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 laws themselves are understood in terms of inference and physical consistency requirements. 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 charge of subjectivity of measurements from quantum mechanics. An illusion of collapse comes from Bayesian updates by observer's continuous outcome data. Spacetime is to be understood in epistemic sense, instead of existing independently of an observer, following spirits of black hole complementarity. Dark matter and dark energy pop up directly as entropic tug-of-war in the reformulation.

Contents

1 Introduction

Section 2 and 3 revolve 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 updates a state vector when her own outcome arrives 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.[1][2] As with state vectors, spacetime is considered to be of epistemic nature, following spirits of black hole complementarity[3].

In fact, the word "reformulation" may even be misleading, as this is in essence modern Copenhagen interpretation, updated from traditional Copenhagen interpretation such that collapse of a state vector is just understood as a Bayesian update with new observer outcomes. What follows thus is what hidden consequences we can get from this modern Copenhagen interpretation by explicitly stating it out. However, the reformulation works with other consistent interpretations of quantum mechanics, as they all are supposed to give us same predictions.

Section 2 provides some intuitions for what is to come up in section 3. Details of a state vector and its Bayesian update is described in section 3. How the Born rule and Schrödinger equation may be derived from combination of physical and inference requirements is discussed in the section as well. In essence, this is just usual quantum mechanics, with exception of the principle of maximum entropy, compatible with all consistent interpretations of quantum mechanics. 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.

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 4. "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 5 discusses how one may recover spacetime from entanglement data from the principle of maximum entropy and area law.

Section 6 attempts to resolve the measurement problem completely. Section 2 already resolves the basis ambiguity problem of the measurement problem, but completeness of Bayesian inference is left to be demonstrated. The transition probability arithmetic fills in this hole.

The essential parts of this writing are: subsection 3.2 of section 3, section 5 and section 6. Together, they demonstrate that the objective Bayesian re-formulation of QM is consistent and illuminating.

1.1 Terminology

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.

A state vector represents degree of uncertainty an observer has about the entire system, assumed in this writing to be always pure. Thus, initially, different observers may have different state vectors. We distinguish states of the entire system from states of a subsystem that cannot be further factorized into subsystems. The latter states will be referred to as outcomes, instead of states for clarity of writing. An observer is always assumed to be a subsystem that cannot be factorized into other subsystems, unless qualified as a macroscopic observer. Observer outcomes thus refer to states of an observer.

Total entropy of the (entire) system will solely refer to sum of von Neumann entropy of subsystems: $\sum_i S_i$, where S_i refers to von Neumann entropy of subsystem i that cannot be factorized in other 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.

2 Epistemic nature of quantum mechanics

2.1 Area law

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 state 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[4][5][6][7] 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 state we are in, even if we know the state vector of a system. The Born rule provides an exact rule of uncertainty quantification (probability) from the state vector. But it does not say that spacetime is constructed out of ignorance, since post-measurements, spacetime would be coupled to definite matter states, not uncertainty.

In what follows, we will show that the "traditional" understanding is a wrong way to think of quantum mechanics - in fact it is not even orthodox understanding in modern sense. From an epistemic view of quantum mechanics, there is no problem. Spacetime arises out of epistemic requirements, just as black hole complementarity[3] mandates.

2.2 "Reality" of an observer

An observer can only notice and picture exterior world by variations of her outcomes. Even spacetime is visible only through these variations. Effectively, she "constructs" spacetime such that some reality can be seen from her epistemic state vector. Notice that while the word "constructs" is used, the word is used only for convenience, as there is no active measurement. In fact, as one expects, in physics, there is no place for subjective measurements. Thus this "construction" of spacetime is done automatically without requiring active and subjective measurements.

The above seems to go counter to anti-realist theorems of quantum mechanics such as Hardy's paradox, but this is a misunderstanding. In fact, the above is just minimal Copenhagen interpretation, with the collapse postulate updated in modern sense such that "collapse" really is just an observer checking her own outcomes to measure exterior world, which is updating a state vector by outcomes. But because all consistent interpretations of quantum mechanics provide same predictions, the reformulation will be compatible with all these interpretations as well. It is just that Copenhagen interpretation, suitably modernized, provides most minimal set of assumptions for formulating quantum mechanics. For Hardy's paradox, this is about what is being measured, not an observer. Thus even if an observer remains "real" while other subsystems remain "anti-real", there really is no problem. In fact, this aspect is what presents an illusion of observer supremacy and omniscience in modern Copenhagen interpretation. Thus, anti-realist theorems present no problem in upholding "reality" of an observer.

