Tree-Based Models for Random Distribution of Mass

David Aldous *
Department of Statistics
University of California
Berkeley CA 94720

Abstract

A mathematical model for distribution of mass in d-dimensional space, based upon randomly embedding random trees into space, is introduced and studied. The model is a variant of the *superBrownian motion* process studied by mathematicians. We present calculations relating to (i) distribution of position of typical mass element (ii) moments of the center of mass (iii) large deviation behavior (iv) a recursive self-similarity property.

Key words. Spatial distribution, random tree, superBrownian process, large deviations, recursive self-similarity.

Journal of Statistical Physics **73** (1993) 625-641.

*Research supported by N.S.F. Grants DMS90-01710 and DMS-92-24857 and by the Miller Institute for Basic Research in Science

1 Introduction

To start with an analogy, it has long been accepted that the mathematically fundamental model for a quantity varying randomly but continuously with time is (mathematical) Brownian motion (alternatively called the Wiener process or integrated white noise). This arises in contexts as diverse as diffusion of particles, stock market fluctuations, noise in electrical networks and neutral genetic theory. By contrast, there is no accepted fundamental model for random continuous "distribution of mass" in d-dimensional space. Such processes also arise in many contexts: classical spatial statistics [5, 18, 27] deals with examples such as the distribution of the total population of a species over its geographical range; physicists study randomly-growing aggregates of particles [28, 29]. We seek a model for the continuous limit of such processes, where we rescale the total number of particles, population size, etc to be 1 unit and refer to this as *unit mass.* Although many models appear in the books cited above, these tend to be not only specific to the particular phenomenon under study but also difficult to analyze mathematically, except via Monte Carlo simulation. The purpose of this paper is to describe a model "ISE" which seems the most fundamental from the mathematical viewpoint. A direct and combinatorial description will be given in section 2, emphasizing that apparently different discrete constructions have the same continuous limit process ISE. The main content of the paper is section 3 in which we present calculations relating to

- distribution of position of typical mass element
- moments of the center of mass
- large deviation behavior
- a recursive self-similarity property.

ISE is defined using random branching structures. There are many examples of naturally-occurring high density branching structures (a referee suggests lightning patterns, neural networks, river systems, human arterial networks). Unlike Brownian motion, ISE does not seem directly applicable as a realistic model of any specific such real-world phenomenon, but may prove useful as a theoretical building-block for constructing more realistic models.

As described in section 4 our model may be considered as a variant of the superBrownian process which has been studied intensely by mathematical probabilists in recent years (ISE is an acronym for *integrated superBrownian excursion*).

2 Abstract trees and their embeddings into d-space

The material here concerning abstract trees is presented in greater detail in [2].

2.1 Discrete trees

A discrete tree consists of a finite number n of vertices and $n-1$ edges, such that any pair of vertices is linked by a unique path of distinct edges. Regarding edges as having length 1, the distance $d(x,y)$ between vertices x and y is the number of edges on the linking path. We often distinguish one vertex called the *root*. These are "abstract" trees, in that the vertices are not given positions in space. But we can embed a tree into d-space by sending the root to the origin 0 and either
(i) regarding each edge as a step of the form $(0,\ldots,0,\pm1,0,\ldots,0)$, so that each vertex is sent to a vertex of the lattice Z^d; or
(ii) regarding each edge as a vector in R^d of unit length, so that each vertex gets sent to a point in d-space R^d.
In either case, different vertices of the abstract tree may be sent to the same point in d-space. By putting mass $1/n$ on each vertex of the abstract tree we get an abstract unit mass distribution, and then an embedding gives a unit mass distribution in d-space.

We now introduce randomness into both the abstract tree-structure and the embedding. It is easy to specify the latter: for each edge we make an independent random choice, either (case (i)) uniformly over the $2d$ possible directions, or (case (ii)) uniformly over all directions in d-space. Below are three different (at first sight) models of random n-vertex abstract trees.

Combinatorial models. Here we assume all n-vertex trees to be equally likely. Because there are different conventions about when two trees are to be regarded as "the same", and because we may impose restrictions on allowable degrees of vertices, there are quite a number of different models of this type.

Conditioned branching processes. Consider a population process, starting with one individual, in which each individual has a random number of offspring (mean 1, variance $0<\sigma^2<\infty$). This process has some random total population size: conditional on this total population size being n, the family tree is a random n-vertex tree.

Combinatorial aggregation. Start with n vertices and no edges. Repeatedly add an edge, chosen uniformly at random from the edges whose addition would not create a cycle. After $n-1$ edges have been added, a random n-vertex tree has been created.

Although these constructions look different, it is known that combinatorial models (with certain conventions) are exactly the same as conditioned branching process models (with certain offspring distributions). See e.g. [2] sec. 2.1. And there is compelling evidence ([1] and an interchange-of-limits argument) that, in the $n \to \infty$ limits we are concerned with here, the combinatorial aggregation model coincides with the other models.

2.2 Continuum trees

The previous section described random discrete mass distributions in d-space, and our ISE model will be the continuous rescaled limit of these. But mathematically it is more useful to interchange the order of embedding / taking continuous limits. That is, we first describe an abstract "continuum random tree" (CRT) and then it is easy to describe how to embed it randomly into d-space.

