Entanglement entropy on a fuzzy sphere

Matrix Quantum Mechanics simulations using Deep Learning

by

Tomasz Andrzejewski

MPhys Mathematical Physics, The University of Edinburgh, 2019

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

The Faculty of Graduate Studies

(Physics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2022

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

Entanglement entropy on a fuzzy sphere:
Matrix Quantum Mechanics simulations using Deep Learning

submitted by Tomasz Andrzejewski in partial fulfillment of the requirements for the degree of MASTER OF SCIENCE IN PHYSICS.

Examining Committee:

Joanna L. Karczmarek,
Associate Professor, Physics and Astronomy, UBC
Supervisor

Gordon W. Semenoff,
Professor, Physics and Astronomy, UBC
Supervisory Committee Member

Abstract

The Ryu-Takayanagi formula [1], discovered in the context of the AdS/CFT correspondence, revealed that entanglement entropy encodes the information about geometry. In order to learn about non-locality structure of QFTs on fuzzy spaces we calculate entanglement entropy and mutual information for a massive free scalar field on a noncommutative (fuzzy) sphere, using standard methods for finding entropy of coupled harmonic oscillators. When computing these quantities we use two different methods of factorizing quantum mechanical Hilbert spaces, i.e. two different constructions of projection matrices. We find that our results are largely dependent on which projection matrix we used.

We further use machine learning techniques to find variational wavefunction for scalar field theory with a quartic interaction on a fuzzy sphere. The theory is realized by a matrix model, where the matrix size plays the role of an ultraviolet cutoff. We use variational quantum Monte Carlo with deep generative flows to search for ground state energy of this matrix model. We find that, depending on the projection matrix used, entropy stays the same or behaves differently as we vary the parameter of quartic interaciton.

Lay Summary

In string theory, which seeks to unify quantum and gravitational physics, one can describe spacetime geometry in terms of microscopic degrees of freedom. Rather than a fundamental concept, spacetime is then said to be emergent from the collective behavior of quantum objects, not unlike gas temperature and pressure in thermodynamics that are emergent from the collective dynamics of particles.

We aim to understand how space emerges in matrix quantum mechanics where the underlying microscopic objects are matrices. To obtain the quantum-mechanical wave function that describes our model we use machine learning.

Preface

This thesis is composed of original, unpublished work by the author. The project was proposed by Joanna Karczmarek and aims to examine disparity in results for entanglement entropy on a fuzzy sphere between [2] and [3].

For section 4.1.1, Matlab code which calculates entanglement entropy in the vacuum state for scalar free field theory on a fuzzy sphere and was used to obtain the results in [3] was provided by Hong Zhe Chen. The author built on the code to compute entanglement entropy using different projection matrices to factorize Hilbert space. The author also added functions to find Rényi-n entropy using the method described in [4].

Machine learning code used to obtain results in section 4.2 and 4.3 was adapted from Xizhi Han's code used in [2] [1] The author modified the code that was applied to a gauged matrix model with 3 matrices to an ungauged matrix model with a single matrix. Supplementary Mathematica code, that was not made publicly available, used to compute projection matrices was provided by Xizhi Han on author's request. The thesis's author optimized it so it could be used for much larger sizes of matrices.

In addition to comparing ground state energies in section 4.2, Joanna Karczmarek suggested another benchmark of neural network variatonal wavefunction i.e. comparing variatonal wavefunction to the expected Gaussian wavefunction for a scalar free field theory. Joanna Karczmarek also pointed out how to benchmark scalar interacting field theory using the method developed in [5] and described in 4.3. The author performed all the numerical analysis.

[1] It's publicly available at `https://github.com/hanxzh94/matrix-model`.

Table of Contents

Appendices

List of Figures

Chapter 1

Introduction

The main motivation for this work stems from trying to understand spacetime as an emergent concept. Rather than given a priori as a manifold in general relativity or a quantum field as described by quantum mechanics, spacetime can be understood as arising effectively from the underlying microscopic degrees of freedom. A quantity that gives us most insight into the distribution of degrees of freedom and therefore emergent geometry is entanglement entropy (see e.g. [6–9]).

We talk about entanglement entropy whenever we have an observer with access only to a subset of the complete set of observables associated to a quantum system. It is defined as the von Neumann entropy of the reduced density matrix

$$S(A) = -\operatorname{Tr}_A \rho_A \log \rho_A, \quad \rho_A = \operatorname{tr}_{A^c} \rho \,, \tag{1.1}$$

where ρ is the density matrix with respect to the wavefunction of the whole system and ρ_A is the density matrix reduced to a region A by tracing over the degrees of freedom lying outside that region.

In practice, we often compute Rényi entropies which are given by

$$S_n(A) = \frac{1}{1-n} \log \operatorname{tr} \rho_A^n \,, \tag{1.2}$$

where $n \geq 0$ denotes an order of Rényi entropy. In the limit where $n \to 1$ we get entanglement entropy.

The reason why entanglement entropy is a useful quantity to study is most clear in the context of the AdS/CFT correspondence, through the Ryu-Takayanagi formula [1]:

$$S(A) = \frac{\operatorname{Area}(\gamma_A)}{4G_N} \,, \tag{1.3}$$

where $S(A)$ is the entanglement entropy of a region A in a CFT and γ_A is an extremal surface in the bulk of the dual space-time that has the same boundary as A. This formula connects a classical geometric quantity like an area

of extremal surface in the gravity picture (AdS) to a quantum phenomenon, namely the entanglement entropy of spatial subsystems in the dual CFT.

Here, we shall focus on the study of entanglement entropy on noncommutative (fuzzy) spaces in the hope that it gives us insight into UV-IR connection and the scrambling behaviour of black hole horizons [10] and ultimately help in understanding the nature of spacetime at very short distances. The main advantage of working with noncommutative spaces is that the field theory is UV-finite and there is a natural infra-red (IR) ultraviolet (UV) connection. Noncommutative geometry is also interesting in the context of string theory as it is the natural description of the position of D-branes [11, 12]. Furthermore, field theories in noncommutative spaces [12, 13] are realized in the matrix quantum mechanics models [14–16], which give nonperturbative definitions of string theory. One of the simplest fuzzy spaces we can study is a fuzzy sphere S^2_N [17], described by its radius R and an integer "noncommutativity" parameter N. It approaches the classical sphere in the limit $N \to \infty$ for fixed R. A scalar field theory on the fuzzy sphere is given by a matrix model, where the matrix size N plays the role of a UV cutoff. In matrix models we have $D+1$ dimensional emergent geometry from a $0+1$ quantum mechanical system, with the geometry being noncommutative at finite N. In other words, spacetime curvature emerges from the collective dynamics of matrix degrees of freedom.

Generally, local field theories generally obey area law [18, 19], where the leading divergence in the entanglement entropy of a spatial region is proportional to the area of the boundary of that region. That is, $S(A) \sim |\partial A|/\epsilon^{d+1}$, where $|\partial A|$ is the area of the boundary of the region A, ϵ is a UV cutoff, and d is the space dimension. Violations of area law are oberved in some fermionic systems [20, 21] and in QFTs with nonlocal interactions (if the region considered has size smaller than the length scale of the nonlocality) [22, 23]. Entanglement entropy is therefore a good measure of non-locality in QFTs and we expect it to violate area law on a fuzzy sphere. Indeed, departure from the area law on noncommutative spaces was shown both through holographic calculations [24, 25] and field theory calculation [26]. In [3, 27], it was observed that for entanglement entropy on a fuzzy sphere there is a transition from extensive behaviour for small regions to area-law behaviour for larger regions. In other words for small sizes of the polar cap, $S(A) \sim |A|/\epsilon^{d+1}$ ($\sim 1-\cos\theta$); while for larger polar caps $S(A) \sim |\partial A|/\epsilon^{d+1}$ ($\sim \sin\theta$). In this work we try to address why recent results of [2] do not exhibit a similar departure from area-law despite a non-local nature of mini-BMN matrix model. Possible reasons are: presence of interaction terms, presence of a gauge and different factorization of Hilbert space.

This thesis is organized as follows. In chapter 2 we define field theory on a fuzzy sphere and discuss how to compute entanglement entropy. In chapter 3 we introduce the neural network approach of finding ground-state energy for a field theory, and subsequently computing entanglement entropy. In chapter 4 we present our results and compare two different projection matrices that we use to factorize Hilbert space when computing entanglement entropy. The code used to generate the data and make the figures can be accessed at `https://github.com/TomaszAnd/msc-thesis`

Chapter 2

Methodology

In this chapter we discuss how to compute entanglement entropy for a scalar field theory on a fuzzy sphere using so called real time approach [4] .

2.1 Scalar field theory on a fuzzy sphere

In this work we shall study one of the simplest nontrivial noncommutative field theories, i.e. the theory for a free scalar on a noncommutative (or fuzzy) sphere. Let us start by considering a free field theory on a normal commutative sphere S^2. It has the following Hamiltonian:

$$H = \frac{1}{2}\int d\Omega \left(\dot{\phi}^2 - R^{-2}\left(\mathcal{L}_i\phi\right)^2 + \mu^2\phi^2\right) , \tag{2.1}$$

where the dot denotes the time derivative, R is the radius of the sphere and $\mathcal{L}_i, (i = 1, 2, 3)$ are the angular momentum operators:

$$\begin{aligned} &\mathcal{L}_\pm \equiv \mathcal{L}_1 \pm i\mathcal{L}_2 = e^{\pm i\phi}\left(\pm\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\phi}\right) , \\ &\mathcal{L}_3 = -i\frac{\partial}{\partial\phi} , \end{aligned} \tag{2.2}$$

with θ, ϕ being the spherical coordinates with standard relation to Cartesian coordinates. The noncommutative sphere is obtained by replacing Cartesian coordinates x_i with operators X_i that are proportional to the generators of the N-dimensional irreducible representation of $SU(2)$ of spin $l = \frac{N-1}{2}$:

$$X_i = R\frac{L_i}{\sqrt{l(l+1)}}, [L_i, L_j] = i\epsilon_{ijk}L_k . \tag{2.3}$$

Notice that $L_iL_i = l(l+1)I = \frac{\left(N^2-1\right)}{4}I$, so that $X_iX_i = R^2I$, just like we have $x_ix_i = R^2$ for a commutative sphere. Since L_i generate rotations, the Laplacian $\nabla^2\phi$ for commuting sphere is replaced by the $SU(2)$ Casimir operator $\mathcal{L}_i^2 = \mathcal{L}_1^2 + \mathcal{L}_2^2 + \mathcal{L}_3^2$ for a fuzzy sphere with action $\mathcal{L}_i(\phi) = [L_i, \phi]$, $\mathcal{L}_i^2(\phi) = [L_i, [L_i, \phi]]$. Similarly, integration on the fuzzy sphere is a trace

$$\frac{4\pi R^2}{N} \operatorname{Tr}(\cdot) , \tag{2.4}$$

with the prefactor ensuring that the identity function maps to the unit matrix.

The Hamiltonian for a real scalar field theory on a fuzzy sphere is therefore given by

$$H = \frac{4\pi R^2}{N} \frac{1}{2} \operatorname{Tr}\left(\dot{\Phi}^2 - R^{-2}\left[L_i, \Phi\right]^2 + \mu^2 \Phi^2\right) , \tag{2.5}$$

where Φ is an $N \times N$ Hermitian matrix that represents the scalar field of mass μ. The theory has one dimensionless parameter, i.e. mass measured in units of radius $m = R\mu$. In the following we will take $R = 1$, yielding the dimensionless parameter to simply be $m = \mu$. We will work in a basis where L_3 is diagonal.

2.2 Entanglement Entropy in Free Scalar Field Theory

In this section we discuss how to calculate entanglement entropy for quadratic Hamiltonians, such as the one given by eq. 2.5, using an approach from [18].

2.2.1 Splitting Field Theory into Sectors

As shown in [28], we can treat the free field scalar theory as a collection of coupled oscillators on a lattice of space points by splitting the Hamiltonian 2.5 into a sum of $2N+1$ independent sectors. This simplifies the problem of finding EE for a fuzzy sphere as it allows to use the method for computing EE for quadratic Hamiltonians from [18, 29].

We start by splitting the fields into a symmetric part $\phi^{(1)}_{ij} = \phi^{(1)}_{ji}$ and an antisymmetric part $\phi^{(2)}_{ij} = -\phi^{(2)}_{ji}$ by writing $(\tilde{\phi})_{ij} = \phi^{(1)}_{ij} + i\phi^{(2)}_{ij}$. Then recombine into the real field $\Phi_{ij} = \phi^{(1)}_{ij} + \phi^{(2)}_{ij}$ and write

$$Q^{(m)} = (\Phi_{1,1+m}, \Phi_{2,2+m}, \ldots, \Phi_{N-m,N}) , Q^{(-m)} = (\Phi_{1+m,1}, \Phi_{2+m,2}, \ldots, \Phi_{N,N-m}) , \tag{2.6}$$

so that the degrees of freedom are the fields $Q^{(m)}_a$ and $Q^{(-m)}_a$ ($m \geq 0$). That means that in our calculations we will split the matrix Φ into: the main diagonal $Q^{(0)}$ of size N, diagonals $Q^{(-1)}$ and $Q^{(1)}$ of size $N-1$, diagonals

$Q^{(-2)}$ and $Q^{(2)}$ of size $N-2, \cdots$ and diagonals $Q^{-(N-1)}$ and $Q^{(N-1)}$ of size 1. In summary, we have the following relations:

$$\begin{aligned} 2Q_a^{(m)} &= 2\Phi_{a,a+m} = (1-i)\phi_{a-1,a+m-1} + (1+i)\phi^*_{a-1,a+m-1} \ , \\ 2Q_a^{(-m)} &= 2\Phi_{a+m,a} = (1-i)\phi_{a+m-1,a-1} + (1+i)\phi^*_{a+m-1,a-1} \ , \end{aligned} \tag{2.7}$$

Finally, define

$$\begin{aligned} c_2 &= l(l+1) \ , \\ A_a &= -a + \frac{N+1}{2} \ , \\ B_a &= \sqrt{a(N-a)} \ , \end{aligned} \tag{2.8}$$

where c_2 is the quadratic Casimir of the spin-l irreducible representation of $SU(2)$, A_a and B_a are the non-zero elements of L_3 and $L_\pm$, respectively, and $1 \le a \le N$. Then the original Hamiltonian (2.5) may be written as:

$$H = \sum_{m=-(N-1)}^{N-1} H_m = \frac{1}{2} \sum_{m=-(N-1)}^{N-1} \sum_{a,b=1}^{N-|m|} \left(\pi_a^{(m)} \delta_{ab} \pi_b^{(m)} + Q_a^{(m)} V_{ab}^{(m)} Q_b^{(m)} \right) \ , \tag{2.9}$$

where

$$V_{ab}^{(m)} = \left[2\left(c_2 + \frac{\mu^2}{2} - A_a A_{a+|m|} \right) \delta_{a,b} - B_{a-1} B_{a-1+|m|} \delta_{a-1,b} - B_a B_{a+|m|} \delta_{a+1,b} \right] \ , \tag{2.10}$$

and $\pi_a^{(m)} = \dot{Q}_a^{(m)}$. Since the Hamiltonian H decouples into H_m's that do not depend on the sign of m we may write:

$$H = \sum_{m=-(N-1)}^{N-1} H_m \ . \tag{2.11}$$

.

