Rare events

in physics and finance: Concepts and applications

J.P Bouchaud

with G. Biroli, P. Cizeau, R. Chicheportiche, Y. Fyodorov, M. Mézard, M. Potters

Fat tails and the Central Limit Theorem

- Sum of random variables: $S_N = \sum_i x_i$, with

$$\rho(x) \sim_{|x|\to\infty} |x|^{-1-\mu}$$

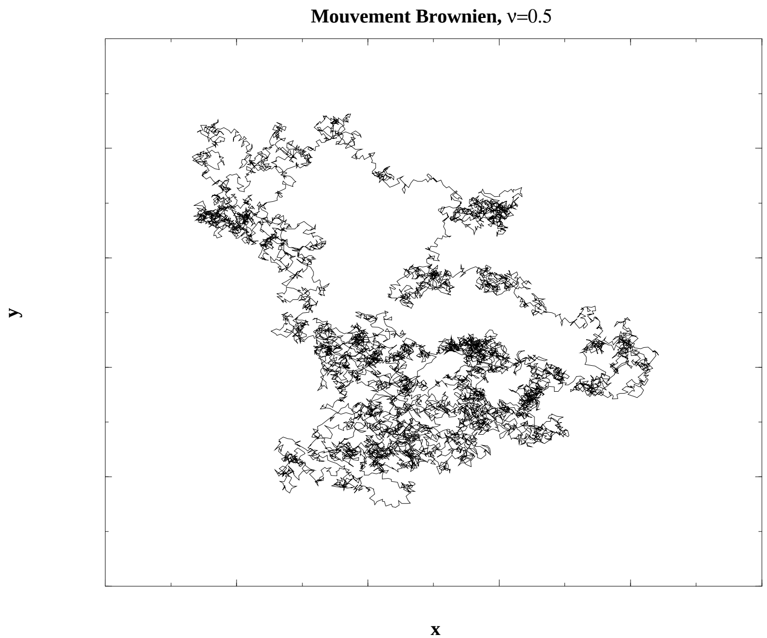
  - $\mu > 2$, finite variance $\rightarrow S_N = uN^{1/2}$ and u Gaussian (in the "central" region)

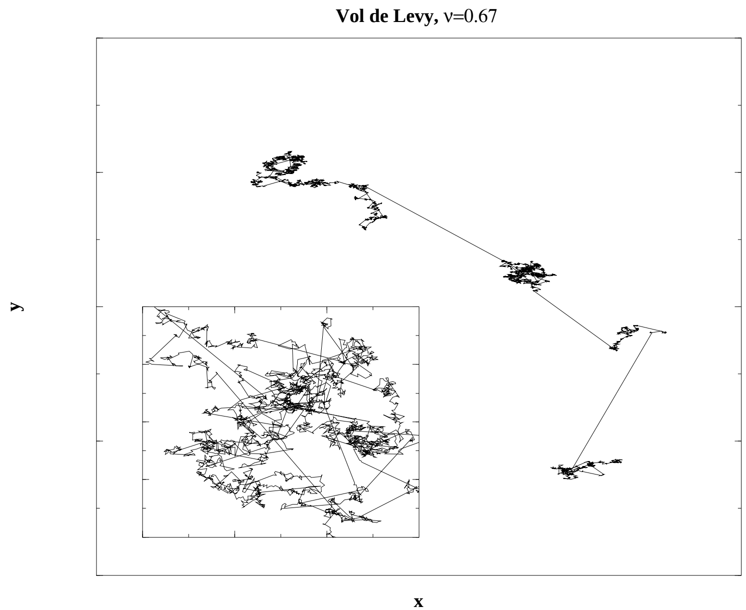
  - $\mu < 2$, infinite variance $\rightarrow S_N = uN^{1/\mu}$ and u Lévy distributed: $L_\mu(u)$

- Random walks and Lévy flights

- Epidemic propagation [Brockmann et al.]

