Analyzing Properties of Driven-Dissipative Quantum Systems from the Mean Field to Machine Learning

by

K. Tucker

B.S., University of Colorado, 2007

M.S., University of Colorado, 2007

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Applied Math

2023

Committee Members:

Ana Maria Rey, Chair

Juan Restrepo

Claire Monteleoni

Tameem Albash

James Meiss

Tucker, K. (Ph.D., Applied Math)

Analyzing Properties of Driven-Dissipative Quantum Systems from the Mean Field to Machine Learning

Thesis directed by Prof. Ana Maria Rey

Quantum systems are poised to play a large role in the emerging technologies of the near future. The areas of application are numerous, with major advances being achieved in recent years in the areas of metrology, sensing, and computing to name just a few. In order for these technologies to realize their full potential, significant challenges must be overcome in the areas of classical simulation, error correction, and device characterization.

In this thesis, we explore a number of approaches to tackling these challenges in the simulation and characterization of driven-dissipative quantum systems. We examine existing techniques for simulating these systems to make new observations about their properties that have potential applications in quantum metrology and sensing. We investigate the emergence of a time crystal in a system of two-level atoms and the conditions under which it arises, as well as the effect of single particle relaxation on spin squeezing in a similar system. We also propose a novel machine learning model for estimating physical parameters, with potential applications for detecting crosstalk in quantum information processors.

Dedication

To all my friends and family who believed in me throughout this long journey

Acknowledgements

I have been very fortunate to have the support of many kind, generous, and brilliant people and this work would not have been possible without them. I want to thank my advisors Ana Maria Rey and Juan Restrepo, who have been so generous with their support, time, and expertise. I also want to thank the other members of my committee, Claire Monteleoni, Tameem Albash, and James Meiss. I've learned so much from you and I know that knowledge will carry me forward throughout my entire career. I'm also extremely grateful for the other members of my research groups, especially the tremendous contributions of Jamir Marino, Robert Lewis-Swan, and Diego Barberena. I consider myself incredibly fortunate to have had the opportunity to work with all of you, and these years have truly been the realization of a dream.

This work would also not have been possible without the support of my colleagues and the leadership at NASK Incorporated. I feel very fortunate to be a part of the important work that we do and to have had the opportunity to continue that work during this process.

Finally, I want to thank my family and friends, most especially my mother Vicki Tucker, my father Gary Tucker, and my stepmother Cheryl Tucker. I wouldn't be where I am without your unwavering love and support, and for your continuing reminders of what is most important in life.

Contents

Chapter

Tables

Table

Figures

Figure

Chapter 1

Introduction

Quantum systems are poised to play a large role in the emerging technologies of the near future. The areas of application are numerous, with major advances being achieved in recent years in the areas of metrology, sensing, and computing to name just a few. When Google demonstrated (with some controversy) quantum supremacy in 2019 [6], an article in Nature likened it to the Wright brothers first flight [133]. In his keynote address at the Quantum Computing for Business conference in 2017, Caltech quantum information scientist John Preskill anticipated this, declaring that the noisy intermediate scale quantum (NISQ) technology era, characterized by computers in the 50-100 qubit range, was imminent [140]. He also sounded a note of caution, however, as significant challenges still remain to make these promising technologies a practical reality.

The concept of entanglement at the quantum scale can be a powerful resource. John Bell demonstrated in 1964 [14] that quantum information is encoded in non-local correlations between components of a quantum system, something that has no counterpart in a classical system [139]. Making use of this resource, however, presents many challenges. The experimentalist is forced to confront the fragility of these correlations, as they can easily be destroyed by interactions between the quantum system and its surrounding environment. For the theorist, the challenge lies in the curse of dimensionality, as the same complexity that makes a quantum system so potentially useful also demands computing resources that generally scale exponentially with the system size when attempting to simulate the system on a classical computer, making a full attempt at simulation impossible for systems of sizes greater than roughly 10 qubits. Another challenge is the inherently

stochastic nature of quantum systems, as any measurement made on the system necessarily modifies it to some degree, and this stochasticity must accounted for in any classical model.

In this thesis, we explore a number of approaches to tackling the challenges in the classical simulation and characterization of quantum systems. Chapter 2 introduces many of the needed physical concepts at the level necessary for someone interested in simulating, though not necessarily building, driven-dissipative quantum systems. Chapters 4 and 5 employ existing techniques for simulating these systems to make new observations about their properties that have potential applications in quantum metrology and sensing.

In chapter 3, we introduce the background of a number of machine learning techniques that are changing the way classical simulation and characterization of quantum systems is being performed. In chapter 6, we make use of these techniques to propose a novel machine learning approach to the characterization of quantum systems, in particular the problem of learning how to extract physical parameters from measurement data. A potential application of this approach is the discovery and quantification of qubit crosstalk in a quantum information processor, a phenomenon that must be detected and mitigated as part of quantum error correction. While the application to quantum systems is both timely and relevant, it is important to note that this approach can be applied to any system where the goal is to discover the parameters of a stochastic differential equation from data.

The following chapters are adapted from the indicated publications. Each one is a slightly longer form that incorporates particularly relevant elements of the supplemental material in the main text, or is a draft of a manuscript to be submitted for publication:

> **Chapter 4**: K. Tucker, B. Zhu, R. J. Lewis-Swan, J. Marino, F. Jimenez, J. G. Restrepo, and A. M. Rey. "Shattered time: can a dissipative time crystal survive many-body correlations?" New Journal of Physics, 20(12):123003, (2018)
>
> **Chapter 5**: K. Tucker, D. Barberena, R. J. Lewis-Swan, J. K Thompson, J. G. Restrepo, A. M. Rey, "Facilitating spin squeezing generated by collective dynamics with single-particle

decoherence", Phys. Rev. A 102, 051701(R) (2020)

Chapter 6: This chapter is a draft of a manuscript to be submitted for publication with authors K. Tucker, A. K. Rege, C. Smith, C. Monteleoni, and T. Albash

Chapter 2

Background - Driven-Dissipative Quantum Systems

In this section, we cover the fundamental physical concepts needed to understand the subsequent chapters of this thesis. This includes the necessary fundamentals of quantum states, the differential equations describing their dynamics, and common computational approaches for the classical simulation of open quantum systems that address the curse of dimensionality arising from the exponential growth of the quantum state with the system size.

2.1 Open Quantum Systems

While the traditional entry point for quantum mechanics states the axioms in terms of a quantum state for a closed system as an element $|\psi\rangle$ of a Hilbert space, all systems discussed in this thesis will be open in the sense that they involve interaction with an environment that cannot be directly modeled or measured, as shown in figure 2.1. We will, therefore, use as a starting point for a quantum state the density operator ρ and proceed from postulates stated in terms of open quantum systems, as is done in section 2.4 of [128].

2.1.1 State Representation and Measurement

The state space of a quantum system is a complex Hilbert space $\mathcal{H}$, and the state of the system is completely described by the density operator ρ, a Hermitian, positive semi-definite operator on the Hilbert space with trace one. Dynamics of the closed system are governed entirely by unitary

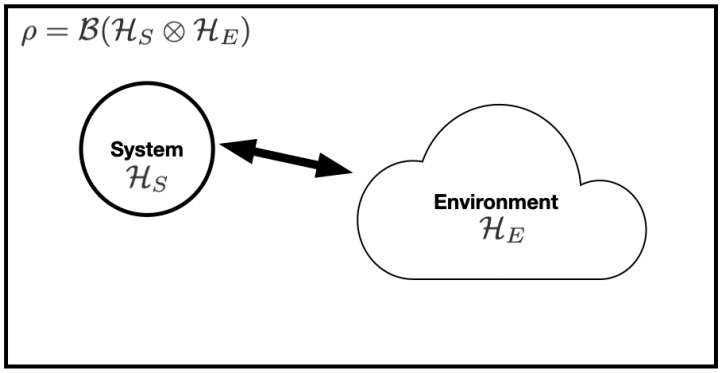


Figure 2.1: An open quantum system is one in which interactions occur with an environment that cannot be measured

transformations where

$$\rho' = U\rho U^\dagger. \tag{2.1}$$

This is equivalent to the Schrödinger equation for ρ given by

$$\frac{d\rho}{dt} = -i[H, \rho] \tag{2.2}$$

which we have normalized such that the Plank constant $\hbar = 1$. Here the Hermitian operator H is the Hamiltonian for the system and $[A, B] = AB - BA$ is the commutator. Dynamics for an open system, that is a system interacting with an environment, are governed by the Lindblad master equation to be discussed in section 2.1.3.

Measurements with N possible outcomes correspond to measurement operators $\{A_n\}_{n=1}^N$ such that

$$\sum_{n=1}^{N} A_n^\dagger A_n = \mathbb{1}. \tag{2.3}$$

If the system is in state ρ, then the probability of measuring outcome n is given by

$$P(n) = \mathrm{Tr}\left[A_n^\dagger A_n \rho\right], \tag{2.4}$$

and the state of the system immediately after measurement is

$$\frac{A_n \rho A_n^\dagger}{\mathrm{Tr}\left[A_n \rho A_n^\dagger\right]}. \tag{2.5}$$

The set of non-negative operators $\left\{E_n \equiv A_n^\dagger A_n\right\}_{n=1}^N$ is known as a positive operator valued measure (POVM).

An observable O in quantum mechanics corresponds to a Hermitian operator on the Hilbert space, with values corresponding to eigenvalues and measurement operators A_n corresponding to orthogonal projections onto the eigenspace of each distinct eigenvalue λ_n. Note that eigenspaces with distinct eigenvalues will be orthogonal due to the Hermiticity of O. The spectral theorem gives us

$$O = \sum_n \lambda_n A_n,$$

so that the expected value of the observable

$$\langle O \rangle \equiv \sum_n \lambda_n \mathrm{Tr}\left[A_n^\dagger A_n \rho\right] = \mathrm{Tr}\left[\rho \sum_n \lambda_n A_n\right] = \mathrm{Tr}\left[\rho O\right], \tag{2.6}$$

where we have used $A_n^\dagger A_n = A_n$ since A_n is an orthogonal projection.

Finally, the state space for a composite of physical systems is given by the tensor product of individual state spaces $\mathcal{H} = \bigotimes_i \mathcal{H}_i$. A density operator on $\mathcal{H}$ that can be written as a convex combination of tensor products of density operators on the subsystems, i.e.

$$\rho = \sum_k p_k \bigotimes_i \rho_{k,i},$$

where $\rho_{k,i}$ is a density on subsystem i and the p_k sum to one, is said to separable. If it cannot be written this way, it is said to be entangled.

2.1.2 Two Level Atom

The simplest possible quantum system is one with a Hilbert space of dimension two. Such a system is often referred to as a qubit, and corresponds to the physical system of the two-level atom. The set of all Hermitian operators on this Hilbert space is spanned by the identity along with the mutually orthonormal (under the trace inner product $\langle A, B \rangle \equiv \mathrm{Tr}\left[A^\dagger B\right]$) operators

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \; \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \; \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.7}$$

These are referred to as the Pauli matrices. They can also simply be denoted as (X, Y, Z) depending on the context.

Given that these operators form an orthonormal basis for all Hermitian operators, the density operator can be written as

$$\rho = \frac{1}{2}\left(\mathbb{1} + \vec{b} \cdot \vec{\sigma}\right), \tag{2.8}$$

where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ is an operator valued vector, $\vec{b}$ is known as the Bloch vector. The set of all Bloch vectors of norm one is known as the Bloch sphere, and the set with norm less than one is known as the Bloch ball. These two sets define all possible states for a two level atom. States on the sphere are rank one and known as pure states, with states in the interior referred to as mixed as they cannot be expressed as an outer product of a single state space element with itself, but rather are given by a statistical mixture. Figure 2.2 shows the Bloch sphere along with the Bloch vector labeled with its cylindrical components (R, ϕ, s).

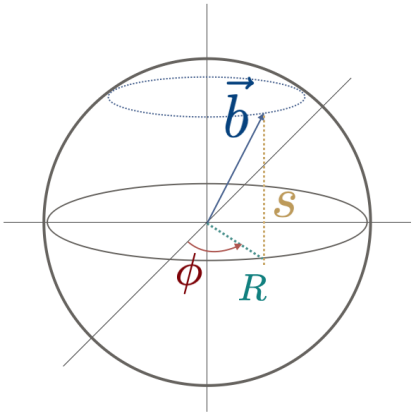


Figure 2.2: The Bloch Sphere

The Hamiltonian of a system is the observable corresponding to energy. In the case of a two-level atom, we can define the difference between the two energy levels of the atom as $\hbar\omega_a \equiv E_e - E_g$, where E_e is the energy of the excited state and E_g the energy of the ground state. The Hamiltonian is then the operator with a difference in eigenvalues of $\hbar\omega_a$. By convention, this is often represented

simply in terms of the Pauli matrix σ_z as

$$H_a = \frac{\hbar\omega_a}{2}\sigma_z. \tag{2.9}$$

This operator has eigenvalues $\pm\hbar\omega_a/2$ so that the energies have the desired splitting $E_e - E_g$ centered at zero. The corresponding eigenvectors are often labeled $|e\rangle \equiv |0\rangle$ and $|g\rangle \equiv |1\rangle$, the excited and ground state, respectively, and choosing a vector representation of the Hilbert space where

$$|e\rangle \equiv \begin{pmatrix} 1 \\ \\ 0 \end{pmatrix} \text{ and } |g\rangle \equiv \begin{pmatrix} 0 \\ \\ 1 \end{pmatrix}$$

is common and gives this choice of basis for the state space the name "computational" basis. The term ω_a is known as the atomic transition frequency where the subscript "a" is for "atom."

2.1.3 Lindblad Master Equation

As discussed above, the density operator is defined as a Hermitian, positive semi-definite operator with trace one. Any operator that governs the dynamics of a quantum system must, therefore, preserve these properties. We refer to such an operator as a Completely Positive Trace-Preserving (CPTP) map. Making a Markov assumption in time, meaning that the state $\rho(t+dt)$ can be written in terms of $\rho(t)$ without direct dependence on $\rho(s)$ where $s < t$, then the most general differential equation governing the dynamics of the system can be written as

$$\frac{d\rho}{dt} = -i[H,\rho] + \sum_{k=1}^{K} \mathcal{D}[O_k](\rho), \tag{2.10}$$

Each $\mathcal{D}[O_k](\rho)$ is a super-operator acting on ρ defined as

$$\mathcal{D}[O](\rho) \equiv O\rho O^\dagger - \frac{1}{2}\left\{O^\dagger O, \rho\right\}, \tag{2.11}$$

where $\{A,B\} = AB + BA$ is the anti-commutator. The $\mathcal{D}[O_k]$ are called Lindblad super-operators, and equation (2.10) is the Lindblad master equation. Note that it is entirely linear in the elements of ρ, and can therefore be written as

$$\frac{d\rho}{dt} = \mathcal{L}\rho, \tag{2.12}$$

where the linear operator $\mathcal{L}$ is referred to as the Liouvillian.

Equation (2.10) is an extension of (2.2) that captures the interaction with the environment in the Lindblad terms. This evolution is mathematically equivalent to allowing a unitary evolution of the composite system including the physical system and the environment, then performing a projective measurement on the environment at each time step, but without knowledge of the measurement record. This interpretation will be important when we discuss quantum trajectories in section 2.4.2.

2.2 Approaches for Classical Simulation

As discussed in section 2.1.1, the density operator ρ for a system of N two-level atoms is an operator on the 2^N-dimensional Hilbert space $\mathcal{H} = \bigotimes_{n=1}^{N} \mathcal{H}_n$. This means that it takes $\mathcal{O}(2^N)$ degrees of freedom to express a full quantum state. This becomes intractable for N around 16 or so, meaning we must find strategies to work around this limitation if we are to perform simulations on a classical computer.

This section covers several common approaches to this. The first two, the mean field approximation and cumulant expansion, make assumptions about the limitations of correlations between qubits in the system. The Dicke basis exploits symmetries in a system to group equivalence classes of degrees of freedom. Finally, the Monte Carlo Wave Function approach to numerical integration exploits these symmetries while also employing parallel operations on separate computing threads to further reduce computation time.

2.2.1 Mean Field Approximation

The mean field approximation makes the simplifying assumption that the density operator will be separable over all subsystems for all time, i.e. $\rho = \bigotimes_{n=1}^{N} \rho_n$. Recall from section 2.1.1 that this corresponds to the case of no entanglement between different elements in the system. Given observables O_i and O_j on two different subsystems, the second order expectation simplifies

according to

$$\langle O_i O_j \rangle = \mathrm{Tr}\left[\rho O_i \otimes O_j\right] = \mathrm{Tr}\left[\rho_i O_i \otimes \rho_j O_j\right] = \mathrm{Tr}\left[\rho_i O_i\right] \mathrm{Tr}\left[\rho_j O_j\right] = \langle O_i \rangle \langle O_j \rangle . \tag{2.13}$$

When writing the equations of motion, or the set of time derivatives of operators that span the space of Hermitian operators on the state space, this reduces the set of equations to just those of first order, or expectations of operators on a single subsystem. This means that the size of the system will be $\mathcal{O}(N)$, which is quite manageable even for very large N. This comes at the cost of being able to model entanglement, meaning all behavior coming from the mean field model will be classical.

Nevertheless, many systems can be well approximated in the mean field under the right conditions, specifically when the correlations between observables are much smaller than the expectations of the individual observables. It can be shown that this is the case in the limit as N becomes very large. This makes the mean field a useful tool, especially for understanding dynamics in the large particle limit.

2.2.2 Cumulant Expansion

The cumulant expansion allows for the expansion of third order operator expectations in terms of first and second order expectations. It can be thought of as similar to the mean field expansion, but one order higher. The cumulant expansion follows by assuming that cumulants of order three and higher are zero, as is the case with the Gaussian distribution. This leads to the following expansion for a third order expectation:

$$\langle O_i O_j O_k \rangle = \langle O_i O_j \rangle \langle O_k \rangle + \langle O_j O_k \rangle \langle O_i \rangle + \langle O_i O_k \rangle \langle O_j \rangle - 2 \langle O_i \rangle \langle O_j \rangle \langle O_k \rangle . \tag{2.14}$$

In a similar fashion to the mean field approximation, this allows us to reduce the dimension of the equations of motion to include only expectations of second order and lower, which implies an $\mathcal{O}(N^2)$ system of ODEs. This allows for the modeling of entanglement without the exponential scaling of the full solution to the master equation.

2.2.3 Dicke Basis

When approaching any classical simulation in quantum mechanics, an important first task is to select a basis for the state space that will determine the representation of Hilbert space elements and operators. This gives numeric vectors and matrices meaning by anchoring them with respect to the chosen basis. It is common with systems of many-level atoms to choose the eigenbasis of sums of angular momentum operators, due to the fact that most observables of interest will be angular momentum operators for each atom and their sums.

2.2.3.1 Addition of Angular Momentum

Much of the material in this section can be found with more detail in [186]. A triplet of Hermitian operators $\{J_i\}_{i=1}^3$ on a Hilbert space is said to satisfy the algebra of angular momentum if

$$[J_i, J_j] = i\hbar\epsilon_{ijk}J_k, \tag{2.15}$$

where ϵ_{ijk} is the Levi-Civita symbol in three dimensions. We have seen such operators before, as one-half times the Pauli operators satisfy this algebra. In addition to the above, we define

$$J_\pm = J_x \pm iJ_y \tag{2.16}$$

$$\boldsymbol{J}^2 = J_x^2 + J_y^2 + J_z^2. \tag{2.17}$$

These operators allow for the following equivalent characterization of the algebra of angular momentum

$$[J_+, J_-] = 2\hbar J_z \tag{2.18}$$

$$[J_z, J_\pm] = \pm\hbar J_\pm, \tag{2.19}$$

together with the statements that $J_\pm^\dagger = J_\mp$ and that J_z is Hermitian.

Given the above definitions, it can be shown that $\boldsymbol{J}^2$ and J_z commute, and are therefore simultaneously diagonalizable. Furthermore, the basis given by the common eigenvectors $|j, m\rangle$

where

$$\boldsymbol{J}^2 |j,m\rangle = \hbar^2 j(j+1) |j,m\rangle \tag{2.20}$$

$$J_z |j,m\rangle = \hbar m |j,m\rangle \tag{2.21}$$

are such that the eigenvalues of $\boldsymbol{J}^2$ and J_z uniquely identify each basis vector, making them a complete set of commuting observables. This basis will be a convenient one for Hilbert space representations of collective spin systems. The above commutation relations between angular momentum operators and the raising and lowering operators $J_\pm$ imply that j must be an integer or half-integer, and for a given j the value of m must range by integers between $-j$ and j, i.e. $m = -j, -j+1, ..., j-1, j$.

When looking at composite systems, such as a collection of multi-level atoms, it can be shown that the result of adding operators satisfying the above algebra from two different systems will satisfy the algebra on the composite system, i.e. if $\{J_i^{(1)}\}_{i=1}^3$ satisfy (2.15) on Hilbert space H_1, and $\{J_i^{(2)}\}_{i=1}^3$ satisfy (2.15) on Hilbert space H_2, then $\{J_i \equiv J_i^{(1)} \otimes \mathbb{1} + \mathbb{1} \otimes J_i^{(2)}\}_{i=1}^3$ will satisfy (2.15) on the composite space $H_1 \otimes H_2$. Each of the individual state spaces can be written as an orthogonal direct sum of their respective total angular momentum multiplets

$$\mathcal{H}_i = \bigoplus_{j_i} \mathcal{H}_i^{(j_i)} \text{ where} \tag{2.22}$$

$$\mathcal{H}_i^{(j_i)} = \bigoplus_{m_i} |j_i, m_i\rangle . \tag{2.23}$$

An argument presented in [186] shows how we can write

$$\mathcal{H}_1^{(j_1)} \otimes \mathcal{H}_2^{(j_2)} = \bigoplus_{j=|j_1-j_2|}^{j_1+j_2} \mathcal{H}^{(j)}, \tag{2.24}$$

where $\mathcal{H}^{(j)} = \text{span}\{|j,m\rangle : m = j, j-1, ..., -j\}$ is the j^{th} multiplet of the total angular momentum sum $\boldsymbol{J}^2$ and z angular momentum sum J_z. The full Hilbert space can now be written in terms of these multiplets by noting that

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 = \bigoplus_{j_1,j_2} \mathcal{H}_1^{(j_1)} \otimes \mathcal{H}_2^{(j_2)} = \bigoplus_{j_1,j_2} \bigoplus_{j=|j_1-j_2|}^{j_1+j_2} \mathcal{H}^{(j)}. \tag{2.25}$$

This result can be repeated to add angular momenta for any number of particles. This process will be clarified with a specific example in the next section.

2.2.3.2 Exploiting Symmetry

For this section we restrict our attention to a system of N spin-1/2 particles. If such a system evolves according to a master equation with particle number permutation invariance, i.e. swapping the index of particles in (2.10) does not change the equation, then it has been shown [46, 10] that a basis with $\mathcal{O}(N^2)$ elements is sufficient to represent the quantum state at all times.

To make the discussion more concrete, consider the specific example of the coupled angular momentum basis for a system of $N = 4$ spin-1/2 particles. To simplify the notation from the last section, we introduce the shorthand $j_i \equiv \mathcal{H}_i^{(j_i)}$. We further suppress the index i when it is clear from the position of the subspace in a sum. Through repeated application of equation (2.24), we get the following:

$$\begin{aligned}
\mathcal{H} &= \frac{1}{2} \otimes \frac{1}{2} \otimes \frac{1}{2} \otimes \frac{1}{2} = (1 \oplus 0) \otimes \frac{1}{2} \otimes \frac{1}{2} \\
&= \left(\left(1 \otimes \frac{1}{2} \right) \oplus \left(0 \otimes \frac{1}{2} \right) \right) \otimes \frac{1}{2} \\
&= \left(\frac{3}{2} \oplus \frac{1}{2} \oplus \frac{1}{2} \right) \otimes \frac{1}{2} \\
&= \left(\frac{3}{2} \otimes \frac{1}{2} \right) \oplus \left(\frac{1}{2} \otimes \frac{1}{2} \right) \oplus \left(\frac{1}{2} \otimes \frac{1}{2} \right) \\
&= (2 \oplus 1) \oplus (1 \oplus 0) \oplus (1 \oplus 0) \\
&= 2 \oplus (1 \oplus 1 \oplus 1) \oplus (0 \oplus 0).
\end{aligned}$$

Summing the dimension of the multiplets, we get $\dim\{\mathcal{H}\} = 5 + (3 + 3 + 3) + (1 + 1) = 16 = 2^4$ as expected for $N = 4$. As discussed in [46, 10], this generalizes to any number of particles N with the number of occurrences of multiplet j satisfying

$$d_N^j = \frac{N!(2j+1)}{(N/2-j)!(N/2+j+1)!}. \tag{2.26}$$

When there is particle permutation invariance in the master equation and the initial conditions, state coefficients will not vary between basis vectors with the same multiplet number j. They can

therefore be grouped together, with values for only one of them being tracked. Since each multiplet j is spanned by $2j+1$ orthonormal vectors, the dimension of the reduced basis is (for even N)

$$\sum_{j=0}^{N/2} 2j+1 = \frac{N}{2}\left(\frac{N}{2}+1\right)+\frac{N}{2}+1 = \frac{1}{4}(N+2)^2,$$

a significant improvement over the 2^N dimension of the full state space. When tracking the collective state of the system $|\psi_C\rangle$ in this representation, we see in [46] that calculating expectations of collective spin operators, i.e.

$$J_k = \sum_{i=1}^{N} J_{i,k},$$

where $J_{i,k}$ is spin operator k for particle i, is equivalent to what one would get if calculating it in the full exponential computational basis.

This reduced dimensional basis is known as the Dicke basis, or irrep (irreducible representation) basis, and will allow for the classical simulation of large systems, particularly when combined with the Monte Carlo Wave Function approach to simulation, discussed in the next section.

2.3 Spin Squeezing

Spin squeezing [93, 175] is a method used to improve accuracy in quantum metrology whereby a state of a collective spin system is created that reduces uncertainty of collective spin measurement in a particular direction orthogonal to the collective Bloch vector at the expense of increasing uncertainty in another direction. This is useful in experiments, for example Ramsey spectroscopy (see [116, 49] for additional examples), where measurement accuracy relies on spin-up population differences and uncertainty is driven by uncertainty in the elevation angle on the Bloch sphere.

The goal behind creating a spin squeezed state is to improve accuracy beyond what is known as the standard quantum limit (SQL). In the context of metrology, the SQL can be seen in the uncertainties of a coherent spin state (CSS), defined for a system of N spin-1/2 particles as the state without entanglement or quantum correlations where all N spins are placed in the same state

$$|\theta,\phi\rangle = \bigotimes_{i=1}^{N} \cos\frac{\theta}{2}|\uparrow_i\rangle + e^{i\phi}\sin\frac{\theta}{2}|\downarrow_i\rangle, \tag{2.27}$$

where θ is the elevation angle on the Bloch sphere and ϕ is the azimuthal angle. If we consider the specific example of $\theta = \pi/2$ and $\phi = 0$, a convention common in spectroscopy where measurements are performed in the z direction, this reduces to

$$|\pi/2, 0\rangle = \bigotimes_{i=1}^{N} \frac{1}{\sqrt{2}} \left(|\uparrow_i\rangle + |\downarrow_i\rangle\right).$$

The collective Bloch vector $\boldsymbol{J} = (\langle J_x\rangle, \langle J_y\rangle, \langle J_z\rangle)$ is simply $(N/2, 0, 0)$, and the uncertainty on, for example J_z, is

$$\Delta J_z = \sqrt{\langle J_z^2\rangle - \langle J_z\rangle^2} = \sqrt{\sum_{i=1}^{N} \frac{1}{4}\mathbb{1}} = \frac{\sqrt{N}}{2},$$

and similarly for J_y. Here and for the remainder of this section, we adopt the convention $\hbar = 1$. An arbitrary spin operator in the plane orthogonal is given by

$$J_\theta = \cos\theta J_z + \sin\theta J_y,$$

where θ in this case can range from $[0, 2\pi)$. The uncertainty is therefore

$$\begin{aligned}(\Delta J_\theta)^2 = \langle J_\theta^2\rangle - \langle J_\theta\rangle^2 = \langle J_\theta^2\rangle &= \cos^2\theta(\Delta J_z)^2 + \sin^2\theta(\Delta J_y)^2 + \cos\theta\sin\theta\langle\{J_z J_y\}\rangle \\ &= \frac{N}{4} + \cos\theta\sin\theta\langle\{J_z J_y\}\rangle.\end{aligned}$$

The lack of quantum correlations ensures that the second term is zero, as $\langle\sigma_i^z\rangle = \langle\sigma_i^y\rangle = 0$ for all i, so that we have

$$\Delta J_\theta = \frac{\sqrt{N}}{2},$$

giving the same uncertainty in all orthogonal directions to the x-axis, otherwise known as an isotropic quasiprobability distribution [93].

This value also saturates the Heisenberg uncertainty bound, which by the algebra of angular momentum (2.15) gives

$$\Delta J_y \Delta J_z \geq \left|\frac{1}{2i}\langle[J_y, J_z]\rangle\right| = \left|\frac{i}{2i}\langle J_x\rangle\right| = \frac{N}{4},$$

making a CSS a minimum uncertainty state. When performing measurements in the z-direction, accuracy is limited by the uncertainty in the elevation angle [49]

$$\Delta\theta = \frac{\Delta J_z}{||\langle\boldsymbol{J}\rangle||_2},$$

which, in the case of a CSS, becomes

$$\Delta\theta_{SQL} = \frac{\sqrt{N}/2}{N/2} = \frac{1}{\sqrt{N}}. \tag{2.28}$$

This is known as the standard quantum limit (SQL).

We have seen how the SQL arises in uncorrelated spin states. The goal of spin squeezing is to create correlations in such a way that an orthogonal direction to the Bloch vector has reduced uncertainty relative to the SQL, thus motivating a metric commonly referred to as the Wineland criterion [175]

$$\xi^2 = \min_{n_\perp} \frac{N(\Delta J_{n_\perp})^2}{||\langle \boldsymbol{J} \rangle||_2^2}, \tag{2.29}$$

where $n_\perp$ is a unit vector orthogonal to the Bloch vector $\boldsymbol{J}$, $J_{n_\perp} \equiv n_\perp \cdot \boldsymbol{J}$ is the spin operator in the direction of $n_\perp$, and the N appearing in the numerator is a normalization factor dividing the minimum orthogonal uncertainty relative to the Bloch vector length by the standard quantum limit. A value $\xi^2 < 1$ is called a squeezed state, and as described above indicates correlations between individual spins, and as such it can be seen as a witness for entanglement.

The original paper proposing spin squeezed states [93] also proposed two Hamiltonian operators that can be used to generate them, the so-called one-axis twisting Hamiltonian

$$H_{OAT} = \chi J_z^2,$$

and two-axis countertwisting Hamiltonian

$$H_{TAC} = \frac{\chi}{2i}(J_+^2 - J_-^2),$$

that reduce the minimum variance orthogonal to the Bloch sphere from $N/4$ to $\frac{1}{2}(N/6)^{1/3}$ and $1/2$, respectively.

2.4 Quantum Trajectories

In this section, we revisit the master equation (2.10) and interpret it as a sequence of measurements being performed at each infinitesimal time interval with an unknown outcome. This

interpretation allows us to simulate the solution as a series of possible realizations of this sequence with known measurement records, a collection of **quantum trajectories**, that can then be averaged to produce the estimated quantum state. We then discuss the topic of **weak measurement**, where only partial information about the system is obtained but where the back-action on the state is reduced, which is modeled by the **stochastic master equation**.

2.4.1 Quantum Channels and the Master Equation

Consider a situation in which we have performed a quantum measurement with respect to the measurement operators $\{A_n\}_{n=1}^N$, but we throw away the outcome without looking at it. What is our knowledge of the state? We know that one of the outcomes

$$\rho_n = \frac{A_n \rho A_n^\dagger}{\mathrm{Tr}\left[A_n \rho A_n^\dagger\right]}$$

has occurred, with probability $p(n) = \mathrm{Tr}\left[A_n \rho A_n\right]$. Given that we do not know which outcome actually took place, the state as we know it is the *statistical mixture*

$$\rho = \sum_{n=1}^{N} p(n)\rho_n = \sum_{n=1}^{N} A_n \rho A_n^\dagger \equiv \mathcal{E}(\rho). \tag{2.30}$$

Together with the completeness relation (2.3), this has the form of what is called a **quantum channel**, also referred to as a completely positive trace preserving (CPTP) map, and it satisfies the following properties for all density operators ρ:

- Linearity in $\rho : \mathcal{E}(a\rho_1 + b\rho_2) = a\mathcal{E}(\rho_1) + b\mathcal{E}(\rho_2)$
- Trace preserving: Due to (2.3), $\mathrm{Tr}\left[\mathcal{E}(\rho)\right] = \mathrm{Tr}\left[\rho\right]$
- Hermiticity preserving: $\mathcal{E}(\rho)^\dagger = \mathcal{E}(\rho)$
- Positivity preserving: $\lambda_i \geq 0$ for all λ_i in the spectrum of $\mathcal{E}(\rho)$

Complete positivity means that $\mathcal{E}$ preserves positivity even when it is acting on a subsystem of a larger, composite system.

With this in mind, as in chapter 3 of [139], we can interpret the master equation (2.10) as a quantum channel performing measurements on the composite system and environment with unknown measurement outcomes. In order for the modeling of the dynamics of an open quantum system in terms of an ordinary differential equation to be possible, we must assume that the evolution of the quantum system is Markovian. This is clear from the ODE itself, which writes $\rho(t + dt)$ as a function of $\rho(t)$, but not $\rho(s)$ for any $s < t$. This assumption holds for a purely dissipative system, where information can flow from system to environment but not back again. This is a reasonable model of what is going on as long as the assumption of *coarse-graining* in time holds. Coarse-graining is a circumstance in which the time Δt_{env} that it takes for the environment to "forget" information passed to it from the system is small compared to the timescale of our model Δt. The example of this given in [55] is that of a hot penny being thrown into a lake, where the penny is the system and the lake is the environment. After a very short time Δt_{env}, the energy from the penny will dissipate out into the lake, which will effectively be in the exact same state as it was when the penny was thrown in, given its size relative to the small amount of energy absorbed from the penny.

As always, modeling via an ODE is only useful if this timescale is small compared to whatever physical phenomenon we wish to observe, e.g. the timescale over which damping occurs Δt_{damp}. To summarize, the Markovian assumption holds if we can confidently say that

$$\Delta t_{env} \ll \Delta t \ll \Delta t_{damp}.$$

Fortunately, this is the case for systems such as an atom interacting with a radiation field, which is the kind of system we are interested in for this thesis.

Given that the Markov assumption holds, consider a situation where we have a composite system and an environment, and we perform repeated measurements on the system over each time interval dt without recording the outcomes. This means that the state of the system after the measurement can be written in terms of a quantum channel

$$\rho(t + dt) = \mathcal{E}(\rho(t)) = \sum_{n=0}^{N} A_n \rho(t) A_n^\dagger. \tag{2.31}$$

Note that, without loss of generality, we have chosen a zero-based index for n this time for reasons that will become clear. In order to arrive at a differential equation, we need this expression to have the form

$$\rho(t+dt) = \rho(t) + \mathcal{O}(dt).$$

Considering only those terms up to $\mathcal{O}(dt)$, we can define

$$A_0 = \mathbb{1} + (-iH + K)dt \tag{2.32}$$

$$A_n = \sqrt{dt}L_n \; n = 1, 2, ..., N, \tag{2.33}$$

where H and K are Hermitian and the separation of A_0 into real and imaginary parts is done to separate out the Hamiltonian evolution from the dissipative. By enforcing the completeness relation (2.3) we can solve for K in terms of the $\{L_n\}$:

$$K = -\frac{1}{2}\sum_{n=1}^{N} L_n^\dagger L_n. \tag{2.34}$$

Putting this together with equation (2.31) gives the familiar Lindblad master equation

$$\frac{d\rho}{dt} = -i[H,\rho] + \sum_{n=1}^{N}\left(L_n\rho L_n^\dagger - \frac{1}{2}\left\{L_n^\dagger L_n \rho\right\}\right).$$

We have, therefore, shown that the master equation can be interpreted as a procedure by which we perform a measurement at each time interval dt, with measurement operators A_n given by equations (2.32)-(2.34), and probabilities given in the usual manner according to Tr $\left[A_n^\dagger A_n \rho\right]$. Note that these probabilities are proportional to dt for $n > 0$, but larger for $n = 0$, making the latter case far more likely at any given step. The cases where $n > 0$ are often referred to as "quantum jumps."

This interpretation is more than just a thought exercise. It has practical importance when we realize that we can actually approximate solutions to the master equation by simulating the measurement procedure described above, which we will discuss in the next section.

