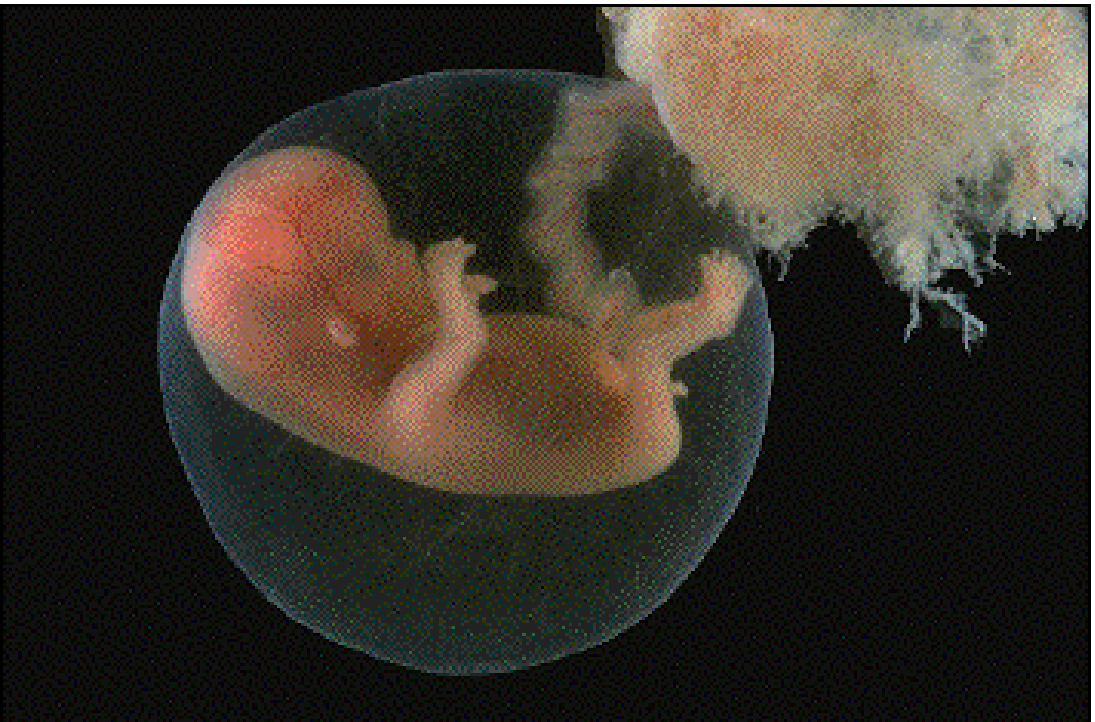
Mechanism for and Detection of Pockets of Predictability in Complex Adaptive Systems

Didier SORNETTE

Institute of Geophysics and Planetary Physics
and Department of Earth and Space Science
University of California, Los Angeles, CA, USA

Laboratoire de Physique de la Matière Condensée
CNRS UMR 6622 and University of Nice, 06108 Nice, France

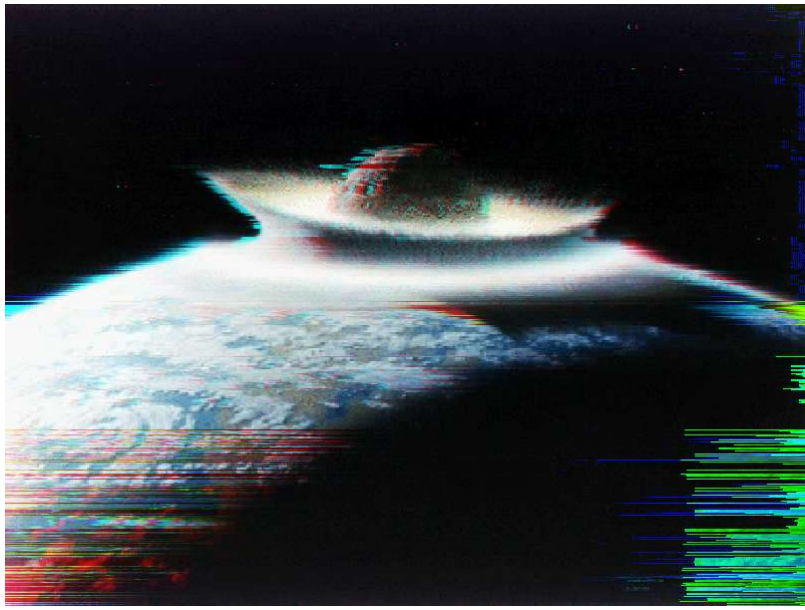
http://www.ess.ucla.edu/faculty/sornette/



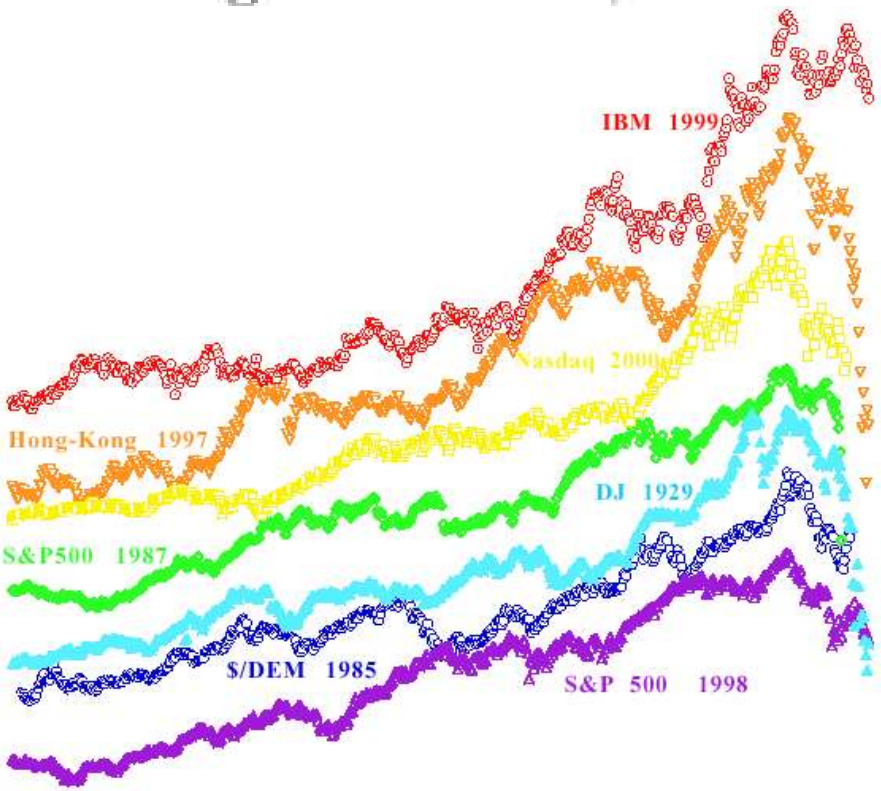
Collaborators:
Y. Ageon (CNRS, France)
J. Andersen (CNRS, France)
D. Darcet (Insight Finance, France)
K. Ide (UCLA)
A. Helmstetter (UCLA)
A. Johansen (Denmark)
Y. Malevergne (Univ. Lyon, France)
V: Pïsarenko (Acad. Sci. Moscow, Russia)
W.-X. Zhou (UCLA)

CRISES

- dramatic and rapid change of a system which is the culmination of a complex preparatory stage.
- fundamental societal impacts
- large natural catastrophes
  1. earthquakes,
  2. volcanic eruptions,
  3. hurricanes and tornadoes,
  4. landslides, avalanches,
  5. lightning strikes,
  6. meteorite/asteroid impacts,
  7. catastrophic events of environmental degradation,



- failure of engineering structures,
- crashes in the stock market,
- social unrest leading to large-scale strikes and upheaval,
- economic drawdowns on national and global scales,
- regional power blackouts,
- traffic gridlock,
- diseases and epidemics, etc.



(useless?) IMPOSSIBILITY THEOREM

Algorithmic complexity theory: most complex systems have been proved to be computationally irreducible, i.e. the only way to decide about their evolution is to actually let them evolve in time.

The future time evolution of most complex systems appears inherently unpredictable.

BUT, Physics works and is not hampered by computational irreducibility because we only ask for answers at some coarse-grained level.

Computational Irreducibility and the Predictability of Complex Physical Systems

Navot Israeli and Nigel Goldenfeld

PhysRevLett.92.074105 (2004)

256 nearest neighbor 1D cellular automata (Wolfram)

Class 3 → Class 1

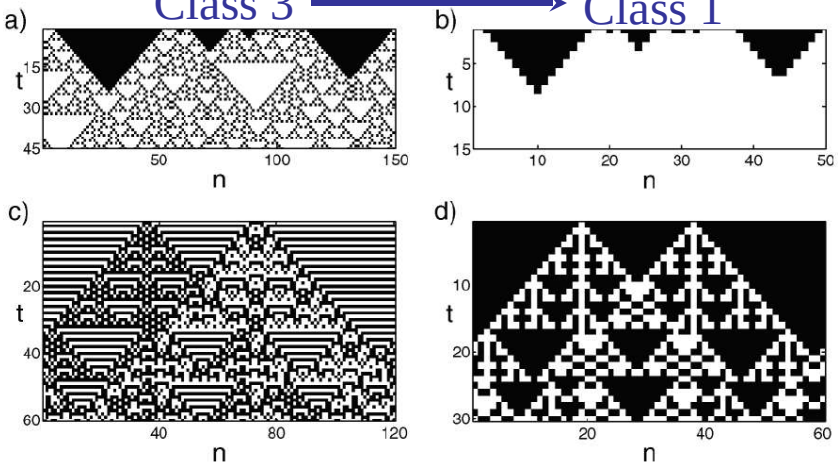


FIG. 1. Examples of coarse-graining transitions. (a) and (b) show coarse-graining rule 146 by rule 128. (a) shows results of running rule 146. The top line is the initial condition and time progress from top to bottom. (b) shows the results of running rule 128 with the coarse-grained initial condition from (a). (c) and (d) show coarse-graining rule 105 by rule 150. (c) shows rule 105 and (d) shows rule 150.

$$C(f_A^{T\cdot t}a(0)) = f_B^t C(a(0)).$$

Namely, running the original CA for Tt time steps and then coarse graining is equivalent to coarse graining the initial condition and then running the modified CA t time steps. The constant T is a time scale associated with the coarse graining.

240 coarse-grainable

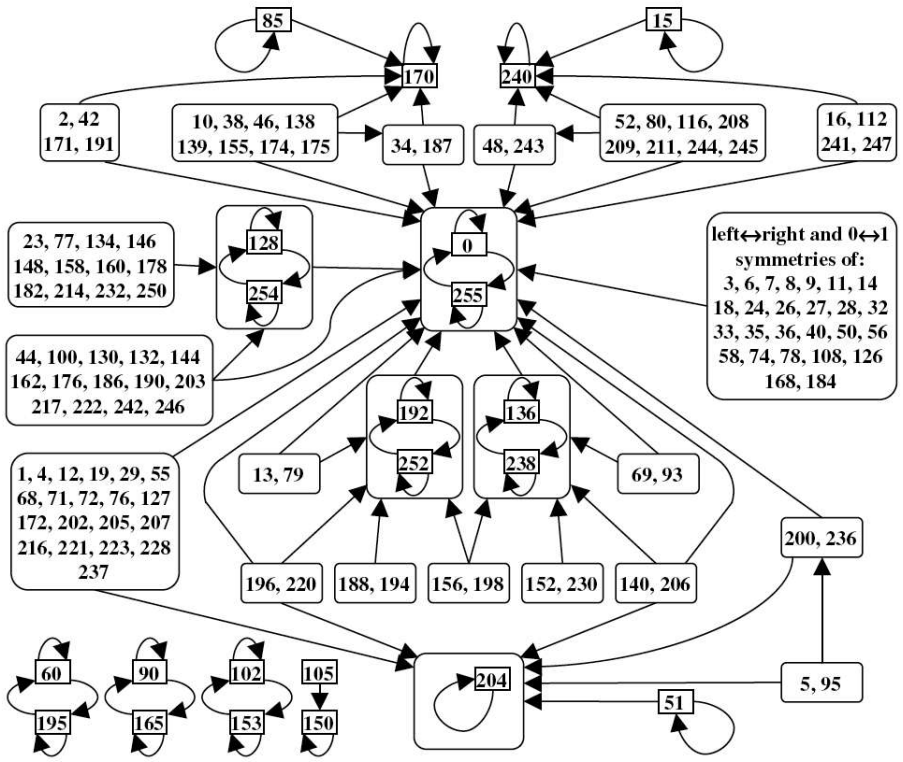


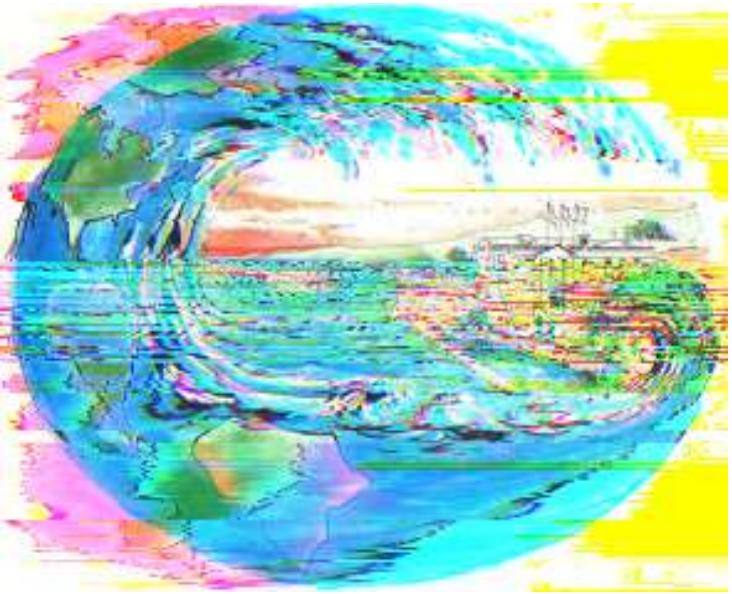
FIG. 2. Coarse-graining transitions within the family of 256 elementary CA. Only transitions with a cell block size $N = 2$, 3, and 4 are shown. An arrow indicates that the origin rules can be coarse grained by the target rules and may correspond to several choices of N and P.

N-block approach with N=2, 3 or 4

Coarse-graining rule 110: CIR => C1

L'vov, V. S., Pomyalov, A., and Procaccia, I. (2001). *Outliers, Extreme Events and Multiscaling*, Physical Review E 6305, 6118, U158–U166.

