



Cite this article: Perc M. 2014 The Matthew effect in empirical data. *J. R. Soc. Interface* **11**: 20140378.
<http://dx.doi.org/10.1098/rsif.2014.0378>

Received: 10 April 2014

Accepted: 10 June 2014

Subject Areas:

biocomplexity, mathematical physics

Keywords:

Matthew effect, preferential attachment, cumulative advantage, self-organization, power law, empirical data

Author for correspondence:

Matjaž Perc

e-mail: matjaz.perc@uni-mb.si

The Matthew effect in empirical data

Matjaž Perc

Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška cesta 160, 2000 Maribor, Slovenia

MP, 0000-0002-3087-541X

The Matthew effect describes the phenomenon that in societies, the rich tend to get richer and the potent even more powerful. It is closely related to the concept of preferential attachment in network science, where the more connected nodes are destined to acquire many more links in the future than the auxiliary nodes. Cumulative advantage and success-breeds-success also both describe the fact that advantage tends to beget further advantage. The concept is behind the many power laws and scaling behaviour in empirical data, and it is at the heart of self-organization across social and natural sciences. Here, we review the methodology for measuring preferential attachment in empirical data, as well as the observations of the Matthew effect in patterns of scientific collaboration, socio-technical and biological networks, the propagation of citations, the emergence of scientific progress and impact, career longevity, the evolution of common English words and phrases, as well as in education and brain development. We also discuss whether the Matthew effect is due to chance or optimization, for example related to homophily in social systems or efficacy in technological systems, and we outline possible directions for future research.

1. Introduction

The Gospel of St Matthew states: ‘For to all those who have, more will be given’ (Matthew 25:29). Roughly, two millennia later, sociologist Robert K. Merton [1] was inspired by this writing and coined ‘the Matthew effect’ for explaining discrepancies in recognition received by eminent scientists and unknown researchers for similar work. A few years earlier, physicist and information scientist Derek J. de Solla Price [2] actually observed the same phenomenon when studying the network of citations between scientific papers, only that he used the phrase cumulative advantage for the description. The concept today is in use to describe the general pattern of self-reinforcing inequality related to economic wealth, political power, prestige, knowledge or in fact any other scarce or valued resource [3]. And it is this type of robust self-organization that goes beyond the particularities of individual systems that frequently gives rise to a power law, where the probability of measuring a particular value of some quantity varies inversely as a power of that value [4]. Power laws appear widely in physics, biology, Earth and planetary sciences, economics and finance, computer science, demography and the social sciences [5–8]. Although there is no single origin of power-law behaviour—many theories and models have in fact been proposed to explain it [9–27]—a strong case can be made for the Matthew effect being responsible in many cases. The purpose of this review is to systematically survey research reporting the Matthew effect in empirical data.

In fairness, the Matthew effect has close ties with several other concepts in the social and natural sciences, and it is debateable whether the name we use predominantly throughout this review is the most fitting. The Yule process, inspired by observations of the statistics of biological taxa [28], was in fact the first in a line of widely applicable and closely related mechanisms for generating power laws that relied fundamentally on the assumption that an initially small advantage in numbers may snowball over time [29]. The Gibrat law of proportional growth [30], inspired by the assumption that the size of an enterprise and its growth are

interdependent, also predates the formal introduction of the Matthew effect. Based on the rule of proportional growth, Simon [31] articulated a stochastic growth model with new entrants to account for the Zipf law [4]. The concept of proportional growth has also been elaborated upon thoroughly in Schumpeter's *The Theory of Economic Development* [32]. In terms of popularity and recent impact, however, preferential attachment would without contest be the most apt terminology to use. Barabási & Albert [16] have reasoned that a new node joining a network can in principle connect to any pre-existing node. However, preferential attachment dictates that its choice will not be entirely random, but linearly biased by the number of links that the pre-existing nodes have with other nodes. This induces a rich-get-richer effect, allowing the more connected nodes to gain more links at the expense of their less-connected counterparts. Hence, over time the large-degree nodes turn into hubs and the probability distribution of the degrees across the entire network follows a power law. Although this set-up is rather frail as any non-linearity in the attachment rate may either eliminate the hubs or generate superhubs [33,34], the concept of preferential attachment, along with the 'small-world' model by Watts & Strogatz [35], undoubtedly helped usher in the era of network science [36–51].

We use the 'Matthew effect' terminology for practical reasons and to honour the historical account of events, even though the famous writing in the Gospel of St Matthew might have had significantly different meaning at the time. It was suggested that 'for to all those who have, more will be given' implied spiritual growth and the development of talents, rather than today's more materialist the 'rich-get-richer and the poor-get-poorer' understanding [3]. However, in present times, the Matthew effect is appreciated also in education [52], so some of the original meaning has apparently been preserved. Whatever the terminology used, the understanding should be that here the Matthew effect stands, at least loosely, for all the aforementioned concepts, including cumulative advantage, proportional growth and preferential attachment. An illustration of the Matthew effect is presented in figure 1.

Already in their seminal work, Barabási & Albert [16] noted that preferential attachment ought to be readily detected in time-resolved data cataloguing network growth. Because of preferential attachment, a node that acquires more connections than another one will increase its connectivity at a higher rate, and thus an initial difference in the connectivity between two nodes will increase further as the network grows, while the degree of individual nodes will grow proportional with the square root of time. This reasoning relates also to the so-called first-mover advantage, which has been found accountable for the remarkable marketing success of certain ahead-of-time products [53], as well as the popular acclaim of forefront scientific research despite the fact that it is often less-thorough than follow-up studies [54]. Scientific collaboration networks, where two researchers are connected if they have published a paper together, were among the first empirical data where the concept of preferential attachments has been put to the test and confirmed [55–60]. Soon to follow were reports of preferential attachment and resulting scaling behaviour in the protein network evolution [61] and the evolution of metabolic networks [62,63], the Internet [57] and World Wide Web [20], the accumulation of citations [57,64–70] and scientific impact

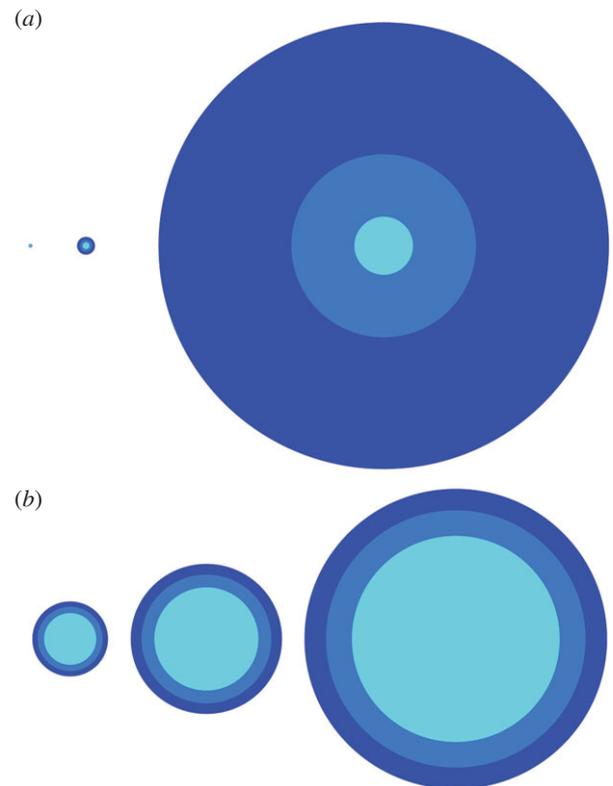


Figure 1. The Matthew effect explained. (a) Starting with three small circles of practically the same size (small dot on the left), over time, the initial differences grow (middle), until eventually they become massive (right). At the beginning, the blue circle has diameter 5, light-blue circle has diameter 4 and the cyan circle has diameter 3. Assuming the growth is proportional to the size, during each time step the circles may become larger by a factor equivalent to their current diameter. After the first time step (middle), this gives us sizes 25, 16 and 9, respectively. Continuing at the same rate, after the second time step (right), we have sizes 625, 256 and 81. Evidently, such a procedure quickly spirals out of easily imaginable bounds. (b) Taking the logarithm of the same diameters over time (and multiplying by 150 for visualization purposes only) reveals that, on the log scale, all the circles grow in diameter linearly by a factor of 2 during each time step from left to right, and the initial relative differences in size remain unchanged over time. This preservation of proportions in logarithmic size manifests as a straight line on a log–log scale—a power-law distribution. In the depicted schematic example, the diameter of the circles can represent anything, from the initial number of collaborators to literacy during formative years. (Online version in colour.)

[71,72], the making of new friends and the evolution of socio-technical networks [57,73–79], population and city size growth [14,15,80], the evolution of source code [81] and the most common English words and phrases [82], in sexual networks [83], as well as the longevity of one's career [84], to name but a few examples. Quantitatively less supported but nevertheless plausible arguments in favour of the Matthew effect also come from education, where there is evidence that early deficiencies in literacy may breed lifelong problems in learning new skills [52], as well as from cognitive neuroscience, where it was hypothesized that the effect could be exploited by means of interventions aimed at improving the brain development of children with low socioeconomic status [85]. We will review observations of the Matthew effect in empirical data thoroughly in the subsequent sections, but first we survey the methodology that is commonly employed for measuring preferential attachment.

2. Measuring preferential attachment

The observation of a power law in empirical data [8] might be an indication for the Matthew effect. Importantly, not finding a power-law distribution or at least a related fat-tailed distribution will falsify the Matthew effect, but the opposite does not necessarily hold. Observing a power-law distribution is consistent with the Matthew effect, but indeed many other processes can also generate power-law distributions [6,7,86]. The probability distribution of a quantity x that obeys a power law is

$$p(x) \sim \frac{1}{x^{1+\mu}} \quad \text{with } \mu > 0, \quad (2.1)$$

where $\alpha = 1 + \mu$ is the scaling parameter. As data in the tail ends of power-law distributions are usually very sparse, one has to be careful with the fitting. The usage of maximum-likelihood fitting methods and goodness-of-fit tests based on Kolmogorov–Smirnov statistics is warmly recommended [8]. Beforehand, there are two ways to get rid of the noise in the tail, at least visually. One option is to bin the data logarithmically, so that the bins appear evenly spaced on a log scale. The second is to use a cumulative distribution function $q(x) \sim 1/x^\mu$, which gives the probability that the quantity is equal to or larger than x . In addition to the fact that the later alleviates statistical fluctuations and does not obscure data as do exponentially wider bins, cumulative distributions can also be used to decide on the presence of a power law. Namely, if the probability density function is a power law with the scaling parameter α , then the cumulative distribution function should also be a power law, but with an exponent $\alpha - 1$. On the other hand, if the probability density function is exponential, the cumulative distribution function will also be exponential, but with the same exponent.

In general, to qualify as a suitable description of empirical data, the probability density function $p(x) \sim 1/x^{1+\mu}$ should hold within a sufficiently large range of x values, extending over at least two or three decades. It is also advisable that one understands the origin of the deviations from the power law, which often appear at both ends of the distribution. It is also worth pointing out that for $\mu = 1$, the power-law distribution is commonly referred to as the Zipf law [4], while the cumulative distribution function is the Pareto law [87,88]. The $\mu = 1$ case is special because it is at the borderline between the converging and diverging unconditional mean of x . While many different physical mechanisms may be at the origin of power laws in complex systems, yielding possibly widely different exponents μ [6,7,86], preferential attachment is certainly one viable candidate.

