The Z-index: A geometric representation of productivity and impact which accounts for information in the entire rank-citation profile

Alexander M. Petersen[1] and Sauro Succi[2,3]

[1] *IMT Lucca Institute for Advanced Studies, 55100 Lucca, Italy*
[2] *Istituto Applicazioni Calcolo C.N.R., Rome, Italy*
[3] *Freiburg Institute for Advanced Studies, Albertstrasse, 19, D-79104, Freiburg, Germany*

We present a simple generalization of Hirsch's h-index, $Z \equiv \sqrt{h^2 + C}/\sqrt{5}$, where C is the total number of citations. Z is aimed at correcting the potentially excessive penalty made by h on a scientist's highly cited papers, because for the majority of scientists analyzed, we find the excess citation fraction $(C - h^2)/C$ to be distributed closely around the value 0.75, meaning that 75 percent of the author's impact is neglected. Additionally, Z is less sensitive to local changes in a scientist's citation profile, namely perturbations which increase h while only marginally affecting C. Using real career data for 476 physicists careers and 488 biologist careers, we analyze both the distribution of Z and the rank stability of Z with respect to the Hirsch index h and the Egghe index g. We analyze careers distributed across a wide range of total impact, including top-cited physicists and biologists for benchmark comparison. In practice, the Z-index requires the same information needed to calculate h and could be effortlessly incorporated within career profile databases, such as *Google Scholar* and *ResearcherID*. Because Z incorporates information from the entire publication profile while being more robust than h and g to local perturbations, we argue that Z is better suited for ranking comparisons in academic decision-making scenarios comprising a large number of scientists.

1. Introduction

The most commonly used quantitative measure of a scientist's publication portfolio is Hirsch's h-index, which was designed to measure both productivity and impact simultaneously [1]. However, there have been many criticisms claiming that the h-index leads to inconsistencies in scientific ranking, represents productivity or impact but not both, is non-decreasing and hence cannot be used as a short-term evaluation metric, and oversimplifies the publication portfolio [2–9]. As such, despite its increasing use in diverse decision-making processes, it may not be optimally suited for career evaluation scenarios. By way of example, it has recently been implemented by the National Agency for the Evaluation of Universities and Research Institutes (ANVUR) of Italy as selection and pay-scale criterion in the most recent large-scale national "habilitation" competition [10]. Furthermore, without proper normalization of c, the number of citations to a given paper, in order to account for variations across time, group size, academic discipline, and even academic sub-discipline, the practice of comparing citation counts without normalizing is highly questionable [2, 5, 11–15].

The Hirsch index integer h counts the number of publications in a scientists's portfolio which satisfy the criterion of having h or more citations each. A scientist's rank-citation profile, $c_i(r)$, is calculated by ranking the N_i total publications of a given scientist i in decreasing order of citations, so that $c_i(1) \geq c_i(2) \geq \cdots \geq c_i(N_i)$. The significance threshold h is chosen somewhat arbitrarily by using the "fixed point" relation, corresponding to $c_i(h_i) = h_i$, which is most easy to communicate by graphically visualizing the rank-citation distribution. Fig. 1(a) illustrates how h corresponds to the intersection of the line $c = r$ with $c_i(r)$. The mathematical definition of h is that the paper of rank h has no less than h citations:

$$c_i(h_i) \geq h_i \ , \tag{1}$$

which insignificantly alters the graphical definition[1].

However, the arbitrary choice of the quality threshold h, whereby papers with only h or more citations are counted, implicitly makes this more a measure of *productivity* conditioned to a quality threshold. Hence, a second measure of impact is required, since an obvious feature of a 2-dimensional representation is that there be at least two independent degrees of freedom. Here, we propose the two-dimensional representation of productivity and impact in the z-plane defined by the coordinates $(h, \sqrt{C})$, which incorporates the total citations $C = \sum_r c(r)$ as a complementary impact measure to h, which is principally a productivity measure.

Moreover, the simple geometric combination

$$Z \equiv \frac{\sqrt{h^2 + C}}{\sqrt{5}} \tag{2}$$

computed by using the information typically provided in online career profile databases and better incorporates the non-trivial information contained in the entirety of a scientist's rank-citation profile $c_i(r)$. In what follows, using analytic and empirical demonstration based on real data for 964 scientists, we will show several basic properties of Z which preserve the conveniences of h while gaining robustness to perturbations in $c(r)$ and also providing a better representation of the net impact of the *entire* rank-citation profile.

---

[1] For the purpose of brevity and compactness, we shall tend to suppress the author index i from career measures throughout the rest of the paper, unless explicit reference to the author dependence is crucial.

