Universal fluctuations and extreme-value statistics

2001 J. Phys. A: Math. Gen. 34 11193

(http://iopscience.iop.org/0305-4470/34/50/302)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **34** (2001) 11193–11200 PII: S0305-4470(01)27669-0

Universal fluctuations and extreme-value statistics

Kajsa Dahlstedt[1,2] **and Henrik Jeldtoft Jensen**[1,3]

[1] Department of Mathematics, Imperial College, 180 Queen's Gate, London SW7 2BZ, UK
[2] Department of Physics, Royal Institute of Technology (KTH), SE-100 44 Stockholm, Sweden

E-mail: h.jensen@ic.ac.uk

Received 6 August 2001, in final form 24 September 2001
Published 7 December 2001
Online at stacks.iop.org/JPhysA/34/11193

Abstract

We study the effect of long-range algebraic correlations on extreme-value statistics and demonstrate that correlations can produce a limit distribution which is indistinguishable from the ubiquitous Bramwell–Holdsworth–Pinton distribution. We also consider the square-width fluctuations of the avalanche signal. We find, as recently predicted by Antal *et al* for logarithmic correlated $1/f$ signals, that these fluctuations follow the Fisher–Tippett–Gumbel distribution from uncorrelated extreme-value statistics.

PACS numbers: 05.65.+b, 05.40.-a, 05.50.+q, 68.35.Rh

1. Introduction

Three years ago Bramwell, Holdsworth and Pinton (BHP) [1] published the remarkable discovery that the same functional form that describes the fluctuation spectrum of the energy injected in an experiment on turbulence also describes the fluctuations in the magnetization of the finite-size two-dimensional XY equilibrium model in the critical region below the Kosterlitz–Thouless transition temperature. Since three-dimensional turbulence and a two-dimensional magnetic equilibrium system appear to have very little in common, BHP made the reasonable suggestion that the origin of the identical functional form for the fluctuation spectra should be sought in the one thing the two systems appear to share, namely, *scale invariance*. This suggestion was supported by the subsequent finding that a long list of scale invariant (or nearly scale invariant) non-equilibrium as well as equilibrium systems exhibit the same BHP form for the fluctuation spectrum for certain quantities [2].

Nevertheless, not all critical systems fluctuate according to the BHP spectrum. This was made very explicit by Aji and Goldenfeld [3], who proffered the interesting suggestion that the reason the two-dimensional XY-model and the driven turbulence experiment exhibit the same fluctuation spectrum could be that the two-dimensional XY-model *is* the effective model for the turbulence experiment. Even if this is correct we still lack an explanation of why so

[3] Author to whom correspondence should be addressed.

many disparate systems [2] do exhibit BHP fluctuations. In the present paper we elaborate on the suggestion already made in [2] that the ubiquity of the BHP spectrum may be the result of extreme-value statistics acting together with correlations.

The BHP functional form is similar to one of the asymptotic forms for extreme-value statistics: the asymptote first discussed by Fisher and Tippett [5] and often referred to as the Gumbel distribution [6]. The BHP distribution is, however, not identical to the Fisher–Tippett–Gumbel (FTG) asymptote. Consider T independent and identically distributed stochastic variables. Under certain conditions (essentially exponential tail) the kth largest of the T variables will be distributed according to the FTG asymptote with k entering as a parameter. The BHP form can be thought of as corresponding to the somewhat uninterpretable case of $k = \pi/2$ [2].

As already alluded to in [2] the deviation between the BHP and the FTG form may be related to correlations. Our main aim in the present paper is to study this point in detail. We do that by simulating the so-called Sneppen depinning model [7]. The power spectrum of the Sneppen model behaves like $1/f^{\beta}$ with $\beta \simeq 0.5$ for low frequencies corresponding to a very slow algebraic decay of the autocorrelation function (see e.g. [8]). We use the Sneppen model in the next section to demonstrate that extreme-value statistics of T strongly correlated exponentially distributed variables may follow the BHP form. Remarkably, Antal *et al* [4] demonstrated analytically that, at least for a certain class of $1/f$ signals (periodic signals), the width-square fluctuations (w_2) of the signal follow the FTG distribution from extreme-value statistics, though it is not clear why this should be the case. Inspired by this finding we study in section 3 the width fluctuations of the avalanche signal in the Sneppen model. We find, contrary to the $1/f$ case studied by Antal *et al* that the probability density function (PDF) for w_2 in the Sneppen model is very well represented by the FTG distribution even for non-periodic signals. Section 4 contains a discussion and our conclusions.

