Born in an infinite universe: A cosmological interpretation of quantum mechanics

Anthony Aguirre and Max Tegmark

Born in an Infinite Universe: a Cosmological Interpretation of Quantum Mechanics

Anthony Aguirre

Dept. of Physics & SCIPP, University of California at Santa Cruz, Santa Cruz, CA 95064; aguirre@scipp.ucsc.edu

Max Tegmark

Dept. of Physics & MIT Kavli Institute, Massachusetts Institute of Technology, Cambridge, MA 02139; tegmark@mit.edu

We study the quantum measurement problem in the context of an infinite, statistically uniform space, as could be generated by eternal inflation. It has recently been argued that when identical copies of a quantum measurement system exist, the standard projection operators and Born rule method for calculating probabilities must be supplemented by estimates of relative frequencies of observers. We argue that an infinite space actually renders the Born rule *redundant*, by physically realizing all outcomes of a quantum measurement in different regions, with relative frequencies given by the square of the wave function amplitudes. Our formal argument hinges on properties of what we term the quantum confusion operator, which projects onto the Hilbert subspace where the Born rule fails, and we comment on its relation to the oft-discussed quantum frequency operator. This analysis unifies the classical and quantum levels of parallel universes that have been discussed in the literature, and has implications for several issues in quantum measurement theory. Replacing the standard hypothetical ensemble of measurements repeated *ad infinitum* by a concrete decohered spatial collection of experiments carried out in different distant regions of space provides a natural context for a statistical interpretation of quantum mechanics. It also shows how, even for a single measurement, probabilities may be interpreted as relative frequencies in unitary (Everettian) quantum mechanics. We also argue that after discarding a zero-norm part of the wavefunction, the remainder consists of a superposition of indistinguishable terms, so that arguably "collapse" of the wavefunction is irrelevant, and the "many worlds" of Everett's interpretation are unified into one. Finally, the analysis suggests a "cosmological interpretation" of quantum theory in which the wave function describes the actual spatial collection of identical quantum systems, and quantum uncertainty is attributable to the observer's inability to self-locate in this collection.

I. INTRODUCTION

Although quantum mechanics is arguably the most successful physical theory ever invented, the century-old debate about how it fits into a coherent picture of the physical world shows no sign of abating. Proposed responses to the so-called measurement problem (e.g., [1]) include the Ensemble [2, 3], Copenhagen [4, 5], Instrumental [6–9], Hydrodynamic [10], Consciousness [12–15], Bohm [16], Quantum Logic [17], Many-Worlds [18, 19], Stochastic Mechanics [20], Many-Minds [21, 22], Consistent Histories [23], Objective Collapse [24], Transactional [25], Modal [26], Existential [27], Relational [28], and Montevideo [29] interpretations. Moreover, different proponents of a particular interpretation often disagree about its detailed definition. Indeed, there is not even consensus on which ones should be called interpretations.

While it is tempting to relegate the measurement problem to "philosophical" irrelevancy, the debate has both informed and been affected by other developments in physics, such as the understanding of the importance [30] and mechanisms [21, 31–33] of decoherence, experimental probes of quantum phenomena on ever larger scales, and the development of cosmology.

Putting quantum mechanics in the context of cosmology creates new issues. How can we apply the measurement postulate at times before measuring apparatuses – or, in some interpretations consciousness – existed? How can we split the universe into a system and measuring apparatus if the system is the entire universe itself? And cosmology may even be intertwined with quantum measurements here-and-now. In an infinite universe, our *observable* universe (the spherical region from which light has had time to reach us during the past 14 billion years) may be just one of many similar regions, and even one of many *exact* copies – and in fact many versions of the widely-accepted standard inflationary cosmology provide just such a context [34], as we shall discuss. Don Page has recently argued that pure quantum theory and the textbook Born rule cannot produce outcome probabilities for multiple replicas of a quantum experiment [35–37], and must therefore be augmented in a cosmological scenario in which such replicas exist.

It is therefore timely to further investigate how such a cosmology impacts the quantum measurement problem, and this is the aim of the present paper. We will argue that cosmology is not merely part of the problem, but also part of the solution. Page essentially showed that to calculate the probability that a measurement in a collection of N identical experiments has a given outcome, one must supplement the standard Born rule for assigning probabilities to measurement outcomes, based on projection operators, by a rule that assigns probabilities to each of the events "the experiment was performed by the kth observer, $k = 1...N$". We will go further, and argue that

if one identifies probabilities with the relative frequencies of experimental outcomes in three-dimensional space, the measurement postulate (Born's rule) becomes superfluous, as it emerges directly from the quantum Hilbert space formalism and the equivalence of all members of an infinite collection of exact replicas. We will see that this is intimately linked to classic frequency operator results [38–42], except that a fictitious infinite sequence of identical measurements is replaced by an actually existing spatial collection. It is also closely related to arguments[1] by [43, 45].

This potential cosmological connection between probabilities in quantum mechanics and the relative frequencies of actual observers is relevant to most of the above-mentioned quantum interpretations. It is particularly interesting for Everett's Many Worlds Interpretation (MWI), where we will argue that it eliminates the perplexing feature that, loosely speaking, some observers are more equal than others. To wit, suppose a spin measurement that should yield "up" with probability $p = 0.5$ is repeated $N = 10$ times. According to the MWI, the final wavefunction has $2^N = 1024$ terms, each corresponding to an equally real observer, most of whom have measured a random-looking sequence of ups and downs. This suggests that quantum probabilities can be given a simple frequentist interpretation. However, for an unequal probability case such as $p = 0.001$, the final wavefunction still has 2^N terms corresponding to real observers, but now most of them have measured approximately 50% spin up and concluded that the Born rule is incorrect. (We are supposed to believe that everything is still somehow consistent because the observers with a smaller wave function amplitude are somehow "less real".) We will show that in an infinite inflationary space, probabilities can be given a frequentist interpretation even in this case.

The rest of this paper is organized as follows. In Section II, we describe the cosmological context in which quantum mechanics has found itself. We then investigate how this yields a forced marriage between quantum probabilities and relative frequencies, in both a finite space (Section III) and an infinite space (Section IV). Rather than launching into an intimidating mathematical formalism for handling the most general case, we begin with a very simple explicit example, then return to the rather unilluminating issue of how to generalize the result in Appendix B. In Section V we describe how to formally describe measurement in this context, then discuss possible interpretation of our mathematical results in Section VI. We discuss some open issues in Section VII, and summarize our conclusions in Section VIII.

II. THE COSMOLOGICAL CONTEXT

When first applying General Relativity to our universe, Einstein assumed the Cosmological Principle (CP): our universe admits a description in which its large-scale properties do not select a preferred position or direction. This principle has served cosmology well, supplying the basis for the open, flat, and closed universe metrics that underly the highly successful Friedmann-Lemaître-Robertson-Walker (FLRW) Big-Bang cosmology. We shall argue that this principle and the interpretation of QM may be closely intertwined, with the theory of cosmological inflation as a central player. In particular, we will discuss how *eternal* inflation naturally leads to a universe obeying a strong version of the CP [43, 44], in which space is infinite and has *statistically uniform*[2] properties. In this context, any given finite region is replicated throughout the infinite space, which in turn requires a re-appraisal of quantum probabilities.

A. The Cosmological Principle and infinite spaces

In a finite space, the CP has a curious status: with a single realization of a finite space, there is no meaningful way for the statistical properties to be uniform. There would, for example, always be a unique point of highest density. We could compare our realization to a hypothetical ensemble of universes generated assuming a set of uniform statistical properties, but we could never *recover* these putative statistical properties beyond a certain degree of precision. In this sense the CP in a finite space is really nothing more than an assumption (as by Einstein) that space and its contents are "more or less" homogeneous on large scales; a precise description would require the specification an enormous amount of information.

---

[1] In brief, these works argue that a complete description of the universe does not single out a particular place. So instead of describing what happens "here" it describes an ensemble (in Gibbs' sense) of identical experiments uniformly scattered throughout an infinite (expanding) space (a cosmological ensemble). It follows from a description of the measuring process in which the measuring apparatus is assigned a definite macrostate but not a definite microstate that the measurement outcomes and their relative frequencies in the cosmological ensemble coincide with those given by the measurement postulate. (A similar idea was independently put forward in schematic form by [46].)

[2] The matter distribution in space is customarily modeled as evolved from some random initial conditions, so properties in any given region must be described statistically, e.g., by the mean density, 2-point and higher correlation functions, etc. By *statistically uniform*, we mean that all such statistics are translationally invariant. (For example, the homogeneous and isotropic Gaussian random fields generated by inflation – and anything evolved from such initial conditions – are statistically uniform in this sense.) Another way to look at this is that the probability distribution for different realizations in a given region of space is the same as the distribution across different spatial regions in a given realization. (See [45] for an extended discussion of this point and its implications.)

An infinite space is quite different: by examining arbitrarily large scales, its statistical properties can in principle be assessed to arbitrarily high accuracy about any point, so there is a precise sense in which the properties can be uniform. Moreover, if (as the holographic principle suggests) a region of some finite size and energy can only take on a fixed finite set of possible configurations, then the full specification of a statistically-uniform infinite space would require *only* those statistics. This implies [43, 44] that in contrast to a finite system, there would be only one possible realization of such a system, as any two systems with the same statistical properties would be indistinguishable.

The CP might be taken as postulated symmetry properties of space and its contents, consistent with the near uniformity of our observed universe. In an infinite (open or flat) FLRW universe, these symmetries can be *exact* in the above sense, and such a postulated cosmology would support the arguments of this paper beginning in Section II D, or those of [45].

Alternatively, we might search for some physical *explanation* for the near-uniformity of our observable universe. This was a prime motivation for cosmological inflation. Yet inflation can do far more than create a large uniform region: in generic models inflation does, in fact, create an *infinite* uniform space.

B. Infinite spaces produced by eternal inflation

Inflation was devised ([47]; see [48] for some history) as a way to grow a finite-size region into an extremely large one with nearly uniform properties, and if inflation is realized in some region, it does this effectively: the exponential expansion that inflates the volume also dilutes or stretches into near homogeneity any particles or fields within the original region. The post-inflationary properties are then primarily determined not by cosmic initial conditions, but by the *dynamics* of inflation, which are uniform across the region; although particular initial conditions are required for such inflation to arise, once it does, information about the initial conditions is largely inflated away.

