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. The firewall paradox, analyzed by a parallel with Hardy's paradox, is used as an example supporting necessity of the reformulation and its consequential results. Other clues pointing to this reformulation are analyzed. The reformulation, with the addition of novel transition probability arithmetic, eliminates basis ambiguity and the collapse postulate, thereby eliminating subjectivity of measurements from quantum mechanics, and resolving the measurement problem completely.

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 and 3 are there to establish the need for the theory of theories - that is, the law of which theory (Hamiltonian) to use for an observer at each time. Along the way, the collapse postulate, requiring subjectivity of measurements, is shown to be problematic in the firewall paradox context. Analysis is done through a parallel (two incompatible descriptions) between the firewall paradox and Hardy's paradox. Because of this parallel, a consistent theory of theories satisfactorily resolves the basis ambiguity part of the measurement problem.

A basis is considered always fixed in the reformulation of quantum mechanics. Thus basis transformation is disallowed - if there is to be basis transformation, it is absorbed by transforming a theory instead.

Once this is established, consistency requirements of Bayesian inference are transported to quantum mechanics in section 4, now viewed as an inference framework. The theory of theories complements on traditional postulates of quantum mechanics minus the collpase postulate by the principle of maximum entropy and invariance of Z, cast in terms of renormalization. The epistemological problem of quantum mechanics is that we do not know state vector evolution but have laws of how state vector must evolve, with free parameter being a Hamiltonian that is conserved across time. Despite imposed conservation of Hamiltonian, a learned Hamiltonian may be different over time. After all, if learned Hamiltonians are forced to be invariant over time, then the principle of maximum entropy is not operational, and requirement of the theory of theories, established in section 2 and 3, would not be met.

Technically, this writing can be read from section 4 section, as far as one accepts that quantum mechanics is just a Bayesian inference framework. The idea of the theory of theories naturally rises out of physical constraints and the principle of maximum entropy, viewed just as a consistency requirement.

In section 5, how spacetime arises from entanglement is reviewed, also hinting toward analysis established in section 4. An important inspiration is that in AdS/CFT, bulk works as quantum error correction for variations in boundary states. This suggests that despite strong complementarity, observers share effectively same description, up to some deviations in descriptions. When error correction is impossible, there goes a black hole. This responds to a potential criticism of strong complementarity. It will be shown that entanglement equilibrium literature provides a direct connection to the principle of maximum entropy as well.

A foundation for thermodynamics and fundamental irreversibility is reviewed in section 6, and is found deficient, still conflicting somewhat with reversibility in quantum mechanics, despite uses of thermodynamics still proven to be sound. How the theory of theories fixes this deficiency is explored. Frauchinger-Renner thought experiment[1] is mentioned as an example necessitating irreversibility of nature.

Section 8 provides a complete resolution of the measurement completely and thereby completing the fully objective formulation of quantum mechanics without the collapse postulate. This is done by providing a transition probability arithmetic from one outcome at some time t to an outcome at another time $t + dt$. The arithmetic couples Born's rule with quantum decoherence understanding of measurements.

Additional potential concerns to the theory of theories and strong complementarity are addressed in section 7.

That is, reformulation of quantum mechanics from Bayesian inference requirements without the collapse postulate and subjectivity of measurements would be done in this writing. The by-product of this approach is the theory of theories - or the principle of maximum entropy - which is the center of focus. Subsection 4.6 presented the completed reformulation, along with further discussion as regards to deriving Born's rule and Schrödinger equation from inference requirements in subsection 4.10.

2 Traditional weak (black hole) complementarity

The key point of some traditional variants of black hole complementarity is that even if Alice and Bob have two seemingly different descriptions of the world, they are not actually inconsistent. In fact, two descriptions are different only in disguise - they are equivalent by some mapping. This resolves black hole information paradox automatically.

Except things are not that easy. This is because we expect that semiclassical effective field theory is valid around event horizon of a black hole, and that an observer entering a black hole would see nothing special at event horizon. Thus, one has to check whether our effective descriptions would be consistent. A simplified verification mechanism by a qubit model is whether an observer can notice a quantum cloning of a same qubit by exploiting on two different descriptions - if so, two descriptions are inequivalent - contradiction.

This is what many thought experiments on black hole complementarity set out to do, the AMPS thought experiment[2] belonging to this category.

2.1 The firewall paradox

An exposition of the firewall paradox is given in Bousso (2013)[3], which will be used as a reference of the paradox for the rest of the writing. The word "theory" in Bousso (2013) would translate to "description" in this writing.

The key issue, to quote Bousso, is that both Alice (infalling observer) and Bob (distant observer away from a black hole) can stay outside a black hole until key point of time to measure one set of entanglement. But right before Alice jumps in, the causal diamond of Alice contains all of late Hawking radiations. Also, Alice can picture herself measuring interior modes later, which establish another incompatible set of entanglement. That is, Alice can see two incompatible entanglements existing in her causal diamond.

The key argument of course is that Alice has to agree with Bob's description even if Alice did not observe all of what Bob can see. That interpolation of what Bob is about to see is not the center of the controversy - the central controversy is on whether unmeasured outcomes in quantum mechanics really have meaning.

2.2 Why not go for very strong complementarity?

Before exploring a parallel between Hardy's paradox and the firewall paradox, it is wise to explore why a very strong notion of complementarity is not widely accepted. It may be argued that because Alice did not actually see all late Hawking radiations, she is saved from contradictions.

However, complementarity that is this strong causes a problem of what probability in quantum mechanics really stands for. Suppose there are two "paths" - one path has Alice never measuring object O and the other path has Alice measuring some object O. In case only measurements have results - then probability of an unmeasured outcome does not even make sense. This seems to require a heavy modification of quantum mechanics.

