Workshop On
PROBABILITY THEORY and its APPLICATIONS

Peter Eichelsbacher, Matthias Löwe

Contents

Preface

This preprint is a collection of abstracts of talks held on the workshop on " PROBABILITY THEORY and its APPLICATIONS" held on October, 9th and October, 10th 1997 in Bielefeld.
The workshop was supported by the University of Bielefeld and the SFB 343 "Diskrete Strukturen in der Mathematik" and the "Schwerpunkt Interagierende stochastische Systeme von hoher Komplexität".

We thank all participants of the workshop for their coming and for contributing to this collection.

FLUCTUATIONS OF THE FREE ENERGY IN THE KAC-VERSION OF THE SHERRINGTON-KIRKPATRICK MODEL AT HIGH TEMPERATURES[#]

Anton Bovier [1]

Weierstrass –Institut
für Angewandte Analysis und Stochastik
Mohrenstrasse 39, D-10117 Berlin, Germany

1. Introduction and results

Recently, there has been a revival of interest in a class of models introduced more than 30 years ago by Marc Kac, and which we will generically call Kac models. They are characterized by interactions that contain a parameter that can be tuned to let the range of the interaction tend to infinity while the strength tends to zero in a measured way. The main motivation behind these model is that they allowed to give a rigorous derivation of the van der Waals theory, including the Maxwell construction [KUH], as an asymptotic theory. Our interest has been mainly connected to the question to what extent one may use Kac models in in the theory of *disordered systems* to link mean field results to properties of lattice systems. Some results on this line in the context of Hopfield models can be found in [BGP1,BGP2,BGP3].

One of the most disputed issues in the theory of spin-glasses is the question as to what extent the results obtained for the mean-field Sherrington-Kirkpatrick (SK) model [SK] are relevant for finite dimensional short-range spin glasses (for a recent review and interesting discussions on this issue see [NS1,NS2]). It is thus natural to consider the Kac-version of the SK model. Given that even in the SK model, virtually all rigorous results known so far are high-temperature results [ALR,FZ,CN,Co,T1,T2], our ambitions therefore are limited to the question to what extend these reflect properties of finite range spin glasses at high temperatures. More precisely, our aim is to

Invited talk presented at the Workshop "Probability Theory and its Applications", Bielefeld, October 1997

1 e-mail: bovier@wias-berlin.de

apply the techniques developed in [CN] and in [T1] to study the free energy and its fluctuations to the Kac-SK-model. Let us mention that the Kac-SK-model was first introduced (to our knowledge) By Fröhlich and Zegarlinski [FZ2].

Let us define the Kac-SK model. Let $\sigma_i \in \{-1,1\}$, $i \in \mathbb{Z}^d$ be Ising spins. Let G_{ij}, $i,j \in \mathbb{Z}^d$ be a family of i.i.d. random variables with mean zero and variance 1. To avoid complications, in this paper we only consider the case where G_{ij} are Gaussian, but more general distributions could be considered. Let $J_\gamma(i)$ be a Kac-kernel, i.e. as positive function $\mathbb{Z}^d \to \mathbb{R}^+$ such that $J_\gamma(i) = \gamma^d J(\gamma i)$ where $\int d^d x J(x) = 1$ and J has compact support. Then the Hamiltonian of our model is defined, for any $\Lambda \subset \mathbb{Z}^d$ by

$$H_{\Lambda,\gamma}(\sigma)[\omega] = -\frac{1}{\sqrt{2}} \sum_{i,j\in\Lambda} \sqrt{J_\gamma(i-j)} G_{ij}\sigma_i\sigma_j \tag{1.1}$$

We define the partition function as

$$Z_{\Lambda,\gamma,\beta}[\omega] \equiv \mathbb{E}_\sigma \exp\left(-\beta H_{\Lambda,\gamma}(\sigma)[\omega]\right) \tag{1.2}$$

and the free energy as

$$f_{\Lambda,\gamma,\beta}[\omega] \equiv -\frac{1}{\beta|\Lambda|} \ln Z_{\Lambda,\gamma,\beta}[\omega] \tag{1.3}$$

For notational convenience we also introduce the *quenched* and *annealed* free energies as $f^q_{\Lambda.\gamma,\beta} \equiv \mathbb{E} f_{\Lambda,\gamma,\beta}[\omega]$, respectively $f^a_{\Lambda.\gamma,\beta} \equiv -\frac{1}{\beta|\Lambda|} \ln \mathbb{E} Z_{\Lambda,\gamma,\beta}[\omega]$. As is well known, it follows from the subadditive ergodic theorem that for all $\gamma > 0$, for almost all ω,

$$\lim_{\Lambda\uparrow\mathbb{Z}^d} f_{\Lambda,\gamma,\beta}[\omega] = \lim_{\Lambda\uparrow\mathbb{Z}^d} f^q_{\Lambda,\gamma,\beta} = f_{\gamma,\beta} \tag{1.4}$$

Then our first result can be formulated as follows:

Theorem 1: *For all $\beta < 1$*

$$\lim_{\gamma\downarrow 0} f_{\gamma,\beta} = -\frac{\beta}{4}, \quad a.s. \tag{1.5}$$

Remark: From the work of [ALR] we know that the free energy of the SK model equals to $-\frac{\beta}{4}$ for $\beta \leq 1$. Theorem 1 thus says that in the high temperature regime, the free energy of the Kac-SK model converges to that of the SK model in the Lebowitz-Penrose limit.

Our second result concerns the fluctuations of the free energy.

Theorem 2: *Let $\beta < 1$. Assume moreover that β is such that for all γ sufficiently small, the Kac-SK model has a unique infinite volume Gibbs state for all $\beta' \leq \beta$. Then* *(i) If γ is small enough, there exists a constant $c_{\gamma,\beta}$ such that*

$$\sqrt{|\Lambda|}\left(f_{\Lambda,\gamma,\beta}[\omega]-f^q_{\Lambda,\gamma,\beta}\right)\xrightarrow{\mathcal{D}} c_{\gamma,\beta} g \tag{1.6}$$

where g is a standard gaussian random variable and $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution (or "in law").

(ii) For Lebesgue almost all such β

$$\lim_{\gamma\downarrow 0} c_{\gamma,\beta}=0 \tag{1.7}$$

Remark: In the SK model it was proven that $|\Lambda|\left(f_{\Lambda,\gamma,\beta}[\omega]-f^q_{\Lambda,\gamma,\beta}\right)$ converges in distribution to a standard normal r.v. so that $\sqrt{|\Lambda|}\left(f_{\Lambda,\gamma,\beta}[\omega]-f^q_{\Lambda,\gamma,\beta}\right)$ converges to zero. Thus, on the level of the "normal" fluctuations on the scale $\sqrt{|\Lambda|}$ our theorem states that the properties in the Kac model converge to those of the SK model.

Remark: The weak uniqueness of the Gibbs measure follows from the work of Fröhlich and Zegarlinski [FZ2] for $\beta<\beta_c$, with some $\beta_c>0$, independent of γ (say for $\gamma\in[0,1/4]$). Of course we would expect this to be true for $\beta<\beta_c(\gamma)$, where $\beta_c(\gamma)\uparrow 1$, as $\gamma downarrow 0$, but except in $d=1$, this would require some extra work.

The two theorems above have been proven in [B], and details of the proofs can be found there. Below I give a brief outline of the proofs. In the presentation of the proof of Theorem 2 below I will emphasize a slightly different point of view than in the original paper.

2. Proof of Theorem 1

The proof of Theorem 1 follows closely Talagrand's proof in the SK model that appeared in [T1]. I just give a very brief sketch. There are three key ingredients:

Lemma 2.1: *For any β and γ,*

$$I\!\!E Z_{\Lambda,\gamma,\beta}=e^{|\Lambda|\beta^2/4} \tag{2.1}$$

Lemma 2.2: *For any β and γ,*

$$\frac{I\!\!E Z^2_{\Lambda,\gamma,\beta}}{\left[I\!\!E Z_{\Lambda,\gamma,\beta}\right]^2}=\exp\left(-\beta|\Lambda|\left(F^{\mathrm{kac}}_{\gamma,\beta^2}+o(1)\right)\right) \tag{2.2}$$

where $F^{\mathrm{kac}}_{\gamma,\beta}$ is the free energy of the Kac-Ising model (for a precise definition, see below). Moreover, if $\beta\leq 1$, then

$$\lim_{\gamma\downarrow 0} F^{\mathrm{kac}}_{\gamma,\beta^2}=0 \tag{2.3}$$

Lemma 2.3: *For all β and all γ,*

$$\mathbb{P}\left[|f_{\Lambda,\gamma,\beta} - \mathbb{E}f_{\Lambda,\gamma,\beta}| > x\right] \leq 2e^{-x^2|\Lambda|} \tag{2.4}$$

Remark: The proofs of Lemmata 2.1 and 2.2 are simple computations, plus, in Lemma 2.2 some well-known facts on the Kac-Ising model. Ther are more details in [B]. Lemma 2.3 is an immediate consequence of standard Gaussian concentration inequalities (see e.g. [LT]).

We now have all the tools ready to apply Talgrands idea to the Kac-SK model. The Paley-Szygmund inequality (see [T1]; the proof of this inequality is elementary) asserts that

$$\mathbb{P}\left[Z_{\Lambda,\gamma,\beta} > \frac{1}{2}\mathbb{E}Z_{\Lambda,\gamma,\beta}\right] \leq \frac{1}{4}\frac{\left[\mathbb{E}Z_{\Lambda,\gamma,\beta}\right]^2}{\mathbb{E}Z^2_{\Lambda,\gamma,\beta}} = \frac{1}{4}\exp\left(-\beta|\Lambda|\left[F^{\text{kac}}_{\gamma,\beta^2} + o(1)\right]\right) \tag{2.5}$$

On the other hand,

$$\begin{aligned}
&\mathbb{P}\left[Z_{\Lambda,\gamma,\beta} > \frac{1}{2}\mathbb{E}Z_{\Lambda,\gamma,\beta}\right] = \mathbb{P}\left[\ln Z_{\Lambda,\gamma,\beta} - \mathbb{E}\ln Z_{\Lambda,\gamma,\beta} > \ln \mathbb{E}Z_{\Lambda,\gamma,\beta} - \mathbb{E}\ln Z_{\Lambda,\gamma,\beta} - \ln 2\right]\\
&= \mathbb{P}\left[-f_{\Lambda,\gamma,\beta} + \mathbb{E}f_{\Lambda,\gamma,\beta} > -f^a_{\Lambda,\gamma,\beta} + f^q_{\Lambda,\gamma,\beta} - \tfrac{\ln 2}{\beta|\Lambda|}\right]\\
&\leq \exp\left(-|\Lambda|\left[f^q_{\Lambda,\gamma,\beta} - f^a_{\Lambda,\gamma,\beta} - \tfrac{\ln 2}{\beta|\Lambda|}\right]^2\right)
\end{aligned} \tag{2.6}$$

where we used that by Jenssen's inequality, $\ln \mathbb{E}Z_{\Lambda,\gamma,\beta} \geq \mathbb{E}\ln Z_{\Lambda,\gamma,\beta}$. Comparing (2.5) with (2.6) we find

$$\left[f^q_{\Lambda,\gamma,\beta} - f^a_{\Lambda,\gamma,\beta} - \tfrac{\ln 2}{\beta|\Lambda|}\right]^2 \leq \beta\left[F^{\text{kac}}_{\gamma,\beta^2} + o(1)\right] \tag{2.7}$$

and so

$$\left|f^q_{\gamma,\beta} + \frac{\beta}{4}\right| \leq \sqrt{\beta F^{\text{kac}}_{\gamma,\beta^2}} \tag{2.8}$$

Since $F^{\text{kac}}_{\gamma,\beta^2}$ tends to zero with γ if $\beta \leq 1$, the claim of Theorem 1 follows.$\diamondsuit$

3. Proof of Theorem 2

Theorem 2 was proven in [B]. In that paper, we tried to give an elementary proof avoiding stochastic calculus as much as possible. Here we will, on the contrary, give a short proof using explicitly standard results from stochastic calculus (see e.g. [RY]), as was done in [CN,C].

The main idea is that the Gaussian couplings G_{ij} can be represented as Brownian motions $B_{ij}(t)$ at $t = 1$. We will denote by $\mathcal{F}_t$ the sigma-algebra generated by the random variables $\left(B_{i,j}(s); i,j \in \mathbb{Z}^d, s \in [0,t]\right)$.

Let us first see what becomes of the approach of Comets and Neveu when applied to our model. Let us introduce the t-dependent partition function

$$Z_{\Lambda,\gamma,\beta}(t) = I\!\!E_\sigma e^{\frac{\beta}{\sqrt{2}} \sum_{i,j\in\Lambda} \sqrt{J_\gamma(i-j)}\sigma_i\sigma_j B_{ij}(t)} \tag{3.1}$$

Not that by Brownian scaling $Z_{\Lambda,\gamma,\beta}(t)$ and $Z_{\Lambda,\gamma,\beta\sqrt{t}}$ have the same distribution. We set

$$\widehat{Z}_{\Lambda,\gamma,\beta}(t) \equiv \frac{Z_{\Lambda,\gamma,\beta}(t)}{I\!\!E Z_{\Lambda,\gamma,\beta}(t)} \tag{3.2}$$

It is easy to see that $\widehat{Z}_{\Lambda,\gamma,\beta}(t)$ is a martingale of mean value one. One can represent this martingale in the form

$$\widehat{Z}_{\Lambda,\gamma,\beta}(t) = \exp\left(\sqrt{|\Lambda|}M_\Lambda(t) - \frac{|\Lambda|}{2} < M_\Lambda(t) >\right) \tag{3.3}$$

where $M_\Lambda(t)$ is a martingale of mean zero and $< M_\Lambda(t) >$ is the unique increasing process that makes this formula correct [RY].

Note that for finite Λ, we might drop the $|\Lambda|$ terms in this formula, and indeed in [CN] they are not present. However, as we are interested in the limit $\Lambda \uparrow Z\!\!\!Z^d$, and the key point is the question whether the bracket $< M_\Lambda(t) >$ converges in that limit. In the SK-model, Comets and Neveu prove that this convergence takes place without putting the volume terms in the definition (3.2) (this constitutes the major part of their work). However, as we will see, for non-zero γ we cannot expect such a convergence without the normalization used in (3.2). This is the essential difference that arises in the Kac-SK model. Namely, while, as we will see, $< M_\Lambda(t) >$ converges to a a.s. constant, and while this implies that the martingale $M_\Lambda(t)$ converges to a centered Gaussian process, this has no direct application to the fluctuations of the free energy, since there is no obvious relation between the two quantities. While this looks disappointing at first sight, it will still be useful to look more closely at the bracket $< M_\Lambda(t) >$ and its limit. This may be motivated by the fact that this object has an immediate connection to the quenched and annealed free energies, namely an immediate computation reveals that

$$f^q_{\Lambda,\gamma,\beta} - f^a_{\Lambda,\gamma,\beta} = \frac{1}{2\beta} I\!\!E < M_\Lambda(1) > \tag{3.4}$$

Thus Theorem 1 gives us control over the behaviour of this bracket as γ tends to zero. Now, by standard computations totally analogous to those in [CN], we find that

$$\frac{d}{dt} < M_\Lambda(t) >= \frac{1}{|\Lambda|} \sum_{i,j\in\Lambda} \langle\langle \sigma_i\sigma'_i J_\gamma(i-j)\sigma_j\sigma'_j\rangle\rangle_{\Lambda,\gamma,\beta}(t) \tag{3.5}$$

where $\langle\langle\cdot\rangle\rangle_{\Lambda,\gamma,\beta}(t)$ denotes the expectation w.r.t. two copies of the Gibbs measure with identical couplings $B_{ij}(t)$. The point is that this quantity converges:

Lemma 3.1: *Assume that γ, β, t are such that almost surely, the infinite volume Gibbs state at temperature $\beta\sqrt{t}$ is unique. Then*

$$\lim_{\Lambda\uparrow \mathbb{Z}^d} \frac{1}{|\Lambda|} \sum_{i,j\in\Lambda} \langle\langle \sigma_i\sigma_i' J_\gamma(i-j)\sigma_j\sigma_j'\rangle\rangle_{\Lambda,\gamma,\beta}(t) = I\!E\langle\langle \sigma_0\sigma_0' \sum_{i\in\mathbb{Z}^d} J_\gamma(i)\sigma_i\sigma_i'\rangle\rangle_{\infty,\gamma,\beta\sqrt{t}}, \quad \textit{in Prob..} \tag{3.6}$$

Proof: Note that in distribution $\langle\langle\cdot\rangle\rangle_{\Lambda,\gamma,\beta}(t) = \langle\langle\cdot\rangle\rangle_{\Lambda,\gamma,\beta\sqrt{t}}$. Since the infinite volume Gibbs state is assumed to be unique, it is also translation covariant (and the same holds trivially for the replicated state), the convergence to a constant in probability is assured by the ergodic theorem. $\diamondsuit$

An immediate corollary we get the following representation for the difference between quenched an annealed free energies:

$$f^q_{\gamma,\beta} - f^a_{\gamma,\beta} = \frac{1}{2\beta}\int_0^1 I\!E\langle\langle \sigma_0\sigma_0' \sum_{i\in\mathbb{Z}^d} J_\gamma(i)\sigma_i\sigma_i'\rangle\rangle_{\infty,\gamma,\beta\sqrt{t}} \tag{3.7}$$

It remains to see how we can relate to the CLT for the free energy. This is based on the representation

$$f_{\Lambda,\gamma,\beta} - f^q_{\Lambda,\gamma,\beta} = \int_0^1 dI\!E\left(f_{\Lambda,\gamma\beta}|\mathcal{F}_t\right) \tag{3.8}$$

where

$$dI\!E\left[f_{\Lambda,\gamma\beta}|\mathcal{F}_t\right] = \frac{1}{|\Lambda|}\sum_{i,j\in\Lambda} dB_{ij}(t) I\!E\left[\langle \sigma_i\sigma_j\sqrt{J_\gamma(i-j)}\rangle_{\Lambda,\gamma,\beta}|\mathcal{F}_t\right] \tag{3.9}$$

which gives together with (3.8) a representation of the fluctuations of the free energy as a stochastic integral. Let us define the quadratic variation

$$F_{\Lambda,\gamma,\beta} = \int_0^1 \frac{1}{|\Lambda|^2}\sum_{i,j\in\Lambda} J_\gamma(i-j)\left(I\!E\left[\langle\sigma_i\sigma_j\rangle_{\Lambda,\gamma,\beta}|\mathcal{F}_t\right]\right)^2 \tag{3.10}$$

Then for any u

$$I\!E\exp\left(u\sqrt{|\Lambda|}\left(f_{\Lambda,\gamma\beta} - f^q_{\Lambda,\gamma\beta}\right) - |\Lambda|\frac{u^2}{2}F_{\Lambda,\gamma,\beta}\right) = 1 \tag{3.11}$$

(where the volume dependence was proven in view that we want to prove a CLT). We can make use of this formula, provided we can show that $|\Lambda|F_{\Lambda,\gamma,\beta}$ converges as $\Lambda\uparrow\mathbb{Z}^D$. But this is precisely the case, and for precisely the same reasons as in Lemma 3.1! Moreover, by the Schwartz inequality,

$$|\Lambda|F_{\Lambda,\gamma,\beta} \leq int_0^1\frac{1}{|\Lambda|}\sum_{i,j\in\Lambda} J_\gamma(i-j) I\!E\left[\langle\langle\sigma_i\sigma_j\sigma_i'\sigma_j'\rangle\rangle_{\Lambda,\gamma,\beta}|\mathcal{F}_t\right] \tag{3.12}$$

let us collect these observations:

Lemma 3.2: *Assume that γ, β are such that almost surely, the infinite volume Gibbs state is weakly unique. Then there exists a constant $c_{\gamma,\beta}$ such that*

$$\lim_{\Lambda\uparrow\mathbb{Z}^d} |\Lambda|F_{\Lambda,\gamma,\beta} = c_{\gamma,\beta}, \quad \textit{in Prob..} \tag{3.13}$$

Moreover,

$$c_{\gamma,\beta} \leq \sum_{i\in\Lambda} J_\gamma(i) \mathbb{E}\langle\langle\sigma_0\sigma_0'\sigma_i\sigma_i'\rangle\rangle_{\infty,\gamma,\beta} \tag{3.14}$$

A a corollary, it follows that for any u

$$\lim_{\Lambda\uparrow\mathbb{Z}^d} \mathbb{E} \exp\left(u\sqrt{|\Lambda|}\left(f_{\Lambda,\gamma\beta} - f^q_{\Lambda,\gamma\beta}\right)\right) = \exp\left(\frac{u^2}{2} c_{\gamma,\beta}\right) \tag{3.15}$$

which implies the first assertion of Theorem 2 (the CLT).

To prove the second assertion, it suffices to note that the upper bound (3.14) together with (3.7) allows to relate $c_{\gamma,\beta}$ to the difference between the quenched and the annealed free energies. Namely,

$$\int_0^1 c_{\gamma,\beta\sqrt{t}} dt \leq 2\beta\left(f^q_{\gamma,\beta} - f^a_{\gamma,\beta}\right) \tag{3.16}$$

From here the second assertion of Theorem 2 follows immediately. $\diamondsuit\diamondsuit$

Acknowledgements: I would like to thank Peter Eichelsbacher and Mattias Löwe for organizing this nice workshop and for encouraging the writing of these notes.

References

[ALR] M. Aizenman, J.L. Lebowitz, and D. Ruelle, "Some rigorous results on the Sherrington-Kirkpatrick spin glass model". Commun. Math. Phys. **112**, 3-20 (1987).

