Tsallis distribution in hadronization and in citation

Physica A474(2017)355; Phys.Rev.E95(2017)032130; PLOSONE 12(2017)0179656; arXiv: 1707.07912, 1711.02364, 1711.11331

T.S. Biró[1] Z. Néda[2] A. Telcs[1]

[1]MTA Wigner Research Centre for Physics, Budapest, Hungary

[2]Hungarian Institute for Physics UBB Cluj, Romania

Version: December 2, 2017

Talk by T.S.Biró at ZImanyi School, Budapest, Dec 4-8, 2017.

Outline

Problems, projects, propaganda

- Unidirectional Growth and Resetting Model
- Negative binomial multiplicity distribution
- Single particle energy distribution
- Citation popularity distribution

Master Equation

general form and properties

$$\dot{P}_n = \sum_m w_{nm} P_m \tag{1}$$

would not conserve normalization. Therefore

$$\dot{P}_n = \sum_m \left[w_{nm} P_m - w_{mn} P_n \right]. \tag{2}$$

Stationary distribution is subject to a **total balance**:

$$0 = \sum_m \left[w_{nm} Q_m - w_{mn} Q_n \right]. \tag{3}$$

This is W equations for W unknowns.

detailed balance constraints the transition rates:

$$\frac{w_{nm}}{w_{mn}} = \frac{Q_n}{Q_m}, \tag{4}$$

W constraints for W^2 unknowns.

Stability

entropic distance

$$\lim_{t\to\infty} P_n(t) = Q_n. \tag{5}$$

Exists Lyapunov functional: **Kullback–Leibler divergence**

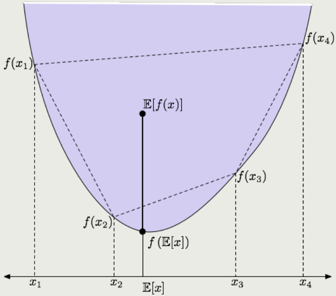
$$\rho[P,Q] \equiv \sum_n Q_n \ln \frac{Q_n}{P_n}. \tag{6}$$

For all $w_{nm} > 0$ transition rates we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho[P,Q] \le 0. \tag{7}$$

Jensen inequality

for $f'' > 0$



$$\sum_{n=0}^{W} p_n f(x_n) \geq f\left(\sum_{n=0}^{W} p_n x_n\right) \tag{8}$$

Jensen inequality

special case

For $p_1 = 1/2, p_2 = 1/2$ and $f(x) = -\ln x$ with $f''(x) = 1/x^2 > 0$:

$$-\frac{1}{2}(\ln a + \ln b) \geq -\ln\left(\frac{a+b}{2}\right) \tag{9}$$

$$-\ln\sqrt{ab} \geq -\ln\left(\frac{a+b}{2}\right) \tag{10}$$

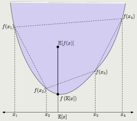
$$\ln\sqrt{ab} \leq \ln\left(\frac{a+b}{2}\right) \tag{11}$$

$$\sqrt{ab} \leq \frac{a+b}{2} \tag{12}$$

geometric mean $\leq$ aritmetic mean

Convexity

Jensen inequality for $f'' > 0$



$$\sum_{n=0}^{W} p_n f(x_n) \geq f\left(\sum_{n=0}^{W} p_n x_n\right) \tag{13}$$

Apply to $f(x) = -\ln x$, $x_n = P_n/Q_n$:

$$\rho[P,Q] \geq -\ln\left(\sum_n Q_n \frac{P_n}{Q_n}\right) = 0. \tag{14}$$

Conclusion with uniform:

$$\rho[U,Q] = \ln W + \sum_{n=0}^{W} Q_n \ln Q_n \geq 0. \tag{15}$$

Particular Master Equation

Unidirectional Growth + Reset

Dynamics of probabilities (fractions)

