Journal of Statistical Mechanics: Theory and Experiment
An IOP and SISSA journal

Extreme value problems in random matrix theory and other disordered systems

Giulio Biroli[1,2], Jean-Philippe Bouchaud[2,3] and Marc Potters[2]

[1] Service de Physique Théorique, Orme des Merisiers—CEA Saclay, F-91191 Gif sur Yvette Cedex, France
[2] Science and Finance, Capital Fund Management, 6 Bd Haussmann, F-75009 Paris, France
[3] Service de Physique de l'État Condensé, Orme des Merisiers—CEA Saclay, F-91191 Gif sur Yvette Cedex, France
E-mail: biroli@spht.saclay.cea.fr, jean-philippe.bouchaud@cea.fr and marc.potters@cfm.fr

Received 12 February 2007
Accepted 6 March 2007
Published 24 July 2007

Online at stacks.iop.org/JSTAT/2007/P07019
doi:10.1088/1742-5468/2007/07/P07019

Abstract. We review some applications of central limit theorems and extreme values statistics in the context of disordered systems. We discuss several problems, in particular concerning random matrix theory and the generalization of the Tracy–Widom distribution when the disorder has 'fat tails'. We underline the relevance of power-law tails for directed polymers and mean-field spin glasses and we point out various open problems and conjectures on these matters. We find that, in many instances, the assumption of Gaussian disorder cannot be taken for granted.

Keywords: extreme value problems, random matrix theory and extensions, spin glasses (theory)

ArXiv ePrint: cond-mat/0702244

 1742-5468/07/P07019+15$30.00

Contents

1. Introduction

Most statistical models of disorder start by assuming that randomness has Gaussian statistics—from the classic Brownian motion to the Edwards–Anderson (or Derrida) models of spin-glasses, Kraichnan models of turbulent flows, KPZ models of surface growth, Black–Scholes models of financial markets, etc. Thanks to the outstanding mathematical properties of Gaussian random variables, this assumption is often technically very convenient and allows one to use powerful analytical techniques: stochastic calculus and Ito's lemma, field theory and replicas, etc. The real (and often implicit) justification is, however, the existence of a central limit theorem. This should ensure that on large enough lengthscales or timescales, the physical results are universal, independent of the details of the microscopic randomness—which can therefore, for simplicity and congeniality, be chosen as Gaussian. The paradigm of such a mechanism is the Brownian motion; in this case, provided elementary hops are sufficiently decorrelated from one another, it is well known that the sum of a very large number of these small displacements leads to a Gaussian diffusion profile, quite independently of the distribution of elementary hops—whenever its second moment is finite. If the second moment diverges, the walk becomes a Lévy flight, with anomalous diffusion described by the generalized central limit theorem of Lévy and Gnedenko which again ensures a certain degree of universality [1]: only the extreme tails of the microscopic distribution matter in the macroscopic limit. Although this dichotomy between finite and infinite variance is asymptotically rigorous, finite time or size effects can be strong and lead to effective violations of these central limit theorems. An important example is when the distribution of elementary hops has a finite variance but power-law tails. In this case, fat tail effects are persistent and convergence towards Gaussian diffusion is very slow. This is particularly relevant in finance, where significant deviations from Gaussian statistics are observed even for long time lags [2].

Sums of N i.i.d. random variables therefore provide a beautiful illustration of universality and universality classes, a concept that extends far beyond this simple, exactly soluble example. The existence of generalized central limit theorems for more complicated

J. Stat. Mech. (2007) P07019

(non-linear) problems involving random variables should be generic, again leading to some universality—universality classes should, however, be determined on a case-by-case basis and might be different from the Lévy–Gnedenko classification. A well-known example is the statistics of extreme values, say the largest of N independent random variables x_i. In this case again, the limiting distribution becomes to some degree universal; one has to distinguish three different cases, depending on the 'microscopic' distribution $p(x)$: Weibull (for distributions $p(x)$ which strictly vanish beyond a finite value x^*), Gumbel–Fisher–Tippett (for distributions decaying faster than any power law) and Fréchet (for power-law distributions) [3]. Interestingly, it is possible to formulate a problem which interpolates between *sums* of random variables and *extremes* of random variables, by considering the following quantity:

$$S_q = \left[\sum_{i=1}^{N} x_i^q\right], \tag{1}$$

where one assumes for simplicity that x_i's are all positive. Clearly, $q = 1$ corresponds to a simple sum, whereas when $q \to \infty$ at fixed N, $S_q^{1/q}$ converges to the largest element $x_{\max}$. Defining $x_i \equiv \exp -\varepsilon_i$, it is clear that S_q plays the role of a partition function and q is the inverse temperature for a generalized random energy model (REM), where the energies ε_i's are not necessarily Gaussian. This problem was considered in [4, 5] and has, beyond the REM interpretation, many different applications. For example, suppose ε_i is a growth rate of species i in the population, or the return of asset i in a portfolio, and q is the time. Then S_q is the total population after time q or the total value of the portfolio after time q; the detailed statistics of these objects is therefore quite interesting[4]. The result depends on the relative value of q and N when both diverge to infinity. More precisely, taking for simplicity ε_i to be Gaussian with variance σ^2, the relevant parameter is $\mu = \sqrt{2\ln N}/q\sigma$. The statistics of $S_q(N)$ only depends on μ and, quite interestingly, closely follows the above Gauss/Lévy dichotomy: for $\mu > 2$, S_q is Gaussian; for $\mu < 2$ it becomes Lévy distributed and more and more dominated by extreme values. In fact, as soon as $\mu < 1$, the whole sum S_q is well approximated by a finite number of terms, whereas when $\mu \to 0$, only the largest survives. The transition at $\mu = 1$ corresponds exactly the glass transition in the REM. The above results can be extended to any distributions of ε_i in the Gumbel class, up to a redefinition of μ [4, 5]; interestingly, the detailed statistics of S_q is precisely encoded in the one-step replica symmetry broken solution of the generalized random energy model [4, 7].

As the above example illustrates, it is clear that low temperature/long time properties of disordered systems are sensitive to extreme values rather than to typical values; this change of focus means that one should *a priori* be specially weary about universality classes and the influence of the choice of distribution on the physical results. The aim of this paper is to discuss several problems within this perspective, reviewing some recent and older results, and pointing out several open problems and technical challenges well worth investigating in the future.