Now that we have mutually uncoupled diagonals $Q^{(m)}$, we can calculate entanglement entropy for each diagonal separately [4, 18], giving

$$S = \sum_{m=-(N-1)}^{N-1} S^{(m)} = S^{(0)} + 2 \sum_{m=1}^{N-1} S^{(m)} \ , \tag{2.12}$$

with $S^{(m)}$ being the entanglement entropy of the $m-$th sector [2]. In the next section we give a prescription for finding $S^{(m)}$ in the case of entanglement entropy and Rényi entropy.

[2] We can write $S^{(-m)} = S^{(m)}$ as we consider symmetric regions

2.2.2 Entropy of Quadratic Hamiltonians

Entanglement entropy

We find $S^{(m)}$ by writing down the ground state of H_m explicitly in terms of $V^{(m)}$ [4, 18]. Each sector H_m in eq. 3.18 has $N-|m|$ degrees of freedom as it represents $N-|m|$ coupled harmonic oscillators. Its normalized ground state is therefore:

$$\psi_0^{(m)}\left(Q^{(m)}\right) = \pi^{-N/4}(\det W^{(m)})^{1/4}\exp\left[-\frac{1}{2}Q_a^{(m)}W_{ab}^{(m)}Q_b^{(m)}\right] , \qquad (2.13)$$

where $W^{(m)}$ is the square root of $V^{(m)}$: if $V^{(m)} = U^\intercal V_D^{(m)} U$, where $V_D^{(m)}$ is diagonal and U is orthogonal, then $W^{(m)} = U^\intercal [V_D^{(m)}]^{1/2} U$. The corresponding density matrix is

$$\rho^{(m)}\left(Q^{(m)}, Q^{(m)\prime}\right) = \left[\det\frac{W^{(m)}}{\pi}\right]^{1/2}\exp\left[-\frac{1}{2}W_{ab}^{(m)}\left(Q_a^{(m)}Q_b^{(m)} + Q_a^{\prime(m)}Q_b^{\prime(m)}\right)\right] , \qquad (2.14)$$

and a reduced density matrix is obtained by integrating over the unavailable degrees of freedom Q_{A^c}:

$$\rho_{\mathrm{A}}^{(m)}\left(Q_{n+1}^{(m)}, Q_{n+1}^{(m)}, \ldots, Q_{n+1}^{\prime(m)}, Q_{n+2}^{\prime(m)}, \ldots\right) = \int\prod_{\alpha=1}^{n} dQ_\alpha^{(m)}\rho\left(Q^{(m)}, Q^{\prime(m)}\right) . \qquad (2.15)$$

The associated von Neumann entropy of ρ_A, defined by $S = -\operatorname{Tr}\rho_A\log\rho_A$, is the entanglement entropy which can be shown to be [18, 29]:

$$S^{(m)} = \sum_i\left[\log\left(\frac{1}{2}\sqrt{\lambda_i^{(m)}}\right) + \sqrt{1+\lambda_i^{(m)}}\log\left(\frac{1}{\sqrt{\lambda_i^{(m)}}} + \sqrt{1+\frac{1}{\lambda_i^{(m)}}}\right)\right] , \qquad (2.16)$$

where $\lambda_i^{(m)}$ are the eigenvalues of the matrix

$$\Lambda_{i,j}^{(m)} = -\sum_{\alpha=1}^{n}\left[W_{i\alpha}^{(m)}\right]^{-1}W_{\alpha j}^{(m)} , \qquad (2.17)$$

and $W_{\alpha j}^{(m)}$ and $\left[W_{i\alpha}^{(m)}\right]^{-1}$ are elements of $W^{(m)}$ and $\left[W^{(m)}\right]^{-1}$ respectively with $n+1 \le i,j \le N$ and $1 \le \alpha \le n$ ($\Lambda^{(m)}$ is an $(N-n)\times(N-n)$ matrix). And so we have outlined how to compute entanglement entropy of a fuzzy sphere.

Rényi entropy

To compute Rényi entropy we follow the real time approach [4] to the vacuum state for N coupled harmonic oscillators with a quadratic Hamiltonian

$$H = \frac{1}{2}\sum_{i=1}^{N}\pi_i^2 + \frac{1}{2}\sum_{i,j=1}^{N} x_i V_{ij} x_j \ . \tag{2.18}$$

Using two point correlators $[X_A]_{ij}$ and $[P_A]_{ij}$ for the region A defined by indices $1 \leq i, j \leq n$:

$$\begin{aligned} [X_A]_{ij} &= \langle x_i x_j \rangle = \frac{1}{2}\left(V^{-\frac{1}{2}}\right)_{ij} \ , \\ [P_A]_{ij} &= \langle \pi_i \pi_j \rangle = \frac{1}{2}\left(V^{\frac{1}{2}}\right)_{ij} \ , \end{aligned} \tag{2.19}$$

let us define

$$[C_A]_{ij} = \left[\sqrt{X_A P_A}\right]_{ij} = \frac{1}{2}\sqrt{\sum_{k=1}^{n}\left(V^{-1/2}\right)_{ik}\left(V^{1/2}\right)_{kj}} \ , \tag{2.20}$$

Then we may show that the entanglement entropy and Rényi entropy in region A are given in terms of the (positive) eigenvalues of $C_A = \sqrt{X_A P_A}$:

$$S\left[\rho_A\right] = \mathrm{tr}\left[\left(C_A + \frac{1}{2}\mathbf{1}\right)\log\left(C_A + \frac{1}{2}\mathbf{1}\right) - \left(C_A - \frac{1}{2}\mathbf{1}\right)\log\left(C_A - \frac{1}{2}\mathbf{1}\right)\right] \ , \tag{2.21}$$

$$S_n\left[\rho_A\right] = \frac{1}{n-1}\,\mathrm{tr}\log\left[\left(C_A + \frac{1}{2}\mathbf{1}\right)^n - \left(C_A - \frac{1}{2}\mathbf{1}\right)^n\right] \ . \tag{2.22}$$

The above equations are derived in [4] and we give a brief summary in the appendix B.

Partition of Hilbert space using projection matrices

The problem of computing entropies reduces then to separating the square root of the potential, $W^{(m)}$, into degrees of freedom inside and outside region A. We do that by writing degrees of freedom $Q^{(m)}$ as column vectors inside and outside region A in the eigenbasis of V; we do change of basis from $Q^{(m)}$ to diagonal $\tilde{Q}^{(m)}$.

$$W^{(m)} = [\tilde{Q}_1^{(m)}...\tilde{Q}_n^{(m)}\tilde{Q}_{n+1}^{(m)}...\tilde{Q}_N^{(m)}]^\intercal \begin{bmatrix} \omega_1^{\frac{1}{2}} & & \\ & \ddots & \\ & & \omega_{N-m}^{\frac{1}{2}} \end{bmatrix} [\tilde{Q}_1^{(m)}...\tilde{Q}_n^{(m)}\tilde{Q}_{n+1}^{(m)}...\tilde{Q}_N^{(m)}] \,, \tag{2.23}$$

where ω_i are eigenvalues of V. Similarly,

$$W^{-1} = \left[\tilde{Q}_1^{(m)}...\tilde{Q}_n^{(m)}\tilde{Q}_{n+1}^{(m)}...\tilde{Q}_{N-m}^{(m)}\right]^\intercal \begin{bmatrix} \omega_1^{-\frac{1}{2}} & & \\ & \ddots & \\ & & \omega_{N-m}^{-\frac{1}{2}} \end{bmatrix} \left[\tilde{Q}_1^{(m)}...\tilde{Q}_n^{(m)}\tilde{Q}_{n+1}^{(m)}...\tilde{Q}_{N-m}^{(m)}\right] \,. \tag{2.24}$$

To split $\tilde{Q}_a^{(m)}$ into degrees of freedom inside $(\tilde{Q}_1^{(m)}, \cdots, \tilde{Q}_n^{(m)})$ and outside $(\tilde{Q}_{n+1}^{(m)}, \cdots, \tilde{Q}_N^{(m)})$ region A, we use projection matrices from [2]. We explain how to find them in the next section. Note that we will have to put the projections in the block-diagonal form with blocks of size $N-m$ for $0 \leq m \leq N-1$.

2.2.3 Projections

To calculate entanglement for quantum states we need to be able to factorize the Hilbert space into the region inside and outside the polar cap A, i.e. $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_{A^c}$. In this section we review how to do that for a Hilbert space of states on a fuzzy sphere following [2]. Since a quantum state is a function from the configuration space $\mathcal{Q}$ to complex numbers, the Hilbert space of all quantum states is the square integrable functions $\mathcal{H} = L^2(\mathcal{Q})$. The Hilbert space $L^2(\mathcal{Q}) = L^2(\mathcal{Q}_A) \otimes L^2(Q_{A^c})$ may be then factorized via an orthogonal decomposition $\mathcal{Q} = \mathcal{Q}_A \oplus \mathcal{Q}_{A^c}$ defined by the orthogonal projection $P : \mathcal{Q} \to \mathcal{Q}$, such that $\mathcal{Q}_A = \operatorname{im} P$ and $\mathcal{Q}_{A^c} = \ker P$.

Normally to project any function f on a sphere to a polar cap A we could simply multiply by χ_A, $f \to f\chi_A$, where χ_A is the function on the sphere that is 1 on A and 0 otherwise. However, this is a valid choice for a projection only in the limit of $j_{\max} \to \infty$. With a finite cutoff $j_{\max}$, multiplying by χ_A generally takes the function out of the subspace of functions with $j \leq j_{\max}$. We therefore need an orthogonal projection $P_A^{j_{\max}} : \mathcal{Q}^{j_{\max}} \to \mathcal{Q}^{j_{\max}}$ that lives in the space of functions on the sphere spanned by spherical harmonics $Y_{jm}(\theta, \phi)$ with $j \leq j_{\max}$, call it $\mathcal{Q}^{j_{\max}}$. We define $P_A^{j_{\max}}$ as a projection that minimizes the distance $\|P_A^{j_{\max}} - P_A^{\infty}\|$ where P_A^{∞} is the multiplication by χ_A and $\|\cdot\|$ is chosen to be the Frobenius norm.

Let us now briefly present how to minimize Frobenius distance, in particular how to find an orthogonal projection operator P such that $\|P - R\|$ is minimal given another Hermitian operator R. If we let $R = UR'U^\dagger$ where U is a unitary matrix and R' is diagonal, then the following P minimizes $\|P - R\|_F$ among orthogonal projection operators:

$$P = UP'U^\dagger, \quad P' \text{ is diagonal with } P'_{ii} = 1 \text{ if } R'_{ii} > \frac{1}{2}, \text{ and 0 otherwise.} \tag{2.25}$$

This is the unique minimum if none of the eigenvalues of R is $\frac{1}{2}$.

So in summary, the projection operator P we will choose is the one that is both within the subspace spanned by spherical harmonics $Y_{jm}(\theta, \phi)$ with $j \leq j_{\max}$ and closest to P_A^∞. In other words we are looking for $P = P_A^{j_{\max}}$ that minimizes $\|P - P_A^\infty\|$, with the region A being a spherical cap with polar angle θ_A [3]. We compute P as follows:

1. Since for $j_1, j_2 \leq j_{\max}$, the projection $[P]_{j_1 m_1, j_2 m_2}$ should converge to its value at $j_{\max} \to \infty$, we start by finding the matrix elements $[P_A^\infty]_{j_1 m_1, j_2 m_2}$. This is what we will take as R in eq. 2.25; its matrix elements are that of multiplication by the unit function χ_A on a polar cap A:

$$[P_A^\infty]_{j_1 m_1, j_2 m_2} = \frac{1}{4\pi} \int_0^{\theta_A} d\theta \sin\theta \int_0^{2\pi} d\phi Y^*_{j_1 m_1}(\theta, \phi) Y_{j_2 m_2}(\theta, \phi) \ , \tag{2.27}$$

   where χ_A restricts the θ integral to $[0, \theta_A]$. For example P_A^∞ for $N = j_{\max} - 1 = 3$ has the following matrix elements:

[3] Note that for mini-BMN model there is an additional constraint for choosing the projection which is that it must preserve gauge directions $P'(G) \subset G$ with

$$G = \left\{ i \left[Y, L^i \right] : Y \text{ is Hermitian} \right\} \tag{2.26}$$

.

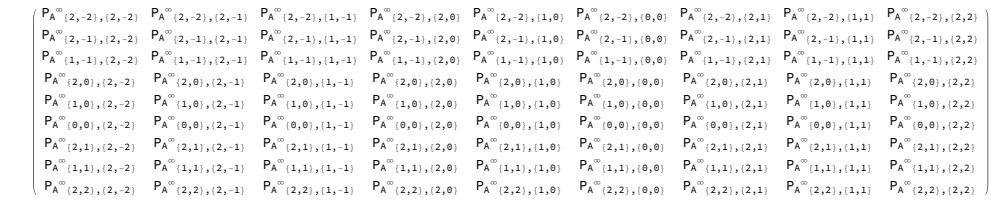

Figure 2.1: A "pseudoprojection" P_A^{∞} for $N = 3$ with matrix elements given by eq. 2.27. From there a projection $P_A^{j_{\max}}$ is obtained using eq. 2.25.

