Evaluation of Causality Measures Based on Non-Uniform Embedding Schemes with Application to the Cardiovascular System

Petroula Laiou¹, Ralph G. Andrzejak¹ and Dimitris Kugiumtzis²

¹ Department of Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain
and
² Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Greece
e-mail: dkugiu@auth.gr http://users.auth.gr/dkugiu

ESGCO 2014
Trento, 25-28 May 2014
Outline

1. Granger causality measures
2. The Mutual Information from Mixed Embedding (MIME)
3. The Conditional Entropy from Non-uniform Embedding (Faes-Nollo-Porta, FNP)
4. Simulations and Evaluation of MIME and FNP
5. MIME and FNP applied to the Cardiovascular system
Granger causality measures
1. Granger causality measures

2. The Mutual Information from Mixed Embedding (MIME)

3. The Conditional Entropy from Non-uniform Embedding (Faes-Nollo-Porta, FNP)

4. Simulations and Evaluation of MIME and FNP

5. MIME and FNP applied to the Cardiovascular system
Outline

1. Granger causality measures
2. The Mutual Information from Mixed Embedding (MIME)
3. The Conditional Entropy from Non-uniform Embedding (Faes-Nollo-Porta, FNP)
1. Granger causality measures
2. The Mutual Information from Mixed Embedding (MIME)
3. The Conditional Entropy from Non-uniform Embedding (Faes-Nollo-Porta, FNP)
4. Simulations and Evaluation of MIME and FNP
Outline

1. Granger causality measures
2. The Mutual Information from Mixed Embedding (MIME)
3. The Conditional Entropy from Non-uniform Embedding (Faes-Nollo-Porta, FNP)
4. Simulations and Evaluation of MIME and FNP
5. MIME and FNP applied to the Cardiovascular system
Cardio-Respiratory-Vascular system

“apnea?”

“normal?”

Coupling between heart rate, respiration and blood oxygen concentration?

Direction of coupling (driving)?

Granger causality

Laiou, Andrzejak and Kugiumtzis

Non-Uniform Embedding Causality Measures
Cardio-Respiratory-Vascular system

“apnea?”

“normal?”

Coupling between heart rate, respiration and blood oxygen concentration?
Cardio-Respiratory-Vascular system

“apnea?”

“normal?”

Coupling between heart rate, respiration and blood oxygen concentration?
Direction of coupling (driving)?
Cardio-Respiratory-Vascular system

“apnea?”

“normal?”

Coupling between heart rate, respiration and blood oxygen concentration?

Direction of coupling (driving) ? ⇒ Granger causality

Laiou, Andrzejak and Kugiumtzis

Non-Uniform Embedding Causality Measures
1. Granger causality measures

2. The Mutual Information from Mixed Embedding (MIME)

3. The Conditional Entropy from Non-uniform Embedding (Faes-Nollo-Porta, FNP)

4. Simulations and Evaluation of MIME and FNP

5. MIME and FNP applied to the Cardiovascular system
## Granger Causality measures

<table>
<thead>
<tr>
<th>Linear</th>
<th>X → Y</th>
<th>X → Y</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Granger Causality Index (GCI)</td>
<td></td>
<td></td>
<td>Conditional (Partial) GCI (CGCI)</td>
</tr>
<tr>
<td>Directed Coherence (DC)</td>
<td></td>
<td></td>
<td>Partial DC (PDC)</td>
</tr>
<tr>
<td>Directed Transfer Function (DTF)</td>
<td></td>
<td></td>
<td>direct DTF (dDTF)</td>
</tr>
<tr>
<td>Nonlinear</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Directionality Index</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inter-dependence measures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Conditional Recurrence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transfer entropy (TE)</td>
<td></td>
<td></td>
<td>Partial TE (PTE)</td>
</tr>
<tr>
<td>Mutual Information from Mixed Embedding (MIME)</td>
<td></td>
<td></td>
<td>Partial MIME (PMIME)</td>
</tr>
<tr>
<td>Conditional Entropy from Non-uniform Embedding (FNP)</td>
<td></td>
<td></td>
<td>Conditional Entropy from Non-uniform Embedding (FNP)</td>
</tr>
</tbody>
</table>
Embedding parameters and nonlinear causality measures

time series \( \{x_t, y_t\}_{t=1}^n \) driving system: \( X \), response system: \( Y \),
Embedding parameters and nonlinear causality measures

time series \( \{x_t, y_t\}_{t=1}^n \)  

driving system: \( X \), response system: \( Y \),

**State space reconstruction**

\[
x_t = [x_t, x_{t-\tau}, \ldots, x_{t-(m-1)\tau}]', \quad y_t = [y_t, y_{t-\tau}, \ldots, y_{t-(m-1)\tau}]',
\]

\( \tau, m \): embedding parameters (generally different for \( X \) and \( Y \))
Embedding parameters and nonlinear causality measures

time series \(\{x_t, y_t\}_{t=1}^n\) driving system: \(X\), response system: \(Y\),

State space reconstruction
\(x_t = [x_t, x_{t-\tau}, \ldots, x_{t-(m-1)\tau}]', y_t = [y_t, y_{t-\tau}, \ldots, y_{t-(m-1)\tau}]',\)

