

Theory of minimum spanning trees. I. Mean-field theory and strongly disordered spin-glass model

T. S. Jackson^{*} and N. Read[†]*Department of Physics, Yale University, P.O. Box 208120, New Haven, Connecticut 06520-8120, USA*

(Received 20 February 2009; revised manuscript received 10 November 2009; published 25 February 2010)

The minimum spanning tree (MST) is a combinatorial optimization problem: given a connected graph with a real weight (“cost”) on each edge, find the spanning tree that minimizes the sum of the total cost of the occupied edges. We consider the random MST, in which the edge costs are (quenched) independent random variables. There is a strongly disordered spin-glass model due to Newman and Stein [Phys. Rev. Lett. **72**, 2286 (1994)], which maps precisely onto the random MST. We study scaling properties of random MSTs using a relation between Kruskal’s greedy algorithm for finding the MST, and bond percolation. We solve the random MST problem on the Bethe lattice (BL) with appropriate wired boundary conditions and calculate the fractal dimension $D=6$ of the connected components. Viewed as a mean-field theory, the result implies that on a lattice in Euclidean space of dimension d , there are of order W^{d-D} large connected components of the random MST inside a window of size W , and that $d=d_c=D=6$ is a critical dimension. This differs from the value 8 suggested by Newman and Stein. We also critique the original argument for 8, and provide an improved scaling argument that again yields $d_c=6$. The result implies that the strongly disordered spin-glass model has many ground states for $d>6$, and only of order one below six. The results for MSTs also apply on the Poisson-weighted infinite tree, which is a mean-field approach to the continuum model of MSTs in Euclidean space, and is a limit of the BL. In a companion paper we develop an $\varepsilon=6-d$ expansion for the random MST on critical percolation clusters.

DOI: [10.1103/PhysRevE.81.021130](https://doi.org/10.1103/PhysRevE.81.021130)

PACS number(s): 02.50.-r, 89.75.Fb, 75.10.Nr

I. INTRODUCTION

A. Motivation and approach

The minimum spanning tree (MST) problem is one of the oldest and best-studied problems of combinatorial optimization [1–5] and has found application to the physics of random systems [6–9]. To define the problem, we consider an undirected, connected graph G with vertex set V , edge set E and a real-valued cost ℓ_e assigned to each edge $e \in E$. A spanning tree is then defined as a subset of the edges of G that connects all the vertices and contains no cycles: in other words, it is a tree and it spans V . Such a tree must exist because the graph is assumed connected. A *minimum* spanning tree \mathbf{T} is a spanning tree such that the sum of the costs of its edges,

$$\ell(\mathbf{T}) = \sum_{e \in \mathbf{T}} \ell_e, \quad (1)$$

is minimized over the set of all spanning trees on G . If the costs ℓ_e are strictly positive, then any spanning subset of the edges that has minimum cost is automatically a tree.

If we view the cost [Eq. (1)] as an energy, then this is a problem of finding the ground state of a classical system in which the configurations are spanning trees. If the costs of the edges are taken to be random variables, then this becomes a classical system with quenched disorder: the minimum must be found for a fixed set of costs, before any averaging over realizations of the edge costs is performed.

In the present paper and its companion [10], we begin a program to develop an analytical theory of the statistical geometry of random MSTs on a lattice \mathbb{Z}^d in d -dimensional Euclidean space, where the edge costs are independently and identically distributed (iid) with a continuous probability distribution. (For continuous distributions of iid edge costs, the MST on a finite graph is unique with probability one.) We will primarily be interested in this model’s long-range scaling properties (to be defined in the following section) because these are predicted to be the same for all models in the same universality class: in a quantitative sense, changing microscopic details of the model such as the type of lattice will not affect these properties (see e.g., Ref. [11] for further discussion). We will argue below that, due to universality, our results also apply to related problems, such as the continuum model [5,12]. In this model, the vertices are points in Euclidean space \mathbb{R}^d which are Poisson distributed with uniform density, so the graph is the infinite complete graph, and the cost assigned to each edge is the Euclidean distance between its end points. Results concerning the expectation value of the cost $\ell(\mathbf{T})$ were previously given in Ref. [13].

It is simple to solve computationally an instance of the MST problem, and many efficient algorithms exist [3,14–17]. However, we are interested in the statistical properties of the random MST, which is relevant to physical disordered systems. MSTs play a role in transport in disordered networks [18–22], for example in current flow in random resistor networks [13,23–25]. There exist in the literature several numerical simulations [6,7,26–28] determining D_p , the fractal dimension of the optimal path.

Furthermore, Newman and Stein (hereafter referred to as NS) have shown [8,9] that in the strong disorder limit, the problem of finding the ground state of an Ising spin glass

^{*}thomas.s.jackson@yale.edu

[†]nicholas.read@yale.edu

may be directly mapped on to the MST problem. Some of the results presented here bear directly on the discussion by these authors, which relate to fundamental questions about the ground state structure of spin glasses, and, by extension, solutions to hard optimization problems [29–31]. Related considerations also arose in the quantum spin glass (or random transverse field Ising model) [32]. These connections will be discussed at length below.

Our program for MSTs takes a form similar to those for critical phenomena in, for example, Ising spin systems, or percolation. It proceeds in stages. The first stage is a mean-field theory, which we develop in the present paper by introducing and solving a version of random MST on the Bethe lattice (BL) (or Cayley tree), with suitable boundary conditions. The BL model can be solved exactly in some cases, or exactly for asymptotic, universal properties in more cases. We also investigate the related Poisson-weighted infinite tree (PWIT) model [12,33,34], which can be obtained as a limit of the BL as the coordination number goes to infinity and which can be viewed as the mean-field theory of the continuum model defined above. The subsequent stages would be to develop a full statistical field theory (or at least, a perturbation expansion) for the geometric properties in Euclidean space (either for the lattice or continuum models). Then a renormalization group approach should first show that corrections to the mean-field theory do not change the universal asymptotic properties above some critical dimension d_c , and identify d_c . A third stage would be calculation of universal properties at $d < d_c$ in an epsilon expansion in powers of $\varepsilon = d_c - d$. A fourth stage would be to show that the epsilon expansion is Borel summable, so that it defines the true results. The second and later stages will not be completed in this paper.

In developing an approach to the MST problem using Kruskal's greedy algorithm [14], which is related to bond percolation, we will be led to consider the process of finding the minimum spanning forest (MSF) on the clusters of bond percolation as a function of the probability p that an edge is occupied. (A spanning *forest* is a spanning, vertex-disjoint collection of trees.) We call this process MSF(p) [note that MSF(1) is the same as the MST]. We will argue that the universal properties of MSF(p) at any $p > p_c$, where p_c is the percolation threshold of the model, are the same. At this time we have been able to develop the perturbation techniques as mentioned above only for MSF(p) with $p \leq p_c$. The expansion at $p > p_c$ is more difficult. However, it has been argued that some properties for $p > p_c$ are the same as those for $p = p_c$ [20,23–25]. In the absence of a perturbation theory treatment, it is not clear if this is true, but certainly we find many indications on the BL that the vicinity of p_c dominates the behavior in the region $p > p_c$. In any case, in the companion (to be referred to as II) to the present paper we develop first a small- p expansion that is exact for any p on a finite graph, and then a perturbation expansion using modified Feynman diagram techniques valid for $p \leq p_c$. By using renormalization-group techniques, this yields an epsilon expansion for the fractal dimension of a path on MSF(p_c).

B. Outline and discussion of main results

We now describe the main results of this work. The BL is an infinite tree with fixed coordination number (degree) at

each vertex. As it is itself a tree, the MST on such a graph would be simply the whole graph, so attention to the definition of boundary conditions in the spirit of [8,9] is essential in producing nontrivial behavior. The boundary conditions can be introduced first on a finite version of the BL, and then the infinite size limit can be taken. Specifically, as discussed in detail in Sec. II, we adopt a wired boundary condition, which has the result that instead of the MST, the minimum object is a spanning forest. [This MSF produced by the wired boundary condition should not be confused with that in the process MSF(p) mentioned above; in that language, at present we are considering the minimum objects at $p=1$.] We are interested in the statistical geometry of this nontrivial random forest. As the size of the lattice goes to infinity, the statistical properties of the MSF have a well-defined limit. We call the number of vertices that are connected to the central site and lie within m steps on the BL the “mass” $M(m)$ within m steps. We can then calculate (among other things) its expectation value $\overline{M(m)}$, and we find that it scales as

$$\overline{M(m)} \sim m^3 \quad (2)$$

as $m \rightarrow \infty$. The same result holds for the PWIT.

Employing this result as a mean-field theory, the standard method (see, e.g., [8,9,35]) for transferring results from the BL to a Euclidean lattice entails that distance m on the BL corresponds to distance squared, $m \sim R^2$ on the Euclidean lattice. We find that the expected mass within distance R of the origin scales as

$$\overline{M(R)} \sim R^6, \quad (3)$$

so that the tree has fractal dimension $D=6$. As the trees fill the lattice, this means that the expected number of connected components that intersect a ball of radius R , denoted $N(R)$, scales as

$$N(R) \sim R^\# \sim R^{d-6}, \quad (4)$$

as $R \rightarrow \infty$, so the tree “proliferation exponent” [36,37] $\# = d - D = d - 6$. Note that on the BL, $N(m)$ increases exponentially with m , but there is a power-law correction factor of m^{-3} which produces the behavior relevant for the Euclidean lattice.

Two points should be explained here. One is that $N(R)$ may be subject to boundary effects at the surface of the ball, so that the number of components intersecting the ball is larger, maybe R^{d-1} . Our result is expected to describe the number of *large* components intersecting the ball, say those whose intersection with the ball is of linear size $R/2$ [36,37].

The second, very important, point is that the MST on a finite connected portion of a Euclidean lattice is again a connected object by definition (for conventional “free” boundary conditions). But as this portion becomes larger and approaches the whole lattice, the path between any two vertices on the tree may make larger and larger excursions, so that in the limit, viewed locally on any finite length scale R , the tree appears to be a forest of many connected components [8,9]. The use of the wired boundary condition is intended to simulate this possible effect, by producing such a forest in a finite system, though the properties near the wired boundary may

differ from those near the boundary of a ball of radius R inside a system of size much larger than R .

The proliferation exponent $\#$ cannot become negative, so even if the mean-field theory is indeed valid in high dimensions, it must break down in sufficiently low dimensions. These notions parallel some in other problems of random fractal clusters, such as in percolation at threshold (the critical point). A nonzero proliferation exponent $\#$ is the geometric counterpart to the violation of hyperscaling relations in critical phenomena at $d > d_c$ [36,37]—hyperscaling is obeyed when $\#=0$. In critical percolation, the clusters at threshold are non-space-filling, and their fractal dimension is $D=4$ for $d \geq d_c$, while $d_c=6$. The MST, on the other hand, is an example of the subclass of such problems in which the union of the clusters fills space, so that $D+\#=d$. This then suggests that $d_c=D=6$ is the upper critical dimension, below which exponents (such as D) must change with d . Below $d=6$ it is plausible that there is only a single connected component (in the local sense described above) with probability one, or at least that, as $R \rightarrow \infty$, $N(R) < R^\alpha$ for any $\alpha > 0$. Then $D=d$ and $\#=0$, and

$$D = \begin{cases} 6 & (d \geq 6) \\ d & (d \leq 6), \end{cases} \quad (5)$$

$$\# = \begin{cases} d-6 & (d \geq 6) \\ 0 & (d \leq 6). \end{cases} \quad (6)$$

This result contrasts with that of NS, who suggested that $d_c=8$ for MSTs [8,9], in the sense that $\# > 0$ for $d > 8$, $\#=0$ for $d < 8$. NS appear to have believed that the connected components of the MST have dimension 4, but their arguments would actually imply that $D=8$ also; their arguments will be discussed in more depth in Sec. II. (In fact, they showed that $d_c \leq 8$.) The question of the number of connected components was raised in a construction of a minimum spanning forest in the continuum model directly in infinite volume by Aldous and Steele [5,12], who suggested that the forest has a single connected component in all dimensions d . It has been proved that there is a single connected component for this model in $d=2$ [38], but for larger d there has not so far been agreement even at a heuristic level. We note that for a different model of random trees on a lattice, that of uniform spanning trees (each spanning tree on a finite graph is given equal probability), similar behavior of the dimensions was proven, except that $d_{c\text{UST}}=4$ and $D_{\text{UST}}=4$ for $d > 4$ [39]. Hence the universal properties of MSTs and uniform spanning trees are distinct, at least in sufficiently high dimensions [40,41].