[B] A. Bovier, "The Kac version of the Sherrington Kirkpatrick model at high temperatures", to appear in J. Stat. Phys. **91** (1998).

[BGP1] A. Bovier, V. Gayrard and P. Picco),"Large deviation principles for the Hopfield and the Kac-Hopfield model", Probab. Theory Rel. Fields **101**, 511-546 (1995).

[BGP2] A. Bovier, V. Gayrard and P. Picco, "Distribution of overlap profiles in the one-dimenisonal Kac-Hopfield model", Commun. Math. Phys. **186**, 323-379 (1997).

BGP3] A. Bovier, V. Gayrard and P. Picco, "Typical profiles of the Kac-Hopfield model", in "Mathematical aspects of spin glasses and neural networks", A. Bovier and P. Picco (eds.), Progress in Probability **41**, (Birkhäuser, Boston-Basel-Berlin 1998).

[CN] F. Comets and J. Neveu, "The Sherrington-Kirkpatrick model of spin glasses and stochastic calculus, the high temperature case", Commun. Math. Phys. **166**, 549-564 (1995).

[Co] F. Comets, "The martingale method for mean-field disordered systems at high temperatures", in "Mathematical aspect of spin glasses and neural networks", A. Bovier and P. Picco (Eds.), Progress in Probab. **41**, (Birkhäuser, Basel-Boston, 1998).

[FZ] J. Fröhlich and B. Zegarlinski, "Some comments on the Sherrington-Kirkpatrick model of spin glasses", Commun. Math. Phys. **112**, 553-566 (1987).

[FZ2] J. Fröhlich and B. Zegarlinski, "The high-temperature phase of long-range spin glasses", Commun. Math. Phys. **110**, 547-560 (1982).

[GR] S. Goulart-Rosa, "The thermodynamic limit of quenched free energy of magnetic systems with random interactions, J. Phys. **A 15**, L51-L54 (1982).

[KUH] M. Kac, G. Uhlenbeck, and P.C. Hemmer, "On the van der Waals theory of vapour-liquid equilibrium. I. Discussion of a one-dimensional model" J. Math. Phys. **4**, 216-228 (1963); "II. Discussion of the distribution functions" J. Math. Phys. **4**, 229-247 (1963); "III. Discussion of the critical region", J. Math. Phys. **5**, 60-74 (1964).

[LT] M. Ledoux and M. Talagrand, "Probability in Banach spaces", (Springer, Berlin-Heidelberg-New York, 1991).

[NS1] Ch.M. Newman and D.L. Stein, "Non-mean-field behaviour in realistic spin glasses", Phys. Rev. Lett. **76**, 515-518 (1996).

[NS2] Ch.M. Newman and D.L. Stein, "Thermodynamic chaos and the structure of short range spin glasses", in "Mathematical aspects of spin glasses and neural networks", A. Bovier and P. Picco (Eds.), Progress in Probability Vol. 41, (Birkhäuser, Basel-Boston, 1997).

[RY] D. Revuz and M. Yor, "Brownian motion and continuous martingales", (Springer, Berlin-Heidelberg-New York, 1992).

[SK] D. Sherrington and S. Kirkpatrick, "Solvable model of a spin glass", Phys. Rev. Lett. **35**, 1792-1796 (1972).

[T1] M. Talagrand, "Concentration of measure and isoperimetric inequalities in product space", Publ. Math. I.H.E.S., **81**, 73-205 (1995).

[T2] M. Talagrand, "The Sherrington-Kirkpatrick model: A challenge for mathematicians", preprint 1996, to appear in Prob. Theor. Rel. Fields.

Julia Brettschneider

Shannon–Mac Millan Type Theorems for Lattice Models

1 Information and entropy for random fields

Consider $\Omega = \Upsilon^{\mathbb{Z}^2}$, where Υ is a finite state space. Let P be a random field, that is a probability measure on Ω. Given any $v \in \mathbb{Z}^2$, a transformation on Ω is induced by $\theta_v\omega(u) = \omega(u+v)$ $(u \in \mathbb{Z}^2)$. For any subset V of $\mathbb{Z}^2$ define $\Omega_V := \Upsilon^V$, ω_V as the projection of ω on V, $\mathcal{F}_V := \sigma(\omega_V)$ and P_V as the distribution of ω_V w.r.t. P.

The *information* of P restricted to V ist given by

$$\mathcal{I}(P_V)(\omega) = -\log P(\omega_V),$$

and the information of P_V conditioned on $\mathcal{F}_W$ for some $W \subset \mathbb{Z}^2$ is of the form

$$\mathcal{I}\big(P_V(\cdot|\mathcal{F}_W)(\omega)\big) = -\log P(\omega_V|\,\omega_W).$$

Integration leads to these formulae for the corresponding entropies:

$$\begin{aligned} H(P_V) &= -E[\log P(\omega_V)] \\ H(P_V(\,\cdot\,|\mathcal{F}_W)) &= -E[\log P(\omega_V|\,\omega_W)]. \end{aligned}$$

For $n \in \mathbb{N}$ let V_n denote the set of all lattice points which are contained in a box of side length n around the origin and V^- those which are smaller than 0 w.r.t. the lexicographical ordering on $\mathbb{Z}^2$. In what follows, we assume P to be stationary, i.e. invariant w.r.t. θ_v $(v \in \mathbb{Z}^2)$. The *specific entropy* of P is given by

$$h(P) = \lim_{n\to\infty} \frac{1}{|\,V_n|} H(P_{V_n}). \tag{1}$$

Its existence is assured by a subadditivity property. The Shannon-Mac Millan theorem leads to the representation formula (see [Fö73])

$$h(P) = E[H\big(P_0(\cdot|\mathcal{F}_{V^-})(\omega)\big)],$$

where $H\big(P_0(\cdot|\mathcal{F}_{V^-})(\omega)\big)$ denotes the entropy of the probability measure $P_0(\cdot|\mathcal{F}_{V^-})(\omega)$ on Υ. It states further that there is a $\mathcal{L}^1$-convergence of the sequence of rescaled information behind the limit in (1), i.e.

$$\frac{1}{|\,V_n|}\mathcal{I}(P_{V_n}) \longrightarrow E\big[H\big(P_0(\cdot|\mathcal{F}_{V^-})(\omega)\big)\big|\mathcal{J}\big] \qquad \text{in } \mathcal{L}^1(P), \tag{2}$$

where $\mathcal{J}$ is the σ-algebra of all sets which are invariant w.r.t. the transformations θ_v $(v \in \mathbb{Z}^2)$. In the ergodic case the limit in (2) turns out to be $h(P)$.

Our aim is to derive theorems of this type, but for objects of surface order instead of boxes. This question is motivated by the investigation of large deviations of Gibbs measures in the case of phase transitions, where surface order terms come into play (see [Olla88], [Schon87], [FöOrt88], [DKS92], [Ioffe94] and [Ioffe95]).

2 A Shannon Mac-Millan type theorem along a line

Our purpose is to construct a *specific surface entropy,* to be denoted by $h_\lambda(P)$, of a random field P along the lattice approximation of a line with slope λ. Furthermore, we are interested in obtaining a Shannon-Mac Millan type theorem and a corresponding representation formula for this entropy. In the case of the coordinate axis, this is carried out in [FöOrt88]. We want to identify $h_\lambda(P)$ as the limit of the rescaled entropies along successively increased parts of the lattice approximation of the line. Furthermore we investigate the $\mathcal{L}^1$-convergence of the corresponding sequence of information.

$$\begin{aligned} l_{\lambda,a} : \ \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \lambda x + a \end{aligned}$$

describes the line with slope λ and y-intercept a. Using $[x]$ for the integer and $\{x\}$ for the fractional part of $x \in \mathbb{R}$, the bilateral sequences

$$\big([l_{\lambda,a}(z)]\big)_{z\in\mathbb{Z}} \quad \text{and} \quad \big(\{l_{\lambda,a}(z)\}\big)_{z\in\mathbb{Z}}$$

are, at any step z, the line's integer and fractional parts, respectively. Furthermore, for $a \in \mathbb{R}$ and $I \subset \mathbb{R}$ define $L_\lambda(a, I) := \big\{(z, [l_{\lambda,a}(z)]) \,|\, z \in I \cap \mathbb{Z}\big\}$.

Every rational λ can be represented uniquely by $\lambda = \frac{p}{q}$ with $p \in \mathbb{Z}$ and $q \in \mathbb{N}$ that have no common divisor. In these terms we can formulate the first

Theorem 2.1
If P is invariant w.r.t. $\theta_{\binom{q}{p}}$ and $H(P_0)$ is finite then for all $a \in \mathbb{R}$,

$$\frac{1}{n+1}\mathcal{I}(P_{L_\lambda(a,[0,n])}) \longrightarrow \frac{1}{q}\sum_{i=0}^{q-1} E\big[H\big(P_0(\cdot|\mathcal{F}_{L_\lambda(\{a+i\lambda\},(-\infty,-1])})(\omega)\big)\big|\mathcal{J}\big],$$

P-a.s. and in $\mathcal{L}^1(P)$, where $\mathcal{J}$ is the σ-algebra of all sets which are invariant w.r.t. the transformation $\theta_{\binom{q}{p}}$. In particular, if P is ergodic w.r.t. $\theta_{\binom{q}{p}}$, then

$$\lim_{n\to\infty}\frac{1}{n+1}\mathcal{I}(P_{L_\lambda(a,[0,n])}) = \frac{1}{q}\sum_{i=0}^{q-1} E\big[H\big(P_0(\cdot|\mathcal{F}_{L_\lambda(\{a+i\lambda\},(-\infty,-1])})(\omega)\big)\big] \qquad \forall\, a \in \mathbb{R}.$$

Of course, by translation invariance, the theorem could be stated for other suitable sequences instead of $([0,n])_{n\in\mathbb{N}}$.

In the case of an irrational slope, the above method fails, because of the lack of periodicity. In the context of a directional entropy for cellular automata, a comparable problem occurred when [Mil86] introduced a (volume order) directional entropy and, by this, aroused a debate on the question whether it is continuous w.r.t. the directional parameter (see [Sinai85], [Park95] and [Park96]).

For our purpose, we have to focus on the information rather than on the entropy. In order to make the problem accessible for methods of ergodic theory, we need to find a transformation that follows a stair climbing pattern along the lattice approximation of the line. This will be achieved by memorizing not only the integer part but also the fractional part in each step. Let $\mathbb{T}$ be the interval $[0,1)$ considered as a torus and equipped with the Borel σ-algebra and the Lebesgue measure μ. Then the translation by λ modulo 1, $t \mapsto \{t+\lambda\}$, defines an invariant transformation on $\mathbb{T}$, which is ergodic for every irrational λ. Consider the product space $\mathbb{T} \times \Omega$, equipped with the product σ-algebra and the product measure $\overline{P} = \mu \otimes P$. The transformation

$$\begin{aligned} S: \quad \mathbb{T} \times \Omega &\longrightarrow \mathbb{T} \times \Omega \\ (t, \omega) &\longmapsto (\{t+\lambda\}, \theta_{(1,[t+\lambda])}\omega) \end{aligned}$$

turns out to follow the desired path, as we see in the following

Lemma 2.2

$$S^n(a, \omega) = \big(\{t+n\lambda\}, \theta_{(n,[l_a(n)])}\omega\big) \qquad \forall\ n \in \mathbb{N}_0.$$

This leads to

Theorem 2.3
Let $\lambda \in \mathbb{R} \setminus \mathbb{Q}$, P be stationary and $H(P_0)$ be finite. Then for all $a \in \mathbb{R}$,

$$\frac{1}{n+1}\mathcal{I}(P_{L_\lambda(a,[0,n])}) \longrightarrow \overline{E}\big[H\big(P_0(\cdot|\mathcal{F}_{L_\lambda(t,(-\infty,-1])})(\omega)\big)\big|\mathcal{J}\big],$$

P-a.s. and in $\mathcal{L}^1(P)$, where $\mathcal{J}$ is the σ-algebra of all sets which are invariant w.r.t. the transformation S.

In order to achieve a nicer expression for the limit, we would like S to be ergodic. This question will be discussed in a more general setting in the next section.

Corollary 2.4
Assuming the same as in Theorem 2.3 *and further that S is ergodic, we obtain*

$$\lim_{n\to\infty} \frac{1}{n+1}\mathcal{I}(P_{L_\lambda(a,[0,n])}) = \int_0^1 E\big[H\big(P_0(\cdot|\mathcal{F}_{L_\lambda(t,(-\infty,-1])})(\omega)\big)\big]\, dt,$$

in P-a.s. and $\mathcal{L}^1(P)$.

3 A class of skew products

Let α be a measure preserving transformation on the probability space $(M, \mathcal{B}, \mu)$ and ϑ a measure preserving transformation on the probability space $(\Omega, \mathcal{F}, P)$. The function κ is assumed to be a $\mathcal{B}$-measurable function on M with values in $\mathbb{N}_0$. Being composed, those ingredients lead to the following skew product on the product space $(M \times \Omega, \mathcal{B} \otimes \mathcal{F}, \mu \otimes P)$

$$\begin{aligned} S: \quad M \times \Omega \quad &\longrightarrow M \times \Omega \\ (t, \omega) \quad &\longmapsto \big(\alpha(t), \vartheta^{\kappa(t)}\omega\big). \end{aligned}$$

So κ, depending on t, controls the action of ϑ. Setting $\kappa \equiv 1$ we obtain an uncoupled product. Obviously, S preserves the product measure $\overline{P} = \mu \otimes P$ and we obtain for its iterates

$$S^n(t, \omega) = \big(\alpha^n(t), \theta^{\kappa_n(t)}\omega\big), \qquad \text{where} \qquad \kappa_n = \sum_{i=0}^{n-1} \kappa \circ \alpha^i. \tag{3}$$

We are now looking for assumptions on the single transformations that assure the ergodicity of S w.r.t. $\overline{P}$. Having in mind the application discussed in the previous section, we want to assume that α is ergodic only, but we would be more flexible with mixing conditions for ϑ. Of course, remembering the case of an uncoupled product, we would need ϑ to be at least weakly mixing. In our case, we would need to take κ additionally into consideration.

Definition 3.1
ϑ is called weakly mixing along the sequence $(k_n)_{n \in \mathbb{N}}$, if

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \big| P\big(A \cap \theta^{-\kappa_i} B\big) - P\big(A\big)\, P\big(B\big) \big| = 0 \quad \forall A, B \in \mathcal{F}. \tag{4}$$

Theorem 3.2
If α is ergodic w.r.t. μ and ϑ is weakly mixing along $(\kappa_n(t))_{n \in \mathbb{N}}$ for μ–f.a. $t \in M$. Then S is ergodic w.r.t. $\overline{P}$.

Given those assumptions, we could not expect S to be more than ergodic, because, by projection, any further mixing condition would be carried over to α. Under slightly stronger assumptions we can derive a corollary which is more convenient to use. If ϑ is *strongly mixing* w.r.t. P, i.e.

$$\lim_{n \to \infty} P\big(A \cap \vartheta^{-n} B\big) = P\big(A\big)\, P\big(B\big) \quad \forall A, B \in \mathcal{F},$$

this property translates to all sequences $(k_n)_{n \in \mathbb{N}} \subset \mathbb{N}_0$ which tend to ∞ for $n \to \infty$, i.e.

$$\lim_{n \to \infty} P\big(A \cap \vartheta^{-k_n} B\big) = P\big(A\big)\, P\big(B\big) \quad \forall A, B \in \mathcal{F}.$$

In particular, we get (4), and thus the following

Corollary 3.3
If α is ergodic w.r.t. μ, ϑ strongly mixing w.r.t. P and $\kappa_n(t)$ tends to ∞ for μ-a.a. $t \in M$. Then S is ergodic w.r.t. $\overline{P}$.

Literature:

- [DKS92] R. L. Dobrushin, R. Kotecký, S. Shlosman: *Wulff construction: a global shape from local interaction,* AMS, Providence, 1992
- [Fö73] H. Föllmer: *On entropy and information gain in random fields,* Z. Wahrscheinlichkeitstheorie verw. Gebiete 26, 207-217 (1973).
- [FöOrt88] H. Föllmer, M. Ort: *Large deviations and surface entropy for Markov fields,* Astérisque 157-158 (1988)
- [Ioffe94] D. Ioffe: *Large deviations for the 2D Ising model: a lower bound without cluster expansion,* J. Stat. Phys. 74 (1994)
- [Ioffe95] D. Ioffe: *Exact large deviation bounds up to T_c for the Ising model in two dimensions,* Probab. Theory Relat. Fields 102, No.3 (1995)
- [Mil86] J. Milnor: *Directional entropies of cellular automaton-maps,* Disordered systems and biological organization (Les Houches, 1985), 113–115, NATO Adv. Sci. Inst. Ser. F: Comput. Systems Sci., 20, Springer, Berlin-New York, 1986.
- [Olla88] S. Olla: *Large deviations for Gibbs random fields,* Probab. Th. Rel. Fields 77 (1988)
- [Park95] Park, K.: *Continuity of directional entropy for a class of Z^2-actions,* J. Korean Math. Soc. 32, No.3, 573-582 (1995)
- [Park96] Park, K.: *Entropy of a skew product with a Z^2-action,* Pac. J. Math. 172, No.1, 227-241 (1996)
- [Schon87] R. Schonmann: *Second order large deviation estimates for ferromagnetic systems in the phase coexistence region,* Comm. Math. Phys. 112 (1987), no. 3
- [Sinai85] Ya. G. Sinai: *An answer to a question of J. Milnor,* Comment. Math. Helvetici 60 (1985)

RELAXATION OF DISORDERED SYSTEMS OF ROTORS

IN THE GRIFFITHS' REGIME

Pietro Caputo

Techinsche Universität Berlin

The material presented in this note is part of a joint work with Fabio Martinelli and Filippo Cesi, [3]. We consider bounded continuous spins on a lattice, such as the *plane rotor* (or XY) model and the Heisenberg model. The interaction is given by a *random* nearest neighbour potential, and the dynamics is an infinite dimensional diffusion process with reversible measure given by the Gibbs equilibrium state. We study the relaxation of the system in the so-called *Griffiths' regime*. In the case of discrete (Ising-type) spins these questions were recently answered by the works [4, 5], where the authors obtain essentially optimal bounds on the deacy to equilibrium of a single spin-flip Glauber dynamics. Much less advanced is the corresponding analysis for continuous spins. Previous results in this direction can be found in [9, 10].

1 The Model

For the purposes of this note we confine ourselves to the *plane rotor* model. To each site $x \in \mathbb{Z}^d$ we assign a spin variable $\boldsymbol{\sigma}_x \in S^1$, where S^1 is the unit circle in $\mathbf{R}^2$. We consider nearest neighbour random interaction: to each pair of nearest neighbour sites x, y, denoted by $\langle xy \rangle$, we assign a real-valued random variable J_{xy}. The couplings J_{xy} are assumed to be indipendent and identically distributed, and we call $\mathbf{P}$ their distribution. Further assumuptions on $\mathbf{P}$ will be described later.

Given a finite set $\Lambda \subset \mathbb{Z}^d$, for a fixed realization J of the couplings, the Hamiltonian is given by

$$H_\Lambda^{J,\tau} = -\sum_{\langle xy \rangle \in \Lambda^*} J_{xy}\, \boldsymbol{\sigma}_x \cdot \boldsymbol{\sigma}_y - \sum_{\langle xy \rangle \in \partial\Lambda} J_{xy} \boldsymbol{\sigma}_x \cdot \boldsymbol{\tau}_y \tag{1}$$

where '$\cdot$' denotes the scalar product, τ is a boundary condition, Λ^* denotes the set of bonds $\langle xy \rangle$ such that $x, y \in \Lambda$, and $\partial\Lambda$ is the boundary of Λ, i.e. the set of $\langle xy \rangle$ with $x \in \Lambda$ and $y \notin \Lambda$. As usal, one defines the finite volume Gibbs measures, or local specifications, by

$$d\mu_\Lambda^{J,\tau}(\sigma) = \begin{cases} (Z_\Lambda^{J,\tau})^{-1} \mathrm{e}^{-H_\Lambda^{J,\tau}(\sigma)}\, \nu_\Lambda(d\sigma_\Lambda) & \text{if } \sigma_x = \tau_x \; \forall x \notin \Lambda \\ 0 & \text{otherweise} \end{cases} \tag{2}$$

where ν_Λ is Lebesgue measure on the product $(S^1)^\Lambda$, and $Z_\Lambda^{J,\tau}$ is the normalizing factor, or partition function. We define a probability measure μ on $\Omega \equiv (S^1)^{\mathbf{Z}^d}$ (equipped with the product topology) to be *Gibbs*, if the *DLR* conditions are satisfied. That is if, for all local functions f (i.e. f only depends on finitely many spins $\boldsymbol{\sigma}_x$), and for all finite Λ,

$$\mu(\mu_\Lambda^{J,\cdot}(f)) = \mu(f). \tag{3}$$

To define the dynamics on finite volumes, for a fixed realization J of the disorder, consider the differential operator

$$L_\Lambda^{J,\tau}\phi \;=\; \sum_{x\in\Lambda} \Delta_x \phi \;-\; \nabla_x H_\Lambda^{J,\tau} \cdot \nabla_x \phi, \qquad \phi \in \mathcal{C}^\infty(\Omega). \tag{4}$$

The closure of $L_\Lambda^{J,\tau}$ in $L^2(\mu_\Lambda^{J,\tau})$ is a self adjoint operator, which we denote again by $L_\Lambda^{J,\tau}$. The semigroup $S_\Lambda^{J,\tau}(t) = \exp tL_\Lambda^{J,\tau}$ is the Markov semigroup of a finite-dimensional diffusion process whose unique invariant measure is given by $\mu_\Lambda^{J,\tau}$. In a similar way one can construct the infinite volume semigroup $S^J(t)$, see e.g. [9]. With $\mathbf{P}$-probability one $S^J(t)$ is a well defined Markov semigroup, with the following interpretation: if σ_t denotes the configuration of the system at time t, then

$$S^J(t)f(\sigma) \;=\; \mathrm{E}_\sigma^J[f(\sigma_t)], \quad f \in \mathcal{C}(\Omega) \tag{5}$$

where E_σ denotes expectation with respect to the process starting in the configuration σ. Moreover, all (infinite-volume) Gibbs measures are invariant measures for this process. The dynamics is said to be ergodic if it has a unique invariant measure and and if it converges to it for any distribution of the initial configuration.

