

Korrelationssumme:

$$C_m(\varepsilon) = \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{i=1}^{N-1} \sum_{j=i+1}^N H(\varepsilon - |\mathbf{v}_m(i) - \mathbf{v}_m(j)|)$$

$H = \text{Heaviside-Funktion} \begin{cases} H(a) = 0, a < 0 \\ H(a) = 1, \text{sonst} \end{cases}$

\Rightarrow Korrelationsdimension:

$$D_2 = \lim_{\varepsilon \rightarrow 0} \frac{\log C(\varepsilon)}{\log \varepsilon}$$

Annahme:

linearer Zusammenhang

Realität:

hochgradig nichtlinear

Beeinflussende Faktoren

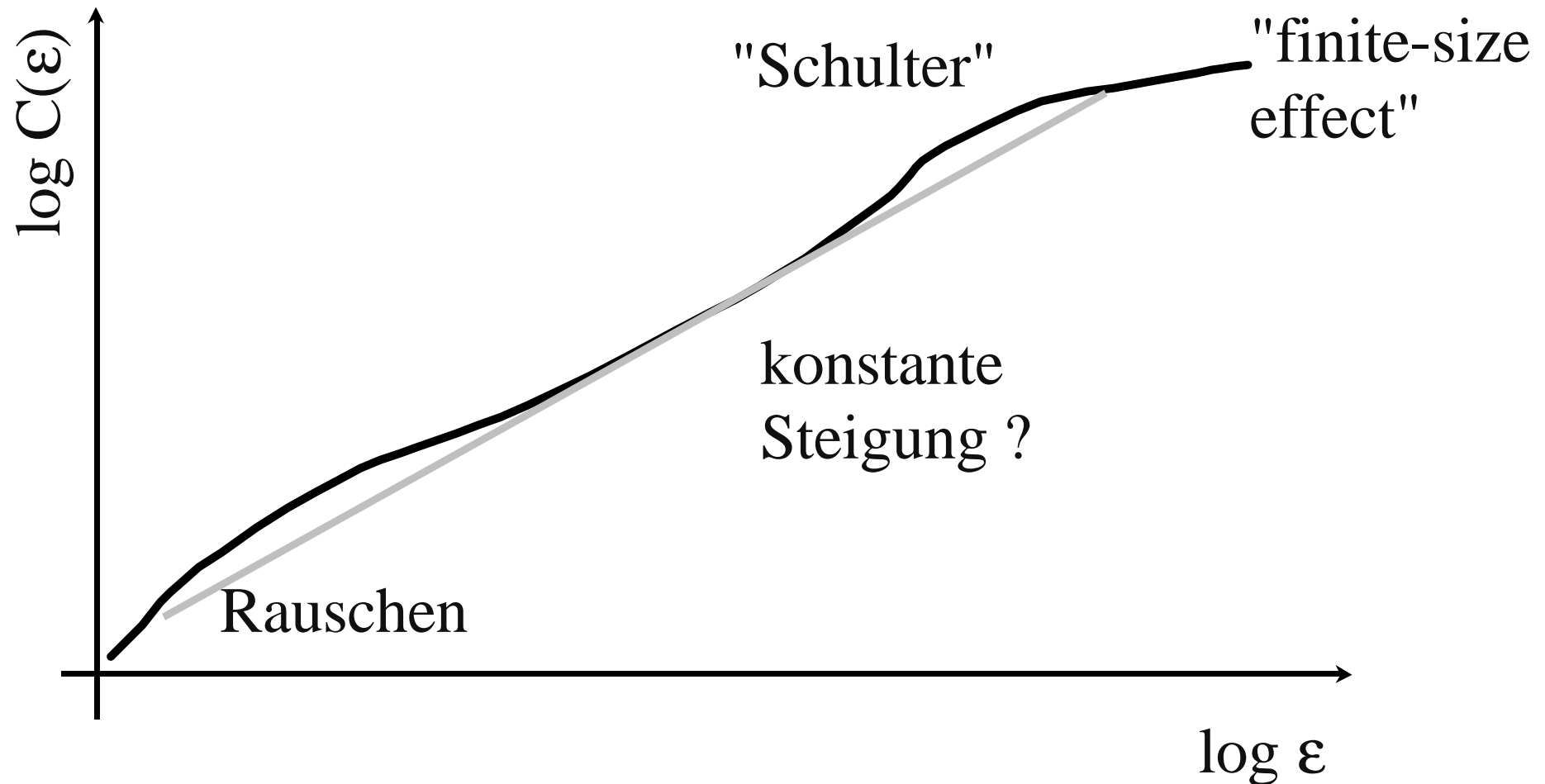
Korrelationsdimension

- Anzahl der Datenpunkte / Besetzungsdichte / Stationarität
- Rauschen (weiß, farbig) / Korrelationen mit Signal
- Datengenauigkeit
- Einbettungsdimension
- Zeitverzögerung / multivariate Rekonstruktion
- zeitliche Korrelationen (Theiler-Korrektur)

Beeinflussende Faktoren

Korrelationsdimension

Abschätzung der Korrelationsdimension aus Korrelationssumme:



Abschätzung der Korrelationsdimension aus Korrelationssumme:

traditionelles Regressionsmodell:

$$\log C(\varepsilon) = \alpha + \beta \log \varepsilon + u \quad \beta \approx D_2$$

u : additiver Fehlerterm, normalverteilt, unabh. von ε

$$\text{var}(u) \Rightarrow \text{var}(\beta)$$

adäquateres Regressionsmodell:

random coefficient regression model

$$\log C(\varepsilon) = \alpha(\varepsilon | N, m) + \beta(\varepsilon | N, m) \log \varepsilon + u$$

Beeinflussende Faktoren

Korrelationsdimension

- *random coefficient regression model* im strengen Sinne falsch
- konstanter Term α abh. von Einbettungsdimension m

$$C(\varepsilon) \cong \text{const.} \varepsilon^{D_2} e^{-m\tau K_2}$$

\Rightarrow

$$\log C(\varepsilon) \cong \underbrace{\log(\text{const.})}_{\alpha} - m \underbrace{\tau K_2}_{\alpha} + \underbrace{D_2}_{\beta} \log \varepsilon$$

$\Rightarrow \alpha$ monoton fallend mit m für $K_2 \neq 0$

Gründe: Besetzungsdichte, Fraktalität, Lakunarität

Ramsey & Yuan (Dept. of Economics, NYU)

Phys Lett A 134, 287, 1989

Untersuchung der Fehler bei Abschätzung der Korrelationsdimension aus Korrelationssumme:

- einfache Modellsysteme (Hénon, 5-Torus)
- Rauschen (uniform, normal, exponentiell)
- Variation von N und m
- τ fest
- nichtlineares Regressionsmodell (N, m):

$$\ln(\text{const.} + D_2) = \gamma_1 + \gamma_2 N^{\gamma_3} + \gamma_4 N^{\gamma_5} \left(e^{-\frac{\gamma_6}{m^{\gamma_7}}} - 1 \right)$$

Ergebnisse der Untersuchungen von Ramsey und Yuan:

- Modell adäquat ($R^2 > 0,999$)

