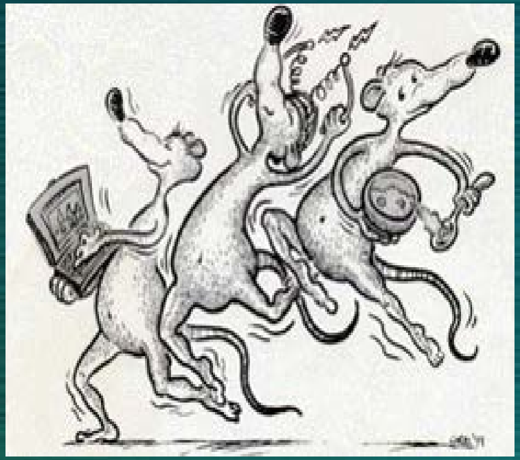
Signal conditioning for learning in the wild

Ayon Borthakur (ab2535@cornell.edu)

Thomas A. Cleland (tac29@cornell.edu)

Computational Physiology Laboratory
Cornell University

http://cplab.net



The olfactory “hard problem”: *learning in the wild*

Stronger odor of interest
Low background interference

Weak odor of interest
High background interference

Odor-specific sensor activation patterns are severely disrupted by interference from other environmental odorants, and by other uncontrolled physical variables.

What is learning in the wild ? Why is it hard ?

Data sampling in the wild

- Variable environment: unpredictable concentrations, overlapping odors, temperature, humidity…
- Incoming data may have missing values.
- Training sets may not be labelled.
- Sensor drift (over time, or due to contamination).

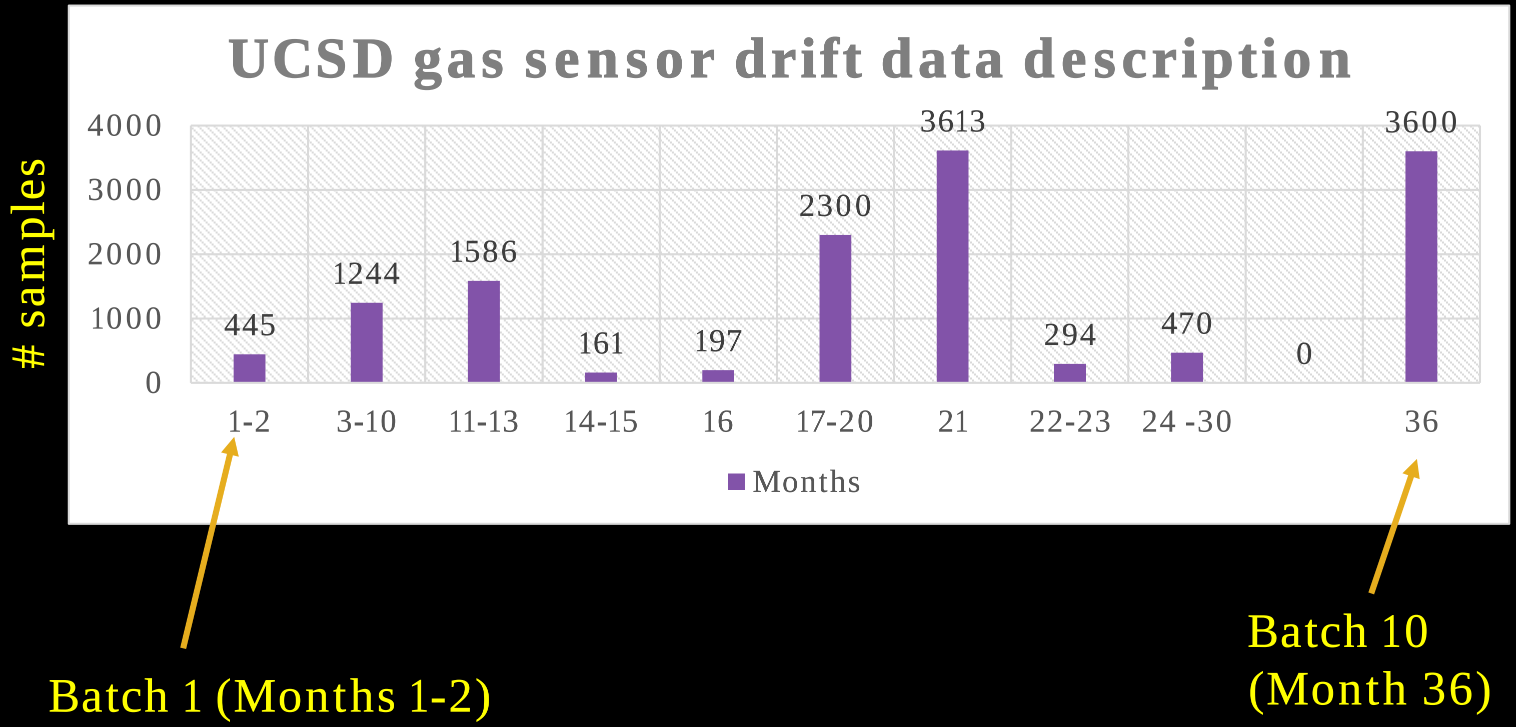
Algorithm requirements

- Must exhibit concentration tolerance (while also providing an estimate of concentration).
- Must be able to identify the signatures of known odors despite substantial interference/variance.
- Must exhibit rapid one- or few-shot learning of novel stimuli.
- Must support online learning (no catastrophic forgetting, no storage of training data).
- Must exhibit semi-supervised/unsupervised learning.
- Must provide a “none of the above” option (classifier confidence).
- Must have a solution for sensor drift (owing to time and/or contamination)
- Must be robust to “wild”, poorly-behaved inputs without parameter re-tuning.

Data set

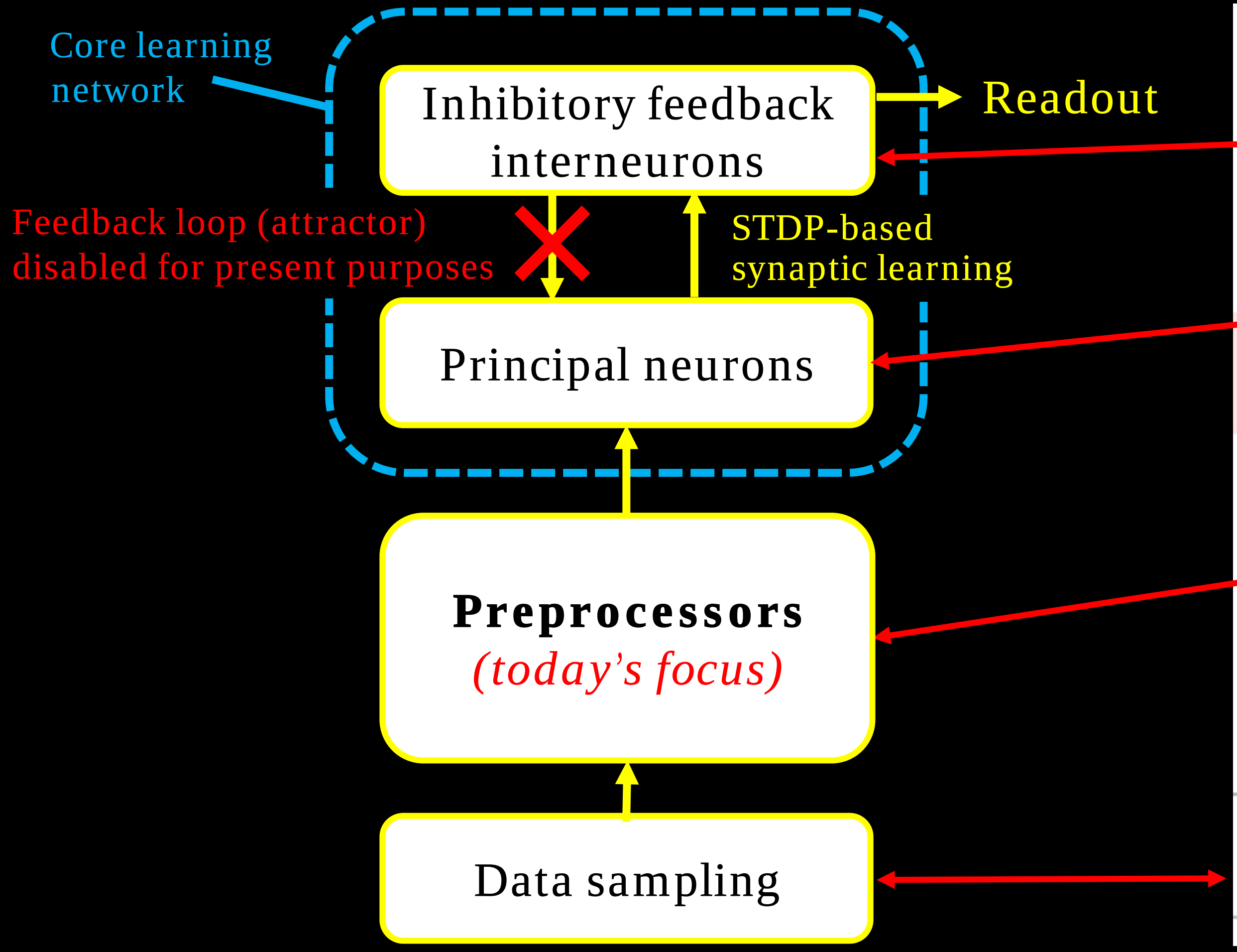
- UCSD gas sensor drift dataset*.
- 6 odors, each presented at a wide range of concentrations.
  - Ammonia
  - Acetaldehyde
  - Acetone
  - Ethylene
  - Ethanol
  - Toluene
- 13190 samples
  - Split into 10 batches.

