Extreme value statistics of correlated random variables

Lectures by Satya N. Majumdar[1] and Notes by Arnab Pal[2]

[1] *CNRS, LPTMS, Orsay, Paris Sud*
[2] *Raman Research Institute, India*
(Dated: June 20, 2014)

Extreme value statistics (EVS) concerns the study of the statistics of the maximum or the minimum of a set of random variables. This is an important problem for any time-series and has applications in climate, finance, sports, all the way to physics of disordered systems where one is interested in the statistics of the ground state energy. While the EVS of 'uncorrelated' variables are well understood, little is known for strongly correlated random variables. Only recently this subject has gained much importance both in statistical physics and in probability theory. In this note, we will first review the classical EVS for uncorrelated variables and discuss few examples of correlated variables where analytical progress can be made.

I. INTRODUCTION

Extreme events are ubiquitous in nature. They may be rare events but when they occur, they may have devastating consequences and hence are rather important from practical points of view. To name a few, different forms of natural calamities such as earthquakes, tsunamis, extreme floods, large wildfire, the hottest and the coldest days, stock market risks or large insurance losses in finance, new records in major sports events like Olympics are the typical examples of extreme events. There has been a major interest to study these events systematically using statistical methods and the field is known as Extreme Value Statistics (EVS) [1–4]. This is a branch of statistics dealing with the extreme deviations from the mean/median of probability distributions (for a recent review on the subject see [5] and references therein). The general theory sets out to assess the type of probability distributions generated by processes which are responsible for these kind of highly unusual events. In recent years, it has been realized that extreme statistics and rare events play an equally important role in various physical/biological/financial contexts as well–for a few illustrative examples (by far not exhaustive) see [5–50]. A typical example can be found in disordered systems where the ground state energy, being the minimum energy, plays the role of an extreme variable. In addition, the dynamics at low temperatures in disordered are governed by the statistics of extremely low energy states. Hence the study of extremes and related quantities is extremely important in the field of disordered systems [6, 7, 9, 10, 15, 40, 51–54]. Another important physical system where extreme fluctuations play an important role correspond to fluctuating interfaces of the Edwards-Wilkinson/Kardar-Parisi-Zhang varities [12, 14, 18, 19, 22, 29, 55–57]. Another exciting recent area concerns the distribution of the largest eigenvalue in random matrices: the limiting distribution [58, 59] and the large deviation probabilities [60–62] of the largest eigenvalue and its various applications (for a recent review on the largest eigenvalue of a random matrix, see [63]). Extreme value statistics also appears in computer science problems such as in binary search trees and the associated search algorithms [8, 9, 13, 16, 17].

In the classical extreme value theory, one is concerned with the statistics of the maximum (or minimum) of a set of *uncorrelated* random variables. In contrast, in most of the physical systems mentioned above, the underlying random variables are typically *correlated*. In recent years, there have been some advances in the understanding of EVS of correlated variables. In these lectures, we will first review the classical EVS of uncorrelated variables. Then we will discuss the EVS of *weakly* correlated random variables with some examples. Finally few examples of *strongly* correlated random variables will be discussed.

It should be emphasised that this is not a review article, rather lecture notes based only on two lectures. It would thus be impossible to cover, in only two lectures, all aspects of this important and rapidly advancing field of extreme value statistics with an enormous range of applications spanning across disciplines–from engineering sciences all the way to physics. In these two lectures we will focus only on some basic and key concepts. Consequently we will not attempt to provide an exhaustive list of references in this broad subject and any indevertent omission of relevant references is apologised in advance.

II. EXTREME VALUE STATISTICS: BASIC PRELIMINARIES

In a given physical situation, one needs to first identify the set of relevant random variables $\{x_1, x_2, \ldots, x_N\}$. For example, for fluctuating one dimensional interfaces, the relevant random variables may denote the heights of the interface at different space points. In disordered systems such as spin glasses, $\{x_i\}$'s may denote the energy of different spin configurations for a given sample of quenched disorder. Once the random variables are identified, there are subsequently two basic steps involved : (i) to compute explicitly the joint distribution $P(\{x_i\})$ of the relevant random variables (this is sometimes very difficult to achieve) and (ii) suppose that we know the joint distribution $P(\{x_i\})$ explicitly–then from this, how to compute the distribution of some observables, such as the sample mean or the sample maximum, defined as:

$$\text{Mean} \quad \bar{X} = \frac{x_1 + x_2 + \ldots + x_N}{N} \tag{1}$$

$$\text{Maximum} \quad M = \max\left(x_1, x_2, \ldots, x_N\right). \tag{2}$$

Particular simplications occur for IID (independent and identically distributed) random variables, where the joint distribution $P(\{x_i\})$ factorises, i.e., $P(x_1, x_2, \ldots, x_N) = p(x_1)p(x_2)\ldots p(x_N)$, where each variable is chosen from the same parent density $p(x)$. Knowing the parent distribution $p(x)$, one can then easily compute the distributions, e.g., of $\bar{X}$ and of M.

For example, let us first consider $\bar{X}$. One knows that irrespective of the choice of the parent distribution (with finite variance) the PDF of the **mean** of the IID random variables tends to a Gaussian distribution for large N namely,

$$P(\bar{X}, N) \xrightarrow{N \to \infty} \frac{1}{\sqrt{2\pi\sigma^2/N}} e^{-\frac{N}{2\sigma^2}(\bar{X} - \mu)^2} \tag{3}$$

where μ, σ^2 are the mean and the variance of the parent distribution respectively. This is known as the **Central Limit Theorem** [64] and this Gaussian form is universal. However, for correlated variables, one does not know, in general, how to compute the distribution of $\bar{X}$ and consequently, there is no known universal law for the limiting distribution of the mean of a set of correlated random variables.

Similar question about universality also arises for the distribution of extremes, e.g., that of M. We will see below that, as in the case of the mean $\bar{X}$, there exist universal limit laws for the distribution of the maximum M for the case of IID variables. However, for *strongly* correlated variables, the issue of universality is wide open.

Suppose that we know the joint distribution $P(\{x_i\})$ explicitly. Then to compute the distribution of the maximum M, it is useful to define the cumulative distribution of M which can be easily expressed in terms of the joint distribution

$$Q_N(x) = \text{Prob}[M \leq x, N] = \text{Prob}[x_1 \leq x, \ x_2 \leq x, ..., x_N \leq x] \tag{4}$$

$$= \int_{-\infty}^{x} \int_{-\infty}^{x} ... \int_{-\infty}^{x} dx_1 dx_2 ... dx_N P(x_1, x_2, ..., x_N) \tag{5}$$

and the PDF of the maximum can be obtained by taking the derivative i.e. $P(M, N) = Q_N'(x)$.

III. INDEPENDENT AND IDENTICAL RANDOM VARIABLES

For IID random variables, the joint PDF factorises and we get

$$Q_N(x) = [\int_{-\infty}^{x} dy \ p(y)]^N = [1 - \int_{x}^{\infty} dy \ p(y)]^N . \tag{6}$$

This is an exact formula for the cumulative distribution of the maximum for any N. Evidently, $Q_N(x)$ depends explicitly on the parent distribution $p(x)$ for any finite N. The question is: as in the CLT of the sum of random variables discussed before, does any universality emerge for $Q_N(x)$ in the large N limit? The answer is that indeed a form of universality emerges in the large N limit, as we summarize below (for a recent review, see [5]).

