Foundations Summer School (ETH Zurich)

An Introduction to Epistemic Models of Quantum statistics

Angela Karanjai

karanjai@physics.usyd.edu.au

Contents

Chapter 1

Ontological models

1.1 Introduction

In this series of lectures, we will be concerned with studying different models one can construct to explain Quantum statistics. Before beginning our discussion on the motivation behind studying this topic, let me explain what I mean by a model of Physical data. Any experiment may be broken up into two procedures : *preparations* and *measurements*. One can think of this as a list of instructions about the settings of a physical apparatus, as a black box, with a knob on it that controls its settings. Two distinct knob settings correspond to two distinct preparation procedures. This black box spits out an object, that can be sent to another black box. The setting of the knob on the second black box is a choice of measurement, one wants to perform on the object prepared by the first box. The second box spits out an outcome, which is recorded. This experiment is repeated several times, and the probability of getting a measurement outcome for each preparation procedure is recorded in a data table. A model for this data is a set of rules with which one can reproduce the data table. Such a model could be as simple as "look up the data table" or even "redo all the experiments", however one might attempt to come up with a model that could reproduce the outcome statistics more efficiently.

For the purposes of these lectures, we will assume that the data table is filled with probabilities predicted by quantum theory. Given such a data table, one can use it to come up with a model to explain what might be happening inside the black boxes. One model for such a data table is quantum mechanics, where one can calculate the probabilities using the Born rule. The reason for trying to come up with other models is two-fold. Firstly using experimental data to understand more about a physical system is the very pursuit of Physics and such efforts have lead to our understanding of the universe. The second reason is that one could hope to come up with a model which could be used to predict the outcomes of a subset of experiments without having to look it up in a very big data table. For example, we have found certain sub-theories of Quantum mechanics, that is, data tables made of a restricted set of preparation and measurement procedures, which can be simulated efficiently using classical statistical models that do not require explicitly calculating the Born rule. Thus, we know that, if one was to spend billions of dollars building a device that could perform tasks, which are impossible by any classical computer, then they would have to include preparation and measurement procedures that lie strictly outside of these sub-theories. Alternatively, general constraints on all models attempting to reproduce a particular data table can be used to identify resources for quantum computing or communication. Bell's theorem is an example of such a constraint. Given the data table of a set of preparations and measurements with which one can violate a CHSH inequality, we know that no local model could reproduce the data, thus we know that any device that can produce such outcome statistics has an advantage over devices that are known to be local. The Kochen-Specker theorem makes a similar type of statement about constraints on models of quantum statistics and its link to computational advantage of quantum devices is an area of active research.

All such models, can be studied under the framework of ontological models, without any loss of generality. I will therefore introduce this framework in the next section, which will allow me to formalise the notion of preparation procedures, measurements and data tables. However, I will concentrate on a particular subset of models known as Epistemic models. I will offer two examples of when a quantum sub-theory is reproduced using an epistemic model; *Epistemically restricted Liouville (ERL) mechanics*, which reproduces Gaussian quantum mechanics and *the 8-state model*, which reproduces the quantum stabilizer sub-theory for a single qubit. I shall then introduce Kochen-Specker theorem and other forms of contextuality and briefly talk about their implications on epistemic models and classical simulations of quantum systems.

1.2 Ontological models framework

The ontological models framework provides a general framework to study all models of quantum statistics under. It provides a useful way of studying both hidden variable models and quantum simulation algorithms in a unified framework. I will start by outlining the minimal requirements on models which reproduce quantum statistics which will give us the most general class of ontological models and then proceed to classify different subclasses of such models.

The first ingredient of an ontological model is a sample space Λ, the elements of which represent the state of the physical system in the model. We refer to Λ as the *ontic space* and $\lambda \in \Lambda$ as an *ontic state*. One then defines a probability measure $\mu_P : \Lambda \to [0,1]$ [1] for every preparation procedure P. The only constraint on Λ is that it is measurable with respect to all these measures.The standard interpretation of $\mu_P(\lambda)$ is the probability of the ontic state of the system being λ given the system has been prepared using preparation procedure P. Thus we require that

$$\sum_{\lambda \in \Lambda} \mu_P(\lambda) = 1 \ \forall P \tag{1.1}$$

Since, μ_P is a probability distribution over the ontic states, it represents our knowledge of the state of the system and thus is known as the *epistemic state* of the system.

