

OPEN ACCESS

The non-equilibrium nature of culinary evolution

To cite this article: Osame Kinouchi *et al* 2008 *New J. Phys.* **10** 073020

View the [article online](#) for updates and enhancements.

You may also like

- [Absorbed dose distribution visualization for superficial treatments through the Fricke Xylenol Gel dosimeter \(FXG\)](#)
M Alva, F G A Sampaio, M V Moreira et al.
- [System for measuring the transmission spectrum of `in vitro' corneas](#)
Liliane Ventura, Sidney J F Sousa, André M V Messias et al.
- [Two-dimensional sample entropy: assessing image texture through irregularity](#)
L E V Silva, A C S Senra Filho, V P S Fazan et al.

The non-equilibrium nature of culinary evolution

Osame Kinouchi¹, Rosa W Diez-Garcia², Adriano J Holanda¹,
Pedro Zambianchi³ and Antonio C Roque^{1,4}

¹ Departamento de Física e Matemática, FFCLRP, Universidade de São Paulo,
14040-901 Ribeirão Preto, SP, Brazil

² Departamento de Clínica Médica, Curso de Nutrição e Metabolismo,
Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo,
14040-901 Ribeirão Preto, SP, Brazil

³ Departamento de Física, Universidade Tecnológica Federal do Paraná,
Av. Sete de Setembro 3165, Centro, 80230-901 Curitiba, PR, Brazil
E-mail: antonior@ffclrp.usp.br

New Journal of Physics **10** (2008) 073020 (8pp)

Received 2 April 2008

Published 10 July 2008

Online at <http://www.njp.org/>

doi:10.1088/1367-2630/10/7/073020

Abstract. Food is an essential part of civilization, with a scope that ranges from the biological to the economic and cultural levels. Here, we study the statistics of ingredients and recipes taken from Brazilian, British, French and Medieval cookery books. We find universal distributions with scale invariant behaviour. We propose a copy-mutate process to model culinary evolution that fits our empirical data very well. We find a cultural ‘founder effect’ produced by the non-equilibrium dynamics of the model. Both the invariant and idiosyncratic aspects of culture are accounted for by our model, which may have applications in other kinds of evolutionary processes.

Almost every human group has its own cuisine [1, 2], usually one of the dearest aspects of its cultural identity. The diversity of foodstuffs that constitute a cuisine as well as their social and symbolic aspects are of permanent concern to anthropologists, psychologists and sociologists [3]–[5]. From another perspective, the chemical physics of culinary processes has attracted the interest of physicists and chemists [6, 7].

Here, we propose a statistical [8] and complex networks [9]–[11] approach to model culinary diversity and evolution. Culinary recipes are prime examples of cultural algorithms with strong capacity for stabilization, innovation and transmission. In literate cultures their standardized reproduction and widespread use are favoured by cookery books, which

⁴ Author to whom any correspondence should be addressed.

Table 1. Numerical characteristics for the considered cookery books.

Cookery book	No. of recipes	No. of ingredients	Average recipe size
<i>Dona Benta</i> (1946)	1786	491	6.7
<i>Dona Benta</i> (1969)	1894	657	7.1
<i>Dona Benta</i> (2004)	1564	533	7.4
<i>Larousse Gastronomique</i> (2004) ^a	1200	1005	10.8
<i>New Penguin Cookery Book</i> (2001)	878	489	9.5
<i>Pleyn Delit</i> (Medieval)	380	219	9.7

^aRecipes have been sampled. The total number of recipes is above 3000.

represent their culinary traditions. Cookery books provide statistical information about cuisines, indicating the relative importance of foodstuffs, preparation methods and combinations of them in a cuisine. These elements have made traditional cookery books a useful source of information for food researchers [12]–[15].

In this work, we examine only a particular aspect of cookery books, namely the relationship between its recipes and ingredients, and neglect for the moment the very important role played by culinary preparations [6]. We study the statistics of ingredient usage in different countries and cultures in search of common statistical patterns or differences between them.

