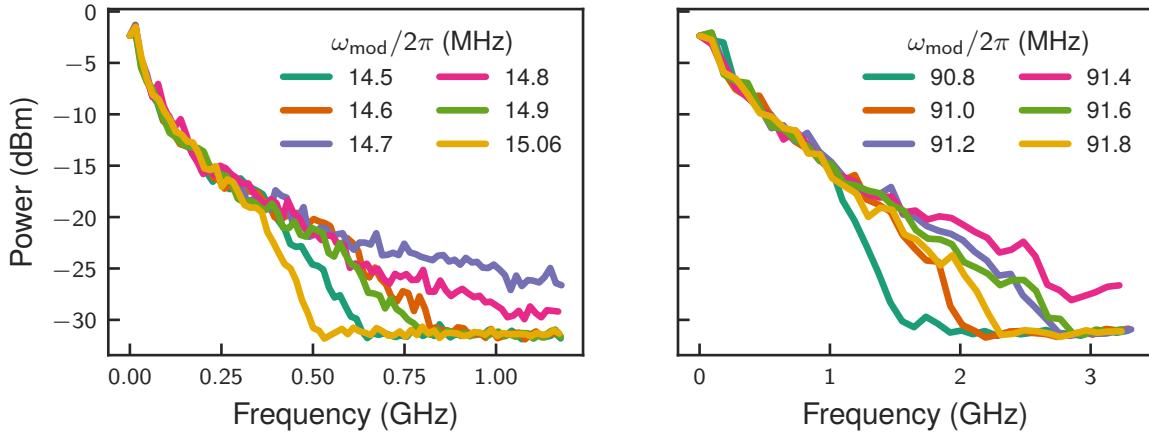


Figure 3.9: Measurement of the cavity resonance by maximizing transport response at lasing threshold. The cavity was modulated around a window of $\omega_{\text{mod}} \approx 12 \times \Omega$ (Left), and $\omega_{\text{mod}} \approx 75 \times \Omega$ (Right).

To measure the FSR of our cavity down to 10 Hz, we modulated the EOM with multiple tones, which increased the sensitivity due to interferences between different paths taken over multiple roundtrips of the cavity. For example, a modulation consisting of two tones at Ω and $10 \times \Omega$ will maximize transport if 10 hoppings along the nearest-neighbor modulation produces is coherent with a single hop along $10 \times \Omega$ [135]. Figure 3.10 shows the transport response for a 144×144 2D lattice instantiated by two tone modulations at frequencies ω_{mod} and $144 \times \omega_{\text{mod}}$. When detuned from the FSR (left and right panels of Fig. 3.10), we see the injected light reaching out then oscillating back in. Closer to resonance (middle panel of Fig. 3.10), the light propagates symmetrically outward. This effect is related to Bloch oscillations (see Section 3.5), expected to occur in modulated cavity systems where the modulation is detuned [134]. Suppressing these oscillations allowed us to find the FSR to 5 decimal places, down to 10 Hz.

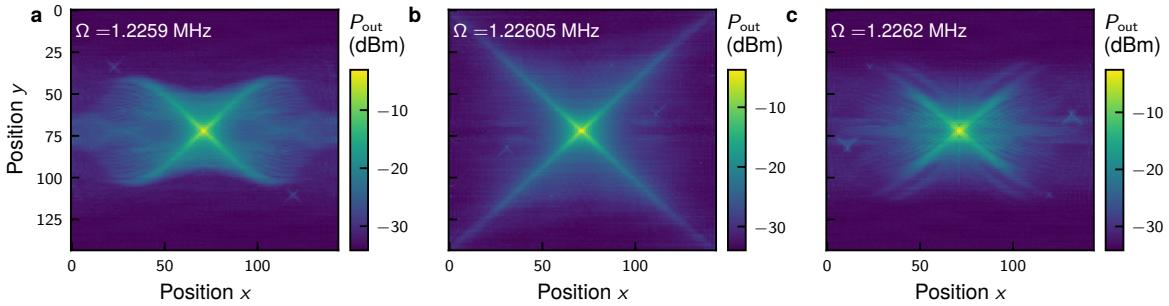


Figure 3.10: Response to single site injection in a 2D square lattice with 20,000 sites as a function of detuning between the modulation frequency and the cavity free spectral range. By optimizing the detuning for transport, the cavity FSR was measured to be 1.22605 MHz. Panels A and B show the presence of Bloch oscillations (see text)

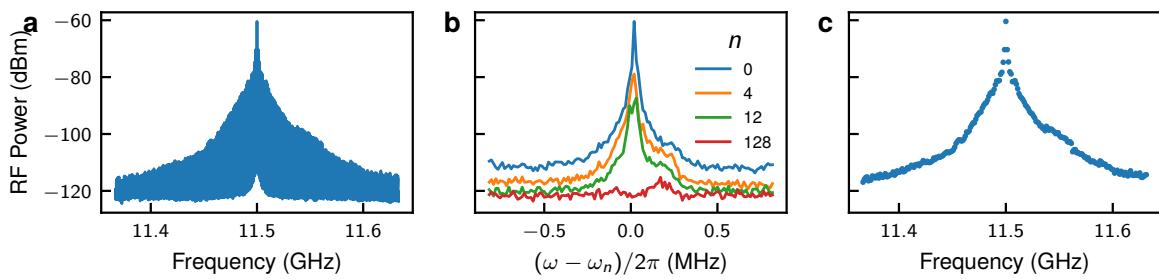


Figure 3.11: Processing raw spectrum data. (a) RF spectral cavity response due to a single site injection in a 1D nearest-neighbor chain. (b) closeup of mode excitation for various lattice sites overlayed on the same scale. n is the distance away from the injection measured in number of lattice sites. (c) Constructed lattice plot from peaks of lattice site excitation.

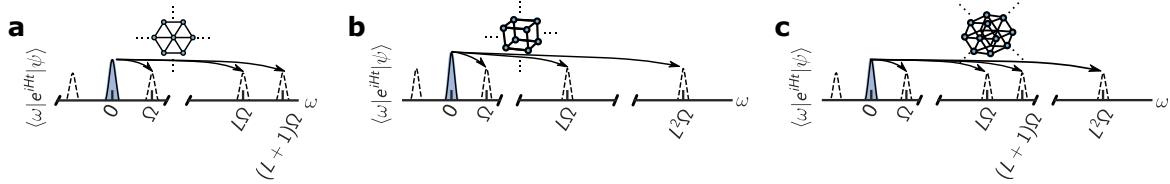


Figure 3.12: **Modulation schemes for the lattices presented in Figure 2 of the main text** (a) 2D triangular lattice, (b) 3D square lattice, (c) 3D triangular lattice. The modulation scheme for a 2D square lattice is shown in Fig. 1 of the main text.

3.4.2 Real-space-occupation measurements

Here we provide the methods used in measurements of the steady state tomography in various lattices. The output of the cavity is sent to an RF detector, whose current measures the optical power. This RF signal is then sent to an RF spectrum analyzer which integrates the signal over 100s of round trips. The spectrum analyzer measures RF power, related to the optical power via $P_{\text{RF}} = \eta^2 P_{\text{opt}}^2 \times 50\Omega$, where η is the responsitivity of the photodetector (measured to be about 0.77 A/W). All spectral plots are plotted in terms of optical power leaving the cavity before the EDFA, with the exception of Fig. 3.11, which plots the raw RF power spectrum.

To map the power spectrum to the occupation at a particular lattice site, the peak of the power around a neighborhood of the expected frequency of the cavity mode was picked. Figure 3.11A shows the raw heterodyne response data for a 1D lattice due to a single-site injection. Figure 3.11B shows several cavity modes imposed on the same axis for a neighborhood of 1 MHz just below threshold. Here, the linewidth of the cavity is about 0.5 MHz. From these windows, the peaks were picked to construct the final lattice response, displayed in the right pane of Fig. 3.11.

To generate lattice plots in two dimensions and higher, we employed the above prescription to retrieve the peaks from a 1D chain, then reshaped the data appropri-

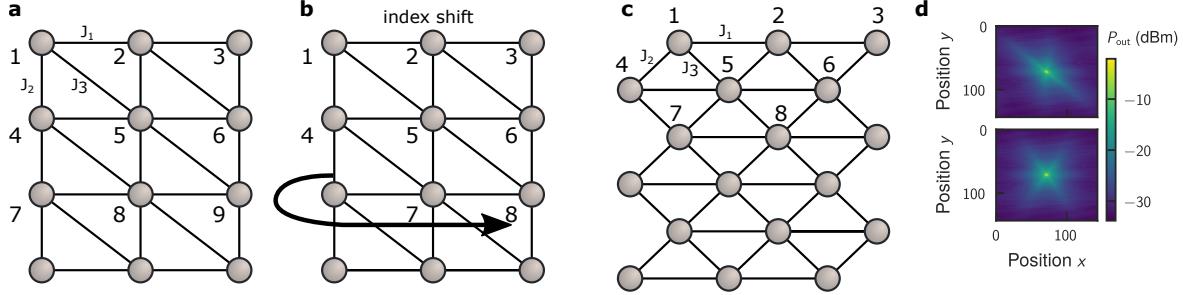


Figure 3.13: Mapping the frequency spectrum to a 2D triangular lattice geometry. First, the frequency modes are mapped to points in (a), similar to a 2D square lattice, though with an extra hopping. Beginning with every other row, the indices are shifted over by one in order to preserve the local geometry of the triangular lattice (b). Finally, every other row is shifted by half a "pixel" in all spatial representations (c). (d) Compares the unshifted lattice (top) with the shifted (bottom).

ately depending on the type of lattice we were realizing. For example, if we measure the response in a 2D square lattice, where we modulate the cavity at Ω and $L\Omega$, we pick $N = L \times L$ peaks of the spectrum, and wrap them in a $L \times L$ matrix. The response of the electro-optic modulator to the voltage drive (characterized by V_π – the voltage needed to induce a phase shift of π) is generally frequency-dependent. Thus with signals composed of far separated tones, the modulation amplitudes of these tones had to be calibrated to realize uniform coupling. As an example, for a 144×144 2D lattice, the modulation amplitude for the $144 \times \Omega$ tone was 1.1 times larger than that for Ω . Given the small mode spacing, and the large bandwidth of our modulator (40 GHz), the dependence was quite weak.

For a triangular lattice, we modulate the intra-cavity EOM at three frequencies $\{\Omega, L\Omega, (L+1)\Omega\}$ (Fig. 3.12A). Following the methods employed in the 2D square lattice geometry, this produces a lattice depicted in Fig. 3.13A – a 2D square lattice with an extra diagonal connections. While the connectivity of this graph is exact with the connectivity of a triangular lattice, the six fold symmetry of this lattice is not captured when presented in regular euclidean space. Intuitively, this can be seen in Fig. 3.13C,

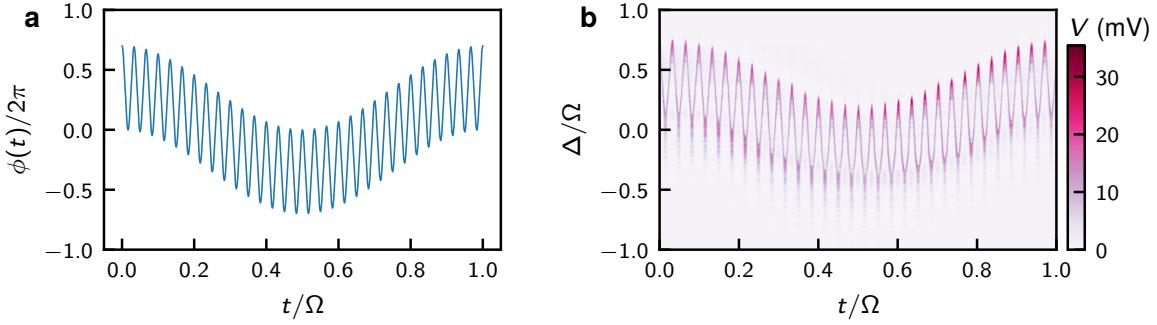


Figure 3.14: Bandstructure measurement of a 30×30 2D square lattice using methods introduced in [28]. (a) Modulation signal sent to the cavity EOM. (b) Experimental measurement of time-domain response of the cavity broken up in chunks of the roundtrip time $\tau = 1/\Omega$. The vertical axis on the experimental plots is the detuning Δ of the injection away from the cavity mode, normalized by the cavity spacing Ω .

where a traversal down along J_2 followed by a traversal along J_3 should result in an occupation that is directly “below” the original site geometrically. In Fig. 3.13 however, this procedure results in a shift along the horizontal direction. Thus, beginning with every even row of our lattice, we shift the indices over by one as depicted in Fig. 3.13B, resulting in the expectant traversals. Doing so preserves the local connectivity of the triangular lattice without physically altering the connectivity. Finally, when plotting these lattices on a 2D plane, we plot the values of the modes on a planar lattices that has the geometry depicted in Fig. 3.13C. The difference in these aesthetic changes are apparent in Fig. 3.13D.

Here, we discuss some of the limitations in the above procedure for realizing high-dimensional lattices. While the procedure is valid for local excitations, the approximation breaks down for either non-local excitations that are on the order of the lattice size L , or for large J/γ – the ratio of the coupling strength to the gain-loss balance. The latter issue could be resolved by increasing the lattice size up to the modulation bandwidth, or by decreasing J/γ . In either case, large lattices faithful to the approximated Hamiltonian can be realized. On the other hand, the above procedure, while

faithful to systems without boundary effects, prohibit the study of physics where boundary effects become important, as is the case for systems with nontrivial topological phases. While the setup can implement topological Hamiltonians (such as the Hatano-Nelson model, see Figs 3.24 and 3.25 in Section 3.5), observing edge states is not possible in this platform which contains no edges. To remedy this, one could couple a second cavity with a cavity mode spacing that is a integer multiple M of the main cavity. As demonstrated in [29], this will produce hybridized modes M sites apart that are off-resonant from the frequency comb, generating a defect that suppresses any coupling to that site modulated by the cavity spacing. To generalize this to two dimensions, one would modulate at Ω and $L\Omega = M\Omega$: Coupling nearest neighbors, and sites L apart, while blocking a site that is positioned next to a defect from crossing over the twisted boundary condition. Generalizing this to three dimensions would require a third cavity and so on. This modification would enable the study of a variety of systems with boundary effects in high dimensions, as long as the second cavity does not introduce much loss to the main cavity.

3.4.3 Band-structure measurements

To measure band structure, the output of the cavity is amplified and filtered with a 0.1-1 GHz-bandpass filter. The amplifier increases the signal to noise ratio needed for the single shot readout for the band structure measurements. Single shot read-out is needed due to the phase walk-off between the injection and the cavity. For details of the band structure measurement, we refer to [28], but we briefly summarize the procedure here, and outline our extensions for measurements of band structures in 2D and higher.

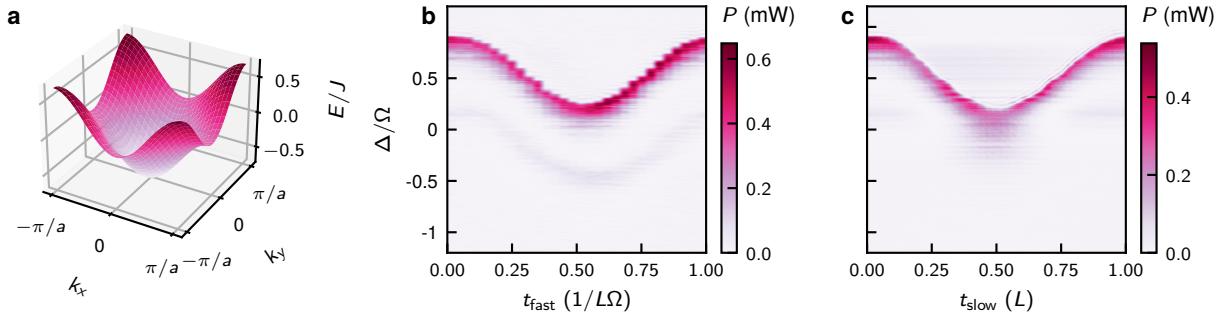


Figure 3.15: Decomposition of band structure of a 2D square lattice along edges of the Brillouin zone. (a) Analytic band structure for a 2D square lattice, with edges of the Brillouin zone highlighted with purple and red planes. By slicing the measured band structures into L chunks along the fast time axis each of length $t_{\text{fast}} = \Omega/L$, we can reconstruct the full 2D band structure. Slices of the reconstructed 2D band structure corresponding to $k_x = \pi/a$ (b), and $k_y = \pi/a$ (c).

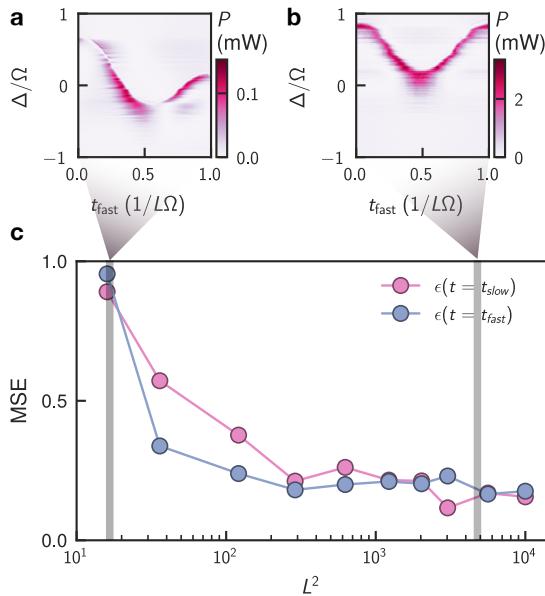


Figure 3.16: Band structure error as a function of lattice size. Slices of the band structure of a quasi-2D square lattice with $L = 3$ (a), and $L = 70$ (b) at the edge of the Brillouin zone. As the lattice size shrinks, the cosine dispersion becomes skewed as it the two momentum directions are no longer independent. (c) Mean squared error (MSE) of 2D band structure measurements of a square lattice as a function of lattice size, along edges of the Brillouin zone (see Fig. 3.15 in Section 3.4.3). MSE was calculated with respect to the analytic function convoluted with a Lorentzian with a finite width that was a free parameter, but fixed for all lattice sizes.

Figure 3.14 shows the modulation signal sent to the cavity to realize a 30×30 2D square lattice, and the measured band structure spectroscopy. The plot on the right is produced by taking a linear scan of the injection signal over a cavity mode, and measuring the response. The raw measurement is a 1D time series, which is then divided up into chunks set by the cavity roundtrip time $\tau = 1/\Omega$. The vertical axis is set by the scanning speed, here normalized by the cavity mode spacing Ω . Here, we scanned over one mode in 1 ms. These two time scales were observed to be well separated enough to allow the laser to equilibrium with the continuously changing scanning frequency. The widths of the band structures are proportional to $\Omega V/V_\pi$, here V is the modulation amplitude, and V_π is the pi-voltage of the phase modulator. For all band structure measurements, we drove the EOM very close to the pi-voltage in order to get wide band structures.

A real 2D square lattice has a band structure of $E(k) = -2J \cos(k_x a) - 2J \cos(k_y a)$. Here, we have a single dimension. To construct the full 2D band structure from this signal, we further separate the time-domain response signal into two further time scales by decomposing the horizontal axis of Fig. 3.14B into L chunks, so that each chunk is of length $t_{\text{fast}} = 1/L\Omega$, where L is the secondary long range coupling used to instantiate a 2D square lattice. The left plot of Fig. 3.15 shows the first chunk of the data plotted in Fig. 3.14. The secondary time scale is synthetically formed by looking at points separated by $1/L\Omega$. In other words, if we reconstruct the full 2D band structure by appending chunks of length t_{fast} , t_{slow} is the orthogonal direction pointing along the different chunks. These two timescales $t_{\text{fast}}, t_{\text{slow}}$ map to the two independent momenta k_x , and k_y , when the time scales are well separated enough, as is the case for $L \gg 1$. Figures 3.15B & C show the band structures along t_{fast} and t_{slow} , corresponding to borders of the Brillouin zone with $k_x = \pi/a$ and $k_y = \pi/a$, respectively.

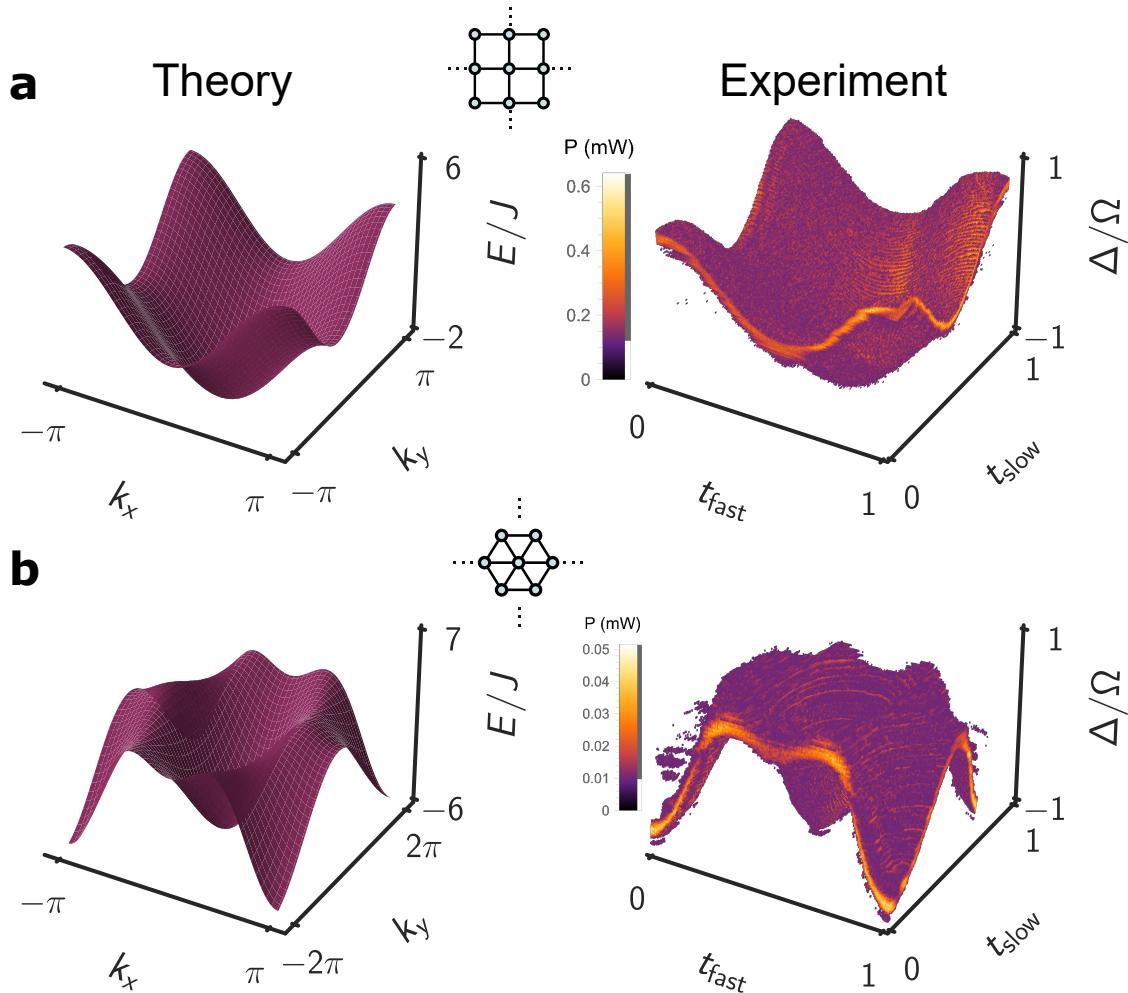


Figure 3.17: **Full reconstruction of band structure measurements for modulations instantiating lattices with $> 10,000$ lattice sites.** (a) band structure for a 2D square lattice comparing theory with experiment. (b) band structure for a 2D triangular lattice comparing theory with experiment. The experimental data is in effect a function of three variables. A opacity cutoff has been applied at low signal for clarity (shown in the grey bar adjacent to the colorbar).

The above procedure to measure bandstructure of multi-dimensional lattices is only justified if the time scales are separated, or equivalently, if $L \gg 1$. This is intuitively understood in the frequency domain according to equation 8 in the methods: the ratio of the coupling strength to the gain-loss in the system must be less than the linear dimension L , otherwise each local injection would have enough ‘momentum’

to propagate around the twisted boundary condition. Figure 3.16 shows the mean-squared error of the band structure of a two-dimensional square lattice as a function of system size along slices at the edges of the Brillouin zone.

Finally, Fig. 3.17 shows the full reconstructed 2D band structure for a square and triangular lattice, along with the analytic band structures. For clarity, an opacity filter was applied to the experimental data for the 3D plots in Fig. 3.17 such that only points greater than a certain power level are plotted. The band structures shown in the main text are slices of the full reconstructed band structure along high symmetry points. The theory curves were computed from the analytically solvable band structures, namely $E_{2D\text{ square}}(\vec{k}) = -J \cos(k_x) - J \cos(k_y)$, $E_{2D\text{ tri}}(\vec{k}) = -J \cos(k_x) - J \cos(k_x - k_y \sqrt{3}/2) - J \cos(k_x + k_y \sqrt{3}/2)$, $E_{3D\text{ square}}(\vec{k}) = -J \cos(k_x) - J \cos(k_y) - J \cos(k_z)$, and $E_{3D\text{ tri}}(\vec{k}) = -J \cos(k_x) - J \cos(k_x - k_y \sqrt{3}/2) - J \cos(k_x + k_y \sqrt{3}/2) - J \cos(k_z)$ for a 2D square, 2D triangular, a 3D square, and a 3D triangular lattice respectively. The theoretical curves for the density of states were computed using the Kwant code [45]. The presence of non-uniform response in the band structure is reflected partially in the density of states measurements, and indicates the multiplicities of projections onto the resonant modes [77].

In order to take paths through the high symmetry points of the Brillouin zone, high resolution experimental data was needed. For the band structure measurements, the cavity was modulated at a slightly detuned frequency ($\sim 10 - 20$ Hz) in order to match with a multiple of the sampling rate of the oscilloscope. This also prevented a walk-off in the reconstructions when the scope was sampling rate was detuned from the cavity mode spacing, leading to a linear shift in the band structure along t_{slow} when digitized.

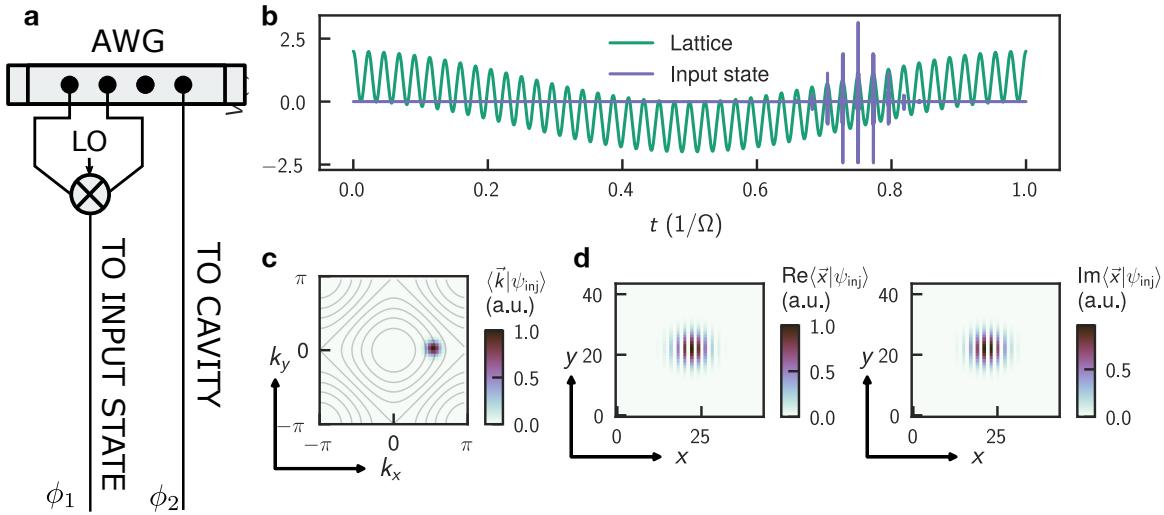


Figure 3.18: Overview of input state preparation in a 2D square lattice. (a) Two channels of an AWG form the in-phase and out of phase quadratures of a signal up-converted with a 12 GHz LO. A third channel is sent to the intra-cavity EOM. These channels are locked to the same clock, therefore the phase difference between the input state preparation and the cavity phase modulation is given by the cable length. (b) Time-domain signal sent to the cavity phase modulator (green) and the injection phase modulator (blue). (c) Two dimensional representation of the input state in momentum/time space with level curves of the 2D band structure plotted in gray. (d) Real and imaginary parts of the injected signal in coordinate/frequency space.

