# ABSTRACT

Title of Dissertation: SPIN GLASS THEORY AND ITS APPLICATIONS TO LEARNING IN CLASSICAL AND QUANTUM NEURAL NETWORKS

Richard Barney
Doctor of Philosophy, 2025

Dissertation Directed by: Professor Victor Galitski
Department of Physics

Spin glasses are, in essence, simply systems of magnetic moments with disordered interactions. Despite this basic premise, spin glasses have become relevant in a wide variety of fields, including condensed matter physics, complexity and optimization theory, error correction, quantum computing, machine learning, and even neuroscience. The pervasiveness of spin glasses stems from the fact that they are paradigmatic examples of how global complexity emerges from disordered interactions between a large number of basic constituents. The inclusion of quantum effects in spin glasses creates an even richer picture, due to the interplay between the energy barriers arising from disorder and quantum fluctuations which allow for tunneling.

This thesis initially focuses on quantum glasses and how their initial states explore their Hilbert spaces. Two complementary approaches are used. One is comparison of the spectral statistics, particularly the spectral form factor, with the spectral statistics of random matrix

ensembles, a hallmark of quantum chaotic behavior. The other approach uses the Thouless-Anderson-Palmer mean-field equations to determine the number of local minima which appear in the free energy landscape.

This thesis then applies methods from spin glass theory to the field of machine learning. Neural networks, like spin glasses, are systems in which remarkable behavior emerges from interactions between basic constituents. By studying the phase diagrams of spin models analogous to snapshots of neural networks throughout their training process, insight is gained into the mechanisms of learning.

Finally, this thesis investigates how quantum effects can lead to advantages in machine learning. It is shown, through classical simulation and experiments on both trapped-ion and superconducting quantum computing hardware, that a natural quantization of a classical neural network can demonstrate superior generalization performance due to the injection of quantum uncertainty.

# SPIN GLASS THEORY AND ITS APPLICATIONS TO LEARNING IN CLASSICAL AND QUANTUM NEURAL NETWORKS

by

Richard Barney

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2025

Advisory Committee:
Professor Victor Galtiski, Chair/Advisor
Professor Jay Sau
Professor Nicole Yunger Halpern
Professor Michael Gullans
Professor Ichiro Takeuchi, Dean's Representative

© Copyright by
Richard Barney
2025

# Acknowledgments

I thank Victor Galitski, my advisor, for his mentorship. I'm grateful for his patience and support, allowing me to investigate topics I personally found interesting, and his guidance when I came up against obstacles.

I thank Brian Swingle for sharing his enthusiasm for physics, which was enormously helpful to me when I was first starting my research experience in graduate school.

I'm grateful for the mentorship of two people who I first knew as postdoctoral researchers at the University of Maryland: Christopher L. Baldwin and Yunxiang Liao. Chris did the hard work of getting me up to speed when I was first starting out and Yunxiang was always an invaluable resource and collaborator.

I also thank all the people I knew through the Galitski research group. These include Andrey Grankin, Amit Vikram Anand, Alireza Parhizkar, Gautam Nambiar, Abu Saleh Musa Patoary, Masoud Mohammadi-Arzanagh, Jeet Shah, Laura Shou, Preston Ohanuka, Jeremy Shuler, Rohan Panchwagh, Paul Argyle, Aranya Chakraborty, and especially my frequent coauthor Michael Winer. The discussions we had together were invaluable.

I thank all the wonderful mentors I had at Brigham Young University, including Manuel Berrondo, Mark Transtrum, and especially my advisor Jean-François S. Van Huele. They all created a truly welcoming environment that fostered my talent and creativity.

I thank the wonderful people of the College Park Ward for giving me a spiritual home and

community while I carried out these studies. I especially am grateful to Clark Ridge for his leadership, example, and invitations to serve others.

I thank my parents, Sarah and Dean Barney, for their constant support throughout my life and for teaching me how to do hard things.

Finally, I thank my wife Delaney for her unwavering belief in me and for all the energy she has poured into helping me accomplish my goals.

## Summary of Research Contributions

The work contained within this thesis was conducted by the present author in collaboration with Victor Galitski, Brian Swingle, Norbert M. Linke, Alaina M. Green, Sarah H. Miller, Christopher L. Baldwin, Yunxiang Liao, Michael Winer, Djamil Lakhdar-Hamina, and Xinxing Liu. This work originally appeared as Refs. [1–6]. In compliance with the guidelines of the University of Maryland and of the Physics Department's Graduate Director, Chapters 2-7 and Appendices A-C of this thesis reproduce these papers.

Below is the report of the present author's contribution to each chapter of this thesis.

Chapter 1 contains introductory material. It is solely the work of the present author.

Chapter 2 and Appendix A originally appeared as Ref. [1]. The calculations and writing were primarily done by the present author. Michael Winer and Christopher L. Baldwin verified the calculations. All authors, including Brian Swingle and Victor Galitski, contributed to the revision of the text.

Chapter 3 and Appendix B originally appeared as Ref. [2]. Christopher L. Baldwin and the present author assisted Michael Winer in the calculations and writing. All authors, including Brian Swingle and Victor Galitksi, contributed to the revision of the text.

Chapter 4 and Appendix C originally appeared as Ref. [3]. The present author contributed the construction of the model, numerical simulations, and the calculation of the SFF. Yunxiang Liao contributed the calculation of the evolution of the disorder averaged density matrix. The

present author and Yunxiang Liao both contributed to the writing. Victor Galitski provided the initial concept of a glassy quantum circuit. All authors contributed to the revision of the text.

Chapter 5 originally appeared as Ref. [4]. Michael Winer contributed the replica analysis of the multilayer Sherrington-Kirkpatrick model. The present author performed the TAP analysis, numerical simulations, data analysis, and the writing. Victor Galitski contributed the initial concept of spin model analogues of neural networks. All authors contributed to the revision of the text.

Chapter 6 originally appeared as Ref. [5]. The present author contributed the construction of the model, the coding and training of the model, and all calculations. Djamil Lakhdar-Hamina provided assistance with the code. All authors, including Victor Galitski, contributed to the revision of the text.

Chapter 7 originally appeared as Ref. [6]. Djamil Lakhdar-Hamina collected experimental data from IBM quantum machines. Xinxing Liu and Alaina M. Green collected experimental data from the Gates Lab trapped ion experiment. The present author provided the construction of the model architecture, the classical training of the model, and theoretical support. Djamil Lakhdar-Hamina, Xinxing Liu, and the present author contributed to the writing. Norbert M. Linke, Sarah H. Miller, and Victor Galitski contributed to the conceptualization and interpretation of the experiments. Victor Galitski also secured the credits from IBM to run the experiment on their platform. All authors contributed to the revision of the text.

Chapter 8 contains a summary of the findings reported in this thesis and a discussion of future directions. It is solely the work of the present author.

# Table of Contents

# List of Tables

# List of Figures

# List of Abbreviations

| | |
|---|---|
| 1RSB | One-step Replica Symmetry Breaking |
| BGS | Bohigas-Giannoni-Schmit |
| BQNN | Benchmark Quantum Neural Network |
| BRP | Block Rosenzweig-Porter |
| COE | Circular Orthogonal Ensemble |
| CUE | Circular Unitary Ensemble |
| CSE | Circular Symplectic Ensemble |
| EA | Edwards-Anderson |
| ETH | Eigenstate Thermalization Hypothesis |
| GOE | Gaussian Orthogonal Ensemble |
| GUE | Gaussian Unitary Ensemble |
| GSE | Gaussian Symplectic Ensemble |
| iid | independently and identically distributed |
| JPDF | Joint Probability Density Function |
| MBL | Many-Body Localization |
| MLP | Multilayer Perceptron |
| MNIST | Modified National Institute of Standards and Technology |
| MSK | Multilayer Sherrington-Kirkpatrick |
| NN | Neural Network |
| OTOC | Out-of-Time Ordered Correlator |
| PSM | $p$-spin Spherical Model |
| QASM | Quantum Assembly Simulation |
| QNN | Quantum Neural Network |

RP Rosenzweig-Porter
RSB Replica Symmetry Breaking

SFF Spectral Form Factor
SGD Stochastic Gradient Descent
SK Sherrington-Kirkpatrick

TAP Thouless-Anderson-Palmer

WD Wigner-Dyson

# Chapter 1: Introduction

## 1.1 Motivation

Spin glasses have proven to be relevant to an impressively large number of fields. These include, but are not limited to, condensed matter physics, statistical mechanics, complexity and optimization theory, machine learning, error correction, quantum computing, and even neuroscience. This ubiquity is due to the fact that spin glasses are paradigmatic examples of complex systems that arise from the disordered interactions of a large number of very basic constituents. A great number of real-world complex systems that are of practical interest share at least qualitative similarities with a spin glass model. So, the study of spin glasses is essentially the study of emergent complexity.

The inclusion of quantum effects allows for even richer phenomena to emerge due to the interplay between barriers in phase space due to disorder and quantum fluctuations which allow for quantum tunneling through such barriers. As nature is fundamentally quantum, understanding this interplay is of fundamental importance. Additionally, these studies are of practical interest as efforts continue to ramp up to build technologies that take advantage of quantum resources. A main theme of this thesis is understanding the ways in which quantum glasses explore their Hilbert spaces, or fail to do so. This will be explored through the lenses of spectral statistics and the structure of glassy energy landscapes.

Recently there have been many advances in the realm of artificial intelligence, which has shown capabilities that were once thought to require human intelligence. These advances have been underpinned by neural network technologies. As with spin glasses, these networks are comprised of very basic elements, the neurons. The remarkable abilities of these models are due to the way these neurons are trained to interact with each other. The other theme of this thesis is applying techniques from spin glass theory to understand the learning dynamics of both classical and quantum NNs.

In the remainder of this introduction I present primers on the subjects of classical and quantum spin glasses, classical and quantum chaos, spectral statistics within random matrix theory, and the construction and training of neural networks.

## 1.2 Spin Glasses

### 1.2.1 Classical Spin Glasses

At the most basic level, spin glasses are systems of magnetic moments with disordered interactions. Despite this apparent simplicity, spin glasses are a fascinating testbed for exploring the phenomena that emerge from large-scale quenched disorder and frustration. Frustration is when a loop exists in a system such that the product of the interactions along the loop is negative. Conceptually, this means that there is no configuration along the loop that satisfies all the interactions.

For example, suppose that we have three Ising spins arranged in a triangle, as shown in Fig. 1.1. Each of these spins can be in either the "up" $(+1)$ or "down" $(-1)$ state. Between each pair of spins is some interaction. If this interaction is positive, i.e. ferromagnetic, the pair will

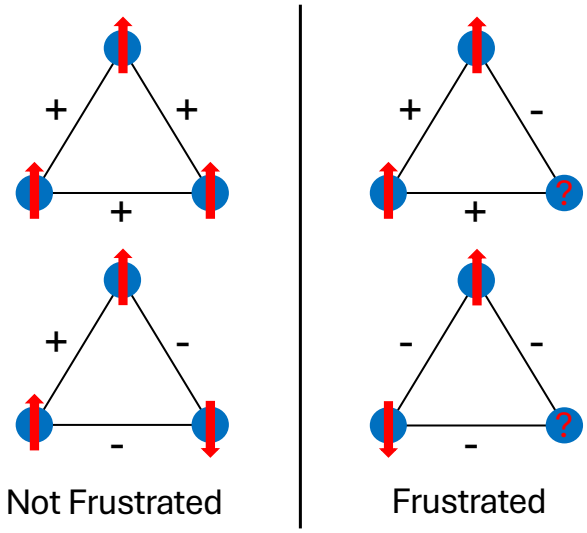



Figure 1.1: (Left) When the product of the interactions along the loop is positive, no frustration occurs. (Right) When the product of the interactions along the loop is negative, frustration does occur.

want to align at low temperature. Conversely, if the interaction is negative, i.e. antiferromagnetic, the pair will want to be in opposite states. This simple system can be described by the Hamiltonian

$$H = -\frac{1}{2}\sum_{i \neq j} J_{ij} s_i s_j, \tag{1.1}$$

where the $s_i$ are the spin states and the $J_{ij}$ are the interactions. If there are an even number of antiferromagnetic interactions, as in the left of Fig. 1.1, each of the interactions may be satisfied and the system has a single ground state, putting aside the $\mathbb{Z}_2$ symmetry of the system by which we see that flipping every spin simultaneously does not change the energy. On the other hand, if there is an odd number of antiferromagnetic interactions, as in the right of Fig. 1.1, it is impossible to find a configuration that satisfies all the interactions. This makes the ground state nontrivial. In fact, if the interactions all have the same magnitude, the ground state becomes three-fold degenerate.

While frustration is not too difficult to handle in small systems, consider what happens when we have a thermodynamically large number of spins described by the Hamiltonian

$$H = -\sum_{(i_1,\ldots,i_p)} J_{i_1\ldots i_p}\sigma_{i_1}\cdots\sigma_{i_p}, \tag{1.2}$$

where the $\sigma_{i_k}$ are some variety of spin variables, not necessarily Ising, and the sum is over all possible sets of $p$ spins. The interactions are drawn from some probability distribution such that each $J_{i_1\ldots i_p}$ may be positive or negative. Since the interactions do not evolve in time or fluctuate with the spins, this is an example of quenched (as opposed to annealed) disorder. Such a system may have high levels of frustration, leading to a very rugged energy landscape in its configuration space. Such a landscape is shown schematically in Fig. 1.2. Instead of having a single minimum, as in the case of an ordered state, it may have many local minima about which the dynamics can be constrained for a time.

Here is where the connection to structural glasses is apparent, explaining the "glass" part of spin glasses. Structural glasses are in amorphous configurations at temperatures below their melting point. This means that, while some of the atoms of the glass are jammed into place, others are somewhat free to move about. This partitions the configuration space into sectors separated by energy barriers which can constrain the dynamics.

Eq. (1.2) may be considered the general formulation of a classical spin glass model. The most commonly studied models are special cases of this construction. A few are listed here:

- The Edwards-Anderson Model [7]:

  This model is defined on a $d$-dimensional lattice, with interactions only between nearest

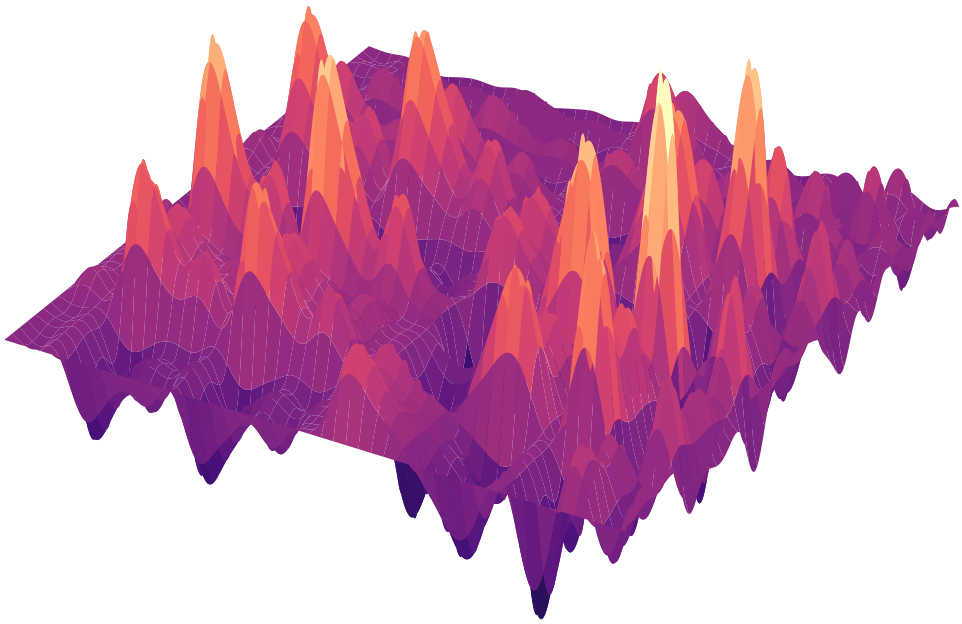


Figure 1.2: A schematic representation of a rugged energy landscape, such as those that exist for spin glass models.

neighbors. The Hamiltonian for this model may be written more simply as

$$H = -\sum_{\langle i,j \rangle} J_{ij} s_i s_j, \tag{1.3}$$

where $\langle i, j \rangle$ indicates the sum over all nearest neighbors. The interactions are typically drawn from a Gaussian or binary distribution. This model is the standard for studying spin glasses in finite dimensions. Its spatial structure makes it more physically realistic than the following models, but less tractable analytically.

- The $p$-spin Model [8, 9]:

  This is the prototypical model for studying mean-field spin glasses. For $p \geq 3$ it provides a clean setting for probing one-step replica symmetry breaking (1RSB). The interactions are

taken to be Gaussian with variance

$$\text{Var}(J_{i_1 \ldots i_p}) = \frac{p!}{2N^{p-1}} \tag{1.4}$$

to ensure that the energy scales extensively. A particularly tractable variation is the $p$-spin spherical model (PSM) [10, 11], in which the spins are treated as continuous "position" variables, while the spin-like nature of the model is somewhat maintained by enforcing the condition

$$\sum_{i=1}^{N} s_i^2 = N. \tag{1.5}$$

This means that the system is constrained to a hypersphere in the configuration space. Calculating the spectral form factor for a quantized version of this model is the subject of Chapter 3.

- The Sherrington-Kirkpatrick Model [12]:

  This model may be considered the restriction to $p = 2$ of the $p$-spin model above. Its Hamiltonian is

$$H = -\sum_{i<j} J_{ij} s_i s_j. \tag{1.6}$$

  This is the most commonly studied mean-field spin glass model and is exactly solvable in the thermodynamic limit. It is also the setting in which Parisi's replica method was first applied [13]. Interactions are taken to be Gaussian with a variance on the order of $N^{-1}$ to ensure the energy scales extensively. A variation of this model, termed the multilayer Sherrington-Kirkpatrick (MSK) model, is studied in Chapter 6. In this model, spins are organized into layers and interactions only exist between spins in adjacent layers. This spin

glass model is found to provide useful insights into the learning process in feed-forward neural networks.

- The Random Energy Model [8, 14]:

  This model focuses on the rugged energy landscapes that exist for glass models. In a system of $N$ Ising spins, independent and identically distributed (iid) Gaussian random energies are assigned to each of the $2^N$ possible configurations. Despite the fact that energy levels are not independent in most glass models, the phase diagram of the random energy model shares many qualitative features with more-physical spin glass models. Additionally, the $p$-spin model becomes the random energy model in the limit $p \to \infty$.

### 1.2.2 Quantum Spin Glasses

At this point, we have only discussed classical spin glasses, which are a rich field in themselves. Quantum spin glasses support even richer phenomena and are therefore of great interest as well. Quantum fluctuations act somewhat in opposition of the energy barriers that arise from the system's disorder, as the fluctuations introduce the possibility of tunneling through these barriers. The quantum models explored in this thesis are straightforward generalizations of classical models described above. For example, Chapter 3 quantizes the PSM by defining conjugate "momentum" operators $\pi_i$ such that the commutation relationship with the spin operators is

$$[\sigma_i, \pi_j] = i\delta_{ij}. \tag{1.7}$$

The quantum Hamiltonian is then generalized from Eq. (1.2) by adding a kinetic energy term such that

$$H = \sum_i \frac{\pi_i^2}{2\mu} - \sum_{(i_1,\ldots,i_p)} J_{i_1\ldots i_p}\sigma_{i_1}\cdots\sigma_{i_p}, \tag{1.8}$$

where $\mu$ is a parameter that controls the strength of the quantum fluctuations. In the same spirit, classical spin models may be quantized by promoting the classical spins to Pauli operators and adding terms to the Hamiltonian containing non-commuting Pauli operators.

### 1.2.3 Thermalization Behaviors

Chaotic systems thermalize quickly. That is, they quickly exhibit ergodic behavior across their Hilbert spaces and the expectation value of a generic observable will relax to its thermal value in a short amount of time. This is not the case for integrable systems. In such systems thermalization is prevented by an extensive number of conserved quantities, which lead to the retention of information about the system's initial state.

Spin glasses exist as a third option between these two extremes. Due to their high degree of frustration their energy landscapes can be very rugged, leading to the formation of a large number of local minima. Metastable states can form in the neighborhoods of these minima and can possibly last for times exponentially long in the system size. Thermalization then occurs at an extremely slow rate due to rare large thermal fluctuations or tunneling between local minima in the quantum case. This results in rich thermalization behaviors such as slow dynamics, multi-stage relaxation, and aging, in which older systems respond to perturbations more slowly [15–18].

### 1.2.4 The Replica Approach

To calculate the disorder averaged thermal expectation value of some observable $A$ we need to calculate

$$\overline{\langle A \rangle} = \overline{Z^{-1} \operatorname{tr} A(\sigma) e^{-\beta H(\sigma, J)}}, \tag{1.9}$$

where $\langle \cdot \rangle$ indicates a thermal expectation, $\bar{\cdot}$ indicates a disorder average over $J$, and $Z$ is the system's partition function. However, the disorder average is often intractable in this form due to the fact that $Z$ and $H$ are both dependent on $J$.

A useful approach is to make use of the replica "trick." That is, we use the fact that

$$\overline{\log Z} = \lim_{n \to 0} \frac{1}{n} \log \overline{Z^n}. \tag{1.10}$$

In physical terms, we create $n$ exact replicas of the system in question. We can then disorder average this replicated partition function and later take the limit of 0 replicas. This allows us to rewrite Eq. (1.9) in the much more approachable form

$$\overline{\langle A \rangle} = \lim_{n \to 0} \overline{Z^{n-1} \operatorname{tr} A(\sigma) e^{-\beta H(\sigma, J)}} \tag{1.11}$$

$$= \lim_{n \to 0} \operatorname{tr} A(\sigma_1) \overline{e^{-\beta H(\sigma_1, J) \cdots - \beta H(\sigma_n, J)}}, \tag{1.12}$$

where the trace in the last equality is now over all replicas.

Some mathematical peculiarities remain with this method. It's not clear how to interpret the limit of 0 replicas in a physical way. Additionally, we are often most interested in disorder averaged thermal expectation values of systems in the thermodynamic limit, where the number

of spins goes to infinity. However, such replica calculations often make the implicit assumption that

$$\lim_{N\to\infty}\lim_{n\to 0}\frac{1}{n}\log\overline{Z^n} = \lim_{n\to 0}\lim_{N\to\infty}\frac{1}{n}\log\overline{Z^n}. \tag{1.13}$$

Despite these irregularities, the replica method has been found to give correct answers across a wide variety of models and was put on more rigorous mathematical footing by, e.g. Refs. [19, 20].

A great advantage of the replica method is that it allows us determine the glass transition by observing when the replica symmetric solution becomes unstable. After disorder averaging a replicated partition function, the overlap matrix

$$Q^{\alpha\beta} = N^{-1}\sum_i \sigma_i^\alpha \sigma_i^\beta \tag{1.14}$$

arises naturally. $\alpha$ and $\beta$ are replica indices in this quantity, so $Q^{\alpha\beta}$ is a measure of how similar the states of two replicas are.

At sufficiently high temperatures, in which we justifiably expect to see ergodic paramagnetic behavior, we find that the stable saddle point of the free energy is when the overlap matrix takes replica symmetric form. That is

$$Q^{\alpha\beta} = q_0 + (1-q_0)\delta^{\alpha\beta}, \tag{1.15}$$

where $q_0 < 1$. This makes sense for an ergodic phase. Due to the chaotic behavior, different replicas will have small overlaps with each other.

However, when temperature drops below the glass transition, we find that the replica sym-

metric solution of the saddle point equations becomes unstable. The new stable solutions are ones in which the overlap matrix takes on a richer form. For example, for a model with 1RSB, such as the $p$-spin model, the overlap matrix will take the block form

$$Q = \begin{pmatrix} Q_d & Q_0 & Q_0 & \cdots \\ Q_0 & Q_d & Q_0 & \cdots \\ Q_0 & Q_0 & Q_d & \\ \vdots & \vdots & & \ddots \end{pmatrix}, \tag{1.16}$$

$$Q_d^{\alpha\beta} = q_1 + (1 - q_1)\delta^{\alpha\beta}, \tag{1.17}$$

$$Q_0^{\alpha\beta} = q_0, \tag{1.18}$$

$$q_0 < q_1 < 1. \tag{1.19}$$

This physical interpretation is thus: in a glassy phase, different replicas will settle into different minima of the rugged energy landscape. Some of these will cluster together, forming a block with large values in the overlap matrix, while others are far apart.

Some models will have higher degrees of RSB, leading to more hierarchical block levels in the overlap matrix. Some models, such as the SK model, even have infinite levels of RSB. In these cases, clusters can form superclusters, which can form even higher-level clusters, and so on. The structure of states in these models is ultrametric, meaning that, for each replica $\gamma$, it holds that

$$Q^{\alpha\beta} \geq \min(Q^{\alpha\gamma}, Q^{\beta\gamma}). \tag{1.20}$$

This is essentially saying that the replicas cluster in a hierarchical fashion.

Note that RSB solutions can appear at temperatures higher than that at which the replica symmetric solution becomes unstable. In fact, there exist discontinuous glasses for which this is the case [21, 22]. For example, in a system with a 1RSB transition, the parameter $q_1$ in Eq. (1.17) might have a discontinuous transition when the replica symmetric solution becomes unstable. This indicates that the states that emerge at this transition are already tight. This is an indication that these states already existed as metastable states above the transition. For this reason, we should understand the transition obtained from the replica method as a static transition, due to the thermodynamic nature of the calculation, which is concerned with equilibrium properties. Other methods, such as the Thouless-Anderson-Palmer approach discussed in the next section, can tell us more about the possibility of a dynamic transition away from the paramagnetic phase at some temperature above the static transition temperature. Below this dynamic transition the system may get trapped in metastable states for long times without necessarily changing the system's equilibrium properties.

### 1.2.5 The Thouless-Anderson-Palmer Equations

Although we have invoked energy landscapes often in the discussion thus far, we have not yet made the concept concrete. While the replica method discussed above is powerful, especially in dealing with disorder averages, it is through the complementary approach developed by Thouless, Anderson, and Palmer (TAP) [23] that we are able to actually get a grasp on this landscape. An additional benefit to this approach is that much progress can be made for single realizations of disorder, instead of immediately averaging disorder out. This makes the TAP approach very valuable for problems in which the distribution of the interactions is not known, such as analogues of

trained neural networks, as examined in Chapter 6.

The TAP equations are a set of self-consistent equations that arise from a mean-field description of a system's free energy. This TAP free energy is most generally defined as a Legendre transformation of the free energy with respect to some order parameters. In the classical case one performs a Legendre transform with respect to the magnetizations $m_i$. In this case, the TAP free energy is

$$F_{\text{TAP,cl}} = -\beta^{-1} \log \operatorname{tr} \exp \left( -\beta H + \beta \sum_i h_i (\sigma_i - m_i) \right), \tag{1.21}$$

where $\beta$ is the inverse temperature and the local fields $h_i$ are Lagrange multipliers that enforce the constraint $\langle \sigma_i \rangle = m_i$. In the quantum case, quantum fluctuations can be understood as fluctuations along an imaginary time dimension by taking the continuous limit of a Suzuki-Trotter transformation [24]. Autocorrelations in imaginary time must also be addressed in order to describe the metastable properties, and so we must also Legendre transform with respect to these. So, in the quantum case, the TAP free energy is

$$\begin{aligned} F_{\text{TAP,quant}} = -\beta^{-1} \log \operatorname{tr} \exp \Bigg[ & - \int_0^\beta d\tau H + \int_0^\beta d\tau \sum_i h_i(\tau) \left( \sigma_i(\tau) - m_i(\tau) \right) \\ & + \int_0^\beta d\tau \int_0^\beta d\tau' \sum_i \Lambda(\tau, \tau') (C(\tau, \tau') - \sigma_i(\tau) \sigma_i(\tau')) \Bigg], \end{aligned} \tag{1.22}$$

where $\Lambda(\tau, \tau')$ is a Lagrange multiplier enforcing the constraint on the autocorrelation $\langle \sigma_i(\tau) \sigma_i(\tau') \rangle = C(\tau, \tau')$. These TAP free energies can also be obtained through a cavity calculation [25].

With this description of the free energy landscape we can define TAP states, which are simply the neighborhoods around the local minima of the TAP free energy. This size of these neighborhoods depends on the temperature. It is these TAP states which describe the metastable

states we have discussed previously. We can then decompose the partition function as

$$Z = \operatorname{tr} e^{-\beta H(\sigma)} = \sum_a Z_a, \tag{1.23}$$

$$Z_a = \operatorname{tr} \delta_{\sigma \in a} e^{-\beta H(\sigma)}, \tag{1.24}$$

where $a$ labels a TAP state.

Within the TAP formalism it is straightforward to determine the static transition away from the paramagnetic phase. It occurs when the paramagnetic solution $m_i = 0$ becomes unstable. This is the same static transition that might also be obtained through the replica approach described previously. However, in addition to the static transition, one may also determine the dynamical transition by finding the point at which additional local minima of the TAP free energy appear.

For discontinuous spin glass models, such a dynamical transition may be found above the static transition. In this case an exponentially large in system size number of such metastable states appear, though the result for the free energy is the same as would be expected if the system were still in its paramagnetic phase [21].

The number of free energy minima that emerge at a transition contains information about the nature of that transition. A glassy transition will have an exponentially large in the system size number of such minima appear, while a transition to an ordered phase, such as a ferromagnet will have just one. This fact will be useful in our analysis of spin model analogues of neural networks in Chapter 6.

## 1.3 Chaos Theory

### 1.3.1 Classical Chaos

For a time after the formulation of Newtonian mechanics and its refinements through Lagrangian and Hamiltonian formalisms, it was believed that, having a set of measurements fully describing any system at some moment in time, the state of that system could, in principle, be calculated at any other moment in time. However, Poincaré's seminal work on the three-body problem in celestial mechanics demonstrated that this is not always the case. Although the three-body problem follows deterministic Newtonian mechanics, Poincaré discovered that it does not admit a closed-form solution [26]. This insight is often considered the beginning of the field of chaos theory.

Chaos theory took on its modern aspect due to the work of Lorenz concerning computer simulations of atmospheric models [27]. Unintentionally, he observed that rounding the values defining the initial conditions of a system led to extremely different dynamics. Due to a scientific address by Lorenz [28], this extreme sensitivity to initial conditions become colloquially known as the "butterfly effect." It is now known that this chaotic behavior is more the rule than the exception in physically significant models.

The butterfly effect is made mathematically meaningful through the definition of Lyapunov exponents. Suppose we have a model that we evolve from two very close initial conditions: $\mathbf{x}(0)$ and $\mathbf{x}'(0) = \mathbf{x}(0) + \delta\mathbf{x}(0)$, where $\mathbf{x}(t)$ and $\mathbf{x}'(t)$ are vectors of coordinates in phase space and $\delta\mathbf{x}(0)$ is very small. If the model is chaotic, we will observe that, at short times, the two

trajectories will diverge exponentially. That is

$$|\delta \mathbf{x}(t)| = |\mathbf{x}'(t) - \mathbf{x}(t)| \sim e^{\lambda t} |\delta \mathbf{x}(0)| \tag{1.25}$$

for some Lyapunov exponent $\lambda > 0$. Formally, this exponent is defined in terms of the Jacobian matrix $M$ as

$$\lambda = \lim_{t \to \infty} t^{-1} \log ||M||_2, \tag{1.26}$$

$$M_{ij} = \frac{\partial x_i(t)}{\partial x_j(0)}, \tag{1.27}$$

where $||M||_2$ is the spectral norm of the Jacobian.

However, a positive Lyapunov exponent is not, by itself, evidence of chaotic behavior. In fact, it is easy to create a mapping with exponential sensitivity to initial conditions, but no chaotic behavior. For example

$$x_n = a x_{n-1}, \quad a > 0. \tag{1.28}$$

This map has a positive Lyapunov exponent, but its behavior is extremely simple. All trajectories, except the one beginning at 0, diverge to positive or negative infinity.

An additional aspect of chaos that is extremely important in the context of statistical mechanics is that of topological mixing. Formally, a mapping $f$ is topologically mixing if, given open sets $A$ and $B$ in phase space, there exists an integer $N$ such that

$$f^n(A) \cap B \neq \emptyset \tag{1.29}$$

for all $n > N$. Though this is defined in terms of discrete mappings, it can easily be applied to continuous-time systems through a stroboscopic approach. Physically, this indicates that, after a certain amount of time has elapsed, a number of points initially confined to a part of the phase space will spread across the entire phase space.

This is closely related to the concept of ergodicity, which is when each point in phase space will eventually visit each other point. Technically, ergodicity does not necessarily imply mixing. Consider, for example, a particle moving in a 2-dimensional periodic space at a constant velocity, where the velocity components along the two dimensions are incommensurate. The particle will eventually visit every point in the space, but this does not mean many such particles starting at different positions will mix together, since they are all moving in parallel. Despite the difference in definition, mixing and ergodicity are most often found together in physical systems, and so the two terms are often used interchangeably, including in this thesis.

Mixing is extremely important in statistical mechanics as it bears on the process of thermalization. For a many-body system, regardless of what the initial configuration of the particles is, we can assume that, after some period of time, the system will thermalize. The distribution of states will be uniform over the region of phase space which is permitted by conservation laws. This allows for the calculation of the expectation values of macroscopic observables that would be intractable on the microscopic level. For example, suppose we want to find the pressure inside a balloon filled with some gas. Due to the absolutely enormous number of atoms in the gas we cannot measure the position and velocity of each atom at an instant of time, nor can we carry through the calculations to determine their later states. Thanks to mixing and thermalization, neither are necessary, as we can just average over the appropriate region of phase space.

### 1.3.2 Quantum Chaos

At first, it does not seem that chaotic behavior should exist at all in quantum systems. After all, quantum wavefunctions are governed by the Schrödinger equation

$$i\partial_t \psi = H\psi, \tag{1.30}$$

which is linear. Since the Hamiltonian is Hermitian, the time evolution operator

$$U(t) = e^{iHt} \tag{1.31}$$

is unitary, meaning distances in phase space are always preserved. Additionally, due to Heisenberg uncertainty, the concept of trajectories we used in defining classical chaos becomes ill-defined.

Some progress can be made by considering a semiclassical quantization of a classical chaotic system. Take one wavefunction $\psi(0)$ which is sharply peaked at $\mathbf{x}(0)$ and another wavefunction $\psi'(0)$ which is sharply peaked at $\mathbf{x}(0) + \delta\mathbf{x}(0)$, where $\mathbf{x}$ is a vector of coordinates. As the system evolves, classical chaos theory tells us that

$$|\langle\psi(t)|\mathbf{x}|\psi(t)\rangle - \langle\psi'(t)|\mathbf{x}|\psi'(t)\rangle| \sim e^{\lambda t}|\delta\mathbf{x}(0)|. \tag{1.32}$$

This is despite the fact that $\langle\psi(t)|\psi'(t)\rangle$ does not change. This indicates that quantum chaos should be understood through operators, rather than the evolution of wavefunctions.

One approach to quantum chaotic behavior is to determine the growth of out-of-time or-

dered correlators (OTOCs) like

$$\langle [V(t), W(0)]^2 \rangle \sim e^{2\lambda t}, \tag{1.33}$$

where $V$ and $W$ are some local operators and $\langle \cdot \rangle$ indicates a thermal average. In this way a quantum Lyapunov exponent is defined. We can see in this definition a close relationship with the classical definition of the Lyapunov exponent in Eq. (1.26), where the partial derivatives can be written as Poisson brackets for a Hamiltonian system, which are then promoted to commutators when quantizing.

A very important concept in quantum chaos is the Eigenstate Thermalization Hypothesis (ETH) [29–31], which is a way of explaining how an isolated quantum system which is initially prepared in a state that is far from equilibrium can move to a state that appears to be in thermal equilibrium, making it possible to describe the system using statistical mechanics. That this happens at all is somewhat surprising since quantum systems do not evolve in time to uniformly sample all states in phase space with the same energy.

ETH shows that this self-thermalizing behavior can be explained when the matrix elements of physical observables in the energy eigenbasis take the form

$$O_{ab} = \langle E_a | O | E_b \rangle \approx O_{\text{th}}(\bar{E})\delta_{ab} + e^{-S(\bar{E})/2} f(\bar{E}, \omega) R_{ab}, \tag{1.34}$$

$$\bar{E} = \frac{1}{2}(E_a + E_b), \tag{1.35}$$

$$\omega = E_a - E_b, \tag{1.36}$$

where $O_{\text{th}}(\bar{E})$, and $f(\bar{E}, \omega)$ are smooth functions, $S(E)$ is the entropy at energy $E$, and $R$ is a matrix of complex standard normal random variables satisfying $R_{ab} = R^*_{ba}$. We then find that the

long-time average of the expectation value is

$$\overline{\langle O \rangle} = \lim_{T\to\infty} \frac{1}{T} \int_0^T dt \langle \psi(t)|O|\psi(t)\rangle \tag{1.37}$$

$$= \lim_{T\to\infty} \frac{1}{T} \int_0^T dt \left( \sum_{a,b} c_a^* c_b O_{ab} e^{-i\omega t} \right) \tag{1.38}$$

$$= \sum_a |c_a|^2 O_{aa} + i \lim_{T\to\infty} \frac{1}{T} \left( \sum_{a\neq b} \frac{c_a^* c_b O_{ab}}{\omega} \left( e^{-i\omega T} - 1 \right) \right) \tag{1.39}$$

$$\approx O_{\text{th}}(E). \tag{1.40}$$

In the last line we used the fact that the diagonal elements of the operator matrix are approximately constant over a small energy window. In this way, we see how the microcanonical thermal expectation is reached.

ETH has been found to hold for typical quantum systems. It has been found to not hold for many-body localized systems, integrable systems, and, of particular interest for this thesis, glassy systems. A great advantage of ETH is that it can distinguish between systems that thermalize and those which do not by examining statistical properties of their energy eigenstates. That is, when quantum chaos is present, the energy eigenstates behave like random superpositions of basis states.

This is closely connected with the Bohigas-Giannoni-Schmit (BGS) conjecture [32], which focuses on the energy spectrum rather than energy eigenstate statistics. The BGS conjecture grew out of studies of quantized chaotic billiard problems. For a particular billiard one may calculate the energy spectrum, then define the normalized level spacings

$$s_n = \frac{E_{n+1} - E_n}{N^{-1} \sum_{n=0}^{N-1} (E_{n+1} - E_n)}, \tag{1.41}$$

where there are $N$ levels in the spectrum. BGS found that, for billiards which are chaotic in the classical limit, the distribution of the level spacings followed a Wigner-Dyson (WD) distribution, which is the distribution expected for Gaussian random matrices. In contrast, for integrable systems, the distribution of level spacings is found to be Poissonian. From these observations come the (paraphrased, and somewhat generalized) BGS conjecture: quantum systems which are chaotic in their classical limit exhibit spectral fluctuations consistent with those predicted by random matrix theory.

The ETH and BGS conjectures both describe hallmarks of quantum chaos in relation to aspects of random matrix theory. These ideas have been further extended to quantum systems without classical limits. In fact, random-matrix-like behavior has come to be considered by many as the defining hallmark of quantum chaos. The next section is devoted to providing more details on what is meant by random-matrix-like behavior.

## 1.4 Spectral Statistics in Random Matrix Theory

### 1.4.1 The Gaussian Ensembles

Though random matrices have long been of interest to the field of statistics, their connection to physics was first appreciated in the context of the spectra of heavy nuclei. In the 1950's, though a great deal of experimental data had been collected about these spectra, they did not appear to follow any mathematical form, instead appearing to be random. Wigner first had the idea to model the spectra using random matrices [33, 34]. He found that the energy level spacings of heavy nuclei very closely matched the spacing distribution of a real symmetric random matrix. This is now known as one of the standard Wigner-Dyson (WD) distributions. In the years since

Wigner's insight, the analogy to random matrices has found application to a huge variety of quantum chaotic systems.

Wigner outlined three random matrix ensembles of particular physical relevance, called the Gaussian random matrix ensembles. These are summarized below. Each is denoted by its Dyson index $\beta$, which is the number of real variables needed to define an entry of the matrix. This number plays a role analogous to inverse temperature.

- Gaussian Orthogonal Ensemble (GOE), $\beta = 1$:

  GOE matrices are real, symmetric, $N \times N$ matrices. A GOE matrix $H$ may be defined as

  $$H = (2N)^{-1/2}(G + G^{\top}), \tag{1.42}$$

  where $G$ is a matrix with real entries that are independently and identically distributed (iid), drawn from the standard normal distribution $\mathcal{N}(0, 1)$. This ensemble is invariant under orthogonal conjugation, modeling Hamiltonians with time-reversal symmetry.

- Gaussian Unitary Ensemble (GUE), $\beta = 2$:

  GUE matrices are complex, Hermitian, $N \times N$ matrices. A GUE matrix $H$ may be defined as

  $$H = (2N)^{-1/2}(G + G^{\dagger}), \tag{1.43}$$

  where $G$ is a matrix with complex entries that are iid, drawn from the complex standard normal distribution $\mathcal{CN}(0, 1)$. This ensemble is invariant under unitary conjugation, modeling Hamiltonians without time-reversal symmetry.

- Gaussian Symplectic Ensemble (GSE), $\beta = 4$:

GSE matrices are quaternionic, self-adjoint, $N \times N$ matrices. A GSE matrix $H$ may be defined as

$$H = (2N)^{-1/2}(G + G^{\dagger}), \tag{1.44}$$

where $G$ is a matrix with quaternionic entries that are iid, drawn from the quaternionic standard normal distribution. This ensemble is invariant under symplectic conjugation, modeling Hamiltonians with time-reversal symmetry and Kramers degeneracy (e.g. electrons with no external field).

More generally, we can define a matrix $H$ from one of these ensembles as being drawn from the distribution

$$p_{\beta}(H) \sim e^{-\frac{N\beta}{4} \operatorname{tr} H^2}, \tag{1.45}$$

where the elements of $H$ are on the field appropriate to the Dyson index $\beta$. By a change of variables, it is straightforward to find that the joint probability distribution for the eigenvalues of a Gaussian random matrix is

$$p_{\beta}(\lambda_1, \ldots, \lambda_N) \sim \exp\left(-\frac{N\beta}{4} \sum_i \lambda_i^2\right) \prod_{i<j} |\lambda_i - \lambda_j|^{\beta}. \tag{1.46}$$

Wigner [35] and Dyson [36] found that the eigenvalues behave as a one-dimensional gas confined in a quadratic potential with logarithmic repulsive interactions. This repulsion can be seen directly in the equation above, as we see that the distribution vanishes as $\lambda_i - \lambda_j \to 0$ for any $i \neq j$. This level repulsion is a hallmark of quantum chaotic systems.

Integrating out all the eigenvalues but one from the joint probability distribution and taking

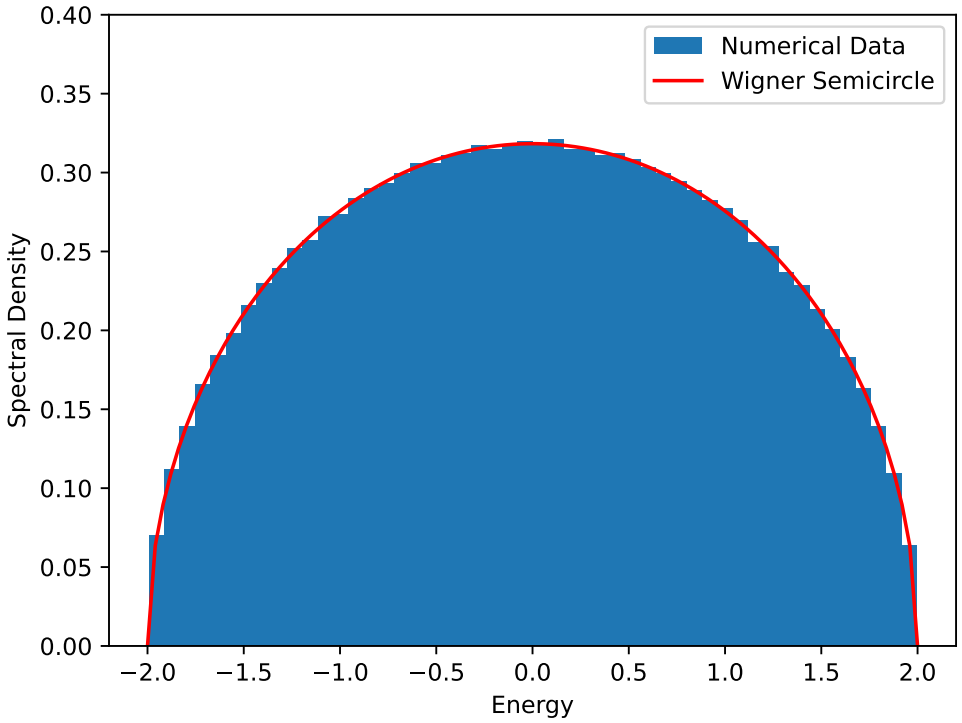



Figure 1.3: A histogram of the spectral density of a numerically generated $10^4 \times 10^4$ GOE matrix compared to the Wigner semicircle.

the limit $N \to \infty$ gives us the celebrated semicircle result

$$p(\lambda) = (2\pi)^{-1}\sqrt{4-\lambda^2}, \quad |\lambda| \leq 2 \tag{1.47}$$

by Wigner for the spectral density [37]. (Though, despite its name, this distribution is technically a semi-ellipse.) This distribution is shown in Fig. 1.3. Due to the repulsion between the levels, the spectrum exhibits spectral rigidity. That is, for finite $N$, the eigenvalue density in any supported region will have much smaller fluctuations than if the eigenvalues were iid from the semicircle distribution.

Concerning the level spacing distributions, in connection with the BGS result, the distribu-

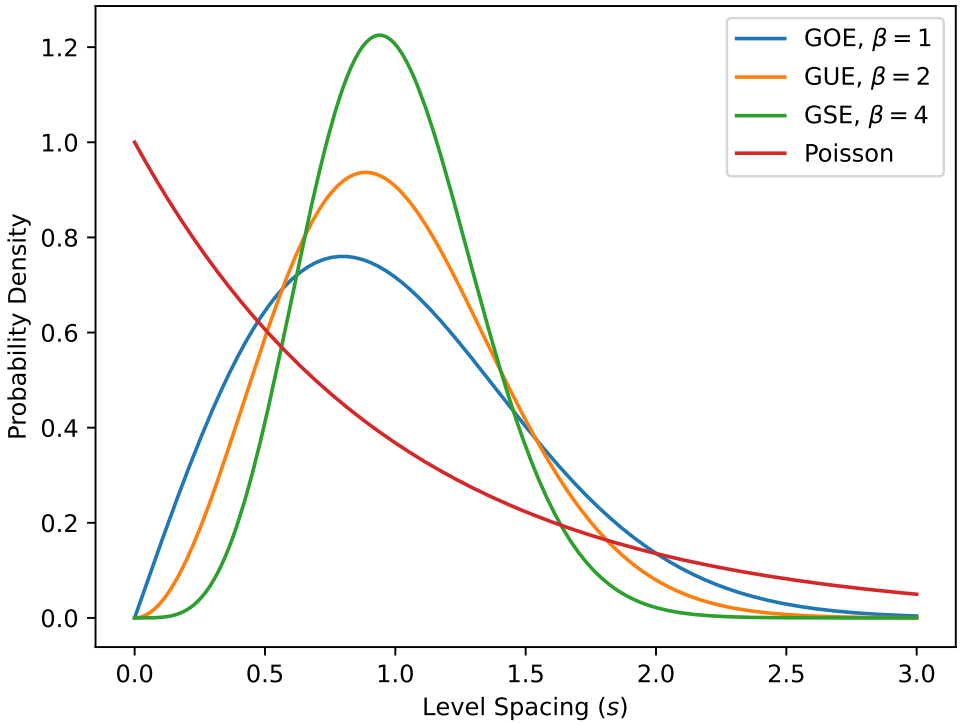



Figure 1.4: The Wigner-Dyson level spacing distributions for the Gaussian random matrix ensembles compared to the Poisson distribution.

tions for the Gaussian random ensembles are

$$p(s) = \begin{cases} \frac{\pi}{2} s e^{-\frac{\pi}{4}s^2}, & \beta = 1 \\ \frac{32}{\pi^2} s^2 e^{-\frac{4}{\pi}s^2}, & \beta = 2 \\ \left(\frac{64}{9\pi}\right)^3 s^4 e^{-\frac{64}{9\pi}s^2}, & \beta = 4 \end{cases} \cdot \tag{1.48}$$

These, along with the Poisson distribution, are shown in Fig. 1.4.

### 1.4.2 The Circular Ensembles

The Gaussian ensembles we have been discussing are analogous to ensembles of chaotic quantum Hamiltonians. However, we may wish to find analogies instead to unitary time-evolution

operators. In fact, this becomes necessary for certain quantum systems which are defined in terms of unitary operations, with no reference to Hamiltonians, such as quantum circuits. Dyson defined the circular ensembles [36] analogous to the Gaussian ensembles that are perfectly suited to this purpose. These ensembles are:

- Circular Orthogonal Ensemble (COE), $\beta = 1$:

  If $U$ is a Haar-random unitary, $U^\top U$ is a random element of the COE ensemble. COE matrices are symmetric and therefore have real elements. This ensemble is invariant under orthogonal conjugation.

- Circular Unitary Ensemble (CUE), $\beta = 2$:

  This ensemble is the Haar measure on the unitary group. Elements of CUE matrices are complex. This ensemble is invariant under unitary conjugation.

- Circular Symplectic Ensemble (CSE), $\beta = 4$:

  As in the case of the GSE, the elements of the CSE are quaternionic. If we represent quaternions by complex $2 \times 2$ matrices, we can generate representations of CSE matrices in terms of Haar-random unitaries. That is, if $U$ is a $2N \times 2N$ Haar-random unitary, then $SU^\top S^{-1}U$, where $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \otimes I_N$, is a random element of the CSE. This ensemble is invariant under symplectic conjugation.

Because the elements of the circular ensembles are unitary, their eigenvalues are complex values with unit magnitude. That is to say that they are all of the form $e^{i\theta}$ where $0 \leq \theta < 2\pi$. It is then most convenient to work in terms of eigenphases. The joint probability density of the

eigenphases from a circular ensemble is

$$p(\theta_1, \ldots, \theta_N) \sim \prod_{i<j} |e^{i\theta_i} - e^{i\theta_j}|^\beta. \tag{1.49}$$

As with the Gaussian ensembles, we can observe level repulsion from the fact that the joint probability density function (JPDF) vanishes as $\theta_i - \theta_j \to 0$ for any $i \neq j$. Unlike for the Gaussian ensembles, there is no quadratic potential containing the eigenphases. However, the eigenphases are only supported on the interval $[0, 2\pi)$, meaning the spectral width is necessarily $O(1)$. When integrating out all eigenphases but one, we find that the spectral density is uniform.

### 1.4.3 The Spectral Form Factor

While level spacing statistics are good at finding level repulsion between nearby levels, we may also ask to what degree repulsion exists between levels that are further away. A particularly useful diagnostic of chaotic random-matrix-like spectral statistics across a range of level separations is the spectral form factor (SFF). The disorder averaged SFF is defined as

$$K(t) = \sum_{i,j} \overline{e^{i(\lambda_i - \lambda_j)t}}. \tag{1.50}$$

We can gain greater insight into the physical significance of the SFF by writing it in terms of the density of states.

$$K(t) = \overline{\left| \int dE\ \rho(E) e^{-iEt} \right|^2}, \tag{1.51}$$

$$\rho(E) = \sum_i \delta(E - \lambda_i). \tag{1.52}$$

In this form, we can see that the disorder averaged SFF is the magnitude squared of the Fourier transform of the density of states. This indicates that the SFF at time $t$ is a measure of the fluctuations in the density of states on the scale of $t^{-1}$.

In addition, we can write the SFF in terms of the time evolution operator as

$$K(t) = \overline{|\operatorname{tr} U(t)|^2}, \tag{1.53}$$

where $U(t) = e^{iHt}$ for a Hamiltonian system. This means the SFF also contains information about the time dynamics of systems. We will use this fact later in this thesis to determine relaxation behaviors.

The random-matrix SFF for the ensembles discussed in this section have 3 distinct time regimes, outlined below.

- The dip (early time):

  The initial value of the SFF at $t = 0$ is clearly $N^2$, where $N$ is the matrix size. Then, as $t$ grows, the SFF decays rapidly. At these early times, the SFF probes correlations between levels that are very far from each other, and therefore uncorrelated. So the dip is given by the disconnected part of the SFF

  $$K_{\text{discon}}(t) = \left|\overline{\sum_i e^{i\lambda_i t}}\right|^2. \tag{1.54}$$

  As the different terms in the sum develop different phases there is a rapid loss of constructive interference, causing the dip. The dip can essentially be thought of as the non-universal noise that occurs before RMT behavior sets in. If we are interested only in the universal

features we may instead work with the connected SFF

$$K_{\text{con}}(t) = K(t) - K_{\text{discon}}(t). \tag{1.55}$$

The exact form of the slope depends on the spectral density.

- The ramp (intermediate time):

  The time at which the ramp appears is often called the Thouless time $t_{\text{Th}}$. For the ensembles discussed here the ramp is positive and linear in $t$. The appearance of a linear ramp is seen as an unmistakable hallmark of random-matrix-like behavior and quantum chaos. The ramp continues until the Heisenberg time $t_{\text{H}}$, which is on the scale of the inverse mean level spacing. In this intermediate time regime the SFF is probing correlations between levels much closer than the spectral width but much larger than the mean level spacing.

- The plateau (late time):

  The appearance of the plateau is simply a result of the discrete nature of the system. The SFF at these late times is probing for correlations at spacings smaller than the mean level spacing, for which none exist. The contribution of all the $i \neq j$ terms in the SFF come to 0, leaving only the diagonal correlations. So we see that the late-time value of the SFF is $N$.

Calculating the SFF for different quantum systems will be a central theme of Chapters 2-4 in this thesis.

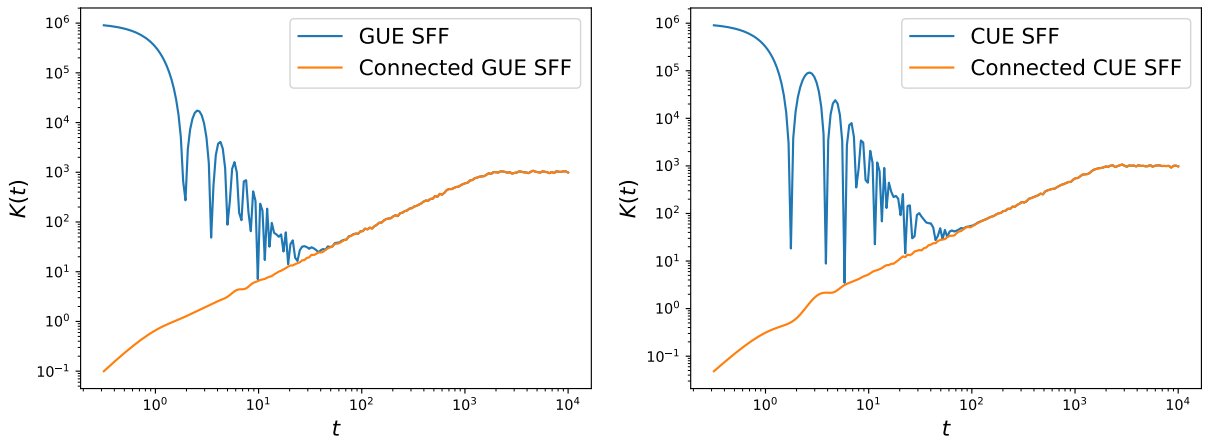



Figure 1.5: The SFFs and connected SFFs for the GUE (left) and CUE (right). Results are averaged over $10^3$ samples of $10^3 \times 10^3$ matrices.

## 1.5 Neural Networks

Recent years have been marked by excitement, extending to the general public, concerning advances in machine learning and artificial intelligence. New models have demonstrated capabilities that were once considered to require human input, such as coherent text, image, and video generation, natural language processing, and complex classification tasks. These advances are underpinned by the technology of large artificial neural networks (NNs).

Artificial NNs take inspiration from biological networks, such as the human brain. These networks are comprised of relatively simple constituents, neurons that essentially move between on and off states. Despite this simplicity, large biological neural networks are, of course, capable of impressive cognitive tasks (e.g. consciousness), proving the maxim that "more is different" [38]. Or, to put it another way, the magic lies in the structure of the connections between the neurons.

This observation makes the synergy between NNs and statistical mechanics immediately apparent. The goal of statistical mechanics is describing how large numbers of very simple

particles can interact to display rich emergent collective phenomena such as universality and phase transitions. Some NN models, such as Hopfield models [39] and Boltzmann machines [40] are directly analogous to spin models. In addition, many NNs have rich rugged loss landscapes in their parameter spaces reminiscent of the free energy landscapes of spin glasses. This opens the door to understanding learning dynamics in NNs in terms of the phases of analogous spin models, a connection explored in Chapter 5 of this thesis.

### 1.5.1 Multilayer Perceptrons

Much of the recent progress in this field has been in deep feed-forward neural networks. These networks (unlike biological networks) have their neurons grouped into layers. The network then operates by having the first layer initialized to some input state, then the information flows unidirectionally from layer to layer until the state of the final layer is observed. Feed-forward networks come in many varieties, such as convolutional neural networks [41], transformers [42], and, of particular relevance to this thesis, multilayer perceptrons (MLPs) [43, 44]. The main differences between these varieties are the connection structure between layers, the functions that determine the state of a layer in terms of the state of the previous layer, and the methods to train them.

MLPs are a particularly straightforward type of feed-forward network, as each neuron is connected to all the neurons in the layers adjacent to it. They consist of an input layer, some number of hidden layers, and an output layer. This structure is shown schematically in Fig. 1.6. The state, also called the activation, $d_i^k$ of the $i^{\text{th}}$ neuron in the $k^{\text{th}}$ layer is

$$d_i^k = \phi\left((W^k \mathbf{d}^{k-1})_i + b_i^k\right), \tag{1.56}$$

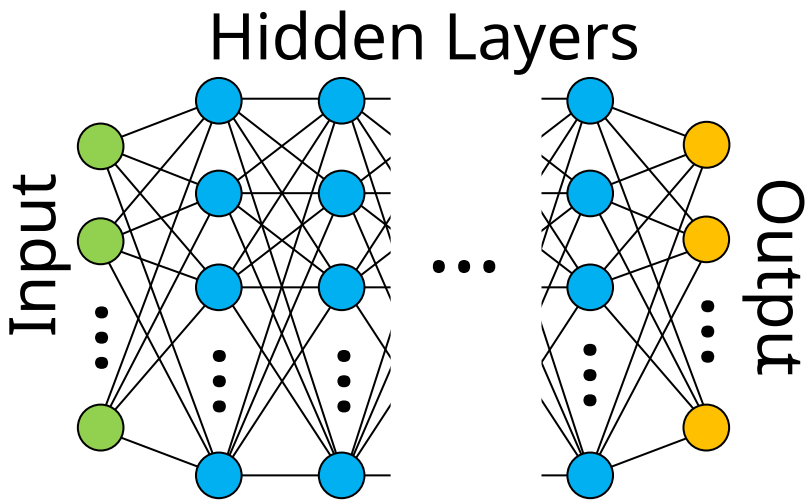



Figure 1.6: A schematic of a multilayer perceptron. Data is fed into the input layer. The state of each following layer is determined by the state of the preceding layer. Finally, the state of the output layer is read and fed into a loss function or used directly.

where the $W^k$ are matrices of trainable weights and the $b_i^k$ are trainable biases. $\phi$ is some nonlinear function called the activation function. In words, the activation of a layer is given by a linear operation on the state of the previous layer, which is then passed through the activation function element-wise. It has been shown that MLPs with at least one hidden layer are universal function approximators [45].

### 1.5.2 Training and Learning

Neural networks are trained to perform some kind of task. One of the most important aspects of training NNs is determining how to best quantify the model's performance. This is done by defining some loss function $\mathcal{L}$, that takes as its argument the state of the NNs output layer. The training task then becomes a matter of minimizing the loss function.

Feed-forward NNs are most often trained using some variety of stochastic gradient descent (SGD). In its most basic form, the steps of this training process are outlined below.

1. Initialize the trainable parameters ($W$'s and $b$'s), which we represent with the vector $\boldsymbol{\theta}$. This is most often done randomly to avoid unintentionally starting at a fixed point of the loss function.

2. Randomly select a batch of training data.

3. Calculate the average loss of the model on that batch.

4. Calculate the gradient of the loss function with respect to the trainable parameters. That is calculate

$$\mathrm{g} = \nabla_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta}). \tag{1.57}$$

   For deep networks, naively calculating these gradients will be prohibitively inefficient. However, better approaches to backpropagation [46, 47], which store information during the forward pass for later use in the gradient calculations, allow this to be done efficiently.

5. Take a step in parameter space such that

$$\boldsymbol{\theta} \mapsto \boldsymbol{\theta} - \eta \mathrm{g}. \tag{1.58}$$

   $\eta$ is called the learning rate and controls how large of a step to take in parameter space.

6. Repeat steps 2-6 until satisfactory performance is achieved.

Many variations of SGD exist that may increase convergence rate and generalization. These include adding a notion of momentum to the update rule [47, 48], looking ahead in parameter space before calculating gradients [49], and having the learning rate decay over time [50]. However, the main concept of SGD as outlined here remains.

It is important to recognize that good performance by an NN on some training set does not mean the NN will perform well generally. The situation when training performance is high but generalization performance is low is known as overtraining. Essentially, it is possible for the NN to "memorize" the training data to such an extent that it cannot perform well on data it has never seen. For this reason it is necessary to divide the available data into separate training and validation sets. Performance on the validation set, which the NN has not seen during training, is the true measure of an NN's capability. Finding ways to regularize NNs to avoid overtraining is itself an active subfield. Chapters 6 and 7 of this thesis address NN regularization through quantum effects.

### 1.5.3 Quantum Neural Networks

As a result of recent advances in quantum computing, there is a natural interest in exploring whether implementing NNs on quantum computers can lead to any sort of advantage over classical NNs. Among the existing approaches to quantum neural networks (QNNs) are quantum Hopfield networks [51–53], quantum Boltzmann machines [54–56], quantum convolutional NNs [57–61], and parameterized quantum circuits [62–65]. The last is perhaps most closely tied to current approaches to quantum computing.

Quantum circuits are the dominant paradigm for describing and implementing quantum algorithms. A circuit is comprised of a register of qubits, a sequence of unitary gates (each gate acting on one or more qubits), and projective measurements that extract classical information. Sometimes a classical register of bits is also included that can control the quantum elements of the circuit. An example of a quantum circuit is shown in Fig. 1.7.

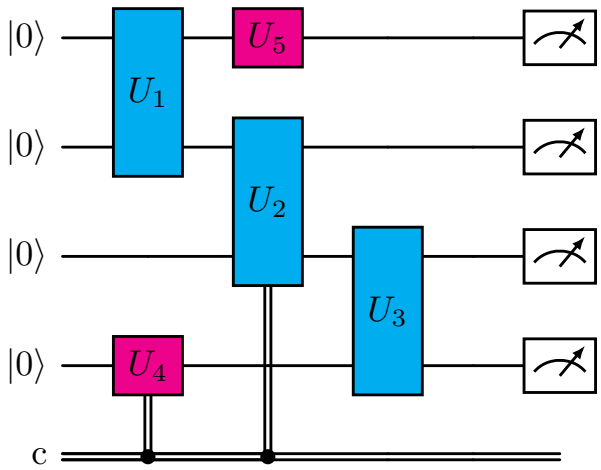



Figure 1.7: An example of a quantum circuit. Quantum gates may act on one or more qubits. Gates may have classical controls. At the end of the circuit projective measurements collect information.

A QNN implemented by a quantum circuit will have some gate structure, where the gates are defined by some set of parameters. After projective measurements are made at the end of the circuit and fed into a loss function, the parameters will be iteratively updated to minimize the loss function. A major drawback of this approach is that unitary operations are linear, in contrast to the nonlinear activations used in classical NNs, which may greatly limit the expressivity of such a QNN. Chapters 6 and 7 explore an alternative formulation of a QNN that introduces nonlinearities through mid-network measurements.

## 1.6 Outline of the Remainder of This Thesis

The remainder of this thesis is comprised of six chapters adapted from Refs. [1–6] by the present author and collaborators.

Chapter 2 is devoted to calculating the SFF of a minimal quantum glass model. Physical glass models have very rugged energy landscapes, causing the Hilbert space to be partitioned

into sectors. At early times in the glassy phase, the system's dynamics is confined to the sectors, with thermalization between sectors only possible at much later times. With this in mind, a random matrix model called the block Rosenzweig-Porter model (BRP) is defined, which captures the essence of this segmented Hilbert space structure, with tunable weak interactions between sectors. This model is a generalization of the standard Rosenzweig-Porter (RP) model, which supports localized, non-ergodic extended, and ergodic phases. Using the Itzykson-Zuber identity and a nonlinear sigma model approach, the exact SFF in the thermodynamic limit is calculated. The results show how the strength of inter-sector couplings affects thermalization timescales.

Chapter 3 is concerned with the study of the spectral statistics for a more physical quantum glass model: the quantum spherical $p$-spin model. The Hamiltonian for this model is of the form of Eq. (1.8). It is constructed by treating the spin variables as position coordinates, while maintaining some of the model's spin nature by constraining the dynamics to the surface of an $N$-dimensional hypersphere. The model is quantized from its classical counterpart by defining operators conjugate to the position variables and adding a "kinetic energy" term to the Hamiltonian. At early times the SFF for this model has a ramp that is linear in time, but the coefficient of this ramp is exponentially large in system size. It is shown that this ramp coefficient measures the number of TAP states that emerge at the dynamic transition away from paramagnetism.

Chapter 4 takes the ideas of the BRP model explored in Chapter 2, and translates them to the context of quantum circuits. In this context there is no Hamiltonian describing the system. Behavior is instead determined by a random Floquet time-evolution operator. It is shown how a sector structure analogous to that of the BRP model can be built into the construction of this time-evolution operator. Using an effective theory for random quantum circuits and a diagrammatic method of integration over the unitary group, the SFF and relaxation behavior of this circuit is ex-

plored. Just like the BRP model, this Floquet model's thermalization time shows dependence on the degree of inter-sector coupling. This study opens up possibilities for exploring rich quantum thermalization dynamics through quantum circuits.

In Chapter 5, the techniques of spin glass theory are applied to NNs. Spin model analogues of MLPs are explored. In this analogy neurons are mapped to Ising spins and weights are mapped to spin-spin interactions. Using the replica approach, it is found that an Ising spin model, called the multilayer Sherrington-Kirkpatrick (MSK) model due to its layered structure like an MLP, exhibits a transition from paramagnet to spin glass. The nature of this transition is then explored throughout an MLP's training process. That is, for an MLP, a sequence of analogous spin model Hamiltonians is defined, parameterized by training time. It is found that training changes the nature of the melting transition. After a very short amount of training, the transition changes from a glass transition to an ordered-phase transition, reminiscent of ferromagnetism, but with no apparent global order. This study demonstrates that statistical mechanical approaches yield valuable insights into the learning process.

Chapters 6 and 7 are then devoted to explorations in quantum machine learning. In Chapter 6 a quantum circuit is constructed which implements a quantum generalization of an MLP. The novelty of this approach is that it allows for adiabatic tuning of a NN between classical and quantum modes of operation, allowing for an "apples-to-apples" comparison between classical and quantum models. These models are trained on a subset of the MNIST dataset of images of handwritten numerical digits [66]. It is found, through classical simulation, that adding a certain amount of uncertainty into the model by turning up its "quantumness" regularizes the model, preventing overtraining and improving generalization performance.

In Chapter 7 these ideas are implemented on actual quantum hardware. Training of the

networks is still done through classical simulation, but classification of the images is done on both trapped-ion and superconducting quantum computing platforms. As in simulations, it is found that a certain degree of quantum uncertainty is beneficial. Additionally, the physical noise that is inevitable in these platforms is also found to be helpful in some contexts. This introduces the possibility of using this simple quantum MLP to benchmark performance across different quantum computing platforms.

# Chapter 2: Spectral Statistics of a Minimal Quantum Glass Model

**Authors:** Richard Barney, Michael Winer, Christopher L. Baldwin, Brian Swingle, and Victor Galitski

**Abstract:** Glasses have the interesting feature of being neither integrable nor fully chaotic. They thermalize quickly within a subspace but thermalize much more slowly across the full space due to high free energy barriers which partition the configuration space into sectors. Past works have examined the Rosenzweig-Porter (RP) model as a minimal quantum model which transitions from localized to chaotic behavior. In this work we generalize the RP model in such a way that it becomes a minimal model which transitions from glassy to chaotic behavior, which we term the "Block Rosenzweig-Porter" (BRP) model. We calculate the spectral form factors of both models at all timescales larger than the inverse spectral width. Whereas the RP model exhibits a crossover from localized to ergodic behavior at the Thouless timescale, the new BRP model instead crosses over from glassy to fully chaotic behavior, as seen by a change in the steepness of the ramp of the spectral form factor.

## 2.1 Introduction

Random Hermitian matrices provide a simple model for the energy levels of a wide variety of quantum systems, including complex nuclei [36, 67–71], systems with a chaotic classical

limit [30, 32, 72–75], strongly interacting quantum field theories [76–80], and much more. The wide prevalence of random-matrix-like energy levels is known as random matrix universality [32, 81, 82], and it is one of the key manifestations of quantum chaos. Given an ensemble of Hamiltonians $\{H\}$, one way to characterize random matrix universality is in terms of the statistical correlations of eigenvalues. Letting $\rho(E)$ denote the density of eigenvalues of a particular random $H$, then although the average density $\overline{\rho(E)}$ is not universal, the pair-correlation $\overline{\rho(E)\rho(E')}$ does turn out to be. Indeed, one finds that the pair-correlation is closely related to the pair-correlation of a Gaussian random matrix of the appropriate symmetry.

However, while many systems ultimately exhibit random-matrix-like behavior at the finest energy scales, i.e. for sufficiently small $E-E'$, real systems typically have additional structure in their energy spectrum which is not random-matrix-like. This structure may be eventually washed out at the finest energy scales, but it does cause a deviation from the random matrix behavior of $\overline{\rho(E)\rho(E')}$ when $E-E'$ is of some intermediate size. Perhaps the simplest such structure occurs when the Hamiltonian breaks up into approximately decoupled blocks labelled by some almost-conserved quantity. This situation is common and is broadly related to the presence of slow dynamics, e.g. a slowly diffusing charge or almost-frozen glassy dynamics. In the simplest case, each block is statistically independent and the blocks are connected by some additional weak perturbations. Then as a function of $E-E'$, the pair-correlation can exhibit a crossover from multiple small random blocks to a single large random block. Recently, a set of effective theories have been formulated to describe this crossover [2]. Here we present a new solvable model in which this crossover can be analytically verified and studied. The results are consistent with Ref. [2], but they also offer new insights into various regimes.

This new model is a generalization of the unitary Rosenzweig-Porter (RP) model, which

is itself a slight generalization of the original model constructed by Rosenzweig and Porter to describe complex atomic spectra [83]. The RP Hamiltonian is

$$H = A + \frac{\lambda}{N^{\gamma/2}} V, \tag{2.1}$$

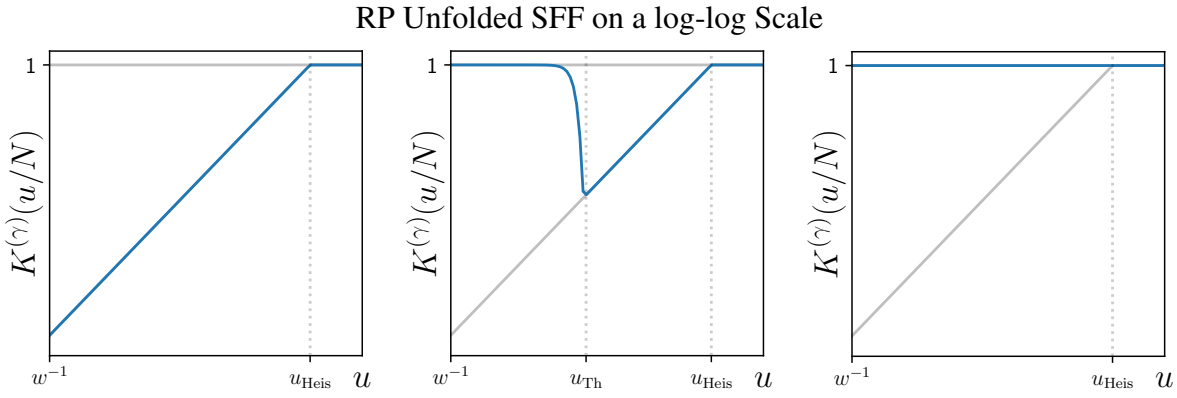
where $A$ is a diagonal matrix with independent and identically distributed elements, and $V$ is a random matrix drawn from the Gaussian Unitary Ensemble (GUE) [84] whose matrix elements have unit variance. $N$ is the size of the Hilbert space. Several studies [85–99] have examined this model at different values of $\gamma$. The RP model has also been generalized in several different ways to create a family of interesting random matrix ensembles [100–106].

Motivated by the problem of many-body localization, Kravtsov et al. [89] studied this RP ensemble as a simple random matrix model with both localization and ergodicity-breaking transitions. Consistent with previous results [85–88, 107], they found Poissonian statistics for $\gamma > 2$, indicating Anderson localized behavior, and GUE statistics for $\gamma < 1$, indicating chaotic behavior. They found that intermediate values of $\gamma$ led to a non-ergodic extended phase in which eigenstates are neither localized nor fully ergodic.

One useful tool for studying spectral statistics is the unfolded spectral form factor (SFF), which is the Fourier transform of the unfolded two-point correlation function. Unfolding refers to the process of rescaling the spectrum in such a way that the local mean level spacing becomes unity, bringing the universal features of the SFF to the fore. Ref. [89] provides a calculation of the RP SFF for all $\gamma > 1$ at the Thouless timescale, i.e., the timescale at which random matrix statistics first appear. Other timescales are then examined by rescaling time in this result.

We can form schematic expectations for the behavior of the RP SFF at different values of

$\gamma$ by comparing the Thouless time to the two other cardinal times: the inverse spectral width $w^{-1}$ and the Heisenberg time, which is the inverse mean level spacing. These schematics are shown in Fig. 2.1. In the figure $u/N$ is the unfolded time, that is the real time divided by the level density. When the Thouless time is smaller than the inverse spectral width, which occurs when $\gamma < 1$, we expect GUE statistics indicating chaotic behavior. When it is larger than the Heisenberg time, which occurs when $\gamma > 2$, we expect Poissonian statistics indicating localized behavior. If the Thouless time is larger than the inverse spectral width but smaller than the Heisenberg time we expect to see a crossover between the two behaviors. Phase transitions occur when the Thouless timescale coincides with one of the cardinal timescales.




(a) $\gamma < 1$, $u_{\text{Th}} \ll w^{-1}$ (b) $1 < \gamma < 2$, $w^{-1} \ll u_{\text{Th}} \ll u_{\text{Heis}}$ (c) $\gamma > 2$, $u_{\text{Th}} \gg u_{\text{Heis}}$

Figure 2.1: Schematic representation on a log-log scale of the unfolded SFF for the RP model in each of the three identified phases. The lower gray curve is the GUE SFF while the upper gray curve is the Poissonian result.

As an initial result of the present work, we directly calculate the SFF of the RP model at all timescales larger than the inverse spectral width. This approach has the added benefit of more closely paralleling the calculation of the SFF for the new model we examine. The result is shown in Eq. (2.66). We find that this result does follow the schematic expectations shown in Fig. 2.1 and is in good agreement with the numerical results we obtain. All these findings are in

agreement with Ref. [89].

In the bulk of this work, we generalize the RP model to obtain a random matrix model which transitions from chaotic to glassy behavior, where glassiness is identified by the presence of multiple thermalization timescales. This is accomplished by redefining $A$ in Eq. (2.1) to be the block-diagonal matrix

$$A = \bigoplus_{i=1}^{P} A^{(i)}, \tag{2.2}$$

where each block $A^{(i)}$ is an independent GUE matrix of size $M = N/P$ whose elements have variance $M^{-1}$ (so that the eigenvalues of each $A^{(i)}$ are finite at large $M$). We call this generalization of the RP model the Block Rosenzweig-Porter (BRP) model. Because of the the normalization of each $A^{(i)}$ the system thermalizes in $O(1)$ time within each block but can fully thermalize only at times diverging with $N$ (if at all). Due to the different structure of the $A$ matrix, the BRP model will have a localized phase different from that of the RP model. In this phase the system will be confined to the subspaces corresponding to the blocks of $A$ instead of individual states.

This generalization of the RP model is motivated by our recent work examining the SFF of a canonical quantum spin glass model [2]. In that work we found that the SFF has a linear ramp at times sub-exponential in the system size, as would a chaotic system. However, the ramp is steeper by a factor of the number of distinct metastable states. This enhancement of the ramp is a hallmark of glassy behavior. It indicates that, at these early timescales, the Hamiltonian can be viewed as a direct sum of independent random matrices. However, it has been argued that, for appropriate values of the coupling between spins and at a sufficiently late timescale, the system will escape from the metastable configurations and fully thermalize [108–110]. At this timescale we would expect the SFF to experience a crossover from the enhanced ramp to the GUE ramp.

Due to the challenge of calculating the SFF of canonical quantum spin glass models at late times with current methods, we turn to the BRP model as a simpler model of a quantum glass for which the SFF can be calculated at late timescales and the crossover from glassy to chaotic behavior can be seen.

As we did for the RP model, we form schematic expectations for the behavior of the SFF at different values of $\gamma$ by comparing the Thouless time with the other cardinal times. These schematics are shown in Fig. 2.2. When the Thouless time is less than the inverse spectral width, which we find occurs when $\gamma < 1$, we expect GUE statistics indicating chaotic behavior. When it is larger than the Heisenberg time, which we find occurs when $\gamma > 2$, we expect the statistics of uncoupled GUE blocks, indicating localization within the blocks. Note that in this case the SFF reaches its plateau value at a time earlier than the Heisenberg time. This is the block Heisenberg time, which is the inverse mean level spacing for a single block of $A$. This means that the block Heisenberg time is smaller than the full Heisenberg time by a factor of $P$, the number of blocks. The block Heisenberg time is an additional cardinal time in the BRP model.

If the Thouless time is larger than the inverse spectral width and smaller than the Heisenberg time we expect to see a crossover between the block localized and chaotic behaviors at the Thouless timescale. Further, if the Thouless time is smaller than the block Heisenberg time, which occurs when $\gamma < 1 + d$ where $d = \log_N M$, this crossover will occur before the SFF can reach its plateau value at the block Heisenberg time. If the Thouless time is greater than the block Heisenberg time, which occurs when $\gamma > 1 + d$, the SFF will reach its plateau value at the block Heisenberg time, drop down to the GUE result at the Thouless timecale, then reach the plateau value again at the Heisenberg time.