2.4.2 Monte Carlo Wave Function

Suppose we knew the outcomes of the measurements of the composite system and environment described in the previous section. If we did, we would have a series of states ρ_k and the associated

measurement outcome μ_k at every time step $t_k = k\Delta t$. This record of states and measurements is called a **quantum trajectory**. This can easily be simulated according to the following algorithm at each time t_k:

- Generate a uniform random number ϵ on $I = [0, 1)$
- Partition the interval I according to the probabilities $p_n = \text{Tr}\left[A_n^\dagger A_n \rho_k\right]$ with each A_n given by equations (2.32)-(2.34)
- If ϵ falls in the partition of width p_0, we say that there is "no quantum jump", and we evolve according to $A_* = A_0$
- If ϵ falls in p_n for $n > 0$, we say that we perform a "quantum jump" and evolve according to $A_* = A_n$
- Set $\rho_{k+1} = A_* \rho_k A_*^\dagger / \text{Tr}\left[A_* \rho_k A_*^\dagger\right]$ and advance to $t_{k+1} = t_k + \Delta t$

In the case of no jump, we say that the state evolves according to the *effective* Hamiltonian $H + iK$, which is effective due to the anti-Hermitian part iK, which is there to account for re-normalization in the event that no jumps occur.

As discussed in the last section, however, we don't actually *have* a measurement record, so our knowledge of the state at each time step is once again a statistical mixture over M simulated trajectories:

$$\rho(t_k) \approx \tilde{\rho}(t_k) \equiv \sum_{i=1}^{M} p(i)\rho_{i,k} = \frac{1}{M}\sum_{i=1}^{M} \rho_{i,k} \tag{2.35}$$

where $\rho_{i,k}$ is the simulated state for trajectory i at time step t_k and the probability of each trajectory $p(i) = 1/M$ since all trajectories are equally likely. These sample means at each time point follow the central limit theorem, meaning that for large M the approximation $\tilde{\rho}$, and operator expectations derived using its trace, will be Gaussian distributed with means equal to the true values and standard deviations proportional to the operator standard deviations multiplied by a factor of $1/\sqrt{M}$.

This is known as the Monte Carlo Wave Function (MCWF) simulation [122], and it has several advantages over direct numerical integration of the master equation. First, if we start in a pure state $\rho_0 = |\psi_0\rangle\langle\psi_0|$, then the probabilities and state updates at each t_k are given by

$$p_n = ||A_n |\psi_k\rangle||^2 \text{ and } |\psi_{k+1}\rangle = \frac{A_* |\psi_k\rangle}{||A_* |\psi_k\rangle||}.$$

This means that we need only store the pure state $|\psi_k\rangle$ rather than the full density ρ_k, which saves space by a factor of d, where d is the dimension of the state space. A second advantage is that the quantum trajectories can be simulated independently of one another. Practically speaking, this allows us to perform the simulation of each trajectory on separate threads, in separate executables, or even on separate nodes of a cluster, allowing for massive parallelization. Finally, as described in [184], in certain systems with favorable symmetry, only a single of many irreducible linear subspaces that comprise the state space will be occupied at any given time by a single trajectory, further reducing the maximum dimension of the tracked state. All of these advantages will come into play in chapter 5, where the MCWF method will enable the simulation of systems with tens of thousands of two level atoms.

2.4.3 Cavity QED

In this section, we provide details on an open quantum system frequently considered in this thesis, a two-level atom interacting with a single mode of an electromagnetic field in a cavity. More details can be found in [159]. We have already discussed the Hamiltonian of the two-level atom H_a in (2.9) in section 2.1.2. This Hamiltonian neglects any momentum the atom may have, so we have assumed that it is being held in place, for example by an optical lattice.

The Hamiltonian for a single mode of an electromagnetic field in a cavity is given by the simple harmonic oscillator Hamiltonian

$$H_f = \hbar\omega \left(a^\dagger a + \frac{1}{2}\right), \tag{2.36}$$

where $a^\dagger$ (a) is the creation (annihilation) operator for the field mode. The operator $a^\dagger a$ is often

called the field number, and it is indicative of the field photon count. This number will rise and fall by increments as the field exchanges a photon with the atom.

As always, the atom-field state space is the tensor product of the two-dimensional state space of the atom and the countably infinite dimensional state space of the field, and operators for each individual space appearing on their own, such as σ_z or a, are understood to be in a tensor product with the identity on the other space, i.e. $\sigma_z \otimes \mathbb{1}$ and $\mathbb{1} \otimes a$.

The full Hamiltonian for the atom-field system is given by

$$H = H_a + H_f + H_{af}, \tag{2.37}$$

where H_{af} is the atom-field interaction Hamiltonian

$$H_{af} = -\boldsymbol{d} \cdot \boldsymbol{E}.$$

The atomic dipole operator $\boldsymbol{d}$ is a vector of operators on the composite space, with entries given by

$$\boldsymbol{d} = \langle g| \boldsymbol{d} |e\rangle \otimes (|g\rangle \langle e| + |e\rangle \langle g|) \equiv \boldsymbol{d}_{ge}(\sigma_- + \sigma_+), \tag{2.38}$$

where $\boldsymbol{d}_{ge}$ is a 3-D vector of operators on the field space and $\sigma_+ \equiv |e\rangle \langle g|$ and $\sigma_- \equiv |g\rangle \langle e|$ are the atomic raising and lowering operators mapping the ground state of the atom to the excited and vice-versa.

The 3-D vector of field operators $\boldsymbol{E}$ is given by

$$\boldsymbol{E}(\vec{r}) \equiv -\sqrt{\frac{\hbar\omega}{2\epsilon_0}} \left(\boldsymbol{f}(\vec{r})a + \boldsymbol{f}^*(\vec{r})a^\dagger \right), \tag{2.39}$$

where $\boldsymbol{f}$ is called the normalized spatial mode profile, and is a function of the atomic position $\vec{r}$. Since we are assuming that the atom is in a fixed position, we can combine all of this into a constant g and write

$$H_{af} = \hbar g \left(\sigma_- + \sigma_+\right) \left(a + a^\dagger\right).$$

When the magnitude of the detuning between the atom and the cavity $|\omega - \omega_a| \ll \omega + \omega_a$, we can make one more simplifying assumption, called the rotating wave approximation (RWA), where

terms that don't conserve energy, such as $\sigma_- a$ and $\sigma_+ a^\dagger$, where both the atom and field lose or gain a photon respectively, are assumed to be zero. In this case we are left with only the terms where the atom and field exchange a photon:

$$H_{af} = \hbar g \left(\sigma_- a^\dagger + \sigma_+ a \right). \tag{2.40}$$

Adopting the convention of dropping terms in the Hamiltonian proportionate to the identity, since they do not affect the physical dynamics, and putting it all together, we have what is known as the Jaynes-Cummings model [87] for a two-level atom interacting with a field:

$$H = H_a + H_f + H_{af} = \frac{\hbar\omega_a}{2}\sigma_z + \hbar\omega a^\dagger a + \hbar g \left(\sigma_- a^\dagger + \sigma_+ a \right). \tag{2.41}$$

The generalization to a collection of N atoms in a cavity is straightforward, as we can just sum the individual atom and interaction Hamiltonian terms

$$H = \sum_{i=1}^{N} \frac{\hbar\omega_{i,a}}{2}\sigma_{i,z} + \hbar\omega a^\dagger a + \sum_{i=1}^{N} \hbar g_i \left(\sigma_{i,-} a^\dagger + \sigma_{i,+} a \right), \tag{2.42}$$

where i subscripts indicate the atom index for each operator, and g_i now depends on the individual atomic position. It is often the case that all atomic transition frequencies are identical ($\omega_{i,a} \equiv \omega_a$), and that the variations in the g_i do not affect the physics, so we can let $g_i \equiv g$. In this case, the Hamiltonian simplifies to

$$H = \frac{\hbar\omega_a}{2}S_z + \hbar\omega a^\dagger a + \hbar g \left(S_- a^\dagger + S_+ a \right), \tag{2.43}$$

where $S_{x,y,z} = \sum_i \sigma_{x,y,z}$ and $S_\pm = \sum_i \sigma_\pm = S_x \pm i S_y$ are collective spin operators satisfying the algebra of angular momentum $[S_i, S_j] = i\epsilon_{ijk} S_k$ as discussed in section 2.2.3.1.

2.4.4 Stochastic Master Equation

A more general form of the master equation can be derived in a similar manner to the Lindblad form in section 2.1.3 if we generalize the operator in (2.32) to be

$$A_0 = \mathbb{1} + (-iH + K)dt + C dW, \tag{2.44}$$

where C is an operator and dW is the Itô differential. As was done in [159, 3], we can now proceed to derive the stochastic master equation (SME) in a manner similar to that of section 2.1.3.

Restricting our attention, for the moment, to the action of just the operator A_0 in (2.44), i.e. considering

$$\rho(t+dt) = A_0 \rho A_0^\dagger,$$

we expand to first order in dt, following the rules of Itô calculus that give $dW^2 = dt$, resulting in

$$d\rho \equiv \rho(t+dt) - \rho = -i[H,\rho]dt + \{K,\rho\}\,dt + C\rho C^\dagger dt + \left(C\rho + \rho C^\dagger\right) dW. \tag{2.45}$$

Taking the ensemble average, or the average over all possible Wiener processes, and invoking the property of Itô calculus that states $\langle\langle \rho dW \rangle\rangle = 0$, gives

$$d\langle\langle\rho\rangle\rangle = -i[H,\langle\langle\rho\rangle\rangle]dt + \{K,\langle\langle\rho\rangle\rangle\}\,dt + C\langle\langle\rho\rangle\rangle C^\dagger dt. \tag{2.46}$$

Enforcing that the map must be trace preserving means that

$$d\mathrm{Tr}\left[\langle\langle\rho\rangle\rangle\right] = \mathrm{Tr}\left[d\langle\langle\rho\rangle\rangle\right] = 0.$$

Combing this with (2.46) gives

$$\begin{aligned} 0 &= 2\mathrm{Tr}\left[K\langle\langle\rho\rangle\rangle\right]dt + \mathrm{Tr}\left[C\langle\langle\rho\rangle\rangle C^\dagger\right]dt \\ &= \mathrm{Tr}\left[\left(2K + C^\dagger C\right)\langle\langle\rho\rangle\rangle\right]dt, \end{aligned} \tag{2.47}$$

where the cyclical permutation property of the trace gives $\mathrm{Tr}\left[[A,B]\right] = 0$ and $\mathrm{Tr}\left[\{A,B\}\right] = 2\mathrm{Tr}\left[AB\right]$. Equation (2.47) only holds for general $\langle\langle\rho\rangle\rangle$ if

$$K = -\frac{C^\dagger C}{2}.$$

Putting this back in (2.45) gives

$$d\rho = -i[H,\rho]dt + C\rho C^\dagger dt - \frac{1}{2}\left\{C^\dagger C,\rho\right\}dt + \left(C\rho + \rho C^\dagger\right)dW \tag{2.48}$$

We still aren't quite finished. While we enforced the trace preserving condition on the equation in the ensemble average, we have to make sure that it still applies to the full form that includes the

Itô differential. This time requiring $\mathrm{Tr}\left[d\rho\right] = 0$ gives

$$\mathrm{Tr}\left[d\rho\right] = \mathrm{Tr}\left[C\rho + \rho C^\dagger\right] dW = \mathrm{Tr}\left[\rho\left(C + C^\dagger\right)\right] dW = \left\langle C + C^\dagger \right\rangle dW = 0.$$

In [3], this is interpreted as a restriction on the operator C, but in [159] they allow C to be general and instead modify the differential to explicitly normalize the state at each time step. To first order in dt, this is equivalent to just subtracting off the trace of $d\rho$ from (2.48), resulting in

$$d\rho = -i[H,\rho]dt + \left(C\rho C^\dagger - \frac{1}{2}\left\{C^\dagger C, \rho\right\}\right) dt + \left(C\rho + \rho C^\dagger - \left\langle C + C^\dagger \right\rangle\right) dW. \tag{2.49}$$

This is the simplest form of the stochastic master equation (SME), where only one operator C is being measured. Here we see the familiar Lindblad super-operator $\mathcal{D}[C](\rho)$ in the dt term, and a new super-operator $\mathcal{H}[C](\rho)$ in the dW term which reflects our knowledge of the weak measurement record of C. Note that this latter term is non-linear in ρ, which distinguishes the stochastic master equation from the unconditioned master equation, the name often used in the context of the SME for the regular Lindblad master equation (2.10). This name derives from the fact that the unconditioned equation is obtained by taking the SME in the ensemble average, and is equivalent to what one would know of the dynamics if the measurement record of C is unknown.

The SME can be generalized in two ways. First, we can reintroduce the dissipative terms $\{L_n\}_{n=1}^N$ for which no measurement record is known that appear in the Lindblad master equation. Second, we can generalize to multiple weak measurement operators $\{C_m\}_{m=1}^M$. The full stochastic master equation is then given by

$$d\rho = -i[H,\rho]dt + \sum_{n=1}^{N} \mathcal{D}[L_n](\rho)dt + \sum_{m=1}^{M} \mathcal{D}[C_m](\rho)dt + \mathcal{H}[C_m](\rho)dW_m. \tag{2.50}$$

2.4.5 Photon Detection

A quantum trajectory can be defined as the time evolution of a quantum system conditioned on known measurement outcomes [176]. The focus of this section is to discuss three ways in which this can be accomplished, and the master equations associated with each one. These are direct, homodyne, and heterodyne detection. While the procedure described in section 2.4.2 is a

way to simulate a quantum trajectory, where each Lindblad operator corresponds to a generalized measurement that is being performed, this section will describe various detection schemes in terms of a stochastic master equation, by looking at what happens in the limit of large number of detections. In this limit, the jump process of section 2.4.2 becomes a Gaussian process and can be described by an Itô stochastic differential equation.

In the case of homodyne and heterodyne detection, this occurs by combining the output of the atomic system with a local oscillator in a beam splitter and taking the amplitude of the local oscillator to infinity. The resulting measurements are often referred to as "weak" measurements, as only some of the information being collected comes from the system, while the rest is coming from the local oscillator. Further details of this procedure and the following derivations of the corresponding stochastic master equations can be found in [159, 176].

2.4.5.1 Direct Detection

Consider the simple case of a single two-level atom where we are interested in detecting whether or not it emits a photon. In the case of direct detection where the atom is emitting photons at a rate Γ (units of Hz), the operator corresponding to the generalized measurement is $C = \sqrt{\Gamma}\sigma_-$, since the impact on the system should a photon be measured is to drop it from the excited to the ground state, precisely the action of σ_-. Furthermore, the average number of photons detected in an interval of time dt is given by the probability of detecting a photon in a given time increment times dt:

$$\mathrm{Tr}\left[C\rho C^\dagger\right] dt = \left\langle C^\dagger C \right\rangle dt = \Gamma \left\langle \sigma_+\sigma_- \right\rangle dt.$$

This corresponds to a generalized measurement operator (A_1 in (2.32)) of $A_1 = \sqrt{\Gamma dt}\sigma_-$. The resulting master equation (2.10) or the procedure in section 2.4.2 can be used to simulate a direct detection quantum trajectory depending on whether or not we want to simulate a measurement record or just consider dynamics in the ensemble average.

2.4.5.2 Homodyne Detection

While the case of direct detection described above is one way to measure photons coming from the atom, a practical concern is that our photo-detector could be subject to background noise which could interfere with our measurement record. Homodyne detection is a method for boosting the signal coming from the atom to overcome this noise without sacrificing information being gleaned from the system.

This is accomplished by combining the output of the system with a local oscillator in the space of the cavity field, and passing them both through a beam splitter with field reflection coefficient $r \in [0, 1]$.

Analogous to the case of direct detection, the operator corresponding to the generalized measurement of a photon coming from the local oscillator is $C_{lo} = \sqrt{\Gamma}a$, where a is the photon count lowering operator for the field of the oscillator. The beam splitter makes it so that the operator on the composite system being measured is

$$C_r = rC + \sqrt{1-r^2}C_{lo} = \sqrt{\Gamma}\left(r\sigma_- + \sqrt{1-r^2}a\right).$$

The quantum analog of a classical electromagnetic field with a single frequency is called a coherent state and denoted by $|\alpha\rangle$ for a photon flux of $\Gamma|\alpha|^2$, for some complex number α. The important property for this discussion is that it is an eigenstate of the lowering operator such that $a|\alpha\rangle = \alpha|\alpha\rangle$. Since the local oscillator is in this state, the action of C_r effectively becomes

$$C_r = \sqrt{\Gamma}\left(r\sigma_- + \sqrt{1-r^2}\alpha\right).$$

Defining $\beta \equiv \alpha\sqrt{1-r^2}$, we are interested in the case where $r \to 1$ and $|\alpha| \to \infty$. The first condition is because we don't want to waste energy through the splitter that we are not detecting, and the latter case is effectively turning up the local oscillator so we can still see it through our highly reflective beam splitter. This results in $C_r \to C_\beta \equiv \sqrt{\Gamma}(\sigma_- + \beta)$ and makes the average rate of photon detection

$$\left\langle C_\beta^\dagger C_\beta \right\rangle dt = \Gamma\left[\langle\sigma_+\sigma_-\rangle + \langle\beta^*\sigma_- + \beta\sigma_+\rangle + |\beta|^2\right] dt.$$

Ignoring the cross term, we see that our detection rate has been increased by the $|\beta|^2$ term, this helps the signal stand out from the noise and is the whole reason for doing homodyne detection.

In accordance with our definition of measurement, whenever a photon is detected our state changes

$$\rho \to \frac{(\sigma_- + \beta)\,\rho\,(\sigma_+ + \beta^*)}{\langle (\sigma_- + \beta)\,(\sigma_+ + \beta^*)\rangle},$$

which is a mixture of what happens in direct detection of a photon plus the unchanged state. This is due to the fact that we do not know whether the photon came from the atom or the local oscillator. This uncertainty is why this is sometimes referred to as a "weak" measurement. There is more uncertainty, but also less direct impact on the state as a result of the measurement.

We now follow the steps in [159] to produce the stochastic master equation by substituting C_β into the master equation for direct detection. First, we write it in terms of the counting process dN which is one if a photon is detected in dt and zero otherwise to account for the state after the measurement. This modified form is

$$d\rho = -i[H,\rho]dt + \mathcal{D}[\sqrt{\Gamma}\sigma_-](\rho)dt + \left(\frac{\sigma_-\rho\sigma_+}{\langle\sigma_+\sigma_-\rangle} - \rho\right)dN, \tag{2.51}$$

where the last term effectively replaces ρ with the state assuming a photon was measured should dN be one. This occurs with probability

$$P(dN = 1) = \Gamma\,\langle\sigma_+\sigma_-\rangle\,dt,$$

or the probability that a photon is directly measured. The conditioned form of the master equation in (2.51) assumes we keep the information about the measurement record rather than discarding it, and is equivalent to the procedure outlined in section 2.4.2 to simulate a single trajectory.

We obtain the stochastic master equation in the form of (2.50) by first letting $\sigma_- \to C_\beta = \sigma_- + \beta$ in (2.51), noting that the unconditioned form, or the form in the ensemble average, must recover the original unconditioned master equation since the measurement record is not retained in that case. This leads to the additional replacement

$$H \to H - \frac{i\Gamma}{2}\,(\beta^*\sigma_- - \beta\sigma_+)$$

in (2.51). In addition, we take $|\beta| \to \infty$. This amounts to turning up the amplitude of the local oscillator so that the photon count becomes large, and it allows us to take the large N limit for our counting process so that, letting $\mu \equiv \langle\langle dN/dt\rangle\rangle = \Gamma\langle(\sigma_+ + \beta^*)(\sigma_- + \beta)\rangle$,

$$dN \to \mu dt + \sqrt{\mu}\, dW, \tag{2.52}$$

where dW is the Itô differential. Making the substitutions and simplifying leads to the familiar form

$$d\rho = -i[H,\rho]dt + \Gamma\mathcal{D}[\sigma_- e^{i\phi}](\rho)dt + \sqrt{\Gamma}\mathcal{H}[\sigma_- e^{i\phi}](\rho)dW, \tag{2.53}$$

where $\beta = |\beta|e^{-i\phi}$ and $\mathcal{H}[C](\rho)$ is the homodyne super-operator

$$\mathcal{H}[C](\rho) \equiv C\rho + \rho C^\dagger - \left\langle C + C^\dagger \right\rangle. \tag{2.54}$$

The current of the photo-detector is given by

$$I(t) = Q_{ph}\frac{dN}{dt},$$

where Q_{ph} is the charge of a photon and dN is given by the SDE in (2.52). Subtracting off the current of the local oscillator $Q_{ph}\Gamma|\beta|^2$ and keeping only the lowest-order terms in $|\beta|^{-1}$ gives the SDE for the current coming from the atom. Defined to be normalized by $Q_{ph}|\beta|$, this is

$$dr = \Gamma\left\langle \sigma_- e^{i\phi} + \sigma_+ e^{-i\phi} \right\rangle dt + \sqrt{\Gamma}dW. \tag{2.55}$$

The value of $r(t)$ can be interpreted as proportional to the total charge accumulated by the photo-detector, or the total photon count.

2.4.5.3 Heterodyne Detection

The SME for heterodyne detection is a simple modification of the SME for homodyne detection (2.53). In the derivation of the latter SDE, we assumed that the phase ϕ of the local oscillator did not vary with time. Physically, this corresponds to the case where the frequency of the oscillator is the same as that diriving the dynamics of the atom. Heterodyne detection allows the two

frequencies to differ such that $\Delta \equiv \omega_{lo} - \omega_a \neq 0$. This amounts to the substitution $\phi \to \Delta t$ in (2.53), giving the SME

$$d\rho = -i[H, \rho]dt + \Gamma\mathcal{D}[\sigma_-](\rho)dt + \sqrt{\Gamma}\mathcal{H}[\sigma_- e^{i\Delta t}](\rho)dW, \tag{2.56}$$

with the corresponding measurement

$$dr = \Gamma \left\langle \sigma_- e^{i\Delta t} + \sigma_+ e^{-i\Delta t} \right\rangle dt + \sqrt{\Gamma} dW. \tag{2.57}$$

Note that the $e^{i\Delta t}$ does not appear in the Lindblad super-operator. This is due to the fact that it really didn't appear in (2.53) either since the operator argument of $\mathcal{D}$ always appears in conjugate pairs. It is just presented that way for symmetry by convention in the equation for homodyne detection.

A motivating example for the use of heterodyne detection is given in section 18.2.7 of [159], where it is noted that often times in the lab measurement noise is proportionate to $1/\omega_a$, which motivates using a higher frequency. In addition, by demodulating the measurement, a record for $\langle \sigma_+ \rangle$ can be obtained thus providing information on both $\langle \sigma_x \rangle$ and $\langle \sigma_y \rangle$.

Chapter 3

Background - Machine Learning

Recent years have seen a number of applications for machine learning applied to quantum systems. Neural networks and generative models have been used to represent scalable quantum states [37, 42, 145, 41], and for learning parameters in quantum dynamics [65, 97, 68], among others. In chapter 6, we'll be presenting a machine learning model for estimating parameters in a master equation from weak measurement records. This chapter covers the background needed to understand the components of that model.

3.1 Autoencoders

An autoencoder is a neural network model that learns to produce a copy of its input as its output [71]. The utility of this lies in the internal structure of the model, where the input vector $x \in \mathbb{R}^n$ is mapped to a **latent vector** $z \in \mathbb{R}^d$ by an encoder $z = f_\theta(x)$, before being mapped to the output $y \in \mathbb{R}^n$ by a decoder $y = g_\phi(z) = g_\phi(f_\theta(x))$ (see figure 3.1). The vectors θ and ϕ are the model parameters to be learned during training. This allows the model to learn a representation of the input space with reduced dimension in the case where $d < n$, or increased when $d > n$. It is also important to note that it is not expected to be a universal identity map on the input space. It is expected that it will only work for a subset of this space where training data resides, but as such the latent space will provide useful insights about this type of data.

As with all feed-forward neural networks, each horizontal layer of figure 3.1 represents a vector of real (or potentially complex) numbers, with the connecting lines representing a linear

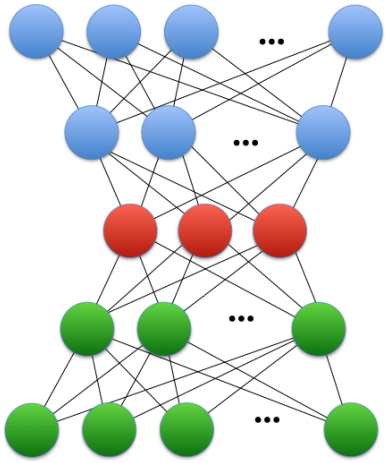


Figure 3.1: Basic neural network architecture for an autoencoder, with the encoder in green, the latent variable layer in red, and the decoder in blue

transformation followed by an optional activation function, which may be nonlinear. Formally, layer i is given by $x_i \in \mathbb{R}^{n_i}$ with $x_i = f_i(W_i x_{i-1} + b_i)$, where W_i is a weight matrix, sometimes called a kernel, b_i is a bias vector, and f_i is the activation function. Common activation functions include the sigmoid function

$$f(x) = \frac{1}{1 + e^{-x}}$$

or the rectified linear unit (ReLU) function

$$f(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0. \end{cases}$$

These functions are applied to the vectors element-wise.

Training is commonly performed using the mean squared error (MSE) loss function

$$L_{MSE}(x_i, y_i) = \frac{1}{n} \sum_{k=1}^{n} |x_{i,k} - y_{i,k}|^2 \tag{3.1}$$

where $x_{i,k}$ is the k^{th} element of x_i (and similarly for y_i), and the total loss over a training set $\{x_i\}_{i=1}^{N}$ is just the mean over all of the individual training example losses

$$L = \frac{1}{N} \sum_{i=1}^{N} L_{MSE}(x_i, y_i) = \frac{1}{Nn} \sum_{i=1}^{N} \sum_{k=1}^{n} |x_{i,k} - y_{i,k}|^2. \tag{3.2}$$

Training is typically performed by minimizing the loss via stochastic gradient descent where batches of training data are randomly selected and used to perform an update on the model parameters with a configurable learning rate. Each pass through the full training data set is called an epoch, and training proceeds for a number of epochs until the loss saturates or until the loss on a separate validation set starts to rise relative to that of the training set. The learning rate is a scalar multiplied by the negative gradient of the loss with respect to the model parameters when computing each step, and is typically decreased with each epoch to allow for fine scale optimization at later epochs. A common practice is to use an exponential decay where

$$\eta_k = \lambda^k \eta_0, \tag{3.3}$$

where η_k is the learning rate for epoch k and $\lambda \in (0, 1]$ is a configurable decay rate. A formula such as (3.3) is referred to as a learning rate schedule.

3.1.1 Denoising Autoencoders

Suppose that the input data for the autoencoder is the result of some stochastic process with probability density $p_{data}(x)$ that includes noise. This means that instead of seeing clean data elements in our training set, we actually have a collection of data elements $\{\tilde{x}_i\}_{i=1}^N$ corrupted by this noise. We would like to train an autoencoder to remove or to see though this noise to reproduce and infer properties of the corresponding uncorrepted data elements $\{x_i\}_{i=1}^N$. This can be accomplished if we have a training set that provides examples of both corrupted and clean training examples, in which case we still use the loss function in (3.1), but using the corrupted examples $\{\tilde{x}_i\}$ as input to the autoencoder, i.e.

$$L_{MSE}(x_i, g_\phi(f_\theta(\tilde{x}_i))) = \frac{1}{n} \sum_{k=1}^{n} |x_{i,k} - g_\phi(f_\theta(\tilde{x}_{i,k}))|^2, \tag{3.4}$$

where f_θ is the encoder function and g_ϕ is the decoder. The denoising autoencoder can thus be seen to be attempting to undo the corruption due to the noise.

As discussed in sections 14.2.2 and 14.5 of [71], the source of the noise could be due to a corruption process $C(\tilde{x}|x)$. If this process is known, samples of uncorrupted data can be used to

simulate corrupted data for training. In other circumstances, such as the case of quantum trajectories, what is observed is the corrupted data emerging according to a not necessarily known distribution $\tilde{x}_t \sim p(\tilde{x}_t, t|x_t)$ where x_t is the measurement sequence in the ensemble average, or the solution to the unconditioned master equation. This means that uncorrupted data can be approximated by taking the mean over a large number of trajectories. Combined with the trajectories that comprised the mean, we have the datasets $\{\tilde{x}_{t,i}\}_i$ and $\{x_{t,i}\}_i$ needed to train the denoising autoencoder. Given the large amount of noise that might be present in a single trajectory, it may be desirable to let each $\tilde{x}_t$ itself be a mean of trajectories, and to let x_t be a mean of means.

A useful interpretation of what is being learned by an autoencoder is provided in section 14.5 of [71]. The output $z = f_\theta(\tilde{x}_t)$ can be thought of as a lower dimensional embedding of the data distribution $p_{data}(x_t)$. The idea is that the training data $\{x_{t,i}\}_i$ actually occupies a manifold of dimension much smaller than that of the data space, and it is the autoencoder's job to learn this manifold. In the case of a denoising autoencoder, it is learning a function $r(\tilde{x}_t) \equiv g_\phi(f_\theta(\tilde{x}_t))$ such that the reconstruction pointing vector $r(\tilde{x}_t) - \tilde{x}_t$ will point in the direction of x_t over a fairly large neighborhood around x_t. The function $r(\tilde{x}_t)$ will therefore have a small gradient near the manifold of clean data examples, and a large derivative in directions orthogonal to that manifold in order to kick corrupted data back onto the true manifold.

In the case of estimating the physical parameters of a stochastic master equation, as will be discussed in chapter 6, the manifold is the set of all solutions to the unconditioned master equation as parameterized by the physical parameters in the ODE. The autoencoder will learn to map a noisy trajectory $\tilde{x}_t$ onto its counterpart x_t on the manifold, and in the process it will learn to output the manifold coordinates of x_t, given by the physical parameters of interest.

3.1.2 Variational Autoencoders

In the preceeding section, we were interested in learning deterministic maps f_θ and g_ϕ that map data points to coordinates of a lower dimensional embedding and back again. Suppose instead that we are interested in finding a distribution with density $p(x)$ that approximates a data distribu-

tion $p_{data}(x)$. If we could do this, then we would be able to sample from p to approximate sampling from p_{data}, and we could approximate expected values taken with respect to the data distribution as a result. Such a model is called a **generative model**, and it is typically trained to maximize the log-likelihood of our training data, presumably sampled according to the data distribution, with respect to our estimated density. Formally, this results in the optimization problem

$$\hat{\theta} = \max_{\theta} \sum_{i=1}^{N} \log(p_\theta(x_i)), \tag{3.5}$$

for a training data set $\{x_i\}_{i=1}^N$, and a parameterized density $p_\theta(x)$.

In the case of autoencoders, we can generalize the model in section 3.1 by replacing the deterministic functions $z = f_\theta(x)$ and $y = g_\phi(z)$ for our encoder and decoder with the distributions $p_\theta(z|x)$ and $q_\phi(y|z)$, respectively. The latent variable vector z is now a random vector with conditional density $p_\theta(z|x)$. In practice, we won't be able to optimize the log-likelihood in (3.5) if z is not a differentiable function of the parameters in θ. To accomplish this, we employ what is known as the "reparameterization trick," whereby we represent z conditional on a data value x as

$$z = \mu_\theta(x) + \sigma_\theta(x)\epsilon,$$

where $\mu_\theta(x)$ and $\sigma_\theta(x)$ are mean and standard deviation functions given by the encoder, and typically $\epsilon \sim \mathcal{N}(0, \mathbb{1})$ is a multivariate standard normal.

While a direct maximization of (3.5) is not tractable in general, we can optimize a lower bound on it, known as the variational lower bound, by using the following loss for a given data element x_i:

$$L_i(\theta, \phi) = -\mathbb{E}_{z\sim p_\theta(z|x_i)} \left[\log q_\phi(x_i|z)\right] + \boldsymbol{KL}\left(p_\theta(z|x_i), p(z)\right), \tag{3.6}$$

where $\boldsymbol{KL}$ is the Kullback-Leibler divergence

$$\boldsymbol{KL}(P(x), Q(x)) = \sum_x P(x)\left[\log P(x) - \log Q(x)\right],$$

and $p(z)$ is a pre-determined prior distribution on the latent space, commonly chosen to be a standard normal. The total loss is just taken to be the sum of individual losses

$$L(\theta, \phi) = \sum_i L_i(\theta, \phi). \tag{3.7}$$

The first term on the right-hand side of (3.6) is a direct maximum likelihood estimation of the data with respect to the latent variable as sampled from the distribution determined by the encoder. The second term is a regularization term that keeps the latent variable well-behaved according to a known distribution. Once trained, sampling can be achieved by sampling $z \sim p(z)$ and then running it through the decoder to sample $x \sim q_\phi(x|z)$.

This model is directly applicable to quantum systems when we consider that a quantum state ρ determines the distributions associated with all measurements, and that when certain distributions are known we can uniquely reconstruct the quantum state. As discussed in [42], a positive operator valued measure (POVM) $\{E_a\}$, as defined in section 2.1.1, is considered informationally complete if any Hermetian operator on the state space can be represented as a linear combination of the operators in $\{E_a\}$. In this case, the authors show how learning the distribution for this POVM allows a reconstruction of the state sufficient to infer any expectation of or probability associated with a local operator. See also [145, 41, 37, 36] for more on using generative models to represent quantum states.

The following example shows what happens when we train a VAE with a 2D latent space $z \in \mathbb{R}^2$ using samples provided by the QuCumber tutorial [141] from the transverse-field Ising Hamiltonian for qubits on a 1D lattice

$$H = -J \sum_{i=1}^{N-1} Z_i Z_{i+1} - h \sum_{i=1}^{N} X_i, \tag{3.8}$$

with $N = 10$ qubits, and $J = h = 1$. Training data was obtained by sampling in the computational basis (measuring the observables $\{Z_i\}_{i=1}^{10}$) from a system in the ground state of H. Figure 3.2 plots the values of $q_\phi(Z_i = 1|z)$ as a square tile for each latent variable (z_1, z_2) pair in a mosaic over a range of values for this two dimensional latent space (note that the lower case z is the latent variable vector, not to be confused with the Pauli operators Z_i). In the figure we are manually feeding a grid of latent variable pairs into the decoder to see what distribution they produce and how it varies with each latent variable to see if we can interpret what physical trait, if any, each variable is correlated with. What we see is that the model has learned to correlate the first latent

dimension with spin direction, where the number of qubits that are spin-up increases as the value of z_1 increases. It has learned to treat the second latent dimension as position on the lattice, where positions that hold a given spin value swapping from left to right with increasing value.

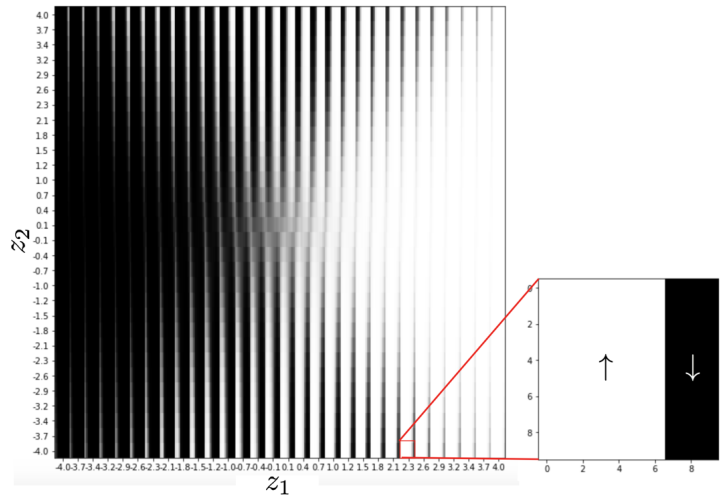


Figure 3.2: The latent space learned by a VAE trained on data sampled from the ground state of an Ising Hamiltonian for 10 qubits on a 1D lattice. It can be seen that the horizontal axis represents spin direction and the vertical axis represents lattice position

This example shows how generative models can learn latent embeddings with physical interpretations from data alone. Note also that the dimension of a VAE will scale polynomially with the number of qubits in the system, in contrast with the state that grows exponentially. Assuming they are applied to systems that inhabit a manifold of sufficiently small dimension relative to that of their state space, this gives generative models the potential to scale to much larger systems than can be simulated with traditional methods.

3.2 Recurrent Neural Networks

A recurrent neural network (RNN) is a model designed to process sequential data. Commonly, as will be the case for quantum trajectories, the sequence is in time, but any ordered stream of data can be processed. An architecture diagram of a standard RNN is shown in figure 3.3. At each time step t, the input vector $x_t \in \mathbb{R}^n$ is fed into the recursive block R, which provides an

output vector for that time $h_t \in \mathbb{R}^d$, but also provides *state* information to be used within R for the next time step. This can be thought of as a recursive loop, as shown on the left side of the figure, or as a sequence of neural network operations as shown on the right. Critically, the weights in the recursive block R are learned during training and then held fixed for each element of the time sequence during prediction. It is the internal state that changes in time, not the trainable weights.

