Super-Brownian Motion and Critical Spatial Stochastic Systems

Ed Perkins, University of British Columbia

0. Introduction

This lecture is a terse introduction to a stochastic process called super-Brownian motion. This process has a number of fascinating properties and a number of connections to other topics in probability and analysis. Our focus will be on its emerging role as a universal limit for random spatially distributed systems near criticality. It is a topic particularly well-suited to a meeting of the Canadian Mathematical Society as Canadian mathematicians have played a leading role in many of these developments. The analogous universal limit for the time evolution of a single random point in space is Brownian motion.

1. Brownian Motion

Consider a particle moving in $\mathbb{Z}^d$ starting at the origin and taking independent random steps (W_i) to each of its $2d$ neighbours with equal probability. After n steps its random position is $S_n = \sum_{i=1}^n W_i$. Now look at this random walk from afar. In order to see a non-trivial picture we need to rescale time by $1/N$ and space by $1/\sqrt{N}$. This leads to the rescaled random walk: $B^{(N)}(t) = S_{[Nt]}/\sqrt{N}$, $t \geq 0$. The classical Central Limit Theorem states that the probability that this rescaled random walk at time 1 is in an interval, I, converges to the integral of a suitably scaled Gaussian density over I:

Theorem 1.1 (CLT). For each interval I,

$$\lim_{N\to\infty} P(B^{(N)}(1) \in I) = \int_I e^{-d|x|^2/2}(2\pi/d)^{-d/2}dx.$$

M. Donsker showed that if one views the entire time evolution of $B^{(N)}$ as a random function then the probability distribution of this random function converges to the probability distribution of a random function which will serve as our definition of Brownian motion. This result is known as the Functional Central Limit Theorem.

Theorem 1.2 (FCLT, Donsker (1951)). $\lim_{N\to\infty} P(B^{(N)} \in A) = P(d^{-1/2}B \in A)$ "for any" set A of paths from $\mathbb{R}_+$ to $\mathbb{R}^d$ where B is standard Brownian motion.

The precise nature of this convergence is weak convergence of probability measures on the space $D([0,\infty),\mathbb{R}^d)$ of right continuous R^d-valued paths with left limits equipped with the Skorokhod J_1 topology. This puts some mild restrictions on the sets of paths A considered above. We will ignore such routine technicalities here and in what follows.

The **universality of B** refers to the fact that the above theorem admits a huge number of extensions:

• If $\{W_i\} \in \mathbb{R}^d$ are any repeated independent random quantities with mean 0 and covariance $\sigma^2 I_{d\times d}$ one gets σB in the limit. Hence the particular local dynamics only affects the limit through the parameter, σ, the standard deviation of each coordinate.
• The entire field of "Invariance Principles" is based on extensions of the above theorem to a large variety of appropriate dependent sequences $\{W_i\}$ such as martingale difference sequences, and stationary sequences under a variety of hypotheses. These results all have the flavour that under fairly minimal conditions one gets Brownian motion as the rescaled limit.
• Instead of $t = i/N$ may take steps "at rate N", i.e. after independent random time intervals $T_i - T_{i-1}$ satisfying $P(T_i - T_{i-1} > t) = e^{-Nt}$ (mean $1/N$ exponential inter-step times). This is a trivial extension which will be useful to us later.

The universality of B leads to its use as a stochastic model for such phenomena as a particle suspended in liquid undergoing molecular bombardment of the medium as observed by Brown (1828) or fluctuations of the stock market as was proposed by Bachelier (1900). Perhaps more important is its use as building block in Itô's theory of stochastic differential equations (Itô 1946). Such solutions behave locally like a Brownian motion with state dependent drift and variance parameters and provide more realistic stochastic models for the above (the Ornstein-Uhlenbeck process and geometric Brownian motion, respectively) and other random phenomena.

Before leaving Brownian motion for now, lets ask the obvious question suggested by Theorem 1.1: Why does the Gaussian integral arise? Assuming that there is a limit in Theorem 1.1, let Z_1, Z_2 and Z denote independent copies of this limit, then

$$(1.1) \qquad \sqrt{t_1}Z_1 + \sqrt{t_2}Z_2 \equiv \sqrt{t_1+t_2}Z \text{ (both sides have the same distribution.)}$$

This is easily seen by grouping together the first $[Nt_1]$ and next $[Nt_2] - [Nt_1]$ summands and applying the CLT to each summation. It turns out that this equivalence forces Z_i to have a Gaussian density (see p. 186 of Breiman (1968)).

2. Super-Brownian Motion (SBM)

Whereas Brownian motion models the time evolution of a single random point in Euclidean space, super-Brownian motion will model the time evolution of an entire random distribution of mass over $\mathbb{R}^d$. Brownian motion was constructed as the limit of rescaled random walks. Super-Brownian motion will now be constructed as a limit of rescaled branching random walks. This involves alternating the two random mechanisms of Brownian migration and near-critical reproduction.

Fix three parameters: $b > 0$ (branching rate), $\sigma^2 > 0$ (diffusion rate), and $g \in \mathbb{R}$ (growth rate). Let $W \in \mathbb{R}^d$ denote a generic random displacement with mean 0, covariance $\frac{\sigma^2}{b} I_{d\times d}$. Now start $O(N)$ particles in $\mathbb{R}^d$. With rate bN a particle at x dies and with rate $bN+g$ particle at x gives birth to particle at x$+W/\sqrt{N}$. The particles evolve independently of one another. We record the state of the population through its empirical measure, namely the finite measure, X_t^N, on $\mathbb{R}^d$, given by by

$$X_t^N(A) = \frac{1}{N}\#(\text{particles in } A \text{ at time } t).$$

The above space-time scaling is designed to ensure that if we keep track of the ancestral lineage leading up to a particle alive at time t then the limit of this lineage should be a Brownian path by Donsker's FCLT. The fact that the number of initial particles is $0(N)$ leading to our mass scaling factor of $1/N$ will be explained below.

