Approximate and Exact Boltzmann Machine Learning for the US Stock Market

Stanislav Borysov, Yasser Roudi and Alexander Balatsky

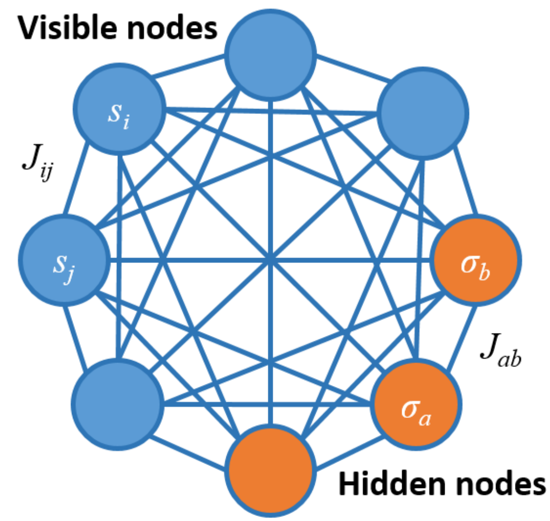
@ Computational Challenges in Nuclear and Many-Body Physics 2014

NORDITA, Stockholm, Sweden
Royal Institute of Technology (KTH), Stockholm, Sweden
The Kavli Institue for Systems Neuroscience, NTNU, Trondheim, Norway
Los Alamos National Laboratory, New Mexico, USA

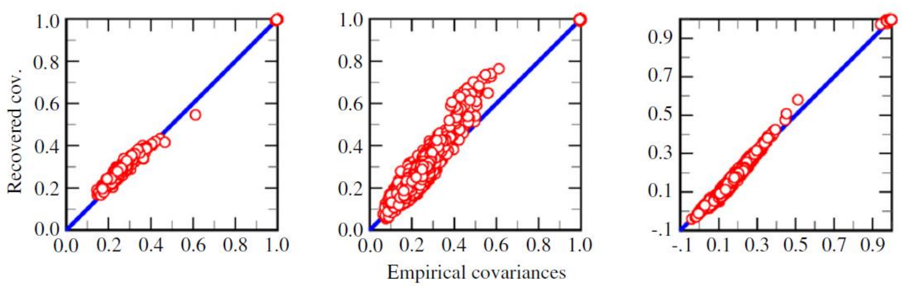
1. Overview

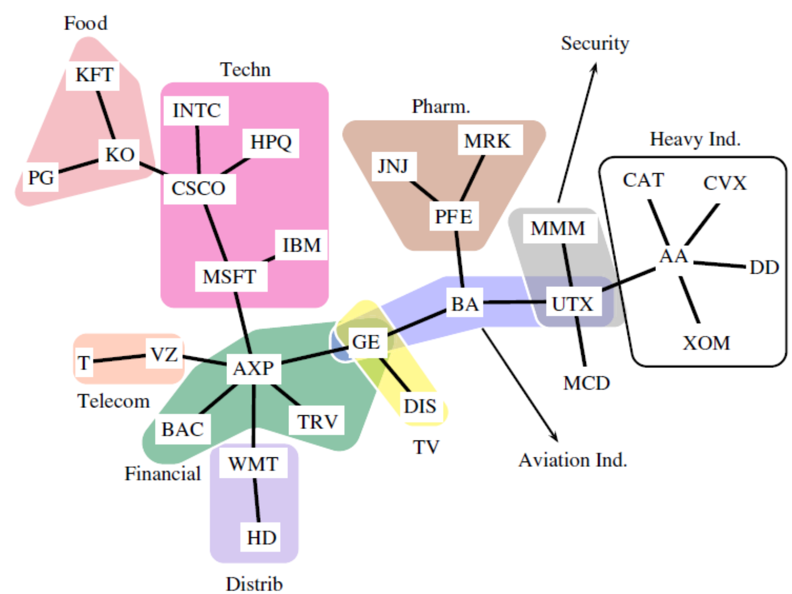
Pairwise interaction models:

1. **Ising models**
2. **Spin glasses**
3. **Neuroscience**
4. **Finance**
5. **Machine learning (Boltzmann machines)**



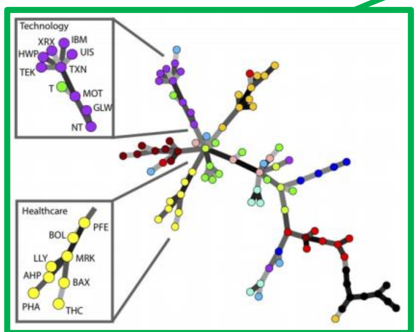
In Finance

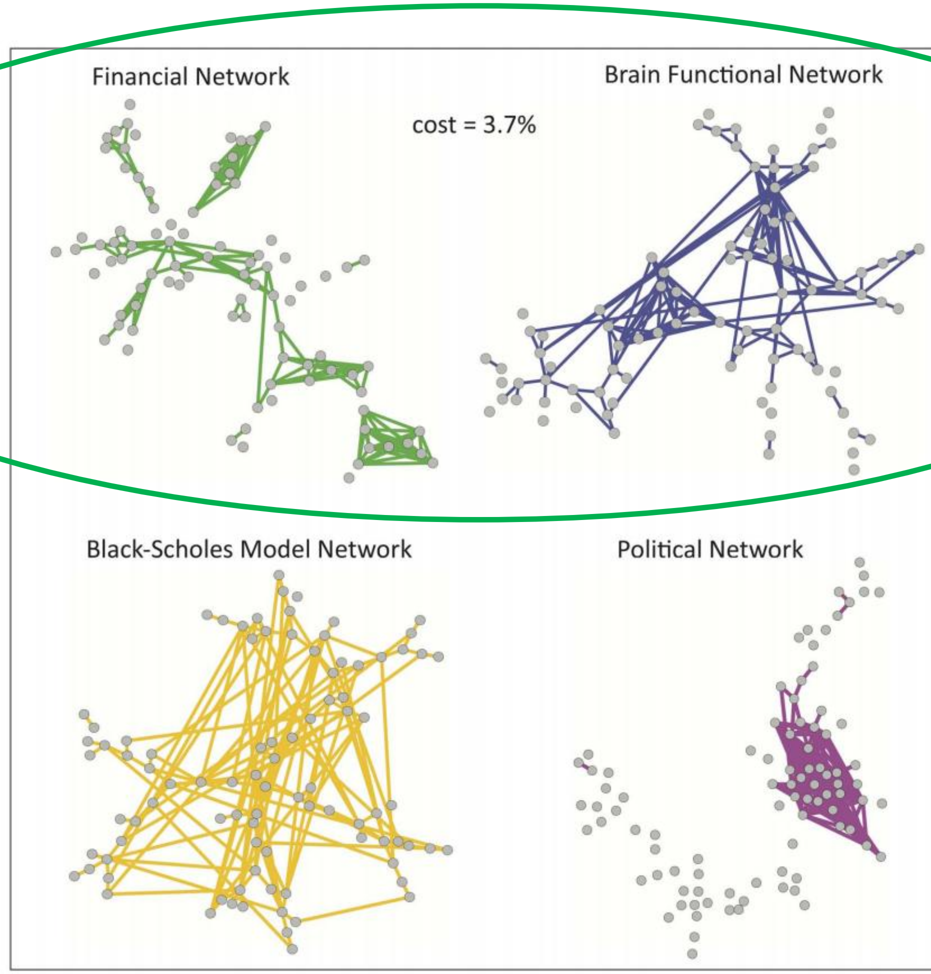


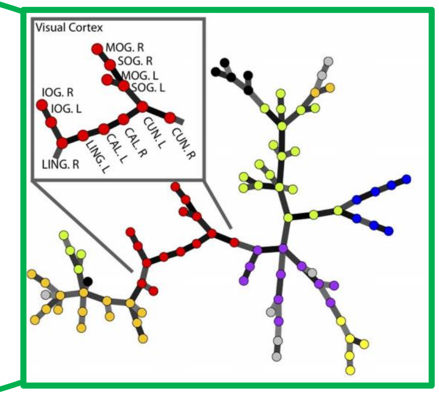


[T. Bury, Eur. Phys. J. B 86, 89 (2013)]
[T. Bury, Physica A 392, 1375 (2013)]
[H. Zeng et al, arXiv:1311.3871v1 (2013)]