3.4.4 Input-state preparation

To prepare arbitrary input states, high modulation bandwidth and high preparation fidelity with no spurious images or modes are required. To this end, we implemented an image rejection IQ mixer in the optical domain by combining a 12 GHz-phase modulator with a fiber Bragg grating as a filter (see Fig. 3.6). The Bragg grating is a 4 GHz-bandpass filter, which enabled the programmability of around 4000 lattice sites, while rejecting the spurious sidebands in addition to the carrier frequency. The filter is centered 12 GHz away from the injection seed, so the modulator was driven with a signal with a center frequency of 12 GHz. The baseband signal around the 12-GHz sideband was encoded in a voltage signal from an AWG as I and Q pairs. These two were then sent into an (electronic) IQ mixer upconverted with a 12-GHz

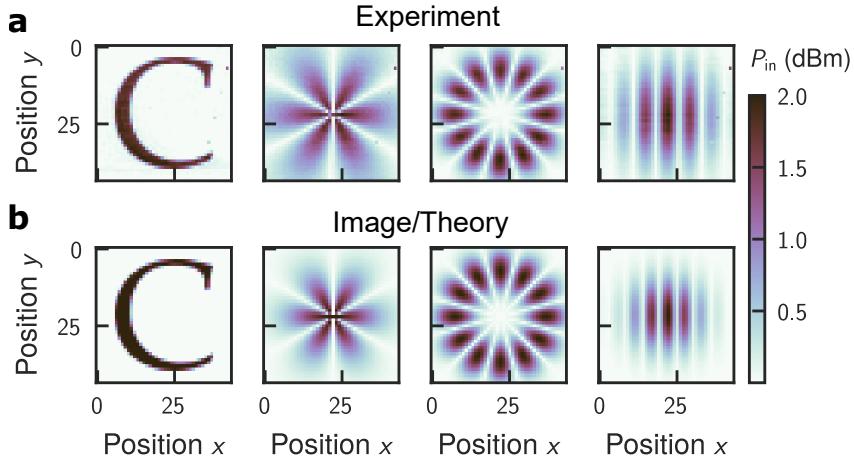


Figure 3.19: **Comparison of experimentally prepared input states with the digital image.** (a) Experimentally measured input states measured with heterodyne with a detector before the cavity (see Figure 3 of the main text). Here, a portion of the input spectrum has been folded into a 44×44 image. (b) Digital image of the input states fed as input into the phase modulator.

local oscillator (Fig. 3.18). The resultant upconverted signal was then sent to drive the external EOM, enabling both phase and amplitude programmability of the input state at every lattice position.

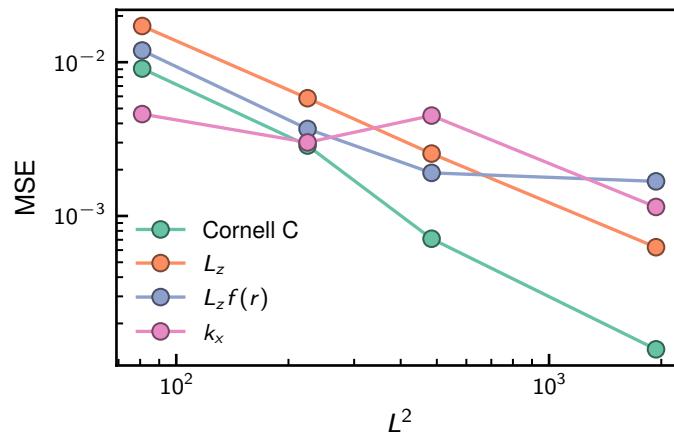


Figure 3.20: **Mean-squared error (MSE) of various input states displayed in the main text as compared to the input image as a function of 2D grid size.** Here, before calculating the error with respect to the linear power, both the theoretical and experimental traces were normalized.

To account for spectral inhomogeneities in the input chain, we calibrated the modulation by measuring the light before entering the cavity. If the modulation has some inhomogeneity, such that a voltage signal $V(t) = \sum_n V_n \sin(n\Omega t)$ is modified to $V(t) = \sum_n (V_n \eta_n) \sin(n\Omega t)$, the spectrum analyzer output modes are scaled by η_n^2 . To compensate, we injected light that had been uniformly modulated at all integer multiples of the FSR within the spectral region of interest, i.e. a top hat distribution defined on lattice sites. This region of interest is roughly up to 2000 modes for the input states presented in Fig. 3B in the main text. The square root of the response of this measurement gives us approximately η_n , which we used to apply an envelope function to the modulation signal. If successful, a top hat modulation multiplied with this envelope function will produce a clean flat spectrum. Otherwise, this process can be iterated for higher order. The measured input states shown in Fig. 3.19 were produced with just a single iteration of the above procedure. The input states were prepared with high fidelity with using this procedure. The mean-squared error between experimentally measured input states versus the original image are shown in Fig. 3.20, calculated from the magnitudes of both, normalized to unity.

The phase modulator preparing input states and the phase modulator programming the cavity interactions are synced to the same clock, however, the voltage signal driving the input state modulator is upconverted before hitting the EOM (Fig. 3.18). This imparts a phase difference between the input state and the cavity due to the different cable lengths. As shown in Fig. 3.18B, the phase difference between the injected light and the cavity defines the average momentum of the excitation. To account for this delay, we prepared a state with net-zero momentum (in the frame of the outgoing signal) into a tight-binding lattice, and tuned the phase of the cavity signal until we observed no transport.

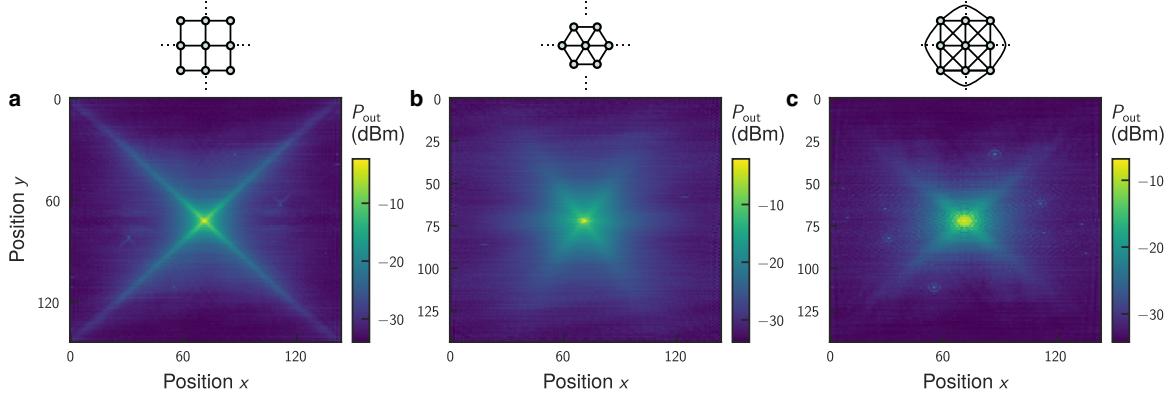


Figure 3.21: Response of single site injection in a 2D square lattice with nearest neighbor connections. (a), 2D triangular(b), and 2D square with second and third-nearest connections(c).

3.5 Supplementary results

Here we present additional measurements of transport in a variety of lattices, showcasing the programmability of our photonic system. The top row of Fig. 3.21 shows the steady state cavity response for a 2D square lattice with nearest neighbor connections, a 2D triangular lattice, and a 2D square lattice with nearest and next-nearest neighbor connections. Beneath these are the simulations of the corresponding lattices with a tight-binding Hamiltonian without the twisted boundary conditions.

Figure 3.22 shows the correlations of a simple 1D lattice with either nearest neighbor hoppings, or next-nearest neighbor hoppings. We see that in Fig. 3.22, as the strength of the nearest-neighbor coupling becomes small relative to the next-nearest neighbor coupling, the correlation length broadens and becomes tessellated, indicating that nearest neighbors are no longer correlated. The inhomogeneity along the diagonals is due components in the readout chain, e.g. detector, amplifier, and spectrum analyzer, as well as input chain, e.g. bandpass filters, amplifier, and injection. These matrices are constructed from the measured spectral response of the lattice at

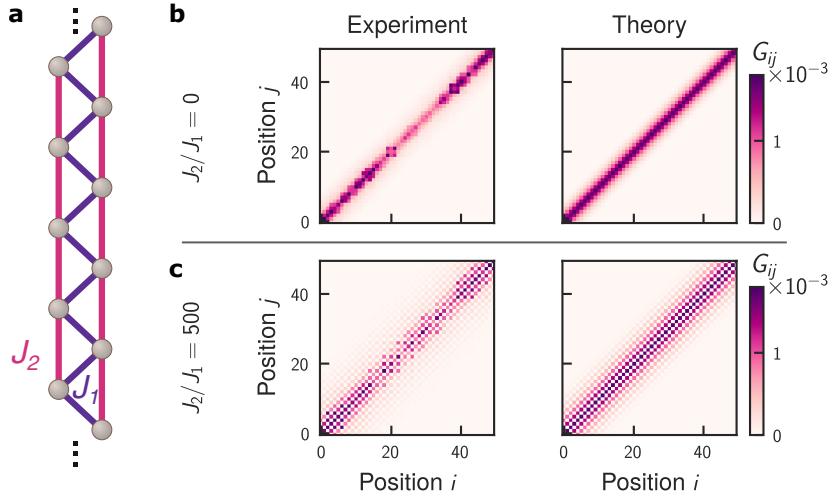


Figure 3.22: A 1D lattice with both nearest neighbor and next-nearest neighbor connections. (a). With only nearest neighbor connections (b), the correlations decay exponentially. As the next-nearest hopping is increased (c), the correlation length doubles on average, and nearest neighbor sites become uncorrelated.

a given injection site. That is,

$$G_{ij} = \langle n(i)^* n(j) \rangle - \langle n(i)^* \rangle \langle n(j) \rangle, \quad (3.3)$$

where $n(j)$ denotes the population at site j , and $n(i)^*$ denotes the population at site i given the injection was made at site i , and the $\langle \dots \rangle$ brackets denote the average over injection sites. We denote these quantities as the non-equilibrium correlation matrices, and find these measurements capture correlations of systems also found in other literature [100].

Additionally, we extend previous works in realizing synthetic electric fields via the observation of Bloch oscillations [10, 102], shown here in a 1D nearest-neighbor chain for arbitrary input states. As in Refs. [78, 19, 134, 106, 105], we realize a synthetic voltage by modulating the cavity with a slightly detuned frequency, as depicted in Fig. 3.23. A general time-dependent phase modulation for a nearest-neighbor cou-

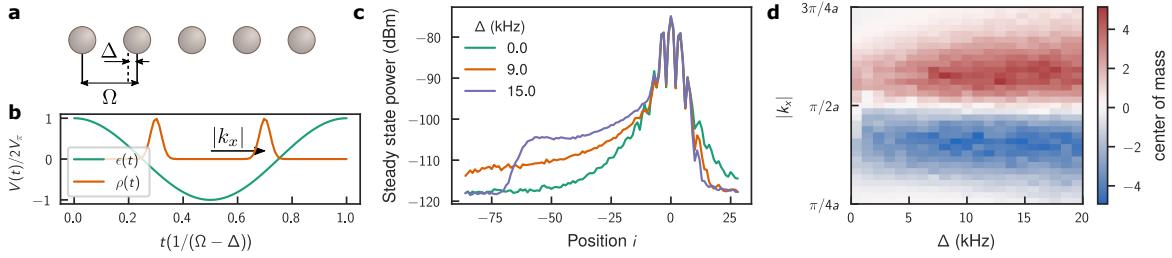


Figure 3.23: Synthetic electric fields and Bloch oscillations via detuned cavity modulations. (a) Coupling nearest neighbor sites with a detuned drive implements an effective electric field (see text). (b) Green and orange curves are the voltage signals driving cavity and input phase modulators respectively. The input state is chosen to be a superposition of two Gaussian wavepackets with equal and opposite momenta k_x , forming a standing wavepacket. (c) Steady state output cavity spectra for different values of detuning. With no detuning (green) the response is symmetric. As detuning is increased, the light will traverse to a maximum amplitude given by the inverse of the detuning. (d) Full experimentally measured phase diagram of Bloch oscillations captured by measuring the center of mass of the output spectra (Eq. (3.6))

pled chain realizes the Hamiltonian

$$H = \sum_n a_{n+1}^\dagger a_n e^{i\theta(t)} + \text{H.c.} \quad (3.4)$$

By performing the gauge transformation $|\Psi\rangle = \sum_n C_n a_n^\dagger |0\rangle \rightarrow \sum_n C_n e^{i\theta(t)} a_n^\dagger |0\rangle$, the Hamiltonian becomes [133]

$$H = \sum_n (a_{n+1}^\dagger a_n + \text{H.c.}) + \sum_n n \dot{\theta}(t) a_n^\dagger a_n. \quad (3.5)$$

By simply detuning the modulation by Δ , such that $\theta(t) = \Delta t$, we implement a linearly increasing voltage, or equivalently, a constant electric field in the 1D chain. These give rise to Bloch oscillations [134], which have been seen in photonic simulators, but are difficult to observe in real material systems, requiring very clean samples. The effects of Bloch oscillations on bandstructure have also been measured in photonic systems [78].

As an example, we prepare a 1D lattice and inject superpositions of equal but opposite momentum wavepackets, i.e. standing wavepackets (Fig. 3.23B). In the ab-

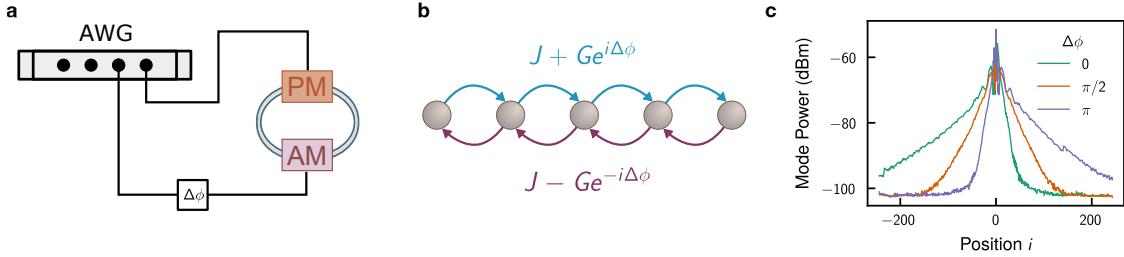


Figure 3.24: Non-Hermitian transport in the 1D generalized Hatano Nelson model with the addition of an amplitude modulator inside the cavity. (a, b) Coupling nearest neighbors with both an amplitude and phase modulator enables the realization of lattices with dissipative and asymmetric coupling, with the degree of non-Hermiticity encoded by the phase delay between the signals sent to the modulators (shown here as a physical delay, but implemented electronically in the experiment). (c) Experimental measurements of the 1D generalized Hatano-Nelson model, for $\Delta\phi = 0, \pi/2$, and π .

sence of an electric field, these states will diffuse, growing the Gaussian envelope of the packet. In the presence of an electric field however, the mirror symmetry of the lattice is broken, and the packet gains some overall momentum, shown in the measurements in Fig. 3.23C. In Fig. 3.23D, we measure the center of mass

$$\langle x \rangle = \int_{\text{lattice}} x \langle x | \psi_{\text{out}} \rangle dx \quad (3.6)$$

over all values of standing wave momenta values, as well as detunings up to 20 kHz.

The above procedure could be generalized to higher dimensions as well as to time-dependent electric fields. In the results presented so far for two-dimensional lattices, the coupling in the second dimension was just an integer multiple of the first, however, these can be independently detuned from each other, instantiating an electric field in either direction. Additionally, by varying the detuning as a function of time, one can implement AC electric fields by frequency modulating the voltage signal driving the phase modulator over a kHz time scale.

Finally, we utilize the system to explore non-Hermitian phenomena. We modify

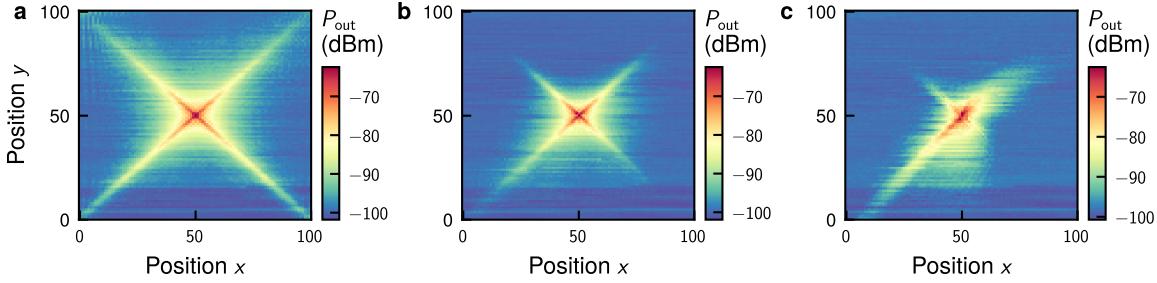


Figure 3.25: **Hatano Nelson ($\Delta\phi = 0$) model in 2D for asymmetry factor $G/J = 0.01$.** (a), 0.05 (b), and 0.2 (c).

the experiment shown schematically in Fig. 3.6 by adding an amplitude modulator. The action of the amplitude modulator subject to a periodic drive is

$$a_{t+\tau} = (1 + \sum_m \beta_m \sin(\omega_m t + \theta_m)) a_t, \quad (3.7)$$

where the G_m 's are proportional to the tone amplitudes of the voltage drive, and the θ_m 's are the phases of the drive up to some constant delay for each tone due to cable delay. Both the β_m 's and the θ_m 's can be finely tuned.

It can be shown, that by modulating the amplitude modulator at multiples of the cavity frequency spacing Ω , the generalized Hatano-Nelson Hamiltonian

$$H = \sum_{m,n} (J_m + G_m e^{i\Delta\phi_m}) a_{n+m}^\dagger a_m + (J_m - G_m e^{-i\Delta\phi_m}) a_{n-m}^\dagger a_m \quad (3.8)$$

can be realized [122]. The case where $\Delta\phi_m = 0$ for all m , or $\Delta\phi_m = \pi$ for all m corresponds to the regular Hatano-Nelson model [49]. To implement this model, we added an amplitude modulator to the cavity. We drive this modulator with the same signal as is sent to the phase modulator, but with an adjustable phase shift and amplitude scaling (Fig. 3.24a). This allows us to implement non-Hermitian lattices in 1D and higher.

Figure 3.24 shows experimental measurements of the HN model in a 1D chain. By tuning the phase delay $\Delta\phi$ between the signals, we can explore different regimes

of the model, shown in Fig. 3.24c for $\Delta\phi = 0, \pi/2$ and π . The first and last curves correspond to the regular HN model, realized highly asymmetric transport, breaking time-reversal symmetry.

Using the procedure to implement 2D and higher Hermitian lattices, we implement a 2D Hatano Nelson model. Figure 3.25 shows the experimental measurements of transport due to a single-site injection at the center of the non-Hermitian lattice. Here, the phase delay was chosen to be 0 for both modulation tones to maximize the asymmetric transport in both directions, while the strength of the asymmetry was varied. To our knowledge, this is the first demonstration of a multi-dimensional Hatano-Nelson model in an optical platform.

CHAPTER 4

MICROWAVE SIGNAL PROCESSING USING AN ANALOG QUANTUM RESERVOIR COMPUTER

This chapter is based on the following publication:

- Senanian, A., Prabhu, S., Kremenetski, V., Roy, S., Cao, Y., Kline, J., Onodera, T., Wright, L.G., Wu, X., Fatemi, V. and McMahon, P.L., 2023. Microwave signal processing using an analog quantum reservoir computer. arXiv preprint arXiv:2312.16166. (To appear in Nature Communications).

4.1 Introduction

Over the last decade, researchers in quantum information processing have broadly divided their efforts into two distinct but complementary directions. In one, the focus has been on realizing the building blocks for large-scale, fault-tolerant quantum processors [73, 65, 17], which would enable running algorithms such as Shor’s or Grover’s at meaningful scale. In the other, there has been a push to realize quantum systems comprising tens to hundreds of qubits or qumodes, but without error correction, and to explore what can be done with such noisy, pre-fault-tolerance systems—often denoted as noisy, intermediate-scale, quantum (NISQ) devices [104]. Quantum computational supremacy with such NISQ devices has been demonstrated [2, 136], but there has been much less progress on achieving quantum advantage in practically relevant applications [90]. There have been many NISQ studies on quantum machine learning [11], and in this area too, quantum advantage for problems of broad practical interest has remained elusive [111, 18]. A key challenge in quantum neural networks

realized with parameterized quantum circuits has been training the parameters when the optimization landscape suffers from barren plateaus [86, 123, 84, 1]. A major open question is whether one can achieve any practically relevant advantage for machine learning with NISQ systems.

Inspired by the framework of reservoir computing [80, 110, 68, 42] in classical machine learning, quantum reservoir computing (QRC) [39, 38, 43, 44, 66, 115] has emerged as an approach to quantum machine learning that entirely avoids barren plateaus by performing all learning in a final, linear layer. The key idea of a QRC is that a quantum system (called a *quantum reservoir*) can generate nonlinear, high-dimensional features of inputs to it, and that these features can be used to perform machine-learning tasks purely by training a classical linear transformation. However, experimental demonstrations to date have been performed with digital quantum circuits [103, 20, 72, 89, 130, 117, 55] that have limited the complexity of tasks that can be performed, in part due to an input bottleneck imposed by the use of discrete gates to input temporal data using a series of separate, imperfect gates.

The aim of our work is to demonstrate a proof-of-principle for a new application of and approach to quantum machine learning with NISQ devices that overcomes or sidesteps the challenges in training and inputs noted above. We use the driven, continuous-time analog quantum nonlinear dynamics of a superconducting microwave circuit as a quantum reservoir to generate features for classifying weak, analog microwave signals (Fig. 4.1a). We use repeated measurements of the reservoir both to extract features that contain information about temporal correlations in the input data, as well as to induce non-unitary dynamics. Our use of a continuous-variable system in our quantum reservoir grants us access to a substantially larger Hilbert space than would be the case with a qubit-only system with equally many

hardware components. Our approach is similar to proposals for analog NISQ processors and simulators [98, 41, 22], which aim to avoid the overhead caused by imposing a discrete-time abstraction. Analog operation grants us an even more important ability however, which fundamentally distinguishes our work from prior experimental demonstrations of quantum machine learning on circuit-model quantum processors: it allows our device to directly, natively receive weak analog microwave signals, and to immediately leverage analog quantum information processing to extract relevant features of the signals for classification. Our experiments do not address the question of whether a QRC can achieve a quantum computational advantage, since our experimental device is small enough to be easily classically simulable. However, our demonstrations suggest a route to achieving a quantum advantage of a different kind: an advantage in the quantum detection and processing of weak microwave signals, allowing quantum hardware to extract complex information of interest from dim, analog signals in ways that would be noisier with a conventional classical approach. This type of quantum advantage, arising from a combination of quantum sensing with extraction of complex features about the sensed signal, is discussed in general terms as a route to quantum advantage with quantum machine learning in Ref. [18]. Our work shows that when classical signals comprising just a few photons have entered an analog quantum reservoir, they can be classified using our QRC approach. The signals we classify are synthesized at room temperature and pass through 60 dB of attenuation before reaching our device. However, if instead one combines this analog quantum processing with a sensitive quantum detector of microwave radiation, as has already been previously demonstrated using superconducting circuits [125, 124, 4, 3, 26], then one can construct a system that achieves a quantum advantage in the task of combined sensing and signal processing of high temperature signals.

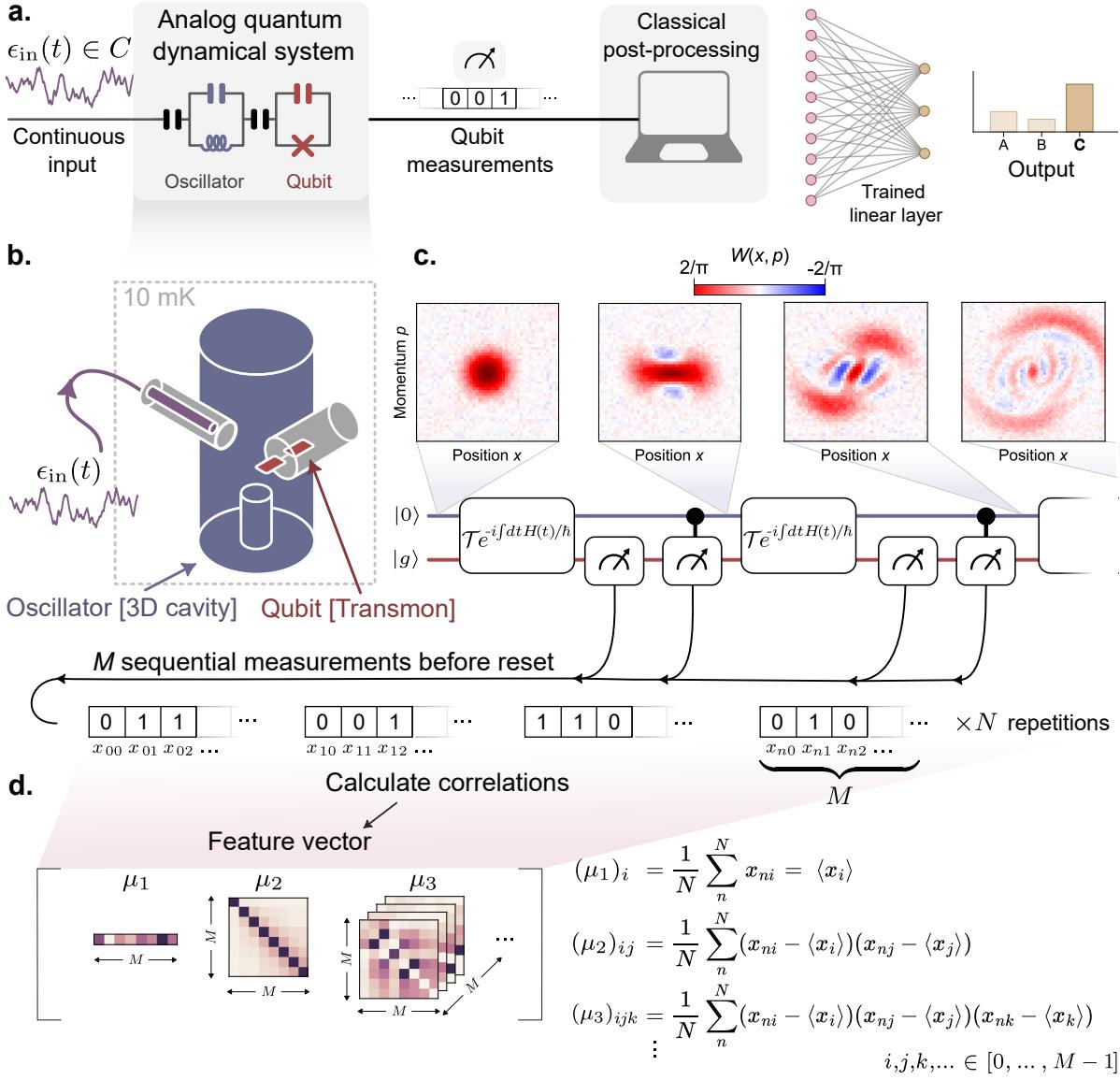


Figure 4.1: Analog signal classification with a continuous-variable quantum reservoir computer (QRC) using measurement trajectories. (a.) We perform machine learning using a quantum system consisting of an oscillator coupled to a qubit. (b.) The signals interface directly with the qubit-oscillator system, composed of a 3D aluminum cavity (blue) hosting a transmon qubit (red). (c.) Wigner tomography performed on the oscillator state through various stages of the reservoir dynamics. The dynamics include entanglement-generating unitary evolution, and projective measurements of both the qubit and oscillator. The balance of measurements and unitaries lead to complex correlations in the measurement trajectories. (d.) The digital linear layer performs classification based on a feature vector, which we construct using the expectation values of the central moments μ_p ($p = 1, 2, 3, \dots$).

4.1.1 Experimental setup and protocols

Our quantum reservoir, composed of a cavity resonator coupled to a transmon (Fig. 4.1b), can be modeled with the following qubit-oscillator Hamiltonian in the rotating-frame,

$$H/\hbar = -\chi|e\rangle\langle e|a^\dagger a + \epsilon(t)a^\dagger + \Omega(t)|e\rangle\langle g| + \text{H.c.}, \quad (4.1)$$

where $|g\rangle$ and $|e\rangle$ define the qubit subspace of the transmon, a is the photon annihilation operator of the oscillator mode, and χ is the nonlinear interaction strength (see Section 4.5.1 for details). The third term of Eq. 4.1 describes the unitary control of the qubit with a time-dependent drive $\Omega(t)$, and the second term describes both the encoding of the input data $\epsilon_{\text{in}}(t)$, and unitary control of the oscillator mode, i.e., $\epsilon(t) = \epsilon_{\text{in}}(t) + \epsilon_{\text{control}}(t)$. Equation 4.1 describes the unitary dynamics, which is complemented by non-unitary dynamics generated by the back-action from qubit measurements interspersed throughout the evolution.