$w_{nm} = \mu_m \delta_{n-1,m} + \gamma_m \delta_{n,0}$

$\dot{P}_0 = \langle \gamma \rangle - (\gamma_0 + \mu_0) P_0$ and ☺

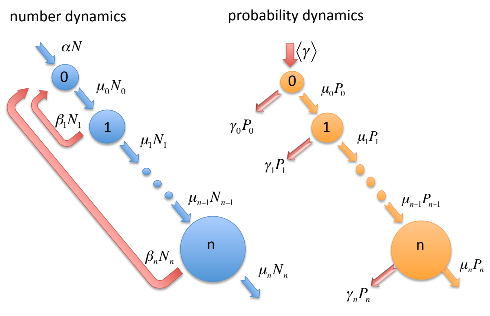
$$\dot{P}_n = \mu_{n-1} P_{n-1} - (\mu_n + \gamma_n)\ P_n \quad (16)$$

Dynamics of numbers

$N(t) = \sum N_n(t)$

$\dot{N} = -\gamma N$ and ☺

$$\dot{N}_n = \mu_{n-1} N_{n-1} - \mu_n N_n \quad (17)$$



Short step-up + long hops to zero:
stationary distribution

Stationary limit: $P_n(t) \to Q_n$, from $\dot{Q}_n = 0$ one obtains

$Q_0 = \langle\gamma\rangle_Q / (\gamma_0 + \mu_0)$ and stationary ☺

$$Q_n = \frac{\mu_{n-1}}{\mu_n + \gamma_n} Q_{n-1} = \cdots = \frac{\mu_0 Q_0}{\mu_n} \prod_{j=1}^{n} \left(1 + \frac{\gamma_j}{\mu_j}\right)^{-1}. \quad (18)$$

Constant rates

$\rightarrow$ exponential distribution

Assume $\mu_j = \sigma$, attachment rate independent of number of links.

$$Q_n = Q_0 \prod_{j=1}^{n} \frac{\sigma}{\sigma + \gamma} = Q_0 \left(1 + \gamma/\sigma\right)^{-n}. \tag{19}$$

Geometrical sum for normalization. We obtain

Boltzmann–Gibbs exponential ☺

$$\boldsymbol{Q_n = \frac{1}{1 + \sigma/\gamma} \, e^{-n \cdot \ln(1+\gamma/\sigma)}}. \tag{20}$$

Linear preference, constant loss rate

$\rightarrow$ Waring distribution

Linear preference in attachment: $\mu_j = \sigma(j+b) \quad (b>0)$.

$$Q_n \;=\; Q_0 \prod_{j=1}^{n} \frac{j-1+b}{j+b+\gamma/\sigma} \;=\; Q_0\, \frac{(b)_n}{(c)_n}. \tag{21}$$

with $c = b + 1 + \gamma/\sigma$. Norm:
$\sum_n Q_n = Q_0\,(c-1)/(c-1-b) = 1$.

Pochhammer ratio (Waring) ☺

$$\boldsymbol{Q_n \;=\; \frac{c-1-b}{c-1}\,\frac{(b)_n}{(c)_n}} \tag{22}$$

Matthias principle: tail of Waring

$\rightarrow$ power-law!

The above result in the $n \to \infty$ limit:
Since

$$\lim_{n\to\infty} n^{c-b} \frac{\Gamma(n+b)}{\Gamma(n+c)} = 1, \tag{23}$$

we obtain

Pochhammer in $n \to \infty$ limit: power-law! ☺

$$Q_n \to \frac{\gamma}{\gamma + b\sigma} \frac{\Gamma(c)}{\Gamma(b)} n^{-1-\gamma/\sigma}. \tag{24}$$

Avalanche dynamics in the **large n limit!**

continuous variable: $x = n \cdot \Delta x$

- $P_n(t) = \Delta x \cdot P(n \cdot \Delta x, t)$ ensures $\sum\limits_{n=0}^{\infty} P_n(t) = \int\limits_{0}^{\infty} P(x,t) dx$.
- $\mu_n = \frac{1}{\Delta x} \cdot \mu(n \cdot \Delta x)$ and $\gamma_n = \gamma(n \cdot \Delta x)$ lead to

Continuum Master: ☺

$$\frac{\partial}{\partial t} P(x,t) = -\frac{\partial}{\partial x} \left(\mu(x)\, P(x,t) \right) - \gamma(x) P(x,t). \tag{25}$$

with the stationary distribution

$$Q(x) = \frac{K}{\mu(x)} \, e^{-\int\limits_0^x \frac{\gamma(u)}{\mu(u)} \, du}. \tag{26}$$

Particular continuous stationary distributions

with constant $\gamma(x) = \gamma$.

