Learning to navigate amid uncertainties

UCSD
Physics

Massimo Vergassola

A. Celani
(ICTP)

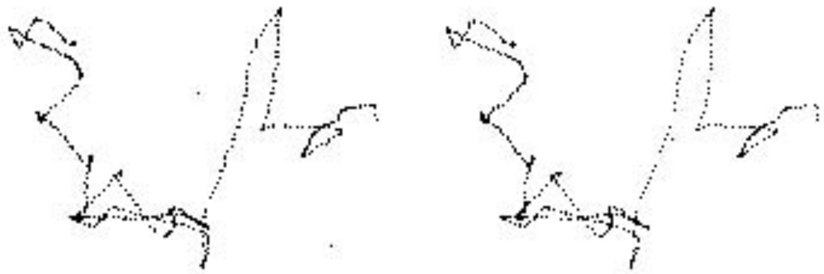
G. Reddy (UCSD)

T. Sejnowski
(Salk)

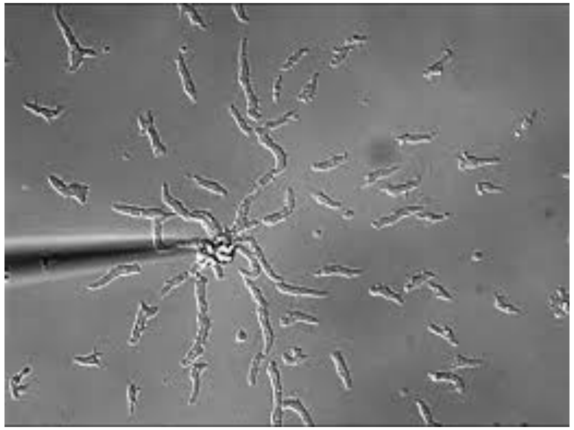
J. Wong-Ng
(UCSD)