The idea of causal diamond complementarity[4][3] is free of the aforementioned trouble. The idea is that as long as O is in the causal diamond of Alice, it makes sense to talk of Alice not measuring O and discuss probability of such a (non-)event. Thus, complementarity is simultaneously tamed and embraced by restricting applications of quantum mechanics to a causal diamond of an observer. But we have seen that the firewall paradox presents a problem for causal diamond complementarity.

So if one does not like alternatives given for black hole complementarity - for example, one may hesitate to adopt the state-dependence idea[5][6][7] because of counter-arguments such as Harlow (2014)[8] and Marolf-Polchinski (2016)[9] - what should one search for? (However, this rejection is only for motivations - eventually, we will see ideas inside the alternative coming back, and thus there actually is no rejection.)

2.3 A parallel between Hardy's paradox and the firewall paradox

In examples like Hardy's paradox, the lesson of "unmeasured outcomes have no result" is often given. How far can this be applied against the Harlow-Bousso argument for initial validity of the firewall paradox? But to do that, we need to establish common aspects of two paradoxes.

Despite the setting of a black hole creating an illusion of more complexity, the core problem is the same. Hardy's paradox demonstrates that two descriptions of the same system, related by basis transformation, can never simultaneously hold for an observer. This is the lesson we want for the firewall paradox as well. In fact, the connection hints that possible descriptions in the firewall paradox be related by basis transformation. The difference between Hardy's paradox and the firewall paradox is that in former, observers have access to only one description at each time. In the firewall paradox, while an individual observer still has access to only one description at each time, observers do have simultaneous access to multiple descriptions. Thus, if faithfully implementing the connection, one can say that the firewall paradox is Hardy's paradox generalized.

There is one problem to establishing the connection - quantum mechanics is silent about which basis would be selected for observations. Thus, if we port Hardy's paradox to the firewall paradox, subjectivist interpretations of quantum mechanics require both descriptions of a black hole to be available for an observer inside or outside a black hole.

But the question of why a particular black hole description must be used for an observer is indeed a valid one. Thus, the aforementioned problem does not serve as a good reason to reject the connection.

The above consideration suggest that if quantum mechanics is complete, then it would very likely have to be objectively so - disallowing freedom of observers. (The objective-subjective question is independent of whether quantum mechanics is epistemic or ontic.) This requires a rule on how a basis is chosen for an observer, which eventually will be elevated to a need for a theory of theories.

The objective completeness also requires filling in details of transition probability of outcomes in one different description at $t = \Delta t$ from some measured outcome in another description at $t = 0$. This question will be deferred until section 8.

2.4 From the basis selection issue to the theory of theories

The aforementioned consideration suggests that one needs to provide a theory of descriptions - how one description/basis is chosen for an observer at each point of its dynamics. But any unitary basis change U from description D_1 to D_2 at time t can be transformed into $e^{-iVt} = Ue^{-iHt}$, by which one instead transforms Hamiltonian H originally shared by D_1 and D_2 into Hamiltonian V for description D_2, assuming that D_1 and D_2 shared the same basis at time $t = 0$, with invariance of Hamiltonian assumed for each description.

Why even transform a basis change into a theory change? A basis naturally should not have any physical interpretation. It is just a device that one uses to map physical outcomes to mathematical forms. If a basis selection issue is so fundamental, why not deal it properly in a theory? Adding a theory of theories is much better than a theory of basis selection. This leads to **strong complementarity** - different observers may even face different Hilbert spaces, and even an individual observer is attached with a series of Hilbert spaces, instead of just one Hilbert space. The latter aspect of strong complementarity can also be mapped to a requirement that an observer must be allowed to be "given" different bases over time for measurements. (It is implausible to think that an observer would measure some subsystem in position basis - over time, it may be measured in momentum basis, and then later position basis and so on.) Again, it is worth emphasizing that the different-bases analogy is only an analogy from now on, as a basis is now assumed to be fixed - the concept of different bases is eliminated.

But an important criticism may remain: that the basis selection issue is already resolved by quantum decoherence literature. I will argue that this is not the case. In fact, this is not unacknowledged by quantum decoherence literature - see Zurek (2009)[10], where despite progress in understanding the measurement problem, a complete resolution is not yet there.

Quantum Darwinism would be considered part of quantum decoherence literature for rest of the paper.

3 Quantum decoherence literature does not resolve the basis selection issue

The argument is quite simple. If one adds nothing to traditional postulates of quantum mechanics, it is silent about which basis must be chosen. Then how does quantum decoherence literature can talk of some pointer basis being "chosen" for measurements out of quantum mechanics? It involves at least one of the below list:

- Subjectivist elements. An observer has some conscious controls over what measurements to be performed. Some basis or observables may be better for an observer to learn about some system.
- An implicit additional law added to quantum mechanics so that basis selection can be done.
- Quantum decoherence literature does not intend to resolve the basis selection issue - it shows how one can represent emergence of measurements in our classical measurement, but it does not show how that emergence is forced.

If the last of the list is the case, then no further discussion is needed. The first option will be rejected because of the discussions above. Thus, let me focus on the second option.

The problem would not be that an additional law is given - rather, it is that the law is not explicitly spelled out. Recognizing that one is providing a new law in addition to traditional postulates of quantum mechanics (in replacement for the collapse postulate), the question shifts to whether a particular law would be justified.

The main huddle in implicitly proposed laws - for example, choosing a basis that brings reduced density matrix closest to being diagonal[11] or one that requires least amount of access of environment for measurements[10] - is that subsystems may have different privileged bases. But clumping together different basis descriptions is not possible because of results in Hardy's paradox.

Thus a cleaner law is proposed: a theory that maximizes entropy of an observer under constraints is used. A theory of theories is proposed instead of basis selection, because a basis is now assumed to be fixed.

