CCS 2023 Satellite Session
Recent Advances in Learning and
Data-Driven Modeling of Complex Systems
19th October 2023, Salvador, Bahia, Brazil

INFERENCE OF DIRECTED INTERACTIONS IN COLLECTIVE DYNAMICS THROUGH INFORMATION-THEORETIC METRICS

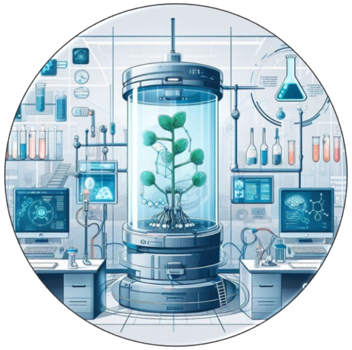
Pietro De Lellis
University of Naples Federico II

(joint-work with Maurizio Porfiri, New York University,
Manuel Ruiz Marin, Universidad Politécnica de Cartagena)

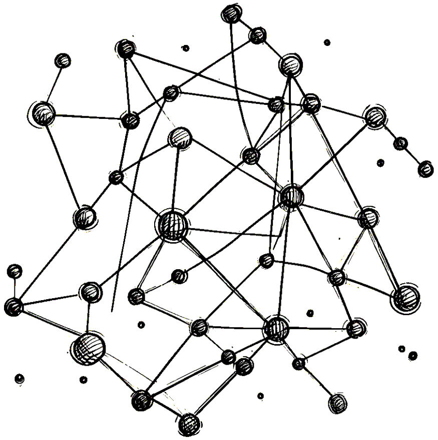
Reconstructing Directed Interactions

- Pairwise interactions are critical to collective dynamics of natural and technological complex systems





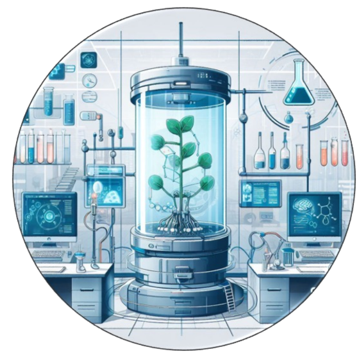

- Information theory has been widely employed to reconstruct such interactions from time series of the units composing the system



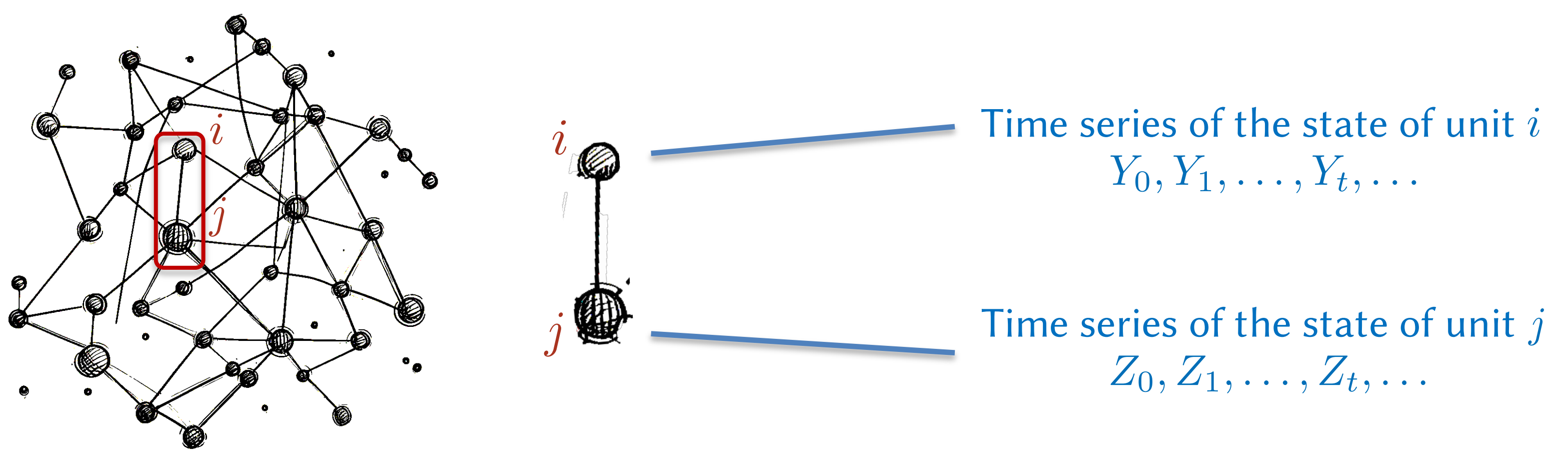
Reconstructing Directed Interactions

- Pairwise interactions are critical to collective dynamics of natural and technological complex systems





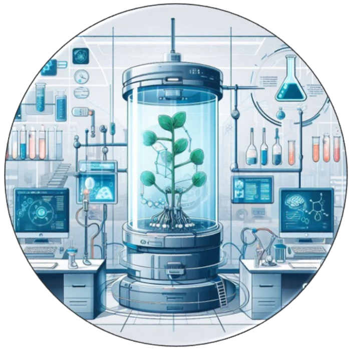

- Information theory has been widely employed to reconstruct such interactions from time series of the units composing the system



