Potsdam Institute for
Climate Impact Research
PIK

About human activity, long-term memory, and Gibrat's law

Diego Rybski

Sergey V. Buldyrev, Shlomo Havlin,
Fredrik Liljeros, Hernán A. Makse

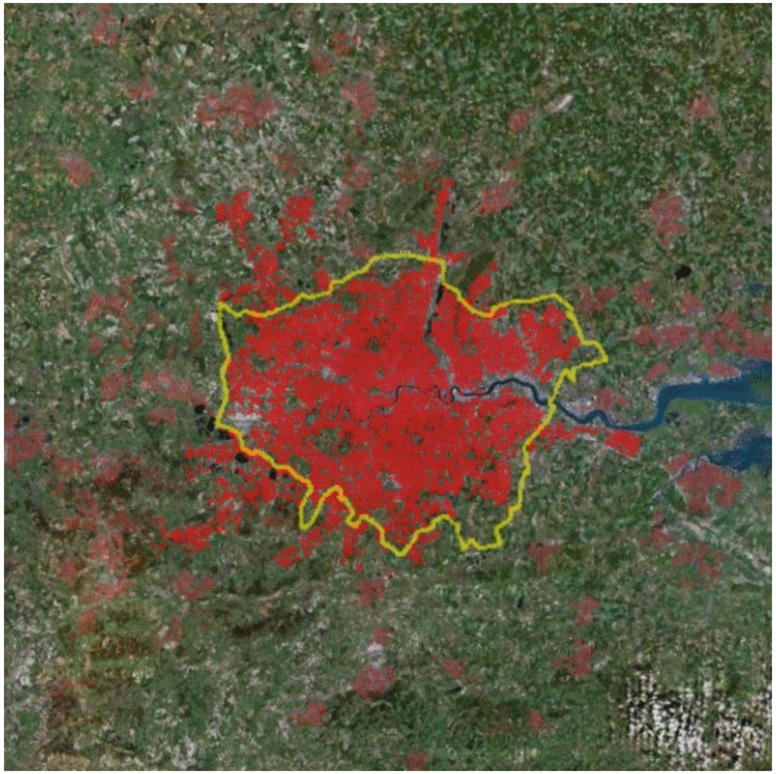
11.1.2011
14:00
Room 1D1

Colloquium, Nonlinear Dynamics, MPI-PKS Dresden

Motivation

City Growth, see:

- Rozenfeld HD, et al., PNAS 105, 2008
- Rozenfeld HD, et al., AER, 2011



Pioneering work

Scaling behaviour in the growth of companies

Michael H. R. Stanley*, Luís A. N. Amaral*, Sergey V. Buldyrev*, Shlomo Havlin*†, Heiko Leschhorn*, Philipp Maass*, Michael A. Salinger‡ & H. Eugene Stanley*

* Center for Polymer Studies and Department of Physics, Boston University, Boston, Massachusetts 02215, USA
† Minerva Center and Department of Physics, Bar-Ilan University, Raman Gan, Israel
‡ Department of Finance and Economics, School of Management, Boston University, Boston, Massachusetts 02215, USA

A SUCCESSFUL theory of corporate growth should include both the external and internal factors that affect the growth of a company[1–18]. Whereas traditional models emphasize production-related influences such as investment in physical capital and in research and development[18], recent models[10–20] recognize the equal importance of organizational infrastructure. Unfortunately, no exhaustive empirical account of the growth of companies exists by which these models can be tested. Here we present a broad, phenomenological picture of the dependence of growth on company size, derived from data for all publicly traded US manufacturing companies between 1975 and 1991. We find that, for firms with similar sales, the distribution of annual (logarithmic) growth rates has an exponential form; the spread in the distribution of rates decreases with increasing sales as a power law over seven orders of magnitude. A model wherein the probability of a company's growth depends on its past as well as present sales accounts for the former observation. As the latter observation applies to companies that manufacture products of all kinds, organizational structures common to all firms might well be stronger determinants of growth than production-related

$$p(r \mid s_0) = \frac{1}{\sqrt{2}\sigma(s_0)} \exp\left(-\frac{\sqrt{2}|r - \bar{r}(s_0)|}{\sigma(s_0)}\right) \qquad (1)$$

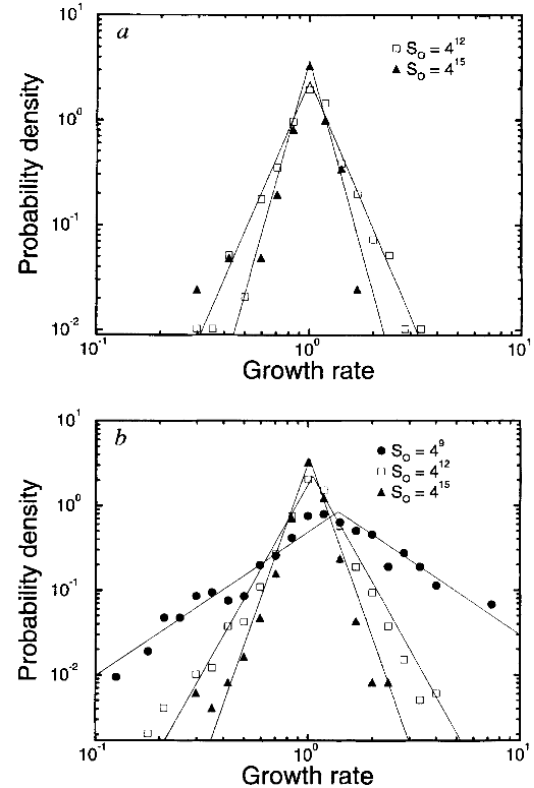


FIG. 1 *a*, Probability density $p(r \mid s_0)$ of the growth rate $r \equiv \ln(S_1/S_0)$ from year 1990 to 1991 for all publicly traded US manufacturing firms in the 1994 Compustat database with standard industrial classification index of 2000–3999. We examine 1991 because between 1992 and 1994 there are several companies with zero sales that either have gone out of business or are 'new technology' companies (developing new products). We show the data for two different bins of initial sales (with sizes increasing by powers of 4): $4^{11.5} < S_0 < 4^{12.5}$ (squares) and $4^{14.5} < S_0 < 4^{15.5}$ (triangles). Within each sales bin, each firm has a different value of *R*, so the abscissa value is obtained by binning these *R* values. The solid lines are fits to equation (1) (in the text) using the mean $\bar{r}(s_0)$ and standard deviation $\sigma(s_0)$ calculated from the data. *b*, Probability density $p(r \mid s_0)$ of the annual growth rate, for three different bins of initial sales: $4^{8.5} < S_0 < 4^{9.5}$ (circles), $4^{11.5} < S_0 < 4^{12.5}$ (squares) and $4^{14.5} < S_0 < 4^{15.5}$ (triangles). The data were averaged over all 16 one-year periods between 1975 and 1991. The solid lines are fits to equation (1) using the mean $\bar{r}(s_0)$ and standard deviation $\sigma(s_0)$ calculated from all data.