2 Main Hypothesis and Results

In order to state our main hypothesis, we need the following definition. For a local function f, we call S_f its support, i.e. the set of sites on which f depends non-trivially, and denote by $|S_f|$ its cardinality. By $d(f,g)$ we denote the distance $d(S_f, S_g)$ between the supports of two functions f and g, w.r.t. the ℓ_1 lattice distance. Let $s, \alpha > 0$ and $\Lambda \subset \mathbb{Z}^d$. We say that the condition $SMT(\Lambda, s, \alpha)$ holds if, for all local bounded functions f, g, with $S_f, S_g \subset \Lambda$ and $d(f,g) > s$,

$$\sup_{\tau\in\Omega} |\mu_\Lambda^{J,\tau}(fg) \;-\; \mu_\Lambda^{J,\tau}(f)\mu_\Lambda^{J,\tau}(g)| \;\le\; 2|S_f|\,|S_g|\,||f||_\infty\,||g||_\infty\, \mathrm{e}^{-\alpha\, d(f,g)}. \tag{6}$$

Let Q_L denote the cube of side L, $Q_L \equiv [0, L-1]^d \cap \mathbb{Z}^d$. Then our main assumptions on the distribution $\mathbf{P}$ of the disorder are

(H1) There exists $L_0 \in \mathbb{Z}^+$, $\theta > 0$ and $\alpha > 0$, such that for all $L > L_0$,

$$\mathbf{P}\Big\{ SMT(Q_L, L/2, \alpha) \text{ holds} \Big\} \;\ge\; 1 \;-\; \mathrm{e}^{-\theta L}. \tag{7}$$

(H2) There exists $J_\infty \in (0,\infty)$ such that

$$\mathbf{P}\{|J_{xy}| > J_\infty\} \;=\; 0 \tag{8}$$

Remark 1. Assumption (H1) has been shown to be a very natural tool in analysing the Griffiths' regime of disordered spin systems ([4]). It is a general condition which can be seen to be implied by the usual high temperature constraints. In the case one adds a random field to the system, one can show that suitable strong field constraints imply condition (H1), see [6]. In particular all the results we derive will hold in this strong field regime.

Remark 2. The assumption of bounded couplings (H2) can be considerably relaxed (see [4]), and our results can be generalized to the case of couplings satisfying a suitable tail condition (stronger than finite exponential moments).

2.1 Upper Bounds

We show that hypothesis (H1) and (H2) are sufficient to have $\mathbf{P}$-almost sure uniqueness of the Gibbs measure and ergodicity of the semigroup $S^J(t)$. In particular, we prove the following upper bounds on both almost sure (quenched) and disorder-averaged decay.

Theorem 1 *Assume (H1) and (H2).*

(i) for $d \geq 1$, there exists a set of couplings $\boldsymbol{J_0}$, with $\mathbf{P}(\boldsymbol{J_0}) = 1$, such that for all $J \in \boldsymbol{J_0}$ there is a unique Gibbs measure μ^J. Moreover, for all $J \in \boldsymbol{J_0}$ and for all local functions f, there exist constants $t_0 = t_0(J, f) \in (0, \infty)$ and $k_1 = k_1(d, J_\infty) \in (0, \infty)$ such that, for all $t \geq t_0$

$$||S^J(t)\phi \; - \; \mu^J(\phi)||_\infty \; \leq \exp\left(-t \; \exp\left[-k_1(\log t)^{\frac{d-1}{d}}(\log\log t)^{d-1}\right]\right), \tag{9}$$

ii) for $d \geq 2$, for all local functions f, there exist constants $t_0(f) \in (0, \infty)$ and $k_2 = k_2(d, J_\infty) \in (0, \infty)$ such that for any $t \geq t_0$,

$$\mathrm{E}||S^J(t)\phi \; - \; \mu^J(\phi)||_\infty \; \leq \exp\left[-k_2(\log t)^{\frac{d}{d-1}}(\log\log t)^{-d}\right], \tag{10}$$

where E *denotes expectation w.r.t.* $\mathbf{P}$.

Remark 3. The proof of the above estimates is based on the use of spectral gap and logarithmic Sobolev inequalities. It follows essentially the ideas introduced in [4], see also [11] for a review of the main techniques.

Remark 4. In the case of Ising spins, the estimates (9) and (10) describe, apart from the technical factor $\log\log t$, the correct speed of relaxation ([4]). On the other hand, they are believed to be far from optimal in our case. For Heisenberg spins in dimension $d = 3$, for instance, the disorder-averaged decay is predicted to be of order $\exp(-At^{1/2})$ (see [1, 2] for a heuristic derivation). The main difficulty in deriving the correct upper bounds seems to be in the understanding of the mechanism of relaxation of strongly correlated regions. While the low-temperature behaviour of Ising-type Glauber dynamics is well understood ([11]), the corresponding analysis for the diffusion dynamics is much less advanced.

2.2 Lower Bounds

We present here a lower bound on the deacy to equilibrium of the two-dimensional plane rotor model. Consider the *pure* model, i.e. the case where the couplings J_{xy} have all the same value $\beta > 0$. It is well known that for β large enough this model is in a massless phase (so-called Kosterlitz-Thouless phase), where spin-spin (static) correlations have a power-law decay, see [8]. Let $\beta_1 > 0$ be such that for all $\beta < \beta_1$ the pure system is in the regime of *complete analyticity* ([7]).

Theorem 2 *Let $d = 2$. Suppose (H1) and (H2) are satisfied. Suppose moreover there exist $p_1 > 0$ and $p_0 > 0$ such that*

$$\mathbf{P}\{\, J_{xy} > \beta_0 \,\} = p_1$$
$$\mathbf{P}\{\, J_{xy} \in [0, \beta_1] \,\} = p_0,$$

where β_0 is such that the pure model at inverse temperature β_0 is in the massless phase. Then there exists a constant k depending on p_0, p_1, such that for t large enough we have

$$\mathrm{E}\,||S^J(t)\boldsymbol{\sigma}_0 \;-\; \mu^J(\boldsymbol{\sigma}_0)||_{L^2(\mu^J)} \;\geq \exp\left(-kt^{1/2+\rho}\right), \tag{11}$$

where $\boldsymbol{\sigma}_0$ denotes the spin at the origin, and ρ is a small parameter depending on β_0, with $\rho \to 0$, as $\beta_0 \to \infty$.

Remark 3. In the above Theorem we estimate from below the decay in the L^2 norm, which clearly gives a lower bound in the uniform norm. The main tool in the proof is the use of correlation inequalities to obtain upper estimates on the spectral gap of the generator of the process, as in [9]. This method can be used to prove a lower bound of the order $\exp(-Bt^{1/2})$ in the case $d = 3$, thereby confirming the predictions of [1, 2], see [3].

References

[1] A. J. Bray, Dynamics of dilute magnets above T_c, *Phys. Rev. Lett.*, **60**, 1988

[2] A. J. Bray, Upper and lower bounds on dynamic correlations in the Griffiths phase, *J. Phys. A: Math. Gen.*, **22**, 1989

[3] F. Cesi, P. Caputo, F. Martinelli, Relaxation of Disordered Systems of Rotors in the Griffiths' Regime, *in preparation*

[4] F. Cesi, C. Maes, F. Martinelli, Relaxation of Disordered Magnets in the Griffiths Regime, *Comm. Math. Phys.*, to appear.

[5] F. Cesi, C. Maes, F. Martinelli, Relaxation to Equilibrium for two dimensional Disordered Ising Systems in the Griffiths Phase, *Comm. Math. Phys.*, to appear.

[6] H. von Dreyfus, A. Klein, J. F. Perez, Taming Griffiths' Singularities: infinite Differentiability of quenched Correlation Functions, *Comm. Math. Phys.*, **170**, 1995

[7] R.L. Dobrushin, S.B. Shlosman, Completely analytical Interactions, Constructive Descritpion. *J. Stat. Phys.* **46**, 1987

[8] J. Fröhlich, T. Spencer, Kosterlitz-Thouless transition in 2-D XY model, *Comm. Math. Phys.* **81**, 1981.

[9] A. Guionnet, B. Zegarlinski, Decay to Equilibrium in Random spin systems on a lattice I, *Comm. Math. Phys.*, **181**, No. 3, 1996

[10] A. Guionnet, B. Zegarlinski, Decay to Equilibrium in Random spin systems on a lattice II, *J. Stat. Phys.* **86**, No. 3-4, 1997

[11] F. Martinelli, Lecture Notes on Glauber Dynamics, Ecole de Probabilite St. Flour, 1997

Large and Moderate Deviations for the Local Time of a Recurrent Markov Chain on $\mathbb{Z}^2$

N. Gantert and O. Zeitouni

Let (X_n) be a recurrent Markov chain on $\mathbb{Z}^2$ with $X_0 = (0,0)$, and let $g(n) := \sum_{k=0}^{n} P[X_k = (0,0)]$ be the truncated Green function. We can extend g to a continuous, increasing function $g(t), t \geq 0$. Since (X_n) is recurrent, $g(t) \to \infty$ for $t \to \infty$.

We will assume throughout that, for some positive constant C,

$$P[X_k = (0,0)] \leq \frac{C}{k}\,, \tag{1}$$

hence $g(n) \leq C \log n$. We will also assume throughout that

$$g \text{ is slowly varying at } \infty\,, \tag{2}$$

that is $g(tx)/g(x) \underset{t\to\infty}{\longrightarrow} 1$ for any $x > 0$. Note that (1) is satisfied for symmetric random walks on $\mathbb{Z}^2$, i.e. if $P[X_1 = (y,z)] = P[X_1 = -(y,z)]$, see [6, Proposition 2.14], while (2) holds for a random walk if X_1 is in the domain of attraction of an $\mathbb{R}^2$-valued Gaussian variable. Since our results depend only on (1) and (2), they also apply to symmetric recurrent random walk on $\mathbb{Z}$ in the domain of attraction of a Cauchy random variable.

We denote by L_n^0 the local time of X at $(0,0)$, i.e. $L_n^0 := |\{0 \leq k \leq n : X_k = (0,0)\}|$, and $L_0^0 = 0$. Let $\rho_0 = 0$, $\rho_k = \min\{j : j > \rho_{k-1}, X_j = (0,0)\}$, $k = 1,2,3,\cdots$. It is known, see [6], (and will follow from the proof of Theorem 1), that $L_n^0/g(n)$ converges in distribution to an exponential distribution, i.e.

$$P\left[\frac{L_n^0}{g(n)} \geq y\right] \underset{n\to\infty}{\longrightarrow} e^{-y}\,. \tag{3}$$

Our goal is to investigate the fluctuations of L_n^0, and associated functional laws.

Theorem 1 (Moderate Deviations) *Let $\psi(n)$ be a positive function such that $\psi(n) \underset{n\to\infty}{\longrightarrow} \infty$ and*

$$\gamma_n := \frac{n}{\psi(n)g(n)} \underset{n\to\infty}{\longrightarrow} \infty\,.$$

Then $L_n^0/g(n)\psi(n)$ satisfies a large deviation principle with speed $\psi(n)g(n)/g(\gamma_n)$ and rate function $I(y) = y$.

Theorem 1 is a moderate deviation principle since the speed can vary without changing the rate function I. Further, the rate function does not depend on the distribution of ρ_1.

The next theorem gives a large deviation principle for the distributions of $\frac{L_n^0}{n}$, with rate function which does depend on the distribution of ρ_1.

Theorem 2 (Large Deviations) *Let* $\Lambda^*(y) = \sup_{\lambda\le 0}(\lambda y - \log E[e^{\lambda\rho_1}])$ *and*

$$J(y) = \begin{cases} y\Lambda^*\left(\frac{1}{y}\right), & 0 < y \le 1 \\ 0, & y = 0 \\ +\infty, & \textit{otherwise} \end{cases}$$

Then the distributions of $\frac{L_n^0}{n}$ *satisfy a LDP with speed* n *and rate function* J.

Remarks

1. Comparing with Theorem 1, the large deviation principle holds for $\psi(n) = \frac{n}{g(n)}$. In this case, $\gamma_n = 1$ and Theorem 1 does not apply. Considering the proof of Theorem 2, it is easy to show that we have a LDP whenever $\gamma_n \underset{n\to\infty}{\longrightarrow} \alpha$, $0 < \alpha < 1$.

2. Let $p_0 := P[X_1 = (0,0)]$. Then we have $J(1) = -\log p_0$ if $p_0 > 0$ and $J(1) = \infty$ otherwise.

As usual, we can derive an Erdös-Renyi law from the large deviation principle:

Corollary 1 *Let* $c > 0$ *and* $\eta_{n,j} := \frac{1}{c\log g(n)}\left(L^0_{j+\lfloor c\log g(n)\rfloor} - L^0_j\right)$, $j = 0,1,2,\cdots,n-\lfloor c\log g(n)\rfloor$. *Then* $\lim_{n\to\infty}\sup_{j=0,1,\cdots,n-\lfloor c\log g(n)\rfloor}\eta_{n,j} = d_c$, *a.s., where* $d_c = \inf\left\{y:\ J(y)\ge\frac{1}{c}\right\}$.

For a random walk on $\mathbb{Z}$, this complements results of [4].

We next turn to the appropriate functional statements. Let $\psi(n)$ and γ_n be as in the statement of Theorem 1, and let $t(n,x)$ be a sequence of positive, increasing (in n, x) functions satisfying, for any $x \in]0,1]$,

$$\lim_{n\to\infty}\frac{g\left(\frac{t(n,x)}{g(n)\psi(n)}\right)}{g(\gamma_n)} = x > 0\,. \tag{4}$$

For example, if $g(n) \sim C\log n$, and $\frac{\log\psi(n)}{\log n} \underset{n\to\infty}{\longrightarrow} 0$, we can take $t(n,x) = n^x$. If $g(n) \sim C\log n$ and $\psi(n) = n^\beta$, $(0<\beta<1)$, we can take $t(n,x) = n^{x(1-\beta)+\beta}$. If $g(n) \sim C\log_2 n$ and $\frac{\log\psi(n)}{\log n} \underset{n\to\infty}{\longrightarrow} 0$, we can take $t(n,x) = e^{(\log n)^x}$ (here and throughout, $\log_k n$ denotes the $k-$th iterated logarithm function). If $g(n) \sim C\log_2 n$ and $\psi(n) = n^\beta$, $(0<\beta<1)$, we can take $t(n,x) = n^\beta e^{(\log n)^x}$.

It is straightforward to check, using (4), that for $0 \le x_1 < x_2 \le 1$, we have

$$\frac{t(n,x_1)}{t(n,x_2)} \underset{n\to\infty}{\longrightarrow} 0\,. \tag{5}$$

Let

$$\overline{L}_n(x) := \frac{L^o_{t(n,x)}}{g(n)\psi(n)}, \quad x \in [0,1]\,.$$

Note that $\overline{L}_n(x) \in M_+$, the space of non-negative Borel measures on $[0,1]$. Equip M_+ with the topology of weak convergence. Our main functional statement is the following:

Theorem 3 (Functional Moderate Deviations) *$\overline{L}_n(x)$ satisfies in M_+ a large deviation principle with speed $g(n)\psi(n)/g(\gamma_n)$ and rate function*

$$I(m) = \begin{cases} \int_0^1 \frac{1}{x} m(dx) & , \quad \frac{1}{x} \in L_1(m) \\ \infty & , \quad \textit{otherwise.} \end{cases}$$

As in the one-dimensional case, we can deduce convergence in distribution from our large deviation bounds, taking $\psi(n) \equiv 1$.

Theorem 4 (Functional Limit Law) *Let $t(n,x)$ be such that $g(t(n,x)) \sim xg(n)$, $x \in [0,1]$. The distributions of $\left(\frac{L^0_{t(n,x)}}{g(n)}\right)_{0\leq x\leq 1}$ converge weakly to $\mu \in M_1(M_+)$, the distribution of the process $(Z_x)_{0\leq x\leq 1}$ with increasing paths and independent increments given by*

$$P[Z_{x_2} - Z_{x_1} \in B] = \frac{x_1}{x_2}\delta_o(B) + \left(1 - \frac{x_1}{x_2}\right)\int_B \frac{1}{x_2} e^{-\frac{1}{x_2}u}\, du\,, \tag{6}$$

for any $0 \leq x_1 < x_2 \leq 1$, B Borel subset of $[0,\infty[$.

J. Bertoin kindly pointed out to us that in fact the process $(Z_x)_{0\leq x\leq 1}$ in Theorem 4 is a pure jump process which can be constructed from an inhomogeneous Poisson point process. Indeed, one may construct a Poisson process $N(x,z)$ on $[0,1] \times \mathbb{R}_+$ with intensity $n(x,z)dxdz = x^{-2}\exp(-z/x)dxdz$ and define $Y_x = \int_0^\infty z d_z N(x,z)$. Obviously, $(Y_x)_{0\leq x\leq 1}$ possesses increasing paths and independent increments. Moreover, it is not hard to check that for any $\lambda \geq 0$,

$$E\Big(\exp\left(-\lambda(Y_{x+y} - Y_x)\right)\Big) = \frac{1+\lambda x}{1+\lambda(x+y)} = E\Big(\exp\left(-\lambda(Z_{x+y} - Z_x)\right)\Big)\,,$$

proving that the processes $(Z_x)_{0\leq x\leq 1}$ and $(Y_x)_{0\leq x\leq 1}$ have the same law.

We finally mention that the functional moderate deviations of Theorem 3 are strong enough to derive by standard arguments the following Strassen law of the iterated logarithm presented in [1, Theorem 5]. Obtaining such a derivation was actually the original motivation for this work.

Theorem 5 (E. Csáki, P. Révész and J. Rosen) *Let $t(n,x)$ be such that $g(t(n,x)) \sim xg(n)$, $x \in [0,1]$. The set $\left(\frac{L^0_{t(n,x)}}{g(n)\log_2 g(n)}\right)_{0\leq x\leq 1}$, n large enough, is relatively compact in M_+ with limit points K, where $K = \{m : I(m) \leq 1\}$.*

We refer to [5] for proofs.

References

[1] E. Csáki, P. Révész and J. Rosen, *Functional laws of the iterated logarithms for local times of recurrent random walks in $\mathbb{Z}^2$*, preprint, 1997.

[2] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, Jones and Bartlett, 1993.

[3] N. C. Jain and W. E. Pruitt, *Lower tail probability estimates for subordinators and non-decreasing random walks*, Ann. Prob 15 (1987), pp. 75–101.

[4] N. C. Jain and W. E. Pruitt, *Maximal increments of local time of a random walk*, Ann. Prob 15 (1987), pp. 1461–1490.

[5] N. Gantert and O. Zeitouni, *Large and Moderate Deviations for the Local Time of a Recurrent Markov Chain on* $\mathbb{Z}^2$, preprint, 1997.

[6] M. Marcus and J. Rosen, *Laws of the iterated logarithm for the local time of recurrent random walks on* $\mathbb{Z}^2$ *and of Lévy processes and random walks in the domain of attraction of Cauchy random variables*, Ann. Inst. H. Poincaré 30 (1994), pp. 467–499.

A CENTRAL LIMIT THEOREM FOR THE MEAN MAGNETIZATION IN THE CURIE–WEISS MODEL WITH RANDOM COUPLINGS

Barbara Gentz

ETH Zürich

Abstract. We investigate the fluctuations of the mean magnetization in a Curie–Weiss model with ferromagnetic random couplings. In the high-temperature region, the limiting fluctuations as the system size tends to infinity are almost surely Gaussian as in the standard Curie–Weiss model. The most interesting case included is the randomly diluted Curie–Weiss model, where the independent random couplings are Bernoulli random variables and the model becomes the Curie–Weiss model on a random graph. The probability p_N for a bond to be present in the random graph is allowed to tend to zero as the system size N increases with the only constraint on the speed at which p_N tends to zero being $\sum_{N\geq 1} 1/(N^2 p_N) < \infty$.

The Curie–Weiss model being *the* mean-field model of a ferromagnet, it seems quite natural to study the diluted ferromagnet by taking random ferromagnetic couplings into account. To my knowledge, this has first been done by A. Bovier and V. Gayrard in [2], were the authors investigate the thermodynamics in the Curie–Weiss model with positively biased random couplings with emphasis on the case of the randomly diluted model. The following note continues this analysis by discussing the fluctuations of the order parameter in the high-temperature region under the additional assumption of not only positively biased but ferromagnetic couplings.