It was soon discovered, however, that in generic models, inflation is *eternal*: although inflation eventually ends with probability unity at any given location, the exponential expansion ensures that the total inflating volume always increases exponentially (see [48–50] for recent reviews.) In many cases, one may think of this as a competition between the exponential expansion $\exp(Ht)$, and the "decay" from inflation to non-inflation with characteristic time t_{decay}. This means that an initial inflating volume V has, at some later time, inflating volume $\sim V \exp(3Ht) \exp(-t/t_{\text{decay}}) = V \exp[(3H - t_{\text{decay}}^{-1})t]$; for inflation to work at all requires the expansion to win for a number of e-foldings, implying a positive exponent;

but in this case expansion will tend to win forever.[3] The result is that eternal inflation does provide post-inflationary regions with the requisite properties, but as part of an ultimately infinite spacetime.

It might seem that a given post-inflationary region is necessarily finite, because no matter how long inflation goes on, it can only expand a given finite initial region into a much larger yet still finite space. But this is not the case. General Relativity forbids any fundamental choice of time variable, but there is a physically preferred choice, which is to equate equal-time surfaces with surfaces of constant inflaton field value (and hence constant energy density), so that the end of inflation occurs at a single time. In eternal inflation, this choice leads to multiple disconnected surfaces on which inflation ends, *each one* generally being both infinite and statistically uniform. Likewise, in each region and in these coordinates, the ensuing cosmic evolution occurs homogeneously.[4]

This occurs in all three basic types of eternal inflation: "open" inflation (involving quantum tunneling, and driven by an inflaton potential with multiple minima), in "topological" inflation (driven be an inflaton field stuck around a maximum in its potential), and "stochastic" inflation (in which upward quantum fluctuations of the field can overwhelm the classical evolution of the field toward smaller potential values). These three particular scenarios are discussed in more detail in Appendix A, where we also provide heuristic arguments as to why infinite, statistically uniform spaces are a generic product of eternal inflation, by its very nature.[5]

Thus eternal inflation, if it occurs, provides a causal mechanism for creating a space (or set of spaces) obeying a form of the CP, in the sense that each space is infinite

————

[3] This is not to say that *every* inflation model has eternal behavior: it is not hard to devise non-eternal versions; but the need to do so deliberately in most cases suggests that eternal behavior is more generic. (An exception is hybrid inflation, which is generically non-eternal [51]; such models however tend to predict a scalar spectral index $n > 1$ [52], which is in some conflict with current constraints [53].) In scenarios where inflation might take place in parallel in different parts of a complicated potential energy "landscape", regions of the landscape with eternal inflation will naturally outcompete those with non-eternal inflation, predicting by almost any measure that the region of space we inhabit was generated by eternal inflation. On the other hand, it has been argued that inflation eternal inflation ([55, 56]) and perhaps even inflation (e.g. [54]) may be difficult to realize in a landscape that is generated as a low-energy effective potential from a true high-energy quantum gravity theory.

[4] Moreover, a given point on the spatial surface at which inflation ends will occur an enormously or infinitely long duration after any putative initial conditions for inflation. Thus, insofar as inflation makes these initial conditions irrelevant, they are arguably *completely* irrelevant in eternal inflation.

[5] In a cosmology with a fundamental positive cosmological constant, this issue becomes more subtle, as some arguments suggest such a cosmology should be considered as having a finite total number of degrees of freedom (see, e.g. [57]). How this can be understood consistently with the semi-classical spacetime structure of eternal inflation is an open issue.

and has uniform properties determined on average by the classical evolution of the inflaton, with statistical variations provided by the quantum fluctuations of the field during inflation.

C. Infinite statistically uniform space, and probabilities, from inflation

The fact that post-inflationary spacetime is infinite in eternal inflation leads to some rather vexing problems, including the "measure problem" of how to count relative numbers of objects so that statistical predictions for the cosmic properties surrounding those objects can be made (see, e.g., [58] for a recent review.) This paper is not an attempt to solve that problem. In particular, we do not address the comparison of observer numbers across regions with a different inflationary history and hence different gross properties. Rather, we will ask about what happens when we apply the formalism of quantum theory to a system in the context of a single infinite space with uniform and randomly-determined statistical properties.[6]

One of the greatest successes of cosmological inflation is that small-scale quantum fluctuations required by the Heisenberg uncertainty relation get stretched with the expanding space, then amplified via gravitational instability into cosmological large-scale structure just like that we observe in, *e.g.*, the galaxy distribution and in the cosmic microwave background [53, 59]. In a given finite cosmic region, this process creates pattern of density fluctuations representing a single realization of a statistical process with a probability distribution governed by the dynamics of inflation and the behavior of quantum fields within the inflating space.

Eternal inflation also creates infinitely many other nearly-homogeneous regions with density fluctuations drawn from the same distribution (because the dynamics are just the same) that evolve independently of each other (because the regions are outside of causal contact if they are sufficiently widely separated, where "widely" means being farther apart than the horizon scale during inflation, say 10^{-24} m). The resulting space then has statistically uniform properties, and the probability distribution governing the fluctuations in any single region is recapitulated as the relative frequencies of these fluctuation patterns across the actually-existing spatial collection of regions.

Now, these inflationary fluctuations constitute the classical cosmological "initial" conditions that determine the large-scale variation of material density and thus, e.g., the distribution of galaxies. Smaller-scale details

of the current matter distribution (such as what you ate for breakfast) were determined by these same inflationary initial conditions, augmented by subsequent quantum fluctuations amplified by chaotic dynamics, etc. Because such small-scale processes (and any microscopic "initial" conditions connected with them) are decoupled from the super-horizon large-scale dynamics giving rise to the infinite space, the overall space should again be statistically uniform, here in the sense that the probability distribution of microstates in each finite region depends *not* on its location in space, but only on its macroscopic properties, which are themselves drawn randomly from a region-independent statistical distribution.

In short, inflation creates an infinite set of cosmic regions, each with "initial conditions" and subsequently-evolving properties that are characterized (and *only* characterized) by a statistical distribution that is independent of the choice of region.

D. Quantum mechanics and replicas

Let us now make our link to everyday quantum mechanics. For a simple example that we shall follow throughout this paper, consider a spin 1/2 particle and a Stern-Gerlach experiment for measuring the z-component of its spin, which has been prepared in the state $\psi = \alpha | \downarrow \rangle + \beta | \uparrow \rangle$. Here α and β are complex numbers satisfying the usual normalization condition $|\alpha|^2 + |\beta|^2 = 1$. If we assume that a finite volume region with a roughly flat background metric has a finite set of possible microscopic configurations[7] (as suggested by, e.g., the holographic principle and other ideas in quantum gravity), and that our system plus experimenter configuration evolved from one of finitely many possible sets of initial conditions drawn from the distribution governing the statistically uniform space at some early time, then it follows that this configuration must be replicated elsewhere.[8] That is, there are infinitely many places in this space where an indistinguishable experimenter has prepared the same experiment using a classically indistinguishable procedure, and therefore uses the same α

---

[6] While it is our assumption for present purposes, it is not a given that these questions are inseparable, as some "global" measures would also "induce" a measure over the otherwise uniform subspaces we are considering.

---

[7] Meaning a finite number of meaningfully distinct ways in which the state can be specified. Note that although the real number coefficient α would seem to allow an uncountably infinite set of specifications, this is misleading: the maximum von Neumann entropy $S = -\mathrm{Tr}\rho \log \rho$ for our system is just $\log 2$, and at most two classical bits of information can be communicated using a single qubit. We should note, however, that while our assumption is quite standard, the precise way in which the continuous α would over-specify the state is a subtle question that we do not address here.

[8] This does assume some additional subtleties. For example it is argued in [43] that "statistical predictions do not prescribe all the properties of infinite collections. ...Any outcome that occurs a finite number of times has zero probability." In particular, an outcome that is consistent with physical laws could in principle occur in only one observable universe.

and β to describe the initial wavefunction of her particle. The rather conservative estimate in [63] suggests that the nearest indistinguishable copy of our entire observable universe ("Hubble volume") is no more than $10^{10^{115}}$ meters away, and the nearest subjectively indistinguishable experimenter is likely to be much closer.[9] We will now argue that the quantum description of this infinite set of systems sheds light on the origin of probabilities in quantum mechanics.

III. PROBABILITIES FOR MEASUREMENT OUTCOMES IN A FINITE REGION

A. The problem

In a statistically uniform space, consider a finite region that is large enough to contain N identical copies of our Stern-Gerlach experiment prepared in the simple above-mentioned state.

The state of this combined N-particle system is simply a tensor product with N terms. For example, $N = 3$ gives the state

$$|\psi\rangle = (\alpha|\downarrow\rangle + \beta|\uparrow\rangle) \otimes (\alpha|\downarrow\rangle + \beta|\uparrow\rangle) \otimes (\alpha|\downarrow\rangle + \beta|\uparrow\rangle) =$$
$$= \alpha^3|\downarrow\downarrow\downarrow\rangle + \alpha^2\beta|\downarrow\downarrow\uparrow\rangle + ... + \beta^3|\uparrow\uparrow\uparrow\rangle. \quad (1)$$

If we order the 2^N basis vectors of this 2^N-dimensional Hilbert space by increasing number of "up" vectors, the vector of wavefunction coefficients takes the simple form

$$\begin{pmatrix} \langle\downarrow\downarrow\downarrow\,|\psi\rangle \\ \langle\downarrow\downarrow\uparrow\,|\psi\rangle \\ \langle\downarrow\uparrow\downarrow\,|\psi\rangle \\ \langle\uparrow\downarrow\downarrow\,|\psi\rangle \\ \langle\uparrow\uparrow\uparrow\,|\psi\rangle \\ \langle\uparrow\uparrow\downarrow\,|\psi\rangle \\ \langle\uparrow\downarrow\uparrow\,|\psi\rangle \\ \langle\uparrow\uparrow\uparrow\,|\psi\rangle \end{pmatrix} = \begin{pmatrix} \alpha^3 \\ \alpha^2\beta \\ \alpha^2\beta \\ \alpha^2\beta \\ \alpha\beta^2 \\ \alpha\beta^2 \\ \alpha\beta^2 \\ \beta^3 \end{pmatrix} \quad (2)$$

for our $N = 3$ example. For general N, there are $\binom{N}{n}$ terms with n spins up, each with coefficient $\alpha^{N-n}\beta^n$.

B. Probabilities in the finite region

Suppose we would like to ask the core question: "given that I have prepared the quantum system as described, what is the probability that I will measure $\uparrow$?" This is more subtle than it would appear, because "I" might be part of any one of the N indistinguishable experimental setups assumed. As argued by Page [35–37], if one wants to consider these N experiments as a single quantum subsystem of the universe, this is problematic because in this situation there is no set of projection operators that can assign outcome probabilities purely via the Born rule.[10] Thus it seems that the quantum formalism by itself is insufficient, and must be in some way supplemented by additional ingredients.