\(\tau, m\): embedding parameters (generally different for \(X\) and \(Y\))

\(y_{t+1}\): future state of \(Y\)
time series \( \{x_t, y_t\}_{t=1}^n \) driving system: \( X \), response system: \( Y \),

State space reconstruction

\[
x_t = [x_t, x_{t-\tau}, \ldots, x_{t-(m-1)\tau}]', \quad y_t = [y_t, y_{t-\tau}, \ldots, y_{t-(m-1)\tau}]'
\]

\( \tau, m \): embedding parameters (generally different for \( X \) and \( Y \))

\( y_{t+1} \): future state of \( Y \)

**Entropy**: information from each sample of \( X \) (assume proper discretization of \( X \)):

\[
H(X) = \sum_x p_X(x) \log p_X(x)
\]

**Mutual Information**: Information on \( X \) from \( Y \) and vice versa:

\[
I(X; Y) = H(X) + H(Y) - H(X, Y)
\]
Embedding parameters and nonlinear causality measures

time series \( \{x_t, y_t\}_{t=1}^n \) driving system: \( X \), response system: \( Y \),

State space reconstruction
\[
\mathbf{x}_t = [x_t, x_{t-\tau}, \ldots, x_{t-(m-1)\tau}]', \quad \mathbf{y}_t = [y_t, y_{t-\tau}, \ldots, y_{t-(m-1)\tau}]',
\]
\( \tau, m \): embedding parameters (generally different for \( X \) and \( Y \))
\( y_{t+1} \): future state of \( Y \)

Entropy: information from each sample of \( X \) (assume proper discretization of \( X \)):
\[
H(X) = \sum_x p_X(x) \log p_X(x)
\]

Mutual Information: Information on \( X \) from \( Y \) and vice versa:
\[
I(X; Y) = H(X) + H(Y) - H(X, Y)
\]

Transfer Entropy (TE) [Schreiber, 2000]

Measure the effect of \( X \) on \( Y \) at one-step ahead, accounting (conditioning) for the effect from its own current state.

\[
\text{TE}_{X \rightarrow Y} = I(y_{t+1}; \mathbf{x}_t | y_t) = H(x_t, y_t) - H(y_{t+1}, x_t, y_t) + H(y_{t+1}, y_t) - H(y_t)
\]
Embedding parameters and nonlinear causality measures

time series \( \{x_t, y_t\}_{t=1}^n \) driving system: \( X \), response system: \( Y \),

State space reconstruction
\[
x_t = [x_t, x_{t-\tau}, \ldots, x_{t-(m-1)\tau}]', \quad y_t = [y_t, y_{t-\tau}, \ldots, y_{t-(m-1)\tau}]',
\]
\( \tau, m \): embedding parameters (generally different for \( X \) and \( Y \))
\( y_{t+1} \): future state of \( Y \)

Entropy: information from each sample of \( X \) (assume proper discretization of \( X \)):
\[
H(X) = \sum_x p_X(x) \log p_X(x)
\]

Mutual Information: Information on \( X \) from \( Y \) and vice versa:
\[
I(X; Y) = H(X) + H(Y) - H(X, Y)
\]

Transfer Entropy (TE) [Schreiber, 2000]
Measure the effect of \( X \) on \( Y \) at one-step ahead, accounting (conditioning) for the effect from its own current state.
\[
TE_{X \rightarrow Y} = I(y_{t+1}; x_t|y_t) = H(x_t, y_t) - H(y_{t+1}, x_t, y_t) + H(y_{t+1}, y_t) - H(y_t)
\]

Joint entropies (and distributions) can have high dimension!

Significance test using randomization
1. Granger causality measures
2. The Mutual Information from Mixed Embedding (MIME)
3. The Conditional Entropy from Non-uniform Embedding (Faes-Nollo-Porta, FNP)
4. Simulations and Evaluation of MIME and FNP
5. MIME and FNP applied to the Cardiovascular system
The idea: [Vlachos & Kugiumtzis, PRE, 2010]
The idea: [Vlachos & Kugiumtzis, PRE, 2010]

1. Find a mixed embedding of varying delays from $X$ and $Y$ that explains best the future of $Y$. 
The idea: [Vlachos & Kugiumtzis, PRE, 2010]

1. Find a mixed embedding of varying delays from $X$ and $Y$ that explains best the future of $Y$.
2. Quantify the information on $Y$ ahead that is explained by the $X$-components of the mixed embedding vector.
The mixed embedding scheme

- Start with an empty embedding vector $w^0_t$, future vector (sample) of $Y$, $y_{t+1}$, and maximum lag $L_x$ for $X$ and $L_y$ for $Y$. 

$$W_t = \{x_t, \ldots, x_{t-L_x-1}, y_t, \ldots, y_{t-L_y-1}\}$$
The mixed embedding scheme

- Start with an empty embedding vector $w_t^0$, future vector (sample) of $Y$, $y_{t+1}$, and maximum lag $L_x$ for $X$ and $L_y$ for $Y$.
  $W_t = \{x_t, \ldots, x_{t-L_x-1}, y_t, \ldots, y_{t-L_y-1}\}$
- First embedding cycle: $w_t^1 = \text{argmax}_{w \in W_t} I(y_{t+1}; w)$, and $w_t^1 = (w_t^1)$
The mixed embedding scheme