- (1) Überschätzung der Dimension von Attraktoren
- (2) Unterschätzung der Dimension von Rauschen
- (3) Verringerung des Fehlers mit steigender Datenpunktanzahl
- (4) Vergrößerung des Fehlers mit steigender Einbettungsdimension
- (5) Varianz der Dimensionsschätzung:
nichtlineares/lineares Regressionsmodell $\approx 1/64$
- (6) Varianzverringerung mit steigender Datenpunktanzahl
- (7) Varianzvergrößerung mit steigender Einbettungsdimension

Beeinflussende Faktoren

Korrelationsentropie

Wiederholung
importance sampling

falls Skalierungsregion nachweisbar:

$$\varepsilon^{D_q} \approx \text{const.}$$

Korrelations-Entropie ($q=2$) dann bestimmbar gemäß:

$$h_q(m) = H_q(m+1, \varepsilon) - H_q(m, \varepsilon) = \lim_{\varepsilon \rightarrow 0} \lim_{m \rightarrow \infty} \log \left[\frac{C_m(\varepsilon)}{C_{m+1}(\varepsilon)} \right] = K_2$$

Korrelationsentropie aus Korrelationssumme ableitbar

⇒ vergleichbare Probleme wie bei Korrelationsdimension

- $N \rightarrow \infty, \varepsilon \rightarrow 0$
- Skalierungsregion

zusätzlich:

- $m \rightarrow \infty$

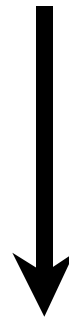
(Verlust von ε^{-K_2} Nachbarn beim Übergang von $m \rightarrow m+1$)

⇒ vergleichbare (wenn nicht schlimmere) Über-/Unterschätzungen der Korrelationsentropie

Beeinflussende Faktoren

Lyapunov-Exponenten

Spektrum der Lyapunov-Exponenten i.A. schwierig zu bestimmen
(Besetzungsdichte, Rauschen, Fehlerfortpflanzung, ...)



Beschränkung auf größten Lyapunov-Exponenten

Wolf-Algorithmus: (Wolf et al. Physica D, 16, 285, 1985)

- verfolge erstes Paar *nächster Nachbarn* über *einige* Zeitschritte
- nach "Entwicklungszeit": finde *neue nächste Nachbarn*
Forderung: Abstand *klein genug*, Differenzvektor hat *ähnliche* Ausrichtung wie erstes Paar
- Wiederholung bis *Konvergenz*-Kriterium erreicht

- Berechnung des Lyapunov-Spektrums äquivalent
(betrachte Differenz(hyper)flächen statt Differenzvektoren)

Beeinflussende Faktoren

Lyapunov-Exponenten

begrenzte Leistungsfähigkeit des Wolf-Algorithmus
(wird noch vielfach verwendet !)

- schlechte Statistik (nur eine Differenz)
- schwer zu erfüllende Voraussetzungen
- sensible Abh. von Einbettungsdimension
- verhängnisvolle Fehlerfortpflanzung (Lyapunov-Spektrum)

neuere Algorithmen

(Rosenstein et al., Physica D, 65, 117, 1992; Kantz, Phys Lett A 185, 77, 1994)

- betrachte Tangentenvektoren $\mathbf{y}(t) = \frac{\Delta \mathbf{x}(t)}{|\Delta \mathbf{x}(0)|}$
- *Näherung* für λ_1

$$l_t \approx C e^{\lambda_1 t}$$

l : Lyapunov-Zahl, abh. von Zeitschritten t (Distanz zwischen anfänglich benachbarten Phasenraumpunkten)

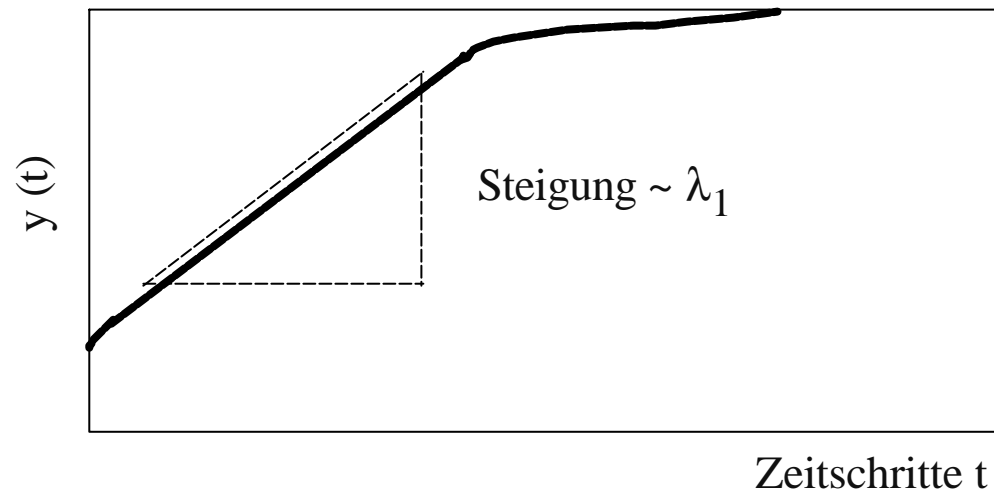
C : Normierungskonstante

ANNAHMEN: größter Lyapunov-Exponent dominiert Verhalten von l (andere Exponenten vernachlässigbar); nur endliche Distanzen; kein Grenzübergang $t \rightarrow \infty$

mit:

$$y(t) = \sum_{i=1}^{K_{Dist.}} \log l_i(t)$$

betrachte Distanz-Plot



weitere Verbesserungen:

- betrachte alle Nachbarn in gegebener ε -Umgebung
- Theiler-Korrektur

mehr dazu:

Jochen Wegner: Zeitreihenanalyse synthetischer Daten sowie hirnelektrischer Aktivität mit Hilfe des größten Lyapunov-Exponenten. Diplomarbeit (1998)

- *Standarddefinitionen* der Dimensionen, Lyapunov-Exponenten und Entropien haben bei realen Daten (N klein, Rauschen) *keine Bedeutung* mehr
- Rauschen dominiert auf kleinen Längenskalen
 - $\Rightarrow (D, \lambda, K) \rightarrow \infty$
- falsch Rauschen klein genug:
 - Skalenverhalten auf größeren Längenskalen beobachtbar
 - *effektive Maße* (erheblich Abweichung von wahren Maßen)

Beeinflussende Faktoren

Determinismus

Vektorfeld im Phasenraum



Parallelität
Kontinuität
Stetigkeit



← Statistik

Determinismus / Stochastizität

- Phasenraumrekonstruktion (m, τ)
- grobe Partitionierung (b^m Kuben)
- Kuben-Besetzungsdichte (N)
- Rauschen (weiß, farbig) / Korrelationen mit Signal
- zeitliche Korrelationen
- Stationarität

mehr dazu:

