Tweedie convergence: A mathematical basis for Taylor's power law, $1/f$ noise, and multifractality

Wayne S. Kendal[1,*] and Bent Jørgensen[2,†]
[1]*Division of Radiation Oncology, University of Ottawa, 501 Smyth Road, Ottawa, Ontario, Canada K1H 8L6*
[2]*Department of Mathematics and Computer Science, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark*
(Received 15 July 2011; revised manuscript received 18 November 2011; published 27 December 2011)

Plants and animals of a given species tend to cluster within their habitats in accordance with a power function between their mean density and the variance. This relationship, Taylor's power law, has been variously explained by ecologists in terms of animal behavior, interspecies interactions, demographic effects, etc., all without consensus. Taylor's law also manifests within a wide range of other biological and physical processes, sometimes being referred to as fluctuation scaling and attributed to effects of the second law of thermodynamics. $1/f$ noise refers to power spectra that have an approximately inverse dependence on frequency. Like Taylor's law these spectra manifest from a wide range of biological and physical processes, without general agreement as to cause. One contemporary paradigm for $1/f$ noise has been based on the physics of self-organized criticality. We show here that Taylor's law (when derived from sequential data using the method of expanding bins) implies $1/f$ noise, and that both phenomena can be explained by a central limit-like effect that establishes the class of Tweedie exponential dispersion models as foci for this convergence. These Tweedie models are probabilistic models characterized by closure under additive and reproductive convolution as well as under scale transformation, and consequently manifest a variance to mean power function. We provide examples of Taylor's law, $1/f$ noise, and multifractality within the eigenvalue deviations of the Gaussian unitary and orthogonal ensembles, and show that these deviations conform to the Tweedie compound Poisson distribution. The Tweedie convergence theorem provides a unified mathematical explanation for the origin of Taylor's law and $1/f$ noise applicable to a wide range of biological, physical, and mathematical processes, as well as to multifractality.

DOI: 10.1103/PhysRevE.84.066120 PACS number(s): 89.75.Da, 05.40.Ca

I. INTRODUCTION

Plants and animals tend to aggregate within their habitats such that the variance of their density relates to the mean in accordance with a power function relationship [1]. Taylor, who first observed this effect with viruses, protozoa, insects, mollusks, vertebrates, and plants [1,2], explained it in terms of intraspecies behavior [3]; others have postulated demographic mechanisms [4], stochastic variations in reproductive rates [5], and interspecies interactions [6]. Taylor's power law also manifests within nonecological systems such as HIV epidemiology [7], regional organ blood flow [8], the genomic distributions of single nucleotide polymorphisms (SNPs) [9], and genes [10], as well as within physical and econometric systems where it has been called fluctuation scaling [11,12]. When applied to sequential data, with expanding enumerative bins, Taylor's law also implies long-range correlations and $1/f$ noise [13]. Taylor's law, and consequently $1/f$ noise, can be shown to have their origin through a mathematical convergence effect related to the central limit theorem (CLT) [14]. In this regard there exists a class of probabilistic models, the Tweedie exponential dispersion models, characterized by Taylor's law, that act as convergence foci for a broad range of non-Gaussian data [15]. It will be argued here that virtually any statistical model designed to produce Taylor's law must, on mathematical grounds, converge to a Tweedie distribution. This convergence theorem provides a mechanistic explanation for the disparate biological, physical, and mathematical manifestations of both Taylor's law and $1/f$ noise, based upon the statistical theory of errors.

Taylor used the empirical power function $\hat{\sigma}^2 = a\hat{\mu}^p$, between the variance $\hat{\sigma}^2$ of the spatial density of organisms and its corresponding mean $\hat{\mu}$, to describe aggregation within their environments (a and p are constants). In most ecological cases p tends to range between 1 and 2, which indicates clustering. The logarithm of this power function yields a linear relationship, with p now expressed as a slope, thus providing a simple demonstration of what has become known as Taylor's power law.

II. SELF-SIMILAR PROCESSES

We will focus here on Taylor's power law within sequential data, where in the description of self-similar processes it manifests implicitly. Much of the early development of self-similar processes was provided by Leland [16], which we summarize here: Consider the discrete sequence $Y = (Y_i: i = 0,1,2,...,N)$, which could represent time series data, spatial measurements, or abstract mathematical sequences. Given the mean $\hat{\mu} = E[Y_i]$ and the deviations from the mean $y_i = Y_i - \hat{\mu}$ one may construct an autocorrelation function

$$r(k) = \mathrm{E}[y_i y_{i+k}]/E\left[y_i^2\right], \tag{1}$$

with the lag k. The corresponding variance is also

$$\mathrm{var}[Y] = \mathrm{var}[y] = \hat{\sigma}^2 = E\left[y_i^2\right]. \tag{2}$$

Self-similar processes have, by their definition, long-range autocorrelations of the form

$$r(k) \sim k^{-\beta} L(k), \qquad k \to \infty, \tag{3}$$

*wkendal@ottawahospital.on.ca
†bentj@stat.sdu.dk

 ©2011 American Physical Society

where the exponent β is a real-valued constant bounded by $0 < \beta < 1$, and $L(k)$ is a slowly varying function for large values of k. One can construct a cover of equal-sized, adjacent and nonoverlapping enumerative bins of integer size m to produce the new sequences $Y^{(m)}$ with a reproductive property,

$$Y_i^{(m)} = \frac{1}{m}(Y_{im-m+1} + \cdots + Y_{im}), \qquad i > 1. \tag{4}$$

Here m is chosen so that N/m is an integer. The mean $\hat{\mu}$ and variance $\hat{\sigma}^2$ of Y can be regarded as constants, provided that the initial sequence remains unaltered during the analysis, $\hat{\mu} = E[Y] = E[Y^{(m)}]$. The reproductive sequences $Y^{(m)}$ will obey the variance–bin-size relationship,

$$\mathrm{var}[Y^{(m)}] = \hat{\sigma}^2 m^{-\beta}, \tag{5}$$

if and only if the autocorrelation function of the primary sequence takes the form [17]

$$r(k) = \tfrac{1}{2}[(k+1)^{2-\beta} - 2k^{2-\beta} + (k-1)^{2-\beta}]. \tag{6}$$

This autocorrelation can be shown to have the limiting behavior,

$$\lim_{k\to\infty} \frac{r(k)}{k^{-\beta}} = \frac{1}{2}(2-\beta)(1-\beta), \tag{7}$$

as per Eq. (3). The constant β given here relates to the Hurst parameter H introduced by Mandelbrot and van Ness to describe fractional Brownian motion [18], such that $\beta = 2(1-H)$. H is a real-valued number within the interval (0,1); if $H = 1/2$, the process represents Brownian motion, $H > 1/2$ implies a positive correlation in the increments of the process, and $H < 1/2$ implies a negative correlation.

One can also construct a second set of sequences that are additive:

$$Z_i^{(m)} = (Y_{im-m+1} + \cdots + Y_{im}). \tag{8}$$

These reproductive and additive sequences are related to each other by the equation $Z_i^{(m)} = mY_i^{(m)}$, from which we have the relationships between their means and variances: $E[Z^{(m)}] = mE[Y^{(m)}]$ and $\mathrm{var}[Z^{(m)}] = m^2\,\mathrm{var}[Y^{(m)}]$. Additive sequences, derived by the method of expanding enumerative bins, thus have the variance function

$$\mathrm{var}\left[Z_i^{(m)}\right] = m^2\,\mathrm{var}[Y^{(m)}] = (\hat{\sigma}^2/\hat{\mu}^{2-\beta})E\left[Z_i^{(m)}\right]^{2-\beta}. \tag{9}$$

Since $\hat{\mu}$ and $\hat{\sigma}^2$ are constants, Eq. (9) expresses Taylor's power law with the exponent $p = 2-\beta$, and thus $H = p/2$. Consequent to the biconditional relationship between Eqs. (5) and (6), any sequence that reveals Taylor's law, by the method of expanding bins, will also express long-range autocorrelations that approximate $r(k) \sim c_1 k^{-\beta}$, where c_1 is a constant.