The oscillator and qubit control drives used in this paper realize a reservoir that consists of a series of entangling unitaries interleaved with qubit and oscillator measurements (Fig. 4.1c). The analog input is sent resonantly to the cavity and results in a time varying conditional displacement of the oscillator, which streams in concurrently with control drives. The cavity resonator hosting the oscillator mode has a resonance frequency of 6 GHz and a 2-kHz linewidth. The combination of the input and control drives implement a unitary that encodes the input into the state of the oscillator and generates entanglement between the qubit and the oscillator. Following the unitary evolution, we perform a qubit measurement, and then a parity measurement of the oscillator state [119, 50] (see Section 4.6.3). The parity measurement projects the oscillator state into super-positions of either even or odd Fock states, a highly non-Gaussian measurement allowing one to sense changes in the photon-number distri-

bution. Additionally, the entangling dynamics between the measurements effectively implement a sequence of non-commuting measurements (see Section 4.5.2), generating correlated measurement distributions that can then be used as complex output features. Finally, after four rounds of applying the unitary and the qubit-oscillator measurements, we reset the system before repeating the scheme so that we may collect many samples of the measurement trajectory. The reset, which occurs at a rate much faster than the decoherence rate of the oscillator, additionally ensures that our system remains coherent.

The measurement outcomes are used to construct output feature vectors to be fed into the linear layer, but this can be done in a few different ways. When performing repeated measurements on our system, we generate a sample bitstring of length M describing the quantum trajectory over M measurements. After M measurements are performed, we reset the system and repeat the procedure, each time generating a bitstring $\vec{x}_n = [x_{n0}, x_{n1}, \dots, x_{nM-1}]$, where n refers to the n th sample (Fig. 4.1c). The outcomes can be counted to directly form a sample probability distribution $p(\vec{x}|\epsilon_{\text{in}}(t))$ over measurement trajectories, which can then be used as a high-dimensional output feature vector after obtaining a sufficient number of samples N . While this approach has the benefit of capturing all information in the measurement distribution [55], it can generally suffer from poor scaling in sampling noise, requiring $N \sim 2^M$ shots in the worst case [128].

Here, we construct an output feature vector from estimates of successive central moments $\mu_1, \mu_2, \mu_3, \dots$ of the underlying distribution $p(\vec{x}|\epsilon_{\text{in}}(t))$ (Fig. 4.1d). For example, the first-order central moment μ_1 is a M -dimensional vector representing the average over all measured bitstrings, i.e. $\mu_1 = [\langle x_{n0} \rangle, \langle x_{n1} \rangle, \dots]$, the second-order central moment μ_2 is the covariance matrix with elements $(\mu_2)_{ij} = \langle x_{ni} x_{nj} \rangle - \langle x_{ni} \rangle \langle x_{nj} \rangle$,

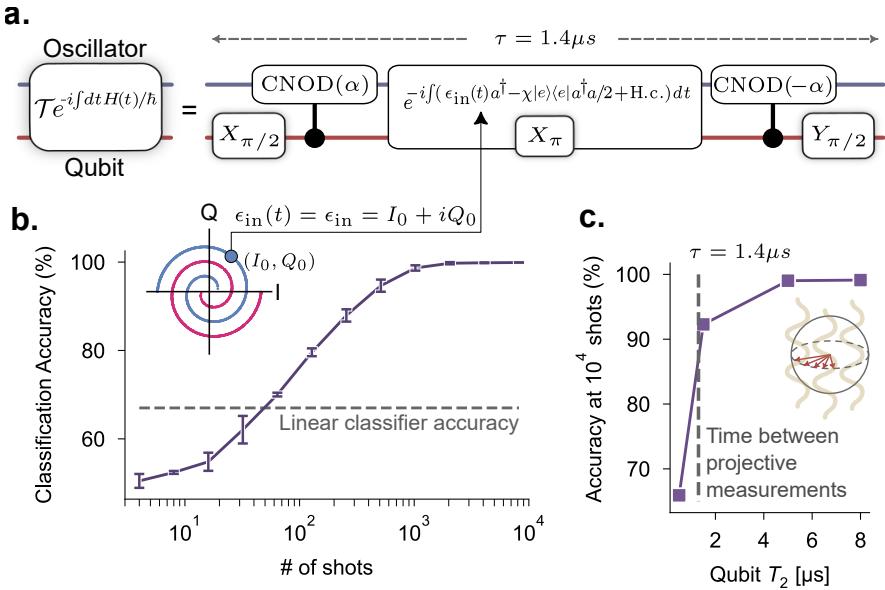


Figure 4.2: Reservoir protocol overview with an example time-independent classification task **(a)** The unitary dynamics in our reservoir are generated by control pulses that serve to entangle the qubit with the oscillator before the analog input is received by the oscillator. For tasks where the analog data is time-independent, the dynamics are fully gate-based, and the oscillator is dis-entangled with the qubit before the qubit and oscillator measurements. For details of the motivation behind the particular unitaries implemented for our reservoir, see Section 4.5. **(b)** (Inset) An illustrative machine learning example is the classification of time-independent signals from two arms of a Spiral distribution defined in the signal $I - Q$ plane. For quantum machine learning, unlike classical, the performance is unavoidably impacted by sampling noise. Here, we plot the classification accuracy of the spiral task against an increasing number of shots. Also plotted is the performance of a linear layer acting directly on the two-dimensional I, Q data, indicating that non-linearity is required to perform this task with sufficient accuracy. **(c)** Classification accuracy at 10^4 shots as a function of qubit coherence time that we tune via resonator-induced dephasing during the classification (see Section 4.6.4). While we see a large drop in classification performance when the qubit coherence time is heavily suppressed and the system is completely disentangled, the performance only begins to suffer once the qubit T_2 approaches the duration between measurements.

and so on. Here, the expectation value is taken over the sample index n . This approach, inspired by Ref. [68], has the benefit of leveraging the hierarchy of noise in the central moments, while capturing the essential correlations in the dynamics to achieve high accuracy even in the few-sample regime. Furthermore, the output feature vector dimension only scales polynomially with the number of measurements,

where the highest polynomial power is given by the order of the highest central moment, which we restrict to 3 for all tasks in this work. Finally, given finite memory in our reservoir, we further restrict the output vector by choosing to only calculate correlations between measurements at most 3 measurements apart. These truncated moments are then flattened and concatenated to construct our output feature vectors. In all, for the $M = 8$ measurements we use in this work, the resultant output feature vector size with this prescription is 94. For a detailed discussion of the construction of our reservoir output features with comparisons of different encodings, see Section 4.7.1.

4.2 Results

4.2.1 Classification of time-independent signals

To illustrate the scheme proposed in this work, we begin with an example classification using our quantum reservoir by performing binary classification task of time-independent signals. Figure 4.2a describes the control drives in more detail. For time-independent input data, the two-dimensional input data is encoded as the I and Q quadratures of an analog signal resonant with the cavity resonance frequency. In the rotating frame of the system (Eq. 4.1), this is effectively a time-independent signal, i.e. $\epsilon_{\text{in}}(t) = \epsilon_{\text{in}} = I + Q$, which results in a displacement of the oscillator state. For such time-independent tasks, the signal bandwidth is set by its duration which, in general, can make the resultant displacement conditioned on the qubit state due to the cross-Kerr interaction (see first term of Eq. 4.1).

The unitary encoding the input displacement is complemented by control drives

that entangle the qubit and oscillator via conditional displacements [25] and qubit rotations (Fig. 4.2a). The entangling conditional displacements are applied before and after the unknown input is fed into the system, and the qubit is rotated by π or $\pi/2$ pulses before, during, and after the input. Due to the qubit-state-dependent shift of the oscillator frequency by $-\chi$, these qubit rotations serve to make the oscillator sensitive to the input signal independent of the state of the qubit at the start of each round of input. Additionally, when combined with conditional displacements on the oscillator, the control and input scheme impart a geometric area enclosed by the oscillator trajectory onto the qubit, such that the phase of an unknown time-independent input signal can be extracted via a qubit measurement (see Section 4.5 for details of this unitary). In Section 4.10, we show the ability of the set of unitaries implemented here to be able to approximate any scalar function of the input signal when the signal is time-independent. For all results presented, we implement our reservoir unitary with these control drives across all tasks, with 4 applications of the unitary interleaved with qubit and oscillator-parity measurements.

The binary classification task we perform here is: Two distributions of time-independent signals, completely characterized by the signal’s in-phase (I) and quadrature (Q) components, are distributed along two separate “arms of a spiral” in the $I - Q$ plane (Fig. 4.2b). Given a displacement described by the points I and Q sampled from either signal distribution, one must figure out which distribution the signal came from. The maximum amplitude of the input signal distribution $\max(|\epsilon_{in}|)$ (i.e. the points in the spiral arms furthest away from the origin in Fig. 4.2b) was chosen such that the amount of displacement of the oscillator state initialized in vacuum would result in a coherent state with $\bar{n} = 0.3$ photons per round of input ($\sim 1 \mu\text{s}$). This input amplitude was needed in order to perform the classification with sufficient accuracy in a reasonable amount of shots. Our QRC solved the spiral classi-

fication task with $> 97\%$ accuracy at 10^3 shots (Fig. 4.2b). This simple task has the feature that, if one feeds in the inputs directly into a linear layer, the classification accuracy would reach no more than 67%—just above the random guessing accuracy of 50%. As a point of comparison with non-linear digital reservoirs, we found that a 64-dimensional, two-layer digital reservoir was needed to achieve the same performance as our quantum reservoir for this task (see Section 4.11.2 for details of this comparison).

To probe the role of quantum coherence in our reservoir, we performed the same classification task, but with reduced coherence time in the qubit during the reservoir execution (Fig. 4.2c). This was achieved by populating the lossy readout resonator with photons that send the qubit to the center of the Bloch-sphere when the readout resonator is traced out (see Section 4.6.4). With $T_2 \rightarrow 0$, we effectively removed all entanglement with the oscillator, and observed two things: a dramatic reduction in classification performance, and importantly, T_2 only began affecting the performance once it was on the order of the reservoir duration, after which the qubit is projected to a pure state.

4.2.2 Classification of radio-frequency (RF) communication modulation schemes

Next, to highlight the ability to perform classification of higher dimensional data, we classified time-dependent radio-frequency (RF) signals. The microwave signals in this dataset encode digital information using one of 10 different digital modulation schemes, a standard benchmark task in RF machine learning [96, 61]. Digital modulation schemes encode binary information in discrete ‘symbols’ encoding in sequential

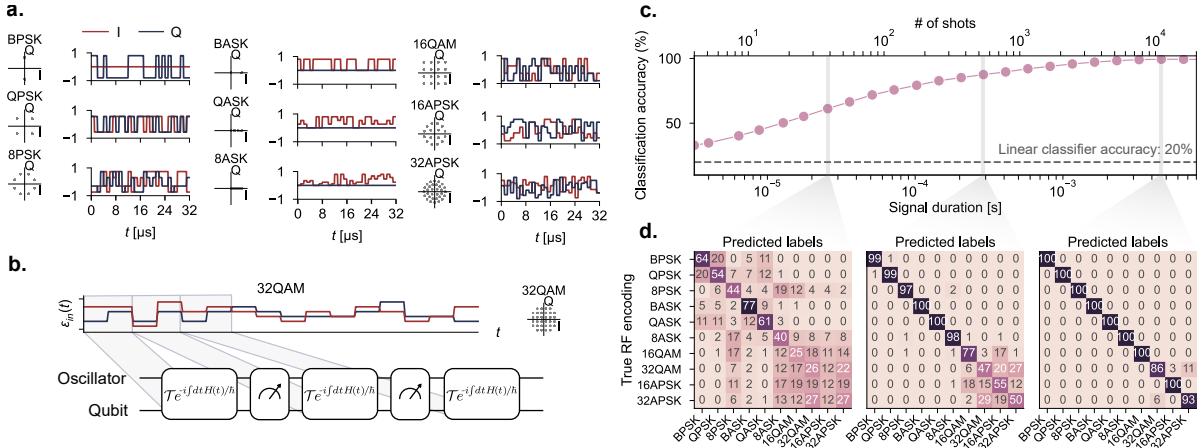


Figure 4.3: Classification of radio-frequency (RF) communication modulation protocols. (a) Description of the dataset for digital modulation schemes used in this experiment. In conventional digital modulation schemes, one encodes data in the amplitude and phase of the signal. The modulation schemes can be represented by a “constellation diagram” in (I, Q) space (left), where point represents one of the possible choice of (I, Q) values to encode a symbol, with example time traces (right). (b) These signals are broken up and fed into our reservoir. (c) The performance of the reservoir as a function of the shots taken in real time (see text). The top row contains the corresponding duration of the radio frequency signal required. For context, a classical linear classifier applied directly on the input data achieves only 20% accuracy, independent of the duration of the signal. (d) Confusion matrix for the QRC at 32, 512, and 10^4 shots, showing that the reservoir confuses only a few classes at the highest shots.

time-bins. For example, Binary Phase-Shift Keying (BPSK) encodes binary data in discrete phase jumps of a signal, such that a symbol 0 (1) maps to a phase flip of 0 (π). Other modulation schemes can encode more bits per symbol. BPSK and other encodings can be represented in a constellation diagram (Fig. 4.3a), which denotes the potential (I, Q) values a signal can take for each symbol. A given string of digital data can then be encoded in a time-domain signal by sequentially choosing points in the constellation diagram with a given symbol rate. For typical WiFi signals this is around 250 kHz per subchannel [64].

For this task, we generated RF signals by encoding random digital strings into the 10 different modulation schemes with a fixed symbol rate of around 2 symbols per μs ,

or with a sampling rate of 2 MSps. The duration of these signals typically lasts much longer than the reset period of our system. Importantly, we did not repeat the same signal to artificially reduce the sampling noise associated with each input data, as this would not typically be applicable in a real-world setting. Instead, the measurement statistics were generated by sampling the signal in real time. Consequently, what we refer to as ‘shots’ in a real-time task does not correspond to identical repetitions of the experiment, but instead, is the number of resets we performed while acquiring the signal, which changed from shot to shot. In effect, each different encoding scheme produces a unique “fingerprint” distribution over measurement outcomes, and the goal of the linear layer is to separate these distributions with as high accuracy as possible.

Figure 4.3c shows the accuracy in classifying digitally modulated RF signals with increasing number of shots, compared with the performance of a linear classifier. We note that in less than a millisecond, or with less than 2000 symbols, the reservoir was able to classify which of the 10 classes a given signal belongs to with > 90% accuracy when using 8 qubit-oscillator measurements. A linear classifier can only achieve 20% classification accuracy for this task, even with infinite symbols. The confusion matrix between the different classes at 32, 512, and 10^4 shots is displayed in Fig 4.3d, the latter two of which are nearly diagonal.

4.2.3 Classification of filtered noise

Finally, to demonstrate the performance of our QRC on continuous-time data¹, and with a task that requires both long-term and short-term memory in the quantum

¹The previous time-dependent task, RF-modulation-scheme classification, concerns discrete-time data.

reservoir, we performed the following classification task: input data assumed to have come from a source of white noise is filtered using a moving-average filter having one of three filter shapes (Gaussian, Lorentzian and inverse-power-law), and one of two window widths (50 ns and 600 ns), and the task is to identify both the filter shape and window width (Fig. 4.4a). The resultant dataset consisting of six classes of noisy signals was designed to probe the ability of our QRC to process high dimensional data with bandwidths larger than the cavity linewidth. Additionally, this task allowed us to probe the memory of our QRC and its ability to be sensitive to fluctuations in time, a key feature that enable temporal signal processing in QRCs [108, 85]. The filter functions were normalized so that the photon-number distributions generated by the time-dependent displacements are identical up to the filter width. This normalization was applied to ensure that the task is not trivially solvable by just measuring the mean photon number (see Section 4.8.2).

Because all the signals used in this dataset are noise with zero mean, a linear classifier would do no better than random guessing. By contrast, Figure 4.4b visually shows (using Singular Value Decomposition (SVD) on the output feature space) that the quantum reservoir was able to peel apart the different noise distributions. On the task of classifying over six different sources of noise, we achieved 93% accuracy (Fig. 4.4c) in only 2000 shots. As seen in the confusion matrix in Fig. 4.4d, the primary confusion at 2000 shots was distinguishing between the 50-ns inverse-power-law noise class and the 600-ns Gaussian noise class, as expected from the overlap in the SVD of the feature space.

Finally, we compared the ability of our reservoir to understand long vs short correlations in input signals. For this, we deconstructed the 6-class classification task into two classification subtasks, where in each subtask, the QRC learned to distin-

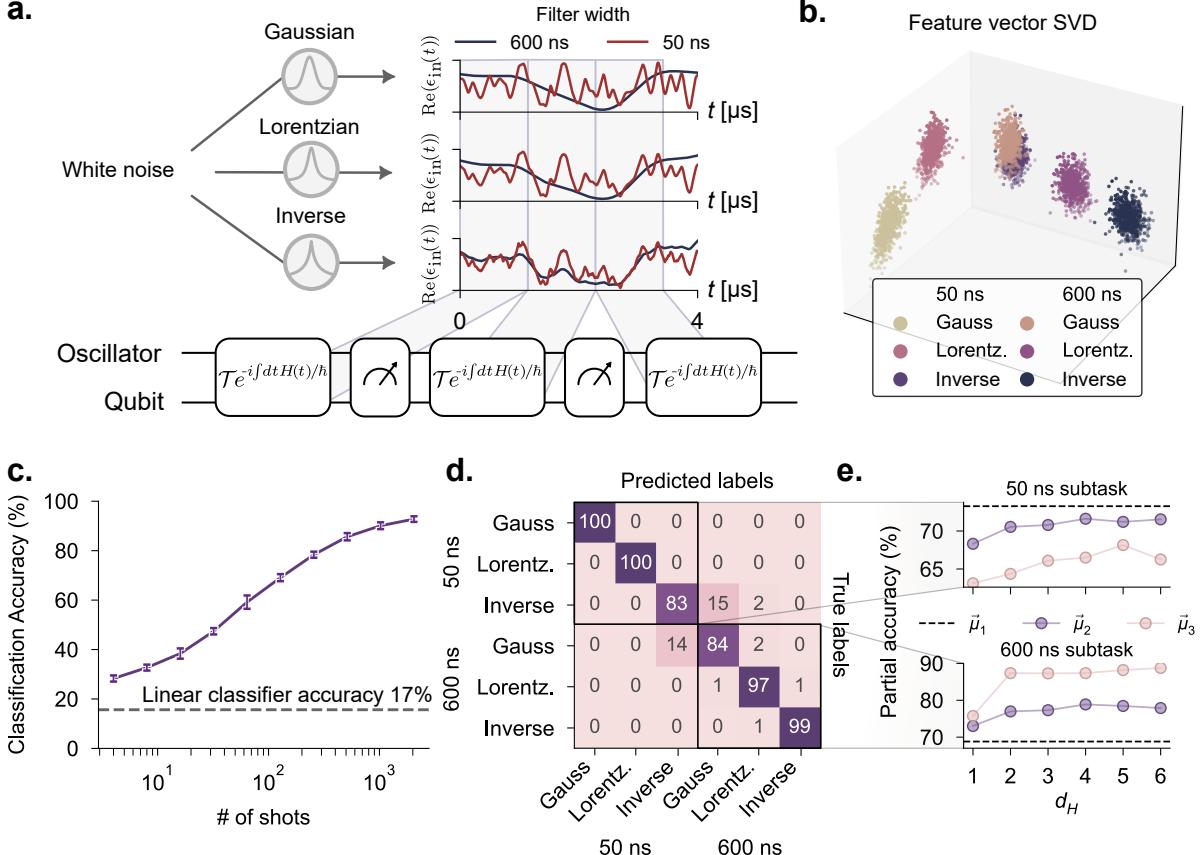


Figure 4.4: Classification of filtered noise **(a)** Three different filters are used for a moving average applied on white-noise signals: a Gaussian filter, a Lorentzian, and an inverse power law. For each filter, we generate stochastic analog signals based on both a 50 ns filter width, and a 600 ns filter width. **(b)** Visualization of the high-dimensional output feature space using Singular Value Decomposition (SVD). Each point corresponds to a different signal over 2000 shots taken in real time (see text). **(c)** Classification accuracy as a function of the number of shots using third-order moments as the output feature. Our reservoir reaches 93% accuracy in about 2000 shots, corresponding to about 10 ms of the signal received. **(d)** Confusion matrix of the task taken at 2000 shots. **(e)** Participation of the mean and the off-diagonal elements of the second- and third-order moments in the classification accuracy within the subtasks of classifying different noise sources with fixed filter width.

guish noisy signals generated from among three different filter window types, but with fixed window widths. The two subtasks differ by the filter window width (see Fig. 4.4d and e). The class of signals with coherence length of 50 ns highlights the convenience of our input encoding scheme, i.e. feeding signals directly into the os-

cillator mode without the need to sample the signal discretely in time. Additionally, the ability for our quantum reservoir to distinguish between signals with correlation times on the order of 50 ns demonstrates the sensitivity to signals which vary on time-scales much faster than the measurement rate. In contrast, classification of the class of signals with coherence lengths of 600 ns requires correlations of the reservoir dynamics beyond that of the measurement rate. To highlight the advantage of our scheme, we simulated the performance of a reservoir with that of a recent gate-based protocol where the input was sampled discretely in time [130]. Our simulation results, in Section 4.9.2, highlight the advantage of our protocol when the sampling rate of the input is slow, which can arise in experiment such as finite pulse durations and latency introduced by the FPGA classical comparison.

Figure 4.4e looks at the participation of the different moments μ_k of the measurements in the classification accuracy of the 50-ns subtask (top), and the 600-ns subtask (bottom). Here, the output features were constructed by the mean μ_1 , or the off-diagonal elements of the moments μ_2 and μ_3 as a function of the distance between measurements d_H , allowing us to probe the contribution of the moments as a function of the locality of the correlations. For the 50-ns subtask, we see that the most important contribution is the mean, with the second-order moment being the next-most important contribution, and the third-order moment being relatively unimportant. In stark contrast, the third-order moment is most important for the 600-ns subtask, surprisingly yielding nearly 90% classification accuracy using non-local third-order correlations alone. The ability to distinguish stochastic signals among the combined six classes demonstrates the ability of our reservoir to capture both slow and fast features of microwave signals.

4.3 Discussion

In summary, we have experimentally realized an analog quantum reservoir computer (QRC) and demonstrated its ability to directly process microwave analog input signals without discretization, achieving high classification accuracy on three different tasks. Previous demonstrations of quantum reservoir computing have used multi-qubit, gate-based quantum reservoirs [130, 103, 20, 117, 72, 89, 55]. In contrast, we perform machine learning directly on analog signals fed into a single oscillator coupled to a transmon qubit. The superconducting-circuits platform not only allows us to leverage projective non-demolition (QND) non-Gaussian measurements to generate correlated output features, but is also well-matched to process microwave signals that can generally be continuous in time. In addition to demonstrating accurate classification of microwave signals in our experiments, we also performed a direct comparison with a state-of-the-art discrete-time, gate-based QRC approach in simulation, and found that a continuous-time reservoir outperforms a discrete-time reservoir when the input signals contain temporal variations fast relative to the discretization time (see Section 4.9.2).

For any quantum neural network, including QRC approaches, a central concern is to what extent one can achieve high accuracy on a particular task without needing an impractical number of shots [128]. Ref. [55] recently reported that certain functions—termed *eigentasks*—can be constructed with low error from quantum reservoirs even when the number of shots is modest, giving evidence that for some tasks, sampling noise need not be overwhelming. In our experiments, we found that it was possible to achieve high accuracy for all the tasks we attempted while needing only 10^3 – 10^4 shots (depending on the task). There is important future work to be done in exploring the trade-offs between reservoir size (e.g., number of oscillators or qubits), number

of measurements M between reservoir resets, feature-vector dimension (dependent both on M and the choice of order of correlators to include), and number of shots required for both training and inference.

With improved quantum hardware, we anticipate that it will be possible to carry out even more sophisticated tasks than what we have already demonstrated. Increasing the coherence time of the oscillator would enable us to perform many more measurements (the qubit’s coherence time is, favorably, less important in our scheme because our protocol involves repeatedly projectively measuring the qubit). While we analytically showed in Section 4.10 the ability of our QRC to be able to approximate any scalar function of the input signal when the signal is time-independent, provided the number of measurements M performed is large enough, there remains the open theoretical question of the expressiveness of the QRC when the input signal is time-dependent. Generalizing our approach to spatial in addition to temporal inputs, as was explored in Ref. [115], would likely support more sophisticated computations. In Section 4.9.4, we explore such extensions in simulations and find a marked improvement in classification accuracy.

It is an open question if QRC—using the type of reservoir we considered in this paper, or any other—can, when implemented with NISQ hardware, achieve a quantum computational advantage over the best classical machine learning approaches, just as it is unclear if any quantum-machine-learning method can [18]. We did not investigate the potential for purely computational quantum advantage: our quantum reservoir is small enough to be easily classically simulable. However, our work opens up the possibility to experimentally achieve a different type of quantum advantage than a purely computational one. If one performs quantum processing on data obtained by a quantum sensor, there is the potential for an advantage that is a hybrid

of being due to the advantage of quantum sensing and of quantum computing [18]. Our work suggests the feasibility of concretely realizing this kind of hybrid quantum sensing-computational advantage, where the quantum sensor is a superconducting circuit that can detect classical microwave radiation with high quantum efficiency and low noise [125, 124, 4, 3]. While the signals classified in this work originate at room temperature and are highly attenuated before reaching the device, our experiments have shown that it is possible to accurately classify signals using a superconducting circuit even when there are only a few photons of signal in the superconducting circuit within any single run. Combining this with a sensitive quantum detector could lead to quantum smart sensors—quantum versions of classical in-sensor processors [138]—that can reliably extract information from weak microwave signals in a way that exceeds the accuracy of any equivalent classical system.

4.4 Experimental setup

The device used in this paper consists of an oscillator, a 3D stub post cavity made from high-purity 4N Aluminum treated with an acid etch, and a transmon qubit. The transmon, made of Niobium, is fabricated on a resistive silicon chip, along with an on-chip readout resonator also made of Niobium. The single chip hosting the transmon and the readout resonator is mounted in the 3D cavity package using copper clamps. The cavity and the copper clamp contain copper films for thermalization directly to gold-plated copper breadboard at the mixing chamber plate of the dilution refrigerator (Fig. 4.6). The device is shielded with Copper coated with Berkeley Black, and two types of magnetic shields: Aluminum, and Cryoperm (Fig. 4.5). The cavity pin is set such that the oscillator mode is undercoupled to the transmission line by a factor of 40. While this reduces the transmission of photons incident on our device

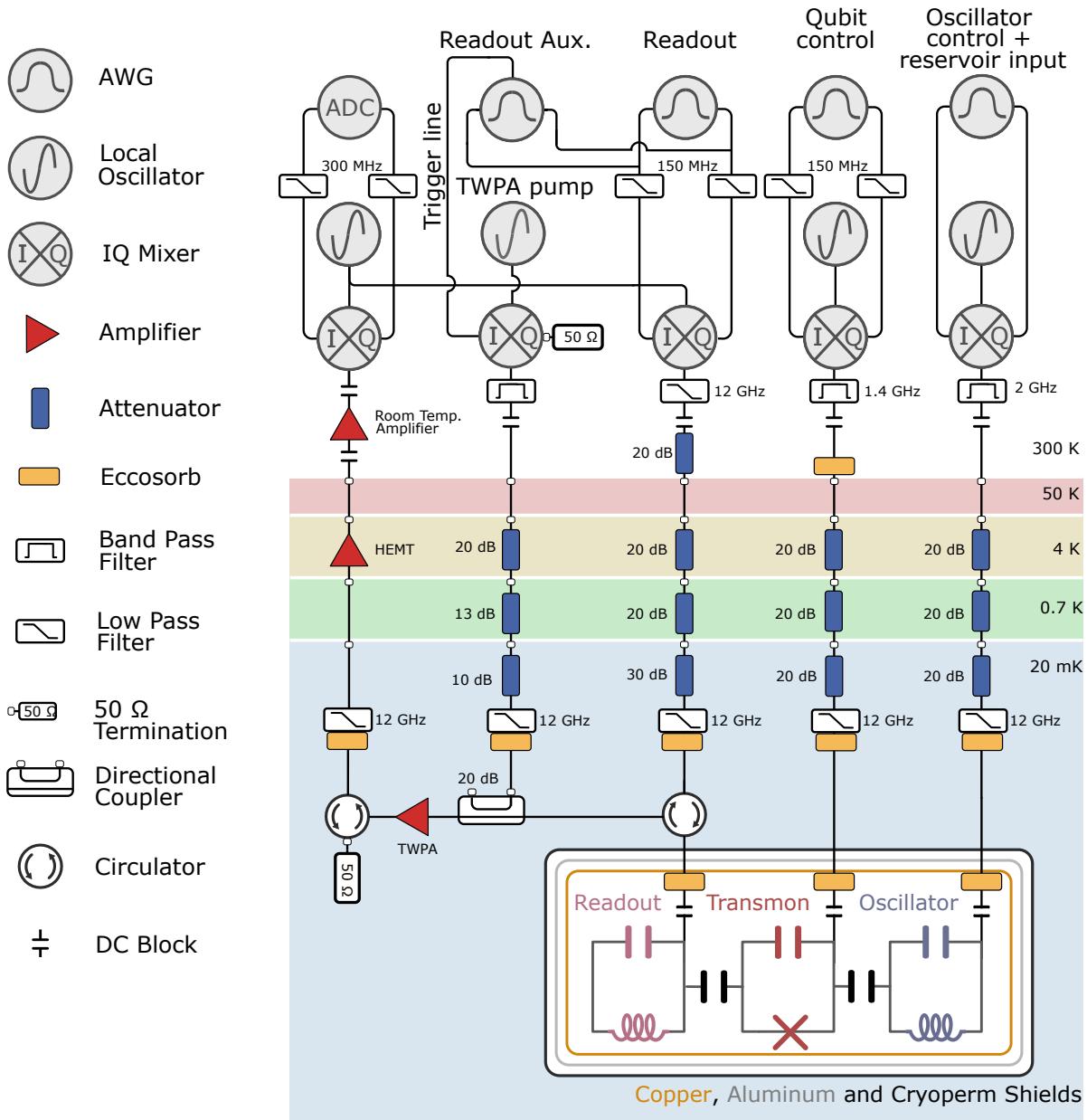


Figure 4.5: **Wiring diagram.** Experimental setup for control hardware, cable routing, and shielding for our device.

by a factor of 40, it keeps the oscillator state thermalized to the fridge rather than the transmission line.

