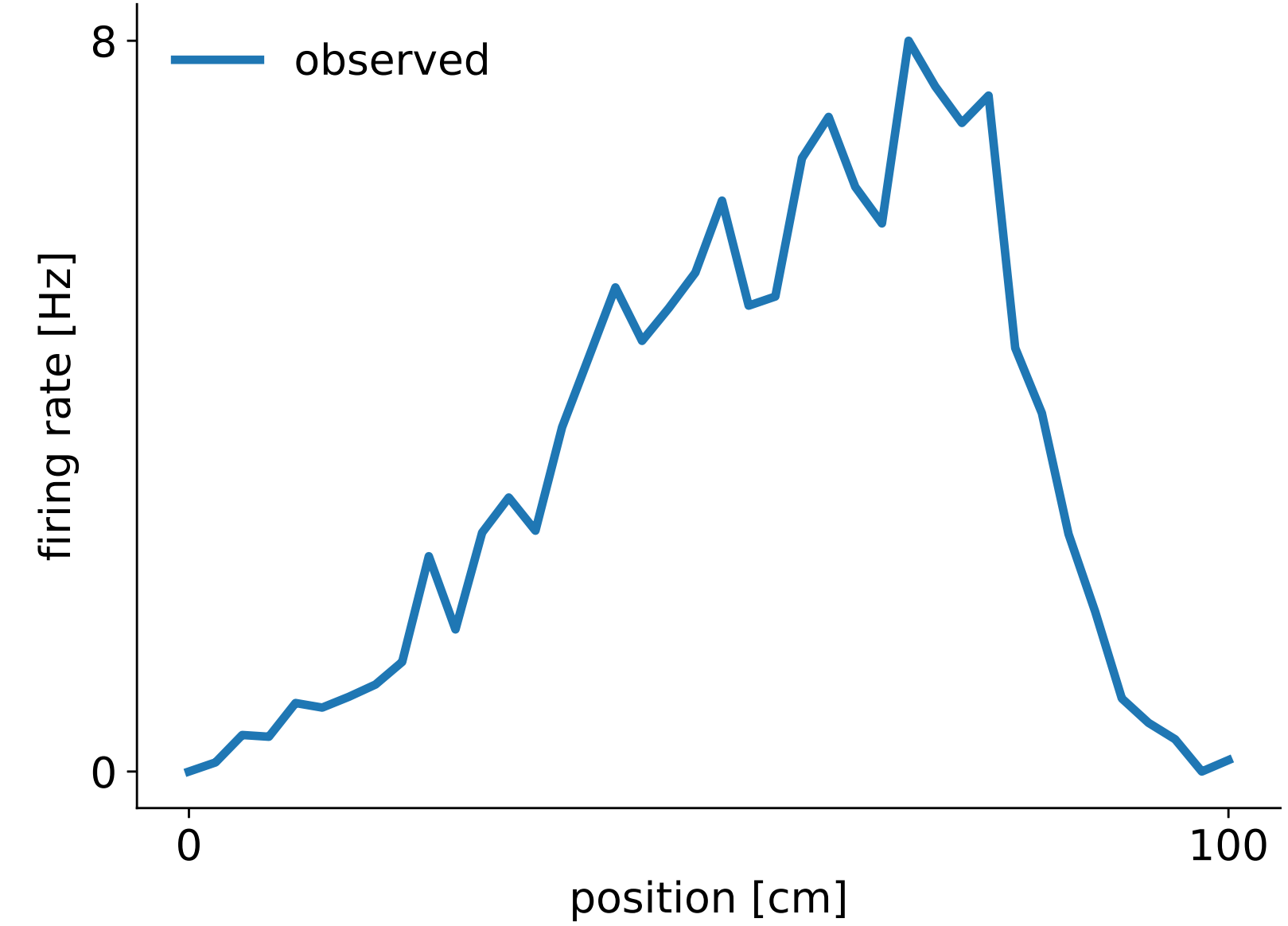
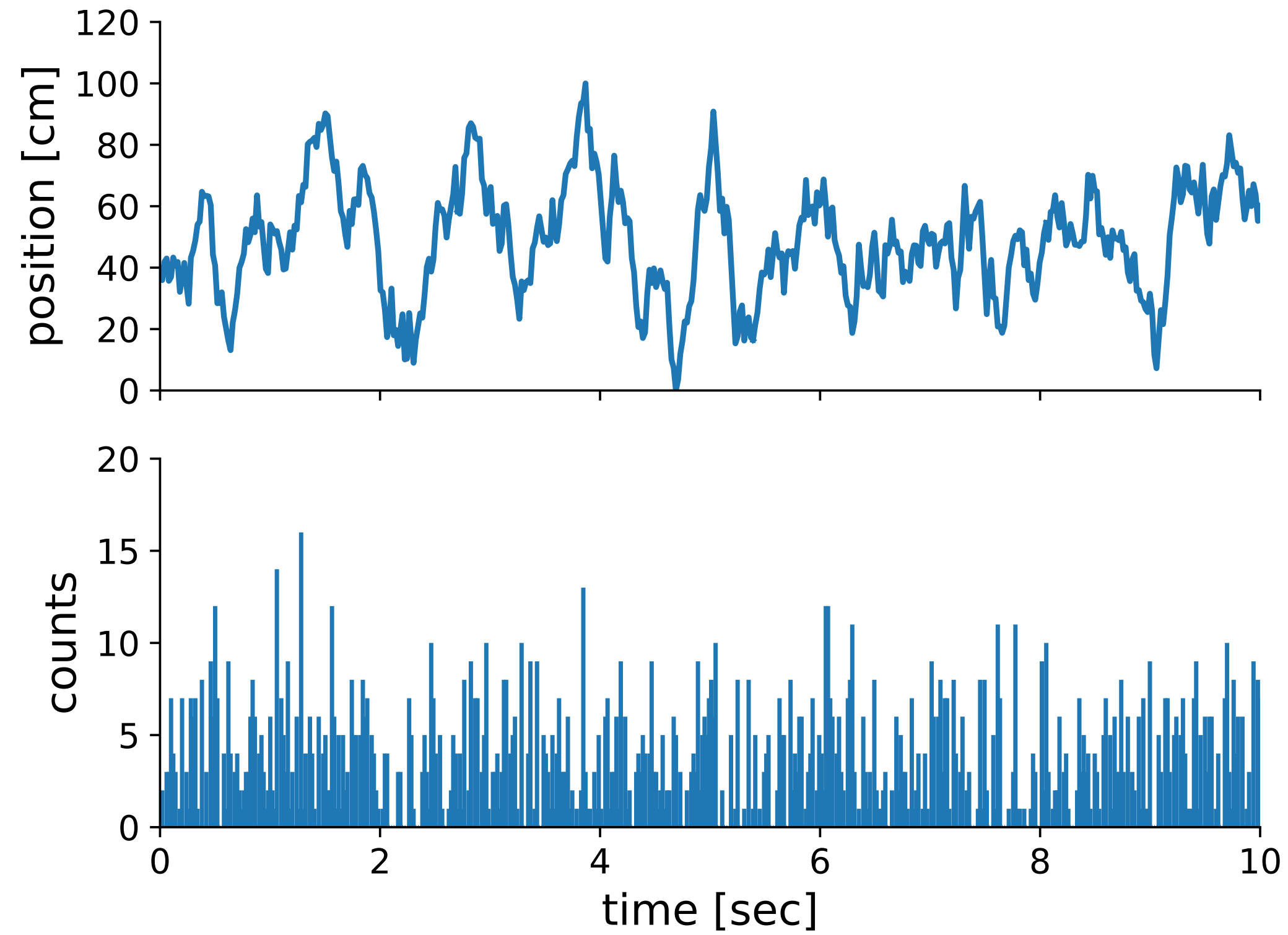


Model Selection and Cross-Validation

Avoid over-fitting by cross-validating your hyper-parameters.

