Fluctuations and Extreme Values in Multifractal Patterns[1]

Yan V Fyodorov

School of Mathematical Sciences
Queen Mary
University of London

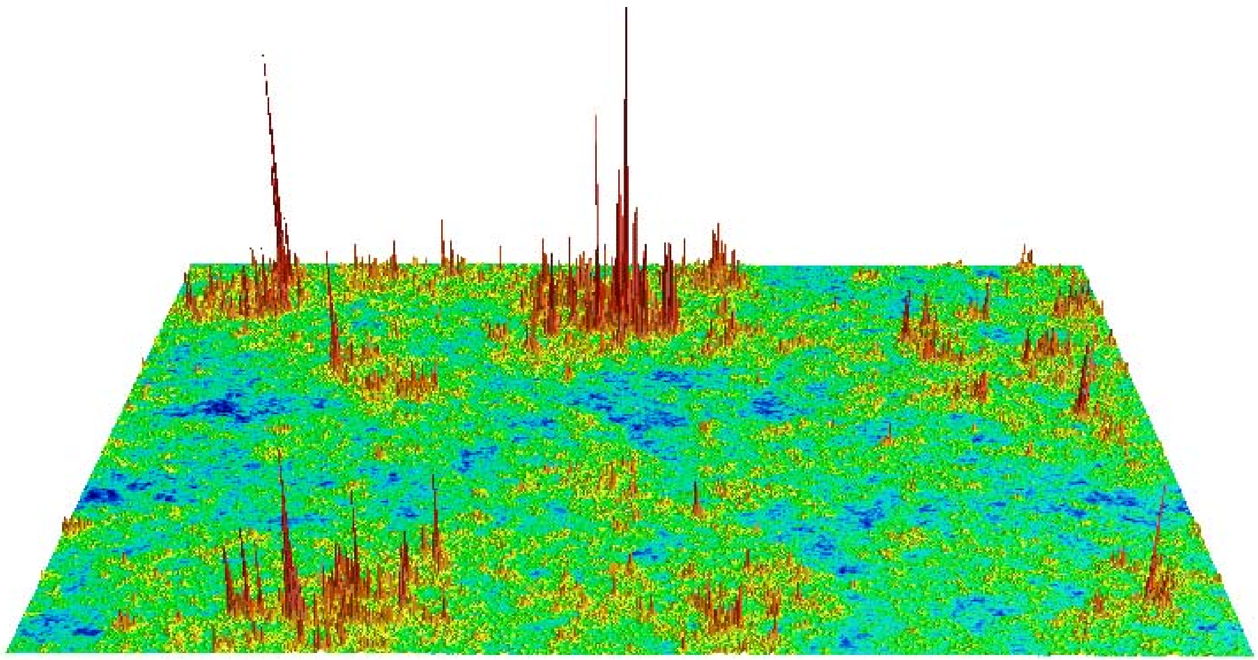
Project supported by the EPSRC grant EP/J002763/1

VIII Brunel-Bielefeld Workshop on Random Matrix Theory, 14th of December 2012

[1]Based on: **YVF**, **P Le Doussal** and **A Rosso** J Stat Phys: **149** (2012), 898-920
YVF, **G Hiary**, **J Keating** Phys. Rev. Lett. 108 , 170601 (2012) & **arXiv:1211.6063**

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.

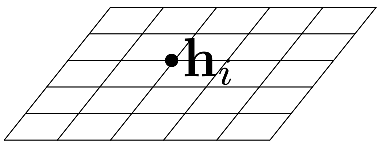


Intensity of a multifractal wavefunction at the point of Integer Quantum Hall Effect.

Courtesy of F. Evers, A. Mirlin and A. Mildenberger.

