New Journal of Physics
The open-access journal for physics

Universal scaling in sports ranking

Weibing Deng[1,2,3], **Wei Li**[1,4,5], **Xu Cai**[1], **Alain Bulou**[3] **and Qiuping A Wang**[2,3]
[1] Complexity Science Center and Institute of Particle Physics, Hua-Zhong (Central China) Normal University, Wuhan 430079, People's Republic of China
[2] LUNAM Université, ISMANS, Laboratoire de Physique Statistique et Systémes Complexes, 44, Avenue Bartholdi, 72000 Le Mans, France
[3] Université du Maine, LPEC, UMR CNRS 6087, 72085 Le Mans, France
[4] Max-Planck Institute for Mathematics in the Sciences, Inselstrasse 22-26, D-04103 Leipzig, Germany
E-mail: liwei@mis.mpg.de

New Journal of Physics **14** (2012) 093038 (15pp)
Received 24 April 2012
Published 20 September 2012
Online at http://www.njp.org/
doi:10.1088/1367-2630/14/9/093038

Abstract. Ranking is a ubiquitous phenomenon in human society. On the web pages of Forbes, one may find all kinds of rankings, such as the world's most powerful people, the world's richest people, the highest-earning tennis players, and so on and so forth. Herewith, we study a specific kind—sports ranking systems in which players' scores and/or prize money are accrued based on their performances in different matches. By investigating 40 data samples which span 12 different sports, we find that the distributions of scores and/or prize money follow universal power laws, with exponents nearly identical for most sports. In order to understand the origin of this universal scaling we focus on the tennis ranking systems. By checking the data we find that, for any pair of players, the probability that the higher-ranked player tops the lower-ranked opponent is proportional to the rank difference between the pair. Such a dependence can be well fitted to a sigmoidal function. By using this feature, we propose a simple

[5] Author to whom any correspondence should be addressed.

toy model which can simulate the competition of players in different matches. The simulations yield results consistent with the empirical findings. Extensive simulation studies indicate that the model is quite robust with respect to the modifications of some parameters.

Contents

1. Introduction

As is well known, ranking is a very interesting and ubiquitous phenomenon in human society [1–5]. On the web pages of Forbes, one can find all kinds of rankings, from the world's most powerful people to the world's richest people, from highest-earning models to America's top colleges, etc. Why does everyone want to top the list? The answer is clear. Being in the top rankings means more power, more wealth and more opportunities. In a ranking system, individuals are not equal anymore, which was confirmed by the famous 80–20 rule [6], namely the Pareto principle, which tells us that for many events, roughly 80% of the effects come from 20% of the causes.

Therefore, the quantities in a ranking system could not be well characterized by their mean values which do not provide any useful information at all. Zipf's law [7–9] states that the frequency of the rth largest occurrence of the event is inversely proportional to its rank, namely $f \sim r^{-\alpha}$, with α approaching 1. Representative examples include word frequencies in text [10–12], people's annual incomes [13–15] and city sizes [16–18].

Our interest then is whether there are some common patterns in the vastly different ranking systems. Moreover, if the answer is yes, can we understand the formalism of such patterns, and unravel some properties of such competition driven systems and human dynamics [19–21]? In order to facilitate our study we choose a specific kind of ranking system, sports ranking, in which data are easily accessible and more suitable for quantitative analysis. Here players' performances in different matches will be used as the basis of their respective rankings, in terms of scores and/or prize money. Amazingly we find that the distributions of scores and/or prize money follow universal power laws, with exponents being nearly identical for different sports. The universal scalings can be reproduced by our toy model in which the key mechanism is concerned with the win–loss probability distribution for any pair of players. This win–loss probability distribution has been verified by the empirical data. Our model is found to be robust with respect to small modifications of minor parts.

2. Empirical results of sports ranking systems

To understand how a certain sports ranking system works [22–24], let us take tennis as an example. Association of Tennis Professionals (ATP) and Women's Tennis Association (WTA) are the world's most successful tennis associations for male and female professionals, respectively. To appear on the ranking systems of ATP or WTA, the number of tournaments a player has to play each year should reach a minimum, say 10. Tournaments have been divided into several categories, such as grand slams, premier tournaments, international tournaments and year-ending tour championships, mainly based on the scale of prize money. For the most important tournaments such as grand slams, the main draw only consists of 128 players. The entry rule is that if you are highly ranked, then you have more chance of being accepted. On the other hand, a player's good performance will improve their rankings which will in turn entitle them to a greater chance of playing tournaments. Since there are so many tournaments each year, for both ATP and WTA, the ranking list of scores and/or prize money may change very frequently. Here we are not interested in which player is world no. 1 in certain sports, but instead in the statistical distribution of performance, measured by scores and prize money, of all the players. What is the form of such a distribution? Is it stable over different time periods? Is it universal?

Our data sets cover 12 different sports, such as tennis, golf, snooker and volleyball, etc. All the data are up-to-date to February 2011 (www.atpworldtour.com, www.wtatennis.com, www.pga.com, www.lpga.com, www.ittf.com, www.fivb.org, www.fifa.com, www.worldsnooker.com, www.fie.ch, www.bwfbadminton.org, www.fiba.com, www.ibaf.org, www.fih.ch, www.ihf.info). We adopt a cumulative distribution due to small system sizes.

2.1. Cumulative distribution of scores

A player's score or prize money is a direct measure of his/her performance in different matches. The higher the score, the better the performance. The statistical distribution of scores or prize money reflects the profile of the performance of all the members belonging to the same association. Every sport has its own scoring system, hence the orders of magnitude of scores are usually different. In order to make the distributions of scores or prize money comparable for different sports, we rescale the quantities of interest. That is,

$$R_S = S/S_{\max}, \tag{1}$$

where S denotes the values of quantities considered, e.g. scores or prize money, and $S_{\max}$ is the maximum value of S in the sample, which pertains to the no. 1 player in the ranking list by using S.