Navigation in uncertain environments

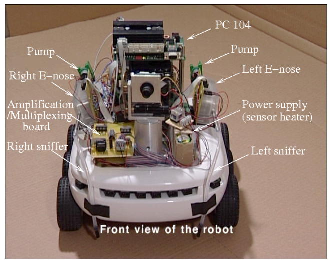
H. Berg's lab image



Bacterial chemotaxis

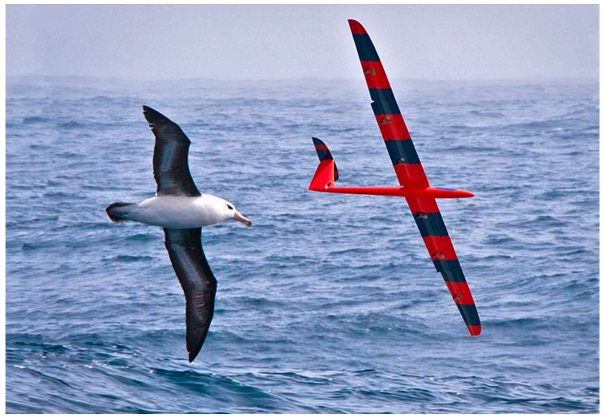


Eukaryotic chemotaxis

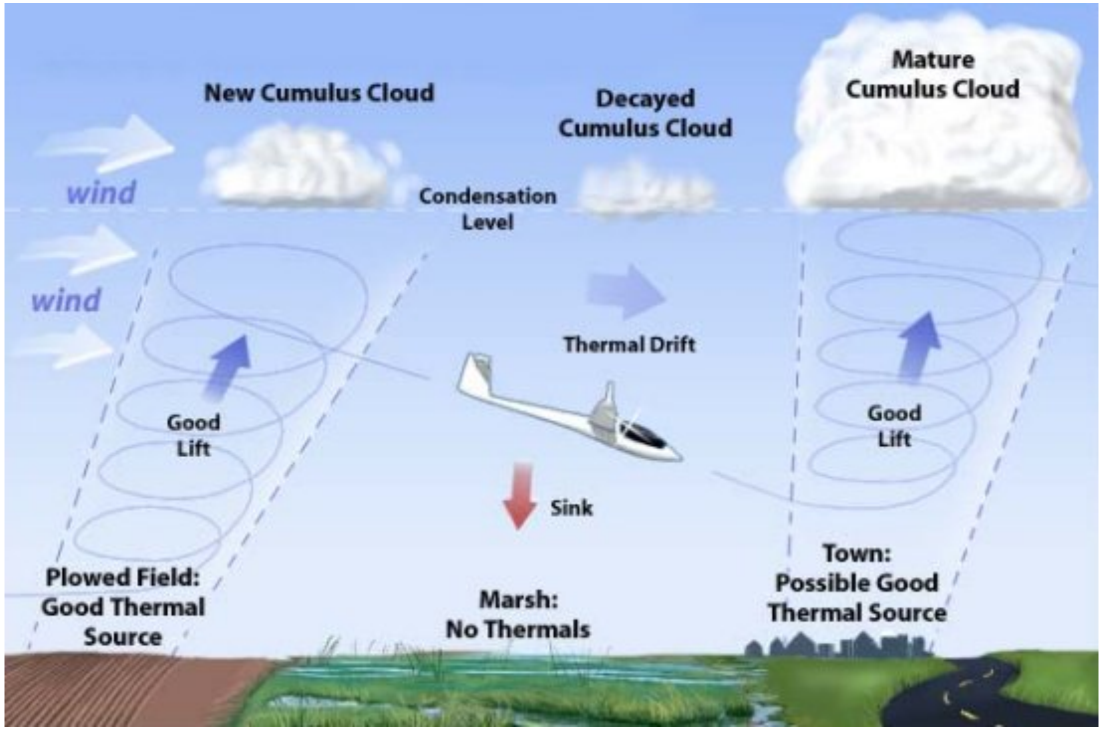


Olfactory navigation by insects and sniffers

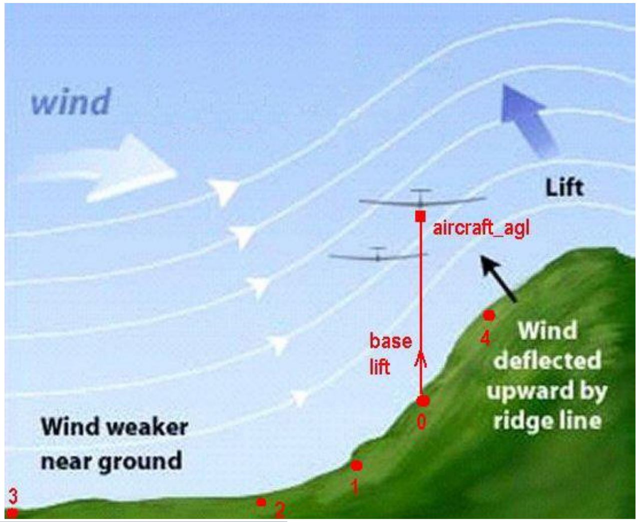
Olfactory navigation by rodents



Soaring by birds and gliders

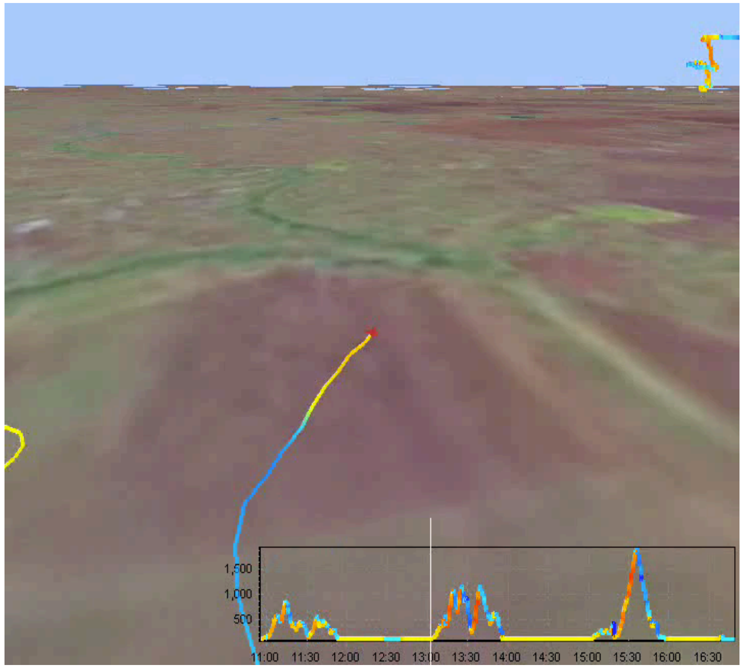


Thermal soaring by flying vehicles

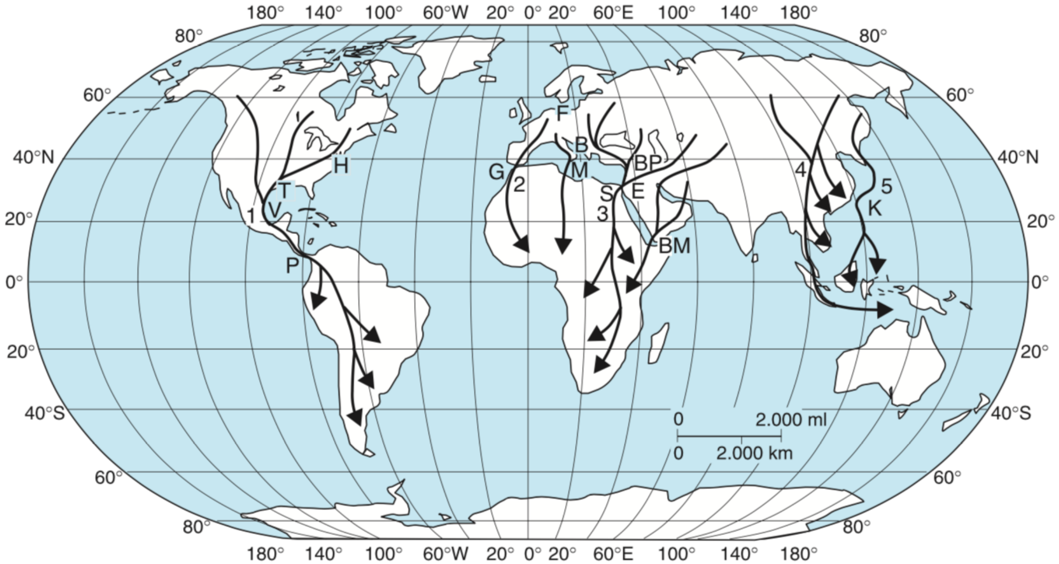


Ridge soaring

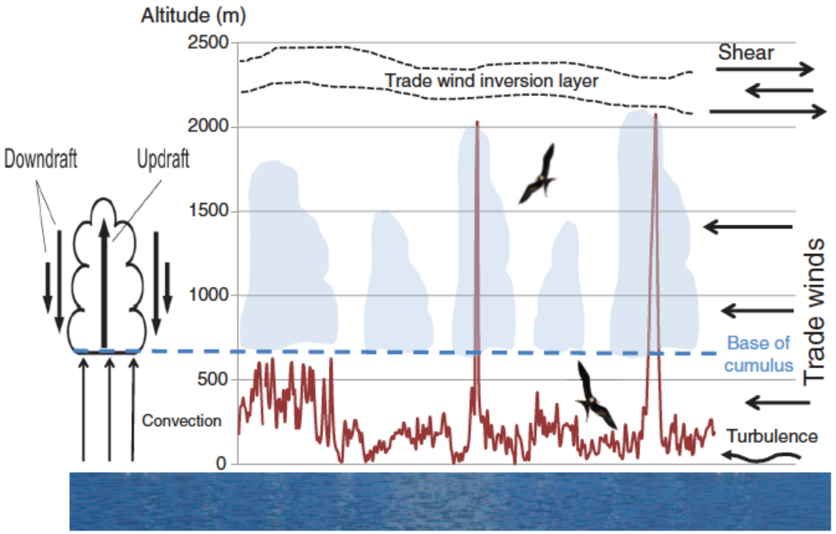
Soaring in birds



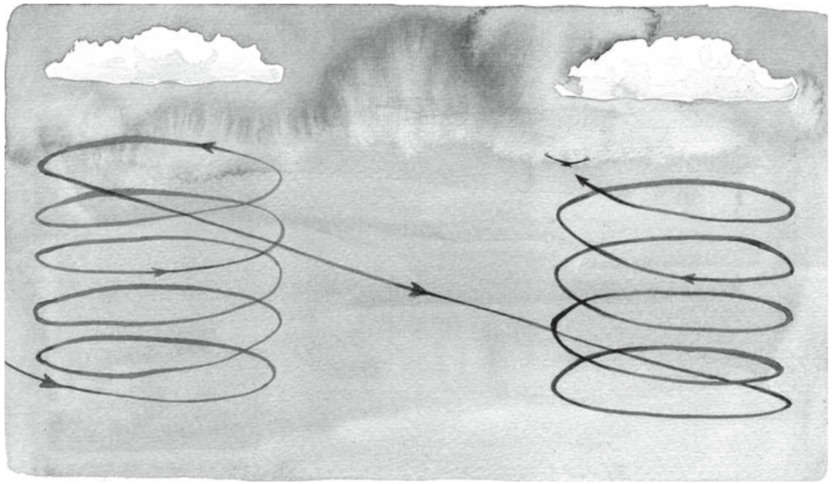
Akos, Nagy, Vicsek, PNAS, ‘08



Migration Ecology of Birds, Ian Newton



Weimerskirch et al. Science ‘16



Soaring energetics

$$\dot{E} = \mathrm{v}\,(T - D) + mg\mathbf{w}_z - m\mathbf{v}.\frac{d\mathbf{w}}{dt}$$

Steppe eagle

If flapping: $T = D \approx \frac{1}{2}\rho S C_D \mathrm{v}^2$

For migration of a 1000km:

$$\Delta E \approx 5000 \text{ kcal}$$

$$S \approx 0.5\mathrm{m}^2$$
$$C_D \approx 0.3$$
$$\mathrm{v} \approx 15 \text{ m/s}$$
$$m \approx 2 \text{ kg}$$

About 500 grams of fat
-> 25% of body mass

A bit of history…

$$\dot{E} = -\mathrm{v}D + mg\mathbf{w}_z - m\mathbf{v}.\frac{d\mathbf{w}}{dt}$$

Lord Rayleigh, 1883: "…Whenever therefore a bird pursues his course for some time without working his wings, we must conclude either

1. that the course is not horizontal,
2. that the wind is not horizontal, or
3. that the wind is not uniform.

It is probable that the truth is usually represented by (1) or (2); but the question I wish to raise is whether the cause suggested by (3) may not sometimes come into operation."