It turns out that in the scaling limit when N is large, x is large, with a particular scaling combination (see below) fixed, $Q_N(x)$ approaches a limiting form:

$$Q_N(x) \xrightarrow[z=(x-a_N)/b_N \text{ fixed}]{x\to\infty, \ N\to\infty} F(\frac{x - a_N}{b_N}) \tag{7}$$

$$\text{equivalently} \quad \lim_{N\to\infty} Q_N(a_N + b_N z) = F(z) \tag{8}$$

where a_N, b_N are non-universal scaling factors that depend on the parent distribution $p(x)$, but the scaling function $F(z)$ can only be of three possible varieties $F_{I,II,III}(z)$ depending only on the large x tail of the parent distribution $p(x)$. This is known as the **Gnedenko**'s classical law of extremes (1943) [3].

A. Parent distributions with a power law tail

We consider the IID random variables whose parent distribution has a power law convergence $p(x) \sim x^{-(1+\alpha)}$ with $\alpha > 0$. We denote the scaling function as $F_1(z)$ and this is found to be

$$F_1(z) = \begin{cases} e^{-z^{-\alpha}} & \text{for } z \geq 0 \\ 0 & \text{for } z \leq 0 \end{cases} \tag{9}$$

The PDF is given by

$$f_1(z) = \frac{\alpha}{z^{\alpha+1}} e^{-z^{-\alpha}}, \qquad z \in [0, \infty) \tag{10}$$

Here one can identify $a_N = 0$, $b_N = N^{1/\alpha}$. This is the famous **Fréchet** distribution.

B. Parent distributions with a faster than power law tail

We consider the parent distributions with tails which decay faster than power law but unbounded such as $p(x) \sim e^{-x^{\delta}}$ with $\delta > 0$. In this case, one finds the scaling function to be

$$F_2(z) = e^{-e^{-z}} \tag{11}$$

where as the PDF is given by

$$f_2(z) = e^{-z-e^{-z}}, \qquad z \in (-\infty, \infty) \tag{12}$$

This is the famous **Fisher-Tippett-Gumbel** distribution [1, 2]. Here one finds $a_N = (\ln N)^{1/\delta}$, $b_N = \frac{1}{\delta}(\ln N)^{1/\delta-1}$. Since a_N is the typical value of the maximum and is defined by $N \int_{a_N} p(y)dy = 1$, the weight will be to the right of a_N. In the following we will compute the scaling functions for the parent distributions having the exponential and the Gaussian tails.

1. $p(x) \sim e^{-x}$ for $x \geq 0$

One finds that

$$Q_N(x) = [1 - e^{-x}]^N = e^{N \log[1-e^{-x}]} \sim e^{-Ne^{-x}} = e^{-e^{-(x-\log N)}} = F_2(z) \tag{13}$$

with $z = x - \log N$ and one identifies $a_N = \ln N$, $b_N = 1$.

2. $p(x) \sim e^{-x^2/2}$

In the case of Gaussian tails, we find that

$$Q_N(x) \xrightarrow{x \to \infty, \ N \to \infty} e^{-N \int_x^\infty p(y)dy} \sim e^{-Ne^{-\frac{x^2/2}{x\sqrt{\pi}}}} \sim e^{-e^{-(\frac{x^2}{2}-\log N)}} = F_2(z) \tag{14}$$

which is peaked around $x = \sqrt{2 \log N}$. One again identifies this to be the Fisher-Tippett-Gumbel distribution with $a_N = \sqrt{2 \ln N}$, $b_N = \frac{1}{\sqrt{2 \ln N}}$.

C. Parent distributions with a upper bounded support

We now consider the parent distributions with the bounded tails such as $p(x) \xrightarrow{x \to a} (a - x)^{\beta-1}$ with $\beta > 0$. In this case

$$F_3(z) = \begin{cases} e^{-(-z)^\beta} & \text{for } z \leq 0 \\ 1 & \text{for } z \geq 0 \end{cases} \tag{15}$$

The PDF is therefore given by

$$f_3(z) = \beta(-z)^{\beta-1}e^{-(-z)^\beta}, \qquad z \in (-\infty, 0] \tag{16}$$

This is well-known as the **Weibull** distribution.

Let us now summarize the results for the limiting distribution of the maximum for IID random variables in the following table.

parent distribution $p(x)$	scaling function $F(z)$	PDF of maximum $f(z)$	Nomenclature
$x^{-(1+\alpha)}$; $\alpha > 0$	$e^{-z^{-\alpha}}\theta(z)$	$\frac{\alpha}{z^{\alpha+1}}e^{-z^{-\alpha}}$; $z > 0$	**Fréchet**
e^{-x^δ}	$e^{-e^{-z}}$	$e^{-z-e^{-z}}$	**Fisher-Tippett-Gumbel**
$(a-x)^{\beta-1}$; $\beta > 0$	$e^{-(-z)^\beta}\theta(-z) + \theta(z)$	$\beta(-z)^{\beta-1}e^{-(-z)^\beta}$; $z < 0$	**Weibull**

So far, we discussed about the limiting laws in the limit of large sample size N. It turns out however that the convergence to these limiting laws is extremely slow and in simulations and experiments, it is very hard to see these limiting distributions [65, 66]. A renormalization group treatment has been recently developed that describes how these EVS distributions for IID variables approach their limiting fixed point distributions. Interested readers may discuss Refs. [65, 67–70].

D.   Order statistics

An interesting generalisation of the statistics of the global maximum corresponds to studying the statistics of successive maxima, known as the 'order' statistics [71, 72]. For a recent review on order statistics, see [5]. Consider again the set of IID random variables $\{x_1, x_2, ..., x_N\}$ and arrange them in decreasing order of their values. So, if we denote them by $M_{k,N}$ where k is the order and N is the number of variables, then

$$M_{1,N} = \max (x_1, x_2, ..., x_N)$$
$$M_{2,N} = \text{secondmax} (x_1, x_2, ..., x_N)$$
$$...$$
$$M_{k,N} = \text{k-thmax} (x_1, x_2, ..., x_N)$$
$$...$$
$$M_{N,N} = \min (x_1, x_2, ..., x_N) \tag{17}$$

and henceforth by definition $M_{1,N} > M_{2,N} > ... > M_{N,N}$. It is of interest to study the statistics of the k-th maximum $M_{k,N}$ and the statistics of the gap defined by $d_{k,N} = M_{k,N} - M_{k+1,N}$. We consider the parent distribution to be $p(x)$. Then we define the upward and the downward cumulative distributions respectively

$$p_>(x) = \int_x^\infty p(y)dy \tag{18}$$

$$p_<(x) = \int_{-\infty}^x p(y)dy \tag{19}$$

Then the cumulative probability $Q_{k,N}(x)$ that the k-th maximum stays below x is given by

$$Q_{k,N}(x) = \text{Prob}[M_{k,N} \le x] = \sum_{m=0}^{k-1} \binom{N}{m} [p_>(x)]^m [p_<(x)]^{N-m} \tag{20}$$

where m is the no. of maximums above x. The exact PDF is then given by

$$P_{k,N} = \text{Prob}[M_{k,N} = x] = N\binom{N-1}{k-1} p(x) \, [p_<(x)]^{N-k}[p_>(x)]^{k-1} \tag{21}$$