We also propose a copy-mutate algorithm to model cuisine growth from a few initial recipes, which fits our empirical data very well. The model suggests an evolutionary dynamics where idiosyncratic ingredients are preserved in a manner akin to the founder effect in biology [16].

The culinary corpus considered by us consists of four different cookery books: the Brazilian *Dona Benta* [17], the French *Larousse Gastronomique* [18], the British *New Penguin Cookery Book* [19] and the medieval *Pleyn Delit* [20]. We have constructed a database for each one of them containing its entire population of recipes and ingredients. The only exception is the database for the French cookery book, which contains a random sample (40%) of the whole population of recipes. To examine temporal effects, we have considered three very different editions of *Dona Benta* [17] (1946, 1969 and 2004) which largely differ in recipes and ingredients repertoire. The numbers of recipes and ingredients in each database are given in table 1.

For each cookery book database, we counted the number of recipes in which each ingredient appears. The ingredients were ordered according to descending frequency of usage. This allows the representation of the statistical hierarchy of ingredients in a cookery book by a frequency-rank plot, as introduced by Zipf [21]. Figure 1(a) gives the frequency-rank plots for the Brazilian (1969 edn), British, French and medieval cookery books, showing a remarkable similarity in their statistical patterns.

Time invariance is supported by the data shown in figure 1(b), which give the frequency-rank plots for the editions of the Brazilian cookery book. The structure of ingredient rantage in the Brazilian cookery book remained stable amidst the change from a regional to a more globalized food consumer profile that took place in the last 50 years.

All these curves exhibit a power-law behaviour which can be well fitted by a Zipf–Mandelbrot law [22, 23] with an exponential cut-off [24] to capture finite size effects,

$$f(r) = \frac{C}{(a+r)^\beta} \exp(-r/r_c). \quad (1)$$

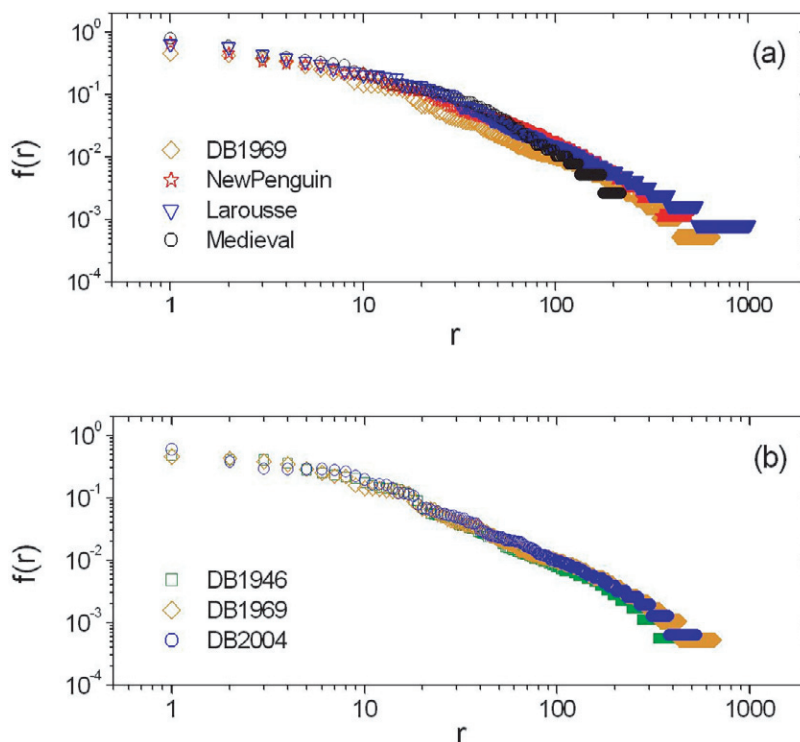


Figure 1. (a) Cultural invariance. Frequency-rank plot for different cookery books: *Pleyen Delit* (circles), *Dona Benta* 1969 (squares), *New Penguin* (stars) and *Larousse Gastronomique* (triangles). (b) Temporal invariance. Frequency-rank plot for *Dona Benta* 1946 (squares), 1969 (lozenges) and 2004 (circles).

