Fluctuations and Extreme Values in Multifractal Patterns[1]

Yan V Fyodorov

School of Mathematical Sciences

Queen Mary
University of London

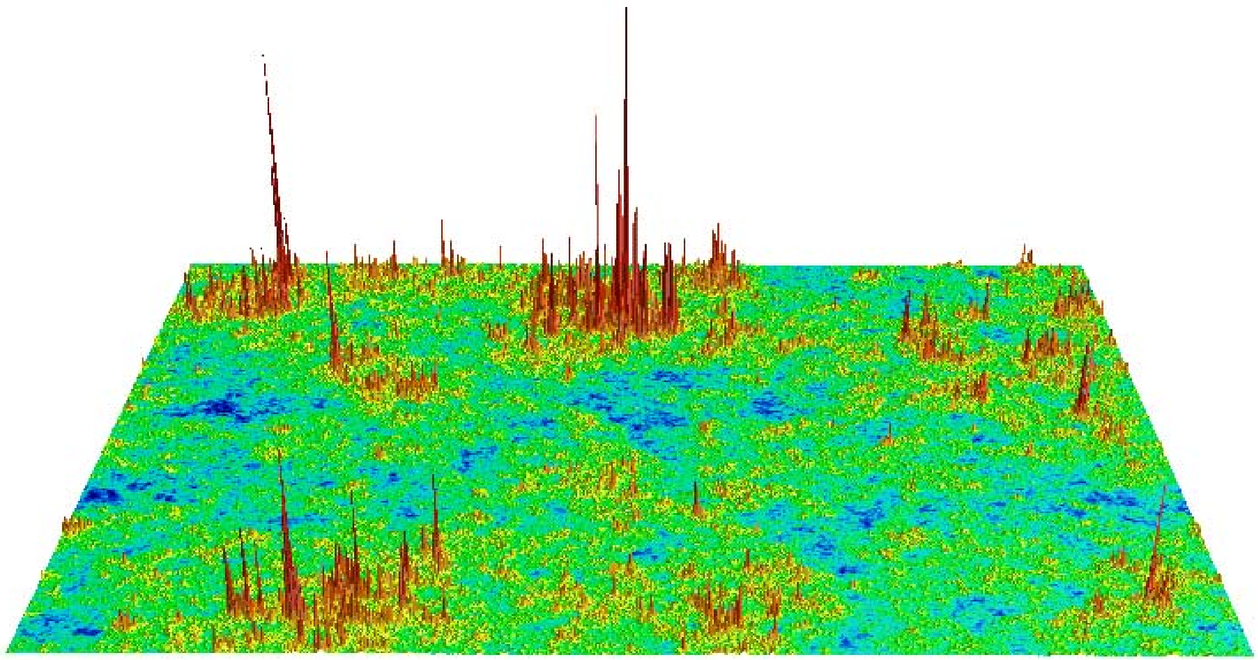
Project supported by the EPSRC grant EP/J002763/1

D.I. Diakonov memorial symposium, St. Petersburg, 13/08/2013

[1] Based on: **YVF**, **P Le Doussal** and **A Rosso** J Stat Phys: **149** (2012), 898-920

Disorder-generated multifractals:

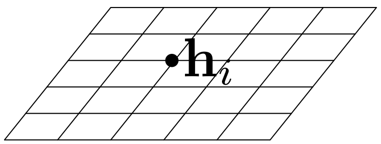
Disorder-generated multifractal patterns display high variability over a wide range of space or time scales, associated with huge fluctuations in intensity which can be visually detected. Another common feature is presence of certain long-ranged **powerlaw-type correlations** in data values.



Intensity of a multifractal wavefunction at the point of Integer Quantum Hall Effect.
Courtesy of F. Evers, A. Mirlin and A. Mildenberger.

