Bayesian reasoning in eternal inflation: A solution to the measure problem

Justin Khoury and Sam S. C. Wong

*Center for Particle Cosmology, Department of Physics and Astronomy, University of Pennsylvania,
Philadelphia, Pennsylvania 19104, USA*

Probabilities in eternal inflation are traditionally defined as limiting frequency distributions, but a unique and unambiguous probability measure remains elusive. In this paper, we present a different approach, based on Bayesian reasoning. Our starting point is the master equation governing vacuum dynamics, which describes a random walk on the network of vacua. Our probabilities require two pieces of prior information, both pertaining to initial conditions; a prior density $\rho(t)$ for the time of nucleation, and a prior probability p_α for the ancestral vacuum. For ancestral vacua, we advocate the uniform prior as a conservative choice, though our conclusions are fairly insensitive to this choice. For the time of nucleation, we argue that a uniform prior is consistent with the time-translational invariance of the master equation and represents the minimally informative choice. The resulting predictive probabilities coincide with Bousso's "holographic" prior probabilities and are closely related to Garriga and Vilenkin's "comoving" probabilities. Despite making the least informative priors, these probabilities are surprisingly predictive. They favor vacua whose surrounding landscape topography is that of a deep funnel, akin to the folding funnels of naturally occurring proteins. They predict that we exist during the approach to near-equilibrium, much earlier than the mixing time for the landscape. We also consider a volume-weighted $\rho(t)$, which amounts to weighing vacua by physical volume. The predictive probabilities in this case coincide with the GSVW measure. The Bayesian framework allows us to compare the plausibility of the uniform-time and volume-weighted hypotheses to explain our data by computing the Bayesian evidence for each. We argue, under general and plausible assumptions, that posterior odds overwhelmingly favor the uniform-time hypothesis.

I. INTRODUCTION

Two separate developments in fundamental physics have led to the seemingly inescapable conclusion that our observable Universe is part of a vast multiverse. The first development is the discovery 40 years ago of eternal inflation [1–5]. It is now understood that eternal inflation is a robust phenomenon that arises for a very wide class of scalar field potentials. The second development is the realization that string theory admits a vast landscape of metastable states [6,7]. Much remains to be understood about the string landscape, in particular the subtle constraints that quantum gravity might place on low-energy physics [8–10], but it seems unambiguous that it is comprised of a myriad of metastable states, giving rise to a rich slew of effective field theories. Eternal inflation offers a mechanism to dynamically populate these vacua, resulting in the multiverse.

As an inhabitant of the multiverse, how should we reason probabilistically about the expected physical properties of our observable universe? Probabilities in eternal inflation are usually defined in terms of frequencies. Concretely, the relative probabilities for two types of events A and B are defined as the ratio of their respective number of instances,

$$\frac{P(A)}{P(B)} = \frac{N_A}{N_B}. \tag{1}$$

The problem is that N_A and N_B are both infinite in an eternally-inflating universe, hence their ratio requires a regularization prescription (or "measure") to be well-defined. This is usually done by defining a finite spacetime region with a geometric cutoff, evaluating N_A and N_B in this region, and taking the limit that the region becomes infinite. Unfortunately the result depends sensitively on the choice of regulator. This is the measure problem.

A variety of different proposals have been put forth over the years (see [11] for a review). Two broad approaches have been pursued:

(i) Global measures define a global foliation of spacetime specified by a global time coordinate. One counts events on a late-time cutoff hypersurface

$t = t_c$, and then lets $t_c \to \infty$ [12–22]. The resulting probabilities are well-defined and independent of initial conditions, consistent with the attractor property of inflation. A major problem is that the result depends sensitively on the choice of time variable [12–14]. For instance, two natural choices, proper time [12,13] and scale factor (or e-folding) time [12,13,20–22], give drastically different probabilities.

The proper-time measure also suffers from the "youngness paradox" [16,23–25]. This problem does not arise with scale factor time, which has led to some convergence that this may be the phenomenologically-sensible foliation.[1] A more radical solution to the youngness paradox is the stationary measure [24,25], which proposes to count different events once they reach the stationary regime.

(ii) Local measures focus on a space-time region around a timelike observer. Examples include the causal diamond measure [27,28], a region bounded by the apparent horizon [29], and a spacelike region around a worldline [18,19, 30,31]. Most important for our purposes is the quasistationary measure of Garriga, Vilenkin, Schwartz-Perlov, and Winitzki (GSVW) [18]. Such constructions are manifestly gauge-invariant. However, because a typical geodesic will eventually enter an AdS or Minkowski vacuum, generally assumed to be terminal, all but a measure zero of watchers will sample a finite number of events, which implies a sensitivity to initial conditions. The ambiguity then lies in defining an ensemble of worldlines.

Interestingly, by choosing suitable initial conditions, it is possible for local measures to agree their global counterpart. This gives rise to the so-called global-local duality [28,32,33].

Aside from the issue of cutoff/regulator dependence, it seems that (1) is ambiguous even in a very large but finite universe. Consider the relative probabilities $P(A)/P(B)$ to inhabit vacuum A or B. It is unclear *a priori* whether one should count the number of bubbles of each type, the fraction of comoving or physical volume for each vacuum, or something else entirely. Simply put, what are the physical observables whose frequencies we are supposed to compare?

In this paper we present an approach to the measure problem that instead relies on the Bayesian framework for probabilities. We adopt the notion of probability theory as an extension of deductive logic. As elegantly enunciated in the classic treatise by Jaynes [34], this is the process of

reasoning by which one draws uncertain conclusions (or "reasonable expectations" [35]) from limited information.

The Bayesian approach in eternal inflation is natural and fruitful for two reasons. First, as emphasized by Hartle and Srednicki [36,37], in a situation like eternal inflation where our data is replicated at (infinitely many) other spacetime locations, a prior assumption must be made about our location within the multiverse. Even if the fundamental theory handed us an unambiguous measure, such an assumption would still be necessary to translate third-person probabilities (what the measure predicts) to first-person probabilities (what *we* are most likely to observe). Thus one is inevitably led to work with first-person probabilities. A second, more practical reason is that all approaches to define a semiclassical measure (including ours) necessarily rely on certain working assumptions. The Bayesian scheme allows one to make all assumptions explicit through careful specification of prior information.

For concreteness, in most of this paper we focus on false-vacuum eternal inflation, though our analysis can be generalized to stochastic (slow-roll) eternal inflation. Our starting point is the master equation governing the probabilities to occupy different vacua [17,18]. This equation describes the Markov process along a given worldline for transitions between transient de Sitter (dS) vacua and into terminal anti–de Sitter (AdS) or Minkowski vacua. The occupational probabilities are normalized at all times and are time-reparametrization invariant.

A unique solution to the master equation requires specifying two pieces of prior information. The first piece of information is the *time of nucleation*. Since eternal inflation is geodesically past-incomplete [38], our bubble universe was nucleated a finite time t after the onset of inflation. The second piece of information is the initial condition. Along our worldline, false-vacuum eternal inflation started in some particular vacuum, which we call the ancestral vacuum α. Thus t and α are model parameters for eternal inflation, and we must specify a prior probability distribution for each. Different approaches to the measure problem amount to different choices for these priors. Consistency requires that our priors reflect all information at hand, but should otherwise be minimally informative.

A. Prior for the ancestral vacuum

The nature of the initial state in quantum cosmology has the subject of much debate over the years. Notable proposals include the Hartle-Hawking state [39,40], which exponentially favors the lowest-energy (highest-entropy) dS vacuum, and the tunneling wave function [41–44], which exponentially favors high-energy (low-entropy) dS vacua. For our purposes, it seems prudent to follow Laplace's principle of indifference and assign a uniform prior for ancestral vacua. Since high-energy dS vacua are expected to vastly outnumber low-energy vacua in the landscape, a uniform prior is statistically equivalent to a

---

[1]The 4-volume cutoff measure [26] proposed recently improves on certain technical drawbacks of the scale factor measure while otherwise making similar predictions.

prior favoring high-energy/low-entropy initial conditions, such as the tunneling wave function. We will discuss this issue in some detail in Sec. IV A.

B. Prior for the time of nucleation

Specifying a prior density $\rho(t)$ is trickier, for the usual reason that an (improper) uniform prior for a continuous parameter is not reparametrization invariant. However the symmetries of the problem at hand offer a natural solution. As argued by Jaynes [34], consistency requires that $\rho(t)$ should be invariant under all transformations that leave our state of knowledge unchanged. Importantly, the master equation, from which our probabilities are derived, is invariant under time translations. A time translation can justifiably have one of two effects on our state of knowledge:

(i) It leaves our state of knowledge unchanged, reflecting complete ignorance about the time of nucleation. This leads us to impose an (improper) uniform prior over proper (or e-folding) time, akin to a temporal Copernican principle.[2] The uniform-time distribution is the least informative prior.

(ii) Alternatively, one might argue that a time translation increases the number of observers proportional to the growth in volume, hence $\rho(t)$ should grow exponentially in time. This effectively corresponds to weighing occupational probabilities by physical volume. For $\rho(t)$ to be well-defined, it is necessary to impose a cutoff time t_c. The resulting prior density is exponentially peaked near t_c, to the extent that it can be well-approximated by a delta function.[3] Thus the late-time or volume-weighted prior is maximally informative.

These two prior densities reflect the assumptions implicit in most approaches to the measure problem. The late-time/volume-weighted prior is closely related to measures based on the late-time, quasistationary distribution [12–15,17, 18,45]. The assumption in this approach is that the evolution of the multiverse has been going on for an exponentially long time, much longer than the mixing time of the landscape. Correspondingly, we will find in Sec. V B that the probabilities with late-time prior coincide with the GSVW quasistationary measure [18]. Because the prior peaks at late times, we refer to this case as the late-time hypothesis $\mathcal{H}_{\text{late}}$.

The uniform-time prior, on the other hand, is representative of local measures, which focus on a spacetime region around a timelike observer [27–29,31,46]. Because a typical worldline ends in a terminal vacuum well-before the mixing time for the landscape, the resulting probabilities differ markedly from the quasistationary distribution. Relatedly, an alternative framework developed recently [47–51] suggests that we live during the approach to equilibrium. (See [52] for related ideas.) The resulting early-time measure [50] favors vacua that can be accessed early on in the evolution, instead of vacua that are frequently generated in the asymptotic future. We will show in Sec. V A that the probabilities with uniform-time prior coincide with Bousso's prior probabilities [27],[4] and are closely related to Garriga and Vilenkin's "comoving" probabilities [18,45]. We refer to this case as the uniform-time hypothesis $\mathcal{H}_{\text{uni}}$.

Our framework allow us to perform the three main operations of Bayesian inference:

(1) By marginalizing over the model parameters t and α, we obtain in Sec. V the prior predictive distributions $P(I|\mathcal{H}_i)$ for each hypothesis. These inform us on which vacua are statistically favored, without taking our data into consideration. They also allow us to compute probability distributions for various physical parameters, in particular the cosmological constant (CC).

(2) The uniform-time and late-time hypotheses will compete in Sec. VI by computing the posterior odds ratio $\frac{P(\mathcal{H}_{\text{late}}|D)}{P(\mathcal{H}_{\text{uni}}|D)}$. Assuming comparable prior odds for the two hypotheses, we will find, under general and reasonable assumptions about transition rates, that posterior odds exponentially favor the uniform-time hypothesis. That is, the GSVW measure [18] is exponentially disfavored compared to the holographic prior probabilities [27]. This is a key result of our analysis. There are possible caveats and loopholes in our analysis, and we will carefully spell these out in Sec. VI D.

(3) Conditioning on our data D, we will perform parameter inference in Sec. VII. We will be specifically interested in the posterior probability distribution $P(t|D, \mathcal{H}_i)$ for the time of nucleation. Focusing on $\mathcal{H}_{\text{uni}}$, we will find that the average time for occupying vacua compatible with our data is much shorter than mixing time, confirming the assumptions underlying the early-time approach [47–51].

---

[2]A uniform prior over the real line is of course ill-defined, so we will need to introduce a late-time cutoff. In the limit that the cutoff is removed, the posterior probabilities will be well-defined and time-reparametrization invariant.

[3]This embodies a version of the youngness paradox [16,23–25], though, as we will see, in our approach there is nothing paradoxical about this choice of prior.

---

[4]As explained in Sec. V A, Bousso first considers the ensemble of possible future "histories" of a worldline. The probability of different histories coincides with our uniform-time predictive probabilities. Bousso then considers the causal diamond of each worldline in the ensemble, and calculates the fraction of observers making different observations within this causal diamond. Because we are interested in prior predictive probabilities, without any anthropic conditioning, this second aspect in Bousso's construction is not relevant for our purposes. In particular, unlike the causal diamond construction, we do not predict that time must "end" in our part of the Universe [53].

We believe that the uniform-time measure is the correct objective approach to inductive reasoning in the multiverse. It consistently reflects our current state of knowledge about the ancestral vacuum and time of nucleation—we simply do not know at this point how/when eternal inflation started in our past. Importantly, despite adopting the least informative priors, the resulting measure is surprisingly predictive. It favors vacua that can be accessed through a sequence of downward transitions, from a large basin of high-energy vacua. Thus such vacua belong to a landscape region with the topography of a funnel [47,48,50], akin to folding funnels of proteins [54]. Furthermore, by predicting that we exist at early times in eternal inflation, the measure implies that we are "normal" observers as opposed to freak observers (Boltzmann brains) [55–59] produced on exponentially longer timescales.

Our approach to the measure problem does not require *ad hoc* geometric constructions, nor are we counting anything. In phrasing the problem in terms of a prior $\rho(t)$ for the time of existence, we were inspired by Caves' elegant approach [60] to the so-called Doomsday paradox [61–63] using similar Bayesian reasoning. Ultimately, despite the conceptual minefield inherent to eternal inflation, landscape dynamics reduce after suitable coarse graining to a linear Markov process, i.e., a random walk on the network of vacua. Such a mathematically simple problem ought to have a simple answer. Indeed, the probabilities obtained with the uniform-time prior are intuitively clear. They favor vacua that are easily accessed under the random walk.

In a forthcoming paper [64] we will show that the measure favors regions of the landscape that are close to the directed percolation phase transition. In other words, the measure selects regions of the landscape that are nearly tuned at criticality.[5] Furthermore, it translates to a probability distribution for the CC that favors a naturally small and positive vacuum energy.

Coincidentally with our paper, Ref. [66] appeared on the arXiv. Although their approach to the measure problem, based on the local Wheeler-de Witt equation, is quite different than the Bayesian method pursued here, the resulting measure appears to be quite similar.

We close this Introduction with some brief remarks about anthropic reasoning. From a Bayesian perspective, anthropic reasoning is intermediate between prior predictive probabilities $P(I|\mathcal{H})$, which are unconditioned, and posterior probabilities $P(\mathcal{H}|D)$, which are conditioned on our data. Anthropic conditionalization is an in-between [67,68], whereby one attempts to condition on the existence of observers. An immediate difficulty, of course, lies in defining "observers" in sufficient generality,[6] so one is forced in practice to condition on some observational proxy for the existence of observers. For instance, one approach is to consider $P(I|\mathcal{H} \cap \mathcal{A})$ [68], which are conditioned on the set of observational proxy $\mathcal{A}$ within the anthropic window. In this sense, within a hypothesis, only the set of universes within this anthropic window is considered. The problem is that the choice of proxy (fraction of baryons per galaxy [69], entropy production [27], or number of observations [70]) is ambiguous and can lead to very different results [70]. For this reason, we avoid any anthropic conditionalization in this work. Our focus is on extracting as much information from prior predictive probabilities.

