

Properties of the scale factor measureRaphael Bousso,^{*} Ben Freivogel,[†] and I-Sheng Yang[‡]*Department of Physics and Center for Theoretical Physics, University of California, Berkeley, California 94720, USA
and Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*

(Received 14 January 2009; published 17 March 2009)

We show that in expanding regions, the scale factor measure can be reformulated as a local measure: Observations are weighted by integrating their physical density along a geodesic that starts in the longest-lived metastable vacuum. This explains why some of its properties are similar to those of the causal-diamond measure. In particular, both measures are free of Boltzmann brains, subject to nearly the same conditions on vacuum stability. However, the scale factor measure assigns a much smaller probability to the observed value of the cosmological constant. The probability decreases further, similar to the inverse sixth power of the primordial density contrast, if the latter is allowed to vary.

DOI: 10.1103/PhysRevD.79.063513

PACS numbers: 11.27.+d

I. INTRODUCTION**A. The measure problem**

Similar to the cosmological constant problem [1–3], the measure problem arises purely within the regime of validity of semiclassical gravity. All that is needed is a long-lived vacuum state [4], or a sufficiently flat scalar field potential [5,6], with positive vacuum energy. Under these conditions, a finite spatial region will inflate eternally, generating an unbounded four-volume. All possible events will occur infinitely many times, and a cutoff is needed to compute probabilities.

Unlike the cosmological constant problem, the measure problem leaves a loophole: it is possible that the conditions for eternal inflation do not actually occur in nature. But increasing evidence indicates otherwise.

Slow-roll inflation is the dominant paradigm explaining the origin of structure and the large-scale homogeneity and flatness of our Universe. If we are prepared to believe that the moderately fine-tuned scalar field potential necessary for driving slow-roll inflation can arise in nature, it is hard to imagine that the more generic feature of a local minimum cannot exist.

Moreover, the observed Universe has positive vacuum energy [7–9]. Unless our vacuum is unstable on a time scale of order 10×10^9 years—which would require remarkable tuning—it alone suffices to generate eternal inflation.

Finally, the only extant explanation [10–12] of the smallness of the observed dark energy—the cosmological constant problem—is the existence of a multidimensional landscape of metastable vacua in string theory. This is empirical evidence for string theory, and, in particular, for eternal inflation.

A theory of everything, if it gives rise to eternal inflation, should eventually allow us to derive a unique prescription

for computing probability amplitudes from first principles. Yet our understanding of string theory, especially in cosmological settings, remains woefully incomplete. For now, a top-down solution to the measure problem seems elusive. Thus, we advocate following the traditional, phenomenological approach.

B. Phenomenological approach

As with any other theory, a compelling measure should be reasonably well defined, simple, and general. In tackling a problem so vast and unfamiliar, it is natural to seek more specific guiding principles. For example, lessons from the black hole information paradox motivated the causal-diamond measure [13,14]. We would be ill advised, however, to turn our intuition into dogma, insisting absolutely on *theoretical* properties that “any reasonable” measure “must” obey. Not only would we run the danger of putting in by hand the answers we wish to get; worse, our wishes may be misguided. Surely, for example, any reasonable measure must reward the volume expansion during slow-roll inflation in any given vacuum? In fact, this intuitive requirement invites conflict with experiment [15,16]. The causal-diamond measure abandoned volume weighting, and the requirement seems now to have lost its dogmatic status [17,18].

This brings us to the one property of a measure that we can insist on absolutely: that its *experimental* predictions not conflict with observations. In fact, many innocent-looking measures do conflict with observation, and violently so. Some conflicts are extremely robust, arising almost independently of the properties of the underlying vacuum landscape. Thus, they allow us to falsify measures even while we still have much to learn about the landscape.

For example, the proper time cutoff [19–24] predicts a very hot universe [25–32] with probability 1 (the “Boltzmann babies” or youngness paradox). Measures that involve counting the number of observers per baryon [33–39] predict an empty, cold universe [40,41] with probability 1 (the “Boltzmann brain” paradox). Such para-

^{*}bousso@lbl.gov[†]freivogel@berkeley.edu[‡]jinking@berkeley.edu

doxes are important tools for testing and eliminating measure theories. We say “paradox,” as if there were any doubt about the culprit. In fact, the above paradoxes represent fatal failures: the measure assigns zero probability to the observations we actually make. Such measures are experimentally ruled out and must be discarded.

Another useful test is the more aptly named “ Q -catastrophe” [15,16] alluded to earlier. In measures that reward the volume expansion during inflation, such as Ref. [38], inflationary model parameters generically receive exponential pressure toward extreme values. In this case, anthropic constraints do not suffice to explain the moderate values we observe for, say, the primordial density perturbation, Q .

Other problems are more subtle, or depend on the detailed structure of the vacuum landscape. The “staggering problem” can arise in the measure of Ref. [38]: for some landscape models, the measure assigns such unequal probabilities to the cosmological production of different vacua that most observers live in extreme environments where their existence is an unlikely fluctuation [42–46]. In particular, the cosmological constant problem cannot be solved in this case.

Finally, the value it predicts for the cosmological constant, Λ , is an important test of any measure. This particular observable is special for two reasons: Its statistical distribution in the landscape is understood well enough to make detailed quantitative predictions. And its value has been measured:

$$\Lambda = 1.5 \times 10^{-123}, \quad (1)$$

in Planck units.

At present, the causal-diamond measure is the most successful proposal phenomenologically. It avoids Boltzmann babies, Boltzmann brains, and the staggering problem. (The absence of Boltzmann brains requires that all vacua decay faster than they produce such rare fluctuations [41]—a nontrivial condition on the vacuum landscape, which, however, may well be satisfied by the string landscape [47].) It does not reward volume expansion, thus avoiding the Q -catastrophe and raising the possibility of the detection of open spatial curvature by future experiments [48].

The causal-diamond measure predicts a value of the cosmological constant consistent with observations. That is, the observed value lies close to the mean and is highly typical in the predicted probability distribution [49]. This successful prediction is robust against variations of Q [49,50].

C. Summary and outline

In this paper, we investigate the scale factor measure [17], which cuts off the Universe at the time η when a (randomly chosen) congruence of timelike geodesics has expanded by a volume factor $\exp(3\eta)$ along each geodesic.

Relative probabilities of different observational outcomes can be defined by computing the ratios of the number of times such outcomes occur in the regulated four-volume, and then taking the regulator $\eta \rightarrow \infty$.

Section II contains review material and establishes most of our notation. In Sec. II A, we give a detailed definition of the scale factor measure. A separate prescription is needed in regions where geodesics contract and intersect with each other. We review the prescription chosen by De Simone *et al.* [17], who were the first to formulate one carefully. In Sec. II B, we collect other useful definitions and results; in particular, we review the solution found by Garriga *et al.* [38] for the volume distribution of different vacua, which also applies to the scale factor measure.

In Sec. III, we determine the shape of the scale factor cutoff hypersurface in a homogeneous, isotropic, open universe formed by bubble nucleation inside a parent de Sitter vacuum. This is nontrivial, because the local scale factor along each geodesic, and the average scale factor in a bubble universe, are two different objects.

In Sec. IV, we compute the number of observations below the cutoff surface by integrating over all bubbles produced prior to the cutoff. We begin, in Sec. IV A, by approximating the bubble interior metric as homogeneous, *ignoring local gravitational collapse such as occurs during galaxy formation*. We find that in this “no-collapse approximation,” the scale factor measure leads to a very simple result: The probability of an observation in some vacuum is proportional to the product of the probability that a given geodesic in the congruence will enter that vacuum, times the number density, per physical volume, at which observations occur. It does not depend directly on the time at which they occur, nor on the amount of volume expansion since the bubble was produced.

Our result generalizes the results of Ref. [17], encompassing in a single formula the following important properties of the (no-collapse) scale factor measure established by De Simone *et al.*¹:

- (1) No youngness paradox, since the physical density of Boltzmann babies is negligible.
- (2) No reward for excessive volume expansion during slow-roll inflation: Once inflation has made the Universe flat enough for structure formation, any extra e -foldings will not affect the physical density of observers. Thus, the scale factor measure avoids the Q -catastrophe.
- (3) The probability distribution for Λ is in excellent agreement with the observed value, Eq. (1): If Λ

¹De Simone *et al.* applied their prescription for collapsing regions to the homogeneous collapse of bubbles with negative cosmological constant, but implicitly used a no-collapse approximation elsewhere, ignoring the turnaround and collapse of geodesics in structure-forming regions. This explains any discrepancies between our papers, in particular, our less favorable conclusion concerning the value of Λ predicted by the scale factor measure.

had dominated much before the time when observations are made (i.e., now), galaxies would now be exponentially dilute and the average density of observations would be highly suppressed. This result is stable against variations of the primordial density contrast, Q .

A simple result should have a simple explanation, which we provide in Sec. IV B: In the no-collapse approximation (i.e., if all geodesics in the congruence are always expanding), the scale factor measure is equivalent to the following prescription²: Consider a single geodesic that starts out in the longest-lived metastable vacuum of the landscape, $*$. Compute the expected number of observations of type μ , $d\langle N_\mu \rangle$, occurring along a fixed physical volume dV transverse to the geodesic. The relative probability of observations μ and ν is $d\langle N_\mu \rangle / d\langle N_\nu \rangle$. In practice, this can be accomplished by summing the probabilities that the geodesic will enter each vacuum i multiplied by the physical density of observations of type μ in vacuum i , so this more general formula reduces to our earlier result.

We thus reformulate the no-collapse scale factor measure as a local measure. By this we mean a measure that involves averaging over the statistical ensemble defined by the different possible histories along a single worldline, without necessarily assembling them into a global geometry. It would be natural to use the local formulation as the general definition of the scale factor measure, since it can be applied to collapsed regions without requiring additional rules.

The causal-diamond measure is another example of a local measure in this sense. It differs from the no-collapse scale factor measure only in two ways: (1) the transverse volume included along the geodesic is not constant, but is set by the size of a causally connected region; and (2) initial conditions are not determined by the causal-diamond measure, but are considered a logically independent question. This explains the pattern of similarities between the measures that have emerged, but it also clarifies how they differ.

In Sec. IV C, we go beyond the no-collapse approximation and investigate how the De Simone *et al.* prescription for collapsed regions affects the formulas for probabilities in the scale factor measure. We find that it can be incorporated by a simple substitution: Instead of using the physical density of observations, what matters is the (potentially far greater) density those observations *would have had*, if they had occurred at the time when the first structures formed that would later merge into the objects hosting the observations. In other words, we can incorporate gravitational collapse by mapping each observation back to the earliest time when a geodesic on which it lies began to collapse. Our own observations, for example, are thus condensed by

inverse of the expansion factor of our Universe since the formation of the first dark matter halos.

Since neither the inflationary era nor the early postinflationary universe contains collapsed regions, this modification has no bearing on the results (1) and (2) above, concerning the youngness paradox and the Q -catastrophe. In Sec. V, we investigate how the inclusion of gravitational collapse affects the prediction for the cosmological constant.