- Start with an empty embedding vector $w_t^0$, future vector (sample) of $Y$, $y_{t+1}$, and maximum lag $L_x$ for $X$ and $L_y$ for $Y$.
  
  \[ W_t = \{x_t, \ldots, x_{t-L_x-1}, y_t, \ldots, y_{t-L_y-1}\} \]

- First embedding cycle: $w_t^1 = \arg\max_{w \in W_t} I(y_{t+1}; w)$, and $w_t^1 = (w_t^1)$

- At embedding cycle $j$ suppose $w_t^{j-1} = (w_t^1, w_t^2, \ldots, w_t^{j-1})$. 

\[ \text{Estimation of } I(y_{t+1}; w|w_t^{j-1}) \text{ using nearest neighbors} \]
The mixed embedding scheme

- Start with an empty embedding vector \( w_t^0 \), future vector (sample) of \( Y, y_{t+1} \), and maximum lag \( L_x \) for \( X \) and \( L_y \) for \( Y \).

\[
W_t = \{x_t, \ldots, x_{t-L_x-1}, y_t, \ldots, y_{t-L_y-1}\}
\]

- First embedding cycle: \( w_t^1 = \arg\max_{w \in W_t} I(y_{t+1}; w) \), and \( w_t^1 = (w_t^1) \)

- At embedding cycle \( j \) suppose \( w_t^{j-1} = (w_t^1, w_t^2, \ldots, w_t^{j-1}) \). Add the component \( w_t^j \in W_t \setminus w_t^{j-1} \) that maximizes mutual information to \( y_{t+1} \) conditioning on the current components in \( w_t^{j-1} \).
The mixed embedding scheme

- Start with an empty embedding vector \( w_t^0 \), future vector (sample) of \( Y, y_{t+1} \), and maximum lag \( L_x \) for \( X \) and \( L_y \) for \( Y \).
- \( W_t = \{x_t, \ldots, x_{t-L_x-1}, y_t, \ldots, y_{t-L_y-1}\} \)
- First embedding cycle: \( w_t^1 = \arg\max_{w \in W_t} I(y_{t+1}; w) \), and \( w_t^1 = (w_t^1) \)
- At embedding cycle \( j \) suppose \( w_t^{j-1} = (w_t^1, w_t^2, \ldots, w_t^{j-1}) \). Add the component \( w_t^j \in W_t \setminus w_t^{j-1} \) that maximizes mutual information to \( y_{t+1} \) conditioning on the current components in \( w_t^{j-1} \),

\[
w_t^j = \arg\max_{w \in W_t \setminus w_t^{j-1}} I(y_{t+1}; w | w_t^{j-1})
\]
The mixed embedding scheme

- Start with an empty embedding vector $w_t^0$, future vector (sample) of $Y$, $y_{t+1}$, and maximum lag $L_x$ for $X$ and $L_y$ for $Y$. 
  $W_t = \{x_t, \ldots, x_{t-L_x-1}, y_t, \ldots, y_{t-L_y-1}\}$

- First embedding cycle: $w_t^1 = \text{argmax}_{w \in W_t} I(y_{t+1}; w)$, and $w_t^1 = (w_t^1)$

- At embedding cycle $j$ suppose $w_t^{j-1} = (w_t^1, w_t^2, \ldots, w_t^{j-1})$. Add the component $w_t^j \in W_t \setminus w_t^{j-1}$ that maximizes mutual information to $y_{t+1}$ conditioning on the current components in $w_t^{j-1}$,
  
  $$w_t^j = \text{argmax}_{w \in W_t \setminus w_t^{j-1}} I(y_{t+1}; w|w_t^{j-1})$$

- Terminate if $I(y_{t+1}; w|w_t^{j-1})$ is zero.
The mixed embedding scheme

- Start with an empty embedding vector $w_t^0$, future vector (sample) of $Y$, $y_{t+1}$, and maximum lag $L_X$ for $X$ and $L_Y$ for $Y$. 
  $W_t = \{x_t, \ldots, x_{t-L_X-1}, y_t, \ldots, y_{t-L_Y-1}\}$

- First embedding cycle: $w_t^1 = \text{argmax}_{w \in W_t} I(y_{t+1}; w)$, and $w_t^1 = (w_t^1)$

- At embedding cycle $j$ suppose $w_{t}^{j-1} = (w_t^1, w_t^2, \ldots, w_t^{j-1})$. Add the component $w_t^j \in W_t \setminus w_t^{j-1}$ that maximizes mutual information to $y_{t+1}$ conditioning on the current components in $w_t^{j-1}$,
  
  $$w_t^j = \text{argmax}_{w \in W_t \setminus w_t^{j-1}} I(y_{t+1}; w|w_t^{j-1})$$

- Terminate if $I(y_{t+1}; w|w_t^{j-1})$ is zero.

