The shape of multidimensional Brunet–Derrida particle systems
NATHANAËL BERESTYCKI AND LEE ZHUO ZHAO 651

On directional derivatives of Skorokhod maps in convex polyhedral domains
DAVID LIPSHTZ and KAVITA RAMANAN 688

Spatial Gibbs random graphs . . . JEAN-CHRISTOPHE MOURRAT AND DANIEL VALESIN 751

On the stability and the uniform propagation of chaos properties of Ensemble Kalman–Bucy filters . . . . . . . . . . . . . . . . . . P. DEL MORAL AND J. TUGAUT 790

A large scale analysis of unreliable stochastic networks
REZA AGHAJANI, PHILIPPE ROBERT AND WEN SUN 851

Reflected backward stochastic differential equations with resistance
ZHONGMIN QIAN AND MINGYU XU 888

Zooming in on a Lévy process at its supremum . . . . . . . . . . . JEVGENIJS IVANOVS 912

Spectral gap of random hyperbolic graphs and related parameters
MARCOS KIWI AND DIETER MITSCHE 941

A phase transition regarding the evolution of bootstrap processes in inhomogeneous random graphs . . NIKOLAOS FOUNTOULAKIS, MIHYUN KANG, CHRISTOPH KOCH AND TAMÁS MAKAI 990

Sharp thresholds for contagious sets in random graphs
OMER ANGEL AND BRETT KOLESNIK 1052

The sample size required in importance sampling
SOURAV CHATTERJEE AND PERSI DIACONIS 1099

Uniqueness and propagation of chaos for the Boltzmann equation with moderately soft potentials . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . LIPING XU 1136

A random matrix approach to neural networks
COSME LOUART, ZHENYU LIAO AND ROMAIN COUILLET 1190

Phase transitions in the one-dimensional Coulomb gas ensembles . . TATYANA S. TUROVA 1249

Random cluster dynamics for the Ising model is rapidly mixing
HENG GUO AND MARK JERRUM 1292
The purpose of the Institute is to foster the development and dissemination of the theory and applications of statistics and probability.
THE SHAPE OF MULTIDIMENSIONAL BRUNET–DERRIDA
PARTICLE SYSTEMS

BY NATHANAËL BERESTYCKI¹ AND LEE ZHOU ZHAO

University of Cambridge

We introduce particle systems in one or more dimensions in which particles perform branching Brownian motion and the population size is kept constant equal to $N > 1$, through the following selection mechanism: at all times only the $N$ fittest particles survive, while all the other particles are removed. Fitness is measured with respect to some given score function $s : \mathbb{R}^d \to \mathbb{R}$. For some choices of the function $s$, it is proved that the cloud of particles travels at positive speed in some possibly random direction. In the case where $s$ is linear, we show under some mild assumptions that the shape of the cloud scales like $\log N$ in the direction parallel to motion but at least $(\log N)^{3/2}$ in the orthogonal direction. We conjecture that the exponent $3/2$ is sharp. In order to prove this, we obtain the following result of independent interest: in one-dimensional systems, the genealogical time is greater than $c(\log N)^3$. We discuss several open problems and explain how our results can be viewed as a rigorous justification in our setting of empirical observations made by Burt [Evolution 54 (2000) 337–351] in support of Weismann’s arguments for the role of recombination in population genetics.

REFERENCES


MSC2010 subject classifications. 60K35, 92B05.

Key words and phrases. Brunet–Derrida particle systems, branching Brownian motion, random travelling wave, recombination.


