Scaling Limits and Super-Brownian Motion

Gordon Slade

with illustrations by Bill Casselman

Students and teachers are well aware that random data tends to lie on a bell curve. The tendency is so widespread that the term "normal distribution" has acquired the technical meaning of a distribution governed by the Gaussian density

$$(1)\quad p_t(x) = \left(\frac{d}{2\pi t}\right)^{d/2} e^{-d|x|^2/2t} \quad (x \in \mathbb{R}^d,\ t > 0).$$

For $d = 1$, the graph of $p_t(x)$ is a bell curve with breadth measured by $\sqrt{t}$, centred for convenience at 0. More generally, for all $d \geq 1$, (1) has been normalised to have variance $\int_{\mathbb{R}^d} |x|^2 p_t(x)dx = t$ and hence standard deviation $\sqrt{t}$. The mathematical explanation of why the Gaussian density is "normal" in the colloquial sense is given by the central limit theorem. Although its impact can become dulled by familiarity, the central limit theorem is a startlingly remarkable assertion. As explained in more detail below, it says that a sum of n independent and identically distributed random variables X_i with finite variance, properly rescaled, has approximately a Gaussian distribution when n is large. As long as the variance of X_i is finite, the limit is the same. It makes no difference whether X_i is a discrete or a continuous random variable—you always get a bell curve. This is an example of what physicists call *universality*: The domain of attraction of the limiting distribution contains all examples obeying a certain regularity condition, which in this case is the finiteness of the variance.

The universality goes further. Consider the sequence of partial sums $S_n = \sum_{i=1}^n X_i$, with n regarded as a time variable. Then $\{0, S_1, S_2, \dots\}$ gives a random walk, with ith step X_i. Properly interpolated and scaled with n, the random walk converges to Brownian motion in the *scaling limit* $n \to \infty$. As long as the variance of X_i is finite, the distribution of X_i is again irrelevant. Brownian motion is an example of a process with the Markov property: What the process does next depends on its current state and not otherwise on its previous history. The state space of the Markov process is the set $\mathbb{R}$, or, more generally, the set $\mathbb{R}^d$ of possible positions for the motion.

The convergence to Brownian motion is an example of a limit theorem for random spatial processes. Such limit theorems derive much of their beauty and power from their universality, and they are among the most important results in probability theory. A large class of random spatial processes involves populations that undergo branching and death in addition to random spatial motion in $\mathbb{R}^d$. Often, the branching occurs on all time scales, including very small ones. Super-Brownian motion, a relatively recent and exotic Markov process, is the scaling limit of many such processes. In contrast to Brownian motion, where the state space is $\mathbb{R}^d$, *the state space of super-Brownian motion is a set of finite measures on* $\mathbb{R}^d$, and the state of the system represents the mass density of particles alive at time t. The state evolves randomly according to the random subsequent branching and particle motions.

Super-Brownian motion was introduced independently by Watanabe in 1968 and Dawson in 1975. The name "super-Brownian motion" was coined by Dynkin in the late 1980s. The subject of super-Brownian motion and of more general superprocesses, or measure-valued diffusions, has considerable depth and breadth, and there is now a large literature on the subject. This article is an introduction to some aspects of super-Brownian

Gordon Slade is professor of mathematics at the University of British Columbia, Canada. His email address is `slade@math.ubc.ca`.

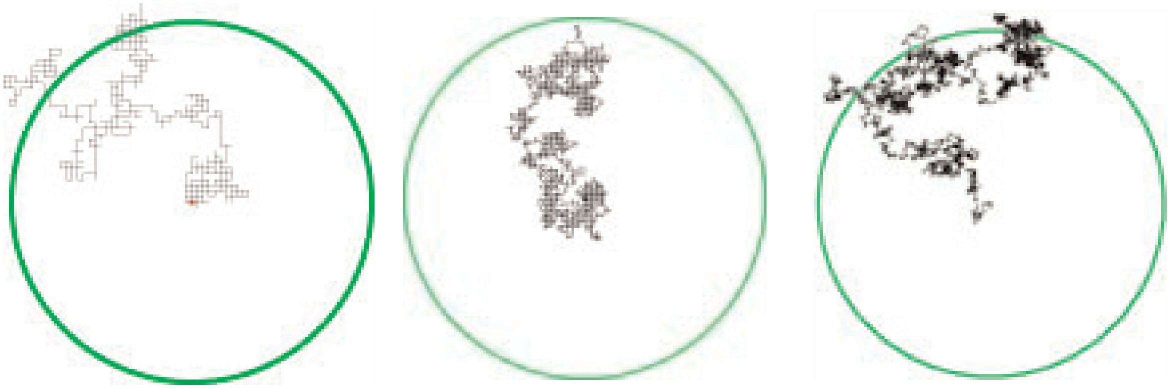

Figure 1. Nearest-neighbour random walks on $\mathbb{Z}^2$ taking n = 1,000, 10,000, and 100,000 steps. The circles have radius $\sqrt{n}$, in units of the step size of the random walk.

motion, centred on recent work showing that in high spatial dimensions its domain of attraction includes models of interest in combinatorics (lattice trees), statistical mechanics (critical percolation), and interacting particle systems (voter model and contact process). We begin with a review of the central limit theorem and Brownian motion.

Part I: Brownian Motion

The Central Limit Theorem and Random Walk

The central limit theorem implies a statement about the scaling limit of the endpoint of a random walk in d dimensions. Let $X_1, X_2, X_3, \ldots$ be any sequence of independent and identically distributed random vectors in $\mathbb{R}^d$. Then $S_n = X_1 + \cdots + X_n$ denotes the position of a random walk in $\mathbb{R}^d$ after n steps. For simplicity, we assume that X_i has mean zero and variance 1 and that the components of X_i are uncorrelated. It easily follows that the expected value of $|S_n|^2$ is n, for all $d \geq 1$. This suggests that the scaling $n^{-1/2}S_n$ is appropriate when analysing the limit $n \to \infty$. The central limit theorem confirms this scaling. It states, in the present context, that

$$\lim_{n\to\infty} \mathbb{P}(n^{-1/2}S_n \in A) = \int_A p_1(x)\, dx, \tag{2}$$

for any Lebesgue measurable set $A \subset \mathbb{R}^d$ whose boundary has measure zero. In other words, the probability that S_n lies in the set $n^{1/2}A$ is given, in the limit $n \to \infty$, by the integral over A of the standard Gaussian density p_1.

For concreteness, we focus in what follows on the *nearest-neighbour* walk in $\mathbb{Z}^d$, in which X_i is equally likely to be any one of the $2d$ nearest neighbours of the origin in $\mathbb{Z}^d$.

Brownian Motion

Equation (2) states that after n steps the endpoint of a random walk, scaled by $n^{-1/2}$, has a standard Gaussian distribution in the limit. It is natural to ask not just about the scaling limit of the endpoint, but the scaling limit of the entire walk. Consider the nearest-neighbour walk. As t varies over the interval $[0,1]$, $S_{\lfloor nt \rfloor}$ jumps from $S_0 = 0$ to S_1 to S_2 and so on until reaching S_n. We obtain a random element of the set $C[0,1]$ of continuous functions from $[0,1]$ to $\mathbb{R}^d$ by linearly interpolating between the S_i and then scaling space by $n^{-1/2}$. There are $(2d)^n$ such continuous functions, corresponding to the $(2d)^n$ nearest-neighbour walks. For each n, the random walk induces the discrete probability measure μ_n on $C[0,1]$ for which each of these $(2d)^n$ continuous functions has equal probability $(2d)^{-n}$. The question of the scaling limit of the random walk can then be rephrased as the question of whether the measures μ_n converge to a limiting measure μ as $n \to \infty$.

In 1951, Donsker proved that the answer is yes. The limiting Wiener measure μ on $C[0,1]$ is the mathematical formulation of Brownian motion. Under μ, the probability that a Brownian path at time t is in a Borel set $A \subset \mathbb{R}^d$ is equal to $\int_A p_t(x)\, dx$. The convergence of μ_n to μ is weak convergence, meaning that for any bounded continuous function f on $C[0,1]$,

$$\lim_{n\to\infty} \int_{C[0,1]} f\, d\mu_n = \int_{C[0,1]} f\, d\mu. \tag{3}$$

On the left side of (3), the integral is simply a finite sum, but this is not the case on the right side.

The Wiener measure μ is supported on irregular continuous paths that are nowhere differentiable and that have Hausdorff dimension 2 for $d \geq 2$. The irregularity of the paths is apparent from Figure 1, as is the fact that $n^{-1/2}$ is the correct spatial scaling.