As examples of Taylor's law from sequential data we present here the $N \times N$ random symmetric matrices with Gaussian distributed elements known as the Gaussian unitary ensemble (GUE) and the Gaussian orthogonal ensemble (GOE) [19]. The GUE and GOE, being purely mathematical constructs, are germane to our argument that the demonstration of Taylor's law and $1/f$ noise would be best attributable to a universally applicable mathematical mechanism rather than to the *ad hoc* behavioral, biological, or physical processes that have been conventionally considered.

The GUE represents complex Hermitian matrices, invariant under unitary transformations; the GOE are real symmetric matrices, invariant under orthogonal transformations. The diagonal elements H_{nn} of a $N \times N$ GOE matrix are Gaussian distributed with mean of 0 and variance of 1; for $m < n$ the elements H_{nm} are Gaussian distributed with mean of 0 but variance of $\frac{1}{2}$; elements for which $n > m$ are given by $H_{mn} = H_{nm}$. In the GUE the diagonal elements are the same; however, for $m < n$ the elements H_{nm} have independent real and imaginary components that are Gaussian distributed with mean of 0 and variance of $\frac{1}{2}$; those for $n > m$ are given by the conjugate transpose $H_{mn} = H_{nm}^*$.

The ranked eigenvalues from such random matrices are not uniformly distributed. Wigner proposed that the average density of the ranked eigenvalues $\bar{\rho}(E)$ should obey a semicircular probability density as $N \to \infty$ so that

$$\bar{\rho}(E) = \begin{cases} \sqrt{2N - E^2}/\pi, & |E| < \sqrt{2N} \\ 0, & |E| > \sqrt{2N} \end{cases} \tag{10}$$

for eigenvalues of magnitude E [19]. This semicircular rule can be integrated to obtain the number of eigenvalues on average less than the value E,

$$\bar{\eta}(E) = \frac{1}{2\pi}\left[E\sqrt{2N - E^2} + 2N\,\sin^{-1}\left(\frac{E}{\sqrt{2N}}\right) + \pi N\right]. \tag{11}$$

This function can be used to unfold (renormalize) individual eigenvalues E_n by means of the equation

$$e_n = \bar{\eta}(E_n) = \int_{-\infty}^{E_n} dE'\,\bar{\rho}(E'). \tag{12}$$

Unfolding effectively removes the main trend of the sequence from the fluctuating portion, and allows for a comparison between different parts of the fluctuating portion. In the unfolded metric the average of the spacing interval, $s_n = e_{n+1} - e_n$, between consecutive ranked eigenvalues is 1, and the individual spacing values will obey specific distribution functions for each ensemble. These nearest neighbor spacing distributions indeed have been used to model the spectral fluctuations of quantum systems [20].

We will focus on the ranked eigenvalues E_1, $E_2 \ldots E_N$ of these $N \times N$ random matrices to estimate the deviations between the actual and the expected cumulative number of eigenvalues,

$$\bar{D}_n = n - \bar{\eta}(E_n). \tag{13}$$

This particular measure has also been employed in number theoretic studies to study the fluctuations of the prime counting function $\pi(x)$ from predictive formulas for the prime number positions [21], as well as in random matrix theory (RMT) to study the energy levels of complex systems [22].

We begin our examples with an empirical examination of the absolute values of these deviations $|\bar{D}_n|$ as obtained from $10\,000 \times 10\,000$ matrices of the GUE [Fig. 1(a)] and GOE [Fig. 2(a)]. The plots of these sequences demonstrated regions where larger deviation values appeared to be clustered together. A closer inspection of these deviations indicated the existence of cusplike patterns suggestive of multifractal singularities.

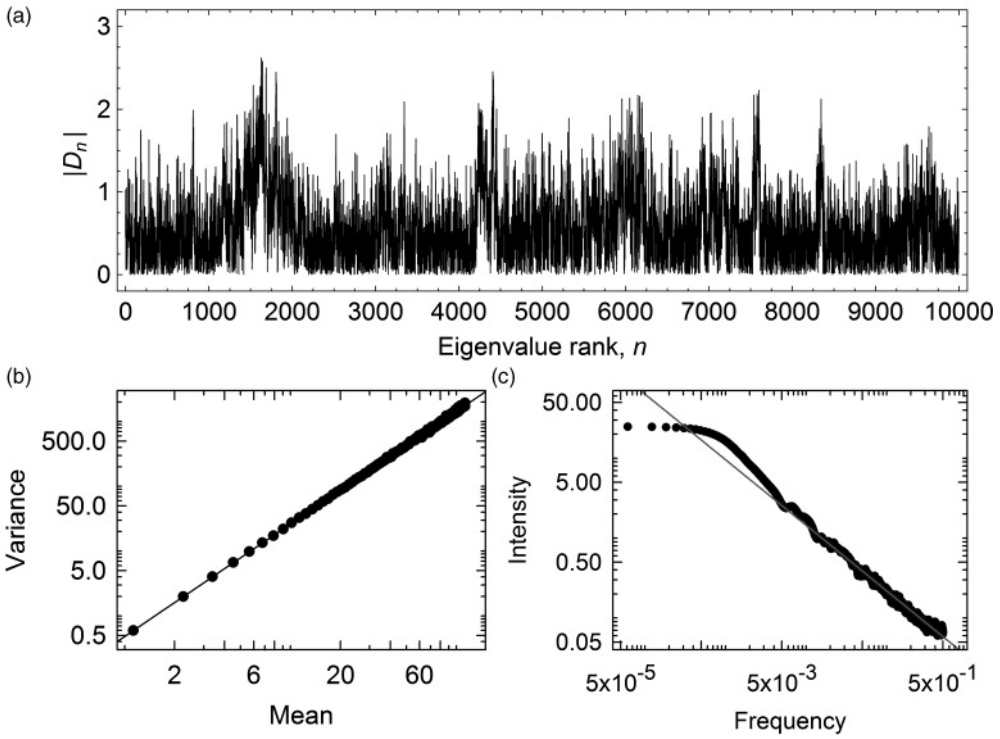

FIG. 1. (a) GUE deviations. The deviations $|\bar{D}_n|$ between the observed and predicted number of eigenvalues below the size E_n were estimated for a 10 000 × 10 000 matrix from the GUE, and are plotted here. (b) Variance function. The data sequence $|\bar{D}_n|$ was divided into a series of sequential, nonoverlapping and equal-sized counting bins; the values within each bin were summed and the mean and variance of these summed values were determined for the series of bins. The process was repeated for larger and larger bin sizes and the variances so obtained were plotted against their respective means on logarithmic axes. A power function relationship was obtained with $a = 0.483$, $p = 1.736$, and $r^2 = 0.999$. (c) Power spectrum for a GUE ensemble. For each of ten repetitions the sequence of deviations was packed with zeros to yield 16 384 points and the mean and trend were subtracted. A 49 point Hamming window was used to smooth these data and a fast Fourier transform was used to estimate the power spectrum by squaring the modulus of the transform. An ensemble average of the spectra was computed and then a log-log plot of these data revealed a power spectrum of the form $1/f^{0.82}$ on linear regression. An initial flattening is evident over about the lowermost ten points of the spectrum, comprising about 0.1% of the data sequence. This flattening can be attributed to an artifact of the discrete Fourier transform caused by the low frequency component introduced by the rectangular data window.

