Open Research Online

The Open University's repository of research publications
and other research outputs

On the continuing relevance of Mandelbrot's non-ergodic fractional renewal models of 1963 to 1967

Journal Item

How to cite:

Watkins, Nicholas W. (2017). On the continuing relevance of Mandelbrot's non-ergodic fractional renewal models of 1963 to 1967. The European Physical Journal B, 90(12), article no. 241.

For guidance on citations see FAQs.

oro.open.ac.uk

**THE EUROPEAN
PHYSICAL JOURNAL B**

Regular Article

On the continuing relevance of Mandelbrot's non-ergodic fractional renewal models of 1963 to 1967[*]

Nicholas W. Watkins[1,2,3,a]

[1] Centre for the Analysis of Time Series, London School of Economics and Political Science, London, UK
[2] Centre for Fusion, Space and Astrophysics, University of Warwick, Coventry, UK
[3] Faculty of Science, Technology, Engineering and Mathematics, Open University, Milton Keynes, UK

Abstract. The problem of "$1/f$" noise has been with us for about a century. Because it is so often framed in Fourier spectral language, the most famous solutions have tended to be the stationary long range dependent (LRD) models such as Mandelbrot's fractional Gaussian noise. In view of the increasing importance to physics of non-ergodic fractional renewal models, and their links to the CTRW, I present preliminary results of my research into the history of Mandelbrot's very little known work in that area from 1963 to 1967. I speculate about how the lack of awareness of this work in the physics and statistics communities may have affected the development of complexity science, and I discuss the differences between the Hurst effect, "$1/f$" noise and LRD, concepts which are often treated as equivalent.

1 Introduction: ergodic and non-ergodic solutions to the problem of $1/f$ noise

The pioneering, and highly influential, work of Montroll and Weiss [1,2] on the continuous time random walk (CTRW) from 1965 onwards is commemorated in this special issue of EPJB. My contribution is about historical and conceptual aspects of one of the many problems to which the CTRW and its close relatives have been applied-the "$1/f$" paradox. This is "the infrared catastrophe" in sample periodograms- studied both as long range dependence (LRD) and $1/f$ noise-which has long been seen as a theoretical puzzle by time series analysts, statisticians, and physicists. There is also a closely related theoretical question of how many different classes of model can share the common property of the $1/f$ spectral shape, e.g. [3,4]).

Since the late 1960s, a highly visible (and still controversial) line of investigation has centred on a model, Mandelbrot's fractional Gaussian noise, which is now widely understood to differ fundamentally from the CTRW, as I shall recap below. However, relatively recently, fractional renewal models that are in the CTRW family have attracted physical interest in physics as examplars of the weak ergodicity breaking [5] seen in, for example, blinking quantum dots [6–9]. This new focus has led to a recent proposal [7] that they can solve the "$1/f$ paradox".

Fractional renewal processes are closely related to the CTRW's with waiting time distributions that decay so slowly as to have infinite variance (and sometimes even infinite mean) introduced by Scher, Montroll, Lax, Kenkre and Shlesinger in the early 1970s (see the historical account of [2]), and characterised by Mandelbrot as "fractal time" processes [10]. My own paper's intent is twofold, to expand on my historical research into Mandelbrot's own still very little-known work on such fractional renewal processes [11], and the "$1/f$ paradox", *in the mid 1960s*, and to use these findings to better classify and clarify the current approaches to $1/f$ noise and LRD.

The physicist's "problem of $1/f$ noise" has been with us since the pioneering work of Schottky and Johnson in the early 20th century on fluctuating currents in vacuum tubes [11–13]. The paradox is usually framed in power spectral terms, i.e. "how can the Fourier power spectral density $S'(f)$ of a stationary process take the form $S'(f) \sim 1/f$ and thus be singular at the lowest frequency", or equivalently "how can the autocorrelation function "blow up" at large lags and thus not be summable?". This framing of the problem in spectral terms has, as we will see below, subtly conditioned the type of solutions sought.

In the 1950s an analogous time domain effect (the Hurst phenomenon) was seen in the statistical growth with the observation time scale τ of Hurst's rescaled range $R/S(\tau)$ calculated on the minima of levels of the Nile river [12]. Rather than the expected dependence $\sim \tau^{1/2}$ many time series including the Nile data showed a dependence $\sim \tau^J$ where J, the "Hurst exponent" was typically greater than 0.5. This rapidly presented a conceptual problem because Feller soon proved that an iid sequence

must asymptotically have $J = 1/2$. Although in fact many observed instances of the Hurst effect may indeed arise from pre-asymptotic effects, gross non-stationarity, or one of several other possibilities, the desire for a stationary solution to the problem with a satisfying level of generality remained.

It was thus an important step forward when, in 1965–67, Mandelbrot presented a stationary process, fractional Gaussian noise (fGn), which could be used as a time series model which exhibits both the Hurst effect and $1/f$ noise. The LRD fGn process is the formal differential of H-self similar fractional Brownian motion (fBm), and was subsequently developed by him with Van Ness and Wallis, particularly in a hydrological context [12,14]. fGn is a stationary ergodic process, for which a power spectral density is a natural, well-defined concept, the paradox here residing in the singular behaviour of $S'(f)$ at zero frequency (in the $H > 1/2$ case).

Skepticism about fGn-and another related LRD process, Granger and Hosking's autoregressive fractionally integrated moving average (ARFIMA, [12]) as a universal explanation for observed Hurst effects remained (and remains) considerable, though, because of the highly non-Markovian properties of these processes. Many authors, particularly in statistics and econometrics, have found models based on change points to be better motivated, not least because many datasets are unambiguously known to have change points that need to be handled. If such change points occur at random intervals, and the time series occupies discrete levels, such a model is already of CTRW form (see for example, the right hand panel of fig. 1 of [15]).