In the companion paper II, we formulate a perturbation expansion for the geometry of the MSF(p) process in Euclidean space for $p \leq p_c$. This is based on an exact small- p expansion. It is used to calculate the fractal dimension of a path on the MST on a critical percolation cluster. Here we will mention only that our calculations can in principle be extended to other exponents, such as those defined in [40], or carried to higher orders in ε . They can also be extended to include statistical properties that involve the cost of the MSF(p_c). There do not appear to be any scaling relations that

relate the geometric exponents for MSTs or MSF(p_c) to those for percolation, unlike those found for the costs in [13], even though the critical dimension $d_c=6$ is the same for all of them. In Ref. [13], it was stated that there are different critical dimensions for properties involving the costs and for those involving only the geometry (6 and 8, respectively). This remark was based on NS's results for the geometry, and is superseded by the present paper.

C. Structure of the paper

The remainder of the paper parallels the discussion of the preceding section. In Sec. II, we discuss the general properties of MSTs that we exploit in our calculations, and explain the relation to percolation. We introduce Prim's and Kruskal's algorithms for computing the MST; these are related to invasion percolation and ordinary bond percolation, respectively. Similarly, the continuum model of MST is related to continuum versions of these. The strongly disordered spin-glass model of NS is defined, and we provide a critique of their results for MSTs. We also give a scaling argument that suggests that the correct answer is $d_c=6$, $D=6$, which is confirmed in the following section. In Sec. III, we solve the MST problem statistically on the BL using the connection between Kruskal's algorithm and percolation; this defines a mean-field theory. We also discuss the limit that gives MSTs on PWITs. Implications of our results for the strongly disordered spin-glass model are discussed in the Conclusion.

II. MINIMUM SPANNING TREES, PERCOLATION, AND STRONGLY DISORDERED SPIN GLASSES

In this section, we describe basic properties and techniques for solving a MST problem. There are two main simple algorithms to consider, Prim's and Kruskal's. Both are "greedy" in nature, and when the edge costs are assumed to be iid random variables with a continuous distribution, both are related [42] to models of percolation. Prim's algorithm, which seems to be more popular in the physics literature on MSTs, is connected with invasion percolation. Kruskal's algorithm (also sometimes referred to simply as the greedy algorithm) is related to bond percolation. As the latter is much more tractable from the statistical mechanics point of view, this is the approach that we will use. Here we review these connections and basic properties of percolation. We review the strongly disordered spin-glass model of Newman and Stein (NS), and critique their argument that the critical dimension of the model is $d_c=8$. In NS, this critical dimension is that dimension d above which there is an infinite number of ground states in the thermodynamic limit of the model. We give a scaling argument for $d_c=6$.

A. Basic properties and algorithms

A property that is useful in finding the MST of an edge-weighted finite graph is the following [2]: suppose a spanning forest is given, that is not necessarily minimum cost (some connected components of such a forest may consist of a single vertex and no edges). If we choose a connected component, and then find the edge e of minimum cost among

all those edges with just one end in that component, then among all the spanning trees that contain the given spanning forest (as a subset of the edges of the tree), the one of minimum cost contains the edge e . Thus starting from any given forest, one can greedily add edges of minimum cost (among those that leave any one connected component at each stage) and arrive at the spanning tree that has minimum cost among all those containing the given forest. In particular, starting with the forest in which each tree is a single vertex and no edges, one can find the MST. This still leaves many ways to proceed in selecting the connected component at each step. Two particular ways of doing so are of interest here: (a) the Kruskal or greedy algorithm [14], in which at each step one adds the cheapest edge not yet occupied that does not create a cycle when added to the set of occupied edges; (b) the Jarník-Prim-Dijkstra algorithm [15–17] (usually referred to as Prim’s), in which one starts by selecting a vertex, and then adds the cheapest edge leaving the connected component containing the initial vertex at each step. In either case, the algorithm stops when a spanning tree is formed, that is when $|V|-1$ edges are occupied. From the preceding result, both of these ultimately produce the MST. (We have neglected to specify how to break any ties that arise from edges of equal cost, which can lead to nonunique MSTs, as these are not of interest in this paper.) The difference between the two in terms of the geometry of the connected components or “clusters” is that in Kruskal’s algorithm, at a typical stage there are several trees containing more than one vertex, whereas in Prim’s there is always only one. Both procedures have similar running times, and can also be improved [3].

The graph and the set of edge costs defines the MST via minimization of the total cost, which is the sum of the costs of the occupied edges as in Eq. (1). From the above algorithms, it is clear that the precise values of the costs are not important in finding the geometry of the MST (and hence neither is the probability distribution for them [26]). Only the rank ordering of the set of costs is important. Indeed, even this much information is not required, as evidenced by the fact that the most efficient known deterministic algorithm for computing the MST on an arbitrary graph [43] has significantly faster asymptotic running time than that for sorting all the edges of that graph by cost.

B. Relation of Kruskal’s algorithm with bond percolation

The standard way to define bond (or Bernoulli) percolation on a graph is to declare that edges are either occupied with probability p or not occupied, independently. We are then interested in the geometric properties of the clusters (connected components) formed by the occupied edges. For a graph that is a portion of a lattice in Euclidean space, the basic problem is to find the *percolation threshold*, the value of p above which there exists a path on a single cluster from one side of the system to the other. When the graph is the complete graph, the percolation problem becomes the theory of “random graphs” (these graphs being the clusters). The threshold is then the value of p at which a cluster of size of order $|V|$ occurs, as $|V| \rightarrow \infty$. In both cases, the behavior near the threshold is a critical phenomenon, and there are critical

exponents. The complete graph model is one popular way of obtaining a mean-field treatment of percolation.

A random edge-weighted graph gives rise to a sample of bond percolation, as follows. We suppose that the weights or costs on the edges are iid random variables, with a probability density $P(\ell_e)$ of the cost of any one edge e . If we say that all edges with costs less than or equal to ℓ_0 are occupied, then this gives a sample of bond percolation in which $p = p_0$ is given by

$$p_0 = \int_{-\infty}^{\ell_0} d\ell_e P(\ell_e). \quad (7)$$

In particular, if $P(\ell_e) = 1$ for ℓ_e in the interval $[0, 1]$ and zero outside, then $p_0 = \ell_0$.

Now consider Kruskal’s algorithm for such a random edge-weighted graph. At each step, one must “test” the edges not tested earlier, and find the cheapest one that does not create a cycle when added to those occupied. We say that such an edge is accepted. Thus we can view the algorithm as working through the edges in order of increasing cost. If the algorithm is terminated at cost ℓ_0 (at which stage there may be only a spanning forest, not yet a spanning tree), then the set of edges already tested forms a sample of percolation, as just described. Thus, if we had accepted all edges instead of only those that did not form a cycle, we would have obtained a sample of bond percolation. We can thus view percolation, as well as Kruskal’s algorithm, as a dynamical process in which clusters are grown beginning from the set of vertices and no occupied edges, and eventually obtaining a cluster spanning the graph. Kruskal’s algorithm gives a variant on this percolation process in which as ℓ_0 increases, an edge is accepted only if adding it does not form a cycle [44]. The process does not depend on the probability density chosen for the costs, because as mentioned above only the ordering is important for the tree. The process is fully characterized by the variable p_0 , and distinct probability densities $P(\ell_e)$ give rise to the same probability measure on MSTs, as emphasized in Ref. [26]. Without loss of generality, we can consider ℓ_e to be uniformly distributed between 0 and 1, so $p_0 = \ell_0$, as above. We will use the name $\text{MSF}(p_0)$ for the random spanning forest produced by halting the Kruskal algorithm at some value p_0 ; the true MST is obtained at $p_0 = 1$, or by halting when $|V|-1$ edges are occupied.

Later in the paper we develop analytical techniques to study MSTs using the theory of bond percolation. Here we mention some of the basic properties of percolation, for the limit of an infinite system [36]. The percolation threshold p_c is nonuniversal; it depends on the lattice or class of graphs considered. For hypercubic graphs in dimension d (i.e., \mathbb{Z}^d , with edges connecting nearest neighbors only), p_c lies strictly between 0 and 1 for $d > 1$, while for $d = 1$, $p_c = 1$. For $p < p_c$, there are finite clusters only. The typical size ξ of a cluster (the correlation length) diverges as $\xi \sim (p_c - p)^{-\nu_{\text{perc}}}$ as $p \rightarrow p_c$ from below. For $p > p_c$ (so assuming $d > 1$), there is a single infinite cluster with probability one, and for $p_c < p < 1$ a nonzero density of finite clusters (clusters not connected to infinity). The typical size of the finite clusters now defines ξ , and $\xi \sim (p - p_c)^{-\nu_{\text{perc}}}$ as $p \rightarrow p_c$ from above, with the

same exponent ν_{perc} . The correlation length exponent ν_{perc} equals $1/2$ for $d \geq 6$, but deviates from this value for $d < 6$. At criticality, $p = p_c$, there is a power-law distribution of cluster sizes up to infinite size. For $d \geq 6$ these clusters have fractal dimension $D_{\text{perc}} = 4$, while this dimension D_{perc} deviates from this value below $d = 6$. For $d \geq 6$, there are of order $R^{\#_{\text{perc}}}$ large such clusters intersecting a ball of radius R , where $\#_{\text{perc}} = d - 6$, while $\#_{\text{perc}} = 0$ for $d \leq 6$. All the exponents may be calculated by considering suitable correlation functions. (All power laws may be subject to logarithmic corrections when $d = 6$, which we do not consider.)

C. Relation of Prim's algorithm with invasion percolation

The arguments of the previous subsection can be repeated for Prim's algorithm. From Prim's algorithm, one obtains a dynamical process in which at each step a single edge is added to growing cluster (tree), which contains the initial vertex \mathbf{x} . The edge added is the cheapest one that borders the current tree and does not form a cycle. In a finite system, this cluster eventually spans all the vertices, and at that point becomes the MST. In an infinite system, if the "time" in the process is identified with the number of edges added, then the tree continues to grow indefinitely until an infinite tree is obtained after infinite time. We should note that it is not obvious that this infinite tree is spanning, and for $d > 1$ it will not be. Because the tree depends on the starting vertex \mathbf{x} say, we denote this tree obtained after infinite time as $T_\infty(\mathbf{x})$; it is a random object that depends on the costs of all the edges that were tested in growing the tree. The latter edges are those on the tree, together with those that have at least one end connected to \mathbf{x} by $T_\infty(\mathbf{x})$.

Prim's algorithm is connected to a type of percolation called invasion percolation [45–47] analogously to how Kruskal's is connected to bond percolation; for Prim's algorithm this connection has been noted repeatedly in the physics literature [6–9,19,26]. Using the edge-weighted graph with iid costs as before, a model of invasion percolation is obtained by modifying Prim's algorithm to accept the cheapest edge that borders but is not on the current cluster (i.e., by neglecting the no-cycle condition). Notice that in both the Prim and invasion process, the costs of the accepted edges do not increase monotonically as they did in the case of the Kruskal or percolation processes, because when the cluster first enters a region of space, additional edges become available, which may be cheaper than those accepted earlier. Indeed, in an infinite system, the edges accepted after long times have (with high probability) costs $\leq p_c$ of the corresponding bond percolation problem (in the model with costs uniformly distributed between 0 and 1) [45–48]. It is believed [45–48] that the invasion cluster is a fractal very similar to the critical percolation clusters, with the same universal properties, so its fractal dimension is $D_{\text{inv}} = D_{\text{perc}} = 4$ for $d \geq 6$. Clearly, the set of vertices on the invasion cluster and on the invasion tree produced by Prim's algorithm are the same, so the same fractal dimensions apply to the trees $T_\infty(\mathbf{x})$.

