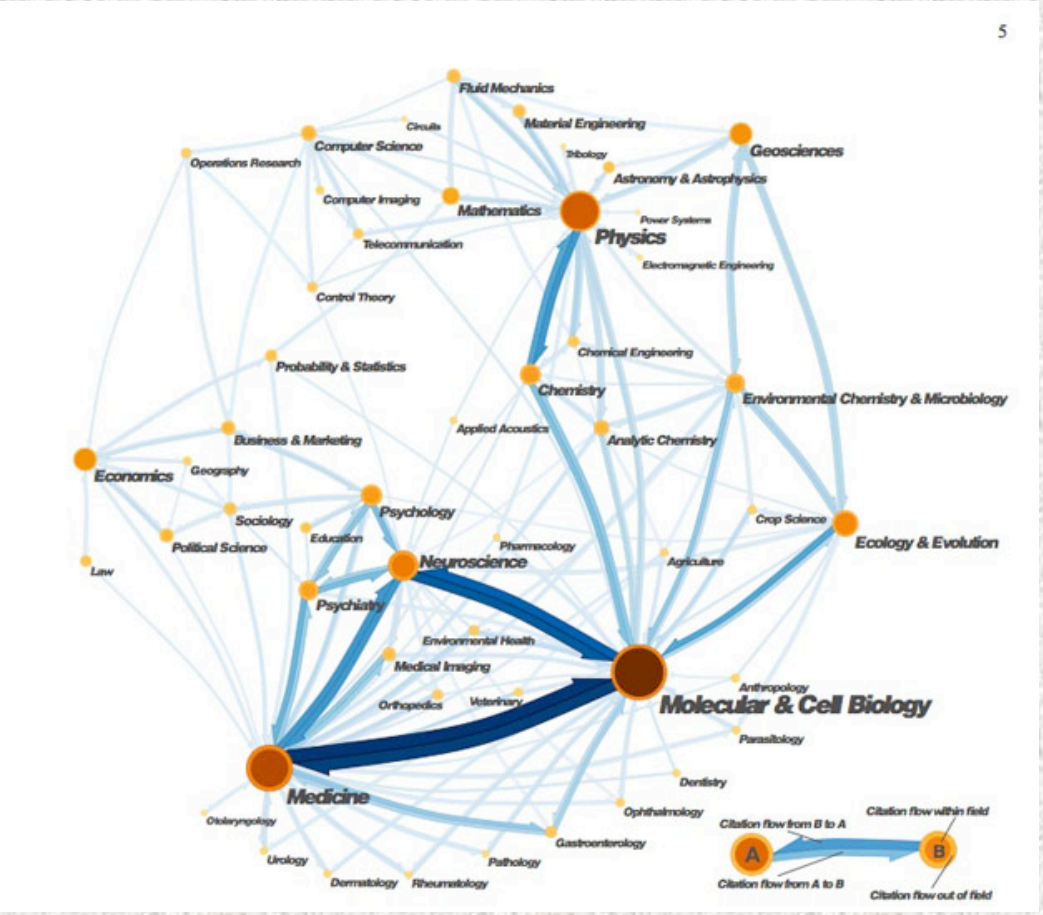
Citation networks: unfolding the dynamics of scientific impact

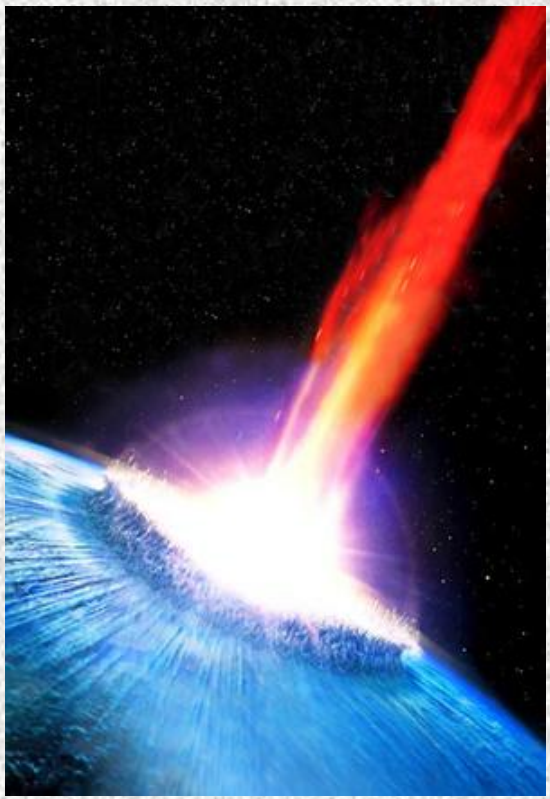
Santo Fortunato

Spectral properties of complex networks
ECT* Trento, July 23-27, 2012

Aalto University in winter ...

Citations = impact?





Outline

- Universality of citation distributions
- The World Citation Network
- The World Collaboration Network
- Citation boosts: the rise of Nobel laureates

Citation statistics

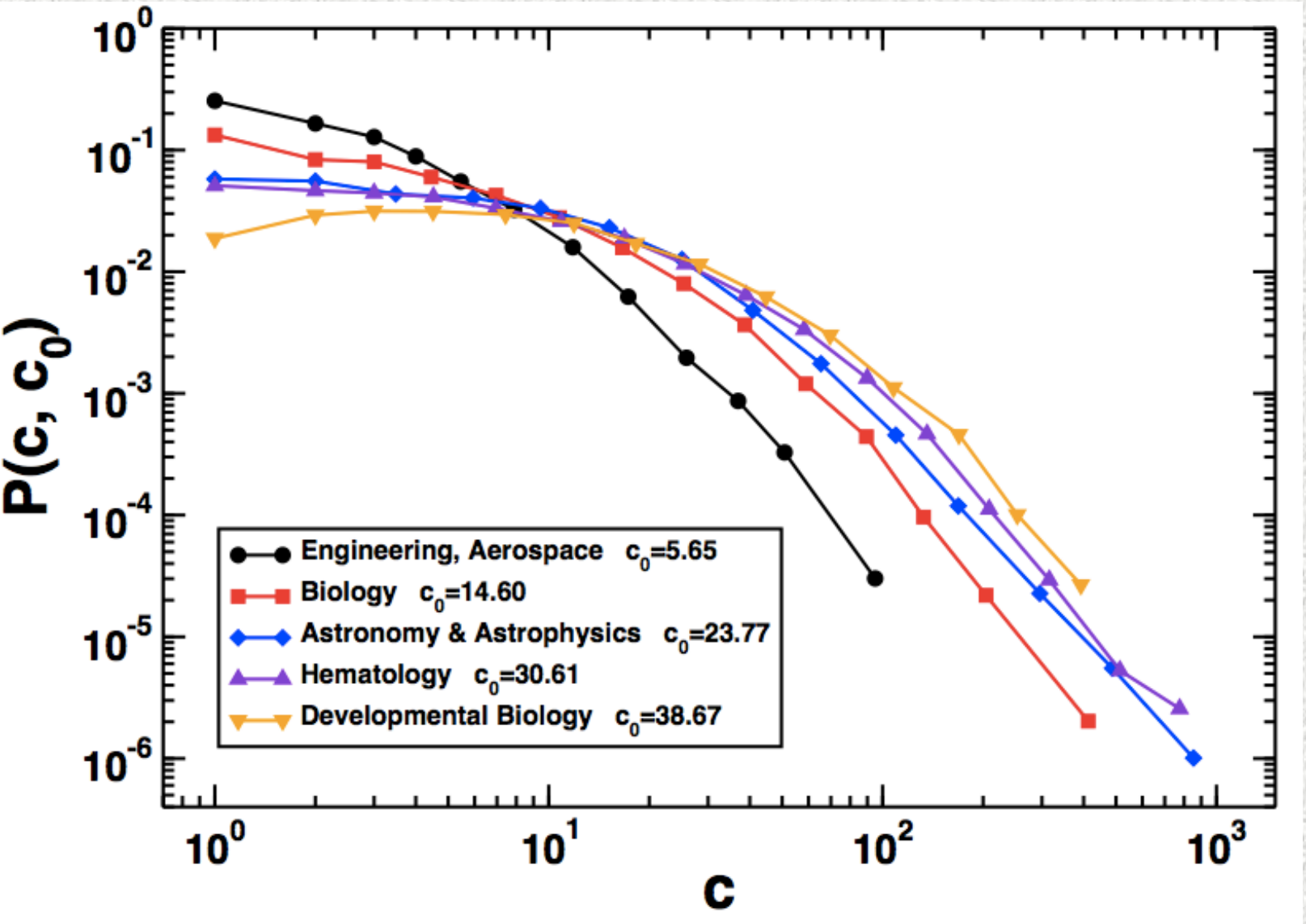
Source of data

ISI Web of Knowledge℠

Index	Subject category	Year	N_p	c_0	c_{max}	σ^2	χ^2/df
1	Agricultural economics and policy	1999	266	6.88	42	1.0 (1)	0.007
2	Allergy	1999	1,530	17.39	271	1.4 (2)	0.012
3	Anesthesiology	1999	3,472	13.25	282	1.8 (2)	0.009
4	Astronomy and astrophysics	1999	7,399	23.77	1,028	1.1 (1)	0.003
5	Biology	1999	3,400	14.6	413	1.3 (1)	0.004
6	Computer science, cybernetics	1999	704	8.49	100	1.3 (1)	0.004
7	Developmental biology	1999	2,982	38.67	520	1.3 (3)	0.002
8	Engineering, aerospace	1999	1,070	5.65	95	1.4 (1)	0.003
9	Hematology	1990	4,423	41.05	1,424	1.5 (1)	0.002
10	Hematology	1999	6,920	30.61	966	1.3 (1)	0.004
11	Hematology	2004	8,695	15.66	1,014	1.3 (1)	0.003
12	Mathematics	1999	8,440	5.97	191	1.3 (4)	0.001
13	Microbiology	1999	9,761	21.54	803	1.0 (1)	0.005
14	Neuroimaging	1990	444	25.26	518	1.1 (1)	0.004
15	Neuroimaging	1999	1,073	23.16	463	1.4 (1)	0.003
16	Neuroimaging	2004	1,395	12.68	132	1.1 (1)	0.005
17	Physics, nuclear	1990	3,670	13.75	387	1.4 (1)	0.001
18	Physics, nuclear	1999	3,965	10.92	434	1.4 (4)	0.001
19	Physics, nuclear	2004	4,164	6.94	218	1.4 (1)	0.001
20	Tropical medicine	1999	1,038	12.35	126	1.1 (1)	0.017

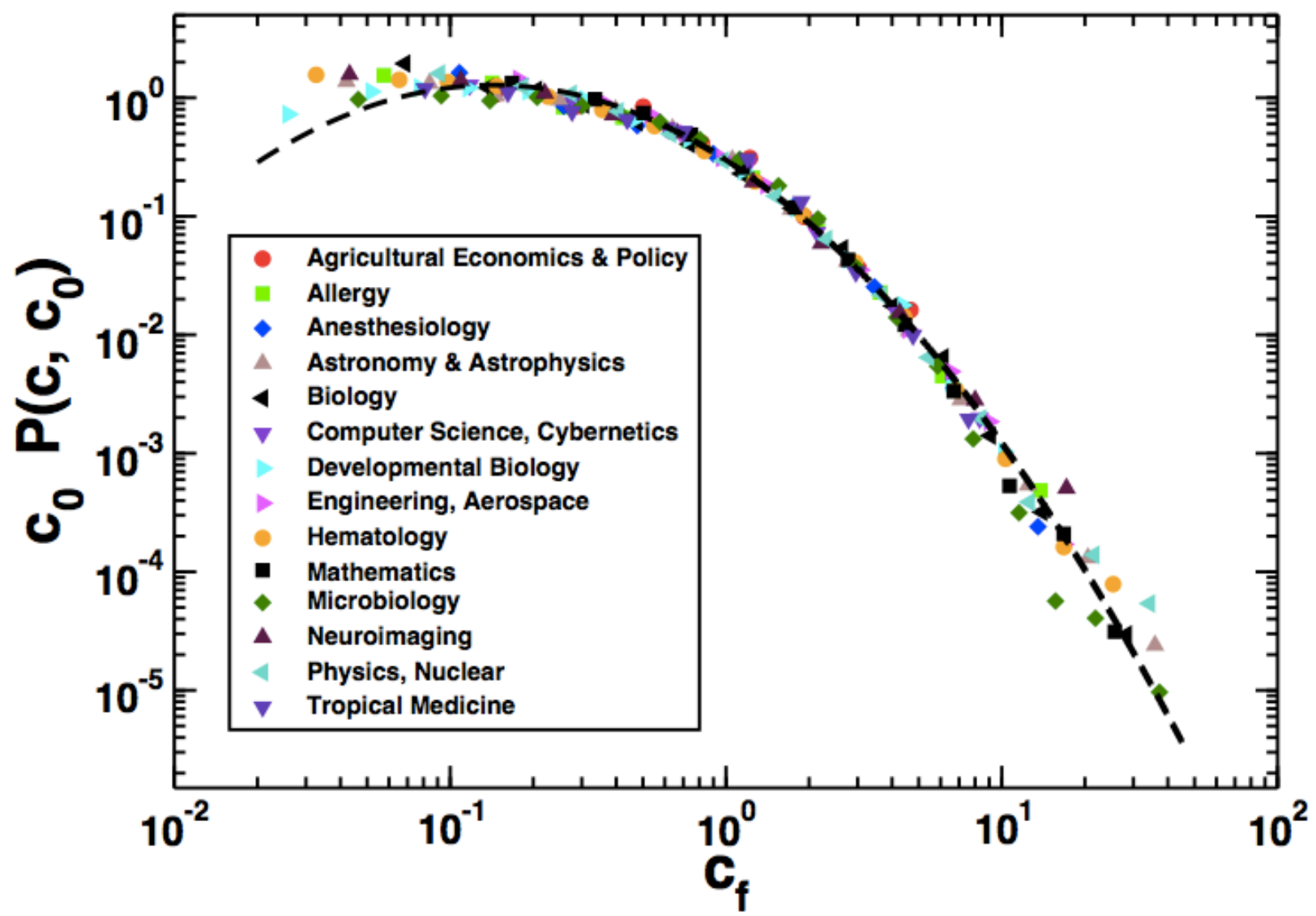
Papers are classified in 172 scientific disciplines (from *Acoustics* to *Zoology*)

Distribution of cites?

