## CONTENTS (continued)

<table>
<thead>
<tr>
<th>Authors</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARINO, V., NOURDIN, I., NUALART, D. and SALAMAT, M.</td>
<td>Limit theorems for integral functionals of Hermite-driven processes</td>
</tr>
<tr>
<td>KOVCHEGOV, Y. and ZALIAPIN, I.</td>
<td>Invariance and attraction properties of Galton–Watson trees</td>
</tr>
<tr>
<td>BOBBIA, B., DOMBRY, C. and VARRON, D.</td>
<td>The coupling method in extreme value theory</td>
</tr>
<tr>
<td>GNEDIN, A. and SEKSENBAYEV, A.</td>
<td>Asymptotics and renewal approximation in the online selection of increasing subsequence</td>
</tr>
<tr>
<td>JAKUBOWSKI, J. and WISNIEWOLSKI, M.</td>
<td>A convolution formula for the local time of an Itô diffusion reflecting at 0 and a generalized Strroock–Williams equation</td>
</tr>
<tr>
<td>FAN, W.-T.L.</td>
<td>Stochastic PDEs on graphs as scaling limits of discrete interacting systems</td>
</tr>
<tr>
<td>BREHMER, J.R. and GNEITING, T.</td>
<td>Scoring interval forecasts: Equal-tailed, shortest, and modal interval</td>
</tr>
<tr>
<td>CÉNAC, P., CHAUVIN, B., NOÜS, C., PACCAUT, F. and POUYANNE, N.</td>
<td>Variable Length Memory Chains: Characterization of stationary probability measures</td>
</tr>
<tr>
<td>HUANG, L.-J., MAJKA, M.B. and WANG, J.</td>
<td>Approximation of heavy-tailed distributions via stable-driven SDEs</td>
</tr>
<tr>
<td>KUTOYANTS, Y.A.</td>
<td>On multi-step estimation of delay for SDE</td>
</tr>
<tr>
<td>NICOLUSSI, F. and CAZZARO, M.</td>
<td>Context-specific independencies in stratified chain regression graphical models</td>
</tr>
<tr>
<td>KABLUCHKO, Z., PROCHNO, J. and THÄLE, C.</td>
<td>A new look at random projections of the cube and general product measures</td>
</tr>
<tr>
<td>BAHMANI, S.</td>
<td>Nearly optimal robust mean estimation via empirical characteristic function</td>
</tr>
</tbody>
</table>
CONTENTS

KRAMPE, J., KREISS, J.-P. and PAPARODITIS, E. 1441
Bootstrap based inference for sparse high-dimensional time series models

LAW, M. and RITOV, Y. 1467
Inference without compatibility: Using exponential weighting for inference on a parameter of a linear model

KRAAIJ, R.C. 1496
Flux large deviations of weakly interacting jump processes via well-posedness of an associated Hamilton–Jacobi equation

BETZ, V., MÜHLBAUER, J., SCHÄFER, H. and ZEINDLER, D. 1529
Precise asymptotics of longest cycles in random permutations without macroscopic cycles

CATOR, E. and DON, H. 1556
Explicit bounds for critical infection rates and expected extinction times of the contact process on finite random graphs

ABI JABER, E. 1583
Weak existence and uniqueness for affine stochastic Volterra equations with $L^1$-kernels

AGAPIOU, S., DASHTI, M. and HELIN, T. 1616
Rates of contraction of posterior distributions based on $p$-exponential priors

CARDONA-TOBÓN, N. and PALAU, S. 1643
Yaglom’s limit for critical Galton–Watson processes in varying environment: A probabilistic approach

KREISS, A. 1666
Correlation bounds, mixing and $m$-dependence under random time-varying network distances with an application to Cox-processes

KREBS, J. 1695
On the law of the iterated logarithm and strong invariance principles in stochastic geometry

HUANG, D. and TROPP, J.A. 1724
From Poincaré inequalities to nonlinear matrix concentration

FONTES, L.R., GOMES, P.A. and SANCHIS, R. 1745
Contact process under heavy-tailed renewals on finite graphs

(continued)

The papers published in Bernoulli are indexed or abstracted in Current Index to Statistics, Mathematical Reviews, Statistical Theory and Method Abstracts-Zentralblatt (STMA-Z), and Zentralblatt für Mathematik (also available on the MATH via STN database and Compact MATH CD-ROM). A list of forthcoming papers can be found online at http://www.bernoulli-society.org/index.php/publications/bernoulli-journal/bernoulli-journal-papers
BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Aims and Scope
BERNOULLI is the journal of the Bernoulli Society for Mathematical Statistics and Probability, issued four times per year. The journal provides a comprehensive account of important developments in the fields of statistics and probability, offering an international forum for both theoretical and applied work.

Bernoulli Society for Mathematical Statistics and Probability
The Bernoulli Society was founded in 1973. It is an autonomous Association of the International Statistical Institute, ISI. According to its statutes, the object of the Bernoulli Society is the advancement, through international contacts, of the sciences of probability (including the theory of stochastic processes) and mathematical statistics and of their applications to all those aspects of human endeavour which are directed towards the increase of natural knowledge and the welfare of mankind.

Meetings: http://www.bernoulli-society.org/index.php/meetings
The Society holds a World Congress every four years; more frequent meetings, coordinated by the Society’s standing committees and often organised in collaboration with other organisations, are the European Meeting of Statisticians, the Conference on Stochastic Processes and their Applications, the CLAPEM meeting (Latin-American Congress on Probability and Mathematical Statistics), the European Young Statisticians Meeting, and various meetings on special topics – in the physical sciences in particular. The Society, as an association of the ISI, also collaborates with other ISI associations in the organization of the biennial ISI World Statistics Congresses (formerly ISI Sessions).

