Geometric Quantum Information Theory

DHRUVA KARKADA

Advised by: Fabio Anza and James P. Crutchfield
UC Davis Physics NSF REU, 2020

Abstract

Understanding the physical nature of quantum information is critical to develop ways to exploit quantum systems, either to perform computation or do work. In particular, the phenomena of entanglement and decoherence play a central role in many-body quantum dynamics and quantum computation. In this report, I explain the work I did this summer to help develop a new formalism we call *geometric quantum mechanics*. This work helps demonstrate the power this formalism has to encode system-environment interactions in the system's local description. I also explain work I did to quantify the ways that quantum measurements extract information from a restricted class of quantum processes.

Motivation

In the past few decades, work in statistical mechanics, dynamical systems theory, and information theory have revealed that information is a dynamical quantity that has a fundamental role in physics[1–3]. Many classical and thermodynamic phenomena can be better understood through the lens of information theory; a pertinent example is the advent of quantum information science in recent years. This summer, I explored various ways to extend formalisms from classical information theory into the quantum domain.

Several quantum information-theoretic theorems exist that prove bounds on what can't be done. For example, the no-cloning theorem tells us that physics forbids us from duplicating an unknown quantum state[4]. On the other hand, the no-hiding theorem tells us that quantum information that is "lost" due to decoherence is in fact just dissipated into the larger environment. Therefore, quantum information is neither created nor destroyed – it's a conserved quantity.

However, like energy, it's a dynamic quantity which flows in and out of the system in question.

This naturally leads us to the theory of *open quantum systems*, an area of active research. Central questions involve understanding the ways in which information, entropy, and entanglement (all of which are intimately related quantities) are stored and transferred between coupled quantum subsystems[5–8]. Of course, isolated quantum systems undergo unitary evolution under the Schrödinger equation. But when a system is coupled to an environment, it may evolve under stochastic dynamics*. In this general case, we would like to understand how to probe a quantum system to extract information about its underlying dynamics, whether they be unitary or stochastic.

Although much of the prior work has focused on understanding open systems through the lens of quantum statistical mechanics, we attempt to take a different, complementary approach. By rephrasing this situation in the language of classical probability theory and dynamical systems theory, we can probe the same questions using the existing information-theoretic frameworks available in the classical setting. Unfortunately, it is a nontrivial task to reframe quantum mechanics in this manner. This summer, I explored a formalism we call *geometric quantum mechanics* to help accomplish this task.

Another formalism in classical information theory is known as *computational mechanics*[1,2]. To motivate it, let's ask a somewhat meta-scientific question: how do we learn about the natural world? We may probe dynamical systems with measurement instruments, extracting information from these systems. From time-series measurements of dynamical systems, it is be possible to construct models of the underlying system, from which we extract physics[9,10]. This task lies

---

*namely, the Lindblad equation and the quantum Langevin equations[4].

in the domain of computational mechanics, which itself is rooted in information theory. In computational mechanics, the goal is to construct optimal models of processes from time series data, and then extract physics from these optimal models. But when the process takes on quantum aspects, its information-theoretic properties may be drastically changed. I spent some time this summer exploring this.

In this report, I will introduce basic concepts in information theory, motivate the formalism of geometric quantum mechanics, and explain work I did to help understand the interplay between the two.

Background

Geometric Quantum Mechanics

In undergraduate quantum mechanics, we learn that quantum states live in Hilbert space[11]. In a finite-dimensional Hilbert space, each coordinate is represented by complex amplitude. This space of complex vectors is endowed with an equivalence relation: vectors which are complex multiples of each other correspond to the same physical state. Therefore, the Hilbert space with dimension $n+1$ is in fact the complex projective space $\mathbb{CP}^n$, which is a Kähler manifold carrying the Fubini-Study metric[12,13]. Understanding these terms is not crucial to this report – the takeaway is that the manifold of pure states carries a Riemannian structure (which defines a notion of distance) as well as a symplectic structure (which defines a notion of dynamic).