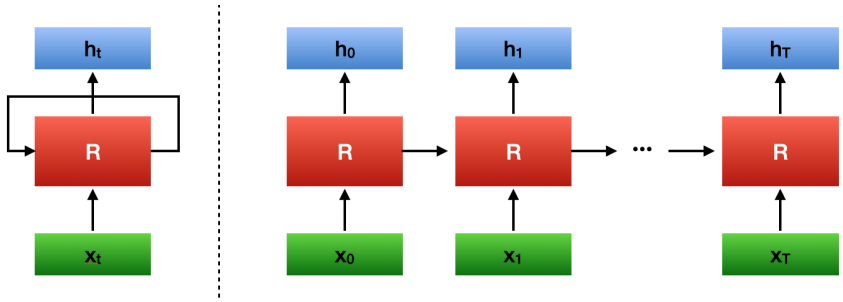


Figure 3.3: A diagram of an RNN in its recursive representation (left) and unrolled representation (right)

RNNs can offer performance advantages, as their recursive nature allows for loop optimization with the same code just operating on different data at each point in time. Training is done using an algorithm called back-propagation in time, a generalization of standard back-propagation for neural networks optimized for a sequential architecture. The details of the architecture inside the recursive block R can vary depending on the type of RNN with three variations having emerged as the most common [132].

In a basic RNN, the state is simply the same as the output vector h_t. The recursive block R uses a feed-forward neural network to map the input x_t and the previous state h_{t-1} to the output h_t, which is then passed on as the state to the R block for the next time step. Formally, this looks like

$$h_t = f(x_t, h_{t-1}; \theta), \tag{3.9}$$

where θ is the set of trainable parameters that do not vary with t.

The important thing to note when it comes to applying RNNs to quantum systems, or to any dynamical system, is that (3.9) is Markovian in the sense that h_t can be calculated from h_{t-1} alone, and is not dependent on h_s for $s < t-1$ when given h_{t-1}. The state at time t contains all information about the past, and therefore everything the model uses to move forward. This limits the utility of a basic RNN if we want to account for non-Markovian dynamics, which is often precisely what we need a machine learning model for, since physical models such as the master equation (2.10) can already account for Markovian dynamics. Fortunately, RNN architectures have been developed to deal precisely with this problem. The two most common examples are the Long Short Term Memory (LSTM) and Gated Recurrent Unit (GRU) architectures. We will be focusing on the LSTM for the algorithms discussed in this thesis, but other architectures are available and can be just as effective.

The LSTM is an alternative design for the recurrent cell R that is designed to allow the RNN to retain information from the past for longer and to draw upon it to perform updates and produce output. It accomplishes this through the introduction of a *carry state* c_t that is passed to the next time step along with the output state h_t. A diagram of the standard construction is seen in figure 3.4, which is a zoomed in view of the RNN cell R. The layout is adapted from [132].

The carry state passes through the top of the cell in the diagram. It is updated first, then used in the update of the output state h_t. Each of the sigmoid activated layers serves as a "forget" gate, as it outputs values that are typically close to zero or one, and it is always element-wise multiplied by another vector in the LSTM. The first sigmoid layer uses the previous state and input to determine a forget gate for the carry state. It drops certain elements of the vector, zeroing them out, while keeping others. The second sigmoid does the same for the output of the tanh layer, which is the traditional output of a basic RNN. These two states get summed together, before passing through a tanh activation and one more forget gate, producing the output state h_t. The carry state c_t, now updated with a forget gate and information from the input and previous state, gets passed on to the next time step, along with the new output state.

The architecture is complex, but the key takeaway is that the carry state allows information

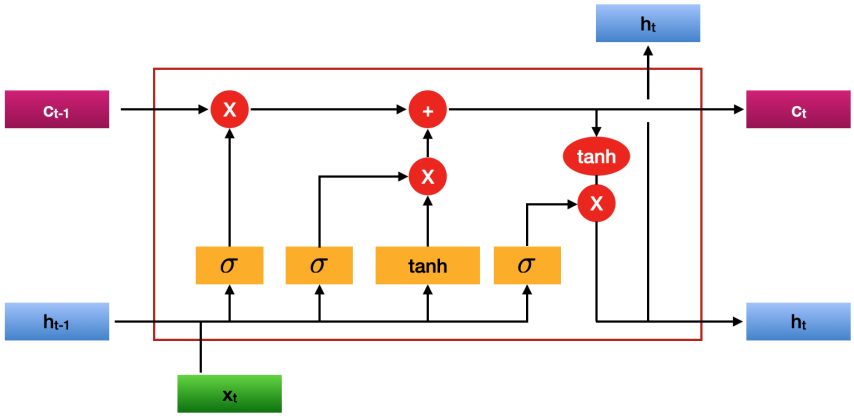


Figure 3.4: A diagram of the LSTM cell architecture, adpated from [132]. The red border represents R in figure 3.3. Each rectangular block inside is a neural network layer labeled by its activation function. Red circles/ellipses represent elemnt-wise operations. Merging black lines are concatenations, and the external rectangles are the input/output of the cell

from the past to be used directly to update the state independently of the information contained in the previous state. This enables non-Markovian behavior within an LSTM, which has been shown to effectively model quantum systems with non-Markovian dynamics [97, 9].

3.3 Neural ODE

If we look closely at what each step of a standard RNN is doing, and (for the moment) ignore the input x_t, we notice it takes a familiar form. Each step can be viewed as updating the state according to a function that takes the form of a neural network

$$h_{t+1} = h_t + f_\theta(h_t),$$

where f_θ is a neural network parameterized by θ. This kind of state update in time is very similar to what is done within a numerical integrator when solving an initial value problem (IVP). A 2018 paper by R. T. Q. Chen et. al [48] proposed a machine learning model to do just that, where a function f is parameterized by a neural network model and trained through observations of data

to learn the derivative of the hidden state, which in the limit of infinitessimal time steps gives

$$\frac{dh(t)}{dt} = f_\theta(h(t), t). \tag{3.10}$$

A 2019 paper by Y. Rubanova et. al [147] took this a step further, incorporating this design into an RNN architecture and extending it to include the input x_t at each time step. The idea is to fill in the state values at times in between observations with the solution, rather than simply assuming the state is constant as in the case of a standard RNN. The full state update algorithm for a single observation time step is therefore

$$h'_{k+1} = ODESOLVE(f_\theta, h_k, t_{k+1}, t_k),$$
$$h_{k+1} = R(h'_{k+1}, x_{k+1}),$$

where $ODESOLVE$ is any numerical ODE solver, R is the usual RNN cell update, and the time is now indexed according to t_k since we do not have to assume a regular time spacing. They refer to this method as ODE RNN, and it has the advantage of allowing larger or irregular time spacing between observations.

For physical systems, such methods can be very useful as they allow us to learn the equation f_θ governing dynamics in a general, model-free way that isn't bound to the assumptions of traditional models such as the master equation (2.10), which assumes both Markovian evolution and linearity. Note that where a standard RNN simply updates the state at each time step, one of the products of neural ODE is the derivative function itself. This is useful for obtaining physical insights into the more complex dynamics of a system, and it also allows us to utilize learned dynamics as a correction to an a priori physical model to account for shortcomings in the model, as discussed in chapter 6. This practice of using machine learning models to augment known physical models is the subject of the next section.

3.4 Discrepancy Modeling

A large amount of progress has been made in recent years combining data-driven techniques from machine learning with mathematical models developed from first principles within a domain,

one example of which is the master equation (2.10) derived from first principles to model the dynamics of open quantum systems. It is well known that the mathematical models of dynamical systems are always a simplification of the physical world to a greater or lesser extent, and discrepancy learning is a means to use data-driven techniques to bridge the gap. The motivating principle is that physical models generally capture the big picture very well, and we just need to correct for smaller effects that slip through the cracks. While much smaller than the known drivers of the dynamics, these discrepancies can nevertheless be large enough to significantly impact the utility of the physical model, for example as a means to make predictions or to estimate the values of certain internal parameters from data.

While physical models may not fully capture all salient features of the true dynamics, data-driven techniques have their shortcomings as well. They are typically opaque and difficult to interpret, offering limited physical insight and struggling to extrapolate beyond the boundaries of their training data sets. Discrepancy modeling attempts to leverage the strength of both approaches, forcing interpretability on machine learning models by keeping them tightly coupled to their physically motivated counterparts, while relying on their flexibility to account for real phenomena that are hard to anticipate with a mathematical model.

In practice, this objective can be accomplished in many different ways. One example is a 2019 paper by Kaheman, Kaiser, Strom, Kutz, and Brunton [89] in which the authors leverage a machine learning technique known as Sparse Identification of Nonlinear Dynamics (SINDy) [33] to correct for model discrepancies in a double pendulum on a cart experiment. Due to the chaotic nature of the dynamics, any deviation between the model and the experiment will be amplified to the point that prediction and control is no longer possible, but the authors use the ML model to successfully develop a feed-forward controller for the system.

In another 2019 example from de Silva, Higdon, Brunton, and Kutz [53], SINDy is once again employed to account for subtle effects in modeling falling objects of different size and mass. Specifically, measurement errors and complex fluid dynamics that are extremely difficult to account for are successfully learned using the SINDy model. The premise behind SINDy is to provide a

library of nonlinear terms that could be present in the derivative function f_θ in (3.10) and to learn from data the coefficients associated with these terms in the actual dynamics. Sparsity of the terms is enforced via the loss function and gating during training. The result is entirely interpretable, and gives physical insight into the dynamics while also improving modeling accuracy.

These principles will be leveraged in chapter 6, where we use a neural ODE based model to correct for dynamics not anticipated in the master equation. This will allow for more accuracy in predicting known Hamiltonian or Lindblad terms in the presence of unanticipated terms, calibration errors, or even non-Markovian and nonlinear effects. While not going so far as to leverage SINDy to determine more specific forms for nonlinear effects, this could be an interesting topic for future work.

Chapter 4

Time Crystals

This chapter is adapted from the publication

> K. Tucker, B. Zhu, R. J. Lewis-Swan, J. Marino, F. Jimenez, J. G. Restrepo, and A. M. Rey. "Shattered time: can a dissipative time crystal survive many-body correlations?" New Journal of Physics, 20(12):123003, (2018)

Prologue

In this chapter, we leverage mean field and cumulant methods to explore the properties of a system of two level atoms that serves as an example of a time crystal. A mean field analysis will give insight into how the system synchronizes, while simulations using a cumulant expansion method as well as an exact solver exploiting particle symmetry establish that the system has the necessary properties to make it a time crystal over a specific region of parameter space.

Abstract

We investigate the emergence of a time crystal in a driven-dissipative many-body spin array. In this system the interplay between incoherent spin pumping and collective emission stabilizes a synchronized non-equilibrium steady state which in the thermodynamic limit features a self-generated time-periodic pattern imposed by collective elastic interactions. In contrast to prior realizations where the time symmetry is already broken by an external drive, here it is only spontaneously broken by the elastic exchange interactions and manifest in the two-time correlation

spectrum. Employing a combination of exact numerical calculations and a second-order cumulant expansion, we investigate the impact of many-body correlations on the time crystal formation and establish a connection between the regime where it is stable and a slow growth rate of the mutual information, signalling that the time crystal studied here is an emergent semi-classical out-of-equilibrium state of matter. We also confirm the rigidity of the time crystal to single-particle dephasing. Finally, we discuss an experimental implementation using long-lived dipoles in an optical cavity.

4.1 Introduction

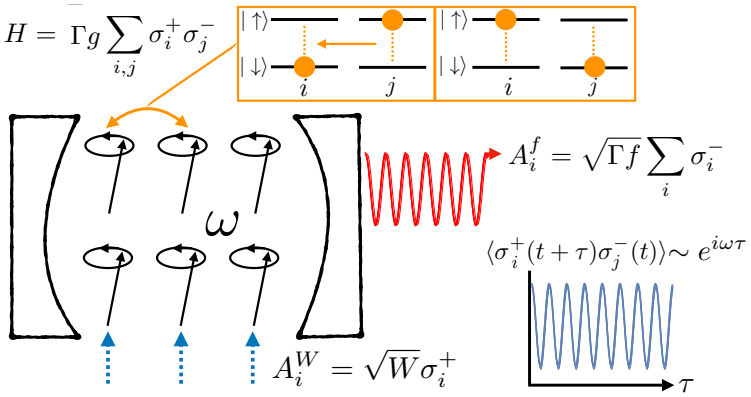
Nature provides many examples of crystalline structures in our day-to-day lives. From the salt and sugar in foods that we eat, to snowflakes falling from the sky, this curious phenomenon is all around us. A crystal is characterized by a spontaneous breaking of spatial symmetry. As molecules are arranged in a lattice, certain spatial dimensions are given preference over others. This occurs due to the properties of the solid itself, rather than according to any external design. These structures also exhibit some amount of robustness to disorder, they do not simply fall apart when acted on by an outside force. Time crystals have been proposed for uses in precise time keeping, quantum metrology, and by Microsoft researchers for use in topological quantum computers.

In 2012, Nobel laureate Frank Wilczek proposed that a similar phenomenon can emerge in time [174]. Specifically, he proposed the existence of phases of quantum matter that spontaneously break time-translational symmetry, without being forced by an external drive, and that are robust to disorder. More formally, a time crystal is a many-body system exhibiting synchronization with an order parameter $\phi(\vec{r}, t)$ having a two-time correlation function converging to a periodic function of time in the thermodynamic limit [173] such that $\langle \phi(\vec{r}, t)\phi(\vec{r}\prime, 0)\rangle \to f(t)$, when the distance $|r - r\prime|$ is sufficiently large. Since being proposed, time crystals have received a considerable amount of attention [156, 131, 149, 104, 61, 91, 92, 168, 169, 179, 180, 69, 7], including a 2015 work by Watanabe et. al [173] showing that time crystals cannot exist in the steady-state, as Wilczek originally suggested, they are instead non-equilibrium phenomena.

A blueprint was developed [179] for the experimental realization of time crystals in periodically driven, interacting many-body quantum systems with spatial disorder, also known as Floquet time crystals [61, 91, 92, 168, 169, 179], and subsequently realized in trapped ions [181] and systems leveraging small imperfections in diamonds [50]. In most cases, time-resolved observables react to the periodic drive to realize subharmonics of the frequency at which they are driven [32, 2, 173, 61, 77, 62, 101, 146, 137, 136, 183, 23]. However, at the time of publication, we are not aware of a realization that spontaneously breaks continuous time symmetry in a quantum many-body system.

In this work, we propose a novel approach that can lead to the first experimental observation of spontaneous continuous time symmetry breaking in a quantum many-body system, by considering an array of incoherently driven long lived dipoles in a cavity that are subject to collective dissipative decay (superradiance) and elastic long range interactions (see figure 4.1). While similar in spirit to recently proposed quantum time crystals with dissipation involving a time periodic steady state in the thermodynamic limit of an open quantum system [69, 85, 45, 34], our system differs in that it employs an incoherent drive, thus avoiding the imposition of an external drive breaking time translational symmetry. Instead, continuous time translational symmetry is spontaneously broken by the subtle interplay between collective interactions and driving processes. In addition, the incoherent drive and disorder result in a system that expands beyond the Dicke subspace of Hilbert space, resulting in certain modeling and simulation challenges but leading to more complex behavior.

Looking at each property of the system in figure 4.1, we see how each impacts the emergence of the periodic steady-state giving rise to the time crystal. Collective dissipation prevents the system from overheating, while the balance between this dissipation and the incoherent drive lead to the stabalization of a syncrhonized non-equilibrium periodic steady-state [120, 19, 185] that is robust to imperfections or environmental disturbances assuming the presence of finite but moderate elastic interactions. The latter interactions are necessary for the formation of the non-equilibrium periodic steady-state, but can destroy it if they are too strong. Thus there is a window of permissible elsastic



Superradiant time crystal

1) incoherent pumping	drives the system out of equilibrium
2) collective emission	prevents heating
3) cavity-mediated elastic interactions	break continuous time symmetry

Figure 4.1: An ensemble of N spin 1/2 particles pumped at rate W, experiencing collective emission at rate $\propto f\Gamma$ and collective spin-exchange interactions (orange), $\propto g\Gamma$, form the basis of the superradiant TC. The elastic interactions imprint collective spin oscillations at frequency ω spontaneously breaking the time translation symmetry (manifested as persistent oscillations in the unequal time spin-spin correlation function in the thermodynamic limit).

interaction strengths that grows as the square root of the particle number. An additional feature is that due to the open nature of the quantum system, the time crystal will emerge regardless of the initial state, in contrast with Hamiltonian systems. Finally, we will show via a spectral analysis of the Liouvillian how these various parameters impact the emergence of the time crystal and the growth of mutual information in the transient dynamics.

4.2 Model

We consider an ensemble of N spin-1/2 particles, whose evolution is described by a master equation for the density matrix $\hat{\rho}$,

$$\frac{\partial \hat{\rho}}{\partial t} = \mathcal{L}[\hat{\rho}] = -i[\hat{H}, \hat{\rho}] + \mathrm{D}[\hat{\rho}], \tag{4.1}$$

$$\hat{H} = g\Gamma \hat{S}^+ \hat{S}^- + \sum_{i=1}^{N} \frac{\delta_i}{2} \hat{\sigma}_i^z, \tag{4.2}$$

where $\hat{S}^{\pm} \equiv \sum_{i=1}^{N} \hat{\sigma}_i^{\pm}$ and $\hat{\sigma}_i^a$ are the Pauli matrices ($a = x, y, z$) acting on spin $i = 1, ..., N$. The first term in the Hamiltonian $\hat{H}$ is a collective spin exchange, an ellastic interaction whereby each particle can exchange a photon with every other one with equivalent rate. We do not normalize the exchange term by N to remain consistent with other experimental work [129, 108]. The second term in the Hamiltonian is the disorder to the detuning from the cavity mode. For simplicity, but without loss of generality, we assume the δ_i's are distributed according to a Lorentzian of width Δ and zero mean.

The dissipator $\mathrm{D}[\hat{\rho}] = \mathrm{L}^W(\hat{\rho}) + \mathrm{L}^f(\hat{\rho})$ encodes two channels via the usual Lindblad superoperator $L[\hat{O}](\hat{\rho}) = \hat{O}\hat{\rho}\hat{O}^\dagger - \frac{1}{2}\left\{\hat{O}^\dagger\hat{O}, \hat{\rho}\right\}$:

$$\mathrm{L}^W(\hat{\rho}) \equiv \sum_i L[\hat{A}_i^W](\hat{\rho})$$

$$\mathrm{L}^f(\hat{\rho}) \equiv \sum_i L[\hat{A}_i^f](\hat{\rho})$$

where local, incoherent pumping is described by $\hat{A}_i^W = \sqrt{W}\sigma_i^+$, and collective emission is described by $\hat{A}^f = \sqrt{f\Gamma}\hat{S}^-$. The parameter Γ sets the scale of the spin-spin interactions, while g and f are dimensionless parameters characterizing the relative strength of their corresponding elastic and dissipative part respectively. The use of a master equation [103, 39] to deal with the dissipative processes is extremely accurate for the experimental systems discussed below [129, 155]. Note that the dissipative part of (4.1) $\mathrm{D}[\hat{\rho}]$ will preserve $U(1)$ phase symmetry, i.e. will be invariant under the transformation $\hat{\sigma}_j^+ \to \hat{\sigma}_j^+ e^{i\phi}$. In addition, this term will allow for the dynamics to proceed through different eigenspaces of collective total spin $\boldsymbol{S}^2$ (with eigenvalues $S(S+1)$), therefore dynamics will not be restricted to the $S = N/2$ manifold of the initial state.

We now turn our attention to establishing that the above system has the needed to properties to be considered a time crystal. To do this, we must establish that it exhibits synchronization, an oscillation in the non-equilibrium steady-state that is necessarily independent of any external drive since pumping is incoherent, and robustness to disorder in the system. We begin by an analysis in the mean-field to establish synchronization before considering quantum correlations to establish the other properties.

4.3 Mean-field analysis

In a mean-field approach, we assume that the many-body density matrix of the system can be factorized as the tensor product of single site density matrices $\hat{\rho} = \bigotimes_{a=1}^{N} \hat{\rho}_a$. We can derive the equations of motion for $\hat{\sigma}_a^+$ and $\hat{\sigma}_a^z$ via the master equation (4.1) closing to first order expectations by noting in this case that $\langle \hat{\sigma}_a^\alpha \hat{\sigma}_b^\beta \rangle = \langle \hat{\sigma}_a^\alpha \rangle \langle \hat{\sigma}_b^\beta \rangle$. The components of the single site density matrix $\hat{\rho}_a$ can be visualized as a Bloch vector, $\vec{\mathcal{S}}_a = \{R_a \cos\phi_a,\, R_a \sin\phi_a,\, s_a\}$ where $\langle \hat{\sigma}_a^+ \rangle \equiv R_a e^{-i\phi_a}$ and $\langle \hat{\sigma}_a^z \rangle \equiv s_a$. This approach results in the system of $3N$ coupled ordinary differential equations

$$\begin{aligned}
\frac{ds_a}{dt} &= -\Gamma R_a \sum_{b\neq a} R_b \left[f\cos(\phi_b - \phi_a) + 2g\sin(\phi_b - \phi_a) \right] \\
&\quad - f\Gamma\left(\frac{1}{2} + s_a\right) + W\left(\frac{1}{2} - s_a\right), \\
\frac{dR_a}{dt} &= -\frac{(f\Gamma + W)}{2} R_a + \Gamma s_a \sum_{b\neq a} R_b \left[f\cos(\phi_b - \phi_a) + 2g\sin(\phi_b - \phi_a) \right], \\
\frac{d\phi_a}{dt} &= -\delta_a + \frac{\Gamma s_a}{R_a} \sum_{b\neq a} R_b \left[f\sin(\phi_b - \phi_a) - 2g\cos(\phi_b - \phi_a) \right].
\end{aligned} \tag{4.3}$$

The equation that determines the evolution of the phase ϕ_a is remarkably similar to the Kuramoto-Sakaguchi model for synchronization of phase oscillators [151], which can be written as

$$\frac{d\phi_a}{dt} = \delta_a + \sum_{b\neq a} \left[f\sin(\phi_b - \phi_a) - g\cos(\phi_b - \phi_a) \right]. \tag{4.4}$$

While the evolution of the phases in Eq. (4.3) is coupled to the other dynamical variables, we find that our model also supports synchronized solutions.

Comparing terms in the mean-field phase equation in (4.3) and the Kuramoto-Sakaguchi model in (4.4), we see that the collective dissipation term f is associated with the sine term in the Kuramoto model [100] responsible for synchronization. For synchronization to occur, the coupling strength per oscillator, here proportional to $f\Gamma s_i > 0$, must be positive and large enough to compensate for the dephasing generated by the different single particle frequencies. This condition is only possible in the presence of incoherent pumping and thus intrinsic to our setup since a coherent drive does not lead to population inversion in the steady-state [38, 59]. The elastic interaction term g is associated with the cosine term of the Kuramoto-Sakaguchi model [151] that

determines the frequency of the collective oscillation. The effective field not only induces a net collective precession but also favors spin alignment and self-rephasing against the depolarization induced by the inhomogeneous field as theoretically and experimentally demonstrated in prior work [60, 144, 54, 95, 119, 129, 28]. Thus we see how both the inelastic and elastic interactions are responsible for the emergence of the time crystal.

We now demonstrate the existence of a synchronized, periodic steady-state in the mean-field in the absence of detunings by setting all of the δ_a values in (4.1) to zero and determining the properties of the steady-state solution. Later we will restore the detunings and show that this state is robust to this kind of disorder. We define an order parameter $S^+ = \sum_a R_a e^{i\phi_a} = NZe^{-i\omega_{\mathrm{MF}}t}$ and focus on its normalized magnitude Z and frequency ω_{MF}. The system synchronizes if we can find a stationary state in the co-rotating frame with frequency ω_{MF} where Z acquires a positive real value, which is self-consistently determined from the system's parameters [120, 185]. To show this, following the analysis in [185], we start with the ansatz $\phi_a = \omega_{\mathrm{MF}}t$, $R_a = R, \dot{R} = 0$, $s_a = s, \dot{s} = 0$. This gives

$$s = \frac{f\Gamma + W}{2fN\Gamma}, \tag{4.5}$$

$$R = \frac{\sqrt{fN\Gamma(W - f\Gamma) - (f\Gamma + W)^2}}{\sqrt{2}fN\Gamma}, \tag{4.6}$$

$$\omega_{\mathrm{MF}} = \frac{g(f\Gamma + W)}{f}, \tag{4.7}$$

and, in this case, we have $Z = R$. The order parameter is nonzero only when $fN\Gamma(W - \Gamma) - (f\Gamma + W)^2 > 0$, which occurs in the finite range $N - 2 - \sqrt{N^2 - 8N} < 2W/(f\Gamma) < N - 2 + \sqrt{N^2 - 8N}$. For $N \gg 1$, this simplifies to $f\Gamma < W < fN\Gamma$. In particular, we note that when $W < f\Gamma$ there is no positive solution. Thus, only intermediate values of pumping can bring the system to a synchronized state. There is a synchronized solution (i.e., $Z > 0$) only if $N \geq 8$, corresponding to an onset at $W = 3f\Gamma$.

We can find an optimal pumping value W_{opt} at which the order parameter achieves a maxi-

mum. In the thermodynamic (large N) limit, the optimal pumping is given by

$$W_{opt} = \frac{fN\Gamma}{2}. \tag{4.8}$$

At optimal pumping, the oscillation frequency is

$$\omega_{\mathrm{MF}}^{opt} = \frac{gN\Gamma}{2}, \tag{4.9}$$

and the order parameter takes the value $Z_{opt} = 1/\sqrt{8}$. Additional results from the mean-field analysis including disorder can be found in appendix A.

4.4 Quantum model

To understand how the time crystal emerges when considering effects beyond the mean-field, we will be interested in analyzing an order parameter given by the two-time correlation function

$$\mathcal{C}(\tau) \equiv \lim_{t\to\infty} \frac{\sum_{i=1,j=1}^{N} \langle {\sigma_i}^+(t+\tau){\sigma_j}^-(t)\rangle}{N^2}. \tag{4.10}$$

Our analysis is based on an efficient exact numerical solution of the master equation (4.1) that uses spin permutation symmetry to drastically improve the exponential scaling of the Liouville space from 4^N to $\mathcal{O}(N^3)$ [67, 178, 154]. $\mathcal{C}(\tau)$ is computed via the linear quantum regression theorem [39], $\langle \hat{\sigma}_i^+(t+\tau)\hat{\sigma}_j^-(t)\rangle = \mathrm{Tr}\left[\hat{\sigma}_i^+ e^{\mathcal{L}\tau}[\hat{\sigma}_j^-\hat{\rho}(t)]\right]$, which is exact for the case of a master equation.

Setting $\tau = 0$ and taking the square-root, we see that

$$Z_Q \equiv \sqrt{\mathcal{C}(0)} = \frac{1}{N}\sqrt{\sum_{i=1,j=1}^{N} \langle {\sigma_i}^+(t){\sigma_j}^-(t)\rangle} = \frac{1}{N}\sqrt{\langle S^+S^-\rangle}$$

is the analogous value of the order parameter Z from the mean-field analysis allowing for quantum correlations, as applying the mean-field assumption to this expression reduces it to $|S^+(t)|/N$. Similarly to its mean-field counterpart, the value of Z_Q is nonzero within a certain window of pumping, reaching an optimal value $Z_Q^{opt} \approx 1/\sqrt{8}$ at W_{opt} approximately independent of g corresponding to maximal synchronization. Looking at $C(\tau)$ for $\tau > 0$, we see a dependence on g similar to the mean-field case in the range $1/N \lesssim g/f \lesssim \sqrt{N}$, where $C(\tau)$ oscillates with the mean-field frequency

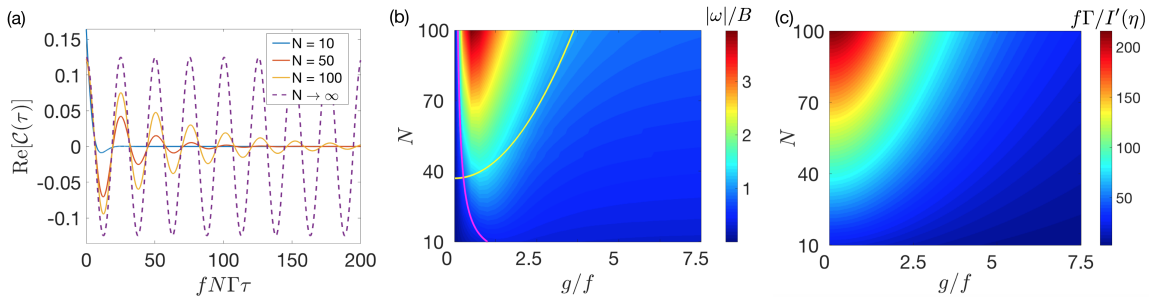


Figure 4.2: (a) Real part of $\mathcal{C}(\tau)$ as a function of the characteristic time $\eta \equiv fN\Gamma\tau$ at optimal pumping and $g/f = 1/2$ for a system of $N = 10$ (blue), $N = 50$ (red), and $N = 100$ (orange) spins along with the finite size scaling prediction in the thermodynamic limit (purple). (b) Extracted ratio of the absolute value of $\mathcal{C}(\tau)$ angular frequency $|\omega|$ over its spectral width B vs system size N and interaction coupling g/f. We also show a frequency contour corresponding to $\omega_{\rm MF}^{opt}/f\Gamma \sim 5$ (purple) and a contour of mutual information growth corresponding to $f\Gamma/I'(\eta = 0.03) \sim 80$ (yellow). (c) Growth rate of two-particle mutual information at short characteristic times (here we set $\eta = 0.03$) starting from a maximally coherent array $\langle\sigma_i^x(0)\rangle = 1$.

$\omega_{\rm MF}^{opt}$. While these oscialltions decay, they appear to become persistent with constant amplitude as N increases, thus being consistent with the requirement for a time crystal that there be stable oscillations in the thermodynamic limit. This can be seen in figure 4.2 (a), where $\mathcal{C}(\tau)$ is plotted as a function of the characteristic time $\eta \equiv fN\Gamma\tau$ for $N = 10$, 50 and $N = 100$.

In contrast with the mean field, which showed unbounded frequency of oscillations as a function of N and g, it is known that large many-body correlations can lead to melting of the time crystal [179, 91]. To consider our time crystal stable, we need to see a non-vanishing number of oscillations before the decay seen in figure 4.2 (a) can eliminate them. We show this region of stability in figure 4.2 (b), which plots the ratio of the angular frequency of the $C(\tau)$ oscillations ω to the decay rate, or bandwidth B, defined as the full-width-at-half-max (FWHM) of the discrete Fourier transform of the numerically determined solution.

We can understand the boundaries of this region by looking at the eigenvalues of the Liouvillian $\mathcal{L}$ defining (4.1), which determines the dynamics of $C(\tau)$ via the quantum regression formula. We can see that the eigenvalues with non-zero imaginary part and smallest real part will determine the oscillatory behavior of the two-time correlator for the stable steady-state solution, while

modes with larger negative real part will decay away relatively quickly and contribute less to the dynamics. In our analysis, the details of which can be found in appendix A, for moderate values of g/f there will exist eigenvalues of $\mathcal{L}$ with non-zero imaginary part and small, negative real part and eigenvectors that are not orthogonal to $\lim_{t\to\infty} S^-\hat{\rho}(t)$, the state that $e^{\mathcal{L}\tau}$ operates on in the quantum regression theorem. These eigenvalues enable an oscillatory solution and determine its frequency. As g/f grows, however, so does the negative real part of these eigenvalues, increasing the decay rate and damping the time crystal.

We can also see a relationship between the stable region of the time crystal and the development of many-body correlations during the transient dynamics, seen in the rate of increase of mutual information $I_{AB} \equiv S_A + S_B - S_{AB}$, where $S_\alpha = -\operatorname{Tr}\left[\hat{\rho}_\alpha \log[\hat{\rho}_\alpha]\right]$ is the Von Neumann entropy computed from the reduced density matrix $\hat{\rho}_\alpha$ of the subsystem $\alpha = A, B, AB$ (AB is the joint subsystem).

As seen in detail in appendix A, the same relationship exists between the eigenvalues of $\mathcal{L}$ and the growth rate of I_{AB} for small times in the transient dynamics as was seen above with regard to damping of the time crystal. Specifically, we find that in the large N limit and short characteristic times $\eta \ll 1$ (respect to the oscillation period), with $\eta \equiv fN\Gamma\tau$, the growth rate of the mutual information approaches $dI_{AB}/d\tau \sim \frac{f\Gamma}{2}\left(\frac{1}{N} + \frac{4g^2}{Nf^2}\eta\right)$ and thus remains irrelevant for $g/f \lesssim \sqrt{N}$. This parameter regime is consistent with the range of g/f values where we observe that the time crystal forms. Outside this region, I_{AB} grows rapidly with increasing g/f (see figure 4.2 (c)).

This consistency between many-body correlations and damping is illustrated in figure 4.2 (b), where a contour of figure 4.2 (c) given by $g/f \propto \sqrt{N}$ is shown bounding the stability region from above. We also illustrate a frequency $\omega_{\mathrm{MF}}^{opt}/f\Gamma$ contour which corresponds to $g/f \propto 1/N$ bounding the region from below as sufficient g/f is needed to sustain oscillations. From these considerations we can conclude that the superradiant crystal only exists in the parameter regime where many-body correlations are subdominant and thus it can be regarded as an emergent semi-classical non-equilibrium state of matter.

Many of the results in this section will prove to be consistent with results employing a second-

order cumulant expansion [22], discussed in the next section, an approximation which is necessary to assess the case with disorder $\Delta \neq 0$.

4.5 Cumulant expansion analysis

To obtain a model that can capture both many-body correlations and nonzero disorder, we turn to a second-order cumulant expansion, which assumes joint cumulants of order three and higher are zero, resulting in the following expansion for third-order expectations:

$$\langle \hat{\sigma}_a^\alpha \hat{\sigma}_b^\beta \hat{\sigma}_c^\gamma \rangle \approx \langle \hat{\sigma}_a^\alpha \hat{\sigma}_b^\beta \rangle \langle \hat{\sigma}_c^\gamma \rangle + \langle \hat{\sigma}_b^\beta \hat{\sigma}_c^\gamma \rangle \langle \hat{\sigma}_a^\alpha \rangle + \langle \hat{\sigma}_a^\alpha \hat{\sigma}_c^\gamma \rangle \langle \hat{\sigma}_b^\beta \rangle - 2 \langle \hat{\sigma}_a^\alpha \rangle \langle \hat{\sigma}_b^\beta \rangle \langle \hat{\sigma}_c^\gamma \rangle. \tag{4.11}$$

In this Section, we present the equations of motion that result, obtain results for the homogeneous system that connect the two-time correlation oscillation frequency to the mean-field frequency, compare the mean-field and cumulant frequencies for systems with nonzero disorder, and perform benchmarking of the cumulant model against exact solutions.

The full equations of motion using the cumulant expansion can be found in appendix A, but they together provide approximate equations of motion as an $\mathcal{O}(N^2)$ non-linear system of ODEs. Assuming $\Delta = 0$, all dependent variables are independent of particle number. This means expectations in the cumulant equations of motion that differ only by particle indices are equal.