FIG. 3.2. Apparent probability distribution function of the square of the fluid velocity, normalized to its time average, in the eleventh shell of the toy model of hydrodynamic turbulence discussed in the text. The vertical axis is in logarithmic scale such that the straight line, which helps the eye, qualifies as an apparent exponential distribution. Note the appearance of extremely sparse and large bursts of velocities at the extreme right above the extrapolation of the straight line. Reproduced from [252].



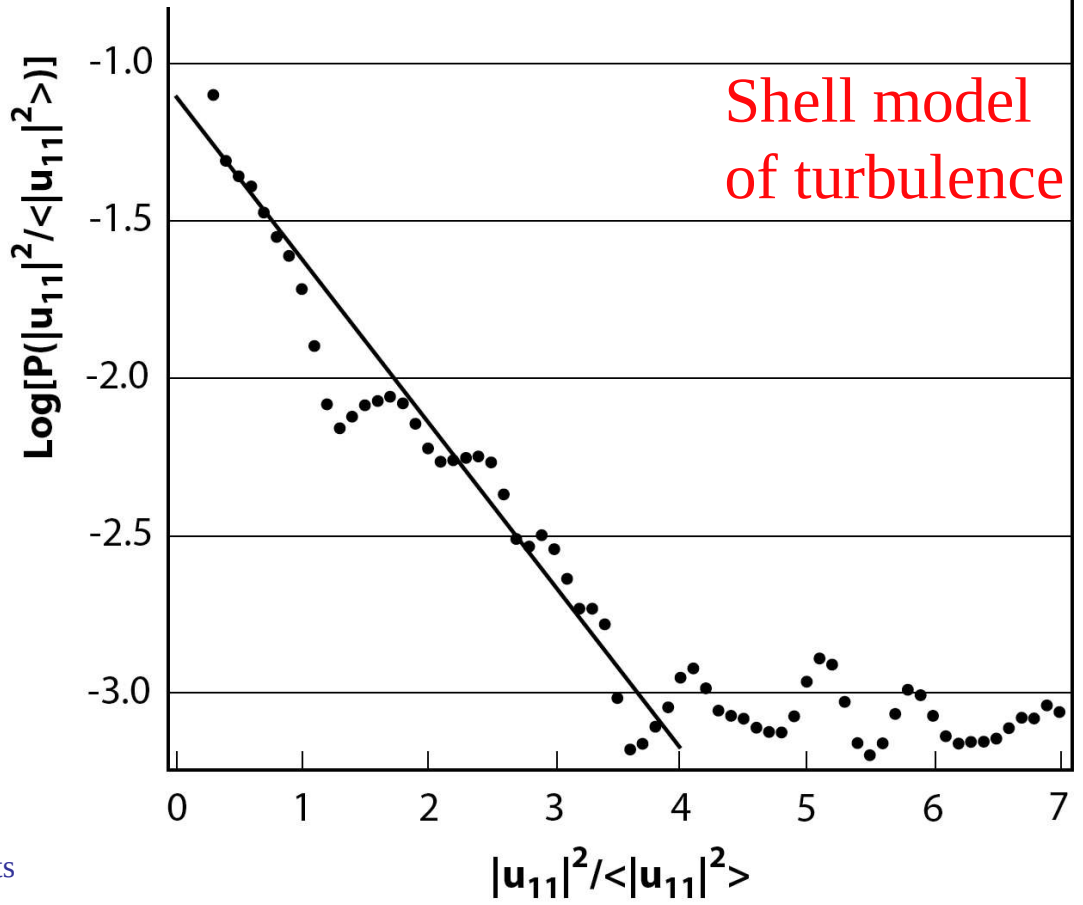
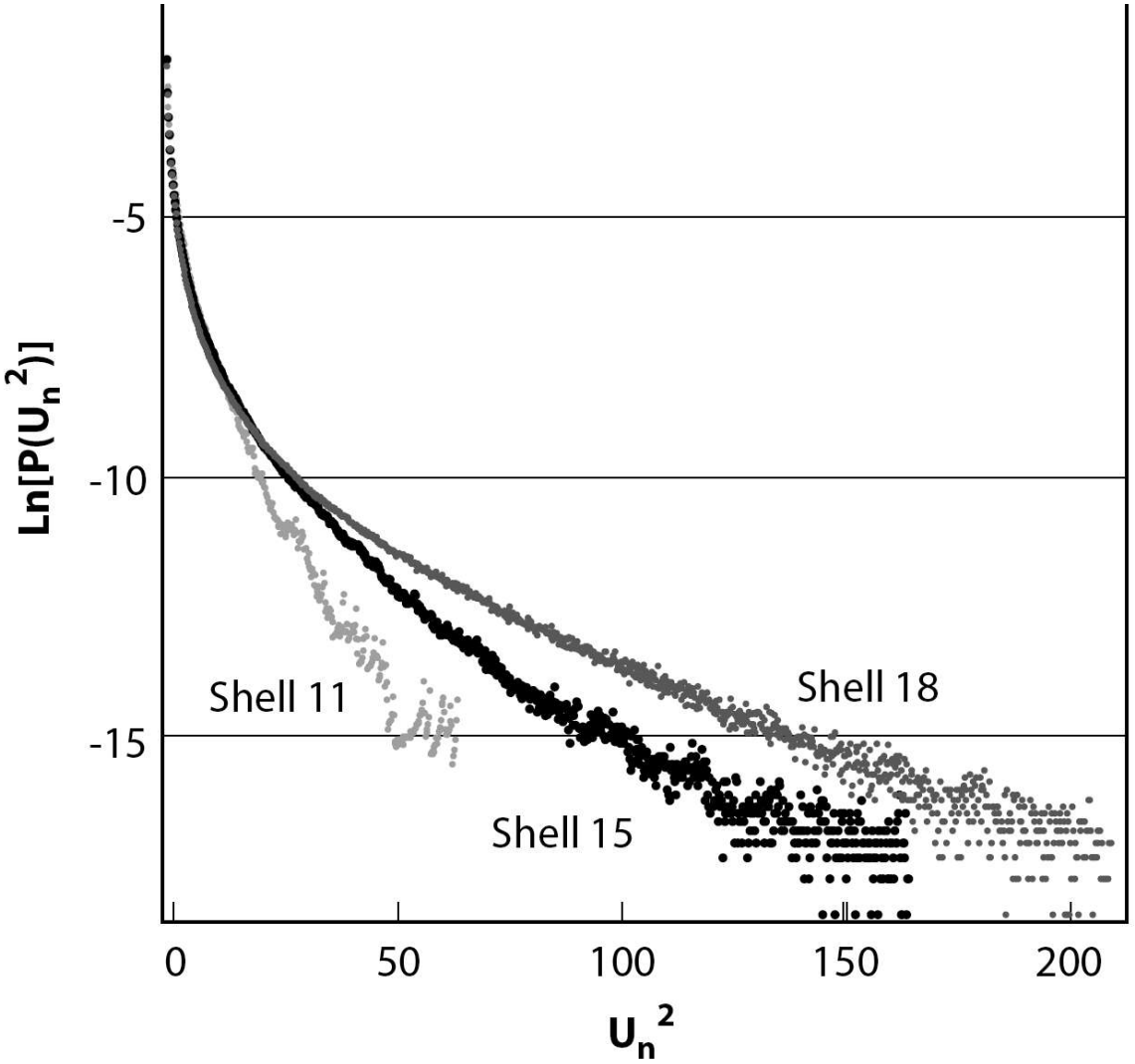


FIG. 3.3. Probability distribution function of the square of the velocity as in Figure 3.2 but for a much longer time series, so that the tail of the distributions for very large fluctuations is much better constrained. The hypothesis that there are no outliers is tested here by "collapsing" the distributions for the three shown layers. While this is a success for small fluctuations, the tails of the distributions for large events are very different, indicating that extreme fluctuations belong to a class of their own,



A Mechanism for Pockets of Predictability in Complex Adaptive Systems

J.V. Andersen and D. Sornette

Europhys. Lett., 70 (5), 697-703 (2005)

The Minority Game (MG) and the $-Game ($G)

Total action of agents

$$A^{\mu}(t) \equiv \sum_{i} a_i^{\mu}(t)$$

Price equation

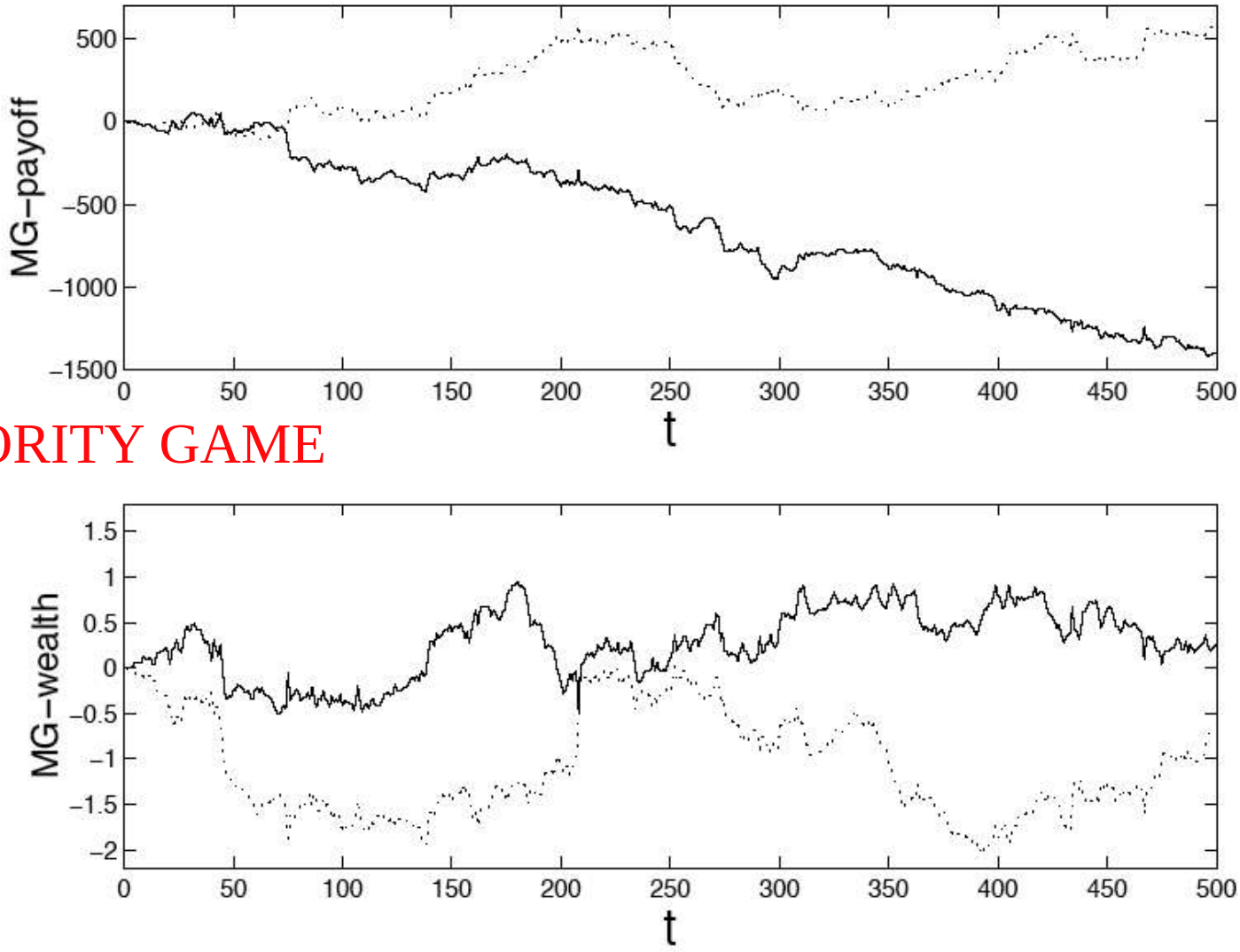
$$\log\left(P(t+1)\right) = \log\left(P(t)\right) + A^{\mu}(t+1/2)/N$$

signal (μ)	*prediction*
000	0
001	0
010	1
011	0
100	1
101	0
110	1
111	0

Example of strategy

MG payoff of strategy i : $g_i(t) = -a_i(t)A(t)$

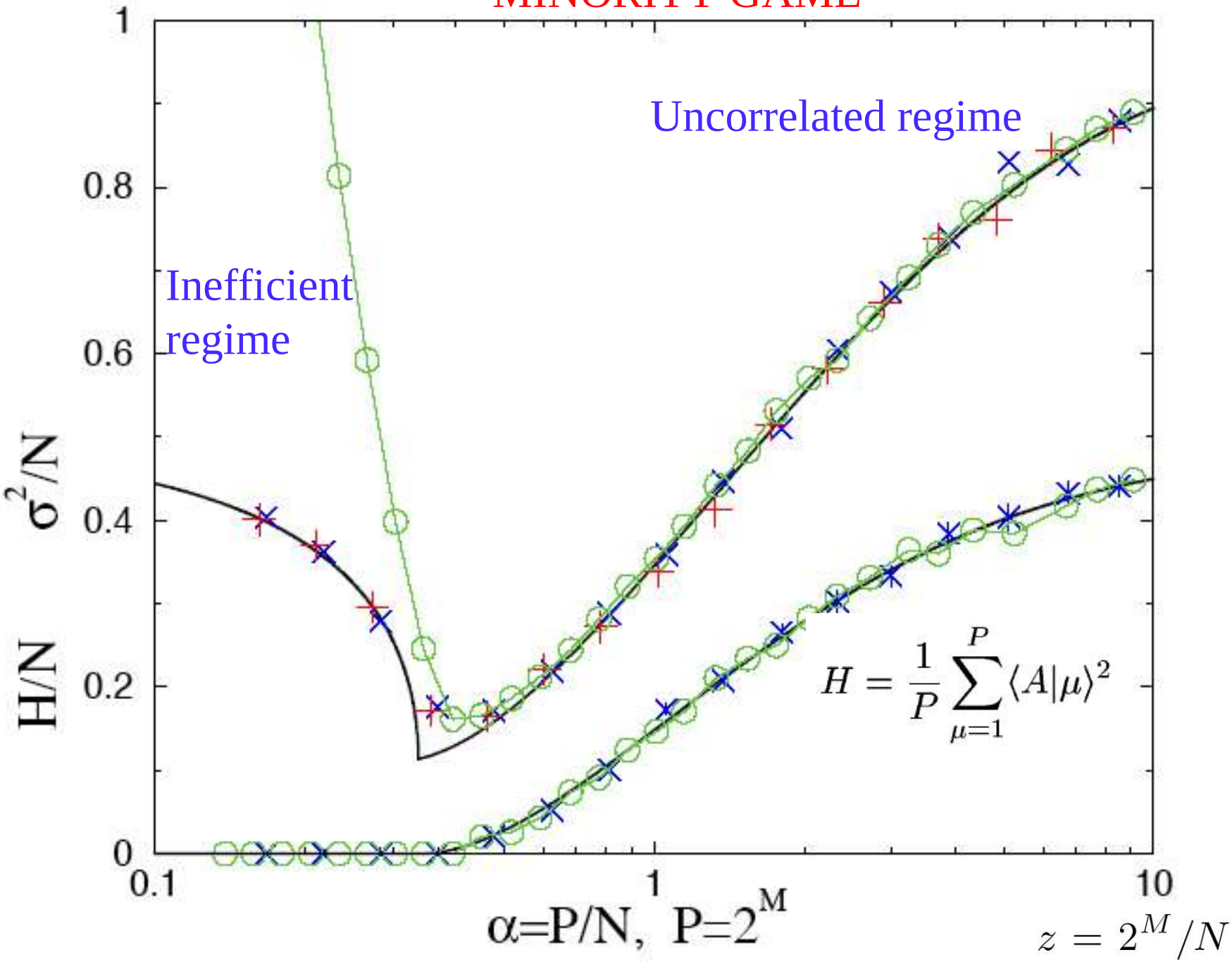
$-game payoff of strategy i : $g_i(t) = a_i(t-1)A(t)$



MINORITY GAME

Fig. 1. Payoff function (1) (upper graph) and wealth (lower graph) for the MG-game showing the best (dotted line) and worst (solid line) performing agent for a game using $N = 501$ agents, memory of $m = 10$ and $s = 10$ strategies per agent. No transaction costs are applied.

