

**Directed percolation criticality in eternal inflation**Justin Khoury and Sam S. C. Wong *Center for Particle Cosmology, Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA*

(Received 27 September 2023; accepted 25 October 2023; published 28 November 2023)

False-vacuum eternal inflation can be described as a random walk on the network of vacua of the string landscape. In this paper we show that the problem can be mapped naturally to a problem of directed percolation. The mapping relies on two general and well-justified approximations for transition rates: (1) the downward approximation, which neglects “upward” transitions, as these are generally exponentially suppressed; and (2) the dominant decay channel approximation, which capitalizes on the fact that tunneling rates are exponentially staggered. Lacking detailed knowledge of the string landscape, we model the network of vacua as random graphs with arbitrary degree distribution, including Erdős-Rényi and scale-free graphs. As a complementary approach, we also model regions of the landscape as regular lattices, specifically Bethe lattices. We find that the uniform-in-time probabilities proposed in our previous work favor regions of the landscape poised at the directed percolation phase transition. This raises the tantalizing prospect of deriving universal statistical distributions for physical observables, characterized by critical exponents that are insensitive to the details of the underlying landscape. We illustrate this with the cosmological constant, and show that the resulting distribution peaks as a power-law for small positive vacuum energy, with a critical exponent uniquely determined by the random graph universality class.

DOI: [10.1103/PhysRevD.108.103541](https://doi.org/10.1103/PhysRevD.108.103541)**I. INTRODUCTION**

Our universe appears to be tantalizingly poised at criticality. Extrapolating the Higgs effective potential reveals that the electroweak vacuum lies within a tiny parameter region of metastability [1–10], a result that is exquisitely sensitive to the top quark and Higgs boson masses. This hinges on an enormous cancellation between the exponentially small decay rate and the exponentially large observable volume of the universe. It is tempting to speculate that there is an intricate relation between various measured quantities: the cosmological constant (CC), which sets the observable volume of the universe, and the Higgs and top quark masses, which set the Higgs effective potential when extrapolated to high enough energy.

Other fine-tuned features of our universe can also be interpreted as near-criticality. In light of the nondetection of supersymmetry at the electroweak scale, the nearly vanishing ratio of Higgs mass to  $M_{\text{Pl}}$  is a result of exponential fine-tuning, which can be interpreted as the boundary of broken/unbroken electroweak symmetry [11]. In cosmology, the CC problem translates to our universe being nearly Minkowski

space, which bifurcates into ever-expanding de Sitter (dS) and crunching anti-de Sitter (AdS) space-times with distinct asymptotics and stability properties [12,13].

In the context of eternal inflation [14–18], universes in causally disconnected Hubble patches possess different physical laws. New universes are being constantly generated in all patches. At the same time, string theory predicts an exponentially large number of metastable vacua with enormously rich low-energy physics [19–21]. False-vacuum eternal inflation is essentially a random walk on the network of vacua of the string landscape. Given these, it is natural to ask if the near-criticality of our universe can be approached from a statistical point of view.

In order to extract predictions in the multiverse, it is intuitive to study statistically the distribution of vacua and their associated physical properties. Although deriving such a statistical distribution may ultimately require a complete understanding of quantum gravity, it is still instructive to approximate it using a semiclassical prescription. Attempts to define semi-classical 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. However, it is well known that defining such a measure is ambiguous, as it is assumption-dependent even under the same framework [22–25].

In a recent paper [26], we presented a general Bayesian framework for probabilistic reasoning in eternal inflation.

---

*Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP<sup>3</sup>.*

Different assumptions about the measure problem amount to different choices of priors to define probabilities. We identified two prior distributions, both pertaining to initial conditions, that must be specified to obtain well-defined occupational probabilities for different vacua. Since eternal inflation is geodesically past-incomplete [27], we know that we exist a finite time  $t$  since the onset of eternal inflation. Our ignorance about the time of existence is captured by a prior density  $\rho(t)$ . Relatedly, along our past world-line eternal inflation must have started within some particular “ancestral” dS vacuum, but we do not know which one. Our ignorance about the ancestral vacuum is parametrized by a probability distribution  $p_\alpha$  over dS vacua. *Different proposed solutions to the measure problem simply amount to different choices for these two priors.*

In [26] we argued that there are two natural and well-justified choices for the time-of-existence prior  $\rho(t)$ :

- (i) Since the number of observers grows with volume, a natural choice is  $\rho(t) \sim a^3$ . This is equivalent to weighing probabilities by physical volume. The resulting “late-time/volume-weighted” probabilities coincide with the measure of Garriga, Vilenkin, Schwartz-Perlov, and Winitzki (GSVW) [28]. This choice of prior reflects the belief that we exist at asymptotically late times in the unfolding of the multiverse, much later than the exponentially long relaxation time for the landscape, such that probabilities have settled to a quasistationary distribution. This assumption is adopted in nearly all existing approaches to the measure problem [22,23,28–32].
- (ii) Alternatively, motivated by the time-translational invariance of the random walk on the landscape, a natural choice is the uniform prior:  $\rho(t) = \text{const.}$  (To be clear, this is uniform in either proper time or  $e$ -folding time.) The resulting “uniform-in-time” probabilities agree with the prior probabilities of [33]. They are closely related to the “comoving” probabilities proposed in [28,32], as well as probabilities derived recently using the local Wheeler-De Witt equation [34]. Importantly, the uniform-in-time probabilities favor vacua that are accessed early on in the evolution of the multiverse, during the *approach to equilibrium*. This is consistent with the early time approach to eternal inflation developed recently [35–39].

The Bayesian framework allowed us to compare the plausibility of the uniform-in-time and late-time measures to explain our data by computing the Bayesian evidence for each. We argued, under general and plausible assumptions, that posterior odds overwhelmingly favor uniform-in-time probabilities [26]. The argument, briefly reviewed in Sec. II, relies on assumptions that have been made in previous studies of the landscape. There are some caveats, of course, and we tried to enunciate them carefully in [26]. For this reason, we believe that the uniform-in-time

measure is the correct objective approach to probabilistic reasoning in the multiverse. In this work we therefore focus on uniform-in-time probabilities.

### A. Directed percolation in eternal inflation

Our goal in this paper is to show how vacuum dynamics on the landscape can be mapped naturally to a problem of directed percolation. The mapping relies on two very general and reasonable assumptions about transition rates between vacua:

- (1) The first assumption is that transition rates between dS vacua satisfy a condition of detailed balance, such that “upward” jumps are exponentially suppressed by  $\frac{\kappa_{\text{up}}}{\kappa_{\text{down}}} \sim e^{-\Delta S}$ , where  $S$  is the dS entropy. This is satisfied by most tunneling instantons, including Coleman-De Luccia (CDL) [40–42]. Thus we are justified to work in the *downward approximation* [43,44], wherein upward transitions are neglected to leading order. In this approximation, the network of vacua becomes a *directed graph*.
- (2) The second assumption rests on the fact that semi-classical tunneling rates depend exponentially on the Euclidean action of the instanton,  $\kappa \sim e^{-S_E}$ . In turn,  $S_E$  depends sensitively on the height and width of the potential barrier. Because of this exponential sensitivity, branching ratios for dS vacua are typically overwhelmingly dominated by a single decay channel. This defines the *dominant decay channel*, in which exponentially subdominant decay channels are neglected.

It should be stressed that these approximations are not strictly necessary to study percolation. They are made for convenience, to simplify the problem, and we will discuss how the analysis can be generalized by relaxing them. In any case, with these approximations, the uniform-in-time probabilities reduce to a simple and intuitive observable in directed graphs. Namely, the probability to occupy a given node  $I$  simplifies to

$$P(I) \sim s_I, \quad (1)$$

where  $s_I$  is the number of ancestors, i.e., nodes that can reach  $I$  through a sequence of directed (downward) transitions. See Fig. 4. Thus the measure favors vacua with a large basin of ancestors. In other words, *regions of the landscape with large probability must therefore have the topography of a deep valley, or funnel* [26,36,37,39]. This is akin to the smooth folding funnels of protein conformation landscapes [45], and those of atomic clusters with Leonard-Jones interactions [46–48]. In the context of deep learning, it has been argued that deep neural networks that generalize well have a loss function characterized by a smooth funnel [49]. Another instance is the “big valley” hypothesis in combinatorial optimization (e.g., the search

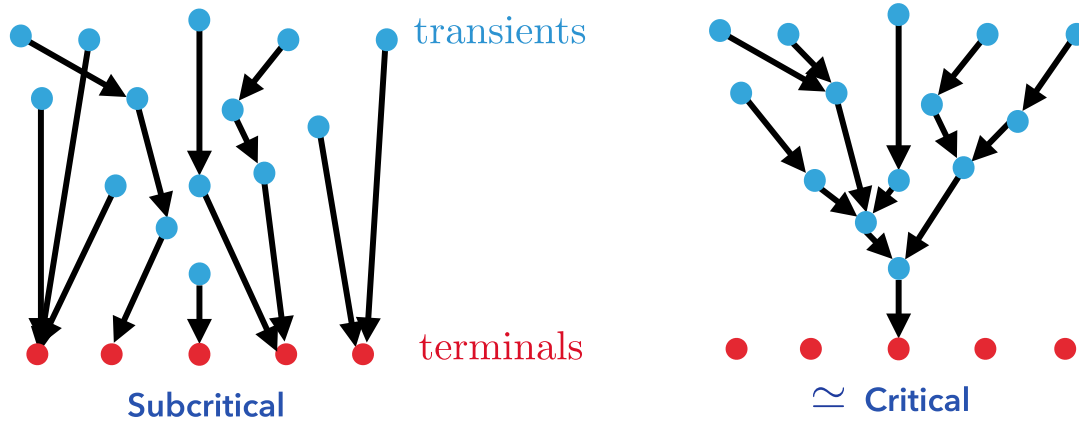


FIG. 1. A region of the landscape, comprised of transient dS vacua (blue nodes) and terminals (red nodes). In the dominant decay channel approximation, each dS vacuum has exactly one outgoing link. Left: if dS vacua decay primarily to terminals, the region breaks up into many small disconnected components and is therefore subcritical. Right: if dS vacua mainly decay to other dS vacua, then a giant connected component can emerge, and the region is near percolation criticality.

space of the traveling salesman problem), where it is conjectured that local optima are clustered around the central global optimum [50]. It is tempting to speculate that funnels are a generic solution to optimization problems on complex energy landscapes.

Equation (1) gives an intuitive and well-justified notion of probability for different vacua. But what can we reasonably assume about the network of vacua, given our limited understanding of the string landscape? Lacking detailed knowledge of the underlying network of vacua, it seems sensible to model regions of the landscape as random graphs. Random graphs have a long and venerable history, going back to the seminal work of Erdős and Rényi graph [51]. In general, they can be defined by specifying a probability distribution that a given node has a certain degree (number of links), with Erdős-Rényi graphs corresponding to the special case of a Poisson degree distribution. In this work we follow [52] and consider arbitrary degree distributions, including scale-free random graphs.

As a complementary approach, we also model landscape regions as a regular lattice, specifically a Bethe lattice (or Cayley tree). This is suitable for local string landscapes in which vacua form a regular network, for instance the axion landscape [53–55]. The directed percolation transition can be studied analytically for both Bethe lattices and (Erdős-Rényi) random graphs [52]. Remarkably, despite being extreme opposites in terms of graph “regularity,” Bethe lattices and Erdős-Rényi graphs belong to the same percolation universality class. Thus it is our hope that, despite being highly simplified and idealized, these two approaches offer important lessons about percolation phenomena on the landscape, that are applicable to more realistic dynamics.

To see how this maps to directed percolation, consider the landscape networks shown in Fig. 1, comprised of dS

transient vacua (blue nodes) and AdS/Minkowski terminal vacua (red nodes).<sup>1</sup> In the downward approximation, all transitions are one-way, and the graphs are directed. Each dS transient has exactly one directed edge emanating from it, corresponding to its dominant decay channel. Now, if dS vacua mainly decay to terminals (left panel), then the region breaks up into many small disconnected components, resulting in  $s_I \sim \mathcal{O}(1)$ , and thus low probability. If, however, dS vacua mainly decay to other dS vacua (right panel), then the graph can include a very large connected component, with low-lying nodes having  $s_I \gg 1$ . This corresponds to the emergence of a giant component at the percolation phase transition.

It is clear from these simple considerations that *vacua with the highest occupational probability reside in landscape regions poised at the directed percolation phase transition*. This is a key result of our analysis. In hindsight, since the uniform-in-time measure is relevant for the approach to equilibrium in landscape dynamics, its relation to directed percolation is perhaps not surprising, as directed percolation is the paradigmatic nonequilibrium critical phenomenon [57].

As usual, the power of criticality lies in universality. Near the percolation phase transition, various quantities assume power-law (scale-invariant) probability distributions, characterized by universal critical exponents that are insensitive to the microscopic details of the system. This raises the tantalizing possibility of deriving universal probability distributions for physical observables without detailed knowledge of the underlying landscape. We illustrate this point concretely with the CC, and briefly mention other potential observables in the conclusions.

<sup>1</sup>We assume throughout that AdS/Minkowski vacua are terminal, acting as absorbing nodes. See [56], however, for a recent discussion of the possibility to up-tunnel out of AdS.

## B. Universal probability distribution for the CC

As reviewed in Sec. V, at criticality the probability  $P_s$  that a randomly chosen vacuum has  $s$  ancestors displays a power-law tail,  $P_s \sim s^{-3/2}$ . (The particular critical exponent of  $-3/2$  holds for the Erdős-Rényi universality class; scale-free graphs have different critical exponents.) Assuming only that the underlying CC probability distribution on the landscape is smooth as  $v \rightarrow 0^+$ , we show in Sec. VII that this translates into a (nonanthropic) universal probability distribution for the CC (which takes into account eternal inflationary dynamics) that is also power-law near the origin:

$$P(v) \sim v^{-3/2}. \quad (2)$$

Thus the uniform-in-time cosmological measure favors small, positive vacuum energy. Quantitatively, the 95% confidence interval for the CC is set by the size of the giant component, which is famously  $\mathcal{O}(N_{\text{dS}}^{2/3})$  for Erdős-Rényi graphs:

$$v \lesssim N_{\text{dS}}^{-2/3}. \quad (3)$$

This can explain the observed CC,  $v_{\text{obs}} \sim 10^{-120}$ , if our vacuum belongs to a funneled region of size  $N_{\text{dS}} \sim 10^{240}$ . To be clear, here  $N_{\text{dS}}$  is the number of dS (transient) vacua in a funnel region near directed percolation criticality, not the total number of dS vacua across the entire landscape. Since the measure favors vacua with the largest number of ancestors, we are likely to inhabit the largest funnel region near percolation criticality, i.e., the near-critical region with largest  $N_{\text{dS}}$ .

Before closing, we should mention other occurrences of percolation criticality in the context of eternal inflation, and how they contrast with our framework. It is well known that the bubbles generated in false-vacuum inflation exhibit a percolation phase transition when the nucleation probability within a Hubble volume,  $\kappa \equiv \Gamma/H^4$ , reaches a critical value somewhere in the range  $10^{-6} \lesssim \kappa_c \lesssim 0.24$  [58]. This transition describes bubble percolation in space-time, as opposed to the percolation phase transition in the network of vacua discussed in this work. Relatedly, in the context of slow-roll inflation, it was shown in [59] that the phase transition to eternal inflation can be described by a Galton-Watson branching process [60], whose critical behavior is equivalent to directed percolation, e.g., [61]. Lastly, we should mention the mechanism of “self-organized localization” [62], 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.

The paper is organized as follows. In Sec. II we briefly review vacuum dynamics as an absorbing Markov process on the network of vacua. In Sec. III we describe the general

Bayesian approach to the measure problem, and review the late-time/volume-weighted and uniform-in-time probabilities as two well-justified choices of priors. We also review the argument, originally given in [26], that posterior odds overwhelmingly favor the uniform-in-time hypothesis. In Sec. IV, we discuss how the mapping of vacuum dynamics to a problem of directed percolation. Section V is a rather comprehensive review of the key notions of percolation on random graphs, both undirected and directed, with general degree distributions. In Sec. VI we apply these notions to the case of interest, namely random networks with terminal (AdS) vacua, and argue that uniform-in-time probabilities favor regions of the landscape poised at directed percolation criticality. In Sec. VII we show how the probability distributions of ancestors and descendants, which assume power-law tails at criticality, translate into universal distributions for the CC with certain critical exponents. We summarize our results and discuss a few avenues of future research in Sec. VIII.