While this conflicts with the idea that quantum theory alone should suffice when applied to the whole universe, we can recover the usual Born rule results for the single system in a fairly straightforward way if we augment the Born rule as applied to the product state for the N systems with probabilities assigned according to relative frequencies among the N systems.[11] In particular, since the 2^N terms are orthogonal (being a basis for the tensor-product state space) we might in principle imagine measuring the whole system, and attribute a quantum probability to each term given by its squared amplitude. Yet even if just one of these terms is "realized",[12] there is still uncertainty as to which spin is measured, because there is complete symmetry between the N indistinguishable measuring apparatuses. You should thus accord a probability for $\uparrow$ given by the relative frequencies of $\uparrow$ and $\downarrow$.

The total probability $P_\uparrow$ of measuring $\uparrow$ would then come from a combination of quantum probabilities and frequentist estimates of probability, using $P(A) = \sum_i P(B_i)P(A|B_i)$ where $P(A|B_i)$ is the conditional probability of A given B_i. Thus

$$P_\uparrow = \sum_{n=0}^{N} \binom{N}{n} (\beta^*\beta)^n (\alpha^*\alpha)^{(N-n)} \frac{n}{N}. \quad (3)$$

Each term in this sum is just the binomial coefficient $f(n; N, p)$ with $p = \beta^*\beta$ (the quantum probability from equation (2) of getting n $\uparrow$-factors) times n/N (the probability that among the N identical observers, you are one of the n who observed $\uparrow$). Mathematically, this sum simply computes $1/N$ times the mean of the binomial distribution, which is Np, giving $P_\uparrow = p$. In this way,

---

[9] The frequency of such repetitions depends on very poorly understood questions such as the probability for certain types of life to evolve, etc.

[10] In particular, define $\hat{P}_{i,\uparrow}$ and $\hat{P}_{i,\downarrow}$ to be operators that project onto $|\uparrow\rangle$ and $|\downarrow\rangle$ for the ith observer (leaving the $N-1$ other components of the product vector unchanged.) Then Page [37] shows that for $N = 2$, there is no state-independent projection operator $\hat{P}_\uparrow$ that gives Born-rule probabilities $P(\uparrow) = \langle\hat{P}_\uparrow\rangle$ for measuring $\uparrow$ (absent information about which particular observer one is) that are a weighted combination (with positive weights) of the probabilities $P_{i,\uparrow}$ for measuring $|\uparrow\rangle$ for each given *known* observer i.

[11] This is essentially 'T5' suggested in [35] as one possible way to restore probabilities.

[12] Readers preferring the Everettian perspective can accord a probability to each of these terms as branches of the wavefunction, and make a similar argument. Note, however, that it is somewhat less satisfying because if we add up the relative frequency of observers across *all* branches, it is by symmetry 50% for $\uparrow$ and 50% for $\downarrow$; this is the uncomfortable issue of some observers being more real than others noted in the introduction.

the standard Born rule probability $\beta^*\beta$ to measure $\uparrow$ is recovered, using a *combination* of the Born rule applied to the 2^N-state superposition, and the relative frequencies of $|\uparrow\rangle$ and $|\downarrow\rangle$ contained in each (product) state in that superposition.

C. Frequency and confusion operators

If the reality of indistinguishable systems has forced us to augment quantum probabilities with probabilities based on observer frequencies, as above, it is very interesting to examine more carefully how these two notions of probability connect. To do so, let us now define two Hermitean operators on this Hilbert space, both of which are diagonal in this basis. The first is the *frequency operator* $\widehat{F}$ introduced by [38, 39, 41, 42], which multiplies each basis vector by the fraction of the arrows in its symbol that point up; for our $N = 3$ example,

$$\widehat{F} = \frac{1}{3}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix} \tag{4}$$

The second is the *confusion operator* $\widehat{\odot}$, which projects onto those basis vectors where the spin-up fraction differs by more than a small predetermined value ϵ from the Born rule prediction $p = |\beta^2|$; for our $N = 3$ example with any $\epsilon < 1/3$,

$$\widehat{\odot} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{5}$$

More generally, the two operators are clearly related by

$$\widehat{\odot} = \theta\left(|\widehat{F} - p| - \epsilon\right), \tag{6}$$

where θ denotes the Heaviside step function, *i.e.*, $\theta(x) = 1$ for $x \geq 0$, vanishing otherwise.

Since both $\widehat{F}$ and $\widehat{\odot}$ are Hermitean operators, one would conventionally interpret them as observables, with $\widehat{F}$ measuring the fraction of spins that are up, and $\widehat{\odot}$ measuring 1 if this fraction differs by more than ϵ from p, 0 otherwise.

Examining the norm of the state multiplied by $(\widehat{F} - p)$,

we find that [39]

$$\begin{aligned} \|(\widehat{F} - p)|\psi\rangle\|^2 &= \langle\psi|(\widehat{F} - p)^2|\psi\rangle = \\ &= \sum_{n=0}^{N} \binom{N}{n}(1-p)^n p^{N-n}\left(\frac{n}{N} - p\right)^2 = \\ &= \frac{p(1-p)}{N}. \end{aligned} \tag{7}$$

Note that as in equation (3), α and β enter only in the combinations $\alpha^*\alpha = (1 - p)$ and $\beta^*\beta = p$ (because $\widehat{F}$ is diagonal), and just as equation (3) is the mean of a binomial distribution (divided by N), here the second line is simply the variance of a binomial distribution (divided by N^2).

As for the confusion operator, using equation (2) and equation (5) we obtain

$$\begin{aligned} \|\widehat{\odot}|\psi\rangle\|^2 &= \langle\psi|\widehat{\odot}|\psi\rangle = \\ &= \sum_{n=0}^{N}\binom{N}{n}(1-p)^n p^{N-n}\theta\left(\left|\frac{n}{N} - p\right| - \epsilon\right) \\ &= \sum_{|n-Np|>N\epsilon}\binom{N}{n}(1-p)^n p^{N-n} \\ &\leq 2e^{-2\epsilon^2 N}, \end{aligned} \tag{8}$$

where θ again denotes the Heaviside step function, and we have used Hoeffding's inequality[13] in the last step. Thus $\|\widehat{\odot}|\psi\rangle\|$ is exponentially small if $N \gg \epsilon^{-2}$; this mathematical result will prove to be important below.

For large enough N, our rescaled Binomial distribution approaches a Gaussian with mean p and standard deviation $\sqrt{p(1-p)/N}$, so

$$\|\widehat{\odot}|\psi\rangle\|^2 \approx \mathrm{erfc}\left[\left(\frac{N}{2p(1-p)}\right)^{1/2}\epsilon\right]. \tag{9}$$

(erfc denotes the complementary error function, *i.e.*, the area in the Gaussian tails).

In summary, for finite N, the product state of our N copies represents a sum of many terms. In some of them, the relative frequencies of up and down states in individual members of the spatial collection closely approximate the corresponding probabilities given by Borns rule. In other terms the frequencies differ greatly from the corresponding Born-rule probabilities. However, as N increases, these confusing states contribute smaller and smaller amplitude, as measured by the quickly diminishing expectation value of $\widehat{\odot}$.

---

[13] Specifically, apply theorem 2 of [11], representing the binomial distribution as the sum of N independent Bernoulli distributions.
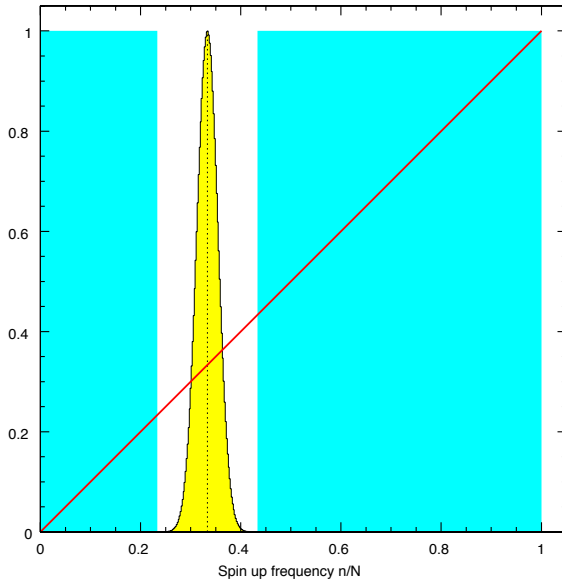
tor $\widehat{F}$ case, since the spectrum of the former is constant (zero) near the limit point.)

Third, let us define the complement of the confusion operator, which projects onto the states where the up fraction is within ϵ of the Born rule prediction p:

$$\widehat{\ddot{\smile}} \equiv \widehat{I} - \widehat{\ddot{\frown}}, \qquad (11)$$

where $\widehat{I}$ is the identity operator. Both $\widehat{\ddot{\frown}}$ and $\widehat{\ddot{\smile}}$ are projection operators, and they by definition satisfy the relations $\widehat{\ddot{\frown}}^2 = \widehat{\ddot{\frown}}$, $\widehat{\ddot{\smile}}^2 = \widehat{\ddot{\smile}}$, and $\widehat{\ddot{\frown}}\widehat{\ddot{\smile}} = \widehat{\ddot{\smile}}\widehat{\ddot{\frown}} = 0$.

Now let us decompose our original state into two orthogonal components:

$$|\psi\rangle = \widehat{\ddot{\frown}}|\psi\rangle + (\widehat{I} - \widehat{\ddot{\frown}})|\psi\rangle = \widehat{\ddot{\frown}}|\psi\rangle + \widehat{\ddot{\smile}}|\psi\rangle. \qquad (12)$$

Substituting equation (10) now implies the important result that

$$|\psi\rangle \to \widehat{\ddot{\smile}}|\psi\rangle \quad \text{as} \quad N \to \infty, \qquad (13)$$

with convergence in the sense that the correction term approaches zero Hilbert space norm as $N \to \infty$. But the right-hand-side $\widehat{\ddot{\smile}}|\psi\rangle$ is a state which by definition is a superposition only of states where the relative frequencies of $|\uparrow\rangle$ and $|\downarrow\rangle$ are in precise accord with Born-rule probabilities, with an up-fraction as close to $p = |\beta|^2$ as we chose to require with our ϵ-parameter. If we now assume that the outcome of a single quantum measurement is one of the measured observables eigenvalues, almost all components of the superposition of pre-measurement states weve been discussing represent post-measurement states in which the relative frequencies of eigenstates are equal to the corresponding Born probabilities for the possible outcomes of a single measurement. (Measurement and decoherence are discussed in the next section.)

