# Universality and Extreme Value Statistics

Term Paper for Physics 464

Spring 2000

Taras Pogorelov

**Abstract**

Universality classes of sums of random variables and those of extreme value statistics (EVS) are presented. Relevance of both to certain physical models based on the recent literature is described. A review of applications of extreme value statistics to random energy models is given.

# 1 Introduction

In the recent years a growing number of evidence for universality in properties of diverse physical systems started to appear. R.Rammal (1985) [3] studied relevance of the extreme value statistics to the description of energy barrier of cluster formation in the percolating systems. J.P.Bouchaund and M.Mezard (1997) [1] used EVS and replica techniques to study random energy models. Finally, S.T.Bramwell (1998,2000) [5] implied relevance of EVS to universality of turbulent and magnetic systems.

In this paper we start with a review of classes and applications of sums of standard random variable. These systems were well studied and many analytic solutions were found, for an excellent review one can use Gnedenko and Kolmogorov book(1968). For physical application the review by Bouchuad and Georges [4] is highly recommended. We follow the latter report and start with the systems where the central limit theorem can be applied,which includes such important systems as free random walks. The cases where the CLT is violated are also quite fruitfull, in particular in studies of self-avoiding random walks (SAW) and Levy flights (SALF). The reproduced 'phase' diagram for the SALF display the range of the applications.

Then the discussion continues with the extreme variables. The modern analytical study of which originated in 1920s. L. von Bortkiewicz attempted to study the distribution of range and the mean range in samples from a Gaussian distribution as a function of sample size. The exact solution was proven to be difficult to find, but some numerical results were received.The beginning with the Gaussian distribution leaded to delays with futher development. Later it was found that for the Exponential distribution the results are much easier to receive.

E.J.Gumbel started to use ESV for application in the middle 1930s. The first application was to old age. He showed, for example that the characteristic limiting old age for a group with a lower average but wider distribution is actually higher then for the one with higher average and narrowier distribution.

In also become evident that to study EVS one need additional characteristics e.g.intensity, return period, besides the usual ones e.g. momemts.

Because of the difficulties in finding an exact solution the problems were approached with asymptotes. It was shown that to satisfy the Stability principle all asymptotic EVS need to fall into three (Exponential) classes: Gumbel, Frechet/Cauchy, Weibull.

We conclude the paper with reproduction of the studies by Bouchaud and Mezard [1] who applied EVS to random energy models and have shown the relation to the replica method. They showed that the models are related to the Gumbel class and have discussed future perspectives fo rthe other universality classes of EVS.

In fact, up to this time all models of physical systems with relevance to EVS belong to the Gumbel class. This leave a large field of potential application to be expored, especially remembering that next to EVS is its generalization the order statistics, which deals with the second, third, etc. extremes.

# 2 Universality classes and Statistics

This section contains an overview of the standard universality classes of sums of random variable and the universality classes for the extremes of random variables.

## 2.1 Universality classes of sums of random variables

### 2.1.1 Central limit theorem (CLT)

Let us state the central limit theorem by considering the example of the one-dimensional random walk [4]. Each time n a jump of length $l_n$, distributed as $P(l)$ is performed. Thus after $N$ independent jumps we have position

$$X_t = \sum_{n=1}^{N} l_n, \tag{1}$$

Assuming that $< l >$ and $< l^2 >$ of $P(l)$ are finite we have linear dependence in time for the mean and the variance.

$$\overline{X}_t = Vt, \overline{X_t^2} - \overline{X}_t^2 = 2Dt, \tag{2}$$

where the velocity and the diffusion constant are

$$V = \langle l \rangle / \tau, D = (2\tau)^{-1}[\langle l^2 \rangle - \langle l \rangle^2], \tag{3}$$

Now the CLT for such random walk can be stated as:

$$\text{Prob}\{u_1 \leq (x_t - Vt)/2\sqrt{Dt} \leq u_2\} \overset{t\to\infty}{\longrightarrow} \frac{1}{\pi} \int_{u_1}^{u_2} \mathrm{e}^{-\xi^2} \mathrm{dx} \tag{4}$$

