

# **Dendritic Ca<sub>2+</sub> as a Predictor of Stimulus Perception and Behavior**

Uni- and Multivariate Analysis

---

Georg Chechelnizki

July 20, 2017

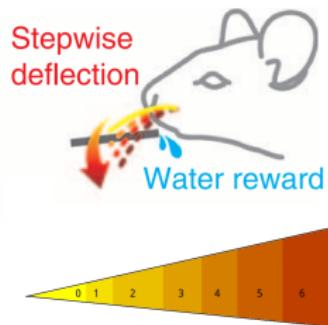
BCCN Berlin

# **Introduction**

---

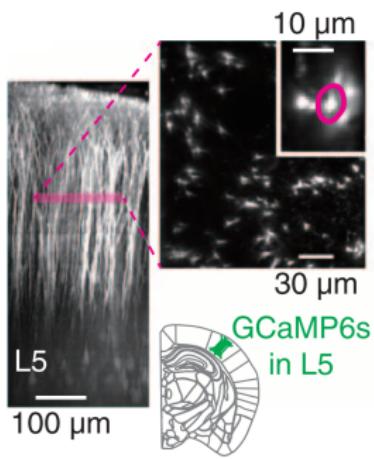
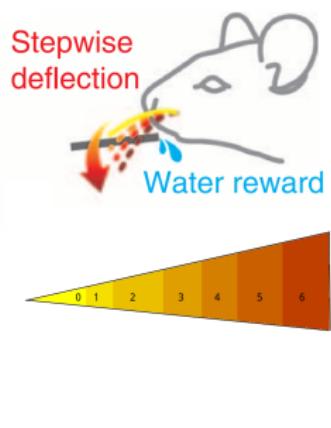
# The Experiment

Takahashi et al. 2016

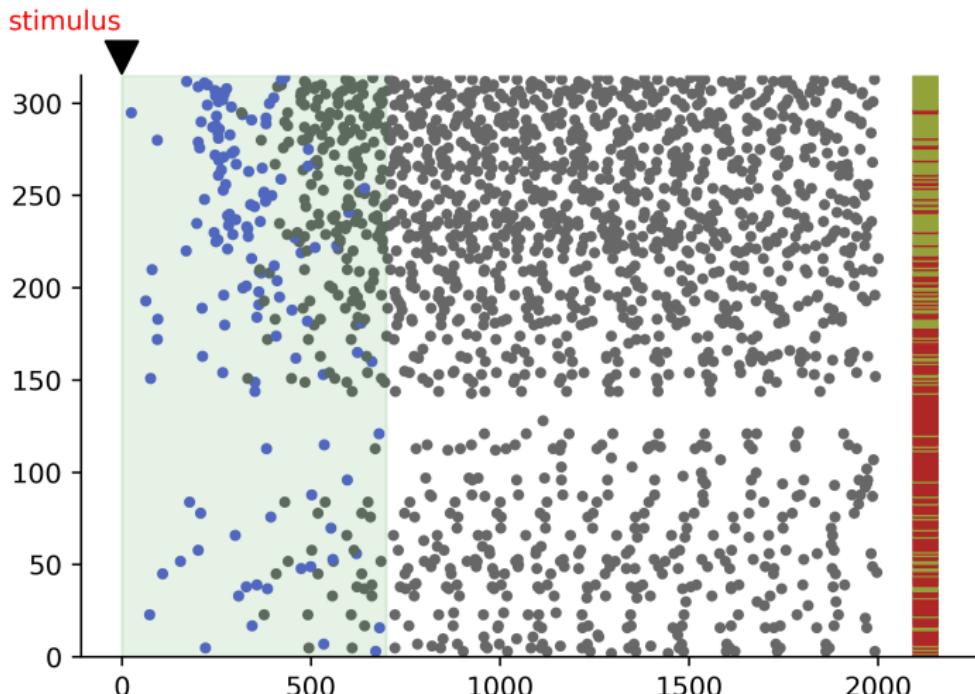


# The Experiment

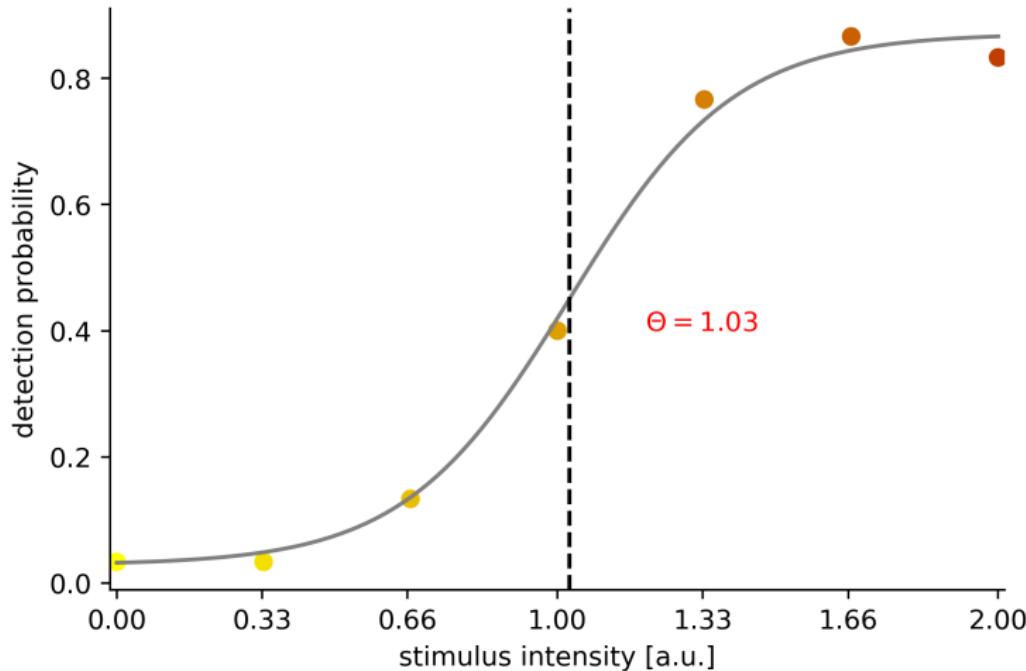
Takahashi et al. 2016



## The Data - Behavioral

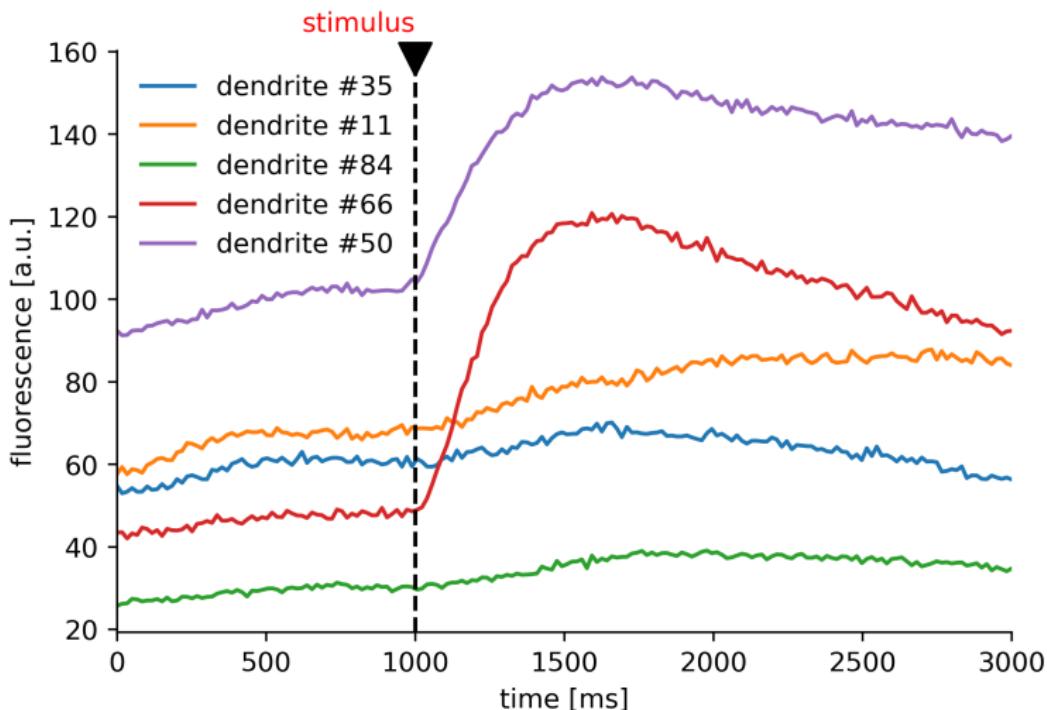


## The Data - Psychometric Curve



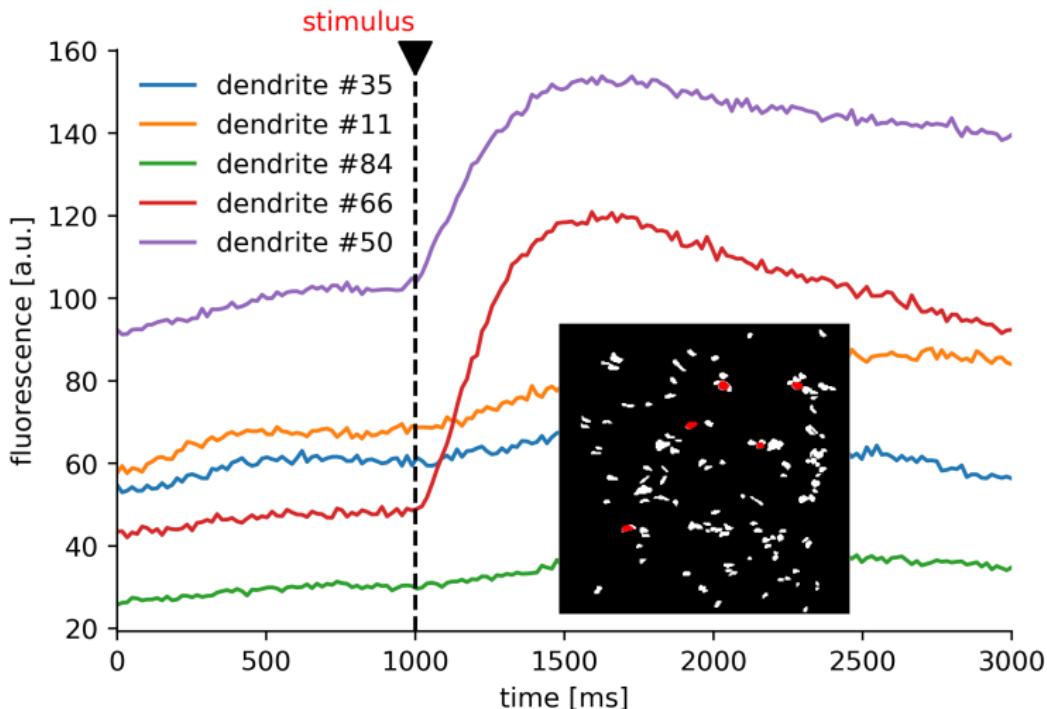
# The Data - Neuronal

Trial-averaged  $\text{Ca}^{2+}$  fluorescence traces of random dendrites



# The Data - Neuronal

Trial-averaged  $\text{Ca}^{2+}$  fluorescence traces of random dendrites



# Goal

---

# Goal

The goal of this project is to investigate the following:

- What is the relationship between  $\text{Ca}^{2+}$  activity and stimulus intensity/behavior

## Goal

The goal of this project is to investigate the following:

- What is the relationship between  $\text{Ca}^{2+}$  activity and stimulus intensity/behavior
- Can we build predictive classifiers for this data?