Taking this into consideration, we arrive at the following system for $\langle\hat{\sigma}_a^z\rangle$, $\langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle$, and $\langle\hat{\sigma}_a^z\hat{\sigma}_b^z\rangle$:

$$\frac{d\langle\hat{\sigma}_a^z\rangle}{dt} = -2\Gamma f(N-1)\mathrm{Re}(\langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle) + 4\Gamma g(N-1)\mathrm{Im}(\langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle) - \langle\hat{\sigma}_a^z\rangle(\Gamma f + W) - \Gamma f + W, \tag{4.12}$$

$$\begin{aligned}\frac{d\langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle}{dt} &= \frac{1}{2}\Gamma(N-2)(f-i2g)\\ &\quad\big(-2\langle\hat{\sigma}_a^+\rangle\langle\hat{\sigma}_a^z\rangle\langle\hat{\sigma}_a^+\rangle^* + \langle\hat{\sigma}_a^z\hat{\sigma}_b^+\rangle\langle\hat{\sigma}_a^+\rangle^* + \langle\hat{\sigma}_a^+\rangle\langle\hat{\sigma}_a^z\hat{\sigma}_b^+\rangle^* + \langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle^*\langle\hat{\sigma}_a^z\rangle\big)\\ &\quad+\frac{1}{2}\Gamma(N-2)(f+i2g)\\ &\quad\big(-2\langle\hat{\sigma}_a^+\rangle\langle\hat{\sigma}_a^z\rangle\langle\hat{\sigma}_a^+\rangle^* + \langle\hat{\sigma}_a^z\hat{\sigma}_b^+\rangle\langle\hat{\sigma}_a^+\rangle^* + \langle\hat{\sigma}_a^+\rangle\langle\hat{\sigma}_a^z\hat{\sigma}_b^+\rangle^* + \langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle\langle\hat{\sigma}_a^z\rangle\big)\\ &\quad+\frac{1}{2}\Gamma f(\langle\hat{\sigma}_a^z\rangle + \langle\hat{\sigma}_a^z\hat{\sigma}_b^z\rangle) + \langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle(-(\Gamma f + W)),\end{aligned} \tag{4.13}$$

$$\begin{aligned}\frac{d\langle\hat{\sigma}_a^z\hat{\sigma}_b^z\rangle}{dt} &= -4\Gamma(N-2)\mathrm{Re}\left((f+i2g)\left(-2\langle\hat{\sigma}_a^+\rangle\langle\hat{\sigma}_a^z\rangle\langle\hat{\sigma}_a^+\rangle^* + \langle\hat{\sigma}_a^z\hat{\sigma}_b^+\rangle\langle\hat{\sigma}_a^+\rangle^* + \langle\hat{\sigma}_a^+\rangle\langle\hat{\sigma}_a^z\hat{\sigma}_b^+\rangle^* + \langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle\langle\hat{\sigma}_a^z\rangle\right)\right)\\ &\quad+4\Gamma f\mathrm{Re}\left(\langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle\right) + 2\langle\hat{\sigma}_a^z\rangle(W-\Gamma f) - 2\langle\hat{\sigma}_a^z\hat{\sigma}_b^z\rangle(\Gamma f + W).\end{aligned} \tag{4.14}$$

We would like to determine steady state values for the various parameters. Simulations and exact solutions for small systems motivate the following ansatz:

$$\langle\hat{\sigma}_a^+\rangle = \langle\hat{\sigma}_a^z\hat{\sigma}_b^+\rangle = \langle\hat{\sigma}_a^+\hat{\sigma}_b^+\rangle = 0, \tag{4.15}$$

$$\langle\hat{\sigma}_a^z\rangle = \alpha, \tag{4.16}$$

$$\langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle = \beta, \tag{4.17}$$

$$\langle\hat{\sigma}_a^z\hat{\sigma}_b^z\rangle = \gamma, \tag{4.18}$$

where α, β, γ are real and time-independent. The above system then reduces to

$$0 = -2\Gamma f(N-1)\beta - (\Gamma f + W)\alpha - \Gamma f + W, \tag{4.19}$$

$$0 = \Gamma f(N-2)\alpha\beta + \frac{1}{2}\Gamma f(\alpha+\gamma) - (\Gamma f + W)\beta, \tag{4.20}$$

$$0 = -4\Gamma f(N-2)\alpha\beta + 4\Gamma f\beta + 2(W-\Gamma f)\alpha - 2(\Gamma f + W)\gamma. \tag{4.21}$$

This nonlinear system can be solved, giving the following steady-state values for $\langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle$ and $\langle\hat{\sigma}_a^z\rangle$:

$$\begin{aligned}\langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle_{SS} &= \frac{f^2\Gamma^2(2-3N) + f\Gamma(N-5)W - W^2}{4f^2\Gamma^2(N-2)(N-1)}\\ &\quad+\frac{\sqrt{\left(f^2\Gamma^2(3N-2) - f\Gamma(N-5)W + W^2\right)^2 - 8f^3\Gamma^3(N-2)(N-1)(f\Gamma - W)}}{4f^2\Gamma^2(N-2)(N-1)},\end{aligned} \tag{4.22}$$

$$\langle\hat{\sigma}_a^z\rangle_{SS} = \frac{f^2\Gamma^2(N+2) + f\Gamma(N+1)W + W^2}{2f\Gamma(N-2)(f\Gamma+W)} - \frac{\sqrt{\left(f^2\Gamma^2(3N-2) - f\Gamma(N-5)W + W^2\right)^2 - 8f^3\Gamma^3(N-2)(N-1)(f\Gamma-W)}}{2f\Gamma(N-2)(f\Gamma+W)}. \tag{4.23}$$

Again using the notation $W = wN$ and taking $N \gg 1$, Eq. (4.22) simplifies to

$$\begin{aligned}\langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle_{SS} &= -\frac{1}{4\Gamma^2 f^2}\left[w^2 - \Gamma f w\right] + \frac{1}{4\Gamma^2 f^2}\sqrt{\left[w^2 - \Gamma f w\right]^2} \\ &= -\frac{1}{4\Gamma^2 f^2} w\left(w - \Gamma f\right) + \frac{1}{4\Gamma^2 f^2}\left|w\right|\left|w - \Gamma f\right| \\ &= \begin{cases} -\frac{1}{2\Gamma^2 f^2} w\left(w - \Gamma f\right) & w \in [0,\, \Gamma f] \\ 0 & w > \Gamma f \end{cases}\end{aligned} \tag{4.24}$$

From here we see that $w_{opt} = f\Gamma/2$ and $\langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle_{opt} = 1/8$. Note that this is the same value for optimal pumping seen in the mean-field case in Eq. (4.8) and the same optimal order parameter, noting that $Z_{opt}^2 = \langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle_{opt} = 1/8$. Taking $N \gg 1$ in Eq. (4.23) gives

$$\langle\hat{\sigma}_a^z\rangle_{SS} = \frac{1}{2f\Gamma}\left[(f\Gamma + w) - |f\Gamma - w|\right] = \frac{w}{f\Gamma}$$

where we have assumed $w < f\Gamma$ in order to be in the synchronized regime according to Eq. (4.24). If we substitute in w_{opt}, we get $\langle\hat{\sigma}_a^z\rangle_{SS} = 1/2$. Note that the steady-state values derived above are independent of g. This is consistent with numerical solutions of the exact system when $\Delta = 0$. We now have a value for optimal pumping and a corresponding steady-state value for $\langle\hat{\sigma}_a^z\rangle$, which will now be used to determine two-time correlation decay rate and frequency.

We can study the analytic properties of $\mathcal{C}(\tau)$ in the case where $\Delta = 0$ via a spectral analysis of the linear equations of motion near the steady-state. Applying the quantum regression theorem to our equations of motion, and assuming zero disorder and particle symmetry results in the following equations for the two-time correlation function in the limit $t \to \infty$:

$$\begin{aligned}\frac{d}{d\tau}\langle\hat{\sigma}_a^+(t+\tau)\hat{\sigma}_b^-(t)\rangle &= -\frac{\Gamma f + W}{2}\langle\hat{\sigma}_a^+(t+\tau)\hat{\sigma}_b^-(t)\rangle + \frac{\Gamma}{2}(f - i2g)\langle\hat{\sigma}_a^z(t)\rangle\langle\hat{\sigma}_a^+(t+\tau)\hat{\sigma}_a^-(t)\rangle \\ &\quad + \frac{\Gamma}{2}(f - i2g)(N-2)\langle\hat{\sigma}_a^z(t)\rangle\langle\hat{\sigma}_a^+(t+\tau)\hat{\sigma}_b^-(t)\rangle, \\ \frac{d}{d\tau}\langle\hat{\sigma}_a^+(t+\tau)\hat{\sigma}_a^-(t)\rangle &= -\frac{\Gamma f + W}{2}\langle\hat{\sigma}_a^+(t+\tau)\hat{\sigma}_a^-(t)\rangle + \frac{\Gamma}{2}(f - i2g)(N-1)\langle\hat{\sigma}_a^z(t)\rangle\langle\hat{\sigma}_a^+(t+\tau)\hat{\sigma}_b^-(t)\rangle,\end{aligned}$$

where we distinguish between diagonal and off-diagonal correlation functions with the subscripts, but no longer particle number, so this is simply a linear system of two equations with the matrix

$$\begin{pmatrix} -\frac{\Gamma f+W}{2} + \frac{\Gamma}{2}(f-i2g)(N-2)\langle\hat{\sigma}_a^z\rangle_{SS} & \frac{\Gamma}{2}(f-i2g)\langle\hat{\sigma}_a^z\rangle_{SS} \\ \\ \frac{\Gamma}{2}(f-i2g)(N-1)\langle\hat{\sigma}_a^z\rangle_{SS} & -\frac{\Gamma f+W}{2} \end{pmatrix},$$

where we have used $\lim_{t\to\infty}\langle\hat{\sigma}_a^z(t)\rangle = \langle\hat{\sigma}_a^z\rangle_{SS}$. Inserting the steady-state value $\langle\hat{\sigma}_a^z\rangle_{SS} = w/\Gamma f$, the eigenvalues are:

$$\lambda_1 = -\frac{w(N+1)+f\Gamma}{2} + i\frac{wg}{f},$$
$$\lambda_2 = -\frac{w+f\Gamma}{2} - i\frac{wg}{f}(N-1).$$

The real parts of these eigenvalues determine the decay rate of the two-time correlation function. Note that $\mathrm{Re}(\lambda_1)$ is negative and scales linearly with N, so it will cause rapid decay and the corresponding mode will not contribute meaningfully to oscillations in the thermodynamic limit. In contrast, the real part of λ_2 is independent of N, so oscillation frequency for large N will be determined by its imaginary part. This will be, for $N \gg 1$,

$$\omega_{\mathrm{C}} = \frac{wgN}{f},$$

which agrees with the mean-field estimate to leading order in $1/N$. Note also that the real part of λ_2 is proportional to $f\Gamma$ at optimal pumping, and therefore so is the decay rate. We also notice, however, that the decay is independent of g. This is not, as we have seen, in agreement with the exact solution. This constrains the window in g where the cumulant can be applied to small values, as disparity in decay rate is small in this regime as we will see below. For all numerically derived results, such as in figure 4.2, spectral width B is used as a proxy for the decay rate given here by $|\mathrm{Re}[\lambda_2]|$.

Further results comparing the cumulant to the mean-field approximation in the case with disorder, as well as benchmarking of the approximation against results from simulation of the exact system in the permutaitionally invariant case, can be found in appendix A.

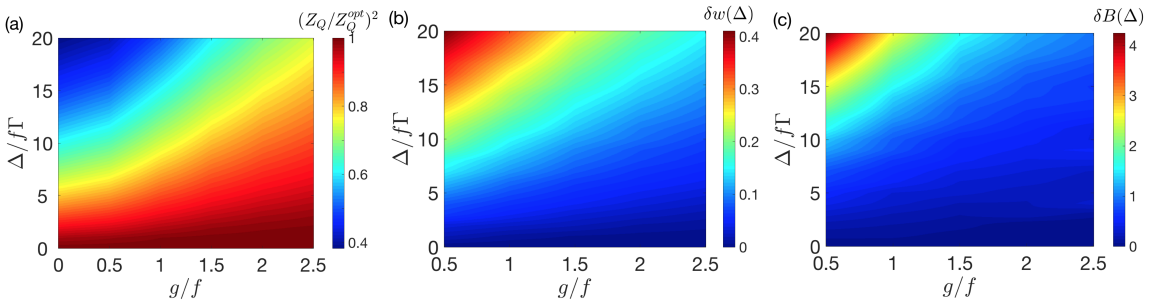


Figure 4.3: Finite interactions protect the TC against weak disorder as can be seen in: (a) The robustness of the averaged $\mathcal{C}(0)$, the insensitivity of the time crystal's (b) frequency, $\delta\omega(\Delta)$, and (c) spectral width, $\delta B(\Delta)$, to weak disorder $\Delta/(f\Gamma)$ for the relevant window of interactions g/f. All plots are computed using the second order cumulant expansion at optimal pumping W_{opt} for $N = 100$ spins.

4.6 Robustness to disorder

We now investigate the case of inhomogeneous dephasing, $\delta_i \neq 0$, to demonstrate the robustness of the time crystal to this form of disorder. We will show that synchronization, frequency, and decay rate all persist in the presence of weak disorder, and that the interactions g/f provide a protective effect increasing robustness over the region of stability of the time crystal. For numerical results shown in this section, we will rely on the cumulant expansion discussed in the last section as exact solvers are not tractable except for small systems where $N \lesssim 15$. Fortunately we find excellent agreement between the cumulant expansion and the exact solution where results are accessible for the relevant parameter regimes as shown in appendix A.

Figure 4.3 (a) shows robustness of the synchronization order parameter $Z_Q \equiv \sqrt{\mathcal{C}(0)}$ to weak disorder $\Delta/f\Gamma$ at optimal pumping W_{opt} over the stable region in g/f. As can be seen, the value is at or near the optimal over the entire region in g/f for small disorder, with a protective effect allowing for more disorder as g/f grows larger. One observes that finite elastic interactions protect the synchronized state against disorder, preserve phase coherence and favor spin alignment. While similar phase locking effects in the transient dynamics have been experimentally reported in cold atom experiments [60, 144, 54, 95, 119, 129, 28], the interesting feature observed here is that the phase locking is achieved in the steady state of a driven dissipative system.

Figure 4.3 (b) and (c) show robustness in the change of frequency $\delta\omega(\Delta) \equiv [\omega(\Delta) - \omega(0)]/\omega(0)$ and bandwidth $\delta B(\Delta) \equiv [B(\Delta) - B(0)]/B(0)$, respectively, of the averaged two-time correlation function $C(\tau)$ as given by the cumulant expansion, where weak disorder induces only a small impact and once again this impact diminishes with increasing g/f. The observed rigidity of the frequency also agrees with the simpler mean-field predictions found in appendix A, which allow us to derive an analytic expression for the protection in the weak disorder limit:

$$\delta\omega(\Delta) \sim \frac{\sqrt{8}\Delta}{N\Gamma\sqrt{f^2 + 2g^2}},$$

where we observe the $1/N$ suppression gained from the collective nature of the elastic and dissipative interactions.

4.7 Experimental realization and outlook

The superradiant crystal can be directly realized using an array of incoherently pumped atomic dipoles tightly trapped by a deep optical lattice that is supported by an optical cavity. The cavity couples two relevant internal states of the atoms, and operates in the bad cavity limit where the bare atomic linewidth γ is significantly smaller than the cavity linewidth κ. In this regime the cavity photons do not participate actively in the dynamics but instead mediate collective dissipative decay (superradiant emission) [73, 120, 19], with $f\Gamma \propto \kappa/(4\delta_c^2+\kappa^2)$, and elastic exchange interactions, with $g\Gamma \propto \delta_c/(4\delta_c^2 + \kappa^2)$, which can be independently controlled by varying the cavity detunning δ_c from the atomic transition. The signature of the TC can then be directly observed in the spectrum of the light leaked from the cavity [129]. A similar implementation can be realized by replacing the cavity photons by phonons in an ion crystal [155]. In the case of the cavity, the order of magnitude for $f\Gamma/2\pi$ and $g\Gamma/2\pi$ is approximately 10^{-4} Hz. For typical atom number in the cavity, $N \approx 10^5$, the TC oscillation frequency approaches ≈ 10 Hz. In the case of the ion crystal, we have $f\Gamma/2\pi = g\Gamma/2\pi \approx 6$ Hz. In this case the TC oscillation frequency for typical ion number $N \approx 10^2$ approaches $\approx 10^3$ Hz.

Having demonstrated the rigidity of the TC to dephasing, now we discuss its rigidity to

variations in the system's parameters. For the proposed implementation, $\omega_{\rm MF}^{opt} \propto N\delta_c/(4\delta_c^2+\kappa^2)$. From this expression, one can see $\omega_{\rm MF}^{opt}$ is not highly sensitive to variations in the cavity linewidth, κ, but on the contrary it is linearly sensitive to variations on δ_c and N. Systematics in the cavity detuning, nevertheless, can be currently controlled at the subhertz level by locking the cavity to a state-of-the-art clock laser [18]. Fluctuations in N can be also suppressed by operating the system in a three dimensional optical lattice in the band or Mott insulator regimes [35] and spectroscopically selecting a fixed region of the atomic array [118].

In summary, we have proposed and analyzed a novel approach for realizing a time crystal in a many-body driven dissipative quantum system that spontaneously breaks continuous time symmetry without a periodic external drive. We have established that it exhibits synchronization, an oscillatory non-equilibrium steady-state, and robustness to disorder. We have also connected the emergence of the time crystal to a parameter regime exhibiting minimal many-body correlations, thus establishing it as an emergent semi-classical non-equilibrium state of matter. Finally, we have proposed a number of potential platforms via which the time crystal may be realized in experiment.

It is important to emphasize that this system is fundamentally distinct from the prototypical laser. This can be seen from the fact that the working mechanism of a laser is stimulated emission, an ingredient absent in our setup. Lasing action is possible even in a single atom system or in the absence of coupling between the atomic dipoles. The superradiant TC, on the contrary, is a genuine many-body phenomenon that happens in the bad cavity limit where the mean photon number in the cavity is less than one. However, even without stimulated emission, superradiance can happen due to collective interactions in a many-body array of long lived atomic dipoles. The superradiant TC is thus a genuine many-body phenomenon which can produce spectrally pure light and might find direct applications in "quantum-interaction enhanced" sensing.

4.8 Acknowledgements

We acknowledge useful discussions with I. Bloch, M. Foss-Feig, M. Holland, A. Shankar, J. Thompson and P. Zoller. The authors acknowledge support from Defense Advanced Research

Projects Agency (DARPA) and Army Research Office grant W911NF-16-1-0576, NSF grant PHY1820885, JILA-NSF grant PFC-173400, and the Air Force Office of Scientific Research FA9550-18-1-0319 and its Multidisciplinary University Research Initiative grant FA9550-13-1-0086, and NIST. JM is supported by the European Union's Framework Programme for Research and Innovation Horizon 2020 2014-2020) under the Marie Sklodowska-Curie Grant Agreement No. 745608 ('QUAKE4PRELIMAT'). B. Zhu was supported by the NSF through a grant for the Institute for Theoretical Atomic, Molecular, and Optical Physics at Harvard University and the Smithsonian Astrophysical Observatory.

Chapter 5

Enhanced Spin Squeezing

This chapter is adapted from the publication

> K. Tucker, D. Barberena, R. J. Lewis-Swan, J. K Thompson, J. G. Restrepo, A. M. Rey, "Facilitating spin squeezing generated by collective dynamics with single-particle decoherence", Phys. Rev. A 102, 051701(R) (2020)

Prologue

In this chapter, we leverage a mean field analysis to establish a phase transition for a system of two level atoms, dividing regions of phase space where we see different amounts of spin squeezing in a collective system. We confirm this transition through a cumulant and Monte Carlo wave function analysis in the Dicke basis, and discover a counter-intuitive relationship between single particle relaxation and spin squeezing in the transient dynamics that has potential applications in quantum sensors.

Abstract

We study the generation of spin-squeezing in arrays of long-lived dipoles subject to collective emission, coherent drive, elastic interactions, and single-particle relaxation. It is found that not only does single-particle relaxation not necessarily degrade the squeezing generated in the collective dynamics, but the interplay of single-particle and collective effects can in fact facilitate the generation of squeezing in a specific parameter regime. This latter behavior is connected to the

dynamical self-tuning of the system through a dissipative phase transition that is present in the collective system alone. Our findings will be applicable to next-generation quantum sensors with an eye towards atomic clocks in cavity-QED set-ups and trapped ion systems.

5.1 Introduction

The preparation of entangled states has become a topic of great interest due to its many applications in quantum metrology [134] and quantum information [24, 25]. Dissipation can often be an obstacle to entanglement generation, motivating conventional protocols to generate entanglement via coherent dynamics and seek to minimize the decoherence induced by couplings to the environment [110, 80]. Recent work, however, has established that it can be achieved in driven dissipative systems, where the environment can prove to be a powerful resource for entanglement generation under the right conditions. For example, it has been shown that an arbitrary entangled pure state of a system can be reached through Markovian dynamics governed by a master equation if the appropriate reservoir and system-reservoir couplings are established [98, 56, 51, 90, 99, 112, 138].

These kinds of systems can, in general, exhibit complex physics but they can be difficult to realize. A convenient platform, however, has emerged in the case of collective spin systems in the form of ultracold atoms coupled to an optical cavity or trapped ion arrays, where the interplay between dissipation and a coherent drive can be realized with precision [26, 129, 125, 109, 52, 166, 79, 8, 96, 13, 27, 20, 150, 182, 88, 12]. In fact, these systems have garnered tremendous theoretical attention for many years [172, 38, 57, 59, 58, 11, 124, 111, 107, 105, 153, 70, 106, 177] given the emergent new behaviors, critical phenomena, and quantum phases of matter that they can feature. For example, dynamical phase transitions in collective models, with steady state entanglement in the form of spin-squeezing [93, 175] near critical points of a transition, have found a broad range of applications in metrology [153, 70, 106, 177, 11]. Collective spin systems are also convenient from a modeling perspective as translational symmetry between atoms can be exploited to drastically reduce the dimension of the Hilbert space [46, 10], while mean-field and cumulant approximations prove accurate over a broad range of parameter space [11].

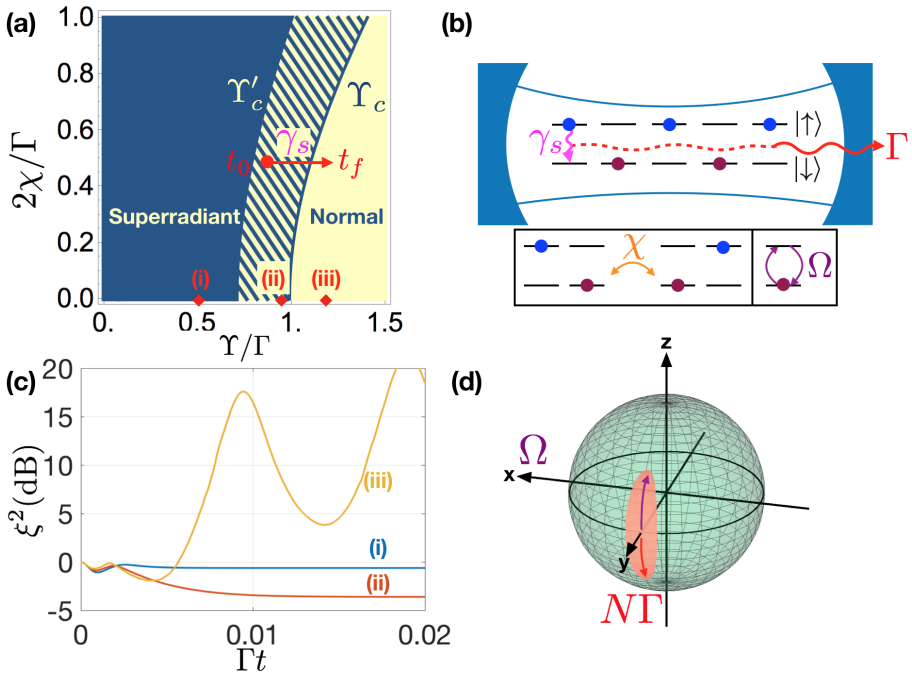


Figure 5.1: (a) Steady-state phase diagram of an ensemble of N spin-1/2 particles subjected to a coherent drive with Rabi frequency $\Omega = N\Upsilon/2$, collective emission at rate Γ, collective spin-exchange interactions χ, and single-particle relaxation at rate γ_s. This system can be engineered using an optical cavity (b) or trapped ion arrays. The spin-1/2 is encoded in a pair of electronic states, while the collective dissipation and global spin-spin interactions are mediated by spin-1/2s exchanging virtual bosons through a common mode. In the absence of single-particle relaxation, the system undergoes a non-equilibrium phase transition (superradiant to normal) signaled by a change in the total steady-state atomic inversion, which serves as an order parameter. Approaching the transition point from the superradiant phase [points (i) and (ii)], the coherent drive (in the $\hat{x}$-direction) and collective emission combine to generate spin-squeezing along $\hat{x}$, as shown in (c) and (d). In the normal phase no squeezing is observed [point (iii)]. For all three graphs in panel (c), $N = 2000$ and all spins are initially polarized along $-\hat{x}$. Panel (d) explicitly displays the Bloch sphere overlaid with a squeezed collective spin distribution of the steady-state (pink). Note that this is for illustrative purposes, and that the actual position and orientation of the squeezing can vary with parameters. Introducing finite γ_s allows the system to dynamically traverse the phase-diagram [red arrow in (a)] and enhances the achievable spin-squeezing in the striped region of panel (a).

However, a drawback is that such states can take an extremely long time to develop [11, 38]. While much of the previous work has focused on the case of negligible single particle effects, the general prediction has been that single particle decoherence would prevent any useful entanglement from developing before the steady-state could be reached making it inaccessible in experiment.

In this work we demonstrate that, for a system of coherently driven atoms in a cavity undergoing both collective and single particle dissipation, the interplay between single particle and collective effects accelerates the rate of entanglement generation and improves the overall amount of transient spin-squeezing, yielding states with more potential metrological utility.

The mechanism driving this phenomenon is the destruction of collective coherence due to single-particle decoherence, which dynamically reduces the effective particle number, allowing the system to dynamically traverse the corresponding non-equilibrium phase-diagram [see Fig. 5.1(a)], and in turn access regimes that may display large transient squeezing. While our analysis of this phenomenon is framed from a cavity-QED perspective, we note that similar conclusions can be drawn in more general models including arrays of trapped ions [155, 153, 20, 150] and superconducting qubits [64, 121].

5.2 Model and definitions

We consider a collection of N atoms in a standing wave optical lattice supported by an optical cavity, illustrated in Fig. 5.1(b). A single common mode of the cavity is coupled to the optical transition of the atoms, which encodes a spin-1/2 degree of freedom in the excited $|\uparrow\rangle$ and ground $|\downarrow\rangle$ states. The cavity is driven by an external coherent field which is tuned resonant to the atomic transition, and upon adiabatic elimination of the rapidly evolving intracavity field [21] the dynamics of the atomic degree of freedom can be described by a master equation for the atomic density operator $\hat{\rho}$ [11]

$$\frac{\partial \hat{\rho}}{\partial t} = -i[\hat{H}, \hat{\rho}] + L_c[\hat{\rho}] + L_s[\hat{\rho}], \tag{5.1}$$

$$\hat{H} = \chi \hat{J}^+ \hat{J}^- + \Omega \hat{J}^x, \tag{5.2}$$

where $\hat{J}_\alpha = \sum_{i=1}^{N} \frac{1}{2}\hat{\sigma}_i^\alpha$ for $\alpha = x, y, z$, $\hat{\sigma}_i^\alpha$ are the Pauli operators on the Hilbert space for spin $i = 1, 2, ..., N$, and $\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y$ are collective raising and lowering operators. The first term in $\hat{H}$ corresponds to a collective exchange interaction realized by detuning the cavity from the

atomic transition and charactized by χ, and the second a coherent drive characterized by Ω. The dissipative part of (5.1) includes a collective decay term given by $L_c[\hat{\rho}] = \Gamma L(\hat{J}^-)[\hat{\rho}]$ that arises due to leakage of the intracavity field via the mirrors. Both χ and Γ are proportional to the single-particle co-operativity of the cavity (see appendix B). We also include a single particle relaxation due to spontaneous emission or other systematic effects such as light scattering [84] given by $L_s[\hat{\rho}] = \gamma_s \sum_{i=1}^{N} L(\hat{\sigma}_i^-)[\hat{\rho}]$. Other types of single-particle decoherence (e.g., dephasing) would result in similar behavior, though we only consider single-particle relaxation here. The Lindblad superoperator is given by $L(\hat{O})[\hat{\rho}] = \{\hat{O}^\dagger \hat{O}, \hat{\rho}\}/2 - \hat{O}\hat{\rho}\hat{O}^\dagger$ for a given operator $\hat{O}$.

Note that there is permutational symmetry between different atoms; however, the spontaneous emission term does not commute with $\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2$ and, therefore, dynamics are not restricted to a single eigenspace of $\hat{J}^2$ when $\gamma_s \neq 0$. The translational invariance allows us to solve for dynamics in the basis corresponding to irreducible representations of the rotation group for the spin ensemble [46, 10], where exact dynamics can be efficiently simulated using Monte Carlo wave function methods [184, 123, 135]. In simulations, we restrict our initial states to be coherent spin states [142], which are eigenstates of $\hat{J}^2$ that exhibit the necessary particle symmetry.

5.3 The collective system

Before discussing the effects of single particle relaxation, we review the behavior of the collective system where $\gamma_s = 0$. As the dynamics is entirely described by collective operators, then the total spin operator $\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2$ is conserved during evolution. For this system, since our initial state will be a coherent spin state, which is an eigenstate of $\hat{J}^2$ with eigenvalue $J(J+1)$ with $J = N/2$, an analytic solution is available for the steady-state density operator $\hat{\rho}_{ss}$ [11, 38]. It has been shown [11, 38] from the resulting collective spin operator expectations that there exists a second-order non-equilibrium phase transition in the steady state as a function of $\Upsilon \equiv \left(2\Omega/N\right)$ and for large N, described by an abrupt change in behavior of the order parameter $\langle \hat{J}_z \rangle$ at a critical value given by

$$\Upsilon_c = \sqrt{\Gamma^2 + 4\chi^2}. \tag{5.3}$$

The critical point separates a superradiant phase for $\Upsilon < \Upsilon_c$ characterized by non-zero inversion $|\langle \hat{J}_z \rangle| > 0$, and a normal phase for $\Upsilon > \Upsilon_c$ with zero inversion $\langle \hat{J}_z \rangle = 0$ [see Fig. 5.1(a)]. The critical point Υ_c also delineates two regions with different steady-state behavior. In particular, there is a change in the steady-state value of the spin-squeezing of the system, defined as [175]

$$\xi^2 = \min_{\mathbf{n}_\perp} \frac{N \Delta \hat{J}^2_{\mathbf{n}_\perp}}{|\langle \hat{\mathbf{J}} \rangle|^2}, \tag{5.4}$$

where $\langle \hat{\mathbf{J}} \rangle = (\langle \hat{J}_x \rangle, \langle \hat{J}_y \rangle, \langle \hat{J}_z \rangle)$ is the collective Bloch vector, $\mathbf{n}_\perp$ is a unit vector orthogonal to $\langle \hat{\mathbf{J}} \rangle$, and $\Delta \hat{J}^2_{\mathbf{n}_\perp} = \langle (\hat{\mathbf{J}} \cdot \mathbf{n}_\perp)^2 \rangle - \langle \hat{\mathbf{J}} \cdot \mathbf{n}_\perp \rangle^2$ is the variance of the collective spin operator in the direction of $\mathbf{n}_\perp$. A state with a value of $\xi^2 < 1$ indicates the presence of entanglement between the atoms of the system, and is referred to as a spin squeezed state [158]. It quantifies the utility of the state for quantum sensing applications [93].

Figure 5.1(c) illustrates that the threshold Υ_c marks the boundary between a region in which the steady-state is squeezed and one in which it is not, by plotting the spin-squeezing versus time on each side of the threshold. In graphs (i) and (ii), on the lower side of the boundary, we see a squeezed steady-state, with an increase in steady-state squeezing as Υ approaches the threshold at Υ_c from below. In graph (iii), where $\Upsilon > \Upsilon_c$, we see that there is no longer a squeezed steady-state. The $\Upsilon < \Upsilon_c$ side of the critical drive frequency is referred to as the superradiant phase, and the $\Upsilon > \Upsilon_c$ side is the normal phase. It should be noted that as Υ approaches the threshold Υ_c from below, a careful selection of initial conditions becomes necessary to reach the squeezed steady-state quickly, and to avoid an oscillatory phase known to exist near the critical point when $|\chi| > 0$ [11]. It is also important to note that the squeezing is predominantly in the azimuthal direction for small $\chi/\Gamma < 1$ (see Fig. 5.1(d) and appendix B).

5.4 Effects of single particle relaxation

When $\gamma_s \neq 0$, $\hat{J}^2$ symmetry is broken and the dynamics are no longer confined to a single manifold of total spin. This means the dynamics are free to explore a larger portion of the full Hilbert space of 2^N states, compared to the limited $N + 1$ states of the collective model. As a

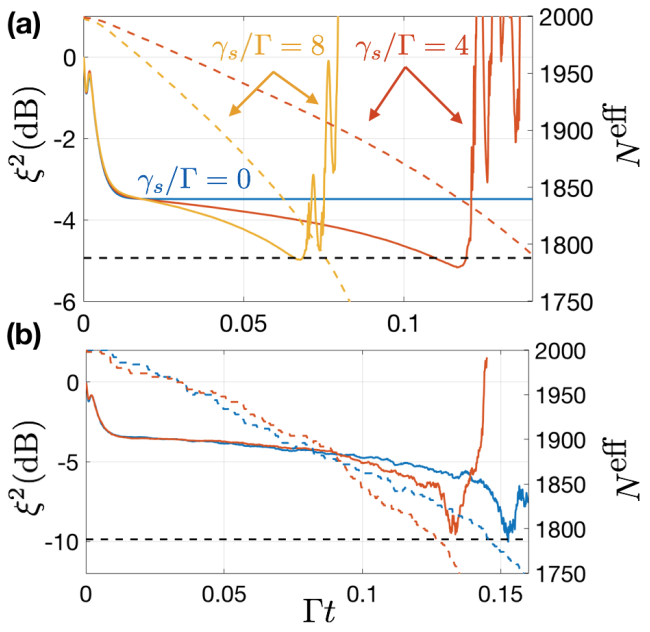


Figure 5.2: (a) Squeezing versus time for $N = 2000$, $\chi/\Gamma = 0$, $\Upsilon/\Upsilon_c = 0.9$, and a range of γ_s/Γ. Solid lines indicate squeezing $\xi^2(t)$ and dashed lines the corresponding time-dependent effective system size $N^{\rm eff}(t)$ (matching colors). The horizontal black line corresponds to the critical effective particle number $N_c = 2\Omega/\Upsilon_c$ for which the transition between superradiant and normal phases occurs. (b) Squeezing $\xi^2(t)$ (solid) and effective system size $N^{\rm eff}(t)$ (dashed) computed from two individual trajectories of the numerical method with $\gamma_s/\Gamma = 4$. For each trajectory, $N^{\rm eff}(t)$ crosses the horizontal line for N_c near the point where its corresponding $\xi^2(t)$ reaches a minimum. In each panel, all spins are initially polarized along $-\hat{x}$.

result of this increased complexity, an analytic formula for the steady-state is not available in this case. However, a mean-field analysis can give insight into the steady-state phase diagram of the system, including the position of critical transitions and transient behavior, and efficient numerical simulations [184, 123, 135] of the master equation giving the full dynamics of the system can confirm these properties, and others such as the presence of spin squeezing.

In the mean-field approximation, we derive equations of motion for the expectations of individual particle Pauli operators, $\partial_t\langle\sigma_i^\alpha(t)\rangle = \mathrm{Tr}[\sigma_i^\alpha\partial_t\hat{\rho}]$, from the master equation [Eq. (5.1)]. Due to the permutational symmetry of the master equation, and assuming the same symmetry applies to the initial state, these equations will be identical for all particles.

Under the mean-field assumption $\rho = \bigotimes_i \rho_i$, where each ρ_i is a single particle density, the

equations can be closed as second-order expectations can be factored as $\langle \sigma_i^\alpha(t)\sigma_j^\beta(t)\rangle = \langle \sigma_i^\alpha(t)\rangle\langle \sigma_j^\beta(t)\rangle$ when $i \neq j$. When $i = j$, second order expectations can be handled in one of two ways. The first approach would be to use commutation relations to resolve the product of Pauli operators into a single operator before taking the expected value, i.e. $\langle \hat{\sigma}_i^\alpha \hat{\sigma}_i^\beta\rangle \to \delta_{\alpha,\beta} + i\epsilon_{\alpha\beta\gamma}\langle \hat{\sigma}_i^\gamma\rangle$. The second approach is to factor in the same way as for unlike particles, i.e. $\langle \hat{\sigma}_i^\alpha \hat{\sigma}_i^\beta\rangle \to \langle \hat{\sigma}_i^\alpha\rangle\langle \hat{\sigma}_i^\beta\rangle$. Previous work [38] has shown that the former approach aligns with the exact solution when $\hat{J}^2$ is not conserved, and that the latter is accurate otherwise. Since we are interested in the case where $\gamma_s > 0$ and total spin is not conserved, we factor only unlike particles. Dropping the subscripts due to particle symmetry and defining $\langle \sigma^+\rangle \equiv re^{i\phi}$, and $\langle \sigma_z\rangle \equiv z$, we arrive at the mean-field equations

$$\dot{r} = -\frac{\Gamma + \gamma_s}{2} r + \frac{\Gamma}{2}(N-1) z\, r - \frac{\Omega}{2} z \sin\phi \tag{5.5}$$

$$\dot{\phi} = -\chi(N-1)z - \frac{\Omega}{2r} z\cos\phi + \chi \tag{5.6}$$

$$\dot{z} = -2\Gamma(N-1)r^2 - (\Gamma + \gamma_s)(1+z) + 2\Omega\, r\sin\phi. \tag{5.7}$$

We determine the steady-state at the mean-field level by setting the LHS of Eqs. (5.5)-(5.7) to zero and solving for (r, ϕ, z). We begin by deriving a steady-state expression for z in terms of the other variables. From the r equation, we get

$$0 = -\frac{\Gamma + \gamma_s}{2} r + \frac{\Gamma(N-1)}{2} z\, r - \frac{\Omega}{2} z \sin\phi$$

which implies that

$$-2(N-1)\Gamma r^2 + 2\Omega\, r\sin\phi = -2(\Gamma + \gamma_s)\frac{r^2}{z},$$

where we have assumed $z \neq 0$. Plugging this into the z equation gives

$$\begin{aligned} 0 =& -2(\Gamma + \gamma_s)\frac{r^2}{z} - (\Gamma + \gamma_s)(1+z) \\ \Rightarrow 0 =& z^2 + z + 2r^2, \end{aligned}$$

and subsequently,

$$z = -\frac{1}{2} \pm \frac{1}{2}\sqrt{1 - 8r^2}. \tag{5.8}$$

Here, we identify that there are two steady-state values of z. A linear stability analysis reveals that the branch with larger absolute value is stable, while the smaller is unstable. We also see that we can expect a bifurcation when r passes through $1/\sqrt{8}$ from below.