The control pulses for the qubit and the storage are synthesized using Zurich Instruments (ZI) HDAWG, which have a baseband bandwidth of 1 GHz. These are

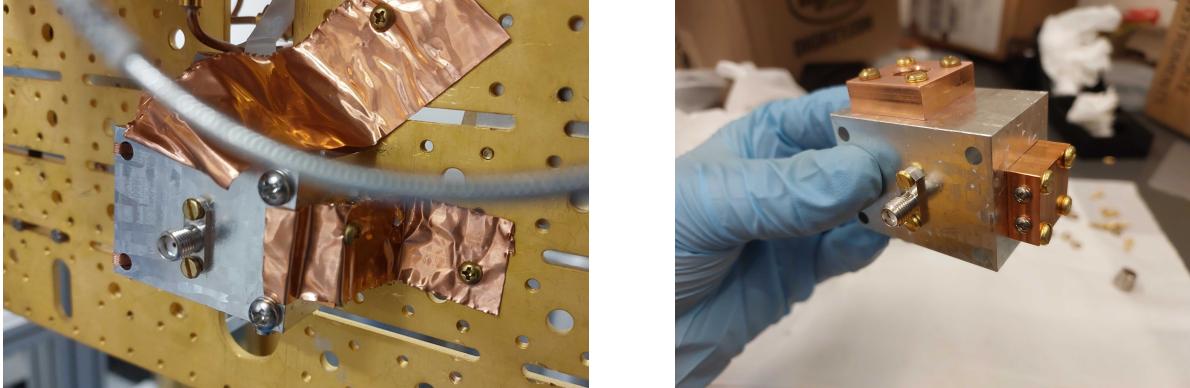


Figure 4.6: Photo of device The device consists of a on-chip transmon and a co-planar waveguide readout resonator mounted inside a high-purity Aluminum cavity. The package is mounted to a gold-plated copper breadboard at the mixing-chamber plate of a dilution refrigerator.

upconverted using Rohde & Schwarz SGS100A, which are signal generators with built-in IQ mixers. These built-in mixers are used for all frequency conversions with the exception of the readout. The readout pulses are synthesized and digitized using ZI UHFQA, and are up-converted and down-converted using Marki mixers (MMIQ-0416LSM-2), with a split LO from a single SGS100A. Readout signals are first amplified with a traveling-wave Josephson Amplifier (TWPA), which is a quantum-limited amplifier. The TWPA typically requires large pump tones, so we gate it with a trigger line from the readout AWG which combines with the CW pump tone in an IQ mixer (as a makeshift fast switch). The readout signals are then further amplified with a High-electron mobility transistor (HEMT) amplifier at the 4K stage, and again amplified with a room temperature amplifier (ZVA-1W-103+ from Mini-Circuits) and filtered. The digitizer on the ZI UHFQA converts to the analog response to a digital signal and integrates it to produce a binary outcome depending on the qubit state.

4.5 System Hamiltonian & Reservoir description

4.5.1 Hamiltonian description

Our transmon-cavity system is well approximated by the Hamiltonian [13]:

$$H/\hbar = \omega_q q^\dagger q + \omega_a a^\dagger a - \chi q^\dagger q a^\dagger a - \chi' q^\dagger q a^{\dagger 2} a^2 - K_q q^{\dagger 2} q^2 - K a^{\dagger 2} a^2 + \Xi(t)(q + q^\dagger) + \xi(t)(a + a^\dagger), \quad (4.2)$$

where a is the annihilation operator for the oscillator mode, and q is the annihilation operator for the qubit mode, ω_a and ω_q are the frequencies of the oscillator and qubit mode respectively, χ and χ' are the dispersive shift and the oscillator state-dependent dispersive shift respectively, K and K_q are the self-Kerr of the oscillator and the transmon anharmonicity respectively. The values for these parameters, as well as values for decay rates, are listed in Table 4.1. The last two terms describe the qubit and oscillator drives in the lab frame. The lab-frame drives are related to the rotating-frame drives in Eq. 4.1 via $\Xi(t) = \Omega(t)e^{i\omega_q t} + \text{H.c.}$ and $\xi(t) = \epsilon(t)e^{i\omega_a t} + \text{H.c..}$ For the design of our drives, we ignore the self-Kerr of the oscillator as well as the higher-order cross-Kerr. We note that these are indeed present, but for the purposes of a quantum reservoir, only add to the complexity of the dynamics. Finally, moving to the rotating frame of the transmon and oscillator mode and truncating to the first two levels of the transmon, we arrive at the Hamiltonian in Eq. 4.1.

4.5.2 Reservoir description for time-independent signals

The advantage of the reservoir computing paradigm is the flexibility in the choice of dynamics. However, simple design principles, motivated by the physics of the

Parameter	Mode(s)	Symbol	Value
Frequency	Transmon g-e	ω_q	$2\pi \times 7.136 \text{ GHz}$
	Oscillator	ω_a	$2\pi \times 6.024 \text{ GHz}$
	Readout	ω_r	$2\pi \times 8.888 \text{ GHz}$
Self-Kerr	Transmon g-e	K_q	$2\pi \times 315 \text{ MHz}$
	Oscillator	K	$2\pi \times 6 \text{ kHz}$
Cross-Kerr	Transmon-Oscillator	χ	$2\pi \times 2.415 \text{ MHz}$
	Transmon-Readout	χ_{qr}	$2\pi \times 1 \text{ MHz}$
Second-order Cross-Kerr	Transmon-Oscillator	χ'	$2\pi \times 19 \text{ kHz}$
Relaxation time	Transmon g-e	T_1	$30 \mu\text{s}$
	Oscillator	T_1^a	$100 \mu\text{s}$
Dephasing time	Transmon g-e	T_2	$25 \mu\text{s}$
Thermal population	Transmon g-e	\bar{n}_{eq}^q	3%
	Oscillator	\bar{n}_{eq}^a	< 0.2%

Table 4.1: **System parameters and dissipation rates.** System parameters were measured using various spectroscopic and time-domain techniques following methods in Ref. [21].

system, can go a long way in engineering a reservoir with high expressive capacity on many tasks. In this section, we provide full details and motivations for the unitaries and measurements in this work, followed by sections outlining characterizations of the device in order to realize the intended dynamics.

The reservoir drives consists of two categories of dynamics: the unitaries and the measurements. In what follows, we will first provide analysis of the dynamics for time-independent input (e.g. the signals in Fig. 4.2). As we will see, the unitary component of the dynamics implemented in this work strives to implement a \cos^2 nonlinearity on the raw input, whereas the measurements generate non-classical features in the state and quantum correlations in the measurement trajectories via measurement backaction.

While measuring the quadratures of some unknown signal is easy with a typical homodyne setup, performing the same measurement of a displacement on an oscillator using only qubit measurements can be non-trivial. Of course, when designing a

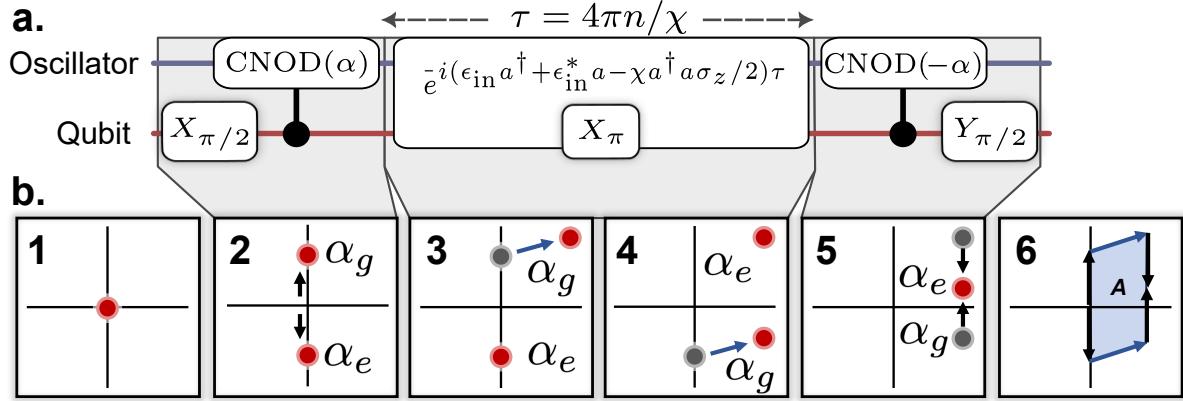


Figure 4.7: Geometric phase unitary to sense phase of unknown displacement (a.) Decomposition of the unitary used throughout the reservoir for time-independent tasks. **(b.)** Schematic representation of the dynamics of the oscillator state under the reservoir drives with time-independent input, highlighting a unitary which implements a geometric phase unitary. (1) At the start of the protocol, the qubit is in the ground state, with the oscillator at vacuum. (2) An initial $X_{\pi/2}$ pulse brings the qubit to the equator. The CNOD(α) unitary conditions the state of the oscillator based on the qubit. For the first reservoir, this is a coherent state. (3) and (4) For time independent inputs, the effective action can be described by a single displacement on the oscillator mode. In this experiment, we operate the displacement at a frequency which, to first order, causes a displacement only on the state conditioned on the ground state of the qubit. A qubit π pulses switches the state in between the two displacements. (5) The final conditional displacement brings the two conditioned states back onto each other. This effectively disentangles the qubit from the oscillator mode. (6) The effective geometrical area enclosed A , which is a function of the input, is imparted onto the qubit.

reservoir, one does not strive to implement the identity, but it is a good starting point – the unitary is thus implemented to approximate the identity. It consists of the input signal data, which is sandwiched on either side by fast conditional displacement gates implemented with CNOD [25] and qubit rotation gates. The broad-overview of the decomposed unitary is given in terms of gates in Fig. 4.7, along with a schematic portrayal of the phase-space trajectory of the oscillator mode initialized in vacuum subject to a time-independent drive.

We begin with an idealized gate-based version decomposition of our reservoir for

time-independent input on resonance with the oscillator conditioned on the qubit being in the ground state. The sequence of gates the reservoir unitary approximates:

$$U_1 = X_{\pi/2} \quad (4.3)$$

$$U_2 = D(\alpha)|g\rangle\langle e| + D(-\alpha)|e\rangle\langle g| \quad \text{CNOT} \quad (4.4)$$

$$U_3 = D(\beta)|g\rangle\langle g| + |e\rangle\langle e| \quad \text{Input} \quad (4.5)$$

$$U_4 = X_\pi \quad (4.6)$$

$$U_5 = U_3 = D(\beta)|g\rangle\langle g| + |e\rangle\langle e| \quad \text{Input} \quad (4.7)$$

$$U_6 = D(-\alpha)|g\rangle\langle e| + D(\alpha)|e\rangle\langle g| \quad \text{CNOT} \quad (4.8)$$

$$U_7 = Y_{\pi/2} \quad (4.9)$$

Ignoring the very first unitary, after applying the sequence of unitaries U_2 through U_7 , we arrive at unitary

$$U_7U_6U_5U_4U_3U_2 = \frac{i}{\sqrt{2}}e^{(\alpha\beta^* - \alpha^*\beta)}D(\beta)(|g\rangle\langle g| - |e\rangle\langle g|) - \frac{i}{\sqrt{2}}e^{(-\alpha\beta^* + \alpha^*\beta)}D(\beta)(|g\rangle\langle e| + |e\rangle\langle e|) \quad (4.10)$$

Let

$$|\psi\rangle = [e^{-i\phi/2}\cos(\theta/2)|g\rangle + e^{i\phi/2}\sin(\theta/2)|e\rangle] \otimes |\text{cavity}\rangle \quad (4.11)$$

be some arbitrary initialized state. Then for $\theta = \pi/2$, we have

$$U_7U_6U_5U_4U_3U_2|\psi\rangle = \frac{1}{\sqrt{2}}D(\beta)[i\sin(A - \phi/2)|g\rangle + \cos(A - \phi/2)|e\rangle] \otimes |\text{cavity}\rangle, \quad (4.12)$$

where $A = 2|\alpha||\beta|\sin(\delta) = i(\alpha\beta^* - \alpha^*\beta)$ is the geometric phase enclosed by the oscillator trajectory which is dependent on the phase difference δ between the known displacement $D(\alpha)$, and the unknown displacement $D(\beta)$ (Fig 4.7b). Thus, for the proper qubit state before the application of $U_2 \dots U_7$, we are able to extract information about the phase of the displacement. We also note that the qubit and the oscillator are disentangled after the unitary, and that the effect of the unitary on the oscillator mode is a simple displacement.

Finally, pre-pending U_1 (Eq. 4.3) to the string of unitaries guarantees that we initialize our qubit state with $\theta = \pi/2$ when following a qubit measurement, independent of that measurement outcome. It also guarantees $\phi = \pi/2$ or $3\pi/2$ depending on the measurement outcome. The probability of measuring the qubit in the excited state conditioned on preparing it e vs g after the entire sequence is then:

$$P_{e|g} = \cos(A - \pi/4)^2 \quad P_{e|e} = \sin(A - \pi/4)^2 \quad (4.13)$$

Thus, with this sequence of unitaries, we are able to extract the phase of some unknown displacement (relative to some known displacement α) by simply measuring the qubit. While for the first run of the reservoir, the qubit will start in the ground state (up to thermal noise), after performing a parity measurement, the qubit state will depend on the previous measurement outcome. See Fig. 4.11 for an experimental implementation of the above results.

In principle, Eq. 4.13 enables us to perform the identity operation on the input x, y points followed by a \cos^2 kernel. Without loss of generality, we take $\arg(\alpha) = 0$, then $i(\alpha\beta^* - \alpha^*\beta) = \text{Im}(\beta) = \beta_x$. Alternating between $\arg(\alpha) = 0$ and $\arg(\alpha) = \pi/2$ allows us to extract $\cos^2(\vec{\beta})$ with two runs of the reservoir.

Whereas all gates besides the input (Eqs. 4.5 and 4.7) are fast and therefore insensitive to the cross-Kerr interaction, the primary deviation from the gate description occurs for the input, which can be very long. This input displacement is conditioned on the qubit being in the ground state. Therefore, in the rotating frame of the qubit-oscillator system, the branch of the oscillator state conditioned on the qubit being in the excited state will rotate at a frequency χ , which in general will break the geometric phase construction that works for time-independent tasks. Therefore, we limit the exposure time of the reservoir to the input signal to be an integer multiple of $4\pi/\chi$, so that the oscillator state conditioned on the qubit being in the excited state will return

to the same point.

The unitary described in Eqs. 4.3-4.9 is followed by a qubit measurement, then a parity measurement Π [50, 123] with projectors P_{\pm} , where

$$\Pi = (-1)^{a^\dagger a} \quad P_{\pm} = \frac{1}{2}(1 \pm \Pi) \quad (4.14)$$

As mentioned above, the effect of the unitary on the oscillator state for time-independent signals is simply a displacement of the input data $D(\beta)$, independent of the qubit measurement outcome. For the following discussion, we will ignore the qubit dynamics, since the qubit and the oscillator are disentangled at the end of the unitary. In effect, the state of the oscillator can be described by a series of alternating displacements and parity measurements:

$$|\text{cavity}\rangle = \dots P_{p_4} D(\beta) P_{p_3} D(\beta) P_{p_2} D(\beta) P_{p_1} D(\beta) |0\rangle, \quad (4.15)$$

where P_{p_n} is the projector of the n th parity measurement with outcomes $p_n = \{+, -\}$. For k runs of the reservoir, we can reorder terms and add pairs of canceling displacements $D(-\beta)D(\beta)$ to rewrite the above as

$$|\text{cavity}\rangle = \left(\prod_n^k P_{p_n}^{n\beta} \right) D(k\beta) |0\rangle. \quad (4.16)$$

Equation 4.16 describes a series of projective measurements after preparing a displaced vacuum state. The projectors and their associated measurements are

$$P_{\pm}^{\alpha} = D(\alpha) P_{\pm} D(-\alpha) \quad \Pi^{\alpha} = D(\alpha) \Pi D(-\alpha) \quad (4.17)$$

The measurements Π^{α} describe parity measurements in displaced frame at α . Incidentally, the expectation value of this operator are proportional to the Wigner function at α [107]. However, importantly, Eq. 4.16 *does not* describe performing Wigner tomography of the state $D(k\beta)|0\rangle = |k\beta\rangle$ at points given by $\beta, 2\beta, 3\beta, \dots$, as the effective measurements Π^{α} do not commute for different values of α . Instead, in general $[\Pi^{\alpha}, \Pi^{\gamma}] \neq 0$. Therefore, in this light, our reservoir construction can be seen to

leverage non-commuting measurements and quantum contextuality to generate conditional and correlated probabilities over measurement trajectories.

4.5.3 Reservoir description for slow varying time-dependent signals

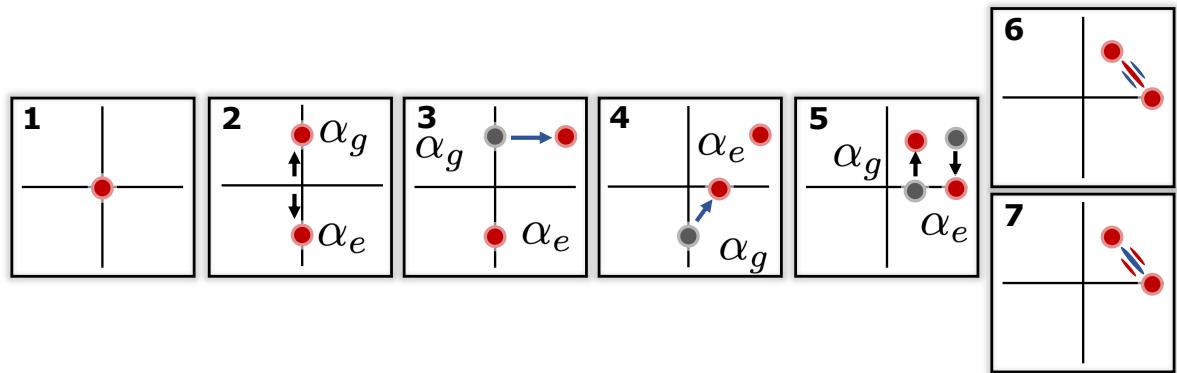


Figure 4.8: **Schematic description of the dynamics of the oscillator, starting in vacuum, for a slow time-varying input.** In such a scenario (as is the case for the task of classifying radio modulation schemes), the signal causes a displacement largely conditioned on the ground state. Generally, the value of the displacement is different before and after the π pulse in between (3) and (4). Unlike the regime for time-independent signals, there is no effective area enclosed in phase space, which leaves the qubit entangled with the oscillator. (6) and (7) describe the state of the oscillator after the qubit measurement. The resulting state of the oscillator is a cat state, where the parity of the cat dependent on whether the outcome of the qubit measurement is ground or excited.

For generic, time-dependent signals, like those classified in Figs. 4.3 and 4.4 in the main text, the geometric unitary described by Eqs. 4.3-4.9 does not in general hold, as the symmetry between panels 3 and 4 in Fig. 4.7 is broken. Additionally, the approximation that the input is displacement conditioned on the qubit in the ground state (Eq. 4.5 and 4.7) will not hold for high bandwidth signals, like those in Fig. 4.4 in the main text. For high-bandwidth signals, the input will also have some contribution

in displacing oscillator conditioned on the qubit being in the excited state, which can lead to complex dynamics in the oscillator. While for generic signals, this can be hard to describe, here we prove a treatment of our reservoir construction for slowly-varying, time-dependent signals, like those in Fig. 4.3 of the main text.

We can follow most of the derivation from the scenario of time independent signals in Section 4.5.2, to describe the dynamics of the QRC for the task of classifying radio frequency modulation schemes. Along with the assumptions in the previous section, we make the slow-varying input approximation, such that the displacement on the oscillator of the reservoir is still effectively conditioned on the ground state of the qubit. The displacement on the oscillator depends on the value of the symbol encoded for the given modulation scheme. Since, in general, the symbol is different before and after the qubit π pulse: the direction of the displacement in the oscillator will be different. Given the timescales of the input signal involved, this essentially corresponds to a displacement on the oscillator conditioned on the ground state of the qubit. When the two displacements are different in magnitude and direction, the qubit remains entangled with the oscillator at the end of the reservoir unitary. The state of the system just before the measurements is (step (5) of Fig 4.8:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(e^{-iA_i}D(\beta_i)|g, \text{cavity}\rangle + e^{iA_j}D(\beta_j)|e, \text{cavity}\rangle), \quad (4.18)$$

where β_i is the displacement before the qubit flip, and β_j is the displacement after. $A_i = 2\text{Im}[\alpha\beta_i^*]$ is the phase acquired after two non-orthogonal displacements. When $\beta_i = \beta_j$, we recover the dynamics for time independent signals. It is straightforward to show that the qubit will be disentangled from the oscillator and that the area A_i , corresponding to the geometrical phase form the area enclosed in phase space will be present as a relative phase difference between the ground and excited state. After a

$Y_{\pi/2}$ gate, we have the following state in our system:

$$|\psi\rangle = \frac{1}{2}[e^{iA_i}D(\beta_i) + e^{-iA_j}D(\beta_j)]|g, \text{cavity}\rangle + \frac{1}{2}[e^{iA_i}D(\beta_i) - e^{-iA_j}D(\beta_j)]|e, \text{cavity}\rangle, \quad (4.19)$$

One can think of this as a cat state in the cavity, with a parity determined by the qubit state. This is schematically shown in (6) and (7) in Fig 4.8. In the limit of very different displacements, the probability of the qubit measurement is the same for both ground and excited states. The goal of this task can be thought of as discriminating probability distribution functions over the (I, Q) plane. Fig 4.3 (a) represents the so-called “constellation” diagram of the modulation schemes considered in this work. Each scheme can take discrete values in (I, Q) space, with even equal probability (we construct the dataset of radio signals encoding random binary strings). Our lack of knowledge of the exact displacement on the oscillator can be mathematically expressed as a density matrix. This is the most apparent in the state of the oscillator after the initial qubit measurement,

$$\rho'_{\text{cavity}} = \sum_{\beta_i \in P} p_i D(\beta_i) \rho_{\text{cavity}}^\dagger D(\beta_i), \quad (4.20)$$

where ρ'_{cavity} is the density matrix representation of the cavity right after the qubit measurement and ρ_{cavity} describes the initial density matrix before the application of the protocol. The set P describes the distribution of possible displacements which can be received from the input. p_i is probability for receiving the symbol corresponding to a displacement β_i . For the task considered in this work, these distributions are uniform, with no contributions from conditional probabilities. However, this description of the reservoir motivates the potential for the QRC to distinguish signals with complex correlations in the symbols of the message encoded.

4.6 Quantum reservoir characterization

4.6.1 CNOD

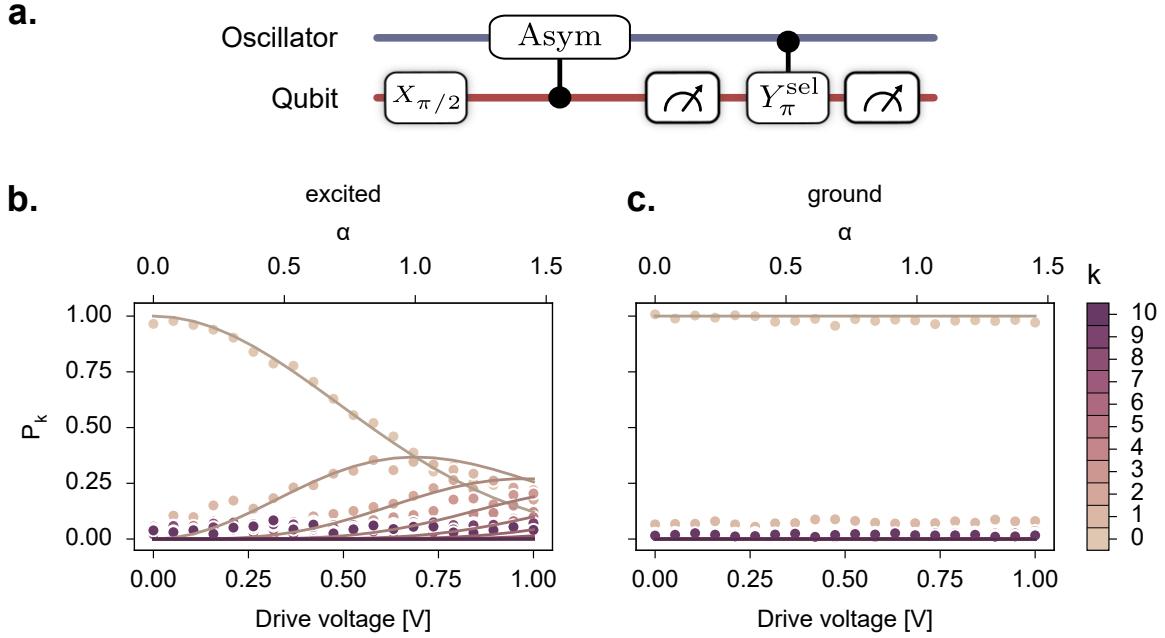


Figure 4.9: **Characterization of conditional displacements using an Anti-symmetric pulse from Ref. [25].** (a.) Pulse control schematic of characterization of conditional displacement. Here, we make use of number-splitting spectroscopy to characterize the state of the cavity after perform a conditional displacement on the state $|+\rangle|\text{vacuum}\rangle$ conditioned on the qubit state. By post-selecting on the qubit state, we can evaluate the effectiveness of the conditional displacement. (b.) Number splitting spectroscopy conditioned on measuring the qubit in the excited state. A single parameter fit is used to capture the behavior of the state of the cavity as a function of amplitude. From the good agreement, we conclude that conditioned on measuring the qubit in the excited state, the cavity is displaced. (c.) Number splitting spectroscopy conditioned on measuring the qubit in the ground state. We see very limited change in the cavity state when measuring the qubit in the ground state.

Here, we provide the calibration of the CNOD unitary [25], one of the components of our reservoir unitary (Fig. 4.7). The CNOD protocol implements the following

unitary

$$\text{CNOD}(\alpha) = D(\alpha)|g\rangle\langle e| + D(-\alpha)|e\rangle\langle g|. \quad (4.21)$$

The protocol is implemented with two ‘Anti-symmetric pulses’ sandwiching a qubit pi-pulse. In the frequency domain, the pulse is composed of two gaussian envelopes offset such that there is a zero-crossing at the qubit ground state frequency, and that the spectrum is anti-symmetric around this point (see Ref. [25]). The Anti-symmetric pulse is a conditional displacement, conditioned on the qubit being in the excited state. The motivation for using CNOD instead of a single tone displacement on resonance with the stark-shifted qubit frequency is that it enables the ability to perform conditional displacements at time scales much smaller than $2\pi/\chi$.

