Fluctuation scaling in complex systems: Taylor’s law and beyond

Zoltán Eisler\textsuperscript{ab*}, Imre Bartos\textsuperscript{ed} and János Kertész\textsuperscript{be}

\textsuperscript{a}Science & Finance, Capital Fund Management, Paris, France;\textsuperscript{b}Department of Theoretical Physics, Budapest University of Technology and Economics, Budapest, Hungary;\textsuperscript{c}Department of Physics of Complex Systems, Loránd Eötvös University, Budapest, Hungary;\textsuperscript{d}Department of Physics, Columbia University, New York, USA
\textsuperscript{e}Physics of Condensed Matter Group, HAS, BME, Budapest, Hungary

(Received 16 August 2007; final version received 28 December 2007)

Complex systems consist of many interacting elements which participate in some dynamical process. The activity of various elements is often different and the fluctuation in the activity of an element grows monotonically with the average activity. This relationship is often of the form \( \frac{\text{fluctuations}}{\text{constant}} \times \frac{\text{average}}{\text{average}} \), where the exponent \( \alpha \) is predominantly in the range \([1/2, 1]\). This power law has been observed in a very wide range of disciplines, ranging from population dynamics through the Internet to the stock market and it is often treated under the names Taylor’s law or fluctuation scaling. This review attempts to show how general the above scaling relationship is by surveying the literature, as well as by reporting some new empirical data and model calculations. We also show some basic principles that can underlie the generality of the phenomenon. This is followed by a mean-field framework based on sums of random variables. In this context the emergence of fluctuation scaling is equivalent to some corresponding limit theorems. In certain physical systems fluctuation scaling can be related to finite size scaling.

Keywords: fluctuation scaling; Taylor’s law; complex systems; scaling

Contents

1. Introduction 91
2. Fluctuation scaling 92
   2.1. Basic notions 92
      2.1.1. Temporal fluctuation scaling 93
      2.1.2. Ensemble fluctuation scaling 94
   2.2. Empirical results: ensemble averages 95
      2.2.1. Pioneering studies 95
      2.2.2. Ecology 96
      2.2.3. Life sciences 97
      2.2.4. Physics 97

*Corresponding author. Email: eisler@maxwell.phy.bme.hu
\textsuperscript{1}Dedicated to the memory of L. R. Taylor (1924–2007).
| Section                                                                 | Page |
|-----------------------------------------------------------------------|------|
| 2.3. Empirical results: temporal averages                              | 98   |
| 2.3.1. Complex networks                                               | 98   |
| 2.3.2. Ecology                                                        | 99   |
| 2.3.3. Life sciences                                                  | 100  |
| 2.3.4. Stock market                                                   | 101  |
| 2.4. New empirical results                                            | 102  |
| 2.4.1. Stock market (ensemble averaging)                              | 102  |
| 2.4.2. Human dynamics                                                | 103  |
| 2.4.3. Precipitation                                                 | 105  |
| 2.5. Corrections to FS                                                | 107  |
| 2.6. Summary of observations                                          | 108  |
| 3. A general formalism                                                | 109  |
| 3.1. The components of fluctuations                                   | 109  |
| 3.2. ‘Universality’ classes                                           | 110  |
| 3.2.1. The case $\alpha = 1/2$                                       | 110  |
| 3.2.2. The case $\alpha = 1$                                         | 111  |
| 3.3. Other values of $\alpha$                                         | 113  |
| 3.3.1. The dependence of $\alpha$ on the time resolution $\Delta t$   | 113  |
| 3.3.2. Impact inhomogeneity                                           | 114  |
| 3.3.3. Examples of impact inhomogeneity                              | 115  |
| 3.3.4. Constituent correlations                                      | 115  |
| 4. Models                                                             | 116  |
| 4.1. Random walks on complex networks                                 | 116  |
| 4.1.1. The model                                                      | 116  |
| 4.1.2. Fluctuation scaling and corrections                           | 117  |
| 4.1.3. The role of node–node interactions and a connection with surface growth | 119  |
| 4.1.4. The role of external driving                                  | 120  |
| 4.1.5. Impact inhomogeneity                                           | 120  |
| 4.2. Critical fluctuations and finite size scaling (FSS)              | 120  |
| 4.3. Scaling and multiscaling                                         | 123  |
| 4.4. Binary forest model                                              | 125  |
| 5. Discussion                                                         | 127  |
| 5.1. Separation of global and local dynamics                          | 128  |
| 5.2. Limit theorems for sums of random variables                     | 130  |
| 5.3. The connection between ensemble and temporal averages            | 131  |
| 5.4. Fluctuation scaling for growth rates                             | 132  |
| 6. Conclusions                                                        | 133  |
| Note added in proof                                                   | 134  |
| Acknowledgments                                                       | 134  |
| Notes                                                                 | 134  |
| References                                                            | 135  |
| Appendix A: The components of the fluctuation $\sigma^2$              | 137  |
| Appendix B: Tweedie models and impact inhomogeneity                   | 138  |
| Appendix C: Fluctuations in the network random walker model           | 140  |
1. Introduction

Interacting systems of many units with emergent collective behaviour are often termed ‘complex’. Such complex systems are ubiquitous in many fields of research ranging from engineering sciences through physics and biology to sociology. An advantage of the related multi-disciplinary approach is that the universal appearance of several phenomena can be revealed more easily. Such generally observed characteristics include (multi-)fractality or scale invariance [1,2], the related Pareto or Zipf laws [3,4], self-organized and critical behaviour.

In this paper we study such a general feature related to the scaling properties of the fluctuations in complex systems. This type of scaling relationship is called Taylor’s law by ecologists after L. R. Taylor and his influential paper on natural populations in 1961 [5]. The law states that for any fixed species the fluctuations in the size of a population (characterized by the standard deviation) can be approximately written as a constant times the average population to a power $\alpha$:

$$\text{fluctuations} \approx \text{constant} \times \text{average}^\alpha$$

for a wide range of the average.

The phenomenon was, to the best of our knowledge, first discovered in 1938 by H. Fairfield Smith [6], who wrote an equivalent formula for the yield of crop fields although his paper has, surprisingly, received much less attention than Taylor’s work. The same relationship was explored recently by de Menezes and Barabási [7] for dynamics on complex networks, and was later termed ‘fluctuation scaling’ (FS) [8] in the physics literature. There, the temporal fluctuations and the averages of the network’s traffic were measured at the different nodes.

Despite the analogous questions, the exchange of ideas between disciplines is very limited. This review attempts to show how general FS is by surveying the literature and the current models, as well as by reporting some new empirical evidence and presenting new model calculations. We also have the aim of stepping beyond mere demonstration and show some basic principles that can potentially underlie the generality of the phenomenon.

This paper is organized as follows. In Section 2 we give a more precise definition of FS. We then give a brief overview of empirical results from the literature, and also some previously unpublished findings.¹ In Section 3 we present a general mean-field formalism based on sums of random variables. This is followed by the interpretation of the scaling exponent $\alpha$, and how it reflects the dynamics of the complex system. We show that $\alpha$ is usually between $1/2$ and $1$, and that these two limiting values can both arise from several, simple types of dynamics. We then describe three scenarios of how intermediate exponents can arise. Owing to this multitude of possibilities, no value of $\alpha$ can be used in itself to uniquely identify the internal dynamics of a system, but it is still possible to exclude many options that would be incompatible with the observed value of $\alpha$. The remaining possibilities can be narrowed down further by analysing the time-scale dependence of $\alpha$ and by the application of our mean field framework. The procedure is demonstrated on some simple models in Section 4, and the relationship between FS, (self-organized) criticality, scaling and multiscaling is explored. Fluctuation scaling can be directly applied to certain physical systems, where one finds a strong connection with finite-size
scaling (FSS). In Section 5 we give a general discussion and directions for future research. Finally, we provide conclusions in Section 6. Some calculations have been deferred to the appendices.

2. Fluctuation scaling

2.1. Basic notions

Throughout the paper we always consider some additive quantity \( f \), and the dependence between its mean and standard deviation. By dependence we mean the behaviour of \( f \) over a multitude of observations. Say, if we can observe the same dynamical variable in several settings where it has different means, how does the standard deviation change with the value of the mean?

In order to determine this dependence one needs many realizations. These can be simultaneous temporal observations for different elements (nodes, subsystems) of a large complex system. The measured means and standard deviations are then calculated in time, and the subsystems are compared: for subsystems with a larger mean \( f \) are the fluctuations larger as well?

In other cases \( f \) is not considered as time dependent, only as a fixed value for every subsystem. Then the averages are taken over an ensemble of subsystems of equal size and the standard deviation characterizes the variation of \( f \) between subsystems of the same size.

We just used the expressions ‘elements’, ‘nodes’, ‘subsystems’ and ‘the same size’, but what do these mean? Imagine that we want to quantify fluctuations in the traffic of Internet routers. It is very straightforward to calculate the mean and the standard deviation of, say, daily data throughput, and the question of whether routers with larger mean amounts of traffic exhibit larger fluctuations can be investigated. However, routers are not ‘subsystems’ of the Internet. They merely represent points of measurement, elements of the system. The traffic is formed as a sum of data packets that are ‘extrinsic’ to the elements. The packets do not belong to the structure of the network, but they carry the dynamics on it. Here we are not interested in the structure of the routers, i.e. the nodes. Instead, the data over which the averages are taken have a temporal structure.

Let us take a different example. Now we want to analyse data on the size of animal populations. A population can be divided into smaller groups, which then consist of individuals, and this gives a true notion of size. Various smaller areas can be naively considered as ‘subsystems’ with respect to the habitat of the species, for example, a continent. These subsystems are not structureless and their population comes about as a sum over their smaller subgroups of individuals.

In this review we call the points of measurement nodes, whether they have a structure or not. The additive quantity under study, be it activity, population, traffic or whatever else, is denoted by \( f_i \), where \( i \) indicates the node of measurement. This will always be decomposed as a sum of random variables, which either represent internal constituents or some events similar to the arrival of the extrinsic ‘packets’ to the Internet routers. In both cases we call these the constituents of the nodes/signals. Their number for node \( i \) will be denoted by \( N_i \), and their respective contributions to \( f_i \) will be denoted by \( V_{i,n} \), where \( n = 1, \ldots, N_i \). Examples of the scheme for building up a system from constituents can be seen in Table 1. Now we turn to more precise definitions.
2.1.1. Temporal fluctuation scaling

Let us assume that during an extended period we can measure an additive quantity \( f_i \) within a system at its nodes (labelled by the index \( i \)). For some finite time duration \([t, t + \Delta t]\) the signal can be formally decomposed as the sum

\[
 f_i^{\Delta t}(t) = \sum_{n=1}^{N_i^{\Delta t}(t)} V_{i,n}^{\Delta t}(t). \tag{1}
\]

Here \( N_i^{\Delta t}(t) \) is the number of constituents of node \( i \) during \([t, t + \Delta t]\). We assume that \( V_{i,n}^{\Delta t}(t) \geq 0 \), so that the time average of \( f_i^{\Delta t} \) does not vanish. For example, if on the stock market during \([t, t + \Delta t]\) there are \( N_i^{\Delta t}(t) \) transactions with the papers of the \( i \)th company, and the \( n \)th of those transactions has a value \( V_{i,n}^{\Delta t}(t) \), then the total trading activity of stock \( i \) can be calculated by this formula.

The time average of (1), which we denote as \( \langle f_i^{\Delta t} \rangle \), can be calculated as

\[
 \langle f_i^{\Delta t} \rangle = \frac{1}{Q} \sum_{q=0}^{Q-1} f_i^{\Delta t}(q\Delta t) = \frac{1}{Q} \sum_{q=0}^{Q-1} \sum_{n=1}^{N_i^{\Delta t}(q\Delta t)} V_{i,n}^{\Delta t}(q\Delta t), \tag{2}
\]

where \( Q = T/\Delta t \) and \( T \) is the total time of measurement. From the definitions it is trivial that \( \langle f_i^{\Delta t} \rangle = \Delta t \langle f_i^{\Delta t=1} \rangle \). We use \( \langle f_i \rangle \) without the upper index to denote this latter quantity.

On any time scale the variance can be obtained as a time average:

\[
 \sigma_i^2(\Delta t) = \langle [f_i^{\Delta t}]^2 \rangle - \langle f_i^{\Delta t} \rangle^2,
\]

this quantity characterizes the fluctuations of the activity of a fixed node \( i \) from interval to interval.

When \( f \) is positive and additive, it is often observed that the relationship between the standard deviation and the mean of \( f \) is given by a power law:

\[
 \sigma_i(\Delta t) \propto [\langle f_i^{\Delta t} \rangle]^{\alpha_f},
\]

where one varies the node \( i \), and \( \Delta t \) is fixed. The dependence of the right-hand side on \( \Delta t \) is trivial, because \( \langle f_i^{\Delta t} \rangle \equiv \Delta t \langle f_i \rangle \). Thus, throughout the paper we use \( \langle f_i \rangle \) as the scaling variable:

\[
 \sigma_i(\Delta t) \propto [\langle f_i \rangle]^{\alpha_f}. \tag{3}
\]
The exponent $\alpha_T$ is usually in the range $[1/2, 1]$. The lower index $T$ in the scaling exponent indicates that the statistical quantities are defined as *temporal* averages as in (2).

Finally, if the $i$-dependence of $\sigma$ and $\langle f \rangle$ is only manifested via a well-defined parameter of the nodes, such as their linear extent ($L$), area ($A$), a fixed number of constituents ($N$) or some other size-like parameter $S$, then we use this quantity as a lower index where possible. For example, temporal standard deviation be denoted by $\sigma_S$.

2.1.2. Ensemble fluctuation scaling

Again imagine that nodes have a well-defined size-like parameter $S$, and it is possible to group them according to that. Furthermore, assume that nodes that fall into the same group have equivalent statistical properties. Then aside from the temporal average given separately for each node, one can also define the average of $f$ within each group. This is a sort of ensemble average over similar nodes: it is denoted by $\langle f \rangle_S$ and can be calculated as

$$f^S_S = \frac{1}{M_S} \sum_{i \in S_S} f_i^{\Delta t}(t).$$

Both $t$ and $\Delta t$ are now fixed; the summation instead goes for those nodes $i$ which have a size $S_i = S$, and $M_S$ is the number of such nodes. In the notation we omit $t$ for simplicity. Variance is given by

$$\overline{\sigma^2}_S(\Delta t) = \langle f^S_S \rangle^2 - \langle f^S_S \rangle^2.$$ 

Fluctuation scaling can also arise here in the form

$$\overline{\sigma}_S(\Delta t) \propto \frac{f^S_S}{S},$$

where we compare different groups by varying $S$, while $\Delta t$ is kept constant. For convenience, we follow the convention of the previous section: on the right-hand side of (5) we write $f_S$, which is short-hand notation for $\langle f^S_S \rangle = 1$. The scaling exponent $\alpha_E$ will always indicate when we use *ensemble* averaging over elements of the same size.

For data analysis the size $S$ very often corresponds to the linear size $L$ or the area $A$ of the node/subsystem and there the lower index is altered accordingly. For example, the classic study of Taylor [5] compares areas of different size $A$, and the measured quantity is the population size of a given species in the area. The constituents can be smaller groups or, as they are usually called, metapopulations. If one considers the number of groups ($N$) and the size of the groups ($V_n$) as random variables, the total population has the same sum form as before:

$$f_i = \sum_{n=1}^{N_i} V_{i,n},$$

which is the analogue of (1).

We call the relationship (3) temporal fluctuation scaling (TFS) and the relationship (5) ensemble fluctuation scaling (EFS). When we do not wish to distinguish between the two cases, we simply use FS and then the exponent is denoted by $\alpha$ without a lower index. There exists a large body of results on these subjects, and the literature is spread over
many disciplines. Therefore, in the following we give a (necessarily incomplete) overview of the results. A summary is presented in Table 2.