ON DIRECTIONAL DERIVATIVES OF SKOROKHOD MAPS IN CONVEX POLYHEDRAL DOMAINS

BY DAVID LIPSHUTZ\(^1\) AND KAVITA RAMANAN\(^2\)

Brown University

The study of both sensitivity analysis and differentiability of the stochastic flow of a reflected process in a convex polyhedral domain is challenging due to the abrupt change in the nature of the dynamics at the boundary and is further complicated because the boundary is not smooth. These difficulties can be addressed by studying directional derivatives of an associated extended Skorokhod map, which is a deterministic mapping that takes an unconstrained path to a suitably reflected or constrained version. In this work, we develop an axiomatic framework for the analysis of directional derivatives of a large class of Lipschitz continuous extended Skorokhod maps in convex polyhedral domains with oblique directions of reflection. We establish existence of directional derivatives at a path whose reflected version satisfies a certain boundary jitter property, and also show that the right-continuous regularization of such a directional derivative can be characterized as the unique solution to a Skorokhod-type problem, where both the domain and directions of reflection vary (discontinuously) depending on the state of the reflected path. A key step in the analysis is the proof of certain contraction properties for a family of (oblique) derivative projection operators. The results of this paper are used in subsequent work to study differentiability of stochastic flows and sensitivity analysis for a large class of reflected diffusions in convex polyhedral domains.

REFERENCES


**MSC2010 subject classifications.** Primary 90C31, 93B35; secondary 60G17, 90B15.

**Key words and phrases.** Extended Skorokhod problem, directional derivative of the Skorokhod map, derivative problem, reflected process, sensitivity analysis, stochastic flow, oblique reflection, boundary jitter property.


SPATIAL GIBBS RANDOM GRAPHS

BY JEAN-CHRISTOPHE MOURRAT AND DANIEL VALESIN

Ecole Normale Supérieure de Lyon and University of Groningen

Many real-world networks of interest are embedded in physical space. We present a new random graph model aiming to reflect the interplay between the geometries of the graph and of the underlying space. The model favors configurations with small average graph distance between vertices, but adding an edge comes at a cost measured according to the geometry of the ambient physical space. In most cases, we identify the order of magnitude of the average graph distance as a function of the parameters of the model. As the proofs reveal, hierarchical structures naturally emerge from our simple modeling assumptions. Moreover, a critical regime exhibits an infinite number of discontinuous phase transitions.

REFERENCES


MSC2010 subject classifications. 82C22, 05C80.

Key words and phrases. Spatial random graph, Gibbs measure, phase transition.


ON THE STABILITY AND THE UNIFORM PROPAGATION OF CHAOS PROPERTIES OF ENSEMBLE KALMAN–BUCY FILTERS

BY P. DEL MORAL AND J. TUGAUT

Center INRIA Bordeaux Sud-Ouest & Institut de Mathématiques de Bordeaux and Université de Lyon, Université Jean Monnet

The ensemble Kalman filter is a sophisticated and powerful data assimilation method for filtering high dimensional problems arising in fluid mechanics and geophysical sciences. This Monte Carlo method can be interpreted as a mean-field McKean–Vlasov-type particle interpretation of the Kalman–Bucy diffusions. In contrast to more conventional particle filters and nonlinear Markov processes, these models are designed in terms of a diffusion process with a diffusion matrix that depends on particle covariance matrices.

Besides some recent advances on the stability of nonlinear Langevin-type diffusions with drift interactions, the long-time behaviour of models with interacting diffusion matrices and conditional distribution interaction functions has never been discussed in the literature. One of the main contributions of the article is to initiate the study of this new class of models. The article presents a series of new functional inequalities to quantify the stability of these nonlinear diffusion processes.

In the same vein, despite some recent contributions on the convergence of the ensemble Kalman filter when the number of sample tends to infinity very little is known on stability and the long-time behaviour of these mean-field interacting type particle filters. The second contribution of this article is to provide uniform propagation of chaos properties as well as $L_2$-mean error estimates w.r.t. to the time horizon. Our regularity condition is also shown to be sufficient and necessary for the uniform convergence of the ensemble Kalman filter.

The stochastic analysis developed in this article is based on an original combination of functional inequalities and Foster–Lyapunov techniques with coupling, martingale techniques, random matrices and spectral analysis theory.

REFERENCES


MSC2010 subject classifications. 60J60, 60J22, 35Q84, 93E11, 60M20, 60G25.