Theorem 2.1 (S. Watanabe 1968). If $X_0^N \to X_0$, then

$$\lim_{N\to\infty} P(X^N \in A) = P_{X_0}(X \in A) \text{ “for any” set } A \text{ of measure-valued paths.}$$

X is a continuous measure-valued Markov process whose law, P_{X_0}, depends only on (X_0, b, σ^2, g).

We call X super-Brownian motion. There are a number of other constuctions possible involving discrete time, explicit Brownian migration, etc. but any reasonable combination of Brownian or random walk migration and near critical branching will produce the same limit. Super-Brownian motion is the central example of a larger class of stochastic processes called Dawson-Watanabe superprocesses (this terminology was introduced by Dynkin in the late 1980's) which may have more general migration and branching mechanisms.

Let $x_1, \ldots, x_{M_N}$ $(M_N = NX_0^N(\mathbb{R}^d))$ denote the positions of the initial ancestors at $t = 0$ and $X_t^{N,i}$ denote the contribution to X_t^N from the descendants of x_i. Since the particles evolve independently, clearly this decomposes X_t^N into a sum of M_N independent clusters:

$$X_t^N = \sum_{x_i} X_t^{N,i}. \tag{2.1}$$

To simplify the arithmetic let us assume for the current argument that $g = 0$, i.e., we are dealing with a critical branching mechanism. It is a sad fact of life that critical branching processes die out. The precise rate of extinction was found by Kolmogorov:

$$P(X_t^{N,i} \neq 0) \sim (Ntb)^{-1} \text{ as } N \to \infty. \tag{2.2}$$

This explains why we needed to start $0(N)$ particles to get a non-trivial limit in Theorem 2.1. Note also that (2.1) has $NX_0^N(\mathbb{R}^d)$ summands each non-zero independently with probability $\sim (Ntb)^{-1}$. If we let $N \to \infty$ in (2.1), one obtains

$$X_t = \sum_{j=1}^{M(t)} X_t^j,$$

where (X_t^j) are independent clusters descending from a single ancestor at $t = 0$, and $M(t)$ is a Poisson random variable, independent of the clusters X_t^j, with mean $\frac{X_0(R^d)}{tb}$. Here we have used the well-known Poisson approximation to the binomial distribution. This shows that for small t there is a large $(0(1/t))$ number of clusters contributing to the population while as $t \to \infty$, the number will eventually be 0 with probability 1. Hence SBM becomes extinct in finite time almost surely.

Remarks on SBM

A. The Canonical Measure $\mathbf{N_0}$

The clusters contributing to X_t are the fundamental building blocks of the process. It is natural to ask about the probability law of "a typical cluster" on the space of measure-valued paths which become extinct in finite time. The problem here is that the clusters with very short lifetimes will dominate any such law, in fact a "typical cluster" would have an infinitesimally small lifetime as clusters of lifetime greater than t should occur with intensity c/t by the above. To make room for all of these short clusters one is forced to introduce an infinite measure, $\mathbf{N_0}$, called the canonical measure, on the space of measure-valued paths with finite lifetime. $\mathbf{N_0}$ satisfies $\mathbf{N_0}(X_t \neq 0) = (tb)^{-1}$ and if the cluster X^j starts from the origin, $P(X_{t+\cdot}^j \in \cdot) = \mathbf{N_0}(X_{t+\cdot} \in \cdot | X_t \neq 0)$. In this way $\mathbf{N_0}$ governs the evolution of the clusters. Its mass on paths alive at time t is the intensity that they contribute to the SBM at time t and conditional on being alive at time t they give the law of the future evolution of these clusters. Probabilists may note the similarity between this canonical measure and excursion measures for Markov processes. In fact Le Gall (1999) (Sec. IV.1) builds $\mathbf{N_0}$ as the excursion measure from zero of his Brownian snake which runs up and down the family trees of all the individuals whoever existed. One can construct $\mathbf{N_0}$ from a branching random walk as follows (see, for example, Section II.7 of Perkins (2002)):

$$(2.3) \qquad \lim_{N\to\infty} NP(X^N \in A | X_0^N = N^{-1}\delta_0) = \mathbf{N_0}(X \in A) \text{ "for any" set of paths } A.$$

Note that the factor N in front of the probability is natural in light of the survival probability (2.2).

To get a sense of what SBM looks like in two spatial dimensions and see the role of this cluster decomposition, Figures 2.1-2.4 below are simulations of branching random walk on the two dimensional torus starting with $N = 40,000$ particles uniformly distributed over the torus, taking on 255 distinct colours. The simulations were provided by Achim Klenke and more can be found on his webpage. The particles are following a nearest neighbour random walk, so $\sigma^2 = 1/2$. The other parameters are $b = 1/2$ and $g = 0$. Time is measured in units of $1/N$ which means we have total initial mass 1. The colours are inherited by successive generations and so should give us an idea of the number of clusters contributing to the entire population. The figures are taken at $t = 0, t = .003, t = .14$ and $t = .56$. At this last time there are four colours present corresponding to four distinct clusters. The actual number of clusters has a Poisson distribution with mean $X_0(\mathbb{R}^2)/bt \sim 3.6$. From the pictures one may guess that the actual support of SBM is Lebesgue null and we will see in a moment that this is the case. One might also ask if the support is totally disconnected and in two dimensions this question remains unresolved.

INSERT FIGURES 2.1-2.4 SBM on a torus.

B. Additive Property

Run 2 independent SBM's X^1, X^2, then clearly $X^1 + X^2$ is a SBM starting at $X_0^1 + X_0^2$. This is clear from the corresponding property for the branching random walks which in turn follows from the independence of the evolutions of individual particles. This property is central to much of the analysis and stochastic analysis associated to SBM.