For arbitrary system size $N \in \mathbb{N}$, we denote by $\{\varepsilon_{ij}^{(N)}\}_{1\leq i<j\leq N}$ the independent, identically distributed random couplings which are assumed to satisfy $\varepsilon_{ij}^{(N)} \geq 0$ and $p_N = \mathbb{E}\,\varepsilon_{ij}^{(N)} > 0$. Let v_N^2 denote the variance of $\varepsilon_{ij}^{(N)}$. For convenience, we denote by $\mathbb{P}$ the joint distribution of the triangular scheme

$$\varepsilon = \{\,\varepsilon_{ij}^{(N)} \,|\, 1 \leq i < j \leq N,\ N \in \mathbb{N}\}.$$

At system size N, the random Hamiltonian of the Curie–Weiss model with couplings $\{\varepsilon_{ij}^{(N)}\}_{1\leq i<j\leq N}$ is given by

$$H_{N,\varepsilon}(\sigma) = -\frac{1}{Np_N} \sum_{1\leq i<j\leq N} \varepsilon_{ij}^{(N)} \sigma_i \sigma_j$$

1991 *Mathematics Subject Classification.* 60F05 (primary); 60K35, 82B44 (secondary).

with Ising spins $\sigma = \{\sigma_i\}_{i=1,\dots,N} \in \{-1,+1\}^N$. For finite temperature $1/\beta \in (0,\infty)$ and system size N, we define the random Gibbs measure $P_{N,\beta,\varepsilon}$ by

$$P_{N,\beta,\varepsilon}(\sigma) = \frac{1}{Z_{N,\beta,\varepsilon}} \, \mathrm{e}^{-\beta H_{N,\varepsilon}(\sigma)} \; \varrho^{\otimes N}(\sigma), \qquad \sigma \in \{-1,+1\}^N,$$

where $\varrho = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$ is the a priori distribution of each spin variable σ_i, and the partition function $Z_{N,\beta,\varepsilon}$ is the appropriate normalization.

In the sequel, we shall be concerned with the asymptotic behaviour of the spin per site or mean magnetization

$$m_N(\sigma) = \frac{1}{N}\sum_{i=1}^{N} \sigma_i, \qquad \sigma \in \{-1,+1\}^N,$$

under the Gibbs measure $P_{N,\beta,\varepsilon}$. In [2], A. Bovier and V. Gayrard prove a law of large numbers type result, namely

$$P_{N,\beta,\varepsilon} m_N^{-1} \Longrightarrow \frac{1}{2}\delta_{-z(\beta)} + \frac{1}{2}\delta_{z(\beta)}$$

$\mathbb{P}$-almost surely as $N \to \infty$, where $z(\beta)$ denotes the largest solution $z \in [0,1)$ of the Curie–Weiss equation $\beta z = \operatorname{artanh} z$. This is proved for positively biased but not necessarily ferromagnetic couplings under appropriate assumptions on the Legendre–Fenchel transform of the logarithmic moment generating function of the random couplings. The main examples are (a) Bernoulli random variables, (b) Gaussian random variables with strictly positive mean, or (c) random variables satisfying uniform Bernstein conditions. While in the cases (a) and (b) the condition $Np_N^2/v_N^2 \to \infty$ suffices, case (c) requires an additional condition on the variance v_N^2.

Note that $z(\beta) = 0$ if and only if $\beta \le 1$, which implies that $P_{N,\beta,\varepsilon} m_N^{-1} \Rightarrow \delta_0$ $\mathbb{P}$-almost surely for $\beta \le 1$. Motivated by this result, we investigate the fluctuations of m_N around $z(\beta) = 0$ for $\beta < 1$ in the case of ferromagnetic couplings.

Theorem. *Let $0 < \beta < 1$ and assume $Np_N \to \infty$ as well as*

$$\sum_{N \ge 1} \frac{v_N^2}{N^2 p_N^2} < \infty.$$

Then, for $\mathbb{P}$-almost all realizations of the random couplings,

$$P_{N,\beta,\varepsilon}(\sqrt{N} m_N)^{-1} \Longrightarrow \mathcal{N}\Big(0, \frac{1}{1-\beta}\Big)$$

as $N \to \infty$.

The proof, which relies on cluster expansions, will be given in [4].

Note that the standard Curie–Weiss model corresponds to the case of a degenerate distribution of the random couplings, namely $\mathbb{P}(\varepsilon_{ij}^{(N)} = 1) = 1$. The preceding theorem shows that the limiting fluctuations are Gaussian with the same variance as in the standard Curie–Weiss model at the same temperature, see [3, Theorem V.9.4]. Recall that the limiting fluctuations at $\beta = 1$ are non-Gaussian in the standard Curie–Weiss model, see [3, Theorem V.9.5].

As mentioned before, the randomly diluted Curie–Weiss model, i. e., the Curie–Weiss model on a random graph, is of special interest. Let the couplings be independent Bernoulli random variables with success probability p_N, i. e.,

$$\mathbb{P}(\varepsilon_{ij}^{(N)} = 1) = p_N = 1 - \mathbb{P}(\varepsilon_{ij}^{(N)} = 0).$$

At system size N, we consider a graph with edges $1, \dots, N$ and call a bond $\{i, j\}$ present in the random graph if and only if $\varepsilon_{ij} = 1$. Then, each bond of the graph is present with probability $p_N \searrow 0$. In this case, the theorem applies as long as $\sum_{N \geq 1} 1/(Np_N^2)$ is finite. This condition on p_N is more restrictive than assuming $Np_N \to \infty$ as in [2]. In particular, it implies that the probability for the random graph with edges $1, \dots, N$ to be connected tends to one as $N \to \infty$, while $Np_N \to \infty$ does not guaranty connectivity in the above sense, see [1, Theorem VII.3] for instance.

Possible extensions of this result include the improvement of the constraint on the speed at which p_N is allowed to tend to zero, positively biased couplings instead of ferromagnetic ones, and the cases $\beta > 1$ and $\beta = 1$. This will be subject to future research.

Acknowledgement. It is a great pleasure to thank E. Bolthausen, A. Bovier, and U. Schmock for fruitful discussions of different approaches to the problem and, in particular, to thank M. Löwe for bringing the problem to my attention.

References

1. Bollobás, B., *Random Graphs*, Academic Press, London, 1985.
2. Bovier, A., and V. Gayrard, *The thermodynamics of the Curie–Weiss model with random couplings*, J. Stat. Phys. **72**, no. 3–4, 643–664.
3. Ellis, R. S., *Entropy, Large Deviations, and Statistical Mechanics*, Springer-Verlag, New York, 1985.
4. Gentz, B., *The fluctuations of the mean magnetization in the Curie–Weiss model with random couplings in the high temperature region*, in preparation.

Department of Mathematics, ETH Zentrum, CH-8092 Zürich, Switzerland

E-mail address: gentz@math.ethz.ch (Barbara Gentz)

Large Deviations for the sums over the columns and the rows of a random matrix

M. Grunwald*
Department of Mathematics, Imperial College
180 Queen's Gate, London SW7 2BZ, UK
e-mail: m.grunwald@ic.ac.uk

March 23, 1998

Abstract

Let ν_s, $s \in \mathbb{R}_0^+$, be the semi-group belonging to some infinitely divisible (i.d.) measure ν_1 on $\mathbb{R}$ and let $J^{(N)}$ be a $N \times N$ real matrix of i.i.d. random variables $J_{i,j}^{(N)}$, $1 \le i,j \le N$, distributed according to $\nu_{\frac{1}{N}}$. Denote by $y_i = \sum_{j=1}^N J_{i,j}^{(N)}$ and $\tilde{y}_j = \sum_{i=1}^N J_{i,j}^{(N)}$ the sum over the i'th row respectively the j'th column of $J^{(N)}$. In this paper we prove a Large Deviations Principle (**LDP**) for the pair of the empirical measures of the y's and the $\tilde{y}$'s, that is for the law of $\left(\frac{1}{N}\sum_{i=1}^N \delta_{y_i}, \frac{1}{N}\sum_{j=1}^N \delta_{\tilde{y}_j}\right) =: (Z_N, \widetilde{Z}_N)$. For a general infinitely divisible measure ν_1 (i.e. a non-Gaussian and a non-Poissonian measure) the LDP is described in terms of a LDP for sums over columns and rows of a random matrix whose elements are Poissonian random measures respectively some pure jump martingales coming from the Lévy-decomposition of an i.d. law. The motivation for this results is an application to disordered mean-field dynamics.

Key words: Large deviations – Infinitely Divisible Measures

1 Introduction

Recently some convergence results were proved for the mean-field dynamics of a spin glass model ([BeGu97], [Gru96] and the article in [BoPi97]). Very important for this physical models is the symmetry of the coupling matrix used for describing the interaction in the

*Supported by EPSRC grant no. GR/K76801

system. This symmetry guarantees that the formulated dynamics are reversible with respect to the stationary state of the system. But unfortunately this symmetry causes the main difficulties in the proofs but leads also to the most interesting aspects of the limiting dynamics [Gru97a]. Essential for all the proofs in the spin-glass dynamics was that the couplings were Gaussian. In this paper we provide the main technical step to get rid of the necessity of Gaussian couplings, although we stay in the somewhat idealises world of infinitely divisible laws. The formulation of the corresponding dynamics will be published elsewhere.

We want briefly describe what we have in mind if we talk about disordered mean-field dynamics. With a mean-field dynamics we want to model the behaviour of a large number N of simple variables $x_i \in S$, $i = 1, \cdots N$, with some state space S. In the spin-glass models $S = \{-1, 1\}$ describes the two possible orientations of a magnetic spin, but we could think of $S = \{0, 1\}$ describing the state of an individuum in a population as being healthy respectively infected with some decease. In a spin-glass model we would use a symmetric matrix $N \times N$ matrix $K^{(N)}$, $K^{(N)}_{i,j} = K^{(N)}_{j,i}$, with i.i.d centred (i.e. zero mean) Gaussian with covariance $\frac{1}{N}$, that is

$$\mathcal{L}(K^{(N)}_{i,j}) = \mathcal{N}_{(0,1/N)} \quad i \leq j \leq N.$$

Then the field

$$\mathrm{h}_i(x) = \sum_{j=1} K^{(N)}_{i,j} x_j, \quad x \in S^N, \tag{1}$$

is the only information the i'st variable has about the state of the full system and it will change it's state with a rate depending on the random quantity $\mathrm{h}_i(x)$, it will try to align to this field. In an infection model we could model with the matrix $K^{(N)}$ a relation 'i is acquainted with j' (or is friend of), that is $K^{(N)}_{i,j} = 1$ if the relation is fulfilled. For obvious reasons the matrix should be symmetric and we could pick the $K^{(N)}_{i,j}$ independent from a Bernoulli distribution $\frac{p}{N}\delta_1 + (1 - \frac{p}{N})\delta_0$, which is the simplest choice to ensure that the average number of friends stays constant [Remark: $\frac{p}{N}\delta_1 + (1 - \frac{p}{N})\delta_0$ is not infinitely divisible, but 'behaves' in the limit $N \to \infty$ as a Poisson distribution with parameter $\frac{p}{N}$]. In this interpretation the field $\mathrm{h}_i(x)$ is now the number of infected persons i knows – the rate for an infection of i will grow with $\mathrm{h}_i(x)$.

In contrast to side disordered dynamical models we have models in mind where the behaviour of one variable is influenced by the state of the system through a averaged quantity – averaged with some random weights. In some applications (microscopic models of stock prices, or infections in a big town) we believe that mean-field models could be quite appropriate and could give explicit, but not necessarily easy to handle, limiting formulas for the dynamics. Proving Large Deviation results for this dynamics (with symmetry) boils essentially down to studying the empirical measures we address in this paper. For details please have a look in a forthcoming paper.

1.1 Statement of the results

Let ν_s, $s \in \mathbb{R}_0^+$ be some semi-group of infinitely divisible (**i.d.**) measures. Let $\left(J_{i,j}^{(N)}\right)_{1\le i,j\le N}$ be a $N \times N$ real matrix of i.i.d random variables with law $\mathcal{L}\left(J_{i,j}^{(N)}\right) = \nu_{\frac{1}{N}}$. We denote by $y_i = \sum_{j=1}^N J_{i,j}^{(N)}$ and $\tilde{y}_j = \sum_{i=1}^N J_{i,j}^{(N)}$ the sum over the i'th row respectively the j'th column of $J^{(N)}$. The y_i's, $i \in \{1, \cdots, N\}$ are independent and have the same law $\mathcal{L}(y_i) = \nu_1$. The same is true for the $\tilde{y}_j$'s, but $(y_1, \cdots, y_N)$ is not independent from $(\tilde{y}_1, \cdots, \tilde{y}_N)$, although the dependence is rather weak. We will use $Z_N := \frac{1}{N}\sum_{i=1}\delta_{y_i}$ and $\widetilde{Z}_N := \frac{1}{N}\sum_{j=1}\delta_{\tilde{y}_j}$ for the empirical measure of the row sums respectively the column sums, where δ_z denotes a Dirac-measure at point z in some space. It is not difficult to see, that the law of $(Z_N, \widetilde{Z}_N)$ converges weakly to a product measure, that is

$$\mathcal{L}(Z_N, \widetilde{Z}_N) \stackrel{\text{w}}{\Longrightarrow} \delta_{\nu_1} \otimes \delta_{\nu_1}.$$

[Actually the convergence is true almost surely with respect to some underlying probability space, which can be seen from the full Large Deviation Principle (**LDP**) Corollary 1.3.] The influence of the dependence is only visible on a large deviation scale but it is quite strong at this level. Motivated by the applications mentioned at the beginning we first proved the result in the case when $\nu_s = \mathcal{N}_{(0,sc)}$ is a centred Gaussian with covariance sc or a Poissonian semi-group $\nu_s = \pi_{sp}$ with parameter sp. In this cases we have the following result:

Lemma 1.1 *$(Z_N, \widetilde{Z}_N)$ satisfies a full LDP on the twofold product $\mathcal{M}_1(\mathbb{R}) \times \mathcal{M}_1(\mathbb{R})$ of probability measures on the Borel σ-field of $\mathbb{R}$ equipped with the weak topology with the good rate function S_{ν_1} defined below in Eq. (3).*

For finite N we have the condition

$$\int_{\mathbb{R}} x\, Z_N(dx) = \frac{1}{N}\sum_{i=1}^N y_i = \frac{1}{N}\sum_{i,j=1}^N J_{i,j}^{(N)} = \frac{1}{N}\sum_{j=1}^N \tilde{y}_j = \int_{\mathbb{R}} x\, \widetilde{Z}_N(dx). \tag{2}$$

This 'conservation law' is reflected in the rate S_{ν_1}. $\mathrm{S}_{\nu_1}(\mu, \tilde{\mu})$ is infinite if

$$\mathrm{m}_\mu^{\mathbb{R}} := \int x\, \mu(dx) \neq \mathrm{m}_{\tilde{\mu}}^{\mathbb{R}}$$

or one of the expectations does not exist. If $\mathrm{m}_\mu^{\mathbb{R}} = \mathrm{m}_{\tilde{\mu}}^{\mathbb{R}}$

$$\mathrm{S}_{\nu_1}(\mu, \tilde{\mu}) = \begin{cases} \mathrm{H}(\mu|\nu_1) + \mathrm{H}(\tilde{\mu}|\nu_1) - \Gamma_{\nu_1}^*(\mathrm{m}_\mu) & \text{if } \Gamma_{\nu_1}^*(\mathrm{m}_\mu) < \infty \\ \infty & \text{otherwise,} \end{cases} \tag{3}$$

where H is the relative entropy

$$\mathrm{H}(\mu|\nu) := \begin{cases} \int \ln \frac{d\mu}{d\nu}\, d\mu & \text{if } \mu \ll \nu \\ \infty & \text{otherwise,} \end{cases}$$

and $\Gamma^*_{\nu_1}$ is the rate function of a Cramér result for ν_1-distributed random variables, that is

$$\Gamma^*_{\nu_1}(y) := \sup_x \left[xy - \ln \int \mathrm{e}^{xu}\, \nu_1(du) \right],$$

that is $\Gamma^*_{\mathcal{N}(0,c)}(y) = -\frac{y^2}{2c}$ respectively $\Gamma_{\pi_p}(y) = y \ln \frac{y}{p} - y + p$ for $y \geq 0$ and ∞ otherwise. This result in the Gaussian and the Poissonian case is quite misleading. You are (we were) tempted to guess that for a general infinitely divisible measure ν_1 the LD result should hold with a similar rate function – this would correspond to some sort of asymptotic conditional independence of the two empirical measures $Z_N, \widetilde{Z}_N$ given the mean $\mathrm{m}^{\mathbb{R}}_{Z_N}$. But this is not true.

The LD result in the general case is not very explicit. We have to use a appropriate parametrisation of the set of all infinitely divisible laws to formulate the result. For this we split the infinitely divisible law ν_1 in its more or less natural parts. The way to characterise an infinitely divisible measure is to write down the characteristic function. The Lévy-Khinchin formula [Str93, Theorem 3.2.20] tells us that

$$\begin{aligned} \widehat{\nu_1}(z) &= \exp\left[\mathrm{i}az - \frac{c}{2}z^2 + \int_{]-1,1[^c} \left(\mathrm{e}^{\mathrm{i}xz} - 1\right) \mathcal{R}(dx) \right. \\ &\quad \left. + \int_{]-1,0[} \left(\mathrm{e}^{\mathrm{i}xz} - 1 - \mathrm{i}xz\right) \mathcal{Z}^-(dx) + \int_{]0,1[} \left(\mathrm{e}^{\mathrm{i}xz} - 1 - \mathrm{i}xz\right) \mathcal{Z}^+(dx) \right], \end{aligned} \tag{4}$$

where $\mathcal{R}$ is a finite positive measure with $\mathcal{R}(]-1,1[) = 0$ and the positive measures $\mathcal{Z}^\pm$ have no mass outside $]-1,0[$ respectively $]0,1[$ and fulfil

$$\int x^2\, \mathcal{Z}^\pm(dx) < \infty.$$

The probabilistic meaning of the parts in the characteristic function in Eq. (4) is given by the Lévy-decomposition. The $\mathrm{i}az$ is just a constant shift. We will assume that $a = 0$ for the rest of the paper. A Gaussian part is represented by $-\frac{c}{2}z^2$ – we know already how to deal with such a part. The new contributions are coming from the integral parts, which we are going to express with the help of a Poissonian random measure and a jump martingale.

Therefore we have to introduce some definitions. Let $(S, \mathcal{S})$ be some measurable space and $\mathcal{G}$ a σ-finite measure on S. Then there exists a Poissonian random measure, that is a random positive measure ω defined on some probability space such that:

- $\mathbf{P}\left[\{\omega(A) = \infty\}\right] = \begin{cases} 0 & \text{if } \mathcal{G}(A) < \infty \\ 1 & \text{if } \mathcal{G}(A) = \infty \end{cases}$ for $A \in \mathcal{S}$
- $\omega(A_1), \ldots, \omega(A_k)$ are independent for disjoint sets $A_1, \ldots, A_k \in \mathcal{S}$
- and for all $A \subset S$ s.t. $\mathcal{G}(A) < \infty$ $\mathcal{L}(\omega(A)) = \pi_{\mathcal{G}(A)}$ has a Poissonian law for with parameter $\mathcal{G}(A)$.

Such a measure could be constructed from a Poissonian field on S constructed from the atom free part of $\mathcal{G}$ [Kin93, Theorem 2.5] and a part

$$\sum_{x\in S:\mathcal{G}(\{x\})>0} \delta_x \xi_x,$$

where $(\xi_x)_{\{x\in S:\mathcal{G}(\{x\})>0\}}$ are (at most countably many) independent Poissonian random variables with corresponding parameter $\mathcal{G}(\{x\})$. Let (S,r) be a separable, metric space and $\mathcal{G}(S)<\infty$, then ω is almost surely finite and the law

$$\mathcal{L}(\omega) =: \Pi_{\mathcal{G}} \in \mathcal{M}_1(\mathcal{M}_+^{\mathcal{I}}(S))$$

defines a probability measure on the set of discrete, finite measures

$$\mathcal{M}_+^{\mathcal{I}}(S) := \left\{\nu = \sum_{l=1}^{k} \delta_{x_l} : k\in \mathbb{N},\, x\in S^k\right\} \cup \{0\} \subset \mathcal{M}_+(S)$$

equipped with the weak topology and the corresponding σ-field.

Let $\Pi_{\mathcal{R}}$ be the law of the Poissonian random measure belonging to $\mathcal{R}$ on $\mathbb{R}$. Denote by

$$\begin{aligned} \phi : \mathcal{M}_+^{\mathcal{I}}(\mathbb{R}) &\longrightarrow \mathbb{R} \\ \mu &\longmapsto \int x\,\mu(dx). \end{aligned} \tag{5}$$

Then the characteristic function of $\Pi_{\mathcal{G}}\circ\phi^{-1}$ is an exponential of the first integral part of the exponent in Eq. (4), that is $\Pi_{\mathcal{G}}\circ\phi^{-1}$ is a way to describe a compound Poissonian $\exp[-\mathcal{R}(\mathbb{R}) + *\mathcal{R}] := \mathrm{e}^{-\mathcal{R}(\mathbb{R})}\left(\delta_0 + \sum_{k>0}\frac{\mathcal{R}^{*k}}{k!}\right)$, where $\mathcal{R}^{*k}$ denotes the k-fold convolution of the measure $\mathcal{R}$.

The last part in the exponent of Eq. (4) can be interpreted as two martingale contributions.

Let ω be a Poissonian random measure given by $\mathcal{Z}^-$ on $[-1,0[$. Then

$$X_\omega(s) := \int_{[-1,s]} u\,\omega(du) - \int_{[-1,s]} u\,\mathcal{Z}^-(du), \quad \text{for } s\in[-1,0[$$

defines a martingale for the natural filtration

$$\mathcal{F}_s := \sigma\left(\{X_\cdot(t) : -1\le t\le s\}\right)$$

generated by the variables $X_\cdot(t)$ up to time s. The martingale is closeable by some variable $X_\cdot(0)$. The characteristic function of the law of $X_\cdot(0)$ is given by an exponential of the second integral part in the exponent of Eq. (4). $X(\cdot)$ is almost surely right continuous with left limits. The law of the process defines therefore a probability measure $\mathcal{L}(X(\cdot)) =: \mathcal{P}^-_{\mathcal{Z}^-}$ on the space $\mathcal{D}_{\mathbb{R}}[-1,1]$ of right continuous functions having left limits (RCLL-functions) equipped with the topology and the corresponding σ-field generated by the Skorohod topology.

