Fields Institute Distinguished Lecture Series In Statistical Sciences

Probabilistic Phenomena in Mathematics and Science

Don Dawson

UBIQUITY OF PROBABILISTIC PHENOMENA

Certain basic statistical patterns and probabilistic structures arise in empirical data and theoretical models from completely different fields.

The Workshop on Current and Emerging Research Opportunities in Probability – 2002 identified the following areas

- Algorithms
- Statistical Physics
- Dynamical and physical systems
- Complex networks
- Mathematical finance, risk and dependency
- Perception in artificial systems
- Genetics and ecology

Outline

FIRST CIRCLE OF IDEAS:

Probability laws and universality classes

SECOND CIRCLE OF IDEAS:

Markovian dynamics and conditional independence.

THIRD CIRCLE OF IDEAS:

Probabilistic modelling of reversible interacting systems

FOURTH CIRCLE OF IDEAS:

Spatially distributed nonreversible stochastic dynamics

FIFTH CIRCLE OF IDEAS:

Hierarchy, Genealogy and History

SIXTH CIRCLE OF IDEAS:

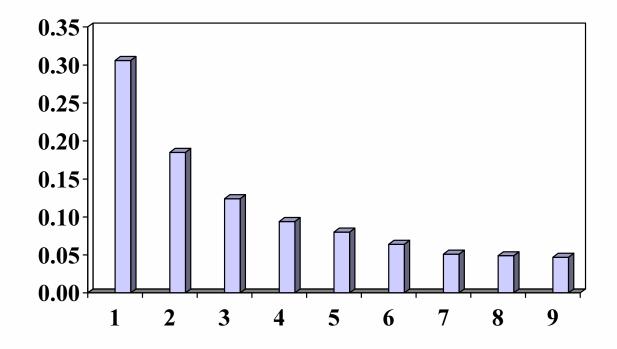
Universality classes of spatial and space-time structures

Empirical discovery of probability laws

Example:Benford's Law

- Simon Newcomb (1881) tables of logarithms.
- Frank Benford (1938) 20,229 data sets (molecular weights of chemical compounds, population sizes, etc.

Probability of first digit



Benford's distribution of first digit

$$p = log_{10}(1+1/d)$$

where p = the probability that the first significant digit is d.

- Statistical derivation T.P. Hill (1996)
 - Introduced idea of Scale invariant and Base invariant distributions

Discovery of the Central Limit Theorem

The Normal (Gaussian) Law

$$P(a < X < b) = \frac{1}{\sqrt{2\pi}} \int_{a}^{b} e^{-\frac{x^{2}}{2}} dx$$

Abraham De Moivre 1733 - aproximation to Binomial - games of chance

Johann Carl Friedrich Gauss 1809 - Law of Errors - astronomical data

Adolphe Quetelet ~1830 "the average man"

Francis Galton ~1860 - biological data - inheritance

$$x_n(t) := rac{\cos(\lambda_1 t) + \cos(\lambda_2 t) + \cdots + \cos(\lambda_n t)}{\sqrt{n}}$$

 $\{\lambda_i\}$ are are linearly independent over the field of rationals.

$$L_n(a,b) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^T 1_{(a,b)}(x_n(t)) dt$$

$$\lim_{n \to \infty} L_n(a,b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx$$

Q. WHY does the Gauss (normal) law arise?

"Even before the climax of our search for the meaning of independence was reached, it became abundantly clear why tout le monde was justified in believing in the loi des erreurs. It proved to be both un fait d'observation and une théorème de mathématiques."

- referring to his work with Steinhaus (1935-38)

Mark Kac - Enigmas of Chance

Basic Concepts

$$X_1,\ldots,X_N$$

zero mean independent identically distributed random variables.