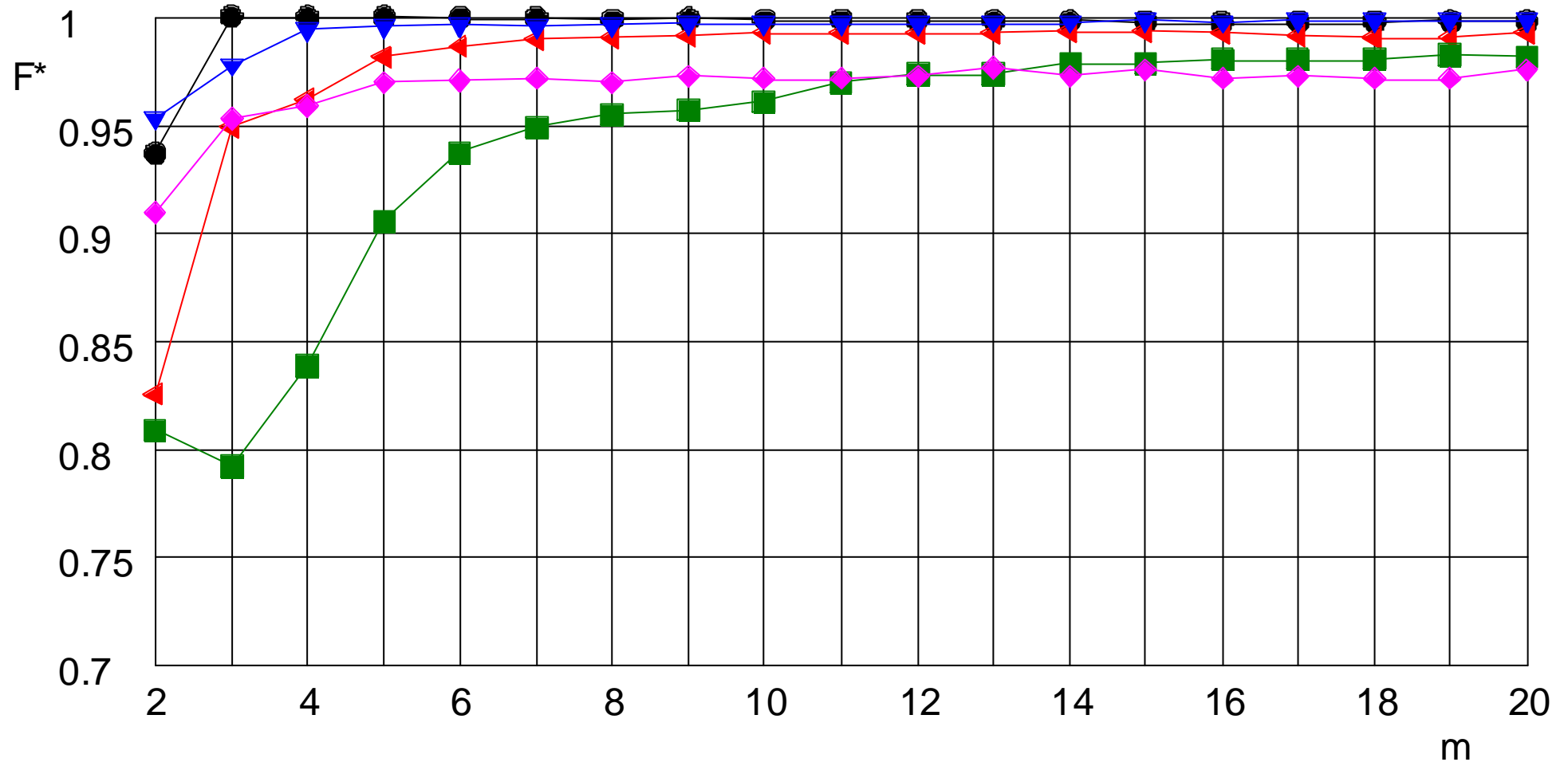
Ralph Andrzejak: Anteile nichtlinearer Determinismen in Zeitreihen:
Theorie und Simulation sowie Anwendung auf hirnelektrische Aktivität
von Patienten der prächirurgischen Epilepsiediagnostik.
Diplomarbeit (1997)

Beeinflussende Faktoren

Determinismus

Einbettungsdimension m ($N=4096$, $\tau=8$, $b=8$)

Modellsysteme



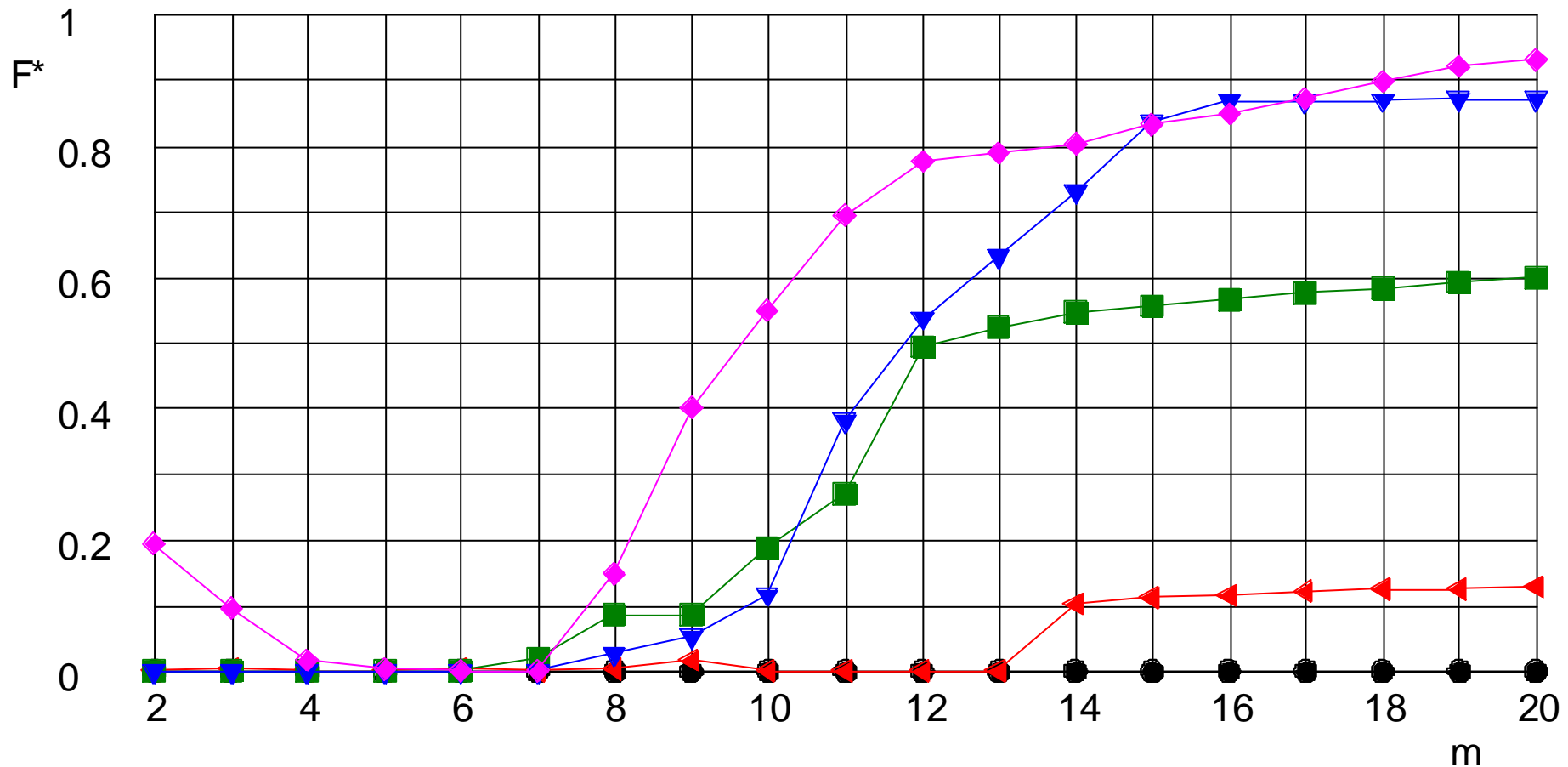
● Mackey-Glass-A ▲ Mackey-Glass-B ■ Mackey-Glass-C ▼ Rössler ◆ Lorenz