We do the same construction with $\mathcal{Z}^+$ on $[1,0[$ (running the martingale in the time-reversed direction) and end up with some other martingale Y and a measure $\mathcal{P}^+_{\mathcal{Z}^+}$ on $\mathcal{D}_{\mathbb{R}}[1,0]$ – we use the 'interval' $[1,0]$ to indicate the reversed time-order. The law of $Y(0)$ corresponds to the last integral part in Eq. (4).

We have now all the pieces together to formulate the LDP for a general infinitely divisible law ν_1. We denote by W the following product space

$$W = \mathbb{R} \times \mathcal{M}^{\mathcal{I}}_{+}(\mathbb{R}) \times \mathcal{D}_{\mathbb{R}}[-1,0] \times \mathcal{D}_{\mathbb{R}}[1,0],$$

equipped with the product topology. W is a Polish space since all factors are Polish. Let $\mu \in \mathcal{M}_1(W)$. Then we denote by m_μ the vector of the 'mean'-value of μ whenever this quantity exists, that is $\mathrm{m}_\mu = \left(\mathrm{m}^{\mathbb{R}}_\mu, \mathrm{m}^{\mathcal{M}}_\mu, \mathrm{m}^{+}_\mu, \mathrm{m}^{-}_\mu\right)$ with

$$\mathrm{m}_\mu := \int_W (u,\rho,x,y)\,\mu(du,d\rho,dx,dy).$$

Let $\left(J^{(N,W)}_{i,j}\right)_{i,j\leq N}$ be now a $N \times N$ matrix of W-valued i.i.d. random variables distributed according to the law

$$\Sigma_{\left(0,\frac{c}{N},\frac{1}{N}\mathcal{R},\frac{1}{N}\mathcal{Z}^-,\frac{1}{N}\mathcal{Z}^+\right)} := \mathcal{N}_{(0,c/N)} \otimes \Pi_{\frac{1}{N}\mathcal{R}} \otimes \mathcal{P}^-_{\frac{1}{N}\mathcal{Z}^-} \otimes \mathcal{P}^+_{\frac{1}{N}\mathcal{Z}^+}.$$

With the map

$$\begin{aligned} \Phi : W &\longrightarrow \mathbb{R} \\ (u,\rho,x,y) &\longmapsto u + \phi(\rho) + x(0) + y(0) \end{aligned} \tag{6}$$

the law of $\Phi\left(J^{(N,W)}_{i,j}\right)$ is our infinitely divisible law $\nu_{\frac{1}{N}}$ in which we are interested. We construct the y^W's, the $\tilde{y}^W$'s and the empirical measures Z^W_N and $\widetilde{Z}^W_N$ as in the real-valued case and can formulate now our main result:

Theorem 1.2 *The law of* $(Z^W_N, \widetilde{Z}^W_N)$ *on* $\mathcal{M}_1(W) \times \mathcal{M}_1(W)$ *(each equipped with the weak topology) fulfils a full LDP with a good rate function* $\mathrm{S}^{\,W}_{\nu_1}$. *When the following conditions are not fulfilled the rate function* $\mathrm{S}^{\,W}_{\nu_1}(\mu,\tilde{\mu})$ *is* ∞*:*

- m_μ *and* $\mathrm{m}_{\tilde{\mu}}$ *have to be well defined*
- $\mathrm{m}_\mu = \mathrm{m}_{\tilde{\mu}}$
- $\widehat{\mathrm{m}}^-_\mu(s) := \int_{[-1,s]} \frac{1}{u}\,d\mathrm{m}^-_\mu(u) + \int_{[-1,s]} \mathcal{Z}^-(du)$, $s \in [-1,0[$, *and* $\widehat{\mathrm{m}}^+_\mu(s) := \int_{[s,0]} \frac{1}{u}\,d\mathrm{m}^+_\mu(u) + \int_{[s,0]} \mathcal{Z}^+(du)$, $s \in]0,1]$ *have to be finite-valued monotone increasing functions (for the* $+$*-function with reversed time) which define then some positive measures* $d\widehat{\mathrm{m}}^{\mp}_\mu$ *on* $[-1,0[$ *respectively* $]0,1]$
- $\int u^2\, d\widehat{\mathrm{m}}^{\mp}_\mu(u) < \infty$

If the above conditions are fulfilled the rate function is:

$$\begin{aligned} \mathrm{S}_{\nu_1}^{W}(\mu,\tilde{\mu}) &= \mathrm{H}\left(\mu \middle| \Sigma_{\left(\mathrm{m}_\mu^{\mathbb{R}},c,\mathrm{m}_\mu^{\mathcal{M}},d\widehat{\mathrm{m}}_\mu^-,d\widehat{\mathrm{m}}_\mu^+\right)}\right) + \mathrm{H}\left(\tilde{\mu} \middle| \Sigma_{\left(\mathrm{m}_\mu^{\mathbb{R}},c,\mathrm{m}_\mu^{\mathcal{M}},d\widehat{\mathrm{m}}_\mu^-,d\widehat{\mathrm{m}}_\mu^+\right)}\right) \\ &\quad + \mathrm{H}\left(\Sigma_{\left(\mathrm{m}_\mu^{\mathbb{R}},c,\mathrm{m}_\mu^{\mathcal{M}},d\widehat{\mathrm{m}}_\mu^-,d\widehat{\mathrm{m}}_\mu^+\right)} \middle| \Sigma_{(0,c,\mathcal{R},\mathcal{Z}^-,\mathcal{Z}^+)}\right), \end{aligned}$$

for $(\mu,\tilde{\mu}) \in \mathcal{M}_1(W) \times \mathcal{M}_1(W)$.

For the original sequence of empirical measures $(Z_N, \widetilde{Z}_N)$ constructed from ν_s on $\mathbb{R}$ we can now formulate the LDP as a contraction from Theorem 1.2:

Corollary 1.3 *The laws of the empirical measures* $(Z_N, \widetilde{Z}_N)$ *on* $\mathcal{M}_1(\mathbb{R}) \times \mathcal{M}_1(\mathbb{R})$ *obey a full LDP with the good rate function* S_{ν_1}*, which can be described as the contraction of the rate function in Theorem 1.2:*

$$\mathrm{S}_{\nu_1}(\kappa,\widetilde{\kappa}) = \inf\left[\mathrm{S}_{\nu_1}^{W}(\mu,\tilde{\mu}) : (\mu,\tilde{\mu}) \text{ s.t. } (\mu\circ\Phi^{-1},\tilde{\mu}\circ\Phi^{-1}) = (\kappa,\widetilde{\kappa})\right].$$

Proof. The map $\phi : \mathcal{M}_+^{\mathcal{I}}(\mathbb{R}) \to \mathbb{R}$ is continuous and the map $\Phi : W \to \mathbb{R}$ is continuous by the definition of the topology on $\mathcal{D}_{\mathbb{R}}[-1,1]$ and $\mathcal{D}_{\mathbb{R}}[1,0]$. Therefore the map $\mathcal{M}_1(W) \times \mathcal{M}_1(W) \ni (\nu,\tilde{\nu}) \mapsto (\nu\circ\Phi^{-1},\tilde{\nu}\circ\Phi^{-1}) \in \mathcal{M}_1(\mathbb{R}) \times \mathcal{M}_1(\mathbb{R})$ is continuous and Corollary 1.3 follows directly from the contraction principle [DeZe93, Theorem 4.2.1]. ■

Remark 1.4 *The rate function* S_{ν_1} *has not much to do with the rate guessed from the Gaussian and Poissonian case. It is neither bigger nor smaller than the guessed one. Knowing the general result the Gaussian and the Poissonian case are some 'lucky' coincidences. In the expression for* S_{ν_1} *is optimised over the set of infinitely divisible laws evaluated with the third entropy part in* $\mathrm{S}_{\nu_1}^{W}$*. In the Gaussian and Poissonian case the set of laws with a finite entropy are one-dimensional and the parameter is the mean (as shift in the Gaussian case and expectation in the Poisson result).*

References

[BeGu97] G. Ben Arous, A. Guionnet [1997]: Symmetric Langevin spin glass dynamics, Ann. Probab. Vol. **25**, No 3, 1367-1423.

[BoPi97] A. Bovier, P Picco, eds.: [1997] Mathematics of spin glasses and neural networks, "Progress in Probability", Birkhäuser.

[DeZe93] A. Dembo, O. Zeitouni [1993]: Large Deviations Techniques. Jones and Bartlett Publisher, Boston-London.

[Gru96] M. Grunwald [1996]: Sanov results for Glauber spin-glass dynamics, Probab. Theory Relat. Fields **106**, 187-232.

[Gru97a] M. Grunwald [1997]: Convergence results for Glauber spin-glass dynamics, Preprint 545, Technische Universität Berlin, Berlin.

[Gru97b] M. Grunwald [1997]: Proving a Sanov result on 'path'-space from a product result, preprint.

[Kin93] J.F.C. Kingman [1993]: Poisson Processes, Oxford Studies in Probability, Clarendon Press, Oxford.

[Str93] D.W. Stroock [1993]: Probability Theory, an analytic view, Cambridge University Press, Cambridge.

Branching Random Walk in a Catalytic Medium

Achim Klenke
(joint work with A. Greven and A. Wakolbinger)

Universität Erlangen-Nürnberg

Abstract

We consider a model of (critical binary) branching random walk where the local branching rate is determined by a random medium. This medium may vary in space and time.

We focus on the case where the medium is itself an autonomous branching random walk and investigate the ergodic theory of this model.

This exposition announces results of a joint work of Andreas Greven, Anton Wakolbinger and the author [GKW].

1 Branching Random Walk

We start with the definition of classical branching random walk on $\mathbb{Z}^d$ and with collecting some of the basic facts. Branching random walk (BRW) $(\eta_t)_{t\geq 0}$ on $\mathbb{Z}^d$ is a continuous time (infinite) particle system, where particles evolve according to the following rules:

- Each particle has a random exponential mean 1 lifetime.
- During its lifetime the particle performs a continuous time random walk on $\mathbb{Z}^d$ with random walk kernel $a(\cdot,\cdot)$.
- At the end of its lifetime the particle is replaced by either 0 or 2 offspring particles (at the same site); each possibility occurs with probability $\frac{1}{2}$.
- The new particles behave (independently) according to the same rules.
- All random mechanisms are independent.

For $A \subset \mathbb{Z}^d$ we write

$$\eta_t(A) = \#\{\text{particles in } A \text{ at time } t\}.$$

More formally, we can define $(\eta_t)_{t\geq 0}$ as Markov process on (a suitable subspace of) $\mathcal{N}(\mathbb{Z}^d)$, the space of non-negative integer valued measures on $\mathbb{Z}^d$. We write δ_i for the Dirac measure on $i \in \mathbb{Z}^d$. Then $(\eta_t)_{t\geq 0}$ is the Markov process with the generator

$$\mathcal{G}_{\text{brw}} = \mathcal{G}_{\text{walk},a} + \mathcal{G}_{\text{branch}},$$

where $\mathcal{G}_{\text{walk},a}$ and $\mathcal{G}_{\text{branch}}$ are defined (for suitable $f : \mathcal{N}(\mathbb{Z}^d) \to \mathbb{R}$ and $\eta \in \mathcal{N}(\mathbb{Z}^d)$) by

$$\mathcal{G}_{\text{walk},a} f(\eta) = \sum_{i,j\in\mathbb{Z}^d} a(i,j)\eta(\{i\})[f(\eta+\delta_j-\delta_i) - f(\eta)] \tag{1}$$

$$\mathcal{G}_{\text{branch}} f(\eta) = \sum_{i\in\mathbb{Z}^d} \eta(\{i\}) \left[\frac{1}{2} f(\eta+\delta_i) + \frac{1}{2} f(\eta-\delta_i) - f(\eta)\right]. \tag{2}$$

We collect some basic facts of the longtime behavior of $(\eta_t)_{t\geq 0}$ where we start with η_0 a Poisson field $\mathcal{H}(\theta)$ on $\mathbb{Z}^d$ with intensity θ. This is, under $\mathcal{H}(\theta)$ the components of η_0 are iid mean θ Poisson random variables.

The longtime behavior is dominated by a competition of opposed tendencies:

- The branching mechanism leads to local extinction (a.s.) and rare large families (clustering).
- The random walk part of the dynamics tries to transport particles from densely populated areas into evacuated regions (smoothing).

The greater the mobility of the particles is the greater is the impact of the smoothing. The crucial property is the recurrence behavior of the symmetrized kernel

$$\widehat{a}(i,j) = \frac{1}{2}[a(i,j) + a(j,i)].$$

We say that $\widehat{a}$ is recurrent (transient) if a random walk on $\mathbb{Z}^d$ with transition kernel $\widehat{a}$ is recurrent (transient).
Let $\mathcal{L}$ denote the law of a random variable and denote by "$\Longrightarrow$" weak convergence (of finite dimensional distributions).

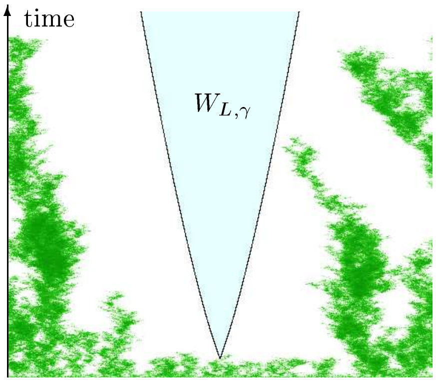

Figure 1 BRW on the torus $\mathbb{Z}/(500)$ and the empty parable $W_{L,\gamma}$

Theorem 1 *(a). If $\widehat{a}$ is transient then there exists a translation invariant random variable η_∞^θ with values in $\mathcal{N}(\mathbb{Z}^d)$ and expectation $\mathbf{E}[\eta_\infty^\theta(\{i\})] = \theta$, $i \in \mathbb{Z}^d$, and*

$$\mathcal{L}^{\mathcal{H}(\theta)}[\eta_t] \overset{t\to\infty}{\Longrightarrow} \mathcal{L}[\eta_\infty^\theta].$$

(b). If $\widehat{a}$ is recurrent, then for all finite sets $A \subset \mathbb{Z}^d$

$$\mathbf{P}^{\mathcal{H}(\theta)}[\eta_t(A) > 0] \overset{t\to\infty}{\longrightarrow} 0.$$

(c). Let $d = 1$ and let $a(\cdot,\cdot)$ be the kernel of symmetric simple random walk on $\mathbb{Z}$. For $\gamma > 1$ and $L > 0$ let (see Figure 1)

$$W_{L,\gamma} = \big\{(t,i) \in [0,\infty) \times \mathbb{Z} : \ t > L \cdot (1 + |i|)^\gamma\big\}.$$

Then

$$\mathbf{P}^{\mathcal{H}(\theta)}[\eta_t(\{i\}) > 0 \textit{ for some } (t,i) \in W_{L,\gamma}] \overset{L\to\infty}{\longrightarrow} 0. \tag{3}$$

In particular, for all finite sets $A \subset \mathbb{Z}$,

$$\eta_t(A) \overset{t\to\infty}{\longrightarrow} 0, \qquad \mathbf{P}^{\mathcal{H}(\theta)} - a.s. \tag{4}$$

(d). Let $d = 2$ and $a(\cdot,\cdot)$ be symmetric simple random walk. Let $(\varrho_t)_{t\geq 0}$ be super Brownian motion on $\mathbb{R}^2$ and let $\mathcal{L}^\theta$ denote its law if $\varrho_0 = \theta \times$ Lebesgue measure. Then

$$\mathcal{L}^{\mathcal{H}(\theta)}[T^{-1}\eta_{tT}(T^{1/2}\cdot))_{t\geq 0}] \overset{T\to\infty}{\Longrightarrow} \mathcal{L}^\theta[(\varrho_t)_{t\geq 0}]. \tag{5}$$

In particular, for $A \subset \mathbb{Z}^2$ non-void this implies

$$\limsup_{t\to\infty} \eta_t(A) = \infty \qquad \mathbf{P}^{\mathcal{H}(\theta)} - a.s. \tag{6}$$

The statements of (a) and (b) are standard (see, e.g., [LMW]). Only (c) and (d) are new, though the implications (4) and (6) are well known. See [GKW] for a proof. If you are not familiar with super Brownian motion (SBM), (5) can be read as a definition. SBM is the measure valued process that arises as the scaling limit of BRW (see also [DHV]). As a reference for SBM we recommend [D].

2 Catalytic Branching Random Walk

Catalytic branching random walk (CBRW) is a Markov process $(\xi_t)_{t\geq 0}$ defined similarly as $(\eta_t)_{t\geq 0}$. The particles evolve according to the rules described above (now with a random walk kernel b), however the clocks that count the *remaining* lifetimes now run at a variable speed. This speed is proportional to the local abundance of a certain "catalytic matter" at a given site and time. This catalytic matter may itself be a random process. We focus on the case where the catalytic medium is the BRW $(\eta_t)_{t\geq 0}$ introduced above. Hence we consider a Markov process $(\eta_t, \xi_t)_{t\geq 0}$ with values in (a suitable subspace of) $\mathcal{N}(\mathbb{Z}^d)\times\mathcal{N}(\mathbb{Z}^d)$. $(\eta_t)_{t\geq 0}$ is a BRW with random walk kernel a and is independent of $(\xi_t)_{t\geq 0}$ (more precisely: autonomous). $(\xi_t)_{t\geq 0}$ depends on $(\eta_t)_{t\geq 0}$. For a given realization of $(\eta_t)_{t\geq 0}$ it is BRW with random walk kernel b and with local branching rate proportional to η_t. We call (η_t) the *catalyst* and (ξ_t) the *reactant*.

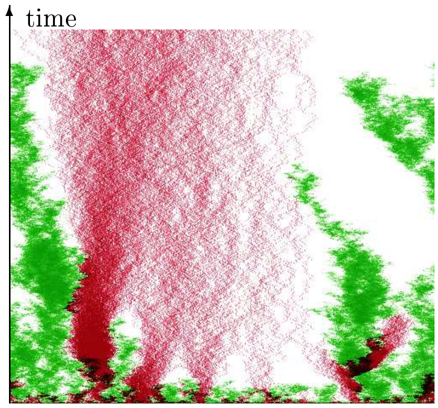

Figure 2 CBRW on the torus $\mathbb{Z}/(500)$. The catalyst (green) is the same realization as in Figure 1.

More formally, $(\eta_t, \xi_t)_{t\geq 0}$ is the Markov process with generator $\mathcal{G}_{\mathrm{cbrw}}$ defined by (for suitable $f: \mathcal{N}(\mathbb{Z}^d)\times\mathcal{N}(\mathbb{Z}^d)\to\mathbb{R}$ and $\eta,\xi\in\mathcal{N}(\mathbb{Z}^d)$)

$$\mathcal{G}_{\mathrm{cbrw}} = (\mathcal{G}_{\mathrm{brw}} f(\cdot,\xi))(\eta) + (\mathcal{G}_{\mathrm{walk},b} f(\eta,\cdot))(\xi) + (\mathcal{G}_{\mathrm{c-branch}} f)(\eta,\xi). \tag{7}$$

Here $\mathcal{G}_{\mathrm{walk},b}$ is defined similarly as $\mathcal{G}_{\mathrm{walk},a}$ and $\mathcal{G}_{\mathrm{c-branch}}$ is defined by

$$\mathcal{G}_{\mathrm{c-branch}} f(\eta,\xi) = \sum_{i\in\mathbb{Z}^d} \eta(\{i\})\xi(\{i\}) \left[\frac{1}{2} f(\eta,\xi+\delta_i) + \frac{1}{2} f(\eta,\xi-\delta_i) - f(\eta,\xi)\right]. \tag{8}$$

We will be interested in the longtime behavior of (η_t,ξ_t) started in a product of Poisson fields with intensities θ_η and θ_ξ:

$$\mathcal{L}^{\mathcal{H}(\theta_\eta,\theta_\xi)}[(\eta_0,\xi_0)] = \mathcal{H}(\theta_\eta,\theta_\xi) := \mathcal{H}(\theta_\eta)\otimes\mathcal{H}(\theta_\xi).$$

Again the symmetrized kernels $\widehat{a}$ and $\widehat{b}$ will play the prominent rôles.

Our first theorem is concerned with the case where the catalyst is *stable*, this is, in the regime of Theorem 1(a). In this case the catalyst converges toward an ergodic field and we see for the reactant the same dichotomy between extinction and survival as for the catalyst in Theorem 1(a) and (b):

Theorem 2 *Let $\widehat{a}$ be transient.*

(a). If $\widehat{b}$ is transient then there exists a translation invariant random variable $(\eta_\infty^{\theta_\eta}, \xi_\infty^{\theta_\xi})$ with $\mathbf{E}[(\eta_\infty^{\theta_\eta}(\{i\}), \xi_\infty^{\theta_\xi}(\{i\}))] = (\theta_\eta, \theta_\xi)$, $i \in \mathbb{Z}^d$, such that

$$\mathcal{L}^{\mathcal{H}(\theta_\eta,\theta_\xi)}[(\eta_t, \xi_t)] \overset{t\to\infty}{\Longrightarrow} \mathcal{L}[(\eta_\infty^{\theta_\eta}, \xi_\infty^{\theta_\xi})]. \tag{9}$$

(b). If $\widehat{b}$ is recurrent then

$$\mathcal{L}^{\mathcal{H}(\theta_\eta,\theta_\xi)}[(\eta_t, \xi_t)] \overset{t\to\infty}{\Longrightarrow} \mathcal{L}[(\eta_\infty^{\theta_\eta}, \mathbf{0})]. \tag{10}$$

For the case where $\widehat{a}$ is recurrent we cannot make a statement that is as general as the last one. One point is that we have to distinguish between two situations that exhibit fundamentally different behavior:

(a). The catalyst dies out locally almost surely. This is the regime of Theorem 1(c). Since larger and larger areas become eventually vacated from the catalyst we expect here that the reactant can evolve as independent random walks in these areas. This would lead to a Poisson field as limit for the reactant.