Although Brownian motion is a Markov process, its domain of attraction includes an interesting example that is not Markovian.

The Self-Avoiding Walk

An n-step self-avoiding walk is a mapping $W : \{0, 1, \ldots, n\} \to \mathbb{Z}^d$ such that $W_0 = 0$, $|W_{i+1} - W_i| = 1$, and $W_i \neq W_j$ for all $i \neq j$. Let c_n be the number of n-step self-avoiding walks. For each n, we declare each of these c_n self-avoiding walks to be equally likely. This does not define a random walk that can be described in terms of transition probabilities. It

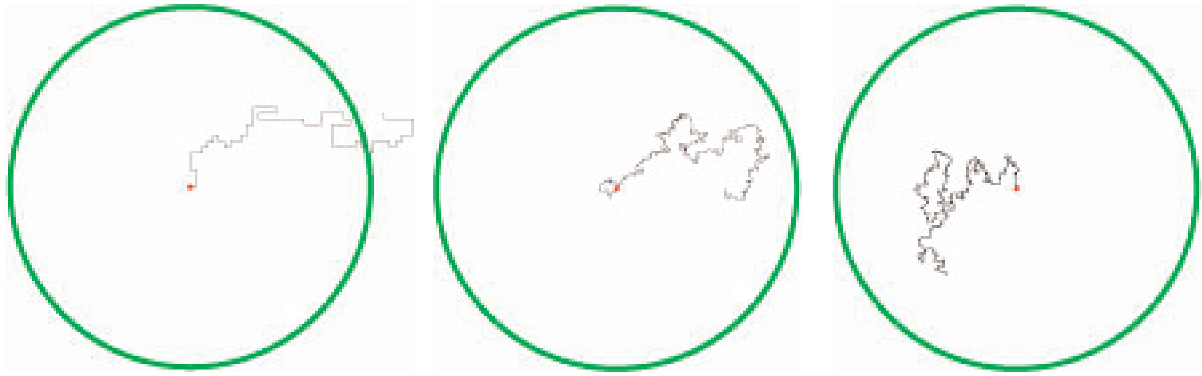

Figure 2. Nearest-neighbour self-avoiding walks on $\mathbb{Z}^2$ taking $n = 100, 1{,}000$, and $10{,}000$ steps. The circles have radius $n^{3/4}$, in units of the step size of the self-avoiding walk. The walks were generated using the pivot algorithm, an algorithm invented by Lal and developed by Madras and Sokal.

is certainly not Markovian, since the path must avoid its entire past history.

Random 2-dimensional self-avoiding walks are depicted in Figure 2. They do not look like Brownian motion. Exciting recent work by Lawler, Schramm, and Werner predicts, but has not yet proved, that with spatial scaling $n^{-3/4}$ the scaling limit for $d = 2$ is described by a process called the stochastic Loewner evolution with parameter $\frac{8}{3}$ ($SLE_{8/3}$). This 2-dimensional work relies heavily on a conjectured conformal invariance property of the scaling limit. For $d = 3$, no process has yet been proposed for the scaling limit, although $n^{-0.588...}$ has been predicted for the correct spatial scaling. For $d = 4$, the scaling limit with spatial scaling $n^{-1/2}(\log n)^{-1/8}$ is believed to be Brownian motion, but this has not yet been proved despite recent partial progress by Brydges and Imbrie using renormalisation group methods. For dimensions $d \geq 5$, Hara and Slade used a method called the *lace expansion* to prove that the scaling limit is Brownian motion in the following sense. Define a measure ρ_n on $C[0,1]$ by assigning mass c_n^{-1} to each of the c_n paths obtained by first interpolating the W_j and then rescaling by an appropriate constant multiple of $n^{-1/2}$. Then ρ_n converges weakly to the Wiener measure in the sense that (3) holds when μ_n is replaced on the left side by ρ_n.

The dimension $d = 4$ is referred to as the upper critical dimension, above which Gaussian behaviour is observed. A rough argument predicting $d = 4$ as the upper critical dimension is the following. Brownian paths are 2-dimensional, and two 2-dimensional sets generically do not intersect above $4 = 2 + 2$ dimensions. This suggests that there is enough room in dimensions above 4 that the self-avoidance constraint is unimportant, and hence, despite its highly non-Markovian nature, the self-avoiding walk should have the same Brownian scaling limit as random walk. The lace expansion exploits this intuition and treats the self-avoiding walk as a small perturbation of simple random walk, in dimensions $d \geq 5$.

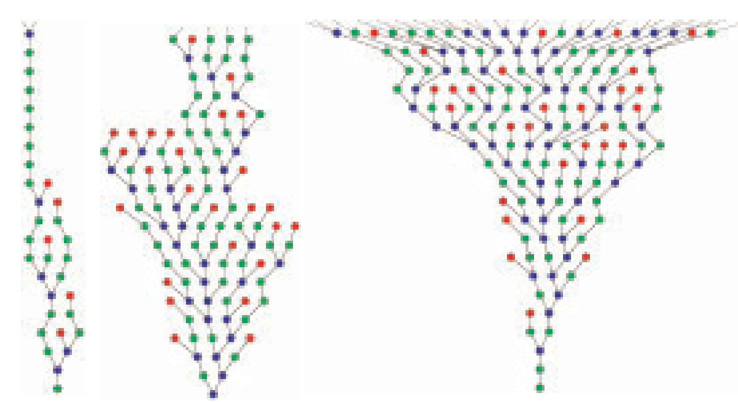

Figure 3. Random trees with binomial offspring distribution $\mathbb{P}(\xi = i) = \binom{2}{i} p^i (1-p)^{2-i}$ ($i = 0, 1, 2$) which survive for at least 20 generations, in the subcritical ($p = 0.47$), critical ($p = 0.50$), and supercritical ($p = 0.53$) regimes.

Part II: Super-Brownian Motion

Galton-Watson Trees

Super-Brownian motion combines random spatial motion with random branching. In preparation, we first consider pure branching processes, without any spatial motion.

A Galton–Watson branching process is defined as follows. Let ξ be a random variable taking values in $\{0, 1, 2, \dots\}$, with finite variance σ^2. A single individual at time 0 has a random number of offspring, with the distribution of ξ, and then dies. At time 1, each of these offspring has a random number of offspring, each with the distribution of ξ, and then dies. At time 2, these grandchildren of the original individual have their own offspring and die, and so on. The offspring random variables are all independent. At present, we do not associate any spatial location to individuals. The process continues forever or until the family dies out. So will it die out? It depends on the mean m of ξ. If $m < 1$, then on average a generation will give birth to a smaller subsequent generation and the family will almost surely die out. If $m > 1$, then a population explosion can occur and there is a positive probability of survival for all time. For $m = 1$, Kolmogorov proved in 1938 that the process will die out almost surely, with the probability of survival to time n going to zero as $n \to \infty$ asymptotically as $2\sigma^{-2}n^{-1}$. The case $m = 1$ is referred to as critical, with $m < 1$ subcritical and $m > 1$ supercritical. It is clear from Figure 3 that trees look qualitatively different in the three regimes.

Our focus is on the rare, long-lived critical family trees. Roughly speaking, in the critical case, the width of a tree tends to be proportional to its height n, so the total number of vertices is proportional to n^2. The large critical trees of Figure 4 suggest that a theory of continuously branching trees would be useful. Such a theory was developed by Aldous and by Le Gall in the early 1990s.

Galton-Watson Trees Embedded in Space

We introduce a spatial element to the trees as follows. Fix a large N, and consider the probability distribution of critical Galton–Watson trees conditioned to have total number of vertices equal to N. The height of a critical tree is roughly the square root of the total number of vertices, so these trees live for about $n = N^{1/2}$ generations. Given a random tree drawn from this conditional distribution, we embed it randomly into $N^{-1/4}\mathbb{Z}^d$ subject to the rules: (i) The individual at generation 0 is mapped to the origin of $N^{-1/4}\mathbb{Z}^d$; (ii) an individual is mapped at random to one of the $2d$ neighbours of its parent. See Figure 5. There are thus two sources of randomness in the embedded tree, as both the tree and the embedding of the tree are random. A tree with N vertices has $(2d)^{N-1}$ equally likely embeddings, since the root is embedded at the origin and each child can be located at any of the $2d$ neighbours of its parent.