A function $\xi_{M,k} : \Lambda \to [0,1]$ is then defined for every outcome k of a measurement M. This is the probability of getting outcome k for a measurement M, given that the ontic state of the system is λ. Since, every measurement outputs some outcome we require that

$$\sum_{k} \xi_{M,k}(\lambda) = 1 \ \forall \lambda \in \Lambda \tag{1.2}$$

This is called the *indicator function*. Now, the ontological model can be used to calculate the probability of getting an outcome k given a measurement M is carried out on a system prepared using preparation procedure P.

$$Pr(k|M,P) = \sum_{\lambda \in \Lambda} \xi_{M,k}(\lambda)\mu_P(\lambda) \tag{1.3}$$

Finally each measurement outcome may have an associated disturbance, that is, may change the ontic state of the system. This is captured by a function $f_{M,k} : \Lambda \times \Lambda \to [0,1]$. $f_{M,k}(\lambda', \lambda)$ is the probability of the ontic state of the system being λ' after a measurement M with outcome k given its initial ontic state was λ. Since, we want to able to describe the state of

[1] $\mu_P : \mathcal{P}(\Lambda) \to [0,1]$, where $\mathcal{P}(\Lambda)$ is the power set of Λ. However, in the interests of those unfamiliar with measure theory, I have adopted the simpler albeit less rigorous notation

the system within the same model even after a measurement we require that

$$\sum_{\lambda'} f_{M,k}(\lambda', \lambda) = 1 \tag{1.4}$$

Using this, one can update the epistemic state of the system too. Let P' be the preparation procedure, which is doing measurement M on a systems prepared in P and post-selecting the system with outcome k. Then we can describe the epistemic state of the system as

$$\mu_{P'}(\lambda') = \sum_{\lambda \in \Lambda} f_{M,k}(\lambda', \lambda)\mu_P(\lambda) \tag{1.5}$$

The left hand side of equation (1.3) is the experimental statistics that one is trying to model. There might be several different models for the same experimental statistics and the choice of a preferred model can depend on what one wants to gain from modelling the statistics. For example, one might want a model that has the most realistic ontology or one might just want a model which describes the statistics in the most efficient manner, by which I mean a model that has the smallest state space. The choice of model can depend on whether one is looking to use it for explanatory reasons or as a tool to simulate the statistics. In these notes, I will not assume either motivation but just look at the different type of models that reproduce quantum statistics. In order, for the ontological model to reproduce quantum statistics, we require that the probabilities predicted by the ontological model agree with quantum statistics. Thus we require that

$$\sum_{\lambda \in \Lambda} \xi_{M,k}(\lambda)\mu_P(\lambda) = Tr(M_k^\dagger M_k \rho) \tag{1.6}$$

where M_k is the measurement operator used to describe the k outcome of a measurement M in quantum mechanics and ρ is the quantum state of the system after preparation procedure P. Since one would like to be able to describe the system within the model even after a measurement, we require that the epistemic state of the system is updated to one that corresponds to the updated quantum state. By this I mean that if after the measurement, the quantum state of the system is ρ', then the epistemic state of the system should be $\mu_{P'}$ where P' is a preparation procedure that is described by ρ' within quantum theory.

These are the minimal requirements of any model that reproduces quantum statistics. I will now present examples of such models, the first one being quantum mechanics itself.

1.3 The Beltrametti-Bugaski Model

Quantum mechanics itself can be described within the ontological framework. By this I mean that the state of the system in the model is the quantum state, the indicator function is the Born rule and the state update rule is the quantum measurement update rule. The ontic space in the model is $\Lambda = \mathcal{PH}$, the projective Hilbert space, that is each $|\psi\rangle \in \mathcal{H}$ is represented by $|\psi\rangle\langle\psi| \in \mathcal{PH}$ in the model. The epistemic state of any preparation procedure which prepares $|\psi\rangle$ is

$$\mu_{P_\psi}(|\lambda\rangle\langle\lambda|) := \delta(|\lambda\rangle\langle\lambda| - |\psi\rangle\langle\psi|) \tag{1.7}$$

The indicator function of a measurement M with outcome k is

$$\xi_{M,k}(|\lambda\rangle\langle\lambda|) := Tr(M_k^\dagger M_k|\psi\rangle\langle\psi|) \tag{1.8}$$

Lastly we have

$$f_{M,k} := \delta(\frac{M_k|\lambda\rangle\langle\lambda|M_k^\dagger}{Tr(M_k^\dagger M_k|\lambda\rangle\langle\lambda|)} - |\lambda'\rangle\langle\lambda'|) \tag{1.9}$$

One can see that this model trivially satisfies the conditions for reproducing quantum statistics. This is presented as an example to illustrate the generality of the ontological framework to include the interpretation of quantum mechanics where the quantum state is considered to be the state of reality.

1.4 Types of Ontological Models

This section will briefly outline the different types of ontological models of Quantum statistics. The aim of this section will be to acquaint the reader with terminology used when talking about these models. The debate of whether the quantum state described the *real* physical state of a system or merely our knowledge of its physical state is one that has been central to the study of foundations. One can begin to explore this within the the ontological models framework, by first asking whether, the model with which we chose to explain the statistics we see is a ψ-*ontic* model or a ψ-*epistemic* model.

We know that it is impossible to experimentally completely distinguish two systems prepared using two distinct preparation procedures P_ψ and P_ϕ, preparations that prepare pure states $|\psi\rangle$ and $|\phi\rangle$ respectively, if $\langle\psi|\phi\rangle \neq 0$. However, if we have access to multiple pairs of such systems, where each pair consists of a system prepared using each preparation procedure, then we can statistically distinguish between them. In other words, there is no single measurement that can distinguish two preparations procedures, which correspond to two non-orthogonal pure states, however the two preparation procedures can be distinguished by their measurement statistics. One intuitive explanation for this indistinguishability is that perhaps P_ψ and P_ϕ represent two statistical distributions over the ontic space, which have some overlap, that is, there exists some ontic state, which the system can end up in after either preparation procedure, however the epistemic states corresponding to the two preparation procedures are distinct. Such models are referred to as ψ-epistemic models. A ψ-ontic model is one where all pure quantum states have disjoint supports on the ontic space and thus the ontology does not offer a similar intuitive explanation for the indistinguishability between the quantum states.