Executive Committee
Detailed information about the members of the Executive Committee can be found on http://www.bernoulli-society.org/who-is-who

The papers published in Bernoulli are indexed or abstracted in Current Index to Statistics, Mathematical Reviews, Statistical Theory and Method Abstracts-Zentralblatt (STMA-Z), Thomson Scientific and Zentralblatt für Mathematik (also available on the MATH via STN database and Compact MATH CD-ROM). A list of forthcoming papers can be found online at http://www.bernoulli-society.org/index.php/publications/bernoulli-journal/bernoulli-journal-papers

©2021 International Statistical Institute/Bernoulli Society

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the Publisher.
In 2021 Bernoulli consists of 4 issues published in February, May, August and November.
Bootstrap based inference for sparse high-dimensional time series models

JONAS KRAMPE\textsuperscript{1}, JENS-PETER KREISS\textsuperscript{2} and EFSTATHIOS PAPARODITIS\textsuperscript{3}

\textsuperscript{1}University of Mannheim, Germany. \textit{E-mail: j.krampe@uni-mannheim.de}
\textsuperscript{2}Technische Universität Braunschweig, Germany. \textit{E-mail: j.kreiss@tu-bs.de}
\textsuperscript{3}University of Cyprus, Cyprus. \textit{E-mail: stathisp@ucy.ac.cy}

Fitting sparse models to high-dimensional time series is an important area of statistical inference. In this paper, we consider sparse vector autoregressive models and develop appropriate bootstrap methods to infer properties of such processes. Our bootstrap methodology generates pseudo time series using a model-based bootstrap procedure which involves an estimated, sparsified version of the underlying vector autoregressive model. Inference is performed using so-called de-sparsified or de-biased estimators of the autoregressive model parameters. We derive the asymptotic distribution of such estimators in the time series context and establish asymptotic validity of the bootstrap procedure proposed for estimation and, appropriately modified, for testing purposes. In particular, we focus on testing that large groups of autoregressive coefficients equal zero. Our theoretical results are complemented by simulations which investigate the finite sample performance of the bootstrap methodology proposed. A real-life data application is also presented.

\textit{Keywords:} De-sparsified estimators; testing; vector autoregressive models

\textbf{References}


Inference without compatibility: Using exponential weighting for inference on a parameter of a linear model

MICHAEL LAW* and YA’ACOV RITOV†

Department of Statistics, University of Michigan, Ann Arbor, USA.
E-mail: *mmylaw@umich.edu; †yritov@umich.edu

We consider hypotheses testing problems for three parameters in high-dimensional linear models with minimal sparsity assumptions of their type but without any compatibility conditions. Under this framework, we construct the first \( \sqrt{n} \)-consistent estimators for low-dimensional coefficients, the signal strength, and the noise level. We support our results using numerical simulations and provide comparisons with other estimators.

Keywords: Lasso; compatibility condition; exponential weighting; inference

References

Flux large deviations of weakly interacting jump processes via well-posedness of an associated Hamilton–Jacobi equation

RICHARD C. KRAAIJ

E-mail: r.c.kraaij@tudelft.nl

We establish uniqueness for a class of first-order Hamilton–Jacobi equations with Hamiltonians that arise from the large deviations of the empirical measure and empirical flux pair of weakly interacting Markov jump processes. As a corollary, we obtain such a large deviation principle in the context of weakly interacting processes with time-periodic rates in which the period-length converges to 0.

Keywords: Weakly interacting jump processes; empirical measure and flux; large deviations; Hamilton–Jacobi equation

References

Precise asymptotics of longest cycles in random permutations without macroscopic cycles

VOLKER BETZ\textsuperscript{1,*}, JULIAN MÜHLBAUER\textsuperscript{1}, HELGE SCHÄFER\textsuperscript{1,†} and DIRK ZEINDLER\textsuperscript{2}

\textsuperscript{1}Fachbereich Mathematik, Arbeitsgruppe Stochastik, Schlossgartenstrasse 7, 64289 Darmstadt, Germany. E-mail: * betz@mathematik.tu-darmstadt.de; † hschafer@mathematik.tu-darmstadt.de

\textsuperscript{2}Lancaster University, Mathematics and Statistics, Fylde College, Bailrigg, Lancaster LA1 4YF, United Kingdom. E-mail: d.zeindler@lancaster.ac.uk

We consider Ewens random permutations of length $n$ conditioned to have no cycle longer than $n^{\beta}$ with $0 < \beta < 1$ and study the asymptotic behaviour as $n \to \infty$. We obtain very precise information on the joint distribution of the lengths of the longest cycles; in particular we prove a functional limit theorem where the cumulative number of long cycles converges to a Poisson process in the suitable scaling. Furthermore, we prove convergence of the total variation distance between joint cycle counts and suitable independent Poisson random variables up to a significantly larger maximal cycle length than previously known. Finally, we remove a superfluous assumption from a central limit theorem for the total number of cycles proved in an earlier paper.

Keywords: Random permutations; Ewens measure; long cycles; functional limit theorem; Total variation distance; cycle structure

References


Explicit bounds for critical infection rates and expected extinction times of the contact process on finite random graphs

E. CATOR* and H. DON†

Faculty of Science, Radboud University Nijmegen, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands.
E-mail: *e.cator@math.ru.nl; †h.don@math.ru.nl

We introduce a method to prove metastability of the contact process on Erdős–Rényi graphs and on configuration model graphs. The method relies on uniformly bounding the total infection rate from below, over all sets with a fixed number of nodes. Once this bound is established, a simple comparison with a well chosen birth-and-death process will show the exponential growth of the extinction time. Our paper complements recent results on the metastability of the contact process: under a certain minimal edge density condition, we give explicit lower bounds on the infection rate needed to get metastability, and we have explicit exponentially growing lower bounds on the expected extinction time.