The gap distribution is also very straightforward to obtain from here

$$P_{k,N}(d) = \text{Prob}[d_{k,N} = d] = N(N-1)\binom{N-2}{k-1} \int_{-\infty}^\infty dx \, p(x)p(x-d)[p_<(x)]^{N-k-1}[p_>(x)]^{k-1} \tag{22}$$

As for the cumulative distribution for order statistics, we can again do the same analysis as before and extract the leading scaling behavior

$$Q_{k,N}(x) = \text{Prob}[M_{k,N} \le x] \xrightarrow[z=(x-a_N)/b_N \text{ fixed}]{x\to\infty, \ N\to\infty} G_k\Big(\frac{x - a_N}{b_N}\Big) \tag{23}$$

where the scaling function $G_k(z)$ is given by

$$G_k(z) = F_\mu(z) \sum_{j=0}^{k-1} \frac{[-\ln F_\mu(z)]^j}{j!} = \frac{1}{\Gamma(k)} \int_{-\ln F_\mu(z)}^\infty e^{-t} \, t^{k-1} \, dt \tag{24}$$

where $F_\mu(z)$, with $\mu = 1, 2, 3$, denote respectively the Fréchet, Fisher-Tippett-Gumbel and Weibull scaling functions for the global maximum discussed in the previous subsection. One can clearly see that for $k = 1$ (global maximum), indeed, $G_1(z) = F_\mu(z)$. As an example, if we choose the parent distribution $p(x)$ from the subclass B where $F_2(z) = e^{-e^{-z}}$ is the Fisher-Tippett-Gumbel distribution, we find $G_k(z) = e^{-e^{-z}} \sum_{j=0}^{k-1} \frac{e^{-jz}}{j!}$ and the PDF is $G'_k(z) = \frac{e^{-kz-e^{-z}}}{(k-1)!}$. This is often known as the generalized Fisher-Tippett-Gumbel law. One can also derive exactly the limiting scaling distribution of the k-th gap, $d_{k,N} = M_{k,N} - M_{k+1,N}$ (see Ref. [5] for details).

IV. CORRELATED RANDOM VARIABLES

In this section we study the extreme statistics of the correlated random variables. In the first subsection we revisit the random variables where the correlation is weak and then we study the 'strongly' interacting random variables. There is no general framework to study the statistics of the correlated random variables. However, we show in the following that for the weakly correlated variables, one can provide a a rather general renormalization group type of argument to study the extreme statistics. should hold in general. This argument does not work when the variables are strongly correlated and one has to study case by case different models to gain an insight.

A. Weakly correlated random variables

Suppose that we have a set of random variables that are not independent, but correlated such that the connected part of the correlation function decays fast (say exponentially) over a certain finite correlation length ζ

$$C_{i,j} = \langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle \sim e^{-|i-j|/\zeta} \tag{25}$$

Clearly, when two variables are separated over a length scale larger than ζ, i.e., when $|i-j| >> \zeta$, then they essentially get uncorrelated. Now weak correlation implies that $\zeta << N$, where N is the total size of the sample.

For such weakly correlated variables one can construct a heuristic argument to study the extreme statistics [19], as we describe now. Consider $N' = \zeta << N$ and break the system into identical blocks each of size ζ. There are thus N/ζ no. of blocks. While the random variables inside each box are still strongly correlated, the variables that belong to different boxes are approximately uncorrelated. So, each of these boxes are non-interacting. Now, for each box i, let y_i denote the 'local maximum', i.e., the maximum of all the x-variables belonging to the i-th block, where $i = 1, 2, \ldots, N' = N/\zeta$. By our approximation, y_i's are thus essentially *uncorrelated*. So, we have

$$M = \text{Max}[x_1, x_2, ..., x_N] = \text{Max}[y_1, y_2, ..., y_{N'}] \tag{26}$$

So, in principle if one knows the PDF of y, then this problem is essentially reduced to calculating the maximum of N' uncorrelated random variables $\{y_1, y_2, \ldots, y_{N'}\}$, which has already been discussed before. So, we know that depending on the tail of $p(y)$, the limiting distribution of M of N weakly correlated variables will, for sure, belong to one of three (Fréchet, Fisher-Tippett-Gumbel or the Weibull class) limiting extreme distributions of IID random variables. To decide the tail of $p(y)$, of course one needs to solve a *strongly* correlated problem since inside each block the variables are strongly correlated. However, one can often guess the tail of $p(y)$ without really solving for the full pdf of $p(y)$ and then one knows, for sure, to which class the distribution of the maximum belongs to. As a concrete example of this procedure for weakly corelated variables, we discuss in the next section the Ornstein-Uhlenbeck stochastic process where one can compute the EVS exactly and demonstrate that indeed this heuristic renormlaization group argument works very well. To summarize, the problem of EVS of weakly correlated random variables basically reduces to IID variables with an effective number $N' = N/\zeta$ where ζ is the correlation length. So, the real challenge is to compute the EVS of strongly correlated variables where $\zeta \geq O(N)$, to which we now turn to below.

B. Strongly correlated random variables

Strongly correlated means that $\zeta >> N$ i.e. a correlation prevails over the whole system and the idea of block spins will no longer hold. A general theory for calculating the EVS, such as in case of IID or weakly correlated variables, is currently lacking for such strongly correlated variables. In absence of a general theory, one tries to study different exactly solvable special cases in the hope of gaining some general insights. There are examples, though their numbers are unfortunately few, where the EVS of a strongly correlated system can be computed exactly. In the rest of the lecture, we will discuss a few examples of exactly solvable cases.

As a first example of a strongly correlated system, we will consider the one dimensional Brownian motion, for which the distribution of maximum can be computed explitcitly. Next, we will discuss another process called the Ornstein-Uhlenbeck (OU) process, which represents the noisy motion of a classical particle in a harmonic well. We will see that the OU process represents a weakly correlated system, for which one can compute the distribution of the maximum explicitly, demonstrating the power of heuristic argument presented in the previous subsection for weakly correlated variables. Then we will present the problem of the maximum height distribution of a fluctuating $(1+1)$-dimensional interface in its stationary state. Finally, we will discuss the statistics of the largest eigenvalue for Gaussian random matrices.

1. One dimensional Brownian motion

We consider the case where the random variables $\{x_i\}$ represent the positions of a one dimensional random walker at discrete time-step i, starting from $x_0 = 0$. We are interested in the maximum position of the walker up to step n. Even though the discrete problem can be solved explicitly, for simplicity we will consider below the continuous-time version of the random walk, i.e., a one dimensional Brownain motion whose position $x(\tau)$ evolves via the stochastic Langevin equation

$$\frac{dx}{d\tau} = \eta(\tau) \tag{27}$$

starting from $x(0) = 0$ and η represents a Gaussian white noise with $\langle \eta(\tau) \rangle = 0$ and $\langle \eta(\tau)\eta(\tau') \rangle = 2D \ \delta(\tau - \tau')$. We are interested in the PDF of the maximum $M(t)$ of this Brownian motion $x(\tau)$ over the time window $[0, t]$: $M(t) = \max_{0 \le \tau \le t} [x(\tau)]$.