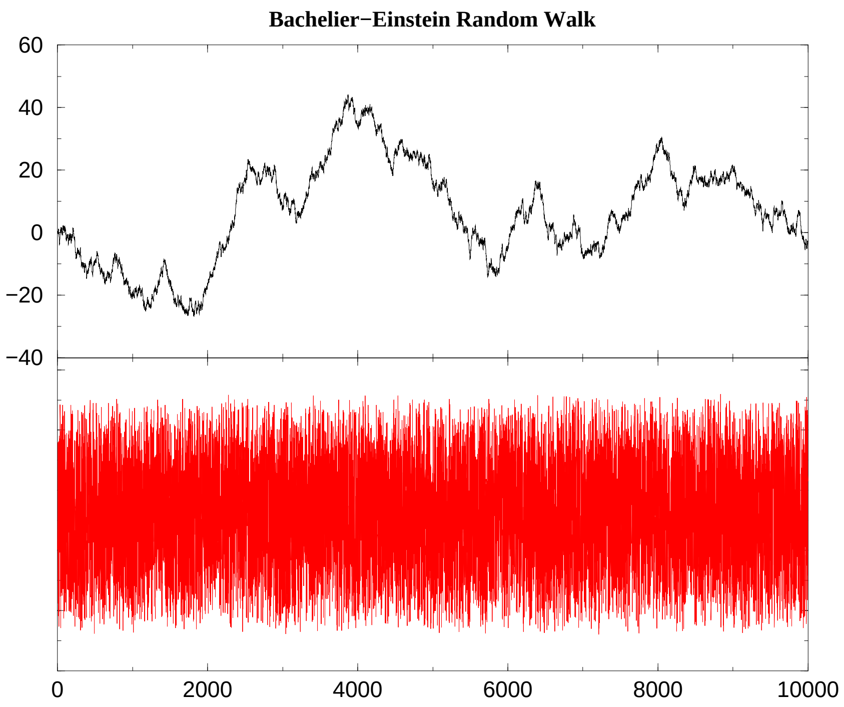
Brownian and Lévy flights

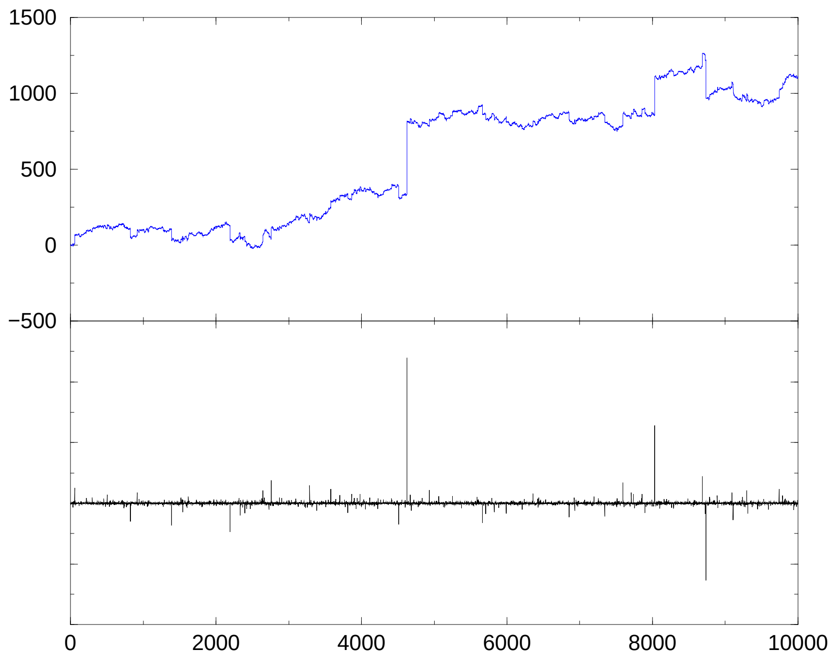




$$\mu = \infty \longrightarrow \mu = 3/2$$

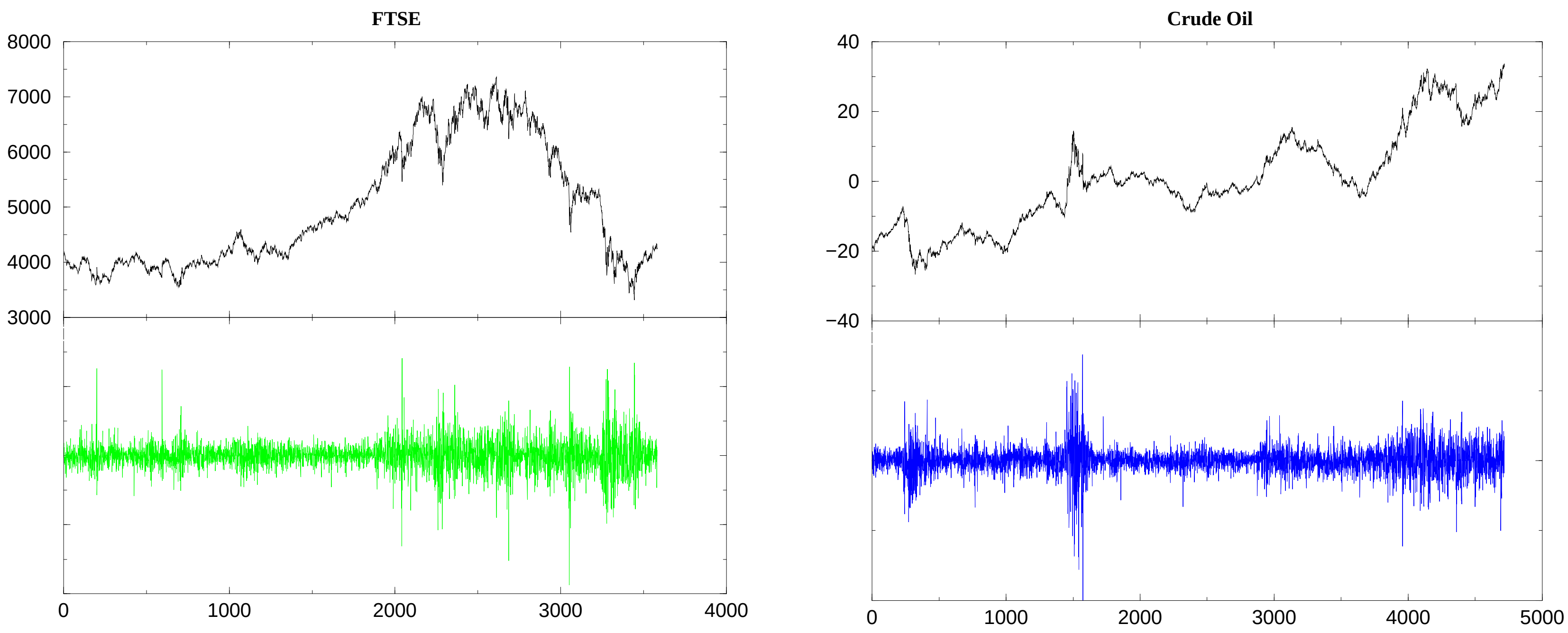
Brownian and Lévy flights





$$\mu = \infty \longrightarrow \mu = 3/2$$

Diffusion of prices



$\mu \approx 3 - 4$ but strong intermittency/clustering

Fat tails and condensation

- Sum of positive random variables: $Z_N = \sum_i w_i$, with

$$\rho(w) \sim_{w\to\infty} |w|^{-1-\mu}$$

- Examples: trapping times, weights (classical $e^{-\beta E}$/quantum mechanical $|\psi|^2$, etc.), wealths, population sizes, etc.

  – $\mu > 1$, finite mean $m \to Z_N \sim mN$; the contribution of each term $\pi_i = w_i/Z_N \to 0$ when $N \to \infty$

  – $\mu < 1$, infinite mean $\to Z_N \sim N^{1/\mu}$; the contribution of some terms $\pi_i = w_i/Z_N = O(1)$ when $N \to \infty$

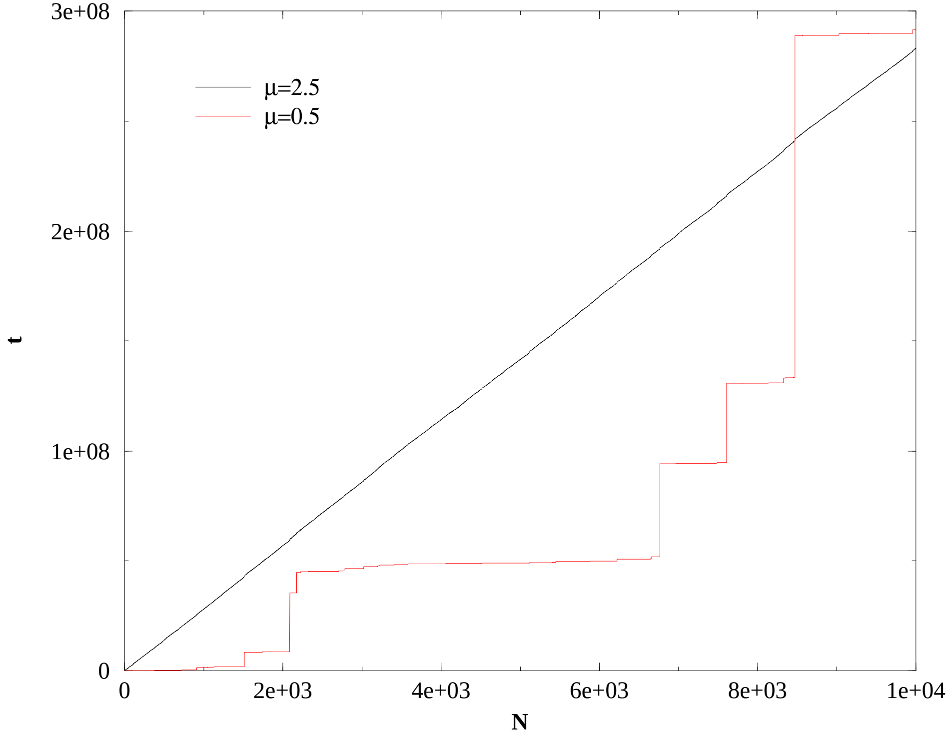
- Indicators of condensation/localization/inequality: participation ratio/Herfindahl index/Simpson index

$$Y = \sum_i \pi_i^2 \to_{N\to\infty} 0 \quad (\mu \geq 1); \quad O(1) \quad \text{with} \quad \overline{Y} = 1-\mu, \quad (\mu < 1)$$

Herfindahl index and antitrust laws

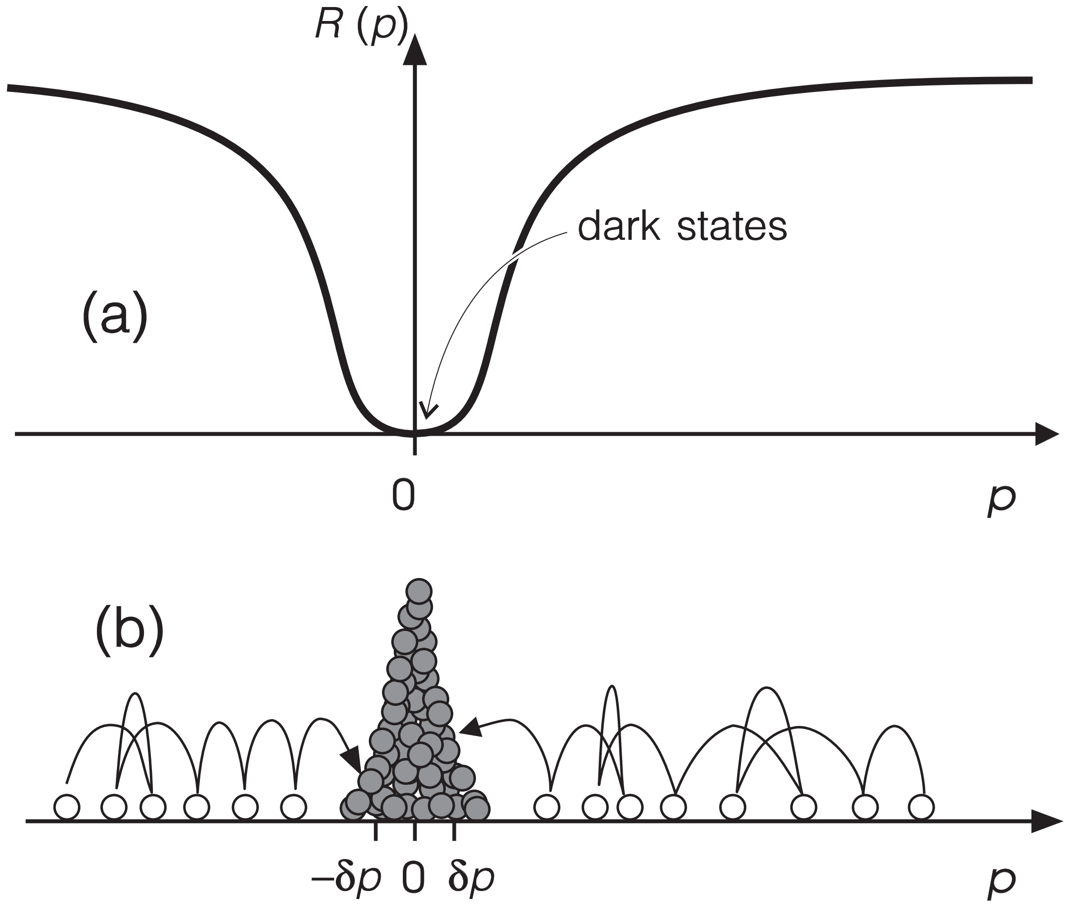
- *Markets in which the Herfindahl index is between* 0.10 *and* 0.18 *are considered to be moderately concentrated, and those in which the index is in excess of* 0.18 *are considered to be concentrated. Transactions that increase the index by more than* 0.01 *in concentrated markets presumptively raise antitrust concerns under the Horizontal Merger Guidelines issued by the U.S. Department of Justice and the Federal Trade Commission.*