NATURE · VOL 379 · 29 FEBRUARY 1996

Pioneering work

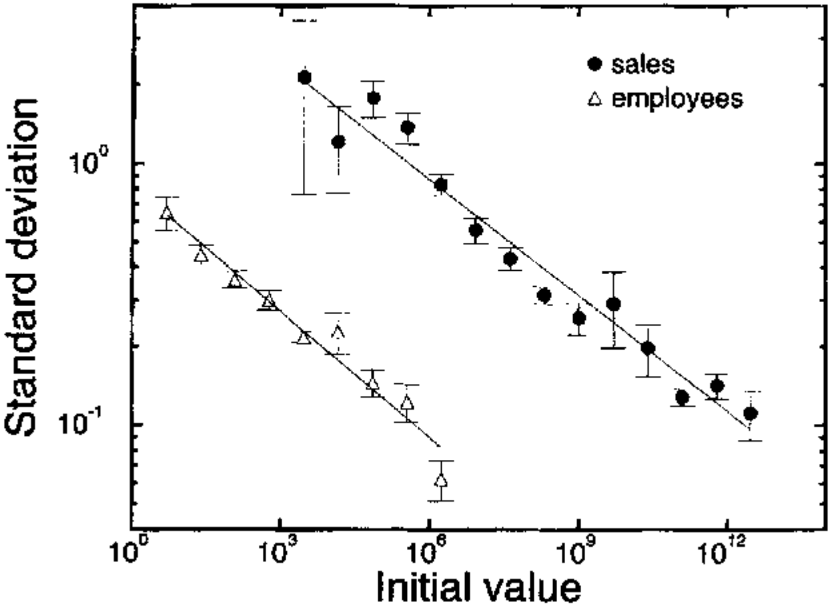


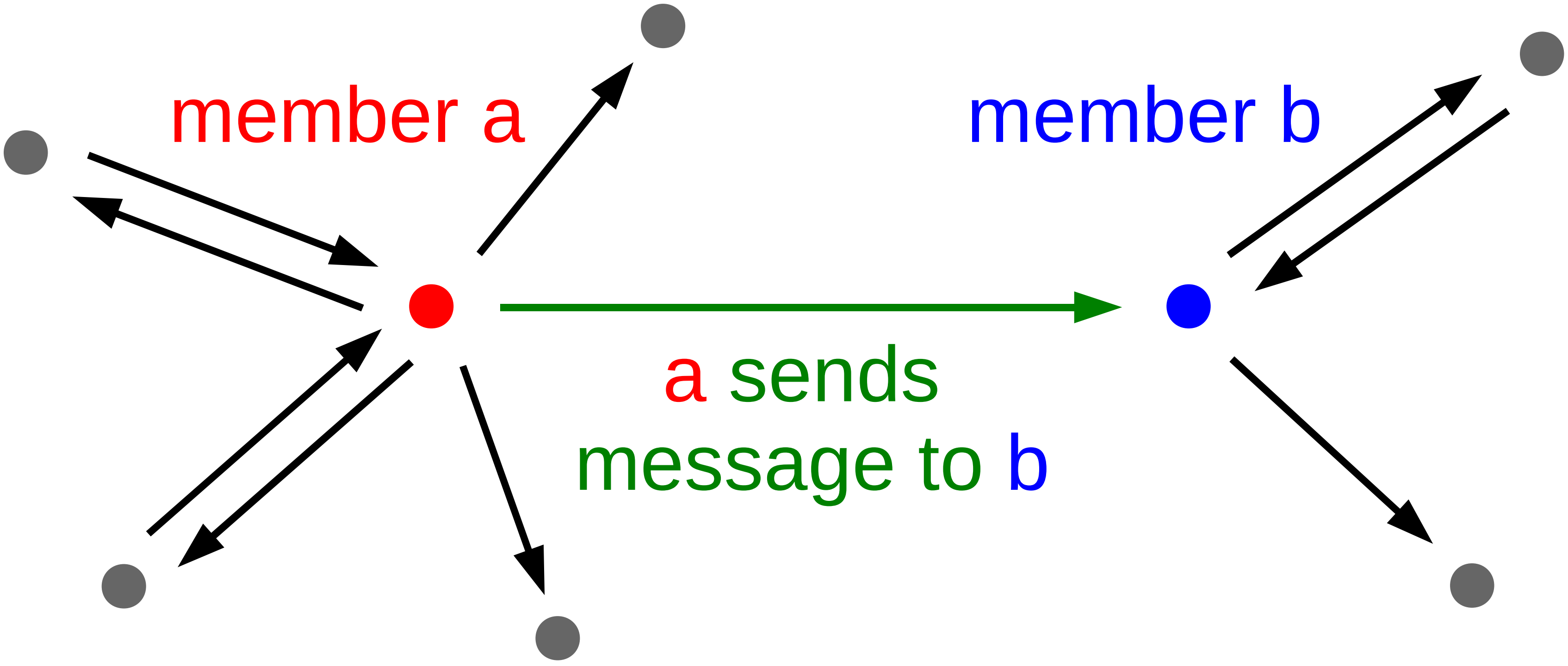
FIG. 2 Standard deviation of the one-year growth rates of the sales (circles) and of the one-year growth rates of the number of employees (triangles) as a function of the initial values. The solid lines are least-square fits to the data with slopes $\beta = 0.15 \pm 0.03$ for the sales and $\beta = 0.16 \pm 0.03$ for the number of employees. We also show error bars of one standard deviation about each data point. These error bars appear asymmetric as the ordinate is a log scale.

NATURE · VOL 379 · 29 FEBRUARY 1996

M.H.R. Stanley et al.