Remarks: (i)At large times the limit distribution (4) is a function of first two moments $\langle l \rangle$ and $\langle l^2 \rangle$; (ii) the CLT guarantees the limit distributions only within the scaling region, where $(X_t - Vt)/\sqrt{Dt}$ is finite. Outside this region the limit distribution $P(x,t)$will have 'non-universal' tails, not described by (4), which depend on all the details of the original distribution $P(l)$. The standard form of the CLT, described above applies assuming the random variables satisfy the following two conditions:(i)the distributions of the summed random variables are not 'too broad' (e.g. the second moment is finite); (ii)the random variables are not 'long-range correlated'. Let us now consider those cases.

## 2.2 Broad distributions and CLT

### 2.2.1 Sums of independent random variables

Consider [4]

$$X_t = \sum_{n=1}^{N} l_n, \tag{5}$$

when the distribution $P(l)$ is 'broad', i.e. decreases more slowly than $l^{-3}$ for large $l$.

The problem of characterization of the limit distribution was solved by Levy and Khintchine. Here we will give only a number of quantitative statements.

Let us consider $P(l)$ which for large $l$ decreases as $l^{-(1+\mu)}$ $(\mu < 0)$. Then:

-For $0 < \mu < 1$, $X_N$ behaves as $N^{1/\mu}$ (for $\mu = 1$ as $N \log N$). $\langle l \rangle$ and $\overline{X}_N$ are infinite in this case.

-For $1 < \mu < 2$, $\langle l \rangle$ is finite and $X_N = \langle l \rangle N$. Also $X_N - \overline{X}_N$ behaves as $N^{1/\mu}$ (or as $\sqrt{N \log N}$ for $\mu = 2$. The variance $\overline{X_N}^2 - \overline{X}_N^2$ is infinite.

-For $\mu > 2$, $\langle l^2 \rangle$ is finite and we have the situation of the previous section.

Or more precisely the variable $Z_N = X_N / N^{1/\mu}$ for $0 < \mu < 1$ [or $(X_N - \langle l \rangle N)/N^{1/\mu}$ for $1 < \mu < 2$ ] has the following limit distribution:

$$\mathrm{Prob}\{u_1 \le Z_N \le u_2\} \overset{N \to \infty}{\longrightarrow} \int_{u_1}^{u_2} \mathrm{L}_{\mu,\beta}(\mathrm{u}) \mathrm{du} \tag{6}$$

where $L_{\mu,\beta}$ are called Levy (or 'stable') laws of index $\mu$. They are fully characterized by the two parameters $\mu$ and $\beta$ $(0 < \mu < 2, -1 \le \beta \le 1)$. The parameter $\beta$ characterizes the degree of symmetry.

To better understand the behavior of the sum $X_N$ let us find the largest value $l_c(N)$ among the $N$ terms of the sum (5).

Clearly, $l_c(N)$ can be estimated as

$$N \int_{l_c(N)}^{\infty} P(l) dl \simeq 1 \tag{7}$$

Thus the value larger than $l_c(N)$ occurred at most once in $N$ trials. Hence

$$l_c(N) \sim N^{1/\mu}, N \longrightarrow \infty \tag{8}$$

We can cut off the distribution $P(l)$ at $l \sim l_c(N)$ because for a large finite number $N$ of trials the value $X_N$ is insensitive to events with $l \gg l_c(N)$.

Thus:

-For $0 < \mu \le 1$, the typical value of $X_N$ can be estimated by the mean value of the sum (5)with this cut off.

$$X_N \sim N \int_{l_c(N)} P(l) dl \sim \begin{cases} N (N^{1/\mu})^{1-\mu} = N^{1/\mu}, & \mu < 1 \\ N \log N, & \mu = 1 \end{cases} \tag{9}$$

-For $1 < \mu < 2$, the typical value of $X_N - \overline{X}_N$ can be estimated by the mean value of the sum (5).

$$(X_N - \overline{X}_N)^2 \sim N \int_{l_c(N)} (l - \langle l \rangle)^2 P(l) dl \sim \begin{cases} N(N^{1/\mu})^{2-\mu} = N^{2/\mu}, & \mu < 2 \\ N \log N, & \mu = 2 \end{cases} \tag{10}$$

-For $\mu > 2$ the integral in (9)converges and we have linear dependence on $N$. Clearly, when $\mu < 2$, the typical value of the sum $X_N$ is dominated by its largest term $l_c(N)$, i.e. the sum $X_N$ has manifest self-similar nature.