Finally, as $N \to \infty$, $|\psi\rangle$ approaches (again in the sense that the correction term approaches zero Hilbert space norm) a state where every element in the grand superposition becomes statistically indistinguishable from the others, in the following sense. Suppose, as above, that each term in the grand superposition is taken to represent a collection of identical apparatuses that have each registered a definite outcome corresponding to either $|\uparrow\rangle$ or $|\downarrow\rangle$. Consider a volume of enormous but finite radius R that contains some large number M of our identical experiments, order them any way you like and write down their readings as a sequence such as ... $\uparrow\uparrow\downarrow\downarrow\uparrow\downarrow$ ..., and let $n_\downarrow$ and $n_\uparrow$ denote the number of readings of $\downarrow$ and $\uparrow$, respectively ($n_\downarrow = M - n_\uparrow$). In our infinite space, just as there are infinitely many realizations of a single experiment, there will be infinitely many such spherical regions, in each of which one of the 2^M possible outcomes is realized. Across this collection, these outcomes will be realized with frequencies that are within ϵ of the Born rule predictions $(1-p)^{M-n_\uparrow}p^{n_\uparrow}$, except for a correction term with zero Hilbert space norm. This follows from

FIG. 1: As a function of the spin up fraction n/N, the figure shows the scaled binomial distribution for $(n,p) = (500, 1/3)$ (shaded curve) together with the spectra of the frequency operator (diagonal line) and the $\epsilon = 0.1$ confusion operator (shaded, equal to 1 for $|n/N - p| > \epsilon$).

IV. PROBABILITES IN AN INFINITE SPACE

As $N \to \infty$, four key things happen.

First, the binomial distribution governing the spread in values of the frequency operator in Eq. 7 approaches a δ-function centered about p, so that $\|(\widehat{F} - p)|\psi\rangle\|^2 = p(1-p)/N \to 0$.

Second, the norm of the state projected by $\widehat{\ddot{\frown}}$ vanishes,

$$\|\widehat{\ddot{\frown}}|\psi\rangle\|^2 \to 0, \qquad (10)$$

with exponentially rapid convergence as per equation (6). This is illustrated in Fig. 1: in the equation (8) sum, the θ-term vanishes outside the shaded area, whereas the remaining binomial distribution factor (the plotted curve) gets ever narrower as $N \to \infty$, eventually ending up almost entirely inside the region where $\theta = 0$. (That $|\psi\rangle$ is an eigenvector of $\widehat{\ddot{\frown}}$ also follows directly from $|\psi\rangle$ being an eigenvector of $\widehat{F}$ and the fact that $\widehat{\ddot{\frown}}$ is a function of $\widehat{F}$ as per equation (6). However, the figure also illustrates that convergence is much faster (indeed exponential) for the confusion operator $\widehat{\ddot{\frown}}$ than for the frequency opera-

the exact same argument given above, generalized to the case of multiple outcomes as we do below in Appendix B. Since we must have $(1-p)^{M-n_\uparrow}p^{n_\uparrow} > 0$ for any sequence that actually occurs in some sphere in some branch of the wavefunction, a corollary is that this exact same sequence will occur in *every* branch of the wavefunction (after neglecting the correction term with zero Hilbert space norm), regardless of how vast this sphere is. Moreover, the two-outcome result implies that the relative frequencies in a randomly selected volume in a given branch give no information whatsoever about which branch it is in, because they all have the same average frequencies. Finally, the multiple outcomes result tells us that even if we compute any finite amount of statistical data from a truly infinite volume (by computing what fraction of the time various combinations of outcomes occur in the infinite volume), the different branches remain statistically indistinguishable. In just the sense of Sec. II A, the different branches are equivalent to different realizations of an infinite universe with the same statistical properties, and therefore cannot be told apart.

V. MEASUREMENT AND DECOHERENCE

In our discussion above we have mentioned "outcomes" of experiments. It is generally agreed that in a measurement process, the post-measurement state of the measurement system is encoded in the degrees of freedom of a macroscopic device (say the readout of a Stern-Gerlach apparatus, the position of a macroscopic pointer, or the brain state of an observer). We can sketch how this process plays out in the simple example and cosmological context of this paper by considering, along with the set of replica quantum systems (each a single-spin system represented by $\alpha|\uparrow\rangle + \beta|\downarrow\rangle$), a corresponding set of indistinguishable measuring devices in a 'ready' state just prior to measurement. Following the scheme of Von Neumann [60], if the apparatus is in a particular 'ready' state $|a_r\rangle$, independent of the system's state, then interaction between the system and the apparatus causes the combined system to evolve into an entangled state:

$$(\alpha|\uparrow\rangle + \beta|\downarrow\rangle))|a_r\rangle \longrightarrow \alpha|\uparrow\rangle|a_\uparrow\rangle + \beta|\downarrow\rangle|a_\downarrow\rangle, \qquad (14)$$

where $|a_\uparrow\rangle$ and $|a_\downarrow\rangle$ are states of the apparatus in which it records an 'up' or 'down' measurement.

Let us consider a set of N perfect replicas of this setup[14] that exist in an infinite statistically uniform space for just the same reason that there are copies of $\alpha|\uparrow\rangle + \beta|\downarrow\rangle$. Following the same reasoning as in Section III, we can consider the product state, which looks

———————

[14] In appendix C we generalize this discussion to the arguably more relevant case where the measuring devices are *macroscopically* indistinguishable, and in particular, described by the same density matrix.

like

$$|\psi\rangle \;=\; (\alpha|\downarrow\rangle + \beta|\uparrow\rangle))|a_r\rangle \otimes (\alpha|\downarrow\rangle + \beta|\uparrow\rangle))|a_r\rangle \otimes \ldots (15)$$

Now, after the interaction between the system and apparatus described by equation (14), and following the exact same reasoning as Section III and Section IV, when $N \to \infty$, our product state becomes an infinite superposition of terms, all of which (except for a set of total Hilbert space norm zero) look like

$$\ldots|\uparrow\rangle|a_\uparrow\rangle \otimes |\uparrow\rangle|a_\uparrow\rangle \otimes |\downarrow\rangle|a_\downarrow\rangle \otimes |\uparrow\rangle|a_\uparrow\rangle \otimes |\downarrow\rangle|a_\downarrow\rangle\ldots, \quad (16)$$

where the relative frequencies of the terms $|\uparrow\rangle|a_\uparrow\rangle$ and $|\downarrow\rangle|a_\downarrow\rangle$ are given by $|\alpha|^2$ and $|\beta|^2$, respectively.

In this way, the interaction between system and apparatus has evolved $|\psi\rangle$ from a superposition containing infinitely many identical apparatuses into one of statistically indistinguishable terms, where each term describes two different sets of apparatuses: one in which each apparatus reads 'up', and one in which each reads 'down', with relative frequencies $|\alpha|^2$ and $|\beta|^2$.

Now, each system described by equation (14) will further interact with the degrees of freedom making up its environment, which we assume have a random character. This causes the local superposition to undergo decoherence [21, 31, 33]. This decoherence is typically quite rapid, with timescales of order 10^{-20} seconds being common [32, 61, 62]. Let us consider the effect of this decoherence on the density matrix ρ of the full N-particle system. Initially, $\rho = |\psi\rangle\langle\psi|$, where $|\psi\rangle$ is the pure N-particle state given by equation (15). Since each apparatus is rapidly entangled with its own local environment, all degrees of freedom of the full N-particle system decohere on the same rapid timescale. This means that when we compute the resulting N-particle reduced density matrix ρ by partial-tracing the global density matrix over the other (environment) degrees of freedom, it becomes effectively diagonal in our basis (given by Equation 2). The vanishing off-diagonal matrix elements are those that connect different pointer states in the grand superposition, and also therefore in the individual systems; thus any quantum interference between different states becomes unobservable.

In other words, decoherence provides its usual two services: it makes quantum superpositions for all practical purposes unobservable in the "pointer basis" of the measurement [31], and it dynamically determines which basis this is (in our case, the one with basis vectors like the example in Equation (16)).

VI. INTERPRETATION

Application of quantum theory to the actually existing infinite collection of identical quantum systems that is present in an infinite statistically uniform space (such as provided by eternal inflation) leads to a very interesting quantum state. As long as we are willing to neglect a part

of the wavefunction with vanishing Hilbert-space norm, then we end up with a superposition of a huge number of different states, each describing outcomes of an infinite number of widely separated identical measurements in our infinite space. In *all* of them, a fraction $p = |\alpha|^2$ of the observers will have measured spin up.[15]

In this way, the quantum probabilities and frequentist observer-counting that coexisted in the finite-N case have merged. Born's rule for the relative probabilities of $\uparrow$ and $\downarrow$ emerges directly from the relative frequencies of actual observers within an unbounded spatial volume; Born's rule as applied to the grand superposition is superfluous since all give the same predictions for these relative frequencies. In particular, the "quantum probabilities" assumed in Eq. 3 to be given by $\alpha^*\alpha$ and $\beta^*\beta$ are replaced by the assumption that two vectors in Hilbert space are the same if they differ by a vector of zero norm.

One of the most contentious quantum questions is whether the wavefunction ultimately collapses or not when an observation is made. Our result makes the answer to this question anti-climactic: insofar as the wavefunction is a means of predicting the outcome of experiments, it doesn't matter, since all the elements in the grand superposition are observationally indistinguishable. In fact, since each term in the superposition is indistinguishable from all the others, it is unclear whether it makes sense to even call this a superposition anymore.[16] Since each state in the superposition has formally zero norm, there is no choice but to consider classes of them, and if we class those states with indistinguishable predictions together, then this group has total Hilbert space norm of unity, while the class of "all other states" has total norm zero. In term of prediction, then, the infinite superposition of states is completely indistinguishable from one quantum state (which could be taken to be any one of the terms in the superposition) with unity norm. In this sense, a hypothetical "collapse" of the wavefunction would be the observationally irrelevant replacement of one statevector with another functionally identical one.

In more Everettian terms we might say that in the

cosmic wavefunction, each of the "many worlds" are the same world, where "world" here refers to the state of an infinite space. In the terminology of [63, 64], the Level I Multiverse is the same as the Level III Multiverse (and if inflation instantiates more than one solution to a more fundamental theory of physics, then the Level II Multiverse is the same as the Level III Multiverse).