$\mu < 1$ and concentration



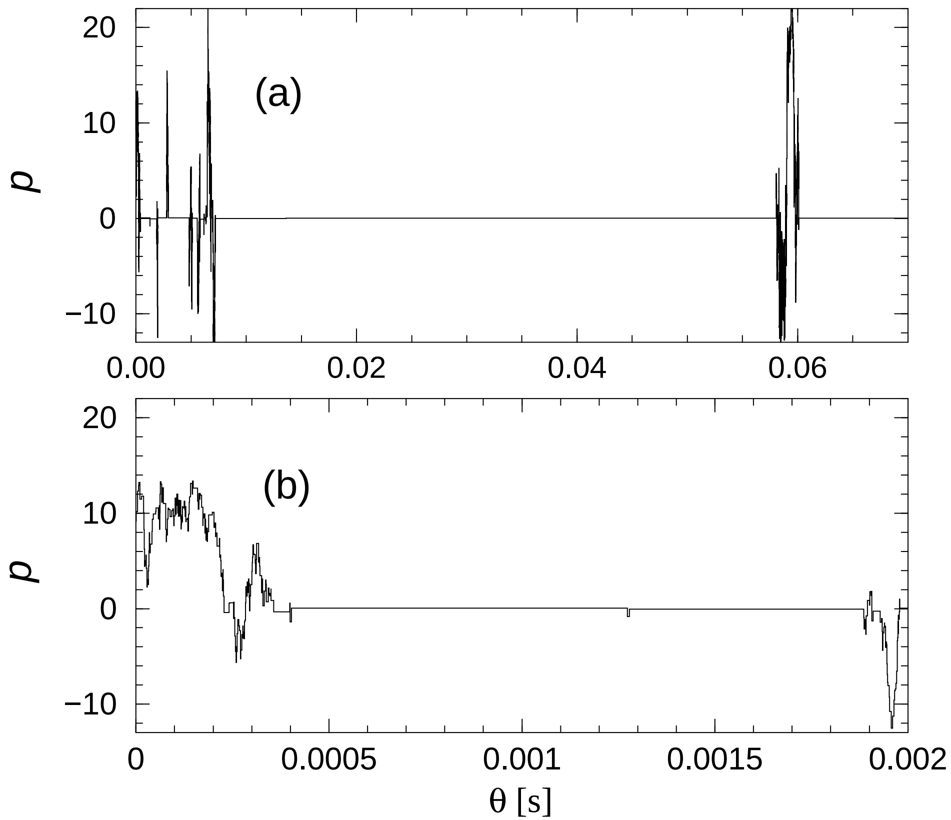
The Grocer's sales vs. the Jeweler's sales as a function of the number of sold products

Anomalous trapping and laser cooling



Atoms that land close $p = 0$ stay there for $\tau \sim p^{-2}$ (with [Aspect, Bardou, Cohen-Tannoudji])

Anomalous trapping and laser cooling



$\mu = 1/2$: The longer one waits, the longer the trapping time

(aging) $\rightarrow T = 2\text{nK}$!

The S&P500 and lack of self-averaging

- The S&P 500 is a market cap. weighted sum of individual returns r_i

$$I = \sum_i \pi_i r_i,$$

- The market cap is Pareto (power-law) distributed with $\mu \approx 1$

- How representative is the index? Suppose the r_i's are IID with average m

$$\langle I \rangle = m; \qquad \langle (I-m)^2 \rangle = \sigma^2 \sum_i \pi_i^2 \longrightarrow_{N\to\infty} 0 \quad \text{iff}\, \mu \geq 1,$$

but remains finite if $\mu < 1$

- Of course, most $\langle r_i r_j \rangle > 0 \rightarrow$ see below

Extreme value statistics

- Extreme value distributions: $M_N = \max_i\{x_i\}$

  $M_N = a_N + b_N u$ with:

  - Exponential variables: Gumbel – $G(u) = \exp[-u - \exp[-u]]$
  - Power-law variables: Fréchet – $F(u) = \mu \exp[-u^{-\mu}]/u^{1+\mu}$
  - Bounded variables: Weibull – $H(u) = \mu u^{\mu-1} \exp[-u^{\mu}]$

- Example: when $\rho(x)$ is Gaussian, $a_N \approx \sigma\sqrt{2 \ln N}$, $b_N \approx a_N/(2 \ln N)$, and the right tail $u \gg 1$ is a *pure exponential*

- Note: beware of (a) the scaling region and (b) the slow convergence ($1/\ln N$)

From sums to max: sums of exponentials

- A generic problem: $Z_N(t) = \sum_{i=1}^{N} e^{t\gamma_i}$,

  γ_i random, t: a parameter

- Examples where t is time:

  - Growth of total population of N species with different growth rates γ_i

  - Growth of population in N cities with different growth rates γ_i

  - Growth of portfolio value made of N assets with different yield rates γ_i, or of the total wealth of N individuals with different investment strategies, etc.

  - Note: $\gamma_i = \frac{1}{t} \sum_{t'=1}^{t} \eta_i(t')$

From sums to max: sums of exponentials

- A generic problem: $Z_N(t) = \sum_{i=1}^{N} e^{t\gamma_i}$, γ_i random, t: a parameter

- <u>Other examples:</u>
  - $(-\gamma_i)$ is a (random) energy of configuration i: Z_N is partition function of the Random Energy Model of glasses ([Derrida]) at temperature $T = 1/t$.
  - γ_i is a random attenuation rate: Z_N is the total transmitted intensity of a bundle of N weakly flawed optical fibers of length L ([Lifshitz & Pastur])
  - γ_i is an inverse localisation length of a 1-d disordered quantum chain

From sums to max: sums of exponentials

- Clearly:
  - CLT when $N \to \infty$ at fixed t, but
  - $Z_N(t) \approx e^{t\gamma_{\text{max}}}$ for fixed N when $t \to \infty$

- What happens in between ?

A Random Condensation Transition

- A universal scenario when γ in the Gumbel class, i.e. $\ln \rho(\gamma) \sim -\gamma^s$ when $\gamma \to \infty$ – e.g. $s = 2$: $\rho(\gamma)$ Gaussian of mean zero and variance Σ^2.

- Then, if $N \to \infty$, $t \to \infty$ at fixed

$$\mu = \frac{\sqrt{2 \ln N}}{t\Sigma},$$

one finds exactly the 3 phases ([Ben Arous et al.]):

  - $\mu > 2$, Z_N is Gaussian distributed
  - $1 < \mu < 2$, Z_N is Lévy distributed, but no condensation $Y = 0$
  - $\mu < 1$: Z_N is Lévy distributed, and condensation $Y > 0$

A Random Condensation Transition

- A true phase transition when $N \gg 1$ fixed, as t increases beyond $t_c = \sqrt{2 \ln N}/\Sigma$ (Note: different scalings for $s \neq 2$)

- Interpretation (for the different examples)

- In the random growth cases, the population/wealth ends up being concentrated within a finite number of species/cities/assets /individuals when $t \to \infty$

- In the random fiber bundle case, the transmission is overwhelmingly due to a few "transparent" fibers

A Random Condensation Transition

- The REM: the thermodynamics of $N = 2^M$ states of energies with $\Sigma \sim \sqrt{M}$ at low temperatures $T < T_c$ is dominated by a finite number of configurations, where the system spends most of its time – a scenario for the glass transition: condensation in a few amorphous states

- Note 1: the distribution of trapping times is power law distributed with $\mu = T/T_c < 1 \rightarrow$ aging

- Note 2: Y is the probability that two randomly chosen configurations end up in the same state after a long time

- Note 3: Gumbel statistics are equivalent to "1-Step Replica Symmetry Breaking"

Sums of exponentials, extensions

- Extension A: what happens when some "mixing" is present – e.g. inter-city migrants, portfolio rebalancing, wealth re-distribution, etc. ?

- Extension B: what happens when the γ_i's are correlated ?

A. Sums of exponentials with mixing

- Previous model: $Z_N(t) = \sum_i w_i(t)$ with $w_i(t+1) = e^{\eta_i(t)} w_i(t)$, with $\eta \sim \mathcal{N}(0, \sigma)$

- Random mixing (without geometry): take a fraction f of w_i and redistribute equally among z cities/assets/individuals...

$$w_i(t+1) = e^{\eta_i(t)} w_i(t) + f \left[\sum_{j \in V_t(i)} \frac{w_j(t)}{z} - w_i(t) \right]$$

- Random mixing (with geometry); think of the i's on a d-dimensional grid $\vec{x}$, and redistribute among neighbours:

$$w_x(t+1) = e^{\eta_x(t)} w_x(t) + \frac{f}{2} \left[w_{x+1}(t) + w_{x-1}(t) - 2 w_x(t) \right]$$

A. Sums of exponentials with mixing

- Random growth + Diffusion – How is the condensation phenomenon affected?

A. Sums of exponentials with mixing

- Without geometry: in the large N, t regime, there is a critical value $f_c(z, \sigma)$ below which $(/\sigma_c(z, f)$ above which) there is condensation of the REM type above

- In the fully connected case $z = N - 1$, there is no condensed phase and $\mu = 1 + \frac{f}{\sigma^2}$ (cf. cities $\mu \approx 1$)

- Redistribution tempers the extreme condensation that occurs for $f = 0$ at large times

- A variation on the theme: consider companies that grow randomly and sometimes merge, with rate $\lambda \longrightarrow$ same phenomenology, same condensation transition, Pareto tail in the size distribution

A. Sums of exponentials with mixing

- With geometry: in the large N, t regime, the situation depends on the dimensionality d

- For $d = 1, 2$, one is always in the ultra-condensed phase $\mu \to 0$ when $t \to \infty$, for any f, σ,

  For $d \geq 3$, there is a transition point σ_c above which the system is condensed

- There is a topological aspect to redistribution – chain geometries are not "mixing" enough, and population "peaks" on favored (but time dependent) spots

A. Sums of exponentials and Directed Polymers

- Note: $Z_N(t)$ is the partition function of the so-called "Directed Polymer in Random Media" problem

$$E(\text{path}) = \sum_{(x,t)\in\text{path}} \eta_x(t)$$

- For $d = 1$, $t \to \infty$, $Z_N(t) \approx e^{E_{\max}}$ where $E_{\max}$ is the "reward" of the best path of length t, with random 'bounties' on each site:

$$E_{\max} = t\varepsilon + t^{1/3}u$$

where u has an exactly known Tracy-Widom distribution: max of non trivially correlated variables!

Fat tails and Directed Polymers

- Case where $\rho(\eta) \sim \eta^{-1-\mu}$? Naive guess: $\mu = 2$ should play a role

- In fact: the Derrida-Spohn solution on a tree loses its meaning as soon as $\mu < \infty$...