Still, one may object that the above view of quantum mechanics does leave subjectivity as to when an observer updates a state vector. This is again a misunderstanding. An observer does not "choose" to see outcomes - an observer is forced to see outcomes and update them into state vector. In fact, we will assume reasonably that variations of observer outcomes are continuous over time. Thus, an observer continuously updates her state vector. This completely gets rid of charges of subjectivity in quantum measurements. An observer does not know what outcome she will see at the end of the experiment, thus usual quantum mechanical calculations are saved in the reformulation.

2.3 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. It provides why despite measurements of an observer herself continuously existing at all time, only few are actually perceived by a macroscopic observer.

2.4 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. 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 that one can follow from Page curve[8]: 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, as a state vector is of her own.)

3 State vector from objective Bayesianism

3.1 Principle of maximum entropy and state vector update

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[1][2][9].

The question is how to update a state vector with new observer outcomes consistent with the principle of maximum entropy. The answer of course is the Bayes rule, but it does not provide a way of gathering and comparing together past outcomes and present outcomes. This requires probabilistic modeling of how reality evolves, because we need to think of how past and present states are related.

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

$$i\frac{d}{dt}|\Psi(t)\rangle = H(t)|\Psi(t)\rangle \tag{3.1}$$

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

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 present outcomes. Prior distributions already contain all past data. In case of quantum mechanics, this needs to be refined with preceding posterior probabilistic distributions (state vector) evolved by Schrödinger equation and updating with a present-time observer outcome, but rest remain the same.

While infinitesimal arguments are not rigorous - and we will remedy this in subsection 3.2 - the general intuitions can be captured as follows:

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.

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 individual 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 $Z = tr[e^{-\beta H(t)}]$ ($\beta = 1/\left[k_B T(t)\right]$, with k_B Boltzmann constant, $T(t)$ temperature) has to stay invariant across time. One may ask why.

A natural motivation comes from studies of renormalization group in thermodynamics, where the idea is capturing same physics but with different succeeding effective descriptions. There, it is the above Z that is constant. This is what we want as well.

Initial state vector of the universe The initial state vector of the universe at the beginning of the universe is not assumed to be set by the principle of maximum entropy. This is because entanglement is about correlation, but at the beginning, there should not have been any epistemic correlation between subsystems. A subsystem only just started learning about other subsystems - thus the initial state vector of the universe must be a tensor product of pure state vectors that describe each irreducible subsystem. The consequence is that the initial state vector of the universe allows one to identify individual irreducible subsystems by Hilbert space factorization.

Basis ambiguity does not exist It is said that quantum mechanics does not provide a rule for a basis on which an outcome comes out of a state vector. But in the reformulation, it is observer outcomes that determine the basis. Or better wording would be that a new outcome would be projected onto a selected basis vector for representation, as physics does not care which basis is selected. Thus, this part of the measurement problem is resolved.

One may argue that this runs counter to an observer having some freedom on what basis some subsystem would be measured in an experiment - but there is nothing contradictory here. In classical physics, humans do not really have "freedom" to build a computer - physics dictates that we will build a computer. But one can imagine as if a human has freedom to do something when the universe is approximated.

3.2 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 irreducible 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, $i\hbar d|\Psi(t)\rangle/dt = H(t)|\Psi(t)\rangle$.
- "Budget" constraint: given constant Z where $Z = tr[e^{-\beta H}]$, $\beta = 1/[k_B T(t)]$.
- Observer outcome trajectory (Bayesian update constraint) that constrains possible state vectors at each time: the observer outcome at t must have probability of 1 in state vector at t.

3.3 Derivation of quantum mechanical laws

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

Why the Born rule? Fortunately, a derivation of the Born rule from Bayesian inference requirements was already done in Sebens-Carroll (2016) [10], 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.