Estimation of $I(y_{t+1}; w|w_t^{j-1})$ using nearest neighbors

[Kraskov et al, PRE, 2004]
Example: Coupled Mackey-Glass system
\[ \Delta = 17, 30, 100, \quad N = 4096 \]
\[ y_t^T = \{ y_{t+1}, y_{t+\tau_1}, y_{t+\tau_2} \}, \quad L_x = L_y = 50 \]

solid line: driving system \quad dashed line: response system
Terminate if $I(y_{t+1}; w|w^{j-1}_t)$ is \textbf{zero}. 

Termination of the mixed embedding scheme

Progressive vector building stops at step $j (w_t = w^{j-1}_t)$ if $\frac{I(y_{t+1}; w|w^{j-1}_t)}{I(y_{t+1}; w_j)} > A$ for a threshold $A < 1$ (here $A = 0.95$).

Randomization significance test for $H_0$: $I(y_{t+1}; w|w^{j-1}_t) = 0$.

Surrogate time series by shuffling randomly the components of the vector $w_j t$ and the rows of the matrix $w^{j-1}_t$. 

Laiou, Andrzejak and Kugiumtzis 

Non-Uniform Embedding Causality Measures
Termination of the mixed embedding scheme

Progressive vector building stops at step $j$ ($w_t = w_t^{j-1}$) if

$$I(y_{t+1}; w_t^{j-1}) / I(y_{t+1}; w_t^j) > A$$

for a threshold $A < 1$ (here $A = 0.95$).

Terminate if $I(y_{t+1}; w|w_t^{j-1})$ is zero.
Termination of the mixed embedding scheme

- Progressive vector building stops at step \( j \) (\( w_t = w_{j-1}^t \)) if
  \[ I(y_{t+1}; w_{j-1}^t) / I(y_{t+1}; w_j^t) > A \]
  for a threshold \( A < 1 \) (here \( A = 0.95 \)).

- Terminate if \( I(y_{t+1}; w|w_{j-1}^t) \) is statistically not significant.
  [Kugiumtzis, PRE, 2013]
Mutual Information from Mixed Embedding - 3

Terminate if $I(y_{t+1}; w|w_t^{j-1})$ is zero.

**Termination of the mixed embedding scheme**

- Progressive vector building stops at step $j$ ($w_t = w_t^{j-1}$) if
  
  $I(y_{t+1}; w_t^{j-1}) / I(y_{t+1}; w_t^j) > A$

  for a threshold $A < 1$ (here $A = 0.95$).

- Terminate if $I(y_{t+1}; w|w_t^{j-1})$ is statistically not significant. [Kugiumtzis, PRE, 2013]

  Randomization significance test for $H_0$: $I(y_{t+1}; w|w_t^{j-1}) = 0$. 

Terminate if $I(y_{t+1}; w|w_t^{j-1})$ is \textbf{zero}.

**Termination of the mixed embedding scheme**

- Progressive vector building stops at step $j$ ($w_t = w_t^{j-1}$) if
  \[ I(y_{t+1}; w_t^{j-1}) / I(y_{t+1}; w_t^j) > A \]
  for a \textbf{threshold $A < 1$} (here $A = 0.95$).

- Terminate if $I(y_{t+1}; w_t^{j-1})$ is \textbf{statistically not significant}.
  \[ [\text{Kugiumtzis, PRE, 2013}] \]

Randomization significance test for $H_0$: $I(y_{t+1}; w_t^{j-1}) = 0$.

Surrogate time series by shuffling randomly the components of the vector $w_t^j$ and the rows of the matrix $w_t^{j-1}$.
Terminate if $I(y_{t+1}; w_t^{j-1})$ is zero.

Termination of the mixed embedding scheme

- Progressive vector building stops at step $j$ ($w_t = w_t^{j-1}$) if
  
  
  $$I(y_{t+1}; w_t^{j-1})/I(y_{t+1}; w_t^j) > A$$

  for a threshold $A < 1$ (here $A = 0.95$).

- Terminate if $I(y_{t+1}; w|w_t^{j-1})$ is statistically not significant. [Kugiumtzis, PRE, 2013]

Randomization significance test for $H_0$: $I(y_{t+1}; w|w_t^{j-1}) = 0$.

Surrogate time series by shuffling randomly the components of the vector $w_t^j$ and the rows of the matrix $w_t^{j-1}$.
The non-uniform embedding vector of lags of all $X, Y$ for explaining $y_{t+1}$:

$$w_t = \left( x_{t-\tau x_1}, \ldots, x_{t-\tau x_{mx}}, y_{t-\tau y_1}, \ldots, y_{t-\tau y_{my}} \right)$$
The non-uniform embedding vector of lags of all $X, Y$ for explaining $y_{t+1}$:

$$
\mathbf{w}_t = (x_{t-\tau_{x1}}, \ldots, x_{t-\tau_{xmx}}, y_{t-\tau_{y1}}, \ldots, y_{t-\tau_{ymy}})
$$

The causality measure MIME

$$
R_{X\rightarrow Y} = \frac{I(y_{t+1}; \mathbf{w}^x_t \mid \mathbf{w}^y_t)}{I(y_{t+1}; \mathbf{w}_t)}
$$

- $R_{X\rightarrow Y}$: information on the future of $Y$ explained only by $X$-components of the embedding vector (given the components of $Y$), normalized with the mutual information of the future of $Y$ and the embedding vector.
The non-uniform embedding vector of lags of all $X$, $Y$ for explaining $y_{t+1}$:

$$w_t = \left( x_{t-\tau_1}, \ldots, x_{t-\tau_{x_m}}, y_{t-\tau_1}, \ldots, y_{t-\tau_{y_m}} \right)$$

The causality measure MIME

$$R_{X \rightarrow Y} = \frac{I(y_{t+1}; w^x_t | w^y_t)}{I(y_{t+1}; w_t)}$$

- $R_{X \rightarrow Y}$: information on the future of $Y$ explained only by $X$-components of the embedding vector (given the components of $Y$), normalized with the mutual information of the future of $Y$ and the embedding vector.