All this suggests that we take a different and radically more expansive view of the statevector for a finite system: this quantum state describes not a particular system "here", but rather the spatial collection of identically prepared systems that already exist. This provides a real collection rather than fictitious ensemble for a statistical interpretation of quantum mechanics. It also allows quantum mechanics to be unitary in a very satisfactory way. Rather than the world "splitting" into a decohered superposition of two outcomes as seen by an experimenter, infinitely many observers already exist in different parts of space, a fraction $|\alpha|^2$ of which will measure one outcome, and a fraction $|\beta|^2$ of which will measure the other. The uncertainty represented by the superposition corresponds to the uncertainty before the measurement of *which* of the infinitely many otherwise-identical experimenters the observer happens to be; after the measurement the observer has reduced this uncertainty. Moreover, the "partially real" observers in the Everett picture have vanished: observers either exist equally if they are part of the wavefunction with unity, or don't exist if they are part of the zero norm branch that has been discarded.

VII.   DISCUSSION

The above-mentioned results raise interesting issues that deserve further work, and we comment on a couple below.

A.   Levels of indistinguishability

Considering our finite or infinite sequences, in which ways are two such sequences distinguishable, and what does this mean physically? For finite N, if each system is *labeled*, then each term in the superposition is different. However, if we consider just *statistical* information such as the relative frequencies of $|\uparrow\rangle$ and $|\downarrow\rangle$, then many sequences will be statistically indistinguishable. Moreover, if we do not label the terms, so that sequences can be re-ordered when they are compared, then *only* the relative frequencies are relevant.

Now for an infinite product state, we have argued above that once we discard a zero-norm portion, the remaining states are statistically indistinguishable using any finite amount of statistical information. If the elements are unlabeled, this statistical information is simply p (the $|\uparrow\rangle$ fraction). We can ask, however, if these states are *mathematically* distinguishable. To see that indeed

---

[15] For readers who are concerned whether infinity should be accepted as a meaningful quantity in physics, it is interesting to also consider the implications of a very large but finite N. In this case, the Hilbert space norm of the wavefunction component where the Born rule appears invalid is bounded by $2e^{-2\epsilon^2 N}$, so although it is not strictly zero, it is exponentially small as long as $N \gg \epsilon^{-2}$. For example, if $N = 10^{1000}$ (a relatively modest number in many inflation contexts), then the Born rule probability predictions are correct to 100 decimal places except in a wavefunction component of norm around $10^{-10^{800}}$.

[16] As well as giving identical predictions, this replacement is irrelevant because decoherence has removed any practical possibility of interference. This does not mean that the superposition has *mathematically* gone away, however, any more than when decoherence is applied to a single quantum system. Nor does it mean that mathematically the sequences are necessarily equivalent; see Section VII A.

they are, consider two such states such as ...$|\uparrow\rangle|\downarrow\rangle|\uparrow\rangle|\uparrow\rangle$..., and ...$|\downarrow\rangle|\downarrow\rangle|\uparrow\rangle|\uparrow\rangle$... and represent them as binary strings with $1 = |\uparrow\rangle$ and $0 = |\downarrow\rangle$. Now for each state, arbitrarily select one system, and enumerate all systems as $i = 1, 2, ...$ by increasing spatial distance from this central element. Each state then corresponds to a real number in binary notation such as $0.a_1 a_2...$ where $a_i = 0, 1$. Because the selected system is arbitrary given the translation symmetry of the space, we can consider our two sequences as indistinguishable if these real number representations match for any choice of the central element of each sequence. But there are only countably many such choices, and uncountably many real numbers, so a given sequence is indeed distinguishable from some (indeed almost all) other sequences in this sense.

Now, *physically*, we can ask two key questions. First, is the difference between (finitely) *statistically indistinguishable* and *mathematically indistinguishable* important, given that any actual operation will only be able to gather a finite amount of statistical information? Second, given that these systems are by assumption identical and indistinguishable, and far outside of each others' horizon, is it meaningful to think of them as labeled (even if there *were* a preferred element in terms of their global distribution)?

If the answer to either question is negative, then it becomes unclear what purpose is served by distinguishing the elements in the post-measurement superposition, and one might ask whether in some future formulation of quantum cosmology they might be meaningfully identified, thereby rendering the issue wavefunction collapse fully irrelevant.

B. The Chicken-and-Egg Problem of Quantum Spacetime

We have argued that an infinite statistically uniform space can naturally emerge in modern cosmology, and place the quantum measurement problem in a very different light. Yet quantum processes affect spacetime in any theory, and in inflation are responsible for the large-scale density fluctuations. Moreover, some versions of eternal inflation themselves *rely* on quantum processes: stochastic eternal inflation is eternal solely due to quantum fluctuations of the inflaton, and in open eternal inflation, inflation *ends* due to quantum tunneling. If quantum probabilities (and rates, etc.) are to be understood by making use of a cosmological backdrop, how do we make sense of the quantum processes involved in creating that backdrop?[17] There are at least four alternative ways in which we might view this chicken-and-egg problem.

First, we might from an Everettian perspective consider such processes as simply parts of the unitary evolution (or non-evolution, if considering the Wheeler-DeWitt equation) of a wave-function(al). Processes such as measurement of quantum systems by apparatuses and observers would not be meaningful until later times at which a classical approximation of spacetime had already emerged to describe an infinite space, in which such quantum measurement outcomes can be accorded probabilities via the arguments of this paper. How this view connects with the existing body of work on the emergence of classicality is an interesting avenue for further research.

Second, we might retain the Born rule as a axiomatic assumption, and employ it to describe such processes; then, for quantum measurements that exist as part of a spatial collection, the Born rule would be superfluous, as the probabilities for measurement outcomes are better considered as relative frequencies as per the arguments above.

Third, we might consider a classical spacetime description as logically prior to the quantum one, and in particular postulate certain symmetries that would govern the spacetime in the classical limit. In this perspective we could simply assume an FLRW space obeying the CP [45]. Or, we might include inflation and eternal inflation, but postulate an appropriately generalized inflationary version of the CP (or *perfect* cosmological principle) governing the semi-classical universe (see [65, 66] for ideas along these lines). Within this context, quantum events such as bubble nucleations could also be considered as part of a spatial collection of identical regions, and the whole set of arguments given herein could be applied to accord them probabilities. This would, in the language of [63], unify the "Level II" (inflationary) multiverse with the "Level III" (quantum) multiverse.

Fourth, we might imagine that a full theory of quantum gravity in some way fundamentally changes the quantum measurement problem, and that the considerations herein, based on standard quantum theory, apply only processes in a well-defined background spacetime.

This subtle issue is similar to the – possibly related – matter of Mach's principle in General Relativity (GR): applications of GR almost invariably implicitly assume a background frame that is more-or-less unaccelerated with respect to the material contents of the application; yet this coincidence between local inertial frames and the large-scale bulk distribution of matter is almost certainly of cosmological origin. Should Mach's principle simply be assumed as convenient, or explained as emerging in a particular limit from dynamics that do not assume it, or does it require specification of cosmological boundary conditions, or must new physics beyond GR be introduced?

---

[17] Similar considerations might apply to other fields – determined by quantum events but stretched into homogeneity by inflation – that go into making up the "background" in which the quantum experiment is posed, and to which the CP applies.

VIII. CONCLUSIONS

Modern inflationary cosmology suggests that we exist inside an infinite statistically uniform space. If so, then any given finite system is replicated an infinite number of times throughout this space. This raises serious conceptual issues for a prototypical measurement of a quantum system by an observer, because the measurer cannot know which of the identical copies she is, and must therefore ascribe a probability to each one [35–37, 40]. Moreover, as shown by Page [35], this cannot be seamlessly done using the standard projection operator and Born rule formalism of quantum mechanics; rather, it implies that quantum probabilities must be augmented by probabilities based on relative frequencies, arising from a measure placed on the set of observers. We have addressed this issue head-on by suggesting that perhaps it is not observer-counting that should be avoided, but quantum probabilities that should *emerge* from the relative frequencies across the infinite set of observers that exist in our three-dimensional space.

To make this link between quantum measurement and cosmology, we have built on the classic work concerning frequencies of outcomes in repeated quantum measurements [38, 39, 41, 42]. Our goal has not been to add further mathematical rigor (see [67–70] for the current state-of-the-art[18]), but instead to develop these ideas in the new context of the physically real, spatial collection provided by cosmology. The argument shows that the product state of infinitely many existing copies of a quantum system can be rewritten as an infinite superposition of terms. Because each term has zero norm, these must be grouped in terms of what they predict. Projecting these states with a "confusion operator" shows that a grouping with total Hilbert space norm unity consists of terms all of which are functionally indistinguishable, and contain relative frequencies of measurement outcomes in precise accordance with the standard Born rule. The remaining states, which would yield different relative frequencies, have total Hilbert space norm zero.

Predictions of measurement outcomes probabilities in this situation are, then, provided entirely by relative frequencies; if conventional quantum probabilities enter at all, it is only to justify the neglect of the zero-norm portion of the global wavefunction. Because all terms in the unity-norm portion are indistinguishable, quantum interpretation must be done in a *cosmological* light. Any "collapse" of the wavefunction is essentially irrelevant, since collapse to any of the wavevectors corresponds to exactly the same outcome. In Everettian terms, the "many worlds" are all the same; moreover, frequentist statistics emerge even for a single quantum measurement[19] (rather than a hypothetical infinite sequence of them). In this "cosmological interpretation" of quantum mechanics, then, quantum uncertainty ultimately derives from uncertainty as to which of many identical systems the observer actually is.

In conclusion, the quantum measurement issue is fraught with subtlety and beset by controversy; similarly, infinite and perhaps diverse cosmological spaces raise a host of perplexing questions and potential problems. We suggest here that perhaps combining these problems results not in a multiplication of the problems, but rather an elegant simplification in which quantum probabilities are unified with spatial observer frequencies, and the same infinite, homogeneous space that provides a real, physical collection also provides a natural measure with which to count. Further comprehensive development of this quantum-cosmological unification might raise further questions, but we hope that it may unravel further theoretical knots at the foundations of physics as well.

Appendix A: Infinite statistically uniform spaces from eternal inflation

In this appendix, we describe in greater detail how eternal inflation produces infinite spaces. Open inflation is perhaps the most well-studied case. The Coleman-DeLuccia instanton [72] describes the spacetime and field configuration resulting from the nucleation of a single bubble where the inflaton field has lower energy. In single-field models, each constant-field surface is a space of constant negative curvature, so that the bubble interior can be precisely described as an open Friedmann universe [72]. Bubble collisions (which are inevitable) complicate this picture; see [73] for a detailed review. In

---

[18] The core mathematical question is how to deal with the measure-zero set of confused freak observers; but as emphasized in [71], this issue is not unique to quantum mechanics, but occurs also in classical statistical mechanics and virtually other theory involving infinite ensembles.