The ESP idea is simple. Consider the universe consisting of system A, observer O and environment E. Now assume that the state of the combined system of $A + O$ is exactly known at $t = t_0$ - which means it will be represented by a pure state vector. This implies that the state of the environment is in a pure state vector as well at $t = t_0$. At $t = t_0$ then, how one assigns probability to a state vector of $A + O$ should be independent of what the state vector of the environment E is (candidates differing by unitary transformation). However, it must be simultaneously true that a state vector of the whole $A + O + E$ must represent probability of states of the entire universe as well. Resolving these requirements results in the Born rule.

Why Schrödinger equation? 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[11]. In Parwani (2005), linearity of Schrödinger equation follows from physical constraints motivated from classical physics (but those that also seem to apply to quantum systems as well) and the principle of maximum entropy. One can argue that the physical constraints do not apply for non-classical systems, though, and this would be up to empirical evidences to check out. So far, these constraints and Schrödinger equation held up empirically, so linearity in Schrödinger equation is sound.

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 statistics. 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 - because converged final Hamiltonian can always be used to make predictions. There, whether one is a subjective Bayesian, objective Bayesian or classical statistician does not matter.

Evidence: The law of maximum entropy production In non-equilibrium classical thermodynamics, Swenson and others[12] 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.

4 Consequences of the QM reformulation

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:

$$\sum_x P(x)P(x \to z) \neq P(z) \tag{4.1}$$

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 probabilistic inference completeness - the inequality in Equation 4.1 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.

4.1 Strong complementarity

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.

Firewall paradox Because of strong complementarity and epistemic nature of quantum mechanics, the firewall paradox[13] 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, as we will see in section 5.

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.

Notion of distance: why same H in usual circumstances? Observers largely share same H because they are very close to each other. When subsystem X comes to extract most of available mutual information with its complement $\bar{X}$ from my point of view, X must be very close to me via locality - which just says information flows to me from nearby. This connection allows one to define notion of distance from entanglement. And consistency of underlying reality mandates that approximately same H should be shared between me and X.

Now this story seems to be opposite to the usual quantum field theory story[6]: entanglement of subsystems usually decreases with increasing distance. Thus we expect more entanglement whenever distance decreases. This is not the case. From my point of view, observer Y far away from me is more entangled with subsystems near Y than subsystems near me. What we need to remember here is that a state vector and entanglement must be understood from an observer point of view. Thus, there is no inconsistency.

How locality is related to mutual information will be discussed in depth when talking of "spacetime from entanglement", specifically in section 5.

4.2 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.[14] The novelty here is matching dark matter with outcome-induced entropy and dark energy with the principle of maximum entropy.

4.3 Basis, decoherence and causal diamond complementarity

This writing was not initially inspired by Bousso-Susskind (2012)[15], but it became clear that there is a heavy connection. Let us review the two ideas in Bousso-Susskind (2012).

Observer trajectory First, an observer trajectory used to define a series of causal diamonds in Bousso-Susskind (2012). This appeared in this writing as data for a state vector update.

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 may seem strange.

But we believe 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 in Everettian interpretation.

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. This has been the fundamental point in the reformulation of quantum mechanics.

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

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

As observers come 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, as we come to learn more about each other. This requires that classical reality of different observers dominates, as expected. This restricts the number of plausible bases in which an observer outcome can arrive as a subsystem basis vector, because different observers must agree epistemically as well.

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.[16] 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.

5 Spacetime from entanglement

5.1 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 inspires the conjecture of area equaling mutual information.

It is widely believed that while the conjecture holds approximately, it does not hold exactly. However, there is value in approximating spacetime reasonably, especially in case the conjecture does hold exactly. Thus we will proceed as if the conjecture holds exactly.

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:

$$A(X;\bar{X}) = \frac{1}{2\alpha} I(X;\bar{X}) \tag{5.1}$$

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 $\delta[\sum_i S_i] = 0$. The area perturbation data allow one to recover emergent metric - recovery details are left to Cao-Carroll (2018)[6], though in this writing the RC state restriction is dropped without any adverse consequences, because recovery itself is independent of the restriction. In "entanglement equilibrium", $\delta S_i = 0$ is recovered from the principle of maximum entropy, as in [6].

