Reformulation of quantum mechanics and strong complementarity from Bayesian inference requirements

William Heartspring

Abstract: This paper provides an epistemic reformulation of quantum mechanics (QM) in terms of inference consistency requirements of objective Bayesianism, which include the principle of maximum entropy under physical constraints. Physical constraints themselves are understood in terms of consistency requirements. The by-product of this approach is that QM must additionally be understood as providing the theory of theories. Strong complementarity - that different observers may "live" in separate Hilbert spaces - follows as a consequence, which resolves the firewall paradox. Other clues pointing to this reformulation are analyzed. The reformulation, with the addition of novel transition probability arithmetic, resolves the measurement problem completely, thereby eliminating subjectivity of measurements from quantum mechanics. An illusion of collapse comes from Bayesian updates by observer's continuous outcome data. Dark matter and dark energy pop up directly as entropic tug-of-war in the reformulation.

Contents

1 Introduction

Hamiltonian formalism and Schrödinger picture of quantum mechanics are assumed throughout the writing, with time $t \in \mathbb{R}$. Whenever the word entropy is mentioned without additional qualification, it refers to von Neumann entropy. $|x\rangle$ becomes an outcome vector in some given basis if it is a basis vector.

Section 2 revolves around the reformulation of quantum mechanics from objective Bayesian principles. A state vector simply encodes state of uncertainty knowledge about other subsystems from an observer point of view. An observer continuously updates a state vector with her own outcome arriving continuously by the Bayes rule - this creating an illusion of state vector collapse. Because outcomes of past need to be used to update today's state vector, Hamiltonian and Schrödinger equation enter. Hamiltonian is selected by the principle of maximum entropy, which is the key principle of objective Bayesianism. Reasons why spacetime is of epistemic nature are discussed in subsection 2.1, and details of a state vector and its update is described in subsection 2.2. How the Born rule and Schrödinger equation may be derived from combination of physical and inference requirements is discussed in the subsection as well.

Under repeated experiments, Bayesian and classical statistics converge toward the same answer - this is why we can use the same theory to make reliable predictions, despite the principle of maximum entropy imposed for choosing Hamiltonian. Furthermore, despite being consistent with calculations of QM, the reformulation requires that underlying reality exists at all time - which one can call as classical reality. Thus, quantum mechanics is re-interpreted under Bayesianism as a statistical inference framework on classical reality.

The principle of maximum entropy drives strong complementarity and the firewall paradox, which is not a paradox at all from the reformulation point of view.

Different consequences of the reformulation are discussed in section 3. "Dark matter" is created as a result of entropy reduction due to an arrival of an outcome, and "dark energy" is created because of increase in entropy due to the principle of maximum entropy.

Section 4 discusses how one recovers spacetime from entanglement data from the principle of maximum entropy and area law. If correct, it completes the project of quantum gravity.

Evidences for the reformulation are discussed in section 5.

Section 6 completely resolves the measurement problem. Section 2 already resolves the basis ambiguity problem simply from Bayesian update perspectives, but how Bayesian probability is consistent with transition structure of reality was left unresolved there. The transition probability arithmetic fills in this hole.

The essential parts of this writing are: paragraph "Summary: maximization problem" of section 2, section 4 and section 6. Together, they demonstrate that the objective Bayesian re-formulation of QM is necessary and consistent.

2 Epistemic nature of quantum mechanics

2.1 Area law, quantum information and spacetime

Area law for von Neumann entropy of some subsystem is given by:

$$S = \alpha A + S_{sub} = -tr\left[\rho \ln \rho\right] \tag{2.1}$$

where ρ is density matrix of a subsystem, tr is trace, S is entropy of a subsystem and k is some constant. S_{sub} is a sub-correction term to the area law. Consider the following pure two-qubit state vector:

$$|\Psi_m\rangle = (|00\rangle + |11\rangle)/\sqrt{2} \tag{2.2}$$

We notice that entropy of either qubit is maximal. Ignoring sub-correction, given state vector $|\Psi_m\rangle$, one should obtain that the area interface between two qubits is maximal from Equation 2.1.

However, note that this is before some outcome is measured for each qubit. When either qubit is measured, the outcome is either $|00\rangle$ or $|11\rangle$. In either case, an area interface collapses to zero.

This is a rapid change. Before measurement, we had maximal area interface, and now we have zero area interface, assuming sub-corrections are very insignificant. If one reads the area law other way around, the rapid change comes because reality ("area") is constructed out of epistemic ignorance ("entropy"). But can reality really be constructed from quantification of ignorance?

Recently, spacetime-from-entanglement mini-revolution[1][2][3][4] has been under the way. But these approaches are all subject to the above criticism. Traditional quantum (gravity) understandings say there is uncertainty (quantified by probability) about which outcome we are in, even if we know the state vector of a system. The Born rule provides an exact rule of uncertainty quantification from the state vector. But it does not say that (gravitational) outcomes themselves are constructed out of ignorance.

I will now present an alternative to traditional views that can make sense of the area law.

Epistemic view of quantum mechanics In this writing, an epistemic view of quantum mechanics is adopted, because that makes discussions easier here. A state vector represents degree of uncertainty an observer has about the entire system.

Outcome of an observer Recall Equation 2.2. There, an outcome of each qubit (observer) is $|0\rangle$ or $|1\rangle$. An outcome is equivalent to reality or a measurement of an observer herself.

States In this writing, the word "states" would refer to those of the entire system or universe. A state vector only represents probability of states of the entire system at some given time t, not across time.

Spacetime and nature of observations The area law suggests that spacetime is indeed of epistemic nature. And it can be made sense in a fundamental way such that one can piggyback on known results from spacetime-from-entanglement literature to recover classical spacetime in appropriate circumstances.