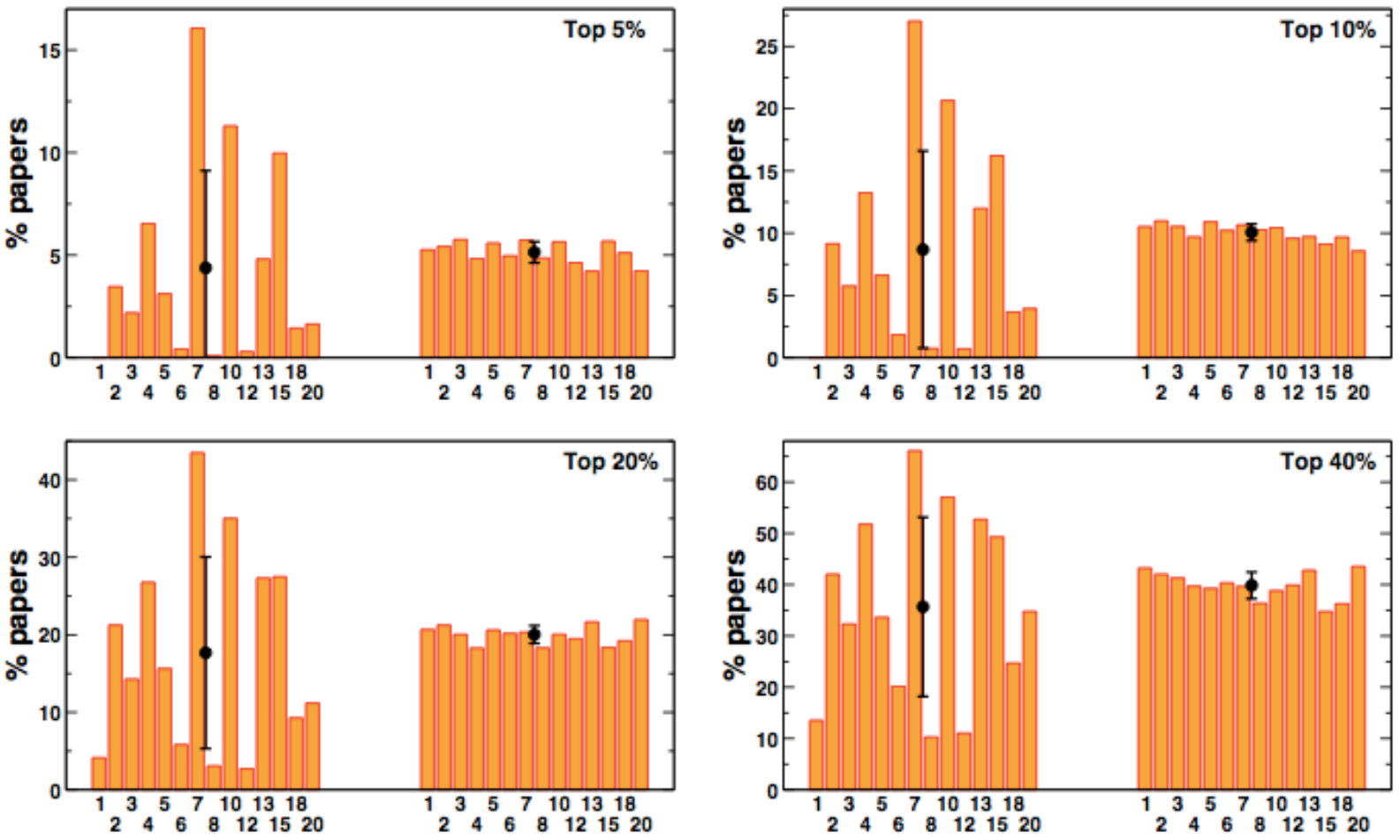


Dependence on field (ISI category)!

Could c_0 be the reason of the discrepancy?

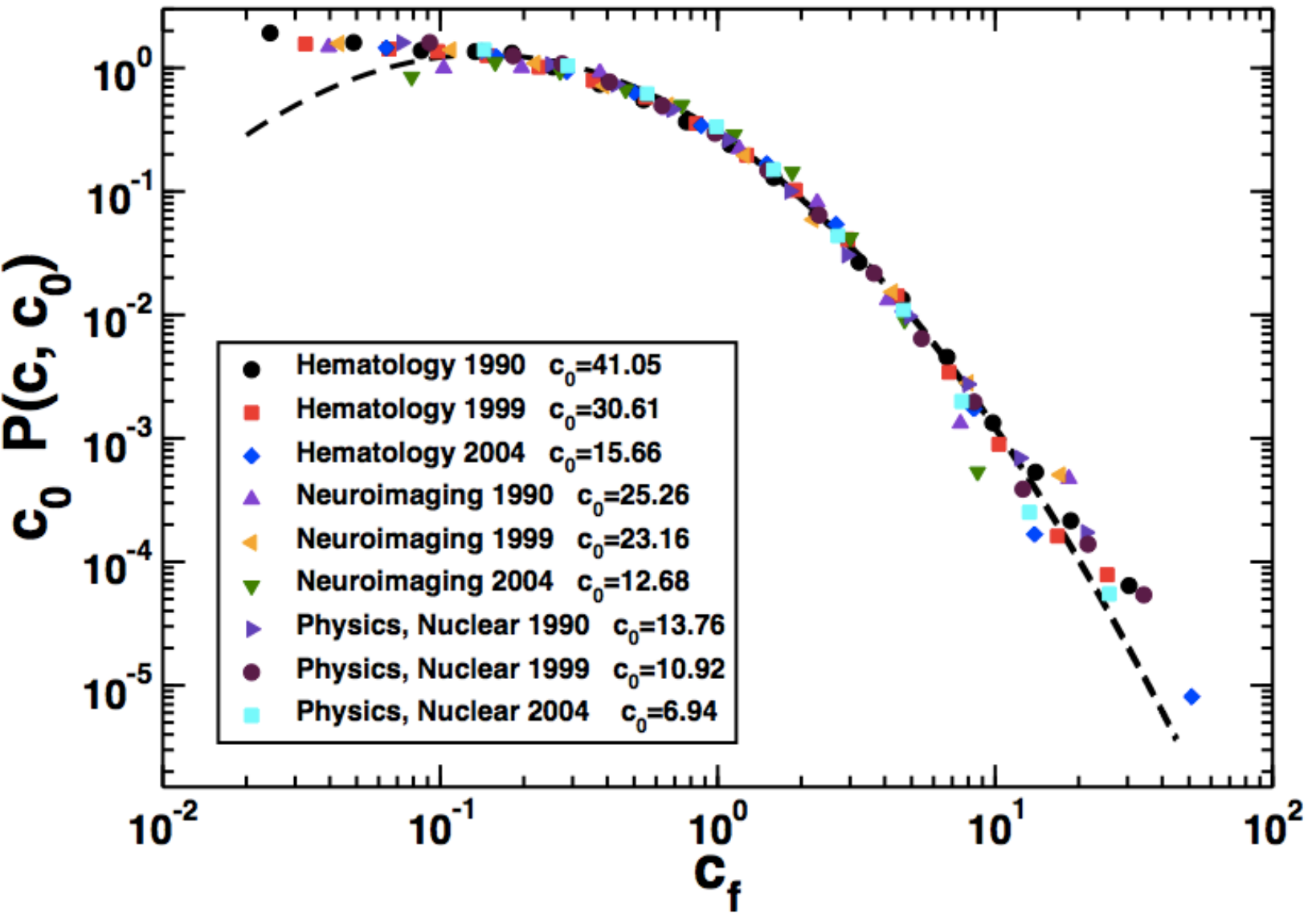


F. Radicchi, S.F. and C. Castellano,
Proc. Natl. Acad. Sci. USA 105, 17268 (2008)



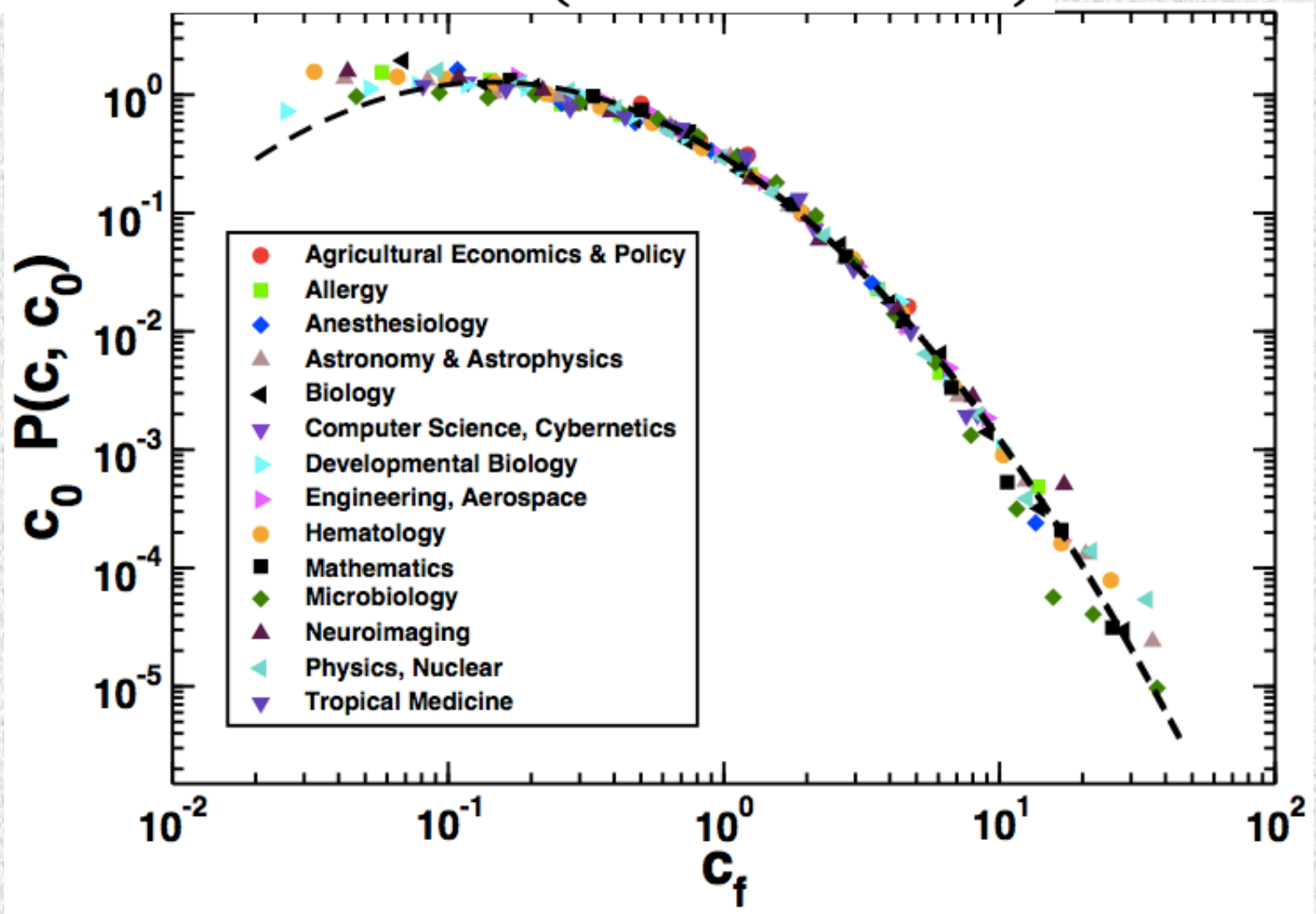
F. Radicchi, S.F. and C. Castellano,
Proc. Natl. Acad. Sci. USA 105, 17268 (2008)

The universal distribution is stable in time!



Fitting the universal distribution

$$Q(x) = \frac{1}{\sqrt{2\pi}\, x\, \sigma} \exp\left\{ -\frac{\left[\ln(x) + \sigma^2/2\right]^2}{2\,\sigma^2} \right\}, \quad x = c/c_0$$



Summary I

- The distribution of the number of citations of papers in the same discipline, normalized by the average citation score, is universal!
- It is possible to compare the impact of papers in different disciplines in an objective way
- Relative citation indicators could lead to more reliable indices of individual performance than, say, the H-index

The World Citation Network

Goal: studying the geographic distributions and correlations of citation flows

Data: Thomson Reuters (ISI Web of Science) database, from 2003 until 2010

Author affiliations

New Journal of Physics > Volume 9 > June 2007 Create ToC alert RSS this journal

Jukka-Pekka Onnela *et al* 2007 *New J. Phys.* **9** 179 doi:10.1088/1367-2630/9/6/179

Analysis of a large-scale weighted network of one-to-one human communication

FOCUS ON COMPLEX NETWORKED SYSTEMS: THEORY AND APPLICATION

Jukka-Pekka Onnela[1,2], Jari Saramäki[1], Jörkki Hyvönen[1], Gábor Szabó[3,4], M Argollo de Menezes[3], Kimmo Kaski[1], Albert-László Barabási[3,4] and János Kertész[1,5]

Hide affiliations

kertesz@phy.bme.hu

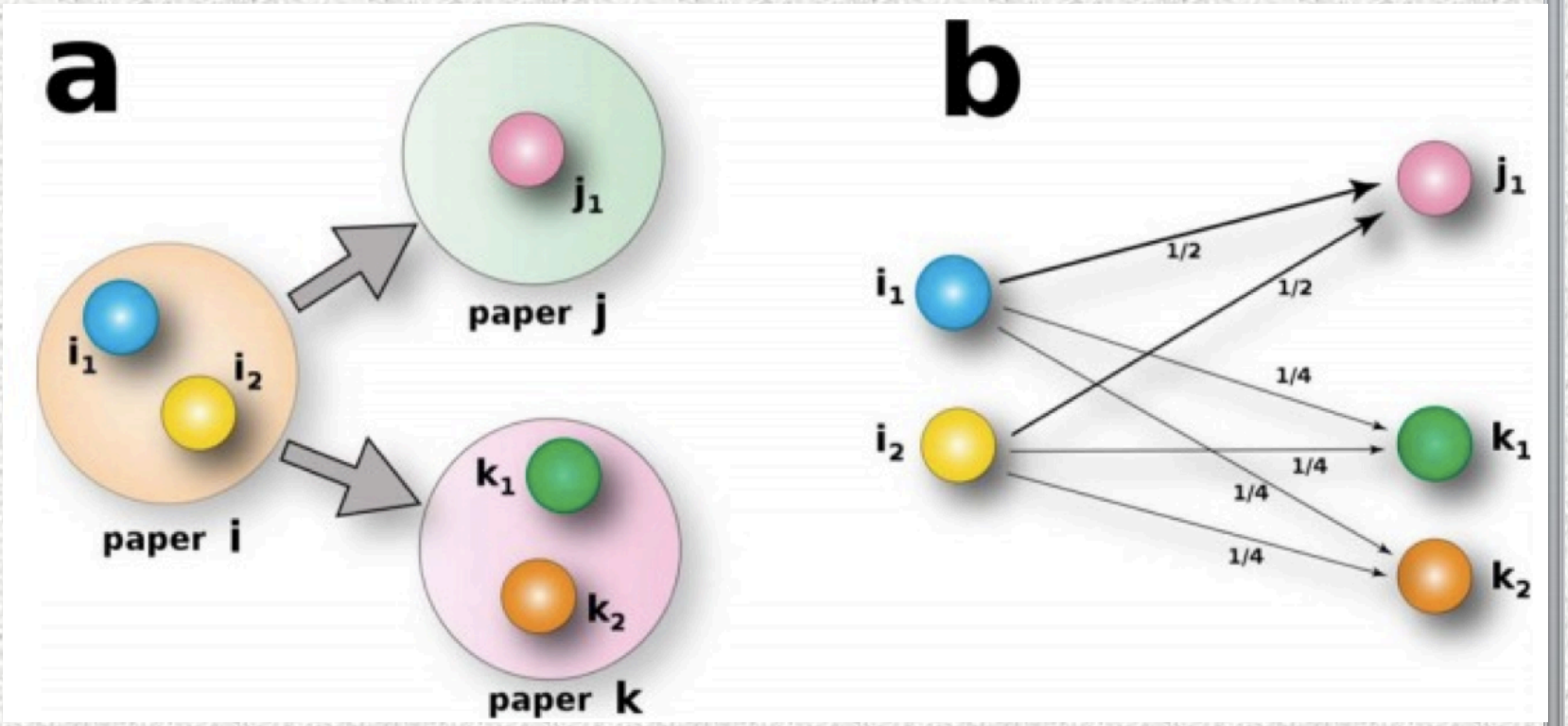
[1] Laboratory of Computational Engineering, Helsinki University of Technology, Finland
[2] Clarendon Laboratory, Physics Department, Oxford University, Oxford, UK
[3] Department of Physdics and Center for Complex Networks Research, University of Notre Dame, IN, USA
[4] Center for Cancer Systems Biology, Dana Farber Cancer Institute, Harvard University, Boston, MA, USA
[5] Department of Theoretical Physics, Budapest University of Technology and Economics, Budapest, Hungary