Neural vs. financial networks





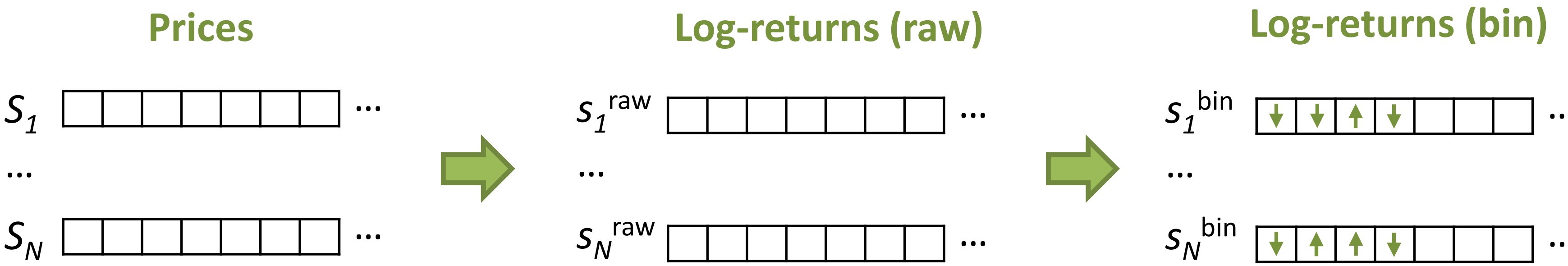


Similar:
Non-random, small-world, modular, hierarchical, fat-tailed degree distribution

However financial networks:
1. More efficient and more modular (better information processing)
2. Less robust to disintegration (“to big to fail” nodes)

[Vértes PE, et al, Front. Syst. Neurosci. 5:75 (2011)]

2. Definitions

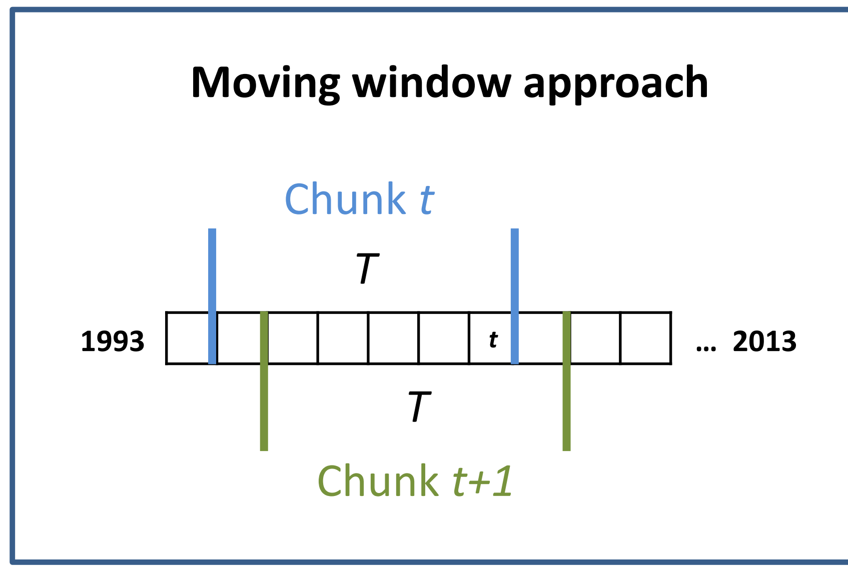


Raw log-returns

$$s_i^{\mathrm{raw}}(t) = \ln[S_i(t)/S_i(t-1)]$$

Binarized log-returns

$$s_i^{\mathrm{bin}} = \mathrm{sign}(s_i^{\mathrm{raw}})$$



Simple Moving Average (SMA) at time t for window T

$$\langle s_i \rangle := \frac{1}{T} \sum_{t=0}^{T-1} s_i(t)$$

Covariance matrix

$$C_{ij} \equiv \sigma_{ij} = \langle s_i s_j \rangle - \langle s_i \rangle \langle s_j \rangle$$

Variance

$$\sigma_i^2 \equiv \sigma_{ii}$$

Correlation matrix

$$Q_{ij} = \sigma_{ij} / \sigma_i \sigma_j$$

Skewness

$$\gamma_{1i} := \left\langle \left(\frac{s_i - \langle s_i \rangle}{\sigma_i} \right)^3 \right\rangle$$

Kurtosis

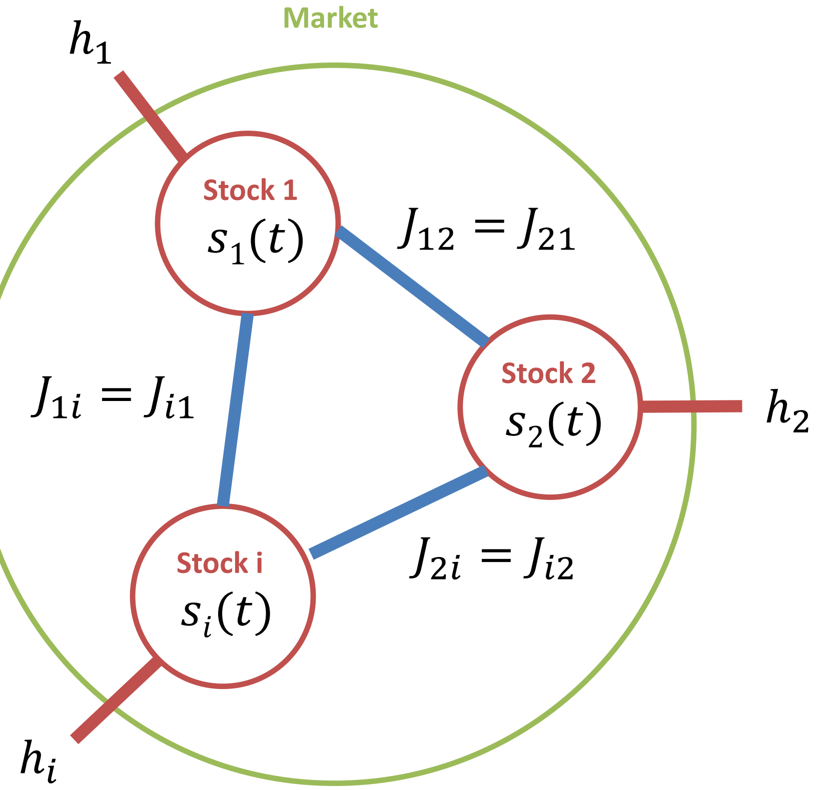
$$\gamma_{2i} := \frac{\left\langle \left(s_i - \langle s_i \rangle \right)^4 \right\rangle}{\sigma_i^4} - 3$$