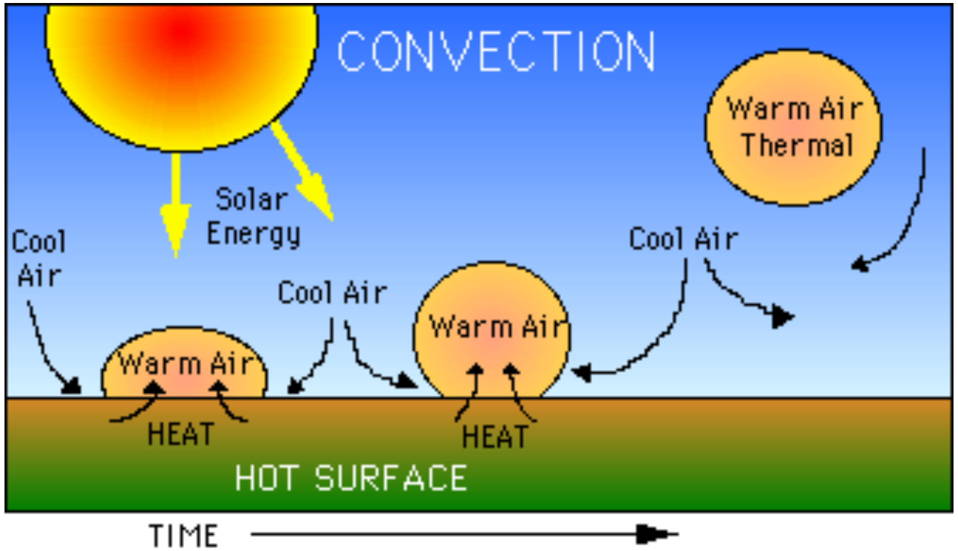
Thermal and ridge soaring

Dynamic and gust soaring

da Vinci, ca 1513-1515

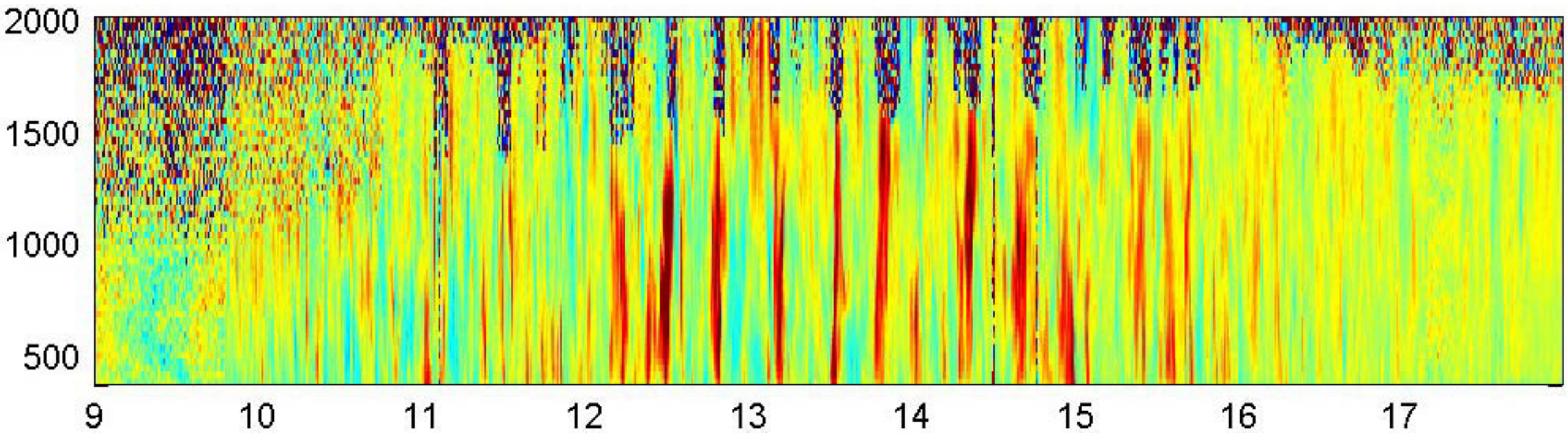
P. Richardson, "da Vinci's discovery of the dynamic soaring by birds in wind shear", 2018

Thermals in the atmospheric boundary layer

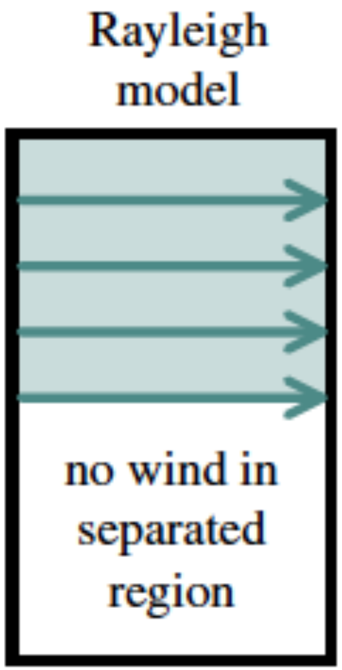


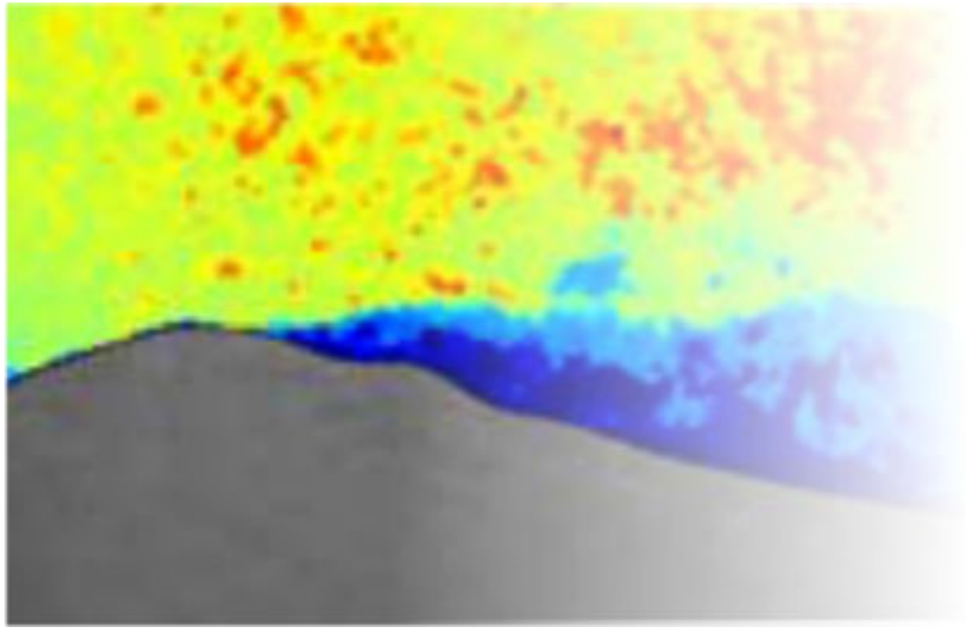
Basics of thermals:
the ground and low-lying layer of air are heated by the sun and tend to raise

The profile of vertical velocity *vs* hour of the day

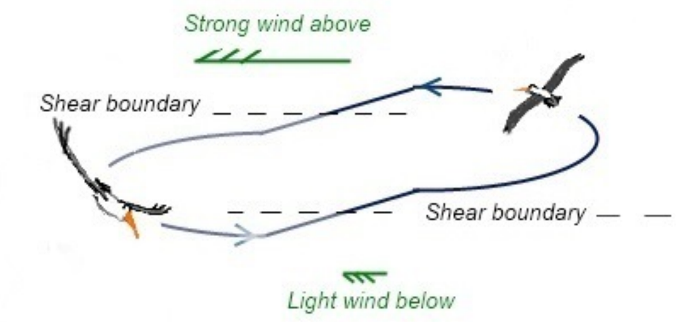


Dynamic soaring





Rayleigh Cycle



Rayleigh, Nature, 1883
Richardson, Progr. Oceanography, 2011
Bousquet et al, J. R. Soc. Interface, 2017

Questions

All these flows are highly fluctuating

What quantities should a bird sense?
vertical velocities, temperature, gradients, etc?

How should the bird respond to these cues?

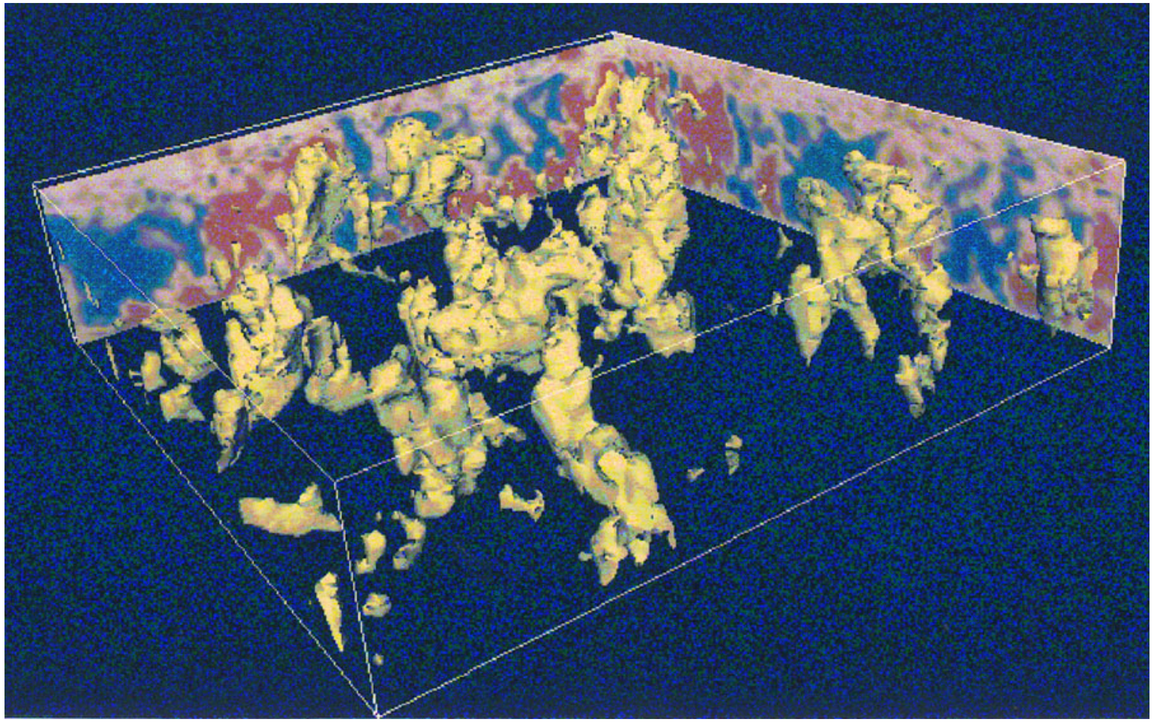
Experiments hard to control and strategies difficult to infer