Author affiliations

Finland, UK, USA, USA, Hungary

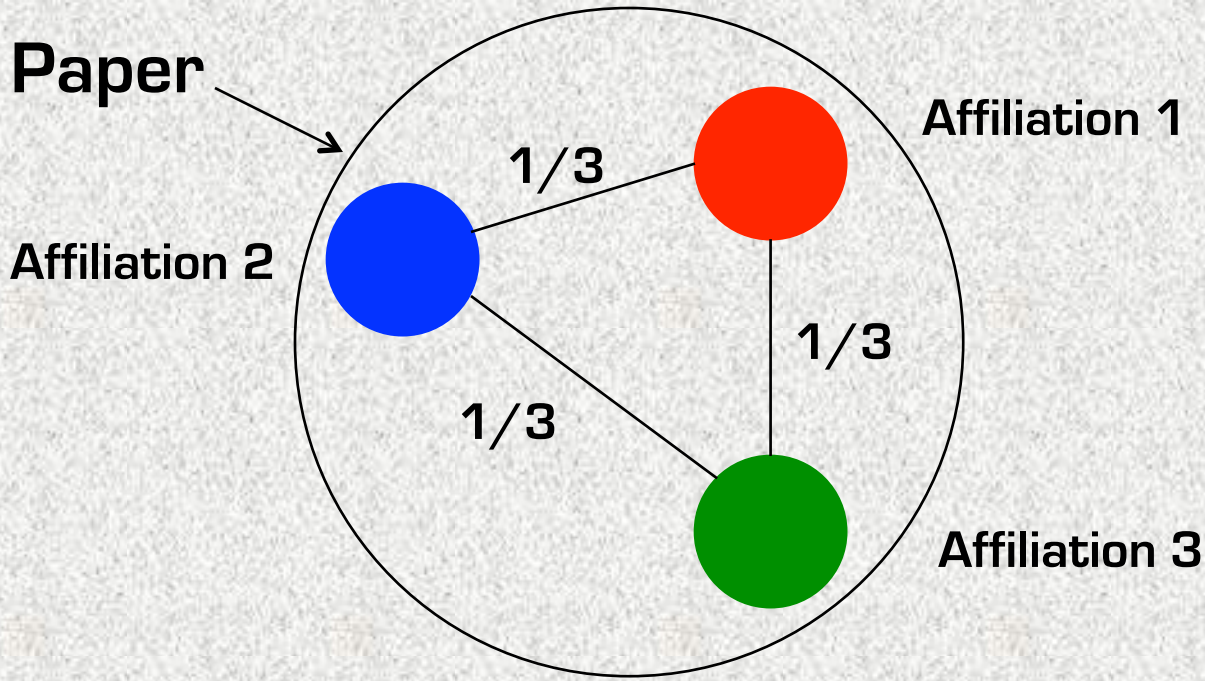
Espoo, Oxford, South Bend, Cambridge (MA), Budapest

Citation networks



Citations are split among the cited authors, and then they are attributed to the countries/cities of the authors

Collaboration networks

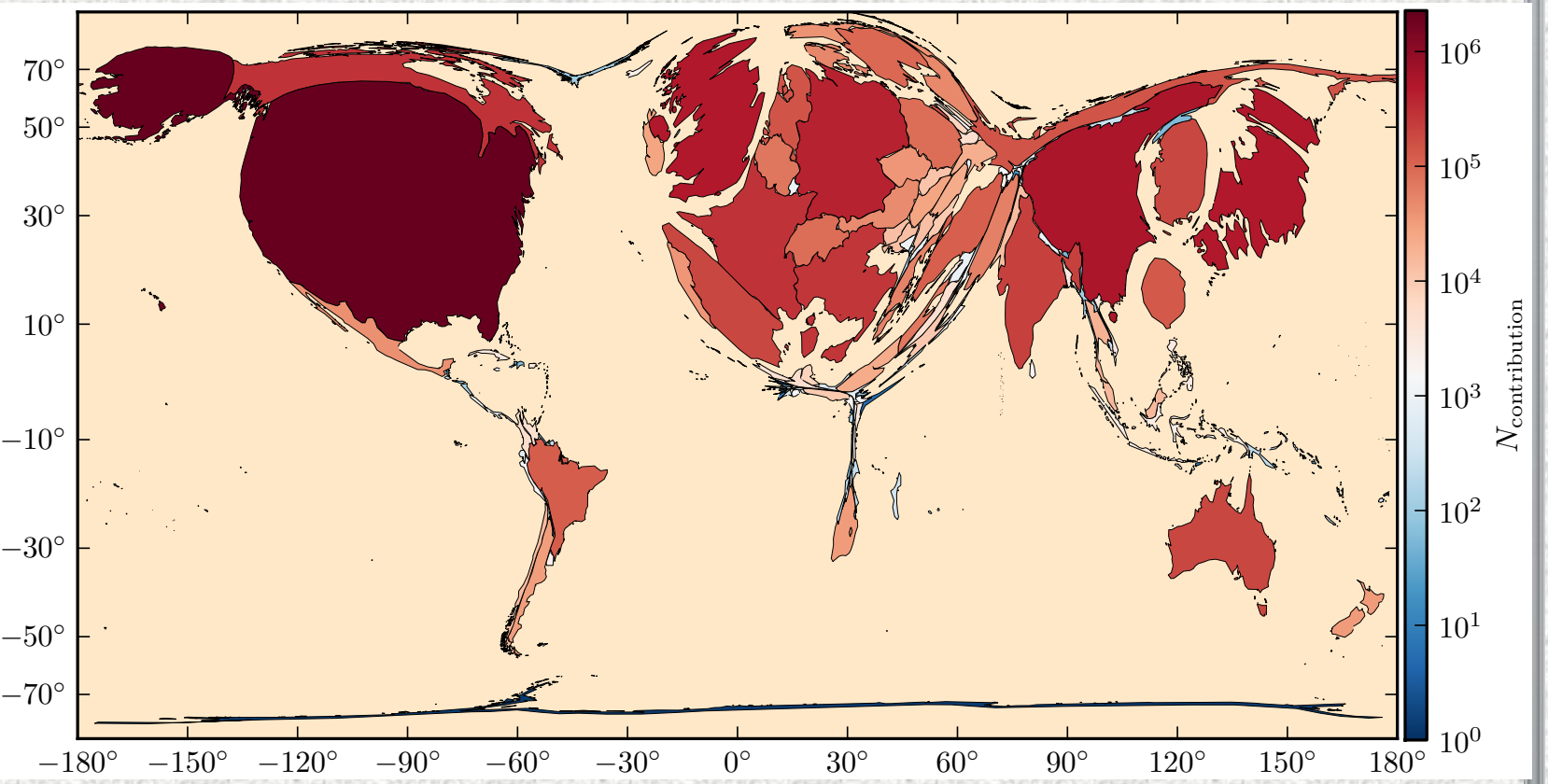


Collaboration links receive a weight of 2/[n(n-1)], where n is the number of different countries/cities involved in the paper

Summary of statistics

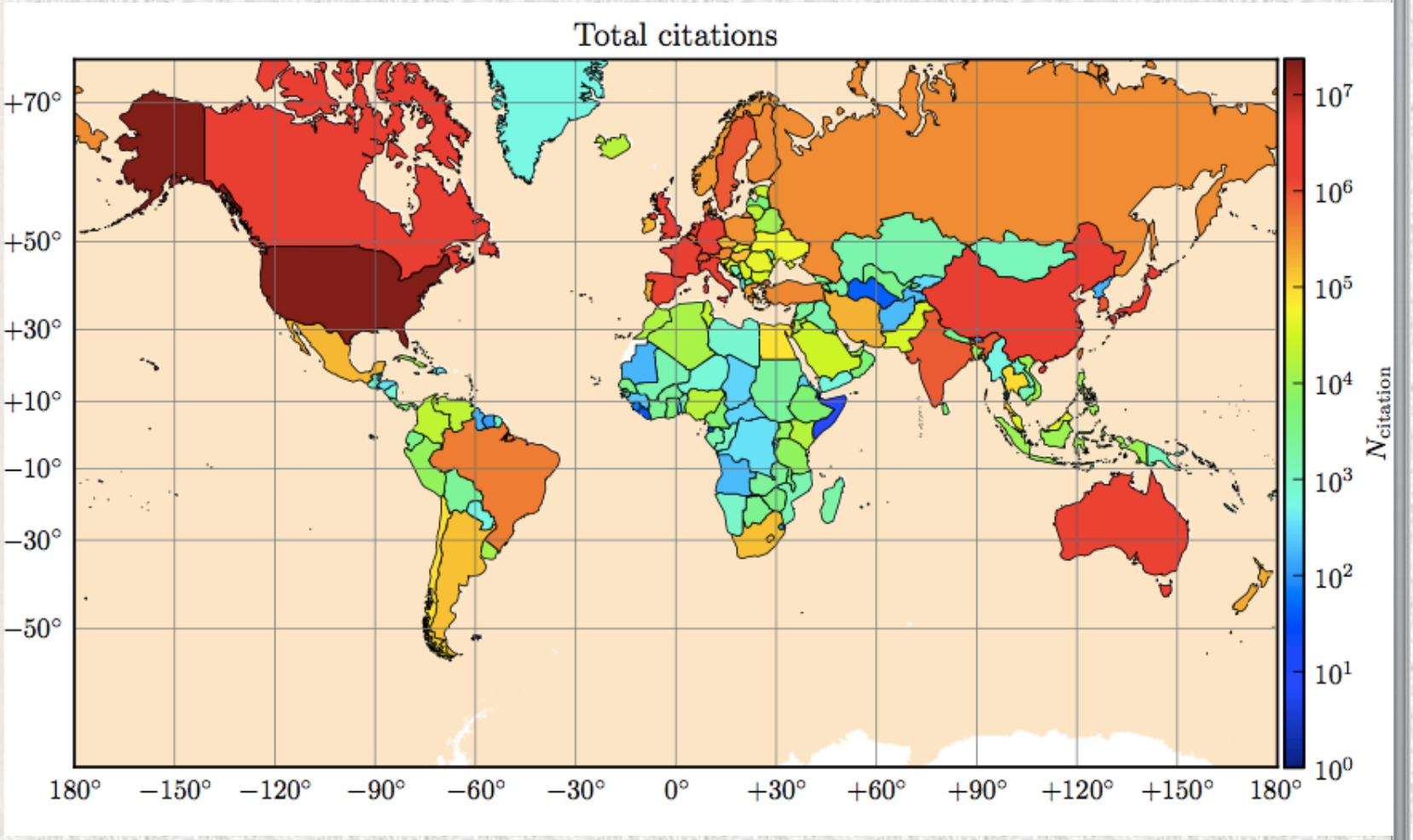
Continent	$N_{\text{Countries}}$	N_{Cities}	f_{Pub} (in %)	f_{Cite} (in %)	Avg. Cites	Country name	f_{Pub} (in %)	f_{Cite} (in %)	Avg. Cites	$f^{\text{Pub}}_{\text{intra-city}}$ (in %)	$f^{\text{Pub}}_{\text{intra-country}}$ (in %)
Africa	57	749	1.32	0.65	5.00±0.05	South Africa	0.43	0.25	5.94±0.08	45.6	54.6
						Egypt	0.29	0.13	3.78±0.04	51.7	63.0
						Tunisia	0.10	0.04	3.34±0.14	36.3	46.3
						Nigeria	0.13	0.03	2.83±0.23	54.2	73.6
						Kenya	0.04	0.03	7.58±0.26	15.3	18.8
Asia	49	3853	27.36	17.71	5.58±0.01	Japan	6.46	5.94	7.68±0.03	45.6	76.3
						China	7.22	4.30	5.06±0.02	52.7	70.8
						South Korea	2.51	1.58	5.39±0.03	47.5	73.6
						India	2.73	1.40	4.34±0.03	65.4	80.8
						Taiwan	1.67	1.04	5.19±0.03	52.7	80.0
						Israel	0.86	0.84	8.86±0.09	43.0	58.6
						Turkey	1.45	0.67	3.89±0.03	65.1	83.2
						Russia	1.87	0.62	3.92±0.04	55.9	63.7
						Singapore	0.52	0.46	7.29±0.08	54.9	54.9
						Iran	0.75	0.31	3.30±0.03	61.6	79.0
Europe	47	6625	33.69	35.25	9.29±0.01	United Kingdom	6.51	7.45	9.92±0.04	44.3	57.7
						Germany	5.13	6.30	10.41±0.04	36.8	50.6
						France	3.61	4.03	9.67±0.04	31.2	48.6
						Italy	3.41	3.26	8.59±0.04	39.8	59.1
						Netherlands	1.83	2.33	11.10±0.07	35.2	50.7
						Spain	2.48	2.26	8.10±0.05	43.9	57.8
						Switzerland	1.11	1.60	12.38±0.09	32.0	38.7
						Sweden	1.23	1.44	10.60±0.08	35.8	48.2
						Belgium	0.92	1.00	10.01±0.08	34.4	43.4
						Denmark	0.66	0.84	11.43±0.10	32.4	45.4
						Finland	0.64	0.67	9.61±0.13	35.7	52.1
North America	37	5346	32.40	42.33	10.36±0.02	United States	28.12	38.22	10.67±0.02	50.6	74.0
						Canada	3.62	3.73	9.15±0.04	46.2	57.8
						Mexico	0.52	0.29	5.54±0.10	39.7	54.8
						Puerto Rico	0.04	0.03	7.68±0.27	28.4	32.5
Oceania	21	844	2.89	2.67	8.22±0.05	Australia	2.45	2.30	8.35±0.05	41.7	59.2
						New Zealand	0.43	0.35	7.61±0.10	43.3	51.7
South America	14	782	2.34	1.39	5.75±0.04	Brazil	1.55	0.87	5.21±0.03	44.8	68.4
						Argentina	0.40	0.26	6.32±0.09	41.5	54.1
						Chile	0.19	0.14	7.42±0.16	32.0	38.7
						Colombia	0.07	0.03	5.67±0.18	27.7	32.1
						Venezuela	0.06	0.03	6.09±0.34	34.9	42.0