The cumulative distributions of players' scores or prize money have been shown in figure 1 for 40 data samples of 12 different sports ranking systems. Amazingly, all the distributions share a very similar trend, and it should also be noticed that for the same field, all the curves nearly collapse with each other. Therefore, the main task now is to determine which statistical distribution is favored over the others, or equivalently, which statistical distribution is ruled out by the observed data, while others are not.

There are several common statistical distributions [25], such as the power law with exponential decay distribution, $p(x) \sim x^{-\alpha}\,\mathrm{e}^{-\lambda x}$, the exponential distribution, $p(x) \sim \mathrm{e}^{-\lambda x}$, the

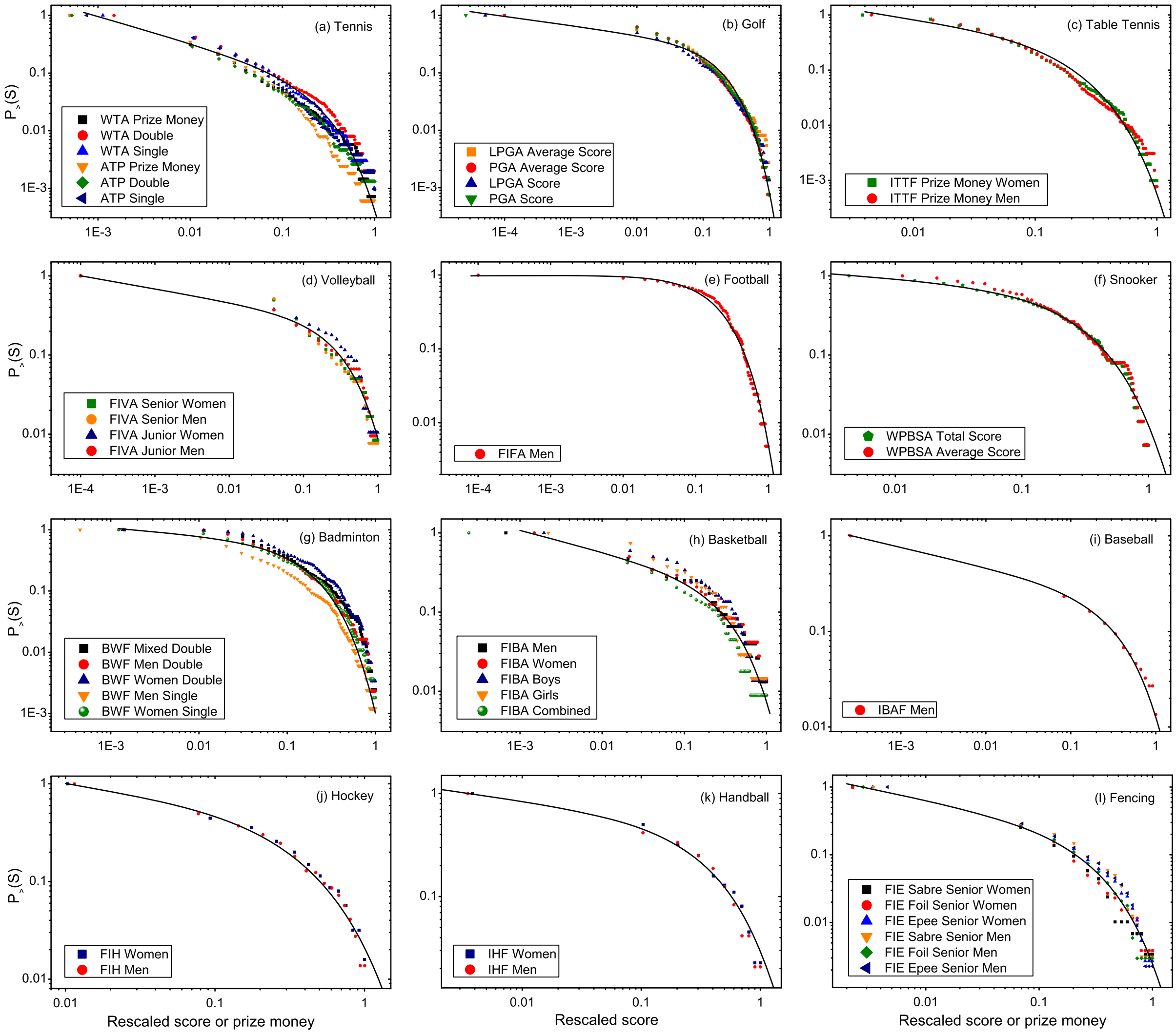

Figure 1. Cumulative distributions of scores and/or prize money for 12 different sports. (a) Tennis: ATP and WTA. (b) Golf: Professional Golfers' Association (PGA) and Ladies Professional Golf Association (LPGA). (c) Table tennis: International Table Tennis Federation (ITTF). (d) Volleyball: International Federation of Volleyball (FIVB). (e) Football: International Federation of Football Association, commonly known as FIFA. (f) Snooker: World Professional Billiards and Snooker Association (WPBSA). (g) Badminton: Badminton World Federation (BWF). (h) Basketball: International Basketball Federation, more commonly known as FIBA. (i) Baseball: International Baseball Federation (IBAF). (j) Hockey: International Field Hockey Federation (FIH). (k) Handball: International Handball Federation (IHF). (l) Fencing: International Fencing Federation (FIE). All the black solid curves in the figures are the power laws with exponential decay, $P_>(S) \propto S^{-\tau} \exp(-S/S_c)$, where τ is the power-law exponent and S_c corresponds to the characteristic turning point of the exponential decay. The values of τ and S_c for different sports are provided in table 1.

stretched exponential distribution, $p(x) \sim x^{\beta-1}\,\mathrm{e}^{-\lambda x^{\beta}}$, and the log-normal distribution, $p(x) \sim \frac{1}{x}\exp[-\frac{(\ln x-\mu)^2}{2\sigma^2}]$, etc. Here, we employ the methods of goodness-of-fit tests in reference [25] to quantify which hypothesis distribution is favored over the others in fitting the data. To do this, we would first determine the least square fitting to the data. Secondly, we calculate the corresponding Kolmogorov–Smirnov (KS) statistics for the goodness-of-fit test of the best-fit hypothesis distribution, then repeat the calculation of the KS statistics for a large number of synthetic data sets. Lastly, we calculate the p-value as the fraction of the KS statistics for the synthetic data sets whose value exceeds the KS statistic for the real data. If the p-value is sufficiently small (say $p < 0.1$), then the hypothesis distribution can be ruled out.