D. Strongly disordered model of a spin glass

NS [8,9] (see also Refs. [6,7]) defined a strongly disordered spin-glass model and showed that the problem of find-

ing the ground state maps onto a MST problem. We now briefly describe this model. In the next subsection, we discuss the arguments of NS, which motivated many of the questions addressed in this paper.

The Edwards-Anderson (EA) Ising spin-glass model [29] is defined by the Hamiltonian

$$\mathcal{H} = - \sum_{i < j} J_{ij} s_i s_j \quad (8)$$

for Ising spins $s_i = \pm 1$, where J_{ij} are quenched random variables. The positions i, j are taken as lattice points in a portion (say, a cube of side L) of a d -dimensional hypercubic lattice, and for the edges $\langle ij \rangle$ of the corresponding graph (i.e., "nearest-neighbor bonds"), the J_{ij} 's are iid variables with a distribution independent of the portion of the lattice chosen (in particular, independent of the number $|V| = L^d$ of spins), for example a Gaussian distribution with mean zero and standard deviation J_0 . The J_{ij} 's are quenched random variables, meaning that thermodynamic quantities must be calculated with a fixed sample of J_{ij} 's, and then averages (or moments, etc) taken at the end. The NS strongly disordered model differs from the EA model in the distribution of J_{ij} 's. While they are still iid, the width of the distribution is assumed to be extremely large, and depends on the number of spins, in a fashion to be specified below. The distribution is symmetric, so that the sign of J_{ij} , $\epsilon_{ij} \equiv \text{sgn } J_{ij}$, is ± 1 with probability $1/2$ for either case, independently of the random magnitude $K_{ij} = |J_{ij}|$. NS focus on ground-state properties, and specify a boundary condition that the spins on the boundary are an arbitrary set of values ± 1 , chosen independently of the J_{ij} 's on the edges in the interior. Clearly, similar models can be defined on other graphs, including the complete graph (infinite-range model), or with different boundary conditions.

The central idea in the use of a broad distribution of disorder (that is, of the K_{ij} 's) is that for such a broad distribution, for any subset of edges, there is always a single K_{ij} that dominates all others in the set, or even dominates the sum of all the others. Thus the width of the distribution must simply be chosen large enough that this is so, and that is why the width must increase with the system size. In this case the problem of finding the ground state of the model with given bonds J_{ij} and boundary spin values may be solved by a greedy procedure. NS chose to use a procedure similar to Prim's algorithm. To find the orientation (i.e., the value ± 1) in the ground state of a given spin located at \mathbf{x} , first find the largest K_{ij} among the edges leaving \mathbf{x} . The relative orientation, $s_i s_j$ for i corresponding to \mathbf{x} , is clearly then determined, but not s_i itself. Then find the largest K_{ij} leaving this cluster, which has one end on the cluster, but not the other. This fixes the relative orientation with a further spin k , say. This process can be repeated, adding edges to the cluster until a boundary spin is encountered, at which point all the spins on the cluster are now determined. Edges that connect two sites of a cluster that are already connected need not be considered (or accepted), since the relative orientation of those spins is already determined. Thus the edges accepted form a tree. The process is clearly identical to Prim's algorithm, until the growing invasion tree touches the boundary. The process can then be repeated starting with any spin not already fixed,

until all spins have been determined. For the trees after the first, the process is defined to restart from another vertex not already connected if the growing tree encounters an earlier tree, as well as if it encounters the boundary, since again such an encounter fixes the spin orientations on that growing tree.

The process of repeatedly growing trees until every vertex is part of a tree that touches the boundary exactly once produces a spanning forest on the graph (or portion of the lattice). From the percolation point of view, similar use of boundary conditions is called a wired boundary condition, and has also been used in MST problems [5]. We may imagine that the edges that connect the boundary vertices to one another have costs less than all those in the interior, so in Prim's algorithm, once the first tree encounters the boundary, it is immediately connected to all other boundary sites along the boundary. (The precise ordering of the boundary costs among themselves is not of interest because we are not interested in the edges of the MST on the boundary.) Then as all earlier trees are now viewed as connected in a single tree that includes all the boundary vertices, the result when the process terminates is a spanning tree. It is in fact the MST on the given graph with costs $-K_{ij}$, together with costs $\sim -\infty$ for the boundary edges. The process used by NS is a variant on the Prim process that occasionally restarts from a different vertex that is not connected to any others. This is among the many different ways to find the same MST, as can be seen using the general fact from Sec. II A. Hence *any* valid construction of the MST with this wired boundary condition produces the same spanning tree, and hence also produces the same ground state of the NS model. In particular, we can use the Kruskal algorithm.

Thus finding the ground state of the NS model is a MST problem. Once the MST with the costs $-K_{ij}$ in the interior, and wired boundary conditions, has been found, the spin orientations follow using the signs ϵ_{ij} and the boundary spins. Because this is now essentially solving a spin-glass model on a forest, there is effectively no frustration left at this stage. We recall that frustration in an Ising spin model with Hamiltonian of the form of \mathcal{H} means that there exist cycles of the graph such that the product of signs ϵ_{ij} of the edges on the cycle is negative; generally this means that all the details of the magnitudes of the couplings must be studied in order to find the ground state. For the EA model in more than two dimensions, finding the ground state is computationally costly. Indeed, in $d > 2$ the problem of determining whether the ground-state energy (cost) of an Ising spin-glass Hamiltonian \mathcal{H} is less than some bound (budget) is NP complete [49], and so presumably cannot be solved in polynomial time (for a discussion of NP completeness, see Ref. [50]). (For $d=2$, or for planar graphs, the spin-glass ground state can be mapped to a network flow problem, and solved in polynomial time.) By contrast, for the NS model, the ground state can be found in polynomial time using an MST algorithm.

E. Ground states of the NS model and fractal dimensions

We now turn to NS's analysis of their model. The model was constructed so as to be soluble. The use of fixed bound-

ary spins, or the wired boundary condition on the MST, was motivated by a deep view of the meaning of a thermodynamic state in a spin system, or in the present case a ground state [8,9,31]. In an infinite system, a ground state can be defined as a spin configuration the energy of which cannot be lowered by reversing the values of any *finite* set of spins. A ground-state spin configuration in any bounded portion of the lattice (such a configuration is simply that of minimum energy) is completely determined by the values of the spins on its boundaries. As the size of this portion goes to infinity (keeping the bonds J_{ij} the same in the interior as fresh bonds and spins are added at the boundary), there should exist sequences of boundary conditions such that the spin configuration seen in any "window" (subregion of the system) converges to a limit. When this is done for a sequence of windows diverging in size (such that the spin configurations agree where the windows overlap), then a ground state of the infinite system is obtained. It is then clear that this ground state is determined by a choice of boundary conditions infinitely far away. The use of arbitrary boundary spins on the boundary of a finite region (a hypercube of side L , say) as in NS is then part of this process, and approximates the ground states of the thermodynamic limit.

As we have seen, NS determine the ground state of their model, using boundary spin values and a set of invasion trees. Then the spins in the box lie on a spanning forest in which each connected component touches the boundary just once. All spins on such a connected component will be reversed if the value of the corresponding boundary spin is reversed. Thus the logarithm of the number of possible distinct ground states obtained inside the box of size L by varying the boundary spin values is bounded by $O(L^{d-1} \ln 2)$ (it is a bound, not the actual number, because some boundary spins may not be connected to any interior spins, and these boundary spins are not to be considered when counting configurations). But it is better to consider a window of side W within the box, with $W \ll L$, and ask how many distinct configurations can be obtained within the window as the boundary conditions are varied, preferably without counting configurations that differ only just inside the surface of the window. The logarithm of this number is given by the number of connected components of the minimum spanning forest that intersect the window (neglecting those whose linear size is less than $W/2$, say). (There may be some ambiguity here concerning connected components that intersect the window more than once, and are connected outside the window but not inside, as these do not give four ground states, but only two, however we will assume this is not significant.) Thus this number $\ln \mathcal{N}(W)$ is the same as $N(W)$ as defined in Sec. I,

$$\ln \mathcal{N}(W) = N(W) \ln 2. \quad (9)$$

We have arrived at the same question about MSTs that was already introduced in Sec. I: the behavior of the number of connected components of a MST intersecting a window, or alternatively the fractal dimension of the connected component(s) of an infinite MST.

NS studied this question using the Prim algorithm. To find whether two vertices, \mathbf{x} , \mathbf{y} , say, are on the same connected

component, we may grow the invasion tree from each of them, and see whether they intersect. More exactly, we could first grow the tree from \mathbf{x} to infinity, then begin again from \mathbf{y} , stopping if $T_\infty(\mathbf{x})$ is encountered. Since each invasion tree is a fractal of dimension $D_{\text{inv}}=4$ (for $d \geq 6$), NS were able to show rigorously that when $d > 2D_{\text{inv}}=8$, there is a nonzero probability that the two trees “miss” and never intersect [8,9]. This event implies that the two points are on distinct connected components, and so if the critical dimension d_c is defined as that above which the MST has more than one connected component in the thermodynamic limit (with probability one), then NS proved the upper bound $d_c \leq 8$.

To determine the actual value of d_c , NS stated that they need a converse result, in other words a lower bound. This converse statement would result if the probability that the two invasion trees do intersect is of order 1 for $d < 8$. Such behavior is natural if one has two independently grown fractals of dimension 4 (or more generally, if $2D > d$ for two fractals of dimension D). Now for the invasion trees, independence does hold when the clusters are sufficiently far apart, because the edges considered when growing either one are only those bordering it, not all those in the system. (We recall that the edges bordering the tree are those with one end on the tree, and one end off it.) More precisely, they are independent as long as they do not intersect and the sets of edges bordering one are disjoint from the set of edges bordering the other. This leads to NS’s result $d_c \leq 8$. However, if the trees do become close together so that they share one or more bordering edges, for example if $T_\infty(\mathbf{x})$ is already grown first, and we are growing the tree from $\mathbf{y} \notin T_\infty(\mathbf{x})$, then they are correlated. From the invasion process, we can see that edges bordering $T_\infty(\mathbf{x})$ must be of cost higher than p_c (again, in the model where edge costs ℓ or p are uniform between 0 and 1), and not arbitrary like those bordering the tree from \mathbf{y} when they are *first* encountered. This very strong correlation effect reduces the probability that any such edge that forms the connection between $T_\infty(\mathbf{x})$ and the tree from \mathbf{y} will be accepted on $T_\infty(\mathbf{y})$. Hence we expect that $d_c < 8$.

More quantitatively, as the tree $T_\infty(\mathbf{x})$ has dimension $D_{\text{inv}}=4$ for $d > 6$, the probability $G(\mathbf{z}-\mathbf{x})$ that a point \mathbf{z} will be on $T_\infty(\mathbf{x})$ should behave as

$$G(\mathbf{z}-\mathbf{x}) \sim |\mathbf{z}-\mathbf{x}|^{D_{\text{inv}}-d} \tag{10}$$

for large $|\mathbf{z}-\mathbf{x}|$ [8,9,36]. If the two trees were grown independently, the probability that they intersect would behave as

$$\int d^d \mathbf{z} G(\mathbf{z}-\mathbf{x}) G(\mathbf{z}-\mathbf{y}) \sim |\mathbf{x}-\mathbf{y}|^{2D_{\text{inv}}-d}. \tag{11}$$

For $d < 2D_{\text{inv}}(d)=8$ this is instead of order one. This would give the converse result of NS. It would also imply that connected components of the minimum spanning forest have dimension at least $2D_{\text{inv}}=8$ for $d \geq 8$ (though NS seem to have believed that this dimension would be 4). But because the trees are not independent when they intersect, neither result is correct.