Multifractal Ansatz:

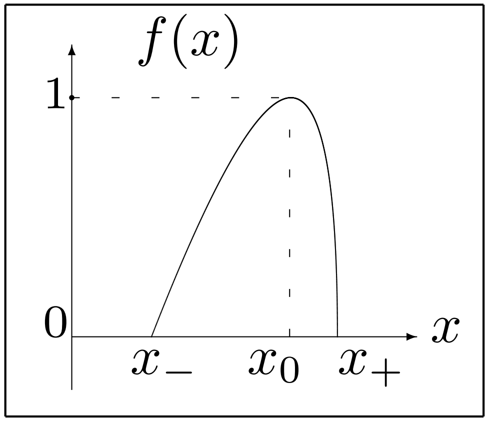
Consider a certain (e.g. hypercubic) lattice of linear extent L and lattice spacing a in $d-$dimensional space, with $M \sim (L/a)^d \gg 1$ being the total number of sites in the lattice. The multifractal patterns are then usually associated with a set of non-negative "heights" $h_i \geq 0$ attributed to every lattice site $i = 1, 2, \ldots, M$ such that the heights scale in the limit $M \to \infty$ differently at different sites:



$$h_i \geq 0, \quad h_i \sim M^{x_i}$$

with exponents x_i forming a dense set such that

$$\rho_M(x) = \sum_{i=1}^{M} \delta\left(\frac{\ln h_i}{\ln M} - x\right) \approx c_M(x)\sqrt{\ln M}\, M^{f(x)}$$



We will refer below to the above form of the density as the **multifractal Ansatz**.

The major effort in the last decades was directed towards determining the shape and properties of the **singularity spectrum** function $f(x)$. In contrast, our main object of interest will be understanding the sample-to-sample fluctuations of the prefactor $c_M(x)$ in disorder-generated multifractal patterns like those in the field of **Anderson localization**. Such fluctuations are reflected in statistics of the number of lattice points i satisfying $h_i > M^x$ which is given by the **counting function**

$$\mathcal{N}_M(x) = \int_x^{\infty} \rho_M(y)\, dy.$$

Thermodynamic formalism for multifractals:

When dealing with multifractal patterns it is frequently more convenient to characterize them by the set of exponents ζ_q describing the large-M scaling behaviour of the so-called **partition functions**

$$Z_q = \sum_{i=1}^{M} h_i^q \sim M^{\zeta_q}\,, \quad \ln M \gg 1$$

The counting function $\mathcal{N}_M(x)$ is related to the partition function Z_q via the common density $\rho_M(y)$ as

$$\mathcal{N}_M(x) = \int_x^{\infty} \rho_M(y)\,dy, \quad Z_q = \int_{-\infty}^{\infty} M^{qy}\rho_M(y)\,dy$$

Substituting for $\rho_M(y)$ the **multifractal Ansatz** form $\rho_M(y) \approx c_M(y)\sqrt{\ln M}\,M^{f(y)}$ we get for $\ln M \gg 1$

$$\mathcal{N}_M(x) \approx \frac{c_M(x)}{|f'(x)|\sqrt{\ln M}}\,M^{f(x)}, \quad Z_q \approx c_M(y_*)\sqrt{\frac{2\pi}{|f''(y_*)|}}\,M^{\zeta_q}$$

where $f'(y_*) = -q$ and $\zeta_q = f(y_*) + q\,y_*$ are related by the **Legendre transform**.

Conclusion: The fluctuation properties of the **counting function** $N_M(x)$ and the **partition function** Z_q can be related to each other via the statistics of the common prefactor $c_M(x)$.

From disorder-generated multifractals to log-correlated fields:

As discussed e.g. in **Duplantier, Ludwig** 1991 disorder-generated multifractal patterns of intensities $h(\mathbf{r})$ are typically **self-similar**

$$\mathbb{E}\left\{h^q(\mathbf{r_1})h^s(\mathbf{r_2})\right\} \propto \left(\frac{L}{a}\right)^{y_{q,s}} \left(\frac{|\mathbf{r}_1-\mathbf{r}_2|}{a}\right)^{-z_{q,s}}, \quad q,s\geq 0, \quad a \ll |\mathbf{r}_1-\mathbf{r}_2| \ll L$$

and **spatially homogeneous**

$$\mathbb{E}\left\{h^q(\mathbf{r_1})\right\} = \frac{1}{M} Z_q, \quad \text{with } M=\left(\frac{L}{a}\right)^d \text{and} \quad Z_q=\sum_{\mathbf{r}} h^q(\mathbf{r}) \propto \left(\frac{L}{a}\right)^{d\zeta_q}$$