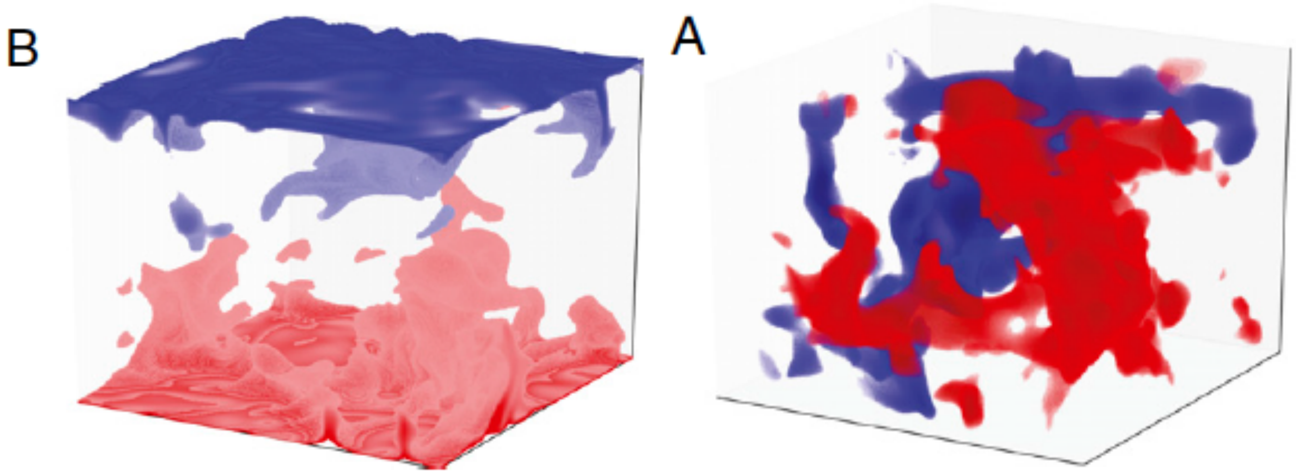
We will instead ask: how does an optimal agent navigate thermals? and what sensorimotor cues are most useful for the task?

(most focus on navigating a single thermal)

Model of turbulent flow

Instantaneous profiles are strongly turbulent

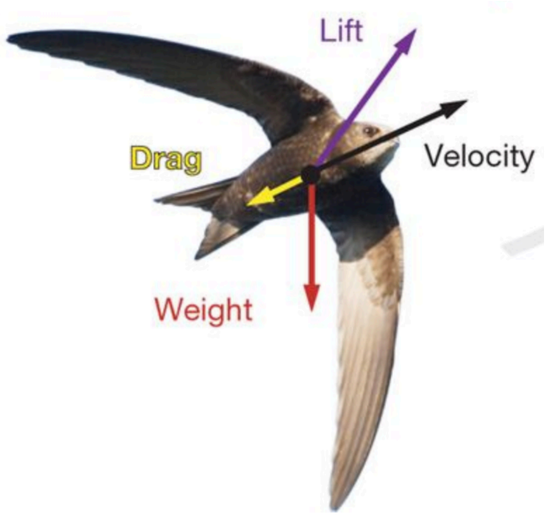


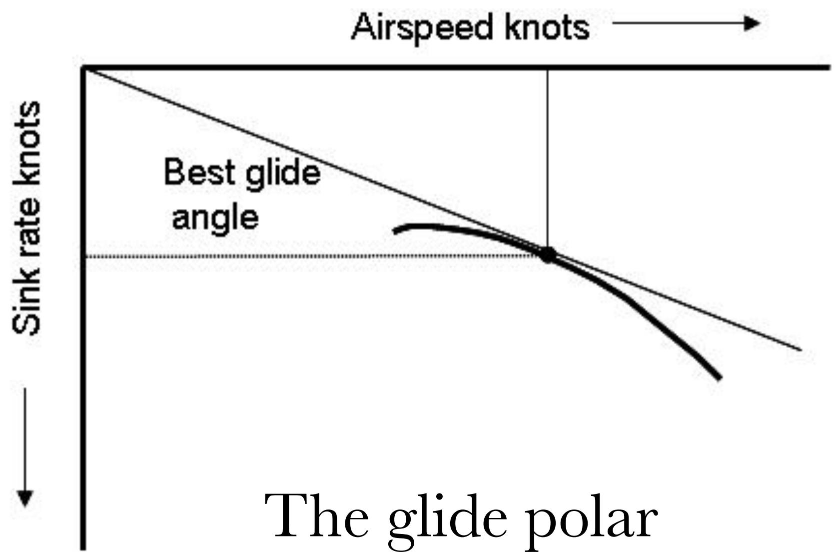


$$\frac{\partial u}{\partial t}+u\cdot\nabla u=-\nabla P+\left(\frac{\mathrm{Pr}}{\mathrm{Ra}}\right)^{1/2}\nabla^2 u+\theta\hat{z},$$

$$\frac{\partial\theta}{\partial t}+u\cdot\nabla\theta=\frac{1}{(\mathrm{Pr}\ \mathrm{Ra})^{1/2}}\nabla^2\theta,$$

Aerodynamics of a gliding object



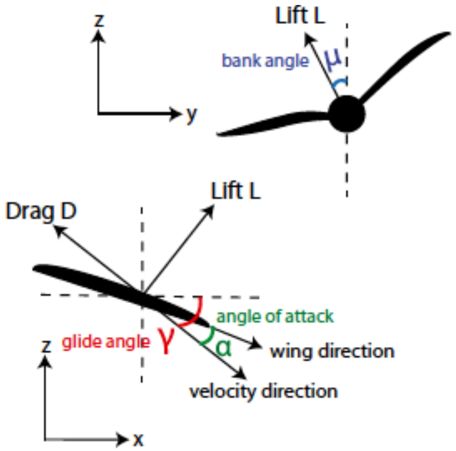


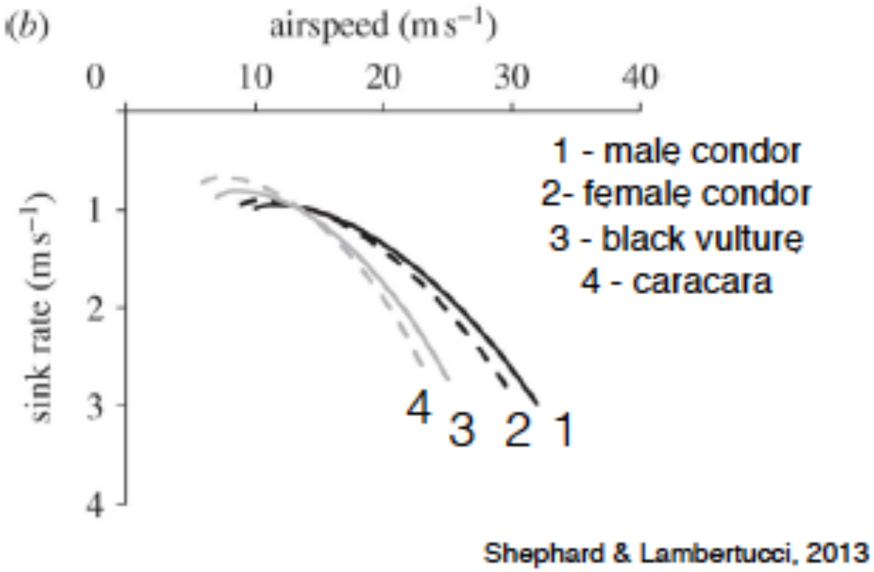
The glide polar

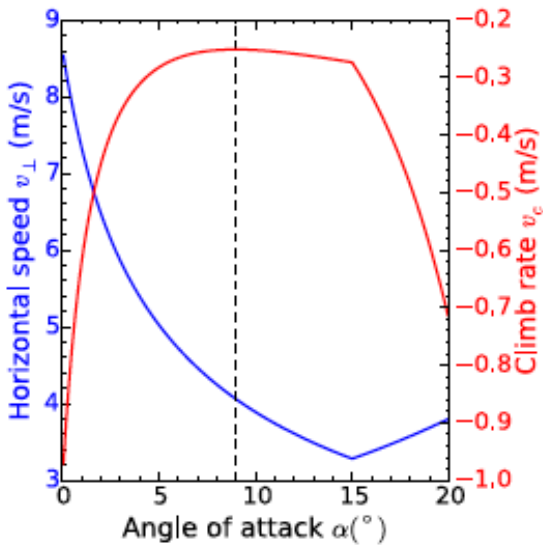
Glide-to-sink ratio

Swift	12:1
Hangglider	15:1
Frigatebird	20:1
Sailplanes	40:1

Learning of a flying agent







Agent exposed to an ensemble of turbulent flows tries to learn a policy for soaring, i.e., **what "short-term" sensorimotor cues and what response should it use to gain height?**