Key words and phrases. Ensemble Kalman filter, Kalman–Bucy filter, Riccati equations, ill-conditioned systems, mean-field particle models, sequential Monte Carlo methods, interacting particle systems, random covariance matrices, nonlinear Markov processes.


A LARGE SCALE ANALYSIS OF UNRELIABLE STOCHASTIC NETWORKS

BY REZA AGHAJANI∗, PHILIPPE ROBERT† AND WEN SUN†,1

University of California, San Diego* and INRIA Paris†

The problem of reliability of a large distributed system is analyzed via a new mathematical model. A typical framework is a system where a set of files are duplicated on several data servers. When one of these servers breaks down, all copies of files stored on it are lost. In this way, repeated failures may lead to losses of files. The efficiency of such a network is directly related to the performances of the mechanism used to duplicate files on servers. In this paper, we study the evolution of the network using a natural duplication policy giving priority to the files with the least number of copies.

We investigate the asymptotic behavior of the network when the number \(N\) of servers is large. The analysis is complicated by the large dimension of the state space of the empirical distribution of the state of the network. A stochastic model of the evolution of the network which has values in state space whose dimension does not depend on \(N\) is introduced. Despite this description does not have the Markov property, it turns out that it is converging in distribution, when the number of nodes goes to infinity, to a nonlinear Markov process. The rate of decay of the network, which is the key characteristic of interest of these systems, can be expressed in terms of this asymptotic process. The corresponding mean-field convergence results are established. A lower bound on the exponential decay, with respect to time, of the fraction of the number of initial files with at least one copy is obtained.

REFERENCES


MSC2010 subject classifications. Primary 60J27, 60K25; secondary 68M15.
Key words and phrases. Stochastic networks with failures, mean-field models, nonlinear Markov processes, reliability.


REFLECTED BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS WITH RESISTANCE

BY ZHONGMIN QIAN¹ AND MINGYU XU²

University of Oxford and Chinese Academy of Science

In this article, we study a class of reflected backward stochastic differential equations (introduced in El Karoui et al. [Ann. Probab. 25 (1997) 702–737], RBSDE for short) with nonlinear resistance by means of Skorohod’s equation. The advantage of this approach lies in its pathwise nature and, therefore, provides additional information about solutions of RBSDE. As an application of our approach, we will consider reflected backward problems with resistance as well. This class of RBSDEs possess significance in the super-hedging with wealth constraint.

REFERENCES


MSC2010 subject classifications. Primary 60H10, 60J45.

Key words and phrases. Brownian motion, local time, optional dual projection, reflected BSDE, Skorohod’s equation.


ZOOMING IN ON A LÉVY PROCESS AT ITS SUPREMUM

BY JEVGENIJS IVANOVS

Aarhus University

Let $M$ and $\tau$ be the supremum and its time of a Lévy process $X$ on some finite time interval. It is shown that zooming in on $X$ at its supremum, that is, considering $((X_{\tau+\varepsilon} - M)/a_\varepsilon)_{\varepsilon \in \mathbb{R}}$ as $\varepsilon \downarrow 0$, results in $(\xi_t)_{t \in \mathbb{R}}$ constructed from two independent processes having the laws of some self-similar Lévy process $\hat{X}$ conditioned to stay positive and negative. This holds when $X$ is in the domain of attraction of $\hat{X}$ under the zooming-in procedure as opposed to the classical zooming out $[\text{Trans. Amer. Math. Soc.} \ 104 \ (1962) \ 62–78]$. As an application of this result, we establish a limit theorem for the discretization errors in simulation of supremum and its time, which extends the result in $[\text{Ann. Appl. Probab.} \ 5 \ (1995) \ 875–896]$ for a linear Brownian motion. Additionally, complete characterization of the domains of attraction when zooming in on a Lévy process is provided.

REFERENCES


MSC2010 subject classifications. Primary 60G51, 60F17; secondary 60G18, 60G52.