For constant rate $\mu(x) = \sigma$ **exponential**:

$$Q(x) = \frac{\gamma}{\sigma} \mathrm{e}^{-\frac{\gamma}{\sigma} x}. \tag{27}$$

For linear preference $\mu(x) = \sigma(x+b)$ **Tsallis–Pareto**:

$$Q(x) = \frac{\gamma}{\sigma\, b} \left(1 + \frac{x}{b}\right)^{-1-\gamma/\sigma}. \tag{28}$$

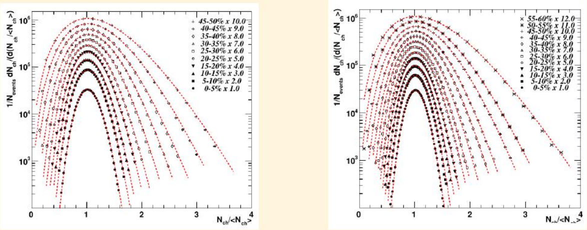
For exponential dispreference $\mu(x) = \sigma \mathrm{e}^{-ax}$ **Gompertz**

$$Q(x) = \frac{\gamma}{\sigma} \mathrm{e}^{ax + \frac{\gamma}{a\sigma}(1-\mathrm{e}^{ax})}. \tag{29}$$

Hadronization

From QGP to n hadrons: NBD PHENIX, PRC 78 (2008) 044902

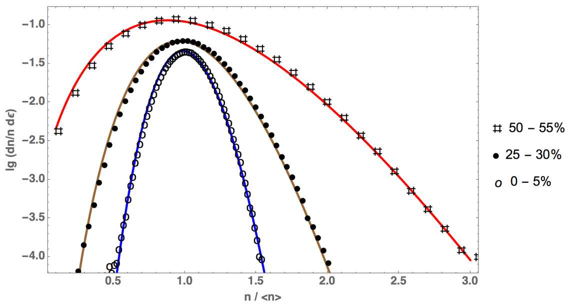
Au + Au collisons at $\sqrt{s_{NN}} = 62$ (left) and 200 GeV (right). Total charged multiplicities.



$$\gamma_n = \sigma(n-kf),\ \mu_n = \sigma f(n+k); \qquad Q_n = \binom{n+k-1}{n} f^n (1+f)^{-n-k}.$$

Hadronization

Our fit to public PHENIX DATA



(a) $0 - 5\%: \quad \langle n \rangle = 61.0, k = 255, f = 0.24$

(b) $25 - 30\%: \quad \langle n \rangle = 27.4, k = 50, f = 0.55$

(c) $50 - 55\%: \quad \langle n \rangle = 8.5, k = 17, f = 0.50$

Energy sharing

pick up one particle from n

Kinetic models:

1. 1-dim gas, extreme relativistic (jet):

$$E = c \sum_{j=1}^{n} |p_j| . \tag{30}$$

2. 2-dim gas, non-relativistic (mid-rapidity hadrons):

$$E = \frac{1}{2m} \sum_{j=1}^{2n} p_j^2 . \tag{31}$$

These total energies define an N-ball ($N = n$ and $N = 2n$) in the L_p-norm ($p = 1$ and $p = 2$) in the phase space.

Phase space volume

and energy shell

$$\Omega_N(E) = \int \delta(E - E')d\Gamma_N = \frac{1}{dE/dR} \int \delta(R - R')\, d\Gamma_N. \tag{32}$$

Fixed radius defines a surface; it is the derivative of volume against the radius:

$$\Omega_N(E) = \frac{1}{dE/dR} \frac{\mathrm{d}}{\mathrm{d}R} V_N(R) = \frac{\mathrm{d}}{\mathrm{d}E} V_N(R(E)). \tag{33}$$

The volume of an N-ball with radius R in L_p-norm is given as

$$V_N^{(p)} \left[\left(\sum_j |x_j|^p \right)^{1/p} \leq R(E) \right] = \frac{\Gamma(1+1/p)^N}{\Gamma(1+N/p)} (2R)^N. \tag{34}$$