Ten “batches” of data, taken over 3 years

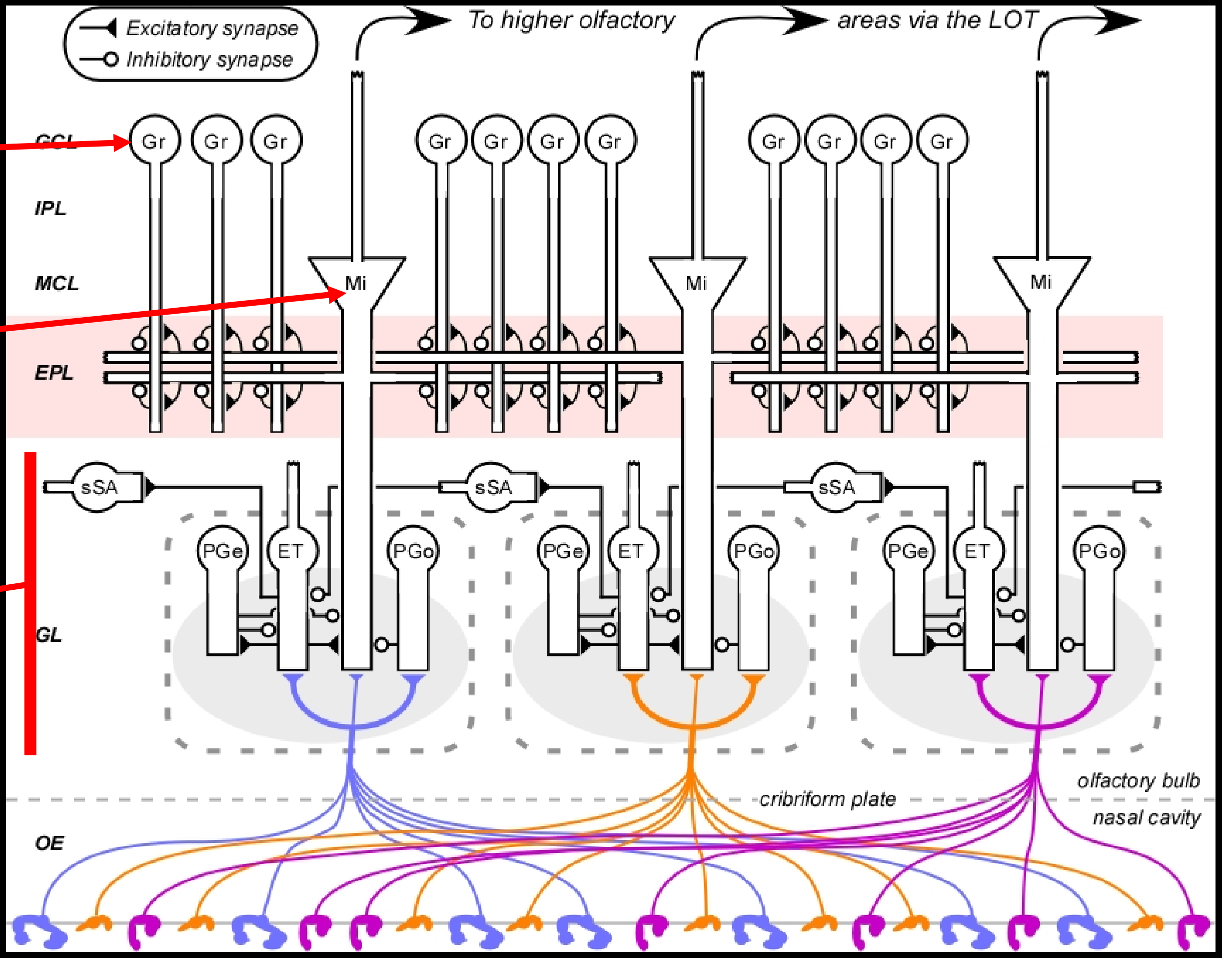


*Vergara et al., 2012; Rodriguez-Lujan et al., 2014

Biomimetic model schematic



Mammalian olfactory bulb network



- ➢ SNN; the architecture is part of the algorithm; local synaptic learning.
- ➢ Spike phase coding in the core feedback loop; exciting but not being discussed today.

Adapting the algorithm for learning in the wild

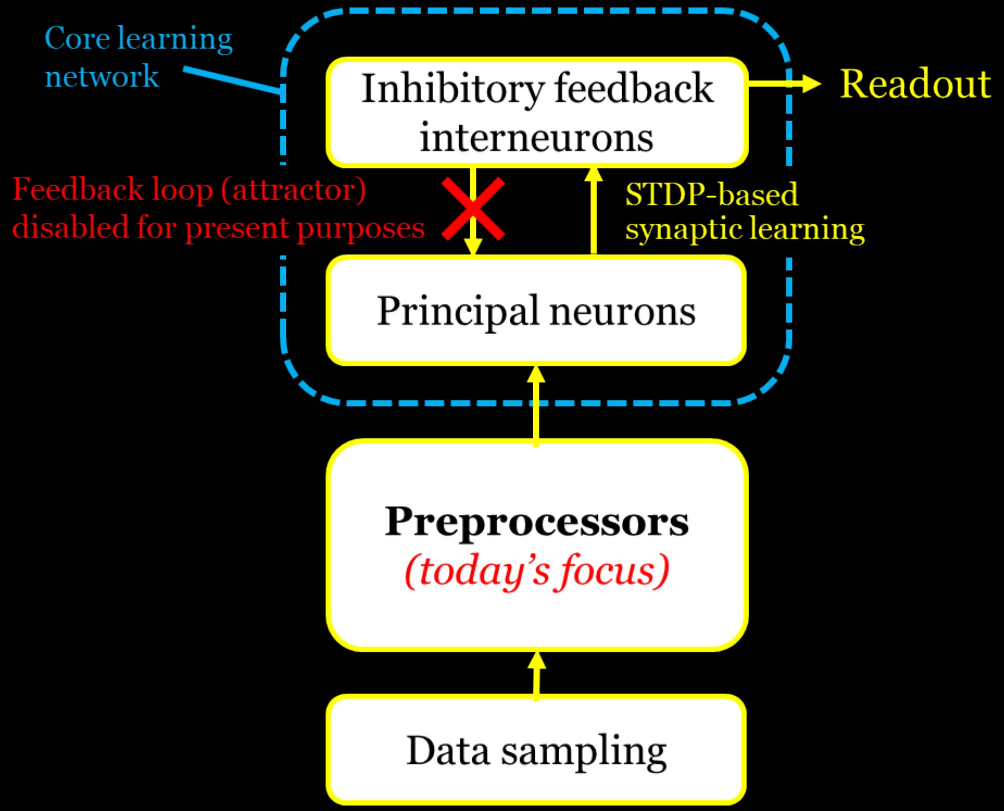
- Core learning network* comprises *principal neurons or mitral cells* (MCs) reciprocally coupled with *interneurons or granule cells* (GCs) in the external plexiform layer (EPL).

- Gist: MC spike time patterns recruit GCs to learn feature combinations via asymmetric STDP;
GC activity then is deployed as feedback inhibition to shape MC spike patterns (attractor).

- Like all networks, for optimal performance, *the core learning network requires that its sensory input patterns adhere to constraints of amplitude, and statistical structure.*

- *Learning in the wild* requires us to overcome this limitation, so that *the network can learn and respond productively to any input source.*

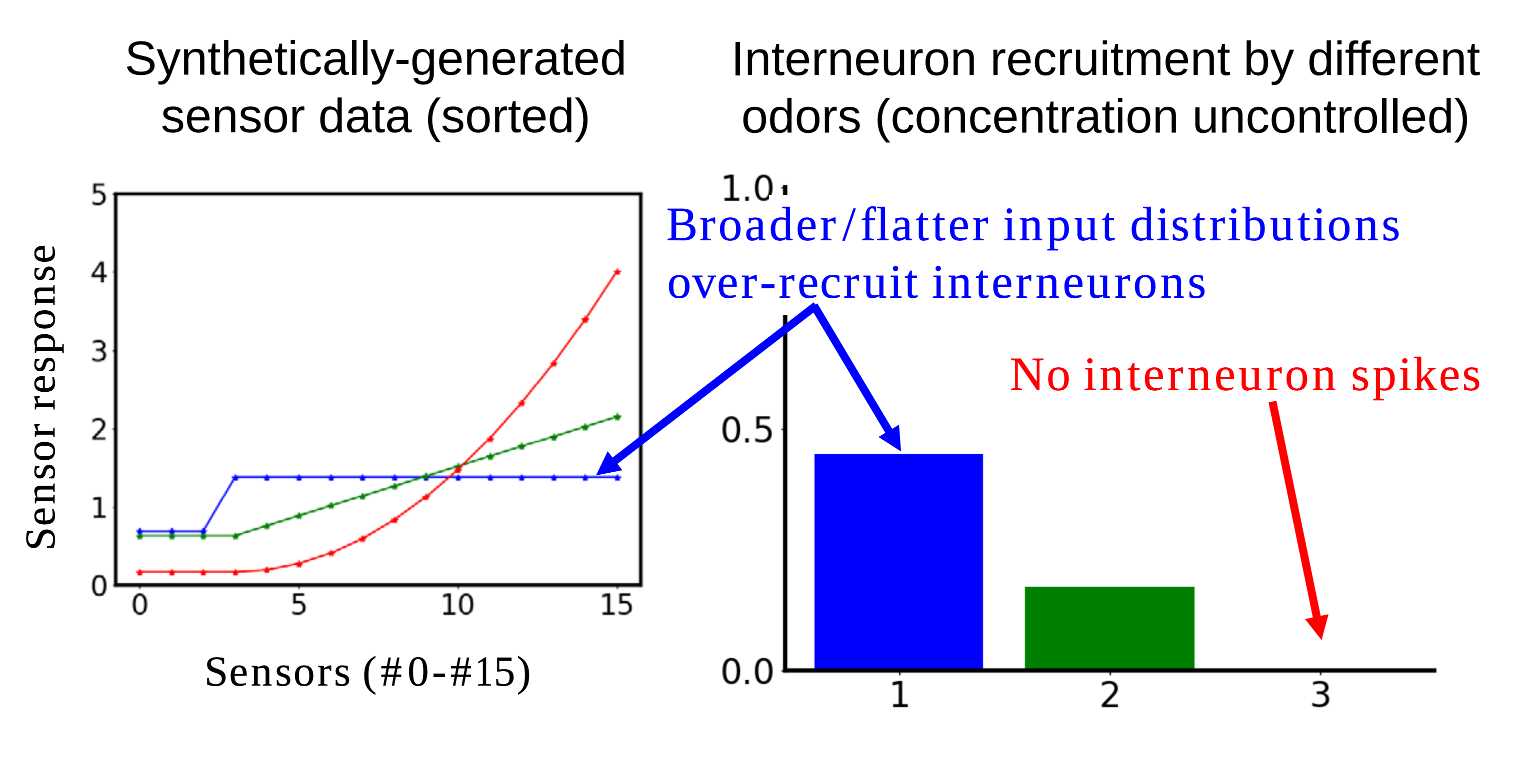
- Readout for classification: assessment of interneuron (GC) activation patterns.



*Nabil Imam; Thomas Cleland

Signal conditioning for learning in the wild

- Inputs from arbitrary sets of sensors in natural environments can be diverse and unpredictable.
- Network parameters can be retuned for different input statistics, but parameter tuning is slow and costly.
- Solution: apply signal conditioning so that the network will "just work" on arbitrary datasets without parameter retuning. Success is indicated when a wide diversity of stimuli each recruit similar numbers of interneurons.