Figure 4.9a displays the protocol for characterizing the anti-symmetric pulse. First, the qubit is unconditionally brought to the equator of the bloch sphere, with a wide-band $X_{\pi/2}$ pulse. After this, the anti-symmetric pulse acts on the cavity, followed by an qubit measurement, collapsing the cavity state to either $D(\alpha)|0\rangle$ or $|0\rangle$. After collapsing the state, we perform a number-splitting spectroscopy on the cavity. This is performed with a conditional Y_π , conditioned on the k th cavity Fock state [40, 112] followed by a second qubit measurement. By post-selecting on the first qubit measurement outcome, we can characterize the cavity state for each branch. Figure 4.9b and c show the number-splitting spectroscopy for the cavity state conditioned on the qubit being in the ground state vs excited, as a function of pulse amplitude. These curves are fitted with a single parameter scaling parameter that defines the relationship between pulse amplitude voltage and the amount of displacement α .

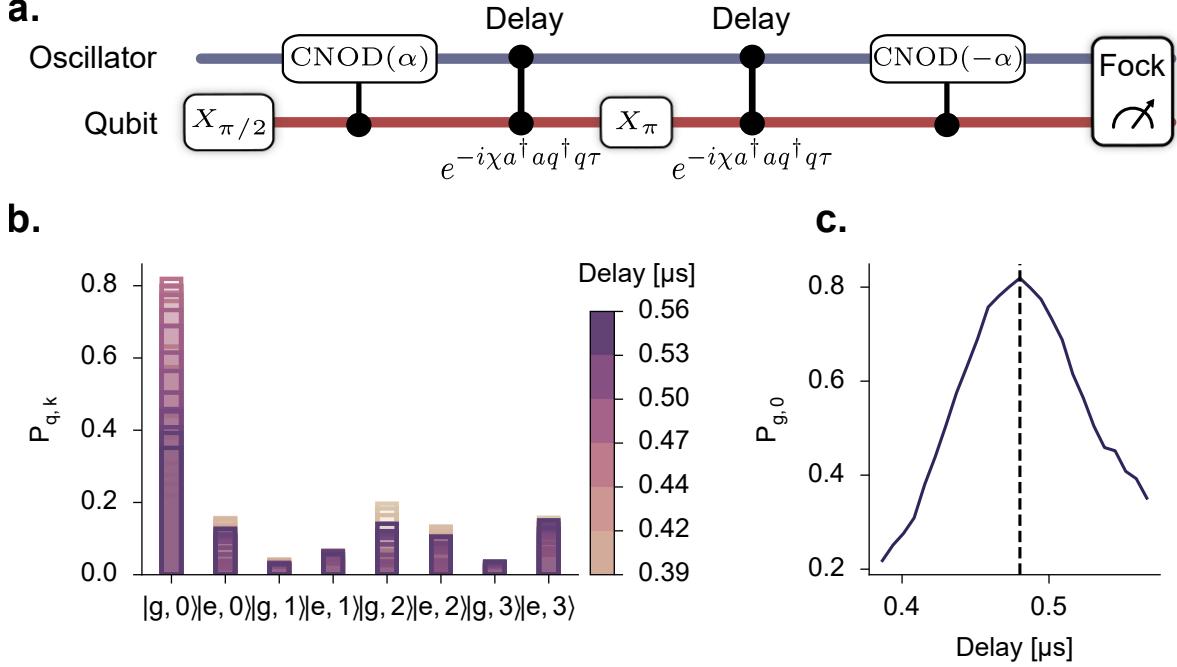


Figure 4.10: **Characterization of input signal duration** (a.) Diagram of control sequence to calibrate length of input signal duration toward implementing the geometric phase unitary in Eqs. 4.3-4.9. Here, a double conditional displacement is performed after sending the qubit to the equator of the Bloch sphere. After this, a variable delay is added before undoing that displacement. Finally, the qubit and the Fock distribution of the cavity is sampled using methods from Refs. [119, 24]. (b.) Overlapping histograms showing Fock distribution of cavity state conditioned on qubit state as a function of delay. (c.) Cavity state overlap with the vacuum state as a function of delay. At a particular value of the delay, the two displacements interfere and cancel each other out.

4.6.2 Reservoir unitary characterization

With our rotation gates and CNOD's calibrated, we describe in this section the calibration of signal drives toward the implementation of Eqs. 4.3-4.9. We begin with a calibrating the duration of time our reservoir is exposed to the input signal. As discussed in Section 4.5, calibrating this delay is crucial for a faithful implementation of the geometric phase detection unitary introduced in this work. While it may seem that this restriction in the input signal duration is contrived in a real-world setting

where the signal is unknown, this restriction would be implemented via a fast switch that exposes our device to the unknown signal periodically.

Figure 4.10a schematically describes the experimental protocol for calculating the delay between the two CNOD pulses. Here, we effectively try to undo a double conditional displacement via second double conditional displacement. Due to the dispersive shift, after the first conditional displacement, the state of the cavity conditioned on the excited state of the qubit will start rotating with respect to the state of the cavity conditioned on the ground state. After a period of $2\pi/\chi$, the this will return to the same position as the start. Undoing the displacement at this point in time will send the cavity state to vacuum. Figure 4.10b shows the Fock distribution of the cavity as a function of the waiting time, and Fig. 4.10c shows the cavity state overlap with the vacuum state as a function of the waiting time.

Next, we implement the full unitary given by Eqs. 4.3-4.9, where the section corresponding to the input data displacement (Eqs. 4.5 and 4.7) is given by the duration found in the results above. For this calibration, we implement the full unitary given by the diagram in Fig. 4.7a by varying the angle of the input displacement and looking at the dependence.

Figure 4.11a shows the schematic overview of the calibration procedure. The geometric phase unitary is parameterized by a long displacement, whose angle we sweep. After the unitary we perform a qubit measurement, followed by a parity measurement. This calibration experiment is essentially identical to the time-independent reservoir computing experiments in terms of the control protocol. Here, instead of sending data from different distributions for the system to classify, we only vary the phase and amplitude of some input displacement to get the phase dependence we want.

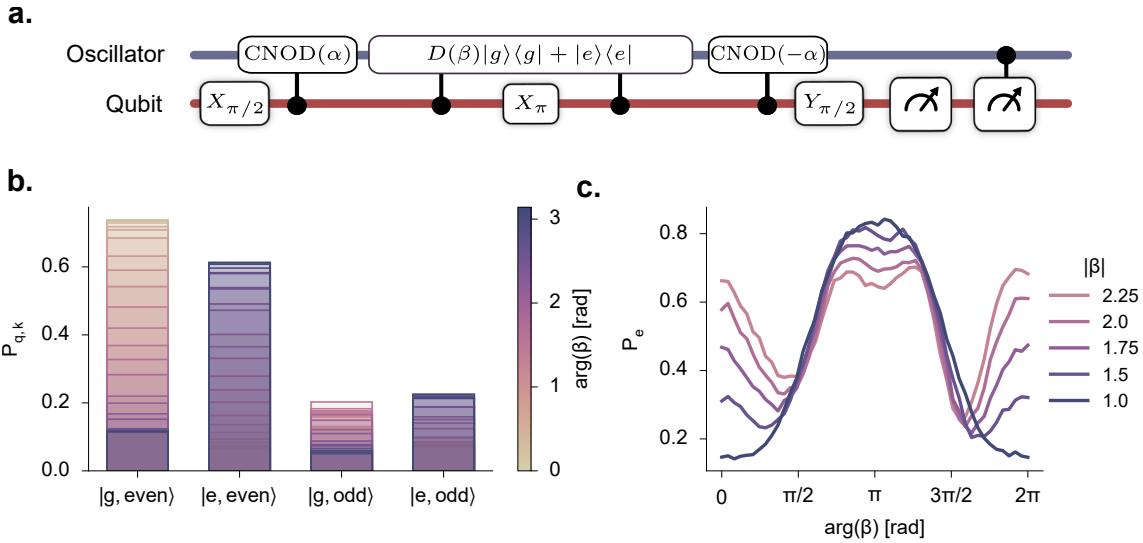


Figure 4.11: Full geometric phase unitary calibration. (a.) Pulse control protocol for the unitary defined by Eqs 4.3-4.9. This protocol is identical to that of Fig. 4.7a. (b.) Overlapping histograms showing the probability over outcomes of measuring the qubit and oscillator-parity as a function of the phase of the input signal with $|\beta| = 0.25$. As the phase is varied, the probability of the qubit being excited increases. (c.) The probability of measuring the qubit excited as the phase of the input displacement β is varied, plotted for different values of $|\beta|$

Figure 4.11b shows the distribution of measurement outcomes from measuring the qubit and the oscillator parity after the unitary is applied with $\alpha = 1$ and $\beta = 0.25$. As the angle of the input is swept, the qubit probability of the qubit being found in the ground state shifts to being found in the excited state. This is more evident in Fig.4.11c where we plot the probability of measuring the qubit in the excited state P_e as a function of the phase of β for different amplitudes of β . In comparison we find great qualitative agreement with the expected result $P_e = \cos(2|\alpha||\beta| \cos(\delta) + \pi/4)^2$, where $\delta = \arg(\alpha) - \arg(\beta)$ (see Eq.4.13), though we find an extra reduction in the dynamic range in P_e for increasing β due to qubit overheating.

For our quantum reservoir tasks, we choose α to be quite small, near 0.2. The effect of this is a severe reduction in the dynamic range of P_e , but one that is easily distinguishable at 1000 shots. For all of our tasks, this was the minimum number of

shots needed to get 100%. Keeping $|\alpha|$ small allows for a greater sensitivity in $|\beta|$ without worrying about qubit overheating.

4.6.3 Qubit & parity measurements

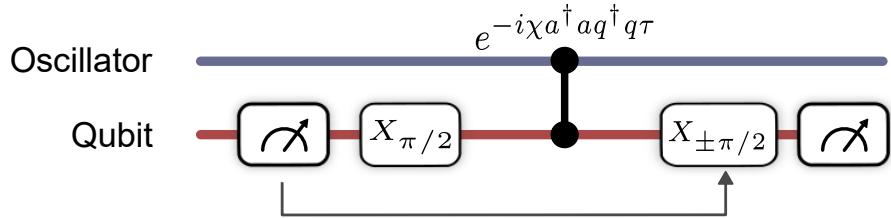


Figure 4.12: **Qubit and oscillator-parity measurements in the quantum reservoir computer.** For our reservoir construction, the state of the qubit is not generally known before a parity measurement. We apply feedback to change the parity measurement based on the preceding qubit measurement to faithfully capture the oscillator parity.

The qubit and parity measurements performed in this work are the standard pulse schemes used in many previous works, with one change. The typical procedure of measuring the parity of a cavity state is similar to a Ramsey experiment (and perhaps more closer still to a ‘qubit-revival’ experiment [21]), and importantly requires knowledge of the state of the qubit before the measurement is performed. In a quantum reservoir setting where measurement trajectories can be unknown, measuring the parity of the cavity is not straight-forward without post-selection or feedback. Here, since we perform a qubit measurement just before the parity measurement, we apply simple feedback that conditions the parity unitary on the measurement outcome of the preceding qubit measurement. The condition is such that the parity measurement outcome is now independent of the preceding measurement outcome. This reduces the order of correlations required to gain the same information: attaining the parity of the cavity only requires information about the parity measurement, whereas previously, second-order correlations between the qubit and parity measurement was

required. A further refinement to reduce trivial correlations in the measurement history would reset the qubit after the oscillator parity, however, due to limitations in the FPGA software, this was not implemented.

4.6.4 Tuning T_2 via resonator-induced dephasing

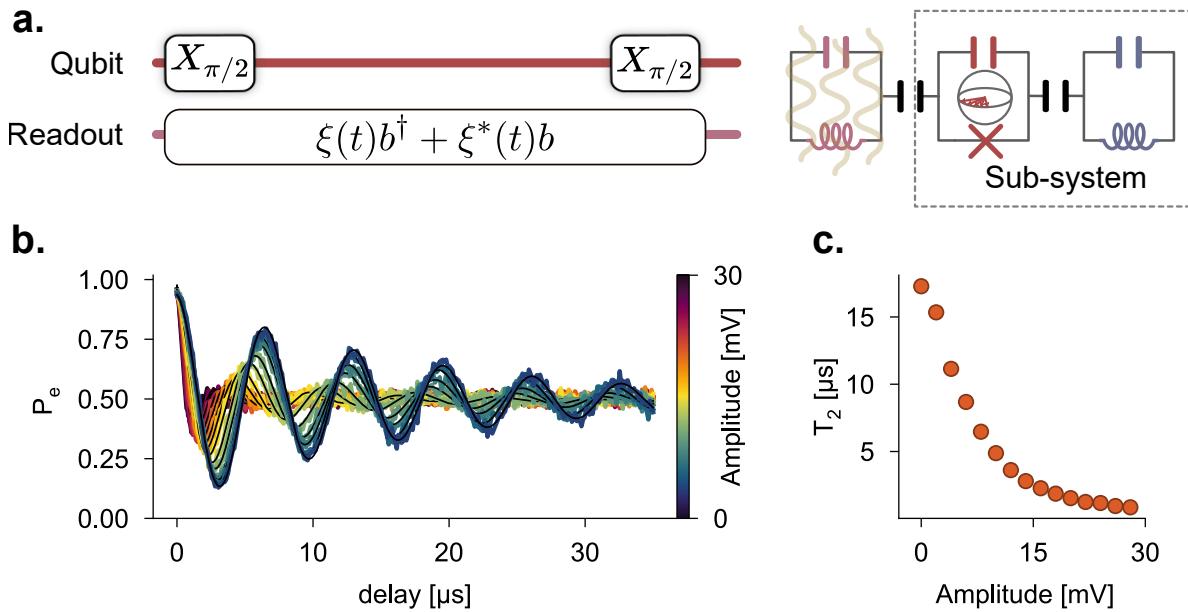


Figure 4.13: **Resonator-induced dephasing via pumping on the readout resonator** (a.) Protocol for characterizing the effect of a readout pump on qubit decoherence: we perform a typically Ramsey experiment while populating the readout resonator to measure the T_2 . (b.) Ramsey curves as a function of readout pump amplitude. These curves are fit using Eq. 4.22 to produce estimates of the qubit coherence. (c.) Extracted T_2 values for each of the curves in part (b)

Here we describe the experiment to reduce the qubit coherence time by pumping the readout resonator with photons during our reservoir experiments (see Fig. 4.2d). The calibration of this experiment involves performing a standard Ramsey T_2 experiment, modified with a pump on the readout resonator (Fig. 4.13a). Once populated, the resonator photons induce a dispersive shift, which sends the qubit to the center

of the Bloch sphere once the readout resonator is traced out. In principle, this interaction is coherent, and the qubit should see a revival. However, due to the leaky nature of the readout resonator by design, a coherent revival is not observed. As remarked at the end of Section 4.4, this experiment required an auxiliary AWG line. Figure 4.5 denotes this as the ‘Readout Auxiliary’ line.

Figure 4.13b shows the results of the Ramsey calibration with the readout pump on, for varying pump powers. We see a steady decrease in the qubit coherence time as the pump amplitude is increased as expected. The curves are fit to the equation

$$P_e = \cos(2\pi\delta t)e^{-t/T_2}, \quad (4.22)$$

where δ is an intentional detuning. Here, a Gaussian pulse was used as the readout pump. We expect that due to the construction of the reservoir, a flattop pulse may be more detrimental to the classification performance, since the Gaussian pulse has little amplitude during the CNOD unitaries shown in Fig. 4.7a. Finally, we note that the maximum T_2 shown in Fig. 4.13 differs from the value quoted in Table 4.1. After preliminary calibration data corresponding to those in Fig. 4.13, the experiments in Fig. 4.2c were performed, after which the qubit T_2 was suddenly lowered. However, all experiments presented in this manuscript, with the exception of Fig. 4.13, were performed where the qubit T_2 matched that of Table 4.1. Given the conclusion that the qubit T_2 does not impact classification accuracies until it approaches the time between measurements, we decided to include the higher quality data presented in Fig. 4.13, rather than the preliminary data used to calibrate the results in Fig. 4.2c.

4.7 Machine learning with the quantum reservoir

4.7.1 Output feature encoding

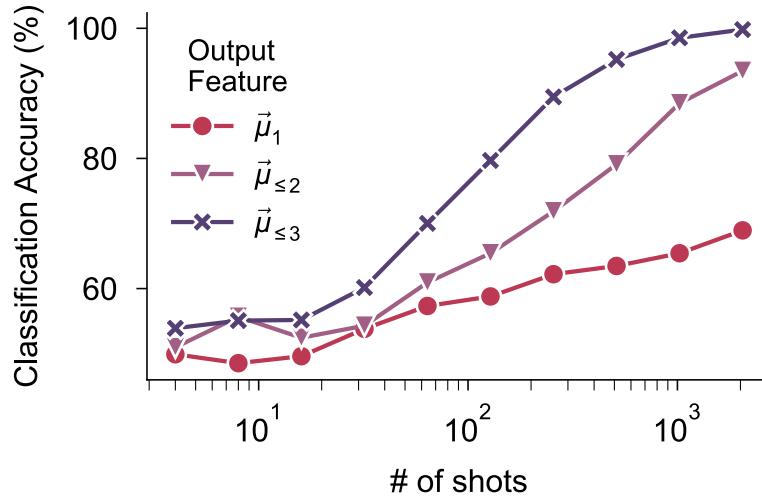


Figure 4.14: **Comparison of feature vectors in spiral classification performance**
Here, classification accuracy on the spiral task is considered for different output encodings. Particularly, we compare included higher and higher correlations. $\vec{\mu}_{\leq p}$ describes a feature vector containing all central moments up to and including the p -th central moment (see text).

In this work, we use measurement correlations as the output feature vectors from which the trained linear layer of our reservoir performs the classification. In this section, we provide details in how these were constructed from measurement results, as well as motivations and comparisons with other output encodings. As described in the main text, measurements of our reservoir involve two measurements following every data input: a qubit measurement and a parity measurement. The qubit measurement, which follows just after the input unitary, either extracts information about the input displacement (if the signal is time-independent), or performs some nontrivial back-action on the oscillator state (see Fig. 4.8). The parity measurement,

which follows the qubit measurement, will simply measure the parity of the cavity state post-qubit measurement, and collapse the oscillator state to either even or odd Fock states. It is worth pointing out that measurements of the parity are done with an entangling unitary starting with a known qubit state and then performing a regular qubit measurement (see Section 4.6.3 for details).

In this manuscript, qubit measurements are performed using standard dispersive readout, which we review here, since the process involves a number of nonlinear steps (for a thorough review, see Ref. [13]). Each measurement outcome is the result of integrating a response signal from the readout resonator, and is defined by a single point on the $I - Q$ plane. For sufficiently strong coupling between the readout resonator and the qubit compared with the resonator linewidth, the set of all possible integrated IQ points will form two (or more) localized and well-separated blobs, indicating projective measurement with single-shot fidelity. These two (or more) blobs correspond to different states of the transmon, and single-shot fidelity refers to the ability to discern the state of the qubit using only one readout pulse. With knowledge of the location of these blobs, and which state they correspond to, we perform a threshold the measurement result to either ‘0’ or ‘1’, indicating the qubit ground state or excited state respectively.

From a string of binary measurement outcomes, or bitstring, we form our feature vectors by first calculating the p -th central moment μ_p , defined as

$$(\mu_p)_{ijkl\dots} = \frac{1}{N_{\text{shots}}} \sum_n^{\text{N}_{\text{shots}}} (x_{ni} - \langle x_i \rangle)(x_{nj} - \langle x_j \rangle)(x_{nk} - \langle x_k \rangle)(x_{nl} - \langle x_l \rangle) \dots, \quad (4.23)$$

where the number of indices of μ_p is equal to p . Here x_{ni} is the n th repeated measurement result of observable x_i . In our setting, i labels the i -th measurement in a sequence of correlated measurements before the system is reset. The expectation value $\langle x_i \rangle$ is taken over the shots N_{shots} – counting the number of system resets and

repetitions. Faithful estimates of these moments typically require on the order of 1000 shots for the results presented in this manuscript.

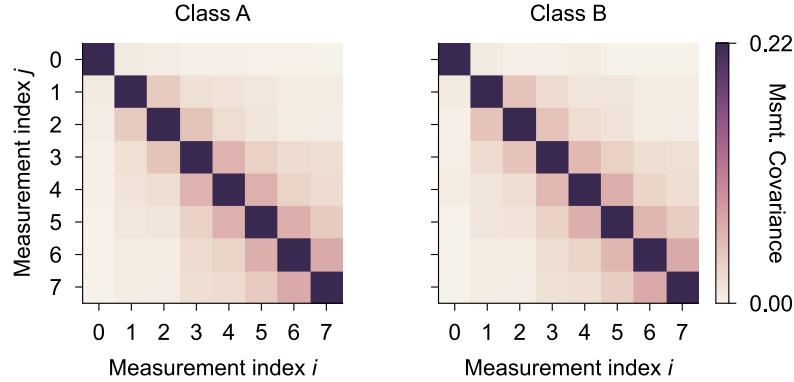


Figure 4.15: Second-order central moment (covariance) of quantum reservoir output over spiral dataset. These correlation matrices were generated from calculating the covariance over measurement outcomes in a reservoir run, then averaged over the entire dataset.

The central moments of Eq. 4.23 are used in the construction of the output feature vector for the linear layer to perform the classification task. Specifically, the feature vector is generated by appending successively more and more central moments. We denote these appended feature vectors as $\vec{\mu}_{\leq p}$ for feature vectors containing up to p central moments, e.g.

$$\begin{aligned}\vec{\mu}_{\leq 2} &= [\vec{\mu}_1, \mu_2] \\ &= [\langle x_0 \rangle, \langle x_1 \rangle, \langle x_2 \rangle, \dots, \langle x_0 x_1 \rangle - \langle x_0 \rangle \langle x_1 \rangle, \langle x_0 x_2 \rangle - \langle x_0 \rangle \langle x_2 \rangle, \dots, \langle x_1 x_2 \rangle - \langle x_1 \rangle \langle x_2 \rangle, \dots]\end{aligned}$$

is a feature vector constructed from appending the flattened covariance to the mean. The first-order moment here is a vector to denote we take the mean over repetitions of different measurements, whereas the covariance is a matrix and thus is not denoted as a vector. Additionally, we only take the independent degrees of freedom of the symmetric covariance matrix equivalent to discarding one of the following redundant

elements $\langle x_i x_j \rangle$ and $\langle x_j x_i \rangle$ for some integers i, j . In general, for arbitrary moments, the number of independent components for M measurements is $\binom{M+p-1}{p}$, where p is the order of the moment. For up to third-order central moments of $M = 8$, this gives a total output feature dimension of $\dim(\vec{\mu}_{\leq 3}) = 8 + 36 + 120 = 164$. This output dimension for the results presented in the main text is further reduced as discussed in the following paragraphs below.

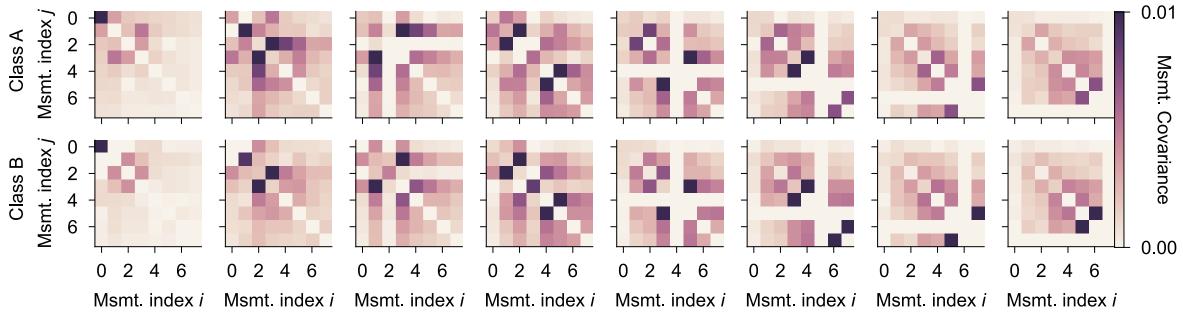


Figure 4.16: Third-order central moment of quantum reservoir output over spiral dataset. These third-order central moments are plotted as an array of 2D matrices, such that the i -th column corresponds to the 2D matrix μ_{ijk} . The two rows corresponds to the two different classes of signals. In the third order, one can begin to see differences between the two classes by eye.

Figure 4.14 contains classification results on the spiral dataset (Fig. 4.2) as a function of the number of shots for the feature vectors $\vec{\mu}_1, \vec{\mu}_{\leq 2}$ and $\vec{\mu}_{\leq 3}$. We see that our quantum reservoir has non-trivial third-order correlations and that the reservoir leverages these correlations to boost classification accuracy. The covariance matrix averaged over the entire spiral dataset is plotted in Fig. 4.15, and the third-order correlations are plotted in Fig. 4.16 – plotted as a set of 2D matrices. In the third-order correlations in particular, we can begin to pick out by eye the differences in the two classes.

This construction generally allows us to construct feature vectors that are smaller than the probability distribution over all possible measurement trajectories, which is

2^M dimensional. However, as can be seen in Fig. 4.15, there is yet redundant information even after taking only the symmetric part - specifically, that the information tends to be very local and that measurements far apart tend not to be correlated. This has the physical interpretation that while measurements are indeed correlated, even possessing higher-order correlations, this correlation tends to be local due to the finite memory of the system. This motivates us to further restrict our feature vector to only capture the essential local correlations.

Figure 4.17 compares the classification performance of feature vectors generated with up to third-order moments, where we truncate the locality of the correlations. That is, the elements of the third order central moment $(\mu_3)_{ijk}$ are set to zero if $|i-j| > d_H$ or $|i-k| > d_H$, for some integer d_H . We note that including third-order correlations between measurements that are up to three ‘sites’ away nearly reproduces the classification accuracy of when you include all third-order central moments. Additionally, we compared the construction of feature vectors using truncated moments up to third-order with that of using the full sampled distribution as the feature vector and found that the former performed much better (Fig. 4.14). These last two statements were found to be true for all tasks presented in this paper. For $M = 8$ measurements, the truncation of long range correlations further reduces the output feature size from 164 down to 94.

4.7.2 Training the linear layer

The only component of the reservoir that was trained to fit the dataset processed by the reservoir was the linear layer applied to the feature that the physical reservoir produced. The linear layer was an $R \times C$ matrix W_{train} and C -dimensional vector v_{train}

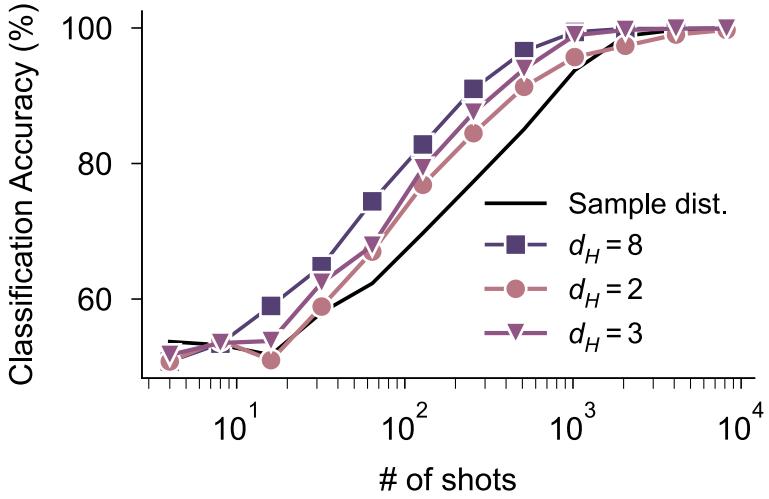


Figure 4.17: Spiral classification accuracy as a function of locality of up-to third-order correlations Classification accuracy for the spiral task using up to third-order central moments truncating correlations to only include correlations up to d_H . We find that we can achieve high accuracy using local third-order correlations, saturating the accuracy when only keeping correlators up to $d_H \leq 3$. The performance of using the sample distribution is also compared which performs worse than using the central moments as output features, despite containing more information (see text).

applied to the R -dimensional reservoir feature x to get

$$y = x^T W_{\text{train}} + v_{\text{train}} \quad (4.24)$$

where the largest of the C elements of y corresponded to the predicted class of the data point (C is the number of classes). To train the linear layer, we chose between two different approaches: the pseudo-inverse method and back-propagation through a softmax function on the output. The two approaches optimize the linear layer over different loss landscapes. This is because our classification method is fundamentally discrete - i.e. we identify the class simply based on whichever output vector entry is the largest - so there is not a perfect correspondence between our loss and the classification inaccuracy.