Independence and the Characteristic Function

$$E(e^{i\theta S_N}) = E(e^{i\theta X_1} \cdots e^{i\theta X_N}) = E(e^{i\theta X_1}) \cdots E(e^{i\theta X_N})$$

where $S_N := (X_1 + \dots X_N)$

Scaling Limit

$$X_n(t) := rac{1}{a_n} S_{\lfloor nt \rfloor}$$

$$\Psi_t^n(\theta) = E(e^{i\theta X_n(t)}) \to \Psi_t(\theta)$$

Properties of the Scaling Limit

$$\Psi_{t+s} = \Psi_t \Psi_s$$

$$= E(e^{i\theta(X_t - X_0)} e^{i\theta(X_{t+s} - X_t)})$$

$$P_{t+s} = P_t \bigstar P_s \quad \text{Convolution Semigroup}$$

• Infinite Divisibility

Khinchin-Levy Representation

• Functional Fixed Point Equation

$$\psi_1(\theta) = F(\psi_1)(\theta)$$

$$F(\psi_1)(\theta) = \psi_1^2(\alpha\theta) \text{ for some } \alpha > 0.$$

Universality Classes Fixed Points and their Domains of attraction

• The Gaussian Law - "Short Tails"

$$\Psi_t(heta) = \exp\left(-rac{ heta^2 t}{2}
ight),$$

$$P_t(dx) = rac{1}{\sqrt{2\pi t}}e^{-rac{|x|^2}{2t}}dx \qquad ext{in } \mathbb{R}^d$$

• The Stable Laws - "Long Tails" - "Noah Effect"

$$\Psi_t(\theta) = \exp(-t|\theta|^{\alpha}), \quad 0 < \alpha < 2$$

The Function Space Perspective - Brownian Motion

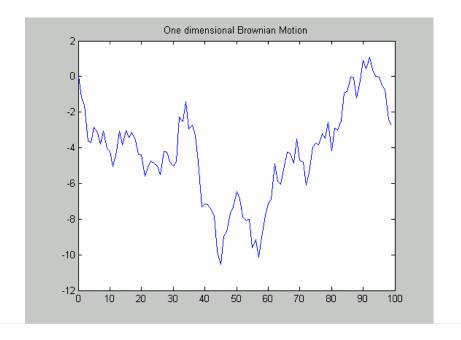
• Independent increments

$$P_{x_0}(B_{t_1} \in dx_1, \dots, B_{t_n} \in dx_n)$$

= $p_{t_1}(x_1 - x_0)p_{t_2 - t_1}(x_2 - x_1) \dots p_{t_n - t_{n-1}}(x_n - x_{n-1})$

• Law of Brownian Motion = Wiener Measure (1923):

$$P_{x_0} = \text{ measure on } C([0, \infty)).$$



• Donsker's Invariance Principle (1952)

$$X_{m{n}}(t) := rac{1}{\sqrt{n}} S_{\lfloor m{n}t
floor}$$

$$X_n(t) \Longrightarrow B(t)$$

B(t) = Brownian motion

• Self-similarity

$$B(ct) = \sqrt{c}B(t)$$

• $t \to B_t$ Almost surely non-differentiable, positive quadratic variation.

Brownian motion is an incredibly rich mathematical object and serves as a key building block of stochastic analysis. Mark Yor:

"I was extremely astonished at the number of very natural questions which have escaped attention until very recently."

One of the questions he raises is "To understand better the ubiquity of Brownian motion in a great number of probabilistic problems"

> SECOND CIRCLE: Dynamics and Conditional Independence

Markov property Conditional independence of past and future

• Markov Semigroup - Brownian Motion

$$T_t f(x) = E_x(f(B_t)) = \int f(y) P_t(x - y) dy$$
 $T_{t+s} = T_t T_s$
 $u(t,x) := T_t f(x)$
 $\frac{\partial u(t,x)}{\partial t} = \frac{1}{2} \Delta u(t,x)$

The Impact of Ito Calculus

$$dF(B(t)) = F'(B(t))dt + \frac{1}{2}F''(B(t))dB(t)$$

Brownian Motion as building block: Stochastic Differential Equations.

Geometric Brownian Motion:

(multiplicative random effects, multiplicative CLT)

$$dx(t) = \mu x(t) dt + \sigma x(t) dw(t)$$
 $x(t) = e^{(\mu - \frac{\sigma^2}{2})t + \sigma w(t)}$

- Black-Scholes formula for option pricing
- σ = volatility, volatility modelling

