Universidad
Carlos III de Madrid

Robotics Lab

Metrics for Assistive Robotics Brain-Computer Interface Evaluation

Martin F. Stoelen, Javier Jiménez, Alberto Jardón, Juan G. Víctores
José Manuel Sánchez Pena, Carlos Balaguer
Universidad Carlos III de Madrid, Spain

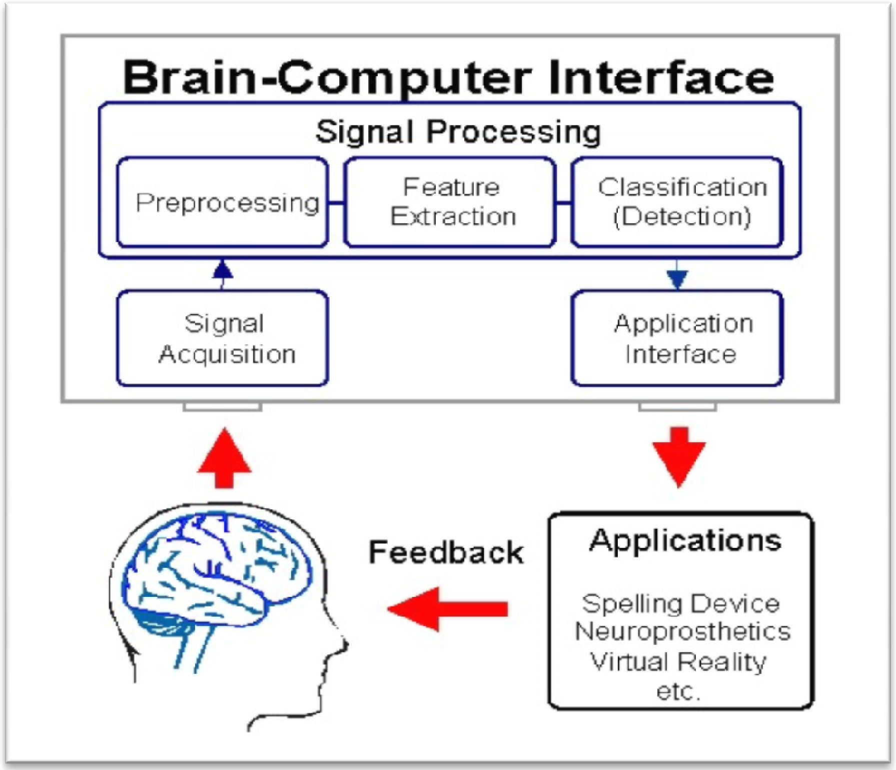
F. Bonsignorio
Heron Robots, Italy and Universidad Carlos III de Madrid, Spain

Workshop on Performance Evaluation and Benchmarking for Intelligent Robots and Systems with Cognitive and Autonomy Capabilities at IROS 2010

Motivation


ROBOT DE ASISTENCIA PERSONAL:
MATS
(Flexible Mechatronics Assistive Technology
System to support persons with special
needs in all their living and working environment
Mechatronics Assistive Technology
System for ederly and disabled people)

Brain-Computer Interfaces (BCI)



Grätzel, Philipp. Brain Tunning. BCI. [Online] (Cited: 07 06, 2009) http://www.brain-tuning.de/2ndlevel/bci/compo1_b.htm

BCI classifications

- Synchronous vs asynchronous
  - External cues (e.g. lights) vs user-initiated action
  - User in “control”: asynchronous
- Dependent vs independent
  - Requires muscle/nerve activity vs EEG only
  - Assistive robotics: mainly independent
- Discrete vs continuous output
  - End-of-trial vs continuous feedback to user
  - Assistive robot low-level control: continuous