[19] For example, the author ordering for this paper was determined by a single quantum measurement, and the order you yourself read is shared by exactly half of all otherwise-indistinguishable worlds spread throughout space.

no case, however, do collisions prevent the existence inside the bubble of an infinite, connected, spatial region[20]

In topological eternal inflation [74, 75], a region of inflating spacetime is maintained by a topological obstruction such as a domain wall or monopole, where the field is at a local maximum of its potential. The causal structure of these models is fairly well-understood if a bit subtle (see, e.g., [74, 77]), and also contains infinite spacelike surfaces after inflation has reheating surfaces.

The precise structure of post-inflation equal-field surfaces in stochastic eternal inflation is less well-defined, as the spacetimes cannot be reliably calculated with analytic or simple numerical calculations. However, they are expected to be infinite (e.g., [34, 78]), with volume dominated by regions in which inflation has emerged naturally from slow roll [79], and thus space is relatively flat and uniform.
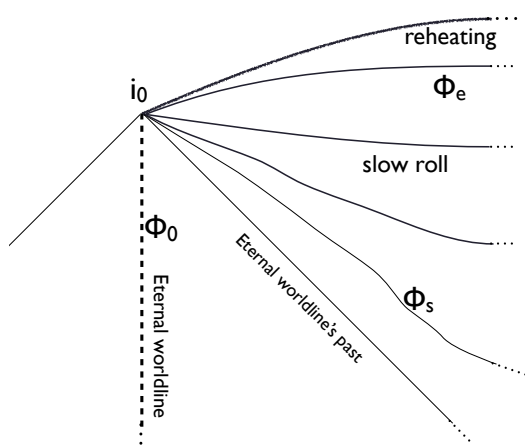

FIG. 2: Generic conformal structure of eternal inflation and post-inflationary reheating surface in open, topological, or stochastic eternal inflation. A region in which inflation is eternal (i.e. contains timelike worldlines that can extend to infinite proper time while remaining in the inflation region) is bounded by an infinite surface of field value ϕ_s at which eternality fails. Between this value and ϕ_e lie some number of e-folds of slow-roll inflation. This slow-roll inflation ends on a spacelike surface that represents a natural equal-time surface for the subsequent evolution including reheating, etc. This surface is spatially infinite, with spatial infinity denoted by i_0. (The future infinities following the reheating surface are not depicted).

That infinite spacelike reheating surfaces are generic in eternal inflation is not surprising, as per the following heuristic argument. Roughly speaking, eternal inflation

occurs when some obstruction prevents the inflaton field ϕ from evolving away from some value ϕ_0 during a typical Hubble time, so that on average, the physical volume containing that field value increases. This implies that there are some (extremely rare) worldlines threading a horizon volume in which $\phi \approx \phi_0$ forever. For such a model to be observationally viable, however, there must exist a route through field-space that crosses a value ϕ_s at which slow-roll inflation begins, then a value ϕ_e at which inflation ends and matter or radiation-domination begins.

This field evolution plays out in spacetime as well, connecting the eternally inflating spacetime region to the post-inflationary region, as sketched in Fig. 2 for a single inflaton field. Whether eternal inflation is open, topological, or stochastic, this diagram must look essentially the same. The surface of constant field ϕ_s, at which eternality fails, is by definition one for which very few worldlines cross back to field values near ϕ_0; this surface must therefore be spacelike nearly everywhere. Morever, slow-roll inflation to the future of this surface exponentially suppresses field gradients, so that surfaces of constant field quickly become uniformly spacelike[21] as the field approaches ϕ_e.

Now, if we consider a spatial region with $\phi = \phi_e$ (soon to evolve to the reheating surface), and follow the surface of constant field $\phi = \phi_s$ in a direction toward the eternal region, we see three things. First, the $\phi = \phi_e$ surface must be infinite, since the $\phi = \phi_s$ surface, for example, is infinite, and further inflates into the constant field surfaces with $\phi_s \leq \phi \leq \phi_e$. Second, by the above argument the $\phi = \phi_e$ surface is *spacelike*, so there is no obstacle to continuing the foliation in our original region as far as we like toward the eternal region. Third, since each point on our surface has roughly the same classical field history since ϕ_s (variations due to the initial field velocity being bounded by the slow-roll condition, and curvature being stretched away), the $\phi = \phi_e$ surface should be uniform up to variations induced by quantum fluctuations during the field's evolution.

Multi-field models are more complex, in that we might imagine many routes for the vector of field valued $\vec{\phi}$ to take from a given (set of) 'eternal' field values $\vec{\phi}_0$ to field values $\vec{\phi}_e$ at which inflation ends. Yet it seems likely that in this case the prime difference would be for the $\phi = \phi_s$ surface to be replaced by an infinite, spacelike surface Σ_s on on which eternality failes, and on which $\vec{\phi}$ is inhomogeneous.[22] Yet for a given value $\vec{\phi}_s$ of this vector on this surface, subsequent evolution would be

---

[20] This region could be delimited by, e.g. defining some criterion by which to identify regions affected by the collision, and excluding them. After this removal, the remaining region would have uniform properties (as would the statistics of the excisions; see below.)

[21] For a more precise specification of this point, see [76, 77].

[22] This is precisely what happens in "Quasi-open" inflation [80]. There, the bubble interior cannot be foliated into constant-field equal-time surfaces. However, there is still an infinite space with infinitely many finite-sized regions in which the field is constant; a subset of these with the same field value could be taken as a statistically uniform (though potentially disconnected) space.

classically similar, so the subset $\Sigma_{\vec{\phi}_s}$ of Σ_s containing this same vector of field values would evolve to a set $\sigma_{\vec{\phi}_s}$ of potentially disconnected regions[23] with uniform properties. Because Σ_s is infinite, but the fields on it can only take a finite range of values, determined randomly, it would seem that each $\sigma_{\vec{\phi}_s}$ must be infinite in volume as in the single field case.

The sense in which the post-inflationary spaces produced by inflation are statistically uniform, in the sense described in Section II C, is rather subtle, as illustrated by the case of "open" eternal inflation. In an open FLRW space created by a single bubble nucleation in a background false-vacuum space, as described by [72], the properties of the space are precisely uniform; this is guaranteed by the de Sitter symmetries obeyed by the progenitor space, which correspond to translations on a spatial slice within the bubble (see [73] and references therein for more detail.) However, a single isolated bubble is not realistic, and we must consider potential bubble collisions. For a given worldline inside a bubble, collisions come in two classes: "early" bubbles that enter the past of the worldline within its first Hubble time, and "late bubbles" that enter later. Early bubbles affect nearly all worldlines, and do *not* obey the CP: there is a distinguished position at which an observer sees the lowest rate of collisions. This defines a frame that breaks the de Sitter symmetry of the background space into the group of rotations and spacetime translations. However, there exists an infinite set of wordlines that see *no* early collisions, and are surrounded by an exponentially large region that is also unaffected. "Late" bubbles do obey the CP in terms of their probabilities, and can affect a fraction of volume that can be either large or exponentially small. In both early and late bubbles, a given type of collision may or may not significantly disrupt the dynamics of inflation within the region it affects, depending upon the details of the collision.

The bottom line is that there exists an infinite spacetime with statistically uniform properties up to some large scale, corresponding to the typical distance between areas affected by bubbles. On these larger scales, the properties of the bubble-collision regions themselves select a preferred position if "early" bubbles are included, but are homogeneous if only "late" bubbles are included. In any case, an infinite set of regions larger than our observable universe, with randomly chosen initial conditions draw from the same probability distribution, exists.

---

[23] If the regions are disconnected, it becomes less obvious how to place a natural measure on the spatial volumes in them, but in some cases it may not make a significant difference. For example, if the spacetime can be foliated into surfaces of constant curvature, these will approximately coincide with the surfaces of constant field-value over an infinite domain sharing the same inflationary history, which could be taken as a statistically uniform space; but these slices would also provide a way to compute spatial volumes without ambiguity.

Is this by itself enough for the arguments of this paper to go through? In principle, the *measure* by which the observers in this collection are counted can affect relative frequencies of observer types. In simple cases, such as a pure FLRW universe as might arise in single bubble nucleation, the spatial rotation and translation symmetries pick out a preferred measure given by spatial volume. For more complicated spaces in which these symmetries are less clear, the measure can be quite important for certain questions. However, this is only really true if the measure depends in some way on the relevant information discriminating these types of observers. For questions of cosmological observations, this can easily be the case. In our case of a local laboratory measurement, it is hard to see how the measure can depend on the outcome of a Stern-Gerlach experiment, thus hard to see how the measure could enter the relative frequencies of observed outcomes in the collection.

Appendix B: Generalization to observations with more than two outcomes

In this section, we generalize the result that $\|\widehat{\widehat{\odot}}|\psi\rangle\|^2 \to 0$ as $N \to \infty$ from 2-state systems to systems with an arbitrary number of states $m \geq 2$. Instead of the two basis vectors $|\downarrow\rangle$ and $|\uparrow\rangle$, we now suppose m orthogonal basis vectors labeled as $|1\rangle, |2\rangle, \ldots, |m\rangle$. Now consider a finite uniform space that is large enough to contain N identical copies of our system (which we will keep referring to as a "particle" for simplicity, even though our proof is valid for an arbitrary m-state system), each prepared in the state

$$|\psi\rangle = \alpha_1|1\rangle + \alpha_2|2\rangle + \cdots + \alpha_m|m\rangle. \qquad (B1)$$

Analogously with equation (1), the state of this combined N-particle system is simply a tensor product with N terms, each of which involves a sum of N basis vectors. The frequency operator $\widehat{F}$ from above is naturally generalized to an m-dimensional vector operator $\widehat{\mathbf{F}}$, whose i^{th} component $\widehat{\mathbf{F}}_i$ measures the frequency of the outcome $|i\rangle$. In other words, equation (4) is diagonal in the m^N-dimensional basis spanned by products of N of our 1-particle states, and the eigenvalues of $\widehat{\mathbf{F}}_i$ are the frequencies of $|i\rangle$ in these basis vectors. For example, for $N = 5$ and $m = 5$,

$$\widehat{\mathbf{F}}|3\rangle|1\rangle|4\rangle|1\rangle|5\rangle = \begin{pmatrix} 2/5 \\ 0 \\ 1/5 \\ 1/5 \\ 1/5 \end{pmatrix} |3\rangle|1\rangle|4\rangle|1\rangle|5\rangle. \qquad (B2)$$

Examining the norm of the state multiplied by $(\widehat{\mathbf{F}} - \mathbf{p})$, where the m-dimensional vector $\mathbf{p}$ is defined by $p_i \equiv$

$|\alpha_i|^2$, equation (7) now generalizes to

$$\langle\psi|(\widehat{\mathbf{F}}_i - p_i)(\widehat{\mathbf{F}}_j - p_j)^t|\psi\rangle =$$

$$= \sum_{n_1...n_m} \binom{N}{n_1...n_m} p_1^{n_1} \ldots p_m^{n_m} \left(\frac{n_i}{N} - p_i\right)\left(\frac{n_j}{N} - p_j\right) =$$

$$= \mathbf{C}_{ij} = \frac{\mathbf{C}_{ij}^*}{N^2}, \qquad (B3)$$

where the sum on the second row is over all m-tuplets of natural numbers $(n_1, ..., n_m)$ such that $\sum_{i=1}^{m} n_i = N$, and the parenthesis following the summation symbol denotes the multinomial coefficient $N!/n_1!...n_m!$. The $m \times m$ matrix $\mathbf{C}^*$ is given by

$$\mathbf{C}_{ij}^* = \begin{cases} Np_i(1-p_i) & \text{if } i = j, \\ -Np_ip_j & \text{if } i \neq j. \end{cases} \qquad (B4)$$

The terms $\binom{N}{n_1...n_m} p_1^{n_1} \ldots p_m^{n_m}$ in equation (B3) are simply the familiar coefficients of a multinomial distribution that has mean vector $N\mathbf{p}$. We recognize the last two lines of equation (B3) as the definition of the covariance matrix $\mathbf{C}^*$ of the multinomial distribution, and the proof of equation (B4) is simply the well-known multinomial covariance derivation.

