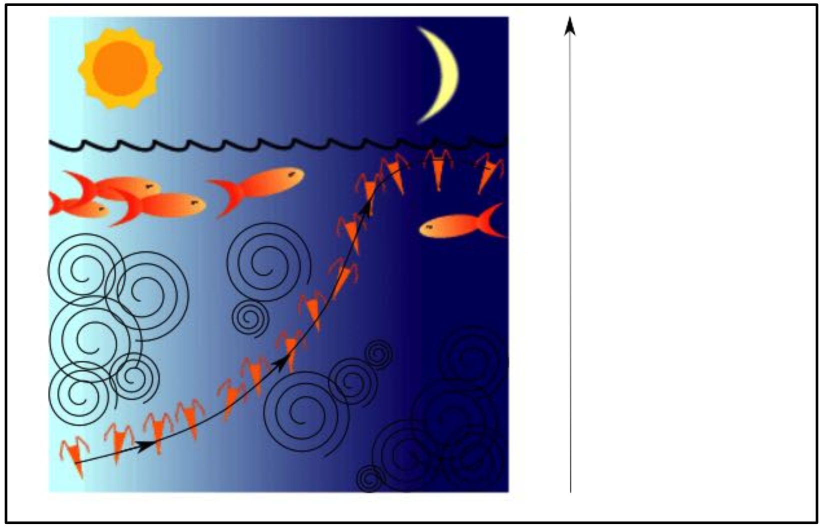
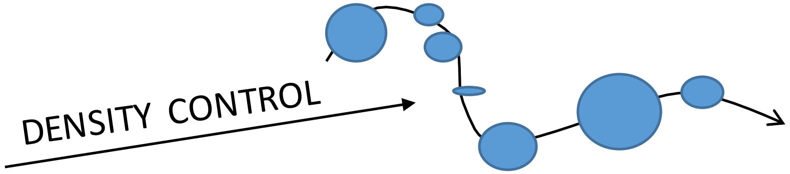
Flow navigation by smart particles via Reinforcement Learning

Luca Biferale
Dept. Physics, INFN & CAST
University of Rome 'Tor Vergata'
biferale@roma2.infn.it
PIML2018 SANTA FE

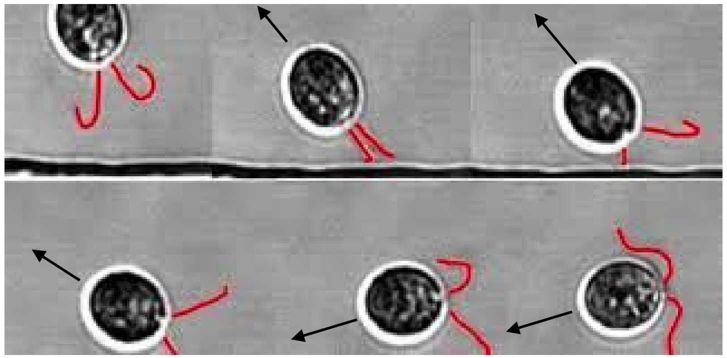


CREDITS: SIMONA COLABRESE (TOR VERGATA UNIV. ROME-IT); ANTONIO CELANI (ICTP TRIESTE-IT); KRISTIAN GUSTAVSSON (GOTHEBORG UNIV. SWEDEN)

cost EuHIT erc European Research Council Established by the European Commission Supporting top researchers from anywhere in the world INFN Istituto Nazionale di Fisica Nucleare Tor Vergata HPC-LEAP EUROPEAN JOINT DOCTORATES



- PARTICLES IN COMPLEX FLOWS I: **SMART INERTIAL PARTICLES**
- PARTICLES IN COMPLEX FLOWS II: **SMART MICROSWIMMERS**

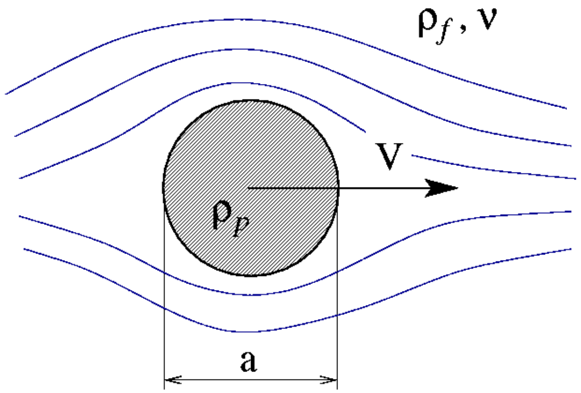


- **Flow navigation by smart microswimmers via reinforcement learning**
S Colabrese, K Gustavsson, A Celani, L Biferale
Physical Review Letters 118 (15), 158004, 2017

-**Smart Inertial Particles**
S Colabrese, K Gustavsson, A Celani, L Biferale
arXiv preprint arXiv:1711.05853, 2017

- **Finding efficient swimming strategies in a three-dimensional chaotic flow by reinforcement learning**
K Gustavsson, L Biferale, A Celani, S Colabrese
The European Physical Journal E 40 (12), 110, 2017

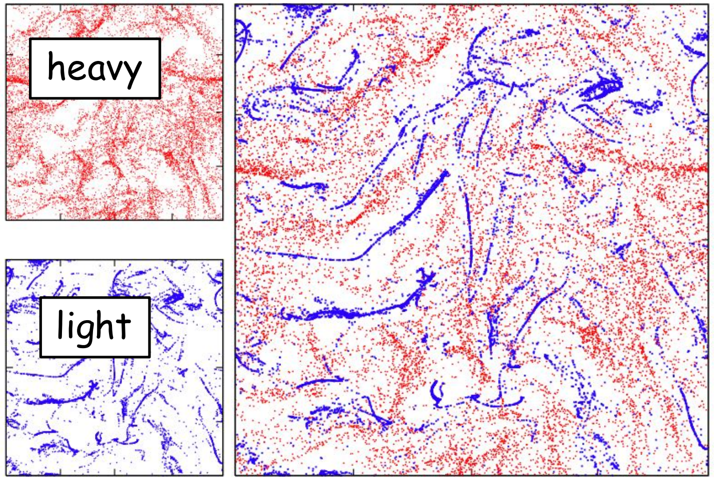
PARTICLES IN COMPLEX FLOWS I: INERTIAL PARTICLES



$$\frac{d\boldsymbol{X}}{dt} = \boldsymbol{V}$$

$$\frac{d\boldsymbol{V}}{dt} = \beta \frac{D\boldsymbol{u}(\boldsymbol{X},t)}{Dt} + \frac{\boldsymbol{u}(\boldsymbol{X},t) - \boldsymbol{V}}{\tau}$$

$$\partial_t \mathbf{u} + (\mathbf{u} \cdot \partial)\mathbf{u} = -\partial P + \nu \nabla \mathbf{u}$$



$$\beta = \frac{3\rho_f}{\rho_f + 2\rho_p}$$

$$\tau = \frac{a^2}{3\nu\beta}$$

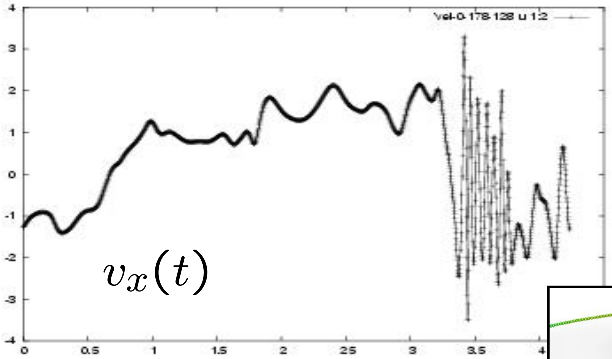
$\beta<1$ heavy particles
$\beta>1$ light particles

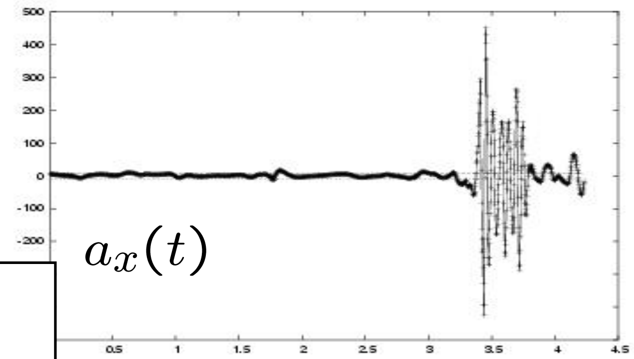
Drag: **Stokes Time**

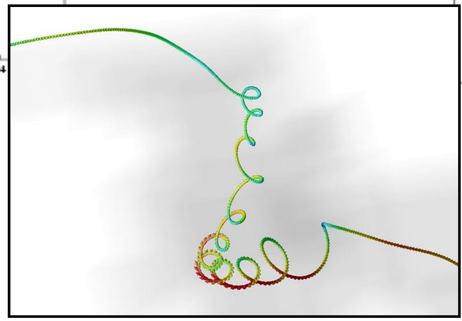
Preferential concentration!

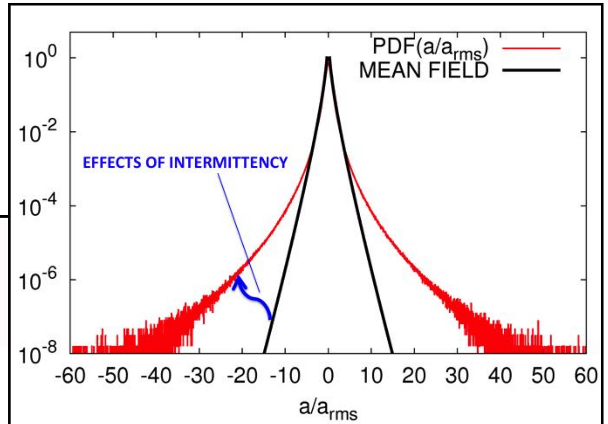
Light(heavy) particles accumulate inside(outside) highly vortical regions

Maxey, *J. Fluid Mech.* **174**, 441 (1987); Falkovich *et al*, *Phys. Rev. Lett.* **86**, 2790 (2001)

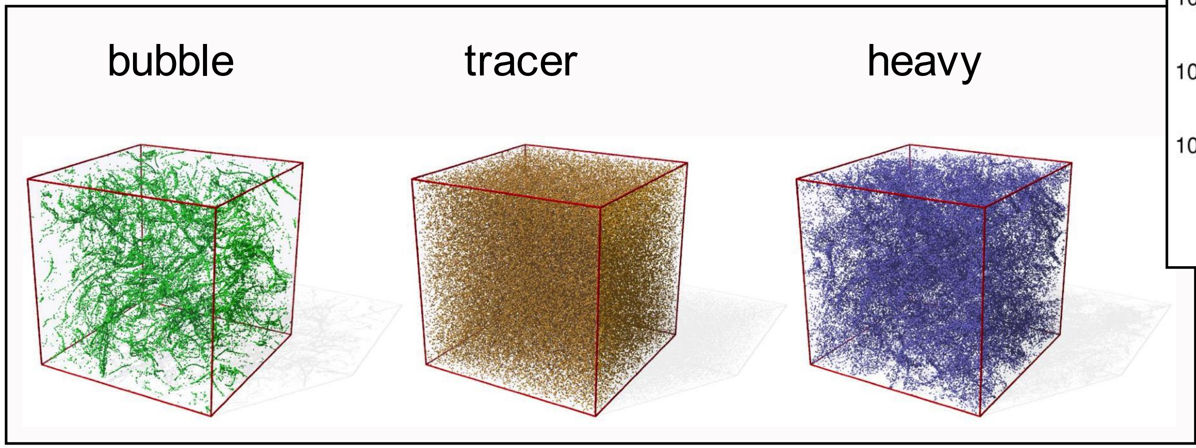








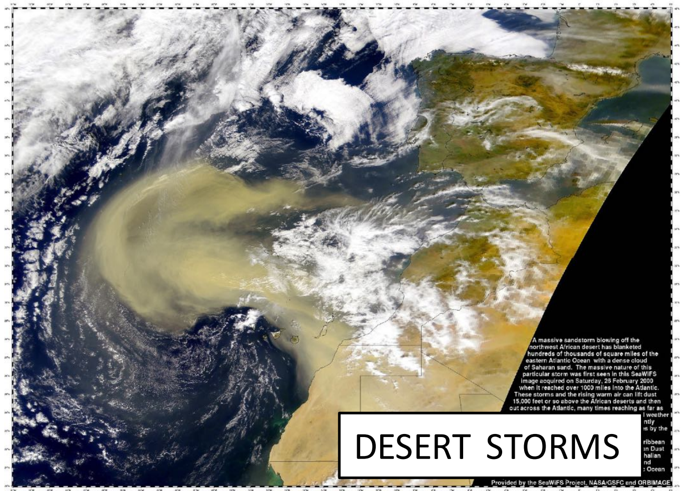
ACCELERATION PROBABILITY DISTRIBUTION FUNCTION (PDF) AT RE ~ 10^5 [Bi04]
COMPARED WITH THE PREDICTION FROM MEAN FIELD (KOLMOGOROV THEORY)