Brownian Motion with Killing - Feynman Kac change of probability measure

$$rac{\partial u}{\partial t} = rac{1}{2}\Delta u - Vu$$
 $u(0) = f \geq 0$

$$T_t f(x) = E_x \left[\exp(-\int_0^t V(x_s) ds) f(x_t)
ight]$$

Mutation-Selection, nonlinear Genetic algorithm:

$$T_t f(x) = rac{1}{Z} E_x \left[\exp(-\int_0^t V(x_s) ds) f(x_t)
ight]$$
 $Z = ext{ normalizing constant}$

$$rac{\partial u}{\partial t} = rac{1}{2}\Delta u - Vu + u\int V(x)u(x)dx$$

Gradient motion in a potential V

$$egin{aligned} rac{\partial u}{\partial t} &= rac{\sigma^2}{2} \Delta u -
abla u.
abla V \ dx(t) &= -
abla V(x(t)) dt + \sigma dw(t) \end{aligned}$$
 Ito SDE

<u>Gibbs Distribution</u>: $\lim_{t\to\infty} T_t f(x) = \frac{1}{Z} \int f(x) \exp(-\frac{1}{\sigma^2} V(x)) dx$

Example: Ornstein-Uhlenbeck

$$V(x)=rac{1}{2}\gamma x^2 \ dx(t)=-\gamma x(t)dt+\sigma dw(t)$$

Method of simulated annealing.

The Mystery of Dimension

• Two dimensional Brownian motion



- Low dimensions
- Intermediate dimensions
- High dimensions

Dynamics of Eigenvalues and Wigner's Semicircle Law

Real symmetric $n \times n$ random matrices: M(t) - the upper triangle entries are independent OU.

Eigenvalues $\lambda_j(t)$:

$$d\lambda_j = \left(rac{1}{n}\sum_{i
eq j}rac{dt}{\lambda_j-\lambda_i}-\lambda_j
ight)dt+\sqrt{rac{1}{n}}dw_j$$

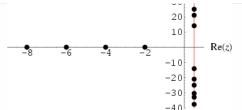
Semicircle Law: Empirical distribution of eigenvalues converges to

$$\lim_{n\to\infty}\frac{1}{n}N_n(\lambda)=\frac{2}{\pi}\int_{-1}^{\lambda}\left(1-u^2\right)^{\frac{1}{2}}du,\ \ |\lambda|\leq 1$$

Universality: $n \times n$ real symmetric matrices with $E(\xi_{ij}^2) = \frac{1}{4}, \ i < j, \ E(\xi_{ii}^2) \leq const + \text{subgaussian decay at infinity}$

From Solid State Physics to the Riemann Zeta Function

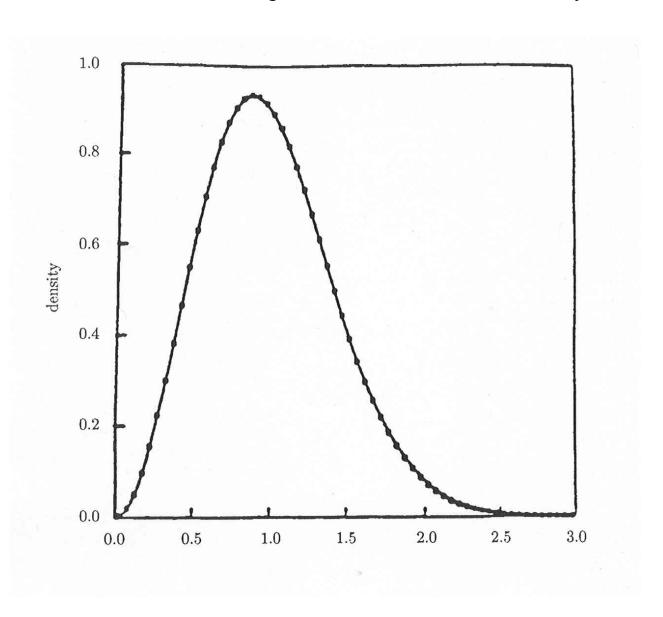
Gaussian unitary ensemble:



Real and complex off-diagonal entries are independent Gaussian.

- "GUE-Hypothesis": (Dyson-Montgomery) the spacings of N successive zeros of the Riemann zeta function and eigenvalues of $N \times N$ Hermitian matrices have the same statistical properties in the $N \to \infty$ limit.
- Persi Diaconis: "Statistical test of hypothesis" Andrew Odlyzko's computation of the 10²⁰th zero of the Riemann zeta function and 175 million of it neighbours.
- Universality Rudnick and Sarnak: universality of the behavior of correlations between successive zeros for a class of L-functions.