To proceed, we first note that $x(\tau) = \int_0^\tau \eta(s) \, ds$ and hence, one can easily compute the mean and the correlator of the process $x(\tau)$: $\langle x(\tau) \rangle = 0$ and $\langle x(\tau)x(\tau') \rangle = 2D \min(\tau, \tau')$. Thus the variables $x(\tau)$ at different times are strongly correlated. The correlation function does not decay and persists over the whole sample size $\tau \in [0, t]$.

It is again very useful to define the cumulative distribution of the maximum

$$\begin{aligned} Q(z, t) &= \text{Prob}[M(t) \le z] \\ &= \text{Prob}[x(\tau) \le z, \ 0 \le \tau \le t] \\ \text{where } M(t) &= \max_{0 \le \tau \le t}[x(\tau)] \end{aligned} \tag{28}$$

To compute $Q(z, t)$, we note that it just represents the probability that the Brownian particle, starting at $x(0) = 0$, stays below the level z up to time t. Let $P(x, t|z)$ denote the probability density for the particle to reach x at time t, while staying below the level z during $[0, t]$. It is then easy to see that $P(x, t|z)$ satisfies the diffusion equation in the semi-infinite domain $x \in [-\infty, z]$ with the following boundary and initial conditions:

$$\begin{aligned} \frac{\partial P}{\partial t} &= D\frac{\partial^2 P}{\partial x^2} \\ \text{with } P(x, 0|z) &= \delta(x) \\ \text{and } P(x \to -\infty, t|z) &= 0, \ P(x = z, t|z) = 0 \, . \end{aligned} \tag{29}$$

The absorbing boundary condition $P(x = z, t|z) = 0$ at $x = z$ gaurantees that the particle does not cross the level at $x = z$. The solution is trivial and can be solved by the method of images [73–75]

$$P(x, t) = \frac{1}{\sqrt{4\pi Dt}}[e^{-\frac{x^2}{4Dt}} - e^{-\frac{(x-2z)^2}{4Dt}}] \tag{30}$$

Therefore the cumulative and the PDF of the maximum are given by

$$Q(z, t) = \int_{-\infty}^z dx \ P(x, t) = \text{erf}\left(\frac{z}{\sqrt{4Dt}}\right); \text{ where erf}(z) = \frac{2}{\sqrt{\pi}}\int_0^z e^{-u^2} \, du \tag{31}$$

$$\text{Prob.}(M = z, t) = Q'(z) = \frac{1}{\sqrt{\pi Dt}}e^{-\frac{z^2}{4Dt}}\theta(z) \tag{32}$$

One easily finds that the mean $\langle M(t) \rangle = \frac{2}{\sqrt{\pi}}\sqrt{Dt}$. This thus represents, perhaps, the simplest example of a strongly correlated system for which one can compute the distribution of the maximum exactly.

For more recent works on the global maximum and the order/gap statistics of discrete-time random walks, Lévy flights, $1/f^\alpha$ signals see Refs. [74–81]. The order/gap statistics for one dimensional branching Brownian motion has also been studied recently and a number of analytical results are available [82–84].

2. Ornstein-Uhlenbeck (OU) process

Consider a Brownian particle in a harmonic potential governed by the following equation

$$\frac{dx}{d\tau} = -\mu x + \eta(\tau) \tag{33}$$

where, as before, $\eta(\tau)$ is a Gaussian white noise with zero mean and is delta correlated, $\langle \eta(\tau)\eta(\tau') \rangle = 2\,D\,\delta(\tau - \tau')$. Assuming that the particle starts at the origin $x(0) = 0$, the correlation function between two times can be trivially computed and is given by

$$C(t_1, t_2) = \langle x(t_1)x(t_2) \rangle = \frac{D}{\mu}[e^{-\mu|t_1 - t_2|} - e^{-\mu(t_1 + t_2)}] \tag{34}$$

Clearly, in the limit $\mu \to 0$, the correlation function reduces to the Brownian limit: $C(t_1, t_2) \to 2\,D \min(t_1, t_2)$ as $\mu \to 0$ and the system becomes strongly correlated. In contrast, for nonzero $\mu > 0$, the correlation function, at large times when $t_1, t_2 > 1/\mu$, decays exponentially with the time-difference: $C(t_1, t_2) = \frac{1}{2\mu}e^{-\mu|t_1 - t_2|}$ with a correlation length $\zeta = 1/\mu$. From our arguments before about weakly correlated random variables, we would then expect to get the limiting Fisher-Tippett-Gumbel distribution for $\mu > 0$. We demonstrate below briefly how actually this Fisher-Tippett-Gumbel distribution emerges by actually solving exactly the EVS problem for the OU process.

As before, let $Q(z, t)$ denote the cumulative distribution of the maximum $M(t)$ of the OU process in the time interval $[0, t]$. The particle starts at the origin $x(0) = 0$ and evolves via Eq. (33). Let $P(x, t|z)$ denote the probability density for the particle to arrive at x and time t, while staying below the level z. This restricted propagator satisfies the Fokker-Planck equation in the domain $x \in [-\infty, z]$

$$\frac{\partial P}{\partial t} = D\frac{\partial^2 P}{\partial x^2} + \mu\frac{\partial}{\partial x}[xP] \tag{35}$$

with the initial condition: $P(x, 0|z) = \delta(x)$ and the boundary conditions: $P(x, t|z) = 0$ as $x \to -\infty$ and also the absorbing condition at level z: $P(x = z, t|z) = 0$ for all t. For simplicity, we will set $D = 1/2$. We note that, unlike in the Brownian case ($\mu = 0$), for $\mu > 0$ we can no longer use the method of images due to the presence of the potential. However, one can solve this equation by the eigenfunction expansion and the solution can be expressed as

$$P(x, t|z) = \sum_\lambda a_\lambda\, e^{-\lambda t}\, D_{\lambda/\mu}(-\sqrt{2\mu}x)\, e^{-\mu\, x^2/2} \tag{36}$$

where $D_p(z)$ is the parabolic cylinder function which satisfies the second order ordinary differential equation: $D_p''(z) + (p + 1/2 - z^2/4)\,D_p(z) = 0$ (out of the two linearly independent solutions, we choose the one that vanishes as $z \to \infty$). The absorbing boundary condition $P(x = z, t|z) = 0$ induces the boundary condition on the eigenfunction: $D_{\lambda/\mu}(-\sqrt{2\mu}z) = 0$, which then fixes the eigenvalues λ's. One gets the spectrum of eigenvalues and they are necessarily positive. At large times t, the summation in Eq. (36) will be dominated by the term involving the smallest eigenvalue $\lambda_0(z)$, which evidently depends on z. For arbitrary z, it is difficult to solve $D_{\lambda/\mu}(-\sqrt{2\mu}z) = 0$ and determine the smallest eigenvalue $\lambda_0(z)$. However, for large z, one can make progress by perturbation theory and one can show that to leading order for large z,

$$\lambda_0(z) \xrightarrow{z \to \infty} \frac{2}{\sqrt{\pi}}\, \mu^{3/2}\, z\, e^{-\mu z^2}. \tag{37}$$