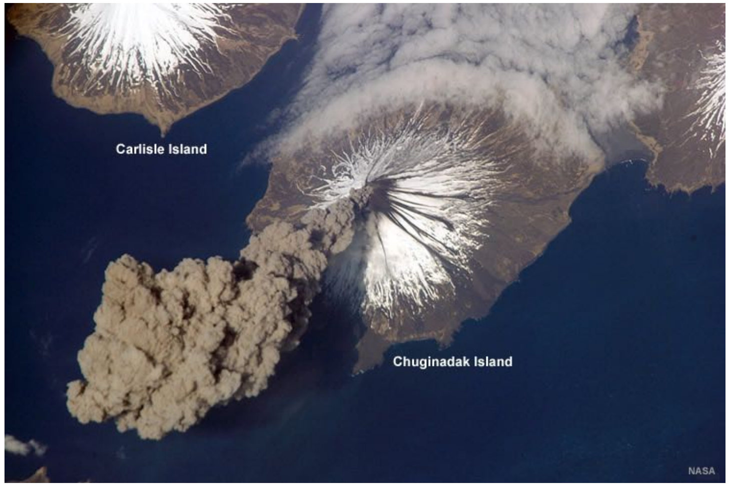


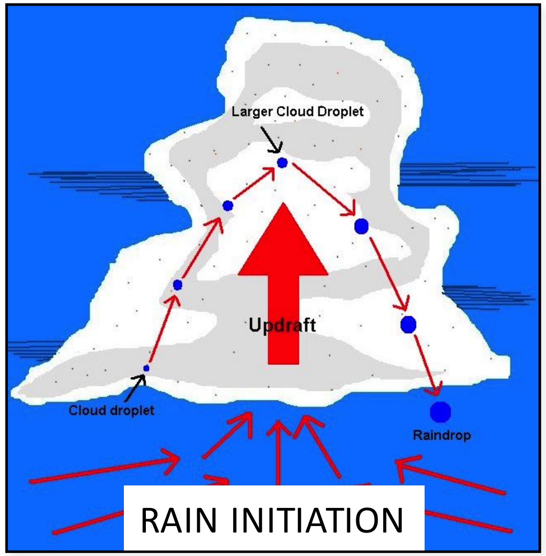
<u>Particle trapping in three-dimensional fully developed turbulence</u>

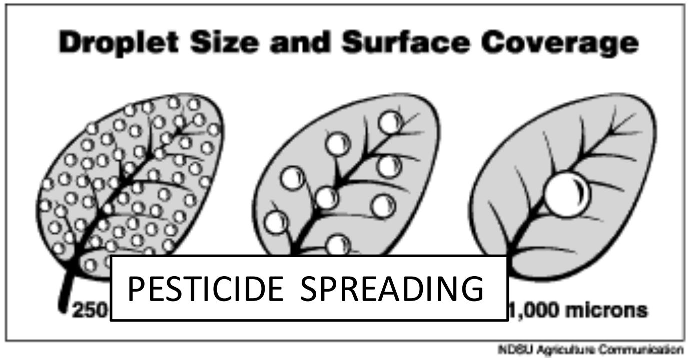
L.B., G Boffetta, A Celani, A Lanotte, F Toschi