MINORITY GAME



$-GAME

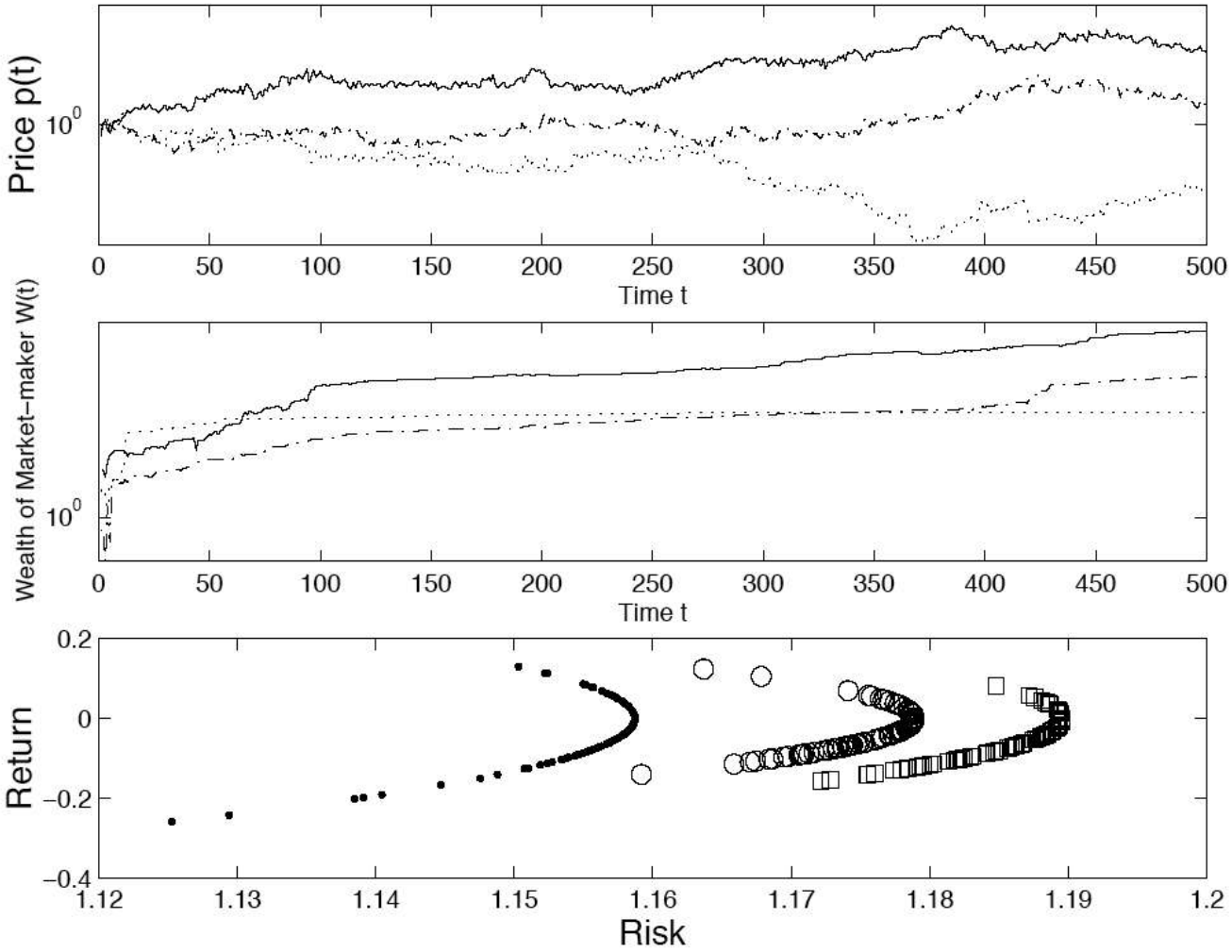


Fig. 2. Price, wealth of market-maker and risk-return plots for three different parameter choices using the payoff function (3) and the constraint that agents can only accumulate one position at a time. Solid line and black circle: $m = 10, s = 4$; dashed-dotted line and circle: $m = 10, s = 10$; dotted line and square: $m = 8, s = 10$.

J.V. Andersen and D. Sornette
The $-game, Eur. Phys. J. B 31, 141-145 (2003)

Predictability of Large Future Changes in a Competitive Evolving Population

Lampert, Howison and Johnson, PRL 88, 017902 (2002)

Third-party game calibration on a black-box game

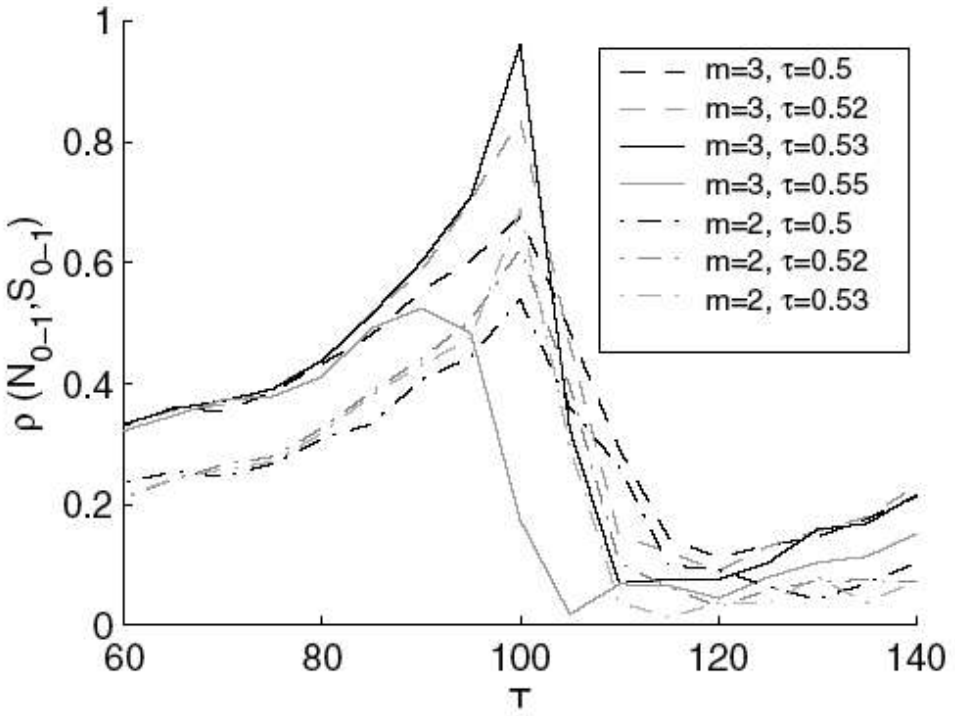


FIG. 1. Estimation of the parameter set for the black-box game. The correlation between N_{0-1} and S_{0-1} is calculated over 200 time steps for an ensemble of candidate third-party games. The third-party game that achieves the highest correlation is the one with the same parameters as the black-box game.

Crash prediction

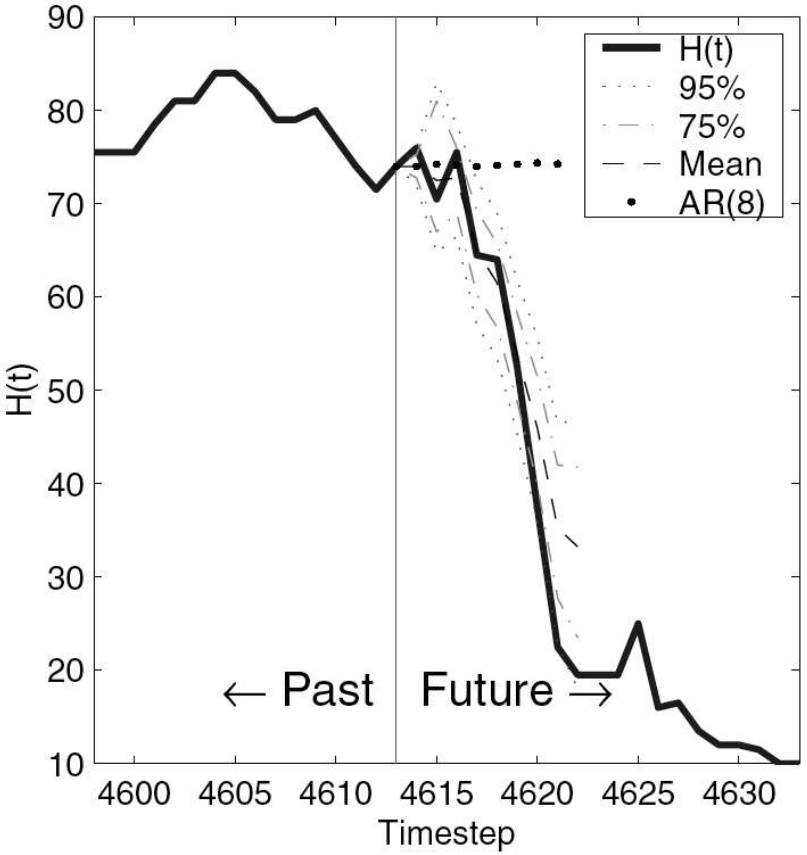


FIG. 3. Comparison between the forecast density function and the realized time series $H(t)$ for a typical large movement. The large, well-defined movement is correctly predicted. An AR(8)-based prediction has been included for comparison.

The concept of decoupled strategies

History: $\mu_m(t) = \{\frac{1+\text{sign}[A^\mu(k)]}{2}; k = t-m+1, ..., t\}$

- A strategy s_j is called $n-$time steps decoupled conditioned on $\mu_m(t)$ if the action $s_j(\mu_m(t+n+1))$ does not depend on $\mu_m(t+1), ..., \mu_m(t+n)$.

Case m=3: $\mu_3(t) = abc$ if $s(bc0) = s(bc1)$

One-step decoupled

The strategy s is two-step decoupled conditioned on $\mu_3(t) = abc$ if $s(c00) = s(c01) = s(c10) = s(c11)$.

If decoupled strategies dominate => predictability since decision independent of next outcome(s)

Decomposition of total action:

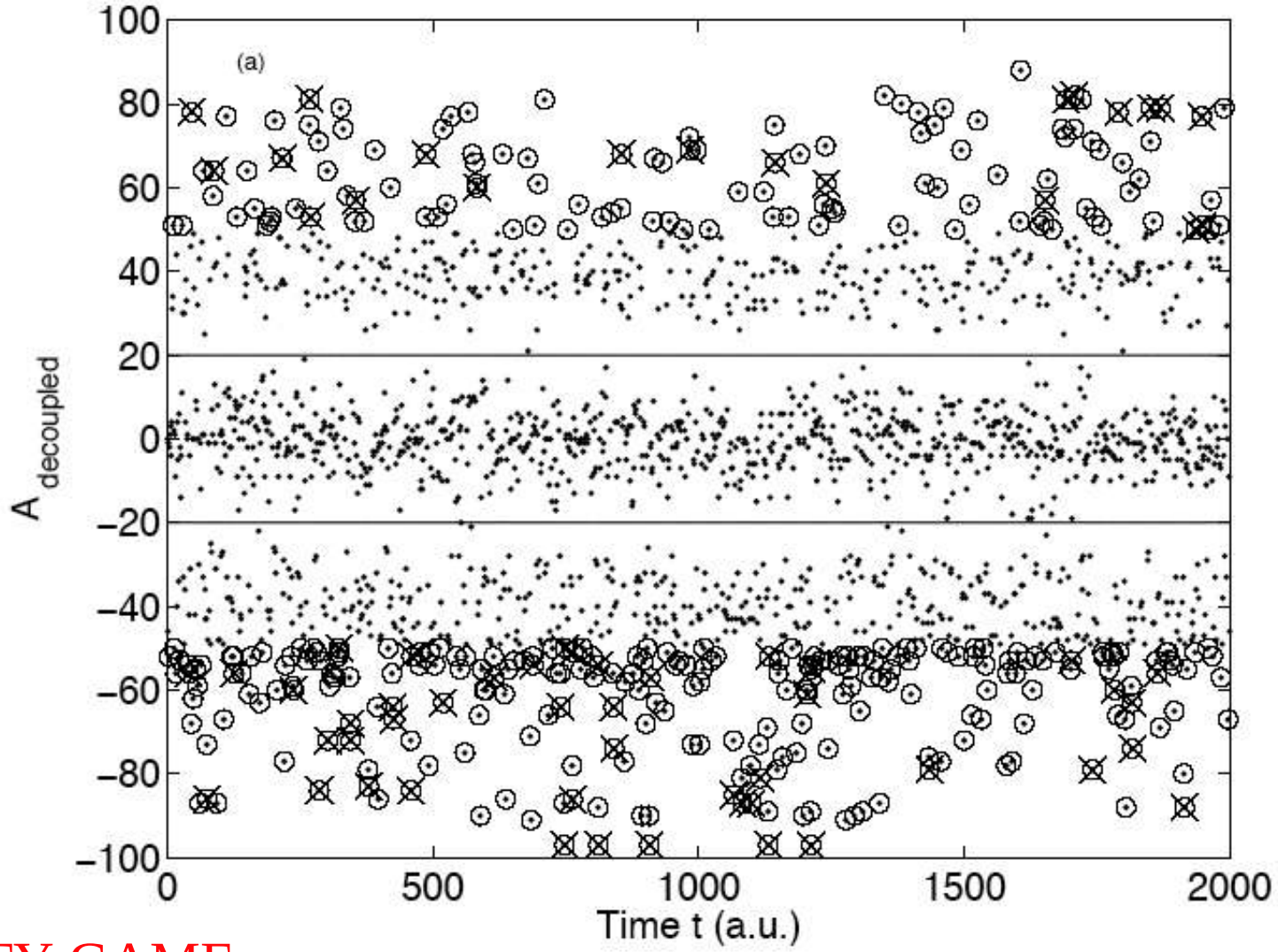
$$A^{\mu_m(t)} \equiv A^{\mu_m(t)}_{\text{coupled}} + A^{\mu_m(t)}_{\text{decoupled}} \qquad (3)$$