Beeinflussende Faktoren

Determinismus

Einbettungsdimension m ($N=4096$, $\tau=8$, $b=8$):
unzureichende Besetzungsdichte

Rauschen



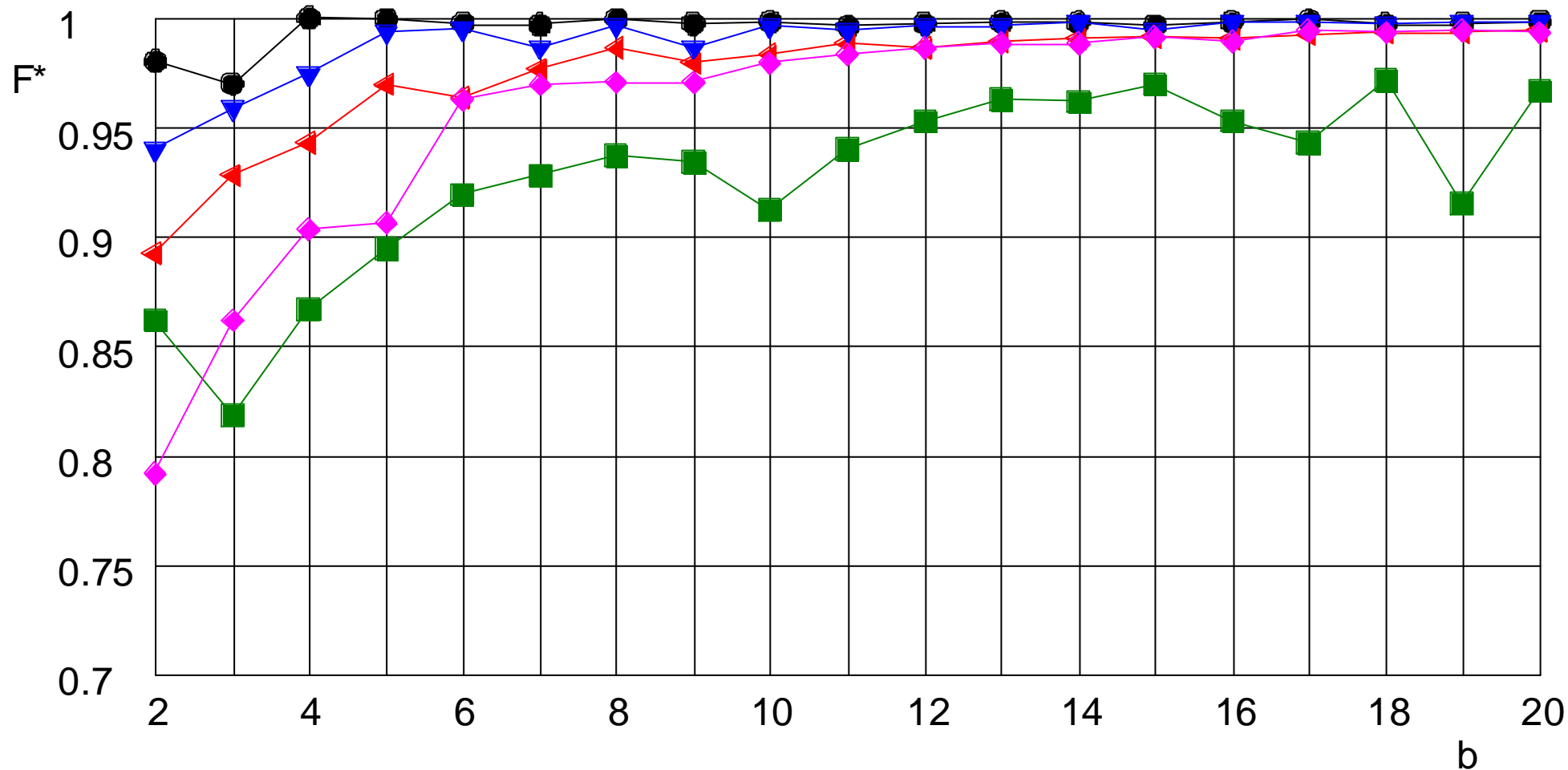
● Weiß ◀ 40 Hz ■ 20 Hz ▼ 10 Hz ◆ 5 Hz

Beeinflussende Faktoren

Determinismus

Kubenzahl b ($N=4096$, $\tau=8$, $m=6$)

Modellsysteme



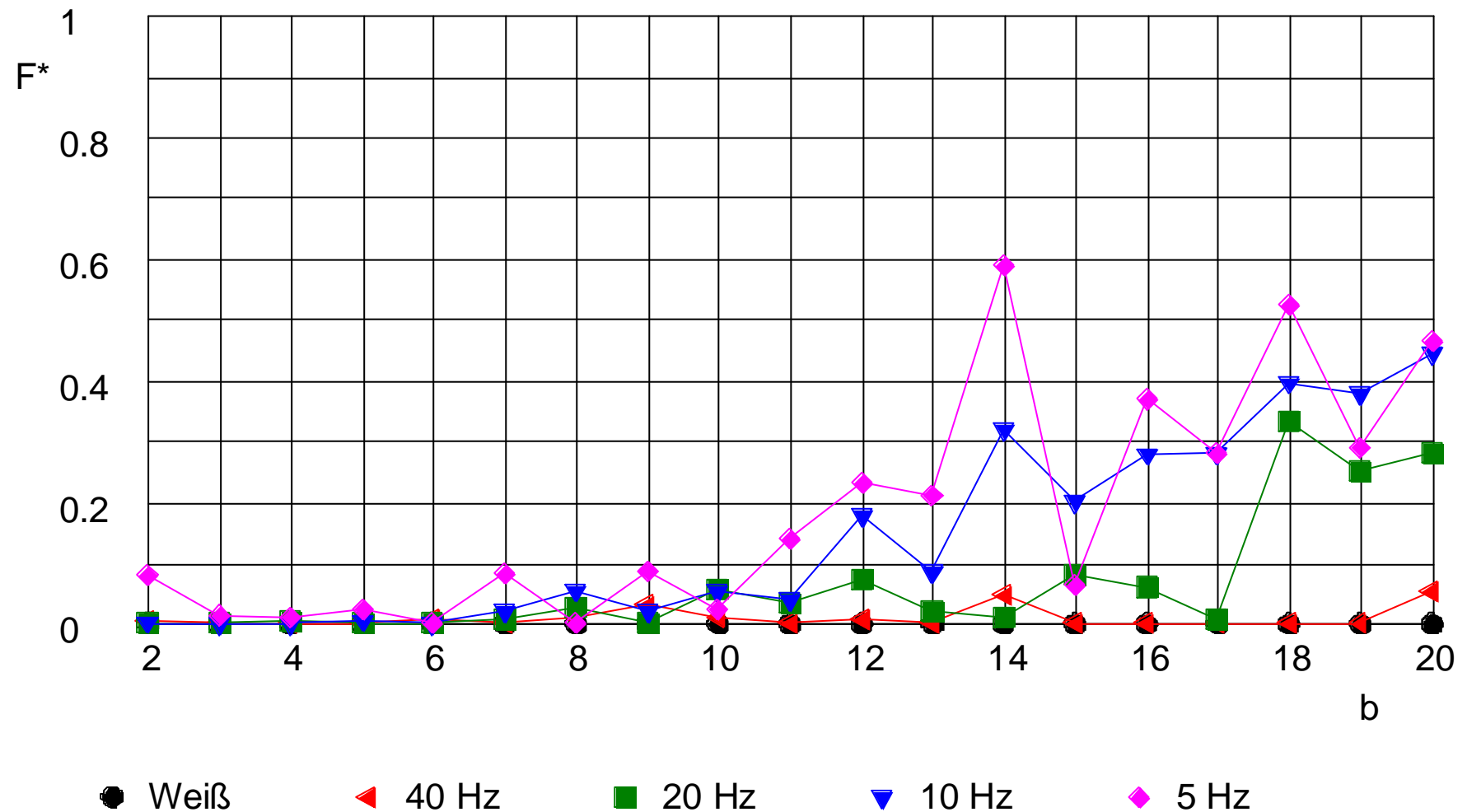
● Mackey-Glass-A ▲ Mackey-Glass-B ■ Mackey-Glass-C ▼ Rössler ◆ Lorenz