This matrix is obtained by writing all possible quantum numbers $\{j, m\}$ with $j \leq j_{\max}$, sorting the indices by m and then taking the outer product of the vector of indices with itself to get a matrix. Note that since the non-zero matrix elements are for $m_1 \neq m_2$, we get a block-diagonal matrix $R = P_A^{\infty}$ with $2N - 1$ blocks $R^{(m)}$ of size $N - |m|$ for each of the sectors. In practice, we only need to compute matrix of first N blocks since $R^{(-m)} = R^{(m)}$.

2. Using singular value decomposition, write R as $R = UR'V$ where R' is diagonal and U, V are unitary. Then the projection matrix is $P = UP'V^{\dagger}$ where

$$P' \text{ is diagonal with } P'_{ii} = 1 \text{ if } R'_{ii} > 1/2, \text{ and 0 otherwise.}$$

This is the projection $P_A^{j_{\max}}$ within the subspace $\mathcal{Q}^{j_{\max}}$ for a polar cap A.

From now on we shall refer to projections found using the method outlined above as $P_A^{\text{Frobenius}}$ as opposed to the projections used in [3] that we shall refer to as $P_A^{\text{Veigenbasis}}$. As demonstrated by fig. 2.2 and 2.3 the traces of these two projectors, which is a measure of modes in a region, have different dependence on the size of the region. Specifically, at large $j_{\max}$, for $\text{Tr } P_A^{\text{Frobenius}}$: $\text{Tr } P_A^{j_{\max}} \propto j_{\max}^2 |A|$ whereas for $P_A^{\text{Veigenbasis}}$: $\text{Tr } P_A^{j_{\max}} \propto j_{\max}^2 |A|^2$. This could be due to some eigenvalues being close to $1/2$ for $P^{\text{Frobenius}}$ before the final step in constrcuing the projection where we replace eigenvalues greater than $1/2$ with 1 and smaller than $1/2$ with 0 (see fig. 2.4).

We claim that the difference in traces is reflected in entropy and is the reason for disparity between the entanglement entropy observed in [2] ($P^{\text{Frobenius}}$) and [3] ($P^{\text{Veigenbasis}}$), i.e. area law ($S_\theta \propto \sin\theta$) and square of

area law ($S_\theta \propto \sin^2\theta$), respectively. (For the explanation why we have the square of area law entanglement entropy in scalar free field theory on a fuzzy sphere see [30]).

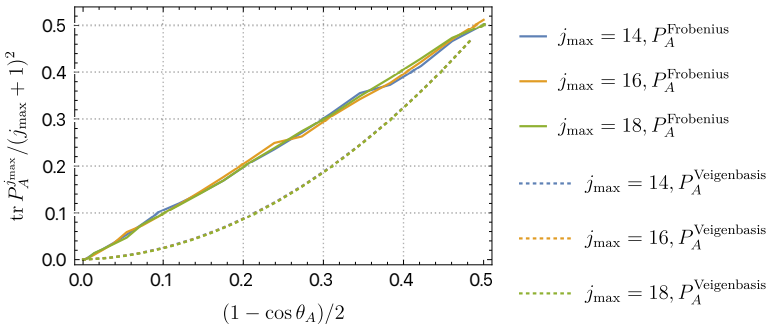

Figure 2.2: Trace of the projection $P_A^{\text{Frobenius}}$ and $P_A^{\text{Veigenbasis}}$ versus fractional area of the region (a spherical cap with polar angle θ_A), with different angular momentum cutoffs $j_{\max} = N - 1$.

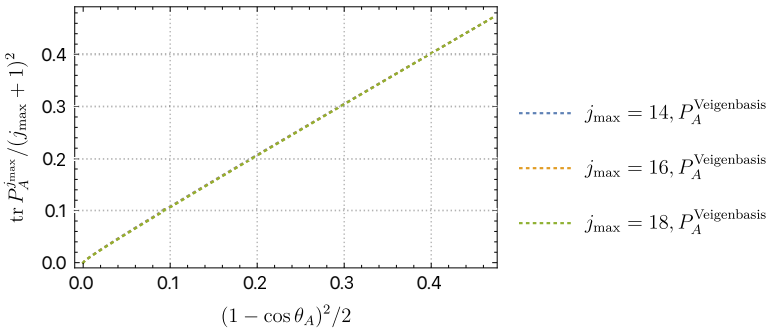

Figure 2.3: Trace of the projection $P_A^{\text{Veigenbasis}}$ versus square of fractional area of the region (a spherical cap with polar angle θ_A), with different angular momentum cutoffs $j_{\max} = N - 1$.

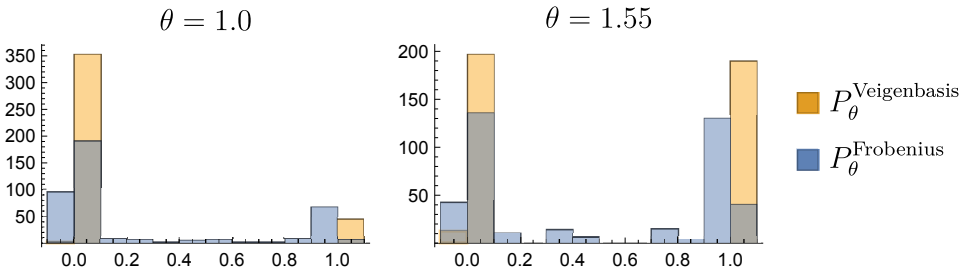

Figure 2.4: Histogram of eigenvalues for $N = 20$ of two pseudoprojections, i.e. before the final step of replacing the eigenvalues larger than a half with 1 and with 0 otherwise.

Chapter 3

Machine Learning Methodology

In this section we give a brief overview of the deep learning approach that was applied to the mini-BMN matrix model in [2]. To train our models we use the variational Monte Carlo, where for a given problem Hamiltonian H, the optimization strategy involves minimizing the expectation value $E_\theta = \langle\psi_\theta|H|\psi_\theta\rangle \geq E_0$ with respect to the variational parameters θ. Here, E_0 is the exact ground state energy of the Hamiltonian H. Starting from some wavefunction ansatz ψ_θ (usually some mixture of normal distributions) we update the variational parameters θ using variants of the gradient descent algorithm with the objective of minimizing the energy expectation value $E_\theta = \langle\psi_\lambda|H|\psi_\theta\rangle$ to find the best parameters θ^* corresponding to ψ_{θ^*} that most closely approximates the lowest energy eigenstate ψ_0. For more applications of variational Monte Carlo with parametrized neural network ansatz see e.g. [31, 32].

3.1 BMN Matrix Model

Let us first introduce the Hamiltonian for which the deep learning approach we will use in this work was developed [2]. The $SU(2)$ (fuzzy) sphere that we considered in a previous section is a saddle-point solution to BFSS matrix model [14]. The BFSS theory is a $0+1$ dimensional supersymmetric matrix theory (the dimensional reduction of the 10d super Yang-Mills) describing the dynamics of D0 branes. BFSS model is interesting to study as it gives a nonperturbative formulation of superstring theory and is a good candidate for testing the gauge/gravity correspondence [33, 34]. In this section we shall focus on its mass deformed version, so-called BMN model [35]. By adding a mass parameter to BFSS matrix model we go from a theory defined on a flat Minkowski spacetime to one defined on a curved spacetime. The advantage of studying BMN model is that it has a discrete energy spectrum and -

unlike BFFS model - a well defined canonical ensemble.[4]

The simplest possible version of BMN model (called mini-BMN [2, 37]) has 3 bosonic matrices and has the following Hamiltonian:

$$H = H_B + \mathrm{tr}\left(\lambda^\dagger \sigma^k \left[X^k, \lambda\right] + \frac{3}{2}\nu\lambda^\dagger\lambda\right) - \frac{3}{2}\nu\left(N^2-1\right) \ , \tag{3.1}$$

where H_B is the bosonic part

$$H_B = \mathrm{tr}\left(\frac{1}{2}\Pi^i\Pi^i - \frac{1}{4}\left[X^i, X^j\right]\left[X^i, X^j\right] + \frac{1}{2}\nu^2 X^i X^i + i\nu\epsilon^{ijk}X^iX^jX^k\right) \ , \tag{3.2}$$

and ν is mass deformation parameter (different from μ in eq. 2.5), σ^k are Pauli matrices, λ are fermionic degrees of freedom (matrices of two-component SO(3) spinors).

3.2 Neural Network Ansatz

As outlined earlier, the basic idea is to use a neural network as the wave function ansatz in the variational quantum Monte Carlo method to find a ground state of a Hamiltonian. This is performed as follows:

1. Represent a quantum state $|\psi_\theta\rangle$ in terms of variational parameters θ. The quantum wavefunction $\psi_\theta(X) = \langle X \mid \psi_\theta\rangle = |\psi(X)|e^{i\theta(X)}$ is a complex function of Hermitian matrices X with norm and phase modeled separately.

2. Estimate the energy from Monte Carlo samples of the wave function

$$E_\theta \equiv \left\langle \psi_\theta|\hat{H}|\psi_\theta\right\rangle = \int dX\,|\psi_\theta(X)|^2 \cdot \frac{\left\langle X|\hat{H}|\psi_\theta\right\rangle}{\psi_\theta(X)} = \mathop{\mathbb{E}}_{X\sim|\psi_\theta|^2}\left[\epsilon_\theta(X)\right] \ , \tag{3.3}$$

where $\mathbb{E}_{X\sim|\psi_\theta|^2}$ is the expectation value, with the random variable X drawn from the probability distribution $|\psi_\theta|^2$, $\epsilon_\theta(X)$ (so-called *local energy*) is defined as $\frac{\langle X|\hat{H}|\psi_\theta\rangle}{\psi_\theta(X)}$ and E_θ is estimated as the mean of $\epsilon_\theta(X)$ from these samples.

[4] This is because there are flat directions [36] in BFFS model while in the BMN model they are absent because of mass term.

3. Compute the gradient of the energy with respect to model parameters θ, noting that the sampling distribution depends on θ as well,

$$\nabla_\theta E_\theta = \mathbb{E}_{X\sim|\psi_\theta|^2}\left[\nabla_\theta \epsilon_\theta(X)\right] + \mathbb{E}_{X\sim|\psi_\theta|^2}\left[\epsilon_\theta(X)\nabla_\theta \ln|\psi_\theta|^2\right] . \quad (3.4)$$

4. Update parameters θ (if using gradient descent) according to the rule:

$$\theta_{t+1} = \theta_t - \alpha\nabla_\theta E_\theta , \quad (3.5)$$

where $t = 1, 2, ...$ denotes the steps of training and the step size $\alpha > 0$ is the *learning rate*. In practice, instead of this simple gradient descent rule, we actually use the Adam optimizer [38] to implement the gradient updates as it gives better results.

5. Repeat steps 2 to 4 until E_θ converges. The (local) minimum found is then a variational upper bound for the ground state energy. Observables of physical interest are evaluated with respect to the optimal parameters after training.

The averages $\mathbf{E}[f(X)]$ are estimated as Monte Carlo sample averages $\sum_{i=1}^{K}\frac{1}{K}f(X_i)$ in stochastic gradient descents; their uncertainties are therefore $\sigma/\sqrt{K}$ where σ is standard deviation.

Training of the model is divided into three epochs, each of which consists of 5000 iterations. The learning rate is set to be 10^{-3} for iterations from 1 to 5000, 2×10^{-4} from 5001 to 10000 and 4×10^{-5} from 10001 to 15000. In each iteration the energy is evaluated from a batch of 10^3 random samples. The final expectation value of energy for the trained variational wavefunction is evaluated from 10^6 Monte Carlo samples.

3.3 Estimating Rényi entropy using Replica Trick

Having found the ground state quantum wavefunction using the neural network ansatz as described in the previous section, we can sample from the wavefunction at different points to estimate Rényi n entropy

$$S_n(\rho) = \frac{1}{1-n}\ln \mathrm{Tr}\,\rho_A^n . \quad (3.6)$$

In the following section we describe how to do this at integer orders $n \geq 2$ via the replica trick [39].

Let x and y denote the coordinates of the subsystem A and its complement A^{c}, determined by the projection matrix P_A. The reduced density matrix is then

$$\rho_A\left(x, x'\right) = \int dy \psi(x+y)\psi^*\left(x'+y\right) \ , \tag{3.7}$$

where $x, x' \in Q_A = \operatorname{im} P_A$ and the integral is over the subspace $Q_{A^{\mathsf{c}}} = \ker P_A$. And so we have:

$$\operatorname{Tr} \rho_A^n = \prod_{i=0}^{n-1} \int dx_i dy_i \psi(x_i, y_i)\psi^*(x_{i+1}, y_i) \ , \tag{3.8}$$

where $x_n = x_0$. This can be estimated using replica trick [39], where we consider the action of Swap_A on two copies of the neural network wave function:

$$\operatorname{Swap}_A \psi(x,y)\psi(x',y') = \psi(x',y)\psi(x,y') \ . \tag{3.9}$$

This operator swaps the degrees of freedom in the region A between the two copies. Rényi-2 entropy

$$S_2\left(\rho_A\right) = -\ln \int dx dx' dy dy' \psi(x+y)\psi^*\left(x'+y\right)\psi\left(x'+y'\right)\psi^*\left(x+y'\right) \ , \tag{3.10}$$

is then simply:

$$S_2\left(\rho_A\right) = -\ln \langle \operatorname{Swap}_A \rangle \ , \tag{3.11}$$

which we can estimate using Monte Carlo sampling:

$$\langle \operatorname{Swap}_A \rangle = \mathbb{E}_{x_0,x_1,y_0,y_1} \frac{\psi^*\left(x_1,y_0\right)}{\psi^*\left(x_0,y_0\right)} \frac{\psi^*\left(x_0,y_1\right)}{\psi^*\left(x_1,y_1\right)} \approx \frac{1}{N_s} \sum_{k=1}^{N_s} \frac{\psi^*\left(x_1^{(k)}, y_0^{(k)}\right)}{\psi^*\left(x_0^{(k)}, y_0^{(k)}\right)} \frac{\psi^*\left(x_0^{(k)}, y_1^{(k)}\right)}{\psi^*\left(x_1^{(k)}, y_1^{(k)}\right)} \ , \tag{3.12}$$

where N_s is the number of samples in Monte Carlo (taken to be 10^5). Hence, to estimate Rényi-2 entropy with replica trick we only need to generate two sets of samples $\{x_0^{(k)}\}_{k=1}^{N_s}$ and $\{x_1^{(k)}\}_{k=1}^{N_s}$ independently from $|\psi_\theta|^2$. If we define

$$\operatorname{Swap}_A^{(k)} = \frac{\psi^*\left(x_1^{(k)}, y_0^{(k)}\right)}{\psi^*\left(x_0^{(k)}, y_0^{(k)}\right)} \frac{\psi^*\left(x_0^{(k)}, y_1^{(k)}\right)}{\psi^*\left(x_1^{(k)}, y_1^{(k)}\right)} \ , \tag{3.13}$$

the statistical error on such estimation is given by [32]:

$$\epsilon = \frac{1}{\langle \mathrm{Swap}_A \rangle} \sqrt{\frac{\mathrm{var}\left(\left\{\mathrm{Swap}_A^{(k)}\right\}\right)}{N_s}} \,. \tag{3.14}$$