Interneuron spike count similarity

No spike penalty

$$g_p = \min(\min(v), 1) * \frac{\sum \frac{v_i}{\max(v)}}{\textit{no of elements in } v}$$

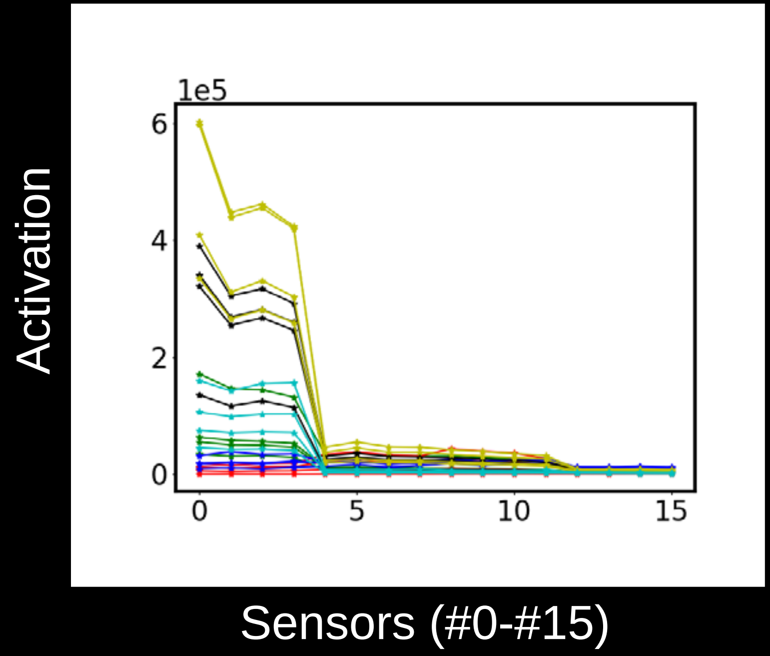
Goodness of preprocessing $g_p \in [0, 1]$

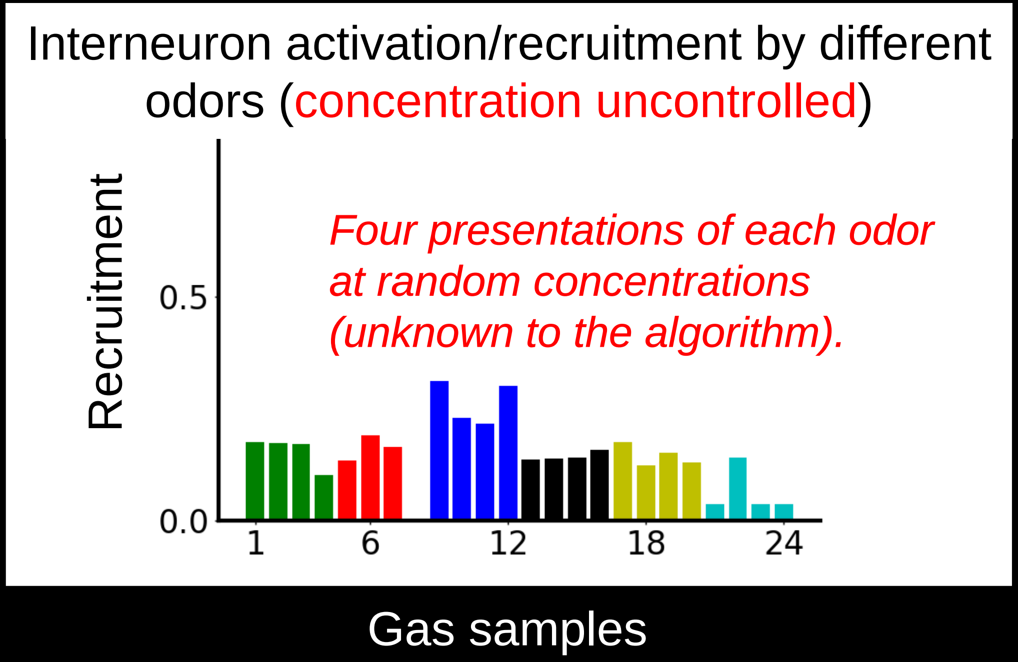
Interneuron spike count vector across samples

Signal conditioning for learning in the wild

- **Raw sensor data:**
- Uniform sensor scaling (from validation set)
- Nonuniform sensor scaling
- Concentration tolerance (glomerular network)
- Heterogenous duplication

$$g_p = 0$$





- Six odorants presented, each four times, at unknown concentrations (overall concentration range: 5 - 1000 ppmv).
- Different mean response amplitudes, different cross-sensor input statistics.

Signal conditioning for learning in the wild

Raw sensor data;
Sorted by amplitude

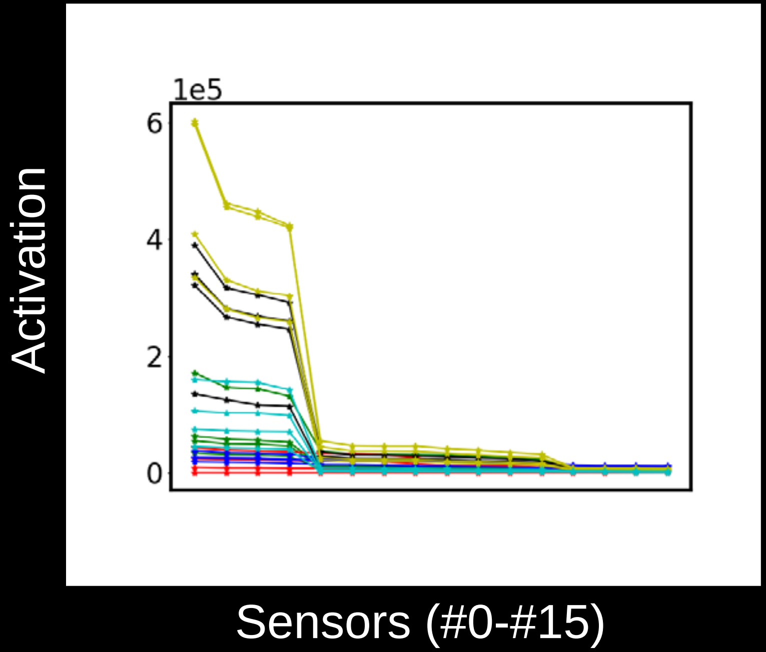
Uniform sensor scaling
(from validation set)

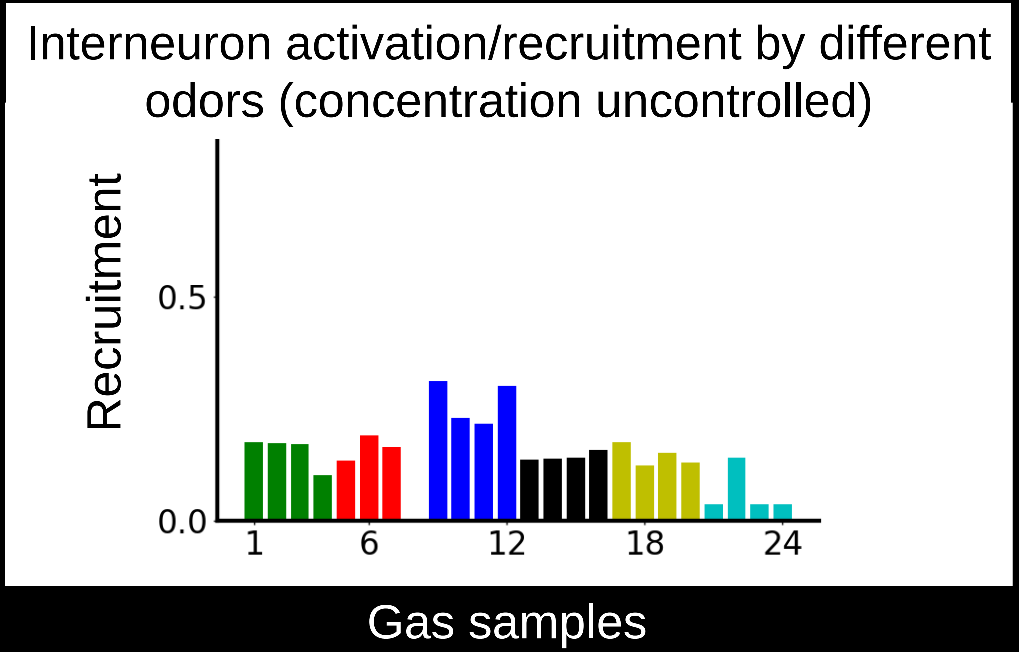
Nonuniform sensor scaling

Concentration tolerance
(glomerular network)

Heterogenous
duplication

$$g_p = 0$$





- All sensor responses sorted by amplitude for illustration.

Signal conditioning for learning in the wild

Raw sensor data; Sorted by amplitude

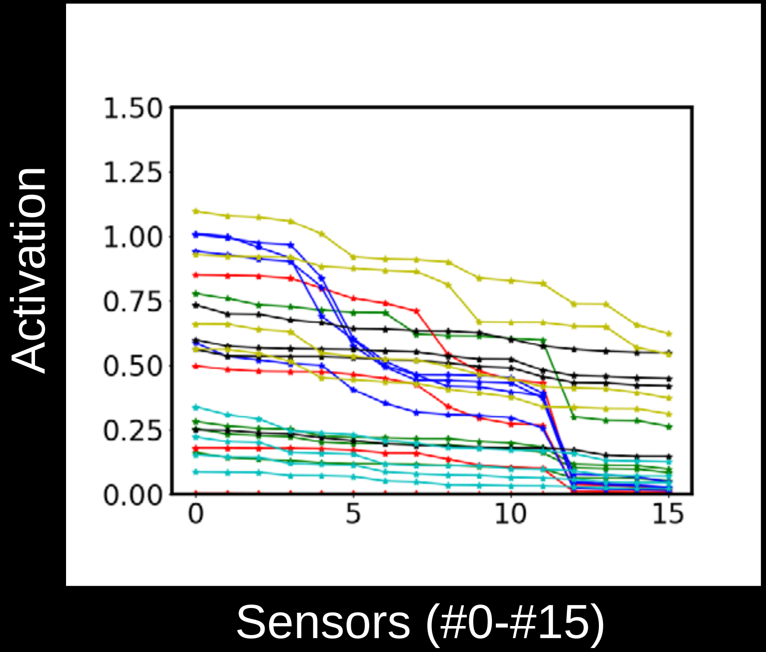
Uniform sensor scaling (from validation set)

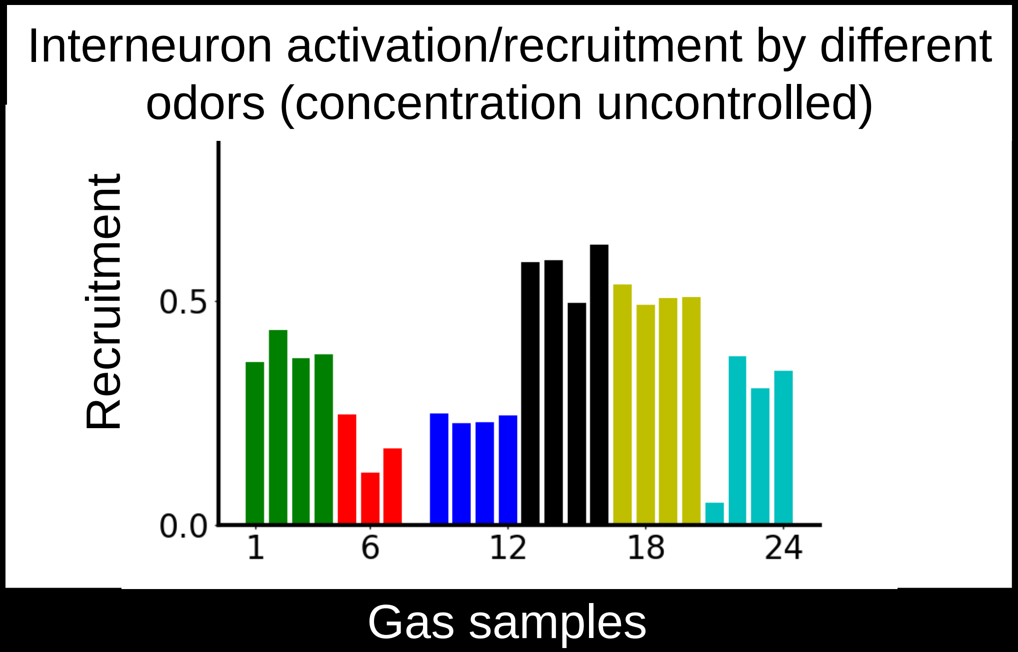
Nonuniform sensor scaling

Concentration tolerance (glomerular network)

Heterogenous duplication

$$g_p = 0$$





- Based on a validation set of samples, the working range of each sensor is scaled to [0,1], such that 1 is a reasonable estimate of the maximum value.
- This simple first step facilitates the use of highly heterogeneous sensors or datasets.