We will further consider possible multifractality within the sequences of deviations later; first, however, we assess the apparent clustering within these sequences by means of Taylor's law. Each sequence was divided into adjacent and nonoverlapping bins, two positions in length. The values of the $|\bar{D}_n|$ within each bin were summed; the mean and variance of these sums were calculated over all the bins. These calculations were repeated for successively expanding bin sizes, and the variances so obtained were plotted versus their corresponding means. Log-log plots of the variance to mean plots from the GUE [Fig. 1(b)] and the GOE [Fig. 2(b)] exhibited linear relationships, in accordance with Taylor's law.

This analysis of the GUE was repeated ten times; each time with independently derived data. The ensemble average of the power-law regressions yielded the constants and 95% confidence intervals (CI), $a = 0.507[0.472 - 0.542]$ and $p = 1.722[1.663 - 1.781]$. Similarly, an ensemble average of ten 10 000 × 10 000 matrices from the GOE yielded the values, $a = 0.612[0.559 - 0.664]$ and $p = 1.679[1.610 - 1.748]$. The degree of clustering, assessed by the magnitude of p, thus appeared similar between the GUE and GOE, with $1 < p < 2$. The behavior of the GUE and GOE variance functions were thus characterized; in all cases a close agreement was obtained with the power law. The statistics indicated that a sampling of ten random matrices of this size could provide a quantifiable assessment of the behavior of the GUE and GOE deviations; the values of p indicated clustering.

To explore the effect of matrix size on p, a number of ensembles from the GUE and GOE were assessed for different sizes (Fig. 3). For matrices ranging from 30 × 30 to 10 000 × 10 000 Taylor's law was evident with its exponent centered on an ensemble mean of $\langle p \rangle = 1.65$, with 90% of the individual estimates falling within the range of 1.29–1.92. We will return to these variations in p again, when we consider multifractality in the deviation sequences. Before this, however, we should discuss how Taylor's law in self-similar processes directly relates to $1/f$ noise and a class of statistical models that inherently express Taylor's law.

III. $1/f$ NOISE FROM TAYLOR'S LAW

When dealing with time series data the concept of frequency f is understood to represent the number of events per unit time sequence. This concept can be extended to nontemporal sequences, where the frequency is now measured in terms of the number of events seen per unit measure of the discrete

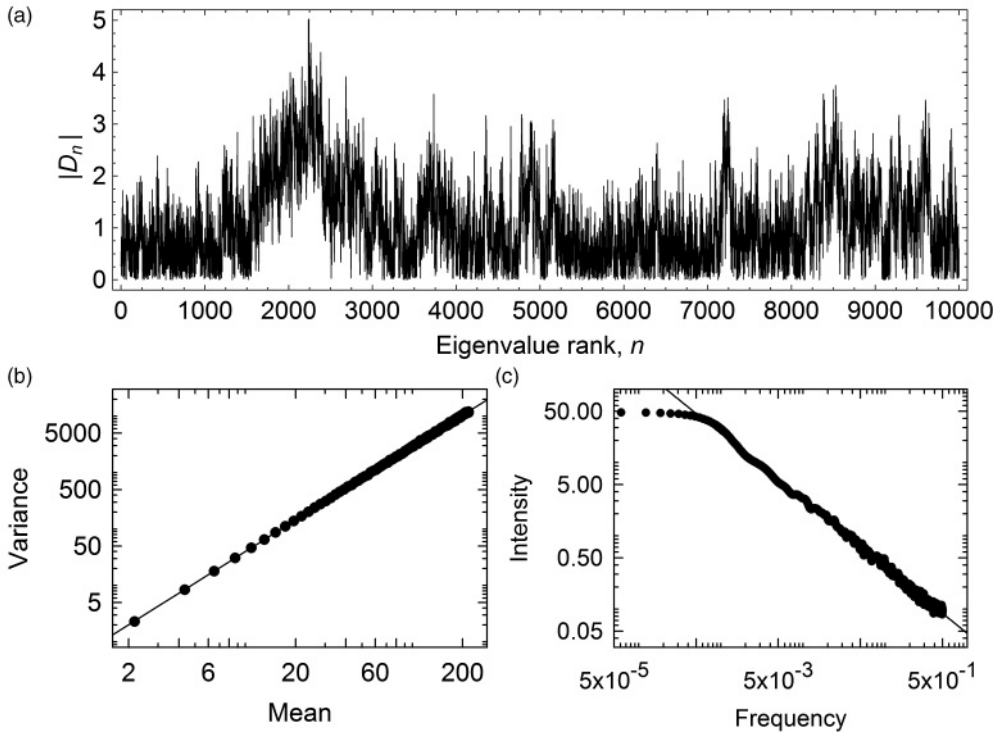

FIG. 2. (a) GOE deviations. The deviations $|\bar{D}_n|$ between the observed and predicted number of eigenvalues below a specified size were estimated from a 10 000 × 10 000 matrix from the GUE, and are plotted here. (b) Variance function. A variance to mean power function was obtained with the values $a = 0.558$, $p = 1.856$, and $r^2 = 1.000$ (c) Power spectrum for a GOE ensemble. Ten data sets were used to estimate an ensemble power spectrum, which on linear regression approximated the form $1/f^{0.91}$.

series. Numerical estimates for autocorrelations tend to be computationally less efficient than that of those for the spectral density function $S(f)$, which is related to the autocorrelation by the Fourier transformation,

$$S(f) = \int_{-\infty}^{\infty} r(k) e^{-2\pi i f k}\, dk. \tag{14}$$

It can be shown that the variance–bin-size relationship, Eq. (5), holds for sequences Y if and only if the corresponding power spectral density $S(f)$ takes the form [17]

$$S(f) \propto |e^{2\pi i f} - 1|^2 \sum_{l=\infty}^{l=\infty} 1/|f + l|^{3-\beta}, \quad -1/2 \leqslant f \leqslant 1/2, \tag{15}$$

where the frequency f is measured from the sequence intervals in analogy to time series sequences.

One may alternatively use the limit from Eq. (7) to approximate the long-range autocorrelation as $r(k) \sim c_1 k^{-\beta}$ and then apply the Wiener-Khintchine theorem [23] to surmise that the corresponding power spectral density should approximate $S(f) \sim c_2 f^{\beta-1}$, within the constant factors c_1 and c_2. We know that power spectra of the form $S(f) \propto 1/f^{\gamma}$, where $0 < \gamma < 2$, are the hallmarks of $1/f$ noise, and consequently we would anticipate a relationship between Taylor's power-law exponent and the spectral exponent such that $p \sim \gamma + 1$. By virtue of this biconditional relationship between Taylor's law [Eq. (9)], and its corresponding autocorrelation function (and the relationship between the autocorrelation and the spectral density), the manifestation of Taylor's law in sequential data implies the existence of $1/f$ noise.

Returning to the examples from the GUE and GOE, since their eigenvalue fluctuations express variance to mean power functions with exponents $p \sim 1.7$, we would anticipate power spectra proportional to about $1/f^{0.7}$. Indeed, direct Fourier analysis of the sequences $|\bar{D}_n|$ from the GUE and GOE was approximately consistent with this prediction [Figs. 1(c) and 2(c)]; the relationship between Taylor's law and $1/f$ noise was thereby supported by numerical analysis of the GUE and GOE.