# Goal

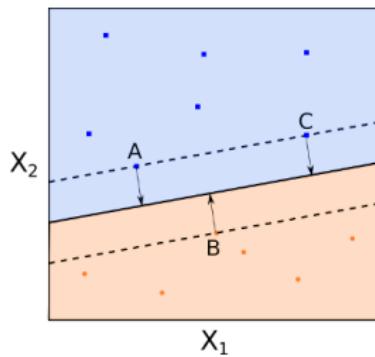
The goal of this project is to investigate the following:

- What is the relationship between  $\text{Ca}^{2+}$  activity and stimulus intensity/behavior
- Can we build predictive classifiers for this data?
- Does a multivariate approach give us any significant advantage over a univariate one?

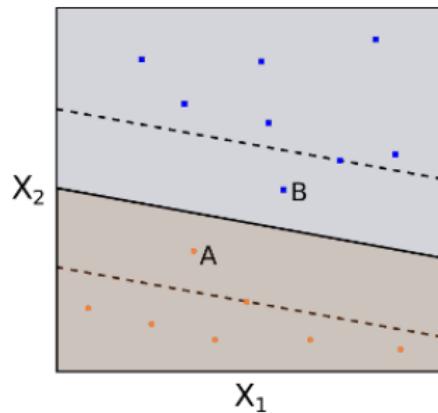
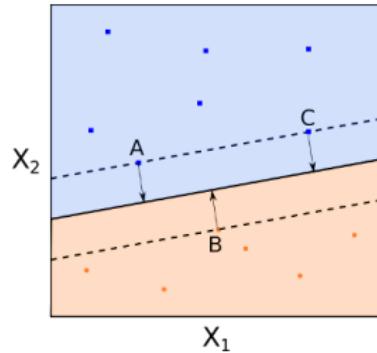
## Univariate Analysis

---

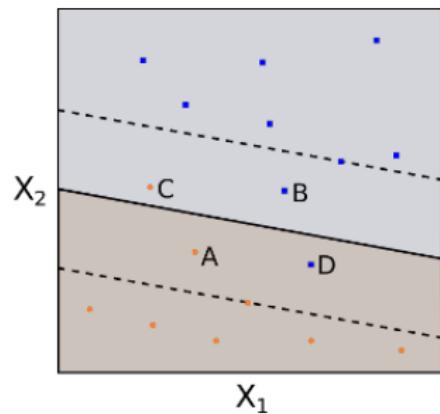
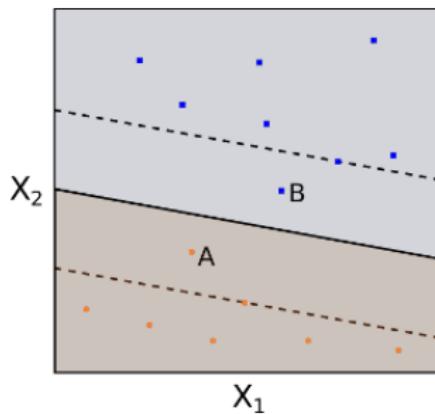
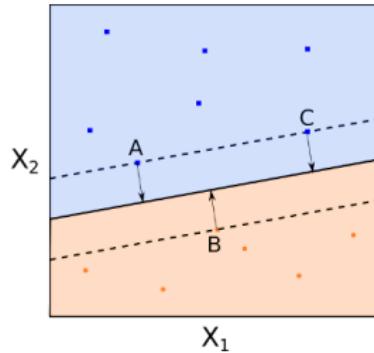
# SVM - What is an SVM?



# SVM - What is an SVM?

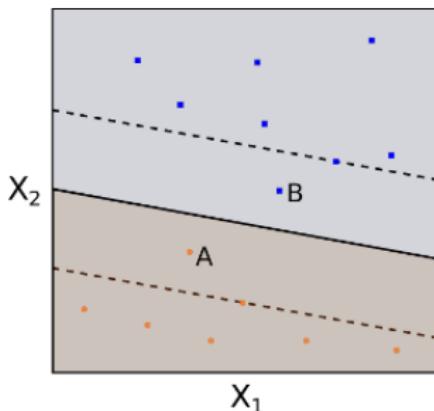
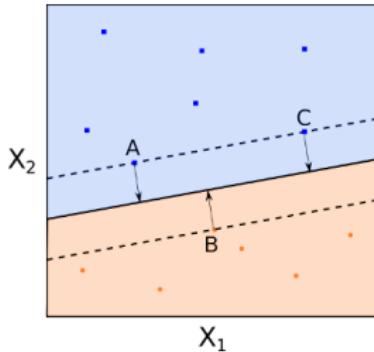


# SVM - What is an SVM?

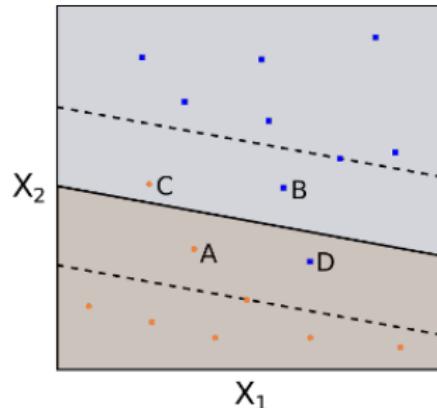


# SVM - What is an SVM?

SVMs...



- are a staple of machine learning
- have a regularization parameter that can combat overfitting
- are good for small samples with many features (in theory)



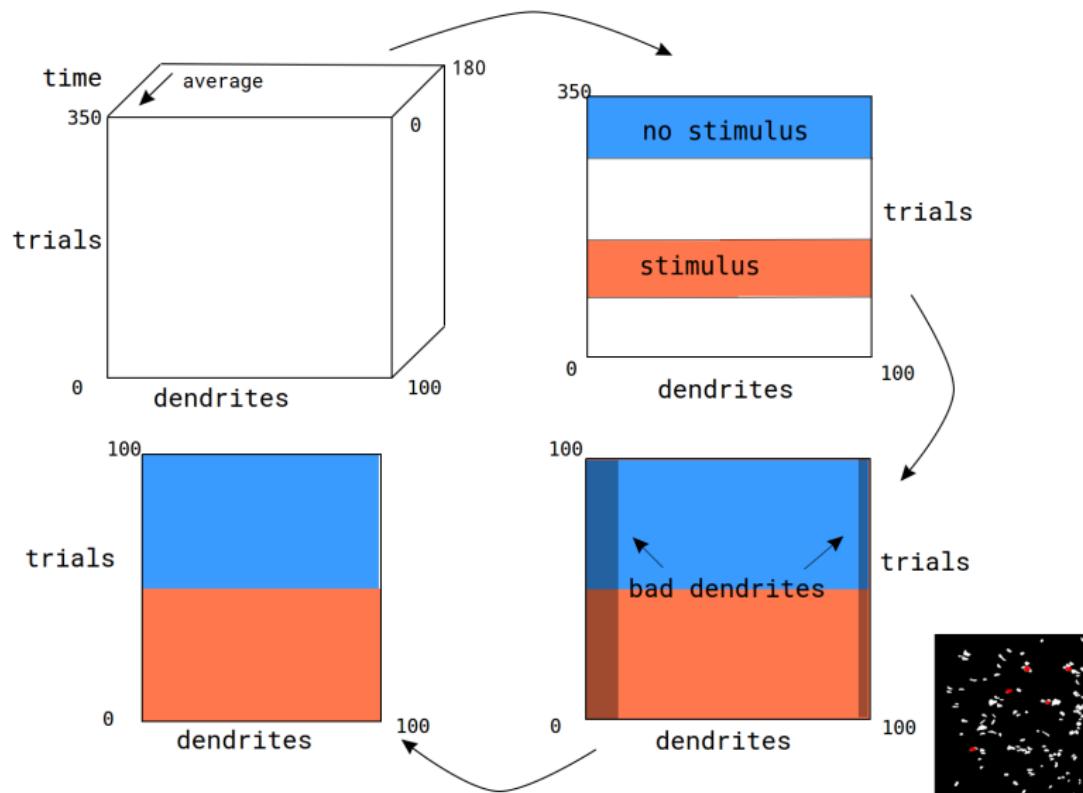
## SVM Specifics

- All data are normalized to zero mean and unit variance

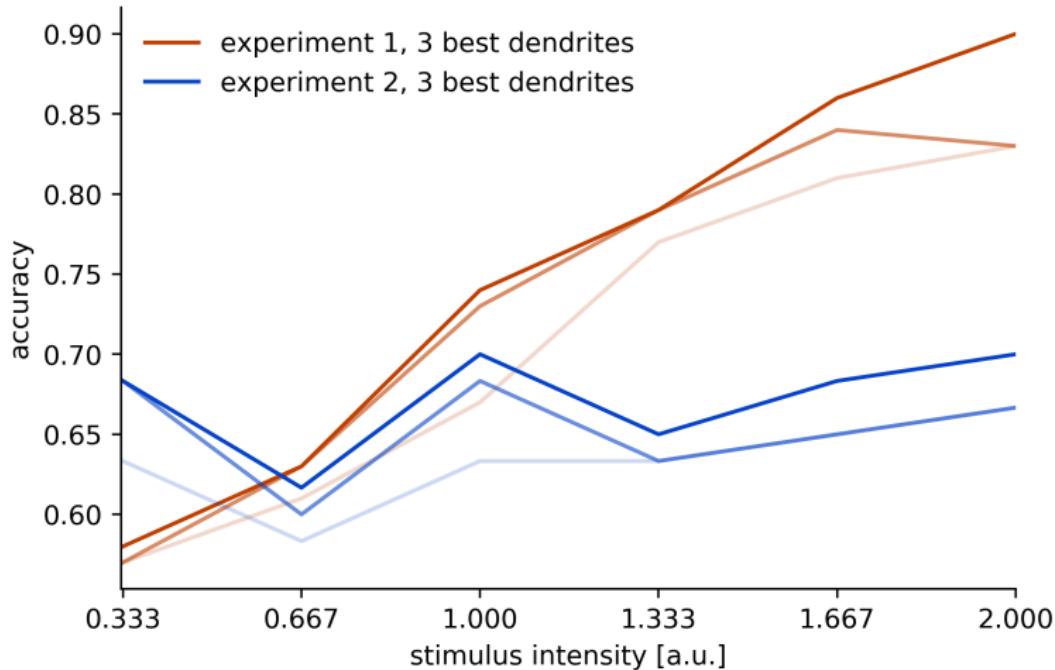
## SVM Specifics

- All data are normalized to zero mean and unit variance
- Crossvalidation is performed to control for overfitting