Problem 1

Fit a GLM to capture tuning to position

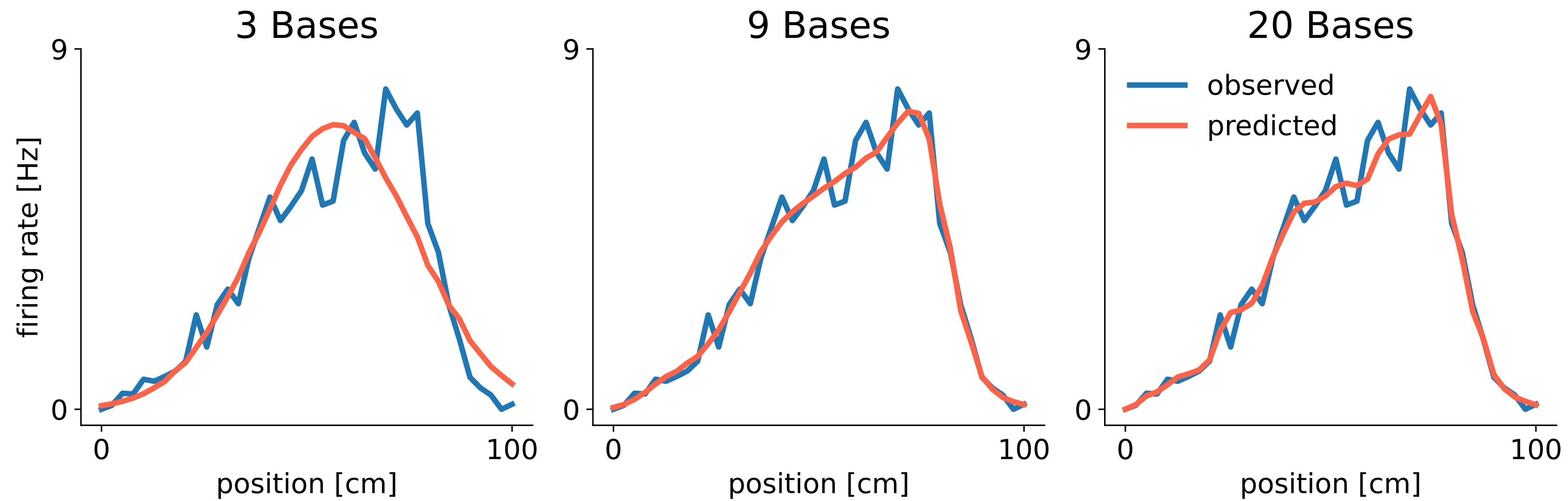


Aim: Parametrize the non-linear map with basis

Problem 1

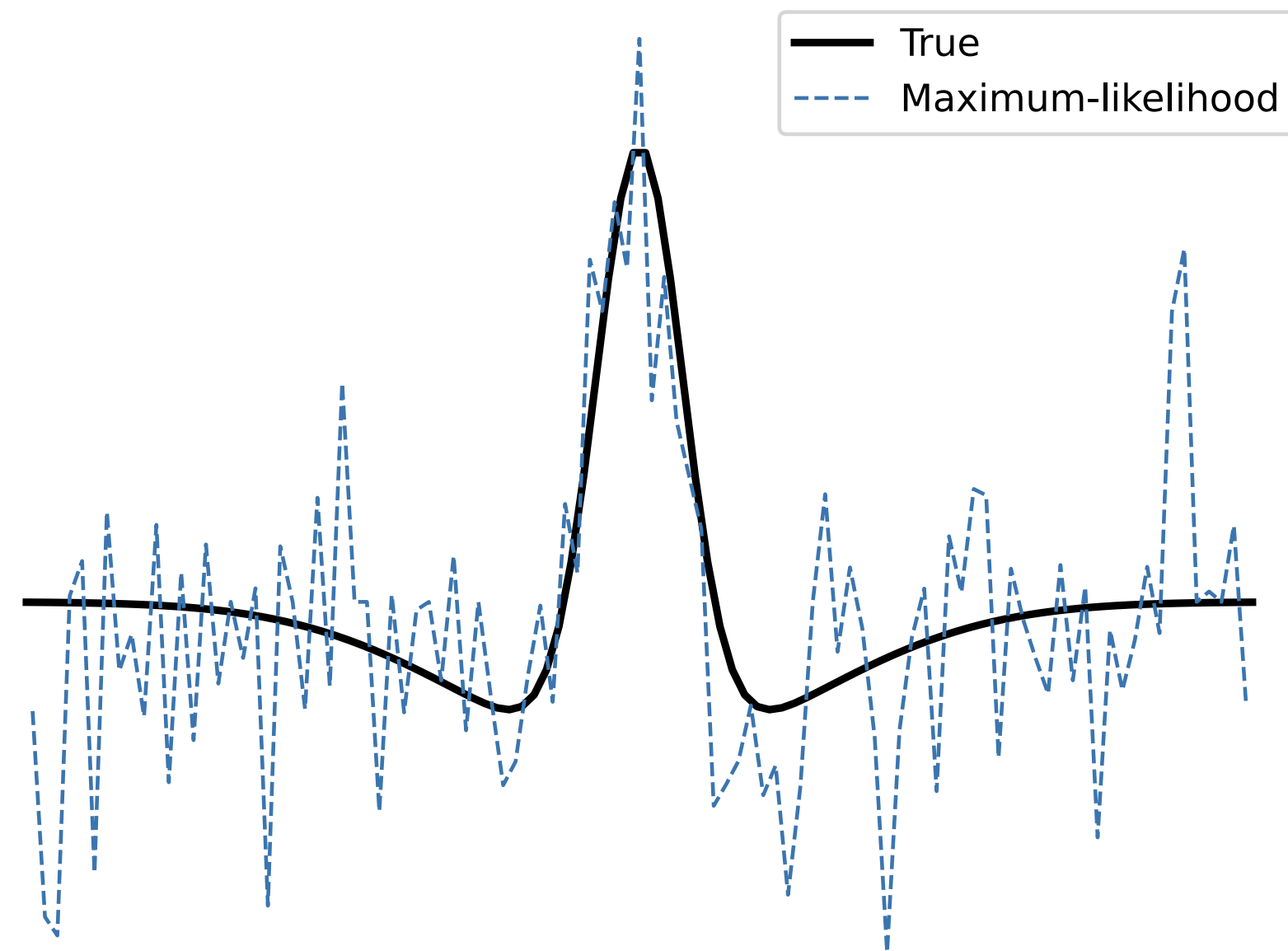
Question:

- What is an optimal number of basis functions?
- How do we decide?



Problem 2

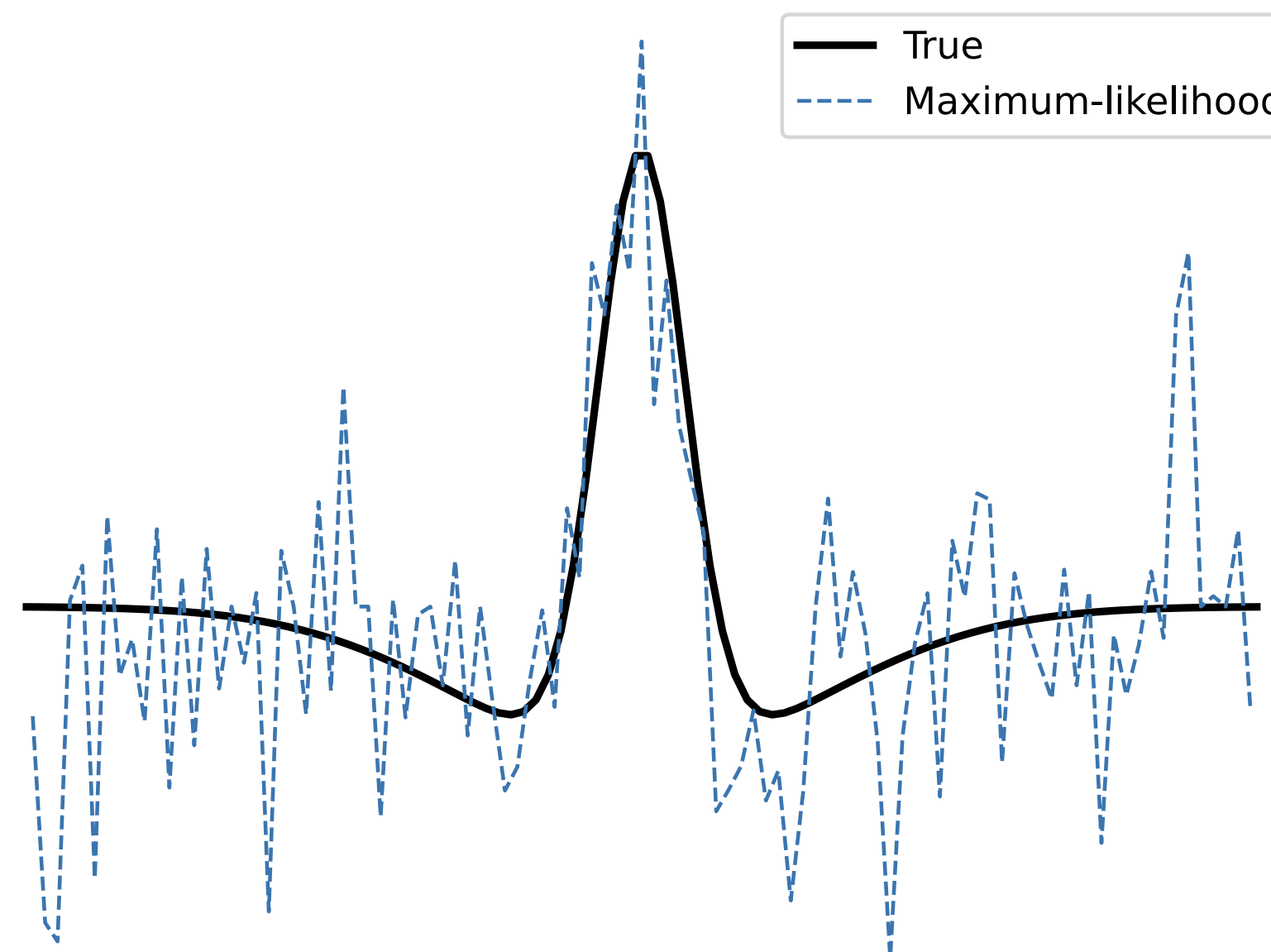
Maximum-Likelihood: $\max_{\mathbf{w}} \log p(\text{counts} | \mathbf{X}, \mathbf{w})$



Problem 2

Maximum-Likelihood: $\max_{\mathbf{w}} \log p(\text{counts} | \mathbf{X}, \mathbf{w})$

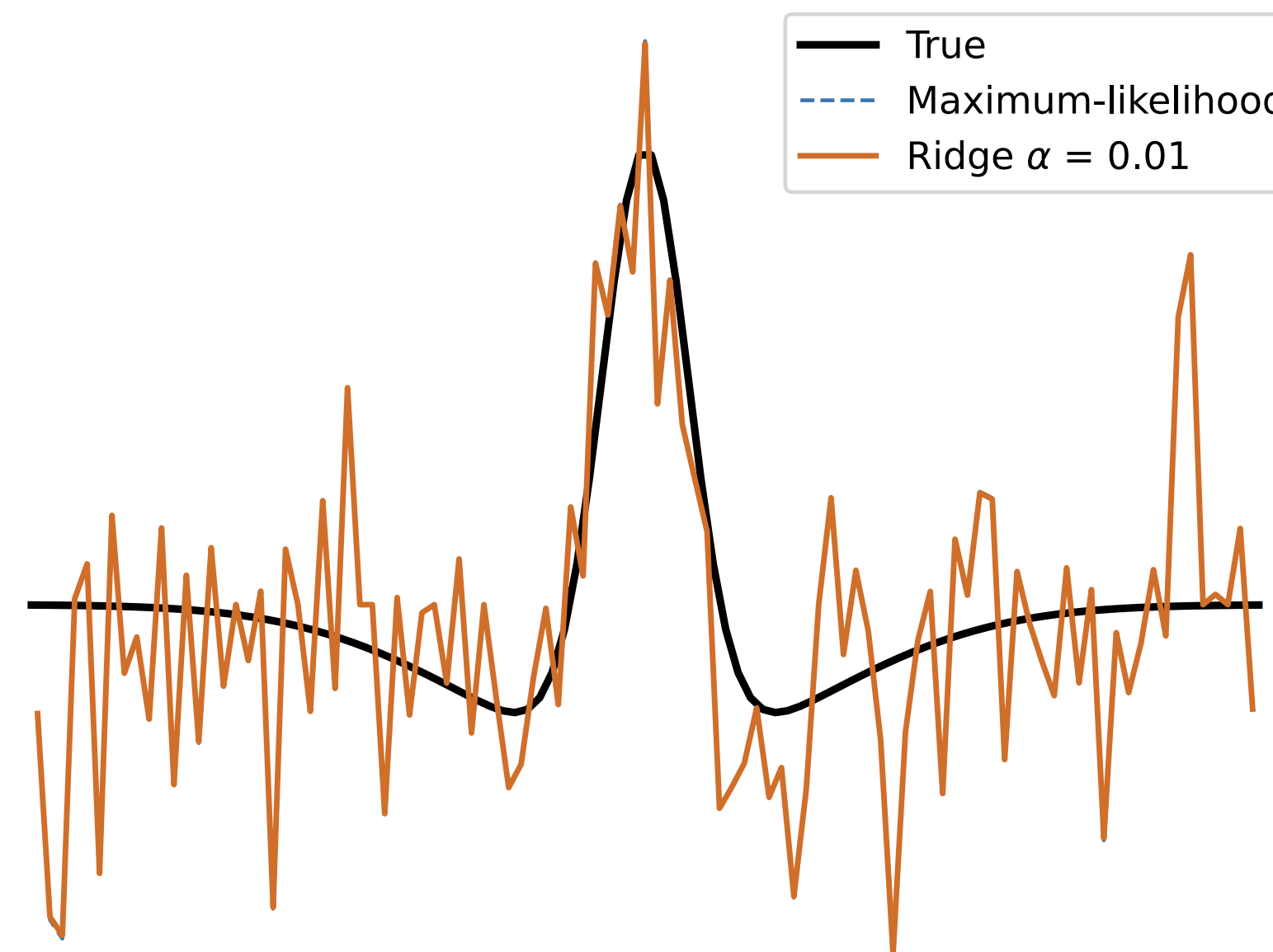
Ridge (L2): $\max_{\mathbf{w}} \log p(\text{counts} | \mathbf{X}, \mathbf{w}) - \alpha (w_1^2 + \dots + w_n^2)$



Problem 2

Maximum-Likelihood: $\max_{\mathbf{w}} \log p(\text{counts} | \mathbf{X}, \mathbf{w})$