For our purposes, we parametrize quantum states belonging to $\mathbb{CP}^n$ using $2n$ real-valued coordinates as follows. Note that each complex amplitude Z_α can be broken into a magnitude and phase:

$$Z_\alpha = \sqrt{p_\alpha}e^{i\phi_\alpha}.$$

We can perform this decomposition for the $n+1$ coordinates in the Hilbert vector. In fact, the last coordinate is completely determined by the others (up to the equivalence relation), so in the end we only have n 2-tuples of real coordinates (p_α, ϕ_α).

This coordinate transformation has benefits and drawbacks. The main drawback is that quantum mechanics is no longer linear. When performing traditional quantum mechanical calculations, it is far more cumbersome to use these geometric coordinates. However, this coordinate system emphasizes the symplectic nature of the underlying manifold. More importantly, the volume element under the Fubini-Study metric in this coordinate system is simply

$$dV_{FS} = dp_1 \, d\phi_1 \, dp_2 \, d\phi_2 \, \ldots \, dp_n \, d\phi_n$$

i.e. the Euclidean volume element. Therefore, it's natural to use this coordinate system to describe probability distributions on $\mathbb{CP}^n$.

Why do we want to describe probability distributions on $\mathbb{CP}^n$? Consider an open quantum system S of dimension $n+1$ interacting with an environment E of finite dimension d_E. Then the system's mixed density matrix ρ_S can be decomposed into d_E pure states living in $\mathbb{CP}^n$:

$$\rho_S = \sum_{i=1}^{d_E} a_i |\chi_i\rangle\langle\chi_i| \qquad (1)$$

where the a_i are positive and sum to 1, and the $|\chi_i\rangle$ are pure states in the system's Hilbert space. It has recently been shown[14] that this decomposition $\{a_i, |\chi_i\rangle\}$ can be chosen such that it encodes the global wavefunction Ψ_{SE}:

$$\Psi_{SE} = \sum_{i=1}^{d_E} \sqrt{a_i}|\chi_i\rangle|e_i\rangle$$

This ensemble $\{a_i, |\chi_i\rangle\}$ is what we call the *geometric quantum state*. Since the a_i are positive and sum to 1, we interpret the geometric state as a discrete probability distribution on $\mathbb{CP}^n$. If the system is a qubit, this amounts to a distribution on the Bloch sphere.

Since the ensemble encodes the whole wavefunction, tracking the open dynamics amounts to computing the time evolution of the probability distribution on only the system Hilbert space. This contrasts the typical picture, namely computing the trajectory of a single state in the global Hilbert space.

Computational Mechanics

Computational mechanics concerns itself with finding and understanding optimal models for physical processes[1,2]. Here, "optimal" is defined in a precise sense: the model should be minimal in size (i.e. satisfying Occam's razor) yet maximal in predictive power. These notions are formalized using information theory.

The foundations of information theory are built upon probability theory[15]. The central information measure, the *Shannon entropy* (often simply referred to as *entropy*) of a random variable, is defined as

$$H(X) = \sum_{x \in \mathcal{X}} p(x) \log \left(\frac{1}{p(x)} \right)$$

where $\mathcal{X}$ is the set of all possible outcomes of X and $p(x)$ is the probability of observing outcome x. Note that if X represents the microcanonical ensemble of a system at fixed energy, this information entropy reduces to the familiar thermodynamic entropy.

The random variable X can be interpreted in a variety of ways, a fact that underlies the flexibility and ubiquitous nature of information theory. For our purposes, it's helpful to think of X as a single symbol in a time series of a stochastic process. In this simple view, the complex dynamics of the process is reduced to the single-symbol statistics of the output (measured) data.