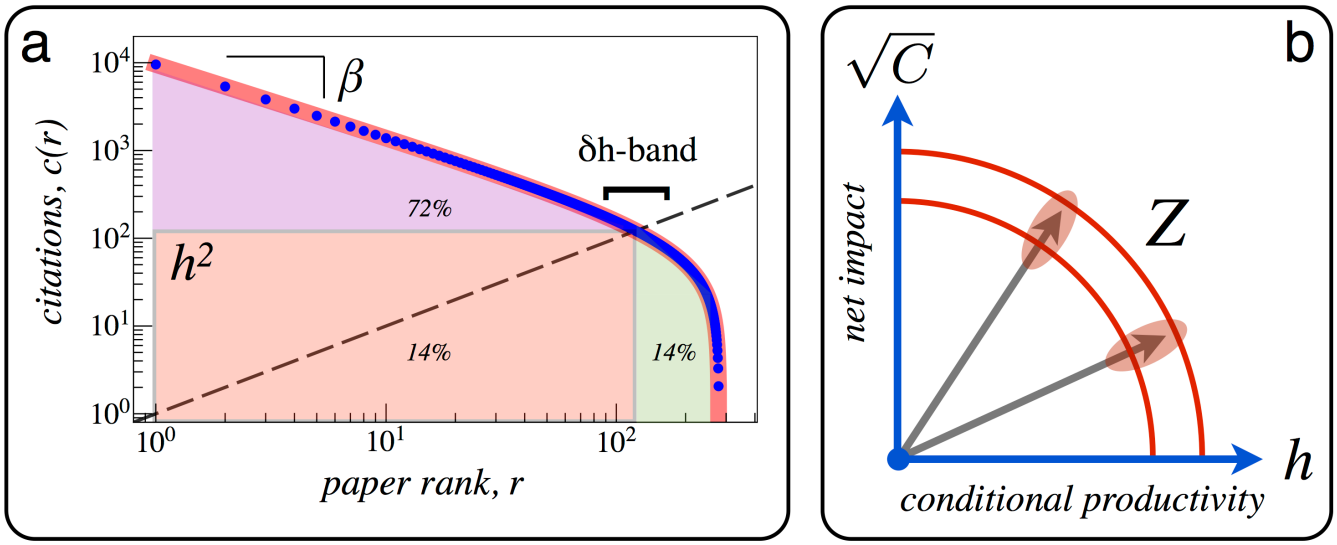

FIG. 1: Understanding the relation between C and h. (a) Schematic illustration of the rank citation profile $c_i(r)$, illustrated as a discrete generalized beta distribution (DGBD) defined in Eq. (8) using $N = 278$, $\beta_i \equiv 0.83$, $\gamma_i \equiv 0.67$, $A_i = 220$, $C_i = 79,342$, and $h_i = 104$ (reproduced from the schematic example in Petersen et al. 16). All single-valued quantitative indicators can be calculated from $c_i(r)$. The citation count h^2 (shaded pink) accounts for only a subset of the total citations C. In fact, for careers similar to the rank-citation profile shown, where the top papers are roughly distributed according to a power law $c_i(r) \sim r^{-\beta_i}$, then h_i^2 represents only a tiny fraction of the total citations C. Indeed, in this example which is characteristic of prestigious scientists, the excess citations of the peak papers contribute to the majority of the total C. The "δh band" is the set of papers with rank $r \in [h \pm \delta h]$ that, once cited, optimally increase the h-index, while only marginally affecting C. For this reason h is unstable with respect to local perturbations in $c(r)$. Because Z is a geometric combination of h and $\sqrt{C}$, the quantity Z is less sensitive to local changes in a scientist's citation profile, and hence less susceptible to direct manipulation. (b) Visualizing the z-plane. The radial isolines for constant $Z = \sqrt{h^2 + C/\sqrt{5}}$ enclose careers with similar Z values. Given the diversity of careers, and the variability in $c(r)$, these isolines capture the two-dimensional features of both conditional (relatively high-impact) productivity as captured by h, and net impact of all publications as measured by C.

2. Why yet another index?

The utility of the h-index is that it provides a remarkably calculable, easy-to-memorize, and comparable quantitative summary of the information contained in the full rank-citation distribution $c(r)$. It captures the difficulty in science of consistently producing highly-cited papers, and discourages voluminous publication strategies which may lack overall quality.

The appreciation of quality over quantity is an issue that must be emphasized as science makes a revolutionary shift away from old system of publication based on peer-review in printed journals with fixed capacity towards a rapid and unlimited online capacity system of the near future. In this future scenario, measures similar to the h-index will be important for extracting the signal from the noise, and set quality thresholds by which to measure productivity.

However, the h-index comes with a number of well-known weaknesses [2–9]. Below we provide a partial list of those that are most relevant to the motivation for Z:

1. The quality threshold c^* used to highlight the representative papers for which $c(r) \geq c^*$, is arbitrarily chosen to be $c^* \equiv h$, corresponding to the fixed point solution of $c(h) \geq h$. Other arbitrary significance thresholds are used in alternative productivity-impact measures, such as the "i10-index", the number of publications with at least 10 citations, which is listed on *Google Scholar* profiles. It is poorly understood how the choice of the significance threshold c^* may alter the overall distribution of the impact indicator across scientists, and whether or not there is a "best" choice for c^*.

2. The h-index severely discounts the impact of the highly cited papers, for which $c(r) \gg h$, in a scientist's publication portfolio. This point is particularly important, since in science, like in sports, many notable distinctions are awarded to recognize top performance rather than overall continued impact.

3. Motivated by the competitive reward system, scientists may begin to adapt strategies that "game the system" of prestige. According to the definition of the h-index as a single point on $c(r)$, it is indeed possible for a scientist who has accurate knowledge of his/her $c(r)$, to selectively self-cite his/her papers in the "δh-band" $r \in \{h - \delta h, h + \delta h\}$, in order to optimally increase his/her h index in the near future (see Fig. 1(a)).

These criticisms are not new, and many alternative metrics have been proposed to mend these weaknesses [7, 8]. Among others, the Egghe g-index, defined by $g^2 \leq \sum_{r=1}^{g} c(r)$, is designed to provide more weight to the highly cited papers within $c(r)$ [17]. Alternatively, it can more readily be appreciated as a fixed point measure of the average number of ci-

tations calculated for the first g papers, $g \leq \sum_{r=1}^{g} c(r)/g = \langle c \rangle_g$. However, as we will show in the later section, g and h are quantitatively related and highly correlated, and hence do not measure remarkably different information contained in $c(r)$. The neglected citation count $C - h^2$ mostly belong to the highly-cited papers, and are the motivation for a complementary excess e-index [18]. Figure 1(a) shows the rank-citation profile of a typical highly-cited scientist with a significant number of highly cited papers. Indeed, the neglected citations $C - h^2$ account in this case for 86% of the total citations! Nevertheless, the h-index is popular because it sends a quick and efficient reputation signal that most practicing scientists can readily appreciate relative to their peers.

3. Data analyzed

Here we analyze real career data for 476 physicist careers and 488 biologist careers. Each dataset contains careers distributed across a wide range of total impact, including the top-100 physicists and top-100 biologists according to total citation counts in high-impact journals. These two datasets will be used as an elite benchmark. We use (i) disambiguated "distinct author" data from *Thomson Reuters Web of Knowledge* (TRWOK), `www.isiknowledge.com/`, using their matching algorithms to identify publication profiles of distinct authors, and (ii) scientist profiles from ResearcherID.com, the open portal for TRWOK which allows individuals to aggregate publications into an online repository which conveniently calculates both h and C.

For the selection of two comparison sets for high-impact physicists, we aggregate all authors who published in *Physical Review Letters (PRL)* over the 50-year period 1958-2008 into a common dataset. From this dataset, we rank the scientists using the citations shares metric defined in [13], and choose the top 100 scientists resulting in dataset [A] (average h-index $\langle h \rangle \pm$ Std.Dev. $= 61 \pm 21$). As a comparative set of highly cited physicists, we also choose from our ranked *PRL* list, approximately randomly, 100 additional highly prolific physicists comprising dataset [B] (average h-index $\langle h \rangle = 44 \pm 15$). We compare the tenured scientists in datasets A and B with 100 relatively young assistant professors from physics in dataset [C] ($\langle h \rangle = 14 \pm 7$). To select dataset [C] scientists, we chose two assistant professors from the top 50 U.S. physics and astronomy departments, ranked according to the magazine *U.S. News*. Further analysis of the publication trajectories and collaboration patterns of physicists in dataset [A,B,C] is provided in [16, 19, 20]. Dataset [D] is comprised of 174 "graphene" scientists with profiles on *ResearcherID.com*, ($\langle h \rangle = 15 \pm 11$). Dataset [E] ($\langle h \rangle = 92 \pm 35$) is comprised of the top 100 scientists who published in the journal *Cell*, using the same ranking method as with dataset [A]. Datasets [F], [G] , and [H] correspond to scientists with profiles on *ResearcherID.com* with the keywords "molecular biology" ($\langle h \rangle = 20 \pm 17$), "neuroscience" ($\langle h \rangle = 17 \pm 13$), and "genomics" ($\langle h \rangle = 18 \pm 14$), comprising 60, 76, and 252 profiles, respectively. Only ResearcherID profiles with more than 7 publications were analyzed.