Definition 1.4.1 *An ontological model is ψ-Ontic iff for every pair of distinct pure quantum states $|\psi\rangle$ and $|\phi\rangle$*

$$\mathrm{Supp}(\mu_{P_\psi}) \cap \mathrm{Supp}(\mu_{P_\phi}) = \emptyset. \tag{1.10}$$

Definition 1.4.2 *An ontological model is ψ-Epistemic iff there exists a pair of distinct pure quantum states $|\psi\rangle$ and $|\phi\rangle$, for which*

$$\mathrm{Supp}(\mu_{P_\psi}) \cap \mathrm{Supp}(\mu_{P_\phi}) \neq \emptyset. \tag{1.11}$$

The next question, one might wish to ask is about the statistical nature of measurements on pure states. By this, I mean that one cannot predict the outcome of a measurement of an observable O on a system which has not been prepared using a preparation consistent with an eigenstate of the Hermitian operator used to describe O. One can only predict the probability of the outcomes of such a measurement. This measurement indeterminism can be attributed entirely to the lack of knowledge of the ontic state of the system. Such is the case for outcome-deterministic models, where having access to the ontic state of the system would allow one to predict the outcome to a measurement of any observable. In outcome-indeterministic models, the statistical nature of measurements is intrinsic and one cannot assign values to all observables despite full knowledge of the ontic state of the system.

Definition 1.4.3 *An ontological model is outcome-deterministic iff for every projective measurement M_k*

$$\xi_{M,k} : \Lambda \to \{0, 1\}. \tag{1.12}$$

Since,projective measurements in quantum mechanics correspond to sharp measurements of observables (measurements with a certain outcome), one can assign values to all observables given the ontic state of the system in an outcome-deterministic model.

This concludes our brief classification of ontological models. This is by no means an exhaustive classification, however it will be sufficient for our discussion of ontological models in these lectures.

Chapter 2

Epistemic Models

Epistemic models offer a very intuitive explanation to many features of quantum mechanics such as the measurement collapse, the statistical nature of quantum mechanics and indistinguishability between non-orthogonal states. In a model, where the quantum state describes the observer's state of knowledge, one can see the measurement collapse as a Bayesian update rule that we are used to from classical probabilistic models. This does away with the measurement problem. However, there are enough no-go theorems to suggest that an epistemic model of all of Quantum mechanics is unlikely. Despite this, it forms an important tool in separating classical sub-theories from quantum ones and provides a way of simulating certain quantum systems.

2.1 Wigner function

The Wigner representation offers a real phase-space representation of quantum mechanics. Quantum states and measurements are represented by normalised, real valued functions over phase-space. The only obstruction to interpreting these real valued functions as probability distributions is that, they can take negative values. This is why we refer to the Wigner representation as a quasi-probability distribution. However, there exist a set of preparations corresponding to pure quantum states which have a non-negative Wigner representation . These states are called Gaussian states (as their phase-space distribution is a Gaussian distribution). If we restrict to the set of measurements, which map all Gaussian states to Gaussian states, we find that the Wigner representation of these measurements is also non-negative . This restricted set of preparations and measurements is known as the *Gaussian quantum sub-theory.* Since the Wigner representation of this sub-theory is real and non-negative, it can be used to construct an ontological model, which looks very similar to a classical statistical model, which we shall get to in the next section. Now, lets start my formally defining the Wigner representation.

A quantum state ρ describing an n-partite system is represented in the Wigner representation as

$$W_\rho(\mathbf{q},\mathbf{p}) = Tr(\rho A_{(\mathbf{q},\mathbf{p})}) \tag{2.1}$$

where

$$A_{(\mathbf{q},\mathbf{p})} = \bigotimes_{i=1}^{n} A_{(q_i,p_i)} \tag{2.2}$$

and

$$A_{(q_i,p_i)} = \frac{1}{2\pi\hbar}\int dy\ e^{-ip_i y/\hbar}|q_i - \frac{1}{2}y\rangle\langle q_i + \frac{1}{2}y|. \tag{2.3}$$

A quantum measurement M_k is represented as

$$W_{M_k}(\mathbf{q},\mathbf{p}) = Tr(M_k^\dagger M_k A_{(\mathbf{q},\mathbf{p})}). \tag{2.4}$$

Both these functions are normalised, that is,

$$\int d\mathbf{p}d\mathbf{q}\ W_\rho(\mathbf{q},\mathbf{p}) = \int d\mathbf{p}d\mathbf{q}\ W_{M_k}(\mathbf{q},\mathbf{p}) = 1. \tag{2.5}$$

The probability of getting outcome M_k given preparation ρ is given my

$$\int d\mathbf{p}d\mathbf{q}\ W_\rho(\mathbf{q},\mathbf{p})W_{M_k}(\mathbf{q},\mathbf{p}) = Tr(M_k^\dagger M_k \rho). \tag{2.6}$$

One can see that W_ρ and W_{M_k} almost satisfy the conditions for the epistemic state and the indicator function over a canonical phase-space. The only condition they do not satisfy is one of positivity.