Foreplanning & Credit assignment

An "optimally" behaving agent: the reinforcement learning framework

$$\cdots\cdots \longrightarrow s_t \xrightarrow[a_t]{r_{t+1}} s_{t+1} \xrightarrow[a_{t+1}]{r_{t+2}} s_{t+2} \xrightarrow[a_{t+2}]{r_{t+3}} \cdots\cdots$$

$$\text{Action-Value :} \quad Q(s_t, a_t) = \underbrace{r_{t+1}}_{\text{reward}} + \underbrace{\beta Q(s_{t+1}, a_{t+1})}_{\text{sum of future rewards}} \,, \qquad 0 \leq \beta < 1$$

$$\text{Policy:} \qquad \pi_s^a = \arg\max_{a'} Q(s, a')$$

Learns empirical statistics and rewards solely through experience

Model-free

https://webdocs.cs.ualberta.ca/~sutton/book/bookdraft2016sep.pdf

States s — sensorimotor cues + bank angle + angle of attack
Actions a — modify angles

Learning by TD

- We use Markov Decision Processes (MDPs) - a framework for modeling decision-making

- Action-Value : $Q(s_t, a_t) = \underbrace{r_{t+1}}_{\text{reward}} + \underbrace{\beta Q(s_{t+1}, a_{t+1})}_{\text{sum of future rewards}} \;, \quad 0 \leq \beta < 1$

- Extracting Value: $Q(s_t, a_t^*) = \max_{a_t}(r_{t+1} + \beta Q(s_{t+1}, a_{t+1}^*))$

- TD update: $Q(s,a) \rightarrow Q(s,a) + \eta(r + \beta Q(s',a') - Q(s,a))$

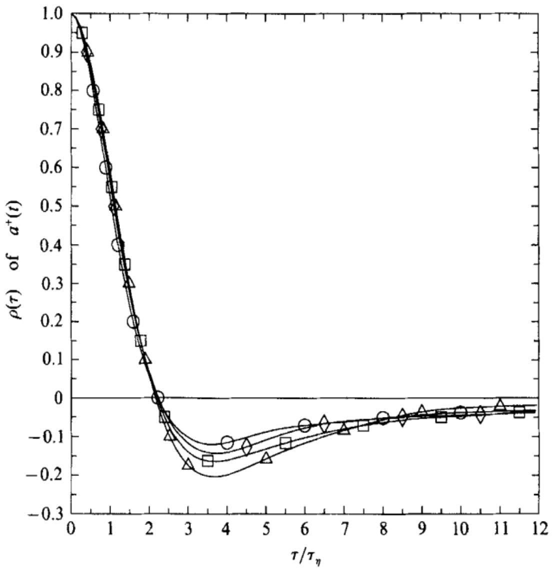
- Policy: $\pi_s^a \propto \exp\left(+ \hat{Q}(s,a)/\tau \right)$

Credit assignment and reward shaping

Vertical velocity as a reward would be the natural choice, yet it does not work while vertical accelerations do

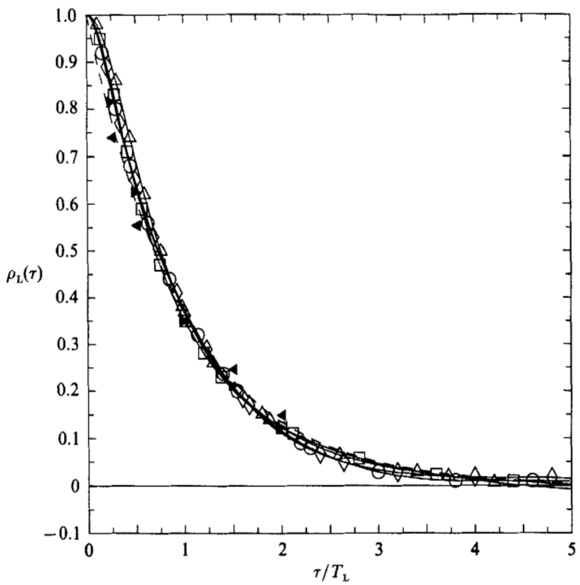
Accelerations

Autocorrelation



t/(kolmogorov scale)

Velocities



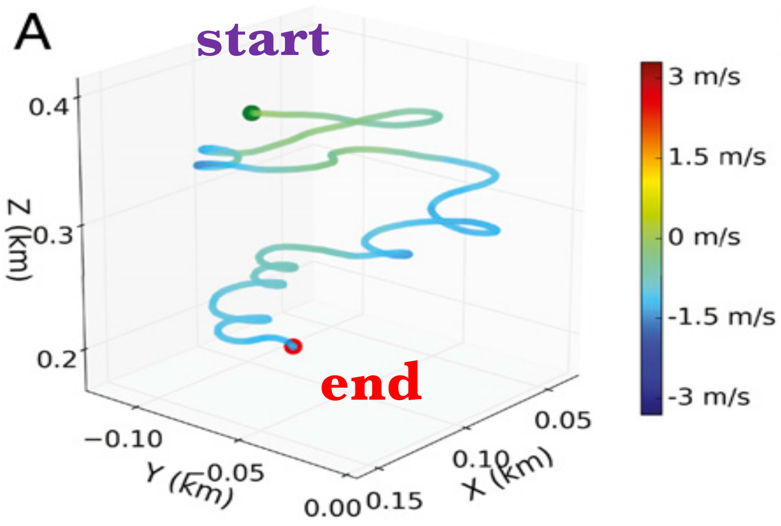
t/(integral scale)

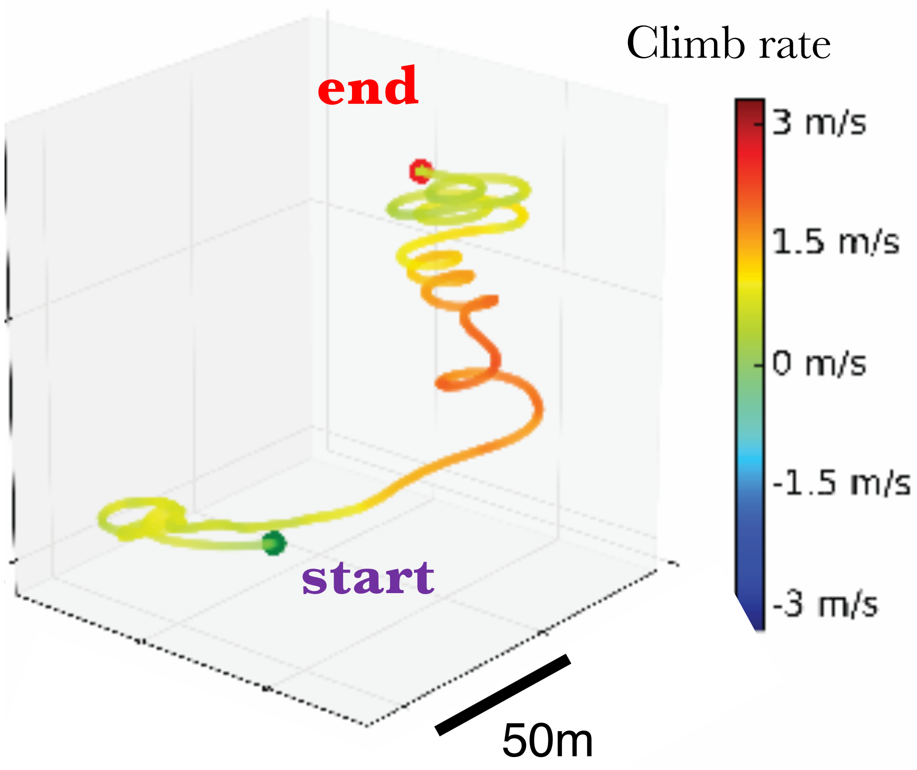
Yeung and Pope, JFM, 1989

Policy optimal for a given reward is also optimal when taking "discrete time derivatives" of the reward

After training in simulations…

After ~200 5-minute “training episodes” in different thermal environments..