II. VACUUM DYNAMICS AS ABSORBING MARKOV PROCESS

The landscape can be modeled as a network (or graph) of nodes representing the various dS, AdS, and Minkowski vacua. We assume as usual that AdS and Minkowski vacua are terminal, acting as absorbing nodes. Links define the network topology and represent all relevant transitions between vacua.[7] In what follows, indices $i, j, \dots$ and $a, b, \dots$ denote dS and terminal vacua respectively, while capital indices $I, J \dots$ refer collectively to all vacua. Greek letters $\alpha, \beta, \dots$ denote the ancestral vacuum along our past worldline.

Following the seminal papers of Garriga, Vilenkin and collaborators [17,18], cosmological evolution on the landscape is described by a Markov process. Because of terminals, this is technically an absorbing Markov process—detailed balance is explicitly violated, hence the dynamics are out-of-equilibrium. Along a given worldline, the probability $f_I(\tau_I)$ to occupy vacuum I as a function of the local proper time τ_I satisfies the master equation

$$\Delta f_I = \sum_J \left(\kappa_{IJ}^{\text{proper}} - \delta_{IJ} \sum_K \kappa_{KJ}^{\text{proper}} \right) \Delta \tau_J f_J, \qquad (2)$$

where $\kappa_{IJ}^{\text{proper}}$ is the $J \to I$ proper transition rate. This equation holds for any tunneling mechanism (Coleman-De Luccia (CDL) [71–73], Hawking-Moss [74], Brown-Teitelboim [75]), but whether it applies to "upward" transitions is an open question [17]. For most of our analysis we will remain agnostic about the nature of transition rates.

The master equation relies on coarse graining over a time interval $\Delta \tau_I$, which should be longer than any transient evolution between epochs of vacuum energy domination.

---

[5]Interestingly, this is complementary to the mechanism of 'self-organized localization' [65], whereby the near-criticality of our Universe arises from quantum first-order phase transitions in stochastic inflation. In contrast, our approach pertains to classical, second-order nonequilibrium criticality.

[6]We would be hard-pressed to offer such a definition within our observable Universe, let alone across a multiverse with varying physical constants.

[7]By "relevant" we mean transitions with non-negligible rates on the timescale of interest.

Within our own bubble universe, for instance, this amounts to coarse-graining over at least the last 14 billion years of radiation and matter domination, until vacuum energy comes to completely dominate. On the other hand, $\Delta\tau_I$ cannot be arbitrarily long. It should be shorter than the lifetime of most metastable dS vacua, for otherwise we would be "integrating out" the transitions we are interested in describing. In practice, the coarse-graining time interval for a given transition to I should satisfy $\Delta\tau_I \gtrsim |H_I|^{-1}\log\frac{H_{\text{parent}}}{|H_I|}$, where H_{parent} is the Hubble rate of the parent dS vacuum (see e.g., [76]). Thus, it suffices to assume

$$\Delta\tau_I = |H_I|^{-1}\log\frac{M_{\text{Pl}}}{|H_I|}. \tag{3}$$

In particular, since bubbles of AdS vacua crunch in a Hubble time, coarse graining spans their entire evolution. An AdS bubble nucleated at a given time crunches and dies within a time $\Delta\tau_I$ later.

It is convenient to define a general time variable t, related to proper time via a lapse function,

$$\Delta\tau_I = \mathcal{N}_I\Delta t. \tag{4}$$

In terms of t, (2) becomes

$$\Delta f_I = \sum_J\left(\kappa_{IJ} - \delta_{IJ}\sum_K\kappa_{KJ}\right)\Delta t f_J, \tag{5}$$

where

$$\kappa_{IJ} \equiv \kappa_{IJ}^{\text{proper}}\mathcal{N}_J. \tag{6}$$

Equation (5) makes two properties of the f_I's manifestly clear: (1) The master equation (5) is manifestly invariant under redefinitions of t, hence the f_I's are time-reparameterization invariant; (2) Because summing the right-hand side over I gives zero, the f_I's can be normalized: $\sum_I f_I = 1$. Thus the $f_I(t)$'s give well-defined, gauge-invariant probabilities to occupy different vacua at time t.

In the continuum limit ($\kappa_{IJ}\Delta t \ll 1$), this reduces to

$$\frac{df_I}{dt} = \sum_J\mathbb{M}_{IJ}f_J, \tag{7}$$

where $\mathbb{M}_{IJ} \equiv \kappa_{IJ} - \delta_{IJ}\sum_K\kappa_{KJ}$ is the transition matrix. We will be primarily interested in the dS component of this equation, given by

$$\frac{df_i}{dt} = \sum_j M_{ij}f_j, \tag{8}$$

where

$$M_{ij} \equiv \kappa_{ij} - \delta_{ij}\kappa_j \tag{9}$$

is the dS → dS transition matrix, and $\kappa_i \equiv \sum_J\kappa_{Ji}$ is the total decay rate of vacuum i. (Note that κ_i includes decay channels into dS as well as terminal vacua.) Our only assumption about M_{ij} is that it is irreducible, i.e., there exists a sequence of transitions connecting any pair of dS vacua. This property has been argued to be valid for the string landscape [77].

Importantly, the form of (7) and (8) makes clear that the master equation is time-translation invariant. To be precise, it is invariant under translations of any time variable t related to proper time via a lapse function $\mathcal{N}_I$ that depends on H_I only. This includes proper time as well as e-folding time. Later on we will invoke this symmetry to justify the uniform-time prior.

A. Green's function

Equation (8) can be solved in terms of a Green's function,

$$f_i(t) = \sum_\alpha(e^{Mt})_{i\alpha}p_\alpha, \tag{10}$$

where $p_\alpha \equiv f_\alpha(0)$ is the initial probability over ancestral vacua. Later on, we will need the Laplace transform of the Green's function,

$$\int_0^\infty dt(e^{Mt})_{ij}e^{-st} = (s - M)_{ij}^{-1}. \tag{11}$$

Using (9), this factorizes as

$$(s - M)_{ij}^{-1} = (s + \kappa_i)^{-1}(\mathbb{1} - T(s))_{ij}^{-1}, \qquad T_{ij}(s) \equiv \frac{\kappa_{ij}}{s + \kappa_j}. \tag{12}$$

In particular, $T_{ij} \equiv T_{ij}(0) = \frac{\kappa_{ij}}{\kappa_j}$ is the branching ratio.

The matrix $(\mathbb{1} - T)^{-1}$ is known as the fundamental matrix for the absorbing Markov chain. In the theory of Markov chains, $(\mathbb{1} - T)_{ij}^{-1}$ gives the expected number of visits to i starting from j before reaching terminals.[8] This is easily seen by expanding it as a geometric series, $(\mathbb{1} - T)^{-1} = \mathbb{1} + T + T^2 + \cdots$, and recognizing that the nth term in the series, $(T^n)_{ij}$, represents a branching probability for the n-step chain $j \to \ldots \to i$, summed over all intermediaries. In other words,

$$(\mathbb{1} - T)_{ij}^{-1} = \sum_{\substack{\text{paths}\\ j\to i}}\eta_{ij}, \tag{13}$$

---

[8] In particular, $\sum_i(\mathbb{1} - T)_{ij}^{-1}$ gives the expected number of steps before reaching terminals starting from j.

where $\eta \equiv \prod_{\text{edges}} T$ is the branching probability for each path connecting j to i. For completeness, let us also define the full branching ratio matrix $\mathbb{T}_{IJ}$ as follows:

$$\mathbb{T}_{ij} = T_{ij} = \frac{\kappa_{ij}}{\kappa_j}; \qquad \mathbb{T}_{Ij} = \frac{\kappa_{Ij}}{\kappa_j}; \qquad \mathbb{T}_{ab} = \delta_{ab}; \qquad \mathbb{T}_{ja} = 0.$$

(14)

This ensures that $\mathbb{T}_{IJ}$ satisfies, for all J,

$$\sum_I \mathbb{T}_{IJ} = 1.$$

(15)

B. Detailed balance and downward approximation

Most of our analysis will hold for general transition rates between vacua. To simplify some of the expressions below, however, it will be convenient at some point to make a very general and reasonable assumption about these rates, namely that transitions between dS vacua satisfy a condition of detailed balance [78],

$$\frac{\kappa_{ji}}{\kappa_{ij}} \sim e^{S_j - S_i},$$

(16)

where $S_j = \frac{8\pi^2 M_{\text{Pl}}^2}{H_j^2}$ is the dS entropy. This assumption is satisfied by CDL, Hawking-Moss, and Brown-Teittleboim tunneling, and has been oft-invoked in earlier works on the measure problem. Notably it is violated by the Farvi-Guth-Guven process [79], though the interpretation of its singular instanton remains an open question [80–84]. It is also violated by the mechanism of nucleating localized, high-energy regions proposed recently [85].

Equation (16) implies that upward transitions, which increase the potential energy, are exponentially suppressed compared to downward tunneling. This allows one to define a "downward" approximation [86,87], in which upward transitions are treated perturbatively. Labeling dS vacua for convenience in order of increasing potential energy, $0 < V_1 \leq \ldots \leq V_{N_{\text{dS}}}$, the transition matrix becomes upper-triangular to zeroth order in this approximation,

$$M \simeq \begin{bmatrix} -\kappa_1 & \kappa_{12} & \kappa_{13} & \cdots \\ & -\kappa_2 & \kappa_{23} & \cdots \\ & & -\kappa_3 & \cdots \\ & 0 & & \ddots \end{bmatrix}.$$

(17)

Thus, the eigenvalues at this order are simply by given by its diagonal entries, i.e., by the decay rates κ_i of individual vacua. In particular, the largest (least negative) eigenvalue λ_1 is set by the most stable vacuum, also known as the dominant vacuum $\star$,[9]

$$q \equiv -\lambda_1 \simeq \kappa_\star.$$

(18)

More generally, it can be shown rigorously that $q \leq \kappa_\star$ [18]. The corresponding dominant eigenvector, denoted by s_j, also admits a simple perturbative expression [87],

$$s_j = \delta_{j\star} + \frac{\kappa_{j\star}}{\kappa_j - \kappa_\star} + \sum_{\ell \neq \star} \frac{\kappa_{j\ell}\kappa_{\ell\star}}{(\kappa_j - \kappa_\star)(\kappa_\ell - \kappa_\star)} + \ldots . \quad (19)$$

Assuming that $\kappa_j \gg \kappa_\star$ for all $j \neq \star$, which is reasonable since rates are typically exponentially staggered, the series can be resummed compactly as

$$s_j \simeq \frac{\kappa_\star}{\kappa_j} (\mathbb{1} - T)_{j\star}^{-1}.$$

(20)

Following [87], we have neglected sequences of transitions which return to $\star$ at least once before reaching j, since these involve additional upward transitions and therefore amount to exponentially small corrections.

C. First-passage statistics

First-passage statistics [88] offer a useful tool to study false-vacuum eternal inflation [47–50], and have also been applied to stochastic inflation [89–92]. We briefly review a few elementary results that will be helpful in our analysis.

The first-passage density, $F_{ij}(t)$ ($i \neq j$), is defined as the probability density that a random walker starting from j visits i for the first time at time t.[10] The first-passage density is related to the Green's function through the well-known result [88],

$$(e^{Mt})_{ij} = \int_{0^-}^t dt' (e^{M(t-t')})_{ii} F_{ij}(t').$$

(21)

This equation is valid for all $t \geq 0$ for $i \neq j$, and for all $t > 0$ for $i = j$ with $F_{ii}(t') = \delta(t')$. Thus the occupational probability at time t is the probability of reaching i for the first time at any earlier time t', multiplied by the loop probability for returning to i in the remaining time. For $i \neq j$, the solution is given in terms of Laplace transforms,

$$\tilde{F}_{ij}(s) = \frac{(s - M)_{ij}^{-1}}{(s - M)_{ii}^{-1}} = \frac{(\mathbb{1} - T(s))_{ij}^{-1}}{(\mathbb{1} - T(s))_{ii}^{-1}},$$

(22)

where we have used (12).

All first-passage statistics can be derived from F. For instance, its integral gives the ever-hitting probability,

$$\mathcal{P}_{ij} \equiv \int_0^\infty dt F_{ij}(t) = \tilde{F}_{ij}(0) = \frac{(\mathbb{1} - T)_{ij}^{-1}}{(\mathbb{1} - T)_{ii}^{-1}}; \qquad (i \neq j).$$

(23)

---

[9]It is conceivable that the landscape features a cluster of dominant vacua, with nearly degenerate decay rates. We ignore this possibility for simplicity.

[10]In this paper we will only need dS → dS first-passage statistics. See [48] for dS → AdS results.

Note that this probability is less than unity because of terminals. Meanwhile, its first moment gives the mean first-passage time (MFPT),

$$\langle t_{ij} \rangle \equiv \frac{1}{\mathcal{P}_{ij}} \int_0^\infty \mathrm{d}t\, t F_{ij}(t) = -\frac{\mathrm{d}\ln \tilde{F}_{ij}(s)}{\mathrm{d}s}\bigg|_{s=0}, \qquad (i \neq j). \tag{24}$$

This gives the average time starting from j and conditioned on hitting the target i. Similar expressions can be obtained for the ever-return probability $\mathcal{P}_{ii}$ and mean first-return time $\langle t_{ii} \rangle$ [48]. These expressions simplify in the downward approximation. To zeroth order in this approximation, the network becomes acyclic, and therefore $(\mathbb{1} - T(s))_{ii}^{-1} \simeq 1$. It follows that

$$\mathcal{P}_{ij} \simeq (\mathbb{1} - T)_{ij}^{-1} \qquad (i \neq j). \tag{25}$$

III. BAYESIAN REASONING IN ETERNAL INFLATION

As a first step in defining probabilities, we must carefully distinguish the elements that are inherent to the eternal inflation hypothesis from those that require additional assumptions in the form of prior information.

(i) Since eternal inflation is not past geodesically complete [38], it started a finite time t in our past. We do not know how long ago that was.

(ii) Along our past worldline, false-vacuum eternal inflation, governed by the master equation (7), started out in some ancestral dS vacuum α. We do not know which one.

(iii) Our bubble universe was nucleated at time t in some parent dS vacuum j, but we do not know which one. We will condition our probabilities on one piece of data, namely that we exist in the transient period before vacuum domination, that is, within a coarse-graining time Δt after nucleation.

Our starting point is to define the conditional probability $\wp(j|t,\alpha)$ to occupy parent vacuum j at given time t, given an ancestral vacuum α. The Green's function offers an unambiguous and gauge-invariant definition of this probability,

$$\wp(j|t,\alpha) = (\mathrm{e}^{Mt})_{j\alpha}. \tag{26}$$

The joint probability $\wp(j,t,\alpha)$ is obtained as usual by multiplying with a prior probability distribution,

$$\wp(j,t,\alpha) = (\mathrm{e}^{Mt})_{j\alpha} P_{\mathrm{prior}}(t,\alpha). \tag{27}$$

Since the time of nucleation and ancestral vacuum correspond *a priori* to logically distinct assumptions, it is natural to assume they are independent,

$$P_{\mathrm{prior}}(t,\alpha) = p_\alpha \rho(t). \tag{28}$$

The p_α's, defined in (10), are discrete probabilities for ancestral vacua. They satisfy $\sum_{\alpha=1}^{N_{\mathrm{dS}}} p_\alpha = 1$. Meanwhile, $\rho(t)$ is a prior probability density for the time of nucleation. It satisfies

$$\int_0^\infty \mathrm{d}t\, \rho(t) = 1. \tag{29}$$

We will discuss at length two justified choices for $\rho(t)$ in Sec. IV B. An analogous prior for the time of observation was discussed by Caves [60] in a different context.