Motivation

online community: members sending messages



either following an existing link $m_{\mathrm{a}} \longrightarrow m_{\mathrm{a}} + 1$

or creating a new one $k_{\mathrm{a}}^{\mathrm{out}} \longrightarrow k_{\mathrm{a}}^{\mathrm{out}} + 1$

=> **growth process**

Outline

0. motivation
1. online community data
2. growth process
3. temporal correlations
4. missing link
5. conclusions

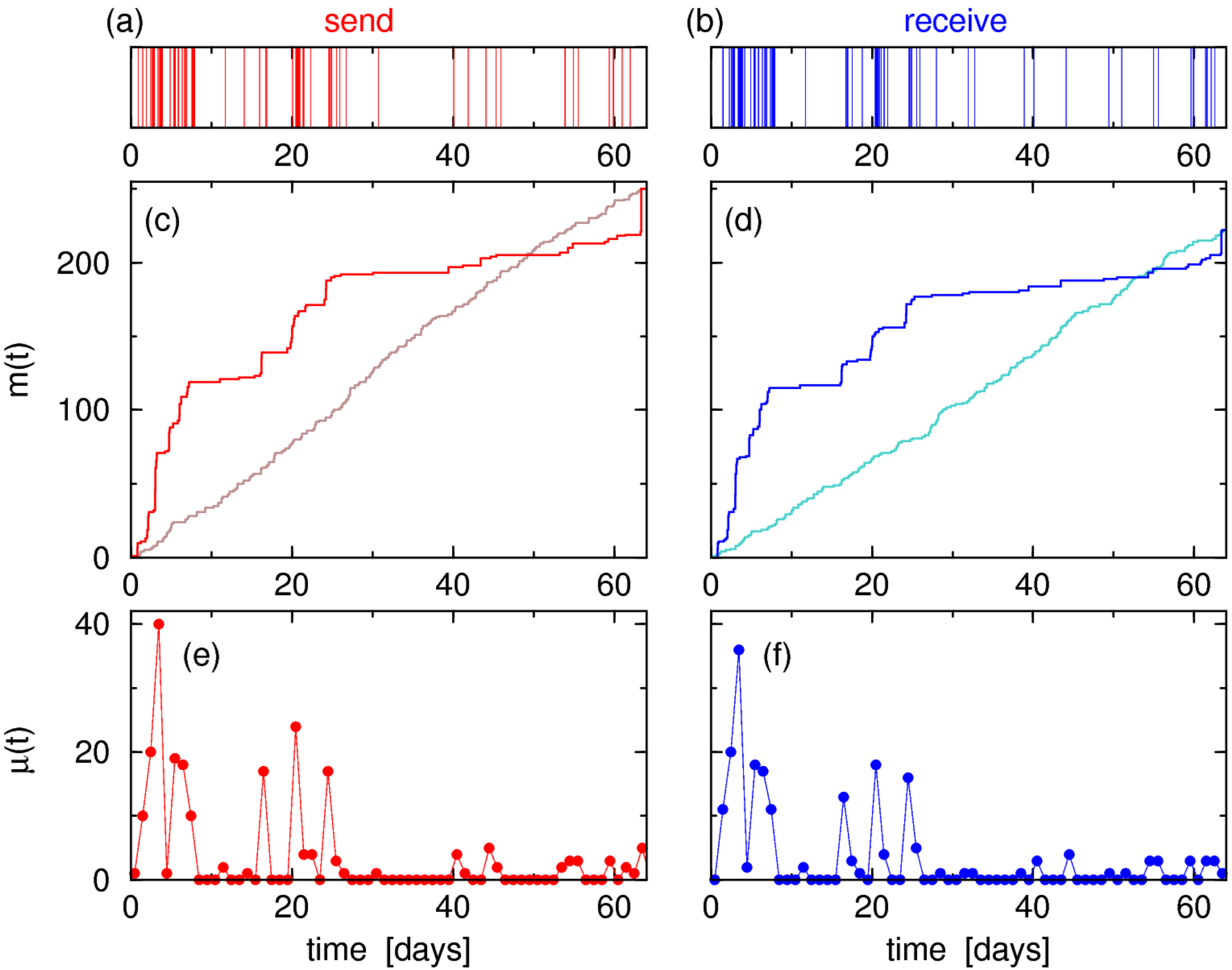
Online community data

online community 1 (OC1):
- 80,000 members
- 12.5 million messages
- 63 days

online community 2 (OC2):
- 30,000 members
- 500,000 messages
- 492 days

both are dating-communities
also used for social interaction in general
completely anonymous

Typical activity (OC1)



Growth process

for each member:

cumulative number of messages $m(t)$

logarithmic growth rate $r = \ln \frac{m_1}{m_0}$

between two time-steps t_0, t_1

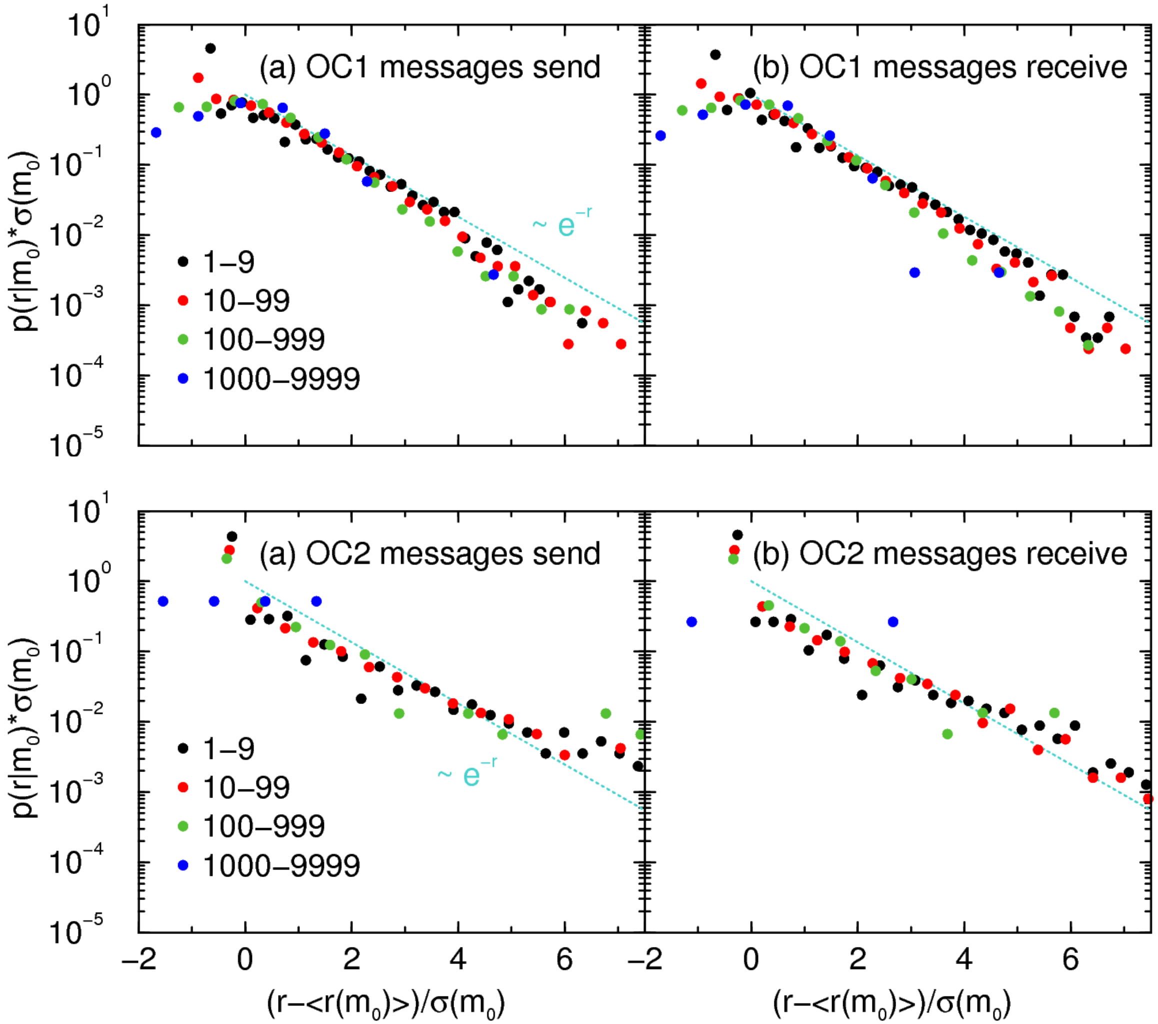
two quantities:

conditional average growth $\langle r(m_0) \rangle = \langle r | m_0 \rangle$

cond. standard deviation $\sigma(m_0) = \sigma(r | m_0)$

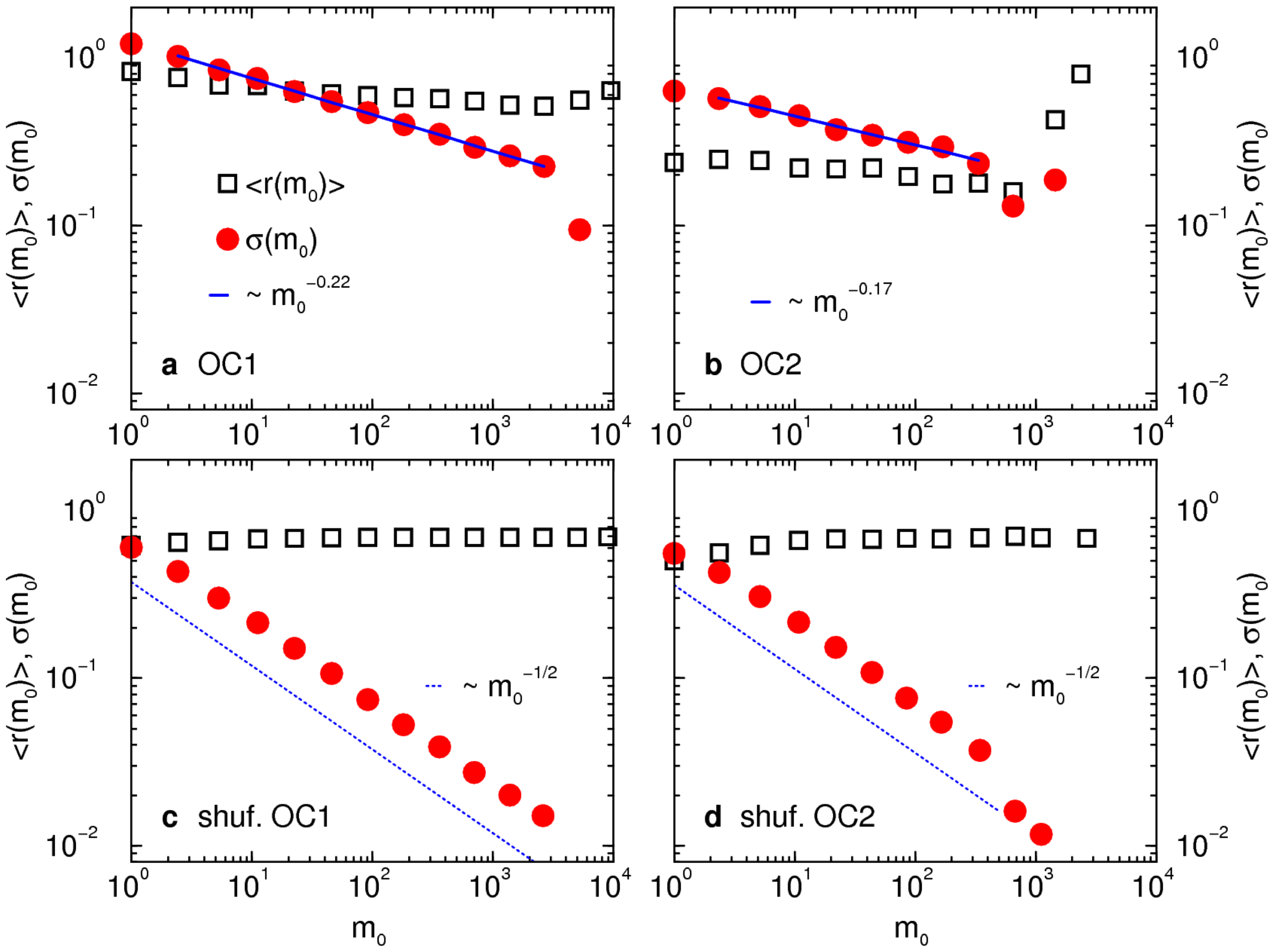
see e.g. M.H.R. Stanley et al., nature, 1996.