We begin by reviewing the predictions for the cosmological constant obtained in various measures: the observers-per-baryon prescription (Sec. VA), the causal-diamond measure (Sec. VB), and the no-collapse scale factor measure (Sec. VC1). In Sec. VC2, we consider the scale factor measure, with the prescription of Ref. [17] for collapsed regions. We find that the cosmological constant is not set by the time scale when observations are made—as it is in the causal-diamond measure and, apparently, in nature. Rather, its value is controlled by the time scale of structure formation. This means that the scale factor measure predicts a value that is up to 5000 times larger than the observed value. By adding more specific anthropic assumptions, the discrepancy can be mitigated, but the observed value remains somewhat atypical.

If the primordial density contrast, Q , is allowed to vary, we find that the preferred value of Λ scales like Q^3 , and the associated probability like Q^6 . This means, for example, that a universe with Q 3 times as large, and Λ 27 times as large, is 729 times as likely. It is difficult to see how anthropic constraints or prior distributions for Q can overcome a pressure so strong. In Sec. VD, we discuss possible modifications of the treatment of collapsed regions in the scale factor measure, which might improve these problematic predictions.

In Sec. VI, we investigate the probability for Boltzmann brains in the scale factor measure.³ Boltzmann brains are observers that arise from rare thermal fluctuations, at a superexponentially small rate per unit four-volume. Some measures overcompensate for this suppression by including superexponentially large regions of empty de Sitter space [33–39] for every region containing ordinary observers. Then the vast majority of observations are made by Boltzmann brains. Since almost none of the observations made by Boltzmann brains agree with our observations, these measures are ruled out. The causal-diamond measure was shown in Ref. [41] to favor ordinary observers, if the decay rate of every vacuum in the landscape is greater than the rate for Boltzmann brains. Recent evidence suggests that this nontrivial condition may be satisfied in the string theory landscape [47].

Our local formulation of the scale factor measure makes it clear that similar conditions will play a role in the scale

²We are using here the approximation of [42]; we will be more precise in the body of the paper.

³A. Vilenkin and collaborators have independently analyzed this question. We understand that their results will appear simultaneously or in the near future.

factor measure. The transverse volume along the defining geodesic is the same in all vacua, whereas in the causal-diamond measure, it is set by the cosmological constant. Since the latter varies at most over an exponentially, but not a superexponentially large range, this difference is negligible in the context of Boltzmann brains.

The only remaining potential difference arises from the initial conditions on the geodesic. In the causal-diamond measure, it was reasonable to assume that these conditions do not pick out vacua with unnaturally small cosmological constant (a necessary condition even for Boltzmann brains), so the initial vacuum received no attention in the analysis of Ref. [41]. In the scale factor measure, however, the initial vacuum is the longest-lived de Sitter vacuum, which might well have a small cosmological constant.

Here, we refine the analysis of Ref. [41] in two respects. In Sec. VI A, we include the contributions from the initial vacuum to the number of Boltzmann brains. In Sec. VI B, we demonstrate (under plausible assumptions on the structure of the landscape) that the production rates of different vacua will not invalidate our earlier criterion for the dominance of ordinary observers, and we find that it is augmented only by the condition that the initial vacuum be completely unable to produce Boltzmann brains (independently of its decay rate). We argue that this condition is likely to be satisfied. Thus, the scale factor measure and the causal-diamond measure are virtually equivalent for the purpose of Boltzmann brains.

II. DEFINITION AND RATE EQUATIONS

In this section we define the scale factor cutoff and review some of its basic properties.

A. Definition of the scale factor cutoff

One approach to regulating the infinities in eternal inflation is through a smooth congruence of timelike geodesics orthogonal to a (nearly arbitrary) finite spacelike surface Σ_0 . The idea is to use the geodesic congruence to define a sequence of cutoff hypersurfaces Σ_η . Then one computes the number N_i of observations of type i , O_i , in the four-volume between Σ_0 and Σ_η . The relative probability of two observations is defined by their relative abundance in the limit where the cutoff is taken away:

$$\frac{\mathcal{P}(O_i)}{\mathcal{P}(O_j)} \equiv \lim_{\eta \rightarrow \infty} \frac{N_i}{N_j}. \quad (2)$$

A simple choice for constructing Σ_η would be to follow each geodesic for the same proper time. But the resulting measure is ruled out at an overwhelming confidence level, and independently of the details of the underlying theory: It predicts that we should observe a much hotter universe [25–32].

1. The no-collapse scale factor measure

A different cutoff on the congruence was recently defined by De Simone *et al.* [17], building on earlier work.⁴ In this proposal Σ_η is, roughly, a surface of constant local scale factor. The *scale factor time* is defined by integrating the local expansion rate relative to infinitesimally nearby geodesics, along each geodesic:

$$\eta(\mathbf{x}, t) \equiv \int_0^t \frac{\theta}{3} dt', \quad (3)$$

where t is the proper time along the geodesic labeled by \mathbf{x} , $\theta = \nabla^\mu \zeta_\mu$ is the expansion, and ζ_μ is the 4-velocity vector field tangent to the congruence. The *local scale factor* is defined by

$$A(\mathbf{x}, t) \equiv \exp \eta(\mathbf{x}, t). \quad (4)$$

Intuitively, these quantities measure the growth of a local volume element δV spanned by infinitesimally nearby geodesics [55]:

$$\theta = 3 \frac{d\eta}{dt} = 3 \frac{dA/dt}{A} = \frac{d(\delta V)/dt}{\delta V}. \quad (5)$$

2. Collapsed regions and other ambiguities

However, Σ_η cannot be defined simply as a surface of constant η or A . In collapsing regions, such as pockets with a negative cosmological constant, or structure-forming regions, geodesics will cease to expand and begin to approach each other. Then $\theta < 0$, and by Eq. (3), the scale factor time decreases locally toward the future. Unless a singularity is encountered, the focusing theorem guarantees that geodesics will eventually encounter caustics: they will intersect with infinitesimally neighboring geodesics. Thereafter, they begin to expand again, and the scale factor time increases once more.

This introduces two ambiguities: First, a given scale factor need not be reached precisely once along each geodesic. It may never be reached, or it may be reached more than once. Which (if any) occurrence defines the actual cutoff? Second, a given event may be threaded by more than one geodesic. Should we count such events once, or multiple times?

⁴Scale factor time is considered in some of the earliest literature on (slow-roll) eternal inflation, which focuses on the distribution of different field values [20–23]. The measure defined by De Simone *et al.* is similar to the “pseudocomoving volume-weighted measure” [24], which stops short of an explicit definition of the probability of different observations, such as Eq. (2), Eq. (1) of Ref. [32] or Eq. (1) of Ref. [17], and of a prescription for collapsing geodesics. It shares some properties with other measures [38,51–54], in which the scale factor (in the guise of approximately equivalent formulations) regulates bubble abundances but different (or no) cutoffs regulate the number of observers per bubble.

To resolve these ambiguities, De Simone *et al.* [17] propose that an event should be included “if it lies on any geodesic prior to the first occurrence” of the specified cutoff on that geodesic. In other words, Σ_η is the hypersurface that maximizes the four-volume V_4 of the congruence subject to the constraint that at every point $p \in V_4$ lies on at least one geodesic at scale factor time less than η .

This definition implies that if the cutoff is never reached along some geodesic, all events on the geodesic are included (no future cutoff). The formulation of De Simone *et al.* suggests, moreover, that any event should be counted at most once; we will adopt this definition here. Note that with this definition, Σ_η need not be a spacelike hypersurface. Rather, it becomes timelike near collapsing geodesics, spiking up toward the future, as shown in Fig. 1.

The prescription for collapsed regions chosen by De Simone *et al.* is not the only possible one. Indeed, we will find that it has phenomenological disadvantages (Sec. V), and we will suggest other possible definitions. In particular, the local formulation of the scale factor measure (Sec. IV B) applies without modification both in collapsing and expanding regions. However, we will not explore these alternatives in detail; we focus on the definition of Ref. [17] in this paper.

There is another ambiguity, however, which has yet to be resolved. When a new vacuum with lower cosmological constant is formed by the Coleman-DeLuccia process, there is a quantum region, where the geometry (in the instanton approximation) jumps sharply and the continuation of geodesics is not well defined. For downward

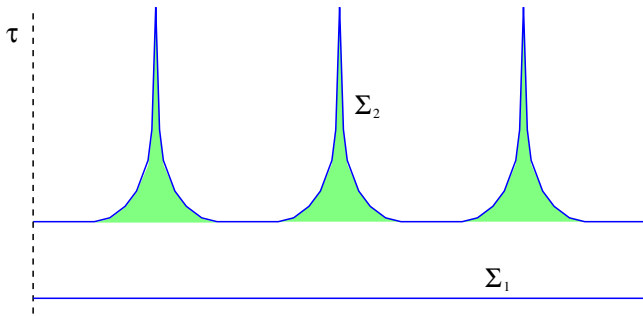


FIG. 1 (color online). Evolution of the scale factor cutoff during structure formation. The constant scale factor time surface Σ_1 lies in the early, approximately homogeneous universe and coincides with a surface of constant FRW time τ . As density perturbations grow, some geodesics decouple from the Hubble flow, stop expanding, and become trapped in collapsed regions such as galaxies. If the scale factor cutoff exceeds the largest scale factor ever reached along such geodesics, then the rule of De Simone *et al.* requires that their entire future evolution be included. (A similar result is obtained if the local formulation of the scale factor measure is used as a general definition; see Sec. V D.) Therefore, the later cutoff surface Σ_2 no longer agrees everywhere with a constant FRW time surface; it includes the entire future of collapsed regions (gray/green), which show up as spikes since the figure is drawn in comoving coordinates.

tunnelings, one can reasonably ignore this problem, since the initial radius of the bubble wall can be much smaller than the de Sitter radius, so only a negligible fraction of geodesics entering the new bubble pass through the bubble interior at the time of nucleation. Moreover, geodesics quickly become comoving in regions with slow-roll inflation, which are of the greatest interest.

De Sitter vacua may also tunnel upward, to vacua with larger cosmological constant. Even though upward tunnelings are very suppressed, the scale factor measure implies that the majority of observers live on worldlines which have been through upward tunnelings [43,46]. But there is no approximate classical geometry describing an upward tunneling. There is no natural way, therefore, of determining the tangent vectors of the portion of the geodesic congruence entering the new bubble.

For the purpose of deriving the rate equations, in Sec. II B, we will follow Ref. [38] and side step this issue by making mathematically convenient assumptions. Once we count observers in Sec. IV, the ambiguity will not reenter: we are only interested in the most recent tunneling, which is necessarily downward for producing ordinary observers.

B. Rate equations and solutions

Before we can count observers in vacuum i , we will need to know the number $n_i(\eta)$ of such bubbles below the cutoff η . This, in turn, requires knowing the physical volume $V_i(\eta)$ occupied by every vacuum. The evolution and distribution of these volumes is determined by the rate matrix Γ_{ij} describing the number of bubbles of vacuum i forming per unit four-volume of vacuum j .