### 2.2. Empirical results: ensemble averages

#### 2.2.1. Pioneering studies

As noted in the introduction, the first observations of FS appeared in two independent studies, well before the widespread recognition of fractality and scaling [9]. The paper of Fairfield Smith [6] was published in 1938, and it was concerned with the yields of agricultural crops. For a fixed size of land \( A \) it is possible to calculate the average yield \( f_A \) of a certain type of crop, and the standard deviation \( \sigma_A \) of the yield between areas of size \( A \). Then the calculation can be done for areas of many different sizes. It was found

| Subject            | System                        | Temporal or Ensemble | References      |
|--------------------|-------------------------------|----------------------|-----------------|
| Networks           | Random walk                   | T                    | [7,31,33]       |
|                   | Network models                | T                    | [34,35]         |
|                   | Highway network               | T                    | [7,31]          |
|                   | World Wide Web                | T                    | [7,31]          |
|                   | Internet                      | T                    | [7,31,32]       |
| Physics            | Heavy-ion collisions          | E                    | [26–28]         |
|                   | Cosmic rays                   | E                    | [29,30]         |
| Social sciences/  | Stock market                  | T                    | [8,55,56,59]    |
| Economics          | Stock market                  | E                    | This review     |
|                    | Business firm growth rates    | E                    | [61,62]         |
|                    | Email traffic                 | T                    | This review     |
|                    | Printing activity             | T                    | This review     |
| Climatology        | River flow                    | T                    | [63,64]         |
|                    | Precipitation                 | T                    | [65]            |
| Ecology/           | Forest reproductive rates     | T                    | [46,47]         |
| population dynamics| Satake–Iwasa forest model     | T                    | [45]            |
|                    | Crop yield                    | T                    | [6]             |
|                    | Animal populations            | T, E                 | [5,10,15,16]    |
|                    | Diffusion Limited population  | E                    | [17]            |
|                    | Population growth             | T                    | [66,67]         |
|                    | Exponential dispersion models  | E                    | [18,21,68]      |
|                    | Interacting population model   | T                    | [37]            |
| Life sciences      | Cell numbers                  | E                    | [20]            |
|                    | Protein expression            | T                    | [54]            |
|                    | Gene expression               | T                    | [69,70]         |
|                    | Individual health             | E                    | [71]            |
|                    | Tumor cells                   | E                    | [21]            |
|                    | Human genome                  | E                    | [22,23]         |
|                    | Blood flow                    | E                    | [68]            |
|                    | Oncology                      | E                    | [21]            |
|                    | Epidemiology                  | T                    | [52,53]         |
that there exists the power law (5) relationship between the two quantities, $E \propto A^{\alpha}$, with $E \approx 0.62$.

Taylor’s 1961 paper [5] stated the scaling law (5) for systems in population dynamics. Similarly to Fairfield Smith, Taylor took an ensemble of areas of the same size, and measured the number of individuals of a certain type of animal. With increasing area size, both the mean and the variance of the population grew, with a power law relationship between the two quantities, see Figure 1. Let us now take a closer look at fluctuations in ecology.

2.2.2. Ecology

Stable populations in a given habitat fluctuate around a typical size called the habitat’s carrying capacity [10–12]. These fluctuations have a very rich internal structure [12]. Both the randomness of birth–death processes (a kind of ‘intrinsic noise’) and external climatic forcing play important roles [12,13]. The effect of climatic factors is so strong that it can synchronize the fluctuations of even non-interacting populations (the so-called Moran effect) [11,14]. To further complicate the situation, individuals of a species interact among themselves, just as well as species interact with each other. These interactions are non-linear and nowadays they are also commonly recognized to have a significant dependence on the population density/size. So interactions, driving and noise all contribute to population dynamics to a certain degree [12,13]. This diversity makes any ecosystem a showcase of complexity; certain regularities are known, but the bigger picture is still missing.

This is the reason why by discovering a universal law, Taylor’s paper [5] triggered growing activity in ecology, with literally 1000 publications to date. Taylor’s results were verified for a wide range of populations, and the value of the exponent was predominantly found to be $1/2 \leq \alpha_E \leq 1$ (see [15]). Despite its generality, the origins of the law and the meaning of $\alpha$ are still much debated. Anderson et al. [16] suggested that the influence of
environmental fluctuations may be responsible for the observed non-trivial exponents. The model of Kendal [17] proposed dynamics similar to diffusion limited aggregation in which self-similarity gives rise to the mean–variance scaling. Another study [18] proposed that the exponents can be described by a class of statistical models, which rely on the interplay between the number of animal clusters in an area and the size of the individual clusters. We discuss these models in detail in Section 3.3.3. For two comprehensive reviews of these (and more) scaling laws in ecological and related systems see Kendal [18] and Marquet et al. [19].

2.2.3. Life sciences

There is a number of findings from cellular and molecular biology regarding FS. Azevedo and Leroi [20] conducted a very extensive study of how the typical cell count of a species is related to its fluctuations from individual to individual. They found that FS holds over almost 10 orders of magnitude in size, between more than 2000 species, see Figure 2. The exponent $\alpha_E$ differs among tissue types, but for entire organisms its value is approximately 1. Kendal [21] presents similar findings for the number of tumor cells in groups of mice, but the exponents vary.

Similarly, Kendal [22] analysed the most common variations in the human genome called single nucleotide polymorphisms (SNPs) [23]. He found that the mean and the variance of the number of SNPs in a DNA sequence scale as different non-trivial powers of the length of the sequence, and thus the variance also scales with the mean.

2.2.4. Physics

FS is also present in the physics literature. Many extensive quantities are known to have equilibrium fluctuations proportional to the square root of the system size, implying the
value $\alpha_E = 1/2$ (see [24,25]). This relationship is a simple consequence of the central limit theorem (see Section 3.2.1). Botet et al. [26–28] found EFS for a wide range of models, and also for the fragment multiplicity measured in heavy-ion collision experiments. Moreover, a linear ($\alpha = 1$) relationship was found between the fluctuations and mean fluxes of cosmic radiation [29,30]. Here the ensemble is formed by cutting a single time series into pieces, and then periods with higher average activity exhibit higher fluctuations.

### 2.3. Empirical results: temporal averages

We now turn to TFS. For the collection of such data it is necessary to have multi-channel measurements, simultaneously monitoring the behaviour of a range of elements $i$. With the unbroken growth of computing infrastructure, many technological networks now offer appropriate datasets, several of which are publicly available.

#### 2.3.1. Complex networks

de Menezes and Barabási [7,31], in part inspired by Taylor’s original paper, found TFS for several complex networks. A good example is the analysis of Internet traffic, which was later revisited by Duch and Arenas [32]. In their study they analysed the traffic of the Abilene backbone network. The nodes $i$ correspond to routers, and the mean and variance of their data flow was calculated. In Figure 3 we show their results for weekly data traffic; the best fit is achieved with $\alpha_T \approx 0.75$. de Menezes and Barabási also analysed web page visitations, river flow, microchip logical gates and highway traffic. They proposed that the datasets should fall into two ‘universality classes’ with $\alpha_T = 1/2$ and 1. There also exists a growing body of literature on transport processes on networks, and the scaling of fluctuations in such systems [32–36].
2.3.2. Ecology

Ecologists have made many advances regarding TFS as well, but the literature is far from unequivocal. The basic concept is to monitor many populations of a given species for an extended period of time, and then for each population $i$ calculate the temporal mean $\langle f_i \rangle$ and standard deviation $\sigma_i$ of abundance. These are typically power law related according to TFS; examples are shown in Figure 4.

Classical population dynamics offers several benchmark models [39–41], but simple deterministic and Markovian models cannot explain the observed $\alpha_T$ values between $1/2$ and 1. After a range of small populations where they show realistic behaviour, they cross over to either $\alpha_T = 1/2$ or 1 [42]. The model of Kilpatrick and Ives [37] suggested that the interaction between species and feedback mechanisms between their fluctuations can give rise to any value of $\alpha_T$. Perry proposed an even simpler chaotic model [43]. Both of these models can yield various exponents, but still only when populations are small enough.

There have been several findings for plant species. In a series of papers Ballantyne and Kerkhoff showed that the reproductive (yearly seed count) variability of trees follows TFS with $\alpha_T = 1$. The same value is supported by the Satake–Iwasa [44] forest model. There the trees are modelled using interacting oscillators which synchronize above a critical value of the coupling [45,46]. The synchronization transition coincides with a transition\(^2\) from $\alpha_T = 1/2$ to 1.

We now briefly describe their empirical study [47]. The dataset [48] consists of yearly observations of the seed production of trees throughout the Northern Hemisphere. In particular, we considered three subsets of the dataset (data available from the author upon request), those collected by Tallqvist [49], Franklin [50] and Weaver and Forcella [51], including 4–17 years of observations for 44, 148 and 28 sites, respectively. The fits for TFS are given in Figure 5(a). The exponents for the three subsets were found to be $\alpha_T = 0.97, 0.93$ and 0.90. Given the quality of the fits it is not
possible to rule out that for all three datasets $\alpha_T = 1$ (as suggested by [47]). However, here we make an attempt to give an argument that predicts otherwise and can be tested.

Simulations of the Satake–Iwasa model already suggested that long-range synchronization can cause $\alpha_T > 1/2$, and the presence of such a tendency is well known for trees. Koenig and Knops [48] conclude that there exists a significant positive correlation between the reproductive activity of trees for distances longer than 1000 km (this phenomenon is called masting in the ecology literature). While [48] is much more precise and detailed, we also outline a simple measurement: in Figure 5(b) we plot the average $C(r)$ cross-correlation coefficients between the sites in the complete dataset as a function of the distance $r$ of the sites. As expected, we find that cross-correlations decay very slowly with distance, and the dependence can be fitted approximately by

$$C(r) \propto r^{-0.40},$$

(although admittedly the fit is not perfect). In Sections 3.3.4 and 4.4 we show that while perfect synchronization à la Satake–Iwasa leads to $\alpha_T = 1$, partial synchronization with the above power-law correlations implies $\alpha_T = 1 - 0.40/2 = 0.8$: see (20). A better quantitative agreement would warrant larger datasets which are not available at present, but there have been some promising attempts along the same lines [46].

2.3.3. Life sciences

Keeling and Grenfell [52] suggested TFS for the size of epidemics, and found both empirically and by a simple Markov chain model of population dynamics that vaccination in general decreases not only the size of epidemics but also the value of $\alpha_T$. TFS was later found by Woolhouse et al. [53] to also hold between different pathogens. TFS has been
found in the cell-to-cell variation of protein transcription by Bar-Even et al. [54], albeit with a crossover and a rather narrow range.

2.3.4. Stock market

In this section we summarize the results of a series of papers by Eisler and Kertész [8,55–57]. The work was based on a TAQ database [58], recording the transactions of the New York Stock Exchange (NYSE) for the years 2000–2002. Very similar results were obtained for the NASDAQ [8] and Chinese markets [59].

We define the activity of stock $i$ as its total traded value, given as

$$f_i^{\Delta t}(t) = \sum_{n=1}^{N_i^{\Delta t}(t)} V_{i,n}^{\Delta t}(t),$$

where $N_i^{\Delta t}(t)$ is the number of transactions of stock $i$ in the period $[t, t + \Delta t)$. The individual values of these transactions are denoted by $V_{i,n}^{\Delta t}(t)$. Data were detrended by the well-known $U$-shaped daily pattern of traded volumes [55].

Then the measurement of mean and variance was carried out. The exponent $\alpha_T$ shows a strong dependence on the window size $\Delta t$, we return to this result in Section 3.3.1. The values range between $\alpha_T = 0.68–0.87$, see Figure 6.

When $\Delta t$ is very small, [8] shows that the individual transactions can be treated as independent events. Moreover, for the stocks that are large enough the average size of transactions ($\langle V_i \rangle$) can be calculated as a power of the mean number of transactions as

$$\langle V_i \rangle \propto \langle N_i \rangle^\beta.$$
with $\beta \approx 0.65$ (see [8]). Equivalents of this property recur in several FS-related contexts. We devote Sections 3.3.2 and 3.3.3 to this observation, which we call impact inhomogeneity [8]. We also show how to map the value of $\beta$ onto non-trivial $\alpha_T$ values. By that method (see (18)) the corresponding $\alpha_T$ value should be 0.70, which is very close to the actual value $\alpha_T(\Delta t \to 0) = 0.69$.

Another general observation [8] is that if $\alpha_T$ is a function of $\Delta t$, then FS can only hold if this dependence is logarithmic (cf. Section 3.3.1). For the stock market this is true in two distinct regimes and those are separated by a crossover. For $\Delta t < 10^3$ s (−sign) and $\Delta t > 3 \times 10^4$ s (+sign) one finds

$$\alpha_{T,\pm}(\Delta t) = \alpha_{T,\pm}^s + \gamma_\pm \log \Delta t,$$

with $\gamma_- \approx 0.00$ and $\gamma_+ \approx 0.06$. On the other hand, the Hurst exponent $H_i$ can be defined as [1,60]

$$\sigma_i(\Delta t) = \left\{ \left[\langle f_i^{\Delta t}(t) \rangle - \langle f_i^{\Delta t}(0) \rangle \right]^2 \right\}^{1/2} \propto \Delta t^{H_i}. \tag{7}$$

For NYSE this equation is found to be valid with

$$H_{i,\pm} = H_{i,\pm}^* + \gamma_\pm \log \langle f_i \rangle.$$

Lower indices indicate the same two regimes, and $\gamma_\pm$ have the same values as for $\alpha_T$ (see [8]).

2.4. New empirical results

In this section we present previously unpublished results for FS. Note that temporal variances were estimated by the partition function of detrended fluctuation analysis [60]. This was necessary in order to (at least partially) remove the non-stationarity from the datasets. All results, including the values of $\alpha_T$, agree qualitatively with those obtained from a direct calculation of variance without detrending, but the accuracy of the estimation is improved.

2.4.1. Stock market (ensemble averaging)

Fluctuation scaling in stock market data has just been discussed, but those earlier results pertained TFS, whereas here we present some new findings on EFS in the same dataset.

We again consider the (daily) trading activity of stocks. The size of companies is often measured by the total value of all of their issued stocks, called the company’s capitalization (C). We take a fixed time period, the day 3 January 2000 (the results are similar for other days). Then we group the stocks according to their capitalization into 35 logarithmic bins. Finally, we calculate the mean $\hat{f}_C$ and the standard deviation $\overline{\sigma_C}$ of the activity in every group. EFS is shown in Figure 7. The fit gives $\alpha_E \approx 0.8–0.9$, although with some deviations from scaling.

In the case of FS in ecology, if the size of the animal population is an extensive quantity, it is justifiable to use the size of the area to parametrize the ensemble: its mean should be exactly proportional to the area. To use capitalization as a parametrization for company size in markets is a different matter. While it is indeed strongly related to the
mean trading activity, there is no one-to-one correspondence between the two [56]. This means that companies of the same capitalization can have different expected trading activities. Thus, our ensemble averaging technique is only approximate.

To circumvent this problem one can apply the following trick. The trading activity of companies fluctuates strongly from day to day, but the expectation value of the distribution is rather stable over time. So let us now take an interval \( t = 1, \ldots, T \) and calculate the time averages \( \langle f_i(t) \rangle \) during this period for each stock. Then, for every stock, take the single value \( f_i(t = 1) \) only, and group the observations according to \( \langle f_i(t) \rangle \). This is equivalent to the measurement before, only instead of \( C_i \), groups are formed with respect to \( \langle f_i(t) \rangle \) (15 logarithmic bins). Then the ensemble mean \( \langle f \rangle \) and variance \( \sigma \) can be calculated in each group. The results from this technique are also indicated in Figure 7; one can see a dramatic decrease in the noise level, while the value of the exponent is approximately preserved, \( \alpha_E = 0.89 \).

It must be emphasized that we use information about the temporal average only for the grouping procedure. The measured \( \alpha_E \) is a true EFS exponent. Moreover, the value of \( \alpha_E \) that we find here is much larger than \( \alpha_T \) (cf. Section 2.3.4). Owing to the intricate statistical properties of markets, the two exponents cannot be expected to coincide.\(^3\)

This example shows that the success of EFS crucially depends on the proper choice of the size parameter. Where possible, physical size or area are good choices, because they are known to be extensive. Otherwise the above trick can be applied, but only if multiple observations are available for every node and the system is close to stationary.

2.4.2. Human dynamics

The analysis of the records of human dynamics has recently seen growing interest [73–75]. Here we discuss two large technological databases of human activity:

(i) Emails from the employees of the company Enron during the year 2000. We used a filtered variant of the original dataset posted by the Federal Energy Regulatory Authority.
We defined \( f_i(t) \) as the number of emails sent by the person \( i \) during the interval \( [t, t + \Delta t] \).

(ii) Data on the printing activity of the largest printer at the Department of Computing at Imperial College London. The files include the complete year 2003; we removed weekends, official holidays and closure times of the computer laboratory (23:00–7:00). We included 987 users who submitted at least three documents during our analysis, except the single largest user who appeared to have different statistical properties from the rest. \( f_i(t) \) is defined as the number of documents submitted to print by user \( i \) in the time interval \( [t, t + \Delta t] \). Further details on the dataset can be found in [76].

Note that multiple copies of the same email/document sent/submitted simultaneously were counted as one item.

