# Lack of Robustness of the Surrogate Data Test:

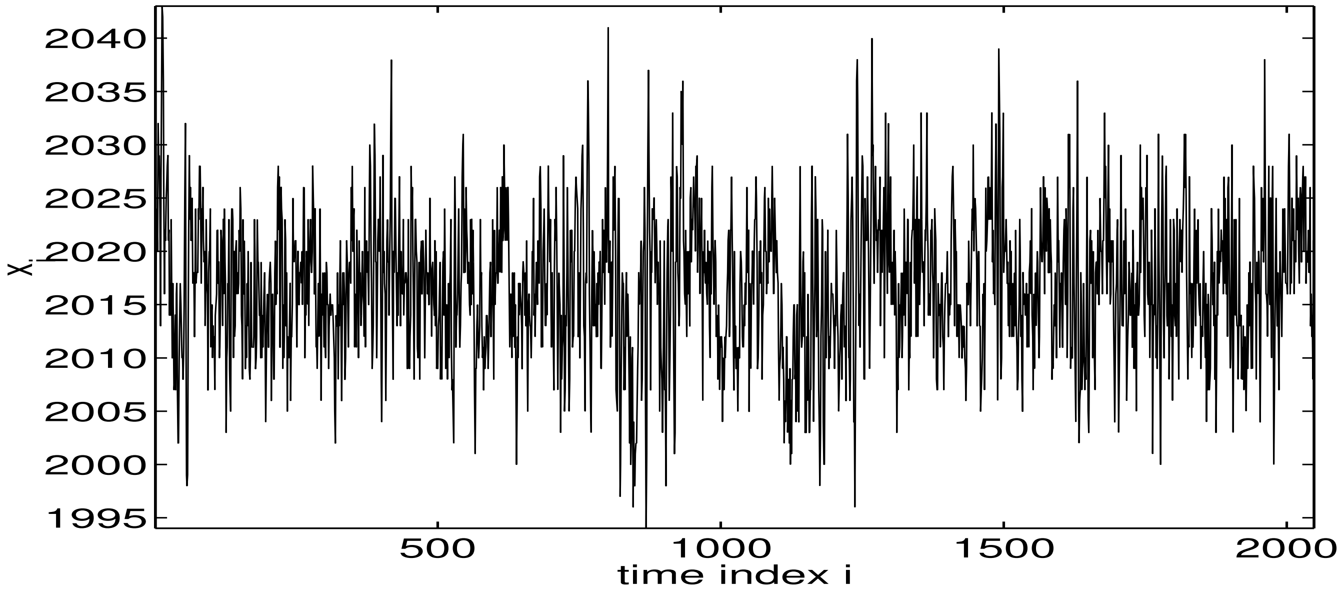
# Examples from Physiology and Finance

*Dimitris Kugiumtzis*

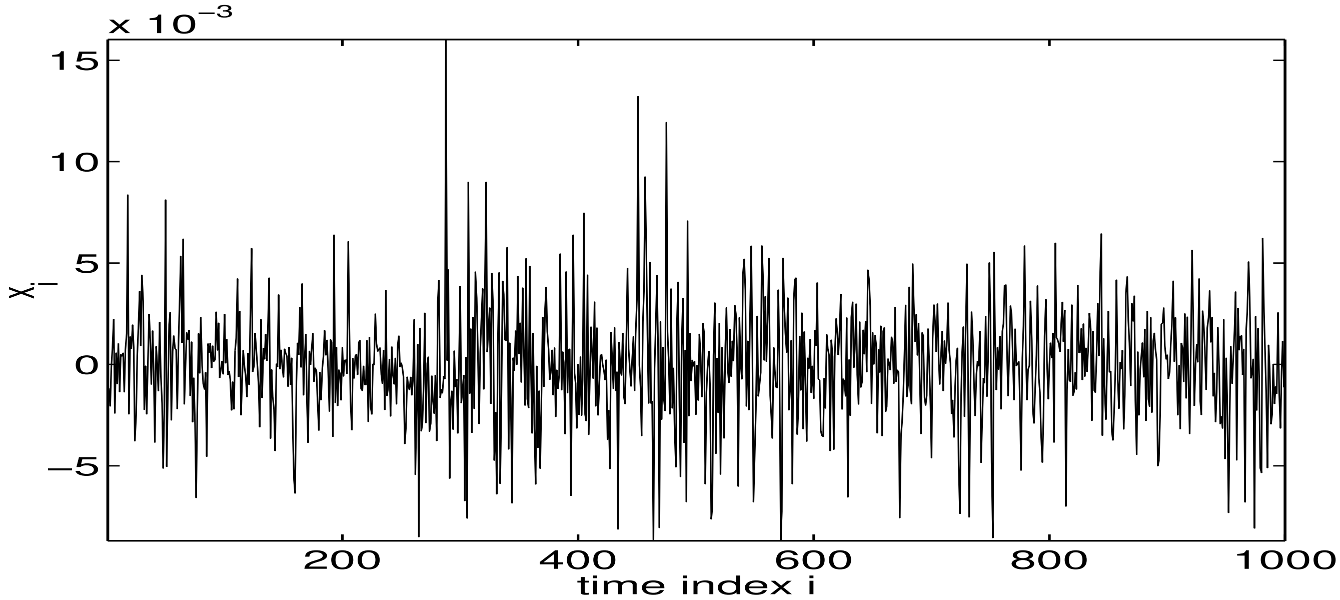
*Talk at the CSNSS'99 Workshop*
*Kazimierz Dolny, Poland, 14-17 October 1999*

## Real time series:
## linear statistical or nonlinear analysis?

EEG time series




First differences of
exchange rate time series (US$ / GBP)



# The surrogate data test

## Null hypothesis:

The original time series $\{x\}$ is a realization of a Gaussian process undergoing a static nonlinear transform: $x_i = h(s_i)$, $\mathbf{s} \sim N(\mu, \Sigma)$

