Credits: Bénédicte Rossi

Understanding the auditory system through its neuronal population activity

Brice Bathellier

Institut de l'Audition, Institut Pasteur, Paris

Purposes of the auditory system ?

- To detect sounds

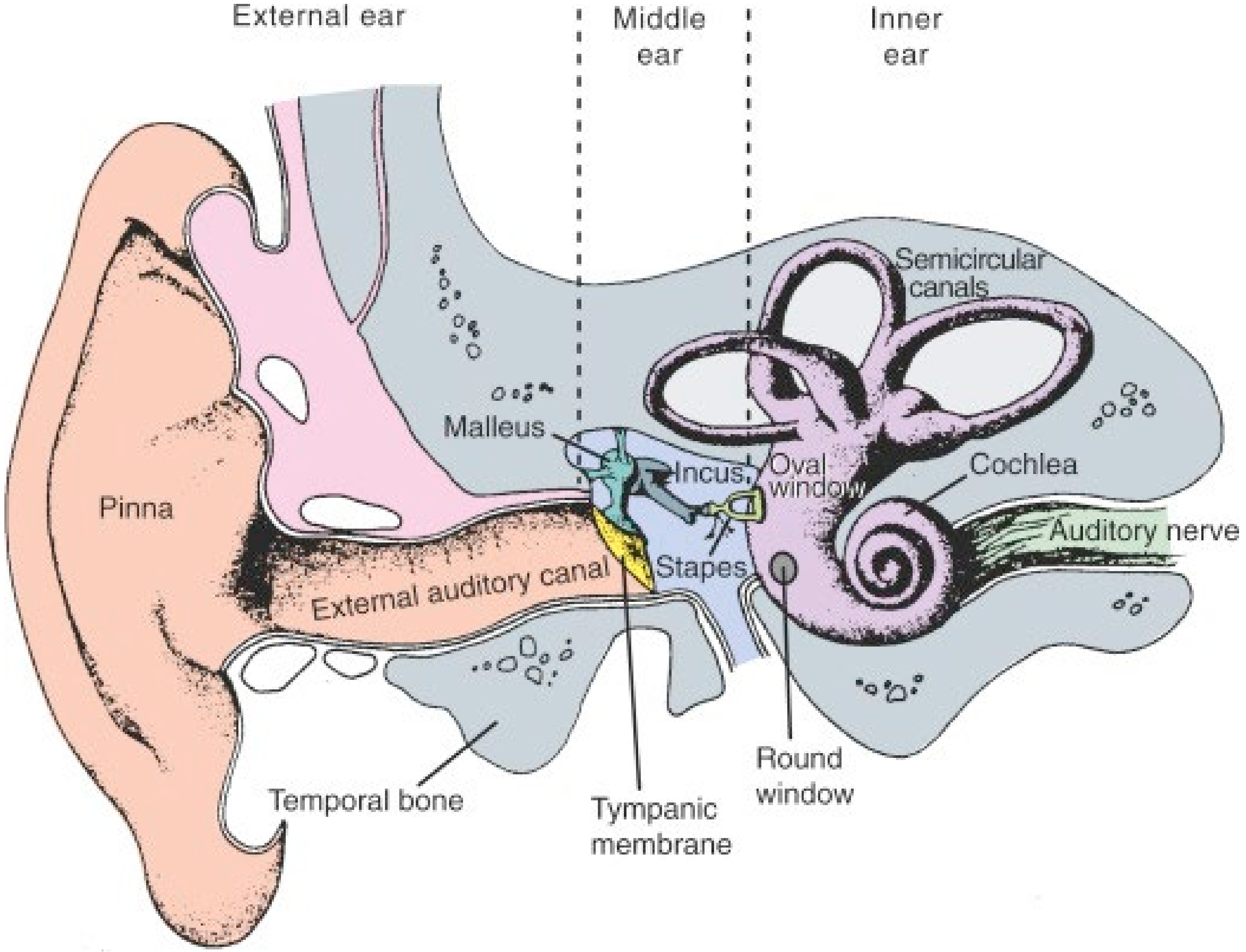
- To localize sounds

- To interpret sounds

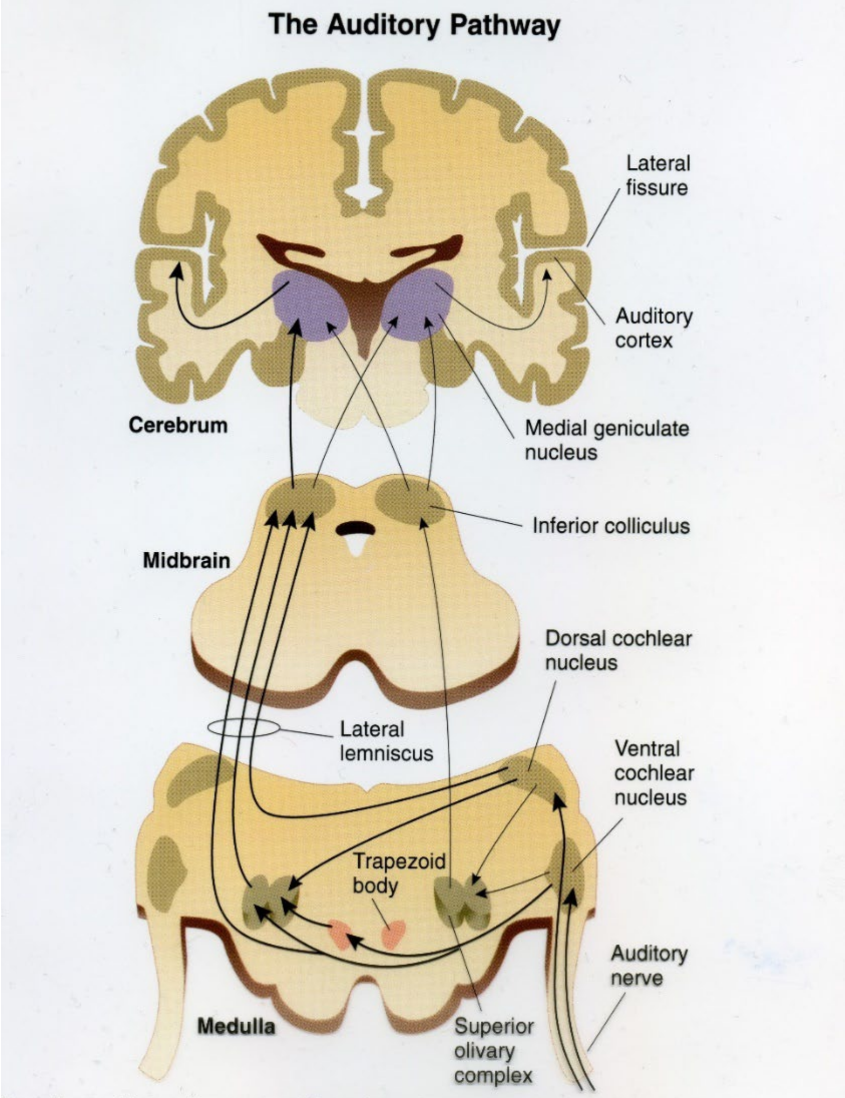
- Identify the origin/cause of sounds
  - ➢ Sound categories
  - ➢ Assign meanings (e.g. bell ringing 12x)
- Identify messages from conspecifics
  - ➢ Language
  - ➢ Vocalizations
- Enable the association of sounds to the appropriate behavioral response

➢ **Like vision, hearing is a sense gathering distant information from propagating waves.**
➢ **Hearing is dedicated to acoustic waves (and vision to electromagnetic waves)**
➢ The source of sounds are mechanical vibrations which transmit through a medium (air, water, …)

The peripheral auditory system



The central auditory system

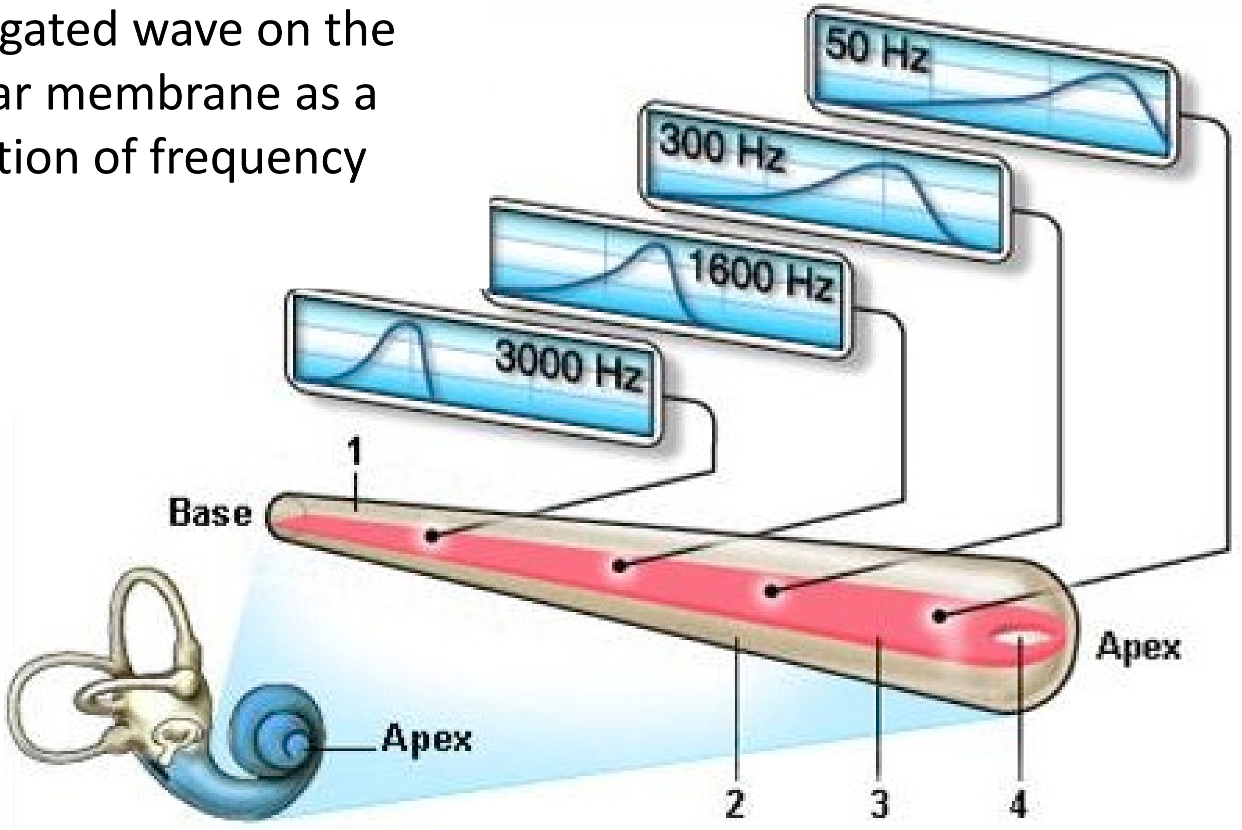


The pathway of the auditory system. The major pathways are indicated by heavy arrows.

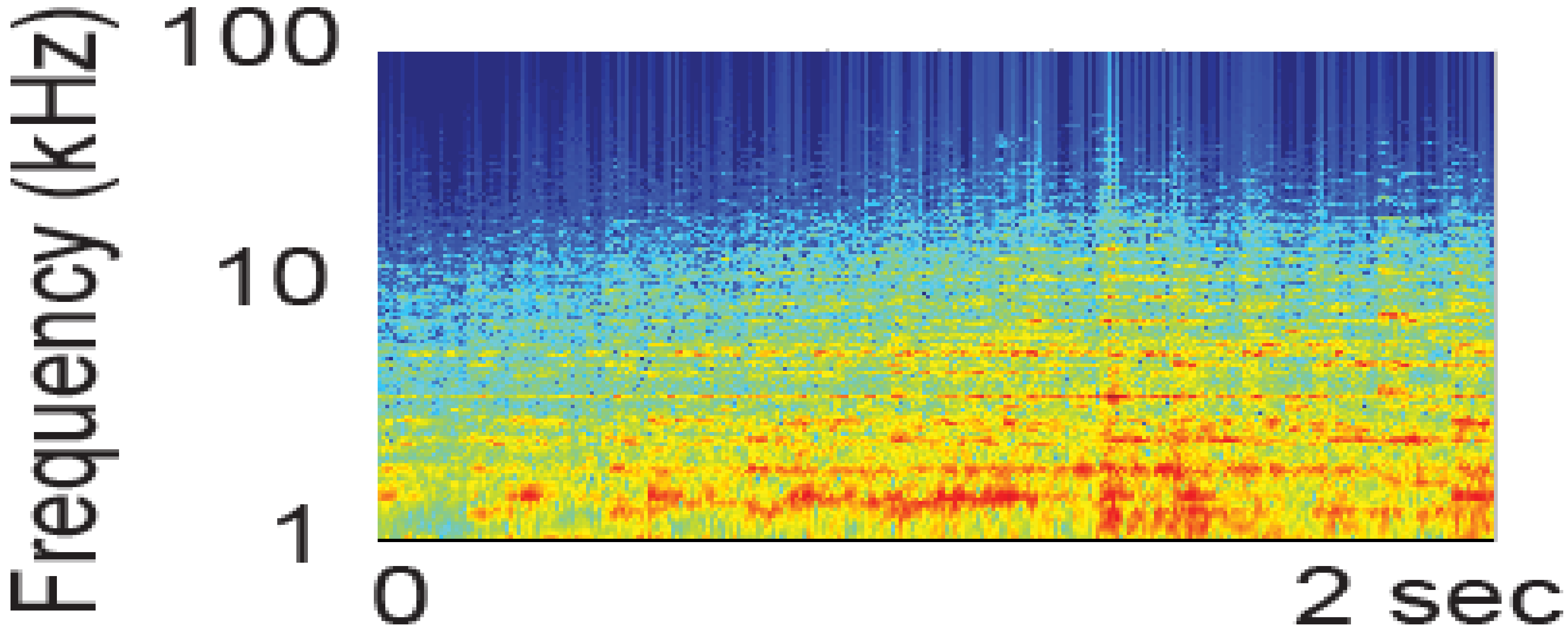
T-46

The cochlea: a frequency analyzer

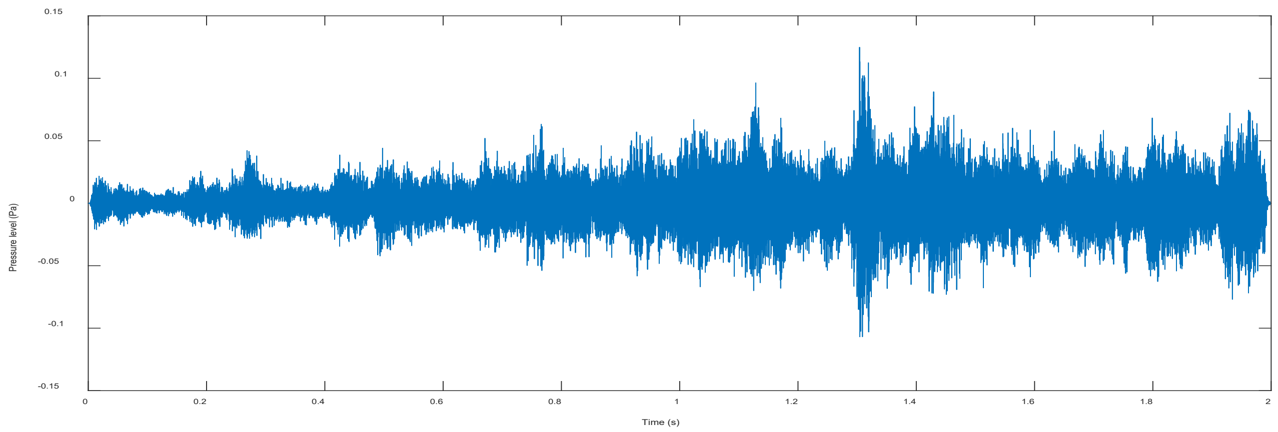
Propagated wave on the basilar membrane as a fonction of frequency



Time-frequency representation: the spectrograms

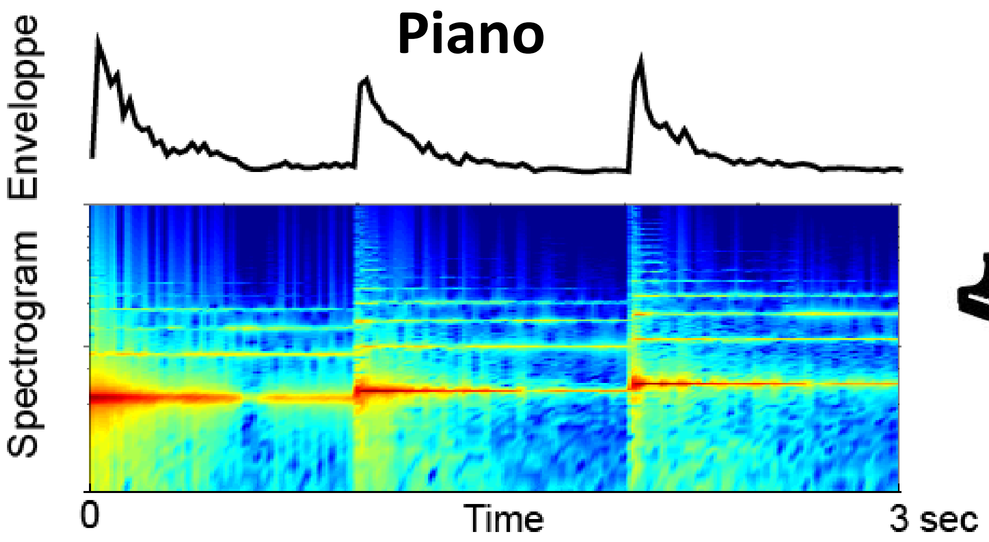


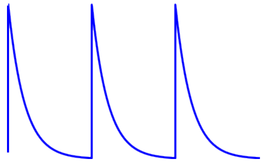
Music (Beethoven)

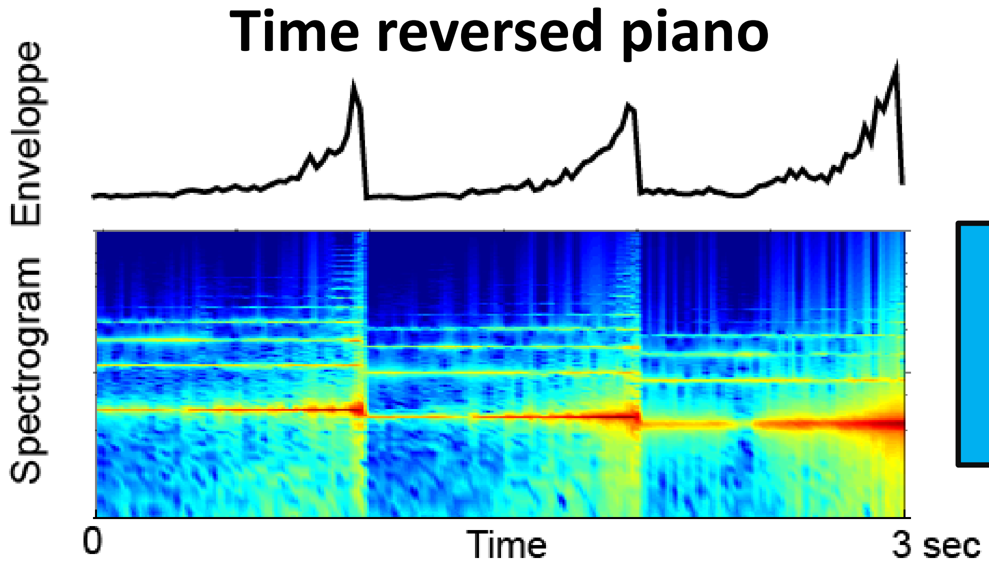


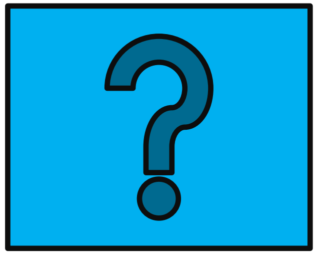
Temporal features are important in for sound interpretation