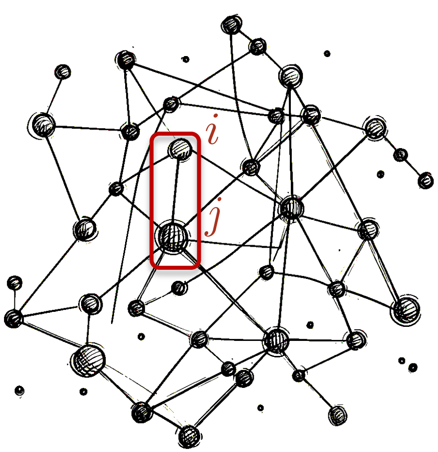
Reconstructing Directed Interactions

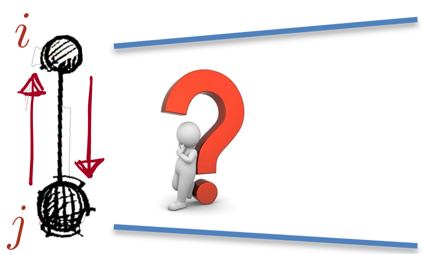
- Pairwise interactions are critical to collective dynamics of natural and technological complex systems






- Information theory has been widely employed to reconstruct such interactions from time series of the units composing the system

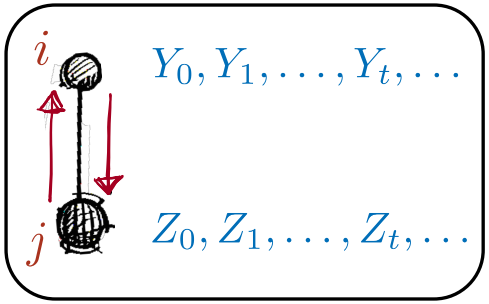




Time series of the state of unit i
$Y_0, Y_1, \ldots, Y_t, \ldots$

Time series of the state of unit j
$Z_0, Z_1, \ldots, Z_t, \ldots$

Information-theoretic Metrics



- The most basic information theoretic tool to study dependencies between two dynamical systems is based on mutual information

$$\mathrm{MI}^{Z \to Y} = I(Z_t; Y_{t+1})$$

$$= \sum_{\substack{y_{t+1} \in \mathcal{Y} \\ z_t \in \mathcal{Z}}} \Pr(Y_{t+1} = y_{t+1}, Z_t = z_t) \log_2 \frac{\Pr(Y_{t+1} = y_{t+1} | Z_t = z_t)}{\Pr(Y_{t+1} = y_{t+1})}$$

- Time-delayed mutual information is symmetric by definition, and a nonzero value can be observed even when Y_{t+1} is not influenced by Z_t, as the result of the memory of past states of Y. Transfer entropy was then introduced

$$\mathrm{TE}^{Z \to Y} = I(Z_t; Y_{t+1} | Y_t)$$

$$= \sum_{\substack{y_t, y_{t+1} \in \mathcal{Y} \\ z_t \in \mathcal{Z}}} \Pr(Y_{t+1} = y_{t+1}, Z_t = z_t, Y_t = y_t) \log_2 \frac{\Pr(Y_{t+1} = y_{t+1} | Z_t = z_t, Y_t = y_t)}{\Pr(Y_{t+1} = y_{t+1} | Y_t = y_t)}.$$

[1] Schreiber, *Physical Review Letters*, 85, 461, 2000.

Transfer Entropy: Pros and Cons

Pros

- Allows to study asymmetric interactions
- Can be readily calculated from time-series
- Hypothesis testing is easy to perform (e.g. via permutation tests[2])
- Can be extended to consider multivariate interactions[3], adapted to be used for short time-series through symbolization[4]

A key limitation

- As pointed out by James *et al.* (*PRL*, 2016; *Science Advances*, 2022), transfer entropy is sensitive to both intrinsic dependencies between Z_t and Y_{t+1}, as well as the dependencies induced by Y_t.
- To filter out these dependencies and precisely assess information flow, Sattari *et al.*[5] proposed to use intrinsic mutual information

[2]Runge, *Chaos*, 28, 075310, 2018
[3]Sun and Bollt, *Physica D*, 267, 49-57, 2014
[4]Staniek and Lehnertz, *Physical Review Letters*, 100, 158101, 2008
[5]Sattari *et al*, *Science Advances*, 8, 1-13, 2022

Intrinsic Mutual Information

$$\mathrm{IMI}^{Z\to Y} = \inf\left\{ \sum_{\substack{\overline{y}_t, y_{t+1}\in\mathcal{Y}\\ z_t\in\mathcal{Z}}} \Pr(Y_{t+1}=y_{t+1}, \overline{Y}_t=\overline{y}_t, Z_t=z_t) \times \log_2 \frac{\Pr(Y_{t+1}=y_{t+1}|\overline{Y}_t=\overline{y}_t, Z_t=z_t)}{\Pr(Y_{t+1}=y_{t+1}|\overline{Y}_t=\overline{y}_t)} : \Pr(Y_{t+1}, Z_t, \overline{Y}_t) = \sum_{y_t\in\mathcal{Y}} \Pr(Y_{t+1}, Z_t, Y_t=y_t)\Pr(\overline{Y}_t|Y_t=y_t) \right\}.$$