Since $\mathbf{C} \equiv \mathbf{C}^*/N^2 \propto 1/N \to 0$ as $N \to \infty$, the multinomial distribution governing the spread in values of the frequency operator vector in Eq. B3 approaches an m-dimensional δ-function centered on the vector $\mathbf{p}$. In particular, the special case $i = j$ implies that

$$\|(\widehat{\mathbf{F}}_i - p_i)|\psi\rangle\|^2 = p_i(1-p_i)/N \to 0, \qquad (B5)$$

so that $|\psi\rangle$ becomes an eigenvector of all of the components of $\widehat{\mathbf{F}}$ (again in the sense that the correction term has zero norm), with eigenvalues corresponding to the coefficients of the vector $\mathbf{p}$, as shown in the seminal paper [39].

Let us now consider the confusion operator $\widehat{\odot}$. Since we are now considering not merely one frequency but m different frequencies p_i, we generalize our definition of $\widehat{\odot}$ to be the projection operator onto those basis vectors where the spin-up fraction differs by more than a small predetermined value ϵ from the Born rule prediction $p_i = |\beta_i^2|$ for *any* $i = 1, ..., m$. It is therefore related to the frequency operator vector $\widehat{\mathbf{F}}$ by

$$\widehat{\odot} = \widehat{I} - \prod_{i=1}^{m} \theta\left(\epsilon - |\widehat{F}_i - p_i|\right), \qquad (B6)$$

where θ again denotes the Heaviside step function and $\widehat{I}$ is the identity operator.

To gain intuition about the confusion operator $\widehat{\odot}$, it is helpful to consider the m-dimensional generalization of Figure 1, the m-dimensional space of vectors $\mathbf{f}$ where the ith component $f_i = n_i/N$ is the frequency of the outcome $|i\rangle$. In the basis where $\widehat{\odot}$ is diagonal (with basis vectors like the example $|3\rangle|1\rangle|4\rangle|1\rangle|5\rangle$), each basis vector maps to a unique point $\mathbf{f}$ in this space which is the vector of eigenvalues exemplified in equation (B2), *i.e.*, a point with coordinates corresponding to the various frequencies $f_i = n_i/N$. In this space, the eigenvalues of $\widehat{\odot}$ equal 0 for eigenvectors falling inside an m-dimensional hypercube of side length 2ϵ centered at the point $\mathbf{p}$, and otherwise equal 1. Analogously to equation (8), we obtain

$$\|\widehat{\odot}|\psi\rangle\|^2 = \langle\psi|\widehat{\odot}|\psi\rangle =$$

$$= \sum_{n_1...n_m} \binom{N}{n_1...n_m} p_1^{n_1}...p_m^{n_m} \left[1 - \prod_{i=1}^{m} \theta\left(\epsilon - \left|\frac{n_i}{N} - p_i\right|\right)\right]$$

$$= \sum_{\text{Outside cube}} \binom{N}{n_1...n_m} p_1^{n_1}...p_m^{n_m}. \qquad (B7)$$

The sum on the second row is again over all m-tuplets sets of natural numbers $(n_1, ..., n_m)$ such that $\sum_{i=1}^{m} n_i = N$, and the sum on the last row is restricted to the subset lying outside of the above-mentioned hypercube where all frequencies are within ϵ of the Born rule prediction.

Since the binomial distribution for the frequency vector $\mathbf{f}$ is known to approach a δ-function $\delta(\mathbf{f} - \mathbf{p})$ as $N \to \infty$, it is obvious from equation (B7) that $\|\widehat{\odot}|\psi\rangle\|^2 \to 0$ in this limit. Just like for the 1-particle case, the convergence will be quite rapid (faster than polynomial) once the width of the multinomial distribution starts becoming significantly smaller than that of the surrounding hypercube. This is because by the central limit theorem, the multinominal distribution becomes well-approximated by a Gaussian in the $N \to \infty$ limit, so we can approximate the sum in equation (B7) by an integral over a multivariate Gaussian with mean $\mathbf{p}$ and covariance matrix $\mathbf{C} = \mathbf{C}^*/N^2$:

$$\|\widehat{\odot}|\psi\rangle\|^2 = \sum_{\text{Outside cube}} \binom{N}{n_1...n_m} p_1^{n_1}...p_m^{n_m} \approx \int_{\text{Outside cube}} \frac{1}{(2\pi)^{N/2}|\mathbf{C}|^{1/2}} e^{-\frac{1}{2}(\mathbf{f}-\mathbf{p})^t \mathbf{C}^{-1}(\mathbf{f}-\mathbf{p})} d^N f$$

$$\leq \int_{\text{Outside cube}} \frac{1}{(2\pi)^{N/2}|\frac{\mathbf{I}}{2N}|^{1/2}} e^{-\frac{1}{2}(\mathbf{f}-\mathbf{p})^t (\frac{\mathbf{I}}{2N})^{-1}(\mathbf{f}-\mathbf{p})} d^N f = \int_{\text{Outside cube}} \frac{1}{(\pi/N)^{N/2}} e^{-N|\mathbf{f}-\mathbf{p}|^2} d^N f$$

$$= 1 - \int_{\text{Inside cube}} \frac{1}{(\pi/N)^{N/2}} e^{-N|\mathbf{f}-\mathbf{p}|^2} d^N f = 1 - \int_{\text{Inside cube}} \frac{1}{(\pi/N)^{N/2}} \prod_{i=1}^{N} e^{-N(f_i - p_i)^2} d^N f$$

$$= 1 - \text{erf}\left[\epsilon N^{1/2}\right]^N = 1 - \left(1 - \text{erfc}\left[\epsilon N^{1/2}\right]\right)^N$$

$$\rightarrow N \text{erfc}\left[\epsilon N^{1/2}\right] \rightarrow 0 \quad \text{as } N \rightarrow \infty. \tag{B8}$$

On the second line, we used the readily proven fact that no eigenvalue of $\mathbf{C}$ can exceed $1/2N$, which means that if we replace the covariance matrix $\mathbf{C}$ by $\mathbf{I}/2N$ (where $\mathbf{I}$ is the identity matrix), then the multivariate Gaussian remains at least as wide in all directions, meaning that the fraction of its integral residing outside of the hypercube is at least as large.

Appendix C: Measurement by the collection of macroscopically indistinguishable apparatuses

In Section V we treated the case in which every measuring apparatus was in an identical quantum state. Let us now generalize this to the arguably more relevant general case where the initial state of the apparatus is described by a density matrix. Specifically, we can write the density matrix of the apparatus in the "ready" state as

$$\rho_a = \sum_i p_i |i_r\rangle\langle i_r|, \tag{C1}$$

where i indexes the "ready" microstates $|i_r\rangle$ of the apparatus that are macroscopically indistinguishable, and the coefficients p_i satisfy $p_i \geq 0$ and $\sum_i p_i = 1$.[24] (The coefficients p_i are commonly interpreted as the probability of finding the apparatus in states $|i_\uparrow\rangle$, but no such interpretation is needed for our argument below.) The initial density matrix for the combined system and apparatus is thus

$$\rho = \sum_i p_i(\alpha|\uparrow\rangle + \beta|\downarrow\rangle))|i_r\rangle\langle i_r|(\alpha^*\langle\uparrow| + \beta^*\langle\downarrow|). \tag{C2}$$

———

[24] The density matrix ρ_a can without loss of generality be written in the diagonal form of equation (C1), because if it were not diagonal, then we could make it diagonal by defining new "ready" basis states that are the eigenvectors of ρ_a; they form an orthogonal basis because ρ_a is Hermitean, and they behave like classical apparatus states because they are superpositions of classical apparatus states (the old basis states) that are macroscopically indistinguishable.

With this notation, the unitary evolution during the measurement process given by equation (14) takes the form

$$(\alpha|\uparrow\rangle + \beta|\downarrow\rangle))|i_r\rangle \longrightarrow \alpha|\uparrow\rangle|i_\uparrow\rangle + \beta|\downarrow\rangle|i_\downarrow\rangle, \tag{C3}$$

where $|i_\uparrow\rangle$ and $|i_\downarrow\rangle$ are the corresponding microstates of the apparatus in which it records an 'up' or 'down' measurement. We thus have two resulting classes of apparatus states $\{|i_\uparrow\rangle\}$ and $\{|i_\downarrow\rangle\}$, where the states in each class are macroscopically indistinguishable, but where the two classes are macroscopically distinguishable (having a macroscopic pointer in different locations, say).