Existing metrics for asynchronous continuous BCI

- Mean Squared Error (MSE)

$$MSE = 1/N \cdot \sum_{t=1}^{N} ((x_t - \hat{x}_t)^2$$

- Correlation Coefficient

$$CC_x = \frac{\sum_i (x_i - \bar{x})(\hat{x}_i - \bar{\hat{x}})}{\sqrt{(\sum_i (x_i - \bar{x})^2)(\sum_i (\hat{x}_i - \bar{\hat{x}})^2)}}$$

- Information Transfer Rate (ITR)
  - Typically applied for discrete BCI output
  - Interesting concept: BCI as a communication channel

BCI for assistive robotics

- Current BCIs lack throughput for application*
- Many interesting challenges **
  - Hybrid BCIs?
  - Simultaneous co-adaptation of user and system?
  - How to improve performance and reliability?
- User and robot in closed-loop on complex tasks
  - Analysis for complete human-bci-machine system?
  - Metrics for complete human-bci-machine system?

**Tonet et al. (2008)*
*** Millán et al. (2010)*

Control system as a directed acyclic graph*

- Current state, X
- Future state, X'
- System controller, C:
  - Sensor, S
  - Actuator, A
- Exists theorems for:
  - Observability
  - Controllability
  - Optimality

Open-loop control

$$I(X;C) = 0$$

Closed-loop control

$$I(X;C) > 0$$

**Touchette and Lloyd (2004)*

Control system as a directed acyclic graph*

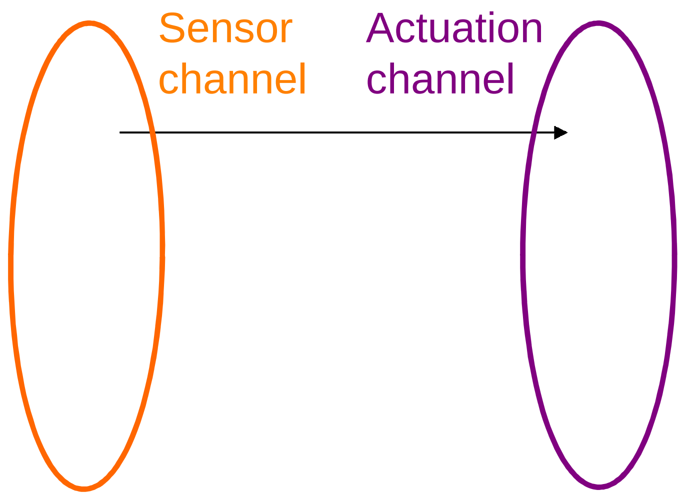
- Current state, X
- Future state, X'
- System controller, C:
  - Sensor, S
  - Actuator, A
- Exists theorems for:
  - Observability
  - Controllability
  - Optimality

Open-loop control

$$I(X;C)=0$$

Closed-loop control

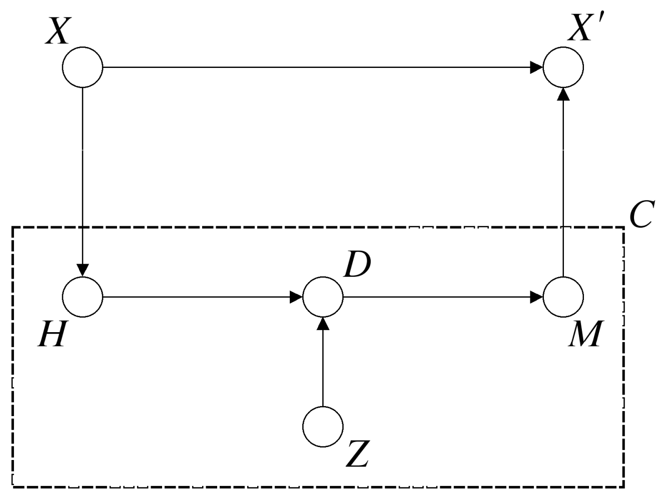
$$I(X;C)>0$$



*Touchette and Lloyd (2004)