*Conditions for the surrogate data* $\{z\}$*:*

1. Same autocorrelation $R_z(\tau) = R_x(\tau)$

2. Same amplitude distribution $A_z(z) = A_x(x)$

## Generate surrogate data:

*Method 1:* Amplitude adjusted Fourier transform (AAFT) *(Theiler, 1992)*

*Method 2:* Iterated (or Improved) amplitude adjusted Fourier transform (IAAFT) *(Schreiber and Schmitz, 1996)*

# Test of the hypothesis

1. Choose a discriminating statistic $q$
   (e.g. CC from a global polynomial fit)

2. Compute $q_0$ for original and $q_1, \ldots, q_M$ for the $M$ surrogates

3. Significance $S$ in sigmas

$$S = \frac{|q_0 - \bar{q}|}{\mathrm{SD}_q}$$

$S > 1.96 \Rightarrow$ rejection with $> 95\%$ confidence

## The AAFT algorithm

*Critical assumption:* $h$ is monotonic

**Step 1:** $\{x\} \to \{y\}$, white Gaussian noise is reordered to match the rank of $\{x\}$ ("Gaussianization") (simulation of $h^{-1}$)

**Step 2:** $\{y\} \to \{y^{\mathsf{FT}}\}$, Fourier transform, phase randomization and inverse Fourier transform

**Step 3:** $\{y^{\mathsf{FT}}\} \to \{z\}$, $\{x\}$ is reordered to match the rank of $\{y^{\mathsf{FT}}\}$ (simulation of $h$).

AAFT gives for $\{z\}$: $A_z = A_x$ and $R_z \approx R_x$

> If $h$ is not monotonic the difference of $R_z$ and $R_x$ can be large

# The IAAFT algorithm

Starting with a random permutation $\{z^{(0)}\}$ of the data $\{x\}$

**Step 1** $\{z^{(i)}\} \to \{y^{(i)}\}$, The power spectrum of $\{z^{(i)}\}$ is brought to be identical to that of $\{x\}$.

**Step 2** $\{y^{(i)}\} \to \{z^{(i+1)}\}$, $\{x\}$ is reordered to have the same rank structure as $\{y^{(i)}\}$.

- Many iterations are made (long computation time)

- The algorithm terminates when reordering in two consecutive steps is the same

IAAFT-1 (the final, from step 2) gives

$$A_z = A_x \text{ and } R_z \approx R_x$$

IAAFT-2 (from step 1 at last iteration) gives
$A_z \approx A_x$ and $R_z \approx R_x$ (closer to identical in $R$)

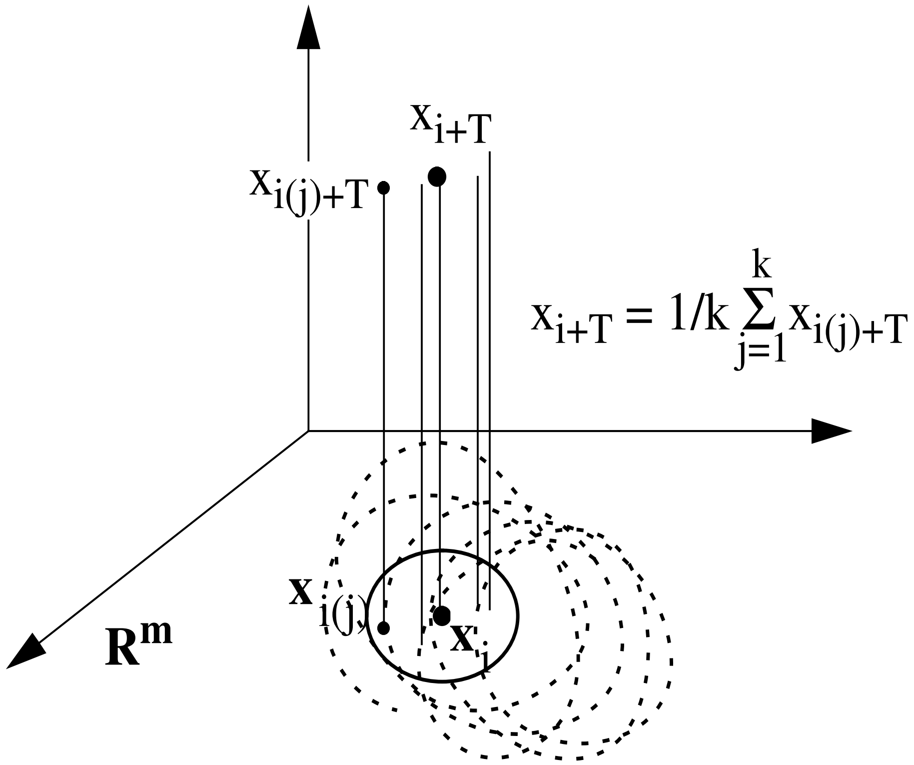
$\boxed{\text{Better matching in } R \text{ than AAFT}}$

# The discriminating statistics

1. **Mutual information (MUT)** measures the general correlation (linear and nonlinear) $I(\tau)$ in the data as a function of the lag $\tau$ *(Fraser, 1986)*.

*discriminating statistics:* $q_\tau = I(\tau)$

2. **Local average mapping (LAM)** estimates the $T$-mapping $x_{i+T}$ of each reconstructed point $\mathbf{x}_i$ from the average of the $T$-mappings of its $k$ neighbor points (here $T = 1$).




*discriminating statistics:* $q_m = \mathrm{CC}(m)$

# The discriminating statistics

3. The **largest Lyapunov exponent (LLE)** measures the rate of divergence of nearby trajectories and thus the complexity of the system.