Keywords: Contact process; critical infection rate; exponential extinction; metastability

References


We provide existence, uniqueness and stability results for affine stochastic Volterra equations with $L^1$-kernels and jumps. Such equations arise as scaling limits of branching processes in population genetics and self-exciting Hawkes processes in mathematical finance. The strategy we adopt for the existence part is based on approximations using stochastic Volterra equations with $L^2$-kernels combined with a general stability result. Most importantly, we establish weak uniqueness using a duality argument on the Fourier–Laplace transform via a deterministic Riccati–Volterra integral equation. We illustrate the applicability of our results on Hawkes processes and a class of hyper-rough Volterra Heston models with a Hurst index $H \in (-1/2, 1/2]$.

Keywords: Stochastic Volterra equations; affine Volterra processes; Riccati–Volterra equations; superprocesses; Hawkes processes; rough volatility

References


Rates of contraction of posterior distributions based on $p$-exponential priors

SERGIOS AGAPIOU$^1$, MASOUMEH DASHTI$^2$ and TAPIO HELIN$^3$

$^1$Department of Mathematics and Statistics, University of Cyprus, Nicosia, Cyprus. 
E-mail: agapiou.sergios@ucy.ac.cy

$^2$Department of Mathematics, University of Sussex, Brighton, United Kingdom. 
E-mail: m.dashti@sussex.ac.uk

$^3$School of Engineering Science, LUT University, Lappeenranta, Finland. 
E-mail: tapio.helin@lut.fi

We consider a family of infinite dimensional product measures with tails between Gaussian and exponential, which we call $p$-exponential measures. We study their measure-theoretic properties and in particular their concentration. Our findings are used to develop a general contraction theory of posterior distributions on nonparametric models with $p$-exponential priors in separable Banach parameter spaces. Our approach builds on the general contraction theory for Gaussian process priors in (Ann. Statist. 36 (2008) 1435–1463), namely we use prior concentration to verify prior mass and entropy conditions sufficient for posterior contraction. However, the specific concentration properties of $p$-exponential priors lead to a more complex entropy bound which can influence negatively the obtained rate of contraction, depending on the topology of the parameter space. Subject to the more complex entropy bound, we show that the rate of contraction depends on the position of the true parameter relative to a certain Banach space associated to $p$-exponential measures and on the small ball probabilities of these measures. For example, we apply our theory in the white noise model under Besov regularity of the truth and obtain minimax rates of contraction using (rescaled) $\alpha$-regular $p$-exponential priors. In particular, our results suggest that when interested in spatially inhomogeneous unknown functions, in terms of posterior contraction, it is preferable to use Laplace rather than Gaussian priors.

Keywords: Bayesian nonparametric inference; non-Gaussian priors; concentration of measure

References


Yaglom’s limit for critical Galton–Watson processes in varying environment: A probabilistic approach

NATALIA CARDONA-TOBÓN and SANDRA PALAU

1Centro de Investigación en Matemáticas. Calle Jalisco s/n. C.P. 36240, Guanajuato, México. E-mail: natalia.cardona@cimat.mx
2IIMAS, Universidad Nacional Autónoma de México. CDMX, 04510, Ciudad de México, México. E-mail: sandra@sigma.iimas.unam.mx

A Galton–Watson process in varying environment is a discrete time branching process where the offspring distributions vary among generations. Based on a two-spine decomposition technique, we provide a probabilistic argument of a Yaglom-type limit for this family processes. The result states that, in the critical case, a suitable normalisation of the process conditioned on non-extinction converges in distribution to a standard exponential random variable.

Keywords: Galton–Watson processes; varying environment; Yaglom’s limit; spines decompositions

References

Correlation bounds, mixing and $m$-dependence under random time-varying network distances with an application to Cox-processes

ALEXANDER KREISS

ORSTAT KU Leuven, Naamsestraat 69, 3000 Leuven, Belgium.
E-mail: alexander.kreiss@kuleuven.be

We will consider multivariate stochastic processes indexed either by vertices or pairs of vertices of a dynamic network. Under a dynamic network, we understand a network with a fixed vertex set and an edge set which changes randomly over time. We will assume that the spatial dependence-structure of the processes conditional on the network behaves in the following way: Close vertices (or pairs of vertices) are dependent, while we assume that the dependence decreases conditionally on that the distance in the network increases. We make this intuition mathematically precise by considering three concepts based on correlation, $\beta$-mixing with time-varying $\beta$-coefficients and conditional independence. These concepts allow proving weak-dependence results, for example, an exponential inequality, which might be of independent interest. In order to demonstrate the use of these concepts in an application, we study the asymptotics (for growing networks) of a goodness of fit test in a dynamic interaction network model based on a Cox-type model for counting processes. This model is then applied to bike-sharing data.

Keywords: Dynamic networks; dependence; survival analysis; nonparametric regression; hypothesis testing

References


On the law of the iterated logarithm and strong invariance principles in stochastic geometry

JOHANNES KREBS

Department of Mathematics, TU Braunschweig, 38106 Braunschweig, Germany.
E-mail: johannes.krebs@tu-braunschweig.de

We study the law of the iterated logarithm (Khinchin (1924), Kolmogorov (1929)) and related strong invariance principles for functionals in stochastic geometry. As potential applications, we think of well-known functionals defined on the $k$-nearest neighbors graph and important functionals in topological data analysis such as the Euler characteristic and persistent Betti numbers.