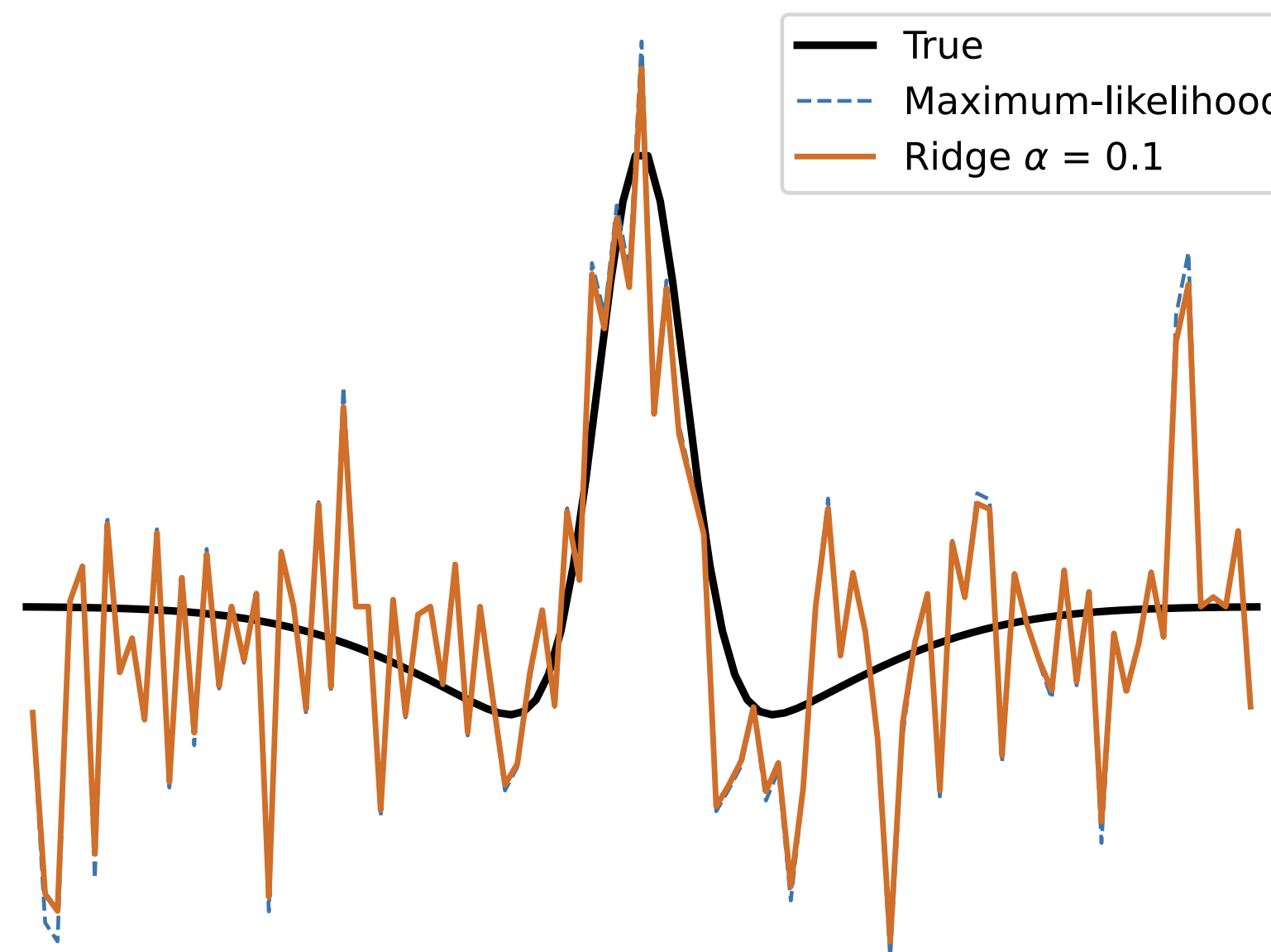
Ridge (L2): $\max_{\mathbf{w}} \log p(\text{counts} | \mathbf{X}, \mathbf{w}) - \alpha (w_1^2 + \dots + w_n^2)$



Problem 2

Maximum-Likelihood: $\max_{\mathbf{w}} \log p(\text{counts} \mid \mathbf{X}, \mathbf{w})$

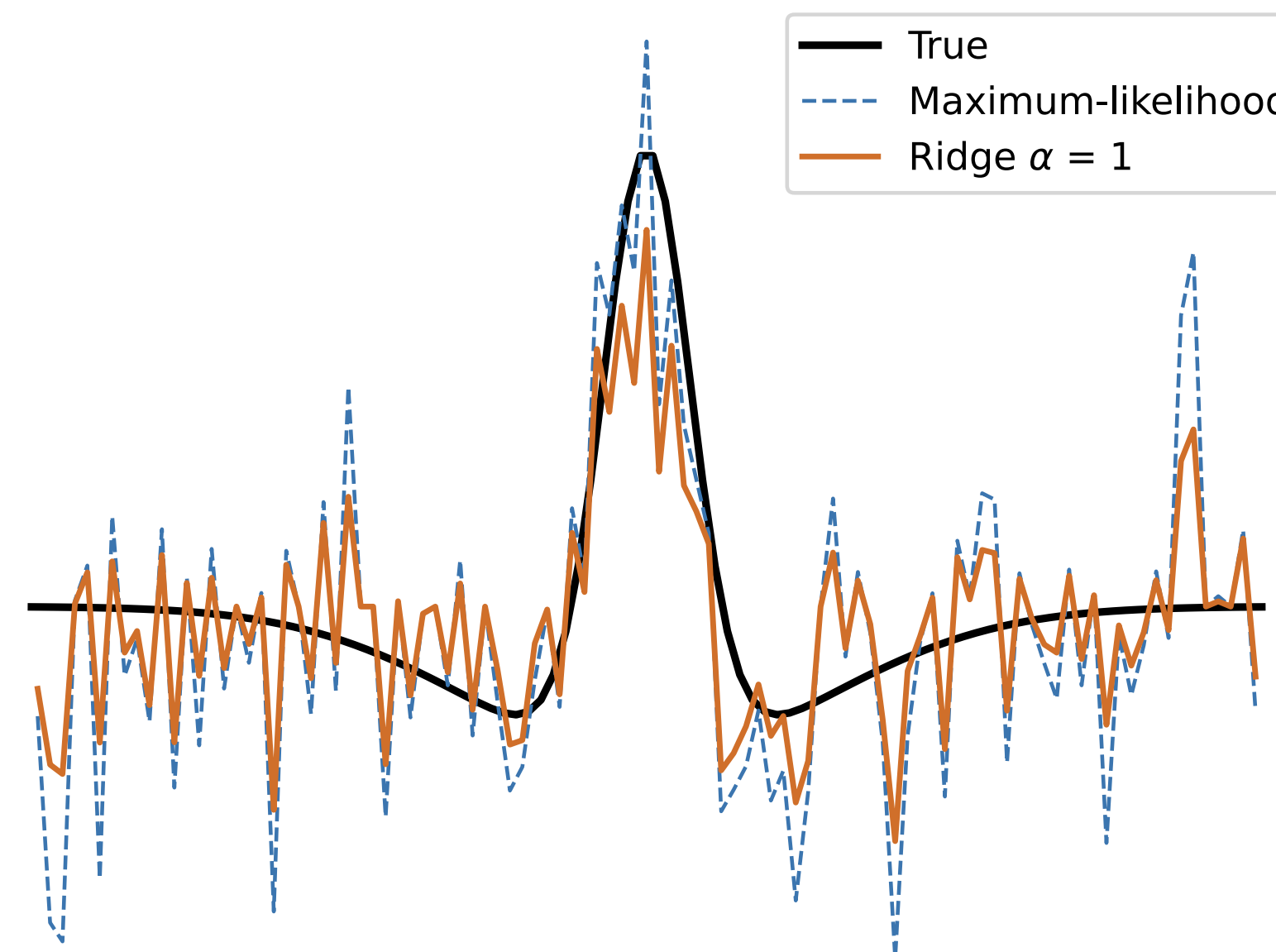
Ridge (L2): $\max_{\mathbf{w}} \log p(\text{counts} \mid \mathbf{X}, \mathbf{w}) - \alpha (w_1^2 + \dots + w_n^2)$



Problem 2

Maximum-Likelihood: $\max_{\mathbf{w}} \log p(\text{counts} | \mathbf{X}, \mathbf{w})$

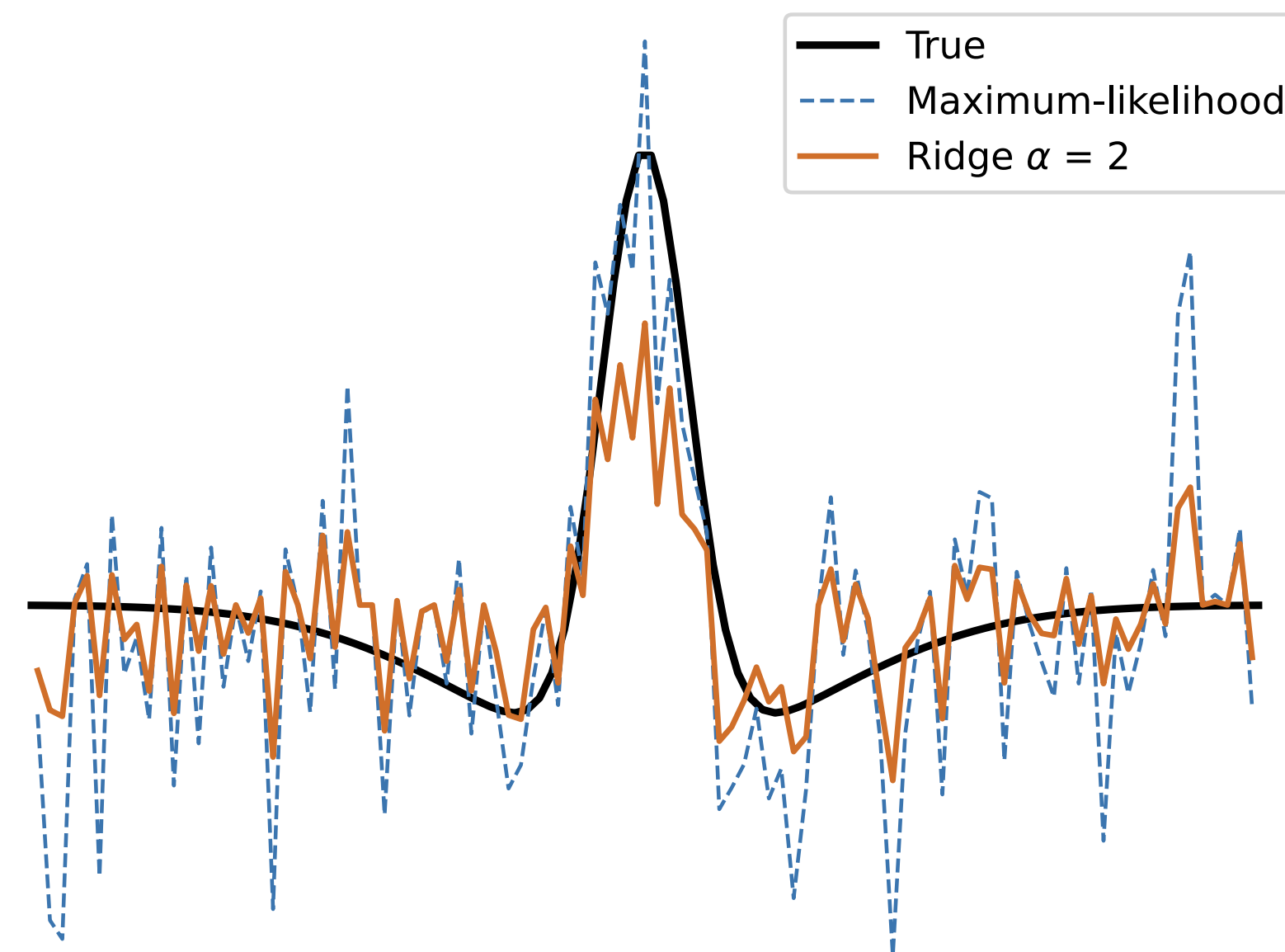
Ridge (L2): $\max_{\mathbf{w}} \log p(\text{counts} | \mathbf{X}, \mathbf{w}) - \alpha (w_1^2 + \dots + w_n^2)$



Problem 2

Maximum-Likelihood: $\max_{\mathbf{w}} \log p(\text{counts} | \mathbf{X}, \mathbf{w})$