[4] More complex situations, for example diffusion of species in a random environment, can be analysed along similar lines [6].

J. Stat. Mech. (2007) P07019

2. Random matrices and top eigenvalues

A remarkable example of universal limit distribution is the eigenvalue spectrum of random symmetric $N \times N$ matrices $\mathbf{M}$ with i.i.d. real elements. Again, as soon as the variance of the matrix elements is finite, the eigenvalue density $\rho(\lambda)$ converges to the Wigner semi-circle, with edges at $\lambda = \pm 2$ when the variance of the entries is normalized to $1/N$. This result can be derived in a way which makes direct use of the central limit theorem for sums of random variables, and in this way makes explicit the mechanism underpinning the universality of the Wigner semi-circle. In line with the above classification, one finds that, where the variance of entries diverges, the eigenvalue spectrum $\rho(\lambda)$ of $\mathbf{M}$ is no longer the Wigner semi-circle. Not surprisingly, the discussion parallels the Lévy–Gnedenko classification for sums of random variables. The case where the distribution of entries decays as a power-law $\sim |M_{ij}|^{-1-\mu}$ (possibly multiplying a slow function) with $\mu < 2$ (such that the variance of entries diverge), the eigenvalue spectrum $\rho(\lambda)$ can be computed exactly [8, 9] and no longer has a compact support, but itself acquires a power-law tail $\rho(\lambda) \sim |\lambda|^{-1-\mu}$, bequeathed from the tails of the matrix entries [8]. The structure of eigenvectors is also quite interesting: whereas for $\mu > 2$ most states are extended, various localization transitions occur, as a function of λ, for $\mu < 2$.

Following the above discussion, it is quite natural to investigate the distribution of extreme eigenvalues as well, and to find the universality classes corresponding to this question. Since the eigenvalues of a random matrix are strongly correlated random variables, one does not expect the result to belong to any of the well known Gumbel–Fisher–Tippett, Weibull and Fréchet cases. One of the most exciting recent results in mathematical physics is the Tracy–Widom distribution of the top eigenvalue of large Gaussian random matrices [10]. The truly amazing circumstance is that the very same distribution also appears in a host of physically important problems [11]: crystal shapes, exclusion processes [12], sequence matching, and, as discussed in the next section, directed polymers in random media.

Again, the Tracy–Widom result is expected to hold for a broad class of random matrices. The precise characterization of this class, as well as the extension of the Tracy–Widom result for other classes, is a subject of intense activity [14, 15]. The case where the distribution of entries decays as a power law $\sim |M_{ij}|^{-1-\mu}$ is expected to fall in a different universality class, at least when μ is small enough. The situation is simple when $\mu < 2$: for large λ, eigenvalues become uncorrelated and, as mentioned above, distributed as $\rho(\lambda) \sim |\lambda|^{-1-\mu}$. Correspondingly, the largest eigenvalues are described by Fréchet statistics [16]. What happens when μ is in the range $]2, +\infty)$, such that the eigenvalue spectrum $\rho(\lambda)$ still converges [8], for large N, to the Wigner semi-circle? We find that Fréchet statistics holds whenever $\mu < 4$, whereas the Tracy–Widom applies asymptotically as soon as $\mu > 4$, with a new limiting family of distributions for $\mu = 4$ [17]. The idea of the method is to start with a real symmetric matrix $\widehat{\mathbf{M}}$ with i.i.d. elements of variance equal to $1/N$, and such that the distribution has a tail decaying as:

$$p(M_{ij}) \simeq \frac{\mu (A N^{-1/2})^{\mu}}{|M_{ij}|^{1+\mu}}, \tag{2}$$

where the tail amplitude ensures that M_{ij}'s are of order $AN^{-1/2}$. As soon as $\mu > 2$, the density of eigenvalues converges to the Wigner semi-circle on the interval $\lambda \in [-2, 2]$,

J. Stat. Mech. (2007) P07019

meaning that the probability to find an eigenvalue beyond 2 goes to zero when $N \to \infty$. However, this does not necessarily mean that the largest eigenvalue tends to 2—we will see below that this is only true when $\mu > 4$. Now, we perturb the matrix $\widehat{\mathbf{M}}$ by adding a certain amount S to a given pair of matrix elements, say (α, β): $\widehat{M}_{\alpha\beta} \to \widehat{M}_{\alpha\beta} + S$ and $\widehat{M}_{\beta\alpha} \to \widehat{M}_{\beta\alpha} + S$. What can one say about the spectrum of this new matrix? Using self-consistent perturbation theory, which becomes exact for large N, one can show [17] that, when $|S| \geq 1$, there is a pair of eigenvectors partly localized on α, β, with eigenvalues $\lambda_{\pm} = \pm(S + 1/S)$ with $|\lambda| \geq 2$, which is therefore expelled from the Wigner sea. When $|S| < 1$, on the other hand, no such eigenvalue exists and the edge of the spectrum remains $\lambda_{\max} = 2$ in this case.

Now consider large matrix entries $|M_{ij}| > 1$. From equation (2), their number is $N^2 \int_1^\infty p(M_{ij})\,\mathrm{d}M_{ij} = A^\mu N^{2-\mu/2}$. In the case $\mu > 4$, it is clear that this number tends to zero when $N \to \infty$. With probability close to unity for large N, no entry is larger than one, in which case the largest eigenvalue is expected to remain Tracy–Widom around $\lambda^* = 2$. With small probability, however, the largest element S of $\mathbf{M}$ exceeds one; its distribution is $A^\mu N^{2-\mu/2}/|S|^{1+\mu}$ and the corresponding largest eigenvalue, using the above analysis, is $\lambda_{\max} = S + 1/S$. For $\mu > 4$ and large but finite N, we therefore expect that the distribution of the largest eigenvalue of $\mathbf{M}$ is Tracy–Widom, but with a power-law tail of index μ that very slowly disappears when $N \to \infty$. Our numerical results are in full agreement with this expectation (see figure 1). When $\mu < 4$, on the other hand, the number of large entries increases with N. However, when μ is larger than 2, such as to ensure that the eigenvalue spectrum still converges to the Wigner semi-circle, the number of rows or columns where two such large entries appear still tends to zero, as $N^{2-\mu}$. Therefore, the above analysis still holds: for each large element S_{ij} exceeding unity, one eigenvalue $\lambda = S_{ij} + S_{ij}^{-1}$ will pop out of the Wigner sea. Even if the eigenvalue *density* tends to zero outside of the interval $[-2, 2]$ when $2 < \mu < 4$, the *number* of eigenvalues exceeding 2 (in absolute value) grows as $N^{2-\mu/2} \ll N$. The largest eigenvalues are then equal to the largest entries and are themselves given by a Poisson point process with Fréchet intensity, as proven by Soshnikov in the case $\mu < 2$ [16]. His result therefore holds in the whole range $\mu < 4$. Finally, the marginal case $\mu = 4$ is easy to understand from the above discussion. The number of entries exceeding one remains of the order of unity as $N \to \infty$; the distribution of the largest entry S is Fréchet with N-independent parameters:

$$P_{\mu=4}(|S|) = \frac{4A^4}{|S|^5} \exp\left[-\frac{A^4}{|S|^4}\right]. \tag{3}$$

The probability that $|S|$ exceeds 1 is $\varphi = 1 - \mathrm{e}^{-A^4}$, in which case $\lambda_{\max} = |S| + |S|^{-1}$; otherwise, with probability $1 - \varphi$, $\lambda_{\max} = 2$. This characterizes entirely the asymptotic distribution of the largest eigenvalue in the marginal case $\mu = 4$: it is a mixture of a δ peak at 2 and a transformed Fréchet distribution. Note that this asymptotic distribution is non-universal since it depends explicitly on the tail amplitude A. Again, all these results are convincingly borne out by numerical simulations, see [17]. The statistics of the second, third, etc., eigenvalues could be understood along the same lines.

One can also consider the case of sample covariance matrices, important in many different contexts. The 'benchmark' spectrum of sample covariance matrix for i.i.d. Gaussian random variables is well known and given by the Marčenko–Pastur distribution [19]. Here again, the spectrum has a well defined upper edge and the

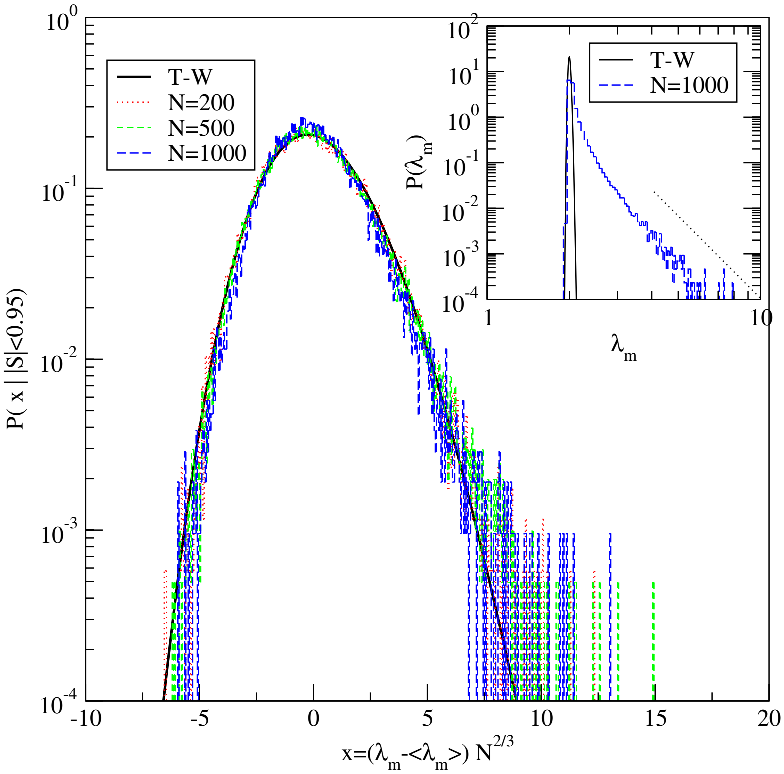

Figure 1. Histogram of $\lambda_{\max}$ conditioned on $|S| < 0.95$ for $\mu = 5$ for $N = 200, 500, 1000$: each eigenvalue has been shifted by the empirical mean and scaled by $N^{2/3}$. For comparison a GOE Tracy–Widom distribution of zero mean and variance adjusted to match $N = 500$ data is also shown (data obtained from [18]). Similar agreement with Tracy–Widom and scaling in $N^{2/3}$ is obtained for any value of μ when conditioned on $|S| < 0.95$. Note that, for the parameters chosen here, the probability of $|S| > 1$ is still quite large (75.2%) at $N = 1000$. Even for such large values of N, the unconditional distribution of $\lambda_{\max}$ has a marked power-law tail of index μ (dotted line) and is very different from the asymptotic Tracy–Widom distribution (inset).

distribution of the largest eigenvalue is Tracy–Widom (see, e.g., [15]). What happens if the random variables have heavy tails? More precisely, we consider N times series of length T each, denoted x_i^t, where $i = 1, \ldots, N$ and $t = 1, \ldots, T$. The x_i^t have zero mean and unit variance, but may have power-law tails with exponent μ. For example, daily stock returns are believed to have heavy tails with an exponent μ in the range 3–5 [2]. The empirical covariance matrix $\mathbf{C}$ is defined as:

$$C_{ij} = \frac{1}{T} \sum_t x_i^t x_j^t. \tag{4}$$

When the time series are independent, and for T and N both diverging with a fixed ratio $Q = T/N$, the eigenvalues of $\mathbf{C}$ are distributed in the interval $[(1 - Q^{-1/2})^2, (1 + Q^{-1/2})^2]$. When $T \to \infty$ at fixed N, all eigenvalues tend to unity, as they should since the empirical covariance matrix converges to its theoretical value, the identity matrix. When N and T

are large but finite, the largest eigenvalue of $\mathbf{C}$ is, for Gaussian returns, a distance $\sim N^{-2/3}$ away from the Marčenko–Pastur edge, with Tracy–Widom fluctuations. When returns are accidentally large, this may cause spurious apparent correlations and substantial overestimation of the largest eigenvalue of $\mathbf{C}$. Let us be more specific and assume, as above, that one particular return, say x_α^τ, is exceptionally large, equal to S. A generalization of the above self-consistent perturbation theory shows that, whenever $S \leq (NT)^{1/4}$, the largest eigenvalue remains stuck at $\lambda_{\max} = (1 + Q^{-1/2})^2$, whereas when $S > (NT)^{1/4}$, the largest eigenvalue becomes:

$$\lambda_{\max} = \left(\frac{1}{Q} + \frac{S^2}{T}\right)\left(1 + \frac{T}{S^2}\right). \tag{5}$$