4. Empirical Results

4.1. Accounting for highly-cited papers by rescaling the h-index

A main weakness of the h-index, its severe neglect of a scientist's highly cited papers, can be remedied by a two-dimensional representation of the net impact $\sqrt{C}$ and conditional productivity h. We propose a representation of authors by the data pair $(h, \sqrt{C})$ in a two-dimensional $z-$plane shown in Fig. 1. We define the index $Z \equiv \sqrt{C + h^2}/\sqrt{5}$, which is simply a vector norm of the coordinate $(h, \sqrt{C})$ using the "natural" units corresponding to $h/\sqrt{5}$ and $C/5$. The factor of 5 is chosen according to statistically robust patterns between h and C which we discuss next.

Since h^2 is a subset of the citations calculated by C, then these two values are highly correlated,

$$C_i = s_i h_i^2 \,, \tag{3}$$

where this slope parameter $s_i \geq 1$, was noted originally by Hirsch (noted as a [1]). Recently, it was shown empirically by S. Redner [21] that the distribution of s across authors is highly peaked around $s \approx 4$. As a result, one may conclude that there is little point of combining two highly correlated indicators. Indeed, the quantity $\sqrt{C}$ is the h-index an individual would have if all of his/her N_i papers had the same number of citations, hence a completely flat $c(r)$. In the following, we shall argue oppositely; namely that the little spread in the slope can nevertheless lead to sensible readjustments between "peaky" and "flat" authors, who deviate significantly away from the characteristic value $s = 4$. For this reason, we define Z using the "natural units" factor $1/\sqrt{5}$, so that the rescaling of h is mild in the region around $s = 4$, but becomes fairly substantial in the tails of the s distribution.

Figure 2 shows the scatter plots of $(h_i, \sqrt{C_i})$ pairs for each scientist i on linear and log-log axes. Colored bands represent isolines of constant Z. The data are confined to a relatively small radial bands of the phase space, mainly because h is highly correlated with C [16, 21]. The fact that data are almost entirely collected between the rescaling lines with slope $Z(s)/h = 5$ and 15, indicates the utility of the polar representation of $(h_i, \sqrt{C_i})$. The angle defining each datapoint is proportional to s, and is not as informative as the overall magnitude Z, which, we argue, captures a great deal of the summary information for each career.

Figure 3 shows the probability distribution $P(s)$ which is peaked around the median $Q_{0.5}(s) = 4.0$. This means that, for the majority of scientists analyzed, the excess citation fraction $(C - h^2)/C = 1 - 1/s$ is distributed closely around the value 0.75, meaning that 75% of a career's citation impact is neglected by h. The standard deviation $\sigma(s)$ is largely dataset