Meanwhile, in the physics literature, in the last two decades, it has increasingly been realised [16] that a CTRW-class model, the nonstationary but bounded fractional renewal process (FRP), gives rise to $1/f$ spectra. These power spectra have been recognised to be nonergodic, possessing a dependence on the observation time which has been proposed as a resolution of the "$1/f$" paradox. To facilitate comparison, in Section 2, I will briefly recap the key properties of the FRP and fGn and the differences between them.

In view of the increased interest in physics of nonergodic FRPs, this paper aims first to offer a brief review of Mandelbrot's second main contribution to the "$1/f$" problem, in parallel with the above work on fBms, his still very little known work on fractional renewal processes in 1963–67. To my great surprise I have found that the still topical dichotomy between ergodic and non-ergodic origins for $1/f$ periodograms was not only recognised but published on by Mandelbrot about 50 years ago, and that he proposed *both* mechanisms as solutions for "$1/f$" signals in different contexts. He developed his FRPs in parallel with his seminal and much more visible work on ergodic, stationary fGn which is thus today very much better known to physicists, geoscientists and many other time series analysts (e.g. [17,18]).

The preprint [11] in which I first reported this work has drawn some attention, and in consequence Mandelbrot's pioneering insights are now being more widely acknowledged (e.g. [19]). This paper thus expands on it to offer more technical detail in order to facilitate the integration of his work into current investigations. I will describe 5 key papers [20–24], and the bridging essays he wrote when he revisited them late in life for (edited) republication in his Selecta volumes [14,25]. In these papers he:

- proposed the use of time series models with heavy-tailed (Pareto) waiting time distributions of the form $Pr(U \geq u) = u^{-\alpha}$, with α in the stable range, to capture hierarchical clustering effects;
- recognised that these were self-similar stochastic point processes;
- noted that such models were not stationary processes with a conventional Wiener–Khinchine interpretation for their power spectrum;
- introduced the concept of a conditionally stationary random variable;
- used conditional stationarity to define a conditional covariance C_W;
- using C_W, showed that their conditional spectral density had an explicit dependence on observation time: $S'(f) \sim T^{1-\alpha} f^{\alpha-2}$;
- explained in detail why this was a resolution of the "infrared paradox" in cases where no intrinsic low end cutoff is seen;
- discussed the nonergodicity of self-similar stochastic point processes.

Mandelbrot's work at IBM was not the only contemporary paper on point processes with heavy tailed waiting times, at least one other example being the work of Mertz [26,27] at RAND on modelling telephone errors, so this article does not attempt to assign priority. Additionally, as noted by Shlesinger [2], one should realise that the idea of introducing waiting times between the transitions between states of a Markov processes had been around for about 10 years when Mandelbrot used it in 1963, having been introduced in work by Paul Lévy in 1954.

The second main purpose of this contribution, in Section 4, aided by this historical perspective, is to clarify the subtle differences between 3 phenomena: the empirical Hurst effect, the appearance of $1/f$ noise in periodograms, and the concept of LRD as embodied by the stationary ergodic fGn model, and to set out their hierarchy with respect to each other. This relatively short paper does not deal directly with multiplicative models (e.g. [3,25]), although they remain a very important alternative source of $1/f$ spectra, particularly those which arise from turbulent cascades. I also do not consider $1/f$-type periodograms arising from nonstationary self-similar walks such as fBm. Such walks are intrinsically unbounded and so the periodogram must already a priori be different from a stationary power spectrum.

I conclude by arguing that the relative neglect of [20–23] at the time of their publication must have had long-term effects, particularly on the nascent field of complexity science as it developed in the 70s and 80s.

2 Fractional Gaussian noise and fractional renewal processes compared

2.1 Fractional Gaussian noise

fGn [18] is effectively a derivative of fractional Brownian motion $Y_{H,2}(t)$:

$$Y_{H,2}(t) = \frac{1}{C_{H,2}} \int_R dL_2(s) \, K_{H,2}(t-s), \qquad (1)$$

which in turn extends the Wiener process to include a self-similar memory kernel $K_{H,2}(t-s)$, such that

$$K_{H,2}(t-s) = [(t-s)_+^{H-1/2} - (-s)_+^{H-1/2}], \quad (2)$$

thus giving a decaying, non-zero weight to all of the values in the time integral over the Gaussian white noise dL_2. In consequence fGn shows long range dependence by construction, and it became the original paradigmatic model for LRD. The attention paid to its $1/f$ spectrum and long-tailed autocorrelation function as diagnostics of LRD has often led to it being forgotten that stationarity is the other essential ingredient for LRD in this sense. Intuitively one can see that without stationarity there can be no LRD because there is no infinitely long past history over which sample values of the process can be dependent. Models like fGn, and also fractionally integrated noise (FIN) and the ARFIMA process, which have been widely studied in the statistics community (e.g. [17,18]) exhibit LRD *by construction*, i.e. stationarity is assumed at the outset in defining them. More subtly, this notion of LRD also appears to require the stronger property of ergodicity, in order that their conventional interpretations can be ascribed to the power spectrum and autocorrelation function.