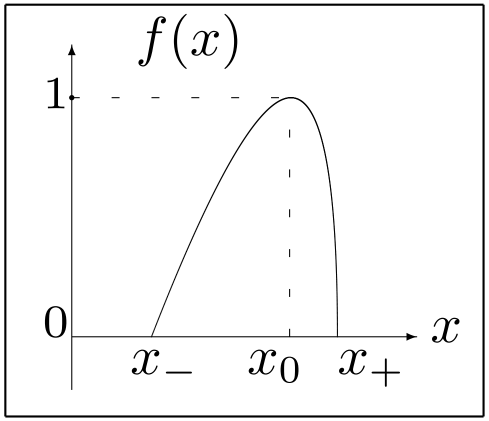
Multifractal Ansatz:

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**

$$\mathcal{N}_M(x) = \sum_i \theta\left(h_i - M^x\right) = \int_x^\infty \rho_M(y)\, dy.$$

From disorder-generated multifractals to log-correlated fields:

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**

$$\mathbb{E}\left\{h^q(\mathbf{r_1})\right\} = \frac{1}{M}\sum_{\mathbf{r}} h^q(\mathbf{r}) \propto \left(\frac{L}{a}\right)^{d(\zeta_q-1)}$$

where ζ_q and $f(x)$ are related by the **Legendre transform**:

$$f'(y_*) = -q \text{ and } \zeta_q = f(y_*) + q\, y_*.$$

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.

To understand statistics of high values and extremes of general logarithmically correlated random fields we consider the simplest $1D$ case of such process: 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

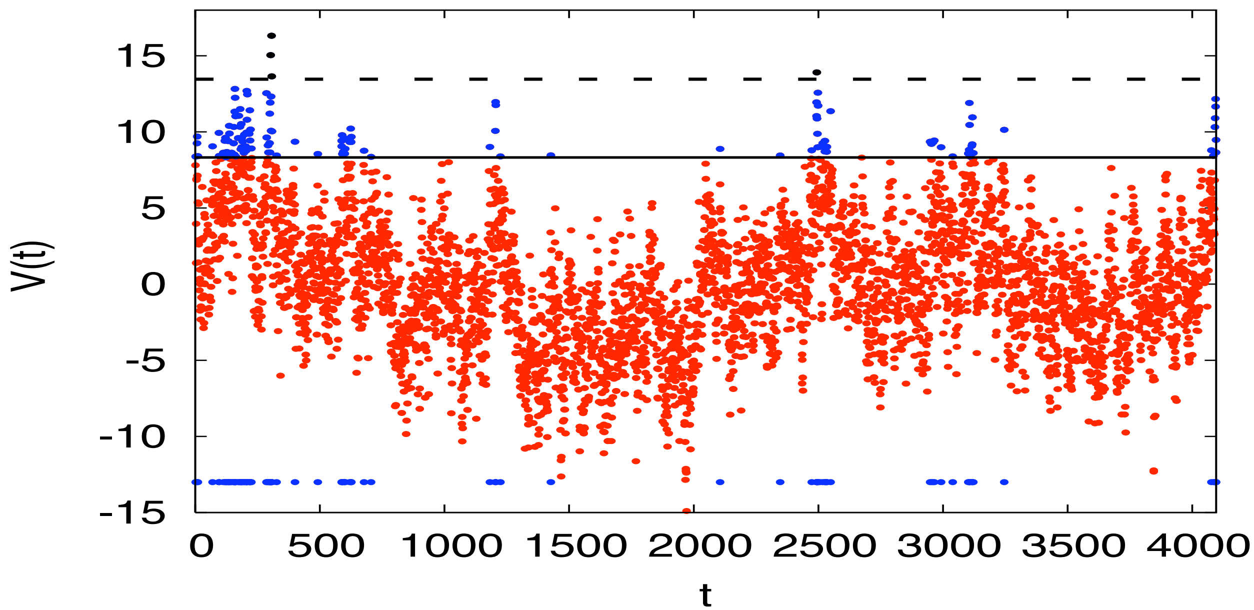
$$C_{k\neq m} = -2\ln|2\sin\frac{\pi}{M}(k-m)|, \quad C_{kk} = \mathbb{E}\left\{V_k^2\right\} > 2\ln M, \quad \forall k = 1, \ldots, M$$

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!