Energy shell sizes

in $1/h$ units in kinetic phase space

Jets: $p = 1$, $N = n$ and $R(E) = E/c$. We obtain

$$\Omega_n^{(1)}(E) = \frac{\mathrm{d}}{\mathrm{d}E} \frac{(2E/c)^n}{n!} = \left(\frac{2}{c}\right)^n \frac{E^{n-1}}{(n-1)!}. \tag{35}$$

Massive hadrons in 2-dim: $p = 2$, $N = 2n$ and $R(E) = \sqrt{2mE}$.

$$\Omega_{2n}^{(2)}(E) = \frac{\mathrm{d}}{\mathrm{d}E} \frac{(2m\pi E)^n}{n!} = (2m\pi)^n \frac{E^{n-1}}{(n-1)!}. \tag{36}$$

Single particle energy spectra

reflect hypershell volume ratio

$g = N/n$ single particle kinetic degrees of freedom.

$$r_{N,g}^{(p)}(\epsilon, E) \equiv \frac{\Omega_g^{(p)}(\epsilon)\,\Omega_{N-g}^{(p)}(E-\epsilon)}{\Omega_N^{(p)}(E)}. \tag{37}$$

This ratio **coincides** for jets and hadrons $\Rightarrow$ a single fit for low and high p_T is possible:

$$r_{n,1}^{(1)} = r_{2n,2}^{(2)} = \frac{n-1}{E}\left(1-\frac{\epsilon}{E}\right)^{n-2}. \tag{38}$$

Note: $\int_0^E r(\epsilon)d\epsilon = 1.$

phsp ratio averaged over NBD

$n \geq 2$

Suppose that the newly made hadrons, those over 2, are NB distributed.

$$\frac{1}{\mathcal{N}} \frac{\mathrm{d}\mathcal{N}}{\mathrm{d}\epsilon} = \sum_{n=2}^{\infty} r_{ng,g}^{(p)}(\epsilon, E)\, P_{n-2}. \tag{39}$$

The result is

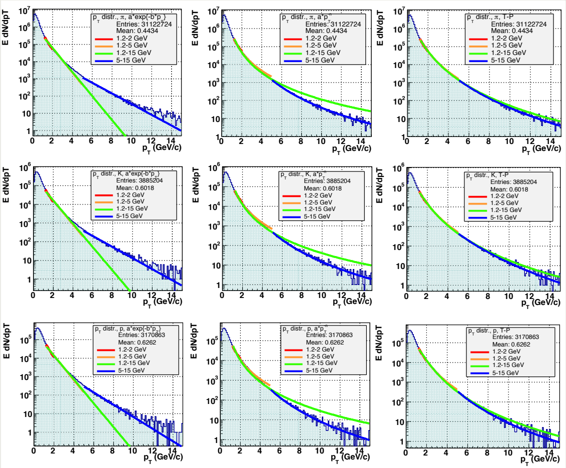
$$\frac{1}{\mathcal{N}} \frac{\mathrm{d}\mathcal{N}}{\mathrm{d}\epsilon} = \frac{1}{E} \left(1 + \frac{\overline{n}}{k} \frac{\epsilon}{E}\right)^{-k-1} \left[1 + \frac{\overline{n}}{k} \frac{\epsilon}{E} + \overline{n} \left(1 - \frac{\epsilon}{E}\right)\right]. \tag{40}$$

For $\epsilon \ll E$ Tsallis–Pareto distribution. For $k \to \infty$ exponential.

