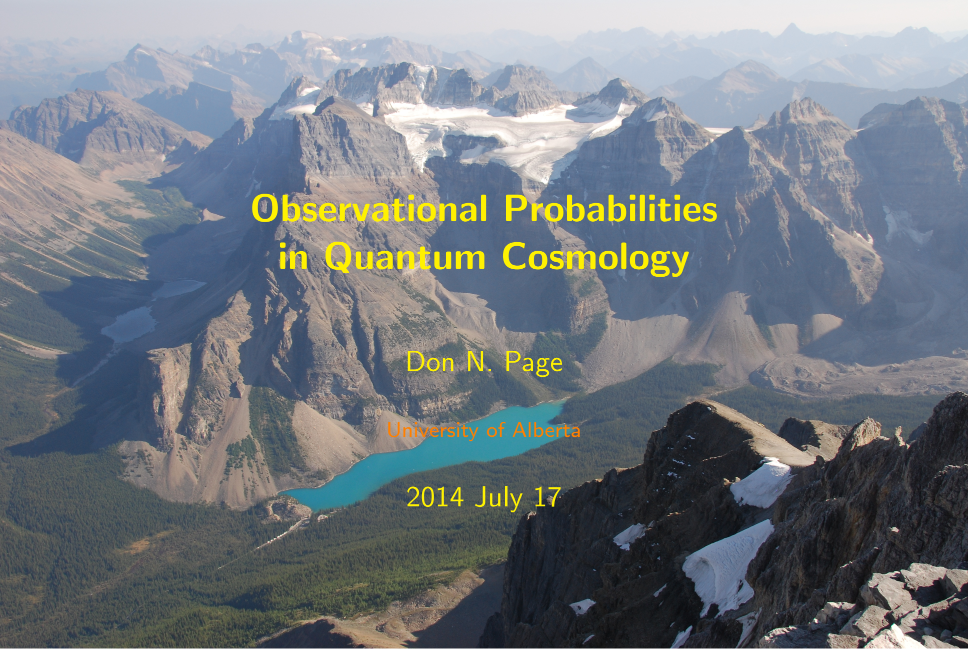

# Observational Probabilities in Quantum Cosmology

Don N. Page

University of Alberta

2014 July 17

# The Measure Problem of Cosmology

The measure problem of cosmology is how to obtain probabilities of observations from the quantum state of the universe.

This is particularly a problem when eternal inflation leads to a universe of unbounded size so that there are apparently infinitely many realizations or occurrences of observations of each of many different kinds or types, making the ratios ambiguous.

There is also the danger of domination by Boltzmann Brains, observers produced by thermal and/or vacuum fluctuations.

The measure problem is related to the measurement problem of quantum theory, how to relate quantum reality to our observations that appear to be much more classical.

An approach I shall take is to assume that observations are fundamentally conscious perceptions or sentient experiences (each perception being all that one is consciously aware of at once).

# Probabilities of Observations

I shall take an Everettian view that the wavefunction never collapses, so in the Heisenberg picture, there is one single fixed quantum state for the universe (which could be a 'multiverse').

I assume that instead of 'many worlds,' there are instead many different actually existing observations (sentient experiences) $O_k$, but that they have different positive measures, $\mu_k = \mu(O_k)$, which are in some sense how much the various observations occur, but they are *not* determined by the contents of the observations.

For simplicity, I shall assume that there is a countable discrete set $\{O_k\}$ of observations, and that the total measure $\sum_k \mu_k$ is normalized to unity for each possible complete theory $T_i$ that gives the normalized measures $\mu_k$ for all possible observations $O_k$.

Then for a Bayesian analysis, I shall interpret the normalized measures $\mu_k$ of the observation $O_k$ that each theory $T_i$ gives as the probability of that observation given the theory, $P(O_k|T_i)$, which for one's observation $O_k$ is the 'likelihood' of the theory $T_i$.

# A Bayesian Analysis for the Probabilities of Theories

Consider theories $T_i$ that for each possible observation $O_k$ give a normalized probability for that observation, $P(O_k|T_i)$, summing to one when one sums over all observations $O_k$ for each theory $T_i$.