Keywords: Binomial process; Euler characteristic; law of the iterated logarithm; persistent Betti numbers; Poisson process; stochastic geometry; strong invariance principles; strong stabilization; topological data analysis

References


From Poincaré inequalities to nonlinear matrix concentration

DE HUANG* and JOEL A. TROPP†

California Institute of Technology, Pasadena, USA.
E-mail: *dhuang@caltech.edu; †jtropp@cms.caltech.edu

This paper deduces exponential matrix concentration from a Poincaré inequality via a short, conceptual argument. Among other examples, this theory applies to matrix-valued functions of a uniformly log-concave random vector. The proof relies on the subadditivity of Poincaré inequalities and a chain rule inequality for the trace of the matrix Dirichlet form. It also uses a symmetrization technique to avoid difficulties associated with a direct extension of the classic scalar argument.

Keywords: Concentration inequality; functional inequality; Markov process; matrix concentration; Poincaré inequality; semigroup

References


Contact process under heavy-tailed renewals on finite graphs

LUIZ RENATO FONTES¹, PABLO ALMEIDA GOMES²,* and REMY SANCHIS²,†

¹Universidade de São Paulo, São Paulo, Brazil.
E-mail: lrfontes@usp.br
²Universidade Federal de Minas Gerais, Belo Horizonte, Brazil.
E-mail: * pabloag@ufmg.br; † rsanchis@mat.ufmg.br

We investigate a non-Markovian analogue of the Harris contact process in a finite connected graph $G = (V, E)$: an individual is attached to each site $x \in V$, and it can be infected or healthy; the infection propagates to healthy neighbors just as in the usual contact process, according to independent exponential times with a fixed rate $\lambda > 0$; however, the recovery times for an individual are given by the points of a renewal process attached to its timeline, whose waiting times have distribution $\mu$ such that $\mu(t, \infty) = t^{-\alpha}L(t)$, where $1/2 < \alpha < 1$ and $L(\cdot)$ is a slowly varying function; the renewal processes are assumed to be independent for different sites. We show that, starting with a single infected individual, if $|V| < 2 + (2\alpha - 1)/((1 - \alpha)(2 - \alpha))$, then the infection does not survive for any $\lambda$; and if $|V| > 1/(1 - \alpha)$, then, for every $\lambda$, the infection has positive probability to survive.

Keywords: Contact process; percolation; phase transition

References

Limit theorems for integral functionals of Hermite-driven processes

VALENTIN GARINO1,*, IVAN NOURDIN1,†, DAVID NUALART2 and MAJID SALAMAT3

1Département de mathématiques, Université du Luxembourg, Esch-sur-Alzette, Luxembourg. E-mail: *valentin.garino@uni.lu; †ivan.nourdin@uni.lu
2Kansas University, Lawrence, KS, USA. E-mail: nualart@ku.edu
3Department of Mathematical Sciences, Isfahan University of Technology, Isfahan, Iran. E-mail: msalamat@iut.ac.ir

Consider a moving average process $X$ of the form $X(t) = \int_{-\infty}^{t} \varphi(t-u) dZ_u$, $t \geq 0$, where $Z$ is a (non Gaussian) Hermite process of order $q \geq 2$ and $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ is sufficiently integrable. This paper investigates the fluctuations, as $T \to \infty$, of integral functionals of the form $t \mapsto \int_0^T P(X(s)) ds$, in the case where $P$ is any given polynomial function. It extends a study initiated in (Stoch. Dyn. 18 (2018) 1850028, 18), where only the quadratic case $P(x) = x^2$ and the convergence in the sense of finite-dimensional distributions were considered.

Keywords: Hermite processes; chaotic decomposition; fractional Brownian motion (fBm); multiple Wiener–Itô integrals

References


Invariance and attraction properties of Galton–Watson trees

YEVGENIY KOVCHEGOV and ILYA ZALIAPIN

1Department of Mathematics, Oregon State University, Corvallis, OR 97331-4605, USA.
E-mail: kovchegy@math.oregonstate.edu; url: https://math.oregonstate.edu/people/view/kovchegy

2Department of Mathematics and Statistics, University of Nevada, Reno, NV 89557-0084, USA.
E-mail: zal@unr.edu

We give a description of invariants and attractors of the critical and subcritical Galton–Watson tree measures under the operation of Horton pruning (cutting tree leaves with subsequent series reduction). Under a regularity condition, the class of invariant measures consists of the critical binary Galton–Watson tree and a one-parameter family of critical Galton–Watson trees with offspring distribution \( \{ q_k \} \) that has a power tail \( q_k \sim Ck^{-(1+1/q_0)} \), where \( q_0 \in (1/2, 1) \). Each invariant measure has a non-empty domain of attraction under consecutive Horton pruning, specified by the tail behavior of the initial Galton–Watson offspring distribution. The invariant measures satisfy the Toeplitz property for the Tokunaga coefficients and obey the Horton law with exponent \( R = (1-q_0)^{-1}/q_0 \).

Keywords: Galton–Watson processes; self-similar trees; Horton–Strahler order; invariant measures; attractor

References

The coupling method in extreme value theory

BENJAMIN BOBBIA *, CLÉMENT DOMBRY † and DAVIT VARRON ‡

Université Bourgogne Franche-Comté, Laboratoire de Mathématiques de Besançon, UMR CNRS 6623, 16 route de Gray, 25030 Besançon Cedex, France.
E-mail: * benjamin.bobbia@univ-fcomte.fr; † clement.dombry@univ-fcomte.fr; ‡ davit.varron@univ-fcomte.fr

A coupling method is developed for univariate extreme value theory, providing an alternative to the use of the tail empirical/quantile processes. We emphasize the Peak-over-Threshold approach that approximates the distribution above high threshold by the Generalized Pareto Distribution (GPD) and compare the empirical distribution of exceedances to the empirical distribution associated with the limit GPD model. Sharp bounds for their Wasserstein distance in the second order Wasserstein space are provided. As an application, we recover standard results on the asymptotic behavior of the Hill estimator, the Weissman extreme quantile estimator or the probability weighted moment estimators, shedding some new light on the theory.