Figure 1 illustrates part of an abstract continuum tree (the reader should mentally add more and more shorter and shorter edges). In such a tree, interior vertices are spread continuously along edges, and there is a unique path (of length $d(x,y)$, say) between any pair x, y of vertices. We also have "unit mass" spread around the vertices, with the mass concentrated on the leaves rather than the interior points of paths.

The particular model of random continuum tree we use, the *Brownian CRT*, was discussed at length in [2], and has several alternative descriptions and constructions. (Keep in mind it is an *abstract* tree.) Up to scaling constants discussed in section 2.3, it is the limit of the random discrete abstract tree models of section 2.1, when edges are scaled to have length $n^{-1/2}$. An intrinsic description is via the spanning subtree $\mathcal{R}_k$ spanned by the root and $(X_1, \ldots, X_k)$, where the (X_i) are independent uniform random vertices chosen from a realization of the Brownian CRT. See figure 2, ignoring the y's for now. This tree has a "shape" $\hat{t}$ and $2k-1$ edge-lengths (l_i), and p.d.f. (probability density function)

$$f(\hat{t}; l_1, \ldots, l_{2k-1}) = (\sum_{i=1}^{2k-1} l_i) \exp(-(\sum_{i=1}^{2k-1} l_i)^2/2). \qquad (1)$$

For discussion of this fundamental formula from the present viewpoint see [2] eq. (13), [3] Lemma 21; see also [16] for a different approach. The "shape" records which existing edge a new vertex and its edge are attached to. Since there are $2j-1$ choices of edge to which the edge to x_{j+1} may be attached, the number of different shapes equals

$$\prod_{j=1}^{k-1}(2j-1). \tag{2}$$

We will use (1) as a starting point for calculations in section 3. We mention that there are two other descriptions of the Brownian CRT: one gives a sequential procedure for growing the tree by adding branches to the existing tree, and the other is a construction in terms of Brownian excursion. Note that expression (1) really is a *probability* density, i.e. integrating over the (l_i) and summing over the $\hat{t}$ gives 1 exactly. Note also that (1) does *not* assert that the total edge-length has Rayleigh density $l\exp(-l^2/2)$.

To embed an abstract continuum tree into d-space, take a sequence of leaves $(x_1, x_2, \ldots)$ dense in the tree. For each k, the root and the leaves $(x_1, \ldots, x_k)$ determine $k-1$ interior branchpoints $(b_1, \ldots, b_{k-1})$ and $2k-1$ edges (e_i) of lengths (l_i), say. Embed the tree into d-space by sending the root to 0 and replacing each edge e_i with a Brownian motion sample path of duration l_i. Figure 3 sketches the embedding of the leaves $(x_1, \ldots, x_5)$ in figure 1 (genuine two-dimensional Brownian motion is much more tangled than the sketch shows).

The object of our study is the Brownian CRT, embedded into d-space as described above. The resulting random distribution of unit mass in d-space we call ISE and write as $\mathcal{M}$. Associated with ISE is its *support* $\mathcal{S}$, the (random) smallest closed set containing all the unit mass (it turns out that $\mathcal{S}$ is a connected set). Figure 4 shows two independent realizations of ISE, by picturing $1,000$ points sampled from each. Note that the tree-structure is hardly visible in two dimensions.

2.3 ISE as a limit of embedded discrete trees

ISE appears as a rescaled limit of embedded discrete trees, as follows. Consider the "conditioned branching process" model of random n-vertex abstract trees, with parameter σ^2 representing offspring variance. Embed into d-space, by making the edges become independent random vectors L with mean 0 and covariance matrix $\theta^2 I_d$, where I_d is the identity matrix. Then the resulting mass distribution is asymptotically ISE, scaled by a factor

$$\theta \sigma^{-1/2} n^{1/4}. \tag{3}$$

Note that the exponent of n is $1/4$, not $1/2$. Note also that in the simplest models of embedding mentioned in section 2.1 we have $\theta = d^{-1/2}$. Finally, note that the different models of random n-tree mentioned in section 2.1 will have the same behavior for the appropriate value of σ.

To explain (3), note first that scaling the abstract Brownian CRT by c is equivalent to scaling the embedded ISE by $c^{1/2}$, by the usual Brownian scaling law. By [3] Theorem 23 the abstract n-tree approximates the Brownian CRT scaled by $\sigma^{-1} n^{1/2}$, so that (in the case $\theta = 1$) the embedded discrete tree approximates the embedded continuum tree, which by the scaling property above is ISE scaled by $(\sigma^{-1} n^{1/2})^{1/2}$. The case of general θ follows by another use of Brownian scaling.