(b). The catalyst dies out only in probability but populates any region at arbitrarily late times almost surely. This is the regime of Theorem 1(d). However this possibility occurs also if $d = 1$ and a is the kernel of simple random walk with a drift. Here it is not a priori clear what the results will be and this is the source for interesting results.

Rather than giving a characterization (which we do not have) we give some examples for the typical possible behaviors.

Let $\delta_{\mathbf{0}}$ denote the Dirac measure on the empty configuration $\mathbf{0} \in \mathcal{N}(\mathbb{Z}^d)$.

Theorem 3 *(a). Assume that $d = 1$ and that $a(\cdot,\cdot)$ is the kernel of symmetric simple random walk. Let $b(\cdot,\cdot)$ be symmetric with a moment of order p for some $p > 1$: $\sum_{i\in\mathbb{Z}} b(0,i)|i|^p < \infty$. Then*

$$\mathcal{L}^{\mathcal{H}(\theta_\eta,\theta_\xi)}[(\eta_t, \xi_t)] \overset{t\to\infty}{\Longrightarrow} \delta_{\mathbf{0}} \otimes \mathcal{H}(\theta_\xi). \tag{11}$$

(b). Let $a(\cdot,\cdot)$ be the kernel of random walk that stands still ($a(i,i) = 1$ and $a(i,j) = 0$ if $i \neq j$). Let $b(\cdot,\cdot)$ be simple random walk with a drift $h \neq 0$. Then

$$\mathcal{L}^{\mathcal{H}(\theta_\eta,\theta_\xi)}[(\eta_t, \xi_t)] \overset{t\to\infty}{\Longrightarrow} \delta_{\mathbf{0}} \otimes \delta_{\mathbf{0}}. \tag{12}$$

The reason why part (a) is true is that for large L the area W_L, γ is catalyst-free with high probability (see Theorem 1 (c)). The moment assumption on b makes sure that a reactant particle that has entered the free area remains in it. Hence the limit of the reactant is a Poisson field. Since there is no loss of intensity at finite times this Poisson field has the full intensity θ_ξ. The actual proof of this statement in [GKW] makes this arguing rigorous.

On the other hand, part (b) is somewhat counter-intuitive; at first glance at least. Although the reactant has a greater mobility it goes extinct. This can be understood if we recall the fact that the catalyst clumps that go along with local extinction are at time t of order t apart. Now since a random walk with drift can cross distances of order t within time t the reactant particles manage to run into the sparser and sparser catalytic clumps until they finally die. This is illustrated in Figure 3 where we chose a to be symmetric simple random random walk instead (this gives a better visualization). In fact, the theorem should be true in this case also but there are a lot of technical problems connected with this case.

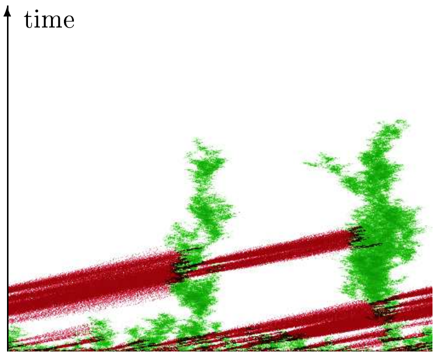

Figure 3 CBRW on the torus $\mathbb{Z}/(500)$. The reactant has a drift to the right.

Now we come to the case where the catalyst populates all sets at arbitrarily large times. We only give the example where both a and b are symmetric simple random walks on $\mathbb{Z}^2$. The catalyst dies out locally in probability. It turns out that the catalyst free regions are of the same order as the range of the reactant random walk. Within these regions the reactant particles evolve as independent random walks and tend towards a Poisson field. However the intensity of the Poisson field is not equal to θ_ξ. In fact, it is not even deterministic but is a random variable ζ that reflects the history of the catalyst. This ζ allows a description in terms of the diffusion limit of CBRW. In [GKW] we show that in this special situation CBRW can be rescaled such that it converges to a measure valued process $(X_t^\varrho)_{t\geq 0}$ on $\mathbb{R}^2$ called *catalytic super Brownian motion*. For a definition and the construction of this process see [FK] and the references given there. It is shown in [FK] that X_1^ϱ has a density

$$\zeta = \frac{X_1^\varrho(dx)}{dx}\bigg|_{x=0}.$$

Let $\mathcal{L}^{\theta_\eta,\theta_\xi}$ denote the law of $(\varrho_t, X_t^\varrho)_{t\geq 0}$ if both ϱ_0 are X_0^ϱ are multiples of Lebesgue measure with intensities θ_η and θ_ξ. It is shown in [FK] that $\mathbf{E}^{\theta_\eta,\theta_\xi}[\zeta] = \theta_\xi$. The next theorem states that the reactant converges toward a mixed Poisson field with random intensity ζ.

Theorem 4 *Let $d = 2$ and assume that $a = b$ is the kernel of symmetric simple random walk on $\mathbb{Z}^2$. Then*

$$\mathcal{L}^{\mathcal{H}(\theta_\eta,\theta_\xi)}[(\eta_t,\xi_t)] \overset{t\to\infty}{\Longrightarrow} \delta_{\mathbf{0}} \otimes \mathbf{E}^{\theta_\eta,\theta_\xi}[\mathcal{H}(\zeta)]. \tag{13}$$

This theorem exhibits a new quality for the longtime behavior of an interacting particle system. While the local structure gets lost there is a global *random* quantity that determines the limit field. A similar behavior has been observed so far only for rescaled occupation time fields. See [CG] for an example.

References

[CG] J.T. Cox, D. Griffeath: Occupation time limit theorems for critical branching brownian motions. Ann. Probab. **13(4)**, 1108-1132, (1985)

[D] D.A. Dawson: Measure-Valued Markov Processes. In: Ecole d'Eté de Probabilités de St.Flour XXI - 1991, LNM **1541**, Springer-Verlag (1993)

[DHV] D. A. Dawson, K. Hochberg and V. Vinogradov: On weak convergence of branching particle systems undergoing spatial motion, Israel Mathematical Conference Proceedings (to appear). (1997)

[FK] K. Fleischmann, A. Klenke: Smooth density field of two-dimensional catalytic super-Brownian motion. Preprint No. 331, Weierstraß Institut, Berlin (1997).

[GKW] A. Greven, A. Klenke, and A. Wakolbinger. The longtime behavior of branching random walk in a catalytic medium. In preparation, (1997)

[LMW] A. Liemant, K. Matthes, A. Wakolbinger: Equilibrium Distributions of Branching Processes, Akademie-Verlag Berlin, (1988.)

Author:
Achim Klenke
Mathematisches Institut
Universität Erlangen-Nürnberg
Bismarckstraße $1\frac{1}{2}$
91054 Erlangen
Germany
e-mail: klenke@mi.uni-erlangen.de
http://www.mi.uni-erlangen.de/~klenke

MOMENT ASYMPTOTICS FOR THE ANDERSON PROBLEM

W. König

1 Cauchy Problem for the Anderson Model

We consider the so-called *Anderson problem*, which is the following parabolic differential equation for $t > 0$ and $x \in \mathbb{Z}^d$:

$$\begin{cases} \frac{\partial}{\partial t} u(t,x) = \kappa \Delta u(t,x) + \xi(x) u(t,x), \\ u(0,x) = 1. \end{cases} \tag{1.1}$$

Here $\Delta f(x) = \sum_{y \sim x} (f(y) - f(x))$ denotes the lattice Laplacian, and $\xi = \{\xi(x); x \in \mathbb{Z}^d\}$ is a random potential consisting of i.i.d. variables. The expectation with respect to the field is denoted by $\langle \cdot \rangle$. We want to address the question: How does $\Lambda_p(t) = \log \langle u(t,0)^p \rangle$ behave as $t \to \infty$ for $p \geq 1$?

There are two competing effects: the potential ξ makes u spatially irregular, the diffusive part $\kappa\Delta$ flattens u. We make two assumptions, the first of which is technical: We suppose that

$$\left\langle \left(\frac{\xi(0)_+}{\log_+ \xi(0)_+} \right)^d \right\rangle < \infty.$$

Secondly, we assume that $H(t) = \log \langle e^{t\xi(0)} \rangle$ is finite for all $t > 0$.

Under these assumptions, Gärtner and Molchanov (1990) proved that the unique nonnegative solution of (1.1) is given by the Feynman-Kac formula

$$u(t,x) = \mathbb{E}_x \exp\left\{ \int_0^t \xi(x(s))\, ds \right\}.$$

Here $(x(s))_{s \geq 0} =$ is a random walk on $\mathbb{Z}^d$ with generator $\kappa\Delta$, started at x.

2 Intermittency

First we explain, on an heuristic level, a certain effect which is expected for the solution of (1.1): as t becomes large, the field $u(t,\cdot)$ shrinks together to small islands with high peaks.

Definition: A random field $u = \{u(t,x) : t > 0, x \in \mathbb{Z}^d\}$ is called *intermittent* if

$$\lim_{t \to \infty} \left(\frac{\Lambda_q(t)}{q} - \frac{\Lambda_p(t)}{p} \right) = \infty \quad \text{for} \quad 1 \leq p < q.$$

Remark: Let u be intermittent, $1 \leq p < q$ and $\Lambda_q(t)/q - f(t) \to \infty$ and $f(t) - \Lambda_p(t)/p \to \infty$. Then

$$\begin{aligned} &\langle u(t,0)^q \mathbb{1}\{u(t,0) \leq e^{f(t)}\} \rangle \\ &\quad \leq e^{qf(t)} = e^{qf(t) - \Lambda_q(t)} \langle u(t,0)^q \rangle = o(\langle u(t,0)^q \rangle). \end{aligned}$$

This shows that the main contribution to the q-th moment of $u(t,0)$ comes from the set $\{u(t,0) > e^{f(t)}\}$, whose complement, however, gives the main contribution to the p-th moment. As an informal conclusion, $u(t,\cdot)$ shrinks together to few small randomly located islands with high peaks, and the islands for the p-th moments are disjoint for different p.

Under the above assumptions, it was shown by Gärtner and Molchanov (1990):

Proposition 1 *If $\xi(0)$ is not constant a.s., then the solution u of (1.1) is intermittent.*

3 Moment Asymptotics

Theorem 1 (Gärtner and Molchanov (1996)) *If there is some $\theta \in [0,\infty]$ such that, for $c \in (0,1)$,*

$$\lim_{t\to\infty} \frac{H(ct) - cH(t)}{t} = \theta c \log c, \tag{3.1}$$

then, as $t \to \infty$, for every $p \in \mathbb{N}$,

$$\langle u(t,0)^p \rangle = \exp\left\{ H(pt) - 2d\kappa\chi\left(\frac{\theta}{\kappa}\right) pt + o(t) \right\},$$

where $\chi : [0,\infty] \to [0,1]$ is strictly increasing, concave and bijective.

Remarks:

(i) If $\gamma = \operatorname{esssup} \xi(0) < \infty$, then $H(t)/t \to \gamma$, and (3.1) is met with $\theta = 0$.

(ii) If $\xi(0)$ is double exponentially distributed with parameter $\theta \in (0,\infty)$, i.e.,

$$\langle 1\!\!1\{\xi(0) > r\}\rangle = \exp\{-e^{r/\theta}\}, \qquad r \in \mathbb{R},$$

then (3.1) is met. The density drops rapidly to zero outside $[-\theta,\theta]$.

(iii) If θ is big then ξ has high peaks; if κ is big then the local maxima of $u(t,\cdot)$ are spread out, "flattened".

Sketch of the proof:

We introduce p independent walks $(x_1(s))_{s\ge 0}, \dots, (x_p(s))_{s\ge 0}$ with local times

$$\ell_t^{(i)}(z) = \int_0^t 1\!\!1\{x_i(s) = z\}\, ds, \qquad i = 1,\dots,p;\ z \in \mathbb{Z}^d$$

and total normalized local time $L_t = \frac{1}{pt}\sum_{i=1}^p \ell_t^{(i)}$. Then we calculate

$$\begin{aligned}
&\langle u(t,0)^p\rangle e^{-H(pt)}\\
&= \mathbb{E}_0^p \left\langle \exp\left\{\sum_{z\in\mathbb{Z}^d}\sum_{i=1}^p \ell_t^{(i)}(z)\xi(z)\right\}\right\rangle e^{-H(pt)}\\
&= \mathbb{E}_0^p \exp\left\{pt\sum_{z\in\mathbb{Z}^d}\frac{H(ptL_t(z))-L_t(z)H(pt)}{pt}\right\}\\
&\approx \mathbb{E}_0^p \exp\left\{pt\theta\sum_{z\in\mathbb{Z}^d} L_t(z)\log L_t(z)\right\}\\
&\approx \exp\left\{pt\sup\left\{\theta\sum_{z\in\mathbb{Z}^d}\mu(z)\log\mu(z)-\kappa\mathcal{S}(\mu)\mid \mu\in\mathcal{P}_{\mathrm{c}}(\mathbb{Z}^d)\right\}\right\}.
\end{aligned}$$

where $\mathcal{S}(\mu)=\sum_{y\sim z}(\sqrt{\mu(y)}-\sqrt{\mu(z)})^2$ is the Donsker-Varadhan functional.

The main technical problem is the compactification of the space, since there is only a *weak* large deviation principle for the family $(L_t)_{t>0}$ of the total occupation times measures with rate function $\kappa\mathcal{S}$. In the above work, two different procedures are given to overcome this problem: The one is to cut the space into compact regular pieces and to work with the solution of the parabolic differential equation (1.1) in these boxes with *periodic* boundary conditions, and the other one uses a similar cutting procedure, but now with *zero* boundary condition (Dirichlet problem). □

4 Refinement: Correlations

The study of the solution of (1.1) was deepened by Gärtner and den Hollander (1997): Put $\varrho=\theta/\kappa$, then $\chi(\varrho)$ is given by

$$\chi(\varrho)=\frac{1}{2d}\inf_{\mu\in\mathcal{P}(\mathbb{Z}^d)}\left(\mathcal{S}(\mu)-\varrho\sum_{z\in\mathbb{Z}^d}\mu(z)\log\mu(z)\right).$$

If ϱ is large, then

$$\mu_\varrho(z_1,\dots,z_d)=\prod_{i=1}^d\frac{v_\varrho(z_i)^2}{\|v_\varrho\|_{\ell^2}^2}$$

is a minimizer, where $v_\varrho:\mathbb{Z}\to(0,\infty)$ is the solution of $\Delta v+2\varrho v\log v=0$ with minimal ℓ^2-norm.

Theorem 2 *If $\varrho=\theta/\kappa$ is large, then, for $x,y\in\mathbb{Z}^d$,*

$$\lim_{t\to\infty}\frac{\langle u(t,x)u(t,y)\rangle}{\langle u(t,0)^2\rangle}=\sum_{z\in\mathbb{Z}^d}\sqrt{\mu_\varrho(x+z)\mu_\varrho(y+z)}=c_\varrho(x,y).$$

Furthermore,

$$\lim_{\varrho\to\infty} c_\varrho(x,y) = \delta_{xy}, \qquad x,y\in\mathbb{Z}^d,$$

$$\lim_{\varrho\to 0} c_\varrho\left(\lfloor \tfrac{x}{\sqrt{\varrho}}\rfloor, \lfloor \tfrac{y}{\sqrt{\varrho}}\rfloor)\right) = e^{-\frac14|x-y|^2}, \quad x,y\in\mathbb{R}^d.$$

5 Continuous Version

We consider now the continuous version which is the same parabolic problem, but now for $t>0$ and $x\in\mathbb{R}^d$:

$$\begin{cases}\frac{\partial}{\partial t}u(t,x) = \kappa\Delta u(t,x) + \xi(x)u(t,x),\\ u(0,x)=1\end{cases} \tag{5.1}$$

Again $\xi=\{\xi(x)\colon x\in\mathbb{R}^d\}$ is a random potential, but now it is assumed to be Hölder-continuous and stationary. Again the function $H(t)=\log\langle e^{t\xi(0)}\rangle$ is assumed to be finite for all $t>0$. Under these assumptions, the minimal nonnegative solution is given by the Feynman-Kac formula

$$u(t,x) = \mathbb{E}_x \exp\left\{\int_0^t \xi(W_s)\,ds\right\},$$

where $(W_s)_{s\geq 0}$ is a d-dimensional Brownian motion, starting at $x\in\mathbb{R}^d$.

In order to state our assumption on the field ξ, we need some positive scaling function $\alpha(t)\to 0$ as $t\to\infty$. Put

$$\xi_t(x) = \alpha^2(t)(\xi(\alpha(t)x)-\xi(0))$$

and $\beta(t)=t/\alpha^2(t)$. We want to make an assumption on the manner in which way the field ξ shrinks together at zero, or, equivalently, how exponential moments of ξ_t, tested against probability measures with compact support, behave as t becomes large. We use the notation $(\mu,f)=\int_{\mathbb{R}^d} f(x)\,\mu(dx)$.

For $\mu\in\mathcal{P}_{\mathrm{c}}(\mathbb{R}^d)$ define $J_t(\mu)\in[0,\infty)$ by

$$\frac{\langle e^{t(\mu,\xi(\alpha(t)\cdot))}\rangle}{\langle e^{t\xi(0)}\rangle} = e^{-H(t)}\langle e^{t\xi(0)}e^{\beta(t)(\mu,\xi_t)}\rangle = e^{-\beta(t)J_t(\mu)}.$$

Assumption: There is a function $J:\mathcal{P}_{\mathrm{c}}(\mathbb{R}^d)\to[0,\infty)$ such that for $\mu_t,\mu\in\mathcal{P}_{\mathrm{c}}(\mathbb{R}^d)$ whose supports are contained in one compact set, we have:

$$\lim_{t\to\infty}\mu_t=\mu \text{ weakly} \quad\Longrightarrow\quad \lim_{t\to\infty} J_t(\mu_t) = J(\mu).$$

Theorem 3 (Gärtner and König, in preparation) *Define the convergence parameter*

$$\chi = \inf\{J(\mu)+\kappa\mathcal{S}(\mu)\,|\,\mu\in\mathcal{P}_{\mathrm{c}}(\mathbb{R}^d)\}>0,$$

then, as $t\to\infty$, for every $p\in[1,\infty)$,

$$\langle u(t,0)^p\rangle = \exp\{H(pt)-\beta(pt)(\chi+o(1))\}. \tag{5.2}$$

Here, $\mathcal{S}$ is the Donsker-Varadhan functional for the occupation times measures of d-dimensional Brownian motion.

Remark: If J is a function of the covariance matrix, i.e., $J(\mu) = j(\operatorname{cov}\mu)$ for some function j on the set $\mathcal{M}_d^+$ of the positive definite symmetric $d \times d$-matrices, then we have the less complex formula

$$\chi = \inf\left\{\frac{\kappa}{4}\operatorname{tr}(\Gamma^{-2}) + j(\Gamma^2) \,|\, \Gamma \in \mathcal{M}_d^+\right\}.$$

6 Examples

6.1 Gaussian potential

Let ξ be a stationary, homogeneous, Hölder-continuous Gaussian field with $\langle\xi(0)\rangle = 0$ and covariance function $B(x) = \langle\xi(x)\xi(0)\rangle$. Abbreviate $B(0) = \sigma^2$ and assume that $B(\cdot)$ is strictly maximal at zero, hence $-B''(0) = \Sigma^2 \in \mathcal{M}_d^+$.

Then $H(t) = t^2\sigma^2/2$, $\quad \alpha(t) = t^{-\frac{1}{4}}$, $\quad \beta(t) = t^{\frac{3}{2}}$ and

$$J(\mu) = \frac{1}{2}\int |\Sigma^2(x - m(\mu))|^2\,\mu(dx) = \frac{1}{2}\operatorname{tr}(\Sigma^2\operatorname{cov}\mu), \qquad \chi = \sqrt{\frac{\kappa}{2}}\operatorname{tr}\Sigma.$$

6.2 Poissonian potential

Let $\Phi(dx) =$ be the distribution of a Poisson process on $\mathbb{R}^d$, and let $\varphi \in C_{2,\mathrm{c}}(\mathbb{R}^d)$ be strictly maximal at zero with $\varphi(0) > 0$ and $\Sigma^2 = -\varphi''(0) \in \mathcal{M}_d^+$. Define the potential by $\xi(x) = \int_{\mathbb{R}^d} \varphi(y + x)\,\Phi(dy)$.