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

By relating moments of the counting function $\mathcal{N}_M(x) = \int_x^\infty \rho_M(y)\,dy$ for log-correlated $1/f$ **noise** to **Selberg integrals** we are able to show that the probability density for the (scaled) counting function $n = \mathcal{N}_M(x)/\mathcal{N}_t(x)$ is given by:

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

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

$$\mathcal{N}_t(x) = \frac{M^{1-x^2/4}}{x\sqrt{\pi \ln M}} \frac{1}{\Gamma(1-x^2/4)} = \mathbb{E}\left\{\mathcal{N}_M(x)\right\} \frac{1}{\Gamma(1-x^2/4)}$$

Note: For $x \to 2$ the **typical** value $\mathcal{N}_t(x)$ of the counting function is parametrically smaller than the **mean** value $\mathbb{E}\left\{\mathcal{N}_M(x)\right\}$. In particular, the position x_m of the typical threshold of **extreme values** determined from the condition $\mathcal{N}_t(x) \sim 1$ is given by

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

In contrast, for **short-ranged** random sequences **mean**$=$**typical**. Had we instead decided to use the condition $\mathbb{E}\left\{\mathcal{N}_M(x)\right\} \sim 1$ that would give

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

From $1/f$ noise to Riemann $\zeta(1/2+it)$:

One can argue that **log-mod** of the Riemann zeta-function $\zeta(1/2+it)$ **locally** resembles a (non-periodic) version of the **1/f noise**. One can exploit this fact to predict statistics of **moments** and **high values** of the Riemann zeta along the critical line using the previously exposed theory (**YVF, Hiary, Keating** 2012).

Our approach to statistics of $\zeta(1/2+it)$:

We expect a **single** unitary matrix of size $N_T = \log(T/2\pi) \gg 1$ to model the Riemann zeta $\zeta(1/2+it)$, statistically, over a range of $T \leq t \leq T+2\pi$. We thus suggest splitting the **critical line** into ranges of **length** 2π, and making the statistics of $\zeta(1/2+it)$ over the many ranges.

There are roughly N_T zeros in each range of length 2π. At each height T we use a sample that spans $\approx 10^7$ zeros yielding $\approx 10^7/N_T$ sample points.

Note: Developing the statistical mechanics analogy further we conjecture the distribution of the **absolute maximum** of the $1/f$ noises which turns out to be different from the double-exponential **Gumbel distribution** $\Phi(x) = \exp\{-e^x\}$ universally valid for short-range correlated random variables.

Our predictions for $\zeta(1/2+it)$ and CUE characteristic polynomials:

For the **maximum value**: $\zeta_{max}(T) = \max_{T\le t\le T+2\pi} |\zeta(1/2+it)|)$ we expect

$$\log \zeta_{max}(T) \approx \log N_T - \frac{c}{2}\log\log N_T + \gamma + [\text{rand. noise } O(1)],$$

with $N_T = \log(T/2\pi)$ and $\gamma = 0.57721\ldots$.

The first numerical test concerns the value of the constant c. We expect the **logarithmic correlations** to lead to $c = \frac{3}{2}$, rather than $c = \frac{1}{2}$, as would be the case if the zeta correlations were short-range. Below we give numerical estimation of c for CUE matrices of size N.