The distribution (27) pertains to parent dS vacua. The joint probability distribution $P(I,t,\alpha)$ to inhabit vacuum I within Δt after a nucleation event at time t, starting from α, is given by

$$P(I,t,\alpha) = \mathcal{N} \sum_j \kappa_{Ij} \Delta t\, \wp(j,t,\alpha)$$
$$= \mathcal{N} \sum_j \kappa_{Ij} \Delta t (\mathrm{e}^{Mt})_{j\alpha} p_\alpha \rho(t). \tag{30}$$

As mentioned above, this probability distribution is conditioned on our bubble being nucleated within the last Δt. The normalization constant $\mathcal{N}$ will be fixed shortly. It is important to stress that $P(I,t,\alpha)$ and $\wp(j,t,\alpha)$ are different probabilities, because the former assumes that I is reached within the last Δt.[11] Notice that I can be either dS or terminal, since we are not conditioning on our data at this stage.

A. Prior predictive distribution

Marginalizing over the model parameters t and α gives the prior predictive distribution,

$$P(I) = \sum_\alpha \int_0^\infty \mathrm{d}t\, P(I,t,\alpha)$$
$$= \mathcal{N} \sum_j \kappa_{Ij} \Delta t \int_0^\infty \mathrm{d}t\, f_j(t) \rho(t), \tag{31}$$

where we have used (10). Thus $P(I)$ is the probability to inhabit vacuum I, averaged over all possible time of existence and ancestral vacua. These probabilities inform us on which vacua are statistically favored, without taking our data into consideration (other than conditioning on our bubble being nucleated within the last Δt). The normalization constant is fixed by demanding $\sum_I P(I) = 1$. This yields the normalized probabilities

---

[11]Relatedly, one may be tempted to infer from (30) that $P(I,t,\alpha) = P(I,\alpha|t)\rho(t)$, but this is incorrect. (For instance, the $P(I,\alpha|t)$'s thus defined would not be normalized over I and α.) Instead, the correct conditional probabilities are $P(I,\alpha|t) = P(I,t,\alpha)/\sum_{J,\beta} P(J,t,\beta)$.

$$P(I) = \frac{\sum_j \kappa_{Ij} \int_0^\infty dt f_j(t) \rho(t)}{\sum_k \kappa_k \int_0^\infty dt f_k(t) \rho(t)}, \qquad (32)$$

where we have used $\sum_J \kappa_{Jk} = \kappa_k$. From the $P(I)$'s, one can make predictions for various observables. For instance, the predictive probability density $\rho(\Lambda)$ for the CC is obtained by summing over all vacua with vacuum energy between Λ and $\Lambda + d\Lambda$,

$$\rho(\Lambda) d\Lambda = \sum_{I \in \Lambda < \Lambda_I < \Lambda + d\Lambda} P(I). \qquad (33)$$

B. Parameter inference

A second operation of interest is to use our data D to infer the model parameters α and t. The data refers to all the information available about our observable Universe, in the form of measured values for various observables $\{O_i\}$. These include the particle content, masses and couplings of the Standard Model, as well as the parameters of the cosmological ΛCDM model. A vacuum compatible with our data should, in particular, have a suitable dark matter candidate with correct relic abundance, its vacuum energy should match the observed CC,[12] and its surroundings should allow for a period of slow-roll inflation compatible with the observed amplitude and spectral tilt of primordial perturbations, etc.

At a more philosophical level, there should be more layers of conditioning [67]. For instance, the so-called observed Higgs mass of 125 GeV is not really a type of data D as the true form of data is simply the 125 GeV bump measured from a huge number of scattering events. The Standard Model including the Higgs boson, even quantum field theory, on the other hand, are in fact part of the hypothesis to be tested by these scattering events. Nevertheless, a confidence level of six sigma means that the hypothesis is verified to an extreme extent under this single measurement. Therefore, we can just treat the 125 GeV boson as part of the data D. Despite this, the remarkable accuracy still cannot not stop us from challenging the Standard Model by comparing it with alternatives through other measurements. When it comes to cosmology, hypothesis testing is far less accurate than that in particle scattering as we are limited to a single sample size, yet it is still possible to compare hypotheses using the same parameter inference strategy. Also, due to a higher level of uncertainty, there is no reason not to think of alternatives to cosmological scenarios.

Parameter inference is performed using the posterior probability $P(t, \alpha|D)$. We will be primarily interested in the dependence on the nucleation time, hence we will

---

[12]Hence, we will only include dS vacua when summing over vacua compatible with D.

marginalize over ancestral vacua. The posterior probability for the time of nucleation conditioned on our data is

$$P(t|D) = \frac{\sum_{i \subset I_D} P(i, t)}{P(D)}, \qquad (34)$$

where the denominator $P(D) = \sum_{j \subset I_D} P(j)$ is the probability of our data, and $P(i, t) \equiv \sum_\alpha P(i, t, \alpha)$. To be clear, here I_D denotes the set of all vacua compatible with our data. One can compute various moments of this distribution. For instance, the average time of nucleation conditioned on our data is

$$\langle t \rangle = \int_0^\infty dt\, t P(t|D) = \frac{1}{P(D)} \sum_{i \subset I_D} \int_0^\infty dt\, t P(i, t). \qquad (35)$$

This is just the mean nucleation time for vacua compatible with our data, weighted by the probability for each.

C. Hypothesis testing and posterior odds

The Bayesian framework allows one to compare the plausibility of different hypotheses/models to explain the data through the Bayes factor. As mentioned already, we will be specifically interested in comparing two hypotheses differing only in their priors $\rho(t)$ for the time of existence; a hypothesis $\mathcal{H}_{\mathrm{uni}}$ with quasiuniform $\rho(t)$ and a hypothesis $\mathcal{H}_{\mathrm{late}}$ with a volume-weighted or late-time prior.

The Bayesian evidence for each hypothesis is given by

$$P(D|\mathcal{H}) = \sum_{i \subset I_D} P(i|\mathcal{H}), \qquad (36)$$

where $P(i|\mathcal{H})$ is given by (32) with priors corresponding to $\mathcal{H}$. The relative plausibility of different hypotheses to account for the data is quantified by the Bayes factor,

$$\frac{P(\mathcal{H}_{\mathrm{late}}|D)}{P(\mathcal{H}_{\mathrm{uni}}|D)} = \frac{P(D|\mathcal{H}_{\mathrm{late}})}{P(D|\mathcal{H}_{\mathrm{uni}})} \frac{P(\mathcal{H}_{\mathrm{late}})}{P(\mathcal{H}_{\mathrm{uni}})}, \qquad (37)$$

where $P(\mathcal{H})$ is the prior for each hypothesis, and $\frac{P(D|\mathcal{H}_{\mathrm{late}})}{P(D|\mathcal{H}_{\mathrm{uni}})}$ is the posterior odds. Assuming equal priors, the Bayes factor coincides with the posterior odds.

IV. PRIOR INFORMATION

The crux of deriving a measure lies in the specification of prior probability distributions p_α and $\rho(t)$. The choice of priors is a long-standing problem in probability theory, for which there is unfortunately no universally applicable rule. The following general guiding principles have proven fruitful in other contexts.

Consistency requires that priors reflect all information at hand. In the case of interest, this includes state-of-the-art understanding of the string landscape, quantum cosmology etc. At the same time, our priors should be minimally

informative. They should incorporate all available information, but should not otherwise be constrained by biases or prejudices. In practice, this is often achieved by applying the principle of indifference or, more generally, the principle of maximal entropy [34]. Lastly, our priors should be objective, in that two physicists presented with the same information should agree upon a set of justified priors.

Fortunately, the problem at hand is sufficiently simple and symmetric that the set of justified priors, we will argue, can be narrowed down to essentially just two possibilities.

A. Prior distribution for the ancestral vacuum

The prior distribution p_α for the ancestral vacuum pertains to the question of initial conditions. The question of the initial state in quantum cosmology remains an open problem and has been the subject of active debate for decades.

A well-motivated proposal for the quantum creation of a closed universe is the Hartle-Hawking (HH) state [39,40],

$$P_{\text{HH}} \sim e^{\frac{24\pi^2 M_{\text{Pl}}^4}{V}}, \tag{38}$$

which exponentially favors the lowest-energy (highest entropy) dS vacuum. Another well-studied proposal is the tunneling wave function [41–44],

$$P_{\text{tunneling}} \sim e^{-\frac{24\pi^2 M_{\text{Pl}}^4}{V}}. \tag{39}$$

Because of the crucial sign difference in the exponent, this wave function instead favors high-energy/low-entropy initial vacua. Thus the tunneling wave function favors (high-energy) inflation, whereas the HH state does not [93]. Without entering this long-standing debate, it is worth noting that the HH state, obtained by analogy with the ground state of the harmonic oscillator, corresponds to the equilibrium state of a dS-only landscape. Since the equilibrium state is reached typically after an exponentially long relaxation time, its motivation as an initial state seems unclear. This is further complicated by the presence of terminal (AdS) vacua in any realistic landscape.

For our purposes, it seems prudent to err on the side of maximal ignorance and apply the principle of indifference,[13]

$$p_\alpha = \frac{1}{N_{\text{dS}}}. \tag{40}$$

Since high-energy dS vacua are expected to vastly outnumber low-energy vacua in the landscape, a uniform prior

is statistically equivalent to a prior favoring high-energy/low-entropy initial conditions, such as the tunneling wave function (39).[14] If it turns out, however, that the HH state (38) is the correct initial conditions for eternal inflation in our past, then this would have important implications for our results. From an information theory perspective, because the HH wave function exponentially favors the absolutely lowest dS vacuum, it amounts to a highly informative prior on the ancestral vacuum. In contrast, the tunneling wavefunction assigns comparable probability to all high energy ($V \sim M_{\text{Pl}}^4$) vacua, hence it represents a minimally informative prior.

We will see that our results are highly insensitive to the prior over ancestral vacua. For all but a very special subset of initial conditions, we will argue that the uniform-time hypothesis is exponentially preferred over the late-time hypothesis. Furthermore, the master equation also implies that the initial time t_0, which we set to be 0, can be traded for a different set of p_α, as one can evolve backward or forward in time to get a different initial distribution in the landscape. This ambiguity is not a big issue as we are comparing the two hypotheses, $\mathcal{H}_{\text{uni}}$ and $\mathcal{H}_{\text{late}}$, under the same prior in ancestral vacua p_α. For the sake of generality we will leave p_α arbitrary in our analysis below, though we implicitly have in mind the uniform prior (40).

B. dS isometries and prior density for the time of existence

Applying the principle of indifference (or maximal entropy) to a continuous variable is tricky, simply because a uniform prior for a continuous variable is not reparametrization invariant. An obvious strategy is to work with a discretized time variable. We will pursue this in Appendix A, and show that the resulting measure agrees exactly with the continuous-time approach developed here. This offers a nontrivial check on the time-reparametrization invariance of the measure.

Sticking with continuous time, it is instructive to consider the symmetries of the problem and apply the notion of group invariance. As argued by Jaynes [34], one must first identify all transformations that leave our state of knowledge unchanged. Consistency then requires that the prior probabilities be invariant under those transformations. For instance, if one's prior state of knowledge is oblivious to a spatial translation $x \rightarrow x + c$, then the prior distribution for x should be uniform. If one's state of knowledge is instead

---

[13]The number of dS vacua may well be infinite in the string landscape [94–96], in which case (40) would represent an improper prior. This is of no concern as the resulting probabilities would nevertheless be well-defined.

[14]In fact, the tunneling wave function is precisely equivalent to a uniform prior if one considers hyperbolic or flat spatial geometry with nontrivial topology (e.g., flat 3-torus), instead of the closed universe originally considered in [41–44]. Since an open/flat universe can evolve classically from $a = 0$ without any tunneling, the resulting wave function is not exponentially suppressed for $V \ll M_{\text{Pl}}^4$ [97,98]. We thank the anonymous referee for suggesting this argument.

invariant under a scale transformation $x \to \lambda x$, then the appropriate choice should be the Jeffreys prior, $\rho(x) \sim 1/x$.

In our case we are guided by the symmetries of the master equation and the isometries of dS space. Consistency requires that $\rho(t)$ be invariant under these transformations. As we will see, these considerations lead to two possible priors, which reflect the well-known dichotomy between local and global approaches to the measure problem.

(i) *Time-translation invariance and uniform-time prior:* As mentioned earlier, a key observation is that the master equation (7) is time-translation invariant. More precisely it is invariant under translations of any time variable t related to proper time via a lapse function $\mathcal{N}_I$ that is a function of H_I only. Without any additional information, the uniform prior is the most intuitive and reasonable choice.

Let us warm up using a simple but well known example: the diffusion problem of a particle in one dimension. Given $P(x|t)$ as the solution to the diffusion equation, which is time-translation invariant, one may ask for the posterior probability $P(t|x)$, which is the distribution of lapsed time t given that the particle is observed at location x. Without further information, the most intuitive way to calculate such quantity is simply

$$P(t|x) = \frac{P(x|t)}{\frac{1}{T}\int_0^T \mathrm{d}t' P(x|t')}. \qquad (41)$$

By consulting Bayes' theorem, $P(t|x)P_x(x) = P(x|t)P_t(t)$, the intuitive answer corresponds exactly to using a uniform prior with a cutoff in time T, which can be sent to infinity. The prior $P_t(t)$ in this simple example can be regarded as the distribution of the time of observation. For instance, if instead one were told that most of the observations were concentrated within a particular time window, it would be appropriate in this case to use a nonuniform $P_t(t)$.

The Markov process described by the master equation $\dot{f}_I = \sum_J \mathbb{M}_{IJ}f_J$ is analogous to the simple diffusion problem. Given a distribution of initial condition p_α, the solution to the master equation,

$$P(I|t) = f_I(t) = \sum_\alpha (e^{\mathbb{M}t})_{I\alpha}p_\alpha, \qquad (42)$$

is exactly the conditional probability of being in vacuum I given time t, as explained in the previous sections. To find the distribution of lapsed time given that the state is observed in vacuum I, $P(t|I) = \frac{P(I|t)P_t(t)}{P(I)}$, without additional information, the most natural choice of prior $\rho(t) = P_t(t)$ with maximal entropy is a uniform prior,

$$\rho(t) = \text{constant}. \qquad (43)$$

Of course, the uniform distribution on the half real line is not normalizable, and a regularization is needed. A mathematically convenient choice is

$$\rho(t) = \epsilon e^{-\epsilon t}, \qquad (44)$$

with a cutoff $\sim\epsilon^{-1}$ in time, therefore

$$P(t|I) = \frac{P(I|t)e^{-\epsilon t}}{\int_0^\infty \mathrm{d}t' P(I|t')e^{-\epsilon t'}}. \qquad (45)$$

At the end of the calculation we will send $\epsilon \to 0$ and obtain well-defined, gauge-invariant posterior probabilities.[15]

A geometric way to interpret this prior is to focus on one of the metastable dS vacua in our past, for instance our parent vacuum. Assuming it is sufficiently long-lived, the parent geometry is approximately invariant under the 10 dS isometries. Taking a random set of worldlines with initial condition p_α, the chance that a randomly chosen worldline is in vacuum I at time t is exactly predicted by the solution to the master equation.