Condition of certain predictability

$$|A^{\mu_m(t)}_{\text{decoupled}}(t+n+1)| > N/2$$

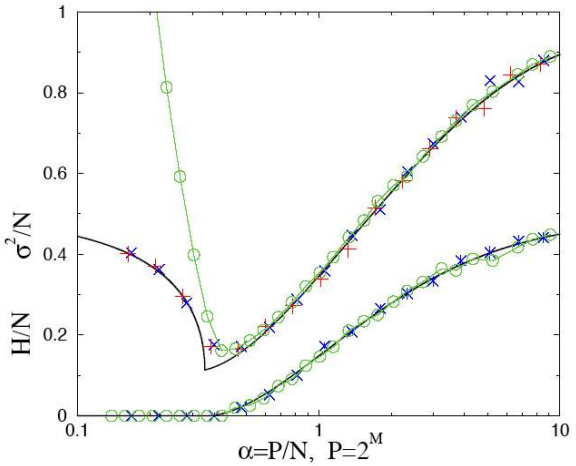
For N=25 and N=102, very small probability for these pockets of predictability to occur by chance (assuming decoupling between agents)

$$\text{Pr}_{\text{pred}} < 7 \cdot 10^{-4} \text{ and } \text{Pr}_{\text{pred}} < 2.5 \cdot 10^{-15}$$



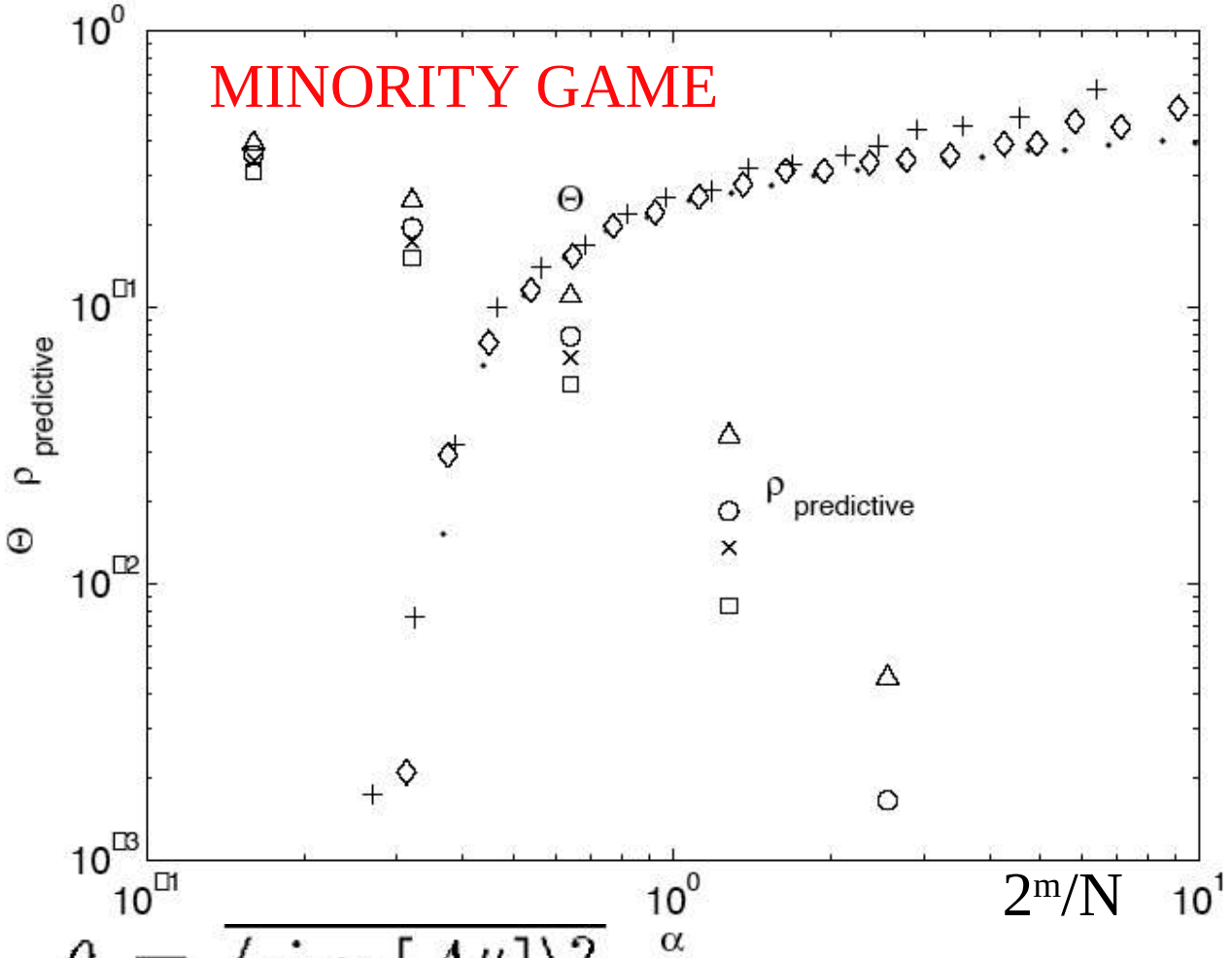
MINORITY GAME

FIG. 1: $A_{\mathrm{decoupled}}$ defined in (3) as a function of time for the MG with $N = 101, s = 12, m = 3$. Circles indicate one-step prediction days, crosses are the subset of days starting a run of two or more consecutive one-step prediction days.



$2^m/N$

MINORITY GAME



Standard measure of predictability $\theta \equiv \overline{\langle \mathrm{sign}[A^\mu] \rangle^2}$ over time and over histories

FIG. 2: $\theta \equiv \overline{\langle \mathrm{sign}[A] \rangle^2}$ and the frequency of predictive days, ρ_{pred} versus $\alpha \equiv 2^m/N$. Simulations for θ were done with $s = 2$, $m = 5$ (+), $m = 6$ (diamonds), $m = 7$ (black dots). Simulations for $\rho_{\mathrm{predictive}}$ were done with $N = 25$, $s = 7$ (squares), $s = 8$ (crosses), $s = 9$ (circles), $s = 12$ (triangles).

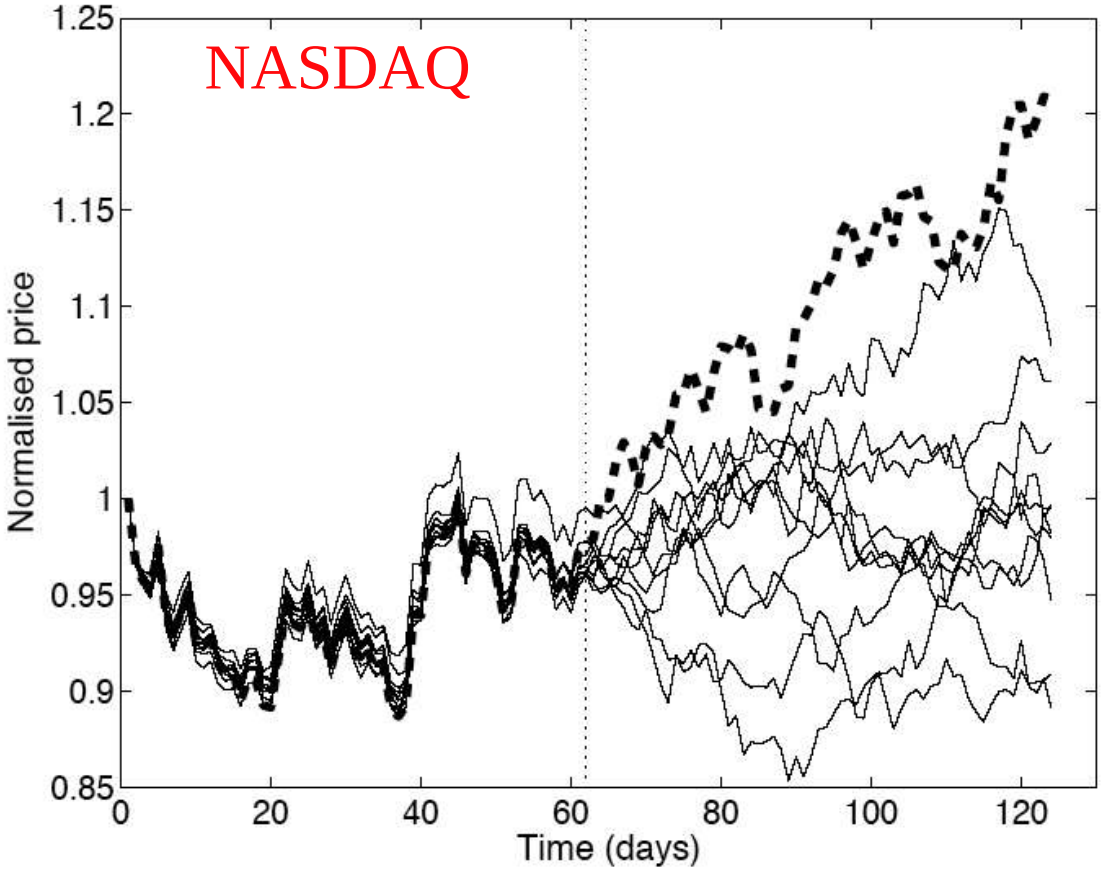


FIG. 3: Fat dashed line: Nasdaq Composite price history (black-box game) as a function of time (days); thin solid lines: ten predicted price trajectories obtained from the third-party games. The first in-sample 61 days are used to calibrate ten third-party games. The days 62-123 are out-of-sample. The third-party games make a poor job at predicting the out-of-sample prices of the Nasdaq Composite index, while table 1 shows that they predict specific pockets of predictability associated with forecasted "prediction days" (see text).

$\|A\|$	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5
%	53	61	67	65	82	70	67	67	100	100
Nb	62	49	39	23	17	10	6	3	2	1

TABLE I: Out-of-sample success rate % (second row) using different thresholds for the predicted global decoupled action (first row) of the third-party \$-games calibrated to the Nasdaq Composite index. Nb (third row) is the number of days from $t = 62$ to 123 which have their predicted global decoupled action $|A_{\text{decoupled}}|$ larger than the value indicated in the first row.

