

Supplementary Material:

Analysis of pairwise interactions in a maximum likelihood sense to identify leaders in a group

Violet Mwaffo¹, Sachit Butail² and Maurizio Porfiri^{1,*}

¹Department of Mechanical and Aerospace Engineering, New York University Tandon School of Engineering, Brooklyn, NY, USA

²Department of Mechanical Engineering, Northern Illinois University, DeKalb, IL, USA

Correspondence*:
Maurizio Porfiri
mporfiri@nyu.edu

1 SUPPLEMENTARY TABLES AND FIGURES

2 1.1 Figures

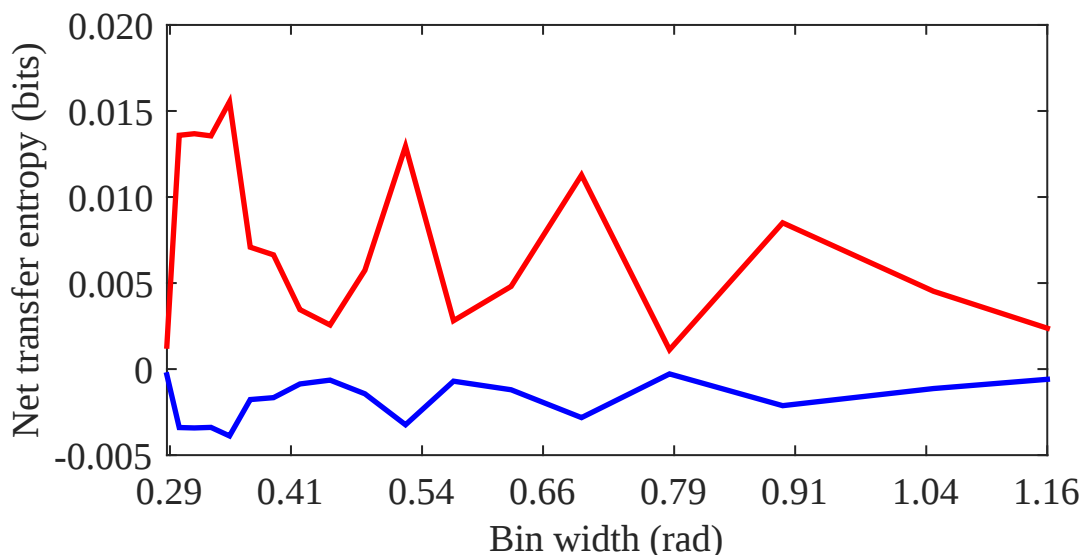


Figure 1. Net transfer entropy computed for group leader and followers as a function of the bin width for the Vicsek model. Markers identify bin widths selected for the analysis in the main document.