3 Mathematical properties

3.1 Joint distributions

One starting point for mathematical analysis is an expression for the joint distribution of the positions $(X^{(1)}, \ldots, X^{(k)})$ of k independent randomly-picked mass elements from ISE. If these are at positions $(x_1, \ldots, x_k)$, there there is an associated tree-structure illustrated in figure 2. There is an abstract tree whose leaves are embedded at $(x_1, \ldots, x_k)$, and whose branchpoints are embedded somewhere: write (y_i) for the embedded positions of the leaves and branchpoints, and (l_i) for the lengths of edges in the abstract tree, and $\hat{t}$ for the "shape" of the abstract tree. Combining (1) and the d-dimensional standard Normal density gives the following joint p.d.f. for the shape, the abstract edge-lengths and the embedded positions of leaves and branchpoints.

$$f(\hat{t}; l_1, \ldots, l_{2k-1}; y_1, \ldots, y_{2k-1}) = (2\pi)^{-(2k-1)d/2} (\sum_e l(e))$$

$$\times \exp\left(-\frac{1}{2} (\sum_e l(e))^2 - \frac{1}{2} \sum_e \frac{(\Delta y(e))^2}{l(e)} \right). \tag{4}$$

Here for each edge e, $l(e)$ denotes its length in the abstract tree and $\Delta y(e) = |y_i^* - y_j^*|$ is the distance between endpoints after embedding. Thus the joint density of $(X^{(1)}, \ldots, X^{(k)})$ at $(x_1, \ldots, x_k)$ can in principle be obtained by summing over $\hat{t}$ and integrating over the $(l_i; y_i)$ for which the positions of the leaves are (x_i).

We first consider the case $k = 1$, i.e. the position $X = (X_1, \ldots, X_d)$ of a single randomly-chosen mass element from ISE. Here (4) can be collapsed to

$$X = H^{1/2} Z \tag{5}$$

where H, the distance in the abstract tree from a randomly-chosen vertex to the root, has p.d.f. $f_H(h) = he^{-h^2/2}$, where Z has standard Normal distribution in d dimensions, and where H and Z are independent. We now list some immediate consequences of (5).

(a) Consistency between dimensions, i.e. the distribution of $(X_1, \ldots, X_j)$ is the same for all $d \geq j$.

(b) X_i and X_j are uncorrelated but not independent.

(c) X_1 has a symmetric distribution whose even moments are

$$EX_1^{2n} = \pi^{-1/2} 2^{3n/2} \Gamma(n + \frac{1}{2}) \Gamma(\frac{n}{2} + 1), \ n \geq 1. \tag{6}$$

(d) X has the spherically symmetric p.d.f.

$$f_X(x) = (2\pi)^{-d/2} \int_0^\infty h^{1-\frac{d}{2}} \exp\left(-\frac{1}{2}\left(h^2 + \frac{|x|^2}{h}\right)\right) \ dh. \tag{7}$$

We do not have any simple explicit formula for the integral in (7), even in the case $d = 1$ where the moments are given by (6). However, by Laplace's method (i.e. expanding the integrand about its maximum) we obtain the asymptotics

$$f_X(x) \sim 2^{-\frac{d+1}{3}} \ 3^{1/2} \ \pi^{\frac{1-d}{2}} |x|^{\frac{2-d}{3}} \exp(-3 \cdot 2^{-5/3} |x|^{4/3}) \text{ as } |x| \to \infty. \tag{8}$$

As an illustration of calculations based upon (4), consider the center of mass $S = (S_1, \ldots, S_d)$ of ISE. (S is a random point in R^d). S has spherically symmetric distribution, and in principle we can compute even moments of S_1. The second moment is more easily calculated by a symmetry argument – see (13) below. The fourth moment is

$$ES_1^4 = \int_0^\infty \ldots \int_0^\infty (45 l_1^2 + 57 l_1 l_2)(\sum_{i=1}^7 l_i) \exp\left(-(\sum_{i=1}^7 l_i)^2/2\right) \ dl_1 \ldots dl_7 \approx 1.40 \tag{9}$$

Unfortunately the method soon becomes impractical, due to combinatorial explosion in both the first term of the integrand and the dimension of the integral.

Here is the calculation for (9). First observe

$$ES_1^4 = E(X^{(1)}X^{(2)}X^{(3)}X^{(4)}) \tag{10}$$

where the $X^{(j)}$ are independent choices of mass elements from the same realization of ISE. Consider the associated tree-structure as in figure 2. Associated with each edge e is an increment $\Delta Y(e)$, and the X_j are the sums of the $\Delta Y(e)$ along the path from the root to x_j. Expanding (10), the only terms with non-zero expectation are those where each $\Delta Y(e)$ has even power, so the contribution to (10) from the tree-shape $\hat{t}$ in figure 2 is

$$E(\Delta Y(e_1)^4 + \quad 4 \text{ terms of form } E(\Delta Y(e_i))^2(\Delta Y(e_j))^2.$$

Writing $\Delta Y(e_i) = l_i^{1/2} Z_i$ for standard Normal Z_i, the contribution is

$$3l_1^2 + \quad 4 \text{ terms of form } l_i l_j. \tag{11}$$

There are 15 tree-shapes $\hat{t}$ (c.f. (2)), of which 12 are topologically the same as figure 2 and the other 3 are a different topological type, for which the contribution (11) has 3 instead of 4 cross-terms. Combining (11) with the density (1) leads to (9).