C. A Stochastic pde for SBM

"$X_t(dx) = X(t,x)dx$" where $X(t,x)$ is the unique solution of the stochastic pde

$$(SPDE) \qquad \frac{\partial X}{\partial t} = \frac{\sigma^2 \Delta X}{2} + gX + \sqrt{2bX}\dot{W}.$$

Here $\dot{W}$ is a space-time white noise. The meaning of the stochastic partial differential equation is as follows (see Walsh (1986) for a good introduction to SPDE's). If $\phi(x)$ is smooth and $\langle X_t, \phi \rangle = \int \phi(x) X_t(dx)$, then

$$(ISPDE) \qquad \langle X_t, \phi \rangle = \int_0^t \langle X_s, \sigma^2 \Delta \phi/2 + g\phi \rangle ds + \int_0^t \int \sqrt{2bX(s,x)}\phi(x) dW(s,x).$$

Here $dW(s,x) = w(s,x)\sqrt{ds}\sqrt{|dx|}$ where "$\{w(s,x)\}$ are independent normal random quantities", that is the above integral is the stochastic integral with respect to white noise. For $d = 1$, the above SPDE was obtained independently by Reimers (1989) and, Konno-Shiga (1988). If $d > 1$, $X_t(dx) \perp dx$ and $X(t,x)$ will not exist. Nonetheless one can still interpret the stochastic integral in (ISPDE) (one has an infinite integrand on an infinitesimal set) and use this to characterize X. This represents one of the rare occasions that a parabolic stochastic pde driven by white noise can be rigorously solved in more than one dimension.

The origins of the diffusion term and linear term (gX) in (SPDE) are clear but the fact that the branching dynamics gives rise to a term of the form $\sqrt{2bX}\dot{W}$ is less obvious. To make the square root function at least plausible, note that by the Additive Property we need

$$\sqrt{2bX^1}\dot{W}_1 + \sqrt{2bX^2}\dot{W}_2 \equiv \sqrt{2(b(X^1 + X^2)}\dot{W},$$

where $\equiv$ means their distributions are the same. This is true since by (1.1), $\sqrt{c_1}Z_1 + \sqrt{c_2}Z_2 \equiv \sqrt{c_1 + c_2}Z$ for independent normals, and we have conditioned on past information to treat $X^i(t,x)$ as constants.

3 Properties of Super-Brownian motion and Brownian Motion

If limit theorems are to be of any use it is important that the limit is reasonably well understood. In the past 20 years a number of methods have led to a fairly detailed understanding of super-Brownian motion. In this section we will discuss a sample of these results and methods by comparing them with their counterparts for ordinary Brownian motion.

A. PDE

1. Brownian Motion

If E_x denotes expection over all Brownian paths starting at x, then $u(t,x) = E_x(\phi(B_t))$ solves $\frac{\partial u}{\partial t} = \frac{\Delta}{2}u$, $u_0 = \phi$. This fact underlies the rich interaction between Brownian paths and elliptic boundary value problems.

2. Super-Brownian Motion

If E_{X_0} denotes expectation over all super-Brownian paths starting at the finite measure X_0, then

$$E_{X_0}(e^{-\langle X_t,\phi\rangle}) = e^{-\langle X_0,v_t\rangle}, \tag{3.1}$$

where

$$\frac{\partial v}{\partial t} = \frac{\sigma^2\Delta v}{2} - bv_t^2 + gv_t,\ v_0 = \phi \geq 0.$$

Note that the fact that (3.1) holds for some function $v(t,x)$ is to be expected from the Additive Property of SBM. That v solves the above nonlinear pde (only the quadratic term perhaps needs some justification) is an exercise in stochastic calculus using (ISPDE). (3.1) gives a convenient characterization of SBM which was used by Watanabe (1968) in his proof of the convergence theorem (Theorem 2.1). More general Dawson-Watanabe superprocesses will lead to other nonlinear pde's with, for example, more general nonlinear terms on the right-hand side. This connection allows one to use nonlinear pde's to derive properties of SBM, e.g., see Iscoe (1986). In the other direction Dynkin (2002), Kuznetsov and Le Gall pursued a program to characterize, and probabilistically represent, solutions to the corresponding nonlinear elliptic equation in terms of their boundary behaviour. Recently Mselati (2002) completed this program for the quadratic nonlinearity discussed here.

B. Longterm Behaviour

1. Brownian Motion

If $d \leq 2$, Brownian motion is neighbourhood recurrent, i.e., $B_t \in G$ infinitely often as $t \to \infty$ for any open set G a.s. If $d \geq 3$, Brownian motion is transient, i.e., $\lim_{t\to\infty} \|B(t)\| = \infty$ a.s.

2. Super-Brownian Motion

If X_0 finite then X becomes extinct in finite time almost surely iff $g \le 0$. If $g = 0$ we have already noted this and if $g > 0$ it is an easy exercise to find the (positive) probability of surviving forever using (3.1) with $\phi = 1$, in which case we have a simple o.d.e. which can be solved explicitly. This survival probability is $1 - \exp(-X_0(\mathbb{R}^d)g/b)$.

In the critical case $g = 0$, Dawson (1977) obtained a finer dichotomy by considering infinite initial conditions such as $X_0(dx) = mdx$:

(a) For $d \le 2$, for any $R > 0$, $P(X_t(|x| \le R) > 0)$ approaches 0 as $t \to \infty$.

(b) For $d \ge 3$, $\lim_{t\to\infty} P(X_t \in A) = P(X_\infty \in A)$ "for any" set of measures A where $E(X_\infty(A)) = mLeb(A)$ and X_∞ are (the only extremal) equilibrium distributions for X.

The fact that these are the only extremal equilibrium measures is due to Bramson, Cox and Greven. This dichotomy may be understood in terms of the recurrence/transience dichotomy for Brownian motion. For $d \ge 3$, the finite SBM consisting of descendants of particles initially in $\{|x| \le R\}$ will become extinct in finite time but the transience of particles near ∞ means there is a steady stream of replacement particles leading to a non-trivial equilibrium state. For $d \le 2$ the recurrence of Brownian motion means that the particles at ∞ cannot effectively replenish the mass in $\{|x| \le R\}$. A related process called the historical process has been used to decompose these equilibrium measures according to ancestral histories (see Dawson and Perkins (1991)).