The p-values of the goodness-of-fit tests for the above hypothesis distributions are given in table 1. As one can see, with the hypothesis distribution being the power law with exponential decay, the p-values are all much larger than 0.1. Whereas for the exponential distribution, the p-values are all smaller than 0.1, so the exponential distribution is ruled out. While for the stretched exponential distribution and the log-normal distribution, the majority of p-values are smaller than 0.1, yet a few of them are a little bit larger than 0.1, which implies these two alternative distributions are just good fits in the very rare cases. Therefore, we can conclude that the case of the power law with exponential decay in its favor is strengthened. With the form

$$P_{>}(S) \propto S^{-\tau}\exp(-S/S_{\mathrm{c}}), \tag{2}$$

where τ and S_{c} are exponents of the power law and the exponential decay, respectively, values are shown in table 1, with $0.01 \leqslant \tau \leqslant 0.39$ and $0.12 \leqslant S_{\mathrm{c}} \leqslant 0.28$. Therefore, by using the goodness-of-fit test and checking the values of the fitting parameters, we can observe the shared feature in the sports systems. The evidence of the power laws in the sports ranking indicates that there is still a significant probability having supermen such as Roger Federer in tennis or Tiger Woods in golf. But the prevalent probability is still players who do not play at the top level. Unlike the human height system, it seems there is no typical player who plays at an average level.

Such distributions have also been widely found in a number of different systems, such as the distribution of wealth, city-sizes, word-frequencies, family names, species and degrees of metabolic networks, etc. In [26–28], it is proposed that the shared feature in these systems could be well characterized by the random group formation (RGF), from which a Bayesian estimate is obtained based on the minimal information cost, given the sole *a priori* knowledge of the total number of elements, groups and the number of elements in the largest group. This estimate predicts a unique distribution of the system, with the form

$$P(k) = A\frac{\exp(-bk)}{k^{\gamma}}, \tag{3}$$

where k denotes the elements of the system, and values of A, b and γ are obtained directly from a set of self-consistent equations, while γ usually takes values in the range of $1 \leqslant \gamma \leqslant 2$ [27]. According to the detailed explanations and calculation processes in [27], we applied the RGF predictions to sports systems, with *a priori* knowledge being the total scores of the system M, the number of players N and the highest scores in the system $k_{\max}$. Table 2 gives the values of M, N and $k_{\max}$ of 19 sports systems described above, which are needed for uniquely determining the RGF prediction for each case. By using the same calculation method in [27], we could obtain the values of A, b and γ of the RGF predictions for each sports system (table 2).

Now, we employ the Kolmogorov–Smirnov test [29] (KS test) to compare the RGF predictions with the original probability distributions of scores in sports systems, in order to

Table 1. System sizes of 40 samples in the 12 different sports ranking systems, p-values for the statistical hypothesis test, values of the exponents τ and S_c in the power law with exponential decay, and the ratio of the Pareto principle test.

Sports ranking systems	Sizes	p_1[a]	p_2	p_3	p_4	τ	S_c	Ratio[b]
ATP single	1763	0.65	0.00	0.03	0.06	0.31	0.12	0.79
ATP double	1516	0.52	0.01	0.02	0.03	0.32	0.18	0.78
ATP prize money	1636	0.56	0.02	0.05	0.03	0.33	0.13	0.79
WTA single	1523	0.62	0.00	0.23	0.18	0.39	0.15	0.78
WTA double	1028	0.75	0.00	0.18	0.20	0.38	0.19	0.80
WTA prize money	1388	0.81	0.00	0.21	0.16	0.39	0.12	0.81
PGA score	1323	0.85	0.00	0.00	0.00	0.16	0.18	0.82
LPGA score	734	0.82	0.00	0.00	0.00	0.18	0.19	0.78
PGA average score	1323	0.76	0.00	0.00	0.00	0.16	0.19	0.79
LPGA average score	734	0.82	0.00	0.00	0.00	0.17	0.20	0.82
ITTF prize money men	1717	0.85	0.00	0.02	0.03	0.32	0.17	0.83
ITTF prize money women	1288	0.73	0.00	0.01	0.02	0.32	0.18	0.82
FIVA junior men	105	0.86	0.00	0.00	0.00	0.16	0.21	0.76
FIVA junior women	95	0.68	0.00	0.00	0.00	0.14	0.20	0.79
FIVA senior men	138	0.69	0.01	0.00	0.00	0.13	0.16	0.78
FIVA senior women	127	0.92	0.00	0.00	0.00	0.11	0.18	0.82
FIFA men	209	0.59	0.01	0.00	0.00	0.01	0.19	0.77
WPBSA total score	97	0.69	0.00	0.00	0.00	0.11	0.27	0.83
WPBSA average score	97	0.58	0.00	0.00	0.00	0.13	0.25	0.78
BWF women single	548	0.68	0.00	0.00	0.00	0.12	0.16	0.80
BWF women double	295	0.53	0.00	0.00	0.00	0.13	0.18	0.78
BWF men single	833	0.62	0.00	0.00	0.00	0.06	0.17	0.82
BWF men double	429	0.75	0.00	0.00	0.00	0.08	0.13	0.81
BWF mixed double	407	0.63	0.00	0.00	0.00	0.07	0.14	0.79
FIBA men	79	0.86	0.00	0.00	0.00	0.19	0.20	0.81
FIBA women	72	0.98	0.00	0.00	0.00	0.18	0.21	0.83
FIBA boys	77	0.62	0.00	0.00	0.00	0.18	0.23	0.82
FIBA girls	72	0.85	0.00	0.01	0.01	0.26	0.22	0.76
FIBA combined	115	0.52	0.01	0.00	0.00	0.23	0.20	0.81
IBAF men	78	0.96	0.00	0.00	0.00	0.20	0.28	0.79
FIH men	73	0.86	0.00	0.00	0.00	0.23	0.26	0.78
FIH women	68	0.83	0.00	0.00	0.00	0.21	0.27	0.81
IHF men	52	0.68	0.00	0.00	0.00	0.16	0.25	0.79
IHF women	46	0.69	0.00	0.00	0.00	0.15	0.27	0.76
FIE sabre senior women	371	0.56	0.00	0.12	0.08	0.34	0.25	0.81
FIE foil senior women	260	0.65	0.00	0.03	0.00	0.32	0.23	0.78
FIE epee senior women	293	0.53	0.01	0.16	0.17	0.36	0.24	0.83
FIE sabre senior men	319	0.67	0.00	0.00	0.02	0.32	0.23	0.78
FIE foil senior men	337	0.56	0.00	0.00	0.00	0.30	0.21	0.82
FIE epee senior men	442	0.72	0.00	0.01	0.00	0.28	0.25	0.81