- If $w_t$ contains no components from $X$, then $R_{X \rightarrow Y} = 0$ and $X$ has no effect on the future of $Y$. 

Laiou, Andrzejak and Kugiumtzis

Non-Uniform Embedding Causality Measures
1. Granger causality measures
2. The Mutual Information from Mixed Embedding (MIME)
3. The Conditional Entropy from Non-uniform Embedding (Faes-Nollo-Porta, FNP)
4. Simulations and Evaluation of MIME and FNP
5. MIME and FNP applied to the Cardiovascular system
FNP makes two non-uniform embeddings, one on the basis of
\[ W_t = \{x_t, \ldots, x_{t-L_x-1}, y_t, \ldots, y_{t-L_y-1}\} \] (as in MIME)
and one on the basis of \[ W_t^Y = \{y_t, \ldots, y_{t-L_y-1}\} \]

[Faes, Nollo and Porta, PRE, 2011]
FNP makes two non-uniform embeddings, one on the basis of
\[ W_t = \{ x_t, \ldots, x_{t-L_x-1}, y_t, \ldots, y_{t-L_y-1} \} \] (as in MIME)
and one on the basis of \( W^Y_t = \{ y_t, \ldots, y_{t-L_y-1} \} \)

[Faes, Nollo and Porta, PRE, 2011]

- First embedding cycle: component minimizing conditional entropy of \( y_{t+1} \), \( w_t^1 = \arg\min_{w \in W_t} H(y_{t+1} | w) \), and \( w_t^1 = (w_t^1) \)
FNP makes two non-uniform embeddings, one on the basis of  
\[ W_t = \{x_t, \ldots, x_{t-L_x-1}, y_t, \ldots, y_{t-L_y-1}\} \] (as in MIME)  
and one on the basis of  
\[ W_t^Y = \{y_t, \ldots, y_{t-L_y-1}\} \]  

[Faes, Nollo and Porta, PRE, 2011]

- First embedding cycle: component minimizing conditional entropy of \( y_{t+1} \),  
  \[ w_t^1 = \arg\min_{w \in W_t} H(y_{t+1} | w) \], and  
  \[ w_t^1 = (w_t^1) \]

- At embedding cycle \( j \) suppose  
  \[ w_t^{j-1} = (w_t^1, w_t^2, \ldots, w_t^{j-1}) \]
Conditional Entropy from Non-uniform Embedding (FNP)

FNP makes two non-uniform embeddings, one on the basis of
\[ W_t = \{x_t, \ldots, x_{t-Lx-1}, y_t, \ldots, y_{t-Ly-1}\} \]
(as in MIME)
and one on the basis of \( W^Y_t = \{y_t, \ldots, y_{t-Ly-1}\} \)

[Faes, Nollo and Porta, PRE, 2011]

- First embedding cycle: component minimizing conditional entropy of \( y_{t+1} \), \( w^1_t = \text{argmin}_{w \in W_t} H(y_{t+1}|w) \), and \( w^1_t = (w^1_t) \)
- At embedding cycle \( j \) suppose \( w^{j-1}_t = (w^1_t, w^2_t, \ldots, w^{j-1}_t) \).
  Add the component \( w^j_t \in W_t \setminus w^{j-1}_t \):
  \[ w^j_t = \text{argmin}_{w \in W_t \setminus w^{j-1}_t} H(y_{t+1}|w; w^{j-1}_t) \]
Conditional Entropy from Non-uniform Embedding (FNP)

FNP makes two non-uniform embeddings, one on the basis of
\[ W_t = \{ x_t, \ldots, x_{t-L_x-1}, y_t, \ldots, y_{t-L_y-1} \} \] (as in MIME)
and one on the basis of \[ W^Y_t = \{ y_t, \ldots, y_{t-L_y-1} \} \]

[Faes, Nollo and Porta, PRE, 2011]

- First embedding cycle: component minimizing conditional entropy of \( y_{t+1} \), \( w^1_t = \text{argmin}_{w \in W_t} H(y_{t+1}|w) \), and \( w^1_t = (w^1_t) \)
- At embedding cycle \( j \) suppose \( w^{j-1}_t = (w^1_t, w^2_t, \ldots, w^{j-1}_t) \).
  Add the component \( w^j_t \in W_t \setminus w^{j-1}_t \):
  \[ w^j_t = \text{argmin}_{w \in W_t \setminus w^{j-1}_t} H(y_{t+1}|w; w^{j-1}_t) \]
- Terminate if \( H(y_{t+1}|w^{j-1}_t) < H(y_{t+1}|w^j_t) \), and then \( w_t = w^{j-1}_t \).
FNP makes two non-uniform embeddings, one on the basis of
\[ W_t = \{x_t, \ldots, x_{t-L_x-1}, y_t, \ldots, y_{t-L_y-1}\} \] (as in MIME)
and one on the basis of \[ W_t^Y = \{y_t, \ldots, y_{t-L_y-1}\} \]