IV. A CLASS OF PROBABILITY DISTRIBUTIONS CHARACTERIZED BY TAYLOR'S LAW

It is not widely recognized that Taylor's law is also a defining feature of a class of probability distributions that are invariant under changes of scale, known as the Tweedie exponential dispersion models (in recognition of the man who first described them) [15]. Exponential dispersion models (EDMs) describe error distributions for generalized linear models and can be employed to analyze a wide variety of non-normal data. There are two classes of these models: additive and reproductive. The additive models are described by a canonical equation based on the interrelated probability measures ν_λ and measurable sets A such that

$$P_{\lambda,\theta}(Z \varepsilon A) = \int_A \exp\left[\theta z - \lambda \kappa(\theta)\right] \nu_\lambda(dz). \tag{16}$$

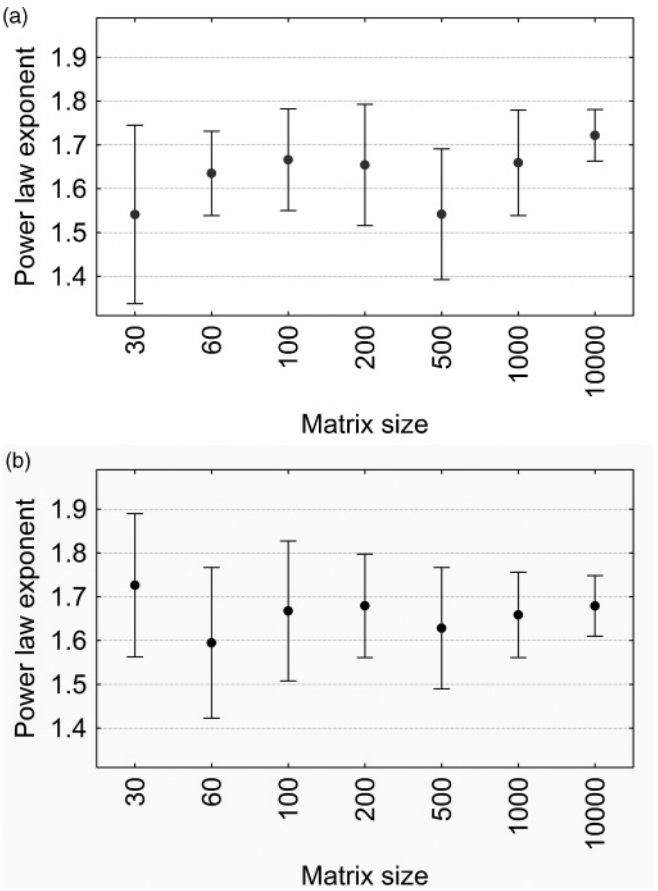

FIG. 3. Influence of matrix size on the power-law exponent. (a) GUE. Taylor's law exponent p was estimated for the deviations $|\bar{D}_n|$ of various sized matrices from the GUE. For each matrix size the calculations were repeated ten times with independently derived data. The mean value of p (solid dots) are provided along with their respective 95% CIs (error bars). For 30×30 matrices through to $10\,000 \times 10\,000$ matrices the mean values for p ranged between 1.54 and 1.72. (b) GOE. Taylor's law exponent p for the GOE deviations $|\bar{D}_n|$ were estimated for a range of matrix sizes and repeated ten times with independently derived data. The mean values of p, and their corresponding 95% CIs, are provided graphically. These mean values ranged between 1.59 and 1.73 indicating that, as in the case of the GUE, the manifestations of Taylor's law were essentially independent of matrix size.

Here θ is the canonical parameter and λ is the index parameter, analogous to the scale and shape parameters of conventional statistics, respectively.The function

$$\kappa(\theta) = (1/\lambda)\log\left[\int e^{\theta z}\nu_\lambda(dz)\right] \tag{17}$$

is known as the cumulant function. The distribution corresponding to the random variable $P_{\lambda,\theta}$ in fact represents a family of additive distributions $E_D^*(\theta,\lambda)$ that are completely determined by θ, λ, and $\kappa(\theta)$. This family has the property that the distribution of the sum of independent random variables, $Z_+ = Z_1 + \cdots + Z_n$ with $Z_i \sim E_D^*(\theta,\lambda_i)$ corresponding to fixed θ and various values of λ, belongs to the family of distributions with the same θ, $Z_+ \sim E_D^*(\theta,\lambda_1 + \cdots + \lambda_n)$.

The cumulant function can be used to construct the cumulant generating function (CGF) $K^*(s)$ for the additive model corresponding to the random variable Z based on the generating function variable s:

$$K^*(s) \equiv \log[E(e^{sZ})] = \lambda[\kappa(\theta + s) - \kappa(\theta)]. \tag{18}$$

The first two derivatives of $K^*(s)$ at $s = 0$ therefore give the mean and variance of Z.

The function $\tau(\theta) = \kappa'(\theta)$ is known as the mean value mapping; it gives the relationship between θ and the mean, $\mu = \kappa'(\theta)$. By means of this mapping we define the variance function $V(\mu) \equiv \tau'[\tau^{-1}(\mu)]$, where $\tau^{-1}(\mu)$ denotes the inverse function of $\tau(\theta)$. The mean and variance of an additive random variable are then $E(Z) = \lambda\mu$ and $\mathrm{var}(Z) = \lambda V(\mu)$.

The class of additive models is related to a second class of reproductive EDMs, described by the random variable $Y = Z/\lambda \sim E_D(\mu,\sigma^2)$, where $\sigma^2 = 1/\lambda$. These reproductive EDMs possess the convolution property such that for n independent random variables $Y_i \sim E_D(\mu,\sigma^2/w_i)$, with weighting factors w_i subject to the summation $w = \sum_{i=1}^n w_i$, we have $1/w\sum_{i=1}^n w_i Y_i \sim E_D(\mu,\sigma^2/w)$. This weighted average, corresponding to fixed μ and σ^2 and various values of w_i, belongs to the family of distributions with the same μ and σ^2. A duality transformation $Y \mapsto Z = Y/\sigma^2$ exists between the additive and reproductive EDMs. Consequently, any theory pertaining to one such class of models can be adapted to the other.

We have seen how EDMs are closed with respect to different types of convolution; there remains a further property where EDMs can be closed with respect to scale transformation. For example, a reproductive EDM $E_D(\mu,\sigma^2)$ can be required to obey $cE_D(\mu,\sigma^2) = E_D(c\mu,c^{2-p}\sigma^2)$, where c is any positive valued real constant and p is a real-valued, unitless constant. Under this scale transformation the new random variable $\hat{Y} = cY$ belongs to the same family of distributions with fixed μ and σ^2 but different values of c. The variance function will now necessarily obey the relationship $V(c\mu) = g(c)V(\mu)$ for some function $g(c)$. Scale invariance implies that $g(c) = V(c)$, and thus $V(\mu) = \mu^p$. This scale invariance becomes the defining feature of a class of EDMs known as the Tweedie models.

The general properties of exponential dispersion models provide two useful differential equations [24]: one is the relationship between the mean value mapping and the variance function,

$$\frac{\partial\tau^{-1}(\mu)}{\partial\mu} = \frac{1}{V(\mu)}. \tag{19}$$

The other shows how the mean value mapping relates to the cumulant function,

$$\frac{\partial\kappa(\theta)}{\partial\theta} = \tau(\theta). \tag{20}$$

These equations can be solved to express the cumulant function corresponding to the Tweedie EDMs for different values of p [24],

$$\kappa_p(\theta) = \begin{cases} e^\theta & \text{for} \quad p = 1 \\ \frac{\alpha-1}{\alpha}\left(\frac{\theta}{\alpha-1}\right)^\alpha & \text{for} \quad p \neq 1,2 \\ -\log(-\theta) & \text{for} \quad p = 2. \end{cases} . \tag{21}$$

The CGFs for the additive Tweedie EDMs also will take different forms depending upon the value of p,

$$K_p^*(s;\theta,\lambda) = \begin{cases} \lambda e^{\theta}\,(e^s - 1) & \text{for} \quad p = 1 \\ \lambda\kappa_p(\theta)\left\{\left(1 + \frac{s}{\theta}\right)^{\alpha} - 1\right\} & \text{for} \quad p \neq 1,2. \\ -\lambda\,\log\left(1 + \frac{s}{\theta}\right) & \text{for} \quad p = 2. \end{cases} \tag{22}$$

The exponent α, employed here for brevity, is related to the power-law exponent p, $\alpha = (p-2)/(p-1)$. For those readers who require a more detailed description of the Tweedie EDMs we recommend the monograph of Jørgensen [24].