To make a crude estimate of the correct result, we use scaling arguments for the growth of $T_\infty(\mathbf{y})$. Given \mathbf{x} and \mathbf{y} , the tree from \mathbf{y} first approaches close to $T_\infty(\mathbf{x})$ when its linear

size ξ is of order $|\mathbf{x}-\mathbf{y}|$. The costs of edges typically accepted when it reaches a size scale ξ will be distributed up to about p_c , with $p-p_c$ not larger than of order $\xi^{-1/\nu_{\text{perc}}}$, where we recall that $\nu_{\text{perc}}=1/2$ for $d \geq 6$. (Here p stands for the cost of the edge, in accordance with the earlier discussion of the relation of MST with percolation.) Indeed, the probability density for an edge to be accepted at this stage will be given by a scaling function of $(p-p_c)\xi^{-1/\nu_{\text{perc}}}$, which goes to one when its argument is large and negative, and to zero when its argument is large and positive. Similarly, the costs of edges bordering $T_\infty(\mathbf{x})$ at distance of order ξ from \mathbf{x} have a probability distribution that is a function of $(p-p_c)\xi^{-1/\nu_{\text{perc}}}$, which vanishes for $p < p_c$ and goes to one for large positive arguments. Multiplying these functions and integrating over p to obtain the probability that the first such edge is accepted, the result is of order $\xi^{-1/\nu_{\text{perc}}}=|\mathbf{x}-\mathbf{y}|^{-2}$ for $d > 6$. Assuming that consideration of subsequent events, which lead to smaller probabilities as they occur on larger scales, leads to a converging sum of terms, this then leads to the reduction of the dimensions by 2, so $d_c=6$ and $D=6$ for $d \geq 6$ (where, throughout this paper, D is the dimension of the connected components of the MST). In the next section, we obtain the same results by much more secure methods on the BL, and a plausible application to Euclidean space also.

The behavior of the probability that two vertices are on the same connected component of the spanning forest is exactly what is needed to count the significantly different ground states of the NS model that are visible in a window. The number of large connected components intersecting the window will scale as W^{d-D} , which by the above scaling argument, and the results below, will be W^{d-6} for $d \geq 6$; thus $\# = d-6$ (thus the use of $\nu_{\text{perc}}(d)=1/2$ in the above scaling argument is justified self-consistently). We mention here that the authors of Ref. [26], who considered the path exponent D_p but not D or $\#$ for MSTs, also stated that the critical dimension for MSTs is 6, apparently because they were using Prim’s algorithm, and its connection to invasion percolation, for which again $d_c=6$.

III. BETHE LATTICE

In this section we define and solve the MST problem on the BL with wired boundary conditions. The BL can be motivated by the desire to find a “mean-field” theory, in which only the mean effect of neighbors of a vertex is included, while effects of correlations that propagate around cycles of the lattice are neglected. The BL fulfils these requirements as it possesses no cycles, but is still homogeneous due to the constant coordination number (degree). For the MST on the BL, we calculate the expectation value of the number of vertices that are connected to the origin within radius m . We show how to define and analyze other correlation functions also. The BL solution forms the basis for a mean-field theory defined on a finite-dimensional lattice in the next section. The BL results imply that the mean-field fractal dimension of paths on the MST in Euclidean space is 2, and the fractal dimension of connected components is 6. This establishes that $d_c=6$ by the argument given above.

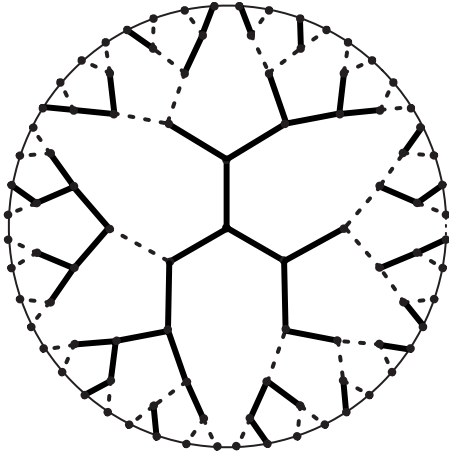


FIG. 1. A sample realization of the MST (solid lines) on the BL with wired boundary condition. The graph shown has $\sigma=2$ and radius $M=5$. Vertices on the boundary are connected with boundary edges of vanishing cost, shown here as thin lines. All vertices in the interior are connected to the boundary only once.

A. Preliminaries

The BL is a tree of degree $z=\sigma+1$ at every vertex, except for the “leaves” at the boundary, which are all a “distance” (distance means minimum number of steps along the tree required to go from one vertex to another) M from the central vertex, which we call the origin. We use the term BL interchangeably to refer to both a finite graph of radius M and the graph obtained in the $M \rightarrow \infty$ limit. Since the MST on a finite tree graph will clearly be the whole tree, we must use wired boundary conditions, as discussed in the previous section, in order to obtain an interesting spanning forest. We will see that the limit of such a forest does exist, at least in terms of its local statistical properties, at finite distances from the origin (hence, far inside the boundary); this is the “weak limit” which has been discussed in Ref. [41]. We expect that the local properties of this forest in the limit mimic the local properties of the MST in Euclidean space in a mean-field sense.

The wired boundary condition means that we connect all vertices on the boundary of the BL to each other with additional “boundary edges” of vanishing cost, so that they are always connected for $p_0 > 0$, as shown in Fig. 1. In practical terms, this means that when we run the Kruskal process, once a connected component (or cluster) of the forest touches the boundary, no edge that would connect it to the boundary by a distinct path can be added later. When the process is run up to $p_0=1$, we obtain a forest spanning the BL, which becomes a tree if the boundary edges are included.

We will use the correspondence with bond percolation discussed in Sec. II. Percolation on the BL has been studied thoroughly [36,51]. The most basic question of interest is to find the probability that a vertex in the interior is connected to the boundary, as a function of the probability p for each edge to be (independently) occupied. A convenient quantity to look at is the probability $F_M(p)$ that a given vertex is *not* connected to the boundary at distance M via a path of edges

through a given outward branch. This obeys the recurrence relation

$$F_{M+1}(p) = 1 - p + pF_M(p)^\sigma, \quad (12)$$

with $F_0(p)=0$. This arises because either (a) the first edge along the branch is unoccupied, which occurs with probability $1-p$, or (b) the first edge is occupied (with probability p) and there is no connection to the boundary through the σ branches further toward the boundary; these two alternatives are disjoint. As the radius M of the lattice goes to ∞ , all vertices a fixed distance from the origin are far from the boundary, and the probability of any of them not being connected to the boundary along a particular outward branch approaches a limit $\lim_{M \rightarrow \infty} F_M(p) = F(p)$, which is given by a *stable* fixed point with $0 \leq F \leq 1$ of the recurrence, that is

$$F(p) = 1 - p + pF(p)^\sigma. \quad (13)$$

This has the trivial solution $F=1$, and its stability is given by linearizing the recurrence about this solution; the eigenvalue of the linearized recurrence for F_M-1 is σp . Thus for $p < 1/\sigma$ the solution $F=1$ is stable. In this regime, no interior vertex is connected to infinity, with probability one. The value $p_c=1/\sigma$ is called the percolation threshold. For $p > p_c$, the solution $F=1$ is unstable, and the stable solution is a nontrivial solution to the fixed-point condition, Eq. (13). The existence, uniqueness, and stability of, and the convergence to, these fixed points is proved in, for example, Ref. [52] (in different notation). For $\sigma=2$, this solution can be found explicitly,

$$F(p) = (1-p)/p \quad (14)$$

for $p > 1/2$ ($\sigma=2$). For general σ , we can expand in powers of $p-p_c$, and find that $1-F(p) = a(p-p_c) + O[(p-p_c)^2]$ as $p \rightarrow p_c$ from above, where $a=2\sigma/(\sigma-1)$. [As $p \rightarrow 1$, $F(p) \rightarrow 0$.] The probability that any vertex is connected to infinity (and hence is on an infinite spanning cluster) is then $P_\infty(p) = 1 - F^{\sigma+1}(p)$, which is zero for $p < p_c$, and turns on with a discontinuous derivative at p_c . This behavior defines the critical exponent $\beta=1$ via $P_\infty(p) \sim (p-p_c)^\beta$ for $p \rightarrow p_c$ from above, which illustrates that percolation is a critical phenomenon with discontinuous properties at p_c .

B. Correlation functions

In using Kruskal’s algorithm to calculate the MST on the BL, we need to keep track of when different vertices become connected to the boundary as p_0 is raised from 0 to 1. As an example, consider the probability $P_{(0;m)}(p_0)$ that the path on the MST from the origin to the boundary has formed by time p_0 , that is when all edges of cost $\leq p_0$ have been tested, and passes through a given vertex a distance m from the origin. (When we discuss such paths, we always mean a path that does not backtrack on itself.) We may consider this as done in a finite system, and we always assume that the limit $M \rightarrow \infty$ is unproblematic. Indeed, we will see that the only ingredient involving properties all the way out to the boundary is just the probability $F_M(p)$ from the percolation problem, which is already known to have the appropriate limit. Note that $P_{(0;m)}(p_0)$ can also be described as the probability that on

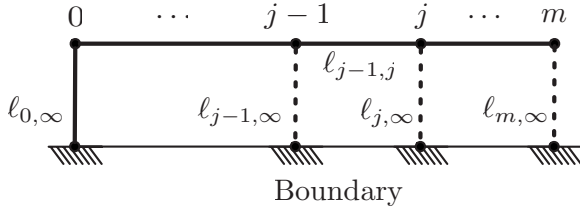


FIG. 2. Construction of a path on the BL starting from the origin (labeled m), passing through the vertex labeled 0, and connecting to the boundary, here shown as a shaded line. For this path to lie on the MST, the edge costs must obey the set of inequalities given in Eq. (15). Note that each vertical line (both solid and dashed) stands for a subtree, not just a single edge, so this diagram depicts the entire BL.

the MSF(p_0) on the infinite BL there is a path to infinity passing through two specified vertices at separation m .

For later convenience, let us label the vertices so that m is the origin, and 0 is the vertex at distance m that the path on the MST passes through. To begin with, consider $m > 0$. For every vertex j ($j = 1, \dots, m-1$) along the path from the origin to our chosen vertex, there is a set of paths on the BL connecting j to the boundary through any of the $\sigma-1$ side branches leaving the path from 0 to m . We define $\ell_{j,\infty}$ as the minimum over this set of the maximum edge cost encountered along each of these paths (see Fig. 2). In the model with ℓ for each edge uniformly distributed in $[0,1]$, $\ell_{j,\infty}$ is the value of p at which a connection to infinity is first formed along any of these $\sigma-1$ side branches from vertex j in the bond percolation process. Similarly, for $j=0$ ($j=m$), define $\ell_{0,\infty}$ ($\ell_{m,\infty}$) as the minimum cost at which connection to infinity is formed along any of the σ branches other than the path to m (0) in the percolation process. Then the desired probability can be written as

$$P_{(0,m)}(p) = \Pr\left[\bigwedge_{j=1}^m \bigwedge_{i=0}^j (\ell_{j,\infty} > \ell_{i-1,i}) \wedge \bigwedge_{j=0}^m (\ell_{j-1,j} < p)\right], \quad (15)$$

where \wedge denotes a logical ‘and’ (conjunction) and we simplify the notation by defining $\ell_{-1,0} \equiv \ell_{0,\infty}$. Notice that in forming the minimum cost tree connected to infinity, it is immaterial whether $\ell_{i-1,i} > \ell_{j,\infty}$ for $i > j$, provided that $\ell_{j',\infty} > \ell_{i-1,i}$ for $j' \geq i$. This is the crucial fact in setting up a recurrence relation for $P_{(0,m)}(p)$. Finally for $m=0$, we must make a special definition. The most natural is that no direction for the path is specified, and $P_{(0,0)}(p)$ is defined to be $P_{(0,0)}(p) = P_{\infty}(p) = 1 - F(p)^{\sigma+1}$.