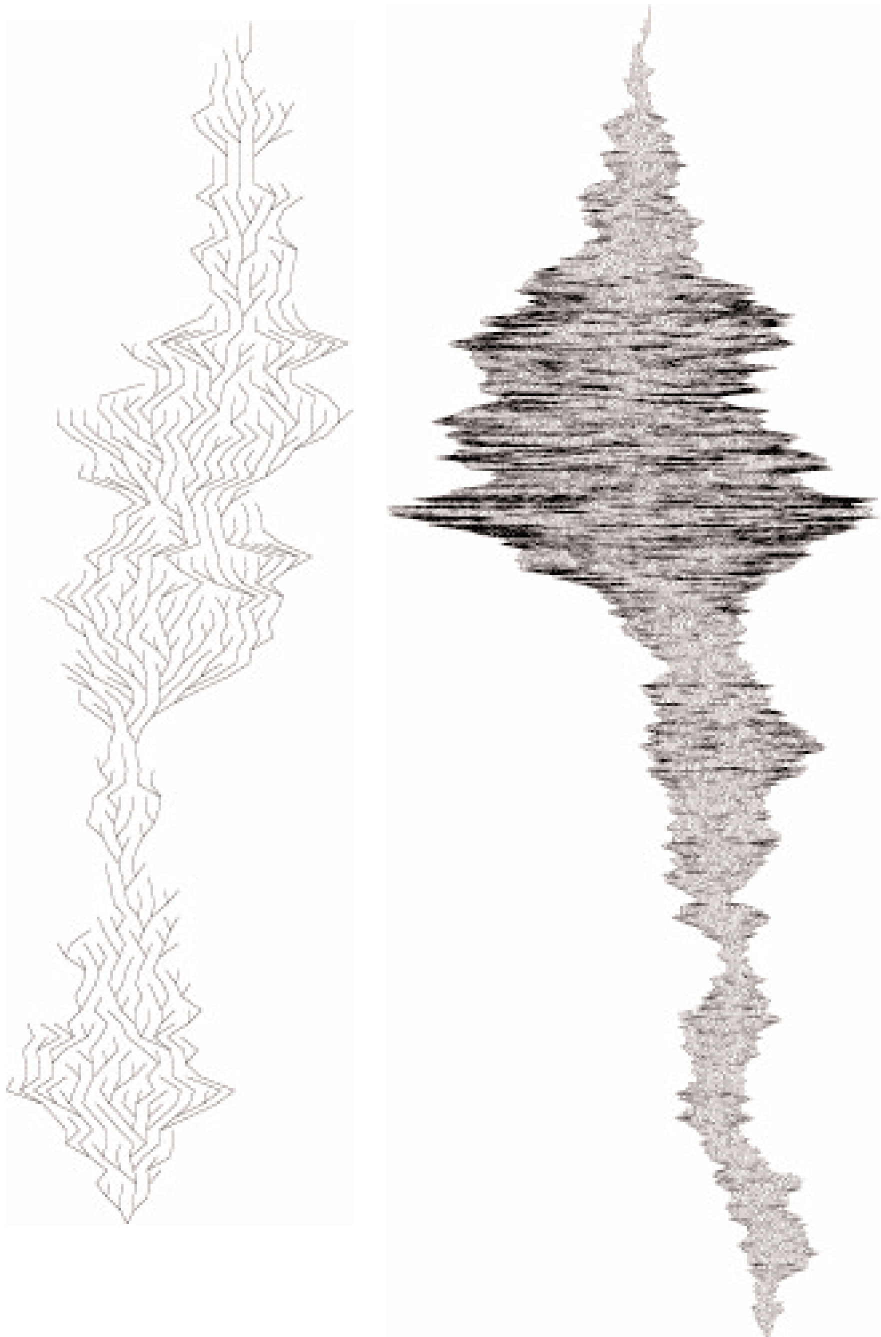

Figure 4. Rare critical trees that survive for (approximately) 100 and 500 generations. The offspring distribution is the same as for the middle tree of Figure 3.

Dynamically, we can think of this as a branching random walk, where at each time step an individual gives birth to a random number of offspring and dies, with the offspring located at random neighbours of their parent's position. An individual born into a generation of order $n = N^{1/2}$ has a line of ancestry that is a random walk path from the origin

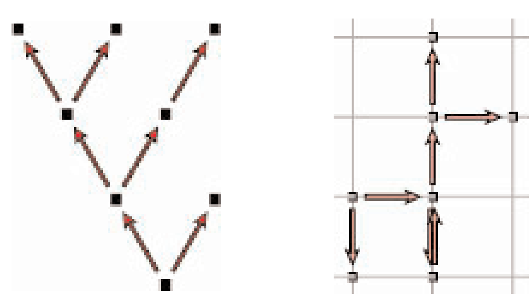

Figure 5. A tree together with an embedding into $\mathbb{Z}^2$.

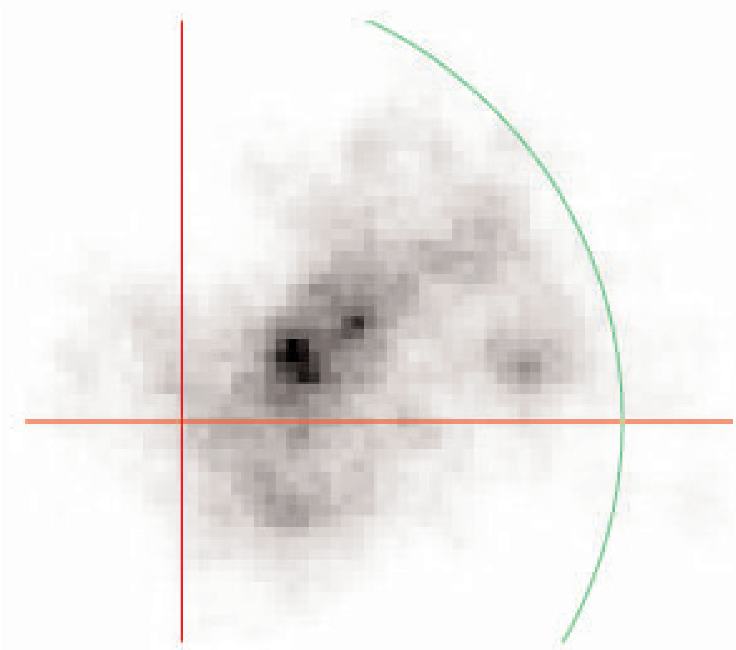

Figure 6. A random mass distribution drawn from $\mathcal{I}_N$, with $N \approx 96,000$. Darker shading represents higher mass, and the arc has radius $N^{1/4}$.

taking about n steps. According to Brownian scaling, the individual is thus located at a distance about $n^{1/2} = N^{1/4}$ from the origin, so $N^{-1/4}\mathbb{Z}^d$ is the natural scaling.

The Scaling Limit: Integrated Super-Brownian Excursion (ISE)

We want to take the limit $N \to \infty$ of the embedded trees and need to decide how much structure to attempt to track. A simple choice is to record only the location and multiplicity of the N embedded vertices and to forget, for now, all trace of ancestral lines.

Thus we assign mass N^{-1} to each embedded vertex, with multiplicity. The total mass is 1, so an embedded tree (T, φ), where T is the tree and φ is its embedding, produces a discrete probability measure τ on $\mathbb{R}^d$, with $\tau(x)$ equal to the embedded mass at $x \in \mathbb{R}^d$. The measure τ represents the mass distribution of the embedded tree. Since τ depends on the random tree T and the random embedding φ of T, τ is a *random measure*. The randomness can be described by a discrete probability measure $\mathcal{I}_N$ on the measures τ, with $\mathcal{I}_N(\tau)$ equal to the probability that the mass distribution τ occurs as the mass distribution of an embedded N-vertex tree. A random τ is depicted in Figure 6.

The measure $\mathcal{I}_N$ is a probability measure on probability measures on $\mathbb{R}^d$. This is fancier language than is really essential to understand finite N, but it is the right framework to describe the limit $N \to \infty$. The limit of $\mathcal{I}_N$ is a probability measure $\mathcal{I}$ on probability measures on $\mathbb{R}^d$, called integrated super-Brownian excursion (ISE). Its initial study and properties were developed by Aldous and by Le Gall. The measures on $\mathbb{R}^d$ represent mass distributions, and the measure $\mathcal{I}$ characterises their randomness. The convergence $\mathcal{I}_N \to \mathcal{I}$ is weak convergence, which means the following. Let $M_1(\mathbb{R}^d)$ denote the space of probability measures on $\mathbb{R}^d$, equipped with the topology of weak convergence of measures on $\mathbb{R}^d$. Then, given any bounded continuous real-valued function f on $M_1(\mathbb{R}^d)$,

$$\lim_{N\to\infty} \int_{M_1(\mathbb{R}^d)} f(\tau)\, d\mathcal{I}_N(\tau) = \int_{M_1(\mathbb{R}^d)} f(\tau)\, d\mathcal{I}(\tau). \tag{4}$$

In other words, the average of a function f of the discrete mass distributions converges to the average of f over random mass distributions that have the probability distribution of ISE. In this way, ISE describes the scaling limit of the embedded trees.