A family of Tweedie models is thus described which includes the extreme stable ($p < 0$), Gaussian ($p = 0$), Poisson ($p = 1$), compound Poisson-gamma (PG) ($1 < p < 2$), gamma ($p = 2$), positive stable ($2 < p < 3$), inverse Gaussian ($p = 3$), positive stable ($p > 3$), and extreme stable ($p = \infty$) distributions [24]. The Tweedie models are therefore quite comprehensive; they include the Poisson, gamma, and Gaussian distributions as well as some less frequently employed distributions.

From these CGFs one can confirm that the variance of any additive Tweedie model will relate to the mean by the power function,

$$\text{var}(Z) = aE(Z)^p; \tag{23}$$

that is, Taylor's power law. For most ecological data as well as for the GUE and GOE deviations, $1 < p < 2$, corresponding to the compound PG distribution.

Pursuant to our examples of the GUE and GOE deviations $|\bar{D}_n|$, we now determine how well the Tweedie PG model corresponds to the empirical cumulative distribution functions (CDFs) from these random matrices. The PG distribution is specified by three independent adjustable parameters α, λ, and θ to be fitted to the data. The additive PG probability density function is given by the equation [24]

$$p^*(z;\theta,\lambda,\alpha) = c^*(z;\lambda)\exp[\theta z - \lambda\kappa(\theta)], \tag{24}$$

where

$$c^*(z;\lambda) = \begin{cases} \frac{1}{z}\sum\limits_{n=1}^{\infty} \lambda^n\kappa^n(-1/z)/\,\Gamma(-\alpha\cdot n)n! & \text{for} \quad z > 0 \\ 1 & \text{for} \quad z = 0. \end{cases} \tag{25}$$

The theoretical CDF was then fitted to the empirical CDFs. The probability-probability plots (Fig. 4) revealed qualitative agreement between the theoretical PG distributions and the data. Moreover, the fitted values for α obtained here were consistent with the exponent p obtained from the corresponding Taylor plot. We would expect for sequential data governed by the PG distribution (assessed by the method of expanding bins) to express Taylor's law with exponent $1 < p < 2$ and to provide $1/f$ power spectra that would give $1/f^{\gamma} \sim 1/f^{p-1}$ with $0 < p - 1 < 1$, approximately.

Earlier we noted that the power-law exponent p from such sequential data could be related to the Hurst parameter H through the behavior of the autocorrelation function, $H = p/2$. Given the range of values for p seen for the deviations $|\bar{D}_n|$ this would imply that $0.5 < H < 1$, indicating a degree

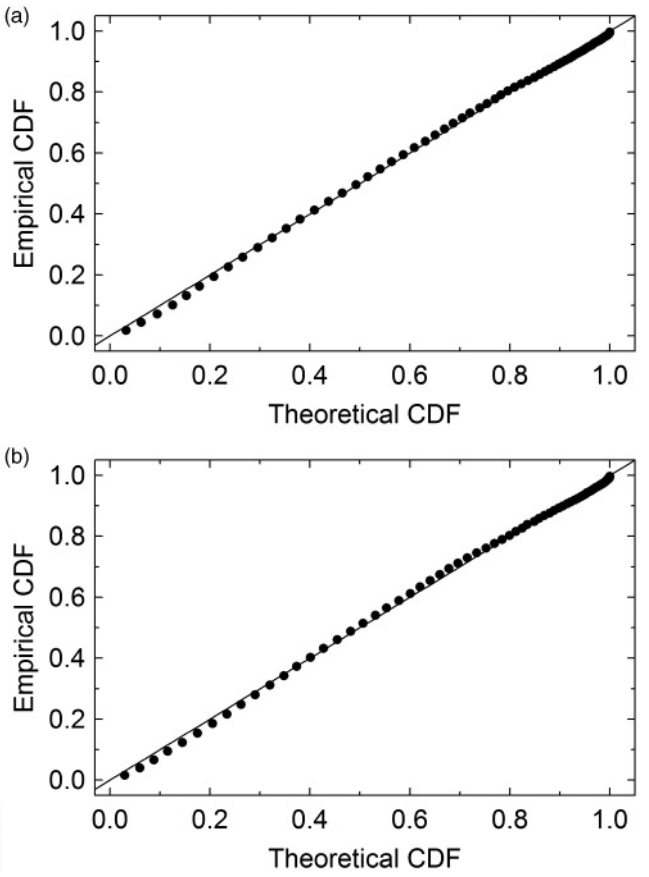

FIG. 4. Probability-probability plots. (a) GUE. An empirical CDF was derived from the data of Fig. 1(a), and a theoretical CDF, based on the Tweedie PG distribution [Eq. (25)], was fitted to the empirical CDF ($\theta = -3.717$, $\lambda = 2.186$, and $\alpha = -0.333$ so that $p = 1.75$). The empirical CDF was plotted versus the theoretical CDF to yield a linear relationship, demonstrating consistency of the Tweedie model with the GUE data. (b) GOE. The empirical CDF derived from the data of Fig. 2(a) was plotted versus a theoretical PG CDF ($\theta = -2.054$, $\lambda = 1.885$, and $\alpha = -0.358$ so that $p = 1.74$). The PG model was consistent with these data.

of positive correlation within the sequences of deviations. The Hurst parameter also relates linearly to the Hausdorff (or fractal) dimension D_H for the sequence of deviations,

$$D_H = \lim_{\varepsilon\to 0}\frac{\log[N(\varepsilon)]}{\log[1/\varepsilon]}, \tag{26}$$

such that for an affine process on an n-dimensional space, $D_H + H = n + 1$ [25], and where $N(\varepsilon)$ is the number of self-similar structures of linear size ε needed to cover the whole structure.

We have also seen, with the many assessments of p from the GUE and GOE deviations, that there existed a degree of variation associated with the values of p (Fig. 3). Two alternative hypotheses can be considered to account for these variations: (1) they result from random numerical error or (2) the exponents p exhibit a spectrum of values as an intrinsic property of the system. This second hypothesis (if true) would indicate a possible inherent multifractality of the sequences of deviations $|\bar{D}_n|$ [26]. Also, as noted earlier, the sequences of deviations seemed to exhibit cusplike

(multifractal) singularities. To further investigate this possibility we performed multifractal analyses on the deviations using the wavelet transform method [27]. Briefly, this method utilized the wavelet transform $T_\psi[f](\tilde{b},\tilde{a})$ of a function f by decomposition into contributions from an analyzing wavelet ψ by means of translations and dilations specified through the real-valued scale and shape parameters $\tilde{a} \in \mathbf{R}^+$ and $\tilde{b} \in \mathbf{R}$,

$$T_\psi[f](\tilde{b},\tilde{a}) = \frac{1}{\tilde{a}} \int_{-\infty}^{+\infty} \psi\left(\frac{x-\tilde{b}}{\tilde{a}}\right) f(x)dx. \tag{27}$$

Wavelets used in our analyses were chosen from successive derivatives of the Gaussian function,

$$\psi^{(N)}(x) = d^N(e^{-x^2/2})/dx^N. \tag{28}$$

This analysis relied on the local Hölder exponent to characterize singularities of a function f at some specified point x_0. The Hölder exponent $h(x_0)$ is the largest exponent for which a polynomial $P_n(x)$ of order n exists to satisfy the relation

$$|f(x) - P_n(x - x_0)| = O(|x - x_0|^h), \tag{29}$$

for x in some neighborhood of x_0. One defines the $D(h)$ singularity spectrum which represents the Hausdorff dimension for which the Hölder exponent carries the particular value h [27],

$$D(h) = \dim_H[x \,|\, h(x) = h]. \tag{30}$$

This singularity spectrum characterizes the intermittent fluctuations in the sequence over a range of Hölder exponent values. Parenthetically, we will add that in certain circumstances both h and $D(h)$ can assume positive and negative real values [28].