We now turn our attention to the steady-state value of r. From the z equation, we see that

$$4r^2\Omega^2 \sin^2\phi = [2(N-1)\Gamma\, r^2 + (\Gamma+\gamma_s)(1+z)]^2, \tag{5.9}$$

and from the ϕ equation

$$\begin{aligned} 4r^2\Omega^2 \cos^2\phi =& 4r^2(1-\sin^2\phi) \\ =& \left[-\chi(N-1)r^2 + \chi\frac{r^2}{z}\right]^2. \end{aligned} \tag{5.10}$$

Combining Eqs. (5.9) and (5.10) we get

$$\begin{aligned} 4r^2\Omega^2 - \left[2(N-1)\Gamma\, r^2 + (\Gamma+\gamma_s)(1+z)\right]^2 \\ = \left[-\chi(N-1)r^2 + \chi\frac{r^2}{z}\right]^2, \end{aligned} \tag{5.11}$$

where z is given in terms of r by Eq. (5.8). Note that if we now restrict our attention to leading order in N, the above implies

$$\begin{aligned} 4r^2\Omega^2 - 4N^2\Gamma^2 r^4 &= N^2\chi^2 r^4 \\ \Rightarrow r^2 = \frac{\Omega^2}{N^2(\Gamma^2+4\chi^2)} &= \frac{\Upsilon^2}{4\Upsilon_c^2}, \end{aligned} \tag{5.12}$$

a familiar result that shows that we expect an increasing value of r as we increase the Rabi frequency of the drive Ω. Recall, however, that Eq. (5.8) predicts a phase transition when $r = 1/\sqrt{8}$. This corresponds to

$$\Upsilon = \frac{\Upsilon_c}{\sqrt{2}} \equiv \Upsilon_c'. \tag{5.13}$$

Beyond this transition, equation (5.8) predicts complex z, which is not a valid steady-state. Recall, however, that equation (5.8) assumed $z \neq 0$. Numerical simulations confirm that for $\Upsilon > \Upsilon_c'$ the steady-state value of z is, in fact, zero. This aligns with the transition between superradiant and normal phases that we were expecting. We will see that this transition exists beyond the mean-field level when we look at numerical solutions to the cumulant expansion equations below.

The mean-field analysis indicates that many of the qualitative features of the collective physics, particularly the steady-state behavior, remain when single-particle relaxation is included. Specifically, for $\gamma_s \neq 0$ there is a critical point $\Upsilon_c' \equiv \Upsilon_c/\sqrt{2}$ delineating superradiant and normal phases characterized by the long-time limit of collective observables. Moreover, numerical simulations of the full quantum dynamics reveal that Υ_c' also marks the boundary between a squeezed steady state in the superradiant phase and the absence of long-time squeezing in the normal phase (see appendix B). This transition is illustrated in Fig. 5.1(a). In the case where $\hat{J}^2$ is conserved, corresponding to the collective system where $\gamma_s = 0$, this transition occurs at Υ_c, as expected from the exact solution $\hat{\rho}_{ss}$ [11].

5.5 Enhanced squeezing

We now consider the effects of adding spontaneous emission to the model, i. e. the case where $\gamma_s > 0$ in Eq. (5.1). In particular, we are interested in how the phase diagram changes and to what extent spin-squeezing can be achieved. While it might be reasonable to expect that single particle dissipation prevents the development of any useful squeezing [154], we will see that, in fact, the phase diagram is split into three regions: one with a minimally squeezed steady-state for $\Upsilon < \Upsilon_c'$, one where squeezing does not develop at all for $\Upsilon > \Upsilon_c$, and a region in between where an amount of squeezing develops in the late transient dynamics that exceeds what is seen in the steady-state when $\gamma_s = 0$. This corresponds to the striped region in Fig. 5.1(a). While the transient dynamics are qualitatively different in these three regions, the steady-state phase diagram is split in two, just as in the collective case, with a boundary at Υ_c' rather than Υ_c between a superradiant phase with a squeezed steady-state and a normal phase without one.

Figure 5.2(a) illustrates the spin squeezing dynamics in the region where $\Upsilon_c' < \Upsilon < \Upsilon_c$ for several values of γ_s/Γ and in the absence of elastic interactions ($\chi/\Gamma = 0$). We observe that, as expected, squeezing does not persist in a steady-state as it does in the purely collective case ($\gamma_s/\Gamma = 0$). However, we see that squeezing develops and persists at timescales $\sim 1/\gamma_s$, and is being increased in the presence of spontaneous emission over what would be seen in the purely collective

steady-state for $\chi = 0$ at the same values of Υ/Υ_c. It can also be seen that the rate at which it develops scales with γ_s/Γ, while the amount of achievable squeezing does not change (i.e., γ_s/Γ controls the rate at which we dynamically traverse the related collective steady-state phase diagram – which ultimately sets the bound on the achievable squeezing). We now turn to the question of how this improved spin squeezing develops.

5.6 Squeezing mechanism

To simulate the exact dynamics when $\gamma_s > 0$, we have used the Monte Carlo Wave Function (MCWF) method [184, 123, 135]. This method unravels the density matrix into an ensemble of pure state wave functions that evolve independently of one another in time, where dissipation is handled by random jumps. The full time evolution of one member of this ensemble is referred to as a trajectory. The time evolution of the density matrix is recovered by taking the average of the pure state density matrices at each point in time, resulting in the mixed state solution to the master equation [Eq. (5.1)]. The advantages of this method are three-fold. First, as shown in [184], each trajectory will lie within a single eigenspace of total spin at any given time, reducing the dimensionality to $\mathcal{O}(N)$ from $\mathcal{O}(N^2)$ for the full particle symmetric Dicke basis. Second, the trajectories evolve in time independently of one another, allowing for the parallel simulation of different trajectories. Finally, analyzing the time evolution of individual trajectories can provide insight that is not altogether obvious from the evolution of the density matrix resulting from the ensemble averages.

Figure 5.2(b) plots squeezing versus time for a number of trajectories where $\Upsilon_c' < \Upsilon < \Upsilon_c$ and $\gamma_s > 0$. We notice that the dynamics of the trajectories are qualitatively similar to the ensemble average, with an increase in squeezing that is abruptly lost. Both panels of the figure also show the time evolution of the effective system size, motivated as follows. When $\gamma_s = 0$, the value of $\langle \hat{J}^2 \rangle = N/2(N/2+1)$ is conserved. When $\gamma_s > 0$, this value will decay, and we can define the

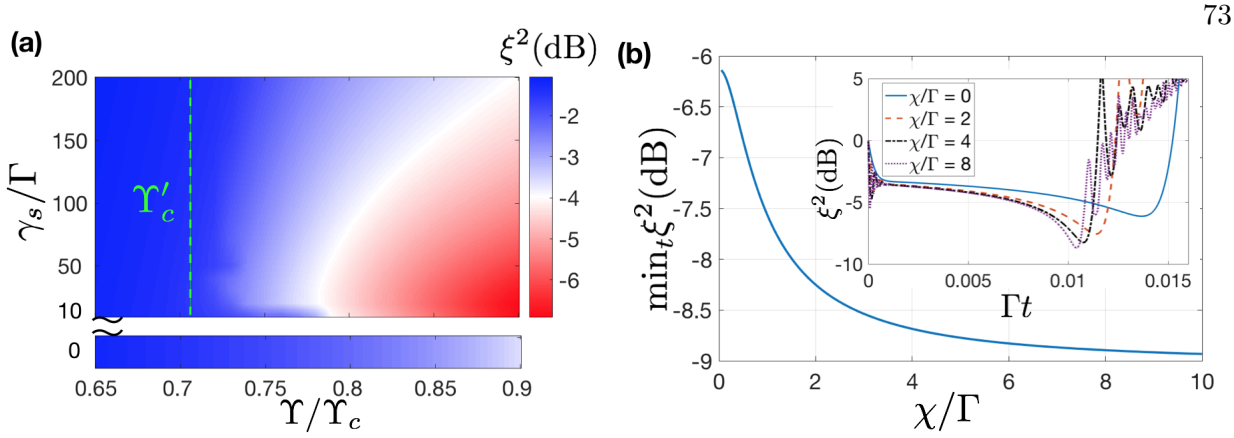


Figure 5.3: (a) Minimum transient spin-squeezing (see text for clarification) as a function of normalized drive amplitude Υ/Υ_c and relaxation rate γ_s/Γ with $\chi/\Gamma = 0$. The strip below the main panel shows a magnified view of the $\gamma_s/\Gamma = 0$ result for comparison. Note the break in the vertical axis, which is required since attainable simulation times cannot capture the squeezing behavior occurring on timescales of $1/\gamma_s$ when this value is very large. (b) Minimum transient spin-squeezing as a function of the interaction strength χ/Γ with fixed $\Upsilon/\Upsilon_c = 0.9$ and $\gamma_s/\Gamma = 50$. Inset: Squeezing versus time for a selection of values of χ/Γ and same Υ, γ_s as main panel. For (a) and (b) we compute the dynamics using a truncated cumulant expansion (see appendix B) and $N = 10^4$. Initial conditions in (a) are the coherent spin state (CSS) in the $-\hat{x}$-direction, and in (b) are taken to be the CSS in the direction of the mean-field steady-state (for each χ/Γ and Υ) when $\gamma_s/\Gamma = 0$ to account for the rotations that result from different values of χ/Γ.

effective system size by solving for N

$$N^{\text{eff}}(t) \equiv 2\sqrt{(1/4) + \langle \hat{J}^2 \rangle(t)} - 1. \tag{5.14}$$

This decrease in effective particle number results in a corresponding increase in the effective drive $\Upsilon^{\text{eff}}(t) \equiv 2\Omega/N^{\text{eff}}(t)$. The behavior is similar to the collective case where the particle number is fixed and Υ is tuned toward the threshold value, as we saw in Fig. 5.1(c), except here one could view it as the threshold moving while the drive frequency Ω is held fixed. As the distance between $\Upsilon^{\text{eff}}(t)$ and Υ_c decreases, the squeezing increases. Once the effective drive frequency exceeds the threshold Υ_c, it is as though the system transitions from the superradiant to the normal phase and the squeezing is lost.

This process can clearly be seen in the trajectories in Fig. 5.2(b). The critical value of N^{eff}, where $\Upsilon^{\text{eff}}(t) = \Upsilon_c$, is shown as a black dashed line. For each trajectory, it can be seen that the

point in time where this threshold is crossed corresponds to the point where squeezing is abruptly lost. In the aggregate across all trajectories, this results in a mixed state that will exhibit an improvement in squeezing until a sufficient number of trajectories have crossed into the normal phase, at which point squeezing will be lost gradually.

Returning to Fig. 5.2(a), we now see that the improvement in squeezing over the collective steady-state value is a result of the fact that the drive frequency is being pushed closer to the effective threshold, where squeezing is greater, before crossing the phase boundary. Finally, it should be mentioned that having a sufficiently strong drive such that $\Upsilon > \Upsilon_c'$ is critical to seeing this behavior since, on the other side of the phase transition, a stable steady-state is reached before N^{eff} can drop below the critical threshold, and the steady-state level of squeezing is small.

Figure 5.3(a) plots the minimum squeezing obtained in the transient dynamics for each value of the normalized drive amplitude Υ and single-particle relaxation rate γ_s/Γ. Due to the number of simulations required, results in this figure are computed using a cumulant expansion approximation that shows close agreement with the exact numerical solution (see appendix B). The introduction of finite $\gamma_s/\Gamma \neq 0$ clearly improves the attainable squeezing within the region of $\Upsilon_c' < \Upsilon < \Upsilon_c$ relative to the collective case ($\gamma_s/\Gamma = 0$, shown in the lower strip). In the case where $\gamma_s/\Gamma = 0$, there is no qualitative change in the minimum squeezing when crossing the threshold at Υ_c', as expected from previous work. The improvement in squeezing occurs for even relatively small values of γ_s, although as γ_s is increased the best transient squeezing, attained for Υ approaching Υ_c, gradually degrades. On the other hand, for $\Upsilon < \Upsilon_c'$, a stable steady-state is quickly reached and squeezing is not enhanced by introducing single-particle decoherence. Crossing the threshold eliminates this stable equilibrium, allowing for the transient enhancement, at the cost of losing the squeezed steady-state by passing into the normal phase. As the drive approaches Υ_c, critical slowing down becomes prominent, preventing the squeezing from developing before dissipative effects make it impossible.

While our qualitative understanding of the mechanism driving squeezing has so far not included discussion of the collective exchange interactions, the achievable squeezing does quantitatively depend on χ for $\Upsilon < \Upsilon_c$. This is demonstrated in Fig. 5.3(b), where we plot the minimum

transient squeezing as a function of χ/Γ for $\gamma_s/\Gamma = 50$. It is apparent that increasing the interaction strength χ leads to an appreciable improvement in the optimal squeezing, particularly in the region $0 < \chi/\Gamma \lesssim 2$. However, the inset of Fig. 5.3(b) indicates that an increased interaction strength does not significantly change the qualitative dynamics of the squeezing, beyond the generation of an earlier transient (absent in the $\chi = 0$ case). This earlier transient that appears at finite χ might be useful for some platforms, but on the other hand for purely metrological applications might be not as practical in cases such as the cavity setup discussed below (see appendix B). There, although feasible, technical challenges come up when quenching χ sufficiently fast to take advantage of the earlier transient squeezing.

5.7 Experimental realization and outlook

The spin model we have discussed could be realized by coupling an optical cavity to the narrow linewidth optical clock transitions available in alkaline earth atoms [129, 125]. We require that $\kappa \gg g\sqrt{N}$ and $\kappa \gg \gamma_s$ (bad cavity limit) with $2g$ the single photon Rabi frequency and κ the cavity linewidth, to ensure that the intracavity field can be adiabatically eliminated and thus realize the desired spin model [Eqs. (1) and (2)]. In this limit, spin-spin interactions can be engineered by detuning the cavity from the atomic transition by Δ_c which leads to a tunable interaction strength $\chi = 4g^2\Delta_c/(4\Delta_c^2+\kappa^2)$. Similarly, the collective dissipation arises due to photon leakage and is characterized by $\Gamma = 4g^2\kappa/(4\Delta_c^2 + \kappa^2)$ [129, 130]. To ensure that decoherence is not too large such that it eliminates any possibility of squeezing, we need to operate in the limit of a large (effective) collective cooperativity $\gamma_s \lesssim N\Gamma$. This condition together with those for κ in the preceding paragraph imply we should work in the hierarchy of energy scales $\gamma_s \lesssim N\Gamma \ll g\sqrt{N} \ll \kappa$ to generate spin squeezing in the cavity platform. The possibility to operate in this regime has previously been demonstrated using both the ^{1}S$_0$-^{3}P$_0$ transition in ^{87}Sr [129, 130] and ^{1}S$_0$-^{3}P$_1$ transition in ^{88}Sr [125]. The former has a natural linewidth of $\gamma \approx 2\pi \times 1$ mHz and $2g = 2\pi \times 8$ Hz [129]. State-of-the-art AMO experiments have demonstrated coherence of the ^{1}S$_0$-^{3}P$_0$ transition of up to $1/\gamma_s \approx 10$ s [84] which corresponds to $\gamma_s/\Gamma \approx 200$ for $\Delta_c = \kappa = 2\pi \times 150$ kHz and thus

$\chi/\Gamma \approx 1$. For $N \sim 10^4$ atoms, dissipatively enhanced squeezing of $\xi^2 \approx 9$ dB is then in principle achievable on timescales $t \sim 2$ s.

A similar implementation can also be realized in trapped ion arrays, where a pseudospin-1/2 is encoded in the hyperfine states of the ion. As carefully shown in Ref. [155], it is possible to engineer in a Penning trap the same collective dissipation that is responsible for superradiance in cavity QED systems. This is achieved by loading two-types of ions (τ and σ) into a shared trap. The two species could be, for example, two different elements, or isotopes of the same element. The τ ions are used to sympathetically cool the normal modes of vibration of the system of ions and generate an effective phonon loss, analogous to κ in the cavity platform. Further, by Doppler cooling the τ ions it is possible to introduce couplings between the normal modes, resulting in a new dressed set of damped normal modes. The σ ions then serve as the effective spins that are squeezed through interactions mediated by the damped phonon modes. The σ ion-phonon coupling can be engineered using an optical dipole force generated via pairs of Raman beams. The detuning of the Raman beams can be set such that predominantly the center of mass (CM) mode is excited (i.e., other mode remain off-resonant), such that the CM mode plays the role of the common cavity mode which mediates both elastic[20] and inelastic collective spin interactions between the ions. Additionally, resonant microwaves can be used to coherently directly drive the spins [150]. As analyzed in detail in Ref. [155], using $^{24}Mg^+$ ions as the τ ions and $^{25}Mg^+$ as the σ ions, it should be possible to achieve an effective $\Gamma \sim 2\pi \times$ Hz in a system of the order of $N = 124$ σ-ions. In this implementation the average single-particle decoherence generated by the Raman beams including effective spontaneous emission, absorption and dephasing is of the order of $\bar{\gamma}_s \sim 2\pi \times$ Hz. In this set-up therefore it should be possible to operate in the regime where $N\Gamma/\bar{\gamma}_s \sim 100$ and reach the conditions required for robust spin squeezing generation.

In summary, we have identified an intriguing and experimentally relevant situation where spin-squeezing can co-exist with relatively large single-particle decoherence as long as collective decoherence remains the dominant dissipative process. We expect our results to have immediate applications for quantum metrology, specifically in the generation of squeezing on long-lived op-

tical transitions for next-generation optical atomic clocks, whilst also being relevant for quantum simulation.

5.8 Acknowledgements

We are grateful for feedback from M. A. Norcia and J. T. Young on the manuscript. This work is supported by the AFOSR grant FA9550-18-1-0319 and its MURI Initiative, by the DARPA and ARO grant W911NF-16-1-0576, the ARO single investigator award W911NF-19-1-0210, the NSF PHY1820885, NSF JILA-PFC PHY-1734006 grants, and by NIST.

Chapter 6

Parameter Estimation via Machine Learning

This chapter is a draft of a manuscript to be submitted for publication with authors K. Tucker, A. K. Rege, C. Smith, C. Monteleoni, and T. Albash

Prologue

In this chapter, we broaden our attention beyond the classical simulation of quantum systems to include estimating system parameters from data. This is done through a novel approach that combines machine learning with traditional classical simulations like those used in previous chapters. Specifically, we propose a machine learning model based on a denoising autoencoder to learn a map from weak measurement input to master equation parameter estimates, then back again to measurement records in the ensemble average. The physical parameters have thus taken on the role of latent variables, with physical interpretability enforced by using a classical simulator in the decoder to map from parameters back to measurement records. This simulator is augmented with a recurrent neural network to account for dynamics in observed data beyond the ability of the simulator to model. These limitations could be due to simplifying assumptions on the states, such as the entanglement limitations entailed in mean field or cumulant expansions, or due to missing or incorrect terms in the physical model itself.

In this way, the ideas in this chapter represent a fusion of traditional classical simulation methods and concepts from machine learning intended to correct for the limitations of these methods. In the field of machine learning, this approach is broadly referred to as discrepancy modeling.

6.1 Introduction

As the scale, complexity, and availability of quantum devices is expected to grow considerably over the coming decades [140], the ability to accurately characterize device parameters has become an urgent need. For example, while great progress has been made in quantum error correction, non-local errors such as crosstalk threaten to derail these approaches in larger systems [152, 162]. Detection of these unwanted effects is a critical first step in mitigating them, with efforts to do so complicated by the exponential scaling of the quantum state with system size and the challenges of modeling increasingly complex systems.

Recent advances in the application of machine learning (ML) tools to quantum systems have shown promise in overcoming these difficulties. Scalable ML models have been used to represent quantum states for state tomography [37, 36, 42, 145, 41, 161] and evolution [161, 9], for learning unknown dynamics [33, 44, 115], as well as for learning device characterization from measurements [68, 65, 97, 5, 43]. Increasingly, a priori knowledge of a physical system has been combined with ML models to improve accuracy and interpretability [143]. The model-free nature of many ML solutions has made them scalable and robust to many of the common pitfalls of physical models, such as non-Markovian dynamics [68, 65, 97, 9]. There are trade-offs, however, as more abstract models are generally less interpretable, which limits the physical insights that can be gleaned from them. Finding a balance between these two competing properties, representability and interpretability, thus becomes an important challenge.

In this work, we build on recent advances in device characterization applying ML models to the continuous measurement of qubits [65, 97, 68], with the goal of making this characterization more scalable and robust to unanticipated dynamics, and to apply it to the specific problem of detecting two qubit interactions, one possible manifestation of crosstalk in multi-qubit systems. In contrast with the parameter estimation approach of [68], where a stochastic master equation is used as a trainable model, our ML model learns a direct map from continuous measurement input to the system parameters of interest. This approach has two advantages: first, in the case

where device parameters corresponding to measurement records are known, the model is completely independent of any state representation, freeing it from the curse of dimensionality that comes from the exponential scaling of the state size. The requirement is rather that the measurements need to provide enough information to estimate the parameters of interest, not the complete quantum state, which in many cases could be less onerous. Second, once the model is trained, parameter estimation can be performed quickly even for systems not seen during training, and can benefit from noise correction capabilities learned from many training examples, thus requiring fewer measurements for estimation.

In the event that device parameters are not provided along with measurement records, an unsupervised approach comparing input to a measurement record reconstructed from parameter estimates can be used. This requires a map from parameter estimates to measurement output, which is accomplished by adding a layer combining an integrator of the physical model with a recurrent neural network (RNN) to provide a model-free correction at every time step to account for unanticipated effects, as long as they are consistent and small relative to the dynamics driven by the parameters of interest. This employs the capabilities of neural ODE found in other works [48, 74, 9], but enhanced to provide a completely model-free correction not bound by assumptions of linearity or Markovianity with a projection step to return the corrected state to the manifold of physical density operators [157]. The output of the model is then the estimated solution to the unconditioned master equation with learned corrections for effects beyond the master equation, which can be compared to the measurement records to update model parameters. While this case is not completely model-free, it also does not make rigid assumptions about the dynamics and remains capable of accurate parameter estimation in the presence of unanticipated effects that would otherwise severely impact accuracy. The ML model is not burdened with learning all of the physics, it just has to correct for lower order effects the model may have missed, an approach known as discrepancy modeling [89, 53].

The full model can be viewed as a denoising autoencoder [71] where the physical parameters being estimated take on the role of the latent space, with interpretability enforced by the presence

of the physical model in the decoder. The encoder is the map from measurement records to parameters, and is the desired product of the training to be used for fast and accurate parameter estimation. While the full unsupervised model clearly has applications for parameter estimation in cases where nothing is known about the parameters of interest, the supervised approach using just the encoder could still find applications for systems where parameters are known at the time training data is generated, but a prediction routine is still required in other circumstances, as would be the case for detecting drift away from device calibration over time.

6.2 Physical System

We consider two qubits with fixed position in a microwave cavity as illustrated in Fig. 6.1(a)(i). A single common mode of the cavity is coupled to the computational degree of freedom of the qubits, and they are coherently driven on resonance with a Rabi drive of frequency Ω. A two-qubit interaction term is present with magnitude ϵ. A weak measurement tone [86, 29] is applied to the cavity to probe the qubit state in one of the $\{X, Y, Z\}$ directions for each qubit, with a measurement back-action dephasing rate of κ. Upon adiabatic elimination of the cavity mode [21], which we assume has dynamics evolving at a rate much faster than time scales relevant to the qubits, we can describe the system with the stochastic master equation (SME) [16, 66]

$$\begin{aligned} d\rho &= -i[H,\rho]dt \\ &\quad + \sum_{i=1}^{2} \mathcal{D}[L_i](\rho)dt + \sum_{i=1}^{2} \sqrt{\frac{\eta}{2}}\mathcal{H}[L_i](\rho)dW_t^{(i)}, \end{aligned} \tag{6.1}$$

$$H = \sum_{i=1}^{2} \frac{\Omega}{2} X_i + \epsilon Z_1 Z_2, \tag{6.2}$$

where $\{X_i, Y_i, Z_i\}$ is the set of Pauli operators for qubit i, $\mathcal{D}[L](\rho) = L\rho L^\dagger - \frac{1}{2}\left\{L^\dagger L\rho + \rho L^\dagger L\right\}$ is the Lindblad super-operator, $\mathcal{H}[L](\rho) = L\rho + \rho L^\dagger - \rho \mathrm{Tr}\left[\rho\left(L + L^\dagger\right)\right]$ is the measurement super-operator, $\eta = 0.1469$ is the efficiency of the measurement, and $L_i = \sqrt{\kappa} C_i$, where $C_i \in \{X_i, Y_i, Z_i\}$ is the weak measurement operator with $\kappa = 3.326$ radians/μs. Here we have adopted the normalization convention $\hbar = 1$. Parameter values are selected to be comparable to those used in [68] and to be

relevant in experiments.

The stochastic differential equations for the measurement records are given by

$$dr_i = \sqrt{\frac{\eta}{2}}\mathrm{Tr}\left[\rho\left(L_i + L_i^\dagger\right)\right]dt + dW_t^{(i)}, \tag{6.3}$$

where $r_i(t)$ is the weak measurement for qubit i and the independent $dW_t^{(i)}$ are the same appearing in (6.1). This is comparable to the system in [68], except generalized to two qubits and with the addition of the two-qubit interaction term in (6.2). Note that results will be examined for cases where the weak measurement operator L_i may be different for each qubit.

6.2.1 Selecting Initial Conditions

Increasing the strength of the weak measurement κ improves signal to noise ratio (SNR) in the measurement records as it increases the drift function in the weak measurement SDE (6.3). The trade-off is that the increased measurement backaction leads to faster decoherence in the state ρ in (6.1) setting up two competing effects on SNR in the full stochastic master equation. The effects of backaction can be mitigated for our purposes if we select the initial state to align with the direction of measurement. To see how, we consider the equation of motion in the ensemble average for the single qubit observable $C_i \in \{X_i, Y_i, Z_i\}$

$$\begin{aligned}\frac{d}{dt}\langle C_i\rangle = \frac{d}{dt}\mathrm{Tr}\left[\rho C_i\right] &= \mathrm{Tr}\left[\frac{d\rho}{dt}C_i\right] \\ &= -i\mathrm{Tr}\left[[H,\rho]C_i\right] + \kappa\sum_{j=1}^{2}\mathrm{Tr}\left[(C_j\rho C_j - \rho)C_i\right] \\ &= -i\mathrm{Tr}\left[[H,\rho]C_i\right].\end{aligned} \tag{6.4}$$

We use the Pauli relation $C_i^2 = \mathbb{1}$ and the invariance of the trace under cyclic permutation to see that the dissipative term is zero, since when $i = j$ we have

$$\begin{aligned}\mathrm{Tr}\left[(C_i\rho C_i - \rho)C_i\right] &= \mathrm{Tr}\left[C_i\rho C_i C_i\right] - \mathrm{Tr}\left[\rho C_i\right] \\ &= \langle C_i\rangle - \langle C_i\rangle = 0,\end{aligned}$$

and when $i \neq j$ we have

$$\begin{aligned}\mathrm{Tr}\left[(C_j \rho C_j - \rho) C_i\right] &= \mathrm{Tr}\left[C_i C_j C_j \rho\right] - \mathrm{Tr}\left[\rho C_i\right] \\ &= \langle C_i \rangle - \langle C_i \rangle = 0\end{aligned}$$

where we have used the fact that operators on different qubits commute. The result is an equation of motion that has no explicit dependence on measurement backaction, such that all dynamics will be proportionate to the Hamiltonian parameters Ω and ϵ. It is therefore useful to start with a maximal value in this direction, so that the Hamiltonian parameters will assert themselves to the greatest possible extent in the dynamics of the measurement observables, and therefore in the slope of the measurement record evolving according to (6.3).

6.3 Machine Learning Model

6.3.1 Model Description

The objective is to train a model that can estimate master equation parameters by observing weak measurement records. Input data is provided for training in the form of weak measurement trajectories averaged over some number of repeated measurements to reduce the effects of diffusive noise. The output of the model is either the physical parameters of interest, in the supervised case where labeled training data is available, or the predicted solution of the measurement SME in the large trajectory limit, or ensemble average, in the unsupervised case where true parameters are not available. The former configuration can be thought of as an encoder mapping noisy trajectories to estimated parameter values, while the latter configuration adds a decoder that takes the estimated parameters as input and produces a predicted solution to the unconditioned master equation. The full architecture is shown in figure 6.1(a) and can be viewed as a denoising autoencoder taking noisy input trajectories and producing output with the noise removed.

The encoder begins with an up-front average pooling layer in the time dimension to further smooth the trend component of the measurement records, followed by a long short-term memory (LSTM) layer [78] to process the sequential input data and consolidate information from multiple

qubits into a single sequence. This precedes a feed-forward neural network of dense layers ending with the parameter layer.

In the unsupervised case, interpretability of the latent variables as parameters is enforced by using them directly in an enhanced numerical ODE integrator in the decoder for the unconditioned master equation. The integrator is implemented as an RNN with a custom cell that combines a single step of Euler's method with a correction produced by a standard LSTM cell. This is shown in figure 6.1(b). The correction term is designed to compensate for calibration errors or other unanticipated dynamics in the physical model used by the numerical integrator. The design leverages the concept of neural ODE [48] or ODE-RNN [74, 147], as it uses a neural network to approximate corrections to the drift function of the unconditioned master equation $d\rho = f(t, \rho; \theta)dt$ where:

$$f\left(t, \rho; \theta\right) = f_{model}\left(t, \rho; \theta\right) + f_{LSTM}\left(t, \rho; \theta\right), \tag{6.5}$$

θ is a set of ODE parameters, and f_{LSTM} is an LSTM cell and as such is a non-linear, model free function of the time, state, and parameters.

The carry-state used in the standard LSTM architecture is passed between evaluations of f_{LSTM} at each time point, which means information from the full history of states is potentially available, allowing for the modeling of non-Markovian effects. Hermiticity of the state is enforced by restricting the free parameters in the correction to have a Hermitian form, trace is preserved by a normalization step, and positivity is preserved by performing an orthogonal projection of the time-evolved state back onto the state space according to the algorithm presented in [157] at each time step. In this respect it is analogous to the approach used by the time-dependent variational principle (TDVP) algorithm [75, 76] for modeling dynamics of matrix product states (MPS).

An additional feature of the correction is that the full drift function f is primarily relying on a known physical model f_{model} and using the ML components in f_{LSTM} only to compensate for discrepancies between this model and the experimental system, which are assumed to be small relative to the known dynamics. As such, it could be categorized along with other approaches in

ML for discrepancy modeling [89, 53] that operate under the principle that small corrections are easier to learn than the full dynamics. A similar approach is taken to modeling non-Markovian dynamics via RNN in [9], though our approach differs in that it combines discrepancy modeling with a completely general, non-linear correction to the state evolution based on neural ODE. In addition, as suggested in the outlook section of [9], we are investigating the situation where only a subset of the measurement information necessary for full state reconstruction is available, so training data volume requirements are substantially reduced.

It should be noted that while the integrator used in the decoder for this work is an Euler integrator for the unconditioned master equation, this piece is completely modular in the design and can be replaced with any integrator for simulating quantum dynamics. For example, a TDVP integrator operating on a matrix product state could be used for larger systems that are approximated well by a MPS ansatz, with the trainable decoder parameters accounting for errors introduced by this approach.

The output of the decoder is the estimated solution to the unconditioned master equation, and the loss is calculated as the mean squared error (MSE) between this estimate and provided approximations of clean measurements used as label values in the case where true parameter estimates are not available. Thus, it is desirable that these labels contain as little noise as possible, but estimation can still take place with noisy labels, as will be seen in 6.4. Clean label values are realized by averaging over many input groups sharing the same parameter values and using this same average for all of the contributing input trajectory groups during training. See 6.3.2 for more details. In this way the model can be viewed as a denoising autoencoder [71], with de-noised output provided for every noisy input example, and with a latent space corresponding to the physical parameters being estimated.

This approach to parameter estimation differs from that of [68] in two key respects. First, the physical parameters we are interested in learning are the output of a machine learning model, rather than parameters to be learned during training. This enables extremely fast parameter estimation once the model is trained even when observing systems with parameter sets not included in the

training data, and it allows us to learn a direct map from measurement input to parameters if labeled training data is available without the need for a physical model, in which case the estimation model size scales polynomially with the system size rather than exponentially. The limitation in this case becomes the amount and type of training data required, which will vary by application and requires further investigation. It also allows for the possibility of fewer measurements being required for prediction versus training as the model learns to account for noise in the input data. The second difference is that when labeled training data is unavailable and a physical decoder is necessary, the flexible correction scheme in the decoder allows the model to compensate for non-Markovian or nonlinear dynamics since it is not bound by a Lindblad form, while still being informed by the master equation for a first order model of the dynamics.

6.3.2 Training and Data

To facilitate training, validation, and testing, $N = 32,000$ measurement trajectories are simulated for $T = 4\mu s$ using an Euler-Maruyama integrator solving (6.3) for both qubits for each of 40 values of ϵ evenly spaced on $[0, 2)$ with fixed $\Omega = 1.395$ radians/μs, and the same number of Ω values evenly spaced on $[1, 5)$ with fixed $\epsilon = 1.0$ radians/μs for a total of $K = 80$ (Ω, ϵ) pairs. Half of this collection is used for training, corresponding to every other parameter pair. The other half, corresponding to parameters midway between training values, is split evenly to be used for validation and testing. This ensures the training set contains a disjoint set of parameter pairs from the validation/test set. Two endpoints of the validation/test set are excluded from each end to ensure that the training data extends slightly beyond the domain of validation/test values.

During training, for each true parameter set $\theta_k = (\Omega_k, \epsilon_k)$, trajectory groups of a preset size d are randomly selected from the full training set and their averages are provided as input values to the model, such that each mini-batch is comprised of $M = N/d$ averaged trajectories $\{\tilde{x}_{j,k}\}_{j=1}^{M}$ where $\tilde{x}_{j,k} = \frac{1}{d}\sum_{i\in I} r_{i,k}$, $r_{i,k}$ is measurement record i for parameter set k, and I is a set of trajectory indices of size d randomly selected from $1, 2, ..., N$ without replacement until all N trajectories are used, which defines one epoch of training. In this way, each mini-batch will consist

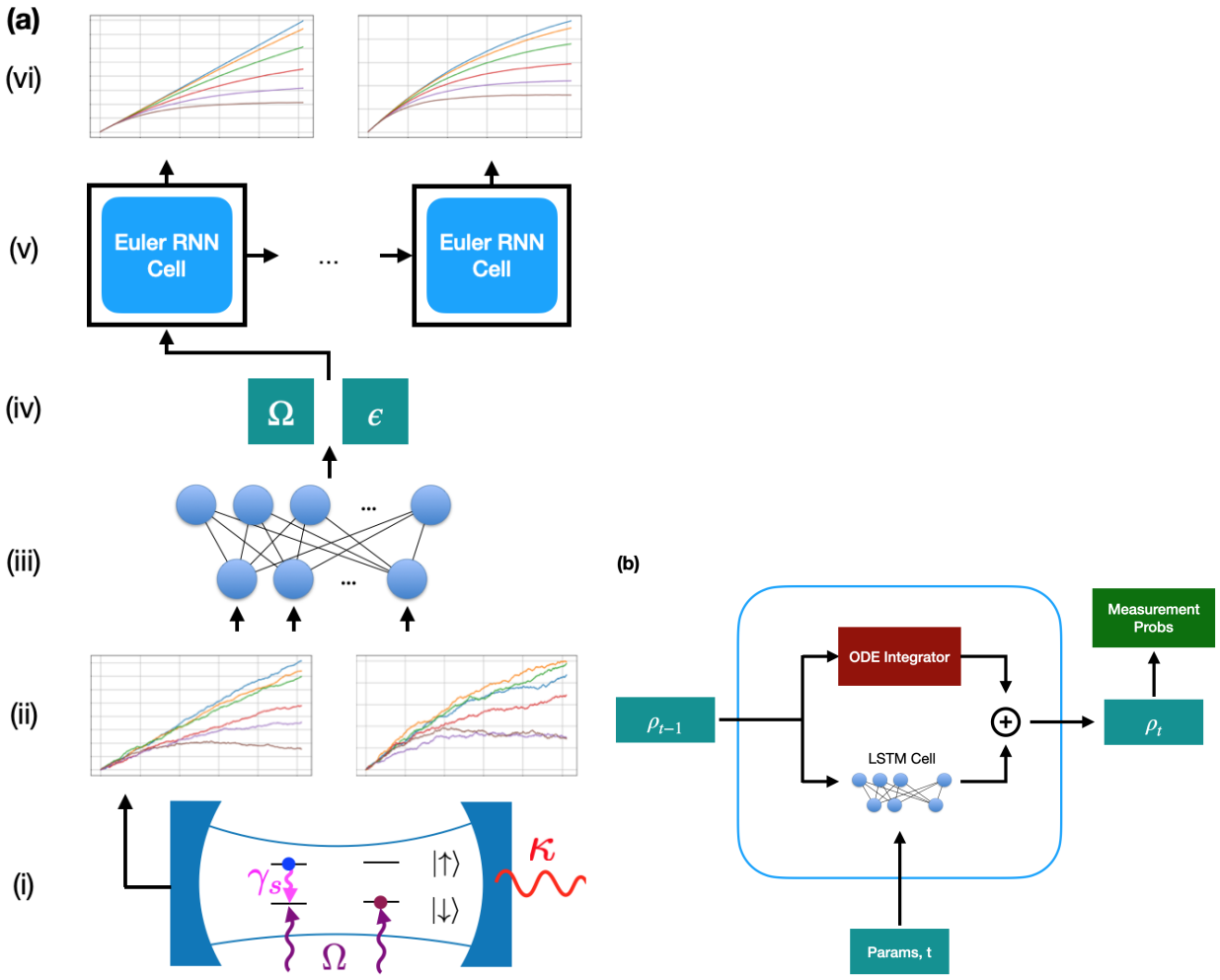


Figure 6.1: (a) Diagram of the ML model and physical system. Measurements are taken from the cavity system (i) producing noisy averaged voltage records (ii), these are sent through a neural network encoder (iii) producing system parameters as output (iv). These parameters are used by the flex integrator decoder (v) to produce noise-free voltage estimates (vi). (b) Details of the decoder RNN cell combining a standard ODE integrator with an LSTM update

of a different set of noisy trajectory groups as input to maximize the diversity of training examples. For all $j = 1, ..., M$, the value x_k to be used in the loss function is the average over the full training set of trajectories corresponding to the true parameter set θ_k, such that $x_k = \frac{1}{N}\sum_{i=1}^{N} r_{i,k}$.