Predictability of large future changes in major financial indices

D. Sornette and W.-X. Zhou

International Journal of Forecasting (in press) cond-mat/0304601

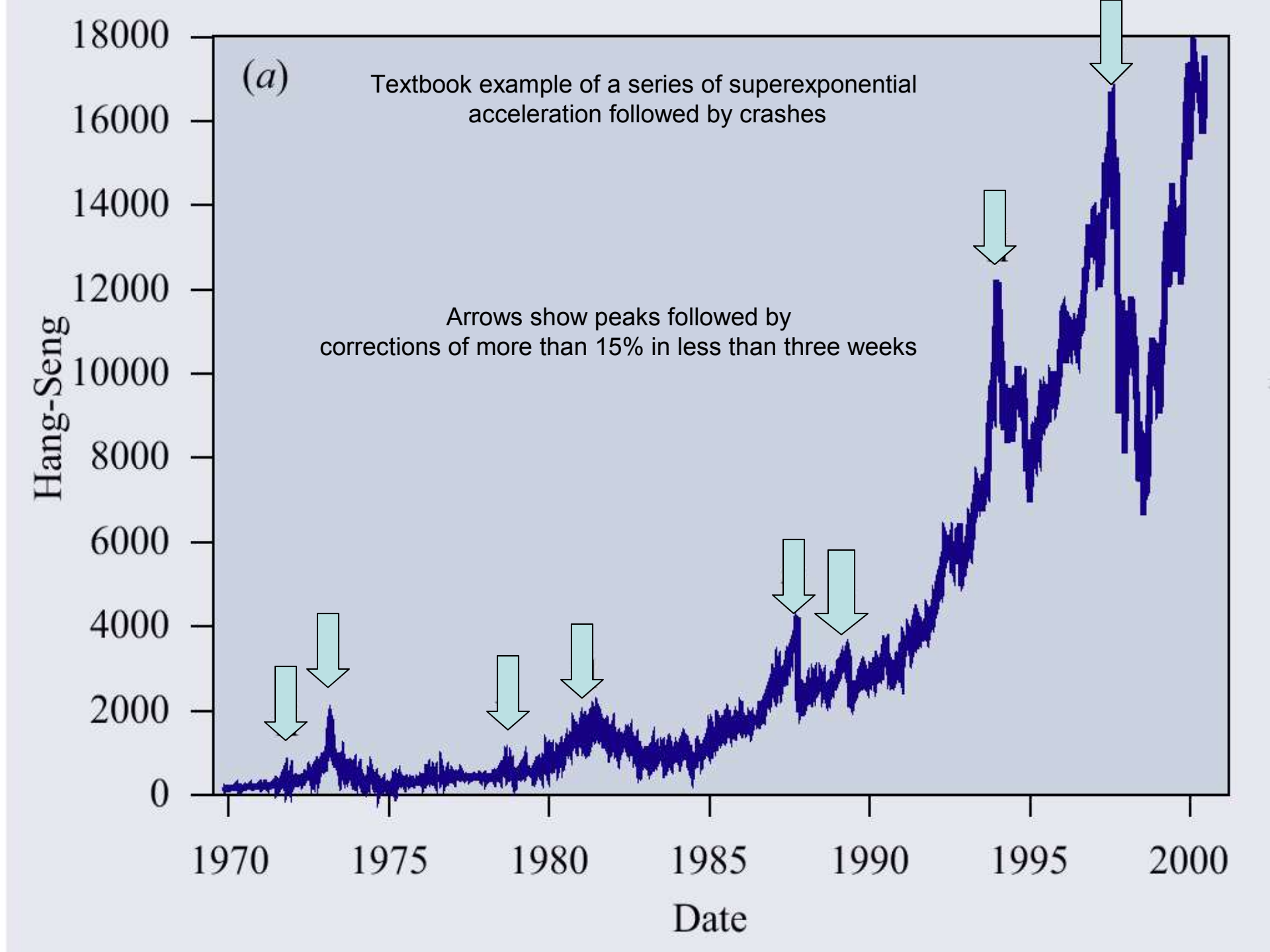
Sparse-data pattern recognition method

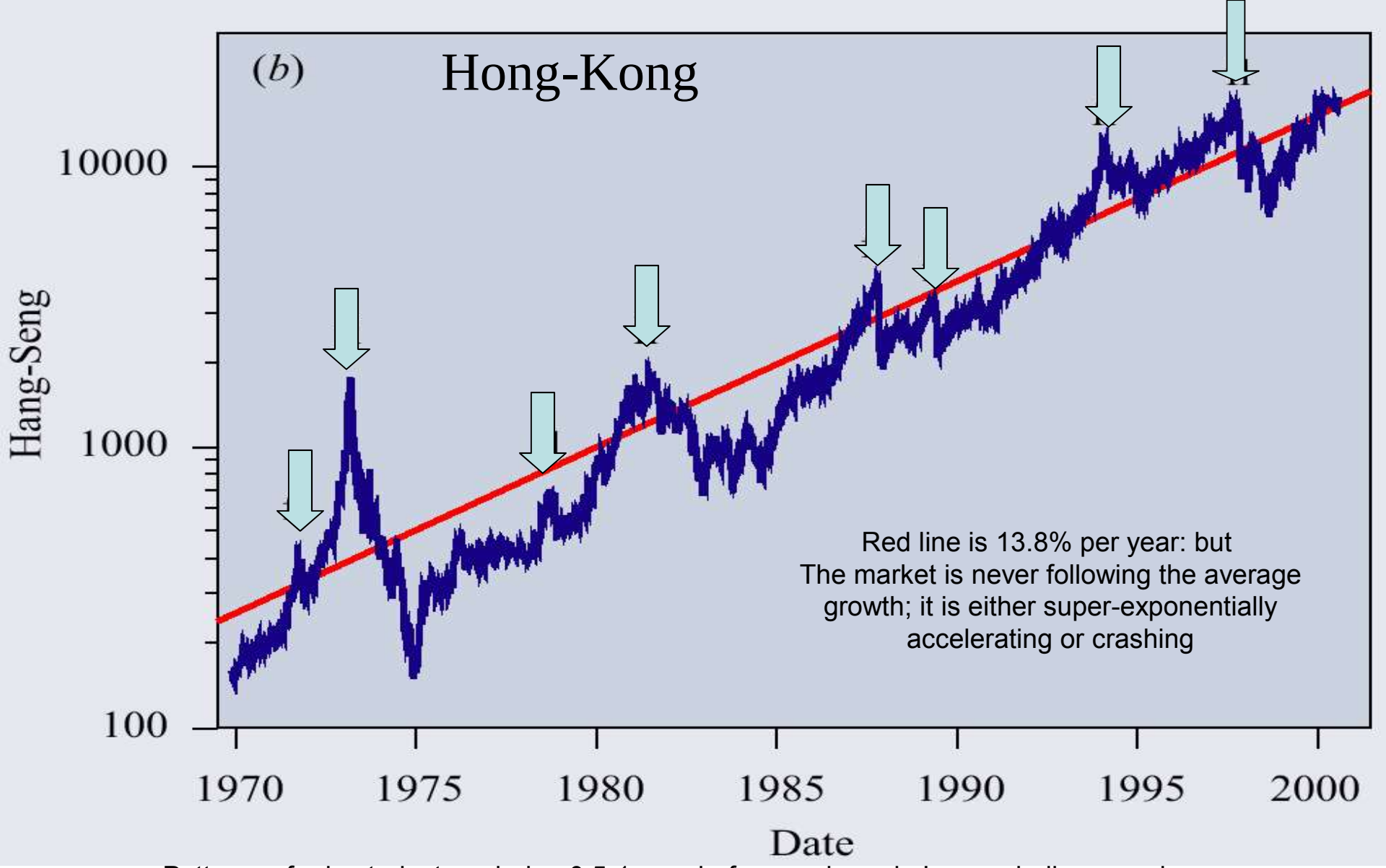
Gelfand et al (1976)

Trait: array of answers to set of questions

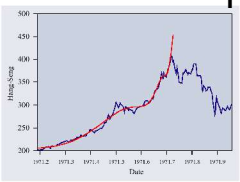
Feature: a treat which is frequent in class I and unfrequent in class II

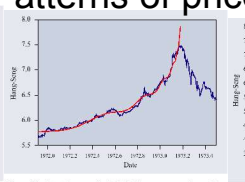
Alarm index(t): moving average of number of features at time t

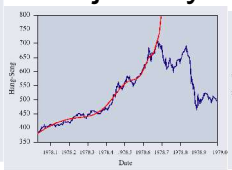

(a)
Textbook example of a series of superexponential acceleration followed by crashes
Arrows show peaks followed by corrections of more than 15% in less than three weeks
Hang-Seng
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
1970
1975
1980
1985
1990
1995
2000
Date

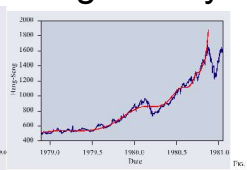


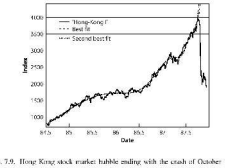
Patterns of price trajectory during 0.5-1 year before each peak: Log-periodic power law

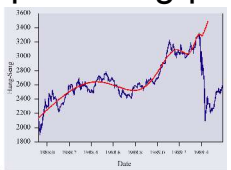


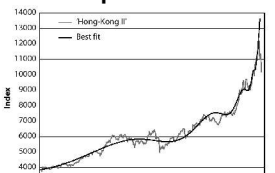


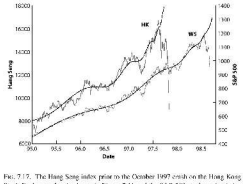












Rational Expectation Bubbles and Crashes

Martingale hypothesis ("no free lunch"):

$$\text{for all } \; t' > t \qquad \mathrm{E}_t[p(t')] = p(t)$$

If crashes are depletions of bubbles:

$$dp = \mu(t)\, p(t)\, dt - \kappa[p(t) - p_1]dj$$

Martingale gives

h(t)=E[dj]

$$\mu(t)p(t) = \kappa[p(t) - p_1]h(t) \ ,$$

i.e., if crash hazard rate $h(t)$ increases, so must the return (bounded rationality)

A. Johansen, D. Sornette and O. Ledoit
Predicting Financial Crashes using discrete scale invariance, Journal of Risk 1 (4), 4, 5-32 (1999)

A. Johansen, O. Ledoit and D. Sornette, Crashes as critical points, International Journal of Theoretical and Applied Finance 3 (2), 219-255 (2000)

Mechanisms for positive feedbacks in the stock market

- Technical mechanisms
  1. Option hedging
  2. Insurance portfolio strategies
  3. Trend following investment strategies

- Behavioral mechanisms
  1. It is rational to imitate
  2. It is the highest cognitive task to imitate
  3. We mostly learn by imitation

Importance of Positive Feedbacks and Over-confidence in a Self-Fulfilling Ising Model of Financial Markets

$$s_i(t) = \text{sign}\left[\sum_{j\in\mathcal{N}} K_{ij}(t)\mathrm{E}[s_j](t) + \sigma_i(t)G(t) + \epsilon_i(t)\right]$$

Imitation

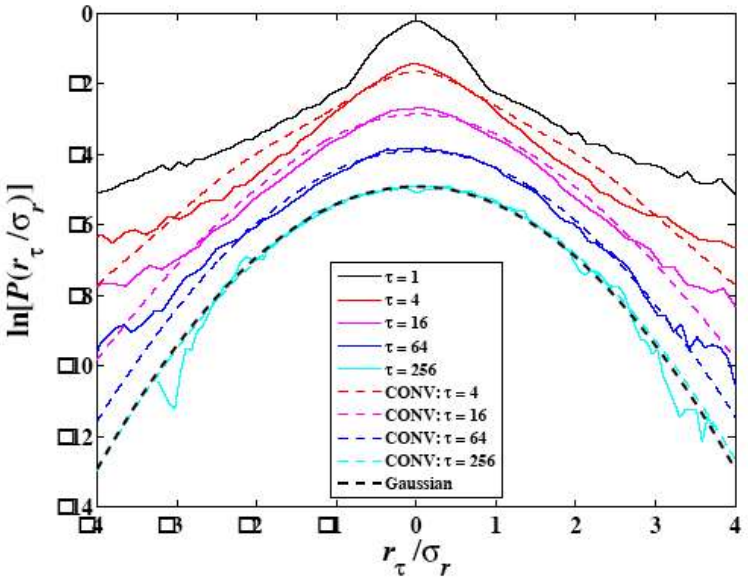
News

Private information

$$K_{ij}(t) = b_{ij} + \alpha_i K_{ij}(t-1) + \beta r(t-1)G(t-1)$$

β<0: rational agents

β>0: over-confident agents



Didier Sornette and Wei-Xing Zhou
Journal of Economic Behavior and Organization
(http://arxiv.org/abs/cond-mat/0503607)

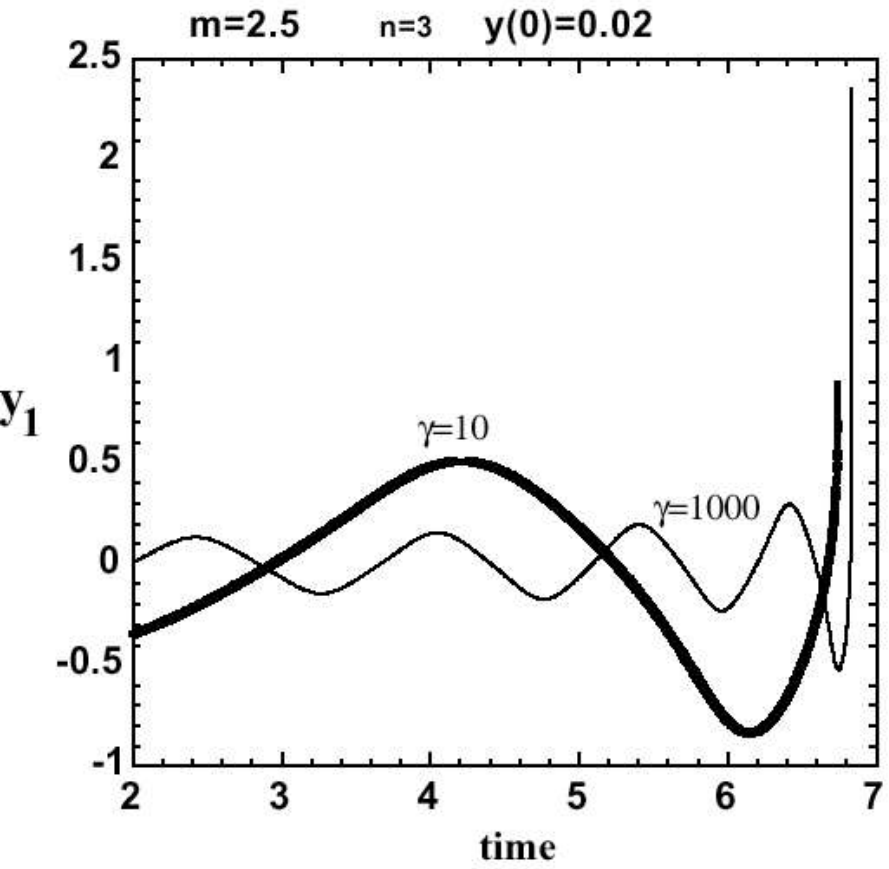
Oscillatory Finite-Time Singularities in Finance, Population and Rupture