*discriminating statistics:* $q_m = \lambda_1(m)$

4. *Simple nonlinear statistics* as

**three point autocorrelation (TPA)**: *(Rao, 1984)*

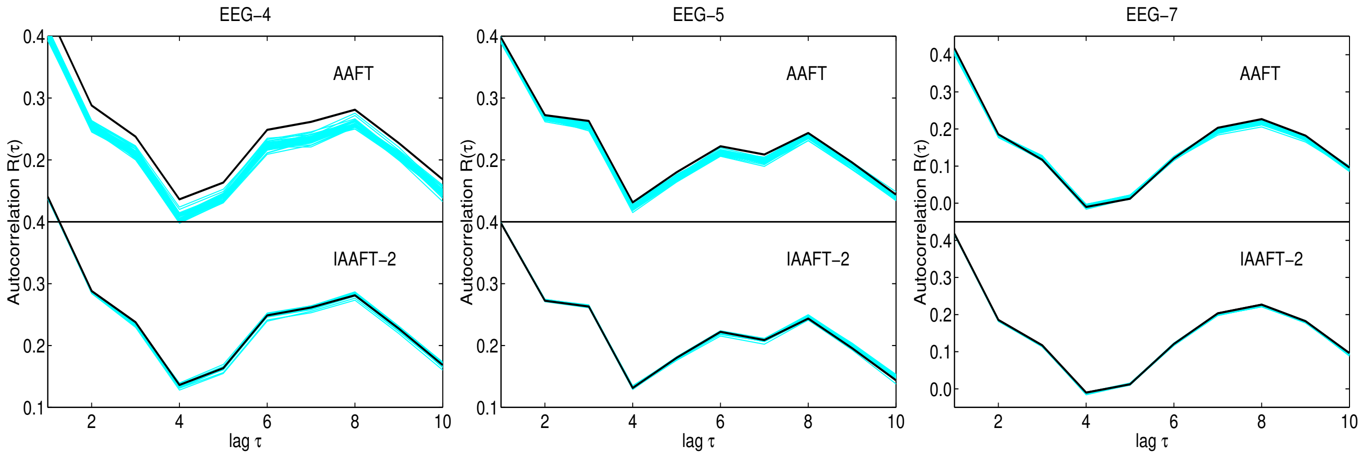
$$q_\tau = \frac{\langle (x_i - \bar{x})(x_{i-\tau} - \bar{x})(x_{i-2\tau} - \bar{x}) \rangle}{\langle (x_i - \bar{x})^3 \rangle}$$

and **time reversal asymmetry (REV)** *(Diks, 1995)*

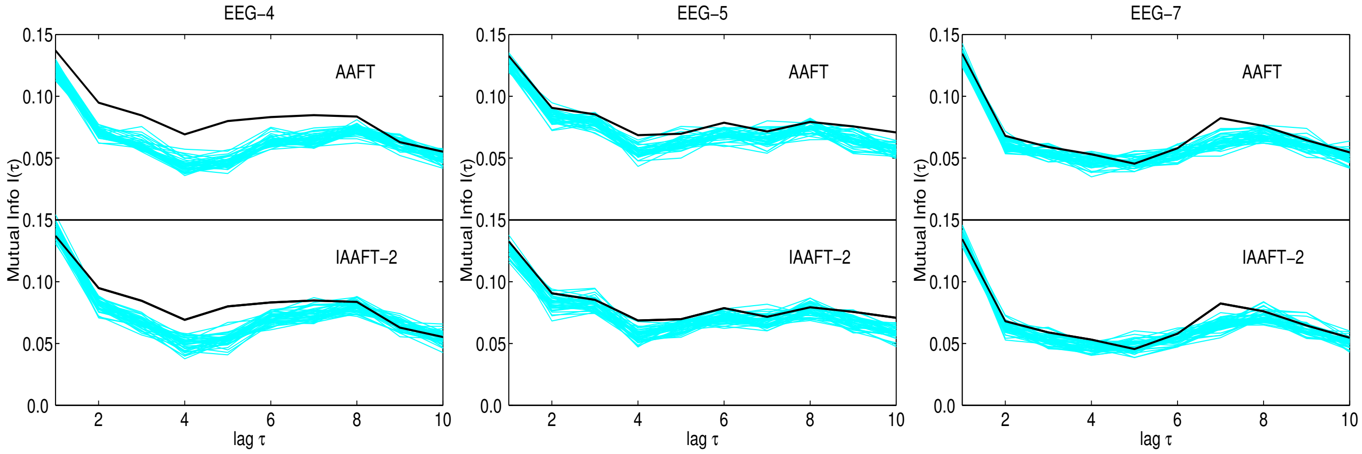
$$q_\tau = \frac{\langle (x_i - x_{i-\tau})^3 \rangle}{\langle (x_i - x_{i-\tau})^2 \rangle}$$

# EEG time series (4,5,7)

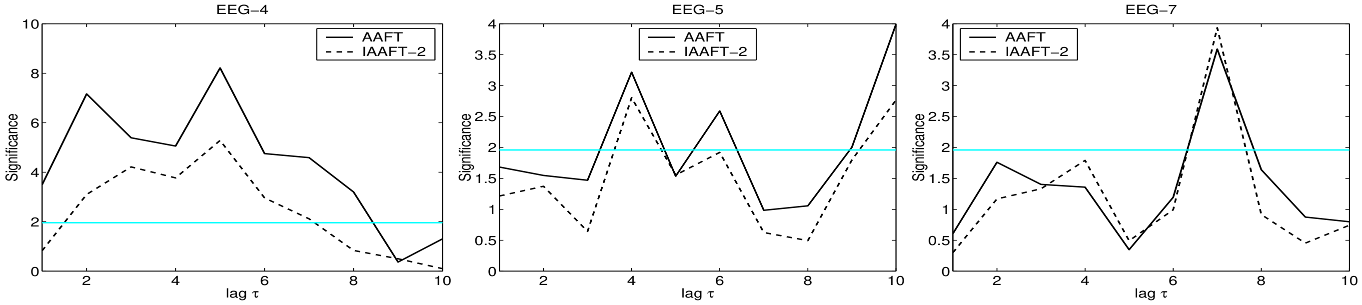
## Autocorrelation (AUT)