Nearest neighbor spacing among 70 million zeros beyond the 10²⁰th zero of the zeta function compared to the Gaussian Unitary Ensemble



THIRD CIRCLE:

Probabilistic Modelling of Interacting Systems

Gibbs Random Fields and Ising Model

Configurations on cube of side K, C_K in \mathbb{Z}^d . $E = \{\pm 1\}^{C_K}$.

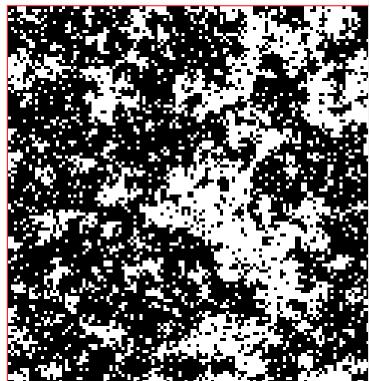
Energy

$$H(\underline{x}) = -\sum_{|i-i|=1} J_{ij} x_i x_j$$

Gibbs Distribution

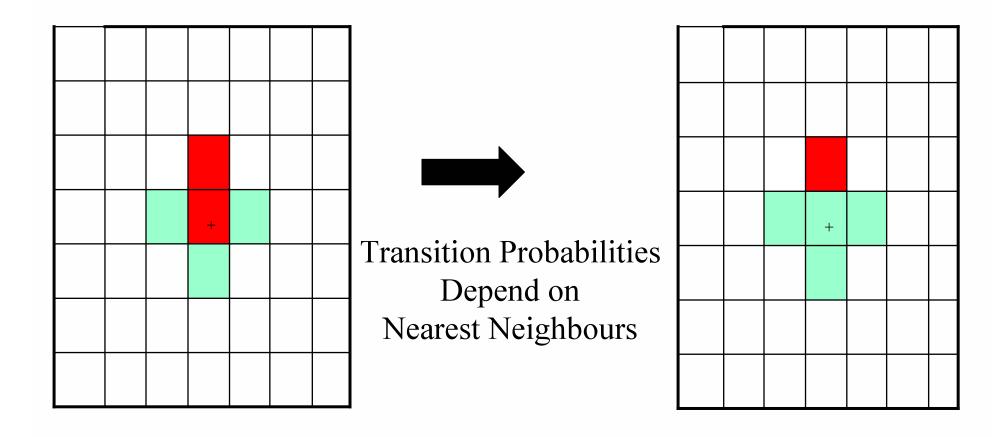
$$P(\underline{x}) = \frac{1}{Z}e^{-\beta H(\underline{x})}$$

 $\beta = \text{inverse temperature}$



Spitzer-Dobrushin

- Markov Random Field (conditional independence)
- Reversible equilibrium, detailed balance.
- Glauber Markovian dynamics: spin-flip



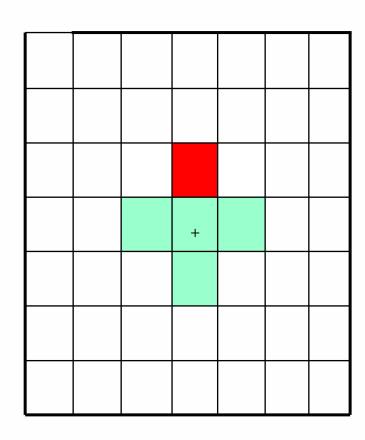
From Statistical Physics to Statistics:

- Phase Transitions, Critical temperature.
- Critical fluctuations d < 4 vs. d > 4.
- Spin Glass Models, disordered medium: random $J_{i,j}$, frustration.
- Markov chain Monte Carlo in statistics and optimization
 - Find a Markov chain whose stationary distribution coincides with the desired distribution
 - * the Gibbs sampler inspired by Glauber dynamics
 - MCMC has many important applications, for example, in Bayesian networks - used e.g. in machine learning.

> FOURTH CIRCLE:

Spatially distributed stochastic dynamics -lattice system on \mathbb{Z}^d

- voter model
- branching random walk
- contact process
- stepping stone model
- random catalytic media
- multiagent systems
- internet dynamics
- spatial economies
- urban growth models.