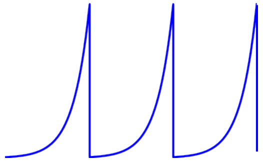
Pure tones







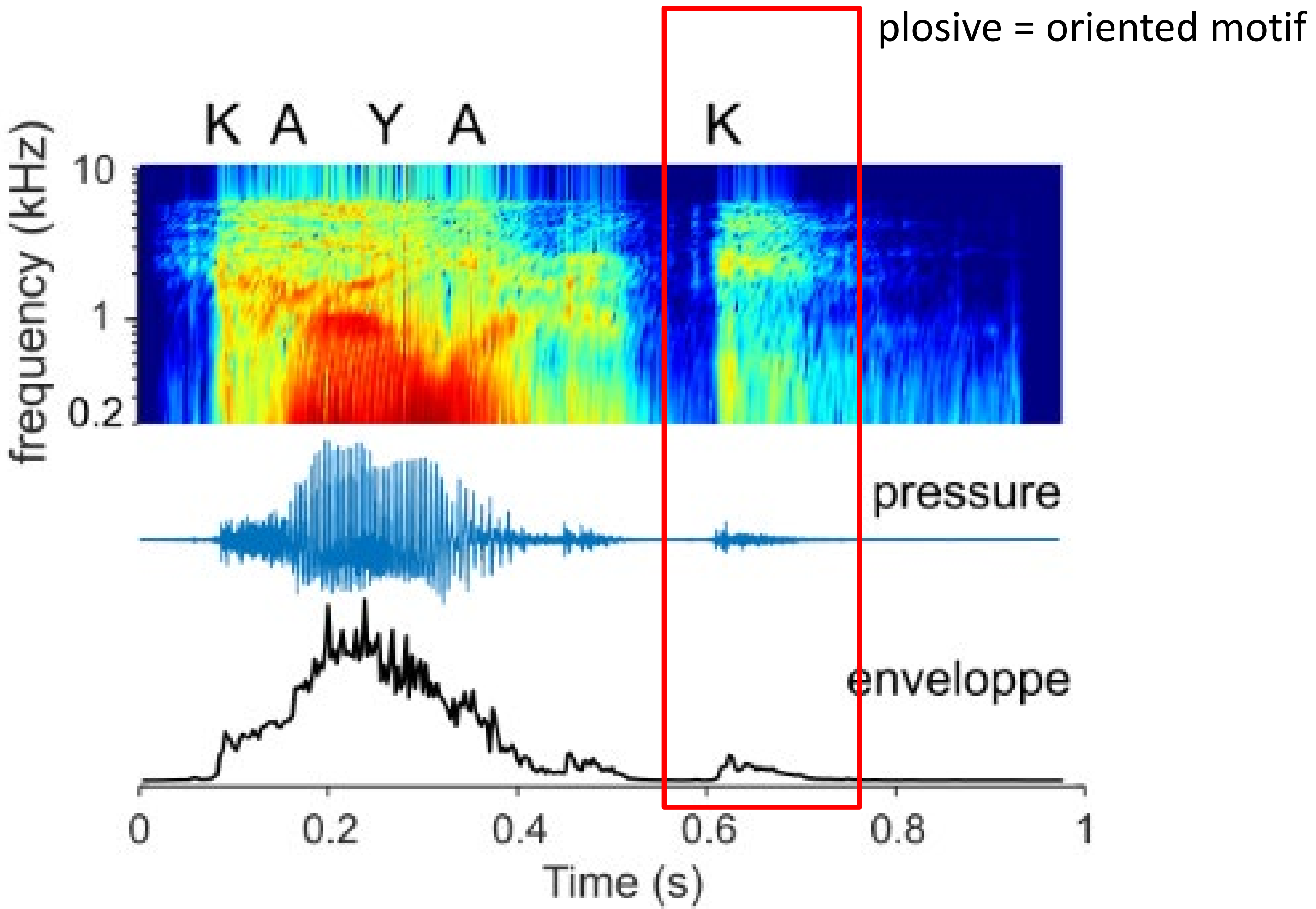




Time dimension in words

PoP **KayaK**

PoP **KayaK**



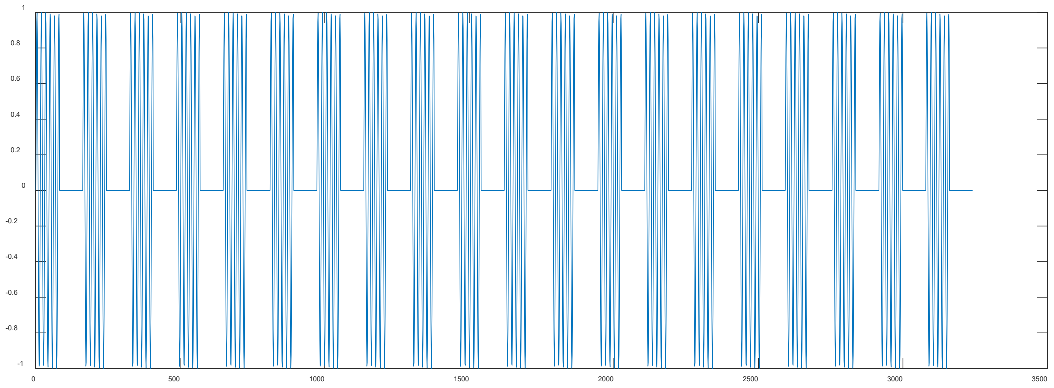
Rythmic amplitude modulations

- Roughness (often unpleasant, arising)

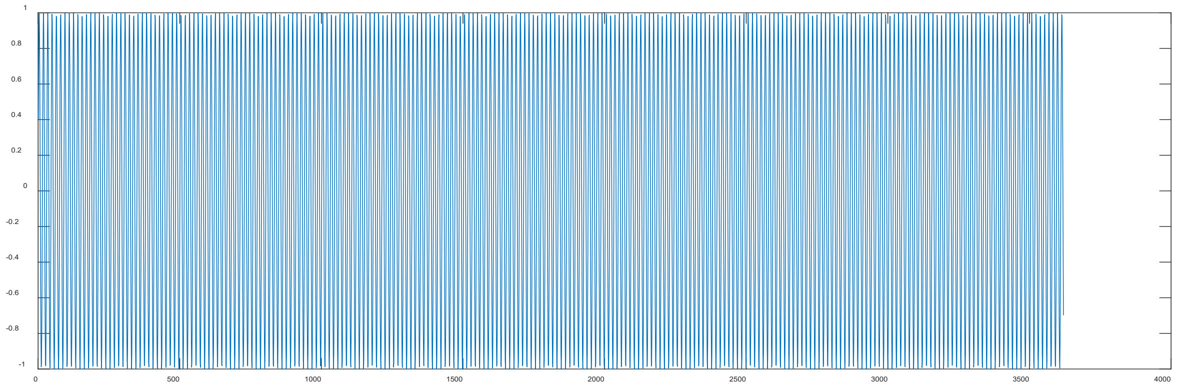
f=540 Hz

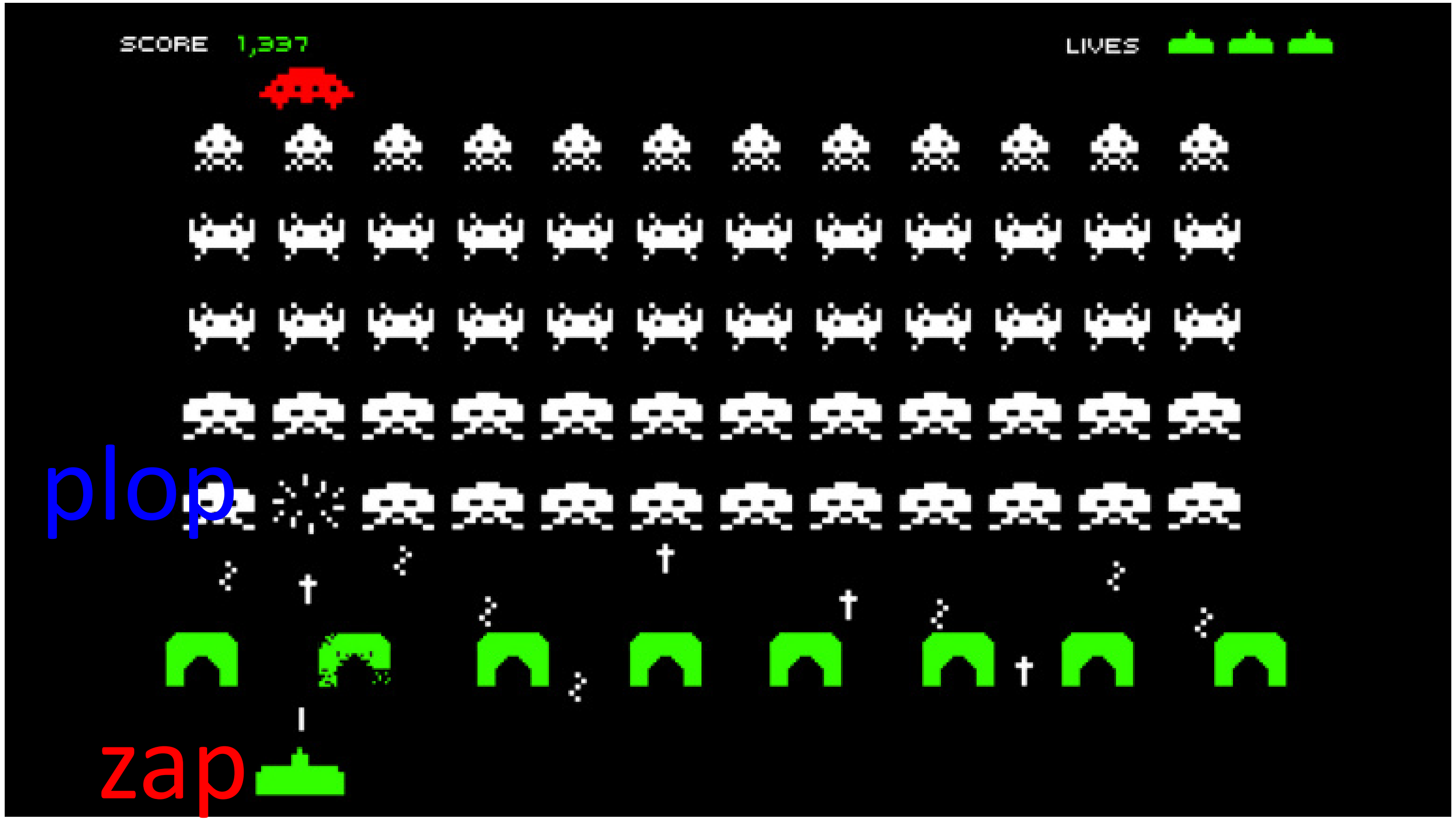
25 Hz AM

50 Hz AM



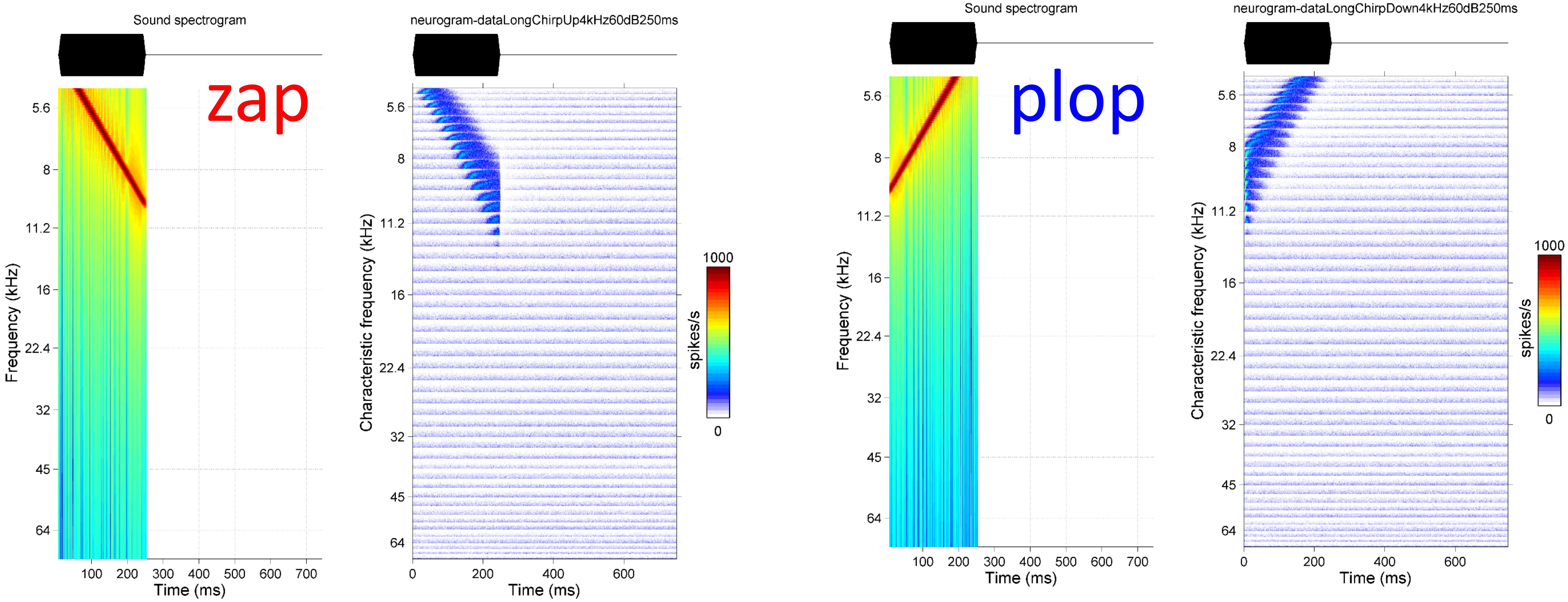
Pure tone without AM





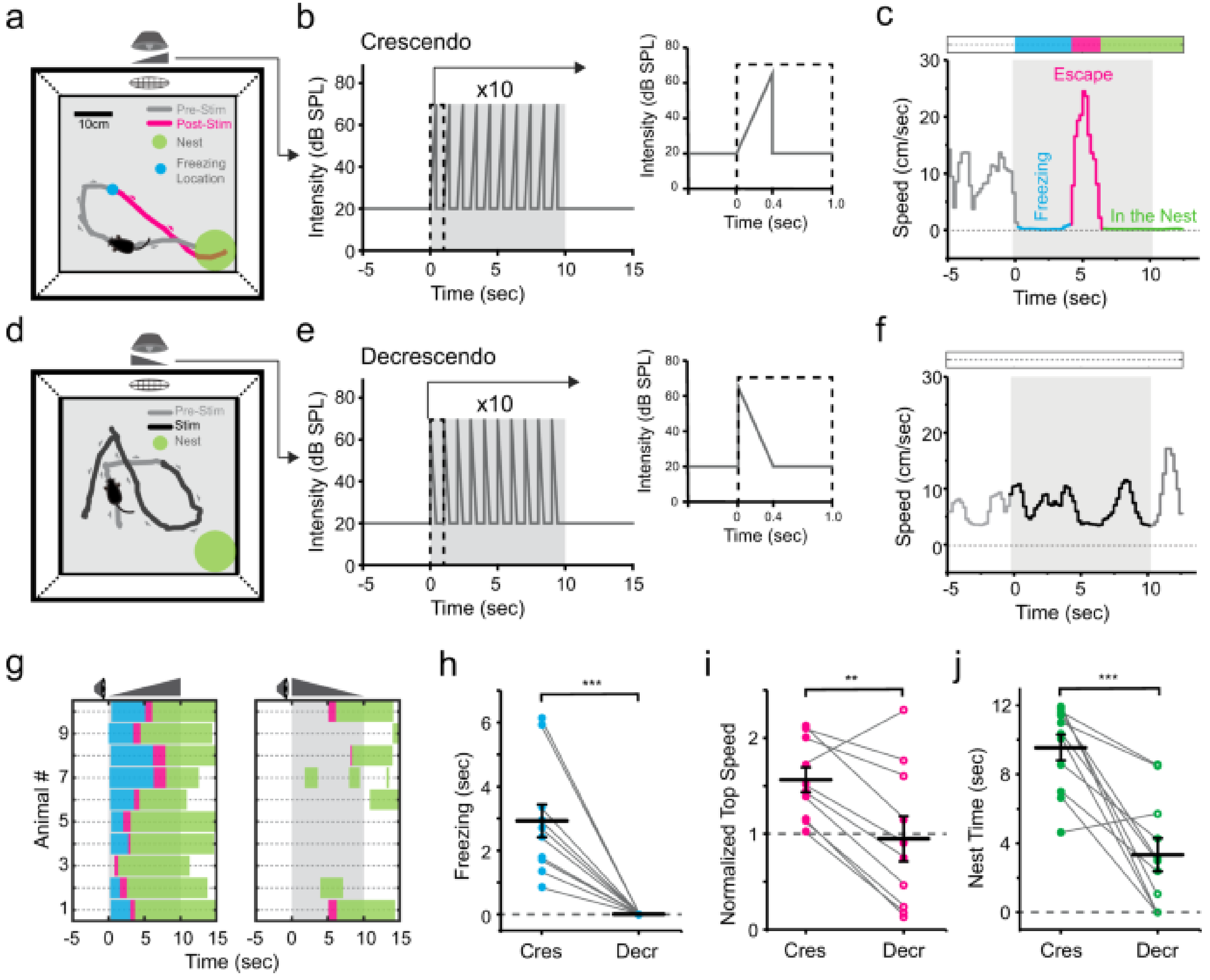
Space invaders

Frequency modulations in time are also features based on which we can recognize and interpret sounds



Biophysical model of cochlea, output nerve simulations, J. Bourien, INM (Montpellier)

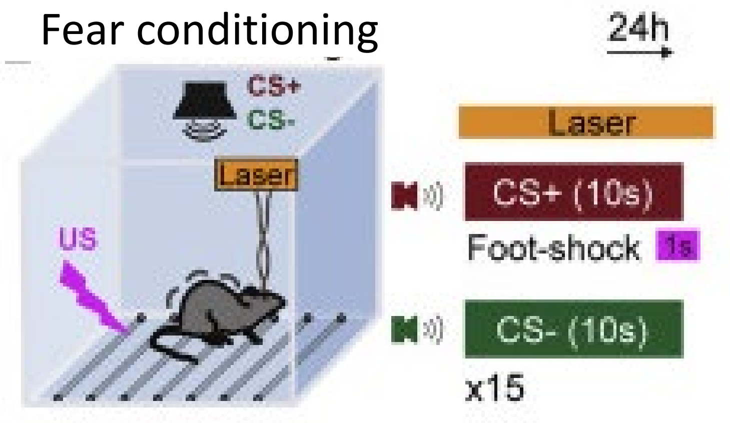
Mice recognize temporal sound features, even inately



Li et al. 2021, Li Zhang’s lab

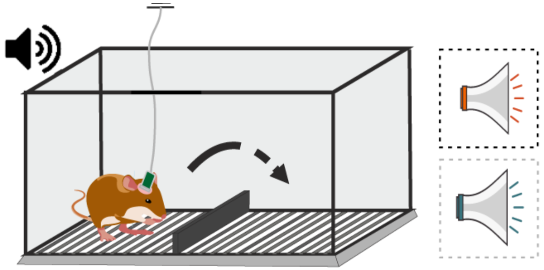
Perception is also often measured through learnt discrimination behavior

Fear conditioning



Letzkus lab

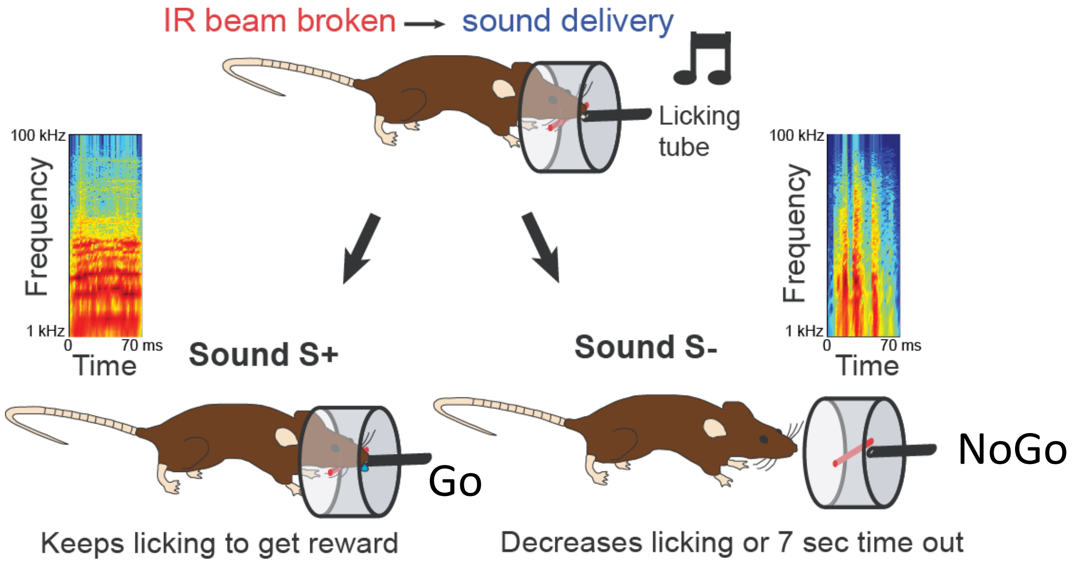
Two-way shuttle task



CS+ = Go

CS- = NoGo

e.g. Ohl & Scheich 1999



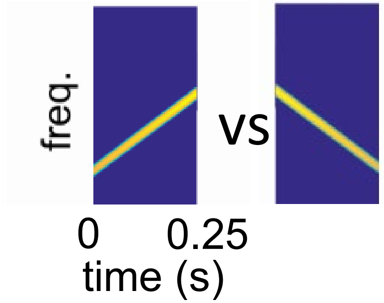
Bathellier et al. 2012

Aversive

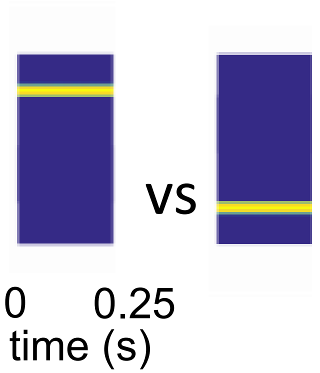
Appetitive

Learnt discrimination of palindromic sounds in gerbils and mice requires auditory cortex