For nonterminal vacua (i.e., those with positive cosmological constant Λ_i), the following definitions will be convenient: the expansion at late times,

$$H_i \equiv \left(\frac{\Lambda_i}{3}\right)^{1/2}; \quad (6)$$

the total decay rate of vacuum i per unit four-volume,

$$\Gamma_j \equiv \sum_i \Gamma_{ij}; \quad (7)$$

the dimensionless decay rate from vacuum j to i ,

$$\kappa_{ij} \equiv \frac{4\pi}{3} \frac{\Gamma_{ij}}{H_j^4}; \quad (8)$$

the total dimensionless decay rate of vacuum j ,

$$\kappa_j \equiv \sum_i \kappa_{ij}, \quad (9)$$

and the branching ratio matrix

$$\beta_{ij} \equiv \kappa_{ij}/\kappa_j. \quad (10)$$

During a time interval $d\eta$, the volume V_i in vacuum i will increase due to intrinsic expansion, due to the produc-

tion of new bubbles of type i , and due to the expansion of the bubble walls into the parent vacua. It will decrease due to decay into other vacua, and due to the growth of such bubble walls after they are produced.

Treating the motion of domain walls in detail is cumbersome, and it is unnecessary if all metastable vacua are long lived ($\kappa_i \ll 1$). This is a reasonable assumption, since $\kappa_i \gg 1$ conflicts with the notion of a vacuum, and $\kappa_i \sim 1$ requires fine-tuning. Such vacua will be too rare in the landscape to play an important dynamical role.

Thus, most of the four-volume of each bubble type is empty de Sitter space, and we can neglect transient effects right after bubble nucleation. One transient is the period between bubble creation and vacuum domination. Therefore, the expansion θ of geodesics in vacuum j can be approximated by the Hubble constant at late times, $\theta \approx H_j \equiv (\Lambda_j/3)^{1/2}$. The four-volume in vacuum j added during the scale factor time $d\eta$ is thus $V_j H_j^{-1} d\eta$.

Another transient is the bubble wall expansion. A bubble nucleated at the time η^{nuc} will eventually occupy a comoving volume in the geodesic congruence that would (in the absence of the decay) have originated from a ball of physical radius H_j^{-1} in the parent vacuum j at the time η^{nuc} . This asymptotic comoving size is reached, to accuracy of order $\exp(\eta - \eta^{\text{nuc}})$, after only a few units of scale factor time. Thus we make a small error by anticipating this growth: We ascribe a physical volume $4\pi/3H_j^3$ to the new vacuum i already at the nucleation time η^{nuc} , and in exchange neglect the bubble wall growth [38].⁵

With these approximations, vacuum i gains

$$d^+ V_i = 3V_i d\eta + \sum_j \left(V_j \frac{d\eta}{H_j} \right) \Gamma_{ij} \left(\frac{4\pi}{3H_j^3} \right) \quad (11)$$

and loses

$$d^- V_i = \sum_j \left(V_j \frac{d\eta}{H_j} \right) \Gamma_{ji} \left(\frac{4\pi}{3H_j^3} \right) \quad (12)$$

in physical volume per scale factor time. Combining inflow and outflow yields the Fokker-Planck equation

$$\frac{dV_i}{d\eta} = 3V_i - \kappa_i V_i + \sum_j \kappa_{ij} V_j. \quad (13)$$

Reference [38] rigorously derives the solution to this equation. Only the behavior of de Sitter vacua will be relevant here. For generic initial conditions with some support in de Sitter vacua, the solution approaches attractor behavior at late times. The volume in de Sitter vacuum i is

⁵This behavior of the bubble wall applies only to downward transitions. During upward transitions, the behavior of the congruence defining the scale factor measure is not well defined in semiclassical gravity. Following Ref. [38], we will choose *ad hoc* to use the same rule in this case.

$$V_i(\eta) = C s_i e^{\gamma \eta}. \quad (14)$$

Here, C is a constant with the dimension of volume that depends on the initial conditions but drops out in all normalized probabilities;

$$\gamma \equiv 3 - q; \quad (15)$$

q is the smallest-magnitude negative eigenvalue of the dimensionless flow matrix $\kappa_{ij} - \delta_{ij}\kappa_i$; and s_j is the associated eigenvector:

$$\kappa_{ij} s_j = (\kappa_i - q) s_i \equiv p_i, \quad (16)$$

where we have defined the vector p_i for later convenience.

To exponentially good approximation [42], the eigenvector is dominated by the longest-lived de Sitter vacuum, $*$:

$$s_j \approx \delta_{j*}; \quad (17)$$

and its eigenvalue is equal to the dimensionless decay rate of this vacuum,

$$q \approx \kappa_* \ll 1. \quad (18)$$

Thus, q is exponentially small in a realistic landscape. By Eq. (2), this attractor solution is all we need to compute probabilities.

A number of other results will be useful below, and we collect them here. The expected number of times a world-line starting in vacuum o will pass through vacuum i can be obtained by summing over the whole branching tree [13]:

$$e_{io} = \delta_{io} + \sum_{\text{paths from } o \text{ to } i} \beta_{i i_n} \beta_{i_n i_{n-1}} \cdots \beta_{i_1 o}, \quad (19)$$

where $i_1 \cdots i_n$ are intermediate vacua connecting o and i , and the branching ratios β_{ij} were defined in Eq. (10). This can be written in matrix form as

$$e_{io} = \left(\sum_{n=0}^{\infty} \mathbf{B}^n \right)_{io}. \quad (20)$$

In the scale factor measure, the initial vacuum is the $*$ vacuum, and we will denote

$$e_i \equiv e_{i*}. \quad (21)$$

If the initial vacuum has relatively small cosmological constant, as one might expect for the longest-lived vacuum, then it is unlikely to be reentered later in the decay chain. Then the sum in Eq. (20) converges rapidly; in particular,

$$e_* \approx 1. \quad (22)$$

In Ref. [44], the vectors p_i and e_i were shown to be closely related:

$$p_i = q\left(\sum_{n=1}^{\infty} \beta^n\right)_{ij} s_j \approx q\left(\sum_{n=1}^{\infty} \beta^n\right)_{i*} \approx qe_i \quad (i \neq *). \quad (23)$$

Note, however, that $p_* \neq qe_*$.

III. THE CUTOFF HYPERSURFACE IN A HOMOGENEOUS BUBBLE

In this section, we investigate the evolution of the cutoff hypersurface in a single bubble of an arbitrary vacuum, nucleated inside a parent vacuum with positive cosmological constant $3H^2$. We will drop the index i while discussing a single vacuum. For now, we will ignore inhomogeneities, such as density perturbations and local gravitational collapse.

Assuming that the tunneling process is very rare, the parent vacuum will have existed for a long time before the nucleation event. This implies [56] that we can take the geodesic congruence to be comoving in the flat slicing of de Sitter space,

$$ds^2 = -dt^2 + e^{2Ht}(dr^2 + r^2 d\Omega^2). \quad (24)$$

We can think of r as labeling a particular shell in the congruence. The geodesics between r and $r + \delta r$ span a volume element

$$\delta V(t, r) = \frac{dV}{dr} \Big|_{t=\text{const}} \delta r = 4\pi e^{2Ht} r^2 \delta r \quad (25)$$

in the parent vacuum.

In the homogeneous approximation, the metric inside the bubble is described by an open Friedmann-Robertson-Walker (FRW) geometry

$$ds^2 = -d\tau^2 + a^2(\tau)(d\xi^2 + \sinh^2 \xi d\Omega_2^2). \quad (26)$$

The comoving geodesics from the host vacuum continue into the bubble along nontrivial trajectories, defining a second coordinate system (t, r) inside the bubble, where t is the proper time along the geodesic passing through an event, and r is the radial position the geodesic had before entering the bubble. We will now review the coordinate transformation between (τ, ξ) and (t, r) , derived in Ref. [32].

Without loss of generality, one can choose coordinates so that the bubble is nucleated at $t = 0, r = 0$. Consider a comoving geodesic which passes through the event $t = 0, r$. If after a proper time t (as measured along the geodesic) it has passed into the bubble, its coordinates (τ, ξ) will be

$$\begin{aligned} \xi(t, r) &= -\log(1 - Hr) + \mathcal{O}\left(\frac{1}{H\tau}\right) \\ \tau(t, r) &= t - \frac{\xi + e^{-\xi} - 1}{H} + \mathcal{O}\left(\frac{1}{H\tau}\right). \end{aligned} \quad (27)$$

Note that H in these formulas refers to the constant expansion rate of the parent de Sitter space. The above equations

are valid for geodesics that have spent more than a proper time H^{-1} inside the bubble.⁶ Physically, this result shows that the geodesics become comoving in the open coordinates after one outside Hubble time, and thereafter the proper time along the geodesics increases at the same rate as the open FRW time.

To construct the cutoff hypersurface, one must determine the scale factor time η as a function of the bubble coordinates τ and ξ . By Eq. (14),

$$\eta(\eta^{\text{nuc}}, \tau, \xi) = \eta^{\text{nuc}} + \frac{1}{3} \log\left(\frac{\delta V[r(\tau, \xi), t(\tau, \xi)]}{\delta V(r, 0)}\right), \quad (28)$$

where η^{nuc} is the scale factor time at $t = 0$, when the bubble is nucleated. The volume element δV orthogonal to the geodesics lies on a hypersurface of constant proper time t [55]. By Eq. (26),

$$\begin{aligned} \delta V &= 4\pi a^2(\tau) \sinh^2 \xi \left[a^2(\tau) \left(\frac{d\xi}{dr}\right)_{t=\text{const}}^2 - \left(\frac{d\tau}{dr}\right)_{t=\text{const}}^2 \right]^{1/2} \delta r \\ &= 4\pi a(\tau)^2 \sinh^2 \xi \left[a(\tau)^2 - \left(\frac{1 - e^{-\xi}}{H}\right)^2 \right]^{1/2} H e^\xi \delta r, \end{aligned} \quad (29)$$

where we have used Eq. (27). We are interested in events much later than the outside Hubble time, so $a^2 \gg H^{-2}$, and we can drop the second term in the square root:

$$\delta V = 4\pi H a^3(\tau) e^\xi \sinh^2 \xi \delta r. \quad (30)$$

After using Eq. (27) to eliminate r from Eq. (25), Eq. (28) yields the scale factor time inside the bubble:

$$\begin{aligned} \eta(\eta^{\text{nuc}}, \tau, \xi) &= \eta^{\text{nuc}} + \log[Ha(\tau)] \\ &\quad + \frac{2}{3} \left[\xi + \log\left(\cosh\frac{\xi}{2}\right) \right]. \end{aligned} \quad (31)$$

IV. COUNTING ORDINARY OBSERVERS

In this section, we apply the scale factor cutoff to counting observations in an eternally inflating multiverse with multiple vacua. We exclude, for now, observations resulting from violations of the second law (Boltzmann brains, treated in Sec. VI). Initially, we will imagine that all observations (if any) in a bubble of type i are made instantaneously⁷ at the FRW time τ_i^{obs} after the vacuum is produced, with number density ρ_i^{obs} per unit physical

⁶We also assume that the cosmological constant inside is much smaller than outside. Corrections due to such approximation are not included in Eq. (27). However, when the inside and outside cosmological constants are the same, one can derive an exact formula which is equivalent to Eq. (31) for all physical questions we considered in this paper.