- Simple argument in 1+1: The path distorts to grasp extreme bounties

$$\eta_{\text{max}} \sim (Wt)^{1/\mu} \quad \text{should balance} \quad W^2/t \rightarrow W \sim t^{\frac{1+\mu}{2\mu-1}}$$

  suggests $\mu = 5$!

- Numerical simulations suggest the argument is exact: $\delta E_{\text{max}} = \varepsilon t^{\frac{3}{2\mu-1}}$ with ε distributed as a geometric sum of Fréchet distribution ([GB,MP,JPB])

Fat tails and Directed Polymers

- Open problem: How to capture these effects within a perturbative FRG formalism, that only seems to care about $\mu = 2$ – could shed light on the unsolved problems, like the upper critical dimension $d_c + 1$

B. Case with correlations

- Recap: $Z_N(t) = \sum_{i=1}^{N} e^{t\gamma_i}$, $\gamma_i = \mathcal{N}(0, \Sigma)$, IID $\longrightarrow$

$$\mu = \frac{\sqrt{2 \ln N}}{t\Sigma}$$

  so if t and Σ are finite, one is always in the delocalized (uncondensed) phase at large N

- What happens with correlations?

- Consider again the problem on a grid and introduce *spatial correlations*:

$$\overline{(\gamma_{\vec{x}} - \gamma_{\vec{y}})^2} = 2d\Sigma^2 F(|\vec{x} - \vec{y}|)$$

  with $N = L^d$.

B. Case with correlations

- Three cases:
  - Short range correlations: $F(r) \to F_0$ when $r \to \infty$: same results as without correlations (Gumbel is forgiving)

    Transition only if $\Sigma \propto \ln N$
  - Long range correlations: $F(r) \to r^{2\theta}$ when $r \to \infty$ ($\theta > 0$) – example $d = 1$, $\theta = 1/2$: Sinai landscape

    Always in the ultra-condensed phase for large N
  - Marginal case: $F(r) \sim \ln(r^2 + a^2)$ – then there is a REM-like transition with

    $$\mu = \frac{1}{t\Sigma}$$

    in all dimensions ([Carpentier-Ledoussal])

B. Case with correlations

Note 1: in this case the largest γ is weakly non Gumbel:

$$\gamma_{\text{max}} = d\Sigma \ln L + u$$

with $P(u) \sim ue^{-u}$ for large us [YF,JPB]

Note 2: Extension to a hierarchical, Parisi landscape by superimposing potentials with different a^2 [YF,JPB]

B. The marginal case: applications

- Stat. Mech. applications: particle in a random, logarithmically correlated potential (e.g. vortices in 2d Josephson arrays), with interesting statics and dynamics

- Finance: building block for a multifractal random walk [Bacry-Muzy-Delour]:

$$\ln \frac{p(t)}{p(0)} = \sigma_0 \sum_{t'=1}^{t} \xi_{t'} e^{\gamma_{t'}} \qquad \xi \sim \mathcal{N}(0,1)$$

A random volatility model where the log-volatility is Gaussian with logarithmic correlation:

$$\overline{(\gamma_s - \gamma_{s'})^2} = \lambda^2 \ln(\tau_0^2 + (s - s')^2)$$

λ^2 is the “vol of vol”

Log-vol correlations

B. The multifractal random walk

- log-price differences is a subordinated Gaussian process, with a random time

$$\tau(t) = \sum_{t'=1}^{t} e^{2\gamma_{t'}}, \qquad \equiv Z_N \quad \text{with} \quad N \to t$$

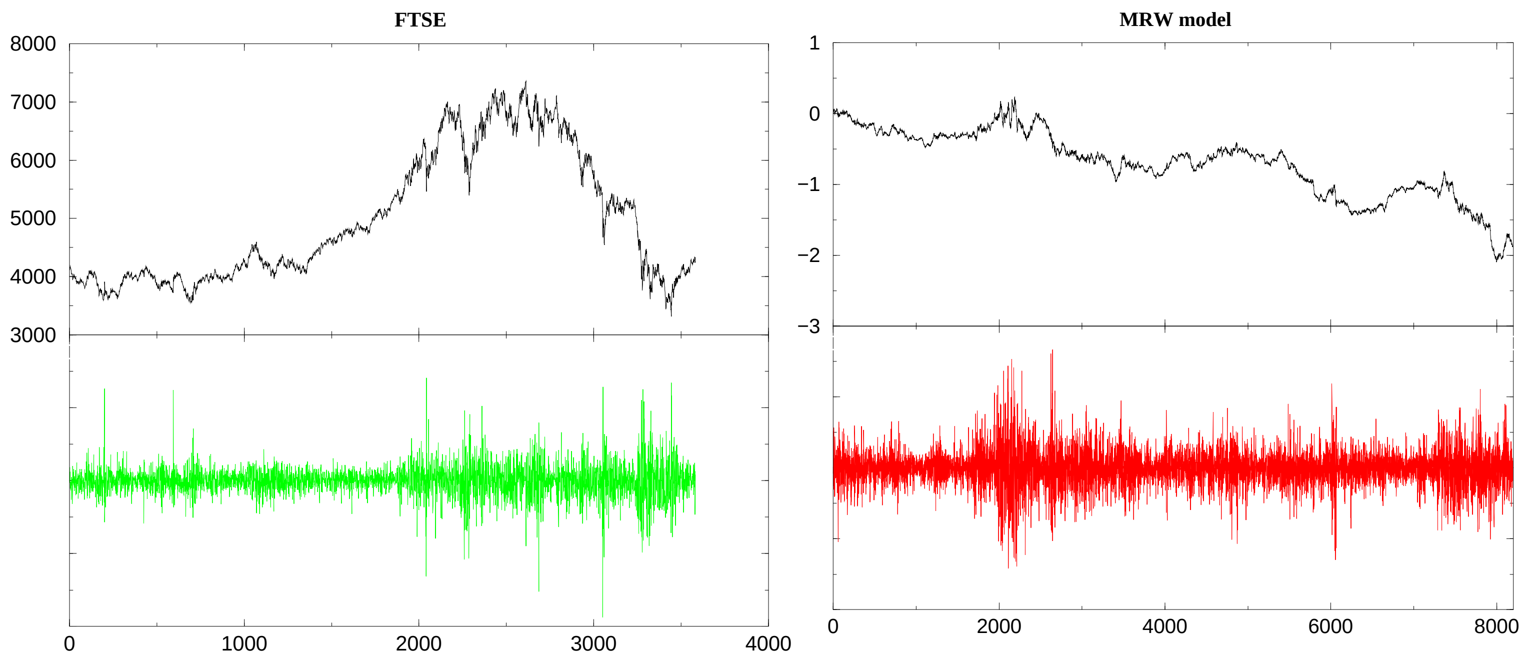
- The process is "multifractal"

$$\overline{[\ln p(t) - \ln p(0)]^{2q}} \propto t^{\zeta(q)} \qquad \zeta(q) = q - 2\lambda^2 q(q-1)$$

- The distribution of returns is power-law with $\mu = 1/\lambda^2$

- Many beautiful results and connections with random cascades, etc. (Mandelbrot and the Verdi trick)

Multifractal fluctuations



An exceptionally "good", parcimonious model:
Mandelbrot-Calvet-Fisher 1997; Bacry-Muzy 2000... – but WHY?

Fat tails and Random Matrix Theory

- Eigenvalue statistics of large real symmetric matrices with Gaussian IID elements x_{ij}:

  Wigner's semi-circle for the "bulk" of the spectrum and Tracy Widom for the largest, "extreme" eigenvalue

- What happens with power-law tails ?

- Eigenvalue density:

  $\mu > 2 \rightarrow$ Wigner semi-circle

  $\mu < 2 \rightarrow$ unbounded density with tails $\rho(\lambda) \sim \lambda^{-1-\mu}$

  <u>Note:</u> $\mu < 2$ non trivial statistics of eigenvectors (localized/delocalized) [PC,JPB]

Fat tails and Random Matrix Theory

- Largest Eigenvalue statistics ([GB,MP,JPB], [Ben Arous et al.])

  - $\mu > 4$: $\lambda_{\max} - 2 \sim N^{-2/3}$ with a Tracy-Widom distribution (max of strongly correlated variables)