ISE represents a random mass distribution on $\mathbb{R}^d$. A natural question is: What is the average amount of mass located in an element dx? The answer is $f^{(1)}(x)dx$, where

$$f^{(1)}(x) = \int_0^\infty t e^{-t^2/2} p_t(x)\, dt, \tag{5}$$

and where we assume for simplicity that the offspring distribution ξ has variance $\sigma^2 = 1$.

The factor $p_t(x)$ has the interpretation that mass arrives at x at time t via a Brownian path. The time t is integrated over $[0, \infty)$ against the density $te^{-t^2/2}$, which accounts for the possible arrival times of mass at x. Similarly, the mean joint mass in $dx_1 dx_2$ is $f^{(2)}(x_1, x_2) dx_1 dx_2$, where

$$f^{(2)}(x_1, x_2) = \int_0^\infty \int_0^\infty \int_0^\infty \int_{\mathbb{R}^d} t e^{-t^2/2} p_{t_1}(y) \times p_{t_2}(x_1 - y) p_{t_3}(x_2 - y)\, dy\, dt_1 dt_2 dt_3 \tag{6}$$

with $t = t_1 + t_2 + t_3$. In (6), the intuition is that x_1 and x_2 have a common ancestor at time t_1. Equations (5) and (6) are the densities of the first and second *mean moment measures* of ISE. Similar formulas apply for the lth mean moment measures, for $l \geq 3$. Together the mean moment measures characterise ISE, and to prove weak convergence of a sequence of probability measures on $M_1(\mathbb{R}^d)$ to $\mathcal{I}$, it suffices to prove convergence of the mean moment measures.

The formulas (5) and (6) suggest that ISE has not completely forgotten that for $\mathcal{I}_N$ the underlying trees have a time variable indexed by generation number. The temporal structure is also apparent in Figure 7. Thus we might ask that the scaling limit retain the ancestral structure of the original trees, rather than just the overall mass distribution. We will come to this later. But first we discuss two examples where ISE arises.

Lattice Trees

A lattice tree on $\mathbb{Z}^d$ is a finite connected graph with no cycles, whose vertices are in $\mathbb{Z}^d$ and whose edges are chosen, for our purposes, from either

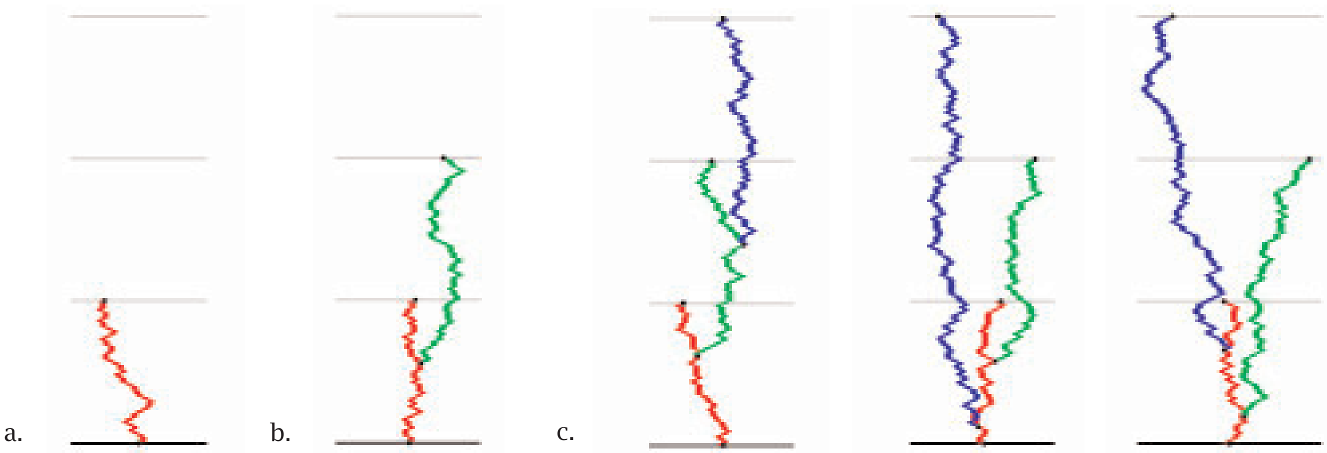

Figure 7. (a) Schematic representation of $f^{(1)}$ of (5), where the path corresponds to $p_t(x)$. (b) Schematic representation of $f^{(2)}$ of (6), where the three factors $p_t(x)$ of (6) correspond to the three paths. (c) For $f^{(3)}$, there are three terms, each with five factors $p_t(x)$.

$$\{\{x,y\} : x,y \in \mathbb{Z}^d,\ |x-y| = 1\} \quad \text{or} \quad \{\{x,y\} : x,y \in \mathbb{Z}^d,\ 0 < \|x-y\|_\infty \le L\}, \tag{7}$$

for some (large) $L \ge 1$. The first model is the *nearest-neighbour model* and the second is the *spread-out model*. We restrict attention to lattice trees that contain the origin, and we define a probability measure on the set of N-vertex lattice trees by declaring them to be equally likely. Lattice trees provide a good model of branched polymers, and they are also natural combinatorial objects. The hypothesis of universality suggests that scaling behaviour should be the same for both the nearest-neighbour and the spread-out models and be essentially independent of L for the spread-out models. Spread-out models are introduced because the small parameter L^{-1} is useful in applying the lace expansion, which is a perturbative method.

To construct a scaling limit of lattice trees, the correct spatial scaling must be used. A natural definition of length scale for N-vertex lattice trees is the root mean-squared radius of gyration[1] r_N. Numerical and other evidence supports the conjecture that r_N grows like a constant multiple of N^ν, for some dimension-dependent critical exponent ν. It is therefore natural to attempt to construct a scaling limit as follows. Given an N-vertex lattice tree in the rescaled lattice $cN^{-\nu}\mathbb{Z}^d$ for a suitable constant $c > 0$, define a discrete probability measure λ on $\mathbb{R}^d$ by placing mass N^{-1} at each vertex of the lattice tree. Let $\mathcal{J}_N$ be the probability measure on $M_1(\mathbb{R}^d)$ induced on the measures λ by the fact that each N-vertex lattice tree is equally likely. Thus λ represents the mass distribution of a random lattice tree in the rescaled lattice, with the randomness described by $\mathcal{J}_N$.

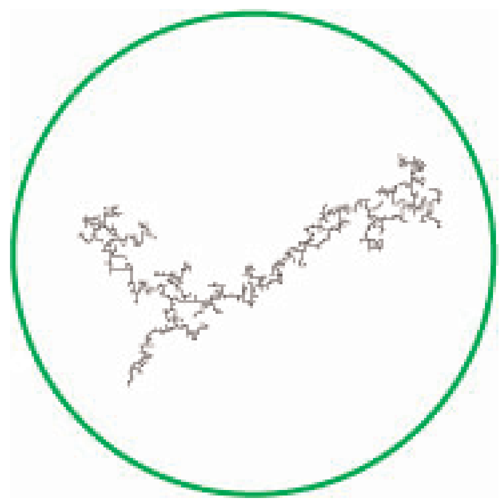

Figure 8. A random nearest-neighbour lattice tree on $\mathbb{Z}^2$ with $N = 1,000$ vertices. The circle has radius $N^{0.64}$. The tree was generated using an algorithm of Janse van Rensburg and Madras.