The best fit exponent $\beta \cong 1.31$ worked reasonably well for all curves (the other parameters are $C = 8.7$, $a = 6.4$ and $r_c = 800$) suggesting that, like Zipf's law in linguistics [21], there exists a statistical universal behaviour which is robust across different cookery books and independent of the culture they refer to, their authors, motivations and even time.

It is important to notice that this is not a Zipfian lexical law, i.e. it is not a measure of appearance frequency of words (ingredient names) in cookery books. The ingredient usage frequency $f(r)$ is the relative number of recipes that use the ingredient with rank r .

A common measure used to characterize complex networks is the degree distribution, which gives the fraction of nodes with k links [9]–[11]. A bipartite network has two kinds of degree distribution, one for the ingredient nodes and the other for the recipe nodes. The degree distribution for the ingredients is the probability distribution $P^I(k)$ that a randomly chosen ingredient appears in k recipes. Plots of the degree distribution for the ingredients of the cookery books show that they are right-skewed with a two-decade power law region well fitted by $P^I(k) \propto k^{-\alpha}$ with $\alpha \cong 1.73$. This value has been found by a fit proportional to $k^{-(\alpha-1)}$ of the complementary cumulative distribution (figure 2),

$$P_C(k) = 1 - \sum_k P^I(k), \quad (2)$$

which gives the probability that an ingredient is used in more than k recipes. It is compatible with the relation $\alpha = 1 + 1/\beta$ valid for pure power laws present in rank plots and degree

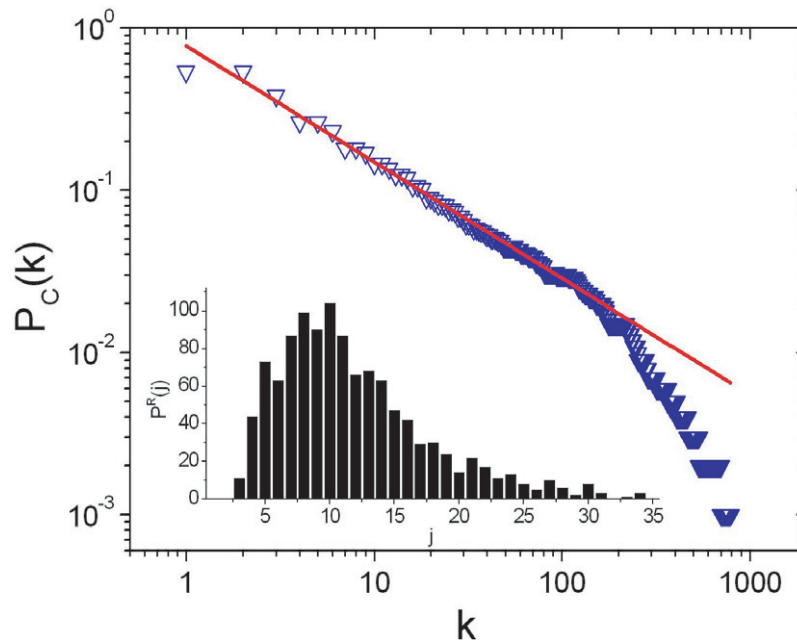


Figure 2. Complementary cumulative degree distribution $P_C(k)$ for ingredients with power law fitting $k^{-0.73}$ (solid line). Inset: degree distribution $P^R(j)$ for French recipes.

distributions [11]. This anomalous exponent cannot be obtained from a general Yule process [8] since in this latter case $\alpha \geq 2$.