Measuring preferential attachment, however, requires time-resolved data. We need to be able to measure the rate at which all the entities (nodes, papers and people) that make up the studied system acquire the measured quantity x (links, citations and wealth). Assuming the change in x over a short time interval Δt is Δx , the mechanism of preferential attachment assumes that

$$\Delta x \sim Ax^\gamma, \quad (2.2)$$

where A is the attachment rate and γ determines the non-linearity of the attachment kernel x^γ . The attachment rate A is time-dependent. In particular, the key assumption underlying the Matthew effect is that A grows proportionally

with the growing value of x , as schematically depicted in figure 1. However, the preferential attachment mechanism will yield a power-law distribution of x values given by equation (2.1) only if $\gamma = 1$, when the attachment kernel is linear [16]. Deviations of γ below or above 1 yield sub-linear and super-linear preferential attachment, respectively. Sub-linear preferential attachment gives rise to a stretched exponential cut-off, while $\gamma > 1$ eventually results in a single entity of the system gaining complete monopoly [33,34]. In the language of growing networks, $\gamma > 1$ implies that a single node will over time connect to nearly all other available nodes, while for the accumulation of citations to scientific papers, the super-linear autocatalytic growth may give rise to immortality by means of a dynamical phase transition that leads to the divergence of the citation lifetime of highly cited papers [70,89]. The differences, created by different forms of preferential attachment, can be spotted at a glance in the structure of the resulting networks, as shown in figure 2.

A direct application of equation (2.2) is problematic because growth governed by preferential attachment is an inherently stochastic process. This statement does not necessarily refer to the origin of preferential attachment—which is subject to a slowly evolving but very interesting debate on whether the Matthew effect is due to dumb luck or optimization [90]—but simply to the fact that, regardless of the origin, there will inevitably be strong irregularities in the way x grows over time for each particular entity of the system. In fact, already Yule's theory of power-law distributions in taxonomic groups [29] and Champernowne's theory of stochastic recurrence equations [91] showed that there are important links between the Zipf law [4] and stochastic growth. More specifically, the autocatalytic growth model actually has the form

$$dx \sim \lambda dt + \sigma dW, \quad (2.3)$$

where $\lambda = Ax^\gamma = \langle \Delta x \rangle / \Delta t$ is the average deterministic growth rate over the ensemble of entities with the same x (indicated by $\langle \cdot \rangle$). Moreover, dW is an increment of the Wiener process with zero mean and standard deviation σ . Note also that while Δx is a discrete variable and equation (2.2) thus essentially a difference equation, λ is a continuous variable and equation (2.3) a stochastic differential equation. To do away with the stochastic fingerprint of autocatalytic growth and to estimate reliably whether the process is governed by linear attachment, one can either employ cumulation or averaging. Both methods have been used successfully in the past, although there appear to be persuasive arguments in favour of the latter [89].

Cumulation was proposed by Jeong *et al.* [57], who used it to test the concept of preferential attachment in a number of different empirical networks. To perform the cumulation, one simply has to calculate

$$\kappa(x) = \int_0^x \Delta x \, dx, \quad (2.4)$$

where within the integral x is the degree of a node up to a certain time t , and Δx is the increase in the degree of that same node until $t + \Delta t$. The integration is performed over all the nodes that at time t have degree at most x . The sensible expectation is that the stochastic fluctuations in Δx will thereby be averaged out, while the key assumption behind the method is that the resulting value of $\kappa(x)$ is the same as if

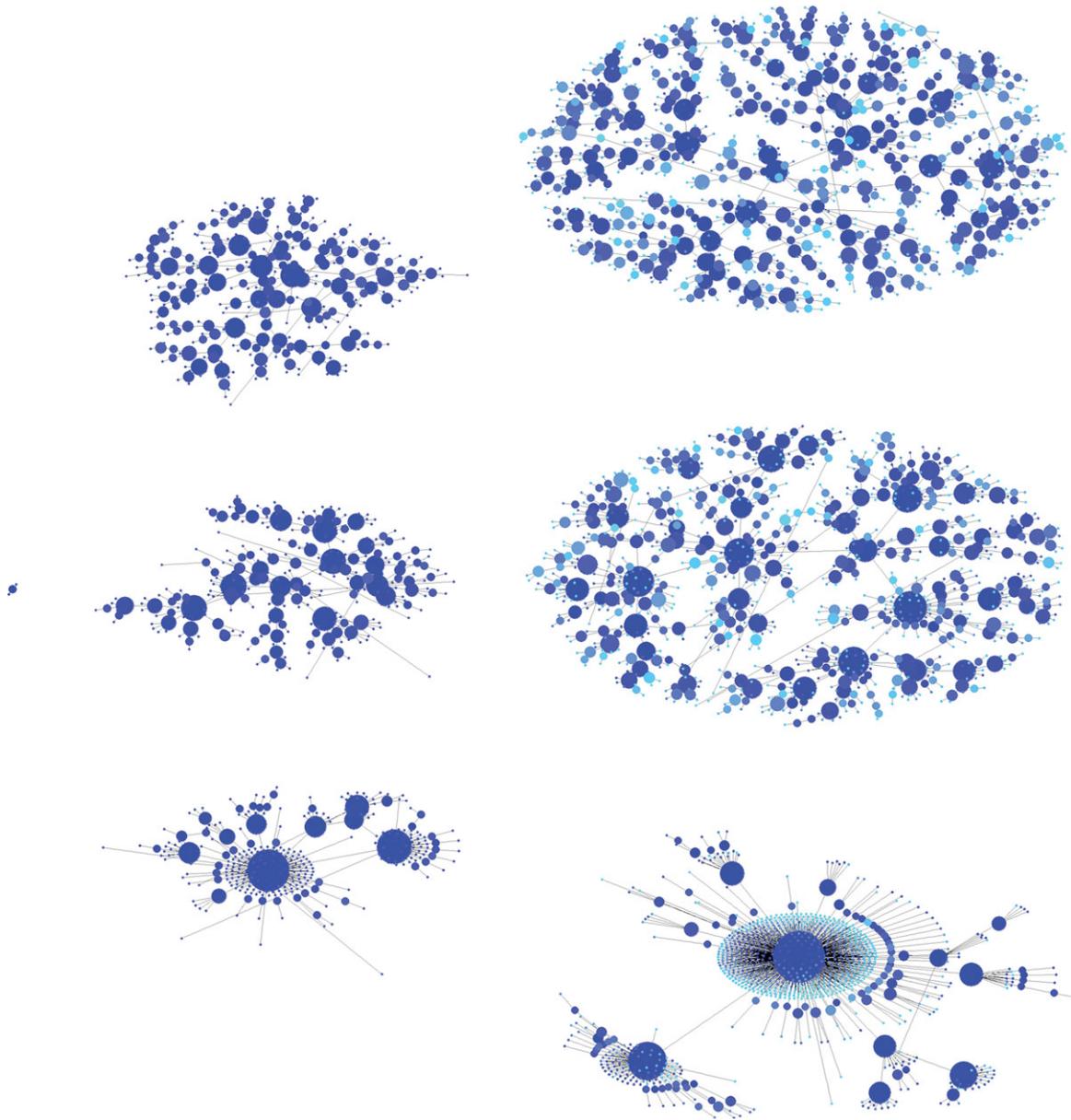


Figure 2. Illustration of network growth by preferential attachment. We start with three nodes, each with a single link to one of the other nodes (small cluster on the left). Subsequently, at each time step, a new nodes arrives and it connects to an existing node with probability proportional to x^γ (see equation (2.2)). Here, x is the degree of nodes. After 300 (centre) and 1000 (right) time steps, sublinear preferential attachment with $\gamma = 0.5$ yields the upper two networks, linear preferential attachment with $\gamma = 1$ yields the middle two networks, while superlinear preferential attachment with $\gamma = 1.5$ yields the lower two networks, respectively. The size and colour (from cyan to blue) of the nodes correspond to their degree in log scale. Sublinear preferential attachment gives rise to a stretched exponential cut-off, thus resulting in somewhat more homogeneous networks than linear preferential attachment. Visually, however, the differences are relatively subtle. Superlinear preferential, on the other hand, clearly favours the emergence of ‘superhubs’, which attract almost all the nodes forming the network. The complete time evolution of the three networks can be viewed at <http://youtu.be/XcGn2KYEmVM>, <http://youtu.be/kfuD53o1yKQ> and <http://youtu.be/vB8yl-WrlRg> for $\gamma = 0.5$, $\gamma = 1$ and $\gamma = 1.5$, respectively. Videos for $\gamma = 0.25$ and $\gamma = 2$, corresponding to even more extreme sublinear and superlinear preferential attachment, are also available at <http://youtu.be/85pZodfi4VM> and http://youtu.be/85R_AGXk2Ko. (Online version in colour.)

equation (2.3) would be integrated directly over x at a fixed time t . Accordingly, we get

$$\kappa(x) \sim Ax^{\gamma+1} \quad (2.5)$$

from where one can readily estimate both A and γ by fitting $\kappa(x)$ in dependence on x . Naturally, we have used the network terminology above only as an example, while of course the same method can be applied on arbitrary time-resolved data to test for preferential attachment [59,61,65,68,76].

Averaging, on the other hand, was proposed by Newman [55], who studied growth and preferential attachment in scientific collaboration networks. In this case, one simply

bins the data over x , calculates the average growth rate $\lambda = \langle \Delta x \rangle / \Delta t$ for each bin over the ensemble of entities for which x falls within a particular bin (indicated by $\langle \cdot \rangle$) and finally compares the resulting histogram with the prediction of equation (2.3). The application of this method requires that one selects the number of bins to cover the interval of x values, and Δt also need not be the finest time-resolution available in the empirical dataset. One can use Δt that are larger to further smooth out the fluctuations that might be due to small and intermittent increments of x across short time intervals. In general, it should be possible to select the number of bins and Δt such that both A and γ could be fitted based on $\lambda = Ax^\gamma$ when plotting λ in dependence

on x . This method or a variation thereof has been used in [60,64,66,71,74,75,82].

While cumulation and averaging are the most frequently applied methods to measure preferential attachment in empirical data, they are not the only ones available. We refer to Golosovsky & Solomon [89] for an in-depth treatment and comparison of the two methods, as well as for an additional control method to check the internal consistency of averaging and cumulation. An additional self-consistent approach to measure preferential attachment in networks has also been proposed in [92], and more recently Markov chain Monte Carlo methodology has been adopted as well [69]. Interested readers will find further details on how it is possible to improve the measurement of preferential attachment if one is in possession of exceptionally detailed data in [81], while here we proceed with the review of the Matthew effect in empirical data that stem from an impressive array of different systems.

3. Scientific collaboration

We begin with scientific collaboration networks, as they were the first empirical data where the conjectured mechanisms for power-law degree distributions in networks have been put to the test [55–57]. Scientific collaboration networks are a beautiful example of social networks [36,39,93,94], where two researchers are considered connected if they have published a paper together. Notably, for a social network to be representative for what it stands—an account of human interaction—a consistent definition of acquaintance is important. And while it may be challenging to define friendship or an enemy in a consistent and precise manner, scientific collaboration is accurately documented in the final product, thus allowing for a precise definition of connectedness and the construction of the social network.

The study of scientific collaboration has been put into the spotlight by the seminal works of Newman [95–98], who constructed networks of connections among researchers by using data from MEDLINE, the Los Alamos e-Print Archive and NCSTRL. Biomedical research, physics and computer science were thus comprehensively covered, which helped reveal that some of the discovered structural properties of these networks have a high degree of universality that is beyond scientific disciplines, while other properties of patterns of collaboration, on the other hand, are field-specific. Most notably, it was shown that collaboration networks form ‘small worlds’ [97], in which randomly chosen pairs of researchers are typically separated by only a short path of intermediate acquaintances [35]. Moreover, the mean and the distribution of the degree of authors revealed the presence of clustering in the networks, which highlighted a number of apparent differences in collaboration patterns between the different fields. The structure of the social science collaboration network has also been studied [58], revealing that a structurally cohesive core in the social sciences has been growing steadily since the early 1960s.