An observer can only notice and picture exterior world by variations of her epistemic state. She "constructs" spacetime such that reality can be seen from this epistemic state - portraying uncertainty into her vision of reality. Notice that while the word "constructs" is used, this does not require active construction by an observer. One can say that for whatever reason, she does that automatically.

This would be met with the following counter-argument: clearly spacetime does not depend on whether we are actively measuring other subsystems or not. But we can define objective epistemic state using outcomes of an observer at every time. This requires assuming that there is definite reality for an observer at every time, or that an observer "measures" herself at every time. I will take and defend the first assumption later, but for now the latter can be assumed instead as well. Even then, these outcomes will be insufficient to determine exact behaviors of exterior subsystems. An observer is effectively condemned to a state vector encoding probabilistic inference, by which she constructs spacetime reality at every time.

Continuous reality of an observer That an observer has reality at every time is refined as continuous reality of an observer. This is required for each measurement not to result in discontinuous entropy change, thereby discontinuously changing spacetime.

“The” philosophy question to be avoided One may ask why an observer experiences (or constructs) spacetime from a state vector. But the question is like asking why we feel conscious when we could have just been experiencing nothing and remain as physical objects. This is a purely philosophical question that cannot really be addressed at the domain of physics.

As far as a state vector is objective, constructed reality is objective.

But why do we only perceive few outcomes of ourselves? But clearly, an observer does not perceive some of her own outcomes. What are circumstances that an observer would perceive outcomes?

The answer relies on the fact that a macroscopic observer actually consists of multiple microscopic observers (or in quantum computation, qubits). These observers together form shared picture of reality. But a final picture, to be presented, would have to be processed. If underlying reality is not “robust,” such as being too volatile, then processing would fail, and we would not notice much.

When is reality guaranteed to be robust? This is answered by quantum decoherence, which we will explore soon. But this provides why despite measurements of an observer herself continuously existing at all time, only few are actually perceived by a macroscopic observer.

Digression: Immanuel Kant The idea that spacetime is not something actually observed but constructed by an observer is nothing new - a central part of Immanuel Kant’s transcendental idealism is about this. From Kant’s view, spacetime is a priori concept used for picturing experiences, not something that objectively exists independently of experience.

While physics work independent of philosophical analysis, we humans desire to understand laws of physics, not just use them and predict results. This is why I brought this philosophical digression, to suggest that there is a wide body of works related to philosophical nature of spacetime that should make one easily understand the epistemic interpretation of spacetime explored in this writing. Furthermore, the connection between interpretations of quantum mechanics by Niels Bohr, one of founders of quantum mechanics, and Kant have been noticed[5], so it is not as if this comes out of nothing.

Black hole evaporation An observer takes away mutual information uncertainty about outcomes of a subsystem, when her own outcome occurs. This reflects back into reduction of area term interface between a black hole (BH) and its non-BH complement. Here, it does not matter whether an observer actually actively noticed her own outcome. As long as reality is constructed by an observer as if an observer did notice her own outcome, objective physics is fine.

Later, it will be seen that the principle of maximum entropy fights against this reduction in entropy. But in case of a black hole, entropy would have been already so maximized at some point that it cannot fight against this loss of entropy due to an observer outcome, because there is no way to increase its entropy. This is a weird but obvious curse: because of entropy maximality, a black hole just has to watch it evaporate completely. (Or more correctly, an observer has to picture a black hole that way, because this is being described from a non-BH observer point of view.)

As said before, all these epistemic understandings depend on possibility of a unique assignment of a state vector. What follows demonstrates that this is indeed the case.

2.2 State vector from objective Bayesianism

Total entropy of the system Total entropy of the system, from now on, will solely refer to sum of von Neumann entropy of subsystems. Thus it does not refer to von Neumann entropy of the entire system, which will anyway be a vacuous measure if the universe is in a pure state. Each observer maintains her own state vector, as a state vector is epistemic.

Principle of maximum entropy Objective Bayesianism suggests that the principle of maximum entropy is to be used to quantify uncertainty an observer faces about the universe. The modern understanding is that the principle arises out of inference consistency requirements[6][7][8].

An important philosophy of objective Bayesianism is that prior probability should not be arbitrarily assigned, and that there is a unique objective way to assign it.

How to update state vector from outcomes? Basically, one wants to preserve the spirit of the principle of maximum entropy but account (or equivalently, update) for new reality unfolding. Effectively, an observer "has" all past and present outcomes of herself, and should use these data when imposing the principle of maximum entropy. That is, a state vector of present time should be determined solely from outcomes of an observer.

But updating requires a state vector at time right before present time as well - after all, we are to reflect all past outcomes of an observer. However, this requires probabilistic modeling of how reality evolves, because we need to think of how past and present states are related.

Furthermore, it is yet unclear how we enforce a state vector to satisfy probability of past observer outcomes being 1, which seems initially reasonable. In fact, I will reject this constraint of past observer outcomes having to be 1 for Bayesian updating.

Nature of (quantum) Hamiltonian We now see two main things an observer must have to experience reality: state vector $|\Psi(t)\rangle$ and Hamiltonian $H(t)$, where $t \in \mathbb{R}$ is time. H is naturally about inference, and thus can change over time - generating strong complementarity, which we will explore soon, and as part of that, black hole complementarity as well.

For now, we will simply assume that Schrödinger equation is the constraint of inference as the law of state vector evolution over time:

$$i\frac{d}{dt}|\Psi(t)\rangle = H(t)|\Psi(t)\rangle \tag{2.3}$$

where H is restricted to self-adjoint operators and $\hbar$ normalized to 1.

I will now argue that there is no reason why a Hamiltonian model must say that past state vectors have these past outcomes as having probability of 1. We do not know whether final $H(\infty)$, which we would take as the universal Hamiltonian that governs evolution for all times including past, would consider the universe as deterministic or probabilistic. In fact, there is even no certainty that Hamiltonian convergence would be there. In this probabilistic view of the universe - considering universe as having some probabilistic distributions, we are condemned to think that even past outcomes are drawn out of these distributions.