Then $H(t) = \int (e^{t\varphi(y)} - 1)\,dy$, $\quad \alpha(t) = t^{\frac{d}{8}} e^{-\frac{t}{4}\varphi(0)}$, $\quad \beta(t) = t^{1-\frac{d}{4}} e^{\frac{d}{2}\varphi(0)}$. Furthermore,

$$J(\mu) = \text{const.} \times \operatorname{tr}(\Sigma^2\operatorname{cov}\mu), \qquad \chi = \text{const.} \times \operatorname{tr}\Sigma.$$

References

[1] J. Gärtner and F. den Hollander, Correlation Structure of Intermittency in the Parabolic Anderson Model, preprint 1997

[2] J. Gärtner and W. König, Moment Asymptotics for the Continuous Anderson problem, in preparation

[3] J. Gärtner and S.A. Molchanov, Parabolic Problems for the Anderson Model. I. Intermittency and Related Topics, Commun. Math. Phys. **132**, 613-655 (1990)

[4] J. Gärtner and S.A. Molchanov, Parabolic Problems for the Anderson Model. II. Structure of High Peaks and Lifshitz Tails, to appear in PTRF, preprint 1996

TU Berlin, Str. des 17. Juni 136, 10623 Berlin, Germany, koenig@math.tu-berlin.de

On Risk Rates and Large Deviations in Finite Markov Chain Experiments

Peter Scheffel and Heinrich von Weizsäcker
Universität Kaiserslautern

1 Introduction

The observation of an ergodic Markov chain asymptotically allows perfect identification of the transition matrix. The information contained in the first n observations, provided the unknown transition matrix belongs to a known finite set is given by the entropy risk. It points out that the entropy risk tends to zero exponentially fast as $n \to \infty$. We are interested in this exponential rate. In the i.i.d. case it was determined in 1952 by H. Chernoff ([Che52]). For the determination of this rate we need to prove large deviations results for the empirical pair measure of ergodic Markov chains.

Let $\Theta = \{1, \ldots, l\}$ be a finite parameter set. For $\theta \in \Theta$ let P_θ be a probability measure on $\Omega = S^{\mathbb{N}_0}$. Let $(X_n)_{n \in \mathbb{N}_0}$ be the coordinate process. For $\omega = (\omega_0, \ldots, \omega_n) \in S^{n+1}$ we write $P_\theta^n(\omega) = P_\theta(X_0 = \omega_0, \ldots, X_n = \omega_n)$. If $\alpha = (\alpha_1, \ldots, \alpha_l)$ is a prior distribution then P_α^n denotes the mixed distribution $\sum_{\theta \in \Theta} \alpha_\theta P_\theta^n$. We introduce the posterior distribution $\hat{\alpha}^n(\omega)$ by the equation

$$\hat{\alpha}^n(\omega) = \left(\frac{\alpha_\theta \, P_\theta^n(\omega)}{P_\alpha^n(\omega)} \right)_{\theta \in \Theta}.$$

Definition 1: *Let $(X_i)_{i \in \mathbb{N}}$ and P_θ, $\theta \in \Theta$ be as before. The entropy risk of α given $(X_0, \ldots, X_n)$ is defined by*

$$H_n(\alpha|X) := \sum_{\omega \in S^{n+1}} P_\alpha^n(\omega) \left(- \sum_{\theta \in \Theta} \hat{\alpha}_\theta^n(\omega) \log \hat{\alpha}_\theta^n(\omega) \right).$$

Similarly

$$B_n(\alpha|X) := \sum_{\omega \in S^{n+1}} P_\alpha^n(\omega) \left(1 - \max_{\theta = 1, \ldots, l} \hat{\alpha}_\theta^n(\omega) \right)$$

is called minimal Bayes risk given $X_0, \ldots, X_n$.

The following Proposition states that in the setting of ergodic Markov chains entropy risk and minimal Bayes risk tends to zero exponentially fast with the same exponential rate. The proof is based on [KW97].

Proposition 1: *Let $\alpha = (\alpha_1, \ldots, \alpha_l)$ be any strictly positive prior. Suppose that for every $\theta \in \Theta = \{1, \ldots, l\}$ the process $(X_i)_{i \in \mathbb{N}_0}$ forms under P_θ a Markov chain with the irreducible transition matrix π_θ where $\theta \mapsto \pi_\theta$ is injective. Then*

a) $\lim_{n\to\infty} B_n(\alpha|X) = \lim_{n\to\infty} H_n(\alpha|X) = 0$

b) $\lim_{n\to\infty} \frac{1}{n} \log H_n(\alpha|X) = \lim_{n\to\infty} \frac{1}{n} \log B_n(\alpha|X) = \lim_{n\to\infty} \frac{1}{n} \log B_n(\alpha'|X)$ *for every other positive prior α', provided one of these limits exist.*

c) Let $r(P_1, \ldots, P_l)$ denote the limit in b). Then $r(P_1, \ldots, P_l)$ is determined by the two parameter subexperiments as follows

$$r(P_1, \ldots, P_l) = max_{1 \le \vartheta \ne \eta \le l}\, r(P_\vartheta, P_\eta)$$

In the i.i.d. case c.) was explicitly shown by Torgersen in [Tor81]. Before we are able to state the main Theorem some more definitions will be needed.
Let S be a finite set. $\mathbb{N}$ is the set of integers and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. We denote by $\mathcal{M}$ the set of all probability distributions on S^2, by $\mathcal{M}_s$ the set of all stationary distributions on S^2, i.e. those probability distributions $Q = (Q_{ij})_{i,j\in S}$ on S^2 whose two marginal distributions coincide. For $Q = (Q_{ij})_{i,j\in S} \in \mathcal{M}_s$ the marginal (i.e. the vector of row -or equivalently column- sums) is denoted by $(Q_i)_{i\in S}$. If $Q \in \mathcal{M}_s$ and $\pi = (\pi(i,j))_{i,j\in S}$ is a transition matrix the symbol $Q \times \pi$ denotes the element of $\mathcal{M}$ defined by $(Q \times \pi)_{ij} = Q_i \pi(i,j)$. For a nonnegative $S \times S-$matrix λ let $\mathcal{M}^\lambda := \{Q \in \mathcal{M} |\, Q \ll \lambda\}$.
By $\mathcal{M}_{s,i}$ we denote the set of all stationary probability distributions on S^2 with the property that there exists a set $I \subset S$ with $Q_{ij} = 0$ whenever $i \notin I$ or $j \notin I$, such that for all $i,j \in I$ there exists an $n \in \mathbb{N}$ with $Q^n_{ij} > 0$.

Let λ be a nonnegative $S \times S$-matrix. We write $i \underset{\lambda}{\rightsquigarrow} j$ (and say i leads to j) iff $\lambda^n(i,j) > 0$ for some $n \in \mathbb{N}$ and $i \underset{\lambda}{\sim} j$ (i communicates with j) iff $i = j$ or ($i \underset{\lambda}{\rightsquigarrow} j$ and $j \underset{\lambda}{\rightsquigarrow} i$). An equivalence class E of the relation $\underset{\lambda}{\sim}$ is called nontrivial if either E contains at least two elements or $E = \{i\}$ with $\lambda(i,i) > 0$.

Definition 2: *We call a set $I \subset S$ a* **$\lambda-$string** *if the following holds: there exist a tuple $(E_1, \ldots, E_k)$ of nontrivial λ-equivalence classes with $I = E_1 \cup \ldots \cup E_k$, such that for every $l \in \{1, \ldots, k-1\}$ there are points $i_l \in E_l,\ j_{l+1} \in E_{l+1}$ with $i_l \underset{\lambda}{\rightsquigarrow} j_{l+1}$.*

Since no state $i \in E_{l+1}$ leads to a state $j \in E_l$ the order of $(E_1, \ldots, E_k)$ is uniquely determined. So a string is the union of a sequence of $\underset{\lambda}{\sim}$ -equivalence classes which can be visited successively by a single path. We want to make clear our definition by the following example.

Example 1:
Consider the following 7×7*-matrix* ($*$ *denotes a positive entry)*

$$\lambda = \begin{pmatrix} 0 & 0 & 0 & 0 & * & 0 & 0 \\ 0 & * & * & 0 & 0 & 0 & 0 \\ 0 & * & * & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & * & * & 0 & 0 \\ * & 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The nontrivial $\underset{\lambda}{\sim}$ *-equivalence classes are* $E_1 = \{2,3\}, E_2 = \{4,5\}, E_3 = \{6\}$ *(*$\{1\}$ *and* $\{7\}$ *are trivial classes). As* λ*-strings we get* $E_1 \cup E_2 = \{2,3,4,5\}$ *because* $\lambda(3,4) > 0$*,* $E_3 \cup E_2 = \{6,4,5\}$ *since* $\lambda^2(6,5) = \lambda(6,1)\lambda(1,5) > 0$ *and of course* E_1, E_2 *and* E_3*.*

2 Main Result

For $x \in S$ let $\mathcal{S}^\lambda(x) := \{I \subset S | I = E_1 \cup \ldots \cup E_k,$ is a $\lambda-$string, with $x \underset{\lambda}{\rightsquigarrow} y$ for some $y \in E_1\}$. For a measure μ on S let $\mathcal{S}^\lambda(\mu) = \{I | I \in \mathcal{S}^\lambda(x),\, \mu(x) > 0\}$. For $I \subset S$ we define further $\mathcal{M}_s^\lambda(I)$ as the set $\{Q \in \mathcal{M}_s \cap \mathcal{M}^\lambda | Q$ is concentrated on $I \times I\}$, $\mathcal{MS}^\lambda(x) := \cup\{\mathcal{M}_s^\lambda(I) | I \in \mathcal{S}^\lambda(x)\}$ and $\mathcal{MS}^\lambda(\mu) := \cup\{\mathcal{MS}^\lambda(x) |\, \mu(x) > 0\}$.

Remark 1: $\mathcal{MS}^\lambda(\mu)$ is the set of those stationary probability measures $Q \in \mathcal{M}$ which can be dominated by the empirical pair measure of a path of a Markov chain whose transitions are possible according to λ and which starts in a point of positive μ-measure. Of course $\mathcal{MS}^\lambda(\mu) = \emptyset$ if there are no λ-strings.
I is an element of $\mathcal{S}^\lambda(\mu)$ if there exist a path starting in x, $\mu(x) > 0$ whose transitions are allowed by λ and which reaches every $\underset{\lambda}{\sim}$ -equivalence class E_k in I.

For a λ-string $I = E_1 \cup \ldots \cup E_k$ we set $\lambda^I := \sum_{l=1}^k \lambda^{E_l}$ where λ^{E_l} is the $S \times S$ matrix which equals λ on $E_l \times E_l$ and is zero otherwise.
Now we are able to formulate our main result. For simplicity we write $K_\theta(Q) = H(Q|Q \times \pi_\theta)$ for $\theta = 0,1$.

Theorem 1: *Let* (P_0, P_1) *be a binary experiment of irreducible Markov chains with different transition matrices* π_0, π_1 *and initial distributions* μ_0, μ_1*. Let*

$$\pi := \pi_0 \wedge \pi_1 = (\min(\pi_0(i,j), \pi_1(i,j)))_{i,j \in S}$$

and

$$\mu := \mu_0 \wedge \mu_1 = (\min(\mu_0(i), \mu_1(i)))_{i \in S}.$$

Then the rate of the entropy risk is given by

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{n}\log H_n(\alpha|X) &= \lim_{n\to\infty}\frac{1}{n}\log B_n(\alpha|X) && (1)\\
&= -\inf_{Q\in\mathcal{MS}^\pi(\mu)} K_0(Q)\vee K_1(Q) && (2)\\
&= \max_{I\in\mathcal{S}^\pi(\mu)}\inf_{0<t<1}\log\rho(\pi_t^I) && (3)
\end{aligned}$$

where $\pi_t(i,j)=\pi_0(i,j)^t\pi_1(i,j)^{1-t}$ *and* ρ *denotes the spectral radius.*

For the proof of Theorem 1 large deviations results for the empirical pair measure of ergodic Markov chains are necessary. They will be provided in the following section.

3 Large Deviations Results

In order to be able to allow zero entries in the transition matrices it is convenient to equip $\mathcal{M}$ with two topologies, the topology τ of coordinatewise convergence and the following topology τ_0 which is finer than τ : A set $U\subset\mathcal{M}$ is τ_0-open iff for every $Q\in U$ there is a $\delta>0$ such that the set

$$U_\delta(Q):=\{P\in\mathcal{M}:P\ll Q \text{ and } |P_{ij}-Q_{ij}|<\delta,\ \text{for all } i,j\}$$

is contained in U. Thus every point $Q\in\mathcal{M}$ has a τ_0-neighbourhood of points which have the same zeroes as Q.
The main point is to get lower bounds for sets which have empty interior in the topology τ. As a first step (Theorem 2) we consider open sets in the finer topology τ_0. This implies the large deviation Theorem 3 based on the λ-strings introduced above. Moreover in Theorem 4 we get large deviation estimates for the probability that the empirical pair measure is in certain thin subsets of the set of stationary measures. Let

$$\hat{m}_n(\omega)_{ij}=\frac{1}{n}\sum_{k=0}^{n-1}1_{\{\omega_k=i,\ \omega_{k+1}=j\}}(\omega)\text{ for } i,j\in S.$$

and P_x^π be the law of the Markov chain with transition matrix π conditioned on start in x.

Theorem 2: *Let* P^π *be the law of a Markov chain with the irreducible transition matrix* π. *Let* $U\subset\mathcal{M}^\pi$ *be open in the topology* τ_0. *Then for all* $x\in S$

$$\liminf_{n\to\infty}\frac{1}{n}\log P_x^\pi(\hat{m}_n\in U)\geq-\inf_{Q\in U_x\cap\mathcal{M}_{s,i}} H(Q|Q\times\pi)$$

where $U_x=\{Q\in U\,|\,Q_x>0\}$.

The following Theorem shows the use of strings in connection with large deviations.

Theorem 3: *Let P^π be the law of a Markov chain with the irreducible transition matrix π. Let λ be a matrix with $\lambda \ll \pi$, $x \in S$.*
a) If U is a relatively open subset of $\mathcal{M}^\lambda$ in the topology of coordinatewise convergence then

$$\liminf_{n\to\infty} \frac{1}{n} \log P_x^\pi(\hat{m}_n \in U) \geq - \inf_{U \cap \mathcal{MS}^\lambda(x)} H(Q|Q\times\pi).$$

b) If $U \subset \mathcal{M}^\lambda$ is closed in the topology of coordinatewise convergence then

$$\limsup_{n\to\infty} \frac{1}{n} \log P_x^\pi(\hat{m}_n \in U) \leq - \inf_{U \cap \mathcal{MS}^\lambda(x)} H(Q|Q\times\pi).$$

Similarly one gets new bounds if the set U is contained in $\mathcal{M}_{s,i}$. The set of irreducible measures is τ_0 open in contrast to the set $\mathcal{M}_{s,i}$. Therefore the estimates of the next result do not follow directly from Theorem 2. Since the empirical pair measure can be stationary only if the path has completed a loop the probability in the following result is nonzero only on a periodic set of times. Therefore the liminf has to be taken along a suitable subsequence of the integers.

Theorem 4: *Let P^π be the law of a Markov chain with irreducible transition matrix π.*
a) Let U be a relatively τ_0-open subset of $\mathcal{M}_{s,i}$ and let d be the least common multiple of the set of periods $\{d_Q : Q \in U\}$. Then with $U_x = \{Q \in U \mid Q_x > 0\}$ we have

$$\liminf_{n\to\infty} \frac{1}{nd} \log P_x^\pi(\hat{m}_{nd} \in U) \geq - \inf_{U_x} H(Q|Q\times\pi).$$

b) Let U be a subset of $\mathcal{M}_{s,i}$. Then

$$\limsup_{n\to\infty} \frac{1}{n} \log P_x^\pi(\hat{m}_n \in U) \leq - \inf_{U_x} H(Q|Q\times\pi).$$

For details we refer to [SW97].

References

[Che52] H. Chernoff. A measure of asymptotic efficiency for tests of a hypothesis on the sum of observations. *Annals of Mathematical Statistics*, 23:493–507, 1952.

[KW97] J. Krob and H.v. Weizsäcker. The Rate of Information Gain in Experiments with a Finite Parameter Set. *Statistics and Decisions*, 15, 1997.

[Tor81] E.N. Torgersen. Measures of Information based on Comparison with total Information and with total Ignorance. *Ann. of Statistics*, 9:638–657, 1981.

[SW97] P. Scheffel and H.v. Weizsäcker. On Risk Rates and Large Deviations in Finite Parameter Markov Chain Experiments. *Math. Methods of Statistics*, 3:293–312, 1997.

ESTIMATING THE VALUE OF THE WINCAT COUPONS OF THE WINTERTHUR INSURANCE CONVERTIBLE BOND*

UWE SCHMOCK**

ABSTRACT. The three annual $2\,{}^1\!/_4\%$ interest coupons of the Winterthur Insurance convertible bond (face value CHF 4 700) will only be paid out if during their corresponding observation periods no major storm or hail storm on one single day damages more than 6 000 motor vehicles insured with Winterthur Insurance. Data for events, where storm or hail damaged more than 1 000 insured vehicles, are available for the last ten years. Using a constant-parameter model, the estimated discounted value of the three WINCAT coupons together is CHF 263.29. A conservative evaluation, which accounts for the standard deviation of the estimate, gives a coupon value of CHF 238.25. However, fitting models which admit a trend or a change-point, leads to substantially higher knock-out probabilities of the coupons. The estimated discounted values of the coupons can drop below the above conservative value; a conservative evaluation as above leads to substantially lower values. Hence, already the model uncertainty is higher than the standard deviations of the used estimators. Consistency, dispersion, robustness and sensitivity of the models can be analysed by a simulation study.

1. INTRODUCTION

The Swiss insurance company Winterthur Insurance has launched a three-year subordinated $2\,{}^1\!/_4\%$ convertible bond with so-called WINCAT coupons, where CAT is an abbreviation for catastrophe. This bond with a face value of CHF 4 700 may be converted into five Winterthur Insurance registered shares[1] at maturity (European-style option) between the 18^{th} and 24^{th} of February 2000. The annual interest coupon of $2\,{}^1\!/_4\%$ will *not* be paid out if on any one calendar day during the corresponding observation period for the coupon more than 6 000 motor vehicles insured with Winterthur in Switzerland are damaged by hail or storm (wind speeds of 75 km/h and over). If the number of insured motor vehicles changes by more than 10%, then the knock-out limit of 6 000 claims will be adjusted correspondingly.

Had Winterthur launched an identical fixed-rate convertible bond, then, according to Crédit Suisse First Boston's brochure [1], the coupon rate would have been

---

1991 *Mathematics Subject Classification.* 62P05 (primary); 90A09 (secondary).

Key words and phrases. WINCAT coupon, Winterthur Insurance, catastrophe bond, storm, hail, model risk, (generalised) Pareto distribution, composite Poisson model, generalised linear model, change-point, peaks over threshold.

*See [4] for the full version of the paper.

**Financial support by Crédit Suisse is gratefully acknowledged.

[1] Due to the merger of Winterthur Insurance and Crédit Suisse Group on December 15^{th}, 1997, the bond may be converted into 36.5 Crédit Suisse Group registered shares at maturity. Due to the conversion right and the rising market value of the Winterthur Insurance registered shares (see [8]), the convertible bond offered a good investment opportunity during its first few months.

around 0.76% lower (approximately 1.49%). In other words, the investor receives an annual yield premium of 0.76% for bearing a portion of Winterthur's damage-to-vehicles risk. This convertible bond is intended as an instrument to diversify portfolios. The WINCAT coupons are very suitable for this purpose, because storm and hail damages have only a very small correlation with traditional financial market risk. The European-style conversion right, however, strongly ties the bond to the financial market. It is the intention of Winterthur Insurance to test the Swiss capital market for such products, make investors acquainted with them, and obtain a partial reinsurance through the financial market by securitizing a portion of its damage-to-vehicles risk.

Within the range of designs of catastrophe bonds, the Winterthur Insurance convertible bond with WINCAT coupons "Hail" belongs to the more conservative ones, namely the principal-protected catastrophe bond. Besides the pure catastrophe bonds, where the coupons and the principal are at risk, another more conservative variant are the deferred catastrophe bonds, where no payment as such is at risk, but the payments may be deferred. This gives the issuer of such a bond an interest-free credit in case of a catastrophe.

Two guiding principles for specifying the conditions of the WINCAT coupons were simplicity and absence of moral hazard. For the purpose of reinsurance, it would have been interesting for Winterthur Insurance to include a knock-out limit connected to the total number of claims during an observation period. To reduce moral hazard, damage arising from a natural cause was chosen as the triggering event, and the knock-out limit is tied to the number of claims and not to the capital necessary to pay full indemnity to the insured. If an event with more than 6 000 claims occurs, then Winterthur Insurance saves the corresponding $2\,{}^1\!/_4\%$ coupon interest payment on 399.5 million Swiss francs, which makes CHF 8 988 750 at the corresponding coupon date. On the other hand, according to Winterthur Insurance, CHF 3 000 have to be paid out per claim on the average for motor vehicles damaged by storm or hail. Therefore, when an event with more than 6 000 claims occurs, Winterthur Insurance can expect to save up to 50% by means of the WINCAT coupons—a profit from a knock-out event seems extremely unlikely. A possible problem with the knock-out limit can be borderline cases of events with about 6 000 claims when a few insured do not know the exact date of the damage (because they have been on holiday, for example). A way to moderate the severity of such a problem would be a linear reduction of the coupon interest rate from $2\,{}^1\!/_4\%$ to 0% between 5 000 and 7 000 claims. However, such a specification would make the product more complex and the statistical analysis for the coupon pricing more involved.

The study [4] was made possible by the willingness of Winterthur Insurance to collect and *publish* the relevant available historical data on the web page [7] as well as in Crédit Suisse First Boston's brochure [1] and thereby to set standards in product transparency, fairness of pricing and investor education. This enables a scientific discussion of such products and their corresponding pricing methodologies, which in turn helps to enhance transparency and acceptance of such products. To satisfy this aim and to build up the confidence of investors, the various sources of risk of such new products should be made explicit to avoid unpleasant surprises. The paper [4] seeks to make a contribution in this direction. Since convertible bonds are well-established securities in the market, a lot of information concerning Winterthur

Insurance in contained in the legally binding prospectus [6], which helps the investor to judge the default risk and the possible profits from the European-style conversion right. However, no information (other than the exact legal specification) for estimating the knock-out probability of the coupons is given in this legally binding prospectus; in particular, there is no historical data on the subject in the prospectus. Apparently, Winterthur Insurance and Crédit Suisse First Boston have been aware of this deficiency; hence their decision to publish [1] and make the historical data available on the web page [7].

