



RESEARCH ARTICLE

Universality of citation distributions: A new understanding

Michael Golosovsky 

The Racah Institute of Physics, The Hebrew University of Jerusalem, 9190401 Jerusalem, Israel

an open access  journal



Citation: Golosovsky, M. (2021). Universality of citation distributions: A new understanding. *Quantitative Science Studies*, 2(2), 527–543. https://doi.org/10.1162/qss_a_00127

DOI: https://doi.org/10.1162/qss_a_00127

Peer Review: https://publons.com/publon/10.1162/qss_a_00127

Received: 22 October 2020
Accepted: 27 January 2021

Corresponding Author:
Michael Golosovsky
michael.golosovsky@mail.huji.ac.il

Handling Editor:
Ludo Waltman

Copyright: © 2021 Michael Golosovsky. Published under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.



Keywords: citation distribution, citation dynamics, fitness

ABSTRACT

Universality of scaled citation distributions was claimed a decade ago but its theoretical justification has been lacking so far. Here, we study citation distributions for three disciplines—Physics, Economics, and Mathematics—and assess them using our explanatory model of citation dynamics. The model posits that the citation count of a paper is determined by its fitness: the attribute, which, for most papers, is set at the moment of publication. In addition, the papers' citation count is related to the process by which the knowledge about this paper propagates in the scientific community. Our measurements indicate that the fitness distribution for different disciplines is nearly identical and can be approximated by the log-normal distribution, while the viral propagation process is discipline specific. The model explains which sets of citation distributions can be scaled and which cannot. In particular, we show that the near-universal shape of the citation distributions for different disciplines and for different citation years traces its origin to the nearly universal fitness distribution, while deviations from this shape are associated with the discipline-specific citation dynamics of papers.

1. INTRODUCTION

Science is an evolving complex network of researchers, research projects, and publications. Citations of scientific publications are the most important links that glue this network together. Analysis of citations was initially focused on journal-based citation distributions. Although these vary from discipline to discipline and from journal to journal, Seglen (1992) noticed that, after proper scaling, citation distributions for different journals collapse onto one universal curve. Radicchi, Fortunato, and Castellano (2008) studied this issue by considering different disciplines/publication years and came forward with the claim of universality of citation distributions. This claim provided a stimulus to look for universality in other complex networks, and, indeed, several dynamic universalities were found there as well (Barzel & Barabasi, 2013; Candia, Jara-Figueroa et al., 2019; Gao, Barzel, & Barabasi, 2016). While significant progress in understanding these universalities has been achieved, the origin of the universality of citation distributions remained elusive. In the context of science of science (Fortunato, Bergstrom et al., 2018; Sugimoto & Larivière, 2018; Zeng, Shen et al., 2017), the striking observation of Radicchi et al. (2008) implies that different research topics develop along similar paths, thus rendering possible such generalizations as Kuhn's paradigm shift theory (Kuhn, 1970). Understanding universality of citation distributions may provide a solid base for this theory, which has been considered as being more like a philosophical idea rather than a quantitative scientific hypothesis.

The work of Radicchi et al. (2008) paved the way for a flurry of empirical studies (Bornmann & Daniel, 2009; Chatterjee, Ghosh, & Chakrabarti, 2016; Evans, Hopkins, & Kaube, 2012; Waltman, van Eck, & van Raan, 2011) aspiring to find a fair indicator that allows quantitative comparison of the performance of papers belonging to different scientific disciplines. In the language of information science, the main achievement of Radicchi et al. (2008) was the demonstration that the variability of citation distributions for different fields is significantly reduced after going to scaled citation distributions, the scaling parameter being the mean of the distribution. The encompassing studies of Bornmann and Daniel (2009) and Waltman et al. (2011) showed deviations from such scaling, especially for research fields with a low mean number of citations. Thus, the limits to the claim of universality of citation distributions have been established. Subsequent studies (Chatterjee et al., 2016; Evans et al., 2012) extended the scaling conjecture of Radicchi et al. (2008) to sets of publications belonging to different journals, institutions, and even to Mendeley readerships (D'Angelo & Di Russo, 2019). In general, these works supported the purported scaling/universality but with some limitations; namely, research fields with a high number of uncited papers showed significant deviations from the universal distribution. To account for these deviations, two-parameter scaling (Radicchi & Castellano, 2011, 2012) was considered as well.

Once universality or near-universality of the scaled citation distributions has been demonstrated, there arises a natural question: What is the functional shape of this purportedly universal distribution? This question is a part of the general debate as to whether degree distribution in complex networks is accounted for by a power-law dependence and is scale-free, or it follows a log-normal distribution, which is not scale free (Broido & Clauset, 2019; Clauset, Shalizi, & Newman, 2009). While the study of Radicchi et al. (2008) suggested that citation distributions have nearly universal shape (whatever it is), universality in the context of complex networks was understood more as a claim of ubiquity of a certain functional form of degree distribution. While early studies, summarized in Barabasi (2015), tended to fit degree distributions in complex networks using a power-law dependence, later studies favored a stretched exponential (Wallace, Larivière, & Gingras, 2009) or a log-normal fit (Radicchi et al., 2008; Stringer, Sales-Pardo, & Amaral, 2008; Thelwall, 2016b),

$$p(K) = \frac{1}{K\sigma\sqrt{2\pi}} e^{-\frac{(\ln K - \mu)^2}{2\sigma^2}}, \quad (1)$$

where K is the number of citations of a paper, μ characterizes the mean of the distribution, and σ is the shape parameter.

In the past, citation distributions were empirically fitted using several functional shapes: stretched exponential, negative binomial, gamma, Weibull distribution, etc., the most popular being the log-normal and power-law distributions. Although the same citation distribution can be fitted by different functions, we followed the works of Radicchi et al. (2008), Stringer et al. (2008), and Thelwall (2016b) and chose the log-normal functional shape due to its simplicity. Under this choice, the universality claim of Radicchi et al. (2008) is more specific than the ubiquity of log-normal distributions in complex networks; it reduces to the statement that citation distributions for different natural science disciplines and journals can be modeled by a log-normal dependence with the *same shape parameter* σ . Indeed, extensive study of citation distributions for different journals by Thelwall (2016b) indicated that they can be described by the log-normal distribution with nearly the same $\sigma = 1$ to 1.2. This is in line with the earlier study of Stringer et al. (2008), who reported the log-normal citation distribution with $\sigma \sim 1$ for hundreds of journals¹. D'Angelo and Di Russo (2019) reported a log-normal distribution with

¹ Obviously, the values of σ listed in Stringer et al. (2008) should be multiplied by $\ln 10 = 2.3025$.

$\sigma \sim 1$ for Mendeley readerships, Evans et al. (2012) and Chatterjee et al. (2016) reported a log-normal distribution with $\sigma = 1.14$ to 1.18 for many journal-based and institution-based publications, and Clough, Gollings et al. (2014) claimed a log-normal distribution with $\sigma = 1.1$ for U.S. patent citations. Thus, a log-normal fit of citation distributions for different journals, fields, and institutions yields more or less the same shape parameter $\sigma = 1$ to 1.2, indicating that although citation distributions may have very different mean numbers of citations, they have nearly the same shape.

After the shape of citation distributions has been empirically established for many scientific disciplines (with some caveats), the quest for the explanatory model of this shape starts. Although there have been many insightful models of citation dynamics of scientific publications, they do not address the nearly universal shape of citation distributions, in such a way that the latter remained an empirical observation lacking theoretical foundation. Our goal is to find an explanatory model accounting for this observation. We have recently developed a quantitative model of citation dynamics (Golosovsky, 2019; Golosovsky & Solomon, 2017) based on our measurements with physics papers. This model can be a good platform for understanding the shape of citation distributions. In this study, we measure and analyze the citation distributions and citation dynamics of physics, economics, and mathematics papers using the same measurement protocol for three disciplines, so that the measurements could be easily compared to the model. Our goal was to find microscopic parameters of citation dynamics for each discipline, to compare them, and to find which of them are discipline specific and which are not. To our surprise, we found that several of these parameters are the same for three disciplines. We trace the near-universal shape of citation distributions to the universality of these parameters.

