Uni- and Multivariate Analysis of Dendritic Ca2+ Data

In a Stimulus Detection Task

Georg Chechelnizki

July 19, 2017

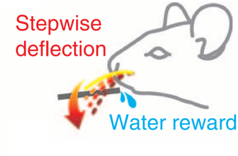
BCCN Berlin

Table of contents

Introduction

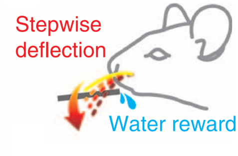
The Experiment

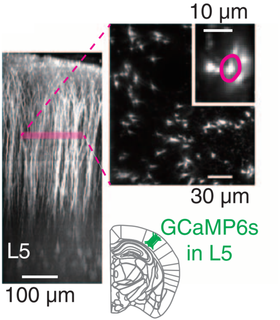
Takahashi et al. 2016



The Experiment

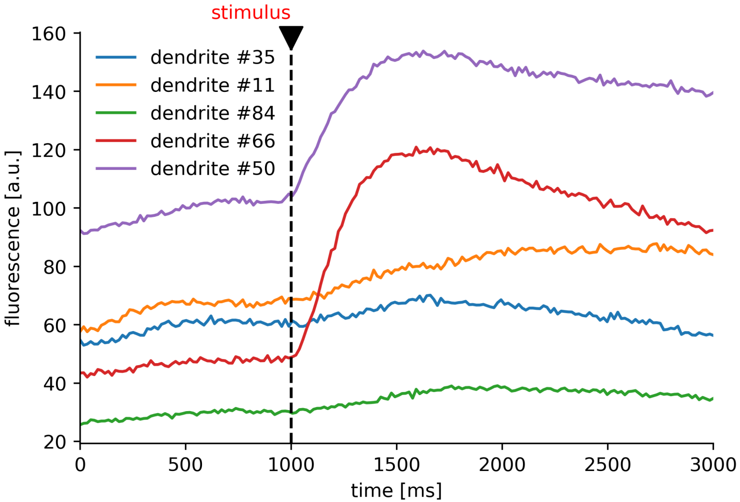
Takahashi et al. 2016





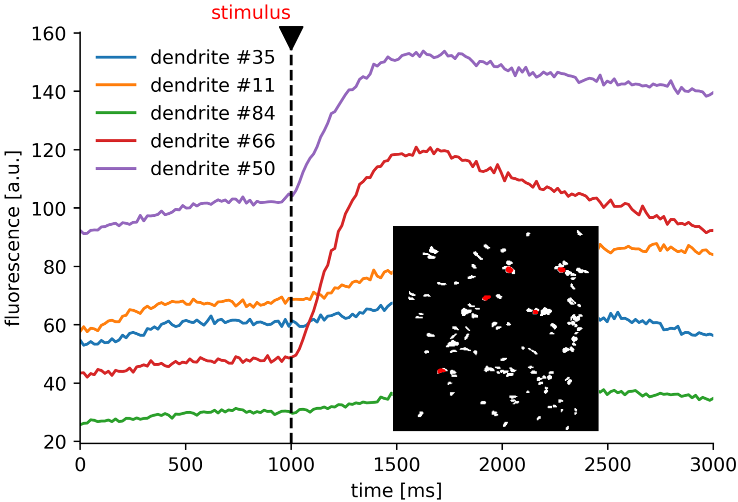
The Data - Neuronal

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

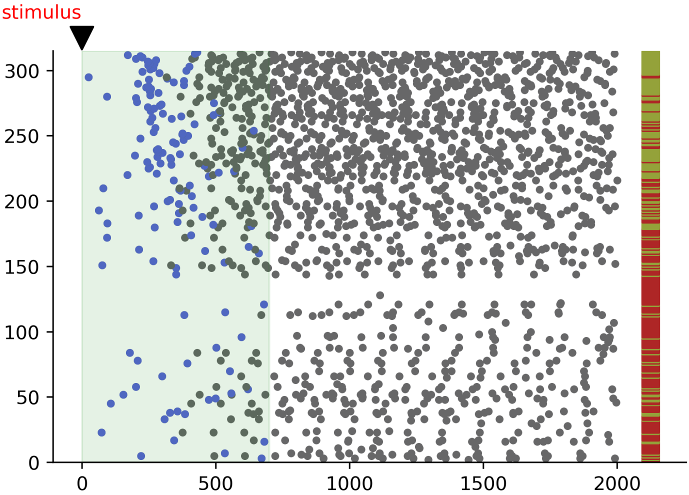


The Data - Neuronal

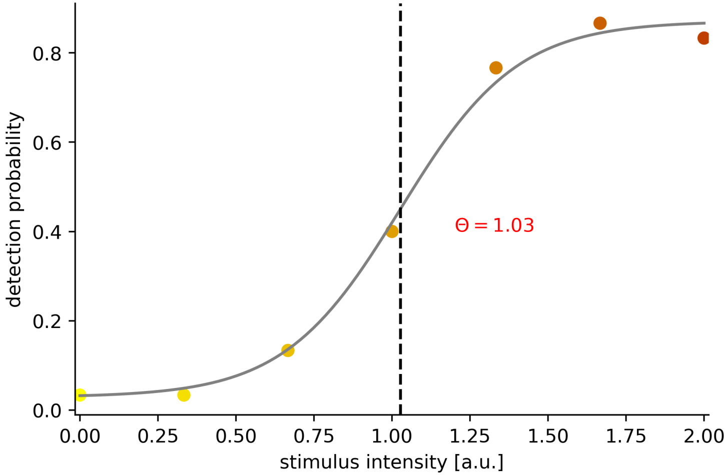
Trial-averaged Ca^{2+} fluorescence traces of random dendrites