2. Extreme-value statistics and the Sneppen model

To investigate the relationship between extreme-value statistics of correlated variables and the BHP probability density we consider now the simple (1 + 1)-dimensional depinning model introduced by Sneppen [7]. The model is imagined to represent a one-dimensional elastic interface moving transverse while acted upon by a set of random pinning forces (see figure 1(a)).

The model, which is discrete and very schematic, consists of L sites in the x-direction and infinitely many sites in the y-direction. Each square on this semi-infinite lattice is assigned a random number, the pinning force, uniformly distributed on the interval [0, 1[. The interface is represented by the set $\{(x, y)|y = h(x, t)\}$, where $h(x, t)$ denotes the height of the interface above the x-axis at time t at site x along the x-axis. The initial configuration is $h(x, 0) = 0$ for $\forall x$. In each time step the interface site, x_{s} with the *smallest* pinning force is located and the interface at this location is moved one step ahead, i.e. $h(x_{\mathrm{s}}, t) \mapsto h(x_{\mathrm{s}}, t) + 1$. This may cause the neighbouring slope $|h(x_{\mathrm{s}}, t) - h(x_{\mathrm{s}} - 1, t)|$ (or $|h(x_{\mathrm{s}} + 1, t) - h(x_{\mathrm{s}}, t)|$) to exceed unity, in which case the interface at site $x_{\mathrm{s}} - 1$ is moved one step ahead, i.e. $h(x_{\mathrm{s}} - 1, t) \mapsto h(x_{\mathrm{s}} - 1, t) + 1$ (and similar for the site $x_{\mathrm{s}} + 1$ if needed). The update of the nearest-neighbour sites of x_{s} may cause the slope on the next-nearest sites to exceed unity, in which case the interface is moved ahead on these sites. This procedure is repeated until all slopes satisfy $|h(x, t) - h(x - 1, t)| \leqslant 1$ once again. The sequence of operations needed to make all slopes smaller than or equal to unity after the update of site x_{s} is denoted *one time step*. The number of sites updated during one time step is called a micro-avalanche. The number of sites being updated during the time step t is called the size of the avalanche $s(t)$ and per definition the duration of each of these avalanches is the same, namely one time step.