The degree distribution for the recipes is the probability distribution $P^R(j)$ that a randomly chosen recipe has j ingredients (what we will call the size j of the recipe). A plot of the degree distribution for the recipes of the French cookery book is shown in figure 2 (inset). The average recipe sizes for all cookery books are given in table 1.

On the basis that culinary recipes are examples of cultural replicators, or ‘memes’ [25, 26], we propose a copy-mutate algorithm to model culinary evolution as a branching process (figure 3 inset). Our attempts to fit all the frequency distributions have shown that our model needs at least five parameters to work: the number T of generations (iterations or ‘time’) until the process is halted, the number K of ingredients per recipe (where K is compared to the average number $\langle K \rangle$ of the cookery book), the number $L (\neq K)$ of ingredients (‘loci’) in each recipe to be mutated, the number R_0 of initial recipes in the cuisine and M , the ratio between the sizes of the pool of ingredients and the pool of recipes. Hence, starting with $t_0 = R_0$, at each time step there are $R(t) = t$ recipes and $MR(t)$ ingredients available to be used.

We tried models without the concept of fitness, with no success. So, the present model ascribes to each ingredient i a random fitness f_i with values uniformly distributed in the interval $[0,1]$. We interpret this fitness as related to intrinsic ingredient properties such as nutritional value, aspect, flavour, versatility, cost and availability. At each iteration, we randomly chose a recipe (a ‘mother’) and copied it. Within this copy, we randomly chose an ingredient (with fitness f_i) to be compared with an ingredient also randomly chosen from the ingredients pool: if $f_j > f_i$, where f_j is the fitness of the ingredient from the pool, we replace ingredient i by ingredient j .

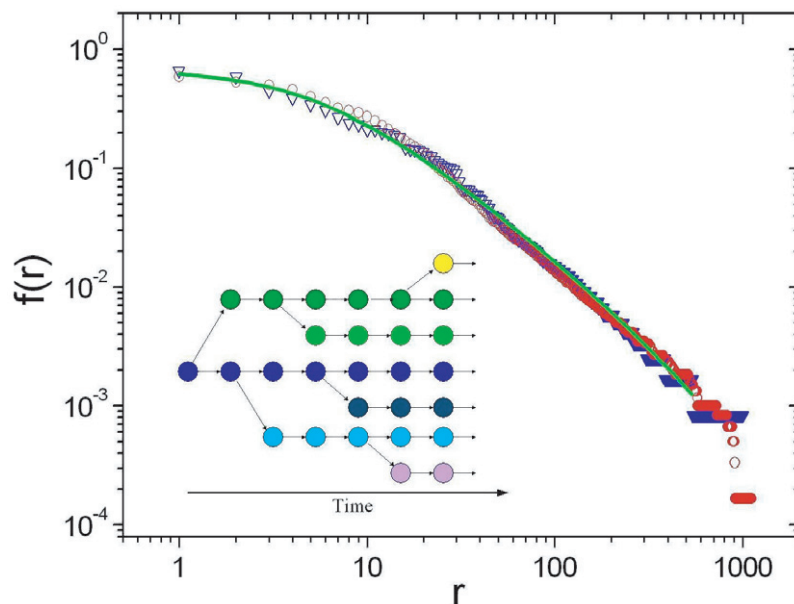


Figure 3. Frequency-rank plot for the Larousse cookery book (triangles) fitted by a Zipf–Mandelbrot curve (solid) and by the model (circles, average of five runs). Inset: schematic view of the copy-mutate growth process. At each time step, a randomly chosen ‘mother’ recipe generates a mutated daughter recipe. Notice that there is no recipe extinction.

This process is repeated L times, thus generating a ‘daughter’ recipe that is added to the recipes pool (it is possible that the daughter remains identical to the mother, which we interpret as a new recipe that differs from the previous one in the cooking procedures but not in their ingredients). Finally, at each time step, new ingredients are introduced in the ingredients pool to maintain the ratio M fixed. A somewhat similar algorithm has been proposed by Ramezanpour [27], but without the introduction of (the very important) fitness-based selection process.

