University of Amsterdam

Rick Quax, Drona Kandhai, Andrea Apolloni, Peter Sloot

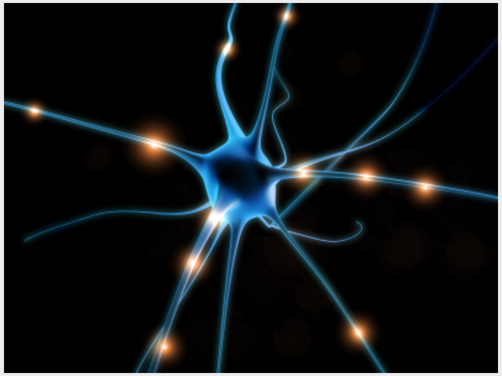
Quantifying systemic instability in networks using information dissipation

Using information theory to study emergent behavior, complexity

Our view of (complex) emergent behavior

node dynamics + interaction network = complex system

problem



+

=

Each node has a state
which it changes over time

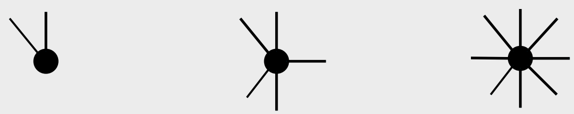
Nodes interact with each other
i.e., their states influence each other

The systemic behavior is complex
compared to an individual node

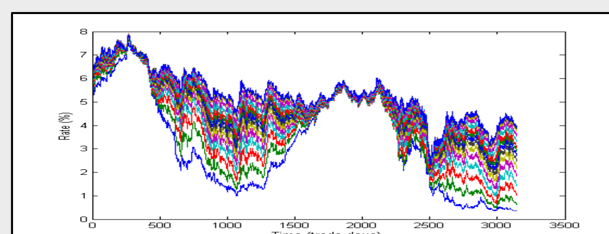
UNIVERSITY OF AMSTERDAM

Research questions

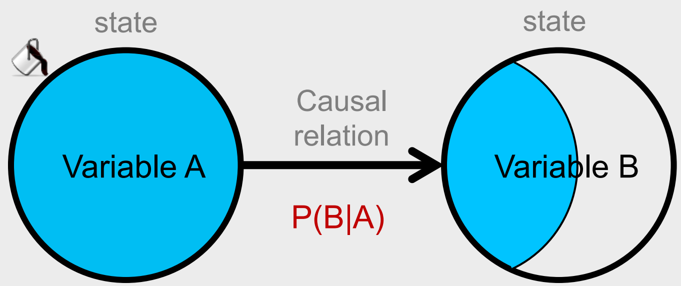
- How much ‘individual behavior’ flows into ‘systemic behavior’? i.e., how much impact has node X on system S
  - → which nodes dominate the systemic behavior?



- How resilient is a systemic behavior?
- Measuring resilience in real data, financial derivatives



First things first: how to measure 'causal impact'

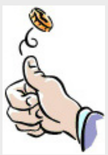


- Suppose an isolated model A→B where one stochastic variable (A) influences another (B)
- P(B|A) encodes the full causality relation
- Unfortunately, for complex systems we cannot solve for the full causality model (given only local rules)
  - This is the birth right of *complexity science*

E.g., P(neuron2|neuron1) → P(brain | neuron1)

- What if we study only *how much* influence? Not *how*?
  - Lesser aspect of full causality, hopefully more feasible
  - Need an impact measure that can handle many types of P(B|A)

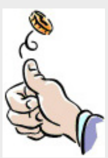
Entropy of a coin flip



$$\begin{cases} 0 & p(X=0)=0.5 \\ 1 & p(X=1)=0.5 \end{cases}$$

The outcome of the coin flip carries **1 bit** of information

I.e., I need 1 bit to fully describe the outcome of the coin flip



$$\begin{cases} 0 & p(X=0)=0 \\ 1 & p(X=1)=1 \end{cases}$$

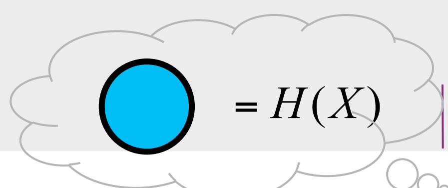
The outcome of the coin flip carries **0 bits** of information

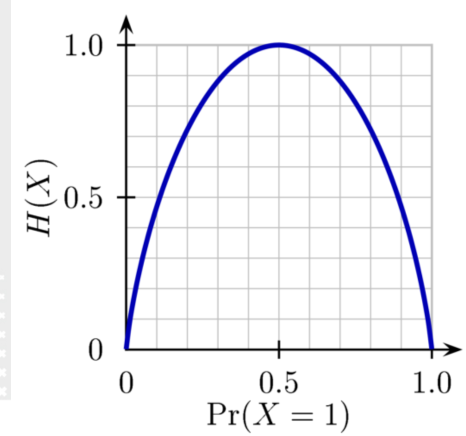
I.e., I need **0 bits** to fully describe the outcome of the coin flip
(it is already known beforehand)

In general:

Entropy of coin flip: $$H(X)=\sum_{x\in\{0,1\}} p(X=x)\cdot\log_2\frac{1}{p(X=x)}$$

Shannon's information theory





Mutual information

X

$$\begin{cases} 0 & p(X=0)=0.5 \\ 1 & p(X=1)=0.5 \end{cases}$$

Communication channel

- Noise
- (Non-linear) function
- …

Transform

$p(Y \mid X)$

Y

inference

X = **0** or **1**?

$p(X \mid Y)$

How much information was transferred? Examples:

$$\begin{cases} p(Y=x \mid X=x)=1 \\ 0 \text{ otherwise} \end{cases}$$ → **1 bit** is transferred (perfect transmission)

$$\begin{cases} p(Y=x \mid X=x)=1/2 \\ 0 \text{ otherwise} \end{cases}$$ → **0 bits** were transferred (no transmission)

Mutual information

X

$$\begin{cases} 0 & p(X=0)=0.5 \\ 1 & p(X=1)=0.5 \end{cases}$$

Communication channel

- Noise
- (Non-linear) function
- …

Transform

$p(Y \mid X)$

Y

inference

X = **0** or **1**?

$p(X \mid Y)$

How much information was transferred? **In general:**

A priori uncertainty

$$I(X:Y) = H(X) - H(X \mid Y)$$

Remaining uncertainty after knowing Y

where: $H(X \mid Y) = \sum_{y \in \{0,1\}} p(Y=y) \cdot H(X \mid Y=y)$

In direct formula: $$I(Y:X) = \sum_{x \in \{0,1\}} p(X=x) \cdot \sum_{y \in \{0,1\}} p(Y=y \mid X=x) \cdot \log \frac{p(X=x) \cdot p(Y=y \mid X=x)}{p(X=x) \cdot p(Y=y)}$$

Mutual information

X

$$\begin{cases} 0 & p(X=0)=0.5 \\ 1 & p(X=1)=0.5 \end{cases}$$

Communication channel

- Noise
- (Non-linear) function
- …

Transform

$p(Y\,|\,X)$

Y

inference

X = **0** or **1**?