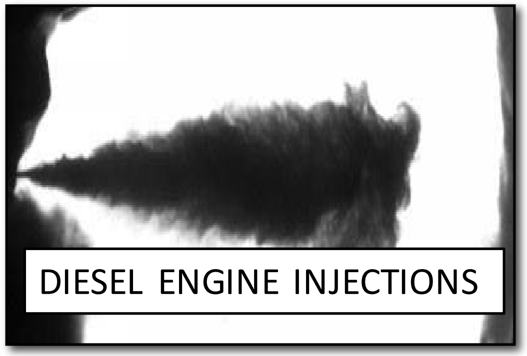
Physics of Fluids 17 (2), 021701

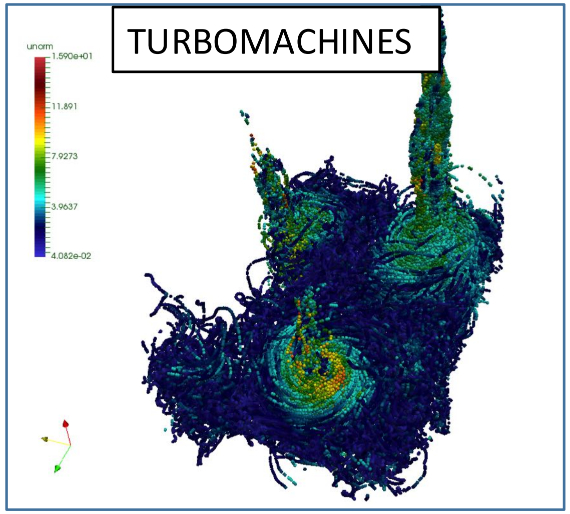






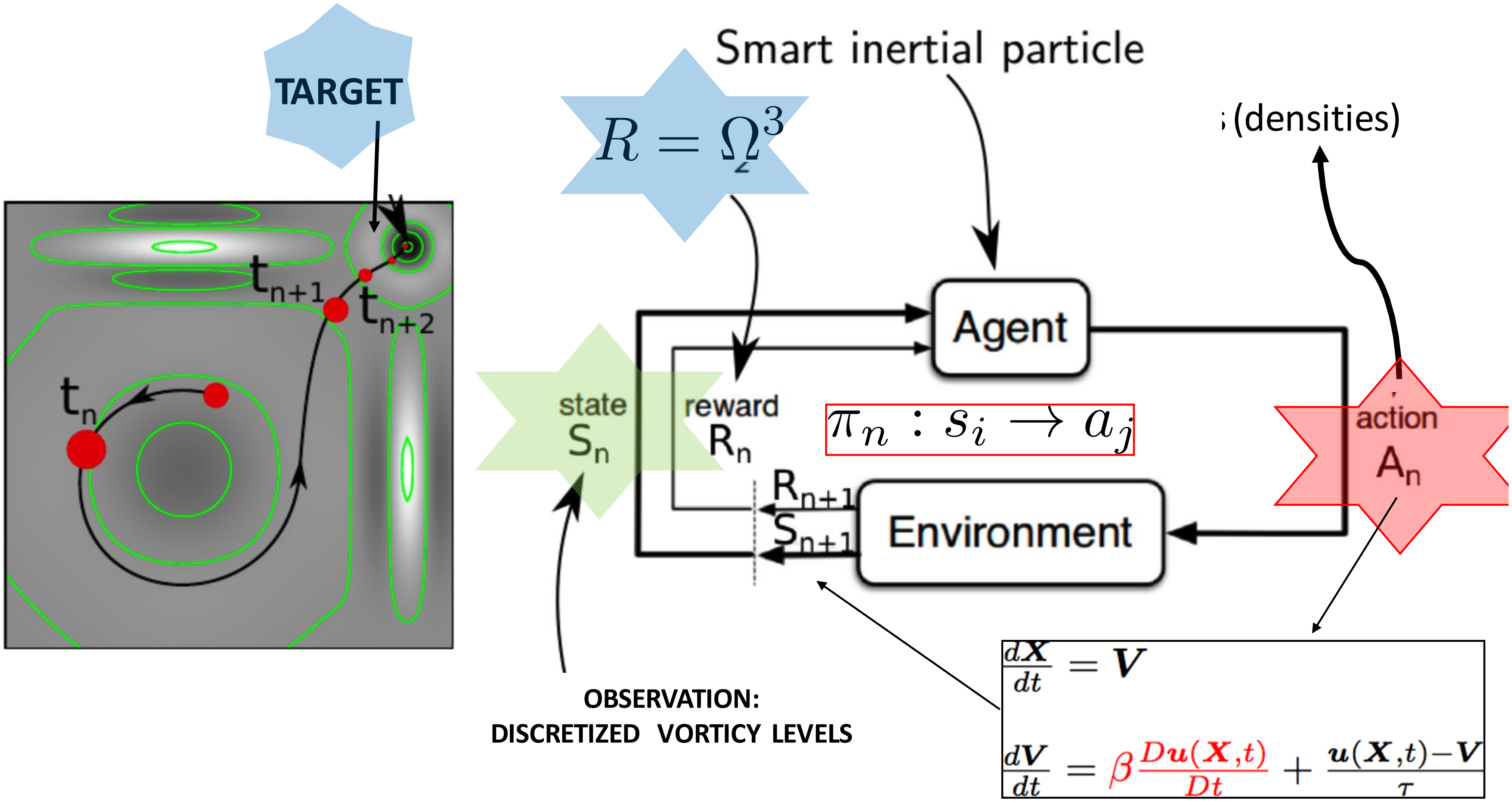




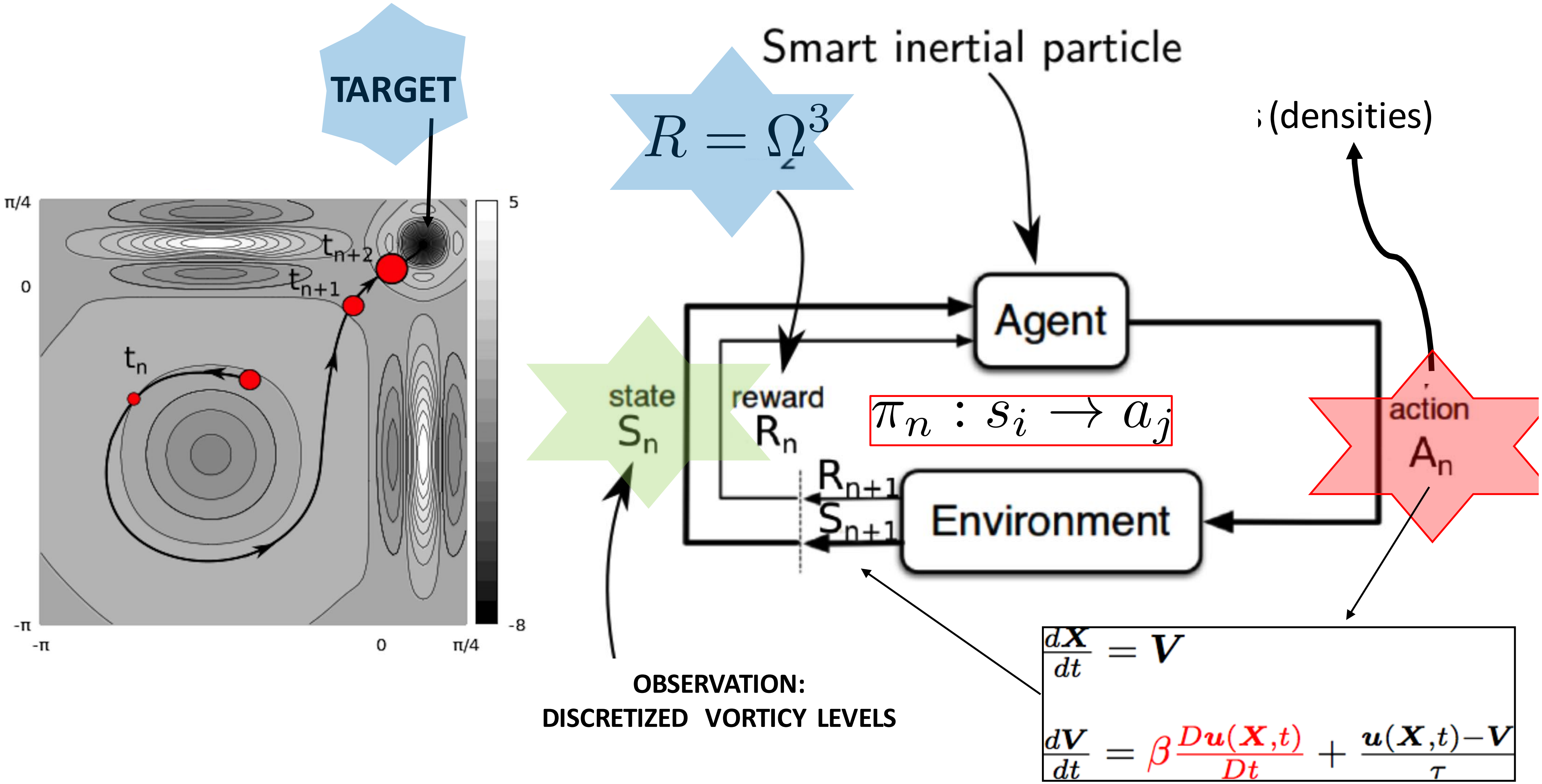


Lagrangian properties of particles in turbulence
F Toschi, E Bodenschatz
Annual Review of Fluid Mechanics 41, 375-404 (2009)

Coherent structures and extreme events in rotating multiphase turbulent flows L.
L.B., F Bonaccorso, IM Mazzitelli, MAT van Hinsberg, AS Lanotte, ...
Physical Review X 6 (4), 041036 (2016)



$$\pi_n \to \pi_{n+1} \to \cdots \pi_{opt}$$



$$\pi_n \to \pi_{n+1} \to \cdots \pi_{opt}$$

TRAINING: Q-LEARNING ALGORITHM

QUALITY MATRIX AT STEP n → $Q_n(a_j, s_i)$

EXPECTED DISCOUNTED FUTURE RETURN IF ACTION **a_j** is taken after observation of state **s_i**

$$\boxed{Q_n(s_i, a_j) = R_n + \gamma R_{n+1} + \gamma^2 R_{n+2} + \gamma^3 R_{n+3} + \cdots = \sum_{t=n}^{\infty} \gamma^t R_t}$$

MYOPIC→ $\gamma = 0$

FAR-SIGHTED→ $\gamma = 1$

GREEDY POLICY AT STEP n:

$$\boxed{\pi_n : a = arg \max_{a'} Q_n(a', s)}$$

$$\begin{array}{cc} s_1 & \\ s_2 & \left[\begin{array}{ccc} 1.2 & 0.3 & 0.1 \\ 2.2 & 4.3 & 10.1 \\ 2.0 & 8.1 & 2.0 \end{array}\right] \\ s_3 & \\ & \begin{array}{ccc} a_1 & a_2 & a_3 \end{array} \end{array}$$

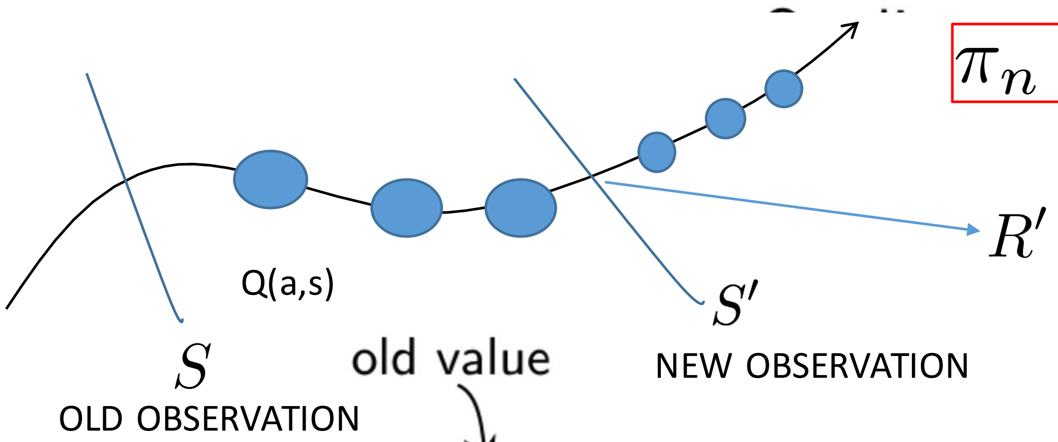
$$s_1 \xrightarrow{\pi_n} a_1$$

$$s_2 \rightarrow a_3$$

$$s_3 \rightarrow a_2$$

Learning through experience: Q-learning algorithm

$$\boxed{\pi_n \to \pi_{n+1}\ ?}$$



old value

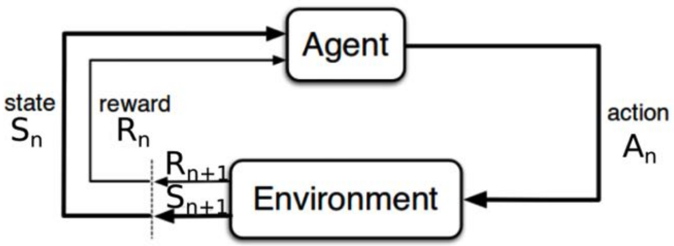
$$Q(s,a) \leftarrow Q(s,a) + \alpha[R' + \gamma \max_{a'} Q(s',a') - Q(s,a)]$$

learning rate

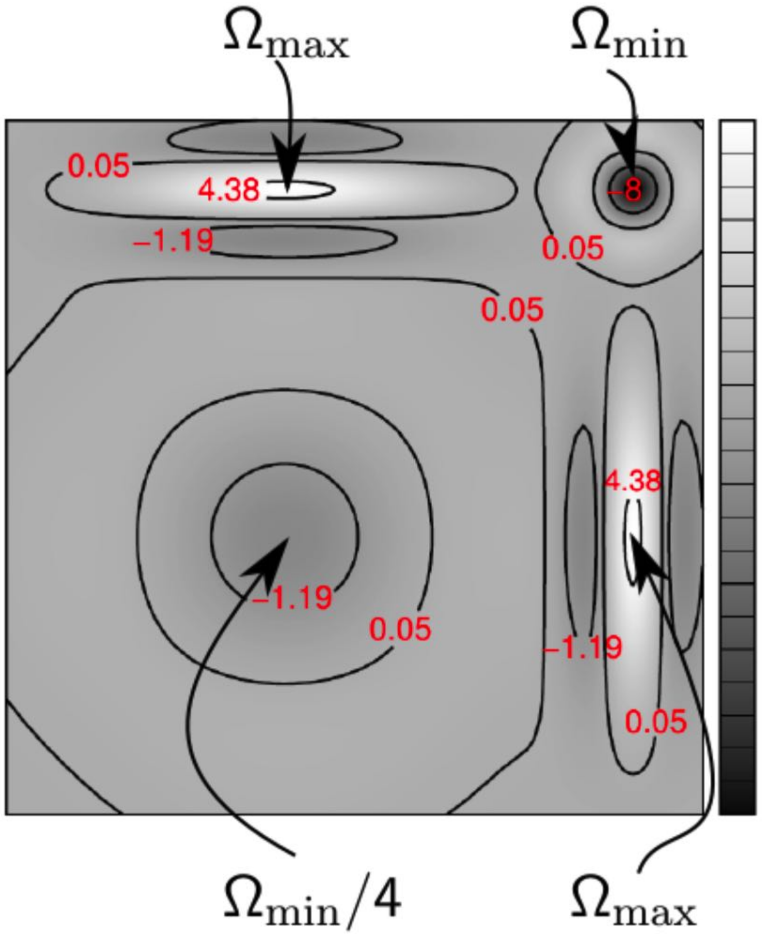
$0 \le \alpha < 1$ fixed or time dependent