⁷In our investigation of the proper time cutoff [32], information about the temporal distribution of observations, $f_i(\tau)$, was crucial to demonstrating the youngness paradox. In the scale factor measure, however, there is no youngness paradox [17], and for the purposes of this paper, it suffices to use only $g_i \equiv \int_0^\infty f_i(\tau) d\tau$ and $\tau_i^{\text{obs}} \equiv g_i^{-1} \int_0^\infty f_i(\tau) \tau d\tau$ as input parameters.

volume:

$$dN_i = \rho_i^{\text{obs}} dv_{\text{ph}}. \quad (32)$$

We will assume, moreover, that these parameters do not depend on the parent vacuum from which i is entered. These assumptions will simplify our treatment, and it will be easy to drop them in the end and state a more general result, Eq. (43).

Our treatment will go beyond that of De Simone *et al.* [17] in that we do not approximate the bubble interior as flat and homogeneous. Our detailed treatment of collapsing regions, in Sec. IV C, has significant implications, invalidating some of the conclusions of Ref. [17] (see Sec. V).

A. Counting observations in the no-collapse approximation

Now we will compute the total number of observations, $N_i(\eta)$, performed in vacua of type i below the cutoff, η . By Eq. (32), this amounts to computing the total physical volume of the $\tau = \tau_i^{\text{obs}}$ hypersurfaces below the cutoff. In any single bubble, a shell of comoving radius ξ and width $d\xi$ contributes a physical volume

$$dv_i^{\text{ph}} = 4\pi(a_i^{\text{obs}})^3 \sinh^2 \xi d\xi, \quad (33)$$

but only if the bubble was nucleated early enough for this volume to be included below the cutoff.

By Eq. (31), the latest nucleation time that allows observers at radius ξ to contribute is

$$\eta_{ij}^{\text{nuc}}(\eta, \xi) = \eta - \log[H_j a_i^{\text{obs}}] - \frac{2}{3} \left[\xi + \log \left(\cosh \frac{\xi}{2} \right) \right]. \quad (34)$$

Note that this time depends on the parent vacuum, j . Hence,

$$N_i(\eta) = \sum_j \int_0^\infty d\xi n_{ij}[\eta_{ij}^{\text{nuc}}(\eta, \xi)] \frac{dv_{\text{ph}}}{d\xi}(\xi) \rho_i^{\text{obs}}, \quad (35)$$

where $n_{ij}(\eta')$ is the total number of bubbles of type i nucleated inside vacuum j by the time η' .

Using

$$dn_{ij} = \Gamma_{ij} V_j d\eta / H_j \quad (36)$$

Eqs. (8) and (14) imply

$$n_{ij}(\eta_{ij}^{\text{nuc}}) = \frac{3C}{4\pi\gamma} H_j^3 \kappa_{ij} s_j \exp(\gamma \eta_{ij}^{\text{nuc}}). \quad (37)$$

Combining the above equations, we find

$$N(\eta) = \frac{3C}{4\pi\gamma} \left[\int_0^\infty d\xi 4\pi \sinh^2 \xi (\exp \xi)^{-2\gamma/3} \times \left(\cosh \frac{\xi}{2} \right)^{-2\gamma/3} \right] e^{\gamma \eta} (a_i^{\text{obs}})^q \left(\sum_j H_j^q \kappa_{ij} s_j \right) \rho_i^{\text{obs}}. \quad (38)$$

The ξ integral is clearly convergent, with support concen-

trated near $\xi \sim O(1)$. (For small q , its value approaches $4\pi/3$.) Physically, this shows that mainly the central curvature volume of the infinite open bubble geometry contributes in the scale factor measure. Therefore, the interesting results of Garriga, Guth, and Vilenkin [56] regarding worldlines at large ξ are not relevant in this measure.

Since q is exponentially small, H_j^q and $(a_i^{\text{obs}})^q$ can safely be neglected. By the discussion at the end of Sec. II, $\kappa_{ij} s_j = q e_i$. Thus we obtain the probability

$$\mathcal{P}_i \propto e_i \rho_i^{\text{obs}} \quad (39)$$

for observing vacuum i . The \propto notation indicates that the universal factor $\approx C q e^{\gamma \eta} / \gamma$ has been dropped [since, by Eq. (2), it does not affect relative probabilities], though the probabilities have not been normalized.

Equation (39) immediately implies several key properties of the scale factor measure as follows:

- (i) The probability for observing vacuum i is simply the product of the number of times a typical worldline can be expected to enter i -bubbles, e_i , and the density, ρ_i^{obs} , of observations per physical volume on the homogeneous time slice τ_i^{obs} at which they are performed.
- (ii) The probability of an observation depends on its FRW time, τ^{obs} , and on the FRW expansion factor at that time, a^{obs} , *only* through ρ^{obs} .
- (iii) As a special case, we recover an important result of De Simone *et al.* [17]: As in the causal-diamond measure, the probability of a vacuum is insensitive to the volume expansion factor during slow-roll inflation. This is a desirable property, because it avoids the Q -catastrophe [15,16]—the overwhelming pressure toward extreme (and counterfactual) inflationary parameter values that results when exponential volume factors are rewarded. Moreover, this property makes it conceivable that inflation was short enough to allow subtle signatures of the preceding era to survive, such as detectable curvature [48].
- (iv) Equation (39) also captures another important result of Ref. [17], which we will examine in Sec. V: the probability of vacua where vacuum energy comes to dominate before τ_i^{obs} is exponentially suppressed, because the matter density, $\rho_{\text{matter},i}^{\text{obs}}$, will have been diluted by the accelerated expansion. This, too, appears to replicate a success of the causal-diamond measure: the suppression of moderately large values of the cosmological constant Λ , which are larger than the observed value but too small to affect the number of observers per matter mass. We will find, however, that this apparent success is an artifact of the no-collapse approximation.

Equation (39) trivially generalizes to the case where observations are made at different times τ_μ , $\mu = 1, 2, 3, \dots$. The total probability to observe vacuum i is

$$\mathcal{P}_i \propto e_i \sum_{\mu} \rho_i^{\text{obs}}(\tau_{\mu}), \quad (40)$$

where $\rho_i^{\text{obs}}(\tau_{\mu})$ is the density of observations taking place at τ_{μ} . In the continuous case,

$$\mathcal{P}_i \propto e_i \int_0^{\infty} d\tau \frac{d\rho_i^{\text{obs}}}{d\tau}, \quad (41)$$

where the integrand is the number of observations made in the FRW time interval $(\tau, \tau + d\tau)$, per unit physical volume of the hypersurface τ . If more differentiated observations are performed (for example, local conditions such as temperature, rather than just a determination of the vacuum), ρ_i can be given additional indices or arguments.

B. The no-collapse scale factor measure as a local measure

In the no-collapse approximation, the scale factor measure can be reformulated as a local measure. By this we mean that the measure can be defined by averaging over an ensemble of individual worldlines, instead of constructing a particular global spacetime. (The causal-diamond measure is an example of a local measure.)

This can be seen by the following argument, illustrated in Fig. 2. The predictions of the scale factor measure are dominated by the late-time attractor behavior of the Universe. (An infinite number of observations are produced during this era, while only a finite number are produced earlier.) Thus, we may as well pick a late-time surface, Σ_{η} with η very large, and choose it as our initial

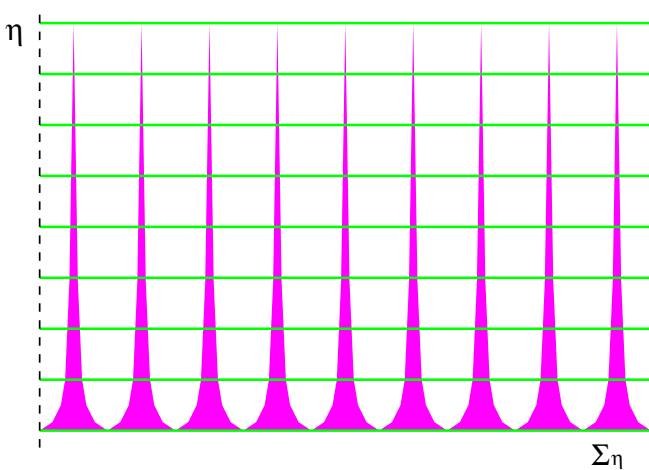


FIG. 2 (color online). The light shaded (green) slices are surfaces of constant scale factor time, Σ_{η} ; they have fixed comoving size but increasing physical volume. Fat geodesics (dark shaded, purple) have fixed physical width and thus decreasing comoving size. If the initial slice is chosen in the attractor regime, the fat geodesics define a representative finite sample of the total four-volume. Thus, the results of the scale factor measure can be reproduced by following a single geodesic of fixed width, starting in the longest-lived metastable vacuum.

surface, Σ_0 . In other words, we redefine $\eta \rightarrow \eta - \Delta\eta$, $\Delta\eta \approx \infty$, making sure we already start in the asymptotic regime. This cuts out irrelevant transients and leads to a very simple picture.

The volume occupied by long-lived metastable vacua on the late-time attractor hypersurface Σ_0 is dominated by empty de Sitter regions, with volume fraction s_i allocated to vacuum i [38]. We can choose Σ_0 to be as large as we like, guaranteeing that it can be allocated the correct attractor volume fractions to arbitrary accuracy.

Now let us begin evolving forward along the congruence orthogonal to Σ_0 . The scale factor measure instructs us to count observations in the four-volume thus generated, taking ratios as $\eta \rightarrow \infty$. Let us instead consider a reduced four-volume, \tilde{V}_4 , which is finite, by dividing the initial surface Σ_0 into volume elements dV . As usual, we follow the geodesic orthogonal to each volume element. But we do not increase dV as the Universe expands. This creates a family of “fat geodesics.” At the time η , the fat geodesics occupy only a fraction $e^{-3\eta}$ of the total volume of Σ_{η} . But for every worldline, no matter in what type of region it started, the missing volume is exactly the same, $1 - e^{-3\eta}$.

Because all fat geodesics expand by the same volume factor, the volume they do occupy on Σ_{η} is a perfectly faithful sample of the hypersurface. (This is the crucial point—note, for example, that it would not hold in the attractor regime of the proper time measure.) Moreover, because this is true for every time interval $d\eta$, the four-volume swept out by the fat geodesics is statistically equivalent to the four-volume between Σ_0 and Σ_{η} .

Therefore, reducing to \tilde{V}_4 will not affect probabilities:

$$\frac{\mathcal{P}(O_i)}{\mathcal{P}(O_j)} \equiv \lim_{\eta \rightarrow \infty} \frac{N_i}{N_j} = \frac{N_i(\tilde{V}_4)}{N_j(\tilde{V}_4)}, \quad (42)$$

where V_4 is obtained by following every geodesic to the asymptotic future.⁸ Instead of thinking about the actual collection of fat geodesics anchored on a large hypersurface, we may equivalently consider a statistical ensemble of single geodesics, with initial conditions weighted by the distribution of regions on a late-time attractor surface. Neglecting the rare regions occupied by recently nucleated bubbles, this means starting out with de Sitter vacuum i with probability s_i . Because the vector s_i is dominated by the * vacuum, to a good approximation [42], this means starting the geodesic in the longest-lived metastable vacuum, *.