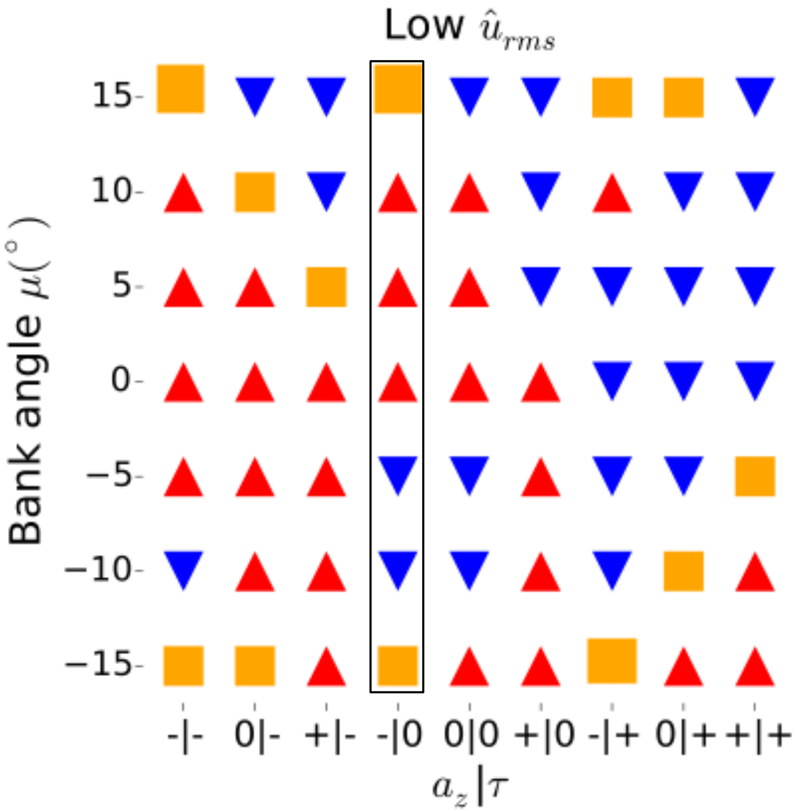




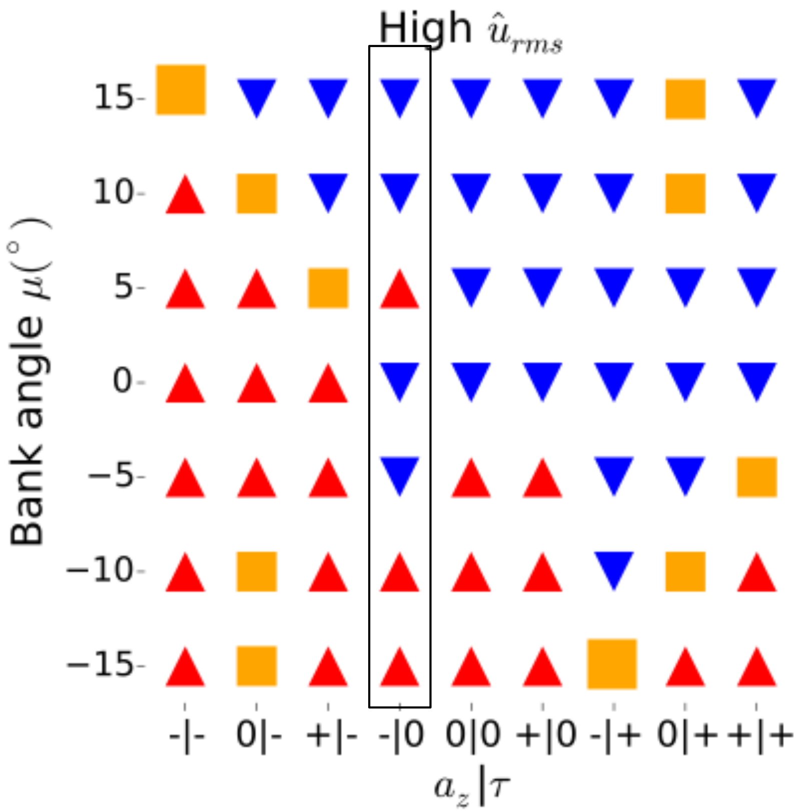
A trajectory in 3D of a glider in simulation after training

Learned flight policy

Weak fluctuations

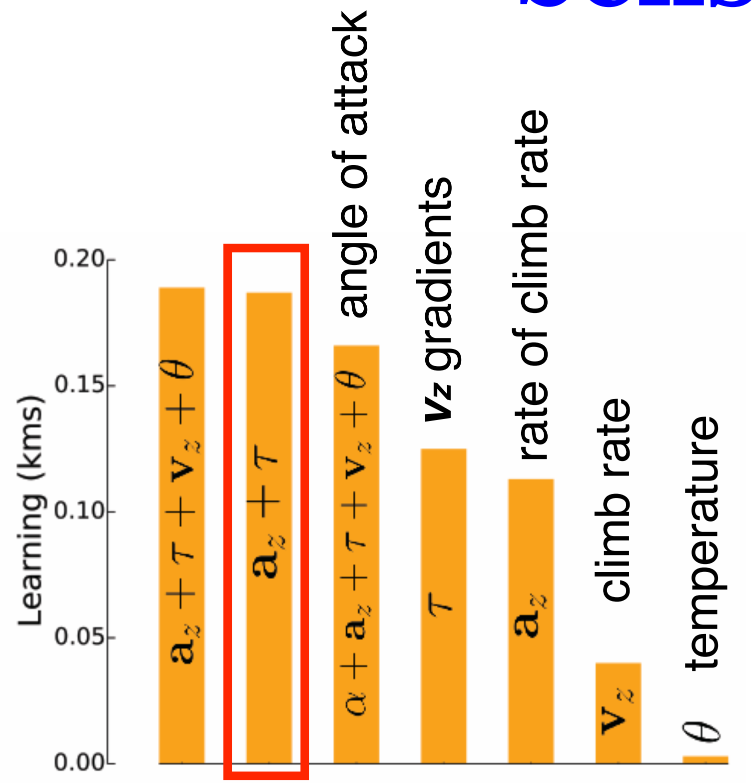


Strong fluctuations

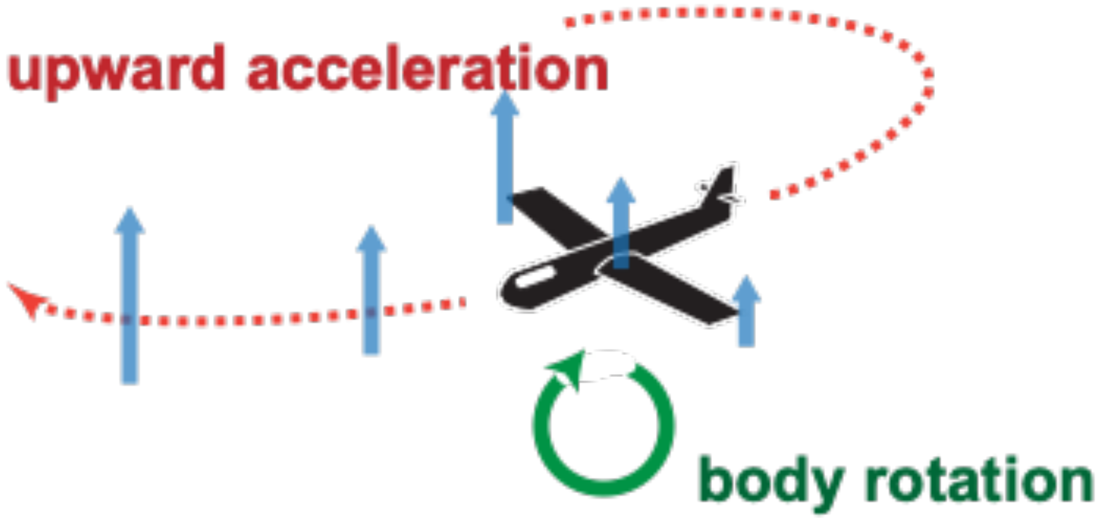


More risk-averse policy for strong fluctuations

Sensory-motor cues



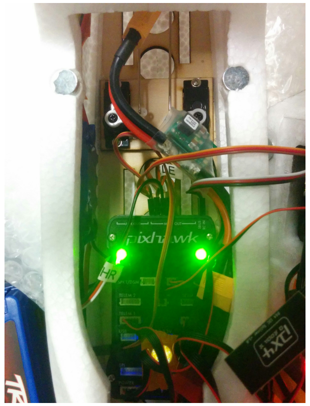
Upward acceleration and body rotation are the most useful cues