## Mutual information (MUT)




## Significance for MUT

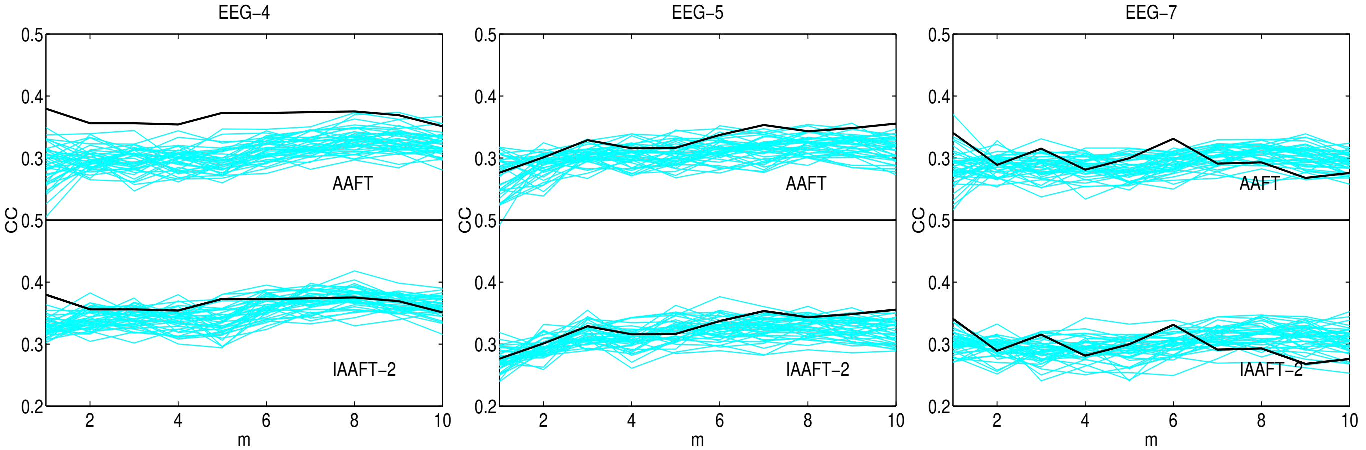



Bias in linear correlations

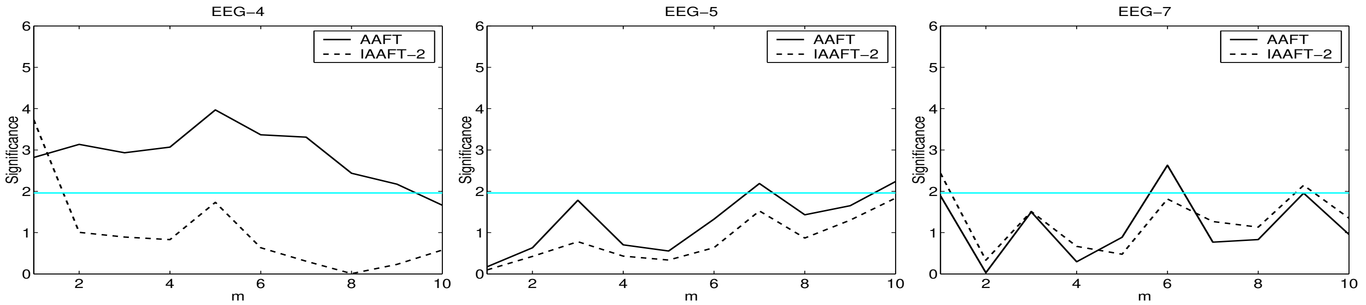
$\Downarrow$

bias for rejection

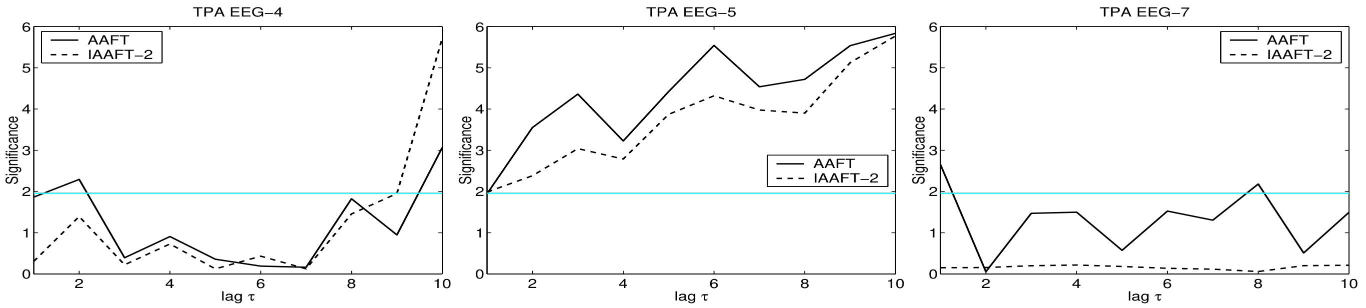
# Local Average Map (LAM)