[a] P-values of goodness-of-fit tests [25], with the hypothesized distribution being the power law with the exponential decay distribution (p_1), the exponential distribution (p_2), the stretched exponential distribution (p_3) and the log-normal distribution (p_4).
[b] Values of the ratio for the test of the Pareto principle.

Table 2. Basic quantities in the RGF predictions of the sports systems. M: the total score of the system; N: the total number of players; $k_{\max}$: the highest score; k_0: the lowest score. A, γ, b and k_c refer to four parameters in the procedure of the RGF prediction [27]. D and p denote the maximum differences D and p-values in the KS tests, while 'BWF M' and 'BWF W' mean 'BWF Men' and 'BWF Women', respectively.

Sports	M	N	$k_{\max}$	k_0	A	γ	b	k_c	D	p
ATP	227 193	1763	7965	6	1.044	1.44	3.91×10^{-4}	6138	0.5098	0.000
WTA	276 279	1523	8835	10	1.261	1.42	3.68×10^{-4}	6864	0.6667	0.000
BWF M	4 777 609	548	81 706	110	0.0018	0.363	7.72×10^{-5}	69 345	0.9048	0.000
BWF W	5 191 108	833	89 002	40	0.0169	0.661	6.37×10^{-5}	74 812	0.7619	0.000
PGA	30 690	1323	384	1	0.158	0.793	0.014 9	325	0.6190	0.000
LPGA	22 779	734	590	1	0.174	0.919	0.007 9	483	0.7143	0.000
ITTF M	195 176	1717	2706	20	3.045	1.501	0.001 5	2193	0.8095	0.000
ITTF W	180 106	1288	2728	23	1.942 1	1.373	0.001 5	2225	0.7097	0.000
FIVA M	2626	138	210	1	0.180	0.826	0.017 6	163	0.5238	0.004
FIVA W	2411	127	200	1	0.174	0.803	0.018 5	155	0.4516	0.002
FIBA M	6921	79	892	1	0.090	0.755	0.003 7	665	0.4762	0.011
FIBA W	6976	72	940	1	0.082	0.733	0.003 5	699	0.5161	0.000
FIH M	36 964	73	2620	30	0.0030	0.058	0.002 0	2122	0.6153	0.000
FIH W	36 079	68	2700	35	0.0029	0.065	0.001 9	2180	0.8571	0.000
IHF M	2 600	52	286	1	0.0381	0.265	0.015 2	224	0.7095	0.000
IHF W	2 326	46	261	1	0.028 3	0.141	0.017 3	205	0.8182	0.000
FIE M	8 593	319	290	1	0.1137	0.622	0.017 4	239	0.5806	0.000
FIE W	9 149	371	294	1	0.1336	0.696	0.016 8	242	0.6364	0.012
IBAF M	7 377	78	986	1	0.091	0.771	0.003 3	731	0.7273	0.000

quantify whether the RGF prediction could characterize the sports data. With the null hypothesis being the sports data that follows the RGF prediction, we calculated the maximum differences D and p-values in the KS tests for the 19 sports systems. From table 2, one can find that all the p-values are much smaller than 0.05, which suggests all the KS tests reject the null hypothesis at the 5% significance level. Therefore, we can draw the conclusion that the RGF predictions could not be used to characterize the sports data.

The possible reason is that the data samples of the sports systems are quite small, which might lead to large uncertainty. We also conjecture that the differences between the two kinds of systems might be caused by different mechanisms of formation. For sports systems, competition is the main driving force. Whether a player's rank will be raised or lowered, depends not only on his own performance but also on the other player's. In sports there is very little of the 'rich-gets-richer' mechanism which is dominant in city sizes, human wealth, etc.

2.2. The Pareto principle

The Pareto principle [6], also known as the 80–20 rule, states that, for many events, roughly 80% of the effects come from 20% of the causes. Pareto noticed that, 80% of Italy's land was owned by 20% of the population. He carried out such surveys on a variety of other countries, and to his surprise, the rule was also found to be true.

The 80–20 rule has also been used to attribute the widening economic inequality, which showed the distribution of global income to be very uneven, with the richest 20% of the world's population controlling 82.7% of the world's income. The 80–20 rule could be applied to many systems, from the science of management to the physical world.

We also check this rule in relation to sports ranking systems. It is interesting to find that 20% of players indeed possess approximately 80% of the scores or prize money of the whole system. The ratios obtained from different sports ranking systems are shown in table 1, the values of the ratios all being very close to 0.8.