Our primary result is the calculation of the SFF of the BRP model at all timescales larger

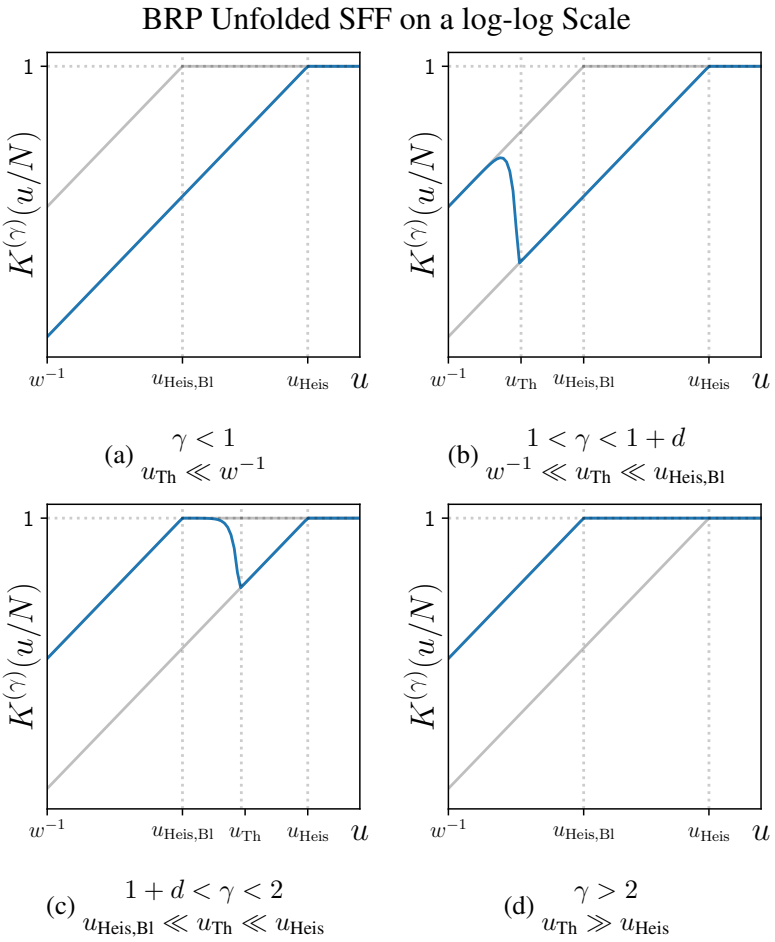



Figure 2.2: Schematic representation on a log-log scale of the unfolded SFF for the BRP model in each of the four identified phases. The lower gray curve is the GUE SFF while the gray curve above it is the SFF for completely uncoupled GUE blocks.

than the inverse spectral width, which is $O(1)$. This result is shown in Eqs. (2.132)-(2.133). It follows the schematic expectations shown in Fig. 2.2 and is in good agreement with the numerical results we obtain. At the Thouless timescale, if it is not of the same order as the Heisenberg time, the SFF decays exponentially over time from its glassy result to the GUE SFF, as shown in Eq. (2.155). In the case where the Thouless time is of the same order as the Heisenberg time, a more complicated crossover behavior emerges, shown in Eqs. (2.156)-(2.157).

Our results indicate that for $\gamma < 1$ the system will immediately thermalize while for $\gamma > 2$ the system will always remain localized within blocks. For intermediate values of $\gamma$ the system

will initially be confined to a single block, but will escape and thermalize at the Thouless time. We also find a transition at $\gamma = 1 + d$. This transition has an important interpretation in terms of the eigenstates. It is the point at which the eigenstates become fully delocalized across the Hilbert space.

The remainder of this work is organized as follows. In section 2.2 we review the spectral form factor's definition, features, and role as a diagnostic of quantum chaos. In section 2.3 we outline the aspects of the construction of the SFF which are common to both the RP and BRP models and discuss the unfolding procedure. In section 2.4 we examine the RP SFF, discussing the results of previous works and directly calculating the SFF at relevant timescales. We compare our analytical results with numerics. Section 2.5 contains the bulk of our results, an examination of the BRP model and the calculation of its SFF, again comparing to numerical results. Finally, we conclude with section 5.4.

## 2.2 Review of the spectral form factor

The spectral form factor (SFF) has a long history as a diagnostic of quantum chaos [32, 81, 84]. Examples of random-matrix-like SFFs in chaotic systems appear in numerous areas of physics, from nuclear systems [36, 111] to condensed matter [112–114] to holographic theories [74, 75]. The SFF diagnoses whether energy levels repel as they do in random matrices [112], have independent Poissonian statistics [115], or have some more exotic behavior [116–118]. The SFF can be written as

$$\text{SFF}(t, f) = \overline{|\operatorname{tr}[e^{-iHt} f(H)]|^2}, \tag{2.3}$$

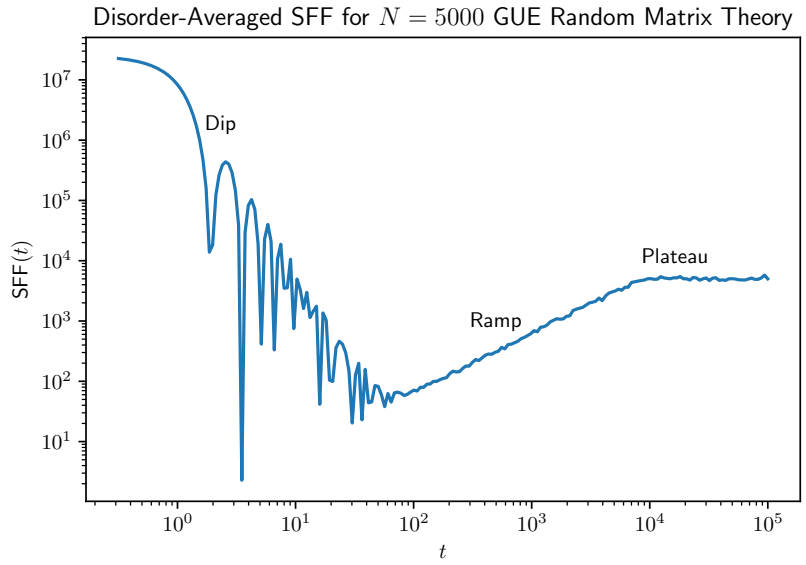



Figure 2.3: A log-log plot of the disorder-averaged SFF for the Gaussian unitary ensemble (GUE). The matrices in this ensemble have dimension $N = 5000$. The SFF was computed numerically by exactly diagonalizing five hundred realizations. The three regimes of the SFF—dip, ramp, plateau—are each labeled.

where $f$ is a filter function used to pick an energy band of interest, and the overline denotes a disorder average over an entire ensemble of Hamiltonians.

The SFF can also be expressed as the two-point function of the density of states. Defining

$$\rho(E, f) = \sum_n \delta(E - E_n) f(E_n) = \mathrm{tr}\left(\delta(H - E) f(H)\right), \tag{2.4}$$

we have

$$\mathrm{SFF}(t, f) = \overline{\int dE_1 dE_2 \rho(E_1, f)\rho(E_2, f) e^{i(E_1 - E_2)t}} \tag{2.5}$$

As a two-point function, the SFF can be broken down into connected and disconnected components. These components have very different behaviors, as discussed below.

For random-matrix-like systems, the spectral form factor has three regimes of interest.

- The "dip," also known as the slope, occurs at early times. It comes from the disconnected piece of the SFF (and thus its precise shape is non-universal and depends on the details of $f$ and the thermodynamics of the system). Its downward nature reflects a loss of constructive interference: at $t = 0$ the terms in $\mathrm{tr}[e^{-iHt} f(H)]$ are all positive, but the different terms of $\mathrm{tr}\, e^{-iHt}$ acquire different phase factors as $t$ increases.

- The "ramp" occurs at intermediate times. It is arguably the most interesting regime, and marks the beginning of the universal behavior in the connected spectral form factor. In the canonical matrix ensembles, it is a consequence of the result [84]

$$\mathbb{E}\left[\rho\left(E+\frac{\omega}{2}\right)\rho\left(E-\frac{\omega}{2}\right)\right]-\mathbb{E}\left[\rho\left(E+\frac{\omega}{2}\right)\right]\mathbb{E}\left[\rho\left(E-\frac{\omega}{2}\right)\right]\sim-\frac{1}{\mathfrak{b}\pi^2\omega^2}, \tag{2.6}$$

  where $\mathfrak{b} = 1, 2, 4$ for the orthogonal, unitary, and symplectic ensembles respectively. The fact that the right hand side is negative is a manifestation of level repulsion [111]. Taking the Fourier transform of Eq. (3.7) with respect to $\omega$ gives a term proportional to $t$ for the connected SFF. Such a linear-in-$t$ ramp is often taken as a defining signature of quantum chaos.

  The exact coefficient of the ramp can tell us a lot about a quantum system. For instance, if $H$ is not a GUE matrix but the direct sum $H_1 \oplus H_2$ of two GUE matrices, the ramp will be enhanced by a factor of two. This enhancement shows up in realistic systems such as the Bunimovich stadium [119, 120], which have different sectors of their Hamiltonian which behave differently under reflection symmetry. It has recently been shown [2] that at times sub-exponential in system size, all-to-all spin glasses exhibit an enhancement of the ramp

equal to the number of effective "sectors," i.e., regions of configuration space rendered dynamically disconnected by large energy barriers. This work serves as an extension of that result for a toy model of spin glasses, going out to late times.

- The "plateau" occurs at late times. It is a signature of the discreteness of the spectrum. It is part of the connected spectral form factor and is completely universal to all systems, whether thermalizing, integrable, glassy, or many-body localized [121, 122]. At times much larger than the inverse level spacing or "Heisenberg time," one expects that all off-diagonal terms in the double-trace of the SFF average to zero, meaning that

$$\mathrm{SFF}(t, f) = \sum_{mn} \overline{e^{-i(E_m - E_n)t} f(E_m) f(E_n)} \sim \sum_n f(E_n)^2. \tag{2.7}$$

  For integrable systems, the plateau is reached very quickly with little to no ramp regime [115, 117, 118], whereas for chaotic systems the plateau isn't reached until a time exponential in system size.

In the bulk of this paper we calculate the unfolded connected SFF for the RP and BRP models at all timescales larger than the inverse spectral width, thus we capture both the ramp and the plateau regimes. We find that when the sectors of the BRP model are uncoupled the ramp of the SFF is enhanced by a factor of the number of sectors. As the coupling strength is increased there will be a crossover to the regular ramp at the Thouless timescale.

## 2.3 Construction of the spectral form factor

In this section we lay out the initial steps of the construction of the SFF which are common to both the RP model and the new BRP model. In fact, the results of this section are applicable to any Hamiltonian which is perturbed by a GUE matrix. Both of the models can be expressed as a Hamiltonian matrix of size $N$ with the form

$$H = A + V, \tag{2.8}$$

where the spectrum of $A$ is centered at 0 and has a width that is $O(1)$ with respect to $N$. $V$ is a random GUE matrix drawn from the distribution

$$p_V(V) \sim \exp\left(-\frac{1}{2\sigma}\operatorname{tr} V^2\right),\ \sigma = \frac{\lambda^2}{N^\gamma}. \tag{2.9}$$

The exact details of $A$ will depend on whether we are working with the RP or BRP ensemble, but are unimportant at this initial stage.

### 2.3.1 The joint probability density function

Our first step is to find the joint probability density function (JPDF) of the eigenvalues of the Hamiltonian $H$. We follow the method of Kunz and Shapiro [107]. For the moment we will hold $A$ constant. The probability distribution of $H$ is then

$$p_H(H) = p_V(H - A) \sim \exp\left[-\frac{1}{2\sigma}\sum_i \left(E_i^2 + a_i^2\right)\right] \exp\left(\frac{1}{\sigma}\operatorname{tr} AUEU^\dagger\right), \tag{2.10}$$

where $\{a_i\}$ are the eigenvalues of $A$, $\{E_i\}$ are the eigenvalues of $H$, $E = \mathrm{diag}(E_1, \ldots, E_N)$ is the diagonal matrix similar to $H$, and $U$ is the unitary matrix which diagonalizes $H$.

We now make the change of variables $H \to \{U, E\}$. The Jacobian of this change is $\Delta^2(E)$, where $\Delta(E) = \prod_{i>j}(E_i - E_j)$ is the Vandermonde determinant. The JPDF is

$$p(E) \sim \Delta^2(E) \exp\left[-\frac{1}{2\sigma}\sum_i \left(E_i^2 + a_i^2\right)\right] \int dU \exp\left(\frac{1}{\sigma}\,\mathrm{tr}\, AUEU^\dagger\right), \tag{2.11}$$

where $dU$ is the Haar measure over the unitary group $U(N)$. We can evaluate the integral over $U$ using the Itzykson-Zuber integral identity [77]

$$\int dU \exp\left(\frac{1}{\sigma}\,\mathrm{tr}\, AUEU^\dagger\right) \sim \frac{\det\exp(a_i E_j/\sigma)}{\Delta(a)\Delta(E)}. \tag{2.12}$$

Applying this identity, we find that the JPDF is

$$p(E) = c(\sigma, N)\frac{\Delta(E)}{\Delta(a)} \exp\left[-\frac{1}{2\sigma}\sum_i \left(E_i^2 + a_i^2\right)\right] \det\exp\left(\frac{a_i E_j}{\sigma}\right), \tag{2.13}$$

with $c(\sigma, N)$ to be determined by normalization.

To normalize we calculate

$$\begin{aligned}
1 =& \int dE\; p(E) \\
=& \frac{c(\sigma,N)}{\Delta(a)} \int dE\; \Delta(E) \exp\left[-\frac{1}{2\sigma}\sum_i \left(E_i^2 + a_i^2\right)\right] \det\exp\left(\frac{a_i E_j}{\sigma}\right) \\
=& \frac{c(\sigma,N)}{\Delta(a)} \int dE\; \Delta(E) \exp\left[-\frac{1}{2\sigma}\sum_i \left(E_i^2 + a_i^2\right)\right] \sum_{\pi\in S_N} \mathrm{sign}(\pi) \prod_{i=1}^{N} \exp\left(\frac{a_i E_{\pi(i)}}{\sigma}\right) \\
=& c(\sigma,N)\frac{N!}{\Delta(a)} \int dE\; \Delta(E) \exp\left[-\frac{1}{2\sigma}\,\mathrm{tr}(E-a)^2\right],
\end{aligned} \tag{2.14}$$

where $a = \mathrm{diag}(a_1, \ldots, a_N)$. In the second line above $S_N$ is the set of permutations of $N$ elements. The third line follows because $\Delta(E)$ is antisymmetric under all transpositions. We now use the identity (proven in the appendix of [107])

$$\int dE\, \Delta(E) \exp\left[-\frac{1}{2\sigma}\,\mathrm{tr}(E-a)^2\right] = \Delta(a)(2\pi\sigma)^{N/2} \tag{2.15}$$

to find that the normalization factor is

$$c(\sigma, N) = \frac{1}{N!}(2\pi\sigma)^{-N/2}. \tag{2.16}$$

We now let $A$ also be a random matrix. From Eqs. (2.13) and (2.16) we see that, for any function $W(E)$ which is symmetric in the eigenvalues of $H$, the ensemble average is

$$\overline{W}(E) = (2\pi\sigma)^{-N/2} \left\langle \frac{1}{\Delta(a)} \int dE\, W(E)\Delta(E) \exp\left[-\frac{1}{2\sigma}\,\mathrm{tr}(E-a)^2\right] \right\rangle, \tag{2.17}$$

where the angle brackets indicate that the average over the distribution of $A$ still needs to be performed. Fortunately the SFF and the correlation functions examined below are such symmetric functions.

### 2.3.2 Correlation functions

We now examine two correlation functions which we will use to construct the SFF. The first is the Fourier transform of the level density

$$\overline{C_1}(t) = \sum_k \overline{e^{itE_k}}. \tag{2.18}$$

Averaging with respect to the JPDF and making use of Eq. (2.15) yields

$$\overline{C_1}(t) = e^{-\sigma t^2/2} \left\langle \sum_k e^{ita_k} \frac{\Delta(a+\delta_k\tau)}{\Delta(a)} \right\rangle = e^{-\sigma t^2/2} \left\langle \sum_k e^{ita_k} \prod_{j\neq k} \left(1 + \frac{\tau}{a_k - a_j}\right) \right\rangle, \tag{2.19}$$

where $\tau = it\sigma$ and $\delta_k$ is the projection matrix onto the $k^{\text{th}}$ dimension. In order to average over the eigenvalues of $A$ we rewrite this as

$$\overline{C_1}(t) = \frac{e^{-\sigma t^2/2}}{\tau} \oint_{\mathcal{C}} \frac{dz}{2\pi i} e^{itz} \left\langle \prod_j \left(1 + \frac{\tau}{z - a_j}\right) \right\rangle, \tag{2.20}$$

where $\mathcal{C}$ is a rectangular integration contour of infinitesimal width in the imaginary direction which encompasses the real axis. Using this form is advantageous because each term within the angle brackets now depends on only a single eigenvalue of $A$.

We now consider the Fourier transform of the two-point correlation function (excepting the $k = l$ terms)

$$\overline{C_2}(t) = \sum_{k\neq l} \overline{e^{it(E_k - E_l)}}. \tag{2.21}$$

Again, averaging with respect to the JPDF and using Eq. (2.15) yields

$$\begin{aligned}\overline{C_2}(t) &= e^{-\sigma t^2}\left\langle \sum_{k\neq l} e^{it(a_k-a_l)}\frac{\Delta\left(a+\tau(\delta_k-\delta_l)\right)}{\Delta(a)}\right\rangle \\ &= e^{-\sigma t^2}\left\langle \sum_{k\neq l} e^{it(a_k-a_l)}\left[1-\left(\frac{\tau}{a_l-a_k-\tau}\right)^2\right]\prod_{j\neq k}\left(1+\frac{\tau}{a_k-a_j}\right)\prod_{j\neq l}\left(1-\frac{\tau}{a_l-a_j}\right)\right\rangle.\end{aligned} \tag{2.22}$$

We can also write this in terms of contour integrals as

$$\begin{aligned}\overline{C_2}(t) = -\frac{e^{-\sigma t^2}}{\tau^2}\oint_{\mathcal{C}}\frac{dz}{2\pi i}\oint_{\mathcal{C}}\frac{dz'}{2\pi i}e^{it(z-z')}\left[1-\left(\frac{\tau}{z'-z-\tau}\right)^2\right] \\ \times\left\langle\prod_j\left(1+\frac{\tau}{z-a_j}\right)\left(1-\frac{\tau}{z'-a_j}\right)\right\rangle.\end{aligned} \tag{2.23}$$

### 2.3.3 The unfolded spectral form factor

We now combine the correlation functions considered above to form the function

$$C(t) = \frac{1}{N}\left(\overline{C_2}(t) - \left|\overline{C_1}(t)\right|^2\right), \tag{2.24}$$

which is the connected SFF apart from the missing $k = l$ terms in $C_2(t)$. We will now demonstrate how $C(t)$ can be used to find the unfolded SFF. We note that

$$C(t) = -\frac{1}{N}\int dz\, dz' e^{it(z-z')}\left(\rho(z)\rho(z') - \rho_2(z,z') + \sum_k \overline{\delta(z-E_k)\delta(z'-E_k)}\right), \tag{2.25}$$

$$\rho(z) = \sum_k \overline{\delta(z-E_k)},\ \rho_2(z,z') = \sum_{k,l}\overline{\delta(z-E_k)\delta(z-E_l)}. \tag{2.26}$$

We now unfold by making the change of variables

$$(z, z') = x \pm \frac{y}{2\rho(x)}, \tag{2.27}$$

$$t = \rho(x)T. \tag{2.28}$$

In these new variables $x$ is the central energy and $y$ is the distance between the energies in units of local mean level spacings at the central energy. With this change we find

$$C(t) = \int dx\, p(x)\left(K(x,T) - 1\right), \tag{2.29}$$

$$K(x,T) = -\int dy\, Y(x,y)e^{iTy}, \tag{2.30}$$

$$Y(x,y) = \frac{1}{[\rho(x)]^2}\left[\rho\left(x + \frac{y}{2\rho(x)}\right)\rho\left(x - \frac{y}{2\rho(x)}\right) - \rho_2\left(x + \frac{y}{2\rho(x)}, x - \frac{y}{2\rho(x)}\right)\right], \tag{2.31}$$

where $p(x) = \rho(x)/N$ is the probability distribution function, $Y(x,y)$ is the unfolded connected two-point correlation function, also called the unfolded two-point cluster function, and $K(x,T)$ is the unfolded connected SFF. Our approach, whichever model we are examining, will be to put $C(t)$ in the form of Eq. (2.29) and extract the unfolded connected SFF, which we will simply call the SFF going forward.

For the models we study in this work it is reasonable to assume that $Y(x,y)$ vanishes in the large $N$ limit for $y \gg 1$, so we can restrict the interval of integration in Eq. (2.30) to be of order 1. For these values of $y$ the disconnected part of $Y(x,y)$ will tend to 1, thus it will not contribute to the SFF for $T \neq 0$. In the calculations that follow we will therefore neglect the disconnected piece of the two-level cluster function. This is an advantage of the unfolding procedure; we no longer need to worry about subtracting out the disconnected part in order to observe the universal

features of the SFF. We will also find that if we unfold the coupling parameter $\lambda$ as well, $Y(x, y)$ and the SFF will become independent of the center energy $x$. So, although $C(t)$ is not universal due to its dependence on $p(x)$, the unfolded connected two-point correlation function and the SFF will be universal.

It will be helpful to use the time variable $u = N|T| = |t|/p(x)$, so $u$ is of the same order as the real time $t$. This means that $u$ is the time unfolded by the level probability density instead of the level density. The modulus may be taken since the SFF is symmetric in time. Because the SFF is the Fourier transform of the two-point cluster function, at time $u$ it probes the correlations between energy levels with separations on the order of $u^{-1}$. For this reason we consider only timescales larger than the inverse spectral width.

## 2.4 The Rosenzweig-Porter model

The Hamiltonian matrix for the RP model has the form of Eq. (2.8), with the additional condition that the eigenvalues of $A$ are all independent and identically distributed. We can consider $A$ to be diagonal such that

$$A = \mathrm{diag}(a_1, \ldots, a_N), \tag{2.32}$$

where each $a_i$ is independently drawn from some distribution $p(a)$ with a variance of order 1 in $N$.

We can view this as an Anderson model [123] of $N$ sites with independent random on-site potentials and all-to-all random couplings on the order of $N^{-\gamma/2}$. From Fermi's golden rule, we can determine that the tunneling rate from one of these sites is on the order of $N^{1-\gamma}$. Thus the Thouless time, the time at which the system will escape a single site and random matrix statistics

first appears, is

$$t_{\text{Th}} \sim N^{\gamma-1}. \tag{2.33}$$

The ergodicity-breaking transition occurs at the point where the tunneling rate is the same order as the spectral width [94, 109, 110, 124], or, equivalently, when the Thouless time becomes of the same order as the inverse spectral width. The spectral width of the GUE matrix $V$ is on the order of $N^{(1-\gamma)/2}$ [84], while the spectral width of $A$ is of order 1. So the spectral width of the RP model is

$$w \sim \max(1, N^{(1-\gamma)/2}), \tag{2.34}$$

This indicates that the ergodicity-breaking transition occurs at $\gamma = 1$. For all $\gamma < 1$ the RP model will exhibit GUE statistics.

The localization transition of the model, on the other hand, can be determined from the Mott criterion [94, 110, 124, 125], i.e., when the number of sites in resonance with a given one becomes finite at large $N$. For the RP model this number of sites is on the order of $N^{1-\gamma/2}$, indicating that the localization transition occurs at $\gamma = 2$. Equivalently, we can understand the localization transition as occuring when the Thouless and Heisenberg times are of the same order. From our result for the spectral width we find that the Heisenberg time, which is the inverse mean level spacing, is

$$t_{\text{Heis}} \sim \min(N, N^{(1+\gamma)/2}). \tag{2.35}$$

The Thouless and Heisenberg times are of the same order when $\gamma = 2$, which matches our earlier reasoning for the localization transition through the Mott criterion. For $\gamma > 2$ the Thouless time is larger than the Heisenberg time. Since the SFF reaches its plateau value at the Heisenberg

timescale, there is no chance for random-matrix-like ramp to appear. The RP model will behave as if $H = A$. That is, Poissonian statistics will emerge. This analysis informs our determination of which values of $\gamma$ lead to which phase in the schematics of Fig. 2.1.

It's clear from the above discussion that the region $1 < \gamma < 2$ is particularly interesting because random-matrix-like behavior should be found, but only after a long period of time has elapsed. This case in which the Thouless time is larger than the inverse spectral width but smaller than the Heisenberg time is the nonergodic extended phase [89]. In this region eigenstates are not localized, but they are not spread sufficiently to be ergodic. For these values of $\gamma$ the calculation of the SFF is nontrivial because the behavior of the SFF depends not only on $\gamma$ but also the timescale. Going forward we will assume that $\gamma > 1$.

As a warm-up we calculate the Fourier transform of the density of states $\overline{C_1}(t)$ to leading order. Due to the independence of the eigenvalues of $A$ we can simplify Eq. (2.20) as

$$\overline{C_1}(t) = \frac{e^{-\sigma t^2/2}}{\tau} \oint_{\mathcal{C}} \frac{dz}{2\pi i} e^{itz} [g_1(z,\tau)]^N, \tag{2.36}$$

$$g_1(z,\tau) = 1 + \tau \left\langle \frac{1}{z - a'} \right\rangle, \tag{2.37}$$

where $a'$ is any eigenvalue of $A$. $\overline{C_1}(t)$ contains information about the average density of states. Expanding with the binomial theorem, we find

$$\frac{1}{N}\overline{C_1}(t) = \frac{1}{N} e^{-\sigma t^2/2} \oint_{\mathcal{C}} \frac{dz}{2\pi i} e^{itz} \sum_{j=0}^{N} \binom{N}{j} (i\sigma t)^{j-1} \left\langle \frac{1}{z - a'} \right\rangle^j. \tag{2.38}$$

Note that we have divided by $N$ to probe the average density of states but we have not performed the unfolding procedure. We want to examine this quantity on the scale of the spectral width, so

we let $t$ be order 1. For $\gamma > 1$, $N\sigma$ goes to 0, meaning only the $j = 1$ term remains. So

$$\frac{1}{N}\overline{C_1}(t) = \oint_{\mathcal{C}} \frac{dz}{2\pi i} e^{itz} \left\langle \frac{1}{z - a'} \right\rangle = \left\langle e^{ita'} \right\rangle . \tag{2.39}$$

This means that the average level density is the same as the probability density of the eigenvalues of $A$ [82].

As discussed above, we can neglect the disconnected contribution[1] so the unfolded SFF at non-zero times is determined entirely by $\overline{C_2}(t)$. In the RP model, Eq. (2.23) simplifies as

$$\overline{C_2}(t) = -\frac{e^{-\sigma t^2}}{\tau^2} \oint_{\mathcal{C}} \frac{dz}{2\pi i} \oint_{\mathcal{C}} \frac{dz'}{2\pi i} e^{it(z-z')} \left[ 1 - \left( \frac{\tau}{z' - z - \tau} \right)^2 \right] [g_2(z, \tau; z', -\tau)]^N, \tag{2.40}$$

$$\begin{aligned} g_2(z, \tau; z', -\tau) &= \left\langle \left( 1 + \frac{\tau}{z - a'} \right) \left( 1 - \frac{\tau}{z' - a'} \right) \right\rangle \\ &= 1 + \tau \left( 1 - \frac{\tau}{z' - z} \right) \left( \left\langle \frac{1}{z - a'} \right\rangle - \left\langle \frac{1}{z' - a'} \right\rangle \right) . \end{aligned} \tag{2.41}$$

Inserting Eq. (2.40) into Eq. (2.24) and neglecting the disconnected contribution, we obtain

$$C(t) = -\frac{e^{-\sigma t^2}}{N\tau^2} \oint_{\mathcal{C}} \frac{dz}{2\pi i} \oint_{\mathcal{C}} \frac{dz'}{2\pi i} e^{it(z-z')} \left\{ \left[ 1 - \left( \frac{\tau}{z' - z - \tau} \right)^2 \right] [g_2(z, \tau; z', -\tau)]^N \right\} . \tag{2.42}$$

We now make the change of variables

$$(z, z') = x \pm \frac{y}{2N} - i(q, q')0^+, \tag{2.43}$$

where $q, q' = \pm 1$, indicating which leg of the contour $\mathcal{C}$ the variables are on. Note that this

[1] Strictly speaking, the argument above for neglecting the disconnected contribution to the SFF only applies for $T \neq 0$, i.e., $t/N \neq 0$. Nonetheless, we make the same approximation for times that scale as a smaller (but still non-zero) power of $N$ as well. We find good agreement between the resulting expressions and numerical calculations for all timescales of interest.

definition of $y$ differs from that of Eq. (2.27) by a factor of $p(x)$, but it has the same scaling with $N$ and the argument for taking $y$ to be order 1 still applies. This change is made purely for convenience. With it we obtain

$$C(t) = -\frac{e^{-\sigma t^2}}{(N\tau)^2} \sum_{q,q'} qq' \int \frac{dx}{2\pi i} \int \frac{dy}{2\pi i} e^{iN^{-1}ty} \times \left[1 - \left(\frac{N\tau}{y - i(q-q')0^+ + N\tau}\right)^2\right] [g_2(z,\tau;z',-\tau)]^N. \quad (2.44)$$

### 2.4.1 The spectral form factor

Kunz and Shapiro [107] used the process outlined above to find the SFF when $\gamma = 2$ and $t \sim N$ (for which the Thouless and Heisenberg timescales coincide). They found that, if the coupling parameter is also unfolded by replacing $\lambda$ with $\Lambda/p(x)$, the two-level cluster function and the SFF become independent of the center energy $x$. Later Kravtsov et al. [89] generalized this result to all $\gamma > 1$ at the Thouless timescale. Additional timescales may be examined by rescaling time in this result. We first review this result and then present our direct calculation of the SFF at all timescales of interest. This calculation more closely parallels that of the SFF for the BRP model presented in section 2.5.

Making the change of variables in Eq. (2.43), we find that

$$\left\langle \frac{1}{z - a'} \right\rangle = \fint da' \frac{p(a')}{x - a'} + i\pi q p(x) + O(N^{-1}), \quad (2.45)$$

where the $\fint$ symbol indicates the principal value of the integral. Using this result in Eq. (2.41), we

determine that

$$g_2(z,\tau;z',-\tau) = 1 + \tau\left[i\pi p(x)(q-q') + O(N^{-1})\right]\left(1 + \frac{N\tau}{y - i(q-q')0^+}\right). \tag{2.46}$$

We examine the Thouless timescale by rescaling time as

$$t = N^{\gamma-1}s, \tag{2.47}$$

where $s$ is independent of $N$. Using this we can rewrite the above equation as

$$g_2(z,\tau;z',-\tau) = 1 - N^{-1}\pi\lambda^2 sp(x)(q-q')\left(1 + \frac{i\lambda^2 s}{y - i(q-q')0^+}\right) + O(N^{-2}). \tag{2.48}$$

With this we determine that, to leading order in $N^{-1}$,

$$[g_2(z,\tau;z',-\tau)]^N = \exp\left[-\pi\lambda^2 sp(x)(q-q')\left(1 + \frac{i\lambda^2 s}{y - i(q-q')0^+}\right)\right]. \tag{2.49}$$

Inserting this result into Eq. (2.44) and then unfolding by setting

$$v = s/p(x), \tag{2.50}$$

$$\Lambda = \lambda p(x) \tag{2.51}$$

yields the result that the Thouless-timescale SFF is

$$K_{\text{Th}}^{(\gamma)}(v) = 1 + e^{-2\pi\Lambda^2 v - N^{\gamma-2}\Lambda^2 v^2}\left[\frac{2I_1(\kappa v^{3/2})}{\kappa v^{3/2}} - \frac{1}{4\pi}\kappa v^{5/2} N^{\gamma-2}\int_0^\infty \frac{d\xi\,\xi}{\sqrt{\xi+1}} I_1(\kappa v^{3/2}\sqrt{\xi+1}) e^{-N^{\gamma-2}\Lambda^2 v^2 \xi}\right], \quad (2.52)$$

where $\kappa = \sqrt{8\pi\Lambda^4 N^{\gamma-2}}$ and $I_1(x)$ is the modified Bessel function of the first kind. See Ref. [89] for full details. Note that we have unfolded using the probability density instead of the level density, which differ by a factor of $N$. Eq. (2.52) may also be used to examine other timescales by rescaling $v$. However, we note that the intermediate step Eq. (2.49) only holds for $s \ll \sqrt{N}$, that is when $t \ll N^{\gamma-1/2}$. Even restricting ourselves to times not larger than the Heisenberg time ($t \not\gg N$), we see that this is not always the case.

We now present a direct calculation of the SFF for the RP model at all relevant timescales. We make the same change of variables $\{z, z'\} \to \{x, y\}$ shown in Eq. (2.43), but we will not set the $N$-dependence of $t$ yet. We can now write

$$C(t) = \int dx\, p(x) e^{-\sigma t^2}(S_1 + S_2), \quad (2.53)$$

$$S_1 = -\frac{1}{2\pi i p(x)(N\tau)^2}\sum_{q,q'} qq' \int \frac{dy}{2\pi i} e^{iN^{-1}ty}[g_2(z,\tau;z',-\tau)]^N, \quad (2.54)$$

$$S_2 = \frac{1}{2\pi i p(x)}\sum_{q,q'} qq' \int \frac{dy}{2\pi i} e^{iN^{-1}ty}\frac{[g_2(z,\tau;z',-\tau)]^N}{[y - i(q-q')0^+ + N\tau]^2}. \quad (2.55)$$

Instead of seeking an asymptotic form for $[g(z,\tau;z',-\tau)]^N$ as in Eq. (2.49), we write

$$\begin{aligned}[g_2(z,\tau;z',-\tau)]^N &= \sum_{j=0}^{N}\binom{N}{j}\frac{\{N\tau^2\left[i\pi p(x)(q-q')+O(N^{-1})\right]\}^j}{\{1+\tau\left[i\pi p(x)(q-q')+O(N^{-1})\right]\}^{j-N}\left[y-i(q-q')0^+\right]^j}\\ &= \sum_{j=0}^{N} p_{\text{bi}}\left(N,j,i\pi\tau p(x)(q'-q)+O(N^{-1}\tau)\right)\left(\frac{N\tau}{y-i(q-q')0^+}\right)^j,\end{aligned} \tag{2.56}$$

where

$$p_{\text{bi}}(N,j,x)=\binom{N}{j}(1-x)^{N-j}x^j \tag{2.57}$$

is the binomial probability density function. We take $i\tau(q'-q)$ to be positive because, when it is not, the integrals over $y$ in Eqs. (2.54)-(2.55) vanish due to a lack of singularities in the half of the complex plane in which the contour of integration may be closed. At this point we restrict ourselves to times not larger than the Heisenberg time, so $|\tau|=\sigma t\ll 1$. This is not a problem since we already know that the SFF goes to 1 for times larger than the Heisenberg time. With this restriction, we see by inspection of the first line of Eq. (2.56) that we can neglect the $O(N^{-1})$ terms when $j\ll N$. The probability density $p_{\text{bi}}(N,j,x)$ is peaked near its mean at $j=Nx$ with a variance of $Nx(1-x)$. Far from the peak the probability density is exponentially small in $N$. So only the $j\sim N|\tau|$ terms will contribute in Eq. (2.56). This means that the $j\sim N$ terms will not contribute and we can therefore neglect the $O(N^{-1})$ terms. With these considerations we can write, for positive $i\tau(q'-q)$,

$$[g_2(z,\tau;z',-\tau)]^N=\sum_{j=0}^{N} p_{\text{bi}}\left(N,j,i\pi\tau p(x)(q'-q)\right)\left(\frac{N\tau}{y-i(q-q')0^+}\right)^j. \tag{2.58}$$

Inserting this into Eq. (2.54) we find

$$S_1 = \frac{1}{N}\sum_q \sum_{j=1}^{N} \binom{N}{j} \frac{(2\pi i N \tau^2 p(x))^{j-1}}{(1+2\pi i \tau q p(x))^{j-N}} \int \frac{dy}{2\pi i} \frac{e^{iN^{-1}tqy}}{(y-i0^+)^j}, \tag{2.59}$$

where we have made the change of variable $y \to qy$. Note that the $j = 0$ term vanishes due to the integrand having no singularities. For the remaining terms the integral over $y$ vanishes when $tq < 0$, so we can remove the sum over $q$ by making the replacement $tq \to |t|$. Performing the integral over $y$, we find that

$$S_1 = \sum_{j=0}^{N-1} \frac{(|\tau||t|)^j}{(j+1)!} p_{\text{bi}}\left(N-1, j, 2\pi|\tau|p(x)\right). \tag{2.60}$$

We now turn our attention to $S_2$. Inserting Eq. (2.56) into Eq. (2.55), we find

$$S_2 = \frac{i}{2\pi p(x)}\sum_q \sum_{j=1}^{N} \binom{N}{j} \frac{(2\pi i N \tau^2 p(x))^{j}}{(1+2\pi i \tau q p(x))^{j-N}} \int \frac{dy}{2\pi i} \frac{e^{iN^{-1}tqy}}{(y+N\tau q)^2 (y-i0^+)^j}, \tag{2.61}$$

where we have again made the change of variable $y \to qy$. The $j = 0$ term vanishes because the integrand has no singularities in the half of the complex plane in which the contour of integration may be closed. For the remaining terms the integral over $y$ vanishes when $tq < 0$, so we again remove the sum over $q$ by making the replacement $tq \to |t|$. Performing the integral over $y$, we find that

$$S_2 = -e^{\sigma t^2} \sum_{j=0}^{N-1} p_{\text{bi}}\left(N-1, j, 2\pi|\tau|p(x)\right) \left[\tilde{\Gamma}(j+2, \sigma t^2) - \frac{\sigma t^2}{j+1}\tilde{\Gamma}(j+1, \sigma t^2)\right], \tag{2.62}$$

where

$$\tilde{\Gamma}(j+1,z) = \frac{1}{j!}\int_z^\infty d\xi\, \xi^j e^{-\xi} = e^{-z}\sum_{k=0}^{j}\frac{z^k}{k!} \tag{2.63}$$

is the regularized upper incomplete gamma function. Note that $\tilde{\Gamma}$ satisfies the recurrence relation

$$\tilde{\Gamma}(j+1,z) = \tilde{\Gamma}(j,z) + \frac{z^j e^{-z}}{j!}. \tag{2.64}$$

We take our results for $S_1$ and $S_2$ in Eqs. (2.60) and (2.62) and insert them into Eq. (2.53), then extract the SFF using Eq. (2.29). Then unfolding time by setting

$$u = |t|/p(x) = N|T| \tag{2.65}$$

and unfolding $\lambda$ using Eq. (2.51), we obtain the SFF

$$K^{(\gamma)}(u/N) = 1 + \sum_{j=0}^{N-1} p_{\text{bi}}(N-1,j,2\pi N^{-\gamma}\Lambda^2 u)\left[\frac{e^{-N^{-\gamma}\Lambda^2u^2}}{(j+1)!}\left(N^{-\gamma}\Lambda^2u^2\right)^j\right.$$
$$\left. -\tilde{\Gamma}\left(j+2,N^{-\gamma}\Lambda^2u^2\right) + \frac{N^{-\gamma}\Lambda^2u^2}{j+1}\tilde{\Gamma}\left(j+1,N^{-\gamma}\Lambda^2u^2\right)\right]. \tag{2.66}$$

With this result we have generalized the Thouless-timescale result of Eq. (2.52) to all times larger than the inverse spectral width. Although it was derived considering only times not larger than the Heisenberg time, it continues to hold for later times, as we will discuss in the following section.

Fig. 2.4 plots Eq. (2.66) for $\gamma = 1.6$ and various values of the coupling parameter $\Lambda$ and compares to numerical results obtained through exact diagonalization. There is good agreement between theory and numerics at not-too-early times (note in particular that the agreement is already good by the Thouless timescale). The discrepancy at early times which are still larger than

the inverse spectral width is a result of calculating the numerical SFF using only the eigenvalues within a finite window in order to perform the unfolding procedure. At times earlier than this inverse window width the SFF probes correlations between eigenvalues at separations larger than the width of the window, but these eigenvalues are cut off in the numerical unfolding procedure. We observe that the SFF decays from the Poissonian result to the GUE SFF over time, with the rate of decay controlled by $\Lambda$. As we vary $\Lambda$ we can see the SFF moving between the behaviors we predicted schematically in Fig. 2.1. For small values of $\Lambda$ we see that the SFF reaches its plateau at times greater (but not much greater) than the Heisenberg time. The agreement between theory and numerics is still very good at these late times.

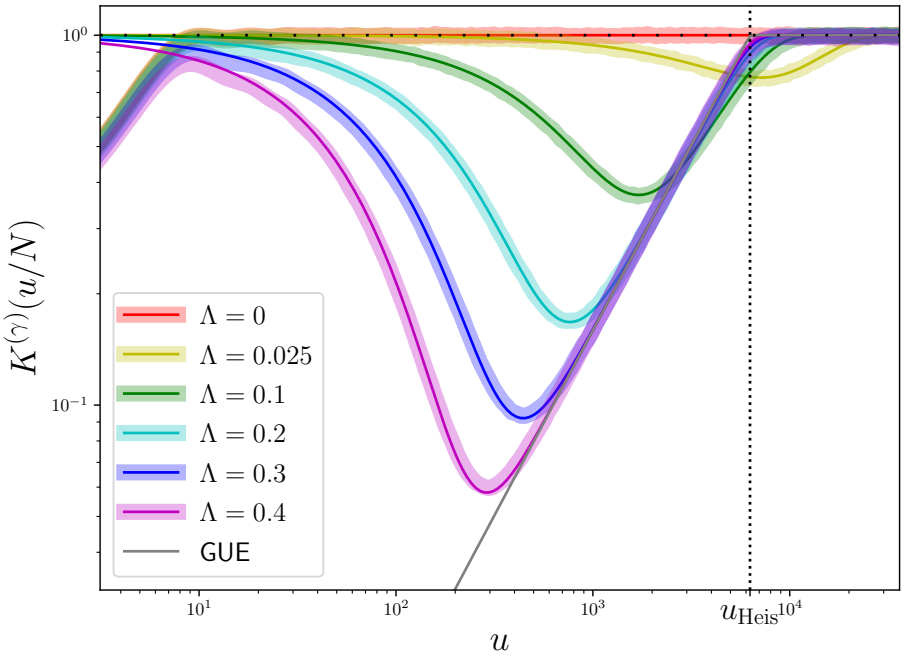



Figure 2.4: The SFF of the RP model when $\gamma = 1.6$ for several values of the coupling parameter $\Lambda$. The dark thin lines show the analytical results while the lighter, thicker lines show numerical results. The dotted line indicates the Heisenberg time. In the numerics the diagonal elements of $A$ are drawn from the Gaussian distribution with 0 mean and a variance of 1. The numerics are obtained through exact diagonalization of size $N = 1000$ random matrices and are averaged over $10^5$ realizations. The numerics are restricted to eigenvalues within the window [-1,1].

If we examine the Thouless timescale by setting $u = N^{\gamma-1}v$ in Eq. (2.66) we can write

$$K_{\text{Th}}^{(\gamma)}(v) = K^{(\gamma)}(N^{\gamma-2}v) = 1 + e^{-2\pi\Lambda^2 v - N^{\gamma-2}\Lambda^2 v^2}\frac{2I_1(\kappa v^{3/2})}{\kappa v^{3/2}} - B_{\text{Th}}^{(\gamma)}(v), \tag{2.67}$$

$$B_{\text{Th}}^{(\gamma)}(v) = e^{-2\pi\Lambda^2 v}\sum_{j=0}^{\infty}\frac{(2\pi\Lambda^2 v)^j}{j!}\left(\tilde{\Gamma}(j+2,\Lambda^2 N^{\gamma-2}v^2) + \frac{\Lambda^2 N^{\gamma-2}v^2}{j+1}\tilde{\Gamma}(j+1,\Lambda^2 N^{\gamma-2}v^2)\right), \tag{2.68}$$

where we have discarded some vanishing terms and used the fact that the modified Bessel function may be written as the infinite series

$$I_1(x) = \sum_{j=0}^{\infty}\frac{(x/2)^{2j+1}}{j!(j+1)!}. \tag{2.69}$$

Writing the regularized upper incomplete gamma functions in the integral form shown in Eq. (2.63), using Eq. (2.69), and making the change of variable $\xi \to \sqrt{\xi+1}$, we recover the Thouless-timescale SFF shown in Eq. (2.52).

### 2.4.2 The infinite matrix size limit

We can now determine the large $N$ limit of the SFF at different timescales. We begin by examining the Thouless timescale where $u \sim N^{\gamma-1}$. When $\gamma > 2$ the regularized incomplete gamma functions in Eq. (2.68) vanish, so

$$B_{\text{Th}}^{(\gamma)}(v) = 0. \tag{2.70}$$

Inserting this result in Eq. (2.67) and noting that the second term of the SFF vanishes for large $N$, we find that the SFF is

$$K_{\text{Th}}^{(\gamma)}(v) = 1, \tag{2.71}$$

indicating localized behavior.

When $\gamma < 2$ the regularized incomplete gamma functions become complete and go to 1. Resumming then gives

$$B_{\text{Th}}^{(\gamma)}(v) = 1 - \frac{T}{2\pi}\left(1 - e^{-2\pi\Lambda^2 v}\right), \tag{2.72}$$

where $T = t/\rho(x) = N^{\gamma-2}v$ is the fully unfolded (including all factors of $N$) time. We again insert this in Eq. (2.67). We note that the second term of the SFF goes to $e^{-2\pi\Lambda^2 v}$, so the SFF for $1 < \gamma < 2$ is

$$K_{\text{Th}}^{(\gamma)}(v) = e^{-2\pi\Lambda^2 v} + \frac{T}{2\pi}\left(1 - e^{-2\pi\Lambda^2 v}\right). \tag{2.73}$$

The second term is subleading, but we include it to make the crossover to GUE behavior for large $\Lambda$ apparent. $K_{\text{Th}}^{(2)}(v)$ does not have a simple limit.

Pulling together all the results for the Thouless timescale SFF at different values of $\gamma$ we have

$$K_{\text{Th}}^{(\gamma)}(v) = \begin{cases} 1, & \gamma > 2 \\ 1 + e^{-\Lambda^2 v(v+2\pi)}\frac{2I_1(\kappa v^{3/2})}{\kappa v^{3/2}} - B_{\text{Th}}^{(2)}(v), & \gamma = 2 \\ e^{-2\pi\Lambda^2 v} + \frac{T}{2\pi}\left(1 - e^{-2\pi\Lambda^2 v}\right), & 1 < \gamma < 2 \end{cases}. \tag{2.74}$$

When $1 < \gamma \leq 2$ the behavior of the SFF is dependent on $\Lambda$. It moves to the Poissonian result for small $\Lambda$ and the GUE result for large $\Lambda$. This can be seen explicitly in the above equation when

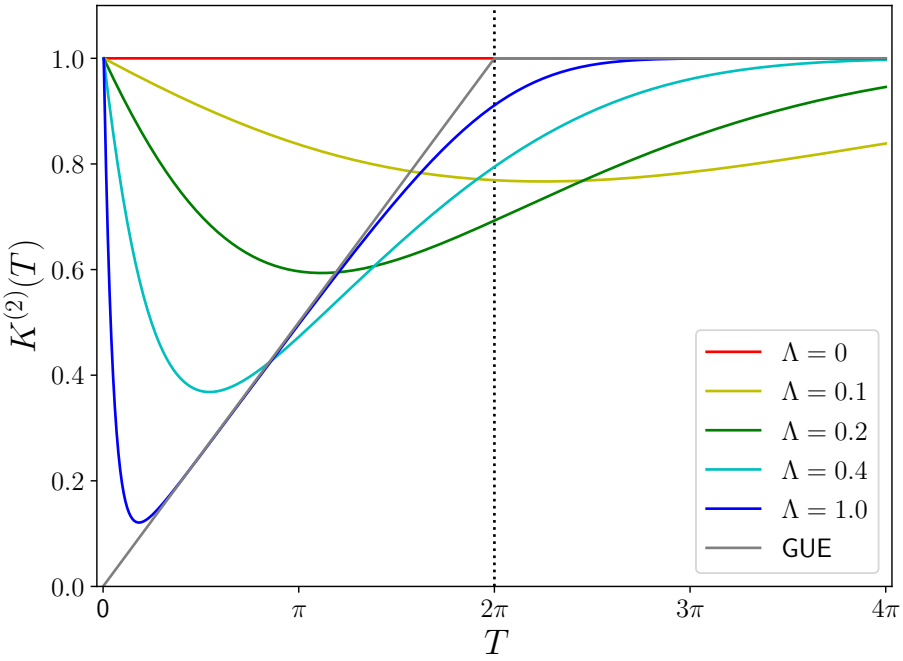



Figure 2.5: The SFF for the RP model when the Thouless and Heisenberg timescales coincide for various values of the coupling parameter $\Lambda$. The dotted vertical line indicates the Heisenberg time.

$1 < \gamma < 2$.

It is not surprising that the SFF does not have a simple limit when $\gamma = 2$ since, in this case, the Thouless and Heisenberg timescales are the same. By plotting this case, shown in Fig. 2.5, we can observe that there is still a crossover from the Poissonian to the GUE result controlled by $\Lambda$. An interesting feature of this crossover is that, for finite $\Lambda$, the SFF saturates at times later than the Heisenberg time. There is a trade-off: increasing the level repulsion at large energy separations decreases the level repulsion at small separations and vice versa. This phenomenon can be shown through the fact that the area between the SFF and its plateau value is independent of the coupling strength for any nonzero coupling. This property is demonstrated in the Appendix.

We can now examine the large $N$ limit of the SFF at other timescales. For convenience we

rewrite the SFF of Eq. (2.66) as

$$K^{(\gamma)}(u/N) = 1 + e^{-N^{-\gamma}\Lambda^2 u^2} \sum_{j=0}^{N-1} \frac{(N^{-\gamma}\Lambda^2 u^2)^j}{(j+1)!} p_{\text{bi}}(N-1, j, 2\pi N^{-\gamma}\Lambda^2 u) - B^{(\gamma)}(u), \quad (2.75)$$

$$\begin{aligned} B^{(\gamma)}(u) = \sum_{j=0}^{N-1} & p_{\text{bi}}(N-1, j, 2\pi N^{-\gamma}\Lambda^2 u) \\ & \times \left[ \tilde{\Gamma}\left(j+2, N^{-\gamma}\Lambda^2 u^2\right) - \frac{N^{-\gamma}\Lambda^2 u^2}{j+1} \tilde{\Gamma}\left(j+1, N^{-\gamma}\Lambda^2 u^2\right) \right]. \end{aligned} \quad (2.76)$$

The regularized upper incomplete gamma function $\tilde{\Gamma}(j, x)$ is the cumulative distribution function for a Poissonian random variable with parameter $x$. This means that these functions in Eq. (2.76) have a crossover from 0 to 1, where the location of the crossover is at approximately $j = N^{-\gamma}\Lambda^2 u^2$ and the width of the crossover is on the order of $N^{-\gamma/2}\Lambda u$. Meanwhile the binomial probability density function in Eq. (2.76) has a mean and variance of approximately $2\pi N^{1-\gamma}\Lambda^2 u$ for large $N$ and $u$ not larger than order $N$.

In the following considerations we will assume that the crossover point of the upper incomplete gamma functions and the mean of the binomial distribution are much larger than the crossover width and the binomial distribution width. Fortunately, for the timescales we are considering, this is only violated when $u \sim N$ and $\gamma = 2$. This is the case where the Thouless and Heisenberg timescales coincide, which we have already considered above.

The binomial distribution is exponentially small far from its mean, so only the terms with $j$ close to $2\pi N^{1-\gamma}\Lambda^2 u$ may not vanish. If the mean of the binomial distribution is less than the crossover point of the regularized upper incomplete gamma functions, which occurs when $u > 2\pi N$ and $\gamma < 2$, we see from Eq. (2.63) that those functions will be exponentially small and can therefore be replaced with 0. If the mean of the binomial distribution is greater than

the crossover point, which occurs when $u < 2\pi N$, they can be replaced with 1. We can also observe that when $\gamma > 2$, the regularized incomplete gamma functions will become complete and approach 1. With these considerations we find, after performing the sum over $j$, that for $u \not\gg N$

$$B^{(\gamma)}(u) = \begin{cases} 1 - \frac{u}{2\pi N}\left[1 - (1 - 2\pi N^{-\gamma}\Lambda^2 u)^N\right], & u < 2\pi N \text{ or } \gamma > 2 \\ 0, & u > 2\pi N \text{ and } \gamma < 2 \end{cases}. \tag{2.77}$$

We now seek to find the large $N$ limit of the SFF for times smaller than the Thouless time and not larger than the Heisenberg time, that is when $u \ll t_{\text{Th}} \sim N^{\gamma-1}$ and $u \not\gg t_{\text{Heis}} \sim N$. Under these conditions we find that $e^{-N^{-\gamma}\Lambda^2 u^2} \to 1$ and the mean and variance of the binomial density in Eq. (2.75) go to 0, leaving only the $j = 0$ contribution nonvanishing in the second term. We also find that $(1 - 2\pi N^{-\gamma}\Lambda^2 u)^N \to 1$. Using these limits in Eqs. (2.75) and (2.77) we find that the SFF is

$$K^{(\gamma)}(T) = 1. \tag{2.78}$$

This matches the result for the SFF at times much later than the Heisenberg time, so this result is valid for all times smaller than the Thouless time but larger than the inverse spectral width.

At these times the SFF indicates Poissonian statistics. This can also be determined directly from Eqs. (2.19) and (2.22). When $\sigma t^2$ and $N\tau$ are both less than order 1, which is the case for times smaller than the Thouless time and not larger than the Heisenberg time, we find in the

$N \to \infty$ limit that

$$\overline{C_1}(t) = \left\langle \sum_k e^{ita_k} \right\rangle, \ \overline{C_2}(t) = \left\langle \sum_{k \neq l} e^{it(a_k - a_l)} \right\rangle, \tag{2.79}$$

indicating that the SFF is simply that of $A$. Since the eigenvalues of $A$ are uncorrelated, Poissonian statistics are found.

We now seek to find the large $N$ limit of the SFF for times larger than the Thouless time but not larger than the Heisenberg time, that is when $u \gg t_{\text{Th}} \sim N^{\gamma-1}$ and $u \not\gg t_{\text{Heis}} \sim N$. We note that the maximum value of the binomial probability density is on the order of the inverse of its standard deviation. This tells us that

$$\sum_{j=0}^{N-1} \frac{(N^{-\gamma}\Lambda^2 u^2)^j}{(j+1)!} p_{\text{bi}}(N-1, j, 2\pi N^{-\gamma}\Lambda^2 u) \ll \sum_{j=0}^{\infty} \frac{(N^{-\gamma}\Lambda^2 u^2)^j}{j!} = e^{N^{-\gamma}\Lambda^2 u^2}. \tag{2.80}$$

Therefore the second term in Eq. (2.75) vanishes. We also note that $(1 - 2\pi N^{-\gamma}\Lambda^2 u)^N \to 0$. With these considerations we find that the SFF is

$$K^{(\gamma)}(T) = \min\left(\frac{T}{2\pi}, 1\right). \tag{2.81}$$

This result for the SFF is equal to 1 for all times later than the Heisenberg time, so it is valid for all times larger than the Thouless time and inverse spectral width. We have recovered the GUE SFF. We know that GUE statistics hold for $\gamma < 1$, so we can conclude that this result is valid for all values of $\gamma$.

Finally we examine the large $N$ behavior of the SFF at times larger than the Heisenberg time, where $u \gg t_{\text{Heis}} \sim N$. Although we derived the SFF assuming earlier times, the numerical

results shown in Fig. 2.4 indicate that our result may also capture the plateau behavior. The way to determine the large $N$ behavior at these late times depends on $\gamma$. When $\gamma > 2\log_N u$ we see that the mean and variance of the binomial distribution in Eqs. (2.75)-(2.76) vanish, meaning the distribution goes to $\delta_{j0}$. Additionally, the regularized incomplete gamma functions become complete in this case and go to 1. Making these replacements we see that the SFF goes to 1. When $\log_N u < \gamma < 2\log_N u$ the crossovers of the regularized incomplete gamma functions are much larger than the mean of the binomial distribution, so the gamma functions may be taken to be 0, meaning $B^{(\gamma)}(u)$ vanishes. Eq. (2.80) holds in this case, so the second term in Eq. (2.75) also vanishes, showing that the SFF goes to 1. In the last case, where $1 < \gamma < \log_N u$, the function $p_{\text{bi}}(N-1, j, 2\pi N^{-\gamma}\Lambda^2 u)$ can no longer be considered a probability distribution since its parameter is larger than 1, causing it to take on negative values for odd $j$. Because $N^{-\gamma}\Lambda^2 u^2 \gg N$ in this case, we see from Eq. (2.63) that the regularized incomplete gamma functions are dominated by their highest order terms. Taking the summand in Eq. (2.66) to leading order we find that the SFF is

$$K^{(\gamma)}(u/N) = 1 - 2e^{-N^{-\gamma}\Lambda^2 u^2} \sum_{j=0}^{N-1} p_{\text{bi}}(N-1, j, 2\pi N^{-\gamma}\Lambda^2 u) \frac{j}{(j+1)!} (N^{-\gamma}\Lambda^2 u^2)^{j-1}. \quad (2.82)$$

The sum in the above expression is also dominated by its highest order term, which grows slower than the exponential factor vanishes. Again the SFF goes to 1. These considerations show that the plateau behavior is recovered for all times much larger than the Heisenberg time. This accounts for the good agreement between the theory and numerics at these late times.

We can visualize the results of Ref. [89] and the new results found here with a diagram for the behavior of the SFF at different timescales and values of $\gamma$, shown in Fig. 2.6. The SFF is 1

before the Thouless timescale and equal to the GUE SFF after. At the Thouless timescale, when it is not greater than the Heisenberg time, there is a crossover between these two behaviors. At times larger than the Heisenberg time the SFF is simply 1.

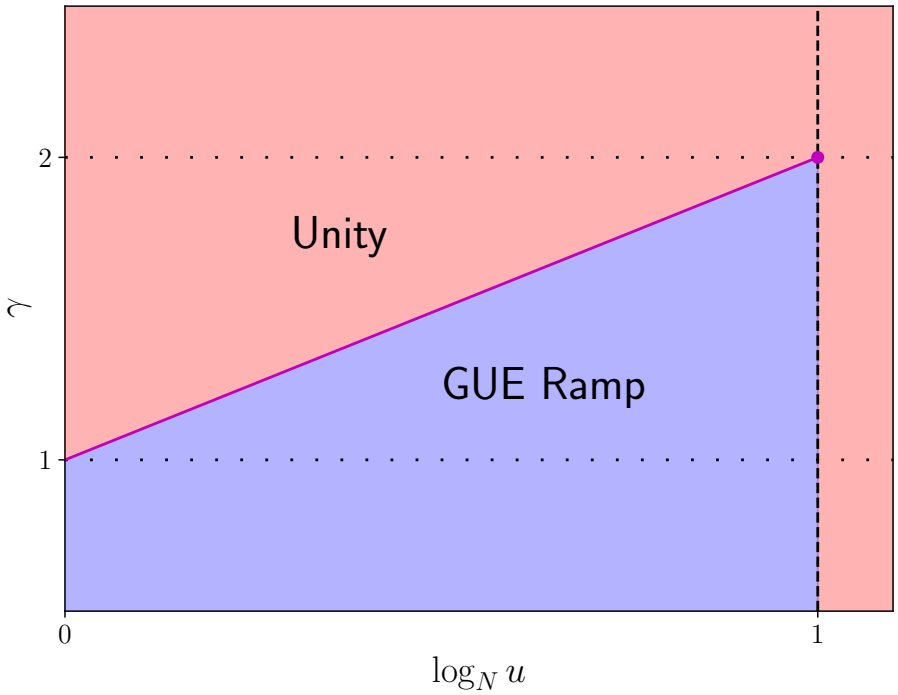



Figure 2.6: The SFF of the RP model at different timescales and values of the coupling parameter $\gamma$. The magenta dot indicates the point where the Thouless and Heisenberg timescales coincide, where a closed form expression for the SFF is not known. The dotted lines indicate the values of $\gamma$ at which there is a phase transition. The dashed line indicates the Heisenberg time.

## 2.5 The block Rosenzweig-Porter model

Motivated by our recent results for the early time ramp of the SFF for a quantum spin glass [2], we construct a minimal quantum model which can also exhibit glassy behavior. We create this model by generalizing the RP model examined in the previous section. The diagonal matrix $A$ in Eq. (2.32) is redefined as a block-diagonal matrix with $P$ independent GUE matrices of size $M = N/P$. Each block in this matrix corresponds to a single sector and the fact that the

block is a GUE matrix means that the system is ergodic within that sector. The Hamiltonian is the sum of this block-diagonal matrix and a size $N$ GUE matrix which couples the sectors.

More formally, the Hamiltonian matrix for the BRP model has the form of Eq. (2.8), but now with the condition that the eigenvalues of $A$ are those of $P$ independent size $M$ GUE matrices drawn from the distribution

$$p\left(A^{(i)}\right) \sim \exp\left[-\frac{M}{2}\,\mathrm{tr}\left(A^{(i)}\right)^2\right]. \tag{2.83}$$

This means that the spectral width of $A$ is order 1 and, if $M$ is large, the eigenvalues of $A$ cannot have magnitude greater than 2 [70]. We can consider $A$ to be block diagonal such that

$$A = \bigoplus_{i=1}^{P} A^{(i)}. \tag{2.84}$$

Similar to our analysis of the RP model, we can use the Fermi golden rule to determine the Thouless timescale and determine the locations of the transitions by comparing to the other cardinal timescales. We are interested in the escape rate from a particular site to sites outside that starting site's sector. The number of such sites to escape to is of order $N$, the coupling is of order $N^{-\gamma/2}$, and the spectral width of $A$ is order 1. It follows that the RP results for the Thouless time, the spectral width, and the Heisenberg time, shown in Eqs. (2.33)-(2.35) respectively, are valid for the BRP model also.

The localization transition occurs when the Thouless and Heisenberg times are of the same order, which is when $\gamma = 2$. Notice that the block-diagonal structure of $A$ leads to a localized phase in the BRP model which is distinct from that of the RP model. Within this phase the system

will be confined to the subspaces corresponding the the blocks of $A$ instead of individual states. The transition to full GUE statistics occurs when the Thouless time is of the same order as the inverse spectral width. This occurs when $\gamma = 1$.

For the BRP model there is another timescale of interest—the block Heisenberg time $t_{\text{Heis,Bl}} = t_{\text{Heis}}/P$. This is the inverse mean level spacing for a single block of $A$. In the large $\gamma$ limit where the coupling between the blocks vanishes, $t_{\text{Heis,Bl}}$ is the time at which the SFF will reach its plateau value. Comparing the Thouless time to the block Heisenberg time, we see that there will be another transition at $\gamma = 1 + d$, where $d = \log_N M$, which is when these timescales are of the same order. This analysis informs our determination of which values of $\gamma$ lead to which phase in the schematics of Fig. 2.2.

Two of the transitions we have discussed have an important interpretation in terms of the size of the support set of the energy eigenstates. Starting with an unperturbed energy eigenstate of $A$, which is spread over $M$ sites in a single block, the perturbation from the $V$ matrix causes each of those $M$ sites to hybridize with other sites within an energy interval on the order of $E_{\text{Th}} \sim t_{\text{Th}}^{-1}$. Assuming this interval is not smaller than the mean level spacing $\delta = w/N$, the number of sites in this interval is $E_{\text{Th}}/\delta$. This means that the size of the support set for the energy eigenstates is on the order of $ME_{\text{Th}}/\delta = N^{2-\gamma+d}$. This is of order $M$ when $\gamma = 2$, indicating the localization transition. For larger $\gamma$, $E_{\text{Th}}$ becomes smaller than $\delta$ and no hybridization occurs. The size of the support set is of order $N$ when $\gamma = 1 + d$, indicating delocalization over the entire Hilbert space. It is an interesting feature of the BRP model that this transition is, in general, separate from the transition to full GUE spectral statistics at $\gamma = 1$.

We note that if the block size $M$ is of order 1, the BRP model becomes equivalent to the RP model. In this case the full eigenstate delocalization transition occurs at $\gamma = 1$ as it does for

the RP model. In terms of the SFF, the eigenvalues of $A$ are uncorrelated on timescales larger than the block Heisenberg time. When $M$ is of order 1 the block Heisenberg time is of the same order as the inverse spectral width. Since the eigenvalues of $A$ are uncorrelated for the relevant timescales, the BRP model reduces to the RP model. Going forward we will consider $M$ to be larger than order 1.

We may now turn our attention to the calculation of the SFF for the BRP model. We will once again assume in our calculations that $\gamma > 1$. With the additional information about the structure of $A$ we can simplify Eqs. (2.20) and (2.23) for $\overline{C_1}(t)$ and $\overline{C_2}(t)$ respectively. Each eigenvalue of $A$ will be correlated only with the eigenvalues from the same block so

$$\overline{C_1}(t) = \frac{e^{-\sigma t^2/2}}{\tau} \oint_{\mathcal{C}} \frac{dz}{2\pi i} e^{itz} [Z_1(z,\tau)]^P, \tag{2.85}$$

$$Z_1(z,\tau) = \left\langle \frac{\det(z+\tau-A')}{\det(z-A')} \right\rangle, \tag{2.86}$$

and

$$\overline{C_2}(t) = -\frac{e^{-\sigma t^2}}{\tau^2} \oint \frac{dz}{2\pi i} \oint \frac{dz'}{2\pi i} e^{it(z-z')} \left[ 1 - \left( \frac{\tau}{z'-z-\tau} \right)^2 \right] [Z_2(z,\tau;z',-\tau)]^P, \tag{2.87}$$

$$Z_2(z,\tau;z',-\tau) = \left\langle \frac{\det(z+\tau-A')\det(z'-\tau-A')}{\det(z-A')\det(z'-A')} \right\rangle, \tag{2.88}$$

where $A'$ is any of the $A^{(i)}$. The functions $Z_1$ and $Z_2$ are actually generating functions. The ensemble averaged Green's function for a size $M$ GUE matrix is

$$\langle G(z) \rangle = \left\langle \mathrm{tr} \frac{1}{z-A'} \right\rangle = \left[ \frac{\partial}{\partial \tau} Z_1(z,\tau) \right]_{\tau=0} \tag{2.89}$$

and the correlator of two Green's functions is

$$\langle G(z)G(z')\rangle = \left\langle \mathrm{tr}\,\frac{1}{z-A'}\,\mathrm{tr}\,\frac{1}{z'-A'}\right\rangle = \left[\frac{\partial^2}{\partial\tau\partial\tau'}Z_2(z,\tau;z',\tau')\right]_{\tau,\tau'=0}. \tag{2.90}$$

We now consider whether the condition on the average level density which we derived in Eq. (2.39) for the RP model holds for the BRP model also. From Eq. (2.86) we find that

$$Z_1(z,\tau) = 1 + M\tau\left\langle\frac{1}{z-a'}\right\rangle + O(\tau^2). \tag{2.91}$$

Using this in Eq. (2.85) and dividing by $N$ while letting $t$ be order 1, the same order as the inverse spectral width, we find that Eq. (2.39) does indeed hold for the BRP model. This means that, when $\gamma > 1$, the average level density is equal to the probability density for the eigenvalues of $A$, which is the average level density for size $M$ GUE matrices. For infinitely large $M$ this is

$$p(a) = \begin{cases} \frac{1}{2\pi}\sqrt{4-a^2}, & |a| \le 2 \\ \\ 0, & |a| > 2 \end{cases}. \tag{2.92}$$

In this section we will make the change of variables

$$(z,z') = x \pm \frac{y}{2N} + i(q,q')0^+. \tag{2.93}$$

This differs from the change of variables of Eq. (2.43) only in the signs of $q$ and $q'$, which is done

purely for convenience. With this change of variables we see that

$$C(t) = -\frac{e^{-\sigma t^2}}{(N\tau)^2} \sum_{qq'} qq' \int \frac{dx}{2\pi i} \int \frac{dy}{2\pi i} e^{iN^{-1}ty} \times \left[1 - \left(\frac{N\tau}{y + N\tau + i(q-q')0^+}\right)^2\right] [Z_2(z, \tau; z', -\tau)]^P. \quad (2.94)$$

As we did for the calculation of the SFF for the RP model, we are neglecting the contribution from the disconnected part of the two-level correlation function under the assumption that correlations between energy levels with separations much larger than the mean level spacing will vanish.

### 2.5.1 The two-point generating function

In this section we calculate the two-point generating function

$$Z_2(z, \tau; z', -\tau) = \left\langle \frac{\det(z + \tau - A') \det(z' - \tau - A')}{\det(z - A') \det(z' - A')} \right\rangle, \quad (2.95)$$

where $A'$ is a GUE matrix of size $M$ drawn from the distribution

$$p(A') \sim \exp\left(-\frac{M}{2} \operatorname{tr} A'^2\right). \quad (2.96)$$

We begin by writing the determinants in the numerator as fermionic integrals and the determinants in the denominator as bosonic integrals. Doing so, we obtain

$$Z_2(z,\tau;z',-\tau) = \pi^{-2M}\left\langle \int D(\bar{\psi},\psi)\exp\sum_{\zeta}\sum_{\alpha}\left\{\sum_i \bar{\psi}^{\alpha}_{\zeta i}A'_{ii}\psi^{\alpha}_{\zeta i}\right.\right.$$
$$+\sum_{i>j}\left(\bar{\psi}^{\alpha}_{\zeta i}\psi^{\alpha}_{\zeta j}+\bar{\psi}^{\alpha}_{\zeta j}\psi^{\alpha}_{\zeta i}\right)\Re(A'_{ij})+i\sum_{i>j}\left(\bar{\psi}^{\alpha}_{\zeta i}\psi^{\alpha}_{\zeta j}-\bar{\psi}^{\alpha}_{\zeta j}\psi^{\alpha}_{\zeta i}\right)\Im(A'_{ij})$$
$$\left.\left.-\left[\left(z^{(\alpha)}-(-1)^{\alpha}\tau\right)\delta_{\zeta F}+z^{(\alpha)}\delta_{\zeta B}\right]\sum_i \bar{\psi}^{\alpha}_{\zeta i}\psi^{\alpha}_{\zeta i}\right]\right\}\right\rangle, \quad (2.97)$$

where $\alpha = 1,2$, $(z^{(1)},z^{(2)}) = (z,z')$, and $\zeta = B,F$, indicating bosonic and fermionic fields respectively.

Performing the GUE average over $A'$, we obtain the following mean-field sigma model:

$$Z_2(z,\tau;z',-\tau) = \pi^{-2M}\int D(\bar{\psi},\psi)\exp\left\{\frac{1}{2M}\operatorname{str}\tilde{\Omega}\right.$$
$$\left.-\sum_{\zeta}\sum_{\alpha}\sum_i\left[\left(z^{(\alpha)}-(-1)^{\alpha}\tau\right)\delta_{\zeta F}+z^{(\alpha)}\delta_{\zeta B}\right]\bar{\psi}^{\alpha}_{\zeta i}\psi^{\alpha}_{\zeta i}\right\}, \quad (2.98)$$

$$\tilde{\Omega} = \begin{pmatrix}\tilde{\Omega}_{BB} & \tilde{\Omega}_{BF}\\ \tilde{\Omega}_{FB} & \tilde{\Omega}_{FF}\end{pmatrix},\ \tilde{\Omega}_{\zeta\zeta'} = \begin{pmatrix}\sum_i \psi^1_{\zeta i}\psi^1_{\zeta' i} & \sum_i \psi^1_{\zeta i}\psi^2_{\zeta' i}\\ \sum_i \psi^2_{\zeta i}\psi^1_{\zeta' i} & \sum_i \psi^2_{\zeta i}\psi^2_{\zeta' i}\end{pmatrix}, \quad (2.99)$$

where $\tilde{\Omega}$ is a $4\times 4$ matrix and $\operatorname{str}\tilde{\Omega} = \operatorname{tr}\tilde{\Omega}_{BB} - \operatorname{tr}\tilde{\Omega}_{FF}$ is the supertrace. We perform a Hubbard-

Stratonovich transformation to remove the quartic terms in the action by inserting the unity

$$1 = \frac{1}{4\pi^4}\int d\Omega \exp\left(-\frac{1}{2M}\,\mathrm{str}\,\tilde{\Omega}^2 - \bar{\Psi}\Omega\Psi - \frac{M}{2}\,\mathrm{str}\,\Omega^2\right), \tag{2.100}$$

$$\Omega = \begin{pmatrix} \Omega_{BB} & \Omega_{BF} \\ \Omega_{FB} & i\Omega_{FF} \end{pmatrix},\ \Omega_{\zeta\zeta'} = \begin{pmatrix} \Omega^{11}_{\zeta\zeta'} & \Omega^{12}_{\zeta\zeta'} \\ \Omega^{21}_{\zeta\zeta'} & \Omega^{22}_{\zeta\zeta'} \end{pmatrix},\ \Psi = \begin{pmatrix} \psi^1_B \\ \psi^2_B \\ \psi^1_F \\ \psi^2_F \end{pmatrix}, \tag{2.101}$$

where $d\Omega = d\Omega_{BB}d\Omega_{FF}d(\Omega_{BF},\Omega_{FB})$. The field $\Omega^{\alpha\alpha'}_{\zeta\zeta'}$ is bosonic when $\zeta = \zeta'$ and fermionic when $\zeta \neq \zeta'$. Inserting this unity and integrating out the $\Psi$ field, we find

$$Z_2(z,\tau;z',-\tau) = \frac{1}{4\pi^4}\int d\Omega[\mathrm{sdet}(\Omega+\mathcal{E})]^{-M}e^{-M\,\mathrm{str}\,\Omega^2/2}, \tag{2.102}$$

$$\mathcal{E} = \mathrm{diag}(z^{(1)}, z^{(2)}, z^{(1)}+\tau, z^{(2)}-\tau). \tag{2.103}$$

where $\mathrm{sdet}\,\Omega = \det(\Omega_{BB}-\Omega_{BF}\Omega^{-1}_{FF}\Omega_{FB})(\det\Omega_{FF})^{-1} = \det\Omega_{BB}[\det(\Omega_{FF}-\Omega_{FB}\Omega^{-1}_{BB}\Omega_{BF})]^{-1}$ is the superdeterminant.

We now make the change of variable $\Omega \to U\omega U^\dagger - \mathcal{E}$, where $U$ is an element of the superunitary group $U(2|2)$ and $\omega = \mathrm{diag}(\omega^1_B, \omega^2_B, i\omega^1_F, i\omega^2_F)$. The Jacobian of this change of variables is the supersymmetric generalization of the Vandermonde determinant [126] $\Delta_s(\Omega) = \det(1/(\omega_{B,i} - i\omega_{F,j}))$, where $\omega_{B,i}$ are the Bose-Bose eigenvalues and $i\omega_{F,j}$ are the Fermi-Fermi eigenvalues. We can now write

$$Z_2(z,\tau;z',-\tau) = \frac{1}{4\pi^4}\int d(\omega-\mathcal{E})\Delta^2_s(\omega)(\mathrm{sdet}\,\omega)^{-M}\int dU e^{-M(U\omega U^\dagger-\mathcal{E})^2/2}, \tag{2.104}$$

where $dU$ is the Haar measure of the superunitary group $U(2|2)$.

We can now use the supersymmetric generalization of the Itzykson-Zuber integral [126]:

$$\int dU e^{-M\,\mathrm{str}(U\omega U^\dagger-\beta)^2} \sim M^k[\Delta_s(\omega)\Delta_s(\eta)]^{-1}e^{-M\,\mathrm{str}(\omega-\eta)^2}, \tag{2.105}$$

where $dU$ is the Haar measure of the superunitary group $U(k|k)$, $\omega$ is diagonal, and $\eta$ is the diagonal matrix similar to $\beta$. Using this integral identity gives us

$$Z_2(z,\tau;z',-\tau) \sim \left(\frac{M}{2\pi}\right)^2 \int d(\omega-\mathcal{E})\frac{\Delta_s(\omega)}{\Delta_s(\mathcal{E})}(\mathrm{sdet}\,\omega)^{-M}e^{-M\,\mathrm{str}(\omega-\mathcal{E})^2/2} = \frac{\det R}{\Delta_s(\mathcal{E})}, \tag{2.106}$$

$$R^{\alpha\alpha'} = \frac{M}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{d\omega_B d\omega_F}{\omega_B + iq^{(\alpha)}0^+ - i\omega_F}\exp\left[-Mf\left(\omega_B + iq^{(\alpha)}0^+, z^{(\alpha)}\right) + Mf\left(i\omega_F, z^{(\alpha')} - (-1)^{\alpha'}\tau\right)\right], \tag{2.107}$$

$$f(\omega,c) = \frac{1}{2}(\omega-c)^2 + \ln\omega, \tag{2.108}$$

with $(q^{(1)},q^{(2)}) = (q,q')$.

The saddle points of the integrand in Eq. (2.107) are at

$$\omega_B = \omega^\alpha_{B\pm} = \frac{1}{2}\left(z^{(\alpha)} \pm i\sqrt{4-\left(z^{(\alpha)}\right)^2}\right), \tag{2.109}$$

$$i\omega_F = i\omega^{\alpha'}_{F\pm} = \frac{1}{2}\left(z^{(\alpha')} - (-1)^{\alpha'}\tau \pm i\sqrt{4-\left[z^{(\alpha')} - (-1)^{\alpha'}\tau\right]^2}\right). \tag{2.110}$$

Since $M$ is large, the eigenvalues of $A'$ lie within the interval $(-2,2)$; we only need to concern ourselves with the case when $|\Re(z)| < 2$. This means that the terms under the square roots in the equations above will always have a positive real part. Note that, due to the presence of

a singularity at $\omega_B = 0$, only one saddle point value is reachable through deformation of the contour of integration for $\omega_B$. This is $\omega^\alpha_{B+}$ when $q = 1$ and $\omega^\alpha_{B-}$ when $q = -1$. We need to consider both possible saddle point values for $i\omega_F$.

The result of the saddle point approximation is

$$R^{\alpha\alpha'} = \sum_\beta L^{\alpha\alpha'}_{q^{(\alpha)}\beta} \exp\left(M F^{\alpha\alpha'}_{q^{(\alpha)}\beta}\right), \tag{2.111}$$

$$F^{\alpha\alpha'}_{\beta\beta'} = -f\left(\omega^\alpha_{B\beta}, z^{(\alpha)}\right) + f\left(i\omega^{\alpha'}_{F\beta'}, z^{(\alpha')} - (-1)^{\alpha'}\tau\right), \tag{2.112}$$

$$L^{\alpha\alpha'}_{\beta\beta'} = \left(\omega^\alpha_{B\beta} - i\omega^{\alpha'}_{F\beta'}\right)^{-1} \left\{\left[1 - \left(\omega^\alpha_{B\beta}\right)^{-2}\right]\left[1 - \left(i\omega^{\alpha'}_{F\beta'}\right)^{-2}\right]\right\}^{-1/2}, \tag{2.113}$$

where $\beta, \beta' = \pm$. Expanding about infinite $N$ and $\tau = 0$, we find that the exponents of the exponential terms in Eq. (2.111) are, apart from the factor of $M$,

$$\begin{aligned} F^{\alpha\alpha'}_{\pm\pm} = \frac{(-1)^{\alpha'+1}}{2N}\left(x \mp 2\pi i p(x)\right)\Big\{(1-\delta^{\alpha\alpha'})[y + i(q-q')0^+] + N\tau\Big\} \\ + O(\tau^2) + O(N^{-1}\tau) + O(N^{-2}), \end{aligned} \tag{2.114}$$

$$\begin{aligned} F^{\alpha\alpha'}_{\pm\mp} = \pm\left[i\pi x p(x) + \ln\left(\frac{x - 2\pi i p(x)}{x + 2\pi i p(x)}\right)\right] \\ + \frac{(-1)^{\alpha'+1}}{2N}\left\{\Big[(1-\delta^{\alpha\alpha'})x \pm \delta^{\alpha\alpha'} 2\pi i p(x)\Big] y + \left(x \pm 2\pi i p(x)\right) N\tau\right\} \\ + O(\tau^2) + O(N^{-1}\tau) + O(N^{-2}). \end{aligned} \tag{2.115}$$

The prefactors are

$$L^{\alpha\alpha'}_{\pm\pm} = \frac{(-1)^{\alpha'} N}{(1-\delta^{\alpha\alpha'})[y + i(q-q')0^+] + N\tau} + O(\tau) + O(N^{-1}), \tag{2.116}$$

$$L^{\alpha\alpha'}_{\pm\mp} = \mp i(2\pi p(x))^{-2} + O(\tau) + O(N^{-1}). \tag{2.117}$$

Writing these quantities in this way is useful for times not larger than the Heisenberg time, where $|\tau| \ll 1$. We can restrict ourselves to these timescales since the SFF is already known to be 1 at later times.

