Data in Collective Motion

The Big Statistical Issues

What this talk is not…

Ranting!

Bayesian v Frequentist

Reproducibility, p-hacking etc

PLoS Computational Biology 2005

Open access, freely available online

Essay

Why Most Published Research Findings Are False

John P. A. Ioannidis

Summary

There is increasing concern that most current published research findings are false. The probability that a research claim is true may depend on study power and bias, the number of other studies on the same question, and, importantly, the ratio of true to no relationships among the relationships probed in each scientific field. In this framework, a research finding is less likely to be true when the studies conducted in a field are smaller; when effect sizes are smaller; when there is a greater number and lesser preselection of tested relationships; where there is greater flexibility in designs, definitions, outcomes, and analytical modes; when there is greater financial and other interest and prejudice; and when more teams are involved in a scientific field in chase of statistical significance. Simulations show that for most study designs and settings, it is more likely for a research claim to be false than true.

factors that influence this problem and some corollaries thereof.

Modeling the Framework for False Positive Findings

Several methodologists have pointed out [9–11] that the high rate of nonreplication (lack of confirmation) of research discoveries is a consequence of the convenient, yet ill-founded strategy of claiming conclusive research findings solely on the basis of a single study assessed by formal statistical significance, typically for a *p*-value less than 0.05. Research is not most appropriately represented and summarized by *p*-values, but, unfortunately, there is a widespread notion that medical research articles

It can be proven that most claimed research findings are false.

should be interpreted based only on

is characteristic of the field and can vary a lot depending on whether the field targets highly likely relationships or searches for only one or a few true relationships among thousands and millions of hypotheses that may be postulated. Let us also consider, for computational simplicity, circumscribed fields where either there is only one true relationship (among many that can be hypothesized) or the power is similar to find any of the several existing true relationships. The pre-study probability of a relationship being true is $R/(R+1)$. The probability of a study finding a true relationship reflects the power $1 - \beta$ (one minus the Type II error rate). The probability of claiming a relationship when none truly exists reflects the Type I error rate, α. Assuming that c relationships are being probed in the field, the expected values of the 2×2 table are given in Table 1. After a research finding has been claimed based on achieving formal statistical significance, PLoS Computational Biology 2005

Open access, freely available online

Essay

Why Most Published Research Findings Are False

John P. A. Ioannidis

Summary

There is increasing concern that most current published research findings are false. The probability that a research claim is true may depend on study power and bias, the number of other studies on the same question, and, importantly, the ratio of true to no relationships among the relationships probed in each scientific field. In this framework, a research finding is less likely to be true when the studies conducted in a field are smaller; when effect sizes are smaller; when there is a greater number and lesser preselection of tested relationships; where there is greater flexibility in designs, definitions, outcomes, and analytical modes; when there is greater financial and other interest and prejudice; and when more teams are involved in a scientific field in chase of statistical significance. Simulations show that for most study designs and settings, it is more likely for a research claim to be false than true. Moreover, for many current scientific

factors that influence this problem and some corollaries thereof.

Modeling the Framework for False Positive Findings

Several methodologists have pointed out [9–11] that the high rate of nonreplication (lack of confirmation) of research discoveries is a consequence of the convenient, yet ill-founded strategy of claiming conclusive research findings solely on the basis of a single study assessed by formal statistical significance, typically for a p-value less than 0.05. Research is not most appropriately represented and summarized by p-values, but, unfortunately, there is a widespread notion that medical research articles

It can be proven that most claimed research findings are false.

should be interpreted based only on

characteristic of the field and can vary a lot depending on whether the field targets highly likely relationships or searches for only one or a few true relationships among thousands and millions of hypotheses that may be postulated. Let us also consider, for computational simplicity, circumscribed fields where either there is only one true relationship (among many that can be hypothesized) or the power is similar to find any of the several existing true relationships. The pre-study probability of a relationship being true is $R/(R+1)$. The probability of a study finding a true relationship reflects the power $1 - \beta$ (one minus the Type II error rate). The probability of claiming a relationship when none truly exists reflects the Type I error rate, α. Assuming that c relationships are being probed in the field, the expected values of the 2 × 2 table are given in Table 1. After a research finding has been claimed based on achieving formal statistical significance,

What this talk is not…

Ranting!

Bayesian v Frequentist

Reproducibility, p-hacking etc

What this talk is…

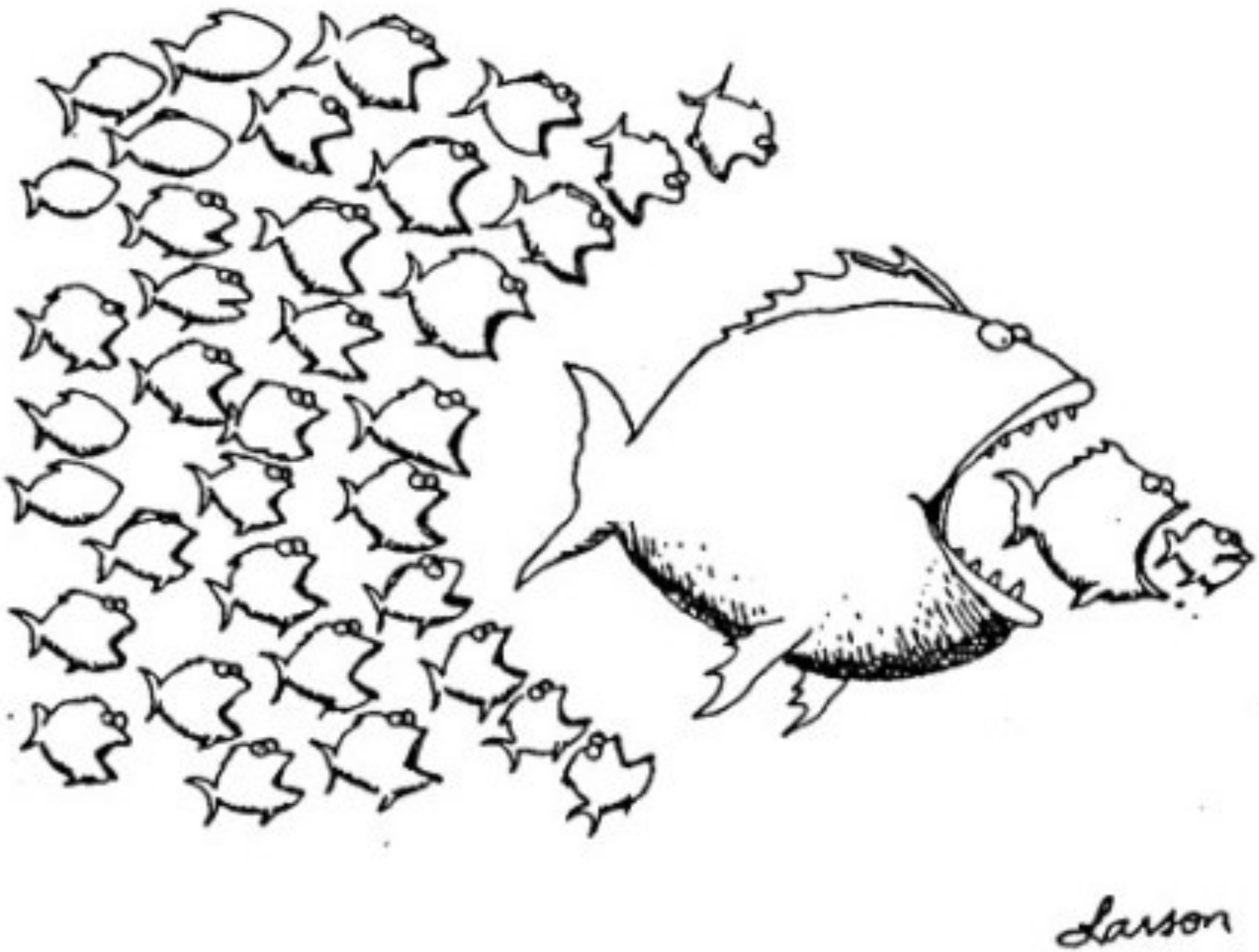
Statistical inference issues that I repeatedly see in collective motion, and some (imperfect) solutions.