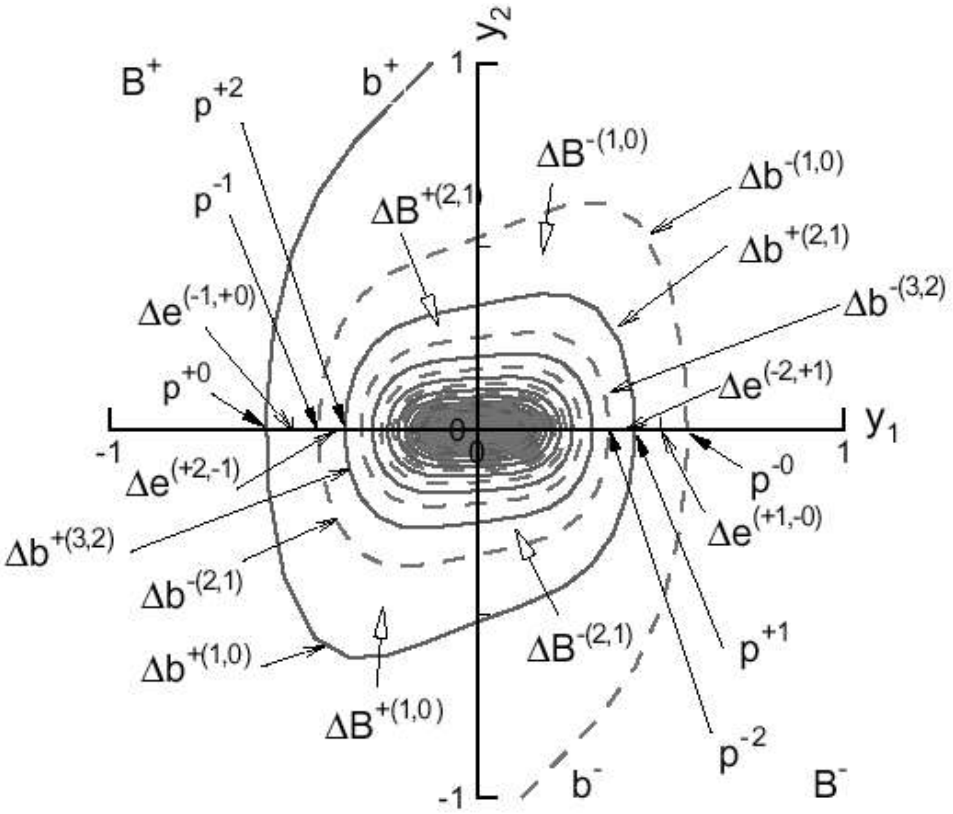
Non-linear fundamental value strategies

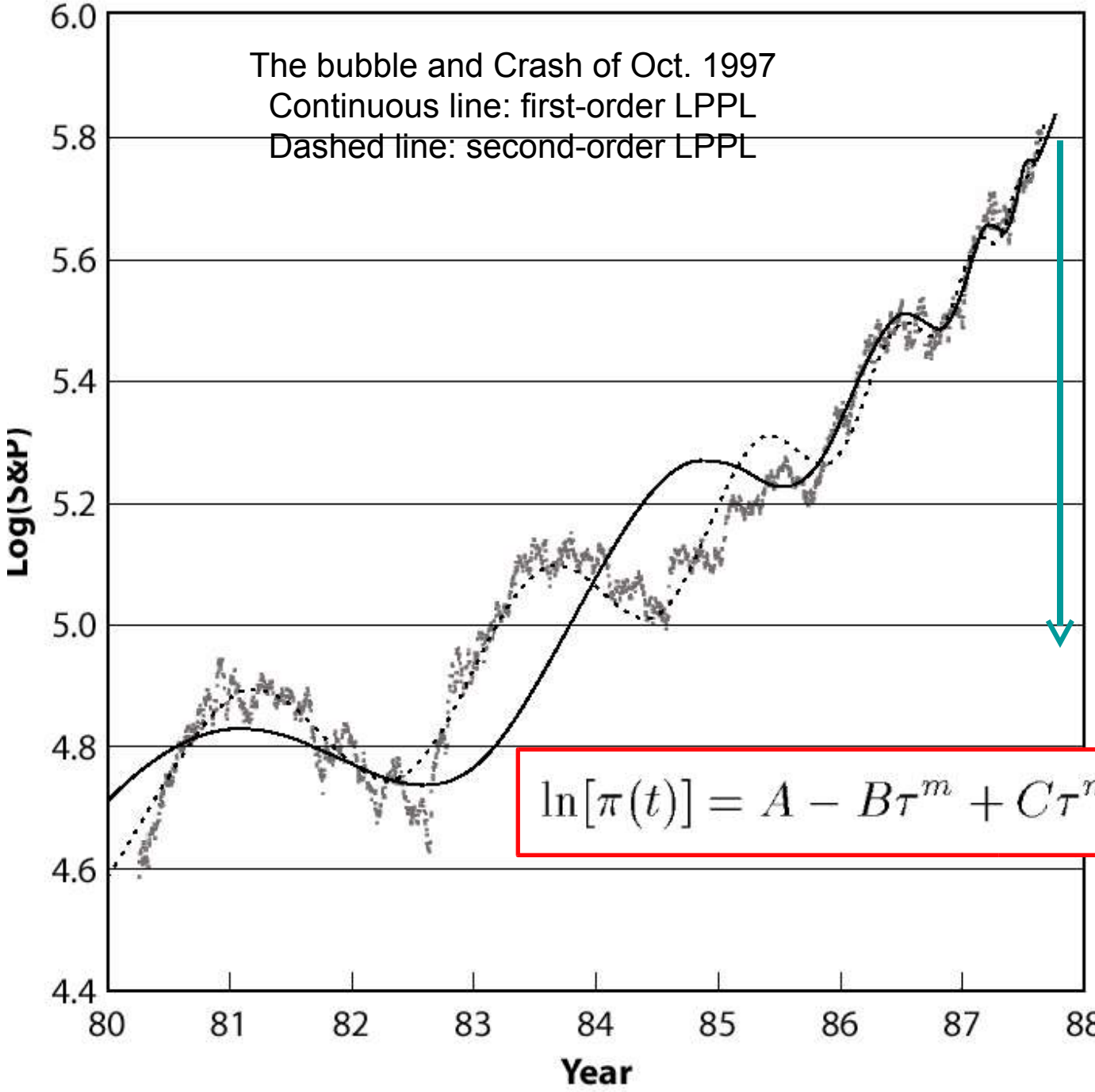
Non-linear technical analysis strategies

Inertia

K. Ide and D. Sornette
Oscillatory Finite-Time Singularities in Finance, Population and Rupture,
Physica A 307 (1-2), 63-106 (2002)

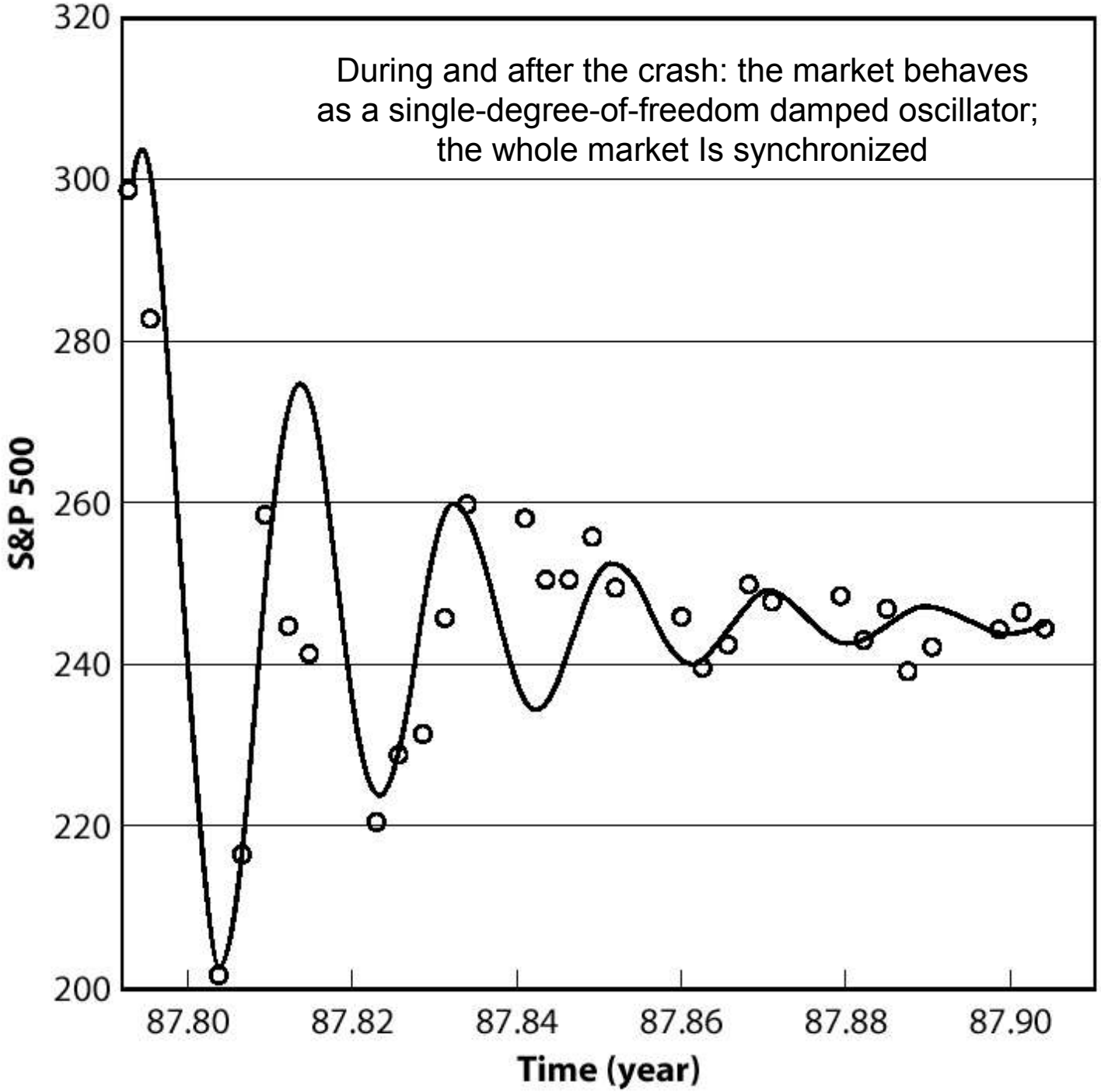


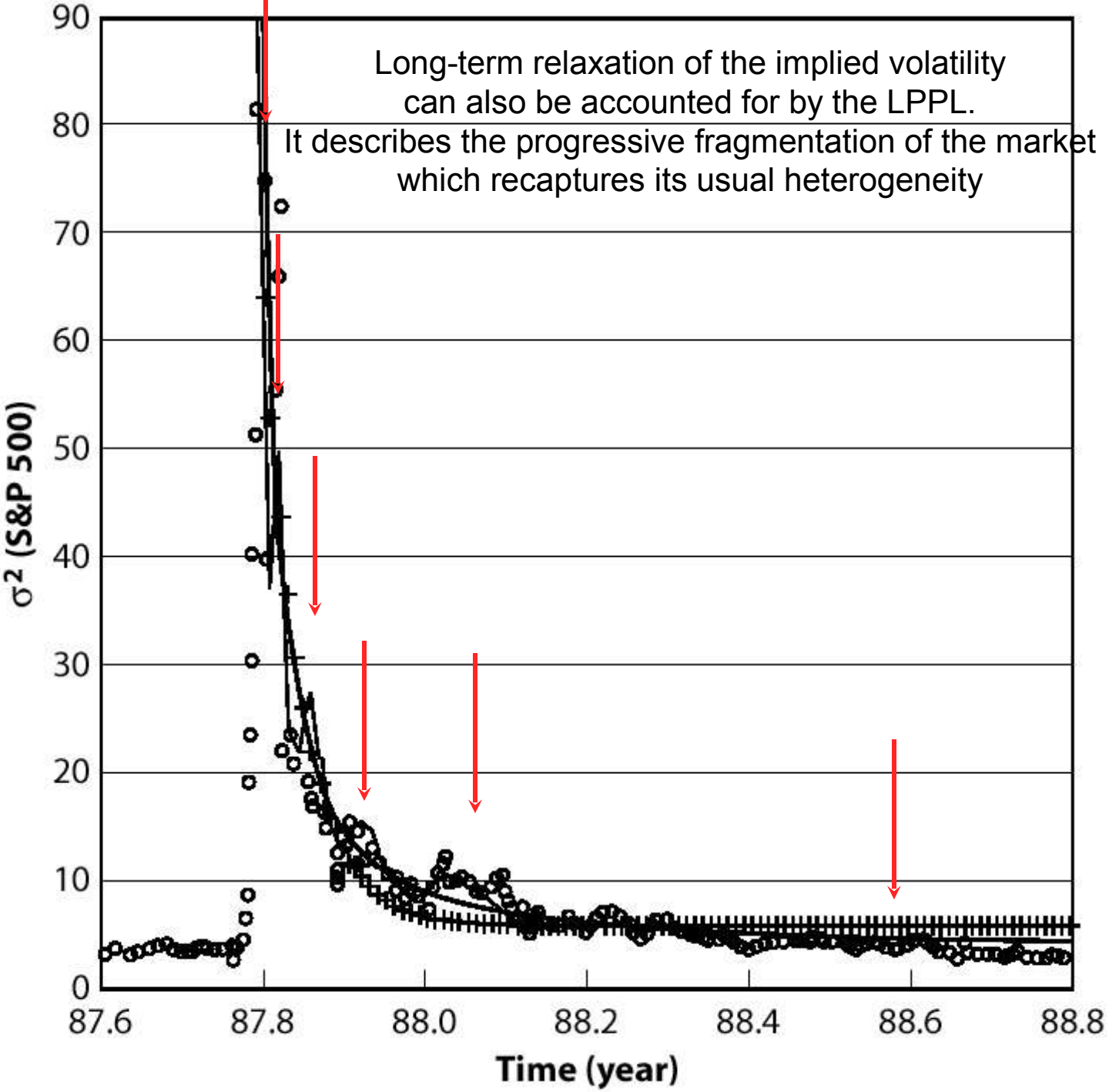




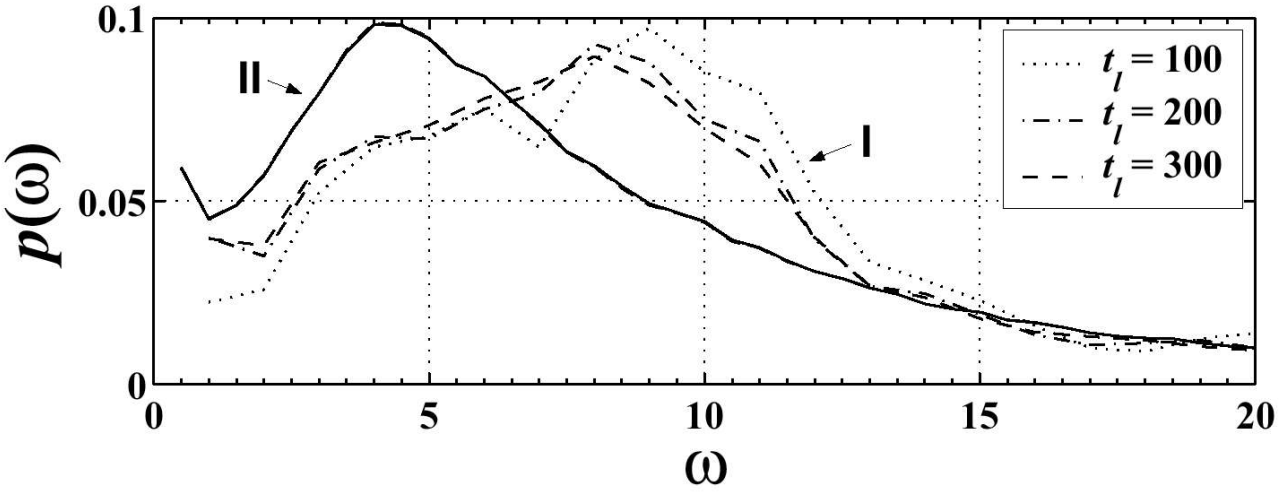
$$\ln[\pi(t)] = A - B\tau^m + C\tau^m \cos[\omega \ln(\tau) - \phi]$$

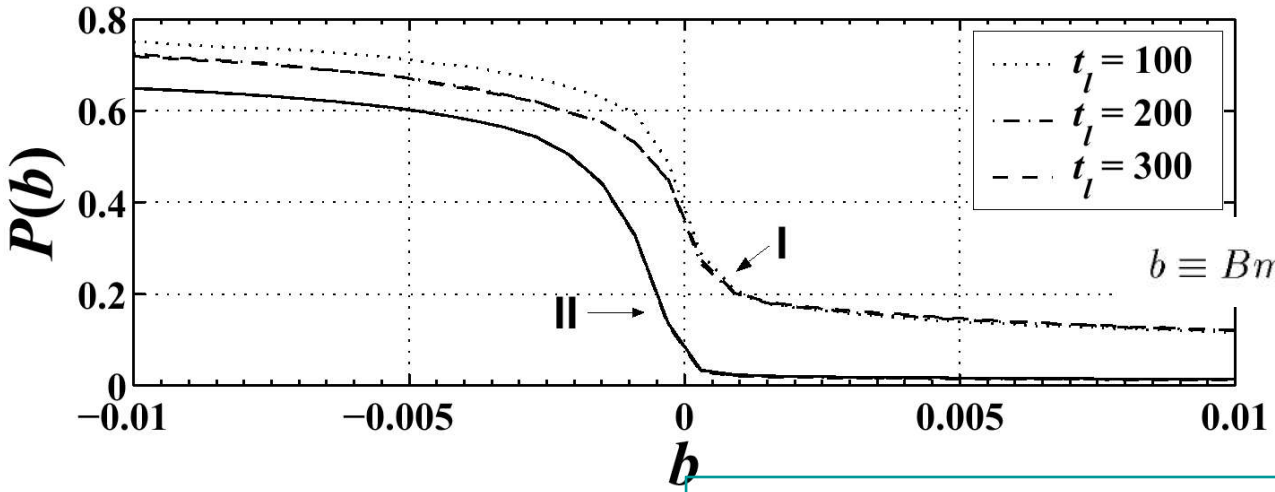
$$\tau = t_c - t$$


During and after the crash: the market behaves
as a single-degree-of-freedom damped oscillator;
the whole market Is synchronized
320
300
280
260
240
220
200
S&P 500
87.80
87.82
87.84
87.86
87.88
87.90
Time (year)


Long-term relaxation of the implied volatility
can also be accounted for by the LPPL.
It describes the progressive fragmentation of the market
which recaptures its usual heterogeneity
σ² (S&P 500)
90
80
70
60
50
40
30
20
10
0
87.6
87.8
88.0
88.2
88.4
88.6
88.8
Time (year)

Determination of relevant "traits" that allow us to distinguish targets from non targets in the Learning process





$h(t) \geq 0$

$b \equiv Bm - |C|\sqrt{m^2 + \omega^2} \geq 0$

Parameter for positivity of crash hazard rate

Figure 1: Density distribution $p(\omega|\mathrm{I\ or\ II})$ of the DSI parameter ω obtained from (1) and complementary cumulative distribution $P(b|\mathrm{I\ or\ II})$ of the constraint parameter b obtained from (2) for the objects in classes I (dotted, dashed, and dotted-dashed) and II (continuous) for three different values of t_l.