Recall that we are interested in the $P^{\text{th}}$ power of the two-point generating function to use in Eq. 2.94. If $|\tau| \gg N^{-1}$, indicating times larger than the Thouless time, the $P^{\text{th}}$ power of the two-point generating function will be dominated by the saddle points that maximize the real part of $F^{\alpha\alpha'}_{\beta\beta'}$, excluding the unphysical combinations of saddle points that lead to the two-point generating function having a complex phase that grows with $N$. Considering this and using Eqs. (2.114)-(2.115), we find that the dominant Fermi-Fermi saddle point in the calculation of $R^{\alpha\alpha'}$ is $i\omega^{\alpha}_{F\beta^*(\alpha')}$, where

$$\beta^*(\alpha') = \delta_{qq'}q + (1-\delta_{qq'})(-1)^{\alpha'+1}. \tag{2.118}$$

Our result for the $P^{\text{th}}$ power of the two-point generating function at these timescales is then

$$[Z_2(z,\tau;z',-\tau)]^P = [\Delta_s(\mathcal{E})]^{-P}\left[\det\left(L^{\alpha\alpha'}_{q^{(\alpha)}\beta^*(\alpha')}\right)\right]^P \exp\left[N\left(F^{11}_{q\beta^*(1)} + F^{22}_{q'\beta^*(2)}\right)\right], \tag{2.119}$$

where we have used the fact that $F^{11}_{\beta_1\beta_2} + F^{22}_{\beta_3\beta_4} = F^{12}_{\beta_1\beta_4} + F^{21}_{\beta_3\beta_2}$ to extract the exponential term from the determinant. We can verify that the normalization is correct by setting $\tau = 0$ and finding that the above expression is equal to 1, as we know it must be from Eq. (2.95).

If $|\tau| \not\gg N^{-1}$, indicating times not larger than the Thouless time, the real parts of the exponent of $R^{\alpha\alpha'}$ will not be much larger than $P^{-1}$, so they will not contribute to the determination of the dominant saddle points for the $P^{\text{th}}$ power of the two-point generating function. The saddle points that will dominate are instead determined by the prefactors $L^{\alpha\alpha'}_{\beta\beta'}$. We see from

Eqs. (2.116)-(2.117) that, at these timescales, the prefactor for the $\beta = \beta'$ case is always greater than that for the $\beta = -\beta'$ case by at least a factor of $N$, so the non-dominant contributions can be neglected. This means that the dominant Fermi-Fermi saddle point in the calculation of $R^{\alpha\alpha'}$ is simply $i\omega_{Fq}^{\alpha'}$. With these considerations we find that the $P^{\text{th}}$ power of the two-point generating function at these timescales is

$$\begin{aligned}[Z_2(z,\tau;z',-\tau)]^P &= [\Delta_s(\mathcal{E})]^{-P}\left\{\det\left[L^{\alpha\alpha'}_{q^{(\alpha)}q^{(\alpha)}}\exp\left(MF^{\alpha\alpha'}_{q^{(\alpha)}q^{(\alpha)}}\right)\right]\right\}^P \\ &= e^{-i\pi N\tau(q-q')p(x)}\left[1-\left(\frac{N\tau}{y+i(q-q')0^+ +N\tau}\right)^2\right]^{-P} \\ &\quad\times\left[1-\left(\frac{N\tau}{y+i(q-q')0^+ +N\tau}\right)^2 e^{i\pi(q-q')p(x)(y+2N\tau)/P}\right]^P,\end{aligned} \tag{2.120}$$

where we have neglected the vanishing terms in the second equality. Once again we can verify that the normalization is correct by noting that the above expression becomes 1 when $\tau = 0$.

### 2.5.2 The spectral form factor

We are now in a position to calculate the SFF. We begin with the case where time is not larger than the Thouless time. Inserting Eq. (2.120) into the correlation function of Eq. (2.94), then extracting the SFF using Eq. (2.29), we obtain

$$\begin{aligned}K^{(\gamma)}(T) = 1+\frac{e^{-\sigma t^2}}{2\pi i(N\tau)^2 p(x)}\sum_q\Bigg\{e^{-2\pi iqN\tau p(x)}\int\frac{dy}{2\pi i}e^{iN^{-1}ty}\left[1-\left(\frac{N\tau}{y+N\tau+iq0^+}\right)^2\right]^{1-P} \\ \times\left[1-\left(\frac{N\tau}{y+N\tau+iq0^+}\right)^2 e^{2\pi iqp(x)(y+2N\tau)/P}\right]^P - \int\frac{dy}{2\pi i}e^{iN^{-1}ty}\left[1-\left(\frac{N\tau}{y+N\tau}\right)^2\right]\Bigg\}.\end{aligned} \tag{2.121}$$

The first term in the braces is the $q = -q'$ contributions while the second term is the $q = q'$ contributions. The integral over $y$ in the second term never has singularities in the half of the complex plane in which the contour of integration may be closed, so the term vanishes.

Making the changes of variable $q \to -q$ and then $y \to qy$ and expanding with the binomial theorem, we obtain

$$K^{(\gamma)}(T) = 1 + \frac{e^{-\sigma t^2}}{2\pi i (N\tau)^2 p(x)} \sum_q e^{2\pi i q N\tau p(x)} \int \frac{dy}{2\pi i} e^{iN^{-1}tqy} \left[\frac{(y+qN\tau)^2}{(y-i0^+)(y+2qN\tau)}\right]^{P-1} \times \sum_{j=0}^{P} \binom{P}{j} \left(\frac{N|\tau|}{y+qN\tau}\right)^{2j} e^{-2\pi i p(x)(y+2qN\tau)j/P}. \quad (2.122)$$

We observe that the integrand contains no singularities in the half of the complex plane in which the contour integral may be closed when $tq < 0$. So we can remove the sum over $q$ by making the substitution $tq \to |t|$. This gives us

$$K^{(\gamma)}(T) = 1 + \frac{e^{-\sigma t^2 - 2\pi N|\tau| p(x)}}{2\pi i (N\tau)^2 p(x)} \sum_{j=0}^{P} \binom{P}{j} \int \frac{dy}{2\pi i} D_j(y), \quad (2.123)$$

$$D_j(y) = (N|\tau|)^{2j} \frac{(y+iN|\tau|)^{2(P-1-j)}}{(y-i0^+)^{P-1}(y+2iN|\tau|)^{P-1}} \exp\left[iN^{-1}|t|y - 2\pi i p(x)(y+2iN|\tau|)j/P\right]. \quad (2.124)$$

By examining the $y$-dependent terms in the exponent of the exponential term of Eq. (2.124), we see that the contour of integration can be closed in the lower half of the complex plane when $|t|/p(x) < 2\pi M j$. In this case there may be contributions from residues located at $y = -iN|\tau|$ and $y = -2iN|\tau|$. When $|t|/p(x) > 2\pi M j$ the contour of integration may be closed in the upper half of the complex plane and there may be a contribution from the residue at $y = i0^+$. Fig. 2.7

shows the contours of integration for each situation and the poles which they enclose. We find that

$$K^{(\gamma)}(T) = 1 + \frac{e^{-\sigma t^2 - 2\pi N|\tau| p(x)}}{2\pi i (N\tau)^2 p(x)} \sum_{j=0}^{P} \binom{P}{j} \begin{cases} -\operatorname*{Res}_{y=-iN|\tau|} D_j(y) - \operatorname*{Res}_{y=-2iN|\tau|} D_j(y), & \frac{|t|}{p(x)} < 2\pi M j \\ \operatorname*{Res}_{y=i0^+} D_j(y), & \frac{|t|}{p(x)} > 2\pi M j \end{cases} . \tag{2.125}$$

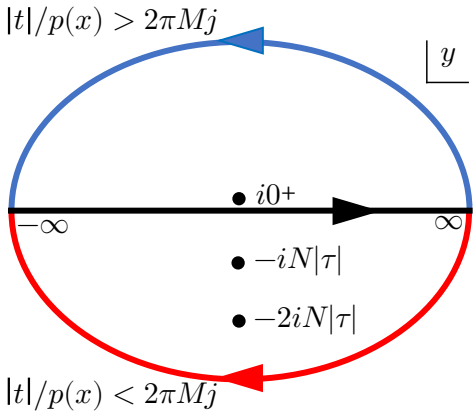



Figure 2.7: The poles and contours of integration for the integrals over $y$ in Eqs. (2.123) and (2.136). If $|t|/p(x)$ is less (greater) than $2\pi Mj$ the contour is closed in the lower (upper) half of the complex plane.

We first find the residue at $y = -iN|\tau|$. Observing that it is nonzero only when $j = P$, we find that

$$\operatorname*{Res}_{y=-iN|\tau|} D_P(y) = i(N|\tau|)^2 (N^{-1}|t| - 2\pi p(x)) e^{\sigma t^2 + 2\pi N|\tau| p(x)}. \tag{2.126}$$

The other two residues are

$$\operatorname*{Res}_{y=-2iN|\tau|} D_j(y) = 2i(-1)^{j+1} N|\tau| e^{2\sigma t^2} \sum_{k=0}^{P-2} J(j,k) \left(\sigma t^2 - 2\pi j M|\tau| p(x)\right)^k , \tag{2.127}$$

$$\operatorname*{Res}_{y=i0^+} D_j(y) = 2i(-1)^{j} N|\tau| e^{4\pi M j|\tau| p(x)} \sum_{k=0}^{P-2} J(j,k) \left(2\pi j M|\tau| p(x) - \sigma t^2\right)^k , \tag{2.128}$$

where

$$J(j,k) = \frac{2^{k-2(P-1)}}{k!} \sum_{k'=0}^{P-2-k} 2^{k'} \binom{2(P-1-j)}{k'} \binom{1-P}{P-2-k-k'}. \tag{2.129}$$

Using Eqs. (2.126)-(2.128) in Eq. (2.125) and unfolding by setting

$$u = |t|/p(x) = N|T|, \tag{2.130}$$

$$\Lambda = \lambda p(x), \tag{2.131}$$

we find that the SFF is

$$K^{(\gamma)}(T) = \min\left(\frac{T}{2\pi}, 1\right) - \frac{e^{-2\pi\Lambda^2 N^{1-\gamma}u}}{\pi\Lambda^2 N^{1-\gamma}u} \sum_{j=0}^{P} (-1)^j \binom{P}{j} Q_j^{(\gamma)}(u), \tag{2.132}$$

$$Q_j^{(\gamma)}(u) = \sum_{k=0}^{\infty} J(j,k) \left(-\Lambda^2 N^{-\gamma} u|u - 2\pi M j|\right)^k \begin{cases} e^{\Lambda^2 N^{-\gamma} u^2}, & u \leq 2\pi M j \\ e^{\Lambda^2 N^{-\gamma} u(4\pi M j - u)}, & u \geq 2\pi M j \end{cases}. \tag{2.133}$$

Note that we have raised the upper limit for the summation index $k$ from $P-2$ to $\infty$. We can do this because $J(j,k) = 0$ for $k > P-2$. Doing so removes the need for separate treatments of the cases where $P < 3$. Although we have derived this result for times not larger than the Thouless time, it turns out to be valid at all relevant timescales, as we will show.

We now examine the SFF at times larger than the Thouless time, where $|\tau| \gg N^{-1}$. We obtain the SFF by inserting Eq. (2.119) into the correlation function of Eq. (2.94) then extracting

the SFF using Eq. (2.29). This yields

$$K(T) = -\frac{e^{-\sigma t^2}}{2\pi i (N\tau)^2 p(x)} \sum_{qq'} qq' \int \frac{dy}{2\pi i} e^{iN^{-1}tq} \left[1 - \left(\frac{N\tau}{y + N\tau + i(q-q')0^+}\right)^2\right] \times [\Delta_s(\mathcal{E})]^{-P} \left[\det\left(L^{\alpha\alpha'}_{q^{(\alpha)}\beta^*(\alpha')}\right)\right]^P \exp\left[N\left(F^{11}_{q\beta^*(1)} + F^{22}_{q'\beta^*(2)}\right)\right]. \quad (2.134)$$

We note from Eq. (2.118) that, for the $q = q'$ terms, the dominant saddle points are the same as for the case examined above for earlier times. This means that, although there may be additional nonvanishing terms in the exponent of the exponential term of the integrand, they will be subleading and the residue analysis is unchanged. There are no singularities in the half of the complex plane in which the contour of integration may be closed. As they did for earlier times, the $q = q'$ terms of the SFF vanish.

We next consider the $q = -q' = 1$ term. From Eqs. (2.114) and (2.116)-(2.117) we find that $(L^{11}_{++}L^{22}_{--} - L^{12}_{+-}L^{21}_{-+})^P$ and $N(F^{11}_{++} + F^{22}_{--})$ have vanishing $y$-dependence. If we consider only positive times, which we can do since the SFF is symmetric in time, we find that the integrand once again has no singularities in the half of the complex plane in which the contour of integration may be closed. For this reason, the $q = -q' = 1$ contribution to the SFF also vanishes.

Finally we consider the only nonvanishing term, for which $q = -q' = -1$. From Eqs. (2.116)-(2.117) we find that

$$L^{11}_{-+}L^{22}_{+-} - L^{12}_{--}L^{21}_{++} = X + \left(\frac{N}{y - i0^+ + N\tau}\right)^2, \quad (2.135)$$

where $X$ is a quantity of order 1 with its largest $y$-dependent terms being of order $N^{-1}$, so they

can be neglected. Expanding with the binomial theorem, we find that the SFF is

$$K^{(\gamma)}(T) = 1 + \frac{e^{-\sigma t^2}}{2\pi i p(x)(N\tau)^2} \sum_{j=0}^{P} \binom{P}{j} \left(|\tau|^2 X\right)^{P-j} \int \frac{dy}{2\pi i} D'_j(y), \tag{2.136}$$

$$D'_j(y) = (N|\tau|)^{2j} \frac{(y + iN|\tau|)^{2(P-1-j)}}{(y - i0^+)^{P-1}(y + 2iN|\tau|)^{P-1}} \exp\left[iN^{-1}|t|y + N(F^{11}_{-+} + F^{22}_{+-})\right]. \tag{2.137}$$

From Eq. (2.115) we learn that that the largest $y$-dependent term in $N(F^{11}_{-+} + F^{22}_{+-})$ is $-2\pi i p(x) y$. This means that the contour of integration can be closed in the lower half of the complex plane when $|t|/p(x) < 2\pi N$. In this case there may be contributions from residues located at $y = -iN|\tau|$ and $y = -2iN|\tau|$. When $|t|/p(x) > 2\pi N$ the contour of integration is instead closed in the upper half of the complex plane and there may be a contribution from the residue at $y = i0^+$. So

$$K^{(\gamma)}(T) = 1 + \frac{e^{-\sigma t^2}}{2\pi i p(x)(N\tau)^2} \sum_{j=0}^{P} \binom{P}{j} \left(|\tau|^2 X\right)^{P-j} \times \begin{cases} -\operatorname*{Res}_{y=-iN|\tau|} D'_j(y) - \operatorname*{Res}_{y=-2iN|\tau|} D'_j(y), & |t|/p(x) < 2\pi N \\ \operatorname*{Res}_{y=i0^+} D'_j(y), & |t|/p(x) > 2\pi N \end{cases}. \tag{2.138}$$

It turns out that only the residue located at $y = -iN|\tau|$ contributes for large $N$. The exponent of the exponential factor of the terms coming from the residue at $y = -2iN|\tau|$ is

$$-\sigma t^2 + \left[iN^{-1}|t|(y - i0^+) + N(F^{11}_{-+} + F^{22}_{+-})\right]_{y=-2iN|\tau|} = \sigma t^2 - 2\pi N|\tau|p(x) + O(N\tau^2), \tag{2.139}$$

which becomes infinitely negative when $|t|/p(x) < 2\pi N$. Similarly, the exponent of the expo-

nential factor of the terms coming from the residue at $y = i0^+$ is

$$-\sigma t^2 + \left[iN^{-1}|t|(y - i0^+) + N(F^{11}_{-+} + F^{22}_{+-})\right]_{y=i0^+} = -\sigma t^2 + 2\pi N|\tau|p(x) + O(N\tau^2), \quad (2.140)$$

which becomes infinitely negative when $|t|/p(x) > 2\pi N$. The remaining residue at $y = -iN|\tau|$ is only nonzero when $j = P$. This residue is, to leading order,

$$\operatorname*{Res}_{y=-iN|\tau|} D'_P(y) = i(N|\tau|)^2(N^{-1}|t| - 2\pi p(x))e^{\sigma t^2}. \quad (2.141)$$

This leads to the same contribution to the SFF as the residue at the same location in the case examined above for earlier times. With this result we find that the SFF for times larger than the Thouless time is

$$K^{(\gamma)}(T) = \min\left(\frac{T}{2\pi}, 1\right). \quad (2.142)$$

Noting that $u \gg N^{\gamma-1}$ at these timescales, we see that this is consistent with the SFF we found for earlier times in Eq. (2.132). Therefore Eq. (2.132) holds for all relevant timescales.

In Fig. 2.8 we compare this result for the SFF to numerical results obtained through exact diagonalization for $\gamma = 1.6$, $P = 20$ blocks, and various values of the coupling parameter $\Lambda$. When numerically calculating the SFF with the full spectrum there is good agreement between theory and numerics at early times, but the agreement is less good at later times. We can trade this early-time agreement with later-time agreement by using only the eigenvalues within a window smaller than the spectral width. We can combine these results to get good agreement at all times of interest by determining the time at which the two approaches converge, then using the full spectrum result before that time and the windowed results at later times. We observe that the

SFF decays from the fully uncoupled result to the GUE SFF over time, with the rate of decay controlled by $\Lambda$. As we vary $\Lambda$ we can see the SFF moving between the behaviors we predicted schematically in Fig. 2.2. As in the RP model, for small values of $\Lambda$ we see that the SFF reaches its plateau at times greater (but not much greater) than the Heisenberg time. The agreement between theory and numerics is still very good at these late times.

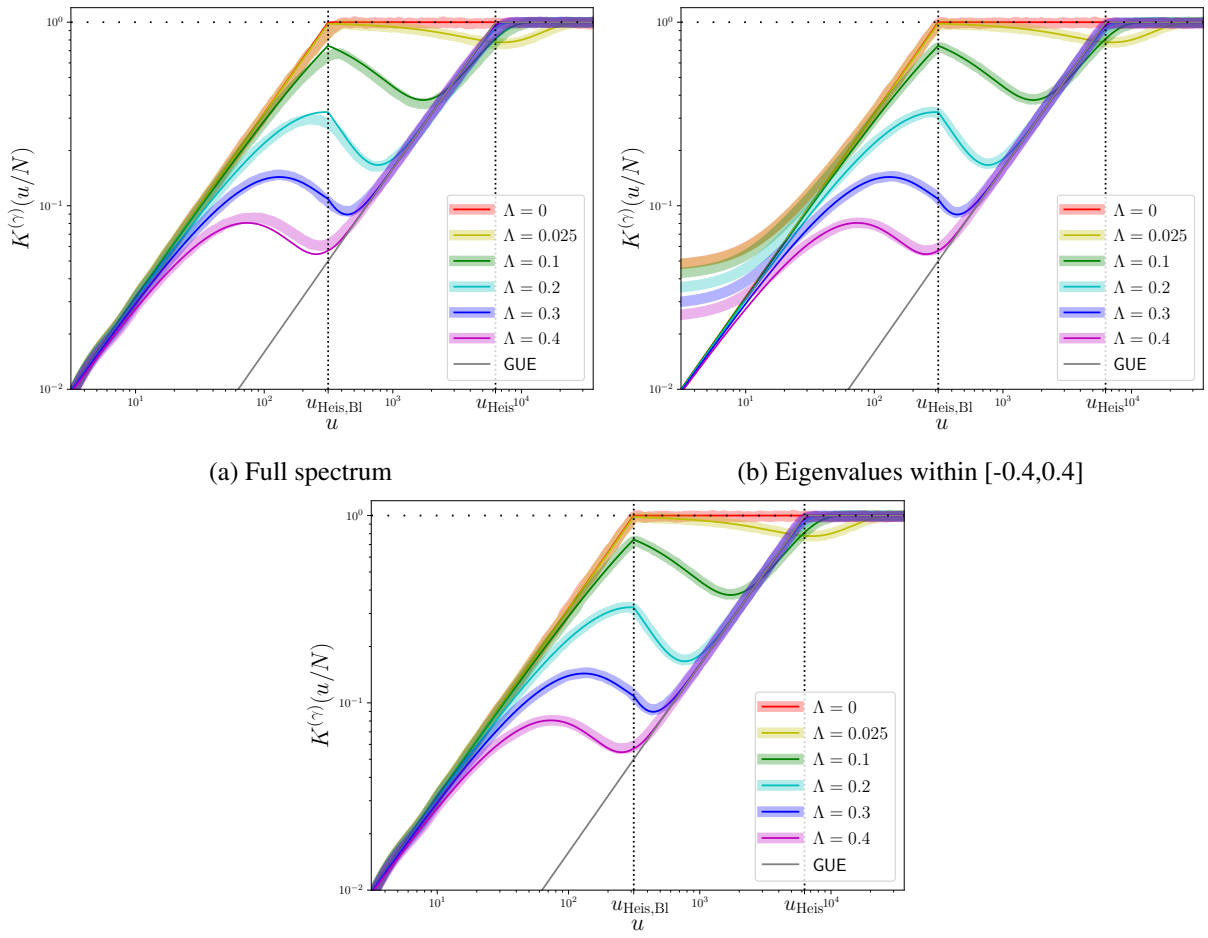
Somewhat similar SFFs with a short-time peak then a crossover to RMT behavior have been found for random quantum circuits [127–129] and the mass-deformed Sachdev-Ye-Kitaev model [130]. However, the enhanced ramps are faster than linear in time in these cases because these models transition from many-body-localized to ergodic behavior, in contrast to the glassy to ergodic transition we see in the BRP model.

### 2.5.3 The infinite matrix size limit

We now seek to find the behavior of the SFF in the large $N$ limit for all timescales of interest. Above we have already considered the SFF at times greater than the Thouless time and argued that several terms vanish at large $N$, so Eq. (2.142) is the large $N$ limit for the SFF at those timescales. Going forward we will consider earlier times. We will assume for now that we are not at the Thouless and Heisenberg timescales simultaneously, which occurs when $\gamma = 2$ and $u \sim t_{\text{Th}} \sim N$. This is a special case which we will examine later.

Under these conditions, the sum of the $k = 0$ contributions to the SFF in the time period $2\pi M j \leq u \leq 2\pi M(j+1)$ is

$$\frac{e^{-2\pi\Lambda^2 N^{1-\gamma}u}}{\pi\Lambda^2 N^{1-\gamma}u}\left[\sum_{l=0}^{j}(-1)^l\binom{P}{l}J(l,0)e^{\Lambda^2 N^{-\gamma}u(4\pi Ml-u)}+e^{\Lambda^2 N^{-\gamma}u^2}\sum_{l=j+1}^{P}(-1)^l\binom{P}{l}J(l,0)\right]. \tag{2.143}$$




(a) Full spectrum

(b) Eigenvalues within [-0.4,0.4]

(c) Hybrid of (a) and (b)

Figure 2.8: The SFF of the BRP model when $\gamma = 1.6$ and there are $P = 20$ blocks for several values of the coupling parameter $\Lambda$. The dark thin lines show the analytical results while the lighter, thicker lines show numerical results. The left (right) dotted line indicates the block (full) Heisenberg time. The numerics are obtained through exact diagonalization of size $N = 1000$ random matrices and are averaged over $10^5$ realizations. In (a) the numerics are calculated with the full spectrum while in (b) only the eigenvalues within the window [-0.4.0.4] are used. In (c) the results of (a) are used at times before the numerical curves for a particular $\Lambda$ converge and the numerical results of (b) are used after.

Since $\gamma > 1$, the exponents of the exponential terms within the brackets are small. Expanding them to first order, we can write this as

$$\frac{e^{-2\pi\Lambda^2 N^{1-\gamma}u}}{\pi\Lambda^2 N^{1-\gamma}u}\sum_{l=0}^{P}(-1)^l\binom{P}{l}J(l,0) + \frac{e^{-2\pi\Lambda^2 N^{1-\gamma}u}}{\pi}\left[\sum_{l=0}^{j}(-1)^l\binom{P}{l}J(l,0)(T-4\pi l/P) - \sum_{l=j+1}^{P}(-1)^l\binom{P}{l}J(l,0)T\right]. \quad (2.144)$$

We note that $J(j,k)$ is a polynomial of order less than $P$ in $j$. So, from the theory of finite differences, we find that [131]

$$\sum_{j=0}^{P}(-1)^j\binom{P}{j}J(j,k) = 0. \quad (2.145)$$

Using this fact, we find that the sum of the $k < 2$ contributions is, to first order in $N^{-1}$,

$$\frac{2}{\pi}e^{-2\pi\Lambda^2 N^{1-\gamma}u}\sum_{l=0}^{j}(-1)^l\binom{P}{l}[J(l,0)+J(l,1)](T-2\pi l/P). \quad (2.146)$$

The sum of the $k \geq 2$ contributions will be subleading. So, to first order, the SFF in this time range is

$$K^{(\gamma)}(T) = \min\left(\frac{T}{2\pi},1\right)+\frac{2}{\pi}e^{-2\pi\Lambda^2 N^{1-\gamma}u}\sum_{l=0}^{j}(-1)^l\binom{P}{l}[J(l,0)+J(l,1)](T-2\pi l/P). \quad (2.147)$$

We now seek to simplify $J(l,0)+J(l,1)$. By shifting the summation index in $J(l,1)$ and using Pascal's rule

$$\binom{n-1}{k}+\binom{n-1}{k-1} = \binom{n}{k}, \quad (2.148)$$

we find that

$$J(l,0)+J(l,1)=2^{-2(P-2)}\sum_{k'=0}^{P-2}2^{k'}\binom{2(P-1-l)+1}{k'}\binom{1-P}{P-2-k'}=J(l-1/2,0). \tag{2.149}$$

We can also find a closed-form expression for $J(l,0)$ by noting that it is the $(P-2)^{\text{th}}$ coefficient in the Maclaurin series expansion

$$J(l,0)=2^{-2(P-1)}[z^{P-2}](1+z)^{1-P}\sum_{k=0}^{\infty}(2z)^k\binom{2(P-1-l)}{k}, \tag{2.150}$$

where $[z^{P-2}]$ is the $(P-2)^{\text{th}}$ coefficient extractor. Now, summing over $k$, writing the coefficient as a residue at $z=0$, and making the change of variable $z\to(\sqrt{4z+1}-1)/2$, we find that

$$J(l,0)=2^{-2(P-1)}\operatorname*{Res}_{z=0}z^{1-P}(1+4z)^{P-l-3/2}=\frac{1}{4}\binom{P-l-3/2}{P-2}. \tag{2.151}$$

Using Eqs. (2.149) and (2.151) in Eq. (2.147), we find, for $2\pi j/P\le T\le 2\pi(j+1)/P$,

$$K^{(\gamma)}(T)=\min\left(\frac{T}{2\pi},1\right)+e^{-2\pi\Lambda^2N^{1-\gamma}u}\begin{cases}(P-1)\frac{T}{2\pi}, & j=0\\ 1-\frac{T}{2\pi}, & 0<j<P\\ 0, & j=P\end{cases}. \tag{2.152}$$

This means that, overall, the SFF is

$$K^{(\gamma)}(T)=\left(1-e^{-2\pi\Lambda^2N^{1-\gamma}u}\right)\min\left(\frac{T}{2\pi},1\right)+e^{-2\pi\Lambda^2N^{1-\gamma}u}\min\left(P\frac{T}{2\pi},1\right). \tag{2.153}$$

This result becomes 1 for times greater than the Heisenberg time, so it is also valid at those times. We conclude that Eq. (2.153) is valid at all relevant times, excepting the case of coincident Thouless and Heisenberg timescales. For times smaller than the Thouless time, where $u \ll N^{\gamma-1}$, we see that

$$K^{(\gamma)}(T) = \min\left(P\frac{T}{2\pi}, 1\right). \tag{2.154}$$

This result is exactly what we would expect in the weak-coupling limit: the GUE SFF with time enhanced by a factor of $P$, the number of blocks.

For the Thouless timescale, while different from the Heisenberg timescale (meaning $\gamma \neq 2$), we obtain

$$K^{(\gamma)}_{\text{Th}}(v) = K^{(\gamma)}(N^{\gamma-2}v) = \left(1 - e^{-2\pi\Lambda^2 v}\right)\min\left(\frac{T}{2\pi}, 1\right) + e^{-2\pi\Lambda^2 v}\min\left(P\frac{T}{2\pi}, 1\right), \tag{2.155}$$

where $v = N^{2-\gamma}T$. We see that at the Thouless timescale there is a crossover from the GUE SFF for large $\Lambda$ to the uncoupled blocks result for small $\Lambda$. For times greater than the block Heisenberg time $2\pi M$, this becomes identical to the RP SFF. The reason for this is that the correlations between the eigenvalues of $A$ do not persist for times larger than the block Heisenberg time. At these timescales we are justified in setting $Z_2(z, \tau; z', -\tau) = [g_2(z, \tau; z', -\tau)]^M$ in Eq. (2.94), which reduces the calculation of the SFF to that of the RP model.

We now examine the $\gamma = 2$ case at the Thouless timescale. From Eq. (2.132) we find that

$$K^{(2)}_{\text{Th}}(v) = K^{(2)}(v) = \min\left(\frac{v}{2\pi}, 1\right) - \frac{e^{-2\pi\Lambda^2 v}}{\pi\Lambda^2 v}\sum_{j=0}^{P}(-1)^j\binom{P}{j}Q^{(2)}_{\text{Th},j}(v), \tag{2.156}$$

$$Q^{(2)}_{\text{Th},j}(v) = \sum_{k=0}^{P-2} J(j,k)(-\Lambda^2 v|v - 2\pi j/P|)^k \begin{cases} e^{\Lambda^2 v^2}, & v \le 2\pi j/P \\ e^{\Lambda^2 v(4\pi j/P - v)}, & v \ge 2\pi j/P \end{cases}. \tag{2.157}$$

Due to the coincidence of the Thouless and Heisenberg timescales, the SFF in this case does not generally have a closed form expression. Although the SFF is unwieldy in general, it does simplify greatly for $P < 3$. When $P = 1$ we find the GUE result

$$K^{(2)}_{\text{Th}}(v) = \min\left(\frac{v}{2\pi}, 1\right), \tag{2.158}$$

while for $P = 2$, the simplest nontrivial case, we find

$$K^{(2)}_{\text{Th}}(v) = \min\left(\frac{v}{2\pi}, 1\right) + \frac{e^{-2\pi\Lambda^2 v}}{2\pi\Lambda^2 v}\begin{cases} \sinh(\Lambda^2 v^2), & 0 \le v \le \pi \\ e^{\Lambda^2 v(2\pi - v)} - \cosh(\Lambda^2 v^2), & \pi \le v \le 2\pi \\ -2\sinh^2(\pi\Lambda^2 v)e^{\Lambda^2 v(2\pi - v)}, & v \ge 2\pi \end{cases}. \tag{2.159}$$

The Thouless-Heisenberg timescale SFF is plotted in Fig. 2.9, for the cases of $P = 4$ and $P = 20$. Also shown are numerical results obtained through exact diagonalization which are in strong agreement with the theory. For times much greater than the block Heisenberg time there is graphical evidence that the SFF moves to that of the RP model. By examining Fig. 2.9, we can see that when the number of sectors $P$ becomes large, making $u \gg 2\pi M$, the SFF moves to the

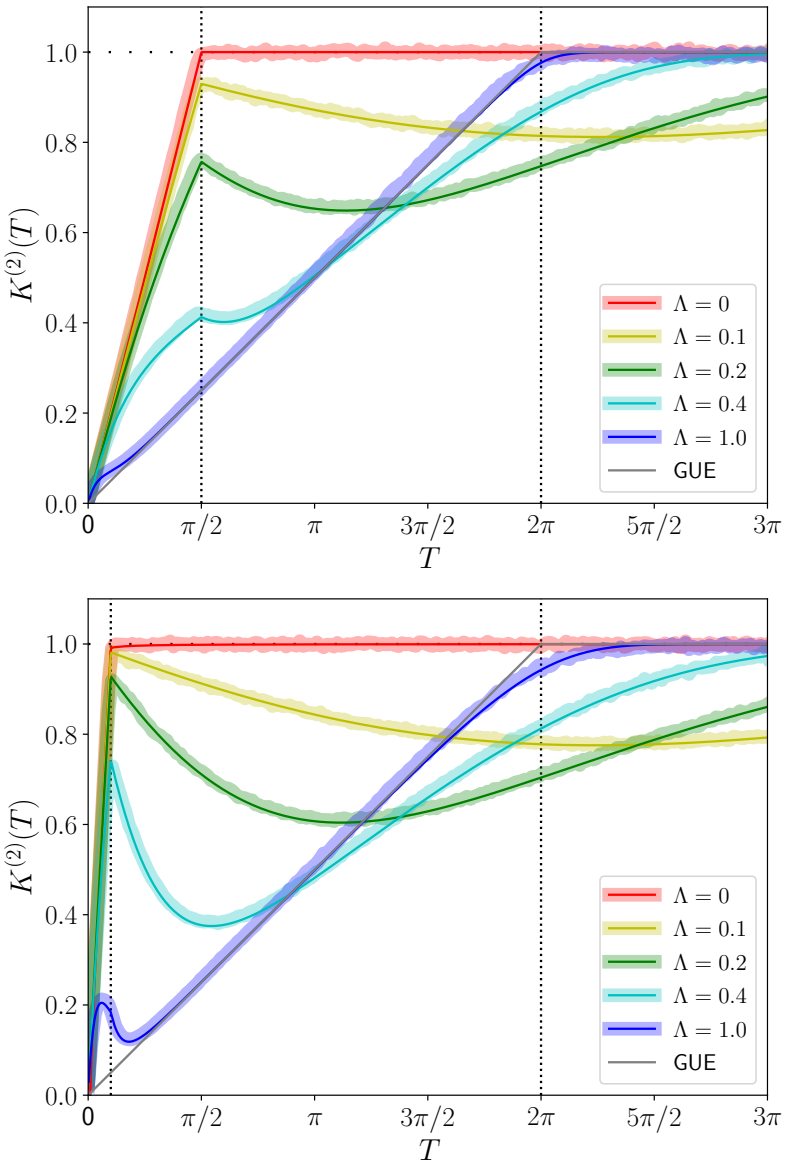
RP SFF shown in Fig. 2.5. As with the RP SFF, we see that the SFF reaches its plateau value at times greater than the Heisenberg time when $\Lambda$ is finite. In this case there is also a trade-off between level repulsion at early and late times. The area between the SFF and its plateau value is independent of the coupling strength for any nonzero coupling between the sectors, as is shown in the Appendix.

Bringing together all the results at the Thouless timescale we have that the SFF is

$$K_{\text{Th}}^{(\gamma)}(v) = \begin{cases} \min\left(P\frac{T}{2\pi}, 1\right), & \gamma > 2 \\ \min\left(\frac{v}{2\pi}, 1\right) - \frac{e^{-2\pi\Lambda^2 v}}{\pi\Lambda^2 v}\sum_{j=0}^{P}(-1)^j\binom{P}{j}Q_{\text{Th},j}^{(2)}(v), & \gamma = 2 \\ \left(1 - e^{-2\pi\Lambda^2 v}\right)\min\left(\frac{T}{2\pi}, 1\right) + e^{-2\pi\Lambda^2 v}\min\left(P\frac{T}{2\pi}, 1\right), & 1 < \gamma < 2 \end{cases} . \tag{2.160}$$

The behavior of the SFF is dependent of $\Lambda$ when $1 < \gamma \leq 2$. For small $\Lambda$ the SFF goes to the result expected for uncoupled sectors, while for large $\Lambda$ it goes to the GUE result. This can be seen explicitly in the above equation for $1 < \gamma < 2$ and can be observed graphically for the $\gamma = 2$ case in Fig. 2.9.

As we did for the RP model, we can create a diagram showing the behavior of the SFF at different timescales and values of $\gamma$, shown in Fig. 2.10. At times larger than the Thouless time the SFF goes to the GUE result. At times smaller than the Thouless time the SFF goes to the GUE SFF with time enhanced by a factor of $P$, the number of blocks. If $\gamma > 1 + d$ this result first reaches its plateau value at the block Heisenberg time, which is smaller than the full Heisenberg time by a factor of $P$. This means that at times larger than the block Heisenberg time the SFF is the same as the RP SFF. At the Thouless timescale, while it is not larger than the Heisenberg timescale, there is a crossover between the enhanced GUE SFF and the standard GUE SFF. The




(a) $P = 4$

(b) $P = 20$

Figure 2.9: The SFF for the BRP model when the Thouless and Heisenberg timescales coincide for various values of the coupling parameter $\Lambda$. The dark thin lines show the analytical results while the lighter, thicker lines show the numerical results. The numerics are obtained through exact diagonalization of size $N = 1000$ random matrices and are averaged over $10^5$ realizations. The numerics are restricted to eigenvalues within the window [-0.75,0.75]. In (a) there are $P = 4$ blocks while in (b) there are $P = 20$. The left (right) dotted line indicates the block (full) Heisenberg time.

SFF is 1 at times larger than the Heisenberg time.

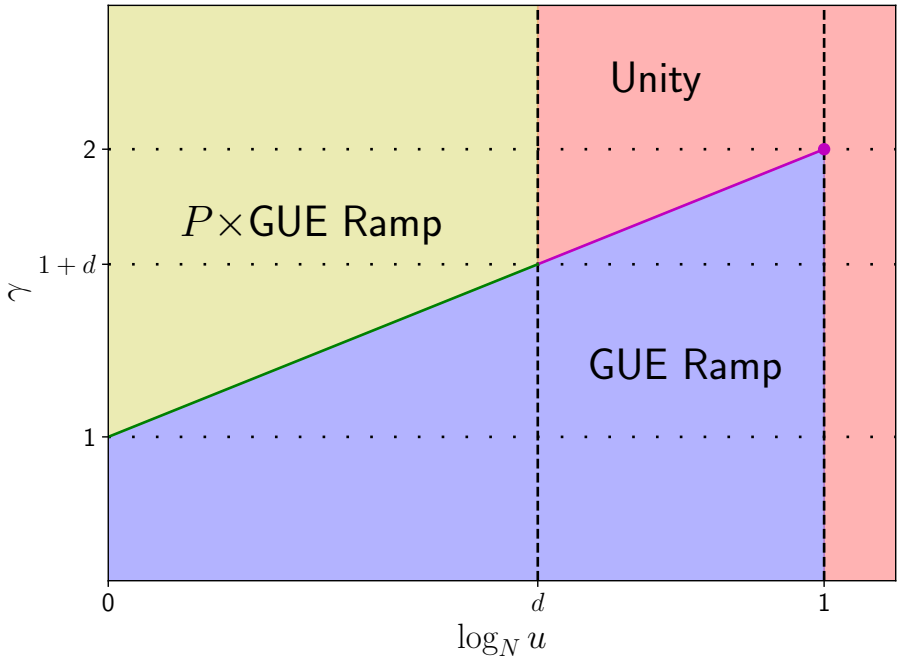



Figure 2.10: The SFF of the BRP model at different timescales and values of the coupling parameter $\gamma$ when the size of the blocks $M$ is much less than $N$. The green line represents a crossover between the enhanced GUE ramp and the regular GUE ramp. The magenta line represents the crossover between the Poissonian result and the regular GUE ramp. The magenta dot indicates the point at which the Thouless and Heisenberg timescales coincide and there is no closed form expression for the SFF. The dotted lines indicate the values of $\gamma$ at which there is a phase transition. The left (right) dashed line indicates the block (full) Heisenberg time.

## 2.6 Conclusion

In this paper we introduce a generalization of the Rosenzweig-Porter (RP) model called the Block Rosenzweig-Porter (BRP) model. This is done by redefining the diagonal matrix in the RP model to be a block diagonal matrix with each block being a GUE matrix. In doing so we obtain a minimal quantum glass model which immediately thermalizes within the blocks but has much slower global thermalization (if it thermalizes at all) depending on the strength of the inter-block coupling. It is known that the RP model is chaotic for $\gamma < 1$ and localized for $\gamma > 2$. For

intermediate values of $\gamma$ the RP model is localized at early times, then thermalizes and becomes chaotic at the Thouless time. We find that the same is true for the BRP model. However, the localized behaviors for the two models are different. Whereas the RP model exhibits localization in single states, the BRP model instead exhibits localization within single blocks.

Our main result is a calculation of the spectral form factor (SFF) at all timescales larger than the inverse spectral width for $\gamma > 1$ in the BRP model. As a lead-in to this, we perform this calculation for the RP model. In the intermediate phase where $1 < \gamma < 2$, the SFF of the RP model indicates Poissonian statistics before the Thouless timescale and GUE statistics afterward. At the Thouless timescale there is a crossover between these behaviors mediated by the unfolded coupling parameter $\Lambda$. These results are consistent with those of Ref. [89]. Within the same range for $\gamma$, the BRP model has statistics consistent with independent GUE blocks before the Thouless timescale. After the Thouless timescale statistics consistent with a single large GUE block are found. As with the RP model, there is a crossover between these behaviors at the Thouless timescale, again mediated by $\Lambda$. This indicates that the system is initially frozen into a single sector, then escapes and thermalizes at the Thouless time.

An important feature of the SFF is the times at which it is equal to its plateau value, as these indicate the disappearance of repulsion between eigenvalues at the energy separations corresponding to these times. We find that if $\gamma < 1 + d$, where $d = \log_N M$, the SFF will first reach its plateau value at or near the Heisenberg time, meaning that eigenvalue correlations first vanish at separations smaller than the mean level spacing, as they must for all discrete quantum systems. However, if $\gamma > 1 + d$ the SFF first reaches its plateau value at the block Heisenberg time, which is smaller than the full Heisenberg time by a factor of $P$, the number of blocks. This means that correlations between eigenvalues first vanish at separations smaller than the mean level spacing

of a single block. If $\gamma < 2$, meaning the system is not in its localized phase, level repulsion may still be present at later times (smaller energy separations), causing the SFF to drop back below its plateau value until the Heisenberg time.

While this work has been concerned with dynamics and associated spectral statistics, one can also ask about the eigenstate properties in the three phases. In the RP model the eigenstates are fully ergodic for $\gamma < 1$ and localized to a single state for $\gamma > 2$. For intermediate values of $\gamma$ the eigenstates are neither localized nor ergodic; they are termed nonergodic extended states [89]. A similar phenomenon occurs in the BRP model. For $\gamma > 2$ the eigenstates are localized to a single sector. However, the eigenstates become fully delocalized across the Hilbert space for $\gamma < 1 + d$ (rather than merely $\gamma < 1$), where $M$ is the sector size. The fact that this eigenstate transition is, in general, separate from the transition to full GUE statistics at $\gamma = 1$ is an interesting feature of the BRP model. Future work may examine this feature and the eigenstate statistics in more detail.

The BRP model we introduce here is significant because it is a solvable random matrix model with a glass transition. The SFF of the BRP model can be calculated at all relevant timescales. In more complex quantum glass models one cannot probe times exponential in the system size with currently available methods. Particularly interesting is the case in which the Thouless time is of the same order as the Heisenberg time. The BRP model is thus a useful starting point in understanding the spectral statistics of systems with slow dynamics.

Due to the simplicity of the BRP model it does not capture some features of more physical models. Each block is taken to be the same size and independent of the others. More generally we can expect the sectors to have some distribution in size with correlations between their matrix elements. The BRP Hamiltonian also has only two levels in its hierarchy, making it a model

with a single nearly conserved quantity or slow mode. More realistic systems may have a richer hierarchy of long timescales, which could be captured by a generalized BRP model with more than two levels of nested blocks. An avenue for future work may be to modify the BRP model to capture these features.

## Acknowledgements

We would like to thank I. M. Khaymovich for useful discussion and comments which improved this manuscript, in particular for pointing out the interpretation of the coupling parameter $\gamma$ in terms of eigenstate properties. This work was supported by the following: The National Science Foundation through the Quantum Leap Challenge Institute for Robust Quantum Simulation (grant OMA-2120757), NSF DMR-2037158, US-ARO Contract No.W911NF1310172, and Simons Foundation (V.G.); the Joint Quantum Institute (M.W.); the Air Force Office of Scientific Research under award numbers FA9550-17-1-0180 (M.W.) and FA9550-19-1-0360 (B.S.); the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Accelerated Research for Quantum Computing program "FAR-QC" (R.B.); the DoE ASCR Quantum Testbed Pathfinder program under award number DE-SC0019040 (C.L.B.); the DoE ASCR Accelerated Research in Quantum Computing program under award number DE-SC0020312 (C.L.B.); the DoE QSA, AFOSR, AFOSR MURI, NSF PFCQC program, NSF QLCI under award number OMA-2120757 (C.L.B.); DoE award number DE-SC0019449 (C.L.B.), ARO MURI, and DARPA SAVaNT ADVENT (C.L.B.). This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE 1840340 (R.B.), and by the National Science Foundation NRC postdoctoral fellowship

program (C.L.B.).

# Chapter 3: Spectral Form Factor of a Quantum Spin Glass

**Authors:** Michael Winer, Richard Barney, Christopher L. Baldwin, Victor Galitski, Brian Swingle

**Abstract:** It is widely expected that systems which fully thermalize are chaotic in the sense of exhibiting random-matrix statistics of their energy level spacings, whereas integrable systems exhibit Poissonian statistics. In this paper, we investigate a third class: spin glasses. These systems are partially chaotic but do not achieve full thermalization due to large free energy barriers. We examine the level spacing statistics of a canonical infinite-range quantum spin glass, the quantum $p$-spherical model, using an analytic path integral approach. We find statistics consistent with a direct sum of independent random matrices, and show that the number of such matrices is equal to the number of distinct metastable configurations—the exponential of the spin glass "complexity" as obtained from the quantum Thouless-Anderson-Palmer equations. We also consider the statistical properties of the complexity itself and identify a set of contributions to the path integral which suggest a Poissonian distribution for the number of metastable configurations. Our results show that level spacing statistics can probe the ergodicity-breaking in quantum spin glasses and provide a way to generalize the notion of spin glass complexity beyond models with a semi-classical limit.

## 3.1 Introduction

An isolated quantum many-body system which reaches an effective thermal equilibrium state starting from an out-of-equilibrium initial state is often called "quantum chaotic." As commonly used, quantum chaos is a loose term referring to a family of phenomena that typically co-occur, including the ability of the system to serve as its own heat bath [29, 30, 132], hydrodynamic behavior of conserved quantities [133–137], and random-matrix-like energy eigenvalues [32, 36, 84, 138]. Given this variety, it is crucial to understand the relationships between different manifestations of quantum chaos [139, 140].

These relationships are complicated and interesting in large part because the systems in question have structure, such as locality and symmetry. For example, if the Hamiltonian has spatial locality, energy conservation implies the existence of slow hydrodynamic modes and an associated long time scale, the Thouless time, such that random-matrix behavior is only present for energy levels closer than the inverse Thouless time [127, 141]. Similarly, if the Hamiltonian possesses a symmetry, then it can be organized into blocks labelled by irreducible representations of the symmetry. One finds random-matrix statistics within each individual block, but full ergodicity is broken because matrix elements between different blocks are forbidden [2, 116, 142–144].

It is natural to ask whether there are other ways in which ergodicity can be lost, and if so, what the resulting spectral statistics of the Hamiltonians are. In particular, we will better understand the relations between different measures of quantum chaos by understanding how they are lost and what replaces them.

Quantum spin glasses provide one well-established context to explore these questions,

since they exhibit a rich phenomenology associated with the inability to fully thermalize [15–18, 145–147] In this paper, we determine the spectral statistics of an analytically tractable spin glass model, the quantum $p$-spherical model. We find that up to times polynomial in the system size, the Hamiltonian can effectively be described as approximately block-diagonal. Each block behaves as a random matrix independent of the others, and the number of blocks depends on the energy per particle. At high energies, there is only one block and the system is ergodic. Below a critical energy density, the Hamiltonian breaks into exponentially many blocks — the average number of blocks jumps discontinuously from the high energy regime and then decreases as the energy density decreases further. We establish these results via a path integral computation of the spectral form factor (SFF), which measures correlations between pairs of energy levels [75, 118, 148–150].

In the remainder of the introduction, we give some physical context by reviewing the spectral form factor and mean-field spin glasses, and then summarize our results. In Sec. 3.2, we review the $p$-spherical model in detail. In Sec. 3.3, we calculate the SFF of this model in the high-temperature ergodic regime, and in Sec. 3.4, we do so in the non-ergodic regime. Finally, in Sec. 3.5, we investigate higher-moment analogues of the SFF. We then discuss implications of these results and directions for future work in Sec. 5.4.

### 3.1.1 Review of the spectral form factor

To study the spectral correlations of a Hamiltonian $H$, a standard tool is the spectral form factor (SFF) [74, 151], defined as

$$\text{SFF}(T) \equiv \left|\text{Tr} e^{-iHT}\right|^2. \tag{3.1}$$

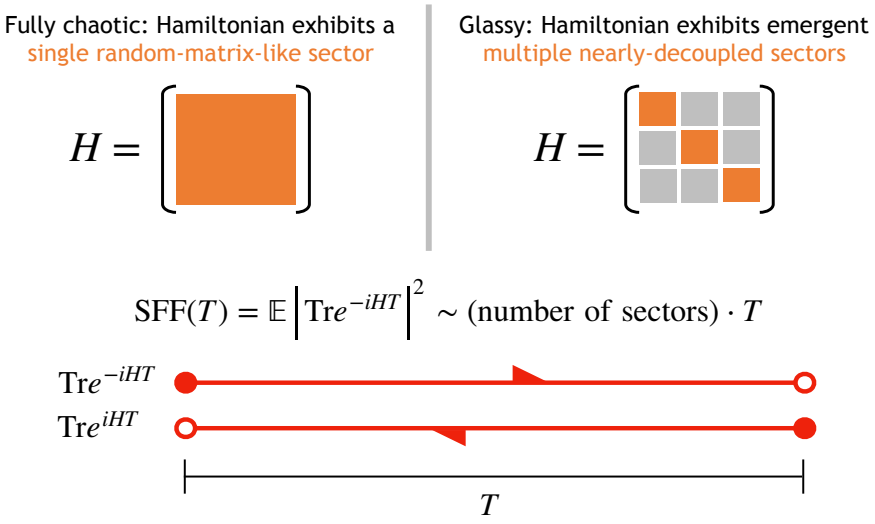



Figure 3.1: (Top left) Fully chaotic systems have energy levels that are statistically similar to a Gaussian random matrix, indicated by the orange block. (Top right) By contrast, quantum spin glasses in the non-ergodic phase have spectral statistics that resemble a collection of many nearly-decoupled random matrices (Bottom) Spectral statistics can be diagnosed via the spectral form factor, denoted SFF$(T)$, which consists of a path integral over a pair of real-time contours as indicated by the red lines. The universal part of SFF$(T)$, which is proportional to $T$, is enhanced by the number of effectively uncoupled sectors (other non-universal contributions are not indicated here).

In situations where the spectrum is unbounded, or when one wishes to concentrate on a portion of the spectrum, the trace in Eq. (3.1) is regulated by a filter function $f(H)$:

$$\text{SFF}(T, f) \equiv \left|\text{Tr} f(H) e^{-iHT}\right|^2. \tag{3.2}$$

One common choice is $f(H) = e^{-\beta H}$[75, 152], and another is $f(H) = e^{-c(H-E_0)^2}$. The latter allows one to study level statistics near a specified energy $E_0$.

For a single Hamiltonian, the SFF is an erratic function of time [151]. Thus one usually considers an ensemble of Hamiltonians and defines the SFF as the average of Eq. (3.2) over the ensemble. Throughout this paper, we use the notation $\mathbb{E}[\,\cdot\,]$ to denote the ensemble average.

The SFF is closely related to the correlation function of the density of states. Formally, the

(filtered) density of states is given by

$$\rho(E, f) \equiv \sum_n f(E_n)\delta(E - E_n) = \mathrm{Tr} f(H)\delta(E - H), \tag{3.3}$$

where $n$ labels the eigenstate of $H$ with eigenvalue $E_n$, and its correlation function is

$$C(E, \omega, f) \equiv \mathbb{E}\left[\rho\left(E + \frac{\omega}{2}, f\right)\rho\left(E - \frac{\omega}{2}, f\right)\right]. \tag{3.4}$$

We have that

$$\begin{aligned}
\mathrm{SFF}(T, f) &= \mathbb{E}\Big[\mathrm{Tr} f(H)e^{-iHT}\mathrm{Tr} f(H)e^{iHT}\Big] \\
&= \int dE d\omega\, e^{-i\omega T}\mathbb{E}\left[\mathrm{Tr} f(H)\delta\left(E + \frac{\omega}{2} - H\right)\mathrm{Tr} f(H)\delta\left(E - \frac{\omega}{2} - H\right)\right] \\
&= \int d\omega\, e^{-i\omega T}\int dE\, C(E, \omega, f).
\end{aligned} \tag{3.5}$$

The SFF is simply the Fourier transform of the correlation function with respect to $\omega$, integrated over $E$ (although the filter function allows one to concentrate on an arbitrary subset of the spectrum).

It is conceptually useful to split the SFF into two contributions:

$$\mathrm{SFF}(T, f) = \left|\mathbb{E}\mathrm{Tr} f(H)e^{-iHT}\right|^2 + \left(\mathbb{E}\left[\left|\mathrm{Tr} f(H)e^{-iHT}\right|^2\right] - \left|\mathbb{E}\mathrm{Tr} f(H)e^{-iHT}\right|^2\right). \tag{3.6}$$

The first term, the disconnected piece of the SFF, comes solely from the average density of states. It is the second term, the connected piece, that contains information on the correlation between energy levels. The assertion of "random matrix universality" [32, 82] can be phrased as

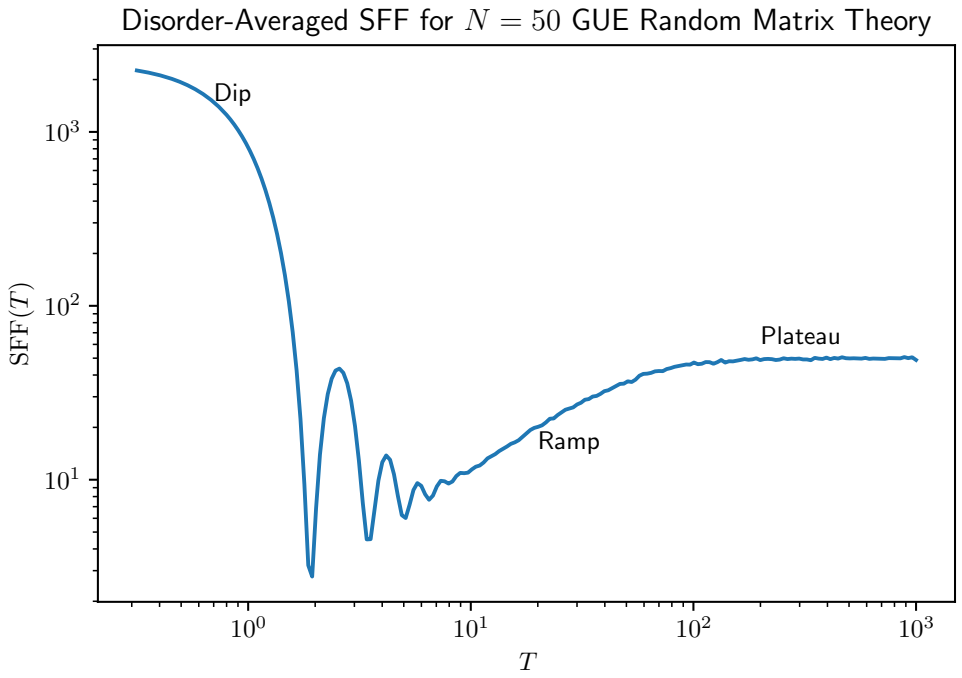



Figure 3.2: The disorder-averaged SFF for the Gaussian unitary ensemble (GUE) of matrix dimension $N = 50$, computed numerically by averaging over ten thousand realizations. The three distinct regimes — dip, ramp, plateau — are indicated.

the statement that an ensemble of quantum chaotic Hamiltonians will generically have the same *connected* SFF as the canonical Gaussian ensembles of random matrix theory [84, 153]. This conjectured universal behavior is illustrated in Fig. 3.2, which plots the disorder-averaged SFF of the Gaussian unitary ensemble (one of the aforementioned canonical ensembles). Note the three distinct regimes:

- The "dip", occurring at short times, comes from the disconnected piece of the SFF (and thus its precise shape is non-universal). It reflects a loss of constructive interference — the different terms of $\mathrm{Tr} e^{-iHT}$ acquire different phase factors as $T$ increases.

- The "ramp", occurring at intermediate times, is arguably the most interesting regime. In

the canonical matrix ensembles, it is a consequence of the result[84]

$$\mathbb{E}\left[\rho\left(E+\frac{\omega}{2}\right)\rho\left(E-\frac{\omega}{2}\right)\right]-\mathbb{E}\left[\rho\left(E+\frac{\omega}{2}\right)\right]\mathbb{E}\left[\rho\left(E-\frac{\omega}{2}\right)\right]\sim-\frac{1}{\mathfrak{b}\pi^2\omega^2}, \tag{3.7}$$

where $\mathfrak{b} = 1, 2, 4$ in the orthogonal, unitary, and sympletic ensembles respectively [84]. The right-hand side being negative is a reflection of the well-known level repulsion in quantum chaotic systems [111]. Taking the Fourier transform with respect to $\omega$ gives a term proportional to $T$ for the connected SFF. Such a linear-in-$T$ ramp is often taken as a defining signature of quantum chaos.

- The "plateau", occurring at late times, results from the discreteness of the spectrum. At times much larger than the inverse level spacing, one expects that all off-diagonal terms in the double-trace of the SFF sum to effectively zero, meaning that

$$\text{SFF}(T,f)=\sum_{mn}e^{-i(E_m-E_n)T}f(E_m)f(E_n)\sim\sum_n f(E_n)^2. \tag{3.8}$$

  As the plateau regime is both challenging to access analytically and not particularly informative, we shall not consider it further in this work.

The bulk of our analysis in this paper is devoted to calculation of the ramp in a well-known quantum spin glass model, the $p$-spherical model (discussed below). The results can be

understood via the elementary observation that when a Hamiltonian is block diagonal,

$$H = \begin{pmatrix} H_1 & 0 & 0 & \\ 0 & H_2 & 0 & \dots \\ 0 & 0 & H_3 & \\ & \vdots & & \ddots \end{pmatrix}, \tag{3.9}$$

then $\mathrm{Tr} e^{-iHT} = \sum_k \mathrm{Tr} e^{-iH_k T}$. If the different blocks are independent, then the variance of $\mathrm{Tr} e^{-iHT}$ is the sum of the variance of each $\mathrm{Tr} e^{-iH_k T}$, i.e., *the SFF is the sum of the SFF for each block.* In particular, the coefficient of the universal linear-in-$T$ ramp is multiplied by the number of independent blocks. Systems with only approximately block-diagonal Hamiltonians, for which there are small matrix elements between blocks, have this enhancement of the ramp up to the transition timescale between blocks. For a more detailed analysis, see Ref. [2].

### 3.1.2 Review of mean-field spin glasses

Broadly speaking, spin glasses are systems in which the magnetic moments $\sigma_i$ are frozen but disordered at low temperatures. However, this definition (much like that of "quantum chaos") encompasses a wide variety of phenomena which are in many ways quite distinct, as is made clear by the literature on the subject [15–18, 145–147]. In the present paper, we focus on what are known as "one-step replica symmetry breaking" (1RSB) spin glass phases [147]. We are specifically interested in quantum spin glasses, but we first review the corresponding classical case, for which configurations are labelled by a list $\sigma \equiv \{\sigma_1, \cdots, \sigma_N\}$ and the Hamiltonian is simply a function of $\sigma$.

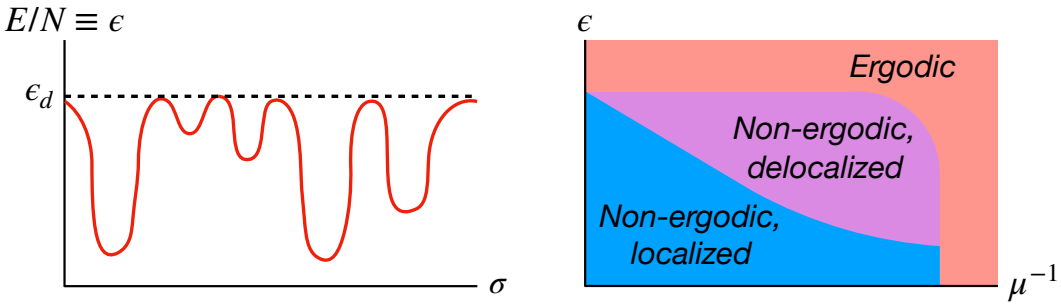



Figure 3.3: (Left) Cartoon of the energy landscape in a 1RSB spin glass. The y-axis is energy per spin, $E/N$, where $E$ is energy and $N$ is the number of spins. Different points on the x-axis represent (very roughly, since the actual configuration space is $N$-dimensional) different spin configurations $\sigma$. The dashed line indicates the energy density $\epsilon_d$ below which the system is non-ergodic. (Right) Sketch of the dynamical phase diagram for a quantum 1RSB spin glass. The x-axis represents parameters controlling the strength of quantum fluctuations, and the y-axis is energy density. Note that many other types of phase transitions are also present, in particular equilibrium transitions, but are not indicated here. See, e.g., Refs. [146, 147, 154] for more information.

While the technical definition of 1RSB is somewhat involved, the qualitative physics is straightforward to understand and captured by the sketch in Fig. 3.3. The energy landscape, i.e., energy as a function of spin configuration, has many deep wells and steep barriers. In particular, the number of wells is $e^{O(N)}$ and the heights of the energy barriers separating wells are $O(N)$, where $N$ is the number of spins. As a result, below a certain energy density $\epsilon_d$, the system is extremely non-ergodic: it remains trapped within an exponentially small fraction of the thermodynamically relevant configuration space until exponentially long timescales. While the 1RSB phenomenon was originally studied in the context of stochastic classical dynamics [11, 155–157], it has recently been shown to imply exponentially long *tunneling* timescales for isolated quantum dynamics as well [108, 158–161].

TAP states (named after Thouless, Anderson, and Palmer [23]) provide a more quantitative description of such "deep wells". Arguably the most general definition (see Ref. [17] for others)

is in terms of the Legendre transform of the free energy with respect to local fields:

$$F\big(\{m_i\}\big) = -\frac{1}{\beta}\log \mathrm{Tr} e^{-\beta H + \beta \sum_i h_i \sigma_i} + \sum_i h_i m_i, \tag{3.10}$$

where $H$ is the Hamiltonian of interest and the fields $\{h_i\}$ are chosen so that $\langle \sigma_i \rangle = m_i$ (where $\langle \cdot \rangle$ indicates a thermal average). TAP states are simply the local minima of $F(\{m_i\})$. Physically, each corresponds to a different "well" of the energy landscape, including thermal fluctuations around the lowest point (thus TAP states do generically depend on temperature). The partition function can be decomposed as a sum over TAP states:

$$Z \equiv \sum_\sigma e^{-\beta H(\sigma)} = \sum_\alpha \left[ \sum_\sigma \delta_{\sigma \in \alpha} e^{-\beta H(\sigma)} \right] \equiv \sum_\alpha Z_\alpha, \tag{3.11}$$

where $\alpha$ denotes a TAP state and $\delta_{\sigma \in \alpha}$ restricts the trace to only those states belonging to TAP state $\alpha$. Note that in this discussion, $\sigma$ can refer to any set of degrees of freedom: Ising spins, vector spins, continuous coordinates, etc. In all cases, Eqs. (3.10) and (3.11) can be interpreted accordingly.

Quantum generalizations of spin glasses are usually obtained by adding non-commuting terms to the Hamiltonian. For example, with an Ising Hamiltonian, one often interprets $\sigma_i$ as the Pauli spin-$z$ operator $\sigma_i^z$ and includes an additional transverse field $\Gamma \sum_i \sigma_i^x$ [162–165]. On the other hand, with systems having continuous degrees of freedom (including the one which we study in this paper), one can interpret $\sigma_i$ as a position coordinate and include the "kinetic energy" $\sum_i \pi_i^2/2\mu$, where $\pi_i$ is the momentum operator conjugate to $\sigma_i$ [166, 167]. Generically, the resulting system has a frozen spin glass phase at low energy and small quantum fluctuations (the

latter being controlled by $\Gamma$ and $\mu^{-1}$ respectively in the examples above), and has a paramagnetic phase at either high energy or large quantum fluctuations. A sketch of the typical phase diagram is shown in Fig. 3.3, with these two phases indicated by "non-ergodic" and "ergodic".

It has recently been noted that quantum 1RSB spin glasses can exhibit *eigenstate* phase transitions which are distinct from the above [110, 168, 169]. Qualitatively speaking, on the low energy/fluctuation side of the eigenstate phase boundary, each eigenstate of the Hamiltonian is localized on a single TAP state. This implies that under the system's internal dynamics alone (i.e., as given by the Schrodinger equation), the system cannot tunnel between TAP states on *any* timescale, even times exponential in the number of spins. On the other side of the phase boundary, each eigenstate is delocalized over many TAP states in accordance with random matrix behavior. As discussed in Ref. [108], while this implies that the system does tunnel between TAP states, the timescale for tunneling is necessarily exponential in system size, analogous to the activation times under open-system dynamics. Only when there exists a single TAP state can one identify the phase as genuinely thermalizing. As a result, one finds phase diagrams like that sketched in Fig. 3.3, with "non-ergodic"/"ergodic" indicating whether multiple TAP states exist and "localized"/"delocalized" referring to the eigenstate properties.

### 3.1.3 Summary of results

In this paper, we calculate the SFF for a particular ensemble of quantum spin glasses, the quantum $p$-spherical model (PSM) [166, 167, 170]. We find that in the ergodic phase, the connected part of the SFF agrees with the expectation from random matrix theory (Eq. (3.62) below), while in the non-ergodic phase, it is enhanced by a factor which is precisely the number

of TAP states (Eq. (3.111)). Given the discussion in Secs. 3.1.1 and 3.1.2, this makes precise and validates the idea that each metastable state (i.e., TAP state) corresponds to a block of the Hamiltonian that is quantum chaotic on its own but is nearly decoupled from all others, thus making the system as a whole non-ergodic [169]. This is the main result of the present work.

We also consider higher moments of the evolution operator and identify a set of saddle points (Eq. (3.136)) which, in addition to confirming the picture that different TAP states have independent level statistics, suggest that at least at low complexity, the number of TAP states at a given energy is Poisson-distributed and independent of other energies. Yet as we shall discuss, since our analysis does not consider the perturbative corrections around each saddle point, this does not constitute a complete calculation and serves more as motivation for future investigation.

## 3.2 Real-time dynamics of the quantum $p$-spherical model

### 3.2.1 The model

The classical $p$-spherical model (PSM) [10] is a disordered spin model with all-to-all $p$-body interactions. It is defined by the classical Hamiltonian

$$H_{\text{cl}} \equiv \sum_{(i_1 \cdots i_p)} J_{i_1 \cdots i_p} \sigma_{i_1} \cdots \sigma_{i_p}, \tag{3.12}$$

where the couplings $J_{i_1 \ldots i_p}$ are independent Gaussian random variables with mean zero and variance

$$\mathbb{E} {J_{i_1 \cdots i_p}}^2 = \frac{J^2 (p-1)!}{C_{i_1 \cdots i_p} N^{p-1}}. \tag{3.13}$$

Here and throughout, $\mathbb{E}$ indicates an average over couplings. The notation $(i_1 \cdots i_p)$ denotes sets of $p$ indices such that $1 \leq i_1 \leq \cdots \leq i_p \leq N$. The sum in Eq. (3.12) is over all such sets. Our treatment differs from the standard convention by including a parameter $J$ for the overall strength of the disorder. To recover the standard expressions, simply set $J^2 = p/2$. We also include the combinatorial factor $C_{i_1 \cdots i_p} = \prod_{1 \leq i \leq N} n_i!$, where $n_i$ is the number of indices set equal to $i$. This term is almost always one, but its inclusion avoids $1/N$ corrections in the action.

The $\sigma_i$ are real, continuous spin variables subject to the spherical constraint

$$\sum_{i=1}^{N} \sigma_i^2 = N, \tag{3.14}$$

which ensures that the system has an extensive free energy. It is apparent that this is a mean-field model without any spatial structure. This allows for infinite free energy barriers around metastable states in the thermodynamic limit, making the model ideal for examining the impact of metastability on the spectral statistics of spin glasses.

In this work, we follow Refs. [166, 167, 170] in generalizing Eq. (3.12) to a quantum Hamiltonian $H$. We treat the $\sigma_i$ as commuting position operators, and define conjugate momentum operators $\pi_i$ which satisfy the commutation relations

$$[\sigma_i, \pi_j] = i\delta_{ij}. \tag{3.15}$$

The *quantum* PSM simply includes a kinetic energy term in the Hamiltonian:

$$H = \sum_{i=1}^{N} \frac{\pi_i^2}{2\mu} + \sum_{(i_1 \cdots i_p)} J_{i_1 \cdots i_p} \sigma_{i_1} \cdots \sigma_{i_p}. \tag{3.16}$$

The mass $\mu$ is an additional parameter controlling the strength of quantum fluctuations. To incorporate the spherical constraint, we take the Hilbert space to be the subspace in which $\sum_i \sigma_i^2$ has eigenvalue $N$.

The quantum PSM may be interpreted as a soft-spin version of the Ising $p$-spin model in an external transverse field — itself the subject of much study [164, 171–174] — where $\mu^{-1}$ is analogous to the transverse field. Alternatively, if we think of $\sigma \equiv \{\sigma_1, \cdots, \sigma_N\}$ as a position vector in $N$-dimensional space, the quantum PSM has a natural interpretation as a particle of mass $\mu$ moving on a hypersphere of radius $\sqrt{N}$. This particle experiences the Gaussian random potential

$$V(\sigma) = \sum_{(i_1 \cdots i_p)} J_{i_1 \cdots i_p} \sigma_{i_1} \cdots \sigma_{i_p}, \tag{3.17}$$

whose correlation function is

$$\mathbb{E}V(\sigma)V(\sigma') = \frac{J^2(p-1)!}{C_{i_1 \cdots i_p} N^{p-1}} \sum_{(i_1 \cdots i_p)} \sigma_{i_1}\sigma'_{i_1} \cdots \sigma_{i_p}\sigma'_{i_p} = \frac{J^2}{pN^{p-1}} \big(\sigma \cdot \sigma'\big)^p. \tag{3.18}$$

Note that there is a very important difference between $p = 2$ and $p > 2$: the former is a Gaussian model, essentially (but for the spherical constraint) a system of linearly coupled harmonic oscillators. It therefore has qualitatively different behavior than the $p > 2$ models, which are genuinely interacting and serve as reasonable toy models for rugged energy landscapes. In this work, we exclusively consider $p > 2$.

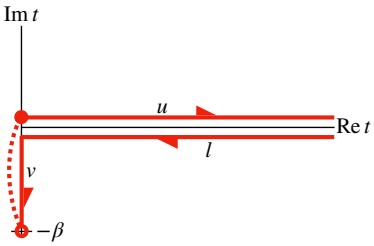
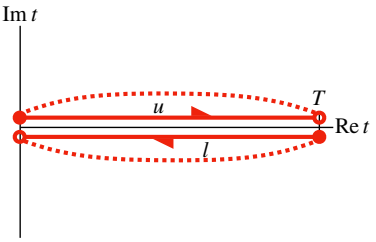
| | Schwinger-Keldysh | Spectral form factor |
|---|---|---|
| Integration contour |  |  |
| Real-time correlation functions | $G_{uu}(t,t') = Z^{-1}\mathrm{Tr}e^{-\beta H}\mathcal{T}\sigma_i(t)\sigma_i(t')$<br>$G_{ul}(t,t') = Z^{-1}\mathrm{Tr}e^{-\beta H}\sigma_i(t')\sigma_i(t)$<br>$G_{lu}(t,t') = Z^{-1}\mathrm{Tr}e^{-\beta H}\sigma_i(t)\sigma_i(t')$<br>$G_{ll}(t,t') = Z^{-1}\mathrm{Tr}e^{-\beta H}\widetilde{\mathcal{T}}\sigma_i(t)\sigma_i(t')$ | $G_{uu}(t,t') = \mathrm{Tr}e^{-iHT}\mathcal{T}\sigma_i(t)\sigma_i(t')\cdot\mathrm{Tr}e^{iHT}$<br>$G_{ul}(t,t') = \mathrm{Tr}e^{-iHT}\sigma_i(t)\cdot\mathrm{Tr}e^{iHT}\sigma_i(t')$<br>$G_{lu}(t,t') = \mathrm{Tr}e^{-iHT}\sigma_i(t')\cdot\mathrm{Tr}e^{iHT}\sigma_i(t)$<br>$G_{ll}(t,t') = \mathrm{Tr}e^{-iHT}\cdot\mathrm{Tr}e^{iHT}\widetilde{\mathcal{T}}\sigma_i(t)\sigma_i(t')$ |
| Equations of motion | $i(\mu\partial_t^2 + z)G(t,t') + \int_{\mathcal{C}} dt'' F(t,t'')G(t'',t') = \delta(t-t'),$<br>$F(t,t') = J^2 G(t,t')^{p-1}, \qquad G(t,t) = 1$ | |

Figure 3.4: Summary of the contours, order parameters, and (at least at high temperature) equations of motion considered in this work. The left column gives the quantities appropriate to the Schwinger-Keldysh path integral, and the right column to the spectral form factor (SFF) path integral.
(Top row) Contours for the respective path integrals. Each of the different branches is labelled, and directions are indicated by arrowheads. Points connected by dashed lines are identified, making the contours periodic.
(Middle row) Relationship between order parameters of the theory and observable quantities. $H$ and $Z$ are the $p$-spin Hamiltonian and partition function respectively. $\mathcal{T}$ and $\widetilde{\mathcal{T}}$ denote time ordering and anti-ordering.
(Bottom row) Equations of motion. These take the same form for both path integrals, differing only in the contour $\mathcal{C}$ being used.

### 3.2.2 Schwinger-Keldysh path integral

Just as other all-to-all models have a saddle-point/mean-field description at large $N$, so too does the PSM. We start with the disorder-averaged (i.e., "annealed") path integral on the Schwinger-Keldysh contour at inverse temperature $\beta$, illustrated in the left column of Fig. 3.4. While it is in general incorrect (often grossly) to disorder-average the path integral itself, it is known that the annealed approximation is accurate in the PSM as long as $\beta$ is less than a critical

value $\beta_s$ [146, 175]. We shall assume that this is true throughout. The annealed path integral is

$$\begin{aligned}\mathbb{E}Z_{\text{SK}} = \int \mathcal{D}\sigma^N \exp\left[\int_{\mathcal{C}} dt \sum_i \left(\frac{i\mu}{2}\big(\partial_t \sigma_i(t)\big)^2 - \frac{iz(t)}{2}\big(\sigma_i(t)^2 - 1\big)\right)\right] \\ \cdot \int dP(J) \exp\left[-i\int_{\mathcal{C}} dt \sum_{(i_1\cdots i_p)} J_{i_1\cdots i_p}\sigma_{i_1}(t)\cdots\sigma_{i_p}(t)\right],\end{aligned} \tag{3.19}$$

where

$$dP(J) \propto \prod_{(i_1\cdots i_p)} dJ_{i_1\cdots i_p} \exp\left[-\frac{N^{p-1}C_{i_1\cdots i_p}J_{i_1\cdots i_p}^2}{2(p-1)!J^2}\right]. \tag{3.20}$$

For brevity, we use $\mathcal{C}$ to denote the entire contour. Thus $\int_{\mathcal{C}} dt$ indicates a contour integral within the complex-$t$ plane. The Lagrange multiplier $z(t)$ is included to enforce the spherical constraint. It can be interpreted as a time-dependent harmonic potential whose value is chosen such that $\sum_i \sigma_i(t)^2 = N$ at all times. Thus the measure $\mathcal{D}\sigma^N$ is simply the product measure over each $\sigma_i$ independently. From here, the same manipulations used to get Schwinger-Dyson equations for the SYK model will give us equations of motion for the PSM.

One can immediately perform the Gaussian integrals over the couplings to obtain

$$\mathbb{E}Z_{\text{SK}} = \int \mathcal{D}\sigma^N e^{-NS'}, \tag{3.21}$$

where

$$
\begin{aligned}
NS' \equiv & \int_{\mathcal{C}} dt \sum_i \left( -\frac{i\mu}{2} \big(\partial_t \sigma_i(t)\big)^2 + \frac{iz(t)}{2} \big(\sigma_i(t)^2 - 1\big) \right) \\
& + \frac{J^2 (p-1)!}{2 C_{i_1 \cdots i_p} N^{p-1}} \sum_{(i_1 \cdots i_p)} \int_{\mathcal{C}} dt dt' \sigma_{i_1}(t) \sigma_{i_1}(t') \cdots \sigma_{i_p}(t) \sigma_{i_p}(t') \\
= & \int_{\mathcal{C}} dt \sum_i \left( -\frac{i\mu}{2} \big(\partial_t \sigma_i(t)\big)^2 + \frac{iz(t)}{2} \big(\sigma_i(t)^2 - 1\big) \right) + \frac{NJ^2}{2p} \int_{\mathcal{C}} dt dt' \left( \frac{1}{N} \sum_i \sigma_i(t) \sigma_i(t') \right)^p .
\end{aligned}
\tag{3.22}
$$

Next introduce a "fat unity",

$$
\begin{aligned}
1 &= \int \mathcal{D}\mathcal{G} \prod_{tt'} \delta\Big( N\mathcal{G}(t,t') - \sum_i \sigma_i(t) \sigma_i(t') \Big) \\
&= \int \mathcal{D}\mathcal{G}\mathcal{D}\mathcal{F} \exp\left[ \frac{N}{2} \int_{\mathcal{C}} dt dt' \mathcal{F}(t,t') \left( \mathcal{G}(t,t') - \frac{1}{N} \sum_i \sigma_i(t) \sigma_i(t') \right) \right] .
\end{aligned}
\tag{3.23}
$$

The integral over the self-energy $\mathcal{F}(t,t')$ runs along the imaginary axis, making the second line simply the identity $\int dp e^{ipx} = 2\pi\delta(x)$ (we absorb factors of $2\pi$ into the measure $\mathcal{D}\mathcal{F}$). However, when we ultimately evaluate the path integral by saddle point, we shall find that the saddle point value of $\mathcal{F}(t,t')$ is real. Inserting Eq. (3.23) into the path integral gives

$$
\mathbb{E} Z_{\text{SK}} = \int \mathcal{D}\mathcal{G}\mathcal{D}\mathcal{F} \int \mathcal{D}\sigma^N e^{-NS''} ,
\tag{3.24}
$$

where

$$
\begin{aligned}
NS'' \equiv & -\frac{iN}{2} \int_{\mathcal{C}} dt z(t) + \frac{N}{2} \int_{\mathcal{C}} dt dt' \left( \frac{J^2}{p} \mathcal{G}(t,t')^p - \mathcal{F}(t,t') \mathcal{G}(t,t') \right) \\
& + \frac{1}{2} \sum_i \left[ \int_{\mathcal{C}} dt \left( -i\mu \big(\partial_t \sigma_i(t)\big)^2 + iz(t) \sigma_i(t)^2 \right) + \int_{\mathcal{C}} dt dt' \sigma_i(t) \mathcal{F}(t,t') \sigma_i(t') \right] .
\end{aligned}
\tag{3.25}
$$

We can now perform the integral over $\sigma_i$, resulting in

$$\mathbb{E}Z_{\text{SK}} = \int \mathcal{D}\mathcal{G}\mathcal{D}\mathcal{F} e^{-NS_{\text{eff}}}, \tag{3.26}$$

where

$$S_{\text{eff}} \equiv -\frac{i}{2}\int_{\mathcal{C}} dt z(t) + \frac{1}{2}\int_{\mathcal{C}} dt dt' \left(\frac{J^2}{p}\mathcal{G}(t,t')^p - \mathcal{F}(t,t')\mathcal{G}(t,t')\right) + \frac{1}{2}\log \text{Det}\Big[i(\mu\partial_t^2 + z) + \mathcal{F}\Big]. \tag{3.27}$$

At large $N$, the remaining path integral can be evaluated within the saddle point approximation. The locations of the saddle points are determined by setting to zero the functional derivatives of Eq. (3.27):

$$\begin{gathered} i\big(\mu\partial_t^2 + z(t)\big)\mathcal{G}(t,t') + \int_{\mathcal{C}} dt'' \mathcal{F}(t,t'')\mathcal{G}(t'',t') = \delta(t-t'), \\ \mathcal{F}(t,t') = J^2\mathcal{G}(t,t')^{p-1}, \qquad \mathcal{G}(t,t) = 1. \end{gathered} \tag{3.28}$$

Keep in mind that the time arguments in Eq. (3.28) are complex and range over the entire Schwinger-Keldysh contour. In particular, although it is hidden in this compact notation, the infinitesimals $dt$ acquire different phases depending on the branch of the contour: $dt$ is a positive real infinitesimal on the upper ("forward") real-time branch, a negative real infinitesimal on the lower ("backward") real-time branch, and a negative imaginary infinitesimal on the thermal branch.