Keywords: Extreme value theory; coupling method; Wasserstein distance

References

Asymptotics and renewal approximation in the online selection of increasing subsequence

ALEXANDER GNEDIN* and AMIRLAN SEKSENBAYEV†

School of Mathematical Sciences, Queen Mary University of London, United Kingdom.
E-mail: *a.gnedin@qmul.ac.uk; †a.seksenbayev@qmul.ac.uk

We revisit the problem of maximising the expected length of increasing subsequence that can be selected from a marked Poisson process by an online strategy. Resorting to a natural size variable, we represent the problem in terms of a controlled piecewise deterministic Markov process with decreasing paths. We apply a comparison method to the optimality equation to obtain fairly complete asymptotic expansions for the moments of the maximal length, and, with the aid of a renewal approximation, give a novel proof to the central limit theorem for the length of selected subsequence under either the optimal strategy or a strategy sufficiently close to optimality.

Keywords: Online selection; monotone subsequence; renewal approximation; dynamic programming

References


A convolution formula for the local time of an Itô diffusion reflecting at 0 and a generalized Stroock–Williams equation

JACEK JAKUBOWSKI* and MACIEJ WIŚNIEWSKI†

Institute of Mathematics, University of Warsaw, Banacha 2, 02-097 Warszawa, Poland.
E-mail: *jakub@mimuw.edu.pl; †wisniewolski@mimuw.edu.pl

A new probabilistic insight into the structure of local time is presented. A convolution formula for the local time at 0 of Itô diffusions reflecting at 0 is obtained. A simple integro-differential equation for the cumulative distribution function of the local time is given. Finally, a probabilistic representation of a generalized Stroock–Williams equation is presented.

Keywords: Itô diffusion; local time; excursions of Markov processes; Stroock–Williams equation

References

Stochastic PDEs on graphs as scaling limits of discrete interacting systems

WAI-TONG LOUIS FAN1,2

1 Department of Mathematics, Indiana University, Bloomington, USA.
E-mail: waifan@iu.edu
2 Center of Mathematical Sciences and Applications, Harvard University, Cambridge, USA.

Stochastic partial differential equations (SPDE) on graphs were recently introduced by Cerrai and Freidlin (Ann. Inst. Henri Poincaré Probab. Stat. 53 (2017) 865–899). This class of stochastic equations in infinite dimensions provides a minimal framework for the study of the effective dynamics of much more complex systems. However, how they emerge from microscopic individual-based models is still poorly understood, partly due to complications near vertex singularities. In this work, motivated by the study of the dynamics and the genealogies of expanding populations in spatially structured environments, we obtain a new class of SPDE on graphs of Wright–Fisher type which have nontrivial boundary conditions on the vertex set. We show that these SPDE arise as scaling limits of suitably defined biased voter models (BVM), which extends the scaling limits of Durrett and Fan (Ann. Appl. Probab. 26 (2016) 3456–3490). We further obtain a convergent simulation scheme for each of these SPDE in terms of a system of Itô SDEs, which is useful when the size of the BVM is too large for stochastic simulations. These give the first rigorous connection between SPDE on graphs and more discrete models, specifically, interacting particle systems and interacting SDEs. Uniform heat kernel estimates for symmetric random walks approximating diffusions on graphs are the keys to our proofs. Some open problems are provided as further motivations of our study.

Keywords: Stochastic partial differential equation; graph; interacting particle system; numerical scheme; duality; scaling limit; population dynamics

References


Is there an analog of Nesterov acceleration for gradient-based MCMC?

YI-AN MA¹, NILADRI S. CHATTERJI², XIANG CHENG³, NICOLAS FLAMMARION⁴, PETER L. BARTLETT³,⁵,* and MICHAEL I. JORDAN³,⁵,†

¹Halıcıo˘glu Data Science Institute, University of California San Diego. E-mail: yianma@ucsd.edu
²Department of Physics, University of California Berkeley. E-mail: chatterji@berkeley.edu
³Department of Electrical Engineering and Computer Sciences, University of California Berkeley. E-mail: x.cheng@berkeley.edu
⁴School of Computer and Communication Sciences, École Polytechnique Fédérale de Lausanne. E-mail: nicolas.flammarion@epfl.ch
⁵Department of Statistics, University of California Berkeley. E-mail: *peter@berkeley.edu; †jordan@cs.berkeley.edu

We formulate gradient-based Markov chain Monte Carlo (MCMC) sampling as optimization on the space of probability measures, with Kullback–Leibler (KL) divergence as the objective functional. We show that an underdamped form of the Langevin algorithm performs accelerated gradient descent in this metric. To characterize the convergence of the algorithm, we construct a Lyapunov functional and exploit hypocoercivity of the underdamped Langevin algorithm. As an application, we show that accelerated rates can be obtained for a class of nonconvex functions with the Langevin algorithm.

Keywords: Markov chain Monte Carlo; Langevin Monte Carlo; accelerated gradient descent; sampling algorithms

References


Scoring interval forecasts: Equal-tailed, shortest, and modal interval

JONAS R. BREHMER¹,²,* and TILMANN GNEITING²,³,†

¹Institute for Mathematics, University of Mannheim, B6, 26, 68159 Mannheim, Germany
²Computational Statistics Group, Heidelberg Institute for Theoretical Studies, Schloss-Wolfsbrunnenweg 35, 69118 Heidelberg, Germany.
E-mail: * jonas.brehmer@h-its.org; †tilmann.gneiting@h-its.org
³Institute for Stochastics, Karlsruhe Institute of Technology (KIT), Englerstraße 2, 76131 Karlsruhe, Germany

We consider different types of predictive intervals and ask whether they are elicitable, that is, are unique minimizers of a loss or scoring function in expectation. The equal-tailed interval is elicitable, with a rich class of suitable loss functions, though subject to translation invariance, or positive homogeneity and differentiability, the Winkler interval score becomes a unique choice. The modal interval also is elicitable, with a sole consistent scoring function, up to equivalence. However, the shortest interval fails to be elicitable relative to practically relevant classes of distributions. These results provide guidance in interval forecast evaluation and support recent choices of performance measures in forecast competitions.