While undeniably important to time series analysis and the development of complexity science, it is obvious from the restriction to stationary processes, that the LRD concept when embodied by fGn might be insufficient to describe the full range of $1/f$ or Hurst behaviour that observations might present us with. Full awareness of this fundamental limitation seems to have been slow, however. I think this has been due to three widespread, deeply-ingrained, but unfortunately erroneous "folk beliefs" (to which I have not been immune): (i) that an observed Fourier periodogram can *always* be taken to estimate a meaningful power spectrum, (ii) that the Fourier transform of an empirically obtained periodogram is *always* a meaningful estimator of an autocorrelation function, and (iii) that the observation of a $1/f$ Fourier periodogram in a time series *must* imply the kind of long range dependence that is embodied in the ergodic fractional Gaussian noise model. The first two beliefs are of course routinely cautioned against in any good course or book on time series analysis, including classics like Bendat's [28]. The third belief remains highly topical, however, because it is only relatively recently being appreciated in the theoretical physics literature just how distinct two of the paradigmatic classes of $1/f$ noise model are, and how these

differences relate not only to LRD but also to the fundamental physical question of weak ergodicity breaking (e.g. [5,7,16]).

2.2 Fractional renewal models: the AFRP and CTRW

The second paradigm for $1/f$ noise mentioned above is the fractional renewal class, which is a descendent of the classic random telegraph model [28]. Its structure is stationary and Markovian, but it has switching times at power law distributed intervals, which in consequence may lack a variance or even a mean. A particularly well studied variant is the alternating fractal renewal process (AFRP, e.g. [29,30]), which is closely connected to the renewal reward process in mathematics. When studied in the telecommunications or other engineering contexts, however, the AFRP has often had a cutoff applied at large values of time to its switching time distribution to allow analytical tractability. The use of an upper cutoff masks some of its most physically interesting behaviour, because when the cutoffs are not used the periodogram, the empirical acf, and *observed* waiting time distributions, all *grow* with the length of time over which they are measured, rendering the process both non-ergodic and non-stationary in a fundamental sense (Mandelbrot preferred his own term "conditionally stationary"). In particular, Mandelbrot stressed that the process no longer obeys the necessary conditions on the Wiener–Khinchine theorem for its empirical periodogram to be straightforwardly interpreted as an estimate of the power spectrum.

The existence of this alternative, nonstationary, nonergodic fractional renewal model makes it clear that there is a difference between the observation of an empirical $1/f$ noise alone, and the presence of the type of LRD that is embodied in the stationary ergodic fGn model. We will develop this point further in Section 4, but will first go back to the 1960s to survey the less well known of Mandelbrot's twin tracks to $1/f$.

3 Mandelbrot's fractional renewal route to $1/f$

Mandelbrot was not only aware of the distinction between fGn and fractional renewal models [14,25], but also published a nonstationary model of the AFRP type in 1965 [21,22] and had explicitly discussed the time dependence of its power spectrum as a symptom of non-ergodicity by 1967 [23].

There are 5 key papers in the development of Mandelbrot's consideration of fractional renewal models and ageing:

3.1 A fractional renewal model: Berger and Mandelbrot [20]

The first, cowritten with his IBM colleague, the physicist Berger [20], appeared in the IBM Journal of Research and Development. Concerned with errors in telephone circuits, its main focus was on the power law distribution of times

u between errors, of the form:

$$Pr(U \geq u) = u^{-\theta}, \tag{3}$$

where θ refers to an exponent more usually denoted as α when it falls in the stable range of 0–2.

The errors were themselves assumed to have discrete states. Switching models, particularly the state dependent ones, were already being considered in order to model the clustering of errors, and the main point of the paper was to demonstrate how apparent clustering could originate in the hierarchy created by a fractal waiting time distribution, without the need to assume state dependence. Berger and Mandelbrot acknowledged that Pierre Mertz of RAND had already studied a power law switching model [26], but Mandelbrot's early exposure to the extended central limit theorem via the lectures of Paul Lévy, and the fact that he was contemporaneously studying heavy tailed models in economics and neuroscience, among other applications, seem to have enabled him to see a broader significance for the FRP class than his peers did.

3.2 Conditional stationarity of self-similar stochastic point processes: Mandelbrot [21]

The second, a sole author paper [21], was in the IEEE Transactions on Communication Technology, and essentially also used the model published with Berger, although it employed a technique reminiscent of the renormalisation group to study how the waiting time distribution would change with the introduction of a coarse-graining time scale ϵ. It assumed that $P(u) = 1$ for $U < \epsilon$ while the above form is retained for $u > \epsilon$, resulting in a distribution of times between observed "boxes" of:

$$Pr(U \geq u) = u^{1-\theta} - (u-1)^{1-\theta}. \tag{4}$$

The abstract notes that it describes:

> ... a model of certain random perturbations that appear to come in clusters, or bursts. This will be achieved by introducing the concept of "self-similar stochastic point process in continuous time." The resulting mechanism presents fascinating peculiarities from the mathematical viewpoint. In order to make them more palatable as well as to help in the search for further developments, the basic concept of "conditional stationarity" will be discussed in greater detail than would be strictly necessary from the viewpoint of the immediate engineering problem of errors of transmission.

The idea of "conditional stationarity" was introduced in this paper to handle the peculiarities of waiting times in the case $0 < \theta < 1$, when not only the variance but the mean of the inter-event intervals are infinite, with expectation value $E(U) = \infty$. Mandelbrot drew a contrast with the conventional definition of stochastic point processes in terms of an indicator function $V(t'_h, t''_h)$ which is 1 if the time interval (t'_h, t''_h) contains a switching event. Such a process is stationary if all the joint probabilities of form

$Pr\{V(t'_h, t''_h)\} = v_h$ are unchanged when $\delta \neq 0$ is added to t'_h and t''_h. In the infinite mean case, the probability of t'_1, t''_1 containing a switching event is $|t''_1 - t'_1|/E(U)$ and thus zero. He thus argued for the use of conditional indicator functions, the subset where V was already known to take the value of 1. Conditional stationarity would thus only refer to time translation of this subset.