[Faes, Nollo and Porta, PRE, 2011]

- First embedding cycle: component minimizing \textbf{conditional}
  entropy of \( y_{t+1} \), \( w^1_t = \text{argmin}_{w \in W_t} H(y_{t+1} | w) \), and \( w^1_t = (w^1_t) \)
- At embedding cycle \( j \) suppose \( w^{j-1}_t = (w^1_t, w^2_t, \ldots, w^{j-1}_t) \).
  Add the component \( w^j_t \in W_t \setminus w^{j-1}_t: \)
  \[ w^j_t = \text{argmin}_{w \in W_t \setminus w^{j-1}_t} H(y_{t+1} | w; w^{j-1}_t) \]
- Terminate if \( H(y_{t+1} | w^{j-1}_t) < H(y_{t+1} | w^j_t) \), and then \( w_t = w^{j-1}_t \).

Find \( w^Y_t \) similarly from \( W_t^Y \).
FNP makes two non-uniform embeddings, one on the basis of

\( W_t = \{ x_t, \ldots, x_{t-L_x-1}, y_t, \ldots, y_{t-L_y-1} \} \) (as in MIME)

and one on the basis of \( W_t^Y = \{ y_t, \ldots, y_{t-L_y-1} \} \)

[Faes, Nollo and Porta, PRE, 2011]

- First embedding cycle: component minimizing conditional entropy of \( y_{t+1} \), \( w^1_t = \arg\min_{w\in W_t} H(y_{t+1}|w) \), and \( w^1_t = (w^1_t) \)
- At embedding cycle \( j \) suppose \( w^{j-1}_t = (w^1_t, w^2_t, \ldots, w^{j-1}_t) \). Add the component \( w^j_t \in W_t \setminus w^{j-1}_t \):
  \[
  w^j_t = \arg\min_{w\in W_t \setminus w^{j-1}_t} H(y_{t+1}|w; w^{j-1}_t)
  \]
- Terminate if \( H(y_{t+1}|w^{j-1}_t) < H(y_{t+1}|w^j_t) \), and then \( w_t = w^{j-1}_t \).

Find \( w^Y_t \) similarly from \( W_t^Y \).

Estimation of conditional entropies using binning and correction (for single occupancies) [Porta et al, Biol Cybern, 1998]
The causality measure FNP

\[ C_{X \rightarrow Y} = 1 - \frac{H(y_{t+1}|w_t)}{H(y_{t+1}|w^Y_t)} \]

- \( C_{X \rightarrow Y} = 0 \): the conditional entropy of \( y_{t+1} \) is the same using the non-uniform embedding vector with or without \( X \)
  \( \implies \) \( X \) has no causal effect on \( Y \)
The causality measure FNP

\[ C_{X \rightarrow Y} = 1 - \frac{H(y_{t+1}|w_t)}{H(y_{t+1}|w^Y_t)} \]

- \( C_{X \rightarrow Y} = 0 \): the conditional entropy of \( y_{t+1} \) is the same using the non-uniform embedding vector with or without \( X \)
  \( \implies \) \( X \) has no causal effect on \( Y \)

- \( C_{X \rightarrow Y} > 0 \) the conditional entropy of \( y_{t+1} \) allowing \( X \) in the non-uniform embedding vector is smaller than when \( X \) is not considered
  \( \implies \) \( X \) has causal effect on \( Y \)
Conditional Entropy from Non-uniform Embedding (FNP)

The causality measure FNP

\[ C_{X \rightarrow Y} = 1 - \frac{H(y_{t+1} | w_t)}{H(y_{t+1} | w_Y^t)} \]

- \( C_{X \rightarrow Y} = 0 \): the conditional entropy of \( y_{t+1} \) is the same using the non-uniform embedding vector with or without \( X \)
  \( \implies X \) has no causal effect on \( Y \)

- \( C_{X \rightarrow Y} > 0 \): the conditional entropy of \( y_{t+1} \) allowing \( X \) in the non-uniform embedding vector is smaller than when \( X \) is not considered
  \( \implies X \) has causal effect on \( Y \)

- \( C_{X \rightarrow Y} = 1 \) if \( H(y_{t+1} | w_t) = 0 \)
  \( \implies y_{t+1} \) is a function of \( w_t \) (unlikely)
The causality measure FNP

\[ C_{X \rightarrow Y} = 1 - \frac{H(y_{t+1}|w_t)}{H(y_{t+1}|w^Y_t)} \]