Keywords: Elicitability; forecast evaluation; interval forecast; modal interval; predictive performance; scoring function

References

Variable Length Memory Chains: Characterization of stationary probability measures

PEGGY CÉNAC¹, BRIGITTE CHAUVIN²,* CAMILLE NOÛS³, FRÉDÉRIC PACCAUT⁴ and NICOLAS POUYANNE²,†

¹Institut de Mathématiques de Bourgogne, IMB UMR 5584 CNRS, 9 rue Alain Savary - BP 47870, 21078 DIJON CEDEX, France.
E-mail: peggy.cenac@u-bourgogne.fr

²Université Paris-Saclay, UVSQ, CNRS UMR 8100, Laboratoire de Mathématiques de Versailles, 78000 Versailles, France.
E-mail: *brigitte.chauvin@uvsq.fr; †nicolas.pouyanne@uvsq.fr

³Cogitamus Laboratory, 78000 Versailles, France.
E-mail: camille.nous@cogitamus.fr

⁴Laboratoire Amiénois de Mathématique Fondamentale et Appliquée, CNRS UMR 7352, Université de Picardie Jules Verne, 33 rue Saint-Leu, 80039 Amiens, France.
E-mail: frédéric.paccaut@u-picardie.fr

Variable Length Memory Chains (VLMC), which are generalizations of finite order Markov chains, are an essential tool to modelize random sequences in many domains, as well as an interesting object in contemporary probability theory. The question of existence of stationary probability measures leads us to introduce a key combinatorial structure for words produced by a VLMC: the Longest Internal Suffix. This notion allows us to state a necessary and sufficient condition for a VLMC to admit a unique invariant probability measure.

This condition turns out to get a much simpler form for a subclass of VLMC: the stable VLMC. This natural subclass, unlike the general case, enjoys a renewal property. Namely, a stable VLMC induces a semi-Markov chain on an at most countable state space. Unfortunately, this discrete time renewal process does not contain the whole information of the VLMC, preventing the study of a stable VLMC to be reduced to the study of its induced semi-Markov chain. For a subclass of stable VLMC, the convergence in distribution of a VLMC towards its stationary probability measure is established.

Finally, finite state space semi-Markov chains turn out to be very special stable VLMC, shedding some new light on their limit distributions.

Keywords: Variable Length Memory Chains; stationary probability measure; Longest Internal Suffix; stable context trees; semi-Markov chains

References


Approximation of heavy-tailed distributions via stable-driven SDEs

LU-JING HUANG¹, MATEUSZ B. MAJKA² and JIAN WANG³

¹College of Mathematics and Informatics, Fujian Normal University, 350007 Fuzhou, P.R. China. E-mail: lujingh@yeah.net
²School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, EH14 4AS, UK. E-mail: m.majka@hw.ac.uk
³College of Mathematics and Informatics & Fujian Key Laboratory of Mathematical Analysis and Applications (FJKLMAA) & Center for Applied Mathematics of Fujian Province (FJNU), Fujian Normal University, 350007 Fuzhou, P.R. China. E-mail: jianwang@fjnu.edu.cn

Constructions of numerous approximate sampling algorithms are based on the well-known fact that certain Gibbs measures are stationary distributions of ergodic stochastic differential equations (SDEs) driven by the Brownian motion. However, for some heavy-tailed distributions it can be shown that the associated SDE is not exponentially ergodic and that related sampling algorithms may perform poorly. A natural idea that has recently been explored in the machine learning literature in this context is to make use of stochastic processes with heavy tails instead of the Brownian motion. In this paper, we provide a rigorous theoretical framework for studying the problem of approximating heavy-tailed distributions via ergodic SDEs driven by symmetric (rotationally invariant) α-stable processes.

Keywords: Stochastic differential equations; symmetric α-stable processes; invariant measures; heavy-tailed distributions; approximate sampling; fractional Langevin Monte Carlo

References


We consider the problem of delay estimation by the observations of the solutions of several SDEs. It is known that the MLEs for these models are consistent and asymptotically normal, but the likelihood ratio functions are not differentiable w.r.t. the parameter, and therefore the numerical calculation of the MLEs encounter certain difficulties. We propose One-step and Two-step MLEs, whose calculation has no such problems and provide an estimator asymptotically equivalent to the MLE. These constructions are realized in two or three steps. First, we construct preliminary estimators which are consistent and asymptotically normal, but not asymptotically efficient. Then we use these estimators and a modified Fisher-score device to obtain One-step and Two-step MLEs. We suppose that its numerical realization is much more simple. Stochastic Pantograph equation is introduced and related statistical problems are discussed.

Keywords: One-step MLE; Two-step MLE; One-step MDE; Stochastic Pantograph equation; delay estimation

References


Context-specific independencies in stratified chain regression graphical models

FEDERICA NICOLUSSI\(^1\) and MANUELA CAZZARO\(^2\)

\(^1\)University of Milan, Via Conservatorio, 7, 20122 Milano MI, Italy. E-mail: federica.nicolussi@unimi.it
\(^2\)University of Milan Bicocca, Via Bicocca Degli Arcimboldi 8, 20126 Milano, MI, Italy. E-mail: manuela.cazzaro@unimib.it

Graphical models are a useful tool with increasing diffusion. In the categorical variable framework, they provide important visual support to understand the relationships among the considered variables. Besides, particular chain graphical models are suitable to represent multivariate regression models. However, the associated parameterization, such as marginal log-linear models, is often difficult to interpret when the number of variables increases because of a large number of parameters involved. On the contrary, conditional and marginal independencies reduce the number of parameters needed to represent the joint probability distribution of the variables. In compliance with the parsimonious principle, it is worthwhile to consider also the so-called context-specific independencies, which are conditional independencies holding for particular values of the variables in the conditioning set. In this work, we propose a particular chain graphical model able to represent these context-specific independencies through labeled arcs. We provide also the Markov properties able to describe marginal, conditional, and context-specific independencies from this new chain graph. Finally, we show the results in an application to a real data set.