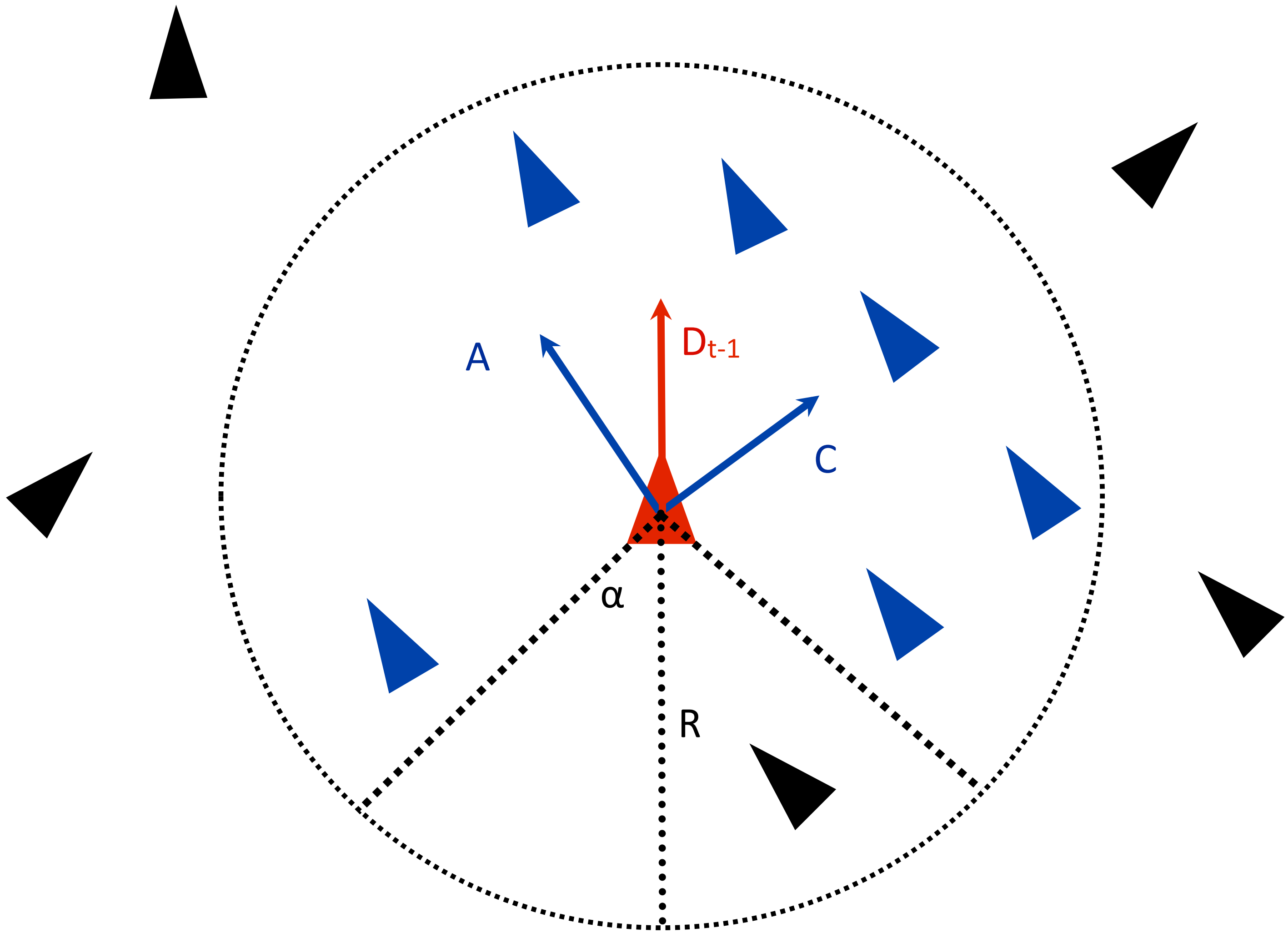
Data in Collective Motion

Some

~~The~~ Big Statistical Issues

Problem 1: Emergence

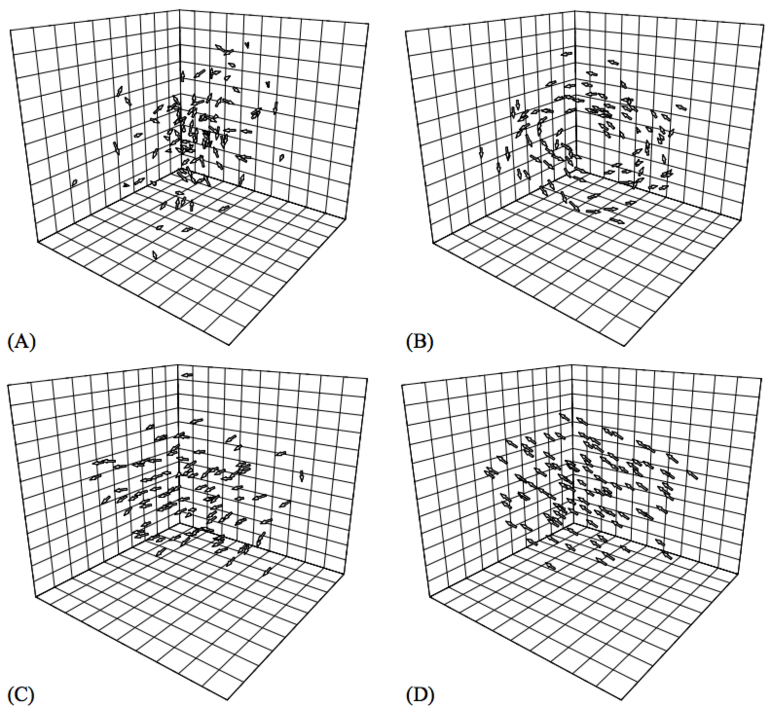




$$D_t = D_{t-1} + \alpha A + c C + \varepsilon$$

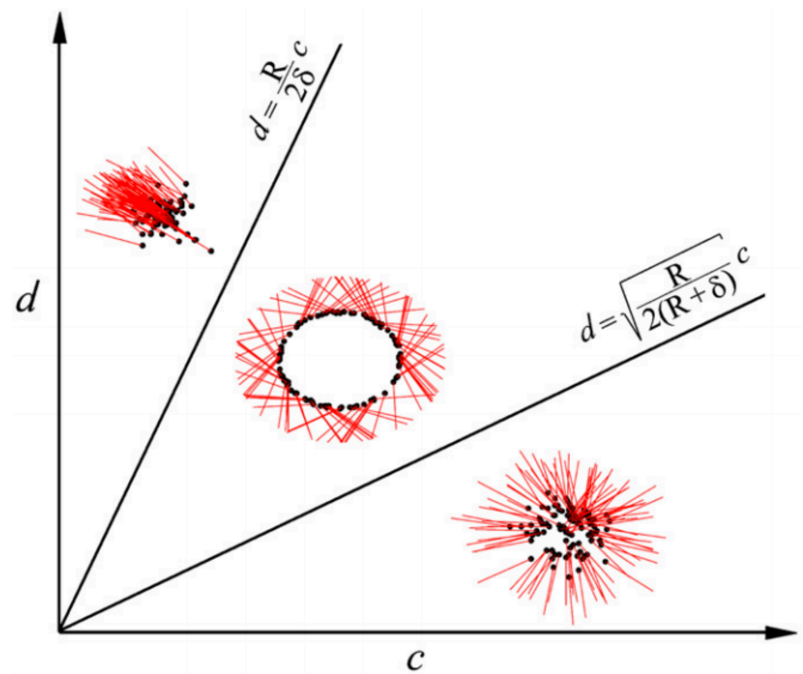
Group level pattern matching

Alignment



Couzin *et al.*: *J. Theor. Biol.* 2002

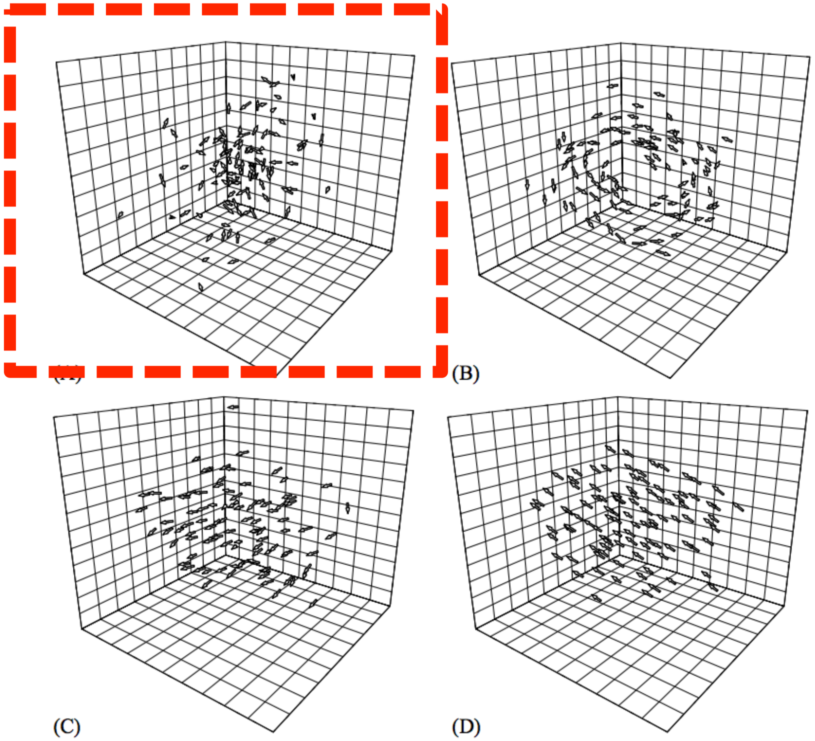
Attraction



Strömbom: *J. Theor. Biol.* 2011

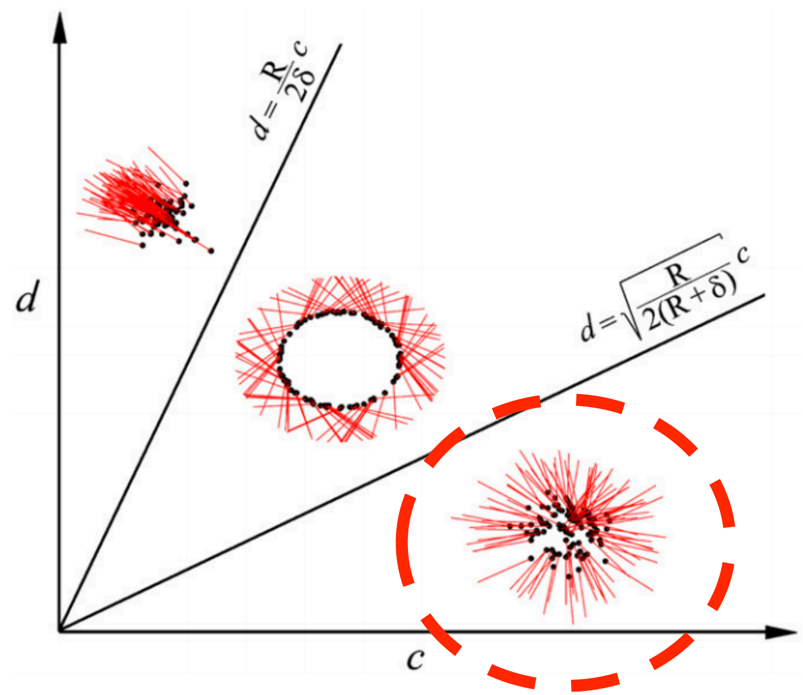
Group level pattern matching

Alignment



Couzin *et al.*: *J. Theor. Biol.* 2002

Attraction

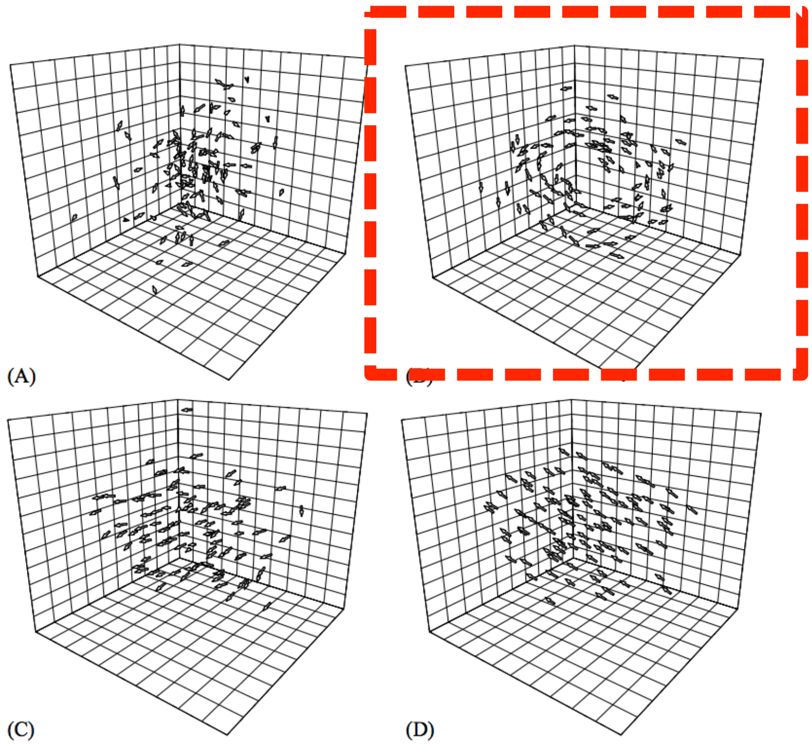


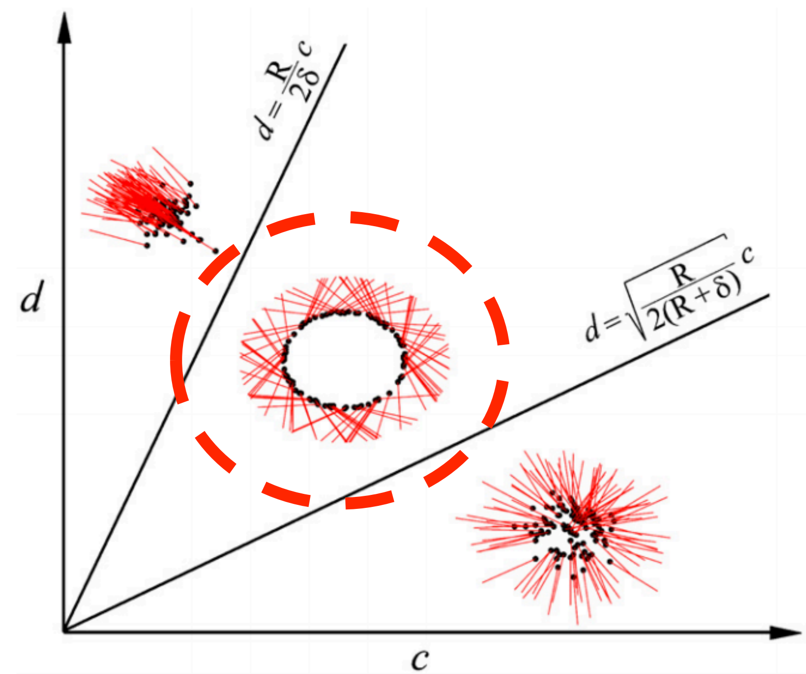
Strömbom: *J. Theor. Biol.* 2011

Group level pattern matching

Alignment

Attraction



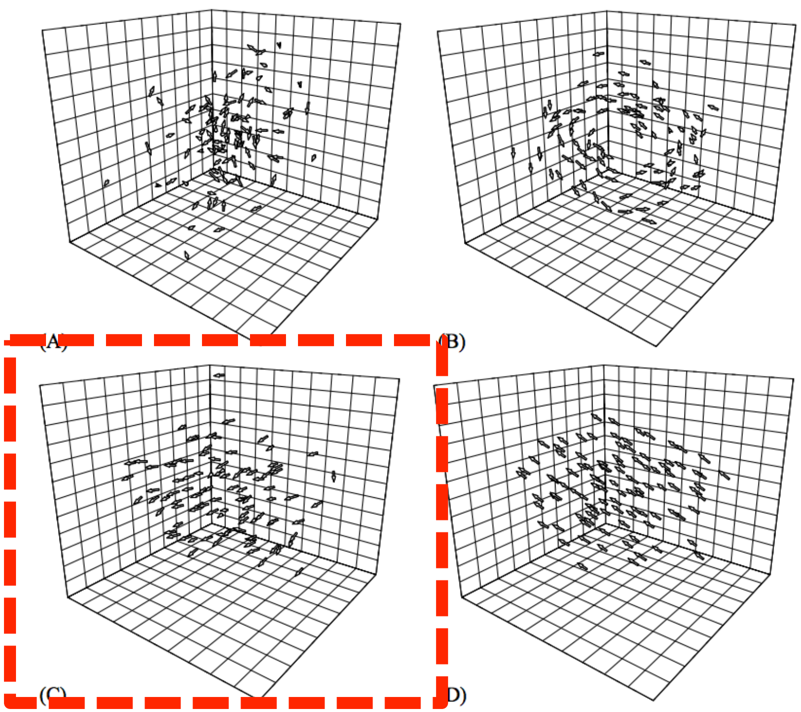


Couzin *et al.*: *J. Theor. Biol*. 2002

Strömbom*: J. Theor. Biol*. 2011

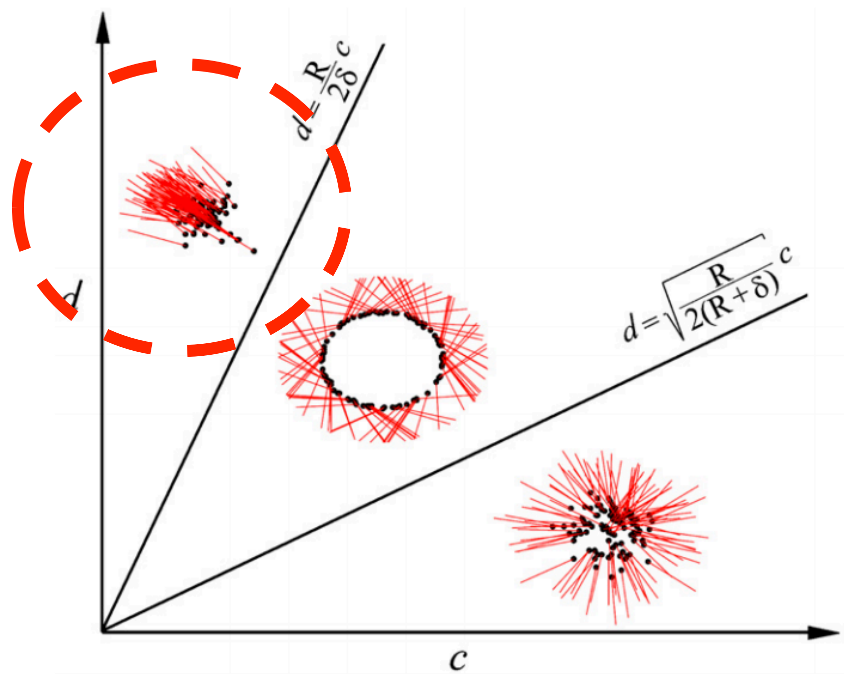
Group level pattern matching

Alignment

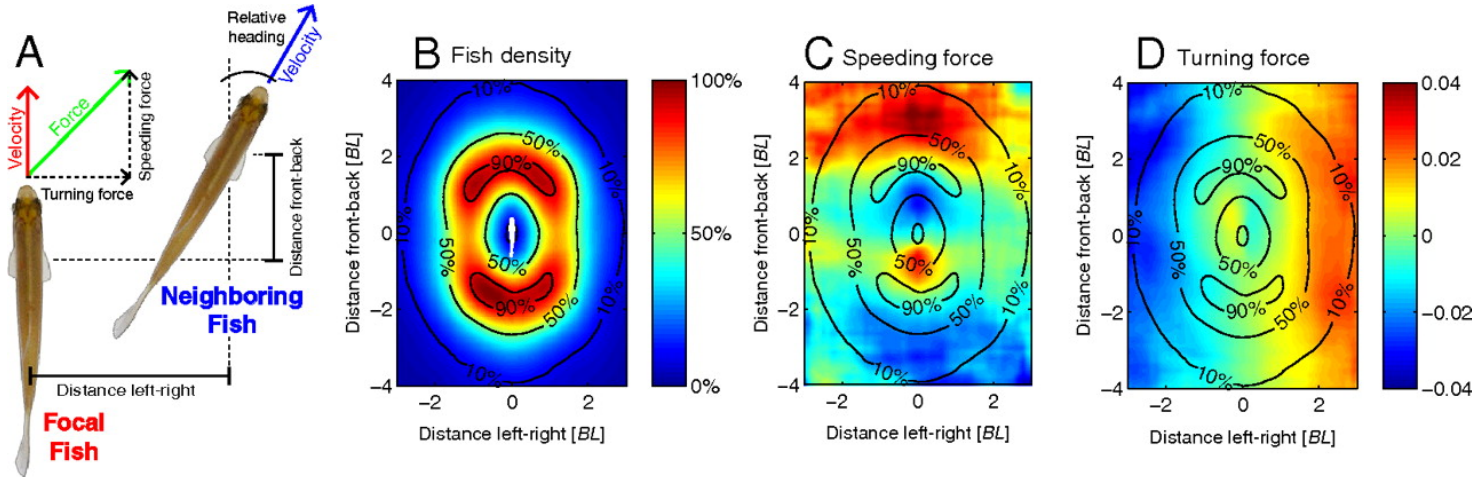


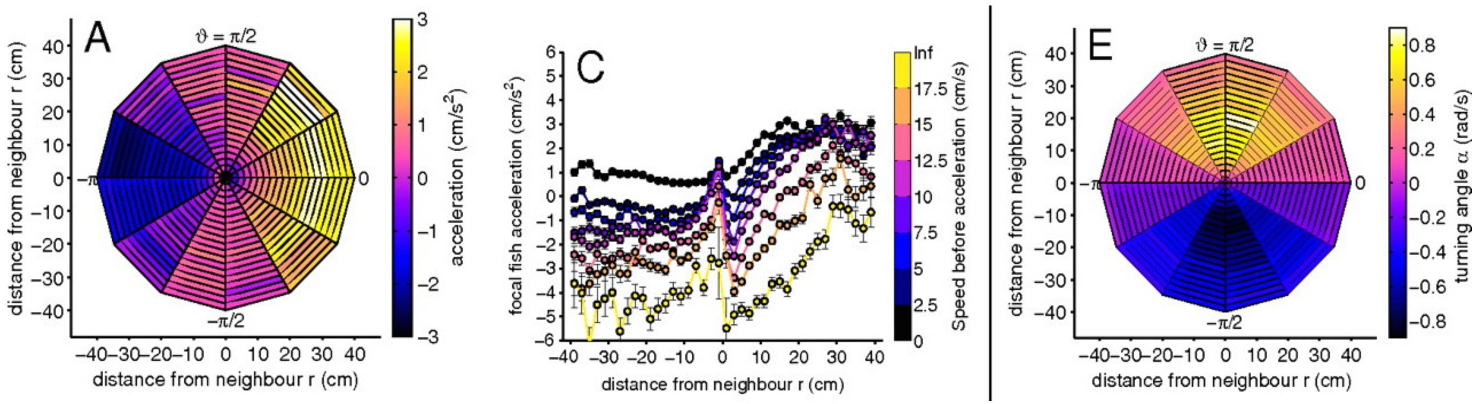