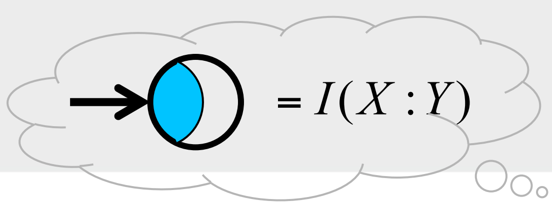
$p(X\,|\,Y)$

How much information was transferred? **In general:**

A priori uncertainty

$$I(X:Y)=H(X)-H(X\,|\,Y)$$

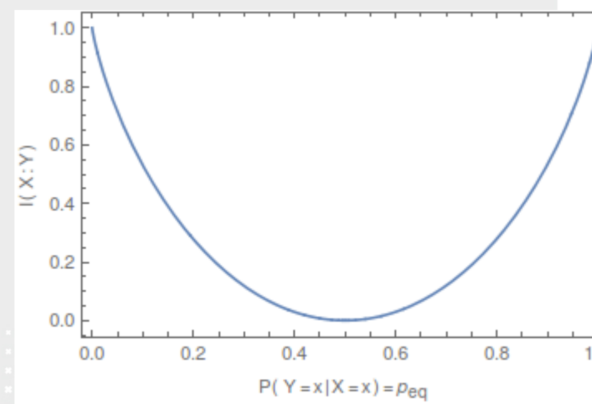
Remaining uncertainty after knowing Y

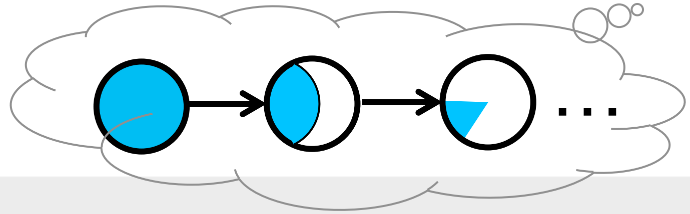


$= I(X:Y)$

$$\begin{cases} p(Y=x\,|\,X=x)=p_{\mathrm{eq}} \\ 1-p_{\mathrm{eq}} \quad \text{otherwise} \end{cases}$$

Assuming p(X=x) = 0.5



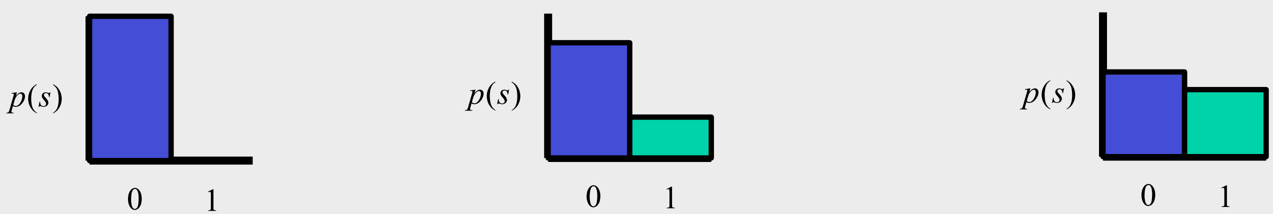


$\forall i : s_i \in \{0,1\}$

Information flow: 1D sequence of variables

$s_1 \xrightarrow{\alpha} s_2 \xrightarrow{\alpha} s_3 \quad \ldots$

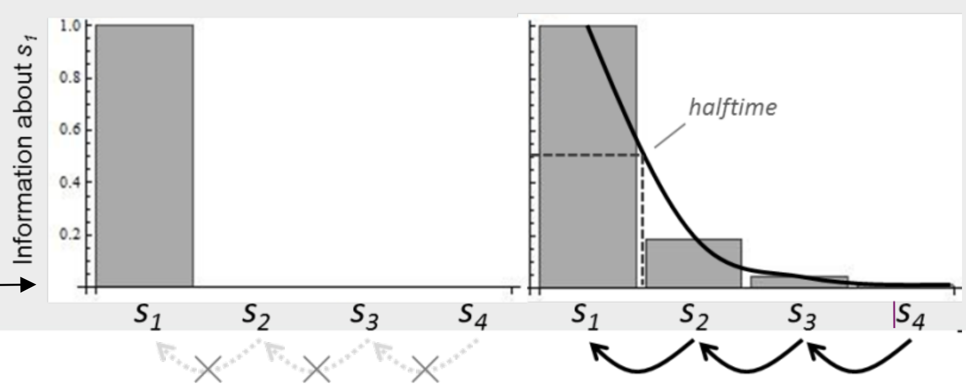
$P(s_1) \longrightarrow P(s_2 \mid s_1) \longrightarrow P(s_3 \mid s_2) \quad \ldots$



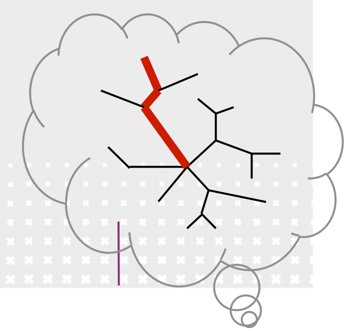
$I(x_i : x_{i-1}) = H(x_i) - H(x_i \mid x_{i-1}):$

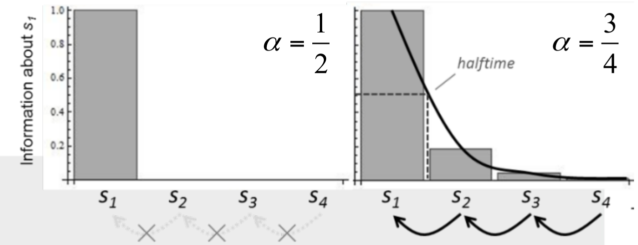
$P(s_2 \mid s_1) \equiv P(s_2 = s_1) = \alpha$

$\alpha = \frac{1}{2} \longrightarrow$



$\longleftarrow \alpha = \frac{3}{4}$





Information dissipation length (or time)

- **How far does the information flow? (before it dissipates)**
- **Measure of distance of causal impact!**

I expect a characteristic decay rate of $1/f$, because all s_x are equivalent

Find an expression for $p(s_1 \mid s_x)$:

$$p(s_1 \mid s_x) = p(s_1 = s_x) = \frac{1+(1-2\alpha)^{x-1}}{2}$$

Then rename $q \equiv 1-2\alpha$ to fit on a slide.

Write down the decay rate at distance x :

$$f = \frac{I(s_1 \mid s_x)}{I(s_1 \mid s_{x+1})} = \frac{H(s_1)-H(s_1 \mid s_x)}{H(s_1)-H(s_1 \mid s_{x+1})}$$

$$= \frac{1+\frac{\left(q+q^x\right)}{2q}\log_2\left[\frac{q+q^x}{2q}\right]+\left(1-\frac{q+q^x}{2q}\right)\log_2\left[1-\frac{q+q^x}{2q}\right]}{1+\frac{\left(q+q^{1+x}\right)}{2q}\log_2\left[\frac{q+q^{1+x}}{2q}\right]+\left(1-\frac{q+q^{1+x}}{2q}\right)\log_2\left[1-\frac{q+q^{1+x}}{2q}\right]}.$$

Take limit of f as $x \to \infty$ (L'Hôpital's rule):

$$\lim_{x\to\infty} f = \lim_{x\to\infty}\frac{\frac{d\ldots}{dx}}{\frac{d\ldots}{dx}} = \ldots = \lim_{x\to\infty}\frac{\left(\log_2\left[\frac{1}{2}-\frac{q^{-1+x}}{2}\right]-\log_2\left[\frac{q+q^x}{2q}\right]\right)}{q\left(\log_2\left[1-q^x\right]-\log_2\left[1+q^x\right]\right)}.$$

Again have to apply L'Hôpital's rule:

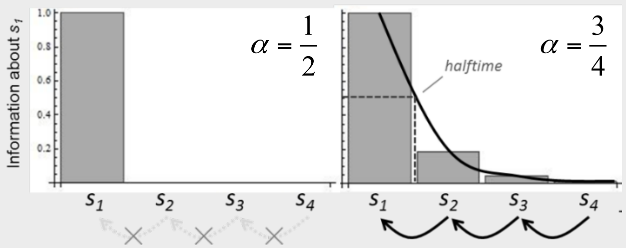
$$\lim_{x\to\infty} f = \lim_{x\to\infty}\frac{-\frac{q^{x-1}\log q}{1-q^{-1+x}}-\frac{q^{x-1}\log q}{1+q^{-1+x}}}{-\frac{q^{1+x}\log q}{1-q^x}-\frac{q^{1+x}\log q}{1+q^x}} = \ldots = \lim_{x\to\infty}\frac{q^{2x}-1}{q^{2x}-q^2}.$$

Use -1<q<1 and substitute back:

$$\lim_{x\to\infty} f = \frac{1}{\left(2a-1\right)^2} \equiv f$$

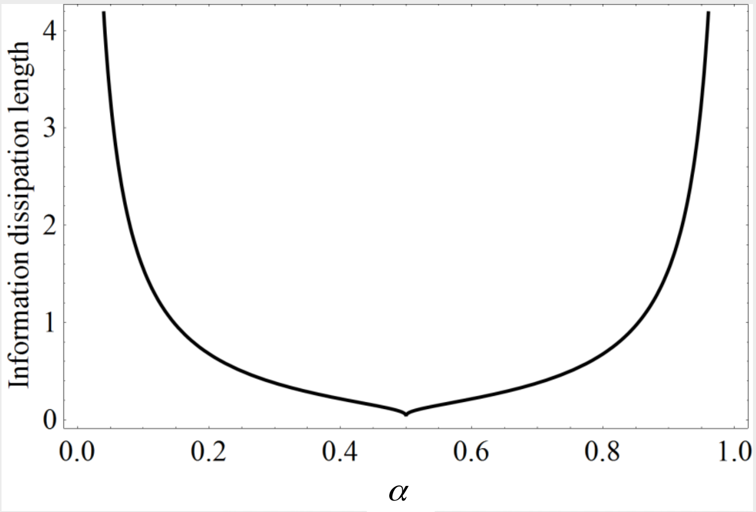
Information dissipation length (or time)

Now we have the rate. Let us define the IDL (IDT) as the distance (time) it takes for information to reach 50% of its original value

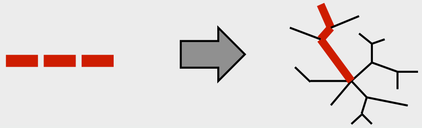


$$\mathrm{IDL} = \log_{1/r} \frac{1}{2} = \frac{-1}{\log_2 (1-2\alpha)^2} \xrightarrow{p=\frac{3}{4}} \frac{1}{2}.$$

It looks like:



Great! Let's go from 1D to a network!

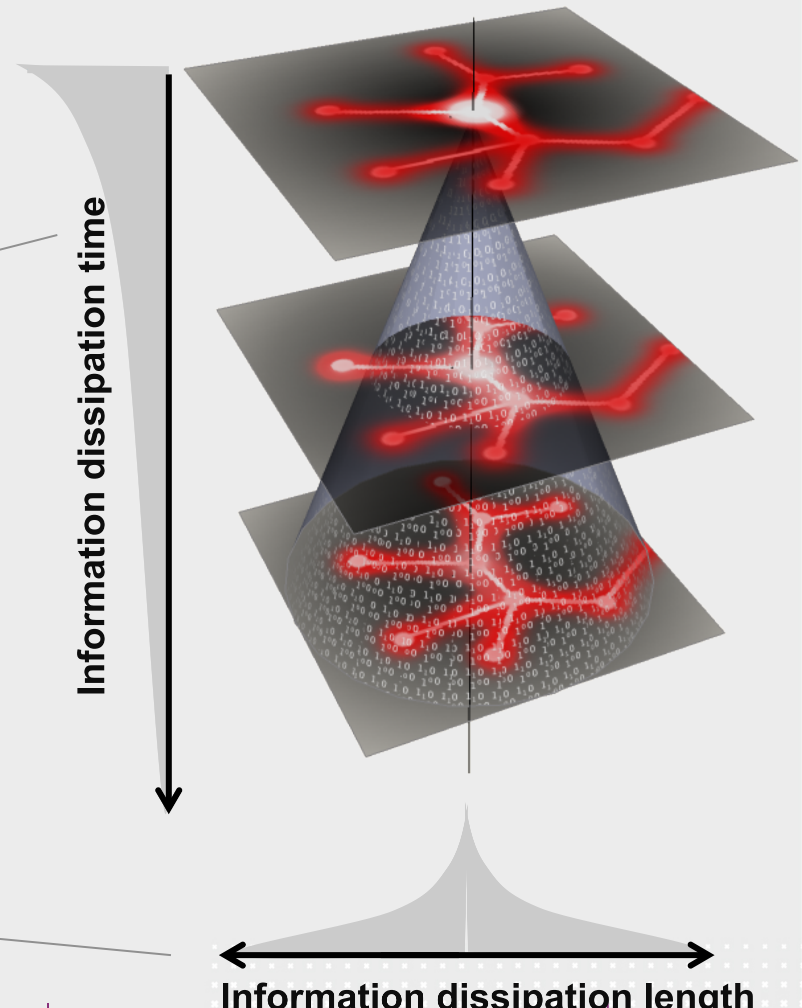


Information dissipation

How long is the information about a node's state retained in the network?

measures of influence of a single node on the behavior of the entire network!

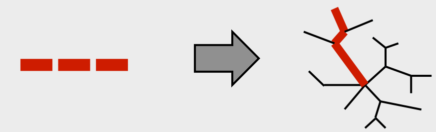
How far can the information about a node's state reach before it is lost?



Interface
The diminishing role of hubs in dynamical processes on complex networks

Rick Quax[1†], Andrea Apolloni[2,†] and Peter M. A. Sloot[1,3,4]

Information dissipation time

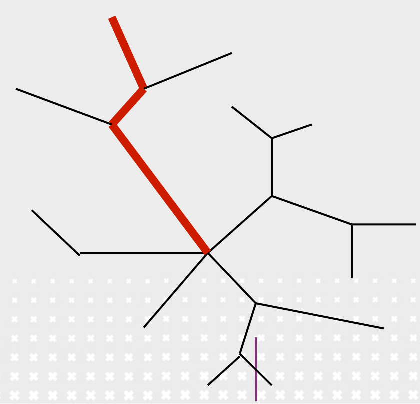


- Now we compute the IDT of each node in a network
  - Intuitively, approximate a network as a set of 1D variable sequences
- Edges represent an interaction potential to which a node can quasi-equilibrate
  - → Node dynamics: (local) Gibbs measure: $p(s_i^{t+1} = x \mid s_j^t, \ldots) \propto \exp \sum_j -E(x, s_j^t)$
  - Can be seen as generalized Ising spin model