Similarly, for Rényi-n entropy we have

$$\mathrm{Tr}\,\rho_A^n = \mathbb{E}_{x_i,y_i}\left[\prod_{i=0}^{n-1} \frac{\psi^*\left(x_{i+1}, y_i\right)}{\psi^*\left(x_i, y_i\right)}\right] \,. \tag{3.15}$$

To see how we separate degrees of freedom into $x \in Q_A$ and $y \in Q_{A^c}$ in 3.10, let us write $z = x + y \in Q$ and $z' = x' + y' \in Q$, so that

$$x = P_A z, \quad x' = P_A z', \quad y = (I - P_A)z, \quad y' = (I - P_A)z' \,. \tag{3.16}$$

Then the integral in eq. 3.10 can be performed over the full configuration space Q instead of Q_{A^c}:

$$S_2\left(\rho_A\right) = -\ln \int dz dz' \psi(z) \psi^*\left(P_A z' + (I - P_A)z\right) \psi\left(z'\right) \psi^*\left(P_A z + (I - P_A)z'\right) \,, \tag{3.17}$$

and then estimated using Monte Carlo sampling:

$$S_2\left(\rho_A\right) = -\ln \mathbb{E}_{z,z' \sim |\psi|^2}\left[\frac{\psi^*\left(P_A z' + (I - P_A)z\right) \psi^*\left(P_A z + (I - P_A)z'\right)}{\psi^*(z)\psi^*\left(z'\right)}\right] \,. \tag{3.18}$$

We use projection matrices described in section 2.2.3 to separate degrees of freedom for a free field theory on a fuzzy sphere.

Chapter 4

Results

4.1 Real-time Approach

In this section we present results of applying real-time approach to calculating entanglement entropy and Rényi entropies with the factorization of the Hilbert space as defined by projections used in [2] and described in section 2.2.3.

4.1.1 Entanglement Entropy

Figure 4.1 shows entanglement entropy for a polar cap region A as a function of the polar angle θ_A for different masses μ and size of matrix N in eq. 2.5. Angles larger than $\frac{\pi}{2}$ are not shown since $S(\theta) = S(\pi - \theta)$ for a pure state such as the vacuum.

For $\mu \leq 1$, $S(\theta_A)/N$ is (roughly) independent of N for both projections $P_\theta^{\text{Frobenius}}$ and $P_\theta^{\text{Veigenbasis}}$ but for $P_\theta^{\text{Frobenius}}$ entropy follows the area law whereas for $P_\theta^{\text{Veigenbasis}}$ entropy is extensive for smaller angles and area law for larger angles [26].

We further studied Rényi entropy which we computed using the method outlined in section 2.2.2. As shown in figure 4.2, Rényi entropy has exactly the same dependence on θ_A as entanglement entropy, with the only difference being in the magnitude. We find that approximately $S \sim 2S_2$.

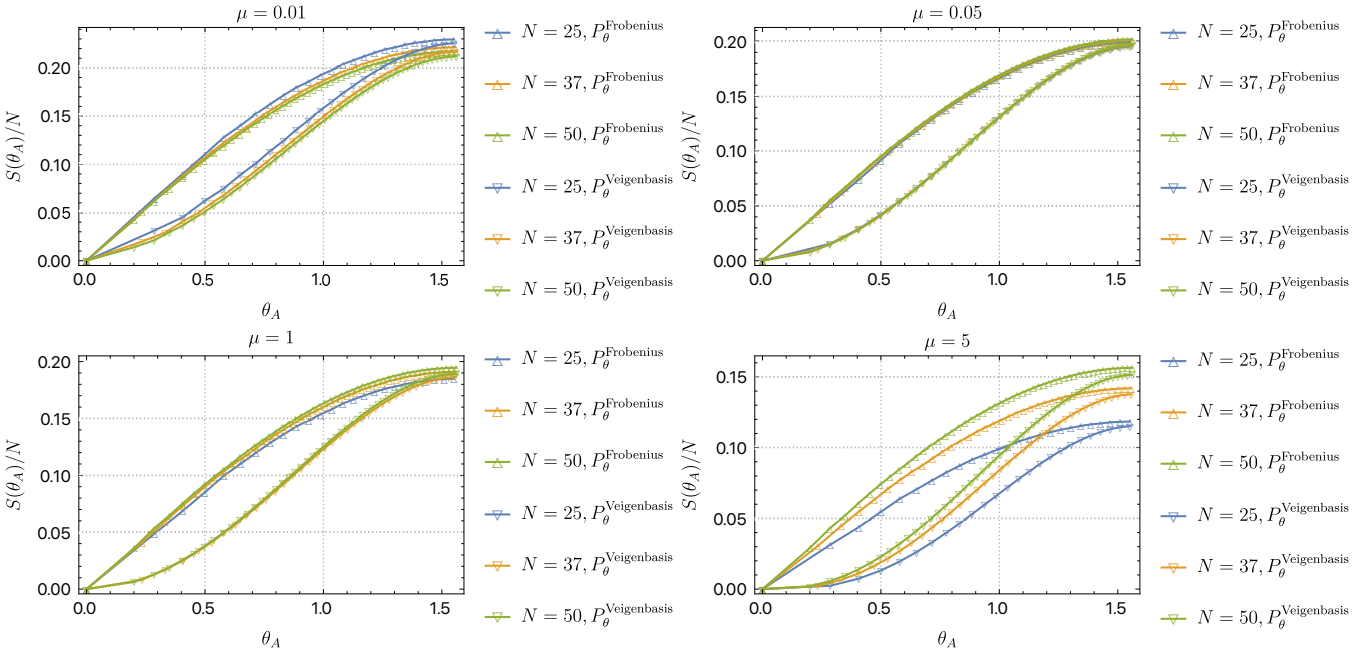

Figure 4.1: Scaled entanglement entropy S/N as a function of angular size θ_A of polar cap A for different μ's and N's.

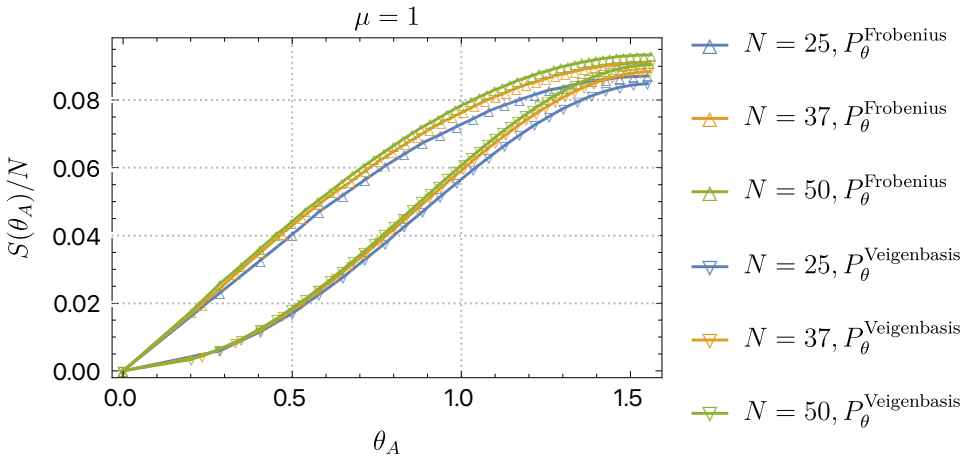

Figure 4.2: Scaled second Renyi entropy S_2/N as a function of angular size θ_A of polar cap A for $\mu = 1$ and different N's.

4.1.2 Mutual Information

A useful UV-finite quantity we can study is mutual information (MI). As it provides a bound on the range of correlations [40], we can think of it as a measure of non-locality over long distances (IR), similarly to entangelement entropy which was a measuere of non-locality over short distances (UV). MI characterizes the total amount of correlation between the subsystems A and B by measuring how the individual entropies of regions A and B differ from their joint entropy:

$$I(A:B) = S(A) + S(B) - S(A \cup B). \tag{4.1}$$

As shown in [27], mutual information is independent of the UV cutoff and it is the same for noncommutative and commutative sphere. In the following we will take two regions to be two polar caps A_1 and A_2 centered at opposite poles of the sphere:

$$I = S(A_1) + S(A_2) - S(A_1 \cup A_2) \ . \tag{4.2}$$

For convenience, we will choose A_1 and A_2 to have the same angular size θ_A. To calculate each term in the above equation we follow the method outlined in section 2.2. Note that to find projection operator $P_{A_1 \cup A_2}$ for $A_1 \cup A_2$ we minimize $\|P_{A_1 \cup A_2} - R\|_F$ just as described in 2.2.3, where the pseudoprojection R is the sum of projections $P_{A_1} + P_{A_2}$ found separately for

each polar cap. A good check on our construction for $P_{A_1 \cup A_2}$ is a trace, in particular we should have $\mathrm{Tr}\, P_{A_1 \cup A_2} = \mathrm{Tr}\, P_{A_1} + \mathrm{Tr}\, P_{A_2} = 2\,\mathrm{Tr}\, P_{A_1}$. This holds true for both projections, $P_\theta^{\text{Veigenbasis}}$ and $P_\theta^{\text{Frobenius}}$, as verified in fig. 4.3 where we plotted $2\,\mathrm{Tr}\, P_{A_1}$ and $\mathrm{Tr}\, P_{A_1 \cup A_2}$.

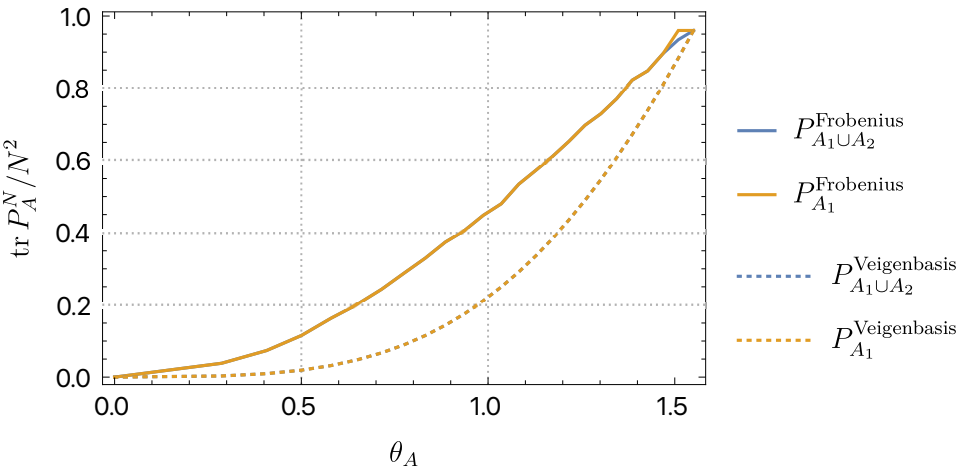

Figure 4.3: Trace of projection for union of north and south caps and double the trace of projection for north cap for both $P_\theta^{\text{Veigenbasis}}$ and $P_\theta^{\text{Frobenius}}$ (for $N = 25$).

As shown in [3, 27] mutual information between two caps should be independent of matrix size N. Figure 4.4 show that this is indeed the case, mutual information is independent of N for all masses μ and for both projections $P^{\text{Veigenbasis}}$ and $P^{\text{Frobenius}}$. Moreover, for small masses ($\mu \leq 0.5$) mutual information has almost the same value for all angles for both projections. This is further demonstrated by fig. 4.5 and 4.6.

At conformal point ($\mu = 0.5$) we also expect the mutual information for fuzzy sphere to be the same as for commutative sphere. In [27] it was shown that on a commutative sphere, mutual information for two polar caps separated by an annulus of angular width δ centered on the equator goes as:

$$I(\delta) = \begin{cases} 0.125 \cot \frac{\delta}{2} & \text{for } \delta \approx 0\ , \\ \frac{1}{12} \cot^2 \frac{\delta}{2} & \text{for } \delta \approx \pi\ , \end{cases} \tag{4.3}$$

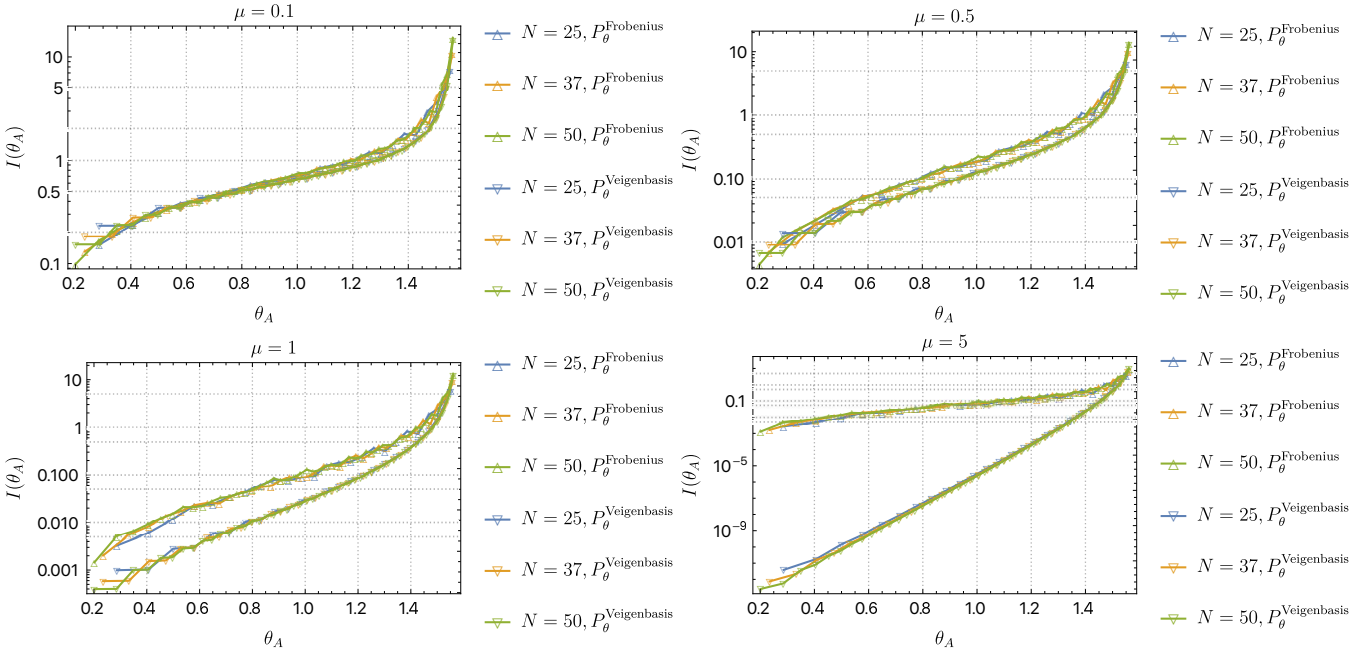

Figure 4.4: Mutual information I (on a logarithmic scale) as a function of common angular size $\theta_{A_1} = \theta_{A_2}$ of two spherical caps A_1 and A_2 centered at opposite poles of the sphere for different μ's and $N = 25, 37, 50$.

or in terms of common angular size of two polar caps $\theta = \frac{\pi-\delta}{2}$:

$$I(\theta) = \begin{cases} \frac{1}{12}\tan^2\theta & \text{for } \theta \approx 0\ , \\ 0.125\tan\theta & \text{for } \theta \approx \frac{\pi}{2}\ . \end{cases} \tag{4.4}$$

As shown by fig. 4.7 the results computed with $P^{\text{Veigenbasis}}$ give a slightly better match with the analytical prediction for a commutative sphere than the ones computed with $P^{\text{Frobenius}}$. Also, as illustrated by fig. 4.8 and 4.9, both $P^{\text{Frobenius}}$ and $P^{\text{Veigenbasis}}$ results are proportional to $\tan^2\theta$ but with different proportionality constants.