If the number of side branches from vertex m ($m > 0$) was the same as those at vertices $j=1, \dots, m-1$, namely, $\sigma-1$, the probability $P_{(0,m)}(p)$ would obey a recurrence relation. Let us define $P'_{(0,m)}(p)$ to be defined the identical way with this modification. (This corresponds to an alternative definition of the BL that is sometimes used, in which the origin alone in the tree has degree σ ; it is used because it simplifies the use of recurrence relations in a similar way as here.) Then Eq. (15) can be explicitly constructed as a recurrence relation, most easily by working with the derivatives

$$\Phi_j(p) = \frac{d}{dp} P'_{(0;j)}(p). \quad (16)$$

The initial condition for the recurrence is

$$\Phi_0(p) = \frac{d}{dp} (1 - F(p)^\sigma). \quad (17)$$

The costs $\ell_{j,\infty}$ for $j=0, \dots, m$, and $\ell_{j-1,j}$ for $j=1, \dots, m$ are all statistically independent (note that $\ell_{m,\infty}$ now refers to a collection of $\sigma-1$ branches, like the others). The probability that $\ell_{j,\infty} < p$ is $1 - F^{\sigma-1}(p)$ for $j=1, \dots, m$. The probability that $\ell_{j-1,j} < p$ is p . The recurrence is then given by

$$\begin{aligned} \Phi_j(p) &= F(p)^{\sigma-1} \left[p \Phi_{j-1}(p) + \int_0^p dp' \Phi_{j-1}(p') \right] \\ &= F(p)^{\sigma-1} \frac{d}{dp} \left[p \int_0^p dp' \Phi_{j-1}(p') \right]. \end{aligned} \quad (18)$$

The factor of $F^{\sigma-1}(p)$ is the probability that the $\sigma-1$ side branches at vertex j are not connected to infinity by p . In the bracket in the first line of Eq. (18), the two terms correspond to the two cases that the most expensive edge connecting the vertex j to infinity on the path, which has cost p , either is not or is the edge connecting vertices $j-1$ and j , respectively. In the first case the latter edge is already occupied (with probability p), and connection occurs at p somewhere in the remainder of path, with probability density described by $\Phi_{j-1}(p)$, while in the second the edge $j-1, j$ is the one that becomes occupied at p , and the probability must be multiplied by the probability $P'_{(0;j-1)}(p)$ that the rest of the path is already formed. We write the recurrence in terms of a kernel K ,

$$\Phi_j(p) = \int_0^1 dp' K(p, p') \Phi_{j-1}(p'), \quad (19)$$

$$\begin{aligned} K(p, p') &= F(p)^{\sigma-1} [\theta(p-p') + p \delta(p-p')] \\ &= F(p)^{\sigma-1} \frac{d}{dp} [p \theta(p-p')], \end{aligned} \quad (20)$$

where $\theta(x)$ is the usual step function.

The kernel $K(p, p')$ in the recurrence has the meaning of a conditional probability density. It is the probability density (in the p variable) that j is first connected to infinity along the specified path passing through $j-1$ at value p (and is not connected to infinity along any side branches), given that $j-1$ is first connected to infinity in the specified manner at value p' . It is clear that this must vanish if $p < p'$. We note that if the factors $F^{\sigma-1}$ are omitted, then we obtain the corresponding conditional probability relevant to percolation, in which connections to infinity along side branches are allowed, instead of that relevant to MSTs. Similar interpretations apply to the iterates of K that we consider next.

The recurrence relation (18), and its generalizations, allow us to compute correlation functions on the BL. Iteration of the recurrence requires iterated integrals of K ,

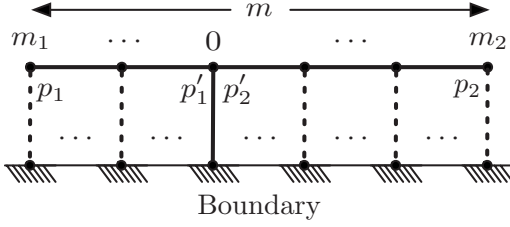


FIG. 3. The set of paths contributing to $P_{(0;m_1,m_2)}(p_0)$, drawn using the same conventions as Fig. 2. For $C^{(2)}(m;p_0)$, we sum over all $m_1 \geq 0, m_2 \geq 0$ such that $m_1 + m_2 = m$.

$$\begin{aligned} K^{*m}(p_1, p_{m+1}) &\equiv K * K \cdots * K(p_1, p_{m+1}) \\ &= \int_0^1 dp_2 \cdots \int_0^1 dp_m K(p_1, p_2) \cdots K(p_m, p_{m+1}). \end{aligned} \quad (21)$$

The explicit step and δ functions in each factor of K guarantee that $p_1 \geq p_2 \geq \cdots \geq p_{m+1}$. From the definition we then have

$$P_{(0;m)}(p_0) = \int_0^{p_0} dp dp' F(p) K^{*m}(p, p') \Phi_0(p'), \quad (22)$$

where the factor $F(p)$ in the final integration restores the correct number of branches at the origin.

We can obtain multipoint correlation functions by introducing branching into the path via additional chains of K 's. For example, the probability that two points at distances $m_1 > 0, m_2 > 0$ from a vertex that we label 0 are both connected to infinity through 0 by time $p = p_0$ is given by (see Fig. 3)

$$\begin{aligned} P_{(0;m_1,m_2)}(p_0) &= \int_0^{p_0} dp_1 dp_2 dp'_1 dp'_2 F(p_1) F(p_2) K^{*m_1}(p_1, p'_1) \\ &\quad \times K^{*m_2}(p_2, p'_2) \Phi_{(0;0,0)}(p'_1, p'_2), \end{aligned} \quad (23)$$

The initial distribution is

$$\Phi_{(0;0,0)}(p, p') = \delta(p - p') \frac{d}{dp} [1 - F^{\sigma-1}(p)], \quad (24)$$

because the connection of 0 to infinity must not use either of the two side branches leading to m_1 and m_2 . We then define the two-point correlation function $C^{(2)}(m, p_0)$ as the probability that two vertices at separation m are connected to each other and to infinity by p_0 . For $m > 0$, it is given by

$$C^{(2)}(m; p_0) = \sum_{m'=1}^{m-1} P_{(0;m',m-m')}(p_0) + 2P_{(0;m)}(p_0). \quad (25)$$

The term $2P_{(0;m)}(p_0)$ covers the cases where the path to infinity on the tree from one vertex passes through the other. For $m=0$, we define $C^{(2)}(0; p_0) = P_{(0;0)}(p_0) = P_{\infty}(p_0)$. This construction of the two-point correlation function can be immediately extended to multipoint correlations, giving the probability that, by time p_0 , several specified vertices are on the same component of $\text{MSF}(p_0)$ and connected to infinity.

These will be used in Sec. III E to analyze the geometry of the trees.

C. Analysis of the iterated kernel

To make further progress, we need to analyze the behavior of the iterated kernel K^{*m} . First, for $P_{(0;m)}(p_0)$ there is a simple result: substituting Eq. (17) into Eq. (18) and using Eq. (13) shows that $\Phi_1(p) = \Phi_0(p)/\sigma$. That is, Φ_0 is a right eigenfunction of K with eigenvalue $1/\sigma$. Then $\Phi_j(p) = \Phi_0(p)/\sigma^j$ for $j=1, \dots, m-1$. Then finally we obtain

$$P_{(0;m)}(p_0) = \frac{[1 - F(p_0)^{\sigma+1}]}{(\sigma+1)\sigma^{m-1}} = \frac{P_{\infty}(p_0)}{(\sigma+1)\sigma^{m-1}}. \quad (26)$$

In particular, if we put $p_0=1$ so that all edges have been tested, then $P_{(0;m)}(1) = \sigma^{-(m-1)}/(\sigma+1)$. The meaning of these results should be clear: due to the isotropy of the BL, the path from the origin to infinity [which exists at p_0 with probability $P_{\infty}(p_0)$] must pass through one of the $(\sigma+1)\sigma^{m-1}$ vertices a distance m away, and all are equally probable. Thus the path is an isotropic random walk on the BL (with no backtracking). In Sec. III F, we will interpret this result in Euclidean space as a random walk with fractal dimension $D_p=2$.

All the right eigenfunctions of K can also be obtained. We will require them to be integrable, in particular at $p=0$. The eigenvalue equation for K , $\int_0^1 dp_2 K(p_1, p_2) v_{\lambda}(p_2) = \lambda v_{\lambda}(p_1)$, becomes a differential equation for $V_{\lambda}(p) = \int_0^p dp' v_{\lambda}(p')$,

$$(\lambda - pF(p)^{\sigma-1}) \frac{d}{dp} V_{\lambda}(p) = F(p)^{\sigma-1} V_{\lambda}(p). \quad (27)$$

Locally in p , the general solution of the differential Eq. (27) is

$$V_{\lambda}(p) = C \exp \int^p dp' \frac{F(p')^{\sigma-1}}{\lambda - p'F(p')^{\sigma-1}} \quad (28)$$

(C is a constant). From the definition, V_{λ} must obey the initial condition $V_{\lambda}(0)=0$, and be continuous at all p . Excluding the solution $V_{\lambda}(p)$ identically zero on $[0,1]$, the continuity of $V_{\lambda}(p)$ implies that $V_{\lambda}(0)=0$ is only possible if λ obeys $\lambda = pF(p)^{\sigma-1}$ for some $p = p_{\lambda} \geq p_c$, and $V_{\lambda}(p)=0$ for $p < p_{\lambda}$ (the solution for $p_{\lambda} \geq p_c$ to $\lambda = pF(p)^{\sigma-1}$ is unique). This means that λ lies in the interval $(0, 1/\sigma]$. As p_{λ} is approached from above, these solutions have a power-law behavior,

$$V_{\lambda}(p) \sim (p - p_{\lambda})^{\alpha_{\lambda}+1} \quad (29)$$

as $p \rightarrow p_{\lambda}$ from above. Using the behavior of $F(p)$ near p_c , we find $\alpha_{\lambda} \sim -2\sigma(p_{\lambda} - p_c)$ as $p_{\lambda} \rightarrow p_c$ from above. Also, $\alpha_{\lambda} \rightarrow -1$ as $p_{\lambda} \rightarrow 1$ ($\lambda \rightarrow 0$), and α_{λ} always lies between 0 and -1 . Thus the eigenfunctions v_{λ} are non-negative and integrable for all $\lambda \in (0, 1/\sigma]$. For $\lambda = 1/\sigma$, $p_{\lambda} = p_c = 1/\sigma$, and $v_{1/\sigma}(p) \propto \Phi_0(p)$.

To make progress with higher-order correlation functions such as Eq. (23), we first introduce the generating function (a discrete Fourier-Laplace transform), for complex w ,

$$\tilde{\Phi}_w(p) = \sum_{j=0}^{\infty} w^j \Phi_j(p). \quad (30)$$

Inserting this into Eq. (19) and summing over j yields

$$\int_0^1 dp' [\delta(p-p') - wK(p,p')] \tilde{\Phi}_w(p') = \Phi_0(p). \quad (31)$$

In order to solve this integral equation, we find the resolvent operator (or Green's function) g_w such that

$$\int_0^1 dp_2 [\delta(p_1-p_2) - wK(p_1,p_2)] g_w(p_2,p_3) = \delta(p_1-p_3). \quad (32)$$

The formal solution of Eq. (32) is simply

$$g_w(p_1,p_2) = \delta(p_1-p_2) + \sum_{j=1}^{\infty} w^j K^{*j}(p_1,p_2). \quad (33)$$

Because the set of eigenvalues of K is the interval $(0, 1/\sigma]$, the series cannot be expected to converge if $|w| > \sigma$. This will be directly confirmed later. From the structure of K , we expect in general that g_w is zero for $p_1 < p_2$. Because of the δ -function term in K , g_w contains a δ -function term as well as a smooth piece. For $|w| < \sigma$, the δ -function part is clearly $g_w(p_1,p_2) = (1-wp_1F(p_1)^{\sigma-1})^{-1} \delta(p_1-p_2) + \dots$, where the omitted parts are ordinary functions. For w real and in the interval $[\sigma, \infty)$, the coefficient of the δ -function blows up at $p_1=p_2=1/w$ and at $p_1=p_2=p_{1/w}$, in terms of the values p_λ related to the eigenvalues λ of K as above.