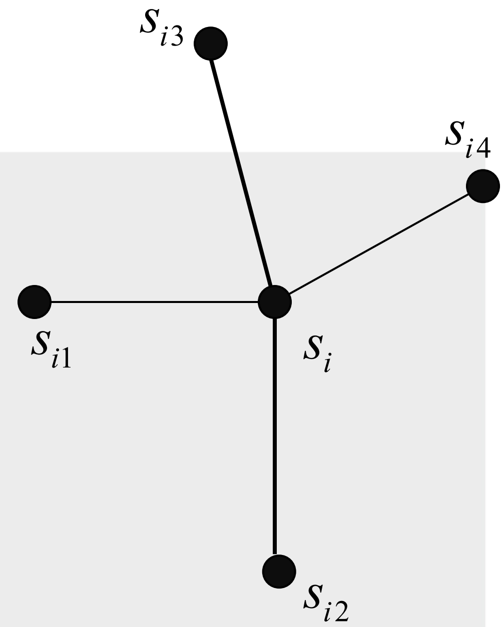
Generalized energy function

- Network structure
  - Locally tree-like (i.e., no short loops)
    - E.g., large and

      no community-structure / modularity
    - Any degree distribution can be chosen



Information dissipation time



$$I_0^k \equiv I(S^t; s_i^t) = I(s_i^t; s_i^t) = H(s_i^t)$$

$$I_1^k \approx I([s_{j1}^t, \ldots, s_{jk}^t]; s_i^t)$$

$$\hat{I} = \sum_m q(m) \cdot I_1^{m+1} / I_0^{m+1}.$$

$$D(s) = \log_{c_{\text{eff}} \cdot \hat{I}} \left[\frac{\varepsilon}{I_1^k} \right] = \frac{\log \varepsilon - \log I_1^k}{\log c_{\text{eff}} + \log \hat{I}}.$$

$$D(s) \propto \text{const} + \log I_1^k,$$

$$I_1^k = U(k) \cdot k \cdot T(k), \text{ where}$$

$$T(k) = \left\langle I(s_j^{t+1}; s_i^t) \right\rangle_{k_j},$$

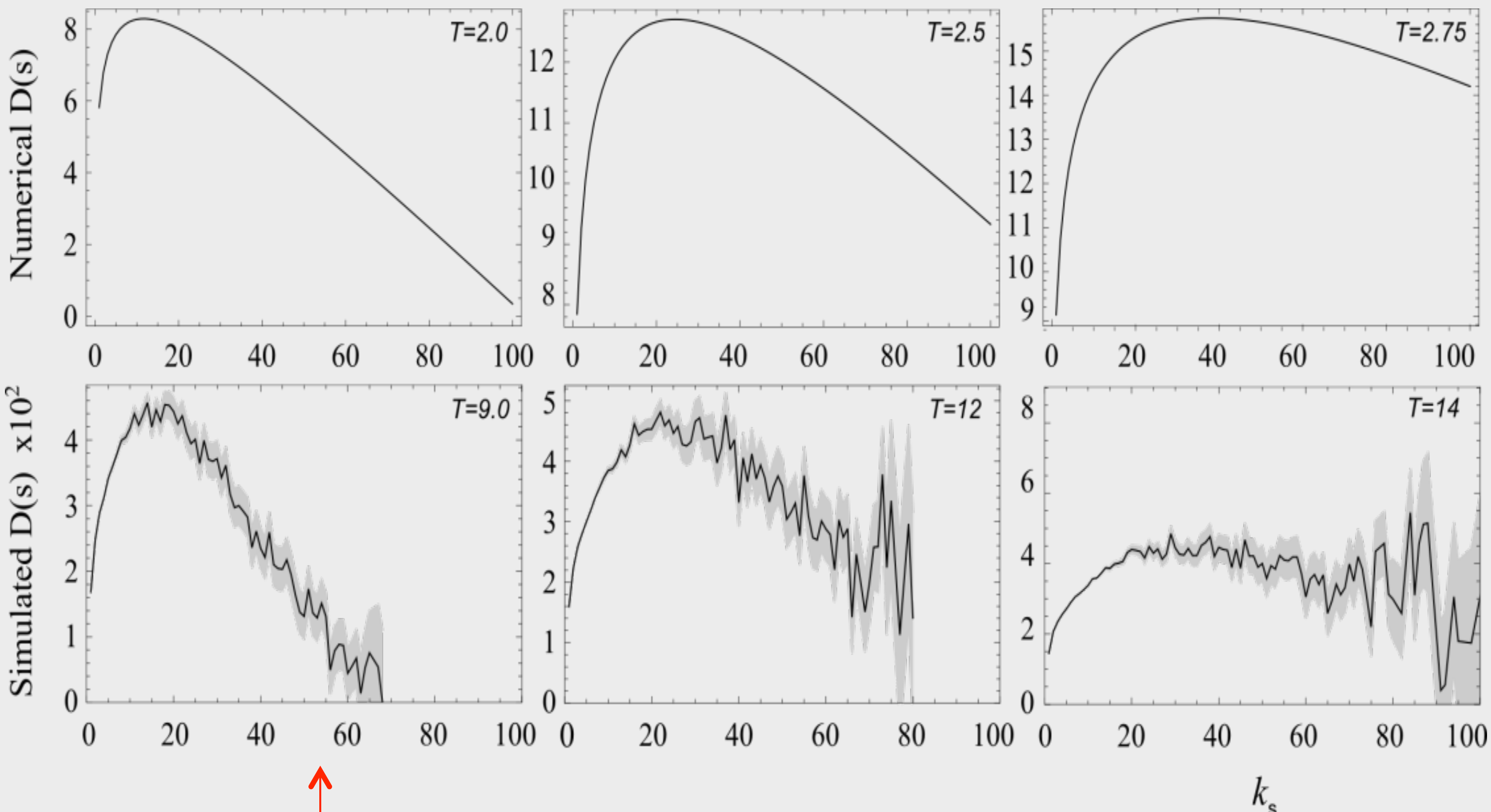
$$T(k) = \left\langle H(s_i^t) - H(s_i^t \mid s_j^{t+1}) \right\rangle_{k_j}.$$

$$\begin{aligned} H(s_i^t) &= -\sum_{q \in \Sigma} p(s_i^t = q) \log p(s_i^t = q) \\ &= -\sum_{q \in \Sigma^+} (1 - b_q{}^{-k}) \log(1 - b_q{}^{-k}) - \sum_{q \in \Sigma^-} b_q{}^{-k} \log b_q{}^{-k} \\ &= -\sum_{q \in \Sigma^+} (1 - b_q{}^{-k}) \log(1 - b_q{}^{-k}) + k \sum_{q \in \Sigma^-} b_q{}^{-k} \log b_q \\ &\approx k \sum_{q \in \Sigma^-} b_q{}^{-k} \log b_q \\ &= O(k \cdot x^{-k}). \end{aligned}$$

$$T(k+1) = \alpha \cdot T(k), \text{ where } \alpha \leq 1.$$

Results: analytical and numerical

Information dissipation time D(*s*) of a node *s*



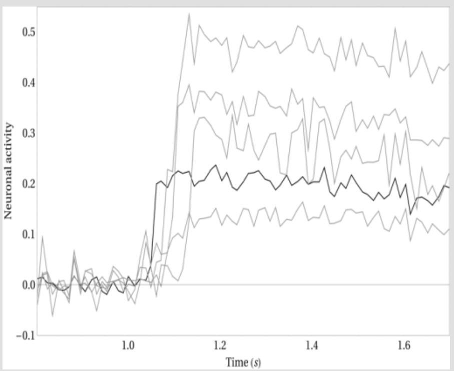
Number of interactions of a node

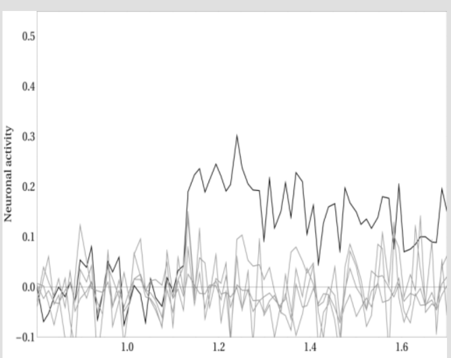
analytical proof: D(*s*) will eventually be a decreasing function of k_s