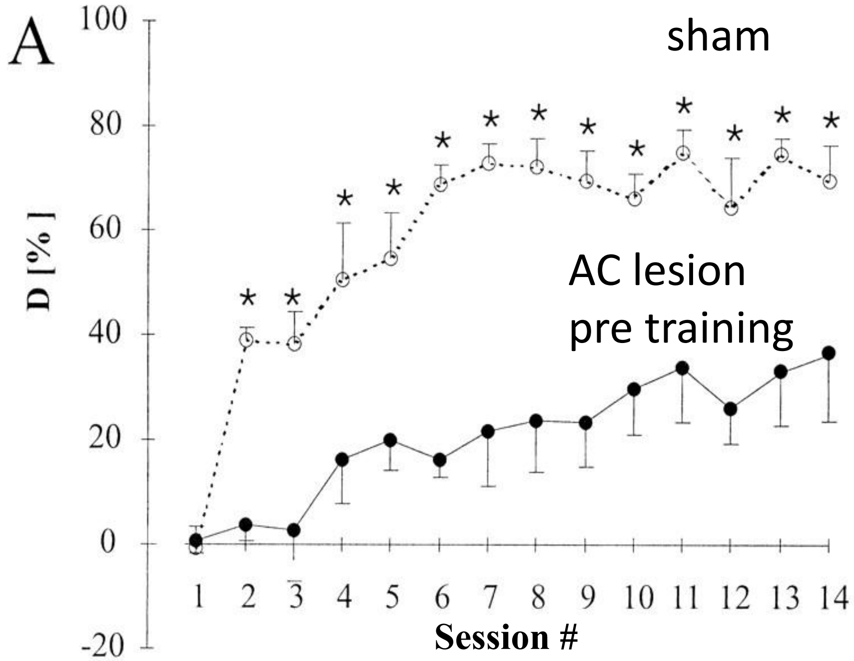
FM sweep



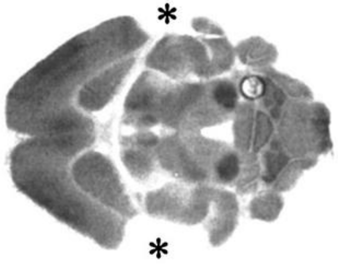
Pure tone

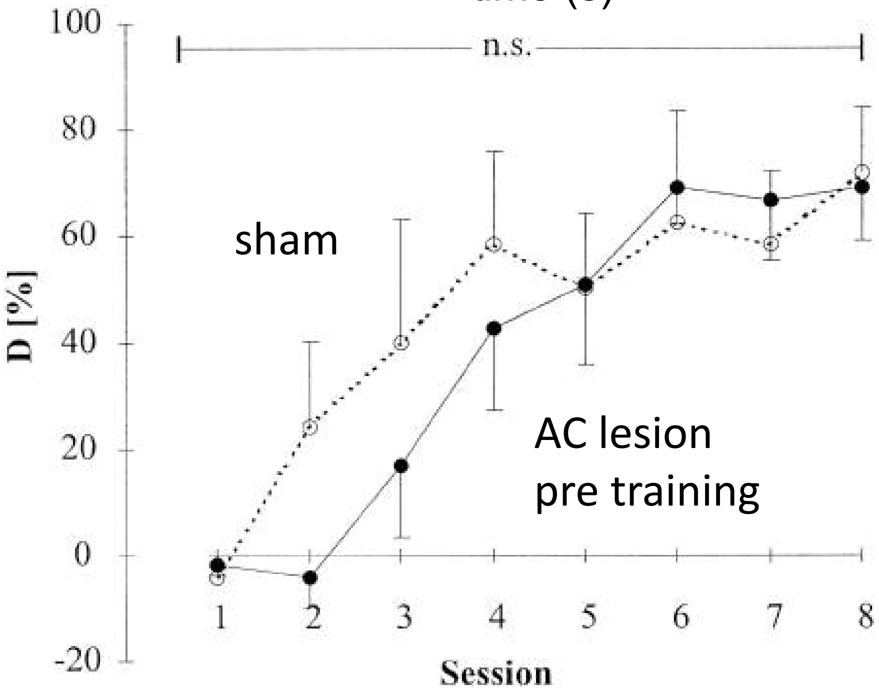


Ohl et al. 1999



Two-way shuttle box task in Gerbils

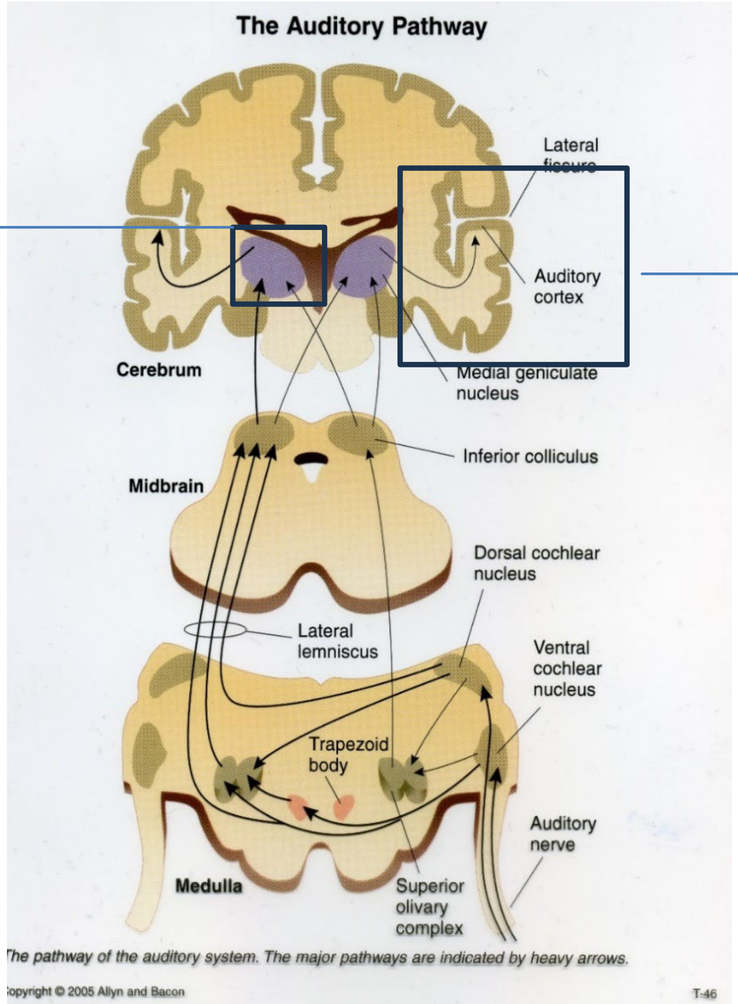




Identical results in mice with optogenetics: see Dalmay et al. 2019 (fear) and Ceballo et al. 2019 (licking)

What are the transformations of sound representations which enable temporal feature discrimination ?

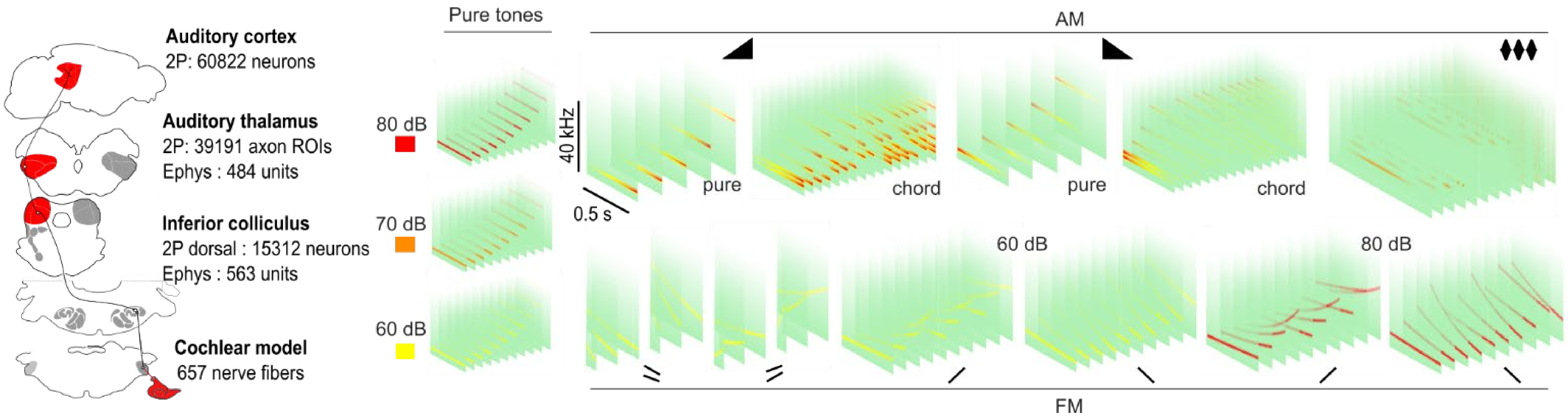
Limited for temporal feature discrimination



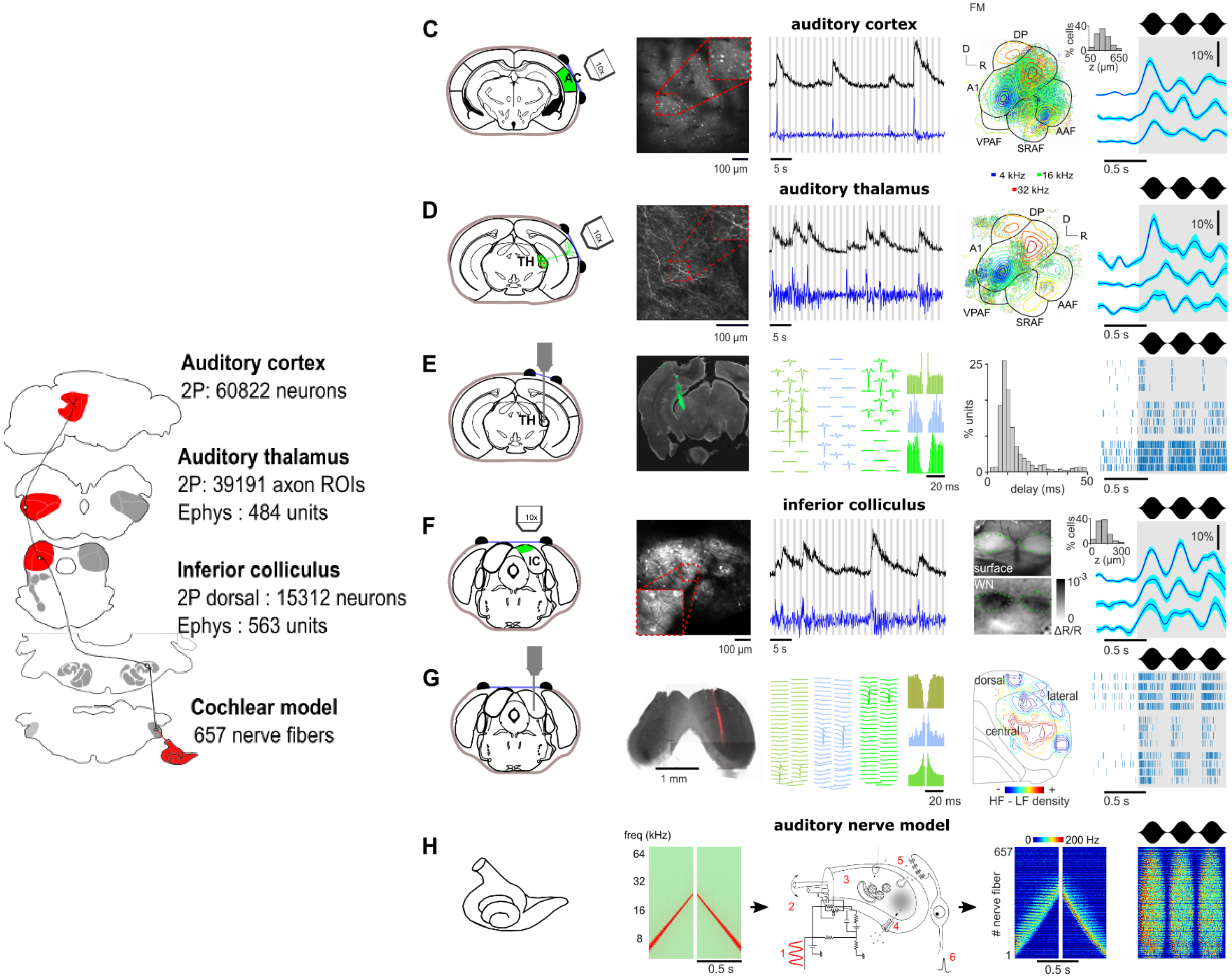
Efficient for temporal feature discrimination

Let’s look at neuronal population across the auditory system and how they encode sounds to find out the difference !

Large-scale sampling of the late auditory system for sounds varying in frequency and time



500 ms sounds



AM 3 Hz

Linear deconvolution
Resolution ~150ms

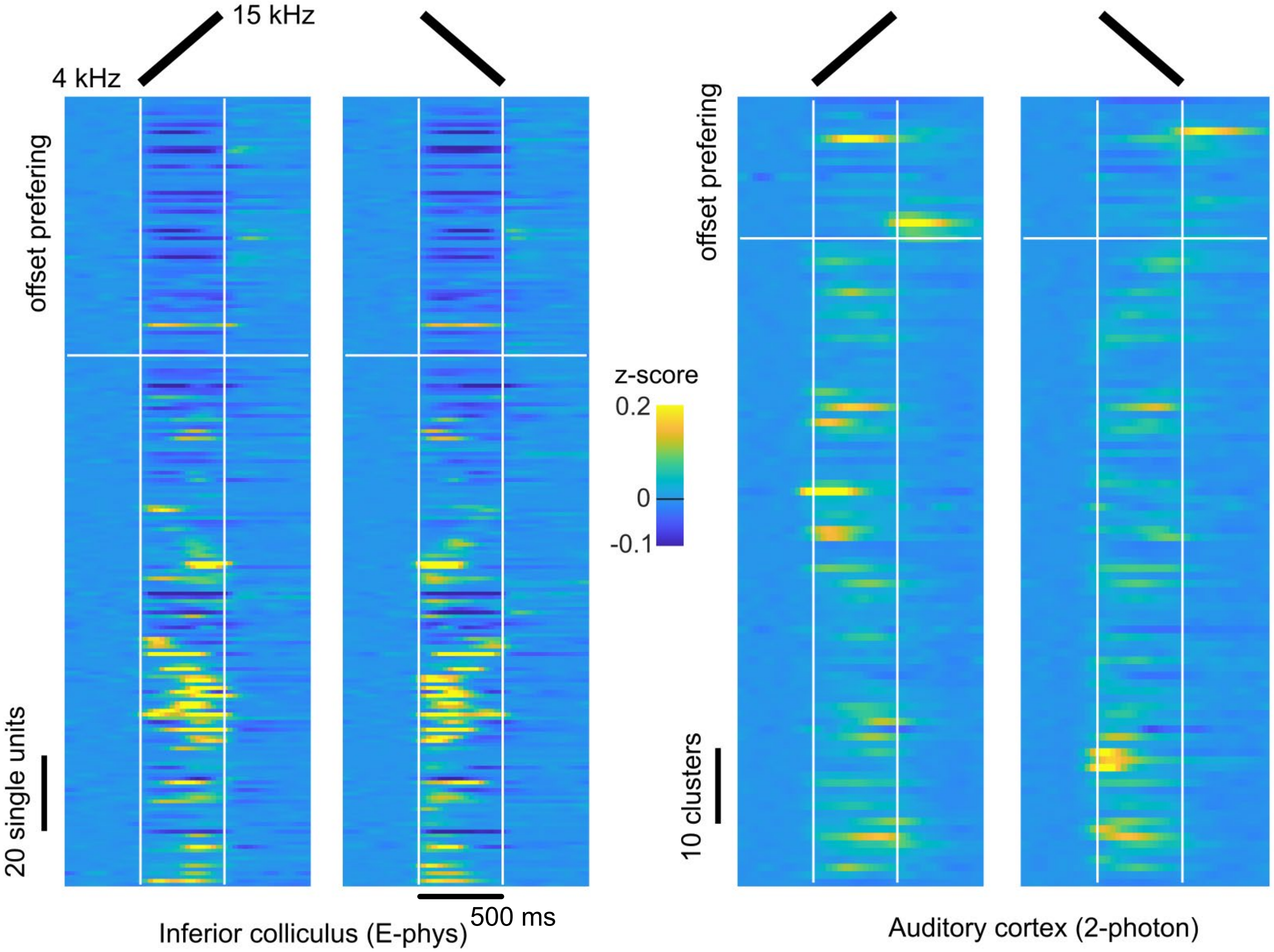
T. Tarpin
A. Kempf
J. Bourg
S. Bagur
J. Bourien (INM)

Transformation of temporal feature representations: the emergence of a spatial code in cortex

Subcortical

You need to know <u>when</u> neurons fire

Efficient temporal code
Inefficient spatial code

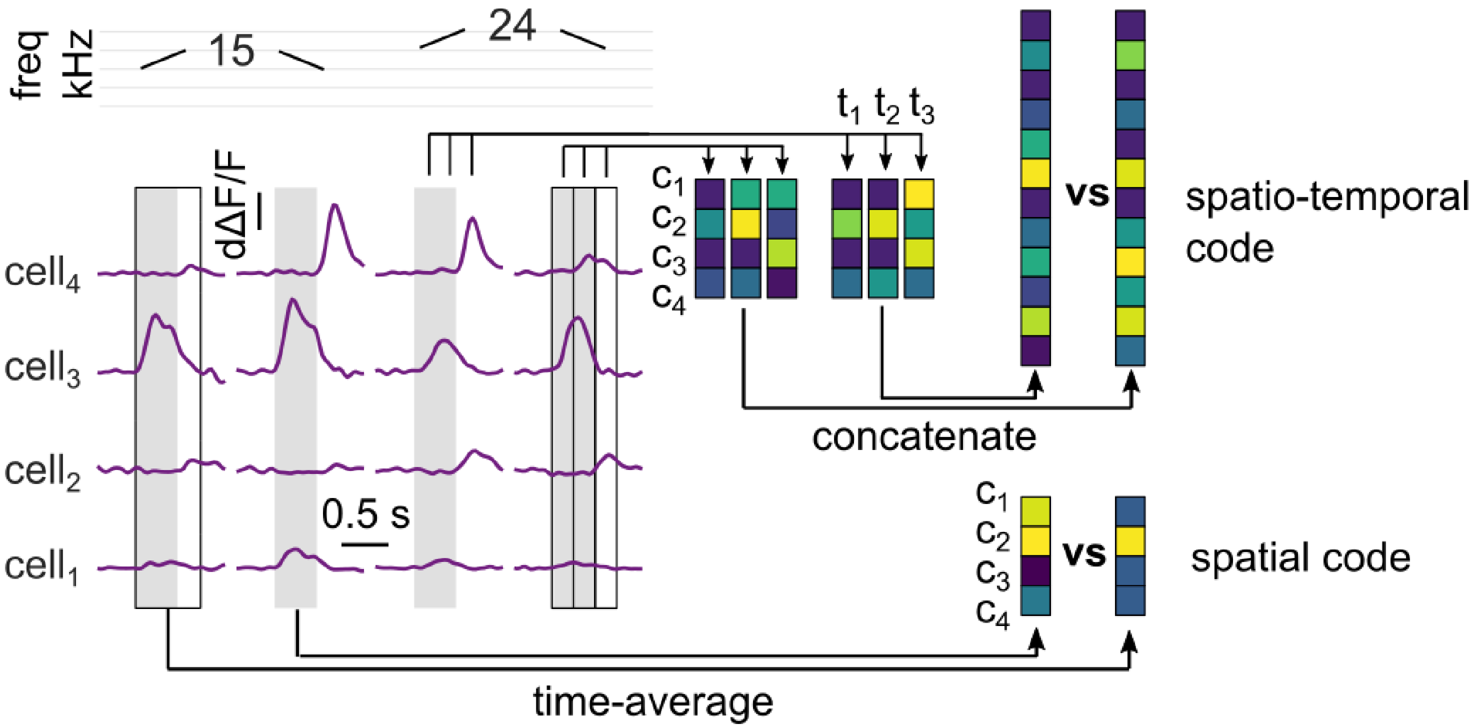


Cortical

You just need to know <u>which</u> neuron fire

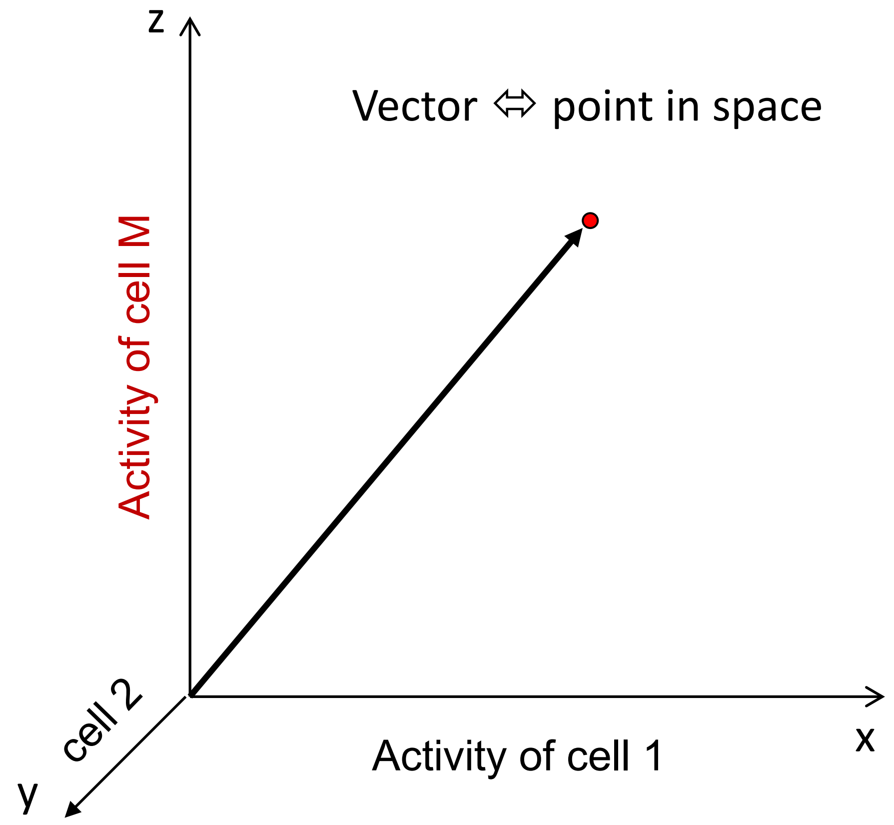
Efficient temporal code
Efficient spatial code