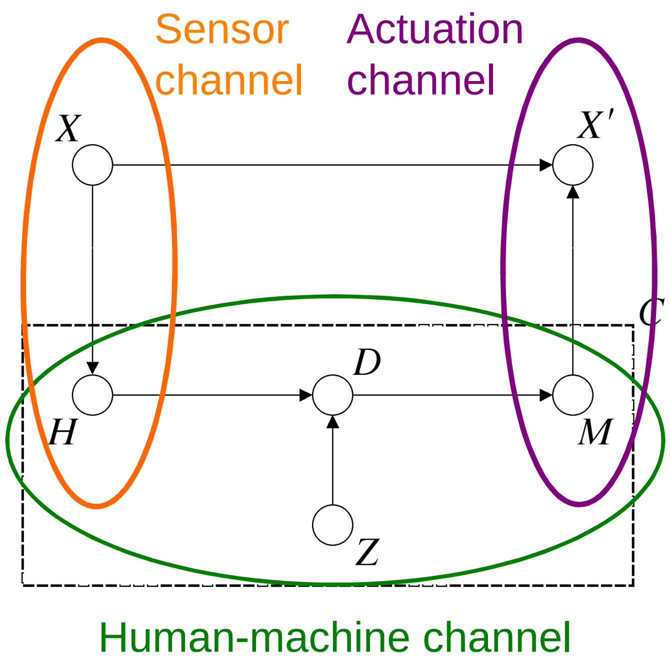
Human-machine system as a directed acyclic graph

- Current state, X
- Future state, X'
- System controller, C:
  - User (human, H)
  - Robot (machine, M)
  - Input device (D)
  - Disability of user (Z)
- Interested in channels - flow of information



Human-machine system as a directed acyclic graph

- Current state, X
- Future state, X'
- System controller, C:
  - User (human, H)
  - Robot (machine, M)
  - Input device (D)
  - Disability of user (Z)
- Interested in channels - flow of information



The human-machine channel

- Objective: **Transmit information of user's intent (H) over noisy channel of capacity C_{HM} with a minimum of errors**

- Entropy*: $S(X) = -\sum_{x \in X} p(x) \log(p(x))$

- A mentally and physically healthy user: $S(H) \leq C_{HM}$

- A mentally healthy, physically disabled user: $S(H) > C_{HM}$

- How to measure C_{HM}? Other channels?

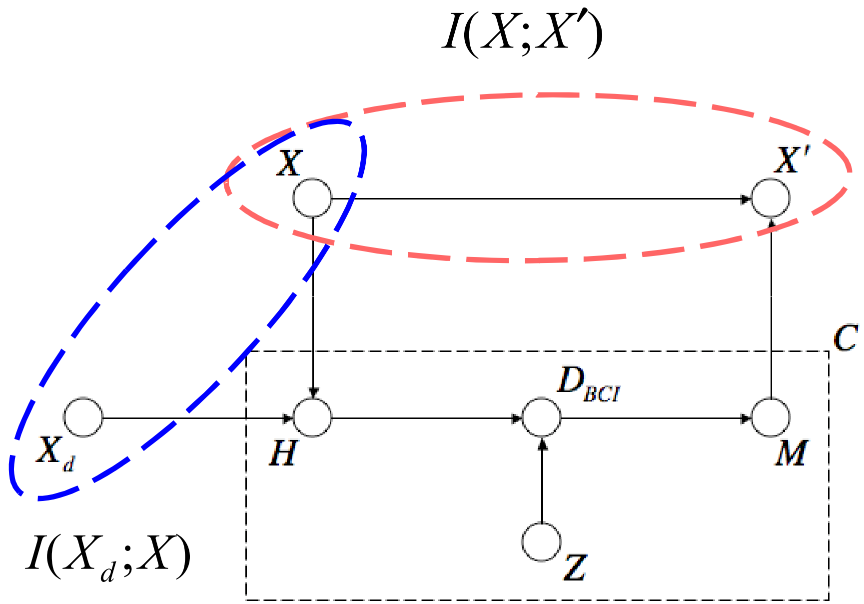
**Shannon (1948)*

Metrics considered

1. “Empowerment”*
   - “… agent’s potential ability to influence the environment…”
   - Information Transfer Rate (ITR)