The consistency of the two conditions for $|\mathbf{r}_1-\mathbf{r}_2| \sim a$ and $|\mathbf{r}_1-\mathbf{r}_2| \sim L$ implies:

$$y_{q,s} = d(\zeta_{q+s}-1), \quad z_{q,s} = d(\zeta_{q+s}-\zeta_q-\zeta_s+1)$$

If we now introduce the field $V(\mathbf{r}) = \ln h(\mathbf{r}) - \mathbb{E}\left\{\ln h(\mathbf{r})\right\}$ and exploit the identities $\frac{d}{ds}h^s|_{s=0} = \ln h$ and $\zeta_0=1$ we arrive at the relation:

$$\mathbb{E}\left\{V(\mathbf{r_1})V(\mathbf{r_2})\right\} = -g^2 \ln \frac{|\mathbf{r}_1-\mathbf{r}_2|}{L}, \quad g^2 = d\frac{\partial^2}{\partial s \partial q}\zeta_{q+s}|_{s=q=0}$$

Conclusion: logarithm of a multifractal intensity is a **log**-correlated random field.

Strategy: To understand statistics of **high values** and **positions of extremes** of general logarithmically correlated random processes and fields we first consider the simplest case of such process which is the **Gaussian** $1/f$ **noise**.

Ideal Gaussian periodic 1/f noise:

We will use a (regularized) model for ideal Gaussian periodic **1/f** noise defined as

$$V(t) = \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} \left[v_n e^{int} + \overline{v}_n e^{-int} \right] \ , \quad t \in [0, 2\pi)$$

where $v_n, \overline{v}_n$ are **complex standard Gaussian i.i.d.** with $\mathbb{E}\{v_n \overline{v}_n\} = 1$. It implies the formal covariance structure:

$$\mathbb{E}\left\{V(t_1)V(t_2)\right\} = -2\ln|2\sin\frac{t_1 - t_2}{2}|, \quad t_1 \neq t_2$$

Regularization procedure (YVF & Bouchaud 2008): subdivide the interval $[0, 2\pi)$ by a finite number M of observation points $t_k = \frac{2\pi}{M}k$ where $k = 1, \ldots, M$, and replace the function $V(t)$, $t \in [0, 2\pi)$ with a sequence of M random mean-zero Gaussian variables V_k correlated according to the $M \times M$ **covariance matrix** $C_{km} = \mathbb{E}\left\{V_k V_m\right\}$ such that the off-diagonal entries are given by

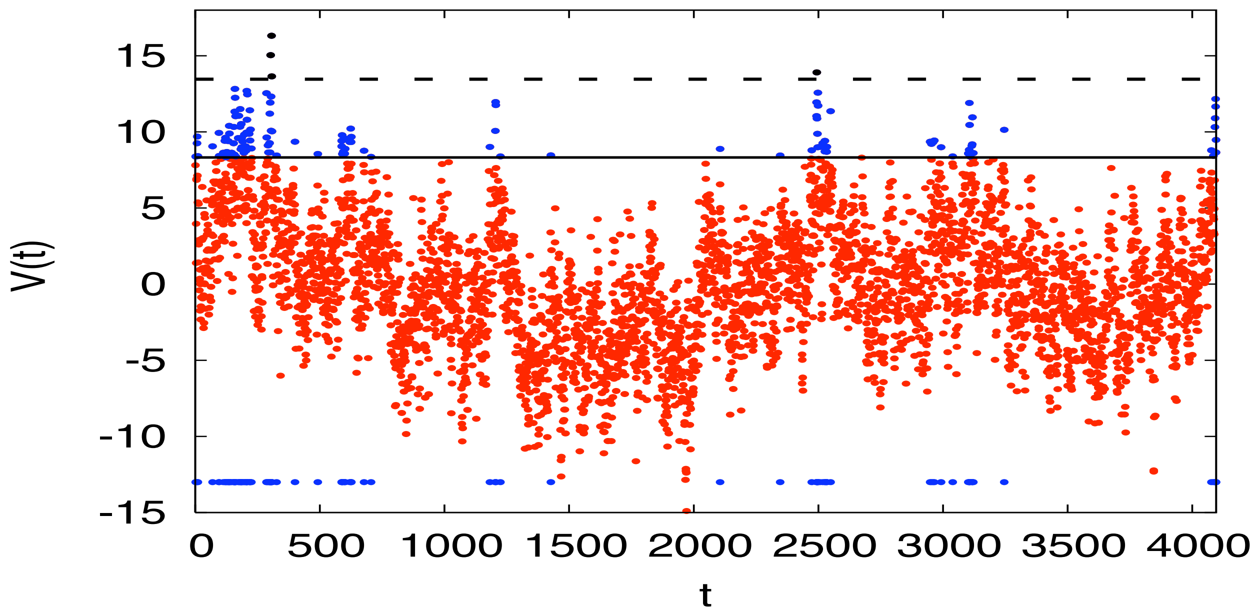
$$C_{k\neq m} = -2\ln|2\sin\frac{\pi}{M}(k-m)|, \quad C_{kk} = \mathbb{E}\left\{V_k^2\right\} > 2\ln M, \quad \forall k = 1, \ldots, M$$