Couzin *et al.*: *J. Theor. Biol.* 2002

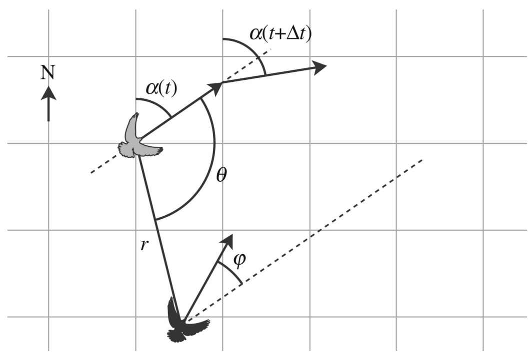
Attraction

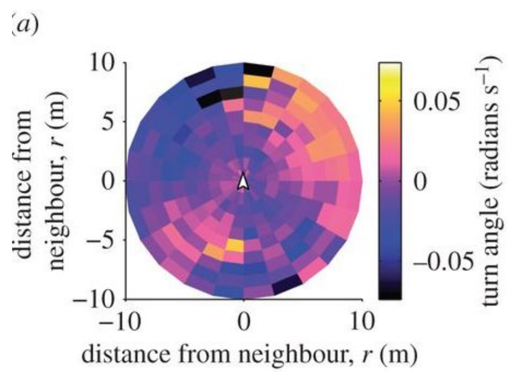


Strömbom: *J. Theor. Biol.* 2011



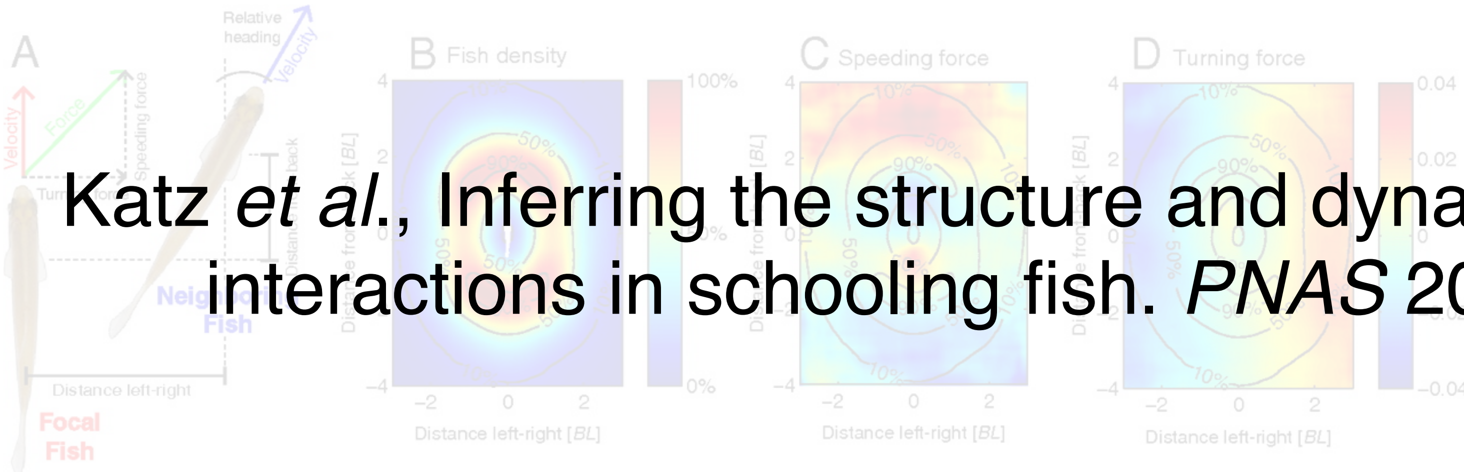


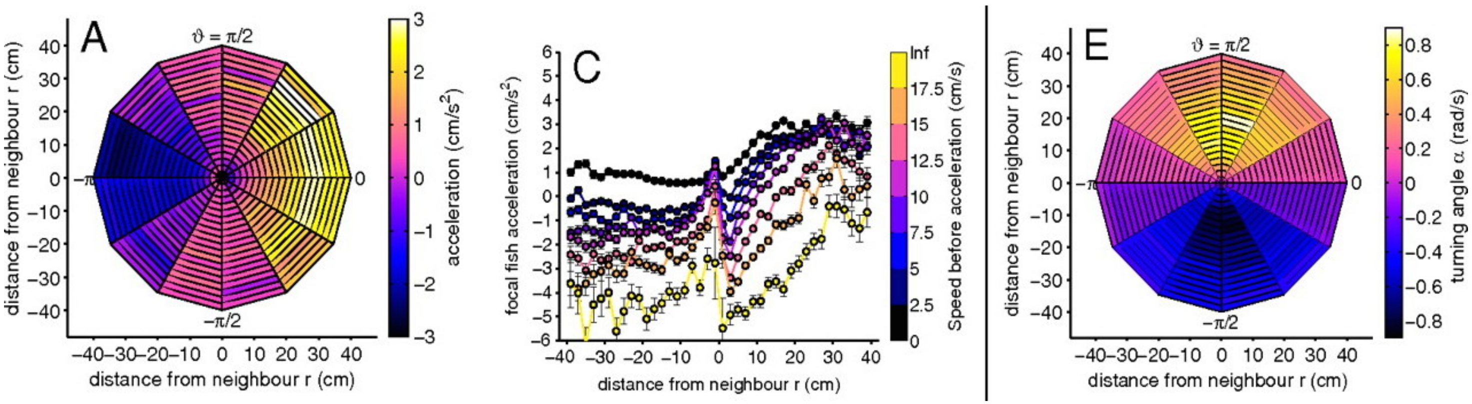


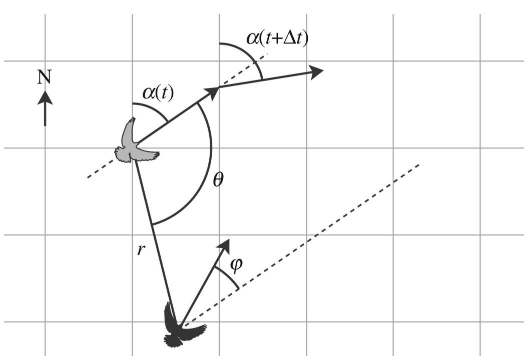


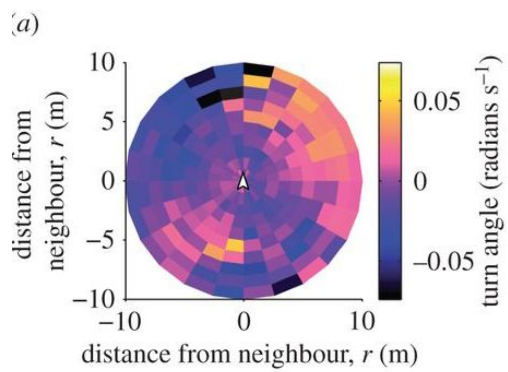
2010--: Many ‘interaction rules’ papers

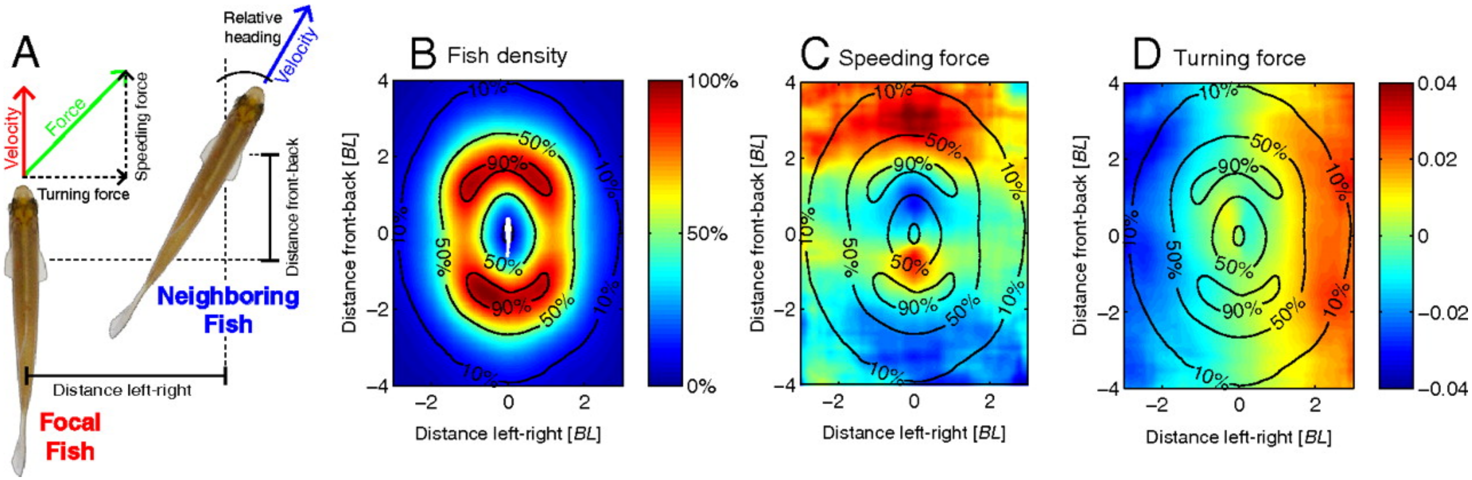
Katz *et al*., Inferring the structure and dynamics of interactions in schooling fish. *PNAS* 2011

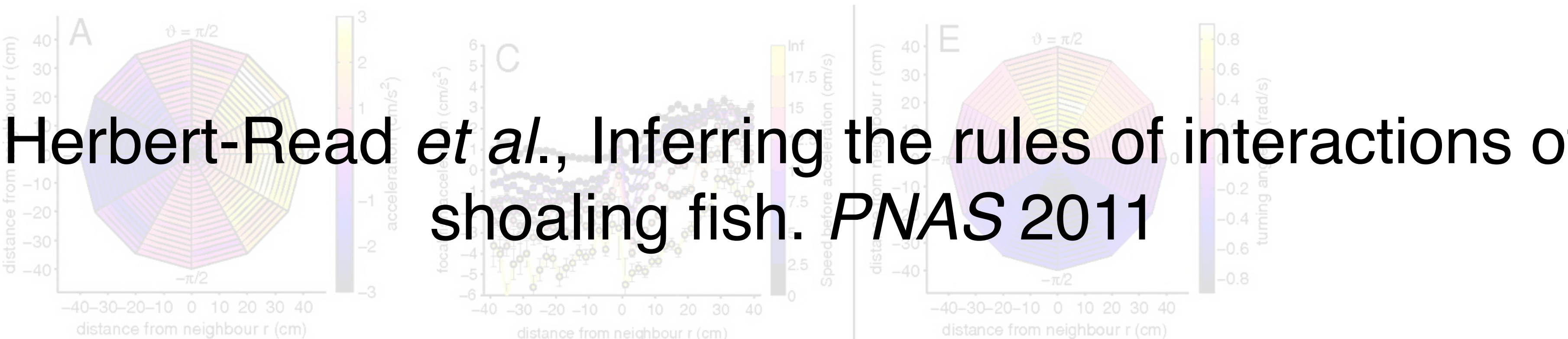




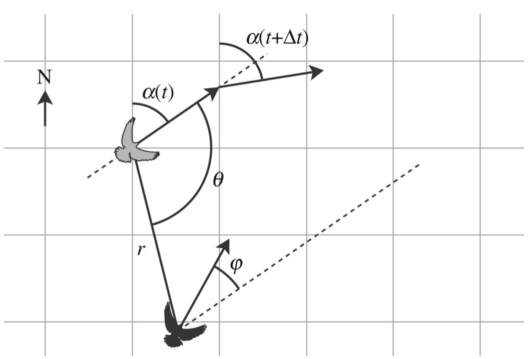


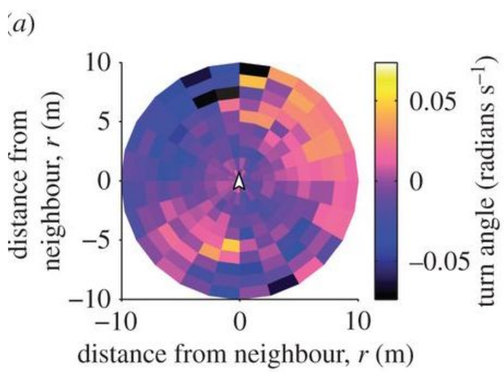


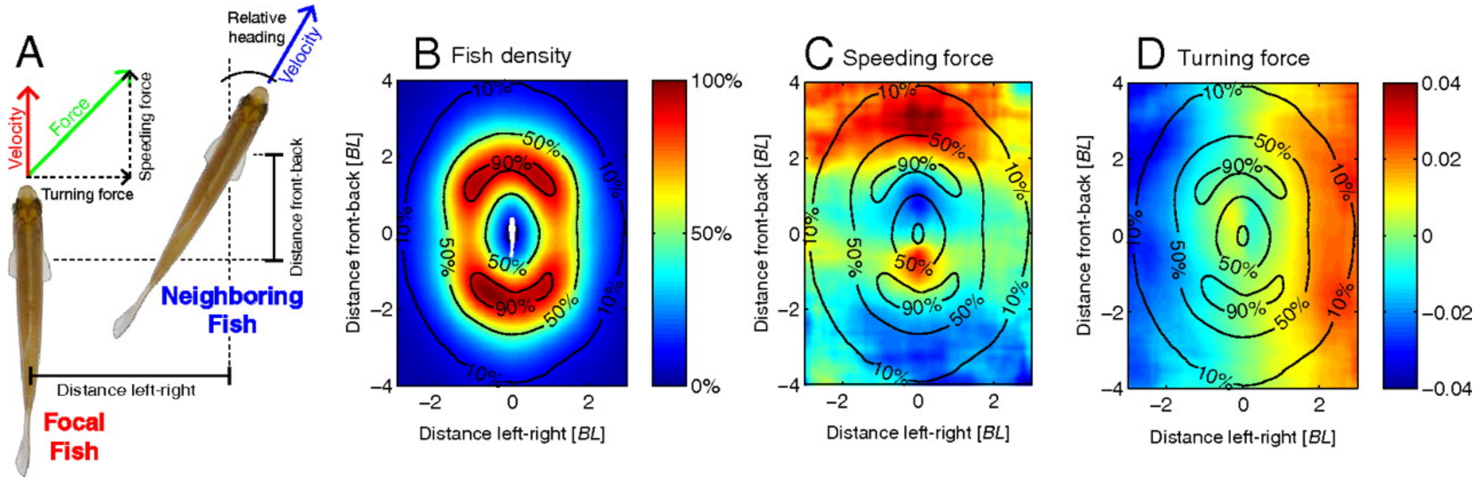


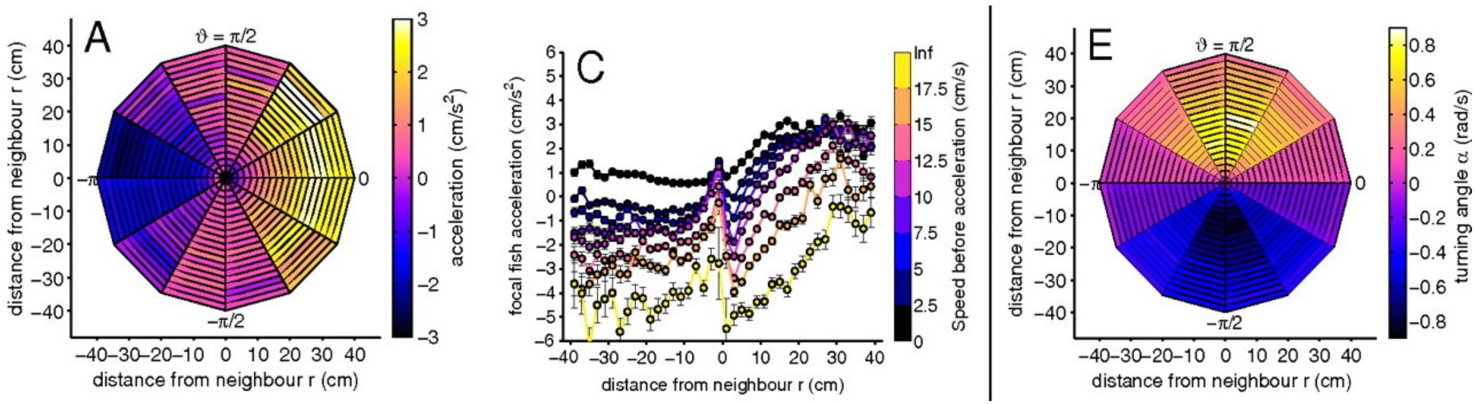


Herbert-Read *et al*., Inferring the rules of interactions of shoaling fish. *PNAS* 2011





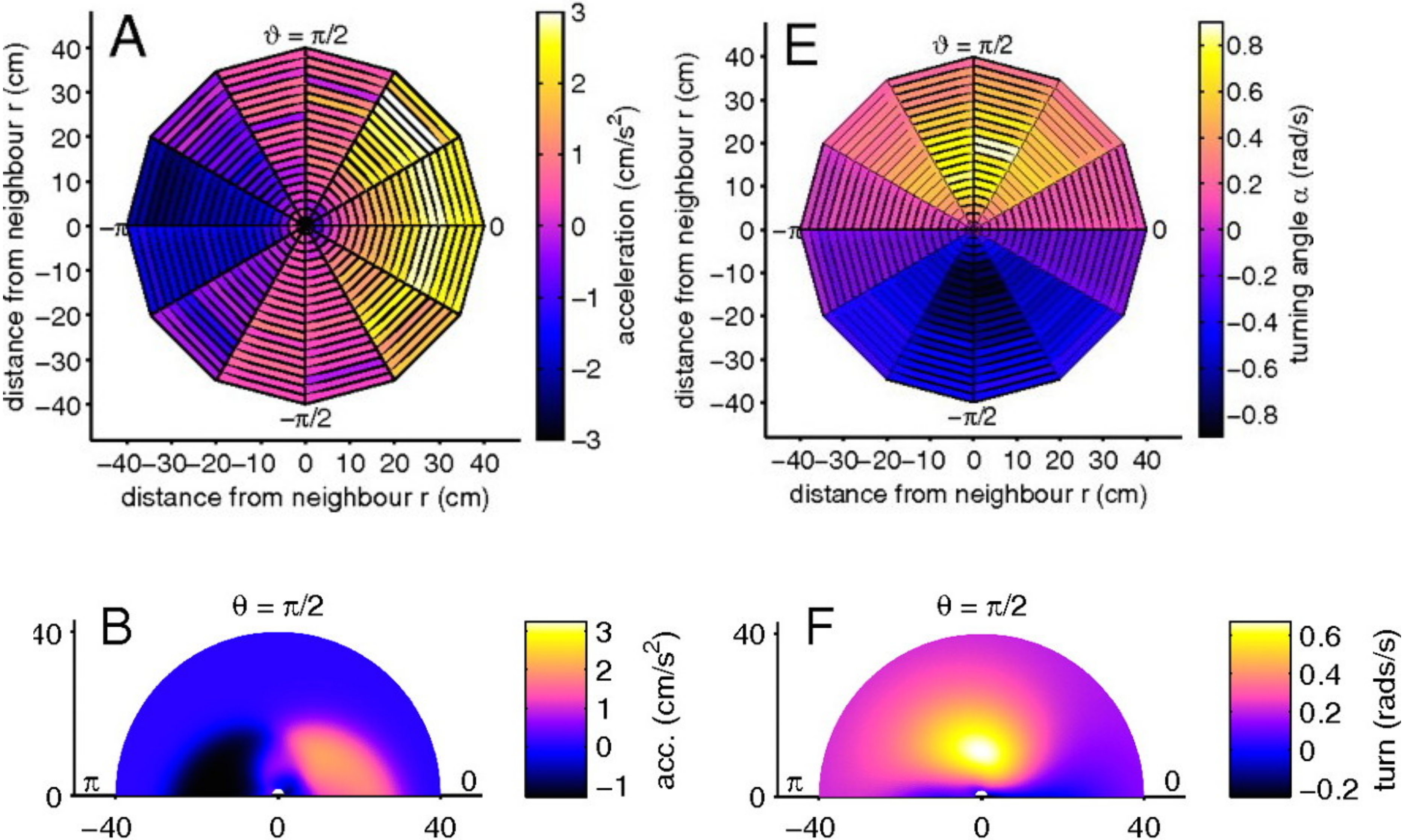




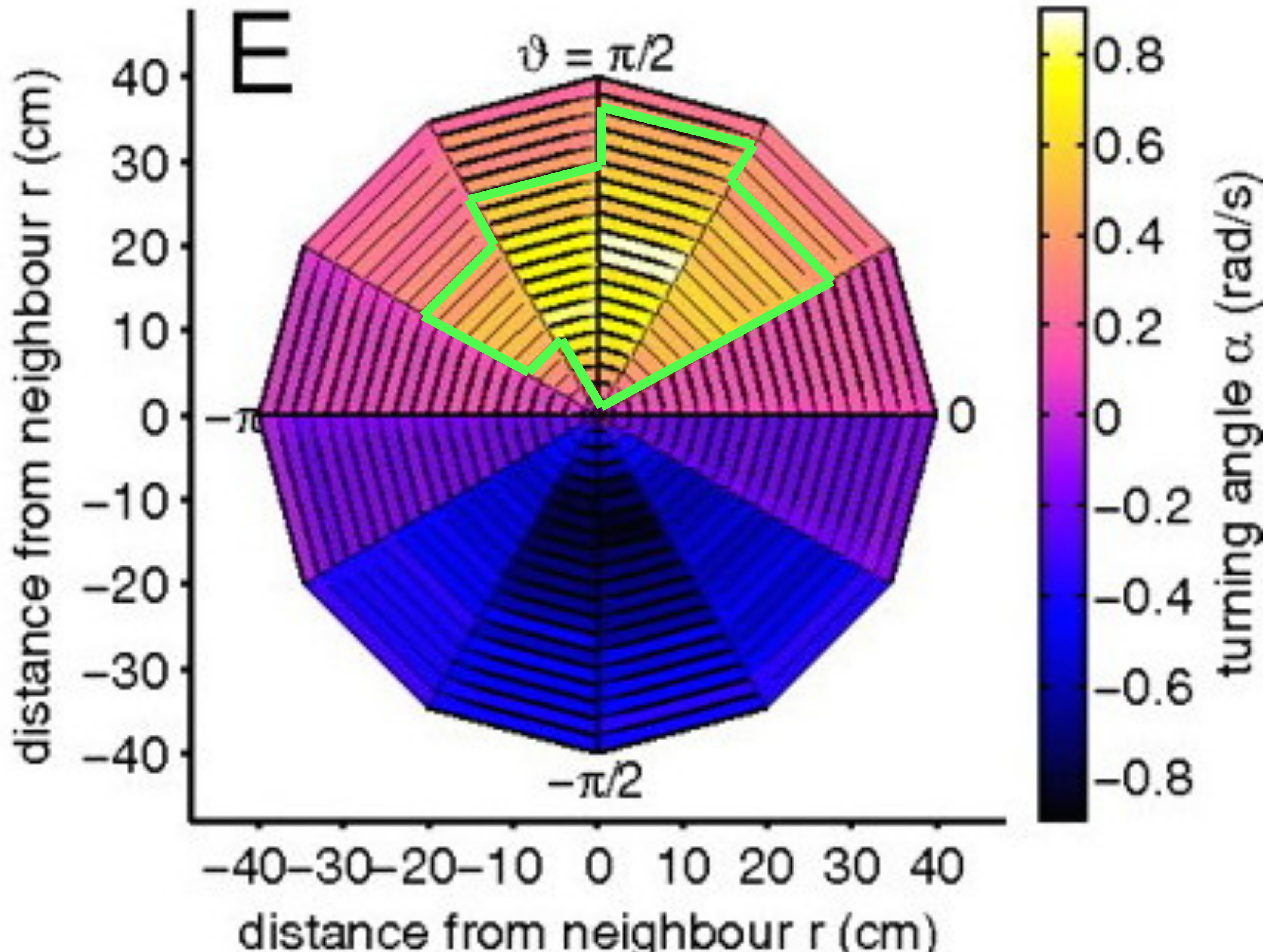
Pettit *et al.* Interaction rules underlying group decisions in homing pigeons. *J R Soc Interface* 2013