We present these two datasets side by side, because they have strong similarities. Both show TFS for window sizes \( \Delta t = 5, \ldots, 2.8 \times 10^4 \) s. The exponent varies between \( \alpha_T = 0.52–0.72 \) (email) and \( \alpha_T = 0.57–0.83 \) (printing). Figure 8 shows the fits for time window sizes \( \Delta t = 10, 10,000 \) s. We find scaling over about 2.5 orders of magnitude, and the exponent depends on \( \Delta t \), as shown in Figure 9. Despite the level of the noise in the data, the dependence appears to be monotonically increasing, with two regimes separated by a crossover near \( \Delta t \sim 4000 \) s (email) and \( \Delta t \sim 1000 \) s (printing). The dependence is close to logarithmic, with the same form as for stock markets (the index − corresponds to the regime below the crossover and + above the crossover):

\[
\sigma_{T,\pm}(\Delta t) = \sigma_{T,\pm}^* + \gamma_{\pm} \log \Delta t,
\]

with \( \gamma_{\text{email}} \approx 0.04, \gamma_{\text{print}} \approx 0.13 \), \( \gamma_{\text{print}} \approx 0.09 \), and \( \gamma_{\text{print}} \approx 0.02 \). Also for the Hurst exponents

\[
H_{i,\pm} = H_{i,\pm}^* + \gamma_{\pm} \log \langle f_i \rangle.
\]

The coefficients are \( \gamma_{\text{email}} \approx 0.04, \gamma_{\text{email}} \approx 0.11, \gamma_{\text{print}} \approx 0.07 \) and \( \gamma_{\text{print}} \approx 0.01 \). The \( \sigma(\Delta t) \) scaling plots are shown in Figure 10, and the Hurst exponents’ dependence on \( \langle f \rangle \) in Figure 11.

![Figure 8. Fluctuation scaling for time resolutions of \( \Delta t = 10 \) s and \( \Delta t = 10,000 \) s. Points were logarithmically binned and \( \log \sigma \) was averaged for better visibility, the error bars represent the standard deviations inside the bins. (a) Results for the number of sent emails; (b) Results for number of printed documents.](image-url)
Finally, note that for email data $\alpha_T(\Delta t)$ tends to $1/2$ with decreasing window size, and the logarithmic tendency appears to saturate. On the other hand, for printing data the logarithmic tendency is markedly present even for short times. By an extrapolation from the trend, one expects $\alpha_T(1\,\text{s}) \approx 0.51$. Section 3.2.1 offers an explanation why for very short times one expects $\alpha_T = 1/2$ in these datasets.

2.4.3. Precipitation

In this section we present a study [65] of the weekly precipitation records of 22,928 weather stations worldwide. The dataset was obtained from the Global Daily Climatology Network (GDCN). For one station, typically 40 years of data are available between 1950 and 1990. TFS is found with $\alpha_T = 0.77$; see Figure 12. However, there are also some significant deviations. These can be interpreted based on geographical information: for every station, the geographical latitude ($l_i$, measured in degrees) and the height $h_i$ measured from sea level was known, and the multiple regression

$$\log \sigma_i = C + \alpha_T \log |f_i| + C_l l_i + C_h h_i + \epsilon_i$$

(8)

with an error term $\epsilon_i$ yields the results $C = 0.896 \pm 0.002$, $\alpha_T = 0.732 \pm 0.002$, $C_l = -(8.79 \pm 0.05) \times 10^{-3}$ and $C_h = (-6 \pm 1) \times 10^{-6}$. All values are significantly different from zero at the 99.98% confidence level. For the single parameter fit of TFS, $R^2 = 0.73$, while for the multiple regression (8) one finds $R^2 = 0.90$ which is a substantial improvement owing to the inclusion of geographical latitude and, in smaller part, to the height above sea level. The remaining error term, although we did not find any appropriate explanatory variable, is still not unsystematic. By plotting $\epsilon_i$ on a map (see Figure 13) one finds a strong geographical clustering with $\epsilon_i > 0$ typically, but not exclusively, in continental areas. This systematic tendency suggests that a well-defined origin might exist for such corrections.

![Figure 9](image_url)

Figure 9. (a) The dependence of the FS exponent in the Enron email database on the size of the time window $\Delta t$. The dependence is logarithmic in two regimes, with the coefficients $\gamma_- \approx 0.04$ for $\Delta t < 4000\,\text{s}$ and $\gamma_+ \approx 0.13$ for $\Delta t > 4000\,\text{s}$; (b) The dependence of the FS exponent in printing data on the size of the time window $\Delta t$. The dependence is logarithmic in two regimes, with the coefficients $\gamma_- \approx 0.09$ for $\Delta t < 300\,\text{sec}$ and $\gamma_+ \approx 0.02$ for $\Delta t > 1000\,\text{s}$. 

Advances in Physics
Figure 11. The dependence of the Hurst exponent of \( f \) on \( h_f \). Points were logarithmically binned and \( H \) was averaged for better visibility, the error bars represent the standard deviations inside the bins. (a) Results for Enron email data. When \( \Delta t \approx 4000 \) s, correlations are weak and their strength increases slowly with greater \( f \). Then, after a crossover regime, for \( \Delta t > 4000 \) s correlations become stronger, with larger difference between the three groups; (b) Results for printing data. For shorter time windows \( \Delta t < 300 \) s, positive correlations exist and their strength increases with greater \( f \). Then after a crossover regime, for \( \Delta t > 1000 \) s correlations become very weak for all three groups.

Figure 10. Scaling plots of \( \log \sigma - 1/2 \times \log \Delta t \) versus \( \log \Delta t \) generated by Detrended Fluctuation Analysis. Users were grouped by their average activity into three groups (\( f \) increasing from bottom to top; see the plot for ranges) and the curves were averaged within groups. A horizontal line would correspond to a complete absence of correlations, and the slopes of the linear regimes are \( H = 0.5 \), where \( H \) are the typical Hurst exponents of groups. (a) Results for Enron email data. For shorter time windows \( \Delta t < 4000 \) s, correlations are weak and their strength increases slowly with greater \( f \). Then, after a crossover regime, for \( \Delta t > 4000 \) s correlations become stronger, with larger difference between the three groups; (b) Results for printing data. For shorter time windows \( \Delta t < 300 \) s, positive correlations exist and their strength increases with greater \( f \). Then after a crossover regime, for \( \Delta t > 1000 \) s correlations become very weak for all three groups.
As for the value of $\alpha_T = 0.77$. Points were logarithmically binned and $\log \sigma$ was averaged for better visibility, the error bars represent the standard deviations inside the bins. The inset shows the same plot and the same axis range, but without binning. One can see that there is a high number of outliers, owing to other, $(f)$-independent corrections to FS.

As for the value of $\alpha_T$, its origin will be analysed in a later study [65]. From preliminary studies it appears that it is not strongly dependent on the choice of the time scale $\Delta t$, and it is always significantly different from both $1/2$ and $1$.

2.5. Corrections to FS

Similarly to precipitation data, there are also significant corrections to TFS in stock markets. These are related to differences between market sectors. In Figure 14 we plot...
The data points were not binned or altered in any way, which makes visible the deviations from the original scaling law, which would correspond to a horizontal line. The lighter points indicate all stocks, while boxes (□) highlight the distribution of points for the three indicated economic sectors. There is a great degree of clustering both horizontally and vertically. Clustering along the log $\langle f \rangle$ axis only suggests that the sector has some typical trading activity. Systematic corrections to FS are indicated by clustering along the log $\sigma_\epsilon/(\langle f \rangle)^{\gamma_{\epsilon}}$ axis. If a sector is clustered in the lower/higher half of the dataset, it means that its trading activity has typically lower/higher fluctuations than the market average. The presence of such sector dependent clustering suggests that the corrections to FS are not purely random.

$\sigma_\epsilon/(\langle f \rangle)^{\gamma_{\epsilon}}$ versus $\langle f \rangle$ to characterize corrections to FS. By highlighting the alignment of three industrial sectors one can see that they form clusters, so the deviations are, to some degree, systematic.

In most real systems FS is found rather as a general tendency than an exact law. The scaling plots have some broadening, which can have several origins. A possible origin of poor fits can be the presence of crossovers and breaks in the scaling plots [7,16,42,46], although these can only be seen clearly in very few studies [26,63]. In other cases the deviations are attributed to the quality of data and short sampling intervals [78]. Still, these corrections can be large and systematic. In Section 2.4.3 we showed that for precipitation fluctuations geographical location and height plays a role. Similarly, for the stock market the market sector matters. These effects could be uncovered, because of the availability of these independent quantities for each station/stock. In some models they can even be calculated analytically, see Section 4.1.2.

The fact that scaling is mostly very well preserved suggests that the investigated complex systems have a robust dynamics characterized by a value of $\alpha$, and that the role of the corrections is not substantial in the formation of fluctuations. There is wide consensus that the exponents are meaningful, and not significantly distorted by the non-scaling corrections.

2.6. Summary of observations

In summary, FS appears to be a surprisingly general concept that can be recognized in virtually any discipline where the proper data are available. The fluctuations of positive additive quantities appear to have the structure

\[
\sigma_\epsilon/(\langle f \rangle)^{\gamma_{\epsilon}} = \text{constant} \times \text{average}^\alpha \times (1 + \text{corrections}).
\]

There is immense literature on the origins of fluctuations in various systems, ranging from gene networks [79] through complexity [31] to animal populations [11–13,80].
The common point of all of these works is that fluctuations originate from two factors: internal and external. Naturally, the dynamics, the structure and the interaction of the nodes vary from case to case. We expect that, for example, the stock market trading activity and the reproduction of trees is very different. The discovery of FS as a common pattern can be a good start to point out further analogies and to build a broader picture.

3. A general formalism

In the following we focus on the temporal variant of FS. In many cases the results continue to apply by simply dropping the time index \( t \) and averaging over an ensemble of systems.

The previous section reviewed ample evidence that FS emerges in a very broad range of areas. Here we attempt to describe many of these by the same formalism. In the following we assume that the systems are stationary. When considering a node \( i \), its activity \( f_i \) will always be decomposed as a sum. In some cases this means summation over the node’s internal constituents, as for forests where reproductive activity was the total of that for all trees. In other cases the nodes themselves are simple, and the signal is the sum of events at the nodes, like the passing of cars at counting stations. If in the time interval \( [t, t + \Delta t] \) there are \( N_i^{\Delta t}(t) \) such constituents/events, and the activity of the \( n \)th contributes \( V_i^{\Delta t}(t) \) to the total activity \( f_i^\Delta(t) \), then

\[
f_i^{\Delta t}(t) = \sum_{n=1}^{N_i^{\Delta t}(t)} V_i^{\Delta t}(t).
\]

These \( V \)'s do not necessarily have to be independent [8,46], but we assume that their (unconditional) distribution does not depend on \( n \). We omit the index \( \Delta t \), where appropriate.

3.1. The components of fluctuations

Throughout this section we analyse the fluctuations of quantities of the form (1), as measured by the standard deviation/variance. Thus, it is important that there exists a simple analytical expression [57] for

\[
\sigma_i^2 = \left\langle \left[ f_i(t) - \langle f_i(t) \rangle \right]^2 \right\rangle.
\]

Appendix A gives a proof that

\[
\sigma_i^2 = \Sigma_{V_i}^2 \left( N_i^{2H_{V_i}} \right) + \Sigma_{N_i}^2 \langle V_i \rangle^2,
\]

where \( \langle V_i \rangle \) and \( \Sigma_{V_i}^2 \) are the mean and the variance of \( V_{i,n} \). Similarly, \( \langle N_i \rangle \) and \( \Sigma_{N_i}^2 \) are the mean and variance of \( N_i \). We also introduced the Hurst exponent \( H_{V_i} \) of the constituents, which is defined as

\[
\Sigma_{V_i}^2(N) = \left\langle \sum_{n=1}^{N} V_{i,n} - \left( \sum_{n=1}^{N} V_{i,n} \right) \right\rangle^2 \propto N^{2H_{V_i}}
\]

for any \( t \) and \( \Delta t \). If for a fixed \( i \) and \( t \) all \( V_{i,n}(t) \) are uncorrelated, then \( H_V = 1/2 \), while if they are long-range correlated \( H_V > 1/2 \) (see [60,81]).

---

*[8] Advances in Physics 109*
3.2. ‘Universality’ classes

In this formalism it is relatively easy to show that there exist two important classes of systems, one with $\alpha = 1/2$ and one with $\alpha = 1$. The existence of such classes was pointed out, for example, by Anderson et al. [16] and later by de Menezes and Barabási [7].

These are not universality classes but rather simple limiting cases, and $\alpha$ is not a universal exponent in the usual sense of statistical physics [24]. Many empirical systems do not belong to either class, and both $\alpha = 1/2$ and $\alpha = 1$ can arise from several types of dynamics. In order to make FS a truly useful tool in the analysis of empirical data, one needs a classification scheme for how different types of dynamics can be mapped onto $\alpha$. Our current understanding of such classification is outlined in this section.

3.2.1. The case $\alpha = 1/2$

We now present two scenarios that can give rise to $\alpha = 1/2$. The arguments are given in the language of time averages, but they can be generalized to ensemble averages in a straightforward way.

1. Let us assume that every node $i$ consists of a fixed number $N_i(t) = N_i$ of constituents, each with a signal $V_{i,n}(t)$ which is independent and identically distributed for all $i, n$ and $t$, with the same mean $\langle V \rangle$ and variance $\Sigma^2_V$. From (9) it is trivial that here $\sigma^2_I = N_i \Sigma^2_V$. Owing to the linearity of the mean, $\langle f_i \rangle = N_i \langle V \rangle$, so

$$\sigma^2_I = \frac{\Sigma^2_V}{\langle V \rangle} \langle f_i \rangle,$$

and hence $\alpha = 1/2$. Of course, in this simple case one can say more, because the central limit theorem [83] is applicable:

$$\frac{\sum_{n=1}^{N_i} V_{i,n}(t) - N_i \langle V \rangle}{\sqrt{N_i \Sigma^2_V}} \to G_i(t), \quad (11)$$

where $G_i(t)$ are independent and identically distributed standard Gaussians and $\rightarrow$ means convergence in distribution for $N_i \to \infty$.

Exactly the same equation can be rewritten to more resemble FS:

$$\frac{f_i(t) - \langle f_i \rangle}{K \langle f_i \rangle} \to G_i(t). \quad (12)$$

The power $\alpha = 1/2$ is exactly the power in FS, and $K = \Sigma^2_V \langle V \rangle^{-1/2}$. The conceptual difference is only that since we know that $\langle f_i \rangle = N_i \langle V \rangle$, we can use $\langle f_i \rangle \langle V \rangle^{-1}$ as a surrogate variable for $N_i$. This is very useful when we only have the time series of $f_i(t)$ available but not $N$, since the limit can be switched to $\langle f \rangle \to \infty$ (cf. Section 5.2).

1. Let us consider an example for scenario (1): a system, where $V_{i,n}(t)$ can only be 1 with probability $p$ and 0 with probability $1 - p$. Scenario 1) still applies because $V$’s are i.i.d., so $\alpha = 1/2$. This binary distribution can be instructive, as one can think of $V_{i,n}(t)$ as independent indicator variables. For example let us take a volume $S$ of ideal gas within a large container. Let the whole system contain $N$ gas atoms, and $V_{i,n} = 1$ if the $n$’th atom is in the container, while $V_{i,n} = 0$ if it is not. Then $f$ is the
number of atoms in the container. The ideal gas is homogeneous and the atoms are independent, every atom having a probability \( p \propto S \) of being in the small container. From here, one can apply the above argument to show that for various containers of different sizes

\[
\sigma_S \propto \langle f_S \rangle^{1/2}.
\]

These are the well-known square-root type fluctuations of equilibrium statistical physics (see, e.g., [25, section XII]). Similar arguments were suggested for the number of animals in an area: if the motion of individuals were independent (gas-like), then their spatial density fluctuations should follow \( \alpha = 1/2 \) (see [84]).

(1") The example of the ideal gas can be given in the language of ensemble averages as well. Simply we take a large number of containers of the same size \( S \) and calculate the mean \( \overline{f_S} \) and standard deviation \( \sigma_S \) of atom counts between these containers. Then we vary the container size, and we recover an analogous relationship:

\[
\sigma_S \propto \overline{f_S}^{1/2}.
\]

Of course, this was expected, because the system is ergodic, so temporal and ensemble averages are equal.

(2) For an even simpler mechanism let us recall the findings of Section 2.4.2. We found that for very short times (\( \Delta t \sim 1 \) s) the number of sent emails/printed documents follow TFS with \( \alpha_T = 1/2 \). It is highly unlikely that someone will send several different emails/print several different documents in the same second (duplicates of the same email to multiple recipients were excluded). Thus, \( f(t) = 0 \) or \( 1 \), and so \( f(t) = f_i^2(t) \). Remember that this is very different from the previous example, where \( f_i(t) \) was allowed to have any value, and only \( V_{i,x}(t) \) were constrained to 0 or 1.

In the email/print data the number of events per second was very low, generally \( \langle f_i \rangle < 4 \times 10^{-3} \text{ s}^{-1} \). The standard deviation is then

\[
\sigma_i^2 = \langle f_i^2 \rangle - \langle f_i \rangle^2 = \langle f_i \rangle - \langle f_i \rangle^2 \approx \langle f_i \rangle,
\]

so \( \alpha = 1/2 \). The same argument holds for the number of trades per second in the stock market [8].