• Interacting systems

$$\{x_i(t)\}_{i\in\mathbb{Z}^d}$$

- Migration between sites via random walk
- Stochastic dynamics with interaction

Critical Binary Branching RW

$$X_t = \sum_{i=1}^{N(t)} \delta_{x_i(t)},$$

Rate 1:
$$\delta_x \to 0$$
, prob. $\frac{1}{2}$

Rate 1:
$$\delta_x \to \delta_x + \delta_x$$
, prob. $\frac{1}{2}$

Large scale viewpoint

Space-time-mass rescaling

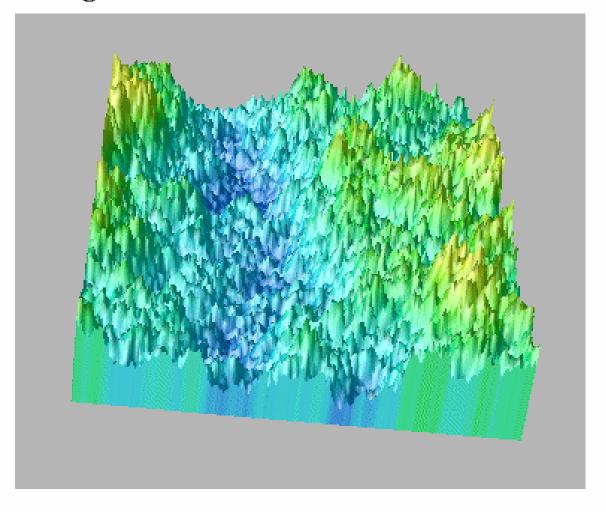
$$X_t^{arepsilon} = arepsilon \sum_{i=1}^{N(t/arepsilon)} \delta_{\sqrt{arepsilon}x_i(rac{t}{arepsilon})},$$

In one dimension $X_t^{\varepsilon} \Longrightarrow X_t(x)dx$ where $X_t(x)$ is solution of SPDE

$$dX_t(x) = \frac{1}{2}\Delta X_t(x) + \sqrt{X_t(x)}W(dt,dx)$$

W(dt, dx) = "Space-time white noise"

Wright-Fisher SPDE



Space-time Evolution

$$dX_t(x) = \frac{1}{2}\Delta X_t(x) + \sqrt{X_t(x)(1-X_t(x))}W(dt,dx)$$

Super Brownian Motion

Dimensions $d \geq 2$. $X_t^{\varepsilon} \Longrightarrow X_t(dx)$ measure-valued process

Laplace Functional:

$$E(e^{-\int f(x)X_t(dx)}) = e^{-\int u(t,x)X_0(dx)}$$

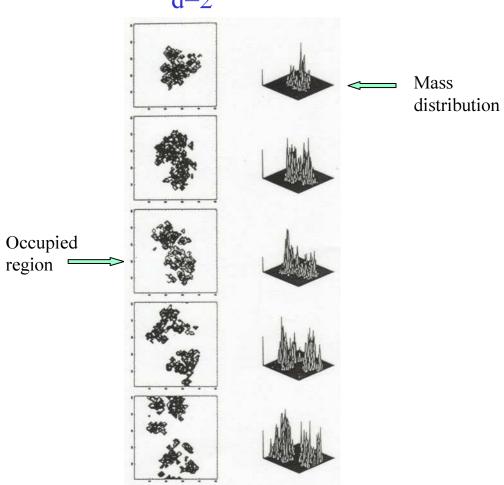
Log-Laplace Equation

$$egin{aligned} rac{\partial u}{\partial t} &= rac{1}{2}\Delta u - u^2 \ u(0) &= f \end{aligned}$$

Integrated Super Excursion:

$$\mathcal{I}:=\int_0^\infty X_t dt, \; X_0=arepsilon \delta_0$$

Super-Brownian Motion in d=2



 $X_t = \text{singular measure!}$

FIFTH CIRCLE: Hierarchy, Genealogy and History Finite Population Sampling and Mutation -Wright-Fisher

- Generation of new generation by finite population sampling
- Current population of size N with K alleles (types) with frequencies p_1, \ldots, p_K
- next generation

$$P(n_1,\ldots,n_K) = rac{N!}{n_1!\ldots n_K!} p_1^{n_1}\ldots p_K^{n_K}$$

- eventual fixation of one type
 - Creation of new types by mutation

The Infinitely Many Alleles Model

- Types labelled by points in [0,1]
- Finite population sampling
- Mutations occur at constant rate and give rise to new type Large population limit X_t is measure-valued diffusion process with state space $M_1([0,1])$