To facilitate the wavelet analysis we employed the partition function

$$Z(\tilde{a},q) = \sum_{l \in \mathcal{L}(\tilde{a})} \left\{ \sup_{(x,\tilde{a}') \in l} |T_\psi[f](x,\tilde{a}')| \right\}^q, \tag{31}$$

where $q \in \mathbf{R}$ is the order of the generalized fractal dimension and $\mathcal{L}(\tilde{a})$ is the set of connected wavelet maxima lines l_i which reach or cross the $\tilde{a}$ scale. Under the limit $\tilde{a} \to 0$ an exponent $\tau(q)$ (this exponent is not related to the mean value mapping from Sec. IV that employed a similar notation) can be determined from the power-law scaling of the partition function, $Z(\tilde{a},q) \sim \tilde{a}^{\tau(q)}$. The $D(h)$ spectrum is determined from the Legendre transform of $\tau(q)$,

$$D(h) = \min_q [qh - \tau(q)]. \tag{32}$$

Wavelet analysis of the deviations $|\bar{D}_n|$ provided multifractal spectra $\tau(q)$ and singularity spectra $D(h)$ for the GUE and GOE (Fig. 5). In both cases the multifractal spectra revealed inflexions and the singularity spectra revealed inverse convex forms, consistent with multifractality.

V. STATISTICAL CONVERGENCE EXPLAINS BOTH TAYLOR'S LAW AND $1/f$ NOISE

The correspondence between the PG distribution and the GUE/GOE eigenvalue fluctuations, as well as its correspondence within other biological systems that have exhibited Taylor's law [8–10,29,30], might be dismissed as an artifact of curve fitting of little consequence were it not for the mathematical role that the Tweedie models have as foci of statistical convergence: For EDMs $E_D(\mu,\sigma^2)$ with unit variance functions of the form $V(\mu) \sim \mu^p$, Jørgensen *et al.* have proven that as $\mu \to 0$ or $\mu \to \infty$ then, within the constant factor c, $c^{-1}E_D(c\mu,\sigma^2c^{2-p})$ will converge to the form of a Tweedie model as either $c \downarrow 0$ or $c \to \infty$ [14]. Since the variance functions for many probability distributions approximate the form $V(\mu) \propto \mu^p$, for small or large values of μ, the Tweedie EDMs act as the foci of convergence for a wide variety of data [14]. This convergence property appears related to stable generalizations of the CLT [24], suggesting that the Tweedie models have a role analogous to that of the Gaussian distribution in statistical theory. The biconditional relationship between Eqs. (5) and (6) connects Taylor's law to $1/f$ noise. The manifestation of $1/f$ noise can thus be viewed as a consequence of the Tweedie convergence theorem.

Up to this point the demonstrations of Taylor's law within the GUE and GOE have been presented as empiricisms. The Tweedie convergence theorem, however, specifies that the distribution functions for the deviations $|\bar{D}_n| = |n - \overline{\eta}(E_n)|$ would be expected to converge to the form of the PG distribution. For this reason we propose that the Tweedie PG distribution can be employed to represent these deviations.

At this point it would be useful to consider the range of applicability of the Tweedie convergence theorem, particularly since it appears so central to explaining the broad manifestations of Taylor's law and $1/f$ noise. The theorem states that any EDM with an asymptotic power variance function will be in the domain of attraction of a Tweedie distribution. It is actually quite difficult to propose an alternative distribution with a power variance function and finite CGF that is not in the Tweedie domain of attraction, and constructions so derived appear quite contrived. Moreover, any distribution function with finite CGF will belong to an EDM, so that effectively a wide range of distributions can be approximated by the Tweedie models.

Other alternative models for Taylor's law might be conceived (including convergence theorems [11]), however, limit theorems for independent and identically distributed variables, as demonstrated with the Tweedie convergence theorem, must be considered as fundamental explanations for the appearance of any given distributional form, and so the Tweedie models would be the first place to look for an explanation. One might still wish to consider alternatively some non-EDM that provides Taylor's law yet appears unrelated to any Tweedie model. But if one considers the space of all distributions with finite moment generating functions, the natural exponential families would form an equivalence relation on this space. Hence, this remaining alternative would be expected to be a member of some natural exponential family, as long as its moment generating function was finite. If we joined together natural exponential families to form EDMs, the same consideration would hold, and we would expect convergence toward a Tweedie model. Also, by the Tweedie convergence theorem, we would expect that a large proportion of all EDMs (i.e., those with power asymptotic variance functions) would be well approximated by Tweedie distributions. By similar arguments one would expect that processes subject