Key words and phrases. Conditioned to stay positive, discretization error, domains of attraction, Euler scheme, functional limit theorem, high frequency statistics, invariance principle, mixing convergence, scaling limits, self-similarity, small-time behaviour.


SPECTRAL GAP OF RANDOM HYPERBOLIC GRAPHS AND RELATED PARAMETERS

BY MARCOS KIWI¹ AND DIETER MITSCHE

University of Chile and Université de Nice Sophia-Antipolis

Random hyperbolic graphs have been suggested as a promising model of social networks. A few of their fundamental parameters have been studied. However, none of them concerns their spectra. We consider the random hyperbolic graph model, as formalized by [Automata, Languages, and Programming—39th International Colloquium—ICALP Part II. (2012) 573–585 Springer], and essentially determine the spectral gap of their normalized Laplacian. Specifically, we establish that with high probability the second smallest eigenvalue of the normalized Laplacian of the giant component of an n-vertex random hyperbolic graph is at least \( \Omega(n^{-(2\alpha-1)/D}) \), where \( \frac{1}{2} < \alpha < 1 \) is a model parameter and \( D \) is the network diameter (which is known to be at most polylogarithmic in \( n \)). We also show a matching (up to a polylogarithmic factor) upper bound of \( n^{-(2\alpha-1)\log n^{1-o(1)}} \).

As a byproduct, we conclude that the conductance upper bound on the eigenvalue gap obtained via Cheeger’s inequality is essentially tight. We also provide a more detailed picture of the collection of vertices on which the bound on the conductance is attained, in particular showing that for all subsets whose volume is \( O(n^{\varepsilon}) \) for \( 0 < \varepsilon < 1 \) the obtained conductance is with high probability \( \Omega(n^{-(2\alpha-1)\varepsilon+o(1)}) \). Finally, we also show consequences of our result for the minimum and maximum bisection of the giant component.

REFERENCES


MSC2010 subject classifications. Primary 60C05; secondary 05C80.

Key words and phrases. Random hyperbolic graphs, spectral gap, conductance.


A PHASE TRANSITION REGARDING THE EVOLUTION OF BOOTSTRAP PROCESSES IN INHOMOGENEOUS RANDOM GRAPHS

BY NIKOLAOS FOUNTOULAKIS*, 1, MIHYUN KANG†, 2, CHRISTOPH KOCH‡, 3 AND TAMÁS MAKAI†, 4

University of Birmingham*, Graz University of Technology† and University of Warwick‡

A bootstrap percolation process on a graph with infection threshold $r \geq 1$ is a dissemination process that evolves in time steps. The process begins with a subset of infected vertices and in each subsequent step every uninfected vertex that has at least $r$ infected neighbours becomes infected and remains so forever.

Critical phenomena in bootstrap percolation processes were originally observed by Aizenman and Lebowitz in the late 1980s as finite-volume phase transitions in $\mathbb{Z}^d$ that are caused by the accumulation of small local islands of infected vertices. They were also observed in the case of dense (homogeneous) random graphs by Janson et al. [Ann. Appl. Probab. 22 (2012) 1989–2047]. In this paper, we consider the class of inhomogeneous random graphs known as the Chung-Lu model: each vertex is equipped with a positive weight and each pair of vertices appears as an edge with probability proportional to the product of the weights. In particular, we focus on the sparse regime, where the number of edges is proportional to the number of vertices.

The main results of this paper determine those weight sequences for which a critical phenomenon occurs: there is a critical density of vertices that are infected at the beginning of the process, above which a small (sublinear) set of infected vertices creates an avalanche of infections that in turn leads to an outbreak. We show that this occurs essentially only when the tail of the weight distribution dominates a power law with exponent 3 and we determine the critical density in this case.

REFERENCES


MSC2010 subject classifications. Primary 05C80, 60K37; secondary 82B26, 82B27.

Key words and phrases. Bootstrap percolation, inhomogeneous random graphs, critical phenomena.