Many useful information measures rely on comparing two such streams of data X and Y. For example, it's natural to ask if there exist correlations between the two; this is measured using mutual information:

$$I(X;Y) = \sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} p(x,y) \log \left(\frac{p(x,y)}{p_X(x) p_Y(y)} \right)$$

where p is the joint probability distribution and p_X and p_Y are the marginal distributions. Note that the fraction inside the logarithm roughly measures the dependence between the two variables X and Y. Intuitively, mutual information measures the amount of information revealed about a process by observing another. Mutual information is symmetric in its arguments, so each process reveals the same amount of information about the other.

Another measure which measures dependence is the conditional entropy:

$$H(Y|X) = \sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} p(x,y) \log \left(\frac{p_X(x)}{p(x,y)} \right).$$

Note that the fraction inside the logarithm is the reciprocal of the conditional probability. This mirrors the expression for $H(X)$. The conditional entropy is intimately related to mutual information:

$$H(X) - H(X|Y) = I(X;Y) = H(Y) - H(Y|X)$$

In addition to these information measures, we need a way to represent models of physical systems. In computational mechanics, we typically use hidden Markov models (HMMs) to represent a process that can generate a time series of data. An HMM is a graph whose nodes represent internal states of the model, and whose directed edges represent transitions between internal states. An example of an HMM is shown in Figure 1.

The extension of computational mechanics to quantum systems poses an additional challenge: the measurement instrument itself will introduce nontrivial dynamics to the system[16]. Acquiring information about a quantum system necessarily disrupts the system (wavefunction collapse), and the information acquired is probabilistic (with outcomes determined by the Born rule). Together, these properties imply that it's generically not possible to extract all the quantum information from a system.
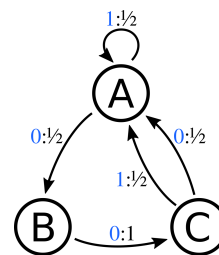

Figure 1: A Hidden Markov Model. Each transition is labelled with a 2-tuple (o, p) where p is the probability of making that transition and o is the emitted output symbol. Of course, for each hidden internal state s, the p's for all outflowing edges must sum to 1 (conservation of probability).

3

Results and Discussion

Visualizing Geometric Quantum States

I began by creating visualizations of geometric quantum states. For qubits (i.e. 2-level quantum systems), our coordinate system consists of a single p and a single ϕ coordinate. Together, they comprise a Cartesian plane, restricted to $p \in [0, 1]$ and $\phi \in [0, 2\pi]$. Traditionally, the manifold of pure states for qubits is represented as the Bloch sphere. The (p, ϕ) plane is constructed by slicing the Bloch sphere along the arc connecting the north and south poles ($|0\rangle$ and $|1\rangle$ respectively) and passing through the $|+\rangle$ state, and unfurling this "sliced peel" into a plane. Note that the unfurling causes the north and south poles to map to the bottom and top edges of the (p, ϕ) plane respectively. This one-to-many mapping reflects the fact that ϕ is undefined on the poles.

Traditionally, qubit mixed states are visualized as points within the Bloch ball. By extending the manifold of states from the sphere to the ball, we can account for mixed states (which extends the family of all pure states). In the geometric formalism, we instead represent mixed states as probability distributions on the manifold of pure states. Thus, we don't change the underlying manifold; we just extend the class of distributions from single Dirac deltas (pure states) to the family of all possible probability distributions. This approach introduces degeneracies: multiple distributions can yield the same density matrix. This is by design; it turns out that the "extraneous" information actually encodes information about the interaction of the system with its environment.

Decades ago, E.T. Jaynes showed that statistical mechanical distributions obey the Principle of Maximum Entropy – they always maximize the information entropy of the distribution subject to certain constraints[17]. If we apply the same principle to qubit mixed states, we can derive the geometric quantum state of highest entropy given a density matrix[14]. This max-entropy distribution is a continuous distribution on the system manifold (the Bloch sphere).