Conditional probability and state vector The point was that state vector update is not about enforcing state vector's evolution to satisfy all past and present outcomes having probability of 1. But it makes sense to discuss instead probabilistic distribution of state of the present-time universe, given my own outcome now - conditional probability - because that is what I am experiencing now. And this conditional probability is what is being represented by a state vector of an observer.

Principle of maximum entropy again: Bayesian updating The magic of Bayesian updating is that one can forget about all past outcomes and simply update on prior probabilistic distributions, which are same as preceding posterior probabilistic distributions, with a present outcome. Prior distributions already contain all past data.

And somewhat amazingly, probabilistic distributions of present-time states conditional on a present-time outcome are just indeed posterior probabilistic distributions in Bayesian statistics.

Thus Bayesian updating is consistent with what we are set out to do.

At time $t = -dt$ (dt is infintesimal time), an observer would have state vector $|\Psi(-dt)\rangle$ and Hamiltonian $H(-dt)$ which is the inferred Hamiltonian at $t = -dt$. This Hamiltonian would be used to evolve state vector such that we get some state vector $|\Psi_0'\rangle$. Notice that Ψ has a prime in superscript. This reflects the fact that an outcome at $t = 0$ has not yet been factored in.

Now an observer factors in her own present-time outcome to obtain $|\Psi(0)\rangle$. This is simply an application of the Bayes rule applied to state vector, and what one can call as "collapse" of a state vector. At this point, there is nothing that traditional quantum mechanics has not done before.

But now $H(0)$ is missing, and this is required to form $|\Psi(dt)'\rangle$. And this is where the principle of maximum entropy enters. It asks us to set $H(0)$ to obtain $|\Psi(dt)'\rangle$ by Schrödnger equation such that total entropy of the system is maximized for $|\Psi(dt)'\rangle$ subject to some constraint.

But again, there exists an additional constraint that has to be factored in. To make an analog to macroeconomics, one specified law of state evolution and utility to be maximized, but one has not specified a budget constraint. And what comes next is this budget constraint.

Euclidean unitarity: partition function invariance And that budget constraint is Euclidean unitarity: that partition function $Z = tr[e^{-\beta H(t)}]$ ($\beta = 1/[k_B T(t)]$, with k_B Boltzmann constant, $T(t)$ temperature) has to stay invariant across time. One may ask why.

This Z is a normalization constant in quantum field theory, but varying it makes a theory behave differently. It is why in renormalization, which in the modern Wilsonian view is roughly about tracing out UV degrees of freedom but maintaining essentially same physics for relevant scales, Z is kept invariant through renormalization group flow.

There exists another way of thinking about this: Z reflects sum of probability of states in a canonical ensemble. While usually 1 is thought of as sum of probability, as long as we keep the sum of probability constant, there is no problem with non-unit probability sum. Also because of Wick rotation relation, one can consider β as imaginary time. This thus allows one to re-interpret different values of Z as providing different time scales.

Summary: maximization problem In form of a control theory problem:

- Objective: $\max \sum_i S_i(t)$ where $S_i(t)$ represents von Neumann entropy of subsystem i at time t, with i indexing subsystems.
- State: $|\Psi(t)\rangle$
- Control: Hamiltonian $H(t)$, temperature $T(t)$
- Initial state vector: $|\Psi(t_0)\rangle$
- State evolution: Schrödinger equation, combined with observer outcome trajectory: $d|\Psi(t)\rangle/dt = Sch[Up[|\Psi(t)\rangle]]$, where Up represents Bayesian update function with outcome trajectory, Sch represents Schrödinger equation.
- "Budget" constraint: given constant Z where $Z = tr[e^{-\beta H}]$, $\beta = 1/[k_B T(t)]$.

An initial state vector $|\Psi(t_0)\rangle$ at $t = t_0$ is generated using the principle of zero entanglement between subsystems. Thus Hilbert space of the universe is factorized into Hilbert spaces of individual subsystems. Control vector is set after this initial state vector is determined.

Why Schrödinger equation? It was so far explained how quantum mechanics is essentially equivalent to Bayesian inference, assuming the Born rule and Schrödinger equation. These two are not something yet derived from Bayesian inference requirements.

Fortunately, a derivation of the Born rule from Bayesian inference requirements was already done in Sebens-Carroll (2016) [9], and despite the title, many-worlds intepretations are not required to derive the Born rule in the article. The required principle is the epistemic separability principle (ESP), which will be accepted as a law in this writing.

Now Schrödinger equation. One can easily derive it using assumption of linearity and unitarity. Unitarity comes from the Born rule requirement and sum of probability being 1, so what is left is linearity. Fortunately, such a consideration was already explored using information-theoretic and thus Bayesian inference requirements[10]. In Parwani (2005), linearity of Schrödinger equation comes from Lorentz invariance. But how could this be, when we know that theories that violate Lorentz invariance exist? The extra ingredients, of course, are information-theoretic and Bayesian inference considerations.

And we of course accept Lorentz invariance of underlying reality - thus we completely derived quantum mechanics from Bayesian inference requirements plus Lorentz invariance.

Basis selection issue does not exist So far the question of basis has been ignored, so I will deal with it here.

The basis selection issue is already not an issue in this Bayesian reformulation of quantum mechanics. If an observer outcome requires basis transformation because a basis you originally chose does not represent that outcome as a basis vector, just transform a state vector by basis transformation and then work out Bayesian updating. It is as simple as that.

Again, quantum mechanics, in this reformulation, is equivalent to the objective Bayesian inference framework that takes in physical requirements. But we have no exact knowledge on what information - that includes basis on which outcomes will be projected - would be revealed to us and can only update perceived degree of uncertainty based on actual outcomes.