It follows from a result of Dawson, Iscoe, and Perkins that a random mass distribution under ISE is almost surely supported on a set of Hausdorff dimension 4, in dimensions above 4. An analogue of the rough argument discussed above for the self-avoiding walk led Aldous to conjecture that lattice trees should have ISE as scaling limit in dimensions $d > 8 = 4 + 4$. The conjecture captures the idea that in dimensions $d > 8$ there should be little qualitative difference between embedded critical Galton-Watson trees that can have many vertices per lattice site and lattice trees that can have only one. Consistent with the conjecture, in 1992 Hara and Slade showed that r_N scales like $N^{1/4}$ in dimensions $d > 8$ for spread-out models with L

[1] *Given a lattice tree T, its squared radius of gyration $r(T)^2$ is the mean-squared distance of vertices in T to the centre of mass of T, and r_N^2 is the average of $r(T)^2$ over N-vertex lattice trees containing the origin.*

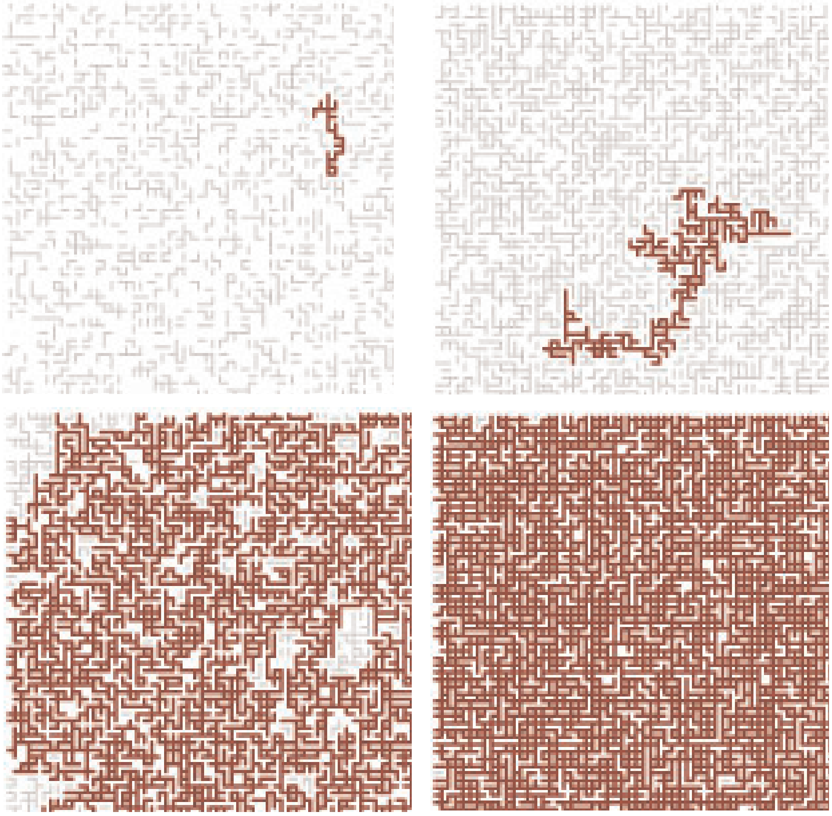

Figure 9. Nearest-neighbour bond percolation configurations on a 60×60 piece of the square lattice $\mathbb{Z}^2$ for $p = 0.25$, $p = 0.45$, $p = 0.55$, and $p = 0.75$. For the square lattice, $p_c = \frac{1}{2}$. The largest cluster in each figure is shown in red.

sufficiently large (depending on d) and for the nearest-neighbour model in dimensions much higher than 8. Under the same assumptions on the dimension, Derbez and Slade proved in 1998 that $\mathcal{J}_N \to \mathcal{I}$ weakly, in the sense of (4). The proof uses an adaptation of the lace expansion to lattice trees to prove convergence of all the moment measures of $\mathcal{J}_N$ to the moment measures of $\mathcal{I}$.

For $d < 8$, the scaling limit is predicted to be dimension-dependent and not to be ISE (contrast the 2-dimensional Figures 6 and 8). The limiting processes are not known.

Percolation

Percolation was introduced by Broadbent and Hammersley in 1957 as a model of fluid flow in a random medium. For our purposes, the setting for bond percolation is the infinite graph with vertex set $\mathbb{Z}^d$ and edge set given by either possibility of (7). Edges are now called *bonds*. To each bond b is associated a random variable n_b that takes the value 1 (b is *occupied*) with probability p and the value 0 (b is *vacant*) with probability $1 - p$. Here p is a parameter in the interval $[0, 1]$, and the random variables n_b are independent. Occupied bonds are interpreted as permitting fluid flow, whereas vacant bonds are blocked. The focus is on the geometry of the connected clusters of occupied bonds. What makes this model interesting is that it undergoes a phase transition. For $d \geq 2$, there is a critical value $p_c \in (0, 1)$ such that for $p < p_c$ there is no infinite cluster with probability 1, while for $p > p_c$ there exists a unique infinite cluster with probability 1 ("percolation occurs"). The phase transition is apparent in Figure 9. It is analogous to the transition in a Galton-Watson branching process as the mean of the offspring distribution is varied through $m = 1$. The conjecture that an infinite cluster occurs with probability zero when $p = p_c$ is widely and strongly believed. However, proofs[2] remain restricted to dimensions $d = 2$ and $d \geq 19$ for nearest-neighbour models and to $d > 6$ for spread-out models with $L \gg 1$. A general proof in arbitrary dimensions remains a central outstanding problem in the field.

Let us concentrate on the critical value $p = p_c$, and assume that percolation does not occur. This presents a delicate situation in which there is no infinite cluster, but the slightest increase in the bond density p would create an infinite cluster. The term "incipient infinite cluster" is used with various meanings to refer to the emerging structures that are poised to appear, when p is exactly critical. One approach to studying this is to investigate the scaling limit of $C(0)$, the random set of vertices in $\mathbb{Z}^d$ that are connected to 0 by paths consisting of occupied bonds; see Figure 10.

Given a set A containing N vertices that is a possible candidate for $C(0)$, we define a discrete probability measure $\pi \in M_1(\mathbb{R}^d)$ by assigning mass N^{-1} to each vertex in $cN^{-1/4}A$, for a suitable constant c. The spatial scaling $N^{-1/4}$ is ISE scaling and is expected to be the correct scaling for $C(0)$ only in dimensions $d > 6$. We then define a probability measure $\mathcal{P}_N$ on $M_1(\mathbb{R}^d)$ by defining $\mathcal{P}_N(\pi)$ to equal the conditional probability that $C(0) = A$ given $|C(0)| = N$. In other words, $\mathcal{P}_N$ gives the distribution of rescaled mass profiles that arise from critical percolation clusters that are conditioned to contain exactly N vertices. Hara and Slade conjectured that $\mathcal{P}_N \to \mathcal{I}$ weakly for $d > 6$ and partially proved the conjecture by showing that for the nearest-neighbour model in dimensions much higher than 6 the first and second moment measures of $\mathcal{P}_N$ converge to those of $\mathcal{I}$ (see (5) and (6)). A weaker statement links $\mathcal{P}_N$ to $\mathcal{I}$ for spread-out models with $d > 6$ and $L \gg 1$. The proof uses the lace expansion. The critical dimension 6 arises in the proof as the dimension above which an ISE cluster and a Brownian path typically do not intersect, which is

[2] *The final step for $d = 2$ was obtained by Kesten in 1980, following earlier work of Harris, Russo, and Seymour and Welsh. The high-dimensional results are due to a combination of theorems of Barsky and Aizenman and of Hara and Slade.*

above dimension $6 = 4 + 2$. This is more subtle than the formula $8 = 4 + 4$ that arises for lattice trees.

It is believed that different, unknown scaling limits apply when $3 \le d < 6$. Celebrated recent results of Smirnov and of Lawler, Schramm, and Werner prove that for critical site percolation on the 2-dimensional triangular lattice the scaling limit has a conformal invariance property and is described by the stochastic Loewner evolution with parameter 6 (SLE_6). Universality suggests that the same should be true for critical bond percolation on $\mathbb{Z}^2$.

Canonical Measure of Super-Brownian Motion

We return now to the question of the scaling limit of embedded Galton–Watson trees. Previously, with ISE, we considered trees consisting of N vertices and kept track of the mass distribution of the embedded vertices. In so doing, we worked with probability measures on $\mathbb{R}^d$ that fixed the overall mass to be 1. Also, we lost information about the time evolution inherent in the underlying trees. Now, we want to revisit the scaling limit in such a way that the limiting object experiences a time evolution and has a random mass.

Let T be a critical Galton–Watson tree. We make no assumption about the total number of vertices in T. Typically, T will have few vertices, but we are most interested in the rare, long-lived trees like the examples in Figure 4. Let T_0 denote the initial individual in the 0th generation, let T_1 denote the offspring of the initial individual, and more generally, let T_m denote the individuals of generation m. Since T is finite, eventually $T_m = \varnothing$. We regard m as a discrete time variable.

Fix a large integer n, the scaling variable. We will select embedded trees that survive for order n generations, in the limit $n \to \infty$. This involves four simultaneous scalings, as follows.