Thus, we reproduce the scale factor measure by following a single geodesic starting in *. The worldline evolves according to local dynamical laws, witnessing the decay of vacua and perhaps the production of observers, until it ends up in a terminal vacuum. We “fatten” the worldline, giving

⁸The no-collapse approximation does not allow us to consider negative cosmological constant regions, but for the purposes of this argument we can set $\Lambda_i \rightarrow 0$ for all vacua with $\Lambda_i < 0$.

it a *fixed physical cross section*, a volume element dV orthogonal to the worldline. Finally, we compute the expectation value (ensemble average) of the differential number of observations of type μ , dN_μ , in the resulting four-volume:

$$\mathcal{P}_\mu = \langle \rho_\mu \rangle, \quad (43)$$

where, for a given worldline in the ensemble,

$$\rho_\mu = \frac{dN_\mu}{dV} = \int dt \frac{dN_\mu}{dV dt}. \quad (44)$$

We showed in Sec. III that geodesics quickly become comoving after entering a new bubble. Thus, the volume element dV lies inside a constant- τ hypersurface of each FRW bubble universe. With the approximations used in Sec. IVA, therefore, ρ_μ is identical to the physical density of observers, ρ_i^{obs} , in vacuum i . The factor e_i is captured by averaging over different worldlines in the ensemble [44], so Eq. (43) reduces to Eq. (39) as a special case. The more general Eqs. (40) and (41), too, are special cases of Eq. (43).

It would be interesting to use Eq. (43) as the *defining* equation of a measure. This would have a number of formal advantages. One geodesic carries much less geometric information than a whole congruence, so we face fewer ambiguities about the treatment of upward jumps. Moreover, Eq. (43) can be applied without modification in collapsing regions, whereas the formulation based on integrating the expansion along a geodesic congruence requires an additional, *ad hoc*, prescription to deal with such regions. Finally, the use of Eq. (43) to define a measure liberates us from the initial conditions s_i picked out by the attractor regime of the geodesic congruence. Starting the fat geodesic with initial conditions that favor Planck-scale vacua, for example, would avoid the potential “staggering problem” associated with the enormous suppression of the upward jumps from the dominant vacuum*.

C. Proper treatment of collapsed regions

At least in our own bubble, some observations are made in collapsed regions. The FRW metric, Eq. (26), does not capture the local geometry of such regions, but provides only an average over scales on which the Universe can be considered homogeneous. Hence, this metric cannot be used to compute the local scale factor, A , and the scale factor time, η , along geodesics entering collapsed regions such as galaxies. Despite the similar names, the FRW scale factor a and the local scale factor A are two completely different objects.

Geodesics that end up in dark matter halos will, by definition, have ceased to expand, and decoupled from the Hubble flow, before the halo formed. After reaching a maximum expansion, they will have turned around and collapsed, with A and η decreasing during the collapse. These geodesics will reach observers, but their maximum

scale factor A_{max} will be unrelated to the FRW scale factor at the time t_{obs} , a_{obs} . Rather, it will be related to the FRW scale factor at the time of structure formation, a_{NL} .

Consider a simple, spherically symmetric model. A dark matter halo forms by the collapse of a spherical overdensity. Then baryons fall into their gravitational well, cool, and condense in the center. Eventually the gas fragments into stars.

Let us focus on the dark matter particles that end up in the halo. Initially, each particle follows one of the geodesics in the congruence that defines the scale factor cutoff. The maximum scale factor is achieved at the time of turnaround, when all the dark matter particles are momentarily stationary. Note that it is smaller than, but on the order of, the FRW scale factor a at the time of turnaround.

After the turnaround, the particles fall toward the center of the halo. Depending on interactions, the particles will eventually stop following the geodesics, but in the spherical model, the geodesics remain very simple. They will focus in the center of the halo, and begin expanding again, back out to the turnaround radius; this pattern will be repeated indefinitely. The entire congruence of geodesics will keep oscillating about the center of the halo, with an amplitude given by the turnaround radius. This radius is twice the virial radius, and perhaps 10 times the eventual galaxy radius. Therefore, the congruence continues to thread the galaxy at all times. Crucially, it will capture observers independently of how long it takes to form them.⁹

With a more realistic structure formation model, the situation would appear to be even more clear cut. Generic geodesics will remain gravitationally bound to the halo, but will be spread chaotically through microlensing. (This means that the same event will lie on multiple geodesics, but as discussed above, we will count each event only once.) Moreover, most large halos do not form directly from a single overdensity, but by mergers and accretion of smaller halos that virialized much earlier. One expects that most geodesics threading merging halos will remain gravitationally bound during the merger, and end up covering the resulting larger galaxy. But then observers will have a local scale factor which is less than the averaged scale factor at the time of the formation of the smallest structures that eventually merge to form our galaxies. Note that these structures need not even be galaxies, i.e., their mass could be below 10^7 solar masses.

The measure defined by De Simone *et al.* instructs us to include all observations reached by geodesics whose maximum scale factor time is below the cutoff. Therefore, all

⁹In the spherical idealization with an empty shell between overdense geodesics and the homogeneous background, there would be a small fraction of events that are missed by the congruence near the time when it is in focus. This would not be an important effect quantitatively, and will be absent in a realistic congruence.

observations in collapsed regions will contribute as soon as the cutoff exceeds the maximum local scale factor of the first collapsing objects that end up constituting the host objects of observations.

To include this effect, one could generalize to a more detailed metric. But it is much simpler to continue working with the homogeneous FRW metric, Eq. (26), and to include collapse effects by pretending that all observations in collapsed regions happen at the time τ^{NL} when those regions decoupled from the Hubble flow. This amounts to projecting observations from the time τ^{obs} back to the earlier time τ^{NL} , and thus, to increasing their number density per physical volume to

$$\hat{\rho}_i^{\text{obs}} \equiv \rho_i^{\text{obs}} \left(\frac{a_i^{\text{obs}}}{a_i^{\text{NL}}} \right)^3, \quad (45)$$

where $a_i^{\text{NL}} \equiv a(\tau_i^{\text{NL}})$.

Thus, we can include the effects of local collapse by replacing Eq. (39) by

$$\mathcal{P}_i \propto e_i \hat{\rho}_i^{\text{obs}}. \quad (46)$$

V. THE COSMOLOGICAL CONSTANT

In this section, we estimate the probability distribution for the cosmological constant in the scale factor measure. We will find that it agrees roughly with the observed value in the no-collapse approximation. Treating collapsed regions according to the prescription of De Simone *et al.*, however, yields a much larger value. We will then consider variations of the scale factor measure and discuss how they might solve this problem. Let us begin by reviewing the history of anthropic predictions for the cosmological constant and how they depend on the measure.

Throughout this paper, we will focus on positive values of Λ , which have a greater potential for discrepancy with observation. Negative values of Λ are bounded (in any nonpathological measure, for example, if there is no youngness paradox) by the order of magnitude of the observed value, but positive values could be much larger [11]. Both the causal-diamond measure [49] and the no-collapse scale factor measure [17] assign somewhat higher integrated probability to negative values than to positive values of Λ , but the imbalance is not large enough to render the observed value unlikely.

A. Λ and observers per baryon

In anthropic approaches to the cosmological constant problem, one computes a probability distribution for *observed* values of the cosmological constant. Originally [33–39] this was implemented by multiplying the underlying distribution (arguably, $dp/d\Lambda = \text{const}$ in the regime of interest) by the “number of observers per baryon,” or per some other reference object. Implicitly, this defines an (incomplete) measure. If one tries to take observers per

baryon seriously as a measure, one finds that any eternally inflating de Sitter vacuum has an infinite number of observers per baryon (if it has any observers at all), and 100% of the observers are Boltzmann brains [40,41].

Even aside from this problem, however, the observer-per-baryon measure was already plagued by some annoying problems. The simplest test of the landscape solution to the cosmological constant problem is to restrict attention to vacua that differ from ours *only* through their value of Λ . In this setting, the observer-per-baryon measure prefers values of Λ about 5000 times larger than the observed value; the observed value is excluded at the 3.5σ level [49]. We find it useful to explain this in terms of the time scale

$$\tau_\Lambda = \sqrt{\frac{3}{\Lambda}} \quad (47)$$

at which the cosmological constant comes to dominate the evolution of the Universe. In the observer-per-baryon measure, the number of observers will be proportional to the number of baryons that are captured by galaxies. The underlying distribution prefers larger values of Λ . There is no penalty until Λ becomes large enough to disrupt galaxy formation ($\tau_\Lambda \ll \tau_{\text{gal}}$); such values will be strongly suppressed. Thus one would expect Λ to be correlated with the time of the formation of the first galaxies, τ_{gal} :

$$\Lambda \sim \tau_{\text{gal}}^{-2}. \quad (48)$$

In fact, however, the observed value of Λ appears to be correlated with the (considerably later) time when observations are made, τ_{obs} :

$$\Lambda \sim \tau_{\text{obs}}^{-2} \sim 10^{-4} \tau_{\text{gal}}^{-2}. \quad (49)$$

This is the coincidence problem. The observers-per-baryon measure (ignoring Boltzmann brains) would have naturally explained a hypothetical coincidence between τ_{gal} and τ_Λ , but it is only barely consistent with the actual coincidence between τ_{obs} and τ_Λ .

Further tests can be made by allowing other parameters (such as the density contrast, curvature, etc.) to vary in addition to Λ , or by integrating out such parameters altogether. This exacerbates the troubles of the observer-per-baryon measure. For example, it strongly prefers larger values of the density contrast, Q , than the observed value $Q \sim 2 \times 10^{-5}$. Since $\tau_{\text{gal}} \propto Q^{-3/2}$, vacua with larger Q can tolerate larger Λ (like Q^3) while still forming a structure. Thus, a greater fraction of such vacua contains observers, and they are strongly preferred.

B. Λ and the causal-diamond measure

The causal-diamond measure [13,14] was the first to solve these problems. Most directly, it solves the coincidence problem, predicting a value of Λ such that $\tau_\Lambda \sim \tau_{\text{obs}}$, or $\Lambda \sim \tau_{\text{obs}}^{-2}$. Greater values of Λ are suppressed mainly because observers become exponentially dilute

within the cutoff [49]:

$$\frac{d\mathcal{P}}{d\log\Lambda} \propto \Lambda^{1/2} \exp\left(-\frac{3\tau_{\text{obs}}}{\tau_{\Lambda}}\right). \quad (50)$$

Using $\tau_{\text{obs}} \sim O(10 \text{ Gyr})$ (or alternatively, and less anthropocentrically, equating the peak observation time with the time of peak entropy production), this relation leads to a prediction of Λ in excellent agreement with the observed value. Because the causal-diamond measure correlates Λ directly with the era when observations are made, and not with the time when structure forms, the Q runaway problem is absent [49,50].