4 Quantum mechanics as an inference framework

Despite quantum mechanics assumed to be an objective framework, it is still open to both epistemic and ontic interpretations. Reasonably, one may expect that whether quantum mechanics is epistemic or ontic really is matter of semantics - they are equivalent as far as actual physics goes, with one interpretation making explanations simpler in some contexts. With this equivalence assumption in mind, I will take the epistemic view of quantum mechanics for rest of the writing.

A state vector in simultaneously objective and epistemic interpretations describes knowledge an observer has. This state vector may not be something actually existing, but the state of knowledge that an observer has about the world is objectively determined and "real", encoded as a state vector.

4.1 Defining an observer subsystem: observer identification problem

Because a basis is assumed to be fixed - thus there is only one basis used - and we are manipulating on theories, it is no longer true that entropy of a subsystem is invariant. A subsystem is defined in terms of the sole basis, and that is all one needs for defining a subsystem. This point would have been obvious, but stated just in case there is confusion. Thus, maximizing entropy of an observer would make sense.

The real problem is how we may epistemically figure out different subsystems. That is, what would count as an observer? If theories selected by an observer differ from another, and we want quantum mechanics to be objective, we would wish to be able to define what would be considered an observer subsystem.

That problem is difficult to answer. We can provide observer information directly, but while this works for matters of calculation, it is unsatisfactory. One optimal possibility is that at asymptotic past infinity, every observer could be said to choose a theory such that its shared Hilbert space is factorized to Hilbert spaces of observers as their tensor product. This allows for clean identification of observers.

4.2 Bayesian inference: why maximize entropy of an observer under constraints?

In Bayesian inference, a prior is often chosen by the principle of maximum entropy, incorporating constraints or information an observer has. The simple justification of the principle usually given is the Wallis derivation given by Edwin Thompson Jaynes. However, more recent approaches, starting from Shore-Johnson (1980)[12] simply cast the principle of maximum entropy as rising out of consistency requirements[13] - that is, consistency of statistical inference requires the principle[14]. There is indeed the question of functional form of entropy - but von Neumann entropy as a functional form of entropy is safe.

Now onto why entropy of "an observer" in particular instead of others. The point is that an observer Alice only sees an external world by noticing variation of herself. Combine these two points, and we get the principle of maximum entropy of an observer under constraints.

A remaining question then would be what constraints are. We can immediately list one of them:

Schrödinger equation under conserved Hamiltonian

But an optimization problem, cast in terms of control and state vector, points to us that one more constraint is required. To avoid going into details until later in the writing, let me give an example of economics: optimization problems in economics provide a law for state vector evolution and a separate budget constraint to provide a unique solution. It is this budget constraint part that would need to be added.

4.3 Invariance of partition function Z

And invariance of partition function Z, which is actually constant for the purpose here, is this budget constraint. Z, in statistical mechanics is defined as:

$$Z = tr\left[e^{-\beta H}\right]$$

where tr is trace, $\beta = 1/(k_B T)$, with k_B Boltzmann constant and T being temperature.

In quantum field theory, Z is a normalization constant, but for correct comparisons and predictions, one does need to keep it constant. In terms of Euclidean field theory by Wick rotation (or almost equivalently statistical mechanics), Z takes a role of unitarity requirement - just that the constant sum of probability no longer is 1.

Effectively, we have formulated quantum mechanics of time in $\mathbb{C}$ with two probability normalization requirements. But this would not matter for rest of the writing - and time will be kept in $\mathbb{R}$.

Also, the principle of maximum entropy of an observer already places quantum mechanics in fully thermodynamic settings, as the second law of thermodynamics is now satisfied. Thus it would be better to take the parallel of quantum mechanics with statistical mechanics very seriously.

4.4 Renormalization

One important clue toward aforementioned analysis is renormalization in statistical mechanics. Simplified, it is about keeping basic properties of a system invariant, while sacrificing on some other properties. In terms of neural networks, it is about compressing data such that relevant information can be extracted readily. In renormalization, by definition, Z is supposed to be kept constant.

Renormalization possesses several ideal characteristics that we may wish to transport to our analysis:

- Renormalization group (RG) flow is irreversible. Nature over time seems irreversible.
- A scale parameter is in $\mathbb{R}$. Time in our Hamiltonian formalism is in $\mathbb{R}$
- Initial and end fixed points of RG flow can be considered. We want different observers to share same learning outcomes asymptotically - something like S-matrix though not necessarily in terms of scattering, especially in that out-states are not to be asymptotically free particle states. This is especially preferable given the observer identification problem.

Thus, we may say that observers learn of nature by doing renormalization with the principle of maximum entropy and Schrödinger equation.

In renormalization, it is usually understood that (entanglement) entropy decreases in direction of RG flow, as renormalization "takes out" UV degrees of freedom. But what one wants here is the opposite - increase of entropy. Reconciling this would be easy - we think of RG flow as going from IR to UV. This is actually a wrong description - it rather reflects underlying duality of nature where what is IR (UV) in one description is UV (IR) in another description. Thus RG flow would still fundamentally flow from UV to IR.

4.5 Objective physics of an observer: why the principle of maximum entropy is still valid

A counter-argument may be made against the principle of maximum entropy: since an observer is obtaining measurements, it seems unreasonable to impose the principle of maximum entropy directly to a state vector without updating for measurements.

The argument itself is invalid, because post-measurement, one can always reset the entire inference process and change initial conditions - for details, see subsection 4.6. It is just as collapse in subjectivist interpretations of quantum mechanics "resets" state evolution.

However, one can go deeper, even though the above is enough for most practical uses. Another issue in the straw-man argument is that it assumes an effectively omniscient observer. We do have to think of whether observers actually can keep records of their measurements and whether they accidentally receive information without active measurements. This requires inference in terms of a state vector of the entire system. However, quantum decoherence, with the basis selection issue resolved, provides why one may effectively reset inference process by setting initial conditions with a measurement outcome.