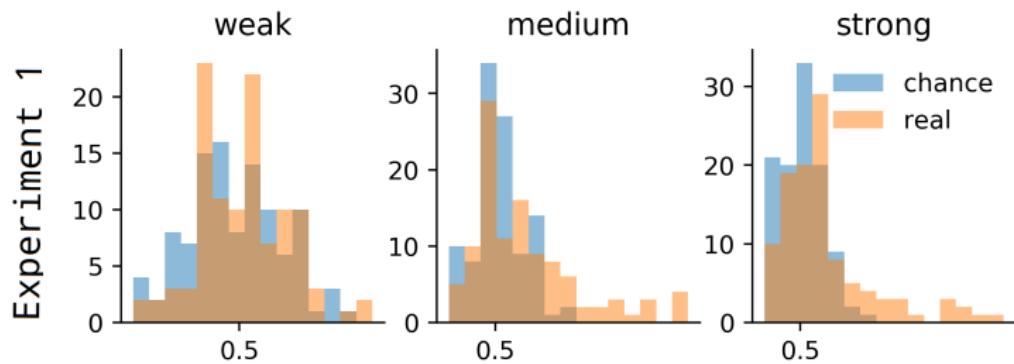
# Stimulus Presence Detection



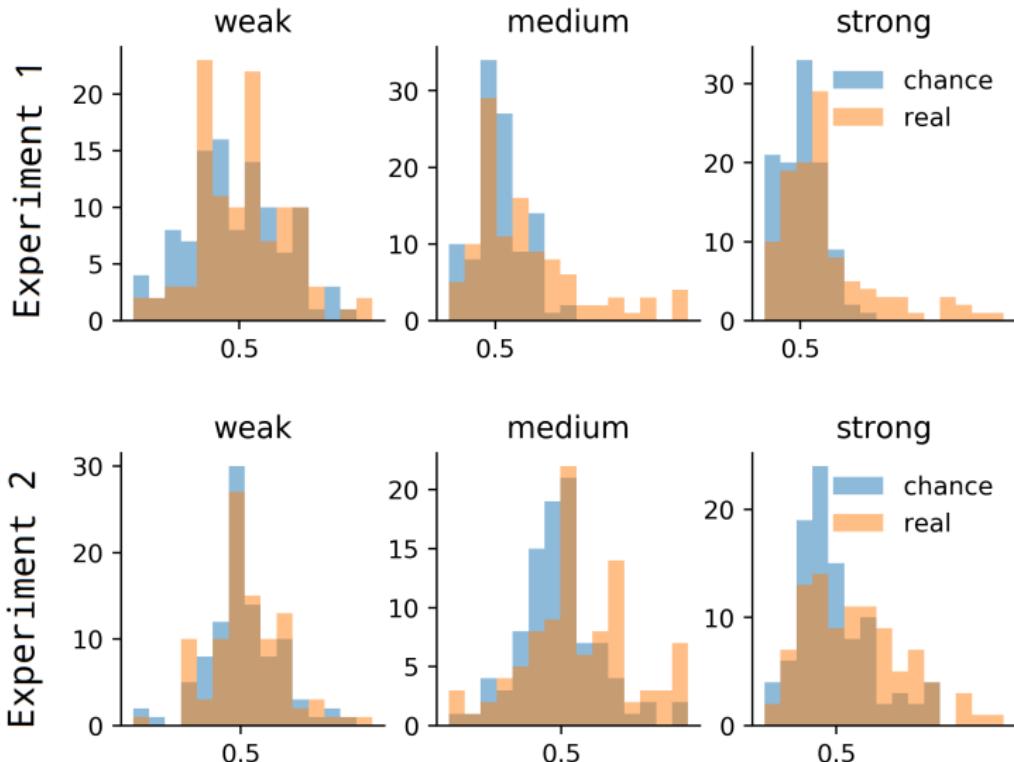
# SVM - Most Accurate Dendrites



# SVM - Statistical Significance



# SVM - Statistical Significance



## Rank Order Correlation

We want to quantify the similarity between two rank orders of dendrites  $R$  and  $Q$ , which both have length  $n$ .

## Rank Order Correlation

We want to quantify the similarity between two rank orders of dendrites  $R$  and  $Q$ , which both have length  $n$ .

Dendrite	$R_i$	$Q_i$	$R_{acc}$	$Q_{acc}$
1	1	2	0.78	0.81
2	2	1	0.72	0.76
3	3	3	0.65	0.70
4	4	4	0.58	0.57
5	5	16	0.51	0.53
6	6	75	0.53	0.48
7	7	8	0.51	0.52
8	8	89	0.52	0.49
9	9	100	0.50	0.51
...	...	...	...	...

## Rank Order Correlation

Standard approach: **Spearman's rank order coefficient  $\rho$**

## Rank Order Correlation

Standard approach: **Spearman's rank order coefficient  $\rho$**

$$\rho(R, Q) = 1 - \frac{6 \sum_i^n D_i^2}{n(n^2 - 1)}$$

## Rank Order Correlation

Standard approach: **Spearman's rank order coefficient  $\rho$**

$$\rho(R, Q) = 1 - \frac{6 \sum_i^n D_i^2}{n(n^2 - 1)}$$

Problem: Many of the dendrites are not significant for classification but affect  $\rho$  greatly. We would like to be able to **weight our R's and Q's**.

## Weighted Rank Order Correlation Coefficient

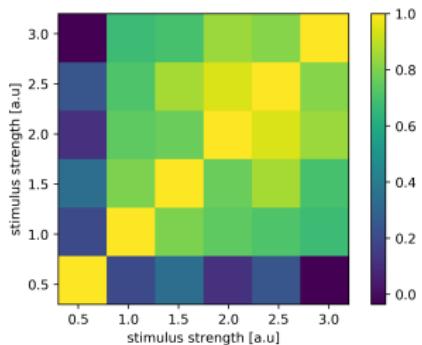
One possible solution:

$$\rho_W(R, Q) = 1 - \frac{-2 \sum_i^n w_i D_i^2}{\sum_i^n w_i (n - 2i + 1)^2}$$

With a monotonicity constraint on  $W$ .

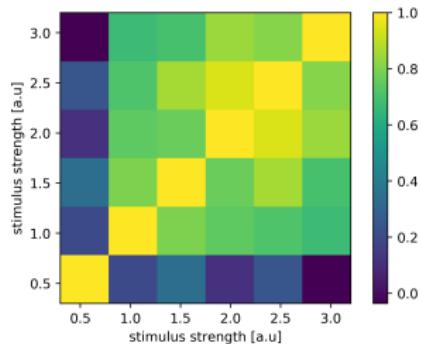
# Rank Order Correlations of Dendrites Over Stimuli

Experiment 1

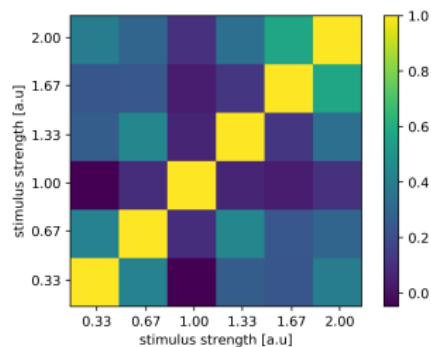


# Rank Order Correlations of Dendrites Over Stimuli

Experiment 1

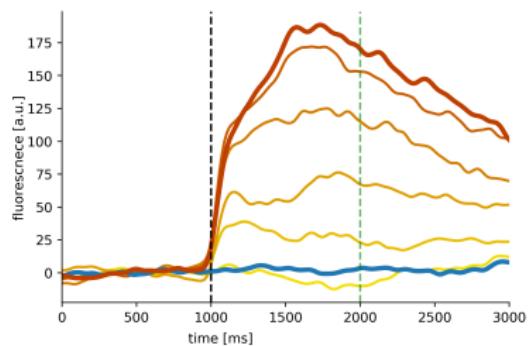


Experiment 2



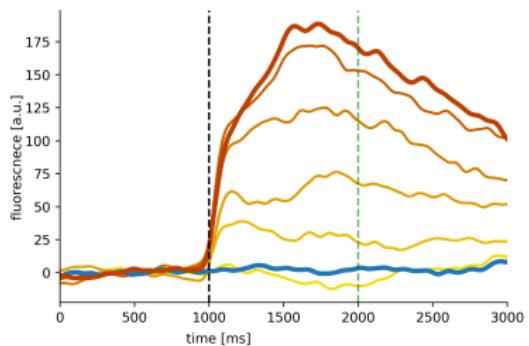
# Responses of One Dendrite to Different Stimuli