Signal conditioning for learning in the wild

Raw sensor data;
Sorted by amplitude

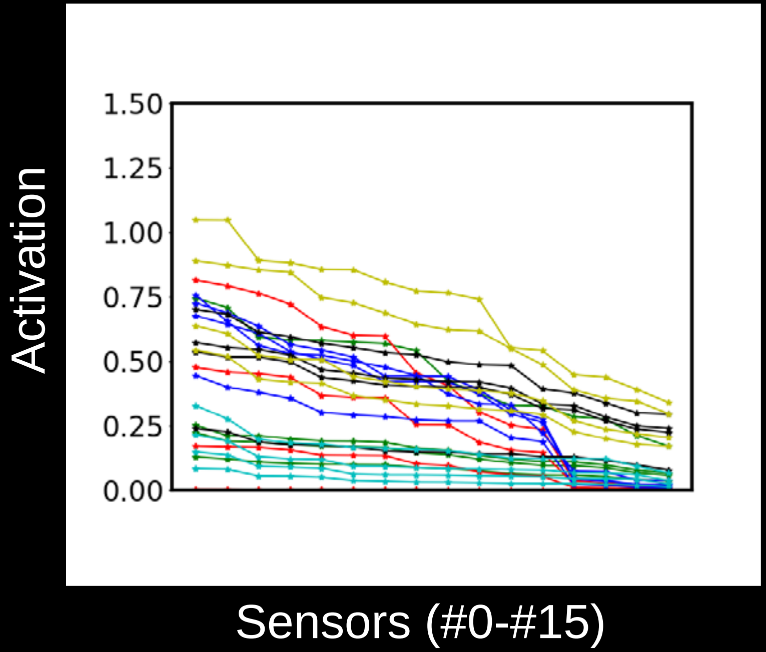
Uniform sensor scaling
(from validation set)

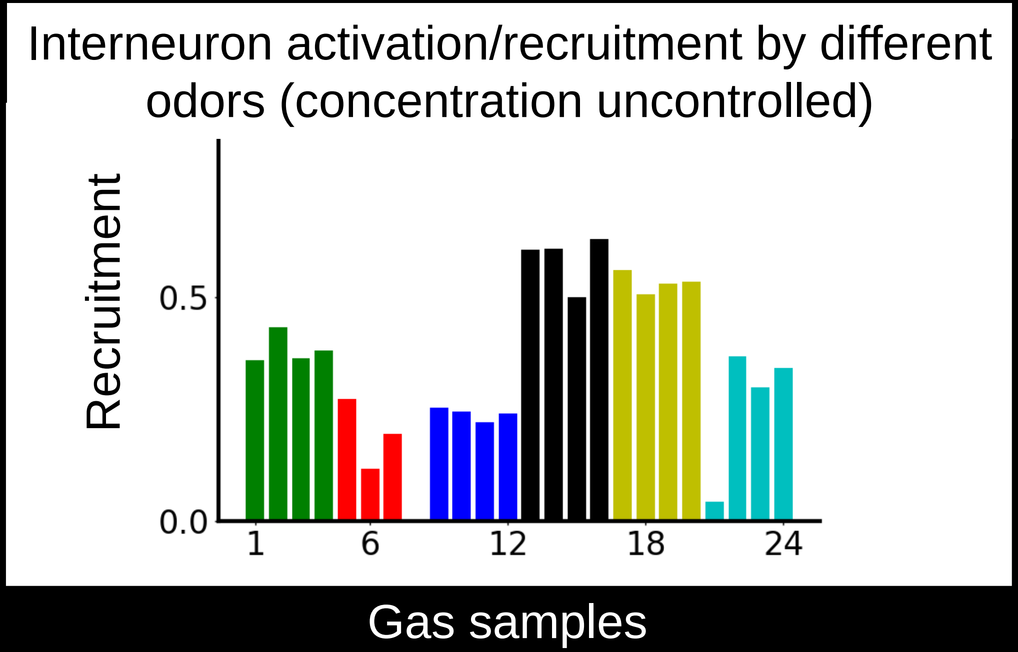
Nonuniform sensor scaling

Concentration tolerance
(glomerular network)

Heterogenous
duplication

$$g_p = 0$$





- Scaling values are further modulated by a equidimensional vector drawn from a uniform distribution.
- This is useful for subsequent preprocessing if the training set is small.

Signal conditioning for learning in the wild

Raw sensor data; Sorted by amplitude

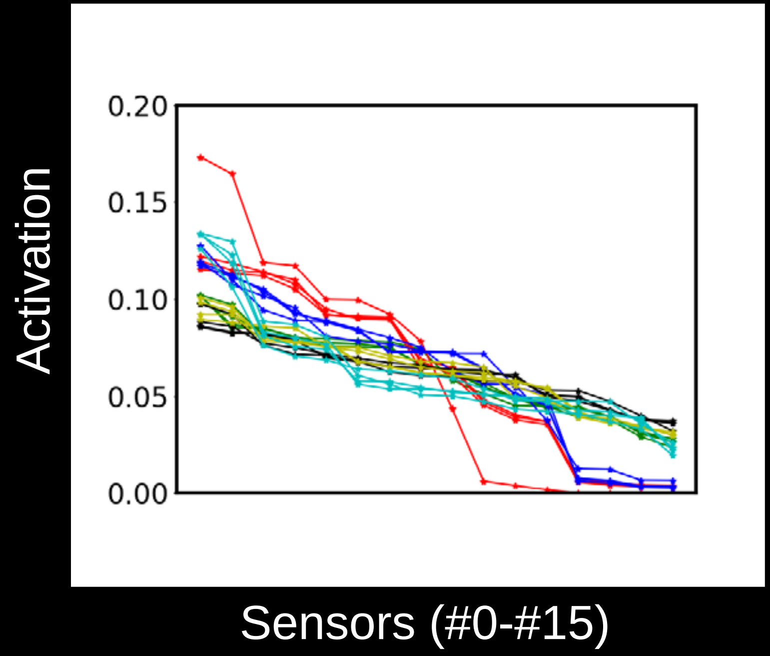
Uniform sensor scaling (from validation set)

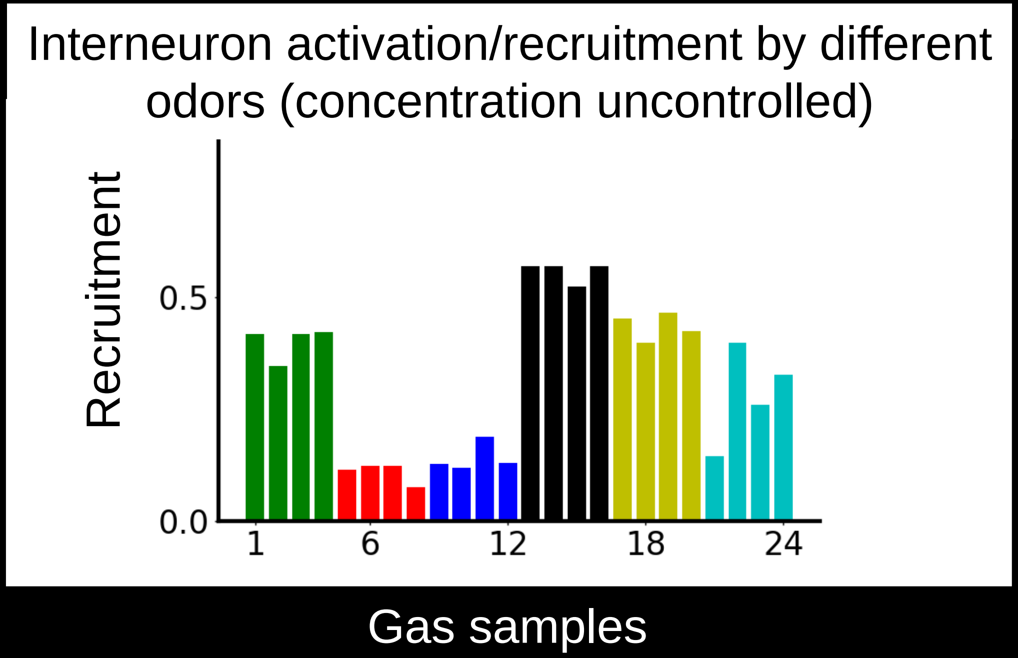
Nonuniform sensor scaling

Concentration tolerance (glomerular network)

Heterogenous duplication

$$g_p = 0.56$$





- Unsupervised concentration tolerance implemented in a biomimetic network.
- An intercolumnar network integrates net input across columns and delivers it uniformly to all columns as inhibition.