Another way to think of the straw-man argument is in terms of recoherence. Even decoherence, which allows for illusion of collapse, has limits - while they hold reasonably under our largely macroscopic environments, unitary evolution in many other circumstances create recoherence after decoherence. Thus, an observer should not always be assumed effectively omniscient.

Of course the argument presupposes rejection of subjectivist interpreations of quantum mechanics, and if one is, at this point, not already convinced of necessity of the rejection, there is nothing more that can be done.

4.6 Optimization problem formulation of the theory of theories

We are now in the position to complete the reformulation of quantum mechanics and thus the theory of theories. Rest of the writing would be on studying clues pointing toward this reformulation and analyzing its details.

Optimization problem

> For subsystem A, entropy $S_{A,t} = tr\left[\rho_{A,t} \log \rho_{A,t}\right]$ of A, where tr represents trace, ρ_A represents reduced density matrix of A, is maximized at time t subject to initial conditions, law of state vector evolution and the "budget" constraint.
>
> Initial conditions: initial state vector $|\Psi_{t_0}\rangle$ at $t = t_0$, partition function constant Z.
>
> Law of state vector evolution: $|\Psi_t\rangle = e^{-iH_t\Delta t}|\Psi_{t_0}\rangle$ with $\Delta t = t - t_0$.
>
> Budget constraint: invariance of $Z = tr\left[e^{-\beta H_t}\right]$.
>
> Controls at each time: H_t, T_t, where H_t represents chosen Hamiltonian at time t required to be self-adjoint, T_t represents temperature at time t (to pipe into β of Z).
>
> State: $|\Psi_t\rangle$

t_0 may be sent to asymptotic past infinity, if we are interested in state vector evolution of the entire system, not just a subsystem of subsystems that are effectively modeled together as residing in pure state vector.

4.7 Quantum mechanics reformulated

From aforementioned analysis, one can re-interpret quantum mechanic (including the theory of theories) as follows:

> Quantum mechanics is a set of procedures to learn state vector's evolution or how world evolves. Given only knowledge of the law of state vector evolution, Schrödinger equation of invariant Hamiltonian, and partition function Z, the principle of maximum entropy, now viewed as the solely consistent way of doing inference, requires us to maximize entropy of an observer under these constraints. Initial conditions provided would be initial state vector and Z. This is to be framed as an observer learning nature via renormalization.

Note that invariance of Hamiltonian for Schrödinger equation does not mean that Hamiltonian learned will be invariant across time.

In "reality", Hamiltonian will be conserved in Schrödinger picture of quantum mechanics. However, we do not know what this Hamiltonian is, and have to learn it. At each time, given constraints, a different Hamiltonian can be learned even when we take the constraint that Hamiltonian is conserved across time.

4.8 AdS/CFT as a neural network

Recently, a paper[15] 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. 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.

That is, the learning mechanism is missing to train the bulk from the boundary data. That missing piece is of course the theory of theories explored before. Furthermore, this Boltzmann machine picture gives another justification into why partition function is kept invariant, and how quantum mechanics relates to statistical mechanics and thermodynamics.

4.9 Equivalence of subjective-objective, epistemic-ontic interpretations

So far, two main dividing lines of interpretations of quantum mechanics were stated: subjective-objective and epistemic-ontic.

Having taken an epistemic view of quantum mechanics, there surely would have been something weird. An observer is seen as a learner, yet quantum mechanics was said to be objective. This of course is matter of semantics. We could say that observers have subjective powers. But these powers are immediately crushed by inference consistency requirements. After all, this is just objective Bayesianism patching on subjective Bayesianism. Thus, all subjective interpretations can be made objective as long as inference consistency requirements constrain subjectivity significantly, and there is not much value distinguishing the two categories.

Epistemic interpretations can be made ontic as well. If epistemic interpretations are about understanding quantum mechanics as recording state of knowledge of an observer, instead of representing existing reality, then we could simply elevate state of knowledge of an observer as reality.

This would not be possible if we only allow for a single universe. But we can allow for multiverse. When observer Alice chooses a different theory as being learned, we could say that she moved from one universe to another, where universe is equated to a theory. Each universe still carries time-invariant Hamiltonian.

Thus equivalence of epistemic, ontic, subjective and objective interpretations of quantum mechanics, and we can use whatever interpretation is more convenient for explanations.

4.10 Theme: formulating quantum mechanics fully from Bayesian inference requirements

The core theme of this writing is reformulating quantum mechanics fully from Bayesian inference requirements. What remains to be tackled for this reformulation is: derivation of Schrödinger equation, conservation of Hamiltonian (and partition function Z) and Born's rule. If we get these three derived from Bayesian inference requirements, then quantum mechanics is completely reformulated in terms of Bayesian inference.

Of course we do not really need to derive the three remaining postulates - we can simply assume and move on. But as will be seen, there are values in deriving them.

General Schrödinger equation that accommodates Hamiltonians that are not conserved over time is easily derived when unitarity of a state vector and linearity of quantum operators are assumed. And unitarity has to do with Born's rule. Thus, if Born's rule can be derived from some Bayesian inference requirements, then unitarity assumption is justified.

Fortunately, the derivation has already been done by Sebens-Carroll (2018)[16]. One important thing to note about the cited paper is that it is not actually restricted to Everettian, or many-worlds, interpretations of quantum mechanics despite the title. The ESP principle (epistemic separability principle) used to derive Born's rule is inherently a Bayesian inference principle that does not require references to many-worlds interpretation. Thus, it can be used safely to complete formulation of quantum mechanics as a Bayesian inference framework.