Beeinflussende Faktoren

Determinismus

Kubenzahl b ($N=4096$, $\tau=8$, $m=6$):
unzureichende Besetzungsdichte

Rauschen



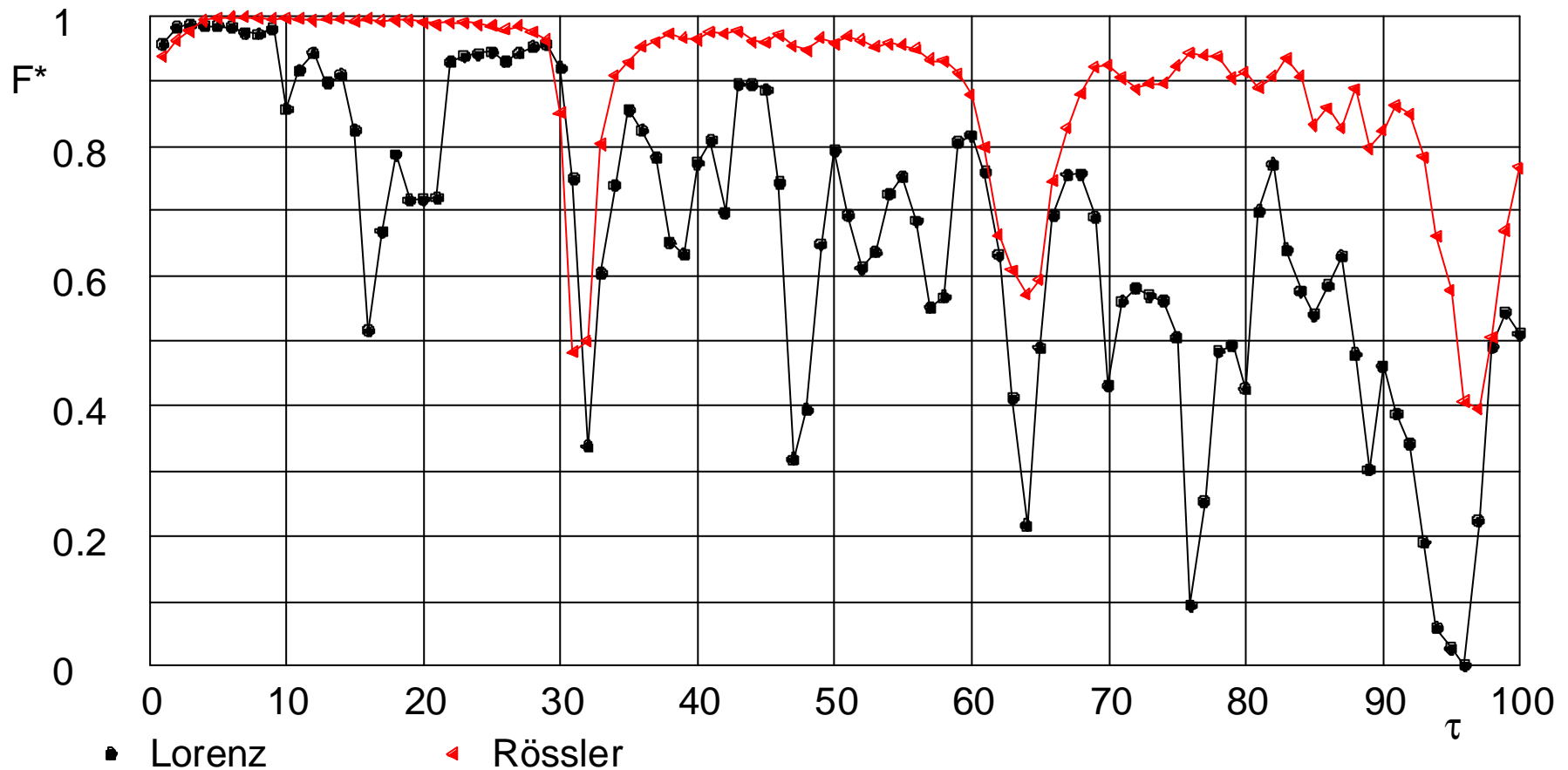
Beeinflussende Faktoren

Determinismus

Time-delay τ ($N=4096$, $m=6$, $b=8$)