C. Local Behaviour

1. Brownian Motion

In more than one dimension, the range of Brownian motion is a random set of Lebesgue measure zero and Hausdorff dimension 2. This is due to Brownian scaling which implies that for small t, $B(t)$ is about distance $t^{1/2}$ from its starting point. More precisely Ray and Taylor showed that if

$$\phi_d(r) = \begin{cases} r^2 \log(1/r) \log\log\log(1/r) & \text{if } d = 2 \\ r^2 \log\log 1/r & \text{if } d \ge 3, \end{cases}$$

then $\phi_d - m(\{B_s : s \le t\}) = c_d t$ for all $t \ge 0$ a.s. Here $\phi_d - m$ denotes Hausdorff ϕ_d-measure and c_d is a positive constant.

2. Super-Brownian Motion

Let $S(X_t)$ = closed support of X_t. If $d \ge 2$, $S(X_t)$ is a Lebesgue null set of Hausdorff dimension 2 for all $t > 0$ a.s. This 2 comes from the Brownian scaling already mentioned and the fact that $S(X_t)$ may be decomposed into the union of $0(1/t)$ clusters for t small as discussed in Section 3. More precisely SBM has the same exact Hausdorff measure functions as Brownian motion (Dawson-Perkins (1991), Le Gall-Perkins (1995)):

$$X_t(A) = C_d \frac{b}{\sigma^2} \phi_d - m(A \cap S(X_t)) \;\forall A \; a.s., \; t > 0.$$

This result indicates that X_t can be recovered from the pictures of $S(X_t)$ in Figures 2.2–2.6 because there is no non-trivial measure of local density of mass. The latter may again be explained using the transience/recurrence dichotomy of Brownian motion. For $d \geq 3$ consider a typical point x in $S(X_t)$ and look at the demographics of the mass in $\{y : |y - x| \leq r\}$ for r small. Due to the transience, particles which branched off from the family tree of x some time ago will not return to this small ball and so the mass we do see will consist of close cousins of x. This means that the actual mass about x and a nearby point x' will depend on independent Brownian increments as they will have disjoint sets of close cousins and so the law of large numbers will guarantee a constant local density when we average these local densities over small sets. If $d = 1$, the strong (point) recurrence of Brownian motion will mean there are strong correlations between the local mass at nearby points and hence we get a random density, $X_t(dx) = X(t,x)dx$, where $X(t,x)$ is the unique solution of (SPDE) as has already been discussed. The two-dimensional case is the most delicate as the neighbourhood recurrence is very weak. In this case, the proof in Le Gall-Perkins (1995) uses the Brownian snake described earlier.

4. Voter Model

The voter model was introduced almost 30 years ago by a number of different researchers and is one of the fundamental examples of interacting particle systems (see Liggett (1985)). Each site x in $\mathbb{Z}^d$ is occupied by an individual with type 0 or 1 (e.g. a Democrat or Republican) and so $\xi_t(x) = 0$ or 1. Each type tries to take over the territory held by the other: with rate 1, the type at x chooses a nearest neighbour at random and imposes its type on it. If the chosen neighbour is of the same type nothing happens but if it has a different type it switches to the type of x. At the same time the neighbours are of course independently trying to impose their type on x, leading to completely symmetric dynamics. One could introduce a biased voter model in which 1's convert neighbouring 0's with rate α and 0's convert neighbouring 1's with rate 1. We analyze the large scale behaviour at the critical value $\alpha = 1$.

Rescaling space and time as for Brownian motion, we set $\xi_t^N(x) = \xi_{tN}(x\sqrt{N})$, $x \in \mathbb{Z}^d/\sqrt{N}$. Let $V_t^N(A) = m_N^{-1} \sum_{x \in A} \xi_t^N(x)$ denote the empirical distribution of 1's, where

$$m_N = \begin{cases} N, & \text{if } d \geq 3 \\ N/\log N & \text{if } d = 2. \end{cases}$$

If (S_n) is a nearest neighbour random walk, let

$$p_{\text{esc}} = \begin{cases} P_0(S_n \neq 0 \ \forall n \geq 1) & \text{if } d \geq 3 \\ \pi = 2\pi\sigma^2 = \lim_N (\log N) P(S_n \neq 0 \ \forall 1 \leq n \leq N) & \text{if } d = 2. \end{cases}$$

Theorem 4.1 (Cox, Durrett, Perkins (2000)). Assume $d \geq 2$ qnd $V_0^N \to X_0$. Then $\lim_{N\to\infty} P(V^N \in A) = P_{X_0}(X \in A)$ "$\forall A$", where X is SBM with $g = 0$, $\sigma^2 = 1/d$ and $b = p_{\text{esc}}$.

"Proof". To make the result plausible, reinterpret the dynamics as follows:
$\xi_t(x) = 1 \iff$ particle at x; $\xi_t(x) = 0 \iff$ no particle at x.
Define $f_0^N(t,x) = \{\text{no. of neighbouring 0's to } x\}/2d$. Then a particle at x dies with rate $Nf_0^N(t,x)$, and with rate $Nf_0^N(t,x)$ produces a child at y chosen at random from the neighbouring sites of type 0. The "death" corresponds to a conversion by one the neighbouring 0's and the "birth" is due to the 1 at x changing one of the neighbouring 0's. This is similar to our earlier description of branching random walk with $g = 0$, $\sigma^2 = 1/d$ and a random $b = f_0^N(t,x)$. Hence the proof must show that if $\xi_t(x) = 1$, then $f_0^N(t,x) \sim p_{\rm esc}$ on average. This is done by a law of large numbers argument providing we can show
(i)$E(f_0^N(t,x)|\xi_t(x) = 1) \sim p_{\rm esc}$
(ii) $f_0^N(t,x)$ and $f_0^N(t,x')$ are asymptotically uncorrelated for $|x - x'| >> 1/\sqrt{N}$.
These conditions require some conditional third moments and these are calculated with the help of a "dual coalescing random walk".