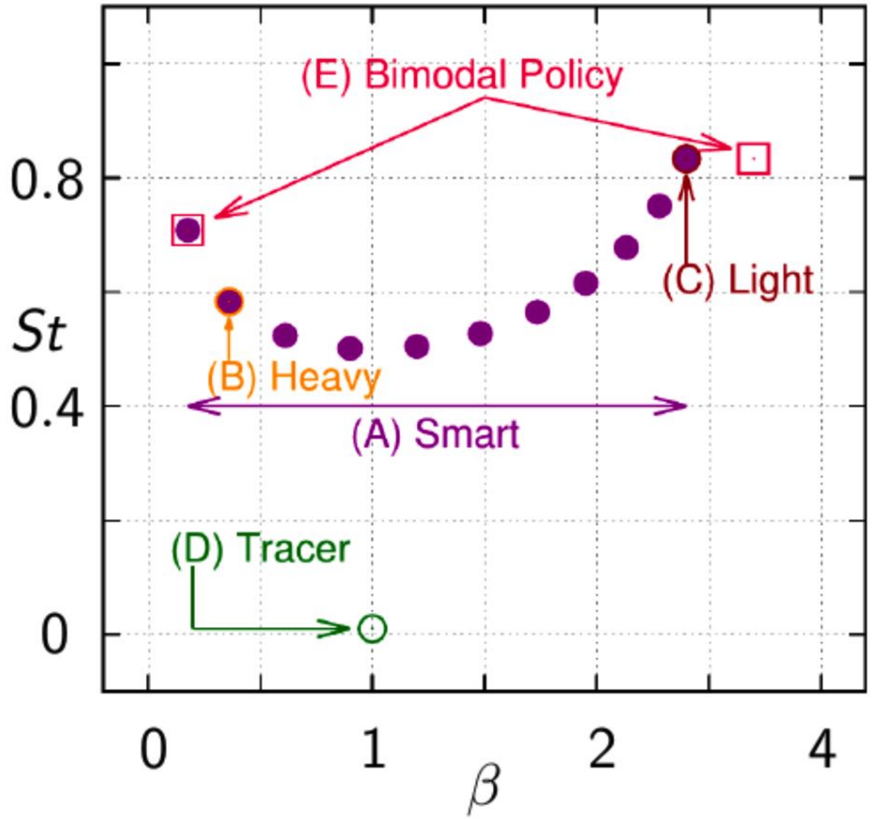
discount factor

$0 \le \gamma < 1$



CHANGING RADIUS b_n





$Ns = 21$

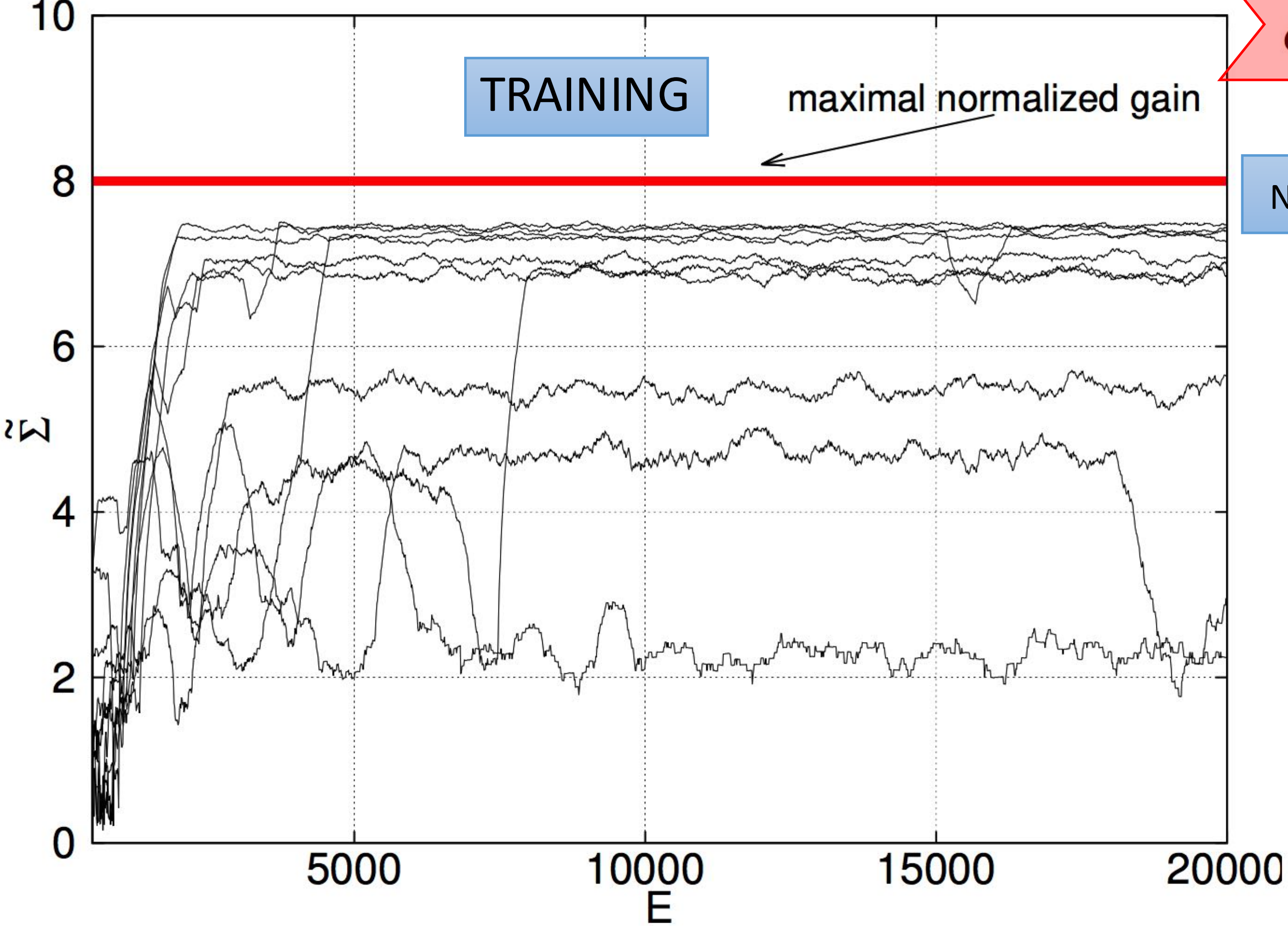
$Na = 11$

$\beta \to \beta(b_n)$

$St \to St(b_n)$

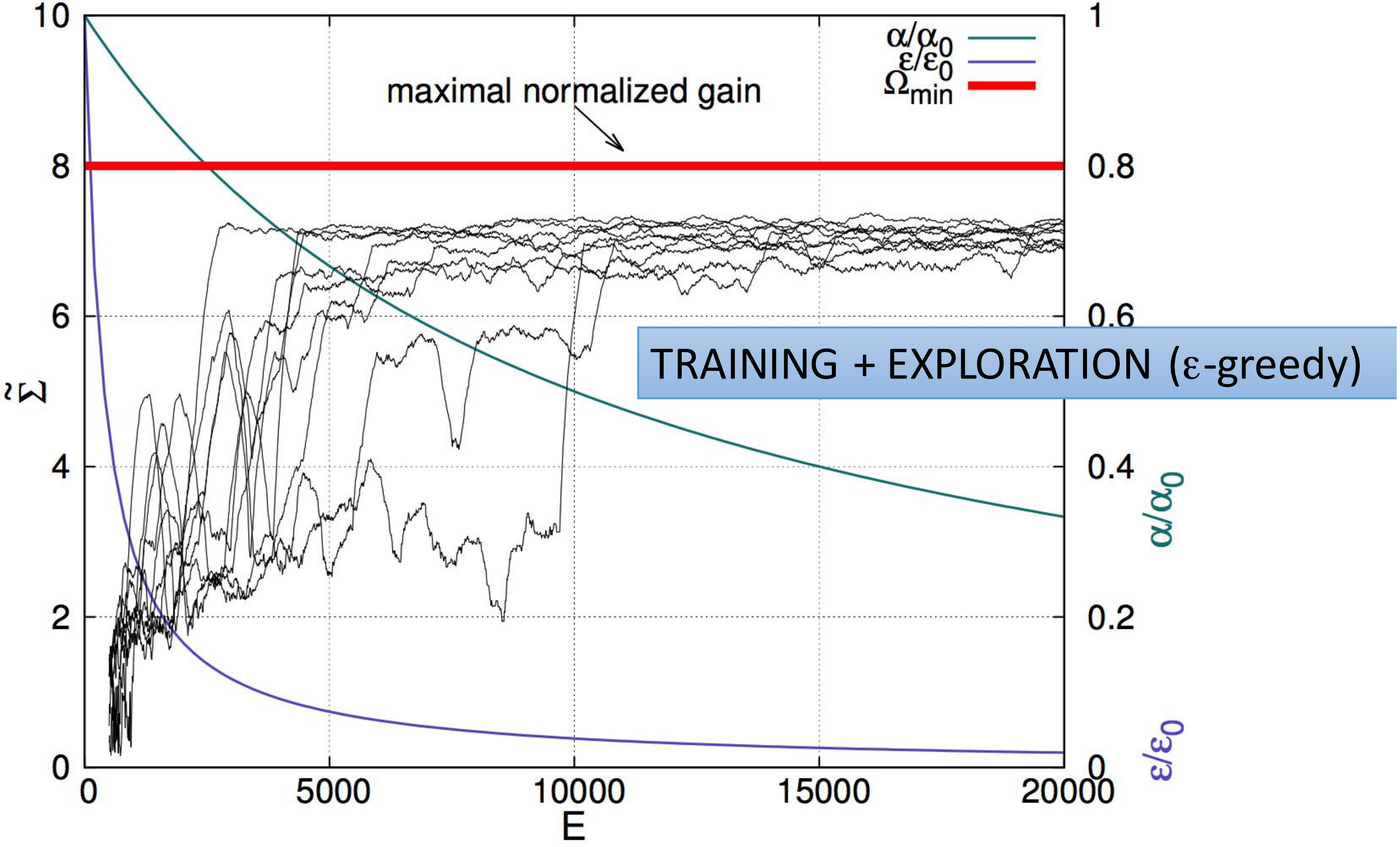
Comparison with naive behaviors

Learning gain $\tilde{\Sigma}(E) = \sqrt[3]{\frac{\sum_{n=1}^{N} R_n}{N}}$



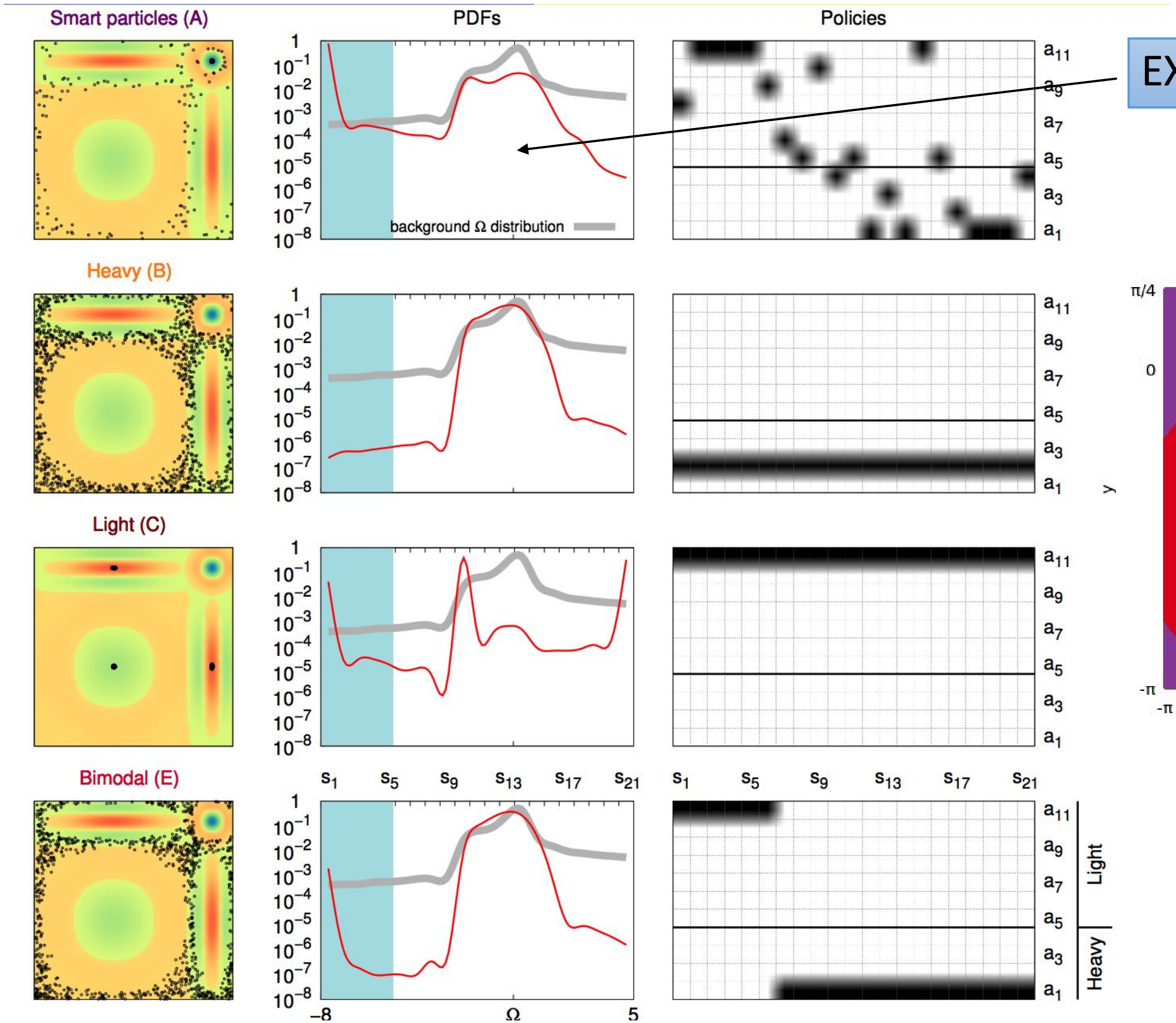
$\epsilon = 0$

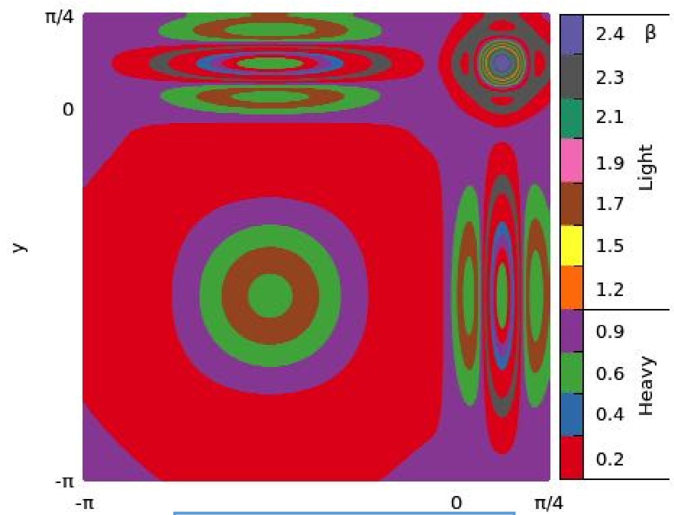
NO EXPLORATION



maximal normalized gain
α/α_0
$\varepsilon/\varepsilon_0$
Ω_{min}
TRAINING + EXPLORATION (ε-greedy)
$\tilde{\Sigma}$
E
0
5000
10000
15000
20000
0
2
4
6
8
10
0
0.2
0.4
0.6
0.8
1