Denoising autoencoders are characterized by a corruption process $C(\tilde{x}|x)$ whereby noisy inputs $\tilde{x}$ are generated for each uncorrupted training example x [71]. In our case, the full trajectory means take on the role of x, approximating the solution to the SME in the ensemble average, while the random selection of much smaller trajectory groups is the process by which the corrupted data

elements $\tilde{x}$ are generated.

In the supervised case, the loss function is the MSE between predicted parameter sets $\tilde{\theta}_k$ and the provided true values θ_k

$$\mathcal{L} = \frac{1}{pK} \sum_k \left|\left| \tilde{\theta}_k - \theta_k \right|\right|_2^2, \tag{6.6}$$

where p is the dimension of θ_k, the number of parameters being estimated. In the unsupervised case, the loss function is the full MSE between each x_k and the estimated clean measurement record

$$\mathcal{L} = \frac{1}{MKN_t} \sum_{j,k,t} |x_k(t) - \mathcal{M}(\tilde{x}_{j,k})(t)|^2 \tag{6.7}$$

where $\mathcal{M}$ denotes the model, t is the time index, and $N_t \equiv T/\Delta t$ is the number of time points excluding the initial condition.

During evaluation, validation and test errors are evaluated in a similar manner to how training mini-batches are selected, with groups of size d randomly chosen from the full validation and test sets, followed by the calculation of the MSE. This process is repeated 100 times and the average is taken as the calculated MSE for each set.

Training is performed for 300 epochs, with a learning rate of 3×10^{-3} and a decay rate of 0.99 per epoch, resetting the learning rate after every 100 epochs, at which point parameter MSE is computed for the validation and test sets. This entire process is performed 100 times, each with a different random initialization of the model parameters. The best model is considered to be the one with the smallest validation loss after the final epoch. Hyperparameter tuning for the model layer sizes was performed using a grid search for a physical system with parameters distinct from those in 6.4. A table of hyperparameter values for the models used to derive the results in the sections below is given in table 6.1.

6.4 Evaluation

To fully assess the performance of the model, we evaluate both the encoder alone to evaluate the case of labeled training data, and the full model for the case where labels are not available. We consider a range of trajectory group sizes d as well as noise-free data derived from an Euler

Encoder LSTM	32
Time distributed dense layers	[32, 16]
Encoder dense layers	[100, 50]
Decoder LSTM size	16

Table 6.1: Model layer sizes for models used in evaluation sections. Layers are listed in the order in which they appear in the model. Lists in the second column indicate stacked layers of the indicated type

integration of the unconditioned master equation. The noise-free case is denoted by ∞ as the number of trajectories in all tables. For results in this section, training mini-batches are comprised of half the total number of available groups, meaning there will be two mini-batches per epoch, except in the case where there is only one total group, e.g. in the noise-free case.

We perform measurements in the X and Y directions for the first and second qubit, respectively. To minimize the impact of the weak measurement back-action via the parameter κ, in 6.4.1 and 6.4.2, the initial state is chosen to be spin-up in the directions of measurement. A different configuration is used in 6.4.3, where more diverse measurements are needed to correct for unanticipated single particle relaxation not explicitly present in the decoder's physical model.

6.4.1 Supervised Learning

First, we study the impact of noise in the training set and how it helps or hinders the model's performance when predicting parameters from noisy data. We do this by performing estimation with two models, one trained using noise-free measurement values, and another set trained using eight groups of $d = 4,000$ averaged measurement records. Both models are then evaluated on a test set of noisy, averaged measurement records with the same group size d as the noisy training set. In both cases, the model with the best validation loss on the noisy data out of 100 randomly initialized models was used.

Figure 6.2 shows the error of the estimated parameter pair $\tilde{\theta}_k$ compared to truth for each $k = 1, 2, ..., 80$. We see that for many specific pairs, and for the overall MSE, the model trained on the noise-free measurements is producing less accurate estimates than the model trained on noisy

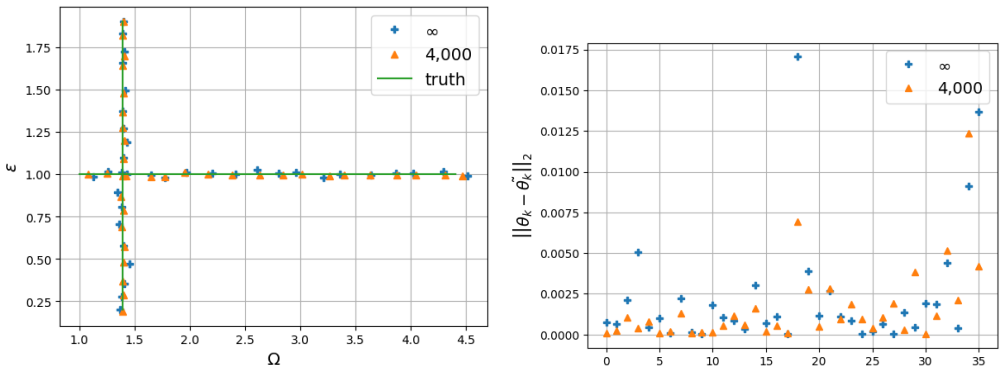


Figure 6.2: Estimations of the parameter pair $\theta_k = (\Omega_k, \epsilon_k)$ (top) and the squared error (bottom) for a model trained on clean data and noisy data, then evaluated on noisy data. Shuffle evaluated test set MSEs are (6.63e-3, 9.47e-4) and (4.37e-3, 6.26e-4) respectively

data. This is consistent in both the Ω and ϵ errors. This suggests that the model learns to account for noise, as expected, and training with noise on the level of measurements used for prediction is beneficial.

Next we examine the impact of the training group size d. Table 6.2 lists test set MSEs for (Ω, ϵ) for a number of training sets where evaluation is taking place with the same group size d used for training. The best and median were taken from the list of results sorted according to validation loss, as this is what would be known in practice during training. Here we see a steady improvement as the number of trajectories in a group increases, and noise decreases, saturating in the case where noise-free data is used for training and evaluation. Note that this test differs from the results in figure 6.2 where the model trained without noise was evaluated on noisy data, as here we consider only the case where training and test group sizes are equal.

d	Best MSE	Median MSE	Mean MSE
2,000	8.01e-3, 9.50e-4	9.15e-3, 1.09e-3	9.88e-3, 1.16e-3
4,000	4.37e-3, 6.26e-4	6.32e-3, 5.59e-4	5.67e-3, 7.07e-4
8,000	2.46e-3, 3.28e-4	3.12e-3, 5.34e-4	3.30e-3, 4.89e-4
16,000	1.70e-3, 2.45e-4	1.32e-3, 2.20e-4	1.81e-3, 3.24e-4
∞	7.83e-5, 3.28e-5	2.25e-4, 2.13e-4	4.60e-4, 2.45e-4

Table 6.2: MSE of (Ω, ϵ) estimates on supervised training set using the encoder only

6.4.2 Unsupervised Learning

We now consider the unsupervised case where labeled training data is not known, but it is assumed that the physical model in the decoder is correct, i.e. we are still not learning parameters for the drift function correction f_{LSTM}, as will be considered in 6.4.3. We are therefore using the measurement record MSE as the loss function as described above.

First we examine how parameter estimate accuracy for a fixed group size $d = 4,000$ varies with the total training set size N. Table 6.3 shows (Ω, ϵ) test set MSEs for the case where Ω is unknown but fixed at a true value of 1.395 radians/μs and ϵ is allowed to vary, for various training set sizes N. The same test set containing $16,000$ trajectories is used for each row. From the table, we see that accuracy improves significantly as the amount of training data increases, even though the groups being presented to the model for parameter estimation remain the same. This indicates that a greater diversity of noisy measurement records when training results in models that can produce more accurate parameter estimates when presented with the same number of measurements when performing prediction. This motivates using $N = 32,000$ going forward.

N	Best MSE	Median MSE	Mean MSE
4,000	6.58e-4, 8.67e-4	8.42e-4, 1.11e-3	8.36e-4, 1.09e-3
8,000	4.69e-4, 7.88e-4	4.95e-4, 9.97e-4	4.52e-4, 8.65e-4
16,000	1.83e-4, 7.49e-4	2.19e-4, 8.43e-4	2.86e-4, 8.32e-4
32,000	7.95e-5, 6.95e-4	1.61e-4, 8.58e-4	1.40e-4, 8.17e-4

Table 6.3: MSE of (Ω, ϵ) when using $d = 4,000$ trajectories to estimate parameters for varying training set sizes N for an unsupervised training set with fixed Ω and varying ϵ. Other model parameters are assumed to be known exactly

Next, we consider the impact of group size d and measurement record time spacing Δt on accuracy. Table 6.4 again shows (Ω, ϵ) test set MSEs for the case where Ω is unknown but fixed at a true value of 1.395 radians/μs and ϵ is allowed to vary. It illustrates how extremely low test set MSEs are achievable for Ω in this case, which is expected given the high volume of training data available for a single value. The accuracy of ϵ estimates depends on both the number of trajectories used to create each input sequence, as well as the time spacing at which measurements

are recorded. To evaluate measurement records with varying time spacing, trajectories simulated with $\Delta t = 2^{-8}\mu$s are sub-sampled to avoid introducing numerical integration error associated with simulations with a larger time step.

Table 6.5 shows test set MSEs for the training set where both Ω and ϵ are allowed to vary. This is a harder task as the model has less training data for each unique Ω value, hence the loss in accuracy for that parameter. The error in ϵ remains roughly the same or better, however, as in the fixed Ω case, with a best case root mean-square error of around one percent of the median test value of the ϵ parameter. Here we see a roughly linear trend in the MSE versus d at smaller time steps, as doubling the number of trajectories in a group roughly halves the MSE. This trend breaks down, however, for the $\Delta t = 2^{-4}\mu$s case, suggesting a permissive time step threshold around $\Delta t = 2^{-6}\mu$s at or below which the expected trend in accuracy versus input data size is realized.

d	$\Delta t = 2^{-8}\,\mu s$	2^{-6}	2^{-4}
2,000	9.60e-5, 1.45e-3	7.99e-5, 1.49e-3	1.05e-4, 2.67e-3
4,000	7.95e-5, 6.95e-4	1.66e-4, 8.32e-4	1.19e-4, 2.09e-3
8,000	1.43e-4, 3.59e-4	1.11e-4, 4.10e-4	8.03e-5, 1.81e-3
16,000	9.99e-5, 1.44e-4	1.31e-4, 2.14e-4	1.19e-4, 1.89e-3
∞	3.66e-5, 1.21e-5	8.75e-6, 1.03e-4	9.86e-6, 1.66e-3

Table 6.4: MSE of (Ω, ϵ) estimates on unsupervised training set with fixed Ω and varying ϵ. Other model parameters are assumed to be known exactly

d	$\Delta t = 2^{-8}\,\mu s$	2^{-6}	2^{-4}
2,000	1.00e-2, 8.73e-4	1.04e-2, 9.37e-4	1.08e-2, 1.91e-3
4,000	5.71e-3, 4.98e-4	4.64e-3, 5.11e-4	5.38e-3, 1.28e-3
8,000	3.03e-3, 2.58e-4	3.12e-3, 3.35e-4	2.74e-3, 1.02e-3
16,000	1.18e-3, 1.26e-4	9.42e-4, 2.78e-4	1.09e-3, 9.41e-4
∞	7.25e-4, 1.29e-5	5.53e-4, 5.60e-5	6.37e-4, 7.63e-4

Table 6.5: MSE of (Ω, ϵ) estimates on unsupervised training set with varying Ω and ϵ. Other model parameters are assumed to be known exactly

6.4.3 Model Correction

In this section we demonstrate the ability of the decoder to correct for dynamics not explicitly considered in the physical model (6.1). This is done by enabling training for the decoder LSTM parameters. For the datasets, we simulate $N = 30,000$ trajectories with fixed $\Omega = 1.395$ radians/μs but with varying ϵ, which is the parameter to be estimated, while adding a dissipative term for each qubit to (6.1) corresponding to the Lindblad operator $\mathcal{D}[\sqrt{\gamma_s}\sigma_i^-](\rho)$, where $\gamma_s = 0.1$ and $\sigma_i^- = \frac{1}{2}(X_i - iY_i)$ is the single particle relaxation operator mapping the excited state to the ground state.

In this case, we take a more diverse set of measurements, simulating 10,000 trajectories measuring in each of the X, Y, and Z directions for both qubits, and relaxing κ to one-fourth the value used in the last section to reduce measurement back-action. Spin-up in the Z direction is the initial state for each qubit. This more cautious approach to measurement is warranted if completely unknown effects are expected to be present. More information about the type of phenomenon, but not necessarily the magnitude, could allow for a more targeted measurement scheme, but here we keep it general. The trajectory group size used for each input was $d = 5,000$, and the best of 20 randomly initialized models was selected for the results in this section.

The results of the parameter estimation with and without the correction are shown in figure 6.3 and table 6.6. Here we see that single particle relaxation has introduced a significant bias to the estimated ϵ parameters when unaccounted for in the model, but the decoder LSTM has successfully corrected for the effect, returning the MSE to a value much closer to where it would have been had the physical model explicitly accounted for it.

Previous work has established that LSTM models are capable of learning dynamics beyond unknown Lindbladian dissipation [97, 9], and it is an interesting subject of future work to examine the robustness of our model to the presence of unanticipated non-Markovian and nonlinear dynamics in the training data.

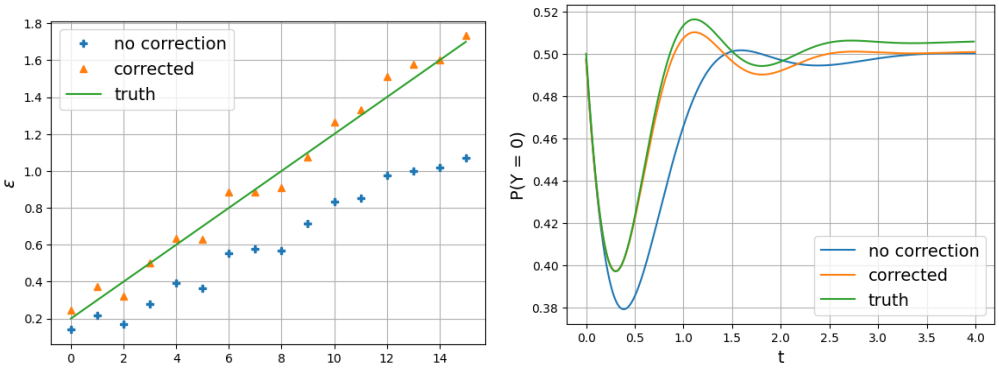


Figure 6.3: Estimate of ϵ and $P(Y_i = 0)$ when $\epsilon = 1.7$ and measuring X with spin-up in Z as the initial state for both qubits for training data simulated with single particle relaxation rate $\gamma_s = 0.1$ but a model that does not explicitly account for γ_s. Results are shown for the physical model alone and with a correction learned by the decoder LSTM

d	$\gamma_s = 0.1$	$\gamma_s = 0.0$	$\gamma_s = 0.0$ + correction
5,000	2.84e-3	0.142	3.82e-3
∞	1.87e-5	0.145	8.67e-4

Table 6.6: MSE of ϵ estimates using different γ_s values in the model. The final column shows results when learning is enabled for the free parameters in the decoder to account for the unanticipated term γ_s

6.5 Summary

We have proposed a machine learning model based on a denoising autoencoder capable of direct estimation of physical parameters in a system modeled by a stochastic master equation from weak measurement records. The model is capable of learning in a supervised or unsupervised context, and can accurately predict parameters for systems not seen in the training data. While leveraging the use of a master equation integrator to enable unsupervised learning, the autoencoder is robust to unanticipated dynamics not included in its physical model. We have demonstrated this in the case of unanticipated Lindblad dissipative terms, but it remains an interesting subject of future work to investigate the impact of other sources of error that cannot be modeled by a master equation. Another potential subject for future investigation is the ability of the model to estimate parameters for much larger systems in the supervised context, as in this case it is not subject to the exponential scaling that often limits other approaches to parameter estimation.

Chapter 7

Summary and Outlook

In summary, we have explored new properties of driven-dissipative quantum systems using established methods for classical simulation of quantum systems, and we have proposed a novel approach for learning parameters from data that leverages classical methods to expand the utility and interpretability of machine learning methods applied to the characterization of such systems.

In chapter 4, we leveraged mean-field and cumulant methods to demonstrate the emergence of a time crystal through the interplay between incoherent pumping, collective emissions, and elastic interactions. The computational approaches in this case are based on the common practice of assuming limits on the amount and type of many-body correlations that can develop within the system. The same types of assumptions are present, for example, in matrix product state ansatz that assume area law entanglement and upper bounds on the Schmidt rank of adjacent subsystems on a one dimensional lattice.

Chapter 5 also makes use of cumulant expansions, although it takes a step further by taking advantage of permutational particle symmetry to dramatically reduce the dimension of the state space, thus enabling an exact simulation free of mean-field or cumulant assumptions. It also leverages the notion of quantum trajectories, allowing for further dimensionality reduction and a tremendous speed-up through parallel simulation of the trajectories. This allowed us to explore the effects of single particle relaxation on the development of spin-squeezing in a cavity system, leading to the counter-intuitive discovery that such dissipative effects can actually improve entanglement in the transient dynamics. This could have important applications in quantum metrology and sensing.

Finally, in chapter 6 we proposed a novel approach for employing machine learning to estimate system parameters from continuous measurement data. This chapter builds on the trajectory approach of chapter 5 to demonstrate that observed continuous measurement trajectories can be used to train a machine learning model capable of predicting the parameters of systems not present in the training data, even when an exact mathematical description of the dynamics of the system is not precisely known. The methods of this chapter have the potential to scale to large systems of qubits, or to account for complex dynamics that go beyond the ability of a master equation to describe. Demonstrating these capabilities remains an interesting area for future work. The applications of this approach include the detection and quantification of crosstalk in quantum information processors, an application that is increasingly important as the size of these systems increases.

Taken together, we hope that the insights and numerical methods explored in this thesis can be of practical utility in a world of emerging quantum technologies. The fundamentally different nature of these technologies presents many challenges, and new approaches to modeling will be vital to realizing their untapped potential.

Bibliography

[1] D. A. Abanin and Z. Papić. Recent progress in many-body localization. ArXiv e-prints, May 2017.

[2] Dmitry A. Abanin, Wojciech De Roeck, and F. Huveneers. Exponentially slow heating in periodically driven many-body systems. Phys. Rev. Lett., 115:256803, Dec 2015.

[3] Stephen L Adler. Derivation of the lindblad generator structure by use of the itô stochastic calculus. Physics Letters A, 265(1):58–61, 2000.

[4] J.B. Altepeter, E.R. Jeffrey, and P.G. Kwiat. Photonic state tomography. In P.R. Berman and C.C. Lin, editors, Photonic State Tomography, volume 52 of Advances In Atomic, Molecular, and Optical Physics, pages 105–159. Academic Press, 2005.

[5] Zheng An, Jiahui Wu, Muchun Yang, D. L. Zhou, and Bei Zeng. Unified quantum state tomography and hamiltonian learning using transformer models: A language-translation-like approach for quantum systems, 2023.

[6] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando G. S. L. Brandao, David A. Buell, Brian Burkett, Yu Chen, Zijun Chen, Ben Chiaro, Roberto Collins, William Courtney, Andrew Dunsworth, Edward Farhi, Brooks Foxen, Austin Fowler, Craig Gidney, Marissa Giustina, Rob Graff, Keith Guerin, Steve Habegger, Matthew P. Harrigan, Michael J. Hartmann, Alan Ho, Markus Hoffmann, Trent Huang, Travis S. Humble, Sergei V. Isakov, Evan Jeffrey, Zhang Jiang, Dvir Kafri, Kostyantyn Kechedzhi, Julian Kelly, Paul V. Klimov, Sergey Knysh, Alexander Korotkov, Fedor Kostritsa, David Landhuis, Mike Lindmark, Erik Lucero, Dmitry Lyakh, Salvatore Mandrà, Jarrod R. McClean, Matthew McEwen, Anthony Megrant, Xiao Mi, Kristel Michielsen, Masoud Mohseni, Josh Mutus, Ofer Naaman, Matthew Neeley, Charles Neill, Murphy Yuezhen Niu, Eric Ostby, Andre Petukhov, John C. Platt, Chris Quintana, Eleanor G. Rieffel, Pedram Roushan, Nicholas C. Rubin, Daniel Sank, Kevin J. Satzinger, Vadim Smelyanskiy, Kevin J. Sung, Matthew D. Trevithick, Amit Vainsencher, Benjamin Villalonga, Theodore White, Z. Jamie Yao, Ping Yeh, Adam Zalcman, Hartmut Neven, and John M. Martinis. Quantum supremacy using a programmable superconducting processor. Nature, 574(7779):505–510, 2019.

[7] S. Autti, V. B. Eltsov, and G. E. Volovik. Observation of a time quasicrystal and its transition to a superfluid time crystal. Phys. Rev. Lett., 120:215301, May 2018.

[8] Markus P. Baden, Kyle J. Arnold, Arne L. Grimsmo, Scott Parkins, and Murray D. Barrett. Realization of the dicke model using cavity-assisted raman transitions. Phys. Rev. Lett., 113:020408, Jul 2014.

[9] Leonardo Banchi, Edward Grant, Andrea Rocchetto, and Simone Severini. Modelling non-markovian quantum processes with recurrent neural networks. New Journal of Physics, 20(12):123030, dec 2018.

[10] Ben Q. Baragiola, Bradley A. Chase, and JM Geremia. Collective uncertainty in partially polarized and partially decohered spin-$\frac{1}{2}$ systems. Phys. Rev. A, 81:032104, Mar 2010.

[11] Diego Barberena, Robert J. Lewis-Swan, James K. Thompson, and Ana Maria Rey. Driven-dissipative quantum dynamics in ultra-long-lived dipoles in an optical cavity. Phys. Rev. A, 99:053411, May 2019.

[12] Julio T. Barreiro, Markus Müller, Philipp Schindler, Daniel Nigg, Thomas Monz, Michael Chwalla, Markus Hennrich, Christian F. Roos, Peter Zoller, and Rainer Blatt. An open-system quantum simulator with trapped ions. Nature, 470(7335):486–491, February 2011.

[13] Kristian Baumann, Christine Guerlin, Ferdinand Brennecke, and Tilman Esslinger. Dicke quantum phase transition with a superfluid gas in an optical cavity. Nature, 464(7293):1301–1306, April 2010.

[14] J. S. Bell. On the einstein podolsky rosen paradox. Physics Physique Fizika, 1:195–200, Nov 1964.

[15] J. Berges, Sz. Borsányi, and C. Wetterich. Prethermalization. Phys. Rev. Lett., 93:142002, Sep 2004.

[16] Alexandre Blais, Arne L. Grimsmo, S. M. Girvin, and Andreas Wallraff. Circuit quantum electrodynamics. Rev. Mod. Phys., 93:025005, May 2021.

[17] Immanuel Bloch, Jean Dalibard, and Wilhelm Zwerger. Many-body physics with ultracold gases. Rev. Mod. Phys., 80:885–964, Jul 2008.

[18] B. J. Bloom, T. L. Nicholson, J. R. Williams, S. L. Campbell, M. Bishof, X. Zhang, W. Zhang, S. L. Bromley, and J. Ye. An optical lattice clock with accuracy and stability at the 10-18 level. Nature, 506:71, Jan 2014.

[19] Justin G. Bohnet, Zilong Chen, Joshua M. Weiner, Dominic Meiser, Murray J. Holland, and James K. Thompson. A steady-state superradiant laser with less than one intracavity photon. Nature, 484:78 EP –, 04 2012.

[20] Justin G. Bohnet, Brian C. Sawyer, Joseph W. Britton, Michael L. Wall, Ana Maria Rey, Michael Foss-Feig, and John J. Bollinger. Quantum spin dynamics and entanglement generation with hundreds of trapped ions. Science, 352(6291):1297–1301, 2016.

[21] R. Bonifacio, P. Schwendimann, and Fritz Haake. Quantum statistical theory of superradiance. i. Phys. Rev. A, 4:302–313, Jul 1971.

[22] M. Bonitz. Quantum Kinetic Theory. Springer Cham, 2015.

[23] Pranjal Bordia, Henrik Lüschen, Ulrich Schneider, Michael Knap, and Immanuel Bloch. Classical time crystals. Nature Physics, 13:460–464, 2017.

[24] Dirk Bouwmeester, Artur Ekert, and Anton Zeilinger, editors. The Physics of Quantum Information. Springer Berlin Heidelberg, 2000.

[25] S.L. Braunstein and A.K. Pati. Quantum Information with Continuous Variables. Springer Netherlands, 2012.

[26] Boris Braverman, Akio Kawasaki, Edwin Pedrozo-Peñafiel, Simone Colombo, Chi Shu, Zeyang Li, Enrique Mendez, Megan Yamoah, Leonardo Salvi, Daisuke Akamatsu, Yanhong Xiao, and Vladan Vuletić. Near-unitary spin squeezing in ^{171}Yb. Phys. Rev. Lett., 122:223203, Jun 2019.

[27] Ferdinand Brennecke, Tobias Donner, Stephan Ritter, Thomas Bourdel, Michael Köhl, and Tilman Esslinger. Cavity QED with a bose–einstein condensate. Nature, 450(7167):268–271, November 2007.

[28] S. L. Bromley, S. Kolkowitz, T. Bothwell, D. Kedar, A. Safavi-Naini, M. L. Wall, C. Salomon, A. M. Rey, and J. Ye. Dynamics of interacting fermions under spinñorbit coupling in an optical lattice clock. Nature Physics, 14:399ñ404, 2018.

[29] Todd A. Brun. A simple model of quantum trajectories. American Journal of Physics, 70(7):719–737, 07 2002.

[30] Patrick Bruno. Comment on "quantum time crystals". Phys. Rev. Lett., 110:118901, Mar 2013.

[31] Patrick Bruno. Comment on "space-time crystals of trapped ions". Phys. Rev. Lett., 111:029301, Jul 2013.

[32] Patrick Bruno. Impossibility of spontaneously rotating time crystals: A no-go theorem. Phys. Rev. Lett., 111:070402, Aug 2013.

[33] Steven L. Brunton, Joshua L. Proctor, and J. Nathan Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. Proceedings of the National Academy of Sciences, 113(15):3932–3937, 2016.

[34] B. Buca, J. Tindall, and D. Jaksch, 2018.

[35] S. L. Campbell, R. B. Hutson, G. E. Marti, A. Goban, N. Darkwah Oppong, R. L. McNally, L. Sonderhouse, J. M. Robinson, W. Zhang, B. J. Bloom, and J. Ye. A fermi-degenerate three-dimensional optical lattice clock. Science, 358(6359):90–94, 2017.

[36] Giuseppe Carleo, Yusuke Nomura, and Masatoshi Imada. Constructing exact representations of quantum many-body systems with deep neural networks. Nature Communications, 9(1):5322, 2018.

[37] Giuseppe Carleo and Matthias Troyer. Solving the quantum many-body problem with artificial neural networks. Science, 355(6325):602–606, 2017.

[38] H J Carmichael. Analytical and numerical results for the steady state in cooperative resonance fluorescence. J. Phys. B: Atom. Molec. Phys., 13:3551–3575, 1980.

[39] H. J. Carmichael. Statistical methods in quantum optics, volume 2. Springer, 1999.

[40] Howard Carmichael. An Open Systems Approach to Quantum Optics. Springer Berlin Heidelberg, 1993.

[41] Juan Carrasquilla and Giacomo Torlai. How to use neural networks to investigate quantum many-body physics. PRX Quantum, 2:040201, Nov 2021.

[42] Juan Carrasquilla, Giacomo Torlai, Roger G. Melko, and Leandro Aolita. Reconstructing quantum states with generative models. Nature Machine Intelligence, 1(3):155–161, 2019.

[43] Leonardo K. Castelano, Iann Cunha, Fabricio S. Luiz, Marcelo V. de Souza Prado, and Felipe F. Fanchini. Physics informed neural networks learning a two-qubit hamiltonian, 2023.

[44] Kathleen Champion, Bethany Lusch, J. Nathan Kutz, and Steven L. Brunton. Data-driven discovery of coordinates and governing equations. Proceedings of the National Academy of Sciences, 116(45):22445–22451, 2019.

[45] Ching-Kit Chan, Tony E. Lee, and Sarang Gopalakrishnan. Limit-cycle phase in driven-dissipative spin systems. Phys. Rev. A, 91:051601, 2015.

[46] Bradley A. Chase and J. M. Geremia. Collective processes of an ensemble of spin-1/2 particles. Phys. Rev. A, 78:052101, Nov 2008.

[47] Yanming Che, Clemens Gneiting, and Franco Nori. Exponentially improved efficient machine learning for quantum many-body states with provable guarantees, 2023.

[48] Ricky T. Q. Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, Advances in Neural Information Processing Systems, volume 31. Curran Associates, Inc., 2018.

[49] Zilong Chen. Breaking Quantum Limits with Collective Cavity-QED: Generation of Spin Squeezed States via Quantum Non-Demolition Measurements. PhD thesis, University of Colorado at Boulder, 2013.

[50] Soonwon Choi, Joonhee Choi, Renate Landig, Georg Kucsko, Hengyun Zhou, Junichi Isoya, Fedor Jelezko, Shinobu Onoda, Hitoshi Sumiya, Vedika Khemani, Curt von Keyserlingk, Norman Y. Yao, Eugene Demler, and Mikhail D. Lukin. Observation of discrete time-crystalline order in a disordered dipolar many-body system. Nature, 543(7644):221–225, 2017.

[51] Emanuele G. Dalla Torre, Johannes Otterbach, Eugene Demler, Vladan Vuletic, and Mikhail D. Lukin. Dissipative preparation of spin squeezed atomic ensembles in a steady state. Phys. Rev. Lett., 110:120402, Mar 2013.

[52] Emily J. Davis, Gregory Bentsen, Lukas Homeier, Tracy Li, and Monika H. Schleier-Smith. Photon-mediated spin-exchange dynamics of spin-1 atoms. Phys. Rev. Lett., 122:010405, Jan 2019.

[53] Brian M. de Silva, David M. Higdon, Steven L. Brunton, and J. Nathan Kutz. Discovery of physics from data: Universal laws and discrepancies. Frontiers in Artificial Intelligence, 3, 2020.

[54] C. Deutsch, F. Ramirez-Martinez, C. Lacroûte, F. Reinhard, T. Schneider, J. N. Fuchs, F. Piéchon, F. Laloë, J. Reichel, and P. Rosenbusch. Spin self-rephasing and very long coherence times in a trapped atomic ensemble. Phys. Rev. Lett., 105:020401, Jul 2010.

[55] Ivan Deutsch. Lecture notes for physics 566 and 581, quantum optics i and ii. `https://cquic.unm.edu/research/research-groups/deutsch-research-group/courses.html`.

[56] S. Diehl, A. Micheli, A. Kantian, B. Kraus, H. P. Büchler, and P. Zoller. Quantum states and phases in driven open quantum systems with cold atoms. Nature Physics, 4:878 EP –, 09 2008.

[57] P. D. Drummond. Observables and moments of cooperative resonance fluorescence. Phys. Rev. A, 22:1179–1184, Sep 1980.

[58] P. D. Drummond and H J Carmichael. Voltera cycles and the cooperative fluorescence critical point. Optics Communications, 27(1), 1978.

[59] P. D. Drummond and S. S. Hassan. Multiple sidebands in cooperative resonance fluorescence: Exact semiclassical results. Phys. Rev. A, 22:662–670, Aug 1980.

[60] X. Du, L. Luo, B. Clancy, and J. E. Thomas. Observation of anomalous spin segregation in a trapped fermi gas. Phys. Rev. Lett., 101:150401, Oct 2008.

[61] Dominic V. Else, Bela Bauer, and Chetan Nayak. Floquet time crystals. Phys. Rev. Lett., 117:090402, Aug 2016.

[62] Dominic V. Else, Bela Bauer, and Chetan Nayak. Prethermal phases of matter protected by time-translation symmetry. Phys. Rev. X, 7:011026, Mar 2017.

[63] Samuel Fernández-Lorenzo and Diego Porras. Quantum sensing close to a dissipative phase transition: Symmetry breaking and criticality as metrological resources. Phys. Rev. A, 96:013817, Jul 2017.

[64] J. M. Fink, R. Bianchetti, M. Baur, M. Göppl, L. Steffen, S. Filipp, P. J. Leek, A. Blais, and A. Wallraff. Dressed collective qubit states and the tavis-cummings model in circuit qed. Phys. Rev. Lett., 103:083601, Aug 2009.

[65] E. Flurin, L. S. Martin, S. Hacohen-Gourgy, and I. Siddiqi. Using a recurrent neural network to reconstruct quantum dynamics of a superconducting qubit from physical observations. Phys. Rev. X, 10:011006, Jan 2020.

[66] Jay Gambetta, Alexandre Blais, M. Boissonneault, A. A. Houck, D. I. Schuster, and S. M. Girvin. Quantum trajectory approach to circuit qed: Quantum jumps and the zeno effect. Phys. Rev. A, 77:012112, Jan 2008.

[67] Martin Gärttner, Philipp Hauke, and Ana Maria Rey. Relating out-of-time-order correlations to entanglement via multiple-quantum coherences. Phys. Rev. Lett., 120:040402, 2018.

[68] Élie Genois, Jonathan A. Gross, Agustin Di Paolo, Noah J. Stevenson, Gerwin Koolstra, Akel Hashim, Irfan Siddiqi, and Alexandre Blais. Quantum-tailored machine-learning characterization of a superconducting qubit. PRX Quantum, 2:040355, Dec 2021.

[69] Zongping Gong, Ryusuke Hamazaki, and Masahito Ueda. Discrete time-crystalline order in cavity and circuit qed systems. Phys. Rev. Lett., 120:040404, Jan 2018.

[70] Alejandro González-Tudela and Diego Porras. Mesoscopic entanglement induced by spontaneous emission in solid-state quantum optics. Phys. Rev. Lett., 110:080502, Feb 2013.

[71] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Deep Learning. MIT Press, 2016. `http://www.deeplearningbook.org`.

[72] M. Gring, M. Kuhnert, T. Langen, T. Kitagawa, B. Rauer, M. Schreitl, I. Mazets, D. Adu Smith, E. Demler, and J. Schmiedmayer. . Science, 337:1318, 2012.

[73] M. Gross and S. Haroche. Superradiance: An essay on the theory of collective spontaneous emission. Physics Reports, 93(5):301–396, 1982.