### 2.2.2 Levy flights and some applications

We can generalize Brownian motion by considering $l_n$ (1) to be the length of jumps at time steps $t = n\tau$ [4]. The mean squared displacement in general will have a power law dependence:

$$\langle X^2(t) \rangle \sim K_\alpha T^\alpha, \tag{11}$$

Dependent of $\alpha$ we have different types of diffusion:

For $0 < \alpha < 1$, subdiffusion

For $\alpha = 1$, normal diffusion

For $1 < \alpha$, super diffusion

For $\alpha = 2$ ballistic diffusion

Example: Polymer adsorption and self-avoiding Levy flight.

Consider and adsorbed polymer made of points in direct contact with an attractive wall, separated by large loops in the bulk. Main point is that the distribution of the size of these loops decay as power law. This induces a broad distribution between two consecutive adsorbed monomers, i.e. the projection of the chain;s conformation on the wall is a self-avoiding Levy flight ('node' avoiding).

The step length distribution reads

$$P(l) \sim l^{-(1+\mu)}, \tag{12}$$

The 'phase' diagram in the plane $(\mu, d)$, where $d$ is the dimension of the wall, for a self-avoiding Levy flight is shown in the figure 1, [4].

### 2.2.3 Long-range correlation

When the second condition for the CLT is not satisfied , i.e. the summed random variables are not independent, the analysis can not be developed as precisely as for the previous case.

The relevance of the correlations can be studied with the correlation function defined as following [4]

$$C(n) = \langle l_k l_{k+n} \rangle - \langle l_k \rangle \langle l_{k+n} \rangle \tag{13}$$

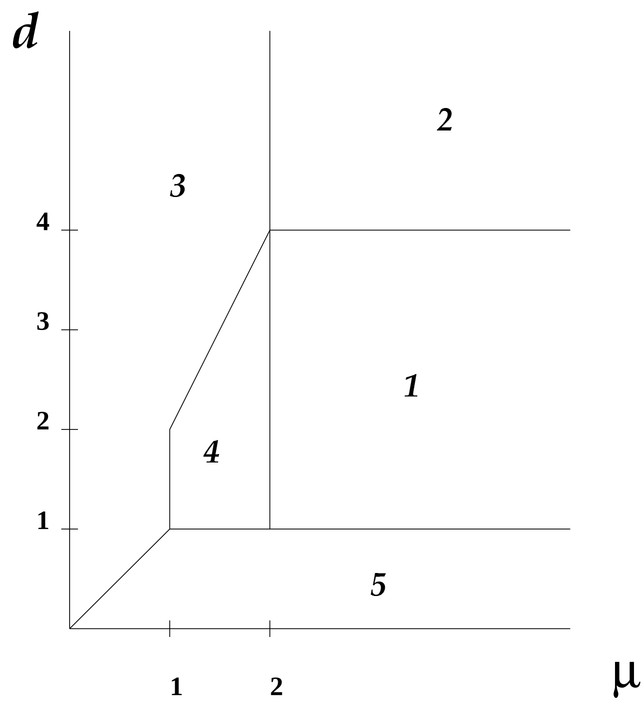


Figure 1: The 'phase diagram' of SALF [4].(1)usual SAW,(2)free random walk,(3)free Levy flight,(4)SALF, (5)collapsed.

The function clearly depend on the difference $n$ (stationary process).