The Data - Behavioral



The Data - Psychometric Curve



Goal

Goal

The goal of this project is to investigate the following:

- What is the relationship between Ca2+ activity and stimulus intensity/behavior

Goal

The goal of this project is to investigate the following:

- What is the relationship between Ca2+ activity and stimulus intensity/behavior
- Can we build good classifiers for this data?

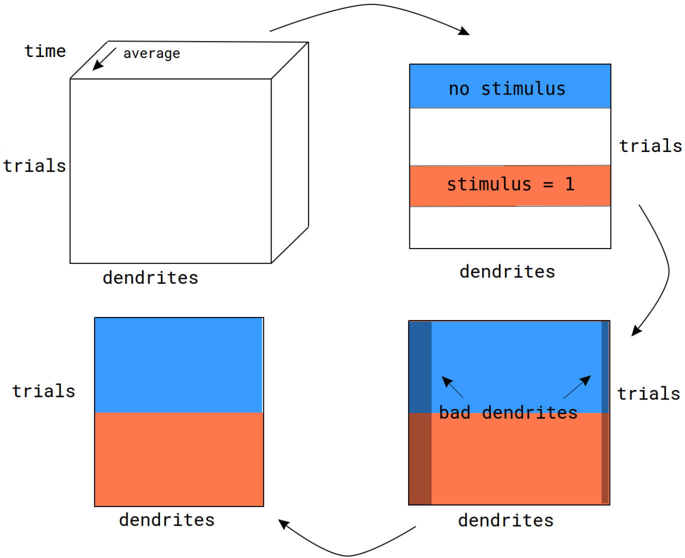
Goal

The goal of this project is to investigate the following:

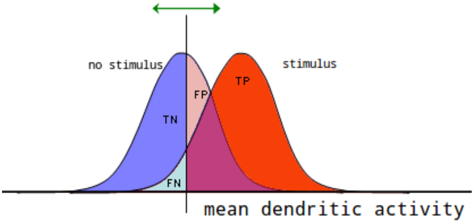
- What is the relationship between $Ca2+$ activity and stimulus intensity/behavior
- Can we build good classifiers for this data?
- Does a multivariate approach give us any significant advantage over a univariate one?