2. Predictive information
   - How “predictable” is system?

3. Mean Squared Error (MSE)



$$I(X;Y) = \sum_{y \in Y} \sum_{x \in X} p(x,y) \log\left(\frac{p(x,y)}{p(x)\,p(y)}\right)$$

Klyubin, Polani and Nehaniv (2008)

BCI apparatus

- Data collection:
  - Amplifier: gMobilab+
  - Electrodes: g.EEGelectrode
- Data processing:
  - BCI2000*
- BCI paradigm used:
  - ERD/ERS paradigm
  - Mu/beta rhythm (8-30 Hz)
  - E.g. left hand vs both feet

**Schalk (2004)*

Pursuit tracking pilot study

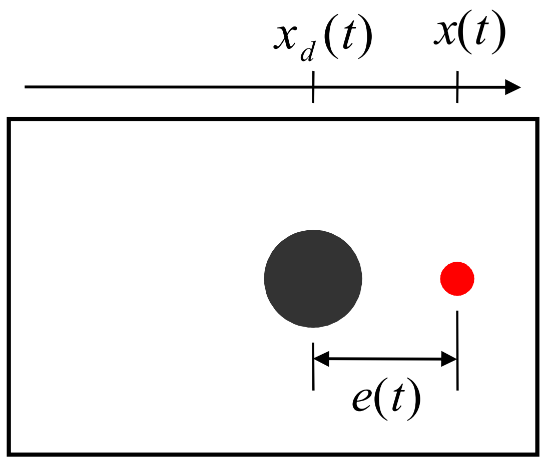
ERS: Event Related Synchronization -> post-movement period and relaxation, a steady rhythm when not thinking of or moving limb of interest

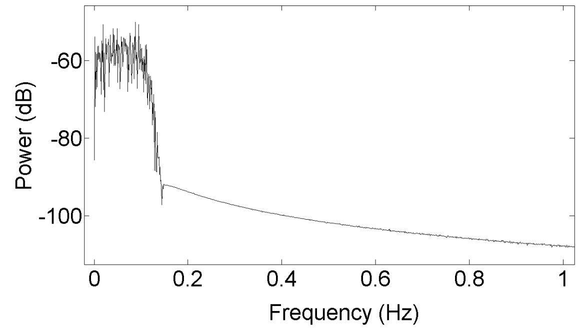
ERD: Event Related Desynchronization -> movement or preparation for movement, a reduction of amplitude of above rhythm when thinking of or moving limb of interest

For our testing we typically used amplitude of signal in area of brain (and at specific frequency) activated by left hand to indicate moving in one direction and both feet in the other direction.

Pursuit tracking pilot study

- Continuous and asynchronous task:
  - *Cursor* and *target* disc
- 3 user signal types:
  - User 1 (real limb movements)
  - User 2 (imaginary limb movements)
  - Rest, no visual (user 1)
- 3 target signal types:
  - Sinusoidal, 0.05 Hz
  - Sinusoidal, 0.1 Hz
  - “Random” (white, low-pass 0.1 Hz)