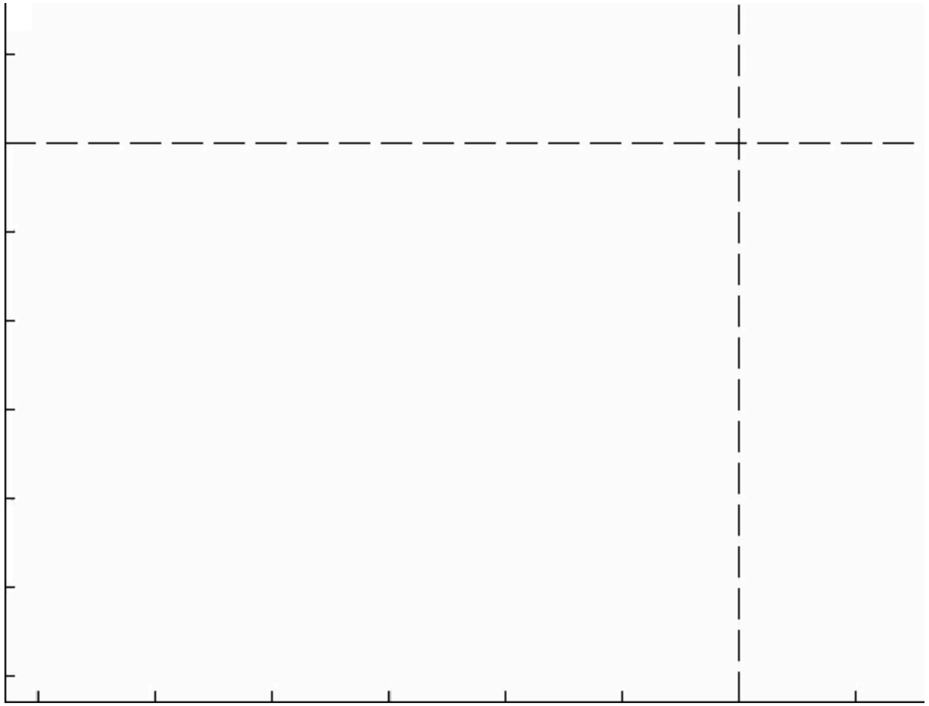
EXAM



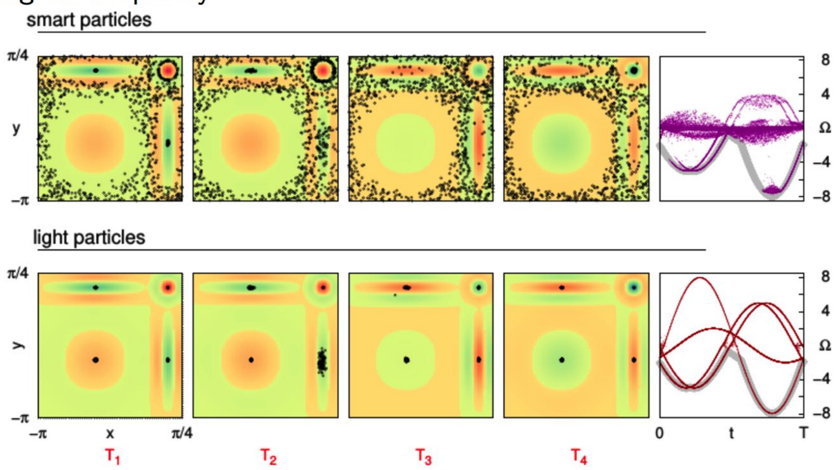


OPTIMAL ACTIONS

TIME DEPENDENT FLOW



Is the algorithm robust for more complex flows?
Steady flow $\rightarrow$ 4 oscillating vortex flow with different phases but same angular frequency

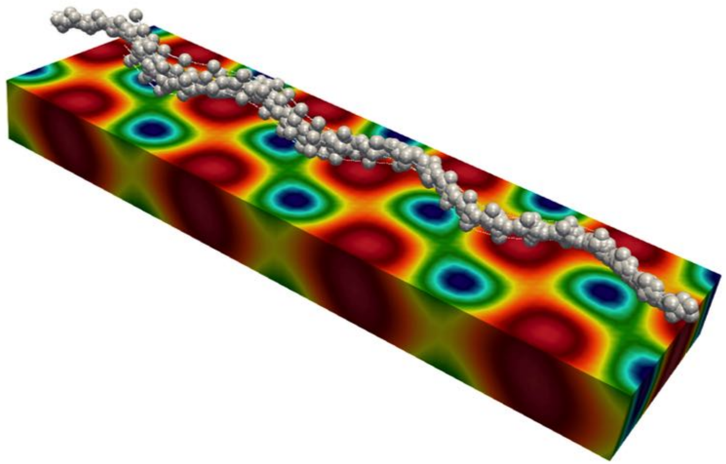


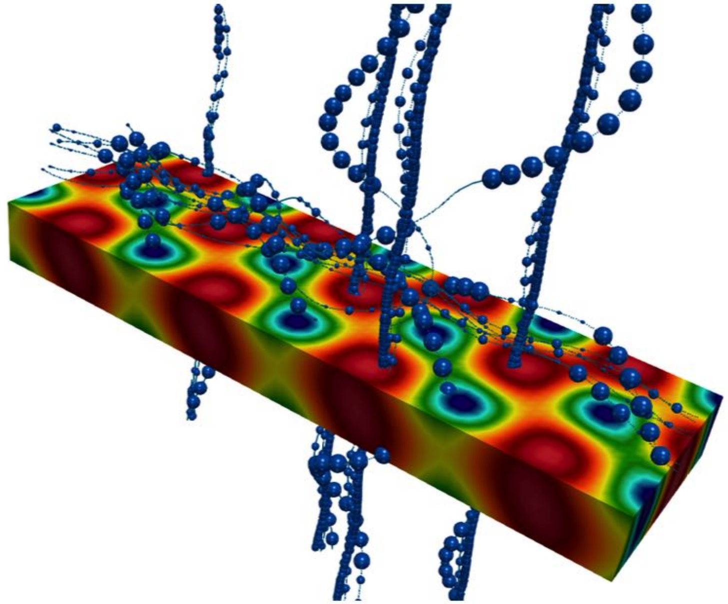
Asymmetric ABC flow

$\boldsymbol{u}(\boldsymbol{x}) = (C\cos y + A\sin z, A\cos z + B\sin x, B\cos x + C\sin y)$ [4A=2B=C=1]

Task: Optimize long-term vorticity $|\boldsymbol{\Omega}|$ by perception of Ω_z or Ω_x

-different setup: *St* fixed



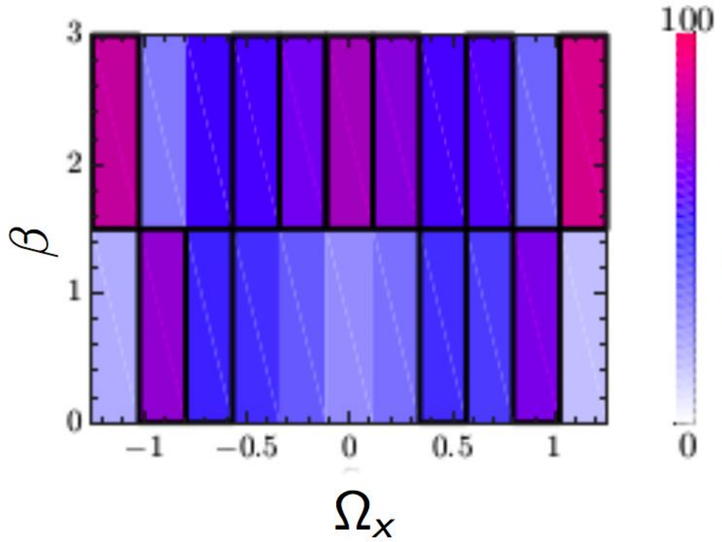


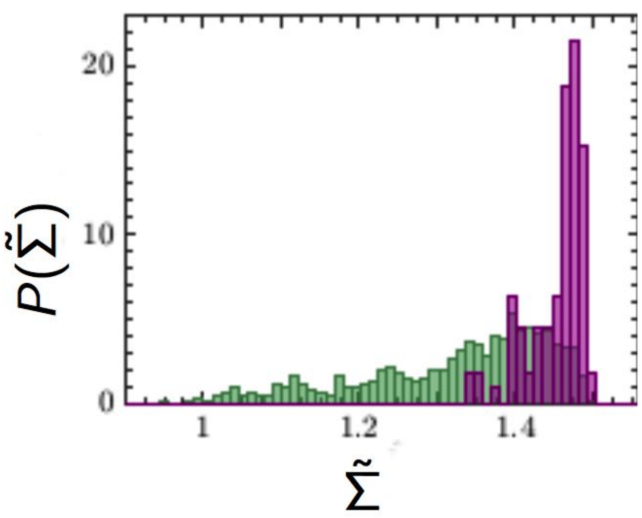
Light particles distribute on minor vortices

Smart particle learns to target principal vortices

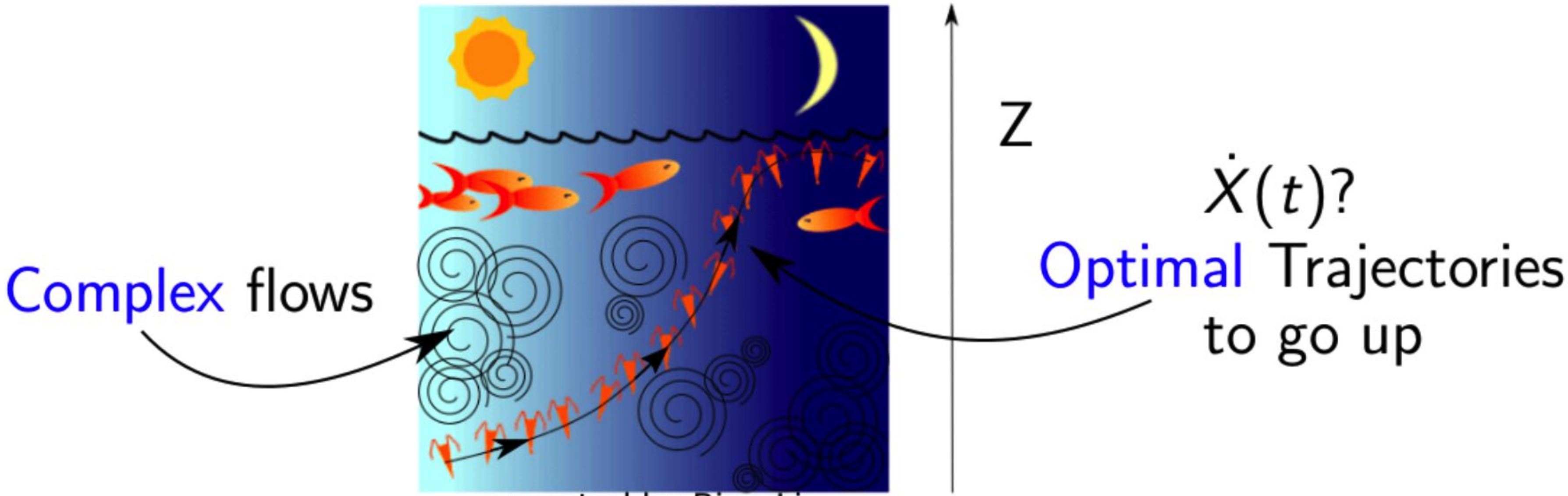
Evaluation of the algorithm

The global optimal policy is not known
Brute force approach by testing all possible Q-matrices [$2^{11} = 2048$]
(reducing the number of states and actions)