$\mathcal{G}(t,t')$ is the order parameter of this theory. As is clear from the manner by which it was introduced (top line of Eq. (3.23)), expectation values of $\mathcal{G}(t,t')$ within the path integral are equivalent to expectation values of $N^{-1}\sum_i \sigma_i(t)\sigma_i(t')$. The latter are simply time-ordered correlation functions. We shall focus on the real-time correlation functions, for which it is more

transparent to explicitly indicate the branches by $\alpha \in \{u, l\}$ and have $t$ be simply a real variable. Formally, we have that

$$\begin{aligned}
\big\langle \mathcal{G}_{uu}(t,t') \big\rangle &= \mathbb{E}\Big[ Z_{\text{SK}}^{-1} \text{Tr} e^{-\beta H} \mathcal{T} \sigma_i(t) \sigma_i(t') \Big], \qquad & \big\langle \mathcal{G}_{ul}(t,t') \big\rangle &= \mathbb{E}\Big[ Z_{\text{SK}}^{-1} \text{Tr} e^{-\beta H} \sigma_i(t') \sigma_i(t) \Big], \\
\big\langle \mathcal{G}_{lu}(t,t') \big\rangle &= \mathbb{E}\Big[ Z_{\text{SK}}^{-1} \text{Tr} e^{-\beta H} \sigma_i(t) \sigma_i(t') \Big], \qquad & \big\langle \mathcal{G}_{ll}(t,t') \big\rangle &= \mathbb{E}\Big[ Z_{\text{SK}}^{-1} \text{Tr} e^{-\beta H} \widetilde{\mathcal{T}} \sigma_i(t) \sigma_i(t') \Big],
\end{aligned} \tag{3.29}$$

where $\mathcal{T}$ denotes time ordering and $\widetilde{\mathcal{T}}$ denotes time anti-ordering. Note that we can omit the sum over $i$ because the different spins (upon disorder-averaging) have equivalent behavior.

A number of formal properties of $\mathcal{G}_{\alpha\alpha'}(t,t')$ are evident from Eq. (3.29). For one thing, $\mathcal{G}_{\alpha\alpha'}(t,t')$ clearly depends only on the time difference $t-t'$, and we shall often write $\mathcal{G}_{\alpha\alpha'}(t)$ with $t'=0$. Since the four components differ only in time ordering, we see that for *any* function $f(x)$,

$$f\Big(\mathcal{G}_{uu}(t)\Big) + f\Big(\mathcal{G}_{ll}(t)\Big) = f\Big(\mathcal{G}_{ul}(t)\Big) + f\Big(\mathcal{G}_{lu}(t)\Big). \tag{3.30}$$

We can further express all four components in terms of a single complex-valued function (equivalently two real-valued functions). For example, write $\mathcal{G}_{lu}(t)$ in terms of its real and imaginary parts as $\mathcal{G}^R(t) + i\mathcal{G}^I(t)$. Since $\mathcal{G}_{lu}(t)^* = \mathcal{G}_{lu}(-t)$, $\mathcal{G}^R(t)$ is even and $\mathcal{G}^I(t)$ is odd. One can easily confirm that

$$\begin{aligned}
\mathcal{G}_{uu}(t) &= \mathcal{G}^R(t) + i\text{sgn}[t]\mathcal{G}^I(t), \qquad & \mathcal{G}_{ul}(t) &= \mathcal{G}^R(t) - i\mathcal{G}^I(t), \\
\mathcal{G}_{lu}(t) &= \mathcal{G}^R(t) + i\mathcal{G}^I(t), \qquad & \mathcal{G}_{ll}(t) &= \mathcal{G}^R(t) - i\text{sgn}[t]\mathcal{G}^I(t).
\end{aligned} \tag{3.31}$$

One of the most important features of $\mathcal{G}_{\alpha\alpha'}(t,t')$ is the limiting behavior at large $|t-t'|$, as a function of the inverse temperature $\beta$. Numerical solution of Eq. (3.28) demonstrates that there

is a critical value $\beta_d$ (which is less than $\beta_s$):

- For $\beta < \beta_d$, $\lim_{|t-t'|\to\infty} \mathcal{G}_{\alpha\alpha'}(t,t') = 0$. We call this the "ergodic" phase ($\mathbb{E}\langle\sigma_i(t)\rangle = 0$ by symmetry regardless of temperature, and so in this phase $\mathbb{E}\langle\sigma_i(t)\sigma_i(t')\rangle \to \mathbb{E}\langle\sigma_i(t)\rangle\mathbb{E}\langle\sigma_i(t')\rangle$).

- For $\beta_d < \beta < \beta_s$, $\lim_{|t-t'|\to\infty} \mathcal{G}_{\alpha\alpha'}(t,t') = q_{\text{EA}} > 0$. We call this the "non-ergodic" phase. The quantity $q_{\text{EA}}$ is referred to as the "Edwards-Anderson" order parameter.

- For $\beta_s < \beta$, our initial annealed approximation is no longer valid. The replica trick is required to obtain accurate results [15, 16], but (at least for finite-time dynamical properties) the behavior is qualitatively similar to that of the non-ergodic phase.

### 3.2.3 TAP equations on the Schwinger-Keldysh contour

The dynamical calculation described above only hints at the complexity of the non-ergodic phase. A more complete picture emerges from a generalization in the spirit of the TAP equations. Our treatment follows that of Ref. [25], which derived TAP equations on the thermal circle for the quantum PSM. While the extension to real-time dynamics is straightforward, we are not aware of any explicit calculation in the literature. Thus we present a detailed derivation of the following equations in App. B.1.

As discussed in Sec. 3.1.2, the TAP free energy (or Gibbs potential) is the Legendre transform of the free energy with respect to local fields. It is therefore a function of the magnetization $m_i$ of each spin. For the free energy of quantum systems, the magnetization should also have an imaginary time index $m_i(\tau)$. The imaginary-time correlation function $\mathcal{G}(\tau,\tau')$ becomes an additional order parameter.

We define the TAP *action* on the Schwinger-Keldysh contour analogously. It is a function

of the magnetizations $m_i(t)$ and the correlation function $\mathcal{G}(t,t')$, with $t$ again being complex-valued and ranging over the entire contour. Specifically,

$$iNS_{\text{TAP}}[m,\mathcal{G}] \equiv \log \int \mathcal{D}\sigma^N \exp\left[ i\sum_i S_i^0 - i\int_{\mathcal{C}} dt \sum_{(i_1\cdots i_p)} J_{i_1\cdots i_p}\sigma_{i_1}(t)\cdots\sigma_{i_p}(t)\right] + \frac{iN}{2}\int_{\mathcal{C}} dt z(t) - i\int_{\mathcal{C}} dt \sum_i h_i(t)m_i(t) + \frac{iN}{2}\int_{\mathcal{C}} dtdt'\Lambda(t,t')\mathcal{G}(t,t'), \quad (3.32)$$

where $\mathcal{C}$ denotes the Schwinger-Keldysh contour and

$$S_i^0 \equiv \int_{\mathcal{C}} dt \left(\frac{\mu}{2}\big(\partial_t\sigma_i(t)\big)^2 - \frac{z(t)}{2}\sigma_i(t)^2 + h_i(t)\sigma_i(t)\right) - \frac{1}{2}\int_{\mathcal{C}} dtdt'\Lambda(t,t')\sigma_i(t)\sigma_i(t'). \quad (3.33)$$

The fields $h_i(t)$ and $\Lambda(t,t')$ are *not* independent parameters. They are instead chosen so that $\langle\sigma_i(t)\rangle = m_i(t)$ and $N^{-1}\sum_i\langle\sigma_i(t)\sigma_i(t')\rangle = \mathcal{G}(t,t')$, just as $z(t)$ is again chosen to enforce $N^{-1}\sum_i\langle\sigma_i(t)^2\rangle = 1$, where the expectation value is with respect to the action in Eq. (3.32).

Due to the Legendre-transform structure of $S_{\text{TAP}}$, we have that

$$N\frac{\partial S_{\text{TAP}}}{\partial m_i(t)} = -h_i(t), \qquad \frac{\partial S_{\text{TAP}}}{\partial \mathcal{G}(t,t')} = \frac{1}{2}\Lambda(t,t'). \quad (3.34)$$

The TAP equations are those for $m_i(t)$ and $\mathcal{G}(t,t')$ which one gets by setting the right-hand sides of Eq. (3.34) to zero. The solutions are therefore the values of magnetization and correlation function which the system can consistently possess "on its own," without any external fields. In this sense, each solution corresponds to a distinct metastable state. There is no reason why there cannot be many self-consistent solutions, and indeed, spin glass models such as the PSM do have many at sufficiently low temperature.

We calculate the TAP equations in App. B.1. They are simplified by the fact that we can take $m_i(t) = m$ and $z(t) = z$. We also define $q_{\text{EA}} \equiv N^{-1}\sum_i m_i^2$. The equations come out to be (together with $\mathcal{G}(t,t) = 1$)

$$i\big(\mu\partial_t^2 + z\big)\Big(\mathcal{G}(t,t') - q_{\text{EA}}\Big) + J^2\int_{\mathcal{C}} dt''\Big(\mathcal{G}(t,t'')^{p-1} - q_{\text{EA}}^{p-1}\Big)\Big(\mathcal{G}(t'',t') - q_{\text{EA}}\Big) = \delta(t-t'), \tag{3.35}$$

$$\begin{aligned} J^2\int_{\mathcal{C}} dt'\Big(\mathcal{G}(t,t')^{p-1} - (p-1)q_{\text{EA}}^{p-2}\mathcal{G}(t,t') + (p-2)q_{\text{EA}}^{p-1}\Big)m_i \\ = -izm_i - i\sum_{(i_1\cdots i_p)} J_{i_1\cdots i_p}\frac{\partial(m_{i_1}\cdots m_{i_p})}{\partial m_i}. \end{aligned} \tag{3.36}$$

Note that Eq. (3.36) is $N$ equations, one for each spin $i$, and that it holds equally for any value of $t$ due to time translation invariance. Defining $\mathcal{F}(t,t') \equiv J^2\mathcal{G}(t,t')^{p-1}$, Eq. (3.35) is quite similar to Eq. (3.28). The only difference is that Eq. (3.35) uses $\Delta\mathcal{G}(t,t') \equiv \mathcal{G}(t,t') - q_{\text{EA}}$ and $\Delta\mathcal{F}(t,t') \equiv \mathcal{F}(t,t') - J^2 q_{\text{EA}}^{p-1}$, which decay to zero at large $|t-t'|$, rather than $\mathcal{G}(t,t')$ and $\mathcal{F}(t,t')$ themselves.

Despite the more involved derivation, $\mathcal{G}(t,t')$ remains a contour-ordered expectation value. Thus, returning to the notation in which $\alpha \in \{u,l\}$ labels branches and $t$ is real, $\mathcal{G}_{\alpha\alpha'}(t-t')$ possesses the same formal properties as discussed in the previous subsection (Eqs. (3.30) and (3.31)). Of particular importance will be the Fourier transform of $\Delta\mathcal{G}_{\alpha\alpha'}(t)$ at zero frequency, denoted $\Delta\widetilde{\mathcal{G}}_{\alpha\alpha'}(0)$, as well as its (matrix) inverse, $\Delta\widetilde{\mathcal{G}}^{-1}_{\alpha\alpha'}(0)$. Also define $L \equiv \int_{-\infty}^{\infty} dt\Delta\mathcal{G}^R(t)$ and $\Lambda \equiv \int_0^{\infty} dt\Delta\mathcal{G}^I(t)$. Then from Eq. (3.31), we see that

$$\begin{pmatrix} \Delta\widetilde{\mathcal{G}}_{uu}(0) & \Delta\widetilde{\mathcal{G}}_{ul}(0) \\ \Delta\widetilde{\mathcal{G}}_{lu}(0) & \Delta\widetilde{\mathcal{G}}_{ll}(0) \end{pmatrix} = \begin{pmatrix} L + 2i\Lambda & L \\ L & L - 2i\Lambda \end{pmatrix}, \tag{3.37}$$

$$\begin{pmatrix} \Delta\widetilde{\mathcal{G}}_{uu}(0) & \Delta\widetilde{\mathcal{G}}_{ul}(0) \\ \Delta\widetilde{\mathcal{G}}_{lu}(0) & \Delta\widetilde{\mathcal{G}}_{ll}(0) \end{pmatrix}^{-1} = \frac{1}{4\Lambda^2} \begin{pmatrix} L - 2i\Lambda & -L \\ -L & L + 2i\Lambda \end{pmatrix}. \tag{3.38}$$

The multiplicity of solutions to the TAP equations comes from Eq. (3.36). By use of Eqs. (3.35), (3.37), and (3.38), it can be written (associating $u$ with 0 and $l$ with 1)

$$\begin{aligned} &\left[(-1)^{\alpha} \sum_{\alpha'} \Delta\widetilde{\mathcal{G}}^{-1}_{\alpha\alpha'}(0) - (p-1)J^2 q_{\text{EA}}^{p-2} \sum_{\alpha'} (-1)^{\alpha'} \Delta\widetilde{\mathcal{G}}_{\alpha\alpha'}(0)\right] m_i \\ &\qquad = \left[\frac{1}{2i\Lambda} - (p-1)J^2 q_{\text{EA}}^{p-2} 2i\Lambda\right] m_i = -i \sum_{(i_1 \cdots i_p)} J_{i_1 \cdots i_p} \frac{\partial(m_{i_1} \cdots m_{i_p})}{\partial m_i}. \end{aligned} \tag{3.39}$$

Eq. (3.39) is identical to that which appears and has been well-studied for the *classical* PSM [10, 11, 146, 176]. Thus we simply quote the following results. In addition to the inverse temperature $\beta$, solutions to Eq. (3.39) are parametrized by the quantity

$$\mathcal{E} \equiv \frac{1}{N J q_{\text{EA}}^{p/2}} \sum_{(i_1 \cdots i_p)} J_{i_1 \cdots i_p} m_{i_1} \cdots m_{i_p}, \tag{3.40}$$

which can be interpreted as a "normalized" potential energy density: each magnetization has a value which is (very roughly) comparable to $q_{\text{EA}}^{1/2}$, and thus the natural scale for the interaction energy is $J q_{\text{EA}}^{p/2}$. The value of $q_{\text{EA}}$ for a given $\mathcal{E} < 0$ is given by the largest solution to

$$-2J q_{\text{EA}}^{p/2-1} \Lambda = \frac{p}{2(p-1)} \left(-\mathcal{E} - \sqrt{\mathcal{E}^2 - \mathcal{E}_{\text{th}}^2}\right), \qquad \mathcal{E}_{\text{th}} \equiv -\frac{2\sqrt{p-1}}{p}, \tag{3.41}$$

where $\Lambda$ depends on $q_{\text{EA}}$ through Eq. (3.35). One can show that solutions to Eq. (3.41) exist only for $\beta > \beta_d$, with $\beta_d$ the same as defined in Sec. 3.2.2. Furthermore, Eq. (3.41) only makes sense if $\mathcal{E} \leq \mathcal{E}_{\text{th}}$. In that case, the number of solutions $\mathcal{N}(\beta, \mathcal{E})$ to Eq. (3.39) — in addition to the trivial

solution $m_i = 0$ — is exponential in system size: $N^{-1} \log \mathcal{N}(\beta, \mathcal{E}) \sim \Sigma(\mathcal{E})$, with[1]

$$\Sigma(\mathcal{E}) = \frac{1}{2}\left(1 + 2\log\frac{p}{2}\right) - \frac{p\mathcal{E}^2}{2} + \frac{p^2}{8(p-1)}\left(\mathcal{E} + \sqrt{\mathcal{E}^2 - \mathcal{E}_{\text{th}}^2}\right)^2 + \log\left(-\mathcal{E} + \sqrt{\mathcal{E}^2 - \mathcal{E}_{\text{th}}^2}\right). \tag{3.42}$$

The exponent $\Sigma(\mathcal{E})$ is referred to as the "complexity" in the spin glass literature.

The connection between this TAP approach and the conventional Schwinger-Keldysh path integral lies in the fact that: i) the inverse temperature $\beta_d$ at which TAP states with non-zero magnetization appear is identical to that at which the autocorrelation function acquires a non-zero late-time limit; ii) the overlap determined by Eq. (3.41) is identical to the late-time value of the autocorrelation function. This strongly suggests the following picture:

- For $\beta < \beta_d$ (the "ergodic" phase), there exists a single equilibrium state with zero magnetization, and the correlation function decays to zero on a finite timescale.

- For $\beta_d < \beta$ (the "non-ergodic" phase), there exist exponentially many metastable states having non-zero magnetization. The number of states is given by the exponential of the complexity $\Sigma(\mathcal{E})$. Dynamically, in the $N \to \infty$ limit, a system prepared in one metastable state will remain in that state for all time. At finite $N$, it is only on a timescale exponential in $N$ that the system can transition between states.

Much more can be said about these phases (in particular how the replica-symmetry-breaking transition at $\beta_s$ appears within the TAP approach), and a large body of literature is devoted to this topic. We refer in particular to Ref. [146] as an excellent starting point.

In the present work, we determine the spectral statistics of the PSM in both the ergodic and

[1] As written, Eq. (3.42) is a bit sloppy. $\mathcal{N}(\mathcal{E})$ is given by Eq. (3.42) when the latter is non-negative and $\beta$ is such that solutions to Eq. (3.41) exist. In all other cases, $\mathcal{N}(\mathcal{E}) = 0$.

non-ergodic phase. Those of the former can be computed very much along the lines of Ref. [75], which we do in Sec. 3.3. Those of the latter, however, require novel calculations which we present in Sec. 3.4. Unsurprisingly, the properties of TAP states shall play an essential role.

## 3.3 The semiclassical ramp in the ergodic phase

To reiterate, we are evaluating

$$\mathrm{SFF}(T,f) \equiv \mathbb{E}\left|\mathrm{Tr} f(H) e^{-iHT}\right|^2 = \mathbb{E}\left[\mathrm{Tr} f(H) e^{-iHT} \mathrm{Tr} f(H) e^{iHT}\right], \tag{3.43}$$

where $H$ is the PSM Hamiltonian (Eq. (3.16)) and $f$ is a filter function as discussed in Sec. 3.1.1. Here we consider the ergodic phase, for which the results are analogous to those of SYK [75]. We then consider the non-ergodic phase in Sec. 3.4.

### 3.3.1 Effective action

The calculation begins by retracing the steps described in Sec. 3.2.2, only on a modified contour. We still have upper and lower branches indicated by $\alpha \in \{u, l\}$ (with $u = 0$ and $l = 1$), but now each is separately periodic. Furthermore, we no longer have a thermal branch. See the right column of Fig. 3.4, as compared to the left column. While some care is required to account for the filter functions (as discussed in Appendix B.3), we ultimately arrive at an expression analogous to Eq. (3.27):

$$\mathrm{SFF}(T,f) = \int \mathcal{D}G\mathcal{D}F\, f\big(\epsilon_u[G]\big) f\big(\epsilon_l[G]\big) e^{-NS_{\mathrm{eff}}[G,F]}, \tag{3.44}$$

$$S_{\text{eff}}[G,F] = \frac{1}{2}\int_0^T dtdt' \sum_{\alpha\alpha'}(-1)^{\alpha+\alpha'}\left(\frac{J^2}{p}G_{\alpha\alpha'}(t,t')^p - F_{\alpha\alpha'}(t,t')G_{\alpha\alpha'}(t,t')\right)$$
$$-\frac{i}{2}\int_0^T dt\sum_{\alpha}(-1)^{\alpha}z_\alpha(t) + \frac{1}{2}\log\text{Det}\Big[i(-1)^{\alpha}\delta_{\alpha\alpha'}\big(\mu\partial_t^2 + z_\alpha\big) + (-1)^{\alpha+\alpha'}F_{\alpha\alpha'}\Big], \tag{3.45}$$

where the "energy density" $\epsilon_\alpha[G]$ is defined as ($0^+$ denotes a positive infinitesimal)

$$\epsilon_\alpha[G] \equiv -\frac{\mu}{2}\partial_t^2 G_{\alpha\alpha}(0^+,0) - \frac{iJ^2}{p}\int_0^T dt\sum_{\alpha'}(-1)^{\alpha'}G_{\alpha\alpha'}(t,0)^p. \tag{3.46}$$

See App. B.3 for details. The saddle point of $S_{\text{eff}}$ is given by the equations (compare to Eq. (3.28))

$$i\big(\mu\partial_t^2 + z_\alpha(t)\big)G_{\alpha\alpha'}(t,t') + \int_0^T dt''\sum_{\alpha''}(-1)^{\alpha''}F_{\alpha\alpha''}(t,t'')G_{\alpha''\alpha'}(t'',t') = (-1)^{\alpha}\delta_{\alpha\alpha'}\delta(t-t'),$$
$$F_{\alpha\alpha'}(t,t') = J^2 G_{\alpha\alpha'}(t,t')^{p-1}, \qquad G_{\alpha\alpha}(t,t) = 1. \tag{3.47}$$

Denoting averages with respect to the path integral of Eq. (3.44) by $\langle\,\cdot\,\rangle$, the expectation value of $G$ is related to the original degrees of freedom as follows (we omit the filter functions here for brevity):

$$\big\langle G_{uu}(t,t')\big\rangle = \mathbb{E}\Big[\text{Tr}e^{-iHT}\mathcal{T}\sigma_i(t)\sigma_i(t')\text{Tr}e^{iHT}\Big], \quad \big\langle G_{ul}(t,t')\big\rangle = \mathbb{E}\Big[\text{Tr}e^{-iHT}\sigma_i(t)\text{Tr}e^{iHT}\sigma_i(t')\Big],$$
$$\big\langle G_{lu}(t,t')\big\rangle = \mathbb{E}\Big[\text{Tr}e^{-iHT}\sigma_i(t')\text{Tr}e^{iHT}\sigma_i(t)\Big], \qquad \big\langle G_{ll}(t,t')\big\rangle = \mathbb{E}\Big[\text{Tr}e^{-iHT}\text{Tr}e^{iHT}\widetilde{\mathcal{T}}\sigma_i(t)\sigma_i(t')\Big], \tag{3.48}$$

where $\mathcal{T}$ denotes time ordering and $\widetilde{\mathcal{T}}$ denotes time anti-ordering. One immediately sees from Eq. (3.48) that:

1. all components of $\langle G_{\alpha\alpha'}(t,t')\rangle$ are time-translation invariant and have period $T$;

2. $\langle G_{uu}(t,t')\rangle$ and $\langle G_{ll}(t,t')\rangle$ are even functions of $t-t'$;

3. $\langle G_{ul}(t,t')\rangle$ and $\langle G_{lu}(t,t')\rangle$ are in fact independent of both time arguments;

4. $\langle G_{uu}(t,t')\rangle^* = \langle G_{ll}(t,t')\rangle$;

5. $\langle G_{ul}(t,t')\rangle^* = \langle G_{lu}(t,t')\rangle$.

Solutions to Eq. (3.47) do *not* necessarily share all these properties, since some of the symmetries may be spontaneously broken.

However, one simple solution that obeys all of the above is to take $G_{ul}(t,t') = G_{lu}(t,t') = 0$. The resulting action is precisely what one would get from averaging each factor of $\mathrm{Tr} e^{-iHT}$ separately, i.e., this solution gives the disconnected contribution to the SFF:

$$\mathbb{E}\Big[\mathrm{Tr} f(H) e^{-iHT} \mathrm{Tr} f(H) e^{iHT}\Big] = \mathbb{E}\Big[\mathrm{Tr} f(H) e^{-iHT}\Big]\mathbb{E}\Big[\mathrm{Tr} f(H) e^{iHT}\Big] + \cdots, \tag{3.49}$$

where $\cdots$ denotes the contribution to the path integral from non-zero $G_{ul}$ and/or $G_{lu}$. Eq. (3.49) holds equally well in the non-ergodic phase, and thus the remainder of this paper will be concerned with determining those additional contributions.

### 3.3.2 Connected solutions

Following Ref. [75], we construct approximate solutions to Eq. (3.47) which become accurate at large $T$. We first present the solutions and justify them afterwards. Take $\mathcal{G}_{\alpha\alpha'}(t,t')$ to be the Schwinger-Keldysh correlation function at inverse temperature $\beta_{\text{aux}}$, exactly as given in Sec. 3.2.2 (Eq. (3.29) in particular). Again define $\mathcal{F}_{\alpha\alpha'}(t,t') \equiv J^2 \mathcal{G}_{\alpha\alpha'}(t,t')^{p-1}$. A solution to the

SFF saddle point equations (up to terms which vanish at large $T$) is

$$G_{\alpha\alpha'}(t,t') = \sum_{n=-\infty}^{\infty} \mathcal{G}_{\alpha\alpha'}(t-t'+\delta_{\alpha\neq\alpha'}\Delta + nT), \tag{3.50}$$

$$F_{\alpha\alpha'}(t,t') = \sum_{n=-\infty}^{\infty} \mathcal{F}_{\alpha\alpha'}(t-t'+\delta_{\alpha\neq\alpha'}\Delta + nT). \tag{3.51}$$

Here $\Delta$ can be any real number between 0 and $T$. Thus Eqs. (3.50) and (3.51) constitute a two-parameter family of solutions, the parameters being $\beta_{\text{aux}}$ and $\Delta$. Every such solution contributes to the SFF.

As for the Lagrange multipliers $z_\alpha(t)$, they are independent of $t$ due to time translation invariance. We further have that $z_u = z_l \equiv z$: both equal the value of the chemical potential needed to satisfy the *equilibrium* spherical constraint, i.e., $N^{-1}\sum_i \text{Tr} Z_{\text{SK}}^{-1} e^{-\beta_{\text{aux}} H}\sigma_i^2 = 1$ (time translation invariance then implies that $\mathcal{G}_{\alpha\alpha}(t,t) = 1$ for all times and both branches).

To justify Eqs. (3.50) and (3.51), it is essential that $\mathcal{G}_{\alpha\alpha'}(t-t')$ decay exponentially to zero as $|t-t'| \to \infty$. Thus these solutions only apply in the ergodic phase. With this in mind, the following comments together establish their validity:

- The sum over $n$ ensures that $G_{\alpha\alpha'}(t-t')$ has period $T$, even though $\mathcal{G}_{\alpha\alpha'}(t-t')$ does not.
- Since $\mathcal{G}_{\alpha\alpha}(t-t')$ decays exponentially, $G_{\alpha\alpha}(0) \sim 1$ up to terms which are exponentially small in $T$.
- The equation $F_{\alpha\alpha'}(t,t') = J^2 G_{\alpha\alpha'}(t,t')^{p-1}$ is satisfied up to exponentially small terms because, when raising Eq. (3.50) to the $p-1$'th power, all cross terms are exponentially

small (as is the sum over them). In other words,

$$\left(\sum_n \mathcal{G}_{\alpha\alpha'}(t-t'+nT+\delta_{\alpha\neq\alpha'}\Delta)\right)^{p-1} \sim \sum_n \mathcal{G}_{\alpha\alpha'}(t-t'+nT+\delta_{\alpha\neq\alpha'}\Delta)^{p-1}. \quad (3.52)$$

- $\mathcal{G}_{\alpha\alpha'}(t,t')$ obeys Eq. (3.28), written explicitly in terms of components as

$$\begin{aligned} &i\big(\mu\partial_t^2+z\big)\mathcal{G}_{\alpha\alpha'}(t-t')+\int_0^\infty dt''\sum_{\alpha''}(-1)^{\alpha''}\mathcal{F}_{\alpha\alpha''}(t-t'')\mathcal{G}_{\alpha''\alpha'}(t''-t') \\ &\qquad -i\int_0^{\beta_{\text{aux}}} d\tau''\mathcal{F}_{\alpha v}(t+i\tau'')\mathcal{G}_{v\alpha'}(-i\tau''-t')=(-1)^\alpha\delta_{\alpha\alpha'}\delta(t-t'), \quad (3.53)\end{aligned}$$

where $v$ denotes the thermal branch of the contour. For $t,t'\gg 1$ (which still allows $t-t'$ to take any value), $\mathcal{G}_{\alpha v}(t+i\tau)$ is exponentially small for all $\tau$ and the last term on the left-hand side can be neglected. We can also take the lower limit of the $t''$ integral to $-\infty$. Thus when checking whether Eq. (3.50) satisfies Eq. (3.47), we have that

$$\begin{aligned} &i\big(\mu\partial_t^2+z\big)G_{\alpha\alpha'}(t-t')+\int_0^T dt''\sum_{\alpha''}(-1)^{\alpha''}F_{\alpha\alpha''}(t-t'')G_{\alpha''\alpha'}(t''-t') \\ &\sim i\big(\mu\partial_t^2+z\big)\mathcal{G}_{\alpha\alpha'}(t-t')+\int_{-\infty}^\infty dt''\sum_{\alpha''}(-1)^{\alpha''}\mathcal{F}_{\alpha\alpha''}(t-t'')\mathcal{G}_{\alpha''\alpha'}(t''-t') \quad (3.54) \\ &\sim (-1)^\alpha\delta_{\alpha\alpha'}\delta(t-t'),\end{aligned}$$

again making use of the fact that $\mathcal{G}_{\alpha\alpha'}(t-t')$ is exponentially small when $|t-t'|$ is large. The equation is indeed satisfied.

- Finally, the off-diagonal components $G_{ul}(t,t')$ and $G_{lu}(t,t')$ contain the parameter $\Delta$ because they break the *separate* time translation symmetries in $t$ and $t'$ (see property iii

above). Thus if any choice of $\Delta$ solves Eq. (3.47), so do all choices of $\Delta \in [0, T)$.

As noted above, we have thus identified a two-parameter family of solutions to the SFF saddle point equations. In what follows it will be more convenient to parametrize the solutions by the equilibrium energy density $\epsilon(\beta_{\text{aux}})$ corresponding to inverse temperature $\beta_{\text{aux}}$. We can express $\epsilon(\beta)$ in terms of $\mathcal{G}$ (and thus $G$) by inserting a factor of $H$ into the Schwinger-Keldysh contour. Since $H$ clearly commutes with the evolution operator $e^{-\beta H} e^{iHt} e^{-iHt}$, it can be inserted at any point, in particular at a late time for which (again because $\mathcal{G}_{\alpha\alpha'}(t-t')$ decays exponentially) the thermal branch can be neglected. By following the same steps as in Appendix B.3, we find that $\epsilon(\beta)$ is given precisely by Eq. (3.46), evaluated on either branch:

$$\begin{aligned} \epsilon &= -\frac{\mu}{2}\partial_t^2 \mathcal{G}_{uu}(0^+) - \frac{iJ^2}{p}\int_{-\infty}^{\infty} dt \Big(\mathcal{G}_{uu}(t)^p - \mathcal{G}_{ul}(t)^p\Big) \\ &= -\frac{\mu}{2}\partial_t^2 \mathcal{G}_{ll}(0^+) + \frac{iJ^2}{p}\int_{-\infty}^{\infty} dt \Big(\mathcal{G}_{ll}(t)^p - \mathcal{G}_{lu}(t)^p\Big). \end{aligned} \tag{3.55}$$

### 3.3.3 Contribution of connected solutions

Having demonstrated that Eqs. (3.50) and (3.51) solve the SFF saddle point equations, it remains to calculate the action (Eq. (3.45)) evaluated at the solutions. First note that, since each solution obeys Eq. (3.47), we can rewrite the action as

$$S_{\text{eff}} = -\frac{J^2(p-1)T}{2p}\int_0^T dt \sum_{\alpha\alpha'} (-1)^{\alpha+\alpha'} G_{\alpha\alpha'}(t)^p - \frac{1}{2}\sum_{\omega} \log \text{Det}\widetilde{G}_{\alpha\alpha'}(\omega), \tag{3.56}$$

where $\omega \in 2\pi\mathbb{Z}/T$ and $\widetilde{G}_{\alpha\alpha'}(\omega) \equiv \int_0^T dt e^{i\omega t} G_{\alpha\alpha'}(t)$. Note that the Lagrange multiplier terms have dropped out since $z_u = z_l$. Furthermore, since $\int dt G_{\alpha\alpha'}(t)^p \sim \int dt \mathcal{G}_{\alpha\alpha'}(t)^p$, the general

relation in Eq. (3.30) implies that the first term of Eq. (3.56) in fact vanishes.

For the second term, note that by Eq. (3.50),

$$\widetilde{G}_{\alpha\alpha'}(\omega) = e^{-i\delta_{\alpha\neq\alpha'}\omega\Delta}\widetilde{\mathcal{G}}_{\alpha\alpha'}(\omega), \qquad \widetilde{\mathcal{G}}_{\alpha\alpha'}(\omega) \equiv \int_{-\infty}^{\infty} dt e^{i\omega t}\mathcal{G}_{\alpha\alpha'}(t). \tag{3.57}$$

The exponential decay of $\mathcal{G}_{\alpha\alpha'}(t)$ implies that $\widetilde{\mathcal{G}}_{\alpha\alpha'}(\omega)$ (and thus $\widetilde{G}_{\alpha\alpha'}(\omega)$) is an infinitely differentiable function of $\omega$. Strictly speaking, since the path integral is regularized by a timestep $\Delta t \to 0$, $\widetilde{\mathcal{G}}_{\alpha\alpha'}(\omega)$ is furthermore periodic with period $2\pi/\Delta t$. The same is true of $\log \mathrm{Det}\widetilde{G}_{\alpha\alpha'}(\omega)$. Thus the Euler-Maclaurin formula [177] gives

$$\sum_{n=-\pi/\Delta t}^{\pi/\Delta t} \log \mathrm{Det}\widetilde{G}_{\alpha\alpha'}\left(\frac{2\pi n}{T}\right) \sim \frac{T}{2\pi}\int_{-\pi/\Delta t}^{\pi/\Delta t} d\omega \log \mathrm{Det}\widetilde{G}_{\alpha\alpha'}(\omega) \to \frac{T}{2\pi}\int_{-\infty}^{\infty} d\omega \log \mathrm{Det}\widetilde{G}_{\alpha\alpha'}(\omega), \tag{3.58}$$

up to terms which vanish faster than any polynomial in $T^{-1}$. Thus $S_{\text{eff}}$ is proportional to $T$, and we only need to evaluate the proportionality constant.

Rather than calculate the integral directly, we follow Ref. [75] and evaluate the derivative $dS_{\text{eff}}/dT$ starting from Eq. (3.45). It is convenient to rescale time as $t \to Tt$, so that $T$ becomes simply another parameter:

$$\begin{aligned} S_{\text{eff}} = \frac{T^2}{2}\int_0^1 dt dt' \sum_{\alpha\alpha'}(-1)^{\alpha+\alpha'}\left(\frac{J^2}{p}G_{\alpha\alpha'}(t,t')^p - F_{\alpha\alpha'}(t,t')G_{\alpha\alpha'}(t,t')\right) \\ + \frac{1}{2}\log \mathrm{Det}\Big[i(-1)^{\alpha}\delta_{\alpha\alpha'}\left(\mu T^{-2}\partial_t^2 + z_\alpha\right) + (-1)^{\alpha+\alpha'}F_{\alpha\alpha'}\Big]. \end{aligned} \tag{3.59}$$

Note that, since $S_{\text{eff}}$ is evaluated at a solution of the saddle point equations, we only need to

differentiate the explicit factors of $T$:

$$\frac{dS_{\text{eff}}}{dT} = T\int_0^1 dtdt' \sum_{\alpha\alpha'}(-1)^{\alpha+\alpha'}\left(\frac{J^2}{p}G_{\alpha\alpha'}(t,t')^p - F_{\alpha\alpha'}(t,t')G_{\alpha\alpha'}(t,t')\right) \\ -\frac{i\mu}{T^3}\int_0^1 dt\sum_{\alpha}(-1)^{\alpha}\partial_t^2\Big[i(-1)^{\alpha}\delta_{\alpha\alpha'}\big(\mu T^{-2}\partial_t^2+z_\alpha(t)\big)+(-1)^{\alpha+\alpha'}F_{\alpha\alpha'}\Big]^{-1}\Big|_{\alpha=\alpha',t=t'^+}. \tag{3.60}$$

Returning to unscaled time and using Eq. (3.47), we have that

$$\frac{dS_{\text{eff}}}{dT} = -\frac{(p-1)J^2}{p}\int_0^T dt\sum_{\alpha\alpha'}(-1)^{\alpha+\alpha'}G_{\alpha\alpha'}(t)^p - \frac{i\mu}{T}\sum_{\alpha}(-1)^{\alpha}\partial_t^2 G_{\alpha\alpha}(0^+) = 0, \tag{3.61}$$

again using Eqs. (3.30) and (3.55). Thus the proportionality constant is in fact zero, i.e., $S_{\text{eff}} = 0$.

### 3.3.4 Evaluation of the SFF

To finally compute the SFF, we simply need to sum over all connected solutions, i.e., integrate over $\epsilon_{\text{aux}}$ and $\Delta$. However, there are additional discrete symmetries which give further solutions: i) we can time-reverse the off-diagonal components, i.e., take $G_{ul}(t) = \mathcal{G}_{ul}(-t)$ and $G_{lu}(t) = \mathcal{G}_{lu}(-t)$; ii) if $p$ is even, we can take $G_{ul}(t) = -\mathcal{G}_{ul}(t)$ and $G_{lu}(t) = -\mathcal{G}_{lu}(t)$. These must be summed over as well, giving an additional factor of $2(1+\delta_{p\text{ even}})$, where $\delta_{p\text{ even}}$ is the indicator function on $p$ being even (1 if true, 0 if false). Thus our final expression is

$$\mathbb{E}\Big[\text{Tr}f(H)e^{-iHT}\text{Tr}f(H)e^{iHT}\Big] \sim \big|\mathbb{E}\text{Tr}f(H)e^{-iHT}\big|^2 + \int\frac{d\epsilon_{\text{aux}}}{2\pi}f(\epsilon_{\text{aux}})^2\int_0^T d\Delta\, 2\big(1+\delta_{p\text{ even}}\big)e^0 \\ = \big|\mathbb{E}\text{Tr}f(H)e^{-iHT}\big|^2 + 2\big(1+\delta_{p\text{ even}}\big)T\int\frac{d\epsilon_{\text{aux}}}{2\pi}f(\epsilon_{\text{aux}})^2. \tag{3.62}$$

The measure $1/2\pi$ can be derived using hydrodynamic methods [2, 75], but its precise value is not essential for our purposes. The key feature is simply that the linear-in-$T$ ramp has emerged.

However, keep in mind that Eq. (3.62) is only valid if the filter function is such that all contributing values of $\epsilon_{\text{aux}}$ lie in the ergodic phase. In the following section we modify this analysis to hold in the non-ergodic phase as well. We shall see that it is necessary to incorporate the structure of multiple TAP states.

## 3.4 The semiclassical ramp in the non-ergodic phase

As we have stressed repeatedly, the results of Sec. 3.3 rely heavily on having an equilibrium correlation function which decays to zero at late times. Thus a new approach is needed to calculate the SFF in the non-ergodic phase, where $\mathcal{G}_{\alpha\alpha'}(t-t') \to q_{\text{EA}} \neq 0$ as $|t-t'| \to \infty$. More specifically, we can no longer neglect the integral over the thermal branch in Eq. (3.28), and $\mathcal{G}_{\alpha\alpha'}(t-t')$ no longer solves the SFF equations of motion (Eq. (3.47)).

However, in the *TAP* equations of motion, Eq. (3.35), we *can* neglect the thermal branch since $\mathcal{G}(t) - q_{\text{EA}}$ does decay to zero exponentially quickly. This suggests that a viable strategy is to construct solutions for the SFF using the TAP correlation function. Since TAP states are parametrized by the quantity $\mathcal{E}$ in Eq. (3.40), it will be necessary to first modify the SFF path integral so as to involve $\mathcal{E}$. We associate the magnetizations and overlap from the TAP approach with time-averaged functions of the spin configuration, namely

$$m_i[\sigma] \equiv \frac{1}{T}\int_0^T dt \sigma_{iu}(t), \qquad q[\sigma] \equiv \frac{1}{T^2}\int_0^T dt dt' \frac{1}{N}\sum_i \sigma_{iu}(t)\sigma_{iu}(t'). \tag{3.63}$$

The choice to use only the upper contour in defining $m_i[\sigma]$ and $q[\sigma]$ will become convenient

in Sec. 3.5, but for now one could equally well use any other combination of branches, say the average of $\sigma_i(t)$ over the lower branch or over both branches symmetrically. With these definitions, we introduce $\mathcal{E}$ via Eq. (3.40).

### 3.4.1 Effective action

To begin, insert an additional fat unity into the path integral:

$$1 = \int d\mathcal{E}_{\text{aux}} \delta \left[ N\mathcal{E}_{\text{aux}} - \frac{1}{Jq[\sigma]^{p/2}} \sum_{(i_1 \cdots i_p)} J_{i_1 \cdots i_p} \left( \frac{1}{T} \int_0^T dt \sigma_{i_1 u}(t) \right) \cdots \left( \frac{1}{T} \int_0^T dt \sigma_{i_p u}(t) \right) \right]. \tag{3.64}$$

With this addition, the full path integral is

$$\begin{aligned} \text{SFF} = & \\ & \int \mathcal{D}P(J) \mathcal{D}\sigma^N d\mathcal{E}_{\text{aux}} d\lambda \exp \left[ \frac{i}{2} \sum_i \int_0^T dt \sum_\alpha (-1)^\alpha \Big[ \mu \big( \partial_t \sigma_{i\alpha}(t) \big)^2 - z_\alpha(t) \big( \sigma_{i\alpha}(t)^2 - 1 \big) \Big] \right] \\ & \qquad \times \exp \left[ -i \sum_{(i_1 \cdots i_p)} J_{i_1 \cdots i_p} \int_0^T dt \sum_\alpha (-1)^\alpha \sigma_{i_1 \alpha}(t) \cdots \sigma_{i_p \alpha}(t) \right] \\ & \times \exp \left[ iN\lambda \mathcal{E}_{\text{aux}} - \frac{i\lambda}{Jq[\sigma]^{p/2}} \sum_{(i_1 \cdots i_p)} J_{i_1 \cdots i_p} \left( \frac{1}{T} \int_0^T dt \sigma_{i_1 u}(t) \right) \cdots \left( \frac{1}{T} \int_0^T dt \sigma_{i_p u}(t) \right) \right]. \end{aligned} \tag{3.65}$$

Proceeding as usual — averaging over disorder, introducing $G_{\alpha\alpha'}(t,t')$ and $F_{\alpha\alpha'}(t,t')$ as before, integrating out spins — we arrive at

$$\text{SFF}(T,f) = \int d\mathcal{E}_{\text{aux}} d\lambda \mathcal{D}G \mathcal{D}F \, f\big(\epsilon_u[\lambda,G]\big) f\big(\epsilon_l[\lambda,G]\big) e^{-NS_{\text{eff}}[\mathcal{E}_{\text{aux}},\lambda,G,F]}, \tag{3.66}$$

$$S_{\text{eff}}[\mathcal{E}_{\text{aux}},\lambda,G,F] = -i\lambda\mathcal{E}_{\text{aux}} - \frac{i}{2}\int_0^T dt\sum_\alpha (-1)^\alpha z_\alpha(t) + \frac{\lambda^2}{2p} + \frac{J\lambda}{pq[G]^{p/2}}\int_0^T dt\sum_\alpha(-1)^\alpha\left(\frac{1}{T}\int_0^T dt' G_{\alpha u}(t,t')\right)^p + \frac{1}{2}\int_0^T dtdt'\sum_{\alpha\alpha'}(-1)^{\alpha+\alpha'}\left(\frac{J^2}{p}G_{\alpha\alpha'}(t,t')^p - F_{\alpha\alpha'}(t,t')G_{\alpha\alpha'}(t,t')\right) + \frac{1}{2}\log\text{Det}\Big[i(-1)^\alpha\delta_{\alpha\alpha'}\big(\mu\partial_t^2 + z_\alpha\big) + (-1)^{\alpha+\alpha'}F_{\alpha\alpha'}\Big], \quad (3.67)$$

where we are denoting $q[G] \equiv T^{-2}\int_0^T dtdt' G_{uu}(t,t')$. The argument of the filter function is modified as well; it is now

$$\epsilon_\alpha[\lambda,G] \equiv -\frac{\mu}{2}\partial_t^2 G_{\alpha\alpha}(0^+,0) - \frac{iJ^2}{p}\int_0^T dt\sum_{\alpha'}(-1)^{\alpha'}G_{\alpha\alpha'}(t,0)^p - \frac{iJ\lambda}{pq[G]^{p/2}}\left(\frac{1}{T}\int_0^T dt G_{\alpha u}(t,0)\right)^p. \quad (3.68)$$

Note that $\mathcal{E}_{\text{aux}}$ enters linearly into the action. Thus if we were to integrate over $\mathcal{E}_{\text{aux}}$ at this point, we would obtain a $\delta$-function forcing $\lambda = 0$. The action would then reduce to the ergodic-phase expression, Eq. (3.45). While reassuring, this would not have accomplished anything, so we instead treat $\mathcal{E}_{\text{aux}}$ as a fixed parameter for now. We obtain saddle point equations by differentiating Eq. (3.67) only with respect to $\lambda$, $z$, $G$, and $F$.

The saddle point equations, assuming time translation invariance from the outset, are

$$i\big(\mu\partial_t^2 + z_\alpha\big)G_{\alpha\alpha'}(t-t') + \int_0^T dt''\sum_{\alpha''}(-1)^{\alpha''}F_{\alpha\alpha''}(t-t'')G_{\alpha''\alpha'}(t''-t') = (-1)^\alpha\delta_{\alpha\alpha'}\delta(t-t'), \quad (3.69)$$

$$F_{\alpha\alpha'}(t) = J^2 G_{\alpha\alpha'}(t)^{p-1} + \frac{J\lambda}{Tq[G]^{p/2}}\left(\delta_{\alpha u} + \delta_{\alpha' u}\right)\left(\frac{\widetilde{G}_{\alpha\alpha'}(0)}{T}\right)^{p-1} - \frac{J\lambda}{Tq[G]^{p/2+1}}\delta_{\alpha u}\delta_{\alpha' u}\sum_{\alpha''}(-1)^{\alpha''}\left(\frac{\widetilde{G}_{\alpha'' u}(0)}{T}\right)^{p}, \tag{3.70}$$

$$\mathcal{E}_{\text{aux}} = -\frac{iJT}{pq[G]^{p/2}}\sum_{\alpha}(-1)^{\alpha}\left(\frac{\widetilde{G}_{\alpha u}(0)}{T}\right)^{p} - \frac{i\lambda}{p}, \tag{3.71}$$

as well as the usual requirement $G_{\alpha\alpha}(0) = 1$. Here $\widetilde{G}_{\alpha\alpha'}(\omega)$ is the Fourier transform of $G_{\alpha\alpha'}(t)$.

### 3.4.2 Connected solutions

With $\mathcal{E}_{\text{aux}}$ fixed, let $\mathcal{G}_{\alpha\alpha'}(t)$ be the solution to the TAP equation of motion (Eq. (3.35)) corresponding to inverse temperature $\beta_{\text{aux}}$. Denote the Edwards-Anderson order parameter at $\mathcal{E}_{\text{aux}}$ and $\beta_{\text{aux}}$ by $q_{\text{EA}}$. Also recall the various auxiliary quantities we defined in Sec. 3.2.3: the self-energy $\mathcal{F}_{\alpha\alpha'}(t) \equiv J^2\mathcal{G}_{\alpha\alpha'}(t)^{p-1}$, the deviations $\Delta\mathcal{G}_{\alpha\alpha'}(t) \equiv \mathcal{G}_{\alpha\alpha'}(t) - q_{\text{EA}}$ and $\Delta\mathcal{F}_{\alpha\alpha'}(t) \equiv \mathcal{F}_{\alpha\alpha'}(t) - J^2 q_{\text{EA}}^{p-1}$, and the quantity $\Lambda \equiv \int_0^\infty dt \Delta\mathcal{G}^I(t)$. We have that $\mathcal{G}_{\alpha\alpha'}(t)$ and $\mathcal{F}_{\alpha\alpha'}(t)$ obey Eq. (3.35), which by taking $t$ and $t'$ to be far from the thermal branch can be written

$$i\left(\mu\partial_t^2 + z\right)\Delta\mathcal{G}_{\alpha\alpha'}(t-t') + \int_{-\infty}^{\infty} dt'' \sum_{\alpha''}(-1)^{\alpha''}\Delta\mathcal{F}_{\alpha\alpha''}(t-t'')\Delta\mathcal{G}_{\alpha''\alpha'}(t''-t') = (-1)^{\alpha}\delta_{\alpha\alpha'}\delta(t-t'). \tag{3.72}$$

We also have, as a result of Eq. (3.39), the relationship

$$\mathcal{E}_{\text{aux}} = \frac{2(p-1)Jq_{\text{EA}}^{p/2-1}\Lambda}{p} + \frac{1}{2pJq_{\text{EA}}^{p/2-1}\Lambda}. \tag{3.73}$$

Finally, recall the expression for the complexity $\Sigma(\mathcal{E})$, the logarithm of the number of solutions to the TAP magnetization equations at $\mathcal{E}$:

$$\Sigma(\mathcal{E}) = \frac{1}{2}\left(1 + 2\log\frac{p}{2}\right) - \frac{p\mathcal{E}^2}{2} + \frac{p^2}{8(p-1)}\left(\mathcal{E} + \sqrt{\mathcal{E}^2 - \mathcal{E}_{\text{th}}^2}\right)^2 + \log\left(-\mathcal{E} + \sqrt{\mathcal{E}^2 - \mathcal{E}_{\text{th}}^2}\right), \tag{3.74}$$

where $\mathcal{E}_{\text{th}}^2 = 4(p-1)/p^2$. Using Eq. (3.73), we can express $\Sigma(\mathcal{E}_{\text{aux}})$ in terms of $\Lambda$ and $q_{\text{EA}}$:

$$\Sigma(\mathcal{E}_{\text{aux}}) = -\frac{p-2}{2p} - \frac{1}{8pJ^2 q_{\text{EA}}^{p-2}\Lambda^2} + \frac{2(p-1)J^2 q_{\text{EA}}^{p-2}\Lambda^2}{p} - \frac{1}{2}\log 4J^2 q_{\text{EA}}^{p-2}\Lambda^2. \tag{3.75}$$

Our solution to the SFF saddle point equations, Eqs. (3.69) through (3.71), is best written in the frequency domain (tildes denote Fourier transforms):

$$\widetilde{G}_{\alpha\alpha'}(\omega) = Tq_{\text{EA}}\delta_{\omega 0} + \Delta\widetilde{\mathcal{G}}_{\alpha\alpha'}(\omega) + \frac{\widetilde{g}_{\alpha\alpha'}(\omega)}{T}, \tag{3.76}$$

$$\widetilde{F}_{\alpha\alpha'}(\omega) = \left(TJ^2 q_{\text{EA}}^{p-1} + ipJq_{\text{EA}}^{p/2-1}\left(\mathcal{E}_{\text{aux}} - 2Jq_{\text{EA}}^{p/2-1}\Lambda\right)\left(\delta_{\alpha u} + \delta_{\alpha' u}\right)\right)\delta_{\omega 0} + \Delta\widetilde{\mathcal{F}}_{\alpha\alpha'}(\omega) + \frac{\widetilde{f}_{\alpha\alpha'}(\omega)}{T}, \tag{3.77}$$

$$\lambda = ip\left(\mathcal{E}_{\text{aux}} - 2Jq_{\text{EA}}^{p/2-1}\Lambda\right) + \frac{\delta}{T}. \tag{3.78}$$

We again take $z_\alpha$ to be the equilibrium value corresponding to $\beta_{\text{aux}}$. The precise form of the correction terms $\widetilde{g}_{\alpha\alpha'}(\omega)$, $\widetilde{f}_{\alpha\alpha'}(\omega)$, and $\delta$ is largely unimportant — the essential feature is simply that they are $O(1)$ and the corrections are thus $O(T^{-1})$. Note that in the time domain, this solution amounts to

$$G_{\alpha\alpha'}(t) = q_{\text{EA}} + \sum_{n=-\infty}^{\infty}\Delta\mathcal{G}_{\alpha\alpha'}(t+nT) + \frac{g_{\alpha\alpha'}(t)}{T}, \tag{3.79}$$

$$F_{\alpha\alpha'}(t) = J^2 q_{\text{EA}}^{p-1} + \sum_{n=-\infty}^{\infty} \Delta\mathcal{F}_{\alpha\alpha'}(t+nT) + \frac{ipJq_{\text{EA}}^{p/2-1}\left(\mathcal{E}_{\text{aux}} - 2Jq_{\text{EA}}^{p/2-1}\Lambda\right)}{T}\left(\delta_{\alpha u} + \delta_{\alpha' u}\right) + \frac{f_{\alpha\alpha'}(t)}{T}. \tag{3.80}$$

The sums are convergent because $\Delta\mathcal{G}_{\alpha\alpha'}(t)$ and $\Delta\mathcal{F}_{\alpha\alpha'}(t)$ decay rapidly to zero as $|t| \to \infty$.

Although we have omitted it for notational simplicity, we can add a term $\delta_{\alpha\neq\alpha'}\Delta$ to the time arguments of $\Delta\mathcal{G}_{\alpha\alpha'}(t+nT)$ and $\Delta\mathcal{F}_{\alpha\alpha'}(t+nT)$ for any $\Delta \in [0, T)$, exactly as in Sec. 3.3. Due to the separate time translation symmetry on each branch of the SFF contour, all such solutions are equally valid and contribute the same action. Thus we shall demonstrate the validity of Eqs. (3.76) through (3.78) and evaluate the action only for $\Delta = 0$, but then integrate over all $\Delta \in [0, T)$ in the final expression for the SFF.

Let us first confirm that our solution satisfies the saddle point equation for $\lambda$, Eq. (3.71). Referring to Eq. (3.37), we have that $\Delta\widetilde{\mathcal{G}}_{\alpha\alpha'}(0) = L + (-1)^{\alpha} 2i\Lambda\delta_{\alpha\alpha'}$. Thus

$$q[G] = q_{\text{EA}} + \frac{L + 2i\Lambda}{T} + O(T^{-2}), \tag{3.81}$$

and Eq. (3.71) becomes

$$\mathcal{E}_{\text{aux}} = 2Jq_{\text{EA}}^{p/2-1}\Lambda - \frac{i\lambda}{p} + O(T^{-1}). \tag{3.82}$$

Solving for $\lambda$ indeed gives Eq. (3.78). The $O(T^{-1})$ terms determine $\delta$ as a function of the other quantities.

Now turn to Eq. (3.70). In the frequency domain, the right-hand side evaluates to[2]

$$\begin{aligned}
&\int_0^T dt e^{i\omega t} J^2 \left( q_{\text{EA}} + \sum_{n=-\infty}^{\infty} \Delta \mathcal{G}_{\alpha\alpha'}(t+nT) \right)^{p-1} \\
&\qquad + ipJq_{\text{EA}}^{p/2-1} \big( \mathcal{E}_{\text{aux}} - 2Jq_{\text{EA}}^{p/2-1}\Lambda \big) \big( \delta_{\alpha u} + \delta_{\alpha' u} \big) \delta_{\omega 0} + O(T^{-1}) \\
&\overset{\text{set}}{=} \left( TJ^2 q_{\text{EA}}^{p-1} + ipJq_{\text{EA}}^{p/2-1} \big( \mathcal{E}_{\text{aux}} - 2Jq_{\text{EA}}^{p/2-1}\Lambda \big) \big( \delta_{\alpha u} + \delta_{\alpha' u} \big) \right) \delta_{\omega 0} + \Delta \widetilde{\mathcal{F}}_{\alpha\alpha'}(\omega) + \frac{\widetilde{f}_{\alpha\alpha'}(\omega)}{T}.
\end{aligned} \tag{3.83}$$

Along the lines of Eq. (3.52), we have that

$$\begin{aligned}
&J^2 \left( q_{\text{EA}} + \sum_{n=-\infty}^{\infty} \Delta \mathcal{G}_{\alpha\alpha'}(t+nT) \right)^{p-1} \\
&\qquad = J^2 q_{\text{EA}}^{p-1} + J^2 \sum_{r=1}^{p-1} \binom{p-1}{r} q_{\text{EA}}^{p-1-r} \left( \sum_{n=-\infty}^{\infty} \Delta \mathcal{G}_{\alpha\alpha'}(t+nT) \right)^r \\
&\qquad \sim J^2 q_{\text{EA}}^{p-1} + J^2 \sum_{r=1}^{p-1} \binom{p-1}{r} q_{\text{EA}}^{p-1-r} \sum_{n=-\infty}^{\infty} \Delta \mathcal{G}_{\alpha\alpha'}(t+nT)^r \\
&\qquad = J^2 q_{\text{EA}}^{p-1} + \sum_{n=-\infty}^{\infty} \Delta \mathcal{F}_{\alpha\alpha'}(t+nT).
\end{aligned} \tag{3.84}$$

Thus, up to $O(1)$, both sides of Eq. (3.83) agree.

Finally, we confirm that Eq. (3.69) is satisfied. At non-zero frequencies, we have

$$i\big( -\mu\omega^2 + z \big) \Delta \widetilde{\mathcal{G}}_{\alpha\alpha'}(\omega) + \sum_{\alpha''} (-1)^{\alpha''} \Delta \widetilde{\mathcal{F}}_{\alpha\alpha''}(\omega) \Delta \widetilde{\mathcal{G}}_{\alpha''\alpha'}(\omega) + O(T^{-1}) \overset{\text{set}}{=} (-1)^{\alpha} \delta_{\alpha\alpha'}, \tag{3.85}$$

[2] Since $\widetilde{g}_{\alpha\alpha'}(\omega) = O(1)$ with respect to $T$, $g_{\alpha\alpha'}(t)$ decays to zero as $|t| \to \infty$ (at least to leading order).

which agrees at $O(1)$ due to Eq. (3.72). At zero frequency, we instead have

$$\sum_{\alpha''}(-1)^{\alpha''}\left(TJ^2q_{\text{EA}}^{p-1}+ipJq_{\text{EA}}^{p/2-1}\left(\mathcal{E}_{\text{aux}}-2Jq_{\text{EA}}^{p/2-1}\Lambda\right)\left(\delta_{\alpha u}+\delta_{\alpha'' u}\right)+\Delta\widetilde{\mathcal{F}}_{\alpha\alpha''}(0)+\frac{\widetilde{f}_{\alpha\alpha''}(0)}{T}\right)$$
$$+iz\left(Tq_{\text{EA}}+\Delta\widetilde{\mathcal{G}}_{\alpha\alpha'}(0)+\frac{\widetilde{g}_{\alpha\alpha'}(0)}{T}\right)\left(Tq_{\text{EA}}+\Delta\widetilde{\mathcal{G}}_{\alpha''\alpha'}(0)+\frac{\widetilde{g}_{\alpha''\alpha'}(0)}{T}\right)\overset{\text{set}}{=}(-1)^{\alpha}\delta_{\alpha\alpha'}. \tag{3.86}$$

The $O(T)$ terms come out to be

$$T\left(izq_{\text{EA}}+\sum_{\alpha''}(-1)^{\alpha''}\left(J^2q_{\text{EA}}^{p-1}\Delta\widetilde{\mathcal{G}}_{\alpha''\alpha'}(0)+q_{\text{EA}}\Delta\widetilde{\mathcal{F}}_{\alpha\alpha''}(0)\right)+ipJq_{\text{EA}}^{p/2}\left(\mathcal{E}_{\text{aux}}-2Jq_{\text{EA}}^{p/2-1}\Lambda\right)\right)$$
$$\overset{\text{set}}{=}0. \tag{3.87}$$

Yet from the TAP magnetization equations, Eq. (3.36), it follows that

$$\sum_{\alpha''}(-1)^{\alpha''}\left(q_{\text{EA}}\Delta\widetilde{\mathcal{F}}_{\alpha\alpha''}(0)-(p-1)J^2q_{\text{EA}}^{p-1}\Delta\widetilde{\mathcal{G}}_{\alpha\alpha''}(0)\right)=-izq_{\text{EA}}-ipJq_{\text{EA}}^{p/2}\mathcal{E}_{\text{aux}}. \tag{3.88}$$

Thus Eq. (3.87) evaluates to

$$-ipJq_{\text{EA}}^{p/2}\mathcal{E}_{\text{aux}}+pJ^2q_{\text{EA}}^{p-1}\sum_{\alpha''}(-1)^{\alpha''}\Delta\widetilde{\mathcal{G}}_{\alpha\alpha''}(0)+ipJq_{\text{EA}}^{p/2}\mathcal{E}_{\text{aux}}-2ipJ^2q_{\text{EA}}^{p-1}\Lambda=0, \tag{3.89}$$

using Eq. (3.37). The $O(1)$ terms of Eq. (3.86) determine $\widetilde{g}_{\alpha\alpha'}(0)$ and $\widetilde{f}_{\alpha\alpha'}(0)$. We have therefore confirmed that all saddle point equations are solved by Eqs. (3.76) through (3.78).

### 3.4.3 Contribution of connected solutions

It remains only to evaluate the action, Eq. (3.67), at the above solution. The action can be written as

$$\begin{aligned} S_{\text{eff}} = -i\lambda\mathcal{E}_{\text{aux}} + \frac{\lambda^2}{2p} + \frac{JT\lambda}{pq[G]^{p/2}}\sum_{\alpha}(-1)^{\alpha}\left(\frac{\widetilde{G}_{\alpha u}(0)}{T}\right)^{p} \\ + \frac{T}{2}\sum_{\alpha\alpha'}(-1)^{\alpha+\alpha'}\int_0^T dt\left(\frac{J^2}{p}G_{\alpha\alpha'}(t)^p - F_{\alpha\alpha'}(t)G_{\alpha\alpha'}(t)\right) \\ + \frac{1}{2}\sum_{\omega}\log\text{Det}\Big[i(-1)^{\alpha}\delta_{\alpha\alpha'}\big(-\mu\omega^2+z\big) + (-1)^{\alpha+\alpha'}\widetilde{F}_{\alpha\alpha'}(\omega)\Big]. \end{aligned} \tag{3.90}$$

Interestingly, we can determine $S_{\text{eff}}$ up to a single additive constant simply by noting that $dS_{\text{eff}}/d\mathcal{E}_{\text{aux}} = -i\lambda$ (recall that $S_{\text{eff}}$ is stationary with respect to variations in all quantities other than $\mathcal{E}_{\text{aux}}$). With $\lambda$ given by Eq. (3.78) and $q_{\text{EA}}^{p/2-1}\Lambda$ given by Eq. (3.41), we can carry out the integral to obtain that

$$S_{\text{eff}}[\mathcal{E}_{\text{aux}}] = \frac{p\mathcal{E}_{\text{aux}}^2}{2} - \frac{p^2}{8(p-1)}\left(\mathcal{E}_{\text{aux}} + \sqrt{\mathcal{E}_{\text{aux}}^2 - \mathcal{E}_{\text{th}}^2}\right)^2 - \log\left(-\mathcal{E}_{\text{aux}} + \sqrt{\mathcal{E}_{\text{aux}}^2 - \mathcal{E}_{\text{th}}^2}\right) + C, \tag{3.91}$$

for some unknown constant $C$. Comparing to Eq. (3.42), this is highly suggestive that $S_{\text{eff}} = -\Sigma(\mathcal{E}_{\text{aux}})$. Of course, we do need to determine the remaining constant, and so we now turn to a more elaborate calculation.

Rather than substitute Eqs. (3.76) and (3.77) into Eq. (3.90), we instead use the simpler

functions

$$\widetilde{G}'_{\alpha\alpha'}(\omega) = Tq_{\text{EA}}\delta_{\omega 0} + \Delta\widetilde{\mathcal{G}}_{\alpha\alpha'}(\omega), \tag{3.92}$$

$$\begin{aligned}\widetilde{F}'_{\alpha\alpha'}(\omega) = \Big(TJ^2q_{\text{EA}}^{p-1} + ipJq_{\text{EA}}^{p/2-1}\big(\mathcal{E}_{\text{aux}} - 2Jq_{\text{EA}}^{p/2-1}\Lambda\big)\big(\delta_{\alpha u} + \delta_{\alpha' u}\big)\Big)\,\delta_{\omega 0} \\ + \Delta\widetilde{\mathcal{F}}_{\alpha\alpha'}(\omega) + \frac{\widetilde{f}_{\alpha\alpha'}(0)}{T}\delta_{\omega 0},\end{aligned} \tag{3.93}$$

and show that the error incurred in doing so vanishes at large $T$.

Let us demonstrate that the error is negligible first. At any non-zero frequency, we have that

$$\widetilde{G}_{\alpha\alpha'}(\omega) = \widetilde{G}'_{\alpha\alpha'}(\omega) + \frac{\widetilde{g}_{\alpha\alpha'}(\omega)}{T}, \qquad \widetilde{F}_{\alpha\alpha'}(\omega) = \widetilde{F}'_{\alpha\alpha'}(\omega) + \frac{\widetilde{f}_{\alpha\alpha'}(\omega)}{T}. \tag{3.94}$$

The partial derivatives of $S_{\text{eff}}$ at nonzero $\omega$ are

$$\frac{\partial S_{\text{eff}}}{\partial \widetilde{G}_{\alpha\alpha'}(\omega)} = \frac{1}{2}(-1)^{\alpha+\alpha'}\int_0^T dt e^{-i\omega t}\Big(J^2G_{\alpha\alpha'}(t)^{p-1} - F_{\alpha\alpha'}(t)\Big), \tag{3.95}$$

$$\frac{\partial S_{\text{eff}}}{\partial \widetilde{F}_{\alpha\alpha'}(\omega)} = \frac{1}{2}(-1)^{\alpha+\alpha'}\left(\Big[i(-1)^{\alpha}\delta_{\alpha\alpha'}\big(-\mu\omega^2+z\big)+(-1)^{\alpha+\alpha'}\widetilde{F}_{\alpha'\alpha}(\omega)\Big]^{-1}_{\alpha\alpha'} - \int_0^T dt e^{-i\omega t}G_{\alpha\alpha'}(t)\right), \tag{3.96}$$

which vanish when evaluated at $\widetilde{G}'_{\alpha\alpha'}(\omega) = \Delta\widetilde{\mathcal{G}}_{\alpha\alpha'}(\omega)$ and $\widetilde{F}'_{\alpha\alpha'}(\omega) = \Delta\widetilde{\mathcal{F}}_{\alpha\alpha'}(\omega)$. Thus the $O(T^{-1})$ difference between $\widetilde{G}_{\alpha\alpha'}(\omega)$ and $\widetilde{G}'_{\alpha\alpha'}(\omega)$, as with $\widetilde{F}_{\alpha\alpha'}(\omega)$ and $\widetilde{F}'_{\alpha\alpha'}(\omega)$, translates only to an $O(T^{-2})$ difference in the action. Even after summing over all $\omega \neq 0$, the total error[3] is only $O(T^{-1})$.

Neglecting non-zero frequencies, $\widetilde{F}'_{\alpha\alpha'}(\omega)$ is identical to $\widetilde{F}_{\alpha\alpha'}(\omega)$ and $\widetilde{G}'_{\alpha\alpha'}(\omega)$ differs only

[3]Since the $G(t)^p$ term is not diagonal in the frequency domain, this argument requires a bit more care. One can easily show that $\partial^2 S_{\text{eff}}/\partial\widetilde{G}(\omega)\partial\widetilde{G}(\omega')$ is $O(T^{-1})$ for $\omega \neq \pm\omega'$ and $O(1)$ for $\omega = \pm\omega'$. Summing over all frequencies, the former case gives a total contribution $O(T^{-3})O(T^2) = O(T^{-1})$ and the latter gives $O(T^{-2})O(T) = O(T^{-1})$. The total error is thus $O(T^{-1})$ as claimed.

by $\widetilde{g}_{\alpha\alpha'}(0)\delta_{\omega 0}/T$. In the time domain, the latter corresponds to

$$G_{\alpha\alpha'}(t) = G'_{\alpha\alpha'}(t) + \frac{\widetilde{g}_{\alpha\alpha'}(0)}{T^2} = q_{\rm EA} + \sum_{n=-\infty}^{\infty} \Delta\mathcal{G}_{\alpha\alpha'}(t+nT) + \frac{\widetilde{g}_{\alpha\alpha'}(0)}{T^2}. \tag{3.97}$$

Yet

$$\frac{\partial S_{\rm eff}}{\partial G_{\alpha\alpha'}(t)} = \frac{T}{2}(-1)^{\alpha+\alpha'}\Big(J^2 G_{\alpha\alpha'}(t)^{p-1} - F_{\alpha\alpha'}(t)\Big) + O(1). \tag{3.98}$$

When evaluated at $G'_{\alpha\alpha'}(t)$ and $F'_{\alpha\alpha'}(t)$, the $O(T)$ contribution vanishes (see Eq. (3.84)). Thus $\partial S_{\rm eff}/\partial G_{\alpha\alpha'}(t)$ is $O(1)$, and an $O(T^{-2})$ change to $G_{\alpha\alpha'}(t)$ leads only to an $O(T^{-1})$ change in the action even after integrating over $t$.

Since all errors are $O(T^{-1})$, we can safely evaluate $S_{\rm eff}$ at Eqs. (3.92) and (3.93) rather than the full solution (we still use Eq. (3.78) for $\lambda$). The first line of Eq. (3.90) can be computed straightforwardly. It comes out to be

$$\frac{p\big(\mathcal{E}_{\rm aux} - 2Jq_{\rm EA}^{p/2-1}\Lambda\big)^2}{2} = \frac{1}{8pJ^2q_{\rm EA}^{p-2}\Lambda^2} - \frac{1}{p} + \frac{2J^2q_{\rm EA}^{p-2}\Lambda^2}{p}, \tag{3.99}$$

where we used Eq. (3.73) to obtain the right-hand side.

Next consider the bottom line. Since $\widetilde{F}'_{\alpha\alpha'}(\omega) = \Delta\widetilde{\mathcal{F}}_{\alpha\alpha'}(\omega)$ for $\omega \neq 0$, while $\widetilde{F}'_{\alpha\alpha'}(0) = \widetilde{F}_{\alpha\alpha'}(0)$, we can write the determinant term as

$$\begin{aligned}
&\frac{1}{2}\sum_{\omega}\log\mathrm{Det}\Big[i(-1)^{\alpha}\delta_{\alpha\alpha'}\big(-\mu\omega^2+z\big) + (-1)^{\alpha+\alpha'}\Delta\widetilde{\mathcal{F}}_{\alpha\alpha'}(\omega)\Big] \\
&\qquad + \frac{1}{2}\log\mathrm{Det}\Big[iz(-1)^{\alpha}\delta_{\alpha\alpha'} + (-1)^{\alpha+\alpha'}\widetilde{F}_{\alpha\alpha'}(0)\Big] \\
&\qquad\qquad - \frac{1}{2}\log\mathrm{Det}\Big[iz(-1)^{\alpha}\delta_{\alpha\alpha'} + (-1)^{\alpha+\alpha'}\Delta\widetilde{\mathcal{F}}_{\alpha\alpha'}(0)\Big].
\end{aligned} \tag{3.100}$$

The top line vanishes by exactly the same reasoning as in Sec. 3.3.3: it is proportional to $T$ by the Euler-Maclaurin formula, and then must be zero since the derivative with respect to $T$ vanishes. Given Eq. (3.37), the bottom line is simply

$$\frac{1}{2}\log \text{Det}\Delta\widetilde{\mathcal{G}}_{\alpha\alpha'}(0) = \frac{1}{2}\log 4\Lambda^2. \tag{3.101}$$

For the middle line we take an indirect approach. We have that $iz(-1)^\alpha\delta_{\alpha\alpha'} + (-1)^{\alpha+\alpha'}\widetilde{F}_{\alpha\alpha'}(0)$ is the matrix inverse to $\widetilde{G}_{\alpha\alpha'}(0)$ (using the full solution for the latter, Eq. (3.76)). Written out,

$$\begin{pmatrix} iz+\widetilde{F}_{uu}(0) & -\widetilde{F}_{ul}(0) \\ -\widetilde{F}_{lu}(0) & -iz+\widetilde{F}_{ll}(0) \end{pmatrix} = \begin{pmatrix} \widetilde{G}_{uu}(0) & \widetilde{G}_{ul}(0) \\ \widetilde{G}_{lu}(0) & \widetilde{G}_{ll}(0) \end{pmatrix}^{-1} = \frac{1}{\text{Det}\widetilde{G}(0)} \begin{pmatrix} \widetilde{G}_{ll}(0) & -\widetilde{G}_{ul}(0) \\ -\widetilde{G}_{lu}(0) & \widetilde{G}_{uu}(0) \end{pmatrix}. \tag{3.102}$$

Rather than this $(u,l)$ basis, express Eq. (3.102) in the $(u+l, u-l)$ basis (called "classical"/"quantum" in the Keldysh literature), denoted $(+,-)$:

$$\begin{pmatrix} \widetilde{F}_{--}(0) & iz+\widetilde{F}_{-+}(0) \\ iz+\widetilde{F}_{+-}(0) & \widetilde{F}_{++}(0) \end{pmatrix} = \frac{1}{\text{Det}\widetilde{G}(0)} \begin{pmatrix} \widetilde{G}_{--}(0) & -\widetilde{G}_{+-}(0) \\ -\widetilde{G}_{-+}(0) & \widetilde{G}_{++}(0) \end{pmatrix}. \tag{3.103}$$

We can read off that $\text{Det}\widetilde{G}(0)^{-1} = \widetilde{F}_{++}(0)/\widetilde{G}_{++}(0)$. Note that we only need $\widetilde{G}_{++}(0)$ and $\widetilde{F}_{++}(0)$ to $O(T)$ in order to calculate the determinant to $O(1)$. Thus the middle line of Eq. (3.100) evaluates to $(\log J^2 q_{\text{EA}}^{p-2})/2$, and the total contribution of the determinant term is

$$\frac{1}{2}\log 4J^2 q_{\text{EA}}^{p-2}\Lambda^2. \tag{3.104}$$

Lastly consider the middle line of Eq. (3.90). Since $G'_{\alpha\alpha'}(t) = \mathcal{G}_{\alpha\alpha'}(t)$ (up to exponentially

small corrections), $\sum_{\alpha\alpha'}(-1)^{\alpha+\alpha'}G'_{\alpha\alpha'}(t)^p = 0$ by virtue of Eq. (3.30). We are left with

$$
\begin{aligned}
&-\frac{T}{2}\sum_{\alpha\alpha'}(-1)^{\alpha+\alpha'}\int_0^T dt F'_{\alpha\alpha'}(t)G'_{\alpha\alpha'}(t) \\
&\quad\sim -\frac{T}{2}\sum_{\alpha\alpha'}(-1)^{\alpha+\alpha'}\int_0^T dt \mathcal{F}_{\alpha\alpha'}(t)\mathcal{G}_{\alpha\alpha'}(t) \\
&\qquad -\frac{1}{2}\sum_{\alpha\alpha'}(-1)^{\alpha+\alpha'}\left(ipJq_{\text{EA}}^{p/2-1}\big(\mathcal{E}_{\text{aux}}-2Jq_{\text{EA}}^{p/2-1}\Lambda\big)\big(\delta_{\alpha u}+\delta_{\alpha' u}\big)+\frac{\widetilde{f}_{\alpha\alpha'}(0)}{T}\right)\widetilde{G}'_{\alpha\alpha'}(0).
\end{aligned}
\tag{3.105}
$$

The first term is again proportional to $\sum_{\alpha\alpha'}(-1)^{\alpha+\alpha'}G'_{\alpha\alpha'}(t)^p = 0$. The second term would appear to be more problematic, since $\widetilde{f}_{\alpha\alpha'}(0)$ (for which we have not given an explicit expression) contributes at $O(1)$ due to $\widetilde{G}'_{\alpha\alpha'}(0)$ being $O(T)$. However, we only need the component $\widetilde{f}_{--}(0)/T = \widetilde{F}_{--}(0)$, and from Eq. (3.103) we see that

$$
\widetilde{F}_{--}(0) = \frac{1}{\text{Det}\widetilde{G}(0)}\widetilde{G}_{--}(0) \tag{3.106}
$$

$$
= \frac{1}{\text{Det}\widetilde{G}(0)}\frac{\text{Det}\widetilde{G}(0)+\widetilde{G}_{+-}(0)\widetilde{G}_{-+}(0)}{\widetilde{G}_{++}(0)} \tag{3.107}
$$

$$
= \frac{1-4J^2q_{\text{EA}}^{p-2}\Lambda^2}{2Tq_{\text{EA}}}+O\left(\frac{1}{T^2}\right). \tag{3.108}
$$

Eq. (3.105) evaluates to

$$
2pJq_{\text{EA}}^{p/2-1}\Lambda\big(\mathcal{E}_{\text{aux}}-2Jq_{\text{EA}}^{p/2-1}\Lambda\big)-\frac{1}{2}+2J^2q_{\text{EA}}^{p-2}\Lambda^2 = \frac{1}{2}-2J^2q_{\text{EA}}^{p-2}\Lambda^2, \tag{3.109}
$$

again using Eq. (3.73).

We finally have the large-$T$ limit of the action, given by the sum of Eqs. (3.99), (3.104),

and (3.109):

$$S_{\text{eff}}[\mathcal{E}_{\text{aux}}] = \frac{p-2}{2p} + \frac{1}{8pJ^2 q_{\text{EA}}^{p-2}\Lambda^2} - \frac{2(p-1)J^2 q_{\text{EA}}^{p-2}\Lambda^2}{p} + \frac{1}{2}\log 4J^2 q_{\text{EA}}^{p-2}\Lambda^2. \tag{3.110}$$

Comparing to the complexity $\Sigma(\mathcal{E})$ given in Eq. (3.75), we see that $S_{\text{eff}}$ is precisely $-\Sigma(\mathcal{E}_{\text{aux}})$.