A more general fitness function could be used,

$$f_i^I = f_i + \sum_j f_{ij} + \sum_{j,k} f_{ijk} + \dots,$$

to account for higher order interactions between ingredients, which would mean that the fitness of an ingredient is contextual and depends on the other ingredients present in the recipe.

We searched in the parameter space (T, K, L, M, R_0) for a good fit for the frequency-rank plot of the Larousse cookery book (figure 3), and obtained the values $T = 1200$, $K = 11$, $L = 4$, $M = 3$ and $R_0 = 20$. They show a good agreement with the actual values of the Larousse database, since the number of recipes is $R = T = 1200 = R_{\text{Larousse}}$ and $K = 11 \cong \langle K \rangle_{\text{Larousse}} = 10.8$. So, the effective free parameters are only the initial number of recipes R_0 (or independent ‘phylogenetic trees’), the mutation rate L and the growth rate M of the ingredients pool, which are not determinable from empirical data. Parameter fitting also gives a good agreement with the curves and values for all the other cookery books considered.

We also define the fitness of the k th recipe as $F^{(k)} = \frac{1}{K} \sum_{i=1}^K f_i$ and a total time-dependent cuisine fitness $F_{\text{Total}}(R(t)) = \frac{1}{R(t)} \sum_{k=1}^{R(t)} F^{(k)}$. In figure 4, we examine the temporal evolution of

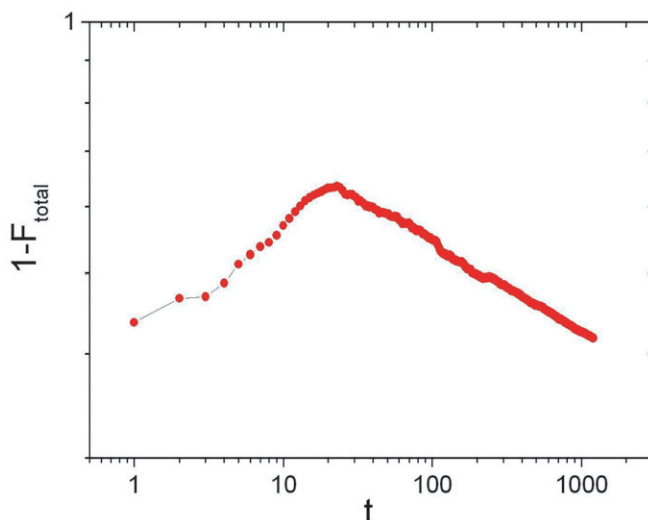


Figure 4. Power law decay with small exponent $\gamma = 0.1$ observed in the convergence of total culinary fitness F_{total} to 1. The rising phase is due to the initial copy of low fitness recipes, which lowers the overall culinary fitness F_{total} .

$1 - F_{\text{Total}}(t)$, which shows a very slow convergence to equilibrium in the form of a power law $1 - F_{\text{Total}}(t) \propto t^{-\gamma}$, with $\gamma = 0.1$. Hence, this kind of historical dynamics has a glassy character, where memory of the initial conditions is preserved, suggesting that the idiosyncratic nature of each cuisine will never disappear due to invasion by alien ingredients and recipes.

It is interesting that the same model presents a ‘stationary’ state (the frequency-rank power law) coexisting with a power law convergence toward a global fitness maximum. We conjecture that this scale invariance arises because the model implements a critical branching process [28]: the branching ratio is $\sigma = R(t+1)/R(t) \cong 1 = \sigma_c$.