Consequently, for large t and large z,

$$Q(z, t) \sim e^{-\lambda_0(z)\,t} \sim e^{-e^{-\mu z^2 + \log(\frac{2t\mu^{3/2}z}{\sqrt{\pi}})}}$$

$$\to F_2\left(\sqrt{4\,\mu \ln t}\left(z - \frac{1}{\sqrt{\mu}}\sqrt{\ln t}\right)\right) \tag{38}$$

where $F_2(y) = \exp[-\exp[-y]]$ is the Fisher-Tippett-Gumbel distribution. As a result, for $\mu > 0$, the average value of the maximum grows very slowly for large t as, $\langle M(t) \rangle \sim \frac{1}{\sqrt{\mu}}\sqrt{\ln t}$, while its width around the mean decreases as $\sim 1/\sqrt{\ln t}$.

Indeed for $\mu > 0$, a full analysis of the mean value of the maximum $\langle M(t) \rangle$ for all t shows that initially it grows as $\sqrt{t}$ (for $t \ll 1/\mu$) where it does not feel the confining potential and hence behaves as a Brownian motion. But for $t \gg 1/\mu$, the particle feels the potential and the mean maximum crosses over to a slower growth as $\sqrt{\ln t}$

$$\langle M(t) \rangle \sim \begin{cases} \sqrt{t} & \text{for } t \ll 1/\mu \\ \sqrt{\ln t} & \text{for } t \gg 1/\mu \end{cases} \tag{39}$$

3. Fluctuating interfaces in 1D

The most well studied model of a fluctuating $(1 + 1)$-dimensional surfaces is the so called Kardar-Parisi-Zhang (KPZ) equation that describes the time evolution of the height $H(x, t)$ of an interface growing over a linear substrate of size L via the stochastic partial differential equation

$$\frac{\partial H}{\partial t} = \frac{\partial^2 H}{\partial x^2} + \lambda \left(\frac{\partial H}{\partial x} \right)^2 + \eta(x, t) \tag{40}$$

where $\eta(x, t)$ is a Gaussian white noise with zero mean and a correlator $\langle \eta(x, t)\eta(x', t') \rangle = 2\delta(x - x')\delta(t - t')$. For $\lambda = 0$, the equation becomes linear and is known as the Edwards-Wilkinson (EW) equation. For nice reviews on fluctuating interfaces, see e.g. [85, 86]. The height is usually measured relative to the spatially averaged height i.e.

$$h(x, t) = H(x, t) - \frac{1}{L} \int_0^L H(y, t) dy \tag{41}$$

$$\text{with } \int_0^L h(x, t) dx = 0 \tag{42}$$

It can be shown that the joint PDF of the relative height field $P(\{h\}, t)$ reaches a steady state as $t \to \infty$ in a finite system of size L. Also the height variables are strongly correlated in the stationary state. Again in the context of the EVS, a quantity that has created some interests recently is the PDF of the maximum relative height in the stationary state, i.e. $P(h_m, L)$ where

$$h_m = \lim_{t \to \infty} \max_x [h(x, t), 0 \leq x \leq L]. \tag{43}$$

This is an important physical quantity that measures the extreme fluctuations of the interface heights [12, 18]. We assume that initially the height profile is flat. As time evolves, the heights of the interfaces at different spatial points grow more and more correlated. The correlation length typically grows as $\zeta \sim t^{1/z}$ where z is the dynamical exponent ($z = 3/2$ for KPZ and $z = 2$ for EW interfaces). For $t << L^z$, the interface is in the 'growing' regime where again the height variables are weakly correlated since $\zeta \sim t^{1/z} << L$. In contrast, for $t >> L^z$, the system approaches a 'stationary' regime where the correlation length ζ approcahes the system size and hence the heights become strongly correlated variables.

Following our general argument for weakly correlated variables, we would then expect that in the growing regime the maximal relative height, appropriately centered and scaled, should have the Fisher-Tippett-Gumbel distribution. In contrast, in the stationary regime, the height variables are strongly correlated and the maximal relative height h_m should have a different distribution. This distribution was first computed numerically in [12] and then it was computed analytically in Refs. [18, 19]. This then presents one of the rare solvable cases for the EVS of strongly correlated random variables. Below, we briefly outline the derivation of this distribution. On the other hand, due to the strong correlation

The joint PDF of the relative heights in the stationary state can be written putting all the constraints together [18, 19]

$$P_{st}[\{h\}] = C(L) e^{-\frac{1}{2} \int_0^L (\partial_x h)^2 \, dx} \times \delta \Big[h(0) - h(L) \Big] \times \delta \Big[\int_0^L h(x, t) dx \Big] \tag{44}$$

where $C(L) = \sqrt{2\pi L^3}$ is the normalization constant and can be obtained integrating over all the heights. Note that this stationary measure of the relative heights is independent of the coefficient λ of the nonlinear term in the KPZ equation, implying that the stationary measure of the KPZ and the EW interface is the same in $(1 + 1)$-dimension. But this is a special property only in $(1 + 1)$-dimension. The stationary measure indicates that the interface locally behaves as a Brownian motion in space [85, 86]. For an interface with periodic bounday condition, one would then have a Brownian bridge in space. However, tt turns out that the constraint $\int_0^L h(x, t) \, dx = 0$ (the zero mode being identically zero), as shown explicitly by the delta function in Eq. (44, plays an important role for the statistics of the maximal relative height [18]. It shows actually that the stationay measure of the relative heights actually corresponds to a Brownian bridge, but with a global constraint that the area under the bridge is strictly zero [18, 19]. This fact plays a crucial role for the extreme statistics of relative heights [18, 19].

We define the cumulative distribution of the maximum relative height $Q(z, L) = \text{Prob}[h_m \leq z]$. The PDF of the maximum relative height is then $P(z, L) = Q'(z, L)$. Clearly $Q(z, L)$ is also the probability that the heights at all

points in $[0, L]$ are less than z and can be formally written in terms of the path integral [18, 19]

$$Q(z, L) = C(L) \int_{-\infty}^{z} du \int_{h(0)=u}^{h(L)=u} \mathcal{D}h(x) e^{-\frac{1}{2} \int_0^L (\partial_x h)^2 \, dx}$$

$$\times \delta \Big[\int_0^L h(x, t) dx \Big] I(z, L) \tag{45}$$

where $I(z, L) = \prod_{x=0}^{L} \theta[z - h(x)]$ is an indicator function which is 1 if all the heights are less than z and zero otherwise. Using path integral technique, this integral can be computed exactly (for details see [18, 19]). It was found that the PDF of h_m has the scaling form for all L

$$P(h_m, L) = \frac{1}{\sqrt{L}} f \left(\frac{h_m}{\sqrt{L}} \right) \tag{46}$$

where the scaling function can be computed explicitly as [18, 19]

$$f(x) = \frac{2\sqrt{6}}{x^{10/3}} \sum_{k=1}^{\infty} e^{-\frac{b_k}{x^2}} b_k^{2/3} U \left(-\frac{5}{6}, \frac{4}{3}, \frac{b_k}{x^2} \right) \tag{47}$$

where $U(a, b, y)$ is the confluent hypergeometric function and $b_k = \frac{2}{27} \alpha_k^3$, where α_k's are the absolute values of the zeros of Airy function: $\mathrm{Ai}(-\alpha_k) = 0$. It is easy to obtain the small x behavior of x since only the $k = 1$ term dominates as $x \to 0$. Using $U(a, b, y) \sim y^{-a}$ for large y, we get as $x \to 0$,

$$f(x) \to \frac{8}{81} \alpha_1^{9/2} x^{-5} \exp \left[-\frac{2\alpha_1^3}{27x^2} \right] \tag{48}$$