- *Scaling of time.* We rescale time to a new time variable $t = \frac{m}{n}$. This fixes attention on the individuals T_m in generations m of order n.
- *Scaling of space.* If we embed a tree into $\mathbb{Z}^d$ according to the previous rules that T_0 is embedded at the origin, and an individual is embedded at a random neighbour of its parent, then an individual in generation n (if T lives so long) is embedded at the endpoint of an n-step random walk path. Such a point is typically at distance order $n^{1/2}$ from the origin, so we rescale the lattice to $n^{-1/2}\mathbb{Z}^d$.
- *Scaling of mass.* A critical tree that survives to a generation of order n typically has order n members in that generation. To obtain a generation mass of order 1, we assign mass n^{-1} to each embedded vertex.
- *Scaling of probability.* By Kolmogorov's theorem, the probability that a critical tree survives for order n generations is of order n^{-1}. The large trees of interest thus have a vanishingly small probability. To compensate, we multiply probabilities by n, so that survival to order n generations will have measure of order 1, rather than probability of order n^{-1}. This produces an unnormalised measure, rather than a probability measure.

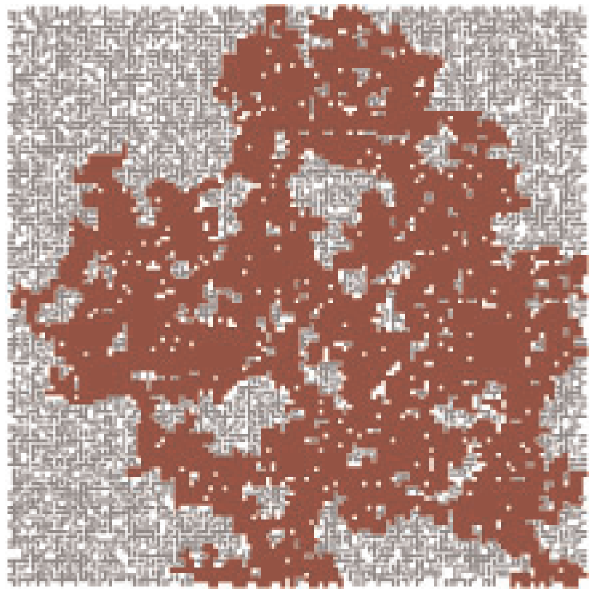

Figure 10. The connected cluster $C(0)$ of the origin consists of vertices that are joined to the origin by a path consisting of occupied bonds. Here $p = p_c = \frac{1}{2}$ for $\mathbb{Z}^2$, and the vertices in $C(0)$ are depicted by red pixels. It is a rare event that $C(0)$ is so large at the critical point.

The above is carried out, in detail, as follows. Given a critical Galton–Watson tree T and an embedding of T into $n^{-1/2}\mathbb{Z}^d$, we denote by $R_n^{(m/n)}$ the mass distribution in $n^{-1/2}\mathbb{Z}^d$ of the mth generation T_m of T. Explicitly, $R_n^{(m/n)}$ is the discrete finite measure on $\mathbb{R}^d$ that places mass n^{-1} at each embedded vertex of T_m, with multiplicity. Note that $R_n^{(m/n)}$ need not be a probability measure. The measure $R_n^{(m/n)}$ is a *random measure*, since T and its embedding are random. Let $\mathcal{R}_n^{(m/n)}$ denote the probability law of this random measure. Thus $\mathcal{R}_n^{(m/n)}$ is a probability measure on the space $M(\mathbb{R}^d)$ of finite measures on $\mathbb{R}^d$, which quantifies how likely it is that a particular mass distribution occurs as the embedding in $n^{-1/2}\mathbb{Z}^d$ of the mth generation of a critical Galton–Watson tree. So far, we have rescaled generation m to time m/n, space to $n^{-1/2}\mathbb{Z}^d$, and vertex mass has been set equal to n^{-1}.

It remains to scale the probability measure $\mathcal{R}_n^{(m/n)}$. Think of m of order n. The random measure $R_n^{(m/n)}$ is the zero measure on $\mathbb{R}^d$ with probability $1 - O(n^{-1})$, since the probability of survival for m generations is order $m^{-1} \approx n^{-1}$. To amplify the rare event of survival to m generations, we consider the measure $n\mathcal{R}_n^{(m/n)}$ on $M(\mathbb{R}^d)$, which has total measure n. This measure places measure $n - O(1)$ on the zero measure on $\mathbb{R}^d$ and the

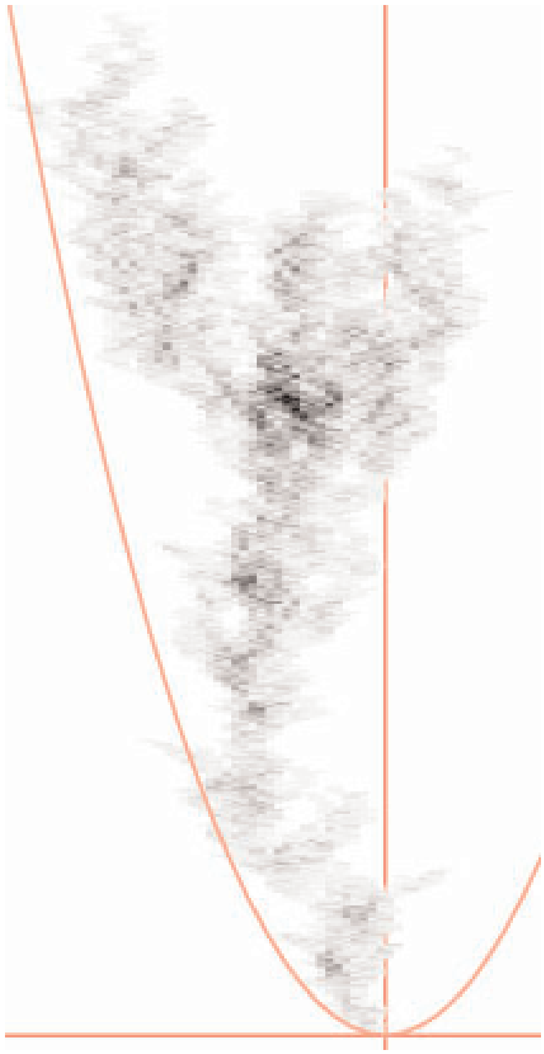

Figure 11. The evolution of a random mass distribution under the canonical measure, with time vertical and 1-dimensional space horizontal. Darker shading represents higher mass, and the parabola $t = x^2$ shows the spatial scaling. The mass distribution arises from an embedding of a large critical tree like those in Figure 4. A superposition of all the mass onto the spatial axis produces a random mass distribution, whose total mass K is a random variable. The superposition, with mass rescaled by K^{-1} and space rescaled by $K^{-1/4}$, has the same distribution as ISE, depicted in Figure 6.

remaining measure $O(1)$ on the nontrivial mass distributions due to the rare trees that survive for m generations. It can be shown that the limit of $n\mathcal{R}_n^{(\lfloor tn\rfloor/n)}$, as $n \to \infty$, consists of an infinite point mass on the zero measure on $\mathbb{R}^d$, plus a nontrivial finite measure $\mathcal{R}^{(t)}$ on $M_0(\mathbb{R}^d)$, the finite measures on $\mathbb{R}^d$ excluding the zero measure. More precisely, there is a measure $\mathcal{R}^{(t)}$ on $M_0(\mathbb{R}^d)$ such that $n\mathcal{R}_n^{(\lfloor tn\rfloor/n)}$ converges weakly to $\mathcal{R}^{(t)}$ plus an infinite point mass on the zero measure on $\mathbb{R}^d$. This means that for every bounded continuous function f on $M_0(\mathbb{R}^d)$,

$$\lim_{n\to\infty} n \int_{M_0(\mathbb{R}^d)} f(R)\,d\mathcal{R}_n^{(\lfloor tn\rfloor/n)}(R) = \int_{M_0(\mathbb{R}^d)} f(R)\,d\mathcal{R}^{(t)}(R). \tag{8}$$

In other words, the average of a function over the interesting, long-lived, random configurations of the discrete mass distribution defined by $\mathcal{R}_n^{(\lfloor tn\rfloor/n)}$ converges to a corresponding average over the random configurations of the continuous mass distribution defined by $\mathcal{R}^{(t)}$. The family of measures $\mathcal{R}^{(t)}$, for $t > 0$, is intimately related to the canonical measure of super-Brownian motion.

The *canonical measure* $\mathcal{N}$ describes the evolution in continuous time of a mass distribution on $\mathbb{R}^d$ arising from a single initial individual at the origin whose progeny survives for some positive time. Thus $\mathcal{N}$ is a measure on functions $Y^{(\cdot)}$ from the time interval $[0,\infty)$ into finite measures on $\mathbb{R}^d$. A representative measure-valued path $Y^{(\cdot)}$ is illustrated in Figure 11. For a specific time $t > 0$, and restricted to configurations that have not died out by time t, the distribution of $Y^{(t)}$ under the canonical measure is given by $\mathcal{R}^{(t)}$.