- From its definition, IMI reduces to time-delayed mutual information when the minimization process yields a constant $\overline{Y}_t$, and to transfer entropy if one gets $\overline{Y}_t = Y_t$
- We then have $\mathrm{IMI}^{Z\to Y} \leq \min\{\mathrm{TE}^{Z\to Y}, \mathrm{MI}^{Z\to Y}\}$
- Albeit it was shown that IMI is more accurate in measuring information flow compared to TE, this does not necessarily imply that it is a better instrument for inferring directional interactions

Comparing IMI, TE, and MI

- A key step in the application of information-theoretic construct to infer directional interactions is hypothesis testing
- This requires contrasting observed values against data obtained under the null hypothesis of independence

Our goal is to

- Clarify the relationship between IMI and the classical metrics of information flow
  - To this aim, we will introduce a minimalistic Boolean model, where IMI, TE, and MI can be exactly computed
- Compare IMI, TE, and MI in terms of their ability to detect leader-follower interactions

Comparing IMI, TE, and MI

- A key step in the application of information-theoretic construct to infer directional interactions is hypothesis testing
- This requires contrasting observed values against data obtained under the null hypothesis of independence

Our goal is to

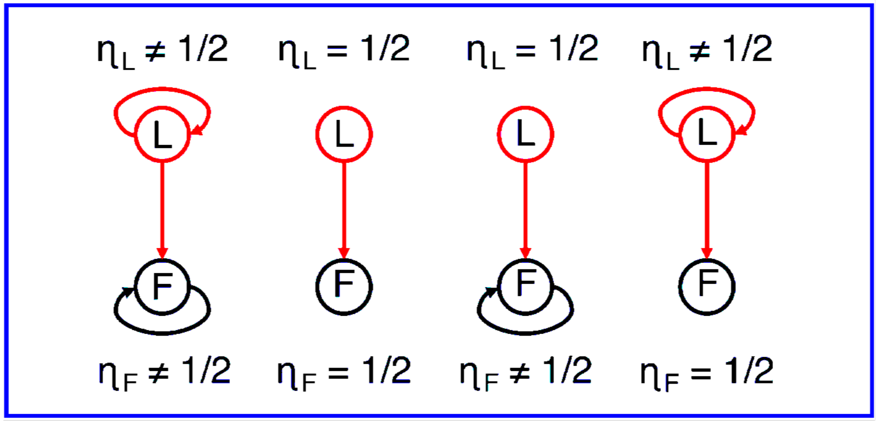
- Clarify the relationship between IMI and the classical metrics of information flow
  - To this aim, we will introduce a minimalistic Boolean model, where IMI, TE, and MI can be exactly computed
- Compare IMI, TE, and MI in terms of their ability to detect leader-follower interactions

Boolean Leader-Follower Model

- We consider two Boolean random processes X_t^L and X_t^F

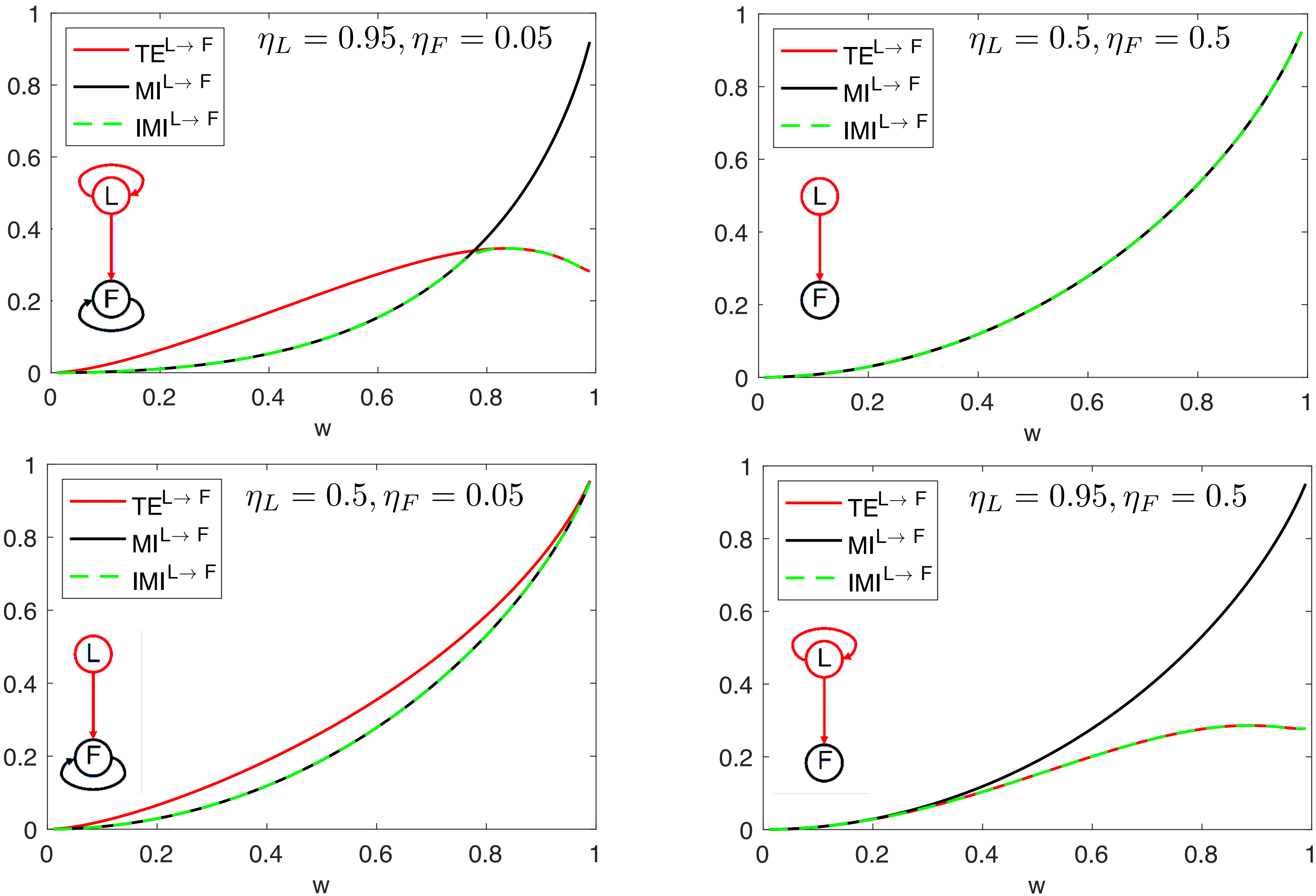
$$X_{t+1}^L = \begin{cases} X_t^L, & \text{with probability } (1-\eta_L), \\ 1 - X_t^L, & \text{with probability } \eta_L, \end{cases}$$