The evolutionary model of copy mutation of recipes along with a selection mechanism generates a scale-free cuisine. This evolution is an out-of-equilibrium process: the number of recipes is never sufficient to fully explore the combinatorial space of ingredients. This means that the invasion of new high fitness ingredients and the elimination of initial low fitness ingredients never end. The latter is related to the ‘founder effect’ [16] known in evolutionary theory: the small number R_0 of initial recipes plays the role of a ‘founder’ population. As a consequence, some low fitness ingredients present in the initial R_0 recipes have a strong difficulty of being replaced and can even propagate during culinary growth. They are like frozen ‘cultural’ accidents difficult to overcome in the out-of-equilibrium regime (see figure 5).

We notice that sometimes in biology, gene fitness is taken as proportional to the relative frequency in the genome. In our model, the fitness variable f_i is an independent quantity that measures intrinsic competitive advantage of an ingredient. This means that a high fitness ingredient can have low prevalence and a low fitness ingredient can sometimes be well ranked. This occurs because the stationary state (the power law) is out-of-equilibrium. Only in an equilibrium state the two definitions of fitness coincide.

In our simulations, we have found that ingredient competition due to differential fitness is essential to obtain good modelling for rank values in the interval $r = 1, \dots, KR_0$. We tested an algorithm without fitness selection and observed that the conservation of the first KR_0 ingredients is so strong that they remain over represented in the population.

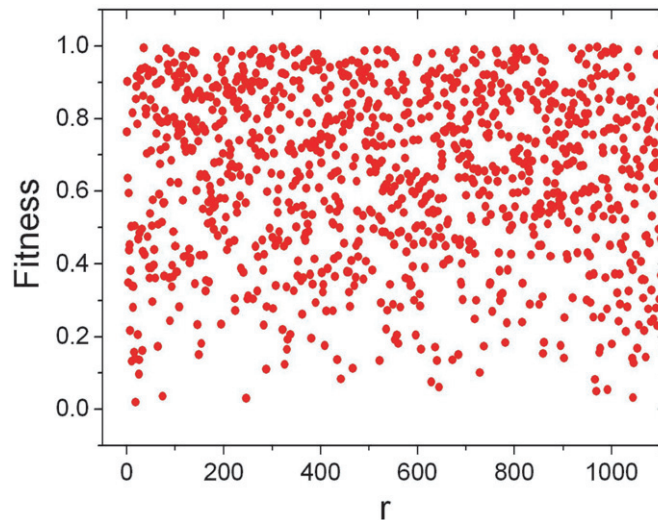


Figure 5. Scatter plot of ingredient fitness versus ingredient rank r . Observe the founder effect for some well ranked ingredients with small fitness (bottom left) and the overall average high fitness due to selective pressure. Notice also that there are several high fitness ingredients with poor ranking (upper right) due to the strong non-equilibrium character of the evolutionary process.

The mathematical modelling of evolutionary processes has attracted the attention of the scientific community [29], with particular emphasis on evolutionary dynamics [30, 31] and emergence of cooperation [32, 33]. As far as we know, the founder effect, though important, has not received comparable attention. In our model, the founder effect emerges naturally as a by-product of the strong non-equilibrium dynamics needed to obtain the highly skewed distributions observed in data. We conjecture that this property may be related to the fact that in our model the branching ratio is one, so that it is equivalent to a critical branching model.

In this work, we have found statistical patterns in culinary data independent of culture and time. Of course, there are other features (specific ingredients and recipes, ingredient ranks, etc) that depend on time and place and define the uniqueness of a cuisine. We have found that those universal statistical patterns can be obtained by a copy-mutate branching model. The model suggests that culinary history has a glassy dynamics very far from equilibrium. This would mean that diffusion of new ingredients within a cuisine is very slow, explaining the permanence and relevance of certain idiosyncratic ingredients as a kind of cultural founder effect.

Acknowledgments

We thank N Caticha, M Copelli, A S Martinez and P Riegler for discussions and encouragement, H This, M Buchanan and C Shalizi for critical comments and M A Bellezi, R F Caron, J G Costa, D M Gonçalves, D G Lidovero, G Marini, S T Pinto and S M F Ribeiro for preliminary statistical results. OK and AR are supported by research grants from CNPq.