First, we will describe the pseudo-inverse method. Let X be a $N \times (R + 1)$ matrix

consisting of R -dimensional reservoir features generated for N training points, with a column of 1's appended (this is to compute both W_{train} and v_{train} at once). Let Y be an $N \times C$ matrix consisting of C -dimensional row vectors that serve as labels for the training points such that $Y_{i,j} = 1$ if j corresponds to the class of the i^{th} training point and zero otherwise. For an $\epsilon > 0$, we construct W'_{train} (W_{train} appended with v_{train}) as :

$$W'_{\text{train}} = (X^T X + \epsilon I)^{-1} X^T Y \quad (4.25)$$

In our case, the value of ϵ was swept to maximize the accuracy of the classification. In the limit of $\epsilon \rightarrow 0$, the pseudo-inverse matrix of Eq. 4.25 is provably optimal for minimizing $\|XW_{\text{train}} - Y\|_2^2$ [12] up to numerical stability, and so has been a popular choice for training the linear layer at the output of reservoirs [93, 116, 42, 109]. However, our goal was to classify the input signals based on the *largest* element of the final output vector. Consequently, the linear layer that resulted in the lowest mean squared error with our labels was not always the linear layer that gave us the best accuracy.

For this reason, we also used a second training method for our linear layer. This approach used softmax, a popular choice for classifiers in neural networks [7] and back-propagation using the automatic differentiation package from PyTorch [99]. Training through back-propagation with an optimizer is now necessary since an exact analytic solution to minimize the loss no longer exists, unlike the case of the pseudo-inverse. In this approach, the prediction vector y from Eq. 4.24 is passed through the “Softmax” activation function:

$$(y_{\text{prediction}})_i = \frac{\exp(y_i)}{\sum_{j=1}^C \exp(y_j)} \quad (4.26)$$

We then computed the mean squared error between the resulting $y_{\text{prediction}}$ and the label for the training point that produced the underlying reservoir feature x . Finally,

we used back-propagation to compute the gradient for our linear layer. The linear layer was then updated using the ADAM optimizer [69] with the default settings of $\beta_1 = 0.9$, $\beta_2 = 0.999$ and a learning rate of 0.01. For our reservoirs, we tried both methods of training the linear layer and used whichever yielded the best accuracies. Empirically, we found that while pseudo-inverse training was better in some cases, training the linear layer with back-propagation often yielded quite large accuracy advantages over pseudo-inverse.

4.8 Supplementary information machine learning tasks

4.8.1 Classification of Radio-Frequency signals

In this section, we discuss about the algorithm for generating the dataset for the classification of digital modulation schemes on radio signals. The digital modulation scheme involves encoding sequences of binary values into the amplitude and phase of a radio signal for a fixed duration. The number of binary values encoded depends on the modulation scheme. For example, for BPSK (binary phase shift key), each symbol (change in property of the signal) encodes one bit of information. For 32QAM (quadrature amplitude key), there are 32 possible values, which allows each symbol to contain 5 bits of information. For this task, we keep the symbol rate fixed across all the tasks. Moreover, the pulses generated by the arbitrary waveform generator (AWG) all occur at the baseband frequency. This signal is then upconverted to the frequency of the cavity before sent to the device. To generate the set of possible sequences, we randomly select each symbol with equal probability. This corresponds to the case of each possible binary string of digits encoded to be equally likely. Due

to memory constraints on the AWG, we cannot output a continuous encoded signal for long durations, corresponding to the regime of large samples of the reservoir. We circumvent this constraint by realizing that, for this task, there are no correlations in the encoded binary digit sequence (since each symbol is equally likely). Therefore, the probability of a long binary digit sequence can be correctly emulated by sampling multiple short binary digit sequences and concatenating them together. For this task, we can simply achieve this by generating a signal with eight symbols, which is the number of symbols enter our QRC before its state is completely reset.

4.8.2 Classification of noisy signals

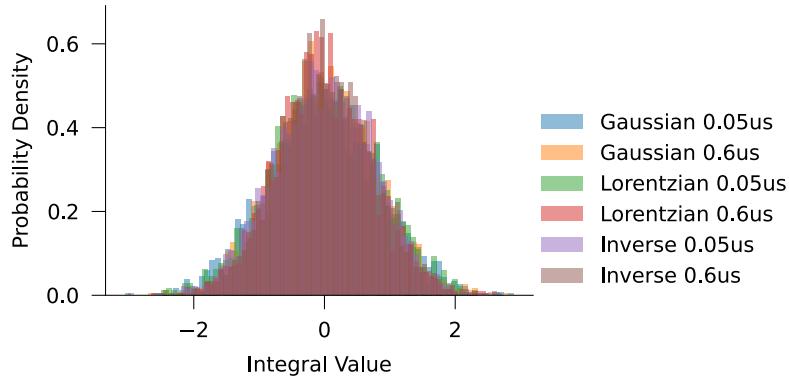


Figure 4.18: Histogram of integrated value of the input signal for the noise classification task. In this work, we enforced a normalization condition on the amplitude of the filter functions. We set the normalization such that the long time integral of the signal corresponds to a value with zero mean and the same standard deviation. This is visually seen from the probability density function from the dataset of the classes of noisy signals. We do this to ensure that any reservoir which simply integrated the signal, before applying a non linear kernel, cannot classify the different signals. The fact that our reservoir is able to solve this task can therefore be associated with the continuous-time processing by the cross-Kerr interaction between the qubit and the cavity. Enforcing this normalization is mathematically equivalent to setting the DC component of the filter functions to be the same value in frequency space.

To generate the dataset describing the task of classifying noisy signals using the

QRC (see Fig. 4.4), we start with emulating white noise. At each time step of the sampling rate of the AWG, we choose a value for in-phase and quadrature signals uniformly between the unit interval (up to an overall normalization). While this is limited to the sampling rate of the AWG (around 2×10^9 samples per second), this is much larger than any relevant time scale of the experiment. Therefore the approximation of broadband white noise is appropriate to describe the effect of the signal on the system. We then apply “kernels” as a convolution in the time domain to each need seed of the white noise generated signal. This can also be thought of a bandpass filtering function in frequency domain. The classification task is then to identify the kernel. Each kernel is defined by a time domain function. The only hyper-parameter to describe each kernel is the overall scaling value. In this work, we set the DC component of this kernel in frequency domain to be the same for all classes (set to unit value without loss of generality). In the time domain, this corresponds to scaling the amplitude such that the area enclosed by the filter function in time is the same for all functions.. We do this to make sure that a direct integration of the signal over a time domain much longer than the correlation length introduced by the kernel, cannot distinguish the signals from each other (see Fig. 4.18). The above normalization ensures the random variable associated with this integrated value is the same for all distributions. Therefore, this ensures that any ability of the reservoir to classify the signals arises intrinsically from its computational capacity to distinguish short-time correlations (in this work we choose a correlation time scale of 50ns and 600ns, with kernel functions of Gaussian, Lorentzian, and the Inverse function: generating a total of 6 classes.).

4.9 Simulation of the quantum reservoir

4.9.1 Introduction

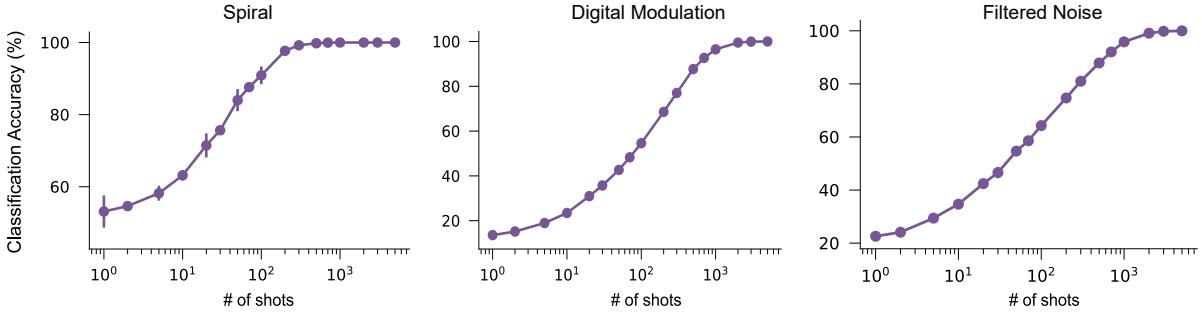


Figure 4.19: **Classification accuracies obtained from numerical simulations of the QRC.** Numerical simulations of the quantum reservoir can help guide the expected performance of the system in experiment. Here, we simulate the three main tasks considered in the paper: Spiral (a time-independent input with the goal of classifying the arms of the spiral), Digital Modulation (a slow-varying, time-dependent signal with the goal of identifying the digital modulation scheme used in the signal), Correlated Noise (a fast-varying, time-dependent signal with the goal of classifying the “kernel” of the correlation function).

Classical simulations of the QRC can provide insight into the expected computational capacity in experiment. For our work, classical simulations of the dynamics of the reservoir were primarily performed with the aid of QuTiP [63]. The algorithm to estimate the classification accuracy for a given task then follows the same technique used in experiment, with training and testing datasets on the measurement outcomes of the simulation. We implement the Hamiltonian in Eq 4.1, by approximating the transmon as a qubit, and introducing a finite dimensional Fock truncation to the cavity subspace. It is important to ensure that the Fock truncation does not introduce any spurious effects, for it can be a source of non-physical non-linearities in the system. For example, a linear cavity, treated as a harmonic oscillator, only performs a linear transformation on an incoming analog radio frequency signal. However, if in simulation, the support of the state of the cavity exceeds the Fock truncation of the simula-

tion, numerical errors introduce non-Gaussian states in the cavity mode. Such effects will depend non-linearly on the input, and hence can effectively act as a “good” (but of course unphysical) reservoir! To ensure this doesn’t happen in simulation, at every step of the unitary evolution, we monitor the probability of the wavefunction on the largest Fock state in the simulation. If this value goes above 1% during the simulation, a warning is raised, and the results of the simulations are discarded.

To make the simulations efficient, we make certain assumptions on the quantum system. Firstly, we treat the reservoir controls of qubit rotations and conditional displacements with a “gate”-based unitary. However, to take into account the analog, continuous dynamical evolution implemented by the cross-Kerr interaction term in the Hamiltonian, the interval of the input into the system is implemented with the full time-dependent Hamiltonian evolution (using QuTiP’s “mesolve” functions). Finally, another approximation we make (in favor of simulation speed) is ignoring decoherence effects. To ensure this approximation is valid, we performed simulation with the stochastic wavefunction approach with photon loss and qubit dephasing rates measured in experiments [23]. We obtain differences in expected classification accuracies within error bars (which are obtained from different datasets from repeated simulations). This also gives us confidence that the role of decoherence in the system plays a minimal role in the computational capacity of the reservoir. The results of the simulations with the third central moment are plotted in Fig. 4.19. Interestingly, the performance as a function of the number of samples agrees to experiment within the same order of magnitude. This gives a good estimate for the experimental time required to produce a classification accuracy versus shots curves in experiments. For all three tasks, the reservoir approaches 100% accuracy with sufficient samples or integration time of the input.

4.9.2 The advantage of continuous-time continuous-variable QRCs over discrete-time qubit-based QRCs

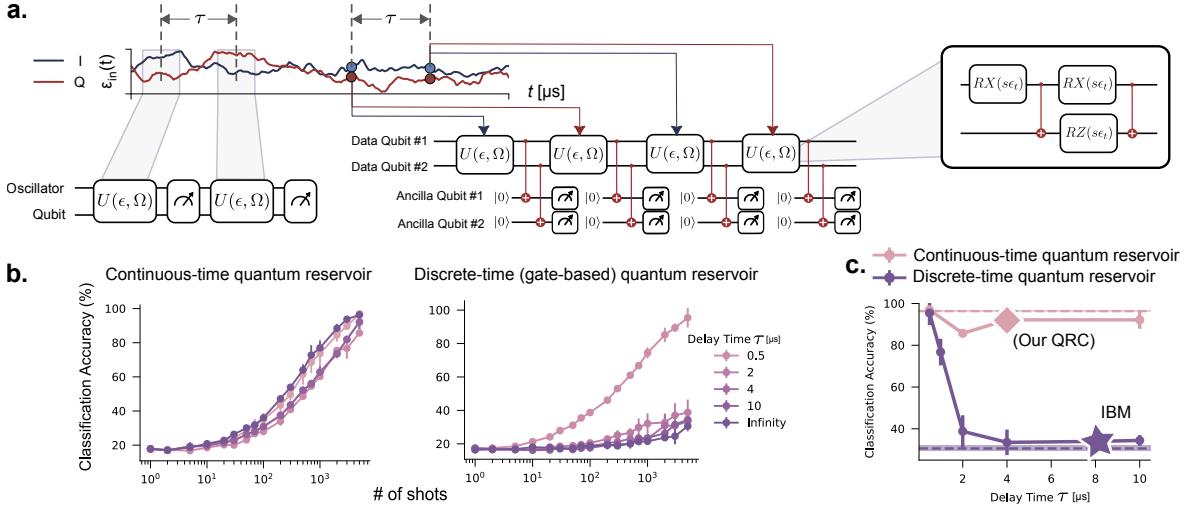


Figure 4.20: **Simulations comparing the performance of a continuous-time quantum reservoir (our work), and a discrete-time qubit-based quantum reservoir, based on a recently implemented protocol [130].** (a.) Schematic reservoir protocol for our experiment (on the left), versus that introduced in [130], involving two qubits as the data qubits, and two as the ancillas. The reservoir consists of single qubit rotations interleaved with CNOT gates between the data qubits and data and ancilla qubits. In this simulation, we simulate the performance of the two reservoirs as a function of the delay τ between two durations of the input. Such a delay can be introduced by the finite pulse durations and latencies introduced by the FPGA. (b.) Classification accuracy curves for the two reservoirs as a function of shots, for different values of delays. The continuous time analog reservoir is much more robust to delays between inputs compared to the discrete time reservoir implementation. (c.) Plot of accuracies at 5000 shots for the two reservoir implementations as a function of delay time. Experimentally relevant times include $4\mu s$ for our experiment, and around $8\mu s$ [57] for the experimental realization of [130] on an IBM quantum computer.

In this section, we benchmark the performance of our continuous-time-continuous variable QRC in comparison with other hardware implementations of reservoirs. To highlight the benefit of our QRC in processing time varying input signals, we compare the simulation of our reservoir with that of a recent QRC scheme involving repeated measurements on a multi-qubit based superconducting circuit

quantum system [130]. For this comparison, we simulate the expected performance of both systems on the task of classifying different noise signals with the classes described in Fig. 4.4. While our reservoir can naturally interface with analog signal, this is not the case with the protocol introduced in [130]. For this simulation, the signal is sampled at discrete times and input to the system as a scalar parameter (one for the in-phase value and one for quadrature value). To highlight the advantage of our QRC, we slightly modify the task introduced in the Fig. 4.4. Here, we normalize the six filter functions such that the integral of the filter function in frequency domain is kept constant. We do this such that the standard deviation associated with the distribution of the sampled signal is the same across all signals. The only information distinguishing the signals is in the correlation between two close samples in time. To elucidate this reasoning, we simulate the performance of the two reservoirs as a function of the time duration in between two samples of the signal (in the case of the discrete qubit based reservoir) and integration windows (for our analog reservoir). Such a finite duration can arise from finite-pulse durations of reservoir protocols, qubit-measurement times, and the finite latency introduced by the classical FPGA processor. For example, for our experiment, this time is around $4\mu s$, mostly arising from the measurement of the qubit and the parity of the cavity. In experiment (Fig. 4.4), we had generated and timed the input wave forms such that the delay between inputs is essentially $0\mu s$. For a typical IBM quantum device with mid-circuit measurement, the protocol used in Ref. [130], the finite latency can be estimated to be around $8\mu s$ [57]. The protocol for the discrete-time quantum reservoir is designed to only act on real-valued input signals. However, for a continuous signal in the rotating frame, we have both the in-phase and quadrature values. For the experimental quadrature, these values correspond to displacements on the oscillator in orthogonal directions. To extend the scheme presented in [130], we do the following minimal

change: we interleave the between sample points of in-phase and quadrature values. We could have chosen these points with the delay of τ in between each. However, this might have had the effect of introducing twice the delay compared to the continuous time reservoir. Therefore, we chose the relaxed constraint of the input such that both the in-phase and quadrature values are chosen at the same point, with just a delay in between two different in-phase and quadrature points.

4.9.3 Comparison to other reservoirs

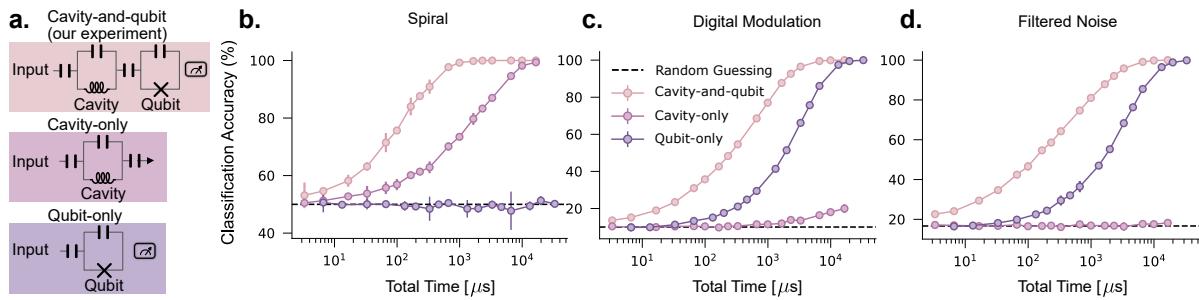


Figure 4.21: Simulations of the performance of the cavity-and-qubit reservoir system with its components: just a cavity reservoir and just a qubit reservoir. (a.) The circuit diagrams of the three quantum reservoir simulated here. 1) The circuit diagram of the experimental QRC in this work: a cavity coupled to a qubit. The input is interfaced to the cavity, and the state of the qubit is measured using the standard dispersive readout technique with a readout resonator. 2) The circuit diagram of a cavity as a reservoir. In this case, the natural output of the cavity is the transmitted signal. 3) The circuit diagram of a qubit reservoir. In this case the input is coupled instead to the qubit. (b.) Simulation of the classification accuracy of the three different reservoirs for the three tasks considered in the paper. As a fair comparison, the output feature vector dimension is kept constant for all three reservoir. The experimental setup drastically outperforms either of its components, just a cavity (in (c.)) and just a qubit (in (d.)), which highlights the important role of entanglement in classification accuracy.

A cavity coupled to a qubit is a hardware efficient quantum system to perform reservoir computing on analog signals. In this section, we motivate this by simulating the performance of other natural choices of quantum reservoirs: a single qubit, and a

single cavity. The protocols for these systems are inspired by what one can naturally perform in experiment. To make a reservoir with a cavity, we couple the input into the cavity (as is the case for the experimental design). To readout the cavity, we perform a transmission style Homodyne measurement, which infers the mean field value of the cavity. This is a continuous form of measurement, where the output feature is a time dependent radio frequency signal at the frequency of the cavity mode. Since the cavity is always in a coherent state, the output time trace is linearly dependent on the incoming signal. For a fair comparison, we only use a handful of values from the time trace (as many as the number of measurements in the experiment). While this might seem restrictive, we process this via the same method as the case of the experimental reservoir, by computing the functional definition of the central moments. This does not necessarily make sense for this protocol, since the outputs do not correspond to samples from a discrete probability distribution, but can nevertheless introduce non-linearities in the representation of the feature vector. These non-linearities can improve the performance of the reservoir beyond a linear layer. This is observed for the case of time-independent Spiral classification, where the cavity reservoir performs better than random. This performance is solely due to the “post-processing” of the output of the reservoir we adopt for our experiment. However, for time dependent tasks, the performance is hardly better than random.

Another natural candidate is a single qubit reservoir. For this case, we directly interface the signal to the qubit. A qubit is able to naturally represent non-linear functions of the input, which can be intuitively seen by visualizing the action of a qubit rotation on a Bloch sphere. As a fair comparison, we choose the same qubit reservoir controls in experiment, which involve qubit pulses before, during and after the continuous input. The output is a string of binary outcomes of qubit measurements, which can be done experimentally with the use of a readout resonator. Each

reservoir of the qubit lasts twice as long as the experimental QRC to obtain the same feature vector size. This is then processed the same way as the cavity-qubit coupled reservoir, before applying a trained linear layer. Interestingly, the qubit fails to perform better than random for the Spiral task. On the other hand, it is able to reach near 100% accuracy for the time dependent signal classification tasks. The ability for even a single qubit to successfully perform a many-class classification task is illuminating at the remarkable processing capabilities of reservoir. However the total input signal required can be more than order of magnitude longer compared to that for the experimental QRC to achieve the same accuracy. The ability of a cavity coupled to a qubit system to perform significantly better than either of its components provides a clear picture of the important role entanglement can play.

4.9.4 Multi-qubit reservoirs

Quantum reservoir computing is a promising paradigm in the NISQ era. It is therefore interesting to consider the potential benefits in performance with larger devices, which are within reach of today's experimental capabilities. As a natural extension of our quantum reservoir, we consider a scenario of one continuous variable cavity mode dispersively coupled to multiple qubits. For simplicity, we assume the dispersive strength of each qubit to the cavity is the same. To motivate the capacity of such a reservoir, we simulate the system for up to four qubits to estimate the classification accuracy for the task of identifying correlated noise signals. The unitary protocol is illustrated in Fig 4.22 (a.). The protocol begins with a $\pi/2$ pulse on each qubit, which brings the state of the qubit onto the equator of the Bloch sphere. To entangle the qubits with cavity, we simulate the action of a generalized, multi-qubit-conditioned cavity displacement. This involves a displacement on the cavity, whose value is dif-

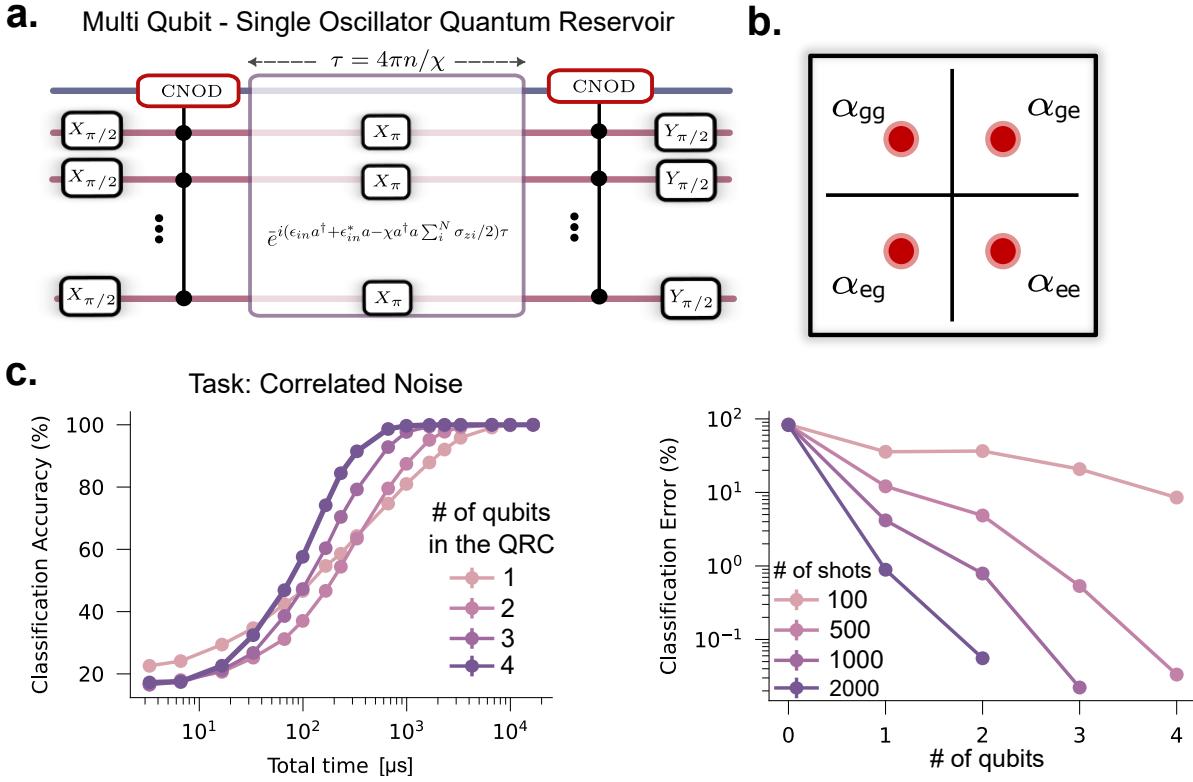


Figure 4.22: Exploring the computational capacity of more complex reservoirs, involving a single cavity mode coupled to multiple qubits. (a.) The reservoir protocol, before measurement. The protocol is inspired by naturally extending the protocol of the experiment. The state of each qubit, along with the parity of the cavity, is sampled afterwards. (b.) Schematic illustration of the state of the cavity after the generalized qubits conditioned cavity displacement implemented in this reservoir, for the case of two qubits. Here the cavity is displaced by a unique value of each 2^N combinations of qubit possibilities, for a reservoir with N qubits. (c.) To illustrate the computational capacity of such reservoir, we simulate the quantum system to obtain the classification accuracy as a function of the duration of signal received for the task of classifying correlated noise signals. Increasing the number of qubits generally increases the performance of the reservoir. For different values of shots of the reservoir, which corresponds a total time of input, we plot the classification error as a function of qubits. The case of zero qubits in the reservoir corresponds to just a cavity reservoir.

ferent for each of 2^N N -qubit possibilities. An example of the action of this operator is depicted in Fig 4.22 (b.) for the case of two qubits. Here, there are four possible qubit states, and each is associated with a displacement value of the four corners of the square. The displacement values were chosen somewhat arbitrarily, but serve to illustrate an efficient multi-component entanglement. For the case of two qubits, the

real and imaginary components of the displacement where either ± 0.5 . For the case of three qubits, there are eight total possible states. The set of displacements chosen form a three-by-three grid state, ranging between ± 1 , excluding the center of this grid (which is centered at the origin). For the case of four qubits, a four-by-four grid uniformly distributed between ± 1.5 covers all sixteen possibilities. The correspondence between qubit states and displacements was somewhat arbitrary - the motivation is that even without much design choice, a reservoir can successfully implement machine learning! For these simulations, each position of the grid is associated with a decimal value, increasing sequentially from left to right, starting from the top left and progressing towards the bottom right (starting with zero). This decimal value is the decimal representation of the qubit-state bit string that the displacement is conditioned on.

After the multi-qubit entangling conditional displacement gate, the cavity is subject to the input. The dynamics of the system are influenced by the cavity-qubit coupled dispersive interactions, where the interaction strength is the same between the cavity and all qubits and set to that of the experiment. Like the experimental QRC, each qubit is flipped with a π pulse in the middle of the input. The protocol ends with the same conditional displacement, before a $\pi/2$ pulse. The output of the reservoir is the measurement of each qubit, along with the parity measurement of the cavity. This protocol is repeated four times, to match the experimental protocol as much as possible.

Fig 4.22 (c.) is the classification accuracy as a function of the total time of input signal received for the reservoirs with different number of qubits, for the task of noise classification. While they all achieve essentially 100% accuracy, the total time required to achieve this accuracy drops significantly with increasing number of qubits. Other

than the single qubit reservoir (the experimental protocol), the performance of the reservoir is similar both at the low and high signal duration regime, differing only in the intermediate regime. The reason for the difference in behavior of the performance for the case of cavity coupled to a single qubit is the slight change in reservoir protocol. To accurately account for the experimental protocol, the state of the qubit is determined by the outcome of the parity measurement of the cavity. This was not implemented in the simulations for multiple qubits. This ends up improving the performance of this reservoir for this task in the low signal duration regime. However, in the higher signal duration regime, increasing the number of qubits increases the accuracy.

The classification error as a function of number of qubits in the reservoir is plotted in Fig 4.22 (d.), including the case for just a cavity (zero number of qubits in the reservoir), for a select number of total shots of the entire reservoir. Very crudely, the error in classification seems to reduce exponentially with every additional qubit in the reservoir.

4.10 Theoretical analysis of the expressivity of our QRC for time-independent signals

The ability of the QRC to perform better than an optimal linear layer on the input lies in the reservoir's ability to express many non-linear functions of the input—its expressivity. Here, we quantitatively characterize the class of functions which can be represented by the oscillator component of the QRC for a time-independent input. In this regime, the input can be represented by two variables: the values of the in-phase and quadrature components. The output feature vector from the QRC is then

a function of these two variables.

We denote $\alpha = |\alpha|e^{i\phi_\alpha}$, $\beta = |\beta|e^{i\phi_\beta}$, and set $|\alpha| = 1/2$. Choosing different values of ϕ_α gives rise to different output features of the QRC. In this experiment, we pick $\phi_\alpha \in \{0, \pi/2\}$, but in principle, one can add to the feature vector with more choices of ϕ_α . For example, one can choose $\phi_\alpha \in \{0, \omega, 2\omega, \dots, (r-1)\omega\}$ where $\omega = \frac{2\pi}{r}$. The final output after the linear layer is an arbitrary linear combination of all the $p_\alpha(\beta)$ functions.