Practically, simultaneously with the research on the structural properties of scientific collaboration networks, research on the time evolution of scientific collaboration networks has been unfolding as well. In [55], Newman has studied empirically the growth of scientific collaboration networks in physics and biology, employing again data from the Los

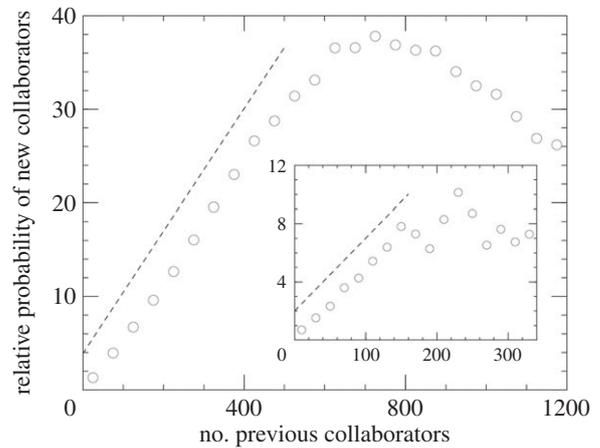


Figure 3. The Matthew effect in scientific collaboration networks. Depicted is the relative probability that a new edge in the network will connect to an author with a given number of previous collaborators. The main panel shows results for the MEDLINE database, whereas the inset shows results for the Los Alamos e-print Archive. In both cases, the relative probability of a new collaborator initially increases linearly with the number of existing collaborators, but there is a field-specific cut-off occurring at around 600 collaborators in biomedicine (main panel) and 150 collaborators in physics (inset), which is related to the fundamental limits of scientific collaboration. (Adapted from [55] with permission from the American Physical Society.)

Alamos e-Print Archive and MEDLINE. It was shown that the probability of a pair of scientists collaborating increases with the number of other collaborators they have in common, and that the probability of a particular scientist acquiring new collaborators increases with the number of his or her past collaborators—a hallmark property of the Matthew effect. As shown in figure 3, which we reproduce from [55], the relative probability of a new collaborator increases practically linearly with the number of existing collaborators. This is particularly true for the initial part of the curve, but since no one can collaborate with an infinite number of people in a finite period of time, the probability falls off as x (here denoting the degree of authors) becomes large. Interestingly, this point appears to be around 150 collaborators in physics (inset) and 600 in biomedicine (main panel), indicating the aforementioned differences in the patterns of collaboration between scientific disciplines.

A closer look at the results presented in figure 3 reveals that the employed averaging method actually yields $\gamma = 1.04$ for MEDLINE and $\gamma = 0.89$ for the Los Alamos e-Print Archive, which in agreement with equation (2.2) corresponds to slightly superlinear and sublinear preferential attachment, respectively. A closely related study that was conducted around the same time by Barabási *et al.* [56], and which was based on all relevant journals in mathematics and neuroscience, also produced evidence for sublinear preferential attachment with $\gamma = 0.8$. The growth of Slovenia’s scientific collaboration network [60] and a co-authorship study based on neuroscience journals [57] also supported the concept of sublinear preferential attachment, both reporting $\gamma = 0.79$. The lowest γ value was reported by Tomassini & Luthi [59], who showed that the time evolution of the genetic programming co-authorship network is governed by $\gamma = 0.76$. However, time-reversing or permutating randomly the order in which the co-authorship networks were constructed within the resolution window Δt yielded $\gamma = 0.88$ and $\gamma = 0.85$, respectively. Taken together, these

results favour the concept of slightly sublinear preferential attachment governing the growth of scientific collaboration networks, but as rightfully pointed out by Newman [55], alternative to linear preferential attachment this difference may have little effect. As shown by Krapivsky *et al.* [33] and reviewed in §2, sublinear preferential attachment gives rise to a stretched exponential cut-off in the resulting degree distribution, but a similar cut-off is already present in the degree distribution as a result of the deviation from linear behaviour for sufficiently large x in figure 3. Indeed, the same deviation has also been reported for the growth of Slovenia's scientific collaboration network [60], thus providing evidence that the sublinear preferential attachment translates fairly accurately into the expected degree distribution.

Irrespective of these details, the overwhelming evidence fully supports the Matthew effect in scientific collaboration networks, indicating that over time initial differences in the number of collaborators are destined to grow and give rise to a strong segregation among authors. Ultimately, some individuals therefore acquire hundreds while others only a handful of collaborators during their scientific career.

4. Socio-technical and biological networks

Scientific collaboration networks reviewed above are obviously also prime examples of social networks and would thus be fit for this section, but we have awarded them a separate section due to their forerunner role in testing preferential attachment in empirical data. There are, however, a number of other socio-technical [57,73–76,78–81] and biological [61,63] networks, where the availability of time-resolved data allowed testing for the Matthew effect. The evolution of socio-technical networks in particular has been in the focus of attention for decades [99]. Recent leaps of progress in the availability of reliable 'big data', mathematical modelling and informatics tools enable increasingly deeper understanding of contagion processes, emerging tipping points, cascading and related nonlinear phenomena that underpin the most interesting characteristics of socio-technical systems [100,101].

The Matthew effect in socio-technical networks was reported first by Jeong *et al.* [57], who at that time also proposed cumulation (see equations (2.4) and (2.5)) to measure preferential attachment in time-resolved data describing network growth. In addition to a scientific collaboration network (§3) and a citation network (§5), they have shown that the evolution of the network of movie actors and the evolution of the autonomous systems forming the Internet are both governed by near-linear preferential attachment. Akin to the definition of a scientific collaboration network, in the movie actor network two actors are connected if they have acted together in a movie. The investigated network was made up of all movies and actors from 1892 till 1999, and it was shown that the growth is characterized by $\gamma = 0.81$. Similarly as by scientific collaboration, here too the slightly sublinear character of preferential attachment can be linked to obvious constraints in the number of co-actors an individual can possibly amass in the course of a lifetime, and this also translates to the expected exponential cut-off in the resulting degree distribution of actors. Notably, preferential attachment in a movie actor network was also reported in [76]. For the Internet, Jeong *et al.* [57] used the data provided

by NLNR, and they have observed slightly superlinear preferential attachment characterized by $\gamma = 1.05$. As evidenced by the examples of network growth depicted in figure 2, however, such small deviations from $\gamma = 1$ lead to hardly recognizable deviations (note that in the depicted examples, we have used $\gamma = 0.5$ for sublinear and $\gamma = 1.5$ for superlinear preferential attachment), and one can thus in good faith conclude to the Matthew effect as a more general description of the mechanism governing the growth of these networks.

In addition to the Internet, the related World Wide Web has also been shown to display striking rich-get-richer behaviour that is driven by the competition of links on the web [20,75]. Interestingly, although the connectivity distribution over the entire web is close to a pure power law, Pennock *et al.* [20] reported that the distribution within sets of category-specific web pages is typically unimodal on a log scale, with the location of the mode, and thus the extent of the rich-get-richer phenomenon, varying across different categories. A simple generative model, incorporating a mixture of preferential and uniform attachment to describe these observations has also been proposed [20].

Online social networks, such as the Internet encyclopaedia Wikipedia [74], bulletin board systems [76], social networking services like Flickr, the obsolete Yahoo! 360° or the now popular Facebook [73,77], as well as longitudinal micro-blogging data [78] also show evidence of the Matthew effect. Wikipedia growth, for example, can be described by local rules such as the preferential attachment mechanism, despite the fact that individual users who are responsible for its evolution can act globally on the network [74]. Research also revealed that triadic closure—if Alice follows Bob and Bob follows Charlie, Alice will follow Charlie—is not such a major mechanism for creating social links in online networks as initially assumed. Longitudinal micro-blogging data reveal more complex strategies that are employed by users when expanding their social circles [78]. In particular, while the network structure affects the spread of information among users, the network is in turn shaped by this communication activity. This suggests a link creation mechanism whereby Alice is more likely to follow Charlie after seeing many messages by Charlie. Weng *et al.* [78] conclude that triadic closure does have a strong effect on link formation, but shortcuts based on traffic are another key factor in interpreting network evolution. Link creation behaviours can be summarized by classifying users in different categories with distinct structural and behavioural characteristics, as shown in figure 4. Users who are popular, active and influential tend to create traffic-based shortcuts, making the information diffusion process more efficient in the network [78]. Notably, the subject of preferential attachment in online networks has recently been surveyed comprehensively in [79], where interested readers will find many further examples and interesting information related specifically to this type of empirical data.

In addition to the vast landscape of online social networks, there are also many socio-technical systems that do not exist solely online, but for which useful data can still be obtained. Rozenfeld *et al.* [80], for example, introduced a method to designate metropolitan areas called the 'City Clustering Algorithm' and used the obtained data to examine the Giblat law of proportional growth [30]. The latter postulates that the mean and standard deviation of the growth rate

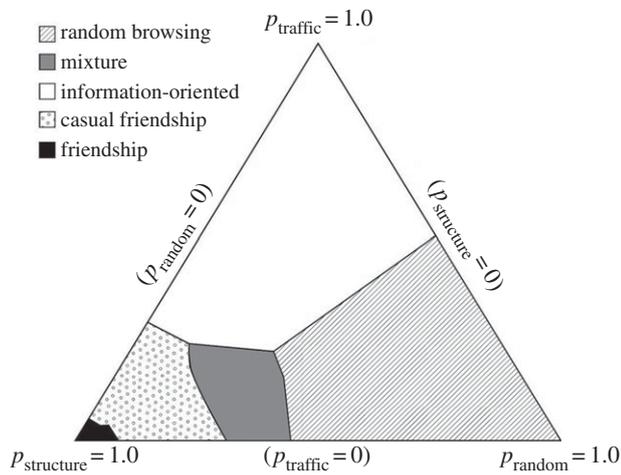


Figure 4. The expansion of online social circles is governed by users that employ many different individual link creation strategies. Indeed, various criteria are taken into account in different proportions when deciding with whom to connect next. The depicted ternary plot encodes the proportions of different link creation strategies for different user types (see legend) in terms of structure ($p_{\text{structure}}$), traffic (p_{traffic}) and chance (p_{random}). Combined, these strategies may give rise to the Matthew effect and lead to strongly heterogeneous social interaction networks. (Adapted from [78] with permission from the ACM.)

of cities are constant, independent of city size. The study revealed that the data deviate from the Gibrat law and that the standard deviation decreases as a power law with respect to the city size. The ‘City Clustering Algorithm’ allowed for the study of the underlying process leading to these deviations, which were shown to arise from the existence of long-range spatial correlations in population growth. Prior to this empirical research, Gabaix [14] and Brakman *et al.* [15] elaborated theoretically on the mechanisms behind city growth, including prominently on the Zipf law.

Maillart *et al.* [81], on the other hand, made use of detailed data on the evolution of open source software projects in Linux distributions. They have showed that the network resulting from the tens of thousands of connected packages precisely obeys the Zipf law over four orders of magnitude, and that this is due to stochastic proportional growth. The study thus delivers a remarkable example of a growing complex self-organizing adaptive system that is subject to the Matthew effect.

Sexual contact networks have also been the subject of research related to the Matthew effect [83,102]. In particular, de Blasio *et al.* [83] have tested the conjecture of preferential attachment by means of a maximum-likelihood estimation-based expectation-maximization fitting technique, which was used to model new partners over a 1-year period based on the number of partners in foregoing periods of 2 and 4 years, as well as the lifetime. The preferential attachment model was modified to account for individual heterogeneity in the inclination to find new partners and fitted to Norwegian survey data on heterosexual men and women. The research revealed sublinear preferential attachment governing the growth of sexual contact networks with $0.5 \leq \gamma \leq 0.7$, which similarly like for scientific collaboration and movie actor networks reviewed above, likely has to do with the physical limits of sexual contacts. Interestingly, the lower value of γ might suggest that the constraints on the maximal feasible number of sexual partners are greater than on the number of collaborators or

co-actors in a movie, thus leading to a stronger exponential cut-off in the corresponding probability distributions—a conclusion that certainly seems to resonate with reality. Moreover, a preceding study by Jones & Handcock [102] concluded that the scaling of sexual degree distributions and the underlying assumption of preferential attachment is actually a very poor fit to the data stemming from several different sexual contact networks. This in turn has important implications for reducing the transmissibility of sexually transmitted diseases, for example by means of condom use or high-activity anti-retroviral therapy, as such interventions could thus bring a population below the epidemic transition, even in populations exhibiting large degrees of behavioural heterogeneity.