This result again enables us to understand the statistics of $\lambda_{\max}$ as a function of the tail exponent μ. For N times series of i.i.d. random variables, of length T each, the largest element is of order $(NT)^{1/\mu}$. For $\mu > 4$, this is much smaller than $(NT)^{1/4}$ and, exactly as above, we expect the largest eigenvalue of $\mathbf{C}$ to be Tracy–Widom, with possibly large finite size corrections [20]. For $\mu < 4$, large 'spikes' in the time series dominate the top eigenvalues, which are of order $\lambda_{\max} \sim N^{4/\mu-1}Q^{2/\mu-1}$ and distributed according to a Fréchet distribution of index $\mu/2$. In the marginal case $\mu = 4$, as above, $\lambda_{\max}$ has a finite probability to be equal to the Marčenko–Pastur value, and with the complementary probability it is distributed according to a transformed Fréchet distribution of index 2, with a T and N independent scale. The structure of the corresponding eigenvectors can also be investigated and is again found to be partly localized when $S > (NT)^{1/4}$. Finally, we expect similar results to hold for the random singular value problem studied in [21], where rectangular matrices corresponding to cross-correlations between different sets of time series are considered.

3. Directed polymers, KPZ/KPP equations and pinned manifolds

Quite remarkably, the Tracy–Widom distribution for the largest eigenvalue of complex sample covariance matrices has deep links with the directed polymer (DP) problem in $(1+1)$ dimension, defined as follows: consider a two-dimensional square lattice, such that on each site one independently draws a random energy $e(x, y)$ from a given distribution $p(e)$. The zero-temperature directed polymer is the directed walk starting from $(0, 0)$ and only allowed to move North-East, such that the sum of encountered energies is minimum. When e is an exponential variable, there is an exact mapping to the Tracy–Widom problem [13]. Thanks to this mapping, the DP in $(1 + 1)$ dimension can be considered as a rare example of an exactly soluble disordered system in finite dimensions, for which not only the scaling exponents but the full distribution of the ground state energy can be completely characterized in terms of the Tracy–Widom distribution. In particular, the scaling between the sample-to-sample fluctuations of the ground state energy ΔE and the length of the path L is $\Delta E \sim L^{1/3}$ and the typical width of the optimal path is given by $W \sim L^{2/3}$. These conjectured scalings, based on physical arguments, are therefore exact for a certain class of random energy distributions $p(e)$. In the spirit of the above discussion, it is quite natural to wonder about the universality class of such results [22].

J. Stat. Mech. (2007) P07019

The question is all the more interesting that the DP problem maps onto a non-linear stochastic partial differential equation describing non-equilibrium surface growth, the so-called Kardar–Parisi–Zhang equation [23]:

$$\frac{\partial h(\vec{r},t)}{\partial t} = \nu \Delta h + \frac{\Lambda}{2}(\vec{\nabla} h)^2 + \eta(\vec{r},t), \tag{6}$$

where h is the height of the interface, $\vec{r}$ the coordinates along the d-dimensional interface and η a white noise term, mapping to the random energy e in the DP problem. (The $(1+1)$ DP problem corresponds to $d = 1$.) This equation can in turn be interpreted as a Burgers equation on the quantity $\vec{u} = -\vec{\nabla} h$, which corresponds to the effective force acting on the end point of the directed polymer. But since the Burgers equation is a toy model for turbulence, the question of the relevance of the distribution of the external forcing term $\vec{\nabla}\eta$ on the statistics of the velocity field is particularly interesting. KPZ-like equations may also describe very different physical situations, such as, for example, the propagation of crack fronts in disordered materials [24]; the robustness of the results with respect to the nature of the randomness is therefore quite important.

It is clear that it is the presence of a non-linear term in the KPZ/Burgers equation, responsible for the appearance of shocks in the velocity field, which makes the problem non-trivial. For a linear KPZ equation $\Lambda = 0$, called the Edwards–Wilkinson equation in this context, it is easy to show that the long time statistics of $h(\vec{r},t)$ is Gaussian provided $\eta(\vec{r},t)$ has a finite second moment, again thanks to the Central Limit Theorem for sums. The interplay between non-linearity and fat tails does, however, lead to rather unexpected results, as suggested by a simple Flory argument in the case where $p(e)$ decays, for large negative e, as $p(e) \sim |e|^{-1-\mu}$. The Flory argument compares the energy of the best 'bounty' site in a volume $V = W^d L$ to the elastic stretching energy the polymer has to pay to get there, of order W^2/L. Using extreme value theory in the Fréchet case, one gets $E_{\min} \sim -V^{1/\mu}$, leading to $W \sim L^{(1+\mu)/(2\mu-d)}$ and $\Delta E \sim E_{\min} \sim L^{(2+d)/(2\mu-d)}$. Of course, this reasoning is only valid if the extreme bounty site is worth the trip, i.e. if the distortion W is larger than in the absence of tails in the distribution of e. Calling ζ_d the exponent relating W to L in the 'thin tail' case, one expects the results to be strongly affected by the power-law tail of $p(e)$ as soon as μ is smaller than:

$$\mu < \mu_{\mathrm{c}} = \frac{1 + d\zeta_d}{2\zeta_d - 1}. \tag{7}$$

The exponent ζ_d is only known exactly in $d = 1$, with only numerical estimates available in $d > 1$ and still a very controversial situation concerning the value of the upper critical dimension d_{c} above which ζ_d takes the trivial random walk value $\zeta = 1/2$. In any case, for $d = 1$, $\zeta_1 = 2/3$, leading to $\mu_{\mathrm{c}} = 5$. In other words, the Flory argument suggests that, as soon as the fifth moment of the local energy distribution diverges, the Tracy–Widom scaling breaks down, contrary to the naive expectation that the DP exponents are universal as long as the variance of the local disorder is finite. The value $\mu_{\mathrm{c}} = 5$ is also different from the critical value $\mu_{\mathrm{c}} = 4$ found for the statistics of the top eigenvalue of random matrices, showing that a general mapping between the two problems, if it exists, is more subtle.