2.3. Dependence of win probability on Δrank

Here we employ the concept of 'win probability' to describe the chances that a player or a team will win when encountering an opponent. For instance, what are the odds that a no. 1 player will beat a no. 100 player? And what are the odds against a no. 2 player? Theoretically, the chance is much higher in the former case than it is in the latter. But the result of a competition is not known until it is over, which mainly depends on how the player performs at that specific match. However, the win probability could be solely based on the previous performance of a player against a certain opponent, which then can be used to predict her future performance against the same opponent. This might have some applications in betting on the result of a match. To simplify the case without loss of generality, we relate the win probability solely to the ranking difference of a pair of players. Suppose we now have two players A and B, with A having a higher rank. We will then need to know how likely it is that A can beat B when they meet. This quantity is related to, but different from, the win percentage we usually refer to. The win percentage depicts the percentage of wins over all previous encounters. We assume that the win probability only depends on the rank difference between two players. This means that the probability that no. 1 beats no. 100 is the same as that for no. 100 beats no. 200. Hence, we have the following definition:

$$P_{\mathrm{win}}(\Delta r) = \frac{N_{\mathrm{win}}(\Delta r)}{N_{\mathrm{total}}(\Delta r)}, \tag{4}$$

where Δr denotes the rank difference (integer), $N_{\mathrm{win}}(\Delta r)$ is the total number of wins for the higher-ranked player when the rank difference is Δr, and $N_{\mathrm{total}}(\Delta r)$ is the total number of matches in which the rank difference between the pair is Δr. We here emphasize again that the win probability is the probability that the higher-ranked player will win when two players meet. When Δr is small, say 1, it is difficult to judge which player will win, and in this case P_{win} might approximately equal 0.5. When Δr is large, for instance 100, P_{win} might approach 1, which means the higher-ranked player is very likely to win.

By using the *Head to Head* records of ATP and WTA, we find that the dependence of P_{win} on Δr can be characterized well by the Bradley–Terry model [30] for paired comparisons as follows:

$$P_{\mathrm{win}} = \frac{1}{1 + \exp\left(-a * \Delta r\right)}, \tag{5}$$

where a is a parameter dependent on the specific systems. For ATP and WTA, a is 0.021 and 0.032, respectively (figure 2). The existence of fluctuations is quite natural since even Roger Federer will not win every match. The value of a can still give us some information about how competitive a certain sport is. The smaller a is, the more competitive the sport will be. Let us

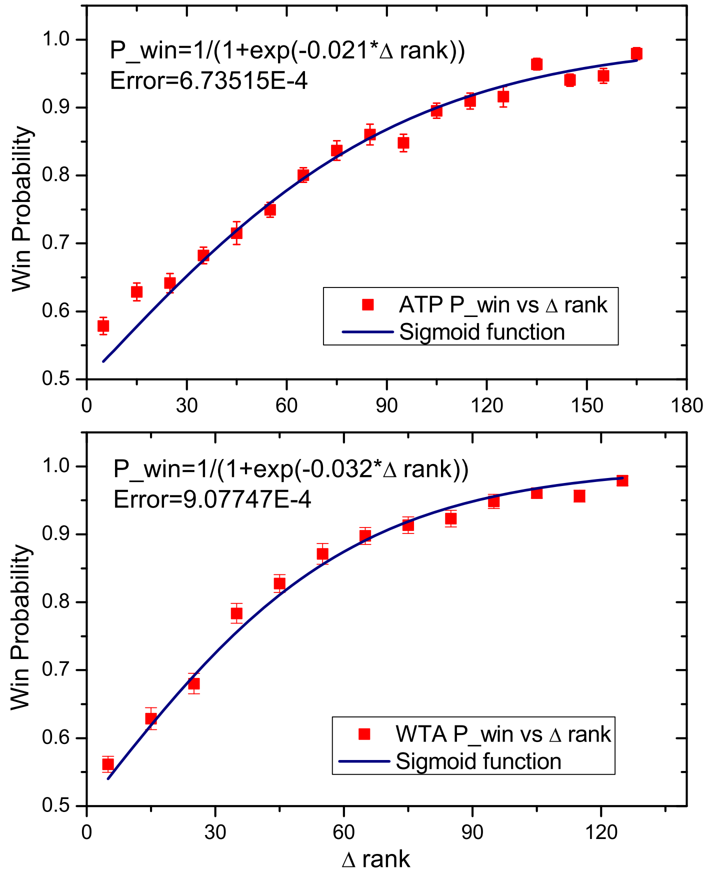

Figure 2. Dependence of win probability on Δr of the players for ATP and WTA, which can be well fitted to the sigmoidal function $P_{\rm win} = 1/(1+\exp(-a*\Delta r))$, with $a = 0.021$ and 0.032 for ATP and WTA, respectively.

take WTA and ATP as two examples. When Δr is 30, the win probability for WTA is nearly 0.7, while the counterpart for ATP is 0.65. This means the game is more unpredictable in ATP than in WTA, which is not strange since the men's game is more competitive than the women's.

The competitiveness parameter a plays a key role in both the empirical analysis of the win probability and the simulations of the toy model, so we explain the differences between the different systems in two respects.

For the empirical part, we really wish to test the empirical findings by checking data from different sports, other than in the tennis field of ATP and WTA. The problem is that the data source for *Head to Head* records is very limited in other sports, to the best of our knowledge. Alternatively, we present here the trend of the functional form of the win probability in figure 3, in which $a = 0.01$, 0.015 and 0.03 may correspond to three different sports systems. As one can see, for the same Δrank, the competition becomes stronger when a gets smaller.

3. A simple toy model of sports ranking systems

What is the origin of the universal scaling in different sports systems? Of course, there have been so many approaches that can explain the origin of power laws. Some mechanisms or theories

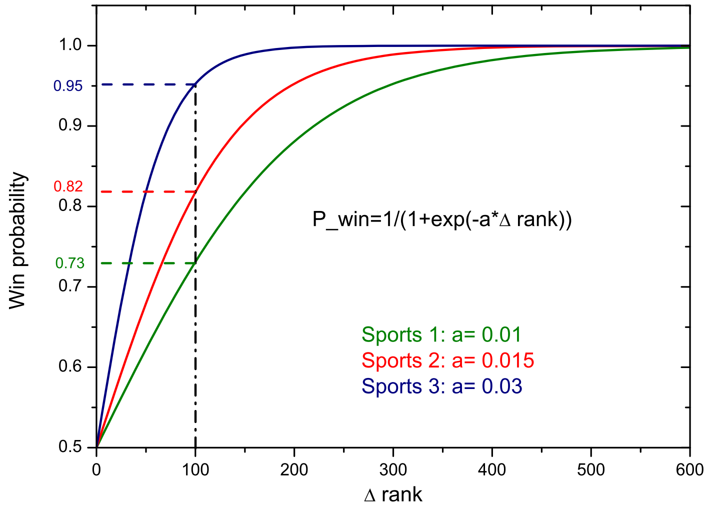

Figure 3. Theoretical curves of the win probability formula, $P_{\text{win}} = 1/(1 + \exp(-a * \Delta r))$, with $a = 0.01$, 0.015 and 0.03 for three different sports systems, respectively.

are elegant, e.g. random walks [31] and self-organized criticality [32, 33], etc. It is, however, difficult to try to apply these frameworks to sports ranking systems. We propose a simple toy model, inspired by tennis. Of course, the model may not suit any sport but it does have some general implications. Most importantly, our model can reproduce robust power laws without having to introduce additional parameters.