Modellsysteme

Minima bei ganzzahligen Vielfachen der System-Periodendauer



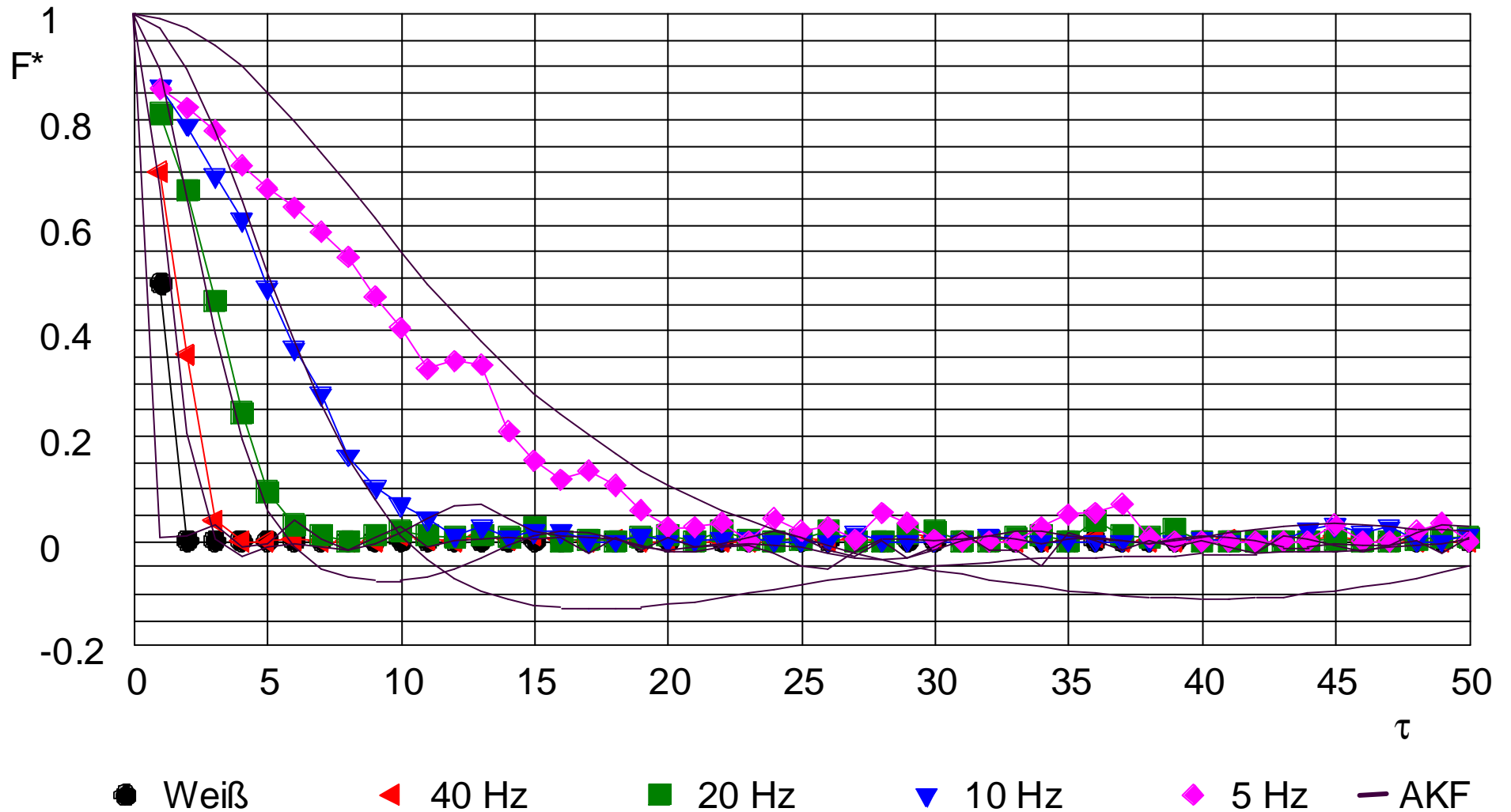
Beeinflussende Faktoren

Determinismus

Time-delay τ ($N=4096$, $m=6$, $b=6$):

Rauschen

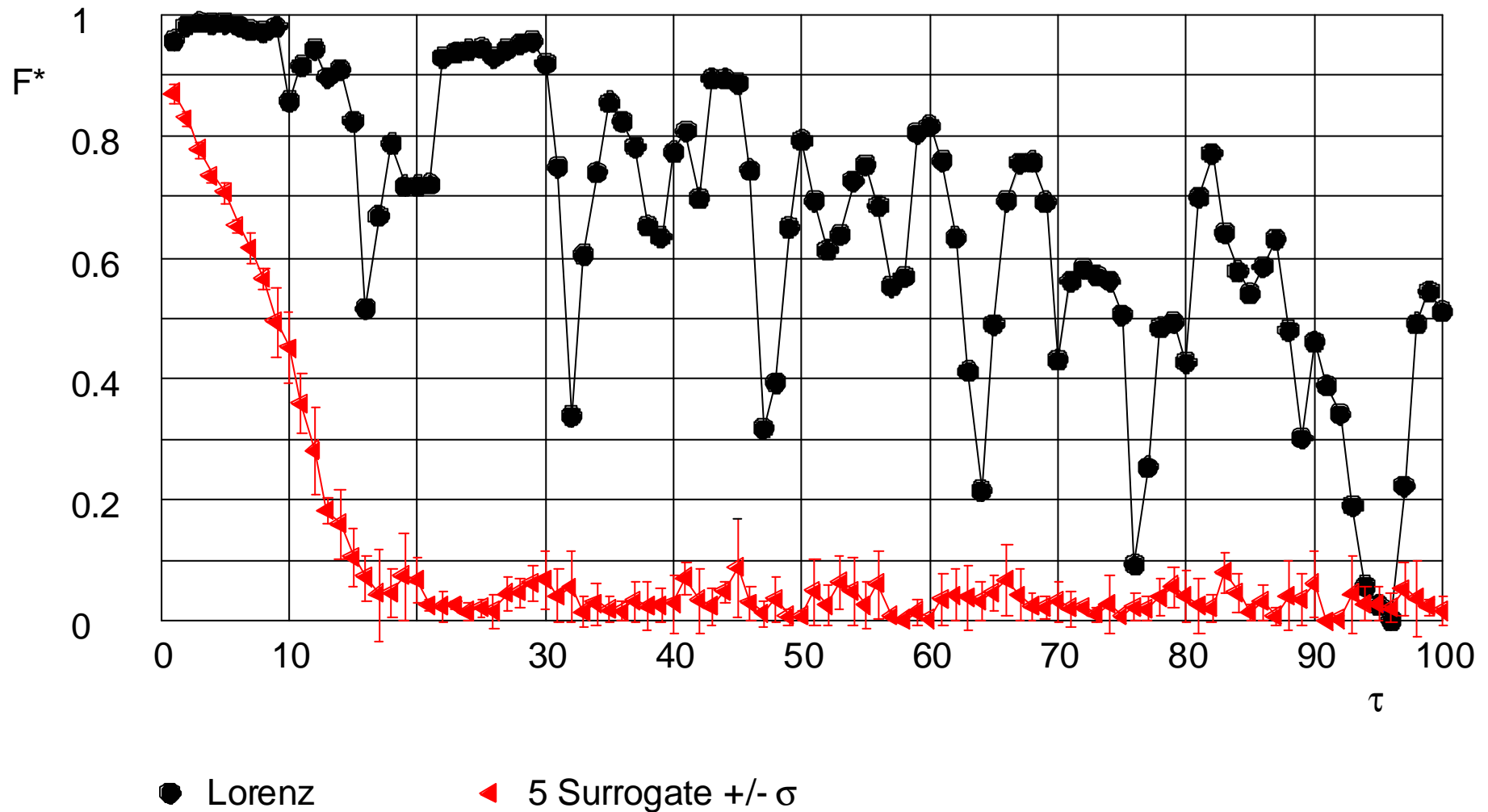
$\tau \neq \text{AKF}(0) \rightarrow$ Umlaufsinn \rightarrow lokaler Fluß \rightarrow Determinismus