3.3 Conditional spectra in fractional renewal processes as a solution to the infrared catastrophe: Mandelbrot [22]

It is clear that by 1965 Mandelbrot had come to appreciate that the application of the Fourier periodogram to conditional stationary processes would give counter-intuitive results if naively treated as an estimate of a Wiener–Khinchine spectrum, noting in [21] that:

> The now classical technique of spectral analysis is inapplicable to the processes examined in this paper but it is sometimes unavoidable that otherwise excellent spectral estimates be applied in this context. Another publication of the author [i.e. Ref 18 in [21]] is devoted to an examination of the expected outcomes of such operations. This will lead to fresh concepts that appear most promising indeed in the context of a statistical study of turbulence, excess noise [i.e. "$1/f$"], and other phenomena when interesting events are intermittent and bunched together (see also [Ref 19 in [21]]).

The "[other] publication ... Ref 18", became the third key paper [22] in the sequence. It resulted from Mandelbrot's talk at the IEEE Communications Convention in Boulder 1965, and is now available in the *post hoc* edited form that all papers take in his Selecta volumes [14,25]. The editing has attracted controversy [31], but with the proviso that the Selecta version may not fully reflect the original content, it nonetheless seems clear that in [22] Mandelbrot discussed a three state, explicitly nonstationary, renewal model. This stochastic process was intended as a "cartoon" to model intermittent turbulence, in which "off" periods (of no activity) were interrupted by jumps to a negative (or positive) "on" (active) state. His key finding was that the traditional Wiener–Khinchine spectral diagnostics would return a $1/f$ periodogram and thus a spectral "infrared catastrophe" when viewed with traditional methods, but, building on the notion of conditional stationarity proposed in [21], a conditional power spectrum $S(f, T)$ could be defined that was decomposable into a stationary part in which no catastrophe was seen, and one that depended on the length of the time series T, multiplying a slowly varying function $L(f)$.

3.4 Explicit calculation of the spectrum of the FRP, and ergodicity breaking: Mandelbrot [23]

The "Reference 19" anticipated in [21] seems likely, from the above description of its subject matter, to have been intended to be a paper in the applied mathematics or physics literature. I have not yet been able to determine
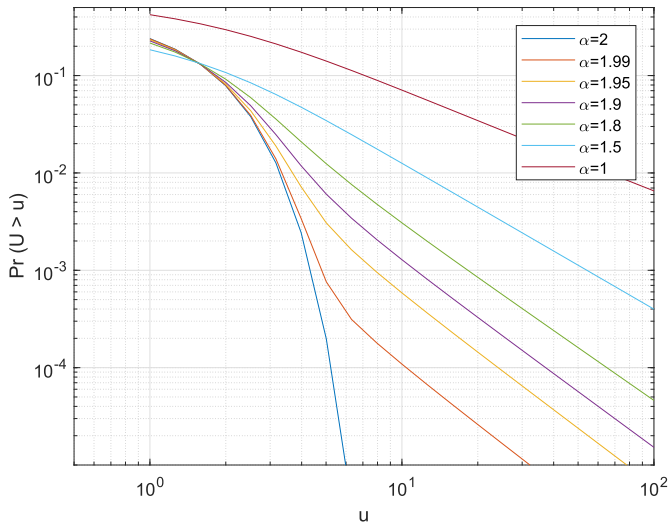
Fig. 1. "The positive portions of Paul Lévy's symmetric stable distributions" in the finite mean $1 < \alpha \le 2$ case, after figure 3 of [23], with the more usual notation α for the stability exponent, rather than Mandelbrot's θ. Note that the paper also considered the infinite mean $0 < \alpha < 1$ range of the stable distributions, not plotted here.

that original version's fate but its role was effectively taken over by the fourth key paper [23]. This instead appeared in an electrical engineering and communications journal and contained a very detailed examination of the fractional renewal process and its implications for Fourier spectra.

In it Mandelbrot generalised his earlier 3-state FRP model to one with an arbitrary number of discrete levels. Although he devoted considerable space to considering slowly varying waiting times, for concreteness he specialised to waiting times drawn from a stable distribution, specifying it by its characteristic function. The resulting probability density decayed for large t as a power law $p(t) \sim t^{-(1+\theta)}$, when θ was chosen in the stable range $0 < \theta < 2$, while $\theta = 2$ was the Gaussian case. The paper distinguished between the region $1 < \theta < 2$, with finite mean waiting times, and the region with infinite mean, $0 < \theta < 1$, which has also been the focus of much recent research (e.g. [7]). He illustrated in the former case in one of the paper's three figures, replotted as my Figure 1. In the former (stationary) case he argued that some but not all of Wiener–Khinchine theory was still applicable, but contended that for the latter a "non-Wieneran" spectral theory would be needed, remarking that:

> [...] the existence of $f^{\theta-2}$ noises challenges the mathematician to reinterpret spectral measurements otherwise than in "Wiener–Khinchin" terms ... [because] operations meant to measure the Wiener–Khinchine spectrum may unvoluntarily measure something else, to be referred to as the "conditional spectrum" of a "conditionally covariance stationary" random function.