The rules of the model are defined in the following way:

1. 2^N players are ranked from 1 to 2^N, being assigned random scores drawn from a Gaussian distribution.
2. For each tournament, all the players have entry permission. Therefore the draw will include 2^N players and in total N rounds. At each round, half of the players will be eliminated when they lose. The rest will enter the next round. The losers at round n will gain a score $2^{(n-1)}$. The final champion wins a score 2^N (figure 4).
3. The key mechanism is to decide which one will lose for a given pair of players. Here our empirical findings will be employed. Namely, when two players meet, the probability that the higher-ranked player will beat the lower-ranked opponent is given by $1/(1 + \exp(-a * \Delta r))$, where Δr is their rank difference, as before.
4. A new tournament opens up and a new draw is made.

In principle, there is only one parameter in our model, that is a. We can simply call it the competitiveness parameter. Of course, there are some shortcomings in the model. First, in the actual tournaments not all the players will be accepted. In grand slams there are only 128 players. Second, tournaments can be divided into many categories and may consist of different players. Third, the scoring systems for different tournaments are a little different. For grand slams the scores and prize money are much higher than other tournaments, if the players are eliminated at the same round. We certainly can add these issues into our model in order to test the resilience of the model. At the moment we do not wish to complicate the model by introducing

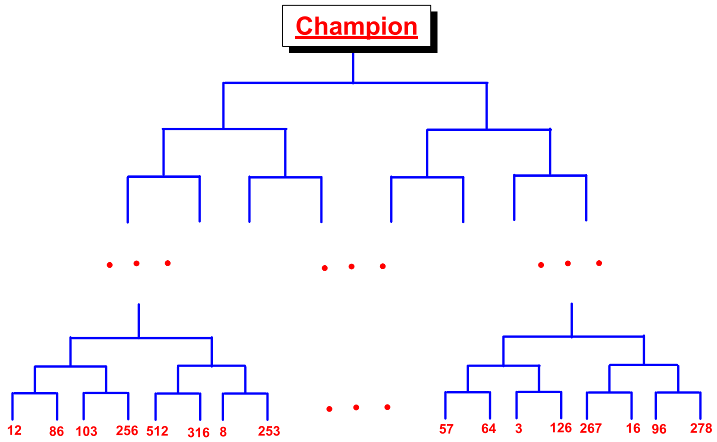

Figure 4. A cartoon of a draw sample. After each round, half the players will be eliminated, the numbers '12, 86 ...' denote the ranks of the players.

additional parameters. What we need here is a skeleton which will allow us to understand some key features of the specific systems. Namely, if the power laws with exponential decay can be reproduced through our model, then it is a feasible model. We need not care about other minor issues.

4. Simulation results and discussions

The most important parameter in our model is a, the so-called competitiveness parameter. The number of players N_p and the number of tournaments N_t only have finite-size effects. It is natural to check the dependence of the simulation results on these parameters, which can reflect the resilience of our model.

First of all, we need to test whether the model can reproduce the power laws of the cumulative distribution of scores. In figure 5, N_p equals 2048, and N_t is 128, while the win probability, $P_{win} = 1/(1+\exp(-a * \Delta r))$, with $a = 0.021$ and 0.032, as given by the empirical data of ATP and WTA, respectively. Here, we also use the same goodness-of-fit test, and p-value equal to 0.85 and 0.91 for the two distributions, respectively. Therefore, the cumulative distributions of scores given by the simulations indeed follow the power-law distributions with exponential decay, $P(S) \propto S^{-\tau} \exp(-S/S_c)$, with $\tau = 0.2, 0.22$, $S_c = 0.23, 0.19$, respectively for these two samples. Here, we notice that the values of the parameters are very close to what are obtained from the empirical data.

In the formula of win probability, smaller values of a correspond to more intensive competition. For instance, when $a = 0.0001$, $P_{win} \leqslant 0.525$ for $\Delta r \leqslant 1000$, which means the higher-ranked player has only a slightly greater chance than the lower-ranked player of winning the match between them. While larger values of a suggest that the higher ranked player would win the match with a much larger probability. For example, when $a = 2.0$, $P_{win} \geqslant 0.88$ for $\Delta\text{rank} \geqslant 1$.

Thus, to analyze the influence of win probability, we simulated our models with different values of a, $0.0001 \leqslant a \leqslant 2.0$. From figure 6, we can find that, as a gets smaller, the values of τ will become larger, while those of S_c will become smaller. When a is very small, such as $a = 0.0001$, the cumulative score distributions change from the power laws with exponential