I wrote code to represent arbitrary geometric quantum states (including these maximum-entropy states). I then wrote code to normalize and visual-
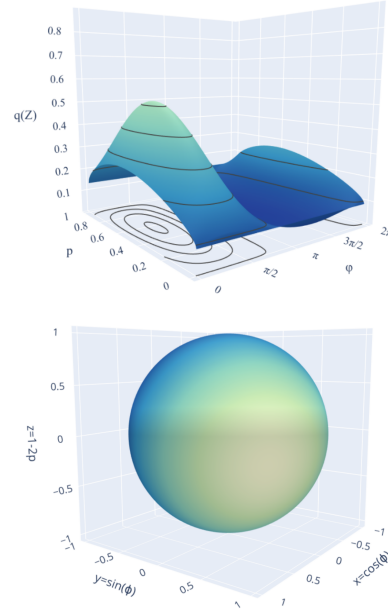

Figure 2: A max-entropy distribution $q(Z)$, constrained to ρ where $\rho_{00} = 1 - \rho_{11} = 0.45$ and $\rho_{01} = \rho_{10}^* = 0.2 - 0.3i$. The distribution is represented both on p, ϕ coordinates (top) and on the Bloch sphere (bottom, color-coded).

ize these distributions on both the Bloch sphere and the (p, ϕ) plane, as shown in Figure 2. I used the `plotly` package to generate interactive 3D plots. I also wrote code to sample pure states from arbitrary geometric quantum states, using the inverse CDF method. Since the inverse CDF method relies on computing marginal distributions, it does not scale well to Hilbert spaces larger than $\mathbb{CP}^1$ (the qubit).

Quantum Dynamics

Sampling from the geometric state allows us to understand qubit isolated dynamics as Hamiltonian flows in the (p, ϕ) plane, where the symplectic structure is evident. Since it can be shown that p and ϕ are canonically conjugated variables, the Schödinger equation in these coordinates takes the form of Hamilton's equations. Initially, I focused on the simple case where the Hamiltonian is time-independent
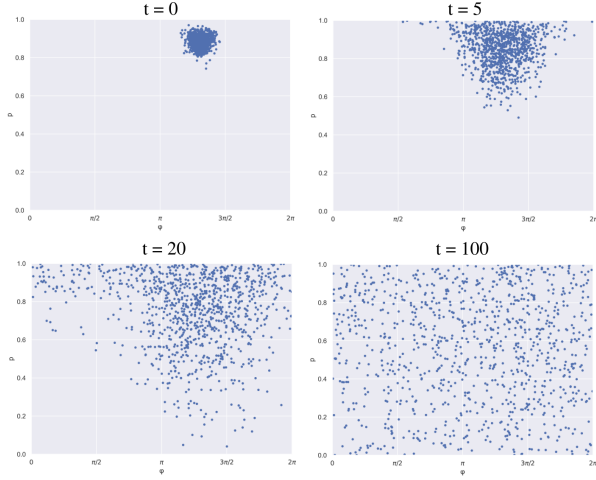
Figure 3: The diffusion of a geometric state undergoing a random walk. The distribution is initially drawn from a maximum-entropy state constrained to a density matrix ρ and an effective temperature T.

and the qubit remains coherent. From introductory quantum mechanics, we know that the qubit will precess around the Hamiltonian's eigenvector (Larmor precession)[11]. By applying this dynamic on each pure state sampled from the geometric state, we can visualize the effect of this dynamic on a mixed state. Unsurprisingly, the geometric state is also periodic in time and the shape of the distribution is time-invariant (i.e. the distribution is just translated around the Bloch sphere).

A slightly more involved toy model simulates a thermalizing process. The idea is to write a simple local stochastic equation of motion, for example a random walk dynamic:

$$|\Psi_{t+1}\rangle = \sqrt{1-\Delta}|\Psi_t\rangle + \sqrt{\Delta}e^{i\gamma_t}|\Psi_t^\perp\rangle$$

where $|\Psi^\perp\rangle$ is the unique orthogonal state to $|\Psi\rangle$, $\Delta \ll 1$ is effectively a step size and $\gamma \in [0, 2\pi]$ is a random variable. If γ is unbiased, i.e. chosen from a uniform distribution, then the overall dynamic on an ensemble of pure states resembles a diffusion process. This is shown in Figure 3. We interpret this to be a toy model of thermalization.