Remarks. 1. If we change the local dynamics so that instead of selecting a nearest neighbour of x at random the particle chooses y with probability $p(y-x)$ for some symmetric finite variance kernel p, then the above result remains valid where $p_{\rm esc}$ is defined in terms of a new random walk with step distribution p. Hence the local dynamics affect the limit only through the values of $p_{\rm esc}$ and σ^2.

2. If $d = 2$, the recurrence of (the coalescing dual) random walk leads to clustering of the voter model and $Ef_0^N(t,x) \sim (\log N)^{-1}$, leading to a deterministic limit. To counteract this we must increase the branching rate by $\log N$ and this is achieved by increasing the mass of each particle by $\log N$–recall for our rescaled branching random walks, the inverse mass and branching rate were proportional.

3. If $d = 1$ consider an initial distribution of 1's over an interval I. For the unscaled voter model it should be easy to see that the set of sites occupied by 1's evolves as an interval whose endpoints are continuous time random walks which evolve independently until they meet and annihilate each other. If $m_N = \sqrt{N}$ then it is not hard to use Donsker's FCLT (Theorem 1.2) to see that our rescaled voter models converge to Lebesgue measure on an interval whose endpoints are given by a pair of annihilating Brownian motions. This is due to R. Arratia.

An Application

Let ξ_t be voter model starting from a single 1 at $x = 0$, take $d \geq 2$, and set $S(\xi_t) = \{x : \xi_t(x) = 1\}$. Bramson, Griffeath (1980) posed the following question:
Conditional on $\xi_t \neq 0$, what is asymptotic shape of $S(\xi_t)$ as $t \to \infty$?
Williams and Bjerknes (1972) had proposed the biased voter model (with $\alpha > 1$) as a stochastic model for tumour growth and Bramson and Griffeath had answered the above question by showing that $S(\xi_t)$ grows linearly to an asymptotic shape of positive volume.

They viewed the corresponding result for the voter model as the asymptotically critical case of the tumour growth result and noted that the result would necessarily be quite different.

It is easy to guess the answer in light of Theorem 4.1. Set $t = N$ and note that $S(\xi_N)/\sqrt{N} = S(V_1^N)$. We therefore expect

$$P(S(\xi_N)/\sqrt{N} \in \cdot | \xi_N \neq 0) \sim P_{\delta_0/N}(S(X_1) \in \cdot | X_1 \neq 0).$$

As we are starting the voter model with a single 1 at the origin we would expect the right-hand side to be the law of a single cluster of SBM, that is SBM conditioned on survival under the canonical measure. These heuristics were proved by Bramson, Cox and Le Gall (2001) where one can also find other applications and extensions of Theorem 4.1.

Theorem 4.2 (Bramson, Cox, Le Gall (2001)). For $d \geq 2$,

$$\lim_{t\to\infty} P(S(\xi_t)/\sqrt{t} \in A | \xi_t \neq 0) = \mathbf{N_0}(S(X_1) \in A | X_1 \neq 0) \text{ “for all” sets of sets } A.$$

It should be stressed that the limiting random set in the above is a well-understood object–we have exact Hausdorff measure functions, precise results on its dynamics and multiple points and characterizations of its polar sets to name only a few of its properties.

5. The Contact Process

The contact process, introduced by Harris (1974) is a stochastic epidemic model. Let $L \in \mathcal{N}$ be the infection range of the process and, rescaling the lattice so that the range becomes 1, let $S_t \subset \mathbb{Z}^d/L$ be the set of infected sites at time t. The dynamics are as follows: a bacterium at $x \in \mathbb{Z}^d$ dies with rate 1, and with rate λ infects $y \in \mathbb{Z}^d/L$ chosen at random in $[x-1, x+1]^d$ **provided y is not infected** (if it is, nothing happens). Distinct bacteria have independent death and infection times.

Basic Fact: $\exists \lambda_c(L) > 1$ such that $\lambda > \lambda_c \iff P(S_t \neq \emptyset \; \forall t | S_0 = \{0\}) > 0$.

The precise value of λ_c is not known although a number of bounds exist in the literature. Note that if we did not suppress infections onto occupied sites (and hence allow multiple occupancies) the above model would be a (biased) branching random walk and λ_c would be 1. This makes it reasonable to expect $\lambda_c > 1$ to compensate for this suppression of “births”. As this latter effect should become negligible as $L \to \infty$ we expect (as is the case) $\lim_{L\to\infty} \lambda_c(L) = 1$

Let $L = L_N \to \infty$, and set $\lambda = 1 + \frac{\theta}{N}$ where θ is a real parameter. Then $\lambda \sim \lambda_c(L_N) \sim 1$ and so we are analyzing the long-range contact process near criticality. Now invoke Brownian space-time rescaling and assign mass per site as for SBM: $S_t^N = S_{Nt}/\sqrt{N}$, $X_t^N(A) = \#(S_t^N \cap A)/N$, $A \subset \mathbb{R}^d$. Therefore a particle at x with rate N dies, and with rate $N\lambda = N + \theta$ gives birth onto y, uniformly chosen in

$$[x - N^{-1/2}, x + N^{-1/2}]^d \cap (\mathbb{Z}^d/(L_N\sqrt{N})) = \mathcal{N}_N(x),$$

but the child is killed if y is occupied. This gives $b = 1$, $\sigma^2 \sim 1/3$ (the variance of each component of a uniform distribution on $[-1,1]^d$) and a random growth rate

$$g = \theta - (N+\theta) f_1^N(t,x), \text{ where } f_1^N \text{ is the frequency of 1's in } \mathcal{N}_N(x).$$

Comparing the above with our construction of SBM, we must choose L_N so that conditional on x being occupied, $f_1^N(t,x)(N+\theta) \sim 0(1)$. If we pretend we are dealing with a branching random walk the (conditional) mean of the above is not hard to find and leads to

$$L_N = \begin{cases} N^{1/d} & d \geq 3 \\ (N \log N)^{1/2} & d = 2 \\ N^{3/2} & d = 1. \end{cases}$$

Let $(U_n, n \in \mathbb{Z}_+)$ denote a random walk with uniformly distributed steps over $[-1,1]^d$ $(U_0 = 0)$.