The paper [4] focuses on estimating the risk arising from the WINCAT coupons, with emphasis on the model risk which is not addressed in [1]. Based on the available historical data, several models are presented and worked out and the discounted value of the WINCAT coupons is calculated in every case for an easy comparison of the various results. The pricing of the European-style option for converting the bond into Winterthur Insurance registered shares can be found in [1]. We should mention here that the current value of the call option depends on the knock-out probability of the last coupon, because the exercise price of the call option is either CHF 4 805.75 (face value of the bond plus last coupon), if the last coupon is paid, or simply the face value of CHF 4 700, if the last coupon is knocked out.

When modelling low-frequency event risks, the scarcity of the available statistical data is a typical problem. If one wants to follow a kind of Bayesian approach, it is desirable to take additional information into account when selecting a model (see [2, Section 6] for such a case study of the correlation of wind storm losses of the Swedish insurance group Länsförsäkringar with wind speed data provided by the Swedish Meteorological and Hydrological Institute). For a fair and transparent pricing of financial products, such information should either be public or should be published together with the introduction of the financial product. For the pricing of the WINCAT coupons, such additional information besides the historical data of Table 2.2 is contained in the study [9] of Winterthur Insurance. This study, as well as the public available report [3], for example, provide information on the variability of the weather; they also describe the development of hail storms, the different frequency of hail storms in the various parts of Switzerland, and the properties of hailstones (size, shape, speed) that cause damage to motor vehicles. The study [9] also points out that damages to agriculture and motor vehicles are mainly caused by different types of hail storms: damage to motor vehicles requires a large momentum of the hailstones (large product of mass and speed), while damage to agriculture can already be caused by small but numerous hailstones. This indicates that the extensive statistical data collected from insured damages to agriculture since 1881 is of limited use when estimating a possible trend in the frequency or severity of damages to motor vehicles caused by hail. Also in the report [3], the severity of hail storms is measured by the number of communities reporting damages to agriculture.

If the knock-out probability P_{CAT} for the WINCAT coupons were known exactly, then a very small risk premium for the investor would suffice, because the investor has the freedom to invest only a small fraction of the capital in the Winterthur Insurance convertible bond thereby diversifying the risk. This small risk premium is the motivation for insurance companies to securitize their catastrophe risk. However, the true knock-out probability P_{CAT} is not known. Therefore, at various places in the paper [4], we follow the procedure used in [1] and add an estimated standard deviation $\hat{\sigma}(P_{\text{CAT}})$ to the estimated knock-out probability to obtain a conservative upper estimate of the knock-out probability thereby adding a risk premium for the investor to account for the uncertainty of P_{CAT}. We could elaborate on this point by using the entire estimated distribution of P_{CAT} and tilt it towards higher values (the paper [5] by G. G. Venter is interesting in this context). Taking investor-dependent utility functions and the current market price of risk into account, a more profound analysis might be possible than the one sketched above. However, since the estimated knock-out probabilities and the corresponding standard deviations will vary substantially with the models used, the model risk should also be taken into account, because is seems to be the dominating one in the present problem. There should be a coherent way to calculate an adequate risk premium which accounts for the variation of the estimated knock-out probability *and* the corresponding model risk. We leave it to future research to develop a rigorous mathematical basis for this purpose and to apply it to the present problem.

2. Presentation and discussion of the data

To estimate the risk of the Wincat coupons, a 10-year history of damage claims is provided in [1] and [7], see Table 2.2. During this period, a total of 17 events with more than 1 000 damaged vehicles were registered. Of these events, 15 happened during the summer and two were winter storms. None of these events occurred between 1987 and 1989. Only two of the events, which happened on the 21st of July 1992 and the 5th of July 1993, caused more than 6 000 claims. Without any sophisticated modelling, this suggests a knock-out probability of 20%, i. e., the expectation of the annual coupon payment would be 80% of the 2 1/4% Wincat coupon, which is an expected annual yield of 1.8%. Of course, as mentioned in [1, p. 11], this estimate has little statistical significance.

Whether a Wincat coupon is paid on February 28th depends on the events happening during the corresponding observation period specified in Table 2.1. The first observation period is shorter than a year so that there are always four months left between the end of the observation period and the coupon payment date. This provides enough time to count the number of claims and to determine whether the corresponding coupon is knocked out. In the 10-year history of damage claims, two events are not within the period from February 28th to October 31st. This is relevant for the first coupon, we shall therefore always reduce the knock-out probability for the first coupon in a deterministic way using the formula

$$P_{\mathrm{CAT}} = 1 - (1 - \tilde{P}_{\mathrm{CAT}})^{15/17}, \tag{2.1}$$

where $\tilde{P}_{\mathrm{CAT}}$ denotes here the knock-out probability if the observation period were a full year. Formula (2.1) is motivated by the used Poisson models, it corresponds to reducing the Poisson parameter by the factor $15/17$. By using (2.1), we neglect

Coupon date	Relevant observation period
February 28, 1998	February 28, 1997 – October 31, 1997
February 28, 1999	November 1, 1997 – October 31, 1998
February 28, 2000	November 1, 1998 – October 31, 1999

Table 2.1. Observation periods for the Wincat coupons according to [6] and the web page [7] of Winterthur Insurance.

Year	Date	Event	Number of claims	Vehicles insured index	Adjusted claims
1987				1.248	
1988				1.204	
1989				1.161	
1990	27. Feb.	Storm	1 646	1.127	1 855
	30. June	Hail	1 395		1 572
1991	23. June	Hail	1 333	1.104	1 472
	6. July	Hail	1 114		1 230
1992	21. July	Hail	8 798	1.098	**9 660**
	31. July	Hail	1 085		1 191
	20. Aug.	Hail	1 253		1 376
	21. Aug.	Hail	1 733		1 903
1993	5. July	Hail	6 589	1.099	**7 241**
1994	2. June	Hail	4 802	1.086	5 215
	24. June	Hail	940		1 021
	18. July	Hail	992		1 077
	6. Aug.	Hail	2 460		2 672
	10. Aug.	Hail	2 820		3 063
1995	26. Jan.	Storm	1 167	1.067	1 245
	2. July	Hail	1 290		1 376
1996	20. June	Hail	1 262	1.000	1 262

TABLE 2.2. Claim numbers of past events which caused over 1 000 adjusted claims as provided in [1] and [7]. In 1987, 1988 and 1989, such events did not occur. Since the number of motor vehicles insured with Winterthur in Switzerland tends to increase, former actual claim numbers are set into relation with the number of insured vehicles to obtain the number of adjusted claims.

the fact that the number of events not occurring in the period from February 28$^{\text{th}}$ to October 31$^{\text{st}}$ is random as well. This simplification, however, is suggested by the lack of data and can be justified by the small influence of this $^{15}/_{17}$-correction (CHF 2.21 for $\tilde{P}_{\text{CAT}} = 20\%$, for example) when compared with the model uncertainty to be discussed. Furthermore, when analysing the numbers of adjusted claims, we assume that the two numbers arising from the winter storms come from the same

Coupon	Discount rate	Discount factor
1.	1.87%	0.9816
2.	2.33%	0.9550
3.	2.57%	0.9267

TABLE 2.3. Assumptions regarding the interest-rate structure taken from [1]. The discount rates correspond to the zero-coupon yield on Swiss Confederation bonds plus a spread of 35 basis points.

underlying distribution as the numbers arising from the hail storms. Again, this simplifying assumption is suggested by the small historical data set.

The number of claims arising from damage by storm or hail have to be set into relation with the number of vehicles insured with Winterthur in Switzerland. The statistical basis is 773 600 insured motor vehicles in 1996. This number includes the motor vehicles insured with Neuenburger Schweizerische Allgemeine Versicherungsgesellschaft, which merged with Winterthur Insurance in 1997. The column *Vehicles insured index* in Table 2.2 gives the number of insured vehicles in 1996 divided by the number of insured vehicles for the respective year. The column *Adjusted claims* in Table 2.2 contains the claim numbers multiplied with the insured-vehicles index. Only events with more than 1 000 adjusted claims are shown in Table 2.2, because other historical data is not provided by Winterthur Insurance.

If the statistical basis of 773 600 insured motor vehicles changes by more than 10%, then, according to [6, Condition 2(e)], the knock-out limit of 6 000 claims will be adjusted accordingly, rounded to the nearest multiple of 100 claims.[2] As the column *Vehicles insured index* of Table 2.2 shows, the statistical basis tends to increase, but it seems unlikely that it reaches the adjustment trigger of 10% within three years without a merger with another insurance company. Apparently, such a scenario slightly increases the risk of the investor. On the other hand, there was a recent change in the Swiss legislation concerning the mandatory motor vehicle insurance, and new competitors are becoming active in the motor vehicles insurance market. Therefore, it is not clear whether the rising trend in the statistical basis will persist. For the further analysis, we assume that the statistical basis stays constant. It should be kept in mind however, that (depending on the model) the estimated coupon values in Table 3.1 can change by up to CHF 10 if the statistical basis changes by as much as ±10% already in the first observation period.

3. Possible models and the corresponding estimated values

In this section we very briefly discuss some of the models which can be used to estimate the discounted value of the three WINCAT coupons. The corresponding results are given in Table 3.1. For a deeper analysis of the models, for statistical tests which examine the significance of time-inhomogeneity, and for conservative estimates of the values of the WINCAT coupons, we refer to [4].

The binomial model already mentioned at the beginning of Section 2 uses the data of Table 2.2 very inefficiently. Already in the first step, the data is reduced to ten yes/no decisions (one bit of information for every year). By taking the mean, this information is further reduced by ignoring the order of the ten yes/no decisions, leading to one out of the eleven possible numbers 0, $1/10$, $1/5$, $3/10$, $2/5$, $1/2$, $3/5$, $7/10$, $4/5$, $9/10$ and 1 for the estimated knock-out probability. This is less then 4 bit of information. Having gone through this bottleneck, not much can be done with a statistical examination afterwards.

To extract more data from Table 2.2, we can use a constant-intensity composite Poisson model. Such models allow to fit the distributions of the event frequency and the event severity separately. For every calendar day in an observation period there is a slight chance of a major storm or hail storm causing more than 1 000 adjusted claims. The data of Table 2.2 as well as common knowledge suggest that

[2] Such an adjustment is based on the number of insured motor vehicles at April 1$^{\text{st}}$.

No.	Coupon value	Corresponding model
1	CHF 244.44	Binomial model
		Constant-parameter Poisson model and a
2	CHF 267.48	— *generalised Pareto distribution*
3	CHF 263.29	— *Pareto distribution*
4	CHF 247.37	— *Bernoulli distribution*
		Peaks-over-threshold model and a
5	CHF 264.00	— *linear trend of the location parameter*
6	CHF 204.34	— *log-linear trend in the scale and location parameter*
		Change-point Poisson model and a
7	CHF 253.80	— *generalised Pareto distribution*
8	CHF 247.99	— *Pareto distribution*
9	CHF 225.28	— *Bernoulli distribution*
		Generalised Pareto distribution and time-dependent Poisson parameter with a
10	CHF 223.88	— *linear trend*
11	CHF 220.53	— *smooth transition*
12	CHF 214.44	— *modified-linear trend*
13	CHF 214.37	— *log-linear trend*
14	CHF 210.86	— *square-root linear trend*
		Pareto distribution and a time-dependent Poisson parameter with a
15	CHF 215.19	— *linear trend*
16	CHF 211.54	— *smooth transition*
17	CHF 204.96	— *modified-linear trend*
18	CHF 204.93	— *log-linear trend*
19	CHF 201.12	— *square-root linear trend*
		Bernoulli distribution and a time-dependent Poisson parameter with a
20	CHF 189.56	— *linear trend*
21	CHF 185.11	— *smooth transition*
22	CHF 177.36	— *modified-linear trend*
23	CHF 177.44	— *log-linear trend*
24	CHF 172.87	— *square-root linear trend*

TABLE 3.1. Comparison of the sum of the three discounted WINCAT coupon values arising from the binomial model, the various composite Poisson models and the two extensions of the peaks-over-threshold model. No explicit risk premium is included. The various estimated coupon values clearly illustrate the model risk.

this slight chance varies with the season: In Switzerland, a storm is more likely to occur in late autumn or winter than in any other season while hail storms usually occur in summer. If the dependence between the different days is sufficiently weak, then the Poisson limit theorem suggests that a Poisson random variable might be a good approximation for the number of those events within an observation period, which cause more than 1 000 adjusted claims. The data of Table 2.2 leads to the estimate $\lambda_{1000}^{\text{const}} = {}^{17}\!/_{10}$ for the Poisson parameter.

Note that Table 2.2 records hail storms for August 20$^{\text{th}}$, 1992, and the following day, hence the assumption of "sufficiently weak dependence" has to be kept in mind. Such two-day events can arise artificially from a single storm due to the dividing line at midnight, or they can arise due to weather conditions favouring a hail storm on two consecutive days. The use of a compound Poisson model however, which allows us to model such two-day events conveniently, does not seem to be appropriate here, because a single observation is not sufficient for a reliable estimate of the corresponding parameter.

The seasonal dependence mentioned above is also the reason why we have chosen the exponent ${}^{15}\!/_{17}$ in the correction formula (2.1). We think that this exponent based on the available data is more appropriate then the exponent ${}^{2}\!/_{3}$ based on the length of the shorter first observation period given in Table 2.1.

Several distributions can be used to obtain an estimate for the probability that an event causing more than 1 000 adjusted claims actually leads to the knock-out of the coupon. These distributions include the Bernoulli distribution, the Pareto distribution (used in [1]) and finally, as suggested by extreme value theory, the generalised Pareto distribution. The corresponding estimated probabilities are $\hat{p}_{6000} = {}^{2}\!/_{17} \approx 0.118$, $\hat{p}_{6000} \approx 0.0857$ and $\hat{p}_{6000} \approx 0.0757$, respectively. The corresponding estimated coupon values are given in Table 3.1.

The constant-parameter models are static ones. They give equal weight to every recorded event and, by construction, do not allow to discover a trend in the data. Every redistribution of the 17 events in Table 2.2 to the ten observation periods would lead to the same result for the coupon values (if we disregard the ${}^{15}\!/_{17}$-correction). However, the tests for over-dispersion and time-inhomogeneity in [4] suggest—although not significantly on the 5 %-level but very close—to consider the possibility of a time-dependent distribution for the event frequency. Such a time-dependence can account for the tendency of over-dispersion considered in [4, Subsection 5.1]. In particular, an investor might want to take a possible trend into account when estimating the discounted value of the WINCAT coupons. Even when a constant-parameter model is preferred for pricing the WINCAT coupons, a model capable to accommodate a possible trend can be useful for risk management. There are several reasons why there might be a trend, for example:

- The variability of the weather could change, due to human influence (increased CO_2-part in the atmosphere) or solar activity, for example.
- Winterthur Insurance might increase its market share in other regions like the French or Italian speaking parts of Switzerland; this can happen in particular when Winterthur merges with another insurance company (like merging with Neuenburger Schweizerische Allgemeine Versicherungsgesellschaft in 1997, for example). Due to the Swiss Alps, the local climate is in general quite different in different regions of Switzerland, so a change in Winterthur's engagement in a particular region can considerably increase or decrease the company's exposure to storm or hail damages.
- Severe damage caused by hail is a local event. If the density of motor vehicles insured with Winterthur increases (due to more cars per inhabitant, more inhabitants per area or a greater market share of Winterthur Insurance), then more insured motor vehicles are likely to be damaged in every single event.

- The habits of the insured might change. They might buy a second or third car for the family without building or renting an additional garage to protect the car in case of bad weather. Or the insured are better off financially and they can afford the deductable, hence they take chances and don't drive the car to a secure place in case of a storm/hail forecast.
- Motor vehicles might get more or less susceptible of hail damage, because the material changes (different kinds of steel, aluminium, different coats of lacquer, for example) or the thickness of the automobile body sheet changes (a thicker sheet can give more protection in case of an accident, a thinner sheet reduces weight and thereby fuel consumption).

In any case—whatever the particular reason—it is a reasonable idea to consider a model which is flexible enough to take a possible trend in the data into account as long as such a possible trend can not be ruled out by additional information concerning all the points mentioned above (and the ones we have not thought of). In [4] we consider and discuss the advantages and drawbacks of five different trend models for the Poisson parameter describing the event frequency: A linear trend, a log-linear trend, a square-root linear trend, a modified-linear trend, and, finally, a smooth transition of the Poisson parameter from one level to another one. When combined with the three distributions of the event severity discussed above, the trend models lead to the estimated coupon values given in Table 3.1.

In contrast to the continuous-trend models considered above, there can also be a sudden change in the expected event frequency. Such a change-point model for the event frequency can be combined with any of the models for the event severity.

The composite Poisson models discussed so far make use of the assumption that the event frequency is independent from the event severity, namely the adjusted numbers of claims arising from these events. The corresponding trend and change-point models take only a varying event frequency into account. Extensions of the peaks-over-threshold method from extreme value theory provide a convenient way to model a possible trend in the event frequency as well as in the event severity. However, when choosing only one additional parameter, those two trends are coupled. Further details on these extensions can be found in [4].

4. Further analysis of the models

When choosing a model, in particular for low frequency event risks, it is of interest to know how sensitive the model reacts to changes of the data. We refrain from manipulating the available historical data of Table 2.2 for this purpose. Instead, we employ a scenario technique in [4] by adding fictitious data for 1997 to the historical data set of Table 2.2. For a favourable scenario, we assume the best possible case, namely that no event with more than 1 000 claims is recorded in 1997. Such an event history happened three times already during the recorded 10-year history. For a stress scenario, we want to add a bad event record for 1997. To remain realistic, we prefer to pick a bad year from the available historical data. While the year 1994 is certainly the worst case with respect to the event frequency, it would not lead to a knock-out of the coupon and therefore counts as a favourable year for the binomial model. Hence we choose the data of the year 1992 as a common stress scenario for all models listed in Table 3.1.

For an easy comparison of the previous results with the coupon values arising from these scenarios, we assume that an identical three-year bond is issued in February 1998, that the observation period for the first coupon is shorter for applying the $^{15}/_{17}$-correction from (2.1), and that the interest-rate structure for the coupons is again given by Table 2.3. Based on the extended 11-year data set, we estimate the discounted value of the corresponding three WINCAT coupons using all models discussed so far. The model-dependent changes of the coupon values are given [4].

In the last section of [4], we check the consistency of the models and investigate the dispersion of the estimated discounted coupon values by a simulation study. For every fitted model—under the assumption that it describes reality correctly—we generate 1 000 new random data sets according to the distribution specified by the fitted model. These data sets replace the actual observations recorded in Table 2.2, and we use the model to estimate the discounted coupon values based on the random data set. In this way we can check whether the model can recover its own features from the simulated data—in particular the mean and the median—and we can see how far the simulated coupon values deviate from the mean. This can help to determine conservative estimates of the coupon values for the models. The mean, the median, the standard deviation and the 15.9%-quantil for all the models are listed in [4]. Instructive are also histograms showing the simulated distribution of the estimated coupon values for the various models.

Acknowledgements. I would like to thank Prof. Dr. F. Delbaen (Mathematical Finance, Department of Mathematics, ETH Zürich), Dr. A. Gisler (Winterthur Insurance) and in particular Dr. A. McNeil (Swiss Re Research Fellow at the ETH Zürich) for fruitful discussions.

REFERENCES

[1] Crédit Suisse First Boston, Fixed Income Research, *Convertible bond Winterthur Insurance with WinCAT coupons "Hail"*, Zürich, January 1997.

[2] H. Rootzén and N. Tajvidi, *Extreme value statistics and wind storm losses: A case study,* Scand. Actuarial J. (1997) 70–94.

[3] H.-H. Schiesser et al., *Klimatologie der Stürme und Sturmsysteme anhand von Radar- und Schadendaten,* Vdf, Hochschulverlag AG an der ETH, Zürich, 1997.

[4] U. Schmock, *Estimating the value of the WinCAT coupons of the Winterthur Insurance convertible bond: A study of the model risk,* Astin Bulletin (to appear).

[5] G. G. Venter, *Premium Calculation implications of reinsurance without arbitrage,* Astin Bulletin **21**, no. 2 (1991) 223–230.

[6] Winterthur Insurance, *Nachrangige Wandelanleihe 1997–2000 von CHF 399'500'000 mit $2\,^1/_4$ % WinCAT-Coupons "Hagel",* Emissions- und Kotierungsprospekt.

[7] Winterthur Insurance, $2\,^1/_4$ *% convertible bond with WinCAT coupon "Hail"*, Online information at URL `http://www.winterthur.com/prod/wincat/index-e.html`.

[8] Winterthur Insurance, *The Winterthur Share,* Online information at URL `http://www.winterthur.com/prod/aktien/index-e.html`.

[9] Winterthur Insurance, *Hagel und Schäden an Fahrzeugen, Ein Kurzbericht über die Bedrohung in der Schweiz,* internal report, January 1997.

(U. Schmock) MATHEMATICAL FINANCE, DEPARTMENT OF MATHEMATICS, ETH ZENTRUM, HG G 51.2, CH-8092 ZÜRICH, SWITZERLAND

E-mail address: `schmock@math.ethz.ch`