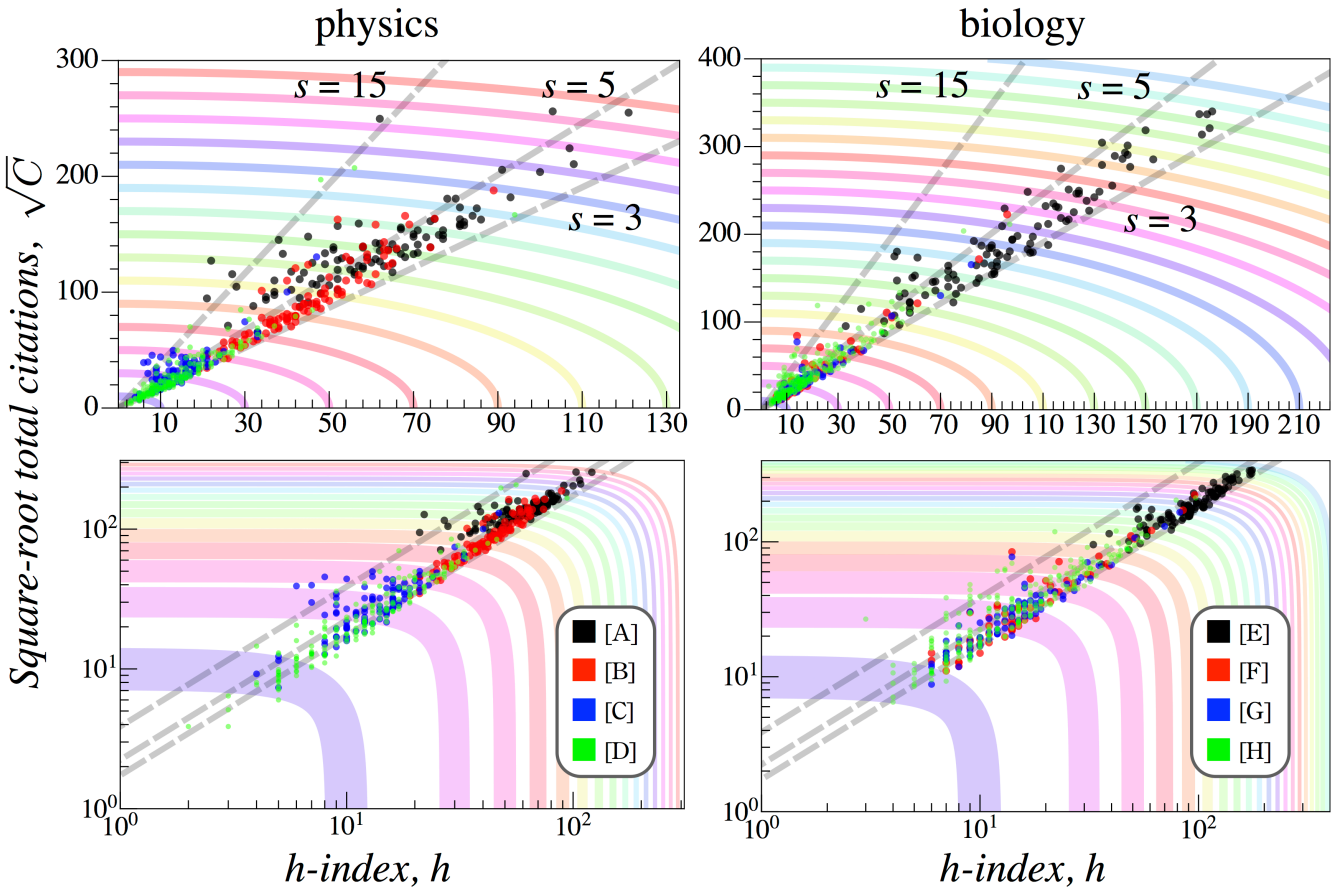

FIG. 2: Scatterplot of empirical $h, \sqrt{C}$ values in the z-plane. The colored curves are constant Z isolines of exponentially decreasing thickness with increasing Z. The straight dashed lines correspond to the rescale factors with slope $R(s)$ using $s = 3, 5, 15$. The middle 50% of the careers are enclosed by the lines with $s = 3$ and 5 (since the quartiles $Q_{0.75} = 5.2$ and $Q_{0.25} = 3.3$). Interestingly, the low-h outliers consisting of many assistant professor profiles (blue data points) tend to be well described by clusters along Z isolines, suggesting that assistant professors are typically hired using criteria that select for relatively large C.

dependent, ranging from 1.2 to 6.6 due to the potential for extremely large s values arising from careers with a large citation difference between the top-cited paper(s) and the rest of the papers. Nevertheless, most careers are contained within a relatively narrow radial band in the z-plane.

Because of the regularities in the statistical distribution of s, we define the norm using the "natural units" for Z, leading to the formulation

$$Z = h\sqrt{\frac{1+s}{5}}\,. \tag{4}$$

The "natural units" normalization factor $1/\sqrt{5}$ means that Z retains the advantages of the scalar index h since careers with $s = 4$ correspond to the traditional h-index, $Z(s = 4) = h$.

So what is gained by using Z instead of h? Mainly, Z does not discount the value of significantly cited papers, those papers from which a scientist derives much of his/her scientific reputation [20]. A second practical advantage of Z is its robustness against perturbations in the z-plane, $(h, \sqrt{C}) \to (h+1, \sqrt{C+1})$. Such perturbations could arise from just the stochastic inflow of citations, or possibly from covert "cosmetic surgery" self-citation strategies, aimed at increasing the h-index.

We calculate the perturbation to Z,

$$\delta Z = \frac{1}{\sqrt{5}}\left(\sqrt{(h+1)^2 + C + 1} - \sqrt{h^2 + C}\right), \tag{5}$$

resulting from a citation landing perfectly on a paper with h citations located in the center of the "h-band" (see Fig. 1). Hence, $\delta Z \sim \frac{h+1}{5Z} \sim 1/5 \ll 1$ for profiles with $Z \approx h$. More generally, from the definition, the change of Z in the generic transition $(h, C) \to (h + \delta h, C + \delta C)$ is given by

$$\delta Z = \frac{1}{5Z}(h\delta h + \frac{1}{2}\delta C) \tag{6}$$

Assuming $Z \sim h$, this expression shows that, in order to increase Z by one unit, the change $\delta h = 1$ must be accompanied by a change in the total citations of the order of $\delta C \sim 8Z$. In other words, C serves as an inertial reservoir of citations, preventing rapid changes due to local adjustments in $c_i(r)$. In summary, a single citation raising h by one unit (not a