### 3.4.4 Evaluation of the SFF

We have shown that, at a given $\mathcal{E}_{\text{aux}}$ and for each value of inverse temperature $\beta_{\text{aux}}$, there is a solution to the SFF saddle point equations with $S_{\text{eff}} = -\Sigma(\mathcal{E}_{\text{aux}})$. The full (connected) SFF is obtained by integrating over all $\mathcal{E}_{\text{aux}}$ and $\beta_{\text{aux}}$, as well as the symmetry-broken order parameter $\Delta$ (which contributes an overall factor of $T$) and an additional factor $2(1+\delta_{p\text{ even}})$ from the discrete symmetries. As in Sec. 3.3, it is more convenient to integrate over the energy density $\epsilon(\mathcal{E}_{\text{aux}}, \beta_{\text{aux}})$. We show in App. B.2 that $\epsilon(\mathcal{E}, \beta)$ comes out to be precisely the argument of the filter function, Eq. (3.68), when evaluated at the saddle point solution.

Our final result is that[4]

$$\text{SFF}(T, f) = \left|\mathbb{E}\text{Tr} f(H) e^{-iHT}\right|^2 + 2\big(1+\delta_{p\text{ even}}\big) T \sqrt{\frac{pN}{2\pi}} \int d\mathcal{E}_{\text{aux}} e^{N\Sigma(\mathcal{E}_{\text{aux}})} \int_{\epsilon_-(\mathcal{E}_{\text{aux}})}^{\epsilon_+(\mathcal{E}_{\text{aux}})} \frac{d\epsilon_{\text{aux}}}{2\pi} f(\epsilon_{\text{aux}})^2, \tag{3.111}$$

where the inner integral runs only over the range $[\epsilon_-(\mathcal{E}_{\text{aux}}), \epsilon_+(\mathcal{E}_{\text{aux}})]$ in which solutions to the TAP equations exist. Furthermore, one can easily generalize Eq. (3.111) by making the filter function $\mathcal{E}$-dependent, i.e., $f(\mathcal{E}, \epsilon_{\text{aux}})$. The resulting quantity is the SFF for the projection of the system into certain TAP states.

[4] The factor $\sqrt{pN/2\pi}$ comes from the integral over fluctuations in $\lambda$ — the variance is $p/N$ (see Eq. (3.90)), and the original fat unity introducing $\mathcal{E}_{\text{aux}}$ comes with a prefactor $N/2\pi$.

Compare Eq. (3.111) for the non-ergodic phase to Eq. (3.62) for the ergodic phase, and recall the discussion of block-diagonal Hamiltonians in Sec. 3.1.1. Our result demonstrates that each metastable (i.e., TAP) state can be thought of its own quantum chaotic subspace, one which is independent of any others. This is the central result of our paper. While the qualitative idea has been proposed in previous work [169], the present analysis both makes it precise and proves it.

## 3.5 Higher moments of the evolution operator

In this final section we consider higher moments of $\mathrm{Tr}e^{-iHT}$, i.e., the quantities

$$\mathrm{SFF}^{(n)}(T,f) \equiv \mathbb{E}\Big[\Big(\mathrm{Tr}f(H)e^{-iHT}\Big)^n\Big(\mathrm{Tr}f(H)e^{iHT}\Big)^n\Big]. \tag{3.112}$$

The saddle points of these higher moments exhibit an interesting structure that will shed further light on the distribution of TAP states, although care must be taken in interpreting the results. We first present the calculation and discuss afterwards.

### 3.5.1 Effective action

The effective action governing the $n$'th moment is derived in exactly the same manner as in Sec. 3.4. The only major difference is that now spins have a "replica" index $a \in \{1,\cdots,n\}$ in addition to a contour index $\alpha \in \{u,l\}$. We also include a separate fat unity defining $\mathcal{E}_{\mathrm{aux},a}$ for each replica. The result is (compare to Eqs. (3.66) and (3.67))

$$\mathrm{SFF}^{(n)}(T,f) = \int d\mathcal{E}_{\mathrm{aux}} d\lambda \mathcal{D}G\mathcal{D}F \prod_{a=1}^{n} f\big(\epsilon_{au}[\lambda,G]\big) f\big(\epsilon_{al}[\lambda,G]\big) e^{-NS_{\mathrm{eff}}[\mathcal{E}_{\mathrm{aux}},\lambda,G,F]}, \tag{3.113}$$

$$
\begin{aligned}
S_{\text{eff}}[\mathcal{E}_{\text{aux}},\lambda,G,F] = &-i\sum_a \lambda_a \mathcal{E}_{\text{aux},a} - \frac{i}{2}\int_0^T dt \sum_{a\alpha}(-1)^\alpha z_{a\alpha}(t) \\
&+ \sum_{aa'} \frac{\lambda_a \lambda_{a'}}{2pq[G_{aa}]^{p/2} q[G_{a'a'}]^{p/2}} \left(\frac{1}{T^2}\int_0^T dt dt' G_{au,a'u}(t,t')\right)^p \\
&+ \sum_{a'} \frac{J\lambda_{a'}}{pq[G_{a'a'}]^{p/2}} \int_0^T dt \sum_{a\alpha}(-1)^\alpha \left(\frac{1}{T}\int_0^T dt' G_{a\alpha,a'u}(t,t')\right)^p \\
&+ \frac{1}{2}\int_0^T dt dt' \sum_{aa'}\sum_{\alpha\alpha'}(-1)^{\alpha+\alpha'} \left(\frac{J^2}{p} G_{a\alpha,a'\alpha'}(t,t')^p - F_{a\alpha,a'\alpha'}(t,t') G_{a\alpha,a'\alpha'}(t,t')\right) \\
&+ \frac{1}{2}\log \text{Det}\Big[ i(-1)^\alpha \delta_{aa'}\delta_{\alpha\alpha'}\big(\mu\partial_t^2 + z_{a\alpha}\big) + (-1)^{\alpha+\alpha'} F_{a\alpha,a'\alpha'}\Big], \quad (3.114)
\end{aligned}
$$

with energy densities

$$
\begin{aligned}
\epsilon_{a\alpha}[\lambda,G] = &-\frac{\mu}{2}\partial_t^2 G_{a\alpha,a\alpha}(0^+,0) \\
&-\frac{iJ^2}{p}\int_0^T dt \sum_{a'\alpha'}(-1)^{\alpha'} G_{a\alpha,a'\alpha'}(t,0)^p - i\sum_{a'} \frac{J\lambda_{a'}}{pq[G_{a'a'}]^{p/2}} \left(\frac{1}{T}\int_0^T dt G_{a\alpha,a'u}(t,0)\right)^p .
\end{aligned} \quad (3.115)
$$

The saddle point equations are therefore

$$
\begin{aligned}
&i\big(\mu\partial_t^2 + z_a\big) G_{a\alpha,a'\alpha'}(t-t') + \int_0^T dt'' \sum_{a''\alpha''}(-1)^{\alpha''} F_{a\alpha,a''\alpha''}(t-t'') G_{a''\alpha'',a'\alpha'}(t''-t') \\
&\qquad = (-1)^\alpha \delta_{aa'}\delta_{\alpha\alpha'}\delta(t-t'), \quad (3.116)
\end{aligned}
$$

$$
\begin{aligned}
F_{a\alpha,a'\alpha'}(t) = &\, J^2 G_{a\alpha,a'\alpha'}(t)^{p-1} \\
&+ \frac{J}{T}\left(\frac{\lambda_a}{q[G_{aa}]^{p/2}}\delta_{\alpha u} + \frac{\lambda_{a'}}{q[G_{a'a'}]^{p/2}}\delta_{\alpha' u}\right)\left(\frac{\widetilde{G}_{a\alpha,a'\alpha'}(0)}{T}\right)^{p-1} + O(T^{-2}), \quad (3.117)
\end{aligned}
$$

$$\mathcal{E}_{\text{aux},a} = -\frac{iJT}{pq[G_{aa}]^{p/2}} \sum_{a'\alpha'} (-1)^{\alpha'} \left( \frac{\widetilde{G}_{au,a'\alpha'}(0)}{T} \right)^p - \frac{i}{pq[G_{aa}]^{p/2}} \sum_{a'} \frac{\lambda_{a'}}{q[G_{a'a'}]^{p/2}} \left( \frac{\widetilde{G}_{au,a'u}(0)}{T} \right)^p . \tag{3.118}$$

Note that Eqs. (3.116) through (3.118) have the following permutation symmetry with respect to replica indices. Suppose that $G$, $F$, and $\lambda$ constitute a valid solution. For any permutation $\pi$ of the set $\{1, \cdots, n\}$, define $\pi_\alpha(a)$ to be the permuted element $\pi(a)$ if $\alpha = l$ but simply the original element $a$ if $\alpha = u$. Then the quantities $\overline{G}$, $\overline{F}$, and $\overline{\lambda}$ defined by

$$\overline{G}_{a\alpha,a'\alpha'}(t,t') \equiv G_{\pi_\alpha(a)\alpha,\pi_{\alpha'}(a')\alpha'}(t,t'), \qquad \overline{F}_{a\alpha,a'\alpha'}(t,t') \equiv F_{\pi_\alpha(a)\alpha,\pi_{\alpha'}(a')\alpha'}(t,t'), \qquad \overline{\lambda}_a = \lambda_a, \tag{3.119}$$

constitute an equally valid solution. This symmetry has a nice graphical interpretation in terms of pairings between upper and lower contours, illustrated in Fig. 3.5: however contour $au$ is correlated with $a'l$ in a given solution, there is an alternate solution in which $au$ has the same correlation with $\pi(a')l$.

One trivial solution to the saddle point equations is to use the solution from Sec. 3.4 for $a = a'$ while setting all cross-replica elements to zero. The action then decomposes into a sum of single-replica actions, which we evaluated in Sec. 3.4. In other words, this contribution to the $n$'th moment is simply $\text{SFF}(T,f)^n$. However, by the permutation symmetry described above, we actually have $n!$ such contributions:

$$\text{SFF}^{(n)}(T,f) = n! \cdot \text{SFF}(T,f)^n + \cdots , \tag{3.120}$$

where the ellipses denote additional solutions.

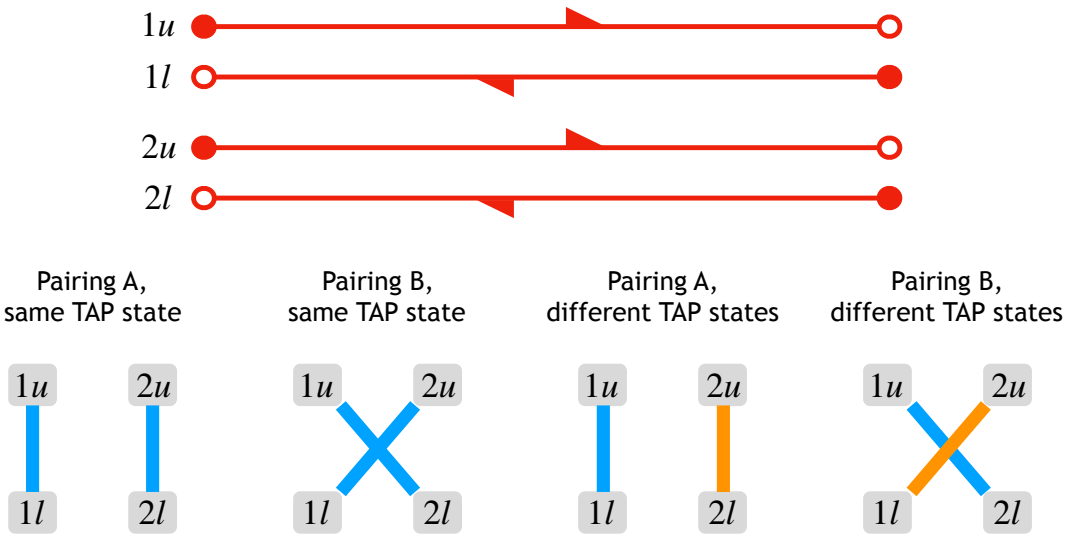



Figure 3.5: Graphical representation of the various saddle point solutions for the $n = 2$ moment. The four contours — $1u$, $1l$, $2u$, $2l$ — are shown at the top. Below are the four varieties of solutions: each upper contour must be paired with a lower contour, but one is free to choose which replicas are paired, and there is further freedom in which TAP state each pair lies within (blue and orange lines indicate two different TAP states).

### 3.5.2 Connected solutions

In general, for arbitrary values of $\mathcal{E}_{\text{aux},a}$, we have been unable to find any further saddle points. However, when some replicas have equal values of $\mathcal{E}_{\text{aux}}$, we can construct additional solutions. Pick any set of inverse temperatures $\beta_a$ (not necessarily equal), and suppose that the replicas $\{1, \cdots, n\}$ partition into groups $A \equiv \{a_1, \cdots, a_{|A|}\}$, such that $\mathcal{E}_{\text{aux},a}$ equals a common value $\mathcal{E}_{\text{aux},A}$ for all $a \in A$. We again take $G_{a\alpha,a\alpha'}(t - t')$ to be the solution from Sec. 3.4. For $a$ and $a'$ in different groups, we still set $G_{a\alpha,a'\alpha'} = 0$. For $a$ and $a'$ in the same group $A$, however, we now set

$$G_{a\alpha,a'\alpha'}(t - t') = \left(q_{\text{EA},a} q_{\text{EA},a'}\right)^{1/2}, \tag{3.121}$$

where $q_{\text{EA},a}$ is the Edwards-Anderson order parameter corresponding to $\mathcal{E}_{\text{aux},A}$ and $\beta_a$. This corresponds to the replicas lying within the same TAP state (see Fig. 3.5). We can write this compactly as

$$G_{a\alpha,a'\alpha'}(t) = \left(q_{\text{EA},a}q_{\text{EA},a'}\right)^{1/2} + \delta_{aa'}\Big(\Delta\mathcal{G}_{a,\alpha\alpha'}(t) + O(T^{-1})\Big). \tag{3.122}$$

Inserting into Eq. (3.118), we have that $\lambda_a$ must obey

$$\sum_{a'\in A}\lambda_{a'} = ip\big(\mathcal{E}_{\text{aux},A} - 2Jq_{\text{EA},a}^{p/2-1}\Lambda_a\big) + O(T^{-1}). \tag{3.123}$$

Note that, by virtue of Eq. (3.41), $Jq_{\text{EA},a}^{p/2-1}\Lambda_a$ is a function solely of $\mathcal{E}_{\text{aux},A}$. Thus Eq. (3.123) is consistent among all $a \in A$. The self-energy is then given by

$$\begin{aligned} F_{a\alpha,a'\alpha'}(t) = J^2\left(q_{\text{EA},a}q_{\text{EA},a'}\right)^{\frac{p-1}{2}} + \delta_{aa'}\Delta\mathcal{F}_{a,\alpha\alpha'}(t) \\ + \frac{J}{T}\left(\sqrt{\frac{q_{\text{EA},a'}^{p-1}}{q_{\text{EA},a}}}\lambda_a\delta_{\alpha u} + \sqrt{\frac{q_{\text{EA},a}^{p-1}}{q_{\text{EA},a'}}}\lambda_{a'}\delta_{\alpha' u}\right) + O(T^{-1}). \end{aligned} \tag{3.124}$$

It remains only to check that Eq. (3.116) can be satisfied. It is automatically solved at non-zero frequencies, since then $\widetilde{G}(\omega)$ and $\widetilde{F}(\omega)$ reduce to $\delta_{aa'}\Delta\widetilde{\mathcal{G}}(\omega)$ and $\delta_{aa'}\Delta\widetilde{\mathcal{F}}(\omega)$ respectively. At zero frequency we confirm that the equation is solved to $O(T)$ (the $O(1)$ terms only determine subleading corrections). Following the same steps as in Sec. 3.4.2, the left-hand side of Eq. (3.116) simplifies to

$$JT\sqrt{q_{\text{EA},a}^{p-1}q_{\text{EA},a'}}\left(\sum_{a''\in A}\lambda_{a''} + 2i(p-1)Jq_{\text{EA},a}^{p/2-1}\Lambda_a + 2iJq_{\text{EA},a'}^{p/2-1}\Lambda_{a'} - ip\mathcal{E}_{\text{aux},A}\right) = 0, \tag{3.125}$$

as desired.

Note that in this solution, only the sum $\sum_a \lambda_a$ is determined — all orthogonal components of the vector $\lambda$ are free to take any values. This does not imply that there are multiple such solutions, however. Returning to the effective action in Eq. (3.114), the fact that the saddle point equations determine only $G$, $F$, and $\sum_a \lambda_a$ means that, if we first integrate over them, the resulting $\lambda$-dependent action is of the form

$$S_{\text{eff}}[\mathcal{E}_{\text{aux}}, \lambda] = S\left[\mathcal{E}_{\text{aux}}, \sum_a \lambda_a\right] - i \sum_{aa'} \sum_{b=2}^{|A|} \lambda_a u_{ab} u_{a'b} \mathcal{E}_{\text{aux},a'}, \tag{3.126}$$

for some function $S$ of the single quantity $\sum_a \lambda_a$ (as well as all $\mathcal{E}_{\text{aux}}$) and for any choice of orthonormal basis vectors $u_{ab}$ orthogonal to the all-1 vector. When we integrate over $\sum_a \lambda_a u_{ab}$, we thus get a $\delta$-function forcing $\sum_{a'} u_{a'b} \mathcal{E}_{\text{aux},a'} = 0$. Together, the $\delta$-functions force all $\mathcal{E}_{\text{aux},a}$ to equal a common value $\mathcal{E}_{\text{aux},A}$. Not only is this consistent with our original assumption, it shows that our construction cannot work for any other values of $\mathcal{E}_{\text{aux},a}$.

### 3.5.3 Contribution of connected solutions

To evaluate the action, note first of all that since the numbers $\beta_a$ define a continuous family of solutions, and since the action is by definition stationary at these solutions, all choices of $\beta_a$ must give the same value of the action. We thus take all $\beta_a$ to equal a common value $\beta$ for simplicity. The action evaluated at this solution still decomposes into a sum over groups, but now

the contribution of a single group $A$ is

$$\begin{aligned} S_{\text{eff}}[\mathcal{E}_{\text{aux}}] = -i\mathcal{E}_{\text{aux}} \sum_{a\in A} \lambda_a + \frac{1}{2p}\left(\sum_{a\in A}\lambda_a\right)^2 + 2iJq_{\text{EA}}^{p/2-1}\Lambda \sum_{a\in A}\lambda_a \\ + \frac{T}{2}\sum_{aa'}\int_0^T dt \sum_{\alpha\alpha'}(-1)^{\alpha+\alpha'}\left(\frac{J^2}{p}G_{a\alpha,a\alpha'}(t)^p - F_{a\alpha,a\alpha'}(t)G_{a\alpha,a\alpha'}(t)\right) \\ + \frac{1}{2}\log \text{Det}\Big[i(-1)^\alpha \delta_{aa'}\delta_{\alpha\alpha'}\big(\mu\partial_t^2 + z\big) + (-1)^{\alpha+\alpha'}F_{a\alpha,a'\alpha'}\Big]. \end{aligned} \tag{3.127}$$

Note that now $\mathcal{E}_{\text{aux}}$, $q_{\text{EA}}$, and $\Lambda$ are all independent of the replica $a$ (within a given group $A$). We are also free to set all $\lambda_a = \lambda$, meaning that our saddle point solution simplifies to (in frequency space)

$$\widetilde{G}_{a\alpha,a'\alpha'}(\omega) = Tq_{\text{EA}}\delta_{\omega 0} + \delta_{aa'}\Big(\Delta\widetilde{\mathcal{G}}_{\alpha\alpha'}(\omega) + O(T^{-1})\Big), \tag{3.128}$$

$$\begin{aligned} \widetilde{F}_{a\alpha,a'\alpha'}(\omega) = \left(TJ^2q_{\text{EA}}^{p-1} + \frac{ipJq_{\text{EA}}^{p/2-1}\big(\mathcal{E}_{\text{aux}} - 2Jq_{\text{EA}}^{p/2-1}\Lambda\big)}{|A|}\big(\delta_{\alpha u} + \delta_{\alpha' u}\big)\right)\delta_{\omega 0} \\ + \delta_{aa'}\Delta\widetilde{\mathcal{F}}_{\alpha\alpha'}(\omega) + O(T^{-1}), \end{aligned} \tag{3.129}$$

$$\lambda = \frac{ip}{|A|}\big(\mathcal{E}_{\text{aux}} - 2Jq_{\text{EA}}^{p/2-1}\Lambda\big) + O(T^{-1}). \tag{3.130}$$

Eq. (3.127) can be evaluated following the same procedure as in Sec. 3.4.3. Directly substituting Eqs. (3.128) through (3.130) gives

$$S_{\text{eff}}[\mathcal{E}_{\text{aux}}] = \frac{p\mathcal{E}_{\text{aux}}^2}{2} - 2pJ^2q_{\text{EA}}^{p-2}\Lambda^2 - \frac{Tq_{\text{EA}}}{2}\sum_{aa'}\sum_{\alpha\alpha'}(-1)^{\alpha+\alpha'}\widetilde{F}_{a\alpha,a'\alpha'}(0) - \frac{1}{2}\sum_{\omega}\log \text{Det}\widetilde{G}_{a\alpha,a'\alpha'}(\omega), \tag{3.131}$$

and we again must determine certain components of $\widetilde{F}(0)$ and $\text{Det}\widetilde{G}(0)$. As before, it is expedient to use the $(+,-)$ basis with respect to contour indices. We also switch to the Fourier basis with

respect to factor indices: from Eq. (3.128),

$$\begin{aligned}\widetilde{G}_{b\alpha,b'\alpha'}(\omega) &\equiv \frac{1}{|A|}\sum_{aa'=1}^{|A|} e^{2\pi i(ab-a'b')/|A|}\widetilde{G}_{a\alpha,a'\alpha'}(\omega) \\ &= T|A|q_{\mathrm{EA}}\delta_{b0}\delta_{b'0}\delta_{\omega 0} + \delta_{bb'}\Big(\Delta\widetilde{\mathcal{G}}_{\alpha\alpha'}(\omega) + O(T^{-1})\Big).\end{aligned} \tag{3.132}$$

Thus $\mathrm{Det}\widetilde{G}(\omega)$ factors with respect to $b$, and furthermore, $\sum_\omega \log \mathrm{Det}\widetilde{G}_b(\omega) \sim 0$ for all $b \neq 0$ as in Secs. 3.3.3 and 3.4.3. For $b = 0$, the determinant is calculated by comparing to the $b = 0$ block of $iz(-1)^\alpha + (-1)^{\alpha+\alpha'}\widetilde{F}(0)$, written in the $(+,-)$ basis (compare to Eq. (3.103)):

$$\begin{pmatrix} \widetilde{F}_{0-,0-}(0) & iz + \widetilde{F}_{0-,0+}(0) \\ iz + \widetilde{F}_{0+,0-}(0) & \widetilde{F}_{0+,0+}(0) \end{pmatrix} = \frac{1}{\mathrm{Det}\widetilde{G}_0(0)} \begin{pmatrix} \widetilde{G}_{0-,0-}(0) & -\widetilde{G}_{0+,0-}(0) \\ -\widetilde{G}_{0-,0+}(0) & \widetilde{G}_{0+,0+}(0) \end{pmatrix}. \tag{3.133}$$

We see that $\mathrm{Det}\widetilde{G}_0(0) = \widetilde{G}_{0+,0+}(0)/\widetilde{F}_{0+,0+}(0) \sim 1/J^2 q_{\mathrm{EA}}^{p-2}$, and $\widetilde{F}_{0-,0-}(0)$ (which is in fact the only element of $\widetilde{F}(0)$ needed in Eq. (3.131)) is given by $\widetilde{G}_{0-,0-}(0)/\mathrm{Det}\widetilde{G}_0(0) \sim (1 - 4J^2 q_{\mathrm{EA}}^{p-2}\Lambda^2)/2T|A|q_{\mathrm{EA}}$. The action evaluates to

$$S_{\mathrm{eff}}[\mathcal{E}_{\mathrm{aux}}] = \frac{p-2}{2p} + \frac{1}{8pJ^2 q_{\mathrm{EA}}^{p-2}\Lambda^2} - \frac{2(p-1)J^2 q_{\mathrm{EA}}^{p-2}\Lambda^2}{p} + \frac{1}{2}\log 4J^2 q_{\mathrm{EA}}^{p-2}\Lambda^2, \tag{3.134}$$

which is again precisely $-\Sigma(\mathcal{E}_{\mathrm{aux}})$.

### 3.5.4 Evaluation of the Higher SFF

In the above calculation, note that we get a single contribution of complexity for the *entire* group $A$. However, there is still a factor $(2T)^{|A|}(1+\delta_{p\text{ even}})^{2|A|-1}$ due to the separate time transla-

tion, time reversal, and reflection symmetries of each replica[5]. Finally, the sum over all connected solutions amounts to a sum over the possible ways of partitioning $n$ elements, in addition to the $n!$ ways of pairing upper and lower contours. Using $P \equiv \{A_1, \cdots, A_{|P|}\}$ to denote a partition, we have that

$$\begin{aligned}\mathrm{SFF}^{(n)}(T,f) = n! \sum_P \prod_{A\in P} 2^{|A|}\big(1+\delta_{p\text{ even}}\big)^{2|A|-1} T^{|A|}\sqrt{\frac{pN}{2\pi}}\int d\mathcal{E}_{\text{aux},A}\, e^{N\Sigma(\mathcal{E}_{\text{aux},A})} \\ \cdot \prod_{a\in A}\int_{\epsilon_-(\mathcal{E}_{\text{aux},A})}^{\epsilon_+(\mathcal{E}_{\text{aux},A})} \frac{\mathrm{d}\epsilon_{\text{aux},a}}{2\pi} f(\epsilon_{\text{aux},a})^2.\end{aligned} \tag{3.135}$$

In particular, suppose the filter function is chosen so as to have a small width $\Delta\mathcal{E} \ll 1/N$ around a certain value $\mathcal{E}$ (as in Sec. 3.4.4, the above calculation can easily be modified to allow for $\mathcal{E}$-dependent filter functions). Then the $n$'th moment simplifies to

$$\begin{aligned}\mathrm{SFF}^{(n)}(T,f) = n! \sum_P \left(\big(1+\delta_{p\text{ even}}\big)^{-1}\sqrt{\frac{pN}{2\pi}} e^{N\Sigma(\mathcal{E})}\Delta\mathcal{E}\right)^{|P|} \\ \times \left(2\big(1+\delta_{p\text{ even}}\big)^2 T\int_{\epsilon_-(\mathcal{E})}^{\epsilon_+(\mathcal{E})}\frac{\mathrm{d}\epsilon_{\text{aux}}}{2\pi} f(\epsilon_{\text{aux}})^2\right)^n.\end{aligned} \tag{3.136}$$

Eq. (3.136) has a nice interpretation as the $n$'th moment of a sum over a Poisson-distributed number of Gaussians. To be precise, suppose we have an infinite sequence of i.i.d. complex Gaussians, $\{Z_i\}_{i=1}^{\infty}$, each with $\mathbb{E}Z_i = 0$ and $\mathbb{E}Z_i Z_i^* = \sigma^2$. Consider the sum $S \equiv \sum_{i=1}^{M} Z_i$, where $M$ is itself a Poisson-distributed random variable with mean $\mu$. The $n$'th moment of $SS^*$,

[5]The contribution $(1+\delta_{p\text{ even}})^{2|A|-1}$, rather than $(1+\delta_{p\text{ even}})^{2|A|}$, is because reflecting all spin configurations does not change the values of any overlaps.

averaging over both Gaussians and $M$, can be written

$$\mathbb{E}\big[S^n S^{*n}\big] = \sum_{m=0}^{\infty} p_\mu(m)\mathbb{E}\left[\left(\sum_{i=1}^{m} Z_i\right)^n \left(\sum_{i=1}^{m} Z_i^*\right)^n\right] = n!\sum_{m=0}^{\infty} p_\mu(m)\big(m\sigma^2\big)^n \tag{3.137}$$

where $p_\mu(m)$ denotes the Poisson distribution of mean $\mu$, and Wick's theorem is used for the latter equality. It is known that the $n$'th moment of a Poisson distribution is $\sum_P \mu^{|P|}$, where the sum is again over all partitions of $n$ elements. Thus

$$\mathbb{E}\big[S^n S^{*n}\big] = n!\sum_{P} \mu^{|P|}\sigma^{2n}. \tag{3.138}$$

If we associate $\sigma^2$ with the SFF of a single TAP state at $\mathcal{E}$,

$$\sigma^2 = 2\big(1+\delta_{p\text{ even}}\big)^2 T\int_{\epsilon_-(\mathcal{E})}^{\epsilon_+(\mathcal{E})} \frac{\mathrm{d}\epsilon_{\text{aux}}}{2\pi} f(\epsilon_{\text{aux}})^2, \tag{3.139}$$

and associate $\mu$ with the number of TAP states,

$$\mu = \big(1+\delta_{p\text{ even}}\big)^{-1}\sqrt{\frac{pN}{2\pi}}e^{N\Sigma(\mathcal{E})}\Delta\mathcal{E}, \tag{3.140}$$

then Eqs. (3.136) and (3.138) are identical.

It is quite tempting to interpret this as saying that the number of TAP states at $\mathcal{E}$ is Poisson-distributed with mean given by Eq. (3.140), and that each TAP state has a Gaussian-distributed value of $\mathrm{Tr}e^{-iHT}$ with variance (i.e., SFF) given by Eq. (3.139). We are not aware of any results in the literature which would contradict such a claim. However, keep in mind that $\mathrm{SFF}^{(n)}$ has perturbative corrections around each saddle point which are suppressed by powers of $N$, whereas

every connected partition in Eq. (3.136) is suppressed *exponentially* relative to the fully disconnected one, whose contribution is given by Eq. (3.120)[6]. Thus we cannot claim to have rigorously computed the $n$'th moment to any level of accuracy beyond the disconnected piece. Nonetheless, the structure of saddle points which we have identified is highly suggestive and warrants further investigation.

## 3.6 Concluding Remarks

The focus of this paper was the derivation of Eq. (3.111), which gives the connected SFF of the quantum PSM in the non-ergodic phase. Our result demonstrates that each metastable state (i.e., TAP state) can be considered as an independent chaotic phase with an independent RMT-like Hamiltonian, at least as far as level statistics are concerned.

It is interesting to compare our result for the PSM with Ref. [75], which performs a similar calculation for the SYK model (the latter is structurally quite similar but with fermionic degrees of freedom). In the ergodic phase, we find an essentially identical result using extremely similar methods. Below the dynamical transition, however, the PSM displays an enhanced ramp quite different from that of the SYK model, which has no analogous phase.

Since we only calculate the SFF up to times polynomial in system size, our results are consistent with but do not test the distinction between localized and delocalized phases shown in Fig. 3.3, which is only relevant beyond the exponentially long timescale corresponding to tunneling between TAP states. We leave it for future work to incorporate such instanton effects into the

[6]At sufficiently low energies, where the function $\Sigma(\mathcal{E})$ is negative, the situation is reversed and the fully connected partition dominates. The issue remains, however, that we do not calculate perturbative corrections around the dominant saddle point. Furthermore, the relevance of these moment calculations to *individual* realizations of the PSM is much more suspect when $\Sigma(\mathcal{E}) < 0$.

path integral, expecting that they will reduce the SFF to the random matrix result precisely in the non-ergodic delocalized phase (and even then only beyond the exponential tunneling timescale). At the same time, consideration of exponential scales can allow one to identify a plethora of additional dynamical phases (see Ref. [160] for an example), and so the structure of instantons in these spin glass models may be quite rich.

In addition to the SFF, we have also considered higher moments of the evolution operator. We identified an important family of saddle points and evaluated its contribution to these higher SFFs (Eq. (3.136)). The results suggest that: i) the number of TAP states at a given energy is Poisson-distributed, and ii) the numbers of TAP states at different energies are independent. However, since we have not evaluated the perturbative corrections around each saddle point, which at finite complexity would generically dominate over any subleading saddle points, we cannot claim to have an accurate calculation. It is another direction for future work to study the distribution of TAP states more systematically.

Our results can be further understood by comparing to Refs. [116] and [178] on one hand and Refs. [179] and [180] on the other. The first set of papers argues that for a system which separates into weakly coupled sectors, the SFF enhancement is the sum of return probabilities over all configurations. If the time evolution can be considered as an effective Markov process with transfer rates between sectors given by some matrix $M$, then the SFF enhancement factor is $\mathrm{Tr} e^{MT}$. The second set of papers argues that for a *classical* spin glass undergoing Markovian stochastic dynamics with generator $M$, the number of TAP states can be calculated — and perhaps even defined — as $\mathrm{Tr} e^{MT}$. In this sense, the present paper can be considered as a "missing link" that extends the results of Refs. [179] and [180] to quantum systems.

The fact that SFF enhancement is related to return probabilities suggests that the spectral

statistics of spin glasses may contain information on aging dynamics as well. Another open question is whether the *equilibrium* replica-symmetry-breaking transition has any consequences for spectral statistics. These, as well as those already mentioned, are all promising directions for future work.

Finally, let us briefly comment on the case $p = 2$, which — being a Gaussian model but for the global spherical constraint (and still integrable in any case [181]) — exhibits very different behavior than the $p > 2$ models considered here. The spectral statistics of the analogous SYK model with two-point interactions among fermions have been studied in the mean-field limit [117, 118], and found to have ramps growing exponentially in time. One explanation for these ramps is the spontaneous breaking of a hidden $SU(2)^k$ symmetry in the saddle point equations, since for $p = 2$ the matrix $G$ has a separate $SU(2)$ conjugation symmetry at each frequency. Inspections show that the $p = 2$ spherical model has the same symmetry at tree level, although it is broken by higher loop effects. A full analysis of this system would be yet another excellent topic for further research.

## Acknowledgements

This work was supported by the following: the U.S. Department of Energy, Office of Science, Basic Energy Sciences under award number DE-SC0001911 (V.G.); the Joint Quantum Institute (M.W.); the Air Force Office of Scientific Research under award numbers FA9550-17-1-0180 (M.W.) and FA9550-19-1-0360 (B.S.); the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Accelerated Research for Quantum Computing program "FAR-QC" (R.B.); the DoE ASCR Quantum Testbed Pathfinder program under

award number DE-SC0019040 (C.L.B.); the DoE ASCR Accelerated Research in Quantum Computing program under award number DE-SC0020312 (C.L.B.); the DoE QSA, AFOSR, AFOSR MURI, NSF PFCQC program, NSF QLCI under award number OMA-2120757 (C.L.B.); DoE award number DE-SC0019449 (C.L.B.), ARO MURI, and DARPA SAVaNT ADVENT (C.L.B.). This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE 1840340 (R.B.), and by the National Science Foundation NRC postdoctoral fellowship program (C.L.B.).

# Chapter 4: Slow Relaxation in a Glassy Quantum Circuit

**Authors:** Richard Barney, Yunxiang Liao, Victor Galitski

**Abstract:** Quantum circuits have become a powerful tool in the study of many-body quantum physics, providing insights into both fast-thermalizing chaotic and non-thermalizing integrable many-body dynamics. In this work, we explore a distinct intermediate class—glassy quantum systems—where thermalization occurs, but over very long timescales. We introduce and analyze a Floquet random quantum circuit that can be tuned between glassy and fully ergodic behavior through a single adjustable parameter. This circuit can be understood as the unitary analog of the block Rosenzweig-Porter model, which is defined by a Hamiltonian. Using an effective field theory for random quantum circuits, we analyze the correlations between quasienergy eigenstates and thereby determine the time evolution of the disorder-averaged density matrix. In the intermediate regime the circuit displays a two-step thermalization process: an initial relaxation within weakly coupled sectors followed by a later, global thermalization. We also show that the ramp of the spectral form factor is enhanced by a factor of the number of sectors in the glassy regime, as well as at early times in the intermediate regime. These results indicate that quantum circuits provide an ideal platform for the exploration of nontrivial thermalization dynamics in many-body quantum systems, offering deeper insights into quantum thermalization.

## 4.1 Introduction

Understanding which type of isolated quantum interacting systems, when starting from an out-of-equilibrium initial state, eventually relax to thermal equilibrium is a foundational question for the validity and applicability of quantum statistical mechanics. This question is partially addressed by the eigenstate thermalization hypothesis (ETH) [29–31, 139], which suggests that quantum chaotic systems, characterized by particular energy eigenstate statistics, can achieve thermal equilibrium under their own dynamics. Compared with earlier attempts which try to explain thermalization in isolated quantum systems, ETH can distinguish systems that can thermalize from those that cannot, based on different statistical properties of their energy eigenstates.

ETH may break down in certain types of quantum systems, including integrable systems with an extensive number of conserved quantities [132, 182–186] and many-body localized (MBL) systems where strong disorder prevents thermalization even in the presence of interactions [187–192]. While there is a substantial number of numerical [193–201] and experimental [202–211] studies supporting MBL in small size systems, the existence of MBL in the thermodynamic limit remains a topic of active debate [212–217]. Additionally, there exists a special type of non-ergodic system where thermalization can occur, but at an extremely slow rate: glasses [2, 7, 12, 15, 16, 18, 145, 146, 164, 167, 218, 219]. The slow relaxation dynamics in these systems remain relatively less explored, particularly concerning the the underlying microscopic mechanism.

Wigner-Dyson energy level statistics [33, 220] has been widely used as a defining property of quantum chaotic systems [75, 149, 221–225], based on the Bohigas-Giannoni-Schmit conjecture [32]. By contrast, integrable systems and MBL systems instead exhibit Poisson statistics

[115, 193]. Recently, it has been demonstrated that a particular quantum spin glass model, the quantum p-spherical model, exhibits energy level statistics distinct from those of quantum chaotic systems and of integrable systems [2]. Its statistical properties are analogous to that of a random matrix model which falls outside the Wigner-Dyson class, known as the block Rosenzweig-Porter model [1]. In particular, the spectral form factors (SFFs) [226, 227]—a diagnostic of energy level statistics—of both the quantum p-spherical model and the block Rosenzweig-Porter model exhibit a linear-in-time ramp, analogous to Wigner-Dyson statistics, but with a coefficient that is enhanced by the number of weakly coupled sectors (metastable configurations).

Quantum circuits have become a widely used platform for the study of many-body quantum physics and can be implemented in digital quantum simulators in various experimental platforms [228–235]. In particular, some of their fine-tuned properties, such as dual unitarity [114, 236–238] or the randomness of the quantum gates [127, 222, 239–246], allow for exact theoretical investigation of the thermalization dynamics as well as other chaotic behaviors which also occur in more realistic quantum chaotic systems. This makes quantum circuits an ideal candidate for the exploration of ergodicity breaking and realization of glassy phases characterized by extremely slow relaxation processes, leading to a deeper understanding of the fundamental question of quantum thermalization.

In this paper, we propose a Floquet random quantum circuit, a random circuit with discrete time-translation symmetry, which can exhibit slow glassy thermalization dynamics. The Floquet operator of this quantum circuit can be considered as a unitary version of the block Rosenzweig-Porter model. The quasienergy eigenstates and eigenvalues of this glassy quantum circuit possess statistical properties distinct from those of quantum chaotic systems. Its relaxation towards thermal equilibrium involves two steps: first thermalization occurs within different sectors that are

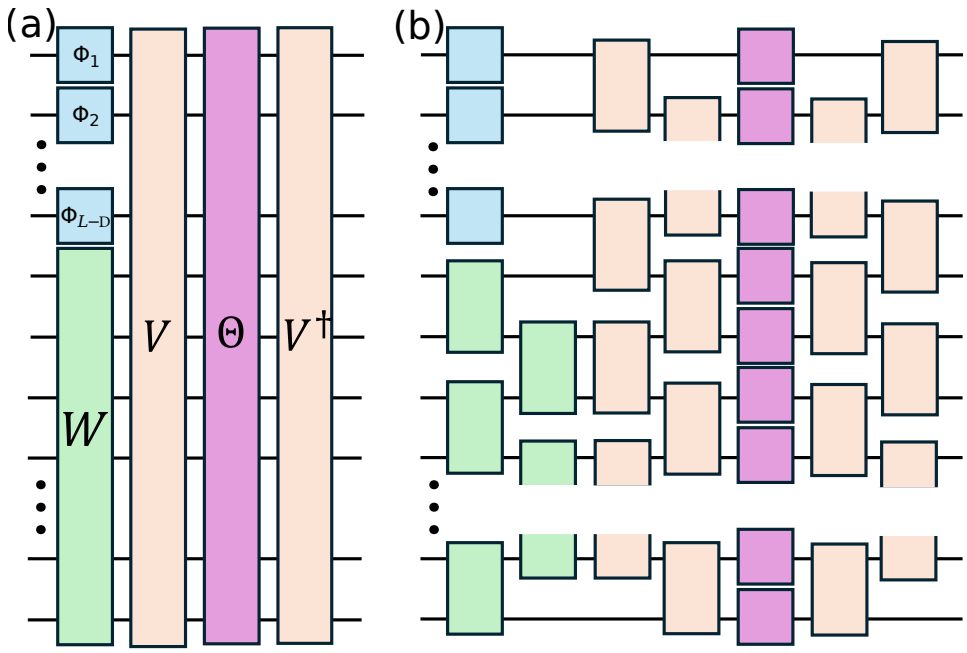



Figure 4.1: (a) The time periodic glassy random quantum circuit's Floquet operator. The first layer involves independent random phase gates $\Phi_i$ acting individually on the first $L - D$ qudits, together with a CUE random matrix $W$ acting on the remaining $D$ qudits. The subsequent layers include an independent CUE matrix $V$, a global quantum phase gate $\Theta$, and $V^\dagger$, all acting on the entire chain of qudits. The random phase gates $\Phi$ and $\Theta$ are diagonal matrices with dimensions $q$ and $q^L$, respectively. The phases in $\Phi$ are independently and uniformly distributed over $(-\pi, \pi]$, while those in $\Theta$ are independently Gaussian distributed with zero mean and variance $\sigma^2$. By tuning $\sigma$, the process of thermalization can be significantly slowed down, and it is completely prevented in the $\sigma \to 0$ limit, in which case only the first layer is effective. (b) A schematic representation of the Floquet operator implemented with local gates. In practice, the global operators $W$, $V$ and $\Theta$ can be implemented using layers of local quantum gates that act on individual or neighboring qudits, as long as the effective time evolution operators exhibit the same statistical properties.

weakly coupled to each other; then this is followed by global thermalization across the entire Hilbert space. Moreover, by tuning a parameter in the model that represents the strength of the coupling between different sectors, the rate of global thermalization can be made extremely slow.

## 4.2 The circuit

The Floquet quantum circuit under consideration consists of a one-dimensional chain of $L$ qudits, each with $q$ internal levels. For our analytic calculation, we focus on the large $q$ limit. The qudits evolve under discrete and periodic applications of unitary quantum gates depicted in Fig. 4.1(a). In particular, the Floquet operator $U$ takes the form of

$$U = V^{\dagger}\Theta V \left( \bigotimes_{i=1}^{L-D} \Phi_i \otimes W \right). \tag{4.1}$$

Here $\Phi_i \equiv \mathrm{diag}\,(e^{i\phi_1^{(i)}}, ..., e^{i\phi_q^{(i)}})$ represents a random phase gate acting on a single qudit at site $1 \leq i \leq L-D$, with the diagonal matrix element $\phi_j^{(i)}$ being independently and uniformly distributed within the range $(-\pi, \pi]$. $\Theta = \mathrm{diag}\,(e^{i\theta_1}, ..., e^{i\theta_{q^L}})$ is another random phase gate that acts on all qudits, where each random phase $\theta_j$ follows an independent Gaussian distribution with zero mean and variance $\sigma^2$. $W$ ($V$) is a Haar-distributed $M \times M$ ($N \times N$) random unitary matrix acting on sites $[L-D+1, L]$ (all sites), where $M \equiv q^D$ ($N \equiv q^L$) is the Hilbert space of $D$ ($L$) qudits. In practice, the random unitary gates $W$ and $V$ can be constructed from layers of two-qudit unitary gates, such that the statistical properties of the resulting effective time evolution operators match those of the circular unitary ensemble (CUE) of the same dimensions [84]. Similarly, the global random phase gate $\Theta$ can also be generated from a series of single-qudit random phase gates (Fig.4.1(b)).

The time evolution operator for the first layer of quantum gates $\left(\bigotimes_{i=1}^{L-D} \Phi_i \otimes W\right) = \bigoplus_{i=1}^{P} A^{(i)}$ has a block diagonal structure with each diagonal block being an $M \times M$ matrix of the form $A^{(i)} = e^{i\phi_i'} W$. Here $P$ is defined as $P \equiv N/M$. All phases $\phi_i'$ are given by

sums of $L - D$ random phases from the set $\left\{\phi_j^{(l)} | j = 1, 2, ..., q; l = 1, 2, ..., L - D\right\}$. Using the fact that $\phi_j^{(l)}$ are independently and uniformly distributed, one can show that the correlation $\left\langle A_{ij}^{(l)} (A^{(l')})_{j'i'}^{\dagger} \right\rangle = \delta_{ll'}\delta_{ii'}\delta_{jj'}/M$, identical to the correlation of statistically independent CUE matrices of dimension $M$. Here we have used the angular bracket to indicate ensemble averaging. In the following, we approximate $A^{(l)}$, $l = 1, 2, ..., P$, by statistically independent CUE matrices of dimension $M$ to simplify the calculation, believing this approximation effectively captures the essential relaxation behavior of the glassy quantum circuit under consideration.

The remaining layers of quantum gates $V\Theta V^{\dagger}$ couple the diagonal blocks $A^{(i)}$, with the coupling strength tuned by the variance $\sigma^2$ of the random phase gate $\Theta$. When $\sigma = 0$, $\Theta$ becomes an identity matrix, making the Floquet operator $U = \bigoplus_{i=1}^{P} A^{(i)}$ block diagonal. In the opposite limit $\sigma \gg 1$, the coupling between $A^{(i)}$ becomes so strong that we expect the random circuit to be quantum chaotic with Wigner-Dyson statistics. In the intermediate regime where $\sigma \sim O(1)$, the Floquet operator $U$ possesses a structure analogous to that of the block Rosenzweig-Porter model in the intermediate regime, which serves as a minimal model for quantum spin glasses. Therefore, similar glassy behaviors in its relaxation dynamics are expected in this intermediate regime.

The correlation of the Floquet operator $U$ is given by,

$$\left\langle U_{ij} U_{j'i'}^{\dagger} \right\rangle = \delta_{ii'}\delta_{jj'} \left( \delta_{B(i),B(j)} \frac{1 + Ne^{-\sigma^2}}{M(N+1)} + \frac{1 - e^{-\sigma^2}}{N+1} \right). \tag{4.2}$$

Here $B(i) \equiv \lceil i/M \rceil$ denotes the block index to which the integer $i$ belongs when the $N \times N$ matrix is divided into $P \times P$ blocks, with each block being an $M \times M$ matrix. When $\sigma = 0$, this correlation function reduces to that of a block CUE, i.e., an $N \times N$ block diagonal matrix

with statistically independent CUE matrices of dimension $M$ as its diagonal blocks. In the large $\sigma$ limit the first term becomes much smaller than the second term, which reduces the correlation function to that of a CUE matrix. However, due to the difference in the Kronecker delta functions of both terms, the first term cannot be simply ignored when investigating the relaxation behavior.

## 4.3 The quasienergy eigenstate correlation function

To investigate the statistical properties of the quasienergy eigenstates and eigenvalues of the Floquet random quantum circuit as well as its relaxation behaviors, we compute the following correlation function of the quasienergy eigenstates

$$C_{nn'm'm}(\omega) = \left\langle \sum_{\mu,\nu} \psi_n^\mu (\psi_{n'}^\mu)^* \psi_{m'}^\nu (\psi_m^\nu)^* \delta_{2\pi} \left(\omega - E_\mu + E_\nu\right) \right\rangle . \tag{4.3}$$

Here $\psi_n^\mu$ denotes the $n^{\text{th}}$ component of the quasienergy eigenstate with quasienergy $E_\mu$, and $\delta_{2\pi}$ represents the $2\pi$-periodic Dirac delta function. The Fourier transform of $C_{nn'm'm}(\omega)$ is the correlation function of the matrix elements of the time evolution operator $U(t) = U^t$,

$$C_{nn'm'm}(t) = \left\langle U_{nn'}(t) U^\dagger_{m'm}(t) \right\rangle . \tag{4.4}$$

From this eigenstate correlation function one can obtain the time evolution of the density matrix $\rho(t)$ starting from an arbitrary initial state $\rho(0)$:

$$\langle \rho_{nm}(t) \rangle = \sum_{n'm'} C_{nn'm'm}(t) \rho_{n'm'}(0), \tag{4.5}$$

as well as the SFF:

$$K(t) \equiv \left\langle |\operatorname{Tr} U(t)|^2 \right\rangle = \sum_{nm} C_{nnmm}(t). \tag{4.6}$$

To compute the eigenstate correlation function for the Floquet quantum circuit under consideration, we employ a sigma model [81, 247, 248] approach developed in Refs. [249, 250]. Since the detailed calculation is analogous to that in Ref. [250] for a family of chaotic quantum circuits, we leave the full derivation to Appendix C.1. By focusing on the quadratic fluctuations around the standard saddle point in the sigma model, we obtain the smoothed eigenstate correlation function $C_{nn'm'm}(\omega)$, where the $2\pi-$periodic Dirac delta function in Eq. 4.3 is replaced by a sharp Lorenzian $L_\eta(\omega) \equiv \sum_{n=-\infty}^{\infty} e^{i\omega n - |n|\eta}/2\pi$ of width $M^{-1} \ll \eta \ll 1$ [248]. This smoothed correlation function measures the overlap of pairs of the quasienergy eigenstates $\mu, \nu$ whose energy separation $E_\mu - E_\nu$ is close to $\omega$, weighted by the Lorenzian $L_\eta(\omega - E_\mu + E_\nu)$. Note that both $\eta$ and $1/q$ serve as the small parameters in our perturbative calculation [248, 251, 252]. To obtain oscillatory corrections to the smoothed correlation function, which reveals the structure of nearby quasienergy eigenstates, or to investigate the small $q$ case, it becomes important to consider the fluctuations beyond the quadratic order. We leave this investigation for future work.

The smoothed quasienergy eigenstate correlation function takes the form:

$$\begin{aligned} C_{nn'm'm}(t) = \delta_{nn'}\delta_{mm'}\left(\delta_{t,0} + c_1(t)\right) + \delta_{nm}\delta_{n'm'}c_2(t) \\ + \delta_{nn'}\delta_{mm'}\delta_{B(n),B(m)}c_3(t) + \delta_{nm}\delta_{n'm'}\delta_{B(n),B(n')}c_4(t), \end{aligned} \tag{4.7}$$

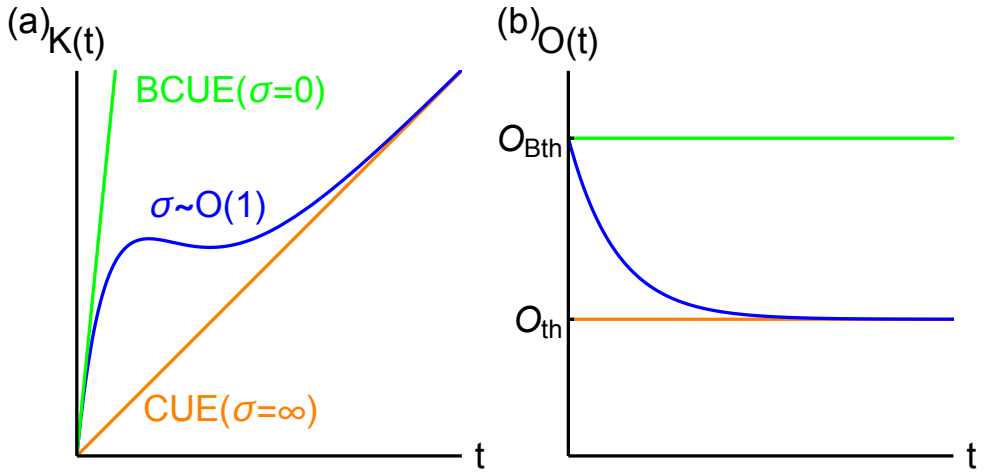



Figure 4.2: The glassy quantum circuit's (a) smoothed SFF $K(t)$ (Eq. 4.9) and (b) the time dependence of the expectation value for a physical observable $O(t)$ in the intermediate time regime (Eq. 4.1). Panel (a) shows the SFF in the intermediate $\sigma$ regime (blue curve), which crosses over from the enhanced ramp for block CUE associated with $\sigma \to 0$ (green curve) at early times to the standard ramp of CUE corresponding to $\sigma \to \infty$ (orange curve) at later times. Panel (b) illustrates the relaxation of the expectation value of physical observable $O(t)$ from the block thermalized value $O_{\text{Bth}}$ to the fully thermalized value $O_{\text{th}}$ using Eq. 4.11. The relaxation rate is also given by $\sigma$. Note here we show only the regime where the terms involving $\rho(0)$ and $\rho_{\text{B}}(0)$, which are important for the early relaxation from $O(0)$ to $O_{\text{Bth}}$, are negligible.

where, to the leading order,

$$
\begin{aligned}
c_1(t) &= \frac{1}{N^2}\left[t(1+e^{-t\sigma^2}) - \frac{1+e^{-\sigma^2}}{1-e^{-\sigma^2}}(1-e^{-t\sigma^2})\right], \\
c_2(t) &= \frac{1}{N}\left(1-e^{-t\sigma^2}\right), \\
c_3(t) &= \frac{1}{N^2}\left[\frac{2Pe^{-\sigma^2}}{1-e^{-\sigma^2}}\left(1-e^{-t\sigma^2}\right) + (P^2-2P)te^{-t\sigma^2} - P^2(e^{-t\sigma^2}-\delta_{t,0})\right], \\
c_4(t) &= \frac{1}{M}\left(e^{-t\sigma^2}-\delta_{t,0}\right).
\end{aligned}
\tag{4.8}
$$

See Eq. C.21 for the corresponding expression in frequency space. We emphasize that this expression, obtained from the Fourier transform of the smoothed eigenstate correlation function in frequency space, applies only at times much shorter than the block Heisenberg time, which is of the order $O(M)$.

## 4.4 The spectral form factor

The smoothed SFF can be derived from Eqs. C.29 and C.27:

$$K(t) = e^{-\sigma^2 t} Pt + (1 - e^{-\sigma^2 t})t + N^2 \delta_{t,0}. \tag{4.9}$$

As expected, this result reduces to that of the CUE when $\sigma \to \infty$, and to that of the block CUE when $\sigma \to 0$. In both cases, the SFF exhibits a linear-in-$t$ ramp, but the ramp in the latter case is enhanced by a factor of the number of blocks $P$. This enhancement is straightforward to understand since the SFF for the block CUE is simply given by the summation of the SFFs for each individual CUE block. Note that the plateau in the non-smoothed SFF occurs at large time and therefore cannot be recovered from the current perturbative calculation. In the intermediate $\sigma$ regime, the SFF of the current model interpolates between the limits of $\sigma \to 0$ and $\sigma \to \infty$, as shown in Fig. 4.2(a). Specifically, it crosses over from the enhanced ramp of the block CUE at small times to the standard ramp of the CUE at larger times, with the rate $\sigma$. This SFF structure resembles that of the block Rosenzweig-Porter model [1]. Notably, a similar enhanced ramp is observed in the SFF of the quantum p-spherical model [2], with the enhanced factor $P$ being the number of metastable states.

We also calculate the SFF to leading order using the diagrammatic method for integration over the unitary group developed by Brouwer and Beenakker [253]. The details of this calculation are contained in Appendix C.2. With this method we find that, for times much less than the

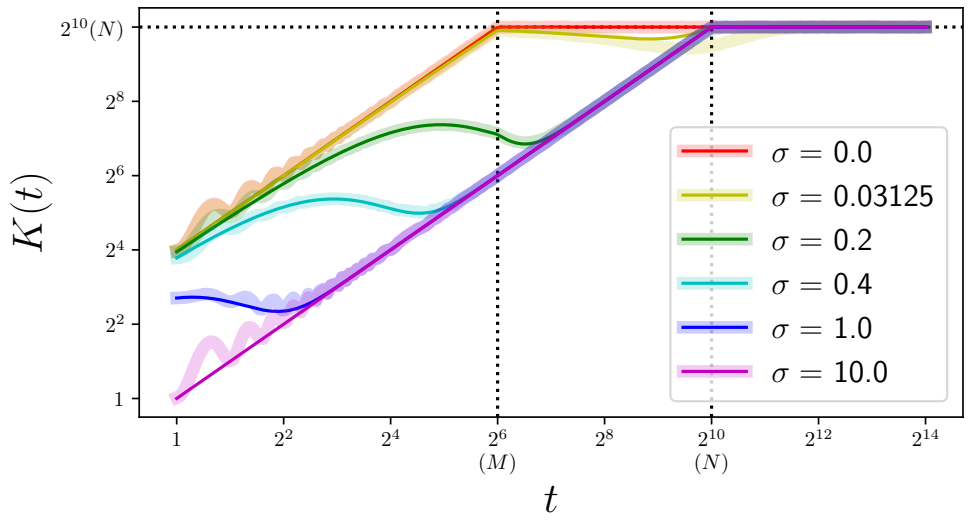



Figure 4.3: Comparison between the diagrammatic result for the SFF (Eq. 4.10) and numerics at different values of $\sigma$ for a system of 10 qubits ($q = 2$, $D = 6$). The thin dark curves are the theoretical results while the wider light curves are the numerical results averaged over $10^5$ realizations of the Floquet operator. The numerical results were obtained by treating the first layer of gates as a direct sum of independent CUE matrices.

Heisenberg time, the SFF is

$$K(t) = e^{-\sigma^2 t} P \min(t, M) + \left(1 - e^{-\sigma^2 t}\right) t + N^2 \delta_{t,0}. \tag{4.10}$$

This is in agreement with the result from the sigma model approach at times less than the block Heisenberg time (Eq. 4.9). Comparison of this diagrammatic result with numerics is shown in Fig. 4.3. We see that there is good agreement at times much less than the Heisenberg time. The oscillations in the numerical results at early times are a result of the finite system size. If we naively extend the diagrammatic result to later times by saying Eq. 4.10 holds for $t < N$ and the SFF is at its plateau for all later times, there is still good agreement with numerics at these late times, so long as the time at which the SFF moves from the glassy result to the CUE result is not on the same order as the Heisenberg time.

## 4.5 Relaxation behavior

From the eigenstate correlation function in Eq. C.27, we can also directly study the the time evolution of the density matrix using Eq. 4.5:

$$\begin{aligned}\langle\rho(t)\rangle =& \left(\delta_{t,0} + c_1(t)\right)\rho(0) + Nc_2(t)\rho_{\mathsf{th}} \\ &+ c_3(t)\rho_{\mathsf{B}}(0) + Mc_4(t)\rho_{\mathsf{Bth}},\end{aligned} \tag{4.11}$$

where $c_i(t)$ are given by Eq. C.28 to the leading order. Here $\rho_{\mathsf{th}} = \hat{1}/N$ is the thermal density matrix at infinite temperature. $\rho_{\mathsf{B}}(0)$ and $\rho_{\mathsf{Bth}}$ are defined separately as

$$\begin{aligned}(\rho_{\mathsf{B}}(0))_{nm} =& \rho_{nm}(0)\delta_{B(n),B(m)}, \\ (\rho_{\mathsf{Bth}}(0))_{nm} =& \frac{\delta_{nm}}{M}\,\mathrm{tr}_{B(n)}\,\rho(0),\end{aligned} \tag{4.12}$$

where $\mathrm{tr}_i$ denotes the trace restricted to the $i$-th diagonal block, defined as $\mathrm{tr}_i\, A \equiv \sum_{j=1}^{M} A_{i(M-1)+j,i(M-1)+j}$. In particular, $\rho_{\mathsf{B}}(0)$ is the initial density matrix $\rho(0)$ with all off-diagonal blocks set to zero, making it block diagonal with the same diagonal blocks as $\rho(0)$. $\rho_{\mathsf{Bth}}$ is the block thermal density matrix to which the system, constituted of independent chaotic blocks of equal dimensions, will eventually relax.

In the limits $\sigma \to \infty$ and $\sigma \to 0$, the time evolution of the current model's density matrix $\langle\rho(t)\rangle$ agrees, to the leading order, with that of CUE ($\rho_{\mathsf{CUE}}$) and block CUE ($\rho_{\mathsf{BCUE}}$), respec-

tively [1],

$$\langle \rho_{\mathsf{CUE}}(t) \rangle = \frac{K_N(t)-1}{N^2-1}\rho(0) + \frac{N^2-K_N(t)}{N^2-1}\rho_{\mathsf{th}}, \tag{4.13a}$$

$$\langle \rho_{\mathsf{BCUE}}(t) \rangle = \delta_{t,0}\rho(0) + \left(\frac{K_M(t)-1}{M^2-1} - \delta_{t,0}\right)\rho_{\mathsf{B}}(0) + \frac{M^2-K_M(t)}{M^2-1}\rho_{\mathsf{Bth}}. \tag{4.13b}$$

Here $K_d(t) = \min(t,d) + d^2\delta_{t,0}$ represents the SFF for CUE of dimension $d = N, M$. For generic quantum chaotic systems, Eq. 4.13a is also expected to hold with $K_N(t)$ replaced by the system's actual SFF, which now contains an early time non-universal slope in additional to the universal ramp and plateau. It is this non-universal slope which determines the detailed relaxation dynamics of the systems [254]. For the block CUE described by Eq. 4.13b, the system now thermalizes within each individual sector and its density matrix eventually relaxes to $\rho_{\mathsf{Bth}}$. Since its sectors are uncoupled from each other, it can never fully thermalize to $\rho_{\mathsf{th}}$.

In the intermediate $\sigma$ regime, substituting Eq. C.28 into Eq. 4.11 shows that the term involving $\rho(0)$ in $\langle \rho(t) \rangle$ instantly becomes negligible while the term involving $\rho_{\mathsf{B}}(0)$ remains small compared to the other terms. This may no longer be the case for small $q$, where $c_1(t)$ and $c_3(t)$ are expected to initially decay at a rate determined by $q$, similar to the slope in the SFF. While Eq. C.28 shows that the magnitudes of $c_{1,3}(t)$ can grow with time, their contributions are expected to remain small. A non-perturbative analysis is required to determine the behaviors of $c_{1,3}(t)$, as well as their contribution to $\langle \rho(t) \rangle$, at large times or for finite $q$. By contrast, terms involving $\rho_{\mathsf{th}}$ and $\rho_{\mathsf{Bth}}$ in $\langle \rho(t) \rangle$ are non-negligible, and also comparable with each other in the intermediate $\sigma$ regime. The corresponding coefficients behave such that $Nc_2(t)$ grows to 1 over time, while

[1] More precisely, the coefficients in front of all $\rho's$ agree with those of the CUE and block CUE to the leading order

$Mc_4(t)$ decays to 0, at the rate $\sigma$. This indicates a relaxation from the block thermalized density matrix $\rho_{\mathsf{Bth}}$ to the globally thermalized one $\rho_{\mathsf{th}}$. This process can be made extremely slow by tuning $\sigma$.

In Fig. 4.2(b), we illustrate schematically the relaxation of the expectation value $O(t) = \langle \mathrm{tr}(\rho(t)O)\rangle$ for an arbitrary physical observable $O$ from the block thermalized value $O_{\mathsf{Bth}} = \mathrm{Tr}(O\rho_{\mathsf{Bth}})$ to the fully thermalized value $O_{\mathsf{th}} = \mathrm{Tr}(O\rho_{\mathsf{th}})$. It is described by

$$O(t) = e^{-\sigma^2 t} O_{\mathsf{Bth}} + \left(1 - e^{-\sigma^2 t}\right) O_{\mathsf{th}}. \tag{4.1}$$

In both Fig. 4.2(b) and Eq. 4.1, we focus on the relaxation from block thermalized $O_{\mathsf{Bth}}$ to the global thermalized $O_{\mathsf{th}}$ and therefore have ignored contribution from terms involving $\rho(0)$ and $\rho_{\mathsf{B}}(0)$ in $\langle\rho(t)\rangle$, which are important for the initial relaxation from $O(0)$ to $O_{\mathsf{Bth}}$. To fully capture the explicit time dependence of $O(t)$ over the entire thermalization process, especially early relaxation from $O(0)$ to $O_{\mathsf{Bth}}$, the higher order terms in $c_{2,4}(t)$ are needed, as their contribution to Eq. 4.11 becomes comparable to that of $c_{1,3}(t)$. However, the existence of a two-step relaxation process is apparent even just from the current leading order result. This process consists of an early time thermalization within individual sectors from $O(0)$ to $O_{\mathsf{Bth}}$, followed by a subsequent global thermalization from $O_{\mathsf{Bth}}$ to $O_{\mathsf{th}}$.

## 4.6 Conclusion

Eq. 4.1 for the thermalization behavior and Eqs. 4.9-4.10 for the SFF have similar structures in that they move from the block CUE result for $t \ll \sigma^{-2}$ to the CUE result for $t \gg \sigma^{-2}$. This tells us that the Thouless time for the circuit, the time it take for the system to escape from a

sector and fully thermalize, is $t_{\text{Th}} \sim \sigma^{-2}$. We find that the circuit exhibits a phase structure analogous to that of the block Rosenzweig-Porter model. In order to see sharp transitions we may reparameterize such that $\sigma^2 \sim N^{-\gamma}$, making the Thouless time $t_{\text{Th}} \sim N^{\gamma}$. If the Thouless time is much smaller than the smallest possible timescale, which is 1 for this circuit with discrete time steps, the system behaves ergodically. So the system is in the ergodic phase for $\gamma < 0$. The Heisenberg time is the longest meaningful timescale for a quantum system. If the Thouless time becomes much larger than the Heisenberg time, the system will never escape a sector and thermalize. This means that the system is in the localized phase for $\gamma > 1$. For $0 < \gamma < 1$ the system will be in the intermediate glassy phase, where thermalization occurs, but only on a timescale exponentially long in the system size.

In this paper, we introduce a Floquet random quantum circuit whose thermalization process can be made extremely slow by tuning a parameter of a specific quantum gate. The statistical properties of quasienergy eigenstates in this model are analogous to those of quantum spin glass models, but are distinct from those of quantum chaotic systems, as well as integrable and MBL systems. We investigate the time evolution of the density matrix for this quantum circuit model in the limit of large local Hilbert space dimensions and for times much shorter than the block Heisenberg time. We find that the relaxation process consists of an initial thermalization within weakly coupled sectors and a subsequent global thermalization over the entire Hilbert space.

The explicit time dependence is described by Eq. 4.11, which is an ensemble averaged result. To show that this result is typical for individual realizations of the quantum circuit in the ensemble [254], it is necessary to examine the statistical fluctuations of $\rho(t)$ by computing higher order correlation functions of the quasienergy eigenstates. A non-perturbative calculation is required to to investigate the small $q$ case and to fully understand the complete behavior of $\rho(t)$,

including the early time thermalization within individual sectors and the later time dynamics around or after the block Heisenberg time. To address these issues, the non-perturbative treatment by Kravtsov and Mirlin [255], initially invented for studying the energy level correlation function in disordered systems, can be applied. We leave these investigations for future work. Due to the specific structure of the Floquet operator, which can be considered as weakly coupled blocks of equal dimensions, our model's relaxation process contains two distinct steps. More realistic quantum systems may display a more complex relaxation behavior, and could be modeled by a generalized quantum circuit whose Floquet operator exhibits a higher level of hierarchy. Generalizing the glassy quantum circuit to incorporate a richer structure in relaxation dynamics is another direction for future work.

## Acknowledgments

This research was sponsored by the Schwinger Foundation, Army Research Office under Grant Number W911NF-23-1-0241, the National Science Foundation under Grant No. DMR-203715, and the NSF QLCI grant OMA-2120757.

# Chapter 5: Neural Networks as Spin Models: From Glass to Hidden Order Through Training

**Authors:** Richard Barney, Michael Winer, Victor Galitski

**Abstract:** We explore a one-to-one correspondence between a neural network (NN) and a statistical mechanical spin model where neurons are mapped to Ising spins and weights to spin-spin couplings. The process of training an NN produces a family of spin Hamiltonians parameterized by training time. We study the magnetic phases and the melting transition temperature as training progresses. First, we prove analytically that the common initial state before training—an NN with independent random weights—maps to a layered version of the classical Sherrington-Kirkpatrick spin glass exhibiting a replica symmetry breaking. The spin-glass-to-paramagnet transition temperature is calculated. Further, we use the Thouless-Anderson-Palmer (TAP) equations—a theoretical technique to analyze the landscape of energy minima of random systems—to determine the evolution of the magnetic phases on two types of NNs (one with continuous and one with binarized activations) trained on the MNIST dataset. The two NN types give rise to similar results, showing a quick destruction of the spin glass and the appearance of a phase with a hidden order, whose melting transition temperature $T_c$ grows as a power law in training time. We also discuss the properties of the spectrum of the spin system's bond matrix in the context of rich vs. lazy learning. We suggest that this statistical mechanical view of NNs

provides a useful unifying perspective on the training process, which can be viewed as selecting and strengthening a symmetry-broken state associated with the training task.

## 5.1 Introduction

The synergy between machine learning and statistical mechanics has been a long-standing topic of research [256–263]. In particular, there is a natural relation between spin systems and neural networks (NNs), including multilayer perceptrons [43], restricted Boltzmann machines [264], and Hopfield networks [39, 265, 266], where neurons correspond to the spin degrees of freedom, the biases to magnetic fields, and the weights to the spin-spin couplings. For a random realization of weights between neurons, spin glass physics naturally becomes relevant, along with the methods of studying glasses such as replica symmetry breaking [13, 15, 16, 267, 268] and the Thouless-Anderson-Palmer (TAP) equations [16, 23, 269]. These methods are potentially useful in transplanting intuition from statistical mechanics (phase transitions, universality, emergent behavior, etc.) to the more practical field of NNs, with the potential to elucidate the underlying mechanisms for when and how they work in accomplishing various tasks.

Another motivation to study this machine learning/statistical mechanics correspondence is the recent progress in the field of neuromorphic computing [270, 271], which employs physical analogue networks rather than algorithmic networks. While still at the experimental state of development, there has been notable recent success in building such NNs with ultracold atoms in cavity QED systems. Unlike feed-forward NNs, there is no unidirectional layered structure in such systems (nor in biological NNs), yet they share similarities with current commercial machine learning platforms and show potential for performing useful tasks. While most work so

far has focused on experimentally probing classical concepts (associative memory [272], replica symmetry breaking [273, 274], etc.) with these systems, they include quantum components from the outset. Hence putting together an underlying theoretical framework to include both classical and quantum NNs is a topic of great interest.

This paper studies two types of NNs with distinct training processes for a classification task from the perspective of statistical mechanics, in particular using the TAP equations. These are conventionally used in the studies of random spin systems to quantitatively characterize their energy landscapes and identify phases and phase transitions. We argue that the TAP equations also provide both a powerful tool to explore traditional NNs and a bridge to their quantum counterpart. The latter is left for future work, while here we study the training process of an NN viewed as a family of classical spin Hamiltonians and explore their phases as a function of "time"—the epoch of the training process.

Specifically, we consider two types of feed-forward NNs, one with step and one with ReLU activations. We train the NNs on an image classification task: recognition of images of handwritten digits from $0$ to $9$ with the MNIST dataset [66]. At each step of training, which we parameterize by training time $t$ measured in epochs, we have a set of weights $J_{ij}(t)$. We associate each such instance with a classical Ising spin Hamiltonian, where the weights are the couplings between the spins, and explore the flow of its energy landscape and phases using the TAP equations.

The conventional starting point of training, at $t = 0$, is a random realization of couplings. We show analytically that its spin counterpart exhibits a spin glass transition with replica symmetry breaking and calculate the melting temperature $T_c$ of the glass. Throughout the training process we numerically calculate $T_c$ by determining the temperature at which nontrivial solutions to the TAP equations appear. We note distinct regimes of power-law scaling of the transition tem-

perature as a function of training, $T_c(t) \propto t^{\alpha}$, where $\alpha$ appears to depend on the rate at which the NN is learning. We also consider the number of TAP solutions that appear at the transition and observe a rapid flow of the spin Hamiltonian from the initial glassy Hamiltonian into one with a $\mathbb{Z}_2$ symmetry breaking phase transition. We note a further enhancement of the hidden order, which manifests itself in the increase in $T_c$ with training. We speculate that the hidden order encodes the classification task, where the outcome—the output layer—is characterized by a set of parameters much smaller than the width of a layer (512 neurons) which encodes a hidden order parameter. Schematically, we observe a phase diagram as shown in Fig. 5.1 for temperatures not much below $T_c$. We determine that these changes can be primarily attributed to the dynamics of the spectrum of the bond matrix $J(t)$. We also comment on the qualitative similarities between the final bond matrix spectrum and the spectra of the weight matrices connecting adjacent layers of NNs in the rich learning regime, as observed in previous works [275–278].

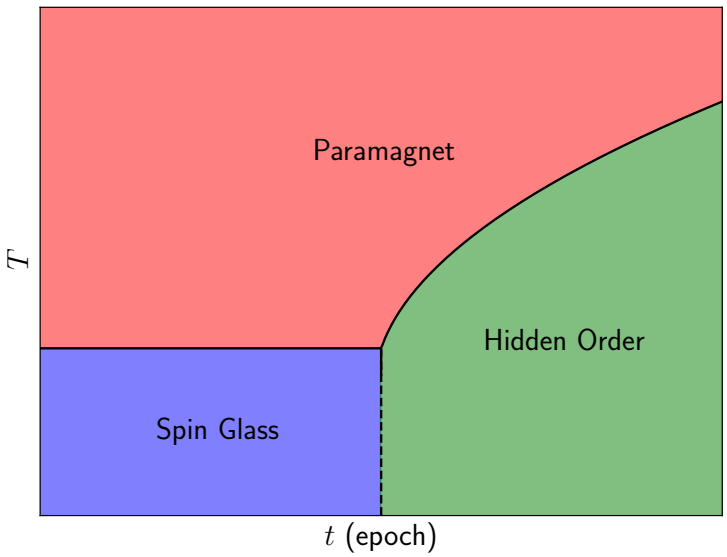



Figure 5.1: A schematic temperature vs. training epoch phase diagram for a multi-layer spin model with bonds determined by neural network training for temperatures not significantly below the melting temperature. The critical temperature grows as a power law with training time.

The remainder of this work is organized as follows. In Section 5.2 we analyze the spin

system corresponding to our NNs before training. In doing so we introduce the multi-layer Sherrington-Kirkpatrick spin glass model and analytically identify the critical temperature at which replica symmetry breaking occurs. We also introduce the TAP equations and describe how they may be used to find the critical temperature for a single instance of a multi-layer spin model with bonds determined by some training process. In Section 5.3 we describe the training procedures for our NNs, discuss their performance, determine how the critical temperature evolves as the training progresses, and discuss the changing nature of the melting transition. We then examine the evolution of the spectrum of the bond matrix, which is the primary driver of the previously discussed changes, and connect to previous work on rich vs. lazy learning. We finish by summarizing our conclusions and future directions in Section 5.4.

## 5.2 The initial state of the neural network

Before any training occurs the weights in our NNs are independently and normally distributed in such as way as to preserve the magnitude of the variance of the activations in the forward pass, as in Kaiming normal initialization [279]. This means that the weights fed into the $k^{\text{th}}$ layer are independently drawn from the normal distribution $\mathcal{N}(0, C_k/N_{k-1})$, where $C_k$ is $O(1)$ and $N_k$ is the number of neurons in the $k^{\text{th}}$ layer. For the sake of generality we allow $C_k$ to be different between layers for now.

Consider an NN with $L+1$ layers labelled $0, \ldots, L$, as shown in Fig. 5.2. By mapping the neurons in the NN onto Ising spins, which can take only the values $\pm 1$, and using the weights as interactions between the spins, we obtain the multi-layer Sherrington-Kirkpatrick

(MSK) model [280] defined by the Hamiltonian

$$H_{\text{MSK}}(\sigma) = -\sum_{k=1}^{L} \sum_{i,j=1}^{N_{k-1},N_k} J_{ij}^{(k)} \sigma_i^{(k-1)} \sigma_j^{(k)}, \tag{5.1}$$

where each $J_{ij}^{(k)}$ is independently drawn from $\mathcal{N}(0, C_k/N_{k-1})$. For large $N_k$ each spin has a high coordination number, so we can expect mean field theory to have good predictive power. MSK is a slight modification of the standard well-known Sherrington-Kirkpatrick (SK) model [12], in which independent and identically distributed interactions exist between every pair of spins in the system, having no layered structure. The MSK model can also be viewed as a specific case of a multi-species spin glass [281].

Later it will become convenient to label the spins with a single index such that

$$\sigma_i^{(k)} = \sigma_{i+\sum_{\ell=0}^{k-1} N_\ell}. \tag{5.2}$$

With this we can write the MSK Hamiltonian in the simpler form

$$H_{\text{MSK}}(\sigma) = -\frac{1}{2} \sum_{i,j=1}^{N} J_{ij} \sigma_i \sigma_j, \tag{5.3}$$

$$J = \begin{pmatrix} 0 & J^{(1)} & 0 & \cdots & 0 \\ J^{(1)T} & 0 & J^{(2)} & \ddots & \vdots \\ 0 & J^{(2)T} & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 & J^{(L)} \\ 0 & \cdots & 0 & J^{(L)T} & 0 \end{pmatrix}, \tag{5.4}$$

where $N = \sum_{k=0}^{L} N_k$.

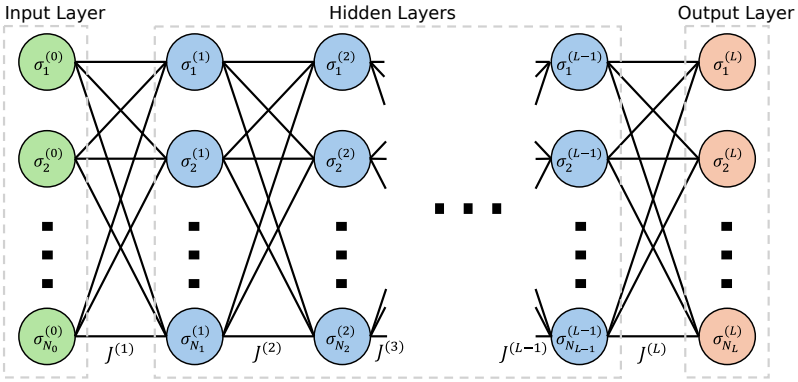



Figure 5.2: Schematic representation of a neural network with $L+1$ layers. The $k^{\text{th}}$ layer contains $N_k$ neurons. $J^{(k)}$ is the weight matrix connecting layer $k-1$ to layer $k$.

### 5.2.1 Transition temperature

For sufficiently high temperatures we reasonably expect the MSK model to be in the paramagnetic phase. That is, the system will behave ergodically. We wish to find the critical temperature at which a non-ergodic phase appears. We expect that, as with the SK model, this low-temperature phase will be a spin glass. This glassy phase is characterized by replica symmetry breaking [13, 15, 16, 267, 268]. If a large number of exact replicas of the system were made with identical realizations of the disorder and all were rapidly cooled, the replicas would settle into a large number of different neighborhoods around local energy minima.