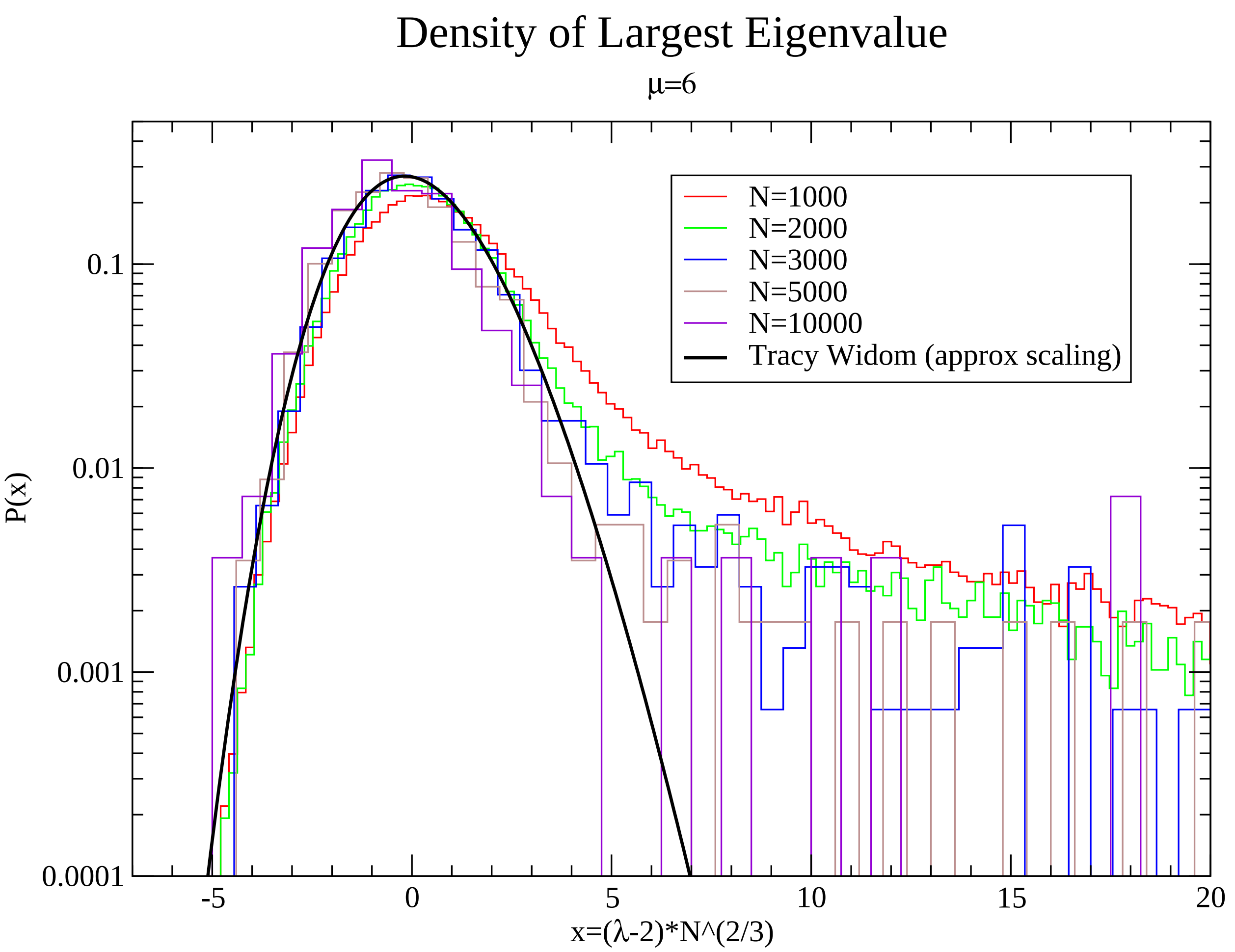
  - $2 < \mu < 4$: $\lambda_{\max} \sim N^{\frac{2}{\mu} - \frac{1}{2}}$ with a *Fréchet* distribution (although the density goes to zero when $\lambda > 2$!!)

  - $\mu = 4$: $\lambda_{\max} \geq 2$ but remains $O(1)$, with a new distribution:

$$P(\lambda_{\max}) = w\delta(\lambda_{\max} - 2) + (1 - w)F(s) \quad \lambda_{\max} = s + \frac{1}{s}$$

- Note: The case $\mu > 4$ still has a power-law tail for finite N

Density for $\mu = 6$



Correlations in complex systems

- First stage of investigation of complex systems: empirical analysis of expressed correlations

- Extension of these RMT results for correlation matrices in complex systems (e.g. finance, economics, etc.)?

Empirical correlations: basics

- Items to be correlated: $i = 1, \ldots, N$ – Realization of the observables: $t = 1, \ldots, T$

- Observed quantity: r_i^t – Examples:

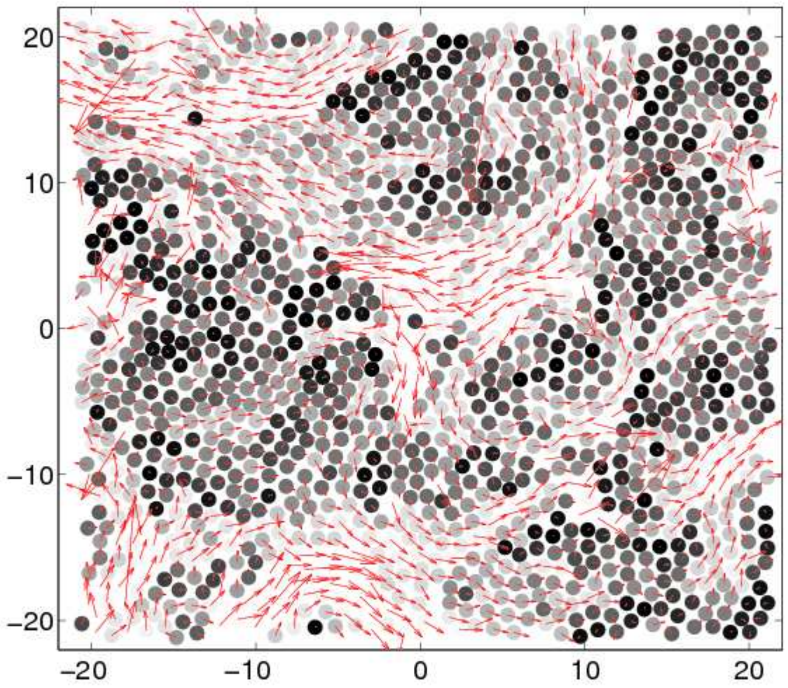
- Displacement of grain i at time t

- Return of stock i at time t

- Value of physiological parameter i for individual t

- Rating of book i by individual t

Dynamic susceptibility close to Jamming



with [Lechenault, Dauchot, Biroli]

Empirical correlations: basics

- Assume (or impose) that r_i^t are zero mean, unit variance

- The empirical correlation matrix is (usually) constructed as:

$$C_{ij} = \frac{1}{T} \sum_{t=1}^{T} r_i^t r_j^t$$

- The diagonalization of this matrix leads to 'Principal Component Analysis'

- In the framework of finance, the eigenvectors are portfolios that have uncorrelated returns, and volatilities given by the corresponding eigenvalues – crucial for portfolio optimisation

Empirical correlations: basics

- A not-so-trivial remark:

- One could have defined another correlation matrix through:

$$C_{ts} = \frac{1}{N} \sum_{i=1}^{T} r_i^t r_i^s$$

- The diagonalization of this matrix leads to the very same non zero eigenvalues (but of course different eigenvectors)

Markowitz Optimization

- Find the portfolio with maximum expected return for a given risk or equivalently, minimum risk for a given return (G)

- In matrix notation:

$$\mathbf{w}_C = G \frac{\mathbf{C}^{-1}\mathbf{g}}{\mathbf{g}^T \mathbf{C}^{-1}\mathbf{g}}$$

- Where all returns are measured with respect to the risk-free rate and $\sigma_i = 1$ (absorbed in g_i).

- Non-linear problem: $\sum_i |w_i| \leq A$ – a "spin-glass" problem!

- Related problem: find the "irreducible" idiosyncratic part of a stock

Risk of Optimized Portfolios

- Let $\mathbf{E}$ be an noisy estimator of $\mathbf{C}$ such that $\langle \mathbf{E} \rangle = \mathbf{C}$

- "In-sample" risk

$$R_{\text{in}}^2 = \mathbf{w}_E^T \mathbf{E} \mathbf{w}_E = \frac{G^2}{\mathbf{g}^T \mathbf{E}^{-1} \mathbf{g}}$$

- True minimal risk

$$R_{\text{true}}^2 = \mathbf{w}_C^T \mathbf{C} \mathbf{w}_C = \frac{G^2}{\mathbf{g}^T \mathbf{C}^{-1} \mathbf{g}}$$

- "Out-of-sample" risk

$$R_{\text{out}}^2 = \mathbf{w}_E^T \mathbf{C} \mathbf{w}_E = \frac{G^2 \mathbf{g}^T \mathbf{E}^{-1} \mathbf{C} \mathbf{E}^{-1} \mathbf{g}}{(\mathbf{g}^T \mathbf{E}^{-1} \mathbf{g})^2}$$

Risk of Optimized Portfolios

- Using convexity arguments, and for large matrices:

$$R^2_{\text{in}} \leq R^2_{\text{true}} \leq R^2_{\text{out}}$$

- Importance of eigenvalue cleaning:

$$w_i \propto \sum_{kj} \lambda_k^{-1} V_i^k V_j^k g_j = g_i + \sum_{kj} (\lambda_k^{-1} - 1) V_i^k V_j^k g_j$$

  – Eigenvectors with $\lambda > 1$ are suppressed,

  – Eigenvectors with $\lambda < 1$ are enhanced. Potentially very large weight on small eigenvalues.

  – Must determine which eigenvalues to keep and which one to correct to avoid over-allocation on pseudo-low risk modes

The large NT problem

- Determining $\mathbf{C}$ requires knowing $N(N-1)/2$ correlation coefficients. Size of data: N series of length T

- For $NT \gg N^2/2$, this should work – this is the usual limit of classical statistics: $T \to \infty$ at fixed N

- But if $NT \ll N^2/2$ there is a problem even when $T \gg 1$!

- Actually, when $T < N$, $\mathbf{E}$ has $N - T$ exact zero eigenvalues

- For $Q = T/N = O(1)$, the correlation matrix is very noisy – this is the case of 'modern' data sources (finance, economics, biology, Amazon, etc.)

The Marcenko-Pastur distribution

- Assume $\mathbf{C} \equiv \mathbf{1}$: no 'true' correlations and Gaussian r's – What is the spectrum of $\mathbf{E}$?