2.2 Epistemically restricted Liouville (ERL) Mechanics

This is the first example of an ψ-epistemic and outcome deterministic model that reproduces the outcome statistics of Gaussian quantum mechanics. It is an example of a sub-theory of quantum mechanics that can be explained using classical probability distributions over states characterised by their position and momentum and classical Hamiltonian dynamics. For details and a proof of equivalence between ERL mechanics and Gaussian quantum mechanics see [1].

[1] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, Phys. Rev. A **86**, 012103 (2012)

Lets start by defining the Gaussian sub-theory. It will be convenient to represent a point in phase-space with a vector $\mathbf{z}$, where $z_i = q_i$ and $z_{2i} = p_i$, where q_i and p_i are the i^{th} position and momentum coordinates respectively (this can be interpreted as the position and momentum of the i^{th} particle or the position and momentum of a single particle in the i^{th} dimension) and then define the operators $\hat{z_i} = \hat{q_i}$ and $\hat{z_{2i}} = \hat{p_i}$. The only preparation procedures allowed with the Gaussian sub-theory are the ones that correspond to quantum states ρ for which

$$W_\rho(\mathbf{z}) \propto e^{-\frac{1}{2}(\mathbf{z}-\mathbf{d}_\rho)^T \gamma_\rho^{-1}(\mathbf{z}-\mathbf{d}_\rho)} \tag{2.7}$$

where $d_{\rho i} = \langle \hat{z_i} \rangle_\rho$ is the vector of means of the i^{th} coordinate and

$$\gamma_{\rho ij} = \langle \hat{z_i}\hat{z_j} + \hat{z_j}\hat{z_i} \rangle_\rho - 2\langle \hat{z_i} \rangle_\rho \langle \hat{z_j} \rangle_\rho \tag{2.8}$$

is the covariance matrix.

The only measurements allowed as the ones that take a system in a Gaussian state to another Gaussian state. These set of measurements also have Gaussian phase-space distribution, that is, for a POVM element $E_k = M_k^\dagger M_k$

$$W_{E_k}(\mathbf{z}) \propto e^{-\frac{1}{2}(\mathbf{z}-\mathbf{d}_{E_k})^T \gamma_{E_k}^{-1}(\mathbf{z}-\mathbf{d}_{E_k})} \tag{2.9}$$

For this restricted sets of preparations and measurements, the Wigner function is non-negative. Any projective measurement in this set must be a measurement of a quadrature of position and momentum. Here, I will only consider such measurements as any POVM can be implemented by a probabilistic implementation of projective measurements.Let us now use this to construct an ontological model.

A preparation that is described by ρ in quantum mechanics is represented as a probability distribution over position-momentum space where the probability of the system having position within $d\mathbf{q}$ of $\mathbf{q}$ and $d\mathbf{p}$ of $\mathbf{p}$ is $W_\rho(\mathbf{z})$. Thus the epistemic state is

$$\mu_\rho(\mathbf{z}) = W_\rho(\mathbf{z}) \tag{2.10}$$

It can be show that the classical covariance matrix defined as

$$\gamma_{ij} = 2\langle z_i z_j \rangle - 2\langle z_i \rangle \langle z_j \rangle \tag{2.11}$$

of μ_ρ is equal to γ_ρ.

Thus if ρ satisfies the generalised quantum uncertainty principle

$$\gamma_\rho + i\hbar\Omega \geq 0 \tag{2.12}$$

where

$$\Omega_{ij} = \delta_{i,j+1} - \delta_{i+1,j} \tag{2.13}$$

then μ_ρ will satisfy

$$\gamma_\mu + i\hbar\Omega \geq 0. \tag{2.14}$$

Thus we know that for any allowed preparation we have a restriction on our knowledge of the state. This restriction of our knowledge of the state of the system given by (2.14) is called the *Epistemic restriction*. What is rather interesting is that for the set of allowed measurements, this epistemic restriction is preserved even if we consider classical Hamiltonian dynamics. Which means that even if we assume classical dynamics, gaining information about one observable will lead the loss of information about another observable. However this kind of measurement disturbance is not surprising if one considers the dynamics of a measurement procedure.

Consider a measurement on a single particle in one-dimension prepared in a Gaussian state. The particle has a definite position q_i and momentum p_i which completely characterises its physical state, exactly like in classical mechanics. Let the observable being measured be an arbitrary quadrature of position and momentum of the particle,

$$O_\theta = \cos\theta q + \sin\theta p. \tag{2.15}$$

In order to measure the particle, it must be coupled to a measurement device. The standard way of doing this is coupling the particle to the pointer of the measurement device such that the change in position of pointer gives us the outcome of our measurement. This set up is simply described classically by the Hamiltonian

$$H = \chi(t) O_\theta P \tag{2.16}$$

where P is the observable representing the momentum of the device pointer and

$$\int_0^{t_0} \chi(t) dt = g \tag{2.17}$$

where the measurement is performed in the time interval $[0, t_0]$. The measurement dynamics requires that we also consider the state of the measurement device. Let the device pointer also be modelled as a particle with position Q_i and momentum P_i.