To conclude this section, we review examples of the Matthew effect in biological networks, where in relation to the socio-technical networks, the examples are comparatively few. The *Saccharomyces cerevisiae* protein–protein interaction network [103] has a scale-free topology, and Eisenberg & Levanon [61] have shown that the older a protein the better connected it is, and that the number of interactions a protein gains during its evolution is proportional to its connectivity. Thus, by using a cross-genome comparison, the study shows conclusively that the evolution of protein networks is governed by linear preferential attachment. Eisenberg & Levanon [61] go on to conclude that preferential attachment is an important concept in the process of evolution, as it dynamically leads to the formation of big protein complexes and pathways, which introduce high complexity regulation and functionality.

The Matthew effect has also been studied in metabolic networks [62,63], which are at the heart of interactions between biochemical compounds in living cells. Light *et al.* [62] have determined the connectivity patterns of enzymes in the metabolic network of *Escherichia coli*, showing that enzymes that have representatives in eukaryotes have a higher average degree, while enzymes that are represented only in the prokaryotes, and especially the enzymes only present in $\beta\gamma$ -proteobacteria, have a lower degree than expected by chance. More importantly, the research revealed that new edges are added to the highly connected enzymes at a faster rate than to the enzymes with low degree, which is consistent with the Matthew effect. The proposed biological explanation for the observed preferential attachment in the growth of metabolic networks was that novel enzymes created through gene duplication maintain some of the compounds involved in the original reaction throughout its future evolution. Although it remains a major challenge in biology to understand the causes and consequences of the specific design of metabolic networks, Pfeiffer *et al.* [63] have shown that the reported empirical observations, in particular the characteristic presence of hub metabolites such as ATP or NADH, could be explained by computer simulations that initially involve only a few multifunctional enzymes. Then, through the selection of growth rates governed by essential biochemical mechanisms, hubs emerge spontaneously through the process of enzyme duplication and specialization.

5. Citations

After the rather extensive but hopefully interesting departure from scientific collaboration networks to socio-technical and biological networks, we may refocus on research, in

particular on the accumulation of citations to scientific papers. Researchers seem to delight in meticulously evaluating their scientific output and its impact. From citation distributions [104–109], co-authorship networks [98] and the formation of research teams [110,111], to the ranking of researchers [112–114] and the predictability of their success [72,115–117]—how we do science has become a science in its own right. Not surprisingly, the patterns of citation accumulation have been, just like the evolution and structure of scientific collaboration networks, studied extensively during the past decade [57,64–68,70,89].

Notwithstanding the seminal observations by Robert K. Merton [1], who actually introduced the Matthew effect based on the discrepancies in recognition received by eminent scientists and unknown researchers for similar discoveries, and the work by Derek J. de Solla Price [2], who was studying the network of citations between scientific papers already in the early 1960s, the first more rigorous test of preferential attachment in the accumulation of citations is again due to Jeong *et al.* [57]. They have shown that the citations to papers published in the *Physical Review Letters* since 1989 accumulate by means of slightly sublinear preferential attachment with $\gamma = 0.95$. Soon thereafter, Redner [64] conducted an analysis of the entire citation history of publications of *Physical Review*, at the time spanning 110 years, and also confirmed that linear preferential attachment appears to account for the propagation of citations. At closer inspection, the analysis even hinted towards slightly superlinear accumulation, although this, as well as the prospect of strictly linear preferential attachment, was in disagreement with the reported lognormal distribution of citations. Two papers by Wang *et al.* [66,67], using as empirical data citations to papers published in the *Journal of Applied Physics* between 1931 and 2005, the *Journal of Experimental Medicine* between 1900 and 2005 and the *IEEE Transactions on Automatic Control* between 1963 and 2005, delivered essentially the same results, reporting $\gamma \approx 1$ to govern the accumulation of citations. Eom & Fortunato [68] also used the full publication history of the *Physical Review* minus *Reviews of Modern Physics* to study the evolution of citation networks, and they have proposed a linear preferential attachment model with time-dependent initial attractiveness that successfully reproduces the empirical citation distributions as well as accounts for the presence of observed citation bursts.

Importantly, the accumulation of citations to scientific papers has recently been revisited by Golosovsky & Solomon [70], who confirmed the hints reported already by Redner [64], namely that the citation dynamics is nevertheless governed by superlinear preferential attachment with $1.25 \leq \gamma \leq 1.3$. The research used as data the citation history of 40 195 physics papers published in 1 year, and it was emphasized that the citation process cannot be described as a memoryless Markov chain as there is a substantial correlation between the present and recent citation rates to a paper. Based on these observations, a stochastic dynamical model of a growing citation network based on a self-exciting point process has been proposed, and it was demonstrated that it accounts perfectly for the measured citation distributions. An intriguing consequence of this result is that the superlinear autocatalytic growth conveys immortality to highly cited papers by means of a dynamical phase transition that leads to the divergence of the citation lifetime—in the language of epidemiology, these papers become endemic [70,89].

Lending further support to the conclusions of Golosovsky & Solomon [70] are several preceding accounts of superlinear preferential attachment in the accumulation of citations, however not to scientific papers, but rather to patents [65,69]. Valverde *et al.* [65], for example, studied the patent citation network resulting from the patents registered by the US Patent and Trademark Office, and in the light of similarities with article citation networks, concluded towards a universal type of mechanism that links ideas, designs as well as their evolution. This mechanism can be broadly classified as the Matthew effect, which governs how credit is amassed by research as well as technological innovations.

Notably, the subject of preferential attachment in the accumulation of citations has recently been surveyed comprehensively in [89], where interested readers will find further interesting information related specifically to this type of empirical data.

6. Scientific progress and impact

The Matthew effect in the evolution of scientific collaboration networks and in the propagation of citations begets the question whether scientific progress and impact in general might be subject to the same effect. The increasing availability of vast amounts of digitized data, in particular massive databases of scanned books [118] as well as electronic publication and informatics archives [119], fuel large-scale explorations of the human culture that were unimaginable even a decade ago. And since science is central to many key pillars of the human culture, the science of science is scaling up massively as well, with studies on World citation and collaboration networks [120], the global analysis of the ‘scientific food web’ [121], and the identification of phylomemetic patterns in science evolution [122], culminating in the visually compelling atlases of science [123] and knowledge [124].

Riding on the wave of increasing availability of digitized data is also the study of scientific impact, which is gaining momentum rapidly [72,116,117,125,126]. Recent research has revealed, for example, that there is ‘no bad publicity’ in science as criticized papers are in fact highly impactful [125], and that atypical combinations in science have a higher chance to make a big impact [126]. Clear limits have also been established on the predictability of future impact in science [116,117], contrary to the overly optimistic predictions reported earlier [115]. Wang *et al.* [72] have recently proposed a mechanistic model for the quantification of long-term scientific impact, which allows the collapse of the citation histories of papers from different journals and disciplines into a single curve, indicating that all papers tend to follow the same universal temporal pattern. The study revealed that the proposed lognormal model without preferential attachment is able to correctly capture only the citation history of small impact papers, while the modelling of the citation patterns of medium- and high-impact papers requires preferential attachment be turned on. In fact, the model has enabled the team to make an analytical prediction of the citation threshold when preferential attachment becomes relevant, which was reported to equal 8.5 [72]. Hence, the impact of papers that surpass this threshold will benefit from the Matthew effect, while papers with fewer citations will not. Wang *et al.* [72] also emphasized that the reported analytical prediction is in close agreement with the empirical

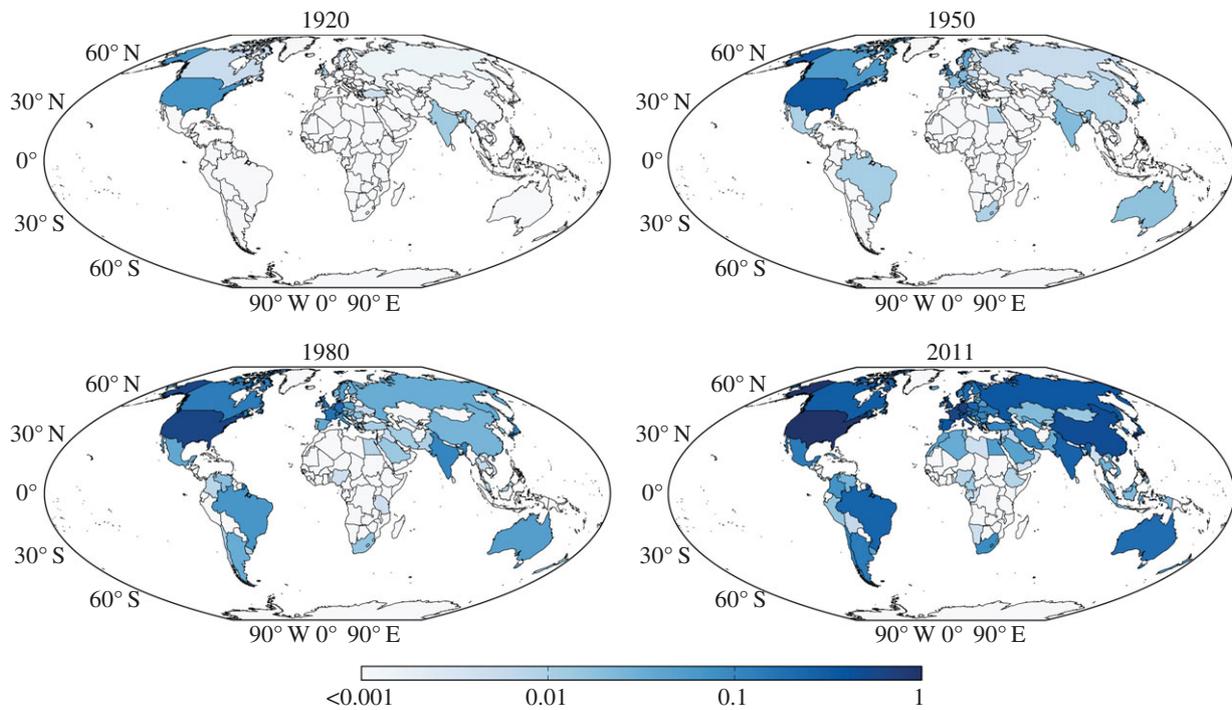


Figure 5. Countries that contribute to research that is published in the *Physical Review*. Colour encodes the average monthly productivity of a country during each displayed year, normalized by the average monthly output of the USA during 2011 (equalling ≈ 565 publications per month—a maximum). All affiliations were used, and in case more than one country was involved on a given publication, all received equal credit. A 12-month moving average was applied prior to calculating the average monthly production for each country. Note that the colour scale is logarithmic. Displayed are World maps for four representative years, while the full geographical timeline can be viewed at <http://youtu.be/0XeySi-EfZs>. (Adapted from [71].) (Online version in colour.)

finding that preferential attachment is masked by initial attractiveness for papers with fewer than seven citations, as reported earlier by Eom & Fortunato [68].