If we assign (subjective) prior probabilities $P(T_i)$ to the theories $T_i$ (presumably higher for simpler theories, by Occam's razor) and use an observation $O_k$ to test the theory, $P(O_k|T_i)$ is then the likelihood of the theory, and by Bayes' theorem we can calculate the posterior probability of the theory as

$$P(T_i|O_k) = \frac{P(T_i)P(O_k|T_i)}{\sum_j P(T_j)P(O_k|T_j)}.$$

We would like to get this posterior probability as high as possible by choosing a simple theory (high prior probability $P(T_i)$) that gives a good statistical fit to one's observation $O_k$ (high likelihood $P(O_k|T_i)$).

# Trade-Off Between Prior Probabilities and Likelihoods

- Prior probabilities $P(T_i)$ for theories (intrinsic plausibilities)
- Conditional probabilities $P(O_k|T_i)$ for observations ('likelihoods' of the theories for a fixed observation)
- Posterior probabilities $P(T_i|O_k) \propto P(T_i)P(O_k|T_i)$

Prior probabilities are subjective, usually higher for simpler theories.

The highest prior probability might be for the theory $T_1$ that nothing concrete (contingent) exists, but then $P(O_k|T_1) = 0$.

$T_2$ might be the theory that all observations exist equally: $P(O_k|T_2) = 1/\infty = 0$ (modal realism or multiverse pantheism).

At the other extreme would be a maximal-likelihood theory giving $P(O_k|T_i) = 1$ for our observation $O_j$, but this seems to require a very complex theory $T_i$ that might be assigned an extremely tiny prior probability $P(T_i)$, hence giving a very low posterior $P(T_i|O_k)$.

Thus there is a trade off between prior probabilities and likelihoods, that is, between intrinsic plausibility and fit to observations.

## Sensible Quantum Mechanics or Mindless Sensationalism

The map from the quantum state to the measures of observations could be nonlinear. However, I assumed a linear relationship in Sensible Quantum Mechanics (which I have also called Mindless Sensationalism because it proposes that what is fundamental is not minds but conscious perceptions, which crudely might be called 'sensations,' though they include more of what one is consciously aware of than what is usually called 'sensations'):

$$\mu(O_k) = \sigma[A(O_k)] = \text{ expectation value of the operator } A(O_k).$$

Here $\sigma$ is the quantum state of the universe (a positive linear functional of quantum operators), and $A(O_k)$ is a nonnegative 'awareness operator' corresponding to the observation or sentient experience (conscious perception) $O_k$. The quantum state $\sigma$ (which could be a pure state, a mixed state given by a density matrix, or a C*-algebra state) and the awareness operators $\{A(O_k)\}$ (along with the linear relationship above and a description of the contents of each $O_k$) are all given by the theory $T_i$.

# The Death of the Born Rule

Traditional quantum theory uses Born's rule with the probability of the observation $O_k$ being the expectation value of $A(O_k) = \mathbf{P}_k$ that is a projection operator ($\mathbf{P}_j \mathbf{P}_k = \delta_{jk} \mathbf{P}_k$, no sum over $k$) corresponding to the observation $O_k$, so

$$P(O_k | T_i) = \sigma_i[\mathbf{P}_k] = \langle \mathbf{P}_k \rangle_i.$$

Born's rule works when one knows where the observer is within the quantum state (e.g., in the quantum state of a single laboratory rather than of the universe), so that one has definite orthonormal projection operators. However, Born's rule does not work in a universe large enough that there may be identical copies of the observer at different locations, since then the observer does not know uniquely the location or what the projection operators are.