Growth process: distribution

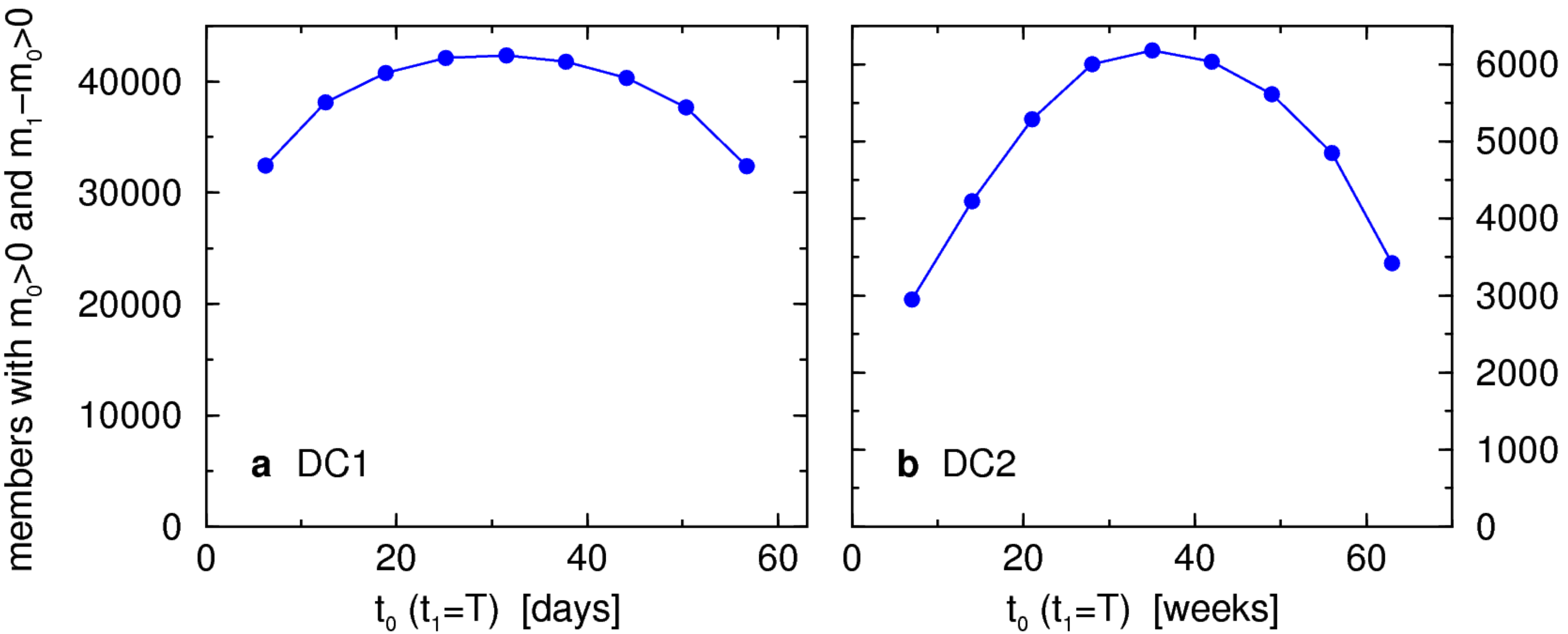


$$p(r|m_0) = \frac{1}{s\sigma(m_0)} \exp\left(-\frac{s|r - \langle r(m_0)\rangle|}{\sigma(m_0)}\right)$$

Growth process: results



Optimal times



Growth process: results

$$\sigma(m_0) \sim m_0^{-\beta}$$

OC1: $\beta_{\mathrm{OC1}} = 0.22 \pm 0.01$
OC2: $\beta_{\mathrm{OC1}} = 0.17 \pm 0.03$
shuffled: $\beta_{\mathrm{rnd}} = 1/2$

Gibrat's law of proportionate growth

multiplicative process
to explain broad distributions (log-normal)

involves assumption: $\langle r(m_0) \rangle = \mathrm{const.}$
$\sigma(m_0) = \mathrm{const.}$

=> $\beta_{\mathrm{G}} = 0$

Temporal correlations

- shuffling destroys temporal correlations, leading to $\beta_{\mathrm{rnd}} = 1/2$

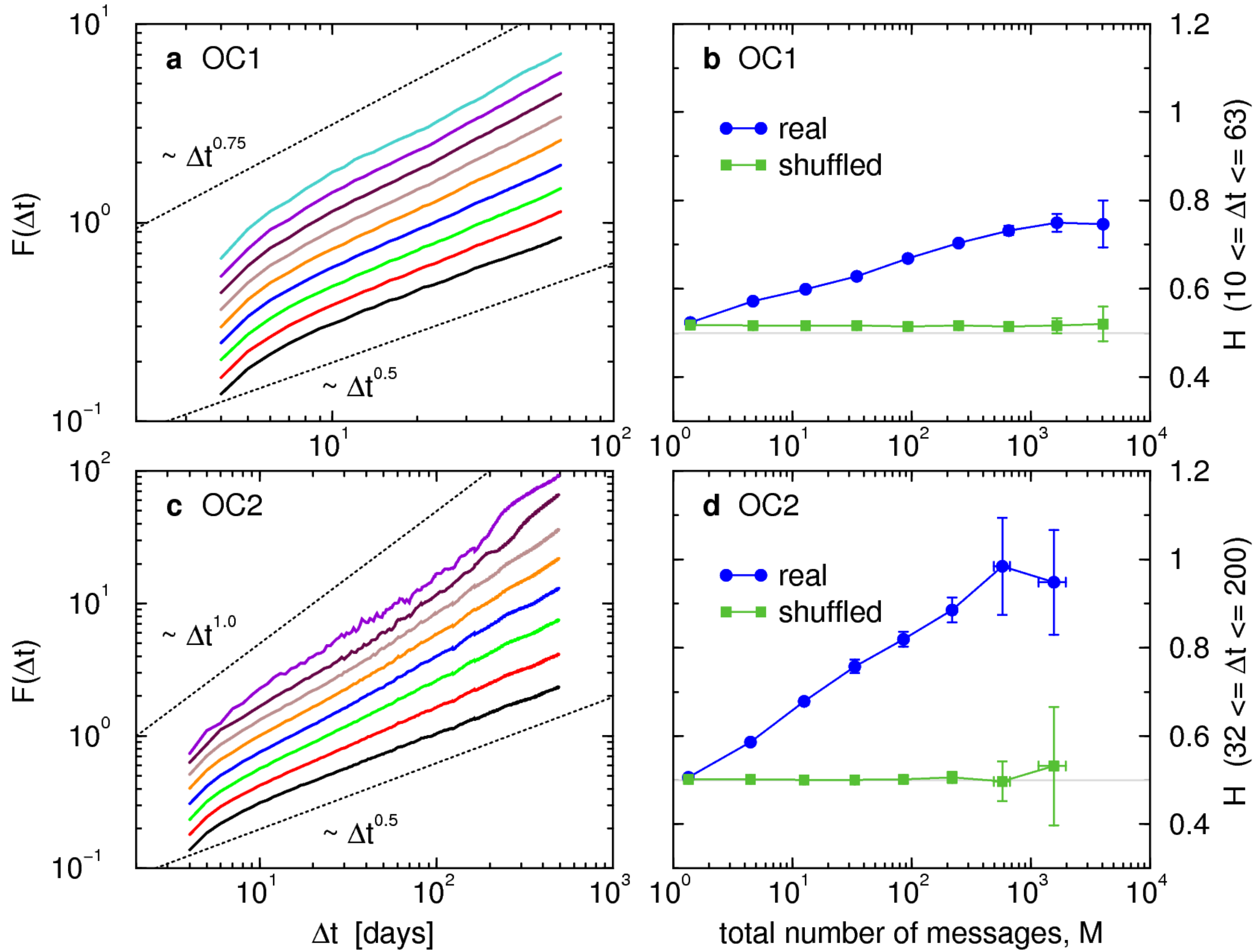
- this suggests $\beta \approx 0.2$ might be due to temporal correlations