In summary the meaning of this scenario is that we are examining the system on such a short time scale that no two events happen in the same time window. Then, the FS exponent tells us nothing about the dynamics of the system, because \( \alpha = 1/2 \) is automatically true.

### 3.2.2. The case \( \alpha = 1 \)

We now present two scenarios that can give rise to the value \( \alpha = 1 \). While (1) is only valid for TFS, (2) can be readily generalized for EFS as well.

(1) It was possible to obtain \( \alpha = 1/2 \) by sums of independent \( V \). In the other extreme case, if every node \( i \) had a fixed number of identical and completely synchronized
constituents, i.e. $N_i(t) = N_i$ and $V_{i,n}(t) = V(t)$, (1) simplifies to

$$f_i(t) = \sum_{n=1}^{N_i} V_{i,n}(t) = N_i V_i(t).$$

Then $\langle f_i(t) \rangle = N_i \langle V_i(t) \rangle$ and $\sigma_i = N_i \Sigma V_i$:

$$\sigma_i = \frac{\Sigma V_i}{\langle V_i \rangle} \langle f_i \rangle \propto \langle f_i \rangle^\alpha,$$

with $\alpha = 1$. The last proportionality only holds if the ratio $\Sigma V_i / \langle V_i \rangle$ is the same for any $i$, for example when the distribution of $V_i$ is independent\(^9\) of $i$.

(1') How is such an argument of any use? The study of Cho et al. [85] reports experimental data on samples of yeast, in which cells were artificially prepared to have almost perfectly synchronized cell cycles. The measured signal $f_i(t)$ is the hourly expression level of various genes $i$ in a sample. If all cells of yeast contribute in the same way to the measured expression level, and they are synchronized, then the value $\alpha_T = 1$ is simply an indicator of such a synchrony. Thus, FS probably tells us nothing about the dynamics of gene transcription, and the exponent is simply a result of the sample preparation.

Nacher et al. [69] proposed a stochastic differential equation model that predicts the same exponent $\alpha_T = 1$ for this dataset ($\alpha_T = 1$ is confirmed by Zikovic et al. [70]). They argue that self-affine temporal correlations are the origin of such a value. Section 3.3.1 will show that self-affine temporal correlations do not contribute to $\alpha_T$ in this way. Instead, our above explanation is simpler, and it suggests that the dataset cannot be used in favour of any proposed model based on the value of $\alpha_T$.

(1'') Real systems are often not closed, but subject to outside forces. In certain cases this driving can be so strong that it can overwhelm the internal dynamics. If the internal structure of the system becomes irrelevant, this must also have an effect on FS. There have been a number of studies discussing how fluctuations in complex systems are formed as the sum of internally generated and externally imposed factors [11–13, 79, 80]. Anderson et al. [16] and de Menezes and Barabási [7, 31] suggested that $\alpha_T = 1$ can arise when the external driving force imposes strong fluctuations in either $V_i(t)$ or $N_i(t)$ (cf. [12]).

When all $V_{i,n}(t)$ (the signals of every constituent at every node) become synchronized, then we are back at scenario (1): $\alpha = 1$, because $f_i(t)/\langle f_i \rangle = V(t)/\langle V \rangle$ which has a universal, $i$-independent distribution.

It is also possible that an external force $W(t)$ affects the number of constituents in the elements so strongly that the fluctuations of $N_i(t)$ become proportional only to this force. In this case $N_i(t) = A_i W(t)$, where $A_i$ are node-dependent constants. One expects that generally $\langle f_i(t) \rangle = A_i \langle W(t) \rangle \langle V_i \rangle$, whereas

$$\sigma_i^2 = \Sigma_{N_i} (V_i)^2 + \Sigma_i^2 \langle N \rangle = A_i^2 \Sigma_i^2 \langle V_i \rangle + \Sigma_i^2 A_i \langle W \rangle.$$

If fluctuations in $W$ are so large that $\langle W \rangle \ll \Sigma_i^2 W$, then only the first term remains. After some algebraic steps

$$\sigma_i^2 \approx \frac{\Sigma_W^2}{\langle W \rangle^2 \langle V_i \rangle} \langle f_i \rangle^2 \propto \langle f_i \rangle^{2\alpha_T},$$
with $\alpha_T = 1$. The last proportionality is true if the distribution of $V_{i,n}$ does not depend strongly on $i$.

(2) $\alpha = 1$ can be a sign of a universal distribution of $f_i(t)/\langle f_i \rangle$, which only varies by a constant multiplicative factor throughout nodes. If this is true, then $f_i(t)$ can be decomposed into this factor $F_i$, and the universal random variable $V_i(t)$, which is identically distributed for all $i$. Naturally $\langle f_i \rangle = F_i \langle V \rangle$, $\sigma_i^2 = F_i^2 \Sigma_i$ and $\sigma_i = \Sigma_i \langle V \rangle^{-1} < f_i >$.

3.3. Other values of $\alpha$

It has been observed that many real systems obey FS with $\alpha$ values that significantly differ from both 1/2 and 1. In this section we summarize the current knowledge of general mechanisms that can give rise to intermediate values.

3.3.1. The dependence of $\alpha$ on the time resolution $\Delta t$

First of all, $\alpha$ can depend on the size of the time window used for its measurement. This phenomenological picture can be used to understand the results of Section 2.3.4 for stock market trading and Section 2.4.2 for human activity.

Let us assume that the activity time series are long time correlated with Hurst exponents $H_i$ that are allowed to depend on the node $i$. The Hurst exponent of the time series $f_i^{\Delta t}(t)$ was previously defined as

$$\sigma_i(\Delta t) = \left[ f_i^{\Delta t}(t) - \langle f_i^{\Delta t} \rangle \right]^{1/2} \propto \Delta t^{H_i}. \quad (7)$$

This definition is almost exactly the same as (10) for $H_V$, the only difference being that now instead of the $N$ number of constituents we consider the time window size $\Delta t$ as the scaling variable. TFS deals with how the variance scales when one moves to stronger (larger $\langle f \rangle$) signals:

$$\sigma_i \propto \langle f_i \rangle^\alpha. \quad (3)$$

Equation (7) takes an alternative point of view and suggests that for a fixed signal, in the presence of long-range temporal correlations, the variance can grow anomalously also by changing the time window.

Following [8], from (7) and (3), it is easy to see that the roles of $\langle f_i \rangle$ and $\Delta t$ are analogous. Since the left-hand sides are the same, one can write a third proportionality between the right-hand sides:

$$\Delta t^{H_i} \propto \langle f_i \rangle^{\alpha(\Delta t)}. \quad (15)$$

After taking logarithm on both sides, and differentiating by $\partial f_i \langle \log \Delta t \rangle \partial (\log \langle f_i \rangle)$, one finds that asymptotically

$$\frac{dH_i}{d(\log \langle f_i \rangle)} \sim \frac{d\alpha(\Delta t)}{d(\log \Delta t)} \sim \gamma. \quad (15)$$
This means that both partial derivatives have the same constant value, which we denote by $\gamma$.

Eisler and Kertész [8] outline three scenarios for this equality to hold:

(I) In systems, where $\gamma = 0$, the exponent $\alpha$, is independent of window size, and the degree of temporal correlations ($H$) is the same at all nodes.

(II) When $\gamma > 0$, $\alpha(\Delta t)$ depends on $\Delta t$ logarithmically: $\alpha(\Delta t) = \alpha^* + \gamma_1 \log \Delta t$. The Hurst exponent of the node also depends on $\langle f \rangle$ logarithmically with the same prefactor: $H_f = H^* + \gamma \log \langle f \rangle$.

(III) It is possible that (15) only holds piecewise, for certain ranges in $\Delta t$. Two regimes are then separated by a crossover between two distinct values $\gamma_{\pm}$, and nodes will have separate Hurst exponents $H_f(i)$ and $H_f(i)$ in the two regimes.

Case (III) was shown for the stock market and human dynamics in Section 2.3.

3.3.2. Impact inhomogeneity

Any value of $\alpha$ can easily arise without dependence on the time window. To better understand the reason how, consider three toy systems with the following elements.

(i) Let us take a fair coin with 0 written on one side and 1 on the other, this will be our group $i = 1$. Then take two such coins for group $i = 2$, three for $i = 3$, etc. In every time step we flip all coins in every group, and let $f_i$ equal the sum of the numbers we flipped in element $i$. Naturally, $\langle f_i \rangle \propto i$ and, if all coins are independent, $\sigma_i \propto i^{1/2}$. Thus, for such a case $\alpha = 1/2$.

(ii) Now let us take another fair coin with 0 written on one side and 1 on the other, this will be our element $i = 1$. For $i = 2$, we again take only one coin with sides 0 and 2. For any $i$, there will be a single coin with sides 0 and $i$. Trivially $\langle f_i \rangle \propto i$, but also $\sigma_i \propto i$. So this time $\alpha = 1$.

(iii) In our final example, let us mix the above two. For the $i$th group there are $i$ coins, each having a side with 0 and a side with $i$. Then $\langle f_i \rangle \propto i^{1/2}$, whereas $\sigma_i \propto i^{1/2} \times i$. We have just constructed a case for $\alpha = 3/4$.

One can unify these examples by introducing impact inhomogeneity. One can write the contribution (impact) of the constituents at a node $i$ as

$$V_{i,n}(t) = \{V_{i,n}\} \cdot X_{i,n}(t),$$  \hspace{1cm} (16)

all $X_{n,i}(t)$ are independent and identically distributed with unit mean. We then allow $\langle V_{i,n} \rangle$ to depend on $\langle N_i \rangle$ as a power law between nodes [33,67]:

$$\langle V_{i,n} \rangle \propto \langle N_i \rangle^\beta.$$  \hspace{1cm} (17)

According to (9) fluctuations can be calculated as

$$\sigma_i^2 = \sum_{i}^2 \langle N_i \rangle + \sum_{i}^2 \langle V_i \rangle^2 = \sum_{i}^2 \langle N_i \rangle^2 \langle V_i \rangle^2 + \langle N_i \rangle \langle V_i \rangle^2 \propto \{f_i\}^{2\alpha},$$
where $\Sigma_1^2 = \langle X^2 \rangle - \langle X \rangle^2$, and

$$\alpha = \frac{1}{2} \left(1 + \frac{\beta}{\beta + 1}\right),$$

(18)

where we introduced the new parameter $\beta$.

As a quick check, the three toy models correspond to $\beta = 0$, $\alpha = 1/2$ (all coins 0 or 1); $\beta = 1$, $\alpha = 3/4$ (the coins value proportional to their number) and $\beta \to \infty$, $\alpha = 1$ (only one coin with growing value). There is always some $\beta \geq 0$ that allows us to reproduce a given value $\alpha \in [1/2, 1)$, whereas the range $\beta < 0$ covers all possibilities of $\alpha < 1/2$ and $\alpha > 1$.

3.3.3. Examples of impact inhomogeneity

The ecology literature has documented [86,87] that empirically there is a strong positive correlation between the typical size of subpopulations $\langle V_i \rangle$ and the number of subpopulations per unit area $\langle N_i \rangle$ or the total population per unit area $\langle f_i \rangle$. The conjecture that these quantities might behave as powers of each other as in (17) was proposed by Keitt et al. [67], both across species and for individual subpopulations of the same species.

Kendal makes a similar suggestion, and shows that it generates non-trivial exponents in EFS for ecological populations [18] and the heterogeneity of blood flow in organs [68]. In fact he does not point out the general mechanism, but instead refers to non-trivial EFS exponents as the property of a class of models, which entail impact inhomogeneity. Here we omit most of the formalism; a proof that Kendal’s approach has impact inhomogeneity can be found in Appendix B.

Let us take the case of animal populations as the example. Kendal proposes that EFS holds with an exponent $1/2 < \alpha < 1$ if the population of an area can be described by the so-called Tweedie exponential dispersion models [88]. These models are widely used, also outside ecology [110], and they emerge from a limit theorem similar to the central limit theorem [110,111]. In the range $1/2 < \alpha < 1$ these assume that (i) an area $i$ contains a Poisson distributed number of animal clusters $\langle N_i \rangle$, (ii) the size of individual clusters $\langle V_{i,n} \rangle$ is independent and identically distributed gamma distributed, and (iii) there is a power-law relationship between the means of these two quantities. Of course, (iii) is the same as (17), along with all of its consequences.

As for blood flow [68], it is measured by the entrapment of radioactive microspheres in capillaries. In a fixed mass of tissue, the number of entrapment sites $N$ is assumed to be Poisson distributed, while the blood flow $V$ of the sites is taken as gamma distributed, along the same lines and with the same conclusions as above.

Finally, Section 2.3.4 suggested impact inhomogeneity also as the origin of non-trivial FS exponents for the traded value on stock markets.

3.3.4. Constituent correlations

There exists a further mechanism to produce any value $1/2 \leq \alpha \leq 1$, without considering the scaling property of impacts. The total output of node $i$ is given by
the equation

\[ f_i(t) = \sum_{i=1}^{N_i} V_{i,n}(t). \]  

(1)

We also fix \( N_i \) as time independent. If we assume that the unconditional distribution of \( V \) is independent from \( N_i \), and also from \( n \), then one can denote the expectation value \( \langle V \rangle = \langle V_{i,n}(t) \rangle \).

The central idea is the introduction of correlations between constituents, i.e. variables with different \( n \). Let us assume for simplicity that the elements are situated on a one-dimensional lattice, and their activity is long-range correlated in space, so that the correlation function decays as a power law,

\[ C(\Delta n) \propto \langle V_{i,n} V_{i,n+\Delta n} \rangle - \langle V_{i,n} \rangle^2 \propto \Delta n^{2H_V-2}. \]  

(19)

Here \( H_V \) is the same Hurst exponent, as defined in (10). Then, positively correlated patterns display \( H_V > 1/2 \), for uncorrelated (or short-range correlated) patterns \( H_V = 1/2 \), and for anticorrelated (antipersistent) patterns \( H_V < 1/2 \).

It follows from (9) that the fluctuation of the combined activity of all constituents is

\[ \sigma^2_i = \Sigma_{V}^{2} N_i^{2H_V} \propto \langle f_i \rangle^\alpha, \]

where

\[ \alpha = H_V. \]  

(20)

This idea was (to the best of our knowledge) first presented by Whittle [112] to describe the yield of crop fields. It was later independently developed by West [89], and demonstrated on surrogate datasets, but it was not applied directly to any new problem. The role of spatial correlations in the formation of FS in the context of ecology was also suggested by Colman-Lerner et al. [90] and more recently by Ballantyne and Kerkhoff [46]. The idea is confirmed by simulations; see Section 4.4.

4. Models

In this section, we discuss some models that can be used to understand basic facts about FS, how it arises and what its limitations are.

4.1. Random walks on complex networks

4.1.1. The model

That random walks can generate TFS was proposed by de Menezes and Barabási [7] in the following way. Let us take a scale-free Barabási–Albert network of \( M \) nodes [91]. We distribute \( W \) independent random walkers (tokens) randomly to the nodes. Then, in every time step these jump from their current node to one of its neighbours randomly. The process is repeated for \( s = 1, \ldots, s_{\text{max}} \) steps, then it is halted and the total number of visits to each node \( i \) is counted. This number defines \( f_i(t = 1) \). Then the deposition and the walk
is repeated, up to $T$ times, giving the time series $f_j(t)$. We ran simulations with the parameters $M = 20,000$, $W = 100$, $s_{\text{max}} = 100$ and $T = 10,000$.

One finds that TFS holds with an exponent $\alpha_T = 1/2$; see Figure 15. This value is the same as the value that arises from sums of independent random variables, so the central limit theorem is a possible origin of FS for random walks. The next part presents an analytical calculation that confirms this conjecture.

### 4.1.2. Fluctuation scaling and corrections

The model can be solved based on a master-equation approach [33,92]. Here we use elementary probability theory instead. The number of walkers on node $i$ can be calculated from their distribution in the previous time step as

$$N_i(s + 1) = \sum_{j \in K_i} \sum_{n=1}^{N_j(s)} \delta_n(j \to i; s),$$

where $\delta_n(j \to i; s)$ is a variable that is 1 if in step $s$ the $n$th token was at node $j$ and then it jumped to node $i$ (which happens with probability $1/k_j$ to all neighbours of node $i$) and 0 otherwise. $k_i$ is the degree of node $i$, $K_i$ is the set of neighbours of node $i$ and $N_j(s = 0)$ corresponds to the initial condition.

Calculations in Appendix C show that for such a model

$$\langle f_i \rangle = s_{\text{max}} \langle N_i \rangle = k_i \frac{s_{\text{max}} W}{\sum_j k_j} = \rho k_i,$$

where $\rho = s_{\text{max}} W / \sum_j k_j$. As the $W$ number of walkers is multiplied by the $s_{\text{max}}$ and divided by the total number of edges; $\rho$ can be understood as the mean number of walkers passing...
any edge during the $s_{\text{max}}$ time steps. Furthermore,

$$\sigma_i^2 = \sum_{j \in K_i} \frac{\sigma_j^2}{k_j} + \langle f_i \rangle. \quad (23)$$

The first term on the right-hand side is a sum over $k_j$ nodes, but every term is multiplied by $1/k_j^2$, thus they can be neglected to a first order. To a leading order

$$\sigma_i^2 = \langle f_i \rangle,$$

thus we find FS with $\sigma_T = 1/2$.