The moment measures of ISE have natural analogues for the canonical measure of super-Brownian motion. The first mean moment measure describes the average amount of mass in an element dx at time t and is equal to $g_t^{(1)}(x)dx$, where

$$g_t^{(1)}(x) = p_t(x). \tag{9}$$

This is consistent with mass arriving at x at time t via a Brownian path started from the origin of $\mathbb{R}^d$. The second mean moment measure describes the average amount of mass in dx_1 at time t_1 and in dx_2 at time t_2 and is equal to $g_{t_1,t_2}^{(2)}(x_1,x_2)dx_1dx_2$, where

$$g_{t_1,t_2}^{(2)}(x_1,x_2) = \int_0^{\min\{t_1,t_2\}} \int_{\mathbb{R}^d} p_s(y) \times p_{t_1-s}(x_1-y)p_{t_2-s}(x_2-y)\,dy\,ds. \tag{10}$$

Here, we have again taken the variance of the offspring distribution ξ to be $\sigma^2 = 1$. Similar explicit formulas can also be given for the higher mean moment measures. Figure 7 provides a schematic representation for the $g^{(l)}$ as well as the $f^{(l)}$. Convergence of the moment measures is one possible statement of convergence to the canonical measure, and we will see it used below.

Super-Brownian Motion

The canonical measure describes a situation in which all particles are descendants of a single initial particle subject to critical branching. In the discrete setting, instead of starting with a single particle and amplifying the rare event of survival to time n by multiplying probabilities by n, we may instead start with order n initial particles located at distances within order 1 from the origin in $n^{-1/2}\mathbb{Z}^d$ (so within order $n^{1/2}$ lattice spacings from the origin). Each initial particle has probability asymptotically proportional to n^{-1} to have progeny that survive to a time of order n and thus remain visible in the scaling limit. In rescaled continuous time, order ϵ^{-1} of the initial particles will have progeny that survive to time ϵ for small ϵ, but only order 1 particles will have progeny that survive to time 1.

After taking the scaling limit, the family tree of each survivor corresponds to an evolution under the canonical measure. The initial state is now given by a finite measure on $\mathbb{R}^d$, rather than a single initial particle. The resulting evolving random measure $X^{(t)}$ is a Markov process with state space $M(\mathbb{R}^d)$, and it is essentially built from independent copies of the canonical measure. It is the Markov process $X^{(\cdot)}$ that is usually referred to as super-Brownian motion. The evolution of $X^{(t)}$ is governed by a nonlinear partial differential equation.[3]

Next, we discuss three more examples that lie in the domain of attraction of super-Brownian motion.

Oriented Percolation

Percolation models can be defined on any infinite graph. A much-studied example is oriented, or directed, percolation. Oriented percolation takes place on the graph with vertex set $\mathbb{Z}^d \times \mathbb{N}$. The bonds are directed bonds of the form $((x,n),(y,n+1))$, with $\{x,y\}$ chosen from one of the two options in (7) and with $n \geq 0$ corresponding to a discrete time variable. Bonds are again occupied with probability p, but now a vertex (x,n) is in the connected cluster $C(0,0)$ of the origin if and only if it can be reached from the origin by a *directed* path consisting of occupied bonds. See Figure 12. For $d \geq 1$, this model undergoes a phase transition as in the nonoriented case, at a critical value $p_c \in (0,1)$. For oriented percolation, it was proved by Bezuidenhout and Grimmett that $C(0,0)$ is almost surely finite when $p = p_c$, for any $d \geq 1$.

The cluster $C(0,0)$ is somewhat similar to the embedding of a Galton–Watson tree. However, an important difference is that for oriented percolation there is no multiple occupation of a vertex by $C(0,0)$ at a given time, whereas a tree can have several distinct vertices simultaneously embedded at

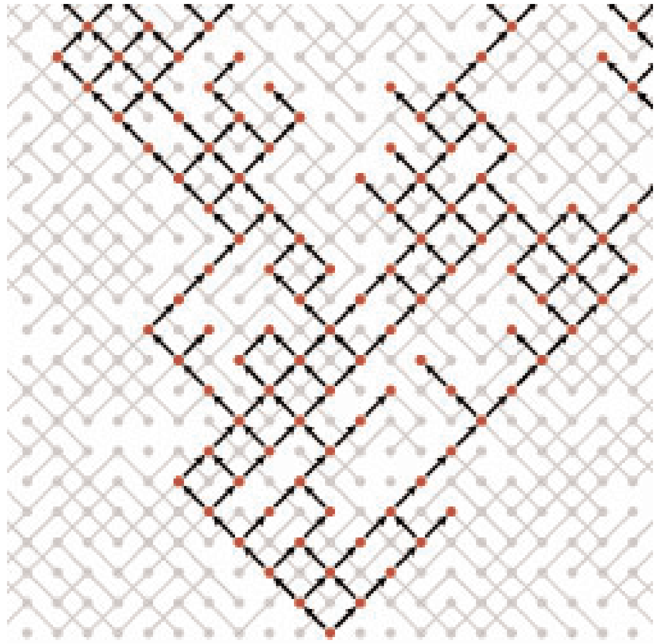

Figure 12. The connected cluster $C(0,0)$ of the origin for oriented percolation. Here $p = 0.7$ is a little above $p_c \approx 0.645$. For the nearest-neighbour model depicted, the lattice decomposes into two noncommunicating lattices. This will not be the case for the spread-out model when $L \geq 2$.

the same location. This difference would appear to diminish as the dimension increases, and one might expect that critical oriented percolation should behave like super-Brownian motion in sufficiently high dimensions. This is the case.

Consider the random cluster $C(0,0)$ of the origin, when $p = p_c$, and apply the four scalings of time, space, mass, and probability that were used in the construction of the canonical measure for super-Brownian motion. For spread-out models in dimensions $d > 4$ with $L \gg 1$, van der Hofstad and Slade used the lace expansion to prove that all the rescaled mean moment measures of critical oriented percolation converge to those of the canonical measure of super-Brownian motion (see (9) and (10)). In the proof, the critical dimension 4 appears as the dimension above which the graphs of independent super-Brownian motion and Brownian motion do not intersect. Different, unknown scaling limits are expected when $d < 4$.

Contact Process

The contact process is a model of an infection that spreads via contact with an infected neighbour. It has been studied for close to thirty years and is a basic example in the field of interacting particle systems. A particle is located at each vertex of $\mathbb{Z}^d$, and the process evolves over the continuous time interval $[0,\infty)$. Each particle is either healthy (state 0) or infected (state 1), so the state of the system at time t is defined by a mapping $\xi_t : \mathbb{Z}^d \to \{0,1\}$. Each

[3] *In this respect, it is noteworthy that the classical representation of the solution of the boundary-value problem for the Laplace equation $\Delta u = 0$ in terms of Brownian motion has an extension, due to Dynkin, to a representation of the solution of the boundary-value problem for the nonlinear equation $\Delta u = u^2$ in terms of super-Brownian motion.*

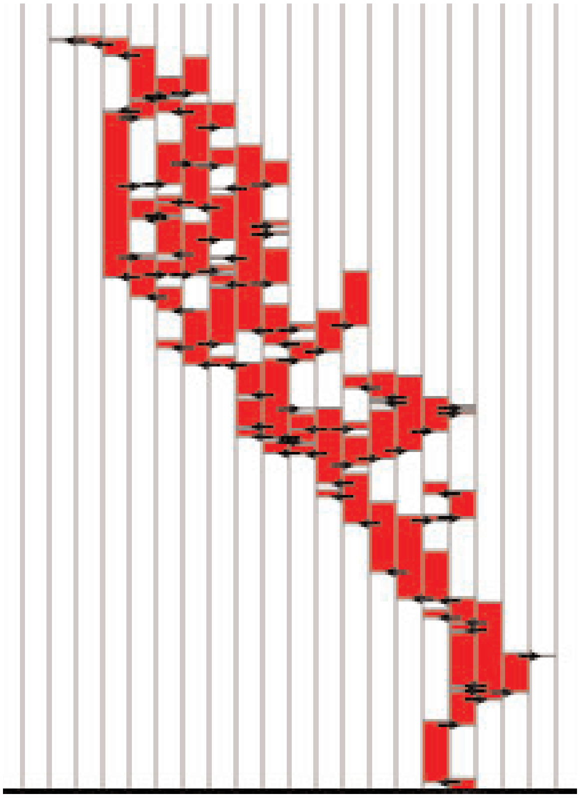

Figure 13. The contact process on $\mathbb{Z}^1$ with nearest-neighbour infections. Here $\beta = 1.6 < \beta_c \approx 1.65$. The legacy of infection of a single infected individual at time zero is shown in red. Arrows show successful attempts at infection.

vertex $x \in \mathbb{Z}^d$ has two alarm clocks:[4] a recovery clock that randomly rings at rate 1 and an infection clock that randomly rings at rate β. All clocks ring independently, and a clock resets itself immediately after ringing. If the particle at x is infected when its recovery clock rings, it becomes healthy spontaneously at that time. If the particle at x is infected when its infection clock rings, it chooses a random neighbour at that moment and infects the neighbour if the neighbour is healthy and does nothing otherwise. Neighbours are defined via either alternative of (7). Nothing happens if one of a healthy particle's clocks rings. The process is depicted in Figure 13.