## II. BRIEF REVIEW OF VACUUM DYNAMICS

Vacuum dynamics on the string landscape are described by a linear Markov process [28,31]. Technically, this is an absorbing Markov process, because of AdS vacua which act as terminals. As a result, detailed balance is explicitly violated, and the dynamics are out of equilibrium. The Markov process describes the probability  $f_I(t)$  along a given world-line to occupy vacuum  $I$  as a function of time. This probability satisfies the master equation

$$\frac{df_I}{dt} = \sum_J \kappa_{IJ} f_J - \sum_K \kappa_{KI} f_I, \quad (4)$$

where  $\kappa_{IJ}$  is the  $J \rightarrow I$  transition rate. (Terminal vacua by definition have  $\kappa_{aI} = 0$ .) While most of the results in this section hold for general tunneling rates, we have in mind transitions mediated by semiclassical instantons, such as Coleman-De Luccia (CDL) [40–42], Hawking-Moss [63], and Brown-Teitelboim [64].

The general time variable  $t$  is related to proper time in vacuum  $I$  via a lapse function:

$$\Delta\tau_I = \mathcal{N}_I \Delta t. \quad (5)$$

The master equation relies on coarse-graining over a time  $\Delta\tau_I$ , which must be longer than any transient evolution between epochs of vacuum energy domination.<sup>2</sup> Since AdS

<sup>2</sup>At the same time,  $\Delta\tau_I$  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  $H_{\text{parent}}$  is the Hubble rate of the parent dS vacuum (see, e.g., [65]).



bubbles crunch in a Hubble time, coarse-graining spans their entire evolution.

The probabilities  $f_I(t)$  have a dual interpretation. They can be interpreted “locally,” as occupational probabilities along a world-line. Or, they can be interpreted “globally,” as the fraction of comoving volume that each vacuum occupies on a spatial hypersurface at time  $t$ . We mainly adopt the former interpretation. Equation (4) makes two properties of the  $f_I$ ’s clear: (1) The master equation (4) is manifestly invariant under redefinitions of  $t$ , hence the  $f_I$ ’s are time-reparameterization invariant; and (2) Because summing the right-hand side over  $I$  gives zero, the  $f_I$ ’s can be normalized:

$$\sum_I f_I = 1. \quad (6)$$

Thus the  $f_I(t)$ ’s offer well-defined, time-reparameterization invariant probabilities to occupy different vacua at time  $t$ .

We will be primarily interested in the dS component of the master equation, given by

$$\frac{df_i}{dt} = \sum_j M_{ij} f_j; \quad M_{ij} \equiv \kappa_{ij} - \delta_{ij} \kappa_j. \quad (7)$$

Here,  $M_{ij}$  is the dS  $\rightarrow$  dS transition matrix, and  $\kappa_i \equiv \sum_j \kappa_{ji}$  is the total decay rate of vacuum  $i$ . 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, a property which has been argued to be valid for the string landscape [66]. Equation (7) can be solved in terms of a Green’s function:

$$f_i(t) = \sum_\alpha (e^{Mt})_{i\alpha} p_\alpha, \quad (8)$$

where  $p_\alpha \equiv f_\alpha(0)$  is the initial probability over ancestral vacua. Since eternal inflation by definition started in a dS vacuum, the initial probabilities satisfy

$$\sum_{\alpha=1}^{N_{\text{dS}}} p_\alpha = 1. \quad (9)$$

### III. BAYESIAN PROBABILITIES

To define Bayesian probabilities, one 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. An important fact is that eternal inflation, while eternal into the future, is not eternal into the past. That is, an eternally inflating space-time is past geodesically incomplete [27]. This has two implications:

- (i) We exist a finite time  $t$  after the onset of eternal inflation, but we do not know how long ago that was.

We must therefore parametrize our ignorance about the time of existence with a prior density  $\rho(t)$ , normalized as  $\int_0^\infty dt \rho(t) = 1$ .

- (ii) Along our past world-line, eternal inflation started in some ancestral dS vacuum  $\alpha$ , but we do not know which one. Our ignorance about the ancestral vacuum is captured by the initial probability distribution  $p_\alpha$ .

Lastly, it is customary to condition probabilities on one piece of observational data. Namely, that we exist in our dS pocket universe during the transient period before vacuum domination. That is, we exist within a coarse-graining time  $\Delta t$  after nucleation of our bubble.

It is then straightforward to write down the joint probability distribution  $P(I, t, \alpha)$  to inhabit a bubble of vacuum  $I$ , nucleated at time  $t$ , starting from an ancestral vacuum  $\alpha$ :

$$P(I, t, \alpha) \sim \sum_j \kappa_{Ij} \Delta t (e^{Mt})_{j\alpha} p_\alpha \rho(t). \quad (10)$$

This is easy to understand. The factor  $(e^{Mt})_{j\alpha} p_\alpha$  is the probability to evolve from ancestral vacuum  $\alpha$  to parent dS vacuum  $j$  at time  $t$ ; while the factor  $\kappa_{Ij} \Delta t$  is the probability to transition from parent dS vacuum  $j$  to vacuum  $I$  in the next  $\Delta t$ . Lastly, we weigh the time of nucleation with  $\rho(t)$ , and sum over all dS parents  $j$ . To our mind, the above joint probabilities are the correct objective approach to inductive reasoning in the multiverse. They accurately encode our ignorance about when and where eternal inflation started in our past. *Different approaches to the measure problem simply amount to different choices for the priors  $p_\alpha$  and  $\rho(t)$ .*

Within this general framework, one can perform the three main operations of Bayesian inference:

- (1) By marginalizing over the model parameters  $t$  and  $\alpha$ , and using (8), we obtain the *prior predictive distribution*:

$$P(I) \sim \sum_j \kappa_{Ij} \Delta t \int_0^\infty dt f_j(t) \rho(t). \quad (11)$$

This distribution informs us on which vacua are statistically favored without taking any data (e.g., value of the CC, particle spectrum etc.) into consideration, other than conditioning on our bubble being nucleated within the last  $\Delta t$ .

- (2) Different hypotheses  $\mathcal{H}_1$  and  $\mathcal{H}_2$ , corresponding to different choices of priors, can be compared by computing the posterior odds:

$$\frac{P(\mathcal{H}_1|D)}{P(\mathcal{H}_2|D)} = \frac{P(D|\mathcal{H}_1) P(\mathcal{H}_1)}{P(D|\mathcal{H}_2) P(\mathcal{H}_2)}, \quad (12)$$

where  $P(\mathcal{H}_i)$  is the prior odds for each hypothesis, and  $\frac{P(D|\mathcal{H}_1)}{P(D|\mathcal{H}_2)}$  is the Bayes factor. The data  $D$  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  $\Lambda$ CDM model.

- (3) Conditioning on our data  $D$  for a given choice of priors, we can perform parameter inference. For instance,  $P(t|D)$  gives the posterior distribution for the time of nucleation.

Each of these operations was studied in detail in [26]. In what follows we will be primarily interested in the prior predictive probabilities (11).

### A. Uniform-in-time measure

As mentioned above, a choice of measure amounts to specifying a choice of priors  $p_\alpha$  and  $\rho(t)$ . Consistency requires that priors reflect *all* information at hand, while at the same time being minimally informative.

Let us first discuss the time of nucleation prior  $\rho(t)$ , as it is most important to determine  $P(I)$ . In general, specifying a prior for a continuous variable is tricky, for the obvious reason that a uniform prior is not reparametrization invariant. Following Jaynes [67], a useful strategy in this case is to identify the symmetries of the problem and apply the notion of group invariance. Logical consistency requires that our prior be invariant under all symmetry transformations.

In the case at hand, a key property of the master equation (4) is that it is *time-translation invariant*. More precisely it is invariant under translations in proper time, as well as any time variable  $t$  related to proper time via a lapse function  $\mathcal{N}_I$  depending on  $H_I$  only (e.g., scale factor/ $e$ -folding time). Without additional information, the most natural choice is the uniform prior<sup>3</sup>:

$$\rho(t) = \text{constant}. \quad (13)$$

To be clear, this prior is uniform in proper time and  $e$ -folding time. In terms of conformal time, however, it corresponds to the Jeffreys prior,  $\rho(\eta) \sim \eta^{-1}$ , consistent with the dS dilation symmetry  $\eta \rightarrow \lambda\eta$ ,  $\vec{x} \rightarrow \lambda\vec{x}$ .

Substituting into (11), we can perform the time integral using the identity

$$\int_0^\infty dt (e^{Mt})_{ij} = -M_{ij}^{-1} = \kappa_i^{-1} (\mathbb{1} - T)_{ij}^{-1}, \quad (14)$$

<sup>3</sup>Of course, a uniform distribution on the half real line is not normalizable, so (13) is technically an improper prior. One can instead work with a regularized prior, such as  $\rho(t) = \epsilon e^{-\epsilon t}$  or  $\rho(t) = 1/T$  over  $0 \leq t \leq T$ . As shown in [26], the resulting probabilities are independent of the regulator as it is removed, i.e., as  $\epsilon \rightarrow 0$  or  $T \rightarrow \infty$ , respectively.

where  $T_{ij} \equiv \frac{\kappa_{ij}}{\kappa_j}$  is the branching ratio. (More generally, the branching ratio matrix has components  $T_{Ij} = \frac{\kappa_{Ij}}{\kappa_j}$ ,  $T_{ab} = \delta_{ab}$ , and  $T_{ja} = 0$ , such that  $\sum_I T_{IJ} = 1$  for all  $J$ .) Equation (11) then gives

$$P(I) \sim \sum_j T_{Ij} \sum_\alpha (\mathbb{1} - T)_{j\alpha}^{-1} p_\alpha. \quad (15)$$

The matrix  $(\mathbb{1} - T)^{-1}$  is known as the fundamental matrix for the absorbing Markov chain. Expanding it as a geometric series,  $(\mathbb{1} - T)_{ij}^{-1} = \delta_{ij} + T_{ij} + \sum_k T_{ik} T_{kj} + \dots$ , it is easily recognized as the total branching probability for all transition paths connecting  $j$  to  $i$ . Thus  $P(I)$  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.

Next, consider the prior  $p_\alpha$  over ancestral vacua. This was discussed in detail in [26], and we briefly mention the salient points. The prior  $p_\alpha$  pertains to the question of the initial state in quantum cosmology, which has been the subject of active debate for decades and remains an open problem. A well-motivated proposal for the quantum creation of a closed universe is the Hartle-Hawking (HH) state [68,69], which exponentially favors the lowest energy (highest entropy) dS vacuum. Another well-studied proposal is the tunneling wave function [70–73], which instead favors high-energy/low-entropy initial vacua. Thus the tunneling wave function favors (high-energy) inflation, whereas the HH state does not [74].

As motivated in [26], a reasonable attitude is to err on the side of maximal ignorance and apply the principle of indifference:

$$p_\alpha = \frac{1}{N_{\text{dS}}}. \quad (16)$$

(If the number of dS vacua in the landscape is infinite [75–77], then (16) would represent an improper prior, which is fine since the resulting probabilities would nevertheless be well-defined.) Because high-energy dS vacua are expected to vastly outnumber low-energy dS 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. If the HH state turns out to be the correct initial conditions for eternal inflation in our past, then this would have important implications for the uniform-in-time probabilities.

Adopting (16), (15) reduces to

$$P(I) \sim \sum_j T_{Ij} \sum_\alpha (\mathbb{1} - T)_{j\alpha}^{-1}. \quad (17)$$

This distribution agrees with the prior probabilities of [33], and is closely related to the “comoving” probabilities

proposed in [28,32]. We will see that this admits a clear and intuitive interpretation with the simplifying assumptions discussed in Sec. IV.

### B. Late-time/volume-weighted measure

Another reasonable choice for  $\rho(t)$  is motivated by the fact that the number of observers grows with volume. Hence,  $\rho(t)$  should grow accordingly<sup>4</sup>:

$$\rho(t) \sim a^3(t). \quad (18)$$

As shown in [26], this is equivalent to weighing occupational probabilities by physical volume.

Because the prior is sharply peaked at late times, the occupational probabilities  $f_j(t)$  can be approximated by their asymptotic form

$$f_j(t) \simeq s_j e^{-qt}. \quad (19)$$

Here  $s_j$  the so-called dominant eigenvector of  $M_{ij}$ , which by definition has the largest (least negative) eigenvalue  $-q$  [28]. Substituting into (11), we obtain in this case

$$P_{\text{late}}(I) \sim \sum_j \kappa_{Ij} s_j. \quad (20)$$

This agrees with the GSVW measure [28] obtained by counting bubbles along a world-line.

The above distribution admits an intuitive explanation in downward perturbation theory, discussed in Sec. IVA below. In this approximation, the dominant eigenvector takes a simple form [44]:

$$s_j \simeq \frac{\kappa_{j\star}}{\kappa_j} (\mathbb{1} - T)_{j\star}^{-1}, \quad (21)$$

where  $\star$  denotes the most stable (i.e., longest-lived) dS vacuum, also known as the dominant vacuum. Thus (20) becomes

$$P_{\text{late}}(I) \sim \sum_j T_{Ij} (\mathbb{1} - T)_{j\star}^{-1}. \quad (22)$$

In other words, this is recognized as the total branching probability from  $\star$  to  $I$ . The late-time/volume-weighted measure is independent of initial conditions (i.e., independent of  $p_\alpha$ ), reflecting the attractor nature of eternal inflation. However, somewhat paradoxically, (22) coincides with (15) for the special choice of initial conditions  $p_\alpha = \delta_{\alpha\star}$ .

<sup>4</sup>For the prior distribution to be normalizable, a regulator is once again necessary. This can be achieved simply by imposing a cutoff time  $t_c$ , and letting  $t_c \rightarrow \infty$  at the end of the calculation. The resulting probabilities are insensitive to the cutoff.

### C. Model comparison favors uniform-in-time probabilities

In [26] we compared the Bayesian evidence for the uniform-in-time and late-time measures by computing the Bayes factor  $\frac{P(D|\mathcal{H}_{\text{late}})}{P(D|\mathcal{H}_{\text{uni}})}$ . We argued, under general and plausible assumptions, that it overwhelmingly favors the uniform-in-time hypothesis. The reason is easily understood intuitively. Since  $\star$  is the most stable vacuum anywhere in the landscape, it is likely that it can only decay via an upward transition, because upward jumps are doubly exponentially suppressed (as discussed in Sec. IVA). Therefore, the branching probability to vacua compatible with our data is also doubly exponentially suppressed. In contrast, for uniform-in-time probabilities, if vacua compatible with our data can be reached from some ancestral vacua via a sequence of downward transitions, then the Bayesian evidence is likely exponentially small, but not doubly exponentially suppressed.

Furthermore, conditioning on our data  $D$ , we performed in [26] parameter inference to determine the most likely time of nucleation. For the uniform-in-time hypothesis, we found that the average time for occupying vacua compatible with our data is much shorter than mixing time for the landscape. This is fully consistent with the ‘‘early time’’ approach to eternal inflation [36–39], which proposes that we live during the *approach to equilibrium* in the unfolding of the multiverse. See [35] for related ideas. This is in contrast with the late-time/volume-weighted distribution, which reflects the belief that the evolution of the multiverse has been going on for an exponentially long time, much longer than the mixing time of the landscape, such that the occupational probabilities have settled to a quasistationary distribution.

We henceforth focus on the uniform-in-time measure (17).

## IV. MAPPING TO A DIRECTED PERCOLATION PROBLEM

The landscape can be modeled as a network (or graph). The nodes represent the different vacua, while the links define the network topology and represent all relevant transitions between vacua. There are two types of nodes: transients (dS) and terminals/absorbing (AdS and Minkowski). Because transition rates are different along each link, the graph is said to be weighted. The master equation (4) describes a random walk on this weighted network. The measures derived above are closely related to network centrality indices: the uniform-in-time measure (17) is analogous to Katz centrality [78]; the late-time (GSVW) measure (22) to eigenvector centrality.