Univariate Analysis

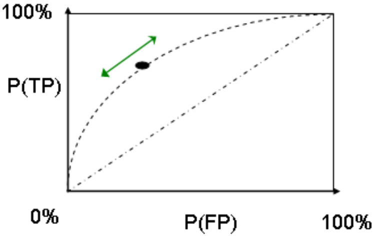
Stimulus Presence Detection



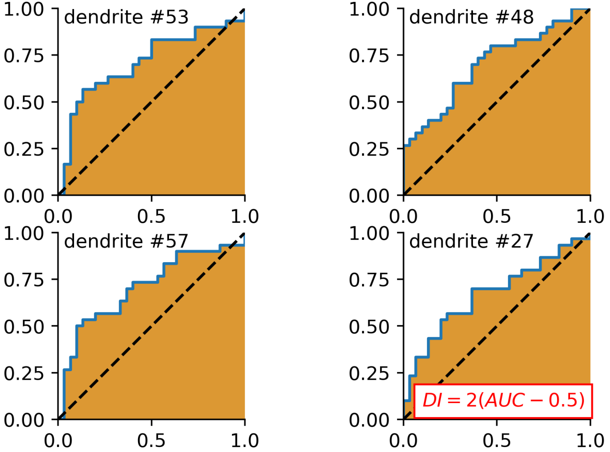
ROC



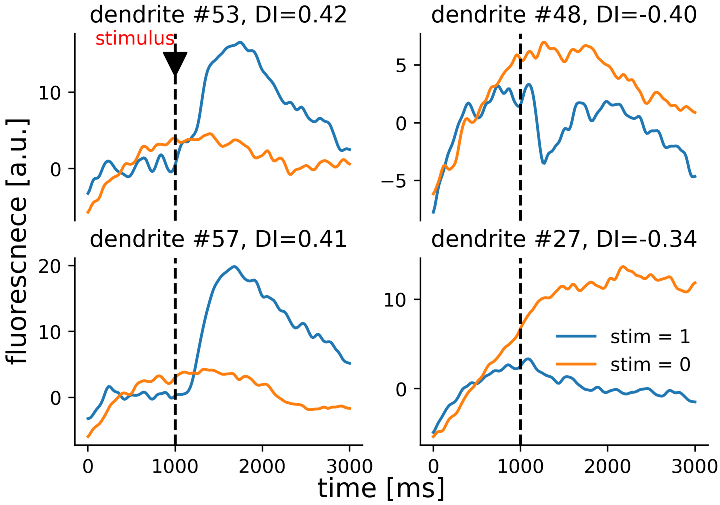
TP	FP
FN	TN
1	1



ROC - Best ROC Curves

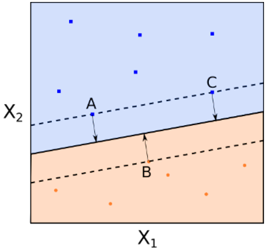


Ca^{2+}-Trace of Largest |DI|-Dendrites

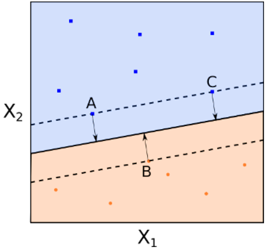


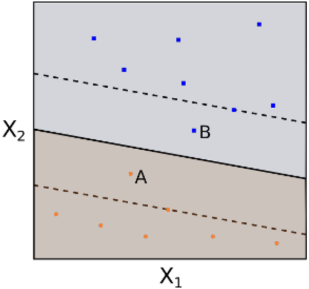
Near threshold-stimulus (≈ 1)

SVM - What is an SMV?

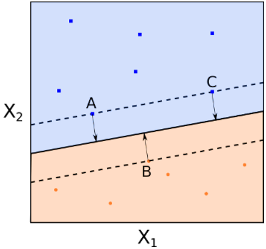


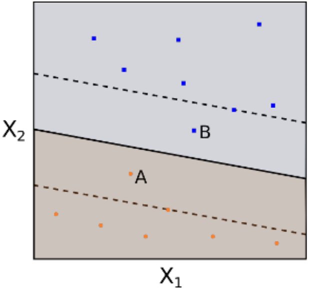
SVM - What is an SMV?

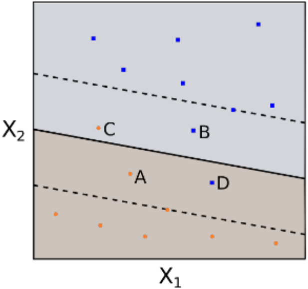




SVM - What is an SMV?







Why SVM?

- staple of machine learning
- regularization parameter limits overfitting
- good for small samples with many features (in theory)

SVM Specifics

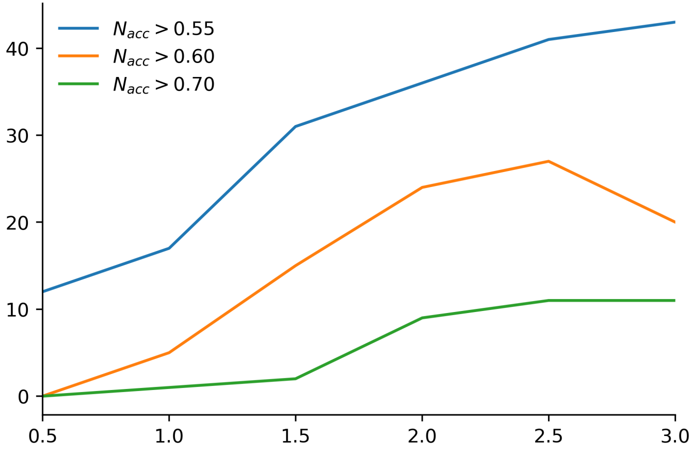
- All data are normalized to zero mean and unit variance
- Crossvalidation is performed to conrol for overfitting

SVM - Most Accurate Dendrites

Stimulus strength 1 (near thereshold)

Dendrite #	μ_{acc}	σ_{acc}
57	0.70	0.12
53	0.68	0.14
48	0.63	0.13
27	0.62	0.08

SVM - Numbers of Classifying Dendrites by Stimulus



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.

Standard approach: Spearman's rank order coefficient

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

where $D_i = R_i - Q_i$.

Rank Order Correlation

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 ρ greatly. We would like to be able to **weight our $\mathbf{D_i}$**'s.

Weighted Rank Order Correlation Coefficient

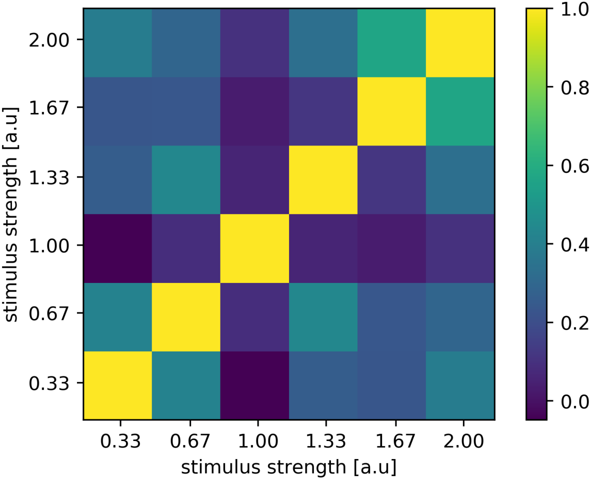
One possible soulution:

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

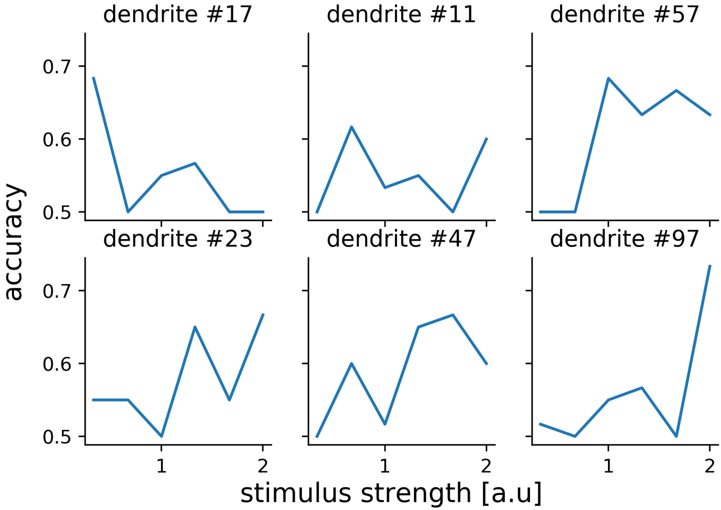
With a monotonicity constraint on W.

Rank Order Correlations of Dendrites Over Stimuli

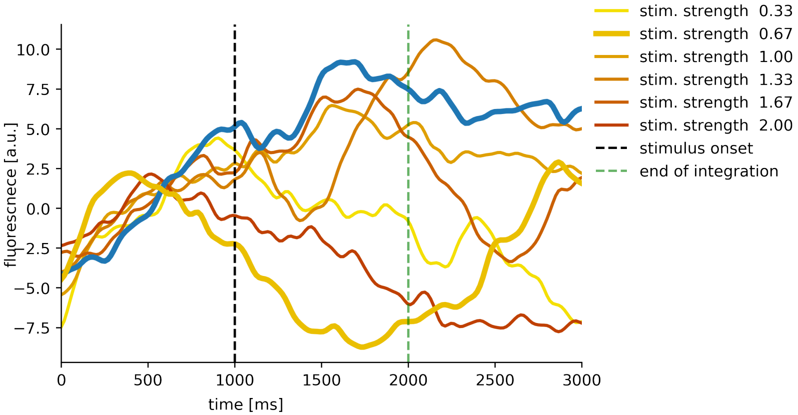
Weighted rank coefficient ρ_ω (todo: find more optimal weights)