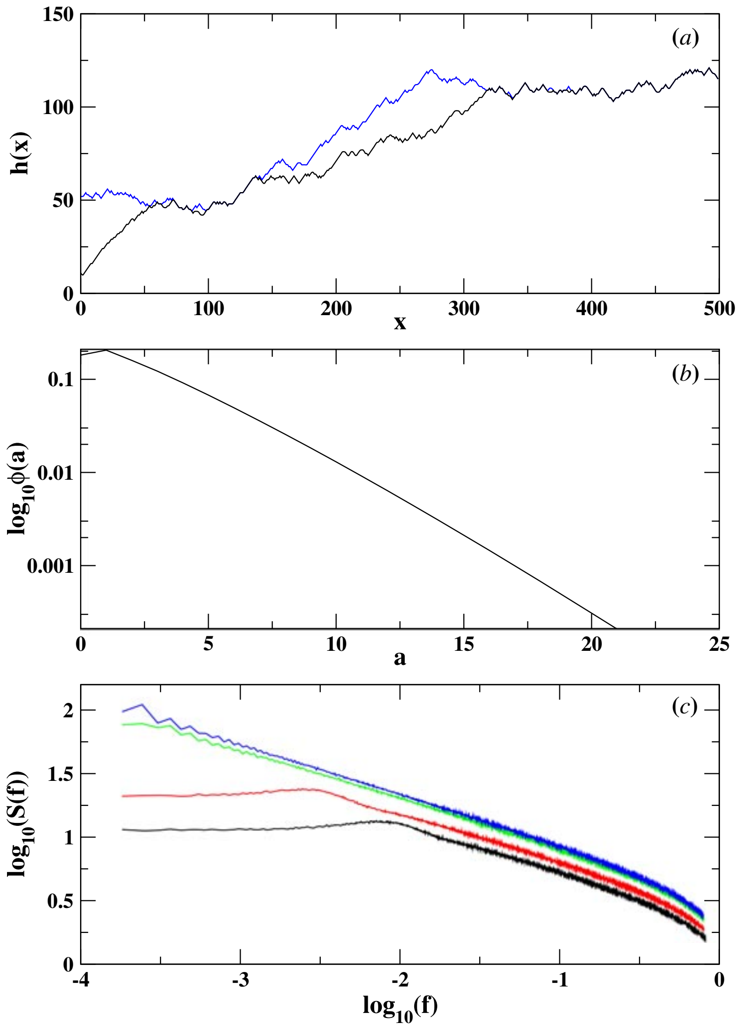

Figure 1. (*a*) Snapshot of the Sneppen interface $h(x)$ at two times 10 000 timesteps apart, system size $L = 500$. (*b*) The PDF of the micro-avalanches in the Sneppen model, system size $L = 500$. (*c*) The power spectrum of micro-avalanches for four different system sizes. From bottom to top $L = 50, 100, 500, 1000$.

(This figure is in colour only in the electronic version)

The PDF for the size of the micro-avalanches is close to an exponential as seen in figure 1(*b*) and first reported in [7]. This narrow distribution of the micro-avalanches suggests that extreme-value statistics on these variables should converge rapidly to the asymptote. This property together with the model simplicity makes the Sneppen model ideal for a numerical study of the effect of correlations on extreme-value statistics. The individual micro-avalanches are strongly correlated in time. This is clearly seen from the power spectrum $S(f)$ of the temporal signal $s(t)$, see figure 1(*c*). The low-frequency behaviour of the power spectrum is approximated $S(f) \propto f^{-\beta}$ with $\beta \simeq 0.6$, indicative of slow algebraic decay of the auto-correlation function $C(t)$ of the signal $s(t)$, namely like $C(t) \propto 1/t^{-\alpha}$ with $\alpha \simeq 0.6$ (see e.g. [8]).

Let us now imagine that the activity of the model is monitored by some device which has only a limited resolution. This might be modelled by assuming that the measured signal $\bar{S}_T(\tau)$, rather than consisting of the microscopic instantaneous activity $s(t)$, is given by the sum of activities within a time window of a certain size T:

$$\bar{S}_T(\tau) = \sum_{t=\tau}^{\tau+T} s(t). \tag{1}$$

This signal was among the set of quantities in [2] found to follow the BHP distribution for a certain range of window sizes T. That is, if T is too small the PDF for $\bar{S}_T(\tau)$ is close to an exponential distribution. The central limit theorem will, however, cause the PDF for $\bar{S}_T(\tau)$ to approach a Gaussian when T becomes large enough that correlations amongst the $s(t)$ entering the sum in equation (1) can be neglected. This deviation can be observed in figure 2(*a*). Presumably the PDF for $S(T)$ would remain close to the BHP form even for $T \to \infty$ if the correlation time of the signal were infinite as it would be expected to be in the limit of infinite system size, see figure 1(*c*).