2. UNIVERSALITY OF CITATION DISTRIBUTIONS AND DEVIATIONS THEREFROM

We illustrate here what is usually meant by the universality of citation distributions. Consider a set of papers published in the same year t_0 , and denote by $K_j(t)$ the number of citations garnered by a paper j from this set from the moment of its publication until the year $t_0 + t$. Next, we consider the citation distribution $p(K; t)$, namely, the number of papers having K citations after t years. Radicchi et al. (2008) introduced the scaled citation distribution $p(x; t)$, where $x = \frac{K(t)}{M(t)}$ and $M(t) = \int_0^\infty Kp(K; t)dK$ is the mean number of citations garnered by a paper during t years after publication. Although citation distributions $p(K; t)$ strongly depend on the number of years after publication, Radicchi et al. (2008) showed that the scaled distributions $p(x; t)$ are very similar and hardly depend on t . The same study found that the scaled citation distributions for different disciplines are very similar as well. In other words, the work of Radicchi et al. (2008) implied that the scaled citation distributions collapse onto one curve, which depends on neither the discipline nor the publication year.

In what follows we present our measurements, which demonstrate only limited support for this claim. Namely, these measurements illustrate the universal shape of citation distributions for some sets of papers, and deviations from this shape for others. In particular, Figure 1 (left panel) shows cumulative citation distributions for all 40,195 physics papers published in 1984. The distributions for different citation years are markedly different. After dividing each distribution by its mean we obtain scaled distributions. While early scaled distributions are very similar and collapse onto a single curve, the scaled distributions in later years do not collapse. Thus, one-parameter scaling, suggested by Radicchi et al. (2008) for papers in one discipline published in 1 year, is valid only for early citation distributions. As time passes, the one-parameter scaling becomes unsatisfactory.

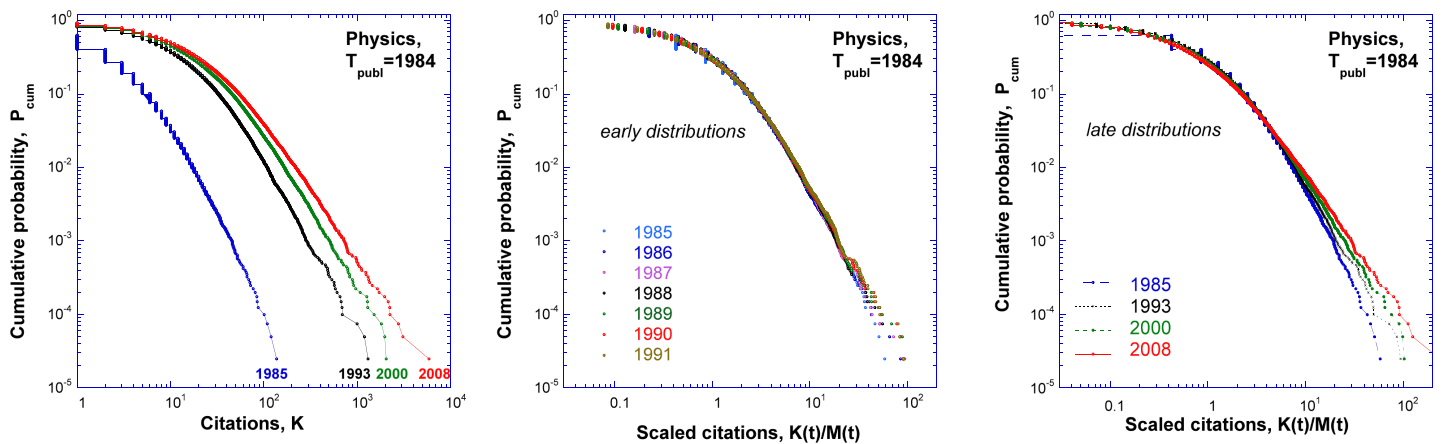


Figure 1. Cumulative citation distributions for 40,195 physics papers published in 1984. t is the number of years after publication. Left: Raw distributions, $\Pi(K; t) = \int_K^\infty p(\kappa; t) d\kappa$. Center: Scaled early distributions $\Pi(x; t)$ for $t = 2-8$ years after publication. Here, $x = \frac{K(t)}{M(t)}$ and $M(t)$ is the mean number of citations in year t . Right: Scaled early and late distributions $\Pi(x; t)$ for $t = 2-25$ years after publication. While early distributions collapse onto a single curve, the late distributions exhibit deviations.

Figure 2 shows citation distributions for the papers belonging to three different disciplines and published in 1 year. After division of each distribution by its mean, they all collapse onto a single curve. Again, this scaling works well for the early citation distributions and breaks for the late distributions (not shown here).

Figure 3 plots citation distributions for economics papers for two publication years, 1984 and 1995. We compare here two citation distributions after the same long 17-year period elapsed after publication. Although the papers published later have more citations due to the explosion in the number of publications in economics, which have been covered in the Web of Science since 2000, the scaled citation distributions again collapse onto a single curve.

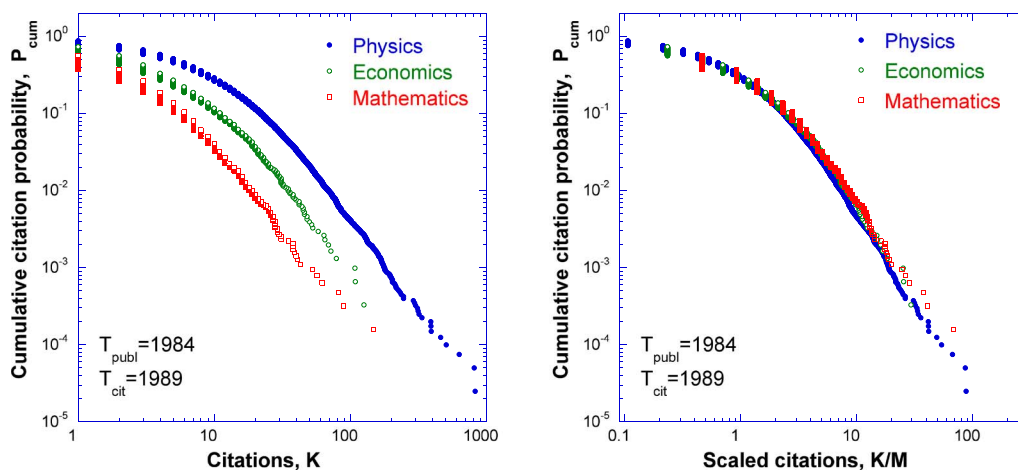


Figure 2. Cumulative citation distributions for 40,195 physics papers, 6,313 pure mathematics papers, and 3,043 economics papers, all published in 1984 and measured in 1990, early after publication. Left: Raw distributions. Right: Scaled distributions collapse onto one curve.

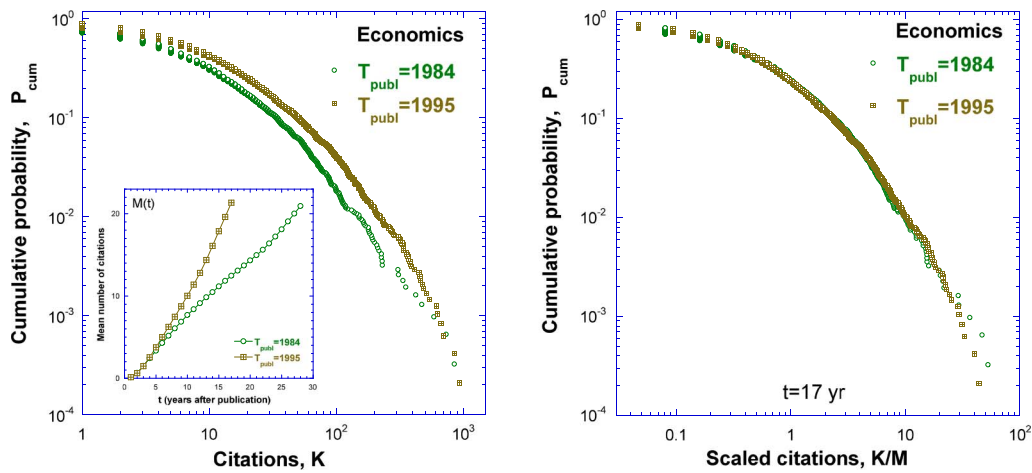


Figure 3. Cumulative citation distributions for all economics papers published in 1984 and cited in 2001; and published in 1995 and cited in 2010. In both cases, citations are counted late after publication, $t = 17$ years. Left: Raw distributions are different. The inset shows the time dependence of the mean number of citations $M(t)$. Right: Scaled distributions collapse onto one curve.

Figure 4 compares citation distributions for the physics papers published in three different journals during the same period. Obviously, these distributions are very different and even after division by the mean number of citations they do not collapse onto one curve. Here, we have intentionally focused on journals with very different mean numbers of citations. If we were considering the scaled citation distributions for the journals with more or less the same mean number of citations (Stringer et al., 2008; Thelwall, 2016a), they would collapse onto one curve.