$$X_{t+1}^F = \begin{cases} X_t^F, & \text{with probability } (1-\eta_F)(1-w) + |1 - X_t^L - X_t^F|w, \\ 1 - X_t^F, & \text{with probability } \eta_F(1-w) + |X_t^L - X_t^F|w, \end{cases}$$

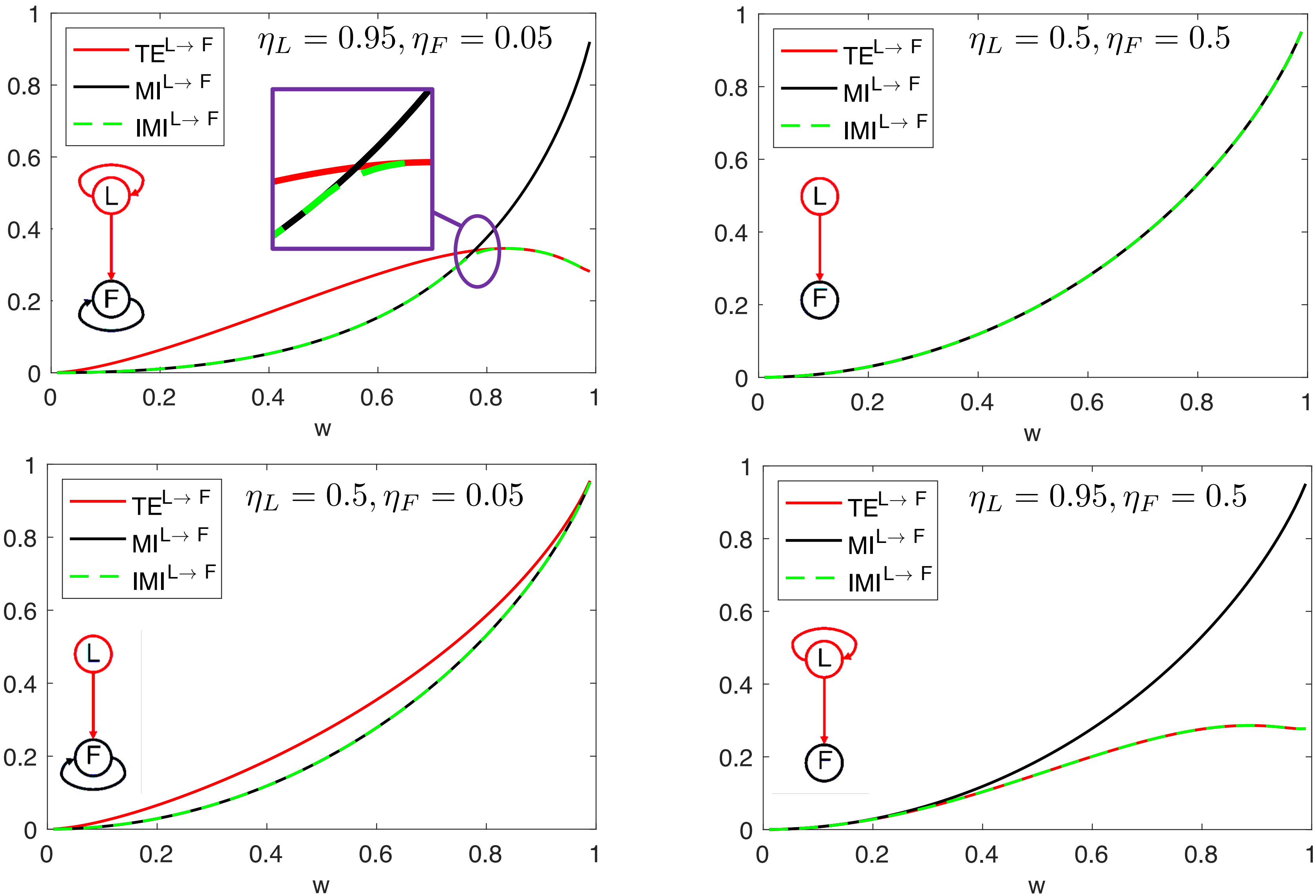


- For this simple model, we can compute exact solutions of the information-theoretic model as function of the model parameters