In this section we show how the problem can be mapped to a problem of directed percolation. Directed percolation is the paradigmatic critical phenomenon for nonequilibrium systems [57]. It is perhaps not surprising that the absorbing Markov process describing vacuum dynamics on the

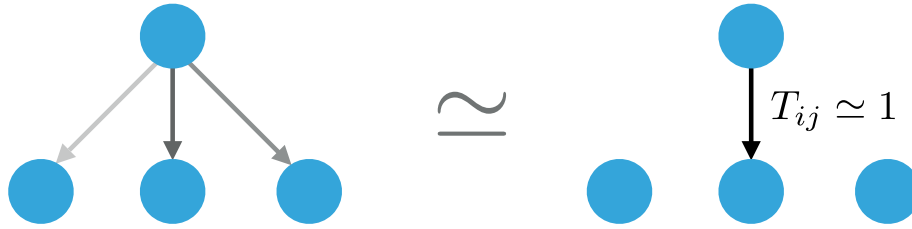


FIG. 2. This explains the dominant decay channel approximation. On the left, in the downward approximation a given node has possibly many allowed decay channels (directed links), but with exponentially staggered branching ratios (different shades of gray). On the right, the approximation amounts to only keeping the link with largest branching ratio. The parent node therefore has out-degree 1.

landscape, which is inherently nonequilibrium, belongs to the universality class of directed percolation. The mapping relies on two very general and reasonable assumptions about transition rates between vacua, discussed respectively in Secs. IV A and IV B. We will see that, with these approximations, the uniform-in-time probabilities reduce to a simple and intuitive observable in directed graphs, namely the number of ancestors of a given node.

### A. Downward approximation

The first assumption is that transitions between dS vacua satisfy a condition of *detailed balance* [79]:

$$\frac{\kappa_{ji}}{\kappa_{ij}} \sim e^{S_j - S_i}, \quad (23)$$

where  $S_j = 8\pi^2 M_{\text{pl}}^2 / H_j^2$  is the dS entropy. This condition is satisfied by CDL, Hawking-Moss and Brown-Teitelboim instantons. More generally, it is consistent with the interpretation of quantum dS space as a thermal state [80].<sup>5</sup> Notice that (23) depends only on the false and true vacuum potential energy—it is insensitive to the potential barrier and does not rely on the thin-wall approximation.

Equation (23) implies that upward transitions, which increase the potential energy, are exponentially suppressed compared to downward tunneling. This allows one to define a “downward” approximation [43,44], in which upward transitions are neglected to zeroth order. (Upward transitions are treated perturbatively at higher order.) In this approximation, the network of vacua reduces to a directed, acyclic graph [88], i.e., without directed loops, whereby a link from  $j$  to  $i$  is only allowed if  $V_j \geq V_i$ . This may be a good place to point out that the validity of the master equation (4) has not been rigorously established for upward transitions. So it may be the case that the

<sup>5</sup>Notably, (23) is violated by the Farvi-Guth-Guven process [81], though the interpretation of its singular instanton is still unsettled [82–86]. It is also violated by the mechanism of nucleating localized, high-energy regions proposed recently [87]. Upward jumps are still suppressed in this case, by at least  $e^{\#M_{\text{pl}}/H_{\text{low}}}$  instead of  $e^{\#M_{\text{pl}}^2/H_{\text{low}}^2}$ .

description of vacuum dynamics as a Markov process is only legitimate in the strict downward approximation.

In any case, dS vacua that can only decay via upward transitions become effectively terminal in this approximation. In other words, in the downward approximation terminals consist both of AdS/Minkowski vacua and dS vacua with upward-only decay channels. Transient nodes are dS vacua with at least one downward decay channel.

### B. Dominant decay channel

The second assumption is motivated by a generic feature of transition rates in quantum field theory, namely that they are exponentially staggered. This is because tunneling rates depend exponentially on the instanton Euclidean action:

$$\kappa \sim e^{-S_E}. \quad (24)$$

For CDL tunneling, for instance,  $S_E$  depends sensitively on the shape of the potential, such as the height and width of the barrier. Because of this exponential sensitivity, branching ratios for dS vacua are typically overwhelmingly dominated by a single decay channel, with  $T_{ij}^{\text{dom}} \simeq 1$ , while other decay channels are comparatively exponentially suppressed, i.e.,  $T_{ij}^{\text{other}} \simeq 0$ .

Hence our second simplifying assumption is that we work in the approximation where  $T_{ij}$  is 0 or 1. In other words, either there is a link between two nodes ( $T_{ij} \simeq 1$ ) or not ( $T_{ij} \simeq 0$ ). Furthermore, if  $T_{ij} \simeq 1$ , then this is the only link emanating from  $j$ , i.e., the out-degree of  $j$  is 1. This is illustrated in Fig. 2. There are of course exceptions, for instance in regular lattices of flux vacua [19]. But we expect that single-channel dominance is justified for random landscapes, which will be our primary interest in Sec. V.

### C. Implications for uniform-in-time probabilities

The two assumptions discussed above greatly simplify the uniform-in-time probabilities (17), and lead to an intuitive interpretation.

- (i) Downward approximation: In this approximation the only contributing paths to a given vacuum  $I$  are those composed of a sequence of downward transitions. It follows that the probabilities (17) favor vacua that



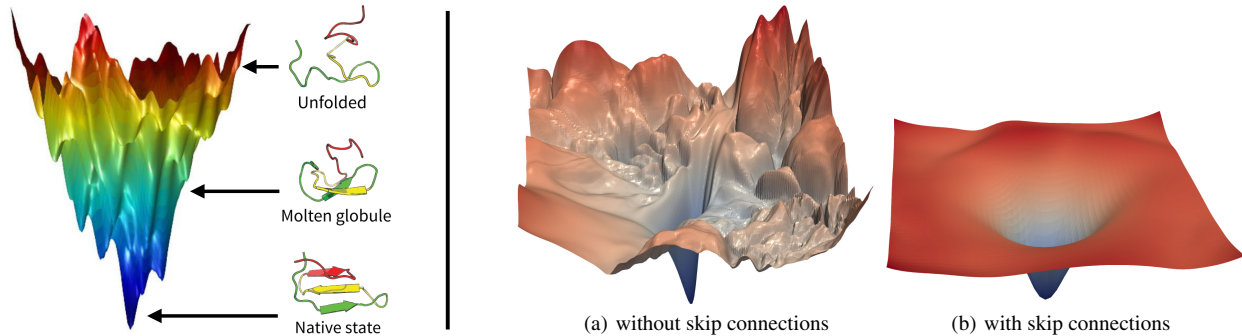


FIG. 3. The uniform-in-time probabilities favor regions of the landscape with the topography of a deep valley, or funnel. This is akin to the free energy landscapes of naturally occurring proteins (Left, reproduced from [89]), which are characterized by a large funnel near the native state. Right: A similar narrative holds in deep learning. This shows the loss surfaces of ResNet-56 with/without skip connections, reproduced from [49]. Skip connections lead to better generalization, and correspond to a loss function characterized by a smooth funnel.

can be accessed through downward transitions from a large basin of ancestors. In other words, *regions of the landscape with large probability must therefore have the topography of a deep valley, or funnel* [26,36,37,39]. This is akin to the smooth folding funnels of protein conformation landscapes [45], as sketched on the left panel in Fig. 3.

A similar narrative holds in deep learning. It has been argued that deep neural networks that generalize well have a loss function characterized by a smooth funnel [49]—see right panel in Fig. 3. Another instance is the “big valley” hypothesis in combinatorial optimization (e.g., the search space of the traveling salesman problem), where it is conjectured that local optima are clustered around the central global optimum [50].

- (ii) Dominant decay channel: In this approximation where  $T_{ij}$  is 0 or 1, the measure (17) simply counts the number  $s_I$  of ancestor vacua that can reach  $I$ :

$$P(I) \sim s_I. \quad (25)$$

This is a key result of our analysis. It entails that the probability of occupying a vacuum is proportional to the number of other nodes that can access it through sequences of unsuppressed ( $T_{ij} \simeq 1$ ), downward transitions. This is illustrated in Fig. 4 for a trivial example.

Thus the problem of determining probabilities on vacua is reduced to a problem of directed percolation. To see this, consider a region of the landscape shown in Fig. 1, comprised of a number of transient dS vacua (blue nodes) and terminals (red nodes). (Recall that in the downward approximation terminals include AdS/Minkowski vacua, as well as dS vacua with only upward decay channels.) Each dS transient has exactly one directed edge emanating from it, corresponding to its dominant decay channel.

Suppose transients decay primarily to terminals, as sketched on the left panel in Fig. 1. In this case, nodes in the region will generically have  $s_I \sim \mathcal{O}(1)$ , corresponding to relatively low probability. From a percolation perspective, the region breaks down into many small disconnected components, and is therefore subcritical. Suppose, on the other hand, that transients decay primarily to other transients, as shown on the right panel. This corresponds to the emergence of a giant directed component, wherein the bottom nodes have  $s_I \gg 1$ , and therefore high probability.

It is clear from these simple considerations that the uniform-in-time probabilities (17) favor regions of the landscape that are close to the directed percolation phase transition [57]. In what follows we will make this precise by studying directed percolation on random graphs and Bethe lattices.

## V. PERCOLATION ON DIRECTED RANDOM NETWORKS

To set up the problem, it is useful to review some essential notions of directed percolation [90,91]. Concretely, in this work we study two simplified approaches for directed

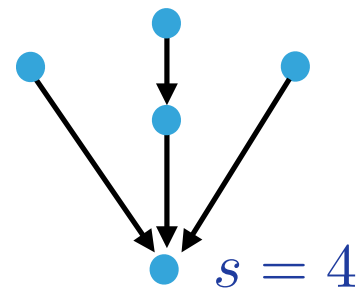


FIG. 4. In the downward and dominant decay channel approximations, the probability for a given node is proportional to the number of its ancestors. In this simple example, the bottom node has  $s = 4$  ancestors.

percolation on the landscape. In the first framework, discussed in this section, we model the a fiducial region of the landscape as a directed random graph with given degree probability distribution. As a special case, the Poissonian degree distribution corresponds to the celebrated Erdős-Rényi graph [51]. In the second approach, discussed in Appendix C, we model the region as a regular lattice, specifically a Bethe lattice.

The directed percolation transition can be studied analytically for both Bethe lattices and (Erdős-Rényi) random graphs [52]. (In fact, they belong to the same universality class, as we will see.) Our focus is on bond percolation, in which the percolation problem on either the Bethe lattice or the Erdős-Rényi random graph is defined by assigning a probability  $p$  that a given edge of the graph is “open.” While the frameworks considered are highly idealized, they allow us to draw important lessons about percolation phenomena on the landscape, which we believe apply more generally to realistic dynamics.

Although much of the analysis is already in the literature, we include it here for completeness. The reader mainly interested in the punchline can skip to (55), which is the main result for our purposes. Our exposition primarily follows [52], which considers random graphs with general degree distributions. For pedagogical purposes, we have also included in Appendix A a review of percolation on undirected random graphs. Many of the results for the undirected case can be easily generalized to directed random graphs.

### A. Directed random graphs

A directed random graph is specified by a joint in-degree and out-degree probability distribution:

$$p_{jk} = \text{probability of randomly chosen node having in-degree } j \text{ and out-degree } k. \quad (26)$$

It is useful to work in terms of its moment generating function,

$$\mathcal{G}(x, y) = \sum_{j,k=0}^{\infty} p_{jk} x^j y^k. \quad (27)$$

Since the distribution is normalized, we have  $\mathcal{G}(1, 1) = \sum_{j,k} p_{jk} = 1$ . Its partial derivatives give in- and-out degree moments of the distribution. For instance, the average in- and out-degrees are given by

$$\begin{aligned} z^{\text{in}} &= \sum_{j,k} j p_{jk} = \left. \frac{\partial \mathcal{G}(x, y)}{\partial x} \right|_{x=y=1}; \\ z^{\text{out}} &= \sum_{j,k} k p_{jk} = \left. \frac{\partial \mathcal{G}(x, y)}{\partial y} \right|_{x=y=1}. \end{aligned} \quad (28)$$

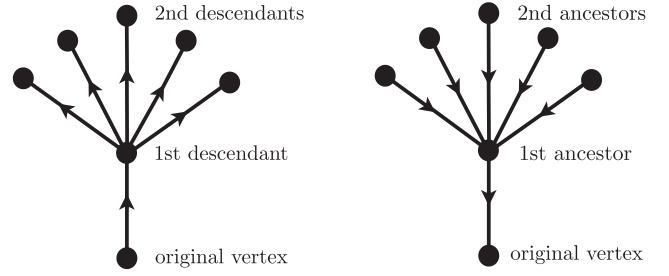


FIG. 5. Left: 1st and 2nd-generation descendants of an original vertex. Right: 1st and 2nd-generation ancestors of that vertex.

Since every link leaving a node terminates at another node, the average in- and out-degrees must be equal:

$$z^{\text{in}} = z^{\text{out}} \equiv z. \quad (29)$$

From  $p_{jk}$ , we can derive the (marginalized) in- and out-degree distributions of a randomly chosen vertex, with generating functions

$$\begin{aligned} F_0(x) &= \mathcal{G}(x, 1) = \sum_{j,k} p_{jk} x^j; \\ G_0(y) &= \mathcal{G}(1, y) = \sum_{j,k} p_{jk} y^k. \end{aligned} \quad (30)$$

In particular, we have  $F_0'(1) = G_0'(1) = z$ .

Now, suppose we start from a randomly chosen vertex, and follow each of its outgoing links to reach its 1st-generation descendants (children), as shown on the left panel of Fig. 5. (In the directed case, ignoring loops, it is natural to distinguish the neighbors of a node as descendants and ancestors.<sup>6</sup>) Let us denote by  $q_k^{\text{out}}$  the out-degree distribution of a 1st descendant. Since we are  $j$  times more likely to arrive at a vertex with in-degree  $j$  than a vertex of degree 1, we have

$$q_k^{\text{out}} = \frac{1}{z} \sum_j j p_{jk}, \quad (31)$$

which is correctly normalized. The corresponding generating function is given by

$$G_1(y) = \frac{1}{z} \sum_{j,k} j p_{jk} y^k = \left. \frac{\partial \mathcal{G}(x, y)}{\partial x} \right|_{x=1}. \quad (32)$$

If the original vertex has out-degree  $k$ , then the number of 2nd descendants is generated by  $(G_1(y))^k$ . (This ignores loops, since their density is  $1/N$ -suppressed close to the percolation threshold for large  $N$ , as argued in Appendix A.)

<sup>6</sup>The terminology is quite apt close to the percolation threshold, given the equivalence with branching processes. See, e.g., [61].

Therefore, the number of 2nd descendants is generated by  $\sum_{j,k} p_{jk}(G_1(y))^k = G_0(G_1(y))$ . For instance, using (28), (30), and (32), the average number of 2nd descendants is

$$z_2 = zG'_1(1) = \left. \frac{\partial^2 \mathcal{G}}{\partial x \partial y} \right|_{x=y=1}. \quad (33)$$

Similarly, suppose we once again start from a randomly chosen vertex, but now follow each of its incoming links in the opposite direction to reach its 1st ancestors (parents). See right panel of Fig. 5. By similar reasoning, the in-degree distribution for a 1st ancestor is generated by

$$F_1(x) = \left. \frac{1}{z} \frac{\partial \mathcal{G}(x, y)}{\partial y} \right|_{y=1}. \quad (34)$$

The number of 2nd ancestors of the original vertex is generated by  $\sum_{j,k} p_{jk}(F_1(x))^j = F_0(F_1(x))$ . For instance, the average number of 2nd ancestors is  $zF'_1(1) = \left. \frac{\partial^2 \mathcal{G}}{\partial x \partial y} \right|_{x=y=1}$ . This is of course identical to (33), given that it has  $x \leftrightarrow y$  symmetry.