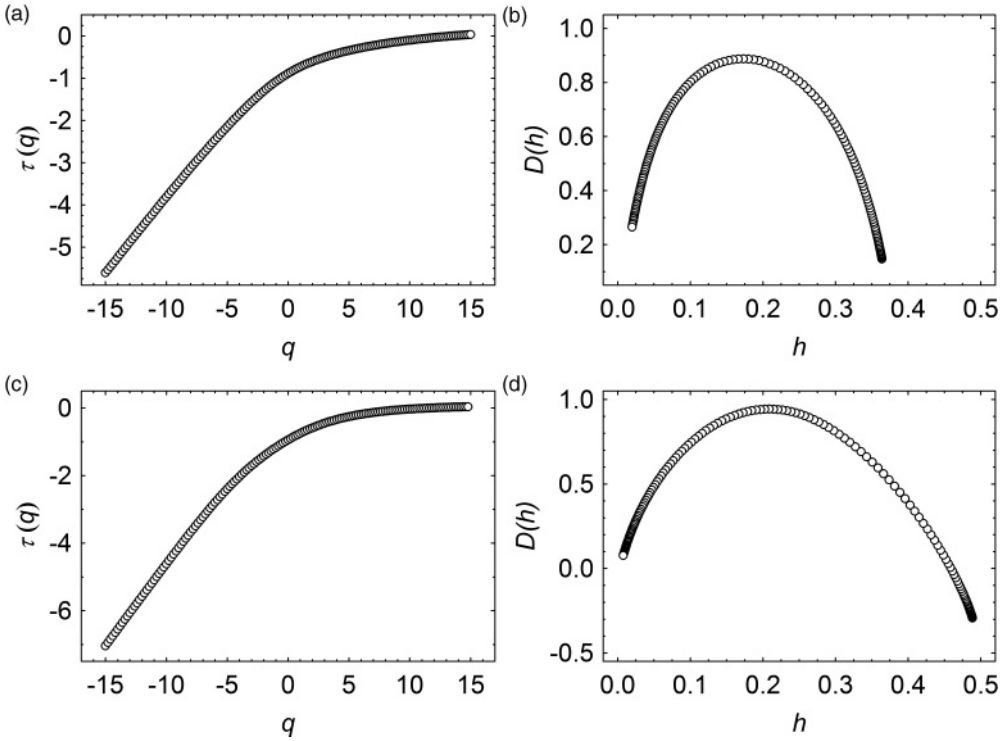

FIG. 5. Multifractal analysis. (a) Multifractal spectrum of the GUE deviations. The deviations $|\bar{D}_n|$ were estimated from a 10 000 × 10 000 matrix of the GUE and a multifractal analysis was conducted using analyzing wavelets based on the sixth derivative of the Gaussian function (software from http://www.physionet.org/physiotools/multifractal/). Here the multifractal scale exponent $\tau(q)$ is plotted versus the dimensional index q. (b) Singularity spectrum of the GUE deviations. The $D(h)$ spectrum was obtained by Legendre transform of the numerical data in (a), and is plotted versus the Hölder exponent h. (c) Multifractal spectrum of the GOE deviations. $|\bar{D}_n|$ were estimated from a 10 000 × 10 000 matrix of the GOE and analysis was conducted using analyzing wavelets based on the sixth derivative of the Gaussian function. (d) Singularity spectrum of the GOE deviations. This spectrum was obtained by Legendre transform of the numerical data in (c).

to many small independent perturbations would also tend to converge toward Tweedie distributions. For these reasons, the simulations and approximations used to model Taylor's law reported in the literature [1,4–6,11,31] would be expected to converge to the form of a Tweedie distribution.

One might ask whether an alternative explanation for Taylor's law could be sought amongst distributions that do not have finite moment generating functions. However such distributions, as a rule, do not have finite moments, and hence do not possess a variance function. But the whole point of Taylor's law is that the first two moments are related by a power law, so alternative explanations basically cannot come from considering distributions without moments. The physics community will also be familiar with distributions like this that exhibit heavy (or power law) tails such as the Pareto and Lévy distributions. Whereas these distributions qualify as EDMs [32], in general, they do not possess finite second moments, and thus they would not express Taylor's law.

On the basis of all these arguments, then, it is reasonable to conclude that virtually any statistical model or simulation, designed to produce Taylor's law, would be mathematically required to converge to a Tweedie distribution, as a consequence of the Tweedie convergence theorem.

The Tweedie convergence theorem provides, in essence, a connection between the CLT (where the power variance exponent is $p = 0$) and the Poisson convergence theorem (where $p = 1$). There then follows a continuous scale connection between these two extremes where a great variety of distributions exist: We have the compound Poisson and gamma distributions, the inverse Gaussian distribution, as well as EDMs generated by positive and extreme stable distributions. On the basis of this convergence theorem we are therefore led to propose the universality of the Tweedie exponential dispersion models for processes which express Taylor's law and $1/f$ noise.

VI. DISCUSSION

It is important to recognize that there are many different facets to the variance to mean power function, which have come under the rubric of Taylor's law. Originally Taylor's law was used to assess spatial aggregation of members of a species within their habitats. In these assessments a region of habitat would be divided into rectangular quadrats, and within each quadrat multiple samples would be obtained to enumerate the number of individuals of the species of interest. Sampling methods might vary for the particular species: Multiple traps might be laid within each quadrat for more mobile species, or insect-infested plants from each quadrat might be sampled. For each quadrat the mean and variance of the number of individuals per sample would be obtained and the variances

and means from the different quadrats employed on a log-log plot to assess for Taylor's law.

The method of expanding bins employed herein to assess sequential data reveals a different facet of Taylor's law for which the derived parameters are distinct from those determined by the original approach. The expanding bin method yields equations that are biconditionally related to the spectral density determination of $1/f$ noise; this biconditional relationship is not applicable to the conventional approach. Notwithstanding these differences, we know that field data used to demonstrate Taylor's law by the original method may also demonstrate Taylor's law through the method of expanding bins [29].

Regardless of the method used to assess for Taylor's law, the Tweedie EDMs are applicable to both approaches. A central tenet proposed here is that the universality of Taylor's law comes from the universality of the Tweedie models that, in turn, is grounded on the limit theorem that establishes these models as foci of statistical convergence. Taylor's law manifests as an inherent and defining feature of the Tweedie models, consequent to the mathematical requirement for their closure with respect to scale transformation.

Much effort by ecologists has been directed toward explaining Taylor's law in terms of animal behavior. Taylor maintained that the balanced immigration and emigration of animals together with their competition for resources could explain his law; he presented simulations to support this hypothesis [3]. Hanski, on the other hand, used simulations to argue that Taylor's law could be explained by the multiplicative effects of reproduction [33]. Then Anderson proposed a model for this law based on Markovian demographics, also supported by simulations [4]. Other models followed, notably that of Perry who demonstrated Taylor's law through simulations based on chaotic processes [31], and Kilpatrick and Ives, who used both simulations and analytic approximation to demonstrate that competitive interactions between species could give rise to this law [6].

None of these population models could account for the nonecological manifestation of Taylor's law, as seen with cancer metastasis [30], regional organ blood flow [8], and the genomic distributions of SNPs [9] and genes [10]. Fronczak and Fronczak attempted to provide a more general explanation for Taylor's law on the basis of a thermodynamic model that employed fluctuation dissipation in the presence of an external physical field [12]. However (at the time of this writing) there exists no compelling observational evidence to support the action of an external physical field in any of the biological systems where Taylor's law has been observed [34]. Moreover, neither this thermodynamic model nor any of the *ad hoc* population models can explain the purely numerical manifestations of Taylor's law reported here with the GUE and GOE, or that with the Mertens function reported elsewhere [34].

We have demonstrated here how the GUE and GOE fluctuations $|\bar{D}_n|$ express both Taylor's law and $1/f$ noise, and how Taylor's law for sequential data implies the manifestation of this noise pattern. $1/f$ noise was first observed with electron emissions in vacuum tubes [35], and a wide range of physical [36], biological [37], and econometric [38] observations of this pattern have since followed. Whereas the term $1/f$ noise has been applied to time series data, it is not necessarily confined to such data. In analogy to time series data, Li and Holste have reported $1/f$ noise within DNA sequences [39], and Faliero *et al.* have similarly reported the presence of $1/f$ noise from RMT [40].

There have been many explanations for $1/f$ noise [41]. Bak, Tang, and Wiesenfeld (BTW) notably proposed a model for $1/f$ noise where self-organized critical states naturally evolve within extended dynamical systems that possess many degrees of freedom [42]. They suggested that $1/f$ spectra should not be considered to reflect random noise, but rather the intrinsic (and deterministic) dynamics of the system at hand [42]. However, the original BTW model was later (and correctly) shown to have a power spectrum that was $1/f^{\,2}$, indicative of one-dimensional Brownian motion [43], not $1/f$ noise. A nonconservative modification of the BTW model was then proposed to explain $1/f^{\gamma}$ noise where the frequency exponent γ was related to the degree of dissipation within the system [44]. Such dynamical explanations, however, would not seem applicable to the GUE or GOE, where $1/f$ noise manifests as a purely mathematical phenomenon. Since $1/f$ noise manifests in the spectra associated with random matrices, this would indicate that $1/f$ noise does not necessarily depend on specific details of particular systems.