[74] Mansura Habiba and Barak A. Pearlmutter. Neural ordinary differential equation based recurrent neural network model. In 2020 31st Irish Signals and Systems Conference (ISSC), pages 1–6, 2020.

[75] Jutho Haegeman, J. Ignacio Cirac, Tobias J. Osborne, Iztok Pižorn, Henri Verschelde, and Frank Verstraete. Time-dependent variational principle for quantum lattices. Phys. Rev. Lett., 107:070601, Aug 2011.

[76] Jutho Haegeman, Christian Lubich, Ivan Oseledets, Bart Vandereycken, and Frank Verstraete. Unifying time evolution and optimization with matrix product states. Phys. Rev. B, 94:165116, Oct 2016.

[77] Wen Wei Ho, Soonwon Choi, Mikhail D. Lukin, and Dmitry A. Abanin. Critical time crystals in dipolar systems. Phys. Rev. Lett., 119:010602, Jul 2017.

[78] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. Neural Comput., 9(8):1735–1780, nov 1997.

[79] O. Hosten, R. Krishnakumar, N. J. Engelsen, and M. A. Kasevich. Quantum phase magnification. Science, 352(6293):1552–1555, 2016.

[80] Jiazhong Hu, Wenlan Chen, Zachary Vendeiro, Alban Urvoy, Boris Braverman, and Vladan Vuletić. Vacuum spin squeezing. Phys. Rev. A, 96:050301, Nov 2017.

[81] Hsin-Yuan Huang, Sitan Chen, and John Preskill. Learning to predict arbitrary quantum processes, 2023.

[82] Hsin-Yuan Huang, Richard Kueng, and John Preskill. Predicting many properties of a quantum system from very few measurements. Nature Physics, 16(10):1050–1057, jun 2020.

[83] Hsin-Yuan Huang, Richard Kueng, Giacomo Torlai, Victor V. Albert, and John Preskill. Provably efficient machine learning for quantum many-body problems. Science, 377(6613), sep 2022.

[84] Ross B. Hutson, Akihisa Goban, G. Edward Marti, Lindsay Sonderhouse, Christian Sanner, and Jun Ye. Engineering quantum states of matter for atomic clocks in shallow optical lattices. Phys. Rev. Lett., 123:123401, Sep 2019.

[85] F. Iemini, A. Russomanno, J. Keeling, M. Schirò , M. Dalmonte, and R. Fazio. Boundary time crystals. Physical Review Letters, 121(3), jul 2018.

[86] Kurt Jacobs and Daniel A. Steck. A straightforward introduction to continuous quantum measurement. Contemporary Physics, 47(5):279–303, 2006.

[87] E. T. Jaynes and F. W. Cummings. Comparison of quantum and semiclassical radiation theories with application to the beam maser. Proceedings of the IEEE, 51:89, 1963.

[88] P. Jurcevic, H. Shen, P. Hauke, C. Maier, T. Brydges, C. Hempel, B. P. Lanyon, M. Heyl, R. Blatt, and C. F. Roos. Direct observation of dynamical quantum phase transitions in an interacting many-body system. Phys. Rev. Lett., 119:080501, Aug 2017.

[89] Kadierdan Kaheman, Eurika Kaiser, Benjamin Strom, J. Nathan Kutz, and Steven L. Brunton. Learning discrepancy models from experimental data, 2019.

[90] M. J. Kastoryano, F. Reiter, and A. S. Sørensen. Dissipative preparation of entanglement in optical cavities. Phys. Rev. Lett., 106:090502, Feb 2011.

[91] Vedika Khemani, Achilleas Lazarides, Roderich Moessner, and S. L. Sondhi. Phase structure of driven quantum systems. Phys. Rev. Lett., 116:250401, Jun 2016.

[92] Vedika Khemani, C. W. von Keyserlingk, and S. L. Sondhi. Defining time crystals via representation theory. Physical Review B, 96(11), sep 2017.

[93] Masahiro Kitagawa and Masahito Ueda. Squeezed spin states. Phys. Rev. A, 47:5138–5143, Jun 1993.

[94] Takuya Kitagawa, Erez Berg, Mark Rudner, and Eugene Demler. Topological characterization of periodically driven quantum systems. Phys. Rev. B, 82:235114, Dec 2010.

[95] G. Kleine Büning, J. Will, W. Ertmer, E. Rasel, J. Arlt, C. Klempt, F. Ramirez-Martinez, F. Piéchon, and P. Rosenbusch. Extended coherence time on the clock transition of optically trapped rubidium. Phys. Rev. Lett., 106:240801, Jun 2011.

[96] Jonathan Kohler, Nicolas Spethmann, Sydney Schreppler, and Dan M. Stamper-Kurn. Cavity-assisted measurement and coherent control of collective atomic spin oscillators. Phys. Rev. Lett., 118:063604, Feb 2017.

[97] G. Koolstra, N. Stevenson, S. Barzili, L. Burns, K. Siva, S. Greenfield, W. Livingston, A. Hashim, R. K. Naik, J. M. Kreikebaum, K. P. O'Brien, D. I. Santiago, J. Dressel, and I. Siddiqi. Monitoring fast superconducting qubit dynamics using a neural network. Phys. Rev. X, 12:031017, Jul 2022.

[98] B. Kraus, H. P. Büchler, S. Diehl, A. Kantian, A. Micheli, and P. Zoller. Preparation of entangled states by quantum markov processes. Phys. Rev. A, 78:042307, Oct 2008.

[99] Hanna Krauter, Christine A. Muschik, Kasper Jensen, Wojciech Wasilewski, Jonas M. Petersen, J. Ignacio Cirac, and Eugene S. Polzik. Entanglement generated by dissipation and steady state entanglement of two macroscopic objects. Phys. Rev. Lett., 107:080503, Aug 2011.

[100] Y. Kuramoto. Chemical Oscillations, Waves, and Turbulence. Dover books on chemistry. Dover Publications, 2003.

[101] T. Kuwahara, T. Mori, and K. Saito. Floquet–magnus theory and generic transient dynamics in periodically driven many-body quantum systems. Annals of Physics, 367:96–124, Apr 2016.

[102] Tim Langen, Thomas Gasenzer, and Jˆrg Schmiedmayer. Prethermalization and universal dynamics in near-integrable quantum systems. J. Stat. Mech., 2016(6):064009, 2016.

[103] Melvin Lax. The lax–onsager regression 'theorem' revisited. Optics Communications, 179(1):463 – 476, 2000.

[104] Achilleas Lazarides and Roderich Moessner. Fate of a discrete time crystal in an open system. Phys. Rev. B, 95:195135, May 2017.

[105] Tony E. Lee and Ching-Kit Chan. Dissipative transverse-field ising model: Steady-state correlations and spin squeezing. Phys. Rev. A, 88:063811, Dec 2013.

[106] Tony E. Lee, Ching-Kit Chan, and Susanne F. Yelin. Dissipative phase transitions: Independent versus collective decay and spin squeezing. Phys. Rev. A, 90:052109, Nov 2014.

[107] Tony E. Lee, Florentin Reiter, and Nimrod Moiseyev. Entanglement and spin squeezing in non-hermitian phase transitions. Phys. Rev. Lett., 113:250401, Dec 2014.

[108] A. Lerose, B. Zunkovic, J. Marino, A. Gambassi, and A. Silva. Impact of non-equilibrium fluctuations on pre-thermal dynamical phase transitions in interacting spin chains. arXiv:1807.09797, 2018.

[109] Ian D. Leroux, Monika H. Schleier-Smith, and Vladan Vuletić. Implementation of cavity squeezing of a collective atomic spin. Phys. Rev. Lett., 104:073602, Feb 2010.

[110] Robert J. Lewis-Swan, Matthew A. Norcia, Julia R. K. Cline, James K. Thompson, and Ana Maria Rey. Robust spin squeezing via photon-mediated interactions on an optical clock transition. Phys. Rev. Lett., 121:070403, Aug 2018.

[111] J Li and G S Paraoanu. Generation and propagation of entanglement in driven coupled-qubit systems. New Journal of Physics, 11(11):113020, nov 2009.

[112] Y. Lin, J. P. Gaebler, F. Reiter, T. R. Tan, R. Bowler, A. S. Sørensen, D. Leibfried, and D. J. Wineland. Dissipative production of a maximally entangled steady state of two quantum bits. Nature, 504(7480):415–418, November 2013.

[113] Netanel H. Lindner, Gil Refael, and Victor Galitski. Floquet topological insulator in semiconductor quantum wells. Nature Physics, 7(6):490–495, 2011.

[114] Andrew D. Ludlow, Martin M. Boyd, Jun Ye, E. Peik, and P. O. Schmidt. Optical atomic clocks. Rev. Mod. Phys., 87:637–701, Jun 2015.

[115] Bethany Lusch, J. Nathan Kutz, and Steven L. Brunton. Deep learning for universal linear embeddings of nonlinear dynamics. Nature Communications, 9(1):4950, 2018.

[116] Jian Ma, Xiaoguang Wang, C.P. Sun, and Franco Nori. Quantum spin squeezing. Physics Reports, 509(2):89–165, 2011.

[117] Katarzyna Macieszczak, Mădălin Guţă, Igor Lesanovsky, and Juan P. Garrahan. Dynamical phase transitions as a resource for quantum enhanced metrology. Phys. Rev. A, 93:022103, Feb 2016.

[118] G. Edward Marti, Ross B. Hutson, Akihisa Goban, Sara L. Campbell, Nicola Poli, and Jun Ye. Imaging optical frequencies with 100 μHz precision and 1.1 μm resolution. Phys. Rev. Lett., 120:103201, Mar 2018.

[119] M J Martin, M Bishof, M D Swallows, X Zhang, C Benko, J Von-Stecher, A V Gorshkov, A M Rey, and Jun Ye. A quantum many-body spin system in an optical lattice clock. Science, 341:632–636, 2013.

[120] D. Meiser, Jun Ye, D. R. Carlson, and M. J. Holland. Prospects for a millihertz-linewidth laser. Phys. Rev. Lett., 102:163601, Apr 2009.

[121] J. A. Mlynek, A. A. Abdumalikov, C. Eichler, and A. Wallraff. Observation of dicke superradiance for two artificial atoms in a cavity with high decay rate. Nature Communications, 5(1):5186, 2014.

[122] Klaus Mølmer. Density matrices and the quantum monte-carlo method in quantum optics. Lectures presented at the Winter School on Quantum Optics, International Centre of Theoretical Physics, Trieste, Italy, 1994.

[123] Klaus Mølmer, Yvan Castin, and Jean Dalibard. Monte carlo wave-function method in quantum optics. J. Opt. Soc. Am. B, 10(3):524–538, Mar 1993.

[124] S Morrison and A S Parkins. Dissipation-driven quantum phase transitions in collective spin systems. Journal of Physics B: Atomic, Molecular and Optical Physics, 41(19):195502, sep 2008.

[125] Juan A. Muniz, Diego Barberena, Robert J. Lewis-Swan, Dylan J. Young, Julia R. K. Cline, Ana Maria Rey, and James K. Thompson. Exploring non-equilibrium phases of matter with a million long-lived optical dipoles in a cavity. Nature, 580:602–607, 2020.

[126] K. Nakatsugawa, T. Fujii, and S. Tanda. Quantum time crystal by decoherence: Proposal with an incommensurate charge density wave ring. Phys. Rev. B, 96:094308, Sep 2017.

[127] Rahul Nandkishore and David A. Huse. Many-body localization and thermalization in quantum statistical mechanics. Annual Review of Condensed Matter Physics, 6(1):15–38, 2015.

[128] Michael A. Nielsen and Isaac L. Chuang. Quantum Computation and Quantum Information: 10th Anniversary Edition. Cambridge University Press, 2010.

[129] Matthew A. Norcia, Robert J. Lewis-Swan, Julia R. K. Cline, Bihui Zhu, Ana M. Rey, and James K. Thompson. Cavity-mediated collective spin-exchange interactions in a strontium superradiant laser. Science, 361(6399):259–262, 2018.

[130] Matthew A. Norcia, Matthew N. Winchester, Julia R. K. Cline, and James K. Thompson. Superradiance on the millihertz linewidth strontium clock transition. Science Advances, 2(10), 2016.

[131] Philippe Nozières. Time crystals: Can diamagnetic currents drive a charge density wave into rotation? Europhysics Letters, 103(5):57008, sep 2013.

[132] Christopher Olah. Understanding lstms. https://colah.github.io/posts/2015-08-Understanding-LSTMs/.

[133] William D. Oliver. Quantum computing takes flight. https://www.nature.com/articles/d41586-019-03173-4, Octover 2019.

[134] Luca Pezzè, Augusto Smerzi, Markus K. Oberthaler, Roman Schmied, and Philipp Treutlein. Quantum metrology with nonclassical states of atomic ensembles. Rev. Mod. Phys., 90:035005, Sep 2018.

[135] M. B. Plenio and P. L. Knight. The quantum-jump approach to dissipative dynamics in quantum optics. Rev. Mod. Phys., 70:101–144, Jan 1998.

[136] Pedro Ponte, Anushya Chandran, Z. Papić, and Dmitry A. Abanin. Periodically driven ergodic and many-body localized quantum systems. Annals of Physics, 353:196–204, 2015.

[137] Pedro Ponte, Z. Papić, F. Huveneers, and Dmitry A. Abanin. Many-body localization in periodically driven systems. Phys. Rev. Lett., 114:140401, Apr 2015.

[138] J. F. Poyatos, J. I. Cirac, and P. Zoller. Quantum reservoir engineering with laser cooled trapped ions. Phys. Rev. Lett., 77:4728–4731, Dec 1996.

[139] John Preskill. John preskill notes on quantum information theory. http://theory.caltech.edu/~preskill/ph229/index.html.

[140] John Preskill. Quantum Computing in the NISQ era and beyond. Quantum, 2:79, August 2018.

[141] Qucumber: Reconstruction of a postive real wavefunction. https://qucumber.readthedocs.io/en/stable/_examples/Tutorial1_TrainPosRealWaveFunction/tutorial_quantum_ising.html.

[142] J M Radcliffe. Some properties of coherent spin states. Journal of Physics A: General Physics, 4(3):313–323, may 1971.

[143] Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. Physics informed deep learning (part ii): Data-driven discovery of nonlinear partial differential equations, 2017.

[144] A. M. Rey, L. Jiang, M. Fleischhauer, E. Demler, and M. D. Lukin. Many-body protected entanglement generation in interacting spin systems. Phys. Rev. A, 77:052305, May 2008.

[145] Andrea Rocchetto, Edward Grant, Sergii Strelchuk, Giuseppe Carleo, and Simone Severini. Learning hard quantum distributions with variational autoencoders. npj Quantum Information, 4(1):28, 2018.

[146] Jared Rovny, Robert L. Blum, and Sean E. Barrett. Observation of discrete-time-crystal signatures in an ordered dipolar many-body system. Phys. Rev. Lett., 120:180603, May 2018.

[147] Yulia Rubanova, Ricky T. Q. Chen, and David K Duvenaud. Latent ordinary differential equations for irregularly-sampled time series. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, Advances in Neural Information Processing Systems, volume 32. Curran Associates, Inc., 2019.

[148] K Sacha and J Zakrzewski. Time crystals: a review. Reports on Progress in Physics, 81:016401, 2018.

[149] Krzysztof Sacha. Modeling spontaneous breaking of time-translation symmetry. Phys. Rev. A, 91:033617, Mar 2015.

[150] A. Safavi-Naini, R. J. Lewis-Swan, J. G. Bohnet, M. Gärttner, K. A. Gilmore, J. E. Jordan, J. Cohn, J. K. Freericks, A. M. Rey, and J. J. Bollinger. Verification of a many-ion simulator of the dicke model through slow quenches across a phase transition. Phys. Rev. Lett., 121:040503, Jul 2018.

[151] Hidetsugu Sakaguchi and Yoshiki Kuramoto. A soluble active rotater model showing phase transitions via mutual entertainment. Progress of Theoretical Physics, 76(3):576–581, 1986.

[152] Mohan Sarovar, Timothy Proctor, Kenneth Rudinger, Kevin Young, Erik Nielsen, and Robin Blume-Kohout. Detecting crosstalk errors in quantum information processors. Quantum, 4:321, September 2020.

[153] S. Schneider and G. J. Milburn. Entanglement in the steady state of a collective-angular-momentum (dicke) model. Phys. Rev. A, 65:042107, Mar 2002.

[154] Nathan Shammah, Shahnawaz Ahmed, Neill Lambert, Simone De Liberato, and Franco Nori. Open quantum systems with local and collective incoherent processes: Efficient numerical simulations using permutational invariance. Phys. Rev. A, 98:063815, Dec 2018.

[155] Athreya Shankar, John Cooper, Justin G. Bohnet, John J. Bollinger, and Murray Holland. Steady-state spin synchronization through the collective motion of trapped ions. Phys. Rev. A, 95:033423, Mar 2017.

[156] Alfred Shapere and Frank Wilczek. Classical time crystals. Phys. Rev. Lett., 109:160402, Oct 2012.

[157] John A. Smolin, Jay M. Gambetta, and Graeme Smith. Efficient method for computing the maximum-likelihood quantum state from measurements with additive gaussian noise. Phys. Rev. Lett., 108:070502, Feb 2012.

[158] A. Sørensen, L. M. Duan, J. I. Cirac, and P. Zoller. Many-particle entanglement with bose–einstein condensates. Nature, 409(6816):63–66, 2001.

[159] Daniel A Steck. Quantum and atom optics. `http://steck.us/teaching`, July 2022.

[160] Andrzej Syrwid, Jakub Zakrzewski, and Krzysztof Sacha. Time crystal behavior of excited eigenstates. Phys. Rev. Lett., 119:250602, Dec 2017.

[161] Giacomo Torlai, Christopher J. Wood, Atithi Acharya, Giuseppe Carleo, Juan Carrasquilla, and Leandro Aolita. Quantum process tomography with unsupervised learning and tensor networks. Nature Communications, 14(1):2858, 2023.

[162] Vinay Tripathi, Huo Chen, Mostafa Khezri, Ka-Wa Yip, E.M. Levenson-Falk, and Daniel A. Lidar. Suppression of crosstalk in superconducting qubits using dynamical decoupling. Phys. Rev. Appl., 18:024068, Aug 2022.

[163] K. Tucker, D. Barberena, R. J. Lewis-Swan, J. K. Thompson, J. G. Restrepo, and A. M. Rey. Facilitating spin squeezing generated by collective dynamics with single-particle decoherence. Phys. Rev. A, 102:051701, Nov 2020.

[164] K Tucker, B Zhu, R J Lewis-Swan, J Marino, F Jimenez, J G Restrepo, and A M Rey. Shattered time: can a dissipative time crystal survive many-body correlations? New Journal of Physics, 20(12):123003, dec 2018.

[165] H. Uys, M. J. Biercuk, A. P. VanDevender, C. Ospelkaus, D. Meiser, R. Ozeri, and J. J. Bollinger. Decoherence due to elastic rayleigh scattering. Phys. Rev. Lett., 105:200401, Nov 2010.

[166] Varun D. Vaidya, Yudan Guo, Ronen M. Kroeze, Kyle E. Ballantine, Alicia J. Kollár, Jonathan Keeling, and Benjamin L. Lev. Tunable-range, photon-mediated atomic interactions in multimode cavity qed. Phys. Rev. X, 8:011002, Jan 2018.

[167] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In Advances in Neural Information Processing Systems, pages 5998–6008, 2017.

[168] C. W. von Keyserlingk, Vedika Khemani, and S. L. Sondhi. Absolute stability and spatiotemporal long-range order in floquet systems. Phys. Rev. B, 94:085112, Aug 2016.

[169] C. W. von Keyserlingk and S. L. Sondhi. Phase structure of one-dimensional interacting floquet systems. ii. symmetry-broken phases. Phys. Rev. B, 93:245146, Jun 2016.

[170] Ronen Vosk, David A. Huse, and Ehud Altman. Theory of the many-body localization transition in one-dimensional systems. Phys. Rev. X, 5:031032, Sep 2015.

[171] D. F. Walls. Squeezed states of light. Nature, 306(5939):141–146, 1983.

[172] D. F. Walls, P. D. Drummond, S. S. Hassan, and H. J. Carmichael. Non-Equilibrium Phase Transitions in Cooperative Atomic Systems. Progress of Theoretical Physics Supplement, 64:307–320, 02 1978.

[173] Haruki Watanabe and Masaki Oshikawa. Absence of quantum time crystals. Phys. Rev. Lett., 114:251603, Jun 2015.

[174] Frank Wilczek. Superfluidity and space-time translation symmetry breaking. Phys. Rev. Lett., 111:250402, Dec 2013.

[175] D. J. Wineland, J. J. Bollinger, W. M. Itano, F. L. Moore, and D. J. Heinzen. Spin squeezing and reduced quantum noise in spectroscopy. Phys. Rev. A, 46:R6797–R6800, Dec 1992.

[176] Howard Wiseman. Quantum Trajectories and Feedback. PhD thesis, University of Queensland, 1994.

[177] Elie Wolfe and S. F. Yelin. Spin squeezing by means of driven superradiance, 2014.

[178] Minghui Xu, D. A. Tieri, and M. J. Holland. Simulating open quantum systems by applying su(4) to quantum master equations. Phys. Rev. A, 87:062101, 2013.

[179] N. Y. Yao, A. C. Potter, I.-D. Potirniche, and A. Vishwanath. Discrete time crystals: Rigidity, criticality, and realizations. Phys. Rev. Lett., 118:030401, Jan 2017.

[180] Norman Y. Yao, Chetan Nayak, Leon Balents, and Michael P. Zaletel. Classical discrete time crystals. Nature Physics, 16(4):438–447, feb 2020.

[181] J. Zhang, P. W. Hess, A. Kyprianidis, P. Becker, A. Lee, J. Smith, G. Pagano, I. D. Potirniche, A. C. Potter, A. Vishwanath, N. Y. Yao, and C. Monroe. Observation of a discrete time crystal. Nature, 543(7644):217–220, 2017.

[182] J. Zhang, G. Pagano, P. W. Hess, A. Kyprianidis, P. Becker, H. Kaplan, A. V. Gorshkov, Z.-X. Gong, and C. Monroe. Observation of a many-body dynamical phase transition with a 53-qubit quantum simulator. Nature, 551(7682):601–604, nov 2017.

[183] Liangsheng Zhang, Vedika Khemani, and David A. Huse. A floquet model for the many-body localization transition. Phys. Rev. B, 94:224202, Dec 2016.

[184] Yuan Zhang, YuXiang Zhang, and Klaus Molmer. Monte-carlo simulations of superradiant lasing. New Journal of Physics, 20, November 2018.

[185] B Zhu, J Schachenmayer, M Xu, F Herrera, J G Restrepo, M J Holland, and A M Rey. Synchronization of interacting quantum dipoles. New Journal of Physics, 17(8):083063, 2015.

[186] Barton Zwiebach. Mit physics 8.05 lecture notes.

Appendix A

Time Crystals

This appendix lists additional results associated with chapter 4.

A.1 Introduction

The dynamics of the system described in the main text are given by the following master equation for the density operator $\hat{\rho}$:

$$\begin{aligned}\frac{d\hat{\rho}}{dt} = \quad & i\sum_a [\frac{\delta_a}{2}\hat{\sigma}_a^z, \hat{\rho}] - i\Gamma \sum_{a\neq b} g[\hat{\sigma}_a^+\hat{\sigma}_b^-, \hat{\rho}] \\ & -\frac{\Gamma}{2}\sum_{a,b} f(\{\hat{\sigma}_a^+\hat{\sigma}_b^-, \hat{\rho}\} - 2\hat{\sigma}_b^-\hat{\rho}\hat{\sigma}_a^+) - \frac{W}{2}\sum_a (\{\hat{\sigma}_a^-\hat{\sigma}_a^+, \hat{\rho}\} - 2\hat{\sigma}_a^+\hat{\rho}\hat{\sigma}_a^-). \end{aligned} \tag{A.1}$$

Here we have dropped self-interaction in the second sum because it gives rise to a negligible single particle term that can be removed by moving to an appropriate rotating frame. Equations of motion for the expectation of a particular observable $\hat{\Omega}$ can be derived from the master equation according to the expression

$$\frac{d}{dt}\langle\hat{\Omega}\rangle = \frac{d}{dt}\mathrm{Tr}\left[\hat{\Omega}\hat{\rho}\right] = \mathrm{Tr}\left[\hat{\Omega}\frac{d\hat{\rho}}{dt}\right],$$

assuming $\hat{\Omega}$ has no explicit time dependence. The density operator $\hat{\rho}$ in the above is a $2^N \times 2^N$ complex valued matrix. The exponential scaling of the size of the density matrix with the number of particles makes exact solutions to the master equation inaccessible for $N \gtrsim 16$. In what follows, we examine the various approaches we have taken to address this issue and also provide detailed derivations of results mentioned in the main text.

We first examine the mean-field approximation and associated results. Next, we look beyond the mean-field to a second order cumulant expansion that allows for the modeling of both disorder and many-body correlations. Finally, we derive results for mutual information growth and explore its relationship with eigenvalues of the Liouvillian operator in the master equation.

A.2 Mean-field analysis

This section includes results for the mean-field analysis with disorder that were referenced in the main text.

A.2.1 Disorder

In this section we consider the effect of a heterogeneous frequency distribution on the time crystal. In particular, we are interested in the robustness of the oscillation frequency. To this end, we first rewrite Eqs. (4.3) in terms of the order parameter $Ze^{i\psi} = \sum_a R_a e^{i\phi_a}/N$, assuming $N \gg 1$:

$$\dot{s}_a = -fN\Gamma R_a Z\cos(\psi-\phi_a) - 2gN\Gamma R_a Z\sin(\psi-\phi_a) + W\left(\frac{1}{2}-s_a\right) - f\Gamma\left(\frac{1}{2}+s_a\right), \tag{A.2}$$

$$\dot{R}_a = -\frac{(W+f\Gamma)}{2}R_a + fNs_a\Gamma Z\cos(\psi-\phi_a) + 2gNs_a\Gamma Z\sin(\psi-\phi_a), \tag{A.3}$$

$$\dot{\phi}_a = -\delta_a - \frac{2s_a gNZ\Gamma}{R_a}\cos(\psi-\phi_a) + \frac{s_a fNZ\Gamma}{R_a}\sin(\psi-\phi_a). \tag{A.4}$$

Now we look for steady rotating solutions of Eqs. (4.3) of the form $\phi_a = \psi - \theta_a$, with R_a, s_a, and θ_a constant, and $\dot{\psi} = \omega_{\mathrm{MF}}$. Inserting this Ansatz in Eqs. (A.2), solving for R_a, s_a, and θ_a, and requiring the self-consistent conditions

$$Z = \sum_a R_a\cos(\theta_a), \tag{A.5}$$

$$0 = \sum_a R_a\sin(\theta_a), \tag{A.6}$$

we obtain

$$Z = \frac{1}{N}\sum_{a=1}^{N}\frac{Z\Gamma(W-\Gamma f)(Nf(\Gamma f+W)-4Ng(\delta_a+\omega_{\mathrm{MF}}))}{(\Gamma f+W)\left[4(\delta_a+\omega_{\mathrm{MF}})^2+\Gamma^2\left(2N^2f^2Z^2+8N^2g^2Z^2+f^2\right)+W^2+2\Gamma W\right]}, \tag{A.7}$$

$$0 = \frac{1}{N}\sum_{a=1}^{N}\frac{-Z(W-\Gamma f)(2Ng(\Gamma f+W)+2Nf(\delta_a+\omega_{\mathrm{MF}}))}{(\Gamma f+W)\left[4(\delta_a+\omega_{\mathrm{MF}})^2+\Gamma^2\left(2N^2f^2Z^2+8N^2g^2Z^2+f^2\right)+W^2+2\Gamma W\right]}. \tag{A.8}$$

When $N \gg 1$, we can approximate the sums by integrals over the distribution of frequencies. Given a distribution of frequencies, the integrals can be evaluated numerically and Z and $\omega_{\rm MF}$ can be obtained using root-finding methods. For simplicity, however, we consider a Lorenzian distribution $H(\delta) = \frac{\Delta}{\pi(\Delta^2+\delta^2)}$. In this case the integrals can be evaluated by contour integration, and we find that the order parameter Z is obtained from the implicit equation

$$Z = \frac{Z\Gamma N f(W-\Gamma f)\sqrt{\Gamma^2\left(2N^2f^2Z^2+8N^2g^2Z^2+f^2\right)+W^2+2\Gamma fW}}{\left[2\Gamma^2N^2f^2Z^2+(W+\Gamma f)^2\right]\left(2\Delta+\sqrt{\Gamma^2\left(2N^2f^2Z^2+8N^2g^2Z^2+f^2\right)+W^2+2\Gamma fW}\right)}, \quad \text{(A.9)}$$

while the frequency is given by

$$\omega_{\rm MF} = \frac{g(\Gamma f+W)\left(2\Delta+\sqrt{\Gamma^2\left(2N^2f^2Z^2+8N^2g^2Z^2+f^2\right)+2\Gamma fW+W^2}\right)}{2f\sqrt{\Gamma^2\left(2N^2f^2Z^2+8N^2g^2Z^2+f^2\right)+2\Gamma fW+W^2}}. \quad \text{(A.10)}$$

While implicit Eq. (A.9) is cumbersome, it can easily be solved with root finding methods. A simplification can be made by considering the $N \gg 1$ limit. In this case, for a given pumping $W = wN$, we obtain, to leading order in $1/N$,

$$\omega_{\rm MF} = \frac{wgN}{f}\left(1+\frac{f\Gamma}{wN}+\frac{2\Delta}{N\sqrt{w^2+2f^2Z^2\Gamma^2+8g^2Z^2\Gamma^2}}\right), \quad \text{(A.11)}$$

showing that in the thermodynamic limit the time crystal frequency is robust to heterogeneity in the oscillator detunings. If we substitute optimal order parameter $Z_{opt} = 1/\sqrt{8}$ and optimal pumping $w_{opt} = f\Gamma/2$ into (A.11), we obtain

$$\omega_{\rm MF}^{opt} = \frac{wgN}{f}\left(1+\frac{f\Gamma}{wN}+\frac{\sqrt{8}\Delta}{N\Gamma\sqrt{f^2+2g^2}}\right). \quad \text{(A.12)}$$

From here we can easily see that $\delta\omega(\Delta) \sim \frac{\sqrt{8}\Delta}{N\Gamma\sqrt{f^2+2g^2}}$ at optimal pumping as discussed in the main text. In the next Section, we will compare this expression with numerically determined values from the cumulant expansion and see that there is close agreement.

Finally, we note that although here we solved for stationary rotating solutions, there can be other types of solutions, including quasiperiodic or chaotic solutions. These other possible solutions, and the bifurcations leading to them, will be explored in future research.

A.3 Cumulant expansion analysis

A second-order cumulant expansion assumes joint cumulants of order three and higher are zero, resulting in the following expansion for third-order expectations:

$$\langle\hat{\sigma}_a^\alpha\hat{\sigma}_b^\beta\hat{\sigma}_c^\gamma\rangle \approx \langle\hat{\sigma}_a^\alpha\hat{\sigma}_b^\beta\rangle\langle\hat{\sigma}_c^\gamma\rangle + \langle\hat{\sigma}_b^\beta\hat{\sigma}_c^\gamma\rangle\langle\hat{\sigma}_a^\alpha\rangle + \langle\hat{\sigma}_a^\alpha\hat{\sigma}_c^\gamma\rangle\langle\hat{\sigma}_b^\beta\rangle - 2\langle\hat{\sigma}_a^\alpha\rangle\langle\hat{\sigma}_b^\beta\rangle\langle\hat{\sigma}_c^\gamma\rangle. \tag{A.13}$$

In this section, we present the equations of motion that result, compare the mean-field and cumulant frequencies for systems with nonzero disorder, and perform benchmarking of the cumulant model against exact solutions.

A.3.1 Equations

The relevant equations of motion can be derived as:

$$\frac{d\langle\hat{\sigma}_a^+\rangle}{dt} = -i\delta_a\langle\hat{\sigma}_a^+\rangle - \frac{\Gamma f+W}{2}\langle\hat{\sigma}_a^+\rangle + \frac{\Gamma}{2}\sum_{a\neq b}(f-i2g)\langle\hat{\sigma}_a^z\hat{\sigma}_b^+\rangle, \tag{A.14}$$

$$\begin{aligned}\frac{d\langle\hat{\sigma}_a^z\hat{\sigma}_b^+\rangle}{dt} = &-\left(i\delta_b+\frac{3\Gamma f+3W}{2}\right)\langle\hat{\sigma}_a^z\hat{\sigma}_b^+\rangle - (\Gamma f-W)\langle\hat{\sigma}_b^+\rangle - \frac{\Gamma}{2}(f+i2g)\langle\hat{\sigma}_a^+\rangle - \Gamma f\langle\hat{\sigma}_a^+\hat{\sigma}_b^z\rangle\\ &+\frac{\Gamma}{2}\sum_{j\neq a,b}(f_{bj}-i2g_{bj})\langle\hat{\sigma}_a^z\hat{\sigma}_b^z\hat{\sigma}_j^+\rangle - \Gamma\sum_{j\neq a,b}(f_{aj}+i2g_{aj})\langle\hat{\sigma}_a^+\hat{\sigma}_b^+\hat{\sigma}_j^-\rangle - \Gamma\sum_{j\neq a,b}(f_{aj}-i2g_{aj})\langle\hat{\sigma}_a^-\hat{\sigma}_b^+\hat{\sigma}_j^+\rangle,\end{aligned} \tag{A.15}$$

$$\frac{d\langle\hat{\sigma}_a^z\rangle}{dt} = -i\Gamma\sum_{a\neq b}2g(\langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle - \langle\hat{\sigma}_a^-\hat{\sigma}_b^+\rangle) - \Gamma\sum_{b\neq a}f(\langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle + \langle\hat{\sigma}_a^-\hat{\sigma}_b^+\rangle) - (\Gamma f+W)\langle\hat{\sigma}_a^z\rangle + (W-\Gamma f), \tag{A.16}$$

$$\begin{aligned}\frac{d\langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle}{dt} = &-i(\delta_a-\delta_b)\langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle - \frac{i\Gamma}{2}g(\langle\hat{\sigma}_a^z\rangle - \langle\hat{\sigma}_b^z\rangle) + \frac{\Gamma}{2}\sum_{j\neq a,b}(f_{aj}-i2g_{aj})\langle\hat{\sigma}_a^z\hat{\sigma}_b^-\hat{\sigma}_j^+\rangle\\ &+\frac{\Gamma}{2}\sum_{j\neq a,b}(f_{bj}+i2g_{bj})\langle\hat{\sigma}_b^z\hat{\sigma}_a^+\hat{\sigma}_j^-\rangle + \frac{\Gamma}{2}f\left(\langle\hat{\sigma}_a^z\hat{\sigma}_b^z\rangle + \frac{\langle\hat{\sigma}_a^z\rangle+\langle\hat{\sigma}_b^z\rangle}{2}\right) - (\Gamma f+W)\langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle,\end{aligned} \tag{A.17}$$

$$\begin{aligned}\frac{d\langle\hat{\sigma}_a^z\hat{\sigma}_b^z\rangle}{dt} = &-i\Gamma\sum_{j\neq a,b}2g_{aj}[\langle\hat{\sigma}_a^+\hat{\sigma}_b^z\hat{\sigma}_j^-\rangle - \langle\hat{\sigma}_a^-\hat{\sigma}_b^z\hat{\sigma}_j^+\rangle] - i\Gamma\sum_{j\neq a,b}2g_{bj}[\langle\hat{\sigma}_b^+\hat{\sigma}_a^z\hat{\sigma}_j^-\rangle - \langle\hat{\sigma}_b^-\hat{\sigma}_a^z\hat{\sigma}_j^+\rangle]\\ &-\Gamma\sum_{j\neq a,b}f_{aj}[\langle\hat{\sigma}_a^+\hat{\sigma}_b^z\hat{\sigma}_j^-\rangle + \langle\hat{\sigma}_a^-\hat{\sigma}_b^z\hat{\sigma}_j^+\rangle] - \Gamma\sum_{j\neq a,b}f_{bj}[\langle\hat{\sigma}_b^+\hat{\sigma}_a^z\hat{\sigma}_j^-\rangle + \langle\hat{\sigma}_b^-\hat{\sigma}_a^z\hat{\sigma}_j^+\rangle]\\ &+(W-\Gamma f)(\langle\hat{\sigma}_a^z\rangle + \langle\hat{\sigma}_b^z\rangle) - 2(\Gamma f+W)\langle\hat{\sigma}_a^z\hat{\sigma}_b^z\rangle + 2\Gamma f[\langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle + \langle\hat{\sigma}_a^-\hat{\sigma}_b^+\rangle],\end{aligned} \tag{A.18}$$

$$\begin{aligned}\frac{d\langle\hat{\sigma}_a^+\hat{\sigma}_b^+\rangle}{dt} = &-i(\delta_a+\delta_b)\langle\hat{\sigma}_a^+\hat{\sigma}_b^+\rangle + \frac{\Gamma}{2}\sum_{j\neq a\neq b}(f_{aj}-i2g_{aj})\langle\hat{\sigma}_a^z\hat{\sigma}_b^+\hat{\sigma}_j^+\rangle + \frac{\Gamma}{2}\sum_{j\neq a\neq b}(f_{bj}-i2g_{bj})\langle\hat{\sigma}_b^z\hat{\sigma}_a^+\hat{\sigma}_j^+\rangle\\ &-(\Gamma f+W)\langle\hat{\sigma}_a^+\hat{\sigma}_b^+\rangle.\end{aligned} \tag{A.19}$$

In these equations, $a \neq b$. Applying the cumulant expansion (A.13), the above equations become closed and can be solved to obtain the dynamics. Note that this a system of equations with a size that grows as $\mathcal{O}(N^2)$.