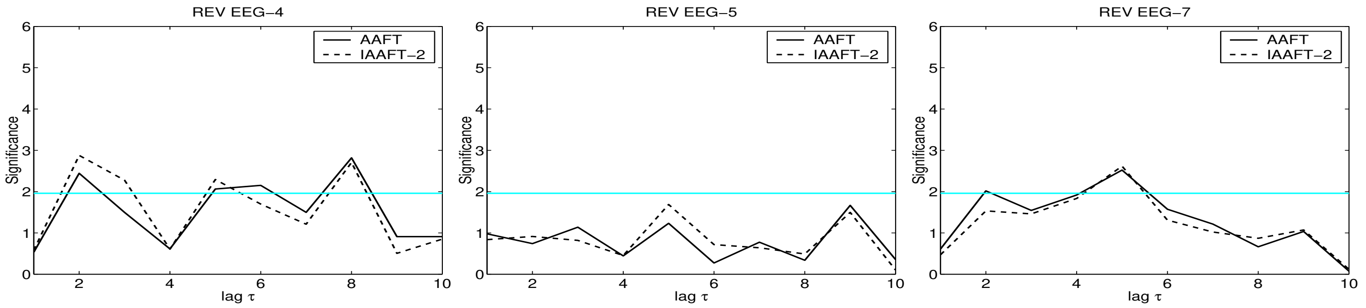

# Significance for LAM




# Significance for three point autocorrelation (TPA)




# Significance for time reversal asymmetry (REV)

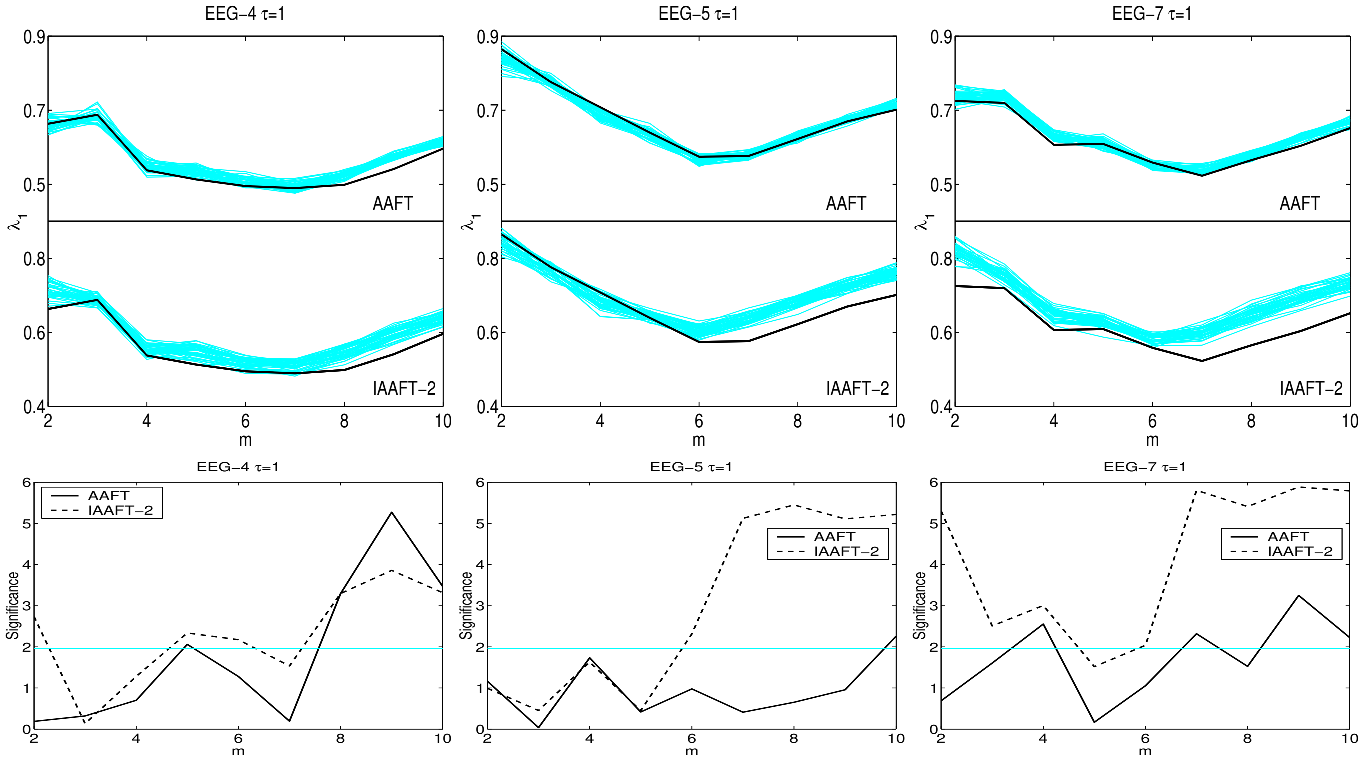



different methods

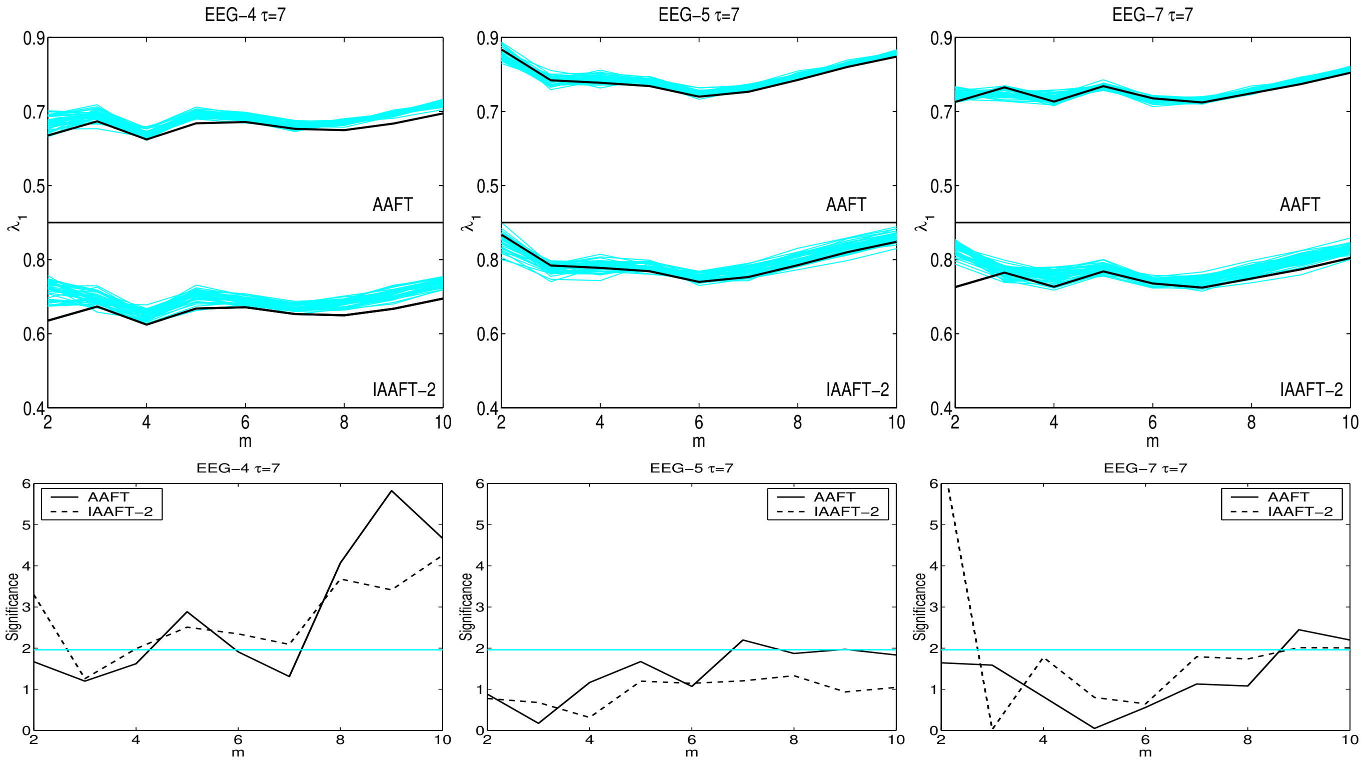
$\Downarrow$

different significance levels