The action for local operators on the worldline is of the general form

$$S = \int \mathrm{d}\tau \mathcal{O}(X^\mu(\tau)), \qquad (46)$$

where the measure $\mathrm{d}\tau$ is proper time along the curve. After sufficient time, the worldline approaches a geodesic comoving ($\vec{x} = \text{const}$) in the dS flat slicing,

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + e^{2Ht}\mathrm{d}\vec{x}^2. \qquad (47)$$

Hence, τ coincides with cosmic proper time t, and the resulting measure $\mathrm{d}\tau \simeq \mathrm{d}t$ on the worldline is time-independent. This simple line of reasoning suggests that the prior density $\rho(t)$ be (proper) time-translation invariant. Note that the prior is also quasiuniform in e-folding time, $\rho(N) = \frac{\epsilon}{H}e^{-\frac{\epsilon}{H}N}$. In terms of conformal time, it corresponds to the Jeffreys prior, $\rho(\eta) \sim \eta^{-1}$, consistent with the dS dilation symmetry $\eta \to \lambda\eta$, $\vec{x} \to \lambda\vec{x}$.

---

[15]One should stress that the nature of ϵ is quite different than the late-time cutoff usually introduced to regularize ratios of number counts, such as (1). In our approach ϵ is only necessary to make the improper uniform prior well-defined in the intermediate steps. Importantly, none of our results are sensitive to the choice of regulator. For instance, choosing $\rho(t) = 1/T$ over a large but finite interval $0 \le t \le T$, yields identical results in the limit $T \to \infty$.

Heuristically, a time-translation invariant prior reflects complete ignorance about the time of nucleation, akin to a temporal Copernican principle. We simply do not know when eternal inflation started along our particular past worldline. Correspondingly, this is the minimally informative prior for the time of nucleation.

(ii) *Volume-weighted prior:* An alternative approach is to consider a finite spatial region in the parent dS geometry. The invariant measure for observables within the region is the usual volume element,

$$\sqrt{-g}\,\mathrm{d}t\,\mathrm{d}^3x = a^3(t)\mathrm{d}t\,\mathrm{d}^3x. \qquad (48)$$

Averaging over $\vec{x}$ yields a measure $\sim a^3(t)\mathrm{d}t$ that grows with volume. This line of reasoning suggests that the prior density $\rho(t)$ should similarly grow with volume,

$$\rho(t) \sim a^3(t). \qquad (49)$$

This choice can also be motivated heuristically as follows. A time translation in dS corresponds to an exponential growth in volume, and therefore an exponential growth in the number of observers. If one abides by the self-indication assumption [99], whereby prior probabilities are weighted by the number of observers produced, then $\rho(t)$ should increase exponentially in time.[16]

For the prior distribution to be normalizable, a regulator is once again necessary. This can be achieved simply by imposing a cutoff time t_c,

$$\rho(t) = \frac{a^3(t)}{\int_0^{t_c} \mathrm{d}t'\, a^3(t')} \quad \text{for} \quad 0 \le t \le t_c, \quad (50)$$

and $\rho(t) = 0$ for $t > t_c$. As in the previous case, we will remove the regulator ($t_c \to \infty$) at the end of the calculation and obtain well-defined, gauge-invariant posterior probabilities. In fact, we will see that *any* growing $\rho(t)$ that dominates after the relaxation timescale of the landscape gives the same probabilities.

Physically, since $\rho(t)$ multiplies the occupational probabilities $f_j(t)$ [see (32)], this choice amounts to weighing probabilities by physical volume. We will show this explicitly in Sec. V C. In standard approaches to the measure problem, volume-weighing

leads to exponential sensitivity on the choice of time variable [12,14,17]. In contrast, because volume weighing is implemented through the prior density $\rho(t)$ in our context, the resulting probabilities will be manifestly gauge invariant.

Since the physical volume of all bubble universes grows in proper time asymptotically as $\sim e^{3H_{\max}t}$, where $H_{\max}$ is the highest Hubble rate of any dS vacuum in the landscape [13,17,101], (49) is effectively equivalent to, in proper time,

$$\rho(t) \sim e^{3H_{\max}t}. \qquad (51)$$

And because $H_{\max}$ is enormous compared to transition rates governing the evolution of the $f_j(t)$'s, (51) is well-approximated by a delta function

$$\rho(t) \simeq 2\delta(t - t_c); \qquad 0 \le t \le t_c. \qquad (52)$$

Thus, because the prior is so sharply peaked, it is maximally informative. We will make this statement precise below once we derive the prior predictive probabilities in this case. The fact that $\rho(t)$ is sharply peaked at the cutoff time is related to the youngness paradox [16,23–25] afflicting global measures based on the proper time cutoff. No such paradox arises in our framework, as far as we can tell.

The two justified prior densities derived above are consistent with the assumptions implicit in most approaches to the measure problem. The volume-weighted prior (50) is reflected in measures based on the late-time, quasistationary distribution [12–15,17,18]. For this reason, we refer to this case as the late-time hypothesis $\mathcal{H}_{\text{late}}$. Meanwhile, the uniform prior (43) is representative of local measures [27–29,31]. It is also consistent with the early-time approach to eternal inflation [47–51], which postulates that we exist well-before the exponentially long mixing time for the landscape. We henceforth refer to this prior as uniform-time hypothesis $\mathcal{H}_{\text{uni}}$.

V. PRIOR PREDICTIVE DISTRIBUTIONS

With the above priors at hand, we are now in a position to calculate probabilities relevant to Bayesian inference. In this Section we focus on the prior predictive distribution $P(I)$, which gives the probability to occupy vacuum I irrespective of the time of existence or ancestral vacuum. We will find that the resulting probabilities $P(I|\mathcal{H}_{\text{uni}})$ and $P(I|\mathcal{H}_{\text{late}})$ coincide with two measures proposed in the literature, respectively the holographic prior probabilities [27] and the quasistationary measure of GSVW [18]. Furthermore, each has a close analog among centrality indices studied in network science.

---

[16]The self-indication assumption can lead to absurd conclusions in other contexts, such as the "presumptuous philosopher" problem [99]. Neal instead advocates applying what he calls full nonindexical conditioning [100]. We will argue in Sec. VI that the late-time hypothesis is disfavored by the data, hence debating the relative merits of these assumptions in eternal inflation is perhaps moot.

A. Uniform-time prior

We first consider the uniform prior for the time of existence. Substituting (44) into (32), and using (12), we obtain

$$P(I|\mathcal{H}_{\text{uni}}) = \frac{\sum_j \mathbb{T}_{Ij}(\epsilon)\sum_\alpha (1 - T(\epsilon))^{-1}_{j\alpha} p_\alpha}{\sum_{J,k} \mathbb{T}_{Jk}(\epsilon)\sum_\gamma (1 - T(\epsilon))^{-1}_{k\gamma} p_\gamma}, \quad (53)$$

where $\mathbb{T}_{Ij}(\epsilon) = \frac{\kappa_{Ij}}{\epsilon + \kappa_j}$. Letting $\epsilon \to 0$ to remove the regulator gives

$$P(I|\mathcal{H}_{\text{uni}}) = \frac{\sum_j \mathbb{T}_{Ij}\sum_\alpha (1 - T)^{-1}_{j\alpha} p_\alpha}{\sum_{k,\beta} (1 - T)^{-1}_{k\beta} p_\beta}, \quad (54)$$

where we have used $\sum_J \mathbb{T}_{Jk} = 1$ to simplify the denominator. The above probabilities only depend on branching ratios, and therefore are invariant under time reparametrizations.[17] They coincide with Bousso's "prior probabilities" [27], derived following a different line of reasoning, and are also closely related to Garriga and Vilenkin's "comoving" probabilities [18,45]. In our approach, these probabilities are an inevitable consequence of justified objective reasoning with complete ignorance about the time of existence.

To be precise, Bousso's holographic measure is the product of two factors. First, one considers a single worldline and studies the ensemble of possible future "histories" of that worldline. The relative probability of different histories, which Bousso calls "prior probabilities", is given by their branching ratio probabilities. This is the first factor in Bousso's measure, and it exactly matches (54). Bousso then considers the causal diamond for each worldline in the ensemble, and calculates the fraction of observers making different observations within this causal diamond. The fraction of observers is the second factor in his measure. Because we are interested in prior predictive probabilities, without any anthropic conditioning, this second factor is not relevant for our purposes.

The prior predictive distribution (54) can be easily understood intuitively. Recall from (13) that $(1 - T)^{-1}_{ij}$ is the sum of branching probabilities over all paths connecting j to i. Thus $P(I|\mathcal{H}_{\text{uni}})$ is naturally interpreted as the sum over all paths connecting ancestral vacua to vacuum I, weighted by the branching probability for each path and averaged over ancestral vacua. It follows that the probabilities (54) are maximized for vacua that are well-connected ($\mathbb{T}_{Ij} \simeq 1$) to parent dS vacua j which themselves are easily accessed.

The above probability distribution has a close analog among centrality indices studied in network science.

Various centrality indices have been proposed in network theory to quantify which nodes in a graph are, in a suitably defined sense, most important [102]. With uniform prior $p_\alpha = N^{-1}_{\text{dS}}$ over ancestral vacua, (54) is similar to the Katz centrality measure [103] on a graph,

$$\vec{C}_{\text{Katz}} = ((1 - \alpha A^{\mathsf{T}})^{-1} - 1)\vec{e}, \quad (55)$$

where $\vec{e} = (1, \ldots, 1)$ is a vector with unit entries, A is the graph adjacency matrix, and $0 < \alpha < 1$ is a so-called attenuation parameter. Intuitively, Katz centrality favors nodes that are well-accessed from other nodes in the network.

Funnel topography: It is also instructive to consider the prior predictive probabilities in the downward approximation (Sec. II B). Using (25) we have

$$\sum_\alpha (1 - T)^{-1}_{j\alpha} p_\alpha \simeq p_j + \sum_{\alpha \neq j} \mathcal{P}_{j\alpha} p_\alpha \equiv p_j + \mathcal{P}_j, \quad (56)$$

where in the last step we have defined $\mathcal{P}_j \equiv \sum_{\alpha \neq j} \mathcal{P}_{j\alpha} p_\alpha$ as the ever-hitting probability to j, averaged over initial conditions. Hence (54) becomes

$$P(I|\mathcal{H}_{\text{uni}}) \simeq \frac{\sum_j \mathbb{T}_{Ij}(p_j + \mathcal{P}_j)}{1 + \sum_k \mathcal{P}_k}. \quad (57)$$

Since this result assumes the downward approximation, the only contributing paths to a given vacuum I are those given by a sequence of downward transitions. Thus the probabilities (57) favor vacua that can be accessed through downward transitions, from a large basin of ancestral vacua. Regions of the landscape with large probability must therefore have the topography of a deep valley, or funnel [47–50]. See Fig. 1. This is akin to the smooth folding funnels of energy landscapes of proteins [54]. Naturally occurring proteins fold efficiently because their free-energy landscape is characterized by a smooth funnel near the native state.

B. Late-time/volume-weighted prior

We next consider the late-time prior (50) for the time of existence, keeping the prior p_α over ancestral vacua once again general. Because the prior is sharply peaked near t_c, which is assumed very large, we can approximate the occupational probabilities $f_j(t)$ by their asymptotic form

$$f_j(t) \simeq s_j e^{-qt}, \quad (58)$$

where s_j the dominant eigenvector of M_{ij} with largest (least negative) eigenvalue $-q$ [18], defined in Sec. II B. With (50) and (58), the prior predictive distribution (32) gives

---

[17]This should be obvious, and to prove it we only need (6) to infer that $\frac{\kappa_{Ij}}{\kappa_j} = \frac{\kappa^{\text{proper}}_{Ij}}{\kappa^{\text{proper}}_j}$.
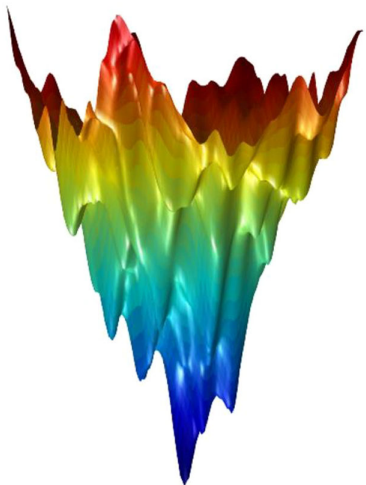
FIG. 1.   The prior predictive probabilities with uniform-time prior favor regions of the landscape where vacua can be accessed through a sequence of downward transitions, from a large basin of ancestral vacua. Their topography is that of a deep valley, or funnel, akin to the smooth folding funnels of energy landscapes of naturally-occurring proteins. (Reproduced from [104].).

$$P(I|\mathcal{H}_{\text{late}}) \simeq \frac{\sum_j \kappa_{Ij} s_j \int_0^{t_c} \mathrm{d}t\, a^3(t) \mathrm{e}^{-qt}}{\sum_k \kappa_k s_k \int_0^{t_c} \mathrm{d}t\, a^3(t) \mathrm{e}^{-qt}} = \frac{\sum_j \kappa_{Ij} s_j}{\sum_k \kappa_k s_k}. \quad (59)$$

Although we assumed the prior density (50), the above makes clear that the result holds for any $\rho(t)$ that peaks at sufficiently late times [including the delta function prior (52)] such that (58) is valid. Thus the above probability distribution is manifestly time-reparametrization invariant. It is also independent of the late-time cutoff, such that the limit $t_c \to \infty$ can be taken. Importantly, the result agrees with the GSVW measure [18] obtained by counting bubbles along a worldline.[18] To be precise, GSVW obtained, working in e-folding time,

$$p_I \sim \sum_j H_j^q \kappa_{Ij} s_j. \quad (60)$$

This differs from (59) only by the factor H_j^q, which is indistinguishable from unity since $q \le \kappa_\star \lll 1$.

The prior predictive distribution also admits an intuitive explanation in downward perturbation theory. Substituting (20) for the dominant vector, (59) becomes

$$P(I|\mathcal{H}_{\text{late}}) \simeq \frac{\sum_j \mathbb{T}_{Ij} (\mathbb{1}-T)_{j\star}^{-1}}{\sum_k (\mathbb{1}-T)_{k\star}^{-1}}. \quad (61)$$

---

[18]GSVW implicitly assumes either that this worldline survives long enough to probe the asymptotically late-time dynamics, or that their measure is the result of averaging over an ensemble of worldlines such that late-time observers dominate the average.

The interpretation is clear—these probabilities are maximized for vacua that are well-connected ($\mathbb{T}_{Ij} \simeq 1$) to parent dS vacua j, which are themselves easily accessed from the dominant vacuum. Remarkably, (61) coincides with the uniform-time probabilities (54) for the special case that $p_\alpha = \delta_{\alpha\star}$, i.e., when the ancestral vacuum is the dominant vacuum. (This is a manifestation of the local/global duality [32,33].) Relatedly, it is often stressed that late-time measures, such as the GSVW measure, reflect the attractor nature of eternal inflation through their independence on initial conditions. This is certainly the case with (59). Paradoxically, however, the late-time distribution (61) appears to be a special case, corresponding to a particular choice of initial conditions, of the seemingly more general distribution (54).