Hadronization

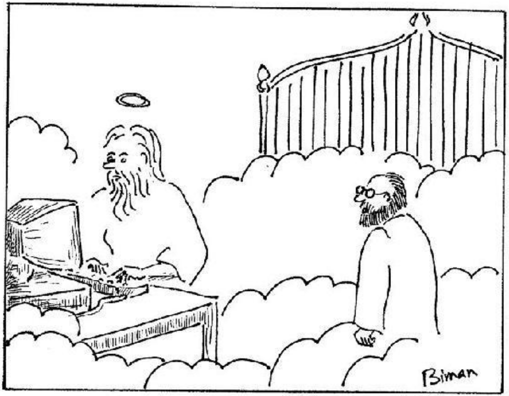
Best fit Tsallis–Pareto

MSc thesis Gábor Bíró 2016



Citations

importance



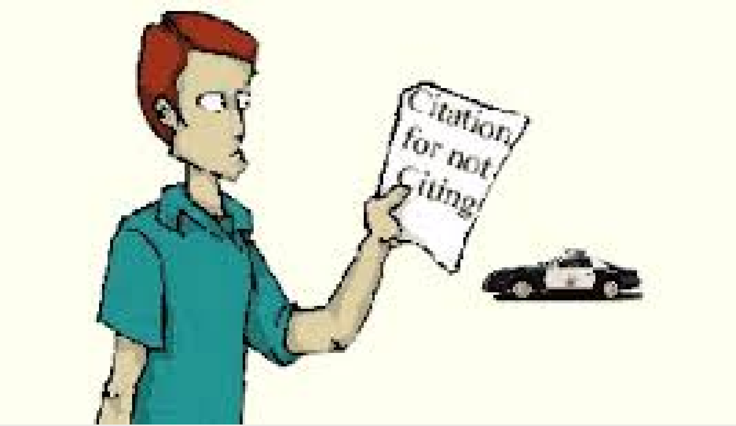
"I'll have to check your citation index first"

Citations

a friend in need is a friend indeed

Citations

consequence of forgetting



Citations

Total number and fraction dynamics

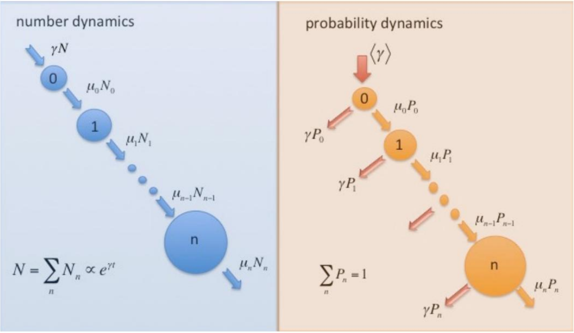


Figure 3. Schematic representation of the coarse-grained random growth model considered in the model. The panel on the left side indicates the growth process in the number of elements with n quanta: N_n. Due to the fact that the total number of elements is exponentially increasing, the probability P_n that an element will have n quanta, experiences the dynamics sketched on the right

Citations

Exponential growth

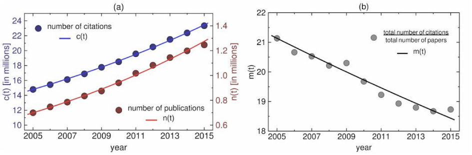


Figure 2. Results for the MEDLINE/PubMed database. **Figure 2a** illustrates the time evolution of the yearly indexed papers, $n(t)$, and the total number of citations, $c(t)$, introduced by them for each year in the 2005-2015 time interval. The trend $n(t)$ can be nicely fitted (red curve) with an exponential curve with $\gamma = 0.06$ using $t_0 = 2005$ and $n_0 = 699915$. Using $t_0 = 2005$, $n_0 = 699915$, $c_0 = 14792864$, $g = 1.4$ and $\gamma = 0.06$ ($\sigma = \gamma / g \approx 0.043$) the trend for $c(t)$ given by equation (2.3) can be fitted by choosing $b \approx 1.6$. **Figure 2b** illustrates the time evolution for the yearly incoming total number of citations divided by the total number of new papers, $m(t)$. Using the parameters from $n(t)$ and $c(t)$ the $m(t)$ trend given by equation (2.1) is plotted by the black curve.