The term with the sum presents corrections to the scaling law. Equations (22) and (23) could be solved numerically, but to gain a qualitative understanding of these corrections it is enough to make a self-consistent solution up to the first non-trivial order. This can be done by taking $\sigma_j^2 = \langle f_j \rangle = \rho k_j$, and substituting it back into the right-hand side of (23), to find

$$\sigma_i^2 = \sum_{j \in K_i} \frac{\rho}{k_j} + \rho k_i = \rho k_i \left( 1 + a \left( \frac{1}{k_{N_i}} \right) \right), \quad (24)$$

where $\langle 1/k_{N_i} \rangle$ is the average inverse neighbour degree of node $i$ and $a = 1$. Simulation results supporting this argument are shown in Figure 16. We find that this formula accounts for a large part of the corrections to FS, only the coefficient is different, $a \approx 3.6$.

The qualitative picture from the above three equations is the following. For simplicity let us consider $\rho = 1$, when the average number of tokens at a node equals its degree. Thus, on average in every step every node transmits one token on each of its edges to its neighbours. Consequently, every node receives typically one token on each edge, so again it will have tokens equal to its degree. These tokens arrive independently, thus the variance

Figure 16. (a) Fluctuation scaling for the random walker model with parameters $M = 20,000$, $W = 100$, $s_{\text{max}} = 100$ and $T = 10,000$. The same as Figure 15, only without the binning procedure. The scaling law with $\sigma_T = 1/2$ holds on average, but there is some systematic broadening. The inset shows 10 randomly selected points from the indicated area, with the average inverse neighbour degree $\langle 1/k_{N_i} \rangle$ indicated for each node. There is a general increasing tendency in $\langle 1/k_{N_i} \rangle$ from bottom to top; (b) The value of $\sigma^2(f)^{1/2}$ plotted versus the average inverse neighbour degree $\langle 1/k_{N_i} \rangle$ of the node. There is an approximately linear relationship of the form $\sigma^2/f \sim 1 + 3.6(1/k_{N_i})$. 

Z. Eisler et al.
is proportional to their number, which implies $\alpha_T = 1/2$. The corrections in (24) imply that nodes with relatively higher degree neighbours (smaller $\langle 1/k_{Ni}\rangle$) exhibit lower fluctuations.\textsuperscript{14} This is because the number of tokens at a neighbouring site with smaller degree is smaller, and thus can have larger relative fluctuations. These fluctuations then affect our site via a stronger variation in incoming tokens.

This argument is important, because it tells us that for random walks FS is only approximately true. The local topology of the network can give significant corrections which cause a broadening in the scaling plots and which are not simply a result of measurement noise. According to (24) the size of the correction term depends on the neighbourhood of the node. As $\langle f_i \rangle \propto k_i$, very large flux nodes also have many neighbours. In an uncorrelated network the term $\langle 1/k_{Ni}\rangle$ will converge to a constant value with growing $k_i$, its node dependence (and, thus, the broadening it causes) is diminished.

4.1.3. The role of node–node interactions and a connection with surface growth

Previously we have shown that $\langle f_i \rangle = \rho k_i$, where $\rho$ is the average number of tokens passing any edge during a time step. The fluctuations of the number of visits to node $i$ come from two sources: (i) the number of such initial tokens at the neighbours; and (ii) how many of the tokens at its neighbours continue their walk to node $i$ in the next step. The number of tokens at a node is coupled with the state of its neighbours in the previous step. This effective interaction between neighbouring nodes is the origin of the corrections to FS in (24). To prove this, de Menezes and Barabási suggest a mean-field model \[7\] which eliminates this interaction as follows.

Instead of a direct contact between nodes, let us completely disconnect the network, and connect every node with its original number of edges to a reservoir. In every step ($s = 1, \ldots, S$ as in the original model) the reservoir sends $W$ tokens, their destination is chosen randomly between the edges. These tokens return to the reservoir in the next step, but simultaneously $W$ new tokens are sent out, etc. It is trivial that in this case fluctuations of the type (ii) are absent: all nodes are neighbours of the reservoir only, which emits the exact same number of tokens every time. Moreover, the distribution of $f_i(t)$ will be Poissonian with mean and variance $\rho k_i$. Thus, exactly

$$\sigma_i = \langle f_i \rangle^{1/2} = \sqrt{\rho k_i},$$

i.e. $\alpha_T = 1/2$ without any corrections. Moreover, both the network topology and the ‘randomness’ of the walk was completely eliminated. As suggested by de Menezes and Barabási \[7\], the remaining model is equivalent to a surface growth problem. Consider a finite one-dimensional lattice with $\sum_i k_i$ sites. At every time step $W$ tokens are deposited on the surface randomly. The Hurst exponent of the resulting surface is equivalent to the $\alpha_T$ of the non-interacting model (cf. (20)).

This example suggests that FS in the random walker model is a mean-field property. The interaction between the nodes is only responsible for higher-order corrections that do not change the scaling exponent in general. Most models in the literature are either non-interacting in this sense \[20,45,52\] or this interaction is not relevant \[7,33\]. At most, complex dynamics is limited to the structure within \[37\] the nodes, but not between the nodes. There exist a few studies of transport models on complex networks where the
interaction between the nodes becomes relevant. In these models FS breaks down and topology-dependent crossovers appear owing to congestion [35] or the presence of multiplicative noise [34].

4.1.4. The role of external driving

Finally, let us briefly remark on the behaviour of the model in the presence of external driving. One can allow the number \( W \) of walkers to fluctuate between the times \( t \) as

\[
W(t) = \langle W \rangle + \Sigma_W \times G(t).
\]

We chose \( G(t) \) as independent and identically distributed standard Gaussians, but the findings are largely independent of the shape of the distribution. If at any time \( W(t) \) became less than zero, we set it to \( W(t) = 0 \). If we restrict ourselves to the mean-field solution, then at every node [7]

\[
\sigma_i^2 = \langle f_i \rangle + \left[ \frac{\Sigma_W}{\langle W \rangle} \right]^2 \langle f_i \rangle^2. \tag{26}
\]

This result implies that when \( \Sigma_W > 0 \), there is a crossover from \( \alpha_T = 1/2 \) to \( \alpha_T = 1 \) around the node strength \( \langle f \rangle \sim \langle W \rangle^2/\Sigma_W^2 \).

The process was simulated with the other parameters set as before. With the increase of \( \sigma_W \) the best fit to (3) yields intermediate effective exponents between 1/2 and 1; see the inset of Figure 15. However, these are not ‘true’ exponents, only signatures of the crossover.

4.1.5. Impact inhomogeneity

The random walker model can be modified [33] to entail (17). This means that when a walker steps onto a site with typically more visitations, it generates a higher impact. As the number of visits is proportional to the degree of the node \( \langle N_i \rangle \propto k_i \), in order to have the impact inhomogeneity relationship \( \langle V_i \rangle \propto \langle N_i \rangle\beta \), one can simply introduce that for a token visiting a node of degree \( k_i \), the impact should be \( \langle V_i \rangle = k_i^\beta \). Simulation results perfectly conform with the theory, \( \alpha_T(\Sigma_W = 0) \) depends on \( \beta \) as expected from (18). The crossover persists to \( \alpha_T = 1 \) when one introduces a large variation in the number of tokens, see Figure 17.

4.2. Critical fluctuations and finite size scaling (FSS)

The mechanism of how (spatial) correlations produce non-trivial values of \( \alpha \) draws on some fundamental knowledge in statistical physics. Critical systems are known to exhibit anomalous fluctuations owing to the presence of strong, but non-trivial correlations. These originate from the interactions of the internal constituents such as Ising spins.

It is instructive to consider the simple ferromagnetic case, such as the nearest-neighbour Ising model [24] on a \( d \)-dimensional square lattice. As this model does not \textit{a priori} have dynamics, its analysis can be understood in the language of ensemble averages.
The number of spins is \( N = L^d \), where \( L \) is the linear size of the lattice. At the critical point local magnetization has a diverging correlation length, and the correlation function becomes of the power law form

\[
C(r) \propto \frac{1}{r^{d-2+\eta}}. \tag{27}
\]

The squared fluctuations of total magnetization \( \overline{\sigma^2(M_L)} \) are known to be proportional to the susceptibility \( \chi \), and for finite systems both quantities diverge at the critical point as

\[
\overline{\sigma^2(M_L)} \propto \chi \propto L^{d+\gamma/\nu}. \tag{28}
\]

This is one of the well-known results of FSS [94].

The susceptibility can be calculated as the integral of the correlation function:

\[
\chi = \frac{N}{k_B T} \int d^d r C(r) \propto N \int_0^L \frac{d^d r}{r^{d-2+\eta}} \propto L^{d+2-\eta}. \tag{29}
\]

It is well known that the exponents in (28) and (29) are related, \( \gamma/\nu = 2 - \eta \) (Fisher’s law [24]). At the critical point, owing to the interactions between the spins, the susceptibility becomes super-extensive, i.e. it grows faster than \( \propto L^d \), a typical sign of criticality.

Let us now consider an ensemble of finite Ising systems at the critical temperature with zero external field and with various linear sizes, and let the signal be the \( N_{L,\uparrow} \) number of ‘up’ spins. Of course, the total number of up and down spins is constant:

\[
N_{L,\uparrow} + N_{L,\downarrow} = L^d,
\]

and their difference gives the magnetization as

\[
M_L = N_{L,\uparrow} - N_{L,\downarrow}.
\]
With the notation \( o(L^\eta)/L^\eta \to 0 \), at the critical point
\[
N_{L,\uparrow} = L^d/2 + o(L^d).
\]
On the other hand, the fluctuations of \( M \) and \( N_\uparrow \) are proportional, because
\[
M_L = 2N_{L,\uparrow} - L^d,
\]
and so
\[
\overline{\sigma(N_{L,\uparrow})^2} = \overline{\sigma(M_L)^2}/2 \propto L^{d+2-\eta} + o(L^{d+2-\eta}).
\]
Consequently, to a leading order, there exists EFS between the fluctuations and the mean of the number of up spins:
\[
\overline{\sigma(N_{L,\uparrow})^2} \propto \overline{N_{L,\uparrow}^{2\alpha_E}}
\]
with
\[
\alpha_E = \frac{1}{2} + \frac{2 - \eta}{2d} = \frac{1}{2} + \frac{\gamma/\nu}{2d}.
\]

The above are true up to the upper critical dimension, which is \( d_c = 4 \) for the Ising model [94]. The mean-field results can be recovered by substituting the corresponding values: \( d = d_c = 4, \gamma_{MF} = 1, \nu_{MF} = 1/2, \eta_{MF} = 0 \). Finally \( \alpha_{E, MF} = 3/4 \), in agreement with the direct mathematical proof of Ellis and Newman [95]. Moreover, at the critical point the susceptibility is super-extensive, so \( \chi \) must grow faster than \( L^d \). This means that in (29) \( d + 2 - \eta > d \), and thus \( \eta < 2 \). On the other hand if \( \eta \) is non-negative, then from the constraints \( 0 \leq \eta < 2 \) and \( d \geq 1 \) it immediately follows that \( 1/2 \leq \alpha_E < 1 \). This range is also valid for the analogous behaviour of all \( n \)-vector models.

This result is two-fold, depending on how we look at it.

(i) The exponent \( \alpha_E \) resembles the FSS exponent of fluctuations/susceptibility. Thus, in this case, \( FS \) is essentially \( FSS \). The difference is that the FS calculation can be performed even when there are no data available regarding ‘system size’. Instead, because \( N_\uparrow \) is a positive extensive quantity, we know that its expectation value will be proportional to the system size, and thus it can act as a surrogate variable for \( L^d \). An anomalous value of the FS exponent can be related to critical behaviour, although (as previous sections suggest) not necessarily. We discuss this question in detail in Section 4.3.

(ii) The finding that when the constituents are long-range correlated gives rise to anomalous values of \( \alpha \), leaves us with a recipe of how to construct simple models that display \( 1/2 < \alpha < 1 \). The simplest scenario is described in detail in Sections 3.3.4 and 4.4.

What is the case with \( N_\uparrow \) off the critical point? In the paramagnetic phase the mean number of up spins is exactly \( N/2 \), while the fluctuations are of the order of \( N^{1/2} \), thus \( \alpha_E = 1/2 \). The ferromagnetic case is a more delicate issue, because the infinite system is not ergodic: spontaneous magnetization is symmetry breaking. For finite systems with a local (e.g. Glauber) dynamics it takes a finite (but very long) time for magnetization to change direction. The phenomenon is more easily interpreted via an (unrestricted) ensemble of
equilibrium ferromagnets. Here still \( \tilde{f} = N/2 \), because configurations magnetized up and down average out. The fluctuations on the other hand are macroscopic, \( \sigma^2 \propto 2|M| \propto L^d \).

Thus, \( \alpha_E = 1 \). In summary, the paramagnet–ferromagnet phase transition is signaled by FS as an abrupt change between the two universality classes (similarly to the Satake–Iwasa forest model [44,45]). At the critical point one finds intermediate exponents that can be calculated from the usual critical exponents. However, it is of fundamental importance that the anomalous FS is not observed in the order parameter \( M \). Instead, it is observed in an extensive quantity, and only whose fluctuations reflect the anomalous fluctuations of the order parameter. FS is there in \( M \), but with a trivial exponent: from FSS \( \bar{M} \propto L^{d-\beta/\nu} \).

This, combined with (29) leads to \( \bar{\sigma}(\bar{M}_L) \propto M^{(\nu+\gamma)/[2(\nu-\beta/\nu)]} \). Owing to the hyperscaling relation \( \gamma + 2\beta = d\nu \) this means that \( \alpha_E = 1 \).

The critical point is a very special state of a system, while FS with \( 1/2 < \alpha < 1 \) occurs very often. To make criticality a viable explanation for these non-trivial values of \( \alpha \) it is important to note that certain types of dynamics under strong external driving can self-organize to their critical state without the fine-tuning of any parameters [2,96,97]. Many real-life systems display the classical signs of self-organized criticality (such as power-law distributions, long-range correlations, etc.) and the value of \( \alpha \) can help to understand the dynamical origins of these observations.

### 4.3. Scaling and multiscaling

Scaling has a fundamental importance in statistical physics. It has found countless successful applications starting with critical phenomena [98], but more recently also outside the classical domain of physics, for example, in ecology [99]. In many cases scaling is not bound to a specific set of system parameters as in the case of critical phenomena, but it is the generic behaviour of the system as in polymers [100], surface growth [101] and self-organized criticality [2]. Mono-scaling or gap scaling means that the probability distribution of a quantity \( f \) depends on the parameter \( L \), usually the system size, as

\[
\mathbb{P}(f; L) = f^{-1} F\left( \frac{f}{L^\Phi} \right),
\]

where \( F \) is a scaling function and \( \Phi \) is some constant. This form can account for a number of observations about power law behaviour in real systems.

Both gap scaling and FS characterize a large number of complex systems. Nevertheless, for the same quantity only one can be true except in a special case: if a quantity shows both gap scaling and FS, then this automatically implies \( \alpha = 1 \). One can reverse this argument: If for a quantity one finds FS with \( \alpha < 1 \) then it cannot exhibit gap scaling.

The proof is straightforward. Any moment of \( f \) can be calculated as

\[
\overline{f^q}_L = \int_0^\infty df f^q \mathbb{P}(f; L) \simeq K_q L^{q\Phi},
\]

where \( \simeq \) denotes asymptotic equality and \( K_q > 0 \). From (32) it follows that

\[
\overline{\sigma^2}_L = \overline{f^2}_L - \overline{f}_L^2 \\
\simeq K_2 L^{2\Phi} - K_1^2 L^{2\Phi} = (K_2 - K_1^2) L^{2\Phi}.
\]
We combine EFS and (33), eliminate \( L \) and find that now \( \overline{\sigma^2} \propto f_L^{-2} \), i.e. \( \alpha = 1 \).

The only possibility for the coexistence of gap scaling (31) and FS (5) with \( \alpha < 1 \) is when the constant factor in the variance vanishes:

\[
(K_2 - K_1^2) = 0.
\]

In this case the gap scaling form does not describe the variance, that is instead given by the next order (correction) terms. Nevertheless, even if it is so, the leading order of the variance is still zero, and consequently \( F \) is proportional to a Dirac-delta:

\[
F\left(\frac{f}{L^\Phi}\right) \propto \delta(f/L^\Phi - K_1).
\]

This case is pathological, and it is usually not considered as scaling. In fact, the previous section contained one such example: the number of up spins in a critical Ising model follows this sort of statistics. Fluctuations scale anomalously (\( \overline{\sigma^2} \propto L^{d+2-\eta} \)), whereas their leading order vanishes because \( \overline{f^2} \sim f^2 \sim L^{2d}/4 \). Such strange scaling arises as a sign of criticality when the scaling variable is an extensive quantity, for which only the fluctuations are connected to those of the order parameter.