This assumes that outcome's basis change occurs continuously, which I believe is justified.

Where Bayesians and classicals all agree A state vector of a quantum system is reliable, if we do not have to care about distinctions between classical and Bayesian interpretations of QM. It is well-known that classicals and Bayesians "effectively" all agree asymptotically on probabilistic distributions - thus the question is whether number of outcomes of the same system (as far as epistemic limitations go) can be generated to guarantee such convergence. This is what allows one to use a quantum theory (Hamiltonian) to make correct predictions. There, whether one is a subjective Bayesian, objective Bayesian or classical statistician does not matter.

Future prediction At time 0, an observer has $H(0)$ and a state vector $|\Psi(0)\rangle$. She does not have her own future outcomes, thus she predicts future state vectors as:

$$|\Psi(t)\rangle = e^{-iH(0)t}|\Psi(0)\rangle$$

3 Consequences of the QM reformulation

Statistical inference on classical reality? Surprisingly, despite all past warnings of realism being inconsistent with quantum mechanics (QM), it turned out that QM is consistent with realism.

In fact, if we take observer outcomes of all observers from initial past to final future, then we effectively have obtained the full picture of classical reality. In this view, the world must have objective reality, though gravity remains purely epistemic. (And notice that gravity is not matter.)

There are then two possible views. One is that despite deterministic laws governing our universe, an individual observer cannot access outcomes of other observers and thus is condemned to probabilistic picture of the universe. The other suggests that nature does carry its probabilistic distributions despite classical reality of outcomes drawn from these distributions.

I believe the former is more reasonable, because continuous evolution of outcomes favors it. Why an observer may be condemned in such a situation is well-known for classical thermodynamics - as an example, see Swendsen for more details[11]. Can states and Hamiltonians of QM then can be turned into discovering possible deterministic theories that actually govern reality, minus gravity?

Geometric quantization literature[12] suggests that this view is not surprising, as quantum states themselves can largely be considered classical states.

Differences from QBism Quantum Bayesianism is not a unified interpretation of quantum mechanics - there are potentially many possible interpretations, some involving objective Bayesianism, some subjective Bayesianism. However, it is true that out of these interpretations, those labeled QBism[13] are dominant.

QBism follows subjective Bayesianism, and thus departs significantly from what is suggested in this writing. Objective outcomes of individual observers seem to be denied for QBism. Furthermore, the principle of maximum entropy does not take the center stage at inference.

While main differences are named against QBism, I think they apply against other Quantum Bayesianism interpretations as well. After all, the principle of maximum entropy, as sensible as it is, initially seems hard to reconcile with quantum mechanics of conserved Hamiltonian often discussed. And because of different supposedly anti-realism results of quantum mechanics, it is hard to imagine that continuous reality can be compatible with quantum mechanics.

While these points are already addressed, I will discuss them again.

Measurement problem The basis ambiguity problem of the measurement problem was shown to be a non-existent problem. What about the rest of the measurement problem?

The measurement problem is:

$$\sum_x P(x)P(x \to z) \neq P(z) \tag{3.1}$$

where x (time 0), z (time dt) all are outcomes at different times in some particular basis (at each time, different basis may be chosen) and $P(x \to y)$ represents transition probability from x to y, assuming $P(x \to y) = |\langle y|e^{-iH(dt)}|x\rangle|^2$.

But this damages traditional notion of probability - the inequality in Equation 3.1 should have been equality.

In fact, the problem is on how transition probability is defined mathematically. What one needs to preserve is the Born rule on individual states, not on transition probability. The full transition probability arithmetic that resolve the measurement problem completely will be provided in section 6.

Strong complementarity and the firewall paradox Strong complementarity is an idea that different observers live in different Hilbert spaces. And in fact, an observer herself may change Hilbert space as Hamiltonian H comes to change. This is because Hamiltonian H is of epistemic nature.

Because of strong complementarity and epistemic nature of quantum mechanics, the firewall paradox[14] is automatically dissolved. Observers have two different epistemic uncertainty, so what? There is nothing wrong about this. Having re-formulated quantum mechanics as Bayesian inference, if the firewall paradox really is a paradox, it would amount to saying that Bayesian inference is conceptually wrong. Of course it is the no drama assumption that is at heart of the firewall paradox - but it is dissolved as long as we recover spacetime of general relativity for an infalling observer, and spacetime-from-entanglement literature suggests it is not a problem.

Another question that arises from the firewall paradox is why observers may face different H in case of a black hole, while in other normal circumstances H is largely shared.

Why same H in usual circumstances? Observers largely share same H because they are very close to each other. When X comes to extract all of available mutual information with its complement $\bar{X}$ from an observer point of view, X must be very close to me via locality, because that means I have learned all I can about X now - and X does not provide more information about other systems now. Via consistency of underlying reality, this mandates that they largely share the same view about the world, as they use the same tools of objective Bayesianism.

Now this story seems to be opposite to the usual quantum field theory story[3]: entanglement of subsystems usually decreases with increasing distance. Thus we expect that more entanglement whenever distance decreases.

However, note that this is just the story already discussed about the quantum state vector in Equation 2.2. There, one sees that unaccounted for measurement, entanglement is maximal - thus we expect distance between qubits to be close to zero, which would be the same conclusion one would arrive for the "collapsed" $|00\rangle$ from "area term decreasing resulting in more closeness".

Notion of distance In fact, both stories are compatible. If Equation 2.2 was in a different form as to exhibit non-maximal entanglement, then an outcome of one qubit would not have completely determined an outcome of the other qubit. Thus, as for a pre-update state vector, that entanglement increases as distance decreases still holds.