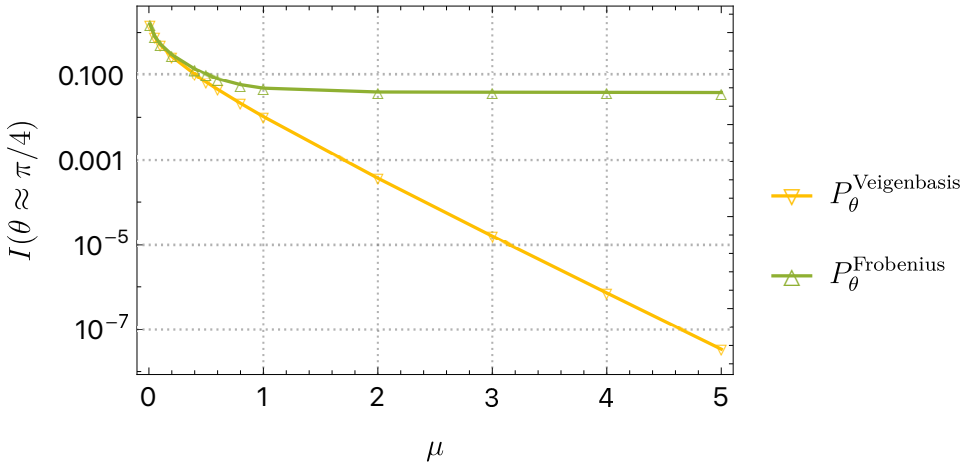

Figure 4.5: Mutual information I at $\theta \approx \pi/4$ for $N = 50$ as a function of μ.

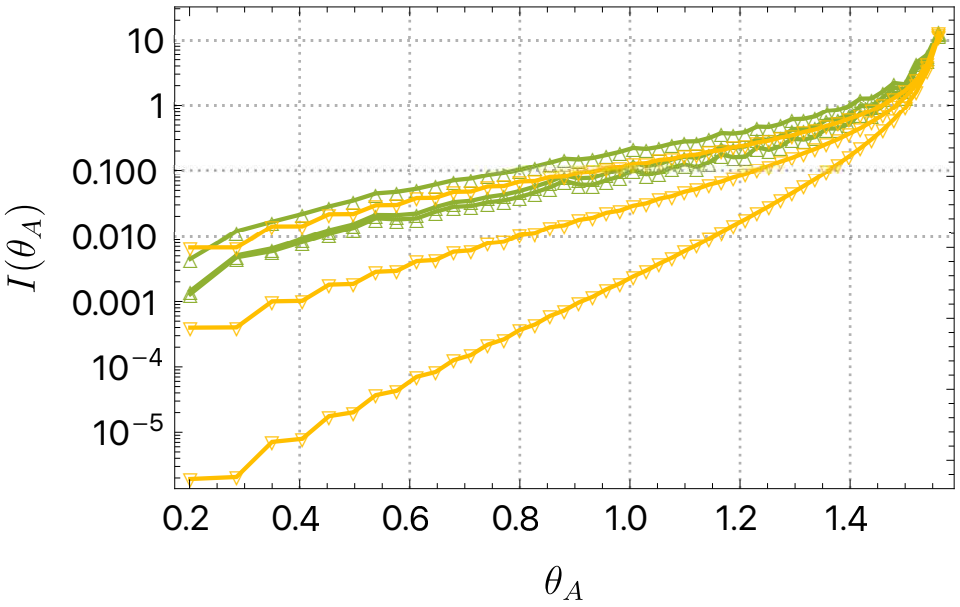

Figure 4.6: Mutual information I for $N = 50$ as a function of θ for masses $\mu = 0.5, 1, 2$ with mutual information decreasing as we increase mass. The results are computed with $P^{\text{Veigenbasis}}$ and $P^{\text{Frobenius}}$, denoted by yellow down triangle and green up triangle respectively.

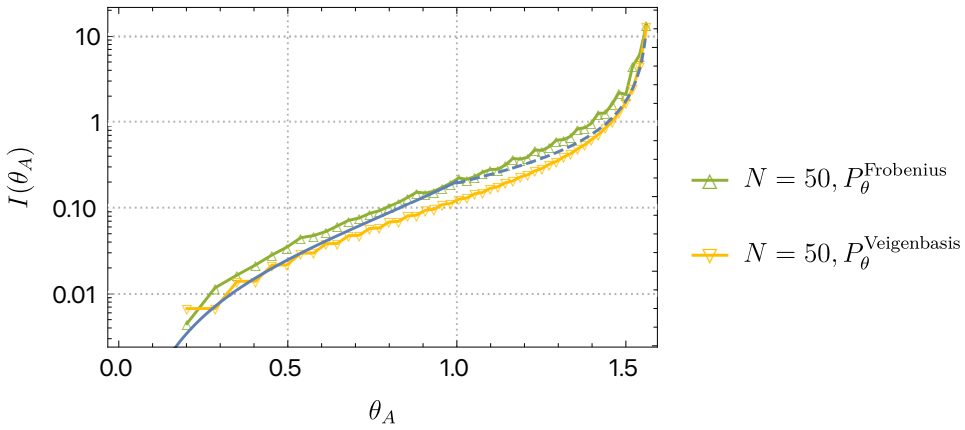

Figure 4.7: Mutual information I (on a logarithmic scale) as a function of common angular size $\theta_{A_1} = \theta_{A_2}$ of two spherical caps A_1 and A_2 centered at opposite poles of the sphere for $\mu = 1$ and $N = 50$. The solid and dashed lines correspond to the analytical predictions (eq. 4.5) for a commutative sphere at small and large θ, respectively.

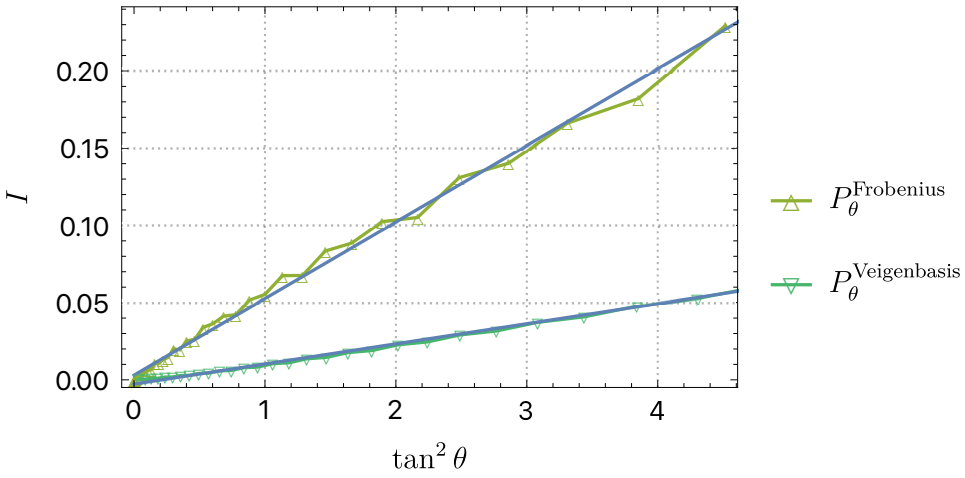

Figure 4.8: Mutual information I as a function of $\tan^2\theta$ of the common angular size $\theta_{A_1} = \theta_{A_2}$ of two spherical caps A_1 and A_2 centered at opposite poles of the sphere for $\mu = 1$ and $N = 50$ along with the best fit lines. Both I computed with $P^{\text{Veigenbasis}}$ and $P^{\text{Frobenius}}$ are proportional to $\tan^2(\theta)$ for $\theta < \frac{\pi}{3}$, i.e. $I(P_\theta^{\text{Veigenbasis}}) \approx 0.051 \tan^2(\theta)$, $I(P_\theta^{\text{Veigenbasis}}) \approx 0.012 \tan^2(\theta)$.

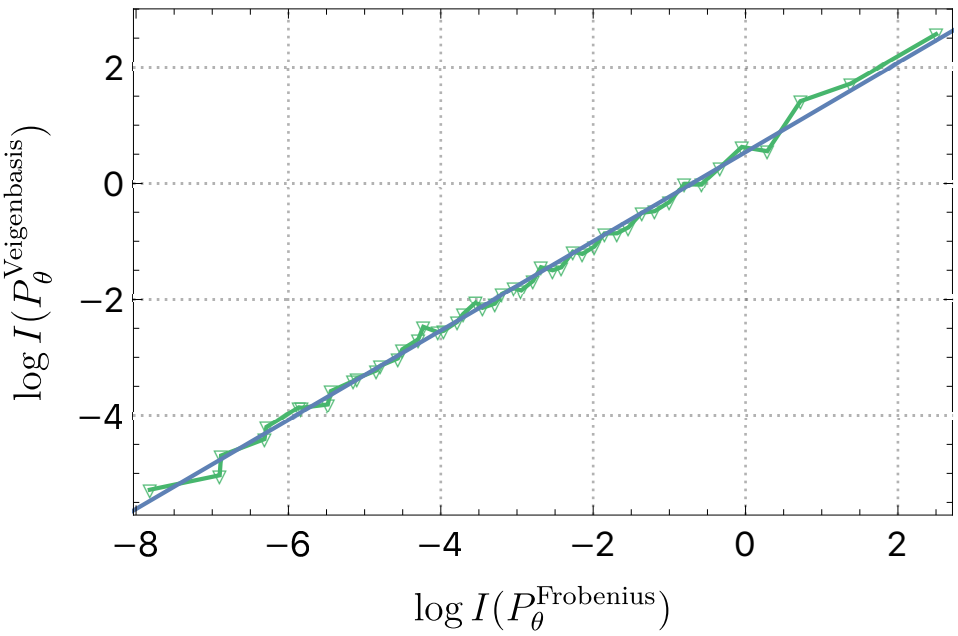

Figure 4.9: Logarithm of mutual information I computed with $P^{\text{Veigenbasis}}$ vs one computed with $P^{\text{Frobenius}}$ for common angular size of two spherical caps $\theta \in (0, \pi/2)$, $\mu = 1$ and $N = 50$ along with the best fit line, found numerically to be $\log I(P^{\text{Veigenbasis}}) = 0.77 \log I(P^{\text{Frobenius}}) + 0.54$.

4.2 Machine Learning

We apply neural network ansatz of [2] described in section 3.2 to a scalar model given by Hamiltonian

$$H = \frac{1}{2}\,\mathrm{Tr}\left((\dot{\phi})^2 - \sum_{i=1,2,3}[L_i,\phi]^2 + \mu^2\phi^2\right)\,, \tag{4.5}$$

where ϕ is an $N \times N$ Hermitian matrix representing a scalar field of mass μ.