For example, in ecology there do exist species with \( \alpha \approx 1 \), for which a gap scaling form of the probability density of \( f \) could be valid. However, this value is by no means universal (cf. Figure 1). Similarly, \( \alpha < 1 \) was observed for Internet router traffic [7] or the traded value on stock markets [55]. These quantities cannot have a gap scaling form.

Instead of gap scaling, one can assume multiscaling, but the results do not change crucially. A probability distribution shows multiscaling if its size dependence is of the form

\[
\ln \mathbb{P}(f, L)/ \ln(L/L_0) = -F[\ln(f/f_0)/ \ln(L/L_0)],
\]

(34)

where \( f_0 \) and \( L_0 \) are appropriately chosen constants. The moments can be calculated by expressing the density function from (34) and substituting into the definition

\[
\overline{f^q}_L = \int_0^\infty f^q \mathbb{P}(f, L) df
= \int_0^\infty f^q \left(\frac{L}{L_0}\right)^{-F[\ln(f/f_0)/ \ln(L/L_0)]} df
\approx f_0^q \left(\frac{L}{L_0}\right)^{-F[a(q)]} \simeq K_q L^{\tau(q)}.
\]

The usual approach is that the value of the integral is dominated by the point \( f^*(q) \) where the integrand is maximal. Then

\[
a(q) = \frac{\ln[f^*(q)/f_0]}{\ln(L/L_0)},
\]

and

\[
\tau(q) = \max_a [qa - F(a)],
\]
or, equivalently, $\tau(q)/a\tau = a$. Now we are back at the same situation as with gap scaling, since

$$f_L \simeq K_1 L^{\tau(1)}$$

and

$$\sigma^2_L = \overline{f^2_L} - \overline{f_L}^2 \simeq K_2 L^{\tau(2)} - K_1^2 L^{2\tau(1)}.$$  

One expects that $\tau(2) \geq 2\tau(1)$, because the variance must remain non-negative for arbitrarily large $L$. If $\tau(2) > 2\tau(1)$ then the first term dominates $\sigma^2_L$, and $\alpha = \tau(2)/2\tau(1)$, but this value is greater than 1. For example Tebaldi et al. [102] report that in the Bak–Tang–Wiesenfeld (BTW) sandpile model of $L$ linear size, the distribution of the number of topplings $f$ in an avalanche follows $f_L \propto L^{\tau(2)}$ with $\tau(1) \approx 2$ and $\tau(2) \approx 4.7$. This results in an $\alpha \approx 1.17$.

The other possibility is again $\tau(2) = 2\tau(1)$ and $\alpha = 1$ (unless the leading order terms in $\sigma^2$ compensate to zero). This solution offers nothing new compared with gap scaling. Such relationships can be seen, for example, in the very same BTW model for the distribution of the area affected by avalanches [102].

The conclusion: if a quantity shows gap scaling with a scaling function which is not fully degenerate (not a Dirac-delta), it must follow $\alpha = 1$. If there is multiscaling, then FS with $\alpha > 1$ is also possible, but such values are rarely observed and should be taken with care.

### 4.4. Binary forest model

In this section we introduce a toy model that can be used to better illustrate the ideas of Sections 4.2–4.3. Moreover, we show that those are in full analogy with the findings of Section 2.3.2 for the reproductive activity of trees. For an easier understanding we present the model in that language.

Let us consider a forest that consists of $N$ trees. For simplicity we also assume that these are situated on a one-dimensional regular lattice, but any higher-dimensional generalization is straightforward. In the year $t$ the reproductive activity (i.e. seed count) of every tree $n$ is characterized by a random variable $V_n(t)$. Again, for simplicity we consider $V$ as binary variables, which are 1 with probability $p$ and 0 with probability $1-p$. As it takes several years for a new tree to reach its full reproductive capabilities, given that the observation period is short enough, we can neglect the changes in $N$ owing to seed production and tree growth.

The year-to-year correlations in seed counts are neglected. On the other hand, it is known [48] that the reproductive activity of forests exhibits long-range spatial dependence, with significant positive correlations for distances of thousands of kilometres. The distance dependence can be fitted approximately by

$$C(\Delta n) \propto \langle V_n V_{n+\Delta n} \rangle - \langle V_n \rangle^2 \propto \Delta n^{2H_v-2},$$  \hspace*{1cm} (19)
see Section 2.3.2. The total seed count is given by the usual form

\[ f_N = \sum_{n=1}^{N} V_n. \]

The standard deviation of the sum of random variables correlated according to (19) scales as

\[ \sigma_N = \sqrt{\langle f_N^2 \rangle - \langle f_N \rangle^2} \propto N^{H_V} \]

with \( H_V \) being the Hurst exponent (cf. (10)), whereas

\[ \langle f_N \rangle = pN. \]

The two equations can be combined into TFS with

\[ \alpha_T = H_V. \tag{20} \]

To the careful reader it should be clear that almost the same model was discussed in Section 4.2. There we argued that, in a critical Ising model, whether any given spin points upwards (1) or downwards (0) is essentially a binary random variable with \( p = 1/2 \). Moreover the spin alignments are power-law correlated in space, such that the power of the decay is related to the FS exponent \( \alpha \). This was expressed by (30), which is essentially equivalent to (20). The binary forest model only differs from the Ising case in that correlations between the random variables are given \textit{a priori}, and not generated by the thermodynamics.

Now we can move on to simulation results. Figure 18 shows the dependence of \( \sigma_N \) on \( N \) and \( \langle f \rangle \), the two plots are basically equivalent owing to \( \langle f \rangle = pN \). Figure 19(a) illustrates that the fluctuations in systems of the same size increase rapidly with \( H_V \). This is a result of strong synchronization of the individual constituents (see Figure 19(b)). The relationship (20) is illustrated in Figure 20.

![Figure 18](image_url)

**Figure 18.** (a) Scaling plots of \( \sigma \) versus \( N \) generated by Detrended Fluctuation Analysis of \( V_n(t) \) in the binary forest model. The slopes correspond to the (spatial) Hurst exponents \( H_V \approx 0.5, \ldots, 0.95 \) from bottom to top; see (10); (b) Scaling plots \( \sigma \) versus \( \langle f \rangle \) for FS in the same data. The slopes correspond to the values of \( \alpha_T \approx 0.5, \ldots, 0.95 \).
5. Discussion

In this section we present our view about unsettled questions related to FS. We also discuss recent, sometimes controversial techniques that might help in understanding FS.

Figure 19. (a) Examples of $f_t(t)$ time series for a ‘forest’ with $N=300$ ‘trees’. The Hurst exponent $H_V$ between the trees was varied: $H_V \approx 0.5, 0.65, 0.8, 0.95$ increasing from bottom to top. The data were shifted by the addition of a constant, but they were not stretched in any way. One can see that owing to the increasing synchronization of the constituents, relative fluctuations increase rapidly; (b) Snapshot of $V_{ln}$ series (at a fixed time $t$) for a forest with $N=300$ elements. The data were shifted by the addition of a constant. The Hurst exponent $H_V$ between the elements was varied: $H_V \approx 0.5, 0.65, 0.8, 0.95$ increasing from bottom to top. Spatial synchronization increases with the growth of the Hurst exponent.

Figure 20. The equality $\alpha_T = H_V$ (dotted line) in simulations of the binary forest model. The measurement points align very closely to the line, with some statistical deviations.
5.1. Separation of global and local dynamics

In Section 3.2 we argued that a system whose internal dynamics can be mapped onto the central limit theorem displays FS with $\alpha = 1/2$. On the other hand, if one imposes a strong external driving on the system, the behaviour crosses over to $\alpha = 1$. One example was shown in Section 4.1 in the case of random walks on complex networks. There the fluctuation was given by (26), which has the structure

$$\sigma_i^2 = \langle f_i \rangle + A^2 \langle f_i \rangle^2,$$

(35)

where $A$ is proportional to the strength of the external driving. If $A \ll 1$ one finds $\alpha = 1/2$, whereas in the strongly driven limit $A \gg 1$ the first term is negligible and $\alpha = 1$.

Now assume that we do not know the strength of external driving and we want to approximate it from data. We can introduce the global activity $F(t)$ of the system as a sum over all constituents:

$$F(t) = \sum_{i=1}^{N} f_i(t).$$

(36)

de Menezes and Barabási [31] suggest that if our system has many elements, then $F(t)$ will be proportional to the external force, because the independent fluctuations of the elements average out in the sum (36), and what remains is only the factor of the common external driving. This argument implicitly assumes, that the external force contributes to the fluctuations of the elements in a coherent way, i.e. $f_i(t)$ can be written in the form

$$f_i(t) = f_i^{\text{int}}(t) + f_i^{\text{ext}}(t),$$

(37)

where

$$f_i^{\text{ext}}(t) = A_i F(t).$$

(38)

This formula is a form of linear response, where: (i) $A_i$ is not allowed to depend on time because of stationarity; and, more importantly, (ii) all nodes are affected by driving instantaneously or with the same constant time lag.

After the summation of (36) we find that it is consistent with (37), if the normalization condition $\sum_i A_i = 1$ is satisfied. In order to keep (35) and (37) consistent in the strongly driven limit, the only possible choice is

$$A_i = \frac{\langle f_i \rangle}{\langle F \rangle}.$$ 

(39)

By this definition, automatically $\langle f_i^{\text{int}} \rangle = 0$ and $\langle f_i^{\text{ext}} \rangle = \langle f_i \rangle$. All time series have finite standard deviations, which are defined in the usual way, for example $\sigma_F = \sqrt{\langle F^2 \rangle - \langle F \rangle^2}$. With these

$$\sigma_i^{\text{ext}} = \frac{\sigma_F}{\langle F \rangle} \langle f_i \rangle,$$

(40)

so the external component follows FS with $\alpha = 1$ in any system. This appears consistent with the fact that in strongly driven systems $f$ itself also shows $\alpha = 1$, not only the external component. However, this is in fact just a trivial consequence of how the external component was defined.
de Menezes and Barabási [31] call the process of assigning the internal and external components ‘noise separation’ and claim that the procedure works well for the random walk model. For each node they define a noise ratio

\[ \eta_i = \frac{\sigma^\text{ext}_i}{\sigma^\text{int}_i}, \]

which is zero in the absence of external driving, and large when the fluctuations of the external component are dominant.

It would be tempting to attribute the real-world observations of \( \alpha \approx 1 \) to external driving, and show that in these cases typically \( \eta \gg 1 \). However, we demonstrate on some examples that noise separation has strong limitations.

de Menezes and Barabási [31] found that the fluctuations of Internet (Abiline backbone) traffic show TFS with \( \alpha = 1/2 \) and typically \( \eta \sim 0.1 \). While this appears very convincing, a more detailed analysis of a subset of the same data [32] instead finds \( \alpha = 0.7–0.8 \). The latter study suggests congestion as the origin of the increase value of \( \alpha \), and does not assume any external driving force.

de Menezes and Barabási [7] report that river-level fluctuations fall into the class \( \alpha = 1 \). It seems plausible that water levels fluctuate owing to rainfall on the river basin, which can be understood as external driving. However, noise separation is impossible here, because the driving is not coherent. The global factor \( F(t) = \sum_i f_i(t) \) is meaningless, because the response times of the water level, and the timing of precipitation vary from river to river. Hence, (38) is not valid.

To the best of our knowledge, our study [55] was the first to reveal FS in the trading activity of stocks. Noise separation was carried out there, finding that the typical value of \( \eta_i \) increases with the time window \( \Delta t \). As \( \alpha \) also shows a similar tendency (cf. Figure 6), we suggested that external driving must play a role in this effect. We also argued that this is because information needs a finite time to spread on the market. On the scale of a few minutes the role of external information is small and localized, whereas on the long run trading is dominated by the external macroeconomic trends and news.

Later we proposed a much simpler explanation [8], which was also summarized in Sections 2.3.4 and 3.3.1 of this review. In the stock market (and human dynamics; see Section 2.4.2) one observes that, for long times, \( \alpha(\Delta t) = \alpha^* + \gamma \log \Delta t \) and \( H_i = H^* + \gamma \log(f_i) \) with some \( \gamma > 0 \). These laws have not yet been related to any external force, even though the possibility cannot be ruled out.

How would noise separation work in this case?

(i) Clearly \( \sigma_F \propto \Delta t^{H_F} \), with \( H_F \approx \max_i H_i \), because \( F \) is the sum of all \( f_i \), and the scaling of the sum is dominated by the highest Hurst exponent.

(ii) We have \( \sigma^\text{ext}_i \propto \sigma_F \propto \Delta t^{H_F} \). On the other hand, if \( \sigma^\text{ext}_i < \sigma^\text{int}_i \), then one expects that qualitatively \( \sigma^\text{int}_i \propto \Delta t^{H_i} \).

(iii) Thus, the ratio \( \eta_i = \sigma^\text{ext}_i / \sigma^\text{int}_i \) should typically grow as long as \( \eta_i < 1 \). This observation of [55] is hence no proof of any particular external influence.

While the presentation of further calculations is not the purpose of this review, we believe that \( \eta \approx 1 \) can arise from spurious effects. A value \( \eta_i > 1 \) consistently, for many nodes, has only been observed in a single study where \( \eta \approx 1.5 \) (see [59]). Our present understanding is that noise separation has a limited range of applicability.
Finally, we would like to point out that the identification of the ensemble average (36) with some external force is somewhat controversial. As we do not have any information about the origin or the physical meaning of the factor $F(t)$, it is probably more appropriate to call this and $f_i^{\text{ext}}$ global and not external factors. Accordingly, $f_i^{\text{int}}$ are better called local rather than internal factors when it is not known how much they represent internal processes at the nodes.

5.2. Limit theorems for sums of random variables

In Section 3.2.1 we briefly remarked on the connection of $\alpha = 1/2$ to the central limit theorem. We recall that $f$ is written as a sum over the constituents (other random variables) whose number $N$ we consider as time independent:

$$f = \sum_{n=1}^{N} V_n.$$ 

Let us assume that a general form of central limit theorem is applicable, so that

$$\frac{\sum_{n=1}^{N} V_n - N\langle V \rangle}{N^{\alpha} \Sigma_V} \rightarrow X,$$  \hspace{1cm} (41)

where $X$ is some random variable and $\rightarrow$ means convergence in distribution for $N \rightarrow \infty$. In the language of FS the same equation reads

$$\frac{\sum_{n=1}^{N} V_n - \langle f \rangle}{K\langle f \rangle^\alpha} \rightarrow X,$$  \hspace{1cm} (42)

where $K$ is a constant. The conceptual difference is only that because we know that $\langle f \rangle = N\langle V \rangle$, we can use $\langle f \rangle \langle V \rangle^{-1}$ as a surrogate variable for $N$.

This analogy tells us that the appearance of FS throughout disciplines might be a result of the generality of certain limit theorems. The trivial example is of course that of independent and identically distributed variables with positive mean and finite variance, leading to the value $\alpha = 1/2$, but there are several other cases.

If the $V$ are independent and identically distributed, but their distribution decays asymptotically as $P(f) \propto f^{-(\lambda + 1)}$ with $0 < \lambda < 2$, then the Lévy–Gnedenko central limit theorem\textsuperscript{16} is applicable [83]. That is similar in spirit to (41), with $\alpha = 1/\lambda$. The difference is that $\sigma$, and if $\lambda \leq 1$ even $\langle f \rangle$, is infinite. However, for $N < \infty$ they will have some finite effective value, which can show apparent FS with some non-trivial value of $\alpha$.

In contrast to the relative simplicity of independent (and possibly identically distributed) random variables, dependent variables can be extremely diverse. They have no general theory, and the number of universality classes/limit theorems is infinite. Their structure is not always fully described by pairwise correlations and Hurst exponents (cf. Sections 4.2–4.4). In these cases sometimes there exists no limit distribution, or, for example, $\alpha < 1/2$ or $\alpha > 1$ in (41) (see [105]).

Even for the usual $1/2 \leq \alpha \leq 1$ values there is only a limited set of results, here we only mention a few inspired by statistical mechanics. In a series of papers Ellis, Newman and Rosen [95,106,107] show that in some statistical mechanical systems physical quantities can obey (41) with $\alpha = 1 - 1/(2k)$, where $k$ is a non-negative integer. For example, in the
Curie–Weiss mean-field model the number of up spins obeys \( k = 2 \) and \( \alpha = 3/4 \) at criticality, and the distribution of \( X \) can also be given explicitly. In Section 4.2 we arrived at the same exponent using heuristic arguments. Baldovin and Stella [108] recently published some more general results on a mean-field theory of strongly correlated random variables. In their model fine-tuning the strength of correlations allows for any \( 1/2 \leq \alpha \leq 1 \), much in the spirit of Section 4.4 and [46].