The model is well defined, and we will actually take $C_{kk} = 2\ln M + \epsilon$, $\forall k$ with $\epsilon \ll 1$. We expect that the statistical properties of the sequence V_k generated in this way reflect for $M \to \infty$ correctly the universal features of the $1/f$ noise.

The **multifractal pattern** of heights is then generated by setting $h_i = e^{V_i}$ for each $i = 1, \ldots, M$.

Circular-logarithmic model (YF & Bouchaud 2008):

An example of the $1/f$ signal sequence generated for $M = 4096$ according to the above prescription is given in the figure.



The upper line marks the typical value of the **extreme value threshold** $V_m = 2\ln M - \frac{3}{2}\ln\ln M$. The lower line is the level $\frac{1}{\sqrt{2}}V_m$ and blue dots mark points supporting $V_i > \frac{1}{\sqrt{2}}V_m$.

Questions we would like to answer: How many points are typically above a given level of the noise? How strongly does this number fluctuate for $M \to \infty$ from one realization to the other? How to understand the typical position V_m and statistics of the **extreme values** (maxima or minima), etc. And, after all, what parts of the answers are **universal** and what is the universality class?

Characteristic polynomial of random CUE matrix and periodic 1/f noise:

Let U_N be a $N \times N$ **unitary matrix**, chosen at random from the unitary group $\mathcal{U}(N)$. Introduce its **characteristic polynomial** $p_N(\theta) = \det\left(1 - U_N\, e^{-i\theta}\right)$ and further consider $V_N(\theta) = -2\log|p_N(\theta)|$. Following **Hughes, Keating & O'Connell** 2001 one can employ the following representation

$$V_N^{(U)}(\theta) = -2\log|p_N(\theta)| = \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} \left[e^{-in\theta} v_n^{(N)} + \text{comp. conj.} \right]$$

where $\quad v_n^{(N)} = \frac{1}{\sqrt{n}} \text{Tr}\left(U_N^{-n}\right)$.

According to **Diaconis & Shahshahani** 1994 the coefficients $v_n^{(N)}$ for any fixed n tend in the limit $N \to \infty$ to i.i.d. complex gaussian variables with zero mean and variance $\mathbb{E}\{|\zeta_n|^2\} = 1$. We conclude that for finite N **Log-Mod** of the characteristic polynomial of CUE matrices is just a **certain regularization** of the stationary random **Gaussian Fourier series** of the form

$$V(t) = \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} \left[v_n e^{int} + \overline{v}_n e^{-int} \right], \quad t \in [0, 2\pi)$$

where $v_n, \overline{v}_n$ are **complex standard Gaussian i.i.d.** with $\mathbb{E}\{v_n \overline{v}_n\} = 1$.

Random characteristic polynomials provide natural models for $1/f$ noise!