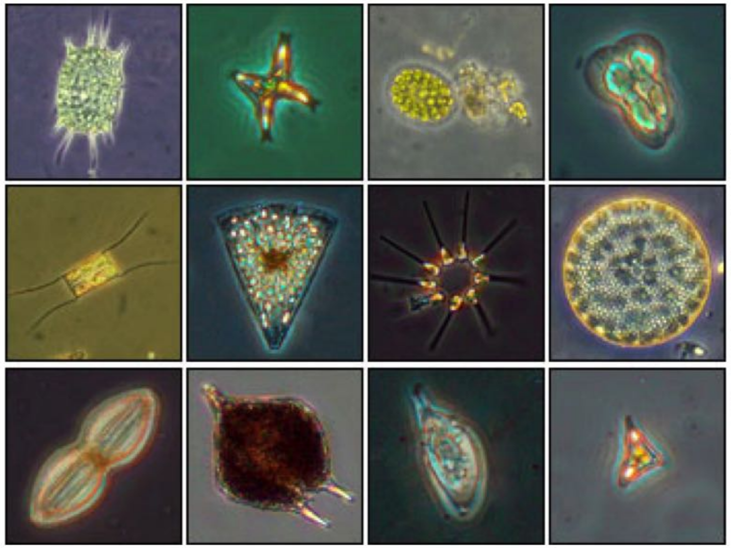
Microswimmers can bias their motion in order to achieve a biologically relevant goal



created by Bing Li

Particles can learn by experience to navigate by using Reinforcement Learning

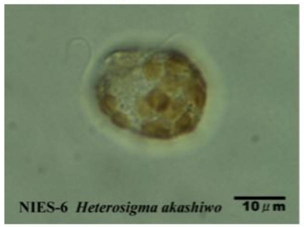
Case study: Gyrotactic Phytoplankton



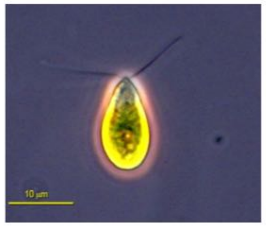
- large diversity of forms
- primary producers in oceans
- ≃50% photosynthetic activity on Earth
- up to 10^4 per milliliter of water
- at the bottom of marine food web
- can form Harmful (toxic) Algal Bloom
- patchiness at different scales
- many species are able to swim, e.g. 90% toxic algae are able to swim

gyrotactic microalgae

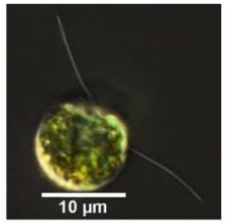
Heterosigma akashiwo



Dunaliella tertiolecta



Chlamydomonas reinhardtii



Advecting flow

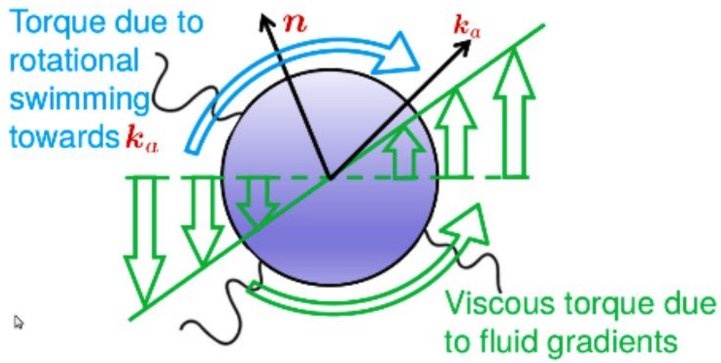
Swimming intensity

Swimming direction

- $\dot{\boldsymbol{x}} = \boldsymbol{u} + \Phi\, \boldsymbol{n}$
- $\dot{\boldsymbol{n}} = \underbrace{\frac{1}{2\Psi}[\boldsymbol{k}_a - (\boldsymbol{k}_a \cdot \boldsymbol{n})\boldsymbol{n}]}_{\text{Particle torque}} + \underbrace{\frac{1}{2}\boldsymbol{\omega} \times \boldsymbol{n}}_{\text{Flow torque}}$

Preferred direction

Strenght of the flow torque



Free Particle parameters

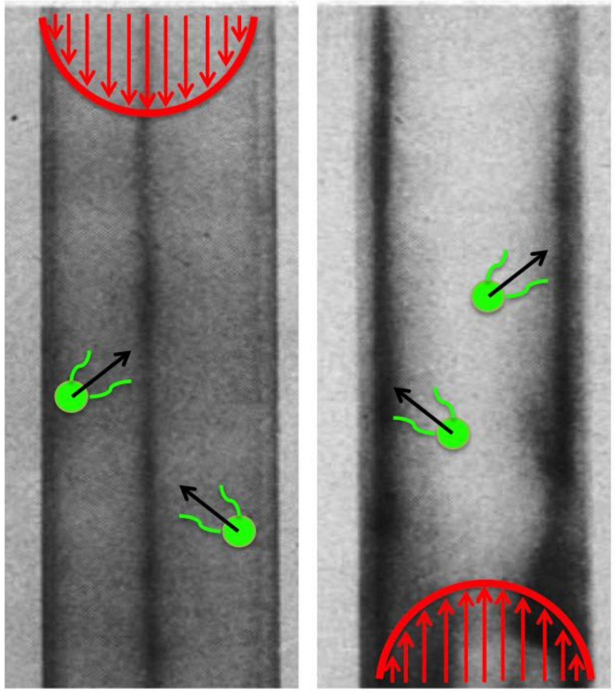
If
$\Phi \to 0,\ \Psi \to \infty$
passive swimmer

Flow properties

Direction to be modified

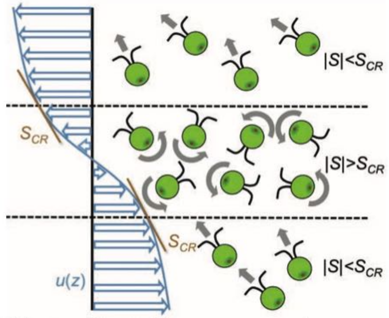
depending on ω

Gyrotactic focusing

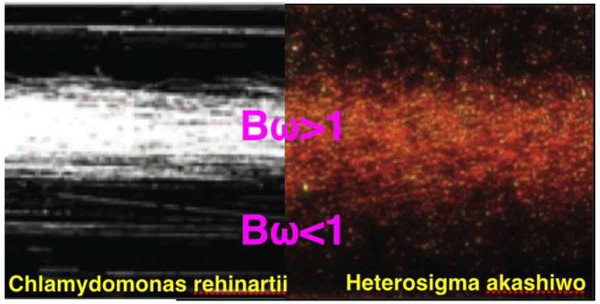


J.O. Kessler Nature (1985)

Gyrotactic trapping



Microfluidic experiments

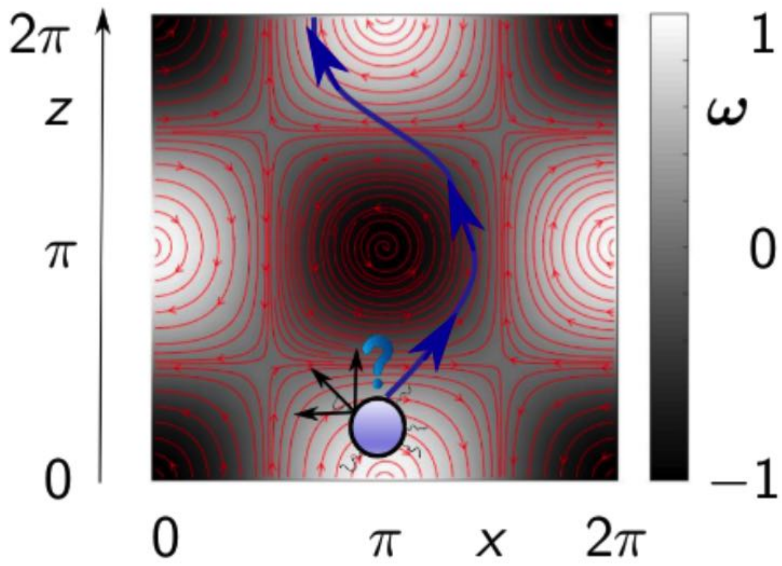


Durham et al Science 2009

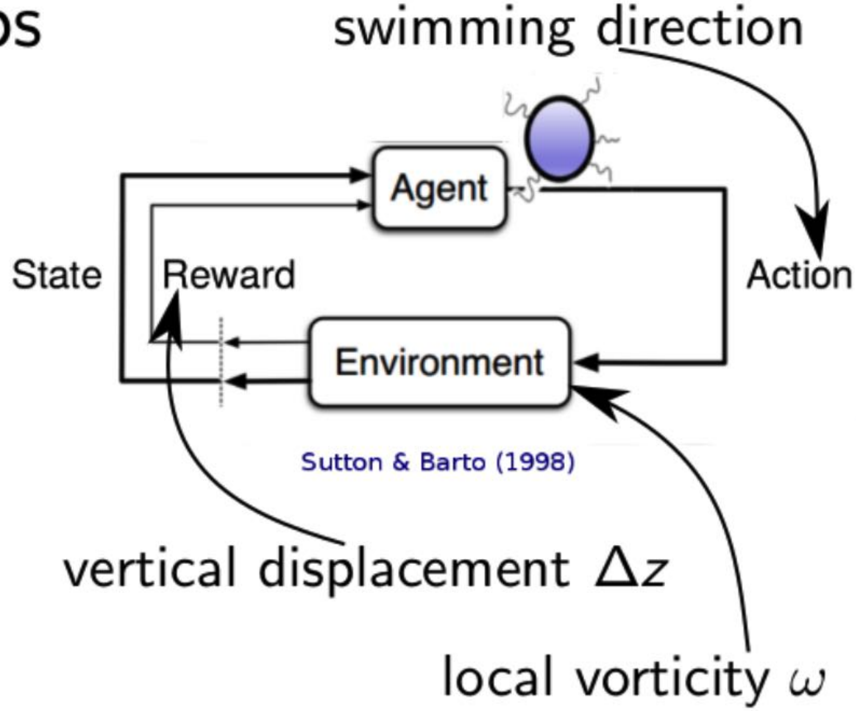
Clustering and turbophoresis in a shear flow without walls
F De Lillo, M Cencini, S Musacchio, G Boffetta
Physics of Fluids 28 (3), 035104 (2016)

Goal: going up as efficient as possible
Difficulty: Hydrodynamic traps

How to choose swimming direction given the underlying vorticity?