Summarizing, the Flory argument suggests that, in $(1+1)$ dimensions, the energy fluctuations should scale as $L^{1/3}$ and by Tracy–Widom for $\mu > 5$, and as $L^{3/(2\mu-1)}$ for

J. Stat. Mech. (2007) P07019

$2 < \mu < 5$ with a new type of limiting distribution (the case $\mu < 2$ corresponds to a complete stretching of the polymer and was recently solved in [25]). We have conducted new numerical simulations of this problem which indeed confirm that, for $\mu > 5$, the ground state energy scales as $L^{1/3}$ with Tracy–Widom fluctuations, while for $\mu < 5$ the above Flory prediction is very accurate. The distribution P of ground state energy can be well fitted by a geometric convolution of Fréchet distributions, suggested by the Flory argument with a finite number of dominant sites: $P = (1-p)(F + pF \star F + p^2 F \star F \star F + \cdots)$, different from the pure Fréchet distributed reported above for the largest eigenvalue for $\mu < 4$.

At this stage, a number of comments should be made. First, equation (7) shows that μ_c diverges when $d \to \infty$. This is perfectly in line with the Derrida–Spohn solution for the DP on a tree [26], which indeed breaks down completely as soon as $p(e)$ decays slower than exponentially when $e \to \infty$. This, in turn, is related to the problem of fronts in the Kolmogorov–Petrovsky–Piscounov (KPP) equation when the initial condition decays too slowly into the unstable phase. Whereas for localized initial conditions, the front between the stable and unstable phase propagates at a well defined velocity (which determines the free-energy of the DP [26, 27]), the case of slowly decaying initial conditions has, to our knowledge, not been carefully investigated. The very notion of a propagating 'front' might disappear altogether, much like in models of epidemic propagation with infected individuals performing Lévy flights, discussed in this context in [28].

Another intriguing property of equation (7) is that μ_c diverges whenever $\zeta_d = 1/2$, independently of dimension. If a critical dimension exists such that the strong disorder fixed point has $\zeta_d = 1/2$ for $d > d_c$, then this fixed point should be unstable against power-law tailed disorder with arbitrarily high exponent μ. Since all moments of the disorder still exist in this case, it is difficult to see how perturbative methods can grasp such a strange behaviour. More generally, it is not obvious to see how perturbative renormalization group methods for pinned systems (including the functional RG) can deal with high moment anomalies which change the scaling exponents. This seems to us to be a very interesting technical challenge; its resolution might indirectly shed light on the still elusive nature of the strong disorder fixed point of the DP/KPZ problem in high dimensions [29]. A case where progress is possible is the problem of pinned manifold in the mean-field limit, which maps onto a deterministic Burgers equation for the effective force acting on the manifold [4]. Disorder is now entirely contained in the initial condition, which represents the (bare) microscopic pinning force. The statistics of the renormalized force, in particular the density and amplitude of the 'jumps' (Burgers shocks) responsible for the famous cusp in the renormalized correlation function predicted by the functional RG, follows in this case the simple extreme value classification for independent random variables [4, 30]. In particular, the case of a Gaussian pinning field is indeed in a different universality class than any power-law tailed disorder. Interestingly, the functional RG has recently been cast in the form of a functional Burgers equation for the effective force in full generality [31]; this might provide a way to understand the importance of the statistics of the microscopic disorder in the general case as well.

Finally, the presence of so-called 'temperature chaos' [32, 33] might be quite sensitive to power-law tails in the microscopic disorder. Temperature chaos is associated with the fact that the free energy of the DP at non-zero temperatures scales as $L^{1/3}$, whereas both energy and entropy fluctuations are dominated by small scale fluctuations and therefore

J. Stat. Mech. (2007) P07019

scale as $L^{1/2} \gg L^{1/3}$. This implies that, when temperature changes by a small amount δT, polymers of length $L > L^*$, with $\delta T L^{*1/2} = L^{*1/3}$ must rearrange and find another equilibrium configuration. One finds $L^* \sim (\delta T)^{-6}$, in good agreement with numerical work [33] and exact results on Migdal–Kadanoff lattices [34]. Interestingly, as long as $\mu > 2$, entropy fluctuations should still scale as $L^{1/2}$, whereas free energy scales as $L^{3/(2\mu-1)}$ when $\mu < 5$. Hence, when $2 < \mu < 7/2$, small temperature changes, or small disorder changes at zero temperature, are not strong enough perturbations to induce large scale rearrangements.

4. Diffusion in a random potential

On the failure of perturbative RG to grasp the potential relevance of 'fat tails', it is interesting to mention the case of Langevin diffusion in a random potential, i.e. the long-time behaviour of $\langle x^2(t)\rangle$, where $\vec{x}$ obeys:

$$\frac{\mathrm{d}\vec{x}}{\mathrm{d}t} = -\vec{\nabla}U(\vec{x}) + \vec{\eta}(\vec{x}, t), \tag{8}$$

where η is again a thermal noise and $U(\vec{x})$ a random potential with short range correlations, over a length ξ. When U is Gaussian, one can show that diffusion is always normal, i.e. $\lim_{t\to\infty}\langle x^2(t)\rangle/t = D > 0$. Exact formulae for the value of D are available in one and two dimensions. Perhaps surprisingly, these formulae are exactly reproduced by a simple RG scheme, as independently proposed by Deem and Chandler [35], and Dean *et al* [36]. Physically, the time spent in a given region of space is $\tau(\vec{x}) \propto \exp -U(\vec{x})/T$; because the potential has Gaussian tails the average of $\tau(\vec{x})$ is finite, and hence the diffusion normal, with $D \sim \xi^2/\langle\tau\rangle$. But as soon as the disorder has tails decaying slower than exponential, the same argument leads to an infinite average trapping time and to subdiffusion. In the special case where $p(U) \sim \exp(-U/T_g)$, the model is similar to the trap model [37, 38]; one expects to find a critical temperature below which diffusion becomes anomalous, and $\langle x^2(t)\rangle \sim t^\mu$ with $\mu < 1$. This result was indeed recently confirmed exactly [39]; however, the RG scheme which leads to such an accurate result for Gaussian potentials is totally blind to the tails of the disorder and is unable to reproduce the above phase transition and subdiffusive behaviour. It would be very interesting to see how to adapt the RG scheme of [35, 36] to subexponential tails; this might be an important technical breakthrough to address quantitatively the issue of activated dynamics in supercooled glasses below the mode-coupling temperature.