N	20	30	40	50	60	70
c	1.43570	1.46107	1.48018	1.49072	1.49890	1.50756

The mean of $\zeta_{\max}(T)$ suggested by the model is $\delta = e^{\gamma} N_T/(\log N_T)^{\frac{c}{2}}$. We give below the numerical values of the ratio of data mean $\tilde{\delta}$ to model mean δ with $c = 3/2$

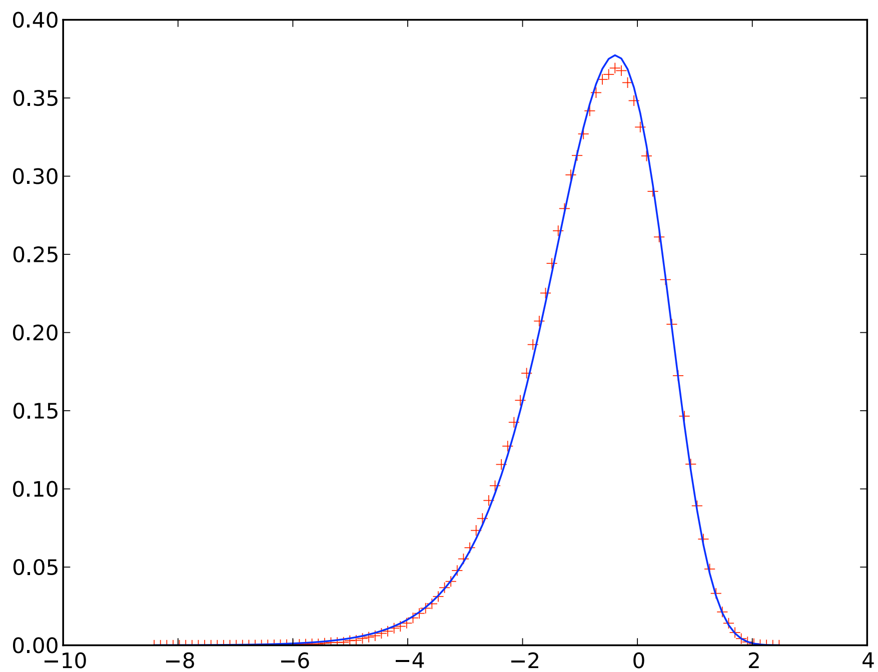
T	N_T	$\left(\tilde{\delta}/\delta\right)_{c=3/2}$
10^{22}	51	1.001343
10^{19}	44	0.992672
10^{15}	35	0.976830
3.6×10^{7}	17	0.930533

Our predictions for $\zeta(1/2+it)$ and CUE characteristic polynomials:

We further expect

$$\log \zeta_{max}(T) \approx \log N_T - \frac{3}{4}\log\log N_T - \frac{1}{2}x, \quad N_T = \log(T/2\pi)$$

where x is distributed with a probability density behaving in the tail as $\rho(x \to -\infty) \approx |x|\, e^x$.