Equilibrium distribution

$$p(\mathbf{s}) = \mathcal{Z}^{-1} \exp\{-\mathcal{H}(\mathbf{s})\}$$

$$\mathcal{H}(\mathbf{s}) = -\mathbf{h}^\intercal \mathbf{s} - \mathbf{s}^\intercal \mathbf{J} \mathbf{s}$$

(no hidden nodes)



infer h_i and J_{ij} so that

$$\begin{aligned} \langle s_i \rangle_{\text{data}} &= \langle s_i \rangle_{\text{model}} \\ \langle s_i s_j \rangle_{\text{data}} &= \langle s_i s_j \rangle_{\text{model}} \end{aligned}$$

[T. Tanaka, Phys. Rev. E 58, 2302 (1998)]
[Y. Roudi and J. Hertz, PRL 106, 048702 (2011)]

Exact learning

$$\begin{aligned}\delta h_i &= \eta_h \left(\langle s_i \rangle_{\text{data}} - \langle s_i \rangle_{\text{model}} \right) \\ \delta J_{ij} &= \eta_J \left(\langle s_i s_j \rangle_{\text{data}} - \langle s_i s_j \rangle_{\text{model}} \right)\end{aligned}$$

Approximate learning

Naïve Mean Field (nMF)

$$\mathbf{J}^{\text{nMF}} = \mathbf{A}^{-1} - \mathbf{C}^{-1}$$

$$A_{ij} = (1 - \langle s_i \rangle^2)\delta_{ij}$$

$$h_i^{\text{nMF}} = \tanh^{-1}\langle s_i \rangle - \sum_{j=1}^{N} J_{ij}\langle s_i \rangle$$

Thouless-Anderson-Palmer (TAP)

$$\left(\mathbf{C}^{-1}\right)_{ij} = -J_{ij}^{\text{TAP}} - 2\left[J_{ij}^{\text{TAP}}\right]^2 \langle s_i \rangle \langle s_j \rangle$$

$$h_i^{\text{TAP}} = h_i^{\text{nMF}} - \langle s_i \rangle \sum_{j=1}^{N} \left[J_{ij}\right]^2 \left\{1 - \langle s_i \rangle^2\right\}$$

3. Results

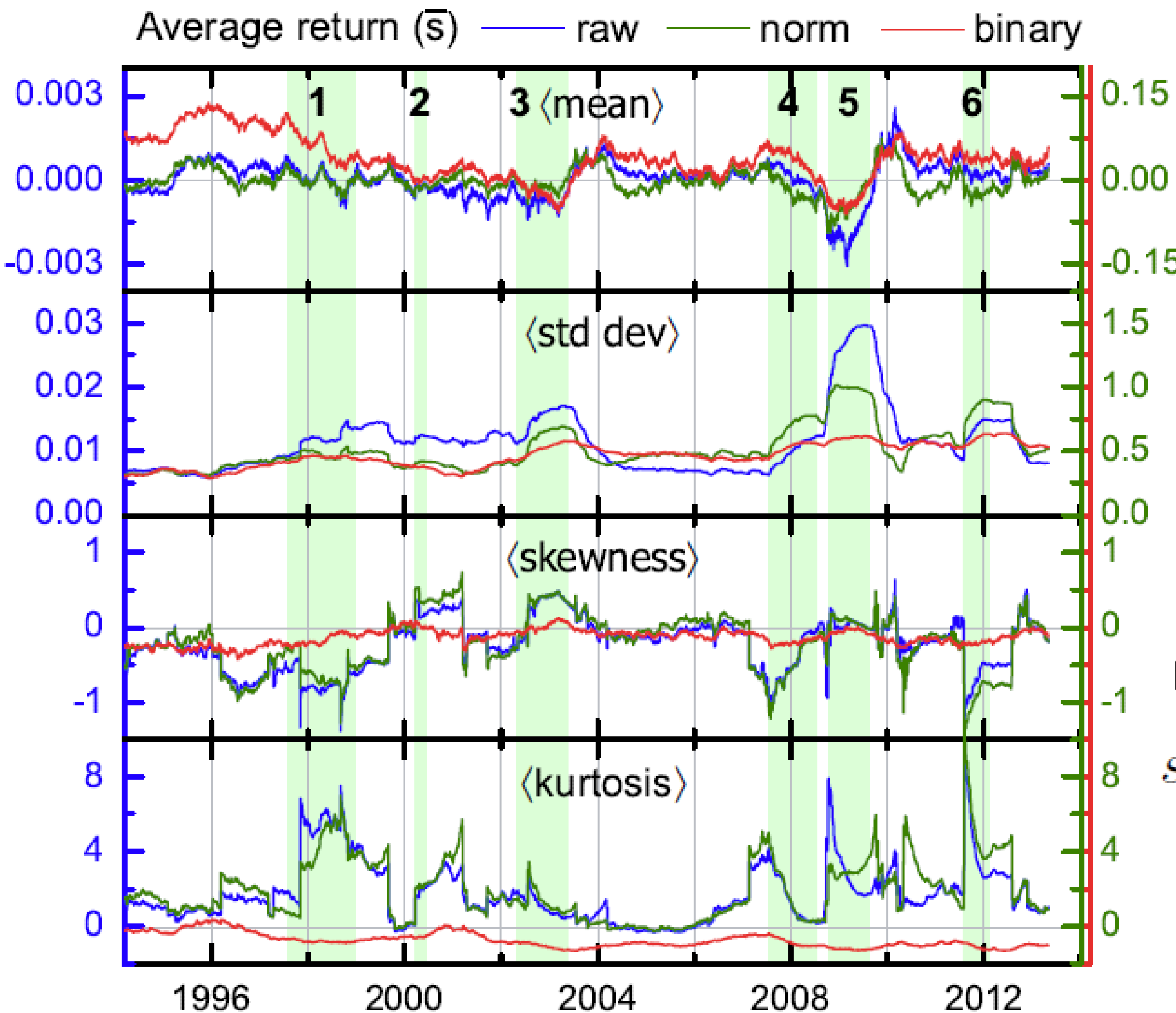
N=71 stocks from the S&P 500 index
Approximately 5000 trading days for 1993-2013
Moving window T=250 days (approximately 1 year)

Portfolio ("market")

$$m(t) = \sum_{i=1}^{N} s_i(t) = N\overline{s}$$

$$\sigma_{\mathrm{m}}^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} C_{ij} = N^2\overline{C}$$

Effect of binarization of returns (1-point moments)