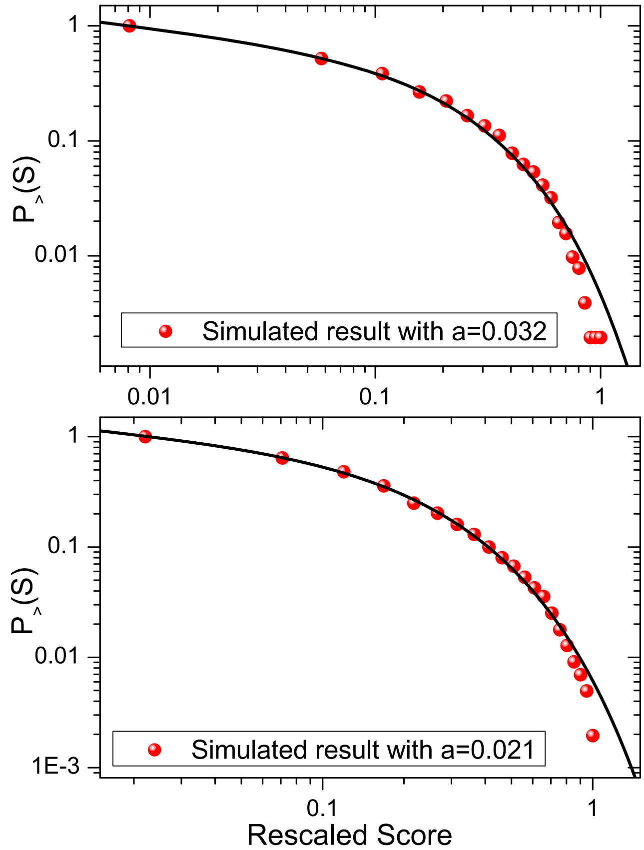

Figure 5. Cumulative distribution of scores from the simulation. For these two samples, the number of players $N_{\mathrm{p}} = 2048$ and the number of total tournaments $N_{\mathrm{t}} = 128$, with $P_{\mathrm{win}} = 1/(1 + \exp(-a * \Delta r))$, $a = 0.032$ and 0.021, respectively.

decay to exponential form. Since, in this case, all players nearly win the match randomly, thus the cumulative probabilities of the scores approximates 1, $1/2$, $(1/2)^2, \ldots$, which results in the exponential distribution.

For a different number of tournaments, $N_{\mathrm{t}} = 64$, 128 and 256, the cumulative distributions of scores are shown in figure 7. As seen, all the cumulative distributions of scores are power laws with exponential decay, values of the exponents τ and S_{c} being also very close to those of the empirical results.

In statistical physics, in order to determine the validity of the statistical approach, we often take the thermodynamic limit, in which the number of components N tends to infinity [34]. However, in real-world networks, the number of vertices or agents can never be that large and therefore we need to study the finite-size effect. For example, even the largest artificial net, the World Wide Web, whose size will soon approach 10^{11}, also shows a qualitatively strong finite-size effect [35].

Therefore, in order to test the influence of the finite-size effect on the final cumulative distribution of scores, we consider the transformed score distribution $P(S) * S^{\tau}$ versus S/S_{c}, where S_{c} is the characteristic turning point of the exponential decay. For four different system

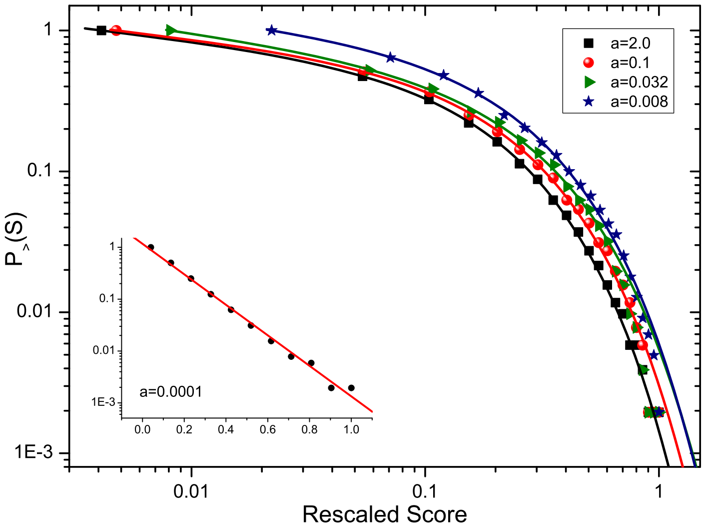

Figure 6. Influence of the critical parameter a on the final cumulative score distributions, values of a ranging from 0.0001 to 2.0.

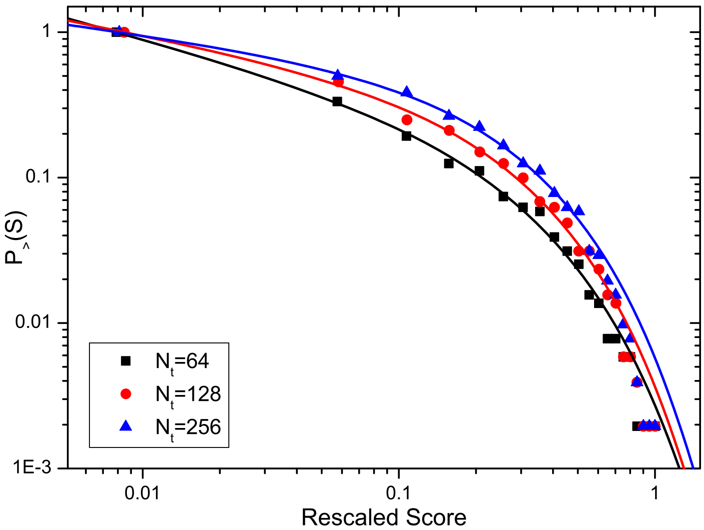

Figure 7. Simulation result of the cumulative score distributions for a different number of tournaments, with $N_{\mathrm{t}} = 64$, 128 and 256.

sizes, such relationships were shown in figure 8, which suggests that the tails of the four curves almost collapse with each other.

As this model is a simple toy model, the major goal is that it could reproduce the trend of empirical findings of cumulative score distributions. Therefore the predictive power of the model is rather modest. We do not think it could be a general framework for all kinds of sports systems. However, we are planning to enrich the model by considering more ingredients so that the model could be more powerful. Of course in doing so we might have to consider the cost of introducing additional parameters. This will be investigated in future work.