We benchmark our model by comparing variational energies with the ground state energies of Hamiltonian 4.5. Recall from section 2.2.1 that we can write it as a sum of $2N-1$ coupled harmonic oscillators

$$H = \sum_{m=-(N-1)}^{N-1} H_m = \sum_{m=-(N-1)}^{N-1}\sum_{a,b=1}^{N-|m|}\left[\frac{1}{2}\left(\pi_a^{(m)}\right)^2 + \frac{1}{2}V_{ab}^{(m)}Q_a^{(m)}Q_b^{(m)}\right]\,. \tag{4.6}$$

The ground state of $N-|m|$ coupled harmonic oscillators with Hamiltonian

$$H_m = \sum_{a,b=1}^{N-|m|}\left[\frac{1}{2}\left(\pi_a^{(m)}\right)^2 + \frac{1}{2}V_{ab}^{(m)}Q_a^{(m)}Q_b^{(m)}\right]\,, \tag{4.7}$$

is found by solving the time-independent Schrodinger equation and is given in terms of of eigenvalues $\omega_i^{(m)}$ of $V_{ab}^{(m)}$:

$$E_m = \frac{1}{2}\sum_{i=1}^{N-|m|}\left[\omega_i^{(m)}\right]^{\frac{1}{2}}\,. \tag{4.8}$$

Hence, ground state energy of a free field theory on a fuzzy sphere (eq. 4.5) is:

$$E = \sum_{m=-(N-1)}^{N-1} E_m = E_0 + 2\sum_{m=0}^{N-1} E_m\,. \tag{4.9}$$

As shown by figure 4.10, we were able to match variational ground state energies found with neural networks to the expected energies. We used a normalizing flow (NF$(1,1)$) with a single layer and a single normal distribution in base mixed distribution. Increasing the number of layers or number of distributions did not significantly improve the results; nor did using an

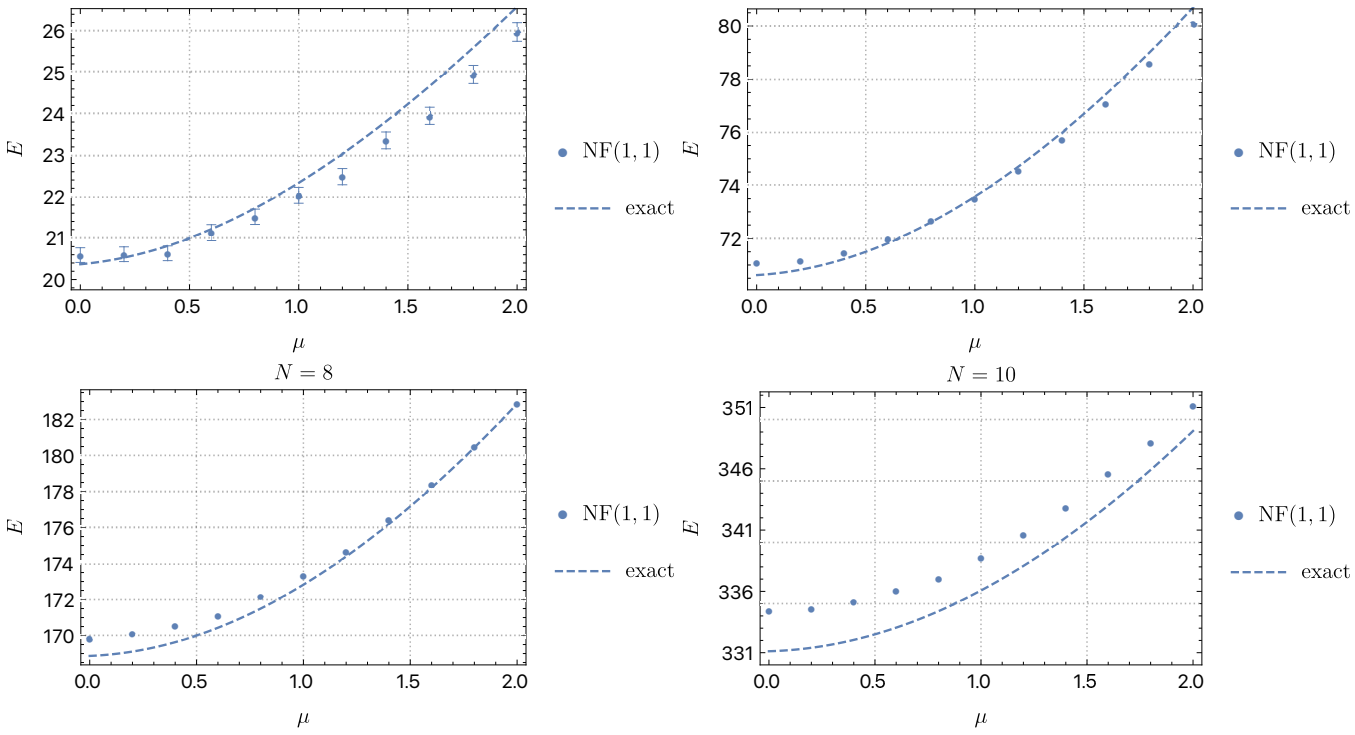

Figure 4.10: Variational ground state energies for free scalar field theory on a fuzzy sphere (eq. 4.5) for different size of the matrix N. As in [2], NF$(1,1)$ denotes a normalizing flow with 1 layer in the neural networks and 1 generalized normal distribution in each base mixed distribution. The dashed line denotes energies computed using eq. 4.9. For $N \geq 6$ uncertainties are below the scale of the markers; in particular the variational energies slightly below the dashed line are within numerical error of the line.

autoregressive flow rather than a normalizing flow (see [2] for more details regarding the neural network architecture).

Another benchmark of our model was to see if the neural network wavefunction is Gaussian. Recall from section 2.2.1 that the Hamiltonian 4.5 can be written as a sum of harmonic oscillators

$$H = \sum_m H_m \ , \tag{4.10}$$

where each H_m has a ground state of

$$\psi_0^{(m)}\left(Q^{(m)}\right) = \pi^{-N/4}(\det W^{(m)})^{1/4}\exp[-Q_a^{(m)}W_{ab}^{(m)}Q_b^{(m)}/2] \ . \tag{4.11}$$

Then the ground state of the field can be characterized by a wave functional which is the product of the ground state wave functions of all modes

$$\begin{aligned}\Psi[\phi] &= \prod_m \psi_0^{(m)}\left(Q^{(m)}\right) = \prod_m \pi^{-N/4}(\det W^{(m)})^{1/4}\exp[-\frac{1}{2}Q_a^{(m)}W_{ab}^{(m)}Q_b^{(m)}] \\ &= \left[\prod_m \pi^{-N/4}(\det W^{(m)})^{1/4}\right]\exp\left[-\frac{1}{2}\sum_m Q_a^{(m)}W_{ab}^{(m)}Q_b^{(m)}\right] \ .\end{aligned} \tag{4.12}$$

In fig. 4.11 we plot $\ln\Psi_{\text{Gaussian}}$ against $\ln\Psi_{\text{NN}}$. The best fit-line is:

$$\ln|\Psi_{\text{Gaussian}}| = a\ln|\Psi_{\text{NN}}| + b \ , \tag{4.13}$$

where slope a is found to be 1 and y-intercept b depends on the size of the matrix N with $b = \{-\frac{1}{2}, -2, ...\}$ for $N = \{2, 4, ...\}$. Therefore, the norms of neural network wavefunctions agree up to some irrelevant factor b: $|\Psi_{\text{Gaussian}}| = e^b|\Psi_{\text{NN}}|$.

We can also quantify the accuracy of the neural network states using the relative error on the ground-state energy:

$$\epsilon \equiv \frac{|E_{\text{NN}} - E_{\text{exact}}|}{|E_{\text{exact}}|} \ , \tag{4.14}$$

and the energy variance

$$\sigma^2 \equiv \left\langle \hat{H}^2 \right\rangle - \langle \hat{H} \rangle^2 \ . \tag{4.15}$$

In appendix B we show how they depend on batch size and number of samples in the neural network and how that correlates with $|\ln\Psi_{\text{NN}}|$.

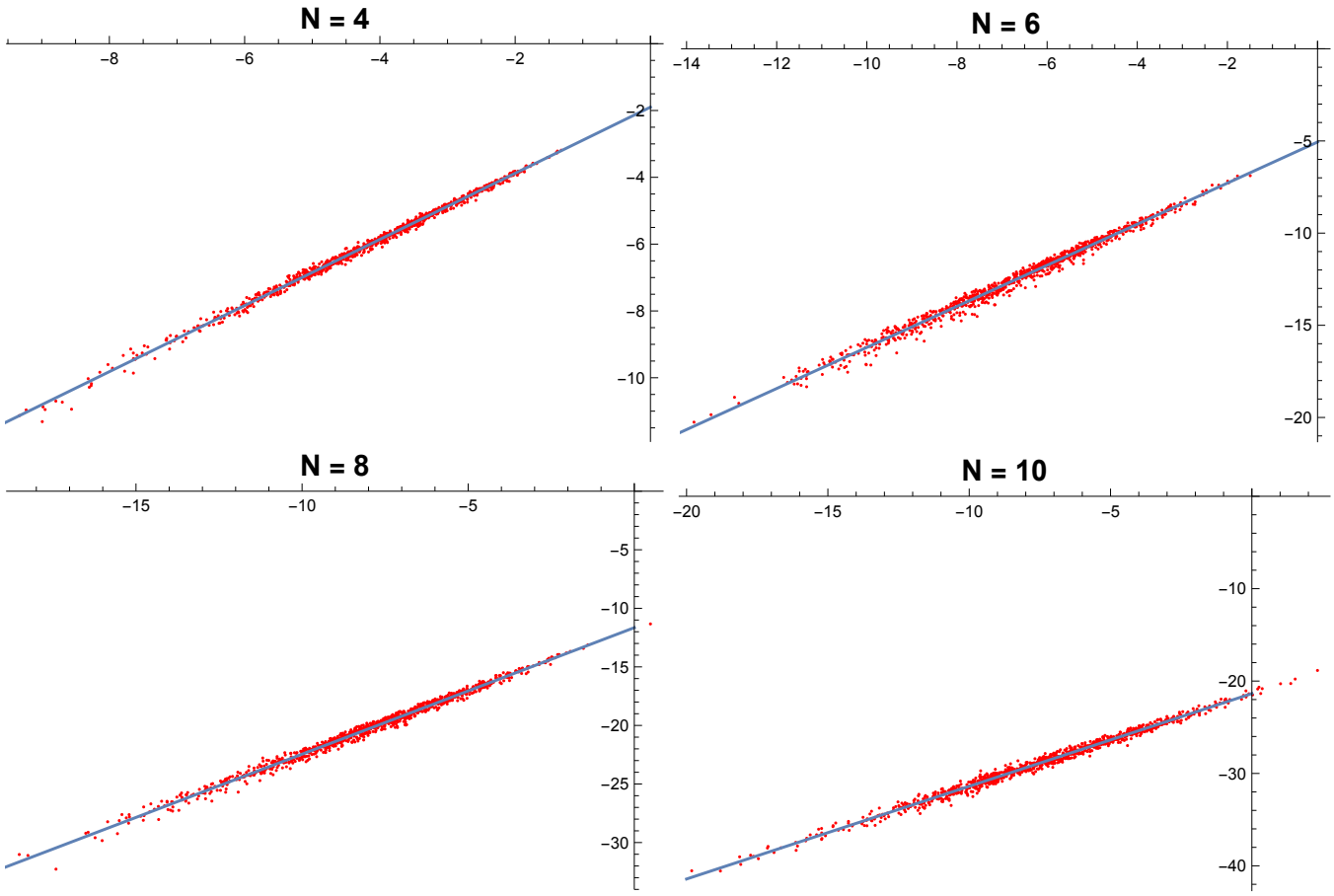

Figure 4.11: Scatter plot of logarithms of norms of wavefunctions: $\ln|\Psi_{\text{Gaussian}}|$ (eq. 4.12) vs. $|\ln\Psi_{\text{NN}}$ (variational wavefunction) for $N = \{4, 6, 8, 10\}$. Solid line is the best fit line with slope of approximately 1 for all N and irrelevant y-intercept that varies with N.

4.3 Matrix model on a fuzzy sphere with a quartic interaction

We also considered eq. 4.5 with an additional quartic interaction determined by the parameter λ:

$$H = \mathrm{Tr}\left(\frac{1}{2}(\dot{\phi})^2 - \frac{1}{2}\sum_{i=1,2,3}[L_i,\phi]^2 + \frac{1}{2}\mu^2\phi^2 + \lambda\phi^4\right) . \tag{4.16}$$

We can think of it as a matrix model which is in-between free field non-interacting single matrix theory (eq. 2.5) and interacting gauged three matrices theory (eq. 2.5). We will use the same neural network variational ansatz that we used for a free field theory. Motivation for studying interacting fields, because the discovered UV/IR anomaly [41, 42] that is a counterpart of the UV/IR mixing in field theories on compact noncommutative manifolds arises from the interactions To benchmark the variational ground state wavefunction for eq. 4.16 we can compare it to the Hamiltonian with no Laplacian ($-\frac{1}{2}\sum_{i=1,2,3}[L_i,\phi]^2$) term. As outlined in [5], an ungauged one-matrix model with a Hamiltonian:

$$H(M) = \mathrm{Tr}\,\frac{1}{2}\dot{M}^2 + \mathrm{Tr}\,V(M) , \tag{4.17}$$

where M is a hermitian $N\times N$ matrix and $V(M)$ is any polynomial function of M, in our case:

$$V(M) = \mathrm{Tr}\left(\frac{1}{2}M^2 + \lambda M^4\right) , \tag{4.18}$$

may be rewritten in terms of eigenvalues λ_i of M:

$$H = \sum_{i=1}^{N}\frac{1}{2}\dot{\lambda_i}^2 + V(\lambda_i) . \tag{4.19}$$

Then we can get the ground state energy eigenvalue and eigenstate by considering the the first N excited states of the analogue Hamiltonian:

$$H = \frac{1}{2}\pi^2 + \frac{1}{2}x^2 + \lambda x^4 . \tag{4.20}$$

(For $\lambda = 0$, this would be a single harmonic oscillator.) The ground state energy is therefore

$$E = \sum_{n=1}^{N} E_n , \tag{4.21}$$

where E_n is the n-th excited eigenvalue of eq. 4.20. The ground state wavefunction of 4.16 is given in terms of n-th excited eigenstates $\psi_n(\lambda_{\sigma(n)})$ of 4.20:

$$\Psi(M) = \Delta(\lambda)^{(-1)} \sum_{\sigma \in S_N} (-1)^{\sigma} \prod_{n=0}^{N-1} \psi_n(\lambda_{\sigma(n)}) \,, \tag{4.22}$$

where S_n is symmetric group and

$$\Delta(\lambda) = \prod_{1 \le i < j \le n} (\lambda_j - \lambda_i) = \prod_{i<ja} (\lambda_j - \lambda_i) \,, \tag{4.23}$$

is called a Vandermonde determinant. Note that $\sum_{\sigma \in S_N} (-1)^{\sigma} \prod_{n=0}^{N-1} \psi_n(\lambda_{\sigma(n)})$ is just the Leibniz formula for the determinant of matrix $\psi_n(\lambda_{n+1})$.

In fig. 4.12 and 4.13 we show that the variational neural network ground state wavefunction for eq. 4.16 is in good agreement with the predicted ground state energy and ground state eigenstate.