One may be curious as to why a more general form of Schrödinger equation is used, instead of the more restricted one used for the optimization problem. But it is easy to realize that this general form is valid as well, because we can map Hamiltonian $H'(t)$ used in general Schrödinger equation to a set of time-invariant Hamiltonians H_t at each time t. With this map, one can reformulate the optimization problem to use the more general form of Schrödinger equation.

That still requires us to derive why Schrödinger equation has to be linear. Again, fortunately, such considerations were already explored using information-theoretic and thus Bayesian inference requirements[17]. 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 if this approach to linearity of Schrödinger equation is correct, then we need to answer why Lorentz invariance must be kept. For now, I would just stop here, but justifying Lorentz invariance is going to be far easier than linearity.

Lastly, why time-invariant Hamiltonian must be used, despite learned Hamiltonians may be different oer time. We would like learning results of different observers on a same system to converge. But the results may only become same in asymptotic past and future infinity. Having justified the general form of Schrödinger equation, one can notice that the additional requirement only makes sense when Hamiltonian is time-invariant. Talking of time-varying Hamiltonian in asymptotic past and future infinity is simply absurd - it means there really is no convergence.

These asymptotic final states (for an entire universe, not necessarily each subsystem) are essentially thermal equilibrium states - thus justifying invariance of Z. This completes full reformulation of quantum mechanics in terms of Bayesian inference (or one may prefer the word information theory).

Of course more substantiating and verification works need to be done on proofs of linearity of Schrödinger equation from inference or information-theoretic principles. But for now at least, indications are good.

5 Quantum redundancy: spacetime from entanglement

5.1 Basis redundancy

It was recently argued (Czech et al. (2019)[18]) 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)[6] as well.

That is, spacetime is sewing together essentially redundant descriptions. This redundancy protection has limits, which the cited paper discusses with reference to error-correcting code literature of holographic duality. Also, the basis selection issue would have to be resolved by an additional clarification, which was provided in this writing.

5.2 Error-correcting code

We have seen the parallel between Hardy's paradox and the firewall paradox in subsection 2.3. One point to rethink here is the question of why in Hardy's paradox we do not see a problem for quantum mechanics, but a problem seems to exist for the firewall paradox despite the parallel.

This is where quantum error-correcting code vision[19] of holographic duality becomes very useful. Basically, the idea is that in the firewall paradox, error correction breaks down such that two observers no longer arrive at the same bulk. But this does not render physics inconsistent - because while error correction was protecting observers from bulk divergence, in fully quantum side, discrepancies were there. Spacetime is thus expressed differently for some sets of observers, but physics is still consistent.

5.3 Entanglement equilibrium

The key clue toward the form of the theory of theories 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[20][21] differs from the approach of this writing in that this is considered solely an equilibrium behavior or holding only in first-order perturbation sense, which means we often only get linearized Einstein equations at the end, which are powerful enough for many purposes.

For Jacobson (2016)[20], it can be understood why the generalization to the principle of maximum entropy was not made, but then subsequent research lifted away many of restrictions of Jacobson (2016). Despite the name "entanglement equilibrium," how far this is applicable has become blurred.

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. Accepting the theory of theories idea, if not the form, it becomes clear that entanglement equilibrium literature, generalized sufficiently as in recent papers, provides a powerful evidence that the form of the theory of theories - centering around the principle of maximum entropy under constraints - is correct.

As stated in Cao-Carroll (2018)[21], 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 Cao-Carroll (2018)[21].

The major question, at this point, is really about how far applicable these results are. And that is unclear. If the theory of theories idea in this writing is correct, and some of spacetime-from-entanglement literature have general applicability far beyond AdS/CFT, then the question of laws of quantum gravity would be finished today. Of course having laws does not equate to full understanding of quantum gravity, as we would still have to seek for an initial state vector that generates our spacetime, but we would know what we must do to finish understanding of quantum physics.

One can do something like Cao-Carroll (2018), but recent works[22] try to generalize AdS/CFT correspondence into dS/dS correspondence[23]. This would allow us to generalize what we have learned in AdS/CFT into examples involving our actual spacetime.

5.4 Quantum decoherence

The key point behind spacetime-from-entanglement literature is redundancy. Why do we such a robust spacetime when quantum states seem so fragile? The answer is redundancy allowing for error correction.

Somewhat amazingly, this emphasis on redundancy is shared by quantum decoherence literature - see Zurek (1982)[24] - with a similar mechanism. But in quantum decoherence, this redundancy would have to be formed, rather than being already there. Just as what one can call a measurement arises after some quantum decoherence, one can say robust spacetime shared by different observers only arises after some redundancy formation has successfully formed. This does not contradict error-correcting code literature of AdS/CFT, and in fact should be expected. Quantum effects to spacetime are supposed to be volatile, and it is quantum decoherence that suppresses full effects of volatility and superposition.

We can combine aforementioned parallels in spacetime-from-entanglement literature into the following story: in machine learning, there are some series of error correction - or simply called data processing - to get information wanted. One can map this sequence of data processing to quantum decoherence over time as well. In this story, all aforementioned redundancy ideas are equivalent, allowing recovery of robust spacetime.

In terms of renormalization, one can imagine a typical quantum experiment and measurement/decoherence process as going from an initial fixed point (which maps to an initial Hamiltonian and state vector) that observers share to a final fixed point, which will not be reached in finite time, but observers would come close to such a fixed point quite fast, allowing for error correction to the shared end result. Learning paths - thus theories used between these fixed points - can be different depending an observer, and divergences can be significant. This is why it is hard to see quantum reality between measurements.

5.5 Evidences and clues for the theory of theories

Having created a case for equivalence of the ideas (neural network, basis redundancy, error-correcting code, quantum decoherence), the glaring missing piece again is a universal learning mechanism. This need again provides another evidence for necessity of the theory of theories explored in this writing.