Equation (32) can be converted into a differential equation by defining $G_w(p_2,p_3) = \int_0^{p_2} dp'_2 g_w(p'_2,p_3)$; then

$$(1-wp_1F^{\sigma-1}(p_1)) \frac{d}{dp_1} G_w(p_1,p_2) - wF^{\sigma-1}(p_1) G_w(p_1,p_2) = \delta(p_1-p_2). \quad (34)$$

For $p_1 \leq p_c$ we have $F(p_1)=1$, and Eq. (34) is solved [using the initial condition $G_w(0,p_2)=0$] by

$$G_w(p_1,p_2) = \frac{\theta(p_1-p_2)}{1-wp_1}, \quad (35)$$

which vanishes if $p_2 > p_c$.

The full solution to Eq. (34), valid for all p_1 and p_2 , is obtained by solving the equation without the δ -function term for $p_1 > p_2$ as in Eq. (27) above (with $\lambda=1/w$), and choosing the constant to obtain the correct discontinuous behavior at $p_1=p_2$ (because of the δ -function). Because of the initial condition $G_w(0,p_2)=0$, we impose the requirement that G_w must vanish for $p_1 < p_2$ for all p_2 . The result is

$$G_w(p_1,p_2) = \frac{\theta(p_1-p_2)}{1-wp_2F^{\sigma-1}(p_2)} \times \exp \int_{p_2}^{p_1} dp' \frac{wF(p')^{\sigma-1}}{1-wp'F(p')^{\sigma-1}}. \quad (36)$$

Then g_w is obtained as $g_w(p_1,p_2) = \partial G_w(p_1,p_2) / \partial p_1$. Both g_w and G_w can be seen to be complex analytic for $w \notin [\sigma, \infty)$. g_w

exhibits singular behavior if w is real and in $[\sigma, \infty)$ (as expected), and nowhere else, which means that the spectrum of K is precisely the interval $(0, 1/\sigma]$. For $w \notin [\sigma, \infty)$, there is no solution of the homogeneous equation obeying the initial condition that could be added to the solution. These results agree with the earlier statements for the two regimes $|w| < \sigma$ and $p_1 < p_c$.

To extract the large m behavior of K^{*m} , we write

$$K^{*m}(p_1,p_2) = \frac{1}{2\pi i} \oint dw \frac{g_w(p_1,p_2)}{w^{m+1}}, \quad (37)$$

where the contour is a small circle around the origin. Because g_w is analytic in w except on $[\sigma, \infty)$, we can increase the radius of the contour until we hit the first singularity in w , which is located on the real axis at the value w_c defined by $1-w_c p_2 F^{\sigma-1}(p_2)=0$. Parameterizing the contour as $w = w_c e^{-i\tau}$, the most strongly divergent piece of the integrand at this singularity will determine the large- m behavior of K^{*m} via

$$\int_{-\pi}^{\pi} d\tau \frac{e^{im}}{(i\tau)^z} = \frac{2\pi m^{z-1}}{\Gamma(z)} + im^{z-1} e^{-i\pi z} \Gamma(1-z; -im\pi) - im^{z-1} e^{i\pi z} \Gamma(1-z; im\pi), \quad (38)$$

where $\Gamma(z; t_0) = \int_{t_0}^{\infty} dt e^{-t} t^{z-1}$ is the incomplete gamma function. The first term on the right-hand side of Eq. (38) comes from the part of the contour coming in from $\tau=0^+ + i\infty$, encircling the singularity, and returning to $\tau=0^- + i\infty$. If z is not an integer, there will be a branch cut from $\tau=0$ to $\tau=+i\infty$, but the same answer is obtained regardless. The last terms come from the segments of the contour running from $\tau=-\infty$ to $\tau=-\pi$ and $\tau=\pi$ to $\tau=\infty$; using the asymptotic form of the incomplete gamma function we see that

$$im^{z-1} (e^{-i\pi z} \Gamma(1-z; -im\pi) - e^{i\pi z} \Gamma(1-z; im\pi)) \sim \frac{2}{m\pi^z} \sin\left(\frac{2\pi m + \pi(1-z)}{2}\right) + O(1/m^2), \quad (39)$$

and so these terms may be neglected relative to the first as $m \rightarrow \infty$.

In order to make progress in closed form we specialize to $\sigma=2$. Using Eq. (13), $F(p)=(1-p)/p$ for $p > p_c=1/2$ and

$$\exp \int_{p_2}^{p_1} dp' \frac{wF(p')^{\sigma-1}}{1-wp'F(p')^{\sigma-1}} = \left\{ \frac{p_2^w [1-w(1-p_1)]}{p_1^w [1-w(1-p_2)]} \right\}^{1/w-1}, \quad (40)$$

taking $p_1 > p_2 > p_c$. Defining $h_w(p) = 1-w(1-p)$ for brevity,

$$g_w(p_1,p_2) = \frac{d}{dp_1} \left\{ \frac{\theta(p_1-p_2)}{h_w(p_2)} \left[\frac{p_2^w h_w(p_1)}{p_1^w h_w(p_2)} \right]^{1/w-1} \right\}. \quad (41)$$

It might appear that Eq. (41) has an essential singularity at $w=1$, but in fact

$$\lim_{w \rightarrow 1} g_w(p_1,p_2) = \frac{d}{dp_1} \left[\frac{\theta(p_1-p_2)}{p_1} e^{(1/p_1)-(1/p_2)} \right], \quad (42)$$

and the singularity with smallest $|w|$ is located at $w_c=1/(1-p_2)$; g_w may be rewritten as

$$g_w(p_1, p_2) = \frac{d}{dp_1} \frac{\theta(p_1 - p_2)}{(w_c - w)^{w/(w-1)}} \left(\frac{p_2^w h_w(p_1)}{p_1^w (1 - p_2)^w} \right)^{1/(w-1)}. \quad (43)$$

Because $1/2 \leq p_2 \leq 1$, we know $1 \leq w_c/(w_c - 1) \leq 2$ and the factor of $(w_c - w)$ is responsible for all divergences as $w \rightarrow w_c$. To find the leading-order divergence of this factor, we use

$$(w_c - w)^{-(w/w-1)} = (w_c - w)^{-(w_c/w_c-1)} + O[(w_c - w)^{-(w_c/w_c-1)+1} \log(w_c - w)], \quad (44)$$

which, with $w = w_c e^{-i\tau}$, diverges as $(i\tau)^{-1/p_2}$. We may now use Eq. (38) to evaluate

$$K^{*m}(p_1, p_2) = \frac{1}{2\pi w_c^m} \int_{-\pi}^{\pi} d\tau e^{i\tau m} g_{w_c e^{-i\tau}}(p_1, p_2). \quad (45)$$

Because of cancellations that take place in Eq. (41) when $p_1 = p_2$, we must take the derivative with respect to p_1 explicitly before the $m \rightarrow \infty$ limit. The result is

$$\begin{aligned} K^{*m}(p_1, p_2) &\sim \delta(p_1 - p_2) (1 - p_2)^m \\ &+ \theta(p_1 - p_2) \frac{(1 - p_1)(1 - p_2)}{p_1(p_1 - p_2)^2} \\ &\times \left(\frac{p_2(p_1 - p_2)}{p_1(1 - p_2)} \right)^{1/p_2} \frac{(1 - p_2)^m m^{1/p_2-1}}{\Gamma(1/p_2)}. \end{aligned} \quad (46)$$

D. Asymptotics of the two-point function and mass

In this section we complete the calculation of the important correlation function $C^{(2)}$. We present two versions: an essentially exact version for $\sigma=2$, and an asymptotic calculation valid for all σ in the region $p_0 - p_c$ small.

If we make no further approximations, when we calculate correlation functions such as Eq. (25) it is clearly easier to integrate over p_2 first and then find the large- m behavior rather than attempt to integrate Eq. (46) directly over p_2 . As an example we now calculate the asymptotic behavior of $C^{(2)}(m; p_0)$. Inserting $\delta_{m, m_1 + m_2}$ in Eq. (25) gives

$$\begin{aligned} C^{(2)}(m; p_0) &= \int_0^{p_0} dp_1 dp_2 \frac{1}{2\pi w_c^m} \int_{-\pi}^{\pi} d\tau \\ &\times e^{i\tau m} \tilde{\Phi}_{w_c e^{-i\tau}, w_c e^{-i\tau}}(p_1, p_2), \end{aligned} \quad (47)$$

where

$$\begin{aligned} \tilde{\Phi}_{w_1, w_2}(p_1, p_2) &= \int_0^1 dp'_1 dp'_2 g_{w_1}(p_1, p'_1) g_{w_2}(p_2, p'_2) \\ &\times \Phi_{(0;0,0)}(p'_1, p'_2). \end{aligned} \quad (48)$$

with $\Phi_{(0;0,0)}(p'_1, p'_2) = \theta(p'_1 - p_c) \delta(p'_1 - p'_2) / p_1^2$ for $\sigma=2$. This integral is done easily,

$$\begin{aligned} \tilde{\Phi}_{w,w}(p_1, p_2) &= \frac{d}{dp_1} \frac{d}{dp_2} \frac{1}{w+1} \left(\frac{h_w(p_1) h_w(p_2)}{p_1^w p_2^w} \right)^{1/(w-1)} \\ &\times \left[\left(\frac{1}{2-w} \right)^{(w+1/w-1)} - \left(\frac{p}{h_w(p)} \right)^{(w+1/w-1)} \right], \end{aligned} \quad (49)$$

where $p = \min(p_1, p_2)$. Clearly, since the integrand in Eq. (48) has no support below p_c , we obtain a branch cut starting at $w_c = 1/p_c = 2$ and this is the maximum radius the contour in w may take. Since the only divergence as $w \rightarrow 2$ comes from the first term in the square brackets, we obtain

$$\begin{aligned} \tilde{\Phi}_{2e^{-i\tau}, 2e^{-i\tau}}(p_1, p_2) &= \frac{d}{dp_1} \frac{d}{dp_2} \frac{h_2(p_1) h_2(p_2)}{p_1^2 p_2^2} \\ &\times \left[\frac{1}{3} \left(\frac{1}{2i\tau} \right)^3 + O\left(\frac{\log \tau}{\tau^2} \right) \right]. \end{aligned} \quad (50)$$

Then, using

$$\int_{-\pi}^{\pi} d\tau \frac{e^{i\tau m}}{(i\tau)^3} \sim m^2 \pi + O(1/m) \quad (51)$$

and

$$\int_{p_c}^{p_0} dp \frac{d}{dp} \frac{h_2(p)}{p^2} = \frac{2p_0 - 1}{p_0^2}, \quad (52)$$

we finally obtain

$$C^{(2)}(m; p_0) \sim \left(\frac{2p_0 - 1}{p_0^2} \right)^2 2^{-m} \left(\frac{m^2}{48} + O(m \log m) \right). \quad (53)$$

We see that $C^{(2)}(m; p_0)$ scales as m^2 / σ^m for large m , and this behavior is obtained for *any* $p_0 > p_c$.