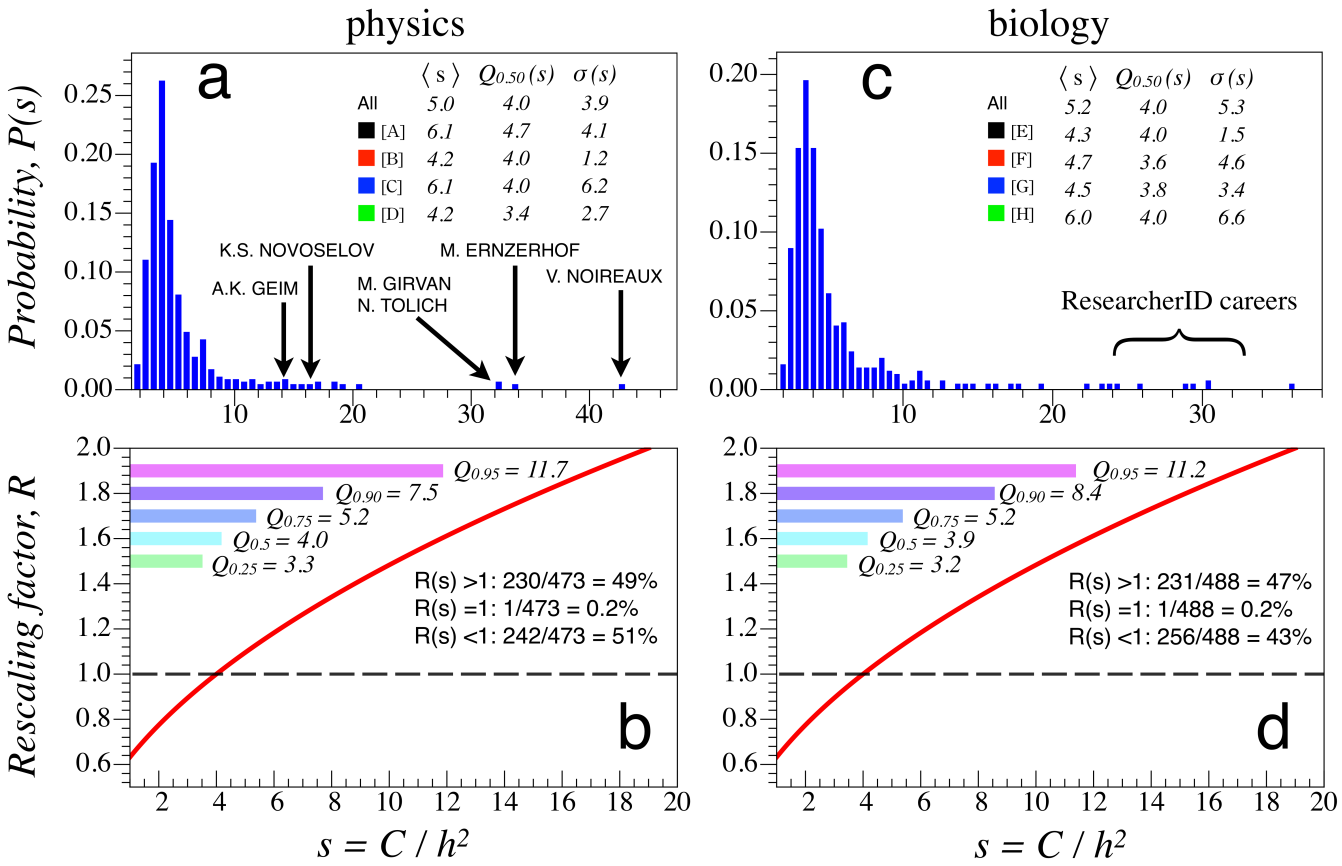

FIG. 3: Rescaling the h-index. (top) Empirical distribution of the scale-factor s calculated for all careers within each disciplinary field aggregated into a single dataset. In panels a, c, each distribution $P(s)$ is peaked around $Median(s) = 4.0$. Mean, standard deviation, and median values are also listed for each individual subset [A]-[D]. Standout careers, such as the recent Nobel Prize in Physics winners KS Novoselov and AK Geim, have large s values due to their seminal publications. Other extreme s_i values can arise if the number of publications N_i is not considerably large and so C_i is dominated by just a few high-impact papers. (bottom) The rescale factor $R(s) = \sqrt{(1+s)/5}$ is a slow function of s around unity for $s = 4$. Bar lines indicate the quantile Q_x calculated from the empirical cumulative distribution of s values. For instance in panel b, 90% of the careers analyzed have $s < 7.5$, and the middle 50% have values in the range $3.3 < s < 5.2$. Shown are the number and % of careers in each dataset with rescale value $R(s) > 1$, $R(s) = 1$, and $R(s) < 1$.

marginal increment, since for a productive scientist h grows by approximately one unit a year) would have no effect on Z, rendering covert self-citation strategies less rewarding.

In the lower panels of Fig. 3 we also plot the rescaling factor

$$R(s) = \frac{Z(s)}{h} = \sqrt{\frac{1+s}{5}} \tag{7}$$

between Z and h, and show the fraction x of scientists in each dataset having R less than various quantile values Q_x. $R(s)$ is a slow monotonically increasing function of s, demonstrated by the perturbation $s = 4 + \Delta$, which varies the rescaling factor as $R \sim 1 + \Delta/10$ for $\Delta \ll 1$. Hence, authors with $\Delta \ll 10$ receive a mild (linear) correction to their h-index, while outliers at both extremes may be significantly affected by the rescaling. For example, Fig. 3(b) indicates that the middle 50% of physics careers analyzed have s values between 3.3 and 5.2, corresponding to $R(3.3) = 0.92$ and $R(5.2) = 1.11$, respectively.

4.2. Accounting for the entire rank-citation profile $c(r)$

It was recently shown, for a large range of careers, that the *entire* citation profile $c_i(r)$ can be quantified with a relatively simple parametric class of rank distribution model, the discrete generalized beta distribution (DGBD)

$$c_i(r) \equiv A_i r^{-\beta_i} (N_i + 1 - r)^{\gamma_i} \, . \tag{8}$$

The DGBD is well-suited for systems with finite number of constituents, as shown recently for rank-ordering of systems in the arts and sciences [16, 20, 22, 23]. The β parameter controls the logarithmic "Zipf-law" slope for small r (high rank) constituents, whereas the γ parameter controls the "exponential-like" cutoff for large r (low rank) constituents.

A schematic example of a characteristic DGBD $c_i(r)$ of an elite scientist is plotted in Fig. 1(a). With only three degrees of freedom, this distribution is able to capture $c_i(r)$ across

the entire range of r, as demonstrated for a broad range of careers in physics, including even assistant professor careers with $N_i \sim 20$. It remains an open problem whether scientists can be robustly classified in terms of the parameters β_i, A_i, and C_i.

It is possible to approximate the coefficient A_i in Eq. (8) using the fixed-point definition $c(h) \equiv h$, which implies that $A/h^\beta \approx h$. Hence, from Eq. (8) it follows that the expected total number of citations can be approximated knowing β and h for a given profile by

$$C_{\beta,h} \equiv \sum_{r=1}^{N} Ar^{-\beta} \approx h^{1+\beta} \sum_{r=1}^{N} r^{-\beta} = h^{1+\beta} H_{N,\beta} \, . \quad (9)$$

Since the *generalized harmonic number* $H_{N,\beta}$ is on the order $O(1)$ for $\beta \approx 1$, then we arrive at the simple scaling relation $C_{\beta,h} \sim h^{1+\beta}$ [16] which agrees with empirical findings that $C \sim h^2$ [21] in the typical case where $\beta \approx 1$. A recent study of an extremely large dataset of more than 30,000 $c_i(r)$ profiles [24] indicates that the global distribution of β values is indeed more concentrated with $\beta > 1$ values, corresponding to $c_i(r)$ that have a sharp decay from the highest cited paper to the bulk of the papers, representative of the majority of non-elite scientists. This is in stark contrast to the $\beta \leq 1$ values found for top scientists reflecting the slow decay in $c_i(r)$ due to a large subset of highly-cited papers.

Furthermore, a similar calculation shows that $g^2 \sim h^{1+\beta}$, and hence $g \sim h$ for $\beta \approx 1$. In these heuristic calculations we neglect the γ_i scaling regime since the low-rank papers typically contribute a relatively small amount to the overall C_i tally as compared to the high-rank papers.

Together, these simple scaling relations show that h and g are closely related, conveying roughly the same overall information. We test this prediction empirically in the upper panels of Fig. 4 which shows for each disciplinary set that $g \sim h^\alpha$ with $\alpha \approx 1.10 \pm 0.01$ for the physicists and $\alpha \approx 1.08 \pm 0.01$ for the biologists. Deviations from unity arise since the β_i scaling exponent is not universal, but varies around $\beta \sim 1$. In the lower panels of Fig. 4 we show the analogous scatter plot and power-law model regression between Z and h, which show slightly smaller R^2, indicative of certain profiles that are extremely penalized according to h, but which are compensated by Z. Since β_i values are significantly more difficult to calculate than h_i, requiring regression or maximum likelihood calculations, we propose that the simple two-dimensional data pair $(h, \sqrt{C})$ conveniently and sufficiently capture scientific production-impact profiles across a broad range of age and prestige.