5.3. The connection between ensemble and temporal averages

Let us now return to the connection between the TFS

\[
\sigma_t(\Delta t) \propto \langle f_t \rangle^{\alpha T},
\]

and the EFS

\[
\sigma_N(\Delta t) \propto \overline{f_N}^{\alpha E},
\]

laws. These correspond to two definitions of the statistical quantities: (i) for (3) the mean and the standard deviation are calculated as temporal averages; (ii) for (5) they are calculated on an ensemble of subsystems of the same size.

For the mere existence of such quantities it is necessary to assume that: (i) signals with the same mean have the same statistical properties, and the processes are stationary; or (ii) systems of the same size can be considered elements of the same statistical ensemble, which is a kind of a homogeneity condition. In real systems neither of these conditions holds exactly, but they often prove to be good approximations. A deviation from these assumptions is one possible source of the observed broadening of the scaling plots. For example, in the case of precipitation data in Section 2.4.3 we found that mean precipitation is not the only determinant of the amplitude of fluctuations. Areas with the same mean precipitation are not equivalent, because they can correspond to very different climates. Factors such as height and geographical position are also relevant.

A related concern is the presence of correlations [109]. The observations may be correlated in space or time. (i) Two nodes (e.g. populations or weather stations) which are located close to each other can have significant cross-correlations. Fits can be biased, because the observations are not independent. (ii) The signals of individual nodes can have strong temporal autocorrelations, which can amplify statistical errors when the time series are not long enough [78].

A more delicate problem is the connection between the two types of FS, which has only been very vaguely investigated in real systems so far. First of all, the two factors cannot be separated completely. McArdle et al. [77] point out this problem through the example of animal populations. (i) The measurement of the number of individuals in an area takes a finite time. There is an inflow and outflow of individuals, so the number fluctuates. Thus, temporal dynamics can affect the results. (ii) If we want to measure the time series of the size of a given population, we have to assign a spatial scale as what to consider as a population. The temporal dynamics will depend on this spatial scale of sampling, possibly in a non-trivial way.

Taylor and Woiwod [84] conducted a very large-scale study of the two (temporal and ensemble) FS laws in animal populations. A systematic comparison is possible when the same sites are sampled at the same time [77]. Taylor and Woiwod calculated the temporal
and spatial means and standard deviations of the abundance of some aphids, moths and birds, then calculated $\alpha_T$ and $\alpha_E$ for each species.

First of all, they found that the temporal and ensemble means of population differ significantly. Thus, it is not surprising that the values of $\alpha$ differ as well. There was absolutely no systematic relationship between $\alpha_T$ and $\alpha_E$, and even the same species can show several such values depending on its natural environment. Rather interestingly, the only systematic dependence between species is the presence of positive correlations between the value of $\alpha$ and average population size. For example, for temporal data this means that across species $\alpha_T$ is correlated with $1/M \sum_{i=1}^M \langle f_i \rangle$. The correlations are present in both cases, although stronger for the temporal variant. Taylor and Woiwod [84] also suggested that the interactions between individuals might contribute to the ensemble law more than to the temporal law.

In some studies such as ecology or climatology the definition of the spatial scale comes naturally. Still, most systems have some hierarchical structure on which a degree of aggregation is possible. For example, it is possible to analyse the fluctuations of Internet traffic at the autonomous system level instead of the router level, which might have different dynamics. On the stock market TFS holds not only for individual stocks, but also when we consider the trading activity of industry sectors (data available from the author upon request).

In summary, the relationship between EFS and TFS is rather unclear in real systems. The exponents $\alpha_E$ and $\alpha_T$ are seldom calculated for the same system, and when they are calculated, they have different values.

### 5.4. Fluctuation scaling for growth rates

To be able to interpret FS for temporal fluctuations one has to assume that the underlying system is stationary. For example, in the binary forest model of Section 4.4 we assumed that the number of trees is constant and we neglect the contribution of reproduction to the population. To depart from stationarity, we can consider a growing population of $N_i(t)$ individuals, all of which can reproduce at a time $t$ ($V_{i,o}(t)$ is the number of offspring), die ($V_{i,d}(t) = -1$) or do nothing ($V_{i,d}(t) = 0$). The population can now be written as a sum

$$N_i(t + 1) = N_i(t) + \sum_{n=1}^{N_i(t)} V_{i,n}(t).$$

Here $N_i(t)$ is obviously not stationary, because its distribution depends on its value in the previous time step. Nevertheless, one can still construct something similar to TFS by using a restricted ensemble average as follows.

Let us define the growth rate of a population as

$$f_i(t) = N_i(t + 1) - N_i(t).$$

Now let us make an ensemble of growth observations when the initial population was $N$. The growth rate in this restricted sample is given by

$$f_N = \sum_{n=1}^N V_n.$$
As $f$ can be negative, to postulate TFS for size dependence it is more convenient to use $N$ and not $\langle f \rangle$ as the scaling variable. We conjecture

$$\sigma_N = \sqrt{\langle f_N^2 \rangle - \langle f_N \rangle^2} \propto N^\alpha$$

(43)

in the spirit of the previous sections, and this is exactly what is found in many systems.

Keitt et al. [66,67] show that the growth rate fluctuations of animal populations scale as a non-trivial power of the initial population $N$. The finding is not specific for population growth, but occurs in many settings where a positive quantity fluctuates by the addition and subtraction of increments. The same behaviour was found by Lee et al. [61] for the growth rates of business firms, and Amaral et al. [62] even presents a model of the complex structure of the business growth process which predicts the correct exponent.

Jánosi and Gallas [63] criticize these results, and show the same distribution of growth rates and the scaling law (43) for the water-level fluctuations of the river Danube, which trivially must have a structure that is very different from business firms. Moreover, they show that the daily absolute change of water level scales with the average water level on the same day, and there is a clear crossover behaviour between two scaling regimes with $\alpha = 1/2$ and $\alpha = 1$. A related study by Dahlstedt and Jensen [64] estimates $\alpha \approx 0.9–1$, and suggests that FS can be decomposed into two distinct scaling laws: $\sigma_A \propto A^a$ and $\langle f_A \rangle \propto A^b$, where $A$ is the area of the river basin.

From the above it is clear that the size-dependent scaling of growth rate fluctuations is a variant of FS for non-stationary (growing) signals. The same formalism can be applied in both cases, and many results could be mutually applied.

6. Conclusions

The aim of this review has been to provide a broader perspective on Taylor’s law and FS, and to encourage the collaboration between disciplines where these phenomena are observed. We have also outlined a classification scheme on the meaning of the FS exponents. The main conclusion is that several types of mechanisms can lead to the same value of $\alpha$. A similar concern was formulated in the 1982 paper of Taylor and Woiwod [84]:

“Extrapolation of dynamic principles from […] observation is likely to be misleading. We find great differences between [species], but the overlap is also very large. Whilst it is improbable that the details of […] behaviour in a bird and an aphid would be alike, there are common elements in the […] structure of their populations.”

While FS alone is not enough to identify the underlying dynamics of a system, it is useful for excluding some possibilities, and for rejecting certain models which would generate unrealistic values for $\alpha$. Empirical data from virtually all fields of science display FS, and so it is possible to make statements about almost any system where such data are available. For this very reason, in order to deepen our understanding of the phenomenon, it is becoming increasingly important to bridge the gap between several disciplines. However, the most puzzling question still remains: why do, for example, email traffic, stock market trading and the printing activity in a computer lab behave in similar, non-trivial ways? Some insights can be gained from the time window dependence of $\alpha$. That can reveal whether on some time scale the behaviour of the system reduces to something
simpler or possibly trivial. One can also make efforts to decompose the signals into well-defined constituents, so that a mean-field model based on sums of random variables can be applied. We believe that a possible common origin of all FS laws is the generality of these underlying mean-field type mathematical structures.

Note added in proof
We would like to thank Bent Jørgensen for calling our attention to the fact, that the Tweedie models, discussed in Section 3.3.3 and Appendix B, emerge under very general conditions. The Tweedie convergence theorem [110,111], mainly related to i.i.d random variables, shows that any exponential dispersion model that is compatible with fluctuation scaling will converge to one of the Tweedie models. As Appendix B shows, these can reproduce any $\alpha$ value within (but also most of those outside) the realistic range $[1/2, 1]$, because they explicitly include impact inhomogeneity.

Acknowledgments
Writing this manuscript would not have been possible without the help of many of people. The authors thank Péter Csermely for advice on Taylor’s law in ecology. They thank Bálint Tóth for discussions of statistical physics and limit theorems. They thank Jari Saramäki for his comments and Jukka-Pekka Onnela for a critical reading and countless useful remarks. They also thank Maya Paczuski, Peter Grassberger and Albert-László Barabási for their ideas on FS. They are grateful to Walter Koenig for data on the reproductive activity of trees, Ricardo Azevedo for cell count data and Jordi Duch and Alex Arenas for their Internet dataset. Ford Ballantyne IV, Marm Kilpatrick and Joe N. Perry are acknowledged for their help with the sections on ecology. The authors also gratefully acknowledge correspondence with Jayanth R. Banavar and Andrea Rinaldo on scaling laws in ecology. Finally, they thank Imre Jánosi for his help and criticism on the analysis of precipitation data. ZE is grateful to Jean-Philippe Bouchaud and for the hospitality of l’Ecole de Physique des Houches. This work was supported by OTKA K60456 and OTKA T049238.

Notes
1 Please note that ‘log’ will always denote 10-base logarithms.
2 A similar synchronization mechanism has also been observed in the reproduction of animals [11].
3 The same two averaging techniques (fixed time and an ensemble of stocks versus a fixed stock and multiple times of observation) were previously introduced for stock market price changes in [72].
4 The full dataset is available at http://www.cs.cmu.edu/~enron/. The filtered data used in this study can be found at http://www.isi.edu/~adibi/Enron/Enron.htm
5 The dataset is available at http://www.doc.ic.ac.uk/~uh/PASTRAMI/Printer/data
6 The dataset is available at http://www.ncdc.noaa.gov/oa/climate/research/gdcn/gdcn.html
7 Climatic fluctuations are well known to have a strong effect on ecological fluctuations [11,12]. It is an interesting fact that the variability of animal populations can also depend on latitudinal position (see [77] and references therein).
8 The Hurst exponent is only related to correlations in this simple fashion because the distribution of $V_{i,n}(t)$ does not depend on $n$ (see [82]).
9 If the dependence is present but weak, then it may cause corrections to FS, but scaling should still hold approximately.
10 This is often called local abundance.
11 This is often called regional abundance.
12 We need to assume that $V_{i,n}$ is stationary as a function of $n$. 
The particular topology is irrelevant from the point of view of $\alpha_T$. The network only has to be connected and the nodes should have a wide range of degrees.

The degree dependence of this correction is related to the assortativity of the network [93].

Except for the trivial case $H_V = 1/2$ (when $V$ are not strongly correlated) the above model is not very straightforward to simulate. We generated a one-dimensional fractional Brownian motion time series by applying the method of Koutsoyiannis [103], and then converted it into a sequence of zeros and ones (for higher dimensions it is necessary to use a more refined method, for example the method introduced by Prakash et al. [104] for the simulation of site percolation on long-range correlated lattices). The conversion slightly decreases the value of the Hurst exponent, which thus had to be measured independently by Detrended Fluctuation analysis [60]. For simplicity, we fixed the number of trees, because the effect of externally imposed noise ($\Sigma_N^2 > 0$) has already been studied in detail in Section 3.2.2 and, e.g., [7,33] for other models.

In fact, the conditions of the Lévy–Gnedenko central limit theorem are somewhat looser.

References

[1] T. Vicsek, *Fractal Growth Phenomena*, World Scientific Publishing, Singapore, 1992.
[2] P. Bak, C. Tang and K. Wiesenfeld, Phys. Rev. Lett. 59 (1987), pp. 381–384.
[3] V. Pareto, *Cours d'économie politique*, Droz, Geneva, Switzerland, 1896.
[4] G.K. Zipf, Harv. Stud. Classic. Philol. 15 (1929), pp. 1–95.
[5] L.R. Taylor, Nature 189 (1961), pp. 732–735.
[6] H. Fairfield Smith, J. Agric. Sci. 28 (1938), pp. 1–23.
[7] M.A. de Menezes and A.-L. Barabási, Phys. Rev. Lett. 92 (2004), Article no. 28701.
[8] Z. Eisler and J. Kertész, Phys. Rev. E 73 (2006), Article no. 046109.
[9] B.B. Mandelbrot, *The Fractal Geometry of Nature*, W.H. Freeman, New York, 1982.
[10] B.A. Maurer and M.L. Taper, Ecol. Lett. 5 (2002), pp. 223–231.
[11] B.T. Grenfell et al., Nature 394 (1998), pp. 674–677.
[12] O.N. Bjørnstad and B.T. Grenfell, Science 293 (2001), pp. 638–643.
[13] B.-E. Saether et al., Science 287 (2000), pp. 854–856.
[14] P.A. Moran, Aust. J. Zool. 1 (1953), pp. 291–298.
[15] D.H. Reed and G.R. Hobbs, Anim. Conserv. 7 (2004), pp. 1–8.
[16] R.M. Anderson et al., Nature 296 (1982), pp. 245–248.
[17] W.S. Kendal, Ecol. Modell. 80 (1995), pp. 293–297.
[18] W.S. Kendal, Ecol. Complex. 1 (2004), pp. 193–209.
[19] P.A. Marquet et al., J. Exp. Biol. 208 (2005), pp. 1749–1769.
[20] R.B.R. Azevedo and A.M. Leroi, Proc. Natl. Acad. Sci. U.S.A. 98 (2001), pp. 5699–5704.
[21] W.S. Kendal, J. Theor. Biol. 217 (2002), pp. 203–218.
[22] W.S. Kendal, Mol. Biol. Evol. 20 (2003), pp. 579–590.
[23] The International SNP Map Working Group, Nature 409 (2001), pp. 928–933.
[24] L.E. Reichl, *A Modern Course in Statistical Physics*, 2nd ed., Wiley, New York, 1998.
[25] L.D. Landau and E.M. Lifshitz, *Course of Theoretical Physics Volume 5: Statistical Physics Part I*, 3rd ed., Pergamon International Library, Oxford, 1980.
[26] R. Botet et al., Phys. Rev. Lett. 86 (2001), pp. 3514–3517.
[27] R. Botet and M. Ploszajczak, Phys. Rev. E 62 (2000), pp. 1825–1841.
[28] R. Botet and M. Ploszajczak, Nucl. Phys. B (Proc. Suppl.) 92 (2001), pp. 101–113.
[29] P. Uttley and I.M. McHardy, Mon. Not. R. Astron. Soc. 323 (2001), pp. L26–L30.
[30] S. Vaughan and P. Uttley, Proc. SPIE 6603 (2007), Article no. 660314.
[31] M.A. de Menezes and A.-L. Barabási, Phys. Rev. Lett. 93 (2004), p. 68701.
[32] J. Duch and A. Arenas, Phys. Rev. Lett. 96 (2006), Article no. 218702.
[33] Z. Eisler and J. Kertész, Phys. Rev. E 71 (2005), Article no. 057104.
[34] S.H. Yook and M.A. de Menezes, Europhys. Lett. 72 (2005), pp. 541–547.
[35] M. Šuvakova and B. Tadić, Physica A 372 (2006), pp. 354–361.
[36] J. Duch and A. Arenas, Eur. Phys. J. ST 143 (2007), pp. 253–255.
[37] A.M. Kilpatrick and A.R. Ives, Nature 422 (2003), pp. 65–68.
[38] L.R. Taylor, J. Anim. Ecol. 55 (1986), pp. 1–38.
[39] R. May, Stability and Complexity in Model Ecosystems, 2nd ed., Princeton University Press, Princeton, NJ, 1974.
[40] A.J. Lotka, Elements of Physical Biology, Williams and Wilkins, Baltimore, MD, 1925.
[41] V. Volterra, Variations and fluctuations of the number of individuals in animal species living together, Animal Ecology, McGraw-Hill, New York, 1925, pp. 409–448.
[42] M.J. Keeling, Theor. Popul. Biol. 58 (2000), pp. 21–31.
[43] J.N. Perry, Proc. R. Soc. Lond. B 257 (1994), pp. 221–226.
[44] A. Satake and Y. Iwasa, J. Theor. Biol. 203 (2000), pp. 63–84.
[45] F. Ballantyne IV and A.J. Kerkhoff, J. Theor. Biol. 235 (2005), pp. 373–380.
[46] F. Ballantyne IV and A.J. Kerkhoff, Oikos 116 (2007), pp. 174–180.
[47] A.J. Kerkhoff and F. Ballantyne IV, Ecol. Lett. 6 (2003), pp. 850–856.
[48] W.D. Koenig and J.M.H. Knops, Am. Nat. 155 (2000), pp. 59–69.
[49] K. Tallqvist, Folia Forestali 364 (1978), pp. 1–60.
[50] J.F. Franklin, Cone production by upper slope conifers, Research Paper No. PNW-60, Pacific NW Forest Range Experiment Station, 1968.
[51] M.J. Weaver and F. Forcella, Cone Production in Pinus albicaulis Forests, General Technical Report INT-203, USDA Forest Service, 1986.
[52] M.J. Keeling and B.T. Grenfell, Philos. Trans. R. Soc. Lond., B, Biol. Sci. 354 (1999), pp. 769–776.
[53] M.E.J. Woolhouse, L.H. Taylor and D.T. Haydon, Science 292 (2001), pp. 1109–1112.
[54] A. Bar-Even et al., Nat. Genet. 38 (2006), pp. 636–643.
[55] Z. Eisler et al., Europhys, Lett. 69 (2005), pp. 664–670.
[56] Z. Eisler and J. Kertész, Eur. Phys. J. B 51 (2006), pp. 145–154.
[57] Z. Eisler and J. Kertész, Europhys. Lett. 77 (2007), Article no. 28001.
[58] New York Stock Exchange, Trades and Quotes Database for 2000-2002, New York Stock Exchange, New York, 2003.
[59] Z.-Q. Jiang, L. Guo and W.-X. Zhou, Eur. Phys. J. B 57 (2007), pp. 347–355.
[60] J.W. Kantelhardt et al., Physica A 316 (2002), pp. 87–114.
[61] Y. Lee et al., Phys. Rev. Lett. 81 (1998), pp. 3275–3278.
[62] L.A.N. Amaral et al., Phys. Rev. Lett. 80 (1998), pp. 1385–1388.
[63] I.M. Jánosi and J.A.C. Gallas, Physica A 271 (1999), pp. 448–457.
[64] K. Dahlstedt and H.J. Jensen, Physica A 348 (2005), pp. 596–610.
[65] Z. Eisler, I. Bartos and I.M. Jánosi, in preparation.
[66] T.H. Keitt and H.E. Stanley, Nature 393 (1998), pp. 257–260.
[67] T.H. Keitt et al., Philos. Trans. R. Soc. Lond., B, Biol. Sci. 357 (2002), pp. 627–633.
[68] W.S. Kendal, Proc. Natl. Acad. Sci. U.S.A. 98 (2001), pp. 837–841.
[69] J.C. Nacher, T. Ochiai and T. Akutsu, Mod. Phys. Lett. B 19 (2005), pp. 1169–1177.
[70] J. Zivković et al., Eur. Phys. J. B 50 (2006), pp. 255–258.
[71] A. Mitninski and K. Rockwood, Mech. Ageing Dev. 437 (2005), pp. 699–706.
[72] F. Lillo and R.N. Mantegna, Phys. Rev. E 62 (2000), pp. 6126–6134.
[73] A.-L. Barabási, Nature 435 (2005), pp. 207–211.
[74] J.G. Oliveira and A.-L. Barabási, Nature 437 (2005), p. 1251.
[75] A. Vázquez et al., Phys. Rev. E 73 (2006), Article no. 036127.
[76] U. Harder and M. Paczuski, Physica A 361 (2006), pp. 329–335.
[77] B.H. McArdle, K.J. Gaston and J.H. Lawton, J. Anim. Ecol. 59 (1990), pp. 439–454.
[78] S.J. Clark and J.N. Perry, Environ. Ecol. Stat. 1 (1994), pp. 287–302.
Appendix A: The components of the fluctuation $\sigma^2$

