Using sparse odor cues to infer the location of their source.

N. Rigolli[1,2], N. Magnoli[2], L. Rosasco[3], A. Seminara[1]
1. Université Côte d'Azur, CNRS 2. DIFI, University of Genoa, INFN
3. MaLGa, DIBRIS, University of Genoa

Bio-inspired motivation

Different animals localize odor sources with high precision even when cues are intermittent.

- **how can they make successfully predictions in such a turbulent environment?**

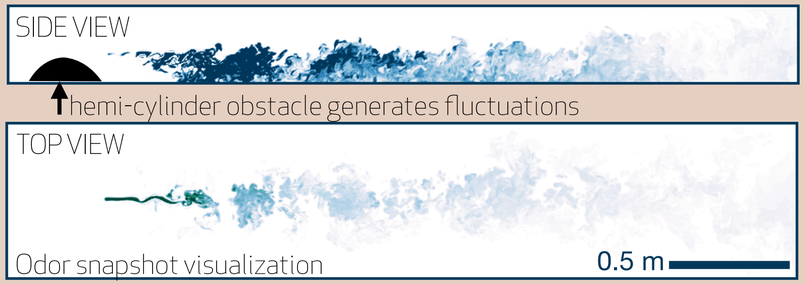
Direct Numerical Simulation (Nek 5000)

We performed a realistic fluid dynamics numerical simulation of odor transport in water motion, solving:

$$\partial_t u + u \cdot \nabla u = -\frac{1}{\rho}\nabla P + \nu \nabla^2 u + f \qquad \text{Incompressible Navier-Stokes equations}$$

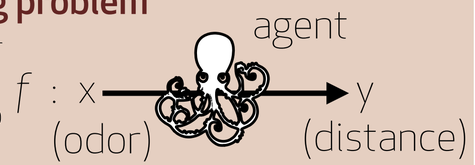
$$\nabla \cdot u = 0$$

$$\partial_t \theta + u \nabla \theta = D \nabla^2 \theta \qquad \text{Passive scalar equation}$$



A supervised learning problem

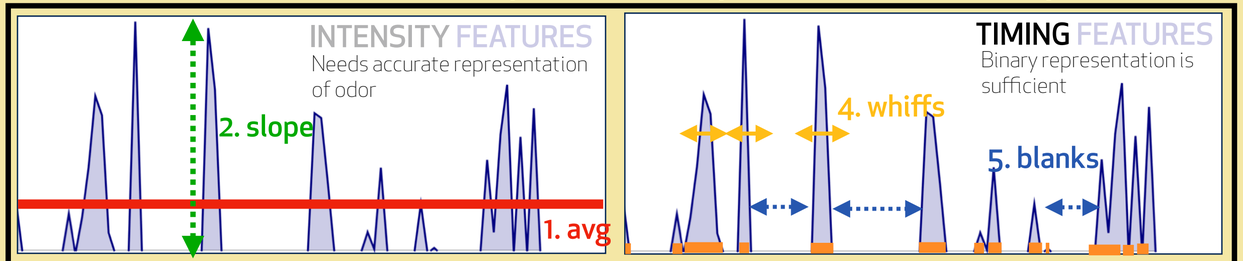
We used the simulated odor fields to find a function f that takes odor as input x to predict the distance from the source (output y).



Supervised learning algorithm rationale:

- Compose many examples of input/output (x_i, y_i) from simulations (=training set)
- Learn function f: $x \rightarrow y$ from training set
- Apply f to new input point to predict its output -> measure prediction error in new datapoints

What is the best definition of input x to best predict the distance y ?



We mapped the odor signal into 5 different features:

1. <concentration> 2. <slope>
3. Intermittency 4. < duration whiffs > 5. < duration blanks >

We used these 5 features as input for the algorithm. Either individually or in pairs or all together.

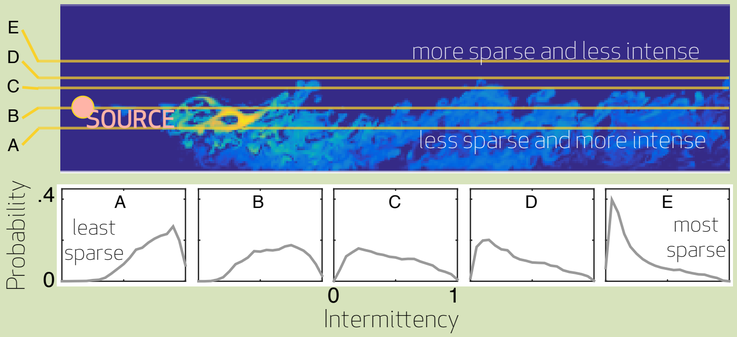
To evaluate algorithm performance, we computed:

$$\chi = \frac{< (f(x) - y_{real})^2 >}{\sigma_y^2}$$

$\chi = 0$ is the perfect prediction

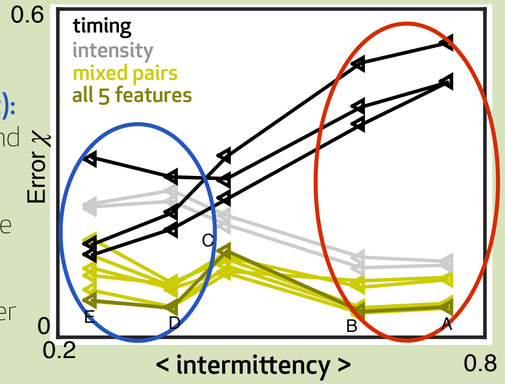
$\chi = 1$ is a trivial prediction

How do ranking of features varies with height?



More intermittent dataset (large height):

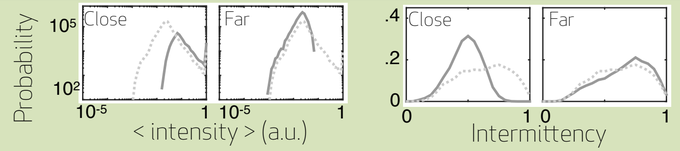
- intensity and **timing** are equal
- pairs are the best
- All 5 do not improve over pairs

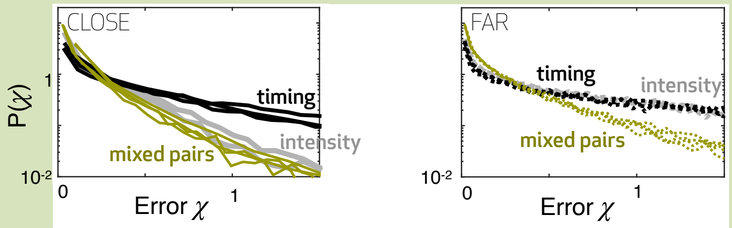


Less intermittent dataset (low height):

- intensity is better than **timing**
- pairs are the best
- All 5 do not improve over pairs

How do ranking of features varies with distance?





Close to the source:

- intensity is better than **timing**
- pairs are the best

Far from the source:

- intensity and **timing** are equal
- pairs are the best

CONCLUSIONS:

1) Ranking of individual features highly depends on dataset.
2) Pairing one intensity feature with one timing feature is the best choice and robust to different conditions.
3) Do 1) and 2) constrain odor representation in the brain ?