See also: Imam and Cleland (2012)

Signal conditioning for learning in the wild

Raw sensor data;
Sorted by amplitude

Uniform sensor scaling
(from validation set)

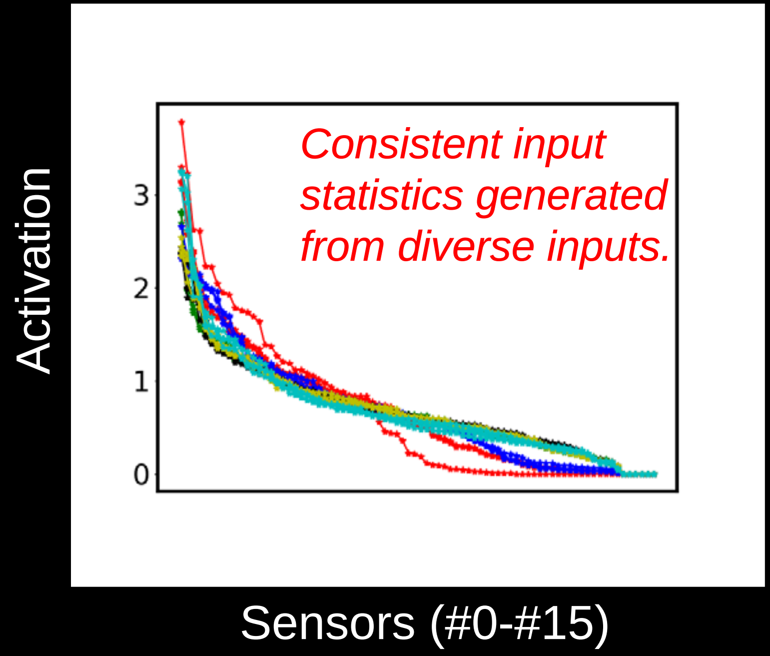
Nonuniform sensor scaling

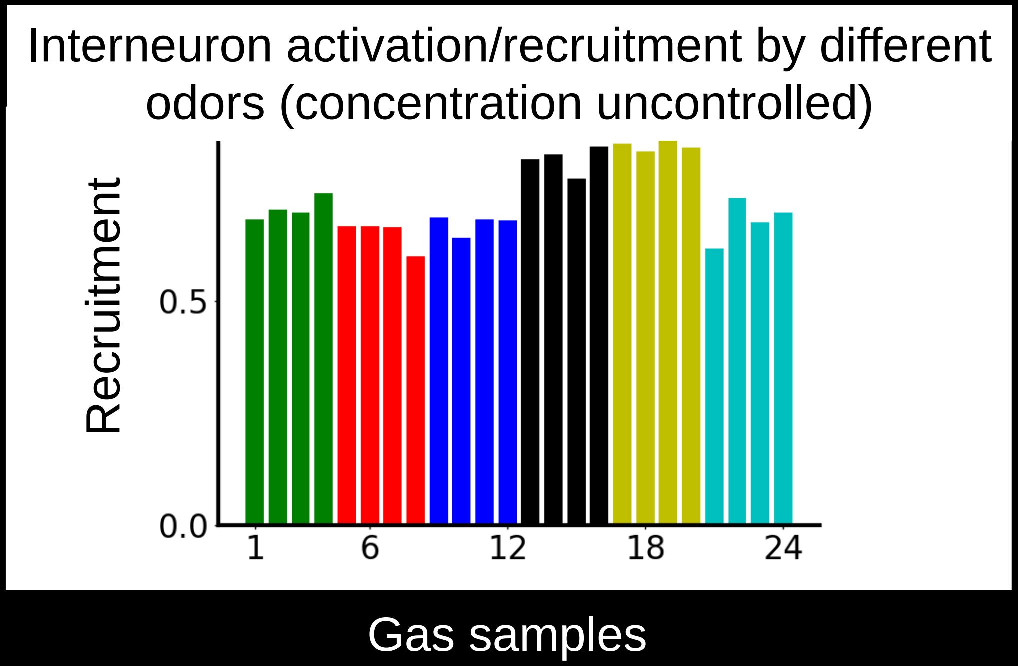
Concentration tolerance
(glomerular network)

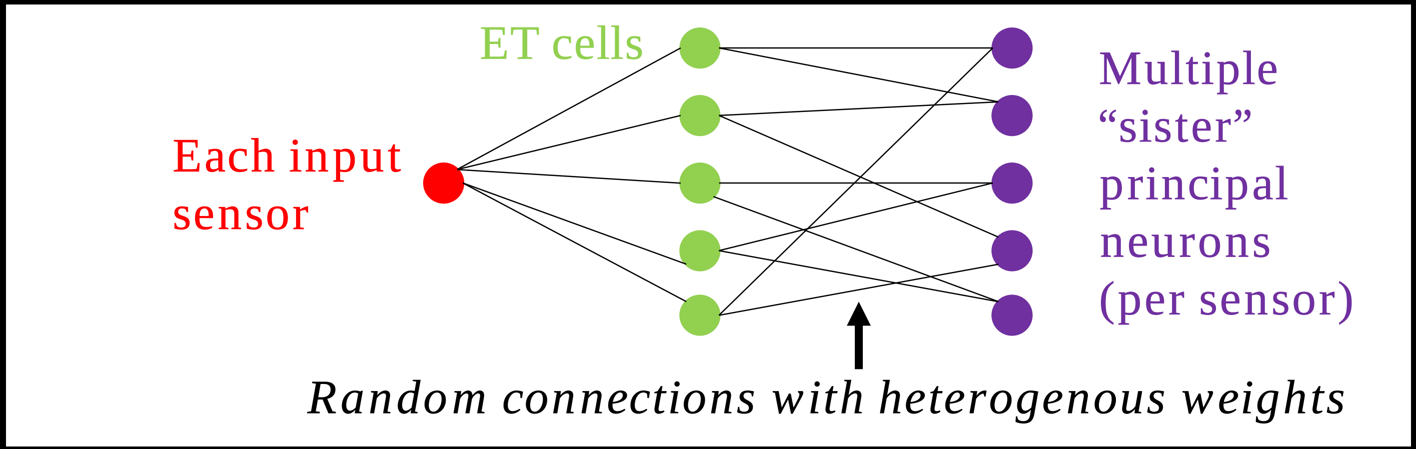
Heterogenous duplication

$$g_p = 0.84$$

28 % improvement







Signal conditioning for learning in the wild

Raw sensor data; Sorted by amplitude

Uniform sensor scaling (from validation set)

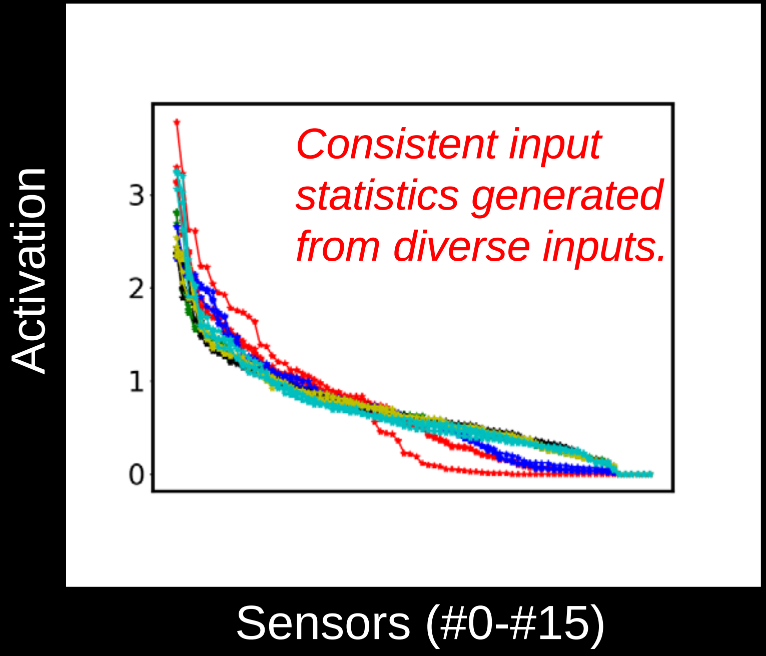
Nonuniform sensor scaling

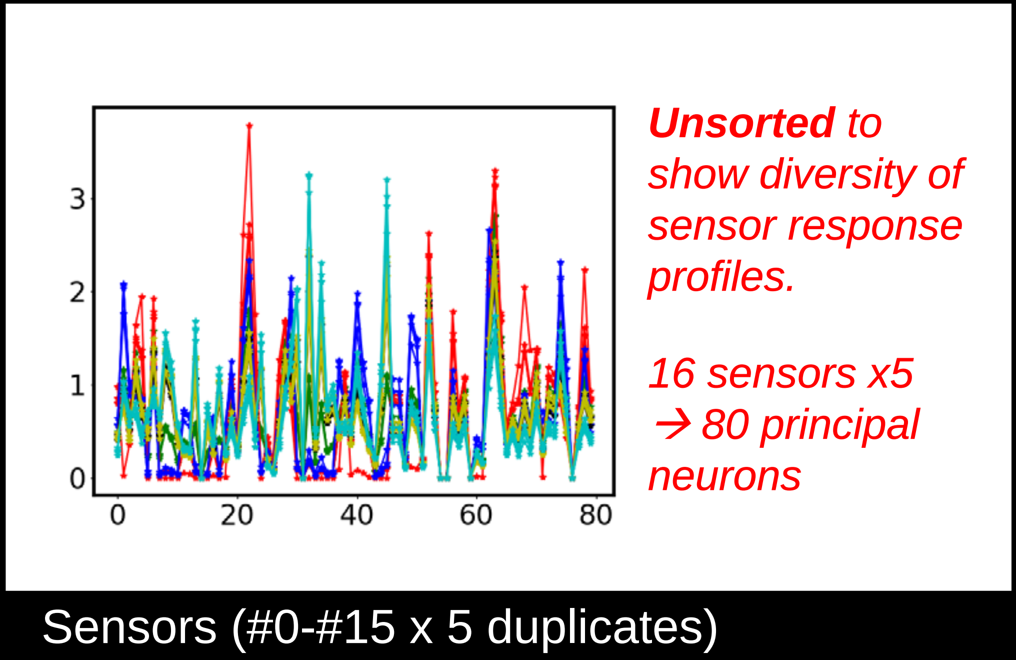
Concentration tolerance (glomerular network)