SHARP THRESHOLDS FOR CONTAGIOUS SETS IN RANDOM GRAPHS

BY OMER ANGEL¹ AND BRETT KOLESNIK²

University of British Columbia

For fixed $r \geq 2$, we consider bootstrap percolation with threshold $r$ on the Erdős–Rényi graph $G_{n,p}$. We identify a threshold for $p$ above which there is with high probability a set of size $r$ that can infect the entire graph. This improves a result of Feige, Krivelevich and Reichman, which gives bounds for this threshold, up to multiplicative constants.

As an application of our results, we obtain an upper bound for the threshold for $K_4$-percolation on $G_{n,p}$, as studied by Balogh, Bollobás and Morris. This bound is shown to be asymptotically sharp in subsequent work.

These thresholds are closely related to the survival probabilities of certain time-varying branching processes, and we derive asymptotic formulae for these survival probabilities, which are of interest in their own right.

REFERENCES


MSC2010 subject classifications. Primary 60K35; secondary 05C80, 60C05, 82B43.

Key words and phrases. Bootstrap percolation, cellular automaton, phase transition, random graph, sharp threshold.


THE SAMPLE SIZE REQUIRED IN IMPORTANCE SAMPLING

BY SOURAV CHATTERJEE\textsuperscript{1} AND PERSI DIACONIS\textsuperscript{2}

Stanford University

The goal of importance sampling is to estimate the expected value of a given function with respect to a probability measure $\nu$ using a random sample of size $n$ drawn from a different probability measure $\mu$. If the two measures $\mu$ and $\nu$ are nearly singular with respect to each other, which is often the case in practice, the sample size required for accurate estimation is large. In this article, it is shown that in a fairly general setting, a sample of size approximately $\exp(D(\nu \parallel \mu))$ is necessary and sufficient for accurate estimation by importance sampling, where $D(\nu \parallel \mu)$ is the Kullback–Leibler divergence of $\mu$ from $\nu$. In particular, the required sample size exhibits a kind of cut-off in the logarithmic scale. The theory is applied to obtain a general formula for the sample size required in importance sampling for one-parameter exponential families (Gibbs measures).

REFERENCES


MSC2010 subject classifications. 65C05, 65C60, 60F05, 82B80.

Key words and phrases. Importance sampling, Monte Carlo methods, Gibbs measure, phase transition.


UNIQUENESS AND PROPAGATION OF CHAOS FOR THE BOLTZMANN EQUATION WITH MODERATELY SOFT POTENTIALS

BY LIPING XU

Université Pierre et Marie Curie (Paris VI)

We prove a strong/weak stability estimate for the 3D homogeneous Boltzmann equation with moderately soft potentials \([\gamma \in (-1,0)]\) using the Wasserstein distance with quadratic cost. This in particular implies the uniqueness in the class of all weak solutions, assuming only that the initial condition has a finite entropy and a finite moment of sufficiently high order. We also consider the Nanbu \(N\)-stochastic particle system, which approximates the weak solution. We use a probabilistic coupling method and give, under suitable assumptions on the initial condition, a rate of convergence of the empirical measure of the particle system to the solution of the Boltzmann equation for this singular interaction.

REFERENCES


MSC2010 subject classifications. Primary 82C40, 60K35.

Key words and phrases. Kinetic theory, Boltzmann equation, stochastic particle systems, propagation of chaos, Wasserstein distance.