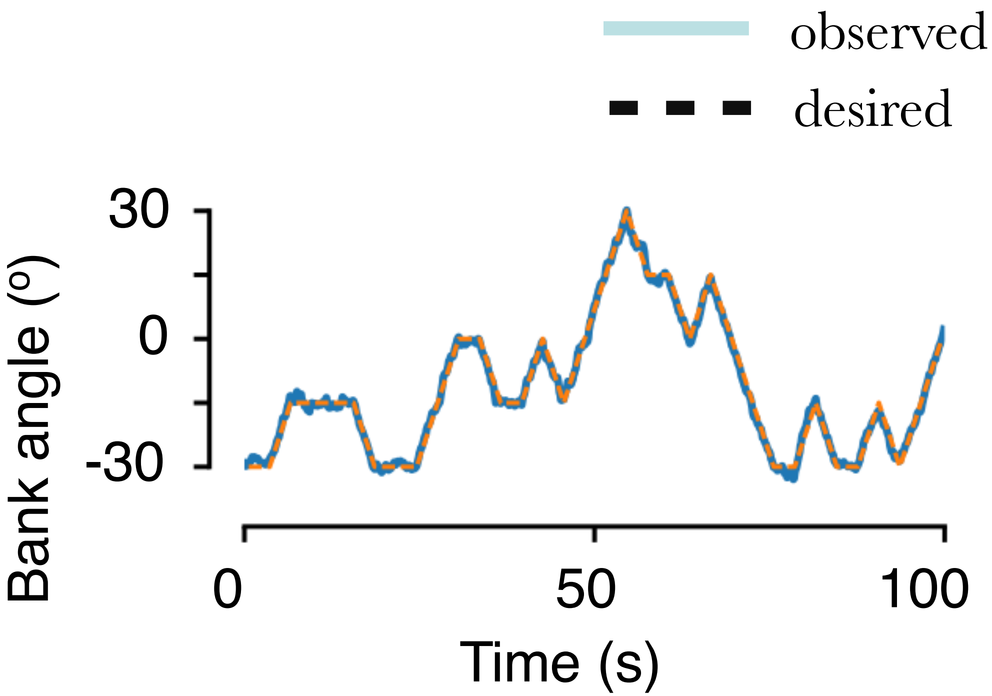


Do birds measure these quantities?
Still works in the field?

In the field


RADIAN




observed
desired
Bank angle (º)
30
0
-30
0
50
100
Time (s)


Google Earth

Measuring vertical wind accelerations

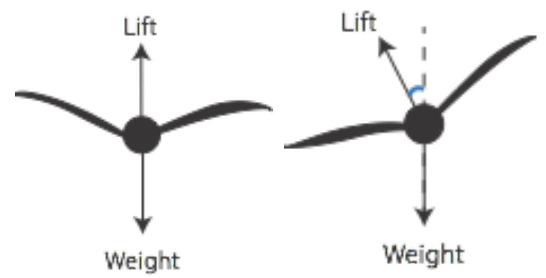
GPS/baro

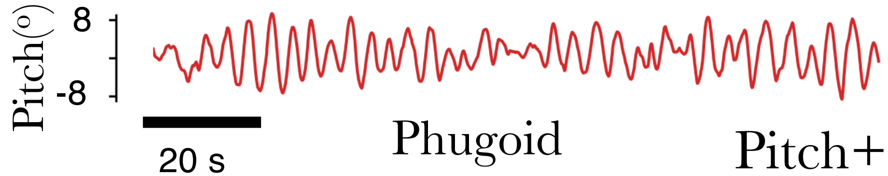
Wind vel. Glider vel.

$$w_z = u_z - v_z$$

Ground vel.