# Why Does the Born Rule Die?

Suppose there are two identical copies of the observer, at locations $B$ and $C$, that can each make the observations $O_1$ and $O_2$ (which do not reveal the location). Born's rule would give the probabilities $P_1^B = \sigma[\mathbf{P}_1^B]$ and $P_2^B = \sigma[\mathbf{P}_2^B]$ if the observer knew that it were at location $B$ with the projection operators there being $\mathbf{P}_1^B$ and $\mathbf{P}_2^B$. Similarly, it would give the probabilities $P_1^C = \sigma[\mathbf{P}_1^C]$ and $P_2^C = \sigma[\mathbf{P}_2^C]$ if the observer knew that it were at location $C$ with the projection operators there being $\mathbf{P}_1^C$ and $\mathbf{P}_2^C$.

However, if the observer is not certain to be at either $B$ or $C$, and if $P_1^B < P_1^C$, then one should have $P_1^B < P_1 < P_1^C$. However, there is no state-independent projection operator that gives an expectation value with this property for all possible quantum states.

No matter what the orthonormal projection operators $\mathbf{P}_1$ and $\mathbf{P}_2$ are, there is an open set of states that gives expectation values that are not positively weighted means of the observational probabilities at the two locations. Thus the Born rule fails in cosmology.

# Awareness Operators as Integrals of Localized Operators

The failure of the Born rule means that in a theory $T_i$, the awareness operators $A_i(O_k)$, whose expectation values in the quantum state $\sigma_i$ of the universe give the probabilities or normalized measures for the observations or sentient experiences $O_k$ as $P(O_k|T_i) \equiv \mu_i(O_k) = \sigma_i[A_i(O_k)] \equiv \langle A_i(O_k)\rangle_i$, cannot be projection operators.

However, the awareness operators could be weighted sums or integrals over spacetime of localized projection operators $\mathbf{P}_i(O_k, x)$ at locations denoted schematically by $x$, say onto brain states there that would produce the observations or sentient experiences.

# The Boltzmann Brain Problem

In local quantum field theory with a definite globally hyperbolic spacetime, any nonnegative localized operator (such as a localized projection operator) will have a strictly positive expectation value in any nonpathological quantum state. Therefore, if such a nonnegative localized operator is integrated with uniform weight over a spacetime with infinite 4-volume, it will give an awareness operator with an infinite expectation value.

If one takes the integral only up to some finite cutoff time $t_c$ and normalizes the resulting awareness operators, then for a universe that continues forever with a 3-volume bounded below by a positive value, the integrals will be dominated by times of the same order of magnitude as the cutoff time. If at late times the probability per 4-volume drops very low for ordinary observers, then most of the measure for observations will be contributed by thermal or vacuum fluctuations, so-called Boltzmann brains. That is, Boltzmann brains will dominate the measure for observations.

# The Problem with Boltzmann Brains

If Boltzmann brains dominate the measure for observations, one might ask, "So what?" Couldn't it be that our observations are those of ordinary observers? Or couldn't it be that our observations really are those of Boltzmann brains?

However, since Boltzmann brain observations are produced mainly by thermal or vacuum fluctuations, it would be expected that only a very tiny fraction of their measure would be for observations so ordered as our observations. This very tiny fraction, plus the even smaller fraction of ordered ordinary observer observations in comparison with the dominant disordered Boltzmann brain observations, would be only a very tiny fraction of the measure of all observations. Thus the normalized probability of one of our ordered observations (which we would use as the likelihood of the theory) would be highly diluted and hence much smaller than those of alternative theories in which Boltzmann brains do not dominate. If these theories do not have prior probabilities that are too small, they would dominate the posterior probabilities.

# The Problem with Boltzmann Brains, Re-Expressed

In summary, Boltzmann brain domination, which is predicted by many simple extensions of current theories (e.g., with the awareness operators or their equivalent being obtained by a uniform integration over spacetime up to a cutoff that is then taken to infinity), gives a *reductio ad absurdum* for such theories, making their likelihoods very small. Surely there are alternative theories that avoid Boltzmann brain domination without such a cost of complexity that their prior probabilities would be decreased so much as the gain in likelihoods from not having the normalized probabilities of our ordered observations highly diluted by disordered Boltzmann brain observations.

The Boltzmann brain problem is analogous to the ultraviolet catastrophe of late 19th century classical physics: Physicists then did not believe that an ideal black body in thermal equilibrium would really emit infinite power, and physicists now do not believe that Boltzmann brains really dominate observations.