Faleiro *et al.* have demonstrated $1/f$ noise from RMT [40] using the δ_n statistic,

$$\delta_n = \sum_{i=1}^{n} s_i - \langle s \rangle = e_{n+1} - e_1 - n, \quad n = 1,...,N-1, \tag{33}$$

which describes the deviation of the $(n+1)$th unfolded eigenvalue from its average value n. By means of analytic expressions for the Fourier transforms from the GUE and GOE they showed that the power spectra of the sequences δ_n would have the form $\langle P_k^{\delta} \rangle \propto 1/k$, for $k \ll N$ and $N \gg 1$ [40]. Our demonstrations of $1/f$ noise from the GUE and GOE here were based on the deviations between the actual and the expected cumulative number of eigenvalues $|\bar{D}_n|$, which relate to Faleiro's statistic through the relation $\delta_n = \bar{D}_n - \bar{D}_{n+1} + 1 - n$. The demonstrations of $1/f$ noise by Faleiro *et al.* and those provided here would thus appear related.

Much work has been done with models of chaotic quantum systems such as nuclear shell models, chaotic quantum billiards, and (as noted above) with classical RMT, which demonstrates $1/f$ spectra [45]. It has been proposed that $1/f$ noise represents an intrinsic property of chaotic quantum systems, yet the origin of this noise pattern, particularly with energy level fluctuations, has remained uncertain [45]. We propose a paradigm, alternative to the dynamical models, whereby $1/f$ noise is explained by the mathematical convergence of complex statistical systems (including chaotic many-body systems) toward the Tweedie distributions in a manner related to the CLT. The Tweedie convergence theorem thus would explain the origin of $1/f$ noise in such systems.

We would additionally conjecture that the Tweedie convergence theorem is applicable to systems where Taylor's exponent p might vary within a certain range. Since, in the case of sequential data, p relates directly to the Hurst parameter, and in turn to the fractal dimension, such a system could have multifractal properties [26]. In this context, and for the GUE

and GOE deviations, the Tweedie convergence theorem would provide insight into the genesis of multifractality.

Taylor's power law was established and confirmed through extensive empirical observations from multiple animal and plant species as well as from many other biological, physical, and econometric processes. There have been many attempts to explain Taylor's law on the basis of process-specific mechanisms, however, no particular explanation has found general acceptance; moreover, such *ad hoc* explanations would individually not be able to account for all the diverse manifestations of Taylor's law. Taylor's law and $1/f$ noise can be shown to arise mathematically through the general convergence behavior of non-Gaussian systems, where the Tweedie PG distribution acts as one of the mathematical foci for convergence. Just as the CLT justifies a primary role in biological and physical processes for the Gaussian distribution, the Tweedie convergence theorem would indicate a universal and fundamental role for the Tweedie distributions, by explaining the apparent ubiquity and generality of both Taylor's law and $1/f$ noise, within biological, numerical, and dynamical systems.

ACKNOWLEDGMENTS

The authors acknowledge scientific grants provided by Patricia Rinaldo and the Danish Natural Science Research Council, and they wish to thank Veit Schwämmle for his scientific guidance given in discussions.

[1] L. R. Taylor, Nature (London) **189**, 732 (1961).
[2] L. R. Taylor, I. P. Woiwod, and J. N. Perry, J. Anim. Ecol. **47**, 383 (1978).
[3] L. R. Taylor and R. A. Taylor, Nature (London) **265**, 415 (1977).
[4] R. Anderson, D. Gordon, M. Crawley *et al.*, Nature (London) **296**, 245 (1982).
[5] I. Hanski and I. P. Woiwod, Oikos **67**, 29 (1993).
[6] A. M. Kilpatrick and A. R. Ives, Nature (London) **422**, 65 (2003).
[7] R. M. Anderson and R. M. May, Nature (London) **333**, 514 (1988).
[8] W. S. Kendal, Proc. Natl. Acad. Sci. USA **98**, 837 (2001).
[9] W. S. Kendal, Mol. Biol. Evol. **20**, 579 (2003).
[10] W. S. Kendal, BMC Evol. Biol. **4**, 3 (2004).
[11] Z. Eisler, I. Bartos, and J. Kertesz, Adv. Phys. **57**, 89 (2008).
[12] A. Fronczak and P. Fronczak, Phys. Rev. E **81**, 066112 (2010).
[13] W. S. Kendal, J. Theor. Biol. **245**, 329 (2007).
[14] B. Jørgensen, J. R. Martinez, and M. Tsao, Scand. J. Stat. **21**, 223 (1994).
[15] M. C. K. Tweedie, in *Statistics: Applications and New Directions. Proceedings of the Indian Statistical Institute Golden Jubilee International Conference, Indian Statistical Institute, Calcutta, India, 1984*, edited by J. K. Ghosh and J. Roy (unpublished), pp. 579–604.
[16] W. E. Leland *et al.*, IEE/ACM Trans. Networking **2**, 1 (1994).
[17] B. Tsybakov and N. D. Georganas, IEEE/ACM Trans. Netw. **5**, 397 (1997).
[18] B. B. Mandelbrot and J. W. van Ness, SIAM Rev. **10**, 422 (1968).
[19] T. Timberlake, Am. J. Phys. **74**, 547 (2006).
[20] O. Bohigas, M. J. Giannoni, and C. Schmit, Phys. Rev. Lett. **52**, 1 (1984).
[21] J. M. Borwein, D. M. Bradley, and R. E. Crandall, J. Comput. Appl. Math. **121**, 247 (2000).
[22] F. Dyson and M. L. Mehta, J. Math. Phys. **4**, 701 (1963).
[23] D. A. McQuarrie, *Statistical Mechanics* (Harper & Row, New York, 1976).
[24] B. Jørgensen, *The Theory of Exponential Dispersion Models* (Chapman & Hall, London, 1997).
[25] T. Gneiting and M. Schlather, SIAM Rev. **46**, 296 (2004).
[26] H. E. Stanley and P. Meakin, Nature **335**, 405 (1988).
[27] J. F. Muzy, E. Bacry, and A. Arneodo, Phys. Rev. E **47**, 875 (1993).
[28] E. Bacry *et al.*, Ann. Appl. Probab. **20**, 1729 (2010).
[29] W. S. Kendal, Ecol. Modell. **151**, 261 (2002).
[30] W. S. Kendal, J. Theor. Biol. **217**, 203 (2002).
[31] J. N. Perry, Proc. R. Soc. London, Ser. B **257**, 221 (1994).
[32] S. K. Bar-Lev and G. Letac, Stat. Probab. Lett. **80**, 1870 (2010).
[33] I. Hanski, Oikos **34**, 293 (1980).
[34] W. S. Kendal and B. Jørgensen, Phys. Rev. E **83**, 066115 (2011).
[35] J. B. Johnson, Phys. Rev. **26**, 71 (1925).
[36] P. Dutta and P. M. Horn, Rev. Mod. Phys. **53**, 497 (1981).
[37] T. Gisiger, Biol. Rev. Cambridge Philos. Soc. **76**, 161 (2001).
[38] A. W. Lo, Econometrica **59**, 1279 (1991).
[39] W. Li and D. Holste, Phys. Rev. E **71**, 041910 (2005).
[40] E. Faleiro, J. M. G. Gómez, R. A. Molina, L. Munoz, A. Relano, and J. Retamosa, Phys. Rev. Lett. **93**, 244101 (2004).
[41] E. Milotti, *Encuentro del Grupo Latinoamericano de Emisión Acústica y 1ro. Iberoamericano*, E-GLEA-2 (2001).
[42] P. Bak, C. Tang, and K. Wiesenfeld, Phys. Rev. A **38**, 364 (1988).
[43] H. J. Jensen, K. Christensen, and H. C. Fogedby, Phys. Rev. B **40**, 7425 (1989).
[44] K. Christensen, Z. Olami, and P. Bak, Phys. Rev. Lett. **68**, 2417 (1992).
[45] J. M. G. Gómez *et al.*, Phys. Rep. **499**, 103 (2011).