The asymptotic behavior of $f(x)$ at large x can be obtained as

$$f(x) \xrightarrow{x \to \infty} e^{-6x^2} \tag{49}$$

It turns out, rather interestingly, that this same function has appeared before in several different problems in computer science and probability theory and is known in the literature as the Airy distribution function (for a review on this function and its appearences in different contexts see Ref. [19] and references therein).

The path integral technique mentioned above for computing the maximal relative height distribution of the EW/KPZ stationary interfaces have subsequently been generalised to more complex interfaces [22, 29, 55–57].

4. Largest eigenvalue in the random matrix theory

Another beautiful solvable example of the extremal statistics of strongly correlated variables can be found in the random matrices [58, 59], see [63] for a recent review from physics perspectives. Let us consider a $N \times N$ Gaussian random matrices with real symmetric, complex Hermitian, or quaternionic self-dual entries $X_{i,j}$ distributed via the joint Gaussian law:

$$\Pr[\{X_{i,j}\}] \propto \exp[-\frac{\beta}{2} N \mathrm{Tr}(X^2)], \tag{50}$$

where β is the Dyson index. The distribution is invariant respectively under orthogonal, unitary and symplectic rotations giving rise to the three classical ensembles: Gaussian orthogonal ensemble (GOE), Gaussian unitary ensemble (GUE) and Gaussian symplectic ensemble (GSE). The quantized values of β are respectively $\beta = 1$ (GOE), $\beta = 2$ (GUE) and $\beta = 4$ (GSE). The eigenvalue and eigenvectors are random and their joint distribution decouple. Integrating out the eigenvectors we focus here only on the statistics of N eigenvalues $\lambda_1, \lambda_2, ..., \lambda_N$ which are all real. The joint PDF of these eigenvalues is given by the classical result

$$P_{\mathrm{joint}}(\lambda_1, \lambda_2, ..., \lambda_N) = B_N(\beta) \exp \left[-\frac{\beta}{2} N \sum_{i=1}^{N} \lambda_i^2 \right] \prod_{i<j} |\lambda_i - \lambda_j|^\beta, \tag{51}$$

where $B_N(\beta)$ is the normalization constant. For convenience, we rewrite the statistical weight as

$$P_{\mathrm{joint}}(\lambda_1, \lambda_2, ..., \lambda_N) = B_N(\beta) \exp \left[-\beta \Big(\frac{N}{2} \sum_{i=1}^{N} \lambda_i^2 - \frac{1}{2} \sum_{i \neq j} \ln |\lambda_i - \lambda_j| \Big) \right], \tag{52}$$

We first note that the eigenvalues are strongly correlated via a logarithmic kind of interaction. Secondly this can also be interpreted as a Gibbs-Boltzmann measure ($\sim e^{-\beta E(\{\lambda_i\})}$) of an interacting gas of charged particles on a line where λ_i denotes the position of the ith charge and β plays the role of the inverse temperature. The energy $E(\{\lambda_i\})$ has two parts: each pair of charges repel each other via a 2-d Coulomb repulsion and each charge is pinned to an external confining parabolic potential. Having the joint PDF, we again ask the same question-whether we can predict some qualitative as well as quantitative features of the largest eigenvalue of the set namely

$$\lambda_{\max} = \max_{1 \leq i \leq N}\{\lambda_i\} \tag{53}$$

There is a competition between the pairwise Coulomb interaction and the external harmonic potential. As a result of which, the system of charges settles down into an equilibrium configuration on average and the average density of the charges will be

$$\rho_N(\lambda) = \frac{1}{N}\Big\langle \sum_{i=1}^{N} \delta(\lambda - \lambda_i) \Big\rangle \tag{54}$$

where the angular brakets denote an average over with respect to the joint PDF. This is a well known result due to Wigner who showed that as $N \to \infty$, the average density approaches an N-independent limiting form which has a semi-circular shape on the compact support $[-\sqrt{2}, \sqrt{2}]$

$$\lim_{N \to \infty} \rho_N(\lambda) = \tilde{\rho}_{sc}(\lambda) = \frac{1}{\pi}\sqrt{2 - \pi^2} \tag{55}$$

where $\tilde{\rho}_{sc}(\lambda)$ is called the Wigner semi-circular law. Hence our first observation is that the maximum eigenvalue resides near the upper edge of the Wigner semi-circle:

$$\lim_{N \to \infty} \langle \lambda_{\max} \rangle = \sqrt{2} \tag{56}$$

However, for large but finite N, $\lambda_{\max}$ will fluctuate from sample to sample and the interseting thing would be to compute the cumulative distribution which is

$$Q_N(z) = \text{Prob}[\lambda_{\max} < z] \tag{57}$$

which can be written as a ratio of two partition functions

$$Q_N(z) = \frac{Z_N(z)}{Z_N(z \to \infty)}, \tag{58}$$

$$Z_N(z) = \int_{-\infty}^{w} d\lambda_1 ... \int_{-\infty}^{w} d\lambda_N \exp\left[-\beta\Big(\frac{N}{2}\sum_{i=1}^{N}\lambda_i^2 - \frac{1}{2}\sum_{i \neq j}\ln|\lambda_i - \lambda_j|\Big)\right] \tag{59}$$

where the partition function describes a 2-d Coulomb gas, confined on a 1-d line and subject to a harmonic potential, in the presence of a hard wall at z. The study of this ratio of two partition functions reveals the existence of two distinct scales correspond to (i) **typical** fluctuations of the top eigenvalue, where $\lambda_{\max} = \mathcal{O}(N^{-2/3})$ and (ii) **atypical** large fluctuations, where $\lambda_{\max} = \mathcal{O}(1)$. It can be shown that the typical fluctuations are governed by

$$\lambda_{\max} = \sqrt{2} + \frac{1}{\sqrt{2}}N^{-2/3}\chi_\beta \tag{60}$$

where χ_β is an N-independent random variable. Its cumulative distribution, $\mathcal{F}_\beta = \text{Prob}[\chi_\beta \leq x]$, is known as the β-Tracy-Widom (TW) distribution which is known only for $\beta = 1, 2, 4$ [58, 59]. For arbitrary $\beta > 0$, it can be shown that this PDF has asymmetric non-Gaussian tails [58, 59],

$$\mathcal{F}'_\beta(x) \approx \begin{cases} \exp\left[-\frac{\beta}{24}|\text{x}|^3\right], & x \to -\infty \\[2mm] \exp\left[-\frac{2\beta}{3}\text{x}^{3/2}\right] & x \to +\infty \end{cases} \tag{61}$$