Thus, in some cases the scaled citation distributions collapse onto one curve, while in other cases they do not. In what follows we explain these observations using our model of citation dynamics (Golosovsky, 2019; Golosovsky & Solomon, 2017).

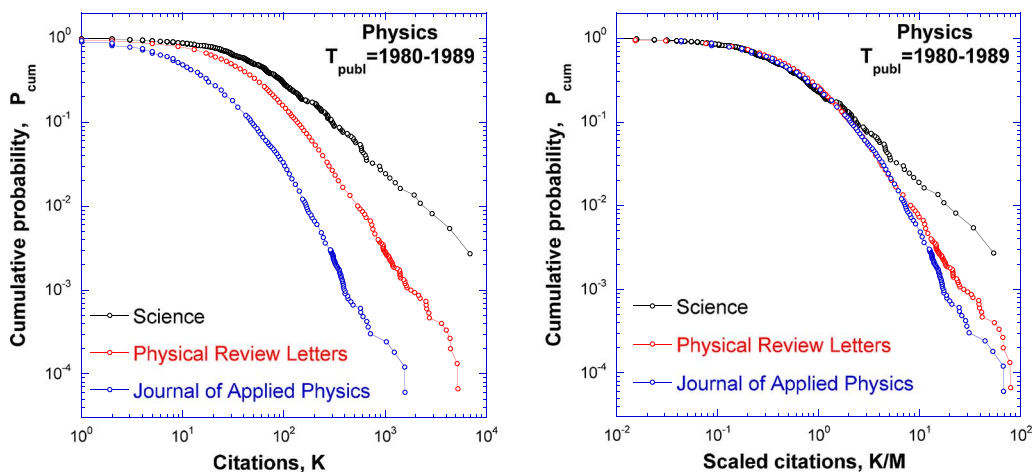


Figure 4. Cumulative citation distributions for physics papers published in three different journals: *Journal of Applied Physics*, *Physical Review Letters*, and *Science* (physics papers only) during the decade from 1980 to 1989. Citations are counted in 2008, late after publication. Left: Raw distributions. Right: The scaled distributions are different and do not collapse onto one curve.

3. MODELING CITATION DYNAMICS AND CITATION DISTRIBUTIONS

3.1. Recursive Search Model

We present here a short summary of our model. Consider a discipline, namely, a network of papers which are densely connected through citations and only loosely connected to the outside world of science (other disciplines)—a community, in the language of complex networks. Consider some paper j published in year t_0 . The author of a new paper that belongs to the same discipline and was published t years later may cite paper j after picking it up from the databases, scientific journals, or following the recommendations of colleagues or news portals. We name this a direct citation. An author of another new paper can pick up paper j from the references of the papers already included in his or her reference list. Such a strategy is known as *copying* or *redirection*, although we call it *indirect citation*. (Our definitions of direct and indirect citations are different from those of Peterson, Presse, and Dill (2010) and Milojevic (2020), who base their models on the preferential attachment mechanism.)

The model assumes that the citation dynamic of a paper follows an inhomogeneous (self-exciting) Hawkes process, namely, the probability of garnering $k_j(t)$ citations in year t is captured by a Poisson distribution, $\frac{\lambda_j^t}{k_j!} e^{-\lambda_j}$, where $\lambda_j(t)$ is the latent citation rate, which, in contradistinction to the rate of the conventional Poisson process, depends on the papers' citation history. It is given by the following expression:

$$\lambda_j(t) = \eta_j R_0 \tilde{A}(t) + \int_0^t m(t-\tau) \frac{T}{R_0} e^{-\gamma(t-\tau)} k_j(\tau) d\tau. \tag{2}$$

The first and the second addends in Eq. 2 capture, correspondingly, the direct and indirect citations, and the time t is counted from the moment of publication. $\tilde{A}(t)$ is the aging function for citations and R_0 is the average reference list length of papers belonging to this discipline and published in the same year. The main individual property of the paper is its fitness η_j —a real number that captures the appeal that this paper makes to readers, in other words, its citation potential. This definition of fitness can be traced to Caldarelli, Capocci et al. (2002) and is very different from that of Bianconi and Barabasi (2001). To determine η_j quantitatively, the papers' citation trajectory must be compared to the model prediction given by Eq. 2. The best proxy to η_j is the initial citation rate (i.e., the number of citations that the paper garners during the first 2–3 years after publication, namely, $\eta_j \propto \frac{K(t)}{t}$, where $t = 2-3$ years).

Each past citation of a paper triggers a cascade of indirect citations. These are captured by the integral in Eq. 2, where $k_j(\tau)$ is the number of citations garnered in year τ (it is also equal to the number of first-generation citing papers published in year τ), $m(t-\tau)$ is the average number of second-generation citing papers garnered by a first-generation citing paper in year t , and $\frac{T}{R_0} e^{-\gamma(t-\tau)}$ is the probability of a second-generation citing paper to cite paper j .

Equation 2 is well known in the context of branching and renewal processes (Feller, 1941) and it yields a probabilistic estimate of the citation trajectory of a paper j with fitness η_j and citation history $k_j(\tau)$. The parameters of the model, which are common for all papers in one discipline and one publication year, are R_0 , γ , T , and the functions $\tilde{A}(t)$ and $m(t)$. The latter one is not an independent function. Indeed, by averaging Eq. 2 over a collection of all papers in one discipline published in 1 year, we obtain

$$m(t) = \eta_0 R_0 \tilde{A}(t) + \int_0^t m(t-\tau) \frac{T}{R_0} e^{-\gamma(t-\tau)} m(\tau) d\tau, \tag{3}$$

where η_0 is an average fitness of the papers in this collection. Equation 3 implicitly defines a single-valued function $m(t)$ and relates it to $\tilde{A}(t)$, R_0 , γ , T , and η_0 .

The requirement of a finite reference list length imposes some constraints on the model parameters and the function $\tilde{A}(t)$. Indeed, one paper's citation is another paper's reference. This translates into the symmetry between synchronous (retrospective) and diachronous (prospective) citation distributions (Nakamoto, 1988; Roth, Wu, & Lozano, 2012). Theoretical understanding of this symmetry yields a reference-citation duality (Glanzel, 2004; Golosovsky & Solomon, 2017; Yin & Wang, 2017), which relates the dynamic of the mean number of citations to the age distribution of references in the reference lists of papers belonging to one discipline. As the growth of the number of publications and of the average reference list length may be crudely approximated by exponentials (Evans et al., 2012; Milojevic, 2012; Sugimoto & Larivière, 2018), namely, $N(t_0) \propto e^{\alpha t_0}$ and $R_0(t_0) \propto e^{\beta t_0}$, the reference-citation duality is captured by the following equation,

$$m(t_0, t_0 + t) \approx R_0(t_0)r(t_0, t_0 - t)e^{(\alpha+\beta)t}, \tag{4}$$

where t_0 is the publication year, $r(t_0, t_0 - t)$ is the average fraction of references of age t in the reference list of papers published in year t_0 , and $\int_0^\infty r(t_0, t_0 - t)dt = 1$, by definition. Equation 4 and the requirement of the finite average reference list length yield a useful relation

$$\int_0^\infty m(t_0, t_0 + \tau)e^{-(\alpha+\beta)\tau}d\tau = R_0. \tag{5}$$

Although citation patterns for different disciplines differ greatly, the referencing practices of the authors are very similar. Thus, Bertin, Atanassova et al. (2015) showed that the reference distribution is invariant with respect to the placement of references in different sections of a paper, while Sinatra, Deville et al. (2015) showed that, at least for physics papers, the function $r(t_0, t_0 - t)$ only weakly depends on the publication year t_0 . The latter study also showed that the function $r(t_0, t_0 - t)$, where t is the argument and t_0 is the parameter, varies with t_0 only on a long time scale, on the order of 10–20 years. We are interested here in the shorter time scales, hence we assume that $r(t_0, t_0 - t)$ does not depend on t_0 . This allows us to drop t_0 from our notation, for clarity. Then Eqs. 3 and 4 yield

$$r(t) = \eta_0 A(t) + \int_0^t r(t - \tau)T e^{-\gamma\tau}r(\tau)d\tau, \tag{6}$$

where

$$A(t) = \tilde{A}(t)e^{-(\alpha+\beta)t} \tag{7}$$

is the aging function for references. Because Eq. 2 contains only the product of the fitness and aging functions, this leaves us some freedom in their definition. We use this freedom to impose the normalization condition,

$$\int_0^\infty A(t)dt = 1. \tag{8}$$

Under this condition, η_0 is equal to the average fraction of direct references in the reference list of papers belonging to this discipline².