Because the late-time prior peaks at the cutoff time, it is a maximally informative prior. To see this concretely, consider the probability density $P(t|I)$ for the time of existence, conditioned on occupying vacuum I:

$$P(t|I) = \frac{P(I,t)}{P(I)}. \quad (62)$$

The joint probability is obtained by marginalizing (30) over α,

$$P(I,t) \sim \sum_j \kappa_{Ij} f_j(t) \rho(t) = \sum_j \kappa_{Ij} s_j \mathrm{e}^{-qt} \rho(t). \quad (63)$$

Combining this with (51), and using the fact that $H_{\max} \gg q$, it is easy to see that

$$P(t|I) \simeq \rho(t). \quad (64)$$

Thus, even if we had we had complete knowledge of which vacuum we inhabit, we would learn nothing more about the time of existence than already assumed with this prior.

The probability distribution (59) also has a close cousin among centrality indices in network theory. Namely, it is analogous to eigenvector centrality, which ranks nodes according to the components of the dominant eigenvector $\vec{\pi}$ of the graph's adjacency matrix A.

C. Alternative derivation of volume-weighted probabilities

Lastly, let us give an alternative derivation of (59) in terms of physical volume. Let $\tilde{f}_i(t) = a^3(t) f_i(t)$ denote the fraction of physical volume in dS vacuum i. Then $\tilde{f}_i$ satisfies the physical-volume master equation

$$\frac{\mathrm{d}\tilde{f}_i}{\mathrm{d}t} = \sum_j M_{ij} \tilde{f}_j + 3\mathcal{N}_i H_i \tilde{f}_i. \quad (65)$$

Following the steps in Sec. III, we arrive at the analog of (32) in terms of $\tilde{\tilde{f}}_i$,

$$P(I|\mathcal{H}_{\text{vol}}) \sim \sum_j \kappa_{Ij} \int_0^\infty \mathrm{d}t\, \tilde{\rho}(t) \tilde{\tilde{f}}_j(t), \qquad (66)$$

where $\tilde{\rho}(t)$ is the prior density for the time of existence in the volume-weighted case.

Importantly, the master equation (65) is time-translation invariant for any choice of lapse function $\mathcal{N}_i$ that only depends on H_i, such as proper time or e-folding time. Hence, a justified choice of prior is the uniform one,

$$\tilde{\rho}(t) \simeq \text{const.} \qquad (67)$$

Using this and the relation $\tilde{\tilde{f}}_i = a^3(t) f_i(t)$, the time integral in (66) gives

$$\int_0^\infty \mathrm{d}t\, \tilde{\rho}(t) \tilde{\tilde{f}}_j(t) = \int_0^\infty \mathrm{d}t\, a^3(t) f_j(t). \qquad (68)$$

Thus (66) agrees with (32) with volume-weighted prior $\rho(t) \sim a^3(t)$. In other words, the comoving probabilities with volume-weighted $\rho(t)$ are precisely equivalent to physical-volume probabilities with uniform $\tilde{\rho}(t)$.

VI. MODEL COMPARISON AND POSTERIOR ODDS

In the previous sections we derived, by applying consistent objective reasoning, two hypotheses for probabilities in eternal inflation, which differ only in their priors $\rho(t)$ for the time of existence; a hypothesis $\mathcal{H}_{\text{uni}}$ with quasiuniform $\rho(t)$, which reflects complete ignorance about the time of existence; and a hypothesis $\mathcal{H}_{\text{late}}$ with volume-weighted $\rho(t)$, which reflects the belief that we exist at asymptotically late times. In this section we compare the plausibility of the two hypotheses to explain our data D by computing the Bayes factor discussed in Sec. III C.

Substituting the prior predictive probabilities (54) and (61), the Bayesian evidence (36) for each hypothesis is

$$P(D|\mathcal{H}_{\text{uni}}) = \frac{\sum_{i \subset I_D, j} T_{ij} \sum_\alpha (\mathbb{1} - T)^{-1}_{j\alpha} p_\alpha}{\sum_{k,\beta} (\mathbb{1} - T)^{-1}_{k\beta} p_\beta}, \quad (69\text{a})$$

$$P(D|\mathcal{H}_{\text{late}}) \simeq \frac{\sum_{i \subset I_D, j} T_{ij} (\mathbb{1} - T)^{-1}_{j\star}}{\sum_k (\mathbb{1} - T)^{-1}_{k\star}}, \qquad (69\text{b})$$

where the latter is valid in downward perturbation theory. Before attempting to estimate (69b), it is useful to get a sense of the hierarchy of transition rates and branching ratios involved.

A. Detailed balance and unsuppressed chains

A generic feature of transition rates in field theory is that they are exponentially staggered. This is because rates depend exponentially on the shape of the potential, such as the height and width of the barrier. Hence branching ratios are typically overwhelmingly dominated by a single decay channel, while other decay channels are comparatively exponentially suppressed.[19]

Because of detailed balance, upward transitions are further suppressed by a double exponential factor. To see this, note from (16) that the upward vs downward rate between two vacua with $V_{\text{low}} \ll V_{\text{high}}$ is

$$\frac{\kappa_{\text{up}}}{\kappa_{\text{down}}} \sim e^{-S_{\text{low}}}. \qquad (70)$$

To get a sense of the suppression, for a vacuum with $V_{\text{low}} = (0.1 M_{\text{Pl}})^4$, i.e., with energy scale just one order of magnitude below the Planck scale, the dS entropy is $S_{\text{low}} \simeq e^{15}$. A vacuum with our observed CC, $V_{\text{low}} = 10^{-120} M_{\text{Pl}}^4$, has $S_{\text{low}} \simeq e^{283}$. The conclusion is that, if a dS vacuum has one or more available downward decay channels, these will typically have overwhelmingly dominant branching ratios.

On the other hand, if a vacuum can only decay via upward transitions, then the branching ratio is necessarily dominated by an upward jump.[20] We will refer to this decay channel as the dominant upward transition. To gain some intuition on what constitutes a dominant upward transition, suppose that a vacuum can either up-tunnel via a high jump to V_{high}, or via a smaller jump to V_{low}, once again with $V_{\text{low}} \ll V_{\text{high}}$. In this case, (16) implies

$$\frac{\kappa_{\text{high jump}}}{\kappa_{\text{low jump}}} \sim \frac{\kappa_{\text{large step down}}}{\kappa_{\text{small step down}}} e^{-S_{\text{low}}}. \qquad (71)$$

The prefactor $\kappa_{\text{large step down}}/\kappa_{\text{small step down}}$ is just the ratio of downward rates to the given vacuum, and its value depends on the details of the potential. It is at best exponentially large and at worse exponentially small. But, in either case, it is generically swamped by the double exponential factor of $e^{-S_{\text{low}}}$, and thus $\kappa_{\text{high jump}} \ll \kappa_{\text{low jump}}$. In other words, the dominant upward transition typically corresponds to the smallest increase in potential energy.

These considerations lead us to define an unsuppressed chain as a sequence of transitions comprised exclusively of downward transitions (whenever these are available) and/or dominant upward transitions (whenever a vacuum can only

---

[19]There are of course exceptions, for instance in regular lattices of flux vacua [6], but one expects that single-channel dominance is justified for random landscapes.

[20]In this case the lifetime is of order the dS recurrence time. It has been conjectured that in string theory all dS vacua have a much shorter lifetime [105], such that the problem of Boltzmann brains is avoided.

decay by jumping upwards). In other words, an unsuppressed chain excludes subdominant upward jumps, whose branching ratio is suppressed by a double exponential. And because downward channels dominate whenever they are available, unsuppressed chains consist mostly of downward transitions. We briefly note a few properties of unsuppressed chains:

(i) There is at least one unsuppressed chain starting from any vacuum, but any number of them (including zero) can arrive at that vacuum.

(ii) If an unsuppressed chain exists from j to i, then $(\mathbb{1} - T)^{-1}_{ij}$ is at worse exponentially suppressed; if no unsuppressed chain exists, then $(\mathbb{1} - T)^{-1}_{ij}$ is doubly-exponentially suppressed.

(iii) If an unsuppressed chain exists from j to i, in general no such chain exists from i to j.

With this intuition and definitions at hand, we can now examine more closely the Bayesian evidence (69) for each hypothesis.

B. Evidence for late-time hypothesis

Let us first examine the Bayesian evidence for the late-time hypothesis, given by (69b). Clearly the answer depends on the nature of the dominant vacuum and its surrounding landscape, which are of course unknown. However, given $\star$'s status as the most stable vacuum anywhere in the landscape, it is safe to assume that: (i) it can only decay via an upward transition (because upward jumps are doubly-exponentially suppressed); and (ii) it has very small potential energy (because the upward rate is suppressed by $e^{-S_\star}$). These assumptions are not new and have been made in other studies of the GSVW measure, e.g., [21,87].

From our earlier discussion, the dominant upward transition from $\star$ likely proceeds in the direction of smallest increase in potential energy. Following this initial jump, the unsuppressed chains that emanate from $\star$ proceed as sequences of downward and dominant upward transitions, until they terminate at terminal vacua. The denominator of (69), $\sum_k (\mathbb{1} - T)^{-1}_{k\star}$, gives the expected number of dS vacua that are visited in the process. All we need for our purposes is that this number is greater than unity, since

$$\sum_k (\mathbb{1} - T)^{-1}_{k\star} = \sum_k (\delta_{k\star} + T_{k\star} + \dots) > 1. \quad (72)$$

Meanwhile, the numerator, $\sum_{i \subset I_D} \sum_j T_{ij}(\mathbb{1} - T)^{-1}_{j\star}$ gives the branching probability to reach any vacuum compatible with our data starting from $\star$. However, because such vacua are (presumably) rare in the landscape, it is highly unlikely that such a vacuum lies along an unsuppressed chain from $\star$. Instead, as argued in [87], the most probable path to our vacuum likely requires additional (subdominant) upward jumps, such that $\sum_{i \subset I_D} \sum_j T_{ij}(\mathbb{1} - T)^{-1}_{j\star} \lesssim e^{-e^Q}$, with $\mathcal{O}(10) \lesssim Q \lesssim \mathcal{O}(100)$. Combined with (72), this implies

$$P(D|\mathcal{H}_{\text{late}}) \lesssim e^{-e^Q}. \quad (73)$$

Therefore, in all likelihood, $P(D|\mathcal{H}_{\text{late}})$ is doubly exponentially suppressed.

C. Evidence for uniform-time hypothesis

We next turn our attention to the Bayesian evidence (69a) for the uniform-time hypothesis. For concreteness, we first assume the uniform prior (40) over ancestral vacua, and then discuss how the results generalize to any prior having nonzero support on high-energy/low-entropy initial vacua.

With the uniform prior $p_\alpha = 1/N_{\text{dS}}$, the Bayesian evidence (69a) becomes

$$P(D|\mathcal{H}_{\text{uni}}) = \frac{\sum_{i \subset I_D}(\sum_\alpha (\mathbb{1} - T)^{-1}_{i\alpha} - 1)}{\sum_{k,\beta}(\mathbb{1} - T)^{-1}_{k\beta}}, \quad (74)$$

where we have used the trivial matrix identity $T(\mathbb{1} - T)^{-1} = (\mathbb{1} - T)^{-1} - \mathbb{1}$. It is convenient to define

$$N_i \equiv \sum_\alpha (\mathbb{1} - T)^{-1}_{i\alpha} \quad (75)$$

as the effective number of vacua that can reach i. Note that $N_i > 1$, since a vacuum can trivially reach itself. In general, $N_i - 1$ can be doubly exponentially small (if no unsuppressed chain reaches i), exponentially small (if some unsuppressed chains reach i, but all involve at least one step with exponentially small T_{ij}), or even order unity or larger (if some unsuppressed chains reaching i solely consist of steps with $T_{ij} \simeq 1$).

In any case, the Bayesian evidence reduces to

$$P(D|\mathcal{H}_{\text{uni}}) = \frac{\sum_{i \subset I_D}(N_i - 1)}{\sum_k N_k}. \quad (76)$$

Since $N_k > 1$, the denominator satisfies $\sum_k N_k > N_{\text{dS}}$. However, since the vast majority of vacua are high-energy vacua, and as such can be reached by few ancestors with non-negligible branching ratio, we expect that $\sum_k N_k$ does not greatly exceed N_{dS}. That is,

$$\sum_k N_k \gtrsim N_{\text{dS}}. \quad (77)$$

Meanwhile, the numerator counts the total effective number of vacua that can reach vacua compatible with our data, other than themselves. Since vacua in I_D all have a tiny CC of $\simeq 10^{-120} M^4_{\text{Pl}}$, it stands to reason that a significant fraction can be accessed by many vacua via unsuppressed chains. That is,

$$\sum_{i \subset I_D}(N_i - 1) \sim N_D, \quad (78)$$

where $N_D = \sum_{i \subset I_D} N_i$ is the total number of dS vacua compatible with our data. The prefactor in (78) is at worse exponentially small, but it could also be $\gg 1$. (This would be the case is a significant fraction of vacua in I_D can be accessed by many ancestors through sequences of transitions with $T_{ij} \simeq 1$.)

It follows that

$$P(D|\mathcal{H}_{\text{uni}}) \sim \frac{N_D}{N_{\text{dS}}}. \qquad (79)$$

Aside from a prefactor which is at worse exponentially small, $P(D|\mathcal{H}_{\text{uni}})$ is given by the fraction of all dS vacua compatible with our data. This fraction is of course unknown, but it is reasonable to expect that it is exponentially small, not doubly exponentially small. For instance, the worst tuning in the Standard Model is the CC. If the underlying CC distribution is approximately uniform, then a fraction of $10^{-120} \simeq e^{-276}$ of all dS vacua would have a CC consistent with the observed value. These considerations lead us to conclude that $P(D|\mathcal{H}_{\text{uni}})$, while likely exponentially small, is not doubly exponentially suppressed like (73).

A similar argument applies to more general prior distributions p_α over ancestral vacua. To simplify the discussion, we work to leading order in the downward approximation. Using (57), the Bayesian evidence becomes

$$P(D|\mathcal{H}_{\text{uni}}) = \frac{\sum_{i \subset I_D} \mathcal{P}_i}{1 + \sum_k \mathcal{P}_k}. \qquad (80)$$

In the special case $p_\alpha = 1/N_{\text{dS}}$, we have $\mathcal{P}_i \simeq \frac{N_i - 1}{N_{\text{dS}}}$, i.e., $\mathcal{P}_i$ measures the effective fraction of other vacua that can reach i. The argument proceeds along similar lines as the uniform case discussed above. Since most dS vacua in the landscape are high-energy vacua, and as such can be reached through downward transitions from a limited set of higher-energy vacua, we have $\sum_k \mathcal{P}_k \ll 1$. Meanwhile, since vacua compatible with our data all have tiny vacuum energy, we expect that a significant fraction are accessible via a sequence of downward transitions.[21] If so, then we are once again led to conclude that $P(D|\mathcal{H}_{\text{uni}})$ is likely exponentially small, but not doubly-exponentially suppressed.

The above analysis leads us to infer that the Bayes' factor $\frac{P(D|\mathcal{H}_{\text{late}})}{P(D|\mathcal{H}_{\text{uni}})}$ in (37) is doubly-exponentially small. Assuming comparable priors for the two hypotheses, $P(\mathcal{H}_{\text{late}}) \sim P(\mathcal{H}_{\text{uni}})$, we are led to conclude that posterior odds overwhelmingly favor the uniform-time hypothesis.