Let us now consider what happens to the product state of all the apparatuses and systems. The initial product state of equation (15) generalizes to a tensor product of N density matrices that are all given by equation (C2), so each of the N factors contains a sum over the local apparatus microstates i. If we expand this product of sums, then this total density matrix takes the form of a weighted average

$$\rho = \sum_{i_1,...,i_N} p_{i_1}...p_{i_N}\rho_{i_1,...i_N}, \tag{C4}$$

over all possible combinations of apparatus microstates. Here, $i_1...i_N$ each run over the set of detector states, p_i are as before, and $\rho_{i_1,...,i_N}$ are pure-state density matrices of the form $|\psi\rangle\langle\psi|$, where $|\psi\rangle$ is a product of the form

$$|\psi\rangle = |\uparrow\rangle|17_r\rangle \otimes |\uparrow\rangle|4711_r\rangle \otimes |\downarrow\rangle|5_r\rangle \otimes |\uparrow\rangle|17_r\rangle \otimes |\downarrow\rangle|666_r\rangle... \tag{C5}$$

(for example, $|666_r\rangle$ denotes the apparatus "ready" microstate $|i_r\rangle$ with $i = 666$). Now, after the interaction between the system and apparatus described by equation (C3), each of these pure-state density matrices evolves into a new $\rho'_{i_1,...i_N} = |\psi'\rangle\langle\psi'|$, where $|\psi'\rangle$ is a product of the form

$$|\psi\rangle = |\uparrow\rangle|17_\uparrow\rangle \otimes |\uparrow\rangle|4711_\uparrow\rangle \otimes |\downarrow\rangle|5_\downarrow\rangle \otimes |\uparrow\rangle|17_\uparrow\rangle \otimes |\downarrow\rangle|666_\downarrow\rangle... \tag{C6}$$

We see that the only difference between these pure states $|\psi'\rangle$ and the ones we considered in equation (16) above

is that the apparatus microstates now vary in a typically random-looking fashion (yet always in such a way that the an $|\uparrow\rangle$ particle state goes with one of the apparatus microstates in the "up" class). Applying the exact same reasoning as Section III and Section IV to one of these pure states $|\psi'\rangle$ when $N \to \infty$, our product state therefore becomes an infinite superposition of terms like equation (C6), all of which (except for a set of total Hilbert space norm zero) have the relative frequencies $|\alpha|^2$ and $|\beta|^2$ for terms with apparatus states in the "up" and "down" classes, respectively. Since this result holds for each $\rho'_{i_1,...,i_N}$, it clearly applies to ρ: it describes an collection of apparatuses, a fraction $|\alpha|^2$ of which measure "up" and a fraction $|\beta|^2$ of which measure "down" (up to a part of zero Hilbert space norm as usual).

Note that this density matrix and the one described in Section V have a different structure; but we can combine them if we also consider the interaction of our set of measuring devices with their local environments. In this case, if there are N particles, then after tracing out the environment, we have an overall density matrix that looks just like Equation (C4), but in which each of the $\rho_{i_1,...,i_N}$ is now a diagonal density matrix of just the type described in Section V, except that each apparatus is labeled by its state i_k, and all such combinations are combined as a weighted sum to get the total density matrix.

[1] M. Jammer, *The Philosophy of Quantum Mechanics: The Interpretations of Quantum Mechanics in Historical Perspective* (New York: Wiley, 1974) & Sons.
[2] M. Born, *Z.Phys.*, **37**, 863 (1926)
[3] L. Ballantine, *Quantum Mechanics: A Modern Development* (Singapore: World, 1998) Scientific Publishing Company
[4] N. Bohr 1985, in *Collected Works*, ed. J. Kalchar (North-Holland: Amsterdam)
[5] W. Heisenberg 1989, in *Collected Works*, ed. W. Rechenberg (Springer: Berlin)
[6] P. A. Dirac, *The Principles of Quantum Mechanics, 4th ed., revised* (Oxford: Clarendon Press, 1967)
[7] E. P. Wigner 1976, Princeton University Lecture Notes on Interpretation of Quantum Mechanics (unpublished).
[8] J. A. Wheeler, and W. H. Zurek, *Quantum Theory and Measurement* (Princeton University Press: Princeton, 1983)
[9] K. Gottfried and Yan T, *Quantum Mechanics: Fundamentals* (Springer: Berlin, 2003)
[10] E. Madelung, *Z.Phys.*, **40**, 322 (1927)
[11] W. Hoeffding, *Journal of the American Statistical Association*, **58 (301)**, 1330 (1963)
[12] J. von Neumann., *Die mathematischen Grundlagen der Quantenmechanik* (Springer: Berlin., 1932)
[13] F. London and E. Bauer 1939, in *Quantum Theory and Measurement (1983)*, ed. J. A. Wheeler and W. H. Zurek (Princeton Univ. Press: Princeton)
[14] E. P. Wigner, *Symmetries and Reflections* (Indiana University Press: Bloomington, 1967)
[15] H. P. Stapp, *In Mind, Matter, and Quantum Mechanics* (Springer: Berlin, 1993)
[16] D. Bohm, *Phys.Rev.*, **85**, 166 (1952)
[17] G. Birkhoff and J. von Neumann, *Ann.Math.*, **37**, 823 (1936)
[18] H. Everett, *Rev. Mod. Phys*, **29**, 454 (1957)
[19] H. Everett 1957, in *The Many-Worlds Interpretation of Quantum Mechanics*, ed. B. S. DeWitt and N. Graham (Princeton: Princeton Univ. Press, available at http://www.pbs.org/wgbh/nova/manyworlds/pdf/dissertation.pdf)
[20] E. Nelson, *Phys.Rev.*, **150**, 1079 (1966)
[21] H. D. Zeh, *Found.Phys.*, **1**, 69 (1970)
[22] D. N. Page, *Int.J.Mod.Phys.*, **D5**, 583 (1996)
[23] R. B. Griffiths, *J.Stat.Phys.*, **36**, 219 (1984)
[24] G. Ghirardi C, A. Rimini, and T. Weber, *PRD*, **34**, 470 (1986)
[25] J. Cramer, *Rev.Mod.Phys.*, **58**, 647 (1986)
[26] B. van Fraassen 1972, in *Paradigms and Paradoxes: The Philosophical Challenge of the Quantum Domain*, ed. R. Colodny (Univ. Pittsburgh Press: Pittsburgh)
[27] W. H. Zurek, arXiv:0707.2832 (1987)
[28] C. Rovelli, *Int. J. of Theor. Phys.*, **35**, 1637 (1996)
[29] R. Gambimi and J. Pullin, arXiv:0903.1859 (2009)
[30] D. Bohm, Quantum Theory, 1950 (1989 Dover edition), Ch. 22
[31] W. H. Zurek, *Nature*, **5**, 181 (2009)
[32] E. Joos and H. D. Zeh, *Z. Phys. B*, **59**, 223 (1985)
[33] M. Schlosshauer, *Decoherence and the Quantum-To-Classical Transition* (Springer: Berlin, 2007)
[34] J. Garriga and A. Vilenkin, *Phys. Rev. D*, **64**, 043511. (2001b)
[35] D. N. Page, *JCAP*, **0907**, 008 (2009)
[36] D. N. Page, arXiv:0907.4152 (2009)
[37] D. N. Page, arXiv:1003.2419 (2010)
[38] D. Finkelstein, *Trans.N.Y.Acad.Sci*, **25**, 621 (1963)
[39] J. Hartle, *Am.J.Phys.*, **36**, 704 (1968)
[40] J. Hartle, and M. Srednicki, *PRD*, **81**, 123524 (2010)
[41] E. Farhi, J. Goldstone, and S. Gutmann, *Ann.Phys.*, **192**, 368 (1989)
[42] S. Gutmann, *PRA*, **52**, 3560 (1995)
[43] D. Layzer, *Cosmogenesis* (Oxford University Press: , 1990)
[44] D. Layzer 1971, "Cosmogonic Processes in Astrophysics and General Relativity", in *Brandeis Summer School Lectures 1968*, ed. M. Chretien, S. Deser, and J. Goldstein
[45] D. Layzer, (2010) ArXiv:1008.1229
[46] J. Knobe, K. D. Olum, A. Vilenkin, *British Journal for the Philosophy of Science*, **57**, 47 (2006)
[47] A. Guth, *PRD*, **23**, 347 (1981)
[48] A. Linde, *Lect.Notes Phys.*, **738**, 1 (2008)
[49] A. Guth, *J. Phys. A*, **40**, 6811 (2007)
[50] A. Aguirre, ArXiv:0705.0164 (2008)
[51] A. Linde, *Physica Scripta*, **40**, (2005)
[52] E. J. Copeland, A. R. Liddle, D. H. Lyth, E. D. Stewart, and D. Wands, *PRD*, **49**, 6410 (1994)
[53] E. Komatsu *et al.*, arXiv:1001.4538 (2010)

[54] M. P. Hertzberg, M. Tegmark, S. Kachru, J. Shelton, Ozcan O, *PRD*, **76**, 103521 (2007)

[55] T. Banks, M. Dine, E. Gorbatov,, *JHEP*, **0408**, 058 (2004)

[56] N. Arkani-Hamed, S. Dubovsky, A. Nicolis, E. Trincherini, G. Villadoro, *JHEP*, **0705**, 055 (2007)

[57] T. Bankt, M. Johnson, arXiv:hep-th/0512141 (2005)

[58] S. Winitzki,, *Lect.Notes Phys.*, **738**, 157 (2008)

[59] M. Tegmark *et al.*, *PRD*, **69**, 103501 (2004)

[60] J. von Neumann, 1932, *Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin)

[61] M. Tegmark, *Found. Phys. Lett.*, **6**, 571 (1993)

[62] M. Tegmark, *PRE*, **61**, 4194 (2000)

[63] M. Tegmark, *Sci. Am.*, **270 (5/2003)**, 40 (2003)

[64] M. Tegmark, astro-ph/0302131 (2003)

[65] A. Aguirre and S. Gratton, *Phys.Rev. D*, **67**, 083515 (2003)

[66] A. Aguirre and S. Gratton, *Phys.Rev. D*, **65**, 083507 (2002)

[67] C. M. Caves and R. Schack, *Ann.Phys.*, **315**, 123 (2005)

[68] N. P. Landsman, arXiv:0804.4849v1 (2008)

[69] Wesep. R. Van A, *Annals of Physics*, **321**, 2438 (2006)

[70] S. Wada, *J.Phys.Soc.Japan*, **76**, 094004 (2007)

[71] M. Tegmark, arXiv:0905.2182 (2009)

[72] S. R. Coleman, L. Deuccia F, *PRD*, **21**, 3305 (1980)

[73] A. Aguirre,M. Johnson, *arXiv*, **0908**, 4105A (2009)

[74] A. Vilenkin, *PRL*, **72**, 3137 (1994)

[75] A. Linde, *PRD*, **50**, 2456 (1994)

[76] V. Vanchurin, A. Vilenkin, S. Winitzki, *PRD*, **61**, 083507 (2000)

[77] N. Sakai K. Nakao, and T. Harada, *PRD*, **61**, 127302 (2000)

[78] S. Winitzki, *PRD*, **65**, 083506 (2002)

[79] S. Gratton, *arXiv*, **1003**, 2409 (2010)

[80] J. Garcia-Bellido J. GarrigaX. Montes, *PRD*, **57**, 4669 (1998)