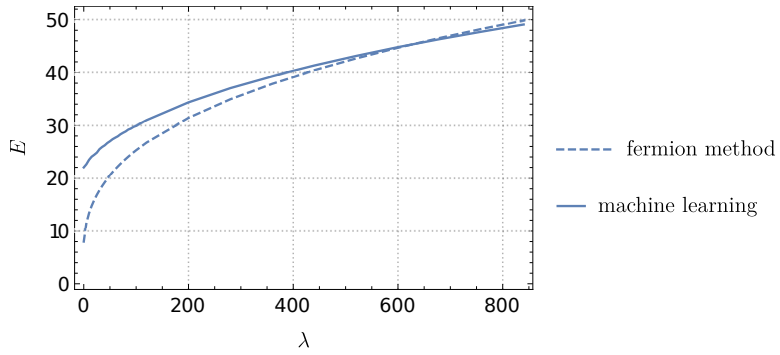

Figure 4.12: Expectation value of ground state energy of scalar field interacting theory on a fuzzy sphere (eq. 4.16) for $N = 4$ for optimized neural network compared to the ground state energy of the same Hamiltonian but with no laplacian term. The latter can be estimated by eq. 4.21. For small λ the laplacian is not negligent so the energies differ.

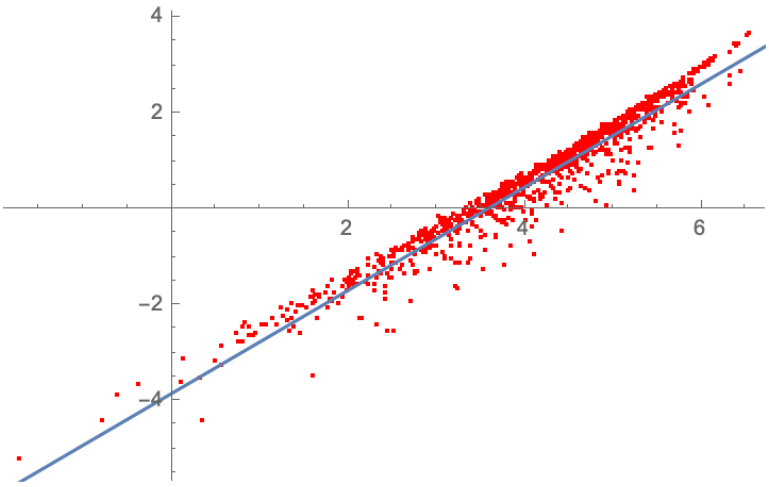

Figure 4.13: Scatter plot of log norms of ground state wavefunction of scalar field interacting theory on a fuzzy sphere (eq. 4.16) for $N = 4$ for optimized neural network vs log norms of the ground state eigenstate of the same Hamiltonian but with no laplacian term. The latter can be estimated by eq. 4.22. The slope is approximately 1.

Having found ground states for a Hamiltonian with a quartic interaction, we computed second Rényi entropy by sampling from the variational wavefunction using Monte Carlo (see section 3.3) and both projections $P^{\text{Veigenbasis}}$ and $P^{\text{Frobenius}}$. As demonstrated by fig. 4.14, for small λ and for S_2 is subextensive for small angles and extensive for large angles in agreement with the behaviour of entanglement entropy described in [26]. However, as we increase λ, Rényi entropy as computed with $P^{\text{Veigenbasis}}$ transitions from the square of area law to the area law. Interestingly, Rényi entropy as computed with $P^{\text{Frobenius}}$ seems independent of λ. Unlike previous Monte Carlo simulations of interacting theory on a fuzzy sphere [30, 43], we observed no change in magnitude of entropy between interacting and free theories (see fig. 4.2 and 4.14).

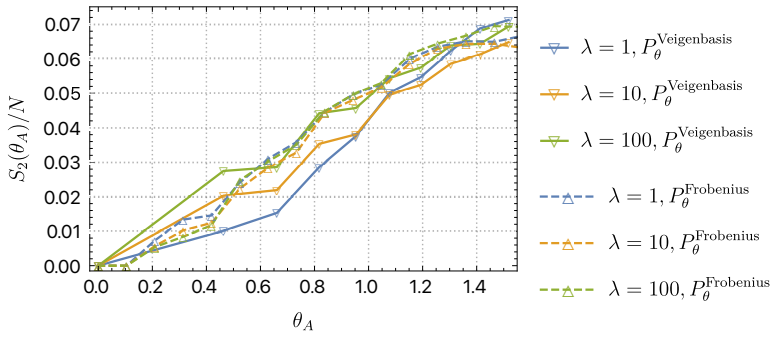

Figure 4.14: Scaled Rényi entropy S_2/N for a free field theory with quartic interaction (eq. 4.16) as a function of angular size θ_A of polar cap A for $N = 10$, $\mu = 1$ and $\lambda = \{1, 10, 100, 1000\}$ computed from variational neural network wavefunctions. The statistical error (eq. 3.14) is below the scale of the markers.

Additionally, we used $P^{\text{Veigenbasis}}$ to compute Rényi entropy for bosonic mini-BMN model in the limit of $\nu = \infty$ and compared it with the results computed with $P^{\text{Frobenius}}$ in [2]. Again, we observe different behaviour for two projections.

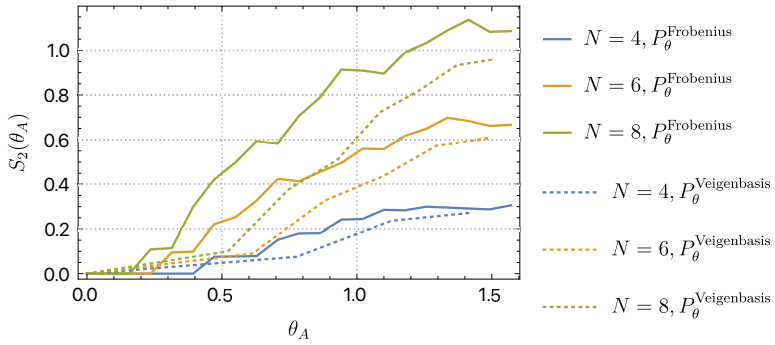

Figure 4.15: The second Rényi entropy S_2 for a spherical cap on the matrix theory fuzzy sphere versus the polar angle as a function of angular size θ_A. These are exact values at $\nu = \infty$ for mini-BMN model with no fermions (eq. 3.2) computed with $P^{\text{Frobenius}}$ (solid) and $P^{\text{Veigenbasis}}$ (dashed).

Chapter 5

Conclusion

In this thesis we explored the disparity of results for entropy in noncommutative theories between [26] and [2]. Reference [2] computed Rényi entropy for a mini-BMN model, which is a gauged model with interaction terms, using the projection matrices described in section 2.2.3 to separate the Hilbert space. The computed entropy followed area law which did not agree with the expected results for a non-local theory: extensive for small angles of a polar cap and area-law for larger angles as followed by scalar free field theory on a fuzzy sphere [26]. We examined why this is the case by applying projection matrices used in [2] to separate degrees of freedom (rather than using symbol map calculation from [26]) and compute EE of a scalar field theory on a fuzzy sphere using standard methods for computing entropy of quadratic Hamiltonians. We found that the results are largely dependent on the projection used. This agrees with different dependence on θ showed by the traces of projections $P_\theta^{\text{Frobenius}}$ and $P_\theta^{\text{Veigenbasis}}$ (see section 2.2.3).

We then studied mutual information (MI) of scalar massive free field theory on a fuzzy sphere and found that it is independent of N for both projections which is a good check for each method of separating degrees of freedom. We found that MI exhibits different dependence on the mass parameter μ for the two projections. For masses below the conformal point ($\mu = 0.5$) MI agrees for the two projections while for large masses they differ by orders of magnitude.

Lastly, we applied neural network ansatz from [2] to find the ground state of a scalar field theory with a quartic interaction on a fuzzy sphere and compute Rényi entropy. When using projection from [2] $P^{\text{Frobenius}}$ to separate degrees of freedom on a fuzzy sphere, Rényi entropy followed area law and was the same regardless of the value of the parameter of quartic interaction λ. When using projection from [3] $P^{\text{Veigenbasis}}$, we recovered the expected square of area law UV divergence in Rényi entropy for scalar free field theory and for small values quartic parameter λ; however for larger λ we obtained area law.

Some possible areas of research, that were beyond the scope of this work, remain to be explored. For example using the variational neural network

ansatz one can compute mutual information of interacting theory on a fuzzy sphere between two polar caps. As suggested in [30], we should find that once interaction is introduced, MI on a noncommutative sphere differs from the one on a commutative sphere [30]. One can also use variational Monte Carlo to approximate entanglement entropy using the best polynomial approximation (BPA) [31] (see appendix C for details). However, we do not expect that entanglement entropy behaves any differently than Renyi-2 entropy. Another interesting problem is to use the variational neural network wavefunction to compute entropy, via both $P^{\text{Frobenius}}$ and $P^{\text{Veigenbasis}}$, for BMN matrix model and its ungauged version [44]. We should get the same results for gauged and ungauged models.

Bibliography

[1] Shinsei Ryu and Tadashi Takayanagi. Holographic Derivation of Entanglement Entropy from AdS/CFT. *Physical Review Letters*, 96(18):181602, May 2006.

[2] Xizhi Han and Sean A. Hartnoll. Deep Quantum Geometry of Matrices. *Physical Review X*, 10(1):011069, March 2020.

[3] Hong Zhe Chen and Joanna L. Karczmarek. Entanglement entropy on a fuzzy sphere with a UV cutoff. *Journal of High Energy Physics*, 2018(8):154, August 2018.

[4] H. Casini and M. Huerta. Entanglement entropy in free quantum field theory. *Journal of Physics A: Mathematical and Theoretical*, 42(50):504007, December 2009.

[5] Joanna L Karczmarek. Scattering in the adjoint sector of the $c = 1$ matrix model. *Journal of High Energy Physics*, 2009(02):011–011, February 2009.

[6] Mark Van Raamsdonk. Building up spacetime with quantum entanglement. *General Relativity and Gravitation*, 42(10):2323–2329, October 2010.

[7] Thomas Faulkner, Monica Guica, Thomas Hartman, Robert C. Myers, and Mark Van Raamsdonk. Gravitation from Entanglement in Holographic CFTs. *Journal of High Energy Physics*, 2014(3):51, March 2014.

[8] J. Maldacena and L. Susskind. Cool horizons for entangled black holes. *Fortschritte der Physik*, 61(9):781–811, 2013.

[9] Veronika E. Hubeny, Mukund Rangamani, and Tadashi Takayanagi. A Covariant Holographic Entanglement Entropy Proposal. *Journal of High Energy Physics*, 2007(07):062–062, July 2007.

[10] Yasuhiro Sekino and Leonard Susskind. Fast Scramblers. *Journal of High Energy Physics*, 2008(10):065–065, October 2008.

[11] Michael R. Douglas and Nikita A. Nekrasov. Noncommutative Field Theory. *arXiv:hep-th/0106048*, October 2001.

[12] Alain Connes, Michael R. Douglas, and Albert Schwarz. Noncommutative Geometry and Matrix Theory: Compactification on Tori. *Journal of High Energy Physics*, 1998(02):003–003, February 1998.

[13] H. Aoki, N. Ishibashi, S. Iso, H. Kawai, Y. Kitazawa, and T. Tada. Noncommutative Yang-Mills in IIB Matrix Model. *Nuclear Physics B*, 565(1-2):176–192, January 2000.

[14] T. Banks, W. Fischler, S. H. Shenker, and L. Susskind. M Theory As A Matrix Model: A Conjecture. *Physical Review D*, 55(8):5112–5128, April 1997.

[15] N. Ishibashi, H. Kawai, Y. Kitazawa, and A. Tsuchiya. A Large-N Reduced Model as Superstring. *Nuclear Physics B*, 498(1-2):467–491, August 1997.

[16] R. Dijkgraaf, E. Verlinde, and H. Verlinde. Matrix String Theory. *Nuclear Physics B*, 500(1-3):43–61, September 1997.

[17] J. Madore. The fuzzy sphere. *Classical and Quantum Gravity*, 9(1):69–87, January 1992.

[18] Mark Srednicki. Entropy and Area. *Physical Review Letters*, 71(5):666–669, August 1993.

[19] J. Eisert, M. Cramer, and M. B. Plenio. Area laws for the entanglement entropy - a review. *Reviews of Modern Physics*, 82(1):277–306, February 2010.

[20] Michael M. Wolf. Violation of the entropic area law for Fermions. *Physical Review Letters*, 96(1):010404, January 2006.

[21] Noriaki Ogawa, Tadashi Takayanagi, and Tomonori Ugajin. Holographic Fermi Surfaces and Entanglement Entropy. *Journal of High Energy Physics*, 2012(1):125, January 2012.

[22] Wei Li and Tadashi Takayanagi. Holography and Entanglement in Flat Spacetime. *Physical Review Letters*, 106(14):141301, April 2011.

[23] Noburo Shiba and Tadashi Takayanagi. Volume Law for the Entanglement Entropy in Non-local QFTs. *Journal of High Energy Physics*, 2014(2):33, February 2014.

[24] Willy Fischler, Arnab Kundu, and Sandipan Kundu. Holographic Entanglement in a Noncommutative Gauge Theory. *Journal of High Energy Physics*, 2014(1):137, January 2014.

[25] Joanna L. Karczmarek and Charles Rabideau. Holographic entanglement entropy in nonlocal theories. *Journal of High Energy Physics*, 2013(10):78, October 2013.

[26] Joanna L. Karczmarek and Philippe Sabella-Garnier. Entanglement entropy on the fuzzy sphere. *Journal of High Energy Physics*, 2014(3):129, March 2014.

[27] Philippe Sabella-Garnier. Mutual information on the fuzzy sphere. *Journal of High Energy Physics*, 2015(2):63, February 2015.

[28] Djamel Dou and Badis Ydri. Entanglement entropy on fuzzy spaces. *Physical Review D*, 74(4):044014, August 2006.

[29] Luca Bombelli, Rabinder K. Koul, Joohan Lee, and Rafael D. Sorkin. Quantum source of entropy for black holes. *Physical Review D*, 34(2):373–383, July 1986.

[30] Mariko Suzuki and Asato Tsuchiya. A generalized volume law for entanglement entropy on the fuzzy sphere. *Progress of Theoretical and Experimental Physics*, 2017(4), April 2017.

[31] Zhaoyou Wang and Emily J. Davis. Calculating Renyi Entropies with Neural Autoregressive Quantum States. *Physical Review A*, 102(6):062413, December 2020.

[32] Mohamed Hibat-Allah, Martin Ganahl, Lauren E. Hayward, Roger G. Melko, and Juan Carrasquilla. Recurrent neural network wave functions. *Physical Review Research*, 2(2):023358, June 2020.

[33] Yuhma Asano, Samuel Kováčik, and Denjoe O'Connor. The Confining Transition in the Bosonic BMN Matrix Model. *Journal of High Energy Physics*, 2020(6):174, June 2020.