3.2 Re-rooting symmetry

One of the combinatorial models of n-vertex trees appearing in section 2.1 is the "uniform random rooted labeled unordered tree", where one starts with n distinguishable vertices then picks uniformly from the n^{n-2} ways to join the vertices into a tree, and then picks uniformly a vertex to be the root. This obviously has the property of "random re-rooting invariance", i.e. that choosing a new root uniformly from the vertices doesn't alter the distribution. This property extends to the Brownian CRT and then to ISE, using the time-reversal property

$$(B_t : 0 \leq t \leq t_0) \stackrel{d}{=} (B_{t_0 - t} - B_{t_0} : 0 \leq t \leq t_0)$$

of Brownian motion. Explicitly, the property for ISE is as follows. From a realization of $\mathcal{M}$ pick uniformly a mass element X, translate the realization so that X is moved to 0, and call the resulting mass distribution $\mathcal{M}^*$. Then (unconditionally) $\mathcal{M}^*$ has the same distribution as $\mathcal{M}$. This immediately implies, in 1 dimension, that the random amount U of mass in $[0,\infty)$ satisfies

$$U \text{ has uniform distribution on } [0,1]. \tag{12}$$

It also enables us to do a simple calculation of second moments for the center of mass S. Writing X as above for a randomly-chosen mass element,

$$E|X|^2 = E|S|^2 + E|X-S|^2 = 2E|S|^2$$

where the second equality uses the random re-rooting property and the first equality is the elementary variance identity, that for any mass distribution μ with center of mass s

$$\int |x|^2 \mu(dx) = |s|^2 + \int |x-s|^2 \mu(dx).$$

Now using (6)

$$E|X|^2 = dEX_1^2 = d\sqrt{\pi/2}$$

implying

$$E|S|^2 = \frac{d}{2}\sqrt{\pi/2}. \tag{13}$$

A few more distributional results about ISE can be deduced from distributional results in [2] about the Brownian CRT, but let us not exhaust the reader. To mention a simple open question, in $d=1$ the mass distribution is supported on some random interval $[L,R]$ containing 0, but there are no known expressions for p.d.f.'s or expectations of R or the length $R-L$.

3.3 Large Deviation behavior

ISE has an interesting large deviation behavior, stated as (16,17) below. For purposes of comparison, and because we re-use some ingredients of the proof, we first recall some known large deviation behavior of *Brownian bridge* $B^0 = (B^0_t : 0 \le t \le 1)$, that is standard d-dimensional Brownian motion conditioned on $B_1 = 0$. Fix a finite set $\mathbf{x} = \{x_1, \ldots, x_k\}$ of points in R^d. Then a precise limit result is

$$\lim_{\delta \downarrow 0} \limsup_{\varepsilon \downarrow 0} \varepsilon^{-1} \log P(\varepsilon^{1/2} B^0_t \text{ visits a } \delta\text{-neighborhood of each } x_i)$$

$$= -\frac{1}{2}(\mathrm{TSP}(\mathbf{x}))^2 \tag{14}$$

where $\mathrm{TSP}(\mathbf{x})$ is the length of the *traveling salesman path* through $\mathbf{x}$ and 0, that is the shortest path starting and ending at 0 which passes through each x_i in some unspecified order. More informally we may rewrite (14) as

$$P(\varepsilon^{1/2}B^0_t \text{ visits a neighborhood of each } x_i)$$
$$\approx \exp\left(-\frac{(\mathrm{TSP}(\mathbf{x}))^2}{2\varepsilon}\right) \text{ as } \varepsilon \to 0. \tag{15}$$

To state an analogous result for ISE, recall $\mathcal{S}$ denotes its support, and write $\varepsilon^{1/2}\mathcal{S}$ for the support after scaling d-space by $\varepsilon^{1/2}$. Then

$$\lim_{\delta\downarrow 0}\limsup_{\varepsilon\downarrow 0} \varepsilon^{-1}\log P(\varepsilon^{1/2}\mathcal{S} \text{ intersects a } \delta\text{-neighborhood of each } x_i)$$
$$= -3\cdot 2^{-5/3}(\mathrm{ST}(\mathbf{x}))^{4/3} \tag{16}$$

where $\mathrm{ST}(\mathbf{x})$ is the length of (i.e. sum of edge-lengths in) the *Steiner tree* on $\mathbf{x}$ and 0, i.e. the tree of shortest length containing $\mathbf{x}$ and 0. More informally we may rewrite (16) as

$$P(\varepsilon^{1/2}\mathcal{S} \text{ intersects a neighborhood of each } x_i)$$
$$\approx \exp\left(-\frac{3\cdot 2^{-5/3}(\mathrm{ST}(\mathbf{x}))^{4/3}}{\varepsilon}\right) \text{ as } \varepsilon \to 0. \tag{17}$$

It is important to distinguish the Steiner tree from the *minimum spanning tree*: the former is allowed to have branchpoints outside the given vertex-set $\mathbf{x}\cup\{0\}$. It is interesting that the large deviation behavior of both the abstract Brownian CRT and the embedding affect the large deviation behavior of ISE. The $\frac{4}{3}$ power law in (16,17) results from the interplay between these two behaviors.