² One may wonder why we introduced a discipline-dependent parameter R_0 into Eq. 2. Indeed, by redefinition of η_j and T , this parameter could have been absorbed there, but then it would pop up in Eq. 6. Our motivation to hold it in Eq. 2 instead of Eq. 6 was driven by the observation that the function $r(t)$, which is a solution of Eq. 6, is almost independent of R_0 , as shown by Roth et al. (2012) and Yin and Wang (2017). Hence, we wished to demonstrate that Eq. 6 does not contain R_0 .

Equation 6 is the counterpart of Eq. 3 for citations. The scenario of the referencing process, which underlies Eq. 6, is as follows. An author, composing the reference list of a new paper, selects several direct references. The probability of choosing a paper j as a direct reference is given by the product $\eta_j A(t)$, where η_j is the paper's fitness and the aging function $A(t)$ is given by Eq. 7. From the reference list of each preselected paper of age τ the author randomly chooses $Te^{-\gamma\tau}$ (indirect) references, copies them, and proceeds recursively. By averaging over all papers in one discipline published in 1 year, we come to Eq. 6.

Our model of citation dynamics and its verification has been described in our previous publication (Golosovsky & Solomon, 2017). This model makes a probabilistic prediction of the citation trajectory of a paper which is based on its citation history. Our model was developed several years after the model of Wang, Song, and Barabasi (2013) and is complementary to it, in the sense that the latter is predictive whereas our model is explanatory and is based on a realistic scenario of the citation process. It should also be noted that the model of Wang et al. (2013) yields citation trajectories of papers using three parameters for each paper, and all three are paper specific. This should be compared to our model, which operates with only one individual parameter (the paper's fitness), while five other empirical parameters and two empirical functions are the same for all papers that were published in 1 year and belong to one discipline.

3.2. Citation Distributions

On the one hand, citation distribution depends on the chosen collection of papers, and on the other, it depends on the number of years after publication. With respect to the factors that determine citation distribution, our model allows separation of the static factors associated with a chosen collection of papers and the time-dependent factors associated with the citation dynamics of papers. Indeed, the model defines a collection of papers in terms of the discipline, publication year, and the fitness distribution $\rho(\eta)$. To analyze dynamic factors appearing in the model, for pedagogical reasons we replace the actual citation rate $k_j(t)$ in Eq. 2 with the latent citation rate $\lambda_j(\eta_j, t)$. Next, we assume temporarily that $T = const$ (we will revise this assumption later). After making corresponding substitutions into Eq. 2, it reduces to a linear integral Volterra equation of the second kind,

$$\lambda_j(t) = \eta_j R_0 \tilde{A}(t) + \int_0^t m(t-\tau) \frac{T}{R_0} e^{-\gamma(t-\tau)} \lambda_j(\tau) d\tau. \tag{9}$$

The solution of this equation is a linear function of η_j and R_0 , namely, $\lambda_j(t) \propto \eta_j R_0$. Next, we introduce the integrated latent citation rate,

$$\Lambda_j(t) = \int_0^t \lambda_j(\tau) d\tau = \eta_j R_0 B(t), \tag{10}$$

where the factor $B(t)$ is the same for all papers in one discipline that were published in 1 year.

According to our model, the statistical distribution of citations for a collection of papers that were published in 1 year is

$$\rho(K) = \int_0^\infty \frac{e^{-\Lambda} \Lambda^K}{K!} \rho(\eta) d\eta, \tag{11}$$

where $\rho(\eta)$ is the fitness distribution for this collection and $\Lambda(\eta)$ is given by Eq. 10. For $K \gg 1$, the Poisson factor $\frac{e^{-\Lambda} \Lambda^K}{K!}$ reduces to the delta-function $\delta(\Lambda - K)$, in such a way that

$$\rho(K) \approx \rho(\eta) R_0 B(t). \tag{12}$$

By averaging Eq. 10 over all papers in the collection, we obtain the mean number of citations

$$M(t) = \overline{K_j(t)} = \overline{\Lambda_j(t)} = \eta_0 R_0 B(t), \quad (13)$$

where η_0 is the average fitness. Because $K \approx \Lambda$ for $K \gg 1$, Eqs. 10 and 13 yield that the ratio $\frac{K_j(t)}{M(t)} \approx \frac{\eta_j}{\eta_0}$ does not depend on time. This prompts us to introduce the reduced fitness $\tilde{\eta}_j = \frac{\eta_j}{\eta_0}$. Next, we note that for $K \gg 1$ the scaled citation distribution is nothing else but the reduced fitness distribution

$$p\left(\frac{K}{M}; t\right) \approx \rho(\tilde{\eta}), \quad (14)$$

and, therefore, does not depend on time.

It should be noted that the conclusion about time-independence of the scaled citation distribution $p(x; t)$ relies on the assumption $T = \text{const}$ (Eq. 2). We will revise this assumption later and demonstrate that Eq. 14 has a very limited range of applicability.

We are now in a position to assess the conjecture of Radicchi et al. (2008). It consists of two separate statements:

- The scaled citation distributions for the same set of papers and for different citation windows collapse onto one curve; namely, the scaled citation distributions do not depend on the time after publication. This statement follows naturally from our model and is captured by Eq. 14.
- The scaled citation distributions for different sets of papers and for the same citation window collapse onto one curve. In the framework of our model, this statement is equivalent to the assertion that different collections of papers are characterized by the same reduced fitness distribution $\rho(\tilde{\eta})$. This assertion is beyond our model, because the latter does not presuppose any particular shape of the fitness distribution.

The model presented above is a synthesis of the fitness model of Caldarelli et al. (2002) and the recursive search/copying/redirection models of Krapivsky and Redner (2005), Vazquez (2001), and Simkin and Roychowdhury (2007). The new ingredient is a realistic rather than cartoon-like representation of the citation habits of authors. Our model is based on two important assumptions:

1. A paper's fitness η_j does not change during the paper's lifetime; in other words, the aging function $\tilde{A}(t)$ is the same for all papers in one discipline published in 1 year.
2. The kernel $\frac{T}{R_0} e^{-\gamma(t-\tau)}$, which characterizes indirect citations in Eq. 2, is the same for all papers in one discipline published in 1 year.

The validity of these assumptions should be verified by measurements. On the one hand, our measurements with physics papers (Golosovsky & Solomon, 2017) validated the first assumption (we found that the time dependence of the direct citations is the same for most papers; the exceptions to this rule being papers with delayed recognition, which constitute only a small fraction of all papers). On another hand, our measurements with physics papers revealed that the second assumption holds only to a certain limit. In particular, we found that the parameter T has a weak logarithmic dependence on the number of accumulated citations K ,

$$T(K) = T_0(1 + b \ln K), \quad (15)$$

where T_0 and b are empirical parameters. The $T(K)$ dependence introduces nonlinearity into Eq. 2 and its derivative, Eq. 9³. Although this nonlinearity is weak, it is important. Indeed, because Eq. 9 is weakly nonlinear, its solution, strictly speaking, cannot be factorized. Thus, Eqs. 10, 13, and 14 are only approximately valid. The nonlinearity, which is captured by Eq. 15, results in deviations from the universality of the scaled citation distributions, the magnitude of these deviations being proportional to the nonlinear coefficient b .

In summary, our model (Eqs. 2–8, and the empirical Eq. 15) provide a framework for the assessment of the purported universality of citation distributions. The model reduces this assessment to the analysis of certain parameters and functions. To find these parameters and functions, we need to perform dedicated measurements of citation dynamics and citation distributions for different collections of papers. We have already reported such measurements for one collection of papers—physics (Golosovsky & Solomon, 2017). Here, we report similar measurements and analysis for two additional collections of papers—mathematics and economics. By analyzing and comparing citation distributions and the corresponding model parameters for three disciplines, we assess the validity of the universality hypothesis of Radicchi et al. (2008).

4. ANALYSIS OF CITATION DISTRIBUTIONS AND COMPARISON TO THE MODEL

4.1. Measurements

Using Clarivate's Web of Science, we pinpointed all physics and pure mathematics papers published in 1984, and all economics papers published in 1984 and in 1995. We considered research papers, letters, notes, and uncited papers, while editorial material and reviews were excluded. We measured the citation dynamics of the papers belonging to these collections during a long period after publication and using a citation window of 1 year. The measurements for physics papers were partially described in our previous publication (Golosovsky & Solomon, 2017), and we compare them here with our new measurements for mathematics and economics papers.

4.2. Fitting Procedure

Figure 5 shows citation distributions for two disciplines and for several citation years. To model these distributions, for each discipline and publication year we considered a synthetic set containing the same number of papers and characterized by a certain fitness distribution. On these synthetic sets, we ran the stochastic numerical simulation based on Eqs. 2 and 15 and tuned the parameters of the simulation to achieve close correspondence between citation dynamics of the real and synthetic sets of papers.