Since, as seen in figure 2(*b*), the functional form of the BHP distribution is close to the first FTG asymptote for extreme-value statistics we proceed to investigate the extreme statistics of the correlated variables $s(t)$ generated by the Sneppen model. For the time windows considered for $\bar{S}_T(\tau)$ we define

$$M_T(\tau) = \max\{s(\tau), s(\tau+1), \ldots, s(\tau+T)\}. \tag{2}$$

In figure 2(*b*) we show the scaled PDF for $M_T(\tau)$. First we compare the distribution of the sum and the max in figure 2(*a*). We see that the distribution for $M_T(\tau)$ remains very close to the BHP distribution for all considered sampling sizes T, while the distribution for $\bar{S}_T(\tau)$ gradually deviates from the BHP as T is increased. In figure 2(*b*) we demonstrate two interesting points. First that for system size $L = 1000$ we are unable to make T so large that the distribution for $M_T(\tau)$ deviates from the BHP form. No difference between the results for $T = 5000$ and 10 000 can be detected. The size dependence is, however, detectable, as illustrated in the inset to figure 2(*b*). We show here the PDF for smaller system sizes $L = 500$ and 50. The spatial extent of the system $L = 50$ is now sufficiently small to destroy the long-range temporal correlations as is seen from the fact that the PDF for $L = 50$, $T = 500$ follows the FTG form for uncorrelated extreme-value statistics. We elaborate on the role of the correlations in the main frame of figure 2(*b*). This figure demonstrates that the deviation between the PDF for $M_T(\tau)$ and the FTG form is indeed caused by the correlations of the Sneppen model. This conclusion is reached in the following way. We generate T uncorrelated stochastic variables χ_i all drawn from the PDF for the individual micro-avalanches $s(t)$ (see figure 1(*b*)). The difference between

$$M_{uc} = \max\{\chi_1, \chi_2, \ldots, \chi_T\} \tag{3}$$

and $M_T(\tau)$ is solely the correlations amongst the primary variables from which $M_T(\tau)$ is generated. Hence, given the exponential form of the PDF for the distribution of the individual micro-avalanches (see figure 1(*b*)) we expect the PDF for the *uncorrelated* extreme M_{uc} to follow the FTG asymptote for large values of T. This is exactly what is found in figure 2(*b*).

3. The square width and extreme-value statistics

In the present section we analyse the statistics of the fluctuations in the width of the avalanche signal $s(t)$ generated by the Sneppen model and of the $1/f$ signal $F(\tau)$ in equation (6) below. Our inspiration is from the beautiful paper by Antal, Droz, Györgysi and Rácz (ADGR) [4].