At the rigorous level, (14) may be derived from Schilder's theorem (e.g. [8] Thm. 5.2.3.), although I do not know an explicit reference. Here is the argument in outline. Consider times $0 = t_0 < t_1 < \ldots < t_k < t_{k+1} = 1$ and points in d-space $0 = x^*_0, x^*_1, \ldots, x^*_k, x^*_{k+1} = 0$ where the (x^*_i) are some permutation of (x_i). The large deviation theorem for the multivariate Normal distribution implies

$$P(\varepsilon^{1/2}B^0_{t_i} \text{ in neighborhood of } x_i \text{ for all } i)$$
$$\approx \exp\left(-\frac{1}{2\varepsilon}\sum_{i=1}^{k+1}\frac{|x^*_i - x^*_{i-1}|^2}{t_i - t_{i-1}}\right). \tag{18}$$

Now the fundamental large deviation paradigm is that the probability of a rare event happening in some unspecified way is (up to subexponential terms) just the probability of it happening in the most likely specified way. Thus the right side of (15) is obtained by minimizing the sum in (18). Fixing the (x_i^*), the minimum over (t_i) is attained by

$$t_i - t_{i-1} = \frac{|x_i^* - x_{i-1}^*|}{A}, \quad A = \sum_i |x_i^* - x_{i-1}^*|$$

and the minimized value equals A^2. And by definition the minimum value of A over permutations (x_i^*) is $\mathrm{TSP}(\mathbf{x})$, establishing (15).

To obtain (17) requires only a slightly more elaborate argument. Given $\mathbf{x}$, use (4) and scaling to see that the probability of an embedded tree-structure as in figure 2, with each y_i in a neighborhood of x_i, is

$$\approx \exp\left(-\frac{\Gamma}{2\varepsilon}\right), \quad \text{where } \Gamma = (\sum_e l(e))^2 + \sum_e \frac{(\Delta y(e))^2}{l(e)}. \tag{19}$$

Appealing to the large deviation paradigm again, we seek the minimum value of Γ over all tree-shapes $\hat{t}$, edge-lengths $(l(e))$ and embedded branch-point positions (y_i) including leaves $\mathbf{x}$. To calculate the minimum, first fix $\hat{t}$ and (y_i) and a constant A, and first minimize over edge-lengths $(l(e))$ with $\sum_e l(e) = A$. Arguing as in the previous analysis, the minimum is attained by

$$l(e) = \Delta y(e)\ \frac{A}{B} \quad \text{where } B = \sum_e \Delta y(e)$$

and the minimized value is

$$\Gamma = A^2 + B^2/A.$$

Minimizing over A gives minimum value $\Gamma = 3 \cdot 2^{-2/3} B^{4/3}$. And by definition the minimum value of B over all trees in d-space with root 0 and leaves $\mathbf{x}$ is $\mathrm{ST}(\mathbf{x})$.

3.4 Dimension and tree-structure of ISE

A realization of $\mathcal{S}$, the support of ISE, has dimension $\min(4, d)$ in any of the usual senses of fractal dimension. This is known rigorously from superprocess results (section 4.1), but here is the intuitive explanation of "4". Choosing k vertices at random in the Brownian CRT, the maximal distance from any point of the CRT to the nearest of the k points is $k^{1/2+o(1)}$, and this order cannot be improved by any other choice of k points. The Brownian embedding of the abstract tree into d-space sends points distance δ apart in the abstract tree to points in R^d distance $O(\delta^{1/2})$ apart. So the maximum distance of any point in $\mathcal{S}$ to the nearest of k randomly-picked mass elements of ISE is $k^{1/4+o(1)}$.

It is important to distinguish the above from a different notion of dimension. If one picks two random points from a continuous density on R^d, the chance that the two points are less than ε apart is $\varepsilon^{d+o(1)}$ as $\varepsilon \to 0$. What if we pick two vertices $X^{(1)}, X^{(2)}$ from ISE? By re-rooting symmetry $X^{(2)} - X^{(1)}$ is distributed as a single random element X, and by (7) this has finite p.d.f. at 0, so

$$P(|X^{(2)} - X^{(1)}| < \varepsilon) = \varepsilon^{d+o(1)} \text{ as } \varepsilon \to 0.$$

In this sense ISE behaves as if it has dimension d.

A related question is: when is $\mathcal{S}$ itself a tree, i.e. when does $\mathcal{S}$ possess the property that there is a *unique* path between any two points? In other words, when does $\mathcal{S}$ have no self-intersections? The heuristic rule is that subsets of dimensions d_1 and d_2 in general position in R^d will not intersect if $d > d_1 + d_2$ but may intersect if $d < d_1 + d_2$. Applying this rule to different branches of $\mathcal{S}$ suggests that $\mathcal{S}$ is a tree in $d \geq 9$ and is not a tree in $d \leq 7$. These results (and the deeper fact that $\mathcal{S}$ is a tree in the critical dimension $d = 8$) also are known rigorously from the superprocess literature (section 4.1).

3.5 Recursive self-similarity

ISE has a property of being the superposition of three randomly rescaled independent copies of itself. To put this in context, recall the familiar recursive constructions of fractals such as the Cantor set or the Sierpinski gasket, which can be decomposed into (deterministically) rescaled copies of themselves. Randomized versions of such constructions have been studied in detail [17]. A somewhat different recursive self-similarity property for the Brownian CRT is given in [4] Theorem 2, and the corresponding property for ISE, which we now state, is an immediate consequence.