We could make use of the replica "trick" [7, 15] to find the disorder averaged expectation

of some variable $A$ as

$$\langle \mathbb{E}(A) \rangle = \left\langle \frac{1}{Z} \sum_{\sigma_1} A(\sigma_1) e^{-\beta H(\sigma_1)} \right\rangle \tag{5.5}$$

$$= \lim_{m \to 0} \left\langle Z^{m-1} \sum_{\sigma_1} A(\sigma_1) e^{-\beta H(\sigma_1)} \right\rangle \tag{5.6}$$

$$= \lim_{m \to 0} \sum_{\{\sigma_\alpha\}_{\alpha=1}^m} A(\sigma_1) \langle Z^m \rangle, \tag{5.7}$$

where $\beta$ is the inverse temperature, $Z$ is the partition function, and the replicated partition function is

$$Z^m = \exp\left[ -\beta \sum_{\alpha=1}^{m} H(\sigma_\alpha) \right]. \tag{5.8}$$

We are therefore interested in the disorder averaged replicated partition function for small $m$. We search for the temperature at which the replica symmetric solution becomes unstable, indicating the breaking of replica symmetry and the appearance of the glassy phase.

For each layer we define the $m \times m$ overlap matrix $Q_{\alpha\beta}^{(k)} = \frac{1}{N_k} \sum_{i=1}^{N_k} \sigma_{i\alpha}^{(k)} \sigma_{i\beta}^{(k)}$. We enforce these conditions with the Lagrange multipliers $\Lambda_{\alpha\beta}^{(k)}$. We then find after disorder averaging that

$$\langle Z^m \rangle = \int \mathcal{D}Q\mathcal{D}\Lambda \; e^S, \tag{5.9}$$

$$\begin{aligned} S = \frac{1}{2}\beta^2 \sum_{k=1}^{L} N_k \sum_{\alpha\beta} C_k Q_{\alpha\beta}^{(k-1)} Q_{\alpha\beta}^{(k)} \\ + \sum_{k=0}^{L} N_k \left[ \sum_{\alpha\beta} \frac{1}{2} \Lambda_{\alpha\beta}^{(k)} Q_{\alpha\beta}^{(k)} + \log \sum_{\{\sigma_\alpha\}} \exp\left( -\frac{1}{2} \sum_{\alpha\beta} \sigma_\alpha \Lambda_{\alpha\beta}^{(k)} \sigma_\beta \right) \right]. \end{aligned} \tag{5.10}$$

We now make the replica symmetric ansatz

$$Q_{\alpha\beta}^{(k)} = q^{(k)} + \delta_{\alpha\beta}(p^{(k)} - q^{(k)}), \tag{5.11}$$

$$\Lambda_{\alpha\beta}^{(k)} = \lambda^{(k)} + \delta_{\alpha\beta}(\mu^{(k)} - \lambda^{(k)}) \tag{5.12}$$

and expand the action to second order in $q$ and $\lambda$. This yields

$$S/m \approx \frac{\beta^2}{2}\sum_{k=1}^{L} N_k C_k (p^{(k-1)}p^{(k)} - q^{(k-1)}q^{(k)}) + \frac{1}{2}\sum_{k=0}^{L} N_k \left(\mu^{(k)}(p^{(k)} - 1) - \lambda^{(k)}q^{(k)} - \frac{1}{2}(\lambda^{(k)})^2\right), \tag{5.13}$$

ignoring higher order terms in $m$ and constant terms. We will evaluate the integrals over $Q$ and $\Lambda$ by the saddle point method. By differentiating the equation for the action above in terms of $\mu^{(k)}$ and $\lambda^{(k)}$ we find the saddle point equations

$$p^{(k)} = 1 \tag{5.14}$$

$$\lambda^{(k)} = -q^{(k)}. \tag{5.15}$$

With these we can write the action as

$$S/m \approx -\frac{\beta^2}{2}\sum_{k=1}^{L} N_k C_k q^{(k-1)}q^{(k)} + \frac{1}{4}\sum_{k=0}^{L} N_k (q^{(k)})^2. \tag{5.16}$$

The Hessian matrix of the action is then proportional to the tridiagonal matrix

$$\mathbf{H} = \begin{pmatrix} N_0 & -\beta^2 N_1 C_1 & 0 & \cdots & 0 \\ -\beta^2 N_1 C_1 & N_1 & -\beta^2 N_2 C_2 & \ddots & \vdots \\ 0 & -\beta^2 N_2 C_2 & N_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -\beta N_L C_L \\ 0 & \cdots & 0 & -\beta N_L C_L & N_L \end{pmatrix}. \tag{5.17}$$

At sufficiently high temperatures this matrix is positive definite and the replica symmetric solution will be stable. This indicates that the system is in the paramagnetic phase. When the temperature decreases to the point that the Hessian is no longer positive definite, the replica symmetric solution is no longer stable and replica symmetry breaking occurs. This indicates the transition to the glassy phase.

For the simple case in which each layer is identical, letting all $N_k = n$ and $C_k = C$, the Hessian becomes a tridiagonal Toeplitz matrix with the eigenvalues

$$n\left[1 - 2\beta^2 C \cos\left(\frac{(k+1)\pi}{L+2}\right)\right], \quad k = 0, \ldots, L. \tag{5.18}$$

The least eigenvalue vanishes at the transition temperature

$$T_c = \left[2C \cos\left(\frac{\pi}{L+2}\right)\right]^{1/2}. \tag{5.19}$$

If the NN is very deep with large $L$, the transition temperature approaches

$$T_c = \sqrt{2C}. \tag{5.20}$$

### 5.2.2 Thouless-Anderson-Palmer Equations

So far our examination of the multi-layer SK model has made use of disorder averaging. That is, we have been examining an ensemble of systems with independent randomly distributed bonds. However, when we are examining systems in which the bonds are determined by some training procedure, we no longer have such an ensemble to average over. We can instead make use of the Thouless-Anderson-Palmer (TAP) approach [23, 269], which allows us to analyze an energy landscape for a single instance.

The mean field free energy for a system with pairwise interactions between spins is

$$\begin{aligned} F = -\frac{1}{2}\sum_{ij} J_{ij} m_i m_j - \frac{1}{4}\beta \sum_{ij} J_{ij}^2 (1-m_i^2)(1-m_j^2) \\ + \frac{1}{2} T \sum_i \left\{ (1+m_i)\log\left[\frac{1}{2}(1+m_i)\right] + (1-m_i)\log\left[\frac{1}{2}(1-m_i)\right] \right\}, \end{aligned} \tag{5.21}$$

where $m_i = \langle \sigma_i \rangle$. This is the Curie-Weiss free energy [282] with the addition of the second term on the right, known as the Onsager term [283]. This term describes the contribution to the effective field felt by a spin due to the spin itself. Since the effective field should be determined only by the other spins, this contribution is subtracted. By minimizing with respect to the $m_i$ we

obtain the TAP equations

$$\tanh^{-1} m_i = \beta \sum_j J_{ij} m_j - \beta^2 \sum_j J_{ij}^2 (1 - m_j^2) m_i. \tag{5.22}$$

The solutions to these equations are the locations of fixed points of the mean field free energy.

At temperatures not much lower than the transition temperature we can take each $m_i$ to be small, allowing us to linearize the TAP equations to

$$m_i = \beta \sum_j J_{ij} m_j - \beta^2 \sum_j J_{ij}^2 m_i. \tag{5.23}$$

We can also express these as the matrix equation

$$(I_N - M)\mathbf{m} = 0, \tag{5.24}$$

$$M_{ij} = \beta J_{ij} - \delta_{ij} \beta^2 \sum_\ell J_{i\ell}^2. \tag{5.25}$$

Above the transition temperature the only solution is the trivial $\mathbf{m} = 0$, indicating the paramagnetic phase. At the transition temperature $I_N - M$ will cease to be positive definite as at least one eigenvalue vanishes and at least one nontrivial solution appears. Due to the layered structure of

our model, $M$ becomes the block tridiagonal matrix

$$M = \beta \begin{pmatrix} R^{(0)} & J^{(1)} & 0 & \cdots & 0 \\ J^{(1)T} & R^{(1)} & J^{(2)} & \ddots & \vdots \\ 0 & J^{(2)T} & R^{(2)} & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & J^{(L)} \\ 0 & \cdots & 0 & J^{(L)T} & R^{(L)} \end{pmatrix}, \tag{5.26}$$

$$R_{ij}^{(k)} = -\delta_{ij}\beta \begin{cases} \sum_{\ell=1}^{N_1} (J_{i\ell}^{(1)})^2, & k = 0 \\ \sum_{\ell=1}^{N_{k-1}} (J_{\ell i}^{(k)})^2 + \sum_{\ell=1}^{N_{k+1}} (J_{i\ell}^{(k+1)})^2, & 1 \leq k \leq L-1 \\ \sum_{\ell=1}^{N_{L-1}} (J_{\ell i}^{(L)})^2, & k = L \end{cases}. \tag{5.27}$$

For not-too-large systems the transition temperature can be found numerically by exact diagonalization, which is the approach we take in this work. For deep NNs with a large number of layers a more efficient approach is to consider that the block tridiagonal matrix $I_N - M$ will be positive definite as long as all $\chi_k$ satisfying the recurrence relation

$$\chi_k = \begin{cases} \beta R^{(0)} - I_{N_0}, & k = 0 \\ \beta R^{(k)} - I_{N_k} - \beta^2 J^{(k)T} \chi_{k-1}^{-1} J^{(k)}, & 1 \leq k \leq L \end{cases} \tag{5.28}$$

are positive definite [284]. The determinant is then

$$\det I_N - M = \prod_{i=0}^{L} \det \chi_k. \tag{5.29}$$

The transition temperature can then be determined by finding the value of $\beta$ for which the deter-

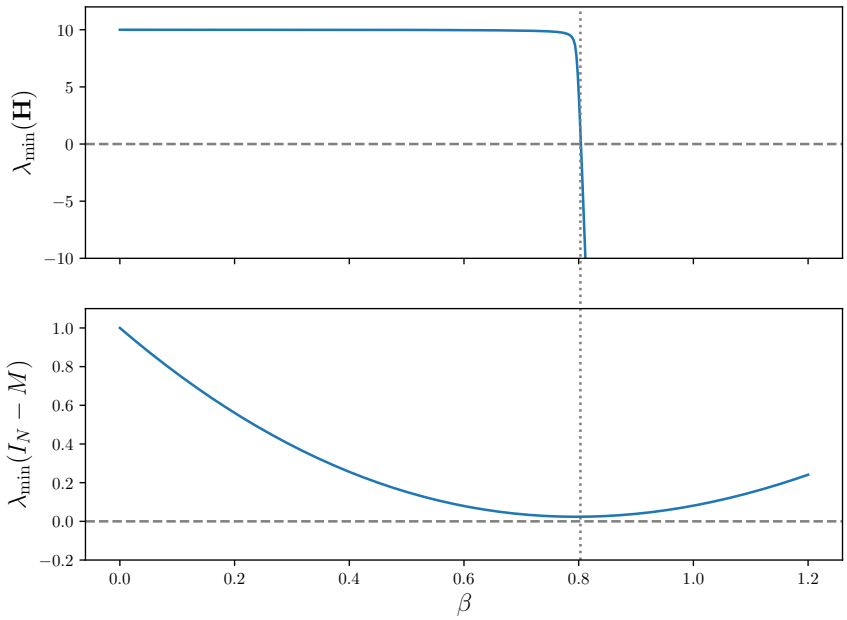



Figure 5.3: (Top) The least eigenvalue of the Hessian of the action [Eq. (5.17)] for the ensemble describing our neural networks before training. (Bottom) The least eigenvalue of $I_N - M$ for a particular system in that ensemble. Both vanish or nearly vanish at the same value of $\beta$.

minant vanishes, meaning that at least one of the $\chi_k$ has a vanishing eigenvalue.

For systems with sufficiently large layers we expect the result from the TAP equations for a single instance to approximate our analytical result for the ensemble. In Fig. 5.3 we show the least eigenvalues of $I_N - M$ for a single system in the ensemble and the Hessian $\mathbf{H}$ of the action for the entire ensemble. We see that the least eigenvalues of both do indeed vanish (or nearly vanish due to finite size effects) at the same temperature. As the system moves to the thermodynamic limit we expect the least eigenvalue of $I_N - M$ to vanish at the single point of its minimum as a function of $\beta$. This indicates that nontrivial TAP solutions appear at the same temperature that replica symmetry breaking occurs in the MSK model.

## 5.3 Evolution with training

### 5.3.1 The training programs

We trained multiple NNs on the MNIST dataset [66]. This dataset is comprised of a $28 \times 28$ grayscale images of handwritten digits 0-9. The training set contains 60,000 images while the test set contains 10,000. The classification task is to label each image with the correct digit. Each of our NNs had three hidden layers containing 512 nodes each.

We analyzed the results of two distinct training programs. The first was designed to closely follow the analogy between neurons in the NN and Ising spins. This was done by training a partially binarized NN. Each neuron in the NN can only take on the values $\pm 1$, but the weights between neurons in adjacent layers are still continuous variables. This is in contrast to NNs which binarize only the weights [285] or fully binarized NNs [286] which binarize both activations and weights. The binarization is enforced by applying the forward activation function

$$\phi_{\text{forward}}(x) = \text{step}(x) = \begin{cases} -1, & x < 0 \\ \\ 1, & x > 0 \end{cases} . \tag{5.30}$$

between each layer of the NN, as well as on the input data. We will refer to this as the binarized NN.

In order to train the binarized NN effectively, we use a modified version of the straight through estimator [287] when calculating the gradients. This amounts to using the hard hyper-

bolic tangent (htanh) activation function

$$\phi_{\text{back}}(x) = \text{htanh}(x) = \begin{cases} -1, & x < -1 \\ x, & -1 < x < 1 \\ 1, & x > 1 \end{cases} \tag{5.31}$$

during the backpropagation step. It has been shown that using such a modified STE in a fully binarized NN allows for near state-of-the-art performance while significantly reducing the computation resources necessary for a forward pass [286]. The loss function for our NN is the $L_2$ multi-class hinge loss [288, 289]. The NN is then trained using stochastic gradient descent (SGD) with momentum.

The second training program is more typical, with neurons able to take on continuous activations. The activation function $\phi(x)$ used in the hidden layers is the popularly used rectified linear unit (ReLU)

$$\phi(x) = \text{ReLU}(x) = \begin{cases} 0, & x \leq 0 \\ x, & x > 0 \end{cases}. \tag{5.32}$$

We will refer to this NN as the standard NN. Once again the loss function is the multi-class hinge loss and training is done by SGD with momentum. In this case the final layer is an $L_2$ support vector machine.

Table 5.1 summarizes the relevant training parameters for both training procedures and the resulting validation error rates. As can be seen in Fig. 5.4, for the larger learning rate, after 60 training epochs the validation error rate of the standard NN reaches approximately 0.019 while the binarized NN reaches approximately 0.027 after 500 epochs. Note that the error rates of both

NNs initally decay as power laws but then eventually plateau.

| NN type | Binarized | | Standard | |
|---|---|---|---|---|
| Forward activation | step | | ReLU | |
| Backward activation | htanh | | ReLU | |
| Learning rate | $10^{-4}$ | $10^{-6}$ | $10^{-4}$ | $10^{-6}$ |
| Training epochs | 500 | 960 | 60 | 960 |
| Validation error rate | 0.0266 | 0.0742 | 0.0189 | 0.0267 |
| Momentum | 0.9 | | | |
| Batch size | 32 | | | |
| # hidden layers | 3 | | | |
| Hidden layer size | 512 | | | |
| Loss function | $L_2$ multi-class hinge | | | |
| Margin | 10 | | | |

Table 5.1: The neural network training details.

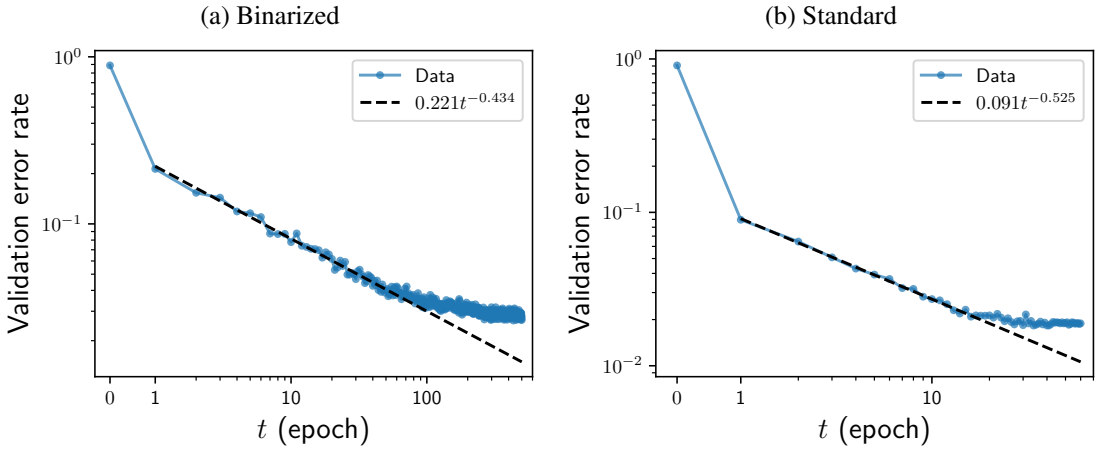



Figure 5.4: The validation error rates as training progresses. All axes are log scaled excepting the linear scaling between 0 and 1 on the horizontal axis. The learning rate for these NNs is $10^{-4}$. Data is collected at the end of each epoch.

### 5.3.2 Evolution of the critical temperature

The weights obtained throughout training allow us to define a family of Hamiltonians for a multi-layer spin system parameterized by the training time $t$:

$$H(\sigma, t) = -\frac{1}{2}\sum_{i,j=1}^{N} J_{ij}(t)\sigma_i\sigma_j. \tag{5.33}$$

A natural question is how the critical temperature of this system changes as training progresses. To accomplish this we determine the least eigenvalue of $I_N - M$ over a range of $\beta$ by exact diagonalization, where $M$ is given in Eqs. (5.26)-(5.27). Fig. 5.5 shows the least eigenvalue vs. $\beta$ before and after training for our two types of NNs. Before training the least eigenvalue vanishes only at its minimum (or nearly does so). As training progresses, the minimum drops and the least eigenvalue first vanishes at a smaller $\beta$. We observe that the least eigenvalue curve drops significantly lower for the standard NN than for the binarized NN. This is likely due to more effective training in the standard NN since the activations are not constrained to be binary and errors are not introduced by having different forward and backward activation functions.

We calculate the transition temperature $T_c$ for each epoch in the training process by finding the value of $\beta$ for which the least eigenvalue curve first vanishes. In order to correct somewhat for finite size effects we find the value of $\beta$ for which the least eigenvalue curve first reaches its minimum value before training. In Fig. 5.6 we see how the critical inverse temperature evolves with training for our two NNs. Before training the critical temperature is near the analytic result obtained in Section 5.2.1. After some transitory behavior at early training times the critical temperature begins growing as a power law. Eventually the power-law behavior of $T_c$ appears to

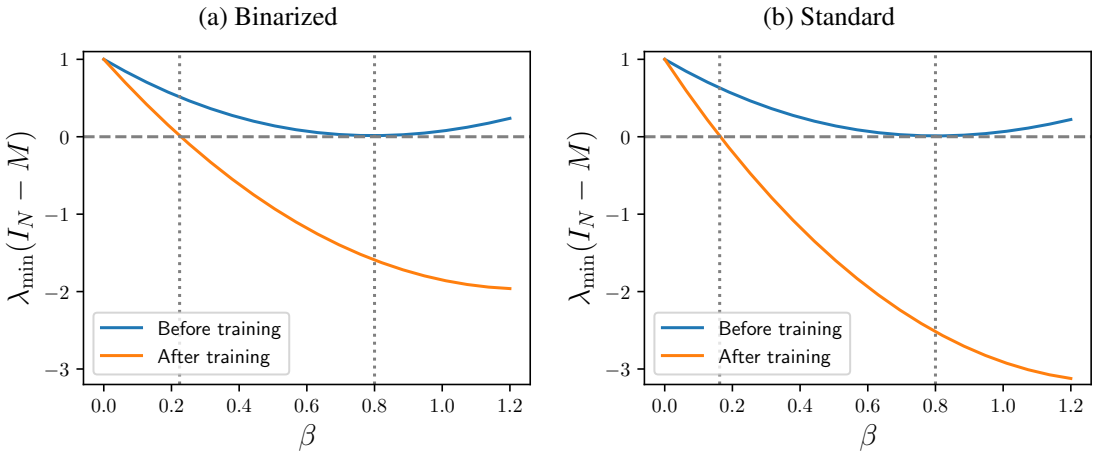



Figure 5.5: The least eigenvalue of $I_N - M$ as a function of inverse temperature $\beta$ before and after training the neural networks. The learning rate for these neural networks is $10^{-4}$.

change to a power-law growth with a smaller exponent. Comparison with Fig. 5.4 shows that this change occurs around the same time at which the validation error rates of the NNs plateau. We surmise that this change in exponents is a result of the NNs reaching their capacity for learning generalizable features of the dataset within their training programs.

### 5.3.3 The nature of the melting transition

In Fig. 5.7 we see the spectrum of $I_N - M$ at $T_c$ before and after training for the two types of NNs at the larger learning rate. As the transition temperature increases with training the spectrum becomes narrower due to the temperature dependence of $M$ seen in Eq. (5.25). More notably, before training the bulk of the spectrum extends all the way to 0 where it has a hard cutoff. This is typical of a glass transition where an extensive number of TAP solutions appear at the transition [290]. After training the spectrum has a small number of eigenvalues that have split off of the bulk and form tails, leading to a much lower spectral density at 0. This suggests that the glassy transition away from the paramagnetic phase that exists before training evolves into a

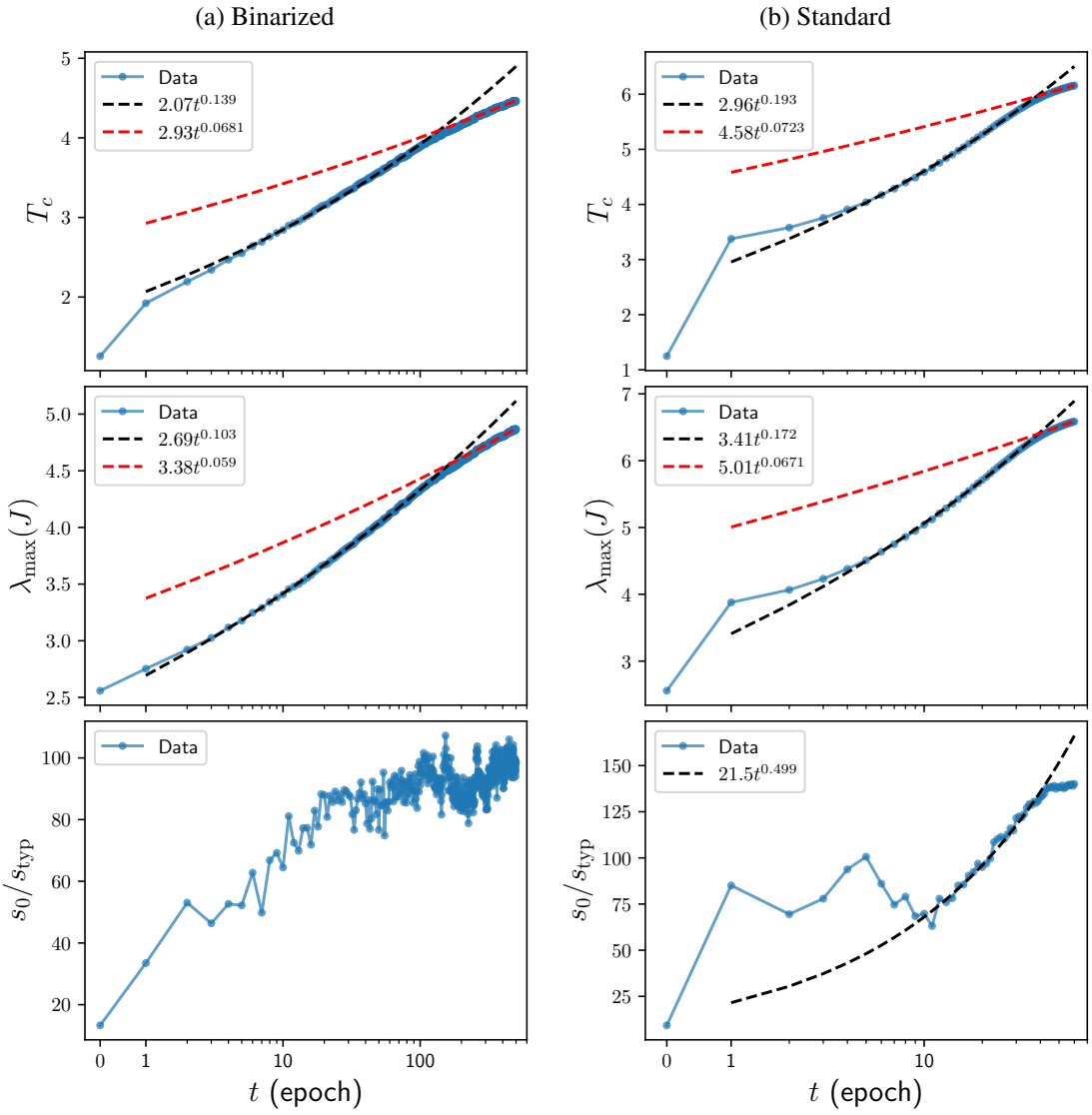



Figure 5.6: From top to bottom, the melting temperature, the largest eigenvalue of the bond matrix $J$, and the normalized first level spacing of $I_N - M$ as training progresses. The learning rate for these neural networks is $10^{-4}$. The horizontal axes are logarithmically scaled except for a linear scaling on the interval [0,1].

different type of transition as training progresses.

In order to further probe this change in the nature of the transition, we track the ratio of the spacing between the smallest eigenvalues $s_0$ to the typical level spacing of the bulk $s_{\text{typ}}$, which we define as the median level spacing. We refer to this ratio as the normalized first level spacing. The evolution of this quantity is shown in Fig. 5.6 for the larger learning rate. For both types of

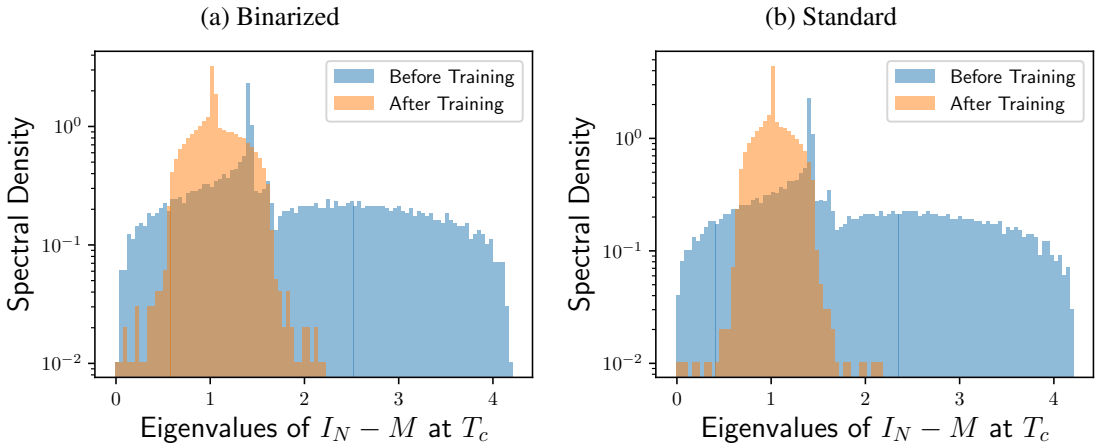



Figure 5.7: The spectrum of $I_N - M$ at the transition temperature before and after training. The learning rate for these neural networks is $10^{-4}$. The vertical axes are log scaled.

NN we observe a large jump in the early training epochs, corresponding with the large increase in $T_c$ at these training times. The standard NN then moves into a period of power-law growth for a time before plateauing. The binarized NN behaves similarly, although no power-law growth for some period of the training is clear.

Since this large jump happens very quickly after training begins, it is difficult to determine the nature of this transition. We find that we can delay the onset of this transition by decreasing the learning rate of our training procedure. Fig 5.8 shows the transition temperature for the two NNs when the learning rate is 100 times smaller than previously. We observe that $T_c$ now stays relatively constant for several epochs before growing. Fig 5.8 also shows the normalized first level spacing for this reduced learning rate. We observe jumps in this quantity at roughly the same times that $T_c$ begins to grow in the two NNs.

Our interpretation of these results is thus. For both NN types the training procedure, which determines the interactions between spins in the multi-layer spin systems, causes eigenvalues to split off from the bulk of the spectrum of $M$. This causes the critical temperature to grow and

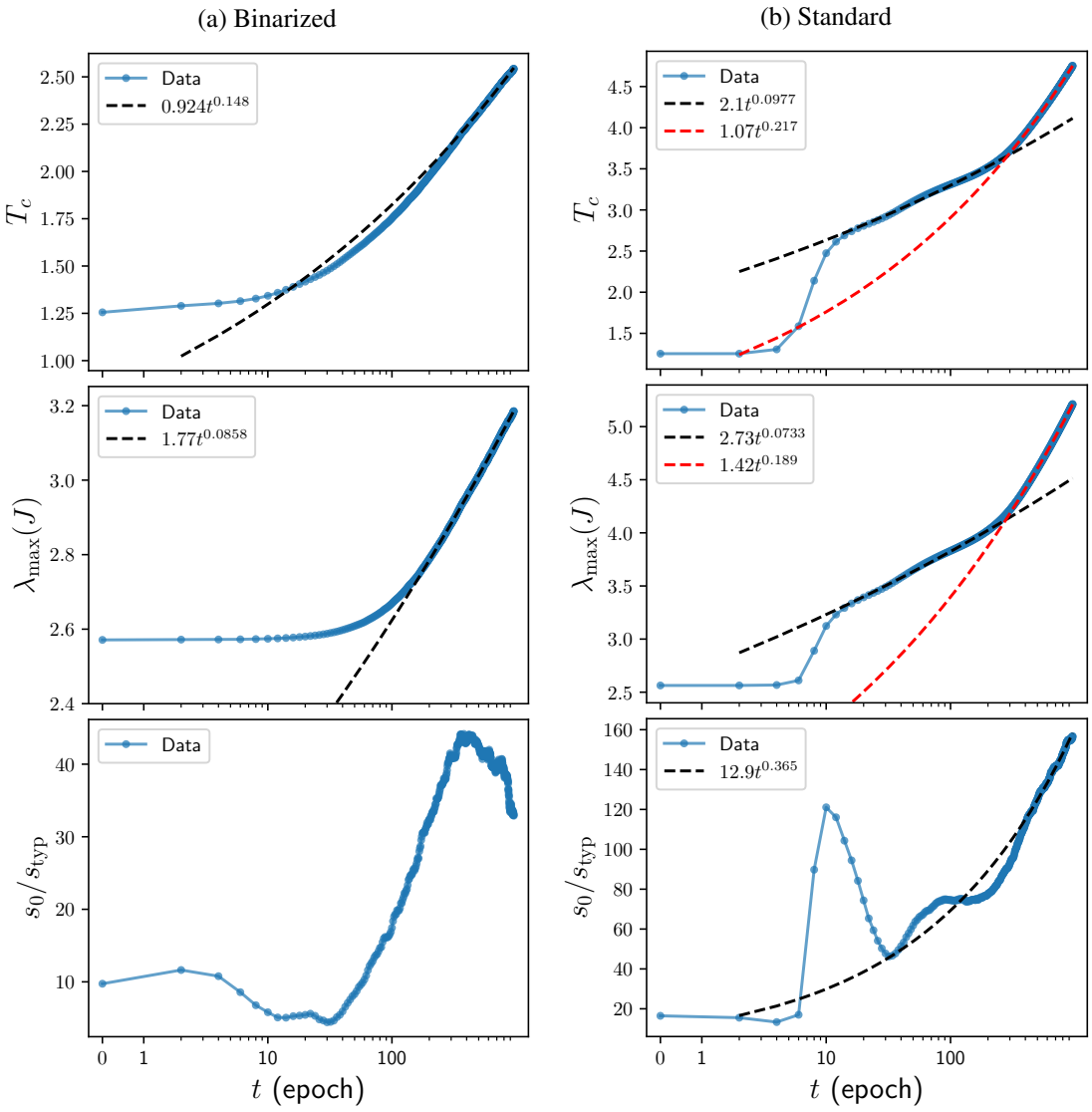



Figure 5.8: From top to bottom, the melting temperature, the largest eigenvalue of the bond matrix $J$, and the normalized first level spacing of $I_N - M$ as training progresses. The learning rate for these neural networks is $10^{-6}$. The horizontal axes are logarithmically scaled except for a linear scaling on the interval [0,1].

greatly reduces the number of TAP solutions which appear at the transition away from the paramagnetic phase, changing the nature of the transition. Instead of a glassy transition at which an extensive number of TAP solutions appear, a single $\mathbb{Z}_2$ symmetry broken solution will appear. This new phase has some hidden order determined by the training data and procedure. Schematically the phase diagram for spin systems with interactions determined by training will look as in

Fig. 5.1.

Note that our approach does not probe temperatures significantly less than the melting temperature. As temperature decreases it is possible for a larger number of TAP solutions to appear in the hidden order phase. However, if the largest eigenvalue of $M$ is greatly isolated from the rest of the spectrum it is likely that there will only be a single TAP solution for lower temperatures, potentially all the way down to $T = 0$ as in a ferromagnet. Such a situation might arise from overtraining the NN.

### 5.3.4 Evolution of the bond matrix spectrum

The transition temperature away from the paramagnetic state is determined by the largest eigenvalue of $M$. We see from Eqs. (5.26)-(5.27) that for high temperatures the largest eigenvalue of $M$ behaves as $\beta\lambda_{\max}(J)$, while for small temperatures it behaves as $-\beta^2\lambda_{\min}(R)$. For the cases we have studied in this work, the transition temperature is within the crossover of these two regimes, making both $\lambda_{\max}(J)$ and $\lambda_{\min}(R)$ quantities of interest. We observed that $\lambda_{\min}(R)$ hardly changed throughout the entire training process. In no case that we examined did it change by more than 1.2% of its original value. For this reason we can understand the training dynamics of the multi-layer spin system primarily through the evolution of the spectrum of the bond matrix $J$, in particular its maximum eigenvalue.

Fig. 5.9 shows the distribution of the eigenvalues of the bond matrix before and after training for the larger learning rate. We see that the bulk of the spectrum maintains the same distribution but a small number of eigenvalues separate from the bulk and form tails. This phenomenon is likely due to the same processes that have an analogous effect on the spectra of the weight

matrices connecting adjacent layers, as studied in other works [275–278]. It has been noted that, as effective training progresses, the spectra of these weight matrices develop tails outside the predictions of Marčenko-Pastur (MP) random matrix theory [291]. Truncated power-law fits to these tails have been used to assess the quality of models without any knowledge of the training or testing data [277].

The formation of these tails has been connected to rich learning [275], as opposed to lazy learning [292–297]. When the NN is in the lazy learning regime the weights change very little and the model performs approximately as its linearization about its initial weights. In this regime the spectra of the weight matrices closely follow the MP prediction with no tails. The appearance of the tails signals that the model is in the rich learning regime which performs significantly better than its linearization, barring overtraining.

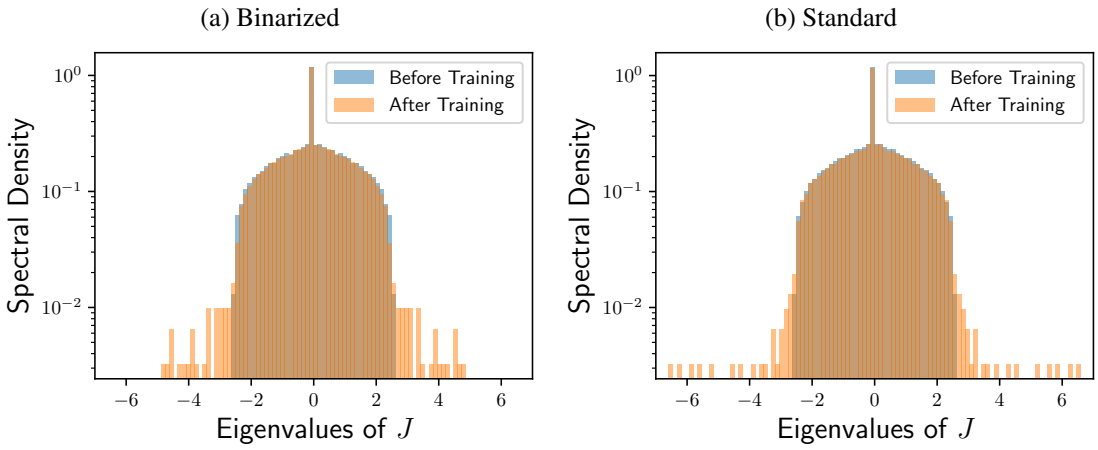



Figure 5.9: The spectral density of the eigenvalues of the bond matrix $J$ before and after training. The learning rate for these neural networks is $10^{-4}$. The vertical axes are log scaled. (Left) Binarized neural networks. (Right) Standard neural networks.

In Figs. 5.6 and 5.8 we see the evolution of the largest eigenvalue of $J$ for the four models we have trained. For the larger learning rate we see that the maximum eigenvalue immediately starts growing and soon behaves as a power-law. This is an indication that the NN has very

quickly moved into the rich learning regime. Eventually the largest eigenvalue starts growing as a different power law with a smaller exponent. The time of these changes correspond to analogous changes in behavior in the transition temperature and the normalized first level spacing of $I_N - M$, as seen in Fig. 5.6. For the smaller learning rate we see that the maximum eigenvalue initially stays constant, but then quickly moves to a power-law growth. This indicates that the NNs remain in the lazy learning regime for a period of training time before moving to the rich learning regime. Again the changes in behavior of the largest eigenvalue of $M$ occur at the same times as the changes in behavior of the transition temperature and the normalized first level spacing of $I_N - M$, as seen in Fig. 5.8.

## 5.4 Conclusion

In this work we have explored a correspondence between NNs and statistical mechanical Ising spin models. We have done this by training two different types of NNs: a standard NN with continuous activations and a partially binarized NN with activations only able to take the values $\pm 1$. At each step in the training process the resulting weight matrices are used to define a bond matrix between spins in a multi-layer spin system analogous to the NN architecture.

We found analytically that, before any training occurs, meaning the weights are all independently and randomly distributed, the multi-layer spin system has a critical temperature separating the ergodic paramagnetic phase at high temperatures from the spin glass phase at low temperatures. Using the TAP equations we numerically determined the transition temperature of the resulting spin system throughout the training. We found that, as training progresses, the transition temperature increases monotonically, exhibiting power-law growth for a time. Additionally,

the spin glass transition at which an extensive number of TAP solutions appear is quickly destroyed in favor of a transition to a phase with a single nontrivial TAP solution determined by the training procedure. This suggests a useful unifying paradigm on the training process: training an NN effectively selects and strengthens a small number of symmetry broken states of the multi-layer spin model, one of which is dominant at the transition.

We note that our approach holds only in the regime in which the TAP equations may be linearized, meaning we do not probe temperatures much lower than the melting temperature. Although numerically determining the number of TAP solutions in this low temperature regime is difficult, it may be possible to analytically calculate the average number of TAP solutions in this regime for bond matrices drawn from a random matrix ensemble whose spectra have tails similar to those we have observed after training. Future work may explore this possibility. It also remains to determine how well our results hold for different architectures, datasets, tasks, etc.

In this study we have used the weights resulting from training an NN as bonds in a classical multi-layer spin system. An interesting alternative would be to consider instead a quantum spin system and observe what differences may arise in comparison to the classical case. Such an examination may shed light on the potential of quantum machine learning [**schuld2014**, 298] and would be directly relevant to recent experimental advances in analogue neuromorphic computing in cavity-QED and Rydberg systems [272–274]. We plan to explore this avenue in a future work.

## Acknowledgements

V.G. is grateful to Ben Lev for useful discussions. This research was sponsored by the Army Research Office under Grant Number W911NF-23-1-0241, the National Science Founda-

tion under Grant No. DMR-203715, and the NSF QLCI grant OMA-2120757. M.W. acknowledges support from the Joint Quantum Institute.

# Chapter 6: Natural Quantization of Neural Networks

**Authors:** Richard Barney, Djamil Lakhdar-Hamina, Victor Galitski

**Abstract:** We propose a natural quantization of a standard neural network, where the neurons correspond to qubits and the activation functions are implemented via quantum gates and measurements. The simplest quantized neural network corresponds to applying single-qubit rotations, with the rotation angles being dependent on the weights and measurement outcomes of the previous layer. This realization has the advantage of being smoothly tunable from the purely classical limit with no quantum uncertainty (thereby reproducing the classical neural network exactly) to a quantum case, where superpositions introduce an intrinsic uncertainty in the network. We benchmark this architecture on a subset of the standard MNIST dataset and find a regime of "quantum advantage," where the validation error rate in the quantum realization is smaller than that in the classical model. We also consider another approach where quantumness is introduced via weak measurements of ancilla qubits entangled with the neuron qubits. This quantum neural network also allows for smooth tuning of the degree of quantumness by controlling an entanglement angle, $g$, with $g = \frac{\pi}{2}$ replicating the classical regime. We find that validation error is also minimized within the quantum regime in this approach. We also observe a quantum transition, with sharp loss of the quantum network's ability to learn at a critical point $g_c$. The proposed quantum neural networks are readily realizable in present-day quantum computers on

commercial datasets.

## 6.1 Introduction

Recent advances in the fields of quantum computing and classical machine learning have fueled the growth of a new area of study examining the intersection of these two approaches to computation—quantum machine learning. Classical artificial neural networks (NNs) have demonstrated remarkable success across a wide range of tasks once thought to require human intelligence, such as complex classification tasks [299, 300], natural language processing [301–303], protein structure prediction [304], and materials data science [305, 306]. At the same time, steady progress has been made in quantum computation [307], which aims to take advantage of the fundamentally quantum aspects of nature, such as superposition of states, interference, and entanglement, using them as resources in computational tasks [308, 309].

Naturally, there is interest in combining these two domains. Artificial neural network models that utilize quantum resources are called quantum neural networks (QNNs), and they have received significant attention in recent years [310–316]. Several formulations of QNNs have been proposed, such as parameterized quantum circuits [62–65], quantum Hopfield networks [51–53], quantum Boltzmann machines [54–56], quantum perceptrons [317–320], and quantum convolutional neural networks [57–61]. Many of these have demonstrated significant learning capabilities. Recent advances in quantum hardware, such as trapped ions [321–323], superconducting qubits [234, 324, 325], and photonic quantum computing [326–328] have further fueled interest in practical implementations of QNNs.

However, the emergence of any quantum advantage by comparison with classical NNs is

often not straightforward. With the limited capabilities of current quantum computers, it is not possible to compare a large classical NN with a similarly large QNN. Differences in activation functions, loss functions, and optimization algorithms are additional confounding factors. In this paper we examine, through classical simulation, the advantages of using quantum resources in neural networks by studying two natural quantization of a standard multi-layer perceptron network [44]. These approaches allow us to tune the network continuously between classical and quantum modes of operation, allowing us to determine the benefit of quantum effects without making significant changes to the network's architecture.

## 6.2 The classical neural network

The classical NN we will quantize is a binarized multi-layer perceptron [329], as shown in Fig. 6.1. The input layer has $M$ neurons. In a forward pass through the network these neurons will be initialized to the elements of a data vector $\mathbf{d}^0$. The network then contains $L$ hidden layers, each containing $N$ neurons, followed by an output layer of size $P$. The network is fully connected, each neuron is connected to all neurons in adjacent layers. The activation of the $i^{\text{th}}$ neuron in the $k^{\text{th}}$ hidden layer is

$$d_i^k = \phi\left(\sum_j W_{ij}^{k-1} d_j^{k-1}\right), \tag{6.1}$$

$$\phi(x) = \text{sgn}(x), \tag{6.2}$$

where the $W^k$ are matrices of learnable weights and $\phi(x)$ is the activation function. The network is binarized in the sense that the activations of the hidden layers can only be $\pm 1$. This is in contrast

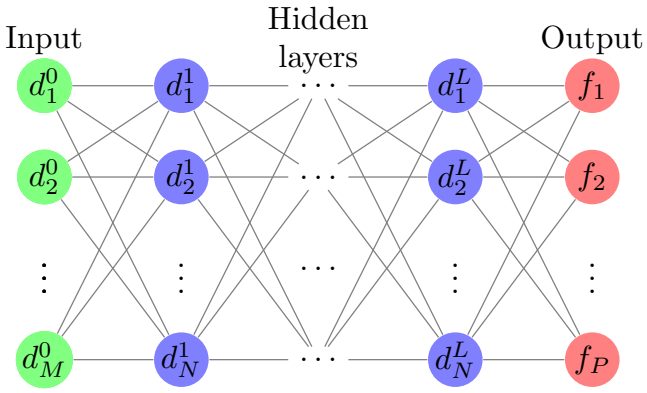



Figure 6.1: The classical binarized multi-layer perceptron network. It consists of an input layer of size $M$, $L$ hidden layers, each of size $N$, and an output layer.

to NNs which binarize only the weights [285] or fully binarized NNs [286] which binarize both activations and weights. A similar binarized network was previously studied by two of the present authors in the context of spin model analogues [4]. The output vector is then computed as $\mathbf{f} = W^L \mathbf{d}^L$. This output vector is then the argument of some loss function that estimates the network's performance on some task.

For this binarized network, a slight change needs to be made the usual backpropagation approach of estimating the gradient of the loss function with respect to the weights [46, 47]. This is necessary because the first derivative of the activation function is either 0 or undefined everywhere. This can be overcome by using a clipped straight-through estimator for the gradient [286, 287]. This amounts to treating the activation function as

$$\phi_{\text{back}}(x) = \text{htanh}(x) = \begin{cases} x, & |x| \leq 1 \\ \text{sgn}(x), & |x| \geq 1 \end{cases} \tag{6.3}$$

when calculating the gradients.

We evaluate the performance of this classical network on the task of classifying images

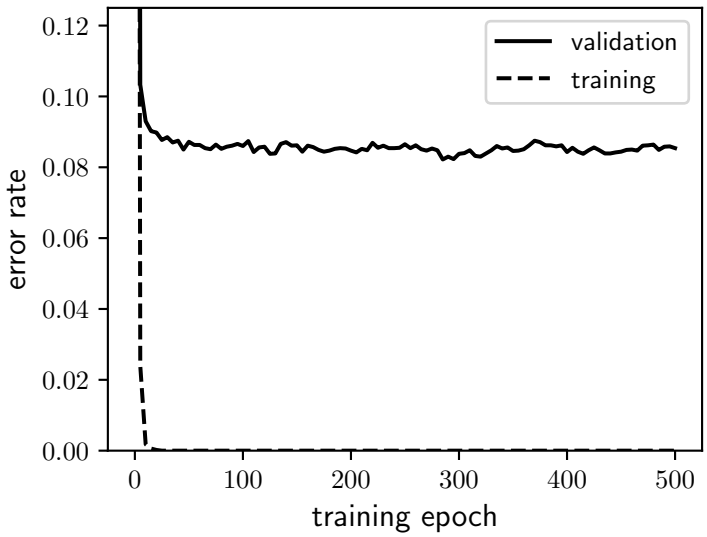



Figure 6.2: The validation and training error rates for the classical network as training progresses. The training error rate quickly vanishes while the validation error rate bottoms out at a nonzero value. This is an indication of overfitting.

from the MNIST dataset [66], which contains images of handwritten digits 0-9. The full MNIST dataset contains 60,000 training images and 10,000 images for validation. In Fig. 6.2 the training curves are shown for the classical network when trained on a subset of 5,000 MNIST training images. The relevant hyperparameters for this network are shown in Table 6.1. We see that, for this training set, overfitting [330] is very apparent in the network. The training error rate quickly vanishes, while the validation error rate bottoms out. This indicates that the model has become trapped in a minimum of the loss function that is unfavorable for generalization, a problem that is common when training on small datasets. In the remainder of this text we will examine how quantization may ameliorate this problem and regularize the network.

| # hidden layers | 3 |
|---|---|
| Hidden layer size | 512 |
| Optimizer | Stochastic gradient descent |
| Loss function | Cross-entropy |
| Learning rate | 0.01 |
| Momentum | 0.9 |
| Batch size | 64 |
| Training epochs | 500 |
| Training dataset size | 5000 |
| Validation dataset size | 10000 |

Table 6.1: The hyperparameters for the training and testing of the classical and quantum networks.

## 6.3 The quantized neural network

The quantized version of the classical NN shown in Fig. 6.1 is implemented by replacing the hidden layers with a quantum circuit, as shown in Fig. 6.3. The circuit is comprised of $N$ qubits and a set of classical channels. Initially, each of the qubits is set to the $|0\rangle$ state. The classical channels initially hold the input data vector $\mathbf{d}^0$. At each step in the forward pass, the $N$ qubits are individually rotated about the $y$-axis and then measured, with the measurement outcomes for the $k^{\text{th}}$ layer being the activations $\mathbf{d}^k$. The single-qubit rotation operator is

$$R_Y(\theta) = e^{-i\frac{\theta}{2}y}, \tag{6.4}$$

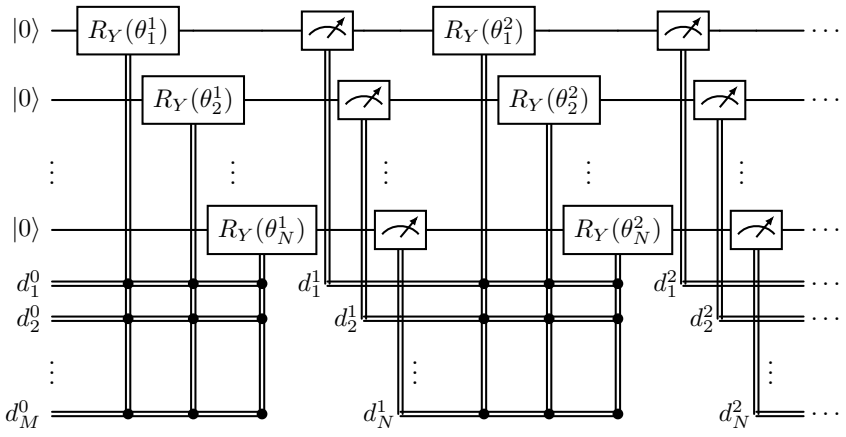



Figure 6.3: The quantum circuit implementing the hidden layers of the quantized network. Each step of the forward pass consists of rotating each qubit by an angle controlled by the classical channels, then measuring each qubit. These measurements are the activations which are then passed to the next layer of the network.

with $y$ being a Pauli operator. The rotation angle for the $i^{\text{th}}$ qubit in the $k^{\text{th}}$ layer is

$$\theta_i^k = \frac{\pi}{2} \begin{cases} 1 - \phi_a\left(\sum_j W_{ij}^{k-1} d_j^{k-1}\right), & k = 1 \\ d_i^{k-1} - \phi_a\left(\sum_j W_{ij}^{k-1} d_j^{k-1}\right), & 2 \le k \le L \end{cases}, \tag{6.5}$$

$$\phi_a(x) = \text{htanh}(x/a) = \begin{cases} x/a, & |x| \le a \\ \text{sgn}(x), & |x| \ge a \end{cases}, \tag{6.6}$$

where $\phi_a(x)$ is the network's activation function. The outputs $\mathbf{f}$ of the network are then calculated as in the classical network.

One can see that, when the limit $a \to 0$ is taken, the activation function reduces to $\phi_0(x) = \text{sgn}(x)$, as in the classical binarized network. In fact, in this limit, the rotation angles all become 0 or $\pm\pi$. One can verify by inspection that, if the output of the activation function is 1, the rotation operator will set the qubit to the $|0\rangle$ state. On the other hand, if the activation is $-1$, the qubit will be set to $|1\rangle$. Since the circuit only moves the qubits between computational basis states

corresponding to the activations of the classical network, classical behavior is recovered in this limit.

Naturally, we are interested to see how model performance changes when tuning the parameter $a$ to nonzero values. Increasing $a$ causes the activation function to be "stretched." This causes each rotation angle to no longer be constrained to be 0 or $\pm\pi$ when the magnitude of its preactivation is less than $a$, meaning that the rotations move the qubits into superpositions of the computational basis states. The outcomes of the projective measurements then become stochastic. In this way, the quantum resource of superposition is used to inject stochasticity into the network. Stochasticity is known to provide regularization in artificial NNs [287, 331–333].

At this point, the question remains of how to best perform inference on the validation dataset using the weights obtained by training the quantum NN. There are two natural ways to do so. The first is to take the weights obtained and run inference deterministically. That is, set $a = 0$ to recover classical behavior for the inference step. The second option we explore is to maintain the same value of $a$ for inference, but to pass each test datum through the network multiple times and select the most common prediction. Fig. 6.4 shows how these two approaches compare for a particular trained network. We see that it only requires a few shots for quantum inference to outperform deterministic inference, though there are diminishing returns when taking more than approximately 10 shots. The validation error rates in the remainder of this text are computed by selecting the most common prediction from 15 shots for each image in the validation dataset.

In Fig. 6.5(a) we see the training curves showing the validation error rate as training progresses for some selected values of $a$. We observe that the quantized network reaches significantly lower error rates than the classical network for some nonzero values of $a$. In Fig. 6.5(b) we see the final validation error rates as a function of $a$. Our best result is obtained for $a = 10^{-1/2} \approx 0.316$,

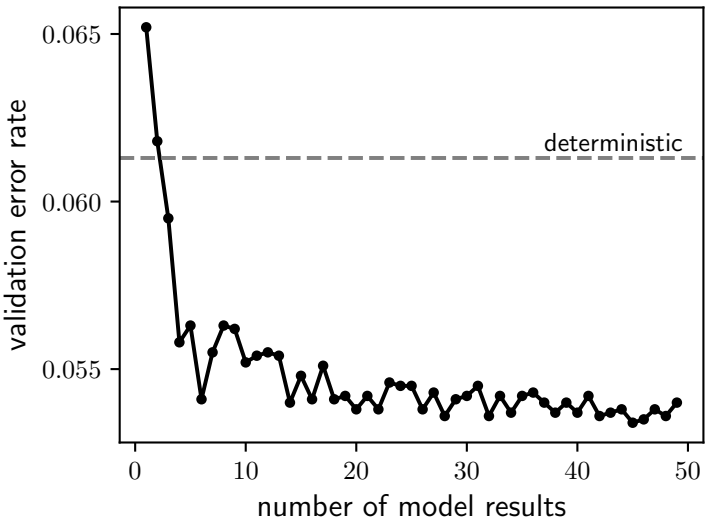
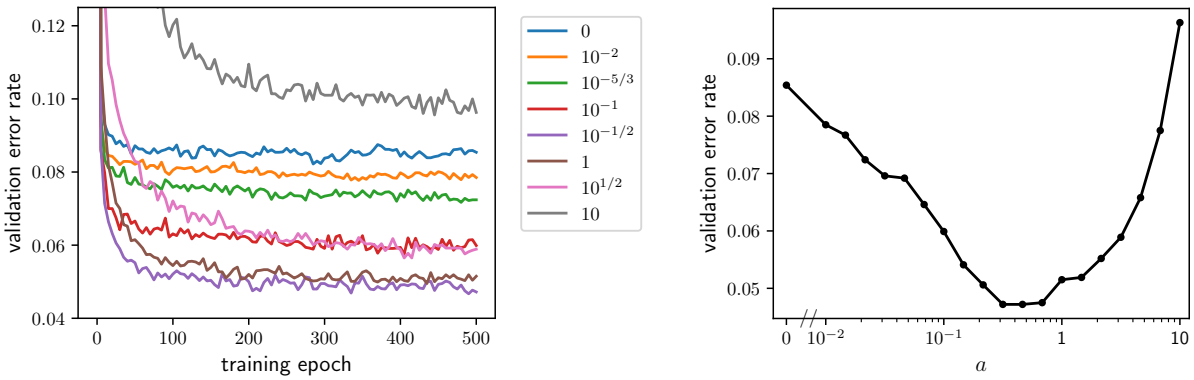



Figure 6.4: The validation error rate for a quantum network ($a = 0.5$) after training as a function of the number of model results used to classify each image in the MNIST validation dataset.

which achieves an error rate of 0.0472. Increasing $a$ further actually reduces performance until the quantum network performs worse than the classical one. This is expected, since introducing too much randomness to the network will inevitably impede its ability to learn.

We can understand the increase in performance for nonzero $a$ in terms of quantum tunneling. The objective of the learning algorithm is to descend down the landscape of the loss function to find a minimum. Doing this purely deterministically can cause the model to become trapped in a local minimum that leads to good performance on the training set, but has poor generalization performance on the validation dataset. This issue becomes more common when training on small datasets. By adding quantum randomness to the model, we make it easier for the model to tunnel out of an unfavorable local minimum and find a more favorable minimum that leads to better generalization.




(a) (b)

Figure 6.5: (a) Validation error rate curves as training progresses in the quantized network for selected values of $a$. (b) The validation error rate as a function of $a$. The best result is achieved for nonzero $a$, indicating that quantum effects improve performance.

## 6.4 Quantization by weak measurements

In addition to the quantization describe above, we examined another method of quantization by way of weak measurements. This may be done by setting $a = 0$ in the previous treatment and replacing the projective mid-circuit measurements with weak measurements. Each of these weak measurements is implemented by entangling each neuron qubit with an ancilla qubit that is prepared in an equal superposition of the computational basis states, and then projectively measuring the ancilla qubit. The entanglement is effected by the two-qubit rotation gate

$$R_{ZY}(-g) = e^{i\frac{g}{2}z\otimes y_{\text{anc}}}, \tag{6.7}$$

where $g$ is a parameter that controls the strength of the entanglement. Fig. 6.6 shows the circuit diagram for this weak measurement protocol.

If the neuron qubit is initially in the state $|\psi_0\rangle = \alpha|0\rangle + \beta|1\rangle$, then, after the entangling

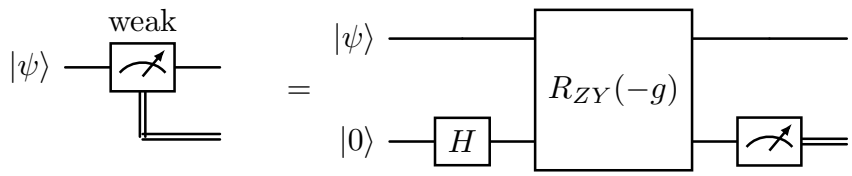



Figure 6.6: The quantum circuit that implements a weak measurement of a neuron qubit by entangling with and then measuring an ancilla qubit.

gate, the two-qubit system will be in the state

$$|\Psi\rangle = \sqrt{(1+\sin g)/2}(\alpha|00\rangle + \beta|11\rangle) + \sqrt{(1-\sin g)/2}(\alpha|01\rangle + \beta|10\rangle). \tag{6.8}$$

We need only consider $g$ within the interval $[0, \frac{\pi}{2}]$. When $g = 0$, the two-qubit system can be factored into a product state; no entanglement is created. On the other hand, when $g = \frac{\pi}{2}$, the system is maximally entangled; the ancilla measurement effectively projectively measures the neuron qubit, recovering classical behavior. Measuring the ancilla qubit will, in general, disturb the neuron qubit. If the measurement outcome is $d$, the state of the neuron qubit will become

$$|\psi_f\rangle = \alpha\sqrt{\frac{1+d\sin g}{1+d\langle z\rangle \sin g}}|0\rangle + \beta\sqrt{\frac{1-d\sin g}{1+d\langle z\rangle \sin g}}|1\rangle, \tag{6.9}$$

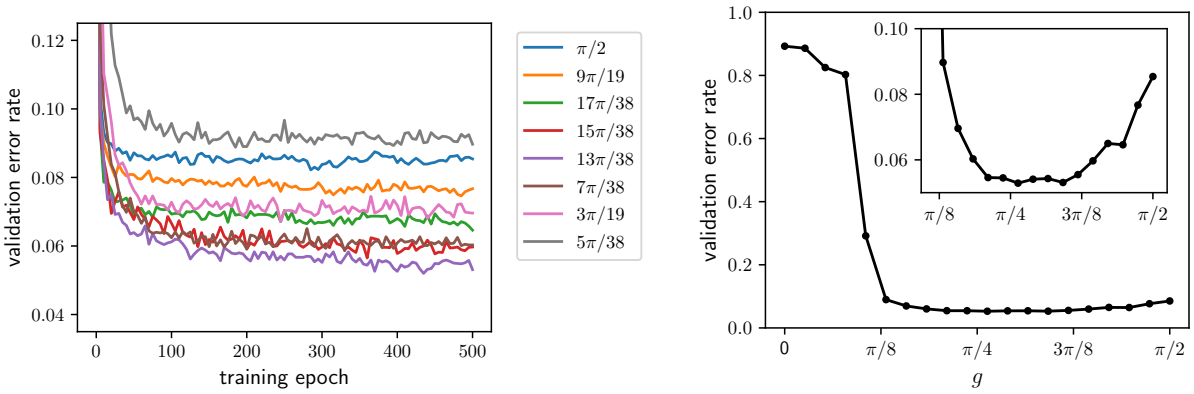
where $\langle z\rangle = |\alpha|^2 - |\beta|^2$ is the expectation value of a $z$ measurement of the neuron qubit before the ancilla measurement.

Since the activations of the network are now determined by measuring the ancilla qubits, each activation has the probability $\frac{1}{2}(1-\sin g)$ to be "wrong" in comparison to what the classical activation would be. This approach provides an alternative way to introduce stochasticity to the network. The nature of the stochasticity introduced by weak measurements is different than the previous approach because each activation has the same probability of being "wrong." In contrast,

the previous approach only has a nonzero probability for an activation to differ from the classical activation when the preactivation has a magnitude less than $a$. In this way, neurons with large preactivations, indicating more certainty, are unaffected by quantum randomness. The advantage of the weak measurement approach is that quantum coherence is not fully destroyed at each step in the forward pass.

In Fig. 6.7(a) we see the training curves showing the validation error rate as training progresses for selected values of $g$. Similar to our previous approach, we see that convergence time is mostly unaffected until $g$ becomes sufficiently small. Before that point, there are values of $g$ for which the model performs significantly better than the classical network. In Fig. 6.7(b) we see the final validation error rates as a function of $g$. We see that, for $g$ less than approximately $\frac{\pi}{8}$, the network quickly loses its ability to learn at all. Our best result was obtained when $g = \frac{5\pi}{19}$, which achieved a validation error rate of 0.0529. This is slightly worse than the best performing model using our previous approach. The lack of discrimination between neurons with large or small preactivations in the weak measurement approach is one possible explanation for this, though the difference is small enough that its significance is uncertain.

It is also natural to examine model performance when our two approaches are combined. That is, to quantize the model by varying both $a$ and $g$. The results of this combined approach are shown in Fig. 6.8. The classical result is shown in the top left corner, where $a = 0$ and $g = \frac{\pi}{2}$. We see that performance is improved by increasing $a$, decreasing $g$, or both simultaneously, up to a point. The best performance we observed was when $a = 10^{-1/3} \approx 0.464$ and $g = \frac{9\pi}{19}$, which achieved a validation error rate of 0.0463, though performance does not vary much along the curved valley floor shown in Fig. 6.8.




(a) (b)

Figure 6.7: (a) Validation error rate curves as training progresses in the quantized network with weak measurements for selected values of $g$. (b) The validation error rate as a function of $g$. The inset is the same but zoomed in near the minimum of the curve. The best result is achieved for $g < \frac{\pi}{2}$, indicating that quantum effects improve performance.

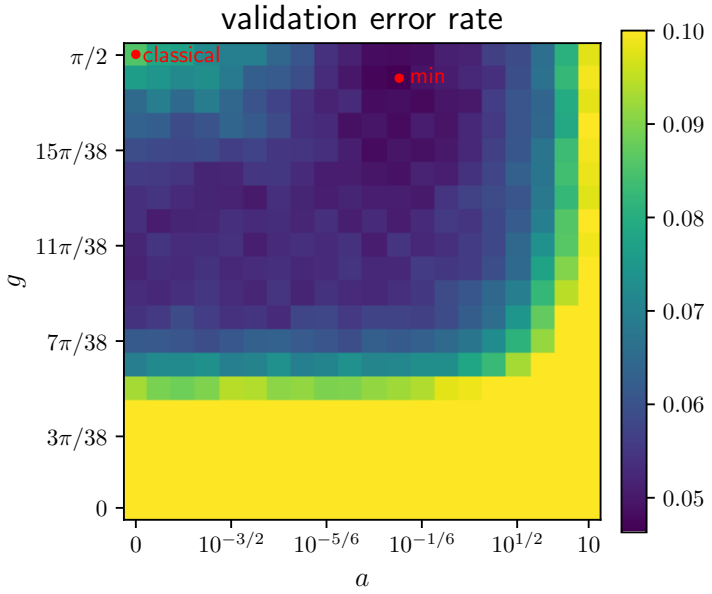



Figure 6.8: Validation error rates results for networks making use of both quantization approaches, varying both $a$ and $g$.

## 6.5 Conclusion

In this paper we explored two approaches to quantizing a binarized multi-layer perceptron model. The first approach uses single qubit rotations to create quantum superpositions for neurons with sufficiently small pre-activations. Each qubit is then projectively measured, leading to stochastic activations for these neurons. The second approach entangles the neuron qubits with ancilla qubits. The activations are then found by projectively measuring the ancilla qubits. This approach is independent of the pre-activation magnitudes, but has the advantage of not completely destroying coherence at each step in the circuit. Both approaches use quantum resources to introduce stochasticity to the network, and both allow for continuous tuning between classical and more quantum modes of operation. We find that, when training on a reduced MNIST dataset, introducing quantum stochasticity leads to a significant reduction in the validation error rate. This improvement can be understood as an enhancement in the model's ability to tunnel away from an unfavorable local minimum of the loss function to a more favorable minimum which leads to better generalization.

In this work we were not primarily concerned with creating state-of-the-art QNNs, instead seeking to understand how performance changes as we move continuously between classical and quantum modes of operation without changing the model architecture. We consider the quantized networks we examined to be a starting point in exploring how quantum resources can be utilized to improve model performance. There are many possible quantum advantages that remain to be explored within our framework, such as representing data more efficiently with multi-qubit superposition states, quantum algorithms for optimizing the loss function, and exploring connections to measurement induced phase transitions [334]. Thus far, we have only

explored the behavior of these networks when trained on classical data. It would be interesting to see what advantages they might have when being trained on data collected from experiments on physical quantum systems. Currently, we are exploring generalizations of the architectures in this paper which entangle the neuron qubits through multi-qubit gates. Additionally, we are working towards implementing our current architectures on actual quantum hardware, extending the results of this paper beyond classical simulation.

## Acknowledgments

We thank Norbert M. Linke and Alaina M. Green for helpful conversations on the topic of practical quantum computation. This research was sponsored by the U.S. Army Research Office under Grant Number W911NF-23-1-0241 and the National Science Foundation under Grant No. DMR-203715.

# Chapter 7: Benchmarking a Tunable Quantum Neural Network on Trapped-Ion and Superconducting Hardware

**Authors:** Djamil Lakhdar-Hamina, Xingxin Liu, Richard Barney, Sarah H. Miller, Alaina M. Green, Norbert M. Linke, Victor Galitski

**Abstract:** We implement a quantum generalization of a neural network on trapped-ion and IBM superconducting quantum computers to classify MNIST images, a common benchmark in computer vision. The network feedforward involves qubit rotations whose angles depend on the results of measurements in the previous layer. The network is trained via simulation, but inference is performed experimentally on quantum hardware. The classical-to-quantum correspondence is controlled by an interpolation parameter, which is zero in the classical limit. Increasing it introduces quantum uncertainty into the measurements, which is shown to improve network performance at moderate values of the interpolation parameter. We then focus on particular images that fail to be classified by a classical neural network but are detected correctly in the quantum network. For such borderline cases, we observe strong deviations from the simulated behavior. We attribute this to physical noise, which causes the output to fluctuate between nearby minima of the classification energy landscape. Such strong sensitivity to physical noise is absent for clear images. We further benchmark physical noise by inserting additional single-qubit and two-qubit gate pairs into the neural network circuits. Our work provides a springboard toward

more complex quantum neural networks on current devices: while the approach is rooted in standard classical machine learning, scaling up such networks may prove classically non-simulable and could offer a route to near-term quantum advantage.

## 7.1 Introduction

There exists a deep analogy between neural networks and spin models in statistical mechanics [4, 39, 256, 257]. In this correspondence, neurons and weights are analogous to classical spins and couplings, respectively. It is natural to quantize these by promoting them to qubits and gates, respectively [5]. This gives rise to a broad class of quantum neural networks. Much recent work in the field of quantum machine learning has focused on searching for "quantum advantage" [298, 312, 335–341]. A variety of network architectures have been discussed in this context, and while some have been simulated under a variety of assumptions, few have been experimentally run on genuine quantum hardware [313, 342, 343], let alone in the context of benchmarking on, *e.g.*, standard image classification tasks [344]. No matter the implementation, such architectures typically do not allow for smooth interpolation between classical and quantum modes, making comparisons between classical neural networks (NN) and quantum neural networks (QNN) difficult.

In the following, we present an architecture for a quantum neural network, which we dub the Benchmark Quantum Neural Network (BQNN), that has several attractive features: (1) It can be adiabatically tuned between the classical and quantum regime. (2) The circuit is deployable on any functional quantum computing platform with minimal demands on qubit count, gate fidelity, and connectivity. (3) Because the classical limit of BQNN reproduces a binary neural network,

it can be benchmarked on any reference data set, enabling direct classical-quantum comparisons. (4) BQNN is generalizable to more complex hardware set-ups, including those not classically simulable [345, 346], by introducing partial mid-circuit measurements and feedback, bridging the field of measurement-induced phase transitions [347] with quantum neural networks.

After reviewing the theory of this BQNN architecture, we present image classification results on both superconducting and trapped-ion hardware. The latter involves two distinct implementations involving gates driven by microwave and laser interactions as discussed below. The experimental results are largely consistent with the expected behavior informed by classical simulations. In particular, in an intermediate quantum regime where quantum uncertainty in measurements is present but does not fully randomize measurements, the BQNN implemented on real devices outperform the classical counterpart. We also examine how device noise affects performance and find that moderate levels of physical noise do not hinder inference. We even find instances where intentional introduction of noise improves performance, up to a limit. Strong gate noise, modeled here via insertion of pairs of single-qubit and two-qubit gates and their inverse, eventually degrades performance to the level of a random guess. The number of noise gates it takes to see a decrease in performance provides another neural-network-specific metric to characterize gate noise. To our knowledge, this study is the first to compare different hardware implementations of QNNs for benchmarking image classification.

## 7.2 The Network Architecture

Ref. [5] presents theory and simulation of a quantum partially binarized multilayer perceptron i.e. the BQNN. See Fig. 7.1a. This network may be smoothly tuned between classical and

quantum modes of operation. This network is partially binarized in the sense that, while each training weight can take on any real value, the activations of the hidden layers are constrained to be $\pm 1$. For a classical network this is done by using the activation function

$$\phi(x) = \text{sgn}(x). \tag{7.1}$$

The activation of the $i^{\text{th}}$ neuron in the $k^{\text{th}}$ layer is

$$d_i^k = \phi\left((W^k \mathbf{d}^{k-1})_i\right), \tag{7.2}$$

where $\mathbf{d}^0 = \mathbf{I}$ is the input vector and each $W^k$ is a matrix of trainable parameters. In this way, information travels through the network layer by layer, in a process called feedforward. The output of the network with $\ell - 1$ hidden layers is

$$\mathbf{o} = W^\ell \mathbf{d}^{\ell-1}, \tag{7.3}$$

which is then the argument of the loss function during training.

This classical network is then quantized in a straightforward way: binary neurons are promoted to qubits and activation functions are implemented by single-qubit rotations followed by projective measurements. By defining the activations as the results of projective measurements, the activations remain binarized. The circuit which implements the hidden layers is shown in Fig. 7.1a.

To tune between quantum and classical regimes, we define the rotation angle of qubit $i$ in

layer $k$ as

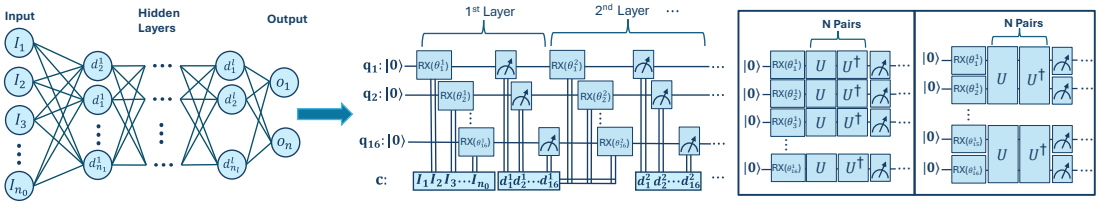
$$\theta_i^k = \frac{\pi}{2} \begin{cases} 1 - \phi_a \left( (W^1 \mathbf{I})_i \right), & k = 1 \\ d_i^{k-1} - \phi_a \left( (W^k \mathbf{d}^{(k-1)})_i \right), & k > 1 \end{cases}, \tag{7.4}$$

$$\phi_a(x) = \text{htanh}(x/a) = \begin{cases} \text{sgn}(x), & |x| \geq a \\ x/a, & |x| < a \end{cases}, \tag{7.5}$$

where $\phi_a$ may be considered a quantum activation function. In the limit $a \to 0$, $\phi_a$ reduces to the classical activation function Eq. 7.1. In this case, each rotation angle is either 0 or $\pm\pi$. This means that, under ideal conditions, the rotation gates will only move the qubits between computational basis states, the behavior of the $a = 0$ network is deterministic, and is equivalent to the standard classical neural net. Non-zero values of $a$ enable arbitrary angels of rotation and hence non-deterministic measurements. We identify $a$ as a parameter describing the "quantumness" of the network.

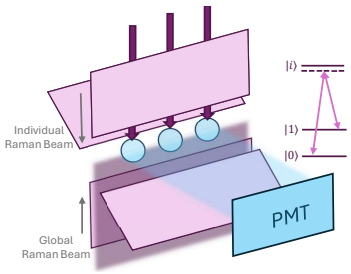
## 7.3 Training

BQNN training was done using stochastic gradient descent with momentum [348]. A difficulty in gradient descent arises for both classical and quantized networks due to the activation function Eq. (7.1) having an undefined first derivative and probabilistic activations correspondingly. To overcome this, a clipped straight-through gradient estimator [287] was used. I.e., when

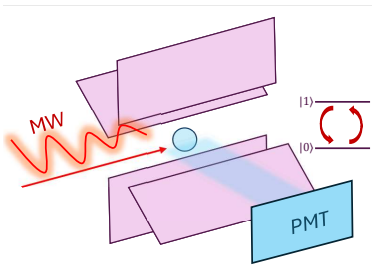



(a) Mapping from classical NN to BQNN.

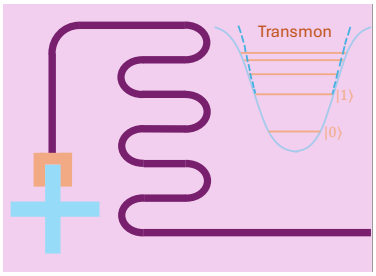
(b) BQNN circuit injecting $U$ and $U^\dagger$ pairs with single- (left) and two-qubit (right) gates.




(c) Trapped-ion quantum computer with laser driven gates.




(d) Trapped-ion quantum computer with microwave driven gates.




(e) Quantum computer based on superconducting qubits.

Figure 7.1: (a) The mapping from the classical binarized multi-layer perceptron network to the quantum circuit implementing the hidden layers of the BQNN. Each step of the forward pass consists of rotating each qubit by an angle controlled by the classical channels, then measuring each qubit. These measurements are the activations which are then passed to the next layer of the network. The networks consist of an input layer, $l$ hidden layers, and an output layer. (b) The BQNN circuit with noise injection by means of inserted $U$ and $U^\dagger$ pairs. (c)-(e) The varieties of quantum hardware on which we run the BQNN: trapped ions with microwave driven gates, trapped ions with laser driven gates, and superconducting qubits.

backpropagating the gradients, the activations are treated as:

$$d_i^k = \phi_{\text{back}}\left((W^k \mathbf{d}^{k-1})_i\right), \tag{7.6}$$

$$\phi_{\text{back}}(x) = \text{htanh}(x) = \begin{cases} \text{sgn}(x), & |x| \geq 1 \\ x, & |x| < 1 \end{cases}. \tag{7.7}$$

This difference between the forward and backward passes introduces error into the gradient estimation, but in practice yields an effective training procedure [285].

An additional difference between classical and quantum networks arises in inference. While

it is only necessary to run inference once for each test datum in the deterministic classical case, in the quantum case, classifications fluctuate from run to run. We take a sample of classifications for a single input and select the outcome by majority vote. Ref. [5] explored both training and inference with a classically-simulated BQNN on the MNIST handwritten digit recognition dataset [349], finding that the classification error rate improved over classical-mode results in an intermediate quantum regime where $0 < a < \mathcal{O}(1)$. This advantage only required on the order of 10 classification samples per image.

## 7.4 Hardware Implementations

In the present work we still train the network using classical simulation, but perform *classification* on two different sets of hardware i.e. IBM circuit-QED processors based on transmon qubits [350, 351], the trapped-ion quantum computer at the Joint Quantum Institute with microwave gates and Raman gates. The network consists of three layers with 16 qubits per layer and is classically trained using a straight-through estimator (STE) on the MNIST dataset.While the circuit in Fig. 7.1b is constructed on 16 qubits according to the network's architecture, the circuit is separable and hence can be executed on as few as one qubit. The training and classification code can be found in the repositories [352, 353]. Below is a brief description of the hardware.

The ion trap system is described in detail in Ref. [354]. It consists of $^{171}\mathrm{Yb}^{+}$ ions confined in a linear Paul trap. Each ion possesses two distinct types of quantum states: internal states and motional states. The internal states, used to encode qubits, correspond to hyperfine levels within the electronic ground-state manifold, $^{2}S_{1/2}$. Specifically, the qubit basis states $|0\rangle$ and $|1\rangle$ are defined as $|F, m_F\rangle = |0, 0\rangle$ and $|1, 0\rangle$ respectively. Here, $F$ indicates the total angular momentum

including both the nuclear spin and total electronic angular momentum in these hyperfine-split states while $m_F$ is its projection onto the quantization axis. The motional states arise from the quantized collective oscillations of the ion chain within the harmonic potential formed by the radio-frequency and static electric fields of the Paul trap. These quantized vibrational modes serve as a shared quantum bus and play a central role in mediating interactions between qubits in trapped-ion quantum systems.

In this experiment, we employ two trapped-ion gate architectures: microwave-driven single-qubit gates and laser-driven Raman transitions which implement both single-qubit and two-qubit entangling gates. For Raman-based operations, three ions are confined simultaneously, though only the left and right ions were used for quantum gates. Single-qubit gates are implemented using stimulated Raman transitions, driven by two beams derived from a 355 nm mode-locked pulsed laser: one beam globally illuminates the ion chain, while the other is split into tightly focused beams for individual ion addressing. The frequency difference matches the qubit splitting and drives coherent transitions between the qubit states. Entangling gates are implemented by simultaneously applying two bichromatic laser beams—one to each ion— tuned near the red and blue sidebands of a shared motional mode. These beams generate spin-dependent forces that couple the ions' internal (spin) states through their collective motion, realizing a high-fidelity Molmer–Sorensen interaction. For microwave gates, a single ion is trapped, and transitions between the states are driven using resonant microwave radiation. This approach enables higher fidelity qubit control independent of the qubit motional state, the latter being a source of error in laser-driven Raman gates. In both modes, the qubit states are read out by collecting state-dependent fluorescence on a photomultiplier tube array.