References

- [1] Flandrin J L, Montanari M and Sonnenfeld A (ed) 2000 *Food: A Culinary History* (New York: Penguin)
- [2] MacLennan R and Zhang A 2004 *Asia Pacific J. Clin. Nutr.* **13** 131–5

- [3] Mintz S W and Du Bois C M 2002 *Annu. Rev. Anthropol.* **31** 99–119
- [4] Rozin P, Kabnick K, Pete E, Fischler C and Shields C 2003 *Psychol. Sci.* **14** 450–4
- [5] Leschziner V 2006 *Theor. Soc.* **35** 421–43
- [6] Weiss G 2001 *Science* **293** 1753–4
- [7] This H 2005 *Nat. Mater.* **4** 5–7
- [8] Newman M E J 2005 *Contemp. Phys.* **46** 323–51
- [9] Strogatz S H 2001 *Nature* **410** 268–76
- [10] Albert R and Barabási A-L 2002 *Rev. Mod. Phys.* **74** 47–97
- [11] Dorogovtsev S N and Mendes J F F 2002 *Adv. Phys.* **51** 1079–187
- [12] Appadurai A 1988 *Comp. Stud. Soc. Hist.* **30** 3–24
- [13] Sherman P W and Hash G A 2001 *Evol. Hum. Behav.* **22** 147–63
- [14] Helstosky C 2003 *Food Foodways* **11** 113–40
- [15] Segers Y 2005 *Appetite* **45** 4–14
- [16] Mayr E 1963 *Animal Species and Evolution* (Cambridge, MA: Harvard University Press)
- [17] *Dona Benta: Comer Bem* 1946 26th edn (São Paulo: Companhia Editora Nacional)
- Dona Benta: Comer Bem* 1969 51st edn Revised (São Paulo: Companhia Editora Nacional)
- Dona Benta: Comer Bem* 2004 76th edn Revised (São Paulo: Companhia Editora Nacional)
- [18] *Larousse Gastronomique* 2004 (Paris: Larousse)
- [19] Norman J 2004 *The New Penguin Cookery Book* (London: Penguin)
- [20] Hieatt C B, Hosington B and Butler S 1996 *Pleyn Delit: Medieval Cookery for Modern Cooks* 2nd edn (Toronto: University of Toronto Press)
- [21] Zipf G K 1949 *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology* (Cambridge, MA: Addison-Wesley)
- [22] Mandelbrot B B 1953 *Communication Theory, the 2nd London Symposium* ed W Jackson (London: Butterworth) pp 486–504
- [23] Mandelbrot B B 1977 *The Fractal Geometry of Nature* (New York: Freeman)
- [24] Amaral L A N, Scala A, Barthélémy M and Stanley H E 2000 *Proc. Natl Acad. Sci. USA* **97** 11149–52
- [25] Dawkins R 2006 *The Selfish Gene 30th Anniversary Edition* (Oxford: Oxford University Press)
- [26] Blackmore S 1999 *The Meme Machine* (Oxford: Oxford University Press)
- [27] Ramezanpour A 2004 *Europhys. Lett.* **68** 316–22
- [28] Kinouchi O and Copelli M 2005 *Nat. Phys.* **2** 348–52
- [29] Drossel B 2001 *Adv. Phys.* **50** 209–95
- [30] Bak P and Sneppen K 1993 *Phys. Rev. Lett.* **71** 4083–6
- [31] de Oliveira P M C, Sá Martins J S, Stauffer D and Moss de Oliveira S 2004 *Phys. Rev. E* **70** 051910
- [32] Nowak M A 2006 *Science* **314** 1560–3
- [33] Szolnoki A and Perc M 2008 *New J. Phys.* **10** 043036