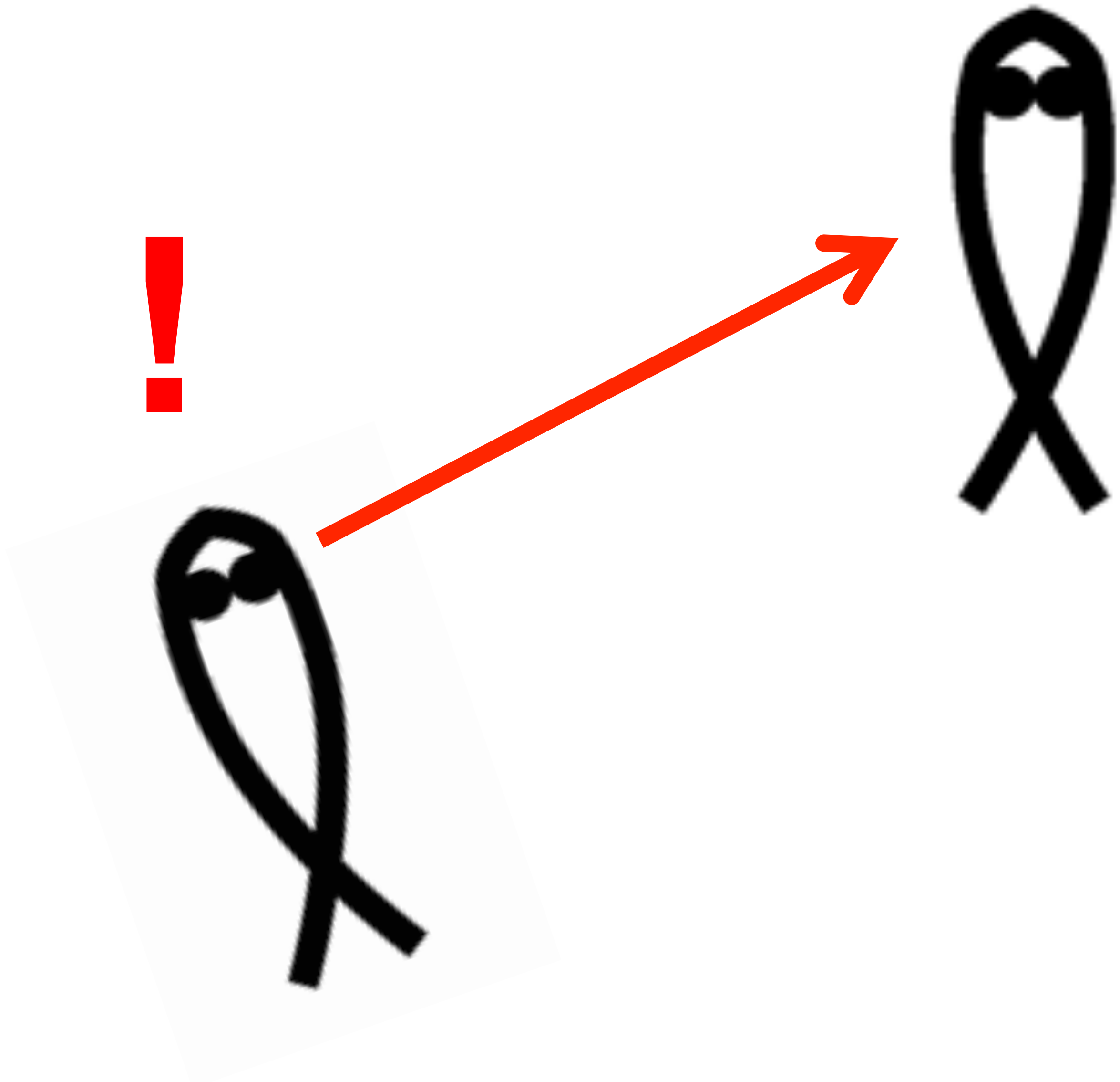
Problem 2: multiple tests

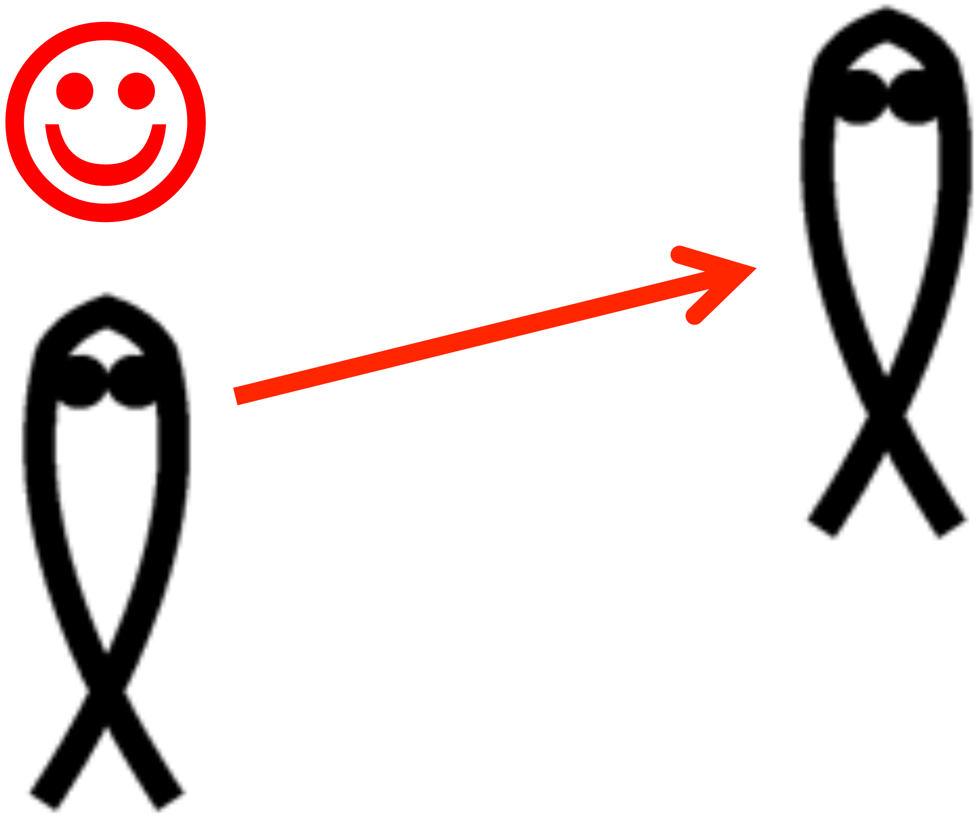
Learn one function, not many



Problem 3: time series


E
ϑ = π/2
−π
0
−π/2
distance from neighbour r (cm)
distance from neighbour r (cm)
40
30
20
10
0
−10
−20
−30
−40
−40 −30 −20 −10 0 10 20 30 40
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
turning angle α (rad/s)





One decision

Four data points

Model decaying acceleration

OPEN ACCESS Freely available online

PLOS | COMPUTATIONAL BIOLOGY

Deciphering Interactions in Moving Animal Groups

Jacques Gautrais[1,2*], Francesco Ginelli[3,4,5], Richard Fournier[6,7], Stéphane Blanco[6,7], Marc Soria[8], Hugues Chaté[3], Guy Theraulaz[1,2]

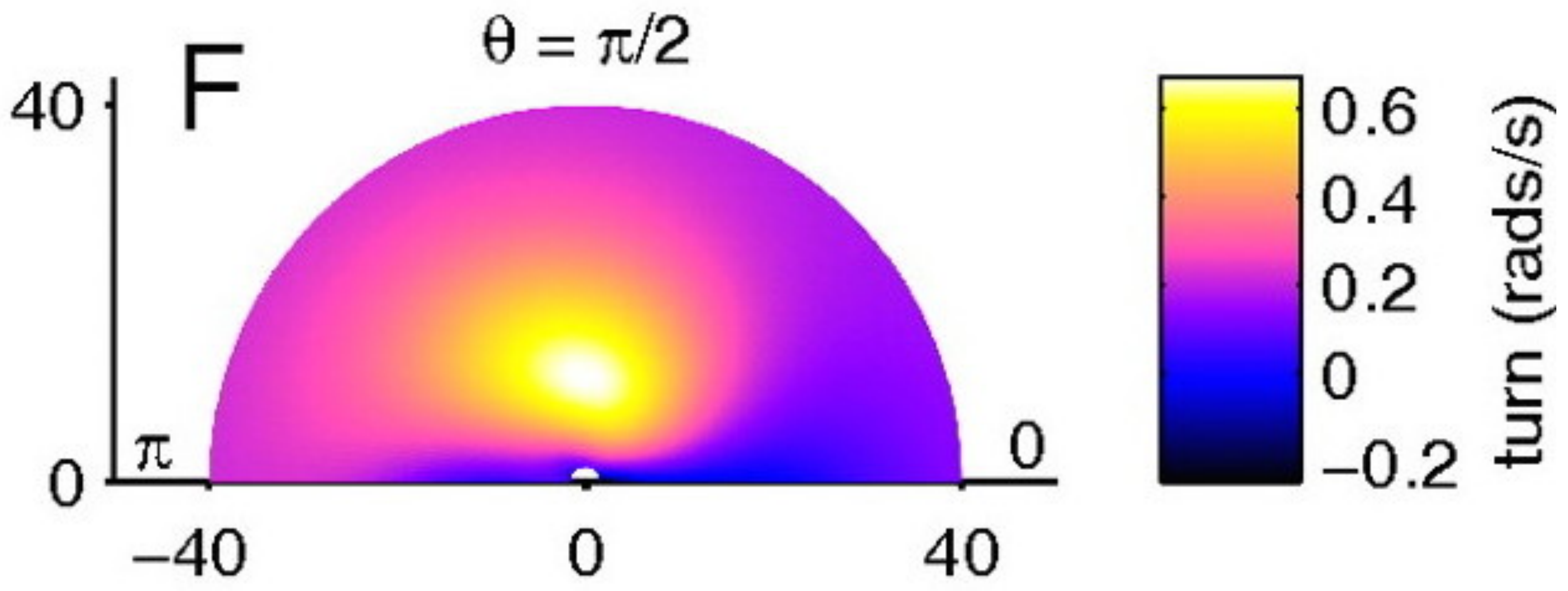
Control for autocorrelation

PNAS

Inferring the rules of interaction of shoaling fish

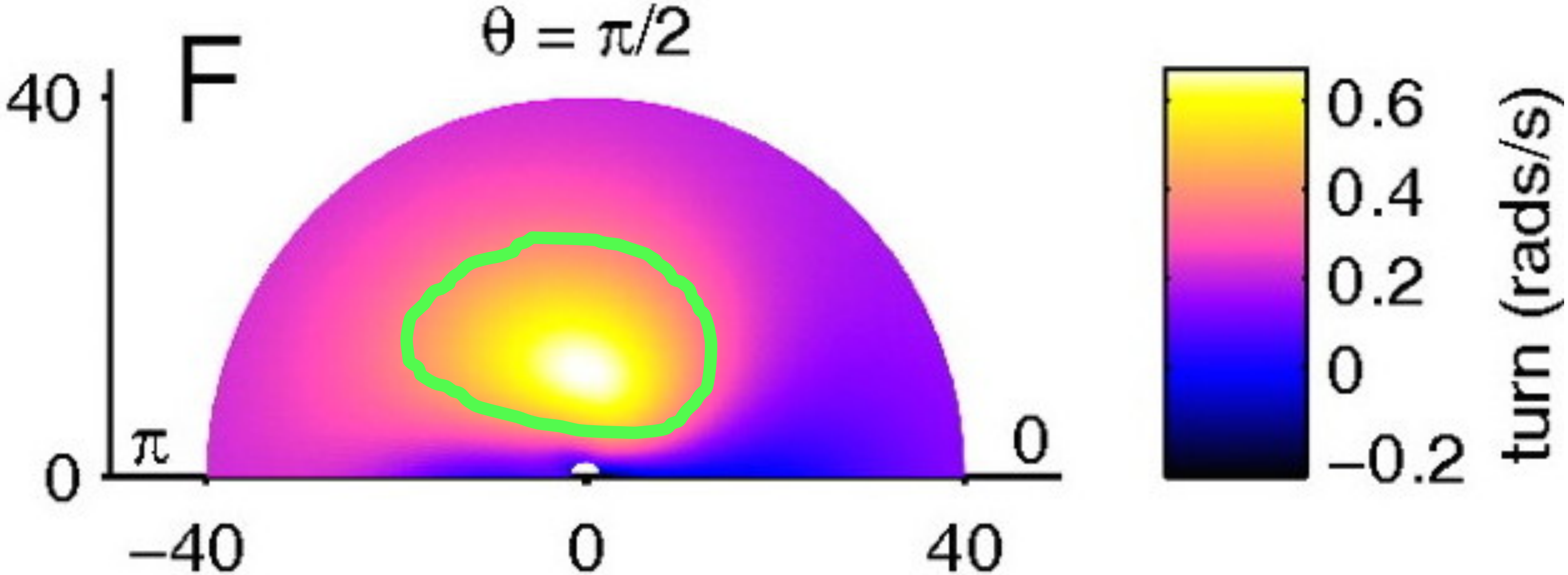
James E. Herbert-Read[a,1,2], Andrea Perna[b,1], Richard P. Mann[b], Timothy M. Schaerf[a], David J. T. Sumpter[b], and Ashley J. W. Ward[a,3]

- Control for autocorrelation
- Fit a single function, with spatial structure
- Use a neural network to define a flexible function space
- Separate different stimuli **(bonus problem)**

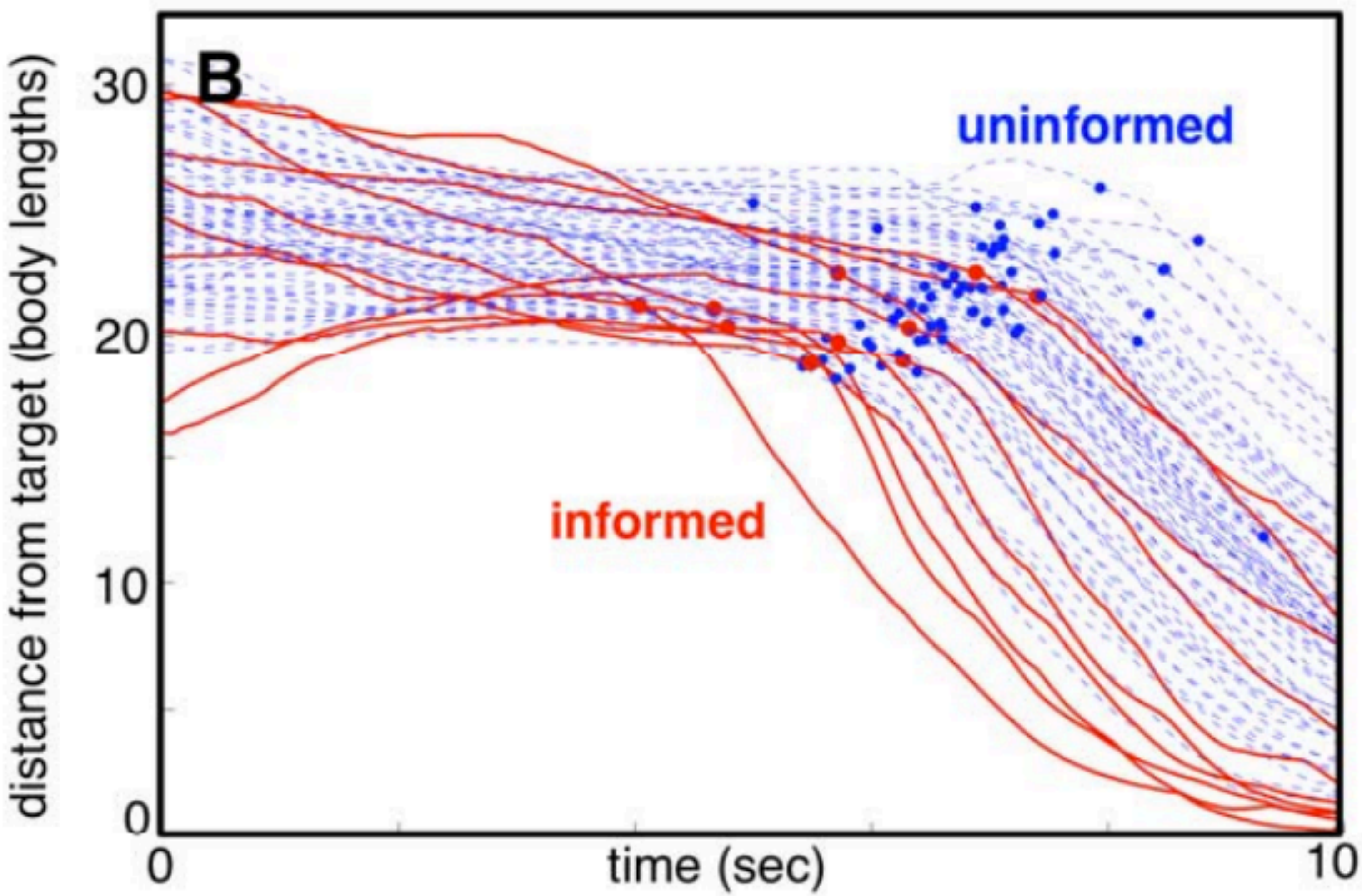


- Control for autocorrelation
- Fit a single function, with spatial structure
- Use a neural network to define a flexible function space
- Separate different stimuli