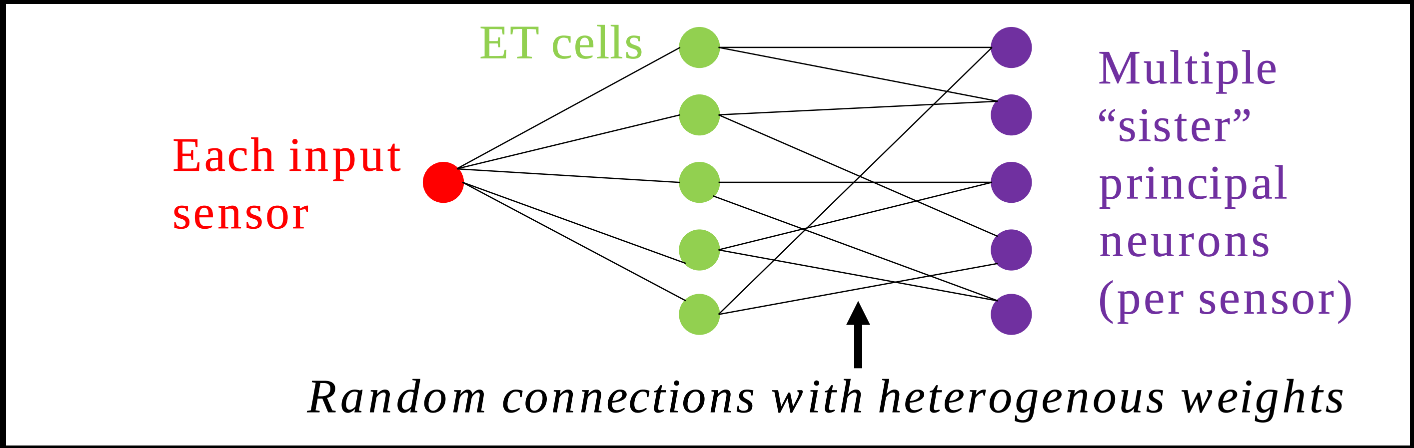
Heterogenous duplication

$$g_p = 0.84$$

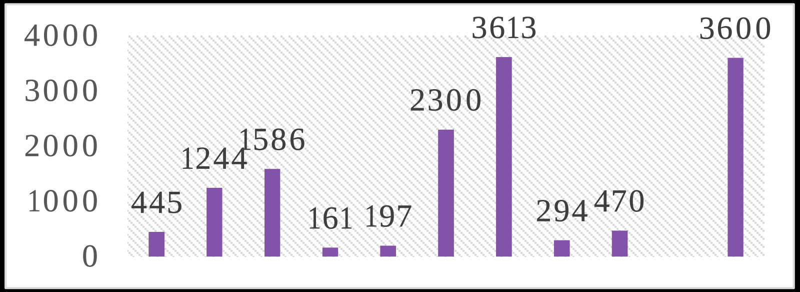
28 % improvement



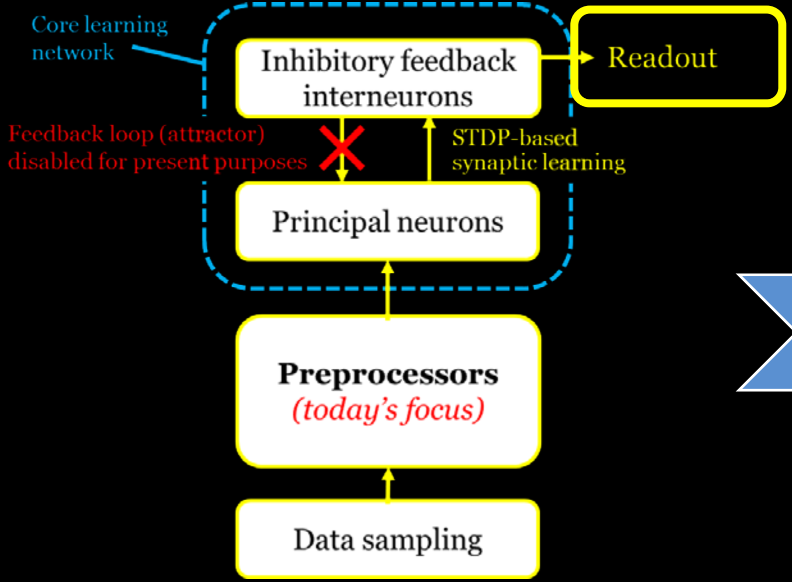




Balanced network learns diverse inputs via online learning

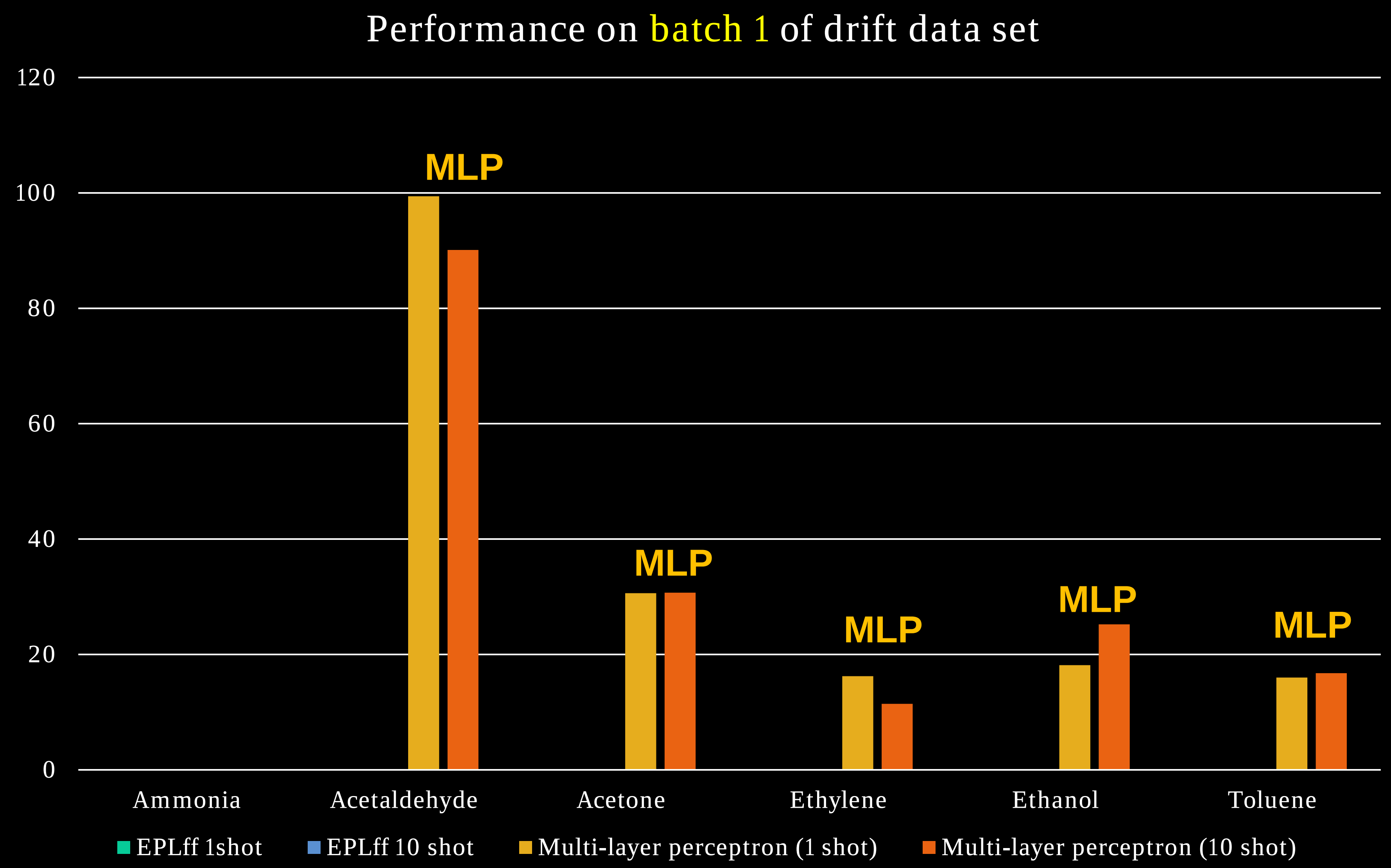


- Train and test using UCSD gas sensor drift dataset:
  - Ten “batches” of data taken over three years of sensor drift
  - Within each batch, train (few-shot) on each of the 5-6 gas types present, *sequentially*, *irrespective of concentrations*.
  - After training on each gas type, using the complete test set, measure classification performance from among all gases trained so far, or “none of the above”.
  - *Feedback loop is here omitted, so classification performance is measured directly from interneurons (“EPLff network”)*



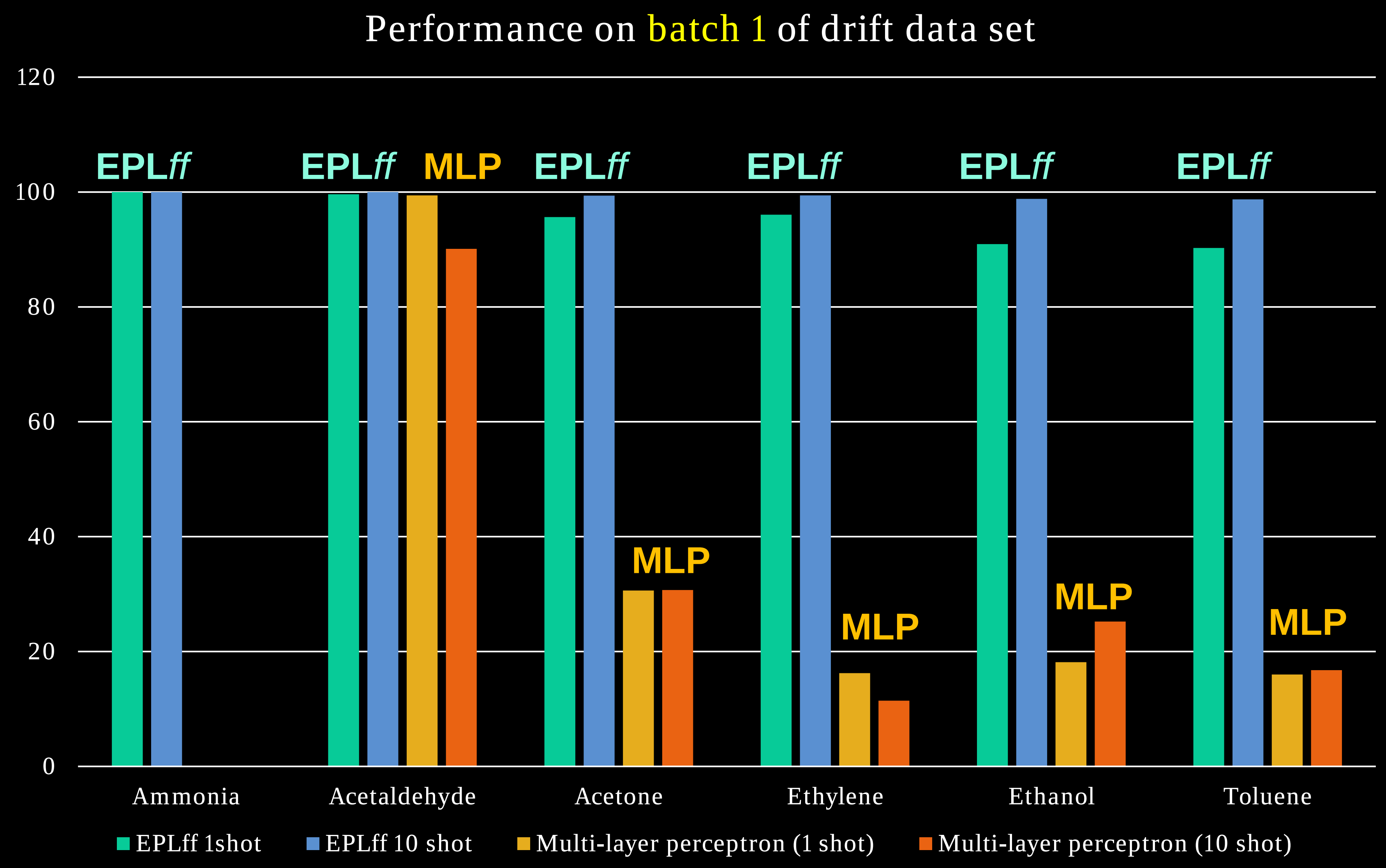
Group 1: Ammonia → Group 2: Acetaldehyde → Group 3: Acetone → Group 4: Ethylene → Group 5: Ethanol → Group 6: Toluene

Online learning performance



Online learning performance in a simple multilayer perceptron (MLP) is shown for illustration.