Qualitative evidence from experiments

Network of neurons cultured in a Petri dish

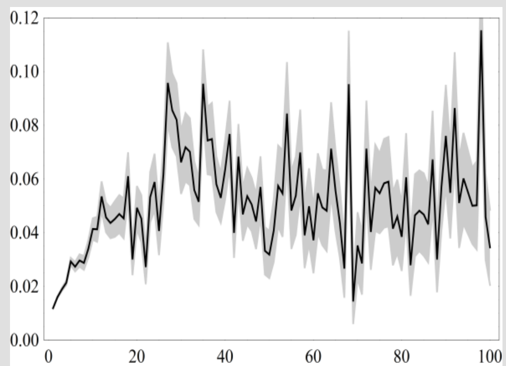
#1

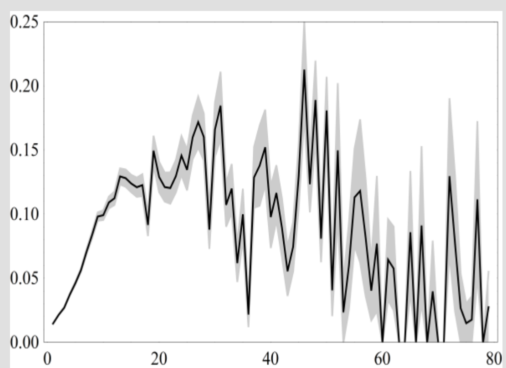




Social network of word-of-mouth marketing

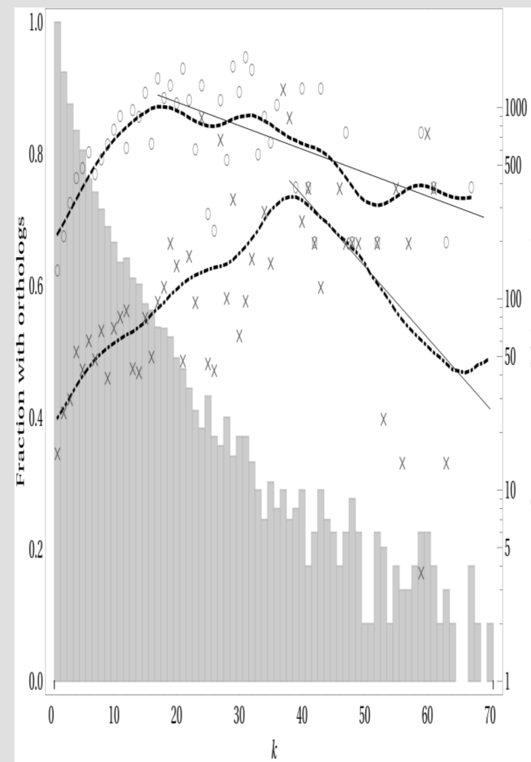
#2



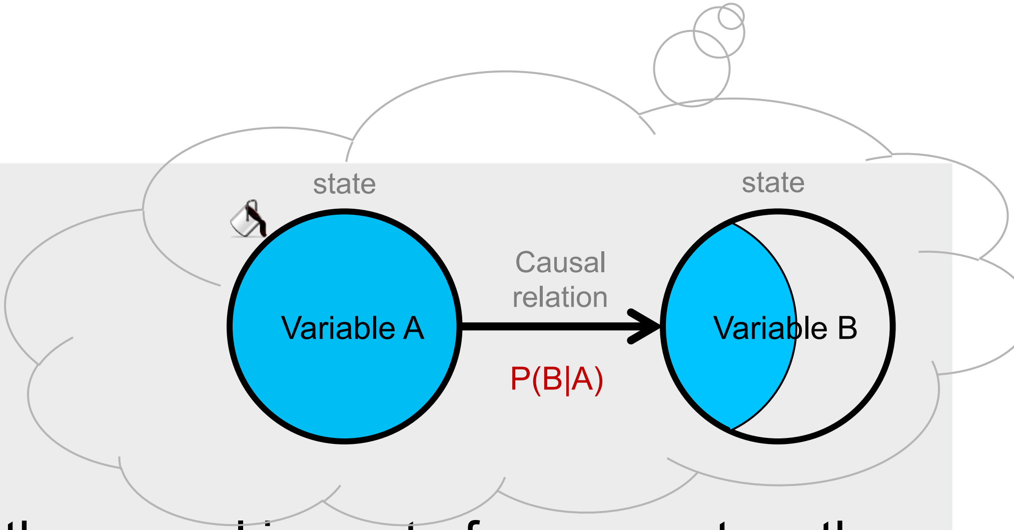


Gene regulation network

#3

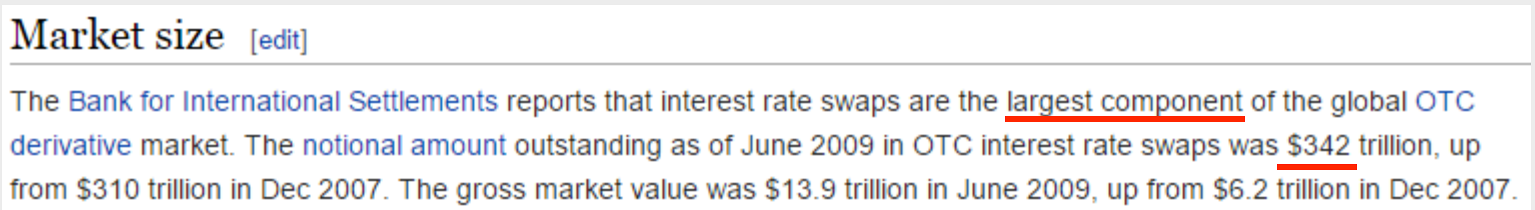


Susceptibility, systemic resilience



- Suppose now I can quantify the causal impact of one agent on the entire networked system
  - Duration of causal impact ~ IDT
  - Distance of causal impact ~ IDL
- This leads to a measure of *susceptibility*, i.e., resilience!
  - The higher the IDT of a single node, the more impact has a small (local) perturbation on the entire system
  - Financial markets? Systemic resilience?
  - Can we estimate it from real data? Lehman Brothers collapse?

Interest-rate swaps (IRS) market

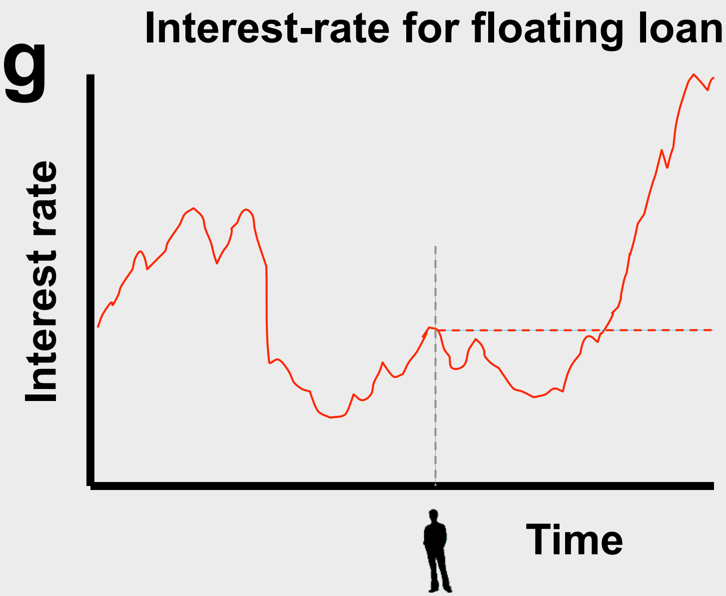

Market size [edit]

The Bank for International Settlements reports that interest rate swaps are the largest component of the global OTC derivative market. The notional amount outstanding as of June 2009 in OTC interest rate swaps was $342 trillion, up from $310 trillion in Dec 2007. The gross market value was $13.9 trillion in June 2009, up from $6.2 trillion in Dec 2007.