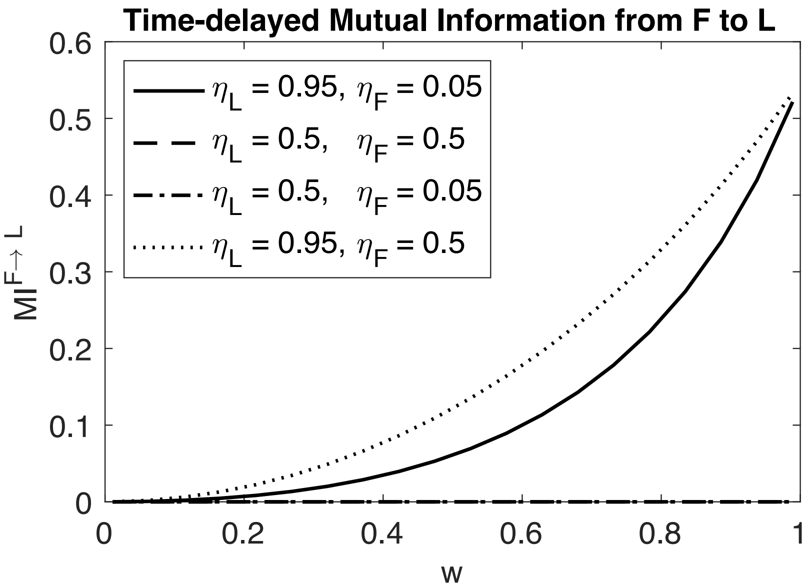
IMI, TE, and MI from the Leader to the Follower



IMI, TE, and MI from the Leader to the Follower

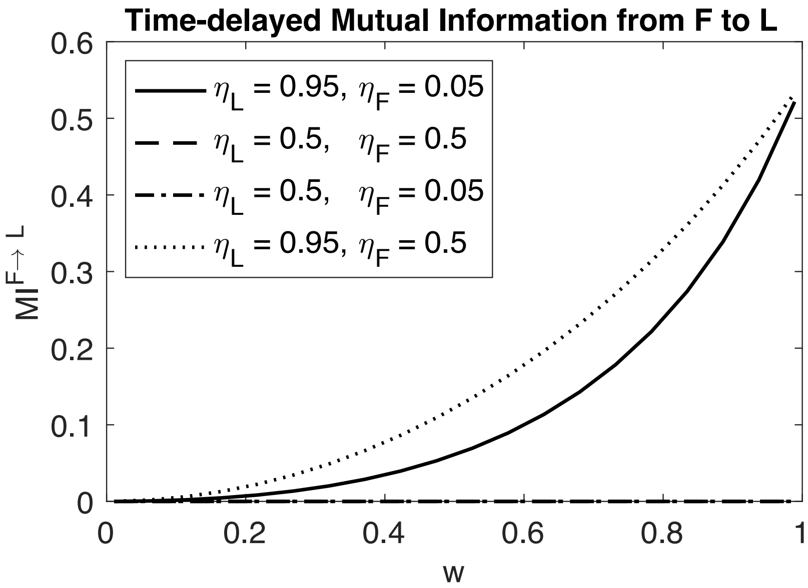


IMI, TE, and MI from the Follower to the Leader



- From the follower to the leader, transfer entropy and intrinsic mutual information are always zero
- Time-delayed mutual information can be instead different from zero due to the shared history between the leader and the follower

IMI, TE, and MI from the Follower to the Leader



- From the follower to the leader, transfer entropy and intrinsic mutual information are always zero
- Time-delayed mutual information can be instead different from zero due to the shared history between the leader and the follower

Overall, intrinsic mutual information is very well approximated by the minimum between time-delayed mutual information and transfer entropy, except for a narrow window of coupling gains in the case when both the leader and the follower have memory

Comparing IMI, TE, and MI

- A key step in the application of information-theoretic construct to infer directional interactions is hypothesis testing
- This requires contrasting observed values against data obtained under the null hypothesis of independence