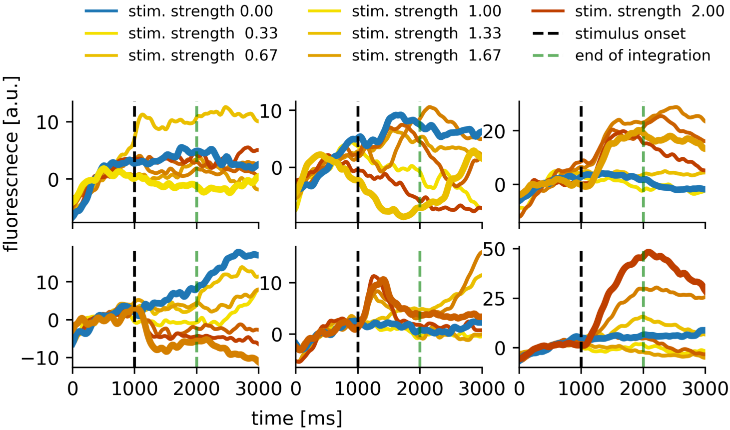
Tuning Curves



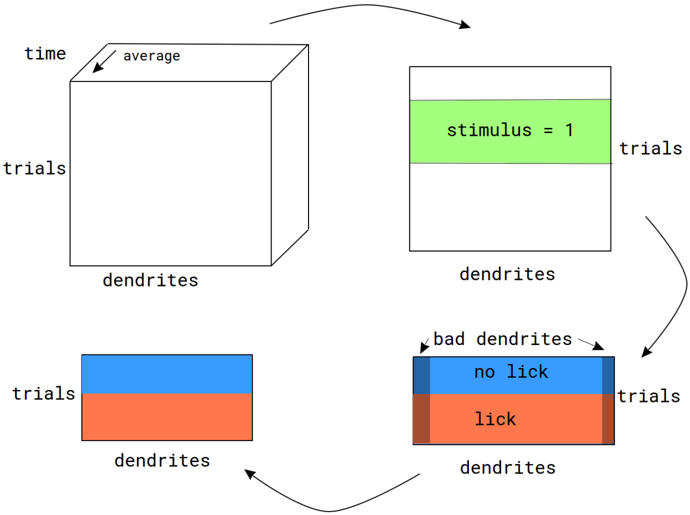
Tuning Curves



Tuning Curves



Bevahior Detection

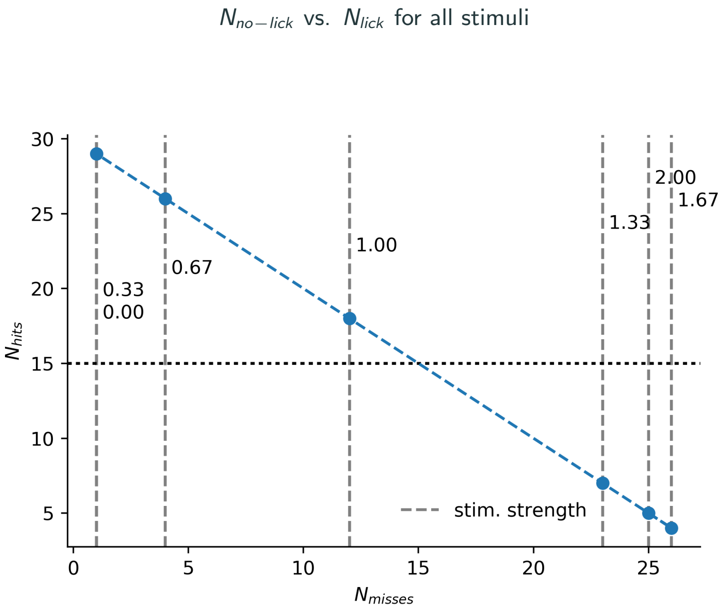


Behavioral Prediction - Most Accurate Dendrites

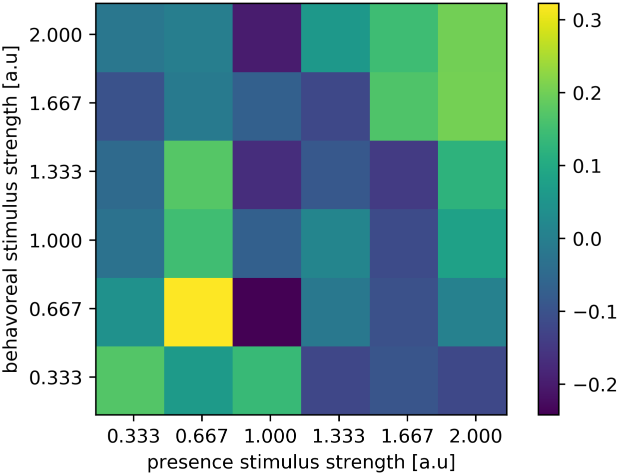
Different dataset - Stimulus strength 1 (near thereshold)

Dendrite #	μ_{acc}	σ_{acc}
88	0.80	0.08
57	0.80	0.08
92	0.78	0.06
56	0.76	0.10

Behavorial Predition - Imbalance



Rank Correlations of Behavoreal and Presence Dendrites

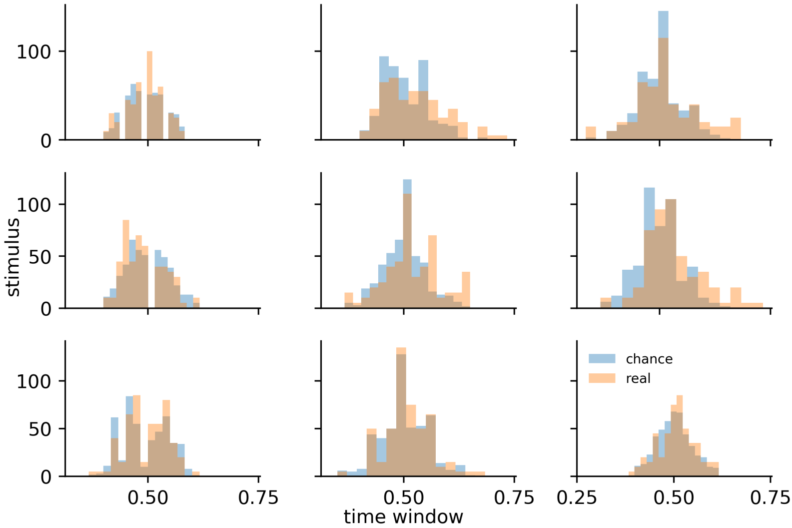


Timing - What if we use different Averaging Windows?

So far we have always averaged over the first second after the stimulus.

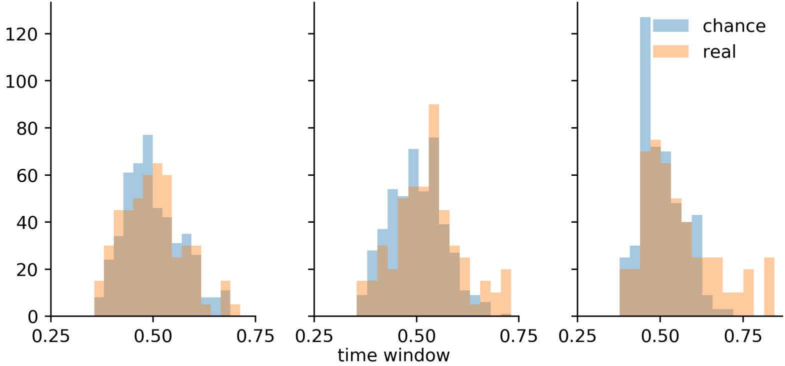
What if we average over the second before or the second second after the stimulus?

Statistical Significance - Stimulus Presence



Statistical Significance - Behavior

At threshold stimulus ≈ 1



Multivariate SMV Analysis

Presence Detection - Feature Selection

Since we have only few datapoints per condition (30-50 per class) in comparison to the number of features (100-150), we are prone to overfit to noise.