Intuitively, the larger r , the more expressive the function space spanned by these features. Furthermore, the higher-order central moments allow the output feature vector to represent powers of this probability: $p_\alpha(\beta)^n$, for moments up-to the n th-order. We have shown that the qubit measurements extract the phase information of the input complex number β . Below we will focus on the oscillator parity measurement which is sensitive to the magnitude of β . Recall that the post-measurement (unnormalized) state of the cavity can be described by a sequence of alternating displacements and parity measurements (Eq. 4.15):

$$|\Psi_{\vec{x}}(\beta)\rangle = P_{x_M}D(\beta) \cdots P_{x_2}D(\beta)P_{x_1}D(\beta)|0\rangle, \quad (4.27)$$

where P_{x_i} is the projector of the i -th parity measurement with outcome $x_i \in \{0, 1\}$, with ‘0’ standing for ‘even’ and ‘1’ for ‘odd’. That is, $P_{x_i} = \frac{I + (-1)^{x_i}\Pi}{2}$, where $\Pi = (-1)^{a^\dagger a}$. The corresponding probability of obtaining $\vec{x} = (x_1, x_2, \dots, x_M)$ as the sequence of measurement results given the input β is

$$\Pr[\vec{x}|\beta] = \langle \Psi_{\vec{x}}(\beta) | \Psi_{\vec{x}}(\beta) \rangle. \quad (4.28)$$

To obtain a simplified expression for $\Pr[\vec{x}|\beta]$, we will make use of the following

formula:

$$P_x D(\beta) P_y = \frac{D(\beta) + (-1)^{x \oplus y} D(-\beta)}{2} P_y, \quad \forall x \in \{0, 1\}, \forall y \in \{0, 1\}, \quad (4.29)$$

which is an easy application of the commutation relation $\Pi D(\beta) = D(-\beta)\Pi$, with the latter being derived from $\Pi a = -a\Pi$. Using Eq. 4.29, we can remove all the explicit parity projectors in Eq. 4.27:

$$|\Psi_{\vec{x}}(\beta)\rangle = \left(\prod_{i=1}^M \frac{D(\beta) + (-1)^{x_i \oplus x_{i-1}} D(-\beta)}{2} \right) |0\rangle, \quad (4.30)$$

where for notational simplicity we have prepended the bit-string \vec{x} by $x_0 \equiv 0$. Note that the order of the product does not matter since the terms commute with each other. It follows that:

$$\begin{aligned} \Pr[\vec{x}|\beta] &= \langle 0 | \left(\prod_{i=1}^M \frac{D(-\beta) + (-1)^{x_i \oplus x_{i-1}} D(\beta)}{2} \right) \left(\prod_{i=1}^M \frac{D(\beta) + (-1)^{x_i \oplus x_{i-1}} D(-\beta)}{2} \right) |0\rangle \\ &= \langle 0 | \left(\prod_{i=1}^M \left[\frac{1}{2} + (-1)^{x_i \oplus x_{i-1}} \frac{D(2\beta) + D(-2\beta)}{4} \right] \right) |0\rangle. \end{aligned} \quad (4.31)$$

There are multiple methods to encode the measurements of the QRC. Representing every binary string of measurement outcomes as the feature, the output of the QRC are all the probabilities $\{\Pr[\vec{x}|\beta]\}_{\vec{x} \in \{0,1\}^M}$. From Eq. 4.31, it is not hard to see that when regarded as functions of β , these 2^M features linearly span a $(M+1)$ -dimensional function space that has the following basis functions:

$$f_k(\beta) := \langle 0 | D(2k\beta) | 0 \rangle = e^{-2k^2|\beta|^2}, \quad k = 0, 1, 2, \dots, M. \quad (4.32)$$

Therefore, the set of all functions realizable by the QRC combined with the linear layer is

$$\{c_0 f_0(\beta) + c_1 f_1(\beta) + \dots + c_M f_M(\beta) : c_0, c_1, \dots, c_M \in \mathbb{R}\}. \quad (4.33)$$

Given the large redundancy of the output feature encoding manifested above, a compact representation can be the centralized moments $\mu_{i_1, i_2, \dots, i_k}(\beta) := \mathbb{E}[(x_{i_1} - \mathbb{E}[x_{i_1}]) (x_{i_2} - \mathbb{E}[x_{i_2}]) \cdots (x_{i_k} - \mathbb{E}[x_{i_k}])]$. These feature functions contain terms like $\mathbb{E}[x_1] \mathbb{E}[x_2]$, $\mathbb{E}[x_1]^2$, $\mathbb{E}[x_1] \mathbb{E}[x_2] \mathbb{E}[x_3]$, and so on. In particular, for any k , $\mathbb{E}[x_1]^k = \left(\frac{1}{2} - \frac{e^{-2|\beta|^2}}{2}\right)^k$ can be written as a linear combination of centralized moments of order less than or equal to k . It follows that the QRC using at most k -th order centralized moments combined with the linear layer can realize (but not limited to) the following vector space of functions:

$$\mathcal{H}_{\text{parity}} := \left\{ c_0 + c_1 e^{-2|\beta|^2} + c_2 \left(e^{-2|\beta|^2}\right)^2 + \cdots + c_k \left(e^{-2|\beta|^2}\right)^k : c_0, c_1, \dots, c_k \in \mathbb{R} \right\}. \quad (4.34)$$

Note that $\mathcal{H}_{\text{parity}}$ is exactly the set of all degree- k polynomials in the variable $w \equiv e^{-2|\beta|^2}$. Suppose that in some classification task, the magnitude of the input has an upper bound, say, $|\beta| \leq 1$, then w takes value in the closed interval $[e^{-2}, 1]$. By the Stone–Weierstrass theorem, in the limit $k \rightarrow \infty$, \mathcal{H}_k approximates all continuous functions of w on $[e^{-2}, 1]$, and hence all continuous functions of $|\beta|$ on $[0, 1]$.

4.11 Leaky Echo State Networks (LESN)

4.11.1 Background

Leaky echo state networks [60] are a generalization of echo state networks (ESN) [59] that were found to outperform their parent design in prediction and classification of slow dynamic systems, noisy time series and time-warped dynamic patterns [116].

Given a sequence of inputs $\{u_n\}_{n=1}^N$, $u_n \in \mathbb{R}^D$, the state of the LESN reservoir after the n^{th} input u_n , x_n , is given by the following equation:

$$x_n = (1 - a\gamma)x_{n-1} + \gamma f(W_{\text{in}}u_n + W_{\text{res}}x_{n-1}). \quad (4.35)$$

Here, a, γ are fixed hyper-parameters in $[0, 1]$, and f is a nonlinear activation function. W_{in} is the $R \times D$ “encoding” matrix whose elements are selected uniformly at random from the interval $[-w_{\text{in}}, w_{\text{in}}]$, where D is the dimension of the input, R is the dimension of the reservoir, and w_{in} is a fixed hyper-parameter. W_{res} is the $R \times R$ “reservoir” matrix. This matrix is constructed by first generating a matrix W_R , which is a random matrix whose elements are chosen to be zero with probability $1 - p_s$ and a number sampled uniformly from the interval $[-1, 1]$ with probability p_s . The largest-magnitude singular value of this matrix, $\lambda_{\max}(W_R)$ is computed, and the reservoir matrix W_{res} is defined as:

$$W_{\text{res}} = \frac{\rho}{|\lambda_{\max}(W_R)|} W_R \quad (4.36)$$

where ρ is a fixed scaling hyper-parameter. Finally, the n^{th} output of the reservoir, y_n , is given by

$$y_n = W_{\text{train}}x_n, \quad (4.37)$$

where W_{train} is a $C \times R$ trainable linear layer, where C is the dimension of the desired output vector.

4.11.2 Digital reservoir comparison

As a way of benchmarking the computational capacity of our physical reservoir, we compared it to the performance of a *digital* reservoir - an LESN - at varying widths and depths. We focused on the accuracy of classifying the spiral, since this is the most direct point of comparison, as the goal was to classify individual points of a

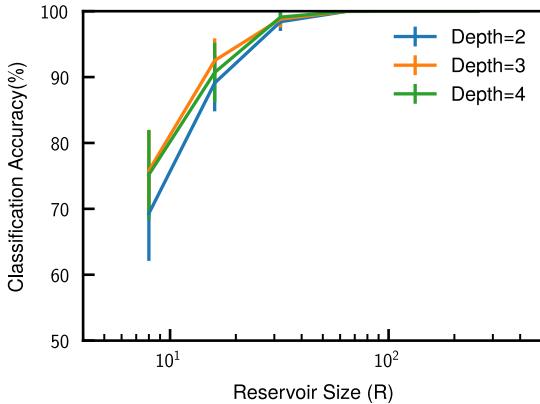


Figure 4.23: Mean spiral classification accuracies and their standard deviations over 100 randomly generated LESN’s.

signal rather than multiple separate signals per-shot as with the time-dependent case. Here, for a depth of N , we sent in N identical two-dimensional data points (x, y) (so $D = 2$) corresponding to the I and Q components of the signal that our experimental reservoir is meant to process, i.e. the spiral point coordinates. We used the rectified linear unit (ReLU) as our nonlinear activation function. Traditionally, sigmoid or tanh activation functions are used for LESN’s [60, 116, 109], but ReLU was found to work better for our application.

To investigate how “trivial” it was to generate a classifier with the same capacity as our experiment, we generated 100 such LESN’s at random, and found their average performance, and standard deviation. Hyper-parameters a , γ , w_{in} , ρ , and sparsity (p_s) were tuned in sweeps to improve performance as much as possible for each width and depth, in order to give the digital reservoir a competitive chance. The reservoir’s computational capacity varies by the number of shots. Comparing Fig 4.2(b) to Fig 4.23, we found that, at around 10^3 shots, our physical reservoir achieved a performance comparable to that of about that of a 32-dimensional LESN reservoir, as seen by the fact that both oscillate around 99% classification accuracy, within about one

percent. In Fig 4.23, a 64-dimensional reservoir was found to be enough to classify the spiral data points with perfect accuracy and a fairly wide choice of parameters. Our reservoir, then, achieved at *least* the capacity of a 64-dimensional LESN reservoir, past around $5 * 10^3$ shots.

In the conventional view of reservoirs, the data must be sent into into a higher dimensional space where linear separability becomes possible [116]. The dimensionality of our Hilbert space, given by the two dimensions of the qubit and the approximately 16 occupied levels of our storage resonator, limits the complex degrees of freedom we have available in encoding our data as the final, measured state. Consequently, we use this total Hilbert space dimension (times two due to complex amplitudes) as a proxy for the quantum resources used in terms of reservoir dimensionality. Given the roughly 2×16 dimensions of Hilbert space used by our reservoir, achieving at *least* the computational capacity of a 64-dimensional LESN reservoir is on the order of what would be expected for a large shot number. Indeed, the point of this comparison is to demonstrate that the computations performed by our reservoir cannot be trivially replicated by a digital reservoir with fewer resources with the same performance.

CHAPTER 5

OUTLOOK

5.1 Toward the construction of a large-scale frequency domain neural network accelerator

In section 3, we described the construction of a frequency domain photonic simulator capable of simulating Gaussian bosonic physics on two- and three-dimensional lattices, with a number of lattice sites over 100,000. The lattices were programmed using microwave tones sent to a nonlinear $\chi^{(2)}$ crystal, that allowed on-the-fly programmability of these lattices, and enabled us to demonstrate exotic phenomena such as non-Hermitian dynamics and time-reversal symmetry breaking. Additionally, we demonstrated the ability to generate fully arbitrary input excitations into the lattices, albeit with support limited to a few thousand sites within the full lattice. The large scale, programmability, and the ability to encode arbitrary information in the inputs motivates one to think about the possibility of construction a physical neural network (PNN) [129] out of the system. Here, we outline the prospects of this system as a laser neural network (LNN).

The laser neural network (LNN) performs machine learning tasks in a frequency-domain data encoding, with the amplitude of longitudinal modes in a fiber laser representing neural activation. Optical gain media, such as Erbium, allow 10 THz of simultaneously-lasing modes. Taken together with a long fiber cavity allows for an enormously complex, controllable nonlinear dynamical system. In the LNN, these dynamics can be controlled for computational purposes with high bandwidth, low power consumption, and cost using mature optical telecommunications hardware.

The processing elements in the LNN are an electro-optic modulator (EOM), nonlinear four-wave mixing through the optical Kerr effect, and the saturating laser gain (see section 2). In a typical neural network architecture, processing is done by alternating sequences of linear and nonlinear transformations. In the LNN, the EOM provides adjustable linear coupling among modes. As outlined in section 2, the action of the electro-optic modulator is a linear operation that can be expressed as a Toeplitz matrix. The Toeplitz matrix can be understood as convolution, which couple modes in a translationally invariant way, up to the bandwidth of the EOM (typically on the order of 10s of GHz). To obtain longer range coupling, one could import into the cavity two offset spectral filters. Together with the Kerr nonlinearity in the fiber, these filters implement artificial saturable absorber (SA), which promotes all-to-all nonlinear coupling between the modes [33]. The addition of this Kerr nonlinearity has the added benefit of adding the nonlinearity, an ingredient in neural networks required for almost all non-trivial tasks.

Increasing the fiber cavity length discussed in section 3 from 100 meters to a 1 km long cavity, the mode spacing shrinks down to ~ 100 kHz. With an optical bandwidth of 10 THz, the number of modes making up the computational space is $N_\omega = 10$ THz/100 kHz = 10^7 modes. By driving a 10 GHz bandwidth EOM in the laser cavity, we can tune the coupling between modes more than 10 GHz away. For a 100 kHz cavity, this provides tunable coupling among groups of $N_L \sim 10^5$ modes. By switching the microwave tones inducing the coupling as the light propagates around the cavity, one can perform time-dependent linear operations, akin to different layers in a neural network. These microwave tones are programmed with high-fidelity DACs and will serve as the programmable knobs that one would tune to train the physical system, using training methods such as physics-aware-training (PAT) [129].

Unlike a conventional neural network, which is traditionally composed of linear layers with on-site nonlinearities, here, the nonlinearity is manifested in long-range coupling. The full equations of motions for this system is

$$\begin{aligned}\dot{a}_n(t) = & \left(\frac{g_n}{1 + \sum_m |a_m|^2 / U_0} - \ell \right) a_n(t) + \sum_m J_{n-m}(t) a_m(t) \\ & + (\Gamma - i\Delta) \sum_{m_1+m_2+m_3=n} a_{m_1}(t) a_{m_2}^*(t) a_{m_3}(t),\end{aligned}\quad (5.1)$$

where ℓ is the loss, Γ and Δ are the nonlinear modulation coefficient associated with the Kerr effect introduced by the saturable absorber. Expressing the above equations of motion in terms of the number of digital operations per second R_{ops} , The potential computationally capability can be characterized as:

$$R_{\text{ops}} = \Delta f \log_2(\Delta f \Omega) + \Delta f + \Delta f N_z \log_2(\Delta f \Omega) \quad (5.2)$$

where Ω is the mode spacing and its inverse is equal to the cavity round-trip time, and sets the timescale for operations. Δf is the optical bandwidth, and $N_z \approx 10L/L_{nl}$, where L is the length of the cavity, and L_{nl} is the length scale over which nonlinear optical effects become important (approximately mm-cmm for typical systems). The first term describes the action of the EOM, which, in the time domain, is a matrix multiplication of a diagonal matrix. The logarithmic factor is thus the cost in performing the fast Fourier transform. The computation needed to simulate the effects of the nonlinear four-wave mixing, which is also simplified by moving to the time-domain.

Equation 5.2 describes the best-case scenario for the computational power of the LNN. While each term has somewhat poor scaling, most being logarithmic in the number of degrees of freedom, the experimental constraints in this system allows for the constant factors to grow to very large numbers with minimal effort. This was demonstrated in the experiment in section 3, where a table-top experiment was able realize simulations of a hundred thousand modes. With suitable engineering, this number can grow into the millions.

5.2 Toward an advantage in sensing room-temperature signals using microwave quantum oscillators

In section 4, we outlined a new application of quantum information processing sandwiched between traditional quantum computing and quantum sensing. We demonstrated how a small quantum system, that is simple enough to easily simulate on a classical computer, can be reliably trained to solve signal processing tasks. We showed how such a device could perform nontrivial processing of analog signals that the system directly interfaces with, and that these analog signals were low enough power to only displace the system by a few photons. Finally, we argued that, based on the results demonstrated, one could construct a quantum device that could perform signal processing on signals with very low SNR that originate at room temperature, and that it could in-principle have an advantage over any classical device. Here, we outline the road toward this computational-sensing advantage in more detail.

Superconducting qubit devices are typically designed to be sensitive to signals that lie within 4-10 GHz. While this is conveniently in the same frequency range that is used in commercial communication applications, the energy scales associated with this frequency range are activated by thermal excitations at temperatures above 100 millikelvin. This presents a fundamental challenge in attempting to interface a quantum device with room temperature signals, as the thermal background noise of a given signal is easily capable of exciting the quantum system. This is the primary reason for superconducting qubits to operate at temperatures of 10s of millikelvin, far below the typical superconducting transition temperatures of the materials in the devices. Operating the device in a dilution refrigerator at millikelvin temperatures

will cool a device down to its ground state, but one must also be able to control the device using signals generated at room temperature. These controls are typically performed using resonant microwave signals that can carry a large number of noise photons, which must also be thermalized to millikelvin temperatures.

In order to thermalize control signals to millikelvin temperatures, one must attenuate them so that the number of noise photons at a particular frequency that can resonantly transition the system is well below one. This might present a challenge to using superconducting qubits to perform computation on ultra-low power signals toward gaining an advantage of classical devices: While attenuators attenuate the signal and the noise equally, attenuators also add additional noise, leading to an overall degradation in the signal-to-noise ratio (SNR). However, it turns out that, by properly distributing the attenuation, one can significantly limit the degradation to the SNR, which would otherwise be detrimental to get a sensing advantage. An attenuator with attenuation $\eta = 100$ effectively is a beam-splitter which admits 1% of the incident signal, and dumps 99% of the incident power as heat. Additionally, since a beam-splitter is a four-port device, the attenuator will also contribute vacuum noise given by the effective temperature of the attenuator T_{att} at which the attenuator is thermalized. The relationship between a coherent mode incident on an attenuator a_{in} and a output mode a_{out} is [114]

$$a_{\text{out}} = \frac{1}{\sqrt{\eta}} a_{\text{in}} + \sqrt{\frac{\eta - 1}{\eta}} h, \quad (5.3)$$

where h accounts for the additional noise added by the attenuator, and is required in order for both a_{in} and a_{out} to satisfy the bosonic commutation relations. The average output power can then be shown to be

$$\langle |a_{\text{out}}|^2 \rangle = \frac{1}{2} \langle \{a_{\text{out}}, a_{\text{out}}^\dagger\} \rangle = \frac{1}{\eta} \langle |a_{\text{in}}|^2 \rangle + \frac{\eta - 1}{\eta} \left(\langle h^\dagger h \rangle + \frac{1}{2} \right). \quad (5.4)$$

Identifying $\langle h^\dagger h \rangle$ as the thermal population at the vacuum port, we can express the

number of outgoing noise photons n_i at attenuator i with attenuation η_i as

$$n_i(\omega) = \frac{n_{i-1}(\omega)}{\eta_i} + \frac{\eta_i - 1}{\eta_i} \left(n_{\text{BE}}(\omega, T_{i,\text{att}}) + \frac{1}{2} \right) \quad (5.5)$$

where n_{BE} is the Bose-Einstein distribution dictating the equilibrium photon number distribution $n_{\text{BE}}(\omega, T) = 1/[\exp(\hbar\omega/k_B T) - 1]$. The first term describes the number of noise photons attenuated incident on the attenuator, and the second term describes the additional noise photons added by the attenuator which leads to a degradation in the signal-to-noise ratio (SNR).

In a typical superconducting qubit experiment, $\eta = 60 \text{ dB} = 10^6$ worth of attenuation is added to signal lines to reduce the number of noise photons produced from room temperature signal generators down to less than 10^{-3} photons, well below the added thermal and quantum noise from the attenuation. In principle, all attenuation would be placed at the lowest temperature possible in order to maximally reduce the number of noise photons incident on the device. In such a configuration, the extra noise added would be $\approx 1/2 + 0.022$ photons—near the quantum limit. In practice however, as was done in the experiment described in section 4, the attenuation needs to be distributed across various temperature stages to reduce the heat load down to a level below the cooling power available at a given temperature stage in the dilution refrigerator. In the experiment described in section 4, 20 dB attenuators were placed at the 4 K stage, the 100 mK stage, and the mixing chamber stage (10 mK) for a total of 60 dB of attenuation across the three stages (see Fig. 4.4). With a distributed attenuation scheme, the number of noise photons is roughly $\approx 1/2 + 0.03$.

The above discussion makes it clear that the added thermal noise from the attenuators can be mitigated by sufficiently attenuating input signals at the mixing chamber, the lowest temperature stage in a dilution refrigerator. On the other hand, the presence of the added half-photon of noise can dominate the reduction of SNR with

sufficiently strong attenuation, especially in regimes used for traditional superconducting qubit experiments. In such experiments, one typically generates signals with SNR ranging in the thousands to millions, and thus any small reduction in SNR has minimal effect. Here, the ultimate goal is generally to achieve high-fidelity gates and protect the qubit. In contrast, in a sensing experiment, one relies on the sensitivity of quantum systems to perform useful tasks.

In order to get an impactful demonstration of computational-sensing advantage at processing room-temperature microwave signals with poor SNR, we envision that one would also need to remove the extra half-photon quantum noise as well. The most obvious way to do this is to get rid of all the attenuators and to couple the cavity at critical coupling. Something missing from the above discussion is the coupling of signals in the transmission line into the oscillator mode, which can also be thought of as a beam-splitter. As mentioned in section 4.4, the cavity used in this thesis was undercoupled by a factor of 40. In line with the above analysis of treating attenuators as beam-splitters, this acts as an additional attenuator with attention of about 16 dB, except that here, the idler is reflected rather than dissipated as heat. At critical coupling, all of the incident energy is converted fully into the oscillator. Thus, with critical coupling and no attenuation, a signal would enter our qubit-cavity system with minimal loss in SNR.

While removing all attenuation preserves the SNR of a given signal, such a configuration is in stark contrast to nearly every superconducting qubit experiment. A transmission line thermalized at 300 K would have thermal excitations of 5 GHz modes in the number of thousands of photons. While some superconducting qubit experiments with setups nearly identical to ours have demonstrated the ability to preserve entanglement between an oscillator with thousands of coherent photons and a

qubit [32], one could instead sacrifice a minimal reduction in SNR in order to significantly reduce the number of noise photons by adding 20 dB attenuation at the lowest temperature stage. This would reduce a 300 K noisy signal down to 3 K, reducing the number of thermal excitations by a factor of a hundred, with the trade-off of introducing additional thermal and quantum noise on the order of 100 mK, which would result in a minuscule reduction of SNR.

Ultimately however, the introduction of thermal excitations in the oscillator would lead to dephasing of the qubit via the cross-Kerr Hamiltonian in Eq. 4.1 at a rate of $\Gamma_\phi = 4\bar{n}\kappa\chi^2/(\kappa^2 + \chi^2)$, where κ is the resonator decay rate (and coupling rate to the transmission line, at critical coupling), and χ is the cross-Kerr interaction strength. Thus, to mitigate the dephasing of the qubit and to preserve a quantum-based advantage, one can use an ultra-high lifetime cavity. One may also reduce the interaction strength between the qubit and oscillator, without sacrificing gate speeds as was demonstrated in Refs. [25, 32]. These could improve the dephasing rate contribution from a thermal cavity by a factor of 100 compared to the experiment in section 4. The requirement for a low-dephasing rates is also much less stringent in our experiment, since an important feature of our scheme involves repeatedly measuring the qubit throughout the dynamics (see Fig. 4.1), which projects the qubit every microsecond. As we found in Fig. 4.2, increasing the dephasing time up to a factor of 10 produced no detriment in the classification accuracy. To summarize, removing enough attenuation so that the SNR is essentially unaffected results in a background noise that is ten to twenty times higher than the experiments performed in section 4. However, with sufficient mitigation strategies, we can overcome the detrimental effects of the thermal population in the oscillator, and preserve the ability to perform signal processing even on ultra-low SNR signals originating at room temperature.

The above discussion outlines the route towards achieving a quantum computational-sensing advantage for processing signals with very poor SNR originating at room temperature. The discussion mainly highlights modifications to the experiment in section 4 to allow for sensing of high temperature signals with minimal reduction in the SNR. However, in order to demonstrate advantage, one must compare the results achievable with this modified system with a classical receiver. While in section 4 we showed the ability of our quantum reservoir to outperform classical reservoir computers, exploring this advantage in the setting of sensing high-temperature signals requires additional study.

BIBLIOGRAPHY

- [1] Andrew Arrasmith, Zoë Holmes, Marco Cerezo, and Patrick J Coles. Equivalence of quantum barren plateaus to cost concentration and narrow gorges. *Quantum Science and Technology*, 7(4):045015, 2022.
- [2] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando GSL Brando, David A Buell, et al. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, 2019.
- [3] Réouven Assouly, Rémy Dassonneville, Théau Peronnin, Audrey Bienfait, and Benjamin Huard. Quantum advantage in microwave quantum radar. *Nature Physics*, 19(10):1418–1422, 2023.
- [4] Kelly M Backes, Daniel A Palken, S Al Kenany, Benjamin M Brubaker, SB Cahn, A Droster, Gene C Hilton, Sumita Ghosh, H Jackson, Steve K Lamoreaux, et al. A quantum enhanced search for dark matter axions. *Nature*, 590(7845):238–242, 2021.
- [5] Zachary Ballard, Calvin Brown, Asad M. Madni, and Aydogan Ozcan. Machine learning and computation-enabled intelligent sensor design. *Nature Machine Intelligence*, 3(7):556–565, June 2021.
- [6] Miguel A Bandres, Steffen Wittek, Gal Harari, Midya Parto, Jinhan Ren, Mordechai Segev, Demetrios N Christodoulides, and Mercedeh Khajavikhan. Topological insulator laser: Experiments. *Science*, 359(6381):eaar4005, 2018.
- [7] Kunal Banerjee, Vishak Prasad C, Rishi Raj Gupta, Karthik Vyas, Anushree H, and Biswajit Mishra. Exploring Alternatives to Softmax Function. *arXiv e-prints*, page arXiv:2011.11538, November 2020.
- [8] Federico Battiston, Enrico Amico, Alain Barrat, Ginestra Bianconi, Guilherme Ferraz de Arruda, Benedetta Franceschiello, Iacopo Iacopini, Sonia Kéfi, Vito Latora, Yamir Moreno, Micah M. Murray, Tiago P. Peixoto, Francesco Vaccarino, and Giovanni Petri. The physics of higher-order interactions in complex systems. *Nature Physics*, 17(10):1093–1098, oct 2021.
- [9] Gregory Bentsen, Tomohiro Hashizume, Anton S Buyskikh, Emily J Davis, Andrew J Daley, Steven S Gubser, and Monika Schleier-Smith. Treelike interactions and fast scrambling with cold atoms. *Physical review letters*, 123(13):130601, 2019.