The distribution of the partition function $Z_q = \sum_{i=1}^{M} h_i^q,\ \ h_i = e^{V_i}$:

For $M \gg 1$ the positive integer moments $\mathbb{E}\{Z_q^n\}$ for $q^2 < 1$ are given by

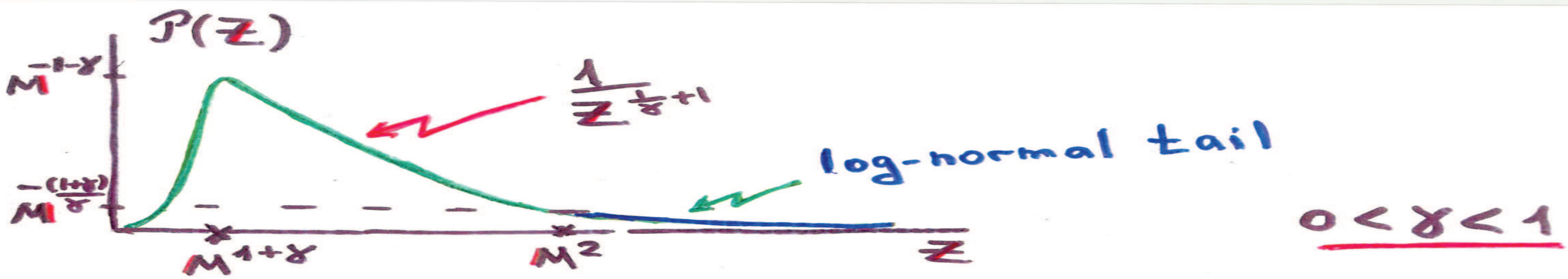
$$\mathbb{E}\left\{Z_q^n\right\}|_{M\gg 1} \approx \begin{cases} e^{n\ln M(1+q^2)}\,\mathcal{S}_n(q^2)\ , & n < 1/q^2 \\ e^{\ln M(1+n^2q^2)}\,O(1)\ , & n > 1/q^2 \end{cases}$$

where

$$\mathcal{S}_n(q^2) = \frac{n!}{\pi^n}\int_0^{\pi} d\theta_1 \int_{\theta_1}^{\pi} d\theta_2 \int_{\theta_2}^{\pi} d\theta_3 \ldots \int_{\theta_{n-1}}^{\pi} d\theta_n \prod_{p<q}^{n}\left[2\sin\left(\theta_p - \theta_q\right)\right]^{-2q^2}\ .$$

$$= \frac{\Gamma(1-nq^2)}{\Gamma^n(1-q^2)}$$ **for $1 < n < 1/q^2$ - Dyson-Morris-Selberg integral.**

One can use the above moments to restore the shape of the probability density $\mathcal{P}_q(Z)$ for the partition function $Z = Z_q$ in the whole domain $\gamma = q^2 < 1$:



$$\mathcal{P}_\gamma(Z) = \frac{1}{\gamma\, Z_e}\left(\frac{Z_e}{Z}\right)^{1+\frac{1}{\gamma}} e^{-\left(\frac{Z_e}{Z}\right)^{\frac{1}{\gamma}}},\quad Z \sim Z_e = \frac{M^{1+\gamma}}{\Gamma(1-\gamma)} \ll M^2$$

$$\mathcal{P}_\gamma(Z) = \frac{M}{\sqrt{4\pi\gamma\ln M}}\frac{1}{Z}F\left(\frac{1}{2}\frac{\ln Z}{\ln M}\right)e^{-\frac{1}{4\ln M\gamma}\ln^2 Z},\quad Z \gg M^2$$

where $F(x) \sim 1$ for $x \sim 1$.

Statistics of the counting function $\mathcal{N}_M(x)$ and threshold of extreme values:

Applying the thermodynamic formalism in our particular case we conclude that the probability density for the (scaled) counting function $n = \mathcal{N}_M(x)/\mathcal{N}_t(x)$ is given by:

$$\mathcal{P}_x(n) = \frac{4}{x^2}\, n^{-\left(1+\frac{4}{x^2}\right)}\, e^{-n^{-\frac{4}{x^2}}}, \quad n \ll n_c, \quad 0 < x < 2\,.$$

with $n_c \to \infty$ for $M \to \infty$ and the **characteristic scale** $\mathcal{N}_t(x)$ given by

$$\mathcal{N}_t(x) = \frac{M^{f(x)}}{x\sqrt{\pi \ln M}} \frac{1}{\Gamma(1-x^2/4)} \text{ with the singularity spectrum } f(x) = 1 - x^2/4.$$