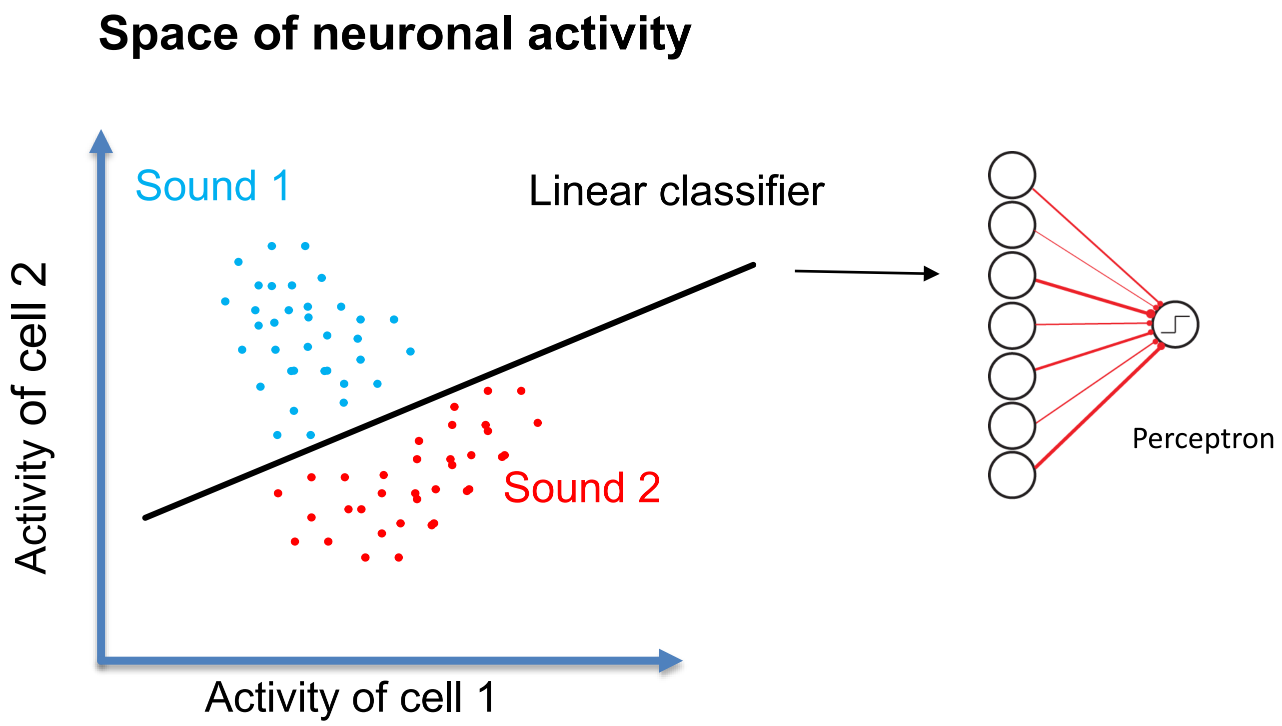
The emergence of a spatial code: quantification at the population level



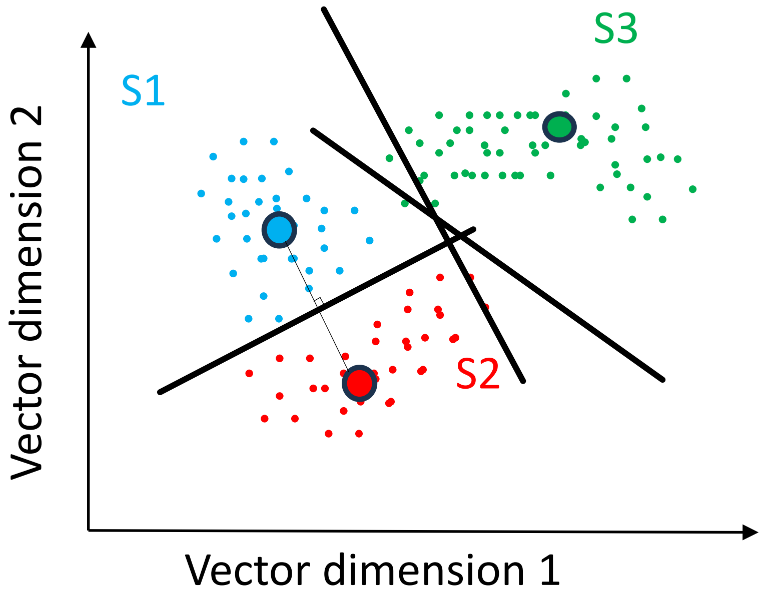
Extraction of spatial or spatio-temporal population vectors

Quantififying discrimination of population vectors with linear classifiers



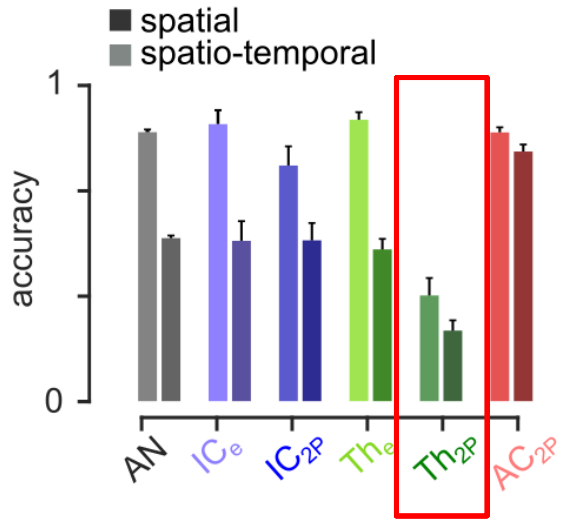


Closest centroid multiclass classifier

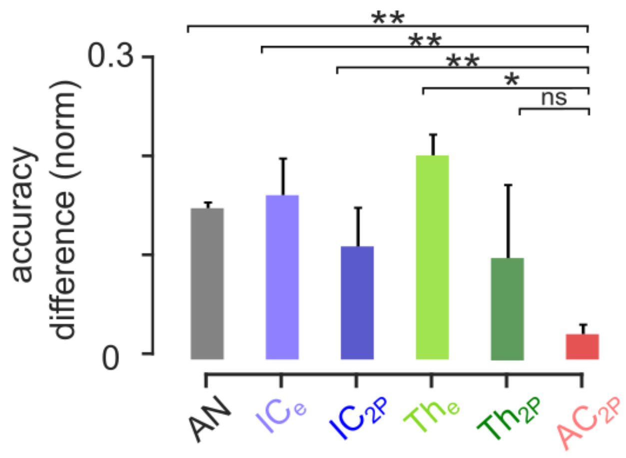


Classifiy single trials is based on the closest centroid. This is equivalent to using the hyperplane that passes thought the middle of the segment joining the two centroids and perpendicular to it. This can be used robustly on many classes with minimal training issues.

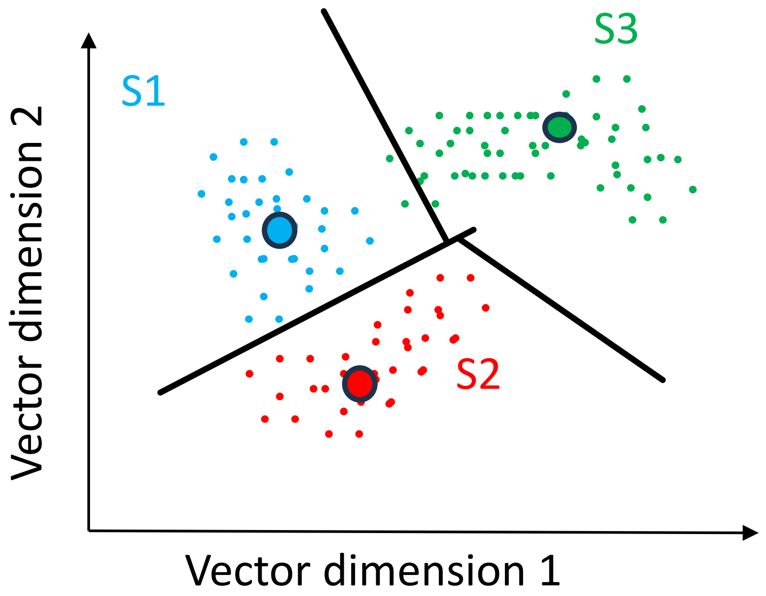
The emergence of a spatial code: quantification of information linearly decodable at the population level



Axonal recordings



Limits of the linear population classifier
(what you should have in mind when using this tool)



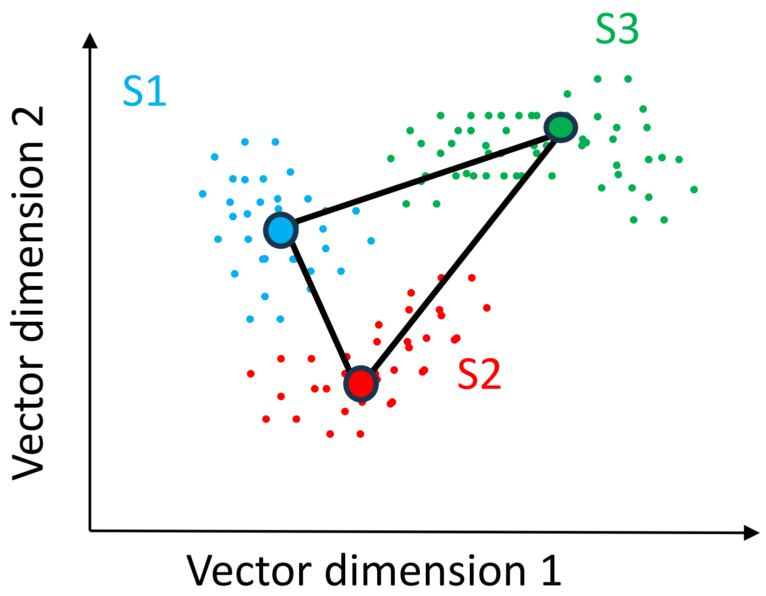
- Cannot separate the contribution of variability/noise (size of the cloud of point) and the contribution of the representation geometry (distance between centroids)

- What is measurement variability and what is neuronal variability ?

- Remains based on the specific assumptions of the classifier which will vary slightly across classification methods (SVM, centroid, etc…)

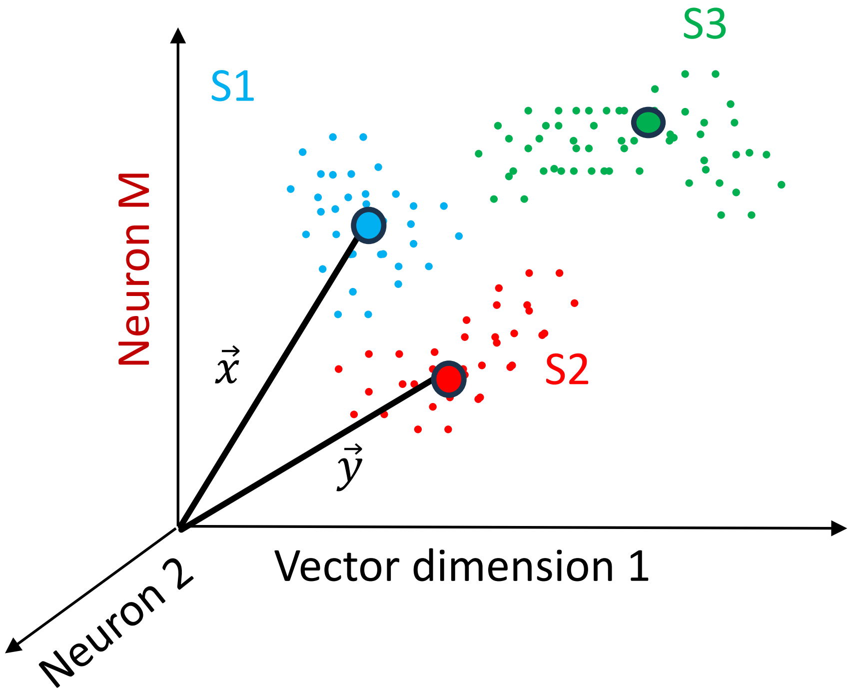
- Sensitive to overfitting and sample size.

Can we evaluate the geometry of the representation in a more generic manner ?



- The distance between the centroids is a key measure that quantifies the similarity between the representations, independant of noise. A change in the geometry indicates a transformation of the information format, e.g. across stages of the auditory system. **See later why this is important.**

Angular distances vs Euclidean distance



- With the **Euclidean distance**, the distance value depends on the size of the dataset, and on the amplitude of the signals.

- To compare patterns independent on amplitude, it is popular to use **angular metrics** that are not distances but that are normalized, such as **the Pearson correlation or the cosine similarity**

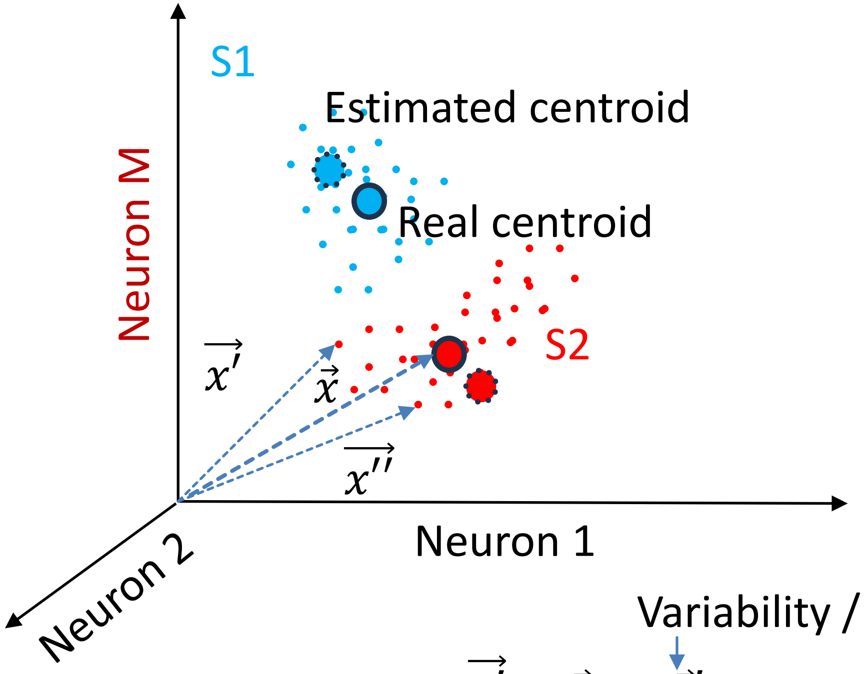
$$d = \sqrt{(\vec{x} - \vec{y}).(\vec{x} - \vec{y})}$$

$$s_{cos} = \frac{\vec{x}.\vec{y}}{\sqrt{\vec{x}^2.\vec{y}^2}}$$

$$\forall i \in (1, .., M), \qquad x_{zi} = x_i - \bar{x}$$

$$s_{cor} = \frac{\overrightarrow{x_z}.\overrightarrow{y_z}}{\sqrt{\overrightarrow{x_z}^2.\overrightarrow{y_z}^2}}$$

Systematic errors in centroid distance measures



Variability / noise

$$\overrightarrow{x'} = \vec{x} + \sigma\vec{\varepsilon}'$$

$$\overrightarrow{x''} = \vec{x} + \sigma\vec{\varepsilon}''$$

Noise estimate $\sigma^2 = \overrightarrow{x'}.\overrightarrow{x'} - \overrightarrow{x'}.\overrightarrow{x''} + O(\frac{1}{M})$

The mean localisation error is proportional to $\sigma/\sqrt{N}$ where σ is the standard deviation of measurement variability and *N* is the number of observations.

This error might be different across dataset.
It must be estimated and corrected if necessary

For the estimation of variability one can use the high dimensionality of the dataset of M neurons with M >> 1.

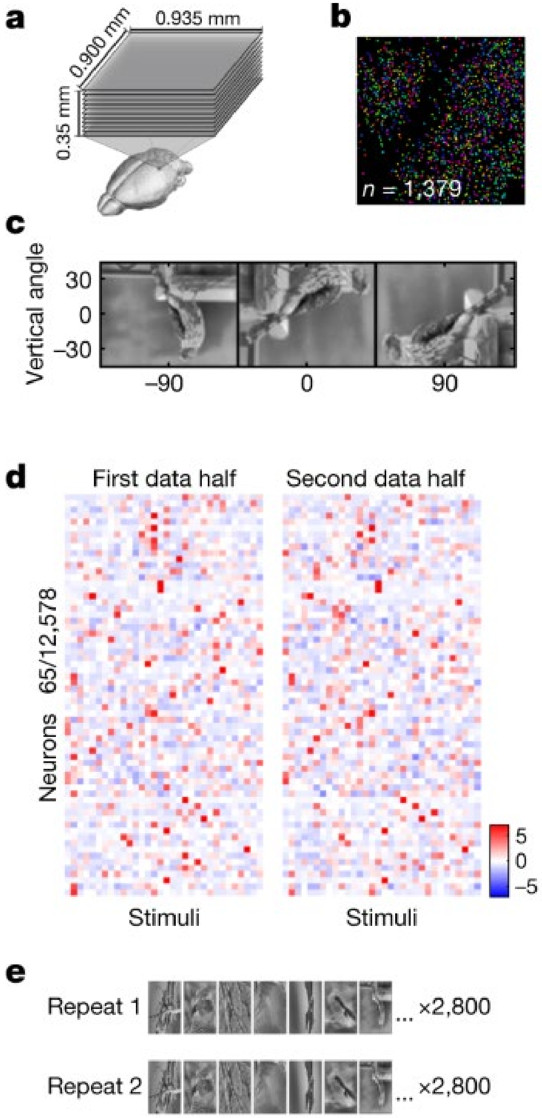
See Stringer et al. 2019

ARTICLE

https://doi.org/10.1038/s41586-019-1346-5

High-dimensional geometry of population responses in visual cortex

Carsen Stringer[1,2,6*], Marius Pachitariu[1,3,6*], Nicholas Steinmetz[3,5], Matteo Carandini[4,7] & Kenneth D. Harris[3,7*]



Geometry of visual representations in V1 with 10,000 neurons / recording

Analysis with systematic correction for trial-to-trial noise

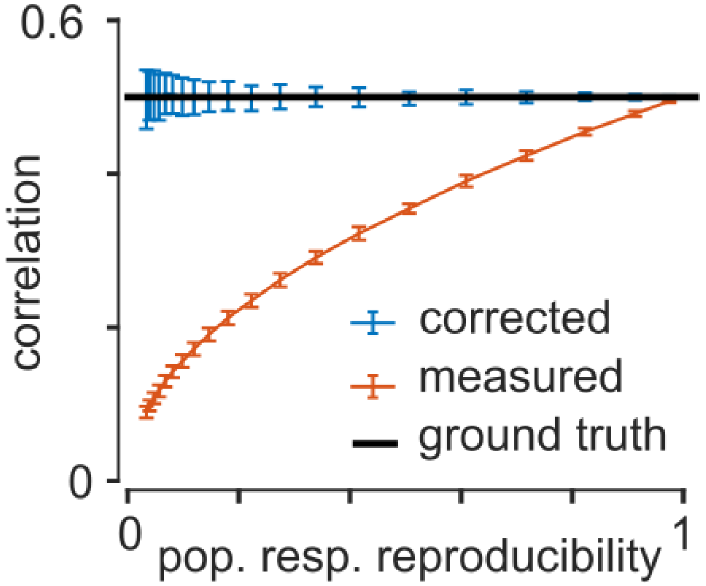
Noise-corrected correlation coefficient

Estimate of noise free Pearson correlation / cosine similarity

$$\rho_{\vec{x}\vec{y}} \approx \frac{\rho_{\vec{x}_r \vec{y}_{r'}}}{\sqrt{\rho_{\vec{x}_r \vec{x}_{r'}} \, \rho_{\vec{y}_r \vec{y}_{r'}}}}$$

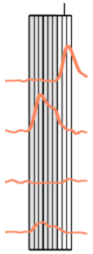
in the limit of large N_{neurons}

Noise-corrected estimate of correlation
(numerically & analytically validated)

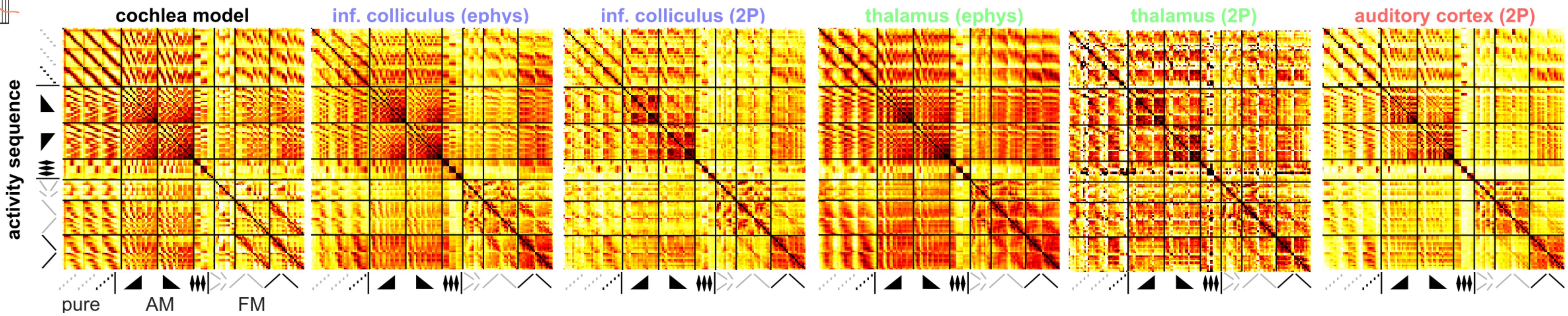


Bagur et al. BioRxiv 2022

Spatio-temporal representations along the auditory system measured through the noise-corrected correlation metric