Say *Brownian-scaling* by c to mean scaling space by $c^{1/2}$ and scaling mass by c. Let $\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3$ be independent copies of ISE. Independently, let $(\Delta_1, \Delta_2, \Delta_3)$ have density

$$f(u_1, u_2, u_3) = (2\pi)^{-1}(u_1 u_2 u_3)^{-1/2} \text{ on } \{u_i > 0; u_1 + u_2 + u_3 = 1\}. \quad (20)$$

Define $\mathcal{M}_i^*$ to be $\mathcal{M}_i$ Brownian-scaled by Δ_i. Pick a random mass element X^* from $\mathcal{M}_1^*$. Translate $\mathcal{M}_2^*$ and $\mathcal{M}_3^*$ by the map $x \to x + X^*$, and superimpose these translated mass distributions and $\mathcal{M}_1^*$ to obtain a random mass distribution $\mathcal{M}^*$. Because $\Delta_1 + \Delta_2 + \Delta_3 = 1$ it follows that $\mathcal{M}^*$ is a *unit* mass distribution, and the recursive self-similarity property of ISE is that $\mathcal{M}^*$ is also ISE.

4 Superprocesses

We have used the phrase "random mass distribution" for what mathematicians call "random probability distribution" or "random measure". The field of *superprocesses* (the papers we cite [6, 7, 10, 11, 13, 15, 22, 23, 26] form only a small part of recent work in this field) studies random measures evolving with time, as the continuous limit of processes of particles undergoing both branching and spatial movement. The *superBrownian* case is where the spatial movement of a particle's ancestral line is a Brownian motion. The simplest set-up for superBrownian motion is to start at time 0 with unit mass at the origin: the total mass varies randomly with time until a random finite extinction time. Our model, *integrated superBrownian excursion*, features two variations on that simplest set-up.

(i) We integrate over time, to get a random measure without "time" explicitly involved.

(ii) We start at time 0 with an infinitesimal mass at the origin, and condition on the total (i.e. integrated over time) mass until extinction being exactly 1 unit.

The first variation has been discussed in the literature [6, 22, 23] but the second has not (though the somewhat related case of conditioning on *never* becoming extinct is treated in [13]). The mathematical literature is rather impenetrable and does not emphasize the kind of concrete calculations in section 3. For instance, calculations of moments via tree-diagrams are performed as means to some other end (e.g. [9, 10]) rather than as natural explicit calculations in themselves. Our combinatorial approach has advantages (e.g. the re-rooting symmetry in section 3.2, which seems more mysterious from the conventional superprocess approach) and disadvantages (e.g. the connection between superprocesses and partial differential equations [12] gets lost).

4.1 Rigorous results

The constructions of the Brownian CRT and of the superBrownian process from Brownian excursion [2, 3, 15, 16] enable us to rigorously connect the definition of ISE in section 2.2 with the superprocess definition above. There is no difficulty in seeing that the calculations in sections 3.1, 3.2 and 3.5 are rigorous. The assertions on dimension and self-intersection in section 3.4 follow from rigorous results on superBrownian motion [6, 26] because conditioning does not affect such "local behavior of sample paths" results. Our calculations in section 3.3 establish (16) as a rigorous *lower* bound on the asymptotic probability in question, but a more thorough argument would be needed to prove rigorously that it is also an upper bound.

4.2 Lattice trees

The most interesting and difficult questions of rigorous proof concern limits of random lattice trees or lattice animals. As an analogy, it has long been believed (and recently proved [19] – see [**?**] for a survey) that above the critical dimension $d = 4$ self-avoiding walk rescales to Brownian motion. It seems equally intuitively clear that random lattice trees, above the critical dimension $d = 8$, should approximate ISE scaled by $c_d n^{1/4}$. Results relating to this $n^{1/4}$ power law have been established by Hara and Slade [20, 21] using lace expansion techniques similar to those used on self-avoiding walk. But the connection with ISE has apparently not been made in the physics literature. Whether the ISE limit can be proved rigorously seems an important open question.

5 Relation to other models

The practical models mentioned in the introduction have definitions intimately tied to the geometry of d-space. On the other hand, axiomatic methods of defining random distributions of mass typically lead to the Dirichlet distributions discussed by Ferguson [14]. These are discrete distributions whose definition pays no attention to the geometry of d-dimensional space. (the occurrence of a Dirichlet distribution at (20) seems merely coincidental). Our model is intermediate between these, and might be called *semi-geometric*. Quite different tree-based models have been proposed by Mauldin et al [25] and Lavine [24], who emphasize statistical modeling issues.

5.1 Acknowledgements

Thanks to Steve Evans and Ed Perkins for pointers to the superprocess literature, and to Gordon Slade for pointers to the physics literature.

References

[1] D.J. Aldous. A random tree model associated with random graphs. *Random Structures and Algorithms*, 1:383–402, 1990.

[2] D.J. Aldous. The continuum random tree II: an overview. In M.T. Barlow and N.H. Bingham, editors, *Stochastic Analysis*, pages 23–70. Cambridge University Press, 1991.