### B. Factorized example

An important example is the case where the in- and out-degree distributions are independent,

$$p_{jk} = p_j^{\text{in}} p_k^{\text{out}}. \quad (35)$$

This includes, as a particular case, Erdős-Rényi random graphs [51]. In a directed Erdős-Rényi graph with  $N$  vertices, any two distinct vertices can be connected with a directed edge with probability  $p$ . Therefore both  $p_j^{\text{in}}$  and  $p_k^{\text{out}}$  are given by a binomial distribution:

$$p_k^{\text{out}} = \binom{N}{k} p^k (1-p)^{N-k} \simeq z^k \frac{e^{-z}}{k!}, \quad (36)$$

and similarly for  $p_j^{\text{in}}$ . The last step follows from taking the limit  $N \rightarrow \infty$ , keeping the average degree  $z = p(N-1)$  fixed, to obtain a Poisson distribution.

The generating function (27) also factorizes,

$$\mathcal{G}(x, y) = G_0(y)F_0(x), \quad (37)$$

with  $G_0(y) = e^{z(y-1)}$  and  $F_0(x) = e^{z(x-1)}$ . An immediate consequence of (37) is that  $G_1(y) = G_0(y)$ , i.e., the out-degree distribution for a 1st descendant is the same as that of the original vertex. Similarly,  $F_1(x) = F_0(x)$ . It follows that the average number of 2nd descendants (or ancestors), given by (33), satisfies

$$z_2 = zG'_1(1) = z^2. \quad (38)$$

This holds for arbitrary factorized distribution (35), including the Erdős-Rényi case.

### C. In- and out-component size distribution and percolation phase transition

Going back to the general case, the generating functions defined earlier allow us to study the size distribution of connected components. In the directed case, we must distinguish between the out-component, comprised of all descendants of a given vertex, and the in-component, comprised of all ancestors of a given vertex.

Let us focus for concreteness on the out-component, and define

$$H_1^{\text{out}}(y) = \text{Gen. fcn for number of descendants following a randomly chosen edge.} \quad (39)$$

Ignoring loops, this generating function satisfies the tree-like consistency condition depicted in Fig. 6(a). From the definition (31) for  $q_k^{\text{out}}$ , the treelike structure implies the consistency condition

$$H_1^{\text{out}}(y) = y \sum_{k=0}^{\infty} q_k^{\text{out}} (H_1^{\text{out}}(y))^k = yG_1(H_1^{\text{out}}(y)), \quad (40)$$

where the last step follows from  $q_k^{\text{out}}$  being generated by  $G_1(y)$ . Note that the factor of  $y$  means that the chosen node is included in the counting.

Similarly, we define

$$H_0^{\text{out}}(y) = \text{Gen. fcn for number of descendants of a randomly chosen vertex.} \quad (41)$$

This generating function satisfies the treelike consistency condition depicted in Fig. 6(b). Using the fact that  $G_0(y)$  is the generating function for the out-degree distribution of a randomly chosen vertex, the consistency condition in this case sums up to

$$H_0^{\text{out}}(y) = yG_0(H_1^{\text{out}}(y)). \quad (42)$$

By definition,  $H_0^{\text{out}}$  describes finite components, i.e., it excludes the giant out-component. As long we work below the percolation threshold, such that there is no infinite cluster, then  $H_0^{\text{out}}(1) = 1$ . Above the percolation threshold,  $H_0^{\text{out}}(1)$  gives the fraction of the vertices that do not belong to the giant component. See Appendix A for more details in the undirected case.

Note that the conditional probability  $P^{\text{in}}(s|k)$  for an in-cluster having size  $s$ , given  $k$  in-degree of a vertex, is generated by

$$\mathcal{P}^{\text{in}}(x|k) = \sum_s x^s P^{\text{in}}(s|k). \quad (43)$$

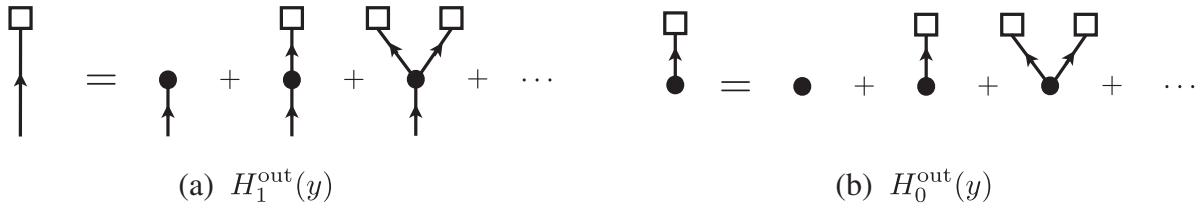


FIG. 6. (a) Treelike structure satisfied by the generating function  $H_1^{\text{out}}$  for the number of descendants following a randomly chosen edge; (b) Same structure, but for the generating function  $H_0^{\text{out}}$  starting from a randomly chosen vertex.

It should be related to the generating function for the number of ancestors  $x^{-1}H_0^{\text{in}}(x) = F_0(H_1^{\text{in}}(x))$  via the definition

$$F_0(H_1^{\text{in}}(x)) = x^{-1}H_0^{\text{in}}(x) = \sum_{k,\ell} p_{k\ell} \mathcal{P}^{\text{in}}(x|k). \quad (44)$$

Therefore by comparison we have

$$\mathcal{P}^{\text{in}}(x|k) = (H_1^{\text{in}}(x))^k. \quad (45)$$

A similar derivation applies to the out-cluster conditional probability.

The algorithm for determining  $H_0^{\text{out}}$  and  $H_1^{\text{out}}$  is then the following. Given a joint degree distribution  $p_{jk}$ , with generating function  $\mathcal{G}(x, y)$ , we can determine  $G_0(y)$  and  $G_1(y)$  using (30) and (32), respectively. Then, the implicit equation (40) can be solved to obtain  $H_1^{\text{out}}(y)$ , and the result is substituted into (42) to obtain  $H_0^{\text{out}}(y)$ .

For general random graphs, it is often difficult in practice to solve (40) analytically. It is, however, straightforward to calculate the moments of the size distribution, in particular the *average size of (finite) connected components*. For instance, consider the average size  $S_{\text{out}}(z)$  of the out-component reached from a random vertex. For simplicity we work below the percolation threshold, such that there is no giant component, and  $H_0^{\text{out}}(1) = H_1^{\text{out}}(1) = 1$ . Using (42), we have

$$S_{\text{out}}(z) = H_0^{\text{out}}(1) = 1 + G_0'(1)H_1^{\text{out}}(1). \quad (46)$$

On the other hand, from (40) we have  $H_1^{\text{out}}(1) = 1 + G_1'(1)H_1^{\text{out}}(1)$ , which implies  $H_1^{\text{out}}(1) = \frac{1}{1-G_1'(1)}$ . Thus,

$$S_{\text{out}}(z) = 1 + \frac{G_0'(1)}{1-G_1'(1)} = 1 + \frac{z^2}{z-z_2}, \quad (47)$$

where we have used (33). Therefore a giant out-component emerges when

$$z_2 = z \quad (\text{directed percolation}). \quad (48)$$

In exactly the same fashion, we can define generating functions  $H_1^{\text{in}}(x)$  and  $H_0^{\text{in}}(x)$  for the in-component size, obtain respectively by following a randomly chosen edge

and starting from a randomly chosen vertex. In doing so, we follow each incoming link in the opposite direction. These generating functions satisfy the implicit relations

$$H_1^{\text{in}}(x) = xF_1(H_1^{\text{in}}(x)); \quad H_0^{\text{in}}(x) = xF_0(H_1^{\text{in}}(x)). \quad (49)$$

Following identical steps as before, it is easy to derive the average size of the in-component:

$$S_{\text{in}}(z) = 1 + \frac{F_0'(1)}{1-F_1'(1)} = 1 + \frac{z^2}{z-z_2}. \quad (50)$$

Therefore a giant in-component emerges when  $z_2 = z$ , which is the same as (48). In other words, the giant in- and out-components emerge simultaneously.<sup>7</sup>

Equation (48) holds for any degree distribution  $p_{jk}$ , with one important assumption—the distribution should have finite mean and variance, i.e., finite  $z$  and  $z_2$ . All degree distributions with this property belong to the Erdős-Rényi percolation universality class. Those that do not, for instance because their degree variance diverges, belong to different universality classes. An important example of the latter are scale-free random graphs, discussed in Sec. V E. In the particular case of factorizable distributions, we can combine (38) and (48) to recover the classic result that percolation occurs when

$$z_c = 1. \quad (51)$$

In particular, for the Poisson distribution (36) of Erdős-Rényi graphs, the critical probability is  $p_c \simeq \frac{1}{N}$ .

<sup>7</sup>We should briefly comment on the meaning of the giant component in the directed case. In general directed graphs, there are three different types of giant components: (1) a strongly connected component, in which every vertex can reach every other vertex; (2) a component that contains vertices reachable from (1) but that cannot reach (1); and (3) a component comprised of vertices that can reach (1) but are not reachable from (1). These can be visualized as a “bow-tie” diagram [92]. Slightly above the percolation threshold, only (1) and (2) are present, since loops are suppressed. In directed acyclic graphs, the strongly connected component never arises, as loops are forbidden altogether.



### D. Tail component size distributions

Of prime importance for our analysis is the tail of the in- and out-component size distributions, defined as

$$\begin{aligned}
 P_s &= \text{probability that a randomly chosen vertex} \\
 &\text{has } s \text{ ancestors;} \\
 P_t &= \text{probability that a randomly chosen vertex} \\
 &\text{has } t \text{ descendants.}
 \end{aligned} \tag{52}$$

Their generating functions are respectively  $H_0^{\text{in}}(x)$  and  $H_0^{\text{out}}(y)$ . Close to the percolation threshold, the tail of these distributions takes the form [52]

$$P_s \sim s^{-\tau} e^{-s/s_{\text{max}}}; \quad (s \gg 1), \tag{53}$$

and similarly for  $P_t$ . As usual the correlation length  $s_{\text{max}}$  diverges as we approach the phase transition, e.g.,  $s_{\text{max}} \sim |p - p_c|^{-1/2}$  for Erdős-Rényi graphs. By studying the behavior of the generating functions near the percolation threshold, one can show that the critical exponent takes the value [52]

$$\tau = \frac{3}{2}. \tag{54}$$

(For completeness, we provide a brief proof of this result in Appendix B.) In other words, at criticality the distributions have the universal power-law tail

$$\boxed{P_s \sim s^{-3/2}; \quad P_t \sim t^{-3/2} \quad (s, t \gg 1)}. \tag{55}$$

Equation (55) is a key result for us, as it will play an important role in deducing the probability distribution for the CC in Sec. VII.

The typical size  $s_*$  of the giant component can be estimated as the value at which the probability is  $1/N$ . In other words,  $P_s(s_*) \sim s_*^{-3/2} \sim 1/N$ , and similarly for  $P_t$ . Thus, at criticality the giant in- and out-components are both of size

$$s_* \sim N^{2/3}. \tag{56}$$

All other components have size  $\mathcal{O}(\log N)$ . This holds for all random graphs with finite  $z$  and  $z_2$ .

### E. Scale-free networks

A class of networks that has attracted much attention in the last two decades are scale-free networks [93,94]. Empirically, many real-world networks exhibit this property, including the World Wide Web, social/collaboration networks and metabolic networks. Scale-free graphs include ‘‘hubs’’—nodes connected to a very large number

of other nodes. We briefly review the percolation structure on such networks, focusing for simplicity on undirected graphs, and refer the reader to [95,96] for details.

Scale-free networks are characterized by a degree distribution with power-law tail

$$p_k \sim k^{-\gamma}. \tag{57}$$

We require  $\gamma > 2$  in order for the distribution to be normalizable and have finite mean. If  $\gamma > 3$ , such that the variance is also finite, then the percolation structure belongs to the Erdős-Rényi universality class. So the interesting regime is

$$2 < \gamma < 3. \tag{58}$$

A concrete example is

$$p_k = \begin{cases} 1 - z \frac{\zeta(\gamma)}{\zeta(\gamma-1)} & k = 0; \\ \frac{z}{\zeta(\gamma-1)} k^{-\gamma} & k \geq 1, \end{cases} \tag{59}$$

where  $\zeta$  is the Riemann zeta function. The fraction of nodes with  $k = 0$  ensures that the distribution is normalized and has mean degree  $z$ . Like Erdős-Rényi graphs, percolation occurs when  $z_c = 1$  [93,94].

At percolation criticality, the distribution of component sizes also exhibits a power-law tail, but with a different critical exponent:

$$P_s \sim s^{-\frac{\gamma}{\gamma-1}}. \tag{60}$$

(Notice that the power matches (54) as  $\gamma \rightarrow 3$ .) Thus each value of  $\gamma$  in the range (58) defines its own universality class, comprised of all degree distributions with a scale-free tail with this particular power. The typical size of the giant component is estimated as before by setting  $P_s(s_*) \sim s_*^{-\frac{\gamma}{\gamma-1}} \sim 1/N$ , which gives

$$s_* \sim N^{\frac{\gamma-1}{\gamma}}. \tag{61}$$

## VI. DIRECTED PERCOLATION IN ETERNAL INFLATION

Lacking detailed knowledge of the underlying string landscape, it is reasonable to model a fiducial landscape region as a random network. With the quantitative results of the previous section at hand, let us briefly recap the approximations underlying the mapping to directed percolation.

- (1) The downward approximation, in which upward transitions are neglected to leading order. Strictly speaking, the downward approximation requires us to study directed acyclic graphs, i.e., without directed loops. However, as argued earlier, at low connectivity the density of cycles is suppressed by

$1/N$ , hence directed random graphs offer a reasonable approximation.

- (2) The dominant decay channel approximation, in which the branching ratio  $T_{Ij}$  is either 0 or 1. This relies on semiclassical transition rates in field theory being exponentially staggered, and therefore generically dominated by a single decay channel.

Let us stress that these approximations are made for convenience, to simplify the problem. It is in principle straightforward to generalize our analysis by relaxing them. For instance, if a landscape region includes vacua that are nearly degenerate, such that the downward approximation is invalid, then the corresponding links would be bidirected. The problem of directed percolation with a finite fraction of bidirected edges was studied in [97], where it was shown that bi-directed edges act as a catalyst for directed percolation. Similarly, if subdominant transitions are not completely negligible, such that the dominant decay approximation is invalid, then the corresponding network would be a random weighted graph [98].

### A. Generating functions

With these provisos in mind, consider a landscape region with  $N_{\text{dS}}$  transient nodes (dS vacua) and  $N_{\text{AdS}}$  terminal nodes. Although the latter also include dS vacua with only upward decay channels, as well as Minkowski vacua, we use the collective ‘‘AdS’’ subscript for simplicity.

The moment generating function (27) can be written as

$$\mathcal{G}(x, y) = \frac{N_{\text{dS}}}{N} \mathcal{G}^{\text{dS}}(x, y) + \frac{N_{\text{AdS}}}{N} \mathcal{G}^{\text{AdS}}(x, y), \quad (62)$$

with  $N = N_{\text{dS}} + N_{\text{AdS}}$ . Since terminals by definition have vanishing out-degree, we have

$$\mathcal{G}^{\text{AdS}}(x, y) = \sum_k P_{0k}^{\text{AdS}} x^j = F_0^{\text{AdS}}(x), \quad (63)$$

where we have used (30). Furthermore, since transients have out-degree 1 in the dominant decay channel approximation, as depicted in Fig. 2, we should set  $\mathcal{G}^{\text{dS}}(x, y) = y \sum_k P_{1k}^{\text{dS}} x^j = y F_0^{\text{dS}}(x)$ . However, we will proceed more generally for now, and specialize to  $z_{\text{out}}^{\text{dS}} \simeq 1$  at the end of the calculation.