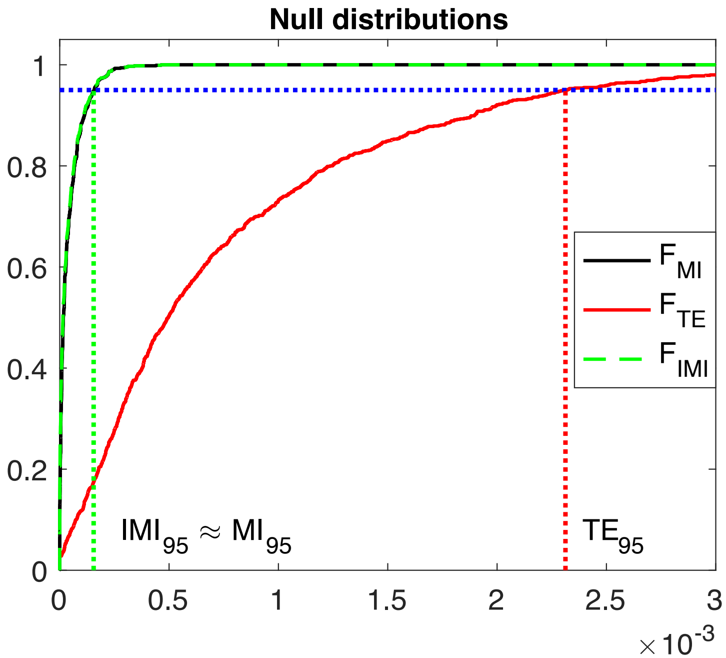
Our goal is to

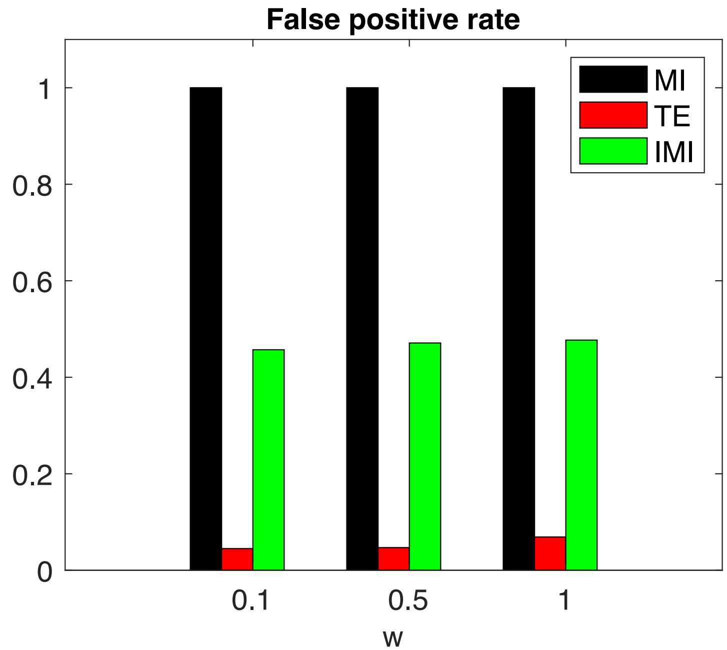
- Clarify the relationship between IMI and the classical metrics of information flow
  - To this aim, we will introduce a minimalistic Boolean model, where IMI, TE, and MI can be exactly computed
- Compare IMI, TE, and MI in terms of their ability to detect leader-follower interactions

IMI, TE, and MI to Infer Directed Interactions

- We explore the feasibility of employing intrinsic mutual information for the inference of the directional coupling between the units
- We contrast its performance with time-delayed mutual information and transfer entropy
- We utilize the time-series of the two units (leader and follower) to estimate all the joint probability distributions in the definition of the information-theoretic metrics
- We focused on the case that showed the richest dependence of IMI on the coupling gain, that is, $\eta_L = 0.95, \eta_F = 0.05$
- The null distributions are estimated by simulating several instances the model with coupling gain $w = 0$

Results of the Inference





- True positive rate (sensitivity) is at the perfection level for all metrics
- Independent of the coupling, we observe an excess of false positives for IMI compared to TE, and totally unacceptable results for MI
- While the poor performance of MI is expected, TE outperforming IMI needs to be further discussed

IMI vs TE: Excess of False Positives

- Counter-intuitively, albeit IMI was introduced to better gauge information flaw compared to TE, it yields a higher number of false positives

- The root-cause can be found in the fact that, for most parameter combinations, $\mathrm{IMI}^{Z\to Y} = \min\{\mathrm{TE}^{Z\to Y}, \mathrm{MI}^{Z\to Y}\}$

Cumulative Distribution Function for IMI, TE and MI

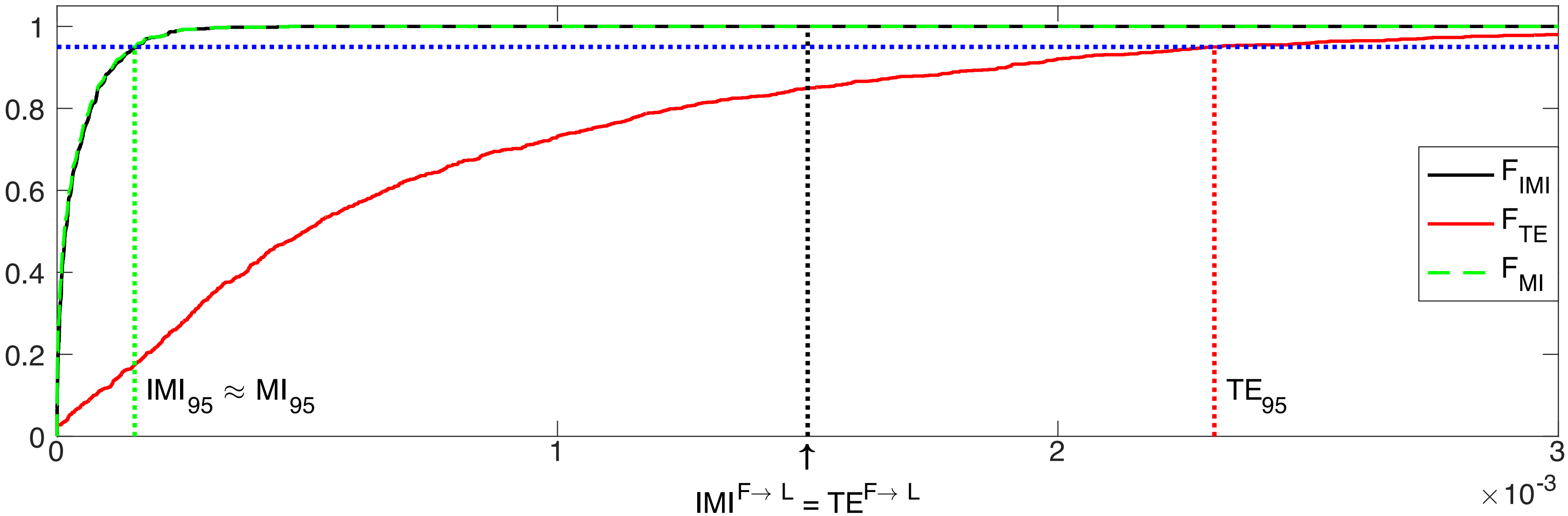
- This assumption implies that $F_{\mathrm{IMI}}(x) \geq \max\{F_{\mathrm{TE}}(x), F_{\mathrm{MI}}(x)\}$, and therefore $\mathrm{IMI}_{95} \leq \min\{\mathrm{TE}_{95}, \mathrm{MI}_{95}\}$