Theorem 5.1 (Durrett-Perkins (1999)). Assume $X_0^N \to X_0$ non-atomic and $d \geq 2$. Then

$$\lim_{N \to \infty} P(X^N \in A) = P_{X_0}(X \in A) \text{ "for all" } A,$$

where X is SBM with $b = 1$, $\sigma^2 = 1/3$, $g = \theta - C(d)$, and

$$C(d) = \begin{cases} 2^{-d} \sum_{n=1}^{\infty} P(U_n \in [-1,1]^d) & d \geq 3 \\ 3/(2\pi) = \lim_N \frac{2^{-2}}{\log N} \sum_{n=1}^{N} P(U_n \in [-1,1]^d) & d = 2. \end{cases}$$

Remark. If $d = 1$ Mueller and Tribe (1994) confirmed an earlier conjecture of Durrett (1988) and showed that if X_0 has a bounded continuous density then $\lim_{N\to\infty} P(X^N \in A) = P_{X_0}(X \in A)$ "$\forall A$", where X is the unique solution of the stochastic pde

$$\frac{\partial X}{\partial t} = X''/6 + \sqrt{2X}\dot{W} + (\theta - X)X.$$

"Proof". From the discussion above, and by comparing our setting with the convergence theorem for branching random walk, we must show that if $\xi_t(x) = 1$, then $f_1^N(t,x) \sim C(d)/N$ on average. The fact that this local density is effectively constant follows from the same reasoning we used to argue that there is no non-trivial local density for SBM for $d \geq 2$. Namely, the only particles contributing to the local density $f_1^N(t,x)$ about an occupied set x are close cousins of the particle at x and so these densities are effectively independent for $|x - x'| >> N^{-1/2}$. (If $d = 1$ there is of course a non-trivial measure of local density for SBM, namely the density $X(t,x)$ of X and if we set $g = \theta - X(t,x)$ and $\sigma^2 = 1/3$ in (SPDE) we obtain the above stochastic pde found by Mueller and Tribe (1994).) The reason we can actually calculate the constant requires another argument which we will not pursue here. The precise value of $C(d)$ comes from corresponding density for branching random walk–one decomposes the particles contributing to the local density according to when they split off from the family tree of the particle at x, leading to the sum over n in the definition of $C(d)$.

An Application

After finding the right order of magnitude of the rate of convergence of $\lambda_c(L)$ to 1 as $L \to \infty$, Bramson, Durrett and Swindle (1989) posed the problem of finding the exact first order asymptotics for $\lambda_c(L)$. A formal interchange of limits readily gives the answer from the limit theorem as we now show. The justification of this interchange of limits does take some work.

Theorem 5.2 (Durrett-Perkins (1999)). If $C(d)$ as above then

$$\lambda_c(L) - 1 \sim \begin{cases} C(d)/L^d & d \geq 3 \\ \frac{3}{\pi} \frac{\log L}{L^2} & d = 2. \end{cases}$$

"Proof". $\theta > C(d) \iff P_{X_0}(X_t \neq 0 \ \forall t > 0) > 0$
$\iff P(X_t^N \neq 0 \ \forall t > 0) > 0$ for N large
$\iff \lambda_c(L_N) < 1 + (\theta/N)$ (recall $\lambda = 1 + (\theta/N)$).

For the first equivalence recall SBM will survive for all $t > 0$ with positive probability iff $g = \theta - C(d) > 0$. The second line is the interchange of limits mentioned above. The third line reflects the fact that if we survive with positive probability starting with X_0^N then we will survive with positive probability starting with a single particle (a superposition of independent contact processes will dominate a single contact process with the same initial condition) and so λ_c must be less than our value of λ. The above equivalence implies $\lambda_c(L_N) \sim 1 + C(d)/N$. Now invert L_N to obtain the above result.

If we are claiming that SBM is a universal limit of spatially distributed random systems near criticality, then it is natural to ask if the above result remains valid (with appropriate parameters) if we fix L and set $\lambda = \lambda_c(L)$. Particles would then be giving birth at rate $\lambda_c(L)N$ and dying with rate N, leading to an asymptotically infinite growth rate which must be exactly compensated by the killing of particles which land on occupied sites. Such a precise cancellation of infinities would be remarkable but ongoing work of Sakai and van der Hofstad suggest this is in fact the case for $d > 4$ and L large enough. The reason for the optimism here is that when $d > 4$ and $L > L_0(d)$, the result has already been proved for the discrete time analogue of the contact process, oriented percolation, by van der Hofstad and Slade (2001).

6. Oriented Percolation

Again fix an interaction range $L \in \mathcal{N}$, and for $x \in \mathbb{Z}^d/L$ let

$$\mathcal{N}_L(x) = ([x-1, x+1]^d - \{x\}) \cap (\mathbb{Z}^d/L).$$

Let $0 < \lambda \le \#\mathcal{N}_L$. If $y \in \mathcal{N}_L(x)$, declare the directed bond from (x, n) to $(y, n+1)$ to be open, independently of the other bonds, with probability $\lambda/\#\mathcal{N}_L$. Now consider a water source at $(0,0) \in \mathbb{Z}^d/L \times \mathbb{Z}_+$ and have water only flow along open bonds in the direction of increasing n. Set $S_n = \{x : (x,n) \text{ is wet}\}$. Another interpretation would be to have the open bonds correspond to the creation of offspring from a parent and cull the offspring so that there is at most one individual per site. S_n then denotes the set of occupied sites in the nth generation. Note that parents then die after producing their children and λ is the mean number of offspring before the culling takes place.

As for the contact process $\exists \lambda_c(L) > 1$ such that $\lambda > \lambda_c \iff P(S_n \ne \emptyset \; \forall n) > 0$. Now rescale space and time in the usual way and let X_t^N be the finite measure which assigns mass $1/N$ to each point in $S_{[Nt]}/\sqrt{N}$.