While the TW density describes the probability of typical fluctuations of $\lambda_{\max}$ around its mean $\langle \lambda_{\max} \rangle = \sqrt{2}$ on a small scale of $\mathcal{O}(N^{-2/3})$, it does not describe atypically the large fluctuations, e.g. of order $\mathcal{O}(1)$ around its mean. The probability of atypically large fluctuations, to leading order for large N, is described by two large deviations (or rate) functions $\Phi_-(x)$ (for fluctuations to the left of the mean) and $\Phi_+(x)$ (for fluctuations to the right of the mean),

which were computed respectively in Ref. [60, 61] and in Ref. [62]. The left and the right tails are equivalent to the physical situations of pushed and pulled Coulomb gas against the wall (see [63] for a review). We now present the asymptotic results which are known for the PDF of the $\lambda_{\max}$ [63]

$$
P(\lambda_{\max} = w, \ N) \approx
\begin{cases}
\exp\left[-\beta N^2 \Phi_-(w)\right], & w < \sqrt{2} \text{ and } |w - \sqrt{2}| \sim \mathcal{O}(1) \\[2mm]
\sqrt{2}N^{2/3}\mathcal{F}'_\beta\left[\sqrt{2}N^{2/3}(w - \sqrt{2})\right], & |w - \sqrt{2}| \sim \mathcal{O}(N^{-2/3}) \\[2mm]
\exp\left[-\beta N^2 \Phi_+(w)\right], & w > \sqrt{2} \text{ and } |w - \sqrt{2}| \sim \mathcal{O}(1)
\end{cases}
\tag{62}
$$

The corresponding large deviation functions as one approaches the critical point from below and above are respectively given by

$$
\Phi_-(w) \sim \frac{1}{6\sqrt{2}}(\sqrt{2} - w)^3, \quad w \to \sqrt{2}
\tag{63}
$$

$$
\Phi_+(w) \sim \frac{2^{7/4}}{3}(w - \sqrt{2})^{3/2}, \quad w \to \sqrt{2}
\tag{64}
$$

We refer to the review [63] and the references therein for more details.

V. SUMMARY AND CONCLUSION

To conclude, we have made a brief overview on the subject of extreme value statistics. We have seen that for uncorrelated or weakly correlated random variables, one has a fairly good understanding of the distribution of extremum and their limiting laws: there exists essentially three limiting classes named **Fréchet, Gumbel, Weibull** respectively. On the other hand, there are very few exact results known for the *strongly* correlated random variables. In these lectures, we have discussed a few of them, but were not able to cover all of them. Most of the theoretical efforts are focussed in finding more and more exactly solvables cases which may shed some light on the issue of the universality classes of EVS for strongly correlated random variables. Identifying the universality classes (if they exist) for strongly correlated random variables is thus a very challenging and outstanding open problem.

Apart from the issue of the universality of the distribution of extreme values for strongly correlated variables, there are other important questions related to extremes that have been studied extensively in the recent past. These include the statistics of the time at which the extremes occur in a given time-series, statistics of record values and their ages in a time-series and many other such questions, with applications in finance, sports, climates, ecology, disordered systems, all the way to evolutionary biology. Unfortunately, all these interesting topics can not be covered in these two lectures. For some of these aspects, we refer the readers to a few recent reviews (from the physics literature) on extremes and related subjects and their applications [5, 25, 39, 49, 54, 63, 74, 75].

[1] R.A. Fisher and L.H.C. Tippett, Proc. Cambridge Phil. Soc. **24**, 180190 (1928).
[2] E.J. Gumbel, *Statistics of Extremes* (Dover, New York, 1958).
[3] B. V. Gnedenko, Annals of Mathematics **44**, 423453 (1943).
[4] M.R. Leadbetter, G. Lindgren, and H. Rootzen, *Extremes and related properties of random sequences and processes*(Springer-Verlag, New York, 1982).
[5] G. Schehr and S.N. Majumdar, "Exact record and order statistics of random walks via first-passage ideas", a book chapter in *"First-Passage Phenomena and Their Applications"*, Eds. R. Metzler, G. Oshanin, S. Redner. World Scientific (2013), also available on arXiv: 1305:0639
[6] B. Derrida, Phys. Rev. B **24**, 2613 (1981).
[7] J.-P. Bouchaud and M. Mezard, J. Phys. A: Math. Gen. **30**, 7997 (1997).
[8] P.L. Krapivsky and S.N. Majumdar, Phys. Rev. Lett. **85**, 5492 (2000).
[9] S.N. Majumdar and P.L. Krapivsky, Phys. Rev. E **62**, 7735 (2000).
[10] D.S. Dean and S.N. Majumdar, Phys. Rev. E **64**, 046121 (2001).
[11] T. Antal, M. Droz, G. Gyorgyi, and Z. Racz, Phys. Rev. Lett. **87**, 240601 (2001).
[12] S. Raychaudhuri, M. Cranston, C. Przybla and Y. Shapir, Phys. Rev. Lett. 87, 136101 (2001).
[13] S.N. Majumdar and P.L. Krapivsky, Phys. Rev. E **65**, 036127 (2002).