Cut-off values for IMI, TE and MI

- From this observation, we can identify three cases in which the inferences performed by TE and IMI may differ

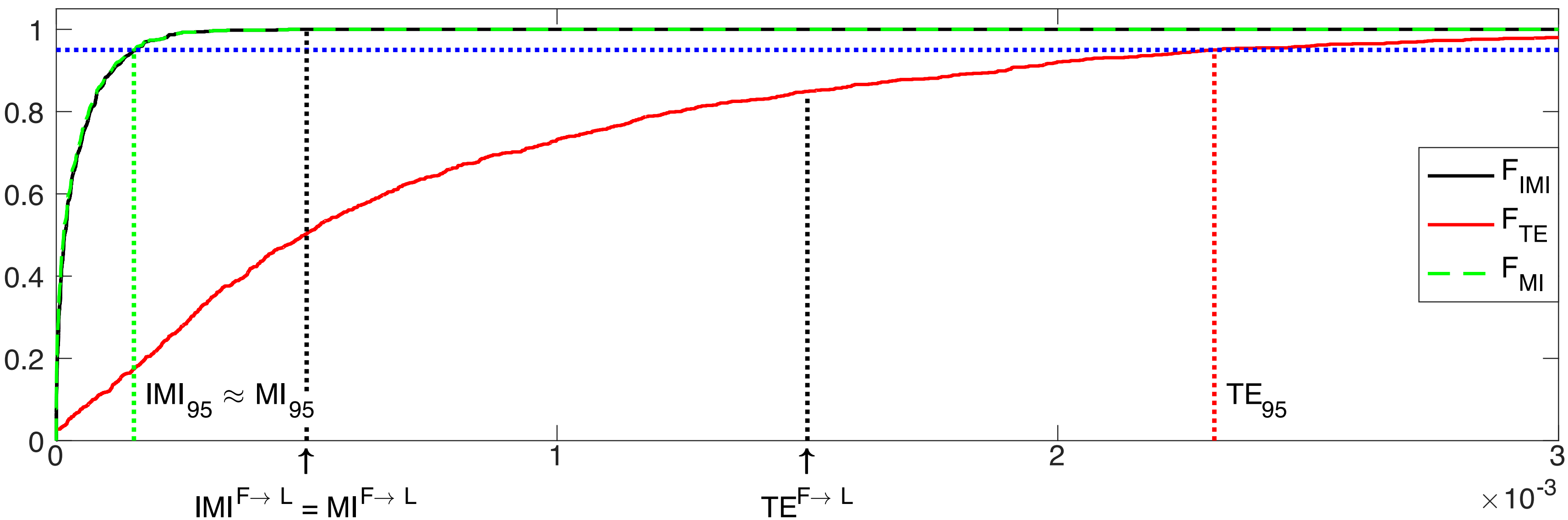
Explaining the Excess of False Positives

- **Case 1**, $\mathrm{IMI}^{F\to L} = \mathrm{TE}^{F\to L}, \mathrm{IMI}_{95} < \mathrm{IMI}^{F\to L} \leq \mathrm{TE}_{95}$: in this case, IMI would reject the null hypothesis and yield a false positive, whereas transfer entropy would correctly infer the absence of a link from the follower to the leader