The availability of digitized text, however, enables also the observation of the textual extension of the Matthew effect in citation rates, or alternatively, the large-scale ‘semantic’ version of the Matthew effect in science [71]. By using information provided in the titles and abstracts of over half a million publications that were published by the American Physical Society during the past 119 years, and by identifying all unique words and phrases and determining their monthly usage patterns, it is possible to obtain quantifiable insights into the trends of physics discovery from the end of the nineteenth century to today (the n -gram viewer for publications of the American Physical Society is available at <http://www.matjazperc.com/aps>). The research revealed that the magnitudes of upward and downward trends yield heavy-tailed distributions, and that their emergence is due to the Matthew effect. This indicates that both the rise and fall of scientific paradigms is driven by robust principles of self-organization, which over time yield large differences in the impact particular discoveries have on subsequent progress. Similar research has also been conducted by Pfeiffer & Hoffmann [127], who analysed the temporal patterns of genes in scientific publications hosted by PubMed. They observed that researchers predominantly publish on genes that already appeared in many publications. This might be a rewarding strategy for researchers, because there is an obvious positive correlation between the frequency of a gene in scientific publications and the impact of these publications [127]. In a way, the Matthew effect can thus be engineered, or at least facilitated, by focusing on the ‘hot topics’ in a specific field of research.

Figure 5 reveals that the Matthew effect in the impact of scientific research translates also to geography [71], where

the USA and large contingents of Europe were able to set the pace in the production of physics research over extended periods of time, interrupted only by periods of war. The collapse of the Soviet Union, the fall of the Berlin Wall and the related changes in World order during the 1980s and 1990s, however, contributed significantly to the globalization, so that today countries like China, Russia, South America and Australia all contribute markedly to the production of physics. However, a beautiful citation map of the world produced by Pan *et al.* [120], where the area of each country is scaled and deformed according to the number of citations received, still reveals a strongly biased geographical distribution of impact. Notably, an in-depth analysis of the scientific production and consumption of physics revealed that even cities can be pinpointed based on their leading positions for scholarly research [128]. Although for now research along this line seems to be focused predominantly on physics, the applied methodology certainly opens up the possibility for comparative studies across different disciplines and research areas, where the Matthew effect is still to be either confirmed or refuted.

7. Career longevity

The overwhelming evidence in favour of the Matthew effect in science, affecting the patterns of collaboration, the propagation of citations and ultimately also scientific progress and impact, probably make it little surprising that the same effect affects also career longevity. Importantly, not just the longevity of scientific careers, but also the longevity of careers in professional sport, as demonstrated in [84].

Career longevity is a fundamental metric that influences the overall legacy of an employee, because for most

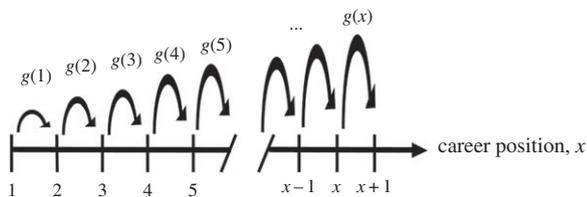


Figure 6. The Matthew effect in professional careers. Progress from career position x to career position $x + 1$ is made with a position-dependent progress rate $g(x) = 1 - \exp[-(x/x_c)^\gamma]$, which increases from approximately zero and asymptotically approaches one over a characteristic time interval x_c . For $x \ll x_c$, the progress rate corresponds to $g(x) \sim x^\gamma$, which for $\gamma = 1$ is the traditional ansatz for linear preferential attachment (see equation (2.2)). As the progress rate increases with increasing x , the essence of the Matthew effect is taken into account in that it becomes easier to make progress the further along the career an individual is. (Adapted from [84].)

individuals the measure of success is closely related to the length of their career. In particular, the more successful an individual, the longer his or her career is going to last. Using this as motivation, Petersen *et al.* [84] analysed publication careers within six high-impact journals, including *Nature*, *Science*, *Proceedings of the National Academy of Sciences*, *Physical Review Letters*, *New England Journal of Medicine* and *Cell*, as well as sports careers within four different leagues, including Major League Baseball, Korean Professional Baseball, the National Basketball Association and the English Premier League. The conducted research delivered testable evidence in favour of the Matthew effect, wherein the longevity and past success of an individual lead to a cumulative advantage in further developing his or her career [84]. From the methodological point of view, it is worth pointing out that for science and professional sports, there exist well-defined metrics that quantify career longevity, success and prowess, which together enable a relatively clear and unbiased assessment of the overall success of each individual employee. In many other professions, however, these criteria are significantly more vague, and thus the same research agenda could be difficult to execute.

To support their quantitative demonstration of the Matthew effect in career longevity, Petersen *et al.* [84] also developed an exactly solvable stochastic career progress model, which is schematically illustrated and summarized in figure 6. Model predictions have been validated on the careers of 400 000 scientists and 20 000 professional athletes. The authors emphasized the importance of early career development, showing that many careers are stunted by the relative disadvantage associated with inexperience. This is closely related to the workings of the Matthew effect in education (§9), where tests suggest that falling behind in literacy during formative primary school years creates disadvantages that may be difficult to compensate all the way to adulthood [52].

8. Common words and phrases

Moving away from scientific production and impact for good, in this section, we review recent research related to the evolution of the most common English words and phrases [82]. Already during the 1960s, the economist Herbert Simon and the mathematician Benoît Mandelbrot had a dispute over the origin of the power-law distribution

of word frequencies in text [4,129–133]. Simon defended the role of randomness and preferential attachment, while Mandelbrot argued in favour of an optimization framework [134]. The original proposal made by Zipf, on the other hand, was that there is tension between the efforts of the speaker and the listener, and it has been shown by means of mathematical modelling that this may indeed explain the origins of scaling in the usage of words [135]. The ecophysics of language change [136]—the application of models from statistical physics and theoretical ecology to the study of language dynamics—has since evolved into a beautiful and vibrant avenue of research [137–144].

A direct test for preferential attachment in the evolution of the most common English words and phrases [82] was made possible by the work of Michel *et al.* [118], which was accompanied by the release of a vast amount of data comprised metrics derived from approximately 4% of books ever published. Raw data, along with usage instructions, are available and updated at <http://books.google.com/ngrams/datasets> as counts of n -grams that appeared in various book corpora over the past centuries with a yearly resolution. By recursively scanning all the files from the English corpus in the search for those n -grams that had the highest usage frequency in any given year, it is possible to determine the most common English words and phrases with a yearly resolution. Tables listing the top 100, top 1000 and top 10 000 n -grams for all available years since 1520 inclusive, along with their yearly usage frequencies and direct links to the Google Books Ngram Viewer, are available at <http://www.matjazperc.com/ngrams>. From this, it is possible to derive evidence in favour of preferential attachment as shown in figure 7, which indicate that the higher the number of occurrences of any given n -gram, the higher the probability that it will occur even more frequently in the future. More precisely, for the past two centuries, the points quantifying the attachment rate follow a linear dependence, thus confirming that the Matthew effect is behind the power-law distribution of word frequencies in text, as argued by Herbert Simon. Evidently, this does not rule out an optimization framework that was favoured by Benoît Mandelbrot, as preferential attachment itself might be the outcome of optimization [24,27,90].

Somewhat related to the study of the most common English words and phrases is also the study of popular memes, which has recently attracted considerable attention [145–151]. According to Dawkins, memes are the cultural equivalent of genes that spread across the human culture by means of imitation [152]. The competition among memes has been studied by Weng *et al.* [150], who by means of an agent-based model accounting for the dynamics of information diffusion, showed that in a world with limited attention only a few memes go viral while most do not. These predictions are consistent with empirical data from Twitter, and they explain the massive heterogeneity in the popularity and persistence of memes as deriving from a combination of the competition for our limited attention and the structure of the social network, without the need to assume different intrinsic values among ideas [150]. The study of how memes compete with each other for the limited and fluctuating resource of user attention has also amassed the attention of physicists, who showed that the competition between memes can bring a social network to the brink of criticality [153], where even minute disturbances can lead to avalanches of events that make a certain meme go viral [151].

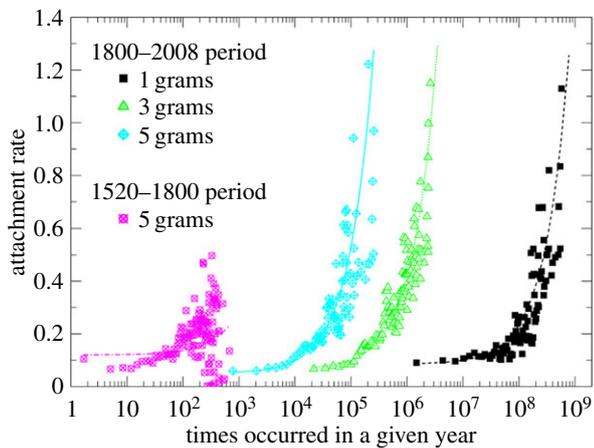


Figure 7. Emergence of linear preferential attachment in the evolution of the most common English words and phrases during the past two centuries. Two time periods were considered separately, as indicated in the figure legend. While preferential attachment appears to have been in place already during the 1520–1800 period, large deviations from the linear dependence (the goodness-of-fit is ≈ 0.05) hint towards inconsistencies that may have resulted in heavily fluctuated rankings. The same analysis for the nineteenth and the twentieth century provides much more conclusive results. For all n , the data fall nicely onto straight lines (the goodness-of-fit is ≈ 0.8), thus indicating that the Matthew effect might have shaped the large-scale organization of the writing of English books over the past two centuries. (Adapted from [82].) (Online version in colour.)

9. Education and beyond

In addition to the above-reviewed examples of the Matthew effect in empirical data, there exist many more, related for example to education [52] and brain development [85], which we here review in passing for a more complete coverage of the subject.

In his synthesis titled *Matthew effects in reading: some consequences of individual differences in the acquisition of literacy* [52], Stanovich presents a framework for conceptualizing the development of individual differences in reading ability, with special emphasis on the concepts of reciprocal relationships—situations where the causal connection between reading ability and the efficiency of a cognitive process is bidirectional, and on organism–environment correlation—the fact that differentially advantaged organisms are exposed to non-random distributions of environmental quality. Foremost, it is explained how these mechanisms operate to create the rich-get-richer and the poor-get-poorer patterns of reading achievement, and the framework is used to explicate some persisting problems in the literature on reading disability and to conceptualize remediation efforts in reading. Owing to the Matthew effect, early deficiencies in literacy may breed lifelong problems in learning new skills, and falling behind during formative primary school years may create disadvantages that could be difficult to compensate all the way to adulthood [52]. It must be noted, however, that the degree to which the Matthew effect actually holds true in reading development is a topic of considerable debate [154–156].

The review by Raizada & Kishiyama [85] on the effects of socioeconomic status on brain development also draws on the Matthew effect, in particular as a potential triggering mechanism for a long-term self-reinforcing trend in training executive function in young children, with improved self-control enabling greater attentiveness and learning, which

would in turn help to make a child's educational experiences more rewarding, thereby facilitating yet more intellectual growth. The authors are sceptical about this rather 'rosy-sounding' scenario, but note that specific interventions aimed at improving the cognitive development of children with low socioeconomic status may well trigger the desired effect. Indeed, Cohen *et al.* [157,158] have shown that even brief self-affirmation writing assignments aimed at reducing feelings of academic threat in ethnic minority high-school students had the effect of producing significant improvements in grade-point average, which endured over a period of 2 years—a potential indication that the Matthew effect might have kicked in.

As noted in the Introduction, the concept today is in wide use to describe the general pattern of self-reinforcing inequality that can be related to economic wealth, political power, prestige and stardom. Although these examples are to a degree rooted in folktales and lack firm quantitative support, they can nevertheless be supported by plausible arguments in favour of the Matthew effect. Being born into poverty, for example, greatly increases the probability of remaining poor, and each further disadvantage makes it increasingly difficult to escape the economic undertow. The Matthew effect also contributes to a number of other concepts in the social sciences that may be broadly characterized as social spirals. Economists speak of inflationary spirals, spiralling unemployment and spiralling debt. These spirals exemplify positive feedback loops, in which processes feed upon themselves in such a way as to cause nonlinear patterns of growth. To make a complete account of such examples exceeds the scope of this review, and so we are content to draw from the recent book *The Matthew effect: how advantage begets further advantage* by Rigney [3], which we warmly recommend to interested readers.