By way of example, consider the following physics careers with various β_i distinguishing how "steep" each $c_i(r)$ is in the highly-cited regime. The average β value calculated across datasets [A,B,C] is $\langle\beta\rangle \approx 0.77$. Hence, scientists with $\beta = 0.78$, similar to A. H. Castro Neto with $(h = 30, C = 3509)$ corresponding to $s = 3.9$ and $R = 0.99$, would not be impacted, as $Z = h$. Scientists with relatively large $\beta = 1.04$, such as R. B. Laughlin with $(h = 32, C = 7751)$ corresponding to $s = 7.6$ and $R = 1.3$, would be significantly upgraded to $Z = 42$. A relatively flat profile with $\beta \sim 0.5$ and $(h = 30, C = 2000)$ corresponding to $s = 2.2$ and $R = 0.8$, would be downgraded to $Z = 24$. For a pathologically extreme case, consider E. Lieberman (not included in datasets [A-D]), with $(N = h = 6, C = 2530)$ corresponding to $s = 70$, $R = 3.8$, and $Z = 23$. Alternatively, consider the "perfectly flat" author with $(N = h = 30, C = h^2)$ corresponding to $s = 1$, $R = \sqrt{2/5}$, and $Z = 19$. We believe that these readjustments represent a fair reward to the highly-cited peak papers which are discounted when considering h alone.

4.3. Testing the rank-stability of h, g, and Z

The rank stability of a system, either with respect to time evolution or variation in the ranking variable, is largely influenced by both the amplitude and the diffusive properties of the noise in the system [25]. In the case of ranking of scientists according to quantitative measured derived from $c_i(r)$, it is important to quantify the rank stability of two comparative indices, say index A and index B. We expect that there will be more rank stability when the two measures being compared representing the same information derived from $c_i(r)$. However, if there is new information entering into index A that is not incorporated into index B, then we expect there to be larger fluctuations in the rankings of the scientists.

We test this underlying assumption using three indexes, h, g, and Z, and compare the rank-stability of each pair. For each disciplinary dataset, we calculate the rank of each scientist, $r_i(A)$ according to each index, h, g, and Z. We then compare the rank of each scientist $r_i(B)$ according to a different index B, resulting in 3 scatter plots for each disciplinary dataset, as shown in Fig. 5.

We quantify the rank-stability of each index pair (A, B) using two methods. First, for the rank pair $(r_i(A), r_i(B))$ for each career i, we calculate the distance

$$\Delta r = |r_i(A) - r_i(B)|/\sqrt{2} \quad (10)$$

from the line $y = x$. The line $y = x$ is the stability benchmark corresponding to the ideal case of perfect rank stability, $r_i(A) = r_i(B)$. In order to investigate the relation between Δr and r, Fig. 5 shows the running averages $\langle\Delta r\rangle$, with 1-standard deviation error bars. Specifically, the $\langle\Delta r\rangle$ is averaged over rank intervals $I_r \equiv [r - \delta r, r + \delta r]$ using $\delta r = 10$ and shown are the running average for $r = [(1 + \delta r)...(S - \delta r)]$. The overall trends are not dependent on the choice of δr.

For the pair $A = g$ and $B = h$ the values of $\langle\Delta r\rangle$ are relatively small over the entire range of r_g, reflecting the property that $g \sim h$, and hence incorporate roughly the same information from $c_i(r)$. However, comparing the ranking for $A = h$ and $A = g$ to $B = Z$ we find that the amplitude of $\langle\Delta r\rangle$ significantly increases for large r_h and r_g, reflecting an increasing rank-instability for the scientists with relatively small publication portfolios, and hence, more "noise" in their productivity-impact scores when measured by h and g. This underscores the importance of using quantitative measures as only complementary factors for the evaluation of early ca-

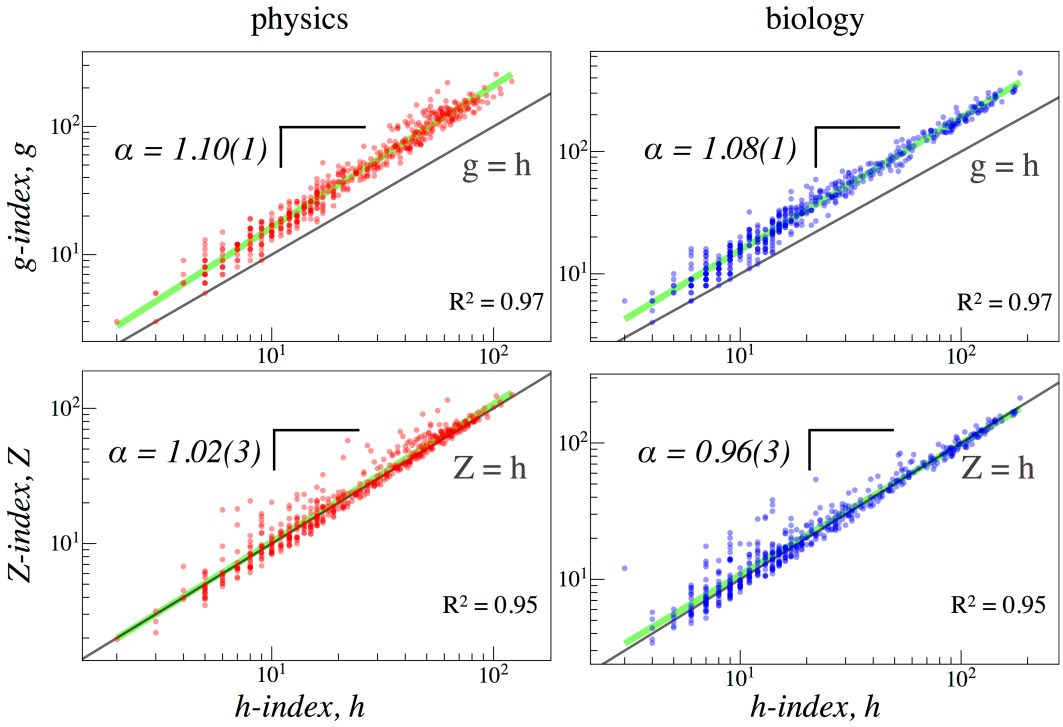

FIG. 4: h and g are nearly linearly related. (Upper panels) Regression of h and g values indicate that these to measures are somewhat redundant. This feature follows from the relation $g_i \sim h_i$ for DGBD $c_i(r)$ with $\beta_i \approx 1$. The high value of the ordinary least squares regression R^2 indicates that the information contained in h is roughly equivalent to the information contained in g. (Lower panels) By construction, h and Z are also approximately linearly related. However, the presence of larger deviations from the $Z = h$ line, as indicated by slightly lower R^2 values, indicates that for certain profiles, the inclusion of C is crucial to account for the entire rank-citation profile.

reers.