---

[21]This is where the prior distribution p_α over ancestral comes in. If, for some reason, the initial conditions strongly favor low-energy vacua, then accessing vacua compatible with our data may require subdominant upward jumps, in which case the numerator in (80) would be doubly exponentially suppressed.

D. Loopholes

The above argument is not ironclad. We can think of a few loopholes that would invalidate our conclusions:

 (i) If the dominant vacuum $\star$ is, miraculously, compatible with our data, this would imply $P(D|\mathcal{H}_{\text{late}}) \simeq 1$. This would not only invalidate our conclusions, it would, more importantly, resurrect the old dream of string theory predicting a unique vacuum. Although in this case, the hierarchy between time of existence (which is longer than the mixing time) and the inhabitation time (lifetime of our universe) may potentially be a problem.

 (ii) A more plausible loophole is that there exists a vacuum compatible with our data that lies along an unsuppressed chain starting from $\star$. This would boost $P(D|\mathcal{H}_{\text{late}})$ to be only exponentially suppressed, and thus comparable to $P(D|\mathcal{H}_{\text{uni}})$.

 (iii) Another plausible loophole is that no vacuum compatible with our data can be reached via a sequence of downward transitions, nor more generally via an unsuppressed chain. This may be because of initial conditions, as mentioned earlier, or because vacua like ours are buried in regions of the landscape that are difficult to access. In either case $P(D|\mathcal{H}_{\text{uni}})$ would be doubly exponentially suppressed, and thus comparable to $P(D|\mathcal{H}_{\text{late}})$.

To determine whether any one of these loopholes is valid would require a more detailed and extensive knowledge of the string landscape. Our conclusion rests on current expectations about the string landscape, to the best of our understanding. We believe that, presented with all the information currently at hand about the landscape and eternal inflation, a bookmaker would set the odds overwhelmingly in favor of the uniform-time hypothesis. Probabilities, after all, are nothing but betting odds [106].

VII. INFERRING THE TIME OF EXISTENCE

The last Bayesian operation of interest is parameter inference, discussed in Sec. III B. We will be primarily interested in using our data D to infer the time of existence t. For concreteness we focus on the uniform-time prior, since it is overwhelmingly favored by the data, but our analysis can be easily be generalized to any other prior of interest.

Using (30) with quasiuniform prior (44), the posterior probability (34) for the time of nucleation becomes

$$P(t|D) = \frac{\sum_{i \subset I_D, j} \kappa_{ij} \sum_\alpha p_\alpha (e^{Mt})_{j\alpha} e^{-\epsilon t}}{\sum_{k \subset I_D, \ell} T_{k\ell} \sum_\beta (\mathbb{1} - T)_{\ell\beta}^{-1} p_\beta}, \qquad (81)$$

where in simplifying the denominator we have used (12) and sent $\epsilon \to 0$. We can compute various moments of this distribution. For instance, the average time of nucleation (35) is

$$\langle t \rangle = \int_0^\infty \mathrm{d}t\, t P(t|D)$$

$$= \lim_{\epsilon \to 0} \frac{\sum_{i \subset I_D, j} \kappa_{ij} \sum_\alpha p_\alpha \int_0^\infty \mathrm{d}t\, t (e^{Mt})_{j\alpha} e^{-\epsilon t}}{\sum_{k \subset I_D, \ell} T_{k\ell} \sum_\beta (\mathbb{1} - T)^{-1}_{\ell\beta} p_\beta}$$

$$= -\frac{\sum_{i \subset I_D, j} \kappa_{ij} \sum_\alpha p_\alpha \times \frac{\mathrm{d}}{\mathrm{d}\epsilon} (\epsilon - M)^{-1}_{j\alpha}|_{\epsilon=0}}{\sum_{k \subset I_D, \ell} T_{k\ell} \sum_\beta (\mathbb{1} - T)^{-1}_{\ell\beta} p_\beta}. \quad (82)$$

Using (12) it is straightforward to obtain

$$\langle t \rangle = \frac{\sum_{i \subset I_D, j} T_{ij} \sum_m (\mathbb{1} - T)^{-1}_{jm} \kappa_m^{-1} \sum_\alpha (\mathbb{1} - T)^{-1}_{m\alpha} p_\alpha}{\sum_{k \subset I_D, \ell} T_{k\ell} \sum_\beta (\mathbb{1} - T)^{-1}_{\ell\beta} p_\beta}. \quad (83)$$

In the downward approximation, together with the assumption that transition rates from most low-lying dS vacua to AdS terminals are always larger than up-tunneling rates, this expression simplifies to

$$\langle t \rangle \simeq \frac{\sum_{i \subset I_D} \langle t_i \rangle}{\sum_{j \subset I_D} \mathcal{P}_j}, \quad (84)$$

where we have defined an average unconditional MFPT to i as

$$\langle t_i \rangle = \sum_\alpha \langle t_{i\alpha} \rangle p_\alpha = \sum_{\ell \neq i, \alpha} \kappa_{i\ell} (M^{(i)})^{-2}_{\ell j} p_\alpha. \quad (85)$$

The technical details of the derivation is given in Appendix B. This result is intuitively clear. The average time to reach vacua i compatible with our data is the average of the characteristic time $\langle t_j \rangle$ to reach a parent vacuum.

The arguments of Sec. VI, which lead us to conclude that the uniform-time hypothesis is overwhelmingly favored, relied on the assumption that there is at least one vacuum compatible with our data which can be accessed through a sequence of downward transitions. Equation (84) also hinges on that assumption. Since downward transition rates are exponentially faster than upward rates [see (70)], the downward MFPT is correspondingly exponentially shorter than the dS recurrence time of low-energy vacuum, and thus certainly exponentially shorter than the mixing time for the landscape.

Therefore, if vacua compatible with our data can be accessed via downward transitions, we most likely exist at early times in the unfolding of the multiverse, well before the mixing time for the landscape. This confirms the assumptions underlying the early-time approach to eternal inflation [47–50]. This also circumvents the issue of Boltzmann brains [55–59], which are produced on exponentially longer timescales. On the other hand, the time of existence for the late time hypothesis is much longer than mixing time and recurrence time of certain vacua. The rate of production of freak observers put stringent constraint on the landscape.

VIII. ETERNAL VS NONETERNAL INFLATION

A useful consistency test for eternal inflationary measures was proposed by Linde and Noorbala [107]. The idea is to consider a scalar potential whose shape can be varied smoothly, such that inflation goes from being eternal to non-eternal in the process. One can then check whether the predictions of a given measure vary continuously through the eternal/noneternal transition.

Following [107], we consider the two toy potentials shown in Fig. 2. The two potentials only differ by their behavior near $\phi = 0$. In both cases, the initial conditions are chosen such that the field starts at $\phi = 0$.

(i) The solid curve displays a local minimum, assumed sufficiently long-lived for eternal inflation to ensue. The tunneling rates across the left and right barriers are assumed equal, by symmetry.

(ii) The dashed curve shows a local peak, assumed sufficiently steep to avoid eternal inflation. The peak is symmetric, which ensures that the field rolls to one or the other with equal probability.

Subsequently the field undergoes a long period of (noneternal) slow-roll inflation on either side of the origin, with different number of e-folds N_+ and N_-, and ends up in the ϕ_+ or ϕ_- vacua. These two vacua have identical physical constants. Both are AdS, with identical tiny negative vacuum energy, such that the regions crunch in a finite time. Furthermore, in the noneternal case, the spatial geometry is assumed compact, either because of positive curvature ($k > 0$) or topological identifications (e.g., flat 3-torus). These assumptions ensure that the noneternal case yields unambiguous, finite answers.

For concreteness, consider the total number $\mathcal{N}^{\mathrm{gal}}_\pm$ of galaxies that form in each vacua. Assuming that the probability to form a certain number of galaxies per comoving volume is the same in each vacua, then the ratio $\mathcal{N}^{\mathrm{gal}}_+ / \mathcal{N}^{\mathrm{gal}}_-$ is simply given by the volume growth during slow-roll inflation [107],

$$\frac{\mathcal{N}^{\mathrm{gal}}_+}{\mathcal{N}^{\mathrm{gal}}_-} = e^{3(N_+ - N_-)}. \quad (86)$$
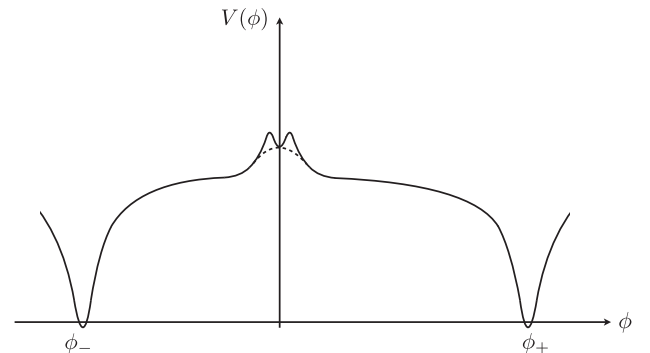

FIG. 2. Eternal vs noneternal test proposed by Linde.

Given the symmetries of the problem, the answer should be the same in the eternally inflating case. This is the consistency check proposed by Linde and Noorbala [107]. The authors found that global measures with proper time [13] or e-folding time [12–14,20,21] cutoff both failed the test. The causal diamond measure [27] can fail the test, depending on the number of e-folds. The stationary measure [24,25] passes the test.

In our case, the full probability distribution of some observable in pocket universes $+$ and $-$, such as the number of galaxies, follows the derivation given in Sec. VI of GSVW [18],

$$P_\pm \sim p_\pm H_{\mathrm{SR}}^3 e^{3N_\pm},\tag{87}$$

where $p_\pm$ is the probability to tunnel to the $\pm$ region, respectively, and H_{SR} is the Hubble constant during slow-roll inflation. Equation (87) should also be multiplied by the probability to form a certain number of galaxies per comoving volume, but since this probability is assumed to be the same in each vacuum, as assumed in the noneternal case, this factor is inconsequential.

Since the branching ratios to the $+$ and $-$ sides are the same, unsurprisingly our measure gives identical tunneling probabilities $p_+ = p_-$. The number of galaxies that form in each vacuum is then given by

$$\frac{\mathcal{N}_+^{\mathrm{gal}}}{\mathcal{N}_-^{\mathrm{gal}}} = \frac{P_+}{P_-} = e^{3(N_+ - N_-)}.\tag{88}$$

Thus our measure passes the consistency test of [107]. (So does the GSVW measure, incidentally.) The same conclusion would be reached if eternal inflation were stochastic, i.e., by considering a flat plateau around $\phi = 0$ instead of a metastable state.

IX. CONCLUSION

Understanding our place within this multiverse is ultimately necessary to make any predictions about physical observables in our Universe. Attempts to define probabilities (or measure) usually rely on limiting frequency distributions. This is perhaps natural, since the infinite ensemble necessary to define frequencies is actually realized in the multiverse. Unfortunately, this approach has failed to yield an unambiguous answer.

In this paper we instead applied Bayesian reasoning to define probabilities. The advantage of this approach is first and foremost a practical one. All attempts to define a semiclassical measure rely on certain assumptions. The Bayesian framework naturally compels one to make all assumptions explicit through prior information. Our approach has been strongly influenced by Jaynes' view of probability theory as an extension of classical logic [34]. Probabilities, in this viewpoint, amount to reasonable expectations [35] drawn from limited information. Our

treatment is also inspired by Caves' elegant resolution [60] to the Doomsday argument [61–63]. Our approach does not rely on *ad hoc* geometric constructions, nor are we counting anything.

The natural starting point to define probabilities is the master equation governing vacuum dynamics, obtained after suitable coarse-graining. Remarkably, this equation describes a linear Markov process, free of the conceptual pitfalls of eternal inflation. This is not a mathematical artifice—physically, the master equation describes the random walk on the network of vacua that "we" have performed since the onset of eternal inflation. The occupational probabilities $f_I(t)$'s are normalized and time-reparametrization invariant, and thus offer well-defined probabilities to occupy different vacua at time t.

Our probabilities require two pieces of prior information; a prior probability density $\rho(t)$ for the time of nucleation; and a prior probability p_α for the ancestral vacuum. Both pertain to initial conditions. We know that eternal inflation started a finite time in our past, but we do not know when, and it started in some particular vacuum, but we do not know which one. Different approaches to the measure problem amount to different choices for these priors.

Consistency requires that our priors reflect all information at hand, but should otherwise be minimally informative. For ancestral vacua, we advocated the uniform prior $p_\alpha = 1/N_{\mathrm{dS}}$ as a conservative choice, though our conclusions are fairly insensitive to this choice. What matters is that the initial conditions have support over high-energy/low-entropy vacua. For the time of nucleation, we argued that a quasiuniform prior is a natural choice, consistent with the time-translational invariance of the master equation. It represents the minimally informative prior. The resulting predictive probability distribution matches the prior probabilities of [27] and is closely related to the "comoving" probabilities discussed in [18,45].

We also considered a volume-weighted $\rho(t)$, which amounts to weighing probabilities for different vacua by their physical volume. This prior peaks at late times, and as such is maximally informative. Interestingly, because volume weighing is implemented as a prior, the resulting probabilities do not suffer from the usual sensitivity to the choice of time variable and the associated paradoxes. Instead, the predictive distribution is time-reparametrization invariant, and agrees with the GSVW measure [18].

The Bayesian framework allowed us to compare the plausibility of the uniform-time and volume-weighted hypotheses to explain our data by computing the Bayesian evidence for each. We argued, under general and plausible assumptions, that posterior odds overwhelmingly favor the uniform-time hypothesis. The argument relies on assumptions about the dominant vacuum that have been made in previous studies of the landscape, e.g., [87]. There are some caveats, of course, and we tried to enunciate them carefully in Sec. VI D.

We believe that the uniform-time measure is the correct objective approach to probabilistic reasoning in the multiverse. The assumed priors are the least informative and consistently reflect our current state of knowledge about how/when eternal inflation started. The prior predictive distribution (54) is very intuitive, to the extent that one could have written down the answer without doing any work. It favors vacua that are easily accessed under the random walk on the landscape.

Despite making the least informative prior assumptions, the probability distribution (54) is surprisingly predictive. It favors vacua lying within deep funnels [47,48,50], wherein they can be accessed through a sequence of downward transitions from a large basin of parent vacua. This is akin to the folding funnels of proteins [54]. As argued in Sec. VII, it also predicts that we exist at times much earlier than the mixing time for the landscape, confirming the intuition behind the early-time approach to eternal inflation [50]. This implies, incidentally, that we are "normal" observers as opposed to Boltzmann brains [55–59], which are produced on exponentially longer timescales.

The analogy with natural selection and protein landscapes is quite apt. The prior predictive probabilities (54) define a fitness function on the string landscape, analogous to the fitness landscape over the space of protein sequences [108]. Sequence space is high dimensional, much like the string landscape. And it is believed that, through evolution, all of sequence space for proteins and genomes has been explored by biology on Earth [109], just like eternal inflation is efficient at populating the entire landscape. Naturally occurring proteins occupy a small region of the fitness landscape, characterized by a large basin of "neutral mutations" [110]. These large basins with high fitness, known as "superfunnels" [111–113], are strikingly similar to the large funnels favored by our measure.