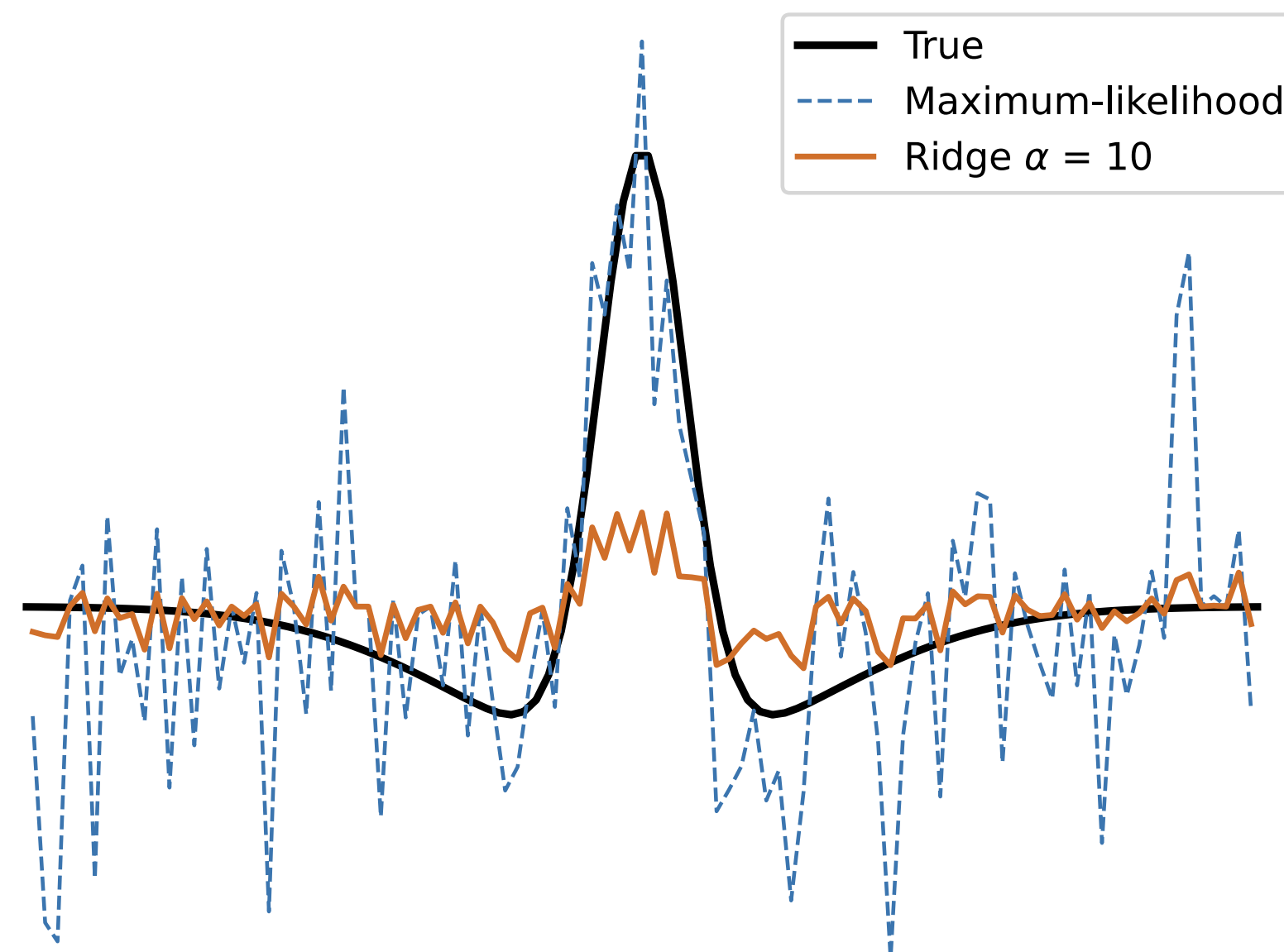
Ridge (L2): $\max_{\mathbf{w}} \log p(\text{counts} | \mathbf{X}, \mathbf{w}) - \alpha (w_1^2 + \dots + w_n^2)$



Problem 2

Maximum-Likelihood: $\max_{\mathbf{w}} \log p(\text{counts} | \mathbf{X}, \mathbf{w})$

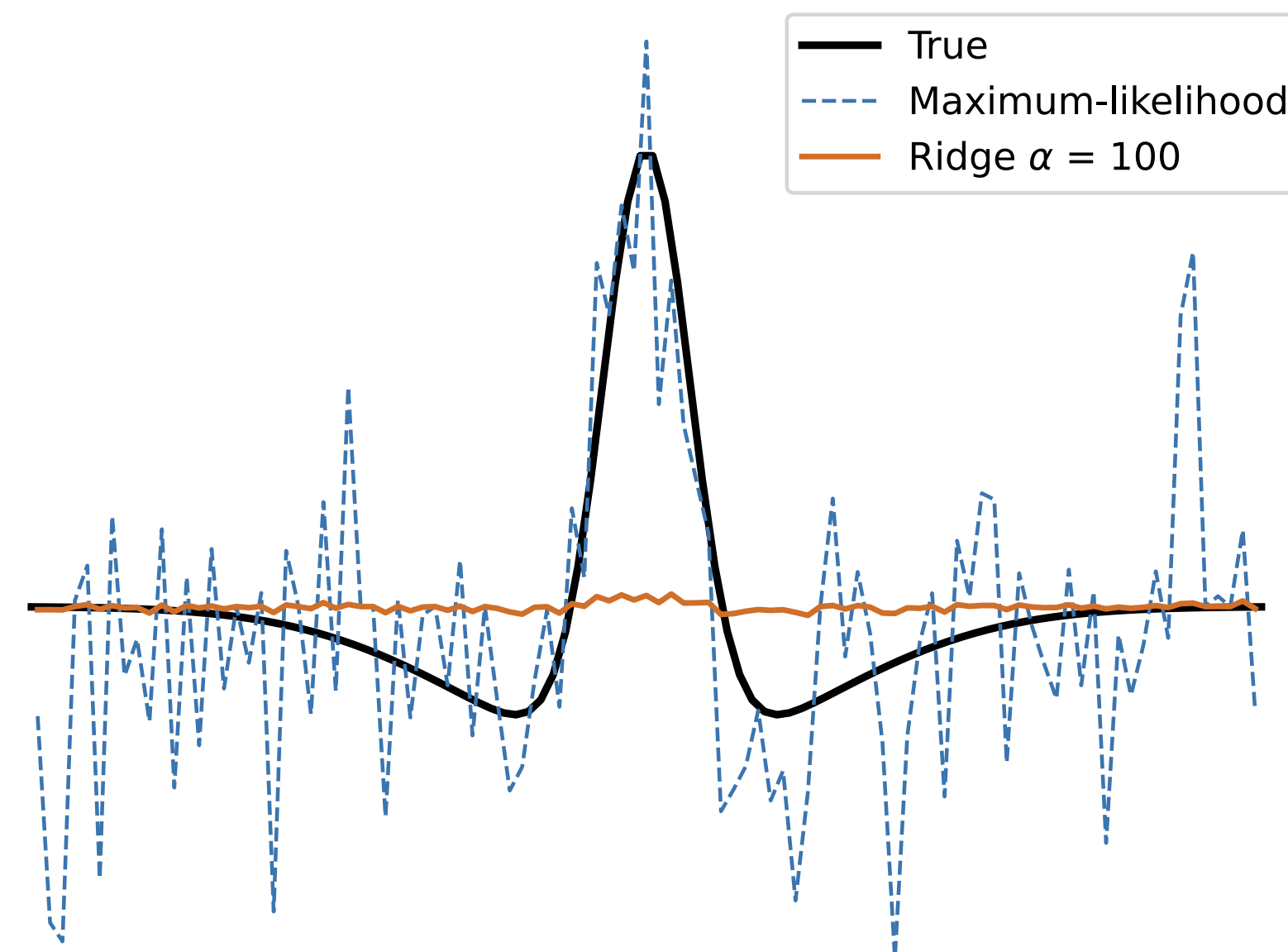
Ridge (L2): $\max_{\mathbf{w}} \log p(\text{counts} | \mathbf{X}, \mathbf{w}) - \alpha (w_1^2 + \dots + w_n^2)$



Problem 2

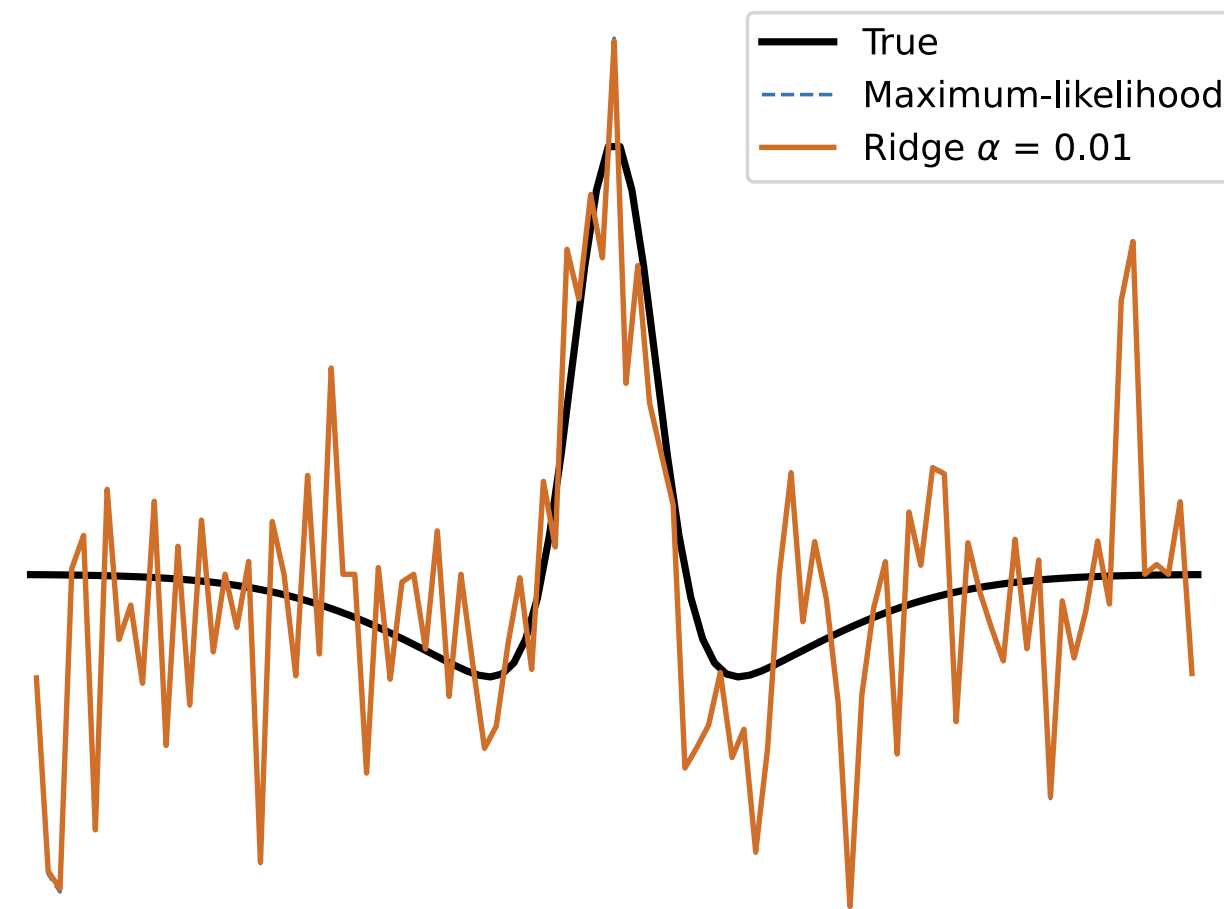
Maximum-Likelihood: $\max_{\mathbf{w}} \log p(\text{counts} \mid \mathbf{X}, \mathbf{w})$