Besides citation distributions, we considered the citation lifetime τ_0 , which was defined from the exponential approximation of the paper's citation trajectory, $K(t) = K_\infty(1 - e^{-\Gamma t})$, where $\Gamma = 1/\tau_0$ is the obsolescence rate (see Figure 6). We tuned the parameters of the numerical simulation to fit not only citation distributions but citation lifetime as well. While the fitting of citation distributions using many parameters leaves some ambiguity, the simultaneous fitting of the citation distributions (snapshot measurements) and citation lifetime (longitudinal measurements), using the same parameters, pinpoints these parameters unambiguously.

³ It should be noted that Eq. 2 with a kernel containing logarithmic nonlinearity pops out not only in the context of citations; Iribarren and Moro (2011) found that the dynamics of viral marketing is captured by a very similar equation.

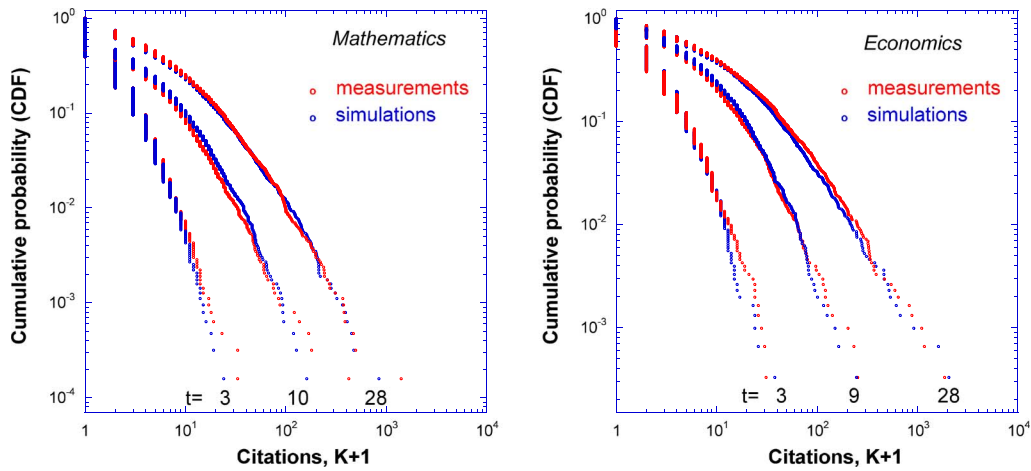


Figure 5. Cumulative citation distributions for two collections of papers published in 1984. Left: Mathematics. Right: Economics. The red circles show our measurements and the blue circles show the results of stochastic numerical simulation based on the inhomogeneous Hawkes process with the rate given by Eq. 2. t is the number of years after publication, the publication year corresponding to $t = 1$.

To perform numerical simulation of the citation distributions, we have to choose some trial fitness distribution. While Eq. 2 contains papers' fitness η_j , our analysis of the scaled citation distributions (Eqs. 13 and 14) focuses on reduced fitness, $\tilde{\eta}_j = \frac{\eta_j}{\eta_0}$, where η_0 is the average fitness. Therefore, in our simulations we used the reduced fitness, which we assumed to follow a log-normal distribution, $\rho(\tilde{\eta}) = \frac{1}{\tilde{\eta}\sigma\sqrt{2\pi}} e^{-\frac{(\ln \tilde{\eta} + \frac{\sigma^2}{2})^2}{2\sigma^2}}$. The mean of this distribution is unity and it is fully defined by its shape parameter σ .

Next, we recast the first term in Eq. 2 as $\tilde{\eta}_i[\eta_0 R_0 \tilde{A}(t)]$ and considered the expression $\eta_0 R_0 \tilde{A}(t)$ as one composite fitting function. Thus, the fitting parameters and functions used in our simulation were σ , the shape parameter of the reduced fitness distribution; the composite function $\eta_0 R_0 \tilde{A}(t)$ which characterizes direct citations; and the parameters $\frac{T_0}{R_0}$ and γ, b which characterize indirect citations.

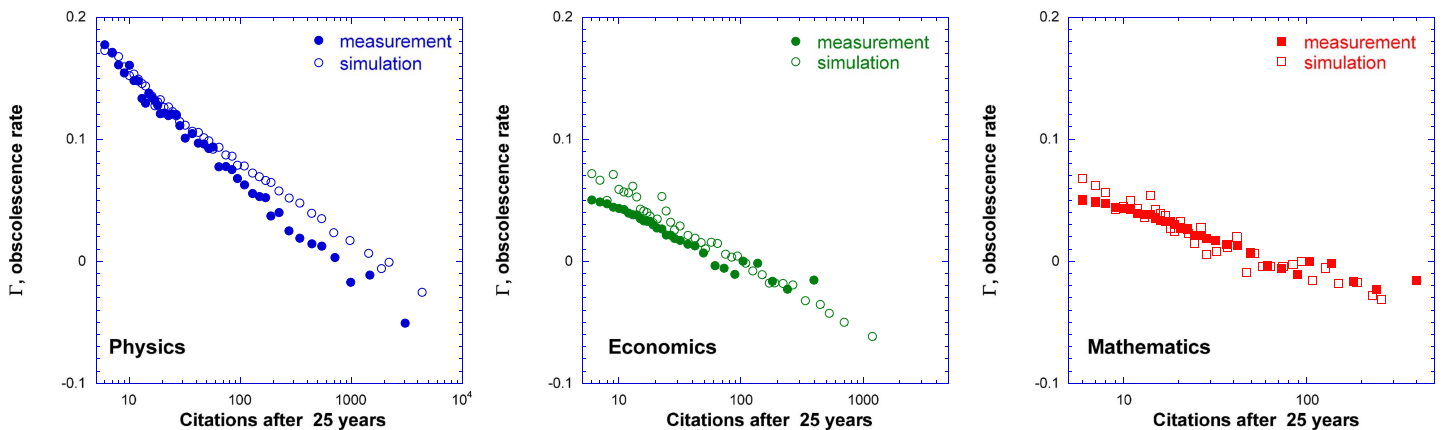


Figure 6. The obsolescence rate $\Gamma(K)$ as a function of the number of citations K . For each discipline, Γ decreases (citation lifetime $\tau_0 = \Gamma^{-1}$ increases) with the number of citations K . Eventually, the obsolescence rate Γ changes sign, indicating the onset of the runaway behavior. The filled symbols show our measurements; the open symbols show the results of numerical simulation.

The fitting procedure was as follows. For each discipline, we found some initial combination of the parameters σ , $\frac{T_0}{R_0}$, γ , b , and $\eta_0 R_0 \tilde{A}(t)$, which satisfactorily fit citation distributions for different citation years. Of these, the parameters σ , $\frac{T_0}{R_0}$, γ , and b were the same for all citation years, while $\eta_0 R_0 \tilde{A}(t)$ was the only parameter that was specific for each citation year. After achieving a reasonable fit for citation distributions, we focused on citation lifetime of papers and its dependence on the number of citations. Because citation lifetime $\tau_0 = 1/\Gamma$ does not depend on the fitness distribution, we used only two parameters, $\frac{T_0}{R_0}$ and b , for fitting while the remaining parameter γ was determined independently from the analysis of the Pearson correlation coefficient for citation fluctuations in subsequent years (not shown here). Figure 6 shows that Γ decreases logarithmically with K (τ_0 increases) and this is a direct consequence of the $T(K)$ dependence captured by Eq. 15. The slope of the $\Gamma(K)$ dependence yields the nonlinear coefficient b . After fitting citation lifetimes, we came back to citation distributions and ran our simulation with the previously found parameters $\frac{T_0}{R_0}$, γ , and b , while fine-tuning σ and $\eta_0 R_0 \tilde{A}(t)$. Then we came back to citation lifetime and fine-tuned the parameters $\frac{T_0}{R_0}$, γ , and b . After several loops of fitting we achieved *simultaneously* a good correspondence between the measured and simulated citation distributions, on the one hand, and the measured and simulated citation lifetimes, on the other.