The position and momentum of the particle and the device pointer change according to the classical equations of motion following

$$\frac{dq}{dt} = \frac{\partial H}{\partial p} \tag{2.18}$$

and

$$\frac{dp}{dt} = -\frac{\partial H}{\partial q} \tag{2.19}$$

Thus after the interaction with the measurement device, the position and momentum of the particle is

$$q = q_i + g \sin\theta P_i \tag{2.20}$$

$$p = q_i - g \cos\theta P_i \tag{2.21}$$

and the position and momentum of the device pointer has changed to

$$Q = Q_i + g(\cos\theta q + \sin\theta p) \tag{2.22}$$

$$P = P_i \tag{2.23}$$

Thus the position of the pointer changes proportional to the observable O_θ which gives you the measurement outcome. However, we can see from equations (2.20) and (2.21) that the position and momentum of the particle changes too. This is the disturbance to the particle associated with the measurement. This is completely classical measurement disturbance as we have only considered the classical equations of motion of the particle and the device pointer. Note that despite the change in the particle's position and momentum, the value of the observable that is being measured remains unchanged:

$$\begin{aligned} O_\theta(q,p) &= \cos\theta(q_i + g\sin\theta P_i) + \sin\theta(p_i - g\cos\theta P_i) \\ &= \cos\theta q_i + \sin\theta p_i \\ &= O_\theta(q_i, p_i) \end{aligned} \tag{2.24}$$

We will discuss more about how one can determine which observable's value might be disturbed as a result of the measurement but first let us look at what effect these dynamics have on the probability distributions of the particle and the device that reflect our knowledge of their state. Since the particle was prepared using an Gaussian preparation procedure, its epistemic state (and quantum state) can be completely characterised by its covariance matrix and its mean position and momentum. In one-dimensional phase-space, the covariance matrix of a particle is

$$\gamma_\mu = \begin{bmatrix} 2\Delta q^2 & 2\langle qp\rangle - \langle q\rangle\langle p\rangle \\ 2\langle qp\rangle - \langle q\rangle\langle p\rangle & 2\Delta p^2 \end{bmatrix}$$

(2.25)

where Δq^2 and Δp^2 are the variance in the position and momentum of the particle's distribution and the off diagonal terms are the system's covariance which is essentially a measure of the correlation between the position and momentum of the system. For simplicity, lets assume the particle is prepared in a state where the off diagonals of the covariance matrix are zero. Thus the distribution corresponding to the particle prior to the measurement is

$$\mu(q,p) = \frac{1}{\pi} e^{-\frac{(q-\langle q\rangle)^2}{2\Delta q^2} - \frac{(p-\langle p\rangle)^2}{2\Delta p^2}} \tag{2.26}$$

In order to stay within the sub-theory, we must also model our device pointer in the same manner. Since the change in position of the device pointer will give us our measurement outcome, it is convenient to have the mean position and momentum of our device at zero. Thus the epistemic state of the device pointer before the measurement is

$$\nu(Q,P) = \frac{1}{\pi} e^{-\frac{Q^2}{2\Delta Q^2} - \frac{P^2}{2\Delta P^2}} \tag{2.27}$$

where ΔQ^2 and ΔP^2 are the variance in the position and momentum of the device pointer. In order to determine the outcome of the measurement perfectly, we require that we know the initial position of the device pointer with complete certainty. This means requiring $\Delta Q \to 0$. However, since we only allow states that satisfy the uncertainty principle in equation (2.14), we have $\Delta P \to \infty$

After the measurement, due to the change in the position and momentum of the particle, the epistemic state of the particle is shifted by the momentum of the device pointer to

$$\mu'(q,p) = \int dQdP\mu(q + g\sin\theta P, p - g\cos\theta P)\nu(Q - g(\cos\theta q + \sin\theta p), P) \tag{2.28}$$

Thus one can see that the more certainty that one tries to gain from the measurement, the more disturbed the epistemic state of the particle becomes as in gaining precision in our knowledge of the measurement outcome, we lose precision in our knowledge of how much the position and momentum of the particle have been shifted by. On the other hand if one wants a small disturbance to the state of the system, one would require that the momentum of the device pointer is distributed tightly about zero, however this would mean a large uncertainty in the position of the device pointer and thus the measurement outcome would have a larger uncertainty. This information-disturbance trade-off arises despite the dynamics of the systems being completely classical purely from the epistemic restriction imposed on the initial states of the system. In fact any Hamiltonian which is at most quadratic in the canonical coordinates correspond to unitary transformations that act linearly on the canonical coordinates. These are known as *linear symplectic* transformations. They are the only transformations that map Gaussian states to Gaussian states and they preserve the covariance matrix and thus the total uncertainty in the system.