Multi-scale approach to critical times

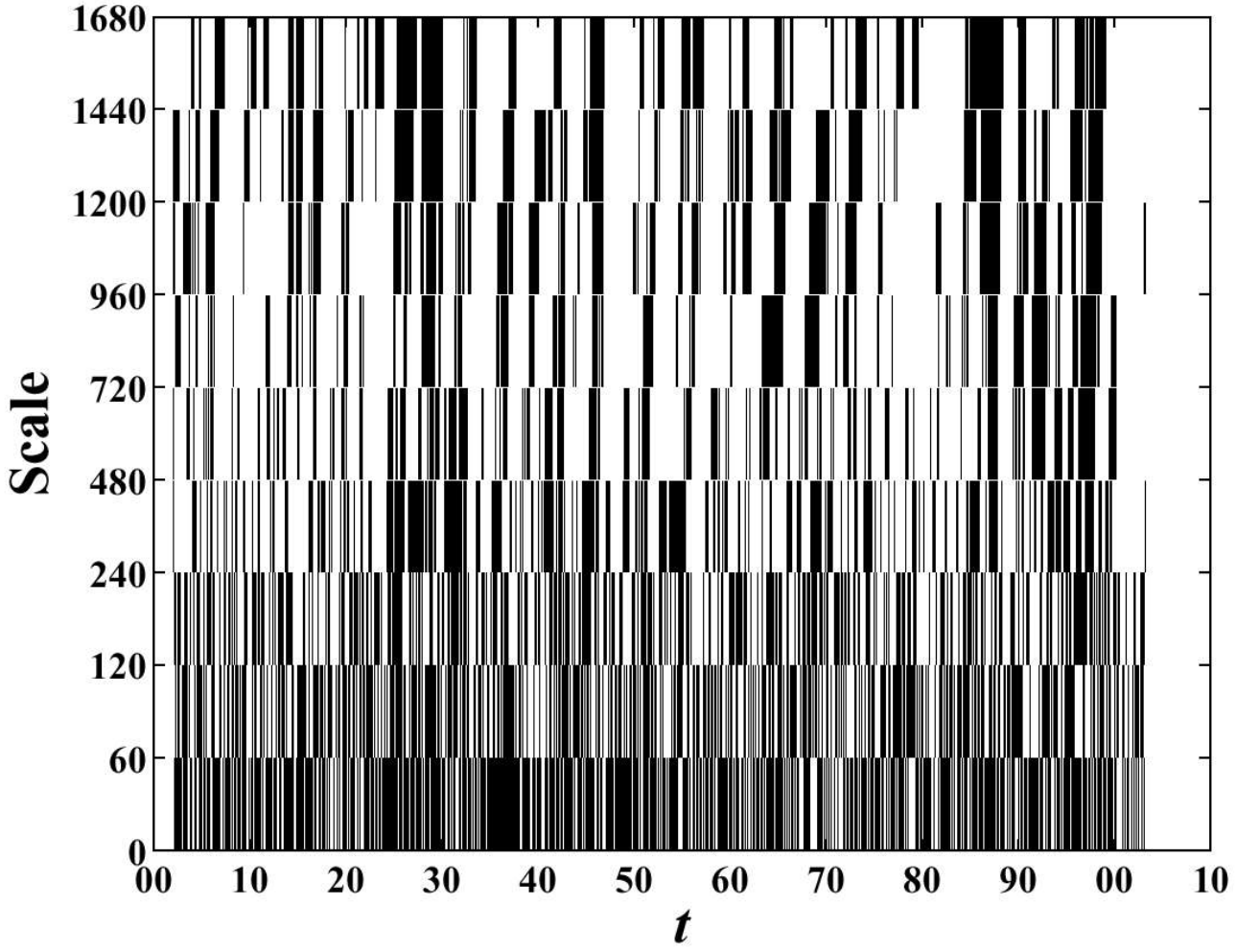


Figure 2: Alarm times t (or dangerous objects) obtained by the multiscale analysis. The alarms satisfy $b \geq 0$, $6 \leq \omega \leq 13$ and $0.1 \leq m \leq 0.9$ simultaneously. The ordinate is the investigation "scale" in trading day unit. The results are robust with reasonable changes of these bounds.

SYNOPSIS OF THE PATTERN RECOGNITION METHOD

- We select a few targets (the three well-known speculative bubbles ending in a crash at Oct. 1929, Oct. 1987 and Aug. 1998 for the US index and two targets for the Hang Seng), which serve to train our system (independently for the two indices).
- An object is defined simply as a trading day ending a block of trading days of a pre-defined duration.
- Those objects which are in a neighborhood of the crashes of the targets are defined to belong to class I. All other objects are said to belong to class II.
- For each object, we fit the price trajectory with expression (1) over a given time scale (defining the duration of the block of days ending with the object used in the fit) and obtain the corresponding values of the parameters, which are considered as characteristic of the object. This step is repeated nine times, once for each of the nine time scales used in the analysis. We keep 5 parameters of the fit by expression (1) for each time scale, thus giving a total of 5×9 parameters characterizing each object.

- We construct the probability density functions (pdf) of each parameter over all objects of class I and of class II separately. Those parameters which are found to exhibit sufficiently distinct pdf's over the two classes are kept as being sufficiently discriminating. In this way, out of the total of 45 parameters, 31 are kept.

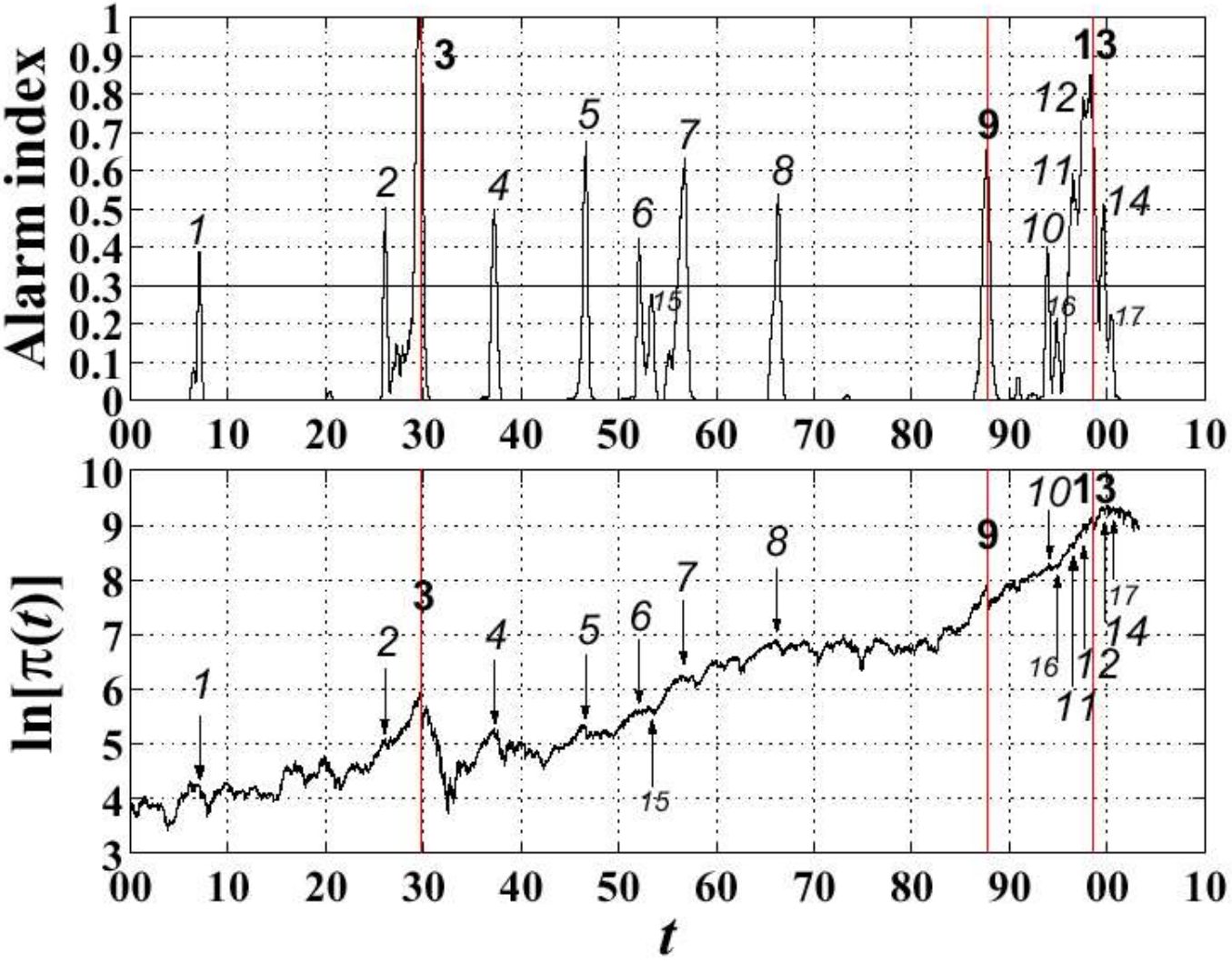
- Each selected parameter gives a binary information, Y or N, as whether its value for a given object falls within a qualifying interval for class I.

- As a compromise between robustness and an exhausive description, we group the parameters in triplets (called traits) to obtain a characterization of objects. Ideally, one would like to use all 31 parameters simultaneously for each object, but the curse of dimensionality prevents doing this.

- We study the statistics of all traits and look for those which are frequent in objects of class I and unfrequent in objects of class II. Such traits are called features.

- The Alarm Index at a given time is defined as a moving average number of distinctive features found at that time. Large values of the Alarm Index are proposed to be predictors of changes of regime in the stock market.

Multiscale Pattern Recognition Method



D. Sornette and W.-X. Zhou, Predictability of Large Future Changes in Complex Systems, (http://arXiv.org/abs/cond-mat/0304601)

Extension to a multi-scale LPPL analysis with Gelfand's method of pattern recognition to predict

Figure 3: (Color online) Alarm index $AI(t)$ (upper panel) and the DJIA index from 1900 to 2003 (lower panel). The peaks of the alarm index occur at times indicated by arrows in the bottom panel.