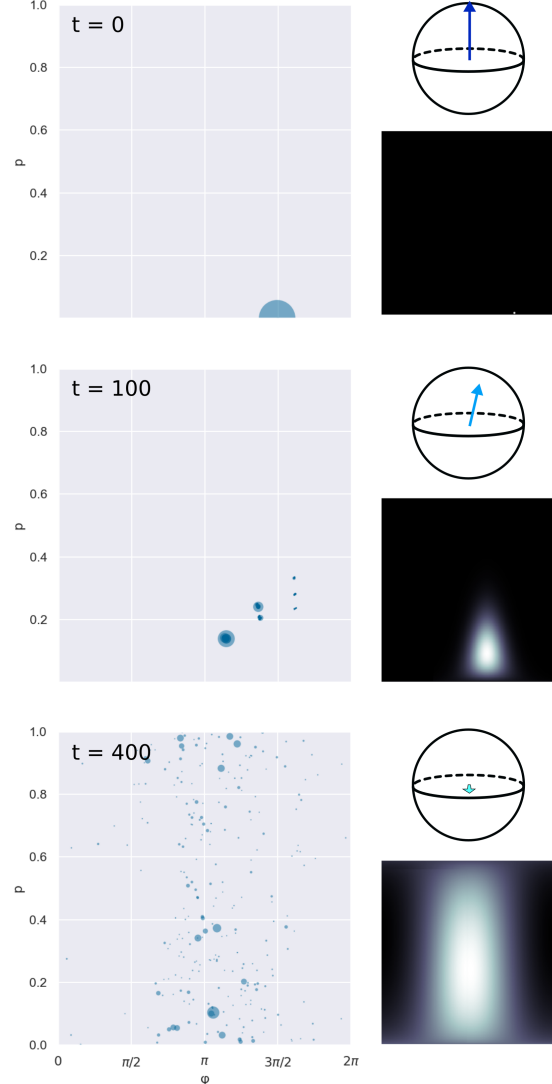To explore time evolution in the non-coherent case, I simulated transverse-field Ising model dynamics on

Figure 4: *Left:* The local geometric state of a qubit interacting with a spin chain under the transverse-field Ising Hamiltonian. Each point is a $|\chi_i\rangle$ from the decomposition in Equation 1, and its radius corresponds to its a_i. *Upper right:* The Bloch vector starts at the north pole $|0\rangle$ and evolves towards the $|+\rangle$ direction while also shortening (i.e. becoming more mixed). *Lower right:* The max-entropy distribution for the geometric state. Note that it's relatively featureless compared to the geometric ensemble. At first, all the weight is concentrated into a single point (i.e. a pure state). As entanglement increases, the geometric state diffuses through the (p, ϕ) plane.

5

a spin chain with periodic boundary conditions. The Hamiltonian for this dynamic is

$$H = -J \left(\sum_i \sigma_i^z \sigma_{i+1}^z + g \sum_i \sigma_i^x \right)$$

where σ_i^z is the operator that locally acts on site i with the Pauli z operator (and likewise for σ_i^x), J is the energy scale, and g is the relative strength of the coupling to the transverse magnetic field compared to the nearest-neighbor interactions. This model exhibits many interesting properties, including nontrivial entanglement dynamics.

While simulating this dynamic, I chose a single qubit and constructed its geometric quantum state as a function of time. As previously mentioned, the geometric state is a decomposition of the local density matrix that encodes the global pure state. I conducted small experiments with 10 qubits; taking the system to be a single qubit leaves 2^9 dimensions in the environmental Hilbert space. This means that the geometric state will consist of 2^9 Dirac deltas. As the system evolves over time, the geometric state also undergoes a time-evolution dynamic, as shown in Figure 4.