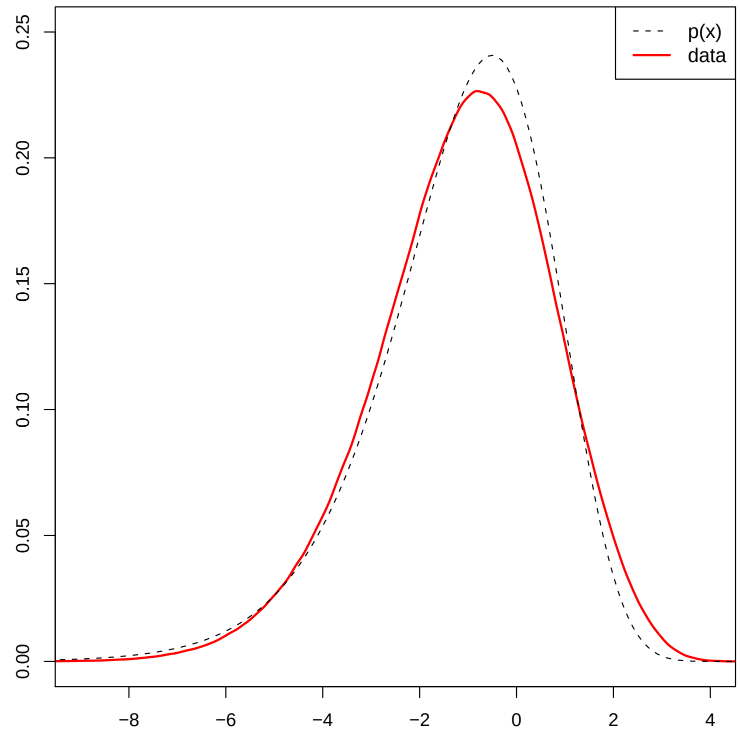


Figure 1: **Statistics of maxima for CUE polynomials (left: $N = 50, 10^6$samples) and $|\zeta(1/2+it)|$ (right: $N_T = 65,\ 10^5$ samples) compared to periodic $1/f$ noise prediction $p(x) = 2e^x K_0(2e^{x/2})$.**

Threshold of extreme values for self-similar multifractal fields:

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 $\&$ zeta-function along the critical line) 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

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

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.

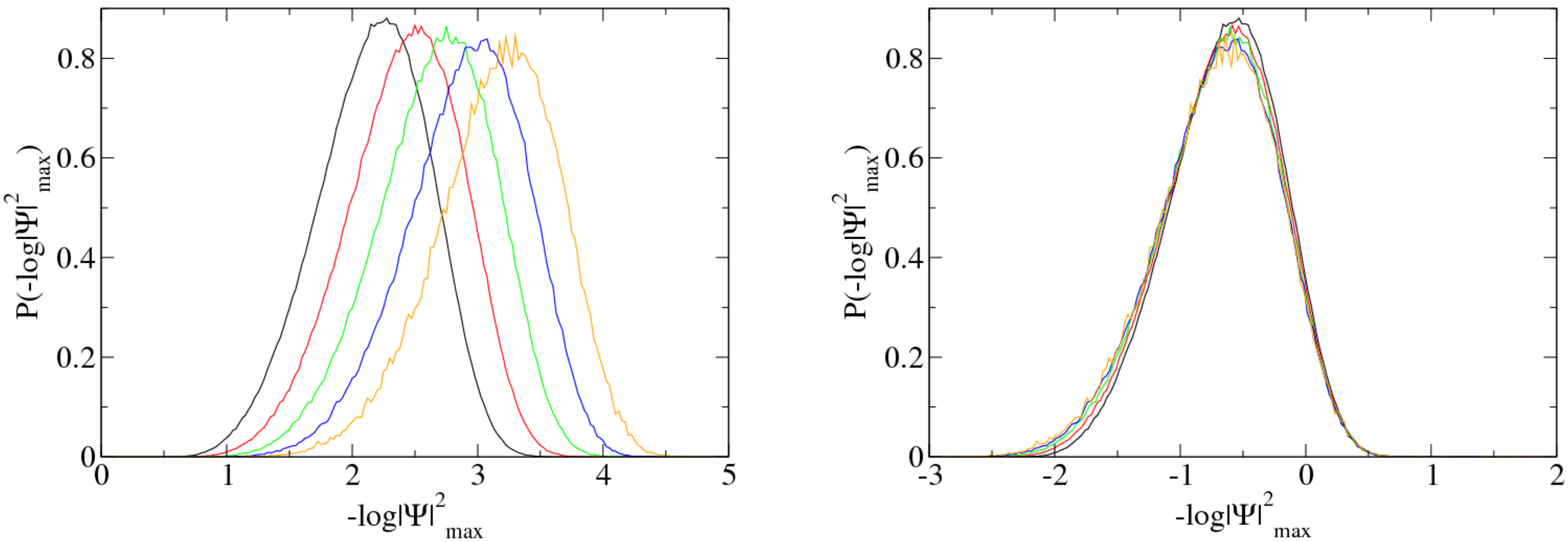


Figure 2: **Statistics of maxima for eigenvectors of RS model for sample sizes $M = 2^n$ with $n = 8, \ldots, 12$.**
left: raw data right: each curve is shifted by $\alpha_- \ln M + \frac{3}{2}\frac{1}{f'(\alpha_-)} \ln \ln M$; data by Olivier Giraud