# Largest Lyapunov Exponent (LLE) $\tau = 1$




# Largest Lyapunov Exponent (LLE) $\tau = 7$



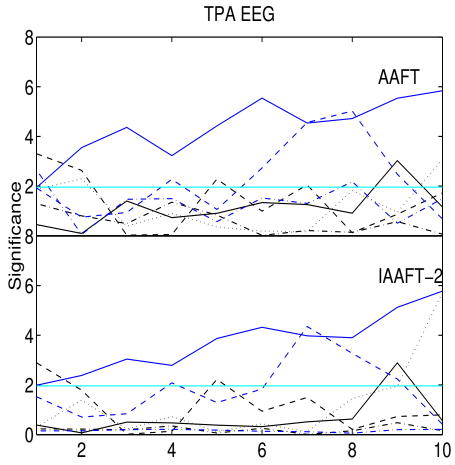

different method parameters

$\Downarrow$

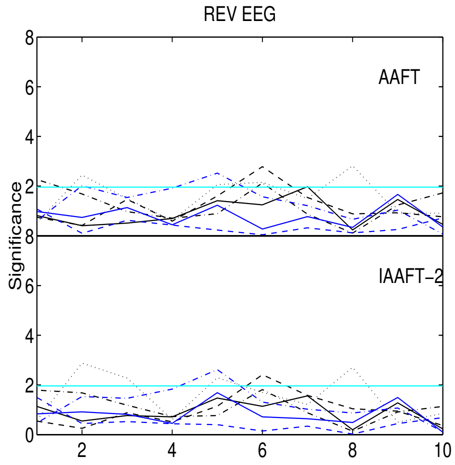
different significance levels

# All EEG time series

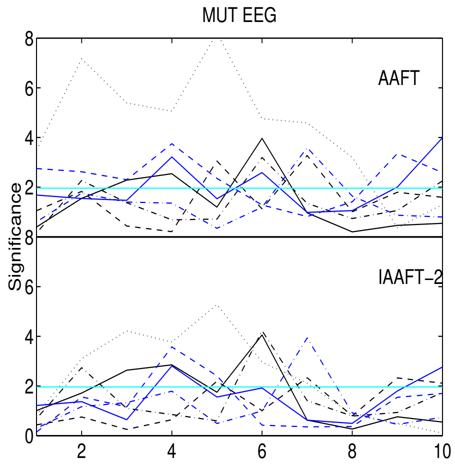
## TPA



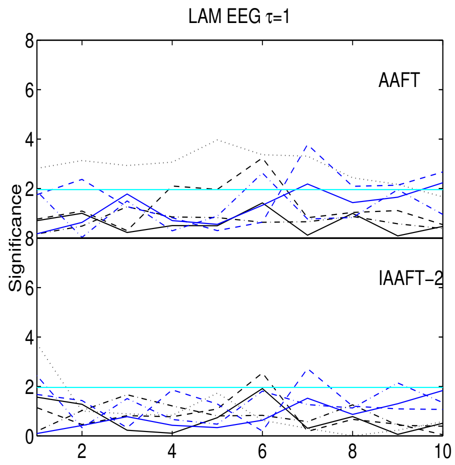

## REV



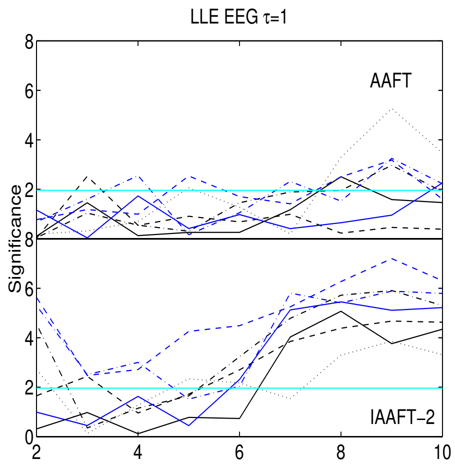