Next we present an alternative calculation valid for all σ when $p_0 - p_c$ is small. We will define $\delta p = p - p_c$ (but e.g., $\delta(p_1 - p_2)$ is a δ -function as usual). We then have

$$pF(p)^{\sigma-1} \simeq p_c - \delta p, \quad (54)$$

valid as $p \rightarrow p_c$ from above. Then, using Eq. (36), G_w can be calculated for $w \notin [\sigma, \infty)$. Because all the singular behavior as $w \rightarrow \sigma$ arises from the denominator of the integrand of Eq. (36), we may approximate $F^{\sigma-1}(p) \simeq 1$ in the numerator and

$$G_w(p_1, p_2) = \theta(p_1 - p_2) \frac{\delta p_1 + 1/w - p_c}{w(\delta p_2 + 1/w - p_c)^2} \quad (55)$$

for $\delta p_1, \delta p_2$ both small and positive. Then if we consider the transform

$$\tilde{C}_w^{(2)}(p_0) = \sum_{m=0}^{\infty} w^m C^{(2)}(m, p_0) \quad (56)$$

we notice that the transform of the sum over m_1 in Eq. (25) becomes simply a product of g_w s inside the integral in Eq. (23), and we neglect the factors $F(p_1), F(p_2)$ as these do not affect the leading m dependence. Further we can neglect $2P_{(0;m)}$ as it falls off faster than the term we keep. Also

$\Phi_{(0;0,0)}(p_1, p_2) = 2\sigma\delta(p_1 - p_2)$ for p_1 above p_c , and zero below. Finally, we will estimate the expected mass inside radius m ,

$$\overline{M(m, p_0)} = C^{(2)}(0; p_0) + \sum_{m'=1}^m (\sigma + 1)\sigma^{m'-1} C^{(2)}(m', p_0) \quad (57)$$

directly, as this is similar to the definition of the transform of $C^{(2)}$: we simply evaluate the transform at $w = \sigma(1 - 1/m)$, which cuts off the sum at around m . This is a value at which G_w is not singular. Thus we have to calculate

$$\begin{aligned} \tilde{C}_w^{(2)}(p_0) &= 2\sigma(\delta p_0 + 1/w - p_c)^2 \\ &\times \int_0^{\delta p_0} d\delta p' \frac{1}{w^2(\delta p' + 1/w - p_c)^4}. \end{aligned} \quad (58)$$

For fixed p_0 , the dominant contribution comes from the lower limit, and contains $(1 - wp_c)^{-3}$ times factors that go to constants as $w \rightarrow \sigma = 1/p_c$. Hence the mass behaves as

$$\overline{M_{\text{soft}}(m, p_0)} \sim \frac{2}{3}\sigma(\sigma + 1)\delta p_0^2 m^3 \quad (59)$$

as $m \rightarrow \infty$ [we inserted a factor $(\sigma + 1)/\sigma$ to account for the number of neighbors $\sigma + 1$ at the first step in $\overline{M(m, p_0)}$, as in Eq. (57)]. This method in fact differs from the definition above in using a soft cutoff for the sum over m' instead of a hard one, $m' \leq m$. Now that the form of the summand is known, we can evaluate it using either form of cutoff. Hence we find that for the *hard* cutoff, the result is smaller by a factor 6,

$$\overline{M(m, p_0)} \sim \frac{1}{9}\sigma(\sigma + 1)\delta p_0^2 m^3. \quad (60)$$

Thus the correlation function behaves as

$$C^{(2)}(m, p_0) \sim \frac{1}{3}(\sigma\delta p_0)^2 m^2 \sigma^{-m}, \quad (61)$$

which agrees with the result for $\sigma = 2$.

Equations (60) and (61) are the main results of this section. From the structure of the expression for G_w , the lower limit always dominates, so this m dependence holds for all $p_0 > p_c$, at sufficiently large m . More precisely, the results are valid only if δp_0 is greater than order $1/m$. This means that m is much larger than the correlation length at p_0 , which is proportional to $1/(\delta p_0)$. The main contribution to the integral is from $\delta p'$ less than of order $1/m$. This is in agreement with the ‘‘superhighways’’ idea [23–25].

Notice also that the factor δp_0^2 is present in both results because the probability that one vertex is connected to infinity is $P_\infty(p_0) \simeq 2\sigma(\sigma + 1)\delta p_0/(\sigma - 1)$, for δp_0 small and positive, and so is $\propto \delta p_0^2$ for two vertices (the two events are uncorrelated, because the two points are separated by more than the correlation length). The dependence on σ is the same within subleading terms of relative order $1/\sigma$.

E. General k th-order correlation functions and moments of the mass

We may extend the method to estimate asymptotics of correlation functions of any order k , that is the probability $C(i_1, \dots, i_k; p_0)$ that some given set of vertices (labeled i_1, i_2, \dots, i_k) are on the same connected component of $\text{MSF}(p_0)$ and that this component is infinite. The procedure is straightforward: to compute the k -point correlation function for a given set of k distinct vertices, one draws the smallest subtree of the BL such that all k given vertices are connected, and choose any vertex on this subtree as the root point, along which the connection to infinity occurs in the $\text{MSF}(p_0)$ (eventually, we will sum over the possible root points). Thus the leaves of the subtree (i.e., the degree 1 vertices) must be among the given k vertices, but if any of the k given vertices are not leaves, they can be anywhere on the subtree. Starting from the root, we propagate out to (or possibly through) each of the k given vertices, along the subtree. The subtree can be viewed as made of chains of edges connected by degree-2 vertices, with the ends of the chains at either (i) the root point, which has degree ≥ 1 , (ii) the leaves of the subtree, or (iii) vertices of degree > 2 other than the root point. For each such chain e of m_e steps, we associate the iterated kernel $K^{*m_e}(p_i, p_j)$, where the labels i, j are associated to the two ends of the chain, with i the end further from the root point. For the initial distribution at the root, if there are n chains leaving it ($n \leq \sigma$), we generalize $\Phi_{(0;0,0)}$ to

$$\Phi_{(0;0,\dots,0)}(p_1, \dots, p_n) = \left[\frac{d}{dp_1} (1 - F^{\sigma+1-n}(p_1)) \right] \prod_{j=2}^n \delta(p_1 - p_j). \quad (62)$$

Similarly, by comparing $dP_\infty(p_0)/dp_0$ with $\Phi_0(p')$ and $\Phi_{(0;0,0)}(p'_1, p'_2)$, we see that at a vertex of the subtree of degree $n \neq 2$, we must associate with it a factor

$$v_n(p_1, p_2, \dots, p_n) = F^{2-n}(p_1) \prod_{j=2}^n \delta(p_1 - p_j). \quad (63)$$

After multiplying together all these factors, we must integrate over all the parameters like p_i between the limits 0 and p_0 . There are two of these parameters for each chain on the subtree; clearly some could be eliminated using the δ functions. Finally, we must sum over all possible root points on the tree. This procedure yields $C(i_1, \dots, i_k; p_0)$ (unlike our earlier description for the $k=2$ case, there are no exceptions to this prescription for cases of vertices coinciding with each other or with the root point).

The higher-point correlation functions can be used to calculate higher moments of the mass, $\overline{M(m, p_0)^k}$. These are the average of the k th power of the sum over positions at distance less than m from the origin of the ‘‘indicator function’’ that is one if and only if the vertex is on the same connected component of $\text{MSF}(p_0)$ as the origin. $\overline{M(m, p_0)^k}$ is equal to the sum of the $k+1$ -point correlation function $C(i_1, \dots, i_{k+1}; p_0)$ over all positions of i_2, \dots, i_{k+1} within m steps of the origin at i_1 .

For m large, the largest contribution to $\overline{M(m, p_0)^k}$ will come from configurations of i_l , ($l=1, \dots, k+1$) for which, in the subtree in the calculation of $C(i_1, \dots, i_{k+1}; p_0)$, all the given vertices are at its leaves, the root point has degree 2, and the vertices of degree >2 have degree 3. For these there are $2k$ chains (iterated kernels) in the subtree. We can estimate the power of m as $m \rightarrow \infty$ using the same approximations as for $k=1$. The sums over position are estimated by using the propagator g_w in place of all K^{*m_e} 's, with $w = \sigma(1 - 1/m)$ in each one. The factors at the vertices of the subtree (other than δ -functions) can be dropped, at least when δp_0 is small (and for larger δp_0 do not affect the scaling behavior). The integrals over the p_i associated with the leaves can be done by using G_w in place of g_w for these chains. The remaining integrals over p_i 's associated with the other vertices and the root point are dominated by the lower limit $\delta p_i = 0$, and can be estimated by power counting. As each additional leaf on the subtree leads to an extra factor $G_w g_w$ and one additional integral similar to that at the root (as for $k=1$ above), this yields finally (neglecting constant factors)

$$\overline{M(m, p_0)^k} \sim m^{3k} \sim [\overline{M(m, p_0)}]^k. \quad (64)$$

As all k th moments scale like the k th power of the first moment, this means that the (random) mass $M(m, p_0)$ does not have a very broad distribution, and its typical behavior is well-described by its expected value. Hence the connected components of MSF(p_0) on the BL are not multifractal.

F. Fractal dimensions

So far we have developed a method for computing correlation functions on the BL, while our real interest is in lattices in Euclidean space of dimension d . For sufficiently high d , we would expect that a mean-field theory holds for quantities such as exponents; this assertion will have to be justified *post-hoc* in a subsequent paper, by a perturbation analysis of corrections due to fluctuations neglected in the mean-field theory. The BL results provide the mean-field theory results, once we have explained how to convert them to apply to Euclidean space.

For a hypercubic lattice on Euclidean space, if we choose a path (starting from the origin) randomly (with equal probability for each), then it behaves as a random walk, and after m steps will be of order \sqrt{m} in Euclidean distance from the origin. In the set of all paths from the origin, any two paths initially coincide but ultimately part company. If we neglect the possibility that they subsequently intersect (and also that a path may intersect itself, including by backtracking), then the union of the paths forms a tree, equivalent to the Bethe lattice with $z=2d$. Hence in this correspondence, separations on the BL behave like distances *squared* on the Euclidean lattice [51],

$$m \sim r^2 \quad (65)$$

which allows us to infer mean-field scaling dimensions from the BL theory. More formally, to apply the Bethe lattice results as an approximation for the lattice, in Eq. (18) we must now sum over all neighbors of the given site, since the path may go through the site in any direction. Equation (30) needs to be replaced by a Fourier transform

$$\tilde{\Phi}_{\mathbf{k}}(p) = \sum_{\mathbf{x}} e^{i\mathbf{k}\cdot\mathbf{x}} \Phi_{\mathbf{x}}(p), \quad (66)$$

which means that in Eq. (31) and all subsequent equations we make the substitution

$$w \rightarrow \sum_{\hat{n}_j} e^{-i\mathbf{k}\cdot\hat{n}_j} = 2d - \mathbf{k}^2 + O(|\mathbf{k}|^4), \quad (67)$$

where $\{\hat{n}_j\}$ are the basis vectors of the hypercubic lattice. This leads to the same relation (65).

We established that the probability the path from any vertex to infinity passes through a given vertex a distance m away behaves as $P_{(0;m)}(p_0) \sim \sigma^{-m}$; summing this over all sites within a ball of radius m on the BL gives

$$P_{(0;0)}(p_0) + \sum_{m'=1}^m (\sigma+1)\sigma^{m'-1} P_{(0;m')}(p_0) \sim m. \quad (68)$$

Then we expect that in Euclidean space, the mass of the path lying within radius R scales as $M_p(R, p_0) \sim R^2$, consistent with the picture of this path (mentioned earlier for the BL) as a random walk, with dimension $D_p=2$. This is the same as the dimension of the backbone of a critical percolation cluster for $d \geq d_c=6$ [36].

Similarly, the mass of the component connected to the origin on the BL, $M(m, p_0) \sim m^3$ becomes $M(R, p_0) \sim R^6$ within radius R in Euclidean space, meaning that a connected component of the MST has a mean-field fractal dimension $D=6$. Because the union of the spanning trees fills the lattice, this strongly suggests that the critical dimension of the MST will be $d_c=6$, as discussed in Sec. II E. This is the same critical dimension as for percolation (at threshold). We emphasize again that the result is valid for p_0 greater than of order $1/m \sim 1/R^2$. Since the correlation length ξ behaves as $|p_0 - p_c|^{-\nu_{\text{perc}}}$, with $\nu_{\text{perc}}=1/2$ for $d > 6$, this means it holds for $R > \xi$, and thus involves distance scales at which ordinary correlations for percolation decay exponentially.

G. Poisson-weighted infinite tree

Another mean-field model for MST (and other random optimization problems), which mimics the continuum model in Euclidean space, is based on the Poisson-weighted infinite tree (PWIT) [33]. This is a tree with infinite degree at each vertex, and the weights (or costs) ℓ for the edges emanating from each vertex are given by a Poisson process on $\ell \geq 0$ with density $\rho(\ell) \propto \ell^{d-1}$. This is the same as the measure on the set of distances between the points of the uniform Poisson process on \mathbb{R}^d . Thus the model is obtained by taking the continuum model, and ignoring all correlations induced by the geometry of space, keeping only the probability distribution for separations of points.