Our next step was to determine each one of the three multipliers in the composite function $\eta_0 R_0 \tilde{A}(t)$. To this end, we analyzed the mean number of citations, $M(t) = \int_0^t m(\tau) d\tau$, in particular, we considered $M_{detrended}(t) = \int_0^t m(\tau) e^{-(\alpha+\beta)\tau} d\tau$ (Figure 7). It follows from Eq. 5 that $\frac{M_{detrended}(t)}{R_0} = \int_0^t r(\tau) d\tau$, where $r(t)$ is the reduced age composition of the average reference list. From the requirement of the convergence of this integral to unity, in the long time limit, we found the sum of the growth exponents $(\alpha + \beta)$ and the reference list length R_0 . From known R_0 and $\frac{T_0}{R_0}$ we found T_0 (see Eq. 15). At the next step, we recast the composite function $\eta_0 R_0 \tilde{A}(t)$, found from the fitting procedure for citation distributions, as $\eta_0 R_0 A(t) e^{(\alpha+\beta)t}$, substituted there R_0 and $(\alpha + \beta)$ found from the analysis of

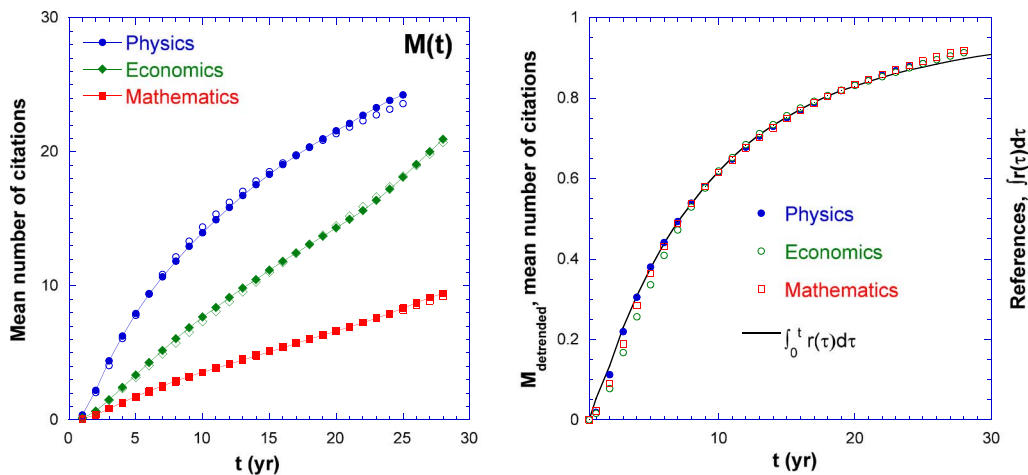


Figure 7. The cumulative mean number of citations for 40,195 physics, 6,313 mathematics, and 3,043 economics papers, all published in 1984. Left: $M(t)$. The symbols show our measurements and the continuous lines show results of the numerical simulations based on Eq. 3. Note the excellent correspondence. Right: Detrended and scaled data, $\frac{M_{detrended}(t)}{R_0} = \int_0^t r(\tau) d\tau$. Although the results of the measurements for the three disciplines are very close, the model does not require them to be identical. The continuous line shows the $\int_0^t r(\tau) d\tau$ dependence found from the analysis of the age composition of the reference lists of physics papers (Golosoosky & Solomon, 2017).

Table 1. Model parameters for collections of papers used in our measurements

Disciplines	Publication year T_{publ}	Number of papers N	Reference list length R_0	Growth exponents $(\alpha+\beta), \text{yr}^{-1}$	Average fitness η_0	Shape parameter σ	Copying exponent γ, yr^{-1}	Indirect citations T_0	Nonlinear coefficient b
Physics	1984	40,195	19.2	0.045	0.48	1.13	1.2 ± 0.2	3.2	0.42
Economics	1995	4,782	8.8	0.12	0.49	1.13	1 ± 0.2	7.2	0.30
Economics	1984	3,043	8.4	0.09	0.45	1.13	1 ± 0.2	6.6	0.29
Mathematics	1984	6,313	3.9	0.092	0.46	1.13	0.8 ± 0.2	5.5	0.25

$M(t)$, and determined the remaining parameter η_0 (average fitness) from the normalization condition $\int_0^\infty A(t)dt = 1$.

4.3. Fitting Parameters and Functions

Table 1 summarizes the parameters of citation dynamics for three disciplines, as determined through the above fitting procedure. Our most important finding is that all three disciplines are characterized by the same reduced fitness distribution, namely, a log-normal distribution with the shape parameter $\sigma = 1.13$. The aging function $A(t)$ for the three disciplines turned out to be nearly identical, as well. Indeed, Figure 8a shows that while the aging functions $\tilde{A}(t)$ differ from discipline to discipline, the detrended aging functions, $A(t) = \tilde{A}(t)e^{-(\alpha+\beta)t}$ (the aging function for references, which captures the proportion of recent references versus old references in the average reference list of papers—see Eq. 6) collapse onto one curve. It should be noted that our model does not presuppose the universality of $\rho(\tilde{\eta})$ and $A(t)$: This unexpected result follows from our measurements for three widely different disciplines.

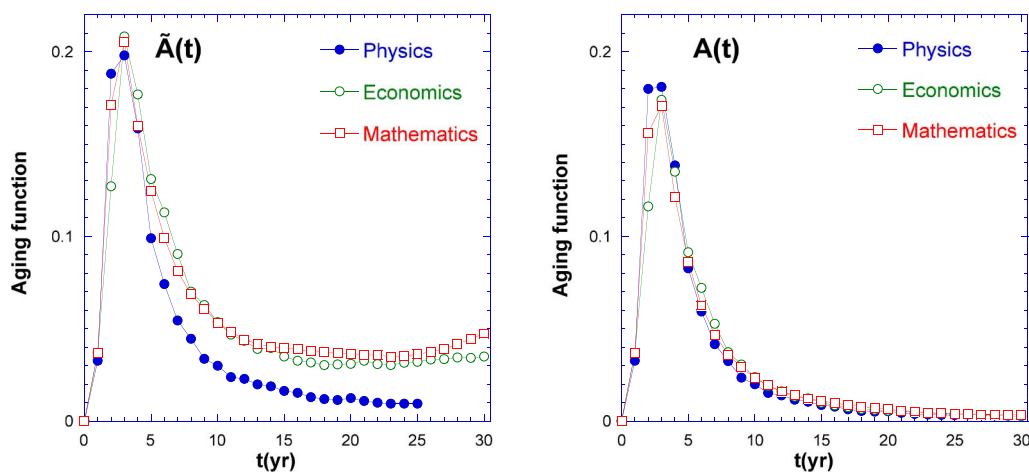


Figure 8. Left: $\tilde{A}(t)$, the aging functions for citations as found from the analysis of citation distributions of the physics, mathematics, and economics papers published in 1984. Right: The detrended aging functions (aging function for references) $A(t)$ (see Eq. 6). While $\tilde{A}(t)$ dependences are different, $A(t)$ dependences nearly collapse onto one curve.

While the average fitness η_0 , which captures the proportion of the direct references in the average reference list, has little variation from discipline to discipline⁴, the remaining parameters (the average reference list length R_0 ⁵, the growth exponents α , β ⁶, and the parameters T_0 , γ , and b , which characterize the indirect citations) differ from discipline to discipline. Thus, they are nonuniversal.

5. DISCUSSION

We are now in a position to assess the empirically established universality of citation distribution and the deviations therefrom. Radicchi et al. (2008) conjectured that properly scaled citation distributions, for the collection of papers published in 1 year and for different citation windows, collapse onto one curve. We assessed this conjecture theoretically, using our model of citation dynamics, and came to the conclusion that, if citation dynamics were linear, the scaled citation distributions would indeed collapse. However, as our measurements show, citation dynamics are nonlinear. Therefore, citation distributions do not exhibit perfect scaling, and there are deviations from one universal curve, such as those presented in Figure 1c. Because the nonlinearity of citation dynamics is associated with viral propagation, namely, with the parameter $T(K)$ (Eq. 15), the magnitude of these deviations is determined by the term $b \ln K$, where b is the nonlinear coefficient and K is the number of accumulated citations. Notably, deviations from the scaling which result from nonlinearity are mostly associated with $K \gg 1$, namely, with the highly cited papers.

For collections of papers belonging to different disciplines and published in the same year, early citation distributions contain a very small number of highly cited papers, hence they obey scaling (Figure 1b); while late citation distributions, containing many highly cited papers, do not scale (Figure 1c). Thus, one source of the deviations from scaling is the *nonlinear* citation dynamics.

Another conjecture that follows from Radicchi et al. (2008) is that properly scaled citation distributions for different sets of papers and for the same citation window collapse onto one curve. This is a much stronger statement. When analyzed in the framework of our model of citation dynamics, this is equivalent to the assertion that different collections of papers are characterized by the same reduced fitness distribution $\rho(\tilde{\eta})$. Our measurements support this remarkable claim for collections of papers belonging to different disciplines and published in 1 year⁷. The same reasoning predicts that the scaled citation distributions for collections of papers

⁴ Our measurements of this fraction in the reference lists of physics papers published in 2010 yielded $\eta_0 = 0.34$, which is slightly lower than η_0 found from the analysis of citations of the physics papers published in 1984. Probably η_0 varies with time.