Experiment 1

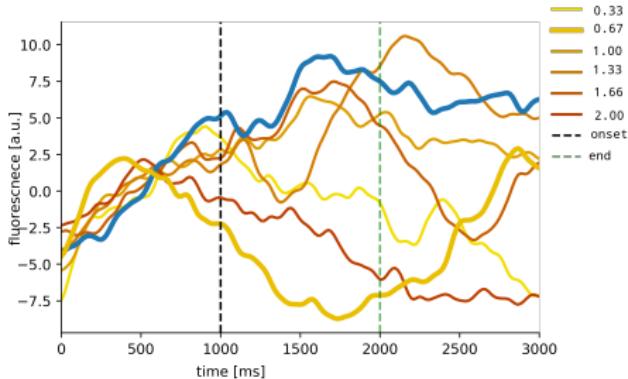


# Responses of One Dendrite to Different Stimuli

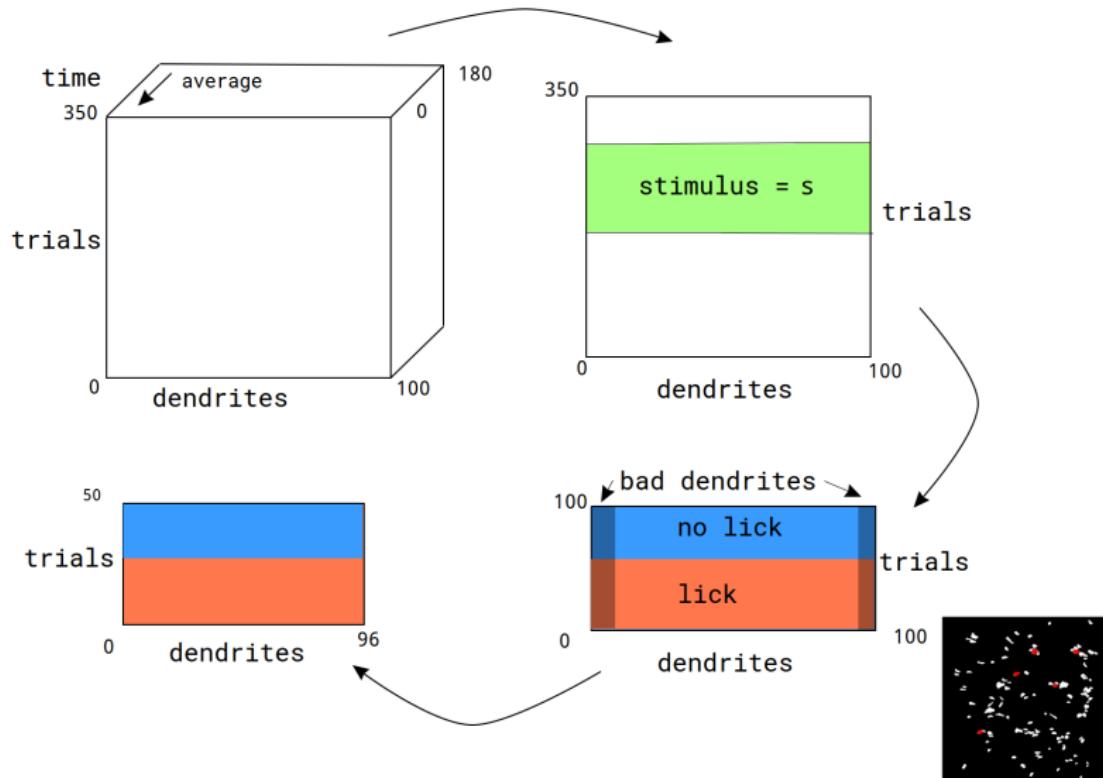
Experiment 1



Experiment 2



# Behavior Detection



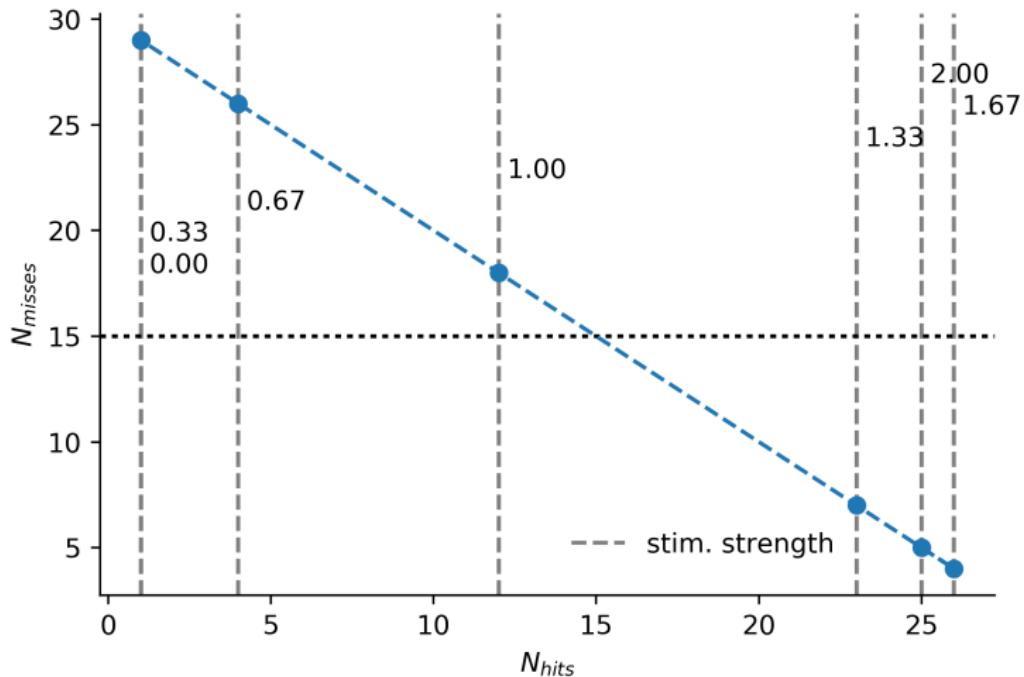
## Behavioral Prediction - Most Accurate Dendrites

Different dataset - Stimulus strength 1 (near threshold)

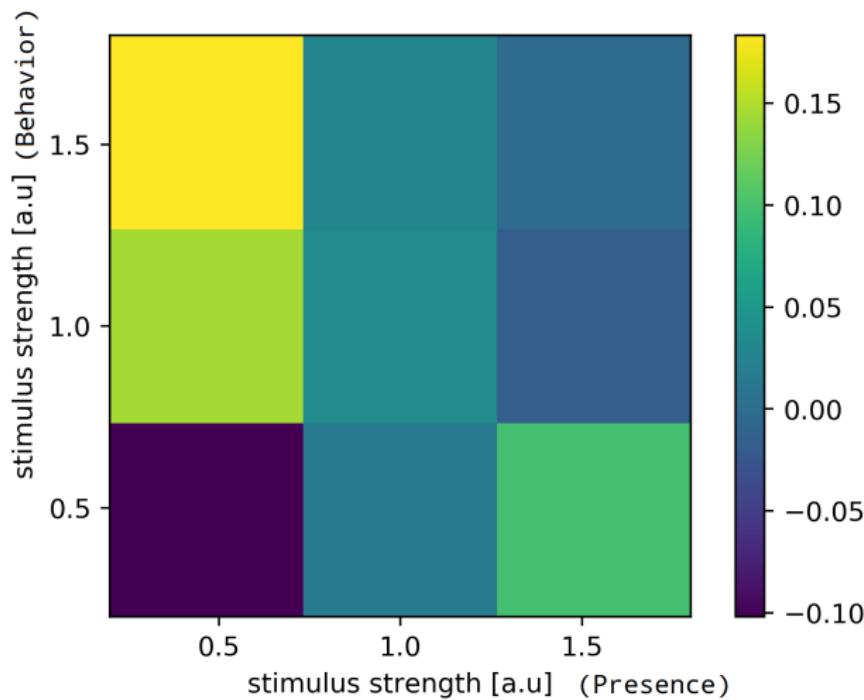
Dendrite #	$\mu_{acc}$	$\sigma_{acc}$
88	0.80	0.08
57	0.80	0.08
92	0.78	0.06
56	0.76	0.10

# Behavioral Prediction - Imbalance

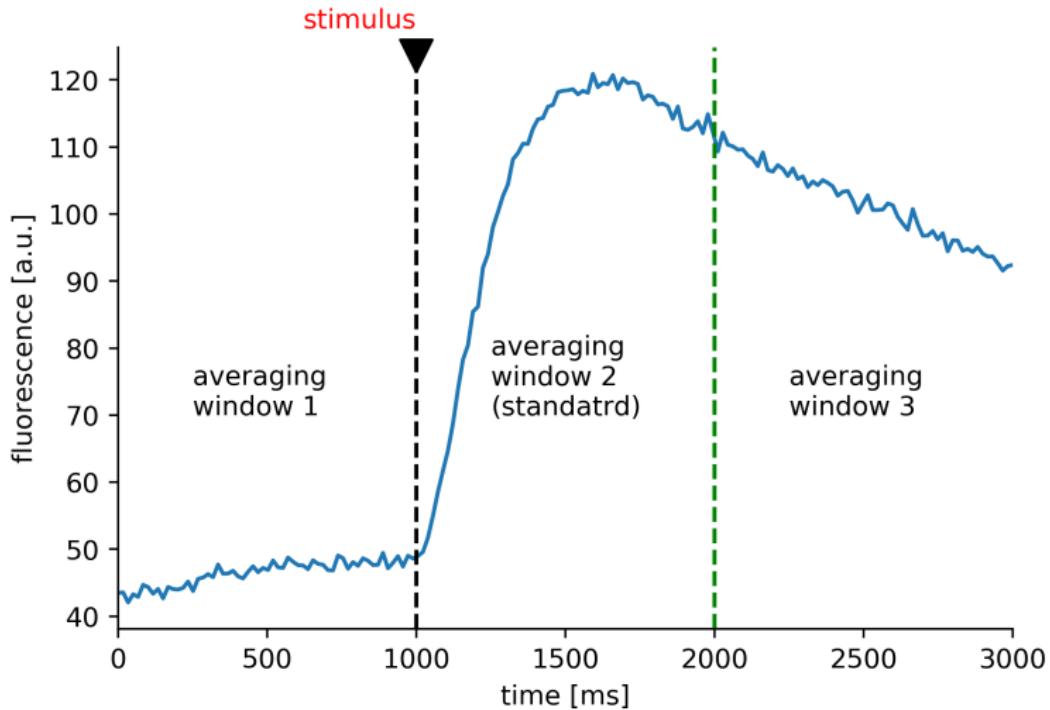
$N_{no-lick}$  vs.  $N_{lick}$  for all stimuli



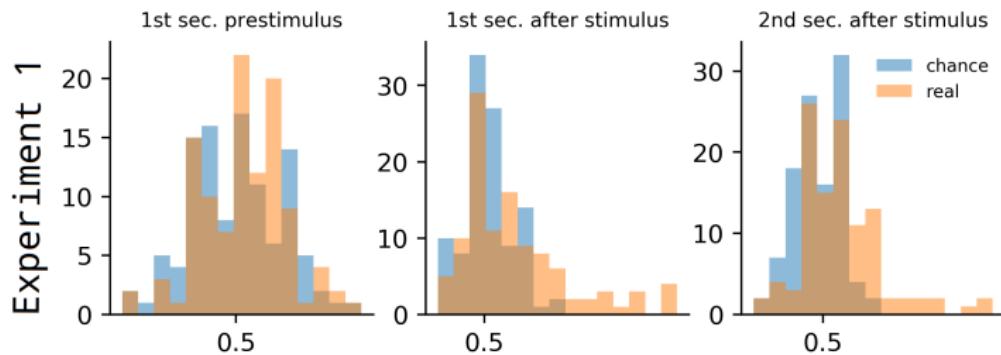
# Rank Order Correlations of Best Behavior and Presence Dendrites



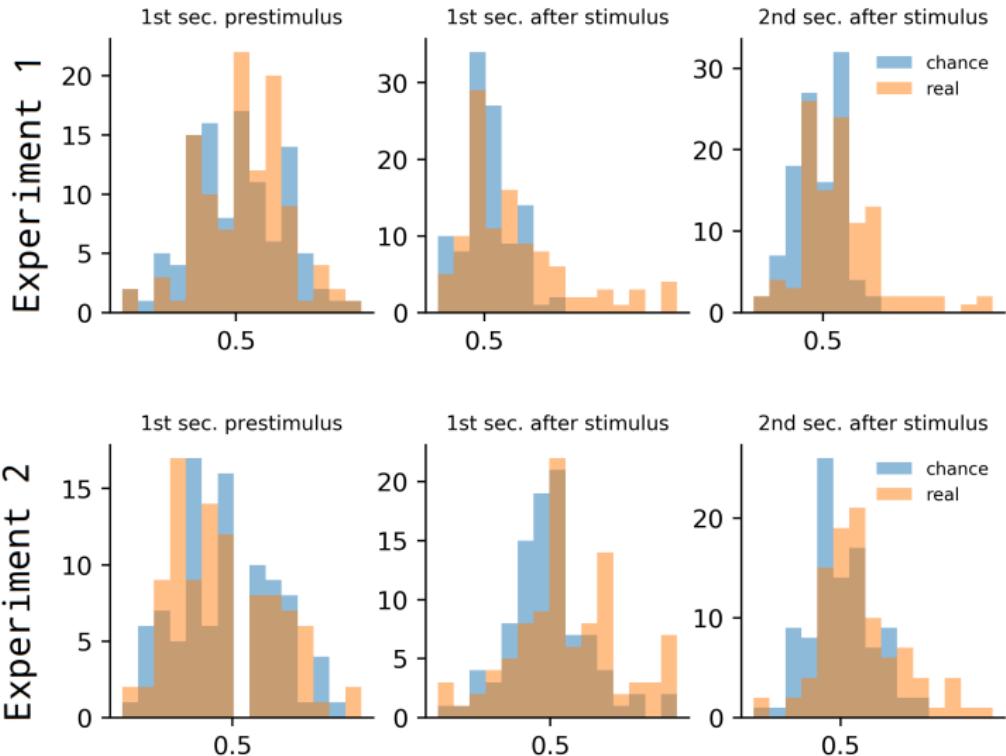
# Timing - What if We Use Different Averaging Windows?