Ridge (L2): $\max_{\mathbf{w}} \log p(\text{counts} \mid \mathbf{X}, \mathbf{w}) - \alpha (w_1^2 + \dots + w_n^2)$



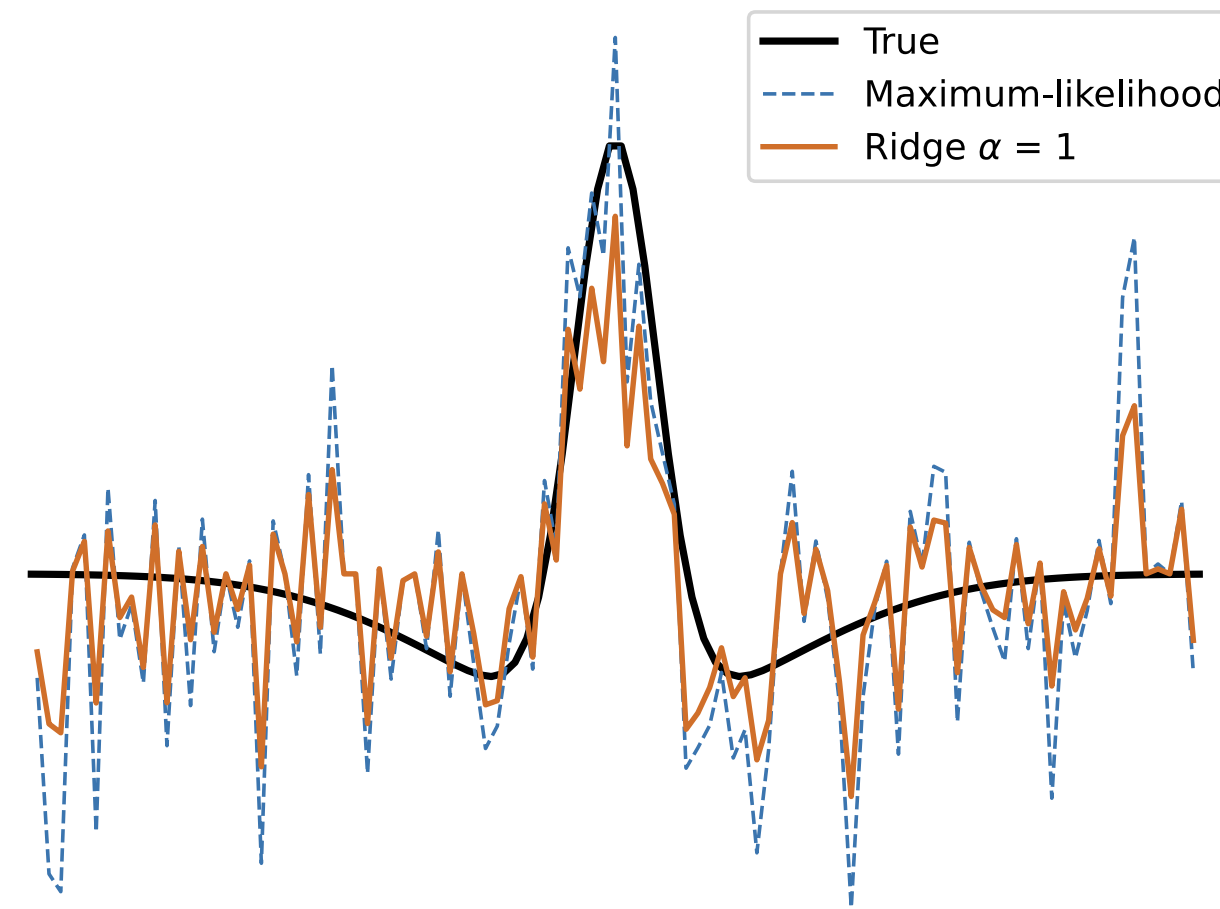
Problem 2

$\alpha = 0.01$

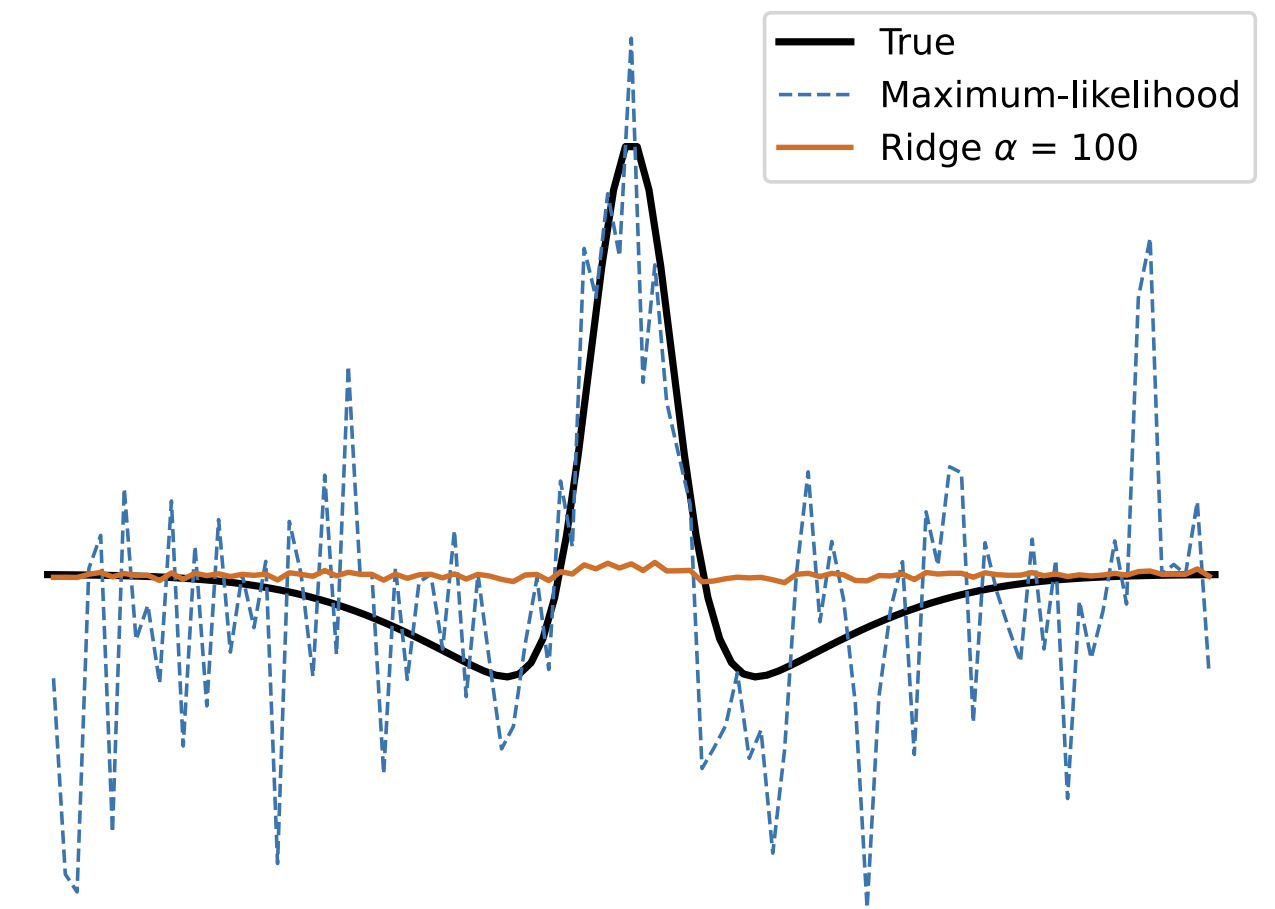


Overfitting
(too much variance)

$\alpha = 1$



$\alpha = 100$



Underfitting
(too much bias towards 0)

How do you select α ?

Problem setting

More generally, we want to answer questions like:

- What features should I include?
- How many bases?
- Which kind of bases?
- ...

Cross-Validation

- **Basic idea:** learn the weights on a subset of your data, test the model on another subset.

Cross-Validation

- **Basic idea:** learn the weights on a subset of your data, test the model on another subset.
- Many different approaches:
 - Leave-One-Out
 - Leave-P-Out
 - Shuffle Split
 - K-Fold
 - Stratified K-Fold
 - ...