⁵ We estimated here the average reference list length R_0 from the measurements of citation dynamics. For physics papers published in 1984, this estimate agrees well with the direct measurements of the reference list length. However, our estimates of R_0 for mathematics and economics are too small. It should be remembered, however, that R_0 , as estimated from citation dynamics through Eq. 4, includes only original research papers and excludes books, conference proceedings, and interdisciplinary references. These constitute a very small proportion of physics references, while they are abundant among mathematics and economics references; hence our estimate of R_0 for these disciplines is smaller than the actual reference list length.

⁶ The data of Sugimoto and Larivière (2018) show that the exponential approximation for the growth of the number of publications is reasonable, and most disciplines exhibit the growth exponent $\alpha \sim 0.04$ in the period 1984–2010; while the exponential growth of the reference list length is a very crude approximation and it grows with time very nonuniformly, in such a way that the corresponding effective growth exponent β depends upon the time window of measurements.

⁷ When we compare citation distributions for different disciplines, we consider here, for clarity, only early citation distributions, for which the nonlinearity associated with viral propagation has not yet developed.

belonging to different journals (Figure 4) would be nonuniversal, as the corresponding fitness distribution $p_{journal}(\eta)$ is journal-specific. Although a set of papers published in a journal is a subset of those belonging to the whole discipline, the sampling performed by each journal is not the same due to different acceptance criteria. For example, while *Science* and *Physical Review Letters* skim the high-fitness tail of the fitness distribution, the *Journal of Applied Physics* samples it more uniformly.

Thus, we have demonstrated that citation distributions for the sets of papers published in 1 year are determined by the fitness distribution, on the one hand, and by the citation dynamics of papers, on the other. While the latter differ from discipline to discipline, the fitness distribution is the same for physics, economics, and mathematics and it is fairly well approximated by a log-normal distribution with shape parameter $\sigma \sim 1.13$. Limpert, Stahel, and Abbt (2001) reviewed the log-normal distributions occurring in nature and demonstrated that the distribution with $\sigma \sim 1$ is one of the narrowest observed. In fact, Ghadge, Killingback et al. (2010) showed that this distribution is something special: It generates a citation network that is a borderline between two classes—a gel-like network and a network consisting of isolated clusters. On the other hand, the log-normal distribution belongs to the class of fat-tailed distributions and is reminiscent of self-organized criticality in sand piles. Such an analogy is not unexpected, as each new paper in the scientific enterprise causes a cascade of citations and, ideally, an avalanche of new and fruitful ideas. Why different disciplines adjust themselves to produce this specific shape of the fitness distribution—a log-normal with $\sigma \sim 1$ —is an intriguing question. This nearly universal fat-tailed distribution probably reveals some facet of science as a self-organizing system.

Our results can be considered from another perspective. In the framework of our recursive search model, the information about a paper propagates in the scientific community in two ways: broadcasting (the authors find this paper after reading news, searching in the Internet, reading the journals, etc.—this corresponds to direct citations) and word-of-mouth (finding this paper in the reference lists of other papers—we name these *indirect citations*). These two modes of propagation are coupled: Each direct citation gives rise to cascades of indirect citations, which can turn viral. Although direct citations are garnered in proportion to the paper's fitness, which captures its intrinsic quality and attributes, indirect citations depend on the structure of the citation network and gauge the paper's fame. The number of citations combines the paper's fitness and fame (Simkin & Roychowdhury, 2013). As indirect citations originate from direct citations, the paper's fitness is the key parameter that determines the overall number of citations. Our results imply that the fitness distributions for different disciplines are very similar whereas citation distributions are not, inasmuch as they are associated with viral propagation of information in the network of communications corresponding to each discipline. In other words, the static attributes of the citation network for each discipline are universal, while the dynamic attributes are not. This differentiation between the dynamic and static attributes can be relevant to other growing complex networks as well.

6. CONCLUSIONS

We explored the conjecture of Radicchi et al. (2008) who claimed that the scaled citation distributions collapse onto one curve, namely, their shape is nearly universal. We found that the scaling holds for collections of papers belonging to one discipline, published in 1 year, and measured several years after publication. We explain this observation using our recently developed model of citation dynamics, which delineates between the static and dynamic factors affecting the citation dynamics of papers. The model attributes the accumulated citations to the paper's fitness, on the one hand, and to the viral propagation of the information about this

paper in the scientific community, on the other. We believe that the underlying reason for the scaling of citation distributions is the universal fitness distribution for scientific disciplines. This claim has been verified by our measurements with physics, economics, and mathematics papers. Although extrapolation from these three disciplines to all science may be too ambitious, because the three disciplines are so different, that is still plausible.

We also found that citation distributions do not scale well when one compares collections of papers many years after publication. In this case, our model traces the deviations from the scaling to the discipline-specific viral propagation. On the other hand, we find that citation distributions for different journals also do not scale. In this case, we attribute deviations from the scaling to the journal-specific fitness distribution which can differ from the universal fitness distribution for a scientific discipline as a whole.

Thus, our model of citation dynamics explains the near-universality of the scaled citation distributions and also accounts for the deviations from this near-universality.

ACKNOWLEDGMENTS

I am grateful to Sorin Solomon for fruitful discussions, to Magda Fontana for stimulating discussions and for the assessment of the economics journals, and to Yakov Varshavsky for the assessment of the mathematics journals.

COMPETING INTERESTS

The author has no competing interests.

FUNDING INFORMATION

This research was not funded.

DATA AVAILABILITY

Citation distributions for the physics, mathematics, and economics papers are available in the supplementary material (<https://doi.org/10.5281/zenodo.4558007>).

REFERENCES

- Barabasi, A.-L. (2015). *Network science*. Cambridge: Cambridge University Press.
- Barzel, B., & Barabasi, A.-L. (2013). Network link prediction by global silencing of indirect correlations. *Nature Biotechnology*, 31, 720–725. DOI: <https://doi.org/10.1038/nbt.2601>, PMID: 23851447, PMCID: PMC3740009
- Bertin, M., Atanassova, I., Gingras, Y., & Larivière, V. (2015). The invariant distribution of references in scientific articles. *Journal of the Association for Information Science and Technology*, 67(1), 164–177. DOI: <https://doi.org/10.1002/asi.23367>
- Bianconi, G., & Barabasi, A.-L. (2001). Bose-Einstein condensation in complex networks. *Physical Review Letters*, 86, 5632–5635. DOI: <https://doi.org/10.1103/PhysRevLett.86.5632>, PMID: 11415319
- Bornmann, L., & Daniel, H.-D. (2009). Universality of citation distributions—A validation of Radicchi et al. relative indicator $c_f = c/c_0$ at the micro level using data from chemistry. *Journal of the American Society for Information Science and Technology*, 60(8), 1664–1670. DOI: <https://doi.org/10.1002/asi.21076>
- Broido, A. D., & Clauset, A. (2019). Scale-free networks are rare. *Nature Communications*, 10(1), 1017. DOI: <https://doi.org/10.1038/s41467-019-08746-5>, PMID: 30833554, PMCID: PMC6399239
- Caldarelli, G., Capocci, A., DeLosRios, P., & Muñoz, M. A. (2002). Scale-free networks from varying vertex intrinsic fitness. *Physical Review Letters*, 89(25), 258702. DOI: <https://doi.org/10.1103/PhysRevLett.89.258702>, PMID: 12484927
- Candia, C., Jara-Figueroa, C., Rodriguez-Sickert, C., Barabasi, A.-L., & Hidalgo, C. A. (2019). The universal decay of collective memory and attention. *Nature Human Behaviour*, 3(1), 82–91. DOI: <https://doi.org/10.1038/s41562-018-0474-5>, PMID: 30932052
- Chatterjee, A., Ghosh, A., & Chakrabarti, B. K. (2016). Universality of citation distributions for academic institutions and journals. *PLOS ONE*, 11, e0146762. DOI: <https://doi.org/10.1371/journal.pone.0146762>, PMID: 26751563, PMCID: PMC4709109
- Clauset, A., Shalizi, C., & Newman, M. (2009). Power-law distributions in empirical data. *SIAM Review*, 51(4), 661–703. DOI: <https://doi.org/10.1137/070710111>