C. Λ and the scale factor measure

De Simone *et al.* [17] have argued that the scale factor measure shares some of these desirable properties. We will now examine this claim, focusing, once more, on the case where only Λ varies, while keeping τ_{obs} fixed. We will find that the conclusions of De Simone *et al.* apply in the no-collapse approximation, but are invalidated by collapse effects.

1. No-collapse approximation

In the no-collapse approximation adopted in Ref. [17], Eq. (39) tells us that

$$\frac{d\mathcal{P}}{d\Lambda} \propto \frac{dp}{d\Lambda} \rho^{\text{obs}}(\Lambda). \quad (51)$$

We assume that $dp/d\Lambda$ is constant in a small interval (say, 10^{-20}) around $\Lambda = 0$, after coarse graining over an even smaller interval (say, 10^{-130}). If this is not the case, then the landscape cannot solve the cosmological constant problem, and the scale factor measure would be ruled out. Hence we may drop this factor and write

$$\frac{d\mathcal{P}}{d\log\Lambda} \propto \Lambda \rho^{\text{obs}}(\Lambda). \quad (52)$$

To understand how the observer density depends on Λ , let us rewrite ρ^{obs} as a fraction of the physical matter density at the time of observations,

$$\rho^{\text{obs}}(\Lambda) = \alpha(\Lambda) \rho_{\text{matter}}^{\text{obs}}(\Lambda), \quad (53)$$

so that

$$\frac{d\mathcal{P}}{d\log\Lambda} \propto \Lambda \alpha(\Lambda) \rho_{\text{matter}}^{\text{obs}}(\Lambda). \quad (54)$$

It is reasonable to assume that α , the number of observers per unit matter mass, is proportional to the Press-Schechter mass function [36], the fraction of matter mass in collapsed objects above a certain critical mass (10^7 solar masses for the smallest galaxies; more if one makes the additional anthropic assumption that observers require particularly large galaxies). To first approximation, α is unaffected by Λ as long as $\tau_{\Lambda} \gg \tau_{\text{gal}}$. Larger values of Λ will

noticeably affect galaxy formation, and for $\tau_{\Lambda} \ll \tau_{\text{gal}}$, α rapidly vanishes.

The matter density at the time of observations is independent of Λ as long as $\tau_{\Lambda} \gg \tau_{\text{obs}}$. But if Λ comes to dominate earlier, then it will drive a period of exponential expansion before observations are made, and the matter density will be far lower. Roughly, we can write

$$\rho_{\text{matter}}^{\text{obs}}(\Lambda) \approx \rho_{\text{matter}}^{\text{obs}}(\Lambda = 0) \exp\left(-3\frac{\tau_{\text{obs}}}{\tau_{\Lambda}}\right). \quad (55)$$

For example, in vacua with $\tau_{\Lambda} \approx 2 \text{ Gyr}$, the density of galaxies at $\tau_{\text{obs}} \approx 13.7 \text{ Gyr}$ will be a factor of order 10^{-9} times smaller than in vacua with $\tau_{\Lambda} \gg 13.7 \text{ Gyr}$. Thus, moderately large values of Λ are totally suppressed, even though they do not disrupt galaxy formation. We conclude that in the no-collapse approximation, the suppression of large Λ is dominated by the dilution of matter, and one obtains

$$\frac{d\mathcal{P}}{d\log\Lambda} \propto \Lambda \exp\left(-3\frac{\tau_{\text{obs}}}{\tau_{\Lambda}}\right). \quad (56)$$

Comparison with Eq. (50) reveals that the exponential factor is the same as that arising in the causal-diamond measure. In both measures, this factor arises primarily from the effect of Λ on the matter density, rather than its effect on galaxy formation. The prefactors differ because the physical volume of the causal diamond also depends on Λ , as with $\Lambda^{-1/2}$ for large Λ . (In both cases the prefactor contains a factor Λ from the Jacobian $d\log\Lambda/d\Lambda$.) Both distributions peak near $\Lambda \sim \tau_{\text{obs}}^{-2}$, in excellent agreement with observation, Eq. (49).

2. Proper treatment of collapsed regions

With the proper treatment of collapsed regions, however, the result changes drastically. Equation (39) is replaced by Eq. (46), and correspondingly, we must replace Eq. (52) by

$$\frac{d\mathcal{P}}{d\log\Lambda} \propto \Lambda \hat{\rho}^{\text{obs}}(\Lambda), \quad (57)$$

where $\hat{\rho}^{\text{obs}}$ is the density, per physical volume, that observations would have had if they had occurred at the time τ_{NL} when the first dark matter halos decoupled from the Hubble flow. By Eqs. (45) and (53), it follows that

$$\frac{d\mathcal{P}}{d\log\Lambda} \propto \Lambda \alpha(\Lambda) \rho_{\text{matter}}^{\text{NL}}(\Lambda), \quad (58)$$

where $\rho_{\text{matter}}^{\text{NL}}$ is the physical matter density at the time τ_{NL} .

This matter density is insensitive to Λ unless Λ is large enough to affect the formation of the earliest dark matter halos, at the time τ_{NL} . Since $\tau_{\text{NL}} < \tau_{\text{gal}}$, the dilution of matter density is now no longer the effect suppressing large Λ . Rather, it is a disruption of galaxy formation: the Press-Schechter factor, which enters through α . It becomes sup-

pressed if Λ exceeds τ_{gal}^{-2} , so the distribution will peak near this value.

Thus, the scale factor measure reproduces the result obtained from the observers-per-baryon measure, Eq. (48), except for the latter measure's manifest Boltzmann brain problem. It prefers a value of Λ that is, depending on the strength of anthropic assumptions, 1 to 4 orders of magnitude larger than the observed value.

Moreover, the scale factor measure suffers from a runaway problem if the strength of initial density perturbations, Q , is allowed to vary. The pressure toward large Q is actually greater than in the observers-per-baryon measure. In both measures, larger Q allows for larger values of Λ , since the maximum vacuum energy is of order of the density at the time of galaxy formation, which is proportional to Q^3 . This enters through the first factor in Eq. (58), suppressing the probability of the observed value of Λ by an additional factor $(Q_0/Q)^3$. But in the scale factor measure, the last factor in Eq. (58) offers an additional reward for large Q : the density at the time of the earliest halo formation also scales similar to the third power of Q . This makes it more difficult to compensate for the runaway problem by prior distributions or anthropic cutoffs.

D. Can the scale factor measure be improved?

We have found that the scale factor measure predicts $\Lambda \sim \tau_{\text{gal}}^{-2}$, while the causal-diamond measure predicts $\Lambda \sim \tau_{\text{obs}}^{-2}$. Can we pinpoint the key difference between the measures leading to these different results? And can the definition of the scale factor measure be modified to yield a more desirable answer?

The causal-diamond measure is defined nonlocally, in terms of the event horizon of a single worldline. Observations inside this horizon will be included, while those outside will not. The horizon size is not affected by local phenomena such as gravitational collapse. If $\tau_{\Lambda} \ll \tau_{\text{obs}}$, then observations will be dilute, and few if any will be captured by the diamond.

The scale factor measure is defined in terms of a local quantity, the expansion and local scale factor of a congruence of geodesics. The local scale factor runs backward once structure formation begins, and a special rule is needed to decide how the cutoff should be specified in this case. By the rule of De Simone *et al.*, future observations occurring on such geodesics are immediately counted at structure formation, so the dilution of galaxies by the cosmological constant cannot affect the probability distribution.

One might attempt to improve the scale factor measure by replacing the local scale factor, A , with the averaged scale factor, a , describing the spatial curvature radius in expanding open FRW bubbles, while a geodesic is passing through matter dominated regions (up to an obvious constant factor that depends on the scale factor time at which the geodesic enters the matter dominated region). The idea

would be to avoid the effects of collapse by averaging over collapsed regions and defining the scale factor time in terms of the averaged Hubble flow.

A problem with this approach is that the FRW scale factor a is only approximately defined. On the scales currently within the horizon, there exists a preferred slicing of constant average density, which defines hypersurfaces of constant a . However, because the Universe is not exactly homogeneous and isotropic, the averaging procedure is necessarily somewhat ambiguous—in the observable Universe, at least at the level of 10^{-5} . At larger scales, or in different bubbles, density fluctuations can be much larger. Bubbles with eternal slow-roll inflation cannot be assigned a preferred FRW slicing at all; neither can the asymptotic regions containing Boltzmann brains. Thus, it is unclear how this idea would lead to a sufficiently general prescription for computing probabilities.

A more promising approach to improving the scale factor measure is to change the rule that deals with collapsing geodesics, so that observers are not counted by geodesics that turn around too early. For example, whenever an event is passed through by multiple geodesics, we might assign its scale factor time to be the *maximum* scale factor time among those geodesics. This is in principle as well defined as the proposal in [17], which corresponds to using the minimum. A closely related idea is to extend each geodesic only until its first caustic, when neighboring geodesics intersect and $\theta \rightarrow -\infty$. Whether such prescriptions lead to a better prediction for Λ is an interesting question, but one beyond the scope of the present paper.

Finally, we may turn to the local formulation of the scale factor measure, Eq. (43), for help. In expanding regions, it is equivalent to the formulation using a geodesic congruence, but as noted at the end of Sec. IV B, the local formulation can be applied without modification in collapsed regions. It would, however, suffer from a similar problem with predicting Λ . Most fat geodesics are captured by dark matter halos around the time when those objects first form. Therefore, the expected density of observations is set by the matter density inside galaxies, and not by the large-scale average of the matter density of the FRW solution at the time of observations, τ_{obs} . As a result, Eq. (43) is insensitive to the value of the cosmological constant unless it is large enough to disrupt galaxy formation. It would predict $\Lambda \sim \tau_{\text{gal}}^{-2}$, in poor agreement with the observed value $\Lambda \sim \tau_{\text{obs}}^{-2}$.

Still, Eq. (43) may turn out a useful starting point for a more successful measure. Instead of fattening the geodesic infinitesimally, for example, we could consider constructing a larger transverse volume. We can then either attempt to take a large volume limit, in which the average density may become sensitive to the dilution caused by early vacuum domination. (Whether this can be done in a well-defined manner is, again, a question beyond the scope of our present work.) Or we could use a finite transverse

volume. To avoid explicitly introducing an arbitrary length scale into the measure, this volume could be defined in terms of the cosmological event horizon surrounding the worldline. This fixes the Λ problem, but it fails to produce a novel measure: With this prescription, we would simply recover the causal-diamond measure (with a particular choice of initial conditions).