5. Spin-glasses and non-standard RSB

In this final section we address the emblematic spin-glass problem, in particular the mean-field Sherrington–Kirkpatrick (SK) model, from the point of view of extreme value statistics. The SK model is defined by the following mean-field Hamiltonian:

$$\mathcal{H} = -\frac{1}{\sqrt{N}}\sum_{i,j} J_{ij} S_i S_j, \qquad S_i = \pm 1, \tag{9}$$

where J_{ij} are independent Gaussian random variables. It is now proven mathematically that the full replica symmetry broken solution invented by Parisi provides the exact solution fo the SK model and encodes in a rather magical way the complexity of the spin-glass phase [40]. Much insight into the meaning of Parisi's hierarchical construction is provided both by the cavity approach and by the random energy model. One understands that RSB is essentially a algebraic way to capture the Gumbel–Fisher–Tippett statistics of low lying energy states, which become relevant below the spin-glass transition [4]. The symptoms signalling that the replica symmetric (RS), one-phase solution is unstable are well known: its entropy becomes negative below a certain temperature (and remains negative at zero temperature) and the spin-glass susceptibility, which measures the correlation of spin fluctuations, assumed to be negligible in the RS description, in fact diverges below the spin-glass transition temperature. One of the striking predictions of the RSB solution is the pseudo-gap in the distribution of local magnetic fields h, which vanishes linearly when $h \to 0$: $P(h) \sim Ah$ at $T = 0$, in contrast with the Gaussian distribution of h found in the RS phase. Interestingly, this pseudo-gap is intimately related to the marginal stability of the ground state configuration [41] and is the close analogue of the Efros–Shklovskii gap in Coulomb glasses [42]. As a consequence of the vanishing of $P(h)$ for small h's, the specific heat of the spin-glass grows as T^2 at small T, rather than the naively expected linear-in-T behaviour.

How are these results affected by non Gaussian disorder? It is easy to be convinced, for example using the cavity method [43], that the SK results are universal provided the central limit theorem holds, i.e. the variance of the J_{ij}s is finite and correlations can be neglected. One knows that introducing strong correlations between the J_{ij}s, as is the case of the random orthogonal model, may change the nature of the solution, for example from full RSB to one-step RSB [44]. When the J_{ij}s are i.i.d. variables with a power law tail $|J_{ij}|^{-1-\mu}$ with $\mu < 2$ such that the variance is infinite, interesting effects appear, related to the fact that some bonds become extremely strong compared to others, thereby decreasing the frustration in the system. For example, there is a 'trivial' spin-glass phase, i.e. a phase where the Edwards–Anderson parameter is non-zero and the RS solution is stable [45]. RS is broken below a second transition temperature T_{AT}; however, the precise nature of the RSB phase is still unknown. The puzzle is that one naively expects the distribution of low-energy states to be governed by Fréchet statistics, since the energy is a sum of power-law distributed random variables. One knows from the study of the REM with power-law distributed energies that the solution in that case cannot be the standard Parisi replica symmetry breaking scheme [4], which—as mentioned above—describes the Gumbel universality class[5]. Things might, however, be more subtle: the energy of low lying states might be of the form $E_0 + e_\alpha$, where E_0 is indeed Fréchet but common to all states, and a Gumbel residue e_α, restoring the Parisi scheme. This should be associated with a distribution of overlaps q which does not extend down to $q = 0$.

Even in the absence of a detailed solution below T_{AT}, one can argue [45] that the distribution of local fields again develops a gap when $1 < \mu < 2$, of the form $P(h) \sim A_\mu h^{\mu-1}$. Correspondingly, the specific heat now grows as T^μ. The situation becomes quite interesting below $\mu = 1$ since in this case replica symmetry appears to be restored at zero temperature, and the gap in $P(h)$ disappears. The hierarchy between bonds is so strong for $\mu < 1$ that only very weak bonds appear to be frustrated in that

[5] The same remark applies to problems where the ground state energy is strictly bounded, such as the number partitioning problem, falling into the Weibull universality class [46].

case, unable to generate a large number of different states. It would be very interesting to characterize in detail the solution of this 'Lévy' SK model, and decide whether or not one needs to invent a non-standard RSB scheme for this problem.

We again end this section by multifarious remarks. First, the Tracy–Widom problem of the largest eigenvalue of random matrices can be seen as the statistics of the ground state energy of the $p = 2$ spherical spin-glass. It is very natural to ask about the statistics of the ground state energy of the (hard-spin) SK model. Interestingly, this problem is still open. Numerically, it is known that the *average* ground state energy per spin converges towards the asymptotic, $N = \infty$ result predicted by the Parisi solution with $N^{-2/3}$ corrections, as for the Tracy–Widom problem. Whether or not this is a coincidence is not understood; one should note that the sample-to-sample fluctuations of this ground state energy are *not* of the order of $N^{-2/3}$ (and not described by a Tracy–Widom distribution [47]) but of order $N^{-\omega}$ with $\omega \approx 3/4 > 2/3$. Again, there are no theoretical consensus on the value of ω [48, 49]; only qualitative arguments for $\omega \equiv 3/4$ exist. The generalization of these arguments to Lévy spin-glasses suggest, for $1 < \mu < 2$, $\omega = 1 - 1/\mu^2$; it would be interesting to have numerical data to test this prediction. Another natural extension of the Tracy–Widom problem within the context of spin-glasses is the statistics of the ground state energy of the spherical p-spin-glass, which leads to a non-linear eigenvalue problem. In the $p = 3$ case, one should find the largest eigenvalue of

$$\sum_{jk=1}^{N} J_{ijk}\phi_j\phi_k = E\phi_i, \tag{10}$$

where J_{ijk} are Gaussian random variables of variance $1/N^2$. It would be interesting to compute the scaling exponent describing the fluctuation of this new type of extreme value problem. Could this problem also be mapped to the directed polymer problem in higher dimension?