In the limit that the environment size goes to infinity, we conjecture that the geometric state approaches a continuous distribution. We also conjecture that probability flows on the reduced system manifold (i.e. the (p, ϕ) plane) for geometric state evolution obey a continuity equation. Further analytical calculations and numerical simulations are required to validate these conjectures.

Nonetheless, these numerical experiments demonstrate that the time evolution of geometric states capture much richer information than the time evolution of the local density matrix. The distribution that exactly encodes the information of the density matrix is the maximum-entropy distribution, which is relatively featureless compared to the geometric state (compare Figure 2 to Figure 4). In other words, the maximum-entropy distribution is a sort of coarse-graining of the detailed geometric state.
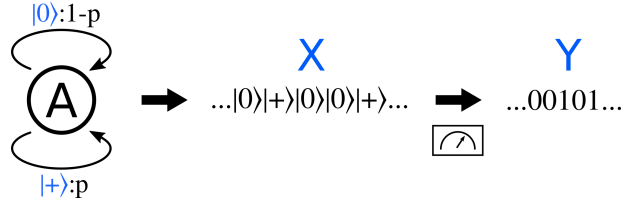

Figure 5: The experiment schematic. An HMM generates quantum states, which are then measured to produce a classical information. Information measures are calculated between X and Y.

Information and Quantum Measurement

I also spent a part of this summer exploring the information-theoretic properties of quantum measurements. To do this, we use a "classically-controlled qubit process" as a toy model of a quantum process. This process is effectively a classical hidden Markov model, with one difference: the emitted symbols are pure qubit states rather than classical bits.

Then, we measure the resulting qubit time-series using a variety of time-independent measurement schemes. This produces a time-series of measurements. Finally, we can numerically compute information measures comparing the quantum time-series with the measurement outcomes. A schematic of this experiment is shown in Figure 5.

The first experiments I conducted involved measuring the time-series using projective measurements. To perform a projective measurement, one must specify a basis; on the Bloch sphere, this is equivalent to picking two antipodal points on the sphere. To simplify the experiment, I restricted the possible measurement bases to the meridian passing through $|+\rangle$ on the Bloch sphere. I chose this meridian because it contains the pure states that are seen in the time-series (namely $|0\rangle$ and $|+\rangle$).

I simulated time series of length 400,000 so that we could minimize statistical noise when computing information measures. Then, I performed measurements on this time series using 50 different measurement schemes. Information measures as a function of measurement basis for the quantum biased coin are shown in Figure 6. Note that the biased coin has no memory – it's a single-state process. Therefore, the
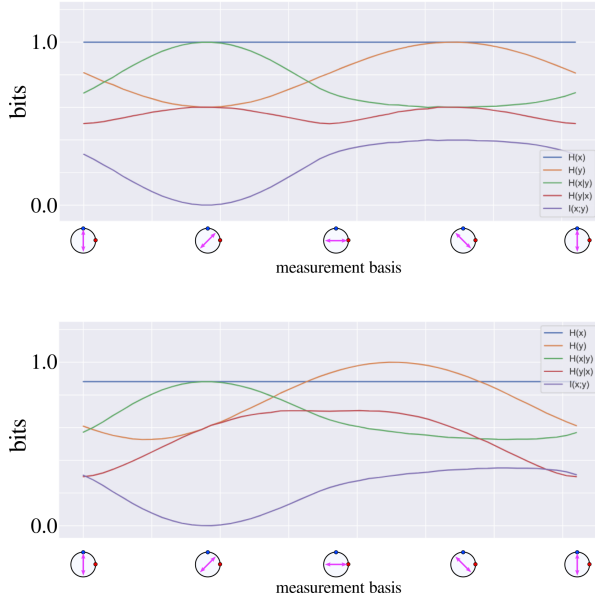
Figure 6: Various information measures for the biased coin with $p = 0.5$ (top) and $p = 0.3$ (bottom), computed as a function of the measurement basis. The basis is graphically depicted as antipodal vectors lying on a cross-section of the Bloch sphere (where the blue dot represents $|0\rangle$ while red represents $|+\rangle$). The blue curve $H(x)$ is the entropy of the quantum states (independent of the measurement angle as expected). The orange curve $H(y)$ is the entropy of the measured series. The green curve $H(x|y)$ is the entropy of the quantum states conditioned on the measurement, and the red curve $H(y|x)$ is vice versa. Finally, the purple curve $I(x; y)$ is the mutual information between the two.