We can combine these two stories by working with area perturbation δA (assuming linear relationship to mutual information) to derive notion of distance (and metric) at each time, which is already how things are handled[3] - see section 4 for more details.

How locality is related to mutual information will be discussed in depth when talking of "spacetime from entanglement" specifically in section 4.

Non-solipsistic objective physics Objective physics is made possible because a same system (as far as our epistemic limitations go) may be probed many times as to allow thinking of classical probability. And solipsism is avoided because same Hamiltonian (theory) is shared by nearby observers.

Dark energy and dark matter A strange observation then can be noticed. Essentially each arrival of an observer outcome, which happens at every time t, reduces von Neumann entropy of a subsystem. The area law suggests that this shrinks spacetime.

But the principle of maximum entropy provides a reason why total entropy of the entire universe may increase.

This seems to echo so heavily with the concept of dark matter (matching with outcome-induced entropy loss) and dark energy (matching with the principle of maximum entropy). If spacetime is of epistemic nature, so should dark energy. Dark matter is of a different beast, but the general point stands.

The vision that dark energy and dark matter be of entropic characteristics is nothing new.[15] The novelty here is matching dark matter with outcome-induced entropy and dark energy with the principle of maximum entropy.

Hardy's paradox: first part

$$|\Psi_h\rangle = \frac{|00\rangle + |01\rangle + |10\rangle}{\sqrt{3}} \tag{3.2}$$

Now define following basis vectors as well (other than $|0\rangle$, $1\rangle$)

$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

$$|-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

When the outcome of the first qubit is $|0\rangle$, then the other qubit measured in the basis of $|+\rangle$ and $|-\rangle$ can only be $|+\rangle$. If the secnod qubit is $|0\rangle$, then the other qubit measured in the non-numeric basis can only be $|+\rangle$ as well.

$|\Psi_h\rangle$ does not contain any $|11\rangle$ component, and thus one of the qubits has to be $|0\rangle$. This means there is zero probability for $|--\rangle$. But QM predicts from $|\Psi_h\rangle$ that probability of $|--\rangle$ is 1/12, which is Hardy's paradox.

But one can immediately see how the argument went wrong. The "wrong" part is "thus one of the qubits has to be $|0\rangle$". But no, it would only be that way if an outcome indeed happens in the numeric basis. QM bans simultaneous cross-basis measurements, and in the reformulation, this arises because underlying reality is being expressed only in one particular basis.

It is often said that Hardy's paradox eliminates possibility of any reality existing in QM - but this is simply not true, as the reformulation is consistent with both existence of reality and QM.

Hardy's paradox: second part However, one may argue that the reformulation is still troublesome, because it assumes some underlying reality. If reality exists, then a state vector of an observer would have to be updated upon that reality. At the end, only an outcome of probability 1 would remain for a qubit in Hardy's paradox.

This is an invalid argument. Of course since an outcome is measured at the end, we would expect such a result. What that someone really finds troubling is continuous resolution to a complete measurement of an outcome, whereas in traditional understandings of QM it is instantaneous collapse that suddenly results in a measurement. But instantaneous collapse is what is really troubling. Thus there is no problem.

Frauchiger-Renner (FR) thought experiment The FR thought experiment[16] is a good example that demonstrates how the reformulation of QM helps clarify nature of QM.

The problem with the FR experiment lies with "collapse, re-cohere and re-collapse in a different basis" strategy involved. In the reformulation of QM, collapse is just an outcome of an observer arriving continuously, which allows some inference about other subsystems. In this context, one can instead say that an observer notices information about outcomes of a random generator, spin or labs.

In the reformulation, QM is statistical inference on essentially classical states. Thus, when one observes some outcome of a subsystem, it is immutable - one cannot change it by some quantum magic, which is what FR attempt to do. The appearance of collapse and inability to measure simultaneously in different bases come from this classical state immutability.

While the FR experiment is not Hardy's paradox Wigner's friend-ified because of a clever spin trick, its result is indeed Hardy's paradox Wigner's friend-ified. In Hardy's paradox, the issue was that $|--\rangle$ should have non-zero probability but have zero probability if calculated in an unapproved way.

The reformulation of QM suggests that in the FR experiment, some observers do not have information about an outcome of some subsystem, while others do. This is what generates the paradox. But to measure a subsystem, an observer needs to get outcome updates. This forces state vector of different observers about some subsystem to converge, as far as underlying classical reality is consistent.

PBR theorem PBR theorem[17] aims to demonstrate that the set of possible epistemic interpretations of QM is heavily restricted. Does this theorem apply to the approach in this writing? The answer is no.

The problem with PBR theorem is that it assumes quantum state $|0\rangle$ represent something solely epistemic. However, it can be ontic as well. Suppose that:

$$|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$$

From perspective of $|+\rangle$, $|+\rangle$ itself is considered epistemic while outcome $|0\rangle$ is considered ontic. Similarly, $|+\rangle$ may be considered ontic whenever it is a basis vector. And this follows from the objective Bayesian reformulation from quantum mechanics.

Now this can be handled in PBR theorem, by stating that the "overlap region" is zero, but then there is no problem with epistemic interpretations of quantum mechanics.

3.1 Basis, decoherence and causal diamond complementarity

The approach of this writing was not initially inspired by Bousso-Susskind (2012)[18], but it became clear that there is a heavy connection. I believe that some physicists would have screamed at causal diamond complementarity being discussed in terms of an observer trajectory and objective decoherence. Let me review these two ideas.

Observer trajectory First, an observer trajectory. This comes very naturally in the Bayesian reformulation, as it is the data for an update. But in traditional understanding of QM, we think that any subsystem would be in quantum fuzziness, however small it is. There, a causal diamond of an observer is somewhat ill-defined. Of course at roughly measurement limits, this works good enough.