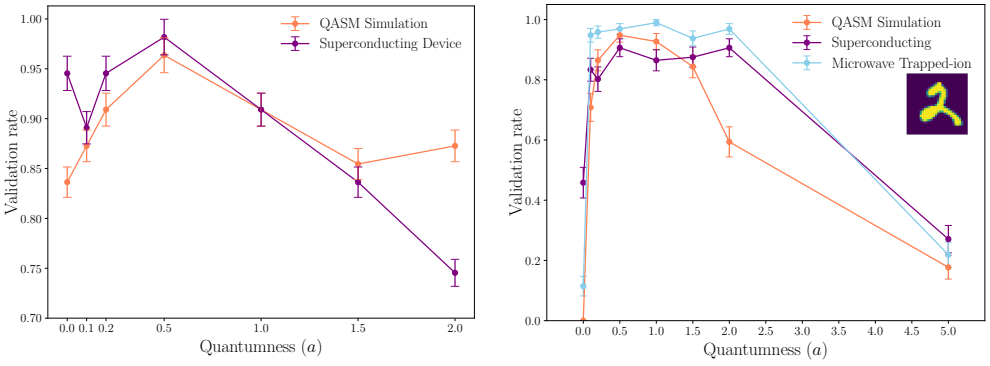
All superconducting-qubit experiments are executed on IBM's commercial Eagle and Heron

processors, which host capacitively coupled transmon qubits arranged in a heavy-hex lattice and controlled with microwave single-qubit rotations, cross-resonance two-qubit gates, and dispersive resonator readout. We access these processors through the IBM Quantum cloud API, submitting circuits and retrieving results remotely.

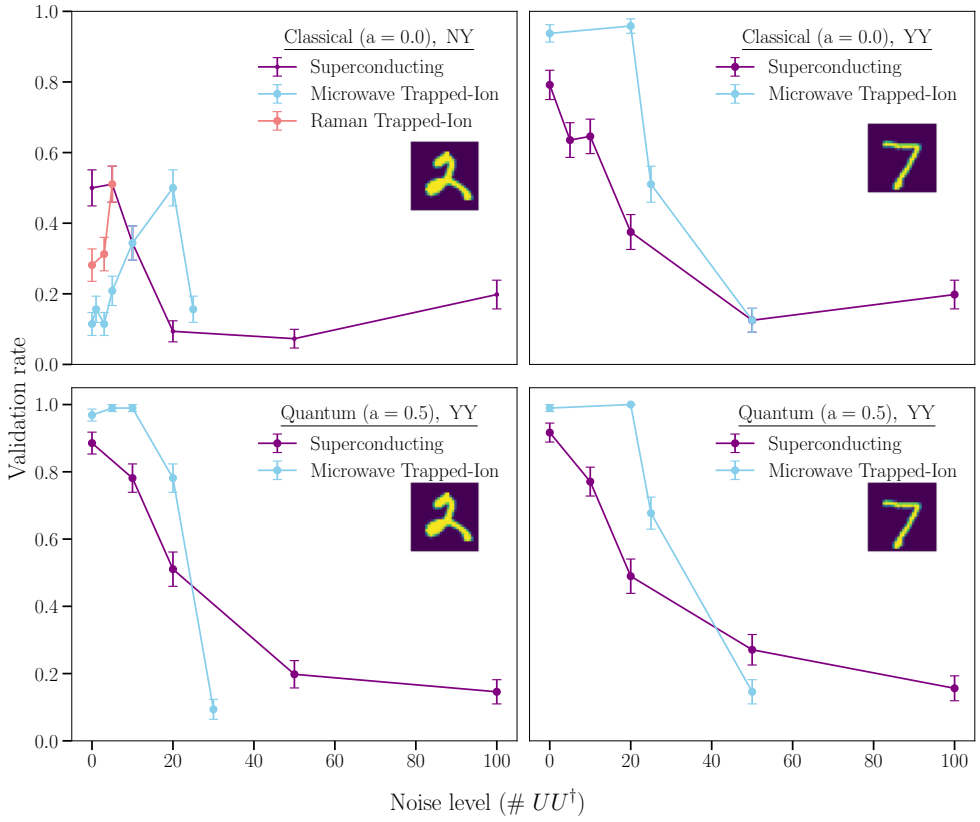
## 7.5 Classification Results

As an initial test of the effectiveness of BQNN, we randomly select and classify 55 images from the MNIST test set, taking 10 samples per image. Due to resource constraints, this initial test is performed only on quantum assembly simulation (QASM) and IBM hardware. The results are depicted in Fig. 7.2a. Our results broadly agree with the prediction of Ref. [5]. In particular, in an intermediate regime ($a \approx 0.5$) we observe the validation rate of the BQNN greater than that of the analogous classical network. Surprisingly, the validation rates observed in the experiment slightly exceed those obtained through classical simulation for $a \leq 0.5$. Although the difference is within error bars, its presence can only be attributed to physical noise and demonstrates that noise is not always hindrance to performance. This initial experiment also explicitly verifies that quantum uncertainty in quantum neural nets can play a role similar to injecting stochastic noise during classical neural-network training, which is known to act as an effective regularizer and boosts performance [355]. We explicitly verify a similar behavior here but induced by intrinsically quantum noise.

To further explore the effect of both quantum randomness and physical noise on image classification, we select a representative image (index 6929) that failed to be correctly classified by the classical model ($a = 0$) but succeeded in simulations with the introduction of quantum




(a) (b)




(c)

Figure 7.2: (a) Validation rates on 55 randomly selected test images obtained from superconducting hardware and simulation for different values of $a$. For each image, ten shots were taken and the most common outcome was chosen as the predicted class. (b) An NY image that was misclassified at $a = 0.0$ shows improved classification accuracy when the quantumness $a$ is increased to some optimal level on the three different platforms, i.e. superconducting, trapped-ion, and classical simulation. (c) For an NY and YY image, taken at either $a = 0.0$ (classical) and $a = 0.5$ (quantum), noise is controlled by the insertion of a number of single-qubit $UU^\dagger = I$ pairs.

noise ($a > 0$). We call such images No-Yes (NY) images, in contrast with Yes-Yes (YY) images, which are correctly classified by both the quantum and classical networks. Fig. 7.2b compares results on the NY image between simulation, IBM hardware, and microwave-gate trapped-ions.

Perhaps the most interesting observation is the difference between the simulated behavior and experimental results at zero and small values of the quantumness parameter, $a$. At $a = 0$, the classical network deterministically fails and so does the idealized quantum simulation of its quantum realization. However, the experimental validation rate at $a = 0$ for this image is 10% for the trapped ion quantum device and as high as 50% for the IBM device. This behavior too can only be attributed to physical noise. A plausible explanation is that there are two nearby minima in the energy landscape of the trained neural net [4] - one representing the erroneous reading, which gets selected classically 100% of the time, and the second being the correct reading. Measurement noise introduces classical stochasticity and errors, which in this case switch to the correct classification in a subset of runs. We note that we have not observed such sensitivity to physical noise at $a = 0$ for clear YY images (where presumably there is a single deep minimum insensitive to small perturbations). This simple observation suggests an interesting use case for noisy quantum neural networks as a way to flag ambiguous data entries. For the MNIST dataset, the data quality is relatively obvious from visual inspection, but this is not necessarily the case for more complex datasets.

The NY images are the source of the boost of performance upon quantization exemplified in Fig. 7.2a. We explore the validation rate as a function of quantumness, $a$, for the particular chosen image in Fig. 7.2b. We observe a steep rise in the validation rate for this image to nearly 100% on the trapped ion device and about 90% on the IBM quantum computer. When $a$ becomes sufficiently large, it introduces nearly random rotations by producing qubit cat-states with almost

equal probabilities for the two computational basis states. This predictably hinders classification performance and we observe validation rates dropping with with $a$. In the $a \to \infty$ limit classifications becomes uniformly random, and the validation rate converges to 0.1 in accordance with Eq. (7.4).

## 7.6 Noise Injection

The validation rate inferred from both quantum platforms shows a subtle maximum at $a \approx 0.5$, but otherwise exhibits a plateau in an intermediate quantum regime before eventual performance decay. Owing to its sensitivity to physical noise, the $a = 0$ and optimal $a$ regimes of the BQNN allow for a loose comparison of the physical noise between platforms. In particular, it is plausible that for these low-depth circuits, physical errors are the largest portion of the error. Hence, devices with relatively high measurement errors, such as superconducting devices, will see a greater noise-induced rise in validation rate [356]. In order to serve as a more stringent benchmark of quantum devices, entanglement and mid-circuit measurement will need to be incorporated into the BQNN.

We further modify the BQNN to benchmark the effects of gate noise on feedforward dynamics. Specifically, $n$ pairs of native unitary gates $UU^\dagger = I$ are injected into the circuit after each layer. While in an ideal system, they have no effect, their application in a real device enhances bit-flip, phase-flip and other types of noise in the case of single-qubit pairs and introduce "parasitic entanglement" and crosstalk for pairs of time-reversed entangling gates.

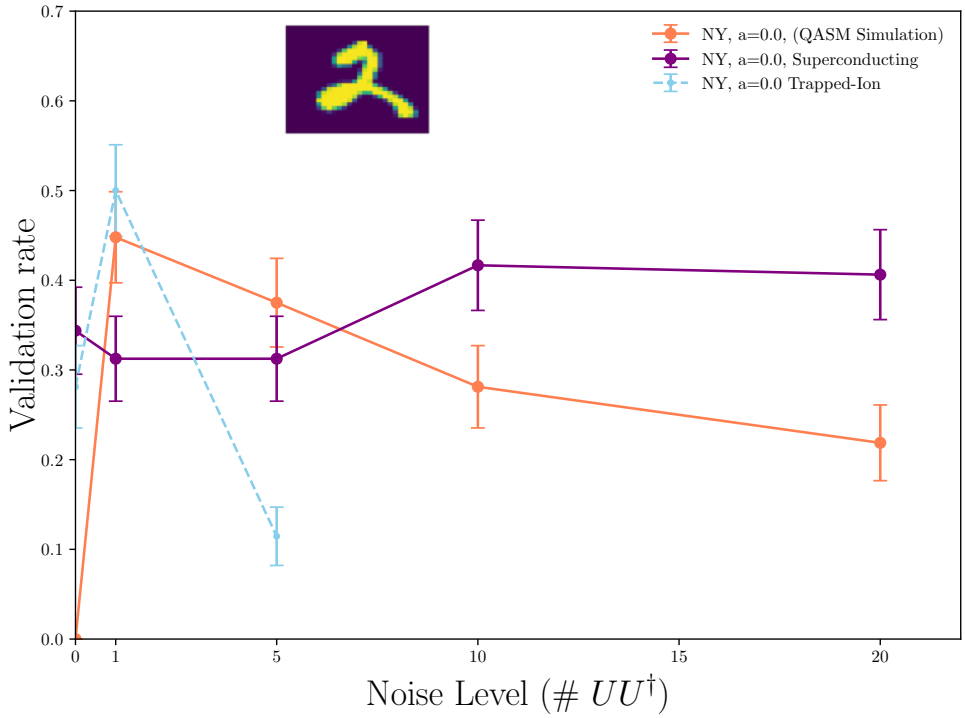



Figure 7.3: For an NY image at $a = 0.0$, noise is controlled by the insertion of a number of two-qubit $UU^\dagger = I$ pairs.

## 7.7 Conclusion

Fig. 7.2c shows results of our single-qubit noise experiments. In the classical regime, we observe for the trapped ion device that even modest amounts of injected noise lead to a clear jump in NY-image accuracy at around 20 noise-gate pairs. We do not see any notable improvement of performance on the YY image with injected noise, which further validates the heuristic picture of two close by minima for the borderline (ambiguous) NY data entries versus a single deep minimum for the clear YY images. This behavior carries over into the moderately quantum regime, where the combination of quantum randomness and added single-qubit noise has only a minor impact on validation rates. By contrast, the superconducting system does not exhibit any enhancement in performance on the NY images with injected noise suggesting that the native

physical measurement noise is already saturated or above the threshold to be beneficial for the BQNN performance. We see the per-image accuracy decline almost monotonically as noise increases.

The final test uses native two-qubit gate pairs and results are depicted in Fig. 7.3. For the trapped-ion experiment it was found that after the introduction of a *single* pair of two-qubit gates, the per-image classification-rate shot up over $40\%$ before deteriorating to random $10\%$ for five gate pairs. The superconducting platform was more resilient to this entanglement-noise and even showed a small improvement in performance with the addition of further pairs, a behavior not consistent with our intuition or simulations.

All in all, our experimental results suggest the possibility that certain types of physical noise, usually thought detrimental to performance of quantum computers, can be beneficial for certain quantum machine learning tasks. While such instances at the inference stage may be accidental or task-specific, we expect a more systematic advantage when training quantum networks with small amounts of background physical noise.

The main obstacle to generalizing this architecture to quantum networks that cannot be efficiently simulated classically, and thus with potential for genuine quantum advantage, is the challenge of implementing mid-circuit measurements and feedback. Without measurements, any circuit is a subset of a unitary matrix which lacks non-linear input. On the other hand, with complete measurement in every layer the quantum state collapses and entanglement is disrupted, hence leading to circuits which can always be simulated classically. The only alternative is entangled (e.g., brick-wall) circuits with partial mid-circuit measurements. This corresponds to circuits designed for measurement-induced phase transitions with the difference being that in neural nets, gates applied in every layer are a function of the partial measurement in the previous layer. Such

measurement-induced quantum neural nets can be trained using reinforcement learning to aid in architecture search, which will be reported elsewhere.

## Acknowledgments

This work is supported by a collaboration between the US DOE and other Agencies. This material is based upon work supported by the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers, Quantum Systems Accelerator under Award No. DE-FOA-0002253 (XL, AMG, NML) and Basic Energy Sciences under Award No. DE-SC0001911 (VG, RB). Additional support is acknowledged from the Army Research Office under Grant Number W911NF-23-1-0241 (DJH) and from the National Science Foundation STAQ project No. PHY-2325080 (NML). We acknowledge the use of IBM Quantum Credits for this work. The views expressed are those of the authors, and do not reflect the official policy or position of IBM or the IBM Quantum team.

# Chapter 8: Summary and Discussion

The body of this thesis began in Chapter 2 by studying the spectral statistics of a random matrix ensemble that captures several features of physical quantum glass models. In quantum glasses the Hamiltonian becomes fractured due to the presence of quenched disorder in the interactions. The BRP model captures this behavior by constructing random matrices with GUE blocks on the diagonal and scalable disordered couplings between these blocks. While calculating the SFF for more physical spin models is often not feasible, the SFF for the BRP was calculated exactly in the limit of infinite matrix size. It was found that, in the glassy and localized phases, the ramp of the SFF is enhanced by a factor equal to the number of weakly coupled sectors. This indicates that the SFF can be used to count the number of metastable states in spin glasses.

This idea is further explored in Chapter 3, where the SFF is calculated for the quantum PSM, a more physical glass model. The classical PSM is known to be a 1RSB glassy model, making the quantized model an ideal place to explore what can be learned about the number of metastable states through the SFF. As for the BRP model, it is found that, below the ergodic transition, the ramp of the SFF is enhanced by a factor equal to the number of metastable states, for times less than exponential in system size. The number of such states is found through the quantum TAP equations, directly connecting the weakly coupled sectors of the Hamiltonian with

local minima that emerge in the free energy landscape.

Chapter 4 takes the concepts of the BRP model and translates them to the context of random quantum circuits. A random Floquet operator is defined which has CUE blocks on the diagonal, with scalable disordered couplings between the blocks. The SFF was calculated with a diagrammatic method for integrating over the unitary group, and the time evolution of the disorder-averaged density matrix was found using an effective field theory approach. This model was shown to exhibit glassy dynamics, such as delayed thermalization. The results of this chapter open the door to using quantum circuits to explore nontrivial thermalization dynamics in many-body quantum systems.

Chapter 5 then applies methods from spin glass theory to another context: neural networks. In this work, snapshots of the training weights are taken for a MLP throughout its training process. This allows for the analysis of the learning process by mapping the network to an analogous Ising spin model. In this analogy neurons are mapped to spins and weights are mapped to two-spin interactions. The TAP equations are then used to determine the melting transition and the nature of that transition. It is found that, before training, the model exhibits a glassy transition at low temperature. However, with training, that transition temperature increases as a power law. Additionally, the transition changes from a glassy transition to that of an ordered phase. This statistical mechanical approach provides an interesting perspective on the training process, where training can be understood as selecting and strengthening some state associated with the network's task.

Chapter 6 explores a simple way to quantize a multilayer perceptron in such a way that the network may be adiabatically tuned between classical and quantum modes of operation. This allows for a direct comparison between classical and quantum performance by tuning a single

parameter, without changing anything about the network's architecture. Through classical simulation the networks are trained to classify handwritten digits on a subset of the MNIST dataset. It is found that, for intermediate "quantumness," performance exceeds that of the classical network. This occurs due to the injection of quantum uncertainty to the network, which provides some regularization in the training process, avoiding overfitting.

These ideas are then implemented on actual quantum hardware in Chapter 7. The networks are trained through classical simulation, but the classification on test data is performed on both superconducting and trapped ion quantum computing platforms. The experimental results are consistent with the simulation results of the previous chapter. Additionally, it is seen that the physical noise that exists in these real-world systems can be advantageous for regularization in some contexts.

The work reported in this thesis opens up many fascinating avenues for further exploration. In terms of quantum neural networks, there are several generalizations of the circuits discussed in Chapters 6 and 7 that may enable even better performance. Although one of the circuits presented makes use of entanglement in order to perform weak measurements on neuron qubits, entanglement is not created between the neuron qubits themselves. True quantum advantage will surely involve nontrivial interactions between these qubits. However, in order to train such a circuit, a different method for optimization must be used. The current method, which estimates gradients of the loss using a clipped straight-through estimator, has been observed to be ineffective when including nontrivial two-qubit gates.

There are a multitude of ways that the random matrix ensembles discussed in Chapters 2 and 4 can be generalized to capture more features of physical glassy systems. The ensembles explored in this thesis are analogous to 1RSB systems with statistically identical sectors. These

could be easily generalized to ensembles with a distribution of sector sizes and higher levels of symmetry breaking. Despite the simplicity of constructing these generalizations, calculating the exact spectral statistics for such models is nontrivial.

Finally, analysis of learning in NNs through statistical mechanical analogues, as in Chapter 5, can also be expanded in a variety of directions. While this thesis uses this approach for a fairly straightforward task and dataset, it would be illustrative to apply it to a problem with nontrivial learning dynamics, such as grokking [357], where generalization performance improves much later than the time at which the model saturates its performance on the training dataset. Such a situation would likely lead to a very different phase diagram.

# Appendix A: Appendix to Chapter 2

Here we show that the area between the SFF and its plateau value is independent of the strength of the coupling between different states, as long as the coupling is nonzero. Consider the integral

$$I = \int_{-\infty}^{\infty} dt(\text{SFF}(t) - N), \tag{A.1}$$

where

$$\text{SFF}(t) = \sum_{mn} \overline{\exp[i(E_n - E_m)t]} \tag{A.2}$$

is the SFF. We can approximate $I$ as

$$I(t_{\text{long}}) = \int_{-\infty}^{\infty} dt(\text{SFF}(t) - N) \exp\left(-\frac{t^2}{2t_{\text{long}}^2}\right), \tag{A.3}$$

which goes to $I$ as $t_{\text{long}} \to \infty$. We find that

$$\begin{aligned} I(t_{\text{long}}) &= \int_{-\infty}^{\infty} dt \sum_{m \neq n} \overline{\exp[i(E_n - E_m)t]} \exp\left(-\frac{t^2}{2t_{\text{long}}^2}\right) \\ &= \sqrt{2\pi} t_{\text{long}} \sum_{m \neq n} \overline{\exp\left(\frac{-(E_n - E_m)^2 t_{\text{long}}^2}{2}\right)}. \end{aligned} \tag{A.4}$$

If there is repulsion between all levels so that each level has a nonzero distance from the others, $I(t_{\text{long}})$ will vanish for $t_{\text{long}}$ larger than the Heisenberg time. This indicates that $I = 0$.

Because the disconnected part of the SFF depends only on the level density, which for the models considered in this work is independent of the coupling strength when $\gamma > 1$, the area between the connected SFF and its plateau value will also be independent of the coupling strength for those values of $\gamma$. We found above that, as $\gamma \to 1^+$ for both models, the SFF goes to the GUE result, which is also the result for all $\gamma \leq 1$. From this we can conclude that the area between the unfolded connected SFF and its plateau value is equal to the GUE result for any nonzero coupling strength. This is simply a triangle with area $\pi N$.

# Appendix B: Appendices to Chapter 3

## B.1 Derivation of Schwinger-Keldysh TAP equations

Here we derive the TAP equations on the Schwinger-Keldysh contour, given by Eqs. (3.35) and (3.36) of the main text. Our derivation is a straightforward generalization of that in Ref. [25] for the thermal circle, but since we are not aware of it appearing in the literature, we aim to make this section as self-contained as possible.

We begin with the expression for the TAP action, given by Eq. (3.32) and reproduced here:

$$iNS_{\text{TAP}}[m,\mathcal{G},\eta] \equiv \log \int \mathcal{D}\sigma^N \exp\left[ i\sum_i S_i^0 - i\eta \int_{\mathcal{C}} dt \sum_{(i_1\cdots i_p)} J_{i_1\cdots i_p}\sigma_{i_1}(t)\cdots\sigma_{i_p}(t)\right] + \frac{iN}{2}\int_{\mathcal{C}} dt z(t) - i\int_{\mathcal{C}} dt \sum_i h_i(t)m_i(t) + \frac{iN}{2}\int_{\mathcal{C}} dtdt' \Lambda(t,t')\mathcal{G}(t,t'), \tag{B.1}$$

where $\mathcal{C}$ denotes the Schwinger-Keldysh contour and

$$S_i^0 \equiv \int_{\mathcal{C}} dt \left( \frac{\mu}{2}\big(\partial_t\sigma_i(t)\big)^2 - \frac{z(t)}{2}\sigma_i(t)^2 + h_i(t)\sigma_i(t)\right) - \frac{1}{2}\int_{\mathcal{C}} dtdt' \Lambda(t,t')\sigma_i(t)\sigma_i(t'). \tag{B.2}$$

Note that we have included an additional parameter $\eta$ governing the strength of interactions. We calculate $S_{\text{TAP}}$ by expanding in $\eta$, ultimately setting $\eta = 1$. One can show that all terms beyond second order vanish in the thermodynamic limit [25, 269], so we need only calculate $S_{\text{TAP}}$ and its

first two derivatives at $\eta = 0$. Then

$$S_{\text{TAP}}[m, \mathcal{G}, 1] \sim S_{\text{TAP}}[m, \mathcal{G}, 0] + \frac{\partial S_{\text{TAP}}[m, \mathcal{G}, 0]}{\partial \eta} + \frac{1}{2}\frac{\partial^2 S_{\text{TAP}}[m, \mathcal{G}, 0]}{\partial \eta^2}. \tag{B.3}$$

We proceed term by term.

## Zeroth order

At zeroth order in $\eta$, we have that

$$\begin{aligned} iN S_{\text{TAP}}[m, \mathcal{G}, 0] = \sum_i \log \int \mathcal{D}\sigma_i e^{iS_i^0} \\ + \frac{iN}{2}\int_{\mathcal{C}} dt z(t) - i\int_{\mathcal{C}} dt \sum_i h_i(t) m_i(t) + \frac{iN}{2}\int_{\mathcal{C}} dt dt' \Lambda(t,t')\mathcal{G}(t,t'). \end{aligned} \tag{B.4}$$

The first term evaluates to

$$\begin{aligned} &\int \mathcal{D}\sigma_i \exp\left[-\frac{i}{2}\int_{\mathcal{C}} dt dt' \sigma_i(t)\left(\delta(t-t')\left(\mu\partial_t^2 + z(t)\right) + \Lambda(t,t')\right)\sigma_i(t') + i\int_{\mathcal{C}} dt h_i(t)\sigma_i(t)\right] \\ &\quad = \text{Det}\Big[i(\mu\partial_t^2 + z) + i\Lambda\Big]^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\int_{\mathcal{C}} dt dt' h_i(t)\Big(i(\mu\partial_t^2 + z) + i\Lambda\Big)^{-1}(t,t')h_i(t')\right]. \end{aligned} \tag{B.5}$$

We further have that

$$\big\langle \sigma_i(t)\sigma_i(t')\big\rangle - \big\langle \sigma_i(t)\big\rangle\big\langle \sigma_i(t')\big\rangle = \Big(i(\mu\partial_t^2 + z) + i\Lambda\Big)^{-1}(t,t'), \tag{B.6}$$

$$\big\langle \sigma_i(t)\big\rangle = i\int_{\mathcal{C}} dt' \Big(i(\mu\partial_t^2 + z) + i\Lambda\Big)^{-1}(t,t')h_i(t'). \tag{B.7}$$

Note that the left-hand sides are constrained to be $\mathcal{G}(t,t') - Q(t,t')$ and $m_i(t)$ (summing over $i$ for the former), where $Q(t,t') \equiv N^{-1}\sum_i m_i(t)m_i(t')$. Thus we can use Eqs. (B.6) and (B.7) to express the action in terms of them. The first term of Eq. (B.2) becomes

$$\frac{N}{2}\log \mathrm{Det}\Big[\mathcal{G} - Q\Big] + \frac{N}{2}\int_{\mathcal{C}} dtdt' \Big(\mathcal{G} - Q\Big)^{-1}(t,t')Q(t,t'), \tag{B.8}$$

and the second line can be written

$$-N\int_{\mathcal{C}} dtdt' \Big(\mathcal{G} - Q\Big)^{-1}(t,t')Q(t,t') - \frac{iN\mu}{2}\int_{\mathcal{C}} dt\partial_t^2 \mathcal{G}(t,t')\Big|_{t'=t} + \frac{N}{2}\int_{\mathcal{C}} dtdt' \Big(\mathcal{G} - Q\Big)^{-1}(t,t')\mathcal{G}(t,t'). \tag{B.9}$$

Combining the two, we have that (up to an unimportant constant)

$$iS_{\text{TAP}}[m,\mathcal{G},0] = \frac{1}{2}\log \mathrm{Det}\Big[\mathcal{G} - Q\Big] - \frac{i\mu}{2}\int_{\mathcal{C}} dt\partial_t^2 \mathcal{G}(t,t')\Big|_{t'=t}. \tag{B.10}$$

## First order

Strictly speaking, since $h_i(t)$ and $\Lambda(t,t')$ are chosen so that $\langle\sigma_i(t)\rangle = m_i(t)$ and $\sum_i\langle\sigma_i(t)\sigma_i(t')\rangle = N\mathcal{G}(t,t')$ at every value of $\eta$, they themselves are functions of $\eta$ (as is $z(t)$). However, the Legendre-transform structure of the TAP action ensures that the total derivative with respect to $\eta$ equals the partial derivative holding fields fixed (for example, the contribution to $\partial S_{\text{TAP}}/\partial\eta$ from

$\partial h_i(t)/\partial\eta$ comes with a factor $\langle\sigma_i(t)\rangle - m_i(t) = 0$). Thus we have simply that

$$\frac{\partial S_{\text{TAP}}[m,\mathcal{G},\eta]}{\partial\eta} = -\frac{1}{N}\int_{\mathcal{C}} dt \sum_{(i_1\cdots i_p)} J_{i_1\cdots i_p}\big\langle\sigma_{i_1}(t)\cdots\sigma_{i_p}(t)\big\rangle. \tag{B.11}$$

At $\eta = 0$, the spins are non-interacting and the expectation value factors:

$$\frac{\partial S_{\text{TAP}}[m,\mathcal{G},0]}{\partial\eta} = -\frac{1}{N}\int_{\mathcal{C}} dt \sum_{(i_1\cdots i_p)} J_{i_1\cdots i_p} m_{i_1}(t)\cdots m_{i_p}(t). \tag{B.12}$$

## Second order

The second derivative requires a bit more work. From Eq. (B.11), we have that

$$\begin{aligned}
\frac{\partial^2 S_{\text{TAP}}[m,\mathcal{G},\eta]}{\partial\eta^2} &= \frac{i}{N}\int_{\mathcal{C}} dtdt' \sum_{II'} J_I J_{I'}\Big(\big\langle\sigma_I(t)\sigma_{I'}(t')\big\rangle - \big\langle\sigma_I(t)\big\rangle\big\langle\sigma_{I'}(t')\big\rangle\Big) \\
&\quad - \frac{i}{N}\int_{\mathcal{C}} dtdt' \sum_{Ii'} J_I \frac{\partial h_{i'}(t')}{\partial\eta}\Big(\big\langle\sigma_I(t)\sigma_{i'}(t')\big\rangle - \big\langle\sigma_I(t)\big\rangle\big\langle\sigma_{i'}(t')\big\rangle\Big) \\
&\quad + \frac{i}{2N}\int_{\mathcal{C}} dtdt' \sum_{Ii'} J_I \frac{\partial z(t')}{\partial\eta}\Big(\big\langle\sigma_I(t)\sigma_{i'}(t')^2\big\rangle - \big\langle\sigma_I(t)\big\rangle\big\langle\sigma_{i'}(t')^2\big\rangle\Big) \\
&\quad + \frac{i}{2N}\int_{\mathcal{C}} dtdt'dt'' \sum_{Ii'} J_I \frac{\partial\Lambda(t',t'')}{\partial\eta}\Big(\big\langle\sigma_I(t)\sigma_{i'}(t')\sigma_{i'}(t'')\big\rangle - \big\langle\sigma_I(t)\big\rangle\big\langle\sigma_{i'}(t')\sigma_{i'}(t'')\big\rangle\Big),
\end{aligned} \tag{B.13}$$

where for brevity, we use $I$ to denote $(i_1\cdots i_p)$ and $\sigma_I(t)$ to denote $\sigma_{i_1}(t)\cdots\sigma_{i_p}(t)$. To evaluate the partial derivatives on the right-hand side, which we only need at $\eta = 0$, note that

$$N\frac{\partial S_{\text{TAP}}[m,\mathcal{G},\eta]}{\partial m_i(t)} = -h_i(t), \qquad \frac{\partial S_{\text{TAP}}[m,\mathcal{G},\eta]}{\partial\rho(t)} = \frac{1}{2}z(t), \qquad \frac{\partial S_{\text{TAP}}[m,\mathcal{G},\eta]}{\partial\mathcal{G}(t,t')} = \frac{1}{2}\Lambda(t,t'), \tag{B.14}$$

where we set $\sum_i \langle \sigma_i(t)^2 \rangle = N\rho(t)$ (even though ultimately $\rho(t) = 1$). Thus, using Eq. (B.12),

$$\begin{aligned} \frac{\partial h_i(t)}{\partial \eta} &= -N \frac{\partial^2 S_{\text{TAP}}[m, \mathcal{G}, 0]}{\partial m_i(t) \partial \eta} = \sum_{(i_1 \cdots i_p)} J_{i_1 \cdots i_p} \frac{\partial \big( m_{i_1}(t) \cdots m_{i_p}(t) \big)}{\partial m_i(t)}, \\ \frac{\partial z(t)}{\partial \eta} &= 2 \frac{\partial^2 S_{\text{TAP}}[m, \mathcal{G}, 0]}{\partial \rho(t) \partial \eta} = 0, \qquad \frac{\partial \Lambda(t,t')}{\partial \eta} = 2 \frac{\partial^2 S_{\text{TAP}}[m, \mathcal{G}, 0]}{\partial \mathcal{G}(t,t') \partial \eta} = 0. \end{aligned} \tag{B.15}$$

The second derivative simplifies to

$$\begin{aligned} \frac{\partial^2 S_{\text{TAP}}[m, \mathcal{G}, \eta]}{\partial \eta^2} = \frac{i}{N} \int_{\mathcal{C}} dt dt' \sum_{II'} J_I J_{I'} \Big( \big\langle \sigma_I(t) \sigma_{I'}(t') \big\rangle - \big\langle \sigma_I(t) \big\rangle \big\langle \sigma_{I'}(t') \big\rangle \Big) \\ - \frac{i}{N} \int_{\mathcal{C}} dt dt' \sum_{II'i'} J_I J_{I'} \Big( \big\langle \sigma_I(t) \sigma_{i'}(t') \big\rangle - \big\langle \sigma_I(t) \big\rangle \big\langle \sigma_{i'}(t') \big\rangle \Big) \frac{\partial \big\langle \sigma_{I'}(t') \big\rangle}{\partial m_{i'}(t')}. \end{aligned} \tag{B.16}$$

Just as one can show that higher-order $\eta$ derivatives vanish in the thermodynamic limit, one can also show that it is safe to replace $J_I J_{I'}$ by its average in Eq. (B.16) [25, 269]. Again using that expectation values factor at $\eta = 0$, we therefore have that

$$\frac{\partial^2 S_{\text{TAP}}[m, \mathcal{G}, 0]}{\partial \eta^2} = \frac{iJ^2}{p} \int_{\mathcal{C}} dt dt' \Big[ \mathcal{G}(t,t')^p - Q(t,t')^p - p \Big( \mathcal{G}(t,t') - Q(t,t') \Big) Q(t,t')^{p-1} \Big]. \tag{B.17}$$

## Full TAP action

Inserting Eqs. (B.10), (B.12), and (B.17) into Eq. (B.3), we find that

$$\begin{aligned} iS_{\text{TAP}} = \frac{1}{2} \log \text{Det} \Big[ \mathcal{G} - Q \Big] - \frac{i\mu}{2} \int_{\mathcal{C}} dt \partial_t^2 \mathcal{G}(t,t') \Big|_{t'=t} - \frac{i}{N} \int_{\mathcal{C}} dt \sum_{(i_1 \cdots i_p)} J_{i_1 \cdots i_p} m_{i_1}(t) \cdots m_{i_p}(t) \\ - \frac{J^2}{2p} \int_{\mathcal{C}} dt dt' \Big[ \mathcal{G}(t,t')^p + (p-1) Q(t,t')^p - p \mathcal{G}(t,t') Q(t,t')^{p-1} \Big]. \end{aligned} \tag{B.18}$$

As discussed in the main text, the TAP equations come from setting to zero the derivatives with respect to $m_i(t)$ and $\mathcal{G}(t,t')$. However, since $\mathcal{G}(t,t)$ is not free to vary (it must equal 1 by the spherical constraint), we include a Lagrange multiplier that we again denote by $z(t)$. The TAP equations are thus

$$\Big(\mathcal{G}-Q\Big)^{-1}(t,t') - i\mu\partial_t^2\delta(t-t') - J^2\big(\mathcal{G}(t,t')^{p-1} - Q(t,t')^{p-1}\big) = iz(t)\delta(t-t'), \tag{B.19}$$

$$\begin{aligned}\int_{\mathcal{C}} dt'\Big(\mathcal{G}-Q\Big)^{-1}(t,t')m_i(t') + i\sum_{(i_1\cdots i_p)} J_{i_1\cdots i_p}\frac{\partial\big(m_{i_1}(t)\cdots m_{i_p}(t)\big)}{\partial m_i(t)} \\ - (p-1)J^2\int_{\mathcal{C}} dt'\Big(\mathcal{G}(t,t')-Q(t,t')\Big)Q(t,t')^{p-2}m_i(t') = 0.\end{aligned} \tag{B.20}$$

Time translation invariance allows us to further simplify by setting $m_i(t) = m_i$ and $z(t) = z$. Note that then $Q(t,t') = N^{-1}\sum_i m_i^2 \equiv q_{\text{EA}}$. After some rearrangement, the TAP equations can be written

$$i\big(\mu\partial_t^2 + z\big)\Big(\mathcal{G}(t,t') - q_{\text{EA}}\Big) + J^2\int_{\mathcal{C}} dt''\Big(\mathcal{G}(t,t')^{p-1} - q_{\text{EA}}^{p-1}\Big)\Big(\mathcal{G}(t'',t') - q_{\text{EA}}\Big) = \delta(t-t'), \tag{B.21}$$

$$\begin{aligned}J^2\int_{\mathcal{C}} dt'\Big(\mathcal{G}(t,t')^{p-1} - (p-1)q_{\text{EA}}^{p-2}\mathcal{G}(t,t') + (p-2)q_{\text{EA}}^{p-1}\Big)m_i \\ = -izm_i - i\sum_{(i_1\cdots i_p)} J_{i_1\cdots i_p}\frac{\partial(m_{i_1}\cdots m_{i_p})}{\partial m_i}.\end{aligned} \tag{B.22}$$

These are precisely Eqs. (3.35) and (3.36) from the main text.

## B.2 Energy of a TAP state

Here we derive an expression for the total energy density $\epsilon$ in terms of a solution $\mathcal{G}(t,t')$ to the TAP equations, Eqs. (B.21) and (B.22).

Note that, despite the complicated derivation, $\mathcal{G}(t,t')$ is simply the contour-ordered expectation value of $N^{-1}\sum_i \sigma_i(t)\sigma_i(t')$. Thus, as we discuss below in App. B.3, the kinetic energy per spin is given by

$$\frac{\mu}{2N\Delta t^2}\sum_i \Big\langle \big[\sigma_i(t+2\Delta t)-\sigma_i(t+\Delta t)\big]\big[\sigma_i(t+\Delta t)-\sigma_i(t)\big]\Big\rangle = -\frac{\mu}{2}\partial_t^2 \mathcal{G}(t^+,t), \tag{B.23}$$

where $t^+$ denotes $t+\Delta t$.

To evaluate the potential energy, we replace $J_{i_1\cdots i_p}$ by $(1+\zeta(t))J_{i_1\cdots i_p}$ in the original path integral. On the one hand,

$$\frac{\partial}{\partial\zeta(t)} S_{\text{TAP}}\Big|_{\zeta=0} = -\frac{1}{N}\sum_{(i_1\cdots i_p)} J_{i_1\cdots i_p}\big\langle \sigma_{i_1}(t)\cdots\sigma_{i_p}(t)\big\rangle. \tag{B.24}$$

On the other hand, it is easy to see how the explicit calculation of $S_{\text{TAP}}$ is modified by $\zeta(t)$:

$$\begin{aligned} S_{\text{TAP}} = S_{\text{TAP}}\Big|_{\zeta=0} &- \frac{1}{N}\int_{\mathcal{C}} dt\zeta(t)\sum_{(i_1\cdots i_p)} J_{i_1\cdots i_p} m_{i_1}(t)\cdots m_{i_p}(t) \\ &+ \frac{iJ^2}{2p}\int_{\mathcal{C}} dtdt'\big(\zeta(t)+\zeta(t')\big)\Big[\mathcal{G}(t,t')^p + (p-1)Q(t,t')^p - p\mathcal{G}(t,t')Q(t,t')^{p-1}\Big] + O(\zeta^2). \end{aligned} \tag{B.25}$$

Thus

$$\frac{1}{N}\sum_{(i_1\cdots i_p)} J_{i_1\cdots i_p}\left\langle \sigma_{i_1}(t)\cdots\sigma_{i_p}(t)\right\rangle = \frac{1}{N}\sum_{(i_1\cdots i_p)} J_{i_1\cdots i_p} m_{i_1}(t)\cdots m_{i_p}(t) \\ -\frac{iJ^2}{p}\int_{\mathcal{C}} dt'\left[\mathcal{G}(t,t')^p + (p-1)Q(t,t')^p - p\mathcal{G}(t,t')Q(t,t')^{p-1}\right], \quad \text{(B.26)}$$

and the total energy per spin is given by

$$\epsilon = -\frac{\mu}{2}\partial_t^2\mathcal{G}(t^+,t) + \frac{1}{N}\sum_{(i_1\cdots i_p)} J_{i_1\cdots i_p} m_{i_1}(t)\cdots m_{i_p}(t) \\ -\frac{iJ^2}{p}\int_{\mathcal{C}} dt'\left[\mathcal{G}(t,t')^p + (p-1)Q(t,t')^p - p\mathcal{G}(t,t')Q(t,t')^{p-1}\right]. \quad \text{(B.27)}$$

Writing out the various components of $\mathcal{G}$ explicitly and using time translation invariance (as well as taking $t$ to be far from the thermal branch), we have that

$$\epsilon = -\frac{\mu}{2}\partial_t^2\mathcal{G}_{\alpha\alpha}(0^+) + Jq_{\text{EA}}^{p/2}\mathcal{E} - \frac{iJ^2}{p}\int dt'\sum_{\alpha'}(-1)^{\alpha'}\left[\mathcal{G}_{\alpha\alpha'}(t-t')^p - pq_{\text{EA}}^{p-1}\mathcal{G}_{\alpha\alpha'}(t-t') + (p-1)q_{\text{EA}}^p\right], \quad \text{(B.28)}$$

where $\mathcal{E}$ is defined as in the main text (Eq. (3.40)). Despite the appearance, this expression is independent of both $t$ and $\alpha$. Compare it to the argument of the filter function from Sec. 3.4 (Eq. (3.68)), reproduced here:

$$\epsilon_\alpha[\lambda, G] = -\frac{\mu}{2}\partial_t^2 G_{\alpha\alpha}(0^+) - \frac{iJ^2}{p}\int_0^T dt'\sum_{\alpha'}(-1)^{\alpha'}G_{\alpha\alpha'}(t')^p - \frac{iJ\lambda}{pq[G]^{p/2}}\left(\frac{1}{T}\int_0^T dt G_{\alpha u}(t,0)\right)^p. \quad \text{(B.29)}$$

Evaluated at the saddle point given in Sec. 3.4.2, Eq. (B.29) comes out to be (up to corrections

small in $T^{-1}$)

$$\epsilon_\alpha[\lambda, G] = -\frac{\mu}{2}\partial_t^2 \mathcal{G}_{\alpha\alpha}(0^+) + J q_{\rm EA}^{p/2}\mathcal{E} - 2J^2 q_{\rm EA}^{p-1}\Lambda - \frac{iJ^2}{p}\int dt' \sum_{\alpha'}(-1)^{\alpha'}\Big(\mathcal{G}_{\alpha\alpha'}(t')^p - q_{\rm EA}^p\Big), \tag{B.30}$$

which, since $2i\Lambda \equiv \int dt' \sum_{\alpha'}(-1)^{\alpha'}\mathcal{G}_{\alpha\alpha'}(t)$, is precisely the energy density of the TAP state, Eq. (B.28).

## B.3 Accounting for filter functions

The purpose of this section is to derive Eqs. (3.44) through (3.46) of the main text. In the original definition of the SFF,

$$\text{SFF}(T, f) = \overline{\text{Tr} f(H) e^{-iHT}\, \text{Tr} f(H) e^{iHT}}, \tag{B.31}$$

we need to be careful in how we treat the filter functions, since $f(H)$ does depend on the specific disorder realization. It is convenient to assume that $f(H)$ is an analytic function of energy *density*, i.e., it can be expanded as

$$f(H) = \sum_{n=0}^{\infty} \frac{c_n}{n!}\left(\frac{H}{N}\right)^n, \tag{B.32}$$

where the coefficients $c_n$ do not depend on $N$. This is not a significant restriction, since we can still choose $f(H)$ to be as tightly concentrated around any given energy density $\epsilon_0$ as we wish.

Inserting Eq. (B.32) into Eq. (B.31) gives

$$\text{SFF}(T, f) = \sum_{n_u n_l} \frac{c_{n_u} c_{n_l}}{n_u! n_l!} \frac{1}{N^{n_u+n_l}} \overline{\text{Tr} H^{n_u} e^{-iHT}\, \text{Tr} H^{n_l} e^{iHT}}. \tag{B.33}$$

We construct the SFF path integral as usual: write $H = H_{\rm q} + H_{\rm cl}$ with $H_{\rm q} \equiv \sum_i \pi_i^2/2\mu$ the kinetic energy and $H_{\rm cl}$ the potential energy (including a Lagrange multiplier for the spherical constraint), approximate

$$e^{iHT} \sim (e^{iH_{\rm q}\Delta t} e^{iH_{\rm cl}\Delta t})^{\frac{T}{\Delta t}}, \tag{B.34}$$

insert resolutions of the identity alternating between the $|\sigma\rangle$ and the $|\pi\rangle$ basis, perform the Gaussian integrals over momenta, and finally take $\Delta t \to 0$. The only subtlety is in where and how we insert the factors of $H$ during this process. While any arrangement has to ultimately give the same answer, we can simplify things by making judicious choices.

First write each $H$ as $\Delta t \sum_t H/T$ and insert one term at every time slice, i.e.,

$$\begin{aligned}
&\mathrm{Tr} H^{n_u} e^{-iHT} \mathrm{Tr} H^{n_l} e^{iHT} \\
&\quad = \left(\frac{\Delta t}{T}\right)^{n_u+n_l} \sum_{t_1 \cdots t_{n_u}} \mathrm{Tr} e^{-iH(T-t_{n_u})} H e^{-iH(t_{n_u}-t_{n_u-1})} \cdots e^{-iH(t_2-t_1)} H e^{-iHt_1} \\
&\qquad\qquad \times \sum_{t_1 \cdots t_{n_l}} \mathrm{Tr} e^{iH(T-t_{n_l})} H e^{iH(t_{n_l}-t_{n_l-1})} \cdots e^{iH(t_2-t_1)} H e^{iHt_1}.
\end{aligned} \tag{B.35}$$

Note that, to leading order in $\Delta t$, no two factors of $H$ coincide at the same time. Furthermore, each factor of $H$ is itself two terms, whose operators we arrange within the Trotterized exponential (Eq. (B.34)) as indicated:

$$\cdots e^{iH_{\rm q}\Delta t} e^{iH_{\rm cl}\Delta t} \left( \frac{1}{2\mu} \sum_i \underbrace{\pi_i} e^{iH_{\rm q}\Delta t} e^{iH_{\rm cl}\Delta t} \underbrace{\pi_i} + e^{iH_{\rm q}\Delta t} e^{iH_{\rm cl}\Delta t} \underbrace{H_{\rm cl}} \right) e^{iH_{\rm q}\Delta t} e^{iH_{\rm cl}\Delta t} \cdots . \tag{B.36}$$

With this arrangement, when we carry out the rest of the path integral procedure, we end up with

the integrand

$$\frac{\mu}{2\Delta t^2}\sum_i \left[\sigma_i(t+2\Delta t)-\sigma_i(t+\Delta t)\right]\left[\sigma_i(t+\Delta t)-\sigma_i(t)\right]+\sum_{(i_1\cdots i_p)} J_{i_1\cdots i_p}\sigma_{i_1}(t)\cdots\sigma_{i_p}(t). \quad \text{(B.37)}$$

While these expressions are for the lower contour, those for the upper contour are exactly analogous. Thus the $n_u n_l$ term of the SFF becomes

$$\begin{aligned}&\mathrm{Tr}H^{n_u}e^{-iHT}\mathrm{Tr}H^{n_l}e^{iHT}\\ &\qquad=\int \mathcal{D}\sigma^N e^{iS}\left(\frac{\Delta t}{T}\right)^{n_u+n_l}\sum_{t_1\cdots t_{n_u}} H_u(t_1)\cdots H_u(t_{n_u})\sum_{t_1\cdots t_{n_l}} H_l(t_1)\cdots H_l(t_{n_l}),\end{aligned} \quad \text{(B.38)}$$

where the action $S$ is, in continuum notation,

$$\begin{aligned}S=\int_0^T dt\sum_\alpha(-1)^\alpha\Bigg[\sum_i\left(\frac{\mu}{2}\big(\partial_t\sigma_i(t)\big)^2-\frac{z(t)}{2}\big(\sigma_{i\alpha}(t)^2-1\big)\right)\\ -\sum_{(i_1\cdots i_p)} J_{i_1\cdots i_p}\sigma_{i_1\alpha}(t)\cdots\sigma_{i_p\alpha}(t)\Bigg],\end{aligned} \quad \text{(B.39)}$$

and we've defined (in a slight abuse of notation) the random function

$$\begin{aligned}H_\alpha(t)\equiv\frac{\mu}{2\Delta t^2}\sum_i\left[\sigma_{i\alpha}(t+2\Delta t)-\sigma_{i\alpha}(t+\Delta t)\right]\left[\sigma_{i\alpha}(t+\Delta t)-\sigma_{i\alpha}(t)\right]\\ +\sum_{(i_1\cdots i_p)} J_{i_1\cdots i_p}\sigma_{i_1\alpha}(t)\cdots\sigma_{i_p\alpha}(t).\end{aligned} \quad \text{(B.40)}$$

Eq. (B.38) must now be averaged over the random couplings. Due to the term of the action

linear in $J_{i_1\cdots i_p}$, each coupling acquires an expectation value:

$$\overline{J_{i_1\cdots i_p}e^{iS}} = -\frac{iJ^2(p-1)!}{C_{i_1\cdots i_p}N^{p-1}}\int_0^T dt' \sum_{\alpha'}(-1)^{\alpha'}\sigma_{i_1\alpha'}(t')\cdots\sigma_{i_p\alpha'}(t'). \tag{B.41}$$

Thus in the average of Eq. (B.38), one contribution is the fully disconnected term in which each insertion of $H_\alpha(t)$ gets replaced by its mean value, denoted $N\epsilon_\alpha(t)$:

$$N\epsilon_\alpha(t) \equiv \frac{\mu}{2}\sum_i \left[\partial_t\sigma_{i\alpha}(t^+)\right]\left[\partial_t\sigma_{i\alpha}(t)\right] - \frac{iNJ^2}{p}\int_0^T dt' \sum_{\alpha'}(-1)^{\alpha'}\left(\frac{1}{N}\sum_i \sigma_{i\alpha}(t)\sigma_{i\alpha'}(t')\right)^p. \tag{B.42}$$

In fact, all other contributions, which come from contractions of pairs of couplings about their means, are necessarily subleading in $1/N$: whereas the disconnected term contributes one factor of $N$ for every insertion of $H$, each contraction accounts for two insertions but contributes only one factor of $N$ between the two. Thus, to leading order,

$$\overline{\mathrm{Tr}H^{n_u}e^{-iHT}\mathrm{Tr}H^{n_l}e^{iHT}} \sim \int \mathcal{D}\sigma^N \left(N\epsilon'_u\right)^{n_u}\left(N\epsilon'_l\right)^{n_l}\overline{e^{iS}}, \tag{B.43}$$

where we've defined $\epsilon'_\alpha \equiv T^{-1}\int dt\epsilon_\alpha(t)$.

Returning to the SFF, we see that we can simply resum the series for the filter functions, but now with the *deterministic* function $\epsilon'$ in place of the random operator $H$:

$$\mathrm{SFF}(T,f) = \int \mathcal{D}\sigma^N f\left(\epsilon'_u\right) f\left(\epsilon'_l\right)\overline{e^{iS}}. \tag{B.44}$$

The average of $e^{iS}$ is treated exactly as in Subsec. 3.2.2, albeit with the extra index $\alpha$. It results in the effective action given by Eq. (3.45) in the main text. Note that, when we introduce the order

parameter $G_{\alpha\alpha'}(t,t')$ (see Eq. (3.23)), we can also substitute it into the expression for $\epsilon'$, giving the function

$$\epsilon_\alpha[G] \equiv -\frac{\mu}{2T}\int dt\partial_t^2 G_{\alpha\alpha}(t^+,t) - \frac{iJ^2}{pT}\int_0^T dtdt' \sum_{\alpha'}(-1)^{\alpha'} G_{\alpha\alpha'}(t,t')^p. \tag{B.45}$$

The only difference from Eq. (3.46) is that the latter assumes $G_{\alpha\alpha'}(t,t')$ is time-translation invariant. This is true of the saddle-point value, and the replacement is justified here because $f(\epsilon[G])$ is independent of $N$ and is thus evaluated solely at the saddle point in the large-$N$ limit. We have therefore arrived at Eqs. (3.44) through (3.46) of the main text.

# Appendix C: Appendices to Chapter 4

## C.1 Sigma model calculation for the eigenstate correlation function

The correlation function $C_{nn'm'm}(\omega)$ defined in Eq. 3 in the main text can be obtained from the generating function [250]

$$\mathcal{Z}[J] = \left\langle \int_0^{2\pi} \frac{d\phi}{2\pi} \frac{\det\left(1-\alpha e^{i\phi}U+J^{+}\right)\left(1-\beta e^{-i\phi}U^{\dagger}+J^{-}\right)}{\det\left(1-\alpha e^{i\phi}U\right)\left(1-\beta e^{-i\phi}U^{\dagger}\right)} \right\rangle, \tag{C.1}$$

by taking derivatives with respect to the matrix elements of sources $J^{\pm}$ and setting them to zero at the end:

$$\begin{aligned} \bar{C}_{nn'm'm}(\alpha\beta) &\equiv \left.\frac{\partial^2 \mathcal{Z}[J]}{\partial J^{+}_{n'n}\partial J^{-}_{mm'}}\right|_{J=0} = \sum_{t=0}^{\infty} C_{nn'm'm}(t)\,(\alpha\beta)^t\,, \\ C_{nn'm'm}(\omega) &= \frac{1}{2\pi}\left(\bar{C}_{nn'm'm}(\alpha\beta) + \bar{C}^{*}_{n'nmm'}(\alpha\beta) - \delta_{nn'}\delta_{mm'}\right). \end{aligned} \tag{C.2}$$

$\alpha$ and $\beta$ are complex numbers whose product is $\alpha\beta = e^{i\omega-\eta}$. The positive infinitesimal $1/M \ll \eta \ll 1$ serves as a small parameter for the following perturbative calculation, and broaden the $2\pi$-periodic Dirac delta function $\delta_{2\pi}(\omega)$ in Eq. 3 of the main text into a Lorentzian function $\sum_{n=-\infty}^{\infty} e^{i\omega n-|n|\eta}/2\pi$ of width $\eta$.

The generating function $\mathcal{Z}[J]$ can be expressed as a Gaussian superintegral, where the

integration variable possesses a structure within advanced/retarded space (contributing to the part associated with forward/backward time evolution $U/U^\dagger$) and within the fermion/boson space (corresponding to the numerator/denominator). This superintegral is then converted by color-flavor transformation into a sigma model representation [248],

$$\begin{aligned} \mathcal{Z}[J] &= \det\left((1+J^+)(1+J^-)\right)\left\langle \int D(\tilde{Z},Z)\exp(-S[\tilde{Z},Z])\right\rangle, \\ S[\tilde{Z},Z] &= -\mathrm{STr}\ln(1-\tilde{Z}Z)+\mathrm{STr}\ln\left(1-\alpha\beta\tilde{Z}M_J^+\mathbf{U}ZM_J^-\mathbf{U}^\dagger\right), \end{aligned} \tag{C.3}$$

which is exact and applicable to any time periodic system with a Floquet operator $U$. Here $M_J^\pm$ is a block diagonal matrix in the boson-fermion space and takes the form of $M_J^\pm = \begin{bmatrix} 1 & 0 \\ 0 & (1+J^\pm)^{-1} \end{bmatrix}_{\mathsf{BF}}$. $Z_{ij}^{ab}$ and $\tilde{Z}_{ij}^{ab}$ are supermatrices which act in both the boson-fermion space (labeled by superscripts $a,b = \mathsf{B},\mathsf{F}$) and the Hilbert space of the entire system (labeled by subscripts $i,j = 1,2,..,N$). They are constrained by the relations: $\tilde{Z}^{\mathsf{FF}} = -Z^{\mathsf{FF}\,\dagger}$, $\tilde{Z}^{\mathsf{BB}} = Z^{\mathsf{BB}\,\dagger}$, and $|Z^{\mathsf{BB}\,\dagger}Z^{\mathsf{BB}}| < 1$. STr is used to indicate the supertrace over the product of boson-fermion space and Hilbert space.

For the current model in the limit of large onsite Hilbert space dimension $q$, we could perform an expansion in powers of $Z$ and $\tilde{Z}$ in $S[\tilde{Z},Z]$ and focus on the Gaussian order fluctuations. The applicability of this perturbative treatment in the current model is analogous to that for the $q$-orbital version of the Wegner model [251, 252] in the limit of large $q$. Additionally, the $\eta = -\mathrm{Re}\ln(\alpha\beta)$ also serves as a small parameter in this perturbative calculation [248]. From the leading order quadratic fluctuations one obtain the smoothed eigenstate correlation function with the $2\pi$-periodic Dirac delta function replaced with a Lorenzian function of width $\eta$.

Up to quadratic order in expansion of $Z$ and $\tilde{Z}$, the action acquires the form

$$S_2[\tilde{Z},Z]=\sum_{i,j,i',j'}\sum_{a,b}s_a\tilde{Z}^{ab}_{ij}Z^{ba}_{j'i'}\Bigg[\delta_{ii'}\delta_{jj'} \\ -\alpha\beta\left(\delta_{ji}-\delta_{b\mathsf{F}}J^{+}_{ji}\right)\left(\delta_{i'j'}-\delta_{a\mathsf{F}}J^{-}_{i'j'}\right)\left(\delta_{B(i),B(j')}\frac{1+Ne^{-\sigma^2}}{M(N+1)}+\frac{1-e^{-\sigma^2}}{N+1}\right)\Bigg]. \quad \text{(C.4)}$$

Here $s_{\mathsf{B/F}}=\pm 1$ and we have ignored terms of higher order in source $J^{\pm}$ which are not relevant to the eigenstate correlation function under study.

This action for the quadratic fluctuations can be rewritten as

$$S_2[X,Y,\tilde{X},\tilde{Y}]=\sum_{a,b=\mathsf{B},\mathsf{F}}\sum_{i\neq j}s_a\tilde{Y}^{ab}_{ij}Y^{ba}_{ji}+\frac{1}{M}\sum_{a,b=\mathsf{B},\mathsf{F}}\sum_{m=1}^{P}\sum_{k=0}^{M-1}s_a\tilde{X}^{ab}_{m}(k)X^{ba}_{m}(k) \\ -\sum_{a,b}\sum_{m,m'=1}^{P}s_a\alpha\beta\left(\delta_{m,m'}\frac{1+Ne^{-\sigma^2}}{M(N+1)}+\frac{1-e^{-\sigma^2}}{N+1}\right) \\ \times\left(\tilde{X}^{ab}_{m}(0)-\delta_{bF}\frac{1}{M}\sum_{k=0}^{M-1}\tilde{X}^{aF}_{m}(k)\sum_{i=1}^{M}e^{-i2\pi ki/M}J^{+}_{(m-1)M+i,(m-1)M+i} \right. \\ \left. -\delta_{bF}\sum_{i\neq j=1}^{N}\tilde{Y}^{aF}_{ij}J^{+}_{ji}\delta_{B(i),m}\right) \\ \times\left(X^{ba}_{m'}(0)-\delta_{aF}\frac{1}{M}\sum_{k=0}^{M-1}X^{bF}_{m'}(k)\sum_{i'=1}^{M}e^{i2\pi ki'/M}J^{-}_{(m'-1)M+i',(m'-1)M+i'} \right. \\ \left. -\delta_{aF}\sum_{i'\neq j'=1}^{N}Y^{bF}_{j'i'}J^{-}_{i'j'}\delta_{B(j'),m'}\right). \quad \text{(C.5)}$$

Here $X_m(k)$ is the Fourier transform of the diagonal matrix elements of $Z$ within the block

$m$, and $Y$ represents the off-diagonal matrix elements of $Z$:

$$X_m^{ab}(k) = \sum_{j=1}^{M} Z^{ab}_{(m-1)M+j,(m-1)M+j} e^{-i2\pi kj/M}, \quad \tilde{X}_m^{ab}(k) = \sum_{j=1}^{M} \tilde{Z}^{ab}_{(m-1)M+j,(m-1)M+j} e^{i2\pi kj/M},$$
$$Y_{ij}^{ab} = (1-\delta_{ij}) Z_{ij}^{ab}, \qquad \tilde{Y}_{ij}^{ab} = (1-\delta_{ij}) \tilde{Z}_{ij}^{ab}. \tag{C.6}$$

We have used the notation that $B(i) \equiv \lceil i/M \rceil$ and $b(i) \equiv i - (B(i)-1)M$.

Setting the source $J = 0$ in Eq. C.5, we obtain the bare propagators for $Y$ and $X(k)$ with nonzero momentum $k$,

$$\begin{aligned} \left\langle Y_{j'i'}^{b'a'} \tilde{Y}_{ij}^{ab} \right\rangle_0 &= s_a s_b \left\langle \tilde{Y}_{ij}^{ab} Y_{j'i'}^{b'a'} \right\rangle_0 = s_a \delta_{ii'} \delta_{jj'} \delta_{aa'} \delta_{bb'}, \\ \left\langle X_{m'}^{b'a'}(k') \tilde{X}_m^{ab}(k) \right\rangle_0 &= s_a s_b \left\langle \tilde{X}_m^{ab}(k) X_{m'}^{b'a'}(k') \right\rangle_0 = s_a M \delta_{kk'} \delta_{aa'} \delta_{bb'} \delta_{mm'}. \end{aligned} \tag{C.7}$$

Here the angular bracket with subscript 0 represents the functional average with respect the quadratic action $S_2$ at $J = 0$. For the zero momentum mode $X(k = 0)$, we instead have

$$\begin{aligned} \left\langle X_{m'}^{b'a'}(0) \tilde{X}_m^{ab}(0) \right\rangle_0 &= s_a s_b \left\langle \tilde{X}_m^{ab}(0) X_{m'}^{b'a'}(0) \right\rangle_0 \\ &= \begin{cases} \delta_{aa'} \delta_{bb'} s_a M \dfrac{(1-\alpha\beta) + \alpha\beta \frac{(1-e^{-\sigma^2})M}{N+1}}{\left(1 - \alpha\beta \frac{1+Ne^{-\sigma^2}}{1+N}\right)(1-\alpha\beta)}, & m = m', \\ \delta_{aa'} \delta_{bb'} s_a M \dfrac{\alpha\beta \frac{(1-e^{-\sigma^2})M}{N+1}}{\left(1 - \alpha\beta \frac{1+Ne^{-\sigma^2}}{1+N}\right)(1-\alpha\beta)}, & m \neq m'. \end{cases} \end{aligned} \tag{C.8}$$

Considering only the contribution from the quadratic fluctuations, the smoothed correlation

function $\bar{C}_{nn'm'm}(\alpha\beta)$ can be obtained from

$$\bar{C}_{nn'm'm}(\alpha\beta) = \left\langle \left( \delta_{nn'}\delta_{mm'} - \frac{\partial^2 S_2}{\partial J^+_{n'n}\partial J^-_{mm'}} + \frac{\partial S_2}{\partial J^+_{n'n}}\frac{\partial S_2}{\partial J^-_{mm'}} - \delta_{nn'}\frac{\partial S_2}{\partial J^-_{mm'}} - \delta_{mm'}\frac{\partial S_2}{\partial J^+_{n'n}} \right) \Big|_{J=0} \right\rangle_0 . \tag{C.9}$$

We now evaluate each term in the expression above separately using Eqs. C.5, C.7, and C.8:

$$\begin{aligned} \left\langle \frac{\partial S_2}{\partial J^+_{n'n}}\Big|_{J=0} \right\rangle_0 &= \delta_{nn'} \sum_a \sum_{i=1}^{P} s_a \alpha\beta \frac{1}{M} \left( \delta_{B(n),i} \frac{1+Ne^{-\sigma^2}}{M(N+1)} + \frac{1-e^{-\sigma^2}}{N+1} \right) \left\langle \tilde{X}^{a\mathsf{F}}_{B(n)}(0) X^{\mathsf{F}a}_i(0) \right\rangle_0 \\ &= 0, \end{aligned} \tag{C.10}$$

$$\begin{aligned} \left\langle \frac{\partial S_2}{\partial J^-_{mm'}}\Big|_{J=0} \right\rangle_0 &= \delta_{mm'} \sum_b \sum_{i=1}^{P} s_F \alpha\beta \frac{1}{M} \left( \delta_{i,B(m)} \frac{1+Ne^{-\sigma^2}}{M(N+1)} + \frac{1-e^{-\sigma^2}}{N+1} \right) \left\langle \tilde{X}^{\mathsf{F}b}_i(0) X^{b\mathsf{F}}_{B(m)}(0) \right\rangle_0 \\ &= 0, \end{aligned} \tag{C.11}$$

$$
\begin{aligned}
\left\langle \frac{\partial^2 S^{(s)}_{\mathsf{eff}}}{\partial J^+_{n'n}\partial J^-_{mm'}}\Big|_{J=0} \right\rangle_0 \;=\; & -(1-\delta_{nn'})(1-\delta_{mm'}) s_F \alpha\beta \left( \delta_{B(n),B(m')} \frac{1+Ne^{-\sigma^2}}{M(N+1)} + \frac{1-e^{-\sigma^2}}{N+1} \right) \\
& \qquad \times \left\langle \tilde{Y}^{\mathsf{FF}}_{nn'} Y^{\mathsf{FF}}_{m'm} \right\rangle_0 \\
& - \delta_{nn'}\delta_{mm'} \sum_{k=0}^{M-1} s_F \alpha\beta \frac{1}{M^2} \left( \delta_{B(n),B(m)} \frac{1+Ne^{-\sigma^2}}{M(N+1)} + \frac{1-e^{-\sigma^2}}{N+1} \right) \\
& \qquad \times \left\langle \tilde{X}^{\mathsf{FF}}_{B(n)}(k) X^{\mathsf{FF}}_{B(m)}(k) \right\rangle_0 e^{-i2\pi k(b(n)-b(m))/M} \\
= \; & -\delta_{nn'}\delta_{mm'}\delta_{B(n),B(m)} (\alpha\beta)^2 \frac{(1+Ne^{-\sigma^2})}{M(N+1)^2} \\
& \qquad \times \frac{(1-\alpha\beta)\frac{1}{M}(1+Ne^{-\sigma^2}) + (2-\alpha\beta)(1-e^{-\sigma^2})}{\left(1-\alpha\beta\frac{1+Ne^{-\sigma^2}}{1+N}\right)(1-\alpha\beta)} \\
& - \delta_{nn'}\delta_{mm'} (\alpha\beta)^2 \frac{(1-e^{-\sigma^2})^2}{(N+1)^2\left(1-\alpha\beta\frac{1+Ne^{-\sigma^2}}{1+N}\right)(1-\alpha\beta)} \\
& \qquad - \delta_{nm}\delta_{n'm'} \alpha\beta \left( \delta_{B(n),B(n')} \frac{1+Ne^{-\sigma^2}}{M(N+1)} + \frac{1-e^{-\sigma^2}}{N+1} \right),
\end{aligned}
\tag{C.12}
$$

$$
\begin{aligned}
\left\langle \frac{\partial S_2}{\partial J^+_{n'n}} \frac{\partial S_2}{\partial J^-_{mm'}}\bigg|_{J=0} \right\rangle_0 &= \delta_{nn'}\delta_{mm'} \sum_{i,j=1}^{P} \sum_{a,b} s_a s_F (\alpha\beta)^2 \frac{1}{M^2} \left( \delta_{B(n),i} \frac{1+Ne^{-\sigma^2}}{M(N+1)} + \frac{1-e^{-\sigma^2}}{N+1} \right) \\
&\times \left( \delta_{B(m'),j} \frac{1+Ne^{-\sigma^2}}{M(N+1)} + \frac{1-e^{-\sigma^2}}{N+1} \right) \left\langle \tilde{X}^{aF}_{B(n)}(0) X^{Fa}_i(0) \right\rangle \left\langle \tilde{X}^{Fb}_j(0) X^{bF}_{B(m)}(0) \right\rangle \\
&+ \sum_{i,j=1}^{P} (\alpha_F \beta_F)^2 \left( \delta_{B(n),i} \frac{1+Ne^{-\sigma^2}}{M(N+1)} + \frac{1-e^{-\sigma^2}}{N+1} \right) \\
&\times \left( \delta_{B(m'),j} \frac{1+Ne^{-\sigma^2}}{M(N+1)} + \frac{1-e^{-\sigma^2}}{N+1} \right) \left\langle X^{FF}_i(0) \tilde{X}^{FF}_j(0) \right\rangle \\
&\times \Bigg[ \delta_{nn'}\delta_{mm'} \frac{1}{M^2} \sum_{k=0}^{M-1} \left\langle \tilde{X}^{FF}_{B(n)}(k) X^{FF}_{B(m)}(k) \right\rangle e^{-i2\pi k(b(n)-b(m))/M} \\
&+ (1-\delta_{nn'})(1-\delta_{mm'}) \left\langle \tilde{Y}^{FF}_{nn'} Y^{FF}_{m'm} \right\rangle \Bigg]
\end{aligned}
$$

$$
\begin{aligned}
&= \delta_{nn'}\delta_{mm'}\delta_{B(n),B(m)} (\alpha\beta)^3 (Ne^{-\sigma^2}+1) \\
&\times \frac{(1-\alpha\beta)\frac{1}{M}(Ne^{-\sigma^2}+1)^2 + (2-\alpha\beta)(1-e^{-\sigma^2})(Ne^{-\sigma^2}+1) + (N+1)\left(1-e^{-\sigma^2}\right)}{(1+N)^3 M \left(1-\alpha\beta\frac{1+Ne^{-\sigma^2}}{1+N}\right)^2 (1-\alpha\beta)} \\
&+ \delta_{nn'}\delta_{mm'} (\alpha\beta)^3 \frac{\left(1-e^{-\sigma^2}\right)^2 \left[(1-\alpha\beta)(Ne^{-\sigma^2}+1) + (N+1)\right]}{(N+1)^3 \left(1-\alpha\beta\frac{1+Ne^{-\sigma^2}}{1+N}\right)^2 (1-\alpha\beta)^2} \\
&+ \delta_{nm}\delta_{n'm'}\delta_{B(n),B(n')} (\alpha\beta)^2 \frac{(Ne^{-\sigma^2}+1)^2}{M(N+1)^2 \left(1-\alpha\beta\frac{1+Ne^{-\sigma^2}}{1+N}\right)} \\
&+ \delta_{nm}\delta_{n'm'} (\alpha\beta)^2 \frac{\left(1-e^{-\sigma^2}\right)\left[(1-\alpha\beta)(Ne^{-\sigma^2}+1) + (N+1)\right]}{(N+1)^2 \left(1-\alpha\beta\frac{1+Ne^{-\sigma^2}}{1+N}\right)(1-\alpha\beta)}.
\end{aligned}
\tag{C.13}
$$

Combining everything and using Eq. C.9, we find that the contribution from the Gaussian

fluctuations assumes the form

$$\bar{C}_{nn'm'm}(\alpha\beta) = \delta_{nn'}\delta_{mm'}\left(1+\bar{c}_1(\alpha\beta)\right) + \delta_{nm}\delta_{n'm'}\bar{c}_2(\alpha\beta) \tag{C.14}$$

$$+ \delta_{nn'}\delta_{mm'}\delta_{B(n),B(m)}\bar{c}_3(\alpha\beta) + \delta_{nm}\delta_{n'm'}\delta_{B(n),B(n')}\bar{c}_4(\alpha\beta),$$

$$\bar{c}_1(\alpha\beta) = (1-e^{-\sigma^2})^2 \frac{\frac{1}{(N+1)^2}(\alpha\beta)^2}{\left(1-\alpha\beta\frac{1+Ne^{-\sigma^2}}{1+N}\right)^2(1-\alpha\beta)^2}, \tag{C.15}$$

$$\bar{c}_2(\alpha\beta) = (1-e^{-\sigma^2}) \frac{\frac{1}{N+1}\alpha\beta}{\left(1-\alpha\beta\frac{1+Ne^{-\sigma^2}}{1+N}\right)(1-\alpha\beta)}, \tag{C.16}$$

$$\bar{c}_3(\alpha\beta) = \frac{1+Ne^{-\sigma^2}}{1+N} \frac{\frac{1}{M^2}(1-\alpha\beta)(\alpha\beta)^2\frac{1+Ne^{-\sigma^2}}{1+N} + \frac{2}{(N+1)M}(\alpha\beta)^2(1-e^{-\sigma^2})}{\left(1-\alpha\beta\frac{1+Ne^{-\sigma^2}}{1+N}\right)^2(1-\alpha\beta)}, \tag{C.17}$$

$$\bar{c}_4(\alpha\beta) = \frac{1+Ne^{-\sigma^2}}{1+N} \frac{\frac{1}{M}\alpha\beta}{\left(1-\alpha\beta\frac{1+Ne^{-\sigma^2}}{1+N}\right)}. \tag{C.18}$$

From now on, we will focus on the regime where $\sigma \ll \sqrt{\ln N}$. From the expression above, one can deduce that outside this regime where $\sigma \gg \sqrt{\ln N}$, the eigenstate correlation function of the current model becomes approximately that of the circular unitary ensemble, and as a result the model becomes chaotic. We then simplify the expression above using $N \gg 1$ and $e^{-\sigma^2} \gg N^{-1}$:

$$\begin{aligned}
\bar{c}_1(\alpha\beta) &= \frac{\frac{1}{N^2}(\alpha\beta)^2(1-e^{-\sigma^2})^2}{\left(1-\alpha\beta e^{-\sigma^2}\right)^2(1-\alpha\beta)^2}, \\
\bar{c}_2(\alpha\beta) &= \frac{\frac{1}{N}\alpha\beta(1-e^{-\sigma^2})}{\left(1-\alpha\beta e^{-\sigma^2}\right)(1-\alpha_F\beta_F)}, \\
\bar{c}_3(\alpha\beta) &= e^{-\sigma^2}\frac{\frac{1}{M^2}(1-\alpha\beta)(\alpha\beta)^2e^{-\sigma^2} + \frac{2}{NM}(\alpha\beta)^2(1-e^{-\sigma^2})}{\left(1-\alpha\beta e^{-\sigma^2}\right)^2(1-\alpha\beta)}, \\
\bar{c}_4(\alpha\beta) &= \frac{\frac{1}{M}\alpha\beta e^{-\sigma^2}}{(1-\alpha\beta e^{-\sigma^2})}.
\end{aligned} \tag{C.19}$$

Making use of this result, we obtain the smoothed correlation function $C_{nn'm'm}(\omega)$ defined by Eq. 3 in the main text but with the $2\pi$ Dirac delta function replaced with the Loranzian of width

$\eta = -\operatorname{Re}\ln(\alpha\beta)$:

$$\begin{aligned} C_{nn'm'm}(\omega) =& \frac{1}{2\pi}\left(\bar{C}_{nn'm'm}(e^{i\omega}) + \bar{C}^{*}_{n'nmm'}(e^{i\omega}) - \delta_{nn'}\delta_{mm'}\right) \\ =& \frac{1}{2\pi}\left[\delta_{nn'}\delta_{mm'}\left(1 + c_1(\omega)\right) + \delta_{nm}\delta_{n'm'}c_2(\omega)\right. \\ & \left. + \delta_{nn'}\delta_{mm'}\delta_{B(n),B(m)}c_3(\omega) + \delta_{nm}\delta_{n'm'}\delta_{B(n),B(n')}c_4(\omega)\right], \end{aligned} \tag{C.20}$$

where

$$c_1(\omega) = -\frac{(1-e^{-\sigma^2})^2}{N^2}\frac{\left(1+e^{-2\sigma^2}\right)\cos(\omega) - 2e^{-\sigma^2}}{\left[\left(e^{-\sigma^2}-\cos(\omega)\right)^2 + \sin^2(\omega)\right]^2\left(1-\cos(\omega)\right)}, \tag{C.21}$$

$$c_2(\omega) = -\frac{1}{N}\frac{(1-e^{-\sigma^2})(1+e^{-\sigma^2})}{\left(e^{-\sigma^2}-\cos(\omega)\right)^2 + \sin^2(\omega)}, \tag{C.22}$$

$$\begin{aligned} c_3(\omega) = 2\frac{e^{-\sigma^2}}{MN}&\left[\left(e^{-\sigma^2}-\cos(\omega)\right)^2 + \sin^2(\omega)\right]^{-2}\left\{Pe^{-\sigma^2}\left[\left(e^{-\sigma^2}-\cos(\omega)\right)^2 - \sin^2(\omega)\right]\right. \\ &\left. + (1-e^{-\sigma^2})\left[2\left(e^{-\sigma^2}-\cos(\omega)\right) + e^{-2\sigma^2} - 1\right]\right\}, \end{aligned} \tag{C.23}$$

$$c_4(\omega) = -2\frac{e^{-\sigma^2}}{M}\frac{e^{-\sigma^2}-\cos(\omega)}{\left(e^{-\sigma^2}-\cos(\omega)\right)^2 + \sin^2(\omega)}. \tag{C.24}$$

We note that the contribution from the fluctuations beyond quadratic order needs to be taken into account to study the correlation function for nearby quasienergy eigenstates with small energy separation $\omega \lesssim \eta$.

It is straightforward to see from its definition that this correlation function of quasienergy

eigenstates $C_{nn'm'm}(\omega)$ is related to the correlation function of the quasienergy:

$$\begin{aligned} R_2(\omega) &\equiv \left\langle \sum_{\nu,\mu} \delta\left(\omega - E_\nu + E_\mu\right) \right\rangle \\ &= \sum_{nm} C_{nnmm}(\omega) \\ &= \frac{1}{2\pi}\left[N^2\left(1 + c_1(\omega)\right) + N c_2(\omega) + NM c_3(\omega) + N c_4(\omega)\right]. \end{aligned} \tag{C.25}$$

Therefore, we can deduce from Eq. C.20 the smoothed quasienergy correlation function

$$R_2(\omega) = \frac{1}{2\pi}\left\{ N^2 + 2(P-1)\frac{e^{-\sigma^2}\left[\left(1 + e^{-2\sigma^2}\right)\cos(\omega) - 2e^{-\sigma^2}\right]}{\left[\left(e^{-\sigma^2} - \cos(\omega)\right)^2 + \sin^2(\omega)\right]^2} - \frac{1}{1-\cos(\omega)} \right\}. \tag{C.26}$$

In the time representation, the eigenstate correlation function for discrete time $t \geq 0$ assumes the form

$$\begin{aligned} C_{nn'm'm}(t) = \delta_{nn'}\delta_{mm'}\left(\delta_{t,0} + c_1(t)\right) + \delta_{nm}\delta_{n'm'}c_2(t) \\ + \delta_{nn'}\delta_{mm'}\delta_{B(n),B(m)}c_3(t) + \delta_{nm}\delta_{n'm'}\delta_{B(n),B(n')}c_4(t), \end{aligned} \tag{C.27}$$

where

$$\begin{aligned} c_1(t) &= \frac{1}{N^2}\left[t(1 + e^{-t\sigma^2}) - \frac{1 + e^{-\sigma^2}}{1 - e^{-\sigma^2}}(1 - e^{-t\sigma^2})\right], \\ c_2(t) &= \frac{1}{N}\left(1 - e^{-t\sigma^2}\right), \\ c_3(t) &= \frac{1}{N^2}\left[\frac{2Pe^{-\sigma^2}}{1 - e^{-\sigma^2}}\left(1 - e^{-t\sigma^2}\right) + (P^2 - 2P)te^{-t\sigma^2} - P^2(e^{-t\sigma^2} - \delta_{t,0})\right] \\ c_4(t) &= \frac{1}{M}\left(e^{-t\sigma^2} - \delta_{t,0}\right). \end{aligned} \tag{C.28}$$

The spectral form factor is given by

$$\begin{aligned} K(t) &= \sum_{nm} C_{nnmm}(t) = N^2\left(\delta_{t,0} + c_1(t)\right) + Nc_2(t) + NMc_3(t) + Nc_4(t) \\ &= N^2\delta_{t,0} + t + (P-1)te^{-\sigma^2 t}. \end{aligned} \tag{C.29}$$

We emphasize that these expressions are valid only at early times and higher order fluctuations are required around and beyond the block Heisenberg time.