The PWIT can be viewed as the limit $\sigma \rightarrow \infty$ of the iid BL model with degree $\sigma+1$ we have used above, with a different probability distribution on the edges (as we have seen, the choice of this does not affect the geometry of the MST). For $\sigma+1$ neighbors, we can cut off the distribution with density $\rho(\ell)$ at $\ell = O[(\sigma+1)^{1/d}]$, then study the finite-degree trees in the limit. For each $d \geq 1$ in this model, the percolation

threshold ℓ_c is nonzero, and because the behavior is dominated by edges close to threshold, the results for the PWIT will be in the same universality class as the BL model above. Indeed, if we set $\ell = (\sigma + 1)p$, the BL model as $\sigma \rightarrow \infty$ coincides with the $d=1$ PWIT. Notice that the limit of our expressions exists, because by writing $F(p) = 1 - G(\ell)/\sigma$, then $F^{\sigma-1} \rightarrow e^{-G}$, and the fixed-point equation for F becomes [34]

$$G(\ell) = \ell(1 - e^{-G(\ell)}). \tag{69}$$

In the transforms, we should also set $s = w/\sigma$. Then the limit of our theory makes sense for the masses M, M_p within m steps of the origin: $\bar{M} \sim \frac{2}{3} \delta \ell^2 m^3$ (again, $\delta \ell = \ell - \ell_c$, where here $\ell_c = 1$). For the PWIT one would naturally wish to express such quantities in terms of the distance ℓ defined as the sum of the ℓ_e 's along a path on the tree, and since most edges accepted are either below, or not far above, ℓ_c , these masses $M(\ell), M_p(\ell)$ scale the same way, and again $\ell \sim R^2$ because the paths are random walks. When this measure of distance is used, the number of connected components inside ℓ is also well behaved. Note that because it is believed that continuum percolation is in the same universality class as bond percolation, we would expect universal properties of the continuum MST to be the same as those of the lattice model anyway, so that all of these conclusions are consistent.

H. Random, locally treelike graphs

The results of the previous sections admit a simple generalization to certain classes of random graphs. We consider a tree where the coordination number of each vertex is an iid random variable distributed according to $\rho(\sigma)$.

In what follows, averages with respect to ρ are denoted by angle brackets. The fixed-point equation for $F(p)$ becomes

$$F(p) = 1 - p + p \sum_{\sigma=0}^{\infty} \rho(\sigma) F(p)^\sigma. \tag{70}$$

Again, $F(p)$ is defined as the smallest solution at fixed p . In what follows we assume $1 < \langle \sigma \rangle < \infty$, which are the conditions necessary for the graph to admit a conventional percolation transition: $F(p) = 1$ for $p \leq p_c = 1/\langle \sigma \rangle$, $F(p) < 1$ for $p > p_c$, and we can construct a nontrivial ensemble of MSTs under wired boundary conditions at infinity.

At this point we find it helpful to multiply the kernel defined in Eq. (20) by a factor of σ . This has the effect of summing the connectedness functions over lattice sites [done in, e.g., Eq. (57)] simultaneously with averaging over edge costs. This modification allows us to relate the spectrum and eigenfunctions of the random graph kernel $\langle \sigma K \rangle$ back to those obtained above: e.g., $\langle \Phi_0(p) \rangle$ is an eigenfunction of $\langle \sigma K \rangle$ but *not* of $\langle K \rangle$.

The entire calculation goes through as before, with occurrences of $F(p)^{\sigma-1}$ in the BL Green's function replaced by

$\langle \sigma F(p)^{\sigma-1} \rangle$ and the transform variable w rescaled by $1/\langle \sigma \rangle$. Under the additional assumption $\langle \sigma^2 \rangle < \infty$, one may repeat the asymptotic analysis of Sec. III D and obtain

$$\overline{\langle M_{\text{soft}}(m, p_0) \rangle} \sim \frac{2}{3} \langle \sigma(\sigma + 1) \rangle \delta p_0^2 m^3. \tag{71}$$

The generalization extends to the higher moments of the cluster mass discussed in Sec. III E.

One may also consider ‘‘quenched’’ moments of cluster masses, of the form

$$\mathcal{M}_{k,\ell}(m, p_0) \equiv \langle [\overline{M(m, p_0)^k}]^\ell \rangle \tag{72}$$

for $\ell > 1$. These quantities cannot readily be calculated with the techniques discussed above and are beyond the scope of this paper.

IV. CONCLUSION

In this paper, we have achieved the following results. For a finite graph, we defined a process $\text{MSF}(p_0)$ which is a random forest that becomes the MST for $p_0 = 1$. Using the Bethe lattice (BL) with wired boundary conditions, and taking the infinite size limit, we showed that the infinite connected components of $\text{MSF}(p_0)$ on the BL contain of order m^3 vertices within m steps of any vertex on this component, for any p_0 greater than the value p_c of the threshold for bond percolation on the BL. This result is essentially rigorous. Transferring it (heuristically) to Euclidean space, this means that the mass of an infinite connected component of $\text{MSF}(p_0)$ within a ball of radius R scales as R^D with $D=6$, for d sufficiently large and p_0 greater than the value p_c of the threshold for bond percolation on the lattice used. This then implies that for $d > 6$ there are of order R^{d-6} large connected components that intersect such a ball. We also gave a non-rigorous second argument for these results, using scaling ideas (this argument directly addresses the critical dimension d_c above which the results hold). The results also hold (rigorously) for the Poisson-weighted infinite tree, and (heuristically) for the continuum MST model in Euclidean space.

Following the reasoning of NS [8,9], these results for the MST imply that the strongly disordered spin-glass model has an uncountable number of ground states for $d > 6$, of which of order $2^{O(R^{d-6})}$ can be distinguished within a ball of radius R . For $d \leq 6$, the logarithm of the number of ground states is smaller than any power of R , and possibly only of order one, or simply one (with probability one).

ACKNOWLEDGMENTS

We would like to thank D. Stein and C. M. Newman for helpful discussions. This work was supported by NSF Grant No. DMR-0706195.

- [1] E. Lawler, *Combinatorial Optimization: Networks and Matroids* (Dover, Mineola, NY, 2001), Chap. 7.
- [2] C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity* (Dover, Mineola, NY, 1998), Chap. 12.
- [3] R. E. Tarjan, *Data Structures and Network Algorithms* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 1983), Chap. 6.
- [4] W. J. Cook, W. H. Cunningham, W. R. Pulleyblank, and A. S. Schrijver, *Combinatorial Optimization* (Wiley Interscience, New York, NY, 1998), Chap. 2 and 8.
- [5] J. M. Steele, *Probability Theory and Combinatorial Optimization* (SIAM, Philadelphia, PA, 1997), Chap. 5.
- [6] M. Cieplak, A. Maritan, and J. R. Banavar, Phys. Rev. Lett. **72**, 2320 (1994).
- [7] M. Cieplak, A. Maritan, and J. R. Banavar, Phys. Rev. Lett. **76**, 3754 (1996).
- [8] C. M. Newman and D. L. Stein, Phys. Rev. Lett. **72**, 2286 (1994).
- [9] C. M. Newman and D. L. Stein, J. Stat. Phys. **82**, 1113 (1996).
- [10] T. S. Jackson and N. Read, Phys. Rev. E **81**, 021131 (2010).
- [11] J. Cardy, *Scaling and Renormalization in Statistical Physics* (Cambridge University Press, Cambridge, 1996).
- [12] D. Aldous and J. M. Steele, Probab. Theory Relat. Fields **92**, 247 (1992).
- [13] N. Read, Phys. Rev. E **72**, 036114 (2005).
- [14] J. B. Kruskal, Proc. Am. Math. Soc. **7**, 48 (1956).
- [15] V. Jarník, Práce Moravské Přírodovědecké Společnosti **6**, 57 (1930).
- [16] R. C. Prim, Bell Syst. Tech. J. **36**, 1389 (1957).
- [17] E. W. Dijkstra, Numer. Math. **1**, 269 (1959).
- [18] A. A. Middleton, Phys. Rev. B **61**, 14787 (2000).
- [19] A.-L. Barabási, Phys. Rev. Lett. **76**, 3750 (1996).
- [20] E. López, S. V. Buldyrev, L. A. Braunstein, S. Havlin, and H. E. Stanley, Phys. Rev. E **72**, 056131 (2005).
- [21] S. V. Buldyrev, S. Havlin, and H. E. Stanley, Phys. Rev. E **73**, 036128 (2006).
- [22] M. Porto, N. Schwartz, S. Havlin, and A. Bunde, Phys. Rev. E **60**, R2448 (1999).
- [23] Z. Wu, L. A. Braunstein, S. Havlin, and H. E. Stanley, Phys. Rev. Lett. **96**, 148702 (2006).
- [24] L. A. Braunstein, Z. Wu, Y. Chen, S. V. Buldyrev, T. Kalisky, S. Sreenivasan, R. Cohen, E. López, S. Havlin, and H. E. Stanley, Int. J. Bifurcation Chaos Appl. Sci. Eng. **17**, 2215 (2007).
- [25] Z. Wu, E. López, S. V. Buldyrev, L. A. Braunstein, S. Havlin, and H. E. Stanley, Phys. Rev. E **71**, 045101(R) (2005).
- [26] R. Dobrin and P. M. Duxbury, Phys. Rev. Lett. **86**, 5076 (2001).
- [27] B. Wieland and D. B. Wilson, Phys. Rev. E **68**, 056101 (2003).
- [28] D. B. Wilson, Phys. Rev. E **69**, 037105 (2004).
- [29] *Spin Glass Theory and Beyond*, edited by M. Mézard, G. Parisi, and M. Virasoro (World Scientific, Singapore, 1987).
- [30] Y. Fu and P. W. Anderson, J. Phys. A **19**, 1605 (1986).
- [31] C. M. Newman and D. L. Stein, J. Phys.: Condens. Matter **15**, R1319 (2003).
- [32] O. Motrunich, S.-C. Mau, D. A. Huse, and D. S. Fisher, Phys. Rev. B **61**, 1160 (2000).
- [33] D. Aldous, Probab. Theory Relat. Fields **93**, 507 (1992).
- [34] D. Aldous, Random Structures Algorithms **18**, 381 (2001).
- [35] G. Grimmett, *Percolation*, 2nd ed. (Springer, New York, NY, 1999), Chap. 10.
- [36] D. Stauffer and A. Aharony, *Introduction to Percolation Theory*, 2nd ed. (Taylor and Francis, London, 1994).
- [37] M. Aizenman, Nucl. Phys. B **485**, 551 (1997).
- [38] K. S. Alexander, Ann. Probab. **23**, 87 (1995).
- [39] R. Pemantle, Ann. Probab. **19**, 1559 (1991).
- [40] M. Aizenman, A. Burchard, C. M. Newman, and D. B. Wilson, Random Struct. Algorithms **15**, 319 (1999).
- [41] R. Lyons, Y. Peres, and O. Schramm, Ann. Probab. **34**, 1665 (2006).
- [42] J. A. Hartigan, J. Am. Stat. Assoc. **76**, 388 (1981).
- [43] B. Chazelle, J. ACM **47**, 1028 (2000).
- [44] A. M. Frieze, Discrete Appl. Math. **10**, 47 (1985).
- [45] R. Lenormand and S. Bories, C. R. Seances Acad. Sci., Ser. B **291**, 279 (1980).
- [46] R. Chandler, J. Koplick, K. Lerman, and J. F. Willemsen, J. Fluid Mech. **119**, 249 (1982).
- [47] D. Wilkinson and J. F. Willemsen, J. Phys. A **16**, 3365 (1983).
- [48] J. T. Chayes, L. Chayes, and C. M. Newman, Commun. Math. Phys. **101**, 383 (1985).
- [49] F. Barahona, J. Phys. A **15**, 3241 (1982).
- [50] C. H. Papadimitriou, *Computational Complexity* (Addison-Wesley, Reading, MA, 1994).
- [51] M. E. Fisher and J. W. Essam, J. Math. Phys. **2**, 609 (1961).
- [52] K. B. Arthreya and P. E. Ney, *Branching Processes* (Springer-Verlag, Berlin, 1972); *Branching Processes* (Dover, Mineola, NY, 2004), Chap. 1, Pt. A.