In particular, the position x_m of the **threshold of extreme values** is determined from the condition $\mathcal{N}_t(x) \sim 1$. This results in

$$x_m = 2 - c\,\frac{\ln\ln M}{\ln M} + O(1/\ln M) \text{ with } c = 3/2.$$

Note that $\mathcal{N}_t(x) = \mathbb{E}\left\{\mathcal{N}_M(x)\right\}\frac{1}{\Gamma(1-x^2/4)}$. Had we instead decided to use the condition $\mathbb{E}\left\{\mathcal{N}_M(x)\right\} \sim 1$ that would give $x_m = 2 - c\,\frac{\ln\ln M}{\ln M} + O(1/\ln M)$ with $c = 1/2$. The latter value is typical for **short-ranged** random sequences. The difference is due to the fact that for $x \to 2$ the **typical** value $\mathcal{N}_t(x)$ of the counting function is parametrically smaller than the **mean** value $\mathbb{E}\left\{\mathcal{N}_M(x)\right\}$. We conjecture that this mechanism is common to all **logarithmically-correlated** processes and fields.

Distribution of the absolute maximum: partition function approach:

Given the sequence $\{V_i,\, i=1,\ldots,M\}$ we are interested in finding the **distribution** of $V_{(m)} = \max(V_1,\ldots,V_M)$ that is

$$P(v) = \text{Prob}(V_{(m)} < v) = \text{Prob}(V_i < v,\, \forall i) = \mathbb{E}\left\{\prod_{i=1}^{M} \theta(v - V_i)\right\}$$

Next we use: $$\lim_{q\to\infty} \exp\left[-e^{-q(v-V_i)}\right] = \begin{cases} 1 & v > V_i \\ 0 & v < V_i \end{cases} \equiv \theta(v - V_i)$$

which immediately shows that:

$$P(v) = \text{Prob}(V_{(m)} < v) = \lim_{q\to\infty} \mathbb{E}\left\{\exp\left[-e^{-qv} Z_q\right]\right\}, \quad \text{where} \quad Z_q = \sum_{i=1} e^{qV_i}$$

From our previous knowledge of statistics of Z_q we can readily extract the function $G_q(v) = \mathbb{E}\left\{\exp\left[-e^{-qv} Z_q\right]\right\}$ for $q < 1$. In the limit $\ln M \gg 1$ that function turns out to be of the form:

$G_q(v) = g_q\left(v - c_q \ln M\right)$ where $c_q = \left(q + \frac{1}{q}\right)$ and $g_q(v) = \int_0^\infty dt\, \exp\left\{-t - e^{-qv} t^{-q^2}\right\}$

One may further notice that not only $c_q = c_{q^{-1}}$ but the whole function satisfies a quite remarkable **duality relation**

$$g_q(v) = 1 + \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \left[e^{-nqv}\Gamma(1 - nq^2) + e^{-n\frac{v}{q}}\Gamma\left(1 - \frac{n}{q^2}\right)\right] = g_{\frac{1}{q}}(x)$$

THIS HOWEVER STILL DOES NOT ALLOW TO CONTINUE TO $q > 1$!

Freezing conjecture and the distribution of extremes:

Using certain analogy with the **Derrida-Spohn** model of polymers on disordered trees we conjecture the following **freezing scenario**: for the $log-$circular model the same sort of **freezing transition** takes place at $q = 1$. Namely, the function

$$g_{q<1}(v) = \int_0^\infty dt \exp\left\{-t - e^{-qv} t^{-q^2}\right\}$$

freezes to the **q−independent** profile $g_{q=1}(v) = 2e^{-v/2} K_1(2e^{-v/2})$ in the whole "glassy" phase $q > 1$.

Consequences:

(**i**) The latter profile then is precisely the distribution $P(v)$ of the (shifted) absolute maximum: $V_m = 2\ln M - \frac{3}{2}\ln\ln M + v$. This distribution is manifestly **non-Gumbel**, and shows the tail behaviour: $P(v \to -\infty) \approx 1 - |v|e^v$

(**ii**) The probability density of the partition function Z_q in the whole regime $q > 1$ must display a power-law forward tail of the form:

$$\mathcal{P}_{q>1}(Z) \propto Z^{-\left(1+\frac{1}{q}\right)} \ln Z$$

This shape, including the meaningful **log-factor**, is believed to be **universal** for the whole class of logarithmically correlated processes.