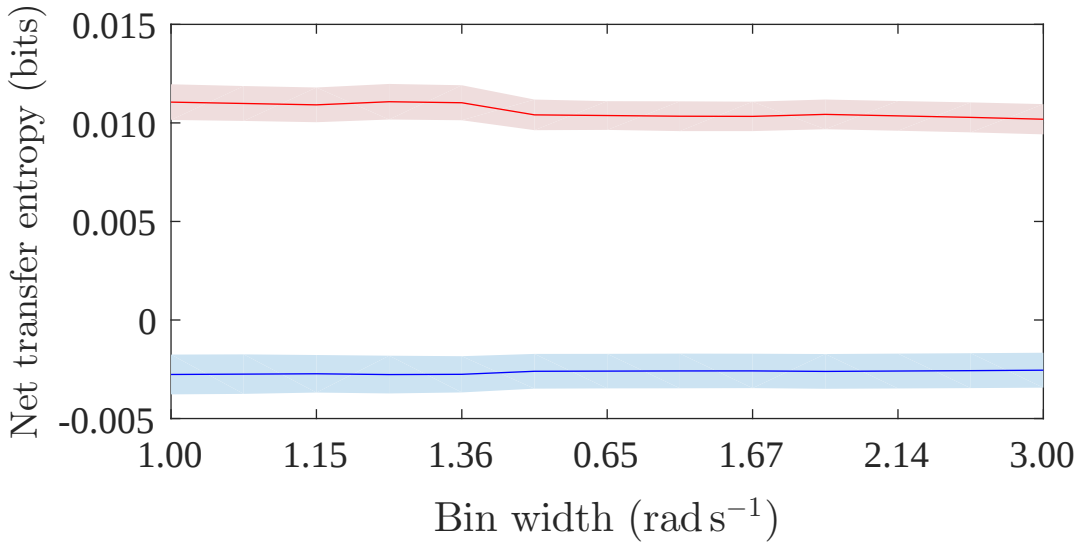


Figure 2. Net transfer entropy computed for group leader and followers as a function of the bin width for the data-driven model of fish social behavior. Errors envelope denote standard errors over 30 simulations. Markers identify bin widths selected for the analysis in the main document

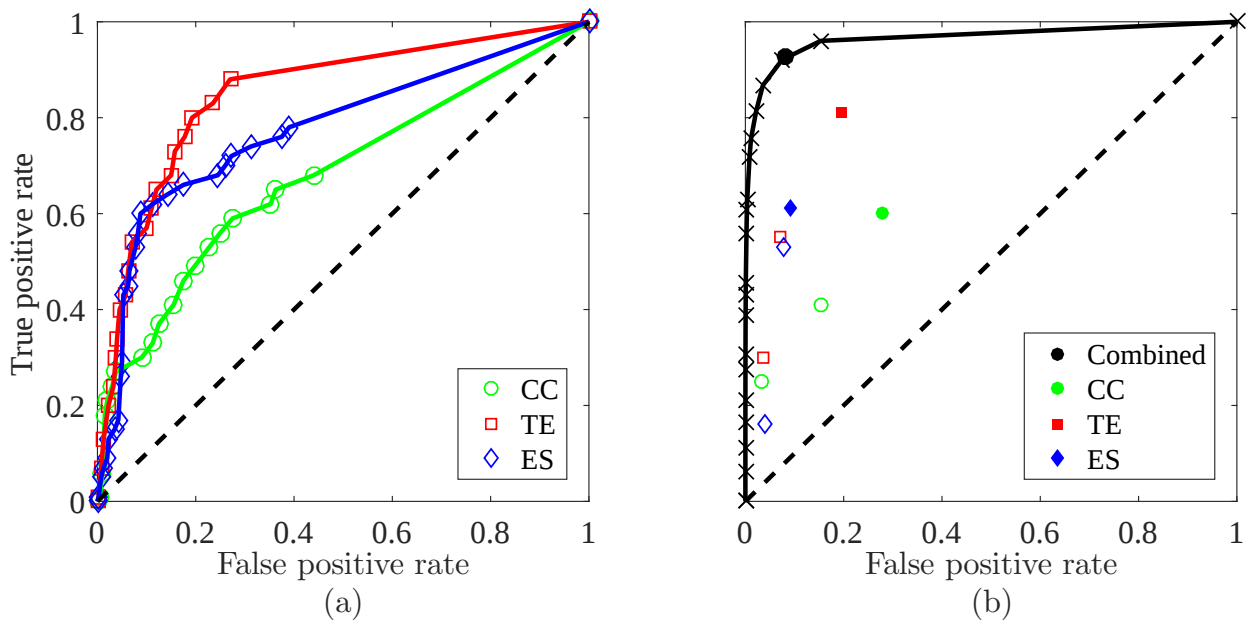


Figure 3. ROC curve for the data-driven model of fish behavior with a single leader for $N = 5$ simulated fish including a single leader by varying the parameter of the fish shoal values within $\pm 10\%$ of their nominal values. The ROC are generated using each of the three classification methods (a) and a combined approach (b), which integrates the ROCs from the three methods and is plotted using 20 out of a possible total of 512 points. The combined ROC is obtained by sampling three points on the ROC for each method. In (a), the AUC for CC, TE, and ES are estimated at 0.68, 0.78, and 0.86 respectively. In (b), the selected cut-off points are chosen such that the first point is just above the 25% quartile, the second is just above the 50% quartile, and the third one is the operating point. The operating point for each individual method is identified as a solid marker and the other two as open markers. The operating point of the combination of the three classifiers is shown as a solid marker and has ROC coordinates (0.08, 0.93). The AUC from the combined method is 0.97.