Cleaner, more powerful inference

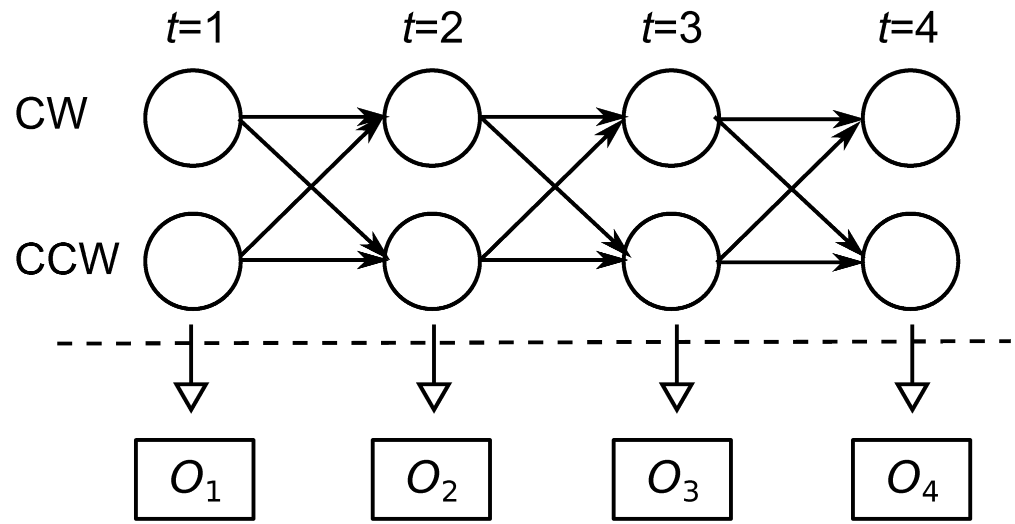


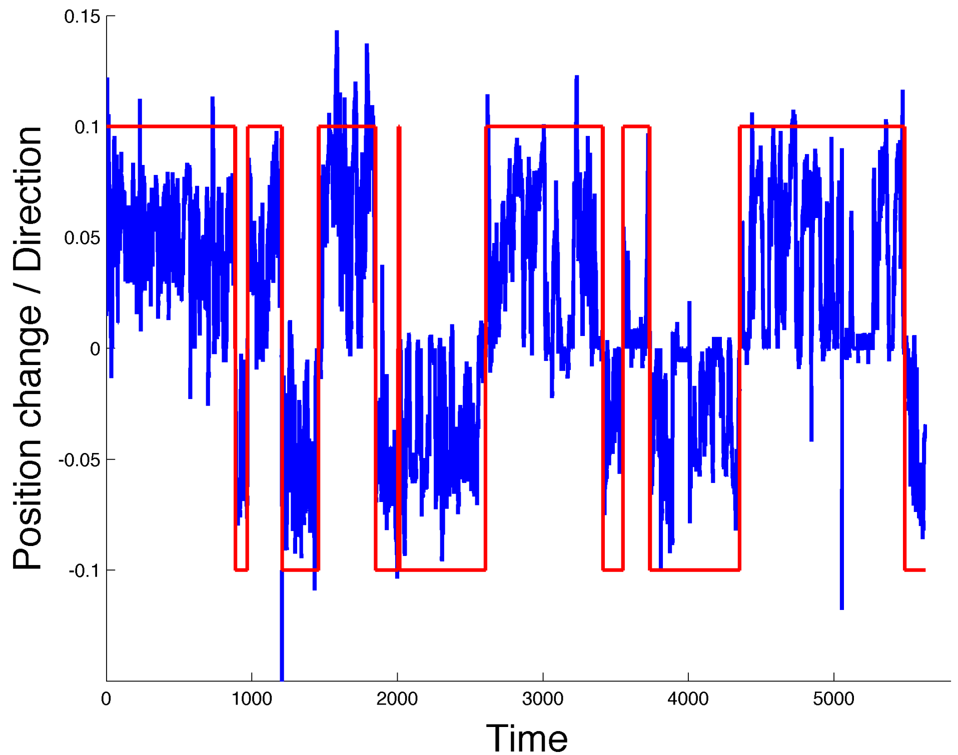
Focus on change points



Strandburg-Peshkin *et al. Current Biology* 2013

Embed the time correlations with a latent space





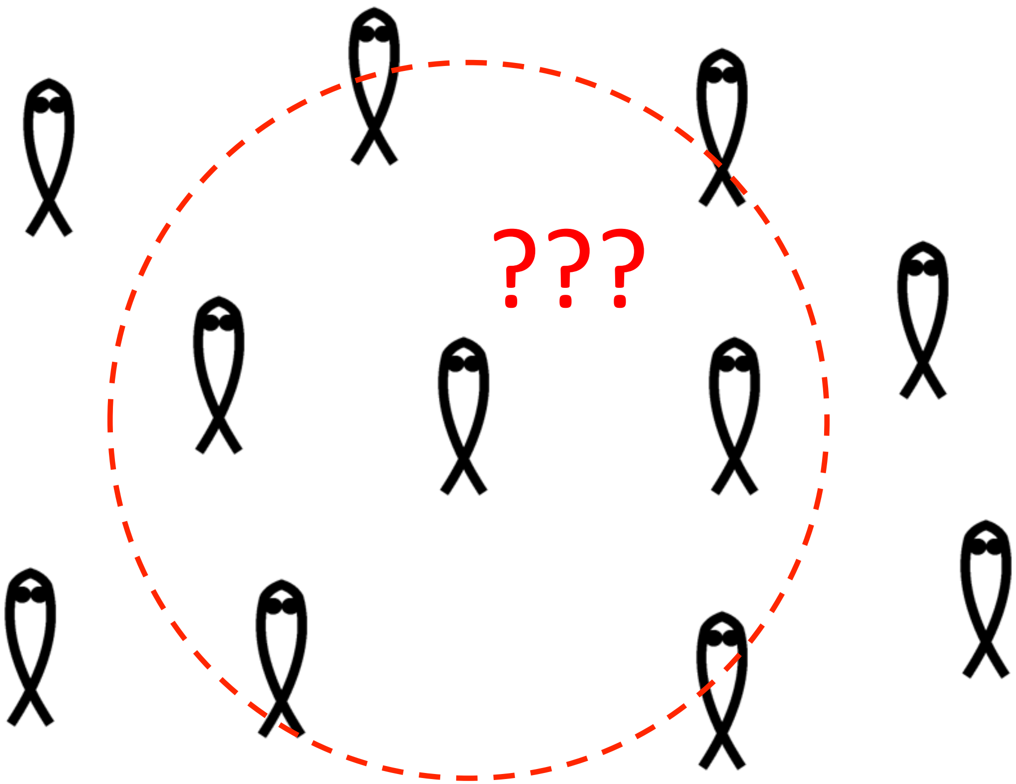
Mann *et al. PLoS Comp. Biol.* 2013

Problem 4: Additivity

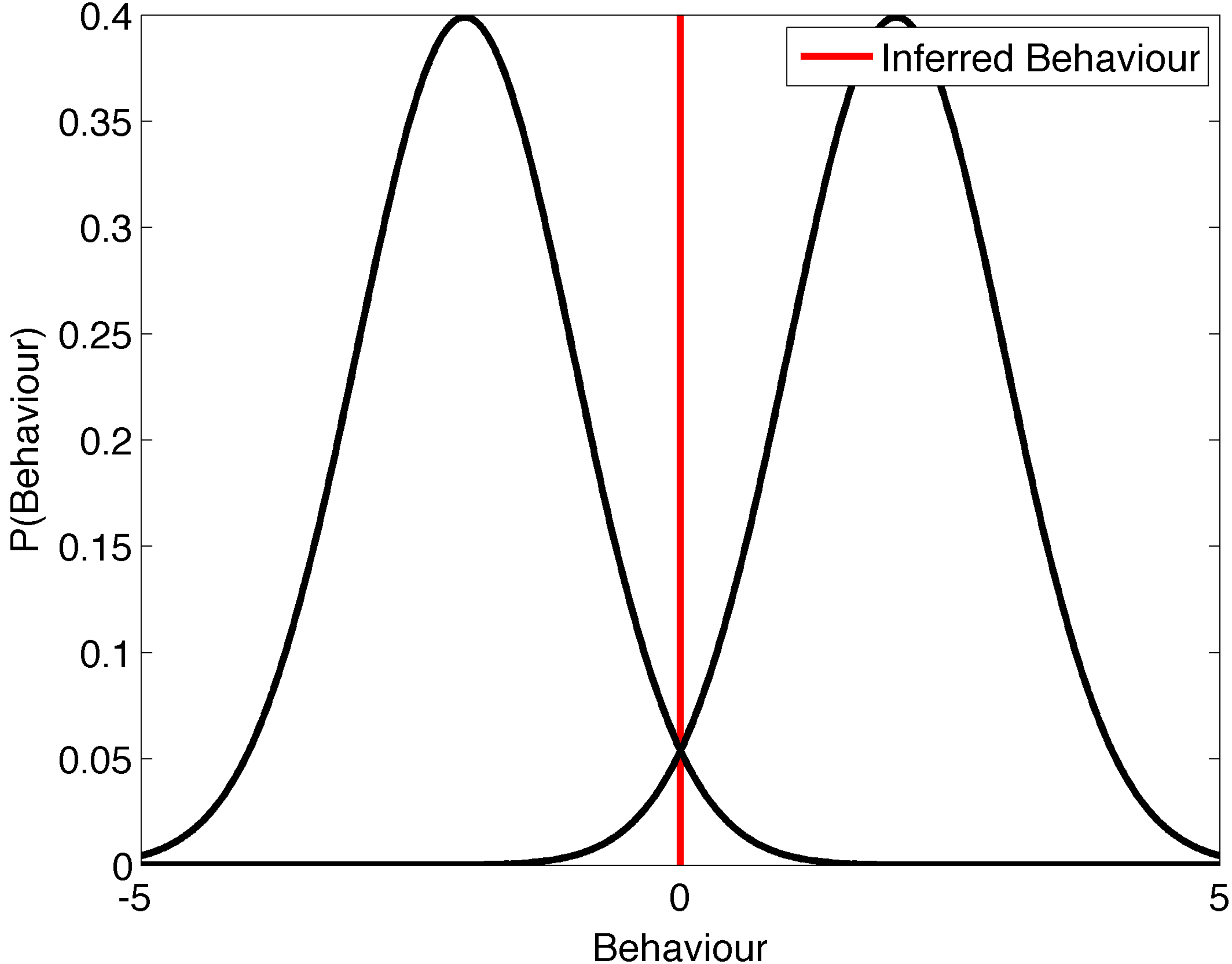
$$\mathbf{F(A \cup B) = F(A) + F(B)}$$

Fish Ain't Physics

Richard P Mann

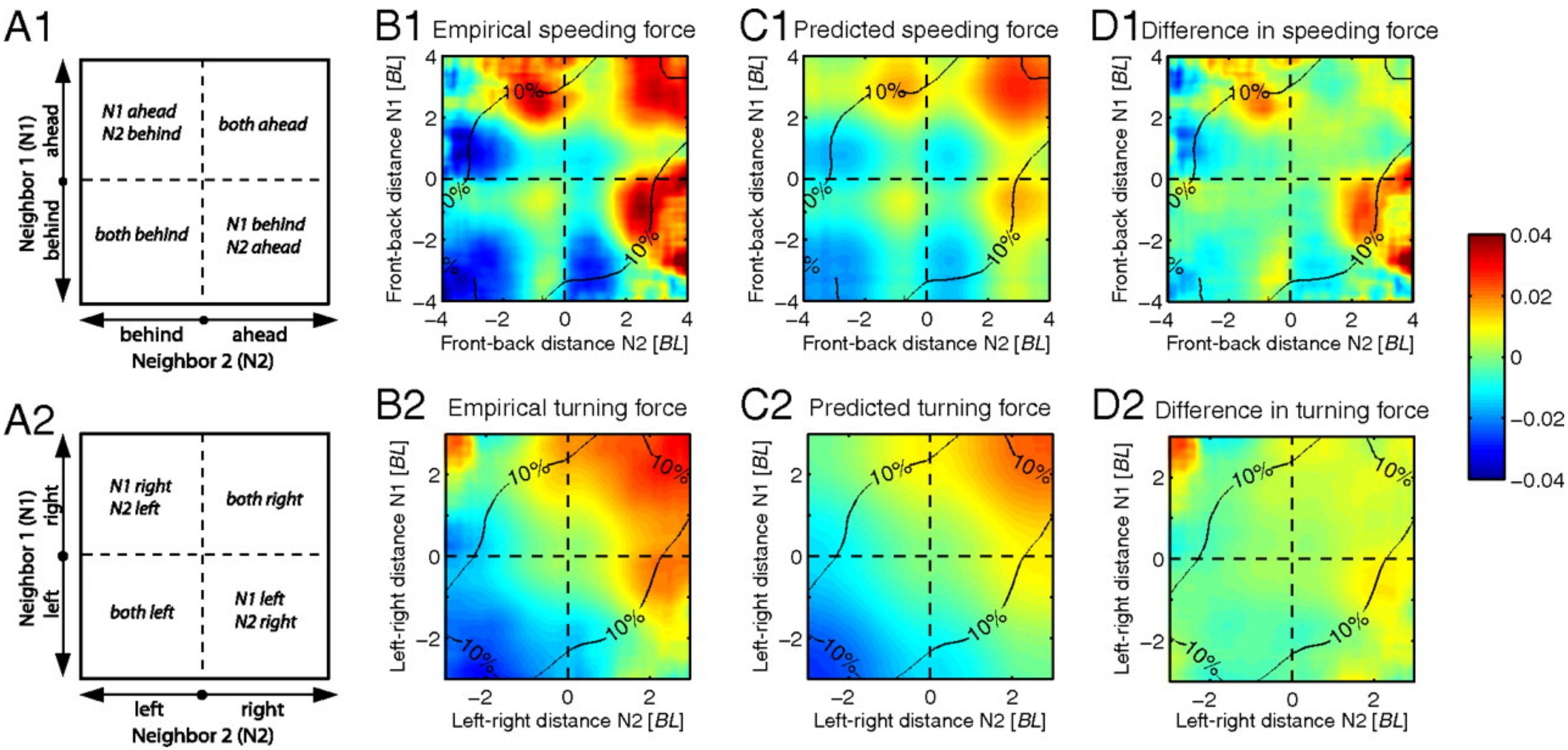


UPPSALA UNIVERSITET

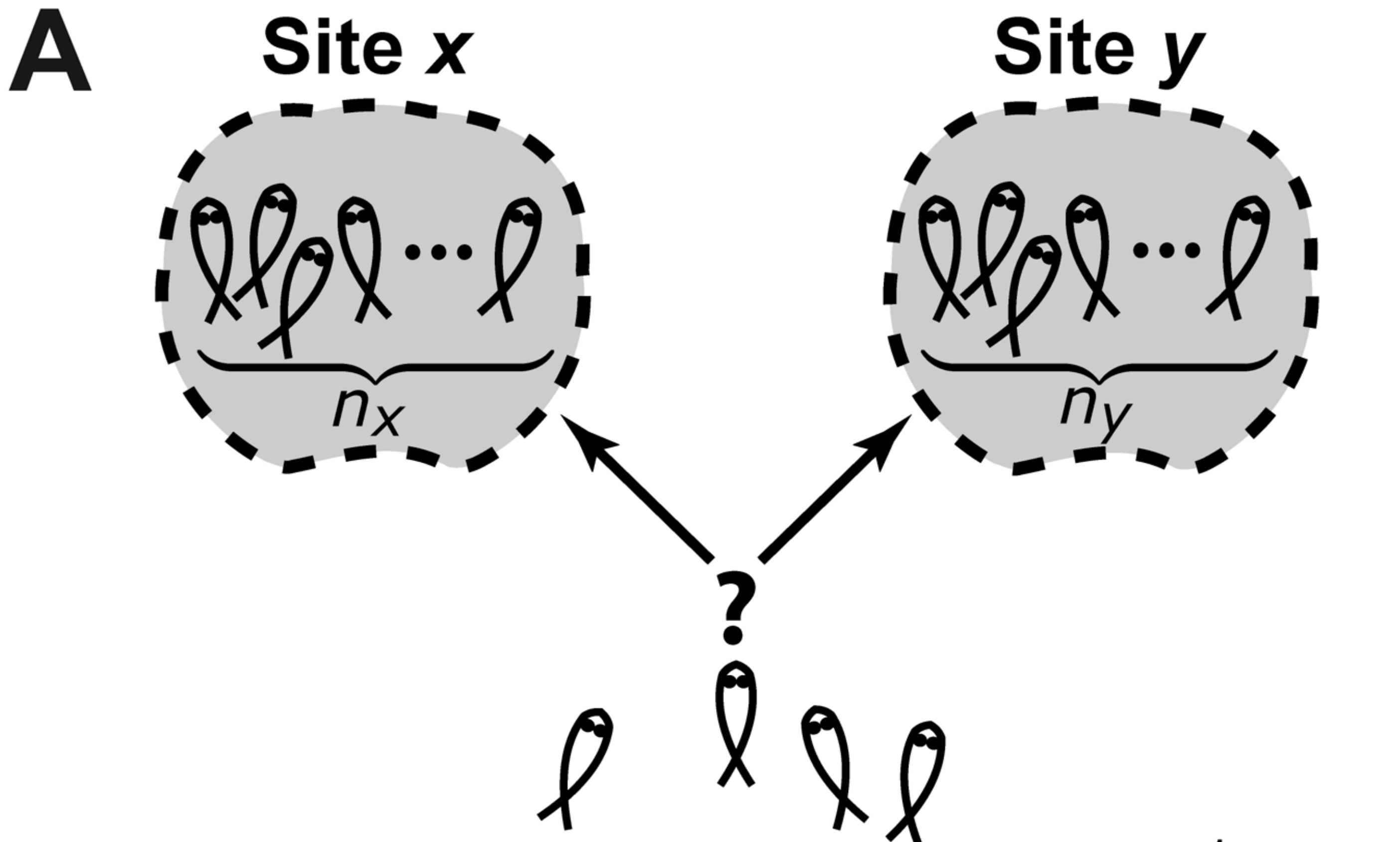


Inferred Behaviour
P(Behaviour)
Behaviour
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-5
0
5

Nonpairwise interactions in three-fish shoals.