Beeinflussende Faktoren

Determinismus

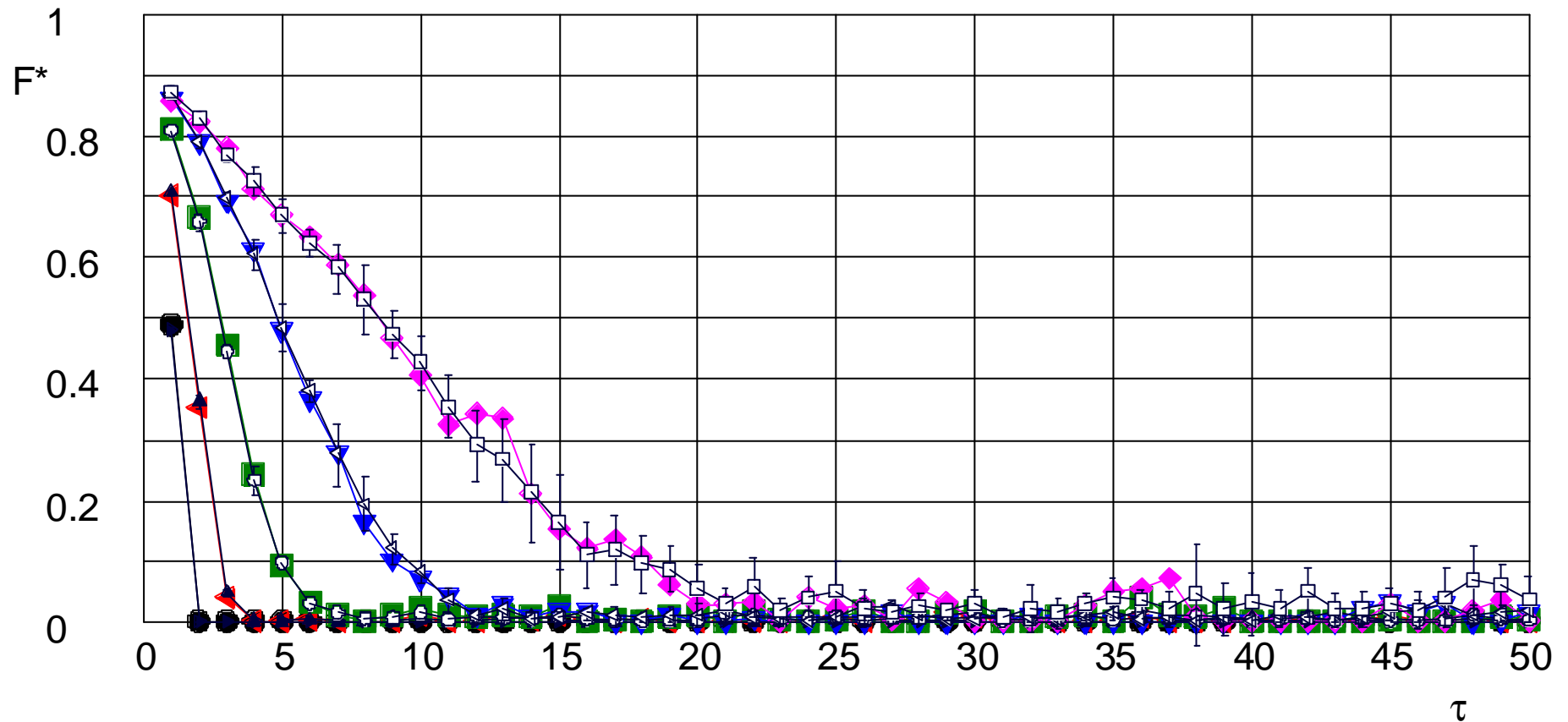
Time-delay τ ($N=4096$, $m=6$, $b=6$):
Modellsysteme + Surrogate



Beeinflussende Faktoren

Determinismus

Time-delay τ ($N=4096$, $m=6$, $b=6$):
Rauschen + Surrogate



- weiß
- ◀ 40 Hz
- 20 Hz
- ▼ 10 Hz
- ◆ 5 Hz
- ▶ ▲ ◻ ◂ ◅ *jeweils* 5 Surrogate $\pm \sigma$

einfache Tests

- Abweichungen von Gaußscher Amplitudenverteilung
- Verlauf der Autokorrelationsfunktion
- Zeitumkehrbarkeit

Phasenraum-basierte Tests

- Zeitumkehrbarkeit
- BDS-Statistik
- Surrogate

statistische Tests

- Datenpunktanzahl
- Stationarität
- Nullhypothese
- diskriminierende Statistik (abh. von Verteilung, ...)
- Signifikanzniveau

keine absoluten Aussagen möglich

Abweichungen von Gaußscher Amplitudenverteilung:

- (1) lineare Dynamik mit nicht-Gaußscher Amplitudenverteilung
(Skewness, Kurtosis)
- (2) lineare Dynamik, jedoch Beobachtung/Messung nichtlineare
statische Transformation → statische Nichtlinearität
- (3) nichtlineare Dynamik

Verlauf der Autokorrelationsfunktion:

- Ordnung der Nichtlinearität
- Varianz des Schätzers (Wiener-Khinchin-Theorem)

Zeitumkehrbarkeit (Zeitbereich):

Weiss, 1975:

Lineare Gaußsche stochastische Prozesse sind reversibel

aber:

- es gibt reversible nichtlineare Prozesse
- es gibt reversible nicht-Gaußsche lineare Prozesse

kein eindeutiger Beweis auf Nichtlinearität bei Verletzung der
Zeitumkehrbarkeit

Zeitumkehrbarkeit (Phasenraum):

Def.:

$\mathbf{v}(t)$ heißt reversibel, wenn Wahrscheinlichkeitsdichteverteilung $\rho(\mathbf{v})$ der Phasenraumvektoren unter der Zeitumkehrung für alle m und τ invariant ist.

- Datenpunktanzahl, Stationarität
- Phasenraumrekonstruktionsparameter
- Besetzungsdichte
- Phasenraumstatistik

BDS-Statistik:

- **wird aus Korrelationssumme abgeleitet**

für unkorrelierte Daten (*independent irregularly distributed*-Prozesse):

$$C_m(\varepsilon) = [C_1(\varepsilon)]^m$$

- **gleiche Einflußfaktoren**
- **zusätzlich: Einbettungsdimension m**

Surrogate:

- geeignete Nullhypothese
- random shuffling von kurzen Segmenten:
Wahl der Segmentlänge
- FT, AAFT, IAAFT:
end-point mismatch, edge-effect

Stationarität

Stationarität und Surrogate:

- singuläre Ereignisse
(intrinsisches vs. extrinsisches Phänomen)
- zyklstationäre Prozesse
Amplitudenmodulation
Frequenzmodulation
- ...