Let $\langle l_k \rangle = 0$ then:

$$\overline{X_N^2} = NC(0) + 2\sum_{k=1}^{N}(N-k)C(k), \tag{14}$$

Thus we have to consider the following two cases:

(i) $\sum_{n=1}^{N} C(n)$ converges as $N \longrightarrow \infty$, that is $C(n)$ decays more rapidly than $n^{-1}$ for large $n$. Thus

$$\overline{X_N^2} \sim N(C(0) + 2\sum_{n=1}^{N} C(n)), N \longrightarrow \infty \tag{15}$$

i.e. $X_N$ behaves as $\sqrt{N}$

(ii)'long-ranged' correlations, when $C(n)$ decays as $1/n$ or more slowly (e.g. $C(n) \sim n^{-y}, y < 1$). Then

$$\overline{X_N^2} \sim N \int_N C(n)dn \sim \begin{cases} N^{2-y}, & y < 1 \\ N \log N, & y = 1 \end{cases} \tag{16}$$

Thus 'diffusion' is enhanced by correlations and the value of $X_N$ is larger than $\sqrt{N}$. Superdiffusion is observed and in the case of perfect correlations, e.g. for $y = 0$ we have 'ballistic' result $X_N \sim N$.

## 2.3 Universality classes for extreme-value statistics

The goal of a statistical theory of extreme values is to analyze observed extremes and to forecast further extremes [2]. The generalization of extreme value statistics (EVS) is order statistics, which is a study of second, third, etc. extremes.

Two types of questions are typically answered by EVS:1)does an individual observation in a sample fall outside what may reasonably be expected? 2)does a series of extreme values exhibit a regular behavior?

The following conditions typically should be fulfilled:1) the distribution from which the extremes have been drawn and its parameters must remain constant in time (or space), 2)the observations from which the extremes are taken should be independent.

Besides the usual tools of analysis i.e. mean, variance, moments, at least two more are needed for analysis of EVS: intensity $\mu(x)$: which is the probability that a value known to be equal to or larger than $x$ be situated between $x$ and $x + dx$. That is:

$$\mu(x)dx = \frac{f(x)dx}{1 - F(x)} \geq f(x)dx \tag{17}$$

Clearly, for a given intensity the probability $F(x)$ is:

$$1 - F(x) = 1 - F(x_0) \exp\left(-\int_{x_0}^{x_1} \mu(z)dz\right) \tag{18}$$

where $x_0$ is an arbitrary value of $x$.

The second tool, specific for EVS, is return period $T(x)$: if an event has probability $p$, we have on the average, $1/p$ trials in order that the event happen once. The mean is the return period $T(x)$.

### 2.3.1 Asymptotic EVS

The study of the exact distributions of the extremes as functions of the sample size is hard. And typically one uses the asymptotic distributions.

All asymptotic distributions of EVS fall into three classes. The limiting distributions $F_k(x)$ are the solutions of the functional equation:

$$F^n(a_n x + b_n) = F(x), n \geq 1 \tag{19}$$

This constitutes the Stability principle: the largest in a sample of $n$ from a distribution with the distribution function $F(x)$ must itself have the distribution $F(x)$.

The three asymptotic classes are:

(i)*Gumbel*

$$F_1(x) = \exp(-e^{-y}), f_1(x) = \alpha_n \exp(-y - e^{-y}), -\infty \leq y \leq \infty \tag{20}$$

where $y = \alpha_n(x - u_n)$, and $u_n$ is the characteristic $n^{th}$ largest value.

(ii)$\underline{Cauchy/Frechet}$

$$F_2(z) = \exp(-(\upsilon_n/z)^k), z > 0, k > 0 \tag{21}$$

with the distribution for the largest value given by

$$f_2(z) = \left(\frac{k}{\upsilon_n}\right)\left(\frac{\upsilon_n}{z}\right)^{k+1} \exp(-(\upsilon_n/z)^k), \tag{22}$$

(iii)$\underline{Weibull}$

$$F_3(x) = \exp\left(-\left(\frac{\omega - x}{\omega - \upsilon}\right)^k\right), x \leq 0, k > 0 \tag{23}$$

$$f_3(x) = \left(\frac{k}{\omega - \upsilon}\right)\left(-\left(\frac{\omega - x}{\omega - \upsilon}\right)^{k-1}\right) F_3(x), x \leq 0, k > 0 \tag{24}$$

### 2.3.2 Some applications

There are a number of publications which study the relation of EVS and physical models. We will give more descriptive account for random energy model in the next next section. Here we just list some additional studies.