Presence Detection - Feature Selection

Since we have only few datapoints per condition (30-50 per class) in comparison to the number of features (100-150), we are prone to overfit to noise.

$\Rightarrow$ we have to do **feature selection**.

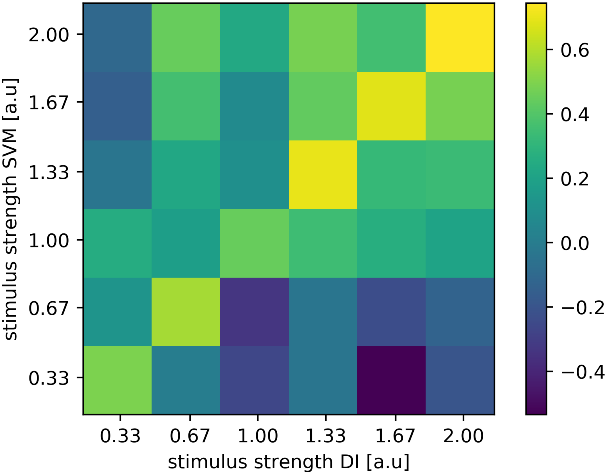
Presence Detection - Feature Selection

Since we have only few datapoints per condition (30-50 per class) in comparison to the number of features (100-150), we are prone to overfit to noise.

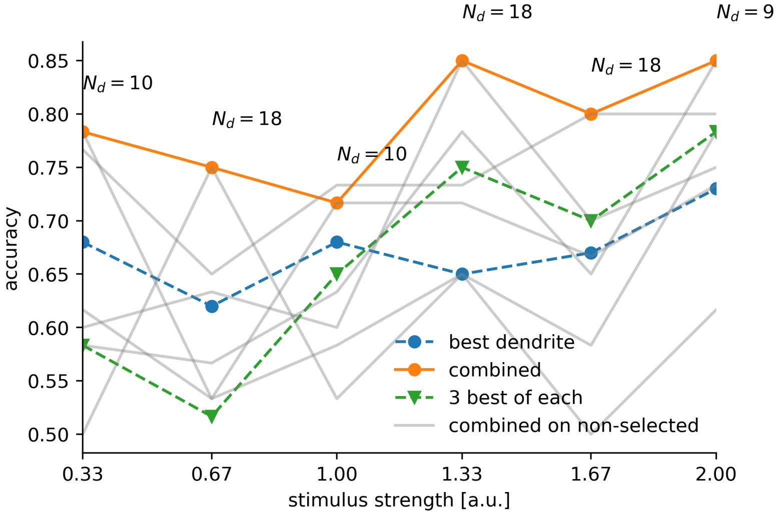
$\Rightarrow$ we have to do **feature selection**.

We can use the previously best discriminating dendrites (SVM/DI).

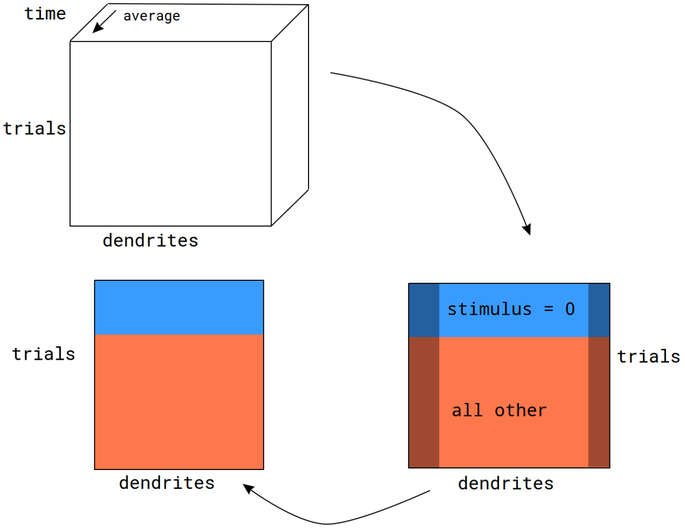
DI and SVM Dendrite Rank Order Correlation



SVM Performance on Combined Dendrites - Presence Detection



SVM Performance on Combined Dendrites - Global Presence



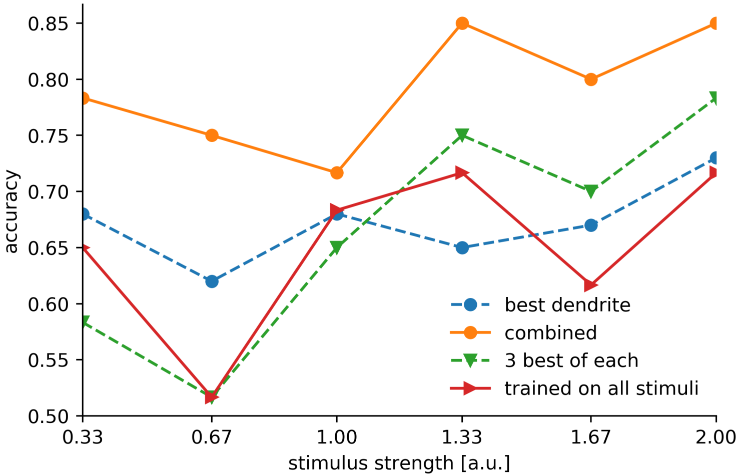
SVM Performance on Combined Dendrites - Global Presence