Ansatz: Nichtstationarität in Nullhypothese aufnehmen

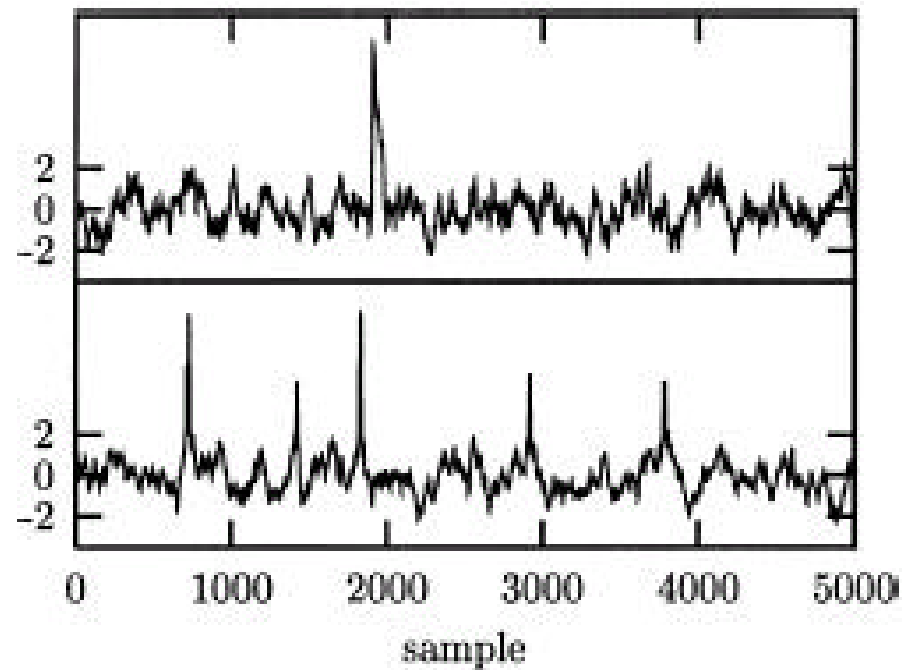


Fig. 23. A single spike is artificially introduced in an otherwise linear stochastic time sequence (upper). In the surrogate time series (lower), this leads to multiple short spikes. Although surrogate data have the same frequency content and take on the same set of values as the data, the remnants of the spike will lead to the detection of nonlinearity.

Zyklostationarität und Surrogate: (Timmer, Phys Rev E 58, 5153, 1998)

AR(2)-Prozeß

$$x(t) = a_1 x(t-1) + a_2 x(t-2) + \varepsilon(t) \quad \varepsilon(t) \sim N(0, \sigma^2)$$

$$a_1 = 2 \cos(2\pi / T) e^{(-1/\tau)}$$

$$a_2 = -e^{(-2/\tau)}$$

Amplitudenmodulation

$$x_{\text{amp}}(t) = \left(1 + \text{Mod}_{\text{amp}} \sin(2\pi / T_{\text{mod}} t)\right) x_0(t)$$

Frequenzmodulation

$$T(t) = T_{\text{mean}} + \text{Mod}_T \sin(2\pi p / T_{\text{mod}} t)$$

$$a_1(t) = 2 \cos(2\pi / T(t)) e^{(-1/\tau)}$$

Zyklostationarität und Surrogate: (Timmer, Phys Rev E 58, 5153, 1998)

- 8192 Datenpunkte
- 50 Surrogate
- vergleichende Statistik mit effektiver Korrelationsdimension:

$$z = \frac{\left| D_{\text{eff}}^2 - \langle D_{\text{eff}}^2 \rangle_{\text{surr}} \right|}{\sigma_{\text{surr}}}$$

- $z = 1.96$ entspricht 5 % Signifikanzniveau

(Timmer, Phys Rev E 58, 5153, 1998)

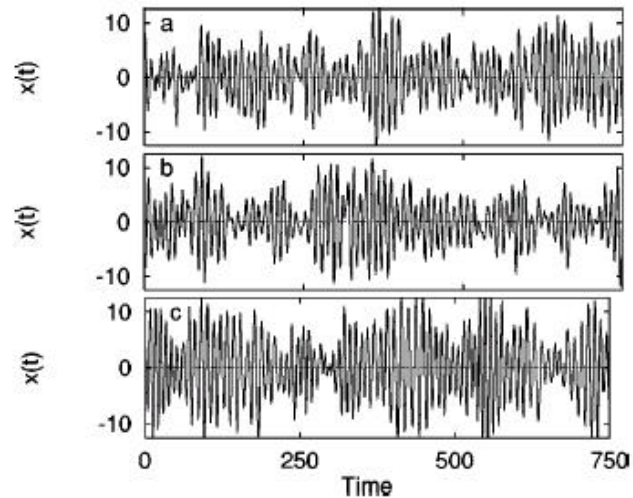


FIG. 1. Realizations of the processes investigated. (a) AR2 process satisfying the null hypothesis. (b) Amplitude modulated process with modulation depth of 0.3. (c) Period modulated process, relative amplitude of modulation is 15%.

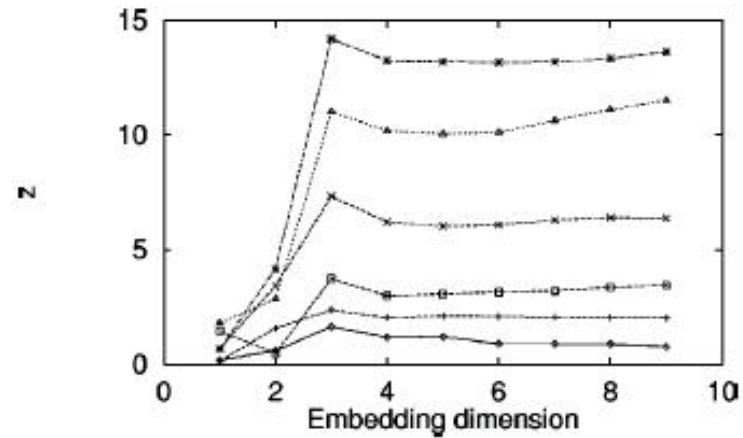


FIG. 3. Results of the simulation study for the amplitude modulated process. Shown is z in dependence on the embedding dimension for different degrees M_{amp} of violation ($\diamond = 0$, $+$ = 0.1, $\square = 0.2$, $\times = 0.3$, $\triangle = 0.4$, $*$ = 0.5).

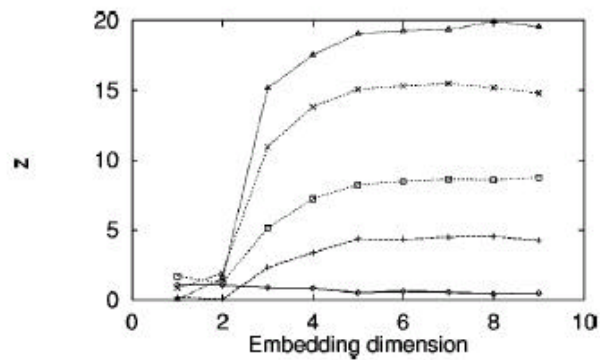


FIG. 4. Results of the simulation study for the period modulated process. Shown is z in dependence on the embedding dimension for different degrees M_T of violation ($\diamond = 0$, $+$ = 1, $\square = 1.5$, $\times = 2$, $\triangle = 3$).

Beeinflussende Faktoren

P 2 C 2 E

Process too complicated to explain

Iff *in*

S. Rushdie: Haroun and the Sea of Stories. Granta Books, London (1990)

Beeinflussende Faktoren

*absolute Aussagen bei realen Daten
nur bedingt möglich*



*möglicher Ausweg:
relative Aussagen*