Objective decoherence Second, objective decoherence. In a way, decoherence is simply about looking at behavior of density matrix of individual subsystems, to see if any coherence in each subsystem is lost. That objective density matrices for the entire causal diamond have special meaning seemed very strange, or at least that is how I used to feel.

Now I think Bousso-Susskind was essentially at the right track. The point is that a single preferred basis may be viewed as being shared across the causal diamond, even when we just trace out an environment, defined as those not in the causal diamond. Of course decoherence is incomplete always, so this is more like an approximation. In this basis, the states of the causal diamond may be considered separately, because there is lack of coherence between them. It is as if the universe has branched into multiple universes.

Why causal diamond? In order for exterior world (relative to an observer) to have any possibility of directly affecting an observer, it must be in the causal diamond of an observer, given a present-time observer outcome. It is echoed by the Bayesian reformulation of QM, which speaks to maximize total entropy of the entire system from an observer's point of view.

Back to decoherence So what is the point about different subsystems in the causal diamond sharing the same basis?

First, as said before, a state in a privileged basis allows one to view it as somewhat classical reality.

Second, as observers become nearby, they come to share almost same Hamiltonian H, and furthermore entropy of an observer from another observer's perspective begins to be reduced significantly. This requires that even in epistemic sense, classical reality dominates. This restricts the number of plausible bases in which an observer outcome can arrive.

This is why quantum decoherence seemed to provide important insights toward resolving the basis ambiguity part of the measurement problem, despite never being able to resolve it completely. It required a deeper understanding of quantum mechanics to see how it all played out.

Macroscopic observer While decoherence is analyzed in cosmological scale, one can restrict to a macroscopic observer (or a human) that consists of microscopic observers. There, macroscopic reality emerges when microscopic observers come to form a robust and redundant picture of the universe.[19] If microscopic observers speak of different pictures that cannot be combined by error correction reliably, then a macroscopic observer would ignore them, while continuing to update her microscopic observers. This is why we do not seem to observe quantum reality, despite them existing.

A macroscopic process involves initial divergence of same H and state vector, but convergence is arrived as decoherence is achieved sufficiently, induced by outcome updates.

4 Spacetime from entanglement

This section largely follows Cao-Carroll (2018), though connections to the reformulation are emphasized and more physical intuitions behind the ideas are discussed.

4.1 Locality: area equals mutual information

The intuition behind locality is if I am very close to you I would be able to know you fairly well - mutual information would have been depleted. If I am far away from you, there will be mutual information still left to be updated by my future outcomes. This is supported by our experiences, so should not really be controversial.

Area is the measure of the interface between a subsystem and its complement. Now let me combine these points. Area of surface between subsystem X and its complement $\bar{X}$ from an observer point of view is:

$$A(X;\bar{X}) = \frac{1}{2\alpha} I(X;\bar{X}) \tag{4.1}$$

which one would set $\alpha = 1/4$ in accordance to Bekenstein-Hawking entropy. $A(X;\bar{X})$ refers to area interface between X and $\bar{X}$, $I(X,\bar{X})$ refers to mutual information between X and $\bar{X}$, $c = G = \hbar = 1$ by Planck units. This allows one to define the area perturbation $\delta A(X,\bar{X})$ from state vector perturbation $\delta|\Psi\rangle$ and the principle of maximum entropy $\delta[\sum_i S(i)] = 0$. These allows one to recover emergent metric - recovery details are left to Cao-Carroll (2018)[3]. Emergent spacetime equations are written in perturbation form, because outcome updates are continuous.

4.2 Story of Big Bang cosmology

Because the initial state vector of the universe exhibits no entangement, it naturally leads to the idea that subsystems (or observers, equivalently) had zero distance between them at the start of the universe. Then entropy maximization kicks in heavily because there are very few outcomes for Bayesian updates. This expands spacetime, generating moments of Big Bang. Outcome updates create shrinking "force" for spacetime - dark matter, while the principle of maximum entropy continuously expands on spacetime - dark energy.

5 Evidences toward the reformulation

AdS/CFT as a neural network Recently, a paper[20] that casts AdS/CFT correspondence in terms of a deep Boltzmann machine has appeared - which presents bulk as learning about boundary.

When we already know boundary exactly, then training can be done to ensure that boundary is learned by bulk almost perfectly. The problem, if one takes an epistemic view of quantum mechanics, is that we do not know boundary exactly. In other words, we do not have training data! Thus, what we are to train against becomes unclear. This question itself is not a problem of the cited paper, but it needs to be asked here. And the required training data are provided by outcomes of an observer in the reformulation.

This Boltzmann machine picture gives another justification into why partition function is kept invariant, and how quantum mechanics relates to statistical mechanics.

The law of maximum entropy production In non-equilibrium classical thermodynamics, Swenson and others[21] proposed the law of maximum entropy production (LMEP) as the fourth law of thermodynamics. The intuition, as stated by Rod Swenson, behind LMEP is simple - let me quote directly from the cited article:

> As an example, R. Swenson considered further the change of the temperature in a house in winter. If all doors and windows are closed, the street and room temperatures equalize through heat conduction, i.e. relatively slowly. Opening of a door or a window provides a new opportunity for equalization of the temperatures through the convective transfer. This mechanism will bring the "house-street" system to the thermal equilibrium faster.

Traditional laws of equilibrium thermodynamics do not explain how equilibrium, or steady state, is reached, and this fourth law fills this missing piece.

And basically, this is a classical thermodynamics application of the principle of maximum entropy. While observer outcomes continuously reduce entropy, maximal generation of entropy ensures maximal entropy states out of feasible ones.

5.1 Quantum redundancy

Basis redundancy It was recently argued (Czech et al. (2019)[4]) that spacetime arises from basis redundancy requirements. The point is essentially what was argued in this writing - that physics should work consistently with different sets of subsystem basis. The paper echoes the mirror operator idea of Papadodimas-Raju (2014)[22] as well.