- Marcenko-Pastur $q = 1/Q$ – true when r's are IID with finite second moment (CLT)

$$\rho(\lambda) = (1-Q)^{+}\delta(\lambda) + \frac{\sqrt{4\lambda q - (\lambda + q - 1)^2}}{2\pi\lambda q} \qquad \lambda \in [(1-\sqrt{q})^2, (1+\sqrt{q})^2]$$

- Two sharp edges ! when $N \to \infty$

- But bleeds in a controlled way for finite N over a region of size $N^{-2/3}$ – detailed knowledge of the statistics of the *largest* eigenvalue (also Tracy-Widom when $\mu > 4$)

Fat tails and the edge of correlation matrices

- Largest eigenvalue empirical correlation matrices:

- $\mu > 4$: $\lambda_{\max} - (1 + \sqrt{q})^2 \sim \sqrt{q} N^{-2/3}$

  Tracy-Widom with a power-law tail as above

- $\mu < 4$: $\lambda_{\max} \sim N^{\frac{4}{\mu}-1} q^{1-2/\mu}$

- Fat tails induce fictitious 'strong' correlations – important for applications in finance where $\mu \approx 3 - 5$.

Multivariate statistics

- Complete description of *simultaneous* observables:

$$P(r_1^t, r_2^t, ...r_i^t, .., r_N^t)$$

- The simplest case: Gaussian multivariate

$$P(\{r_i\}) \propto \exp\left[-\frac{1}{2}\sum_{ij}\sigma_i r_i C_{ij}^{-1}\sigma_j r_j\right] \qquad (\langle r\rangle \approx 0)$$

Maximum likelihood estimator of $\mathbf{C}$ from empirical data:

$$E_{ij} = \frac{1}{T}\sum_t \hat{r}_{it}\hat{r}_{jt}$$

- Immediate problem: No heavy tails (finance)

Multivariate statistics

- A more realistic description: on a given day, all vols. are proportional $\rightarrow$ Elliptic distribution:

$$P(\{r_i\}) \propto \int \mathrm{d}s P(s) \exp\left[-\frac{s}{2}\sum_{ij} \sigma_i r_i C_{ij}^{-1} \sigma_j r_j\right] \qquad (\langle r \rangle \approx 0)$$

- Example: Student multivariate: $P(s) = s^{\mu/2-1} e^{-s} / \Gamma(\mu/2)$ with $\mu \approx 4$ Maximum likelihood estimator of $\mathbf{C}$ from empirical data:

$$E_{ij}^{*} = \frac{T+\mu}{N} \sum_t \frac{\hat{r}_{it}\hat{r}_{jt}}{\mu + \sum_{mn} \hat{r}_{mt} (E^{*-1})_{mn} \hat{r}_{nt}}$$

- When $\mu \to \infty$ for fixed T, Student becomes Gaussian and $\mathbf{E}^* = \mathbf{E}$

Multivariate non-linear correlations

- Standard tool in risk management/portfolio optimisation: the covariance matrix $\langle r_i r_j \rangle$

- Many situations require knowledge of higher order correlations
  - Gamma-risk of option portfolios: $\langle r_i^2 r_j^2 \rangle - \langle r_i^2 \rangle \langle r_j^2 \rangle$
  - Stress test of complex porfolios: correlations in extreme market conditions
  - Correlated default probabilities, etc.

Different correlation coefficients

- Correlation coefficient: $\rho_{ij} = \text{cov}(r_i, r_j)/\sqrt{V(r_i)V(r_j)}$

- Correlation of squares or absolute values:

$$\rho_{ij}^{(2)} = \frac{\text{cov}(r_i^2, r_j^2)}{\sqrt{V(r_i^2)V(r_j^2)}} \qquad \rho_{ij}^{(a)} = \frac{\text{cov}(|r_i|, |r_j|)}{\sqrt{V(|r_i|)V(|r_j|)}}$$

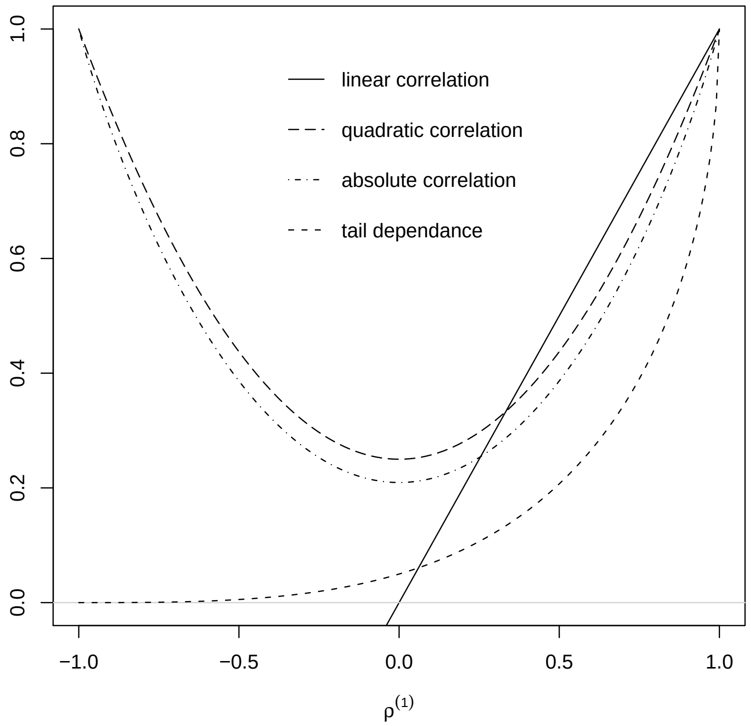
- Tail correlation:

$$\tau_{ij}^{UU}(p) = \frac{1}{p}\text{Prob.}\left[r_i > \mathcal{P}_{>,i}^{-1}(p) \bigcup r_j > \mathcal{P}_{>,j}^{-1}(p)\right]$$

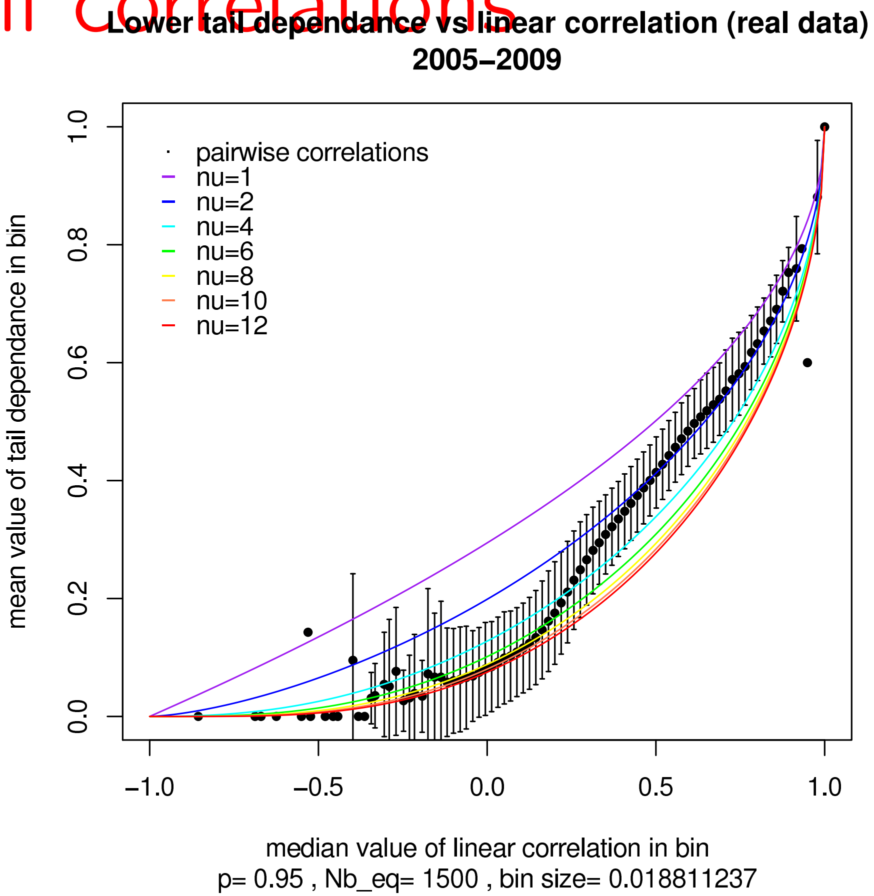
  (Similar defs. for $\tau^{LL}, \tau^{UL}, \tau^{LU}$)

- For Gaussian RV, $\tau_{ij}(p \to 0) = 0$, for any elliptic power-law RV, $\tau_{ij}(p \to 0) > 0$

Student Copulas



Extreme tail correlations



with [Ciliberti, Chicheportiche] – Student ok for highly correlated assets but fail when correlation is low

Copulas

- Sklar's theorem: any multivariate distribution can be "factorized" into
  - its marginals $\mathcal{P}_i(r_i)$
  - a "copula", that describes the correlation structure between N $U[0,1]$ standardized random variables: $c(u_1, u_2, ...u_N)$
- All correlations, linear and non linear, can be computed from the copula and the marginals

Copulas – Examples

- Examples: $(N = 2)$
  - The Gaussian copula: r_1, r_2 bivariate Gaussian $\to$ defines the Gaussian copula $c_G(u_1, u_2|\rho)$
  - The Student copula: r_1, r_2 bivariate Student with tail ν $\to$ defines the Student copula $c_S(u_1, u_2|\rho, \nu)$
  - Archimedean copulas: $\phi(u) : [0, 1] \to [0, 1]$, $\phi(1) = 0$, ϕ^{-1} decreasing, completely monotone

    $$C_A(u_1, u_2) = \int_0^{u_1} \mathrm{d}t_1 \int_0^{u_2} \mathrm{d}t_2 c_A(t_1, t_2) = \phi^{-1}\left[\phi(u_1) + \phi(u_2)\right]$$

    Ex: Frank copulas, $\phi(u) = \ln[e^{\theta} - 1] - \ln[e^{\theta u} - 1]$;
    Gumbel copulas, $\phi(u) = (-\ln u)^{\theta}$, $\theta < 1$.