Then $(X_{\text{equ}}([0, \frac{1}{K})), \dots, X_{\text{equ}}([\frac{K-1}{K}, 1]))$ has the Dirichlet distribution

$$f(p_1,\ldots,p_K) = \Gamma(heta) \prod_{i=1}^K rac{p_1^{rac{ heta}{K}-1}}{\Gamma\left(rac{ heta}{K}
ight)} \ldots rac{p_K^{rac{ heta}{K}-1}}{\Gamma\left(rac{ heta}{K}
ight)}$$

on the simplex $\{p_i \geq 0, \sum p_i = 1\}$.

Poisson-Dirichlet

- Order the types in non-increasing frequency
- Let $K \to \infty$.
- In the limit the vector ordered probabilities $(p_1, p_2, p_3, ...)$ has the Poisson Dirichlet $\mathcal{PD}(\theta)$ distribution.

From Genetics to Number theory

Let N(n) be a randomly chosen integer from 1 to n.

Let $\beta_1(n), \beta_2(n), \beta_3(n), \ldots$ be the multiplicities of the prime divisors p_1, p_2, \ldots of N(n) in non-increasing order

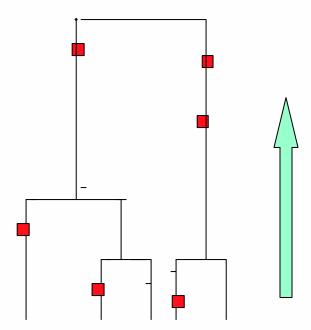
$$lpha_i(n) = rac{eta_i(n) \log p_i}{\log N(n)}$$

$$\{\alpha_1(n), \alpha_2(n), \dots\} \Longrightarrow_{n \to \infty} \text{Poisson-Dirichlet}(1)$$

The Kingman Coalescent, genealogy

Look at common ancestors of individuals in the population:

- Pure death process
$$\mu_k = \begin{pmatrix} k \\ 2 \end{pmatrix}, \quad k = n, \dots, 2$$



Up to last common ancestor

Marks = mutations

Infinitely many sites model

State space $M_1([0,1])^{\mathbb{N}}$.

[0,1] denotes the sites in a DNA string and $\mathbf{x}=(x_1,x_2,\ldots,0,0,\ldots)$ where x_1,x_2,\ldots denote the sites at which mutations have occurred

- $-x_1$ is site of latest mutation,
- $-x_2$ is site of second most recent mutation, ...

This is the "Historical" version of the infinitely many types model and is used to get the distribution of the number of pairwise comparisons and estimate mutation rate.

• Most inference on genetic data involves combination of probability models and computational tools such as MCMC.

Infinitely many types Stepping Stone Model

- Set of sites \mathbb{Z}^d
- Set of types [0,1]
- Population at each site: $M_1[0,1] = \text{distribution of infinitely}$ many types
- Finite population sampling at each site no mutation
- Migration via random walk on \mathbb{Z}^d

Qualitative Behavior

- -d=1,2 Local population becomes unitype loss of diversity
- $-d \ge 3$ Global equilibrium develops with multiple types present locally
- -d=3,4 immortal families (colours) recurrent
- $-d \geq 5$ immortal families (colours) transient

Homozygosity in large space-time scales. Infinitely many types stepping stone model in \mathbb{Z}^d , $d \geq 3$ with $X_0 = \nu$, ν nonatomic. Let $\Delta := \{(x, x) \in [0, 1] \times [0, 1]\}$. Then

$$lim_{L o\infty}(1/(L^d))\sum_{|j|\leq L} < X_{eq}(j) imes X_{eq}(0), I_{\Delta}> = 0$$

Moreover

$$\int_0^\infty < X_{eq}(j) imes X_{eq}(0), I_\Delta > dt < \infty$$

if and only if $d \geq 5$.

Role of Dimension

- Dimensions d = 1, 2. Random walk is recurrent.
- Dimensions d = 3, 4 Random walk is transient but not strongly transient.
- Dimensions $d \geq 5$ Random walk is strongly transient, that is,

$$E(L_B) < \infty$$

where L_B is the last exit time from a bounded set.