10. Discussion

As we hope this review shows, the Matthew effect is puzzling yet ubiquitous across social and natural sciences. It affects patterns of scientific collaboration, the growth of socio-technical and biological networks, the propagation of citations, scientific progress and impact, career longevity, the evolution of the most common words and phrases, education, as well as many other aspects of human culture. The recently acquired prominence of the Matthew effect is largely due to the rise of network science [51], and the concept of preferential attachment in particular [16]. Accordingly, the title of this review might as well have been 'Preferential attachment in empirical data', but since the Matthew effect describes more loosely the general principle that advantage tends to beget further advantage, the age-old Matthew 'rich-get-richer' effect ultimately won the toss.

The theory of evolving networks based on growth and preferential attachment was motivated by extensive empirical evidence documenting the scale-free nature of the degree distribution, from the cell to the World Wide Web, and it was this theory, along with the ever increasing availability of digitized data at the turn of the twenty-first century, that ultimately led to the development of the methodology for measuring preferential attachment and the subsequent application of these methods on a wide variety of complex systems. Although the progress made during the past decade related to data-based mathematical models of

complex systems has been truly remarkable, the data explosion we witness today is surely going to accelerate research along this line even more. Indeed, 'big data' [159] is the keyword for current complex systems research, and the data windfall is also surely going to promote research on the Matthew effect. Especially, data from social media, but also from neuroscience as well as electronic publication and informatics archives, offer many opportunities for fascinating scientific discoveries in the nearest future.

Concepts such as preferential attachment, cumulative advantage and the Matthew effect are at the heart of self-organization in biology and societies, and they give rise to emergent properties that are impossible to understand, let alone predict, at the level of constituent agents. The emergent collective modes of behaviour are due to the heterogeneity of the interaction patterns, the presence of non-linearity and feedback effects, and it is here where the reasons behind the Matthew effect ought to be sought. This, however, raises the question whether the Matthew effect is due to chance or optimization [24,27,90]. While theoretical models in general rely on dumb luck to yield the power laws, in reviewing the subject on empirical data, one finds it difficult to believe that the selection of a collaborator or a sexual partner, or the hiring for a tenure-track position, would be left to chance. These decisions certainly do depend also on unpredictable factors, but predominantly they are nevertheless based on factors such as common appeal, competence and prowess. The argument in favour of randomness gains traction when cognition and reasoning obviously no longer apply—consider the emergence of hubs in protein-interaction networks through gene duplication [160] (see also [61] for a thorough discussion). But more often than not, the line

between chance and thought is much more blurred, like by the propagation of citations. Common sense tells us that credit should be given where credit is due, yet researchers often cite a paper just because it has been cited many times before. An interesting discussion of this was recently delivered by Golosovsky & Solomon [89], who concluded that such spreading of citations and ideas is akin to the epidemiological process [161] and to the copying mechanism [23]. Google Scholar has even been criticized for strengthening the Matthew effect by putting high weight on citation counts in its ranking algorithm [162], by means of which highly cited papers that appear in top positions gain ever more citations while new papers hardly appear in top positions and therefore struggle to amass new citations. Ultimately, one ends up agreeing with Barabási [90], who noted that we do not need to choose between luck and reason in preferential attachment, but simply strive towards a deeper understanding of this puzzling yet ubiquitous force.

The Matthew effect is obviously at the interface of many different fields of research, and while its potential has been realized in the realm of complex systems as being one in a series of fundamental laws that determine and limit their behaviour, the concept deserves also to reach a wider audience and to inform public policy decisions that have an impact on inequality in areas such as taxation, civil rights and public goods [163,164].

Acknowledgements. Special thanks go to Esteban Moro for his insightful tutorial that enabled the creation of figure 2 and the pertaining videos.

Funding statement. This work was supported by the Slovenian Research Agency (Grant P5-0027).

References

- Merton RK. 1968 The Matthew effect in science. *Science* **159**, 53–63. (doi:10.1126/science.159.3810.56)
- de Solla Price DJ. 1965 Networks of scientific papers. *Science* **149**, 510–515. (doi:10.1126/science.149.3683.510)
- Rigney D. 2013 *The Matthew effect: how advantage begets further advantage*. Columbia, UK: Columbia University Press.
- Zipf GK. 1949 *Human behavior and the principle of least-effort*. Reading, MA: Addison-Wesley.
- Reed WJ. 2001 The Pareto, Zipf and other power laws. *Econ. Lett.* **74**, 15–19. (doi:10.1016/S0165-1765(01)00524-9)
- Mitzenmacher M. 2004 A brief history of generative models for power law and lognormal distributions. *Internet Math.* **1**, 226–251. (doi:10.1080/15427951.2004.10129088)
- Newman MEJ. 2005 Power laws, Pareto distributions and Zipf's law. *Contemp. Phys.* **46**, 323–351. (doi:10.1080/00107510500052444)
- Clauset A, Shalizi CR, Newman ME. 2009 Power-law distributions in empirical data. *SIAM Rev.* **51**, 661–703. (doi:10.1137/070710111)
- Miller GA. 1957 Some effects of intermittent silence. *Am. J. Psychol.* **70**, 311–314. (doi:10.2307/1419346)
- Bak P, Tang C, Wiesenfeld K. 1987 Self-organized criticality: an explanation of $1/f$ noise. *Phys. Rev. Lett.* **59**, 381–384. (doi:10.1103/PhysRevLett.59.381)
- Sneppen K, Newman M. 1997 Coherent noise, scale invariance and intermittency in large systems. *Phys. D* **110**, 209–222. (doi:10.1016/S0167-2789(97)00128-0)
- Sornette D. 1998 Multiplicative processes and power laws. *Phys. Rev. E* **57**, 4811–4813. (doi:10.1103/PhysRevE.57.4811)
- Carlson JM, Doyle J. 1999 Highly optimized tolerance: a mechanism for power laws in designed systems. *Phys. Rev. E* **60**, 1412–1427. (doi:10.1103/PhysRevE.60.1412)
- Gabaix X. 1999 Zipf's law for cities: an explanation. *Q. J. Econ.* **114**, 739–767. (doi:10.1162/003355399556133)
- Brakman S, Garretsen H, Van Marrewijk C, Van Den Berg M. 1999 The return of Zipf: towards a further understanding of the rank-size distribution. *J. Reg. Sci.* **39**, 183–213. (doi:10.1111/1467-9787.00129)
- Barabási A-L, Albert R. 1999 Emergence of scaling in random networks. *Science* **286**, 509–512. (doi:10.1126/science.286.5439.509)
- Jin EM, Girvan M, Newman ME. 2001 Structure of growing social networks. *Phys. Rev. E* **64**, 046132. (doi:10.1103/PhysRevE.64.046132)
- Caldarelli G, Capocci A, De Los Rios P, Muñoz MA. 2002 Scale-free networks from varying vertex intrinsic fitness. *Phys. Rev. Lett.* **89**, 258702. (doi:10.1103/PhysRevLett.89.258702)
- Reed WJ, Hughes BD. 2002 From gene families and genera to incomes and internet file sizes: why power laws are so common in nature. *Phys. Rev. E* **66**, 067103. (doi:10.1103/PhysRevE.66.067103)
- Pennock DM, Flake GW, Lawrence S, Glover EJ, Giles CL. 2002 Winners don't take all: characterizing the competition for links on the web. *Proc. Natl Acad. Sci. USA* **99**, 5207–5211. (doi:10.1073/pnas.032085699)
- Vázquez A. 2003 Growing network with local rules: preferential attachment, clustering hierarchy, and degree correlations. *Phys. Rev. E* **67**, 056104. (doi:10.1103/PhysRevE.67.056104)
- Reed WJ. 2003 The Pareto law of incomes—an explanation and an extension. *Phys. A* **319**, 469–486. (doi:10.1016/S0378-4371(02)01507-8)
- Krapivsky P, Redner S. 2005 Network growth by copying. *Phys. Rev. E* **71**, 036118. (doi:10.1103/PhysRevE.71.036118)