modeling



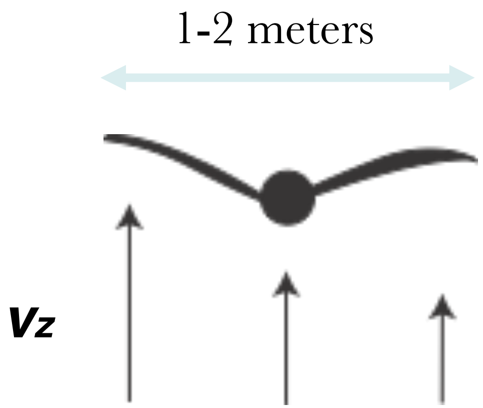


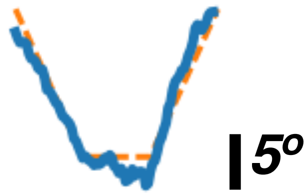
Phugoid

Pitch+angle of attack

$$\omega \approx g/V \sim 1\mathrm{Hz}$$

Measuring vertical wind velocity gradients

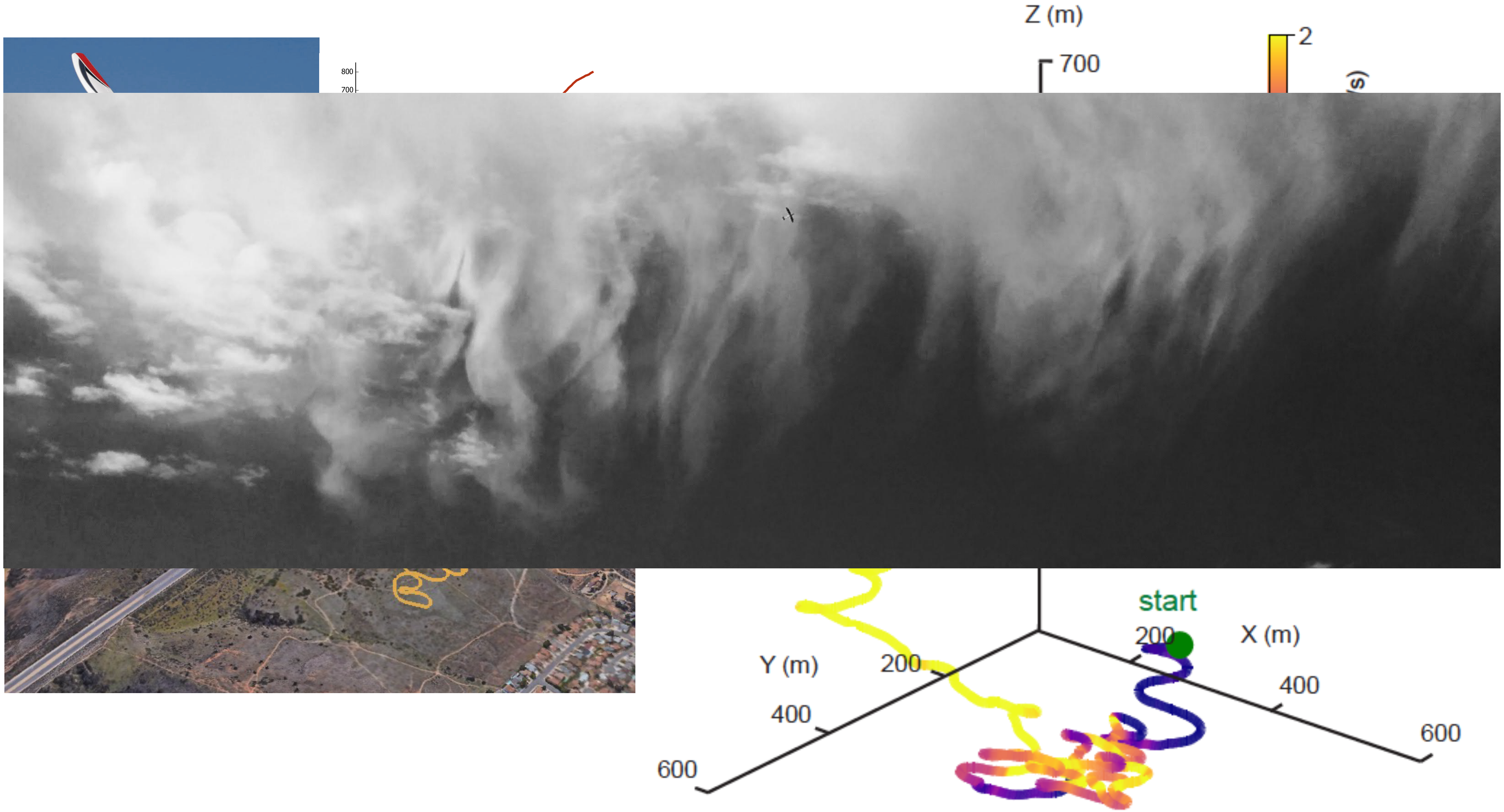




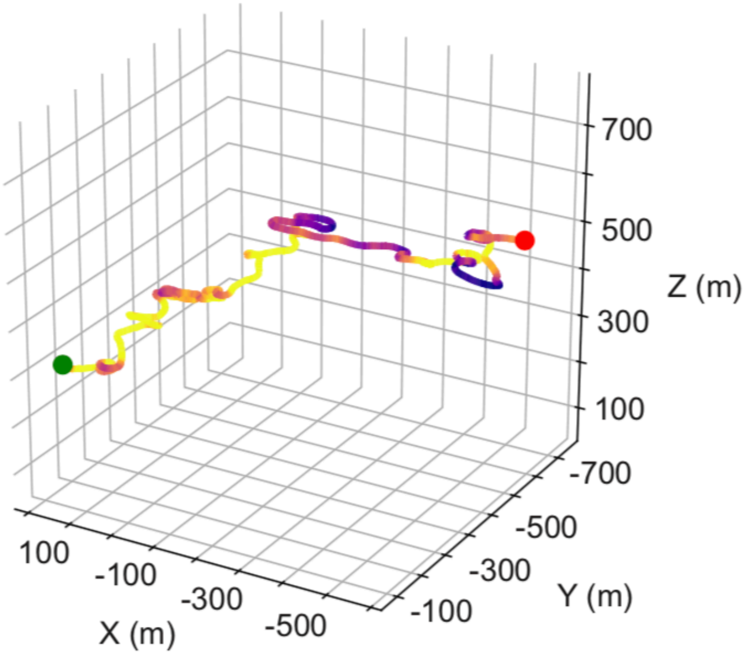
change in bank = feedback control + aerodynamics + wind gradients

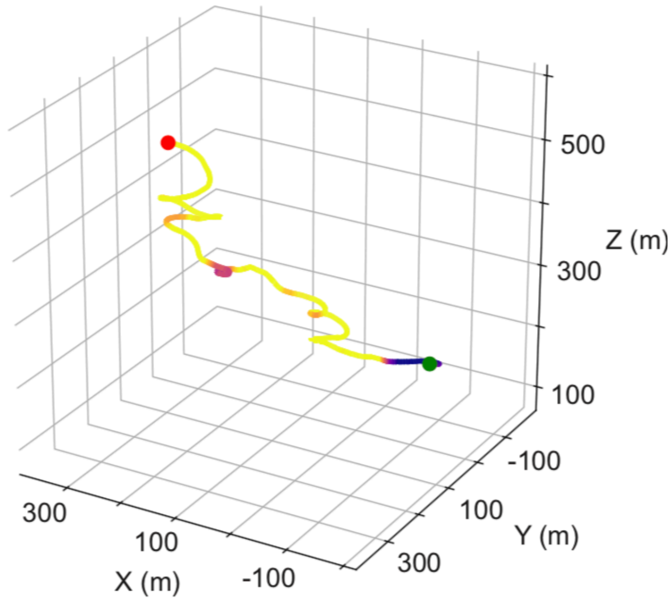
Learning works in the field as well

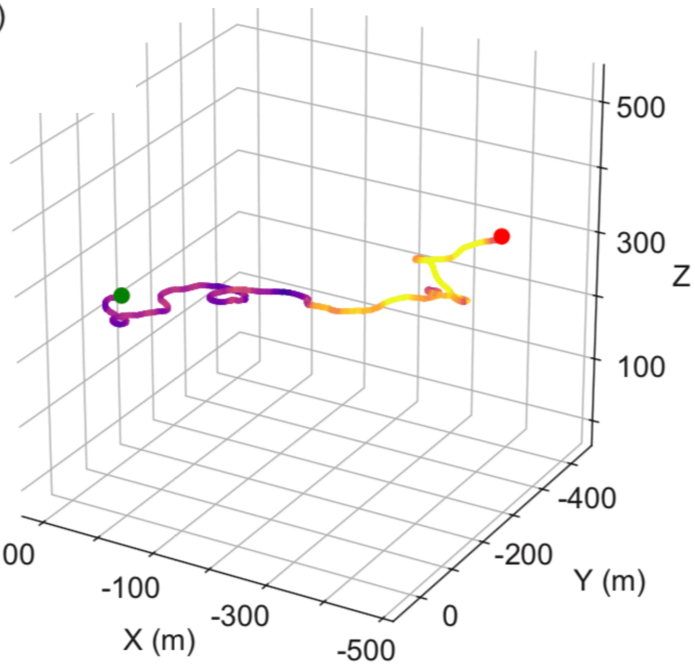
Learning works in the field



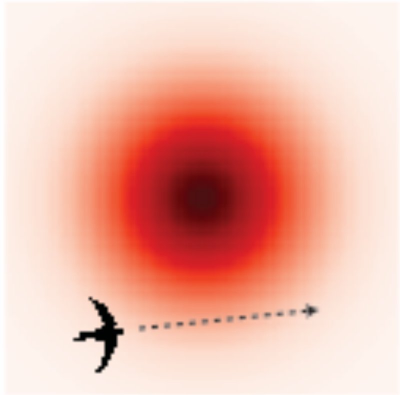
A few more examples





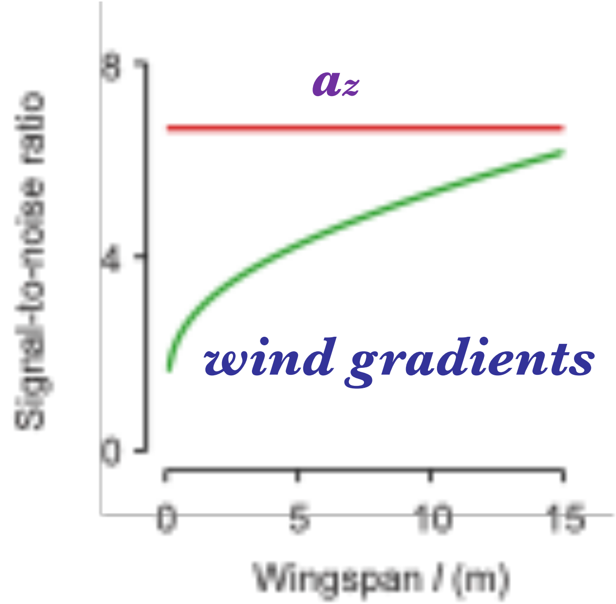


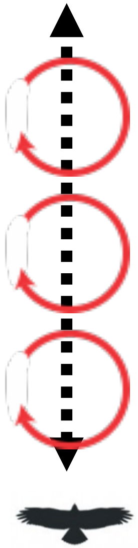
Can a bird sense the relevant cues?



$$\omega \sim l\,?$$

"artistic" view of turbulence





$$\frac{\omega}{\delta v} \sim l^{1/6}(VT)^{1/2}$$

$$\frac{a_z}{\delta a_z} \sim (VT)^{2/3}$$

$$\delta v \sim l^{1/3}$$

$$n \sim VT/l$$

UCSD

SIMONS FOUNDATION

Reading material can be found at

1) Reddy et al., Infomax strategies for an optimal balance between exploration and exploitation, J. Stat Phys, 2016

2) Reddy et al, Learning to soar in turbulent environments, PNAS, 2016

3) Reddy et al, Glider soaring via reinforcement learning in the field, Nature, 2018

ENS

Postdoctoral Junior Research Chairs in Theoretical Biophysics at ENS, Paris

https://academicjobsonline.org/ajo/jobs/14026

Closing date: **October 30, 2019**