Finally, we note that, in finite dimensions, the role of the distribution of J_{ij} on the universality class of the spin-glass transition was numerically investigated by Campbell *et al* [50]; the consensus is, however, that (at least for fast-decaying) distribution, the critical exponents should be universal, with possibly large sub-leading corrections. However, the situation might change for power-law distributed couplings, as suggested by a Migdal–Kadanoff approach to the excitation energy exponent [48].

6. Conclusion and open problems

In summary, we reviewed several examples where the detailed shape of the distribution of randomness matters more than naively anticipated on the basis of the central limit theorem. The tail of the distribution is obviously important when one is concerned by extreme value problems, such as the statistics of the largest eigenvalue of heavy tailed random matrices, a problem we have discussed in detail. Another problem where the tails of the noise are crucial is the problem of barrier crossing and the Arrhenius law. This is not surprising: since barrier crossing is itself a rare event, its occurrence may be much enhanced by the present of anomalous, non-Gaussian tails in the thermal noise. Note that the relevance of extreme value statistics to barrier height in disordered systems dates back to Rammal [51], see also [52].

J. Stat. Mech. (2007) P07019

More surprising, at least at first sight, is the importance of these tails for the statistics of fields obeying a non-linear differential equation, such as the Burgers, KPZ or KPP equations, either with a stochastic forcing or with a stochastic initial condition. This includes the example of the directed polymer in random environments or diffusion in a random potential; the scaling exponents of these problems is extremely sensitive to the tails of the disorder, an unusual feature from a field theoretical point of view. We have insisted on the technical challenges associated with this phenomenon: how should one generalize the (functional) RG to account for non-Gaussian tails? Since the problem of pinned manifold can also be addressed using a replica field theory, one can similarly wonder if and how the Parisi replica symmetry breaking scheme has to be generalized in these cases. We have underlined several other open problems and conjectures, such as the general relation between top eigenvalues and directed polymers, the statistics of the ground state energy of the SK model, the nature of the low temperature phase of Lévy spin-glasses, etc. The solution to these problems, beyond their mere technical interest, could shed some light on the subtleties and surprises of the physics of disordered systems (see [53] for more examples).

Acknowledgments

We thank G Ben Arous, S Boettcher, Z Burda, D Dean, A Lefevre, S Majumdar, O Martin, M Mézard and H Yoshino for many useful discussions on these matters.

References

[1] See e.g. Bouchaud J-P and Georges A, *Anomalous diffusion in random media: statistical mechanisms, models and physical applications*, 1990 *Phys. Rep.* **195** 127
Shlesinger M F, Zaslavsky G M and Frisch U (ed), *Lévy Flights and Related Topics in Physics* (*Lecture Notes in Physics* vol 450) (Berlin: Springer)

[2] See e.g. Bouchaud J-P and Potters M, 2004 *Theory of Financial Risk and Derivative Pricing* (Cambridge: Cambridge University Press)

[3] Galambos J, 1987 *The Asymptotic Theory of Extreme Order Statistics* (Malabar, FL: Krieger)

[4] Bouchaud J-P and Mézard M, *Universality classes for extreme-value statistics*, 1997 *J. Phys. A: Math. Gen.* **30** 7997

[5] Ben Arous G, Bogachev L V and Molchanov S A, *Limit theorems for sums of random exponentials*, 2005 *Probab. Theor. Relat. Fields* **132** 579

[6] Ben Arous G, Molchanov S A and Ramirez A F, *Transition from the annealed to the quenched asymptotics for a random walk on random obstacles*, 2005 *Ann. Probab.* **33** 2149
Ben Arous G, Molchanov S A and Ramirez A F, *Phase transition asymptotics for random walks on a stationary random potential*, 2005 *Preprint* math.PR/0510519

[7] Derrida B, *Non-self averaging effects in sum of random variables*, 1994 *On Three Levels* ed M Fannes, C Maes and A Verbeure (New York: Plenum) p 125

[8] Cizeau P and Bouchaud J P, *Theory of Lévy matrices*, 1994 *Phys. Rev.* E **50** 1810

[9] Burda Z, Jurkiewicz J, Nowak M A, Papp G and Zahed I, *Random Lévy matrices revisited*, 2006 *Preprint* cond-mat/0602087

[10] Tracy C A and Widom H, *Level spacing distributions and the airy kernel*, 1994 *Commun. Math. Phys.* **159** 33

[11] For a review, see Spohn H, *Exact solutions for KPZ-type growth processes, random matrices, and equilibrium shapes of crystals*, 2005 *Preprint* cond-mat/0512011

[12] Majumdar S and Nechaev S, *Exact asymptotic results for the Bernoulli matching model of sequence alignment*, 2005 *Phys. Rev.* E **72** 020901(R)

[13] Johansson K, *Shape fluctuations and random matrices*, 2000 *Commun. Math. Phys.* **209** 437

[14] Soshnikov A, *Universality at the edge of the spectrum in Wigner random matrices*, 1999 *Commun. Math. Phys.* **207** 697

[15] Baik J, Ben Arous G and Péché S, *Phase transition of the largest eigenvalue for non-null complex sample covariance matrices*, 2005 *Ann. Probab.* **33** 1643

[16] Soshnikov A, *Poisson statistics for the largest eigenvalues of Wigner random matrices with heavy tails*, 2004 *Elect. Comm. Probab.* **9** 82

[17] Biroli G, Bouchaud J-P and Potters M, *On the top eigenvalue of heavy-tailed random matrices*, 2007 *Europhys. Lett.* **78** 10001