24. D'Souza RM, Borgs C, Chayes JT, Berger N, Kleinberg RD. 2007 Emergence of tempered preferential attachment from optimization. *Proc. Natl Acad. Sci. USA* **104**, 6112–6117. (doi:10.1073/pnas.0606779104)
25. Poncela J, Gómez-Gardeñes J, Floría LM, Sánchez A, Moreno Y. 2008 Complex cooperative networks from evolutionary preferential attachment. *PLoS ONE* **3**, e2449. (doi:10.1371/journal.pone.0002449)
26. Baek SK, Bernhardtsson S, Minnhagen P. 2011 Zipf's law unzipped. *New J. Phys.* **13**, 043004. (doi:10.1088/1367-2630/13/4/043004)
27. Papadopoulos F, Kitsak M, Serrano MÁ, Boguná M, Krioukov D. 2012 Popularity versus similarity in growing networks. *Nature* **489**, 537–540. (doi:10.1038/nature11459)
28. Willis JC, Yule GU. 1922 Some statistics of evolution and geographical distribution in plants and animals, and their significance. *Nature* **109**, 177–179. (doi:10.1038/109177a0)
29. Yule GU. 1925 A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis F.R.S. *Phil. Trans. R. Soc. Lond. B* **213**, 21–85. (doi:10.1098/rstb.1925.0002)
30. Gibrat R. 1931 *Les inégalités économiques*. Paris, France: Recueil Sirey.
31. Simon HA. 1955 On a class of skew distribution functions. *Biometrika* **42**, 425–440. (doi:10.1093/biomet/42.3-4.425)
32. Schumpeter JA. 1934 *The theory of economic development*. Piscataway, NJ: Transaction Publishers.
33. Krapivsky PL, Redner S, Leyvraz F. 2000 Connectivity of growing random networks. *Phys. Rev. Lett.* **85**, 4629–4632. (doi:10.1103/PhysRevLett.85.4629)
34. Dorogovtsev SN, Mendes JFF, Samukhin AN. 2000 Structure of growing networks with preferential linking. *Phys. Rev. Lett.* **85**, 4633–4636. (doi:10.1103/PhysRevLett.85.4633)
35. Watts DJ, Strogatz SH. 1998 Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442. (doi:10.1038/30918)
36. Watts DJ. 1999 *Small worlds*. Princeton, NJ: Princeton University Press.
37. Amaral LAN, Scala A, Barthélemy M, Stanley HE. 2000 Classes of small-world networks. *Proc. Natl Acad. Sci. USA* **97**, 11149. (doi:10.1073/pnas.200327197)
38. Albert R, Barabási AL. 2002 Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47–97. (doi:10.1103/RevModPhys.74.47)
39. Barabási A-L. 2002 *Linked: the new science of networks*. Cambridge, UK: Perseus.
40. Dorogovtsev SN, Mendes JF. 2002 Evolution of networks. *Adv. Phys.* **51**, 1079–1187. (doi:10.1080/00018730110112519)
41. Newman MEJ. 2003 The structure and function of complex networks. *SIAM Rev.* **45**, 167–256. (doi:10.1137/S003614450342480)
42. Dorogovtsev SN, Mendes JFF. 2003 *Evolution of networks: from biological nets to the internet and WWW*. Oxford, UK: Oxford University Press.
43. Pastor-Satorras R, Vespignani A. 2004 *Evolution and structure of the internet: a statistical physics approach*. Cambridge, UK: Cambridge University Press.
44. Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang D-U. 2006 Complex networks: structure and dynamics. *Phys. Rep.* **424**, 175–308. (doi:10.1016/j.physrep.2005.10.009)
45. Newman MEJ, Watts DJ, Barabási A-L. 2006 *The structure and dynamics of networks*. Princeton, NJ: Princeton University Press.
46. Dorogovtsev SN, Goltsev AV, Mendes JFF. 2008 Critical phenomena in complex networks. *Rev. Mod. Phys.* **80**, 1275–1335. (doi:10.1103/RevModPhys.80.1275)
47. Barrat A, Barthélemy M, Vespignani A. 2008 *Dynamical processes in complex networks*. Cambridge, UK: Cambridge University Press.
48. Arenas A, Diaz-Guilera A, Kurths J, Moreno Y, Zhou C. 2008 Synchronization in complex networks. *Phys. Rep.* **469**, 93–153. (doi:10.1016/j.physrep.2008.09.002)
49. Fortunato S. 2010 Community detection in graphs. *Phys. Rep.* **486**, 75–174. (doi:10.1016/j.physrep.2009.11.002)
50. Holme P, Saramäki J. 2012 Temporal networks. *Phys. Rep.* **519**, 97–125. (doi:10.1016/j.physrep.2012.03.001)
51. Barabási A-L. 2012 The network takeover. *Nat. Phys.* **8**, 14–16. (doi:10.1038/nphys2188)
52. Stanovich KE. 2008 Matthew effects in reading: some consequences of individual differences in the acquisition of literacy. *J. Educ.* **189**, 23–55.
53. Kerin RA, Varadarajan PR, Peterson RA. 1992 First-mover advantage: a synthesis, conceptual framework, and research propositions. *J. Mark.* **56**, 33–52. (doi:10.2307/1251985)
54. Newman MEJ. 2009 The first-mover advantage in scientific publication. *EPL* **86**, 68001. (doi:10.1209/0295-5075/86/68001)
55. Newman MEJ. 2001 Clustering and preferential attachment in growing networks. *Phys. Rev. E* **64**, 025102. (doi:10.1103/PhysRevE.64.025102)
56. Barabási AL, Jeong H, Néda Z, Ravasz E, Schubert A, Vicsek T. 2002 Evolution of the social network of scientific collaborations. *Phys. A* **311**, 590–614. (doi:10.1016/S0378-4371(02)00736-7)
57. Jeong H, Néda Z, Barabási AL. 2003 Measuring preferential attachment in evolving networks. *Europhys. Lett.* **61**, 567–572. (doi:10.1209/epl/12003-00166-9)
58. Moody J. 2004 The structure of a social science collaboration network: disciplinary cohesion from 1963 to 1999. *Am. Sociol. Rev.* **69**, 213–238. (doi:10.1177/000312240406900204)
59. Tomassini M, Luthi L. 2007 Empirical analysis of the evolution of a scientific collaboration network. *Phys. A* **385**, 750–764. (doi:10.1016/j.physa.2007.07.028)
60. Perc M. 2010 Growth and structure of Slovenia's scientific collaboration network. *J. Inform.* **4**, 475–482. (doi:10.1016/j.joi.2010.04.003)
61. Eisenberg E, Levanon EY. 2003 Preferential attachment in the protein network evolution. *Phys. Rev. Lett.* **91**, 138701. (doi:10.1103/PhysRevLett.91.138701)
62. Light S, Kraulis P, Elofsson A. 2005 Preferential attachment in the evolution of metabolic networks. *BMC Genomics* **6**, 159. (doi:10.1186/1471-2164-6-159)
63. Pfeiffer T, Soyer OS, Bonhoeffer S. 2005 The evolution of connectivity in metabolic networks. *PLoS Biol.* **3**, e228. (doi:10.1371/journal.pbio.0030228)
64. Redner S. 2005 Citation statistics from 110 years of *Physical Review*. *Phys. Today* **58**, 49–54. (doi:10.1063/1.1996475)
65. Valverde S, Solé RV, Bedau MA, Packard N. 2007 Topology and evolution of technology innovation networks. *Phys. Rev. E* **76**, 056118. (doi:10.1103/PhysRevE.76.056118)
66. Wang M, Yu G, Yu D. 2008 Measuring the preferential attachment mechanism in citation networks. *Phys. A* **387**, 4692–4698. (doi:10.1016/j.physa.2008.03.017)
67. Wang M, Yu G, Yu D. 2009 Effect of the age of papers on the preferential attachment in citation networks. *Phys. A* **388**, 4273–4276. (doi:10.1016/j.physa.2009.05.008)
68. Eom Y-H, Fortunato S. 2011 Characterizing and modeling citation dynamics. *PLoS ONE* **6**, e24926. (doi:10.1371/journal.pone.0024926)
69. Sheridan P, Yagahara Y, Shimodaira H. 2012 Measuring preferential attachment in growing networks with missing-timelines using Markov chain Monte Carlo. *Phys. A* **391**, 5031–5040. (doi:10.1016/j.physa.2012.05.041)
70. Golosovsky M, Solomon S. 2012 Stochastic dynamical model of a growing citation network based on a self-exciting point process. *Phys. Rev. Lett.* **109**, 098701. (doi:10.1103/PhysRevLett.109.098701)
71. Perc M. 2013 Self-organization of progress across the century of physics. *Sci. Rep.* **3**, 1720. (doi:10.1038/srep01720)
72. Wang D, Song C, Barabási A-L. 2013 Quantifying long-term scientific impact. *Science* **342**, 127–132. (doi:10.1126/science.1237825)
73. Kumar R, Novak J, Tomkins A. 2006 Structure and evolution of online social networks. In *Proc. ACM SIGKDD, 20–23 August, Philadelphia, PA*, pp. 611–617. New York, NY: ACM.
74. Capocci A, Servidio VD, Colaiori F, Burioni LS, Donato D, Leonardi S, Caldarelli G. 2006 Preferential attachment in the growth of social networks: the internet encyclopedia Wikipedia. *Phys. Rev. E* **74**, 036116. (doi:10.1103/PhysRevE.74.036116)
75. Herdāğdelen A, Aygün E, Bingol H. 2007 A formal treatment of generalized preferential attachment and its empirical validation. *EPL* **78**, 60007. (doi:10.1209/0295-5075/78/60007)
76. Eom Y-H, Jeon C, Jeong H, Kahng B. 2008 Evolution of weighted scale-free networks in empirical data. *Phys. Rev. E* **77**, 056105. (doi:10.1103/PhysRevE.77.056105)
77. Nazir A, Raza S, Chuah C-N. 2008 Unveiling Facebook: a measurement study of social network based applications. In *Proc. ACM SIGCOMM, 20–22 October, Vouliagmeni, Greece*, pp. 43–56. New York, NY: ACM.

78. Weng L, Ratkiewicz J, Perra N, Gonçalves B, Castillo C, Bonchi F, Schifanella R, Menczer F, Flammini A. 2013 The role of information diffusion in the evolution of social networks. In *Proc. ACM SIGKDD, 11–14 August, Chicago, IL*, pp. 356–364. New York, NY: ACM.
79. Kunegis J, Blattner M, Moser C. 2013 Preferential attachment in online networks: measurement and explanations. In *Proc. ACM WSC, 2–4 May, Paris, France*, pp. 205–214. New York, NY: ACM.
80. Rozenfeld HD, Rybski D, Andrade JS, Batty M, Stanley HE, Makse HA. 2008 Laws of population growth. *Proc. Natl Acad. Sci. USA* **105**, 18 702–18 707. (doi:10.1073/pnas.0807435105)
81. Maillart T, Sornette D, Spaeth S, Von Krogh G. 2008 Empirical tests of Zipf's law mechanism in open source Linux distribution. *Phys. Rev. Lett.* **101**, 218701. (doi:10.1103/PhysRevLett.101.218701)
82. Perc M. 2012 Evolution of the most common English words and phrases over the centuries. *J. R. Soc. Interface* **9**, 3323–3328. (doi:10.1098/rsif.2012.0491)
83. de Blasio BF, Svensson Å, Liljeros F. 2007 Preferential attachment in sexual networks. *Proc. Natl Acad. Sci. USA* **104**, 10 762–10 767. (doi:10.1073/pnas.0611337104)
84. Petersen AM, Jung WS, Yang JS, Stanley HE. 2011 Quantitative and empirical demonstration of the Matthew effect in a study of career longevity. *Proc. Natl Acad. Sci. USA* **108**, 18–23. (doi:10.1073/pnas.1016733108)
85. Raizada RD, Kishiyama MM. 2010 Effects of socioeconomic status on brain development, and how cognitive neuroscience may contribute to levelling the playing field. *Front. Hum. Neurosci.* **4**, 3.
86. Sornette D. 2006 *Critical phenomena in natural sciences*. Heidelberg, Germany: Springer.
87. Pareto V. 1895 La legge della domanda. *Giornale degli Economisti* **10**, 59–68.
88. Pareto V. 1997 The law of demand. *Rivista di Politica Economica* **87**, 691–700.
89. Golosovsky M, Solomon S. 2013 The transition towards immortality: non-linear autocatalytic growth of citations to scientific papers. *J. Stat. Phys.* **151**, 340–354. (doi:10.1007/s10955-013-0714-z)
90. Barabási AL. 2012 Luck or reason. *Nature* **489**, 507–508. (doi:10.1038/nature11486)
91. Champenowne DG. 1953 A model of income distribution. *Econ. J.* **63**, 318–351. (doi:10.2307/2227127)
92. Massen CP, Doye JP. 2007 A self-consistent approach to measure preferential attachment in networks and its application to an inherent structure network. *Phys. A* **377**, 351–362. (doi:10.1016/j.physa.2006.11.007)
93. Wasserman S, Faust K. 1994 *Social network analysis*. Cambridge, UK: Cambridge University Press.
94. Christakis NA, Fowler JH. 2009 *Connected: the surprising power of our social networks and how they shape our lives*. New York, NY: Little, Brown and Co.
95. Newman MEJ. 2001 Scientific collaboration networks. I. Network construction and fundamental results. *Phys. Rev. E* **64**, 016131. (doi:10.1103/PhysRevE.64.016131)
96. Newman MEJ. 2001 Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Phys. Rev. E* **64**, 016132. (doi:10.1103/PhysRevE.64.016132)
97. Newman MEJ. 2001 The structure of scientific collaboration networks. *Proc. Natl Acad. Sci. USA* **98**, 404–409. (doi:10.1073/pnas.98.2.404)
98. Newman MEJ. 2004 Coauthorship networks and patterns of scientific collaboration. *Proc. Natl Acad. Sci. USA* **101**, 5200–5205. (doi:10.1073/pnas.0307545100)
99. Trist E. 1981 The evolution of socio-technical systems. *Occas. Paper* **2**, 1–67.
100. Vespignani A. 2012 Modelling dynamical processes in complex socio-technical systems. *Nat. Phys.* **8**, 32–39. (doi:10.1038/nphys2160)
101. Borge-Holthoefer J, Baños RA, González-Bailón S, Moreno Y. 2013 Cascading behaviour in complex socio-technical networks. *J. Complex Netw.* **1**, 3–24. (doi:10.1093/comnet/cnt006)
102. Jones JH, Handcock MS. 2003 An assessment of preferential attachment as a mechanism for human sexual network formation. *Proc. R. Soc. Lond. B* **270**, 1123–1128. (doi:10.1098/rspb.2003.2369)
103. Von Mering C, Krause R, Snel B, Cornell M, Oliver SG, Fields S, Bork P. 2002 Comparative assessment of large-scale data sets of protein–protein interactions. *Nature* **417**, 399–403. (doi:10.1038/nature750)
104. Seglen PO. 1992 The skewness of science. *J. Am. Soc. Inform. Sci.* **43**, 628–638. (doi:10.1002/(SICI)1097-4571(199210)43:9<628::AID-ASIS>3.0.CO;2-0)
105. Redner S. 1998 How popular is your paper? An empirical study of the citation distribution. *Eur. Phys. J. B* **4**, 131–134. (doi:10.1007/s100510050359)
106. Lehmann S, Lautrup B, Jackson A. 2003 Citation networks in high energy physics. *Phys. Rev. E* **68**, 026113. (doi:10.1103/PhysRevE.68.026113)
107. Radicchi F, Fortunato S, Castellano C. 2008 Universality of citation distributions: toward an objective measure of scientific impact. *Proc. Natl Acad. Sci. USA* **105**, 17 268–17 272. (doi:10.1073/pnas.0806977105)
108. Stringer MJ, Sales-Pardo M, Amaral LAN. 2010 Statistical validation of a global model for the distribution of the ultimate number of citations accrued by papers published in a scientific journal. *J. Am. Soc. Inform. Sci. Technol.* **61**, 1377–1385. (doi:10.1002/asi.21335)
109. Radicchi F, Castellano C. 2012 A reverse engineering approach to the suppression of citation biases reveals universal properties of citation distributions. *PLoS ONE* **7**, e33833. (doi:10.1371/journal.pone.0033833)
110. Guimerà R, Uzzi B, Spiro J, Amaral LAN. 2005 Team assembly mechanisms determine collaboration network structure and team performance. *Science* **308**, 697–702. (doi:10.1126/science.1106340)
111. Milojević S. 2014 Principles of scientific research team formation and evolution. *Proc. Natl Acad. Sci. USA* **111**, 3984–3989. (doi:10.1073/pnas.1309723111)
112. Hirsch JE. 2005 An index to quantify an individual's scientific research output. *Proc. Natl Acad. Sci. USA* **104**, 16 569–16 572. (doi:10.1073/pnas.0507655102)
113. Radicchi F, Fortunato S, Markines B, Vespignani A. 2009 Diffusion of scientific credits and the ranking of scientists. *Phys. Rev. E* **80**, 056103. (doi:10.1103/PhysRevE.80.056103)
114. Petersen AM, Wang F, Stanley HE. 2010 Methods for measuring the citations and productivity of scientists across time and discipline. *Phys. Rev. E* **81**, 036114. (doi:10.1103/PhysRevE.81.036114)
115. Acuna DE, Allesina S, Kording KP. 2012 Future impact: predicting scientific success. *Nature* **489**, 201–202. (doi:10.1038/489201a)
116. Penner O, Pan RK, Petersen AM, Kaski K, Fortunato S. 2013 On the predictability of future impact in science. *Sci. Rep.* **3**, 3052. (doi:10.1038/srep03052)
117. Penner O, Pan RK, Petersen AM, Fortunato S. 2013 The case for caution in predicting scientists' future impact. *Phys. Today* **66**, 8–9. (doi:10.1063/PT.3.1928)
118. Michel JB *et al.* 2011 Quantitative analysis of culture using millions of digitized books. *Science* **331**, 176–182. (doi:10.1126/science.1199644)
119. Evans JA, Foster JG. 2011 Metaknowledge. *Science* **331**, 721–725. (doi:10.1126/science.1201765)
120. Pan RK, Kaski K, Fortunato S. 2012 World citation and collaboration networks: uncovering the role of geography in science. *Sci. Rep.* **2**, 902. (doi:10.1038/srep00902)
121. Mazloumian A, Helbing D, Lozano S, Light RP, Börner K. 2013 Global multi-level analysis of the 'scientific food web'. *Sci. Rep.* **3**, 1167. (doi:10.1038/srep01167)
122. Chavalarias D, Cointet J-P. 2013 Phylomemetic patterns in science evolution—the rise and fall of scientific fields. *PLoS ONE* **8**, e54847. (doi:10.1371/journal.pone.0054847)
123. Börner K. 2010 *Atlas of science*. Cambridge, MA: MIT Press.
124. Börner K. 2014 *Atlas of knowledge*. Cambridge, MA: MIT Press.
125. Radicchi F. 2012 In science 'there is no bad publicity': papers criticized in comments have high scientific impact. *Sci. Rep.* **2**, 815. (doi:10.1038/srep00815)
126. Uzzi B, Mukherjee S, Stringer M, Jones B. 2013 Atypical combinations and scientific impact. *Science* **342**, 468–472. (doi:10.1126/science.1240474)
127. Pfeiffer T, Hoffmann R. 2007 Temporal patterns of genes in scientific publications. *Proc. Natl Acad. Sci. USA* **104**, 12 052–12 056. (doi:10.1073/pnas.0701315104)
128. Zhang Q, Perra N, Gonçalves B, Ciulla F, Vespignani A. 2013 Characterizing scientific production and consumption in physics. *Sci. Rep.* **3**, 1640.
129. Tsonis AA, Schultz C, Tsonis PA. 1997 Zipf's law and the structure and evolution of languages. *Complexity* **3**, 12–13. (doi:10.1002/(SICI)1099-0526(199705/06)2:5<12::AID-CPLX3>3.0.CO;2-C)