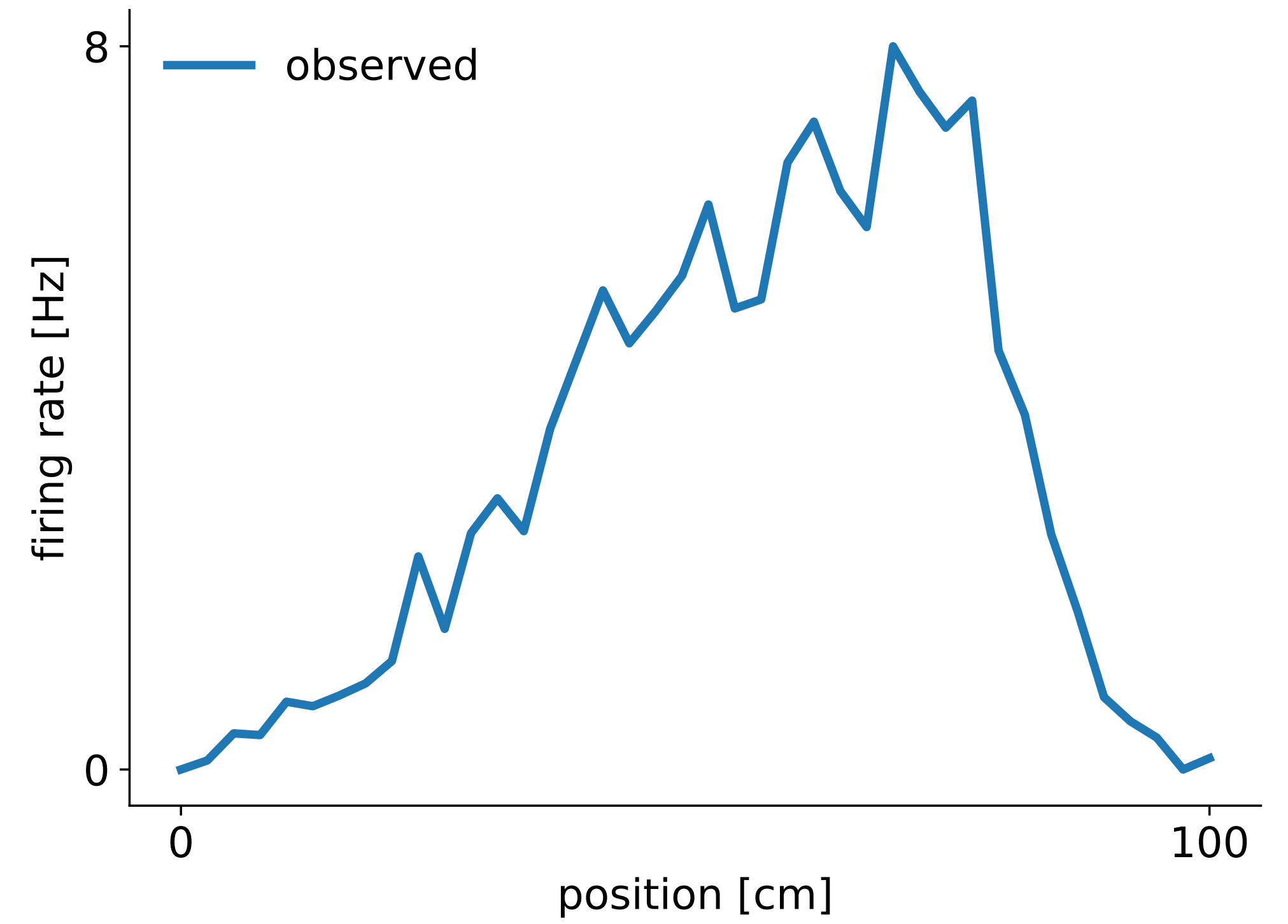
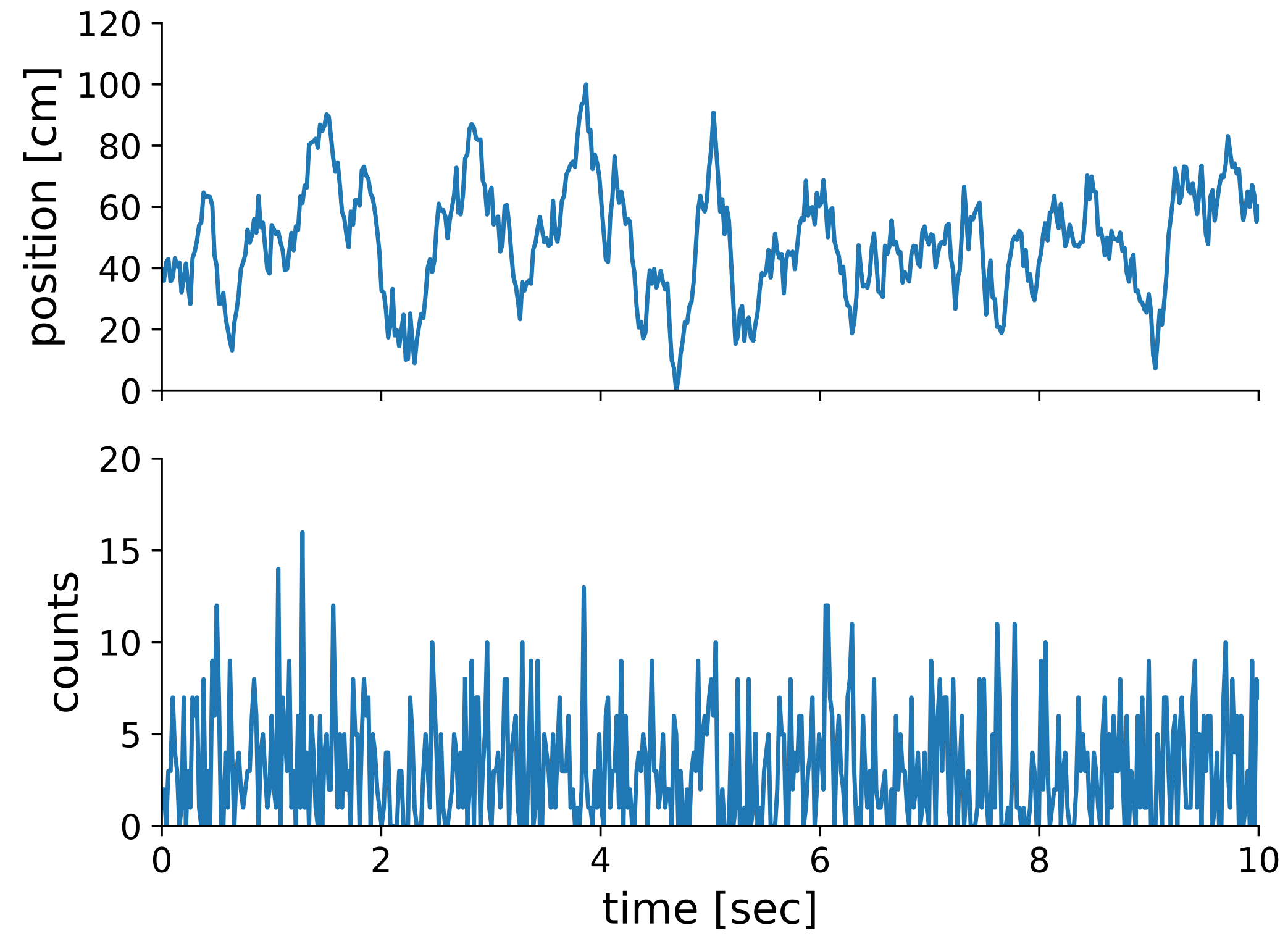
Cross-Validation

- **Basic idea:** learn the weights on a subset of your data, test the model on another subset.
- Many different approaches:
 - Leave-One-Out
 - Leave-P-Out
 - Shuffle Split
 - K-Fold
 - Stratified K-Fold
 - ...
- All these approaches (and more) implemented in scikit-learn.

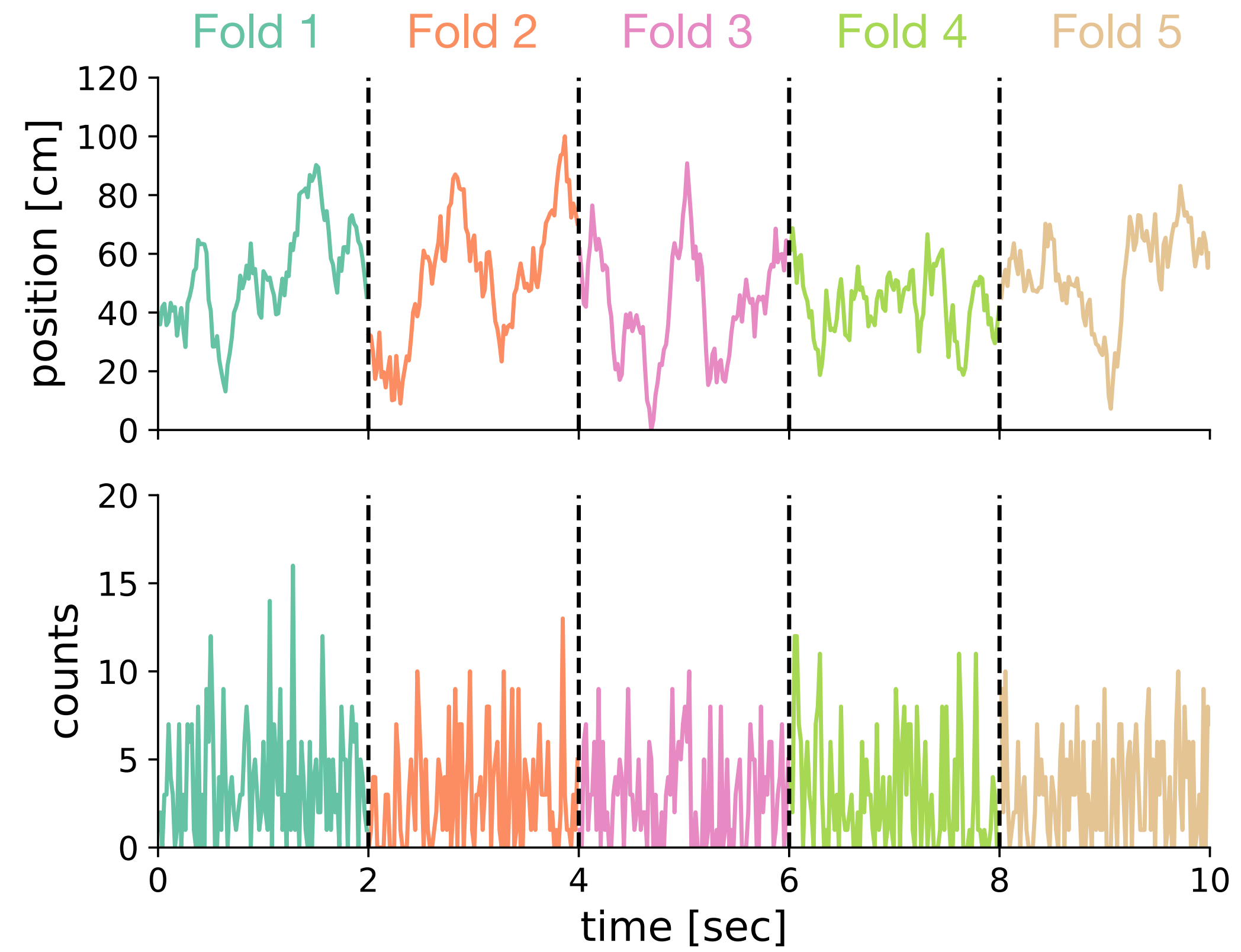
Cross-Validation

- **Basic idea:** learn the weights on a subset of your data, test the model on another subset.
- Many different approaches:
 - Leave-One-Out
 - Leave-P-Out
 - Shuffle Split
 - K-Fold
 - Stratified K-Fold
 - ...
- All these approaches (and more) implemented in scikit-learn.
- NeMoS models are compatible (more on this in the next tutorial).