The generating functions (30) for the marginalized in- and out-degree distributions of a randomly chosen node are given by

$$\begin{aligned} G_0(y) &= \frac{N_{\text{dS}}}{N} G_0^{\text{dS}}(y) + \frac{N_{\text{AdS}}}{N}; \\ F_0(x) &= \frac{N_{\text{dS}}}{N} F_0^{\text{dS}}(x) + \frac{N_{\text{AdS}}}{N} F_0^{\text{AdS}}(x). \end{aligned} \quad (64)$$

The condition (29) for edge conservation gives

$$z = \frac{N_{\text{dS}}}{N} z_{\text{in}}^{\text{dS}} + \frac{N_{\text{AdS}}}{N} z_{\text{in}}^{\text{AdS}} = \frac{N_{\text{dS}}}{N} z_{\text{out}}^{\text{dS}}. \quad (65)$$

Next, the out-degree distribution for a 1st descendant, given by (32), amounts to weighing by the number of edges:

$$G_1(y) = \frac{1}{z_{\text{out}}^{\text{dS}}} \left( \frac{N_{\text{AdS}}}{N_{\text{dS}}} z_{\text{in}}^{\text{AdS}} + \frac{\partial \mathcal{G}^{\text{dS}}(x, y)}{\partial x} \Big|_{x=1} \right). \quad (66)$$

Similarly, the in-degree distribution for a 1st ancestor, given by (34), reduces to

$$F_1(x) = \frac{1}{z_{\text{out}}^{\text{dS}}} \frac{\partial \mathcal{G}^{\text{dS}}(x, y)}{\partial y} \Big|_{y=1}. \quad (67)$$

The number of 2nd ancestors of a given vertex is generated by  $\frac{N_{\text{dS}}}{N} F_0^{\text{dS}}(F_1(x)) + \frac{N_{\text{AdS}}}{N} F_0^{\text{AdS}}(F_1(x))$ . In particular, the average number of 2nd ancestors, which equals the average number of 2nd descendants, is

$$z_2 = \frac{N_{\text{dS}}}{N} \frac{\partial^2 \mathcal{G}^{\text{dS}}}{\partial x \partial y} \Big|_{x=y=1}. \quad (68)$$

### B. Percolation phase transition

The derivation of the directed percolation phase transition given in Sec. V C follows identically in the case of interest. For instance, the generating functions  $H_1^{\text{out}}(y)$  and  $H_0^{\text{out}}(y)$  for the number of descendants satisfy the same implicit relations (40) and (42):

$$H_1^{\text{out}}(y) = y G_1(H_1^{\text{out}}(y)); \quad H_0^{\text{out}}(y) = y G_0(H_1^{\text{out}}(y)), \quad (69)$$

with  $G_0$  and  $G_1$  respectively given by (64) and (66).

From (48), the directed percolation phase transition occurs when  $z_2 = z$ . Using (65) and (68), this means

$$\begin{aligned} \frac{\partial^2 \mathcal{G}^{\text{dS}}}{\partial x \partial y} \Big|_{x=y=1} &= z_{\text{in}}^{\text{dS}} + \frac{N_{\text{AdS}}}{N_{\text{dS}}} z_{\text{in}}^{\text{AdS}} \\ &= z_{\text{out}}^{\text{dS}} \quad (\text{directed percolation}). \end{aligned} \quad (70)$$

As mentioned earlier, consistent with the dominant decay channel approximation we should set  $\mathcal{G}^{\text{dS}}(x, y) = y F_0^{\text{dS}}(x)$ , such that the out-degree of transient vacua is precisely 1. To see how percolation works out, it is instructive to keep things slightly more general by assuming that transients have independent in- and out-degree distributions:

$$\mathcal{G}^{\text{dS}}(x, y) = G_0^{\text{dS}}(y) F_0^{\text{dS}}(x). \quad (71)$$

In this case the percolation condition (70) reduces to

$$z_{\text{in}}^{\text{dS}} = 1. \quad (72)$$

Equivalently, from (65),

$$\boxed{z_{\text{out}}^{\text{dS}} = 1 + \frac{N_{\text{AdS}}}{N_{\text{dS}}} z_{\text{in}}^{\text{AdS}}}. \quad (73)$$

Equation (73) is a key result of our analysis. From the point of view of dS vacua, the presence of terminals pushes the percolation threshold above unity, i.e.,  $z_{\text{out}}^{\text{dS}} > 1$ . This makes sense intuitively, as absorbing nodes inhibit the emergence of a giant component. On the other hand, the dominant decay channel approximation tells us that  $z_{\text{out}}^{\text{dS}} \simeq 1$ . Therefore, in order for a landscape region to be near percolation criticality, it must satisfy

$$N_{\text{AdS}} z_{\text{in}}^{\text{AdS}} \ll N_{\text{dS}}. \quad (74)$$

This is the situation shown in the right panel of Fig. 1, wherein dS vacua decay primarily to other transients, and the region includes a giant funnel of size

$$s_{\star} \sim \begin{cases} N_{\text{dS}}^{2/3} & \text{ER class;} \\ N_{\text{dS}}^{\frac{\gamma-1}{\gamma}} & \text{scale-free, } 2 < \gamma < 3. \end{cases} \quad (75)$$

In contrast, if a significant fraction of dS vacua decay into terminals, such that  $N_{\text{AdS}} z_{\text{in}}^{\text{AdS}}$  is comparable to  $N_{\text{dS}}$ , then the landscape region will be subcritical, as shown in the left panel of Fig. 1.

As argued in (25), in the downward and dominant decay channel approximations, the probability to occupy a node is proportional to the number of its ancestors:  $P(I) \sim s_I$ . In other words, vacua with high occupation probability have large number of ancestors. The probability that a randomly chosen vacuum has  $s$  ancestors is precisely given by  $P_s$ , defined in (52). For subcritical regions, the tail distribution of the component is exponentially cut off, as in (53), and therefore vacua in such regions typically have  $s \sim \mathcal{O}(1)$ . For near-critical regions, however,  $P_s$  displays a power-law tail, given by  $P_s \sim s^{-3/2}$  for the ER universality class, and  $\sim s^{-\frac{\gamma}{\gamma-1}}$  with  $2 < \gamma < 3$  for the scale-free graphs [see (60)]. Correspondingly, near-critical regions include nodes whose number of ancestors is of order the size of the giant component, i.e.,  $s \sim s_{\star}$ .

Thus we arrive at an important realization. To the extent that landscape regions can be modeled as random networks, as we have done, we conclude that *vacua with the highest occupational probability reside in landscape regions that are close to directed percolation criticality*. As usual, near the percolation phase transition, various observables assume power-law (scale-invariant) probability distributions, characterized by universal critical exponents that are insensitive to the microscopic details of the system. As we are about to show, the critical exponent for  $P_s$  translates to a critical exponent for the CC distribution.

## VII. CRITICAL EXPONENT FOR THE COSMOLOGICAL CONSTANT

Let  $f_V(v)$  denote the underlying CC probability distribution function on the landscape, where  $v = \Lambda/M_{\text{Pl}}^4$  is the dimensionless CC. In what follows we will keep  $f_V(v)$  completely general, except for one assumption made at the end, namely that the distribution is smooth as  $v \rightarrow 0^+$ , such that

$$F_V(v) \simeq f_V(0)v; \quad \text{for } 0 < v \ll 1, \quad (76)$$

where  $F_V(v)$  is the cumulative distribution function.<sup>8</sup>

Our task is to derive a probability distribution  $P(v)$  that takes into account the measure factor from cosmological dynamics. For this purpose, we focus on landscape regions close to directed percolation criticality. As argued above, such regions include vacua whose number of ancestors are of order the size of the giant component, i.e.,  $s \sim s_{\star}$ , and therefore have very high probability. Furthermore, since all but one vacuum in the giant component (in any connected component, for that matter) is a terminal, we are justified in focusing on  $v > 0$  to deduce the CC distribution. (This is an obvious consequence of the downward and dominant decay channel approximations.) In other words, vacua in the giant component are overwhelmingly more likely to be dS vacua than AdS. For concreteness we focus on the Erdős-Rényi universality class, and briefly discuss the generalization to scale-free graphs at the end.

Consider a vertex in such a region, and suppose that this vertex has  $s$  ancestors and  $t$  descendants. If the vertex in question has vacuum energy  $v$ , then in the downward approximation its  $s$  ancestors all have larger vacuum energy, while its  $t$  descendants all have lower vacuum energy. In other words, the conditional CC probability distribution  $P(v|s, t)$  follows ordered statistics:

$$P(v|s, t) = \frac{(s+t+1)!}{t!s!} f_V(v) (F_V(v))^t (1-F_V(v))^s. \quad (77)$$

It is convenient to use the cumulative distribution itself as the random variable,  $U \equiv F_V$ , such that

$$du = f_V(v)dv. \quad (78)$$

Clearly  $f_U(u)$  is uniform over  $u \in [0, 1]$ . Indeed, the ordered statistics of  $U$  are simply those of the uniform distribution:

$$\begin{aligned} P(u|s, t) &= \frac{(s+t+1)!}{t!s!} u^t (1-u)^s \\ &= \text{Beta}(t+1, s+1)(u). \end{aligned} \quad (79)$$

<sup>8</sup>Some authors [99–104] have argued that the underlying distribution  $f_V(v)$  for the landscape diverges as  $v \rightarrow 0^+$ . If this is the case, this would make our critical exponent for the CC even more negative.

The desired CC probability distribution is obtained by marginalizing over  $s$  and  $t$ ,

$$P(u) = \sum_{s,t} f_S(s) f_T(t) \text{Beta}(t+1, s+1)(u), \quad (80)$$

where we have used the fact that  $s$  and  $t$  are independent random variables, even for correlated degree distributions [97].

The probability distribution  $f_S(s)$  is given by

$$f_S(s) \sim s P_s, \quad (81)$$

where  $P_s$  is the probability that a randomly chosen vertex has  $s$  ancestors, defined in (52). The factor of  $s$  is the cosmological measure factor (25), which encodes the fact that the probability to pick a node is proportional to the number of ancestors. Since  $P_s \sim s^{-3/2}$  in the tail at criticality,  $f_S(s) \sim s^{-1/2}$  is not normalizable for  $N_{\text{ds}} \rightarrow \infty$ . For the realistic case of finite  $N_{\text{ds}}$ , however, it is regularized by the giant component of size  $s_* \sim N_{\text{ds}}^{2/3}$ . It follows that  $f_S(s)$  has the power-law behavior

$$f_S(s) \simeq \frac{1}{2\sqrt{s_*}\sqrt{s}}; \quad s_* \geq s \gg 1. \quad (82)$$

Meanwhile,  $f_T(t)$  is just the probability distribution  $P_t$  that a randomly chosen vertex has  $t$  descendants. All we will need is its power-law tail behavior (55):

$$f_T(t) = P_t \sim \frac{1}{t^{3/2}}; \quad t \gg 1. \quad (83)$$

### A. Analytic approximation for $P(u)$

We proceed to evaluate (80) analytically. We assume throughout that  $u \ll 1$ , and this will be justified *a posteriori* since the resulting distribution will peak for  $u \ll 1$ . First, let us suppose that the integral is dominated by the tail region  $s, t \gg 1$ , such that  $f_S(s)$  and  $f_T(t)$  are given by (82) and (83). In this regime, the Beta distribution is well-approximated by a Gaussian, and the sums can be approximated as integrals:

$$P(u) \sim \int \frac{ds}{2\sqrt{s_*}\sqrt{s}} \int \frac{dt}{t^{3/2}} \frac{(s+t)^{3/2}}{\sqrt{2\pi st}} \exp\left[-\frac{s+t}{2st}(t-su)^2\right]. \quad (84)$$

The  $t$  integral can be evaluated using Laplace's method. The exponent is stationary for  $t_0 = su$ . Consistency of the tail approximation requires  $t_0 \gg 1$ , and in particular

$$s_* u \gg 1. \quad (85)$$

Evaluating the  $t$  integral, we obtain

$$\begin{aligned} P(u) &\sim \frac{1}{2\sqrt{s_*}} u^{-3/2} \int_{1/u}^{s_*} \frac{ds}{s} \\ &= \frac{1}{2\sqrt{s_*}} u^{-3/2} \ln(s_* u); \quad s_*^{-1} \ll u \ll 1. \end{aligned} \quad (86)$$

This peaks for small  $u$ , as anticipated.

To see that the distribution is well-behaved as  $u \rightarrow 0$ , consider the regime  $s_* u \ll 1$ . In this case we can make the approximation  $(1-u)^s \simeq 1$  for all  $s$ , such that (80) becomes

$$P(u) \simeq \sum_{s,t} f_S(s) f_T(t) \frac{(t+s+1)!}{s!t!} u^t. \quad (87)$$

To proceed, let us assume that  $s \gg t$ , which can be justified *a posteriori*. Thus we obtain

$$P(u) \simeq \sum_{s,t} f_S(s) s f_T(t) \frac{(sv)^t}{t!}. \quad (88)$$

Since  $su \ll 1$ , clearly the sum peaks for  $t = 0$ , in which case

$$P(u) \simeq f_T(0) \sum_s f_S(s) s \simeq \frac{1}{3} f_T(0) s_*; \quad u \ll s_*^{-1}. \quad (89)$$

Thus, as claimed the distribution is smooth for  $u \rightarrow 0$ , and reaches a maximum value of  $\sim s_*$  at the origin.

### B. The CC distribution

Equations (86) and (89) give the general probability distribution for the cumulative variable  $u = F_V(v)$  of a general distribution  $f_V(v)$ . To translate to a probability distribution  $P(v)$  for  $v$  itself, we now make the assumption (76) that  $f_V(v)$  is smooth as  $v \rightarrow 0^+$ , such that

$$u \simeq f_V(0)v. \quad (90)$$

In the ‘‘tail,’’ defined by  $f_V^{-1}(0)s_*^{-1} \ll v \ll \min(f_V^{-1}(0), 1)$ , the CC probability distribution is a power-law

$$P(v) \sim \frac{1}{2\sqrt{f_V(0)s_*}} v^{-3/2} \ln(s_* f_V(0)v). \quad (91)$$

Thus the  $-3/2$  critical exponent for  $s$  and  $t$  translates to an identical critical exponent for  $v$ . In particular, the 95% confidence interval for the CC is

$$v \lesssim \frac{\log 20}{f_V(0)s_*} \sim N_{\text{ds}}^{-2/3}. \quad (92)$$



Therefore the cosmological measure favors small, positive vacuum energy. It can explain our observed CC  $v_{\text{obs}} \sim 10^{-120}$  if our vacuum belongs to a funneled region of size  $N_{\text{dS}} \sim 10^{240}$ .

To be clear  $N_{\text{dS}}$  is the number of dS (transient) vacua in a funnel region near directed percolation criticality, not the total number of dS vacua across the entire landscape. Since the cosmological measure favors vacua with the largest number of ancestors, we are likely to inhabit the largest funnel region near percolation criticality, i.e., the near-critical region with largest  $N_{\text{dS}}$ .

Unfortunately, it is unlikely that the Erdős-Rényi CC distribution derived above offers a viable solution to the CC problem. The reason is that it likely suffers from an empty universe problem. With a Poisson distribution, a given node typically only has a few parents. Their CC distribution is also given by (91), hence they are themselves likely to have a small CC. Since the vacuum energy of the parents sets an upper bound on the energy scale of the last period of slow-roll inflation, the generic outcome is an empty universe. A possible way out is if low-lying nodes have a large number of parents. This requires many other nodes in the giant component to be orphans (i.e., have vanishing in-degree), in order to maintain the criticality condition  $z_{\text{in}}^{\text{dS}} = 1$ .

Scale-free graphs offer a natural realization of this scenario. As discussed in Sec. V E, scale-free graphs include hubs, whose in-degree is of order the size of the giant component. It is straightforward to repeat the analysis for the scale-free degree distribution. The resulting power-law in this case is

$$P(v) \sim v^{-\frac{\gamma}{\gamma-1}} \quad (2 < \gamma < 3), \quad (93)$$

which is the same critical exponent as  $P_s$  itself, cf. (60). Using (61), the 95% confidence interval is

$$v \lesssim N_{\text{dS}}^{\frac{\gamma-1}{\gamma}}. \quad (94)$$

We will comment further on the empty universe problem and scale-free graphs in the Conclusions.