SIXTH CIRCLE: Universality in spatial and space-time structures

- \bullet SBM
- \bullet ISE
- \bullet SLE

The Voter Model in $d \geq 2$

$$\xi_t(x) \in \{0,1\}, \ x \in Z^d$$

Rate $p(x-y): \xi_t(x) \to \xi_t(y)$

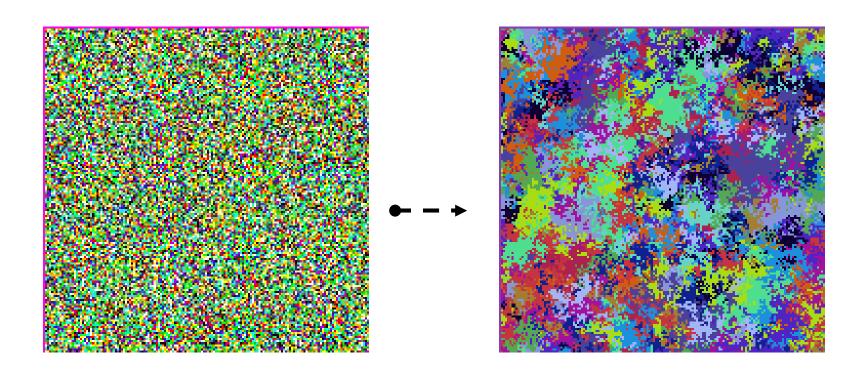
Rescaled Voter Model on \mathbb{Z}^d/\sqrt{N} .

$$\xi_t^N(x) = \xi_{Nt}(x\sqrt{N})$$

$$X_t^N = \frac{1}{N'} \sum_{x \in \frac{Z^d}{\sqrt{N}}} \xi_t^N(x) \delta_x$$

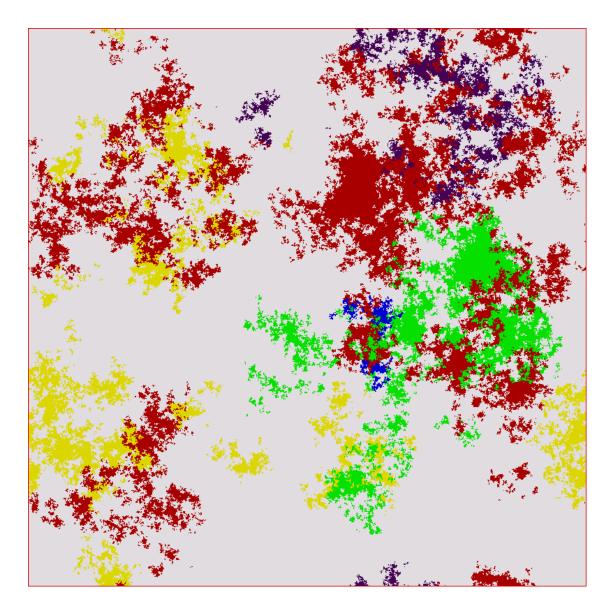
Assume $X_0^N \Longrightarrow X_0 \in M_F(R^d)$ as $N \to \infty$. Then $X_t^N \Longrightarrow X_t$ (SBM) (Cox, Durrett and Perkins, Bramson, Cox and Le Gall.) $(N'=N,\ d\geq 3,\ N'=N/\log N \ \text{if}\ d=2)$

Infinitely many types Voter Model



Initial Configuration

Intermediate time

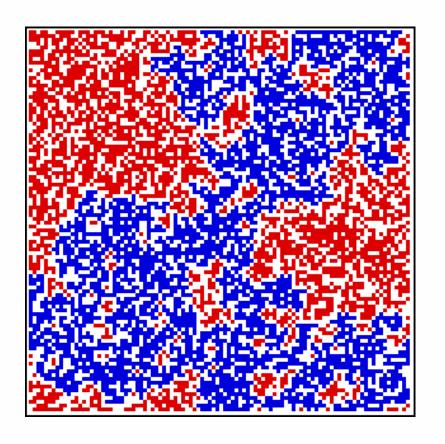


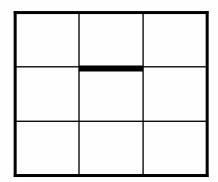
Multitype Voter Clusters - SBM (Palm cluster)

Critical Bond Percolation on \mathbb{Z}^d

p = probability bond closed

 p_c = critical probability for \exists infinite cluster





- High Dimensions. No infinite cluster when $p = p_c$
- Conjecture (Hara and Slade): "Infinite incipient cluster": $C_N(0) = \text{Cluster of size } N$ In dimensions d > 6 as $N \to \infty$.

$$X_N = rac{1}{N+1} \sum_{k \in C_N(0)} \delta_{rac{k}{N^{1/4}}}$$
 $X_N \Longrightarrow_{N o \infty} \mathcal{I} = ISE$
 $\mathcal{I} = \int_0^\infty X_s ds$, conditioned to have total mass one $X_s = ext{SBM}$ excursion from 0 .