Theorem 6.1 (van der Hofstad, Slade (2001)). If $d > 4$, $L > L_0(d)$ and $\lambda = \lambda_c(L)$, there are positive constants A, σ^2 and b such that

$$\lim_N NP(X^N \in \cdot) = A\mathbf{N_0}(X \in \cdot)$$

in the sense of convergence of finite dimensional distributions. Here $\mathbf{N_0}$ is the canonical measure of SBM with parameters σ^2, b and $g = 0$.

The proof uses a recent inductive approach to the lace expansion of van der Hofstad and Slade to directly calculate the rth moment measure $NE(X_{t_1}^N(dx_1)...X_{t_r}^N(dx_r))$ and show it converges to the known moment measures for SBM under the canonical measure.

Remarks. 1. As we are looking at a percolation cluster starting from a single point $(0,0)$ it is natural to get a single cluster of SBM, that is, $\mathbf{N_0}$, in the limit (compare with (2.3)).

2. The lace expansion provides a method for doing complex inclusion-exclusion counting arguments. In the course of the proof certain terms are shown to be neglible in the limit. In the limit these terms will correspond to the probability that a Brownian path collides with an independent super-Brownian cluster. This probability is 0 iff $d \ge 4$, with $d = 4$ critical (see Barlow and Perkins (1994)), and leads to the $d > 4$ condition in the theorem. The fact that $d = 4$ is critical for the non-existence of these collisions can be seen, with a bit of imagination, by recalling from Section 3 that the support of SBM has dimension 2 and the range of the Brownian path has dimension 2, and noting $2 + 2 = 4$. In lower dimensions the scaling limit of oriented percolation is not known but it will not be SBM. (REF?)

3. The inductive approach to the lace expansion mentioned above also produces estimates for the parameters in the Theorem and in particular shows that as $L \to \infty$, $A = 1 + 0(L^{-d})$, $b = 1 + 0(L^{-d})$, $\sigma^2 = 1/3 + 0(L^{-1})$, and $\lambda_c(L) = 1 + 0(L^{-d})$. In fact $\lambda_c(L)$ is characterized as the root of a functional equation and the machinery has the potential of finding higher order expansions for $\lambda_c(L)$ (as it has done in other contexts).

7. Other Models

We have indicated how rescaled voter models, long range contact processes and fixed range oriented percolation all converge to SBM for sufficiently large dimensions. In this section some further connections between SBM and other processes are briefly described.

(a) Rescaled Lattice Trees (Derbez-Slade (1998))

A lattice tree is a connected set of "neighbouring" (range L) bonds in $\mathbb{Z}^d$ with no cycles containing the origin. Pick a such a lattice tree with N^2 vertices at random and let X^N be the random probability which assigns mass N^{-2} to each vertex scaled by $1/\sqrt{N}$. Derbez and Slade (1998) confirmed a conjecture of Aldous (1993) by showing that for $d > 8$ and $L > L_0(d)$, $P(X^N \in \cdot)$ approaches $\mathbf{N_0}(\int_0^\infty X_s ds \in \cdot | \int_0^\infty X_s(\mathbb{R}^d)ds = 1)$. Here $\mathbf{N_0}$ is the canonical measure of SBM with $g = 0$ and appropriate $b, \sigma^2 > 0$. This limiting random probability is called integrated super-excursion (ISE). Note that it corresponds to taking a cluster of SBM and looking at the set of points in space which are traced out by the cluster over its lifetime and then conditioning the total "occupation mass" of this cluster to be one. The criticality of this model (and reason for Aldous' conjecture) arises from the fact that uniform measure on abstract tree shapes with N^2 nodes may be constructed by conditioning a critical Galton-Watson branching process with Poisson offspring distribution to have total progeny size N^2. Again different limits are expected for $d \leq 8$.

(b) Lotka-Volterra Models (Cox-Perkins (2002))

The following interacting particle system was introduced by Neuhauser and Pacala (1999) to model two types competing for resources. All sites in $\mathbb{Z}^d$ are occupied by a particle of one type or the other. The dynamics which depend on the parameters, α_i governing the interspecies competive effect on type i, are as follows:

f_i is local frequency of type i,

$0 \to 1$ with rate $(f_0 + \alpha_0 f_1)f_1$; $\quad 1 \to 0$ rate $(f_1 + \alpha_1 f_0)f_0$.

The first factor above gives the death rate of the particle at a given site (the effect of self-competition has been normalized to 1 for both types) and the second factor represents the immediate invasion of the vacant site by a type chosen at random from the neighbouring sites. If X_t^N is rescaled empirical measure of 1's then in different near-critical regimes of (α_0, α_1) one gets convergence to SBM with non-trivial growth rates depending on (α_0, α_1). The regimes in question are α_i close to 1 or, for the long range setting, α_0 fixed and α_1 close to 1. The methods introduced here appear to extend and unify the arguments used above for the voter model and long range contact process.

(c) Fleming-Viot Models (Etheridge-March (1991), (Perkins (1992))

These represent a large class of stochastic processes used to model the distibution of genotypes in a population undergoing random sampling, selection and mutation. We restrict ourselves to models which parallel SBM although in this setting it would be more natural to consider more general state spaces and mutation opertors (corresponding to more general Dawson-Watanable superprocesses). The results below extend to these more general settings without difficulty. $\mathbb{R}^d$ is now a space of allele types. Migration represents mutation of type of offspring from that of parent.

N is a *fixed* population size due to a finite carrying capacity. Branching now becomes resampling from gene pool: at $t = i/N$ each particle (j) is replaced by k_j offspring of "neighbouring types", where in the absence of selective effects, $(k_1, \ldots, k_N)$ is multinomial $(N; 1/N, \ldots, 1/N)$. Here the shift of the offspring to a "neighbouring type" is due to mutation and the multinomial distribution arises from selecting at random from the effectively infinite set of potential adults, those which reach maturity.

The empirical probability measure of types, V_t^N, is a random probability which converges to V_t, the Fleming-Viot process. Obviously this process cannot be SBM as it has total mass equal to one but it can be constructed from SBM in the most naive manner–simply condition the latter to have total mass one for all time.