Second, for the entire rank pair $(r_i(A), r_i(B))$ sample we calculate a generalized Kullback-Leibler relative entropy [26]

$$d(r_i(A), r_i(B)) \equiv \sum_{i=1}^{S} \Big(r_i(A) - r_i(B)\Big) \ln \Big(r_i(A)/r_i(B)\Big) \tag{11}$$

to quantify the relative change in the rankings. In each $(r_i(A), r_i(B))$ scatter plot in Fig. 5 we list the normalized value

$$\overline{d} \equiv \frac{d(r_i(A), r_i(B))}{S} \tag{12}$$

calculated for all data, and below in parentheses we list the pair $(\overline{d}_-, \overline{d}_+)$, where $\overline{d}_-$ is calculated for the lower range $r \in [1, S/2]$ and $\overline{d}_+$ is calculated for the upper range $r \in [S/2 + 1, S]$.

It is important to first note the significant differences in the constituents of the physics dataset with respect to the biology dataset. For the physics dataset, the top $S/2$ careers are all highly prolific scientists, and as a result there is significant rank instability for $r \in [1, S/2]$. However, for the biologist dataset, for which there is a significant difference between dataset [E] and the other biology datasets [F,G,H], the ranking for $r \in [1, S/2]$ is rather stable, while for $r \in [S/2+1, S]$ there is consistently larger instability with $\overline{d}_+ > \overline{d}_-$. Since the datasets are not well matched, with exception for the 100 top-cited scientists in each, we do not go further into a cross-comparison.

In summary, the largest "information change" occurs for the (Z, g) pair, and the least for the (h, g) pair, indicating that the Z-index is incorporating additional information into the rankings that h and g are neglecting. This is an important consideration for the large number of careers that are not in the top tier who may experience large rank instability if information from their entire $c_i(r)$ is excluded (say using only h) as compared to when it is included (using h in concert with C).

4.4. Discussion

Here we show that a good indicator should incorporate impact information from the entire $c(r)$ while maintaining simplicity. For $c(r)$ well-described by the DGBD, this would correspond to simply knowing A_i, β_i, γ_i (3 parameters) and N_i (known for each scientist i). Because C is strongly related to β and h through the scaling relation $C \sim h^{1+\beta}$ [16], it suffices to know just two of (β, C, and h). Since β is admittedly tedious to estimate, C and h are the simplest parameters to describe the information contained in $c(r)$. Hence, the Z measure appeals to the two criteria of comprehensive yet simple

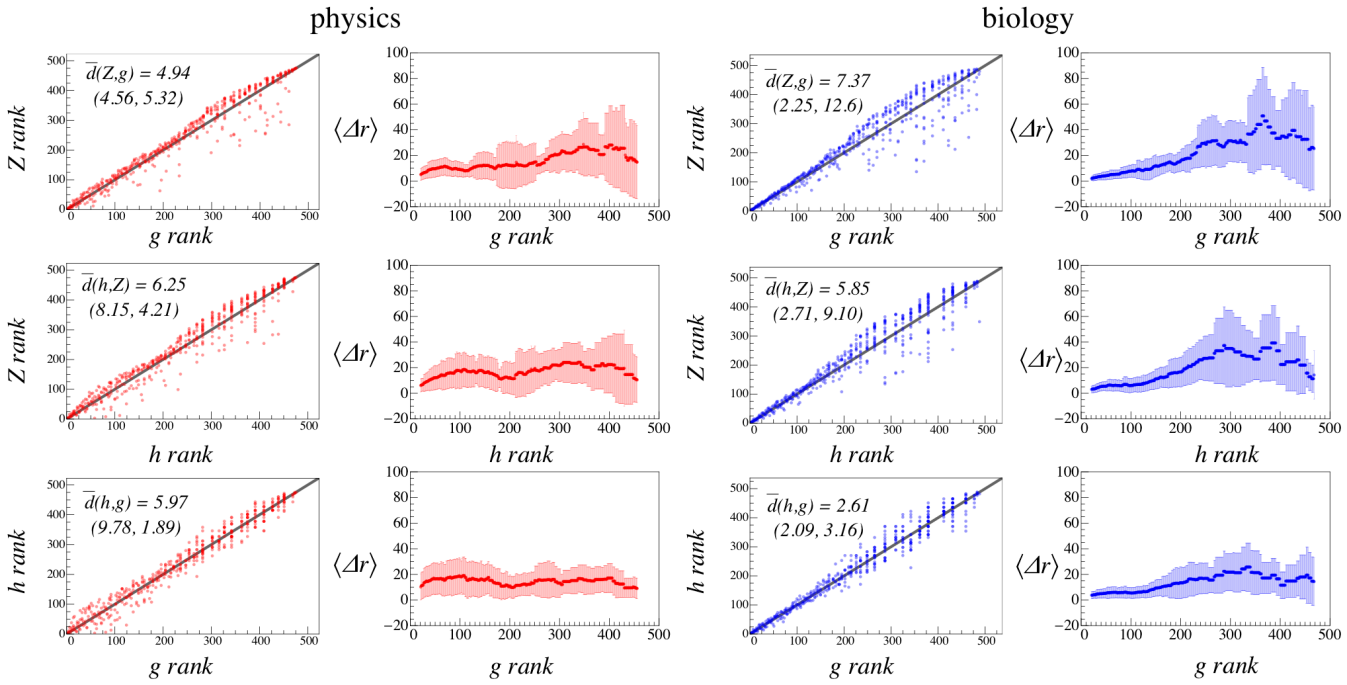

FIG. 5: Quantifying the rank stability of different indices. Panels in columns 1 and 3 show the scatter plots of rank pairs $(r_i(A), r_i(B))$ calculated using the three combinations of rankings according to g_i, h_i, and Z_i calculated for each scientist i. Panels in columns 2 and 4 show the corresponding rank-change distance Δr as a function of the rank value $r_i(A)$. Shown are the running averages $\langle \Delta r \rangle$ calculated over centered bins of width $2\delta r + 1 \equiv 21$ with 1-standard deviation error bars which measure the variations in rank stability. For the index $B = Z$ in the top two rows, there is an increasing rank instability for increasing g and h values, indicating that the Z incorporates information that is neglected, resulting in significant shifts in the overall ranking. We also list the normalized Kullback-Leibler relative entropy $\overline{d}(r_i(A), r_i(B))$ defined in Eq. (12) for each data sample, and in parentheses list the pair $(\overline{d}_-, \overline{d}_+)$ calculated over the lower half and the upper half of each data sample.

(Ockham's razor) by providing a simple geometric representation of the 2-dimensional productivity-impact plane.