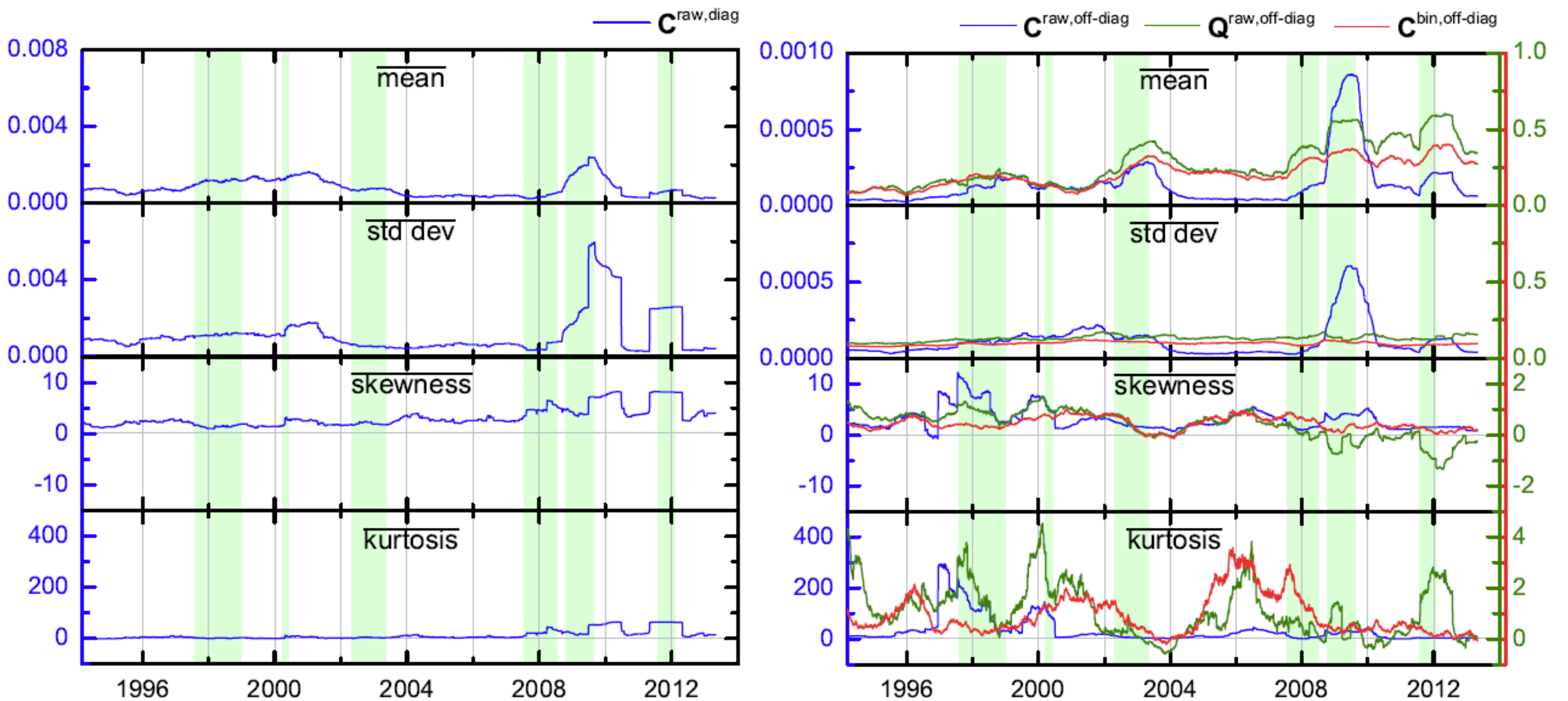
1. Asian and Russian crisis
2. Dot-com bubble
3. US stock market downturn of 2002
4. US housing bubble
5. Global financial crisis
6. European sovereign debt crisis

	$\overline{\text{mean}}$	$\overline{\text{std dev}}$	$\overline{\text{skewness}}$	$\overline{\text{kurtosis}}$
raw, norm	0.711	0.720	0.928	0.842
raw, bin	0.671	0.591	0.565	-0.002
norm, bin	0.476	0.823	0.561	-0.001

Normalized log-returns

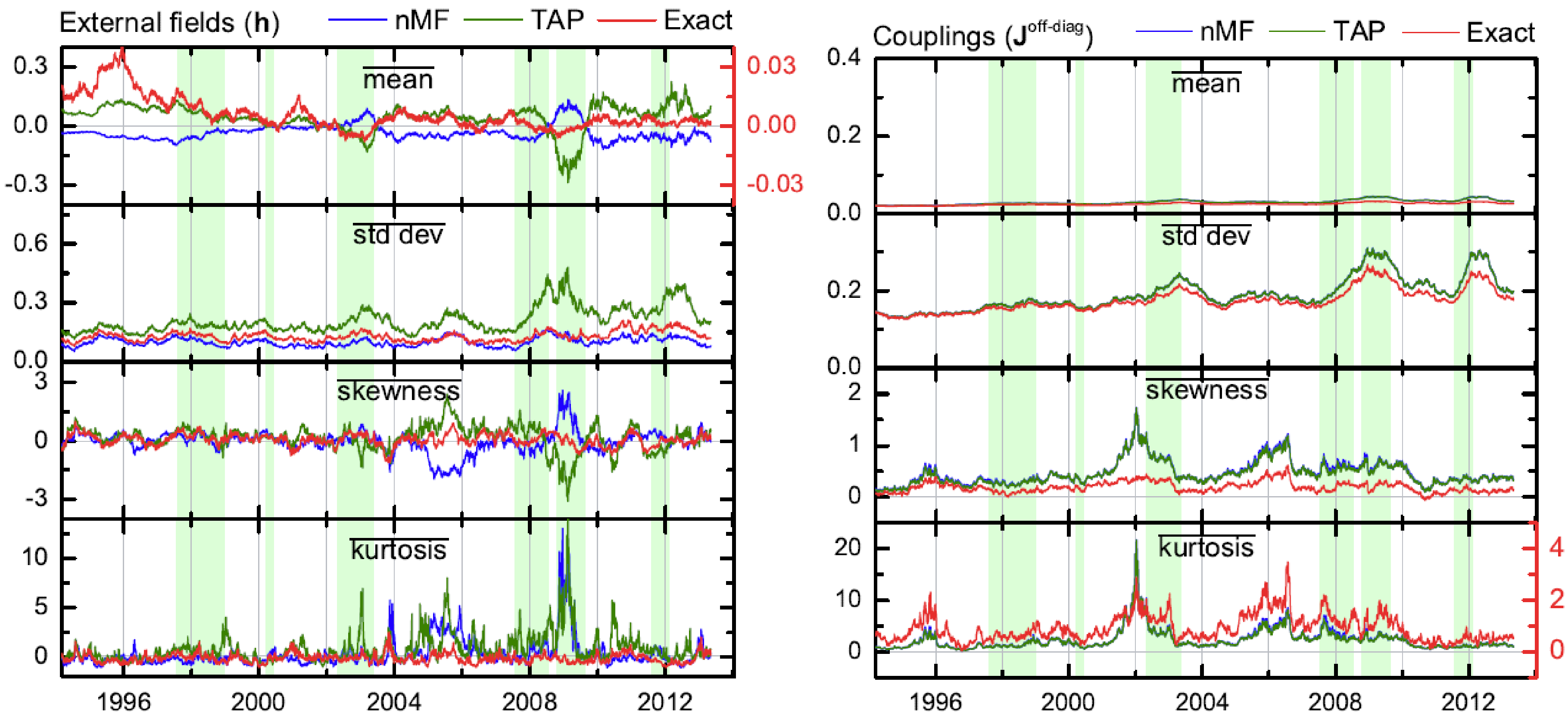
$$s_i^{\text{norm}} := (s_i^{\text{raw}} - \langle s_i \rangle)/\sigma_i$$

Effect of binarization of returns (2-point moments)