 \mathcal{I} is supported on a random subset of dimension 4 in \mathbb{R}^d . (Hara and Slade) In high dimensions two and three point functions (moment measures) converge to those of ISE.

Lattice Trees d > 8

Rescaled random tree T_N with N bonds. (Aldous; Derbez, Slade)

$$X_N = \frac{1}{N+1} \sum_{k \in \mathbb{Z}^d} \delta_{\frac{k}{N^{1/4}}}$$

$$X_N \underset{N \to \infty}{\Longrightarrow} \mathcal{I}$$

(large d or d > 8 with spread-out trees).

Two dimensional phenomena.

- Scaling limit of critical site percolation on two dimensional triangular lattice "boundary path" is equal to chordal SLE₆.
 (Lawler, Schramm, Smirnov and Werner)
- Conformal invariance.

Stochastic Loewner evolution equation

 $SLE(\kappa)$: Let B(t) be one dimensional BM and

$$W(t) = B(\kappa t)$$
$$B(0) = 0$$

$$\frac{\partial}{\partial t}g_t(z) = \frac{2}{g_t(z) - W(t)}$$
 $g_0(z) = z, \ g_t(z) \text{ is a curve in upper half plane}$

Universality Phenomena

- * Why do these different interacting systems have a common scaling limit?
- * How we identify and classify the possible large space-time scale behaviors?
 - · Other universality classes: Fisher-Wright, Fleming-Viot, Mutually catalytic branching
 - · HMF analysis: fixed point of NL integral equation

FUTURE CHALLENGES

- Probabilistic modelling of complex systems arising across the biological and social sciences.
- Integrative approach to the analysis of empirical data and theoretical and simulation studies of mathematical models.
- Identification and classification of possible large space-time scale behaviors?

REFERENCES:

- Current and Emerging Research Opportunities in Probability, at www.math.cornell.edu/durrett/probrep/probrep.html or http://oregonstate.edu/dept/math/research/probrep.html.
- T. Cox, R. Durrett and E. Perkins (1998). Rescaled voter models converge to super-Brownian motion, Annals of Probability.
- D.A. Dawson and E.A. Perkins (1999). Measure-valued processes and Renormalization of Branching Particle Systems, in AMS Surveys and Monographs Vol. 64, pp 45-106.
- D.A. Dawson (2000). Stochastic models of evolving information systems, in *Stochastic Models*, eds. L.G. Gorostiza dn B.G. Ivanoff, CMS Conference Proceedings 26, pp. 1-14.
- P. Diaconis and M. Coram. New tests of the correspondence between unitary eigenvalues and zeros of Riemann's zeta function, preprint.
- R. Durrett (1988) Lecture Notes on Particle Systems and Percolation, Wadsworth Pub.

- A.Greven (2000). Renormalization analysis of interacting spatial stochastic systems, in Infinite Dimensional Stochastic Analysis, Ph. Clément, F. den Hollander, J. van Neerven and B. de Pagter, Amsterdam, 197-216.
- T.P. Hill (1995). Base invariance implies Benford's law, Proc. AMS 12, 887-895.
- M. Kac (1985). Enigmas of Chance, Harper and Row.
- A. Klenke. Computer simulation of stochastic processes, http://www.mi.uni-koeln.de/ klenke/.
- L.C.G. Rogers and Z. Shi. Interacting Brownian particles and the Wigner law, PTRF (95), 1993, 555-570.
- G. Slade (2002). Scaling limits and super-Brownian motion, Notices AMS, 49, 1056-1067.
- G.W. Winkler (1995). Image Analysis, Random Fields and Dynamic Monte Carlo Methods, Springer.
- M. Yor (1997) Some Aspects of Brownian Motion, I,II, Birkhauser.