World publications: density-equalizing map

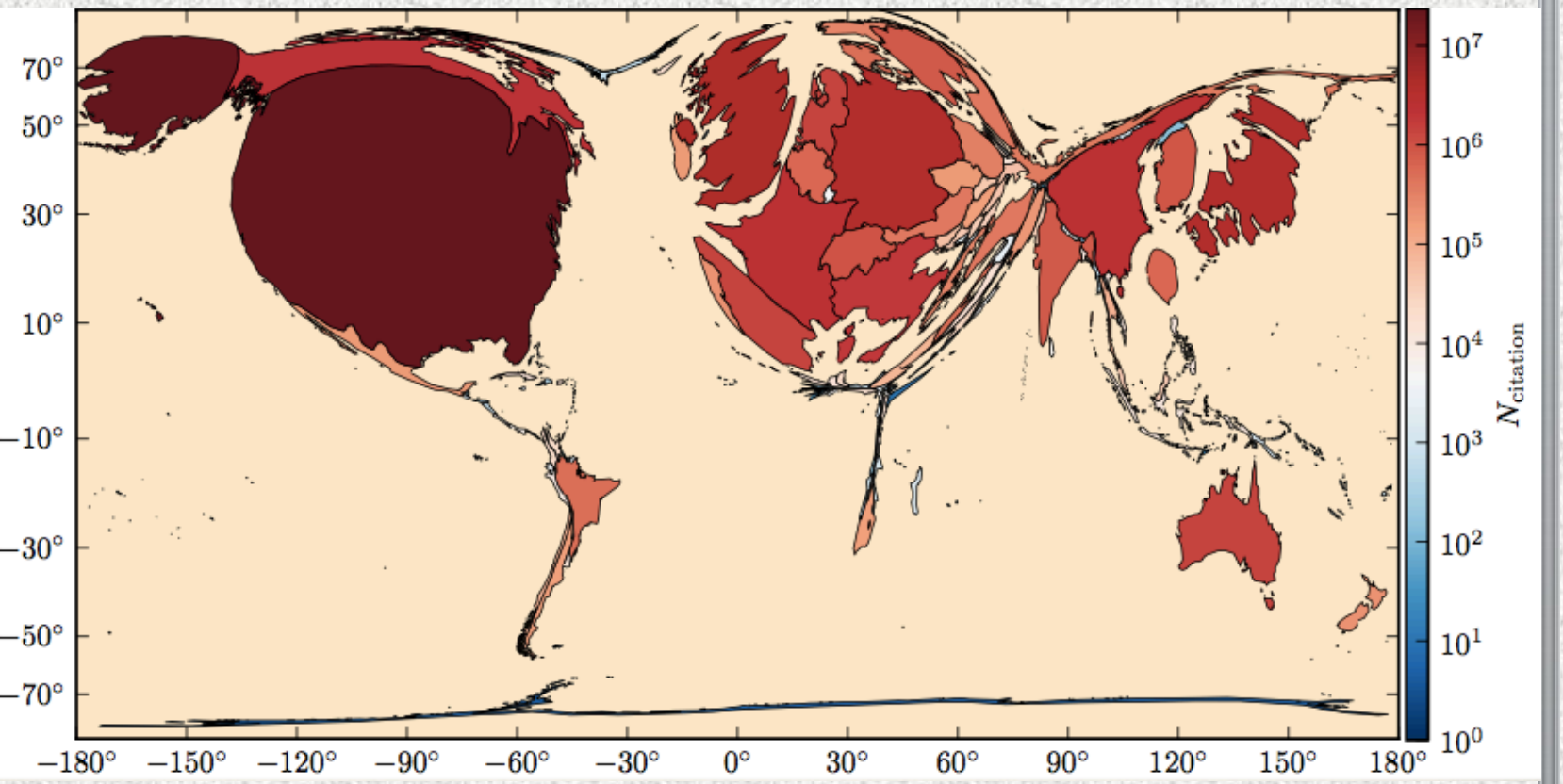


M. Gastner, M. E. J. Newman, Proc. Natl. Acad. Sci. U.S.A. 101, 7499 (2004)

World citations: map

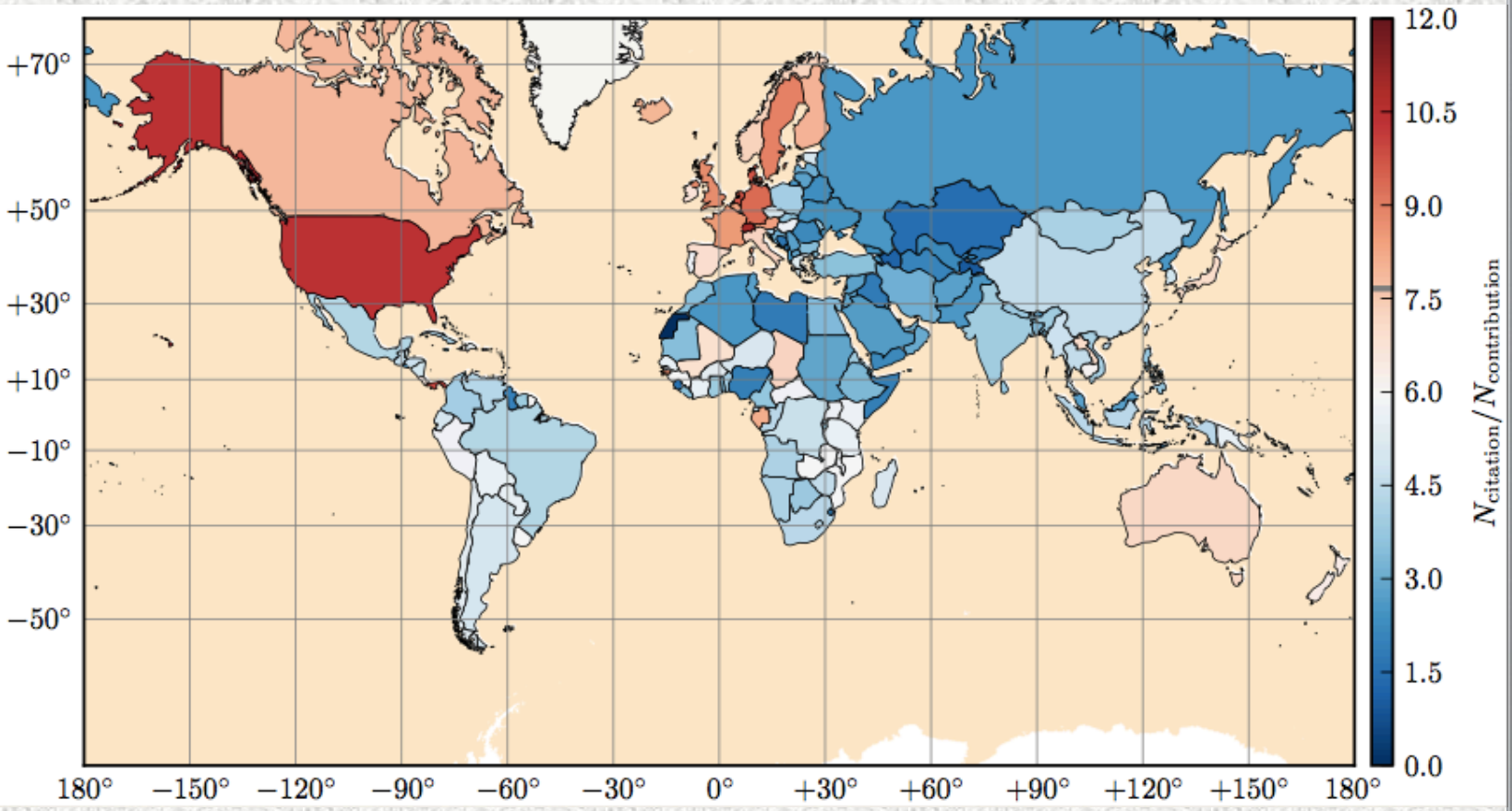


World citations: density-equalizing map

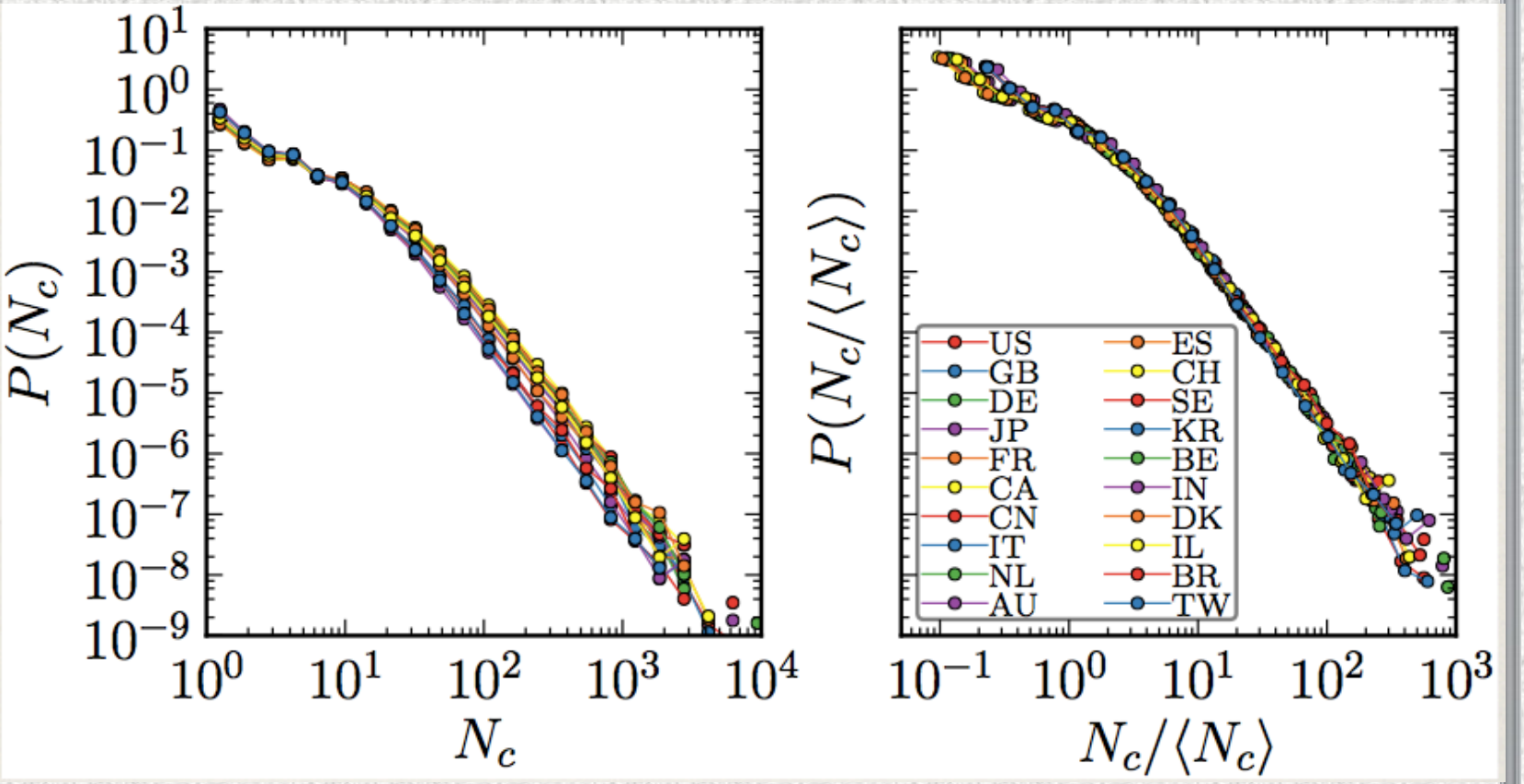


M. Gastner, M. E. J. Newman, Proc. Natl. Acad. Sci. U.S.A. 101, 7499 (2004)