Rammal [3] uses EVS, specifically the Gumbel's class, to study energy barriers of cluster formation in percolation structures.

Bramwell et. al [5] show that the universality of some properties of turbulent and magnetic systems might be related to the Gumbel class.

Finally, Bouchaud and Mezard [1] show relation of EVS and the replica method to REM and the Burger's turbulence. About which we now give a little more details.

# 3 EVS and Physical models

In this section we will discuss on the examples of one- and many-dimensional random energy models (REM) relation of physical models to the universality classes of extreme value statistic.

## 3.1 Scaling

We will need to start with rescaling results of EVS in order to prepare for discussion [1]. Consider $M$ independent identically distributed random variables (iid) $E_i$, $i = 1, ..M$ (energies) with the following probability distribution:

$$P(E) \sim \frac{A}{E^\alpha} \exp(-B|E|^\delta); B, \delta >; E \longrightarrow -\infty \tag{25}$$

We would like to study the statistics of the lowest energy state $E^* = min\{E_1, .., E_M\}$ for large $M$.

Define the repartition function of $E$, $P_<(E)$ as following:

$$P_<(E) = \int_{-\infty}^{E} \mathrm{P(E')dE'} \tag{26}$$

Then

$$P_M(E^*) = MP(E*)(1 - P_<(E^*))^{M-1} = -\frac{\mathrm{d}}{\mathrm{dE^*}}(P_>(E^*))^M \tag{27}$$

Now for large $M$, the minimum $E^*$ will be negative and large:

$$(1 - P_<(E^*))^M \simeq \exp(-MP_<(E^*)) \tag{28}$$

The repartition function of $E^*$ will be very small when $E^* < E_c(M)$, characteristic value of energy defined by $MP_<(E_c) = 1$. In the case of the distribution (25) we have (to logarithmic accuracy):

$$E_c(M) \simeq -\left(\frac{\log M}{B}\right)^{1/\delta} \tag{29}$$

To find the size of the fluctuations of the extreme $E^*$ around $E_c$ we expand (28) in $(E^* - E_c)/E_c$:

$$\Delta(M) = \frac{1}{B\delta(E_c)M} \tag{30}$$

The rescaled minimum energy variable $u \equiv (E^* - E_c(M))/\Delta(M)$ obeys for large $M$ a universal 'Grumbel' distribution [2]:

$$P^*(M) = \exp(u - \exp u) \tag{31}$$

Note that $E_c(M)$ is the most probable value for the extreme energy, since $P^*(u)$ has its maximum at $u = 0$.

This behavior is valid only in the region where the deviation $\epsilon$ of $E_c(M)$ is of the order $\Delta(M)$. It goes to zero with $M$ if $\delta > 1$ and diverges otherwise.

The relative fluctuations $\Delta(M)/E_c(M) \sim 1/\log(M)$. Also important to note that $P^*(u)$ vanishes exponentially for $u \longrightarrow -\infty$. In the scaling region, the probability for a given energy $E_i$ to be $E_i = E_c(M) + \Delta(M)u_i$ behaves as $(c/M)\exp(u_i)$. Thus the low-lying energies are exponentially distributed independent random variables.

## 3.2 Random energy model

Lets consider the following partition function [1]:

$$Z = \sum_{i=1}^{M} z_i, z_i = \exp(-E_i/T) \tag{32}$$

where the energies $E_i$ are distributed as (25). This is a generalization of the original REM of Derrida (Gaussian, $\delta = 2$). Clearly, the independent variables $z_i$ are large for $E_i$ large and negative. In the scaling region the exponential distributions of the rescaled energy translates into a power law decay of $P(z)$ for large $z$:

$$P(z) \propto z^{-1-\mu}, z \longrightarrow \infty \tag{33}$$

where $\mu = T/\Delta(M)$.