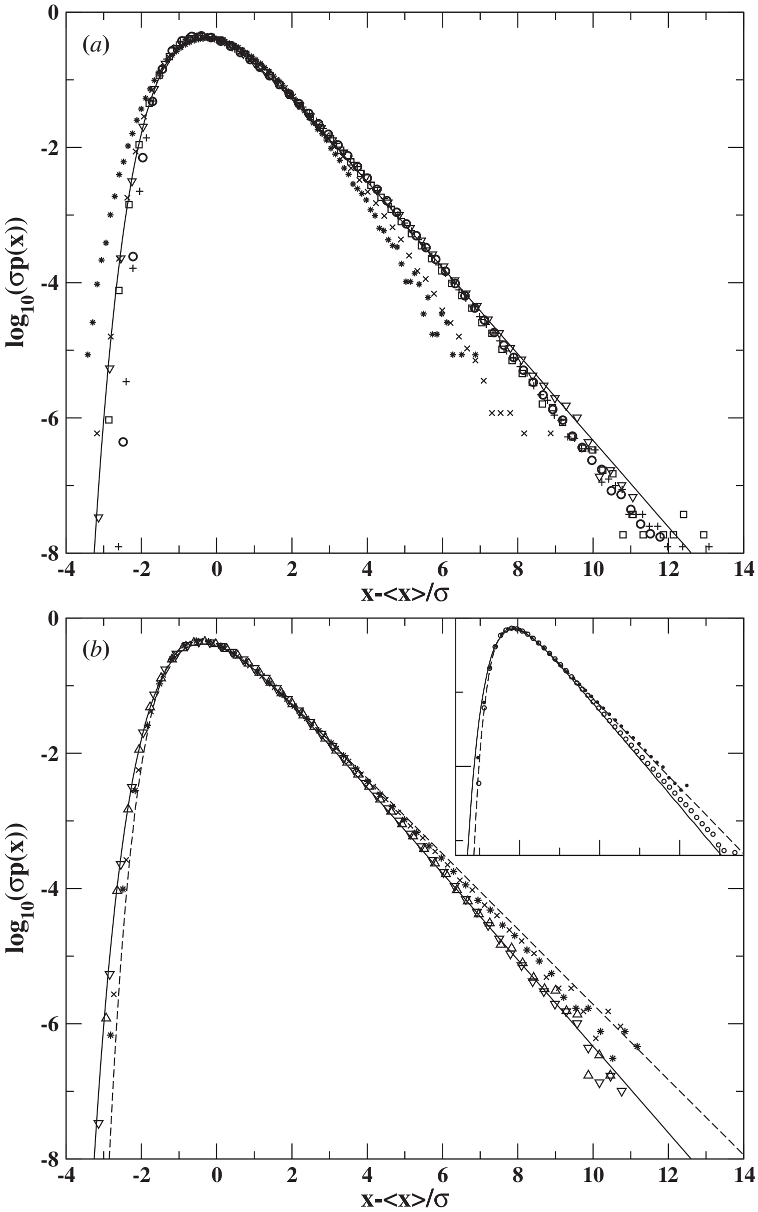

Figure 2. (*a*) The PDF of $\bar{S}_T(\tau)$ for $T = 50(+)$, $T = 500(x)$ and $T = 5000(*)$ and the PDF of $\bar{M}_T(\tau)$ for $T = 50$ (○), $T = 500$ (□) and $T = 5000$ (◇); the solid curve is the BHP PDF. (*b*) The PDF of $M_T(\tau)$ for $T = 5000$ (◇) and $T = 10\,000$ (△) and the PDF of the decorrelated maximum M_uc for $t = 5000$ (∗) and $T = 10\,000$ (+); the solid curve is the BHP form and the dashed curve is the FTG form. Inset, the PDF of $M_T(\tau)$ for system sizes $L = 50$, $T = 50$ (○) and $T = 500$ (●).

These authors demonstrated analytically that at least for a certain class of $1/f$ signals $h(t)$, the width-square

$$w_2(h) = \overline{[h(t) - \bar{h}]^2} \tag{4}$$

is distributed according to the FTG distribution. In equation (4) the over-bar denotes the following time average

$$\bar{h} = \frac{1}{T}\sum_{t=1}^{T} h(t). \tag{5}$$

The result by ADGR is striking since it is unclear how or whether extreme-value statistics is involved in some effective way in determining the distribution of w_2.

As a contribution to the investigation of the generality of the discovery by ADGR we studied w_2 for the signal $s(t)$. This signal is not $1/f$ but rather $1/\sqrt{f}$, and differs in this way from the class studied in [4]. As see in figure 3(a) w_2 for $s(t)$ does follow the FTG functional form except for very small sample sizes. This indicates that the ADGR result is more general than their calculation allows one to conclude. ADGR predicted analytically that this should be the case for periodic $1/f$ signals and they report that deviations are observed in simulations of non-periodic $1/f$ signals.

To address the question about deviations from the FTG form we show in figure 3(b) the fluctuations in w_2 for a $1/f$ signal of the type considered by Antal *et al* [4]. We generated the signal as the 'half integral' of white noise [9]; i.e., the signal is obtained as a convolution between white noise and a propagator $G(t)$ which decays as an inverse square root:

$$F(\tau) = \lim_{T_* \to \infty} \sum_{t=\tau-T_*}^{\tau} G(\tau - t)\chi(t) \tag{6}$$

where $G(x) = 1/\sqrt{x}$. We assume $\chi(t)$ to be uncorrelated and uniformly distributed on the interval $[-1, 1]$ and in our simulations we truncate the sum in equation (6) at $T_* = 2^{20}$.