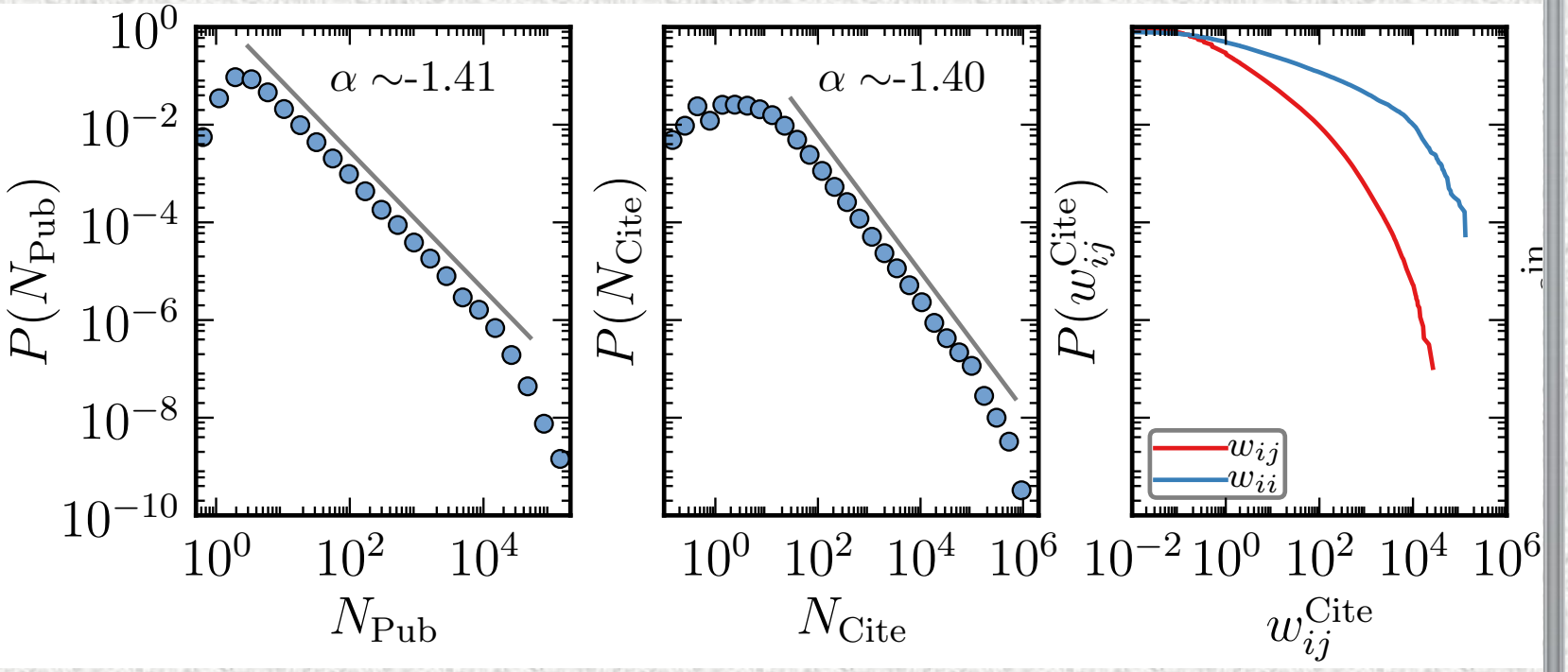
World citation averages



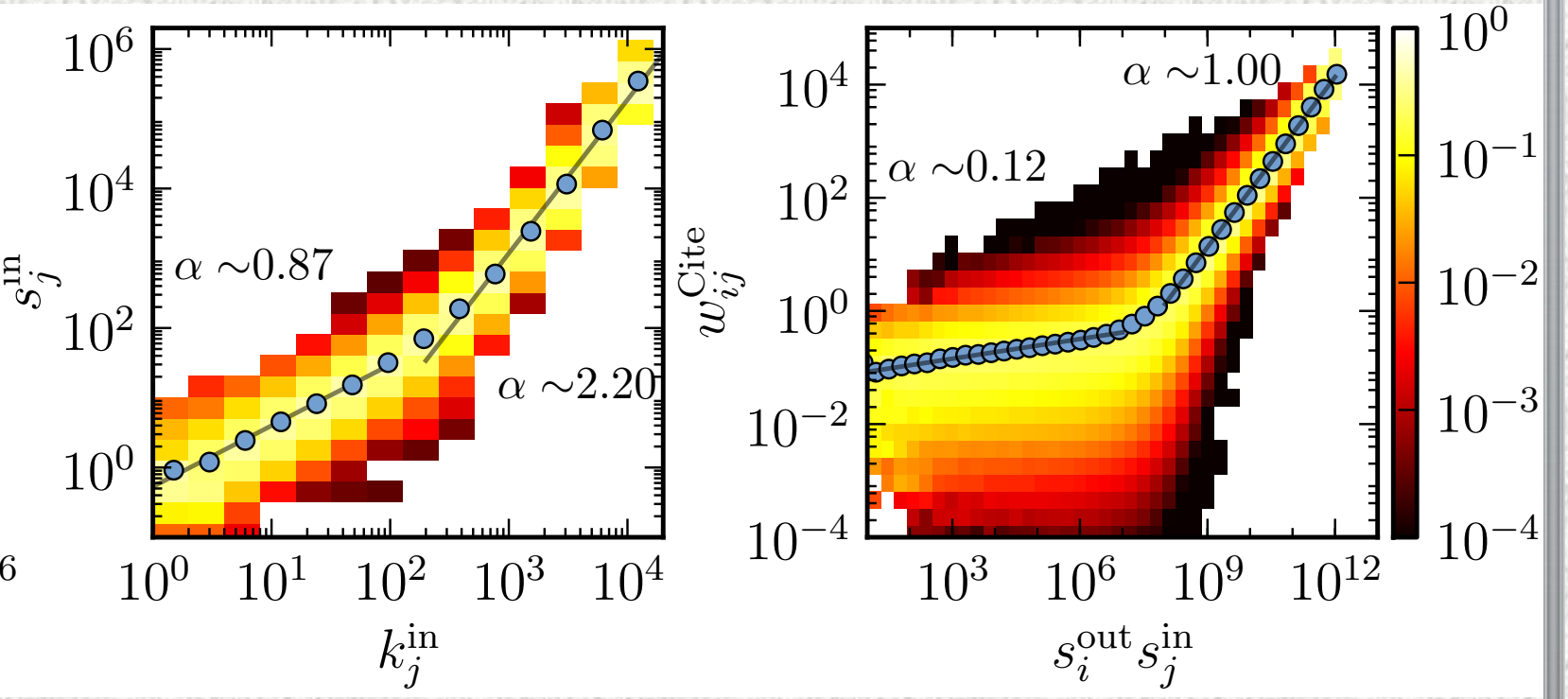
World citation distributions



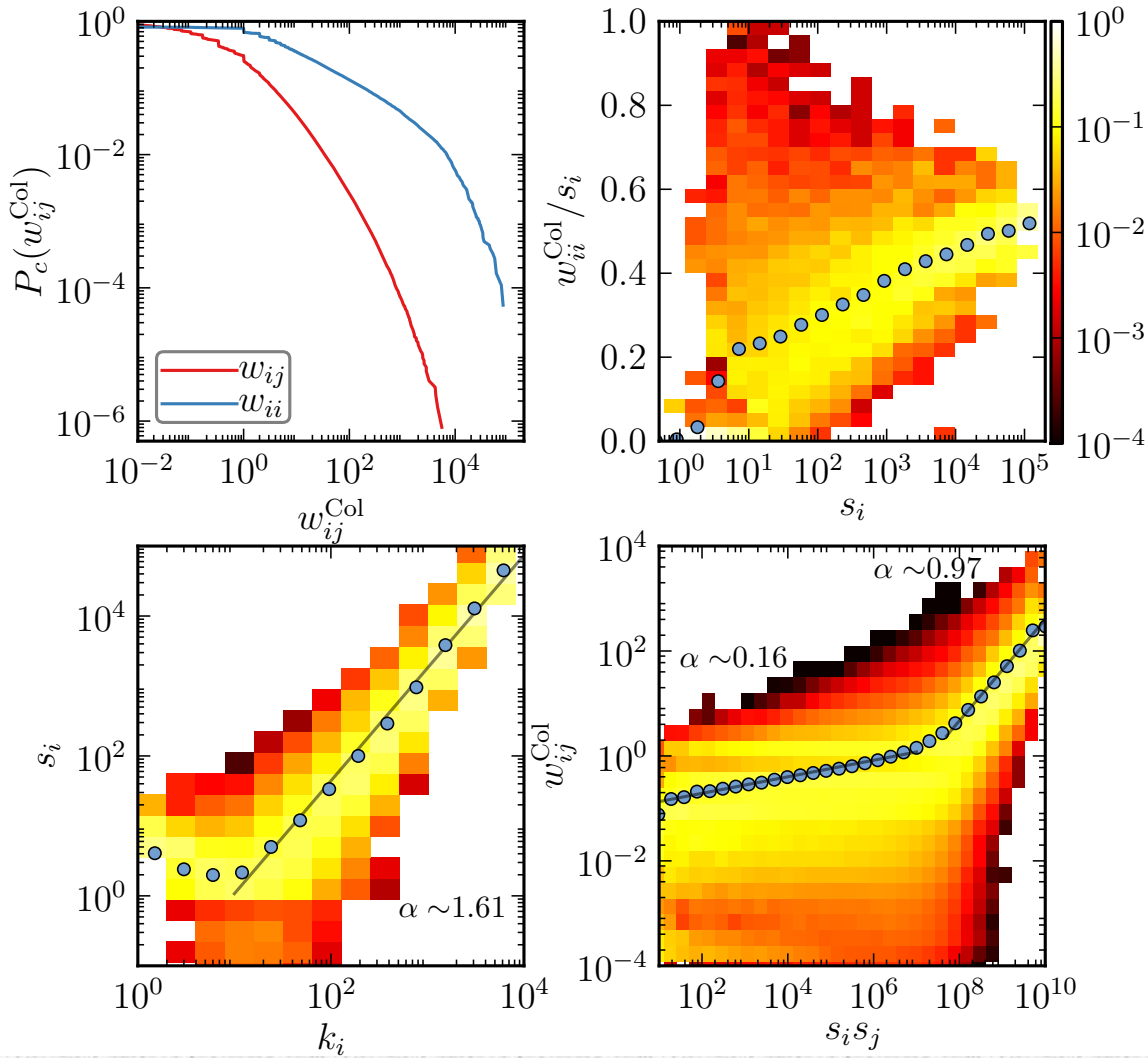
City-level citation distributions



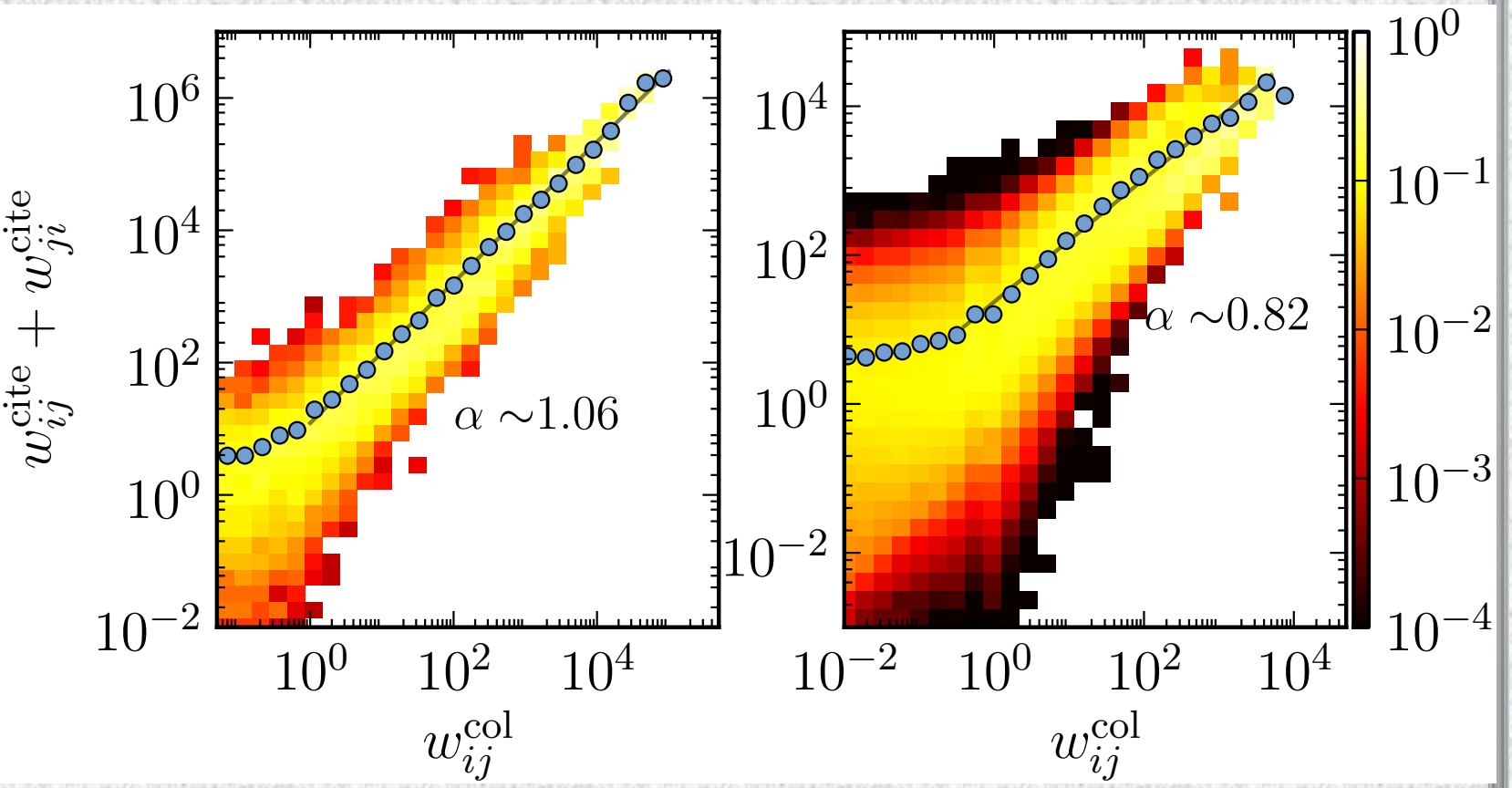
City citation statistics: correlations



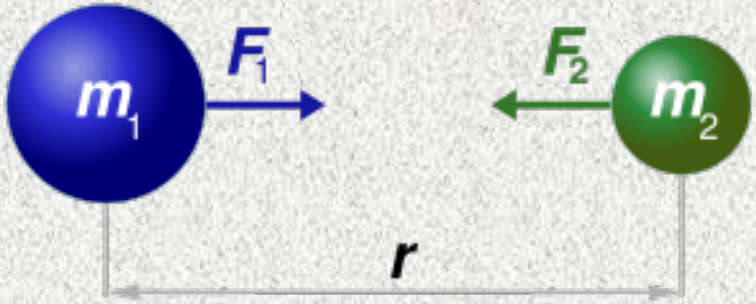
City-level collaborations: distributions and correlations