The partition sum $Z$ behaves differently for $\mu < 1$, where the average value of $Z$ diverges and hence a small number of terms (of order $M^{1/\mu}$) contribute to $Z$. In the region $\mu > 1$, all $M$ terms give a small contribution to $Z$. Hence for $\mu = 1$, i.e.for

$$T_c = \Delta(M) \tag{34}$$

the probability measure concentrates onto a finite number of states, corresponding to the glass transition in these models.

In the REM, $M$ is the number of states $M = 2^N$. To ensure an extensive ground-state energy ($E_c \propto N_c$) and a finite $T_c$ in the large $N$ limit we choose $B = N^{1-\delta}$.

Lets consider the statistics of weights $p_i \equiv z_i/Z$ in the glassy region $T < T_c$:

$$\omega_i = \frac{z_i}{z_i + Z'} \tag{35}$$

where $Z' = \sum_{k(\neq i)} z_k$ is independent of $z_i$ (and of order $M^{1/\mu}$). One can find:

$$\omega_i = \frac{Z'}{(1-\omega)^2} P\left(z = \frac{Z'\omega}{1-\omega}\right) \tag{36}$$

$z_i$ has to be large for $\omega_i$ to be non-zero in the large $M$ limit. Then the asymptotic form (33) for $P(z)$ can be used:

$$P(\omega) = \frac{C}{M}(1-\omega)^{\mu-1}\omega^{-1-\mu}, \omega \gg M^{-1/\mu} \tag{37}$$

where $C$ is constant such that $M \int_0^1 \omega P(\omega) d\omega \equiv 1$.

Now one can deduce the moments $Y_k \equiv \overline{\sum_i \omega_i^k}$, which characterize the extent to which the measure concentrate onto a few states: if all weights are of the same order of magnitude, then $Y_k \sim M^{1-k} \longrightarrow 0$ for $k > 1$. If only a finite number of weights contribute, the moments $Y_k$ remain finite (for $M \longrightarrow \infty$).

In this case for $\mu < 1$:

$$Y_k = \int_0^1 \omega^k P(\omega) d\omega = \frac{\Gamma[k-\mu]}{\Gamma[k]\Gamma[1-\mu]}, k > \mu \tag{38}$$

We have $Y_2$ goes linearly to zero for $T \longrightarrow T_c$ (since $\mu = T/T_c$) and $Y_k = 1 - (\Gamma'[k] - \Gamma'[1])/\Gamma[k])T/T_c$ for $T \longrightarrow 0$. The average energy per degree of freedom of the system is constant in the low-temperature phase ($T < T_c$) and is equal:

$$\overline{E}/N = E_c/N + O(1/N) \sim -(\log 2)^{1/\delta} \tag{39}$$

### 3.3 A d-dimensional REM: extreme value approach

A generalized version of REM is considered [1], where the energy levels are embedded in a Euclidean space. It can be used to model a particle in a disordered environment and is relevant to the study of declining Burgers turbulence.

Each point $x$ of a discretized d-dimensional space is assigned with a potential energy $E(x)$, randomly chosen from a distribution $P(x)$ , the tail of which is given by (25). A deterministic part e.g. $kx^2/2$ is added. The partition function for the energy landscape is

$$Z = \int exp(-V(x)/T)d^d x \tag{40}$$

where $V(x) \equiv kx^2/2 + E(x)$. Remark: the solution of Burgers equation with a random initial velocity field can be expressed as a partition sum of the above equation.

There are two reasons to add the deterministic part. It allows to define a topology in the space of points $x$ ($k = 0$ gives REM). Also the presence of the term allows to deal with this model without the need to introduce a finite box.

One would like to compare the low temperature properties of the system on the limit $k \longrightarrow 0$: typical displacement of the ground state, measured with $\overline{\langle x^2 \rangle}$, the average ground state energy, etc.

In the special case where the energy is Gaussian distributed, this problem has been solved by scaling arguments and with a Gaussian replica variational method.