One can calculate the PDF of the Fourier coefficients of the signal of length T in equation (6) by performing a discrete Fourier analysis (note this amounts to assuming a signal of periodicity T)

$$\hat{F}(\omega) = \sum_{\tau=0}^{T} F(\tau)\mathrm{e}^{-\mathrm{i}\omega\tau}. \tag{7}$$

Averaging over the white noise $\chi(t)$ gives the following PDF for the Fourier amplitudes:

$$p(|\hat{F}(\omega)|^2 = R^2) = \frac{3\omega}{\pi T}\exp\left(-\frac{3}{\pi T}\omega R^2\right) \tag{8}$$

which demonstrates explicitly that the signal in equation (6) is of the type considered by ADGR [4].

We notice that the simulated distribution for w_2 in figure 3(b) for all sample sizes T deviates from the FTG form predicted by Antal *et al* [4]. This was indeed noticed by ADGR (their figure 1) and ascribed to the difference between the assumed periodic boundary condition needed to perform analytical calculation of the PDF of w_2 and the non-periodic signal simulated. This problem appears not to arise in case of the Sneppen model, where figure 3(a) shows that the PDF for w_2 follows the FTG form as soon as T is of order 500 or larger. We recall that the difference between the signal in the Sneppen model and the signal in equation (6) is that the Sneppen signal is algebraically correlated with an autocorrelation function that decays like one over the square root of time whereas the $1/f$ signal in equation (6) decays even more slowly, namely logarithmically. It is likely that this difference in the range of the correlations is the cause of the difference between the two signals sensitivity to boundary condition.