Performance on global presence detection:

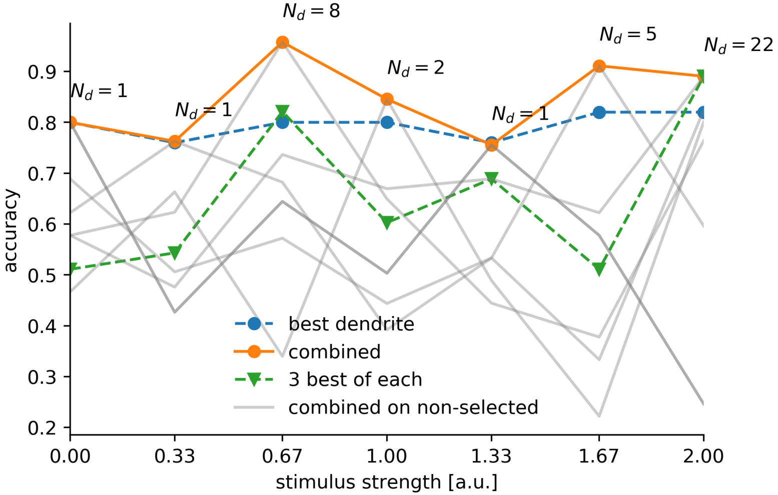
Mean: 0.68

Standard deviation: 0.08

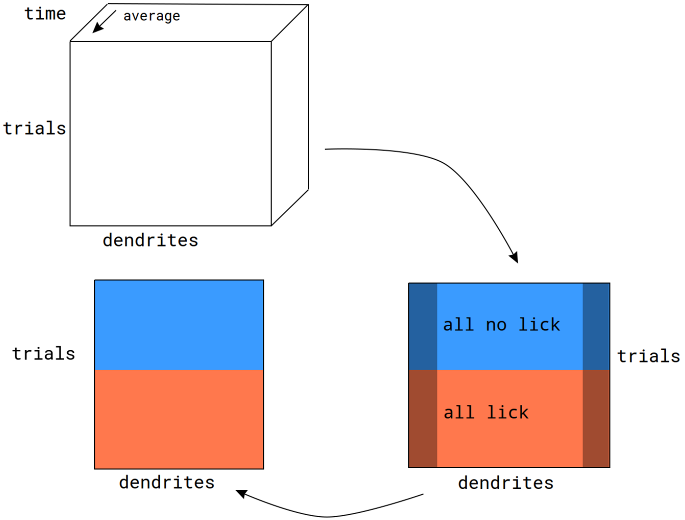
SVM Performance on Combined Dendrites - Presence Detection



SVM Performance on Combined Dendrites - Behavioral



SVM Performance on Combined Dendrites - Global Behavior



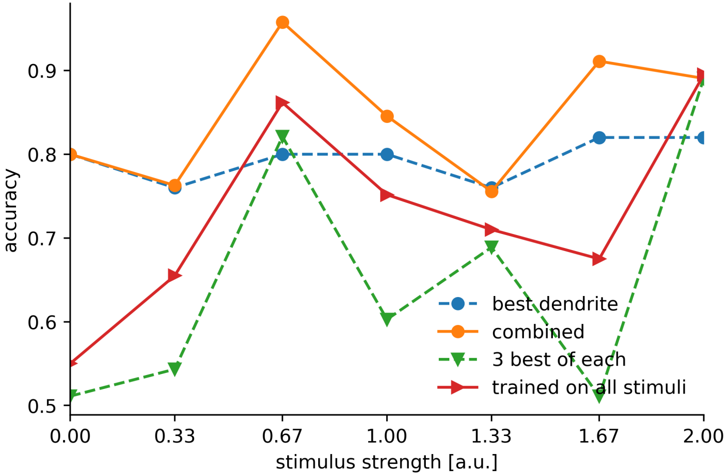
SVM Performance on Combined Dendrites - Global Behavior

Performance on global presence detection:

Mean: 0.71

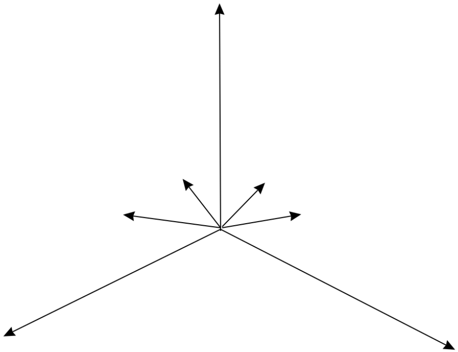
Standard deviation: 0.1

SVM Performance on Combined Dendrites - Behavioral

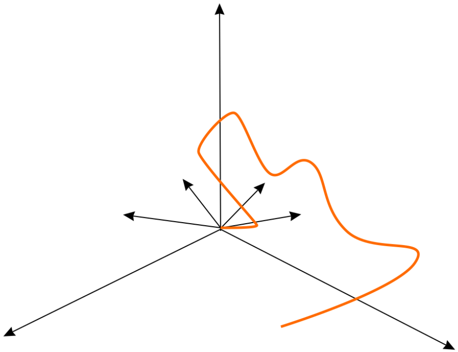


Population Coding

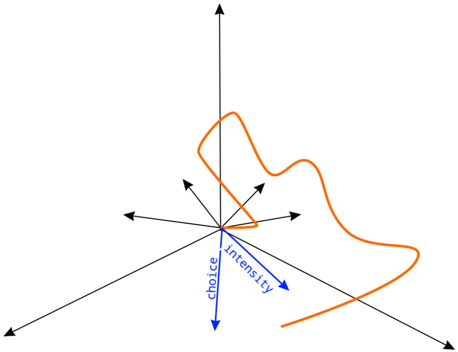
Population Coding - Idea (Mante, Sussillo 2013)



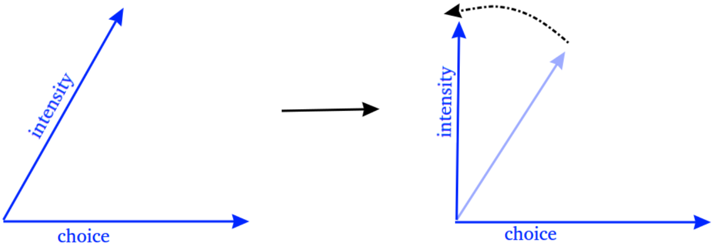
Population Coding - Idea (Mante, Sussillo 2013)



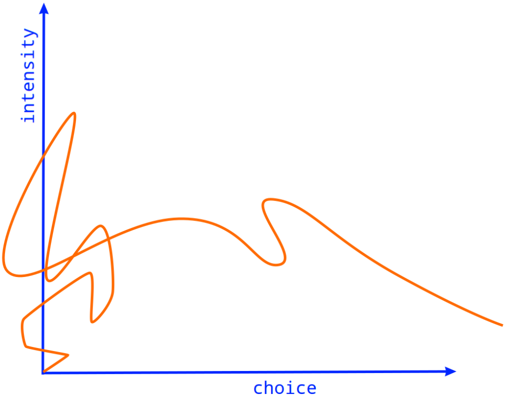
Population Coding - Idea (Mante, Sussillo 2013)



Population Coding - Idea (Mante, Sussillo 2013)

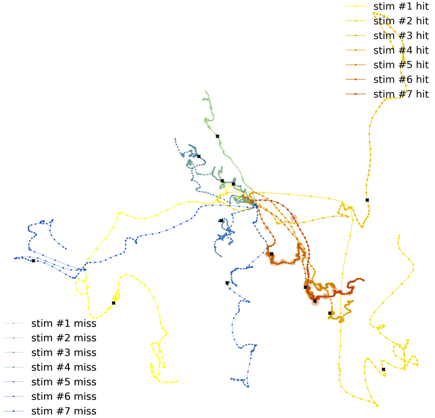


Population Coding - Idea (Mante, Sussillo 2013)



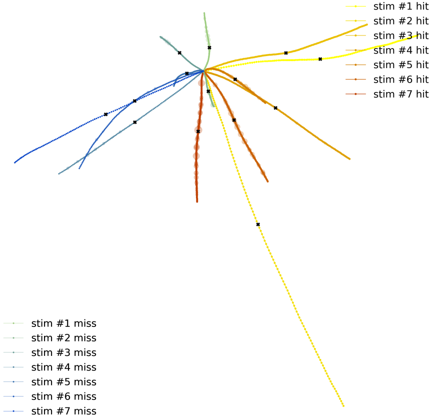
Population Response in Task Variable Space

Solve Problem with combined SVM first



Population Response in Task Variable Space

Solve Problem with combined SVM first



- 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

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 ρ greatly. We would like to be able to **weight our $\mathbf{D_i}$**'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 normaized 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 normaized 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 ν.

Population Response in Task Variable Space

We define

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

and estimate for each dendrite i and timepoint t

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

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:

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

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 β 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

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

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 $\boldsymbol{\beta}^{\boldsymbol{\nu},\boldsymbol{t}}$ as well, which yields the denoised regression vectors

$$\boldsymbol{\beta}^{\boldsymbol{\nu},\boldsymbol{t}}_{\boldsymbol{pca}}$$

Population Response in Task Variable Space

Since the $\beta_{\boldsymbol{pca}}^{\boldsymbol{\nu},\boldsymbol{t}}$ are time variing 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 $\beta^{\nu,t}_{pca}$ are time variing 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:

$$\beta^{\nu}_{max} = \beta^{\nu,t^{\nu}_{max}}_{pca}$$

$$t^{\nu}_{max} = argmax_t ||\beta^{\nu,t}_{pca}||$$

Population Response in Task Variable Space

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

$$\mathbf{B}^{\mathbf{max}} = \begin{bmatrix} \beta^{\mathbf{1}}_{\mathbf{max}} & \beta^{\mathbf{2}}_{\mathbf{max}} \end{bmatrix} = \mathbf{QR}$$

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