- \( C_{X \rightarrow Y} = 0 \): the conditional entropy of \( y_{t+1} \) is the same using the non-uniform embedding vector with or without \( X \) 
  \( \implies X \) has no causal effect on \( Y \)
- \( C_{X \rightarrow Y} > 0 \) the conditional entropy of \( y_{t+1} \) allowing \( X \) in the non-uniform embedding vector is smaller than when \( X \) is not considered
  \( \implies X \) has causal effect on \( Y \)
- \( C_{X \rightarrow Y} = 1 \) if \( H(y_{t+1}|w_t) = 0 \)
  \( \implies y_{t+1} \) is a function of \( w_t \) (unlikely)
- \( C_{X \rightarrow Y} < 0 \) ?
The causality measure FNP

\[ C_{X \rightarrow Y} = 1 - \frac{H(y_{t+1} | w_t)}{H(y_{t+1} | w_t^Y)} \]

- \( C_{X \rightarrow Y} = 0 \): the conditional entropy of \( y_{t+1} \) is the same using the non-uniform embedding vector with or without \( X \)
  \( \implies X \) has no causal effect on \( Y \)
- \( C_{X \rightarrow Y} > 0 \) the conditional entropy of \( y_{t+1} \) allowing \( X \) in the non-uniform embedding vector is smaller than when \( X \) is not considered
  \( \implies X \) has causal effect on \( Y \)
- \( C_{X \rightarrow Y} = 1 \) if \( H(y_{t+1} | w_t) = 0 \)
  \( \implies y_{t+1} \) is a function of \( w_t \) (unlikely)
- \( C_{X \rightarrow Y} < 0 \) ? possible if \( H(y_{t+1} | w_t) > H(y_{t+1} | w_t^Y) \)
  \( (w_t^Y \) is not by construction subset of \( w_t \))
1. Granger causality measures

2. The Mutual Information from Mixed Embedding (MIME)

3. The Conditional Entropy from Non-uniform Embedding (Faes-Nollo-Porta, FNP)

4. Simulations and Evaluation of MIME and FNP

5. MIME and FNP applied to the Cardiovascular system
Example: coupled Henon maps

\[
x_{t+1} = 1.4 - x_t^2 + 0.3x_{t-1}
\]

\[
y_{t+1} = 1.4 - Cx_ty_t + (1 - C)y_t^2 + 0.3y_{t-1}
\]

coupling strength: \( C = 0, 0.1, \ldots, 0.6 \)

\( n = 1000, \) noise 10\%, 100 realizations, \( L = 10 \) MIME(\( k = 10 \)), FNP(bins=6)
Example: coupled Henon maps

\[ x_{t+1} = 1.4 - x_t^2 + 0.3x_{t-1} \]
\[ y_{t+1} = 1.4 - Cx_t y_t + (1 - C)y_t^2 + 0.3y_{t-1} \]

coupling strength: \( C = 0, 0.1, \ldots, 0.6 \)

\( n=1000 \), noise 10\%, 100 realizations, \( L=10 \) MIME(\( k=10 \)), FNP(bins=6)

<table>
<thead>
<tr>
<th>( X \rightarrow Y )</th>
<th>MIME</th>
<th>FNP</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R_{X,Y} )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( C_{X,Y} )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{Mean} )</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ X \rightarrow Y \] MIME

\[ Y \rightarrow X \] MIME

Laiou, Andrzejak and Kugiumtzis
Non-Uniform Embedding Causality Measures
Example: coupled Henon maps - 2

row 1: causality values vs C, row 2: percentage of rejection for the surrogate significance test (time-shifted surrogates)
col 1: \( n = 100 \), noise-free, MIME\((k = 1)\), FNP\((\text{bins}=6)\)
col 2: \( n = 1000 \), noise 10\%, MIME\((k = 10)\), FNP\((\text{bins}=6)\)
blue: MIME, cyan: FNP
Example: Rössler-Lorenz system

coupling strength: $C = 0, 0.5, 1, 1.5, 2, 2.5, 3$

$n = 1000$, noise 10%, 100 realizations, $L = 10$ MIME($k = 10$), FNP(bins=6)

$X \rightarrow Y$ MIME

$X \rightarrow Y$ FNP

$X \rightarrow Y$ Mean

$Y \rightarrow X$ MIME

$Y \rightarrow X$ FNP

$Y \rightarrow X$ Mean

Laiou, Andrzejak and Kugiumtzis  Non-Uniform Embedding Causality Measures
Example: Coupled Mackey-Glass system ($\Delta_1 = \Delta_2 = 30$)

coupling strength: $C = 0, 0.1, \ldots, 0.6$

$n = 1000$, noise 10%, 100 realizations, $L = 10$ MIME($k = 10$),
FNP(bins=6)

\[ X \rightarrow Y \quad \text{MIME} \]

\[ Y \rightarrow X \quad \text{MIME} \]

\[ \text{FNP} \]

\[ \text{Mean} \]

Laiou, Andrzejak and Kugiumtzis
Non-Uniform Embedding Causality Measures
1. Granger causality measures
2. The Mutual Information from Mixed Embedding (MIME)
3. The Conditional Entropy from Non-uniform Embedding (Faes-Nollo-Porta, FNP)
4. Simulations and Evaluation of MIME and FNP
5. MIME and FNP applied to the Cardiovascular system
Inter-dependence of all pairs of
- heart variability ($X$),
- respiration ($Y$),
- blood oxygen concentration ($Z$)

Data from the Santa Fe Competition. [Rigney et al 1993]
Two stationary segments of $X$, $Y$ and $Z$ normalized, $n = 1200$
A: signs of apnea, B: rather normal activity.
Causality in the cardiovascular system

Inter-dependence of all pairs of

- heart variability ($X$),
- respiration ($Y$),
- blood oxygen concentration ($Z$)

Data from the Santa Fe Competition. [Rigney et al 1993]
Two stationary segments of $X$, $Y$ and $Z$ normalized, $n = 1200$
A: signs of apnea, B: rather normal activity.