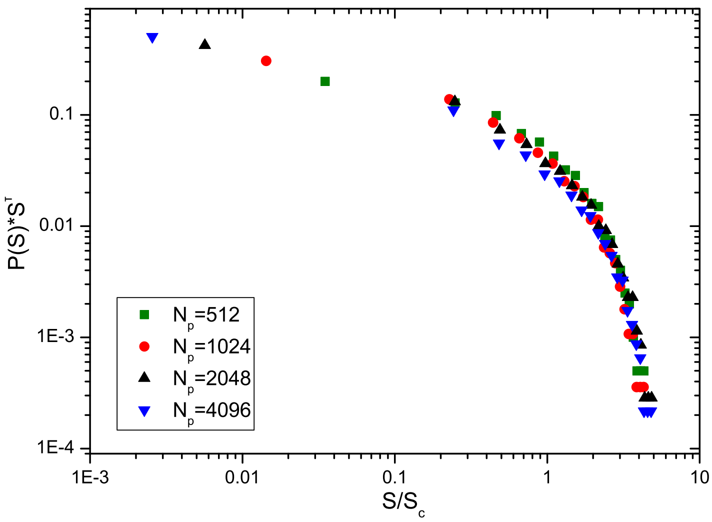

Figure 8. Finite-size effects analysis of the simulation results, with $N_p = 512$, 1024, 2048 and 4096.

5. Conclusion

In summary, to characterize the intrinsic common features and underlying dynamics of ranking systems, we study sports ranking systems. Our main results are as follows. (i) Universal scaling is found in the distributions of scores and/or prize money, with values of the power exponents being close to each other for 40 samples of 12 sports ranking systems. (ii) Players' scores are found to obey the Pareto principle, which means approximately 20% of players possess 80% of the total scores of the whole system. (iii) Win probability is introduced to describe the chance that a higher-ranked player or team will win when meeting a lower-ranked opponent. We relate the win probability solely to the rank difference Δr, and for tennis the win probability has been empirically verified to follow the sigmoid function, $P_{\rm win} = 1/(1 + \exp(-a * \Delta r))$. (iv) By employing the empirical features of win probability, we propose a simple toy model to simulate the process of the sports systems, and the universal scaling could be reproduced well by our model. This result is quite robust when we change the values of parameters in the model.

Acknowledgments

WBD would like to acknowledge Cyril Pujos and Laurent Nivanen for fruitful discussions. This work was supported by National Natural Science Foundation of China (grant nos. 10975057, 10635020, 10647125 and 10975062), the Programme of Introducing Talents of Discipline to Universities under grant no. B08033, and the PHC CAI YUAN PEI Programme (LIU JIN OU (2010) no. 6050) under grant no. 2010008104. Partial support also comes from the Max Planck Society.

References

[1] Squire R A 1981 *Science* **214** 877
[2] Shorrocks A F 1983 *Economica* **50** 3
[3] Kechedzhi K E, Usatenko O V and Yampol'skii V A 2005 *Phys. Rev.* E **72** 046138

[4] Deschâtres F and Sornette D 2005 *Phys. Rev.* E **72** 016112
[5] Radicchi F, Fortunato S, Markines B and Vespignani A 2009 *Phys. Rev.* E **80** 056103
[6] Pareto V and Page A N 1971 *Translation of Manuale di Economia Politica* (*'Manual of Political Economy'*) (New York: A M Kelley)
[7] Zipf G K 1932 *Selected Studies of the Principle of Relative Frequency in Language* (Cambridge, MA: Harvard University Press)
[8] Zipf G K 1949 *Human Behavior and the Principle of Least Effort* (Cambridge, MA: Addison-Wesley)
[9] Perline R 1996 *Phys. Rev.* E **54** 220
[10] Miller G A and Newman E B 1958 *Am. J. Psychol.* **71** 209
[11] Li W T 1992 *IEEE Trans. Inform. Theory* **38** 1842
[12] Ney H, Essen U and Kneser R 1994 *Comput. Speech Lang.* **8** 1
[13] Pareto V 1986 *Cours d'Economie Politique* (Genève: Droz)
[14] Sornette D and Zajdenweber D 1998 *Eur. Phys. J.* B **8** 653
[15] Aoyama H, Souma W, Nagahara Y, Okazaki M P, Takayasu H and Takayasu M 2000 *Fractals* **8** 293
[16] Gabaix X 1999 *Q. J. Econ.* **114** 739
[17] Ioannidesa Y M and Overmanb H G 2003 *Reg. Sci. Urban Econ.* **33** 127
[18] Marsili M and Zhang Y C 1998 *Phys. Rev. Lett.* **80** 2741
[19] Petersen A M, Wang F and Stanley H E 2010 *Phys. Rev.* E **81** 036114
[20] Petersen A M, Jung W S, Yang J S and Stanley H E 2011 *Proc. Natl Acad. Sci. USA* **108** 18
[21] Petersen A M, Stanley H E and Succi S 2011 *Sci. Rep.* **1** 181
[22] Radicchi F 2011 *PLoS One* **6** e17249
[23] Ben-Naim E, Redner S and Vazquez F 2007 *Europhys. Lett.* **77** 30005
[24] Ben-Naim E, Vazquez F and Redner S 2007 *J. Korean Phys. Soc.* **50** 124
[25] Clauset A, Shalizi C R and Newman M E J 2009 *SIAM Rev.* **51** 4
[26] Minnhagen P 2011 *Phys. Today* http://www.physicstoday.org/daily_edition/points_of_view/just_and_unjust_distributions
[27] Baek S K, Bernhardsson S and Minnhagen P 2011 *New J. Phys.* **13** 043004
[28] Baek S K, Minnhagen P and Kim B J 2011 *New J. Phys.* **13** 073036
[29] Massey F J 1951 *J. Am. Stat Assoc.* **46** 253
[30] Bradley R A and Terry M E 1952 *Biometrica* **39** 324
[31] Newman M E J 2005 *Contemp. Phys.* **46** 323
[32] Bak P, Tang C and Wiesenfeld K 1987 *Phys. Rev. Lett.* **59** 381
[33] Bak P, Tang C and Wiesenfeld K 1988 *Phys. Rev.* A **38** 364
[34] Toral R and Tessone C J 2007 *Commun. Comput. Phys.* **2** 177
[35] Dorogovtsev S N, Goltsev A V and Mendes J F F 2008 *Rev. Mod. Phys.* **80** 1275