- In 2008, Lehman Brothers collapsed, ushering in a decade of crisis
- The IRS market is the largest financial derivatives market
  - Basically, agents trade *risk* with other agents (!)
- My colleague Dr. Drona Kandhai works at ING (Risk department)
  - Dataset: a set of timeseries of rates of different types of IRS
    - I have NO network structure, and NO data per agent… And I will never get it either!
  - Question: can we still estimate the susceptibility?

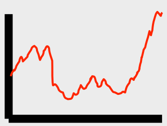
Introduction: risk trading

mortgage?



- If a person gets a € 200,000 mortgage at 5% interest rate that is fixed for 10 years…

→ **The bank runs a risk**

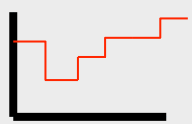


borrow

€ 350,000
at **3%,** fixed
daily

lend

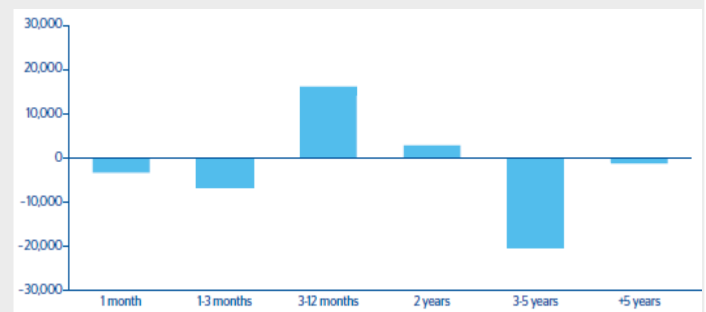
€ 200,000
at **5%,** fixed
for **10** years



lend

€ 150,000
at **4%**, fixed
yearly

Risk profile



Hedge risk → swap

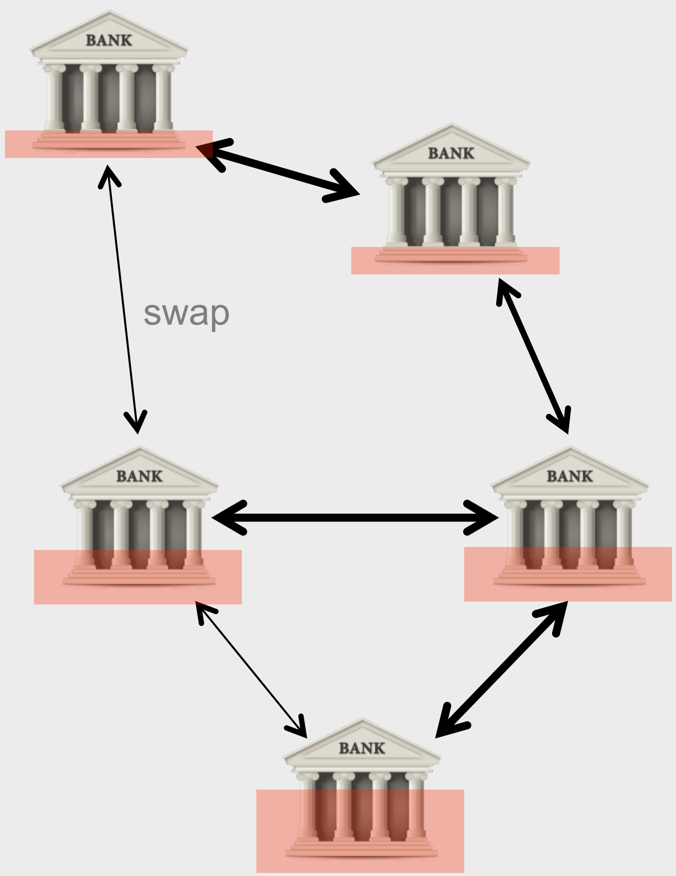
Not only banks! Think of…

- Insurance companies
- Pension funds
- Car dealerships
- …
- (Vestia…)

Interest rate swap → zero risk

(And still make money)

Idea: IRS market → propagating failure



A bank's position in this network of swap contracts makes its risk (almost) zero

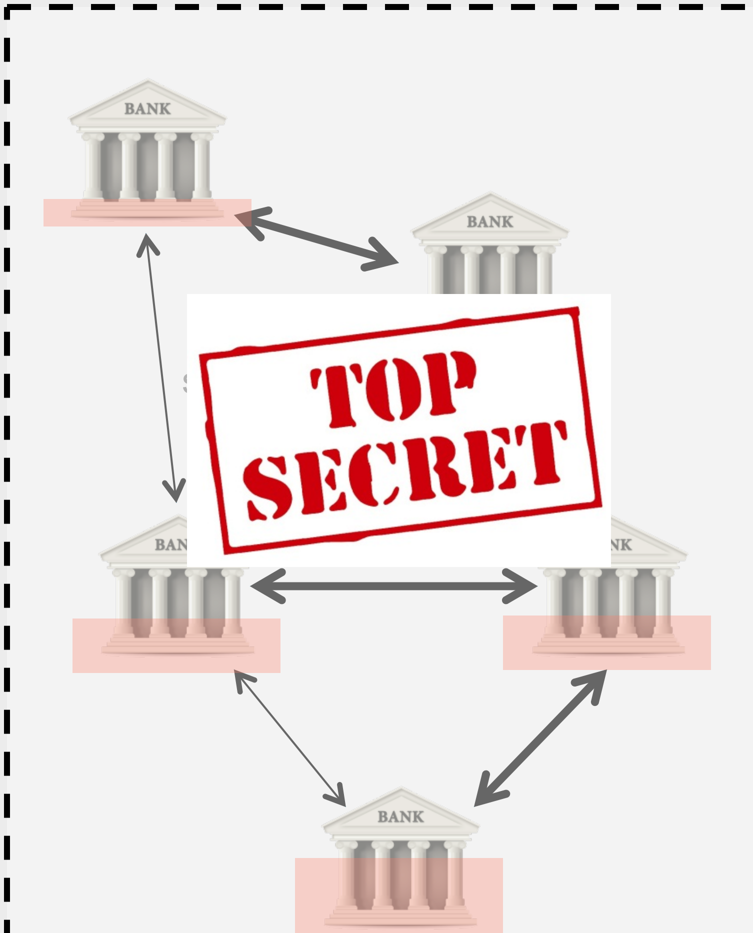
⬇

If one bank gets into trouble, then its neighbors get into trouble to some extent…

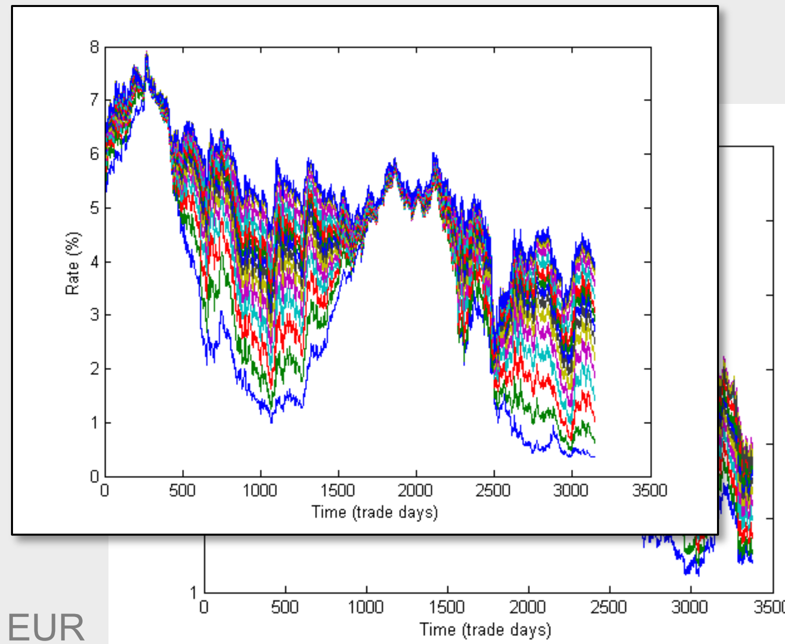
⬇

…And then the neighbors-of-neighbors get into trouble to some lesser extent, and so on