# Volume Weighting versus Volume Averaging

The approach that gives 'awareness operators' as uniform integrals over spacetime of localized projection operators (or equivalently counts all observation occurrences equally, not matter when and where they occur in a spacetime) gives an especially severe Boltzmann brain problem in spacetimes with a positive cosmological constant (as ours seems to have) with the spatial hypersurfaces having 3-volumes that asymptotically grow exponentially, as in the $k = 1$ slicing of the de Sitter spacetime. At each time, counting the number or measure of observations as growing with the volume is called 'volume weighting.'

In 2008 I proposed the alternative of Volume Averaging, which gives a contribution to the measure for an observation from a hypersurface that is proportional to the *spatial density* of the occurrences of the observation on the hypersurface, rewarding the spatial frequency of observation occurrences rather than the total number that would diverge in eternal inflation as the hypersurface volume is taken to infinity.

# What Is Needed Beyond Volume Averaging

Volume Averaging ameliorates the Boltzmann brain problem by not giving more weight to individual spatial hypersurfaces at very late times when Boltzmann brains might be expected to dominate. However, when one integrates over a sequence of hypersurface with a measure uniform in the element of proper time $dt$, one gets a divergence if the time $t$ goes to infinity. One needs some suppression at late times to avoid this divergence.

# Agnesi Weighting with Volume Averaging

In 2010 I proposed Agnesi Weighting, replacing $dt$ by $dt/(1+t^2)$ where $t$ is measured in Planck units. When Agnesi Weighting is combined with Volume Averaging and a suitable quantum state such as my Symmetric-Bounce one, it appears to be statistically consistent with all observations and seems to give much higher likelihoods than current measures using the first extreme view that the measure is just given by the quantum state. It also does not require the unproven hypothesis that bubble nucleation rates for new big bangs are higher than Boltzmann brain nucleation rates or lead to measures dominated by observations of a negative cosmological constant, as do current measures using the second extreme view that the probabilities of observations are essentially independent of the quantum state.

Therefore, for fitting observations, Agnesi Weighting with Volume Averaging seems to be one of the best measures proposed so far.

# Alternatives to Agnesi Weighting

Despite its apparent fit with observations, Agnesi Weighting is admittedly quite *ad hoc*, so there is no obvious reason why it should be right. Its weighting factor favors earlier times or youngness, but in a fairly weak or light way that does not seem to be in conflict with observations, so it might be called a Utility Giving Light Youngness (UGLY) measure.

More recently I have proposed new measures depending on the Spacetime Average Density (SAD) of observation occurrences within a proper time $t$ from a big bang or bounce.

One of these is the Maximal Average Density (MAD) measure, taking the Spacetime Average Density of observation occurrences up to the time when this SAD is maximized.

Another is the Biased Average Density (BAD) measure, which uses a weighting of the SAD over time that depends upon the SAD up to that time in a continuous rather than in a step-function manner.

# A SAD Measure

# Two Extreme Alternative Views

1. The measure is determined nearly uniquely by the quantum state (e.g., Hartle, Hawking, Hertog, and/or Srednicki). For example, they apply the consistent histories or decohering histories (DH) formalism to the Hartle-Hawking no-boundary wave-function of the universe (NBWF). They calculate the probabilities of different observations within each decohering history and then multiply these by the probabilities of these histories given by the NBWF. Within each history, they tend to calculate the probability that a particular observation occurs at least once within it.

2. The measure along with eternal inflation determines the probabilities of observations essentially independently of the quantum state (e.g., Bousso, Freivogel, Guth, Linde, and/or Vilenkin). For example, they look at an apparently state-independent late-time attractor behavior of eternally inflating spacetimes and extract probabilities of observations from that, ignoring any finite number of observations at early times.

# Disadvantages of the First Extreme View

For a history with eternal inflation that gives an infinitely large spacetime, calculating the probability that a particular observation occurs at least once gives unit probability for the vast range of observations that have any positive probability density per spacetime volume of occurring. Therefore, if histories with eternal inflation dominate the quantum measure from the NBWF, a huge set of probabilities for different observations will have nearly equal values (differing only from the contributions of histories without eternal inflation).