Theorem 7.1 (Etheridge-March (1991)).

$$P(V \in \cdot) = \lim_n P_{V_0}(X \in \cdot | \sup_{t \leq n} |X_t(\mathbb{R}^d) - 1| \leq 1/n).$$

The righthand side is SBM ($g = 0$) conditioned to be a probability-valued process.

Here SBM is used as a building block for the Fleming-Viot process V. Recall that in constructing realistic models of random phenomena much of the importance of Brownian motion derives from its role in Itô's program of constructing processes which are locally Brownian but with state dependent drifts and diffusion rates. These arise as solutions of stochastic differntial equations. We have seen that SBM arises as a universal limit of a variety of critical spatially distributed systems. Here the local nature of the interactions and a law of large numbers effect led to constant but non-trivial parameters governing the evolution of the SBM. For many stochastic systems such a law of large numbers effect will not persist and a truly interactive stochastic limit arises:

· mutual attraction/repulsion of particles
· low density competing species or predator-prey systems
· symbiotic/diploid branching models

Such processes can be viewed as behaving locally like SBM but with state dependent coefficients b, σ^2 and g, which may even be singular. They may be modelled by generalized solutions of stochastic pde's of the form (compare to (SPDE)):

$$\frac{\partial X}{\partial t}(x) = (A^*_{X_t} + g(x, X_t))X_t(x) + \sqrt{b(x, X_t)X_t(x)}\dot{W}_t(x),$$

where,

$$A_{X_t}\phi(x) = \frac{1}{2}\sum_{i,j} a_{ij}(x, X_t)\phi_{ij}(x) + \sum_i b_i(x, X_t)\phi_i(x)$$

and $\dot{W}$ is a space-time white noise.

There has been some recent progress here both in some of the specific examples above and in the general theory (the latter with Athreya, Barlow and Bass) but this would be the subject of another lecture.

References.

L. Bachelier (1900), Théorie de la spéculation, Thèse Paris, *Ann. Ec. Norm. Sup. s. 3,* **17** 21-86.

M. Barlow and E. Perkins (1994), On the filtration of historical Brownian motion, *Ann. Prob.* **22**, 1273-1294.

M. Bramson, T. Cox, and J.-F. Le Gall, Super-Brownian limits of voter model clusters, *Ann. Prob.* **29**, 1001-1032.

M. Bramson, R. Durrett, G. Swindle (1989), Statistical mechanics of crabgrass, *Ann. Prob.* **17**, 444-481.

M. Bramson and D. Griffeath (1980), Asymptotics for interacting particle systems on $\mathbb{Z}^d$, *Z. Wahr. verw. Geb.* **53**, 183-196.

L. Breiman (1968) *Probability*, Reading, Addison-Wesley.

R. Brown (1828), Philosophical Magazine N.S. **4**, 161-173.

T. Cox, R. Durrett, and E. Perkins (2000), Rescaled voter models converge to super-Brownian motion, *Ann. Prob.* **28**, 185-234.

T. Cox and E. Perkins (2002), Rescaled Lotka-Volterra models converge to super-Brownian motion, working document.

D. Dawson (1977), The critical measure diffusion, *Z. Wahr. verw. Geb.* **40**, 125-145.

D. Dawson and E. Perkins (1991), Historical Processes, *Mem. Amer. Math. Soc.* **93**.

E. Derbez and G. Slade (1998), The scaling limit of lattice trees in high dimensions, *Comm. Math. Phys.* **193**, 69-104.

M. Donsker (1951), An invariance principle for certain probability limit theorems, *Mem. Amer. Math. Soc.* **6**.

R. Durrett and E. Perkins (1999), Rescaled contact processes converge to super-Brownian motion for $d \geq 2$, *Prob. Th. Rel. Fields* **114**, 309-399.

E. Dynkin (2002), *Diffusions, Superdiffusions and Partial Differential Equations*, AMS Colloq. Pub. **50**, Providence, AMS.

A. Etheridge and P. March (1991), A note on superprocesses, *Prob. Th. Rel. Fields* **89**, 141-147.

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

I. Iscoe (1986), A weighted occupation time for a class of measure-valued branching processes, *Prob. Th. Rel. Fields* **16**, 85-116.

K. Itô (1946), On a stochastic integral equation, *Proc. Imp. Acad. Tokyo*, **22**, 32-35.

N. Konno and T. Shiga (1988), Stochastic differential equations for some measure-valued diffusions, *Prob. Th. Rel. Fields* **78**, 201-225.

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

J.-F. Le Gall and E. Perkins (1995), The Hausdorff measure of the support of two-dimensional super-Brownian motion, *Ann. Prob.* **23**, 1719-1747.

T. Liggett (1985), *Interacting Particle Systems*, Springer-Verlag, New York.

B. Mselati (2002), Thèse de doctorat de L'Université Paris 6.

C. Mueller and R. Tribe (1994), Stochastic pde's arising from the long range contact and long range voter processes, *Prob. Th. Rel. Fields* **102**, 519-546.

E. Perkins (1992), Conditional Dawson-Watanabe processes and Fleming-Viot processes, in *Seminar on Stochastic Processes 1991*, Boston, Birkhäuser, 142-155.

E. Perkins (2002), *Dawson-Watanabe Superprocesses and Measure-Valued Diffusions*, in Ecole d'Eté de Probabilités de Saint Flour 1999.

M. Reimers (1989), One-dimensional stochastic pde's and the branching measure diffusion, *Prob. Th. Rel. Fields* **81**, 319-340.

J. B. Walsh (1986), *An Introduction to Stochastic Partial Differential Eqauations* , in Ecole d'Eté de Probabilités de Saint Flour 1984.

S. Watanabe (1968), A limit theorem of branching processes and continuous state branching processes, *J. Math. Kyoto U.* **8**, 141-167.

T. Williams and R. Bjerknes (1972), Stochastic model for abnormal clone spread through epithelial basal layer, *Nature* **236**, 19-21.