VI. BOLTZMANN BRAINS

In this section, we compute the number of Boltzmann brains, and determine the conditions under which they dominate over ordinary observers. We define Boltzmann brains as observers that arise from local violations of the second law of thermodynamics. This occurs at late times in de Sitter vacua. States of energy $E \gg T$ are produced at the Boltzmann-suppressed rate

$$\Gamma_i(E) = \frac{4\pi}{3} H_i^4 \exp(-E/T_i) \quad (59)$$

where $T_i = H_i/2\pi$ is the Gibbons-Hawking temperature of the de Sitter horizon. The minimum mass of a Boltzmann brain (if any) will depend on the vacuum. But on general grounds [41,57], the exponential factor cannot be larger (though it can be much smaller) than $\exp(-S_{\text{BB}})$, where S_{BB} is the coarse-grained entropy, or number of particles, in the most primitive Boltzmann brain,

$$\Gamma_i^{\text{BB}} < \exp(-S_{\text{BB}}). \quad (60)$$

One can only speculate about the value of S_{BB} , but it is certain to be exponentially large. Thus, Boltzmann brains are double exponentially suppressed.

A. The ratio of Boltzmann brains to ordinary observers

We can neglect crunching vacua, as well as the initial non-de Sitter regime of metastable vacua: because of the enormous suppression of Boltzmann brains, almost all of them will be produced in the asymptotic de Sitter regime. The number of Boltzmann brains produced in vacuum i between the time η and $\eta + d\eta$ is

$$dN_i^{\text{BB}} = \Gamma_i^{\text{BB}} V_i d\eta / H_i. \quad (61)$$

By Eq. (14), the total number of Boltzmann brains produced prior to the cutoff η in vacuum i is therefore

$$N_i^{\text{BB}} = \frac{C e^{\gamma\eta}}{\gamma} \frac{s_i}{H_i} \Gamma_i^{\text{BB}}. \quad (62)$$

By Eqs. (16) and (17), $s_* \approx 1$ for the dominant vacuum and $s_i = p_i/(\kappa_i - q) \approx p_i/\kappa_i$ for all other vacua. Dropping the usual factor $C e^{\gamma\eta}/\gamma$, the total probability for observations by Boltzmann brains is

$$\mathcal{P}^{\text{BB}} \propto H_*^{-1} \Gamma_*^{\text{BB}} + \sum_{i \neq *} H_i^{-1} p_i \frac{\Gamma_i^{\text{BB}}}{\kappa_i}. \quad (63)$$

We are interested in comparing this to the total number of

observations by ordinary observers,

$$\mathcal{P}_i^{\text{OO}} \propto \sum_i p_i \rho_i^{\text{OO}}. \quad (64)$$

The key simplification arising in this comparison is that p_i , Γ_i^{BB} , and Γ_i , are generically double exponentials. That is, they are of the form $\exp(\pm \exp x)$ with $x \gg 1$. We will use a triple inequality sign for such numbers, for example,

$$\Gamma^{\text{BB}} \lll 1. \quad (65)$$

Such numbers obey special laws of arithmetic. For example, for y and z double exponentially large, $y/z \approx y$ if $y > z$. Moreover, if y is a single exponential and z a double exponential, then $zy \approx z/y \approx z$. A double exponential takes the same value in any conventional system of units, though it can be useful to think in terms of Planck units for definiteness.

The landscape contains an exponentially large number of vacua, but so far there are no indications that it might be doubly exponentially large. Thus, H_i is at most a single exponential, and we can write

$$\frac{\mathcal{P}^{\text{BB}}}{\mathcal{P}^{\text{OO}}} \approx \frac{\Gamma_*^{\text{BB}} + \sum_{i \neq *} p_i (\Gamma_i^{\text{BB}} / \Gamma_i)}{\sum_i p_i \rho_i^{\text{OO}}}. \quad (66)$$

Note that we have retained ρ_i^{OO} , since it can be zero or doubly exponentially small in some vacua. Because Eq. (66) involves a ratio of double exponentials, either the numerator or the denominator will completely dominate (depending on the measure, and on the landscape), and the relative probability will be zero or infinity to good approximation. Whoever wins, wins big.

We can restate this result in terms of the expected number e_i of times a worldline starting in the $*$ vacuum will pass through vacuum i (see the discussion at the end of Sec. II). Using $q \approx \kappa_*$, we can include the $*$ vacuum in the sum:

$$\frac{\mathcal{P}^{\text{BB}}}{\mathcal{P}^{\text{OO}}} \approx \frac{\sum_i e_i (\Gamma_i^{\text{BB}} / \Gamma_i)}{\sum_i e_i \rho_i^{\text{OO}}}. \quad (67)$$

At this level of approximation, the causal-diamond measure yields nearly the same result [41]. In the causal-diamond measure, the number of Boltzmann brains in vacuum i is the expected number of times the generating worldline enters vacuum i , e_i , times Γ_i^{BB} , times the expected four-volume the diamond will span in vacuum i (given by the lifetime of vacuum i , $H_i^{-1} \kappa_i^{-1}$, times the de Sitter horizon volume, $4\pi/3 H_i^3$). The expected number of ordinary observers is e_i times the number of ordinary observers inside the de Sitter horizon of vacuum i . Keeping only double exponentials, the relative probability in the causal-diamond measure is again given by Eq. (67).

Aside from negligible non-double-exponential factors, the only way the two measures differ, for the purposes of Boltzmann brains, is through e_i , which depend on initial conditions. *A priori*, the question of initial conditions has nothing to do with the measure problem, and in the causal-diamond measure, they remain separate issues. One imagines that the Universe started, perhaps, in a randomly chosen vacuum, or in an ensemble governed by the tunneling wave function, favoring Planck-scale vacua. Physical probabilities are not strongly affected by this uncertainty [44]. In the scale factor measure, on the other hand, the eternally inflating Universe exhibits attractor behavior, which can be mimicked by choosing the particular initial probability distribution s_i , defined in Sec. II. Roughly, this means starting the worldline in a very particular vacuum: the longest-lived metastable vacuum [42], $*$.

B. Under what conditions do Boltzmann brains dominate?

In this section, we will show that with some reasonable assumptions about the landscape, the ratio of Boltzmann brains to ordinary observers takes an even simpler form. Except for the $*$ term, each term in the numerator of Eq. (66) contains the ratio $\Gamma_i^{\text{BB}}/\Gamma_i$. By the laws of double exponential arithmetic, this ratio is dominated by the larger of the two exponents,

$$\frac{\Gamma_i^{\text{BB}}}{\Gamma_i} \approx \begin{cases} \Gamma_i^{\text{BB}} < \exp(S_{\text{BB}}) & \text{if } \Gamma_i^{\text{BB}} < \Gamma_i, \\ \Gamma_i^{-1} > \exp(+S_{\text{BB}}) & \text{if } \Gamma_i^{\text{BB}} > \Gamma_i, \end{cases} \quad (68)$$

where we have used Eq. (60) for the inequalities.

We will now constrain the factor p_i multiplying this ratio. With no loss of generality, we may restrict the sum in Eq. (66) to the set of vacua containing observers or Boltzmann brains. Let us label the vacua in this set by the index a instead of i . By Eqs. (19) and (23), p_a can be obtained by a sum, over all decay paths, of products of branching ratios.

$$p_a = q \sum_{\text{paths from } * \text{ to } a} \beta_{a i_n} \beta_{i_n i_{n-1}} \cdots \beta_{i_1 *} \quad (a \neq *). \quad (69)$$

This equation holds only for $a \neq *$, but those are precisely the p_a appearing in Eq. (66).¹⁰

We will make three assumptions:

- (1) For every vacuum a , every decay path from $*$ to a contains at least one intermediate de Sitter vacuum. This assumption is very plausible in a multidimensional landscape such as that of string theory, where neighboring vacua have vastly different values of the cosmological constant (a crucial feature for

solving the cosmological constant problem [3,12]). Assuming that distant vacua cannot be accessed with appreciable probability, it follows that $\Delta\Lambda \gg 1/S_{\text{BB}}$ in direct decays, so it is exponentially unlikely that the dominant vacuum can decay to any vacuum large enough to contain observers or Boltzmann brains.

- (2) For every vacuum a , there exists at least one decay path from $*$ to a which does not pass through any de Sitter vacua with horizon entropy bigger than S_{BB} . This assumption, too, is plausible given the scarcity of vacua with $\Lambda \ll 1$, and given our ability to choose any path we like.
- (3) Of the decay paths posited in assumption (2), at least one “dominates” the sum in Eq. (69), in the following extremely weak sense: By dropping all other terms, we change the sum by less than a double exponential factor. Similar to assumption (1), this is plausible in a multidimensional landscape with large step size, where decay chains in the semiclassical regime are short.

We remain agnostic about whether the $*$ vacuum has horizon entropy larger or smaller than S_{BB} . The dominant vacuum is special. It is at least conceivable that its defining properties strongly select for a very small cosmological constant, and we do not wish to prejudice this issue here.

We will be interested in bounding p_a only to within double-exponentially large factors. By assumption (3), we can drop the sum over paths and write

$$p_a \sim q \beta_{a i_n} \beta_{i_n i_{n-1}} \cdots \beta_{i_1 *} \quad (a \neq *). \quad (70)$$

Assumption (1) guarantees that $n \geq 1$, so $\beta_{i_1 *}$ exists in the product and represents the branching ratio from one de Sitter vacuum to another. For two connected de Sitter vacua, detailed balance gives $\kappa_{i_1 *} = \kappa_{* i_1} \exp(S_{i_1} - S_*)$ where S_j is the horizon entropy of vacuum j . Also, the decay rate must be faster than the recurrence time, $\kappa_{i_1 *} > \exp(-S_*)$. We are only considering vacua which eternally inflate, $\kappa_{* i_1} < 1$, so

$$\exp(-S_*) < \kappa_{i_1 *} < \exp(S_{i_1} - S_*). \quad (71)$$

We can divide by κ_* to get a bound on the branching ratio,

$$\frac{1}{\kappa_*} \exp(-S_*) < \beta_{i_1 *} < \frac{1}{\kappa_*} \exp(-S_* + S_{i_1}). \quad (72)$$

For the other branching ratio(s) appearing in Eq. (70), we can use the cruder bound

$$\exp(-S_j) < \frac{1}{\kappa_j} \exp(-S_j) < \beta_{ij} < 1, \quad (73)$$

which holds also if the destination vacuum is terminal (as a might be). Substituting these bounds into Eq. (19), we obtain

¹⁰While our discussion will focus on the scale factor measure, it can be adapted to the causal-diamond measure simply by replacing $*$ with another initial vacuum.

$$\begin{aligned} & \exp(-S_*) \exp(-S_{i_1} - S_{i_2} - \cdots - S_{i_n}) \\ & < p_a < \exp(-S_* + S_{i_1}), \end{aligned} \quad (74)$$

where we have used $\kappa_* \approx q$. By assumption (2) (and assuming that the path is not exponentially long), we have $\sum_{k=1}^n S_{i_k} \ll S_{\text{BB}}$, so

$$\exp(-S_*) \exp(-S_{\text{BB}}) \lll p_a \lll \exp(-S_*) \exp(+S_{\text{BB}}). \quad (75)$$

(We remind the reader that we use the triple inequality sign for the inequality of double exponentials.)