[**Exercise 1**:Write down the covariance matrix of the combined system of the particle and device pointer in the above example before the measurement interaction. Show the covariance matrix of the combined system stays the same after the measurement by explicitly writing down the epistemic state of the combined system after the measurement. Write down the mean position and momentum of the particle and device pointer in terms of their initial mean positions and momenta.]

The measurement disturbance puts a constraint on the sets of observables that can be jointly measured, that is the measurement of one does not change the value of the other. In quantum mechanics, two observables can be jointly measured if they commute relative to their matrix commutator. In this classical theory, two observables can be jointly measured if they commute according to the Poisson bracket:

$$\{O, A\} := \sum_i \left(\frac{\partial O}{\partial q_i}\frac{\partial A}{\partial p_i} - \frac{\partial O}{\partial p_i}\frac{\partial A}{\partial q_i}\right) \tag{2.29}$$

[**Exercise 2**: Show that the obervable O_θ commutes with the measurement Hamiltonian in the above example. This is why its value remains unchanged despite the measurement disturbance.]

Thus we can see that probability distributions over canonical position and momentum evolving under classical equations of motion completely reproduces the operational statistics of the Gaussian quantum sub-theory. This tell us that if one was to try to build a quantum computer only using preparations and measurements from within this sub-theory, such a device would have no advantage over a classical device. Several phenomenon considered "quantum" can also be reproduced within ERL mechanics such as the EPR thought experiment, the no-cloning theorem and teleportation. This is described in detail in [2]

2.3 The 8-state model of Qubits

The most widely used set of states for quantum computation are the qubit stabilizer states. The single qubit stabilizer sub-theory consists single-qubit Pauli measurements, Clifford gates and states that can be prepared using a circuit composed of these operations and classical post-processing. This corresponds to 6 pure stabilizer states and measurement projectors

$$\{|0\rangle\langle 0|, |1\rangle\langle 1|, |+\rangle\langle +|, |-\rangle\langle -|, |i\rangle\langle i|, |-i\rangle\langle -1|\} \tag{2.30}$$

and the single qubit *Clifford* unitaries which are generated by the Hadamard and the Phase gates. The 8-state model is a simple classical probabilistic model that reproduces all the outcome statistics of this sub-theory.

The sub-theory has 3 observables that are statistically independent, that is,

$$Pr(o_i|o_j) = Pr(o_i|o_j, o_k) = Pr(o_i) \quad i, j, k \in \{1, 2, 3\} \tag{2.31}$$

where o_i is the outcome of the i^{th} Pauli observable. Since all the observable have a binary outcome, this can be modelled classically as 3 independent axes, each with two points. This gives us 8 possible value assignments. One can think of this as the vertices of a cube where the vertex coordinates (x, y, z) represent the values of the three observables. Thus the ontic states are represented by 3 dimensional vectors are

$$\Lambda = \{(x, y, z)|x, y, z \in \{+1, -1\}\} \tag{2.32}$$

An ± 1 eigenstate of the observable corresponding to the i^{th} coordinate can then be represented by a uniform distribution over the 4 vertices on the surface of the cube that correspond to the i^{th} coordinate being ± 1. Thus the 6 stabilizer states are represented by uniform distributions over the 6 surfaces of a cube. For example, the preparation that corresponds to $|0\rangle$, the $+1$ eigenstate of Z is represented as a uniform distribution over the 4 vertices $\{(x, y, 1)\}$.

A measurement of an observable simply reveals the value of the coordinate corresponding to that observable. A measurement also has an associated

[2] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, Phys. Rev. A **86**, 012103 (2012).

disturbance which is that it randomizes the value of the other two coordinates. Here is a simple example of how this simple model reproduces the statistics of the single qubit stabilizer sub-theory. Let us consider the preparation where one measures the observable X and selects the $+1$ outcome. This is described by $|+\rangle$ in quantum mechanics. In the classical model, this corresponds to knowing the value of the x coordinate of the system. Thus the epistemic state is a uniform distribution over the 4 vertices that lie on the $x = 1$ surface of the cube, which means that the system could be in any of the 4 ontic states but we have no knowledge of which one it is in. Now if I chose to measure Z, I have equal probability of getting ± 1 as exactly half of the vertices of that plain have the z-coordinate value ± 1. Now imagine the measurement of Z give outcome $+1$, we know that the ontic state is now one with $+1$ z-coordinate, however due to the measurement disturbance, we know longer know the value of its x-coordinate. So again our epistemic state is a uniform distribution over 4 vertices, which now all lie on the $z = +1$ surface of the cube. This is the distribution that represents the $|0\rangle$ state, which is what one expects.

The Hadamard and Phase gates are represented as follows

$$\begin{aligned} H &: \Lambda \rightarrow \Lambda \\ H(x, y, z) &= (z, -y, x) \end{aligned} \tag{2.33}$$

$$\begin{aligned} P &: \Lambda \rightarrow \Lambda \\ P(x, y, z) &= (-y, x, z) \end{aligned} \tag{2.34}$$

and the 24 single-qubit clifford gates are given by the 24 rotational symmetries of a cube.