Citation vs collaborations



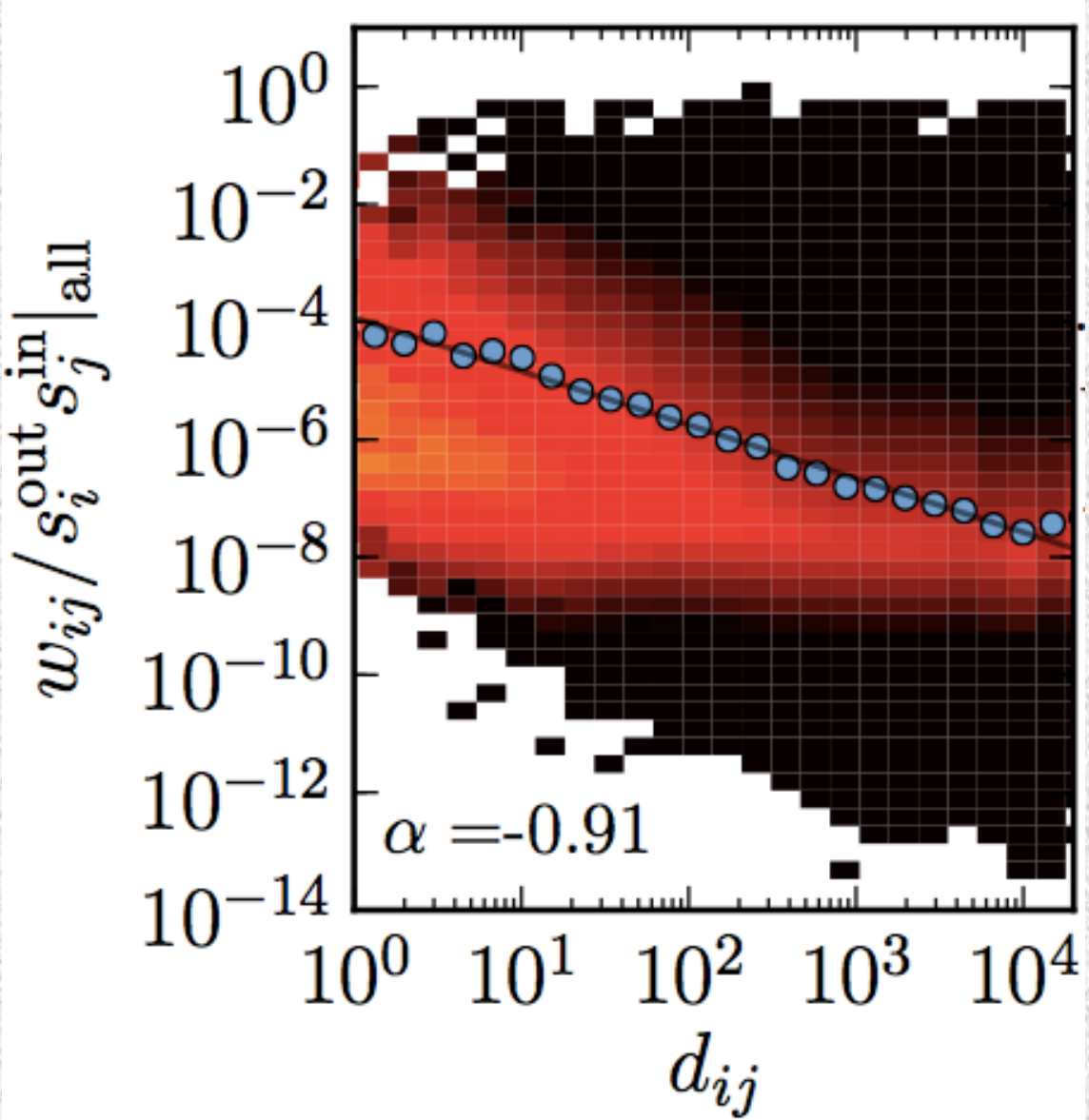
Gravity law



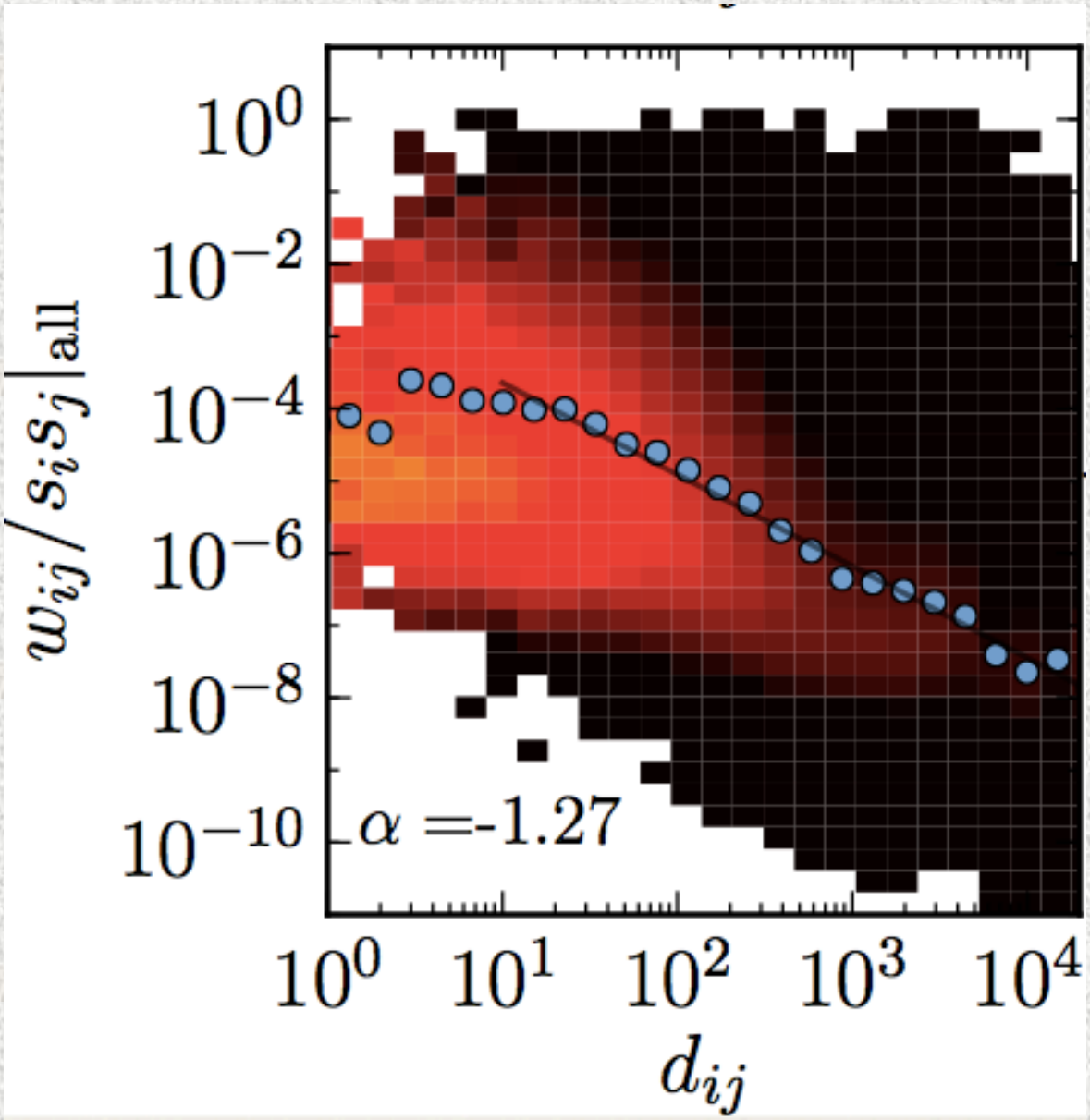
$$F_1 = F_2 = G\frac{m_1 \times m_2}{r^2}$$

$$w_{ij} \propto \frac{s_i^{out} s_j^{in}}{d_{ij}^{\alpha}}$$

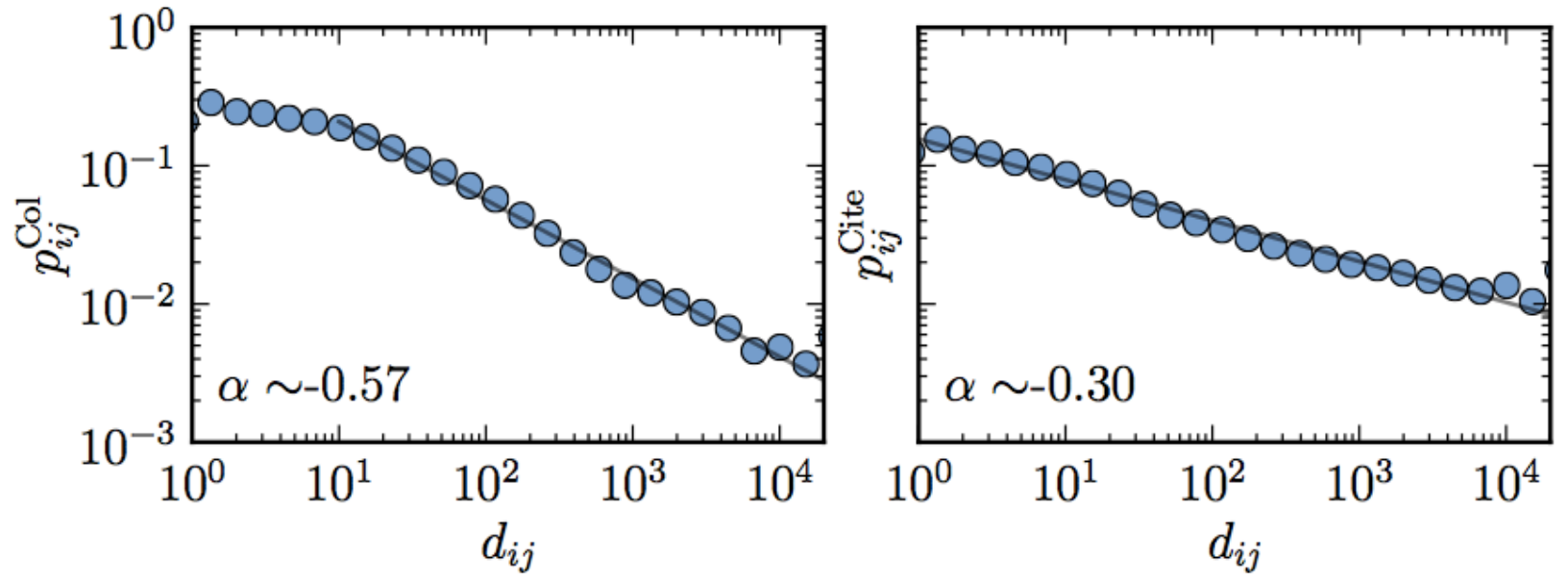
Gravity law: citations



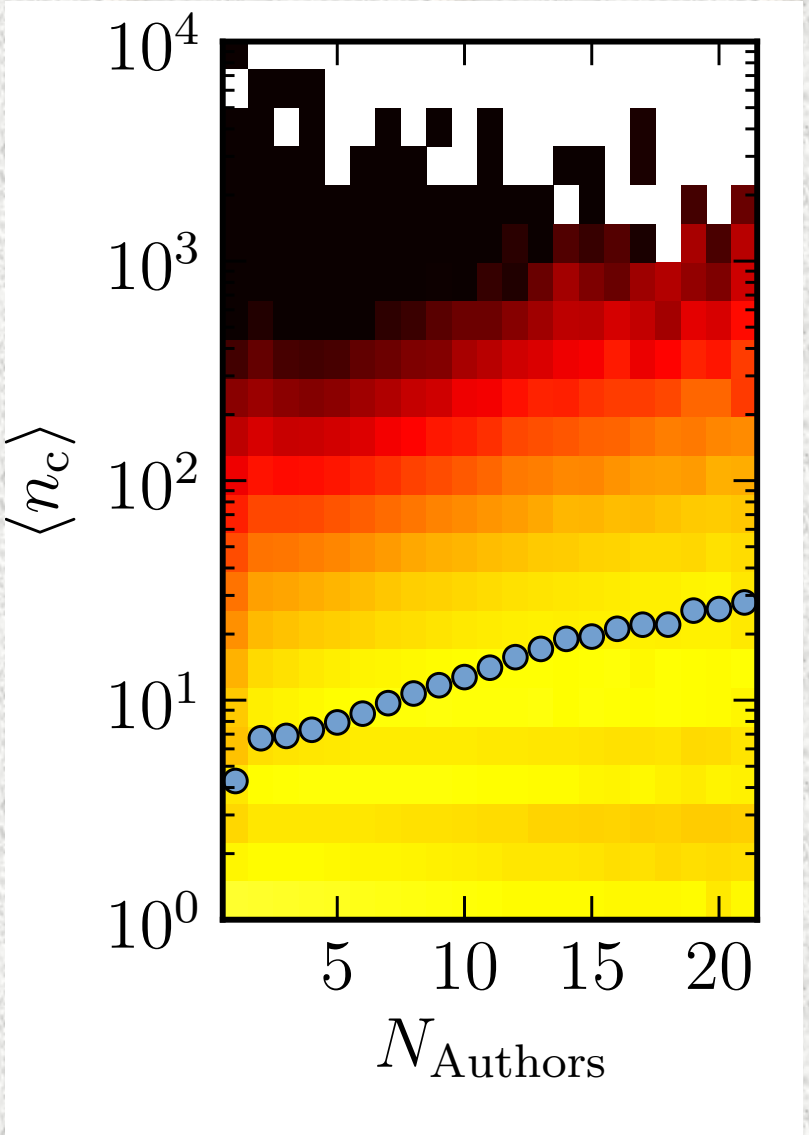
Gravity law: collaborations



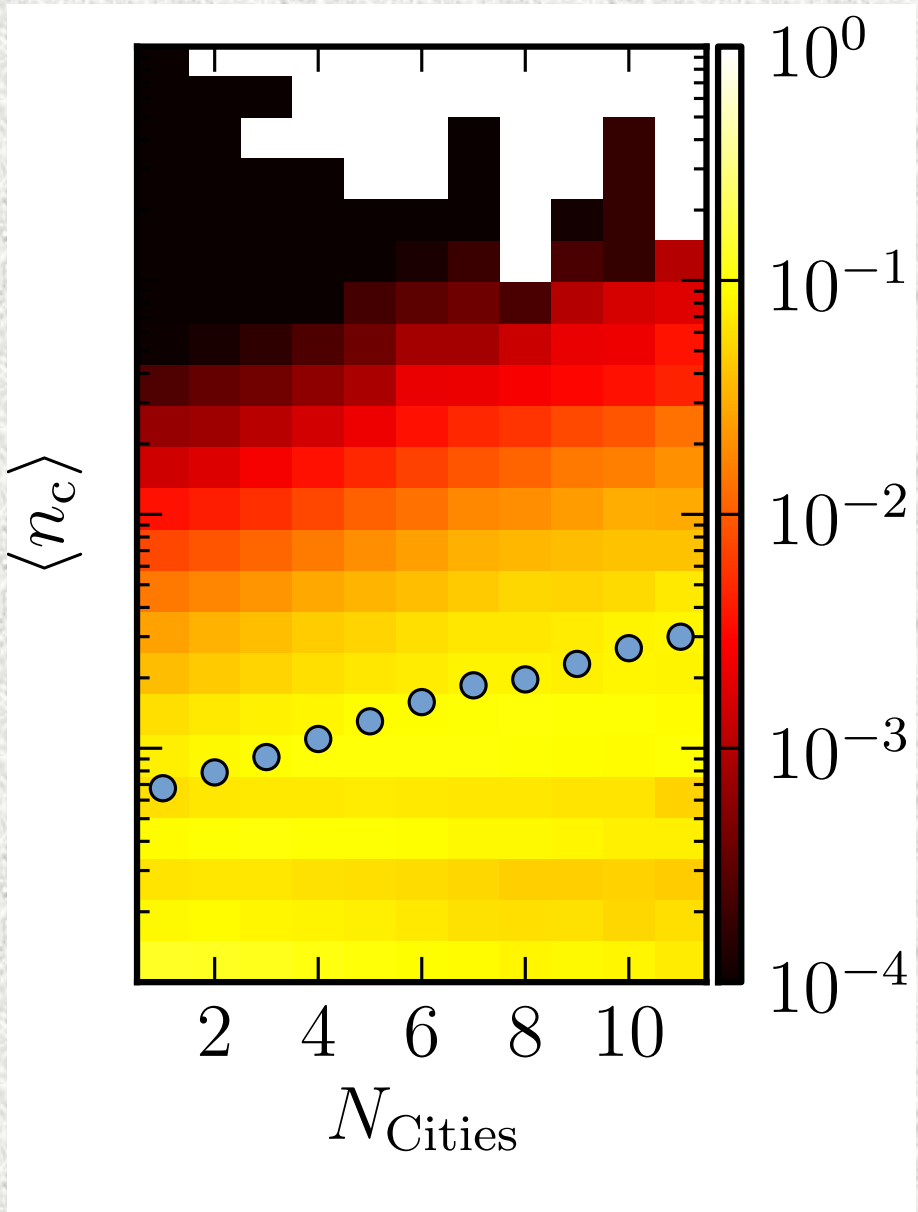
Gravity law: link probability vs distance