We have bounded p_a to within a range small compared to $\exp(S_{\text{BB}})$. Meanwhile, by Eq. (68), the ratio $\Gamma_a^{\text{BB}}/\Gamma_a$ is either larger than $\exp(S_{\text{BB}})$ or smaller than $\exp(-S_{\text{BB}})$. By double exponential arithmetic, therefore, we can neglect the uncertainty in p_a when we multiply these terms:

$$p_a \frac{\Gamma_a^{\text{BB}}}{\Gamma_a} \approx \exp(-S_*) \frac{\Gamma_a^{\text{BB}}}{\Gamma_a}. \quad (76)$$

By similar double exponential reasoning we can do the same thing in the denominator, so the ratio becomes

$$\frac{\mathcal{P}^{\text{BB}}}{\mathcal{P}^{\text{OO}}} = \frac{\Gamma_*^{\text{BB}} + \sum_a \exp(-S_*) (\Gamma_a^{\text{BB}}/\Gamma_a)}{\sum_a \exp(-S_*) \rho_a^{\text{OO}}}. \quad (77)$$

It seems likely that in our vacuum the density of observers is not double exponentially small, and in Planck units $\rho_a^{\text{OO}} < 1$, so $\sum_a \rho_a^{\text{OO}}$ is not double exponential as long as the number of vacua is not double exponential. Therefore,

$$\frac{\mathcal{P}^{\text{BB}}}{\mathcal{P}^{\text{OO}}} = \exp(S_*) \Gamma_*^{\text{BB}} + \sum_a \frac{\Gamma_a^{\text{BB}}}{\Gamma_a}. \quad (78)$$

For the ordinary observers to dominate, the ratio must be less than 1, so every term must be less than 1. This requires

$$\frac{\Gamma_a^{\text{BB}}}{\Gamma_a} < 1 \quad \forall a. \quad (79)$$

Also, if the dominant vacuum can produce Boltzmann brains at all, then it will do so far faster than the recurrence rate $\exp(-S_*)$, so for the first term to be less than 1 we need

$$\Gamma_*^{\text{BB}} = 0. \quad (80)$$

If either of these conditions is violated, the Boltzmann brains dominate.¹¹

¹¹Our result is consistent with the conditions found by Linde [24] for a toy landscape with two de Sitter vacua and a sink. This toy landscape is so small that it violates our assumptions, but nevertheless Eqs. (79) and (80) are necessary and sufficient for ordinary observers to dominate. The condition $\Gamma_{1s} \gg \Gamma_{21}$ (in the notation of Ref. [24]) ensures that $2 = *$; the condition $\Gamma_{1s} \gg \Gamma_{1B}$ corresponds to Eq. (79). Equation (80) was not explicitly spelled out but is implicit in Fig. 4 of Ref. [24].

C. Discussion

Let us summarize our result in general terms. If any vacua have lifetimes longer than their Boltzmann brain time ($\Gamma_a < \Gamma_a^{\text{BB}}$), then by the laws of double exponentials, the Boltzmann brains will dominate—unless there is a conspiracy in the landscape so that the rate of production p_a of every single Boltzmann brain producing vacuum is double exponentially small compared to the p_a for some vacuum which contains mostly ordinary observers. Conversely, if all vacua decay before they produce Boltzmann brains, and the $*$ vacuum does not produce Boltzmann brains, then ordinary observers dominate unless all of the p_a for ordinary observer vacua are double exponentially small compared to the p_a for some vacuum which produces primarily Boltzmann brains.

It is interesting to compare the conditions necessary for the absence of Boltzmann brains with those arising in the causal-diamond measure. For the purposes of Boltzmann brains, the causal-diamond measure and the scale factor measure differ only through the choice of initial conditions, i.e., through the second of the two conditions in Eqs. (79) and (80). In the causal-diamond measure, there is no reason to select initial conditions that favor vacua with extremely small cosmological constant $\Lambda < 1/S_{\text{BB}}$. Thus, it is implausible that Boltzmann brains would dominate via Eq. (80) (with the $*$ vacuum replaced by the relevant initial conditions).

In the scale factor measure, on the other hand, the $*$ vacuum may have very small cosmological constant, for example, because of a correlation between vacuum stability and the degree of supersymmetry breaking. Thus, while the string landscape may well satisfy Eq. (79), it appears that the scale factor measure gives Boltzmann brains a second chance through Eq. (80).

We should not make too much of this difference. Cosmological constant aside, Boltzmann brains are fairly complex objects and will not arise in every imaginable low-energy field theory. Thus, we expect Γ_a^{BB} to vanish exactly in most vacua. In the $*$ vacuum, the necessary conditions on matter content are no more likely to be satisfied than in any other randomly chosen vacuum, since low-energy properties such as particles and fields are unlikely to be correlated with the high-energy features responsible for the vacuum's longevity. Thus it seems very likely that $\Gamma_*^{\text{BB}} = 0$.

ACKNOWLEDGMENTS

We thank A. Guth and A. Vilenkin for discussions. This work was supported by the Berkeley Center for Theoretical Physics, by a CAREER grant (No. 0349351) of the National Science Foundation, and by the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

- [1] S. Weinberg, *Rev. Mod. Phys.* **61**, 1 (1989).
- [2] J. Polchinski, arXiv:hep-th/0603249.
- [3] R. Bousso, *Gen. Relativ. Gravit.* **40**, 607 (2008).
- [4] A. H. Guth and E. J. Weinberg, *Nucl. Phys.* **B212**, 321 (1983).
- [5] A. D. Linde, *Particle Physics and Inflationary Cosmology* (Harwood, Chur, Switzerland, 1990).
- [6] A. S. Goncharov, A. D. Linde, and V. F. Mukhanov, *Int. J. Mod. Phys. A* **2**, 561 (1987).
- [7] S. Perlmutter *et al.*, *Astrophys. J.* **517**, 565 (1999).
- [8] A. G. Riess *et al.*, *Astron. J.* **116**, 1009 (1998).
- [9] J. Dunkley *et al.*, *Astrophys. J. Suppl. Ser.* **180**, 306 (2009).
- [10] A. D. Sakharov, *Sov. Phys. JETP* **60**, 214 (1984).
- [11] S. Weinberg, *Phys. Rev. Lett.* **59**, 2607 (1987).
- [12] R. Bousso and J. Polchinski, *J. High Energy Phys.* **06** (2000) 006.
- [13] R. Bousso, *Phys. Rev. Lett.* **97**, 191302 (2006).
- [14] R. Bousso, B. Freivogel, and I.-S. Yang, *Phys. Rev. D* **74**, 103516 (2006).
- [15] J. Garriga and A. Vilenkin, *Prog. Theor. Phys. Suppl.* **163**, 245 (2006).
- [16] B. Feldstein, L. J. Hall, and T. Watari, *Phys. Rev. D* **72**, 123506 (2005).
- [17] A. De Simone, A. H. Guth, M. P. Salem, and A. Vilenkin, *Phys. Rev. D* **78**, 063520 (2008).
- [18] D. N. Page, *J. Cosmol. Astropart. Phys.* **10** (2008) 025.
- [19] A. Linde, *Phys. Lett. B* **175**, 395 (1986).
- [20] A. Linde, D. Linde, and A. Mezhlumian, *Phys. Rev. D* **49**, 1783 (1994).
- [21] J. García-Bellido, A. Linde, and D. Linde, *Phys. Rev. D* **50**, 730 (1994).
- [22] J. Garcia-Bellido and A. D. Linde, *Phys. Rev. D* **51**, 429 (1995).
- [23] J. García-Bellido and A. Linde, *Phys. Rev. D* **52**, 6730 (1995).
- [24] A. Linde, *J. Cosmol. Astropart. Phys.* **01** (2007) 022.
- [25] A. Linde, D. Linde, and A. Mezhlumian, *Phys. Rev. D* **54**, 2504 (1996).
- [26] A. H. Guth, *Phys. Rep.* **333–334**, 555 (2000).
- [27] A. H. Guth, arXiv:astro-ph/0002188.
- [28] A. H. Guth, arXiv:astro-ph/0404546.
- [29] M. Tegmark, *J. Cosmol. Astropart. Phys.* **04** (2005) 001.
- [30] A. Linde, *J. Cosmol. Astropart. Phys.* **06** (2007) 017.
- [31] A. H. Guth, *J. Phys. A* **40**, 6811 (2007).
- [32] R. Bousso, B. Freivogel, and I.-S. Yang, *Phys. Rev. D* **77**, 103514 (2008).
- [33] C. M. Baugh, E. Gaztanaga, and G. Efstathiou, *Mon. Not. R. Astron. Soc.* **274**, 1049 (1995).
- [34] A. Vilenkin, arXiv:gr-qc/9512031.
- [35] S. Weinberg, arXiv:astro-ph/9610044.
- [36] H. Martel, P. R. Shapiro, and S. Weinberg, *Astrophys. J.* **492**, 29 (1998).
- [37] A. Vilenkin, arXiv:astro-ph/0407586.
- [38] J. Garriga, D. Schwartz-Perlov, A. Vilenkin, and S. Winitzki, *J. Cosmol. Astropart. Phys.* **01** (2006) 017.
- [39] S. Weinberg, arXiv:hep-th/0511037.
- [40] D. N. Page, *Phys. Rev. D* **78**, 063535 (2008).
- [41] R. Bousso and B. Freivogel, *J. High Energy Phys.* **06** (2007) 018.
- [42] D. Schwartz-Perlov and A. Vilenkin, *J. Cosmol. Astropart. Phys.* **06** (2006) 010.
- [43] D. Schwartz-Perlov, *J. Phys. A* **40**, 7363 (2007).
- [44] R. Bousso and I.-S. Yang, *Phys. Rev. D* **75**, 123520 (2007).
- [45] K. D. Olum and D. Schwartz-Perlov, *J. Cosmol. Astropart. Phys.* **10**, (2007) 010.
- [46] D. Schwartz-Perlov, *J. Cosmol. Astropart. Phys.* **10** (2008) 009.
- [47] B. Freivogel and M. Lippert, *J. High Energy Phys.* **12** (2008) 096.
- [48] B. Freivogel, M. Kleban, M. Rodriguez Martinez, and L. Susskind, *J. High Energy Phys.* **03** (2006) 039.
- [49] R. Bousso, R. Harnik, G. D. Kribs, and G. Perez, *Phys. Rev. D* **76**, 043513 (2007).
- [50] J. M. Cline, A. R. Frey, and G. Holder, *Phys. Rev. D* **77**, 063520 (2008).
- [51] R. Easther, E. A. Lim, and M. R. Martin, *J. Cosmol. Astropart. Phys.* **03** (2006) 016.
- [52] V. Vanchurin and A. Vilenkin, *Phys. Rev. D* **74**, 043520 (2006).
- [53] V. Vanchurin, *Phys. Rev. D* **75**, 023524 (2007).
- [54] A. Vilenkin, arXiv:hep-th/0602264.
- [55] R. M. Wald, *General Relativity* (The University of Chicago Press, Chicago, 1984).
- [56] J. Garriga, A. H. Guth, and A. Vilenkin, *Phys. Rev. D* **76**, 123512 (2007).
- [57] R. Bousso, *J. High Energy Phys.* **04** (2001) 035.