Consider $d = 1$ (the extension to higher dimensions is immediate). For temperature going to zero , one would like to find the minimum of all the energies $kx^2/2 + E(x)$. The joint probability that his minimum is achieved on a point $x^*$, with a value $V(x) = kx^{*2} + E$ is given by:

$$P(x^*, E) = P(E) \prod_{x' \neq x^*} \left(1 - P_<(E + kx^{*2} - kx'^2)\right) \tag{41}$$

The continuum limit for $k \longrightarrow 0$ is

$$P(x^*, E) = \frac{P(E)}{1 - P_<(E)} \exp\left(\int \log(1 - P_<(E + kx^{*2} - kx'^2))dx'\right) \tag{42}$$

By integrating over E one gets the probability that the minimum is achieved on $x^*$. For small $k$, the minimum $E$ is expected to be negative and large, and hence only the region where $P_<$ is small will be of importance.

Rescaling with $x^* = \hat{x}^*/\sqrt{k}$ one obtains:

$$P(\hat{x}^*) \simeq \int P(E) \exp\left(-\int P_<(E + \frac{\hat{x}^{*2} - z^2}{2})/\sqrt{k} dz\right) dE \tag{43}$$

For small k the relevant energy region is the one around $E_c$ such that $P_<(E_c) = \sqrt{k}$ or

$$E_c = -\left(\frac{\log(1/\sqrt{k})}{B}\right)^{1/\delta} \tag{44}$$

Expanding the energy around $E_c$ as $E = E_c - \hat{x}^{*2}/2 + \epsilon$ one gets:

$$P_< \left( E + \frac{\hat{x}^{*2} - z^2}{2} \right) \sim \sqrt{k} \exp(\delta B |E_c|^{\delta-1}(\epsilon - z^2/2)) \tag{45}$$

The integral over $z$ in (43) is thus a Gaussian integral. After integration over $\epsilon$:

$$P_<(\hat{x}^*) \sim \exp(-\delta B |E_c|^{\delta-1} \hat{x}^{*2}/2) \tag{46}$$

Therefore the typical distance to the origin of the point $x^*$ corresponding to a minimum energy is

$$\xi = (k\delta B|E_c|^{\delta-1})^{-1/2} = (\log(1/\sqrt{k}))^{1-\delta/2\delta} B^{-1/2\delta}/\sqrt{k\delta} \tag{47}$$

More precisely the distribution of $x^*/\xi$ is a Gaussian of unit variance.

The probability distribution of the ground state energy $\nu^*$ is

$$P(\nu^*) = \int dx \int \delta(\nu^* - kx^2/2 - E) P(x, E) dE \tag{48}$$

where $P(x, E)$ is given by (42). Defining the rescaling energy $u$ as

$$\nu^* = E_c + \frac{1}{2B\delta|E_c|^{\delta-1}} \log(B\delta|E_c|^{\delta-1}/2\pi) + u/(B\delta|E_c|^{\delta-1}) \tag{49}$$

one finds that $u$ is distributed according to the universal Gumbel equation (31)

The extremely deep states are exponentially distributed as $\exp(\mu\nu^*/T)$ with $\nu = TB\delta|E_c|^{\delta-1}$.

Remark: it was shown [1] that there is a well described relation between Parisi's replica symmetry breaking (RSB) scheme and extreme-value statistics.

# 4 Conclusion

We have overviewed some of the aspects of application of mathematical statistics to studies of models of physical system. The statistics of random variable and their extremes appeared to be an effective tool for certain physical models.

Considering the discussed applications it is self-evident that the there is much to explore in applications of the universality classes of both the sums of standard random variables and and those of the extremes.

# References

[1] J.P. Bouchaud and M. Mezard, J. Phys. A 30 (1997) 7997.

[2] E.J. Gumbel, Statistics of Extremes, Columbia University (1958)

[3] A. Rammal, J. Physique 46 (1985) 1837

[4] J.P. Bouchaud, A.Georges, Phys. Rep. 195 (1990) 4-5

[5] S.T. Bramwell, Phys. Rev. Lett. 84 (2000) 3747