That is, spacetime is sewing together essentially redundant descriptions. But if QM is epistemic, then this would immediately point to spacetime being epistemic as well.

Error-correcting code and decoherence Quantum error-correcting code vision[1] of holographic duality becomes very useful, in relation to quantum decoherence (subsection 3.1). A valuable insight provided is that error correction breaks down for a black hole such that two observers no longer arrive at same bulk spacetime. But it is easy to expect this. In fact, quantum decoherence literature[23] suggests that sufficient decoherence takes time, even if little, for quantum redundancy to form, so error correction is not instantaneous.

But why error correction? The intuition is actually clear. If different observers agree (in probabilistic sense) on the system, it would mean that mutual information they expect from each other has aligned. This immediately implies that some error-correcting algorithm would be able to extract "shared agreement" from state vectors of different observers, which would recover spacetime.

Entanglement equilibrium The key clue toward the principle of maximum entropy idea comes from the entanglement equilibrium literature. In essence, entanglement equilibrium is expressed by:

$$\delta S = \delta S_{UV} + \delta S_{IR} = 0$$

where S is an entropy of some subsystem. Usually, the UV-IR factored equilibrium relation is emphasized instead of $\delta S = 0$, which is rendered useless if without factorization in the literature. The entanglement equilibrium literature[2][3] differs from the approach of this writing in that this is considered solely an equilibrium behavior.

The generalization to the principle of maximum entropy of course seemed impossible. How can multiple theories be describing our reality? Ideally, should not we start from one theory and derive everything, with entanglement equilibrium only providing hints toward the ultimate theory of everything?

It was this picture of the theory of everything this writing intended to shatter by providing the epistemic interpretation of QM.

As stated in Cao-Carroll (2018)[3], the connection to error-correcting codes is there when considering dividing entropy into two parts. Now the main problem with utilizing error-correcting code understanding directly is that it is still AdS/CFT correspondence. Our world is not AdS, and that limits applications. The boundary-less spacetime construction is thus attempted in the aforementioned article, which this writing largely references.

But there are other approaches as well. Recent works[24] try to generalize AdS/CFT correspondence into dS/dS correspondence[25]. This would allow us to generalize what we have learned in AdS/CFT into examples involving our actual spacetime.

6 Transition probability arithmetic: resolving the measurement problem completely

The starting system of probability equations Transition probability arithmetic starts from the system of probability equations, reflecting transition, that must hold:

$$\begin{aligned} P_0(x_1)P_u(x_1 \to y_1) + \cdots + P_0(x_n)P_u(x_n \to y_1) &= P_{t_1}(y_1) \\ &\vdots \\ P_0(x_1)P_u(x_1 \to y_n) + \cdots + P_0(x_n)P_u(x_n \to y_n) &= P_{t_1}(y_n) \end{aligned} \tag{6.1}$$

where $P_t(x_1)$ refers to probability of outcome x_1 at time $t = 0, t_1$, $t_1 \to 0$ but $t_1 > 0$, and $P_u(x_i \to y_j)$ represents transition probability from outcome x_i at $t = 0$ to outcome y_j at $t = t_1$. Subscript u refers to undetermined status of transition probabilities, and we derive P_0 and P_{t_1} from a given state vector by Born rule. A state vector is assumed to be n-dimensional, but the conclusion obtained can be generalized to infinite dimension as well. Outcomes satisfy $\langle y_i|y_j\rangle = \delta_{ij}$.

P_0 and P_1 are probability vectors, and this means that P_u is a stochastic matrix with $\sum_j P_u(x_i \to y_j) = 1$. Thus this consistency requirement does not have to be imposed. All that is to be ensured additionally is non-negativity of transition probability.

Note that the above system of equation satisfies objective completeness of quantum mechanics and unitary evolution.

The requirement of transition probability arithmetic A law of transition probability has to satisfy $P_u(x_i \to y_j) = |\langle y_j|e^{-iHt_1}|x_i\rangle|^2$ if $P_0(x_i) = 1$. This reflects the traditional collapse and measurement intuition.

The ratio rule There are n^2 unknown transition probabilities and n equations in Equation 6.1. Thus $n^2 - n$ equations need to be provided to determine transition probabilities. The ratio rule is specified as:

$$\begin{aligned} P_u(x_1 \to y_1) : P_u(x_2 \to y_1) &= |Am(x_1 \to y_1)|^2 : |Am(x_2 \to y_1)|^2 \\ P_u(x_1 \to y_1) : P_u(x_3 \to y_1) &= |Am(x_1 \to y_1)|^2 : |Am(x_3 \to y_1)|^2 \\ &\vdots \\ P_u(x_1 \to y_1) : P_u(x_n \to y_1) &= |Am(x_1 \to y_1)|^2 : |Am(x_n \to y_1)|^2 \\ P_u(x_1 \to y_2) : P_u(x_2 \to y_2) &= |Am(x_1 \to y_2)|^2 : |Am(x_2 \to y_2)|^2 \\ &\vdots \\ P_u(x_1 \to y_2) : P_u(x_n \to y_2) &= |Am(x_1 \to y_2)|^2 : |Am(x_n \to y_2)|^2 \\ &\vdots \\ P_u(x_1 \to y_n) : P_u(x_2 \to y_n) &= |Am(x_1 \to y_n)|^2 : |Am(x_2 \to y_n)|^2 \\ &\vdots \\ P_u(x_1 \to y_n) : P_u(x_n \to y_n) &= |Am(x_1 \to y_n)|^2 : |Am(x_n \to y_n)|^2 \end{aligned} \tag{6.2}$$

where $Am(x_i \to y_j) = \langle y_j | e^{-iHt_1} | x_i \rangle$ refers to transition amplitude from outcome x_i at $t = 0$ to outcome y_j at $t = t_1$. This gives us additional $n^2 - n$ equations needed, and satisfies the requirement. Non-negativity of transition probability is guaranteed as well.