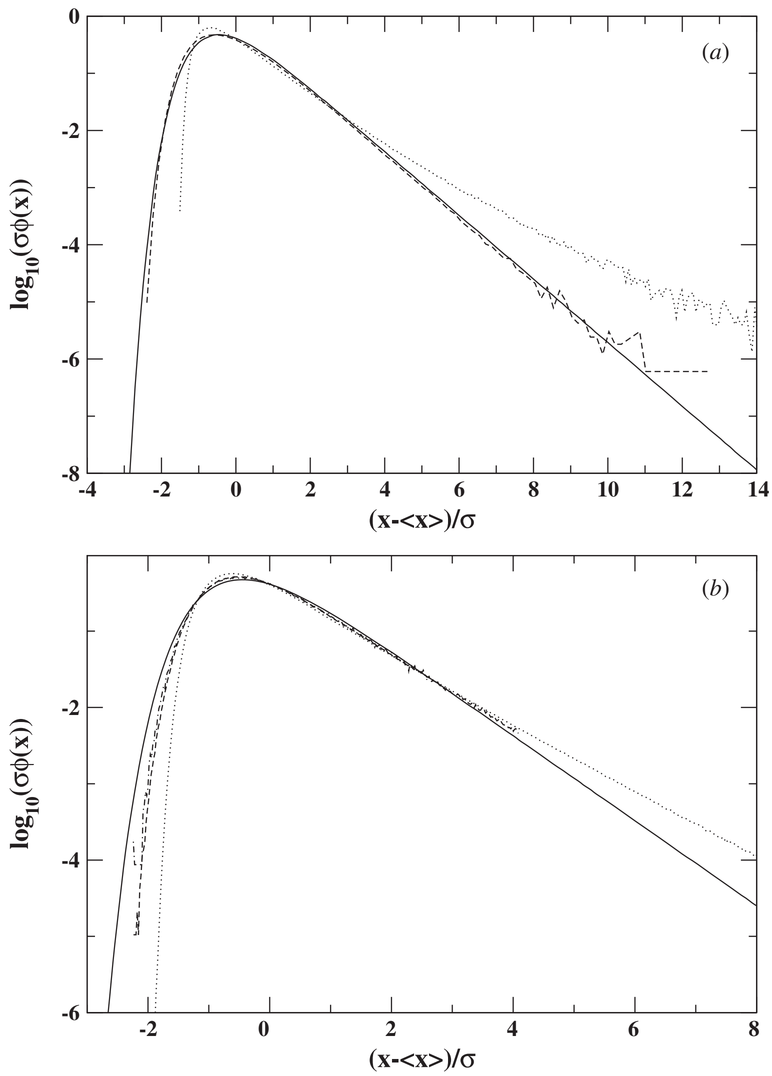

Figure 3. (*a*) The PDF of the width-square for the Sneppen model, sample sizes $T = 50$ (dotted curve) and $T = 500$ (dashed curve); the solid curve is the FTG form. (*b*) shows the PDF of the width-square for the $1/f$-noise, sample sizes $T = 64$ (dotted curve), $T = 4096$ (dashed curve) and $T = 32\,768$ (dashed–dotted curve); the solid curve is the FTG form.

4. Discussion and conclusion

We have studied the effects of correlations on extreme-value statistics. Our aim has been to relate the Bramwell *et al* distribution [1, 2] to correlated extreme statistics. We have achieved this successfully for one specific system: the Sneppen depinning model.

We found that the extreme-value statistics of the algebraically correlated avalanche signal is described by the BHP distribution. Furthermore we studied the width-square of the same signal and found this to be distributed according to the FTG distribution for *uncorrelated* extreme value statistics, though no extremes were explicitly involved for the width-square signal. This finding generalizes a recent result by ADGR [4]. The BHP distribution and the FTG distribution are of similar functional form though they differ significantly for large fluctuations away from the mean.

Extreme-value statistics for $1/f$ correlated signals were studied using renormalization group techniques by Carpenter and Le Doussal [10]. They find that the correlations make the functional form change from the exponential decay of the uncorrelated FTG to the form $y\exp(-y)$ for fluctuations $y = x - \langle x\rangle > 0$ above the average (in the case of maximum statistics). Note this form is also different from the BHP distribution, which decays exponentially like the FTG. Signals with hierarchical correlations were considered in a recent preprint by Dean and Majumdar [11]. They find that correlations typically alter the super-exponential $\exp(-\exp(y))$ found for $y < 0$ in the uncorrelated FTG case, as well as in the BHP distribution.

These results makes it perfectly clear that not all correlations make extreme-value statistics follow the BHP distribution. We have, however demonstrated that in the case of the avalanche signal of the Sneppen model, which has an algebraically correlated exponentially distributed signal, the BHP distribution does describe with great accuracy the extreme statistics generated from the avalanche signal. More research is needed in order to determine the generality of this single example.

Acknowledgments

It is a pleasure to acknowledge stimulating discussion with our friends and colleagues at Imperial, especially K Christensen, D Hamon, M Hall, S Lise, M Nicodemi and G Pruessner. We are also indebted to G Pruessner for pointing out [4]. HJJ greatly appreciates the exchange of general ideas concerning universal fluctuations with J Lopez, P Holdsworth and S Bramwell. KD is supported by KTH.

References

[1] Bramwell S T, Holdsworth P C W and Pinton J-F 1998 *Nature* **396** 552
[2] Bramwell S T *et al* 2000 *Phys. Rev. Lett.* **84** 3744
[3] Aji V and Goldenfeld N 2001 *Phys. Rev. Lett.* **86** 1007
[4] Antal T, Droz M, Györgyi G and Rácz Z 2001 *Preprint* arXiv cond-mat/0105599
[5] Fisher R A and Tippett L H C 1928 *Camb. Phil. Soc.* **28** 180
[6] Gumbel E J 1958 *Statistics of Extremes* (New York: Columbia University Press)
Galambos J 1987 *The Asymtotic Theory of Extreme Order Statistics* (Malabar, FL: Krieger)
[7] Sneppen K 1992 *Phys. Rev. Lett.* **69** 3539
[8] Jensen H J 1998 *Self-Organized Criticality* (Cambridge: Cambridge Univeristy Press)
[9] Press W H 1978 *Comments Astrophys. Space Phys.* **7** 103
[10] Carpentier D and Le Doussal P 2001 *Phys. Rev.* E **63** 026110
[11] Dean D S and Majumdar S N 2001 *Preprint* arXiv cond-mat/0104028