**Keywords:** Graphical models; stratified Markov properties; categorical variables; multivariate regression models; marginal models

**References**


A new look at random projections of the cube and general product measures

ZAKHAR KABLUCHKO¹, JOSCHA PROCHNO² and CHRISTOPH THÄLE³

¹Institut für Mathematische Stochastik, Westfälische Wilhelms-Universität Münster, Germany. E-mail: zakharkabluchko@uni-muenster.de
²Institut für Mathematik & Wissenschaftliches Rechnen, Karl-Franzens-Universität Graz, Austria. E-mail: joscha.prochno@uni-graz.at
³Faculty of Mathematics, Ruhr University Bochum, Germany. E-mail: christoph.thaele@rub.de

A consequence of the celebrated Dvoretzky–Milman theorem is a strong law of large numbers for $d$-dimensional random projections of the $n$-dimensional cube. It shows that, with respect to the Hausdorff distance, a uniform random projection of the cube $[-1/\sqrt{n}, +1/\sqrt{n}]^n$ onto $\mathbb{R}^d$ converges almost surely to a centered $d$-dimensional Euclidean ball of radius $\sqrt{2/\pi}$, as $n \to \infty$. We start by providing an alternative proof of this strong law via the Artstein–Vitale law of large numbers for random compact sets. Then, for every point inside the ball of radius $\sqrt{2/\pi}$, we determine the asymptotic number of vertices and the volume of the part of the cube projected ‘close’ to this point. More generally, we study large deviations for random projections of arbitrary product measures. Let $\nu^\otimes n$ be the $n$-fold product measure of a Borel probability measure $\nu$ on $\mathbb{R}$, and let $I$ be uniformly distributed on the Stiefel manifold of orthogonal $d$-frames in $\mathbb{R}^n$. It is shown that the sequence of random measures $\nu^\otimes n \circ (n^{-1/2} I^*)^{-1}$, $n \in \mathbb{N}$, satisfies a large deviation principle with probability 1. The rate function is explicitly identified in terms of the moment generating function of $\nu$. At the heart of the proofs lies a transition trick which allows to replace the uniform projection by the Gaussian one. A number of concrete examples are discussed as well, including the uniform distributions on the cube $[-1, 1]^n$ and the discrete cube $\{-1, 1\}^n$ as special cases.

Keywords: Cube; Gaussian random matrices; Gaussian projections; high-dimensional probability; Hausdorff distance; large deviations; law of large numbers; random projections; Stiefel manifold

References


Nearly optimal robust mean estimation via empirical characteristic function

SOHAIL BAHMANI

School of Electrical & Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA.
E-mail: sohail.bahmani@ece.gatech.edu

We propose an estimator for the mean of random variables in separable real Banach spaces using the empirical characteristic function. Assuming that the covariance operator of the random variable is bounded in a precise sense, we show that the proposed estimator achieves the optimal sub-Gaussian rate, except for a faster decaying mean-dependent term. Under a mild condition, an iterative refinement procedure can essentially eliminate the mean-dependent term and provide a shift-equivariant estimate. Our results particularly suggests that a certain Gaussian width that appears in the best known rate in the literature might not be necessary. Furthermore, using the boundedness of the characteristic functions, we also show that, except possibly at high signal-to-noise ratios, the proposed estimator is gracefully robust against adversarial “contamination”. Our analysis is overall concise and transparent, thanks to the tractability of the characteristic functions.

Keywords: Mean estimation; robust estimation; characteristic function

References


Bernoulli Forthcoming Papers

GUHA, A., HO, N. and NGUYEN, X.
On posterior contraction of parameters and interpretability in Bayesian mixture modeling

BABIĆ, S., GELBGRAS, L., HALLIN, M. and LEY, C.
Optimal tests for elliptical symmetry: Specified and unspecified location

KASPER, T.M., FUCHS, S. and TRUTSCHNIG, W.
On weak conditional convergence of bivariate Archimedean and Extreme Value copulas, and consequences to nonparametric estimation

PAGE, S. and GRÜNEWÄLDER, S.
The Goldenshluger–Lepski method for constrained least-squares estimators over RKHSs

KLEBANOV, I., SPRUNGK, B. and SULLIVAN, T.J.
The linear conditional expectation in Hilbert space

QIU, H., LUEDTKE, A. and CARONE, M.
Universal sieve-based strategies for efficient estimation using machine learning tools

JIANG, D. and BAI, Z.
Partial generalized four moment theorem revisited

PETE, G. and TIMÁR, Á.
Finite-energy infinite clusters without anchored expansion

KRELL, N. and SCHMISSER, É.
Nonparametric estimation of jump rates for a specific class of piecewise deterministic Markov processes

CHZHEN, E., DENIS, C. and HEBIRI, M.
Minimax semi-supervised set-valued approach to multi-class classification

GONZÁLEZ CÁZARES, J. and IVANOVS, J.
Recovering Brownian and jump parts from high-frequency observations of a Lévy process

BARBER, R.F. and SAMWORTH, R.J.
Local continuity of log-concave projection, with applications to estimation under model misspecification

GANESAN, G.
Minimum spanning trees of random geometric graphs with location dependent weights