[34] Samuel Kováčik, Denjoe O'Connor, and Yuhma Asano. The non-perturbative phase diagram of the bosonic BMN matrix model. *arXiv:2004.05820 [hep-th]*, April 2020.

[35] David Berenstein, Juan Maldacena, and Horatiu Nastase. Strings in flat space and pp waves from ${\cal N}=4$ Super Yang Mills. *Journal of High Energy Physics*, 2002(04):013–013, April 2002.

[36] Miguel S. Costa, Lauren Greenspan, Joao Penedones, and Jorge Santos. Thermodynamics of the BMN matrix model at strong coupling. *Journal of High Energy Physics*, 2015(3):69, March 2015.

[37] Tarek Anous and Cameron Cogburn. Mini-BFSS in Silico. *Physical Review D*, 100(6):066023, September 2019.

[38] Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization, January 2017.

[39] Matthew B. Hastings, Ivan Gonzalez, Ann B. Kallin, and Roger G. Melko. Measuring Renyi Entanglement Entropy with Quantum Monte Carlo. *Physical Review Letters*, 104(15):157201, April 2010.

[40] Berry Groisman, Sandu Popescu, and Andreas Winter. On the quantum, classical and total amount of correlations in a quantum state, February 2005.

[41] P. Castro-Villarreal, R. Delgadillo-Blando, and Badis Ydri. A Gauge-Invariant UV-IR Mixing and The Corresponding Phase Transition For U(1) Fields on the Fuzzy Sphere. *Nuclear Physics B*, 704(1-2):111–153, January 2005.

[42] Chong-Sun Chu, John Madore, and Harold Steinacker. Scaling Limits of the Fuzzy Sphere at one Loop. *Journal of High Energy Physics*, 2001(08):038–038, August 2001.

[43] Shizuka Okuno, Mariko Suzuki, and Asato Tsuchiya. Entanglement entropy in scalar field theory on the fuzzy sphere. *Progress of Theoretical and Experimental Physics*, 2016(2):023B03, February 2016.

[44] Juan Maldacena and Alexey Milekhin. To gauge or not to gauge? *Journal of High Energy Physics*, 2018(4):84, April 2018.

Appendix A

Scaling of resources in neural network wavefunction

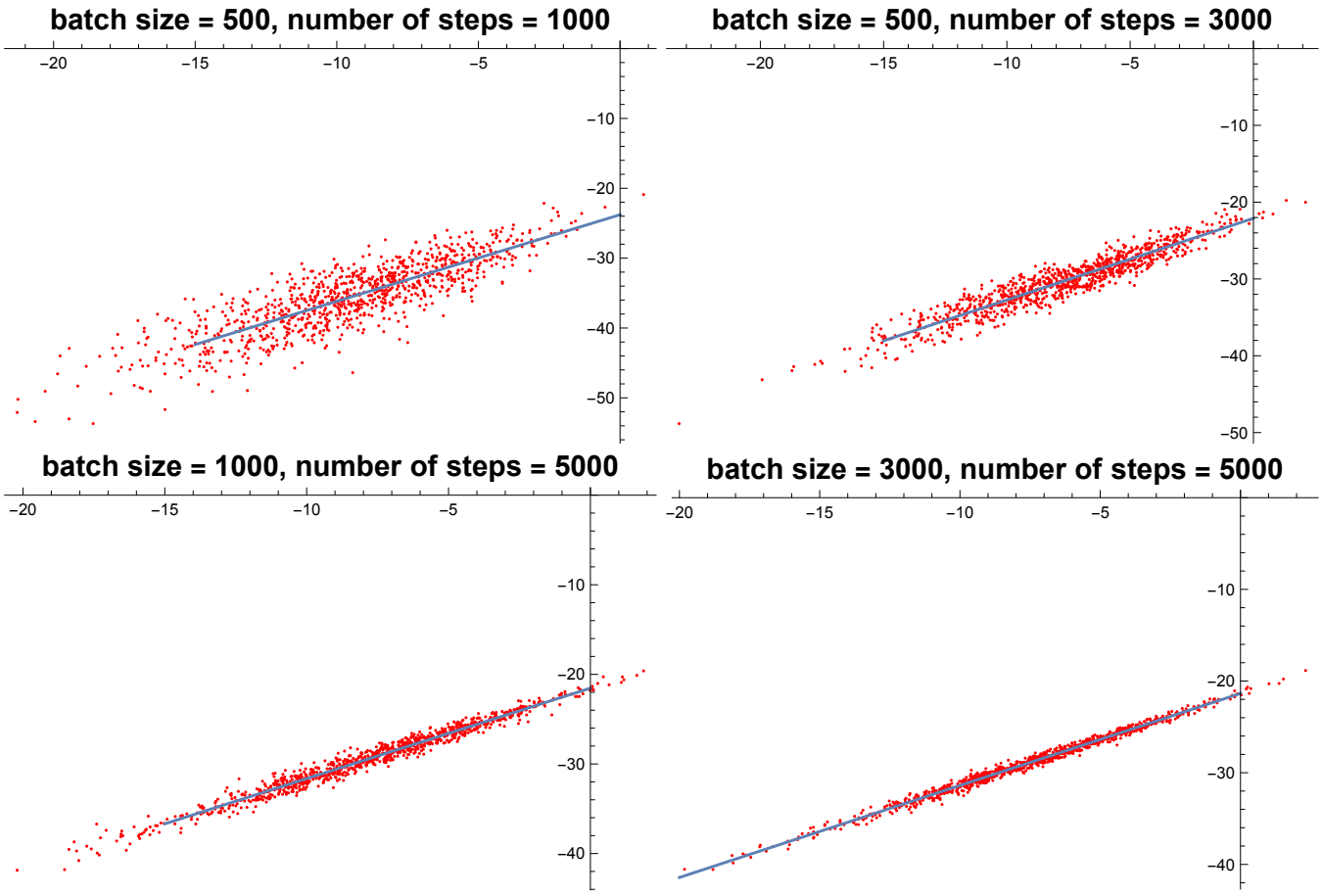

Figure A.1: Scatter plot of logarithms of norms of wavefunctions: $\ln \Psi_{\text{Gaussian}}$ (eq. 4.12) vs. $\ln \Psi_{\text{NN}}$ (variational wavefunction) for $N = \{4, 6, 8, 10\}$. Solid line is the best fit line with slope of approximately 1 for all N and irrelevant y-intercept that varies with N. Relative errors (4.14) and Monte Carlo uncertainties are respectively: $\{0.12, 0.051, 0.013, 0.0048\}$ and $\{0.37, 0.20, 0.12, 0.063\}$ (from top-left to bottom-right).

Appendix B

Entropies of Gaussian states in terms of correlation functions

We use the real time approach to compute the entanglement entropy and Rényi entropy of a compact region V for Gaussian bosonic states. The idea is to express the reduced density matrix ρ_V in terms of the two point correlators restricted to the region V.

The local Hermitian variables ϕ_i and π_j (coordinate and conjugate momentum) obey the canonical commutation relations

$$[\phi_i, \pi_j] = i\delta_{ij}, \quad [\phi_i, \phi_j] = [\pi_i, \pi_j] = 0 \tag{B.1}$$

Let the two ponit correlators inside the region V to be:

$$\begin{aligned} \langle \phi_i \phi_j \rangle &= X_{ij}, \quad \langle \pi_i \pi_j \rangle = P_{ij} \\ \langle \phi_i \pi_j \rangle &= \langle \pi_j \phi_i \rangle^* = \frac{i}{2}\delta_{ij} \end{aligned} \tag{B.2}$$

The reduced density matrix is defined as the unique matrix that satisfies

$$\langle O_V \rangle = \mathrm{tr}\,(\rho_V O_V) \tag{B.3}$$

for any operator O.

A Gaussian state is a state such that all non-zero correlators are obtained from the two point correlators by the prescription

$$\langle \mathcal{O} f_{i_1} f_{i_2} \dots f_{i_{2k}} \rangle = \frac{1}{2^k k!} \sum_\sigma \left\langle \mathcal{O} f_{i_{\sigma(1)}} f_{i_{\sigma(2)}} \right\rangle \dots \left\langle \mathcal{O} f_{i_{\sigma(2k-1)}} f_{i_{\sigma(2k)}} \right\rangle \tag{B.4}$$

Hence, the reduced density matrix must be such that expectation values give the right two point functions and Wick's theorem for the canonical variables.

To get the Wick property for correlators, we propose the following ansatz for the reduced density matrix

$$\rho_V = Ke^{-\mathcal{H}} = Ke^{-\Sigma\epsilon_l a_l^\dagger a_l} \tag{B.5}$$

in terms of independent creation and annihilation operators

$$\left[a_i, a_j^\dagger\right] = \delta_{ij} \tag{B.6}$$

which are linear combinations of ϕ_i and π_j ,

$$\begin{aligned} \phi_i &= \alpha_{ij}^* a_j^\dagger + \alpha_{ij} a_j \\ \pi_i &= -i\beta_{ij}^* a_j^\dagger + i\beta_{ij} a_j \end{aligned} \tag{B.7}$$

where the normalization constant $K = \Pi_l \left(1 - e^{-\epsilon_l}\right)$, leads automatically to the Wick property. Here $\mathcal{H}$ is called modular Hamiltonian.

Hence, the reduced density matrix is a product of independent density matrices for oscillators with mode a_l and the state on each of these independent modes is a thermal state for a harmonic oscillator. The Hilbert space of this mode has a basis on occupation number $|n>$. The density matrix is diagonal on this basis with eigenvalues $e^{-\epsilon_l n}$. Then the normalization constant is $\left(\sum_{n=0}^{\infty} e^{-\epsilon_l n}\right)^{-1} = 1 - e^{-\epsilon_l}$, Hence, the normalization constant for ρ is $K = \Pi_l \left(1 - e^{-\epsilon_l}\right)$. The entropy is the sum of the entropies of each oscillators. These have density matrices with eigenvalues

$$\begin{aligned} S_l &= -\sum_n \left(1 - e^{-\epsilon_l}\right) e^{-\epsilon_l n} \log\left(\left(1 - e^{-\epsilon_l}\right) e^{-\epsilon_l n}\right) \\ &= \left(-\log\left(1 - e^{-\epsilon_l}\right) + \frac{\epsilon_l e^{-\epsilon_l}}{1 - e^{-\epsilon_l}}\right) \end{aligned} \tag{B.8}$$

Plugging the above to eq. B.3 we may compute the two point correlation functions $\mathrm{tr}\left(\rho\phi_i\phi_j\right) = X_{ij}, \mathrm{tr}\left(\rho\pi_i\pi_j\right) = P_{ij}$

$$\begin{aligned} \alpha(2n+1)\alpha^T &= X \\ \beta(2n+1)\beta^T &= P \end{aligned} \tag{B.9}$$

where n is the diagonal matrix of the expectation value of the occupation number

$$n_l = \left\langle a_l^\dagger a_l \right\rangle = \left(e^{\epsilon_l} - 1\right)^{-1} \tag{B.10}$$

Thus, we have

$$\alpha \frac{1}{4}(2n+1)^2 \alpha^{-1} = XP \tag{B.11}$$

which gives the spectrum ϵ_l of the independent oscillators (spectra of the density matrix) in terms of the spectrum of XP. If ν_l are the (positive) eigenvalues of $C = \sqrt{XP}$ then

$$\frac{1}{4}(2n_l + 1)^2 = \nu_l^2 \implies \nu_l = \frac{1}{2}\left(2\left(e^{\epsilon_l} - 1\right)^{-1} + 1\right) = \frac{1}{2}\coth\left(\epsilon_l/2\right) \quad \text{(B.12)}$$

Then, it's easy to see that entropy of the mode l is:

$$S_l = \left(-\log\left(1 - e^{-\epsilon_l}\right) + \frac{\epsilon_l e^{-\epsilon_l}}{1 - e^{-\epsilon_l}}\right) = (\nu_l + \frac{1}{2})\log(\nu_l + \frac{1}{2}) - (\nu_l - \frac{1}{2})\log(\nu_l - \frac{1}{2}) \quad \text{(B.13)}$$

Summing over all the modes l we get the entropy:

$$S = \mathrm{tr}((C + 1/2)\log(C + 1/2) - (C - 1/2)\log(C - 1/2)) \quad \text{(B.14)}$$

Similarly, for Rényi entropies we have for each mode

$$\mathrm{tr}\,\rho^n = \left(1 - e^{-\epsilon_l}\right)^n \sum_{m=0}^{\infty} e^{-\epsilon_l n m} = \frac{(1 - e^{-\epsilon_l})^n}{(1 - e^{-\epsilon_l n})} \quad \text{(B.15)}$$

$$S_l^n = \frac{1}{n-1}\left(n\log\left(1 - e^{-\epsilon_l}\right) - \log\left(1 - e^{-\epsilon_l n}\right)\right) \quad \text{(B.16)}$$

which, after converting to the variable ν_l and summing over the modes l, leads to:

$$\log\left(\mathrm{tr}\,\rho^n\right) = -\,\mathrm{tr}\left[\log\left((C + 1/2)^n - (C - 1/2)^n\right)\right] \quad \text{(B.17)}$$

$$S^n\left[\rho_V\right] = \frac{1}{n-1}\,\mathrm{tr}\left[\log\left((C + 1/2)^n - (C - 1/2)^n\right)\right] \quad \text{(B.18)}$$

Appendix C

Best polynomial approximation (BPA)

Best polynomial approximation (BPA) [31] starts with writing the von Neumann entropy $S_1 = \mathrm{Tr}[\rho_A \ln \rho_A]$ as a sum $-\sum \lambda_i \ln \lambda_i$ where $\lambda_i \in (0, 1]$ are the eigenvalues of ρ_A. If we find a polynomial approximation for $f(x) = x \ln x$ over the range $(0, 1]$ such that $f(x) \approx \sum_n \alpha_n x^n$, then we can approximate entanglement entropy as:

$$S_1 = -\operatorname{Tr}\left[\rho_A \log \rho_A\right] \approx \sum_{n=1}^{n_c} \alpha_n \operatorname{Tr}\left[\rho_A^n\right] \ , \tag{C.1}$$

where n_c is the cutoff polynomial degree. For 1D systems $n_c \approx 7$ is found to be enough [31] however in higher dimensions the required polynomial degree could be very large as it scales with the square root of rank of ρ_A.