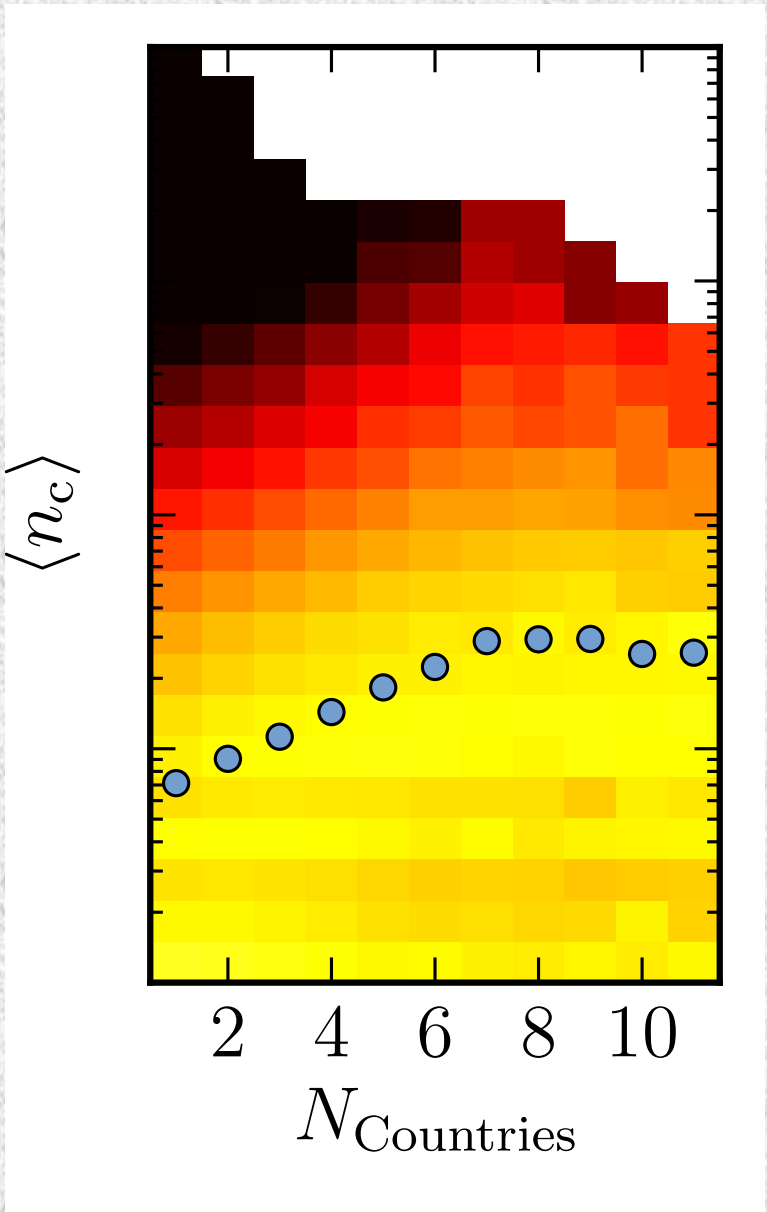
Cites vs number of authors (I)



Cites vs number of authors (II)



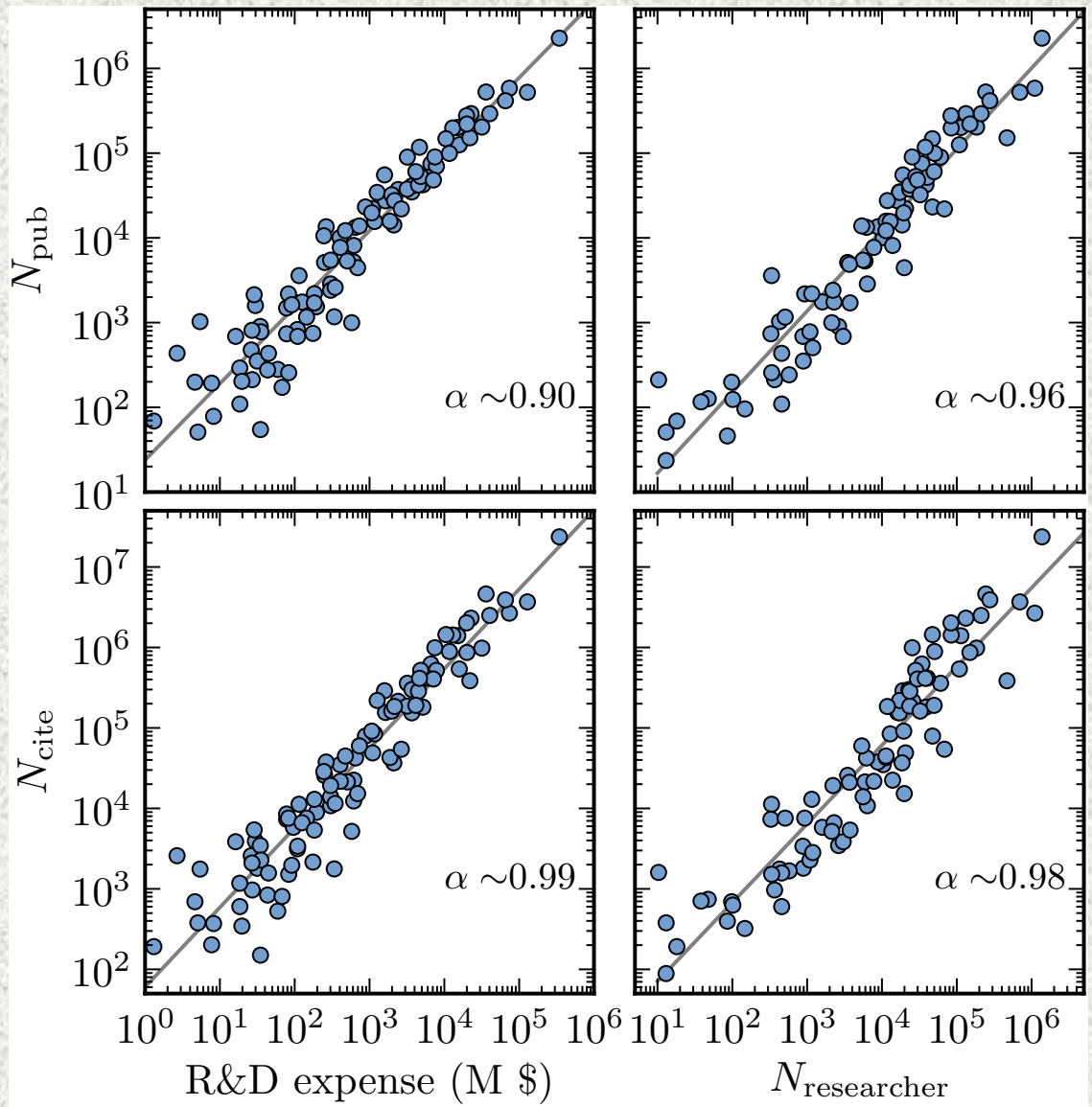
Cites vs number of authors (III)



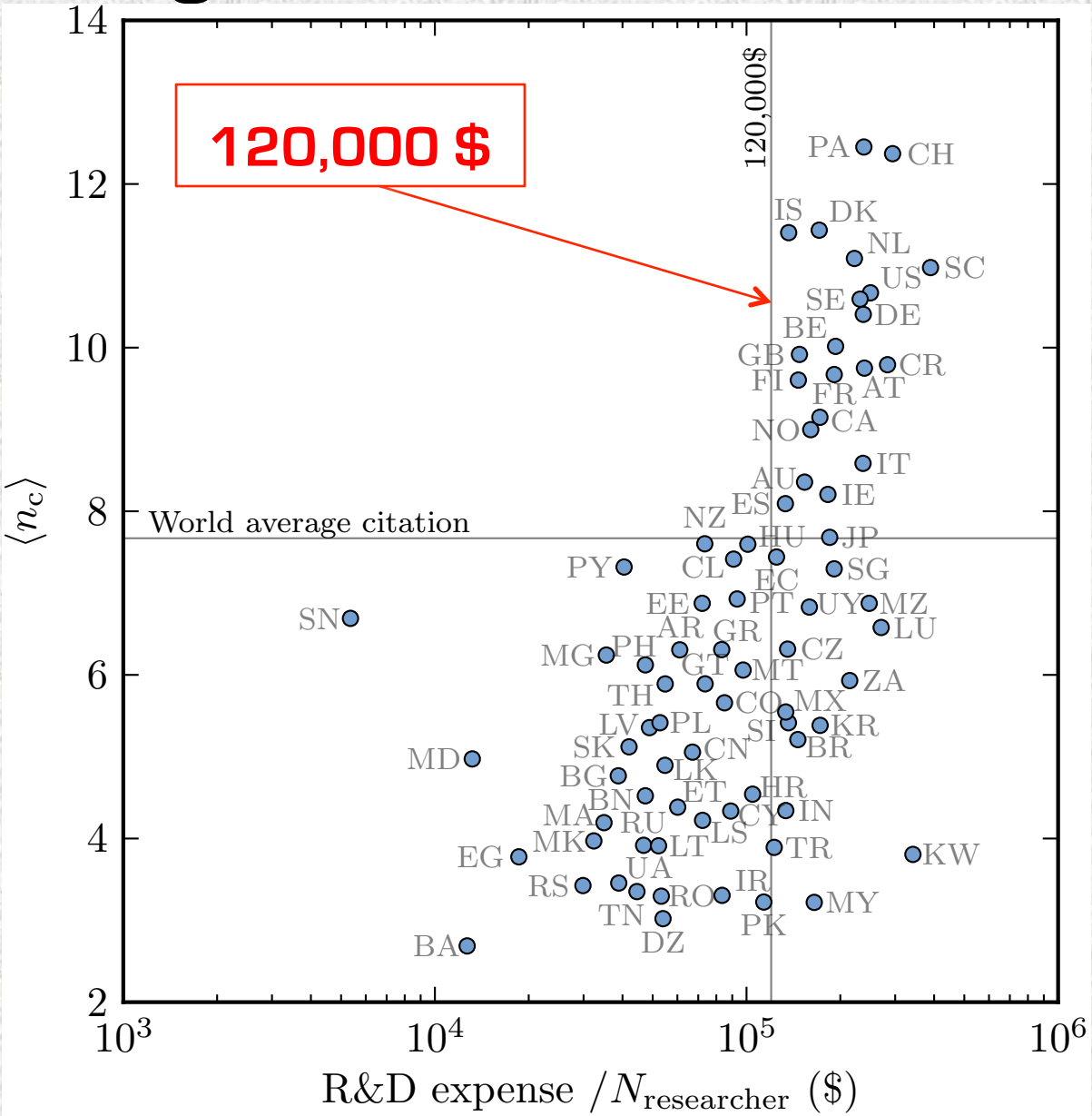
Cites vs number of authors: summary

N_{Authors}	f_{Papers} (in %)	Single City	Multiple City	Multiple Countries
1	13.03	4.25 ± 0.02	4.95 ± 0.12	5.24 ± 0.11
2	19.01	6.80 ± 0.02	6.11 ± 0.04	7.00 ± 0.05
3	18.34	6.92 ± 0.02	6.38 ± 0.03	7.30 ± 0.04
4	14.95	7.19 ± 0.02	7.02 ± 0.03	8.03 ± 0.04
5	11.10	7.62 ± 0.03	7.66 ± 0.03	8.79 ± 0.04
6	8.01	8.13 ± 0.04	8.52 ± 0.05	9.77 ± 0.05
7	5.20	8.85 ± 0.05	9.56 ± 0.07	10.90 ± 0.07
8	3.45	9.50 ± 0.07	10.67 ± 0.09	12.10 ± 0.10
9	2.22	10.23 ± 0.10	11.52 ± 0.12	13.17 ± 0.12
10	1.53	10.57 ± 0.12	12.45 ± 0.14	14.70 ± 0.15
>10	3.17	13.82 ± 0.17	16.64 ± 0.16	21.37 ± 0.17

Funding vs citations

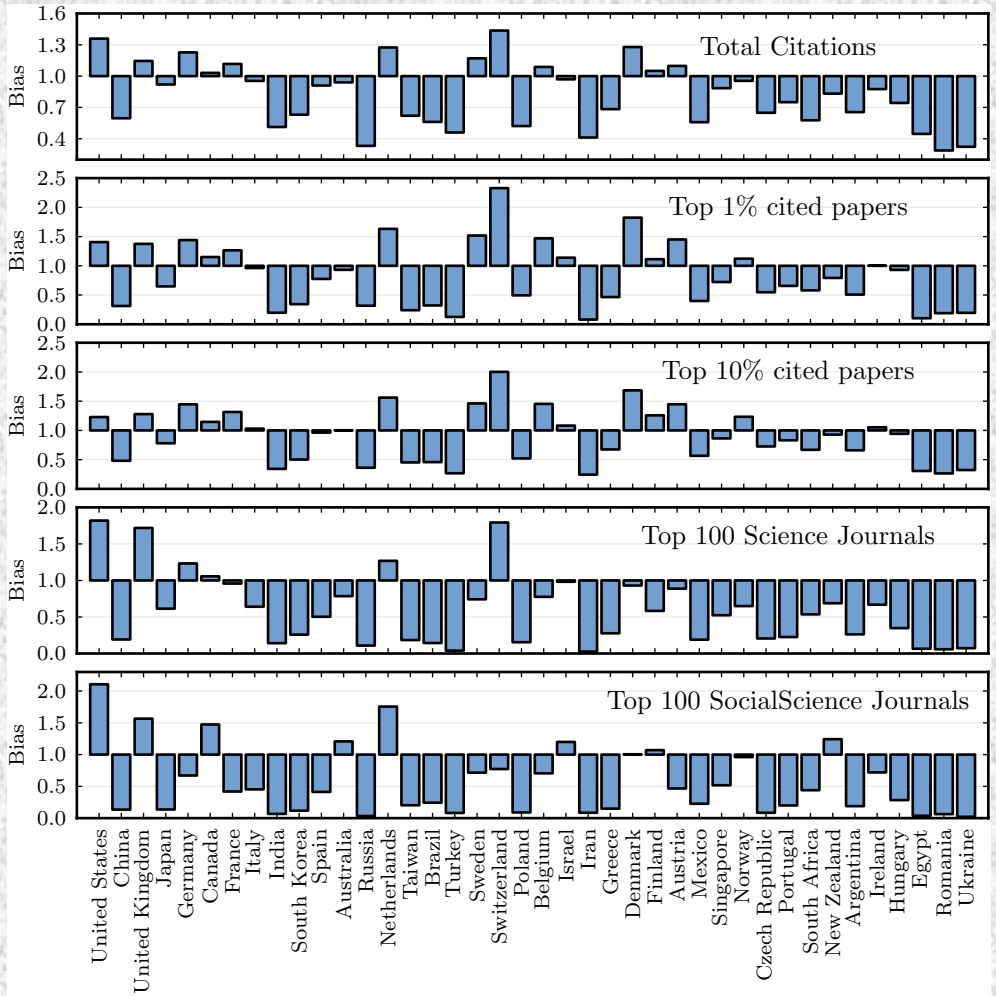


Funding vs citations



Funding vs citations

$$\text{bias for country i} = \frac{N_i^{set}}{\sum_i N_i^{set}} \times \frac{\sum_i N_i}{N_i}$$



Summary II

- Geography plays an important role in the dynamic of citation and collaboration patterns
- The strengths of citation flows and/or collaborations obey gravity laws, i.e. they display a power law decay with distance
- The number of citations of a paper increases with the number of authors, affiliations and countries
- There is a threshold effect in the relationship between research funding and citations: below 120,000 $ per researcher the average number of cites of papers of a country stays below the world average.