[3] D.J. Aldous. The continuum random tree III. *Ann. Probab.*, 21:248–289, 1993.

[4] D.J. Aldous. Recursive self-similarity for random trees, random triangulations and Brownian excursion. *Ann. Probab.*, 22:527–545, 1994.

[5] N.A.C. Cressie. *Statistics for Spatial Data.* Wiley, 1991.

[6] D.A. Dawson, I. Iscoe, and E. A. Perkins. Super-Brownian motion: Path properties and hitting probabilities. *Probab. Th. Rel. Fields*, 83:135–205, 1989.

[7] D.A. Dawson and E. Perkins. Historical processes. *Mem. Amer. Math. Soc.*, 93, 1991. No. 454.

[8] A. Dembo and O. Zeitouni. *Large Deviations and Applications.* Jones and Bartlett, Boston MA, 1992.

[9] E.B. Dynkin. Representation for functionals of superprocesses by multiple stochastic integrals, with applications to self-intersection local times. In *Colloque Paul Levy sur les Processus Stochastiques*, pages 147–171, 1988. Asterisque 157-158.

[10] E.B. Dynkin. Branching particle systems and superprocesses. *Ann. Probab.*, 19:1157–1194, 1991.

[11] E.B. Dynkin. Path processes and historical superprocesses. *Probab. Th. Rel. Fields*, 90:1–36, 1991.

[12] E.B. Dynkin. A probabilistic approach to one class of nonlinear differential equations. *Probab. Th. Rel. Fields*, 89:89–115, 1991.

[13] S.N. Evans and E. Perkins. Measure-valued Markov branching processes conditioned on non-extinction. *Israel J. Math.*, 71:329–337, 1990.

[14] T. S. Ferguson. Prior distributions on spaces of probability measures. *Annals of Statistics*, 2:615–629, 1974.

[15] J.-F. Le Gall. Brownian excursions, trees and measure-valued branching processes. *Ann. Probab.*, 19:1399–1439, 1991.

[16] J.-F. Le Gall. The uniform random tree in a Brownian excursion. *Probab. Th. Rel. Fields*, 96:369–384, 1993.

[17] S. Graf, R. D. Mauldin, and S. C. Williams. The exact Hausdorff dimension in random recursive constructions. *Memoirs Amer. Math. Soc.*, 71(381):1–121, 1988.

[18] D. A. Griffin. *Advanced Spatial Statistics.* Kluwer, 1988.

[19] T. Hara and G. Slade. Critical behavior of self-avoiding walk in five or more dimensions. *Bull. Amer. Math. Soc.*, 25:417–423, 1991.

[20] V. Harison and S.N. Smirnov. Jonction maximale en distribution dans le case Markovien. *Probab. Th. Rel. Fields*, 84:491–503, 1990.

[21] K.J. Harrison and M.W. Short. The last vertex visited in a random walk on a graph. Technical report, Murdoch University, Australia, 1992.

[22] I. Iscoe. Ergodic theory and a local occupation time for measure-valued critical branching Brownian motion. *Stochastics*, 18:197–243, 1986.

[23] I. Iscoe. A weighted occupation time for a class of measure-valued branching processes. *Probab. Th. Rel. Fields*, 71:85–116, 1986.

[24] W. Lavine. Some aspects of Polya tree distributions for statistical modelling. *Annals of Statistics*, 20:1222–1235, 1992.

[25] R. D. Mauldin, W. D. Sudderth, and S. C. Williams. Polya trees and random distributions. *Annals of Statistics*, 20:1203–1221, 1992.

[26] E. Perkins. Polar sets and multiple points for super-Brownian motion. *Ann. Probab.*, 18:453–491, 1990.

[27] B.W. Silverman. *Density Estimation for Statistics and Data Analysis.* Chapman-Hall, 1986.

[28] H.E. Stanley and N. Ostrowsky, editors. *On Growth and Form: Fractal and Non-Fractal Patterns in Physics*, Dordrecht, 1986. NATO ASI series E, Applied Science no. 100, Nijhoff.

[29] T. Vicsek. *Fractal Growth Phenomena.* World Scientific, Singapore, 2nd edition, 1992.

6 Detailed calculations (not for publication)

Laplace's method gives the approximation

$$\int_0^\infty a(t)\exp(-b(t))dt \approx (2\pi b''(t_0))^{1/2} a(t_0)\exp(-b(t_0))$$

where $b(t)$ is minimized at t_0. Applying to (7) for large $x = |x|$,

$$b(t) = \frac{1}{2}(t^2 + \frac{x^2}{t})$$

$$t_0 = 2^{-1/3}x^{2/3},\ b(t_0) = 3 \cdot 2^{-5/3}x^{4/3},\ b''(t_0) = 3$$

and so the integral (7) is approximately

$$(2\pi)^{-d/2}(6\pi)^{1/2}(2^{-1/3}x^{2/3})^{1-\frac{d}{2}}\exp(-3 \cdot 2^{-5/3}x^{4/3}).$$