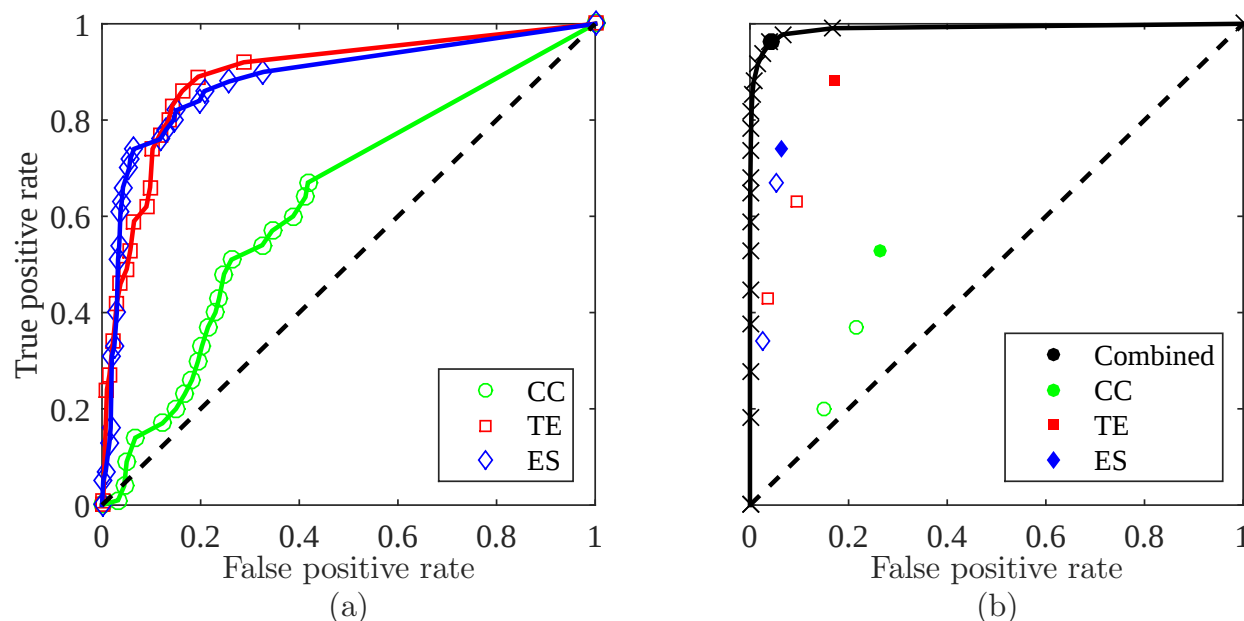


Figure 4. ROC curve for the data-driven model of fish behavior with a single leader for $N = 20$ fish including a single leader generated using each of the three classification methods (a) and a combined approach (b), which integrates the ROCs from the three methods. The combined ROC is obtained by sampling three points on the ROC for each method and is plotted using 20 out of a possible total of 512 points. In (a), the AUC for CC, TE, and ES are estimated at 0.63, 0.89, and 0.90 respectively. In (b), the selected cut-off points are chosen such that the first point is just above the 25% quartile, the second is just above the 50% quartile, and the third one is the operating point. The operating point for each individual method is identified as a solid marker and the other two as open markers. The operating point of the combination of the three classifiers is shown as a solid marker and has ROC coordinates (0.04, 0.96). The AUC from the combined method is 0.99.

3 1.2 Table

Trial	$\neg CC \wedge TE \vee ES$
1	×
2	×
3	×
4	×
5	×
6	×
7	×
8	×
9	×
10	×

Table 1. Leader detection in caffeine-treated zebrafish by implementing the Boolean rule $\neg CC \wedge TE \vee ES$ on each of the ten trials. The symbol $\times$ indicates that the caffeine-treated fish is identified as a leader.