- we use Detrended Fluctuation Analysis (DFA) to quantify long-term correlations in the activity (messages per day): $\mu(t)$

fluctuation function: $F(\Delta t) \sim (\Delta t)^H$

$1/2 < H < 1$ => ltc

Temporal correlations: results



Missing link

derivation leads to:

$$\beta = 1 - H$$

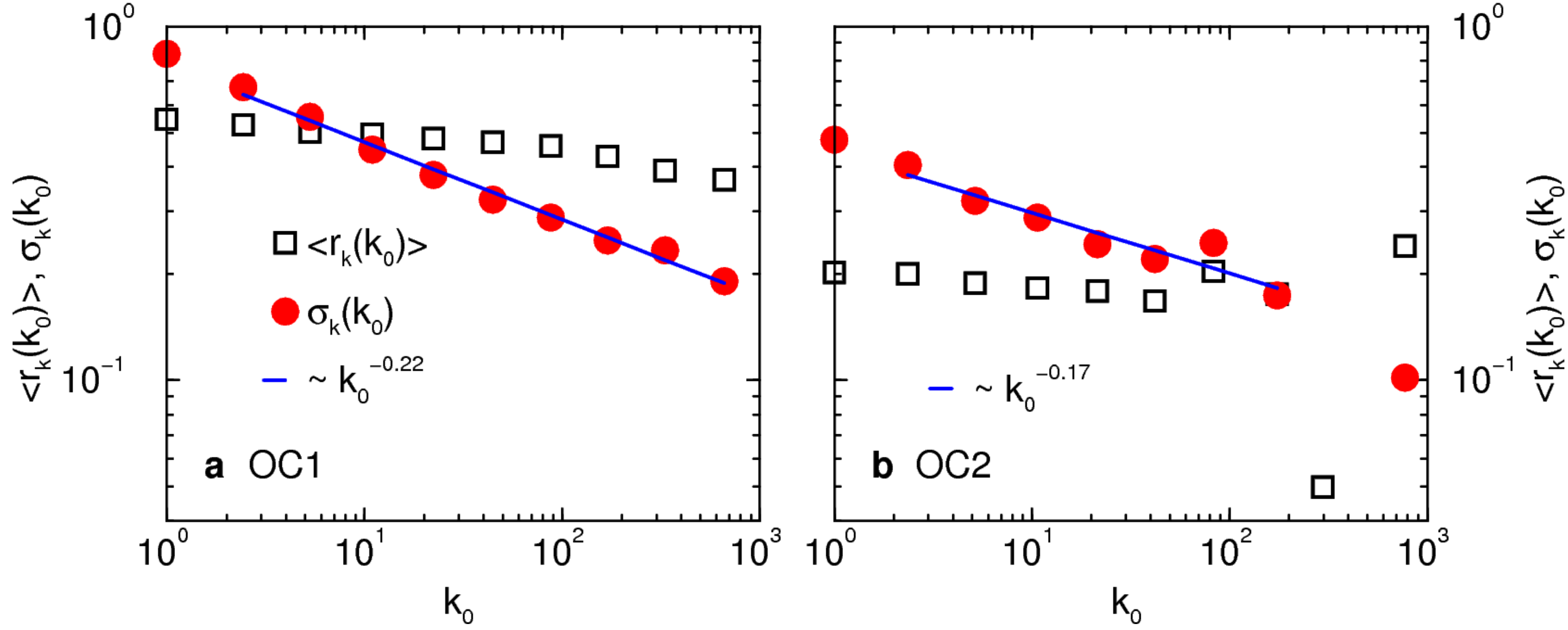
accordingly:

$$\beta \approx 0.2 \quad \Rightarrow \quad H \approx 0.8$$ OCs

$$\beta_{\mathrm{rnd}} = 1/2 \quad \Rightarrow \quad H_{\mathrm{rnd}} = 1/2$$ shuffled