### VIII. CONCLUSIONS

The near-criticality of our universe may be the strongest empirical hint that we are part of a multiverse. A natural arena to realize this ensemble is the vast energy landscape of string theory, together with the dynamics of eternal inflation to instantiate in space-time the different vacua of the landscape. Making robust statistical predictions for physical observables in our own universe is unquestionably a task of fundamental importance in theoretical physics. Yet, how can we ever hope to make progress toward this goal without a detailed understanding of the string landscape?

Fortunately, despite all of its conceptual pitfalls, eternal inflation boils down to a random walk on the network of

vacua. Technically this is an absorbing Markov process, because of terminal (AdS/Minkowski) vacua which act as sinks. Hence the dynamics are inherently nonequilibrium. The Markov process leads to a natural definition of probabilities as occupational probabilities for the random walk.

In this paper we showed how the Markov process governing vacuum dynamics can be mapped naturally to a problem of directed percolation on the network of vacua. The mapping relies on two very general and well-justified approximations for transition rates: (1) the downward approximation, which neglects ‘‘upward’’ transitions, as these are generally exponentially suppressed; and (2) the dominant decay channel approximation, which capitalizes on the fact that tunneling rates are exponentially staggered.

With these simplifying assumptions, we argued that the uniform-in-time probabilities reduce a simple and intuitive observable in directed graphs. Namely, the probability to occupy a particular node is proportional to the number of its ancestors, i.e., how many other nodes can reach it through a sequence of directed (downward) transitions. Thus the probabilities favor vacua with a large basin of ancestors, lying at the bottom of a deep funnel. Funneled landscape topography appears to be a common solution to optimization on complex energy landscapes, including protein folding [45], atomic clusters [46–48], deep learning [49], and combinatorial optimization [50].

Lacking detailed knowledge of the string landscape, we modeled the network of vacua as random graphs with arbitrary degree distributions, including Erdős-Rényi and scale-free graphs. As a complementary approach, we also modeled regions of the landscape as regular lattices, specifically Bethe lattices. Despite representing extreme opposites of graph regularity, Bethe lattices and Erdős-Rényi belong to the same percolation universality class. Thus one may hope that the lessons drawn from studying percolation in these simplified, idealized setups carry over to more realistic landscapes.

The most important result of our analysis is that the uniform-in-time probabilities favor regions of the landscape poised at the directed percolation phase transition. In other words, our vacuum most likely resides within a network of vacua tuned at directed percolation criticality. As usual, the predictive power of criticality lies in universality. This raises the tantalizing prospect of deriving statistical predictions for physical observables that are insensitive to the details of the underlying landscape. More broadly, it suggests a deep and powerful relation between phase transitions in landscape dynamics and the inferred near-criticality of our universe.

To illustrate this point, we derived a probability distribution for the CC. At percolation criticality, the probability distributions for the number of ancestors and descendants of a given node both display power-law tails, with certain critical exponents. Assuming only that the underlying CC distribution function is smooth near the origin, we derived

probability distributions for the CC that are also power-law, with critical exponents that are determined by the universality class (Erdős-Rényi or scale-free). These distributions favor small, positive CC, and can account for the observed CC if our vacuum belongs to a large enough region. In fact, since the measure favors vacua with the largest number of ancestors, we are likely to inhabit the largest funnel region near percolation criticality.

There are many future directions of inquiry worth pursuing. Let us mention two concrete follow-ups:

- (i) It is possible to derive probability distributions for other physical observables. To give one example, consider a node of in-degree  $k$ . The joint probability distribution for its  $k$  parents to have  $s_1, s_2, \dots, s_k$  ancestors,  $P(s_1, \dots, s_k|k)$ , translates to a joint distribution for the potential energy of its parents, given by  $P(v_1, \dots, v_k|k)$ . This statistics informs us on the energy scale of the last period of slow-roll inflation, which of course has immediate bearing on the observational prospects of detecting primordial gravitational waves. This also has immediate bearing on the potential empty universe problem discussed in Sec. VII. For instance, one could condition on sufficiently large  $k$ , i.e., large number of parents.
- (ii) The connection with other complex energy landscapes deserves further exploration. A remarkable aspect of protein folding networks is that they are scale-free, with the native state acting as a hub with very large degree [105]. Protein folding networks also display the small-world property, and are hierarchical. Similar properties are found in atomic clusters with Leonard-Jones interactions [46–48]—their funnel topography is hierarchical (funnels nested within larger funnels), and the degree distribution is scale-free. It may be that such properties are generic to optimization on complex energy landscapes. It will be fascinating to explore their implications in the context of landscape dynamics.

### ACKNOWLEDGMENTS

We thank Giorgos Gounaris, James Halverson, Eleni Katifori, Cody Long, Minsu Park, Anushrut Sharma, and Nathaniel Watkins for helpful discussions. This work is supported in part by the U.S. Department of Energy (HEP) Award No. DE-SC0013528.

### APPENDIX A: PERCOLATION ON UNDIRECTED RANDOM GRAPHS

In this appendix, we give a review of percolation on undirected random graphs, for completeness. An undirected random network is defined by specifying a degree probability distribution:

$$p_k = \text{probability of a randomly chosen node having degree } k. \quad (\text{A1})$$

The moment generating function is

$$G_0(x) = \sum_{k=0}^{\infty} p_k x^k, \quad (\text{A2})$$

with normalization condition  $G_0(1) = \sum_{k=0}^{\infty} p_k = 1$ . Its derivatives give as usual the moments of the distribution, such as the average degree:

$$z \equiv \langle k \rangle = \sum_{k=0}^{\infty} k p_k = G'_0(1). \quad (\text{A3})$$

The next important quantity is the degree distribution of a vertex reached by following a randomly chosen edge. This distribution is not simply  $p_k$ , since we are  $k$  times more likely to arrive at a vertex with degree  $k$  as we are at a vertex of degree 1. Therefore this distribution is proportional to  $k p_k$ . Now, suppose we start from a randomly chosen vertex, and follow each of its links to reach the 1st neighbors. We are interested in the “excess degree” of the 1st neighbors, which excludes the edge we arrived along. Let

$$q_k = \text{probability that 1st neighbor has excess degree } k. \quad (\text{A4})$$

This is shown in Fig. 7. This distribution is related to  $p_k$  by  $q_k = \frac{(k+1)p_{k+1}}{\sum_k k p_k} = \frac{(k+1)p_{k+1}}{z}$ , and its generating function is given by

$$G_1(x) \equiv \sum_{k=0}^{\infty} q_k x^k = \frac{G'_0(x)}{z}. \quad (\text{A5})$$

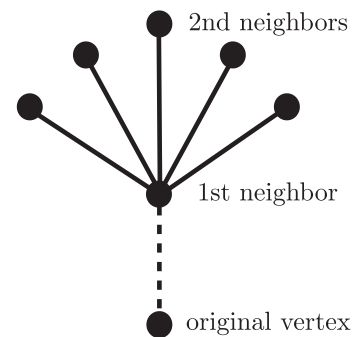


FIG. 7. Starting from a randomly chosen (“original”) vertex, we follow of its links (dashed line) to a 1st neighbor. The probability that this 1st neighbor has  $k$  edges excluding the one we followed is  $q_k$ , with generating function  $G_1(x)$  given by (A5).

Now, if the original vertex has degree  $k$ , then the number of second-nearest neighbors is generated by  $(G_1(x))^k$ . This ignores loops, which are negligible in the large  $N$  limit below the percolation threshold, as argued below. It follows that number of 2nd neighbors is generated by  $\sum_{k=0}^{\infty} p_k (G_1(x))^k = G_0(G_1(x))$ . For instance, the average number of 2nd neighbors is

$$z_2 = \left. \frac{d}{dx} G_0(G_1(x)) \right|_{x=1} = G_0''(1), \quad (\text{A6})$$

where we have used (A3) and  $G_1(1) = 1$ .

### 1. Erdős-Rényi example

An important example is the Erdős-Rényi graph [51]. An undirected Erdős-Rényi graph is a random graph with  $N$  vertices, in which an edge between any two distinct vertices has a probability  $p$  of being included. Therefore the probability distribution  $p_k$  is binomial and tends to the Poisson distribution for large  $N$ , as given by (36). The generating functions (A2) and (A5) in this case are given by

$$G_0(x) = G_1(x) = e^{z(x-1)}. \quad (\text{A7})$$

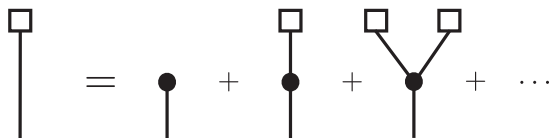
The average number of 2nd neighbors (A6) is  $p^2(N-1)(N-2) \simeq p^2 N^2$ , and therefore

$$z_2 \simeq z^2 \quad (\text{Poisson}). \quad (\text{A8})$$

To see that the probability of having loops is suppressed by  $\frac{1}{N}$  below or close to the percolation threshold, consider the probability of having two 1st neighbors connected to each other. For a random vertex with degree  $k$  there are  $\frac{k(k-1)}{2}$  possible edges between any pair of its neighbors. The chance of having at least one loop is  $1 - (1-p)^{\frac{k(k-1)}{2}}$ . In Erdős-Rényi graphs,  $p \simeq \frac{z}{N}$ , thus the chance of having triangular loops is  $\simeq \frac{zk(k-1)}{2N} \sim \frac{1}{N}$  in the large  $N$  limit.

### 2. Component size distribution

With these tools at hand, we next consider the size distribution of connected components. For this purpose we work below (or close to) the percolation phase transition. We define



(a)  $H_1(x)$

$H_1(x) =$  Gen. fcn for size of components reached  
by following a randomly chosen edge. (A9)

Ignoring loops, as they are suppressed by  $\frac{1}{N}$  close to the percolation threshold, we have the treelike structure depicted in Fig. 8(a). From the definition of  $q_k$ , the treelike structure implies the consistency condition

$$H_1(x) = x \sum_{k=0}^{\infty} q_k (H_1(x))^k = x G_1(H_1(x)), \quad (\text{A10})$$

where the last step follows from  $q_k$  being generated by  $G_1$ . Similarly, we define the distribution of (finite) component sizes reached starting from a randomly chosen vertex:

$P_s =$  probability that (finite) component reached by  
following a randomly chosen vertex has size  $s$ . (A11)

We denote by  $H_0(x)$  the corresponding generating function:  $H_0(x) = \sum_{s=0}^{\infty} P_s x^s$ . It satisfies the consistency relation, depicted in Fig. 8(b),

$$H_0(x) = x \sum_{k=0}^{\infty} p_k (H_1(x))^k = x G_0(H_1(x)), \quad (\text{A12})$$

where the last step follows from  $p_k$  being generated by  $G_0$ . By definition,  $H_0$  describes finite components, i.e., it excludes the giant component. As long we are below the percolation threshold, such that there is no infinite cluster, then  $H_0(1) = 1$ . Above the percolation threshold,  $H_0(1)$  gives the fraction of the vertices that do not belong to the giant component. To be precise, let  $P(z)$  denote the fraction of vertices belonging to the giant component:

$$P(z) = 1 - H_0(1). \quad (\text{A13})$$

### 3. Percolation phase transition

For general random graphs, it is often difficult in practice to solve (A10) analytically, hence we must content ourselves with computing the moments of the size distribution, such as the average size of (finite) connected components  $S(z)$ . First let us work below the phase transition, such that



(b)  $H_0(x)$

FIG. 8. (a) Treelike structure satisfied by the generating function  $H_1$  for the distribution of component sizes starting from a randomly chosen edge; (b) Same structure, but for the generating function  $H_0$  for the component sizes starting from a randomly chosen vertex.

there is no giant component, and  $H_0(1) = H_1(1) = 1$ . Using (A12), we obtain

$$S(z) = H'_0(1) = 1 + G'_0(1)H'_1(1). \quad (\text{A14})$$

On the other hand, from (A10) we have  $H'_1(1) = 1 + G'_1(1)H'_1(1)$ , which implies  $H'_1(1) = \frac{1}{1-G'_1(1)}$ . Thus we obtain

$$S(z) = 1 + \frac{G'_0(1)}{1-G'_1(1)} = 1 + \frac{z^2}{z-z_2}, \quad (\text{A15})$$

where in the last step we have used (A3) and (A6). It is clear that the percolation threshold where a giant component emerges is given by the condition

$$z_2 = z \quad (\text{percolation}). \quad (\text{A16})$$

[Notice that this is identical to (48), hence the percolation structure matches that of the directed case.] This result holds for any degree distribution  $p_k$  with finite mean and variance. More generally, above the phase transition  $S(z)$  gives the average size of finite clusters, i.e., excluding the giant component. Using (A13), the generalization of (A14) is

$$S(z) = \frac{H'_0(1)}{H_0(1)} = 1 + \frac{zH_1^2(1)}{(1-P(z))[1-G'_1(H_1(1))]}, \quad (\text{A17})$$

where  $H_1(1) = u$  is the probability that there is no infinite cluster going down an edge. It is the smallest solution of  $u = G_1(u)$ , with  $0 \leq u \leq 1$ .

As an explicit example, consider Erdős-Rényi graphs. Combining (A8) and (A16), we recover the classic result that percolation occurs at a critical mean degree  $z_c = 1$ , and corresponding critical probability  $p_c \simeq \frac{1}{N}$ . Using (A7), the implicit equations (A10) and (A12) are readily solved in terms of the Lambert  $\mathcal{W}$  function:

$$H_0(x) = H_1(x) = -\frac{1}{z} \mathcal{W}(-zxe^{-z}) \quad (\text{ER graphs}). \quad (\text{A18})$$

The fraction of vertices belonging to a giant cluster (A13) is given by

$$P(z) = 1 - H_0(1) = 1 + \frac{\mathcal{W}(-ze^{-z})}{z}. \quad (\text{A19})$$

Equivalently,  $P(z)$  satisfies the well-known implicit relation  $P(z) = 1 - \exp(-zP(z))$ . As shown in Fig. 9 (blue curve),  $P(z)$  vanishes identically below percolation, starts growing at the percolation threshold  $z_c = 1$ , and approaches unity for large  $z$ . Using the general expression (A17), the average size of finite clusters in this case is given by

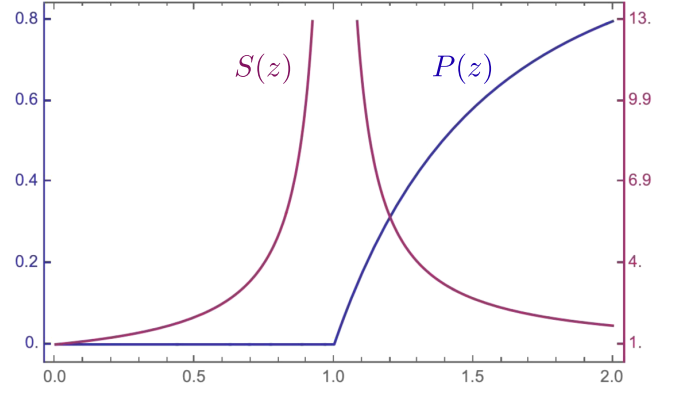


FIG. 9. Percolation criticality for Erdős-Rényi graphs. Blue curve: The fraction of vertices  $P(z)$  belonging to the giant component, given by (A19). Purple curve: The average size of finite clusters,  $S(z)$ , excluding the giant component, given by (A20).

$$S(z) = \frac{1}{1-z(1-P(z))}. \quad (\text{A20})$$

As shown in Fig. 9 (purple curve), this diverges at criticality.

#### 4. Tail component size distribution near percolation threshold

Close to the percolation, the tail of the component size distribution  $P_s$  takes the form (53). Analogously to the directed case, at criticality this distribution has a power-law tail, with the same  $-3/2$  critical exponent [52]:

$$P_s \sim s^{-3/2}. \quad (\text{A21})$$

See Appendix B for a brief proof of this result. A related quantity is  $n_s$ , the probability distribution for the number of clusters of size  $s$ , which also exhibits a power-law tail  $n_s \sim s^{-5/2}$  at criticality. In general,  $n_s$  is related to  $P_s$  by  $P_s = sn_s$ , where the factor of  $s$  accounts for the fact that one is  $s$  times more likely to randomly pick a node belonging to a cluster of size  $s$  than an isolated node. The giant component is of size  $N^{2/3}$  at criticality. All other components have size  $\mathcal{O}(\log N)$ . Above the percolation threshold, the giant component grows to encompass an  $\mathcal{O}(1)$  fraction of all the nodes.