Yael Katz et al. PNAS 2011;108:18720-18725



$$P_y = P(Y|\text{social info}) = \left(1 + s^{-\Delta n}\right)^{-1}$$

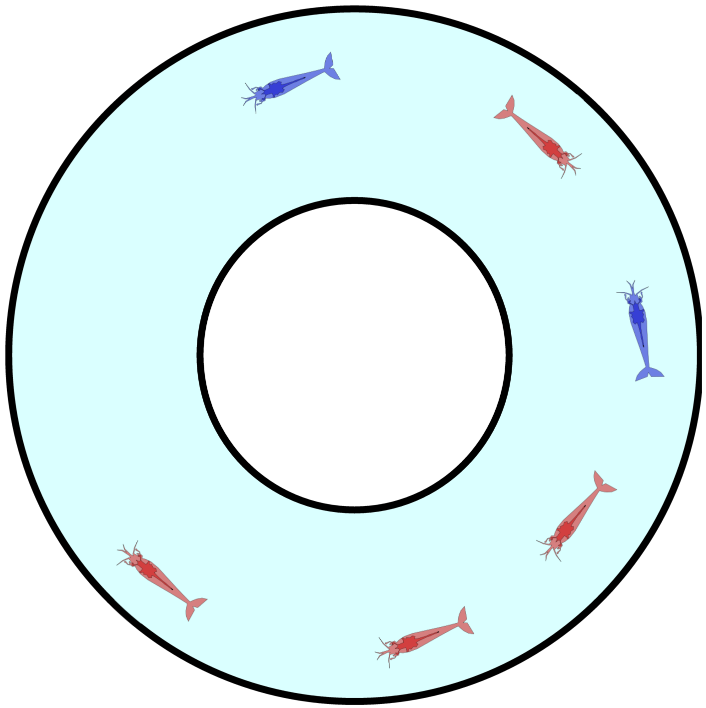
$$P_x = 1 - P_y = \left(1 + s^{\Delta n}\right)^{-1}; \quad \Delta n = n_y - n_x$$

Pérez-Escudero A, de Polavieja GG (2011) Collective Animal Behavior from Bayesian Estimation and Probability Matching. PLoS Comput Biol 7(11): e1002282. doi:10.1371/journal.pcbi.1002282
http://www.ploscompbiol.org/article/info:doi/10.1371/journal.pcbi.1002282

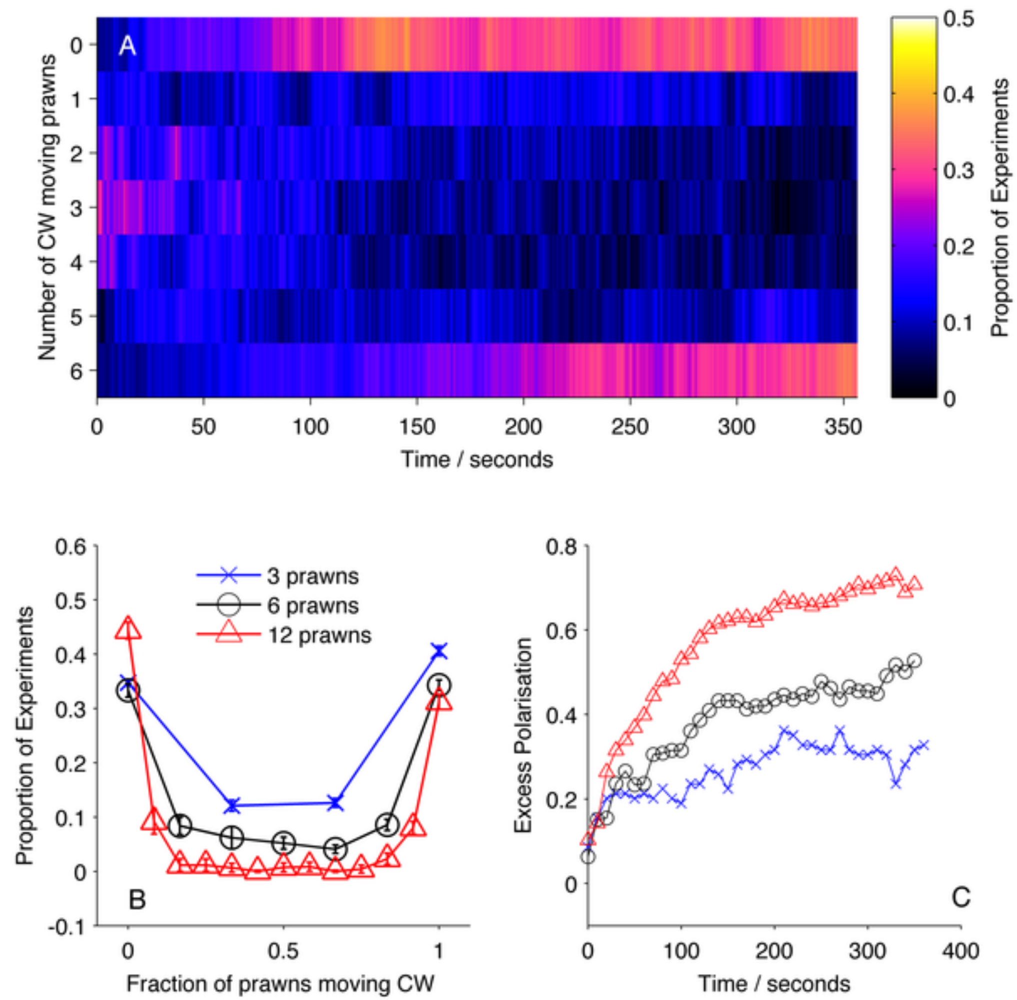
PLoS COMPUTATIONAL BIOLOGY

Problem 5: Emergence (again)

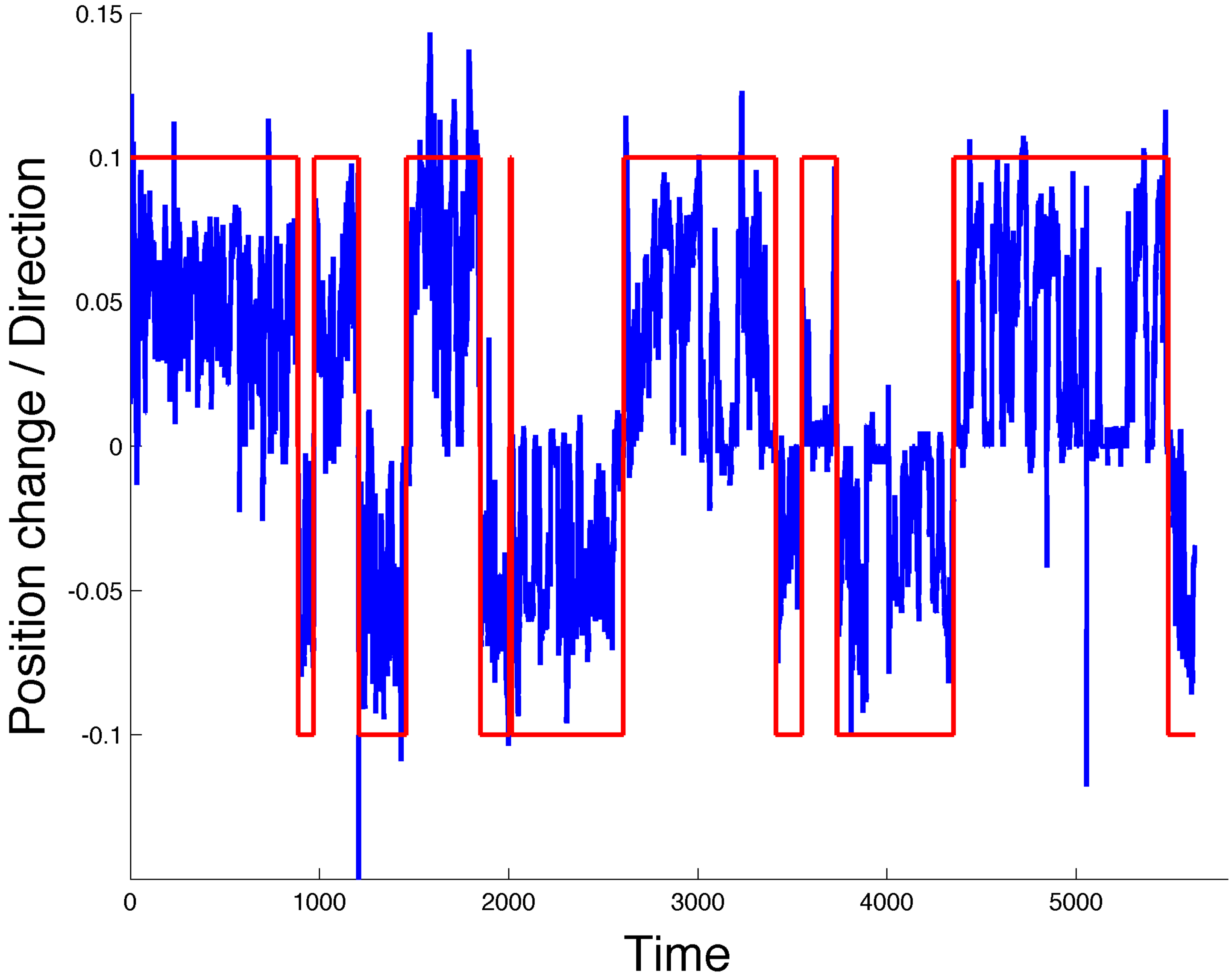
Bringing individual and collective behaviour together