A.3.2 Disorder

Allowing $\Delta > 0$, we can solve for the two-time correlation function numerically and examine its frequency. Fig. A.1 compares this frequency to the derived mean-field value seen in Eq. (A.12) where pumping and synchronization are assumed to be optimal and $N = 100$. As can be seen, there is close agreement, and the two values remain close as Δ increases. It should be noted that for our simulations, we take the $\{\delta_a\}_{a=1}^{N}$ to be spaced at intervals of equal probability according to the Lorentzian probability density function to ensure a zero mean.

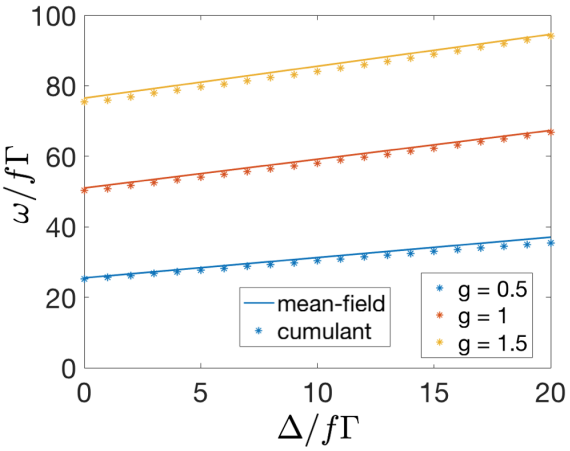


Figure A.1: Comparison of numerically determined cumulant two-time correlation frequency (symbols) and predicted mean-field frequency (solid lines) at optimal pumping when $N = 100$

A.3.3 Benchmarking

Fig. A.2 compares the cumulant expansion solution to that of the exact solver in two cases where the exact solution is accessible. In the left panel, we compare the order parameter Z_Q of the synchronized system in the case where the number of particles is small. In the right panel, we compare the real part of the two-time correlation function $C(\tau)$ for a larger system without

disorder. In both cases, we see close agreement between the cumulant expansion and the exact solution.

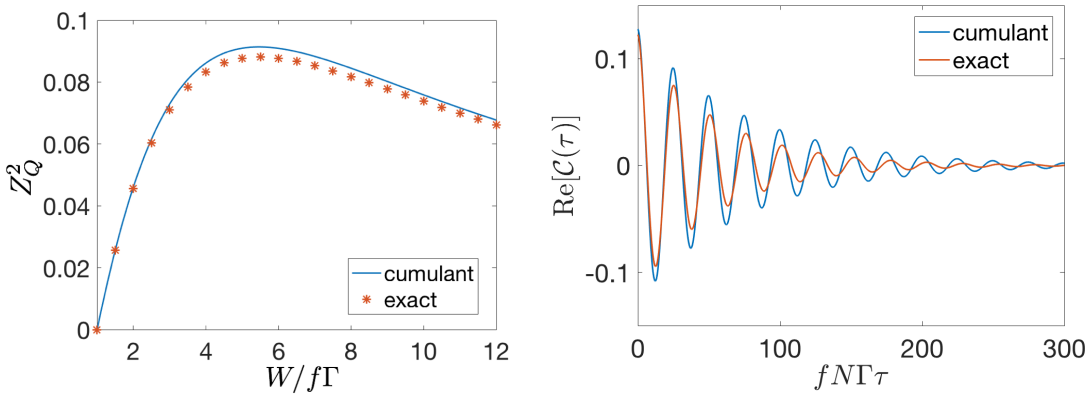


Figure A.2: Comparing the cumulant expansion solution against the exact solver for the order parameter when $N = 5$, $f = 1$, $g = 0.5$, $\Delta = 1$ for a range of $W/f\Gamma$ (left panel) and for $\mathrm{Re}[C(\tau)]$ where $N = 100$, $f = 1$, $g = 0.5$, $\Delta = 0$, and optimal W (right panel)

A.4 Mutual Information

To gain insight into the growth of mutual information in the transient dynamics, we derive an expression for the derivative of the two particle mutual information for zero disorder at small times up to first-order in t. To do this, we write each of the necessary density operators, $\hat{\rho}_{AB}$, $\hat{\rho}_A$, and $\hat{\rho}_B$, in the Pauli basis according to

$$\hat{\rho} = \frac{1}{2^M} \sum_{\vec{\alpha}} c_{\vec{\alpha}} \hat{\rho}_{\vec{\alpha}}, \quad \hat{\rho}_{\vec{\alpha}} \equiv \bigotimes_{j=1}^{M} \hat{\sigma}^{\alpha_j},$$

where $\hat{\sigma}^{\alpha_j} \in \{\mathbf{1}, \hat{\sigma}^+, \hat{\sigma}^-, \hat{\sigma}^z\}$ and M is the number of particles in the subsystem, e.g. $M = 2$ in the case of $\hat{\rho}_{AB}$. Note that $c_{\vec{\alpha}} = \mathrm{Tr}[\hat{\rho}\hat{\rho}_{\vec{\alpha}}] = \langle \hat{\rho}_{\vec{\alpha}} \rangle$, thus we can write $\hat{\rho}_{AB}$ in terms of second-order expectations of Pauli operators, and $\hat{\rho}_A$ and $\hat{\rho}_B$ in terms of first-order expectations. We then expand each density operator to second order in t according to,

$$\hat{\rho}(t) \approx \hat{\rho}\Big|_{t=0} + t\partial_t \hat{\rho}\Big|_{t=0} + \frac{t^2}{2}\partial_{tt}\hat{\rho}\Big|_{t=0}, \tag{A.20}$$

using the cumulant equations when $\Delta = 0$ to expand the derivatives of expectations of Pauli operators to be in terms of the expectations of the operators themselves. This allows for evaluation

at $t = 0$. Eigenvalues can then be determined using first-order perturbation theory. Note that particle symmetry when $\Delta = 0$ ensures that expressions derived for eigenvalues of $\hat{\rho}_{AB}$, $\hat{\rho}_A$, and $\hat{\rho}_B$ are the same for any choice of A and B. The initial condition is taken such that all particles are spin up in the $\hat{x}$-direction at $t = 0$. This approach results in the following expression for the mutual information derivative at small times:

$$\begin{aligned}\frac{dI_{AB}}{dt} = \frac{1}{4}\left[-2(f\Gamma+W)\log\left[t(f\Gamma+W)\right]+(2f\Gamma+W)\log\left[t(2f\Gamma+W)\right]+W\log(tW)\right]\\ +\Bigg[\frac{\Gamma^2}{2}\left(-2f^2-4g^2\right)+W\log\left(tW\right)\left(f\Gamma(N+2)-3W\right)\\ +\log\left[t\left(f\Gamma+W\right)\right]\left(\Gamma^2\left(f^2\left(-\left(N^2-2\right)\right)-4g^2(N-1)\right)+2f\Gamma NW+W^2\right)\\ +\log\left[t\left(2f\Gamma+W\right)\right]\left(\Gamma^2\left(f^2\left(N^2-5\right)+4g^2(N-2)\right)-f\Gamma(3N+4)W+W^2\right)\\ +\frac{1}{2}\left(\Gamma^2\left(3f^2+4g^2\right)+2f\Gamma W+W^2\right)\log\left[t^2\left(\Gamma^2\left(3f^2+4g^2\right)+2f\Gamma W+W^2\right)\right]\Bigg]\frac{t}{4}+\mathcal{O}(t^2).\end{aligned} \tag{A.21}$$

Note that the t in the arguments of the logarithms will cancel and they are in the above expression only for dimensional consistency. The approximation for the density operators in (A.20) will have an error term that is to the same order in t and N. This imposes the restriction $f\Gamma t \ll 1/N$. This works well for our purposes, since we are interested in characteristic time scales of the time crystal oscillations which, as we have seen from ω_{MF} and ω_{C}, have a frequency that grows linearly with N to leading order. Using the change of variables $\eta \equiv fN\Gamma t$ where $\eta \ll 1$, going to first order in time, and expanding to leading order for large N, Eq. (A.21) becomes

$$\frac{dI_{AB}}{dt} \approx \frac{f^2\Gamma^2}{4Nw}+\left(-\frac{3f^3\Gamma^3}{8Nw^2}+\frac{5f^2\Gamma^2}{4Nw}-\frac{9f\Gamma}{4N}+\frac{g^2\Gamma^2}{Nw}+\frac{(f\Gamma-2w)^2}{4w}\right)\eta+\mathcal{O}\left(\frac{1}{N^2}\right). \tag{A.22}$$

Note that this vanishes in the thermodynamic limit only when pumping is optimal, i.e. $w = f\Gamma/2$. At this value of pumping, we get

$$\frac{dI_{AB}}{dt} \approx \frac{f\Gamma}{2}\left(\frac{1}{N}+\left(\frac{4g^2}{f^2}-\frac{5}{2}\right)\frac{\eta}{N}\right)+\mathcal{O}\left(\frac{1}{N^2}\right),$$

thus leading to constant dI_{AB}/dt contours of $g/f \lesssim \sqrt{N}$ as noted in the main text. Using optimal pumping, $\Gamma = f = 1$, $\eta = 0.03$, and plotting $1/I'_{AB}(\eta)$ over a range of g and N produces the

contour plot found in Fig. 2 of the main text. Large values correspond to regions of slow growth in mutual information while small values indicate rapid growth on time scales characteristic of the time crystal oscillations. Extending our analysis using second order perturbation theory produces a result that is very similar at optimal pumping, but includes a correction term on the order of $\eta \log(\eta)/N$ with a coefficient of the same order in g. This reduces the error of the approximation with increasing g to a point where it is very small over the parameter regime of interest without changing the broader analysis given above.

For a small system, we can draw a connection between the growth of mutual information and the eigenvalues of the Liouvillian $\mathcal{L}$ given by the master equation (A.1) which has the form $\partial_t \hat{\rho} = \mathcal{L}[\hat{\rho}]$. To examine frequencies, we restrict our attention to the lowest lying eigenvalues of $\mathcal{L}$ with nonzero imaginary part. Also, since we are looking for contributors to two-time correlation frequency, we consider only those eigenvalues with eigenvectors that are not orthogonal to $\lim_{t\to\infty} S^- \hat{\rho}(t)$, which is what the operator $e^{\mathcal{L}\tau}$ operates on in the quantum regression formula. Fig. A.3 plots the imaginary part of these eigenvalues for a system at optimal pumping where $N = 5$. The blue line is a linear fit of the numerically determined frequency of the two-time correlation function $\mathcal{C}(\tau)$. As expected, the imaginary parts of the eigenvalues align closely with this frequency, particularly for small g/f. In panel (a), the color gives the magnitude of the real part of the Liouvillian eigenvalues, which determines the decay rate of $\mathcal{C}(\tau)$ for the corresponding frequency mode. As can be seen, this decay is gradual for small g/f and increases with this parameter. In panel (b), the color gives the derivative of two particle mutual information at $\eta = 0.03$ in the transient dynamics starting at spin up in the $\hat{x}$-direction. Note that the increase in decay rate aligns closely with the increase in mutual information as we vary g/f.

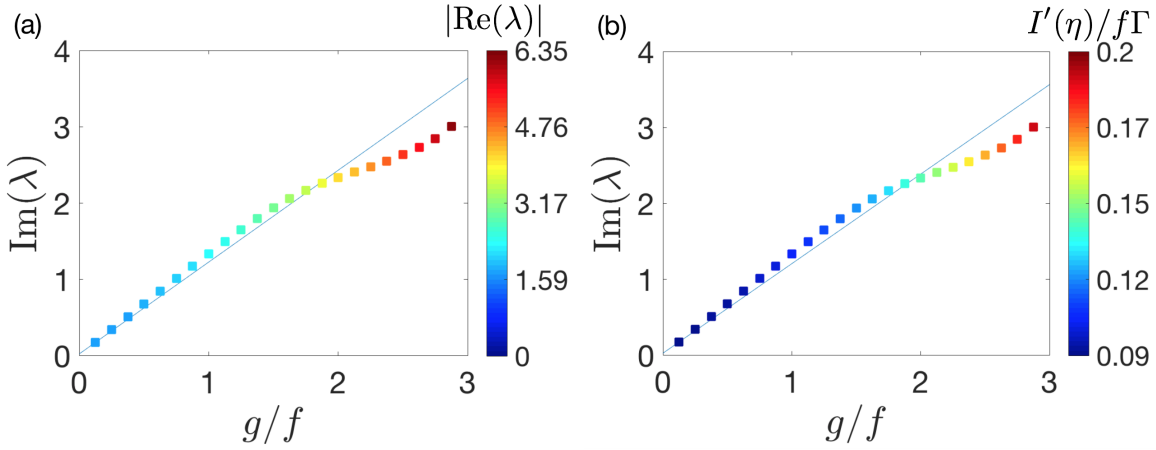


Figure A.3: Imaginary part of the Liouvillian eigenvalue versus g/f plotted with a linear fit of the numerically determined frequency of $\mathcal{C}(\tau)$ (blue). In panel (a), color corresponds to the magnitude of the eigenvalue real part and in panel (b) it corresponds to the mutual information derivative at $\eta = 0.03$ in the transient dynamics

Appendix B

Enhanced Spin Squeezing

This appendix lists additional results associated with chapter 5.

B.1 Cumulant Expansion

To obtain a model that can be numerically integrated efficiently but also allows for quantum correlations, we turn our attention to a cumulant expansion approximation. Specifically, we extend the mean-field analysis by computing equations of motion for expectation values of products of two Pauli operators, e.g., $\partial_t \langle \hat{\sigma}_i^\alpha \hat{\sigma}_j^\beta \rangle = \mathrm{Tr}[\hat{\sigma}_i^\alpha \hat{\sigma}_j^\beta \partial_t \hat{\rho}]$, in addition to the equations of motion for the expectations of individual Pauli operators. Together these form a hierarchy of equations which we truncate by assuming that third order expectations can be factorized into products of lower-order

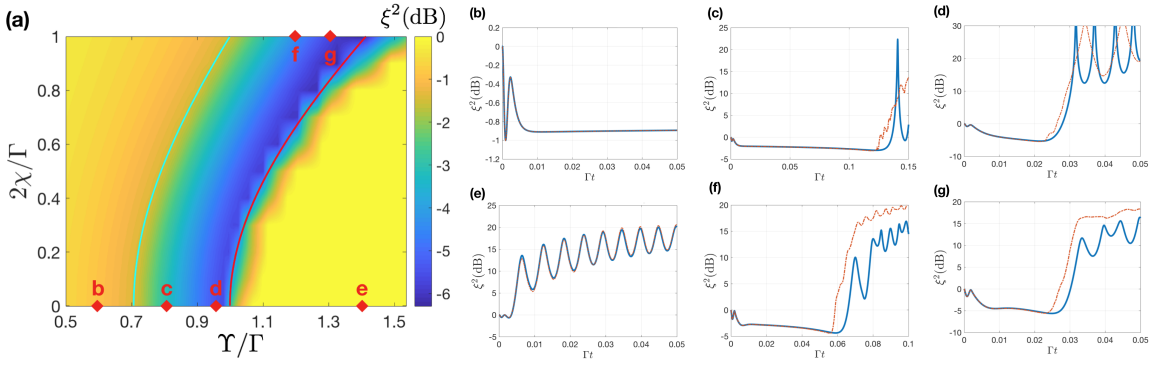


Figure B.1: (a) Minimum squeezing attained from cumulant model after $t_{\min} = 0.003$ where $N = 2000$ and $\gamma_s/\Gamma = 20$ with the phase boundaries Υ_c'/Γ (blue) and Υ_c/Γ (red). (b)-(g) Comparison of squeezing dynamics between cumulant (blue) and MCWF (red) where χ/Γ and $2\Omega/N\Gamma$ values are shown in red in panel (a)

terms. Specifically, we assume they factorize according to

$$\langle\hat{\sigma}_a^\alpha\hat{\sigma}_b^\beta\hat{\sigma}_c^\gamma\rangle \approx \langle\hat{\sigma}_a^\alpha\hat{\sigma}_b^\beta\rangle\langle\hat{\sigma}_c^\gamma\rangle + \langle\hat{\sigma}_b^\beta\hat{\sigma}_c^\gamma\rangle\langle\hat{\sigma}_a^\alpha\rangle + \langle\hat{\sigma}_a^\alpha\hat{\sigma}_c^\gamma\rangle\langle\hat{\sigma}_b^\beta\rangle$$
$$-2\langle\hat{\sigma}_a^\alpha\rangle\langle\hat{\sigma}_b^\beta\rangle\langle\hat{\sigma}_c^\gamma\rangle, \tag{B.1}$$

in the case where a, b and c are distinct. Operator products for the same particle are resolved using the usual Pauli relations. Combining this factorization with the particle symmetry present in the master equation, and assuming all particles start with the same initial conditions, this results in a closed system of six complex ordinary differential equations:

$$\frac{d\langle\hat{\sigma}_a^+\rangle}{dt} = -\frac{\Gamma+\gamma_s}{2}\langle\hat{\sigma}_a^+\rangle + \frac{1}{2}(N-1)(\Gamma-i2\chi)\langle\hat{\sigma}_a^z\hat{\sigma}_b^+\rangle - i\frac{\Omega}{2}\langle\sigma_a^z\rangle + i\chi\langle\sigma_a^+\rangle, \tag{B.2}$$

$$\begin{aligned}\frac{d\langle\hat{\sigma}_a^z\hat{\sigma}_b^+\rangle}{dt} = &-\frac{3}{2}(\Gamma+\gamma_s)\langle\hat{\sigma}_a^z\hat{\sigma}_b^+\rangle - (\Gamma+\gamma_s)\langle\hat{\sigma}_b^+\rangle - \frac{1}{2}(\Gamma+i2\chi)\langle\hat{\sigma}_a^+\rangle - \Gamma\langle\hat{\sigma}_a^+\hat{\sigma}_b^z\rangle \\ &+\frac{1}{2}(N-2)(\Gamma-i2\chi)\langle\hat{\sigma}_a^z\hat{\sigma}_b^z\hat{\sigma}_j^+\rangle - (N-2)(\Gamma+i2\chi)\langle\hat{\sigma}_a^+\hat{\sigma}_b^+\hat{\sigma}_j^-\rangle - (N-2)(\Gamma-i2\chi)\langle\hat{\sigma}_a^-\hat{\sigma}_b^+\hat{\sigma}_j^+\rangle \\ &-i\frac{\Omega}{2}\left(2(\langle\sigma_a\sigma_+\rangle b+ - \langle\sigma_a\sigma_-\rangle b+) + \langle\sigma_a\sigma_z\rangle bz\right) + i\chi\langle\sigma_a\sigma_z\rangle b+,\end{aligned} \tag{B.3}$$

$$\begin{aligned}\frac{d\langle\hat{\sigma}_a^z\rangle}{dt} = &-2i\chi(N-1)(\langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle - \langle\hat{\sigma}_a^-\hat{\sigma}_b^+\rangle) - \Gamma(N-1)(\langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle + \langle\hat{\sigma}_a^-\hat{\sigma}_b^+\rangle) - (\Gamma+\gamma_s)(\langle\hat{\sigma}_a^z\rangle+1) \\ &+2\Omega\mathrm{Im}[\langle\sigma_a^+\rangle],\end{aligned} \tag{B.4}$$

$$\begin{aligned}\frac{d\langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle}{dt} = &\frac{1}{2}(N-2)(\Gamma-i2\chi)\langle\hat{\sigma}_a^z\hat{\sigma}_b^-\hat{\sigma}_j^+\rangle + \frac{1}{2}(N-2)(\Gamma+i2\chi)\langle\hat{\sigma}_b^z\hat{\sigma}_a^+\hat{\sigma}_j^-\rangle + \frac{\Gamma}{2}\left(\langle\hat{\sigma}_a^z\hat{\sigma}_b^z\rangle + \langle\hat{\sigma}_a^z\rangle\right) \\ &-(\Gamma+\gamma_s)\langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle - \Omega\mathrm{Im}[\langle\sigma_a\sigma_z\rangle b+],\end{aligned} \tag{B.5}$$

$$\begin{aligned}\frac{d\langle\hat{\sigma}_a^z\hat{\sigma}_b^z\rangle}{dt} = &-i(N-2)2\chi[\langle\hat{\sigma}_a^+\hat{\sigma}_b^z\hat{\sigma}_j^-\rangle - \langle\hat{\sigma}_a^-\hat{\sigma}_b^z\hat{\sigma}_j^+\rangle] - i(N-2)2\chi[\langle\hat{\sigma}_b^+\hat{\sigma}_a^z\hat{\sigma}_j^-\rangle - \langle\hat{\sigma}_b^-\hat{\sigma}_a^z\hat{\sigma}_j^+\rangle] \\ &-(N-2)\Gamma[\langle\hat{\sigma}_a^+\hat{\sigma}_b^z\hat{\sigma}_j^-\rangle + \langle\hat{\sigma}_a^-\hat{\sigma}_b^z\hat{\sigma}_j^+\rangle] - (N-2)\Gamma[\langle\hat{\sigma}_b^+\hat{\sigma}_a^z\hat{\sigma}_j^-\rangle + \langle\hat{\sigma}_b^-\hat{\sigma}_a^z\hat{\sigma}_j^+\rangle] \\ &-2(\Gamma+\gamma_s)\langle\hat{\sigma}_a^z\rangle - 2(\Gamma+\gamma_s)\langle\hat{\sigma}_a^z\hat{\sigma}_b^z\rangle + 4\Gamma\mathrm{Re}[[\langle\hat{\sigma}_a^+\hat{\sigma}_b^-\rangle] + 4\Omega\mathrm{Im}[\langle\sigma_a\sigma_z\rangle b+],\end{aligned} \tag{B.6}$$

$$\frac{d\langle\hat{\sigma}_a^+\hat{\sigma}_b^+\rangle}{dt} = (N-2)(\Gamma-i2\chi)\langle\hat{\sigma}_a^z\hat{\sigma}_b^+\hat{\sigma}_j^+\rangle - (\Gamma+\gamma_s)\langle\hat{\sigma}_a^+\hat{\sigma}_b^+\rangle - i\Omega\langle\sigma_a\sigma_z\rangle b+ + 2i\chi\langle\sigma_a\sigma_+\rangle b+. \tag{B.7}$$

We numerically integrate these cumulant equations to obtain the dynamics of the driven-dissipative system with $\gamma_s > 0$. In Fig. B.1(a) we plot the best squeezing attained after a given threshold timescale (related to the early transient collective squeezing), for a range of χ/Γ and $2\Omega/N\Gamma$ values with $N = 2000$ and $\gamma_s/\Gamma = 20$ Note that minimum values greater than zero have been cut off at zero so as to avoid saturating the color plot. The two phase boundaries Ω_c and

Ω_c' are clearly visible, bracketing the region where single-particle relaxation enhanced squeezing is allowed to develop.

Panels (b)-(e) of Fig. B.1 show a comparison of results from the cumulant expansion along with those from the MCWF method (which is numerically exact in the limit of infinite trajectories, and further discussed in the following section). It can be seen that there is close agreement up to the crossing of the dynamical phase transition, making the result obtained from the cumulant equations a reasonable indicator of the true extent and timing of maximum squeezing.

B.2 Exact Solver

While the cumulant approximation is attractive as it allows us to obtain numerical results rapidly for a large system, there are also other numerical methods to solve the full quantum dynamics of the system in a reasonably efficient manner. In particular, Ref. [184] recently demonstrated that a Monte Carlo wavefunction approach can be implemented to efficiently solve the dissipative dynamics of spin systems exhibiting permutational symmetry for $N \sim \mathcal{O}(10^3)$ and even up to $N \sim \mathcal{O}(10^5)$ in special cases. While we point the interested reader to Ref. [184] for full details of the numerical method, we summarize here the key aspects and advantages.

The MCWF method unravels the density matrix into an ensemble of pure state wave functions that evolve independently of one another in time, where dissipation is handled by random jumps. The full time evolution of one member of this ensemble is referred to as a trajectory. The time evolution of the density matrix is recovered by taking the average of the pure state density matrices at each point in time, resulting in the mixed state solution to the master equation [Eq. (2.10)].

The advantages of this method are three-fold. First, as shown in [184], even though single-particle relaxation breaks the $\hat{J}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2$ symmetry in the master equation, each quantum trajectory lies within a single eigenspace of total spin at any given time. This means that each trajectory can be efficiently integrated, as the dimensionality of the Hilbert space in which it exists is only $\mathcal{O}(N)$. Second, the trajectories evolve in time independently of one another, allowing for the parallel simulation of different trajectories and thus rapid evaluation of ensemble averages. Finally,

analyzing the time evolution of individual trajectories can provide insight that is not altogether obvious from the evolution of the density matrix resulting from the ensemble averages, as was discussed in the main text.

Following the discussion in the main text, in Fig. B.2 we plot an example of the squeezing and effective particle number versus time from the ensemble average and also individual trajectories. We observe that the squeezing improves until enough trajectories approach the critical N^{eff} threshold. After a sufficient number of trajectories cross the threshold and lose their squeezing, the overall squeezed state is then quickly lost. This critical number of trajectories can be reached even before the ensemble average N^{eff} crosses the critical threshold, as is the case in the figure.

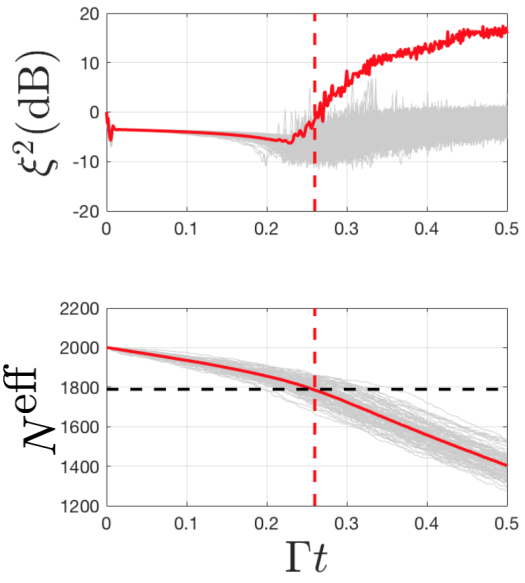


Figure B.2: Spin-squeezing and effective particle number versus time for the full state (red) and a number of individual trajectories (grey) where $N = 2000$, $\Omega/\Gamma = 2000$, $\chi/\Gamma = 1$, and $\gamma_s/\Gamma = 2$. Initial conditions are the coherent spin state in the $-x$-direction. The black horizontal line indicates the critical N^{eff} where $\Upsilon_c^{\text{eff}} = \Upsilon$ and the red vertical line is the point in time where it is crossed

B.3 Technical details of implementation in a cavity-QED platform

In this section we discuss some technical details relevant to a potential realization of our scheme in an optical cavity experiment. Specifically, we discuss some details of the generated

squeezed state which have practical consequences, such as the fact that the dynamically prepared state is phase-squeezed.

B.3.1 Connection between model parameters and cavity co-operativity

As outlined in the manuscript, the spin model we have discussed could be realized by coupling an optical cavity to the narrow linewidth optical clock transitions available in alkaline earth atoms [129, 125]. We require that $\kappa \gg g\sqrt{N}$ and $\kappa \gg \gamma_s$ (bad cavity limit) with $2g$ the single photon Rabi frequency and κ the cavity linewidth, to ensure that the intracavity field can be adiabatically eliminated and thus realize the desired spin model [Eqs. (1) and (2)]. Microscopically, the spin-spin interactions can be engineered by detuning the cavity from the atomic transition by Δ_c which leads to a tunable interaction strength $\chi = 4g^2\Delta_c/(4\Delta_c^2 + \kappa^2)$. Similarly, the collective dissipation arises due to photon leakage and is characterized by $\Gamma = 4g^2\kappa/(4\Delta_c^2 + \kappa^2)$ [129, 130].

It is useful to show that χ and Γ can also be related to the single-particle co-operativity $C = 4g^2/(\kappa\gamma)$ and spontaneous emission rate γ:

$$\chi = \frac{C\gamma\tilde{\Delta}_c}{4\tilde{\Delta}_c^2 + 1}, \tag{B.8}$$

$$\Gamma = \frac{C\gamma}{4\tilde{\Delta}_c^2 + 1}, \tag{B.9}$$

with $\tilde{\Delta}_c = \Delta_c/\kappa$. As the single-particle co-operativity depends on the technical details of the optical cavity, Eqs. (B.8) and (B.9) demonstrate that the collective parameters can be tuned independently of the single-particle spontaneous emission rate γ.

B.3.2 Direction of Spin-Squeezing

Figure B.3 illustrates the direction of the collective Bloch vector (with elevation angle above the equator denoted θ_{el}) and the direction of the minimum variance $(\Delta\hat{J}_{\mathbf{n}_\perp})^2$ in the plane perpendicular to the Bloch vector (denoted ϕ_{min}) as a function of time for the case where $\chi/\Gamma = 0$. The initial state was chosen to be in the $-\hat{z}$-direction. We see that the Bloch vector moves from the south pole toward the positive $\hat{y}$-axis as squeezing builds. The direction of squeezing is parallel to

the equatorial plane throughout the evolution.

B.3.3 Effects of Interaction Strength and Selection of the Initial State

Figure B.4 illustrates how the squeezing generated in the transient dynamics depends on the initial coherent spin state and the presence of elastic interactions. In panel (a), we see that when $\chi/\Gamma = 0$ the selection of the initial state has a minimal impact on the amount of squeezing generated. Furthermore, we see negligible buildup of spin-squeezing in the early dynamics. The only significant squeezing is attained later where we see the interplay between collective and single particle dissipation.

In panel (b), by contrast, we see a buildup in spin-squeezing early in the dynamics when χ/Γ is nonzero; however, the level of squeezing attained is heavily dependent upon the initial state. We also see that certain initial conditions lead to an almost immediate loss of squeezing, behavior consistent with the findings in [11].

B.3.4 Practical utility of the squeezed state

The results of Secs. B.3.2 and B.3.3 for $\chi/\Gamma = 0$ have important practical implications for using the generated squeezing in realistic metrological protocols. Specifically, the important outcome is that an initial coherent spin state prepared along $-\hat{z}$ subject to the dissipative dynamics leads to a **phase**-squeezed state that lies almost perfectly on the equatorial plane. This eliminates the need for any prior characterization of the squeezed state and/or remedial rotations to place the state along the equator in preparation for a Ramsey sequence that can be used to sense rotations about $\hat{z}$. Rotations are a key source of technical error in realistic experiments, particularly those using optical transitions, as small errors can easily couple the anti-squeezed quadrature into the squeezed, quickly eliminating the metrological enhancement of the generated state.

Lastly, we make one further comment on the potential role of interactions in realistic experiments. The most promising candidate for implementation of our protocol is an optical cavity experiment. In this system, one must be mindful that the strength of the interactions χ is controlled

by the detuning Δ_c of the coupled cavity mode from the atomic transition, $\chi = 4g^2\Delta_c/(4\Delta_c^2 + \kappa^2)$. This detuning cannot be rapidly quenched on the relevant time-scales of the dynamics, which means that interactions used to generate the squeezed state would contaminate any subsequent Ramsey sequence to measure rotations. As they cannot be echoed out of the Ramsey protocol, for any practical implementation in an optical cavity one must always use $\chi = 0$.

For completeness, we note that, in contrast, superradiant decay $\propto \Gamma N$ can be effectively turned off independent of the cavity detuning. Specifically, this can be achieved by the following steps [[129]]: i) rapidly dephasing the entangled ensemble with random energy shifts h_j applied via $\hat{H}_{\rm dephase} \propto \sum_j h_j \hat{\sigma}_j^z$, ii) letting the system evolve freely during the interrogation period with $\hat{H}_{\rm rot}$. Note that during that time superradiant emission is suppressed given the incoherent nature of the dephased state and iii) rephasing the ensemble via $-\hat{H}_{\rm dephase}$. As the dephasing $\hat{H}_{\rm dephase}$ commutes with the rotation $\hat{H}_{\rm rot} \propto \hat{J}_z$ it does not degrade the sensitivity of the squeezed state.

However, the $\chi \neq 0$ results studied in this work remain of fundamental interest, not only in the context of understanding the interplay of collective and single-particle effects in the dynamical generation of entanglement, but as it is still relevant for trapped-ion arrays, wherein the spin-spin interactions can be rapidly reversed on relevant experimental time-scales.

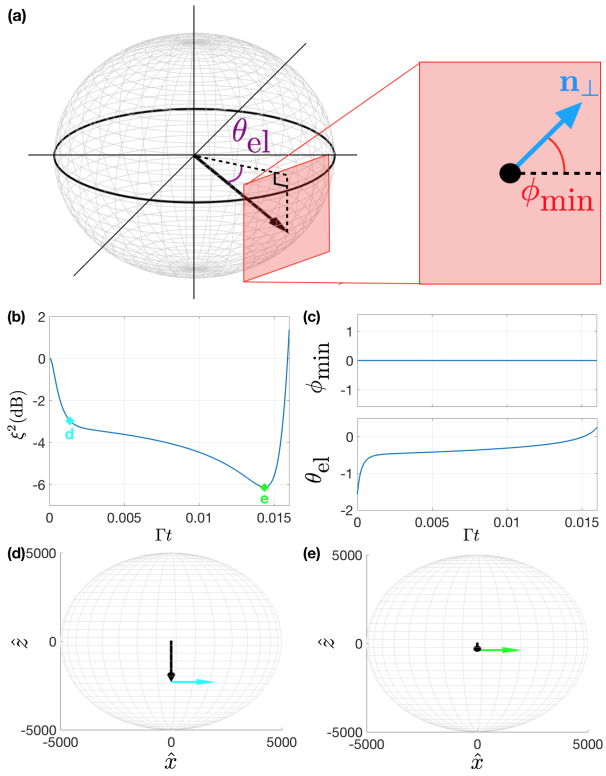


Figure B.3: (a) Diagram illustrating the definition of the angle $\phi_{\rm min}$ of minimum $(\Delta \hat{J}_{\mathbf{n}_\perp})^2$ in the plane perpendicular to the Bloch vector (black arrow) and $\theta_{\rm el}$, the elevation angle of the Bloch vector measured from the equatorial plane. (b) Evolution of spin-squeezing, computed with the cumulant expansion for $N = 10^4$, $\Upsilon = 0.9\Upsilon_c$, $\chi/\Gamma = 0$, and $\gamma_s/\Gamma = 50$. Initial conditions are a coherent spin state in the $-\hat{z}$-direction (c) $\phi_{\rm min}$ (top) and $\theta_{\rm el}$ (bottom) in radians. (d)-(e) The Bloch vector (black) and direction of minimum variance (green)/(blue) at the minimum squeezing in the early/late transient dynamics

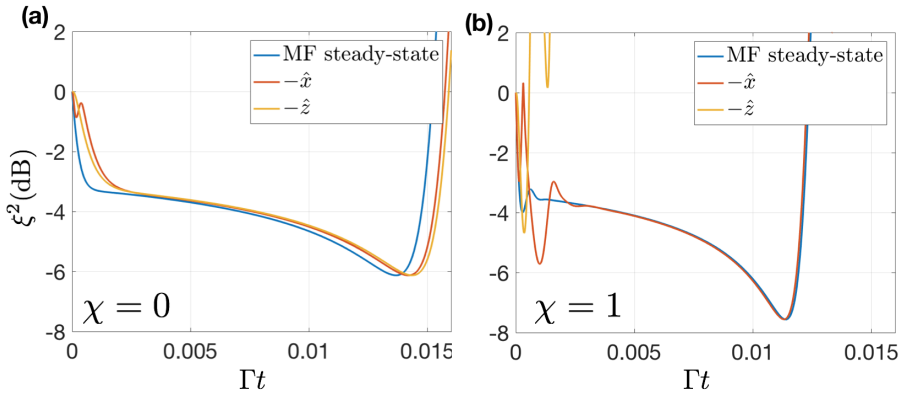


Figure B.4: Spin-squeezing versus time simulated using the cumulant expansion for $N = 10^4$, $\Upsilon = 0.9\Upsilon_c$, and $\gamma_s/\Gamma = 50$ with (a) $\chi/\Gamma = 0$ and (b) $\chi/\Gamma = 1$ for various initial conditions