5.2 Story of Big Bang cosmology

Because the initial state vector of the universe exhibits no entanglement, 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. Note that this story holds independent of which spacetime construction turns out to be general - as far as the area law holds well approximately.

5.3 Alternative spacetime constructions

The spacetime-from-entanglement alternatives to [6] exists - though they are understood in terms of AdS/CFT, such as entanglement holonomies[7], error-correction codes[4] and neural network emergence[17]. Recent dS/dS correspondence[18][19] may allow us to transport results in AdS/CFT to our realistic spacetime eventually, but for now where we stand as to exact recovery of spacetime remains unclear.

5.4 Spacetime from neural network

Recently, a paper[17] that casts AdS/CFT correspondence in terms of a deep Boltzmann machine has appeared - which presents bulk as learning about boundary.

If the idea that trained weights of the neural network create emergent spacetime is correct, then spacetime must be of epistemic nature, as assumed throughout the writing, unless spacetime depends on unknown future outcomes as well.

The bright line behind this idea is instead of trying to first state laws of spacetime construction, we may simply figure out emergent spacetime directly from data. This is much more natural from epistemic perspectives of quantum mechanics. After all, why would spacetime be special, when quantum mechanics simply is Bayesian inference? Furthermore, the "data" for the neural network do not necessarily have to be in style of boundary CFT, technically speaking.

And some ideas, such as entanglement holonomies[7], may not be distinct from the neural network idea, as we need to refine how a neural network would be structured. Those refinements may be provided by these ideas, just as we arrived at the Born rule from epistemic requirements.

5.5 Error-correcting code and decoherence

Quantum error-correcting code vision[4] of holographic duality becomes very useful, in relation to quantum decoherence (subsection 4.3). 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[20] suggests that sufficient decoherence takes time, even if little, for quantum redundancy to form, so error correction is not instantaneous - the main point being that redundancy has to be formed, instead of being just there.

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.

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 in the right direction, then it suggests that we are at least fairly close to completing the project of quantum gravity. Both spacetime and state vector are results of probabilistic inference of an entire system from observer outcomes, with spacetime potentially rising out of epistemic requirements[7][17], just as quantum mechanics arises out of epistemic requirements. We believe that the promising future direction is on finding epistemic requirements to notion of metric to recover spacetime.

Despite this partial (in)completeness of quantum gravity, there are things we newly arrived from the reformulation of quantum mechanics, which may be said as modern Copenhagen interpretation explicitly stated. First, the firewall paradox[13], even if valid in arguments, is not paradoxical because of epistemic nature of quantum mechanics. Second, spacetime does not exist independently of observer state of knowledge. This is a direct consequence of spacetime having to be derived from quantum mechanics. Third, potential candidates for dark energy and dark matter come out directly from the explicit reformulation of quantum mechanics - dark energy from the principle of maximum entropy and dark matter from arrival of observer outcomes. That dark energy and dark matter being entropic has long been suspected[21], and this writing provides a more solid foundation toward that direction.

There is one thing worth mentioning separately. In order to provide realistic predictions, the reformulation of quantum mechanics in this writing, just as with modern quantum gravity theories, requires "regularizing" 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[22]. This work of regularization is non-trivial, and it is yet unclear where our progress is at this point.

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

References

[1] 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.

[2] 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.

[3] L. Susskind, L. Thorlacius and J. Uglum, *The stretched horizon and black hole complementarity*, *Phys. Rev. D* **48** (1993) 3743 [hep-th/9306069].

[4] 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].

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

[6] 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].

[7] B. Czech, J. de Boer, D. Ge and L. Lamprou, *A Modular Sewing Kit for Entanglement Wedges*, 1903.04493.

[8] D. N. Page, *Information in black hole radiation*, *Phys. Rev. Lett.* **71** (1993) 3743 [hep-th/9306083].

[9] 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.

[10] 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].

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

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

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

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

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

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

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

[18] 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.

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

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

[21] E. Verlinde and H. Verlinde, *Black hole entanglement and quantum error correction*, *Journal of High Energy Physics* **2013** (2013) 107 [1211.6913].

[22] 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].