A RANDOM MATRIX APPROACH TO NEURAL NETWORKS

BY COSME LOUART, ZHENYU LIAO AND ROMAIN COUILLET

CentraleSupélec, University of Paris–Saclay

This article studies the Gram random matrix model \( G = \frac{1}{T} \Sigma^T \Sigma \), \( \Sigma = \sigma(WX) \), classically found in the analysis of random feature maps and random neural networks, where \( X = [x_1, \ldots, x_T] \in \mathbb{R}^{p \times T} \) is a (data) matrix of bounded norm, \( W \in \mathbb{R}^{n \times p} \) is a matrix of independent zero-mean unit variance entries and \( \sigma : \mathbb{R} \to \mathbb{R} \) is a Lipschitz continuous (activation) function—\( \sigma(WX) \) being understood entry-wise. By means of a key concentration of measure lemma arising from nonasymptotic random matrix arguments, we prove that, as \( n, p, T \) grow large at the same rate, the resolvent \( Q = (G + \gamma I_T)^{-1} \), for \( \gamma > 0 \), has a similar behavior as that met in sample covariance matrix models, involving notably the moment \( \frac{1}{n} \mathbb{E}[G] \), which provides in passing a deterministic equivalent for the empirical spectral measure of \( G \). Application-wise, this result enables the estimation of the asymptotic performance of single-layer random neural networks. This in turn provides practical insights into the underlying mechanisms into play in random neural networks, entailing several unexpected consequences, as well as a fast practical means to tune the network hyperparameters.

REFERENCES


MSC2010 subject classifications. Primary 60B20; secondary 62M45.
Key words and phrases. Random matrix theory, random feature maps, neural networks.


PHASE TRANSITIONS IN THE ONE-DIMENSIONAL COULOMB GAS ENSEMBLES

By Tatyana S. Turova

University of Lund

We consider the system of particles on a finite interval with pairwise nearest neighbors interaction and external force. This model was introduced by Malyshev [Probl. Inf. Transm. 51 (2015) 31–36] to study the flow of charged particles on a rigorous mathematical level. It is a simplified version of a 3-dimensional classical Coulomb gas model. We study Gibbs distribution at finite positive temperature extending recent results on the zero temperature case (ground states). We derive the asymptotics for the mean and for the variances of the distances between the neighboring charges. We prove that depending on the strength of the external force there are several phase transitions in the local structure of the configuration of the particles in the limit when the number of particles goes to infinity. We identify 5 different phases for any positive temperature.

The proofs rely on a conditional central limit theorem for nonidentical random variables, which has an interest on its own.

REFERENCES


MSC2010 subject classifications. 82B21, 82B26, 60F05.
Key words and phrases. Coulomb gas, phase transitions, Gibbs ensemble.


RANDOM CLUSTER DYNAMICS FOR THE ISING MODEL IS RAPIDLY MIXING

BY HENG GUO\textsuperscript{1,2,*} AND MARK JERRUM\textsuperscript{1,†}

University of Edinburgh* and Queen Mary, University of London†

We show that the mixing time of Glauber (single edge update) dynamics for the random cluster model at $q = 2$ on an arbitrary $n$-vertex graph is bounded by a polynomial in $n$. As a consequence, the Swendsen–Wang algorithm for the ferromagnetic Ising model at any temperature also has a polynomial mixing time bound.

REFERENCES


MSC2010 subject classifications. Primary 68W20; secondary 68Q87.

Key words and phrases. Random cluster, Markov chains, Ising model, Swendsen–Wang dynamics.


BRADLEY EFRON
TREVOR HASTIE

COMPUTER AGE
STATISTICAL
INFERECE

ALGORITHMS, EVIDENCE, AND DATA SCIENCE
Probability on Graphs
Random Processes on Graphs and Lattices

Geoffrey Grimmett

This introduction to some of the principal models in the theory of disordered systems leads the reader through the basics, to the very edge of contemporary research, with the minimum of technical fuss. Topics covered include random walk, percolation, self-avoiding walk, interacting particle systems, uniform spanning tree, random graphs, as well as the Ising, Potts, and random-cluster models for ferromagnetism, and the Lorentz model for motion in a random medium. Schramm–Löwner evolutions (SLE) arise in various contexts. The choice of topics is strongly motivated by modern applications and focuses on areas that merit further research. Special features include a simple account of Smirnov’s proof of Cardy’s formula for critical percolation, and a fairly full account of the theory of influence and sharp-thresholds. Accessible to a wide audience of mathematicians and physicists, this book can be used as a graduate course text. Each chapter ends with a range of exercises.