A large part of this review is concerned with the standard deviation of the sums of random variables. This is defined as

$$\sigma^2 = \left( \left( \sum_{n=1}^{N} V_n \right)^2 - \left( \sum_{n=1}^{N} V_n \right)^2 \right).$$
where \( V_n \) are the individual (not necessarily independent) random variables, and \( N \) is the number of these variables which itself can be random.

Let \( P(N) \) be the probability that the number of variables is \( N \). The sum of \( N \) variables can be written as \( V_N = \sum_{n=1}^{N} V_n \). Let \( P(V_N) \) denote the density function of this sum when \( N \) is fixed. Then the standard deviation of the sum when \( N \) itself is a random variable is

\[
\sigma^2 = \sum_N P(N) \int dV_N P(V_N)V_N^2 - \left( \sum_N P(N) \int dV_N P(V_N)V_N \right)^2 \\
= \sum_N P(N) \left[ \int dV_N P(V_N)V_N^2 - \left( \frac{\int dV_N P(V_N)V_N}{\langle V_N \rangle} \right)^2 \right] + \sum_N P(N) \left( \frac{\int dV_N P(V_N)V_N}{\langle V_N \rangle^2} \right)^2 \\
= \sum_i \sum_N P(N) N + \langle V \rangle^2 \left[ \sum_N P(N) N^2 - \left( \sum_N P(N) N \right)^2 \right] \Sigma_i^2 \\
= \Sigma_i \Sigma_N^2 (N) + \langle V \rangle^2 \Sigma_N^2.
\]

Thus, finally

\[
\sigma^2 = \Sigma_i \Sigma_N^2 (N) + \langle V \rangle^2 \Sigma_N^2. 
\]

In the case when the \( V_n \) are strongly (i.e. power-law) correlated \( \Sigma_i^2 = \Sigma_N^2 N^{2H} \), where \( H_N \) is the Hurst exponent as defined in (10), and so

\[
\sigma^2 = \Sigma_i \Sigma_N^2 \langle N^{2H} \rangle + \langle V \rangle^2 \Sigma_N^2.
\]

The correlations in \( N \) are not reflected directly in this expression. Instead, they affect how \( \Sigma_N \) changes with the time window size \( \Delta t \) as pointed out in Section 3.3.1.

**Appendix B: Tweedie models and impact inhomogeneity**

In this appendix we prove that origin of the non-trivial \( \alpha \) values in the formalism of Kendal [18,68] is essentially impact inhomogeneity (see also [110]). This formalism is based on the so-called Tweedie exponential dispersion models [88]. These form a family of random distributions, and in the range \( 1/2 < \alpha < 1 \) they are characterized by the logarithmic cumulant function (see [88, p. 1516])

\[
K_f(s) = \ln \langle e^{s f} \rangle_f = \lambda_0 [g_0(s) - 1], \tag{B1}
\]

where \( s \) is a constant and \( f \) is the random variable. We use natural logarithms (ln), as opposed to other parts of this review, where we used 10-base logarithms (log). We also introduced the notation \( \langle x \rangle_f = \int_0^\infty dy e^{xy} \). The two terms above are

\[
\lambda = \frac{a - 1}{ka} \left( \frac{k \theta}{1 - a} \right)^a, \tag{B2}
\]
and

\[ g_0(s) = \left(1 + \frac{s}{\theta}\right)^a. \]  \hspace{1cm} (B3)

Here \( \theta > 0 \) and \( a < 0 \).

As pointed out both by Kendall [18] and Bar-Lev and Enis [88], the form (B1) is characteristic of compound Poisson processes [83]. These are distributions of random variables of the following type:

\[ f = \sum_{n=1}^{N} V_n, \]

where \( N \) is Poisson distributed, and \( V_n \) are independent and identically distributed random variables. The proof is straightforward, however, we include it here for completeness. The density function of a compound Poisson variable \( f \) can be written as a complete probability

\[ P(f) = \sum_{N=0}^\infty P(f|N)P(N). \]

The characteristic function is given by

\[
\langle e^{itf} \rangle_f = \int df e^{itf} P(f) \\
= \sum_{N=0}^\infty P(N) \int df e^{itf} P(f|N) \\
= \sum_{N=0}^\infty P(N) \int df e^{itf} P(V_1 + V_2 + \cdots + V_N) \\
= \sum_{N=0}^\infty P(N)\langle e^{itV}\rangle_V^N.
\]

For the last equality we used the property of the characteristic function that \( \langle e^{it\sum_n V_n} \rangle = \langle e^{itV} \rangle^N \). Then, knowing that if \( N \) is Poisson distributed with mean \( \langle N \rangle \) then its characteristic function is \( \langle e^{itN} \rangle = e^{i\langle N \rangle e^{it} - \frac{1}{2}t^2 \langle N \rangle} \),

\[
\sum_{N=0}^\infty P(N)\langle e^{itV}\rangle_V^N = \left\{ e^{i\ln \{\exp(sV)\}_N} \right\}_N \\
= e^{i\langle N \rangle(sV) - \frac{1}{2}t^2 \langle N \rangle} = \langle e^{itV} \rangle.
\]  \hspace{1cm} (B4)

The next step is to compare Equations (B2), (B3) and (B4). One finds that for the Tweedie model

\[ \langle e^{itV} \rangle_V = g_0(s) = \left(1 + \frac{s}{\theta}\right)^a. \]

This is the characteristic function of a gamma distribution. Its moments can be determined as usual:

\[ \langle V \rangle = \left. \frac{\partial}{\partial s} \langle e^{itV} \rangle_V \right|_{s=0} = a\theta^{-1}, \]

\[ \langle V^2 \rangle = \left. \frac{\partial^2}{\partial t^2} \langle e^{itV} \rangle_V \right|_{s=0} = a(a-1)\theta^{-2}. \]

\[ \Sigma_V = \langle V^2 \rangle - \langle V \rangle^2 = -a\theta^{-2}. \]
For the expectation value of the Poisson variable:
\[
\langle N \rangle = \lambda = \frac{a - 1}{ka} \left( \frac{k\theta}{a - 1} \right)^a \propto \theta^a.
\]
Furthermore,
\[
\langle f \rangle = \left[ \frac{\partial}{\partial s} K_i^\gamma(s) \right]_{s=0} = \lambda a \theta^{-1},
\]
and
\[
\sigma^2 = \left[ \frac{\partial^2}{\partial s^2} K_i^\gamma(s) \right]_{s=0} = \lambda a(a - 1) \theta^{-2}.
\]
Let us recall, that also \( \lambda \) contains terms with \( \theta \). Simple calculation yields
\[
\sigma^2 = kf^{2\alpha},
\]
where \( 2\alpha = (a - 2)/(a - 1) \). Consequently, we have the following.

(i) Let us introduce \( \beta = -1/a \). Then
\[
\langle V \rangle \propto \langle N \rangle^\beta,
\]
which is impact inhomogeneity.

(ii) Moreover, \( \sigma \propto \langle f \rangle^\alpha \) with
\[
\alpha = \frac{1}{2} \left( 1 + \frac{\beta}{\beta + 1} \right),
\]
exactly the same relationship as in Section 3.3.2.

Appendix C: Fluctuations in the network random walker model

This section contains calculations starting from the master equation (21). The total number of visitations to node \( i \) is the sum over all steps. The substitution of (21) gives
\[
f_i = \sum_{s=1}^{s_{\text{max}}} N_i(s) = \sum_{s=0}^{s_{\text{max}}-1} \sum_{j \in \mathcal{K}_i} \sum_{n=1}^{N_i(t-1)} \delta_n(j \rightarrow i; s),
\]
where \( \delta_{n,s}(j \rightarrow i; s) \) is a variable which is 1 if the \( n \)th token of node \( j \) in step \( s \) jumps to node \( i \) (happens with probability \( 1/k_j \)) and 0 otherwise. Here \( k_i \) is the degree of node \( i \), \( \mathcal{K}_i \) is the set of neighbours of node \( i \), and \( N_i(t=0) \) corresponds to the initial condition.

For any finite network one can switch the order of the first two sums, and in
\[
f_i = \sum_{s=1}^{s_{\text{max}}} N_i(s) = \sum_{j \in \mathcal{K}_i} \sum_{s=0}^{s_{\text{max}}-1} \sum_{n=1}^{N_i(t-1)} \delta_n(j \rightarrow i; s)
\]
if \( S \) is taken large and because for any fixed \( n \) the variables \( \delta_{n,s}(j \rightarrow i) \) are independent, then owing to the central limit theorem the last two sums converge to independent Gaussians:
\[
f_i = \sum_{s=1}^{s_{\text{max}}} N_i(s) = \sum_{j \in \mathcal{K}_i} \left( \frac{s_{\text{max}} \langle N_j \rangle}{k_j} + \sqrt{\frac{s_{\text{max}} \langle N_j \rangle}{k_j} \mathcal{G}_j(s)} \right),
\]
where $G_j(s)$ are independent and identically distributed standard Gaussians such that

$$\langle G_j(s)G_r(r) \rangle = \delta_{jr}\delta_{sr}, \quad (C3)$$

where the right-hand side has two Kronecker deltas. Consequently

$$\langle f_i(t)f_j(t) \rangle = \langle f_i(t) \rangle \langle f_j(t) \rangle, \quad \text{when } i \neq j. \quad (C4)$$

One can take the expectation value of the left-hand side of (C2). Finally,

$$\langle f_i \rangle = \sum_{j \in K_i} \frac{s_{\text{max}}(N_j)}{k_j}. \quad (C5)$$

By substitution one can check that the solution is

$$\langle f_i \rangle = s_{\text{max}}(N_i) = k\sum_j \frac{s_{\text{max}}(N_j)}{k_j}, \quad (C6)$$

and all the walkers are accounted for: $\sum_i \langle f_i \rangle = s_{\text{max}}W$.

Now let us calculate the standard deviation for both sides of (C2):

$$\sigma_i^2 = \left[ \left( \sum_{j \in K_i} \left( \frac{s_{\text{max}}(N_j)}{k_j} + \sqrt{\frac{s_{\text{max}}(N_j)}{k_j}\bar{G}_j} \right) \right)^2 - \left( \sum_{j \in K_i} \left( \frac{s_{\text{max}}(N_j)}{k_j} \right) \right)^2 \right] = \ldots \quad (a)$$

Here (a) can be replaced by $s_{\text{max}}(N_j)/k_j$, because $\langle \bar{G}_j \rangle = 0$:

$$\sigma_i^2 = \left( \sum_{j \in K_i} \left( \frac{s_{\text{max}}(N_j)}{k_j} \right)^2 \right) + \left( \sum_{j \in K_i} 2 \left( \sum_{l \in K_i} \frac{s_{\text{max}}(N_l)}{k_l} \right) \frac{s_{\text{max}}(N_j)}{k_j} \right) \frac{s_{\text{max}}(N_j)}{k_j} \bar{G}_j \right) \right)^2 \quad \text{(c)}$$

$$+ \left( \sum_{j \in K_i} \frac{s_{\text{max}}(N_j)}{k_j} \right)^2 - \left( \sum_{j \in K_i} \frac{s_{\text{max}}(N_j)}{k_j} \right)^2 \quad \text{(d)}$$

One can use (C4) to write

$$(b) = \left( \sum_{j \in K_i} \frac{s_{\text{max}}(N_j)^2}{k_j^2} \right) + \sum_{j \neq l \in K_i} \frac{s_{\text{max}}(N_j)(N_l)}{k_jk_l};$$

$$(c) = 0, \text{ because of } \langle \bar{G}_j \rangle = 0;$$

$$(d) = \left( \sum_{l \in K_i} \left( \sqrt{\frac{s_{\text{max}}(N_l)}{k_l}} \right)^2 \right) = \left( \sum_{l \in K_i} \frac{s_{\text{max}}(N_l)}{k_l} \right);$$

because of (C3); by changing a summation variable, one can write

$$(e) = \sum_{j,l \in K_i} \frac{s_{\text{max}}(N_j)(N_l)}{k_jk_l} \quad (C7)$$
Combining all the above, one obtains

\[
\left\langle \sum_{j \in K_i} \frac{s_{\text{max}}(N_j)^2}{k_j^2} \right\rangle + \sum_{j \in K_i} \frac{s_{\text{max}}(N_j)(N_i)}{k_j k_i} + \sum_{j \in K_i} \frac{s_{\text{max}}(N_i)}{k_i} - \sum_{j \in K_i} \frac{s_{\text{max}}(N_j)(N_i)}{k_j k_i}.
\]

Combine (f) and (g) and (h) to obtain

\[
\left\langle \sum_{j \in K_i} \frac{s_{\text{max}}(N_j)^2}{k_j^2} \right\rangle - \sum_{j \in K_i} \frac{s_{\text{max}}(N_j)^2}{k_j^2} \equiv \sum_{j \in K_i} \frac{\sigma_j^2}{k_j}.
\]

Then,

\[
\sigma_i^2 = \sum_{j \in K_i} \frac{\sigma_j^2}{k_j^2} + \sum_{j \in K_i} \frac{s_{\text{max}}(N_j)}{k_j}.
\]

The second term can be evaluated from (C5), to find

\[
\sigma_i^2 = \sum_{j \in K_i} \frac{\sigma_j^2}{k_j^2} + \langle f_i \rangle. \quad \text{(C8)}
\]