Example

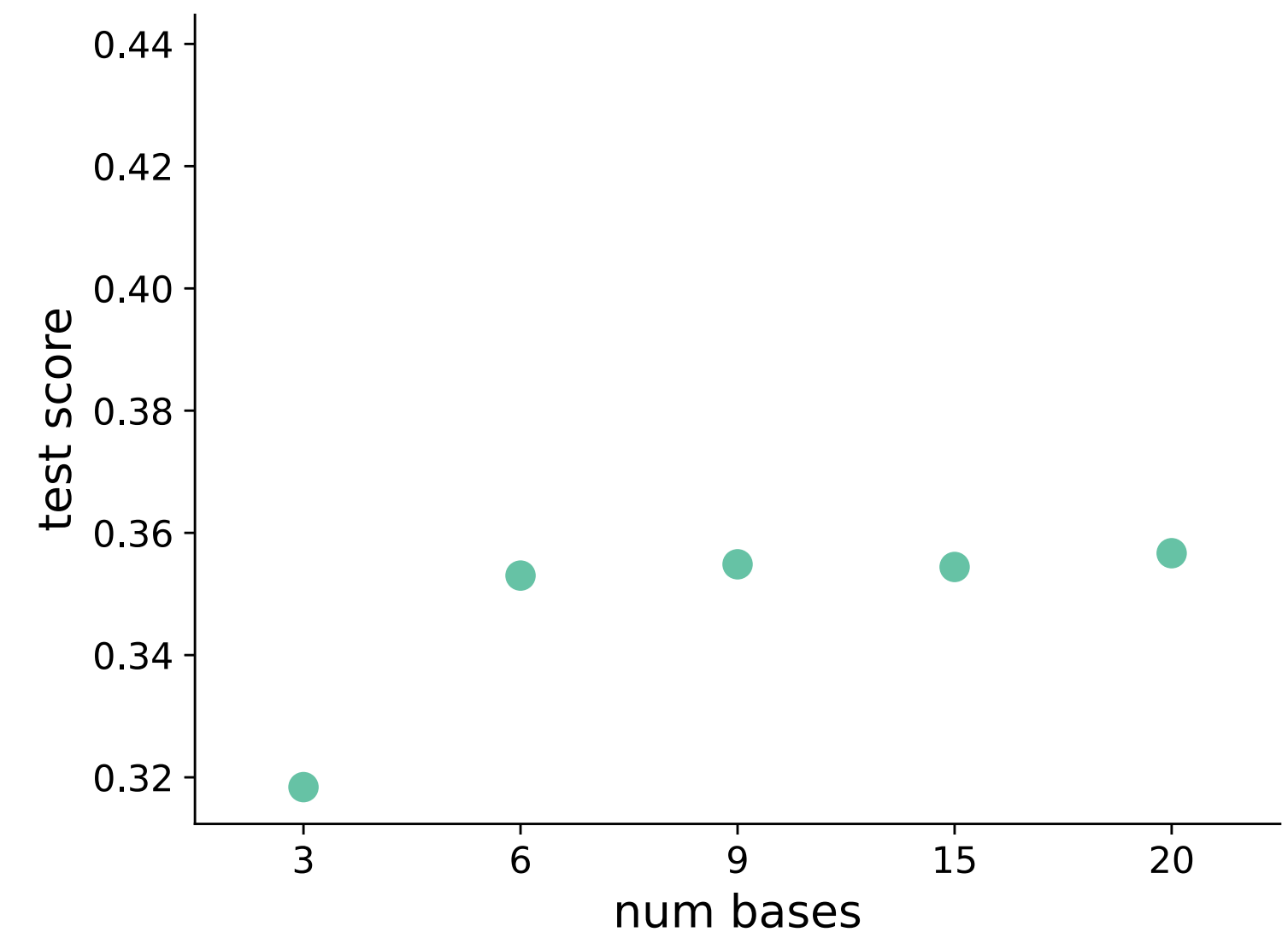
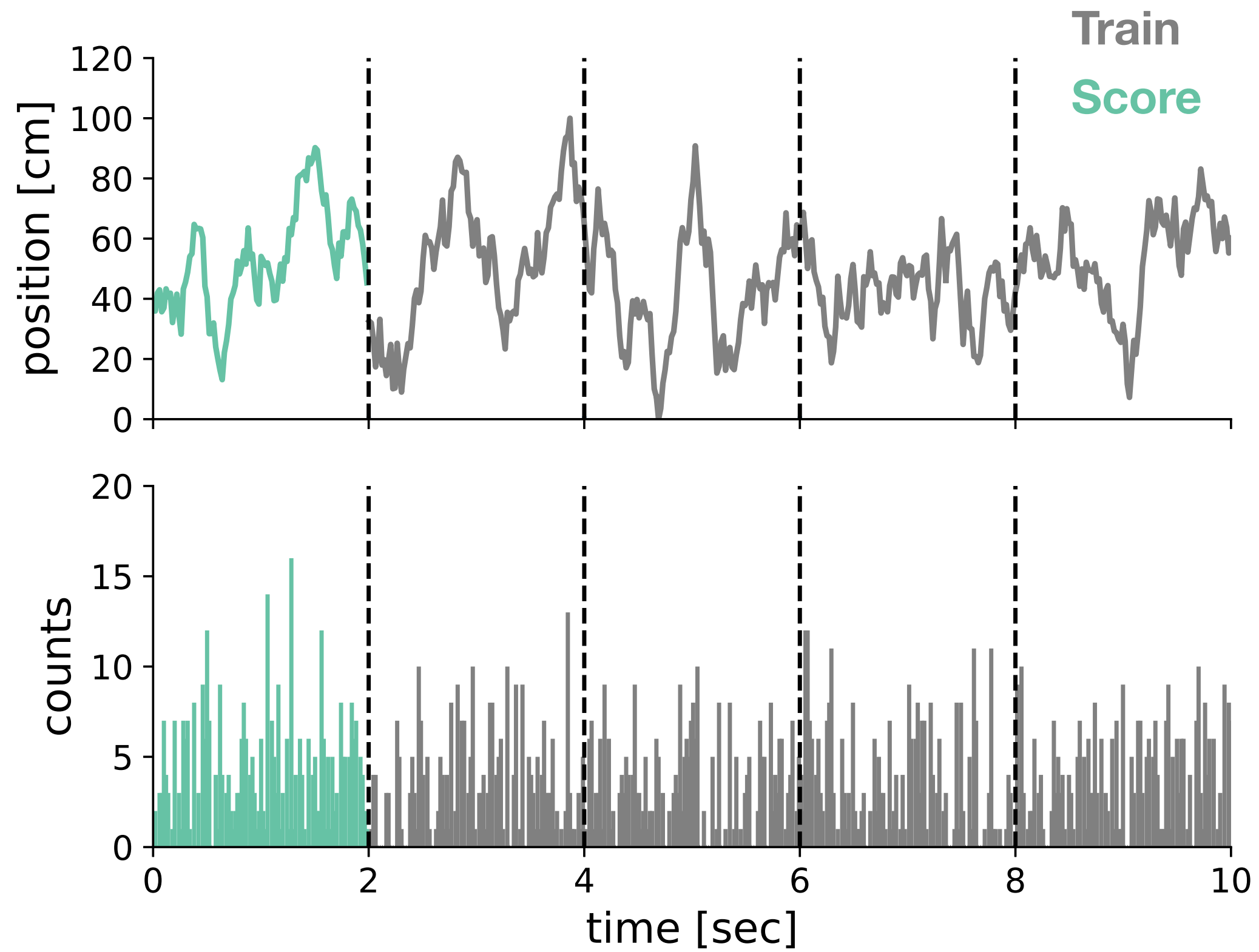


Example: K-Fold



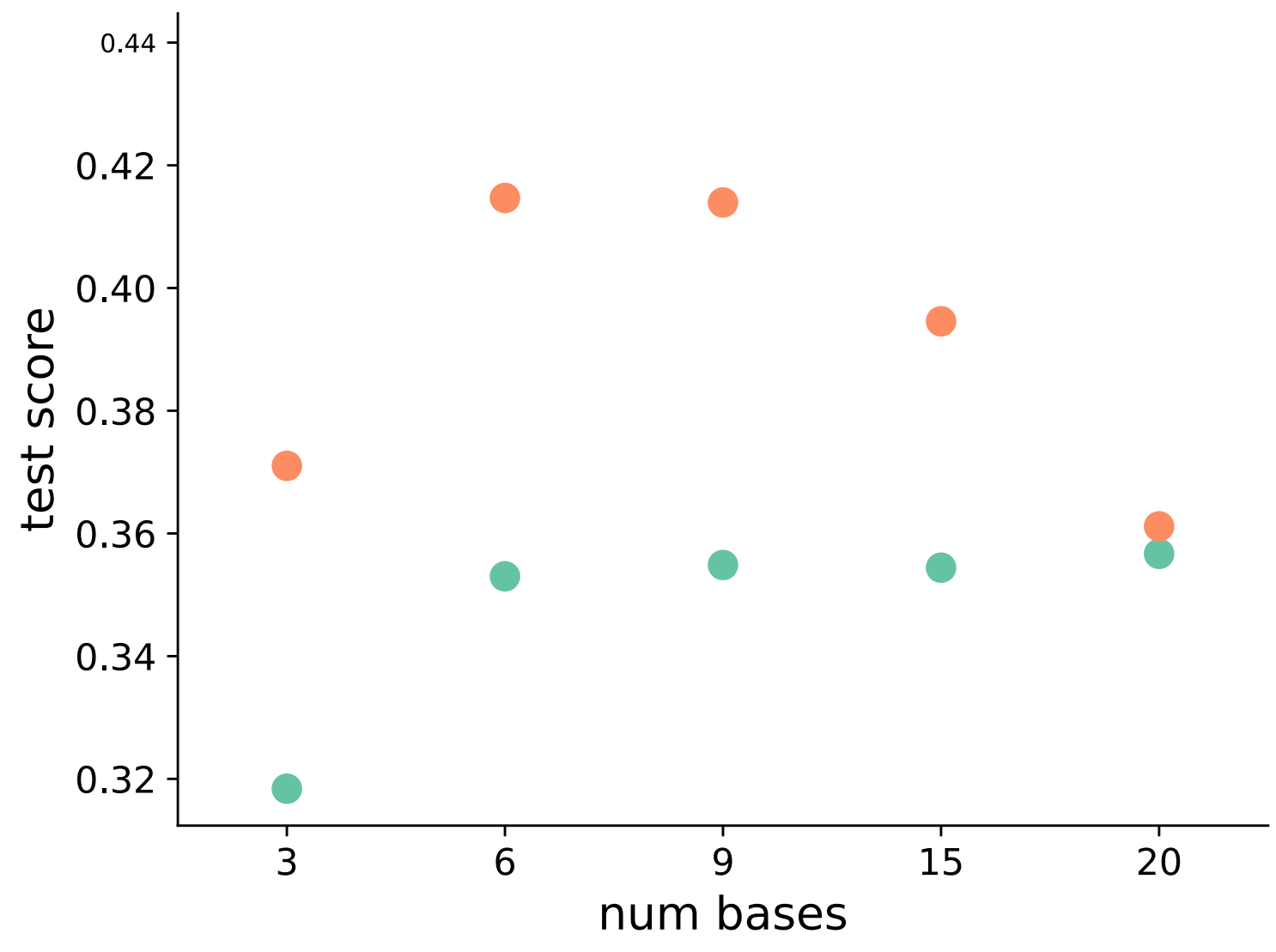
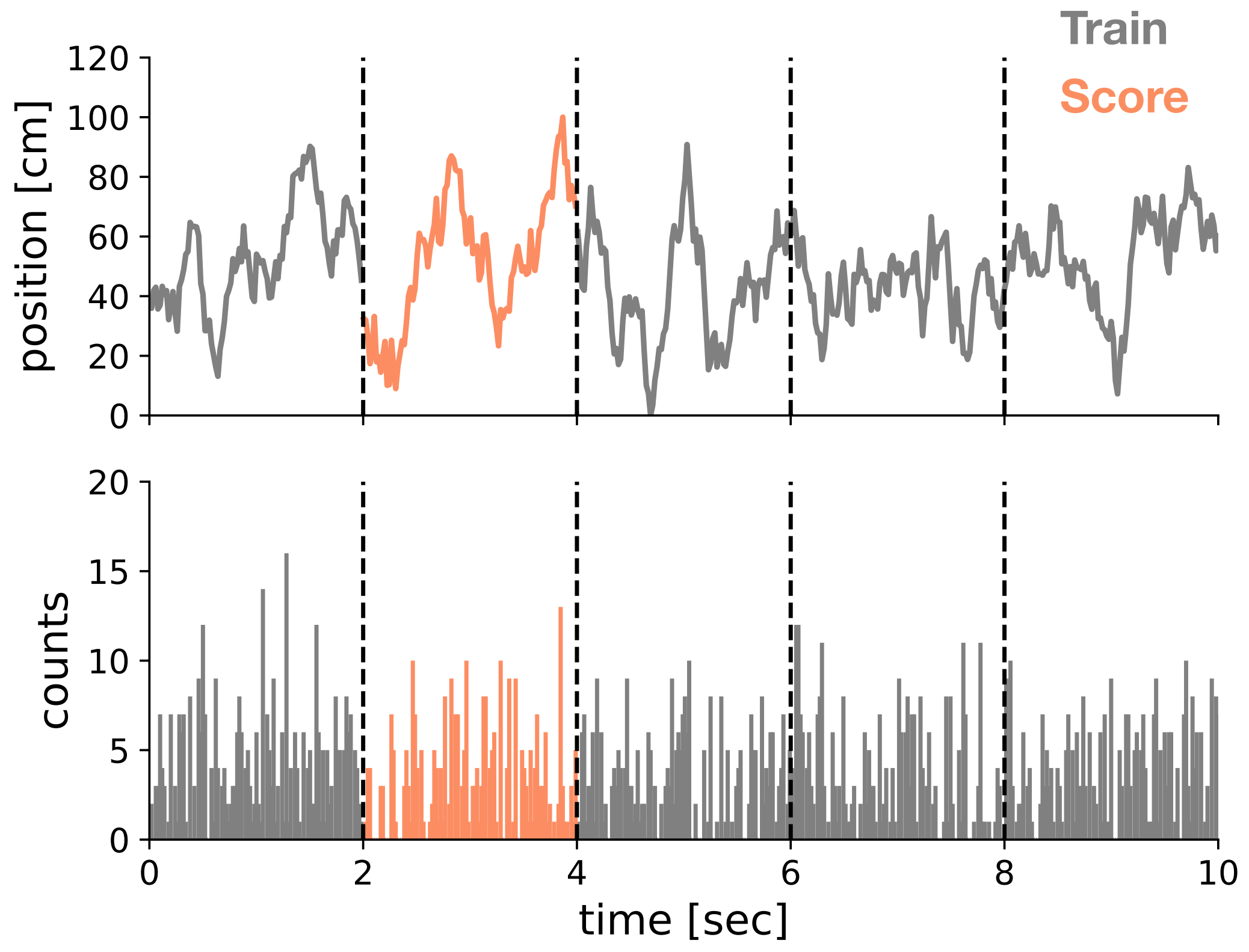
Example: K-Fold