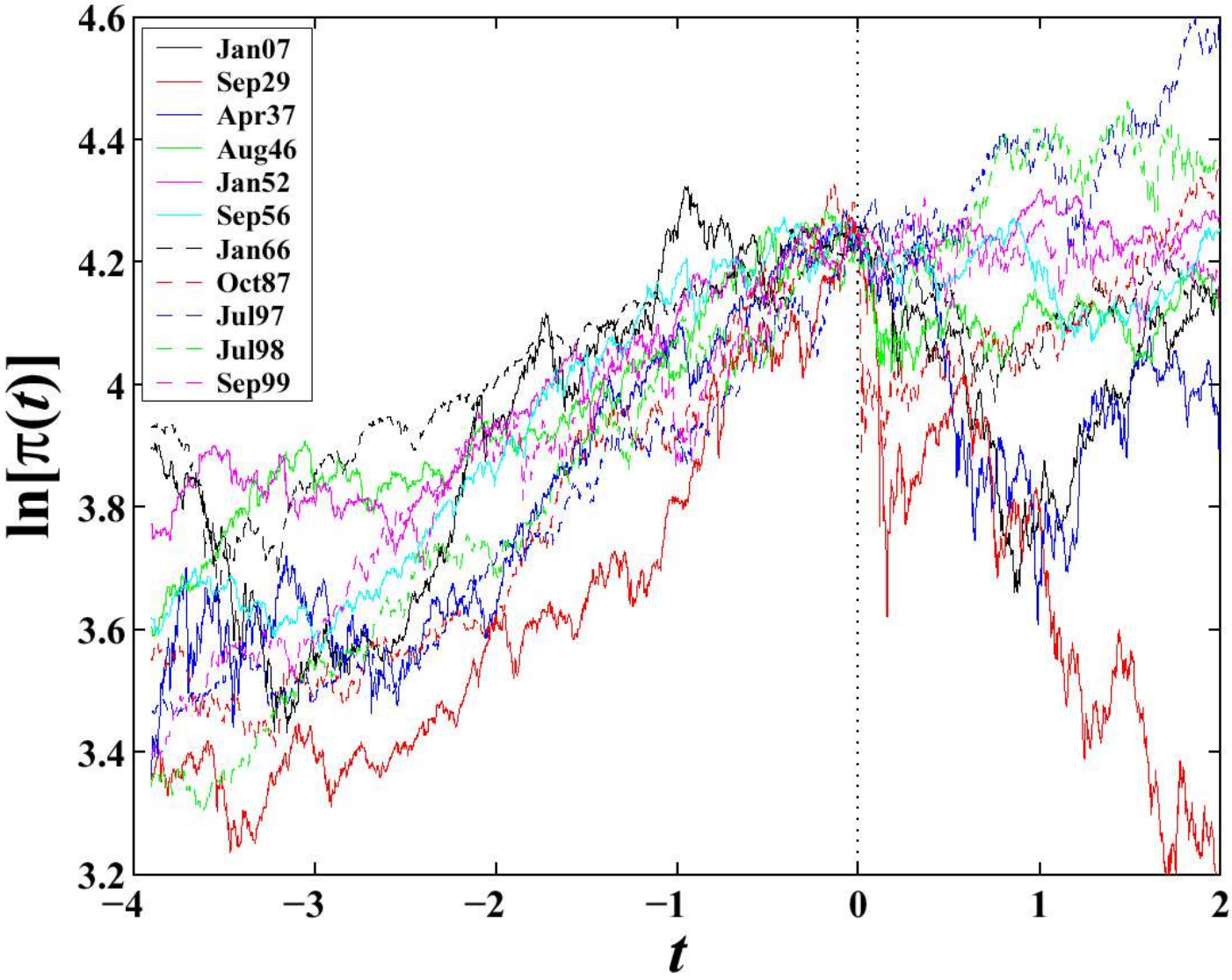
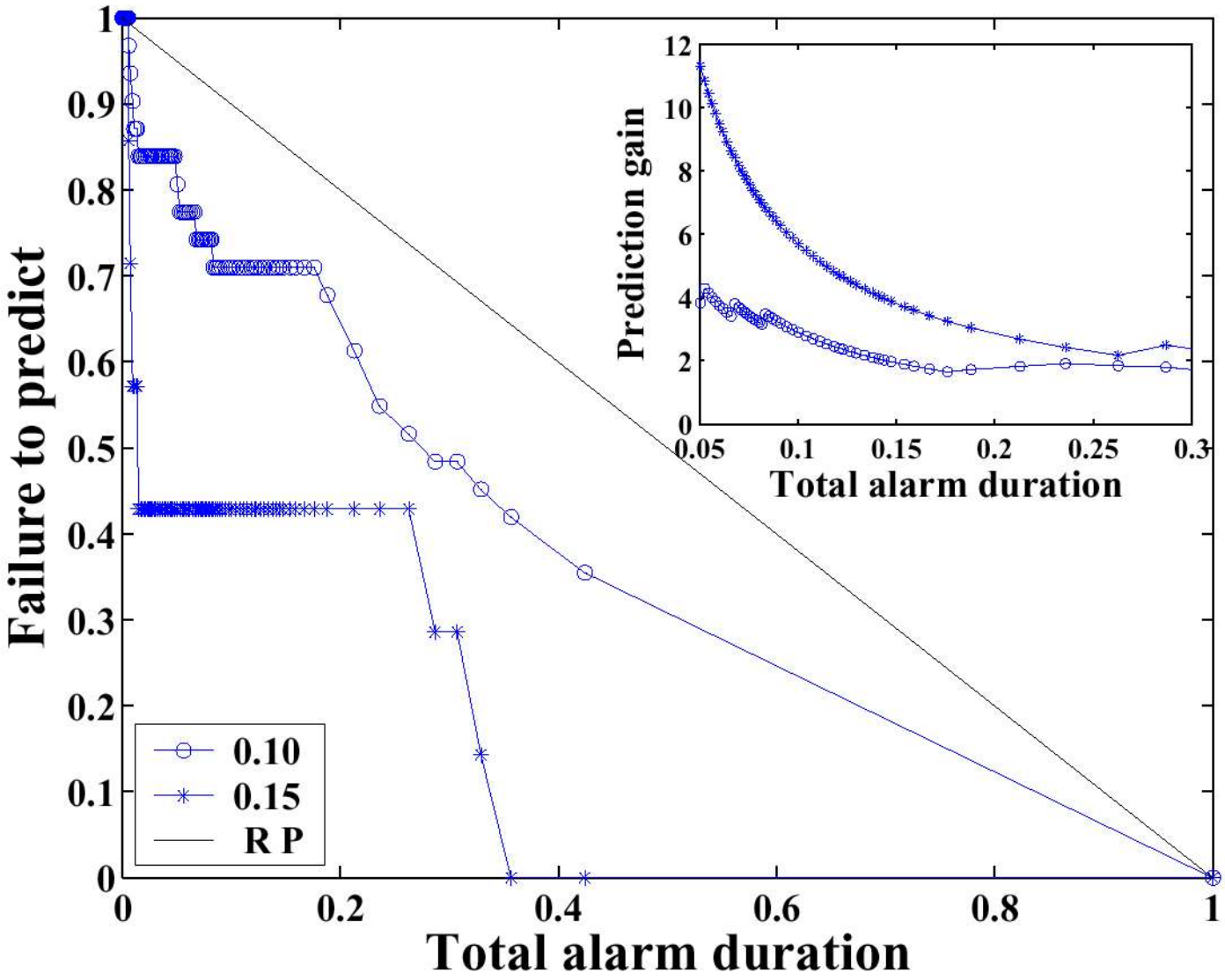


Figure 4: (Color online) Superposed epoch analysis of the 11 time intervals, each of 6 years long, of the DJIA index centered on the time of the maxima of the 11 predictor peaks above $AI = 0.3$ of the alarm index shown in Fig. 3.



We obtain very significant prediction gains

Figure 5: Error diagram for our predictions for two definitions of targets to be predicted $r_0 = 0.1$ and $r_0 = 0.15$ obtained for the DJIA. The anti-diagonal line corresponds to the random prediction result. The inset shows the prediction gain.

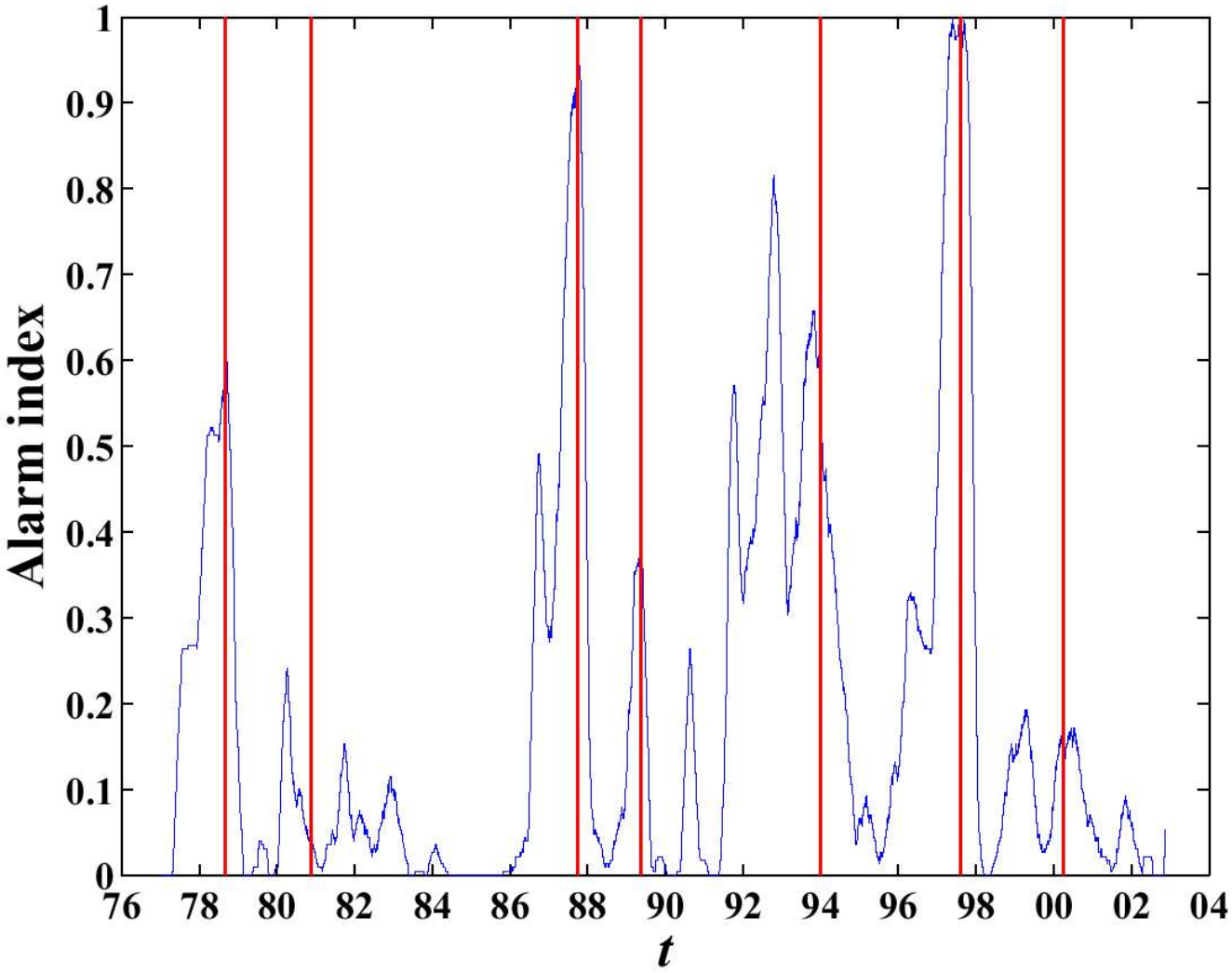


Figure 6: (Color online) Alarm index $AI(t)$ constructed by our algorithm for the Hong Kong Hang Seng composite index since 24-Nov-1969. The vertical lines indicate the timing of the seven crashes that have not been used in the training part. Note that the first two crashes are not included in the analysis since the longest window used in our multiscale analysis is seven years.

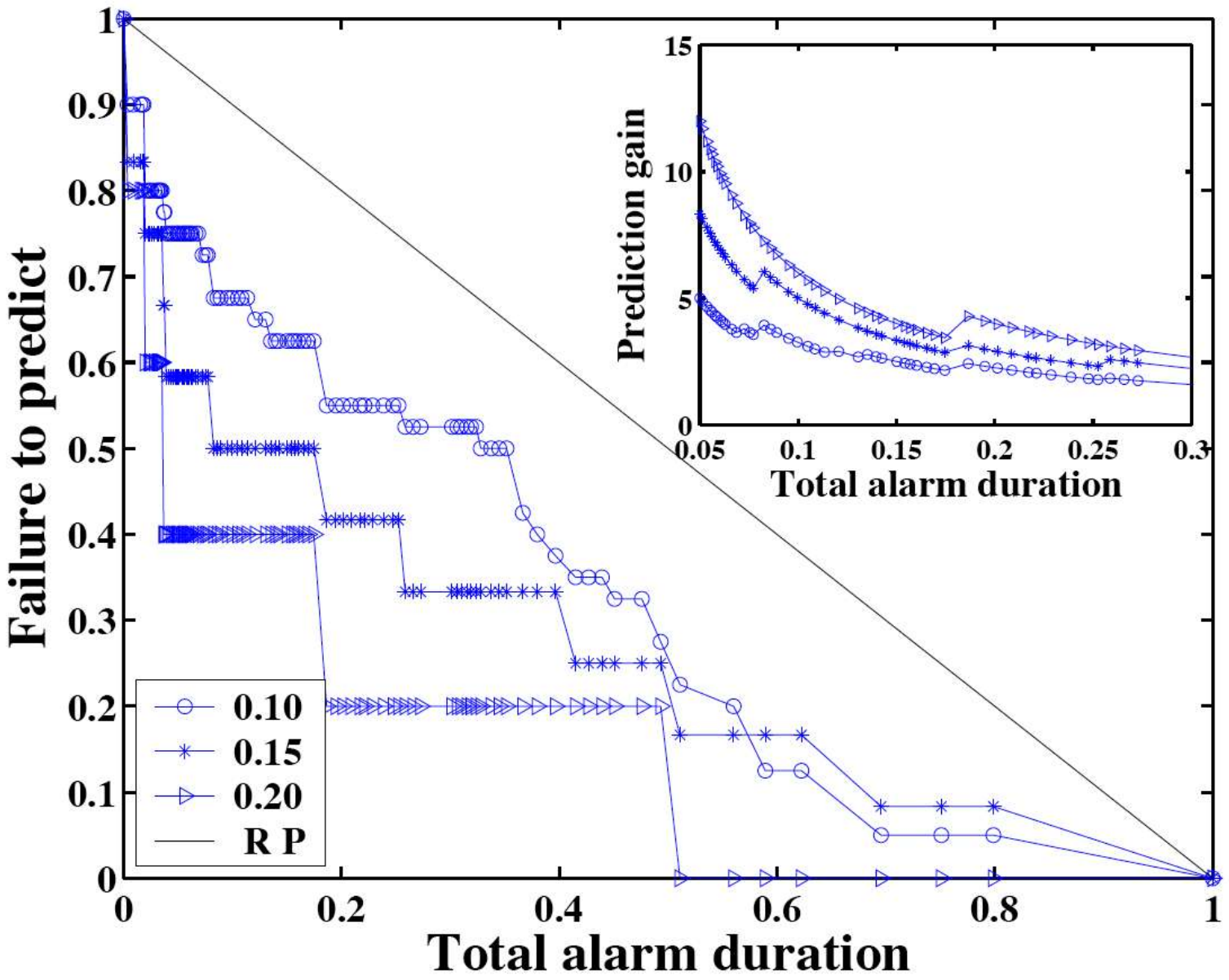


Figure 8: Error diagram of the predictions for three definitions of targets to be predicted with $r_0 = 0.1$, $r_0 = 0.15$, and $r_0 = 0.2$ in regard to HSI. The diagonal line is a random prediction. In the inset shows the prediction gains.

COLLECTIVE BEHAVIOR between AGENTS
(with negative and positive feedbacks)

Basle Committee on Banking Supervision: "In handling systemic issues, it will be necessary to address, on the one hand, risks to confidence in the financial system and contagion to otherwise sound institutions, and, on the other hand, the need to minimize the distortion of market signals and discipline."

A. Greenspan (Aug., 30, 2002):
"We, at the Federal Reserve…recognized that, despite our suspicions, it was very difficult to definitively identify a bubble until after the fact, that is, when its bursting confirmed its existence… Moreover, it was far from obvious that bubbles, even if identified early, could be preempted short of the Central Bank inducing a substantial contraction in economic activity, the very outcome we would be seeking to avoid."

Our conclusion is that the presence of the bubble and its approximate end was predictable.

Challenges

- Small data sets
- Nonlinear dependences vs zero correlation
- Stochastic reconstructions
- Non-stationarity
  - Changes of regime (regime-switching)
  - Conventions
  - Irreversibility, learning, evolution
- Theory of decision making