Furthermore, entanglement equilibrium literature, which stands out from above equivalence despite being part of spacetime-from-entanglement literature, provides a direct clue for the form of the theory of theories.

We will also soon see the example in non-equilibrium classical thermodynamics that points toward the theory of theories idea as well - see subsection 6.2.

Another clue concerns a quantum decoherence process resulting in a clear measurement - the idea involves some entanglement that increases some entropy. Thus, why not go for the principle of maximum entropy? This of course is an aesthetic argument and may not constitute a scientific argument. But there may be something to trust about aesthetics in physics, given that physics in the past decades was motivated by beautiful symmetries.

6 Thermodynamics and Irreversibility

6.1 Do we have a secure foundation for thermodynamics?

This is somewhat the re-hash of subsection 4.5. The main question that thermodynamics often faces is that all physical processes in quantum and classical physics are technically reversible. An equilibrium state may naturally, without intervention, be reversed to a non-equilibrium state.

Proofs that irreversibility of thermodynamics is compatible with fundamental laws of classical or quantum physics have been provided - for now, I will focus on classical thermodynamics, later introducing quantum aspects, following Swendsen (2008)[25].

The logic behind irreversibility is that even under deterministic laws of classical physics, an observer only knows the system in question partially, requiring initial probabilistic distribution understanding of the system. If not interfered, final equilibrium probabilistic distribution, which is obtained with laws of classical physics, cannot be reversed in any imaginable means.

This works whenever initial ignorance of an observer at initial time can never be modified to obtain new probabilistic distributions over time. In practice, the above assumption can be said to hold approximately well. So the use of thermodynamics is more than justified. The problem is when thermodynamics is said to provide foundation for fundamental nature of irreversibility and the arrow of time, no longer remaining as a useful approximation to reality. But in quantum mechanics, initial ignorance has to change because observer Alice herself (and thus Alice's records) can change due to laws of physics, echoing the point made about quantum recoherence after quantum decoherence in subsection 4.5.

Thus a new law is required, if irreversibility is indeed fundamental, and theory of theories presented in this writing exactly fills in that role, because entropy of an observer would always be non-decreasing over time. This of course would matter only for those who believe in irreversibility being fundamental part of nature, but for those do, this would count as a significant evidence for the theory of theories.

6.2 The law of maximum entropy production

In non-equilibrium classical thermodynamics, Swenson and others[26] 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.

The principle of maximum entropy used in the theory of theories can be considered as a proper quantum generalization of LMEP, which serves as another evidence for the theory of theories.

6.3 Frauchinger-Renner thought experiment

But so far, we have not discussed why irreversibility matters so much. And that is left to the Frauchiger-Renner (FR) thought experiment[1] to demonstrate. The details of the thought experiment would have to be referred to the cited article, and I will focus on analysis.

First, that the FR thought experiment is useless because of quantum mechanics is unitary is not really a good argument. It is like saying the firewall paradox should not even be discussed because quantum mechanics is consistent. There is a value in understanding how things really fit into - sometimes revealing a need for something additional.

Second, while it is true that the part of the FR thought experiment is the "Wigner's friend" thought experiment, to call it Hardy's paradox Wigner's friend-ified would be a bit of stretch. This is because two labs are not actually entangled - one spin is simply transported via a quantum communication channel from one agent to another as to correlate results of the labs. This is the real novelty behind the FR thought experiment.

Third, it is the interpretation of Wigner's friend part that is most controversial about the FR thought experiment. A quantum state of Friend-plus-the-qubit-in-question (FQ) that Wigner sees can be stated as:

$$\frac{|F_0\rangle|0\rangle + |F_1\rangle 1\rangle}{\sqrt{2}}$$

Suppose that Wigner's friend observed $|0\rangle$ for the qubit. Having this external knowledge, we can say that Wigner will measure the combined system as $|F_0\rangle|0\rangle$ certainly. Or is it? That is, before Wigner measures anything, should we represent state vector of FQ after the measurement of the qubit by Wigner's friend as $|F_0\rangle|0\rangle$ ("collapse") or the entire state vector ("no collapse")? If former, then there is no paradox. It is the latter case that is complicated.

The traditional understanding of the no-collapse approach is that Wigner's probability of observing $|F_0\rangle|0\rangle$ is $1/2$ regardless of the measurement done by Wigner's friend. The measurement done by Wigner's friend is simply canceled as non-existent. But this is too strong. And even if there was no collapse, possibility that Wigner's probability of observing $|F_0\rangle|0\rangle$ is 1 externally speaking still exists. In language of many-worlds interpretation, while Wigner does not know that it is in the world branch involving $|0\rangle$, we do know that Wigner is in the branch. And Wigner can recognize that this external view exists. And in many-worlds interpretation, state vector never collapses.

From this alternative no-collapse view, the FR thought experiment comes to make sense along with its paradoxical conclusion. If one does not want to abandon any one of semiclassical validity of quantum communication channel, non-collapse completeness of quantum mechanics and effective identity of the theory governing any subsystem involved (or subsystems simply living in same Hilbert space), what one should give up instead is reversibility.

By the spirit of strong complementarity, it seems that it would be better to deny effective identity of the theory "used" by different subsystems. But the FR thought experiment is a low-energy experiment - thus the traditional way of thinking about quantum mechanics would work fine as well, subsystems well-approximately living in a shared Hilbert space, with appropriate basis transformations.

In order to measure some result, quantum decoherence literature suggests that environment, observer and measured subsystems are all required to be entangled. The environment part is what the FR thought experiment does not factor in. Thus, in order for the FR thought experiment to work, each lab must be disentangled to return the state of the lab into the original initial state vector. As said before, this does not mean that a previous measurement outcome is canceled, and many-worlds interpretation is helpful in thinking about this.

But can a lab really be reversed this way? The irreversibility result of the theory of theories says no, thus saving quantum mechanics from paradoxical consequences.