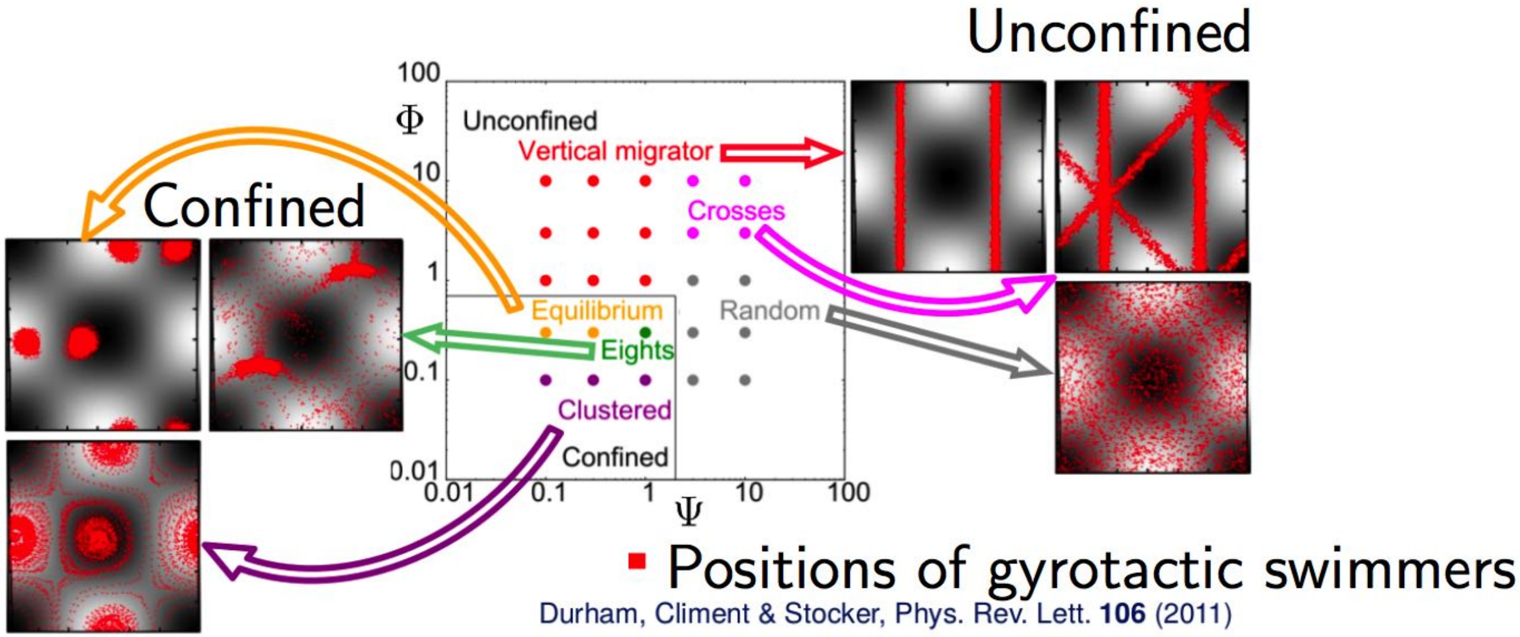
Taylor Green Vortex Flow



Reinforcement Learning

If k_a is fixed, the particle can be trapped in confined region

$$k_a = \hat{z} \uparrow$$

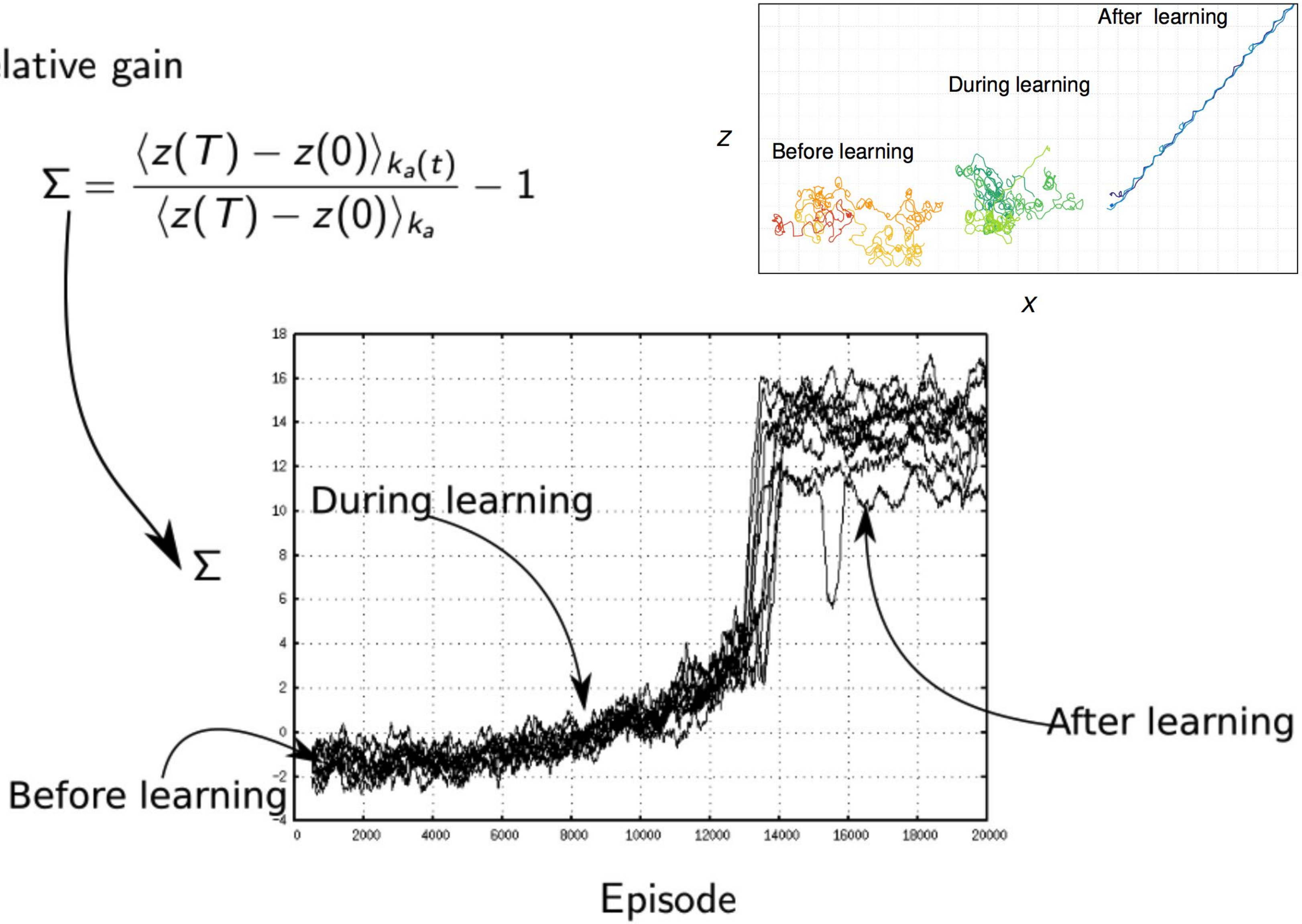


Durham, Climent & Stocker, Phys. Rev. Lett. **106** (2011)

We have to find the optimal $k_a \to k_a(t)$ to swim upward

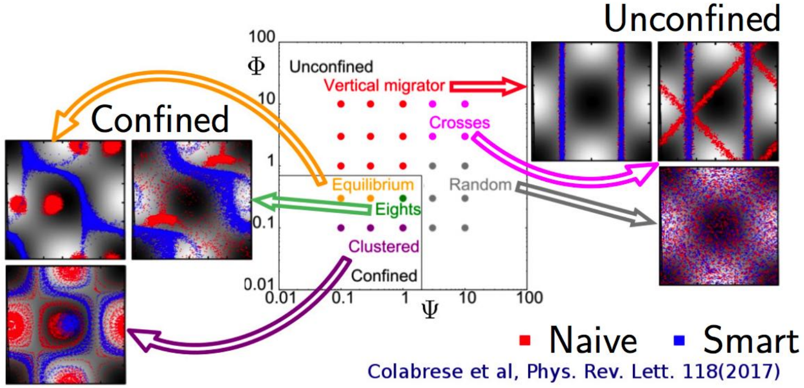
Relative gain

$$\Sigma = \frac{\langle z(T) - z(0) \rangle_{k_a(t)}}{\langle z(T) - z(0) \rangle_{k_a}} - 1$$



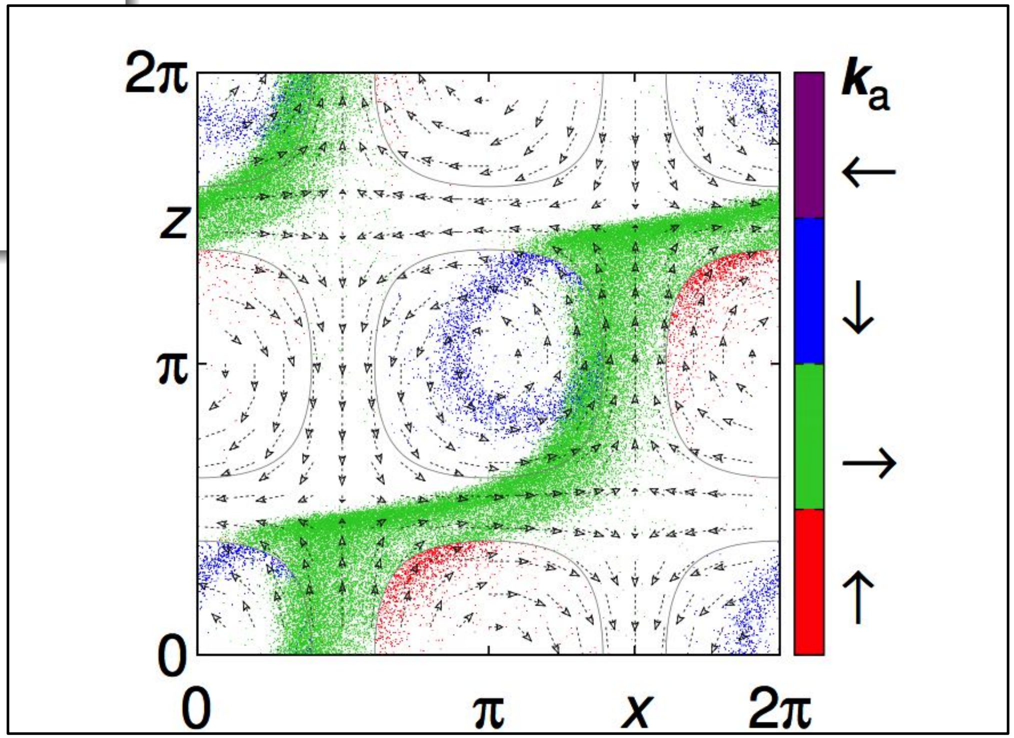
Trajectories after RL training

Comparison between *smart* and *naive* microswimmers trajectories for different parameters values



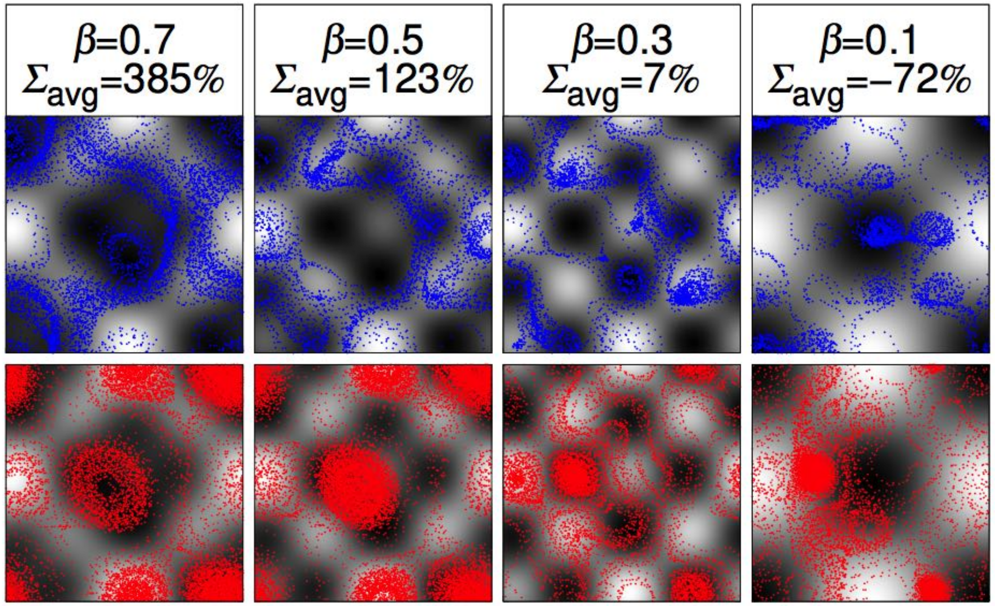
Colabrese et al, Phys. Rev. Lett. 118(2017)