# Timing - Statistical Significance

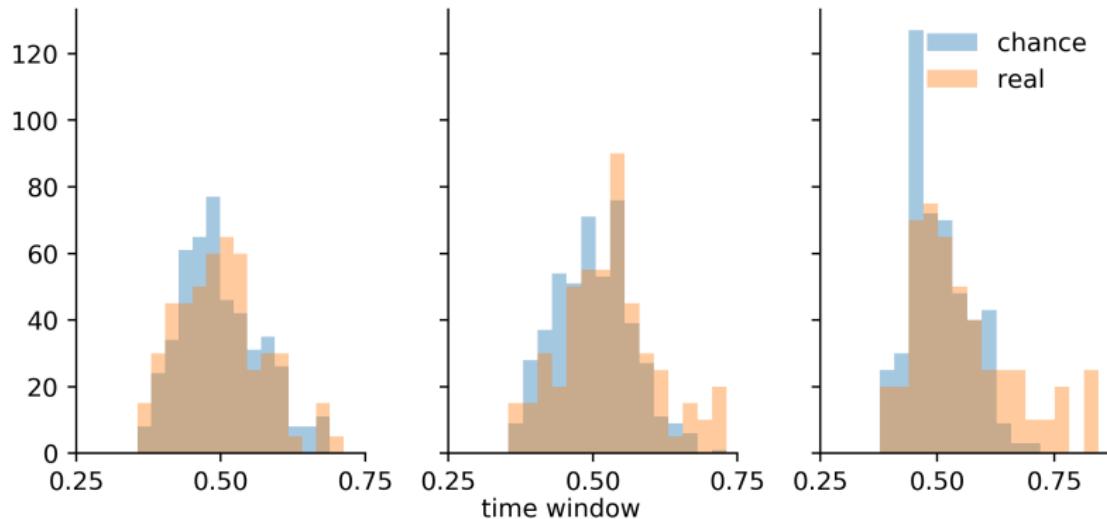


# Timing - Statistical Significance



# Statistical Significance - Behavior

At threshold stimulus  $\approx 1$



## Multivariate SMV Analysis

---

## Presence Detection - Feature Selection

Since we have fewer datapoints per condition (30-50 per class) than features (100-150), we are prone to overfit to noise (which is abundant in our data).

## Presence Detection - Feature Selection

Since we have fewer datapoints per condition (30-50 per class) than features (100-150), we are prone to overfit to noise (which is abundant in our data).

⇒ we have to do **feature selection**.

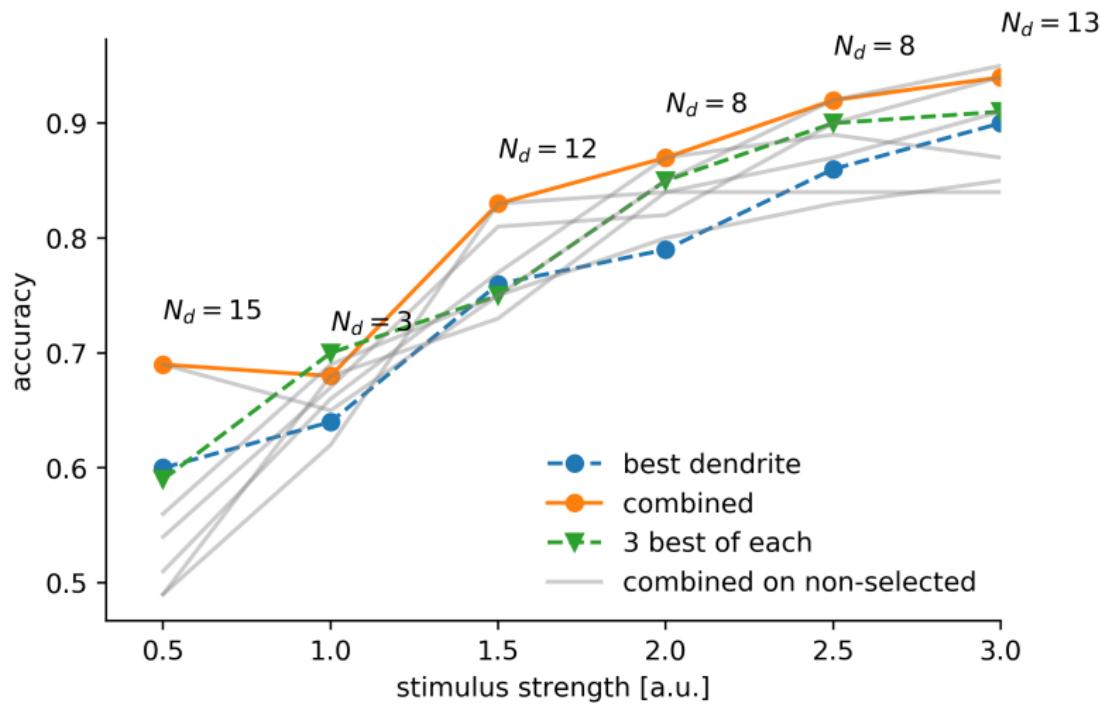
## Presence Detection - Feature Selection

Since we have fewer datapoints per condition (30-50 per class) than features (100-150), we are prone to overfit to noise (which is abundant in our data).

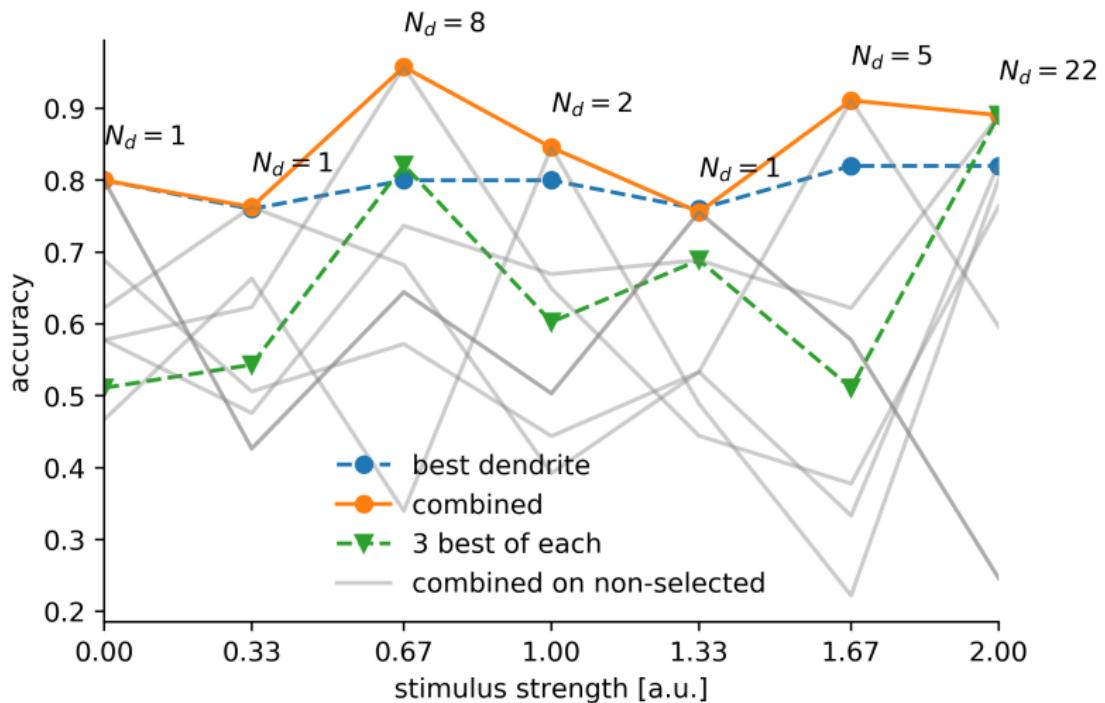
⇒ we have to do **feature selection**.

We can use the previously best discriminating dendrites.

# SVM Performance on Combined Dendrites - Presence Detection, Experiment 1



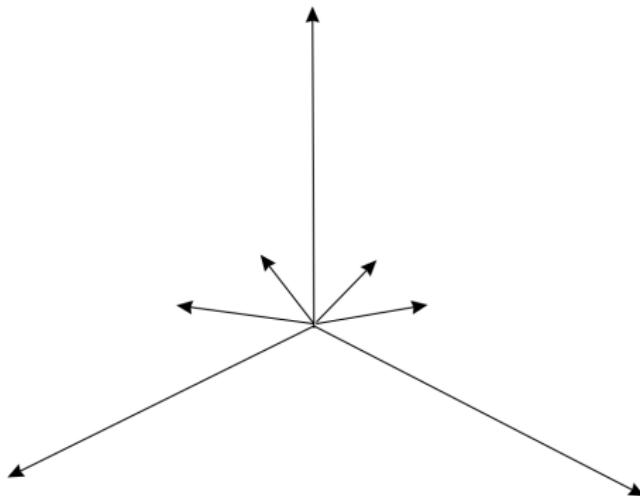
# SVM Performance on Combined Dendrites - Behavioral



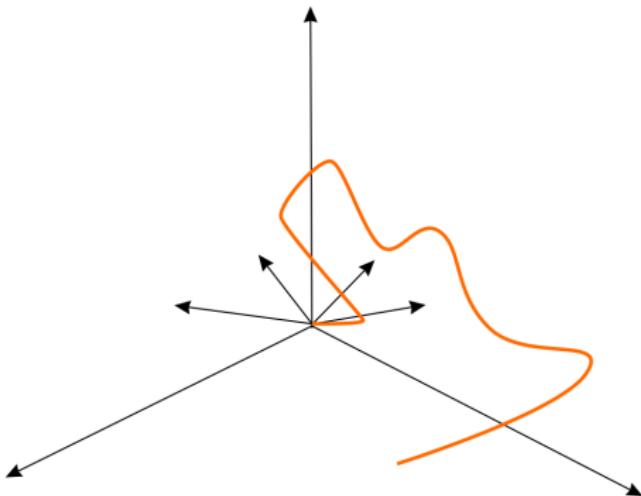
## Population Coding

---

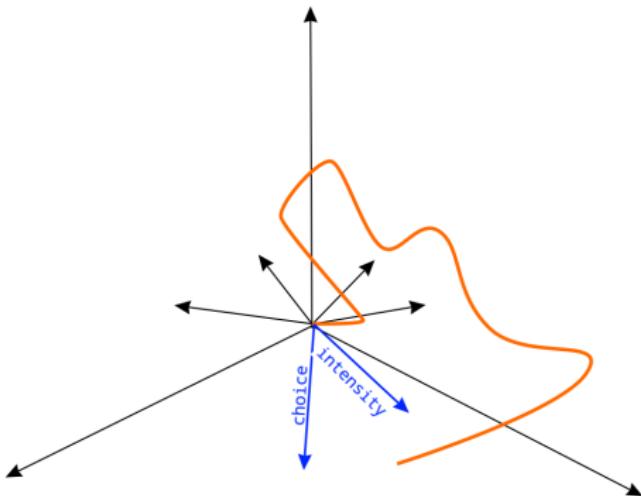
## Population Coding - Idea (Mante, Sussillo 2013)