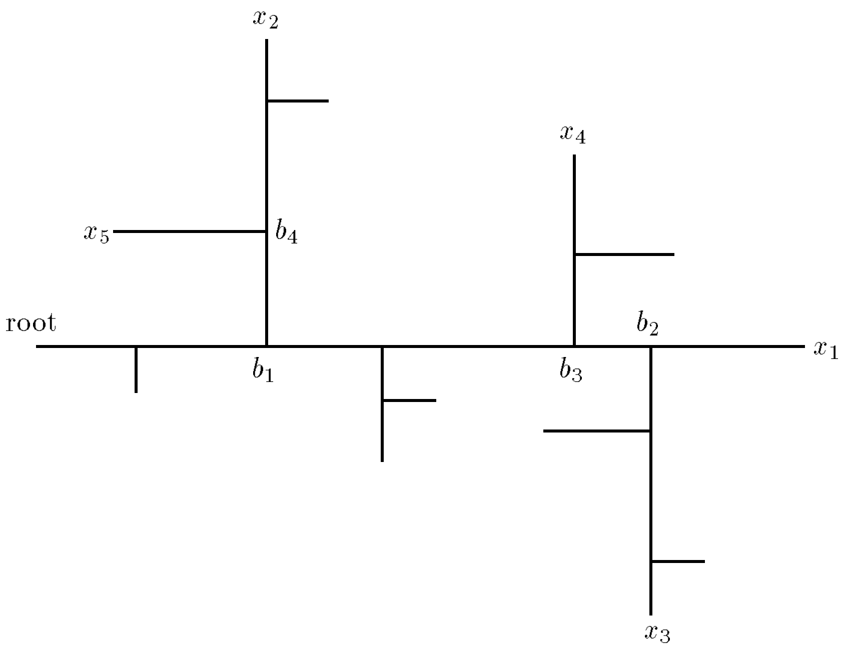

figure 1

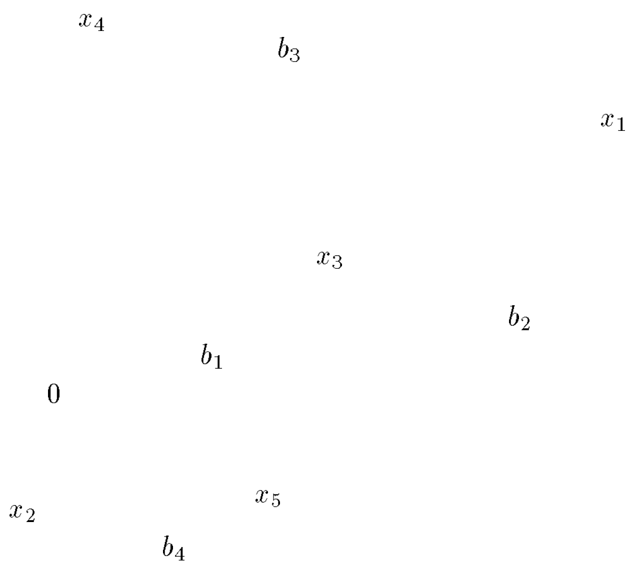

figure 3

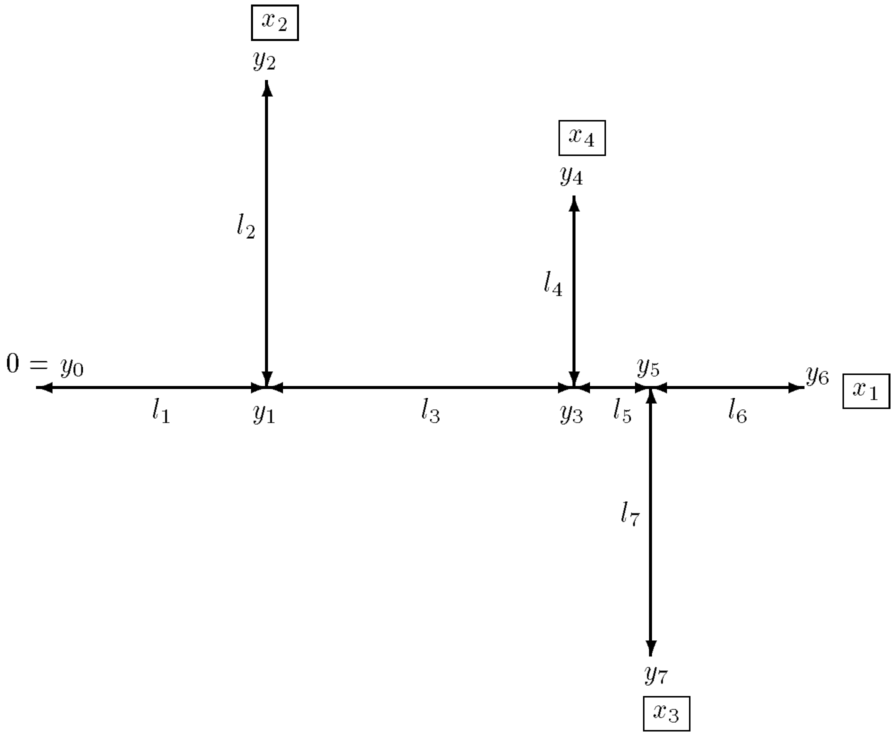

figure 2

figure 4