The probabilistic framework developed here opens up many avenues of inquiry. We mention in closing two particularly interesting directions:

(i) By modeling landscape regions as random networks, we will show in a forthcoming paper [64] that the uniform-time probabilities (54) favor regions that are close to the directed percolation phase transition [114]. Thus landscape dynamics belong to the universality class of directed percolation—the paradigmatic nonequilibrium critical phenomenon. As usual, the predictive power of criticality lies in scale invariant observables characterized by critical exponents. We will argue that the probability distribution for the CC is a power law, that favors a naturally small and positive vacuum energy. Tantalizingly, this hints at a deep connection between nonequilibrium critical phenomena on the landscape and the near-criticality of our Universe.

(ii) The connection with protein folding funnels deserves further exploration. The problem of search

optimization on complex, high-dimensional energy landscapes has already been solved by naturally-occurring proteins. A fascinating result in protein folding is that conformation networks share many properties of real-world networks [115]; their degree distribution is scale free, they enjoy the small-world property, and they are hierarchical. It will be interesting to study the implications of similar properties holding in regions of the string landscape.

APPENDIX A: DISCRETE-TIME DERIVATION

In this appendix we present an alternative derivation of the prior predictive distribution (54) using a discrete-time variable. The latter is defined such that each dS vacuum undergoes a transition (to another dS or to a terminal) at every time step $n = 0, 1, 2, \ldots$ [30].

The probability $X_I(n)$ to occupy vacuum I at time n satisfies the master equation

$$X_I(n+1) = \sum_J \mathbb{T}_{IJ} X_J(n), \qquad (A1)$$

where the full branching matrix $\mathbb{T}_{IJ}$ was defined in (14). Its sum rule (15) ensures that probability is conserved

$$\sum_I X_I(n) = 1. \qquad (A2)$$

The solution to (A1) is given by

$$X_I(n) = \sum_\alpha (\mathbb{T}^n)_{I\alpha} p_\alpha, \qquad (A3)$$

where $p_\alpha \equiv X_\alpha(0)$ denotes as before the initial probability over ancestral vacua. For dS vacua, in particular, (A3) becomes

$$X_j(n) = \sum_\alpha (T^n)_{I\alpha} p_\alpha. \qquad (A4)$$

The joint probability $P(I, n)$ to inhabit vacuum I after a nucleation event at the nth time step is

$$P(I, n) = \mathcal{N} \sum_j \mathbb{T}_{Ij} X_j(n-1) P_{\text{prior}}(n), \qquad (A5)$$

where $\mathcal{N}$ is a normalization constant. This is the discrete-time analog of (30), after marginalizing over α, with the prior $P_{\text{prior}}(n)$ on the discrete time of existence playing the role of $\rho(t)$. The prior predictive distribution is then given by

$$\begin{aligned}
P(I) &= \sum_{n=0}^\infty P(I, n) P_{\text{prior}}(n) \\
&= \frac{\sum_j \mathbb{T}_{Ij} \sum_{n=1}^\infty X_j(n-1) P_{\text{prior}}(n)}{\sum_k \sum_{m=1}^\infty X_k(m-1) P_{\text{prior}}(m)},
\end{aligned} \qquad (A6)$$

where we have fixed the normalization constant to ensure that $\sum_I P(I) = 1$. Note that the sum over n starts at $n = 1$ since, by assumption, there has been at least one nucleation event along our past worldline, i.e., the one giving rise to our bubble.

An improper uniform prior in this case is straightforward and given by $P_{\text{prior}}(n) = $ constant for all $n \geq 1$. This reflects complete prior ignorance about how many nucleation events took place along our past worldline. As in the continuous case, the uniform-time prior must be regularized. A simple prescription is to introduce a large cutoff M,

$$P_{\text{prior}}(n) = \begin{cases} \frac{1}{M} & n = 1, \ldots, M \\ 0 & \text{otherwise.} \end{cases} \qquad (A7)$$

Substituting this prior, together with (A4), the prior predictive distribution (A6) becomes

$$\begin{aligned}
P(I|\mathcal{H}_{\text{uni}}) &= \frac{\sum_{j,\alpha} \mathbb{T}_{Ij} \sum_{n=1}^M (T^{n-1})_{j\alpha} p_\alpha}{\sum_{k,\beta} \sum_{m=1}^M (T^{m-1})_{k\beta} p_\beta} \\
&= \frac{\sum_{j,\alpha} \mathbb{T}_{Ij} \sum_{n=0}^{M-1} (T^n)_{j\alpha} p_\alpha}{\sum_{k,\beta} \sum_{m=0}^{M-1} (T^m)_{k\beta} p_\beta}.
\end{aligned} \qquad (A8)$$

At this point the cutoff can be removed by sending $M \to \infty$, with the result

$$P(I|\mathcal{H}_{\text{uni}}) = \frac{\sum_j \mathbb{T}_{Ij} \sum_\alpha (\mathbb{1} - T)^{-1}_{j\alpha} p_\alpha}{\sum_{k,\beta} (\mathbb{1} - T)^{-1}_{k\beta} p_\beta}. \qquad (A9)$$

This agrees precisely with the continuous-time answer (54). It reaffirms that our probabilities are time-reparametrization invariant.

APPENDIX B: EXACT RELATION BETWEEN TIME OF EXISTENCE AND FIRST PASSAGE TIME

We provide an exact relation between the average time of existence (83) and mean first-passage time in this section. The time of existence (83) derived in Sec. VII can be written in terms of the transition matrix M as

$$\langle t \rangle = \frac{1}{\mathcal{N}} \sum_{i \subset I_D, j, \alpha} \kappa_{ij} (M^{-2})_{j\alpha} p_\alpha, \qquad (B1)$$

where $\mathcal{N} = -\sum_{i \subset I_D, j, \alpha} \kappa_{ij}(M^{-1})_{j\alpha} p_\alpha$ is the normalization factor of the conditional probability $P(t|D)$. Recall that the unconditional MFPT $\langle t_{ij} \rangle$ and the conditional MFPT $\langle t_{ij} \rangle_c$ from j to i are given by [50]

$$\langle t_{ij} \rangle = \sum_{l \neq i} \kappa_{il} (M^{(i)})^{-2}_{lj}; \qquad \langle t_{ij} \rangle_c = \frac{\sum_{l \neq i} \kappa_{il} (M^{(i)})^{-2}_{lj}}{\mathcal{P}_{ij}}, \qquad (B2)$$

where $M^{(i)}$ is the transition matrix with the ith column and row deleted, and $\mathcal{P}_{ij} = -\sum_{k \neq l} \kappa_{il}(M^{(i)})^{-1}_{lj}$ is the ever-hitting probability.

To find a relation between $\langle t \rangle$ and $\langle t \rangle_{ij}$, we can use the relation between M^{-1} and $(M^{(i)})^{-1}$,

$$\begin{aligned}
M^{-1} &= \begin{bmatrix} -\kappa_i & \bar{\kappa}_i \\ \vec{\kappa}_i & M^{(i)} \end{bmatrix}^{-1} = \begin{bmatrix} -Q_i & \bar{\kappa}_i W_i Q_i \\ Q_i W_i \vec{\kappa}_i & (\mathbb{1} - Q_i W_i \vec{\kappa}_i \bar{\kappa}_i) W_i \end{bmatrix}, \\
Q_i &= (\kappa_i + \bar{\kappa}_i W \vec{\kappa}_i)^{-1} = (\kappa_i - \bar{\mathcal{P}}_i \vec{\kappa}_i)^{-1}, \qquad W_i = (M^{(i)})^{-1},
\end{aligned} \qquad (B3)$$

where $[\vec{\kappa}_i]_j = \kappa_{ji}$, $[\bar{\kappa}_i]_j = \kappa_{ij}$ and $[\bar{\mathcal{P}}_i]_j = \mathcal{P}_{ij}$ for $j \neq i$. Note that Q_i has an interesting meaning when expanded out in the following way:

$$Q_i = \kappa_i^{-1} + \kappa_i^{-2} \bar{\mathcal{P}} \vec{\kappa}_i + \kappa_i^{-3} (\bar{\mathcal{P}} \vec{\kappa}_i)^2 + \ldots, \qquad (B4)$$

the nth term $\kappa_i^{-n-1}(\bar{\mathcal{P}} \vec{\kappa}_i)^n$ can be interpreted as the time needed for the trip $i \to (\text{any } j \to i \to)^n \to a$. To compute $\langle t \rangle$ we need

$$M^{-2} = \begin{bmatrix} Q^2(1 + \bar{\kappa} W^2 \vec{\kappa}) & -\bar{\kappa} W Q^2 + \bar{\kappa} W^2 Q - \bar{\kappa} W^2 \vec{\kappa} \bar{\kappa} W Q^2 \\ W \vec{\kappa} Q^2 + W^2 \vec{\kappa} Q - W \vec{\kappa} \bar{\kappa} W^2 \vec{\kappa} Q^2 & Q^2 W \vec{\kappa} \bar{\kappa} W + W^2 - W^2 \vec{\kappa} \bar{\kappa} W Q - W \vec{\kappa} \bar{\kappa} W^2 Q + W \vec{\kappa} \bar{\kappa} W^2 \vec{\kappa} \bar{\kappa} W Q^2 \end{bmatrix}. \qquad (B5)$$

Note that we have omitted the subscript i to avoid clustering. Therefore, the timescale $t_i = \sum_{j,\alpha} \kappa_{ij}(M^{-2})_{j\alpha}p_\alpha$ can be written as

$$t_i = \sum_{j,\alpha} \kappa_{ij}(M^{-2})_{j\alpha}p_\alpha$$

$$= (\bar{t}_i + Q_i^2\bar{\vec{\mathcal{P}}}_i\vec{\kappa}_i\bar{\vec{\mathcal{P}}}_i + Q_i\bar{t}_i\vec{\kappa}_i\bar{\vec{\mathcal{P}}}_i + Q_i\bar{\vec{\mathcal{P}}}_i\vec{\kappa}_i\bar{t}_i$$
$$+ Q_i^2\bar{\vec{\mathcal{P}}}_i\vec{\kappa}_i\bar{t}_i\vec{\kappa}_i\bar{\vec{\mathcal{P}}}_i) \cdot p, \tag{B6}$$

where $[\bar{t}_i]_j = t_{ij}$ is a row vector of MFPTs from j to i. Finally the time of existence $\langle t \rangle$ is just the sum over all i that are compatible with data,

$$\langle t \rangle = \frac{1}{\mathcal{N}}\sum_{i \subset I_D} t_i. \tag{B7}$$

Also note that the normalization factor $\mathcal{N}$ can be written in terms of ever-hitting probabilities,

$$\mathcal{N} = -\sum_{i \subset I_D, j, \alpha} \kappa_{ij}(M^{-1})_{j\alpha}p_\alpha$$
$$= \sum_{i \subset I_D}(\bar{\vec{\mathcal{P}}}_i + Q_i\bar{\vec{\mathcal{P}}}_i\vec{\kappa}_i\bar{\vec{\mathcal{P}}}_i) \cdot p. \tag{B8}$$

In the downward approximation, either

$$[\vec{\kappa}_i]_j = \kappa_{ji} = 0; \qquad [\bar{\vec{\mathcal{P}}}_i]_j \geq 0 \quad (\neq 0 \text{ when there is a directed path from } j \text{ to } i)$$
$$\text{or} \quad [\vec{\kappa}_i]_j = \kappa_{ji} \geq 0; \qquad [\bar{\vec{\mathcal{P}}}_i]_j = 0 \quad (\text{when there are only upwords paths from } j \text{ to } i) \tag{B9}$$

can happen for any node j above or below, therefore the scalar product $\bar{\vec{\mathcal{P}}}_i\vec{\kappa}_i$ goes to zero in the downward approximation. With some effort in analyzing $\bar{t}_i\vec{\kappa}_i$, one finds that it also vanishes in the downward approximation. (This is somewhat counter intuitive as it is an unconditional quantity.) The essential part is that there exist terminals, and the inverse of the matrix is just rational function of its elements. The denominator of $(M^{(i)})^{-2}$ is simply $\det(M^{(i)})^2 = \prod_{\neq i}(-\kappa_j)^2$ in the downward approximation. Given that the transition rate to AdS terminals is faster than

the up-tunneling rate, the denominator would not go to zero in the downward approximation. It is also obvious that the numerator should be of $\mathcal{O}$ (up-tunneling rate) for up-tunneling paths. Therefore, in the downward approximation,

$$\langle t \rangle = \frac{\sum_{i \subset I_D}\langle t_i \rangle}{\sum_{j \subset I_D}\mathcal{P}_j}, \tag{B10}$$

where $\langle t_i \rangle = \bar{t}_i \cdot p$ and $\mathcal{P}_i = \bar{\vec{\mathcal{P}}}_i \cdot p$.

[1] P. J. Steinhardt, Natural inflation, in *Nuffield Workshop on the Very Early Universe* (Cambridge University Press, Cambridge, United Kingdom, 1982), pp. 251–266.

[2] A. Vilenkin, The birth of inflationary universes, Phys. Rev. D **27**, 2848 (1983).

[3] A. D. Linde, Eternal chaotic inflation, Mod. Phys. Lett. A **01**, 81 (1986).

[4] A. D. Linde, Eternally existing self-reproducing chaotic inflationary universe, Phys. Lett. B **175**, 395 (1986).

[5] A. A. Starobinsky, Stochastic de Sitter (inflationary) stage in the early universe, Lect. Notes Phys. **246**, 107 (1986).

[6] R. Bousso and J. Polchinski, Quantization of four form fluxes and dynamical neutralization of the cosmological constant, J. High Energy Phys. 06 (2000) 006.

[7] S. Kachru, R. Kallosh, A. D. Linde, and S. P. Trivedi, De Sitter vacua in string theory, Phys. Rev. D **68**, 046005 (2003).

[8] G. Obied, H. Ooguri, L. Spodyneiko, and C. Vafa, De Sitter space and the swampland, arXiv:1806.08362.

[9] P. Agrawal, G. Obied, P. J. Steinhardt, and C. Vafa, On the cosmological implications of the string swampland, Phys. Lett. B **784**, 271 (2018).

[10] E. Palti, The swampland: Introduction and review, Fortschr. Phys. **67**, 1900037 (2019).

[11] B. Freivogel, Making predictions in the multiverse, Classical Quantum Gravity **28**, 204007 (2011).

[12] A. D. Linde and A. Mezhlumian, Stationary universe, Phys. Lett. B **307**, 25 (1993).

[13] A. D. Linde, D. A. Linde, and A. Mezhlumian, From the big bang theory to the theory of a stationary universe, Phys. Rev. D **49**, 1783 (1994).

[14] J. Garcia-Bellido, A. D. Linde, and D. A. Linde, Fluctuations of the gravitational constant in the inflationary Brans-Dicke cosmology, Phys. Rev. D **50**, 730 (1994).

[15] A. Vilenkin, Predictions from Quantum Cosmology, Phys. Rev. Lett. **74**, 846 (1995).

[16] A. H. Guth, Eternal inflation and its implications, J. Phys. A **40**, 6811 (2007).

[17] J. Garriga and A. Vilenkin, Recycling universe, Phys. Rev. D **57,** 2230 (1998).

[18] J. Garriga, D. Schwartz-Perlov, A. Vilenkin, and S. Winitzki, Probabilities in the inflationary multiverse, J. Cosmol. Astropart. Phys. 01 (2006) 017.