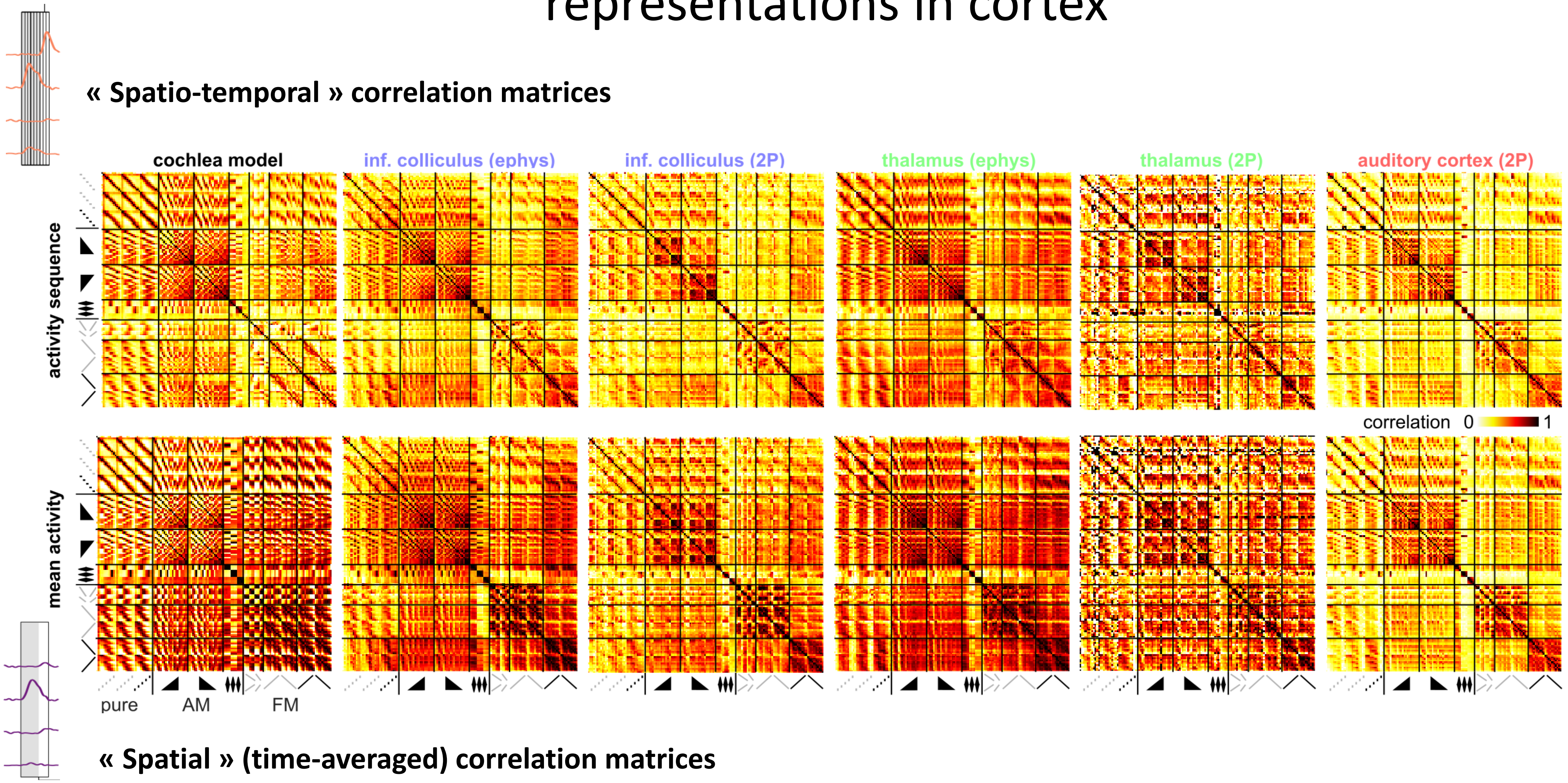


« Spatio-temporal » correlation matrices



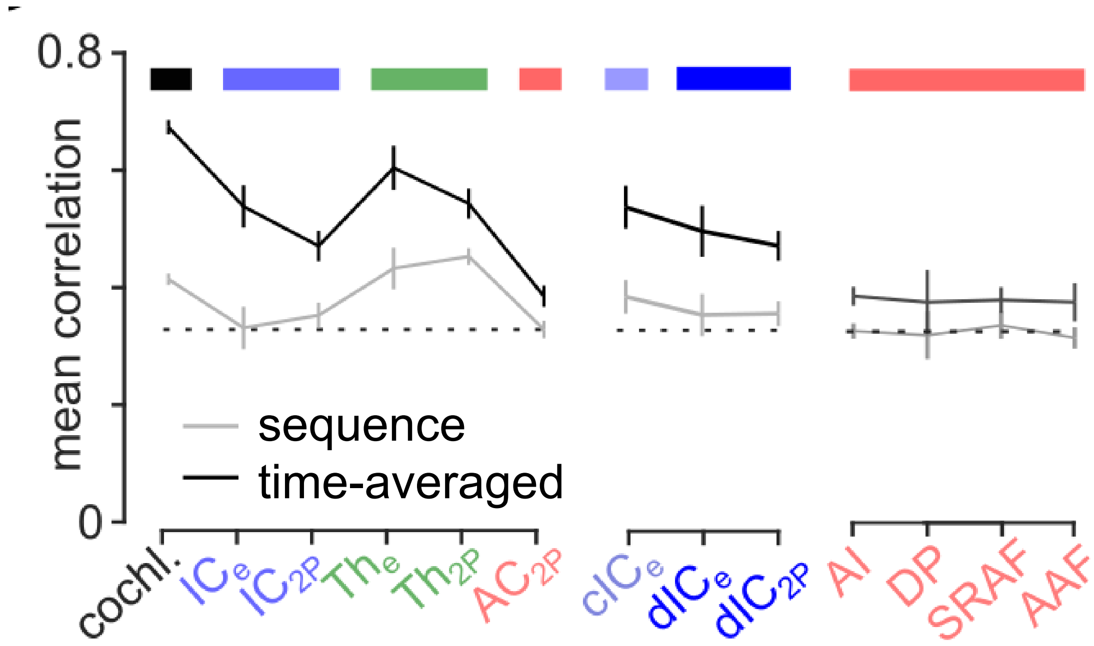
Convergence of spatio-temporal and spatial representations in cortex

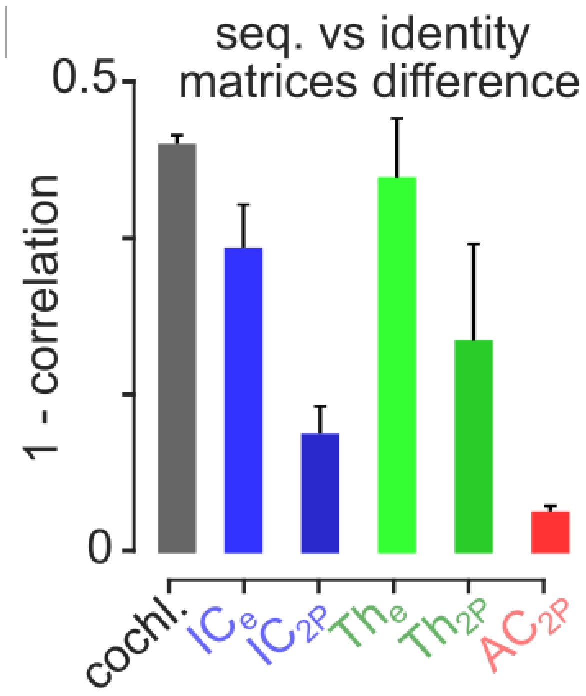
« Spatio-temporal » correlation matrices



« Spatial » (time-averaged) correlation matrices

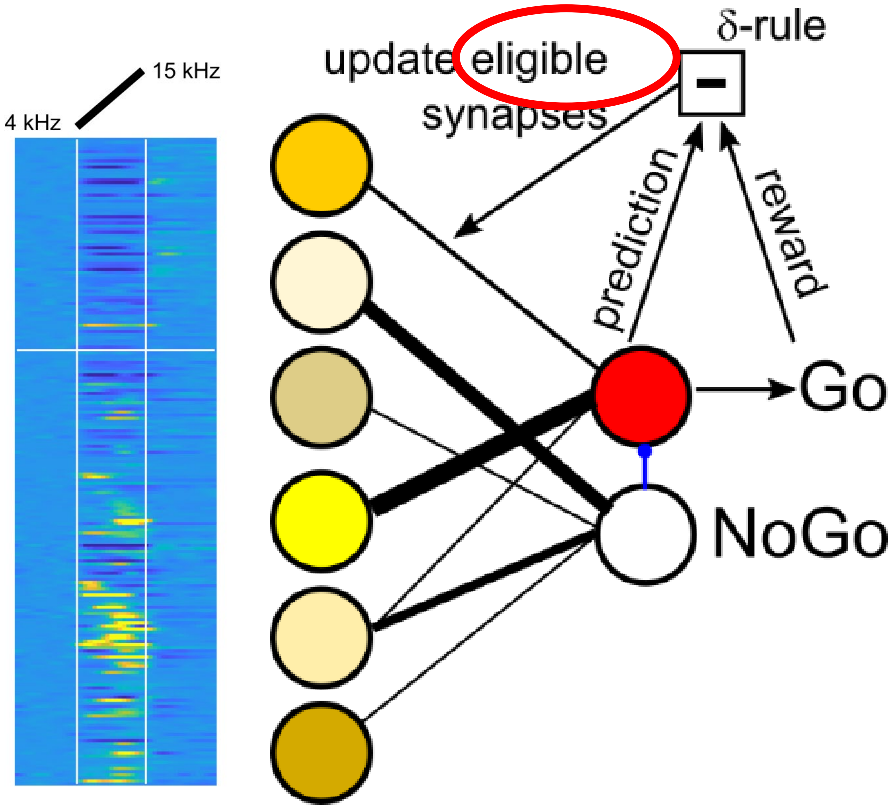
Convergence of spatio-temporal and spatial representations in cortex

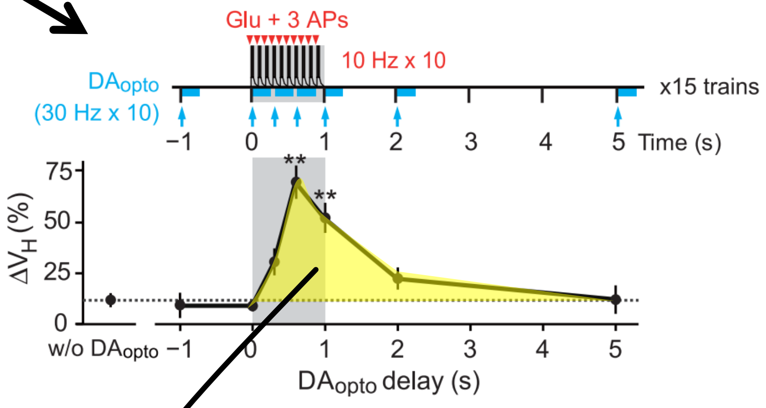




Why is this transformation important ?

Let's model a binary sound discrimination with a perceptron and a Hebbian learning rule

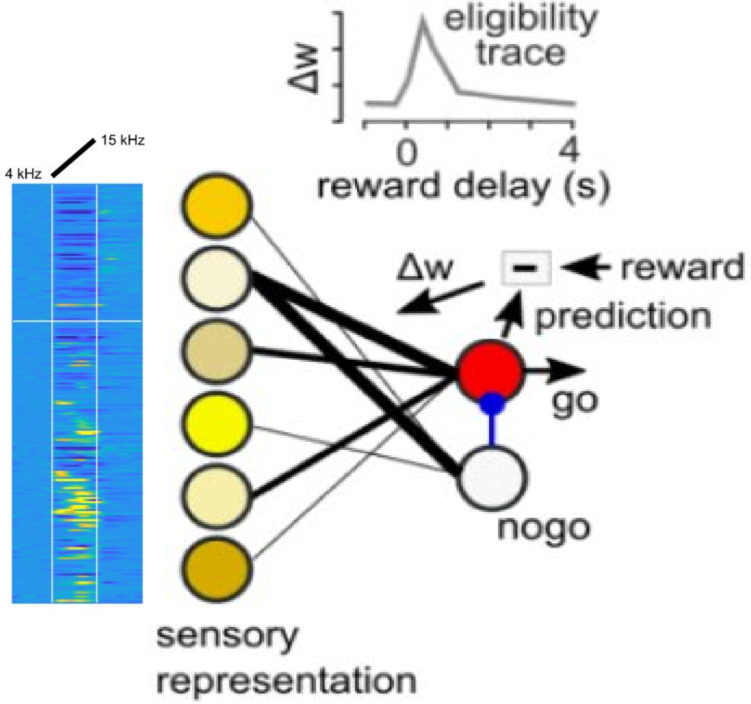




Eligibility = < act$_{pre}$ (t) × act$_{post}$ (t) >$_t$

Temporal average until reward to bridge time scales, as observed in striatum (~1s time window), see Gerstner et al. Frontiers 2018, for review

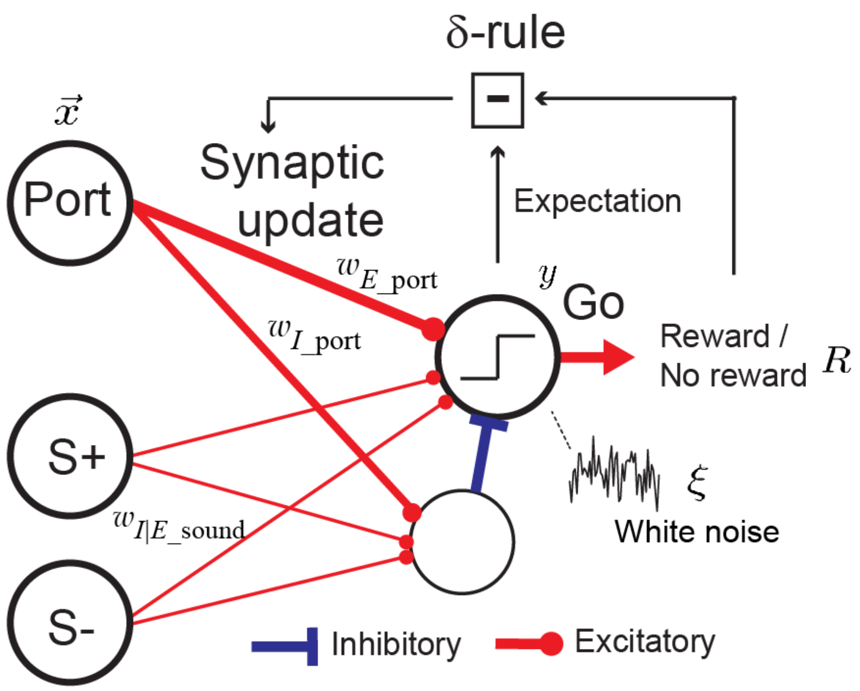
Let's model a binary sound discrimination with a perceptron and a Hebbian learning rule

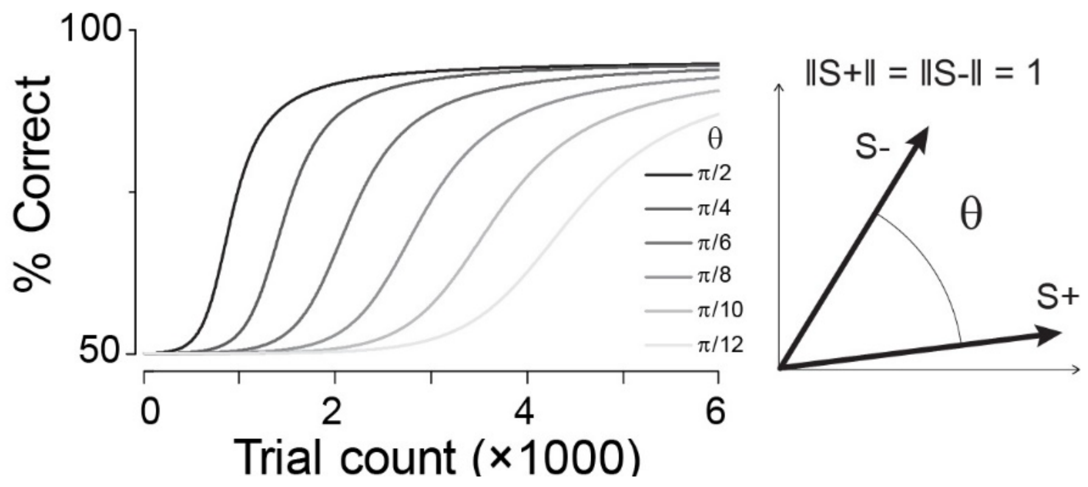


- **Hebbian**: The model flags synapses that were activited together with the firing of the neurons that drive motor response or prevent response.

- **Reinforcement learning**: It updates the synaptic weights of flagged synapses that led to an unexpected favorable outcome (eligibility trace mechanism observed in striatum, review by Gerstner et al. 2018)

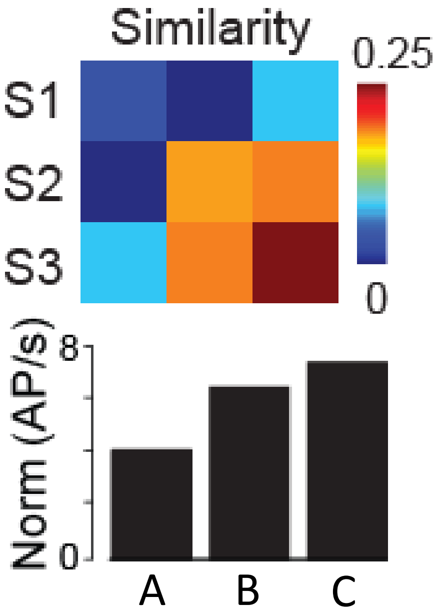
The model learns to discriminate spatial patterns and learn slower when representations are more similar



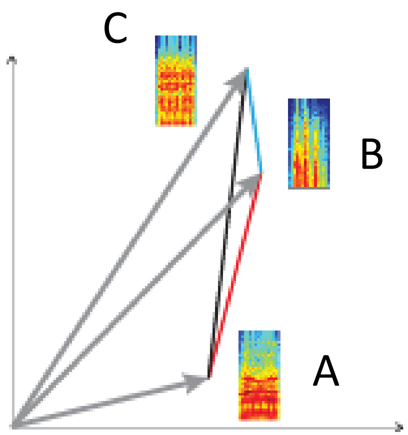


Intuition: More similar representations provide less discriminative neurons & spikes, so less opportunities to learn

Experimental evidence for the impact of similarity

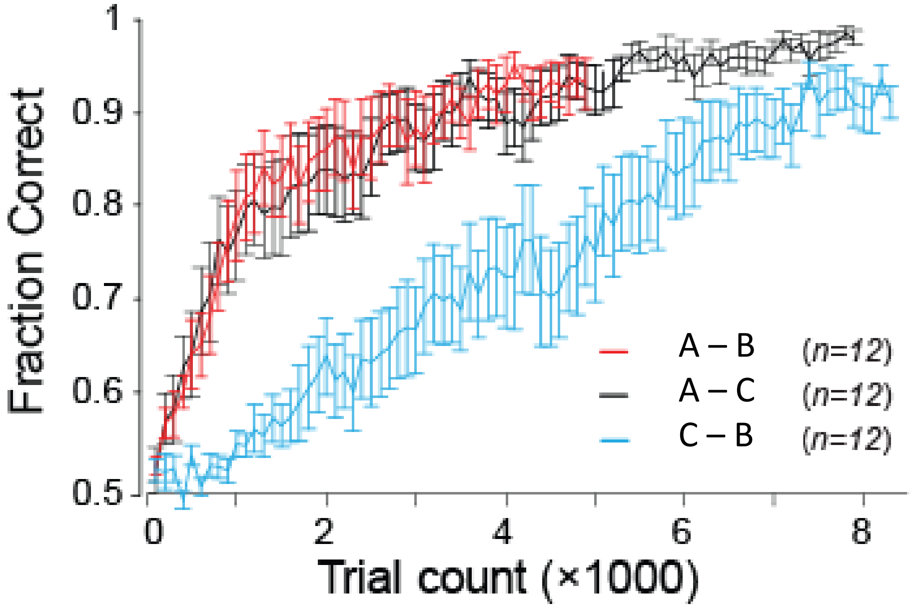


Cortical similarity



Population vectors
(sketch)

Learning behavior for different
sound pairs



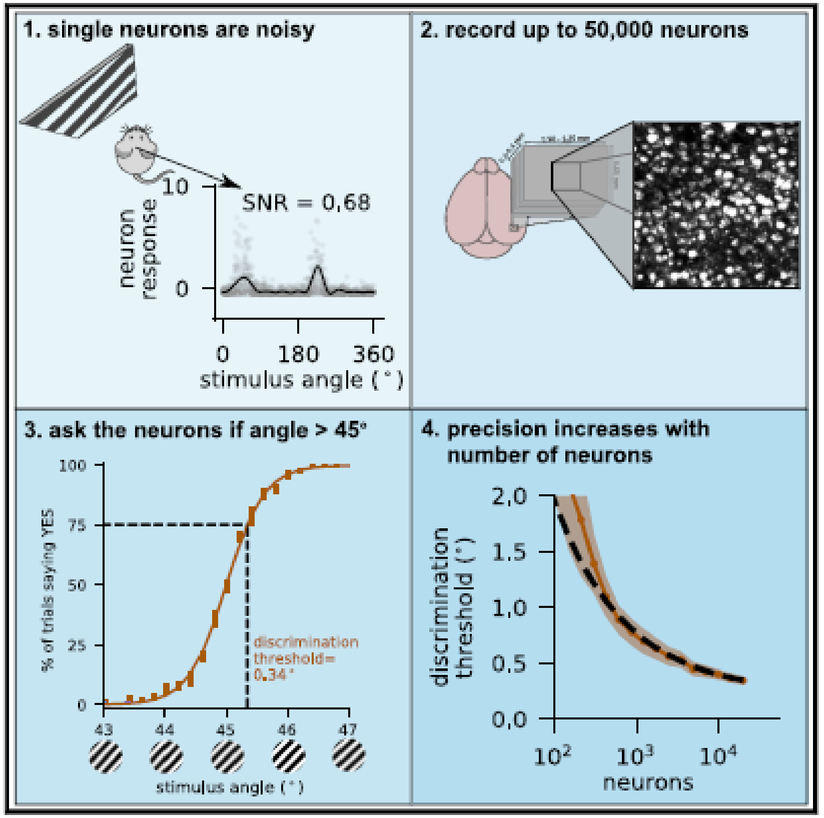
Bathellier et al. 2012

Why is this important?