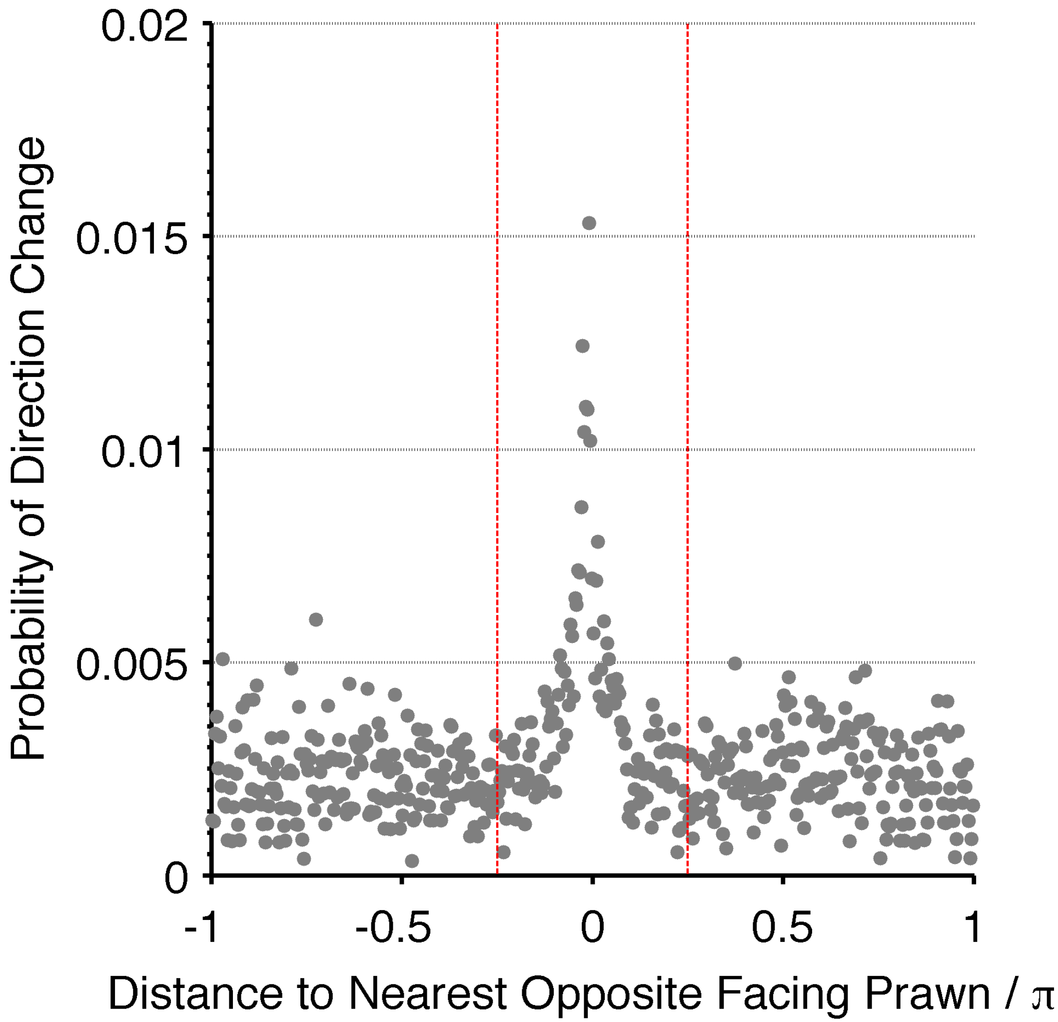


Mann *et al. PLoS Comp. Biol.* 2013

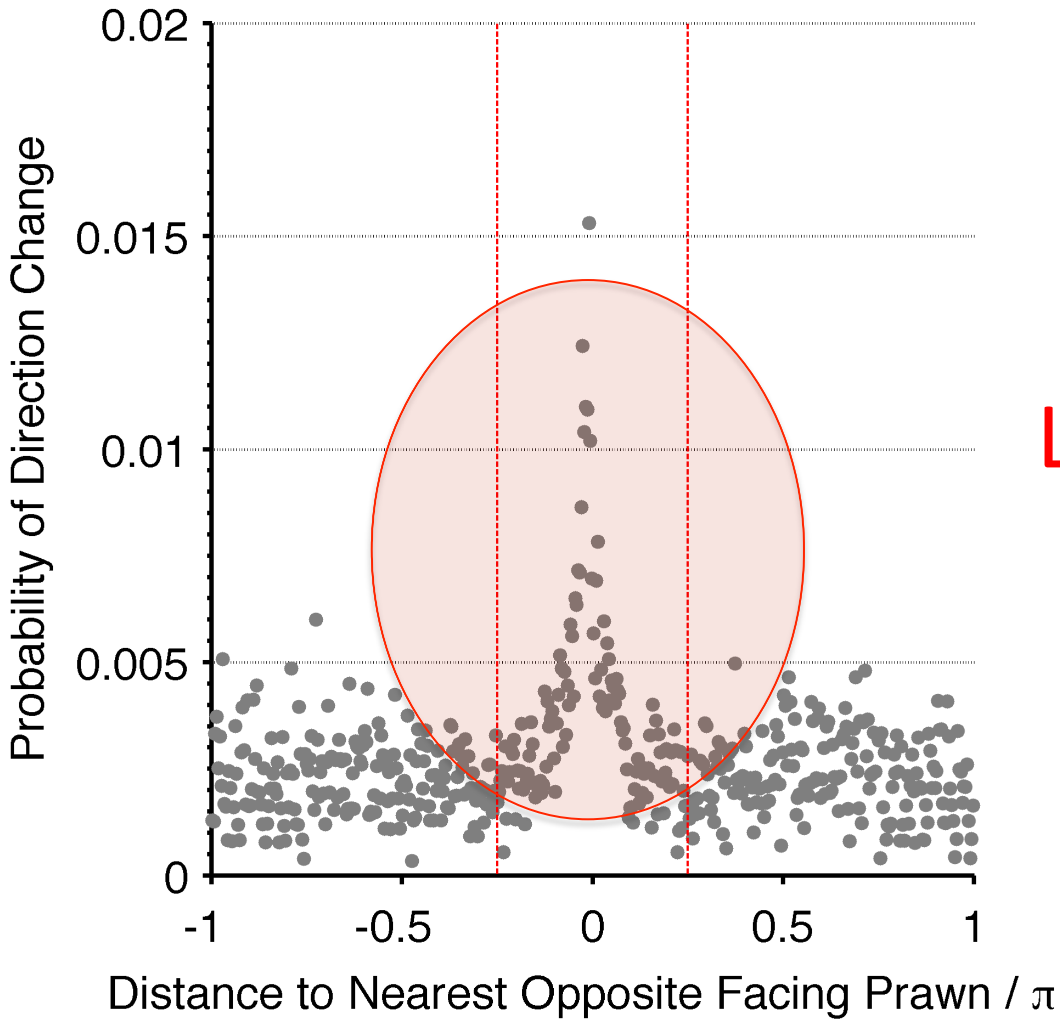


Explain the direction changes

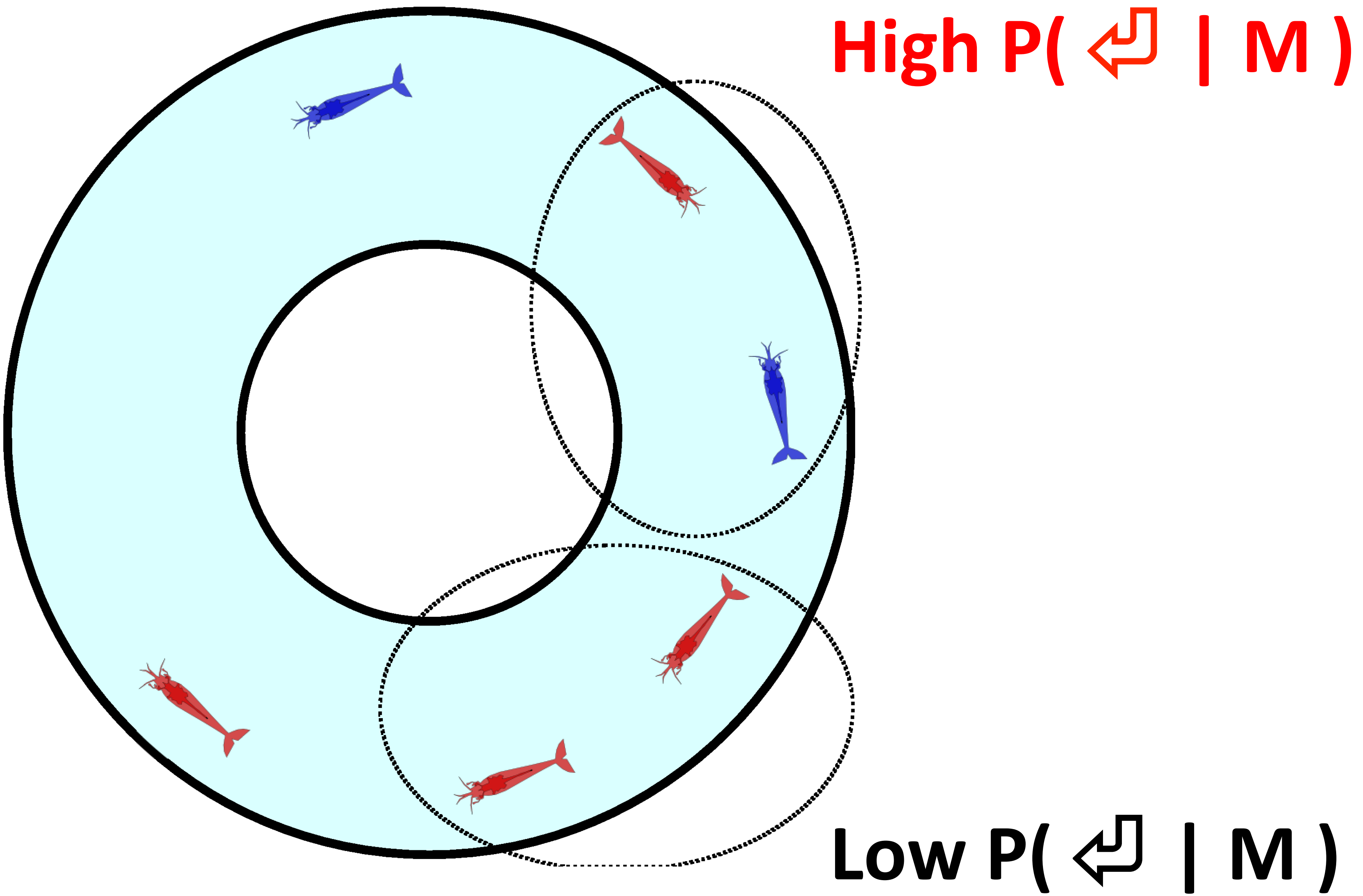


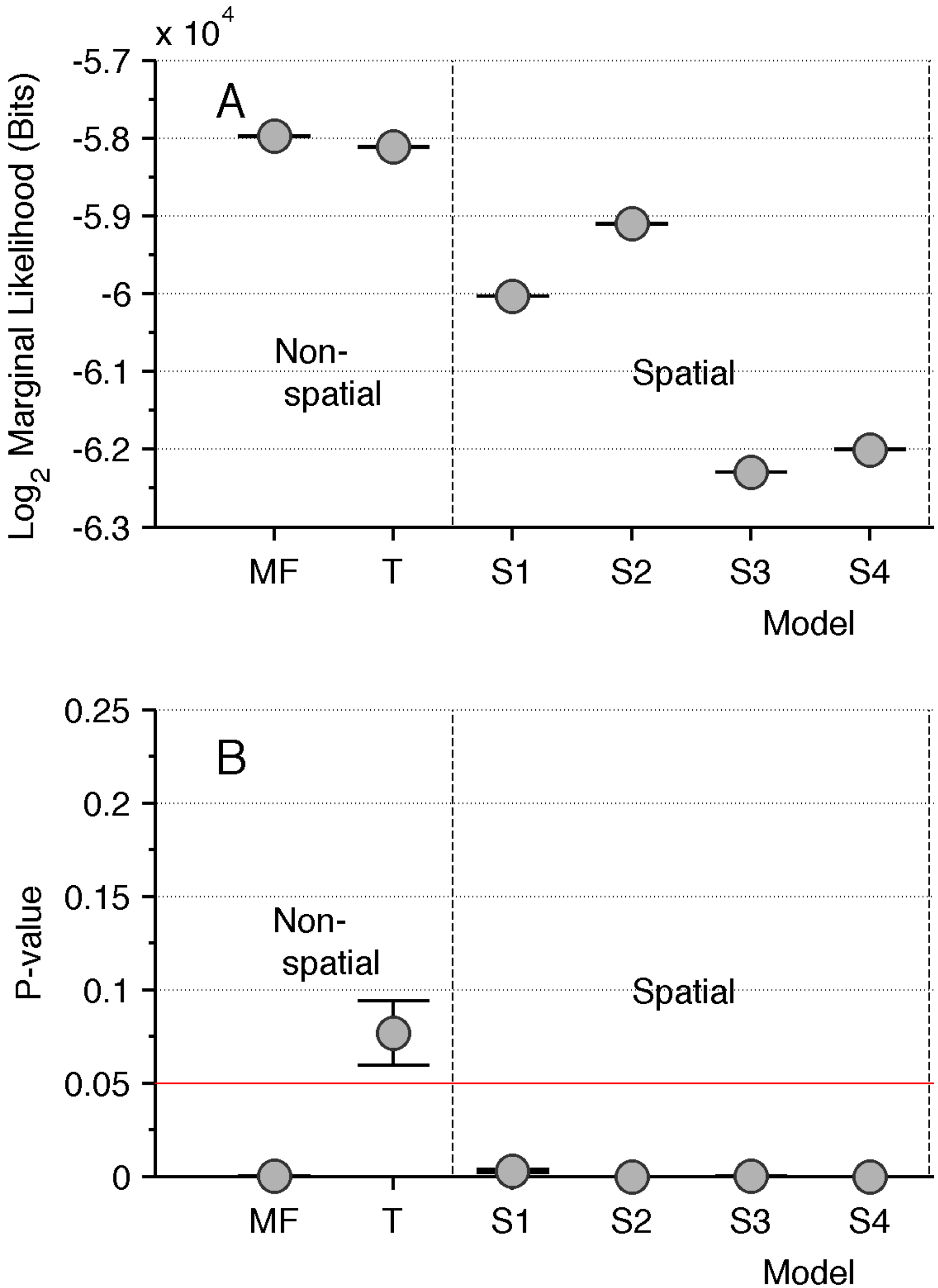


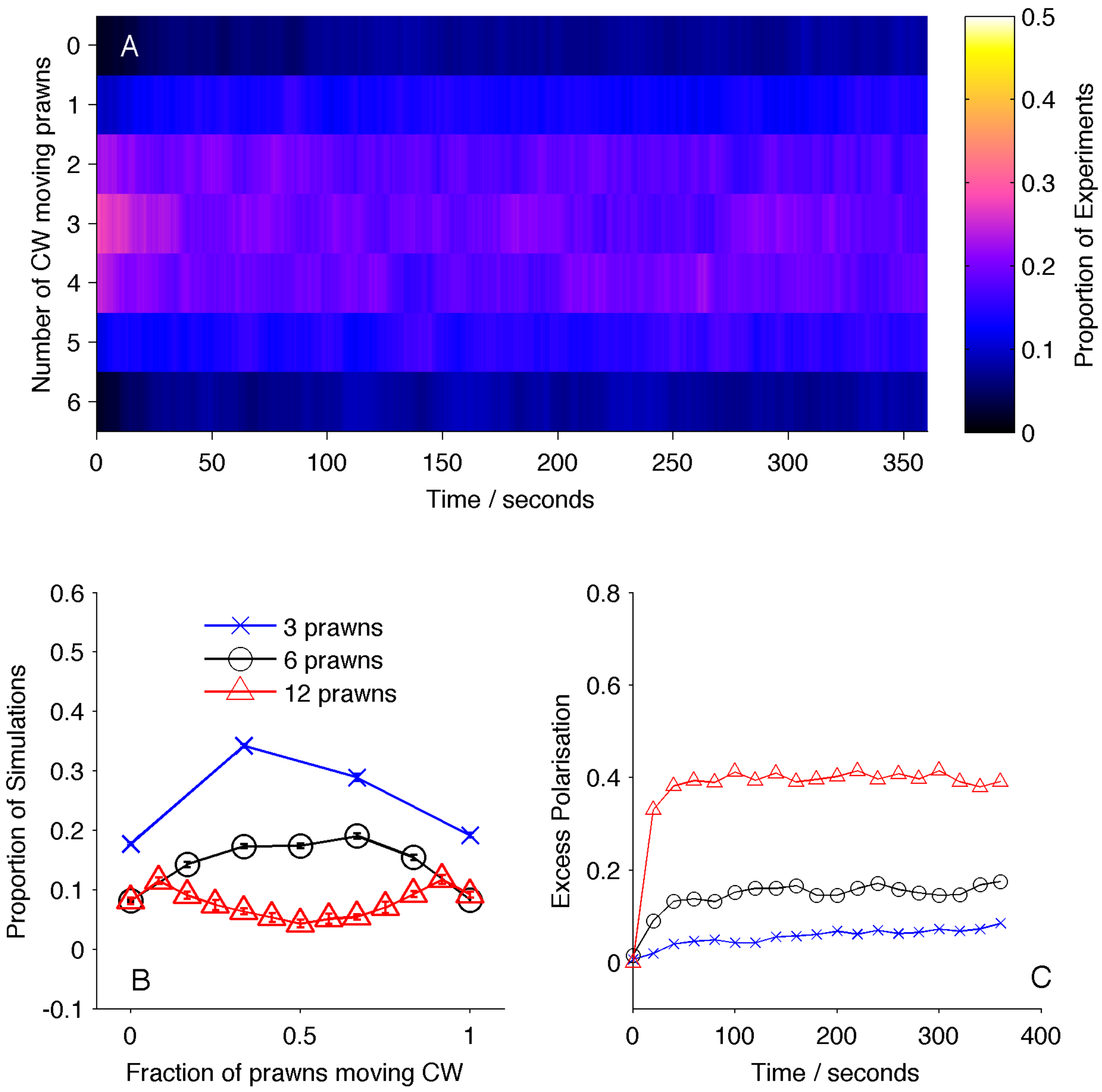
Probability of Direction Change
0.02
0.015
0.01
0.005
0
-1
-0.5
0
0.5
1
Distance to Nearest Opposite Facing Prawn / π



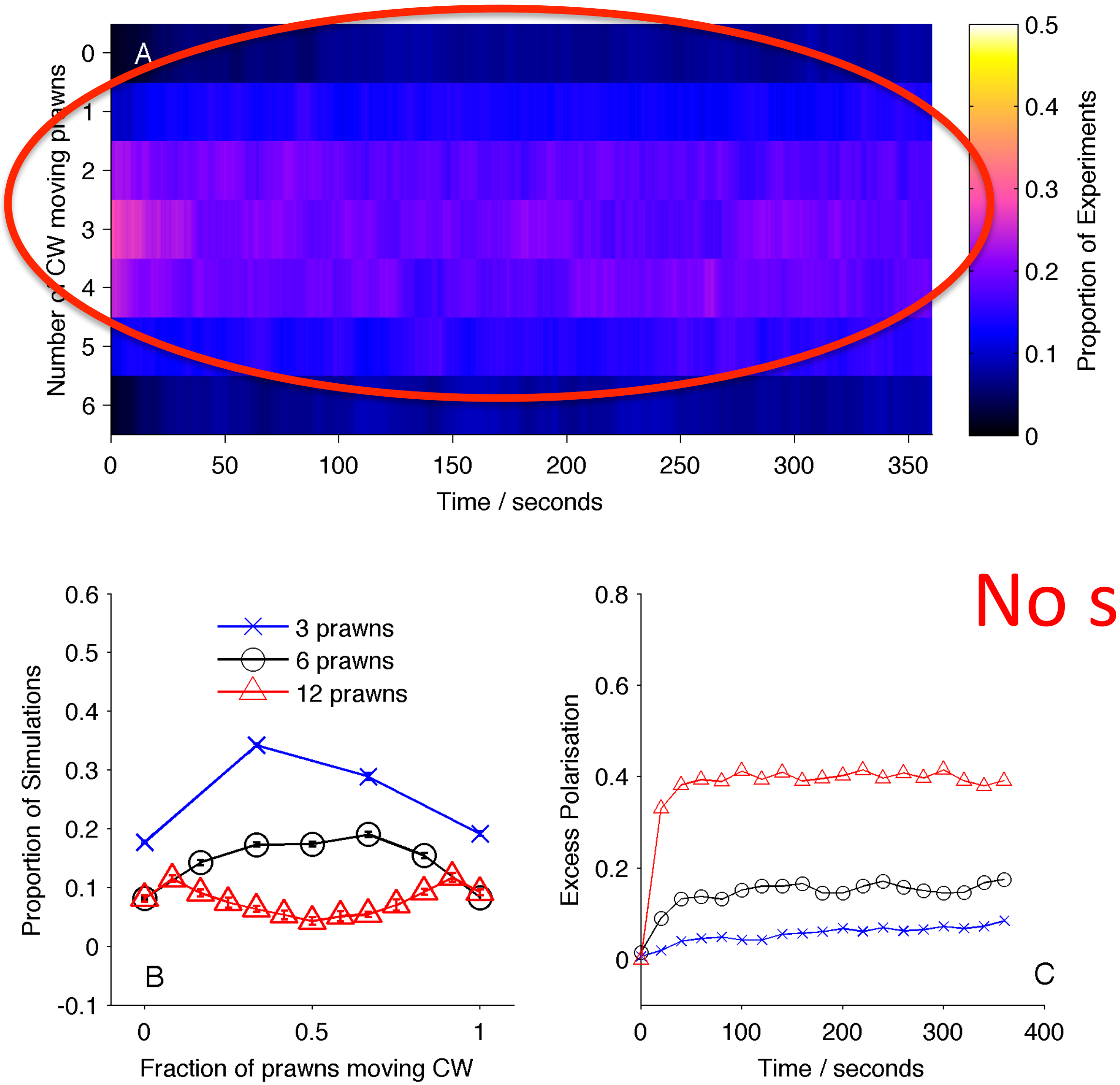
Local interactions


High P(⏎ | M)
Low P(⏎ | M)


x 10^4
Log2 Marginal Likelihood (Bits)
-5.7
-5.8
-5.9
-6
-6.1
-6.2
-6.3
A
Non-spatial
Spatial
MF
T
S1
S2
S3
S4
Model
P-value
0.25
0.2
0.15
0.1
0.05
0
B
Non-spatial
Spatial
MF
T
S1
S2
S3
S4
Model

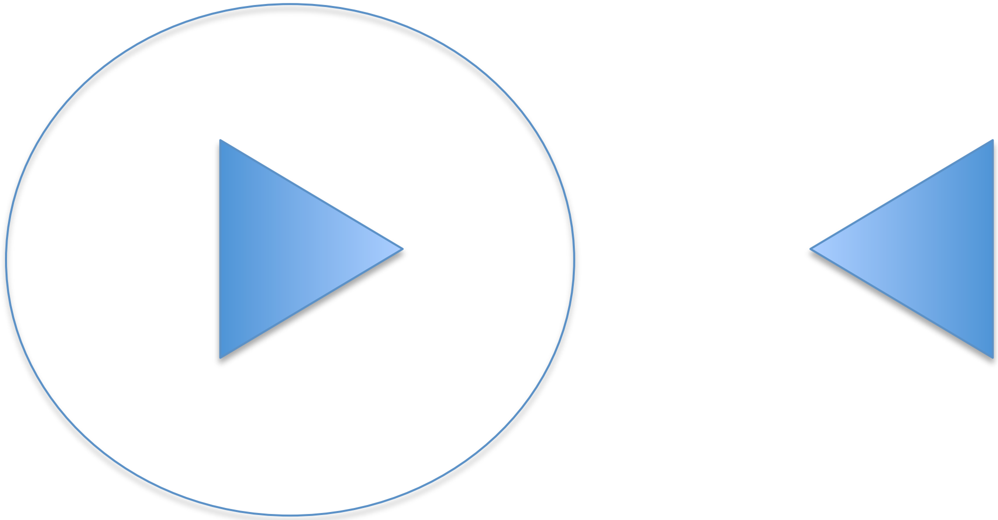


A
Number of CW moving prawns
0
1
2
3
4
5
6
0
50
100
150
200
250
300
350
Time / seconds
Proportion of Experiments
0
0.1
0.2
0.3
0.4
0.5
B
Proportion of Simulations
3 prawns
6 prawns
12 prawns
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
0
0.5
1
Fraction of prawns moving CW
C
Excess Polarisation
0.8
0.6
0.4
0.2
0
0
100
200
300
400
Time / seconds