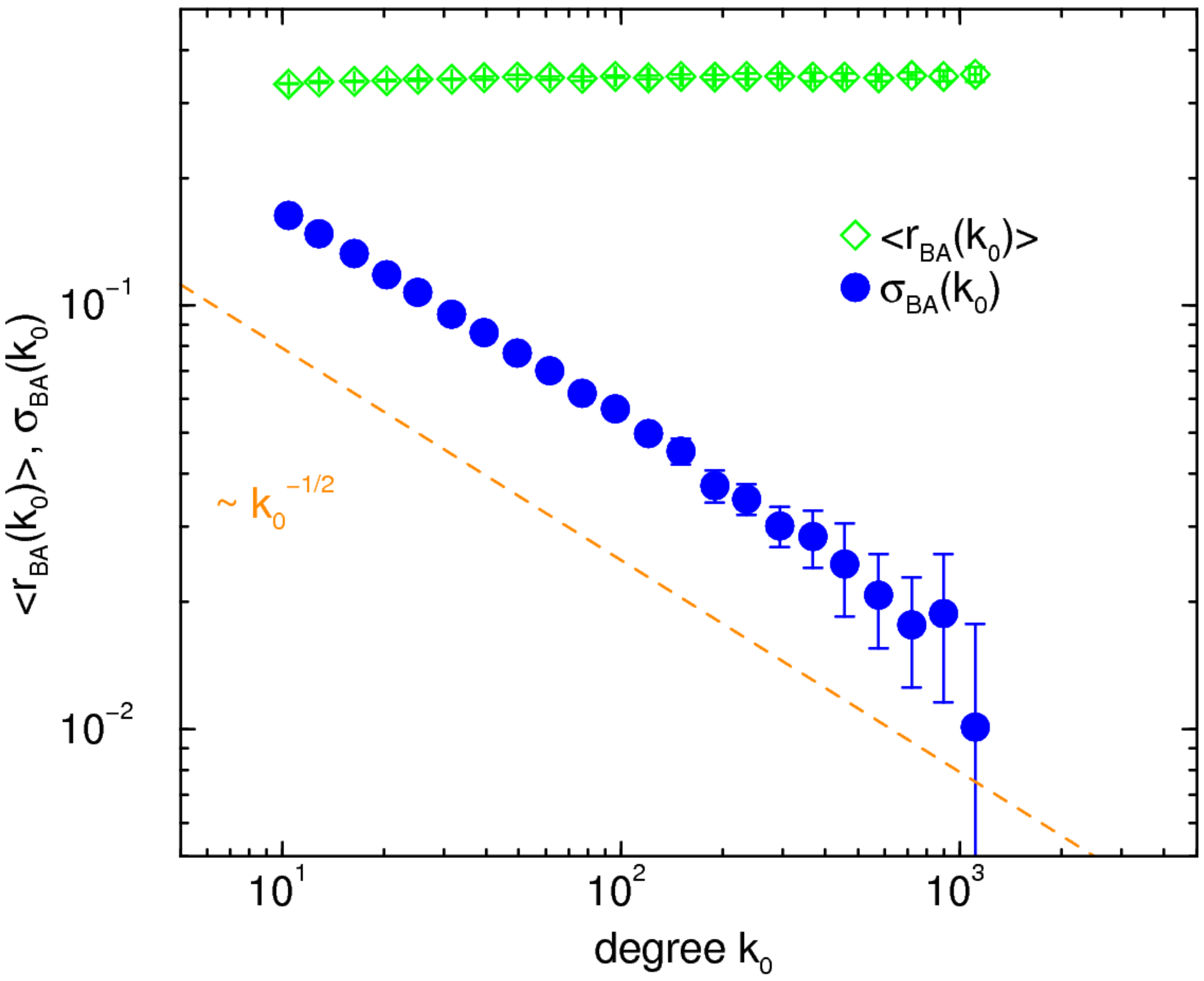
$$\beta_{\mathrm{G}} = 0 \quad \Rightarrow \quad H_{\mathrm{G}} = 1$$ Gibrat's law

Growth process: out-degree



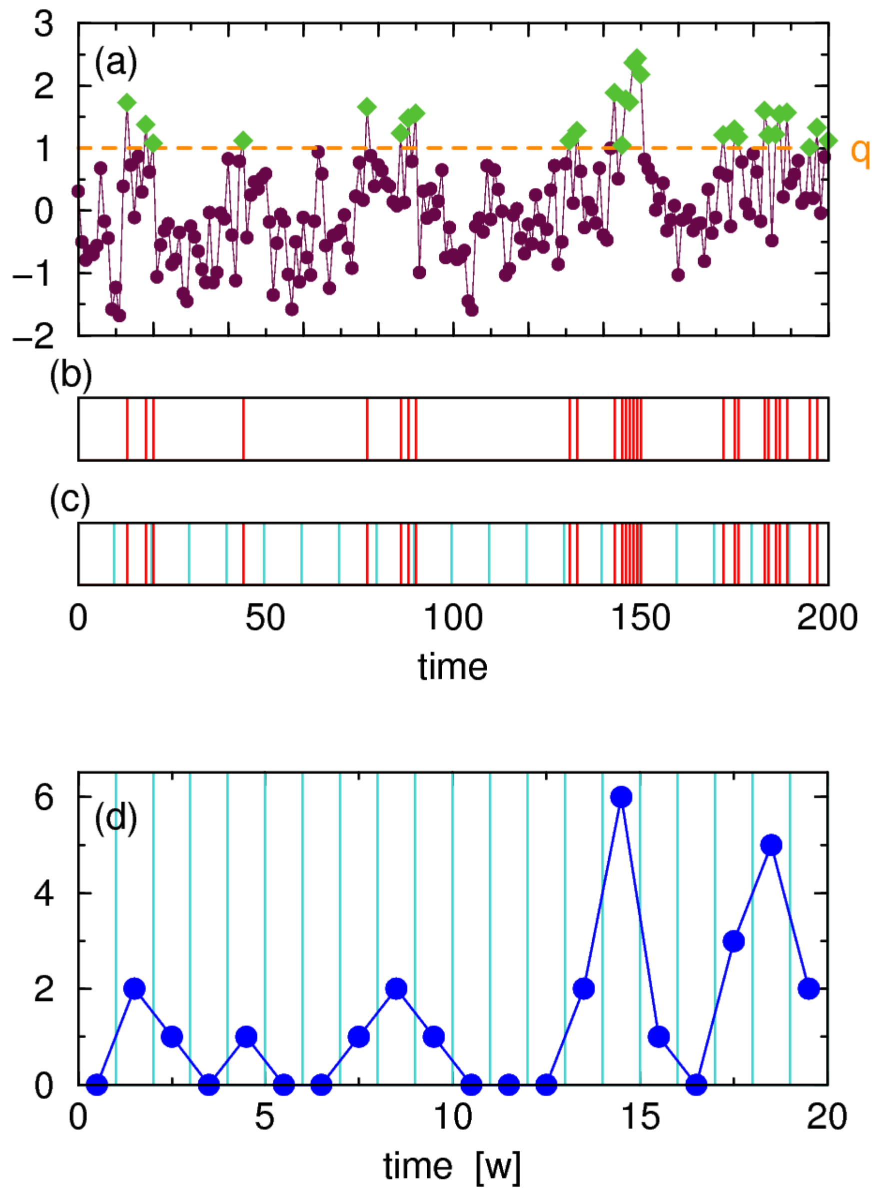
see also: Maillart T, et al., PRL 101, 2008