Numerics for the maxima of CUE characteristic polynomials:

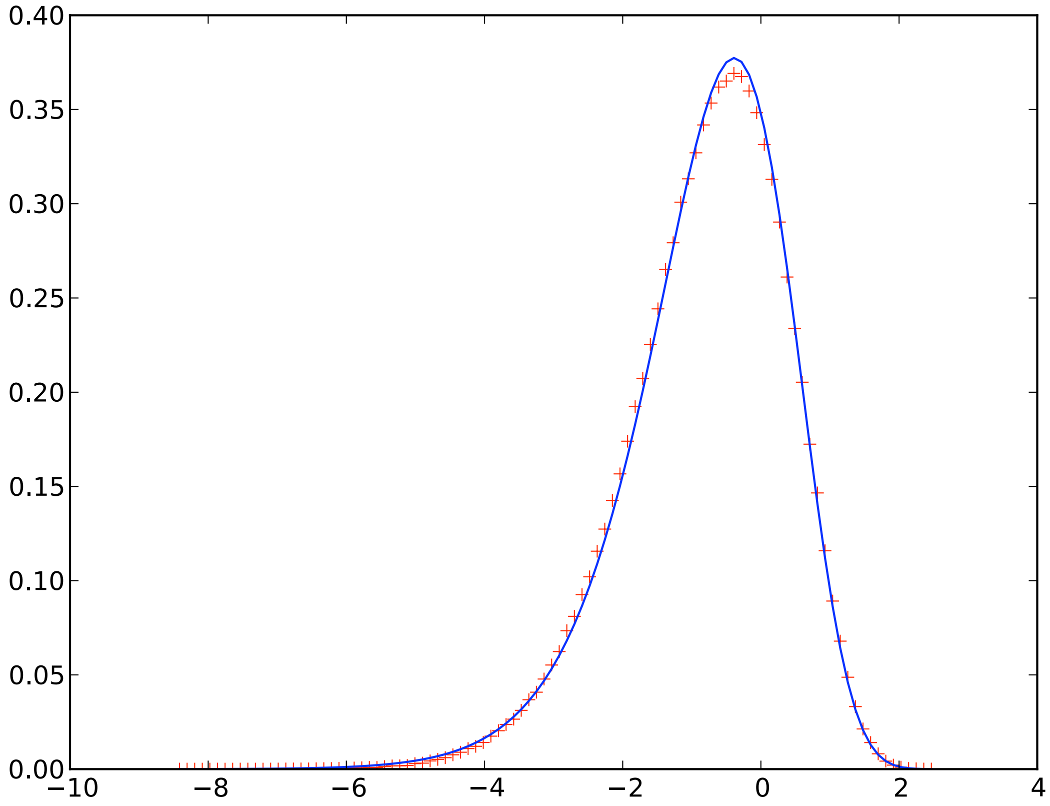


Figure 1: **Density of maxima for CUE polynomials (** $N = 50, 10^6$ **samples) compared to periodic** $1/f$ **noise prediction** $p(v) = 2e^v K_0(2e^{v/2})$**.**

Threshold of extreme values for self-similar multifractal fields:

Conjecture: the value $c=\frac{3}{2}$ is a universal feature of systems with **logarithmic** correlations.

Apart from $1/f$ noise and its incarnations (characteristic polynomials of random matrices, Riemann zeta-function along the critical line, and random Young diagrams sampled with the Plancherel measure) the new universality class is believed to include the $2D$ Gaussian free field, branching random walks $\&$ polymers on disordered trees, some models in turbulence and financial mathematics and, with due modifications the **disorder-generated multifractals**.

Namely, consider a multifractal random **probability measure** $p_i \sim M^{-\alpha_i}$, $i = 1,\ldots,M$ such that $\sum_{i=1}^{M} p_i = 1$ characterized by a general non-parabolic **singularity spectrum** $f(\alpha)$ with the left endpoint at $\alpha=\alpha_- > 0$. Then very similar consideration based on insights from **Mirlin** $\&$ **Evers** 2000 suggests that the **extreme value threshold** should be given by $p_m = M^{-\alpha_m}$, where α_m is given by

$$\alpha_m \approx \alpha_- + \frac{3}{2}\frac{1}{f'(\alpha_-)}\frac{\ln\ln M}{\ln M} \quad \Rightarrow -\ln p_m \approx \alpha_- \ln M + \frac{3}{2}\frac{1}{f'(\alpha_-)}\ln\ln M$$