Online learning performance



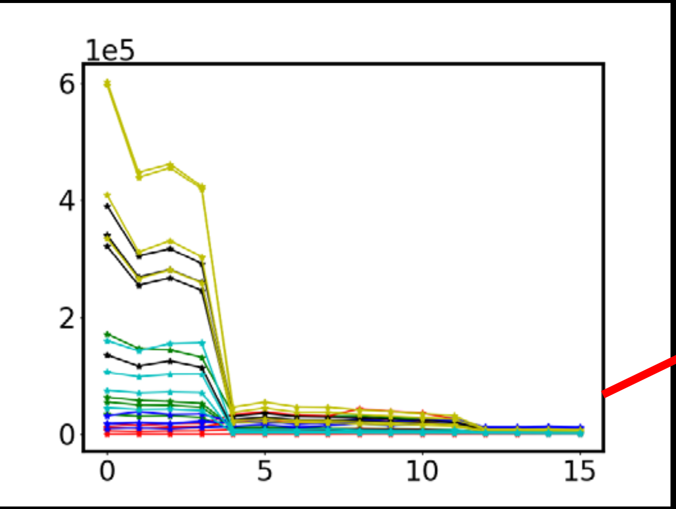
The networks were trained on odorants sequentially, in the order depicted.

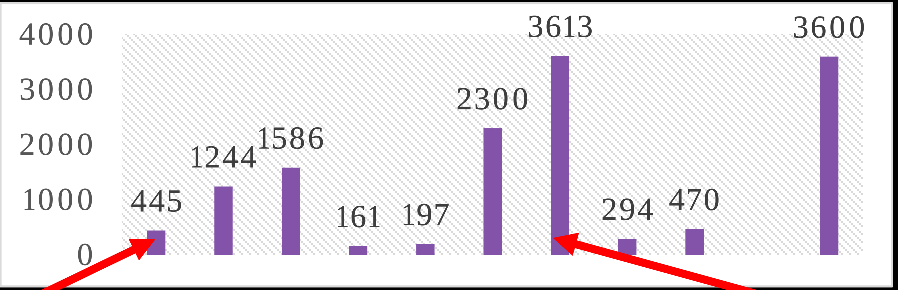
EPLff does not suffer from catastrophic forgetting.

Online learning performance in a simple multilayer perceptron (MLP) is shown for illustration.

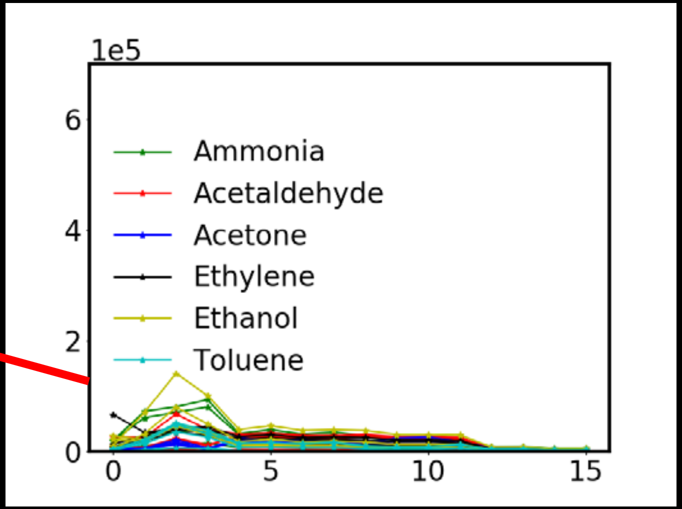
The problem of sensor drift

Batch 1 (Months 1-2)





Batch 7 (Month 21)

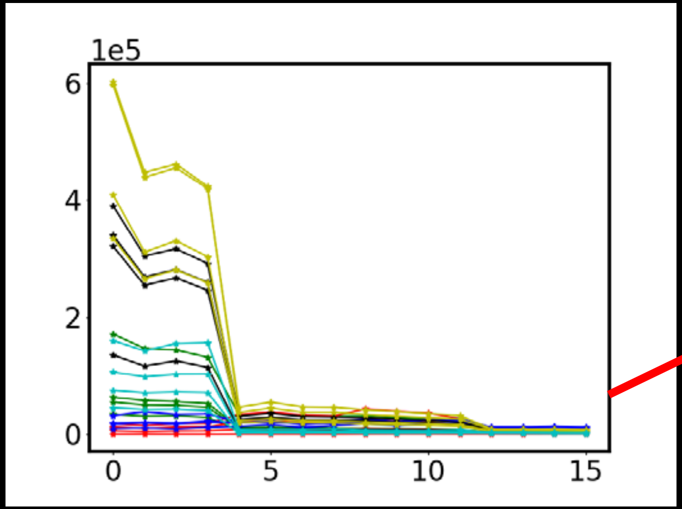


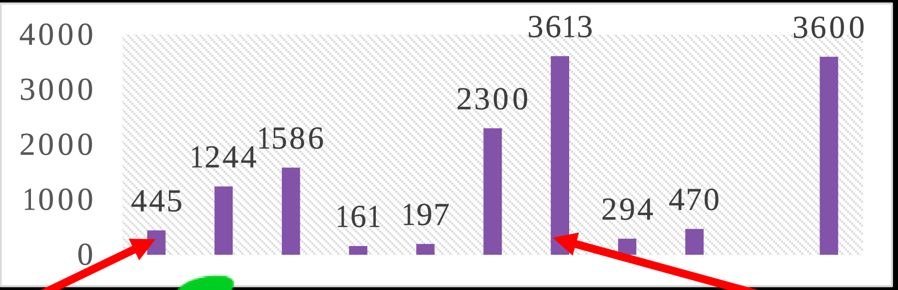
1. Sensor *decay* reduces SNR of inputs.

2. Sensor *drift* renders prior learning obsolete.

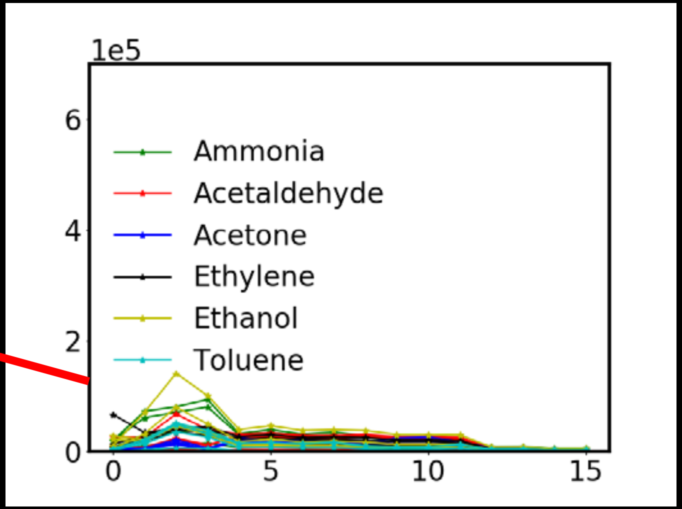
The problem of sensor drift

Batch 1 (Months 1-2)



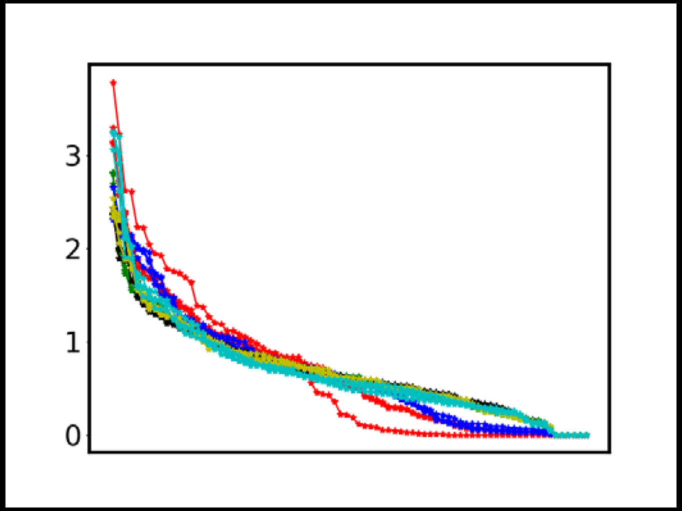


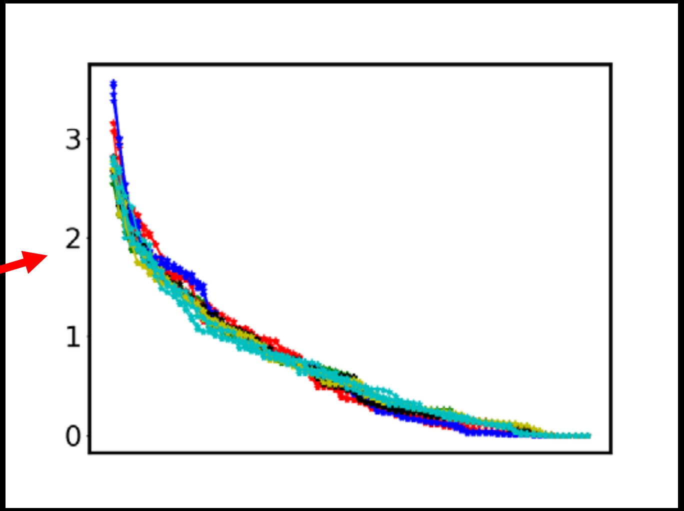
Batch 7 (Month 21)