- Lack of behavioral discrimination may be a cinetic problem (too long to learn)
- Impossibility to show discrimination despite sufficient information in the brain

High-precision coding in visual cortex

Graphical abstract



Authors

Carsen Stringer, Michalis Michaelos, Dmitri Tsyboulski, Sarah E. Lindo, Marius Pachitariu

Correspondence

pachitarium@janelia.hhmi.org

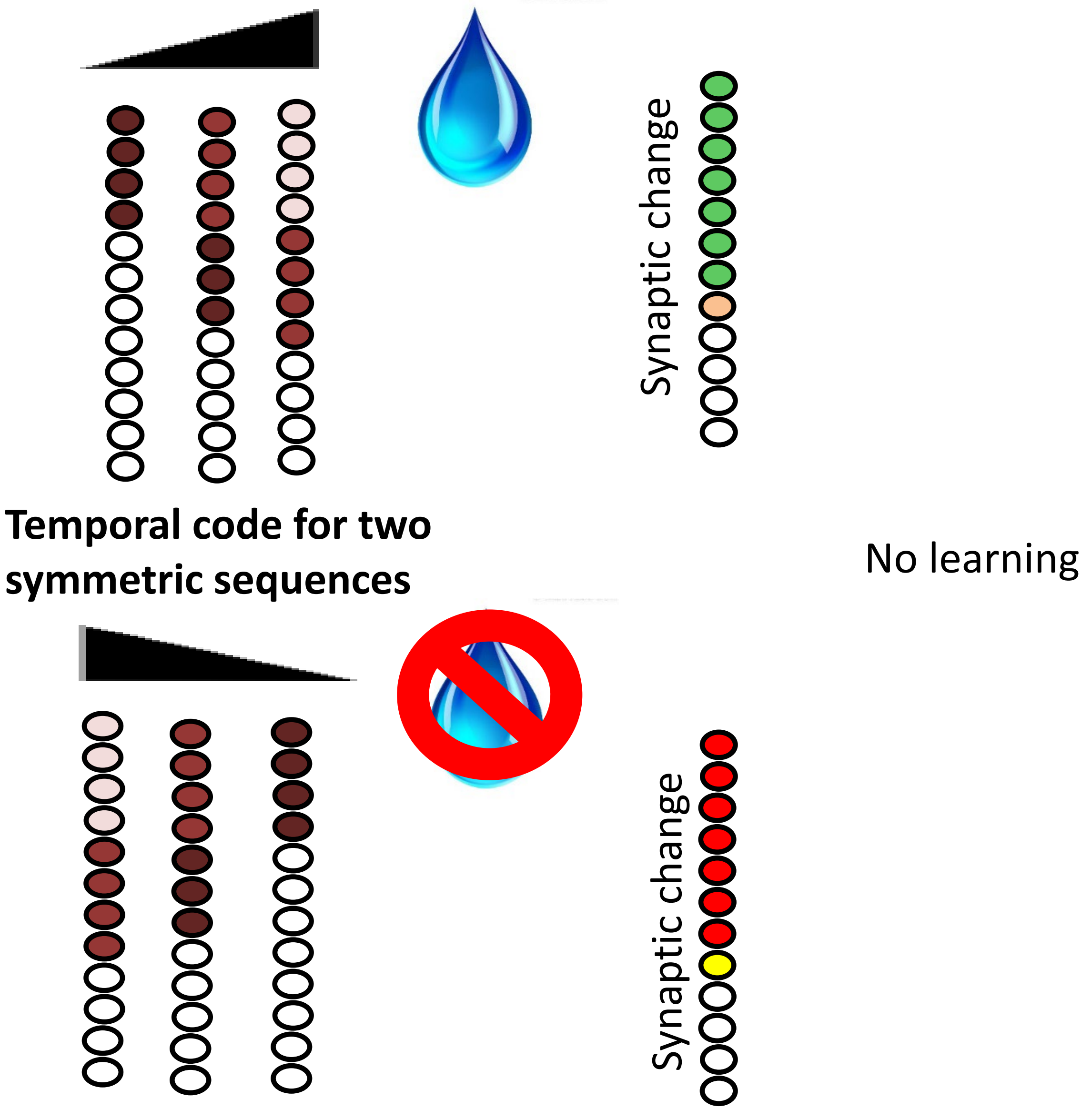
In brief

Large-scale recordings in mouse primary visual cortex and higher order visual areas uncover neural representations more precise than behavioral discrimination thresholds, suggesting visual perception is limited by non-sensory brain networks.

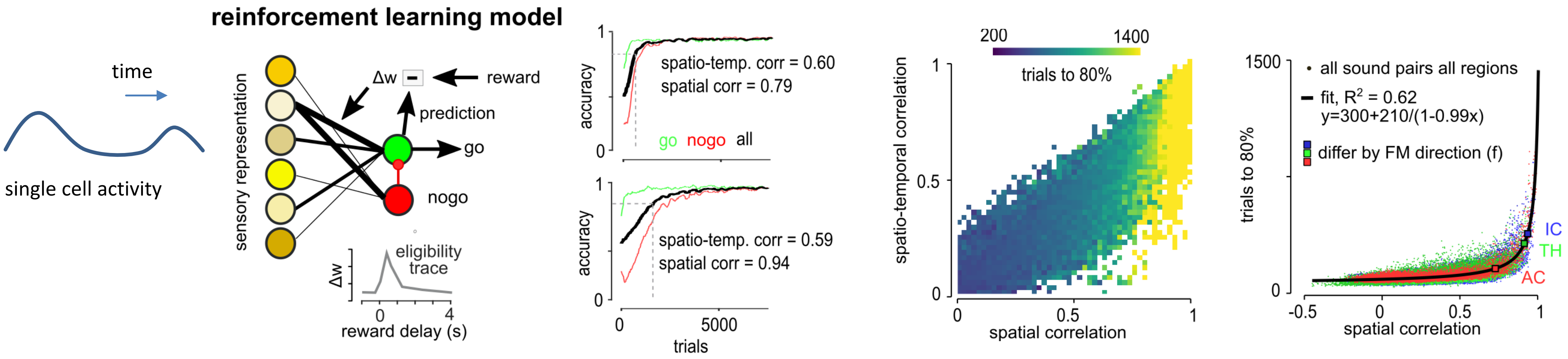
Max discrimination behaviorally: 10 - 30°

Basic Hebbian learning cannot discriminate symmetric sequences

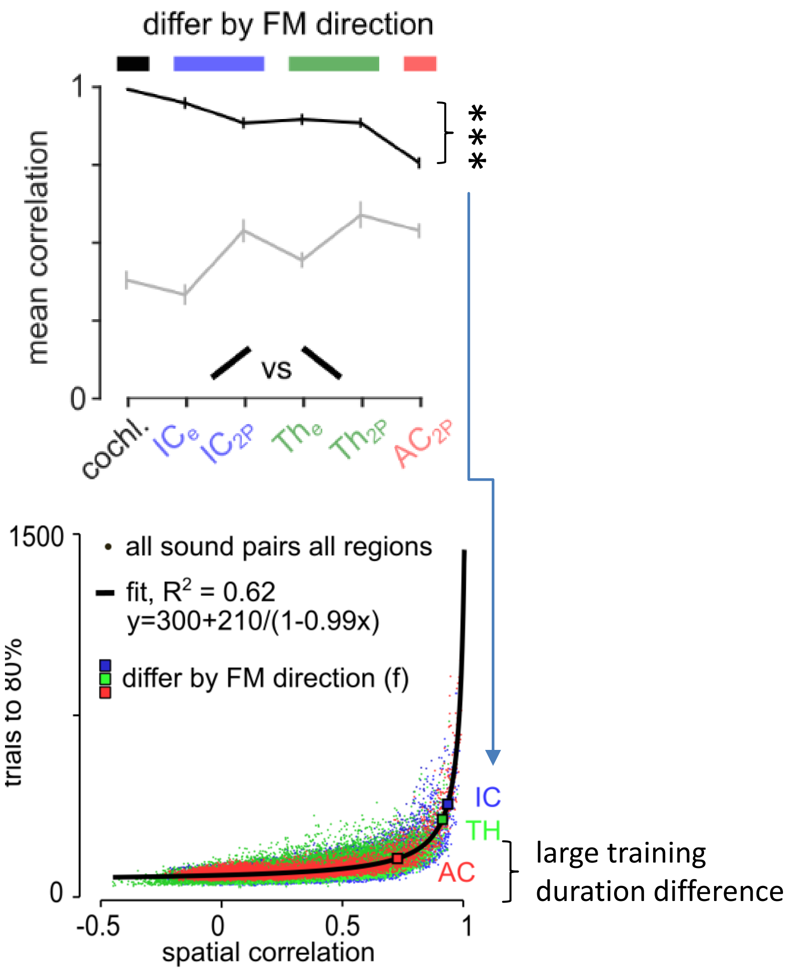
The synapses that are active in one direction are also active in the other direction and will both lead to correct and incorrect Go.

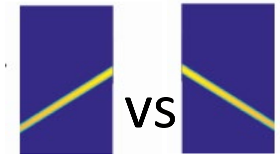


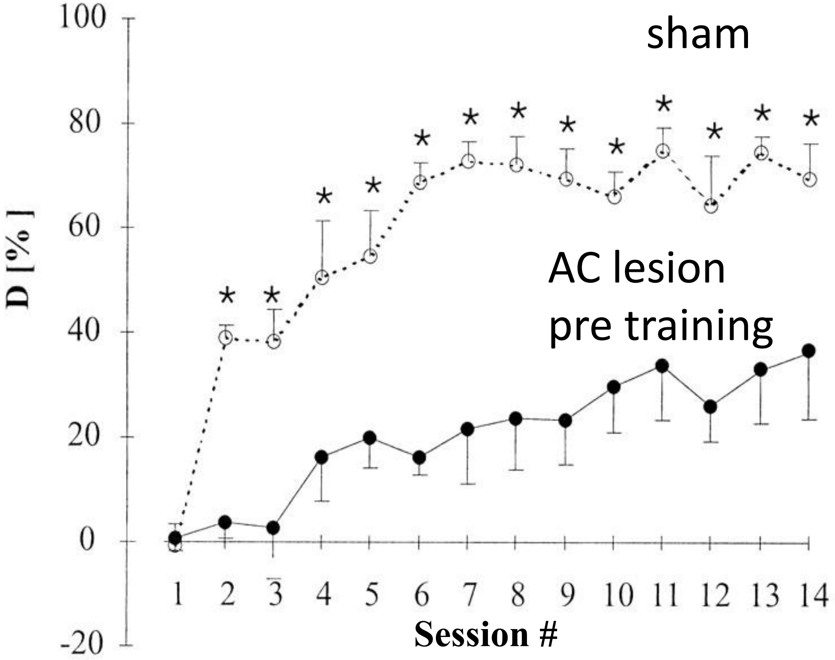
Quantification with the discrimination model: the spatial code determines learning speed

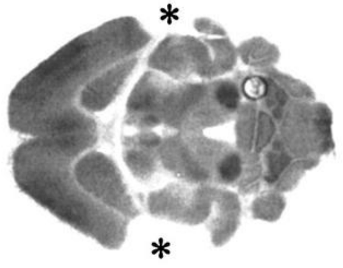


Decorrelation of time-averaged representations explain learning speed difference with and without cortex for time-symmetric chirps



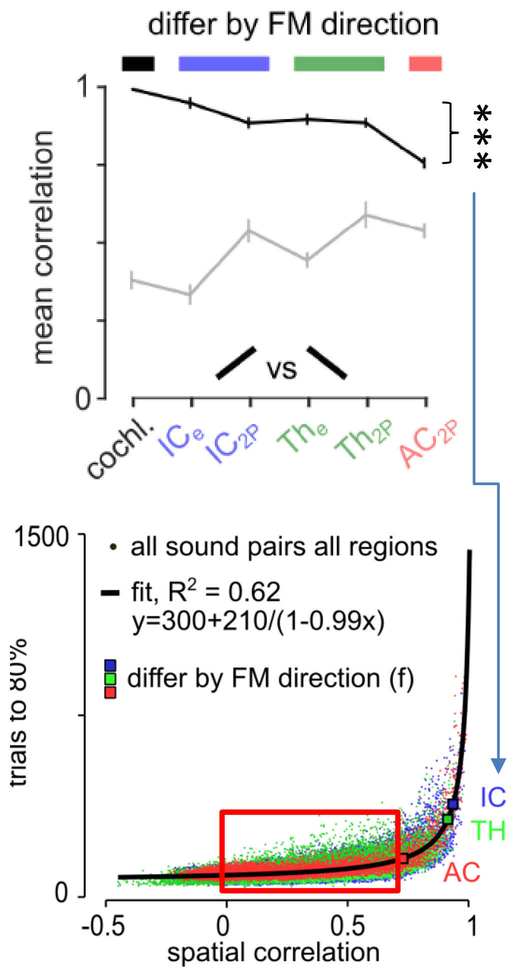


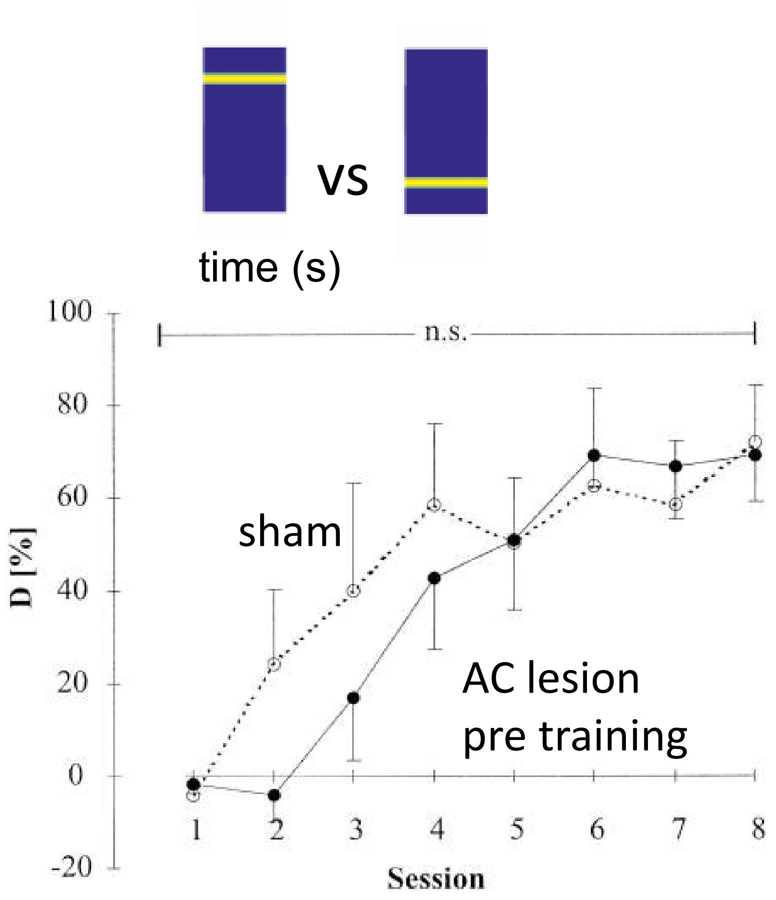


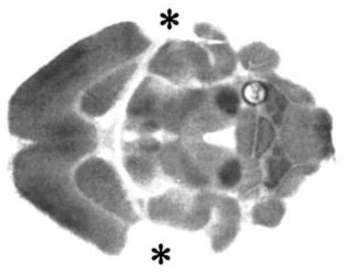


Ohl et al. 1999
see also Dalmay et al.

Little effect of cortical silencing on learning speed for sounds that are everywhere less correlated (e.g. pure tones)