When the probabilities of observations are normalized, the huge number of observations with nearly equal probabilities will lead to each observation having an extremely tiny normalized probability, giving a very low likelihood for any theory of eternal inflation using this particular measure.

# Disadvantages of a Modification of the First Extreme View

If one instead uses a procedure that gives different probabilities for different observations within each history but still assumes that the total probability for all the observations within that history is given purely by the quantum probability for that history, then the bulk of the total probability for all observations will still come from histories that make up the bulk of the quantum probability.

It would seem highly coincidental if these histories that dominate the probabilities tended to have strongly biophilic values for their effective coupling constants (as opposed to weakly biophilic values that allow life and observations barely to occur, at very low densities), so our observations of very strongly biophilic values of the coupling constants would be unexplained by such theories. In particular, the probability for observing coupling constants so strongly biophilic as what we observe would most likely be very low, giving extremely low likelihoods for such theories.

# Disadvantages of the Strongest Version of the Second View

The strongest version of the second extreme view is that the probabilities of observations are absolutely independent of the quantum state. This is logically possible, say by having the probabilities of observations given by the expectation values of the identity operator multiplied by coefficients that will then be the probabilities of the observations for any normalized quantum state.

But then our observations of apparent quantum effects would just be delusions, since the observations would not depend upon the quantum state at all. It seems to me much more plausible that our observations appear to depend upon the quantum state because indeed they do depend upon the quantum state. Surely the set of our observations, with its measure, could be different if the quantum state were different.

# Disadvantages of the Less-Extreme Second View

A less-extreme version of the second view would be that whether eternal inflation occurs depends upon the quantum state, but that if the quantum state is such that eternal inflation does occur, the probabilities of observations do not depend upon further details of the quantum state. This is more nearly plausible, but I find it also a rather implausible view.

If the relative probabilities of observations are given by the expectation values of positive operators that are not just multiples of the identity operator, they would generically be changed by generic changes in the quantum state. That is, it seems very hard to have the relative probabilities insensitive to generic changes in the quantum state if these relative probabilities are nontrivial linear functionals (i.e., not just the expectation values of multiples of the identity operator) of the quantum state, which to me appears to be the simplest possibility, though I do not claim to see any logical necessity against nonlinear functionals for the relative probabilities.

# Disadvantages of Current Versions of the Second View

Popular current versions of the second view include the causal diamond measure, the apparent horizon measure, the fat geodesic measure, the census taker cutoff, the scale factor time cutoff, the lightcone time cutoff, and the new lightcone time cutoff, many of which are equivalent by global-local dualities.

To avoid Boltzmann brain domination, all of these measures require that all vacua in the landscape have faster decay rates than their rates for producing Boltzmann brains. This is a very strong requirement. It may be consistent with the properties of the superstring landscape, but it may turn out to be false. Therefore, it is a disadvantage of the measures using the second extreme view that they seem to require this.

These measures also appear to predict a much higher probability of predicting a negative cosmological constant than a positive one such as what we actually observe. This suppresses the probability of our actual observation and gives these theories lower likelihoods.

# Conclusions

There are many partial answers to the measure problem in cosmology, but none of them so far is completely satisfactory. The main disadvantage is that they are all rather *ad hoc*.

Many alternatives that others have proposed seem to lead also to excessively low probabilities for our observations that include such things as high order, strongly biophilic values for the effective coupling constants, and a positive cosmological constant. Alternatives that avoid Boltzmann brain domination require that all vacua in the landscape decay faster than Boltzmann brain production. Measures that I have proposed are also *ad hoc* but avoid these problems.

I have now made proposals that are MAD, BAD, and UGLY. I am still looking for one that is GOOD in a supreme way of giving a high posterior probability. However, if GOOD were interpreted as simply meaning Great Ordinary Observer Dominance, one could say that the MAD, the BAD, and the UGLY are all GOOD.

# Another SAD Measure