Solution: use ‘public’ data: daily IRS rates



USD



EUR

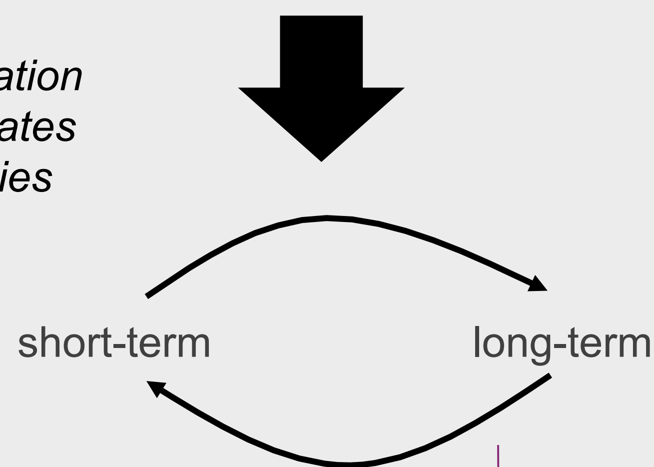
- Swap rates, daily average
- for maturities 1, 2, …, 15, 30
- Data’s time span: 12 years
- Both USD and EUR market

Inferring instability from the data

#1

The heterogeneity of maturities in the market…

…creates a correlation between the IRS rates of different maturities



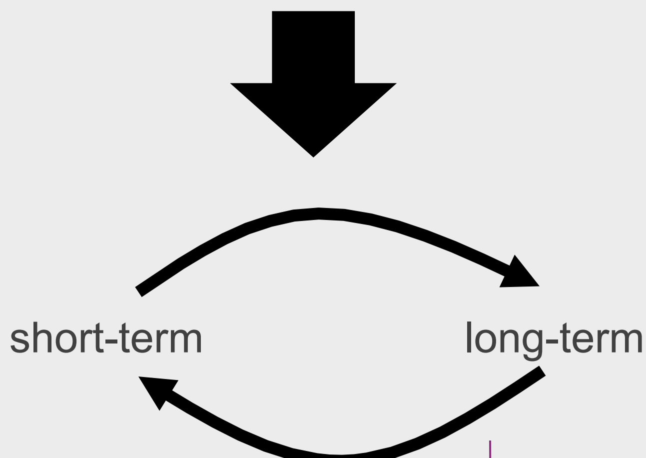
This is our first postulate

Inferring instability from the data

#1

…and the more swaps are traded…

…the stronger this cross-maturity correlation…



This is our first postulate

Inferring instability from the data

#2

swap contract **X years** ≈ swap contract **X-1 years** + swap contract **1 years**

+ forward fee for 1-IRS being X-1 years into future

…and so on…

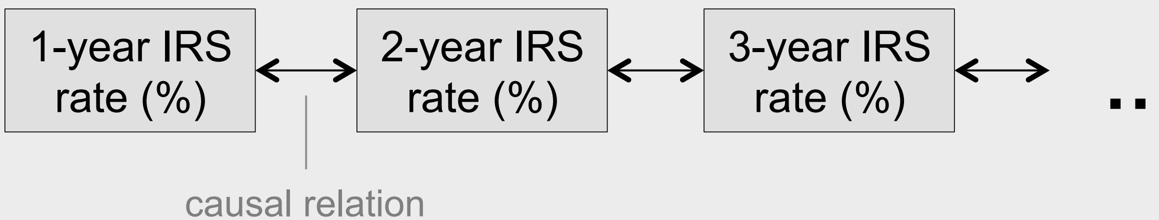
swap contract **X-1 years** ≈ swap contract **X-2 years** + swap contract **1 years**

+ similar fee…

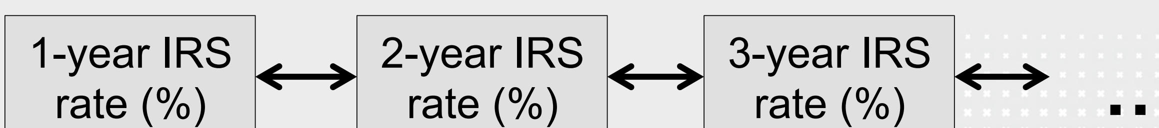
This is our second postulate

Consequence of #1 and #2: 1-dim. system

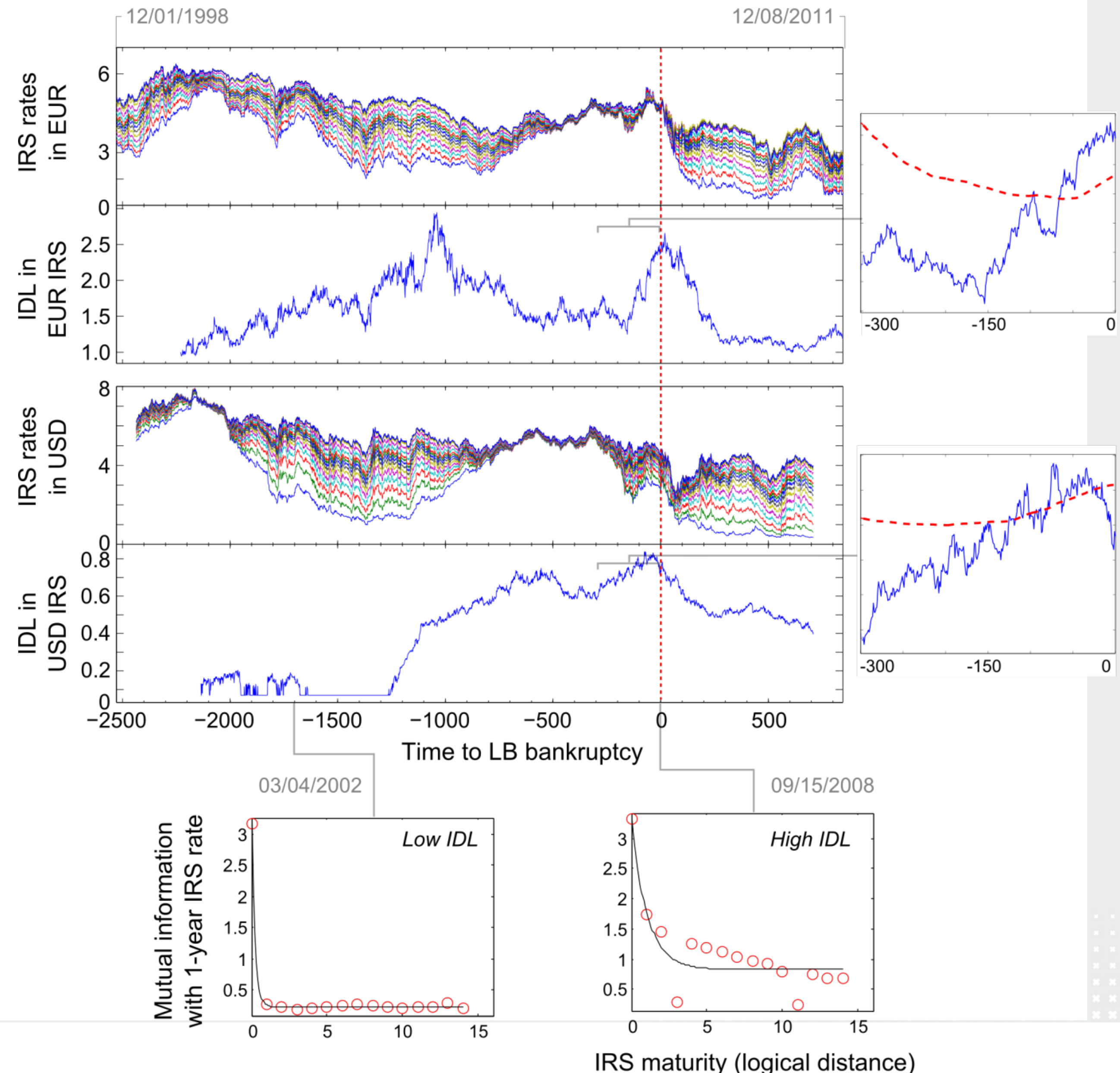
- The rates (%) of the IRS of different maturities form a 1D system:



- The more dense is the ‘swap network’ → the stronger the causal relations between the maturities → the farther a perturbation can travel through this 1D-system of rates



Results

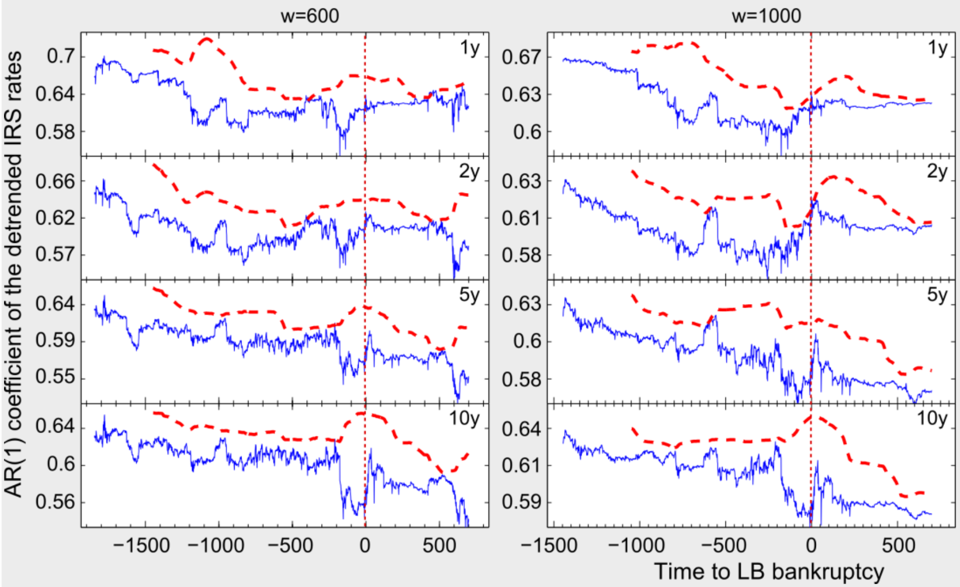


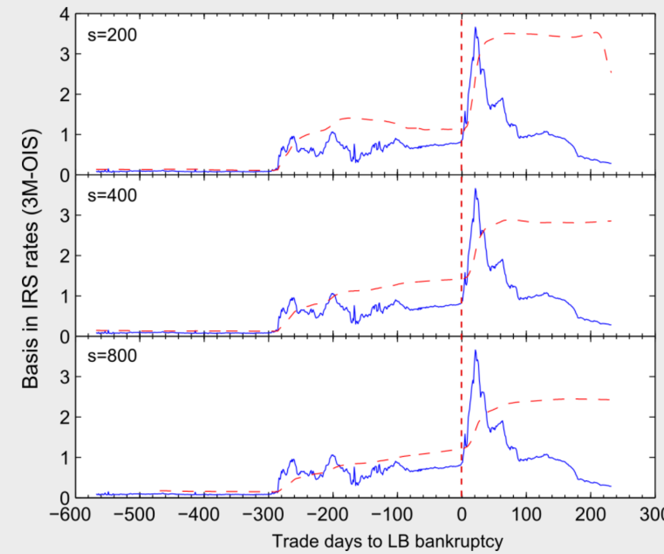
12/01/1998
12/08/2011
IRS rates in EUR
IDL in EUR IRS
IRS rates in USD
IDL in USD IRS
Time to LB bankruptcy
-2500 -2000 -1500 -1000 -500 0 500
-300 -150 0
03/04/2002
09/15/2008
Mutual information with 1-year IRS rate
Low IDL
High IDL
IRS maturity (logical distance)

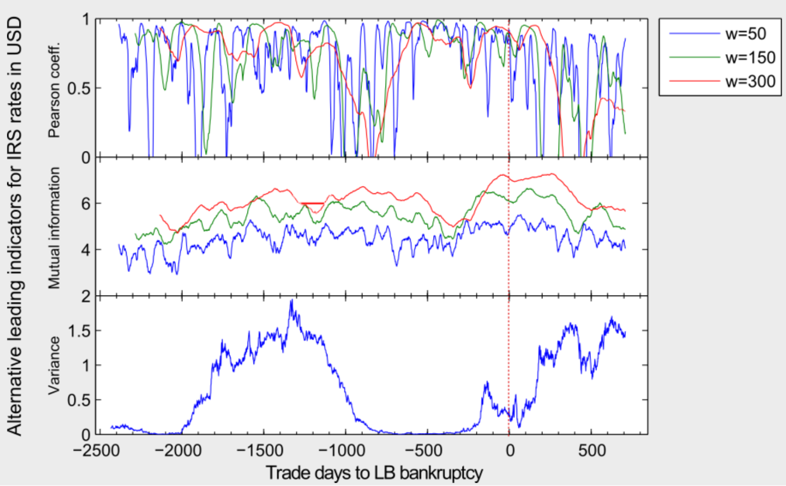
UNIVERSITY OF AMSTERDAM

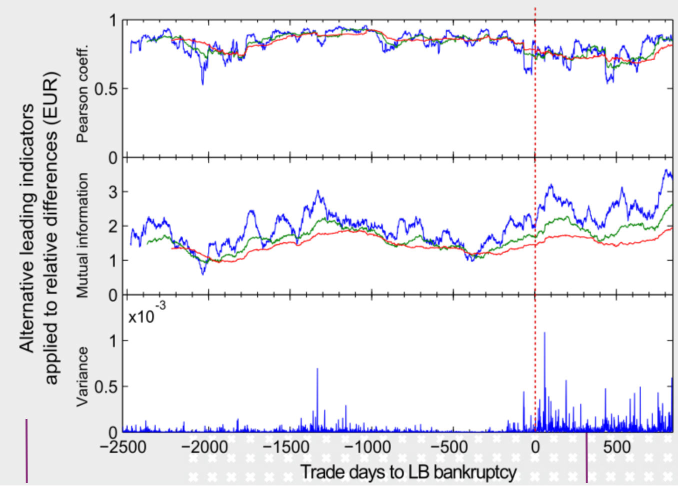
consequence, not cause

Results: well-known indicators fail









Acknowledgements

EU FP7 projects:

DynaNets (finished)

Sophocles

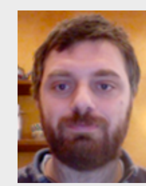
topdrim

Collaborators:

Prof. Peter Sloot

Emiliano Mancini

Andrea Apolloni



Drona Kandhai