The Copula red-herring

- Sklar's theorem: a nearly empty shell – almost any $c(u_1, u_2, ...u_N)$ with required properties is allowed.

- The usual financial mathematics syndrom: choose a class of copulas with convenient mathematical properties and brute force calibrate to data

- Statistical tests are not enough – intuition & plausibile interpretation are required

The Copula red-herring

- Example 1: why on earth choose the Gaussian copula to describe correlation between (positive) default times???

- Example 2: Archimedean copulas: take two $U[0,1]$ random variables s, w. Set $t = K^{-1}(w)$ with $K(t) = t - \phi(t)/\phi'(t)$.

  $$u_1 = \phi^{-1}\left[s\phi(t)\right]; \qquad u_2 = \phi^{-1}\left[(1-s)\phi(t)\right]; \qquad \longrightarrow r_1, r_2$$

  Financial interpretation ???

- Models should reflect some plausible underlying structure or mechanism – the correct multivariate models for stocks is an open problem at this stage ([Chicheportiche,JPB])

More General Correlation matrices

- Non equal time correlation matrices

$$E_{ij}^{\tau} = \frac{1}{T} \sum_t \frac{X_i^t X_j^{t+\tau}}{\sigma_i \sigma_j}$$

  $N \times N$ but not symmetrical: 'leader-lagger' relations

- General rectangular correlation matrices

$$G_{\alpha i} = \frac{1}{T} \sum_{t=1}^{T} Y_\alpha^t X_i^t$$

  N 'input' factors X; M 'output' factors Y

  – Example: $Y_\alpha^t = X_j^{t+\tau}$, $N = M$

- The large N-M-T problem! Sunspots and generalisation of Marcenko-Pastur – See later

RMT: Spectral Transforms

- Stieltjes transform, Green and Blue functions
  - $\rho(\lambda) = N^{-1} \sum_i \delta(\lambda - \lambda_i)$
  - Stieltjes transform:

$$\mathcal{S}(z) = \int \mathrm{d}\lambda \frac{\rho(\lambda)}{\lambda - z} = \frac{1}{N} \mathrm{Tr}\left[(\mathbf{H} - z\mathbf{I})^{-1} \right]$$

  - Green function:

$$G(z) \equiv -\mathcal{S}(z); \qquad \rho(\lambda) = \lim_{\epsilon \to 0} \frac{1}{\pi} \Im \left(G(\lambda - \mathrm{i}\epsilon) \right)$$

  - Blue function: $B[G(z)] = z$

RMT: Spectral Transforms

- R-transforms and S-transforms
  - R-transform: $R(z) = B(z) - z^{-1}$
  - Properties:

$$R_{aH}(z) = aR_H(az)$$

$$R(z) = \sum_{k=1}^{\infty} c_k z^{k-1} \qquad c_k : \qquad \text{Generalized cumulants}$$

  - S-transform:

$$\eta(y) \equiv -\frac{1}{y} G\left(-\frac{1}{y}\right); \qquad S(z) = -\frac{1+z}{z}\eta^{-1}(1+z)$$

RMT: Spectral Transforms

- Example 1: Wigner semi-circle

$$G(z) = \frac{z \pm \sqrt{z^2 - 4}}{2} \qquad R(z) = z$$

- Example 2: Marcenko-Pastur $Q = T/N$, $q = 1/Q$

$$\rho(\lambda) = (1-Q)^{+}\delta(\lambda) + \frac{\sqrt{4\lambda q - (\lambda + q - 1)^2}}{2\pi\lambda q} \qquad \lambda \in [(1-\sqrt{q})^2, (1+\sqrt{q})^2]$$

$$G(z) = \frac{(z + q - 1) - \sqrt{(z + q - 1)^2 - 4zq}}{2zq}, \; R(z) = \frac{1}{1 - qz}, \; S(z) = \frac{1}{1 + qz}$$

Free Random Matrices

- Freeness

  – Freeness is the generalisation of independence for matrices. Two matrices $\mathbf{A}$, $\mathbf{B}$ are said to be free essentially if the eigenvectors of $\mathbf{A}$ are a random rotation of those of $\mathbf{B}$.

  – Examples: $\mathbf{A}$, $\mathbf{B}$ sym. and fixed, $\mathbf{H}$ random GOE matrix, $\mathbf{O}$ a random rotation

$$\mathbf{A} \text{ and } \mathbf{H}; \qquad \mathbf{A} \text{ and } \mathbf{O}^{\mathrm{t}}\mathbf{B}\mathbf{O}; \qquad \mathbf{H_1} \text{ and } \mathbf{H_2}$$

  – Rectangular matrix examples: $\mathbf{A}$ $N \times T$ fixed, $\mathbf{C}$ $N \times N$ fixed, $\mathbf{H}$ $N \times T$ IID Gaussian:

$$\mathbf{A}\,\mathbf{A}^{\mathrm{t}} \text{ and } \mathbf{H}\,\mathbf{H}^{\mathrm{t}}; \qquad \mathbf{H_1}\,\mathbf{H_1^{\mathrm{t}}} \text{ and } \mathbf{H_2}\,\mathbf{H_2^{\mathrm{t}}} \qquad \mathbf{H_1}\,\mathbf{H_1^{\mathrm{t}}} \text{ and } \mathbf{C}$$

Free Random Matrices

- Two powerful composition theorems
  - If $\mathbf{A}$, $\mathbf{B}$ are sym. and free, then the spectrum of $\mathbf{A}+\mathbf{B}$ is such that:

$$R_{A+B}(z) = R_A(z) + R_B(z)$$

  - If $\mathbf{A}$, $\mathbf{B}$ are sym., non negative and free, then the spectrum of $\mathbf{AB}$ is such that:

$$S_{AB}(z) = S_A(z) S_B(z)$$

The Marcenko-Pastur distribution

- Consider the following empirical $N \times N$ correlation matrix

$$E_{ij} = \frac{1}{T} \sum_{k=1}^{T} X_i^k X_j^k \qquad \text{where} \qquad \langle X_i^k X_j^l \rangle = C_{ij} \delta_{kl}$$

- When $\mathbf{C} = \mathbf{1}$, E_{ij} is a sum of rotationally invariant projectors $(X_i^k X_j^k)/T$

$$G_k(z) = \frac{1}{N} \left(\frac{1}{z - q} + \frac{N-1}{z} \right)$$

- Inverting $G_k(z)$ to first order in $1/N$,

$$R_k(x) = \frac{1}{T(1 - qx)} \quad \text{by additivity} \quad R_E(x) = \frac{1}{(1 - qx)}$$

which is the R-transform of the MP distribution

General $\mathbf{C}$ Case

- The general case for $\mathbf{C}$ cannot be directly written as a sum of "Blue" functions.

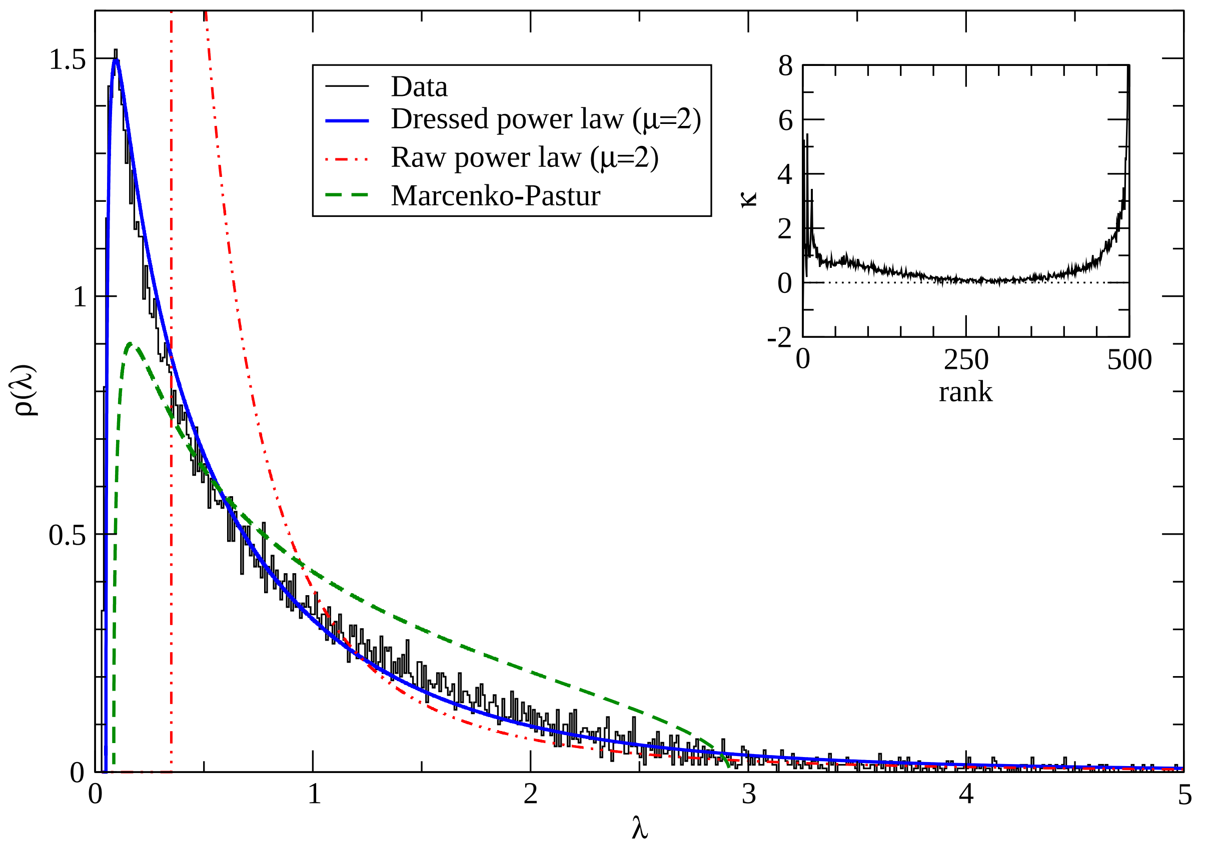
- Solution using different techniques (replicas, diagrams, S-transform:
$$G_E(z) = \int d\lambda\, \rho_C(\lambda) \frac{1}{z - \lambda(1 - q + qzG_E(z))},$$