## MUT



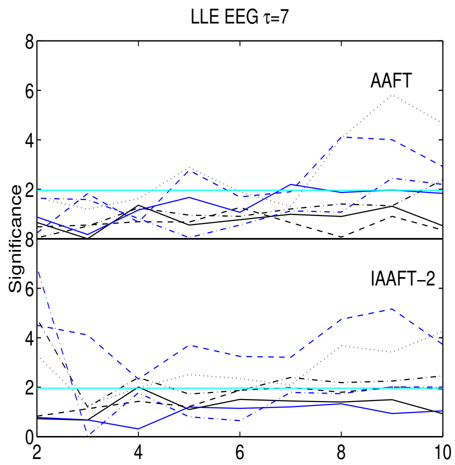

## LAM




## LLE $\tau = 1$




## LLE $\tau = 7$



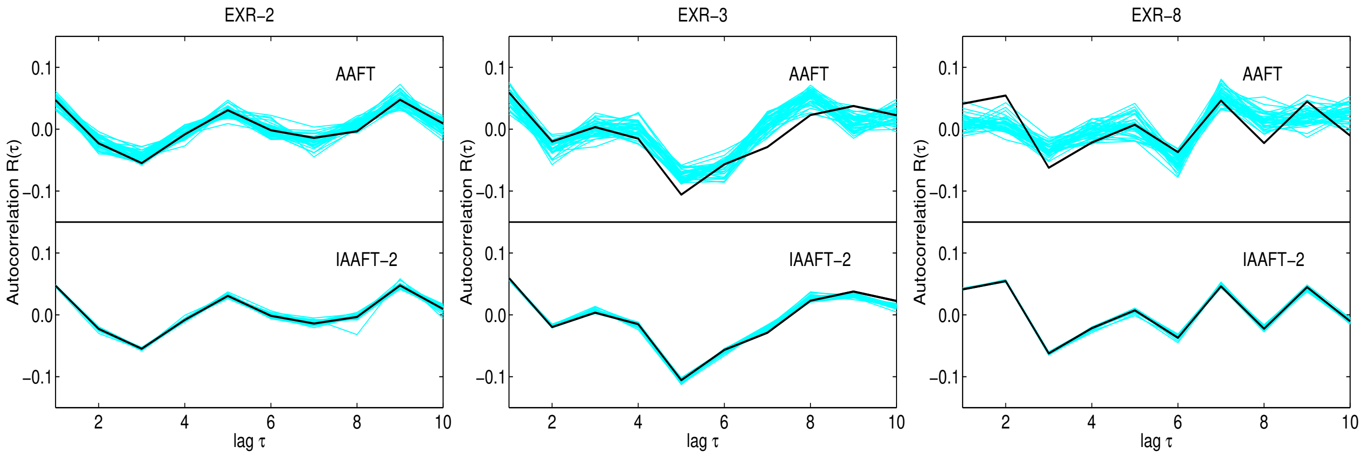

different EEG data sets

$\Downarrow$

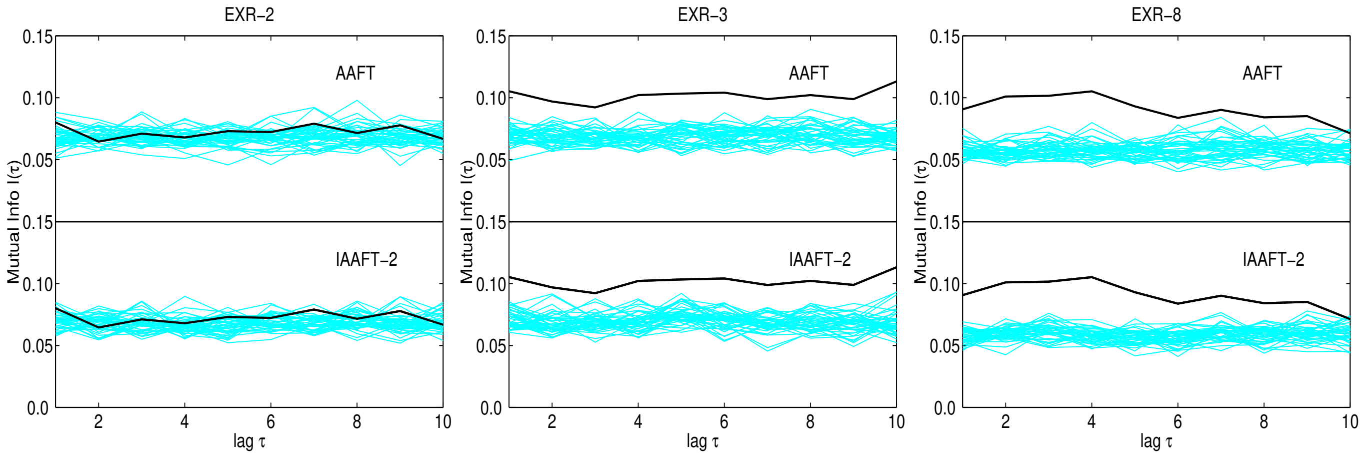
different significance levels

# Exchange rate (EXR) time series (2,3,8)

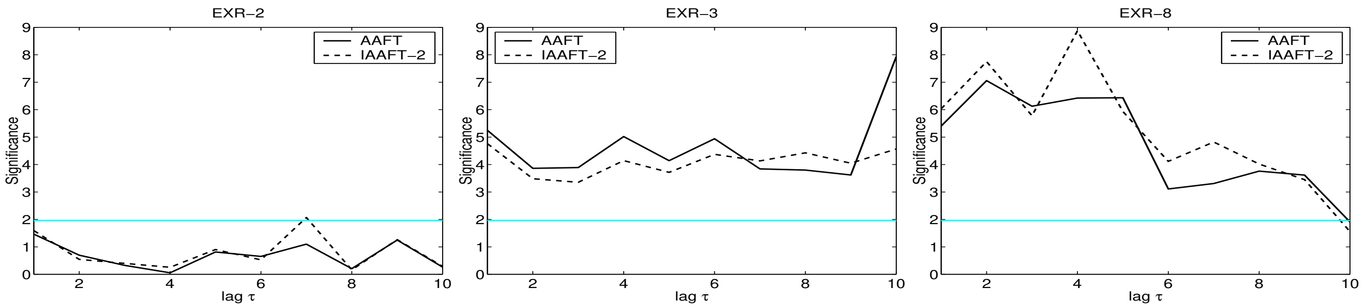
## Autocorrelation (AUT)