He first discussed the $\theta = 0$ case, which corresponds to a situation where a time interval will contain at most one switching event, illustrated in his figure 1 (my Fig. 2). He
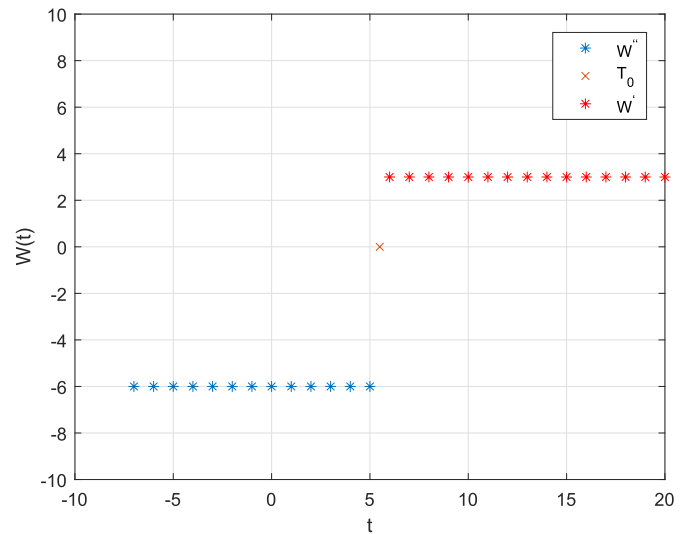

Fig. 2. After figure 1 of [23], the simplest case of a switching model considered by Mandelbrot in which only one transition occurs in the observation interval.

referred to this as the "DC" case because it represented a switch at time T_0 between the two constant values W'' and W'.

The simplicity of this case enabled him to explicitly calculate the conditional convariance function, and thus point out that the conditional spectral density $S'(f,T)$ obeyed

$$S'(f,T) = \frac{d}{df}S(f,T) \sim f^{-2}T^{-1}, \tag{5}$$

and had an explicit dependence on the length of time T over which the series was observed. He also drew attention to the difference between time averages and ensemble averages even in this simplest case.

He then generalised to the $0 < \theta < 1$ case, and illustrated this in his figure 2 (my Fig. 3), finding that:

$$S(f,T) \sim f^{\theta-1}L(f)Q(T), \tag{6}$$

where $Q(T)T^{1-\theta}$ was slowly varying, so that the conditional spectral density $S'(f,T)$ now obeyed

$$S'(f,T) \sim f^{\theta-2}T^{\theta-1}L(f). \tag{7}$$

In the same paper, anticipating [7] by nearly 50 years, he argued that this effect resolved the "$1/f$" paradox. In one of the Selecta essays [25] he described the apparent infrared catastrophe in the power spectral density in the FRP as a "mirage", rather than representing a true singularity in power at the lowest frequencies as is seen in fGn.

Direct experimental evidence for the predicted time dependent prefactor in the power spectrum has only recently been available from experiments on blinking quantum dots [8]. Another pioneering measurement [32]
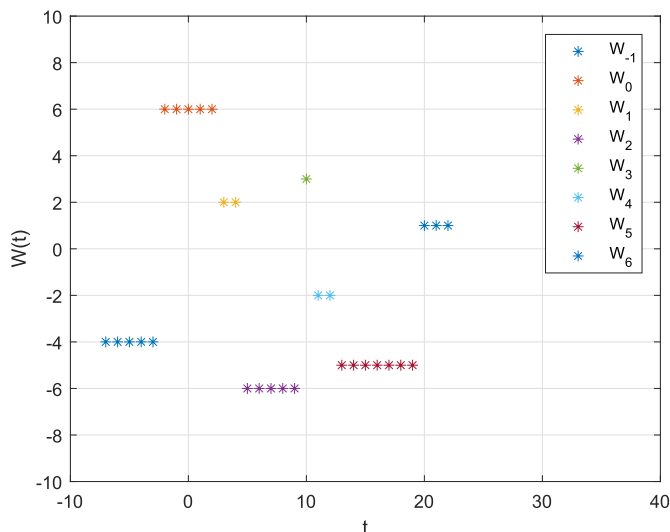
Fig. 3. After figure 2 of [23], the multilevel switching process studied by Mandelbrot where the switching times are drawn from a probability distribution. The paper considered both the finite-mean stable distributions as shown in Figure 1 and their infinite-mean counterpart.

has been that of interface fluctuations in the KPZ universality class (both experiment and simulations). Interestingly, in view of Mandelbrot's original intended application of the fractional renewal process as a caricature of intermittent turbulence, such an approach to $1/f$ has also recently been used in this context by Herault et al. [33]. In their case the mean sojourn time is finite (c.f. Fig. 1) but the variance diverges.

3.5 A mathematical theory for conditional spectra and their fluctuations: Mandelbrot [24]

In [23] Mandelbrot emphasised the clear contrast between his conditionally stationary, non-Gaussian fractional renewal $1/f$ model and his stationary Gaussian fGn model (the 1968 paper about which, with Van Ness, was then in press at SIAM Review):

> Section VI [... of [23] ...] showed that some $f^{\theta-2}L(f)$ noises have a very erratic sampling behavior. Some other $f^{\theta-2}$ noises are Gaussian and, therefore, perfectly "well-behaved;" an example is provided by the "fractional white noise" [i.e. fGn] which is the formal derivative of the process of Mandelbrot and Van Ness 1968 [i.e. fBm]

He was referring here to the behaviour of several quantities calculated from a given realisation of a fractional time process, including $N(t)$, the number of switching events observed in the time interval, and time averages (taken over the interval 0 to t) of his "indicator" and "core" functions V and W. He related these to the skewed Lévy-stable distributions. The calculations were only performed in the finite mean waiting time case, however, and discussion of the infinite mean waiting time case was advertised as being in [24]. This latter is a much

more mathematical contribution, the fruit of what seems to have been a rather bruising (invited) encounter with Berkeley mathematicians in 1965 at the Fifth Symposium on Mathematical Statistics and Probability.

I have been unable so far to fully elucidate its content, but in any event, Niemann et al. [7] have recently given a very precise analysis of the behaviour of the frequency averaged spectra in this infinite mean case, first obtaining its Mittag-Leffler distribution analytically and checking this by simulations (see e.g. their fig. 2).