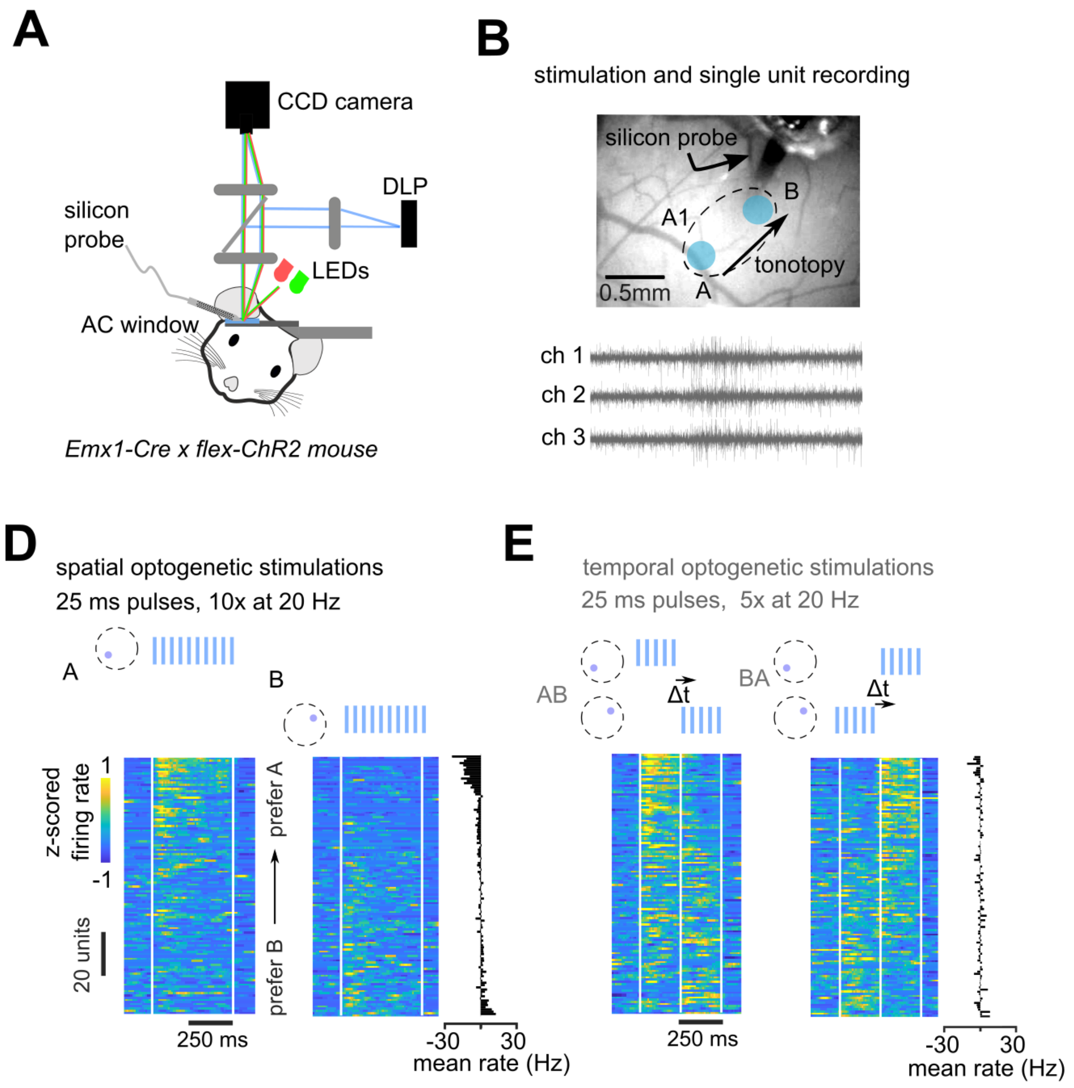




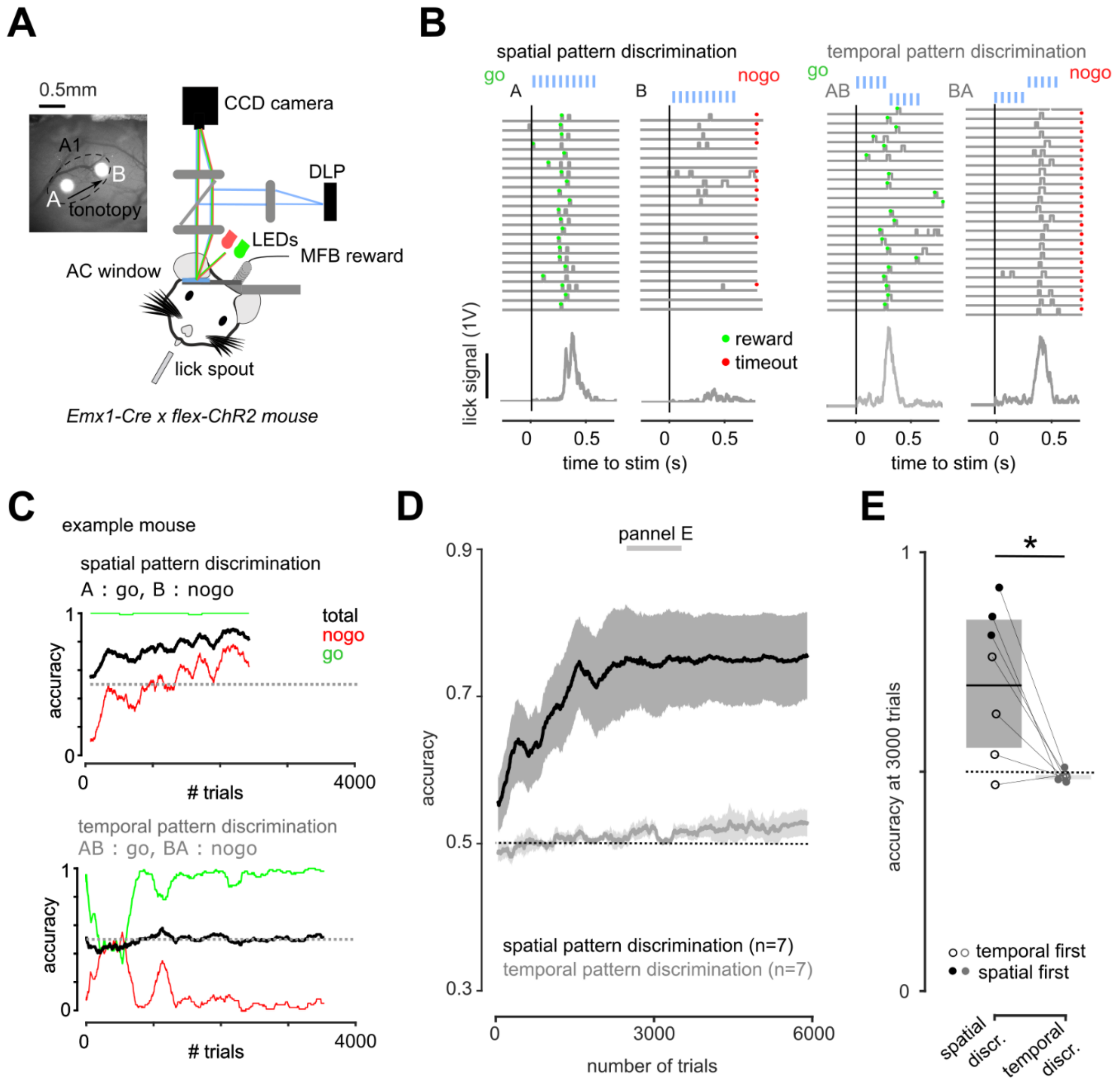


Ohl et al. 1999
see also Dalmay et al. 2019

Causal experiment: artificial spatial and temporal patterns

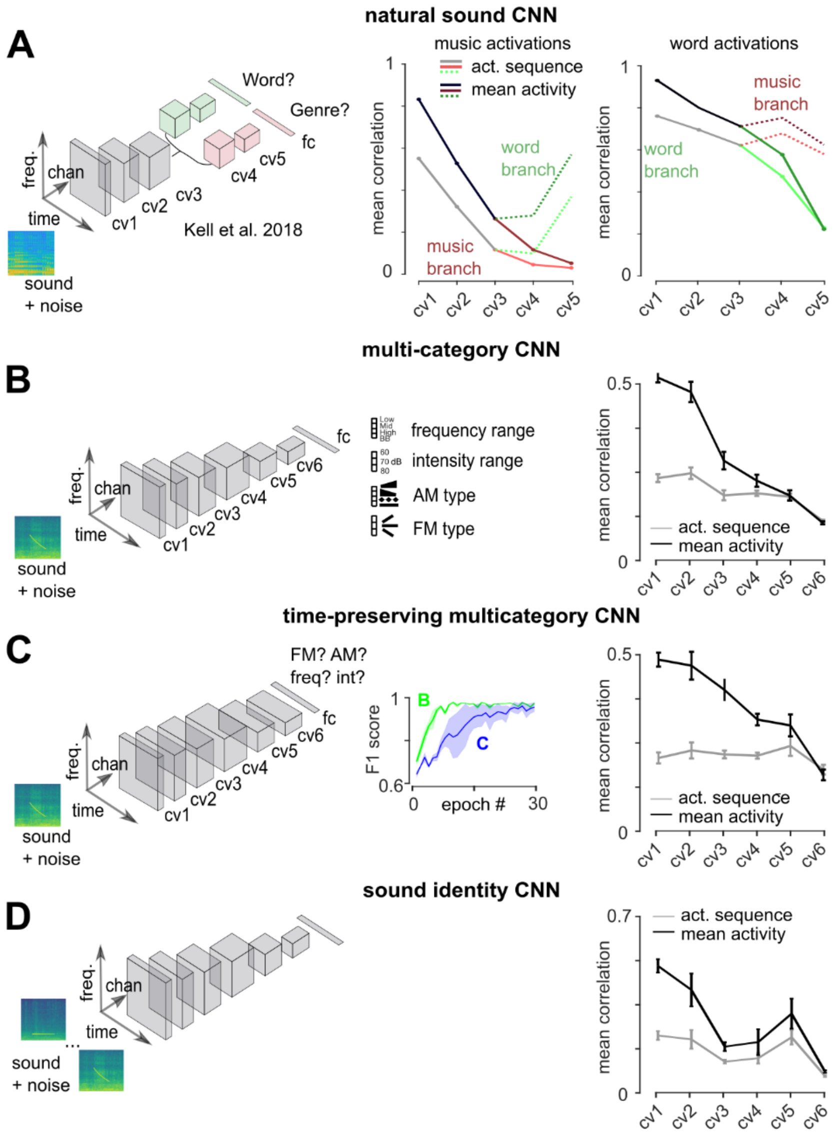


Causal experiment: discrimination of pure temporal patterns is extremely slow

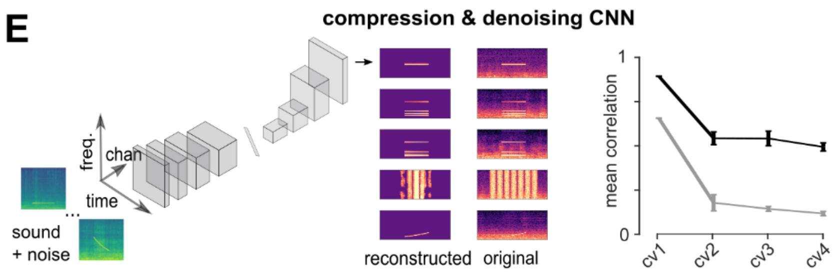


A spatial code for temporal cues in categorisation CNNs

CATEGORISATION



INFORMATION TRANSFER / COMPRESSION



Conclusions

- In auditory cortex (and not before), sequence information over 500ms can be efficiently retrieved from spatial population representations.

- This does not reflect a loss of temporal resolution because sequence information is actually preserved in AC.

- This is important for learning because learning rules for sensory-motor associations have a hard time using temporal information. Our observations constrain the type of learning rules involved in sensory-motor associations (classic Hebbian and not more complex such as tempotron).

- This explains differential involvement of cortex in sound discriminations.

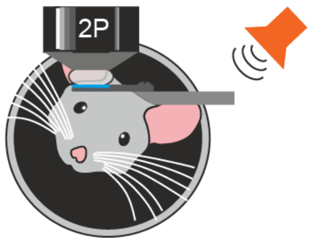
- The same effect occur in CNNs which learn more complex associations / categorisation.

What happens in auditory cortex when there is no perception under anesthesia ?

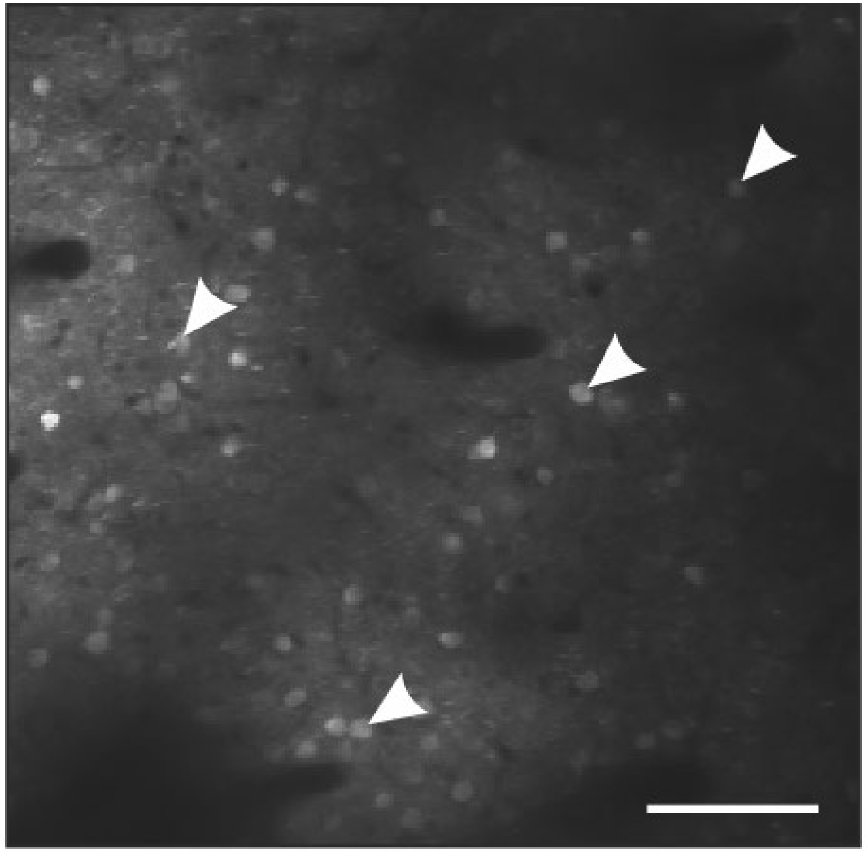
 Anton Filipchuk  Alain Destexhe

Filipchuk et al. 2022

Imaging the same neurons during wakefulness & isoflurane anesthesia

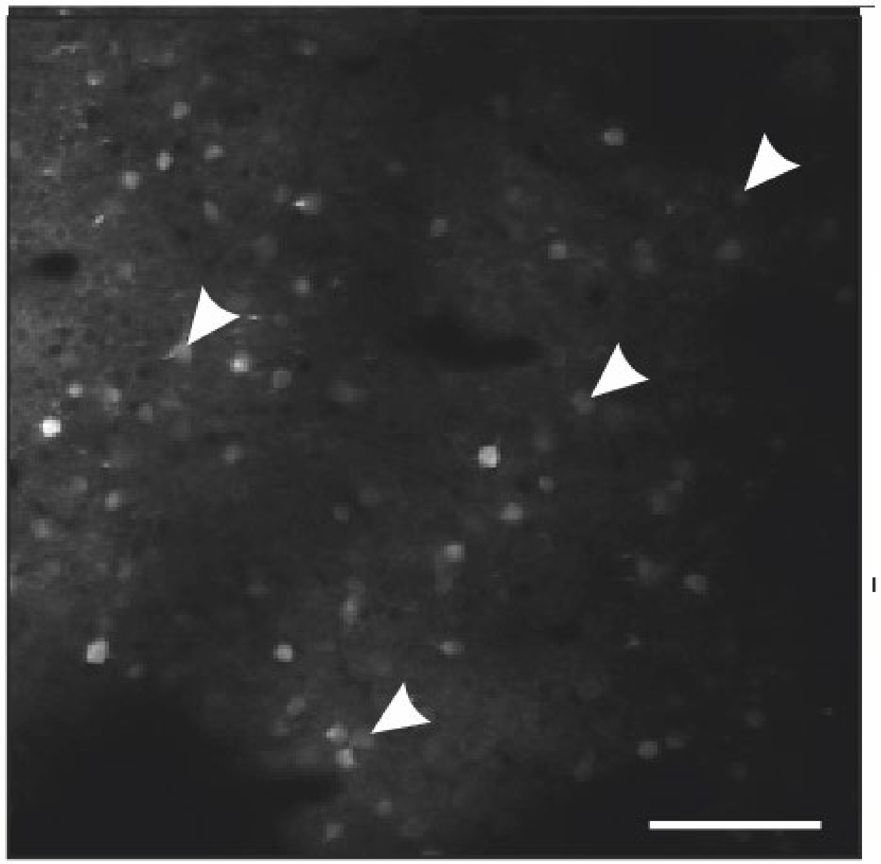


Awake

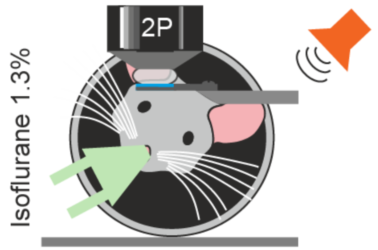


1 mm
Auditory cortex

Anesthesia

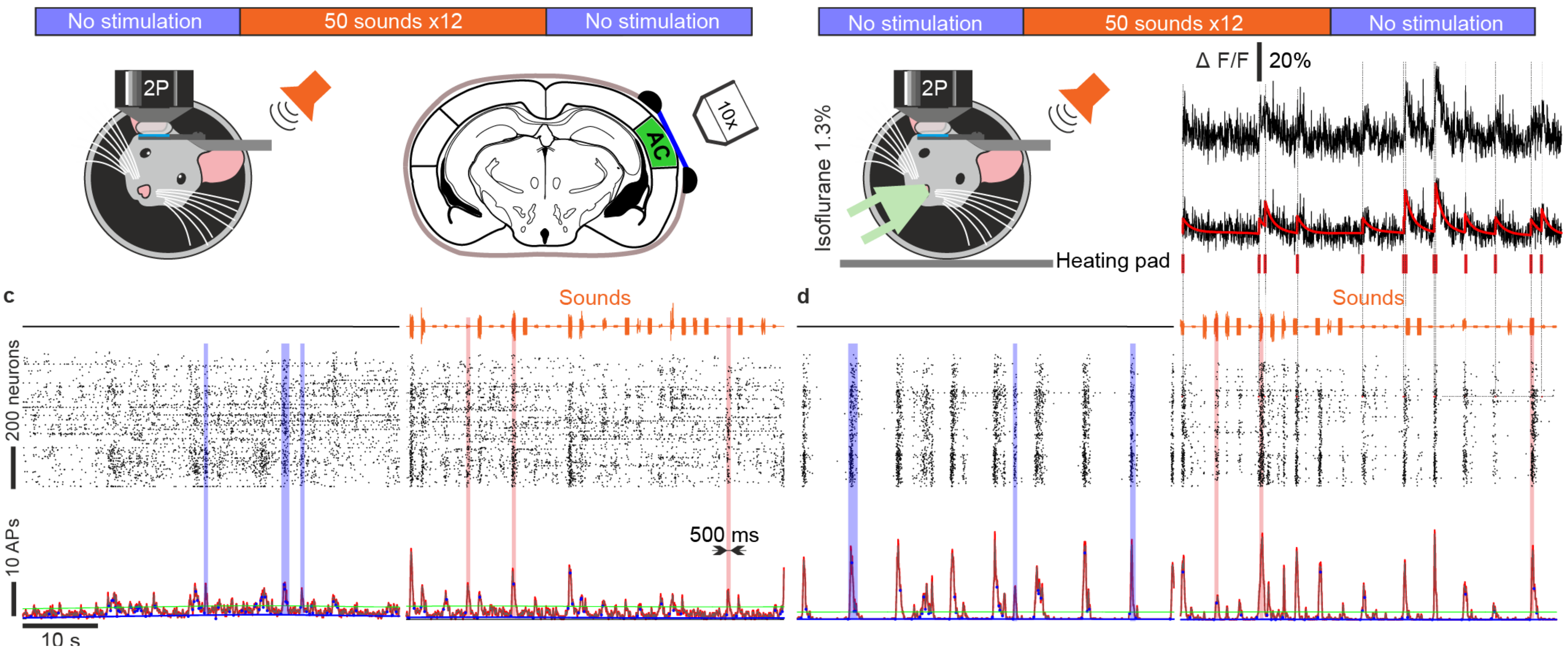


Images of the STD of each pixel



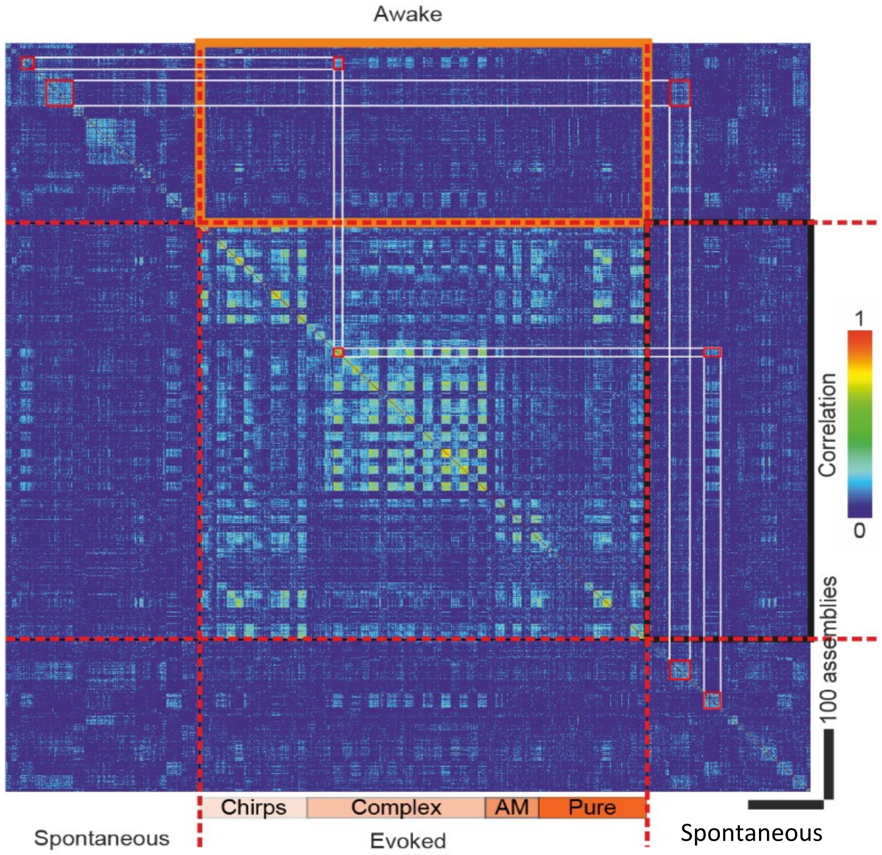
Heat pad

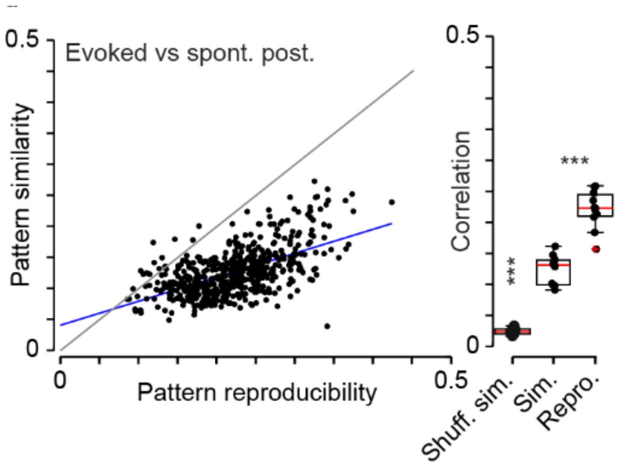
Population activity during wakefulness & anesthesia



ML Spike
for AP
identification

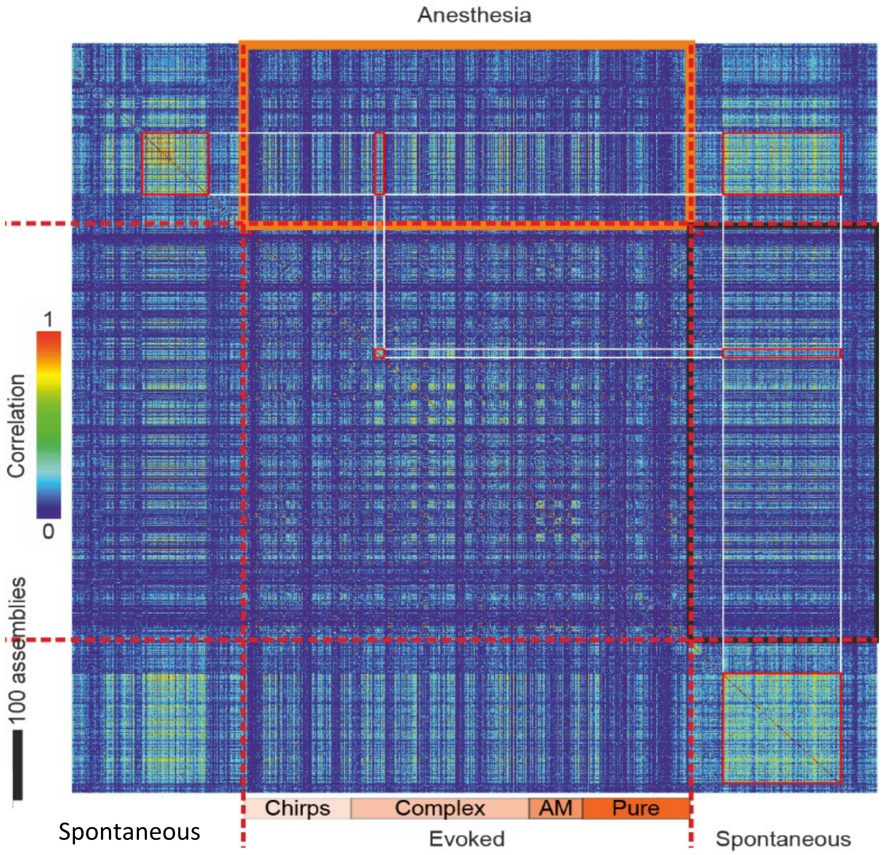
Distinct spontaneous and evoked population events in the awake state