Citations

Fraction of *n* times cited: Facebook and Web of Science

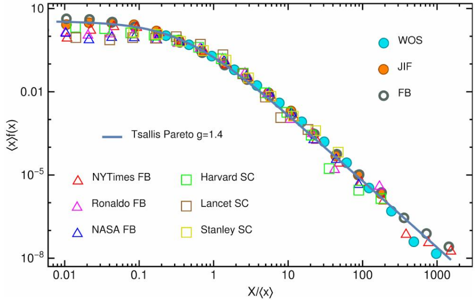
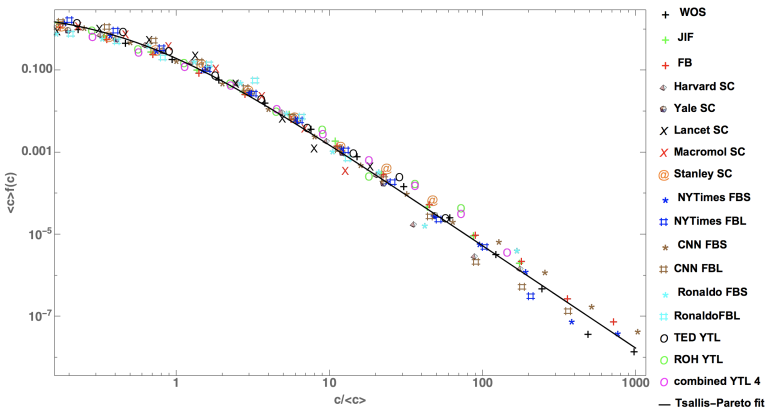


Figure 1. Rescaled distribution of the citation (share) numbers. $f(x)$ is the probability density (PDF) for one paper (post) to have x citations/shares. We present the $\langle x\rangle \cdot f(x)$ value as a function of $x/\langle x\rangle$ ($\langle x\rangle$ the mean value, or first moment of the PDF). For high citation number a clear power-law trend is visible. Different symbols are for different datasets as illustrated in the legend. The considered datasets are described in the Methods section. For high $x/\langle x\rangle$ a clear power-law trend is visible. The entire curve can be well-fitted with a TP distribution (1) with $g \approx 1.4$.

Citations

Fraction of *n* times cited: More Facebook and Web of Science



Summary of Rates and PDF-s

$\mu(x) = \gamma / h(x)$ at constant aging γ

$\gamma_n,\ \gamma(x)$	$\mu_n,\ \mu(x)$	$Q_n,\ Q(x)$
const	const	geometrical $\to$ exponential
const	linear	Waring $\to$ Tsallis/Pareto
const	sublinear power	Weibull
const	quadratic polynomial	Pearson
const	exp	Gompertz
$\ln(x/a)$	σx	Log-Normal
linear	const	Gauss
$\sigma(ax - c)$	σx	Gamma

Deviation shrinks and moves as a soliton: $\dot{x}_c = \mu(x_c)$!

Bibliographical Notes

1. **Tsallis/Pareto degree distribution in huge networks**
Stefan Thurner, Fragiskos Kyriakopoulos, Constantino Tsallis: *Unified model for network dynamics exhibiting nonextensive statistics* Phys. Rev. E **76** (2007) 036111
2. **Avalanche Master Equation for Citations**
A. Schubert, W. Glänzel: *A dynamic look at a class of skew distributions. A model with scientometric applications* Scientometrics **6** (1984) 149-167
3. **Waring distribution** J. O. Irwin: *The Place of Mathematics in Medical and Biological Statistics* J. Roy. Stat. Soc. A **126** (1963) 1-45
4. **Generalized Waring distribution** J. O. Irwin: *The Generalized Waring Distribution Applied to Accident Theory* J. Roy. Stat. Soc. A **131** (1968) 205-225
5. **Rate reconstruction from a truncated integral** W. Glänzel, A. Telcs, A. Schubert: *Characterization by Truncated Moments and Its Application to Pearson-Type Distributions* Z. Wahrscheinlichkeitstheorie verw. Gebiete **66** (1984) 173-183
6. **Avalanche Master eq. for networks** P. L. Krapivsky, S. Redner, F. Leyvraz: *Connectivity of Growing Random Networks* Phys. Rev. Lett. **85** (2000) 4629-4632
7. **Waring Degree Distribution in Network – eq.8 in** P. L. Krapivsky, G. J. Rodgers, S. Redner: *Degree Distributions of Growing Networks* Phys. Rev. Lett. **86** (2001) 5401-5404
8. **Hypergeometrical result in aging site network** S. N. Dorogovtsev, J. F. F. Mendes, *Evolution of networks with aging of sites* Phys. Rev. E **62** (2000) 1842-1845