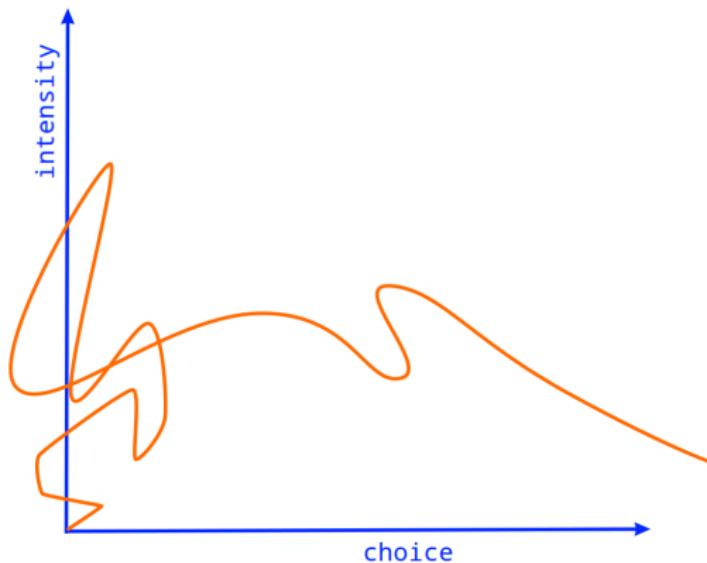
## Population Coding - Idea (Mante, Sussillo 2013)



## Population Coding - Idea (Mante, Sussillo 2013)

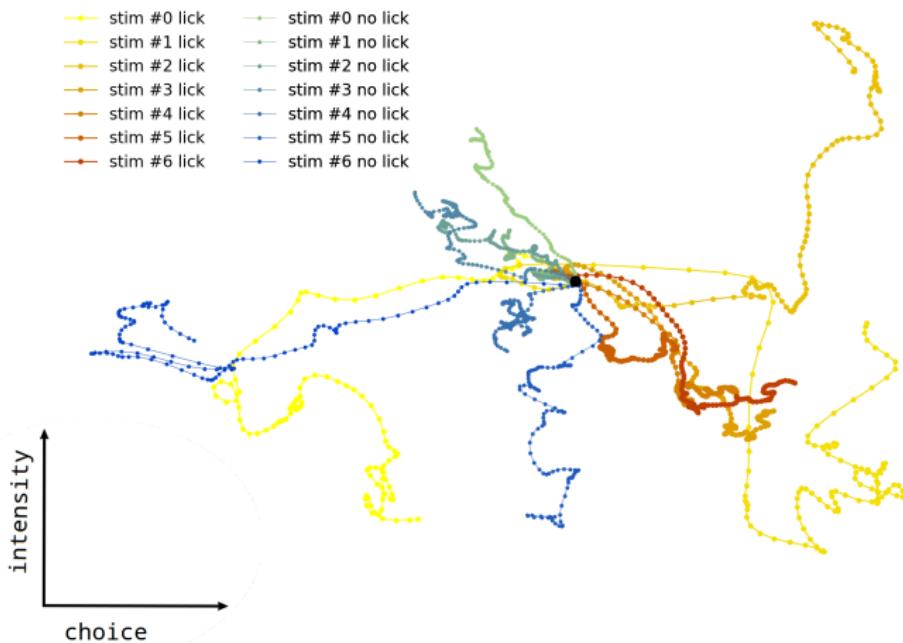


## Population Coding - Idea (Mante, Sussillo 2013)



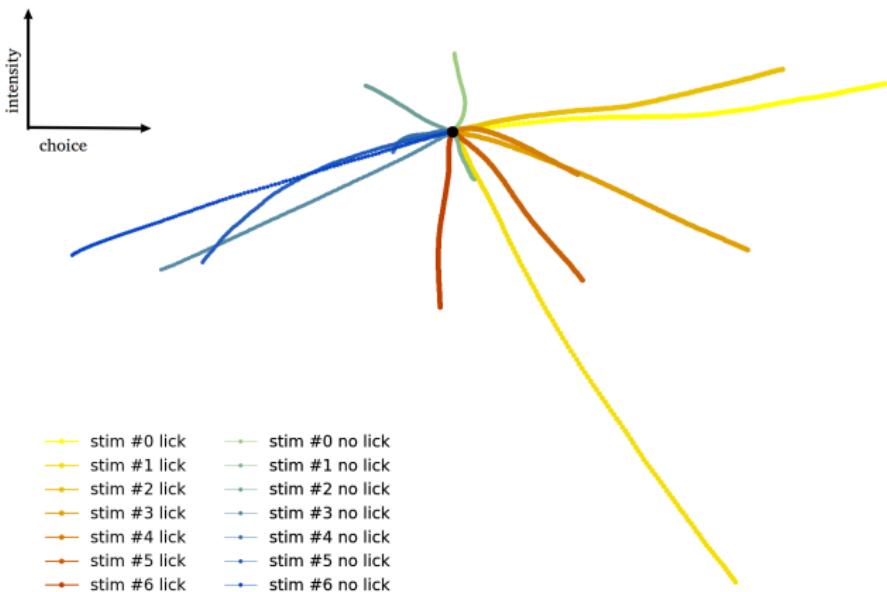
# Population Response in Task Variable Space

Solve Problem with combined SVM first



# Population Response in Task Variable Space

Solve Problem with combined SVM first



# Conclusions

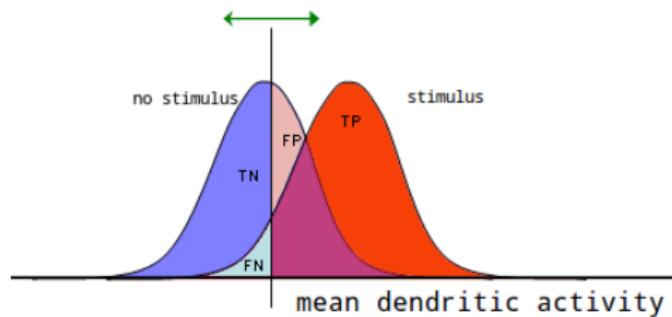
- We can build high accuracy classifiers based on this data
- Multivariate analysis is advantageous
- Prestimulus  $\text{Ca}^{2+}$  activity does not predict the animal's decision
- Different dendrites code for stimulus and behavior
- There are very large differences between different animals/experiments
- We need more data - particularly for multivariate analysis

-  Takahashi, N., Oertner, G. T., Hegemann, P. & Larkum, E. M. Active cortical dendrites modulate perception. *Science* 335, 1587-1590 (2016)
-  Mante V., Sussillo D., Shenoy KV., Newsome WT. Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* 503, 78-84 (2013)
-  Pinto da Costa, J. New Results in Weighted Correlation and Weighted Principal Component Analysis with Applications, Chapter 2 (2015)

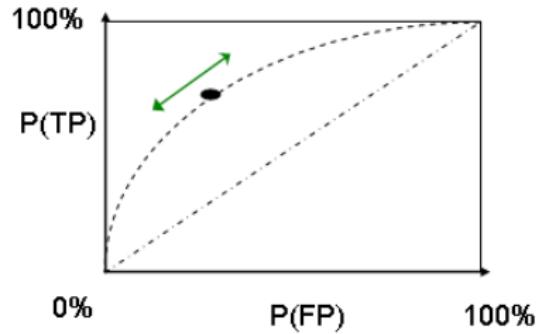
## Extra Slides

---

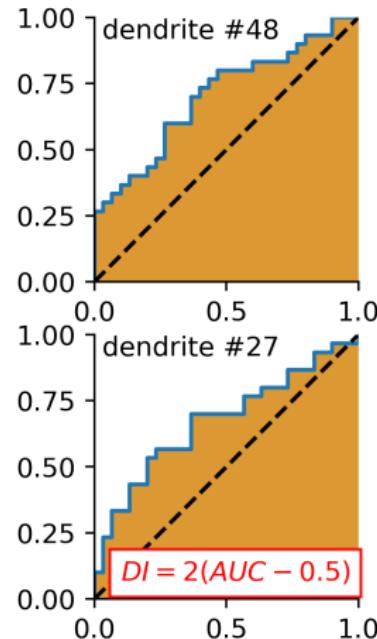
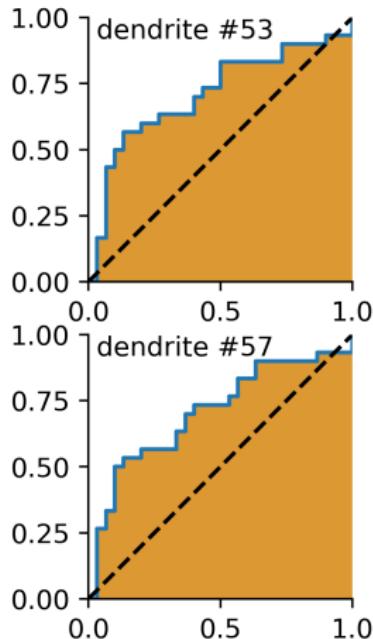
# ROC