- Clough, J. R., Gollings, J., Loach, T. V., & Evans, T. S. (2014). Transitive reduction of citation networks. *Journal of Complex Networks*, 3(2), 189–203. DOI: <https://doi.org/10.1093/comnet/cnu039>
- D'Angelo, C. A., & Di Russo, S. (2019). Testing for universality of Mendeley readership distributions. *Journal of Informetrics*, 13(2), 726–737. DOI: <https://doi.org/10.1016/j.joi.2019.03.011>
- Evans, T. S., Hopkins, N., & Kaube, B. S. (2012). Universality of performance indicators based on citation and reference counts. *Scientometrics*, 93(2), 473–495. DOI: <https://doi.org/10.1007/s11192-012-0694-9>
- Feller, W. (1941). On the integral equation of renewal theory. *Annals of Mathematical Statistics*, 12(3), 243–267. DOI: <https://doi.org/10.1214/aoms/1177731708>
- Fortunato, S., Bergstrom, C. T., Borner, K., Evans, J. A., Helbing, D., ... Barabasi, A.-L. (2018). Science of science. *Science*, 359(6379), eaao0185. DOI: <https://doi.org/10.1126/science.aao0185>, PMID: 29496846, PMCID: PMC5949209
- Gao, J., Barzel, B., & Barabasi, A.-L. (2016). Universal resilience patterns in complex networks. *Nature*, 530, 307. DOI: <https://doi.org/10.1038/nature16948>, PMID: 26887493
- Ghadge, S., Killingback, T., Sundaram, B., & Tran, D. A. (2010). A statistical construction of power-law networks. *International Journal of Parallel, Emergent and Distributed Systems*, 25(3), 223–235. DOI: <https://doi.org/10.1080/17445760903429963>
- Glanzel, W. (2004). Towards a model for diachronous and synchronous citation analyses. *Scientometrics*, 60(3), 511–522. DOI: <https://doi.org/10.1023/B:SCIE.0000034391.06240.2a>
- Golosovsky, M. (2019). *Citation analysis and dynamics of citation networks*. Cham: Springer. DOI: <https://doi.org/10.1007/978-3-030-28169-4>
- Golosovsky, M., & Solomon, S. (2017). Growing complex network of citations of scientific papers: Modeling and measurements. *Physical Review E*, 95(1), 012324. DOI: <https://doi.org/10.1103/physreve.95.012324>, PMID: 28208427
- Iribarren, J. L., & Moro, E. (2011). Branching dynamics of viral information spreading. *Physical Review E*, 84(4), 046116. DOI: <https://doi.org/10.1103/physreve.84.046116>, PMID: 22181236
- Krapivsky, P. L., & Redner, S. (2005). Network growth by copying. *Physical Review E*, 71, 036118. DOI: <https://doi.org/10.1103/PhysRevE.71.036118>, PMID: 15903504
- Kuhn, T. S. (1970). *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press.
- Limpert, E., Stahel, W. A., & Abbt, M. (2001). Log-normal distributions across the sciences: Keys and clues. *BioScience*, 51(5), 341. DOI: [https://doi.org/10.1641/0006-3568\(2001\)051\[0341:LNDATS\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2001)051[0341:LNDATS]2.0.CO;2)
- Milojevic, S. (2012). How are academic age, productivity and collaboration related to citing behavior of researchers? *PLOS ONE*, 7(11), 1–13. DOI: <https://doi.org/10.1371/journal.pone.0049176>, PMID: 23145111, PMCID: PMC3492318
- Milojevic, S. (2020). Towards a more realistic citation model: The key role of research team sizes. *Entropy*, 22(8), 875. DOI: <https://doi.org/10.3390/e22080875>, PMID: 33286646, PMCID: PMC7517479
- Nakamoto, H. (1988). Synchronous and diachronous citation distributions. *Informetrics*, 87/88, 157–163. <https://hdl.handle.net/1942/837>
- Peterson, G. J., Presse, S., & Dill, K. A. (2010). Nonuniversal power law scaling in the probability distribution of scientific citations. *Proceedings of the National Academy of Sciences*, 107(37), 16023–16027. DOI: <https://doi.org/10.1073/pnas.1010757107>, PMID: 20805513, PMCID: PMC2941273
- Radicchi, F., & Castellano, C. (2011). Rescaling citations of publications in physics. *Physical Review E*, 83(4), 046116. DOI: <https://doi.org/10.1103/physreve.83.046116>, PMID: 21599249
- Radicchi, F., & Castellano, C. (2012). A reverse engineering approach to the suppression of citation biases reveals universal properties of citation distributions. *PLOS ONE*, 7(3), e33833. DOI: <https://doi.org/10.1371/journal.pone.0033833>, PMID: 22479454, PMCID: PMC3315498
- Radicchi, F., Fortunato, S., & Castellano, C. (2008). Universality of citation distributions: Toward an objective measure of scientific impact. *Proceedings of the National Academy of Sciences*, 105(45), 17268–17272. DOI: <https://doi.org/10.1073/pnas.0806977105>, PMID: 18978030, PMCID: PMC2582263
- Roth, C., Wu, J., & Lozano, S. (2012). Assessing impact and quality from local dynamics of citation networks. *Journal of Informetrics*, 6(1), 111–120. DOI: <https://doi.org/10.1016/j.joi.2011.08.005>
- Seglen, P. O. (1992). The skewness of science. *Journal of the American Society for Information Science and Technology*, 43(9), 628–638. DOI: [https://doi.org/10.1002/\(SICI\)1097-4571\(199210\)43:9<628::AID-ASI5>3.0.CO;2-0](https://doi.org/10.1002/(SICI)1097-4571(199210)43:9<628::AID-ASI5>3.0.CO;2-0)
- Simkin, M. V., & Roychowdhury, V. P. (2007). A mathematical theory of citing. *Journal of the American Society for Information Science and Technology*, 58(11), 1661–1673. DOI: <https://doi.org/10.1002/asi.20653>
- Simkin, M. V., & Roychowdhury, V. P. (2013). A mathematical theory of fame. *Journal of Statistical Physics*, 151(1), 319–328. DOI: <https://doi.org/10.1007/s10955-012-0677-5>
- Sinatra, R., Deville, P., Szell, M., Wang, D., & Barabasi, A.-L. (2015). A century of physics. *Nature Physics*, 11, 791–796. DOI: <https://doi.org/10.1038/nphys3494>
- Stringer, M. J., Sales-Pardo, M., & Amaral, L. A. N. (2008). Effectiveness of journal ranking schemes as a tool for locating information. *PLOS ONE*, 3(2), e1683. DOI: <https://doi.org/10.1371/journal.pone.0001683>, PMID: 18301760, PMCID: PMC2244807
- Sugimoto, C. R., & Larivière, V. (2018). *Measuring research*. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/wentk/9780190640118.001.0001>
- Thelwall, M. (2016a). Are there too many uncited articles? Zero inflated variants of the discretised lognormal and hooked power law distributions. *Journal of Informetrics*, 10(2), 622–633. DOI: <https://doi.org/10.1016/j.joi.2016.04.014>
- Thelwall, M. (2016b). Citation count distributions for large mono-disciplinary journals. *Journal of Informetrics*, 10(3), 863–874. DOI: <https://doi.org/10.1016/j.joi.2016.07.006>
- Vazquez, A. (2001). Disordered networks generated by recursive searches. *EPL (Europhysics Letters)*, 54(4), 430. DOI: <https://doi.org/10.1209/epl/i2001-00259-y>
- Wallace, M. L., Larivière, V., & Gingras, Y. (2009). Modeling a century of citation distributions. *Journal of Informetrics*, 3(4), 296–303. DOI: <https://doi.org/10.1016/j.joi.2009.03.010>
- Waltman, L., van Eck, N. J., & van Raan, A. F. J. (2011). Universality of citation distributions revisited. *Journal of the American Society for Information Science and Technology*, 63(1), 72–77. DOI: <https://doi.org/10.1002/asi.21671>
- Wang, D., Song, C., & Barabasi, A.-L. (2013). Quantifying long-term scientific impact. *Science*, 342(6154), 127. DOI: <https://doi.org/10.1126/science.1237825>, PMID: 24092745
- Yin, Y., & Wang, D. (2017). The time dimension of science: Connecting the past to the future. *Journal of Informetrics*, 11(2), 608–621. DOI: <https://doi.org/10.1016/j.joi.2017.04.002>
- Zeng, A., Shen, Z., Zhou, J., Wu, J., Fan, Y., Wang, Y., & Stanley, H. E. (2017). The science of science: From the perspective of complex systems. *Physics Reports*, 714–715, 1–73. DOI: <https://doi.org/10.1016/j.physrep.2017.10.001>