#### APPENDIX B: CRITICAL EXPONENT FOR COMPONENT SIZE DISTRIBUTION

In the main text we claimed that, for undirected random graphs, the distribution  $P_s$  of (finite) component sizes from a randomly chosen vertex develops a power-law tail at criticality. In this appendix, for completeness we briefly outline the proof of this result, closely following [52]. We focus on the Erdős-Rényi universality class, which



includes all degree distributions with finite mean and variance.

Near the percolation threshold, the tail of  $P_s$  takes the form

$$P_s \sim s^{-\tau} e^{-s/s_{\max}}. \quad (\text{B1})$$

As usual, the asymptotic behavior of  $P_s$  is encoded in the behavior of its generating function  $H_0(x)$  near its radius of convergence  $|x_\star|$  [106]. (The radius of convergence is defined as the singularity in  $H_0(x)$  closest to the origin.) Concretely, the correlation length  $s_{\max}$  is given by

$$s_{\max} = \frac{1}{\log |x_\star|}. \quad (\text{B2})$$

Given the implicit relation (A12), and using the fact that the first singularity in  $G_0(x)$  lies outside the unit circle, one can deduce that  $x_\star$  also corresponds to the singularity in  $H_1(x)$  closest to the origin.

With these facts at hand, consider the implicit relation (A10) for  $H_1$ :

$$H_1(x) = xG_1(H_1(x)). \quad (\text{B3})$$

Letting  $w \equiv H_1(x)$ , this can be expressed as

$$x = H_1^{-1}(w) = \frac{w}{G_1(w)}. \quad (\text{B4})$$

The singularity at  $x_\star$  entails that  $\frac{dx}{dH_1}|_{x=x_\star} = 0$ , and therefore  $\frac{dH_1^{-1}(w)}{dw}|_{w=w_\star} = 0$ . Using (B4), this last statement implies

$$G_1(w_\star) = w_\star G_1'(w_\star). \quad (\text{B5})$$

Given a solution  $w_\star$  to (B5), we can solve (B4) to find the corresponding  $x_\star$ :

$$x_\star = \frac{w_\star}{G_1(w_\star)}, \quad (\text{B6})$$

and thus  $s_{\max}$  via (B2). It should be stressed that (B5) need not have a solution. When it does not, then the asymptotic behavior of  $P_s$  is not of the form (B1). We will see, however, that a solution to (B5) exists close to the phase transition.

Indeed, at the percolation threshold, defined by  $G_1'(1) = 1$ , (B5) and (B6) are solved by

$$x_\star = w_\star = 1 \quad \text{at percolation.} \quad (\text{B7})$$

Correspondingly,  $s_{\max} \rightarrow \infty$ . To determine the power-law for  $P_s$  at criticality, let us expand (B4) around  $x_\star = w_\star = 1$ :

$$\begin{aligned} x &= H_1^{-1}(w) = \frac{1 + (w-1)}{G_1(1 + (w-1))} \\ &\simeq 1 - \frac{1}{2} G_1''(1)(w-1)^2 + \dots \end{aligned} \quad (\text{B8})$$

where we have used  $G_1(1) = G_1'(1) = 1$ . The derivation clearly relies on  $G_1''(1)$  being finite. From (A5), this amounts to assuming that the first three moments of the degree distribution are finite. This is the case, in particular, for the Poissonian distribution, but not for scale-free graphs. It follows from (B8) that

$$|1-w| = \sqrt{\frac{2}{|G_1''(1)|}} \sqrt{1-x}. \quad (\text{B9})$$

Therefore, since  $w = H_1(x)$ , we deduce that the singular behavior of  $H_1$ , and therefore that of  $H_0(x)$  as well, near  $x = 1$  is given by

$$H_0(x) \sim H_1(x) \sim (1-x)^\beta; \quad \text{with } \beta = \frac{1}{2}. \quad (\text{B10})$$

The critical exponent  $\beta = 1/2$  is related to the critical exponent  $\tau$  for  $P_s$  as follows. On the one hand, from (B10) we have

$$\beta = 1 + \lim_{x \rightarrow 1} (x-1) \frac{H_0''(x)}{H_0'(x)}. \quad (\text{B11})$$

On the other hand, the asymptotic form (B1) implies that  $H_0(x)$  can be expressed as

$$H_0(x) = \sum_{s=0}^{a-1} P_s x^s + C \sum_{s=a}^{\infty} s^{-\tau} e^{-s/s_{\max}} x^s + \dots \quad (\text{B12})$$

where  $C$  is a constant, and  $a$  is sufficiently large. The singular behavior as  $x \rightarrow 1$  is encoded in the infinite sum. Focusing on this term we have

$$\beta \simeq 1 + \lim_{x \rightarrow 1} \frac{x-1}{x} \frac{\sum_{s=a}^{\infty} s^{2-\tau} x^{s-1}}{\sum_{s=a}^{\infty} s^{1-\tau} x^{s-1}}. \quad (\text{B13})$$

In the limit of large  $a$ , i.e., large  $s$ , the infinite sums can be approximated as integrals:

$$\begin{aligned} \frac{\sum_{s=a}^{\infty} s^{2-\tau} x^{s-1}}{\sum_{s=a}^{\infty} s^{1-\tau} x^{s-1}} &\simeq \frac{\int_a^{\infty} ds s^{2-\tau} x^s}{\int_a^{\infty} ds s^{1-\tau} x^s} \\ &= -\frac{1}{\log x} \frac{\int_{-a \log x}^{\infty} dt t^{2-\tau} e^{-t}}{\int_{-a \log x}^{\infty} dt t^{1-\tau} e^{-t}} \\ &= -\frac{1}{\log x} \frac{\Gamma(3-\tau, -a \log x)}{\Gamma(2-\tau, -a \log x)}. \end{aligned} \quad (\text{B14})$$

Substituting into (B13) and taking the limit  $x \rightarrow 1$ , we obtain

$$\beta = \tau - 1. \quad (\text{B15})$$

Since  $\beta = 1/2$ , it follows that  $\tau = 3/2$ , which proves (54).

### APPENDIX C: BETHE LATTICE AND ITS PERCOLATION STRUCTURE

As a second approach to percolation on the landscape, the underlying graph topology is modeled as a Bethe lattice (or Cayley tree). This approach is appropriate whenever transitions are strictly “local” in field space, i.e., they are non-negligible only between “nearest-neighbor” vacua, while transitions to more distant vacua can be safely ignored.

#### 1. Undirected Bethe lattice

A Bethe lattice is loop-free graph where all vertices have the same degree  $k$ . Thus it is a tree with branching number  $k - 1$ . An example is given in Fig. 10. In the undirected case, there is nothing special about the central vertex, since every vertex can be viewed as the root.

Thanks to the self-similar structure of this graph, the percolation problem is easily solved. As usual, one assigns a probability  $p$  for each edge to be included. The percolation phase transition, defined by the emergence of a giant cluster, occurs at a critical  $p_c$ . More precisely, in the  $N \rightarrow \infty$  limit, the probability  $P(p)$  that a vertex is part of an infinite cluster (i.e., connected to infinitely many vertices) vanishes for  $p < p_c$  and is finite for  $p > p_c$ . Clearly  $P(p)$  satisfies

$$P(p) = 1 - Q^k(p), \quad (\text{C1})$$

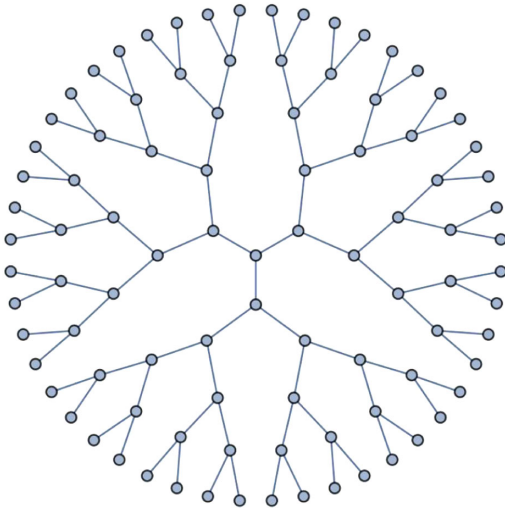


FIG. 10. A Bethe lattice with degree  $k = 3$ .

where  $Q(p)$  is the probability that a neighboring vertex does not connect to infinity on the other side. Due to the similarity of the next layer of neighbor,  $Q(p)$  satisfies the implicit relation,

$$\begin{aligned} Q &= \sum_{\ell=0}^{k-1} \binom{k-1}{\ell} (1-p)^{k-1-\ell} (pQ)^\ell \\ &= (1 - p(1-Q))^{k-1}, \end{aligned} \quad (\text{C2})$$

where the first equality accounts for all combinations of connecting to the next layer.

The critical probability  $p_c$  of percolation is defined such that  $1 > Q(p > p_c) > 0$ , i.e., the boundary of  $p$  in which  $Q(p)$  is nonzero. To obtain an expression for  $p_c$ , let us work near the percolation threshold, where  $1 - Q \ll 1$ . Expanding the implicit relation (C2) for  $1 - Q \ll 1$ , we obtain

$$1 - Q \simeq (k-1)p(1-Q) \Rightarrow p_c = \frac{1}{k-1}. \quad (\text{C3})$$

Thus at percolation criticality each vertex has one connection to the next layer on average. [For  $p < p_c$ , the only solution to (C3), and more generally (C2), is  $Q(p) = 1$ , corresponding to  $P(p) = 0$ .]

The next quantity of interest is the average cluster size,  $S(p)$ , for  $p \leq p_c$ . It satisfies

$$S(p) = 1 + kT(p), \quad (\text{C4})$$

where  $T(p)$  is the expected size of the sub-branch going down an edge. Again, due to the self-similar structure, we have the implicit relation

$$T = p(1 + (k-1)T). \quad (\text{C5})$$

The factor  $p$  accounts for the probability of connection, and  $1 + (k-1)T$  is the vertex itself plus the expected size of the  $k-1$  vertices at the next layer. Solving for  $T$  and substituting (C3) gives

$$T(p) = \frac{p_c p}{p_c - p}. \quad (\text{C6})$$

Substituting into (C4) gives

$$S(p) = \frac{p_c(1+p)}{p_c - p} \quad (p \leq p_c). \quad (\text{C7})$$

As expected, the average cluster size diverges as  $p \rightarrow p_c$  from below. The probability  $P(p)$  and the average size of cluster  $S(p)$  are plotted as a function of  $p$  in Fig. 11.

Near the percolation threshold, the tail of the component size distribution  $P_s$  takes the same form as (B1).

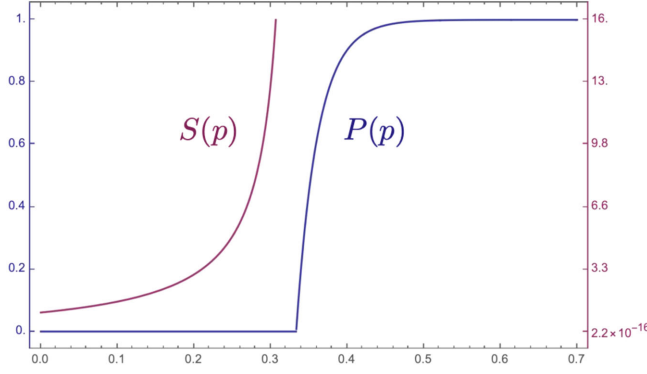


FIG. 11. The probability  $P(p)$  for a vertex to belong to an infinite cluster (blue) and the average cluster size  $S(p)$  (orange) as a function of  $p$  for an undirected Bethe lattice with degree  $k = 4$ , corresponding to  $p_c = 1/3$ .

Remarkably, it can be shown that it has the same critical exponent as Erdős-Rényi graphs,

$$P_s \sim s^{-3/2}, \quad (\text{C8})$$

for any  $k$  [90,91]. To show this, one starts with the quantity

$$n_s(p): \text{ the number of } s\text{-size cluster per site,} \quad (\text{C9})$$

such that  $sn_s(p)$  is the fraction of sites belonging to an  $s$ -size cluster. In general, the chance of forming an  $s$ -size cluster is

$$n(s, p) = g_{s,\alpha} p^{s-1} (1-p)^\alpha, \quad (\text{C10})$$

where  $p^{s-1}$  accounts for  $s-1$  connected edges,  $(1-p)^\alpha$  accounts for  $\alpha$  disconnected edges at the “boundary” of the cluster, and  $g_{s,\alpha}$  is a combinatoric factor. One can easily figure out that  $\alpha = s(k-1) + 2$ , therefore

$$\begin{aligned} \frac{n(s, p)}{n(s, p_c)} &= \frac{p_c}{p} \left( \frac{1-p}{1-p_c} \right)^2 \\ &\times \exp \left[ s \ln \left( (k-1)p \left( \frac{(k-1)(1-p)}{k-2} \right)^{k-2} \right) \right] \\ &\sim \frac{p_c}{p} \left( \frac{1-p}{1-p_c} \right)^2 \exp \left( -\frac{s}{\ell(p)} \right), \end{aligned} \quad (\text{C11})$$

where we only focus on the proximity of the percolation threshold, and

$$\frac{1}{\ell(p)} = \frac{1}{2p_c^2(1-p_c)} (p-p_c)^2 + \mathcal{O}((p-p_c)^3). \quad (\text{C12})$$

Furthermore we can approximate that, for large  $s$  and  $k > 2$ ,

$$n(s, p) \propto s^{-\sigma} \exp \left( -\frac{s}{\ell(p)} \right). \quad (\text{C13})$$

Therefore the average cluster size  $S(p)$ , can be approximated as

$$\begin{aligned} S(p) &= \sum_s s^2 n(s, p) \approx \sum_s s^{2-\sigma} \exp \left( -\frac{s}{\ell(p)} \right) \\ &\simeq \int_1^\infty ds s^{2-\sigma} \exp \left( -\frac{s}{\ell(p)} \right) \\ &\simeq (\ell(p))^{3-\sigma} \Gamma(3-\sigma) \\ &\propto (p-p_c)^{2\sigma-6}. \end{aligned} \quad (\text{C14})$$

At the same time we know that  $S(p) \propto (p_c - p)^{-1}$  from (C7), therefore

$$\sigma = \frac{5}{2}. \quad (\text{C15})$$

By definition, the chance of a randomly chosen site belonging to an  $s$ -size cluster is then,

$$P_s(p) = sn(s, p) \propto s^{-3/2} \exp \left( -\frac{s}{\ell(p)} \right). \quad (\text{C16})$$

At criticality,  $\ell(p)$  diverges per (C12), and we recover (C8).

## 2. Directed Bethe lattice

In the situation of interest, the downward approximation implies that the graph is a *directed Bethe lattice*. This reflects the fact that low-lying (high-lying) vacua have mostly incoming (outgoing) edges. Choosing a fixed vertex as the center, we denote by  $r$  the probability that an edge is directed outward from this center, and by  $s = 1-r$  the probability of the edge being directed inward. Thus the probability of including an inward edge is  $pr$ . For general  $r$ , vertices are no longer indistinguishable since the chosen center vertex is indeed the center. (The exception is  $r = \frac{1}{2}$ , where every vertex is still statistically the same.) Nevertheless, away from the center vertex, layers after layers still have the same self-similarity structure.

Certain properties are insensitive to  $r$ , such as the emergence of the giant component. One can ask about the probability for the existence of a directed path from the center vertex to infinity. The derivation of  $P(p)$  in the undirected case applies just as well to the directed case with the replacement  $p \rightarrow pr$ . Therefore the probability for the existence of a path *to* infinity and the probability for a path *from* infinity are respectively given by

$$\begin{aligned} \text{a path to } \infty &: P(pr); \\ \text{a path from } \infty &: P(ps). \end{aligned} \quad (\text{C17})$$

Similarly, the average size of the outward- and inward-pointing clusters are respectively given by  $S(pr)$  and  $S(ps)$ . Obviously their sum must be strictly smaller than the average cluster size in the undirected case,

$$S(pr) + S(ps) - 1 < S(p), \quad (\text{C18})$$

below the percolation threshold,  $p \leq p_c$ .