[**Exercise 3**: Write down the epistemic state for each quantum stabilizer state and show how they are updated after a Hadamard gate. Check that the epistemic state at the end corresponds to the quantum stabilizer state one expects to get after applying the Hadamard unitary.]

This is another example of how lack of knowledge and measurement disturbance can reproduce a quantum sub-theory in a model where physical systems have well defined values of observables at all times.

Chapter 3

Contextuality

In the previous chapter, we saw that there are sub-theories of quantum mechanics that can be explained perfectly well by stochastic models with classical dynamics where all the indistinguishablity of the quantum states was explained by the overlap between probability distributions corresponding to the quantum states. Such models however can only be constructed in the absence of *Contextuality*. In this chapter, I shall introduce two notions of contextuality and briefly outline the difficulty it poses towards constructing epistemic models.

3.1 Kochen-Specker Contextuality

The two examples of epistemic models in these notes both have the nice feature that the quantum state can be described as the state of the observer's knowledge about a system with well defined physical properties. I am sure, one would agree that this is a much more intuitive picture of reality than the one assumed by the orthodox view of quantum mechanics, which is that, a system does not have well defined values of observables until they are measured. So why do we not interpret all of quantum mechanics in the more intuitive way? The reason for this is that in certain cases, there exists no way of assigning values to set of observables simultaneously, such that quantum statistics are reproduced.

If O is an observable, whose outcome can be described using projectors of rank ≥ 2, then it is possible to infer the outcome of O in several distinct ways by measuring commuting sets of observables. The different ways in which a projector of O can be decomposed into sums of projectors of mutually commuting observables give us the different measurements one can do to infer a particular outcome of O. For example, the value of the observable

YY can be inferred in three different ways using the following relationships:

$$YY = Y_1.Y_2 \tag{3.1}$$

$$YY = (XZ).(ZX) \tag{3.2}$$

$$YY = -(XX).(ZZ) \tag{3.3}$$

[**Exercise 4**: Show that the rank 2 projectors corresponding to the outcomes of YY can be decomposed into rank 1 projectors in three different ways corresponding to (3.1)-(3.3)]

Thus using (3.1)-(3.3) the value of YY can be determined using three different types of experiments,the first, where you measure the Y observable on each of the two qubits, and the other two involve doing two non-local measurements on the system of two qubits. The measurements differ in the set of observables commuting with YY, whose values are known in the process, we say that these three distinct measurements correspond to three different *contexts* of measuring the observable YY. One can define a *Context* of O as a set of observables containing O,where each element of the set commutes with every other element in the set. each of these sets corresponds to a set of observables whose outcomes can be determined simultaneously. If two measurement procedures informing us of the outcome of an observable O, differ in the set of observables commuting with O, whose outcomes they reveal, then we say the two measurements measure O in two different contexts.

It is possible to construct sets of observables such that there is no possible way of simultaneously assigning values to the observables without violating the relationship between observables as predicted by quantum mechanics in a particular context. For example, according to quantum theory

$$(ZZZ).(ZXX).(XZX).(XXZ) = -1. \tag{3.4}$$

Now consider the local context of each of the observables in (3.4), that is

$$ZZZ = Z_1.Z_2.Z_3, \tag{3.5}$$

$$ZXX = Z_1.X_2.X_3, \tag{3.6}$$

$$XZX = X_1.Z_2.X_3, \tag{3.7}$$

$$XXZ = X_1.X_2.Z_3. \tag{3.8}$$

Multiplying the equations (3.5)-(3.6), we see that each local X and Z observable appears exactly twice in the product on the right hand side and thus, there is no possible assignment of values to the local observables such that (3.4) is satisfied. Thus if one tries to predict the values of all these observables simultaneously, there is always a measurement that can be performed which will be at odds with the prediction.

Definition 3.1.1 *A non-contextual value assignment for a set of observables* $O = \{O_i | i = 1, ..n\}$ *is a function* $\nu : O \rightarrow \mathcal{R}$ *such that* $\nu(O_i)$ *is an eigenvalue of the hermitian operator describing* O_j *and* $\nu(O_i O_j) = \nu(O_i)\nu(O_j)$ *if* O_i *and* O_j *commute.*

A proof of contextuality is one where it is possible to show that for a set of observables there exists no non-contextual value assignment.

Definition 3.1.2 *A measurement non-contextual model is one for which*

$$\xi_{M,k}(\lambda) = \xi_{\pi_k}(\lambda) \tag{3.9}$$

for any measurement M *whose* k^{th} *outcome is described as the projector* π_k *in quantum mechanics.*

So what is the implication of a proof of contextuality on an ontological model? A little thought reveals that a proof of contextuality rules out non-contextual models which are outcome deterministic. This is simple to see if one considers that in an outcome deterministic model, knowing λ tells you the outcome for all possible measurements and if the model is also non-contextual then all measurements, whose outcomes correspond to the same projectors are represented by the same indicator function, which implies a non-contextual value assignment. So a proof of contextuality rules out any model for which

$$\xi_{M,k}(\lambda) = \xi_{\pi_k}(\lambda), \ \xi_{\pi_k}(\lambda) \in \{0, 1\} \tag{3.10}$$

Note that if one gets rid out outcome determinism, then a proof of the impossibility of a non-contextual value assignment (proof of contextuality) does not rule out non-contextual models. A simple example is quantum mechanics itself (Beltramati-Bugaski). One can see that it satisfies the condition for a measurement non-contextual model trivially.