4 The Hurst effect vs. $1/f$ vs. LRD

Informed in part by the above historical investigations, the purpose of this section is now to distinguish conceptually between 3 things which are still frequently, and mistakenly, regarded as the same. To recap, the phenomena are:

- the Hurst effect: the *observation* of "anomalous" growth of range in a time series using a diagnostic such as Hurst and Mandelbrot's R/S or detrended fluctuation analysis (DFA)(e.g. [12,18]);
- $1/f$ noise: the *observation* of singular low frequency behaviour in the empirical periodogram of a time series;
- long range dependence (LRD): a property of a stationary model *by construction*. This can only be *inferred* to be a property of an empirical time series if certain additional conditions are known to be met, including the important one of stationarity.

The reason why it is necessary to unpick the relationship between these ideas is that there are three commonly held misconceptions about them.

The first misconception is that observation of the Hurst effect in a time series necessarily implies stationary LRD. This is "well known" to be erroneous, see e.g. the work of [34] who showed the Hurst effect arising from an imposed trend rather than from stationary LRD, but is nonetheless in practice still not very widely appreciated.

The second misconception is that observation of the Hurst effect in a time series necessarily implies a periodogram of power law form. Although less "well known" [35], for example, have shown an example where the Hurst effect arose in the Lorenz model which has an exponential power spectrum rather than $1/f$.

The third misconception is the idea that observation of a $1/f$ periodogram necessarily implies stationary LRD. As noted above, this is a more subtle issue, and although little appreciated since the pioneering work of [21–23] it has now become central to the investigation of weak ergodicity breaking in physics.

4.1 The Hurst effect

The Hurst effect was originally observed as the growth of range in a time series, at first the Nile. The original diagnostic for this effect was rescaled range, or R/S. Using the notation J (not H) for the Joseph (i.e. Hurst) exponent that Mandelbrot latterly advocated [14], the Hurst effect

is seen when the R/S [12,18] grows with time as

$$\frac{R}{S} \sim \tau^J, \tag{8}$$

in the case that $J \neq 1/2$. During the period between Feller's proof that an iid stationary process had $J = 1/2$, and Mandelbrot's papers of 1965–68 on long range dependence in fGn, there was a controversy [12] about whether the Hurst effect was a consequence of nonstationarity and/or a pre-asymptotic effect. The controversy has never fully subsided [12] because Occam's Razor frequently favours at least the possibility of change points in an empirically measured time series (e.g. [36]), and because of the (at first sight surprising) non-Markovian property of fGn.

The latter objection was addressed by Mandelbrot when he was interviewed by physicist Bernard Sapoval for the Web of Stories project in the 1990s. Showing the influence of subsequent developments in the physics on critical phenomena on his worldview, he explained how he had by then come to view LRD models like fGn:

> The consequences of this fundamental idea are hard to accept … [a]nd many people in many contexts have been arguing strongly against it, … If infinite dependence is necessary it does not mean that IBM's details of ten years ago influence IBM today, because there's no mechanism within IBM for this dependence. However, IBM is not alone. The River Nile is [not] alone. They're just one-dimensional corners of immensely big systems. The behaviour of IBM stock ten years ago does not influence its stock today through IBM, but IBM the enormous corporation has changed the environment very strongly. The way its price varied, went up, or went up and fluctuated, had discontinuities, had effects upon all kinds of other quantities, and they in turn affect us. And so my argument has always [sic] been that each of these causal chains is totally incomprehensible in detail, [and] probably exponentially decaying. There are so many of them that a very strong dependence may be perfectly compatible.

A key point to appreciate is that it is easier to generate the Hurst effect over a finite scaling range, as measured for example by R/S, than it is to generate a true $1/f$ spectrum over many decades [35], for example shows how a Hurst effect can appear over a finite range even when the power spectrum is known a priori to not be $1/f$, e.g. in the Lorenz attractor case where the low frequency spectrum is in fact exponential.

4.2 $1/f$ spectra

The term $1/f$ spectrum is usually used to denote periodograms where the spectral density $S'(f)$ has an inverse power law form, e.g. the definition used in [22,23]

$$S'(f) \sim f^{\theta-2}, \tag{9}$$

where θ runs between 0 and 2.

One needs to distinguish here between bounded and unbounded processes. Brownian, and fractional Brownian, motions are unbounded, nonstationary random walks and one can view their $1/f^{1+2J}$ spectral densities as a direct consequence of nonstationarity, as Mandelbrot did (see pp. 78–79 of [25]). In many physical contexts however, such as the on-off blinking quantum dot process [7] or the river Nile minima studied by Hurst [12] the signal amplitude is always bounded and does not grow in time, requiring a different explanation that is either stationary like fGn or "conditionally stationary" like the FRP.

Mandelbrot's best known model for $1/f$ noise remains the stationary, ergodic, fractional Gaussian noise (fGn) that he advocated so energetically in the 1960s. But, evidently aware that this had had received a disproportionate amount of attention, he was at pains late in his life (e.g. page 207 of Selecta Volume N [25], introducing the reprinted [22,23]) to stress that:

> Self-affinity and an $1/f$ spectrum can reveal themselves in several quite distinct fashions … forms of $1/f$ behaviour that are predominantly due to the fact that a process does not vary in "clock time" but in an "intrinsic time" that is fractal. Those $1/f$ noises are called "sporadic" or "absolutely intermittent", and can also be said to be "dustborne" and "acting in fractal time".

He thus clearly distinguished LRD stationary ergodic Gaussian models like fGn from his "conditionally stationary" FRP, noting also that:

> There is a sharp contrast between a highly anomalous ("non-white") noise that proceeds in ordinary clock time and a noise whose principal anomaly is that it is restricted to fractal time.