The FR thought experiment therefore can be re-interpreted as demonstrating fundamental importance of irreversibility, providing another evidence for the theory of theories.

7 Strong complementarity as a requirement of the theory of theories

7.1 Broken invariance of observables?

One potential criticism of strong complementarity, as presented in this writing, is that observables no longer stay invariant for an observer and covariant across different observers. Think of two different theories, related by basis transformation (applied to theories instead of basis). At $t = 0$, one can think of deriving an observable for one theory from the physically same observable of another theory. However, if an observable of one theory is time-invariant and correct, then an observable of another theory would not be time-invariant. This is because two theories are related by a different basis transformation at each time.

This seems to be a strong problem, but it is not actually a problem. The point is that an observer can be thought of as using each theory only once at maximum. Thus, observable attached to a theory is well-defined as far as a reference is clear.

In initial and final state theories, we expect observers agreeing on observations. Observables on those theories must be time-invariant, and there is no problem imposing this.

Essentially, what was explored before is basically a mirror operator idea in [6]. It is just that the idea was massively simplified for the purpose of this writing and cleaner analysis of potential objections. We now see a simple objection that can immediately arise and how this is resolved.

7.2 Satisfaction of postulates of black hole complementarity

Postulates of black hole complementarity[27] go as follow:

- Unitarity (purity) of quantum physics
- No drama: an infalling observer (Alice) crossing the horizon of a black hole notices nothing special. Thus usual spacetime for Alice.
- Effective field theory holds outside the horizon.
- Equality of Bekenstein-Hawking entropy with von Neumann entropy of a black hole, suitably defined.

The real question is whether the theory of theories satisfies the second postulate (no drama), because other postulates were already dealt in this writing. The AMPS paradox was that we should give up no drama, because two incompatible descriptions can be simultaneously accessed. But the theory of theories can be used to the AMPS account of a black hole in that Alice transitions to a new theory that invalidates entanglement relations of an old theory. Otherwise, the AMPS account remains the same - thus all postulates of black hole complementarity are satisfied.

8 Transition probability arithmetic: resolving the measurement problem completely

Moving from some outcome of a theory to a semi-outcome of another theory Because different theories are dealt together instead of a single theory, it must be possible to map outcome $|k\rangle$ at time t of one theory to some vector $|m\rangle$ in another theory at time t.

The idea is simple. Recall that in this writing, moving from one theory (Hamiltonian H) to another theory (Hamiltonian V) is thought as basis transformation. The relevant transformation factor is $Q = e^{-iVt}e^{iHt}$, where it is assumed that all theories share an initial state vector at $t = 0$.

$$|m\rangle = Q|k\rangle \tag{8.1}$$

$|k\rangle$ is a semi-outcome vector that may not be an outcome vector, as it is a weighted superposition of outcome vectors.

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

$$\begin{aligned} P_0(x_1)P_u(x_1 \to y_1) + \cdots + P_0(x_n)P_u(x_n \to y_1) &= P_{t_1}(y_1) \\ &\vdots \\ P_0(x_1)P_u(x_1 \to y_n) + \cdots + P_0(x_n)P_u(x_n \to y_n) &= P_{t_1}(x_n) \end{aligned} \tag{8.2}$$

where $P_t(x_1)$ refers to probability of semi-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 semi-outcome x_i at $t = 0$ to outcome

y_j at $t = t_1$. (Notice the restriction of y_j to an outcome.) 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. Semi-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 $\sum_j P_u(x_i \to x_j) = 1$. 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 8.2. Thus $n^2 - n$ equations need to be provided to determine transition probabilities. The ratio rule is specified as:

$$\begin{aligned}
P_u(x_1 \to y_1) : P_u(x_2 \to y_1) &= |Am(x_1 \to y_1)|^2 : |Am(x_2 \to y_1)|^2 \\
P_u(x_1 \to y_1) : P_u(x_3 \to y_1) &= |Am(x_1 \to y_1)|^2 : |Am(x_3 \to y_1)|^2 \\
&\vdots \\
P_u(x_1 \to y_1) : P_u(x_n \to y_1) &= |Am(x_1 \to y_1)|^2 : |Am(x_n \to y_1)|^2 \\
&\vdots \\
P_u(x_1 \to y_n) : P_u(x_2 \to y_n) &= |Am(x_1 \to y_n)|^2 : |Am(x_2 \to y_n)|^2 \\
&\vdots \\
P_u(x_1 \to y_n) : P_u(x_n \to y_n) &= |Am(x_1 \to y_n)|^2 : |Am(x_n \to y_n)|^2
\end{aligned} \tag{8.3}$$

where $Am(x_i \to y_j) = \langle y_j | e^{-iHt_1} | x_i \rangle$ refers to transition amplitude from semi-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.

9 Conclusion

Let me present the contents of this writing in a different order.

Transition probability arithmetic (TPA) Even without the collapse postulate, it makes sense to talk of transition probability of one outcome $|x_i\rangle$ of a state vector at time t to $|x_j\rangle$ of a state vector at time $t' > t$.

Quantum decoherence It provides a mechanism of how coherence between outcomes in a state vector can be lost such that each outcome can be considered completely separated from other outcomes. However, it still does not resolve basis ambiguity, despite providing some clues toward resolving it.

TPA+quantum decoherence Together, they provide a story on why it is possible to do quantum mechanics both on an entire system and its individual subsystems, without worrying about incompatibility of results, even when the collapse postulate is eliminated. After significant decoherence, understood to happen very quickly, effects of coherence on transition probability, given by TPA, become very close to zero. Thus all traditional quantum calculations are still consistent.