[18] Prähofer M and Spohn H, *Exact scaling functions for one-dimensional stationary KPZ growth*, 2004 *J. Stat. Phys.* **115** 255

[19] Marčenko V A and Pastur L A, *Distribution of eigenvalues for some sets of random matrices*, 1967 *Math. USSR-Sb* **1** 457

[20] For a case where the *asymptotic* eigenvalue spectrum develops fat tails, see: Burda Z, Gorlich T and Waclaw B, *Spectral properties of empirical covariance matrices for data with power-law tails*, 2006 *Preprint* physics/0603186

[21] Bouchaud J-P, Laloux L, Miceli M A and Potters M, *Large dimension forecasting models and random singular value spectra*, 2007 *Eur. Phys. J.* B **55** 201

[22] Zhang Y C, *Growth anomaly and its implications*, 1990 *Physica* A **170** 1

[23] Halpin Healey T and Zhang Y C, *Kinetic roughening, stochastic growth, directed polymers and all that*, 1995 *Phys. Rep.* **254** 189

[24] Bouchaud J P, Bouchaud E, Lapasset G and Planès J, *Models of fractal cracks*, 1993 *Phys. Rev. Lett.* **71** 2240

[25] Hambly B and Martin J B, *Heavy tails in last-passage percolation*, 2006 *Preprint* math.PR/0604189

[26] Derrida B and Spohn H, *Polymers on disordered trees, spin glasses, and traveling waves*, 1988 *J. Stat. Phys.* **51** 817

[27] For an interesting related discussion, see: Carpentier D and Le Doussal P, *Glass transition of a particle in a random potential, front selection in non linear renormalisation group*, 2001 *Phys. Rev.* E **63** 026110
Muzy J-F, Bacry E and Kozhemyak A, *Extreme values and fat tails of multifractal fluctuations*, 2006 *Phys. Rev.* E **73** 066114

[28] Brockmann D and Hufnagel L, *Front propagation in reaction–superdiffusion dynamics*, 2004 *Preprint* cond-mat/0401322

[29] For a recent discussion, see Canet L and Moore M A, *Mode-coupling theory of the Kardar–Parisi–Zhang equation and the functional renormalization group*, 2006 *Preprint* cond-mat/0604301

[30] Kida S, *Asymptotic properties of Burgers turbulence*, 1979 *J. Fluid Mech.* **93** 337

[31] Le Doussal P, *Finite temperature FRG, droplets and decaying Burgers turbulence*, 2006 *Preprint* cond-mat/0605490
See also Middleton A, Le Doussal P and Wiese K J, *Measuring functional renormalization group fixed-point functions for pinned manifolds*, 2006 *Preprint* cond-mat/0606160

[32] Fisher D S and Huse D A, *Directed paths in a random potential*, 1991 *Phys. Rev.* B **43** 10728

[33] Sales M and Yoshino H, *Fragility of the free-energy landscape of a directed polymer in random media*, 2002 *Phys. Rev.* E **65** 066131

[34] da Silveira R A and Bouchaud J-P, *Temperature and disorder chaos in low dimensional directed paths*, 2004 *Phys. Rev. Lett.* **93** 015901

[35] Deem M and Chandler D, *Classical diffusion in strong random media*, 1994 *J. Stat. Phys.* **76** 911

[36] Dean D S, Drummond I and Horgan R, *Perturbation schemes for flow in random media*, 1994 *J. Phys. A: Math. Gen.* **27** 5135

[37] See e.g. Monthus C and Bouchaud J-P, *Trap models and phenomenology of glasses*, 1996 *J. Phys. A: Math. Gen.* **29** 3847

[38] Ben Arous G and Cerny J, *Dynamics of trap models*, 2006 *Preprint* math.PR/0603344

[39] Touya C and Dean D S, *Dynamical transition for a particle in a squared Gaussian potential*, 2006 *Preprint* cond-mat/0610470

[40] Mézard M, Parisi G and Virasoro M A, 1987 *Spin Glass Theory and Beyond* (Singapore: World Scientific)

[41] Anderson P W, *Ill-condensed matter*, 1978 *Les Houches Lecture Notes* (Singapore: World Scientific)

[42] See the very interesting recent paper Mueller M and Pankov S, *Mean field theory for the 3D Coulomb glass*, 2006 *Preprint* cond-mat/0611021

[43] Parisi G, *Chance & Matter*, 1986 *Les Houches Lecture Notes* (Amsterdam: North-Holland)

[44] Parisi G and Potters M, *Mean-field equations for spin models with orthogonal interaction matrices*, 1995 *J. Phys. A: Math. Gen.* **28** 5267

[45] Cizeau P and Bouchaud J-P, *Mean field theory of dilute spin-glasses with power-law interactions*, 1993 *J. Phys. A: Math. Gen.* **26** L187

See also Engel A, *Replica mean-field theory for Levy spin-glasses*, 2007 *Preprint* cond-mat/0701197

[46] See e.g. Bauke H, Franz S and Mertens S, *Number partitioning as a random energy model*, 2004 *J. Stat. Mech.* P04003

[47] Andreanov A, Barbieri F and Martin O C, *Large deviations in spin-glass ground-state energies*, 2004 *Eur. Phys. J.* B **41** 365

[48] Bouchaud J-P, Krzakala F and Martin O C, *Energy exponents and corrections to scaling in Ising spin glasses*, 2003 *Phys. Rev.* B **68** 224404

[49] Goethe M and Aspelmeier T, *Free energy fluctuations in the mean-field Ising spin glass*, 2006 *Preprint* cond-mat/0610228 and references therein

[50] Bernardi L W and Campbell I A, *Critical exponents in Ising spin glasses*, 1997 *Phys. Rev.* B **56** 5271

[51] Rammal R, *Spin dynamics and glassy relaxation on fractals and percolation structures*, 1985 *J. Physique* **46** 1835

[52] Vinokur V, Marchetti C and Chen L W, *Glassy motion of elastic manifolds*, 1996 *Phys. Rev. Lett.* **77** 1845

[53] Juhasz R, Lin Y-C and Igloi F, *Strong Griffiths singularities in random systems and their relation to extreme value statistics*, 2006 *Phys. Rev.* B **73** 224206

J. Stat. Mech. (2007) P07019