single-symbol information measures distill the effect of measurement on the process. Note that the mutual information between quantum states and measured outcomes is always less than the information entropy of either series ($H(X)$ and $H(Y)$ respectively). This is the result of noncommutativity – in other words, it's impossible for a projective measurement to extract all the information from a process that emits non-orthogonal states. However, there does exist an ideal measurement basis (in the sense that mutual information is maximized), and numerical experiments show that the ideal basis is dependent on the bias of the coin.

Finally, I extended these experiments to support POVMs (positive operator-valued measures). POVMs are a more general class of measurement schemes[4]; they can be thought of as projections onto a set of non-orthogonal vectors (in the way that projective measurements are projections onto two orthogonal vectors). Of course, quantum mechanics does not allow non-orthogonal projections; however, we can still implement POVMs by entangling the qubit with a larger system and then performing projective measurement on the larger system. This procedure, known as Stinespring dilation, has the intended POVM effect on the subsystem[4].

The measurement outcomes of POVMs are mathematically specified by a set of positive semi-definite matrices $\{A_i\}$ which sum to the identity matrix. Then, given a state $|\Psi\rangle$, the probability of measuring outcome i is

$$\Pr[i] = \langle\Psi|A_i|\Psi\rangle.$$

When the $\{A_i\}$ are projection operators, the POVM reduces to the familiar projective measurement.

In this experiment, I restricted the POVMs to the class of POVMs that performs optimally in the task of unambiguous quantum state discrimination[18]. The task is to discriminate two states $|\psi\rangle$ and $|\varphi\rangle$; the optimal POVM supports three measurement results: "definitely $|\psi\rangle$", "definitely $|\varphi\rangle$", and "inconclusive".

To investigate this situation, I simulated many POVMs, where $|\psi\rangle$ and $|\varphi\rangle$ were chosen uniformly from the same meridian described earlier. The resulting plot is shown in Figure 7. The dashed lines correspond to POVMs which are physically equivalent to projective measurements. Numerical analysis indicates that the "peaks" of mutual information occur in this subspace (see blue circles). In other words, projective measurements maximize the mutual information between the measured time series and the quantum time series. This suggests that POVMs may not offer significant advantages over projective measurements. However, mutual information is only one information measure; many other measures exist which probe other aspects, such as memory dependence. Further experiments must be done to investigate whether POVM measurements offer advantages in extracting information from quantum processes.
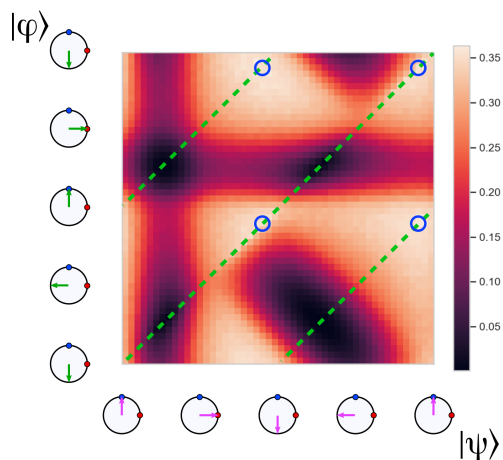
Figure 7: The mutual information between quantum states and measured outcomes for a biased coin ($p = 0.3$) using various POVMs. I use the POVM which unambiguously distinguishes two pure states $|\varphi\rangle$ and $|\psi\rangle$. These states are graphically depicted on the axes as vectors lying on a cross-section of the Bloch sphere (where the blue and red dots represents $|0\rangle$ and $|+\rangle$ respectively). The dashed lines correspond to POVMs which are effectively equivalent to projective measurements. The regions of maximal mutual information are circled in blue.