Elimination of subjectivity and omniscience While the Bayesian update interpretation of collapse can make sense, it is problematic in that an observer is left not modeled by physics. TPA+quantum decoherence eliminate subjectivity, with exception of basis ambiguity. The firewall paradox, analyzed by parallels with Hardy's paradox, demonstrates why basis ambiguity, combined with a subjectivist interpretation of quantum mechanics, can be very problematic. Thus for basis ambiguity, we also want to eliminate aspects of subjectivity.

Epistemic understanding of quantum mechanics (QM) For convenience, an epistemic interpretation of quantum mechanics is used. That is, a state vector in QM describes state of knowledge of a system. This interpretation makes it readily available to objective Bayesian inference tools.

QM from (objective) Bayesian inference consistency The point is that given an initial state vector and partition function constant Z, there is complete uncertainty as to how physical states would evolve, except constraints of invariance of Z - basically conservation of sum of probability - and Schrödinger equation of conserved Hamiltonian, which are reinterpreted as coming out of inference consistency requirements. The Born rule is also understood as a consequence of Bayesian inference consistency requirements.

The principle of maximum entropy of an observer rules in this complete uncertainty circumstance, as consistency axioms in objective Bayesian inference literature dictate. This is because this is essentially equivalent to setting a prior probability. An observer can only know outside by variations on her own states - thus entropy of an observer is the focus.

Hamiltonians that are learned can differ over time, despite the Hamiltonian of a system assumed to be conserved over time. Thus consistent Bayesian inference generates the theory of theories. An observer is effectively forced by objective Bayesianism to use a different description of the world over time in some circumstances. In fact, it is what drives examples like black hole complementarity, despite reality actually being same for any observer.

The standard collapse interpretation is readily accommodated in this re-interpretation of qauntum mechanics: it is about resetting an initial state vector and Z. Then inference is carried out until an observer gets to measure a subsystem, treated effectively as collapse of a quantum state. Post-measurement, the observer again resets initial conditions and so on. This, however, only works as an effective description - eventually, one has to worry about potential recoherence. Thus, completely accurate and objective Bayesian inference requires us not to reset the initial state vector of an entire system.

Understandably, this sounds paradoxical. Why can we not just use a final fixed point theory at every time? That is the theory that every observer is supposed to agree on. But again, we really do not know whether this final fixed point theory really is actual theory of reality. That learning converged does not mean it is the right one. But it does not matter, because every observer shares the same view of the world! I will revisit this issue when I discuss thermodynamics again.

Elimination of basis selection By this point, it would be clear that basis flexibility can safely be eliminated - a basis is fixed. Traditional basis transformation is translated to be theory transformation. This leads to a complaint that an observable is not invariant, which I mainly address by resorting to initial and end (not necessarily thermal) "equilibrium" theories, or fixed points in renormalization group terminology, and defining observables on them. But quantum mechanics is already a relativist theory, so this should not come as much of a surprise. In fact, observables can be replaced by relative understanding in a quantum state. This is just an abridged version of Papadodimas-Raju mirror operators.

Thermodynamics view While I initially rejected Swendsen's justification for a solid foundation of thermodynamics, eventually the view is, at core, correct. Recoherence is not completely eliminated but heavily restricted because of the principle of maximum entropy. It is ignorance of the entire system that drives thermodynamics and irreversibility. The Frauchinger-Renner thought experiment requires irreversibility to eliminate a paradox, providing an evidence for necessity of the Bayesian reformulation of quantum mechanics. Non-equilibrium thermodynamics also supports the reformulation as well, but this is largely because the law of maximum entropy was already demonstrated to be powerful even without objective Bayesian foundations from consistency requirements.

The main key point of Swendsen's argument is that there can be many compatible probabilistic distributions for a system given - after all, it is just about ignorance! An underlying system is deterministic there anyway too. Transporting this into QM, it is now possible that final fixed point theories are not actually equivalent to a theory governing reality. But again, it does not matter, just as classical thermodynamics and classical physics work simultaneously.

Quantum gravity and spacetime It is often said that we are far away from an operational quantum theory of gravity consistent with reality. We may actually be far closer - in fact we may have completed one without us knowing for sure. Even so, identifying an initial asymptotic state of the entire universe, which is required to start inference of theories governing the universe, is a difficult job, and we may still be left without an actual form of the theory of everything. We know laws that govern our universe but not theories.

In section 5, spacetime-from-entanglement literature is shown to provide evidences for the reformulation of quantum mechanics. Here, I will not discuss details again, as they are not at this point important.

Future works This work provides how QM may be linked to questions of cosmology. So it seems natural to port knowledge in cosmology (or more generally astrophysics) back to QM with hopes of extracting the right inferred theory out of obtained initial conditions, mostly coming directly out of observations.

New cross-field works previously unimaginable are made possible by this reformulation of QM, and a new exciting chapter of physics is in front of us.

References

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

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

[3] R. Bousso, *Complementarity is not enough*, *Phys. Rev. D* **87** (2013) 124023 [1207.5192].

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

[5] K. Papadodimas and S. Raju, *An infalling observer in AdS/CFT*, *Journal of High Energy Physics* **2013** (2013) 212 [1211.6767].

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

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

[8] D. Harlow, *Aspects of the Papadodimas-Raju proposal for the black hole interior*, *Journal of High Energy Physics* **2014** (2014) 55 [1405.1995].

[9] D. Marolf and J. Polchinski, *Violations of the Born rule in cool state-dependent horizons*, *Journal of High Energy Physics* **2016** (2016) 8 [1506.01337].

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

[11] W. H. Zurek, *Decoherence and the transition from quantum to classical – revisited*, quant-ph/0306072.

[12] 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.

[13] 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.

[14] 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.

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

[16] 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].

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

[18] 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].

[19] 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].

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

[21] 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].

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

[23] 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.

[24] W. H. Zurek, *Environment-induced superselection rules*, Phys. Rev. D **26** (1982) 1862.

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

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

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