- Remark 1: $-G_E(0) = (1-q)^{-1}$ independently of $\mathbf{C}$

  $\rightarrow R_{\text{in}} = R_{\text{true}}\sqrt{1-q} = R_{\text{out}}(1-q)$

- Remark 2: One should postulate a parametric form for $\rho_C(\lambda)$, for example:
$$\rho_C(\lambda) = \frac{\mu A}{(\lambda - \lambda_0)^{1+\mu}} \Theta(\lambda - \lambda_{\text{min}})$$

Empirical Correlation Matrix



More General Correlation matrices

- Non equal time correlation matrices

$$E_{ij}^{\tau} = \frac{1}{T}\sum_t \frac{X_i^t X_j^{t+\tau}}{\sigma_i \sigma_j}$$

$N \times N$ but not symmetrical: ‘leader-lagger’ relations

- General rectangular correlation matrices

$$G_{\alpha i} = \frac{1}{T}\sum_{t=1}^{T} Y_\alpha^t X_i^t$$

N ‘input’ factors X; M ‘output’ factors Y

– Example: $Y_\alpha^t = X_j^{t+\tau}$, $N = M$

Singular values and Canonical Components

- Singular values: Square root of the non zero eigenvalues of GG^T or G^TG, with associated eigenvectors u_α^k and $v_i^k \rightarrow$ $1 \geq s_1 > s_2 > ...s_{(M,N)^-} \geq 0$

- Interpretation: $k = 1$: best linear combination of input variables with weights v_i^1, to optimally predict the linear combination of output variables with weights u_α^1, with a cross-correlation $= s_1$.

- s_1: measure of the predictive power of the set of Xs with respect to Ys

- Other singular values: orthogonal, less predictive, linear combinations

Benchmark: no cross-correlations

- Null hypothesis: No correlations between Xs and Ys:

  $\langle G \rangle = \mathbf{0}$

- But arbitrary correlations *among* Xs, C_X, and Ys, C_Y, are possible

- Consider exact normalized principal components for the sample variables Xs and Ys:

$$\hat{X}_i^t = \frac{1}{\sqrt{\lambda_i}} \sum_j U_{ij} X_j^t; \quad \hat{Y}_\alpha^t = ...$$

  and define $\hat{G} = \hat{Y}\hat{X}^T$.

Benchmark: no cross-correlations

- Tricks:
  - Non zero eigenvalues of $\widehat{G}\widehat{G}^T$ are the same as those of $\widehat{X}^T\widehat{X}\widehat{Y}^T\widehat{Y}$
  - $A = \widehat{X}^T\widehat{X}$ and $B = \widehat{Y}^T\widehat{Y}$ are mutually free, with n (m) eigenvalues equal to 1 and $1-n$ $(1-m)$ equal to 0
  - "S-transforms" are multiplicative

Benchmark: Random SVD

- Final result:([LL,MAM,MP,JPB])

$$\rho(s) = (m+n-1)^{+}\delta(s-1) + \frac{\sqrt{(s^2-\gamma_-)(\gamma_+ - s^2)}}{\pi s(1-s^2)}$$

  with

$$\gamma_{\pm} = n + m - 2mn \pm 2\sqrt{mn(1-n)(1-m)}, \quad 0 \le \gamma_{\pm} \le 1$$

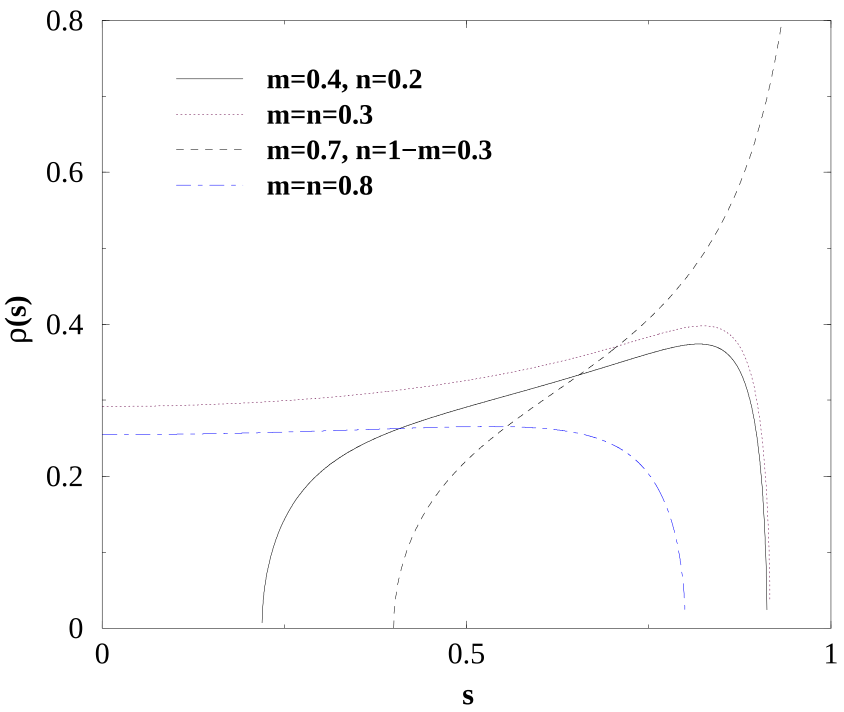
- Analogue of the Marcenko-Pastur result for rectangular correlation matrices

- Many applications; finance, econometrics (‘large’ models), genomics, etc.

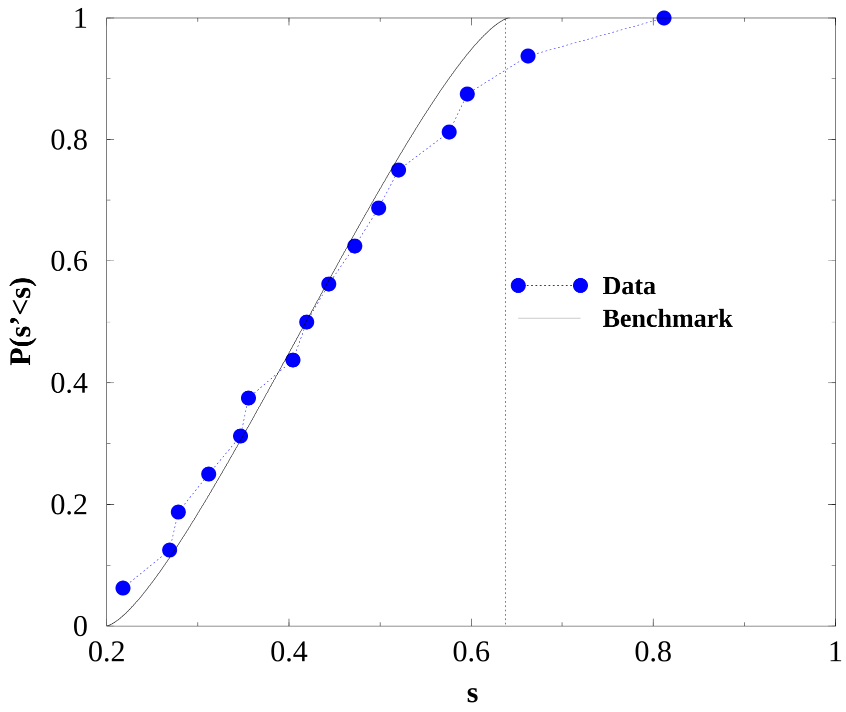
Benchmark: Random SVD

- Simple cases:
  - $n = m$, $s \in [0, 2\sqrt{n(1-n)}]$
  - $n, m \to 0$, $s \in [|\sqrt{m} - \sqrt{n}|, \sqrt{m} + \sqrt{n}]$
  - $m = 1$, $s \to \sqrt{1-n}$
  - $m \to 0$, $s \to \sqrt{n}$

RSVD: Numerical illustration



Inflation vs. Economic indicators



$N = 50, M = 16, T = 265$

Current project: Understanding eigenvectors

- The 'true' correlation matrix is probably time dependent!
- Null model: dynamics of the largest k-eigenvector subspace for a EWMA determination of a fixed correlation matrix
- For the largest eigenvector: Ornstein-Uhlenbeck processes on the unit sphere
- Explicit result for the overlap between eigenspaces for all k (with R. Allez)
- Empirical results show a faster decorrelation $\longrightarrow$ real dynamics of the eigenvectors