BHUDISAKSANG, T. and CARTEA, Á.
Online drift estimation for jump-diffusion processes

CHAUDHURI, R., JAIN, V. and PILLAI, N.S.
Universality and least singular values of random matrix products: A simplified approach

Continues
NING, N., IONIDES, E.L. and RITOV, Y.
Scalable Monte Carlo inference and rescaled local asymptotic normality

WANG, D. and TANG, C.-F.
Testing against uniform stochastic ordering with paired observations

KOHLER, M. and KRZYŻAK, A.
Over-parametrized deep neural networks do not generalize well

BORDA, B.
Equidistribution of random walks on compact groups II. The Wasserstein metric

ČERNÝ, A. and RUF, J.
Pure-jump semimartingales

DUVAL, C. and MARIUCCI, E.
Spectral-free estimation of Lévy densities in high-frequency regime

DENIS, C., DION-BLANC, C. and MARTINEZ, M.
A ridge estimator of the drift from discrete repeated observations of the solution of a stochastic differential equation

ALTMEYER, R.
Approximation of occupation time functionals

KWAK, J., LEE, J.O. and PARK, J.
Extremal eigenvalues of sample covariance matrices with general population

XU, W.
Asymptotic results for heavy-tailed Lévy processes and their exponential functionals

LI, X.-Y., YE, Z.-S. and TANG, C.Y.
Estimating the inter-occurrence time distribution from superposed renewal processes

BRAUNSTEINS, P., HAUTPHENNE, S. and MINUESA, C.
Parameter estimation in branching processes with almost sure extinction

BONNEFONT, M. and JOULIN, A.
A note on eigenvalues estimates for one-dimensional diffusion operators

HAN, Q.
Multiplier U-processes: Sharp bounds and applications

ERNY, X., LÖCHERBACH, E. and LOUKIANOVA, D.
Mean field limits for interacting Hawkes processes in a diffusive regime

IACOBELLI, G., RIBEIRO, R., VALLE, G. and ZUAZNÁBAR, L.
Tree builder random walk: Recurrence, transience and ballisticity

CHÉRIF-ABDELLATIF, B.-E. and ALQUIER, P.
Finite sample properties of parametric MMD estimation: Robustness to misspecification and dependence

Continues
KIM, K. and YI, J.  
Limit theorems for time-dependent averages of nonlinear stochastic heat equations

GROISMAN, P., JONCKHEERE, M. and SAPIENZA, F.  
Nonhomogeneous Euclidean first-passage percolation and distance learning

LEE, C.Y.  
The Hausdorff measure of the range and level sets of Gaussian random fields with sectorial local nondeterminism

HAYASHI, M.  
Information geometry approach to parameter estimation in hidden Markov model

BACKHOFF-VERAGUAS, J. and PAMMER, G.  
Applications of weak transport theory

LEDOIT, O. and WOLF, M.  
Quadratic shrinkage for large covariance matrices

VERGARA, R.C., ALLARD, D. and DESASSIS, N.  
A general framework for SPDE-based stationary random fields

SHAO, Q.M. and ZHANG, Z.-S.  
Berry–Esseen bounds for multivariate nonlinear statistics with applications to M-estimators and stochastic gradient descent algorithms

ELTZNER, B.  
Geometrical smeariness – a new phenomenon of Fréchet means

MOU, W., FLAMMARION, N., WAINWRIGHT, M.J. and BARTLETT, P.J.  
Improved bounds for discretization of Langevin diffusions: Near-optimal rates without convexity

SUN, X., XIE, L. and XIE, Y.  
Strong and weak convergence rates for slow-fast stochastic differential equations driven by $\alpha$-stable process

BENHENNI, K., GIRARD, D.A. and LOUHICHI, S.  
On bandwidth selection problems in nonparametric trend estimation under martingale difference errors

BELLEC, P.C. and ZHANG, C.-H.  
De-biasing the Lasso with degrees-of-freedom adjustment

BURACZEWSKI, D. and DAMEK, E.  
Limit theorems for supercritical branching processes in random environment

EL NAMIER, R., LUSCHGY, H. and PAGÉS, G.  
New approach to greedy vector quantization

PHANDOIDAEN, N. and RICHTER, S.  
Empirical process theory for locally stationary processes

Continues
WANG, Y. and POLITIS, D.
Model-free bootstrap for a general class of stationary time series

COSSO, A. and RUSSO, F.
Crandall–Lions viscosity solutions for path-dependent PDEs: The case of heat equation

GAUTIER, E.
Adaptive estimation in the linear random coefficients model when regressors have limited variation

CAO, H. and WU, W.B.
Testing and estimation for clustered signals

GRAHOVAC, D., LEONENKO, N.N. and TAQQU, M.S.
Growth and intermittency of supOU processes

FRANCQ, C. and ZAKOIAN, J.-M.
Adaptiveness of the empirical distribution of residuals in semi-parametric conditional location scale models

LAM-WEIL, J., LAURENT, B. and LOUBES, J.-M.
Minimax optimal goodness-of-fit testing for densities and multinomials under a local differential privacy constraint

GAYNANOVA, I.
Erratum for prediction and estimation consistency of sparse multi-class penalized optimal scoring

LI, W., WANG, Q., YAO, J. and ZHOU, W.
On eigenvalues of a high-dimensional spatial-sign covariance matrix

LAU, J.W. and CRIPPS, E.
Thinned completely random measures with applications in competing risks models

BHAUMIK, P., SHI, W. and GHOSAL, S.
Bayesian generalized regression in partial differential equation models

ROGERS, D. and WINKEL, M.
A Ray–Knight representation of up-down Chinese Restaurants

HUANG, Q., DUAN, J. and SONG, R.
Homogenization of nonlocal partial differential equations related to stochastic differential equations with Lévy noise