Moreover, the Z-index is a very simple generalization of the h-index, corresponding for most scientists to a "renormalization factor" $R(s)$ that is centered around 1. By accounting for the entire citation count C, the Z-index remedies on of the main weaknesses of h, the potentially excessive penalty on a scientist's high-impact papers, without surrendering the simplicity merit of h. In addition, we have shown that the new Z index is more robust towards local changes in the citation profile, an added feature which protects against potentially excessive self-citation strategies. The calculation of the new Z only requires a square root, and could be readily and effortlessly incorporated within current major databases, such as *Google Scholar* and *ResearcherID.com* profiles, which already include h_i and C_i.

5. Conclusions

The availability of high-resolution career data is opening new avenues in computational social science [27], allowing insights into the social mechanisms underlying productivity, competition, achievement, and reward [28–31]. However, in science there has been a proliferation of indices aimed at measuring simultaneously both productivity and impact, or equivalently, to summarize the entire rank-citation profile $c_i(r)$, with a single number. This direction embraces simplicity with the potentially paradoxical outcome of discounting the most notable career achievements.

Here we take a pragmatic approach to measuring the information contained in a scientist's rank-citation profile $c_i(r)$ using a 2-dimensional representation of total citations C_i and h_i. We propose the 2-component measure Z which is simply a vector norm defined for each coordinate pair $(h_i, \sqrt{C_i})$ in the z-plane. Z does not discount a scientists's extremely highly cited papers, is less sensitive to local perturbations, and can be readily calculated using C_i and h_i, which are commonly reported in CVs, websites, award applications, and online publication profile services such *Google Scholar* and *ResearcherID.com*.

It is also important to note that in the practical scenario of career evaluation, if there is going to be a systematic shift towards quantitative measures, then there should also be measures for the multiple other dimensions of an academic career such as collaborativity, publication of influential books, grant writing, teaching awards, mentoring, administrative leadership, and public and policy-oriented outreach, to name but a few. Nevertheless, single-number indicators for productiv-

ity and impact are commonly used because of their objective nature, simplicity, and immediateness. As the scientific labor force continues to grow, and the incentives for producing high-quality scientific products continues to change [32], it will be increasingly important to understand the evaluation measures underlying career appraisal and their implications on the sustainability of career growth [19, 20, 33, 34].

6. Acknowledgments

We thank the anonymous referees for helpful critique. AMP acknowledges COST Action MP0801 and the PNR "Crisis Lab" project at IMT Lucca.

[1] J. Hirsch, Proc. Natl. Acad. Sci. USA **102**, 16569 (2005).
[2] P. D. Batista, M. G. Campiteli, and A. S. Martinez, Scientometrics **68**, 179 (2006).
[3] L. Bornmann and H.-J. Daniel, JASIST **58**, 1381 (2007).
[4] L. Bornmann, R. Mutz, and H.-J. Daniel, JASIST **59**, 001 (2008).
[5] J. E. Iglesias and C. Pecharromán, Scientometrics **73**, 303 (2007).
[6] L. Bornmann and H.-D. Daniel, JASIST **58**, 1381 (2007).
[7] L. Waltman and N. J. van Eck, ERIM Report **ERS-2009-043-LIS** (2009), URL `http://hdl.handle.net/1765/16556`.
[8] L. Bornmann, R. Mutz, S. E. Hug, and H.-D. Daniel, J. Informetrics **5**, 346 (2011).
[9] L. Waltman and N. J. van Eck, ArXiv: **1108.3901** (2011).
[10] ANVUR, *National agency for the evaluation of universities and research institutes (italy)*, http://www.anvur.org/sites/anvurmiur/files/normalizzazione_indicatori_0.pdf (2012).
[11] F. Radicchi, S. Fortunato, and C. Castellano, Proc. Natl. Acad. Sci. USA **105**, 17268 (2008).
[12] F. Radicchi and C. Castellano, Phys. Rev. E **83**, 046116 (2011).
[13] A. M. Petersen, F. Wang, and H. E. Stanley, Phys. Rev. E **81**, 036114 (2010).
[14] F. Radicchi and C. Castellano, J. Informetrics **6**, 121 (2012).
[15] F. Radicchi and C. Castellano, PloS ONE **7**, e33833 (2012).
[16] A. M. Petersen, H. E. Stanley, and S. Succi, Scientific Reports **1**, 181 (2011).
[17] L. Egghe, Scientometrics **69**, 131 (2006).
[18] C.-T. Zhang, Journal of the American Society for Information Science and Technology **62**, 625 (2010).
[19] A. M. Petersen, M. Riccaboni, H. E. Stanley, and F. Pammolli, Proc. Natl. Acad. Sci. USA **109**, 5213 (2012).
[20] A. M. Petersen, S. Fortunato, R. K. Pan, K. Kaski, O. Penner, M. Riccaboni, H. E. Stanley, and F. Pammolli, ArXiv e-print: 1303.7274 pp. 1–40 (2013).
[21] S. Redner, J. Stat. Mech. **2010**, L03005 (2010).
[22] G. G. Naumis and G. Cocho, Physica A **387**, 84 (2008).
[23] G. Martinez-Mekler, R. A. Martinez, M. B. del Rio, R. Mansilla, P. Miramontes, and G. Cocho, PLoS ONE **4**, e4791 (2009).
[24] F. Radicchi and C. Castellano, Scientometrics **DOI 10.1007/s11192-013-1027-3** (2013).
[25] N. Blumm, G. Ghosal, Z. Forró, M. Schich, G. Bianconi, J.-P. Bouchaud, and A.-L. Barabási, Phys. Rev. Lett. **109**, 128701 (2012).
[26] S. Kullback and R. A. Leibler, Ann. of Math. Stat. **22**, 79 (1951).
[27] D. Lazer, A. Pentland, L. Adamic, S. Aral, A.-L. Barabasi, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, et al., Science **323**, 721 (2009).
[28] F. Radicchi, S. Fortunato, B. Markines, and A. Vespignani, Phys. Rev. E **80**, 056103 (2009).
[29] A. M. Petersen, W.-S. Jung, J.-S. Yang, and H. E. Stanley, Proc. Natl. Acad. Sci. USA **108**, 18 (2011).
[30] A. M. Petersen, O. Penner, and H. E. Stanley, Eur. Phys. J. B **79**, 67 (2011).
[31] C. Franzoni, G. Scellato, and P. Stephan, Science **333**, 702 (2011).
[32] P. Stephan, *How Economics Shapes Science* (Harvard University Press, Cambridge MA, USA, 2012).
[33] O. Penner, A. M. Petersen, R. K. Pan, and S. Fortunato, Phys. Today **66**, 8 (2013).
[34] O. Penner, R. K. Pan, A. M. Petersen, K. Kaski, and S. Fortunato, ArXiv e-print: 1306.0114 pp. 1–8 (2013).