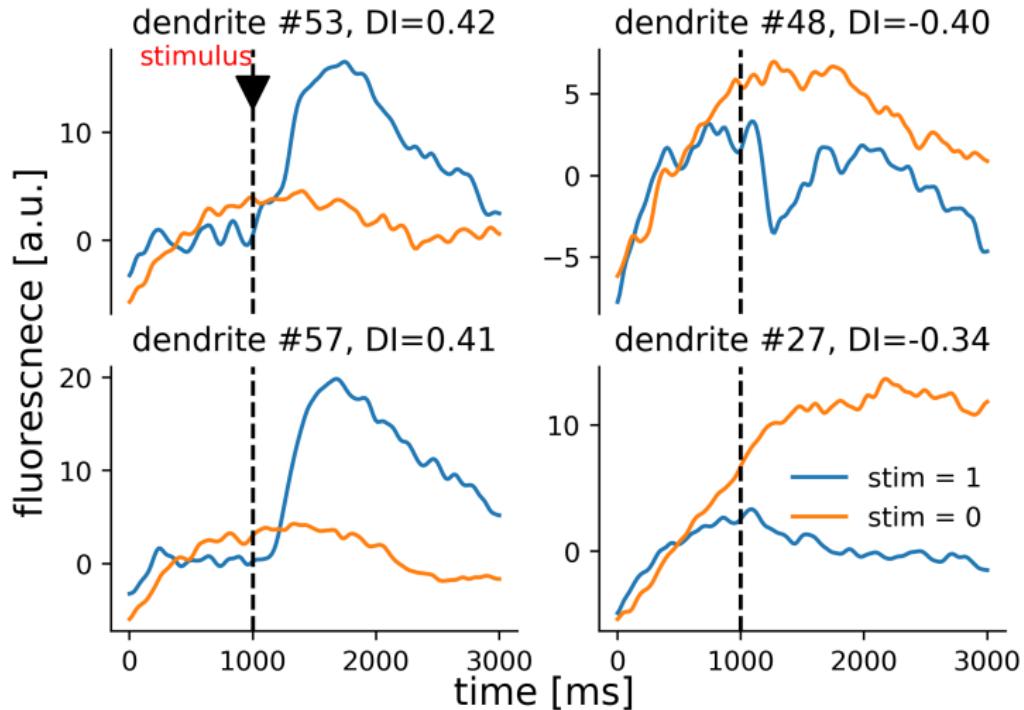
TP	FP
FN	TN
1	1



## ROC - Best ROC Curves

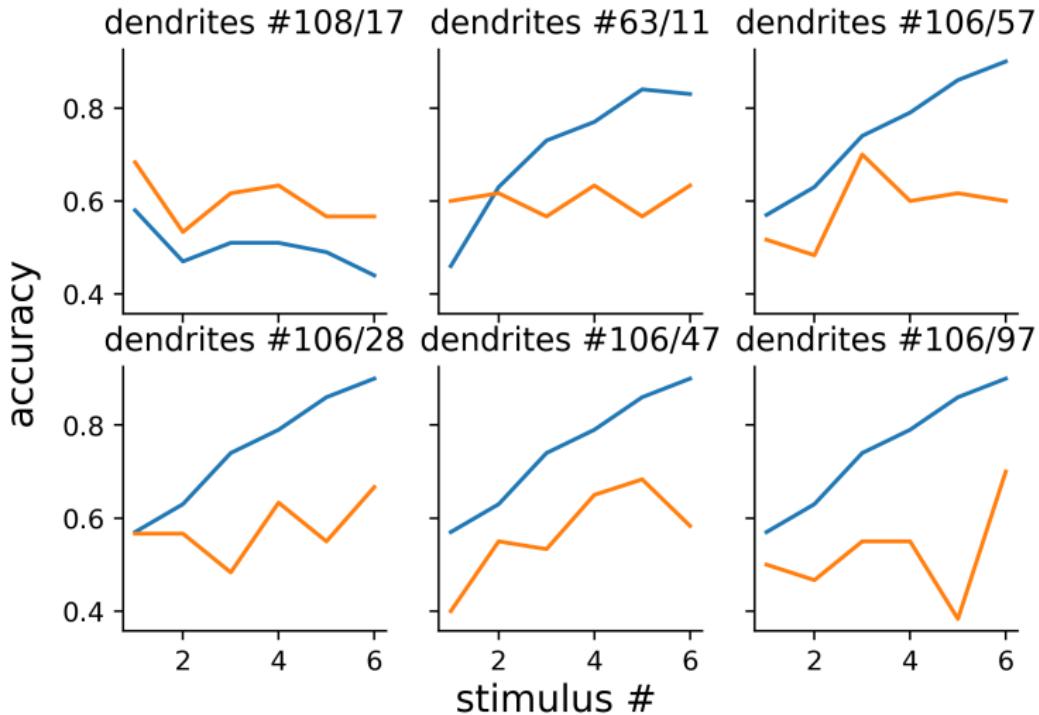


# $\text{Ca}^{2+}$ -Traces of Largest $|\text{AUC} - 0.5|$ -Dendrites

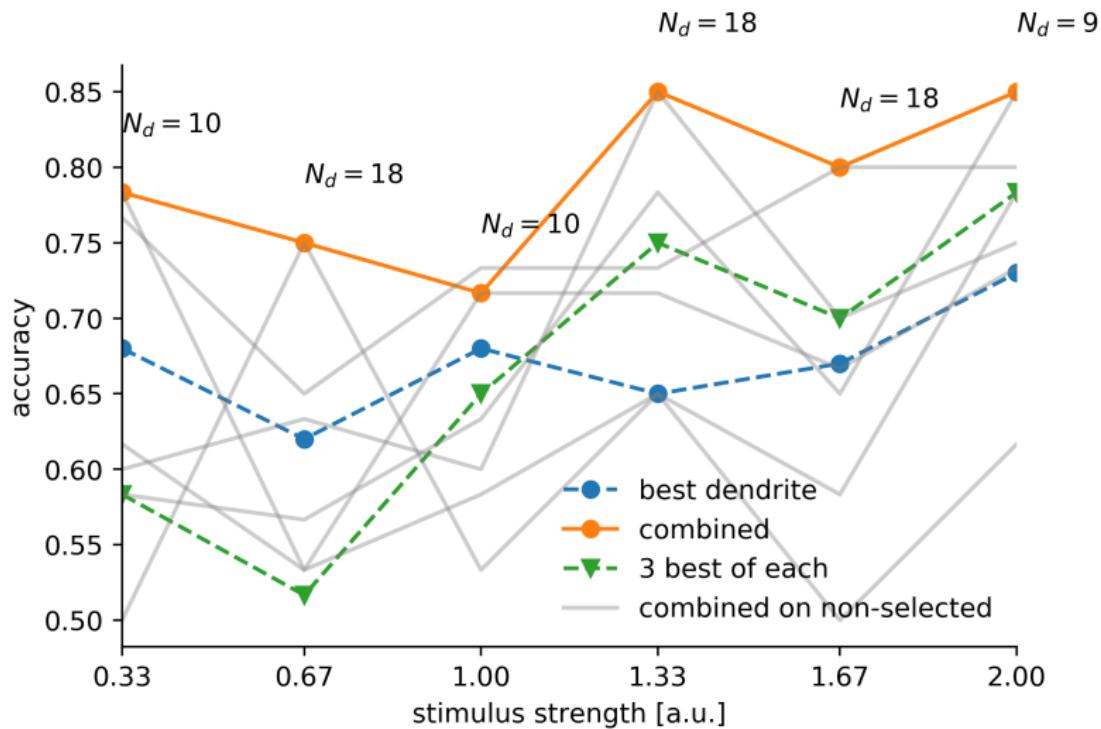


Near threshold-stimulus ( $\approx 1$ )

## Tuning Curves



# SVM Performance on Combined Dendrites - Presence Detection, Experiment 2



## Weighted Rank Order Correlation Coefficient

We want to quantify the similarity between two rank orders of dendrites  $R$  and  $Q$ , which both have length  $n$ .

## Weighted Rank Order Correlation Coefficient

We want to quantify the similarity between two rank orders of dendrites  $R$  and  $Q$ , which both have length  $n$ .

Standard approach: Spearman's rank order coefficient

$$\rho(R, Q) = 1 - \frac{6 \sum_i^n D_i^2}{n(n^2 - 1)}$$

where  $D_i = R_i - Q_i$ .

## Weighted Rank Order Correlation Coefficient

We want to quantify the similarity between two rank orders of dendrites  $R$  and  $Q$ , which both have length  $n$ .

Standard approach: Spearman's rank order coefficient

$$\rho(R, Q) = 1 - \frac{6 \sum_i^n D_i^2}{n(n^2 - 1)}$$

where  $D_i = R_i - Q_i$ .

Problem: Many of the dendrites are not significant for classification but affect  $\rho$  greatly. We would like to be able to **weight our  $D_i$ 's**.

## Weighted Rank Order Correlation Coefficient

We want our rank order coefficient to be an affine linear function of  $\sum_i w_i D_i^2$ :

$$\rho_W(R, Q) = A + B \sum_i^n w_i D_i^2$$

## Weighted Rank Order Correlation Coefficient

We want our rank order coefficient to be an affine linear function of  $\sum_i w_i D_i^2$ :

$$\rho_W(R, Q) = A + B \sum_i^n w_i D_i^2$$

If  $R = Q$ , we want  $\rho_W = 1$ , and if the rankings are reversed, we want  $\rho_W = -1$ .

## Weighted Rank Order Correlation Coefficient

We want our rank order coefficient to be an affine linear function of  $\sum_i w_i D_i^2$ :

$$\rho_W(R, Q) = A + B \sum_i^n w_i D_i^2$$

If  $R = Q$ , we want  $\rho_W = 1$ , and if the rankings are reversed, we want  $\rho_W = -1$ .

The first condition gives us  $A = 1$ , and the second one

$$B = \frac{-2}{\sum_i^n w_i (n - 2i + 1)^2}$$

## Weighted Rank Order Correlation Coefficient

We want our rank order coefficient to be an affine linear function of  $\sum_i w_i D_i^2$ :

$$\rho_W(R, Q) = A + B \sum_i^n w_i D_i^2$$

If  $R = Q$ , we want  $\rho_W = 1$ , and if the rankings are reversed, we want  $\rho_W = -1$ .

The first condition gives us  $A = 1$ , and the second one

$$B = \frac{-2}{\sum_i^n w_i (n - 2i + 1)^2}$$

Thus:

$$\boxed{\rho_W(R, Q) = 1 - \frac{-2 \sum_i^n w_i D_i^2}{\sum_i^n w_i (n - 2i + 1)^2}}$$

## Population Response in Task Variable Space

---

We would like to visualize the time evolution of the population response in a coordinate system in which the axes represent stimulus strength and choice.

## Population Response in Task Variable Space

We would like to visualize the time evolution of the population response in a coordinate system in which the axes represent stimulus strength and choice.

To that end we use linear regression to write the normalized response of dendrite  $i$  at time  $t$  in trial  $k$  as a linear combination of these task variables:

$$r_k^{i,t} = \beta_1^{i,t} choice_k + \beta_2^{i,t} stimulus_k + \beta_3^{i,t}$$

## Population Response in Task Variable Space

We would like to visualize the time evolution of the population response in a coordinate system in which the axes represent stimulus strength and choice.

To that end we use linear regression to write the normalized response of dendrite  $i$  at time  $t$  in trial  $k$  as a linear combination of these task variables:

$$r_k^{i,t} = \beta_1^{i,t} choice_k + \beta_2^{i,t} stimulus_k + \beta_3^{i,t}$$

The regression coefficients  $\beta_\nu^{i,t}$  describe how much the activity of dendrite  $i$  at time  $t$  in trial  $k$  corresponds with variable  $\nu$ .

# Population Response in Task Variable Space

We define

$$\mathbf{F} = \begin{bmatrix} choice_1 & \dots & choice_n \\ stimulus_1 & \dots & stimulus_n \\ 1 & \dots & 1 \end{bmatrix}$$

and estimate for each dendrite  $i$  and timepoint  $t$

$$\beta^{i,t} = (\mathbf{F}\mathbf{F}^T)^{-1}\mathbf{F}r^{i,t}$$

## Population Response in Task Variable Space

In total we have 14 conditions: 7 stimuli X two choices.

## Population Response in Task Variable Space

In total we have 14 conditions: 7 stimuli X two choices.

We define:

$$\mathbf{x}^{c,t}$$

as the (trial-averaged) population response to condition  $c$  (for example lick at stimulus 1.00) at time  $t$ , which is a vector of  $N_{unit}$  length.

## Population Response in Task Variable Space

In total we have 14 conditions: 7 stimuli X two choices.

We define:

$$\mathbf{x}^{c,t}$$

as the (trial-averaged) population response to condition  $c$  (for example lick at stimulus 1.00) at time  $t$ , which is a vector of  $N_{unit}$  length.

The goal is to use  $\beta$  to find a two-dimensional subspace of the dendrite space into which we can transform  $\mathbf{x}^{c,t}$ .