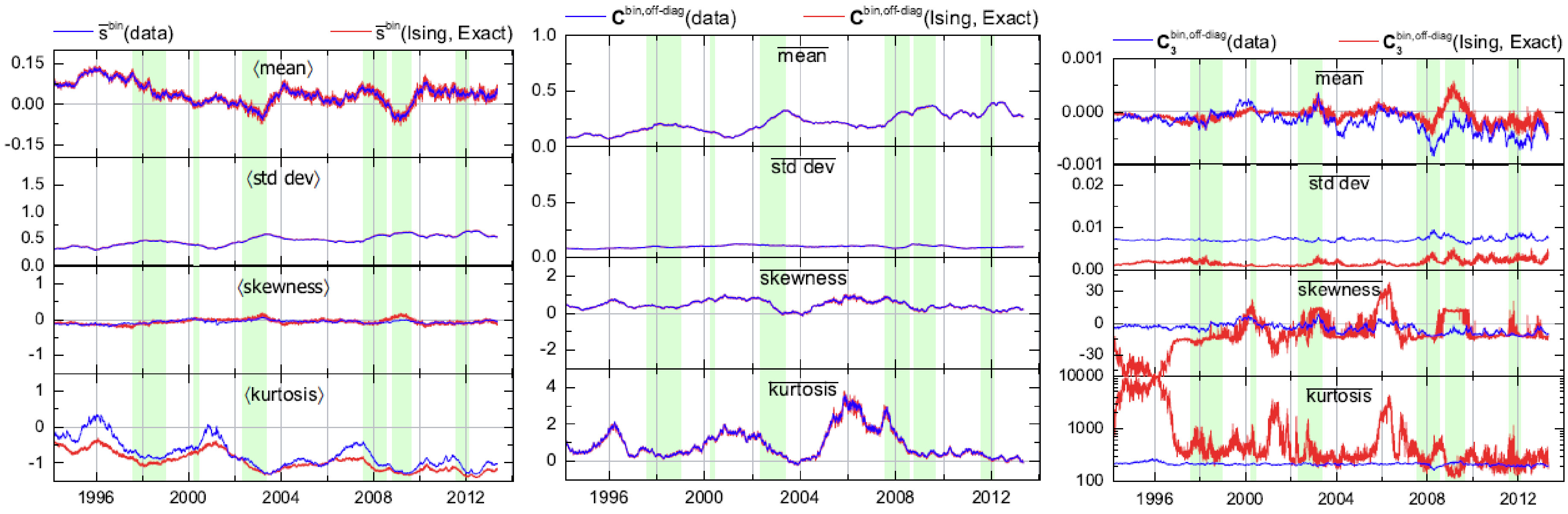


	$\overline{\text{mean}}$	$\overline{\text{std dev}}$	$\overline{\text{skewness}}$	$\overline{\text{kurtosis}}$
raw, norm	0.617	0.191	0.480	0.419
raw, bin	0.561	0.470	0.198	-0.087
norm, bin	0.970	0.668	0.604	-0.281

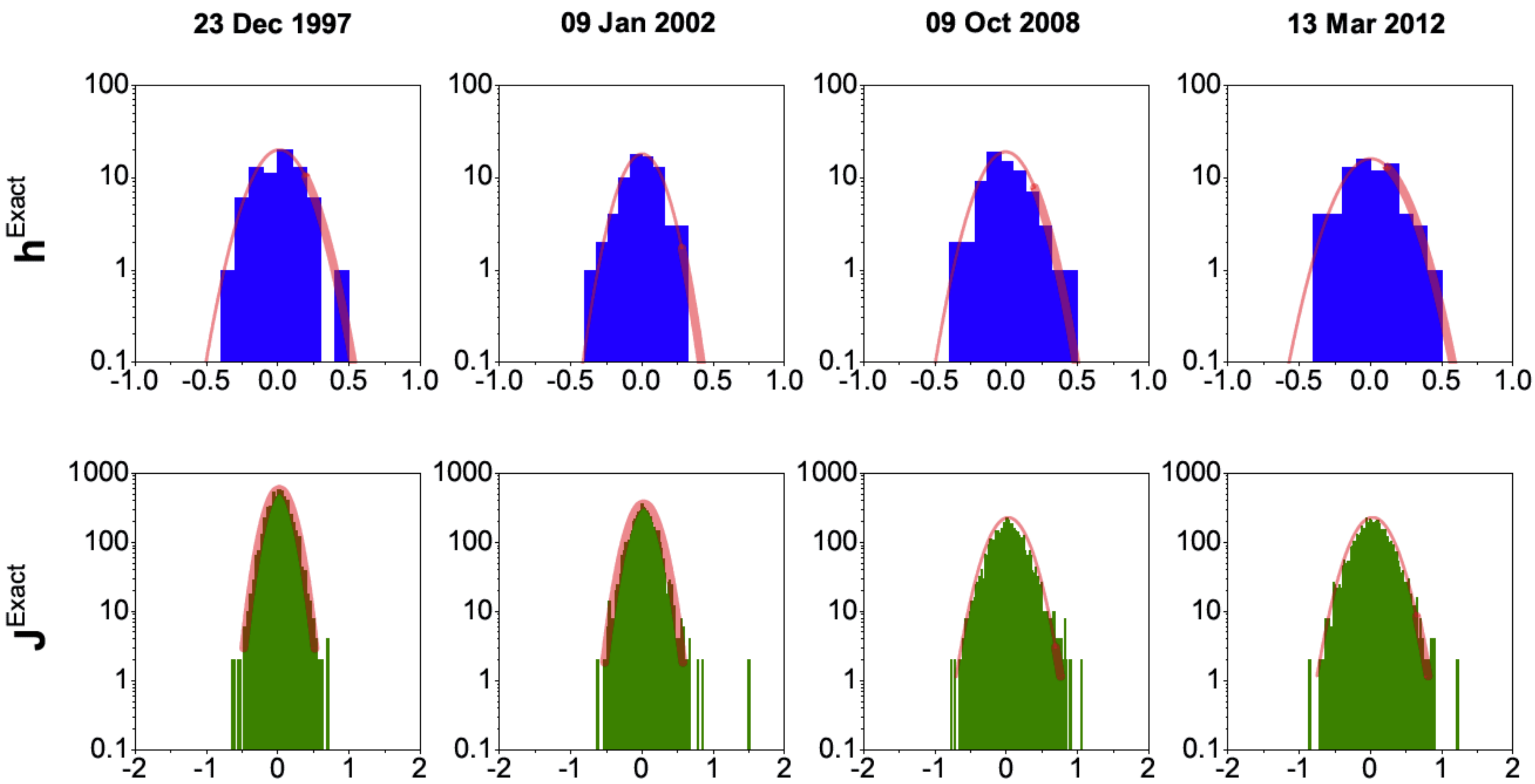
Dynamics of the inferred external fields and couplings