Growth process: preferential attachment

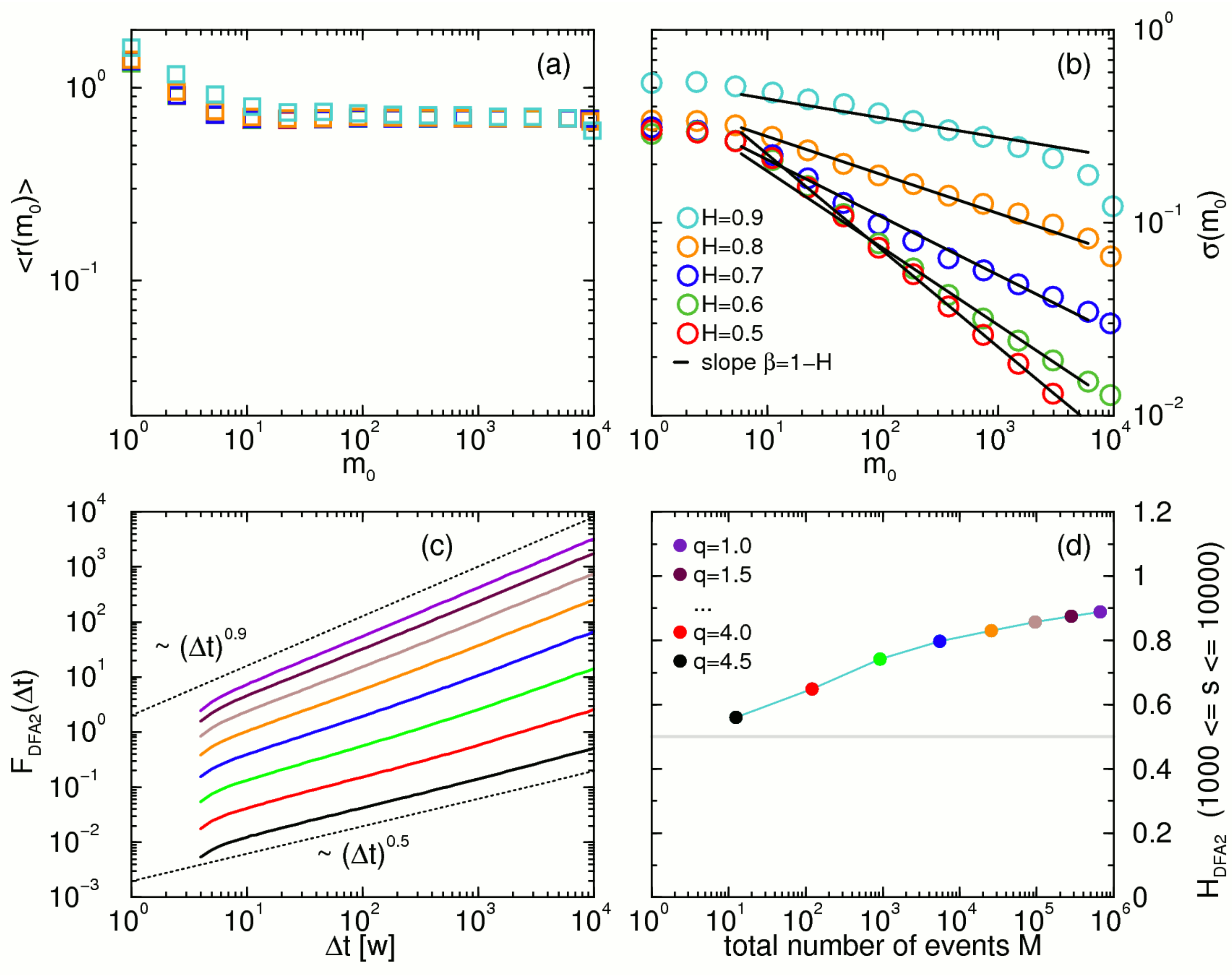


see also: Barabasi AL and Albert R, Science 286, 1999

peak over threshold simulations



peak over threshold simulations



Conclusions

1. scaling in growth of number of messages or out-degree implies that active members are better predictable than less active ones

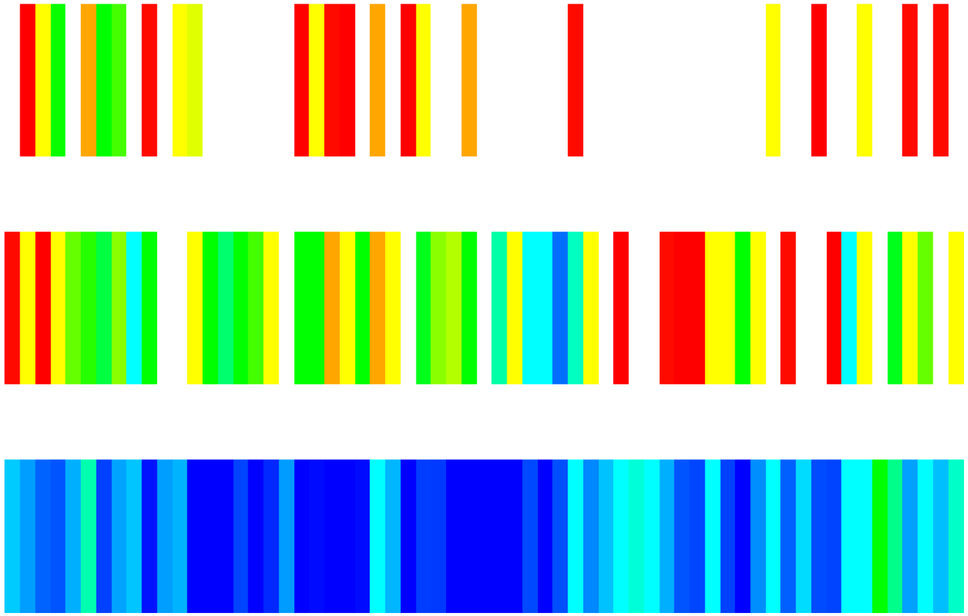
2. human activity sending messages is long-term correlated

3. scaling in growth is due to long-term correlations $\sigma(m_0) \sim m_0^{-\beta}$

=> this may also be the case for other data

D. Rybski et al., PNAS, 2009

Thank you for your attention.



http://www.rybski.de/diego/

http://www.pik-potsdam.de/members/rybski/