# Population Response in Task Variable Space

We then use PCA to denoise the data matrix

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_{c_1, t_1} & \dots & \mathbf{x}_{c_1, t_n} & \dots & \mathbf{x}_{c_m, t_n} \end{bmatrix}$$

# Population Response in Task Variable Space

We then use PCA to denoise the data matrix

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_{c_1, t_1} & \dots & \mathbf{x}_{c_1, t_n} & \dots & \mathbf{x}_{c_m, t_n} \end{bmatrix}$$

and we apply the denoising matrix to all  $\beta^{\nu, t}$  as well, which yields the denoised regression vectors

$$\beta_{pca}^{\nu, t}$$

## Population Response in Task Variable Space

Since the  $\beta_{pca}^{\nu,t}$  are time varying but we would like a coordinate system that is fixed in time, we "freeze" them at the point in time at which there is maximum correlation:

# Population Response in Task Variable Space

Since the  $\beta_{pca}^{\nu, t}$  are time varying but we would like a coordinate system that is fixed in time, we "freeze" them at the point in time at which there is maximum correlation:

$$\beta_{max}^{\nu} = \beta_{pca}^{\nu, t_{max}}$$

$$t_{max}^{\nu} = argmax_t ||\beta_{pca}^{\nu, t}||$$

## Population Response in Task Variable Space

---

We would like these  $\beta_{max}^\nu$  to be the basis vectors of our new coordinate system, however, they are not yet orthogonal.

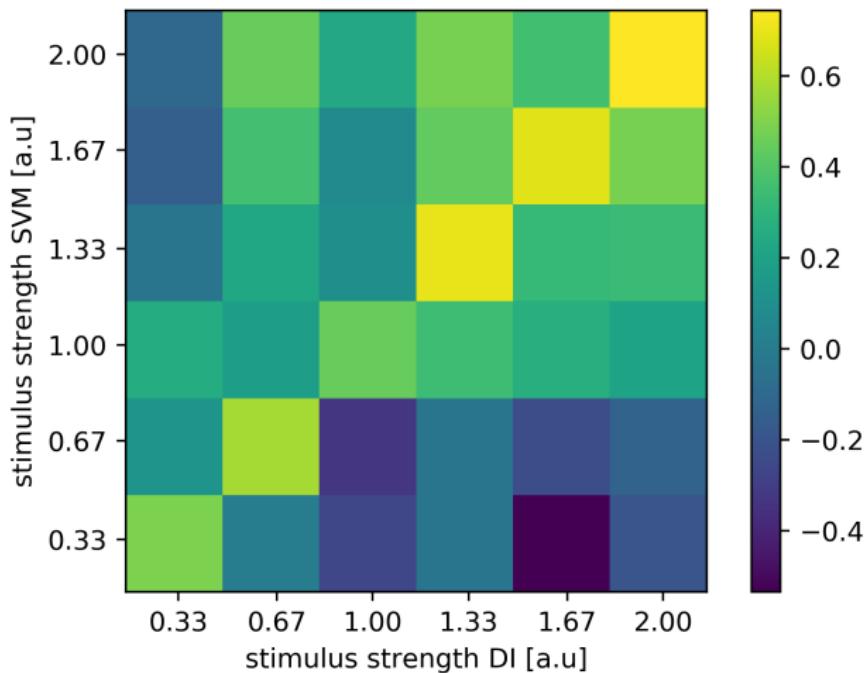
## Population Response in Task Variable Space

We would like these  $\beta_{\max}^\nu$  to be the basis vectors of our new coordinate system, however, they are not yet orthogonal. We fix this by applying QR-decomposition to

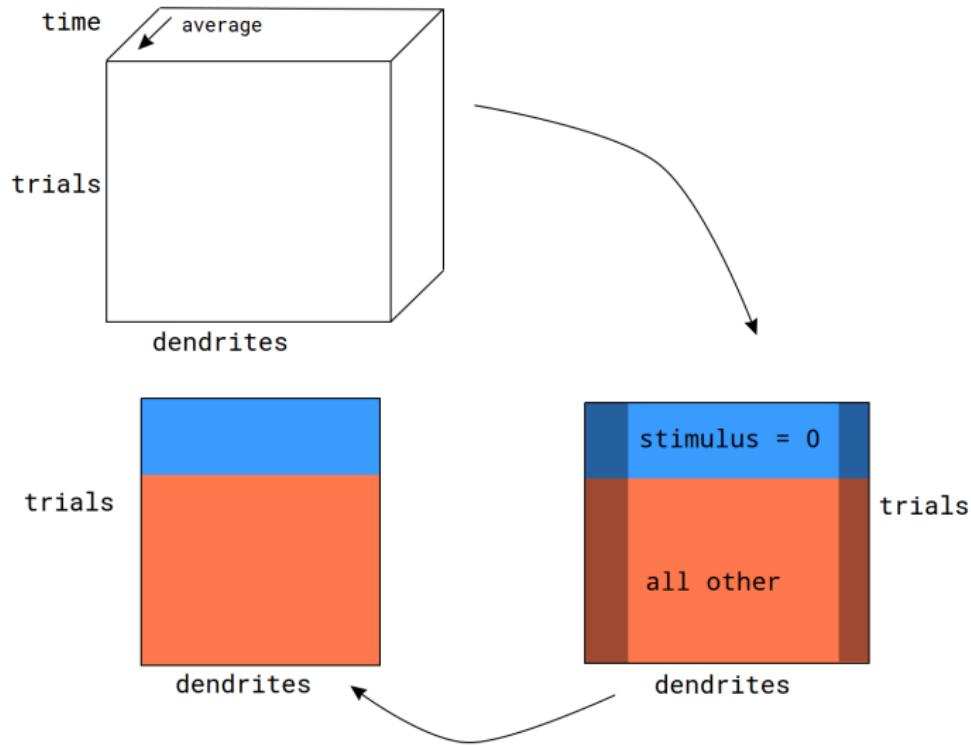
$$\mathbf{B}^{\max} = \begin{bmatrix} \beta_{\max}^1 & \beta_{\max}^2 \end{bmatrix} = \mathbf{Q}\mathbf{R}$$

where  $\mathbf{Q}$  is an orthogonal matrix whose columns  $\beta_\perp^\nu$  are the **basis vectors** of our new coordinate system. We can now transform our data into it.

# DI and SVM Dendrite Rank Order Correlation



# SVM Performance on Combined Dendrites - Global Presence



## SVM Performance on Combined Dendrites - Global Presence

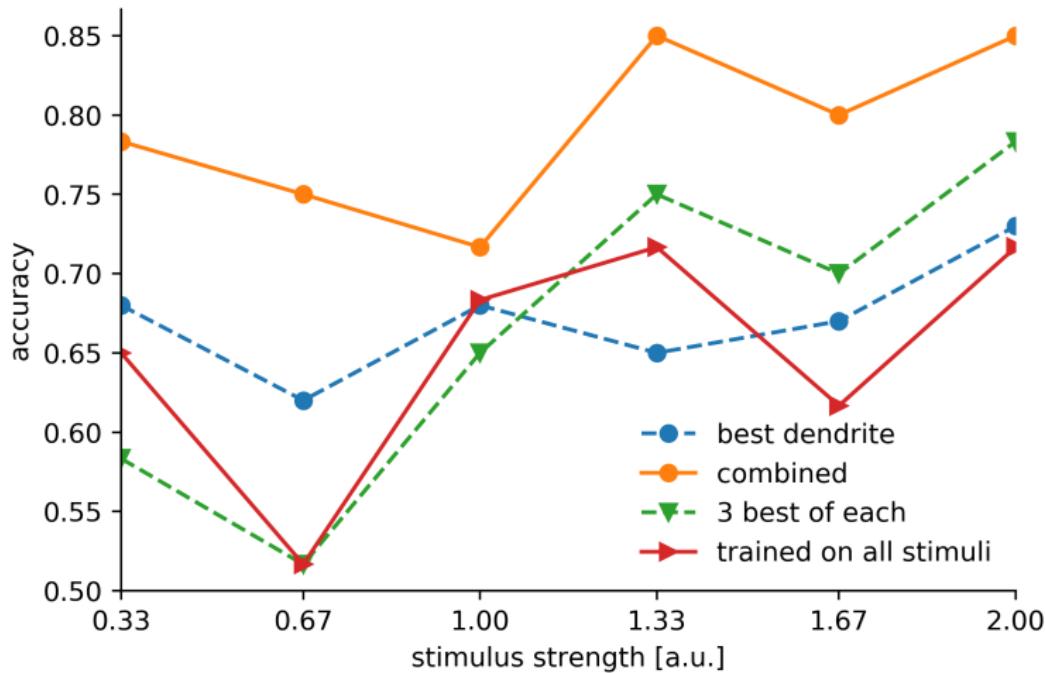
---

Performance on global presence detection:

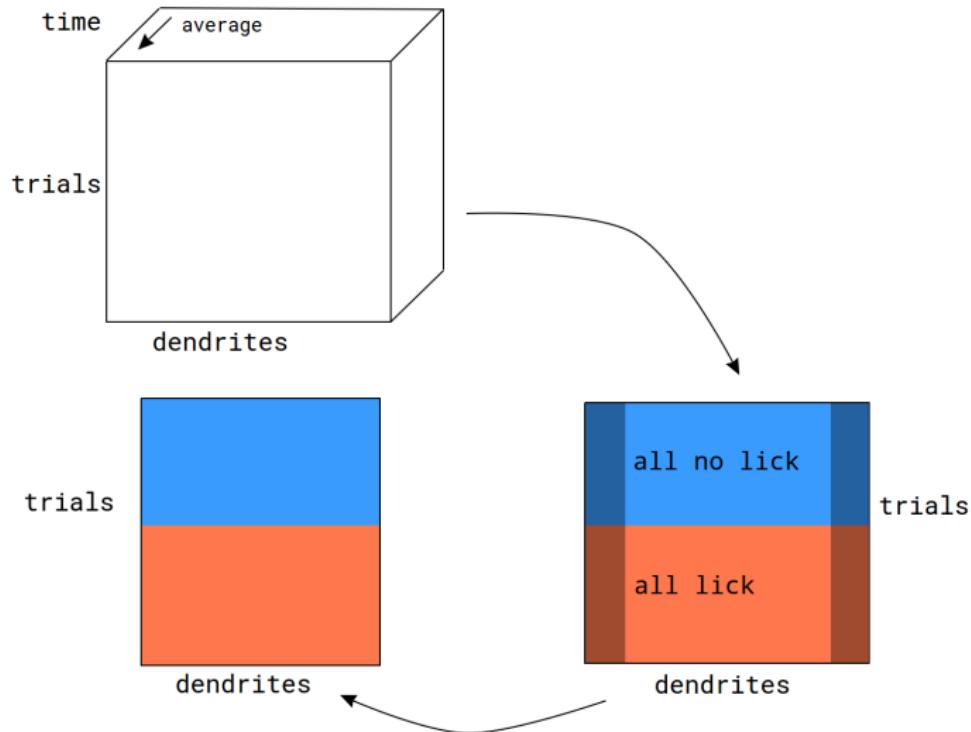
Mean: 0.68

Standard deviation: 0.08

# SVM Performance on Combined Dendrites - Presence Detection



# SVM Performance on Combined Dendrites - Global Behavior



## SVM Performance on Combined Dendrites - Global Behavior

Performance on global presence detection:

Mean: 0.71

Standard deviation: 0.1

# SVM Performance on Combined Dendrites - Behavioral