- [10] Christoph Bersch, Georgy Onishchukov, and Ulf Peschel. Experimental observation of sdpectral bloch oscillations. *Optics Letters*, 34(15):2372, July 2009.
- [11] Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. Quantum machine learning. *Nature*, 549(7671):195–202, 2017.
- [12] Åke Björck. Numerical methods for least squares problems, 1996.
- [13] Alexandre Blais, Arne L. Grimsmo, S.M. Girvin, and Andreas Wallraff. Circuit quantum electrodynamics. *Reviews of Modern Physics*, 93(2), May 2021.
- [14] Daniel J Blumenthal, Rene Heideman, Douwe Geuzebroek, Arne Leinse, and Chris Roeloffzen. Silicon nitride in silicon photonics. *Proceedings of the IEEE*, 106(12):2209–2231, 2018.
- [15] Mark Bohr. A 30 year retrospective on dennard’s mosfet scaling paper. *IEEE Solid-State Circuits Society Newsletter*, 12(1):11–13, 2007.
- [16] David J. Brady, Minghao Hu, Chengyu Wang, Xuefei Yan, Lu Fang, Yiwnheng Zhu, Yang Tan, Ming Cheng, and Zhan Ma. Smart cameras. 2020.
- [17] Earl T Campbell, Barbara M Terhal, and Christophe Vuillot. Roads towards fault-tolerant universal quantum computation. *Nature*, 549(7671):172–179, 2017.
- [18] M Cerezo, Guillaume Verdon, Hsin-Yuan Huang, Lukasz Cincio, and Patrick J Coles. Challenges and opportunities in quantum machine learning. *Nature Computational Science*, 2(9):567–576, 2022.
- [19] Hao Chen, NingNing Yang, Chengzhi Qin, Wenwan Li, Bing Wang, Tianwen Han, Chi Zhang, Weiwei Liu, Kai Wang, Hua Long, et al. Real-time observation of frequency Bloch oscillations with fibre loop modulation. *Light: Science & Applications*, 10(1):48, 2021.
- [20] Jiayin Chen, Hendra I. Nurdin, and Naoki Yamamoto. Temporal information processing on noisy quantum computers. *Phys. Rev. Appl.*, 14:024065, Aug 2020.
- [21] Kevin S Chou. *Teleported operations between logical qubits in circuit quantum electrodynamics*. PhD thesis, Yale University, 2018.
- [22] Andrew J Daley, Immanuel Bloch, Christian Kokail, Stuart Flannigan, Natalie

- Pearson, Matthias Troyer, and Peter Zoller. Practical quantum advantage in quantum simulation. *Nature*, 607(7920):667–676, 2022.
- [23] Jean Dalibard, Yvan Castin, and Klaus Mølmer. Wave-function approach to dissipative processes in quantum optics. *Phys. Rev. Lett.*, 68:580–583, Feb 1992.
- [24] Xiaowei Deng, Sai Li, Zi-Jie Chen, Zhongchu Ni, Yanyan Cai, Jiasheng Mai, Libo Zhang, Pan Zheng, Haifeng Yu, Chang-Ling Zou, Song Liu, Fei Yan, Yuan Xu, and Dapeng Yu. Heisenberg-limited quantum metrology using 100-photon fock states. 2023.
- [25] Asaf A. Diringer, Eliya Blumenthal, Avishay Grinberg, Liang Jiang, and Shay Hacohen-Gourgy. Conditional-not displacement: Fast multioscillator control with a single qubit. *Phys. Rev. X*, 14:011055, Mar 2024.
- [26] Akash V. Dixit, Srivatsan Chakram, Kevin He, Ankur Agrawal, Ravi K. Naik, David I. Schuster, and Aaron Chou. Searching for dark matter with a superconducting qubit. *Phys. Rev. Lett.*, 126:141302, Apr 2021.
- [27] Avik Dutt, Qian Lin, Luqi Yuan, Momchil Minkov, Meng Xiao, and Shanhui Fan. A single photonic cavity with two independent physical synthetic dimensions. *Science*, 367(6473):59–64, 2020.
- [28] Avik Dutt, Momchil Minkov, Qian Lin, Luqi Yuan, David AB Miller, and Shanhui Fan. Experimental band structure spectroscopy along a synthetic dimension. *Nature communications*, 10(1):3122, 2019.
- [29] Avik Dutt, Luqi Yuan, Ki Youl Yang, Kai Wang, Siddharth Buddhiraju, Jelena Vučković, and Shanhui Fan. Creating boundaries along a synthetic frequency dimension. *Nature Communications*, 13:3377, 12 2022.
- [30] Alessio D’Errico, Raouf Barboza, Rebeca Tudor, Alexandre Dauphin, Pietro Massignan, Lorenzo Marrucci, and Filippo Cardano. Bloch–Landau–Zener dynamics induced by a synthetic field in a photonic quantum walk. *APL Photonics*, 6(2):020802, 2021.
- [31] T Eichelkraut, R Heilmann, Steffen Weimann, Simon Stützer, Felix Dreisow, Demetrios N Christodoulides, Stefan Nolte, and Alexander Szameit. Mobility transition from ballistic to diffusive transport in non-Hermitian lattices. *Nature communications*, 4(1):2533, 2013.
- [32] Alec Eickbusch, Volodymyr Sivak, Andy Z. Ding, Salvatore S. Elder, Shan-

- tanu R. Jha, Jayameenakshi Venkatraman, Baptiste Royer, S. M. Girvin, Robert J. Schoelkopf, and Michel H. Devoret. Fast universal control of an oscillator with weak dispersive coupling to a qubit. *Nature Physics*, 18(12):1464–1469, October 2022.
- [33] Nicolas Englebert, Nathan Goldman, Miro Erkintalo, Nader Mostaan, Simon-Pierre Gorza, François Leo, and Julien Fatome. Bloch oscillations of coherently driven dissipative solitons in a synthetic dimension. *Nature Physics*, 19(7):1014–1021, April 2023.
- [34] Nicholas M. Fahrenkopf, Colin McDonough, Gerald L. Leake, Zhan Su, Erman Timurdogan, and Douglas D. Coolbaugh. The aim photonics mpw: A highly accessible cutting edge technology for rapid prototyping of photonic integrated circuits. *IEEE Journal of Selected Topics in Quantum Electronics*, 25(5):1–6, 2019.
- [35] Lingling Fan, Zhixin Zhao, Kai Wang, Avik Dutt, Jiahui Wang, Siddharth Budhiraju, Casey C Wojcik, and Shanhui Fan. Multidimensional convolution operation with synthetic frequency dimensions in photonics. *Physical Review Applied*, 10:34088, 2022.
- [36] T. Freeth, Y. Bitsakis, X. Moussas, J. H. Seiradakis, A. Tsvelikas, H. Mangou, M. Zafeiropoulou, R. Hadland, D. Bate, A. Ramsey, M. Allen, A. Crawley, P. Hockley, T. Malzbender, D. Gelb, W. Ambrisco, and M. G. Edmunds. Decoding the ancient greek astronomical calculator known as the antikythera mechanism. *Nature*, 444(7119):587–591, November 2006.
- [37] Xinian Fu. Architecture technology. In *A History of Chinese Science and Technology*, pages 1–194. Springer Berlin Heidelberg, Berlin, Heidelberg, 2015.
- [38] Keisuke Fujii and Kohei Nakajima. Harnessing disordered-ensemble quantum dynamics for machine learning. *Physical Review Applied*, 8(2):024030, 2017.
- [39] Keisuke Fujii and Kohei Nakajima. *Quantum Reservoir Computing: A Reservoir Approach Toward Quantum Machine Learning on Near-Term Quantum Devices*, pages 423–450. Springer Singapore, Singapore, 2021.
- [40] Jay Gambetta, Alexandre Blais, D. I. Schuster, A. Wallraff, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. Qubit-photon interactions in a cavity: Measurement-induced dephasing and number splitting. *Phys. Rev. A*, 74:042318, Oct 2006.
- [41] Paula García-Molina, Ana Martin, Mikel Garcia de Andoin, and Mikel Sanz. Noise in digital and digital-analog quantum computation. 2021.

- [42] Daniel J. Gauthier, Erik Bollt, Aaron Griffith, and Wendson A. S. Barbosa. Next generation reservoir computing. *Nature Communications*, 12(1), September 2021.
- [43] Sanjib Ghosh, Andrzej Opala, Michał Matuszewski, Tomasz Paterek, and Timothy C. H. Liew. Quantum reservoir processing. *npj Quantum Information*, 5(1), apr 2019.
- [44] L. C. G. Govia, G. J. Ribeill, G. E. Rowlands, H. K. Krovi, and T. A. Ohki. Quantum reservoir computing with a single nonlinear oscillator. *Phys. Rev. Res.*, 3:013077, Jan 2021.
- [45] Christoph W Groth, Michael Wimmer, Anton R Akhmerov, and Xavier Waintal. Kwant: a software package for quantum transport. *New Journal of Physics*, 16:063065, 6 2014.
- [46] Steven S Gubser, Christian Jepsen, Ziming Ji, and Brian Trundy. Continuum limits of sparse coupling patterns. *Physical Review D*, 98(4):045009, 2018.
- [47] F. D. M. Haldane. Model for a quantum Hall effect without Landau levels: Condensed-matter realization of the “Parity Anomaly”. *Physical Review Letters*, 61(18):2015–2018, October 1988.
- [48] Nicholas C Harris, Gregory R Steinbrecher, Mihika Prabhu, Yoav Lahini, Jacob Mower, Darius Bunandar, Changchen Chen, Franco NC Wong, Tom Baehr-Jones, Michael Hochberg, et al. Quantum transport simulations in a programmable nanophotonic processor. *Nature Photonics*, 11(7):447–452, 2017.
- [49] Naomichi Hatano and David R. Nelson. Localization transitions in non-hermitian quantum mechanics. *Phys. Rev. Lett.*, 77:570–573, Jul 1996.
- [50] Reinier W. Heeres, Brian Vlastakis, Eric Holland, Stefan Krastanov, Victor V. Albert, Luigi Frunzio, Liang Jiang, and Robert J. Schoelkopf. Cavity state manipulation using photon-number selective phase gates. *Physical Review Letters*, 115(13), September 2015.
- [51] John L Hennessy and David A Patterson. *Computer Architecture*. The Morgan Kaufmann Series in Computer Architecture and Design. Morgan Kaufmann, Oxford, England, 5 edition, September 2011.
- [52] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. 2020.

- [53] Hossein Hodaei, Absar U. Hassan, Steffen Wittek, Hipolito Garcia-Gracia, Ramy El-Ganainy, Demetrios N. Christodoulides, and Mercedeh Khajavikhan. Enhanced sensitivity at higher-order exceptional points. *Nature*, 548:187–191, 8 2017.
- [54] Mohammad P. Hokmabadi, Alexander Schumer, Demetrios N. Christodoulides, and Mercedeh Khajavikhan. Non-hermitian ring laser gyroscopes with enhanced sagnac sensitivity. *Nature*, 576:70–74, 12 2019.
- [55] Fangjun Hu, Gerasimos Angelatos, Saeed A. Khan, Marti Vives, Esin Türeci, Leon Bello, Graham E. Rowlands, Guilhem J. Ribeill, and Hakan E. Türeci. Tackling sampling noise in physical systems for machine learning applications: Fundamental limits and eigentasks. *Phys. Rev. X*, 13:041020, Oct 2023.
- [56] Yaowen Hu, Christian Reimer, Amirhassan Shams-Ansari, Mian Zhang, and Marko Lončar. Realization of high-dimensional frequency crystals in electro-optic microcombs. *Optica*, 7:1189, 9 2020.
- [57] Fei Hua, Yuwei Jin, Yanhao Chen, Suhas Vittal, Kevin Krsulich, Lev S. Bishop, John Lapeyre, Ali Javadi-Abhari, and Eddy Z. Zhang. Exploiting qubit reuse through mid-circuit measurement and reset. 2023.
- [58] Jimmy S. C. Hung, J. H. Busnaina, C. W. Sandbo Chang, A. M. Vadiraj, I. Nsanzineza, E. Solano, H. Alaeian, E. Rico, and C. M. Wilson. Quantum simulation of the bosonic Creutz ladder with a parametric cavity. *Physical Review Letters*, 127:100503, 9 2021.
- [59] Herbert Jaeger. The “echo state” approach to analysing and training recurrent neural networks-with an erratum note. *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report*, 148(34):13, 2001.
- [60] Herbert Jaeger, Mantas Lukoševičius, Dan Popovici, and Udo Siewert. Optimization and applications of echo state networks with leaky- integrator neurons. *Neural Networks*, 20(3):335–352, 2007. Echo State Networks and Liquid State Machines.
- [61] Anu Jagannath and Jithin Jagannath. Dataset for modulation classification and signal type classification for multi-task and single task learning. *Computer Networks*, 199:108441, 2021.
- [62] Á. Jiménez-Galán, R. E. F. Silva, O. Smirnova, and M. Ivanov. Lightwave control of topological properties in 2d materials for sub-cycle and non-resonant valley manipulation. *Nature Photonics*, 14(12):728–732, November 2020.

- [63] J.R. Johansson, P.D. Nation, and Franco Nori. Qutip: An open-source python framework for the dynamics of open quantum systems. *Computer Physics Communications*, 183(8):1760–1772, August 2012.
- [64] Allan Jones, David Chapman, Helen Donelan, Laurence Dooley, and Adrian Poulton. Exploring communications technology – open university, 2017.
- [65] N Cody Jones, Rodney Van Meter, Austin G Fowler, Peter L McMahon, Jungsang Kim, Thaddeus D Ladd, and Yoshihisa Yamamoto. Layered architecture for quantum computing. *Physical Review X*, 2(3):031007, 2012.
- [66] W. D. Kalfus, G. J. Ribeill, G. E. Rowlands, H. K. Krovi, T. A. Ohki, and L. C. G. Govia. Hilbert space as a computational resource in reservoir computing. *Phys. Rev. Res.*, 4:033007, Jul 2022.
- [67] Amir H Karamlou, Jochen Braumüller, Yariv Yanay, Agustin Di Paolo, Patrick M Harrington, Bharath Kannan, David Kim, Morten Kjaergaard, Alexander Melville, Sarah Muschinske, et al. Quantum transport and localization in 1d and 2d tight-binding lattices. *npj Quantum Information*, 8(1):35, 2022.
- [68] Saeed Ahmed Khan, Fangjun Hu, Gerasimos Angelatos, and Hakan E. Türeci. Physical reservoir computing using finitely-sampled quantum systems. 2021.
- [69] Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. *arXiv e-prints*, page arXiv:1412.6980, December 2014.
- [70] Tobias J Kippenberg, Alexander L Gaeta, Michal Lipson, and Michael L Gorodetsky. Dissipative Kerr solitons in optical microresonators. *Science*, 361(6402):eaan8083, 2018.
- [71] 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.
- [72] Tomoyuki Kubota, Yudai Suzuki, Shumpei Kobayashi, Quoc Hoan Tran, Naoki Yamamoto, and Kohei Nakajima. Quantum noise-induced reservoir computing. 2022.
- [73] Thaddeus D Ladd, Fedor Jelezko, Raymond Laflamme, Yasunobu Nakamura,

Christopher Monroe, and Jeremy Lloyd O'Brien. Quantum computers. *nature*, 464(7285):45–53, 2010.

- [74] Y Lahini, R Pugatch, F Pozzi, M Sorel, R Morandotti, N Davidson, and Y Silberberg. Observation of a localization transition in quasiperiodic photonic lattices. *Physical Review Letters*, 103:013901, 6 2009.
- [75] Nathan RA Lee, Marek Pechal, E Alex Wollack, Patricio Arrangoiz-Arriola, Zhaoyou Wang, and Amir H Safavi-Naeini. Propagation of microwave photons along a synthetic dimension. *Physical Review A*, 101(5):053807, 2020.
- [76] Christian Leefmans, Avik Dutt, James Williams, Luqi Yuan, Midya Parto, Franco Nori, Shanhui Fan, and Alireza Marandi. Topological dissipation in a time-multiplexed photonic resonator network. *Nature Physics*, 18(4):442–449, February 2022.
- [77] Guangzhen Li, Luoja Wang, Rui Ye, Shijie Liu, Yuanlin Zheng, Luqi Yuan, and Xianfeng Chen. Observation of flat-band and band transition in the synthetic space. *Advanced Photonics*, 4(3):036002, 2022.
- [78] Guangzhen Li, Yuanlin Zheng, Avik Dutt, Danying Yu, Qingrou Shan, Shijie Liu, Luqi Yuan, Shanhui Fan, and Xianfeng Chen. Dynamic band structure measurement in the synthetic space. *Science advances*, 7(2):eabe4335, 2021.
- [79] Qian Lin, Meng Xiao, Luqi Yuan, and Shanhui Fan. Photonic Weyl point in a two-dimensional resonator lattice with a synthetic frequency dimension. *Nature communications*, 7(1):13731, 2016.
- [80] Mantas Lukoševičius and Herbert Jaeger. Reservoir computing approaches to recurrent neural network training. *Computer science review*, 3(3):127–149, 2009.
- [81] Eran Lustig and Mordechai Segev. Topological photonics in synthetic dimensions. *Adv. Opt. Photon.*, 13(2):426–461, Jun 2021.
- [82] Eran Lustig, Steffen Weimann, Yonatan Plotnik, Yaakov Lumer, Miguel A Bandres, Alexander Szameit, and Mordechai Segev. Photonic topological insulator in synthetic dimensions. *Nature*, 567:356–360, 3 2019.
- [83] K. F. Mak, K. L. McGill, J. Park, and P. L. McEuen. The valley Hall effect in MoS₂ transistors. *Science*, 344(6191):1489–1492, June 2014.

- [84] Carlos Ortiz Marrero, Mária Kieferová, and Nathan Wiebe. Entanglement-induced barren plateaus. *PRX Quantum*, 2(4):040316, 2021.
- [85] Rodrigo Martínez-Peña, Gian Luca Giorgi, Johannes Nokkala, Miguel C. Soriano, and Roberta Zambrini. Dynamical phase transitions in quantum reservoir computing. *Phys. Rev. Lett.*, 127:100502, Aug 2021.
- [86] 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):4812, 2018.
- [87] Peter L. McMahon. The physics of optical computing. *Nature Reviews Physics*, 5(12):717–734, October 2023.
- [88] Hirokazu Miyake, Georgios A Siviloglou, Colin J Kennedy, William Cody Burton, and Wolfgang Ketterle. Realizing the Harper Hamiltonian with laser-assisted tunneling in optical lattices. *Physical review letters*, 111(18):185302, 2013.
- [89] Zoubeir Mlika, Soumaya Cherkaoui, Jean Frédéric Laprade, and Simon Corbeil-Letourneau. User trajectory prediction in mobile wireless networks using quantum reservoir computing. 2023.
- [90] Masoud Mohseni, Peter Read, Hartmut Neven, Sergio Boixo, Vasil Denchev, Ryan Babbush, Austin Fowler, Vadim Smelyanskiy, and John Martinis. Commercialize quantum technologies in five years. *Nature*, 543(7644):171–174, 2017.
- [91] Naeimeh Mohseni, Peter L McMahon, and Tim Byrnes. Ising machines as hardware solvers of combinatorial optimization problems. *Nature Reviews Physics*, 4(6):363–379, 2022.
- [92] André L. M. Muniz, Martin Wimmer, Arstan Bisianov, Roberto Morandotti, and Ulf Peschel. Collapse on the line – how synthetic dimensions influence nonlinear effects. *Scientific Reports*, 9:9518, 12 2019.
- [93] Kohei Nakajima. Reservoir computing: theory, physical implementations, and applications. *IEICE Technical Report; IEICE Tech. Rep.*, 118(220):149–154, 2018.
- [94] Tomoki Ozawa, Hannah M. Price, Alberto Amo, Nathan Goldman, Mohammad Hafezi, Ling Lu, Mikael C. Rechtsman, David Schuster, Jonathan Simon, Oded Zilberberg, and Iacopo Carusotto. Topological photonics. *Reviews of Modern Physics*, 91:015006, 3 2019.

- [95] Tomoki Ozawa, Hannah M. Price, Nathan Goldman, Oded Zilberberg, and Iacopo Carusotto. Synthetic dimensions in integrated photonics: From optical isolation to four-dimensional quantum Hall physics. *Physical Review A*, 93:043827, 4 2016.
- [96] Timothy James O’Shea, Tamoghna Roy, and T Charles Clancy. Over-the-air deep learning based radio signal classification. *IEEE Journal of Selected Topics in Signal Processing*, 12(1):168–179, 2018.
- [97] Sunil Pai, Taewon Park, Marshall Ball, Bogdan Penkovsky, Michael Dubrovsky, Nathnael Abebe, Maziyar Milanizadeh, Francesco Morichetti, Andrea Melloni, Shanhui Fan, et al. Experimental evaluation of digitally verifiable photonic computing for blockchain and cryptocurrency. *Optica*, 10(5):552–560, 2023.
- [98] Adrian Parra-Rodriguez, Pavel Lougovski, Lucas Lamata, Enrique Solano, and Mikel Sanz. Digital-analog quantum computation. *Phys. Rev. A*, 101:022305, Feb 2020.
- [99] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Köpf, Edward Yang, Zach DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. PyTorch: An Imperative Style, High-Performance Deep Learning Library. *arXiv e-prints*, page arXiv:1912.01703, December 2019.
- [100] Avikar Periwal, Eric S. Cooper, Philipp Kunkel, Julian F. Wienand, Emily J. Davis, and Monika Schleier-Smith. Programmable interactions and emergent geometry in an array of atom clouds. *Nature*, 600(7890):630–635, December 2021.
- [101] Raymond Perrault, Yoav Shoham, Erik Brynjolfsson, Jack Clark, John Etchemendy, Barbara Grosz, Terah Lyons, James Manyika, Saurabh Mishra, and Juan Carlos Niebles. “the ai index 2019 annual report”. *AI Index Steering Committee*, 2019.
- [102] Christopher W Peterson, Wladimir A Benalcazar, Mao Lin, Taylor L Hughes, and Gaurav Bahl. Strong nonreciprocity in modulated resonator chains through synthetic electric and magnetic fields. *Physical review letters*, 123(6):063901, 2019.
- [103] Philipp Pfeffer, Florian Heyder, and Jörg Schumacher. Hybrid quantum-classical reservoir computing of thermal convection flow. *Physical Review Research*, 4(3), September 2022.

- [104] John Preskill. Quantum computing in the nisq era and beyond. *Quantum*, 2:79, August 2018.
- [105] Chengzhi Qin, Luqi Yuan, Bing Wang, Shanhui Fan, and Peixiang Lu. Effective electric-field force for a photon in a synthetic frequency lattice created in a waveguide modulator. *Physical Review A*, 97(6):063838, 2018.
- [106] Chengzhi Qin, Feng Zhou, Yugui Peng, Dimitrios Sounas, Xuefeng Zhu, Bing Wang, Jianji Dong, Xinliang Zhang, Andrea Alù, and Peixiang Lu. Spectrum control through discrete frequency diffraction in the presence of photonic gauge potentials. *Physical Review Letters*, 120(13):133901, 2018.
- [107] Antoine Royer. Wigner function as the expectation value of a parity operator. *Phys. Rev. A*, 15:449–450, Feb 1977.
- [108] Antonio Sannia, Rodrigo Martínez-Peña, Miguel C. Soriano, Gian Luca Giorgi, and Roberta Zambrini. Dissipation as a resource for Quantum Reservoir Computing. *Quantum*, 8:1291, March 2024.
- [109] Simone Scardapane and Aurelio Uncini. Semi-supervised echo state networks for audio classification. *Cognitive Computation*, 9:125–135, 2017.
- [110] Benjamin Schrauwen, David Verstraeten, and Jan Van Campenhout. An overview of reservoir computing: theory, applications and implementations. In *Proceedings of the 15th european symposium on artificial neural networks*. p. 471-482 2007, pages 471–482, 2007.
- [111] Maria Schuld and Nathan Killoran. Is quantum advantage the right goal for quantum machine learning? *PRX Quantum*, 3(3):030101, 2022.
- [112] D. I. Schuster, A. A. Houck, J. A. Schreier, A. Wallraff, J. M. Gambetta, A. Blais, L. Frunzio, J. Majer, B. Johnson, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. Resolving photon number states in a superconducting circuit. *Nature*, 445(7127):515–518, February 2007.
- [113] Alon Schwartz and Baruch Fischer. Laser mode hyper-combs. *Optics Express*, 21:6196, 3 2013.
- [114] Zhimin Shi, Ksenia Dolgaleva, and Robert W Boyd. Quantum noise properties of non-ideal optical amplifiers and attenuators. *Journal of Optics*, 13(12):125201, November 2011.

- [115] Michele Spagnolo, Joshua Morris, Simone Piacentini, Michael Antesberger, Francesco Massa, Andrea Crespi, Francesco Ceccarelli, Roberto Osellame, and Philip Walther. Experimental photonic quantum memristor. *Nature Photonics*, 16(4):318–323, March 2022.
- [116] Chenxi Sun, Moxian Song, Derun Cai, Baofeng Zhang, Shenda Hong, and Hongyan Li. A systematic review of echo state networks from design to application. *IEEE Transactions on Artificial Intelligence*, pages 1–15, 2022.
- [117] Yudai Suzuki, Qi Gao, Ken C. Pradel, Kenji Yasuoka, and Naoki Yamamoto. Natural quantum reservoir computing for temporal information processing. *Scientific Reports*, 12(1), January 2022.
- [118] Aleksandr K Tusnin, Alexey M Tikan, and Tobias J Kippenberg. Nonlinear states and dynamics in a synthetic frequency dimension. *Physical Review A*, 102(2):023518, 2020.
- [119] Christopher S. Wang, Jacob C. Curtis, Brian J. Lester, Yaxing Zhang, Yvonne Y. Gao, Jessica Freeze, Victor S. Batista, Patrick H. Vaccaro, Isaac L. Chuang, Luigi Frunzio, Liang Jiang, S. M. Girvin, and Robert J. Schoelkopf. Efficient multi-photon sampling of molecular vibronic spectra on a superconducting bosonic processor. *Phys. Rev. X*, 10:021060, Jun 2020.
- [120] Kai Wang, Avik Dutt, Charles C. Wojcik, and Shanhui Fan. Topological complex-energy braiding of non-Hermitian bands. *Nature*, 598(7879):59–64, October 2021.
- [121] Kai Wang, Avik Dutt, Ki Youl Yang, Casey C Wojcik, Jelena Vučković, and Shanhui Fan. Generating arbitrary topological windings of a non-Hermitian band. *Science*, 371:1240–1245, 2021.
- [122] Peng Wang, Yuanlin Zheng, Xianfeng Chen, Changming Huang, Yaroslav V Kartashov, Lluis Torner, Vladimir V Konotop, and Fangwei Ye. Localization and delocalization of light in photonic moiré lattices. *Nature*, 577(7788):42–46, 2020.
- [123] Samson Wang, Enrico Fontana, Marco Cerezo, Kunal Sharma, Akira Sone, Lukasz Cincio, and Patrick J Coles. Noise-induced barren plateaus in variational quantum algorithms. *Nature communications*, 12(1):6961, 2021.
- [124] W Wang, Z-J Chen, X Liu, W Cai, Y Ma, X Mu, X Pan, Z Hua, L Hu, Y Xu, et al. Quantum-enhanced radiometry via approximate quantum error correction. *Nature Communications*, 13(1):3214, 2022.

- [125] Zhixin Wang, Mingrui Xu, Xu Han, Wei Fu, Shruti Puri, SM Girvin, Hong X Tang, S Shankar, and MH Devoret. Quantum microwave radiometry with a superconducting qubit. *Physical Review Letters*, 126(18):180501, 2021.
- [126] Sebastian Weidemann, Mark Kremer, Stefano Longhi, and Alexander Szameit. Coexistence of dynamical delocalization and spectral localization through stochastic dissipation. *Nature Photonics*, 15(8):576–581, 2021.
- [127] Logan G. Wright, Demetrios N. Christodoulides, and Frank W. Wise. Spatiotemporal mode-locking in multimode fiber lasers. *Science*, 358:94–97, 10 2017.
- [128] Logan G. Wright and Peter L. McMahon. The capacity of quantum neural networks. 2019.
- [129] Logan G. Wright, Tatsuhiro Onodera, Martin M. Stein, Tianyu Wang, Darren T. Schachter, Zoey Hu, and Peter L. McMahon. Deep physical neural networks trained with backpropagation. *Nature*, 601(7894):549–555, January 2022.
- [130] Toshiki Yasuda, Yudai Suzuki, Tomoyuki Kubota, Kohei Nakajima, Qi Gao, Wenlong Zhang, Satoshi Shimono, Hendra I. Nurdin, and Naoki Yamamoto. Quantum reservoir computing with repeated measurements on superconducting devices. 2023.
- [131] Danying Yu, Bo Peng, Xianfeng Chen, Xiong-Jun Liu, and Luqi Yuan. Topological holographic quench dynamics in a synthetic frequency dimension. *Light: Science & Applications*, 10(1):209, 2021.
- [132] Luqi Yuan, Avik Dutt, and Shanhui Fan. Synthetic frequency dimensions in dynamically modulated ring resonators. *APL Photonics*, 6(7):071102, 2021.
- [133] Luqi Yuan and Shanhui Fan. Three-dimensional dynamic localization of light from a time-dependent effective gauge field for photons. *Physical Review Letters*, 114(24):243901, 2015.
- [134] Luqi Yuan and Shanhui Fan. Bloch oscillation and unidirectional translation of frequency in a dynamically modulated ring resonator. *Optica*, 3(9):1014–1018, 2016.
- [135] Luqi Yuan, Meng Xiao, Qian Lin, and Shanhui Fan. Synthetic space with arbitrary dimensions in a few rings undergoing dynamic modulation. *Physical Review B*, 97(10):104105, 2018.

- [136] Han-Sen Zhong, Hui Wang, Yu-Hao Deng, Ming-Cheng Chen, Li-Chao Peng, Yi-Han Luo, Jian Qin, Dian Wu, Xing Ding, Yi Hu, Peng Hu, Xiao-Yan Yang, Wei-Jun Zhang, Hao Li, Yuxuan Li, Xiao Jiang, Lin Gan, Guangwen Yang, Lixing You, Zhen Wang, Li Li, Nai-Le Liu, Chao-Yang Lu, and Jian-Wei Pan. Quantum computational advantage using photons. *Science*, 370(6523):1460–1463, December 2020.
- [137] Feichi Zhou and Yang Chai. Near-sensor and in-sensor computing. *Nature Electronics*, 3(11):664–671, November 2020.
- [138] Feichi Zhou and Yang Chai. Near-sensor and in-sensor computing. *Nature Electronics*, 3(11):664–671, 2020.