In practise the main importance of this is to caution that, used on its own, even a sophisticated approach to the periodogram like the GPH method [18] cannot tell the difference between a time series being stationary LRD and "just" a $1/f$ noise, unless independent information about stationarity is also available.

One route to reducing the ambiguity in future studies of $1/f$ is to develop non-stationary extensions to the Wiener–Khinchine theorem. An important step [37] has been to distinguish between one which relates the spectrum and the ensemble averaged correlation function, and a second relating the spectrum to the time averaged correlation function. The importance of this distinction can be seen by considering Fourier inverting the power spectrum, i.e. does inversion yield the time or the ensemble average? [E. Barkai, personal communication]. Another is to increase the emphasis on statistical hypothesis testing, where the degree of support between models like ARFIMA and its seasonal or heavy tailed variants is compared (e.g. [38]).

4.3 LRD

Readers will, I hope, now be able to see why I believe that the commonly used spectral definition of LRD has caused misunderstandings. The problem has been that on its own a $1/f$ behaviour is necessary but not sufficient, and stationarity is also essential for LRD in the sense so widely studied in statistics community (e.g. in [17,18]). One may in fact argue that the more crucial aspect of LRD is thus the "loose" one embodied in its name, rather than the formal one embodied in the spectral definition, because *a $1/f$ spectrum can only be synonymous with LRD when there is an infinitely long past.* The fact that fGn exhibits LRD *by construction* because the stationarity property is assumed, and *also* shows $1/f$ noise, and the Hurst effect has led to the widespread misconception that the converse is true, and that observing $1/f$ spectra and/or the Hurst effect *must imply* LRD.

5 Conclusion: beyond Mandelbrot's fractional renewal models

Unfortunately [23] and its predecessors received far less contemporary attention than did Mandelbrot's papers on heavy tails in finance in the early 1960s or the series with van Ness and Wallis in 1968–69 on stationary fractional Gaussian models for LRD, gaining only about 20 citations in its first 20 years. The fact that his work on the AFRP was communicated primarily in the (IEEE) journals and conferences of telecommunications and computer science concealed it from the contemporary audience that encountered fGn and fBm first in SIAM Review and Water Resources Research. Whatever the explanation, it was so invisible that one of his most articulate critics, hydrologist Vit Klemeš, actually proposed [39] an AFRP model as a paradigm for the *absence* of the type of LRD seen in the stationary fGn model, clearly unaware of Mandelbrot's work. Sadly Klemeš and Mandelbrot seem not to have subsequently debated FRP and fGn approaches either, as with the advantage of historical distance, and new theory [19] one can see the importance of both non-ergodic and ergodic models to the $1/f$ question. Leibovich and Barkai [19] have pointed out that there is a fundamental difference between measurement of $1/f$ noise on the single molecule level and measurements of a large ensemble of fluctuating units, in that the former exhibit a time dependent spectrum, and the latter do not. In their view this partially explains why it took 50 years to confirm Mandelbrot's prediction of the form of the time dependent spectrum. The experiments on blinking dots are single molecule experiments, where ensemble averaging is removed.

Although he revisited the 1963–67 fractional renewal papers with new commentaries in the volume of his Selecta [25] that dealt with multifractals and $1/f$ noise, Mandelbrot himself did not mention them explicitly in his popular historical account of the genesis of LRD [40]. It is clear that he saw the FRP and FGn as a representing two different strands from the way each was allocated a separate *Selecta* volume [14,25]. Despite the *Selecta*, the relatively

low visibility seems to have remained to the recent past. Mandelbrot's fractional renewal papers are for example not cited or discussed in Beran et al.'s encyclopaedic book on LRD [18]. Even when cited the FRP papers' actual content seems not always to be known, and I can personally attest to their low visibility in physics, having not come across them until 2014. A notable recent exception to this was a paper by Lenoir [41] which has picked up on the "conditional stationarity" idea.

The relative invisibility of the 1963–67 papers has, however, allowed a fruitful period of independent confirmation by rediscovery which has also seen several key new results not obtained by Mandelbrot, which are developing the field well beyond its origins. These have included:

– experimental confirmation of the time-dependent spectrum [8,32], the absence of which may have contributed to Mandelbrot's relative lack of subsequent emphasis on his fractional renewal models;
– a modern theory [37,42] using scale invariant autocorrelation functions of the form $< I(t)I(t+\tau) >= t^g\phi(\tau/t)$, implying a wider range of models and systems beyond renewal theory;
– extension of the Wiener–Khinchine theorem to this class of processes [37,42];
– explicit calculation of the effect of conditional stationarity on non-ergodicity [7];
– the emphasis of Bouchaud et al. [43] on the effect on the power spectrum of the waiting time t_w between the onset of a nonstationary process and the beginning of a measurement of duration T in the interval $t_w, t_w + T$, as distinct from the previously noted dependence of the spectrum on the measurement interval T. While Mandelbrot considered the case $t_w = 0$, the opposite case $t_w \gg T$ can be physically important.

One long term consequence of the low visibility of non-ergodic solutions to the $1/f$ problem in the physics and statistics literatures may have been to emphasise ergodic mechanisms at their expense. I believe this to have been important, because, for example, Per Bak et al.'s paradigm of Self-Organised Criticality, in which stationary spectra and correlation functions play an essential role, could surely not have been positioned as the unique solution to the $1/f$ problem [44] if it had been widely recognised just how different Mandelbrot's two existing routes to $1/f$ already were. In addition, I think that the route that Mandelbrot took from the fractional renewal models to multifractality (see e.g. pp. 243–246 of [25]) will repay further historical investigation, and may even yield a better physical appreciation for models which are still frequently seen as dauntingly abstract. I also hope to further investigate the idea of conditional stationarity, in order to clarify whether it was an intellectual dead-end or whether it may still have relevance to current work on weak ergodicity breaking and ageing. The adoption of the conditional language by [19] is an encouraging sign in this respect. I thus plan to return to the history of this period (see also [12]) in future articles.