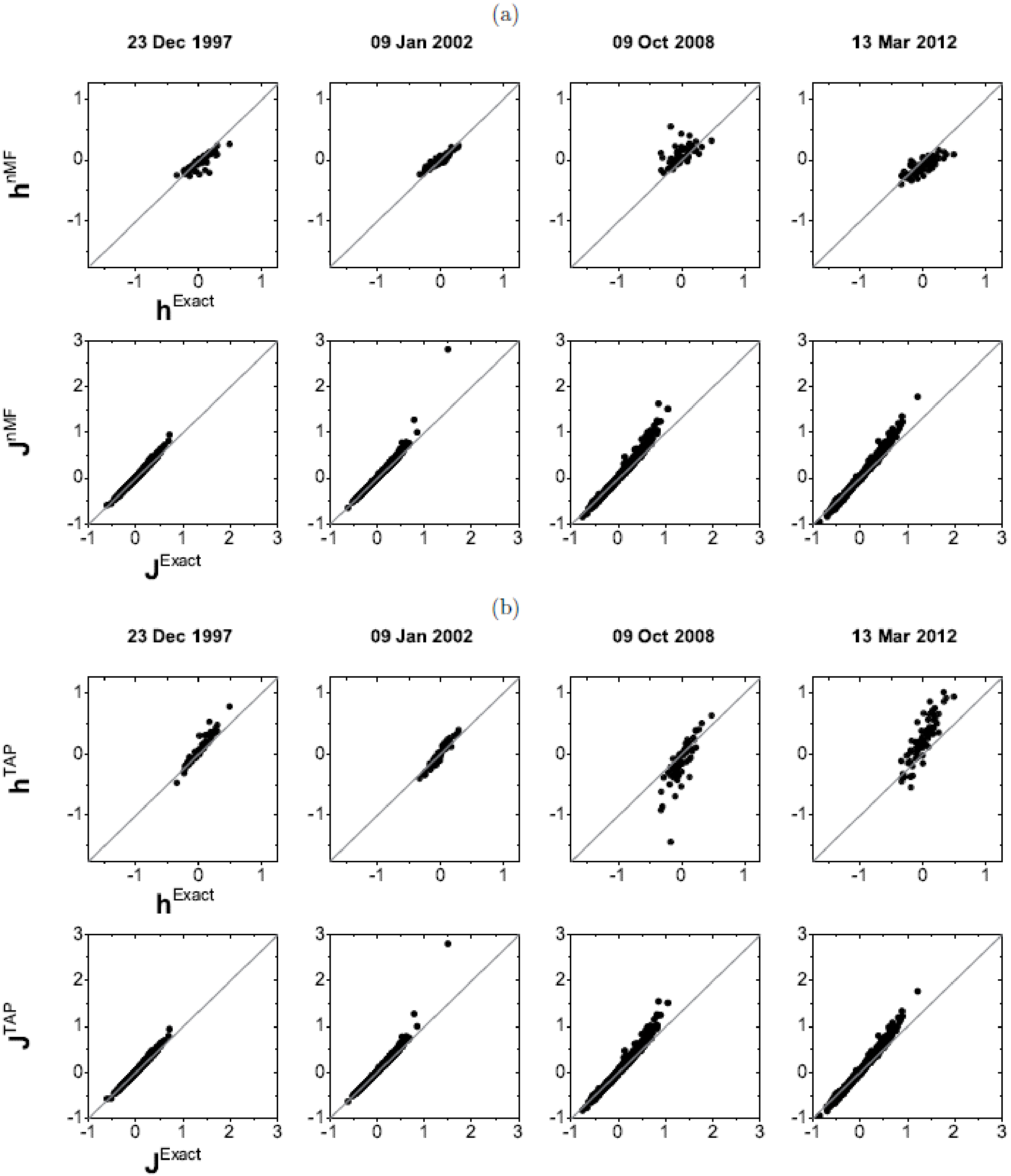
Recovering observables from the model trained using exact learning



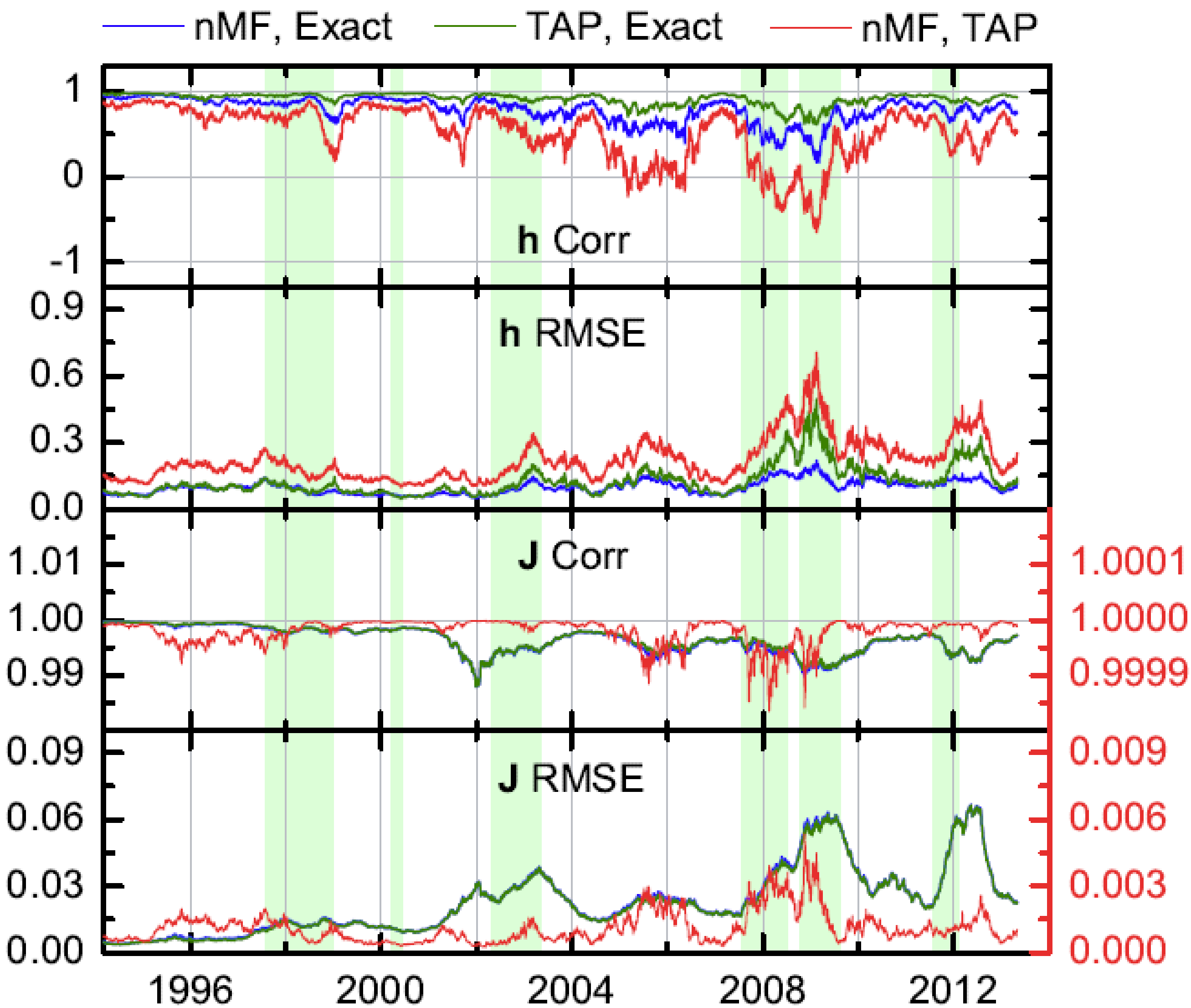
Distribution of the exact external fields and couplings