For branching random walks this was recently rigorously proved: **L. Addario-Berry** $\&$ **B. Reed**2009; **E. Aidekon** 2012

Threshold of extreme values for self-similar multifractal fields:

Work in progress: testing such a prediction for multifractal eigenvectors of a $N \times N$ random matrix ensemble introduced by E. Bogomolny $\&$ O. Giraud, *Phys. Rev. Lett.* **106** 044101 (2011) based on **Rujsenaars-Schneider** model of N interacting particles. Preliminary numerics is supportive of the theory.

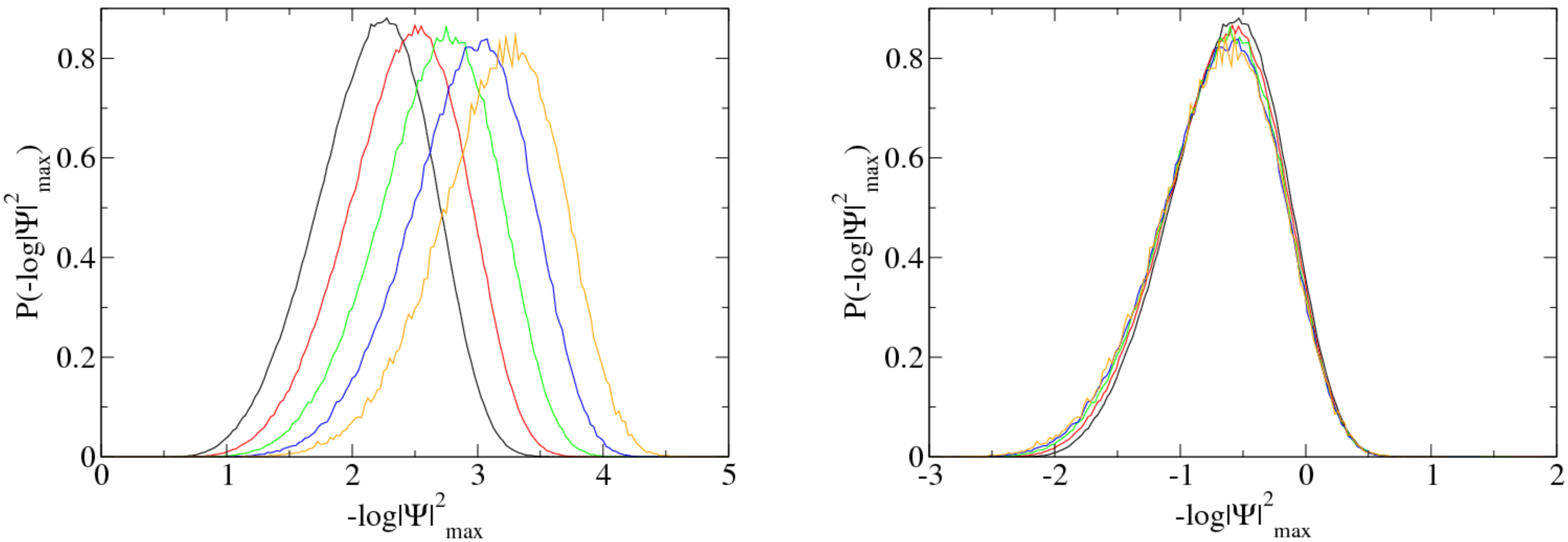


Figure 2: **Statistics of maxima for eigenvectors of RS model for sample sizes $M = 2^n$ with $n = 8, \ldots, 12$.**

left: raw data right: each curve is shifted by $\alpha_- \ln M + \frac{3}{2}\frac{1}{f'(\alpha_-)} \ln \ln M$; data by Olivier Giraud

OTHER FACETS OF THE SAME STORY:

- **Statistics of high values of Riemann** $\zeta(1/2+it)$
  YVF, **G Hiary**, **J Keating**
  Phys.Rev.Lett. 108 , 170601 (2012) $\&$ *arXiv:1211.6063*

- **Fluctuations of the shape of Young diagrams sampled with the Plancherel measure.** **YF** $\&$ **S. Nechaev**, in progress.