References

1. E. Montroll, G.H. Weiss, J. Math. Phys. **6**, 167 (1965)
2. M. Shlesinger, Eur. Phys. J. B **90**, 93 (2017)
3. M.A. Rodriguez, Phys. Rev. E **90**, 042122 (2014)
4. M.A. Rodriguez, Phys. Rev. E **92**, 012112 (2015)
5. J.-P. Bouchaud, J. Phys. I France **2**, 1705 (1992)
6. I. Goychuk, Commun. Theor. Phys. **62**, 497 (2014)
7. M. Niemann, E. Barkai, H. Kantz, Phys. Rev. Lett. **110**, 140603 (2013)
8. S. Sadegh, E. Barkai, D. Krapf, New J. Phys. **16**, 113054 (2015)
9. F.D. Stefani, J.P. Hoogenboom, E. Barkai, Phys. Today **62**, 34 (2009)
10. M. Shlesinger, Math and physics: Lévy flights and drives, in *Benoit Mandelbrot: a life in many dimensions*, edited by M. Frame (World Scientific, Singapore, 2015), pp. 449–457
11. N.W. Watkins, Mandelbrot's 1/f fractional renewal models of 1963–67: the non-ergodic missing link between change points and long range dependence, 2016, arXiv:1603.00738
12. T. Graves, R. Gramacy, N.W. Watkins, C.L.E. Franzke, Entropy **19**, 437 (2017)
13. P. Grigolini, G. Aquino, M. Bologna, M. Lukovic, B.J. West, Physica A **388**, 4192 (2009)
14. B.B. Mandelbrot, *Gaussian self-affinity and fractals: globality, the earth, 1/f noise, and R/S, Selecta volume H* (Springer, Berlin, Heidelberg, 2002)
15. I.M. Sokolov, J. Klafter, A. Blumen, Phys. Today **55**, 48 (2002)
16. G. Margolin, E. Barkai, J. Stat. Phys. **122**, 137 (2006)
17. J. Beran, *Statistics for long-range memory processes* (Chapman and Hall, 1994)
18. J. Beran et al., *Long memory processes* (Springer, Berlin, Heidelberg, 2013)
19. N. Leibovich, E. Barkai, Phys. Rev. E **96**, 032132 (2017)
20. M. Berger, B.B. Mandelbrot, IBM J. Res. Dev. **7**, 224 (1963)
21. B.B. Mandelbrot, IEEE Trans. Commun. Technol. **COM-13**, 71 (1965)
22. B.B. Mandelbrot, Time varying channels, 1/f noises, and the infrared catastrophe: or why does the low frequency energy sometimes seem infinite? in *IEEE Communication Convention, Boulder Colorado, 1965* (1965)
23. B.B. Mandelbrot, IEEE Trans. Inf. Theory **13**, 289 (1967)
24. B.B. Mandelbrot, Sporadic random functions and conditional spectra; self-similar examples and limits, in *Fifth Symposium on Mathematical Statistics and Probability, Berkeley, 1965* (University of California Press, Oakland, 1967), p. 3
25. B.B. Mandelbrot, *Multifractals and 1/f noise: wild self-affinity in physics (1963–1976), Selecta volume N* (Springer, Berlin, Heidelberg, 1999)
26. P. Mertz, IRE Trans. Commun. Syst. **9**, 130 (1961)
27. P. Mertz, Impulse noise and error performance in data transmission, Memorandum RM-4526-PR, RAND Santa Monica, 1965
28. J. Bendat, *Principles and applications of random noise theory* (Wiley, 1958)
29. S.B. Lowen, M.C. Teich, Phys. Rev. E **47**, 992 (1993)
30. S.B. Lowen, M.C. Teich, *Fractal-based point processes* (Wiley, Hoboken, New Jersey, 2005)
31. J. Giles, Nature **432**, 266 (2004)
32. K.A. Takeuchi, J. Phys. A: Math. Theor. **50**, 264006 (2017)
33. J. Herault, F. Pétrélis, S. Fauve, EPL **111**, 44002 (2015)
34. R.N. Bhattacharya, V.K. Gupta, E. Waymire, J. Appl. Prob. **20**, 649 (1983)
35. C.L.E. Franzke, S.M. Osprey, P. Davini, N.W. Watkins, Sci. Rep. **5**, 9068 (2015)
36. T. Mikosch, C. Starica, REVSTAT-Stat. J. **2**, 41 (2004)
37. N. Leibovich, E. Barkai, Phys. Rev. Lett. **115**, 080602 (2015)
38. T. Graves, PhD. thesis, Statistics Laboratory, Cambridge University, 2013
39. V. Klemes, Water Resour. Res. **10**, 675 (1974)
40. B.B. Mandelbrot, R.L. Hudson, *The (mis)behaviour of markets: a fractal view of risk, ruin and reward* (Profile Books, London, 2008)
41. B. Lenoir, Fluct. Noise Lett. **12**, 1350006 (2013)
42. A. Dechant, E. Lutz, Phys. Rev. Lett. **115**, 080603 (2015)
43. J.-P. Bouchaud, L.F. Cugliandolo, J. Kurchan, M. Mézard, in *Out of equilibrium dynamics in spin-glasses and other glassy systems "Spin Glasses and Random Fields"*, edited by A.P. Young (World Scientific, Singapore, 1998), pp. 161–224
44. N.W. Watkins, G. Pruessner, S.C. Chapman, N.B. Crosby, H.J. Jensen, Space Sci. Rev. **198**, 3 (2016)