## Mutual information (MUT)

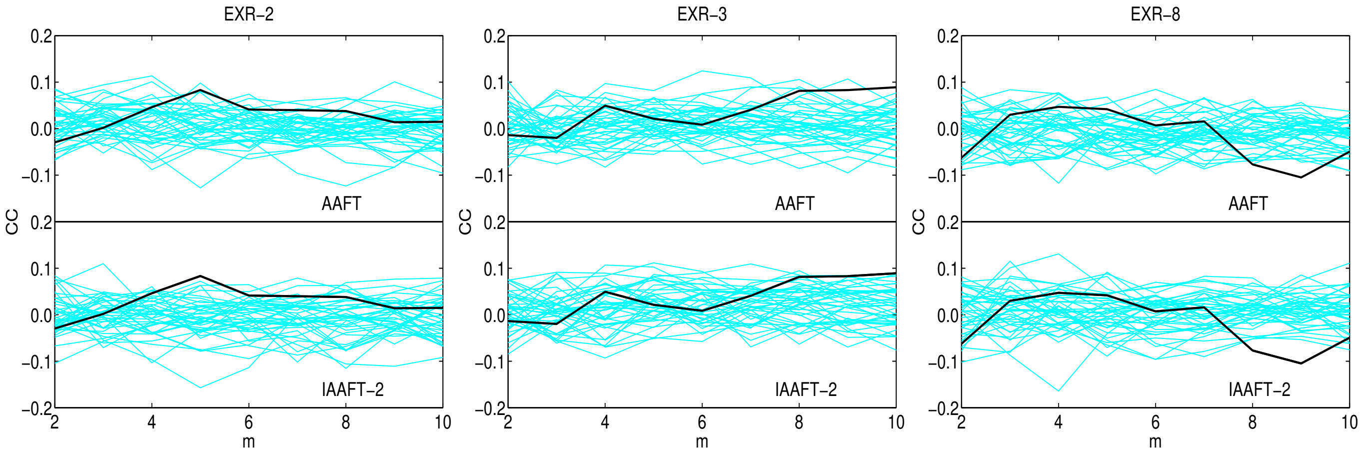



## Significance for MUT

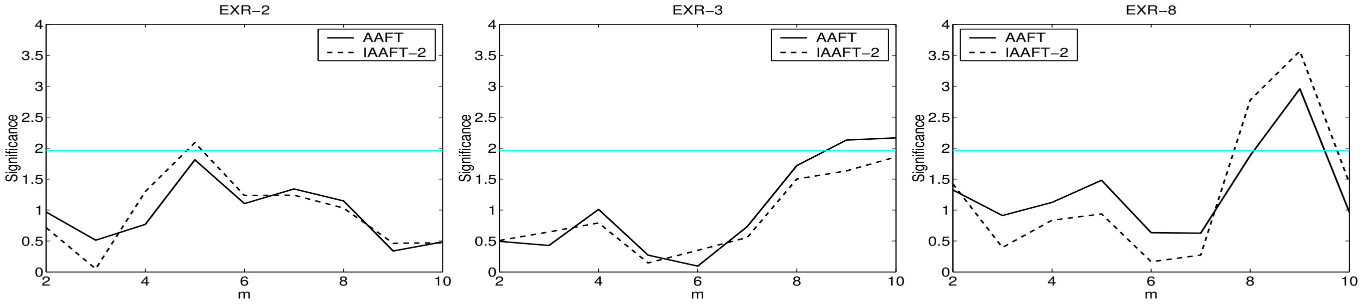



No bias in linear correlations
and
rejection with MUT (nonlinear correlations)

## Local Average Map (LAM)




## Significance for LAM



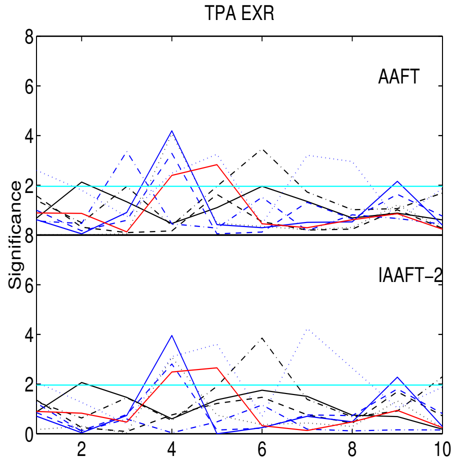

**No predictability with LAM**

$\Downarrow$

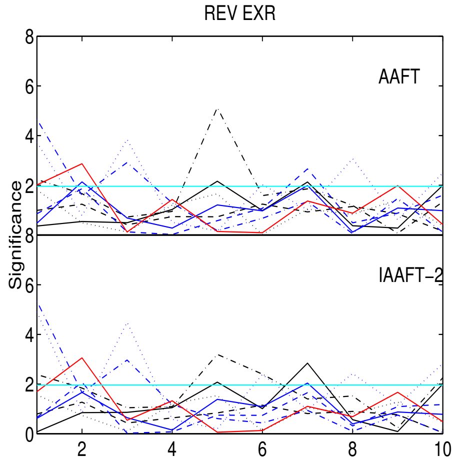
**No nonlinearity**

# All EXR time series

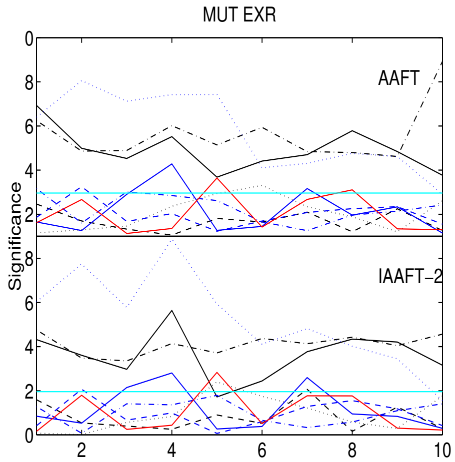
## TPA



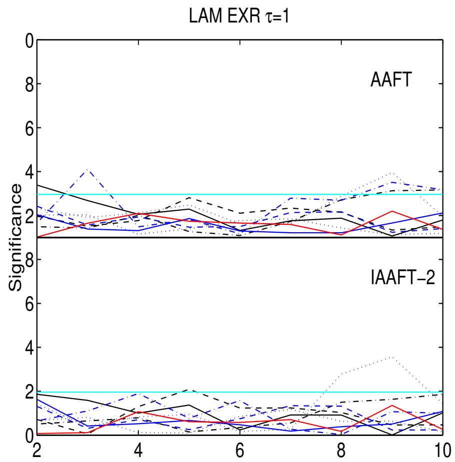

## REV




## MUT




## LAM




different EXR data sets

$\Downarrow$

different significance levels

## Conclusions

- The performance of the surrogate data test is not robust, but depends on

  1. The algorithm for generating the surrogate data (AAFT or IAAFT)

  2. The nonlinear method that provides the diagnostic statistic

  3. The parameters of the method

- Different real data sets of the same type (e.g. EEG or exchange rate) may give different results under the same surrogate data test.
  Is it because of different nonlinear characteristics or because the surrogate data test is not robust?