How to get more citations?

- Write papers with many people
- Get good neighbors
- Do international collaborations
- Get more funding!

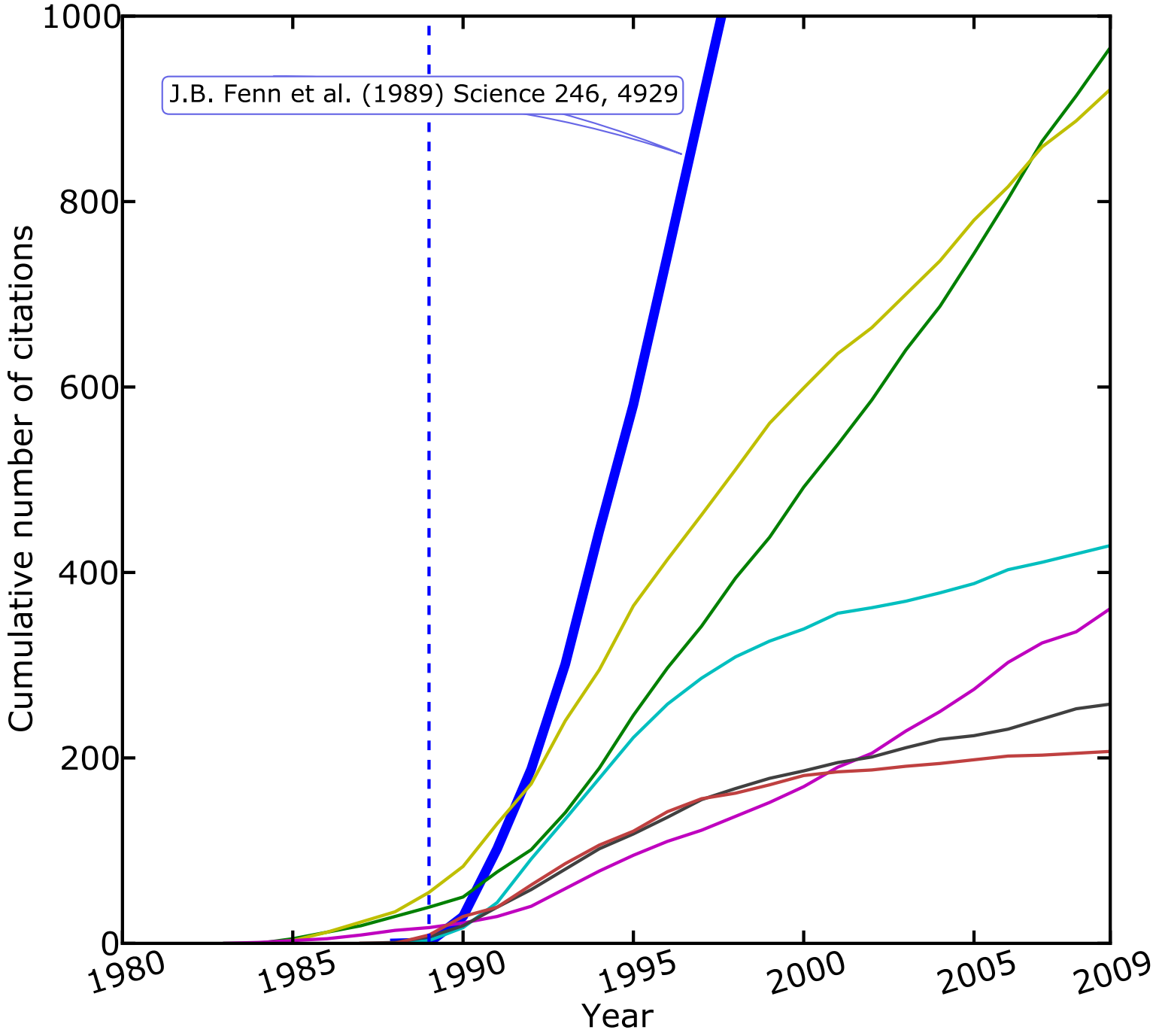
Nobel boosts ...

Goal: studying the occurrence and effects of groundbreaking papers on scientific careers

Focus: Nobel Prize Laureates

Data: ISI Web of Science citations of papers of 124 Nobel Prize Laureates in the last two decades (1990-2009)

A. Mazloumian, Y.-H. Eom, D. Helbing, S. Lozano, S. F., PLoS One 6(5), e18975 (2011)


J.B. Fenn et al. (1989) Science 246, 4929
Cumulative number of citations
Year
0
200
400
600
800
1000
1980
1985
1990
1995
2000
2005
2009

The boost factor

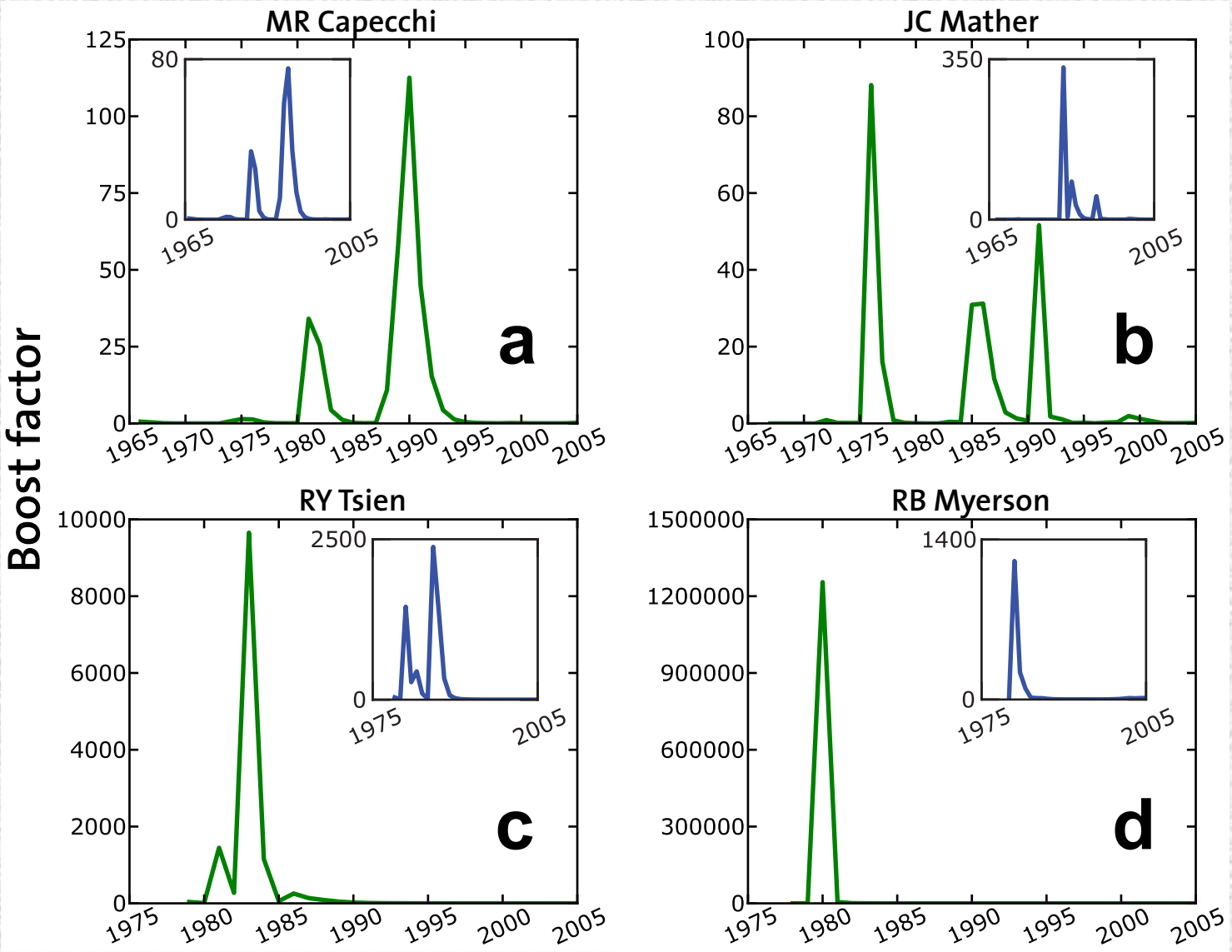
Principle: comparing the citation rates of articles before and after time t of papers published before t

$R_{<t,w} =$ Average number of citations per paper and year received in the period [t-w+1, t]

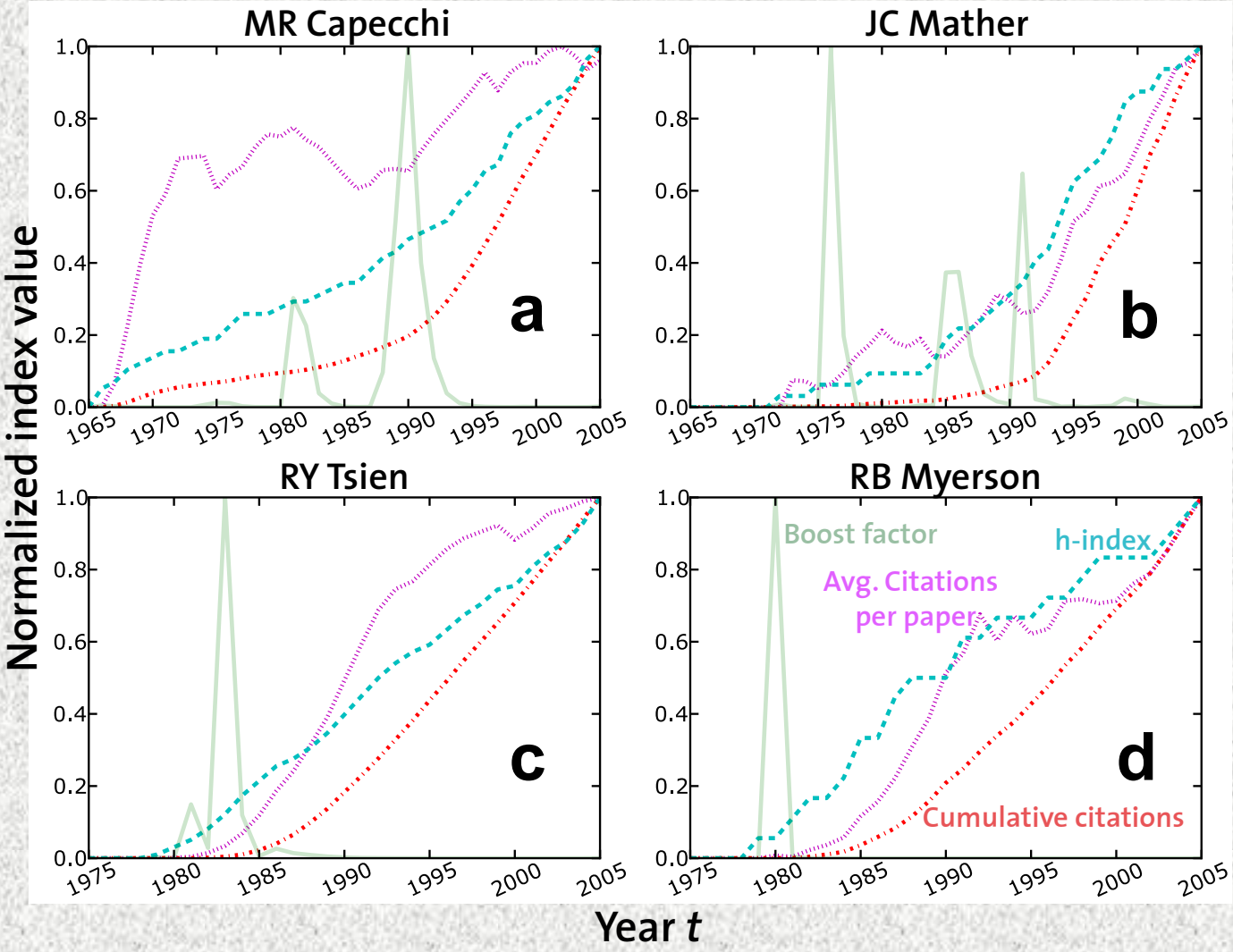
$R_{>t,w} =$ Average number of citations per paper and year received in the period [t+1, t+w]

$$R_w(t) = \frac{R_{>t,w}}{R_{<t,w}}$$

The boost factor



The boost factor vs standard indicators



Summary III

- Groundbreaking scientific papers have a boosting effect on previous publications of their authors, even if they are not topically related to them (*authority effect*)
- The boost factor is able to capture sudden variations of citation rates and to spot potential breakthrough early on
- Peaks in the evolution of the boost factor are not due to the landmark papers themselves but to the citation cascade towards earlier articles
- The boost factor is more effective than traditional scientific metrics

Acknowledgements

Filippo Radicchi

Dirk Helbing

Claudio Castellano

Raj Kumar Pan

Sergi Lozano

Young-Ho Eom

Amin Mazloumian

REVIEWS OF MODERN PHYSICS, VOLUME 81, APRIL–JUNE 2009

Statistical physics of social dynamics

Claudio Castellano*

SMC, INFM-CNR and Dipartimento di Fisica, "Sapienza" Università di Roma, Piazzale A. Moro 2, 00185 Roma, Italy

Santo Fortunato†

Complex Networks Lagrange Laboratory, ISI Foundation, Viale S. Severo 65, 10133 Torino, Italy

Vittorio Loreto‡

Dipartimento di Fisica, "Sapienza" Università di Roma and SMC, INFM-CNR, Piazzale A. Moro 2, 00185 Roma, Italy and Complex Networks Lagrange Laboratory, ISI Foundation, Viale S. Severo 65, 10133 Torino, Italy

(Published 11 May 2009)

Statistical physics has proven to be a fruitful framework to describe phenomena outside the realm of traditional physics. Recent years have witnessed an attempt by physicists to study collective phenomena emerging from the interactions of individuals as elementary units in social structures. A wide list of topics are reviewed ranging from opinion and cultural and language dynamics to crowd behavior, hierarchy formation, human dynamics, and social spreading. The connections between these problems and other, more traditional, topics of statistical physics are highlighted. Comparison of model results with empirical data from social systems are also emphasized.

DOI: 10.1103/RevModPhys.81.591 PACS number(s): 05.10.−a, 89.20.−a, 89.75.−k

CONTENTS

*Electronic address: claudio.castellano@roma1.infn.it
†Electronic address: fortunato@isi.it
‡Electronic address: vittorio.loreto@roma1.infn.it

C. Castellano, S. F., V. Loreto, Rev. Mod. Phys. 81, 591 (2009)

SCIENCE WATCH

HOME ABOUT THOMSON REUTERS PRESS RO

ScienceWatch.com Data & Rankings New Hot Papers

NEW HOT PAPERS

Spectral properties of complex networks
ECT* Trento, July 23-27, 2012