130. Ferrer i Cancho R, Solé RV. 2001 Two regimes in the frequency of words and the origins of complex lexicons: Zipf's Law revisited. *J. Quant. Linguistics* **8**, 165–173. (doi:10.1076/jjul.8.3.165.4101)
131. Ferrer i Cancho R. 2005 The variation of Zipf's law in human language. *Eur. Phys. J. B* **44**, 249–257. (doi:10.1140/epjb/e2005-00121-8)
132. Bernhardtsson S, Correa da Rocha LE, Minnhagen P. 2009 The meta book and size-dependent properties of written language. *New J. Phys.* **11**, 123015. (doi:10.1088/1367-2630/11/12/123015)
133. Serrano MÁ, Flammini A, Menczer F. 2009 Modeling statistical properties of written text. *PLoS ONE* **4**, e5372. (doi:10.1371/journal.pone.0005372)
134. Kornai A. 2008 *Mathematical linguistics*. Berlin, Germany: Springer.
135. Ferrer i Cancho R, Solé RV. 2003 Least effort and the origins of scaling in human language. *Proc. Natl Acad. Sci. USA* **100**, 788–791. (doi:10.1073/pnas.0335980100)
136. Solé RV, Corominas-Murtra B, Fortuny J. 2010 Diversity, competition, extinction: the ecophysics of language change. *J. R. Soc. Interface* **7**, 1647–1664. (doi:10.1098/rsif.2010.0110)
137. Solé RV. 2005 Syntax for free? *Nature* **434**, 289. (doi:10.1038/434289a)
138. Baronchelli A, Felici M, Loreto V, Caglioti E, Steels L. 2006 Sharp transition towards shared vocabularies in multi-agent systems. *J. Stat. Mech.* **6**, P06014.
139. Dall'Asta L, Baronchelli A, Barrat A, Loreto V. 2006 Nonequilibrium dynamics of language games on complex networks. *Phys. Rev. E* **74**, 036105. (doi:10.1103/PhysRevE.74.036105)
140. Loreto V, Steels L. 2007 Social dynamics: emergence of language. *Nat. Phys.* **3**, 758–760. (doi:10.1038/nphys770)
141. Puglisi A, Baronchelli A, Loreto V. 2008 Cultural route to the emergence of linguistic categories. *Proc. Natl Acad. Sci. USA* **105**, 7936–7940. (doi:10.1073/pnas.0802485105)
142. Baronchelli A, Chater N, Pastor-Satorras R, Christiansen MH. 2012 The biological origin of linguistic diversity. *PLoS ONE* **7**, e48029. (doi:10.1371/journal.pone.0048029)
143. Baronchelli A, Loreto V, Tria F. 2012 Language dynamics. *Adv. Complex Syst.* **15**, 1203002. (doi:10.1142/S0219525912030026)
144. Mocanu D, Baronchelli A, Perra N, Gonçalves B, Zhang Q, Vespignani A. 2013 The twitter of babel: mapping world languages through microblogging platforms. *PLoS ONE* **8**, e61981. (doi:10.1371/journal.pone.0061981)
145. Adar E, Adamic LA. 2005 Tracking information epidemics in blogspace. In *Proc. IEEE/WIC/ACM, 19–22 September, Compiègne, France*, pp. 207–214. New York, NY: IEEE.
146. Leskovec J, Backstrom L, Kleinberg J. 2009 Meme-tracking and the dynamics of the news cycle. In *Proc. ACM SIGKDD, 28 June–1 July, Paris, France*, pp. 497–506. New York, NY: ACM.
147. Gomez Rodriguez M, Leskovec J, Krause A. 2010 Inferring networks of diffusion and influence. In *Proc. ACM SIGKDD*, pp. 1019–1028. New York, NY: ACM.
148. Simmons MP, Adamic LA, Adar E. 2011 Memes online: extracted, subtracted, injected, and recollected. In *Proc. ICWSM, 17–21 July, Barcelona, Spain*, pp. 353–360. Menlo Park, CA: AAAI Press.
149. Conover M, Ratkiewicz J, Francisco M, Gonçalves B, Menczer F, Flammini A. 2011 Political polarization on Twitter. In *Proc. ICWSM, 17–21 July, Barcelona, Spain*, pp. 89–96. Menlo Park, CA: AAAI Press.
150. Weng L, Flammini A, Vespignani A, Menczer F. 2012 Competition among memes in a world with limited attention. *Sci. Rep.* **2**, 335. (doi:10.1038/srep00335)
151. Gleeson JP, Ward JA, OSullivan KP, Lee WT. 2014 Competition-induced criticality in a model of meme popularity. *Phys. Rev. Lett.* **112**, 048701. (doi:10.1103/PhysRevLett.112.048701)
152. Dawkins R. 1989 *The selfish gene*. Oxford, UK: Oxford University Press.
153. Stanley HE. 1971 *Introduction to phase transitions and critical phenomena*. Oxford, UK: Clarendon Press.
154. Shaywitz BA, Holford TR, Holahan JM, Fletcher JM, Stuebing KK, Francis DJ, Shaywitz SE. 1995 A Matthew effect for IQ but not for reading: results from a longitudinal study. *Read. Res. Q.* **30**, 894–906. (doi:10.2307/748203)
155. Scarborough HS, Parker JD. 2003 Matthew effects in children with learning disabilities: development of reading, IQ, and psychosocial problems from grade 2 to grade 8. *Ann. Dyslexia* **53**, 47–71. (doi:10.1007/s11881-003-0004-6)
156. Morgan PL, Farkas G, Hibel J. 2008 Matthew effects for whom? *Lear. Disabil. Q.* **31**, 187–198.
157. Cohen GL, Garcia J, Apfel N, Master A. 2006 Reducing the racial achievement gap: a social–psychological intervention. *Science* **313**, 1307–1310. (doi:10.1126/science.1128317)
158. Cohen GL, Garcia J, Purdie-Vaughns V, Apfel N, Brzustoski P. 2009 Recursive processes in self-affirmation: intervening to close the minority achievement gap. *Science* **324**, 400–403. (doi:10.1126/science.1170769)
159. Leskovec J, Rajaraman A, Ullman JD. 2012 *Mining of massive datasets*. Cambridge, UK: Cambridge University Press.
160. Pastor-Satorras R, Smith E, Solé RV. 2003 Evolving protein interaction networks through gene duplication. *J. Theor. Biol.* **222**, 199–210. (doi:10.1016/S0022-5193(03)00028-6)
161. Bettencourt L, Cintrón-Arias A, Kaiser DI, Castillo-Chavez C. 2006 The power of a good idea: quantitative modeling of the spread of ideas from epidemiological models. *Phys. A* **364**, 513–536. (doi:10.1016/j.physa.2005.08.083)
162. Beel J, Gipp B. 2009 Google Scholar's ranking algorithm: an introductory overview. In *Proc. ISSI*, pp. 230–241. Rio de Janeiro, Brazil: International Society for Scientometrics and Informetrics.
163. Pacheco JM, Vasconcelos VV, Santos FC. 2014 Climate change governance, cooperation and self-organization. *Phys. Life. Rev.* (doi:10.1016/j.plrev.2014.02.003)
164. Perc M, Gómez-Gardeñes J, Szolnoki A, Floría Y. 2013 Moreno LM. Evolutionary dynamics of group interactions on structured populations: a review. *J. R. Soc. Interface* **10**, 20120997. (doi:10.1098/rsif.2012.0997)