In the limit $\sigma \to \infty$, the smoothed correlation functions become those of the CUE of dimension $N$:

$$\begin{aligned} c_1(\omega) &= \frac{1}{N^2}\left(1 - \frac{1}{2\sin^2(\omega/2)}\right), & c_1(t) &= \frac{1}{N^2}(t - 1 + \delta_{t,0}), \\ c_2(\omega) &= -\frac{1}{N}, & c_2(t) &= \frac{1}{N}(1 - \delta_{t,0}), \\ c_3(\omega) &= c_4(\omega) = 0, & c_3(t) &= c_4(t) = 0, \\ R_2(\omega) &= \frac{1}{2\pi}\left(N^2 - \frac{1}{2\sin^2(\omega/2)}\right), & K(t) &= N^2\delta_{t,0} + t. \end{aligned} \tag{C.30}$$

On the other hand, in the limit $\sigma \to 0$, the smoothed correlation functions become equivalent to those of a block CUE, i.e., a block diagonal matrix with statistically independent CUE matrices of dimension $M$ as its diagonal blocks,

$$\begin{aligned} c_1(\omega) &= c_2(\omega) = 0, & c_1(t) &= c_2(t) = 0, \\ c_3(\omega) &= \frac{1}{M^2}\left(1 - \frac{1}{2\sin^2(\omega/2)}\right), & c_3(t) &= \frac{1}{M^2}(t - 1 + \delta_{t,0}), \\ c_4(\omega) &= -\frac{1}{M}, & c_4(t) &= \frac{1}{M}(1 - \delta_{t,0}), \\ R_2(\omega) &= \frac{1}{2\pi}\left(N^2 - \frac{P}{2\sin^2(\omega/2)}\right), & K(t) &= N^2\delta_{t,0} + Pt. \end{aligned} \tag{C.31}$$

## C.2 Diagrammatic Calculation of the Spectral Form Factor

In this section we calculate the SFF

$$K(t) = \left\langle |\operatorname{tr} U^t|^2 \right\rangle \tag{C.32}$$

of the Floquet operator

$$U = V^\dagger \Theta V A. \tag{C.33}$$

We use the diagrammatic method for integration over the unitary group developed by Brouwer and Beenakker [253] to integrate over $V$. In this method the $V$ operators are represented by dotted lines connecting a filled dot to an empty dot and the $V^\dagger$ operators are represented by starred dotted lines connecting an empty dot to a filled dot. The dots correspond to operator indices. The $A$, $A^\dagger$, $\Theta$, and $\Theta^*$ operators are then represented by directed solid lines connecting dots of the same type. Calculation of the SFF involves the product of two traces, $\operatorname{tr} U^t$ and $\operatorname{tr}(U^\dagger)^t$. Each trace forms a loop consisting of alternating directed solid lines and dotted lines.

The average over $V$ has the form

$$\left\langle V_{i_1 j_1} \dots V_{i_m j_m} V^*_{i'_1 j'_1} \dots V^*_{i'_m j'_m} \right\rangle = \sum_{\Pi,\Pi'} X_{\Pi,\Pi'} \prod_{k=1}^{m} \delta_{i_k i'_{\Pi(k)}} \delta_{j_k j'_{\Pi'(k)}} \tag{C.34}$$

when the sequence $i'_1 \dots i'_m$ is a permutation of $i_1 \dots i_m$ and $j'_1 \dots j'_m$ is a permutation of $j_1 \dots j_m$. Otherwise the average vanishes. Performing the average requires summing over each of the possible permutations $\Pi$ and $\Pi'$. The coefficients $X_{\Pi,\Pi'}$ turn out to only depend on the cycle structure of the permutation $\Pi^{-1}\Pi'$ [358]. This allows us to write the coefficients as $X_{c_1,\dots,c_k}$,

where the $c_1, \ldots, c_k$ are the lengths of the cycles in the factorization of $\Pi^{-1}\Pi'$.

Each permutation is represented diagrammatically by connecting each filled (empty) dot connected to a dotted line to a filled (empty) dot connected to a starred dotted line with a solid undirected line (colored red in our figures for convenience). This means that at time $t$ there are $[(2t)!]^2$ permutations to sum over to calculate the SFF. This would quickly make the calculation intractable, but we can make progress by considering only the leading order diagrams.

To determine the contribution of each diagram we consider the U-cycles and T-cycles that are formed. The T-cycles are the loops of alternating directed solid lines and undirected solid lines. Each T-cycle corresponds to a trace of the product of the operators in the cycle following the direction of the cycle. The U-cycles are the loops of undirected solid lines and dotted lines. The length of a U-cycle is half the number of dotted lines within it. It is the lengths of the U-cycles which determine the prefactor $X_{c_1,\ldots,c_k}$. These coefficients are given by the recursion relation [358]

$$NX_{c_1,\ldots,c_k} + \sum_{p+q=c_1} X_{p,q,c_2,\ldots,c_k} + \sum_{j=2}^{k} c_j X_{c_1+c_j,c_2,\ldots,c_{j-1},c_{j+1},\ldots,c_k} = \delta_{c_1 1} X_{c_2,\ldots,c_k}. \tag{C.35}$$

We will make use of the large $N$ expansion of these factors

$$X_{c_1,\ldots,c_k} = \prod_{j=1}^{k} X_{c_j} + O(N^{k-4t-2}), \tag{C.36}$$

$$X_c = (-1)^{c-1} N^{1-2c} c^{-1} \binom{2c-2}{c-1} + O(N^{-1-2c}). \tag{C.37}$$

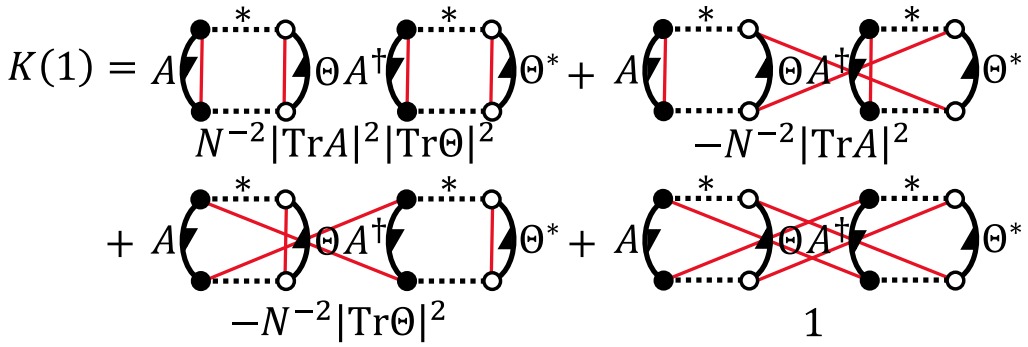



Figure C.1: The diagrams contributing to the spectral form factor at $t = 1$.

Note that $c^{-1}\binom{2c-2}{c-1}$ are the Catalan numbers. In particular we will use that

$$X_1 \approx N^{-1}, \tag{C.38}$$

$$X_2 \approx -N^{-3}. \tag{C.39}$$

We now examine the diagrams contributing to $K(1)$ shown in Fig. C.1. Recall that $A$ is a block diagonal matrix, with each block being a random CUE matrix with some overall phase. This means that a trace involving only $A$ and $A^\dagger$ operators will be much less than $O(N)$ if it does not come to a trace over the identity. On the other hand, traces involving only $\Theta$ and $\Theta^*$ are $O(N)$ for finite $\sigma$. This means that the second diagram in Fig. C.1 is subleading and can be neglected in the large $N$ limit.

Fig. C.2 shows only the leading order diagrams contributing to the SFF at $t = 2$. At this point we can see some patterns emerge that also hold for later times.

- No T-cycles are formed that combine $A$ and $A^\dagger$ operators with $\Theta$ and $\Theta^*$ operators.
- The leading order diagrams do not contain U-cycles with length greater than 2.
- Diagrams that form T-cycles with multiple $\Theta$ or multiple $\Theta^*$ are not of leading order and

can therefore be neglected at large $N$.

- There are always $t$ diagrams with only U-cycles of length 1 that are fully connected in the sense that each dot is connected to a dot in the other trace, as seen in the first row of Fig. C.2. The contribution from these diagrams is $t$.

- There is always one fully disconnected diagram of leading order, as seen in the second row of Fig. C.2. The contribution of this diagram is $N^{-2t}|\operatorname{tr}\Theta|^{2t}|\operatorname{tr}A^t|^2$.

- Excepting the fully connected diagrams mentioned above, each leading order diagram that forms a T-cycle consisting of a $\Theta$ and a $\Theta^*$ operator has a diagram that approximately cancels it for large $N$ that is formed by combining or dividing U-cycles. We see this in the way the diagrams in the last two rows of Fig. C.2 cancel each other at large $N$. This means we can consider only diagrams in which each $\Theta$ and $\Theta^*$ is in a T-cycle by itself.

- For the remaining leading order diagrams there are $t\binom{t}{j}$ diagrams that connect $j$ $A$ operators to $j$ $A^\dagger$ operators for $j = 1, \ldots, t$. These are the diagrams shown in the third and fourth rows. For even $j$ these contributions are positive, while they are negative for odd $j$. The contribution of these diagrams is $N^{-2t}|\operatorname{tr}\Theta|^{2t}\sum_{j=1}^{t}(-1)^j\binom{t}{j} = -N^{-2t}|\operatorname{tr}\Theta|^{2t}t$.

With these considerations we can write the SFF as

$$K(t) = N^{-2t}\left\langle|\operatorname{tr}\Theta|^{2t}\right\rangle\left\langle|\operatorname{tr}A^t|\right\rangle^2 + t\left(1 - N^{-2t}\left\langle|\operatorname{tr}\Theta|^{2t}\right\rangle\right). \tag{C.40}$$

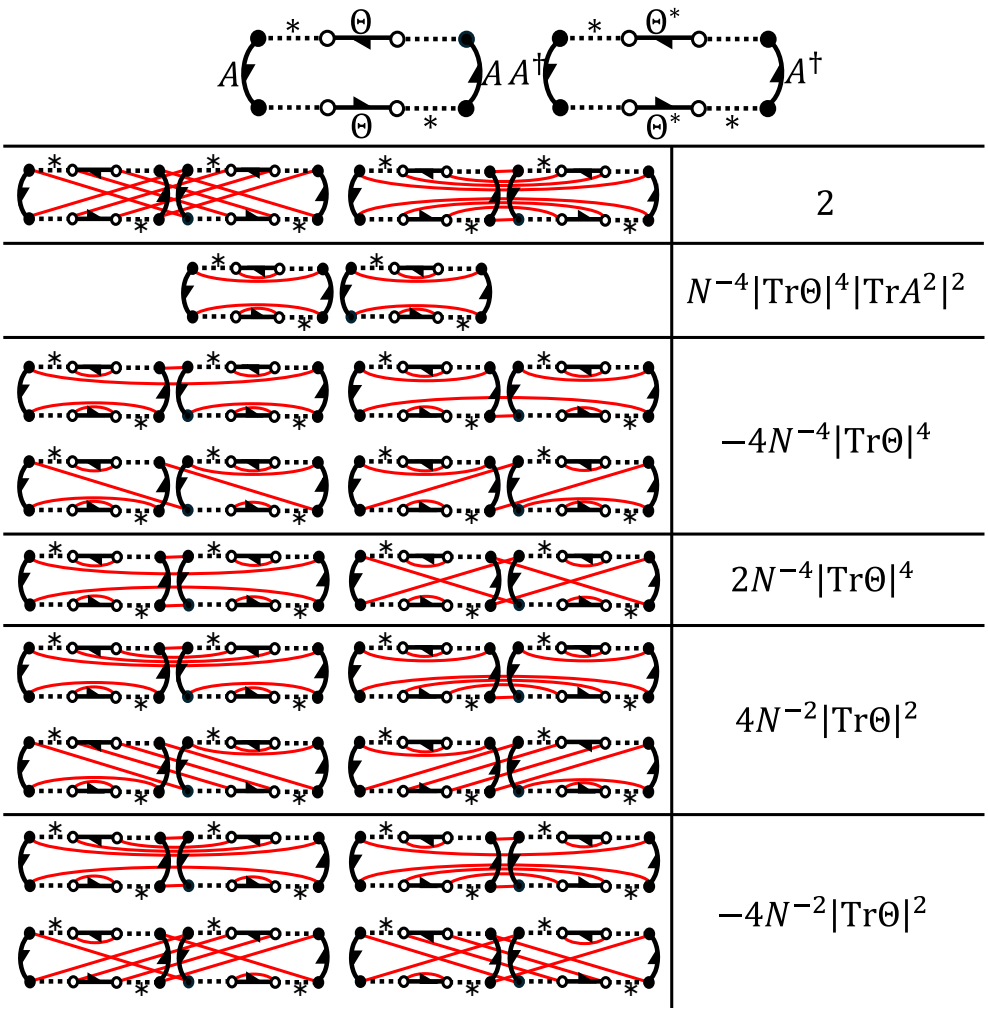



Figure C.2: The leading order diagrams contributing to the spectral form factor at $t = 2$. The base diagram without any contractions is shown at the top with the operators labeled. The labels are dropped in the other diagrams for clarity.

Considering the average over $A$ we see that

$$|\operatorname{tr} A^t|^2 = \sum_{i,j=1}^{P} e^{i(\phi_i - \phi_j)} |\operatorname{tr} W^t|^2. \tag{C.41}$$

Since the $\phi_i$ are uniformly distributed, only the $i = j$ terms do not vanish with averaging. $\langle |\operatorname{tr} W^t|^2 \rangle$ is simply the dimension $M$ CUE SFF, which approaches $\min(t, M)$ for large $M$ and

$t > 0$. So

$$\left\langle |\operatorname{tr} A^t| \right\rangle^2 = P \min(t, M). \tag{C.42}$$

We also note that $N^{-1} \operatorname{tr} \Theta$ has the form of an average, so

$$\lim_{N\to\infty} N^{-2t} \left\langle |\operatorname{tr} \Theta|^{2t} \right\rangle = e^{-\sigma^2 t}. \tag{C.43}$$

Putting this all together yields the SFF

$$K(t) = e^{-\sigma^2 t} P \min(t, M) + \left(1 - e^{-\sigma^2 t}\right) t + N^2 \delta_{t,0}. \tag{C.44}$$

The SFF may be tuned between the glassy result ($\sigma \to 0$) and the CUE result ($\sigma \to \infty$). For finite $\sigma$ the SFF will match the glassy result at early times then cross over to the CUE result at $t \sim \sigma^{-2}$. This is indicative of delayed thermalization.

This calculation of the SFF takes into account only the leading order diagrams, so it need not hold at $O(N)$ times where the lower-order diagrams may collectively make a nonvanishing contribution. We could naively extend the result to later times by saying Eq. C.44 holds for $t < N$ and the SFF is simply at its plateau value $N$ at all later times. Comparison with the numerical results shown in Fig. 3 of the main text does show deviation from this at late times when the thermalization time is of the same order as the plateau time. However, when this is not the case, there is good agreement between this naively extended result and numerics at all times. The oscillations in the numerical results at early times are a result of the finite system size.

# Bibliography


[1] R. Barney, M. Winer, C. L. Baldwin, B. Swingle, and V. Galitski, "Spectral statistics of a minimal quantum glass model", SciPost Phys. **15**, 084 (2023).

[2] M. Winer, R. Barney, C. L. Baldwin, V. Galitski, and B. Swingle, "Spectral form factor of a quantum spin glass", J. High Energy Phys. **2022**, 32 (2022).

[3] R. D. Barney, Y. Liao, and V. Galitski, *Slow relaxation in a glassy quantum circuit*, 2024, arXiv:2410.23281 [cond-mat.dis-nn].

[4] R. Barney, M. Winer, and V. Galitski, *Neural networks as spin models: from glass to hidden order through training*, 2024, arXiv:2408.06421 [cond-mat.dis-nn].

[5] R. Barney, D. Lakhdar-Hamina, and V. Galitski, *Natural quantization of neural networks*, 2025, arXiv:2503.15482 [quant-ph].

[6] D. Lakhdar-Hamina, X. Liu, R. Barney, S. H. Miller, A. M. Green, N. M. Linke, and V. Galitski, *Benchmarking a tunable quantum neural network on trapped-ion and superconducting hardware*, 2025, arXiv:2507.21222 [quant-ph].

[7] S. F. Edwards and P. W. Anderson, "Theory of spin glasses", J. Phys. F **5**, 965 (1975).

[8] B. Derrida, "Random-energy model: limit of a family of disordered models", Phys. Rev. Lett. **45**, 79–82 (1980).

[9] D. Gross and M. Mezard, "The simplest spin glass", Nuclear Physics B **240**, 431–452 (1984).

[10] A. Crisanti and H. J. Sommers, "The spherical p-spin interaction spin glass model: the statics", Z. Phys. B **87**, 341–354 (1992).

[11] A. Crisanti, H. Horner, and H. J. Sommers, "The spherical p-spin interaction spin-glass model", Z. Phys. B **92**, 257–271 (1993).

[12] D. Sherrington and S. Kirkpatrick, "Solvable model of a spin-glass", Phys. Rev. Lett. **35**, 1792–1796 (1975).

[13] G. Parisi, "Toward a mean field theory for spin glasses", Physics Letters A **73**, 203–205 (1979).

[14] B. Derrida, "Random-energy model: an exactly solvable model of disordered systems", Phys. Rev. B Condens. Matter **24**, 2613–2626 (1981).

[15] M. Mezard, G. Parisi, and M. A. Virasoro, *Spin glass theory and beyond* (World Scientific, Singapore, 1987).

[16] K. H. Fischer and J. A. Hertz, *Spin glasses*, Cambridge Studies in Magnetism (Cambridge University Press, 1991).

[17] H. Nishimori, *Statistical physics of spin glasses and information processing* (Oxford University Press, 2001).

[18] D. L. Stein and C. M. Newman, *Spin glasses and complexity* (Princeton University Press, Princeton, U.S.A., 2013).

[19] F. Guerra, "Broken replica symmetry bounds in the mean field spin glass model", Commun. Math. Phys. **233**, 1–12 (2003).

[20] M. Talagrand, *Spin glasses: a challenge for mathematicians*, en, 2003rd ed., Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics (Springer, Berlin, Germany, July 2003).

[21] J. Kurchan, G. Parisi, and M. A. Virasoro, "Barriers and metastable states as saddle points in the replica approach", Journal de Physique I **3**, 1819–1838 (1993).

[22] T. R. Kirkpatrick and P. G. Wolynes, "Stable and metastable states in mean-field potts and structural glasses", en, Phys. Rev. B Condens. Matter **36**, 8552–8564 (1987).

[23] D. J. Thouless, P. W. Anderson, and R. G. Palmer, "Solution of 'solvable model of a spin glass'", Philosophical Magazine **35**, 593–601 (1977).

[24] M. Suzuki, "Relationship between d-dimensional quantal spin systems and (d+1)-dimensional ising systems: equivalence, critical exponents and systematic approximants of the partition function and spin correlations", Prog. Theor. Phys. **56**, 1454–1469 (1976).

[25] G. Biroli and L. F. Cugliandolo, "Quantum Thouless-Anderson-Palmer equations for glassy systems", Phys. Rev. B **64**, 014206 (2001).

[26] J. Barrow-Green, *Poincaré and the three body problem*, en, History of mathematics (American Mathematical Society, Providence, RI, Oct. 1996).

[27] E. N. Lorenz, "Deterministic nonperiodic flow", en, J. Atmos. Sci. **20**, 130–141 (1963).

[28] E. Lorenz, "Predictability: does the flap of a butterfly's wings in Brazil set off a tornado in Texas?", Address at the 139th Meeting of the American Association for the Advancement of Science, Dec 29, 1972.

[29] J. M. Deutsch, "Quantum statistical mechanics in a closed system", Phys. Rev. A **43**, 2046–2049 (1991).

[30] M. Srednicki, "Chaos and quantum thermalization", Phys. Rev. E **50**, 888–901 (1994).

[31] J. M. Deutsch, "Eigenstate thermalization hypothesis", Rep. Prog. Phys. **81**, 082001 (2018).

[32] O. Bohigas, M. J. Giannoni, and C. Schmit, "Characterization of chaotic quantum spectra and universality of level fluctuation laws", Phys. Rev. Lett. **52**, 1–4 (1984).

[33] E. P. Wigner, "On the statistical distribution of the widths and spacings of nuclear resonance levels", Math. Proc. Cambridge Philos. Soc. **47**, 790–798 (1951).

[34] E. P. Wigner, "Random matrices in physics", SIAM review **9**, 1–23 (1967).

[35] E. Wigner, "Statistical properties of real symmetric matrices with many dimensions", Canadian Mathematical Congress Proceedings, 174 (1957).

[36] F. J. Dyson, "The threefold way. algebraic structure of symmetry groups and ensembles in quantum mechanics", J. Math. Phys. **3**, 1199–1215 (1962).

[37] E. P. Wigner, "Characteristic vectors of bordered matrices with infinite dimensions", Ann. Math. **62**, 548–564 (1955).

[38] P. W. Anderson, "More is different", en, Science **177**, 393–396 (1972).

[39] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities", Proc. Natl. Acad. Sci. U. S. A. **79**, 2554–2558 (1982).

[40] D. H. Ackley, G. E. Hinton, and T. J. Sejnowski, "A learning algorithm for boltzmann machines", en, Cogn. Sci. **9**, 147–169 (1985).

[41] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A survey of convolutional neural networks: analysis, applications, and prospects", en, IEEE Trans. Neural Netw. Learn. Syst. **33**, 6999–7019 (2022).

[42] T. Lin, Y. Wang, X. Liu, and X. Qiu, "A survey of transformers", en, AI Open **3**, 111–132 (2022).

[43] L. B. Almeida, "Multilayer perceptrons", in *Handbook of neural computation*, edited by E. Fisler and R. Beale (CRC Press, 1996), pp. C1–2.

[44] R. Kruse, S. Mostaghim, C. Borgelt, C. Braune, and M. Steinbrecher, "Multi-layer perceptrons", in *Computational intelligence: a methodological introduction* (Springer International Publishing, Cham, 2022), pp. 53–124.

[45] G. Cybenko, "Approximation by superpositions of a sigmoidal function", Mathematics of Control, Signals and Systems **2**, 303–314 (1989).

[46] S. Linnainmaa, "Taylor expansion of the accumulated rounding error", BIT Numerical Mathematics **16**, 146–160 (1976).

[47] D. E. Rumelhart and J. L. McClelland, "Learning internal representations by error propagation", in *Parallel distributed processing: explorations in the microstructure of cognition: foundations* (MIT Press, 1987), pp. 318–362.

[48] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors", en, Nature **323**, 533–536 (1986).

[49] I. Sutskever, J. Martens, G. Dahl, and G. Hinton, "On the importance of initialization and momentum in deep learning", in Proceedings of the 30th international conference on machine learning, Vol. 28, edited by S. Dasgupta and D. McAllester, Proceedings of Machine Learning Research 3 (17–19 Jun 2013), pp. 1139–1147.

[50] H. Robbins and S. Monro, "A stochastic approximation method", Ann. Math. Stat. **22**, 400–407 (1951).

[51] P. Rebentrost, T. R. Bromley, C. Weedbrook, and S. Lloyd, "Quantum hopfield neural network", Phys. Rev. A **98**, 042308 (2018).

[52] P. Rotondo, M. Marcuzzi, J. P. Garrahan, I. Lesanovsky, and M. Müller, "Open quantum generalisation of hopfield neural networks", Journal of Physics A: Mathematical and Theoretical **51**, 115301 (2018).

[53] R. Xie and A. Kamenev, "Quantum hopfield model with dilute memories", Phys. Rev. A **110**, 032418 (2024).

[54] M. H. Amin, E. Andriyash, J. Rolfe, B. Kulchytskyy, and R. Melko, "Quantum boltzmann machine", Phys. Rev. X **8**, 021050 (2018).

[55] C. Zoufal, A. Lucchi, and S. Woerner, "Variational quantum boltzmann machines", Quantum Machine Intelligence **3**, 7 (2021).

[56] S. Bangar, L. Sunny, K. Yeter-Aydeniz, and G. Siopsis, "Continuous-variable quantum boltzmann machine", Quantum Machine Intelligence **7**, 33 (2025).

[57] I. Cong, S. Choi, and M. D. Lukin, "Quantum convolutional neural networks", Nature Physics **15**, 1273–1278 (2019).

[58] S. Wei, Y. Chen, Z. Zhou, and G. Long, "A quantum convolutional neural network on nisq devices", AAPPS Bulletin **32**, 2 (2022).

[59] S. Oh, J. Choi, and J. Kim, "A tutorial on quantum convolutional neural networks (qcnn)", in 2020 international conference on information and communication technology convergence (ictc) (2020), pp. 236–239.

[60] J. Herrmann, S. M. Llima, A. Remm, P. Zapletal, N. A. McMahon, C. Scarato, F. Swiadek, C. K. Andersen, C. Hellings, S. Krinner, N. Lacroix, S. Lazar, M. Kerschbaum, D. C. Zanuz, G. J. Norris, M. J. Hartmann, A. Wallraff, and C. Eichler, "Realizing quantum convolutional neural networks on a superconducting quantum processor to recognize quantum phases", Nature Communications **13**, 4144 (2022).

[61] L.-H. Gong, J.-J. Pei, T.-F. Zhang, and N.-R. Zhou, "Quantum convolutional neural network based on variational quantum circuits", Optics Communications **550**, 129993 (2024).

[62] M. Benedetti, E. Lloyd, S. Sack, and M. Fiorentini, "Parameterized quantum circuits as machine learning models", Quantum Science and Technology **4**, 043001 (2019).

[63] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, "Variational quantum algorithms", Nature Reviews Physics **3**, 625–644 (2021).

[64] E. Stoudenmire and D. J. Schwab, "Supervised learning with tensor networks", in Advances in neural information processing systems, Vol. 29, edited by D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett (2016).

[65] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, "Quantum circuit learning", en, Phys. Rev. A (Coll. Park.) **98** (2018).

[66] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition", Proceedings of the IEEE **86**, 2278–2324 (1998).

[67] E. Wigner, "Results and theory of resonance absorption", in Proceedings of the conference on neutron physics by time-of-flight, Gatlinburg, Tennessee (Oak Ridge National Laboratory, 1956).

[68] J. A. Harvey and D. J. Hughes, "Spacings of nuclear energy levels", Phys. Rev. **109**, 471–479 (1958).

[69] R. U. Haq, A. Pandey, and O. Bohigas, "Fluctuation properties of nuclear energy levels: do theory and experiment agree?", Phys. Rev. Lett. **48**, 1086–1089 (1982).

[70] M. Mehta, "On the statistical properties of the level-spacings in nuclear spectra", Nuclear Physics **18**, 395–419 (1960).

[71] C. Porter, *Statistical theories of spectra: fluctuations* (Academic Press, New York, 1965).

[72] A. V. Andreev, O. Agam, B. D. Simons, and B. L. Altshuler, "Quantum chaos, irreversible classical dynamics, and random matrix theory", Phys. Rev. Lett. **76**, 3947–3950 (1996).

[73] M. Gutzwiller, *Chaos in classical and quantum mechanics*, Interdisciplinary Applied Mathematics (Springer, New York, 1991).

[74] J. S. Cotler, G. Gur-Ari, M. Hanada, J. Polchinski, P. Saad, S. H. Shenker, D. Stanford, A. Streicher, and M. Tezuka, "Black holes and random matrices", J. High Energ. Phys. **2017**, 118 (2017).

[75] P. Saad, S. H. Shenker, and D. Stanford, *A semiclassical ramp in SYK and in gravity*, 2019, arXiv:1806.06840 [hep-th].

[76] E. Brézin, C. Itzykson, G. Parisi, and J. Zuber, "Planar diagrams", Comm. Math. Phys. **59**, 35–51 (1978).

[77] C. Itzykson and J. B. Zuber, "The planar approximation. II", J. Math. Phys. **21**, 411–421 (1980).

[78] J. Maldacena, S. H. Shenker, and D. Stanford, "A bound on chaos", J. High Energ. Phys. **2016**, 106 (2016).

[79] Y. Sekino and L. Susskind, "Fast scramblers", J. High Energ. Phys. **2008**, 065 (2008).

[80] M. Srdinšek, T. Prosen, and S. Sotiriadis, "Signatures of chaos in nonintegrable models of quantum field theories", Phys. Rev. Lett. **126**, 121602 (2021).

[81] F. Haake, *Quantum signatures of chaos*, Springer Series in Synergetics (Springer Berlin Heidelberg, 2010).

[82] A. Altland and M. R. Zirnbauer, "Nonstandard symmetry classes in mesoscopic normal-superconducting hybrid structures", Phys. Rev. B **55**, 1142–1161 (1997).

[83] N. Rosenzweig and C. E. Porter, ""Repulsion of energy levels" in complex atomic spectra", Phys. Rev. **120**, 1698–1714 (1960).

[84] M. L. Mehta, *Random matrices*, 3rd ed., Pure and Applied Mathematics vol. 142 (Academic Press, Amsterdam, 2004).

[85] A. Pandey, "Brownian-motion model of discrete spectra", Chaos Solit. Fractals **5**, 1275–1285 (1995).

[86] E. Brézin and S. Hikami, "Correlations of nearby levels induced by a random potential", Nucl. Phys. B **479**, 697–706 (1996).

[87] T. Guhr, "Transition from Poisson regularity to chaos in a time-reversal noninvariant system", Phys. Rev. Lett. **76**, 2258–2261 (1996).

[88] T. Guhr and A. Müller-Groeling, "Spectral correlations in the crossover between GUE and Poisson regularity: on the identification of scales", J. Math. Phys. **38**, 1870–1887 (1997).

[89] V. E. Kravtsov, I. M. Khaymovich, E. Cuevas, and M. Amini, "A random matrix model with localization and ergodic transitions", New J. Phys. **17**, 122002 (2015).

[90] D. Facoetti, P. Vivo, and G. Biroli, "From non-ergodic eigenvectors to local resolvent statistics and back: a random matrix perspective", Europhys. Lett. **115**, 47003 (2016).

[91] M. Pino, J. Tabanera, and P. Serna, "From ergodic to non-ergodic chaos in Rosenzweig–Porter model", J. Phys. A **52**, 475101 (2019).

[92] P. von Soosten and S. Warzel, "Non-ergodic delocalization in the rosenzweig–porter model", Lett. Math. Phys. **109**, 905–922 (2019).

[93] K. Truong and A. Ossipov, "Eigenvectors under a generic perturbation: non-perturbative results from the random matrix approach", EPL **116**, 37002 (2016).

[94] E. Bogomolny and M. Sieber, "Power-law random banded matrices and ultrametric matrices: eigenvector distribution in the intermediate regime", Phys. Rev. E **98**, 042116 (2018).

[95] M. Amini, "Spread of wave packets in disordered hierarchical lattices", EPL **117**, 30003 (2017).

[96] G. D. Tomasi, M. Amini, S. Bera, I. M. Khaymovich, and V. E. Kravtsov, "Survival probability in Generalized Rosenzweig-Porter random matrix ensemble", SciPost Phys. **6**, 014 (2019).

[97] R. Berkovits, "Super-Poissonian behavior of the Rosenzweig-Porter model in the nonergodic extended regime", Phys. Rev. B **102**, 165140 (2020).

[98] M. A. Skvortsov, M. Amini, and V. E. Kravtsov, "Sensitivity of (multi)fractal eigenstates to a perturbation of the Hamiltonian", Phys. Rev. B **106**, 054208 (2022).

[99] D. Venturelli, L. F. Cugliandolo, G. Schehr, and M. Tarzia, *Replica approach to the generalized Rosenzweig-Porter model*, 2022, arXiv:2209.11732 [cond-mat.dis-nn].

[100] I. M. Khaymovich, V. E. Kravtsov, B. L. Altshuler, and L. B. Ioffe, "Fragile extended phases in the log-normal Rosenzweig-Porter model", Phys. Rev. Res. **2**, 043346 (2020).

[101] V. E. Kravtsov, I. M. Khaymovich, B. L. Altshuler, and L. B. Ioffe, *Localization transition on the Random Regular Graph as an unstable tricritical point in a log-normal Rosenzweig-Porter random matrix ensemble*, 2020, arXiv:2002.02979 [cond-mat.dis-nn].

[102] I. M. Khaymovich and V. E. Kravtsov, "Dynamical phases in a "multifractal" Rosenzweig-Porter model", SciPost Phys. **11**, 045 (2021).

[103] C. Monthus, "Multifractality of eigenstates in the delocalized non-ergodic phase of some random matrix models: Wigner–Weisskopf approach", J. Phys. A. **50**, 295101 (2017).

[104] G. Biroli and M. Tarzia, "Lévy-Rosenzweig-Porter random matrix ensemble", Phys. Rev. B **103**, 104205 (2021).

[105] W. Buijsman and Y. Bar Lev, "Circular Rosenzweig-Porter random matrix ensemble", SciPost Phys. **12**, 082 (2022).

[106] G. De Tomasi and I. M. Khaymovich, "Non-Hermitian Rosenzweig-Porter random-matrix ensemble: obstruction to the fractal phase", Phys. Rev. B **106**, 094204 (2022).

[107] H. Kunz and B. Shapiro, "Transition from Poisson to Gaussian unitary statistics: the two-point correlation function", Phys. Rev. E **58**, 400–406 (1998).

[108] C. L. Baldwin and C. R. Laumann, "Quantum algorithm for energy matching in hard optimization problems", Phys. Rev. B **97**, 224201 (2018).

[109] L. Faoro, M. V. Feigel'man, and L. Ioffe, "Non-ergodic extended phase of the quantum random energy model", Ann. Phys. **409**, 167916 (2019).

[110] G. Biroli, D. Facoetti, M. Schiró, M. Tarzia, and P. Vivo, "Out-of-equilibrium phase diagram of the quantum random energy model", Phys. Rev. B **103**, 014204 (2021).

[111] E. Wigner and J. Griffin, *Group theory and its application to the quantum mechanics of atomic spectra*, Pure and applied Physics (Academic Press, 1959).

[112] O. Bohigas and M.-J. Giannoni, "Chaotic motion and random matrix theories", in *Mathematical and computational methods in nuclear physics*, edited by J. S. Dehesa, J. M. G. Gomez, and A. Polls, Lecture Notes in Physics vol. 209 (Springer, Berlin, Heidelberg, 1984), pp. 1–99.

[113] R. Dubertrand and S. Müller, "Spectral statistics of chaotic many-body systems", New Journal of Physics **18**, 033009 (2016).

[114] B. Bertini, P. Kos, and T. Prosen, "Exact spectral form factor in a minimal model of many-body quantum chaos", Phys. Rev. Lett. **121**, 264101 (2018).

[115] M. V. Berry and M. Tabor, "Level clustering in the regular spectrum", Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences **356**, 375–394 (1977).

[116] D. Roy and T. Prosen, "Random matrix spectral form factor in kicked interacting fermionic chains", Phys. Rev. E **102**, 060202 (2020).

[117] Y. Liao, A. Vikram, and V. Galitski, "Many-body level statistics of single-particle quantum chaos", Physical Review Letters **125**, 10.1103/physrevlett.125.250601 (2020).

[118] M. Winer, S.-K. Jian, and B. Swingle, "Exponential ramp in the quadratic sachdev-ye-kitaev model", Phys. Rev. Lett. **125**, 250602 (2020).

[119] M. Winer and B. Swingle, *Emergent spectral form factors in sonic systems*, 2022, arXiv:2211.09134 [cond-mat.stat-mech].

[120] L. A. Bunimovich, "On ergodic properties of certain billiards", Funct. Anal. Its Appl. **8**, 254–255 (1974).

[121] F. Alet and N. Laflorencie, "Many-body localization: an introduction and selected topics", C. R. Phys. **19**, 498–525 (2018).

[122] D. A. Abanin and Z. Papić, "Recent progress in many-body localization", Ann. Phys. **529**, 1700169 (2017).

[123] P. Anderson, "Absence of diffusion in certain random lattices", Phys. Rev. **109**, 1492–1505 (1958).

[124] P. A. Nosov, I. M. Khaymovich, and V. E. Kravtsov, "Correlation-induced localization", Phys. Rev. B **99**, 104203 (2019).

[125] N. F. Mott, "The electrical properties of liquid mercury", Philos. Mag. **13**, 989–1014 (1966).

[126] T. Guhr, "Dyson's correlation functions and graded symmetry", J. Math. Phys. **32**, 336 (1991).

[127] A. Chan, A. D. Luca, and J. Chalker, "Spectral statistics in spatially extended chaotic quantum many-body systems", Phys. Rev. Lett. **121**, 060601 (2018).

[128] A. Chan, A. De Luca, and J. T. Chalker, "Spectral Lyapunov exponents in chaotic and localized many-body quantum systems", Phys. Rev. Res. **3**, 023118 (2021).

[129] S. J. Garratt and J. T. Chalker, "Local pairing of Feynman histories in many-body Floquet models", Phys. Rev. X **11**, 021051 (2021).

[130] D. K. Nandy, T. Čadež, B. Dietz, A. Andreanov, and D. Rosa, "Delayed thermalization in the mass-deformed Sachdev-Ye-Kitaev model", Phys. Rev. B **106**, 245147 (2022).

[131] S. M. Ruiz, "An algebraic identity leading to Wilson's theorem", Math. Gaz. **80**, 579–582 (1996).

[132] M. Rigol, V. Dunjko, and M. Olshanii, "Thermalization and its mechanism for generic isolated quantum systems", Nature **452**, 854–858 (2008).

[133] P. Glorioso and H. Liu, *Lectures on non-equilibrium effective field theories and fluctuating hydrodynamics*, 2018, arXiv:1805.09331 [hep-th].

[134] M. Crossley, P. Glorioso, and H. Liu, *Effective field theory of dissipative fluids*, 2015, arXiv:1511.03646 [hep-th].

[135] S. Grozdanov and J. Polonyi, "Viscosity and dissipative hydrodynamics from effective field theory", Phys. Rev. D **91**, 105031 (2015).

[136] F. M. Haehl, R. Loganayagam, and M. Rangamani, "Effective action for relativistic hydrodynamics: fluctuations, dissipation, and entropy inflow", J. High Energ. Phys. **2018**, 194 (2018).

[137] K. Jensen, N. Pinzani-Fokeeva, and A. Yarom, "Dissipative hydrodynamics in superspace", J. High Energ. Phys. **2018**, 127 (2018).

[138] T. Guhr, A. Müller–Groeling, and H. A. Weidenmüller, "Random-matrix theories in quantum physics: common concepts", Phys. Rep. **299**, 189–425 (1998).

[139] L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, "From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics", Adv. Phys. **65**, 239–362 (2016).

[140] L. F. Santos and M. Rigol, "Onset of quantum chaos in one-dimensional bosonic and fermionic systems and its relation to thermalization", Phys. Rev. E **81**, 036206 (2010).

[141] S. Moudgalya, A. Prem, D. A. Huse, and A. Chan, "Spectral statistics in constrained many-body quantum chaotic systems", Phys. Rev. Research **3** (2021).

[142] M. Schiulaz, E. J. Torres-Herrera, and L. F. Santos, "Thouless and relaxation time scales in many-body quantum systems", Phys. Rev. B **99**, 174313 (2019).

[143] M. Winer and B. Swingle, *Spontaneous symmetry breaking, spectral statistics, and the ramp*, 2021, arXiv:2106.07674 [cond-mat.stat-mech].

[144] D. Roy, D. Mishra, and T. Prosen, *Spectral form factor in a minimal bosonic model of many-body quantum chaos*, 2022, arXiv:2203.05439 [cond-mat.stat-mech].

[145] K. Binder and A. P. Young, "Spin glasses: experimental facts, theoretical concepts, and open questions", Rev. Mod. Phys. **58**, 801–976 (1986).

[146] T. Castellani and A. Cavagna, "Spin-glass theory for pedestrians", J. Stat. Mech. **2005**, P05012 (2005).

[147] M. Mezard and A. Montanari, *Information, physics, and computation* (Oxford University Press, 2009).

[148] P. Saad, *Late time correlation functions, baby universes, and ETH in JT gravity*, 2019, arXiv:1910.10311 [hep-th].

[149] S. Müller, S. Heusler, P. Braun, F. Haake, and A. Altland, "Periodic-orbit theory of universality in quantum chaos", Phys. Rev. E **72**, 046207 (2005).

[150] Y. Chen, *Spectral form factor for free large N gauge theory and strings*, 2022, arXiv:2202.04741 [hep-th].

[151] E. Brézin and S. Hikami, "Spectral form factor in a random matrix theory", Phys. Rev. E **55**, 4067 (1997).

[152] K. Papadodimas and S. Raju, "Local operators in the eternal black hole", Phys. Rev. Lett. **115**, 10.1103/physrevlett.115.211601 (2015).

[153] T. Tao, *Topics in random matrix theory*, Graduate studies in mathematics (American Mathematical Society, 2012).

[154] T. Anous and F. M. Haehl, "The quantum p-spin glass model: a user manual for holographers", J. Stat. Mech. **2111**, 113101 (2021), arXiv:2106.03838 [hep-th].

[155] T. R. Kirkpatrick and D. Thirumalai, "Dynamics of the structural glass transition and the p-spin interaction spin-glass model", Phys. Rev. Lett. **58**, 2091–2094 (1987).

[156] L. F. Cugliandolo and J. Kurchan, "Analytical solution of the off-equilibrium dynamics of a long-range spin-glass model", Phys. Rev. Lett. **71**, 173–176 (1993).

[157] A. Barrat, R. Burioni, and M. Mézard, "Dynamics within metastable states in a mean-field spin glass", J. Phys. A: Math. Gen. **29**, L81 (1996).

[158] B. Altshuler, H. Krovi, and J. Roland, "Anderson localization makes adiabatic quantum optimization fail", Proceedings of the National Academy of Sciences **107**, 12446–12450 (2010).

[159] V. Bapst, L. Foini, F. Krzakala, G. Semerjian, and F. Zamponi, "The quantum adiabatic algorithm applied to random optimization problems: the quantum spin glass perspective", Phys. Rep. **523**, 127–205 (2013).

[160] B. Zhao, M. C. Kerridge, and D. A. Huse, "Three species of Schrödinger cat states in an infinite-range spin model", Phys. Rev. E **90**, 022104 (2014).

[161] V. N. Smelyanskiy, K. Kechedzhi, S. Boixo, S. V. Isakov, H. Neven, and B. Altshuler, "Nonergodic delocalized states for efficient population transfer within a narrow band of the energy landscape", Phys. Rev. X **10**, 011017 (2020).

[162] H. Ishii and T. Yamamoto, "Effect of a transverse field on the spin glass freezing in the Sherrington-Kirkpatrick model", J. Phys. C: Solid State Phys. **18**, 6225–6237 (1985).

[163] D. Thirumalai, Q. Li, and T. R. Kirkpatrick, "Infinite-range Ising spin glass in a transverse field", J. Phys. A: Math. Gen. **22**, 3339 (1989).

[164] Y. Y. Goldschmidt, "Solvable model of the quantum spin glass in a transverse field", Phys. Rev. B **41**, 4858–4861 (1990).

[165] G. Büttner and K. D. Usadel, "Replica-symmetry breaking for the Ising spin glass in a transverse field", Phys. Rev. B **42**, 6385–6395 (1990).

[166] L. F. Cugliandolo and G. Lozano, "Real-time nonequilibrium dynamics of quantum glassy systems", Phys. Rev. B **59**, 915–942 (1999).

[167] L. F. Cugliandolo, D. R. Grempel, and C. A. da Silva Santos, "Imaginary-time replica formalism study of a quantum spherical p-spin-glass model", Phys. Rev. B **64**, 144031–1440326 (2001).

[168] C. R. Laumann, A. Pal, and A. Scardicchio, "Many-body mobility edge in a mean-field quantum spin glass", Phys. Rev. Lett. **113**, 200405 (2014).

[169] C. L. Baldwin, C. R. Laumann, A. Pal, and A. Scardicchio, "Clustering of nonergodic eigenstates in quantum spin glasses", Phys. Rev. Lett. **118**, 127201 (2017).

[170] L. F. Cugliandolo and G. Lozano, "Quantum aging in mean-field models", Phys. Rev. Lett. **80**, 4979–4982 (1998).

[171] E. Gardner, "Spin glasses with p-spin interactions", Nuclear Physics B **257**, 747–765 (1985).

[172] T. M. Nieuwenhuizen and F. Ritort, "Quantum phase transition in spin glasses with multi-spin interactions", Physica A **250**, 8–45 (1998).

[173] V. Dobrosavljevic and D. Thirumalai, "1/p expansion for a p-spin interaction spin-glass model in a transverse field", J. Phys. A: Math. Gen. **23**, L767–L774 (1990).

[174] L. De Cesare, K. Lukierska-Walasek, I. Rabuffo, and K. Walasek, "Replica-symmetry breaking and quantum fluctuation effects in the p-spin interaction spin-glass model with a transverse field", J. Phys. A: Math. Gen. **29**, 1605–1618 (1996).

[175] S. Thomson, P. Urbani, and M. Schiró, "Quantum quenches in isolated quantum glasses out of equilibrium", Phys. Rev. Lett. **125** (2020).

[176] A. Crisanti and H. Sommers, "Thouless-Anderson-Palmer approach to the spherical p-spin spin glass model", J. Phys., I **5**, 805–813 (1995).

[177] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics* (Addison-Wesley Publishing Company, 1994).

[178] M. Winer and B. Swingle, "Hydrodynamic theory of the connected spectral form factor", Phys. Rev. X **12**, 021009 (2022).

[179] G. Biroli and J. Kurchan, "Metastable states in glassy systems", Phys. Rev. E **64**, 016101 (2001).

[180] D. Facoetti, G. Biroli, J. Kurchan, and D. R. Reichman, "Classical glasses, black holes, and strange quantum liquids", Phys. Rev. B **100**, 205108 (2019).

[181] L. F. Cugliandolo, G. S. Lozano, N. Nessi, M. Picco, and A. Tartaglia, "Quenched dynamics of classical isolated systems: the spherical spin model with two-body random interactions or the Neumann integrable model", J. Stat. Mech. Theory Exp. **2018**, 063206 (2018).

[182] M. Rigol, V. Dunjko, V. Yurovsky, and M. Olshanii, "Relaxation in a completely integrable many-body quantum system: an ab initio study of the dynamics of the highly excited states of 1d lattice hard-core bosons", Phys. Rev. Lett. **98**, 050405 (2007).

[183] M. Rigol, "Breakdown of thermalization in finite one-dimensional systems", Phys. Rev. Lett. **103**, 100403 (2009).

[184] A. Polkovnikov, K. Sengupta, A. Silva, and M. Vengalattore, "Colloquium: nonequilibrium dynamics of closed interacting quantum systems", Rev. Mod. Phys. **83**, 863–883 (2011).

[185] J. Eisert, M. Friesdorf, and C. Gogolin, "Quantum many-body systems out of equilibrium", Nat. Phys. **11**, 124–130 (2015).

[186] C. Gogolin and J. Eisert, "Equilibration, thermalisation, and the emergence of statistical mechanics in closed quantum systems", Rep. Prog. Phys. **79**, 056001 (2016).

[187] B. L. Altshuler, Y. Gefen, A. Kamenev, and L. S. Levitov, "Quasiparticle lifetime in a finite system: a nonperturbative approach", Phys. Rev. Lett. **78**, 2803–2806 (1997).

[188] D. Basko, I. Aleiner, and B. Altshuler, "Metal–insulator transition in a weakly interacting many-electron system with localized single-particle states", Ann. Phys. **321**, 1126–1205 (2006).

[189] J. Z. Imbrie, "On many-body localization for quantum spin chains", J. Stat. Phys. **163**, 998–1048 (2016).

[190] D. A. Huse, R. Nandkishore, V. Oganesyan, A. Pal, and S. L. Sondhi, "Localization-protected quantum order", Phys. Rev. B **88**, 014206 (2013).

[191] R. Nandkishore and D. A. Huse, "Many-body localization and thermalization in quantum statistical mechanics", Annu. Rev. Condens. Matter Phys. **6**, 15–38 (2015).

[192] S. Balasubramanian, Y. Liao, and V. Galitski, "Many-body localization landscape", Phys. Rev. B **101**, 014201 (2020).

[193] V. Oganesyan and D. A. Huse, "Localization of interacting fermions at high temperature", Phys. Rev. B **75**, 155111 (2007).

[194] A. Pal and D. A. Huse, "Many-body localization phase transition", Phys. Rev. B **82**, 174411 (2010).

[195] S. Iyer, V. Oganesyan, G. Refael, and D. A. Huse, "Many-body localization in a quasiperiodic system", Phys. Rev. B **87**, 134202 (2013).

[196] D. J. Luitz, N. Laflorencie, and F. Alet, "Many-body localization edge in the random-field heisenberg chain", Phys. Rev. B **91**, 081103 (2015).

[197] R. Mondaini and M. Rigol, "Many-body localization and thermalization in disordered hubbard chains", Phys. Rev. A **92**, 041601 (2015).

[198] P. Sierant and J. Zakrzewski, "Many-body localization of bosons in optical lattices", New J. Phys. **20**, 043032 (2018).

[199] T. Orell, A. A. Michailidis, M. Serbyn, and M. Silveri, "Probing the many-body localization phase transition with superconducting circuits", Phys. Rev. B **100**, 134504 (2019).

[200] N. Macé, N. Laflorencie, and F. Alet, "Many-body localization in a quasiperiodic Fibonacci chain", SciPost Phys. **6**, 050 (2019).

[201] M. Hopjan and F. Heidrich-Meisner, "Many-body localization from a one-particle perspective in the disordered one-dimensional bose-hubbard model", Phys. Rev. A **101**, 063617 (2020).

[202] N. Y. Yao, C. R. Laumann, S. Gopalakrishnan, M. Knap, M. Müller, E. A. Demler, and M. D. Lukin, "Many-body localization in dipolar systems", Phys. Rev. Lett. **113**, 243002 (2014).

[203] M. Schreiber, S. S. Hodgman, P. Bordia, H. P. Lüschen, M. H. Fischer, R. Vosk, E. Altman, U. Schneider, and I. Bloch, "Observation of many-body localization of interacting fermions in a quasirandom optical lattice", Science **349**, 842–845 (2015).

[204] J. Choi, S. Hild, J. Zeiher, P. Schauv, A. Rubio-Abadal, T. Yefsah, V. Khemani, D. A. Huse, I. Bloch, and C. Gross, "Exploring the many-body localization transition in two dimensions", Science **352**, 1547–1552 (2016).

[205] J. Smith, A. Lee, P. Richerme, B. Neyenhuis, P. W. Hess, P. Hauke, M. Heyl, D. A. Huse, and C. Monroe, "Many-body localization in a quantum simulator with programmable random disorder", Nat. Phys. **12**, 907–911 (2016).

[206] 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**, 217–220 (2017).

[207] P. Bordia, H. Lüschen, S. Scherg, S. Gopalakrishnan, M. Knap, U. Schneider, and I. Bloch, "Probing slow relaxation and many-body localization in two-dimensional quasiperiodic systems", Phys. Rev. X **7**, 041047 (2017).

[208] P. Bordia, H. Lüschen, U. Schneider, M. Knap, and I. Bloch, "Periodically driving a many-body localized quantum system", Nat. Phys. **13**, 460–464 (2017).

[209] H. P. Lüschen, P. Bordia, S. S. Hodgman, M. Schreiber, S. Sarkar, A. J. Daley, M. H. Fischer, E. Altman, I. Bloch, and U. Schneider, "Signatures of many-body localization in a controlled open quantum system", Phys. Rev. X **7**, 011034 (2017).

[210] H. P. Lüschen, P. Bordia, S. Scherg, F. Alet, E. Altman, U. Schneider, and I. Bloch, "Observation of slow dynamics near the many-body localization transition in one-dimensional quasiperiodic systems", Phys. Rev. Lett. **119**, 260401 (2017).

[211] K. X. Wei, C. Ramanathan, and P. Cappellaro, "Exploring localization in nuclear spin chains", Phys. Rev. Lett. **120**, 070501 (2018).

[212] W. De Roeck, F. Huveneers, M. Müller, and M. Schiulaz, "Absence of many-body mobility edges", Phys. Rev. B **93**, 014203 (2016).

[213] R. K. Panda, A. Scardicchio, M. Schulz, S. R. Taylor, and M. Žnidarič, "Can we study the many-body localisation transition?", EPL **128**, 67003 (2020).

[214] J. Šuntajs, J. Bonča, T. Prosen, and L. Vidmar, "Quantum chaos challenges many-body localization", Phys. Rev. E **102**, 062144 (2020).

[215] P. Sierant, D. Delande, and J. Zakrzewski, "Thouless time analysis of anderson and many-body localization transitions", Phys. Rev. Lett. **124**, 186601 (2020).

[216] D. Abanin, J. Bardarson, G. De Tomasi, S. Gopalakrishnan, V. Khemani, S. Parameswaran, F. Pollmann, A. Potter, M. Serbyn, and R. Vasseur, "Distinguishing localization from chaos: challenges in finite-size systems", Ann. Phys. **427**, 168415 (2021).

[217] P. Sierant and J. Zakrzewski, "Challenges to observation of many-body localization", Phys. Rev. B **105**, 224203 (2022).

[218] W. Wu, B. Ellman, T. F. Rosenbaum, G. Aeppli, and D. H. Reich, "From classical to quantum glass", Phys. Rev. Lett. **67**, 2076–2079 (1991).

[219] H. Rieger and A. P. Young, "Quantum spin glasses", in Complex behaviour of glassy systems, edited by M. Rubí and C. Pérez-Vicente (1997), pp. 256–265.

[220] F. J. Dyson, "Statistical theory of the energy levels of complex systems. I, II, III", J. Math. Phys. **3**, 140–175 (1962).

[221] M. Sieber and K. Richter, "Correlations between periodic orbits and their role in spectral statistics", Phys. Scr. **T90**, 128 (2001).

[222] A. Chan, A. De Luca, and J. T. Chalker, "Solution of a minimal model for many-body quantum chaos", Phys. Rev. X **8**, 041019 (2018).

[223] P. Kos, M. Ljubotina, and T. Prosen, "Many-body quantum chaos: analytic connection to random matrix theory", Phys. Rev. X **8**, 021062 (2018).

[224] Y. Liao and V. Galitski, "Universal dephasing mechanism of many-body quantum chaos", Phys. Rev. Res. **4**, L012037 (2022).

[225] Y. Liao and V. Galitski, "Emergence of many-body quantum chaos via spontaneous breaking of unitarity", Phys. Rev. B **105**, L140202 (2022).

[226] M. V. Berry, "Semiclassical theory of spectral rigidity", Proc. R. Soc. London, Ser. A **400**, 229–251 (1985).

[227] L. K. Joshi, A. Elben, A. Vikram, B. Vermersch, V. Galitski, and P. Zoller, "Probing many-body quantum chaos with quantum simulators", Phys. Rev. X **12**, 011018 (2022).

[228] M. Lewenstein, A. Sanpera, V. Ahufinger, B. Damski, A. Sen(De), and U. Sen, "Ultracold atomic gases in optical lattices: mimicking condensed matter physics and beyond", Adv. Phys. **56**, 243–379 (2007).

[229] C. Gross and I. Bloch, "Quantum simulations with ultracold atoms in optical lattices", Science **357**, 995–1001 (2017).

[230] F. Schäfer, T. Fukuhara, S. Sugawa, Y. Takasu, and Y. Takahashi, "Tools for quantum simulation with ultracold atoms in optical lattices", Nat. Rev. Phys. **2**, 411–425 (2020).

[231] B. P. Lanyon, C. Hempel, D. Nigg, M. Müller, R. Gerritsma, F. Zähringer, P. Schindler, J. T. Barreiro, M. Rambach, G. Kirchmair, M. Hennrich, P. Zoller, R. Blatt, and C. F. Roos, "Universal digital quantum simulation with trapped ions", Science **334**, 57–61 (2011).

[232] R. Blatt and C. F. Roos, "Quantum simulations with trapped ions", Nat. Phys. **8**, 277–284 (2012).

[233] G. Wendin, "Quantum information processing with superconducting circuits: a review", Rep. Prog. Phys. **80**, 106001 (2017).

[234] M. Kjaergaard, M. E. Schwartz, J. Braumüller, P. Krantz, J. I.-J. Wang, S. Gustavsson, and W. D. Oliver, "Superconducting qubits: current state of play", Annu. Rev. Condens. Matter Phys. **11**, 369–395 (2020).

[235] E. Altman, K. R. Brown, G. Carleo, L. D. Carr, E. Demler, C. Chin, B. DeMarco, S. E. Economou, M. A. Eriksson, K.-M. C. Fu, M. Greiner, K. R. Hazzard, R. G. Hulet, A. J. Kollár, B. L. Lev, M. D. Lukin, R. Ma, X. Mi, S. Misra, C. Monroe, K. Murch, Z. Nazario, K.-K. Ni, A. C. Potter, P. Roushan, M. Saffman, M. Schleier-Smith, I. Siddiqi, R. Simmonds, M. Singh, I. Spielman, K. Temme, D. S. Weiss, J. Vu čkovi ć, V. Vuleti ć, J. Ye, and M. Zwierlein, "Quantum simulators: architectures and opportunities", PRX Quantum **2**, 017003 (2021).

[236] B. Bertini, P. Kos, and T. Prosen, "Exact correlation functions for dual-unitary lattice models in $1 + 1$ dimensions", Phys. Rev. Lett. **123**, 210601 (2019).

[237] B. Bertini, P. Kos, and T. Prosen, "Entanglement spreading in a minimal model of maximal many-body quantum chaos", Phys. Rev. X **9**, 021033 (2019).

[238] S. Gopalakrishnan and A. Lamacraft, "Unitary circuits of finite depth and infinite width from quantum channels", Phys. Rev. B **100**, 064309 (2019).

[239] R. Oliveira, O. C. O. Dahlsten, and M. B. Plenio, "Generic entanglement can be generated efficiently", Phys. Rev. Lett. **98**, 130502 (2007).

[240] M. Žnidarič, "Exact convergence times for generation of random bipartite entanglement", Phys. Rev. A **78**, 032324 (2008).

[241] A. Hamma, S. Santra, and P. Zanardi, "Quantum entanglement in random physical states", Phys. Rev. Lett. **109**, 040502 (2012).

[242] A. Nahum, J. Ruhman, S. Vijay, and J. Haah, "Quantum entanglement growth under random unitary dynamics", Phys. Rev. X **7**, 031016 (2017).

[243] A. Nahum, S. Vijay, and J. Haah, "Operator spreading in random unitary circuits", Phys. Rev. X **8**, 021014 (2018).

[244] C. W. von Keyserlingk, T. Rakovszky, F. Pollmann, and S. L. Sondhi, "Operator hydrodynamics, otocs, and entanglement growth in systems without conservation laws", Phys. Rev. X **8**, 021013 (2018).

[245] T. Zhou and A. Nahum, "Emergent statistical mechanics of entanglement in random unitary circuits", Phys. Rev. B **99**, 174205 (2019).

[246] M. P. Fisher, V. Khemani, A. Nahum, and S. Vijay, "Random quantum circuits", Annu. Rev. Condens. Matter Phys. **14**, 335–379 (2023).

[247] K. Efetov, *Supersymmetry in disorder and chaos* (Cambridge University Press, Cambridge, England, Nov. 2010).

[248] A. Altland, S. Gnutzmann, F. Haake, and T. Micklitz, "A review of sigma models for quantum chaotic dynamics", Rep. Prog. Phys. **78**, 086001 (2015).

[249] Y. Liao and V. Galitski, "Effective field theory of random quantum circuits", Entropy **24** (2022).

[250] Y. Liao and V. Galitski, "Field theory approach to eigenstate thermalization in random quantum circuits", arXiv:2210.06480 (2022).

[251] L. Schäfer and F. Wegner, "Disordered system with n orbitals per site: lagrange formulation, hyperbolic symmetry, and goldstone modes", Z. Phys. B **38**, 113–126 (1980).

[252] F. Wegner, "Disordered electronic system as a model of interacting matrices", Phys. Rep. **67**, 15–24 (1980).

[253] P. Brouwer and C. Beenakker, "Diagrammatic method of integration over the unitary group, with applications to quantum transport in mesoscopic systems", J. Math. Phys. **37**, 4904–4934 (1996).

[254] P. Reimann, "Typical fast thermalization processes in closed many-body systems", Nat. Commun. **7**, 10821 (2016).

[255] V. E. Kravtsov and A. D. Mirlin, "Level statistics in a metallic sample: corrections to the wigner-dyson distribution", JETP-Lett. 60, 645 (1994).

[256] D. J. Amit, H. Gutfreund, and H. Sompolinsky, "Spin-glass models of neural networks", Phys. Rev. A **32**, 1007–1018 (1985).

[257] V. Dotsenko, *An introduction to the theory of spin glasses and neural networks*, Vol. 55, World Scientific lecture notes in physics (World Scientific, Singapore, 1994).

[258] K. Nakanishi and H. Takayama, "Mean-field theory for a spin-glass model of neural networks: tap free energy and the paramagnetic to spin-glass transition", Journal of Physics A **30**, 8085–8094 (1997).

[259] A. Barra and F. Guerra, "About the ergodic regime in the analogical Hopfield neural networks: Moments of the partition function", Journal of Mathematical Physics **49**, 125217 (2008).

[260] A. Barra, G. Genovese, and F. Guerra, "The replica symmetric approximation of the analogical neural network", Journal of Statistical Physics **140**, 784–796 (2010).

[261] A. Barra, G. Genovese, F. Guerra, and D. Tantari, "How glassy are neural networks?", Journal of Statistical Mechanics: Theory and Experiment **2012**, P07009 (2012).

[262] A. Choromanska, M. Henaff, M. Mathieu, G. B. Arous, and Y. LeCun, "The loss surfaces of multilayer networks", in Artificial intelligence and statistics (PMLR, 2015), pp. 192–204.

[263] M. Baity-Jesi, L. Sagun, M. Geiger, S. Spigler, G. B. Arous, C. Cammarota, Y. LeCun, M. Wyart, and G. Biroli, "Comparing dynamics: deep neural networks versus glassy systems", in Proceedings of the 35th international conference on machine learning, Vol. 80, edited by J. Dy and A. Krause, Proceedings of Machine Learning Research (Oct. 2018), pp. 314–323.

[264] N. Zhang, S. Ding, J. Zhang, and Y. Xue, "An overview on restricted Boltzmann machines", Neurocomputing **275**, 1186–1199 (2018).

[265] U.-P. Wen, K.-M. Lan, and H.-S. Shih, "A review of Hopfield neural networks for solving mathematical programming problems", European Journal of Operational Research **198**, 675–687 (2009).

[266] D. Krotov, "A new frontier for Hopfield networks", Nature Reviews Physics **5**, 366–367 (2023).

[267] G. Parisi, "A sequence of approximated solutions to the S-K model for spin glasses", Journal of Physics A: Mathematical and General **13**, L115 (1980).

[268] G. Parisi, "Order parameter for spin-glasses", Phys. Rev. Lett. **50**, 1946–1948 (1983).

[269] T. Plefka, "Convergence condition of the TAP equation for the infinite-ranged Ising spin glass model", J. Phys. A: Math. Gen. **15**, 1971–1978 (1982).

[270] D. Marković, A. Mizrahi, D. Querlioz, and J. Grollier, "Physics for neuromorphic computing", Nature Reviews Physics **2**, 499–510 (2020).

[271] C. D. Schuman, S. R. Kulkarni, M. Parsa, J. P. Mitchell, P. Date, and B. Kay, "Opportunities for neuromorphic computing algorithms and applications", Nature Computational Science **2**, 10–19 (2022).

[272] B. P. Marsh, Y. Guo, R. M. Kroeze, S. Gopalakrishnan, S. Ganguli, J. Keeling, and B. L. Lev, "Enhancing associative memory recall and storage capacity using confocal cavity QED", Phys. Rev. X **11**, 021048 (2021).

[273] R. M. Kroeze, B. P. Marsh, D. A. Schuller, H. S. Hunt, S. Gopalakrishnan, J. Keeling, and B. L. Lev, *Replica symmetry breaking in a quantum-optical vector spin glass*, 2023, arXiv:2311.04216 [quant-ph].

[274] B. P. Marsh, R. M. Kroeze, S. Ganguli, S. Gopalakrishnan, J. Keeling, and B. L. Lev, "Entanglement and replica symmetry breaking in a driven-dissipative quantum spin glass", Phys. Rev. X **14**, 011026 (2024).

[275] M. Thamm, M. Staats, and B. Rosenow, "Random matrix analysis of deep neural network weight matrices", Phys. Rev. E **106**, 054124 (2022).

[276] C. H. Martin and M. W. Mahoney, "Implicit self-regularization in deep neural networks: evidence from random matrix theory and implications for learning", J. Mach. Learn. Res. **22** (2021).

[277] C. H. Martin, T. Peng, and M. W. Mahoney, "Predicting trends in the quality of state-of-the-art neural networks without access to training or testing data", Nature Communications **12** (2021).

[278] Z. Wang, A. Engel, A. D. Sarwate, I. Dumitriu, and T. Chiang, "Spectral evolution and invariance in linear-width neural networks", Advances in Neural Information Processing Systems **36** (2024).

[279] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: surpassing human-level performance on ImageNet classification", in 2015 IEEE International Conference on Computer Vision (ICCV) (2015), pp. 1026–1034.

[280] E. Agliari, L. Albanese, F. Alemanno, and A. Fachechi, *Pattern recognition in deep Boltzmann machines*, 2021, arXiv:2106.08978 [cond-mat.dis-nn].

[281] A. Barra, P. Contucci, E. Mingione, and D. Tantari, "Multi-species mean field spin glasses. rigorous results", Annales Henri Poincaré **16**, 691–708 (2015).

[282] M. Kac, "Mathematical mechanisms of phase transitions", in *Statistical physics, phase transitions, and superfluidity*, Vol. I (Gordon and Breach Science Putlishers, New York, Oct. 1968), pp. 241–305.

[283] L. Onsager, "Electric moments of molecules in liquids", Journal of the American Chemical Society **58**, 1486–1493 (1936).

[284] M. El-Mikkawy, "A note on a three-term recurrence for a tridiagonal matrix", Applied Mathematics and Computation **139**, 503–511 (2003).

[285] M. Courbariaux, Y. Bengio, and J.-P. David, "Binaryconnect: training deep neural networks with binary weights during propagations", in Advances in neural information processing systems, Vol. 28, edited by C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett (2015).

[286] M. Courbariaux, I. Hubara, D. Soudry, R. El-Yaniv, and Y. Bengio, *Binarized neural networks: training deep neural networks with weights and activations constrained to +1 or -1*, 2016, arXiv:1602.02830 [cs.LG].

[287] Y. Bengio, N. Léonard, and A. Courville, *Estimating or propagating gradients through stochastic neurons for conditional computation*, 2013, arXiv:1308.3432 [cs.LG].

[288] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers", in Proceedings of the fifth annual workshop on computational learning theory (1992), pp. 144–152.

[289] Y. Tang, *Deep learning using linear support vector machines*, 2015, arXiv:1306.0239 [cs.LG].

[290] A. J. Bray and M. A. Moore, "Metastable states in spin glasses", Journal of Physics C: Solid State Physics **13**, L469 (1980).

[291] V. A. Marčenko and L. A. Pastur, "Distribution of eigenvalues for some sets of random matrices", Mathematics of the USSR-Sbornik **1**, 457 (1967).

[292] L. Chizat, E. Oyallon, and F. Bach, "On lazy training in differentiable programming", in *Proceedings of the 33rd international conference on neural information processing systems* (Curran Associates Inc., Red Hook, NY, USA, 2019).

[293] S. S. Du, X. Zhai, B. Poczos, and A. Singh, "Gradient descent provably optimizes over-parameterized neural networks", in International conference on learning representations (2019).

[294] Y. Li and Y. Liang, "Learning overparameterized neural networks via stochastic gradient descent on structured data", in Proceedings of the 32nd international conference on neural information processing systems (2018), pp. 8168–8177.

[295] S. Du, J. Lee, H. Li, L. Wang, and X. Zhai, "Gradient descent finds global minima of deep neural networks", in International conference on machine learning (PMLR, 2019), pp. 1675–1685.

[296] Z. Allen-Zhu, Y. Li, and Z. Song, "A convergence theory for deep learning via over-parameterization", in Proceedings of the 36th international conference on machine learning, Vol. 97, edited by K. Chaudhuri and R. Salakhutdinov, Proceedings of Machine Learning Research (Sept. 2019), pp. 242–252.

[297] D. Zou, Y. Cao, D. Zhou, and Q. Gu, "Gradient descent optimizes over-parameterized deep relu networks", Machine Learning **109**, 467–492 (2019).

[298] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, "Quantum machine learning", Nature **549**, 195–202 (2017).

[299] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks", in Advances in neural information processing systems, Vol. 25, edited by F. Pereira, C. Burges, L. Bottou, and K. Weinberger (2012).

[300] Q. Li, W. Cai, X. Wang, Y. Zhou, D. D. Feng, and M. Chen, "Medical image classification with convolutional neural network", in 2014 13th international conference on control automation robotics & vision (icarcv) (2014), pp. 844–848.

[301] J. Hirschberg and C. D. Manning, "Advances in natural language processing", Science **349**, 261–266 (2015).

[302] D. W. Otter, J. R. Medina, and J. K. Kalita, "A survey of the usages of deep learning for natural language processing", IEEE Transactions on Neural Networks and Learning Systems **32**, 604–624 (2021).

[303] I. Lauriola, A. Lavelli, and F. Aiolli, "An introduction to deep learning in natural language processing: models, techniques, and tools", Neurocomputing **470**, 443–456 (2022).

[304] A. W. Senior, R. Evans, J. Jumper, J. Kirkpatrick, L. Sifre, T. Green, C. Qin, A. Žídek, A. W. R. Nelson, A. Bridgland, H. Penedones, S. Petersen, K. Simonyan, S. Crossan, P. Kohli, D. T. Jones, D. Silver, K. Kavukcuoglu, and D. Hassabis, "Improved protein structure prediction using potentials from deep learning", Nature **577**, 706–710 (2020).

[305] K. Choudhary, B. DeCost, C. Chen, A. Jain, F. Tavazza, R. Cohn, C. W. Park, A. Choudhary, A. Agrawal, S. J. L. Billinge, E. Holm, S. P. Ong, and C. Wolverton, "Recent advances and applications of deep learning methods in materials science", npj Computational Materials **8**, 59 (2022).

[306] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, and A. Walsh, "Machine learning for molecular and materials science", Nature **559**, 547–555 (2018).

[307] S. S. Gill, A. Kumar, H. Singh, M. Singh, K. Kaur, M. Usman, and R. Buyya, "Quantum computing: a taxonomy, systematic review and future directions", Software: Practice and Experience **52**, 66–114 (2022).

[308] P. Shor, "Algorithms for quantum computation: discrete logarithms and factoring", in Proceedings 35th annual symposium on foundations of computer science (1994), pp. 124–134.

[309] L. K. Grover, "A fast quantum mechanical algorithm for database search", in Proceedings of the twenty-eighth annual acm symposium on theory of computing, STOC '96 (1996), pp. 212–219.

[310] B. Jaderberg, L. W. Anderson, W. Xie, S. Albanie, M. Kiffner, and D. Jaksch, "Quantum self-supervised learning", Quantum Science and Technology **7**, 035005 (2022).

[311] M. Cerezo, G. Verdon, H.-Y. Huang, L. Cincio, and P. J. Coles, "Challenges and opportunities in quantum machine learning", Nature Computational Science **2**, 567–576 (2022).

[312] M. Schuld, I. Sinayskiy, and F. Petruccione, "An introduction to quantum machine learning", Contemporary Physics **56**, 172–185 (2014).

[313] A. Abbas, D. Sutter, C. Zoufal, A. Lucchi, A. Figalli, and S. Woerner, "The power of quantum neural networks", Nature Computational Science **1**, 403–409 (2021).

[314] K. Beer, D. Bondarenko, T. Farrelly, T. J. Osborne, R. Salzmann, D. Scheiermann, and R. Wolf, "Training deep quantum neural networks", Nature Communications **11**, 808 (2020).

[315] S. K. Jeswal and S. Chakraverty, "Recent developments and applications in quantum neural network: a review", Archives of Computational Methods in Engineering **26**, 793–807 (2019).

[316] Y. Kwak, W. J. Yun, S. Jung, and J. Kim, "Quantum neural networks: concepts, applications, and challenges", in 2021 twelfth international conference on ubiquitous and future networks (icufn) (2021), pp. 413–416.

[317] M. Lewenstein, "Quantum perceptrons", Journal of Modern Optics **41**, 2491–2501 (1994).

[318] A. Kapoor, N. Wiebe, and K. Svore, "Quantum perceptron models", in Advances in neural information processing systems, Vol. 29, edited by D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett (2016).

[319] A. J. da Silva, T. B. Ludermir, and W. R. de Oliveira, "Quantum perceptron over a field and neural network architecture selection in a quantum computer", Neural Networks **76**, 55–64 (2016).

[320] E. Torrontegui and J. J. García-Ripoll, "Unitary quantum perceptron as efficient universal approximator(a)", Europhysics Letters **125**, 30004 (2019).

[321] C. D. Bruzewicz, J. Chiaverini, R. McConnell, and J. M. Sage, "Trapped-ion quantum computing: progress and challenges", Applied Physics Reviews **6**, 021314 (2019).

[322] J. M. Pino, J. M. Dreiling, C. Figgatt, J. P. Gaebler, S. A. Moses, M. S. Allman, C. H. Baldwin, M. Foss-Feig, D. Hayes, K. Mayer, C. Ryan-Anderson, and B. Neyenhuis, "Demonstration of the trapped-ion quantum ccd computer architecture", Nature **592**, 209–213 (2021).

[323] A. Bermudez, X. Xu, R. Nigmatullin, J. O'Gorman, V. Negnevitsky, P. Schindler, T. Monz, U. G. Poschinger, C. Hempel, J. Home, F. Schmidt-Kaler, M. Biercuk, R. Blatt, S. Benjamin, and M. Müller, "Assessing the progress of trapped-ion processors towards fault-tolerant quantum computation", Phys. Rev. X **7**, 041061 (2017).

[324] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis, "Quantum supremacy using a programmable superconducting processor", Nature **574**, 505–510 (2019).

[325] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, "A quantum engineer's guide to superconducting qubits", Applied Physics Reviews **6**, 021318 (2019).

[326] S. Slussarenko and G. J. Pryde, "Photonic quantum information processing: a concise review", Applied Physics Reviews **6**, 041303 (2019).

[327] J. Romero and G. Milburn, *Photonic quantum computing*, 2024, arXiv:2404.03367 [quant-ph].

[328] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn, "Linear optical quantum computing with photonic qubits", Rev. Mod. Phys. **79**, 135–174 (2007).

[329] C. Yuan and S. S. Agaian, "A comprehensive review of binary neural network", Artificial Intelligence Review **56**, 12949–13013 (2023).

[330] X. Ying, "An overview of overfitting and its solutions", Journal of Physics: Conference Series **1168**, 022022 (2019).

[331] T. Raiko, M. Berglund, G. Alain, and L. Dinh, *Techniques for learning binary stochastic feedforward neural networks*, 2015, arXiv:1406.2989 [stat.ML].

[332] J. Sietsma and R. J. Dow, "Creating artificial neural networks that generalize", Neural Networks **4**, 67–79 (1991).

[333] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting", J. Mach. Learn. Res. **15**, 1929–1958 (2014).

[334] B. Skinner, J. Ruhman, and A. Nahum, "Measurement-induced phase transitions in the dynamics of entanglement", Phys. Rev. X **9**, 031009 (2019).

[335] K. L. Pudenz and D. A. Lidar, "Quantum adiabatic machine learning", Quantum Information Processing **12**, 2027–2070 (2013).

[336] S. Lloyd, M. Mohseni, and P. Rebentrost, *Quantum algorithms for supervised and unsupervised machine learning*, 2013, arXiv:1307.0411 [quant-ph].

[337] Z. Li, X. Liu, N. Xu, and J. Du, "Experimental realization of a quantum support vector machine", Phys. Rev. Lett. **114**, 140504 (2015).

[338] C. Ciliberto, M. Herbster, A. D. Ialongo, M. Pontil, A. Rocchetto, S. Severini, and L. Wossnig, "Quantum machine learning: a classical perspective", en, Proc. Math. Phys. Eng. Sci. **474**, 20170551 (2018).

[339] Y. Wang and J. Liu, "A comprehensive review of quantum machine learning: from nisq to fault tolerance", Reports on Progress in Physics **87**, 116402 (2024).

[340] S. Lloyd, M. Schuld, A. Ijaz, J. Izaac, and N. Killoran, *Quantum embeddings for machine learning*, Jan. 2020.

[341] J. Preskill, "Quantum computing in the nisq era and beyond", Quantum **2**, 79 (2018).

[342] H.-L. Huang, Y. Du, M. Gong, Y. Zhao, Y. Wu, C. Wang, S. Li, F. Liang, J. Lin, Y. Xu, R. Yang, T. Liu, M.-H. Hsieh, H. Deng, H. Rong, C.-Z. Peng, C.-Y. Lu, Y.-A. Chen, D. Tao, X. Zhu, and J.-W. Pan, "Experimental quantum generative adversarial networks for image generation", Phys. Rev. Appl. **16**, 024051 (2021).

[343] X. Kong, L. Li, Z. Chen, C. Xue, X. Xu, H. Liu, Y. Wu, Y. Fang, H. Fang, K. Chen, Y. Yang, M. Dou, and G. Guo, *Quantum-enhanced llm efficient fine tuning*, 2025, arXiv:2503.12790 [quant-ph].

[344] E. A. Cherrat, I. Kerenidis, N. Mathur, J. Landman, M. Strahm, and Y. Y. Li, "Quantum Vision Transformers", Quantum **8**, 1265 (2024).

[345] P. Bermejo, P. Braccia, M. S. Rudolph, Z. Holmes, L. Cincio, and M. Cerezo, *Quantum convolutional neural networks are (effectively) classically simulable*, 2024, arXiv:2408.12739 [quant-ph].

[346] E. Gil-Fuster, C. Gyurik, A. Perez-Salinas, and V. Dunjko, "On the relation between trainability and dequantization of variational quantum learning models", in The thirteenth international conference on learning representations (2025).

[347] Y. Li, X. Chen, and M. P. A. Fisher, "Measurement-driven entanglement transition in hybrid quantum circuits", Phys. Rev. B **100**, 134306 (2019).

[348] I. Goodfellow, Y. Bengio, and A. Courville, "Gradient-based learning", in *Deep learning*, http://www.deeplearningbook.org (MIT Press, Cambridge, MA, 2016), pp. 177–191.

[349] National Institute of Standards and Technology, *NIST special database 19: handprinted forms and characters database*, Available at https://www.nist.gov/srd/nist-special-database-19, Accessed: 2025-06-12, 1990.

[350] A. Blais, R.-S. Huang, A. Wallraff, S. M. Girvin, and R. J. Schoelkopf, "Cavity quantum electrodynamics for superconducting electrical circuits: an architecture for quantum computation", Phys. Rev. A **69**, 062320 (2004).

[351] J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, "Charge-insensitive qubit design derived from the cooper pair box", Phys. Rev. A **76**, 042319 (2007).

[352] R. Barney, *Natural quantization of neural networks*, `https://github.com/Galitski-theory-group/quantum-nn`, 2025.

[353] D. Lakhdar-Hamina, *Natural-Quantization: a hybrid classical–quantum neural network framework*, `https://github.com/Galitski-theory-group/natural-quantization`, 2025.

[354] S. Debnath, N. M. Linke, C. Figgatt, K. A. Landsman, K. Wright, and C. Monroe, "Demonstration of a small programmable quantum computer with atomic qubits", en, Nature **536**, 63–66 (2016).

[355] C. M. Bishop, "Training with noise is equivalent to tikhonov regularization", Neural Computation **7**, 108–116 (1995).

[356] T. Tomesh, P. Gokhale, V. Omole, G. S. Ravi, K. N. Smith, J. Viszlai, X.-C. Wu, N. Hardavellas, M. R. Martonosi, and F. T. Chong, "Supermarq: a scalable quantum benchmark suite", in 2022 ieee international symposium on high-performance computer architecture (hpca) (IEEE, 2022), pp. 587–603.

[357] A. Power, Y. Burda, H. Edwards, I. Babuschkin, and V. Misra, *Grokking: generalization beyond overfitting on small algorithmic datasets*, 2022.

[358] S. Samuel, "U(N) Integrals, 1/N, and the De Wit–'t Hooft anomalies", J. Math. Phys. **21**, 2695–2703 (1980).