<table>
<thead>
<tr>
<th></th>
<th>A ($X$)</th>
<th>B ($Y$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIME</td>
<td>FNP</td>
<td>MIME</td>
</tr>
<tr>
<td>$X \rightarrow Y$</td>
<td>0.11</td>
<td>0.01</td>
</tr>
<tr>
<td>$Y \rightarrow X$</td>
<td>0</td>
<td>-0.01</td>
</tr>
<tr>
<td>$X \rightarrow Z$</td>
<td>0</td>
<td>-0.00</td>
</tr>
<tr>
<td>$Z \rightarrow X$</td>
<td>0</td>
<td>0.06</td>
</tr>
<tr>
<td>$Y \rightarrow Z$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$Z \rightarrow Y$</td>
<td>0.12*$</td>
<td>0.02</td>
</tr>
</tbody>
</table>
Both MIME and FNP capture sufficiently the causality effects.

For small time series, $k$ in MIME has to be decreased accordingly.

For very weak coupling, the sensitivity of MIME and FNP varies.

For the cardio-respiratory-vascular system, both MIME and FNP detect correctly respiratory sinus arrhythmia, but they detect other weak couplings differently.

Work in progress...

Evaluation of 4 combinations: MIME and FNP with $k$-nearest neighbor and binning.

Laiou, Andrzejak and Kugiumtzis  Non-Uniform Embedding Causality Measures
Summary

- Both MIME and FNP capture sufficiently the causality effects.
• Both MIME and FNP capture sufficiently the causality effects
• pros and cons for MIME and FNP:

  - FNP gets negative values (when no coupling)
  - For small time series, $k$ in MIME has to be decreased accordingly.
  - For very weak coupling, the sensitivity of MIME and FNP varies.

• For the cardio-respiratory-vascular system, both MIME and FNP detect correctly respiratory sinus arrhythmia, but they detect other weak couplings differently.

Work in progress...

Evaluation of 4 combinations: MIME and FNP with $k$-nearest neighbor and binning.
Summary

- Both MIME and FNP capture sufficiently the causality effects
- Pros and cons for MIME and FNP:
  - FNP gets negative values (when no coupling)

For small time series, $k$ in MIME has to be decreased accordingly.
For very weak coupling, the sensitivity of MIME and FNP varies.

For the cardio-respiratory-vascular system, both MIME and FNP detect correctly respiratory sinus arrhythmia, but they detect other weak couplings differently.

Work in progress...

Evaluation of 4 combinations: MIME and FNP with $k$-nearest neighbor and binning.
Summary

- Both MIME and FNP capture sufficiently the causality effects
- Pros and cons for MIME and FNP:
  - FNP gets negative values (when no coupling)
  - For small time series, $k$ in MIME has to be decreased accordingly.

Evaluation of 4 combinations: MIME and FNP with $k$-nearest neighbor and binning.
Summary

- Both MIME and FNP capture sufficiently the causality effects
- pros and cons for MIME and FNP:
  - FNP gets negative values (when no coupling)
  - For small time series, $k$ in MIME has to be decreased accordingly.
  - For very weak coupling, the sensitivity of MIME and FNP varies.

For the cardio-respiratory-vascular system, both MIME and FNP detect correctly respiratory sinus arrhythmia, but they detect other weak couplings differently.

Work in progress...

Evaluation of 4 combinations: MIME and FNP with $k$-nearest neighbor and binning.
Summary

• Both MIME and FNP capture sufficiently the causality effects
• pros and cons for MIME and FNP:
  - FNP gets negative values (when no coupling)
  - For small time series, $k$ in MIME has to be decreased accordingly.
  - For very weak coupling, the sensitivity of MIME and FNP varies.
• For the cardio-respiratory-vascular system, both MIME and FNP detect correctly respiratory sinus arrhythmia, but they detect other weak couplings differently.

Work in progress...
Evaluation of 4 combinations: MIME and FNP with $k$-nearest neighbor and binning.

Laiou, Andrzejak and Kugiumtzis Non-Uniform Embedding Causality Measures
Summary

- Both MIME and FNP capture sufficiently the causality effects
- pros and cons for MIME and FNP:
  - FNP gets negative values (when no coupling)
  - For small time series, $k$ in MIME has to be decreased accordingly.
  - For very weak coupling, the sensitivity of MIME and FNP varies.
- For the cardio-respiratory-vascular system, both MIME and FNP detect correctly respiratory sinus arrhythmia, but they detect other weak couplings differently.

Work in progress...
Evaluation of 4 combinations: MIME and FNP with $k$-nearest neighbor and binning.
References