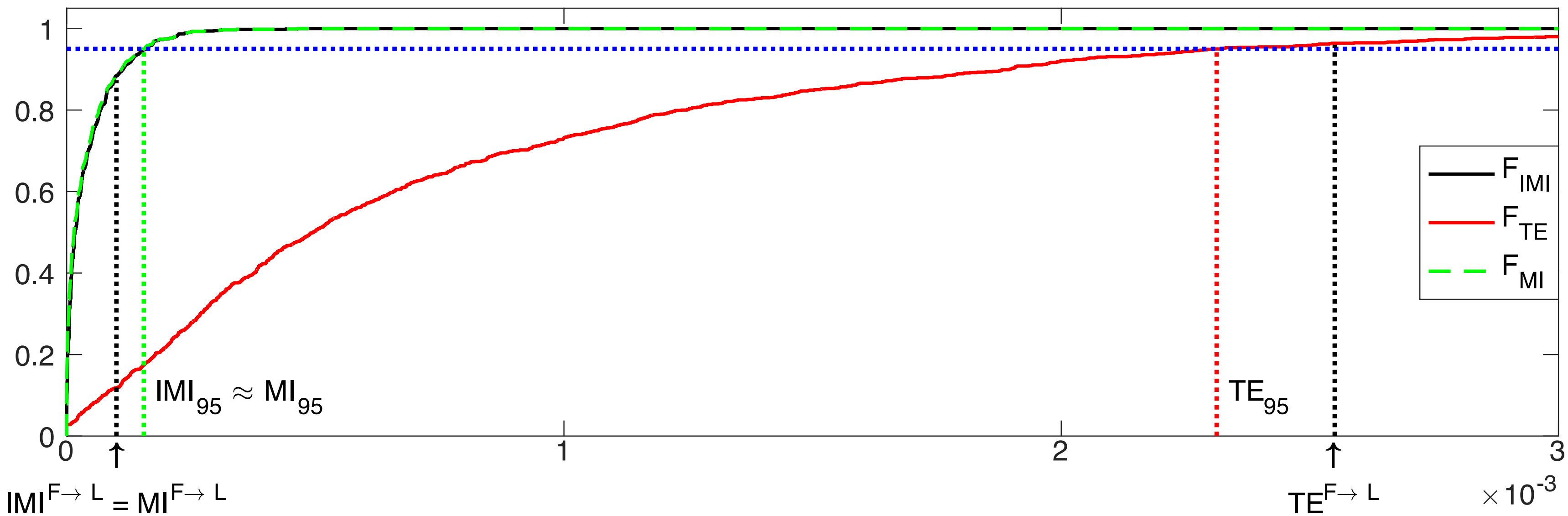
Explaining the Excess of False Positives

- **Case 2**, $\mathrm{IMI}^{F\to L} = \mathrm{MI}^{F\to L}, \mathrm{IMI}^{F\to L} > \mathrm{IMI}_{95}, \mathrm{TE}^{F\to L} \leq \mathrm{TE}_{95}$: again, IMI would reject the null hypothesis and yield a false positive, whereas transfer entropy would correctly infer the absence of a link from the follower to the leader



Explaining the Excess of False Positives

- **Case 3**, $\mathrm{IMI}^{F\to L} = \mathrm{MI}^{F\to L}, \mathrm{IMI}^{F\to L} \leq \mathrm{IMI}_{95}, \mathrm{TE}^{F\to L} > \mathrm{TE}_{95}$: in this case, transfer entropy would reject the null hypothesis and yield a false positive, whereas IMI would correctly infer the absence of a link from the follower to the leader



Explaining the Excess of False Positives

- Case 3 is then the only case in which IMI outperforms transfer entropy in filtering a spurious interaction
- This can occur only if also time-delayed mutual information can filter such a spurious link, which never happens in our study
- Cases 1 and 2 are instead prevalent due to the much fatter tail of the null distribution of TE compared to that of IMI, thus explaining the excess of false positives

Explaining the Excess of False Positives

- Case 3 is then the only case in which IMI outperforms transfer entropy in filtering a spurious interaction
- This can occur only if also time-delayed mutual information can filter such a spurious link, which never happens in our study
- Cases 1 and 2 are instead prevalent due to the much fatter tail of the null distribution of TE compared to that of IMI, thus explaining the excess of false positives
- To support the generality of our findings, we repeated our analysis on a modified leader-follower Vicsek model as in Sattari et al.
- We discretized the phase of the two units in 2, 3, or 4 bins to estimate the distribution probabilities

Inference Results on the Vicsek Model

Numerical analysis on the modified Vicsek model

- Confirm that intrinsic mutual information is well-approximated by the minimum between transfer entropy and time-delayed mutual information (the two quantities being indistinguishable in 99.7% of the 900 cases, at a confidence level of 0.05)
- Cases 1, 2, and 3 are still feasible since the cumulative null distribution of time-delayed mutual information is always above that of transfer entropy

Inference Results on the Vicsek Model

Numerical analysis on the modified Vicsek model

- Confirm that intrinsic mutual information is well-approximated by the minimum between transfer entropy and time-delayed mutual information (the two quantities being indistinguishable in 99.7% of the 900 cases, at a confidence level of 0.05)
- Cases 1, 2, and 3 are still feasible since the cumulative null distribution of time-delayed mutual information is always above that of transfer entropy
- The difference become slimmer as the number of bins increases
- This implies that, as the number of bins increases, the inference performances of TE and IMI become closer, whereas TE outperforms IMI for low number of bins

Conclusions

- Our theoretical and computational results do not point at a practical advantage of intrinsic mutual information versus transfer entropy in the inference of pairwise interactions
- A better appraisal of information flow does not translate into better inference
- While intrinsic mutual information and transfer entropy both display high sensitivity, intrinsic mutual information has considerably lower specificity in the inference of directional interactions
- We warn prudence with the use of intrinsic mutual information as a tool for the discovery of directional interactions

Reference and Contacts

- P. De Lellis, M. Ruiz Marin, M. Porfiri, Inferring directional interactions in collective dynamics: a critique to intrinsic mutual information, Journal of Physics: Complexity, *Focus Issue on Monitoring and Control of Complex Supply Systems*, 4(1), 015001, 2022.

- https://sites.google.com/site/pierodelellis/home

- pietro.delellis@unina.it