7 Conclusion

If what is written in this writing proves to be correct, then we would have a candidate for the complete theory of everything including quantum gravity. But this "complete theory" is not what we usually say as a theory. The "complete theory of everything" in this writing is simply an objective Bayesian inference framework fit to physical constraints. Thus this theory does not give any direct prediction for non-gravitational physics.

It does give a prediction for quantum gravitational physics, but to be able to form predictions for our reality, it requires taming our working quantum field theories so that subsystems do not possess infinite entropy - indeed, we expect Hilbert space of quantum gravity to be locally finite-dimensional[26]. This work of regularization is non-trivial, and it is yet unclear where our progress is at this point. (It does sound somewhat weird: that our own microscopic observers have internalized memory of structure of the universe, but we, macroscopic observers, do not know structure of the universe.)

Furthermore, even in Cao-Carroll (2018)[3], reconstruction algorithms are only worked out partially, and more works would be needed, despites signs pointing to correctness of ideas.

In any case, matters are real (ontic), but spacetime and state vector are purely epistemic, guided by objective Bayesian requirements.

References

[1] A. Almheiri, X. Dong and D. Harlow, *Bulk locality and quantum error correction in ads/cft*, *Journal of High Energy Physics* **2015** (2015) 163 [1411.7041].

[2] T. Jacobson, *Entanglement equilibrium and the einstein equation*, *Phys. Rev. Lett.* **116** (2016) 201101 [1505.04753].

[3] C. Cao and S. M. Carroll, *Bulk entanglement gravity without a boundary: Towards finding einstein's equation in hilbert space*, *Phys. Rev. D* **97** (2018) 086003 [1712.02803].

[4] B. Czech, J. de Boer, D. Ge and L. Lamprou, *A Modular Sewing Kit for Entanglement Wedges*, *arXiv e-prints* (2019) arXiv:1903.04493 [1903.04493].

[5] M. Cuffaro, *The kantian framework of complementarity*, *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* **41** (2010) 309 [1008.2904].

[6] J. Shore and R. Johnson, *Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy*, *IEEE Transactions on Information Theory* **26** (1980) 26.

[7] S. Pressé, K. Ghosh, J. Lee and K. A. Dill, *Principles of maximum entropy and maximum caliber in statistical physics*, *Rev. Mod. Phys.* **85** (2013) 1115.

[8] S. Pressé, K. Ghosh, J. Lee and K. A. Dill, *Reply to c. tsallis' "conceptual inadequacy of the shore and johnson axioms for wide classes of complex systems"*, *Entropy* **17** (2015) 5043.

[9] C. T. Sebens and S. M. Carroll, *Self-locating Uncertainty and the Origin of Probability in Everettian Quantum Mechanics*, *The British Journal for the Philosophy of Science* **69** (2016) 25 [1405.7577].

[10] R. R. Parwani, *Why is Schrödinger's Equation Linear?*, *Brazilian Journal of Physics* **35** (2005) 494 [quant-ph/0412192].

[11] R. H. Swendsen, *Explaining irreversibility*, *American Journal of Physics* **76** (2008) .

[12] A. Ashtekar and T. A. Schilling, *Geometrical formulation of quantum mechanics*, gr-qc/9706069.

[13] C. A. Fuchs, *QBism, the Perimeter of Quantum Bayesianism*, 1003.5209.

[14] A. Almheiri, D. Marolf, J. Polchinski and J. Sully, *Black holes: complementarity or firewalls?*, *Journal of High Energy Physics* **2013** (2013) 62 [1207.3123].

[15] E. P. Verlinde, *Emergent gravity and the dark universe*, *SciPost Phys.* **2** (2017) 016 [1611.02269].

[16] D. Frauchiger and R. Renner, *Quantum theory cannot consistently describe the use of itself*, *Nature Communications* **9** (2018) [1604.07422].

[17] M. F. Pusey, J. Barrett and T. Rudolph, *On the reality of the quantum state*, *Nature Physics* **8** (2012) [1111.3328].

[18] R. Bousso and L. Susskind, *Multiverse interpretation of quantum mechanics*, *Phys. Rev. D* **85** (2012) 045007 [1105.3796].

[19] W. H. Zurek, *Quantum Darwinism*, *Nature Physics* **5** (2009) [0903.5082].

[20] K. Hashimoto, *AdS/CFT as a deep Boltzmann machine*, 1903.04951.

[21] L. Martyushev and V. Seleznev, *Maximum entropy production principle in physics, chemistry and biology*, *Physics Reports* **426** (2006) 1 .

[22] K. Papadodimas and S. Raju, *State-dependent bulk-boundary maps and black hole complementarity*, *Phys. Rev. D* **89** (2014) 086010 [1310.6335].

[23] W. H. Zurek, *Pointer basis of quantum apparatus: Into what mixture does the wave packet collapse?*, *Phys. Rev. D* **24** (1981) 1516.

[24] X. Dong, E. Silverstein and G. Torroba, *De sitter holography and entanglement entropy*, *Journal of High Energy Physics* **2018** (2018) 50 [1804.08623].

[25] M. Alishahiha, A. Karch, E. Silverstein and D. Tong, *The dS/dS Correspondence*, in *The New Cosmology: Conference on Strings and Cosmology*, R. E. Allen, D. V. Nanopoulos and C. N. Pope, eds., vol. 743 of *American Institute of Physics Conference Series*, pp. 393–409, Dec, 2004, hep-th/0407125, DOI.

[26] N. Bao, S. M. Carroll and A. Singh, *The hilbert space of quantum gravity is locally finite-dimensional*, *International Journal of Modern Physics D* **26** (2017) [1704.00066].