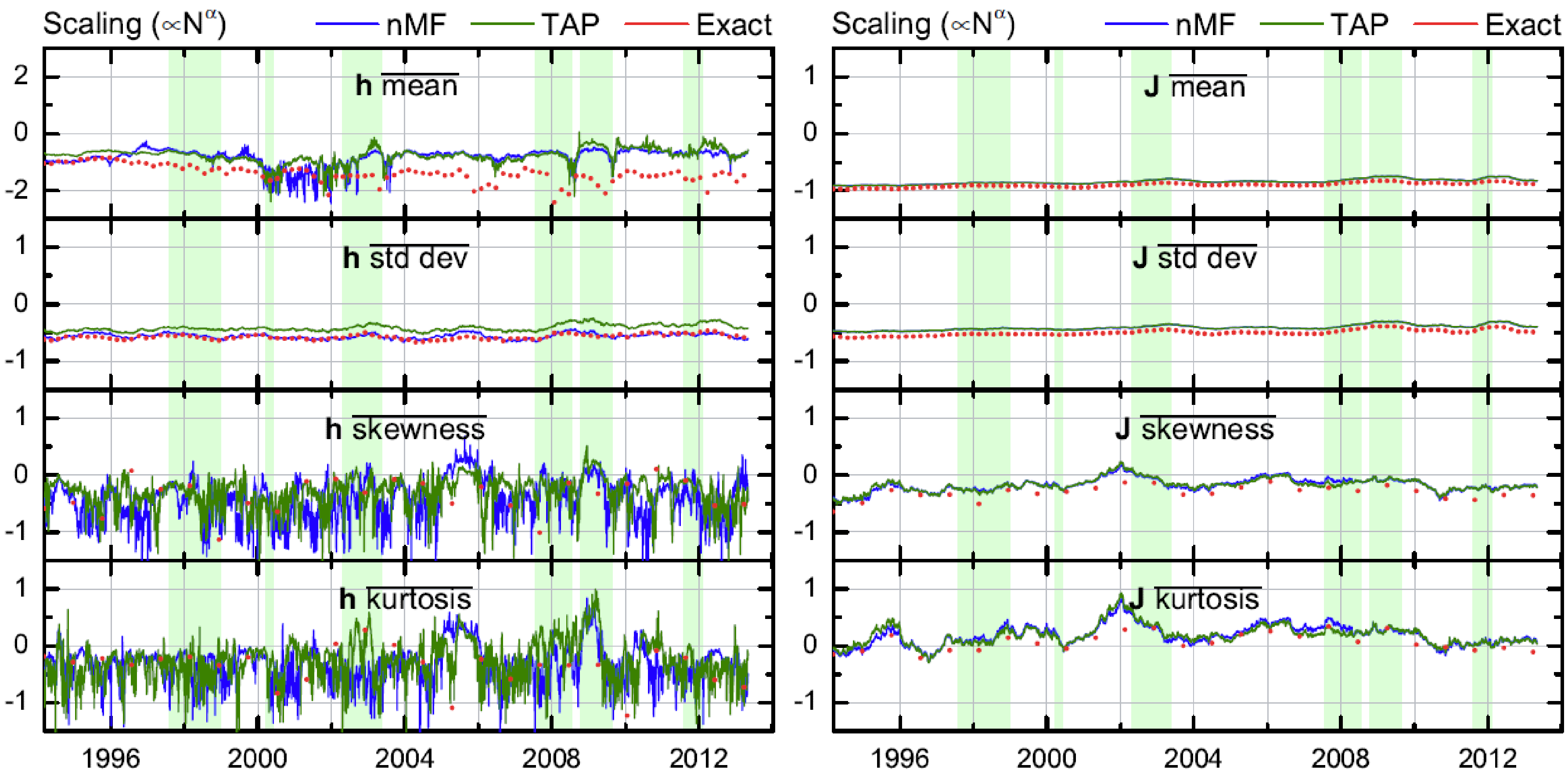
Inference comparison



Inference comparison



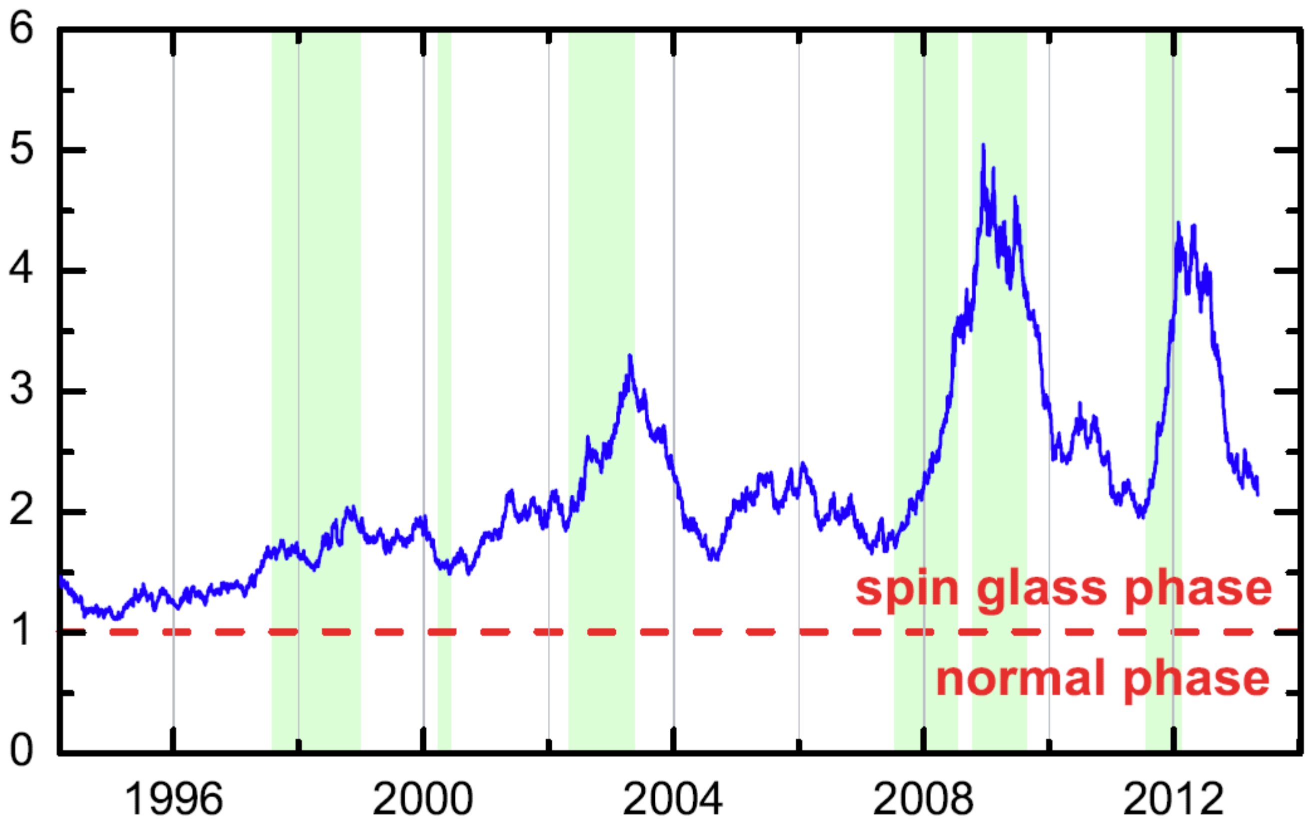
Scaling


Scaling (∝N^α)
nMF
TAP
Exact
h mean
h std dev
h skewness
h kurtosis
J mean
J std dev
J skewness
J kurtosis
2
0
-2
1
0
-1
1996
2000
2004
2008
2012