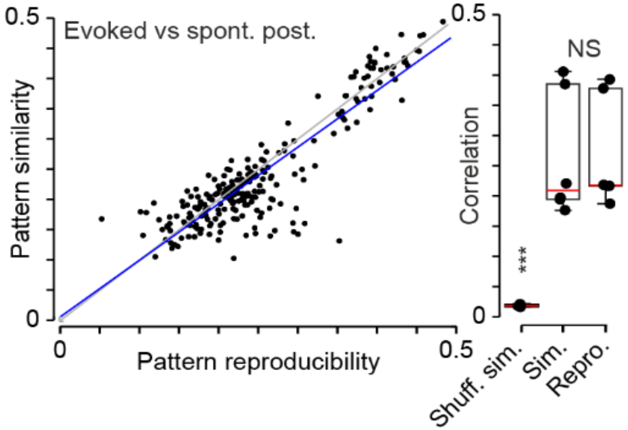




✓ Stringer at al; 2019, visual cortex

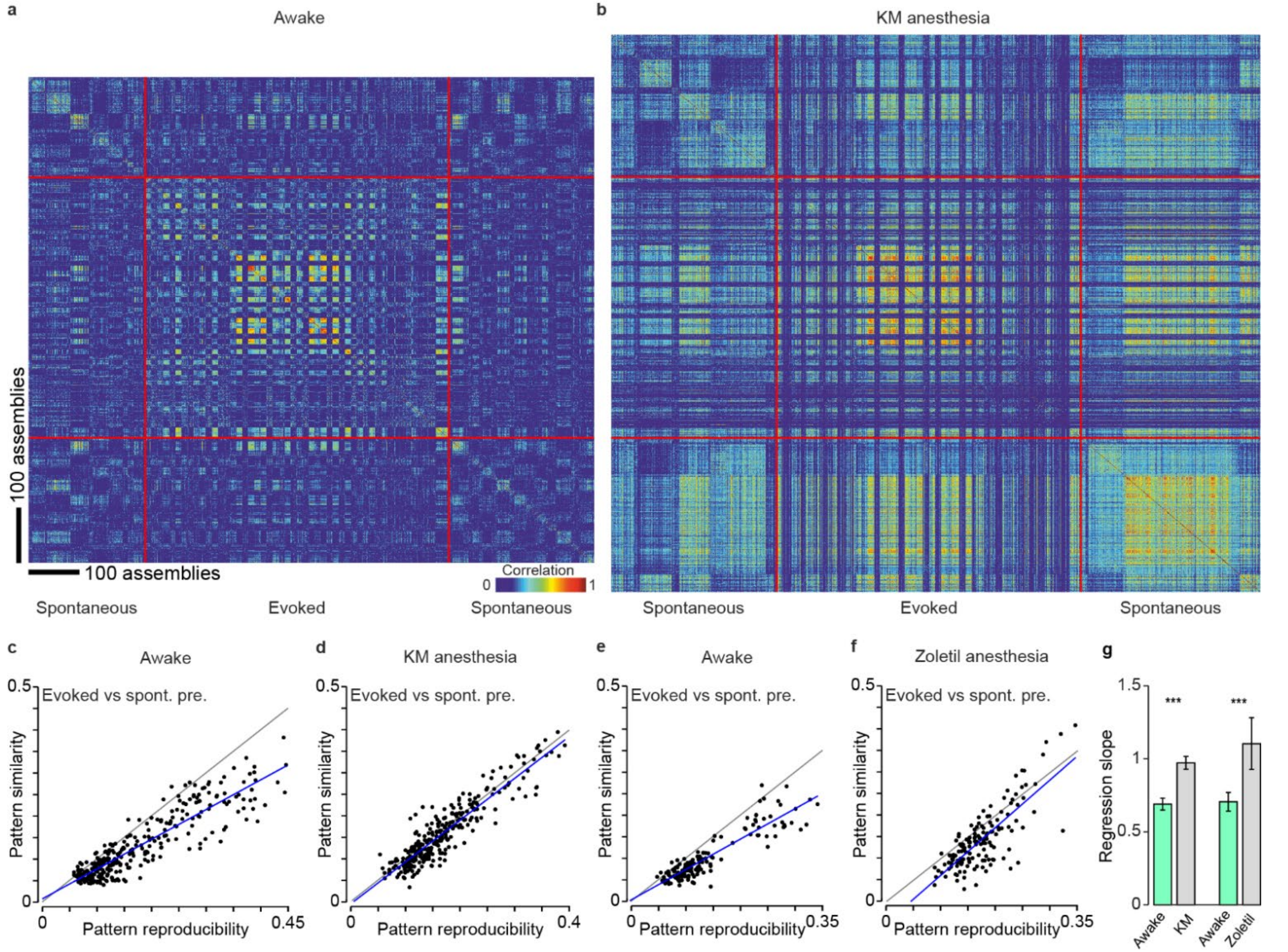
Highly similar spontaneous and evoked population events in the anesthetized state



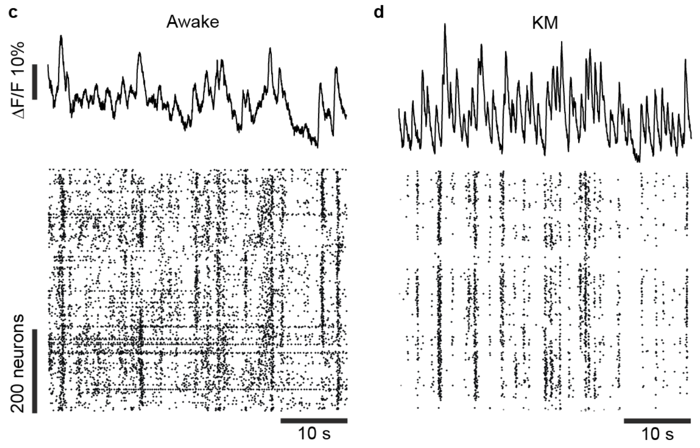


✓ Arieli, Norena, Harris labs (VSD & spikes)

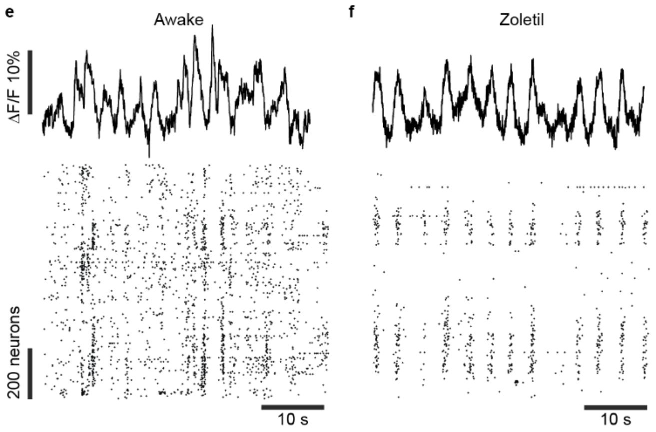
Same effect for ketamine medetomidine & zoletil



KM (50 & 1 mg/kg)

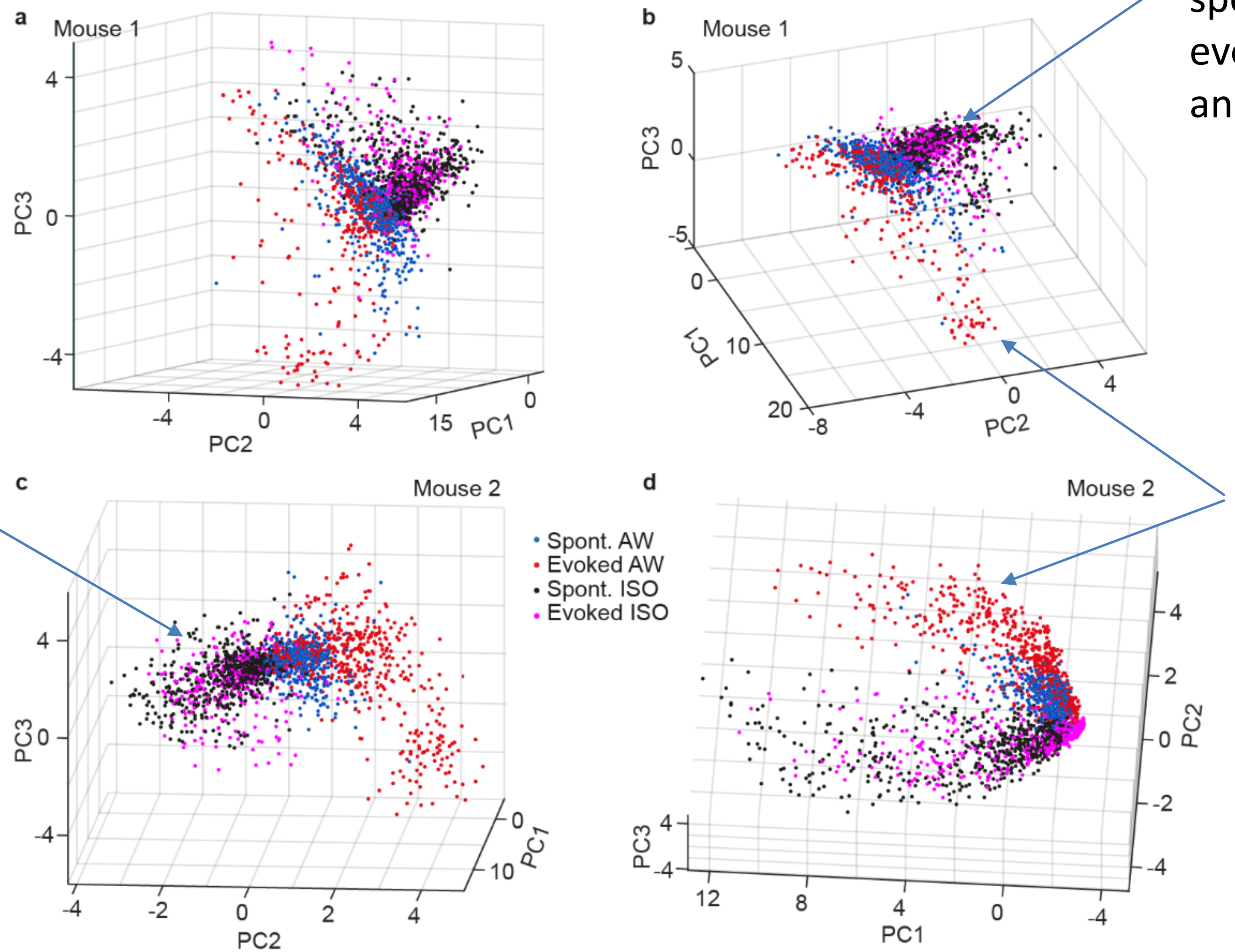


Zoletil (70mg/kg)



Under anethesia, sound representation are in the subspace of spontaneous activity

Overlap of spontaneous and evoked in anesthesia

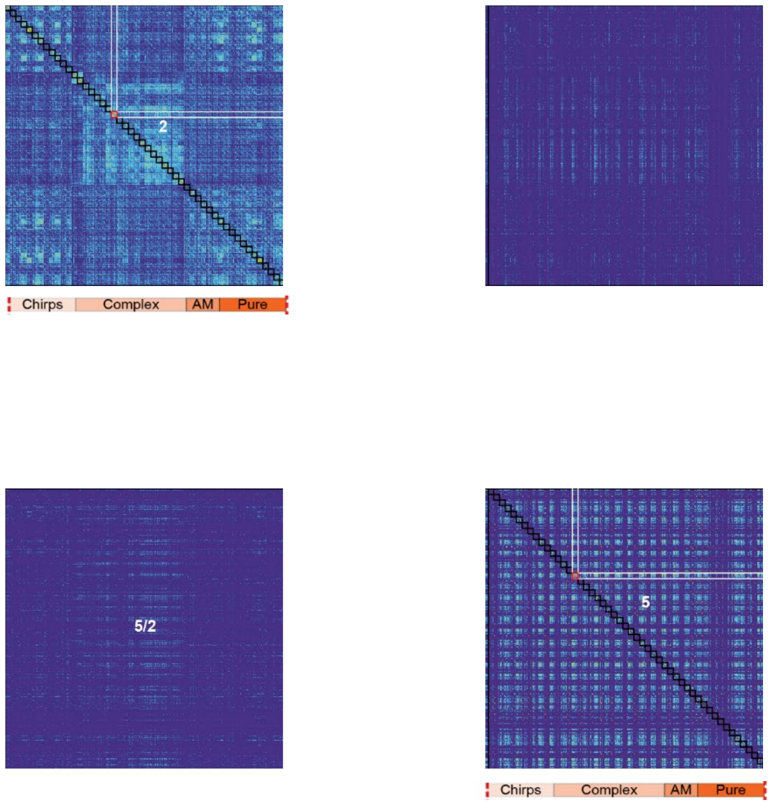


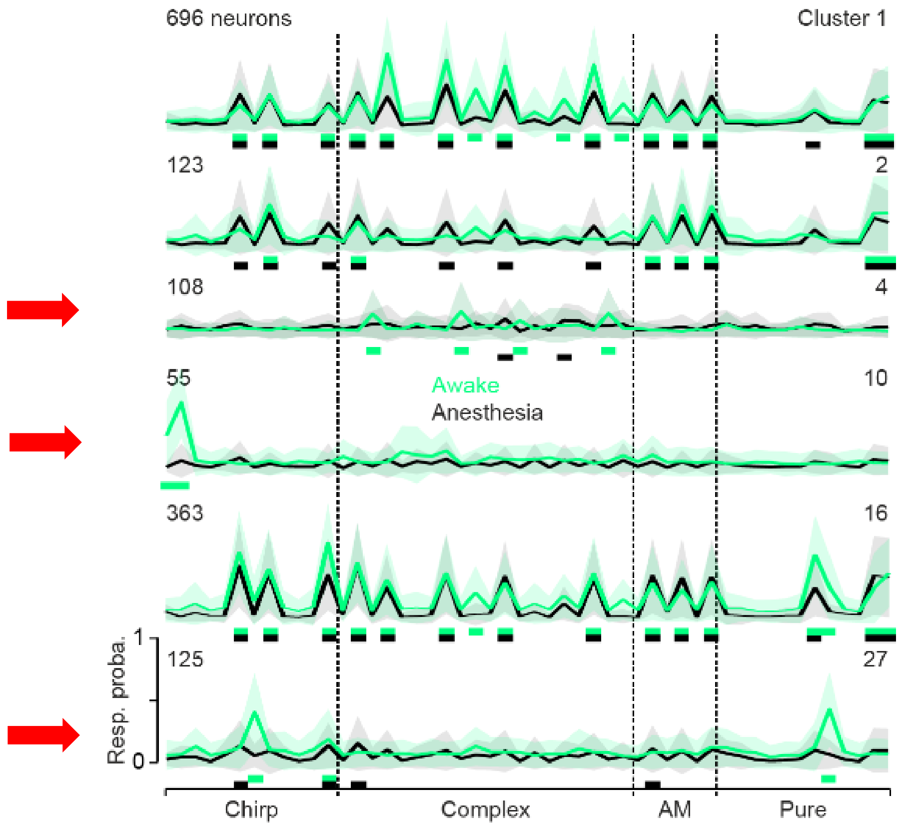
Overlap of spontaneous and evoked in anesthesia

Specific evoked responses in awake

Sound representations also change under anesthesia

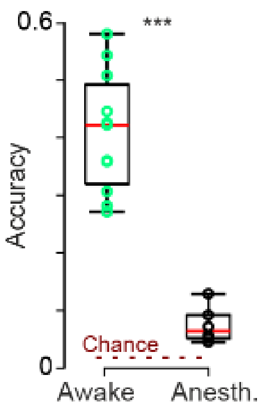
Awake | Isoflurane

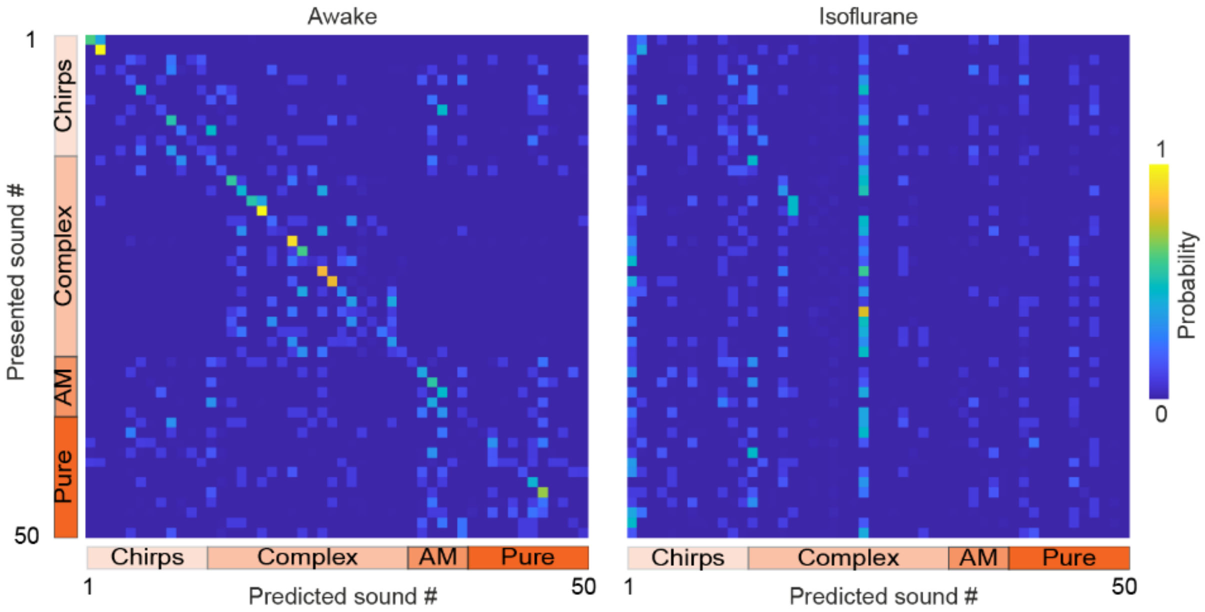




Sound information drastically decreases in anesthesia

classifier output





Conclusions 2

In anesthesia:

- Reformating of population activity
- Drastic loss of information & loss of sparse responses
- High similarity between sound-evoked and spontaneous activity patterns

- *Maybe we don’t perceive sounds in anesthesia because it is hard to discriminate sound responses from spontaneous activity ?*

Acknowledgements

Bathellier lab

Scientists
Sara Jamali
Sophie Bagur
Etienne Gosselin
Antonin Verdier
Elena Kudryavitskaya
Allan Muller
Matteo Pisoni
Clément Lefèvre

Technical support
Noémie Dominique
Déborah Groussard

IdA imaging facility
Yannick Goulam

Alumni
Joanna Schwenkgrub
Simone Azeglio
Anton Filipchuk
Anthony Renard
Jacques Bourg
Sebastian Ceballo
Evan Harrell
Thomas Deneux
Zuzanna Piwkowska
Alexandre Kempf

Funding

Collaborations
Alain Destexhe (NeuroPSI)
Julien Bouvier (NeuroPSI)
Dan Shulz (NeuroPSI)
Jean Marc Edeline (NeuroPSI)
Nicolas Michalski (IdA)

Karim Benchenane (ESPCI)
Srdjan Ostojic (ENS Paris)
Stéphane Dieudonné (ENS Paris)

Robert Prevedel (EMBL)

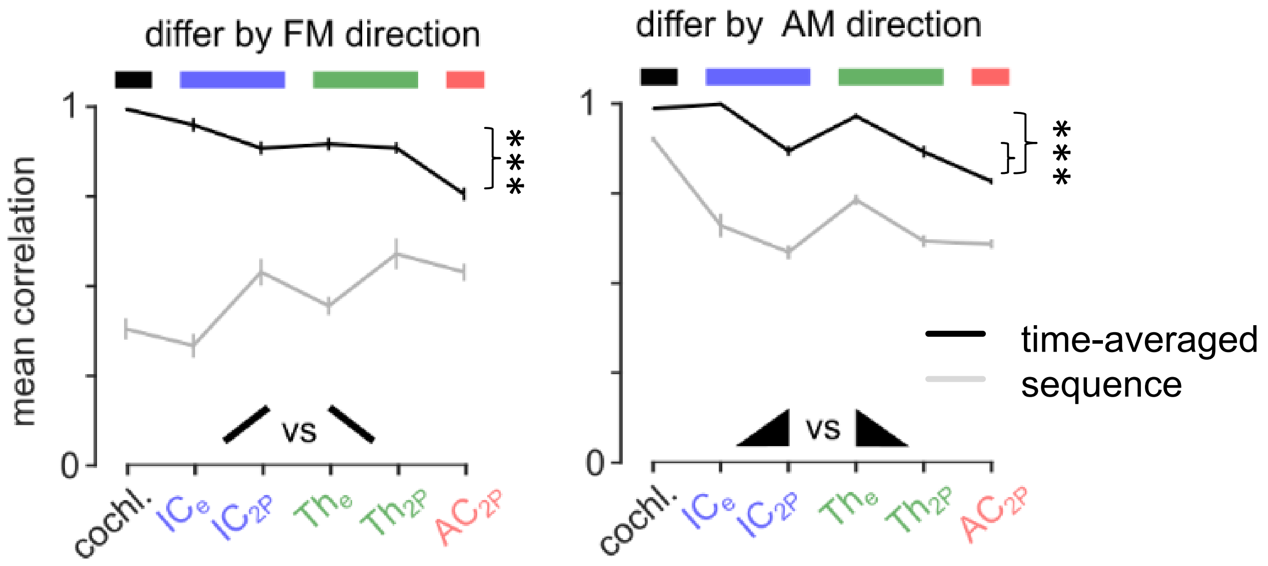
Timo Van Kerkoerle (NeuroSpin)
Stanislas Dehaene (NeuroSpin)

Hearlight
Charles Rezaei (Mines St Etienne)
Keith Mathieson (U. Strathclyde)
Tania Barkat (U. Basel)
John DeMello (NTNU)

Companies
NOVAGAN (Hearlight)
KARTHALA SYSTEMS

Thanks to all staff at the Institut Pasteur & Institut de l'Audition

Time symmetric sounds are more decorrelated in cortex for mean firing rate representations



Trial-to-trial variability is very different across datasets