The contact process undergoes a phase transition in the sense that there is a critical value β_c such that an infection started from a single infected individual will die out in finite time when $\beta < \beta_c$, whereas for $\beta > \beta_c$ there is a positive probability of survival for all time. Bezuidenhout and Grimmett proved that, as in oriented percolation, the infection also dies out when $\beta = \beta_c$. Our focus is on scaling limits for the critical case $\beta = \beta_c$.

[4] *The statement that a clock randomly rings at rate β means that the time between rings is an exponential random variable with mean β^{-1}, i.e., with probability density function $\beta e^{-\beta x}$ $(x \geq 0)$. In other words, the clock rings according to a Poisson process.*

In 1999, Durrett and Perkins studied the scaling limit of the critical spread-out contact process, but with an additional new scaling. The infection range $L = L_n$ of (7) is no longer independent of n but increases as $L_n = n^{1/d}$ for $d \geq 3$, and $L_n = (n \log n)^{1/2}$ for $d = 2$. Time and mass are scaled as usual by n and n^{-1}, and now the spatial lattice is rescaled to $n^{-1/2} L_n^{-1} \mathbb{Z}^d$. Given an initial condition ξ_0 whose rescaled mass distribution converges to a finite measure $X^{(0)}$ as $n \to \infty$, Durrett and Perkins proved a full statement of convergence to super-Brownian motion, going beyond convergence of moment measures, for all dimensions $d \geq 2$. Thus the legacy of infection of the initially infected individuals is described by the evolution of super-Brownian motion, under the above scaling.

Scaling the infection range can affect the limiting distribution, and it is believed that the finite range (fixed L) critical contact process converges to super-Brownian motion above dimension 4 but not below. Oriented percolation can be regarded as a discrete time version of the contact process, and it is likely that the results described above for oriented percolation could be extended to the finite range contact process for $d > 4$, using an observation of Sakai. For finite-range infection in dimensions $d < 4$, it is not known what the scaling limit should be.

Voter Model

The voter model is a model of the spread of opinion, where an individual's opinion is affected by its neighbours' opinions. Like the contact process, the voter model has been studied for about thirty years and is a basic example of an interacting particle system. An individual is located at each vertex $x \in \mathbb{Z}^d$ and holds either opinion 0 or opinion 1. Each individual has an alarm clock that randomly rings at rate 1, and all clocks are independent. When the clock of the individual at x rings, the individual spontaneously adopts the opinion of a randomly selected neighbour, leaving the neighbour's opinion unchanged. No change occurs if the opinions of x and y already agree when the clock rings. The state of the system at time t is thus described by a mapping $\eta_t : \mathbb{Z}^d \to \{0, 1\}$. There is no parameter in the model.

Cox, Durrett, and Perkins studied the scaling limit of the nearest-neighbour voter model, with time speeded up by a factor n and the lattice spacing rescaled to $n^{-1/2}$. In dimensions $d \geq 3$, mass n^{-1} is placed at each vertex x holding opinion 1, whereas for $d = 2$ mass $n^{-1} \log n$ is used instead. Given an initial condition η_0 whose rescaled mass distribution converges to a finite measure $X^{(0)}$ as $n \to \infty$, Cox, Durrett, and Perkins proved that the scaling limit is super-Brownian motion started from $X^{(0)}$. They also proved related results treating long-range voter models. The dimension $d = 2$

arises in this work as the borderline between recurrence and transience of random walk.

Bramson, Cox, and Le Gall have applied the above theorem to relate the voter model to the canonical measure of super-Brownian motion. They consider an initial condition in which the origin has opinion 1 and all other vertices have opinion 0 and condition on the unlikely event that opinion 1 survives until time $t = n$. For $d \geq 2$, under the above scaling, they prove that the probability law of the mass distribution of the individuals holding opinion 1 at time n converges to the canonical measure of super-Brownian motion at time 1, conditioned to survive until time 1.

Concluding Remarks

Super-Brownian motion is a rich and beautiful mathematical structure that models the randomly evolving mass distribution of populations that undergo critical branching and spatial diffusion. The five examples discussed above show that it is a universal object arising as a scaling limit in models from combinatorics, statistical mechanics, and interacting particle systems. Much more can be said, for example, concerning its connections with nonlinear partial differential equations and its representation in terms of Le Gall's Brownian snake. These and other topics can be found in the books and major reviews in the references.

Acknowledgments

Much of my understanding of super-Brownian motion is a result of collaborations with Eric Derbez, Takashi Hara, and Remco van der Hofstad. Timely pointers from David Aldous and Jean-François Le Gall and numerous clarifications from Ed Perkins are also gratefully acknowledged. This work was supported in part by NSERC of Canada.

References for ISE and Super-Brownian Motion

[1] D. Aldous, Tree-based models for random distribution of mass, *J. Stat. Phys.* **73** (1993), 625–41.

[2] C. Borgs, J. T. Chayes, R. van der Hofstad, and G. Slade, Mean-field lattice trees, *Ann. Combinatorics* **3** (1999), 205–21.

[3] D. A. Dawson, Measure-valued Markov processes. In *Ecole d'Eté de Probabilités de Saint-Flour* 1991, Lecture Notes in Mathematics, no. 1541, Springer, Berlin, 1993.

[4] E. B. Dynkin, *An Introduction to Branching Measure-Valued Processes*, American Mathematical Society, Providence, 1994.

[5] A. M. Etheridge, *An Introduction to Superprocesses*, American Mathematical Society, Providence, 2000.

[6] J.-F. Le Gall, *Spatial Branching Processes, Random Snakes, and Partial Differential Equations*, Birkhäuser, Basel, 1999.

[7] ____ , Processus de branchement, arbres et superprocessus. In J.-P. Pier, editor, *Development of Mathematics 1950–2000*, Birkhäuser, Basel, 2000, pp. 763–93.

[8] E. Perkins, Dawson-Watanabe superprocesses and measure-valued diffusions. In P. L. Bernard, editor, *Lectures on Probability Theory and Statistics. Ecole d'Eté de Probabilités de Saint-Flour XXIX-1999*, Springer (to appear).

References for the Examples

[1] M. Bramson, J. T. Cox, and J-F. Le Gall, Super-Brownian limits of voter model clusters. *Ann. Probab.* **29** (2001), 1001–32.

[2] J. T. Cox, R. Durrett, and E. A. Perkins, Rescaled voter models converge to super-Brownian motion, *Ann. Probab.* **28** (2000), 185–234.

[3] ____ , Rescaled particle systems converging to super-Brownian motion. In M. Bramson and R. Durrett, editors, *Perplexing Problems in Probability: Festschrift in Honor of Harry Kesten*, Birkhäuser, Basel, 1999.

[4] E. Derbez and G. Slade, The scaling limit of lattice trees in high dimensions, *Commun. Math. Phys.* **193** (1998), 69–104.

[5] R. Durrett and E. A. Perkins, Rescaled contact processes converge to super-Brownian motion in two or more dimensions, *Probab. Th. Rel. Fields* **114** (1999), 309–99.

[6] G. Grimmett, *Percolation*, second edition, Springer, Berlin, 1999.

[7] T. Hara and G. Slade, The scaling limit of the incipient infinite cluster in high-dimensional percolation, II. Integrated super-Brownian excursion, *J. Math. Phys.* **41** (2000), 1244–93.

[8] R. van der Hofstad and G. Slade, Convergence of critical oriented percolation to super-Brownian motion above $4 + 1$ dimensions, preprint, 2001.

[9] T. M. Liggett, *Interacting Particle Systems*, Springer, New York, 1985.

[10] N. Madras and G. Slade, *The Self-Avoiding Walk*, Birkhäuser, Boston, 1993.

[11] A. Sakai, Mean-field critical behavior for the contact process, *J. Stat. Phys.* **104** (2001), 111–43.

[12] G. Slade, Lattice trees, percolation and super-Brownian motion. In M. Bramson and R. Durrett, editors, *Perplexing Problems in Probability: Festschrift in Honor of Harry Kesten*, Birkhäuser, Basel, 1999.