3.2 Generalised (Spekkens) Contextuality

In [1], a more generalised notion of contextuality was introduced. The notion of contextuality was developed for preparations, transformations and un-sharp measurements. Just like observable outcomes described by projectors with rank ≥ 2 can be decomposed in multiple ways, similarly density matrices with rank ≥ 2 can also be decomposed into lower rank density matrices in multiple ways. These higher rank density matrices correspond to mixed states and their decomposition into rank 1 density matrices correspond to the particular mixture over the pure states represented by the rank 1 density matrices. Let me illustrate this with the following example:

Consider the pure states

$$\rho_a = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$
$$\rho_A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$
$$\rho_b = \frac{1}{4}\begin{bmatrix} 1 & \sqrt{3} \\ \sqrt{3} & 3 \end{bmatrix}$$
$$\rho_B = \frac{1}{4}\begin{bmatrix} 3 & -\sqrt{3} \\ -\sqrt{3} & 1 \end{bmatrix}$$
$$\rho_c = \frac{1}{4}\begin{bmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & 3 \end{bmatrix}$$
$$\rho_C = \frac{1}{4}\begin{bmatrix} 3 & \sqrt{3} \\ \sqrt{3} & 1 \end{bmatrix}$$
(3.11)

The single qubit maximally mixed state is represented as $\rho = \frac{I}{2}$ can be decomposed into sums of these states as follows

[1]R. W. Spekkens, Phys. Rev. A **71**, 052108 (2005)

$$
\begin{aligned}
\frac{I}{2} =& \frac{1}{2}(\rho_a + \rho_A) \\
=& \frac{1}{2}(\rho_b + \rho_B) \\
=& \frac{1}{2}(\rho_c + \rho_C) \\
=& \frac{1}{3}(\rho_a + \rho_b + \rho_c) \\
=& \frac{1}{3}(\rho_A + \rho_B + \rho_C)
\end{aligned} \tag{3.12}
$$

The different decompositions correspond to different ways of preparing the maximally mixed state. For instant the first decomposition implies that it is possible to prepare the maximally mixed by sampling from a uniform distribution of pure states ρ_a and ρ_A but one can also prepare the maximally mixed state by sampling from a uniform distribution of pure state ρ_A ,ρ_B and ρ_C according to the last decomposition. This motivates a definition of preparation non-contextuality.

Definition 3.2.1 *A preparation non-contextual model is one where*

$$\mu_P(\lambda) = \mu_\rho(\lambda) \tag{3.13}$$

for any preparation P that prepares a state described as ρ.

We can use the above example of the maximally mixed state to see how such a preparation non-contextual model is not possible. We start by noting that the supports of orthogonal states must be disjoint and so

$$
\begin{aligned}
\mu_{\rho_a}(\lambda)\mu_{\rho_A}(\lambda) = 0 \\
\mu_{\rho_b}(\lambda)\mu_{\rho_B}(\lambda) = 0 \\
\mu_{\rho_c}(\lambda)\mu_{\rho_C}(\lambda) = 0
\end{aligned} \tag{3.14}
$$

and for a preparation non-contextual model since all preparation procedures that prepare the same ρ must be represented the same we require that

$$
\begin{aligned}
\mu_{\frac{I}{2}} =& \frac{1}{2}(\mu_{\rho_a} + \mu_{\rho_A}) \\
=& \frac{1}{2}(\mu_{\rho_b} + \mu_{\rho_B}) \\
=& \frac{1}{2}(\mu_{\rho_c} + \mu_{\rho_C}) \\
=& \frac{1}{3}(\mu_{\rho_a} + \mu_{\rho_b} + \mu_{\rho_c}) \\
=& \frac{1}{3}(\mu_{\rho_A} + \mu_{\rho_B} + \mu_{\rho_C})
\end{aligned}
\tag{3.15}
$$

and one can see the only solution to equation 3.14 and 3.15 is $\mu_a, \mu_A..\mu_C = 0$. And since this applies to all $\lambda \in \Lambda$, we can see that such a preparation contextual model is not possible.

Similarly, a more generalised notion of measurement non-contextuality is also developed, which takes the previously introduces measurement non-contextuality and generalises it to more general measurements described as POVMs in quantum theory.

Definition 3.2.2 *A measurement non-contextual model is one where*

$$\xi_{M,k}(\lambda) = \xi_{E_k}(\lambda) \tag{3.16}$$

for any measurement M whose k^{th} outcome is described as the POVM element E_k in quantum mechanics.

Thus this is a notion of measurement non-contextuality which is strictly stronger than the measurement non-contextuality introduced in the previous section.