OPTIMAL STRATEGY



Robustness to perturbation of the flow

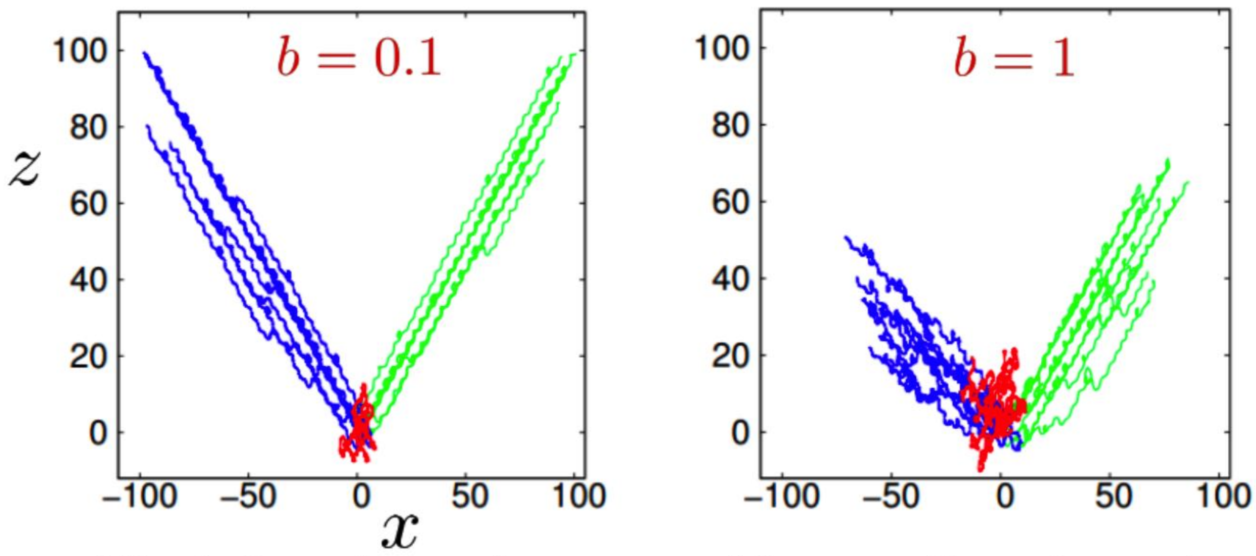
Using strategies found for the unperturbed flow ($\beta = 1$) the particle outperform the naive particle down to moderate β



Time-dependent flow

Robustness even when tracers might have chaotic evolution

Taylor-Green flow with periodic phase shift $\Delta_t = b\cos(u_0 t)$
$\boldsymbol{u} = \frac{u_0}{2}[-\cos(x-\Delta_t)\sin(z-\Delta_t)\hat{\boldsymbol{x}} + \sin(x-\Delta_t)\cos(z-\Delta_t)\hat{\boldsymbol{z}}]$



Weak time dependence *Stronger time dependence*

Strategy learns in the time-dependent flow

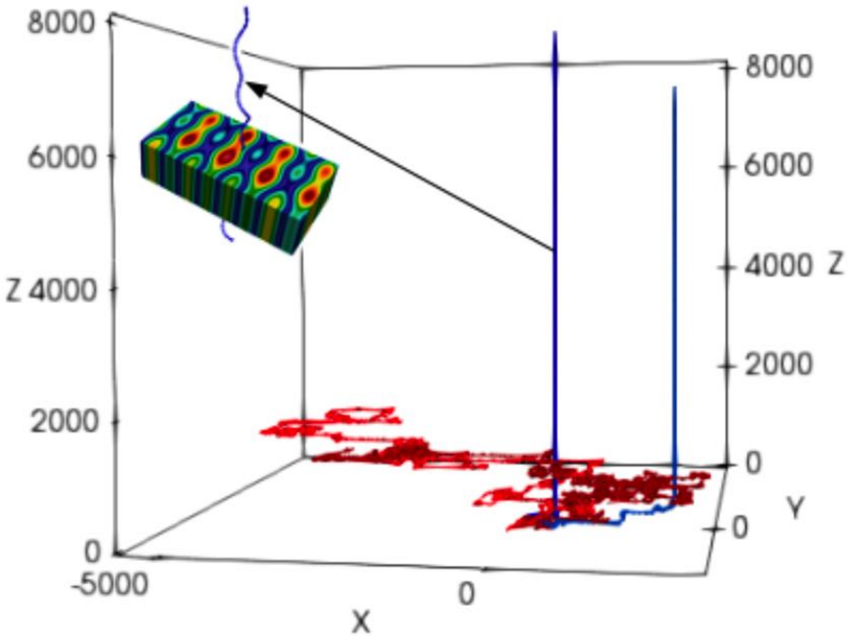
Going up as efficient as possible in ABC flow

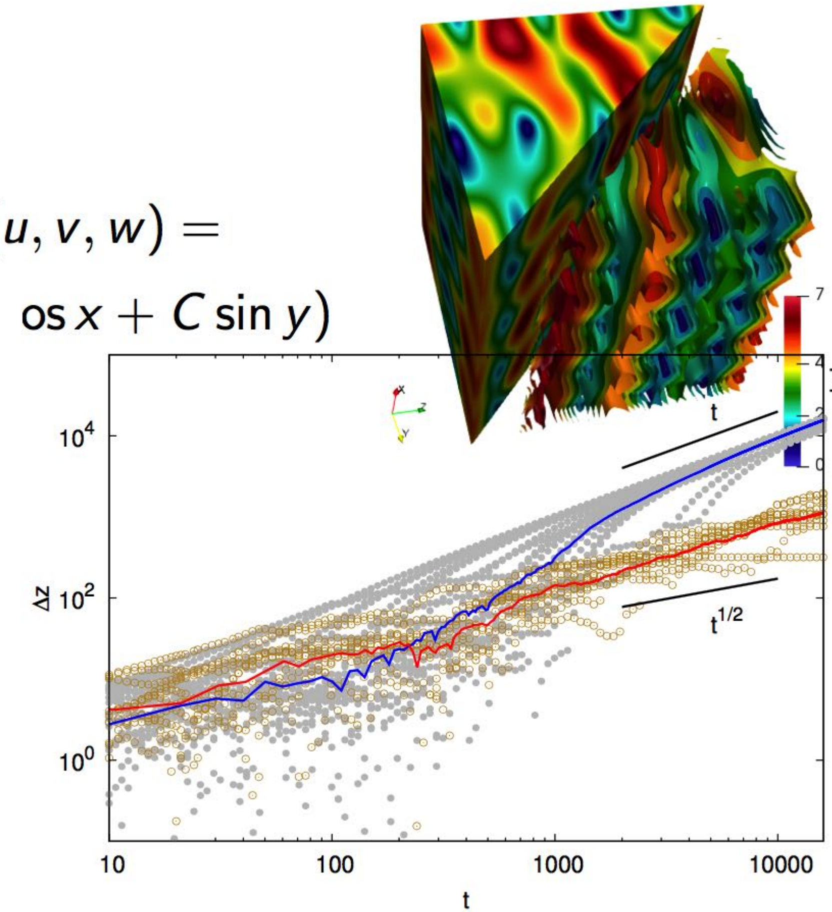
Reddy, G., Celani, A., Sejnowski, T. J., & Vergassola, M. (2016). Learning to soar in turbulent environments. ***Proceedings of the National Academy of Sciences*****, 201606075.**

Comparison between smart and naive microswimmers in a lagrangianly chaotic flow

$(u, v, w) =$

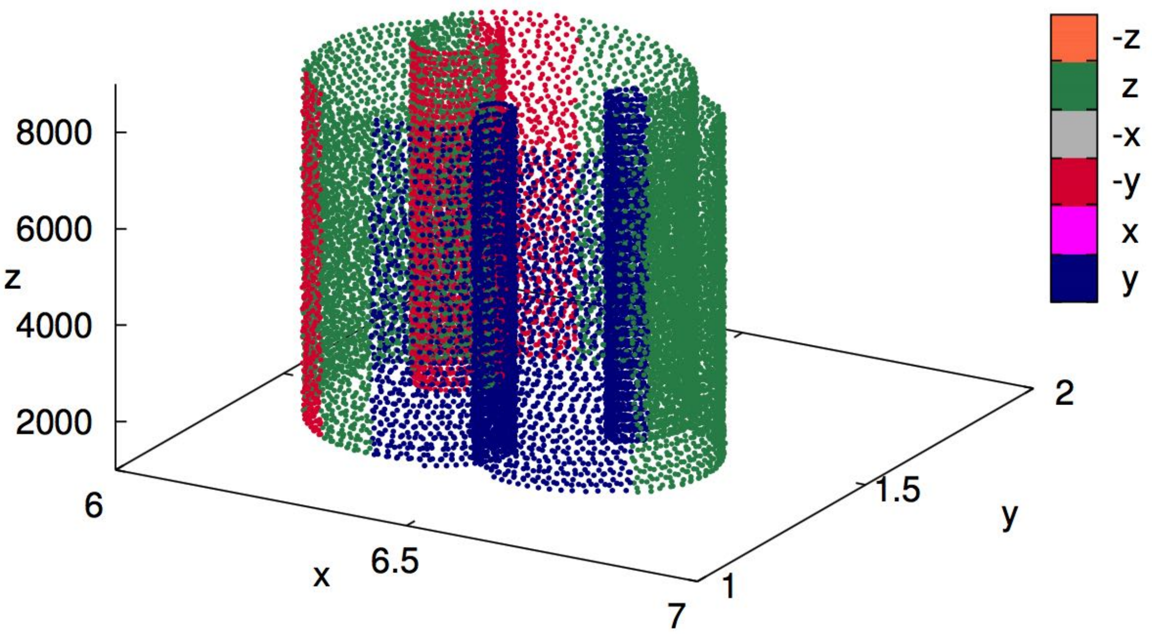
$\text{os}\, x + C \sin y)$





Example of Resulting Strategy

Optimal actions taken during the ascent through an elevator



Conclusions

CREDITS: SIMONA COLABRESE (TOR VERGATA UNIV. ROME-IT); ANTONIO CELANI (ICTP TRIESTE-IT); KRISTIAN GUSTAVSSON (GOTHEBORG UNIV. SWEDEN)

- Reinforcement learning approach is successful to develop strategies for inertial particles to sample specific vortical structures and for microswimmers to ascend two- and three-dimensional vortex flows
- Learnt strategies are robust to spatial or time perturbations of the flow

- **Flow navigation by smart microswimmers via reinforcement learning**
S Colabrese, K Gustavsson, A Celani, L Biferale
Physical Review Letters 118 (15), 158004

-**Smart Inertial Particles**
S Colabrese, K Gustavsson, A Celani, L Biferale
arXiv preprint arXiv:1711.05853

- **Finding efficient swimming strategies in a three-dimensional chaotic flow by reinforcement learning**
K Gustavsson, L Biferale, A Celani, S Colabrese
The European Physical Journal E 40 (12), 110