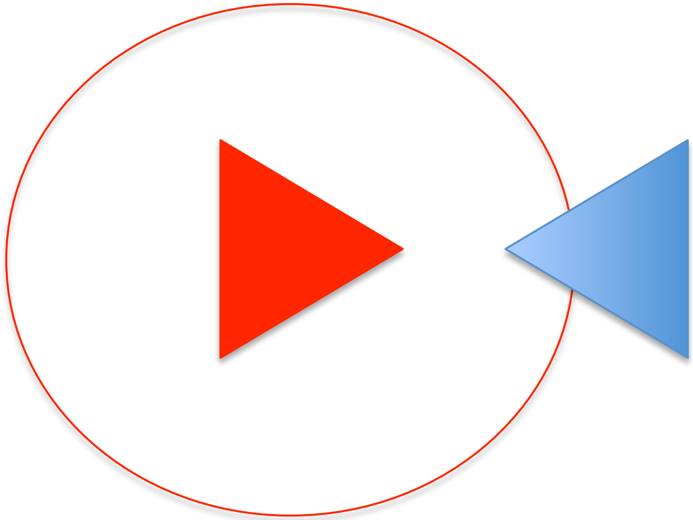
No self-organisation

A non-Markovian interaction



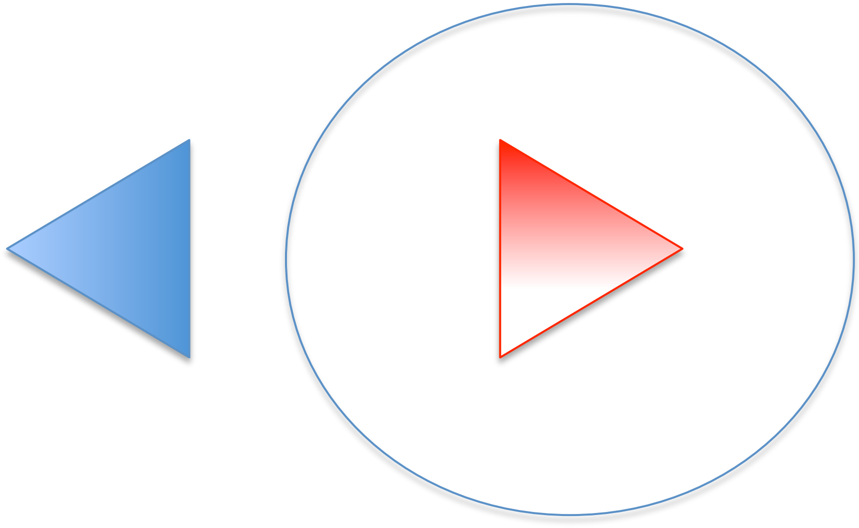
Low probability of changing direction

A non-Markovian interaction

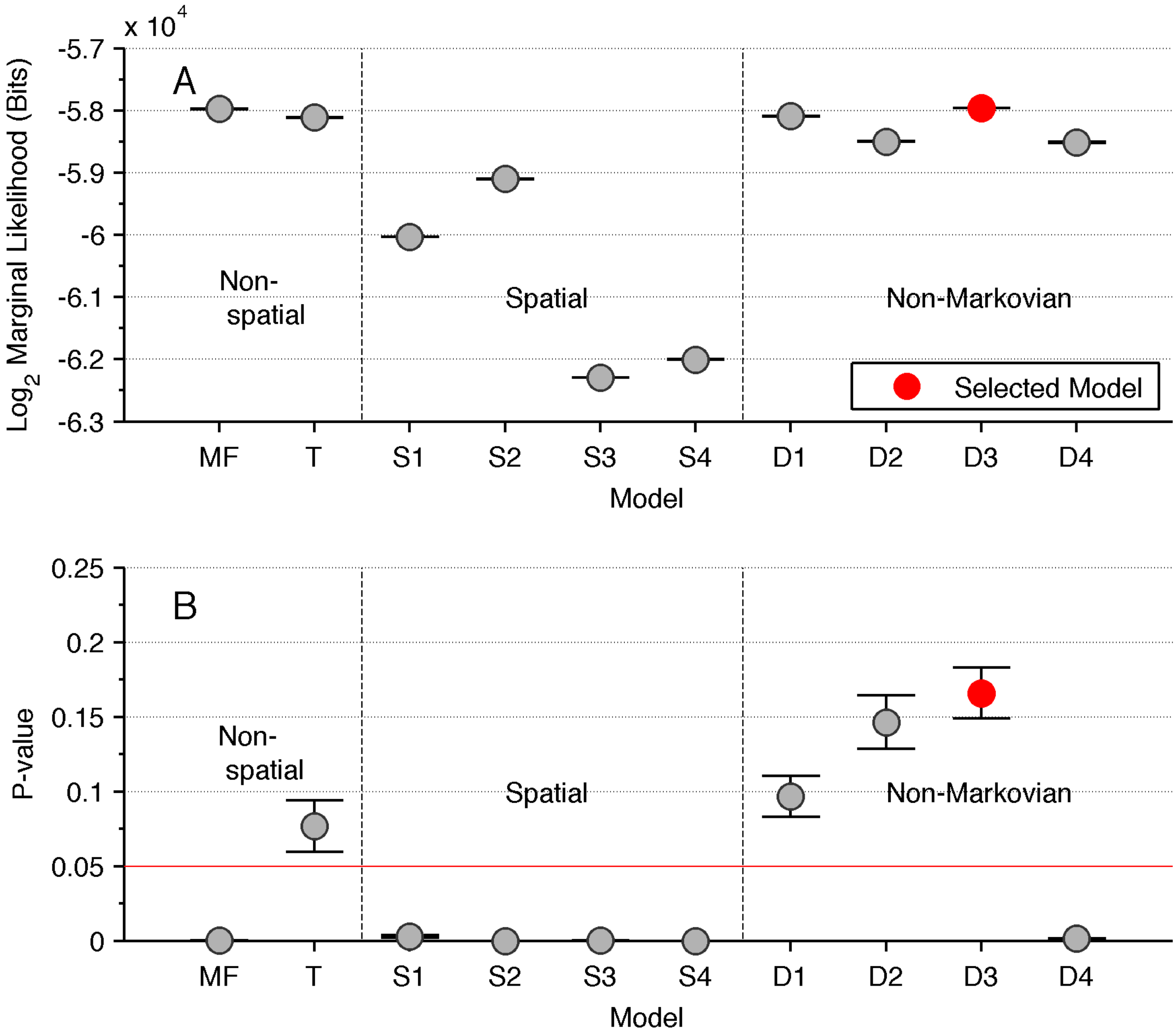


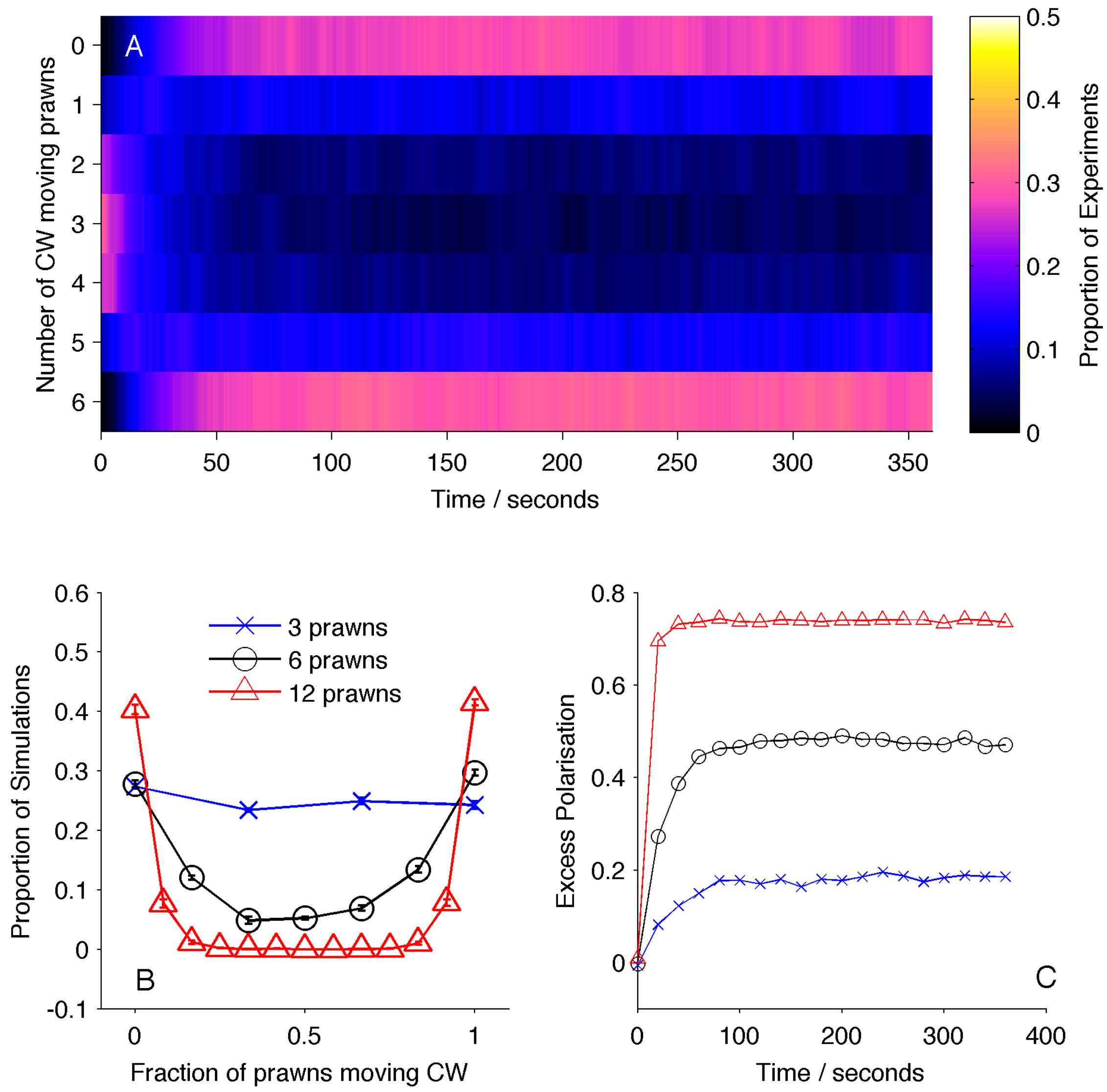
High probability of changing direction

A non-Markovian interaction

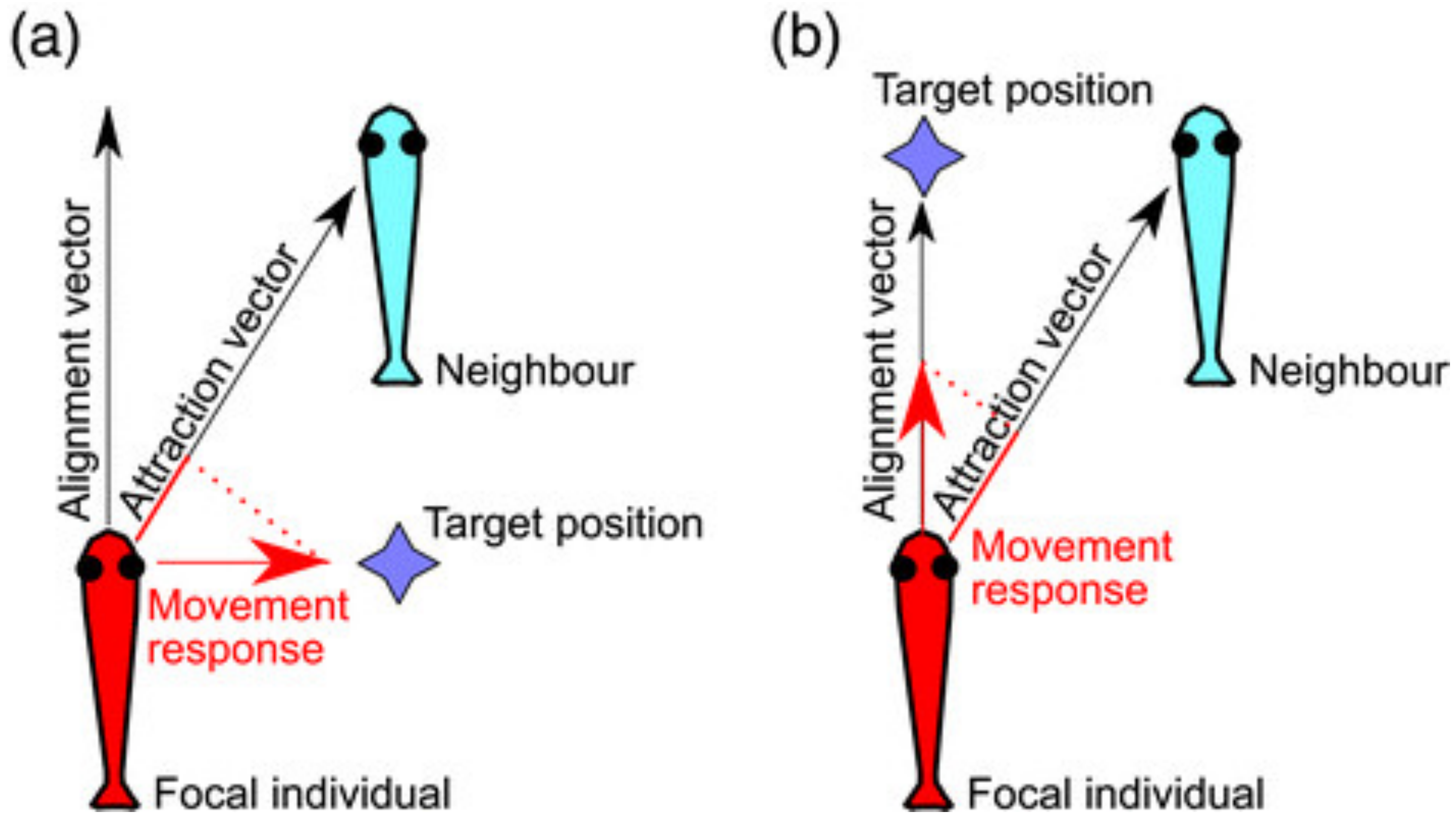


Still raised probability of changing direction


x 10^4
-5.7
-5.8
-5.9
-6
-6.1
-6.2
-6.3
Log2 Marginal Likelihood (Bits)
A
Non-
spatial
Spatial
Non-Markovian
Selected Model
MF
T
S1
S2
S3
S4
D1
D2
D3
D4
Model
0.25
0.2
0.15
0.1
0.05
0
P-value
B
Non-
spatial
Spatial
Non-Markovian
MF
T
S1
S2
S3
S4
D1
D2
D3
D4
Model


A
Number of CW moving prawns
Time / seconds
Proportion of Experiments
B
3 prawns
6 prawns
12 prawns
Proportion of Simulations
Fraction of prawns moving CW
C
Excess Polarisation
Time / seconds

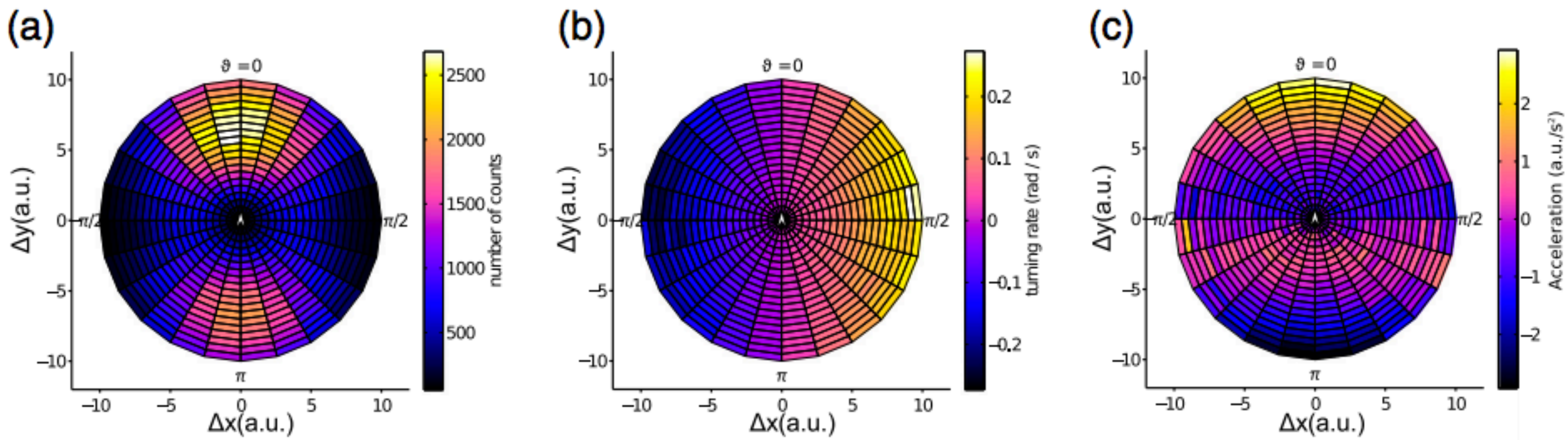
6. Responses are ambiguous



Perna *et al.* On the duality between interaction responses and mutual positions in flocking and schooling. *Movement Ecology* 2014

7. Expected responses mirror regression to the mean

7. Expected responses mirror regression to the mean



Perna *et al.* On the duality between interaction responses and mutual positions in flocking and schooling. *Movement Ecology* 2014


"THE HORROR.."

Return of agent 1 to equilibrium
AFTER agent 2 leaves.

Conclusion: The era of naïve model fitting & validation is over. Time to level up