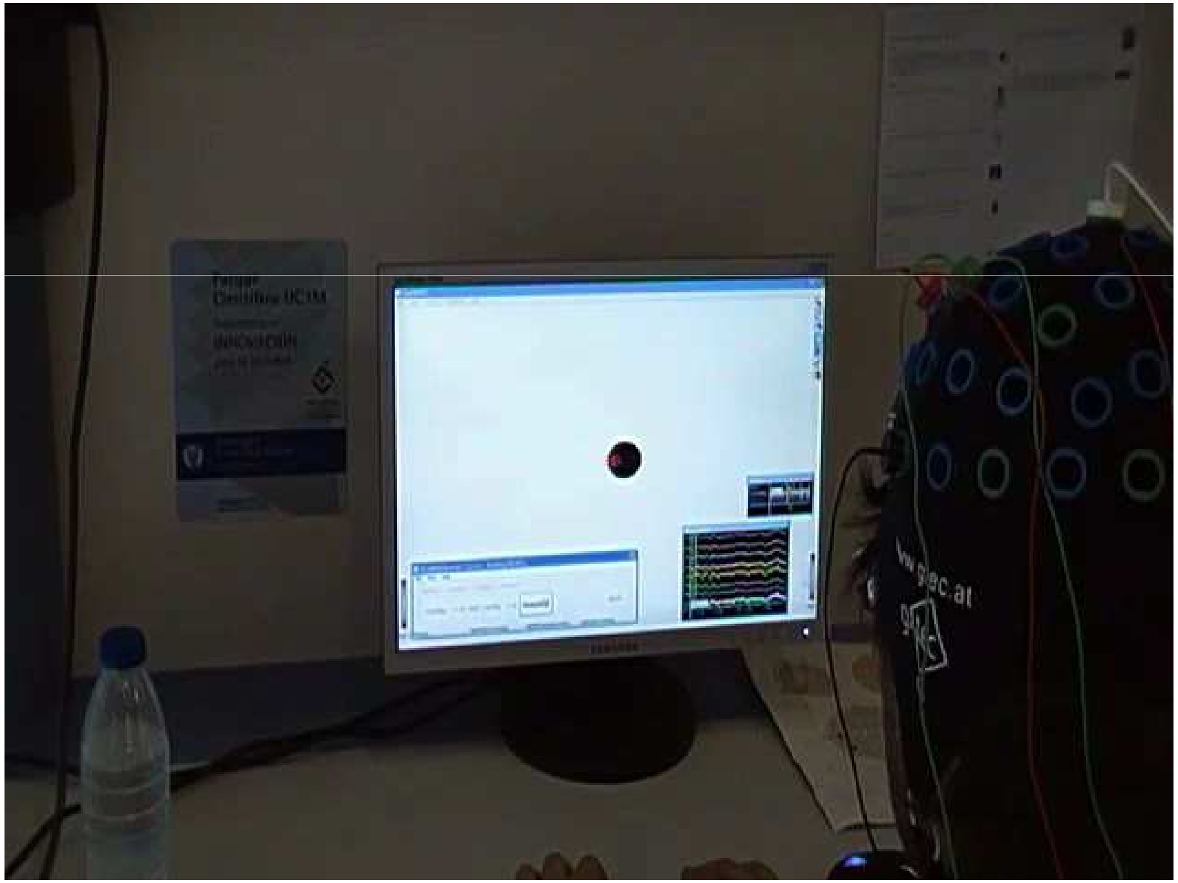


Used Matlab toolbox of Lungarella, Pegors, Bulwinkle and Sporns (2005)

Pursuit tracking pilot study

Although we do provide a desired trajectory, this task is asynchronous in that we do NOT use an external stimuli to improve our chances at interpreting the EEG's. This IS done with for example the P300 spellers that blink one symbol at a time and then measure the response in the visual areas of the brain 300 ms later.

Pursuit tracking pilot study



Preliminary results: Mean Squared Error

	Target signal:		
User signal:	*0.05 Hz*	*0.1 Hz*	*"Random"*
User 1	0.13	0.14	TODO
User 2	0.12	0.10	0.07
Rest, user 1	0.17	0.19	0.10

Data collection and conditioning:
•5 runs with 3 trials each, all trials lasted 26 seconds and was sampled at ~ 30Hz

Discussion

- Note on the low rest vs “random” result:

The “random” target signal tends to stay closer to zero than the sinusoidal target signal. The rest condition should also keep the user signal reasonably close to zero with the continuous online normalization used. This might have affected the low MSE here.