1. Sensor *decay* reduces SNR of inputs.

2. Sensor *drift* renders prior learning obsolete.

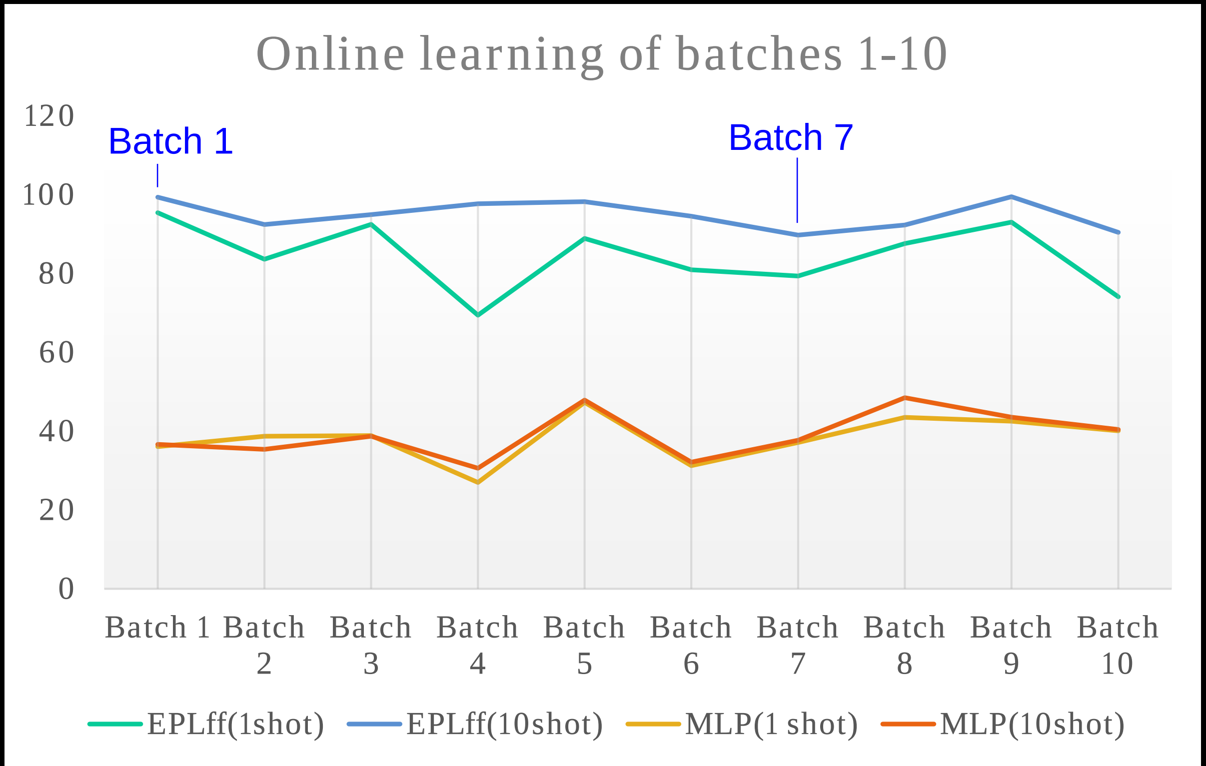




Input preprocessing series recovers well-behaved signals even from degraded sensors

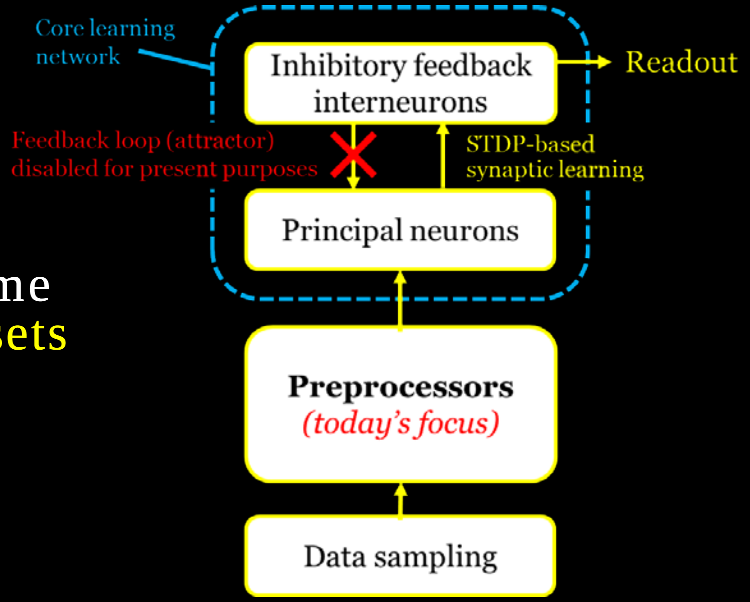
Rapid online learning as a solution for sensor drift

1. Sensor *decay* reduces SNR of inputs.
2. Sensor *drift* renders prior learning obsolete.



- Task: Learn all odors in series within each batch. Test classification of all odors, all concentrations, including "none of the above".
- *Sensor drift solution*: if sensors have drifted, rapidly retrain network (few-shot learning of known odors).
- Does not require hyperparameter re- tuning
- No need to wait for entire train set availability (model can be updated later too).
- Classifier confidence. "None of the above" responses to known trained classes can be used to determine the onset of reset.

Summary: *Learning in the wild*



- Our SNN algorithm supports rapid, few-shot, online learning and robust classification under noise.
- We present a series of signal conditioning preprocessors (some trivial, some novel) that enable this algorithm to usefully process poorly-behaved datasets without hyper-parameter tuning.
  - Signals with dissimilar sensor statistical distributions
  - Signals presented across ranges of intensity (concentration)
  - Signals from low-quality or degraded sensors
- Heterogeneity in the network is useful at multiple stages
  - Non-uniform sensor scaling preprocessor.
  - Heterogeneous duplication of input streams in the preprocessor network enables statistical regularization.
  - Heterogeneity in thresholds (interneurons, sister MCs) improves algorithm performance (not shown).
- Rapid learning (using EPL*ff*) with degraded sensors resolves the problem of sensor drift.

Acknowledgments

Intel Labs

Nabil Imam

http://cplab.net

Neural computation

Guoshi Li
Matt Einhorn
Francesco Cavarretta

Analytical modeling

Jack Cook

MEA slice recordings

Jesse Werth
Shane Peace

Collaboration

Christiane Linster

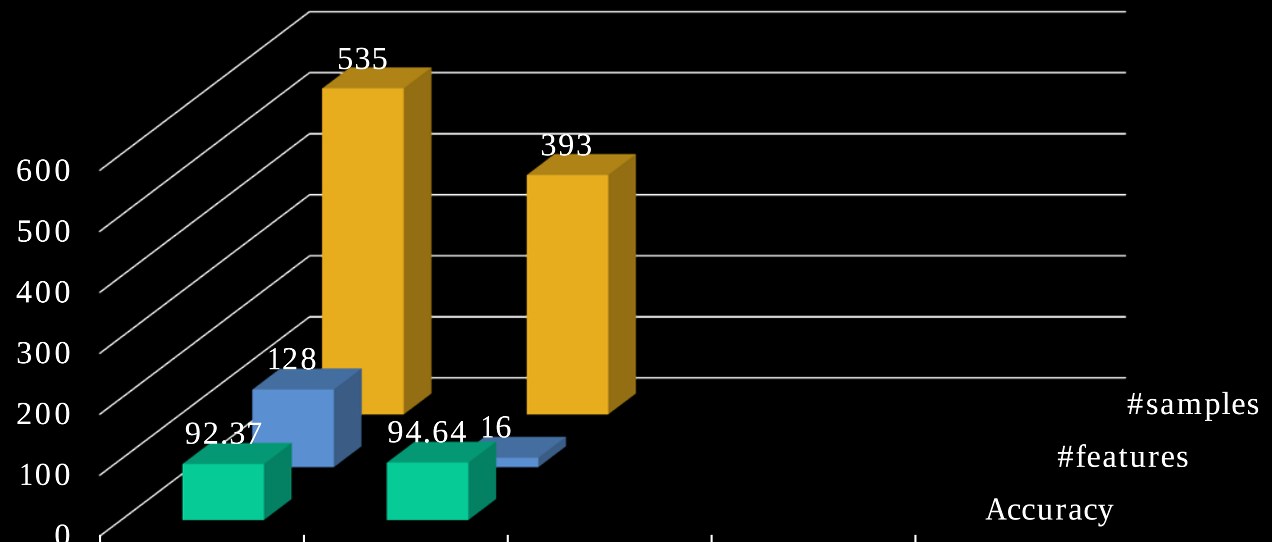
Committee members

Thomas Cleland (Chair)
Alyosha Molnar
Thorsten Joachims
David Field
David Smith

Thank you

Rapid online learning as a solution for sensor drift

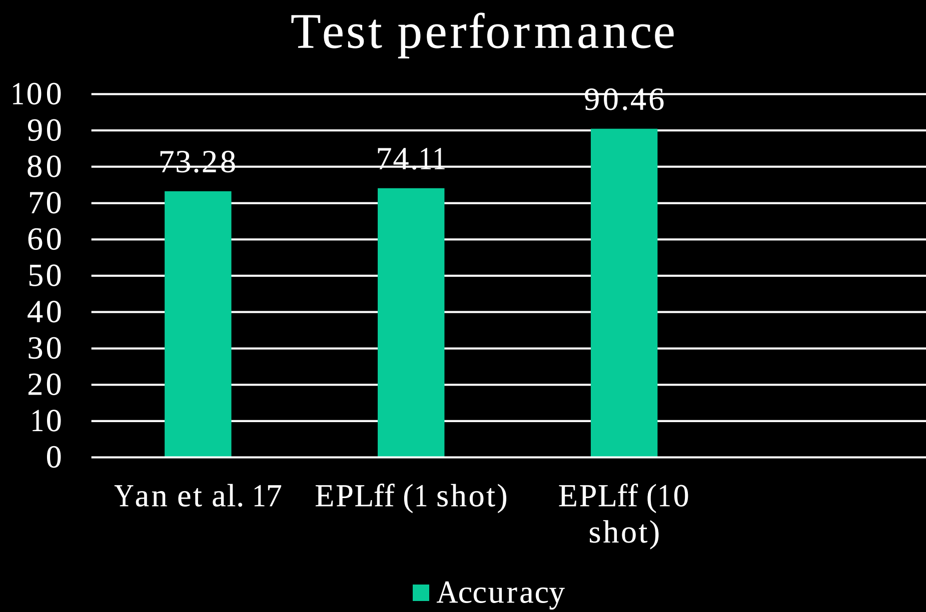
Batch 1-10 : Test performance

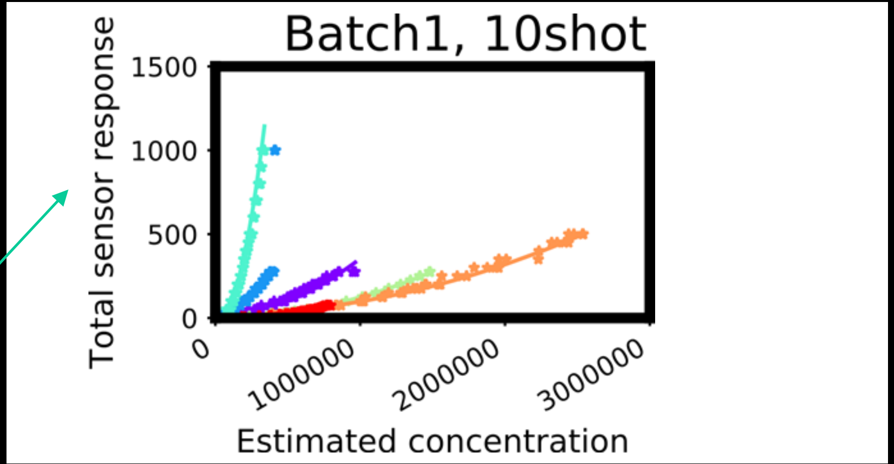


	Yan et al. 17	EPLff		
Accuracy	92.37	94.64		
#features	128	16		
#samples	535	393		

Accuracy #features #samples

Batch 10: Highly contaminated sensors





- ➤ All previous approaches non-online.

EPL*ff* provide concentration estimation.