[14] G. Gyorgyi, P.C.W. Holdsworth, B. Portelli and Z. Racz, PHys. Rev. E **68**, 056116 (2003).
[15] P. Le Doussal and C. Monthus, Physica A **317**, 140 (2003).
[16] S.N. Majumdar and P.L. Krapivsky, Phyica A **318**, 161 (2003).
[17] S.N. Majumdar, D.S. Dean, P.L. Krapivsky, Pramana, Volume 64, Issue 6, 1175-1189, June 2005, also available on arXiv:cond-mat/0410498
[18] S.N. Majumdar and A. Comtet, Phys. Rev. Lett. 92, 225501 (2004).
[19] S.N. Majumdar and A. Comtet, J. Stat. Phys. 119, 777 (2005).
[20] M.J. Kearney and S.N. Majumdar, J. Phys. A: Math. Gen. **38, 4097 (2005)**.
[21] E. Bertin and M. clusel, J. Phys. A: Math. Gen. **39**, 7607 (2006).
[22] G. Gyorgyi, N. R. Moloney, K. Ozogany, Z. Racz, Phys. Rev. E **75**, 021123 (2007).
[23] S. Sabhapandit and S.N. Majumdar, Phys. Rev. Lett. **98**, 140201 (2007).
[24] I. Bena and S.N. Majumdar, Phys. Rev. E **75**, 051103 (2007).
[25] J. Krug, *Records in a changing world*, J. Stat. Mech. P07001 (2007).
[26] J. Randon-Furling and S.N. Majumdar, J. Stat. Mech. P10008 (2007).
[27] C. Sire, Phys. Rev. Lett. **98**, 020601 (2007).
[28] C. Sire, J. Stat. Mech. P08013 (2007).
[29] T. W. Burkhardt, G. Gyorgyi, N. R. Moloney, and Z. Racz, Phys. Rev. E **76**, 041119 (2007).
[30] M.R. Evans and S.N. Majumdar, J. Stat. Mech. P05004 (2008).
[31] S.N. Majumdar, J.-P. Bouchaud, Quant. Fin. **8**, 753 (2008).
[32] S.N. Majumdar, J. Randon-Furling, M.J. Kearney, and M. Yor, J. Phys. A: Math. Theor. **41**, 365005 (2008).
[33] S.N. Majumdar and R.M. Ziff, Phys. Rev. Lett. **101**, 050601 (2008).
[34] C. Godreche and J.M. Luck, J. Stat. Mech. P11006 (2008).
[35] G. Schehr, S.N. Majumdar, A. Comtet, and J. Randon-Furling, Phys. Rev. Lett. **101**, 150601 (2008).
[36] J. Randon-Furling, S.N. Majumdar, and A. Comtet, Phys. Rev. Lett. **103**, 140602 (2009).
[37] P. Le Doussal and K. J. Wiese, Phys. Rev. E **79** 051105 (2009).
[38] C. Godreche, S.N. Majumdar, and G. Schehr, Phys. Rev. Lett. **102**, 240602 (2009).
[39] S.N. Majumdar, A. Comtet, and J. Randon-Furling, *Random convex hulls and extreme value statistics*, J. Stat. Phys. **138**, 955 (2010).
[40] S.N. Majumdar, A. Rosso, and A. Zoia, Phys. Rev. Lett. **104**, 020602 (2010).
[41] S.N. Majumdar, A. Rosso, and A. Zoia, J. Phys. A: Math. Theor. 43, 115001 (2010).
[42] G. Schehr and P. Le Doussal, J. Stat. Mech. P01009 (2010).
[43] J. Neidhart and J. Krug, Phys. Rev. Lett. **107**, 178102 (2011).
[44] J. Rambeau and G. Schehr, Europhys. Lett. **91**, 60006 (2010); Phys. Rev. E **83**, 061146 (2011).
[45] P.J. Forrester, S.N. Majumdar, and G. Schehr, Nucl.Phys. B **844**, 500 (2011).
[46] G. Schehr, J. Stat. Phys. **149**, 385 (2012).
[47] G. Schehr, S.N. Majumdar, A. Comtet, and P.J. Forrester, J. Stat. Phys. **150**, 491 (2013).
[48] E. Dumonteil, S. N. Majumdar, A. Rosso, and A. Zoia, PNAS, **110**, 4239 (2013).
[49] G. Wergen, *Records in stochastic processes – Theory and applications*, J. Phys. A: Math. Theo., **46**, 223001 (2013).
[50] C. Godreche, S.N. Majumdar, and G. Schehr, J. Phys. A: Math. Theor. **47**, 255001 (2014).
[51] Y.V. Fyodorov and J.-P. Bouchaud, J. Phys. A: Math. Theor. **41**, 372001 (2008).
[52] Y.V. Fyodorov, J. Stat. Mech. P07002 (2009).
[53] Y.V. Fyodorov, P. Le doussal, and A. Rosso, J. Stat. Mech. P10005 (2009).
[54] Y.V. Fyodorov, *Multifractality and Freezing Phenomena in Random Energy Landscapes: an Introduction*, Physica A **389**, 4229 (2010).
[55] G. Schehr and S.N. Majumdar, Phys. Rev. E **73**, 056103 (2006).
[56] J. Rambeau and G. Schehr, J. Stat. Mech., P09004 (2009).
[57] J. Rambeau, S. Bustingorry, A.B. Kolton, and G. Schehr, Phys. Rev. E 84, 041131 (2011).
[58] C. A. Tracy, H. Widom, Commun. Math. Phys. 159, 151 (1994).
[59] C. A. Tracy, H. Widom, Commun. Math. Phys. 177, 727 (1996).
[60] D.S. Dean and S.N. Majumdar, Phys. Rev. Lett. **97**, 160201 (2006).
[61] D.S. Dean and S.N. Majumdar, Phys. Rev. E **77**, 041108 (2008).
[62] S.N. Majumdar, M. Vergassola, Phys. Rev. Lett. **102**, 060601 (2009).
[63] S.N. Majumdar and G. Schehr, J. Stat. Mech. P01012 (2014).
[64] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1, 3rd Edition, (Wiley, New Jersey, 1971).
[65] G. Gyorgyi, N. R. Moloney, K. Ozogany, Z. Racz, Phys. Rev. Lett. **100**, 210601 (2008).
[66] M. Taghizadeh-Popp, K. Ozogany, Z. Racz, E. Regoes, and A. S. Szalay, Astrophys. J. **759**, 100 (2012).
[67] G. Gyorgyi, N. R. Moloney, K. Ozogany, Z. Racz, M. Droz, Phys. Rev. E **81**, 041135 (2010).
[68] E. Bertin and G. Gyorgyi, J. Stat. Mech. P08022 (2010) .
[69] F. Angeletti, E. Bertin, P. Abry, J. Phys. A: Math. Theor. **45**, 115004 (2012).
[70] I. Calvo, J.C. Cuch, J.G. Esteve, F. Falceto, Phys. Rev. E **86**, 041109 (2012).
[71] B. C. Arnold, N. Balakrishnan and H. N. Nagaraja, *A first course in order statistics* (Wiley, New York, 1992).
[72] H. N. Nagaraja, H. A. David, *Order statistics* (third ed.) (Wiley, New Jersey, 2003).
[73] S. Redner, *A guide to first-passage processes* (Cambridge University Press, Cambridge, (2001)).
[74] S.N. Majumdar, *Universal first-passage properties of dis crete-time random walks and Lévy flights on a line: Statistics of*

the global max imum and records, Physica A 389, 42994316 (2010).

[75] A.J. Bray, S.N. Majumdar, and G. Schehr, *Persistence and First-Passage Properties in Non-equilibrium Systems*, Adv. in Phys. **62**, 225-361 (2013).

[76] A. Comtet and S.N. Majumdar, J. Stat. Mech. P06013 (2005).

[77] N. R. Moloney, K. Ozogany and Z. Racz, Phys. Rev. E **84**, 061101 (2011).

[78] G. Schehr and S.N. Majumdar, Phys. Rev. Lett. **108**, 040601 (2012).

[79] J. Franke and S.N. Majumdar, J. Stat. Mech. P05024 (2012).

[80] S.N. Majumdar, P. Mounaix, and G. Schehr, Phys. Rev. Lett. **111**, 070601 (2013).

[81] A. Perret, A. Comtet, S.N. Majumdar, and G. Schehr, Phys. Rev. Lett. **111**, 240601 (2013).

[82] E. Brunet and B. Derrida, Europhys. Lett. **87**, 60010 (2009).

[83] E. Brunet and B. Derrida, J. Stat. Phys. **143**, 420 (2011).

[84] K. Ramola, S.N. Majumdar, and G. Schehr, Phys. Rev. Lett. **112**, 210602 (2014).

[85] T. Halpin-Healy and Y.C. Zhang, Phys. Rep. **254**, 215 (1995).

[86] J. Krug, Adv. Phys. **46**, 139 (1997).