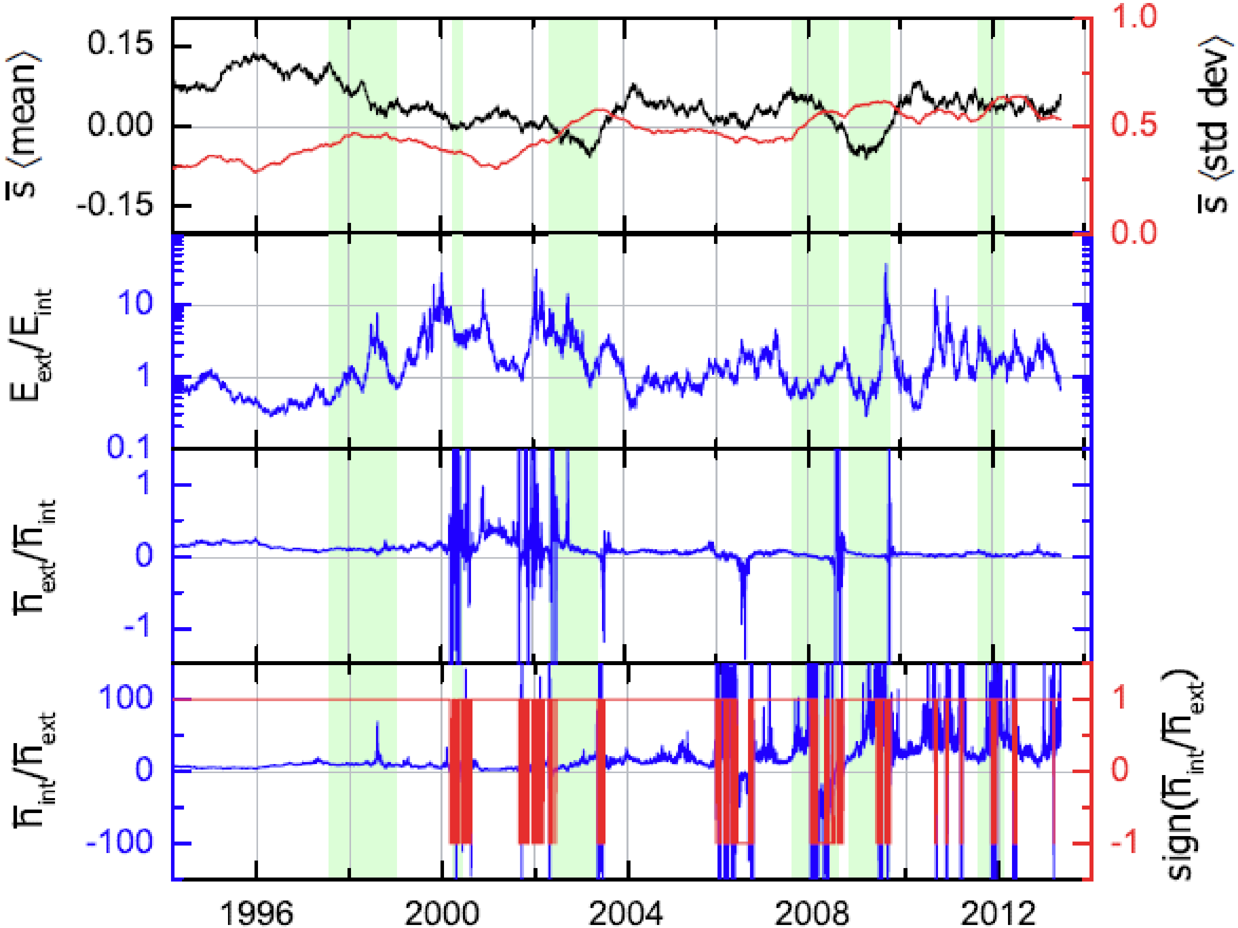
Stability of Sherrington-Kirkpatrick solution

$$\left(\overline{J^2} - \overline{J}^2\right) \sum_i \left(1 - \langle s_i \rangle^2\right)^2 < 1$$

[de Almeida JRL, Thouless DJ, Journal of Physics A: Mathematical and General 11: 983 (1978)]



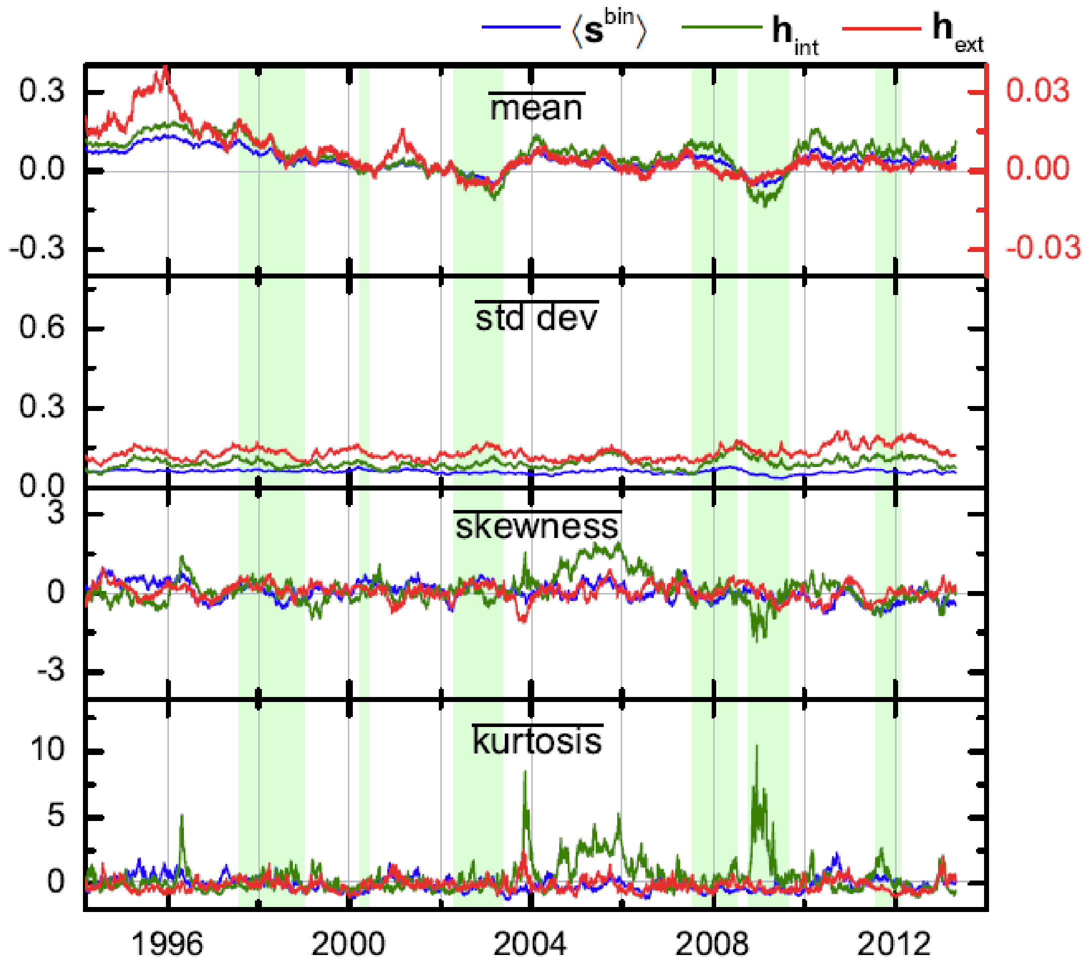
Internal bias vs External bias



$$\mathbf{h}^{\mathrm{ext}} \equiv \mathbf{h},\ \mathbf{h}^{\mathrm{int}} = \langle \mathbf{s}^{\mathsf{T}} \rangle \mathbf{J}$$

$$E^{\mathrm{ext,int}} = -\left(\mathbf{h}^{\mathrm{ext,int}}\right)^{\mathsf{T}} \langle \mathbf{s} \rangle$$

Internal bias vs external bias



	$\overline{\text{mean}}$	$\overline{\text{std dev}}$	$\overline{\text{skewness}}$	$\overline{\text{kurtosis}}$
Internal	0.978	0.440	0.193	0.152
External	0.848	0.203	0.474	0.257

4. Conclusions

1. Accuracy of the mean field inference methods significantly drops in the periods of financial crises.

2. External fields and couplings in the financial market are Gaussian random variables, however outliers are often present.

3. Comparison with the infinite-range spin glass models suggests that the financial market is a disordered magnet with frustrated interactions in spin glass phase and cannot be described by simple order parameters, e.g. mean return or standard deviation. It must be treated within the replica symmetry breaking formalism.

4. External and external fields make different contribution to the market dynamics. Mean market dynamics is very sensitive to mean external field.

Thanks for your attention!