Preliminary results: "Empowerment"

	Target signal:		
User signal:	*0.05 Hz*	*0.1 Hz*	*"Random"*
User 1	0.09	0.03	TODO
User 2	0.12	0.08	0.11
Rest, user 1	0.03	0.02	0.03

Data collection and conditioning:
•5 runs with 3 trials each, all trials lasted 26 seconds and was sampled at ~ 30Hz
•Data normalized and discretized to 25 states

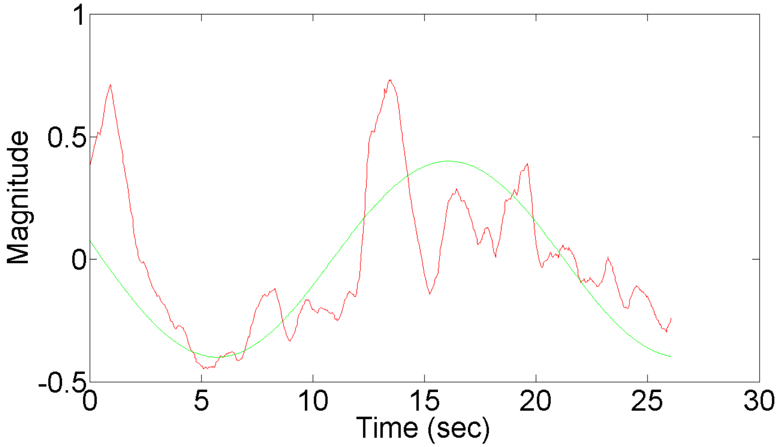
Preliminary results: Predictive Information

	Target signal:		
User signal:	*0.05 Hz*	*0.1 Hz*	*"Random"*
User 1	2.16	1.64	TODO
User 2	2.26	2.08	2.26
Rest, user 1	1.44		

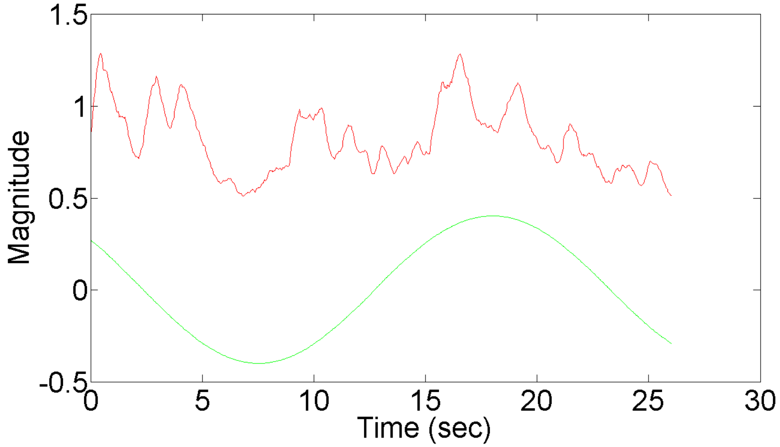
Data collection and conditioning:
- *5 runs with 3 trials each, all trials lasted 26 seconds and was sampled at ~ 30Hz*
- *Data normalized and discretized to 25 states*
- *Time binned at ~10 Hz (3 samples per bin)*

Issue 1: Stationarity of user signal

- Mu/beta rhythms are typically non-stationary
  MSE is sensitive to outliers in data
- “Good” online BCI normalization can be difficult:



A “good” trial

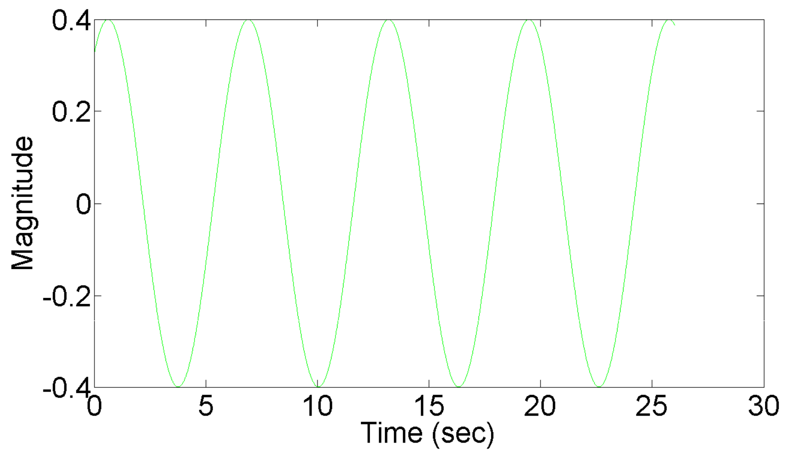


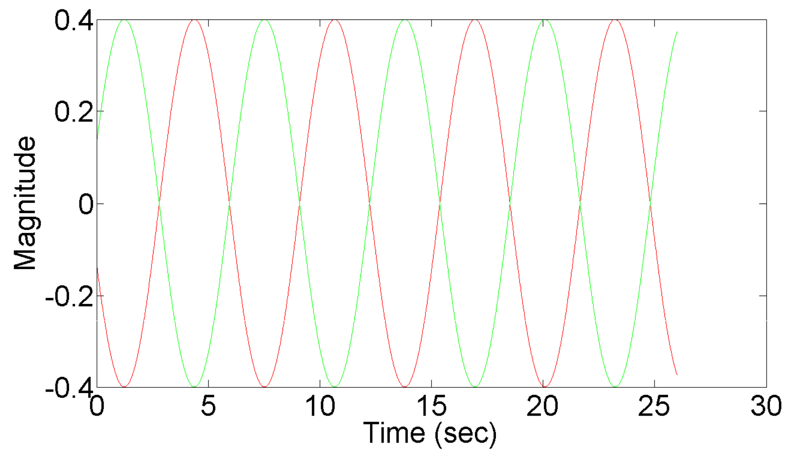
A “bad” trial

Discussion

- we used a normalization module of the BCI2000 application that performed continuous normalization on the BCI output during the trials, to make sure the mean was kept close to zero and the variance unity. The algorithm used all the data received for the current run (3 trials of 26 seconds).

Issue 2: In-phase and anti-phase for periodic tasks





MSE = 0

$I(X;X_d) \approx 3.1$

$I(X;X') \approx 2.9$

MSE ≈ 0.3

$I(X;X_d) \approx 3.1$

$I(X;X') \approx 2.9$

Data collection and conditioning:
- *15 trials, all trials lasted 26 seconds and was sampled at ~ 100Hz*
- *Random phase offset added to each trial (same for user and target signal)*
- *Data normalized and discretized to 10 states*

Discussion

- "empowerment" cannot distinguish between an exact in-phase or anti-phase response. This makes it less suitable for quantifying how "close" the user is to the tracking signal in sinusoidal tracking tasks, especially if there is a lot of lag in the user's response.

Discussion

- In general - not an easy task for BCI used
  - Not sufficient performance nor data for firm conclusions
  - But feasible to distinguish rest vs intentional movement
- Task and purpose dictates metric to use
  - MSE is particularly sensitive to outliers
    - BCI artifacts (e.g. blinking) must be considered with care
  - “Empowerment” is ambiguous for in and anti-phase
    - Preferably applied to non-periodic tasks?
  - Predictive information does not require X_d
    - Potential for use outside experimental setting?

Future Work

- Improve mu/beta rhythm control paradigm
- More extensive controlled experiments
- Apply metrics to actual assistive tasks?
- Apply metrics for motivating online adaptation?