[19] V. Vanchurin and A. Vilenkin, Eternal observers and bubble abundances in the landscape, Phys. Rev. D **74,** 043520 (2006).

[20] R. Bousso, B. Freivogel, and I.-S. Yang, Properties of the scale factor measure, Phys. Rev. D **79,** 063513 (2009).

[21] A. De Simone, A. H. Guth, A. D. Linde, M. Noorbala, M. P. Salem, and A. Vilenkin, Boltzmann brains and the scale-factor cutoff measure of the multiverse, Phys. Rev. D **82,** 063520 (2010).

[22] A. De Simone, A. H. Guth, M. P. Salem, and A. Vilenkin, Predicting the cosmological constant with the scale-factor cutoff measure, Phys. Rev. D **78,** 063520 (2008).

[23] M. Tegmark, What does inflation really predict?, J. Cosmol. Astropart. Phys. 04 (2005) 001.

[24] A. D. Linde, Towards a gauge invariant volume-weighted probability measure for eternal inflation, J. Cosmol. Astropart. Phys. 06 (2007) 017.

[25] A. D. Linde, V. Vanchurin, and S. Winitzki, Stationary measure in the multiverse, J. Cosmol. Astropart. Phys. 01 (2009) 031.

[26] A. Vilenkin and M. Yamada, Four-volume cutoff measure of the multiverse, Phys. Rev. D **101,** 043520 (2020).

[27] R. Bousso, Holographic Probabilities in Eternal Inflation, Phys. Rev. Lett. **97,** 191302 (2006).

[28] R. Bousso, Complementarity in the multiverse, Phys. Rev. D **79,** 123524 (2009).

[29] R. Bousso, B. Freivogel, S. Leichenauer, and V. Rosenhaus, A Geometric Solution to the Coincidence Problem, and the Size of the Landscape as the Origin of Hierarchy, Phys. Rev. Lett. **106,** 101301 (2011).

[30] J. Garriga and A. Vilenkin, Watchers of the multiverse, J. Cosmol. Astropart. Phys. 05 (2013) 037.

[31] Y. Nomura, Physical theories, eternal inflation, and quantum universe, J. High Energy Phys. 11 (2011) 063.

[32] R. Bousso and I.-S. Yang, Global-local duality in eternal inflation, Phys. Rev. D **80,** 124024 (2009).

[33] R. Bousso and D. Mainemer Katz, New local duals in eternal inflation, Phys. Rev. D **86,** 083530 (2012).

[34] E. T. Jaynes, *Probability Theory: The Logic of Science* (Cambridge University Press, Cambridge, UK, 2003).

[35] R. Cox, Probability, frequency and reasonable expectation, Am. J. Phys. **14,** 1 (1946).

[36] J. B. Hartle and M. Srednicki, Are we typical?, Phys. Rev. D **75,** 123523 (2007).

[37] M. Srednicki and J. Hartle, Science in a very large universe, Phys. Rev. D **81,** 123524 (2010).

[38] A. Borde, A. H. Guth, and A. Vilenkin, Inflationary Space-Times are Incompletein Past Directions, Phys. Rev. Lett. **90,** 151301 (2003).

[39] A. Vilenkin, Creation of universes from nothing, Phys. Lett. **117B,** 25 (1982).

[40] J. B. Hartle and S. W. Hawking, Wave function of the universe, Phys. Rev. D **28,** 2960 (1983).

[41] A. D. Linde, Quantum creation of the inflationary universe, Lett. Nuovo Cimento **39,** 401 (1984).

[42] A. D. Linde, Quantum creation of an inflationary universe, Sov. Phys. JETP **60,** 211 (1984).

[43] A. Vilenkin, Quantum creation of universes, Phys. Rev. D **30,** 509 (1984).

[44] A. Vilenkin, Boundary conditions in quantum cosmology, Phys. Rev. D **33,** 3560 (1986).

[45] J. Garriga and A. Vilenkin, A Prescription for probabilities in eternal inflation, Phys. Rev. D **64,** 023507 (2001).

[46] K. Langhoff, C. Murdia, and Y. Nomura, Multiverse in an inverted island, Phys. Rev. D **104,** 086007 (2021).

[47] J. Khoury and O. Parrikar, Search optimization, funnel topography, and dynamical criticality on the string landscape, J. Cosmol. Astropart. Phys. 12 (2019) 014.

[48] J. Khoury, Accessibility measure for eternal inflation: Dynamical criticality and Higgs metastability, J. Cosmol. Astropart. Phys. 06 (2021) 009.

[49] G. Kartvelishvili, J. Khoury, and A. Sharma, The self-organized critical multiverse, J. Cosmol. Astropart. Phys. 02 (2021) 028.

[50] J. Khoury and S. S. C. Wong, Early-time measure in eternal inflation, J. Cosmol. Astropart. Phys. 05 (2022) 031.

[51] J. Khoury and T. Steingasser, Gauge hierarchy from electroweak vacuum metastability, Phys. Rev. D **105,** 055031 (2022).

[52] F. Denef, M. R. Douglas, B. Greene, and C. Zukowski, Computational complexity of the landscape II—Cosmological considerations, Ann. Phys. (Amsterdam) **392,** 93 (2018).

[53] R. Bousso, B. Freivogel, S. Leichenauer, and V. Rosenhaus, Eternal inflation predicts that time will end, Phys. Rev. D **83,** 023525 (2011).

[54] J. D. Bryngelson, J. N. Onuchic, N. D. Socci, and P. G. Wolynes, Funnels, pathways and the energy landscape of protein folding: A synthesis, Proteins-Struct. Funct. Genet. **21,** 167 (1995).

[55] A. Albrecht, Cosmic inflation and the arrow of time, arXiv:astro-ph/0210527.

[56] L. Dyson, M. Kleban, and L. Susskind, Disturbing implications of a cosmological constant, J. High Energy Phys. 10 (2002) 011.

[57] A. Albrecht and L. Sorbo, Can the universe afford inflation?, Phys. Rev. D **70,** 063528 (2004).

[58] D. N. Page, The lifetime of the universe, J. Korean Phys. Soc. **49,** 711 (2006).

[59] D. N. Page, Is our universe likely to decay within 20 billion years?, Phys. Rev. D **78,** 063535 (2008).

[60] C. M. Caves, Predicting future duration from present age: A critical assessment, Contemp. Phys. **41,** 143 (2000).

[61] B. Carter, The anthropic principle and its implications for biological evolution, Phil. Trans. R. Soc. A **310,** 347 (1983).

[62] J. Leslie, Risking the world's end, Bull. Can. Nucl. Soc. 10 (1989).

[63] H. Nielsen, Random dynamics and relations between the number of fermion generations and the fine structure constants, Acta Phys. Pol. B **20,** 427 (1989), https://inspirehep.net/literature/276878.

[64] J. Khoury and S. S. C. Wong, Directed percolation phase transition in eternal inflation (to be published).

[65] G. F. Giudice, M. McCullough, and T. You, Self-organised localisation, J. High Energy Phys. 10 (2021) 093.

[66] B. Friedrich, A. Hebecker, M. Salmhofer, J. C. Strauss, and J. Walcher, A local Wheeler-DeWitt measure for the string landscape, Nucl. Phys. **B992,** 116230 (2023).

[67] A. Aguirre and M. Tegmark, Multiple universes, cosmic coincidences, and other dark matters, J. Cosmol. Astropart. Phys. 01 (2005) 003.

[68] J. B. Hartle, Anthropic reasoning and quantum cosmology, AIP Conf. Proc. **743,** 298 (2004).

[69] S. Weinberg, Anthropic Bound on the Cosmological Constant, Phys. Rev. Lett. **59,** 2607 (1987).

[70] G. D. Starkman and R. Trotta, Why Anthropic Reasoning Cannot Predict Lambda, Phys. Rev. Lett. **97,** 201301 (2006).

[71] S. R. Coleman, The fate of the false vacuum. 1. Semiclassical theory, Phys. Rev. D **15,** 2929 (1977); **16,** 1248(E) (1977).

[72] J. Callan, Curtis G., and S. R. Coleman, The fate of the false vacuum. 2. First quantum corrections, Phys. Rev. D **16,** 1762 (1977).

[73] S. R. Coleman and F. De Luccia, Gravitational effects on and of vacuum decay, Phys. Rev. D **21,** 3305 (1980).

[74] S. W. Hawking and I. G. Moss, Supercooled phase transitions in the very early universe, Phys. Lett. **110B,** 35 (1982).

[75] J. D. Brown and C. Teitelboim, Dynamical neutralization of the cosmological constant, Phys. Lett. B **195,** 177 (1987).

[76] M. P. Salem, Multiverse rate equation including bubble collisions, Phys. Rev. D **87,** 063501 (2013).

[77] A. R. Brown and A. Dahlen, Populating the Whole Landscape, Phys. Rev. Lett. **107,** 171301 (2011).

[78] K.-M. Lee and E. J. Weinberg, Decay of the true vacuum in curved space-time, Phys. Rev. D **36,** 1088 (1987).

[79] E. Farhi, A. H. Guth, and J. Guven, Is it possible to create a universe in the laboratory by quantum tunneling?, Nucl. Phys. **B339,** 417 (1990).

[80] W. Fischler, D. Morgan, and J. Polchinski, Quantum nucleation of false vacuum bubbles, Phys. Rev. D **41,** 2638 (1990).

[81] W. Fischler, D. Morgan, and J. Polchinski, Quantization of false vacuum bubbles: A Hamiltonian treatment of gravitational tunneling, Phys. Rev. D **42,** 4042 (1990).

[82] S. P. De Alwis, F. Muia, V. Pasquarella, and F. Quevedo, Quantum transitions between Minkowski and de Sitter spacetimes, Fortschr. Phys. **68,** 2000069 (2020).

[83] Z. Fu and D. Marolf, Bag-of-gold spacetimes, Euclidean wormholes, and inflation from domain walls in AdS/CFT, J. High Energy Phys. 11 (2019) 040.

[84] M. Mirbabayi, Uptunneling to de Sitter, J. High Energy Phys. 09 (2020) 070.

[85] K. D. Olum, P. Upadhyay, and A. Vilenkin, Black holes and uptunneling suppress Boltzmann brains, Phys. Rev. D **104,** 023528 (2021).

[86] D. Schwartz-Perlov and A. Vilenkin, Probabilities in the Bousso-Polchinski multiverse, J. Cosmol. Astropart. Phys. 06 (2006) 010.

[87] K. D. Olum and D. Schwartz-Perlov, Anthropic prediction in a large toy landscape, J. Cosmol. Astropart. Phys. 10 (2007) 010; 10 (2019) E02.

[88] S. Redner, A Guide to First-Passage Processes (Cambridge University Press, Cambridge, England, 2001).

[89] V. Vennin and A. A. Starobinsky, Correlation functions in stochastic inflation, Eur. Phys. J. C **75,** 413 (2015).

[90] H. Assadullahi, H. Firouzjahi, M. Noorbala, V. Vennin, and D. Wands, Multiple fields in stochastic inflation, J. Cosmol. Astropart. Phys. 06 (2016) 043.

[91] V. Vennin, H. Assadullahi, H. Firouzjahi, M. Noorbala, and D. Wands, Critical Number of Fields in Stochastic Inflation, Phys. Rev. Lett. **118,** 031301 (2017).

[92] M. Noorbala, V. Vennin, H. Assadullahi, H. Firouzjahi, and D. Wands, Tunneling in stochastic inflation, J. Cosmol. Astropart. Phys. 09 (2018) 032.

[93] A. Vilenkin, Quantum cosmology and the initial state of the universe, Phys. Rev. D **37,** 888 (1988).

[94] E. Silverstein, (A)dS backgrounds from asymmetric orientifolds, Clay Mat. Proc. **1,** 179 (2002), arXiv:hep-th/0106209.

[95] A. Maloney, E. Silverstein, and A. Strominger, De Sitter space in noncritical string theory, in Workshop on Conference on the Future of Theoretical Physics and Cosmology in Honor of Steven Hawking's 60th Birthday (2002), pp. 570–591, arXiv:hep-th/0205316.

[96] G. B. De Luca, E. Silverstein, and G. Torroba, Hyperbolic compactification of M-theory and de Sitter quantum gravity, SciPost Phys. **12,** 083 (2022).

[97] D. H. Coule and J. Martin, Quantum cosmology and open universes, Phys. Rev. D **61,** 063501 (2000).

[98] A. D. Linde, Creation of a compact topologically nontrivial inflationary universe, J. Cosmol. Astropart. Phys. 10 (2004) 004.

[99] N. Bostrom, Anthropic Bias: Observation Selection Effects in Science and Philosophy (Routledge, New York, NY, 2002).

[100] R. M. Neal, Puzzles of anthropic reasoning resolved using full non-indexical conditioning, arXiv:math/0608592.

[101] M. Aryal and A. Vilenkin, The fractal dimension of inflationary universe, Phys. Lett. B **199,** 351 (1987).

[102] L. D. F. Costa, F. A. Rodrigues, G. Travieso, and P. R. Villas Boas, Characterization of complex networks: A survey of measurements, Adv. Phys. **56,** 167 (2007).

[103] L. Katz, A new status index derived from sociometric analysis, Psychometrika **18,** 39 (1953).

[104] A. Samarakoon, T. J. Sato, T. Chen, G.-W. Chern, J. Yang, I. Klich, R. Sinclair, H. Zhou, and S.-H. Lee, Aging, memory, and nonhierarchical energy landscape of spin jam, Proc. Natl. Acad. Sci. U.S.A. **113,** 11806 (2016).

[105] B. Freivogel and M. Lippert, Evidence for a bound on the lifetime of de Sitter space, J. High Energy Phys. 12 (2008) 096.

[106] B. de Finetti, Theory of Probability: A Critical Introductory Treatment (Wiley, New York, 1970).

[107] A. Linde and M. Noorbala, Measure problem for eternal and non-eternal inflation, J. Cosmol. Astropart. Phys. 09 (2010) 008.

[108] S. Wright, The roles of mutation, inbreeding, crossbreeding, and selection in evolution, in Proceedings of the Sixth International Congress on Genetics (Brooklyn Botanic Garden, New York, 1932), Vol. 1, pp. 355–366.

[109] D. T. F. Dryden, A. R. Thomson, and J. H. White, How much of protein sequence space has been explored by life on Earth?, J. R. Soc. Interface **5,** 953 (2008).

[110] T. Sikosek and H. S. Chan, Biophysics of protein evolution and evolutionary protein biophysics, J. R. Soc. Interface **11**, 20140419 (2014).

[111] J. Noirel and T. Simonson, Neutral evolution of proteins: The superfunnel in sequence space and its relation to mutational robustness, J. Chem. Phys. **129**, 185104 (2008).

[112] F. Morcos, N. P. Schafer, R. R. Cheng, J. N. Onuchic, and P. G. Wolynes, Coevolutionary information, protein folding landscapes, and the thermodynamics of natural selection, Proc. Natl. Acad. Sci. U.S.A. **111**, 12408 (2014).

[113] Z. Yan and J. Wang, Funneled energy landscape unifies principles of protein binding and evolution, Proc. Natl. Acad. Sci. U.S.A. **117**, 27218 (2020).

[114] G. Odor, Phase transition universality classes of classical, nonequilibrium systems, Rev. Mod. Phys. **76**, 663 (2004).

[115] F. Rao and A. Caflisch, The protein folding network, J. Mol. Biol. **342**, 299 (2009).