Conclusions and Outlook

Geometric quantum mechanics may offer new insights into how quantum information is processed, stored, and transferred between a system and its environment. Visualizations I developed and numerical analyses I performed have helped make a compelling case that density matrices alone are insufficient for understanding the dynamics of quantum information in interacting systems. Further work remains to understand the trifurcation of the temporal evolution of the geometric state at early times in the transverse-field Ising model. We also wish to better understand the infinite-environment limit, where we conjecture that the geometric state becomes a continuous distribution.

I have also performed numerical experiments that suggest that for classically-controlled qubit processes, POVMs do not offer an advantage over projective measurements, at least in the case of mutual infor-

mation (a single-symbol information measure). However, for quantum processes with memory, it may be more appropriate to use different information measures, such as those that are conditioned on the measured history. This may be addressed in future work.

Much work remains in the quest to understand quantum information and the role it plays in physics. This effort may result in technologies which can exploit quantum information phenomena. Although many questions remain unanswered in this exciting field, we hope that the geometric formalism may reveal insights previously unseen.

References

[1] J. P. Crutchfield, "Between order and chaos", Nature Physics **8**, 17–24 (2012).

[2] C. R. Shalizi and J. P. Crutchfield, "Computational mechanics: pattern and prediction, structure and simplicity", Journal of statistical physics **104**, 817–879 (2001).

[3] A. S. Holevo and V. Giovannetti, "Quantum channels and their entropic characteristics", Reports on progress in physics **75**, 046001 (2012).

[4] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information: 10th anniversary edition* (Cambridge University Press, 2010).

[5] X.-M. Lu, X. Wang, and C. Sun, "Quantum fisher information flow and non-markovian processes of open systems", Physical Review A **82**, 042103 (2010).

[6] Á. Rivas, S. F. Huelga, and M. B. Plenio, "Entanglement and non-markovianity of quantum evolutions", Physical review letters **105**, 050403 (2010).

[7] F. Caruso, V. Giovannetti, C. Lupo, and S. Mancini, "Quantum channels and memory effects", Reviews of Modern Physics **86**, 1203 (2014).

[8] B. Vacchini, A. Smirne, E.-M. Laine, J. Piilo, and H.-P. Breuer, "Markovianity and non-markovianity in quantum and classical systems", New Journal of Physics **13**, 093004 (2011).

[9] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, "Geometry from a time series", Physical review letters **45**, 712 (1980).

[10] J. P. Crutchfield and B. S. McNamara, "Equations of motion from a data series", Complex systems **1**, 121 (1987).

[11] R. Shankar, *Principles of quantum mechanics* (Springer Science & Business Media, 2012).

[12] I. Bengtsson and K. Życzkowski, *Geometry of quantum states: an introduction to quantum entanglement* (Cambridge university press, 2017).

[13] D. C. Brody and L. P. Hughston, "Geometric quantum mechanics", Journal of geometry and physics **38**, 19–53 (2001).

[14] F. Anza and J. P. Crutchfield, "Beyond density matrices: geometric quantum states", arXiv preprint arXiv:2008.08682 (2020).

[15] T. M. Cover and J. A. Thomas, *Elements of information theory*, Second (Wiley-Interscience, New York, 2006).

[16] A. E. Venegas-Li, A. M. Jurgens, and J. P. Crutchfield, "Measurement-induced randomness and structure in controlled qubit processes", arXiv preprint arXiv:1908.09053 (2019).

[17] E. T. Jaynes, "Information theory and statistical mechanics", Physical review **106**, 620 (1957).

[18] J. Bae and L.-C. Kwek, "Quantum state discrimination and its applications", Journal of Physics A: Mathematical and Theoretical **48**, 083001 (2015).