- 
- [1] G. Degrassi, S. Di Vita, J. Elias-Miro, J. R. Espinosa, G. F. Giudice, G. Isidori, and A. Strumia, Higgs mass and vacuum stability in the standard model at NNLO, *J. High Energy Phys.* **08** (2012) 098.
- [2] D. Buttazzo, G. Degrassi, P. P. Giardino, G. F. Giudice, F. Sala, A. Salvio, and A. Strumia, Investigating the near-criticality of the Higgs boson, *J. High Energy Phys.* **12** (2013) 089.
- [3] Z. Lalak, M. Lewicki, and P. Olszewski, Higher-order scalar interactions and SM vacuum stability, *J. High Energy Phys.* **05** (2014) 119.
- [4] A. Andreassen, W. Frost, and M. D. Schwartz, Consistent use of the standard model effective potential, *Phys. Rev. Lett.* **113**, 241801 (2014).
- [5] V. Branchina, E. Messina, and M. Sher, Lifetime of the electroweak vacuum and sensitivity to Planck scale physics, *Phys. Rev. D* **91**, 013003 (2015).
- [6] A. Bednyakov, B. Kniehl, A. Pikelner, and O. Veretin, Stability of the electroweak vacuum: Gauge independence and advanced precision, *Phys. Rev. Lett.* **115**, 201802 (2015).
- [7] G. Iacobellis and I. Masina, Stationary configurations of the Standard Model Higgs potential: Electroweak stability and rising inflection point, *Phys. Rev. D* **94**, 073005 (2016).
- [8] A. Andreassen, W. Frost, and M. D. Schwartz, Scale invariant instantons and the complete lifetime of the standard model, *Phys. Rev. D* **97**, 056006 (2018).
- [9] J. Khoury and T. Steingasser, Gauge hierarchy from electroweak vacuum metastability, *Phys. Rev. D* **105**, 055031 (2022).
- [10] T. Steingasser and D. I. Kaiser, Higgs criticality beyond the standard model, [arXiv:2307.10361](https://arxiv.org/abs/2307.10361).
- [11] G. Giudice and R. Rattazzi, Living dangerously with low-energy supersymmetry, *Nucl. Phys.* **B757**, 19 (2006).
- [12] H. Friedrich, On the existence of n-geodesically complete or future complete solutions of Einstein's field equations with smooth asymptotic structure, *Commun. Math. Phys.* **107**, 587 (1986).
- [13] P. Bizon and A. Rostworowski, On weakly turbulent instability of anti-de Sitter space, *Phys. Rev. Lett.* **107**, 031102 (2011).
- [14] P. J. Steinhardt, Natural inflation, in *Nuffield Workshop on the Very Early Universe* (Cambridge University Press, Cambridge, 1982), pp. 251–266.
- [15] A. Vilenkin, The birth of inflationary universes, *Phys. Rev. D* **27**, 2848 (1983).
- [16] A. D. Linde, Eternal chaotic inflation, *Mod. Phys. Lett. A* **01**, 81 (1986).
- [17] A. D. Linde, Eternally existing self-reproducing chaotic inflationary universe, *Phys. Lett. B* **175**, 395 (1986).
- [18] A. A. Starobinsky, Stochastic de Sitter (inflationary) stage in the early universe, *Lect. Notes Phys.* **246**, 107 (1986).
- [19] R. Bousso and J. Polchinski, Quantization of four form fluxes and dynamical neutralization of the cosmological constant, *J. High Energy Phys.* **06** (2000) 006.
- [20] S. Kachru, R. Kallosh, A. D. Linde, and S. P. Trivedi, De Sitter vacua in string theory, *Phys. Rev. D* **68**, 046005 (2003).
- [21] S. Ashok and M. R. Douglas, Counting flux vacua, *J. High Energy Phys.* **01** (2004) 060.
- [22] A. D. Linde and A. Mezhlumian, Stationary universe, *Phys. Lett. B* **307**, 25 (1993).
- [23] 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).
- [24] A. H. Guth, Eternal inflation and its implications, *J. Phys. A* **40**, 6811 (2007).
- [25] B. Freivogel, Making predictions in the multiverse, *Classical Quantum Gravity* **28**, 204007 (2011).
- [26] J. Khoury and S. S. C. Wong, Bayesian reasoning in eternal inflation: A solution to the measure problem, *Phys. Rev. D* **108**, 023506 (2023).
- [27] A. Borde, A. H. Guth, and A. Vilenkin, Inflationary spacetimes are incomplete in past directions, *Phys. Rev. Lett.* **90**, 151301 (2003).
- [28] J. Garriga, D. Schwartz-Perlov, A. Vilenkin, and S. Winitzki, Probabilities in the inflationary multiverse, *J. Cosmol. Astropart. Phys.* **01** (2006) 017.
- [29] 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).
- [30] A. Vilenkin, Predictions from quantum cosmology, *Phys. Rev. Lett.* **74**, 846 (1995).
- [31] J. Garriga and A. Vilenkin, Recycling universe, *Phys. Rev. D* **57**, 2230 (1998).
- [32] J. Garriga and A. Vilenkin, A prescription for probabilities in eternal inflation, *Phys. Rev. D* **64**, 023507 (2001).
- [33] R. Bousso, Holographic probabilities in eternal inflation, *Phys. Rev. Lett.* **97**, 191302 (2006).
- [34] 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).



- [35] F. Denef, M. R. Douglas, B. Greene, and C. Zukowski, Computational complexity of the landscape II—Cosmological considerations, *Ann. Phys. (Amsterdam)* **392**, 93 (2018).
- [36] J. Khoury and O. Parrikar, Search optimization, funnel topography, and dynamical criticality on the string landscape, *J. Cosmol. Astropart. Phys.* **12** (2019) 014.
- [37] J. Khoury, Accessibility measure for eternal inflation: Dynamical criticality and Higgs metastability, *J. Cosmol. Astropart. Phys.* **06** (2021) 009.
- [38] G. Kartvelishvili, J. Khoury, and A. Sharma, The self-organized critical multiverse, *J. Cosmol. Astropart. Phys.* **02** (2021) 028.
- [39] J. Khoury and S. S. C. Wong, Early-time measure in eternal inflation, *J. Cosmol. Astropart. Phys.* **05** (2022) 031.
- [40] S. R. Coleman, The fate of the false vacuum. 1. Semi-classical theory, *Phys. Rev. D* **15**, 2929 (1977); **16**, 1248(E) (1977).
- [41] Curtis G. Callan, Jr. and S. R. Coleman, The fate of the false vacuum. 2. First quantum corrections, *Phys. Rev. D* **16**, 1762 (1977).
- [42] S. R. Coleman and F. De Luccia, Gravitational effects on and of vacuum decay, *Phys. Rev. D* **21**, 3305 (1980).
- [43] D. Schwartz-Perlov and A. Vilenkin, Probabilities in the Bousso-Polchinski multiverse, *J. Cosmol. Astropart. Phys.* **06** (2006) 010.
- [44] K. D. Olum and D. Schwartz-Perlov, Anthropic prediction in a large toy landscape, *J. Cosmol. Astropart. Phys.* **10** (2007) 010; **10** (2019) E02.
- [45] 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).
- [46] J. P. K. Doye, Network topology of a potential energy landscape: A static scale-free network, *Phys. Rev. Lett.* **88**, 238701 (2002).
- [47] J. P. K. Doye and C. P. Massen, Characterizing the network topology of the energy landscapes of atomic clusters, *J. Chem. Phys.* **122**, 084105 (2005).
- [48] J. P. K. Doye and C. P. Massen, Power-law distributions for the areas of the basins of attraction on a potential energy landscape, *Phys. Rev. E* **75**, 037101 (2007).
- [49] H. Li, Z. Xu, G. Taylor, C. Studer, and T. Goldstein, Visualizing the loss landscape of neural nets, [arXiv:1712.09913](https://arxiv.org/abs/1712.09913).
- [50] K. D. Boese, A. B. Boese, and S. Muddu, A new adaptive multi-start technique for combinatorial global optimizations, *Oper. Res. Lett.* **16**, 101 (1994).
- [51] P. Erdős and A. Rényi, On random graphs I, *Publ. Math. Debrecen* **6**, 290 (1959).
- [52] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Random graphs with arbitrary degree distributions and their applications, *Phys. Rev. E* **64**, 026118 (2001).
- [53] T. C. Bachlechner, Axionic band structure of the cosmological constant, *Phys. Rev. D* **93**, 023522 (2016).
- [54] T. C. Bachlechner, K. Eckerle, O. Janssen, and M. Kleban, Multiple-axion framework, *Phys. Rev. D* **98**, 061301 (2018).
- [55] T. C. Bachlechner, K. Eckerle, O. Janssen, and M. Kleban, Axion landscape cosmology, *J. Cosmol. Astropart. Phys.* **09** (2019) 062.
- [56] S. Cespedes, S. de Alwis, F. Muia, and F. Quevedo, Quantum transitions, detailed balance, black holes and nothingness, [arXiv:2307.13614](https://arxiv.org/abs/2307.13614).
- [57] G. Odor, Phase transition universality classes of classical, nonequilibrium systems, *Rev. Mod. Phys.* **76**, 663 (2004).
- [58] A. H. Guth and E. J. Weinberg, Could the universe have recovered from a slow first order phase transition?, *Nucl. Phys.* **B212**, 321 (1983).
- [59] P. Creminelli, S. Dubovsky, A. Nicolis, L. Senatore, and M. Zaldarriaga, The phase transition to slow-roll eternal inflation, *J. High Energy Phys.* **09** (2008) 036.
- [60] T. E. Harris, *The Theory of Branching Process* (Springer-Verlag, Berlin, Heidelberg, Germany, 1963).
- [61] B. Bollobás and O. Riordan, A simple branching process approach to the phase transition in  $G_{n,p}$ , *Electron. J. Comb.* **19**, P21 (2012).
- [62] G. F. Giudice, M. McCullough, and T. You, Self-organised localisation, *J. High Energy Phys.* **10** (2021) 093.
- [63] S. W. Hawking and I. G. Moss, Supercooled phase transitions in the very early universe, *Phys. Lett.* **110B**, 35 (1982).
- [64] J. D. Brown and C. Teitelboim, Dynamical neutralization of the cosmological constant, *Phys. Lett. B* **195**, 177 (1987).
- [65] M. P. Salem, Multiverse rate equation including bubble collisions, *Phys. Rev. D* **87**, 063501 (2013).
- [66] A. R. Brown and A. Dahlen, Populating the whole landscape, *Phys. Rev. Lett.* **107**, 171301 (2011).
- [67] E. T. Jaynes, *Probability Theory: The Logic of Science* (Cambridge University Press, Cambridge, UK, 2003).
- [68] A. Vilenkin, Creation of universes from nothing, *Phys. Lett.* **117B**, 25 (1982).
- [69] J. B. Hartle and S. W. Hawking, Wave function of the universe, *Phys. Rev. D* **28**, 2960 (1983).
- [70] A. D. Linde, Quantum creation of the inflationary universe, *Lett. Nuovo Cimento* **39**, 401 (1984).
- [71] A. D. Linde, Quantum creation of an inflationary universe, *Sov. Phys. JETP* **60**, 211 (1984).
- [72] A. Vilenkin, Quantum creation of universes, *Phys. Rev. D* **30**, 509 (1984).
- [73] A. Vilenkin, Boundary conditions in quantum cosmology, *Phys. Rev. D* **33**, 3560 (1986).
- [74] A. Vilenkin, Quantum cosmology and the initial state of the universe, *Phys. Rev. D* **37**, 888 (1988).
- [75] E. Silverstein, (A)dS backgrounds from asymmetric orientifolds, *Clay Mat. Proc.* **1**, 179 (2002).
- [76] 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.
- [77] G. B. De Luca, E. Silverstein, and G. Torroba, Hyperbolic compactification of M-theory and de Sitter quantum gravity, *SciPost Phys.* **12**, 083 (2022).
- [78] L. Katz, A new status index derived from sociometric analysis, *Psychometrika* **18**, 39 (1953).

- [79] K.-M. Lee and E. J. Weinberg, Decay of the true vacuum in curved space-time, *Phys. Rev. D* **36**, 1088 (1987).
- [80] L. Dyson, M. Kleban, and L. Susskind, Disturbing implications of a cosmological constant, *J. High Energy Phys.* **10** (2002) 011.
- [81] 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).
- [82] W. Fischler, D. Morgan, and J. Polchinski, Quantum nucleation of false vacuum bubbles, *Phys. Rev. D* **41**, 2638 (1990).
- [83] W. Fischler, D. Morgan, and J. Polchinski, Quantization of false vacuum bubbles: A Hamiltonian treatment of gravitational tunneling, *Phys. Rev. D* **42**, 4042 (1990).
- [84] S. P. De Alwis, F. Muia, V. Pasquarella, and F. Quevedo, Quantum transitions between Minkowski and de Sitter spacetimes, *Fortschr. Phys.* **68**, 2000069 (2020).
- [85] 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.
- [86] M. Mirbabayi, Uptunneling to de Sitter, *J. High Energy Phys.* **09** (2020) 070.
- [87] K. D. Olum, P. Upadhyay, and A. Vilenkin, Black holes and uptunneling suppress Boltzmann brains, *Phys. Rev. D* **104**, 023528 (2021).
- [88] B. Karrer and M. E. J. Newman, Random graph models for directed acyclic networks, *Phys. Rev. E* **80**, 046110 (2009).
- [89] 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).
- [90] D. Stauffer and A. Aharony, *Introduction to Percolation Theory* (Taylor & Francis, London, UK, 2003).
- [91] J. W. Essam, Percolation theory, *Rep. Prog. Phys.* **43**, 833 (1980).
- [92] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, Graph structure in the web, *Comput. Netw.* **33**, 309 (2000).
- [93] A.-L. Barabási and R. Albert, Emergence of scaling in random networks, *Science* **286**, 509 (1999).
- [94] R. Albert and A.-L. Barabási, Statistical mechanics of complex networks, *Rev. Mod. Phys.* **74**, 47 (2002).
- [95] R. Cohen, D. ben Avraham, and S. Havlin, Percolation critical exponents in scale-free networks, *Phys. Rev. E* **66**, 036113 (2002).
- [96] N. Schwartz, R. Cohen, D. ben Avraham, A.-L. Barabási, and S. Havlin, Percolation in directed scale-free networks, *Phys. Rev. E* **66**, 015104 (2002).
- [97] M. Boguna and M. A. Serrano, Generalized percolation in random directed networks, *Phys. Rev. E* **72**, 016106 (2005).
- [98] D. Garlaschelli, The weighted random graph model, *New J. Phys.* **11**, 073005 (2009).
- [99] Y. Sumitomo and S.-H. Tye, A stringy mechanism for a small cosmological constant, *J. Cosmol. Astropart. Phys.* **08** (2012) 032.
- [100] Y. Sumitomo and S. H. Tye, A stringy mechanism for a small cosmological constant—Multi-moduli cases, *J. Cosmol. Astropart. Phys.* **02** (2013) 006.
- [101] Y. Sumitomo and S.-H. Tye, Preference for a vanishingly small cosmological constant in supersymmetric vacua in a type IIB string theory model, *Phys. Lett. B* **723**, 406 (2013).
- [102] U. Danielsson and G. Dibitetto, On the distribution of stable de Sitter vacua, *J. High Energy Phys.* **03** (2013) 018.
- [103] Y. Sumitomo, S. Tye, and S. S. Wong, Statistical distribution of the vacuum energy density in racetrack Kähler uplift models in string theory, *J. High Energy Phys.* **07** (2013) 052.
- [104] S. H. H. Tye and S. S. C. Wong, Linking light scalar modes with a small positive cosmological constant in string theory, *J. High Energy Phys.* **06** (2017) 094.
- [105] F. Rao and A. Caflisch, The protein folding network, *J. Mol. Biol.* **342**, 299 (2009).
- [106] H. S. Wilf, *Generatingfunctionology*, 3rd ed. (AK Peters/CRC, Natick, MA, USA, 2006).