Split 1



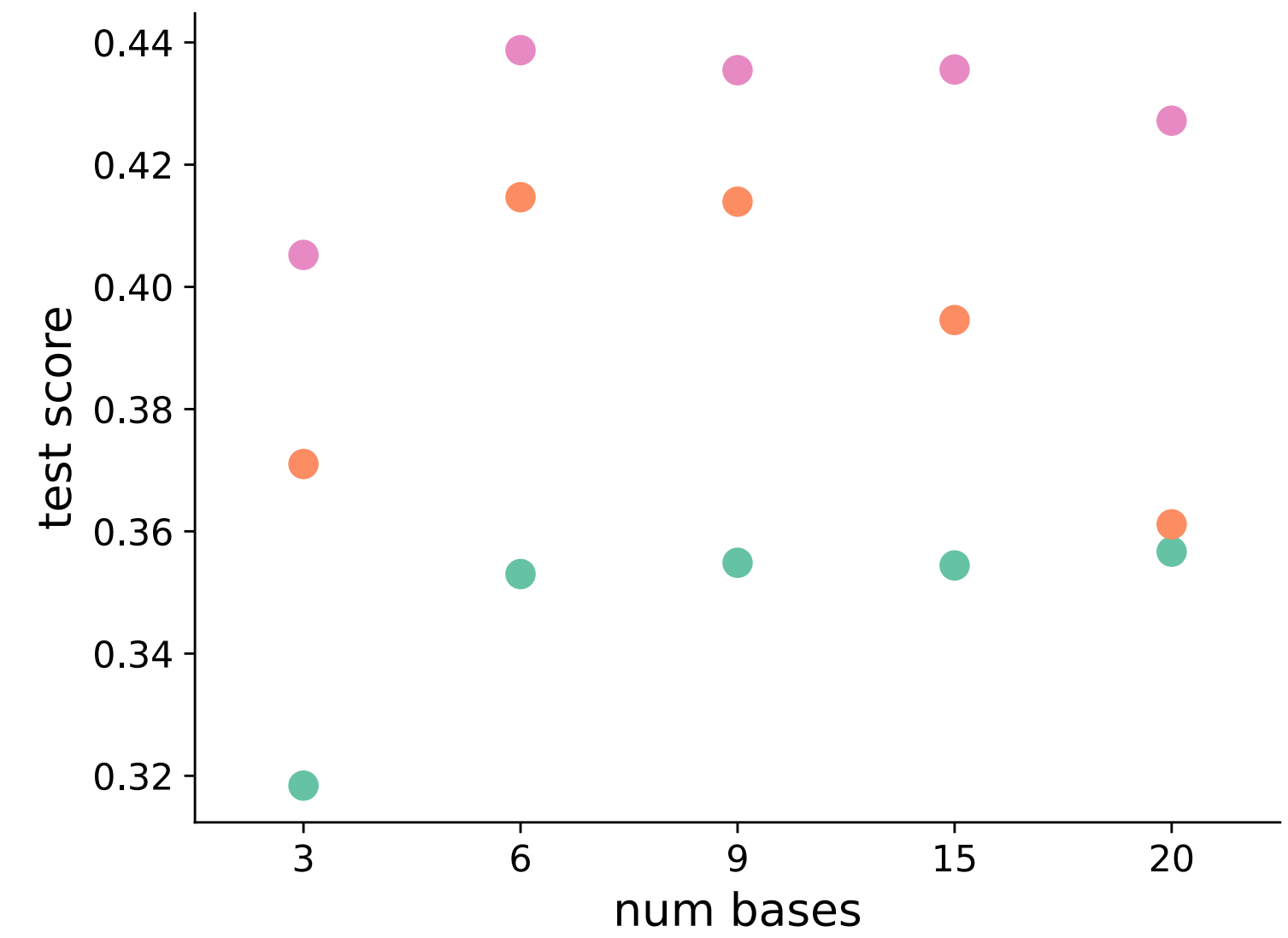
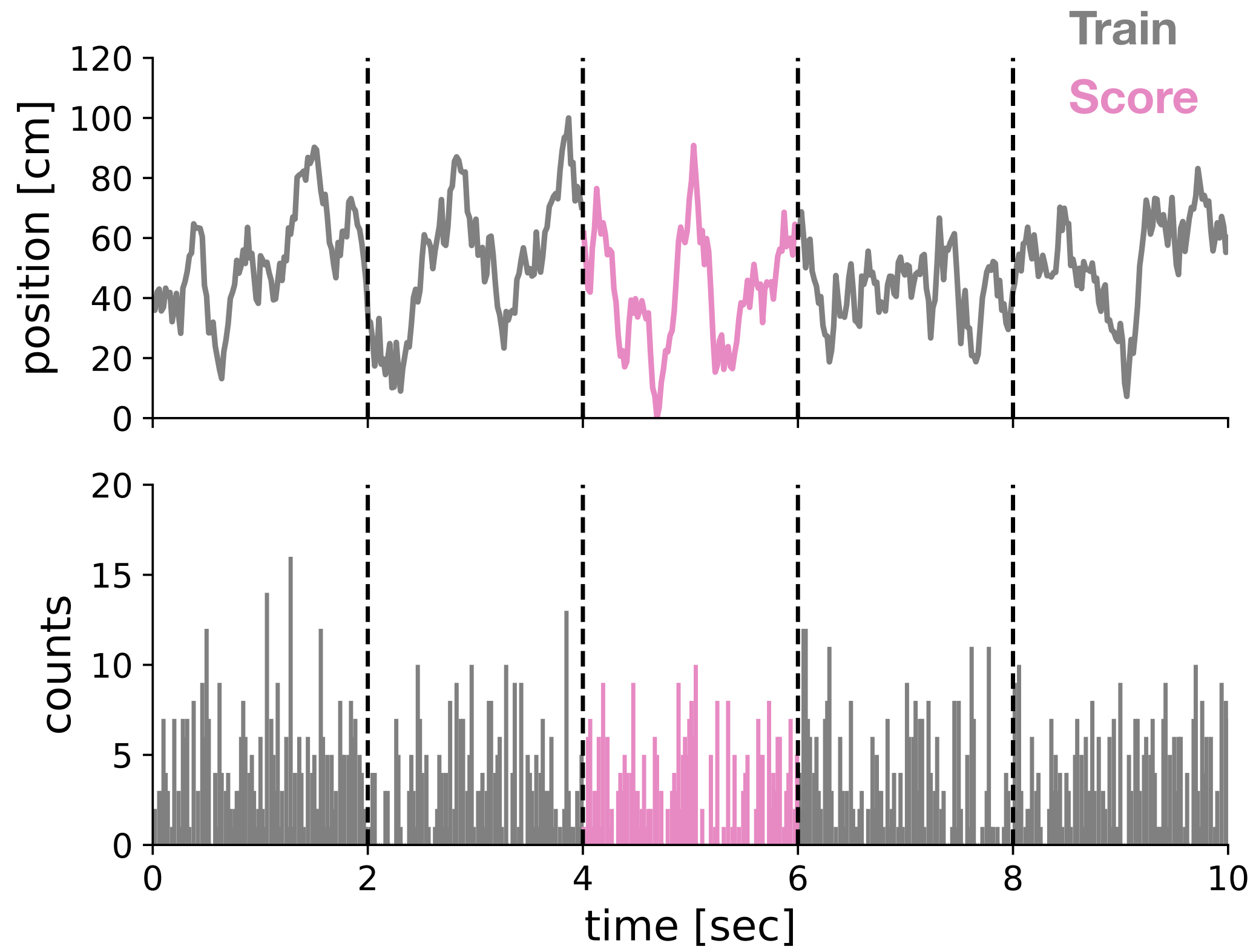
Example: K-Fold

Split 2



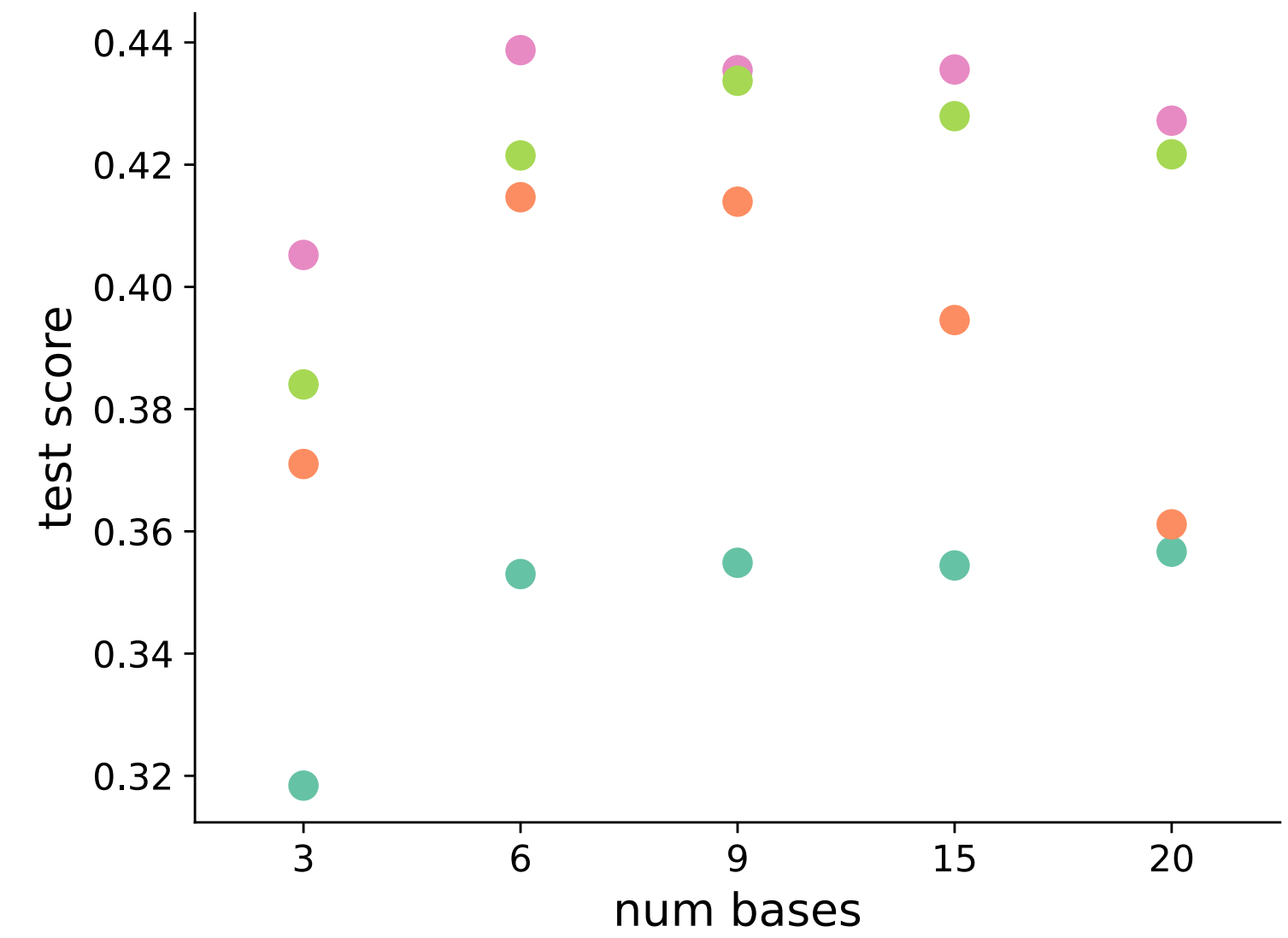
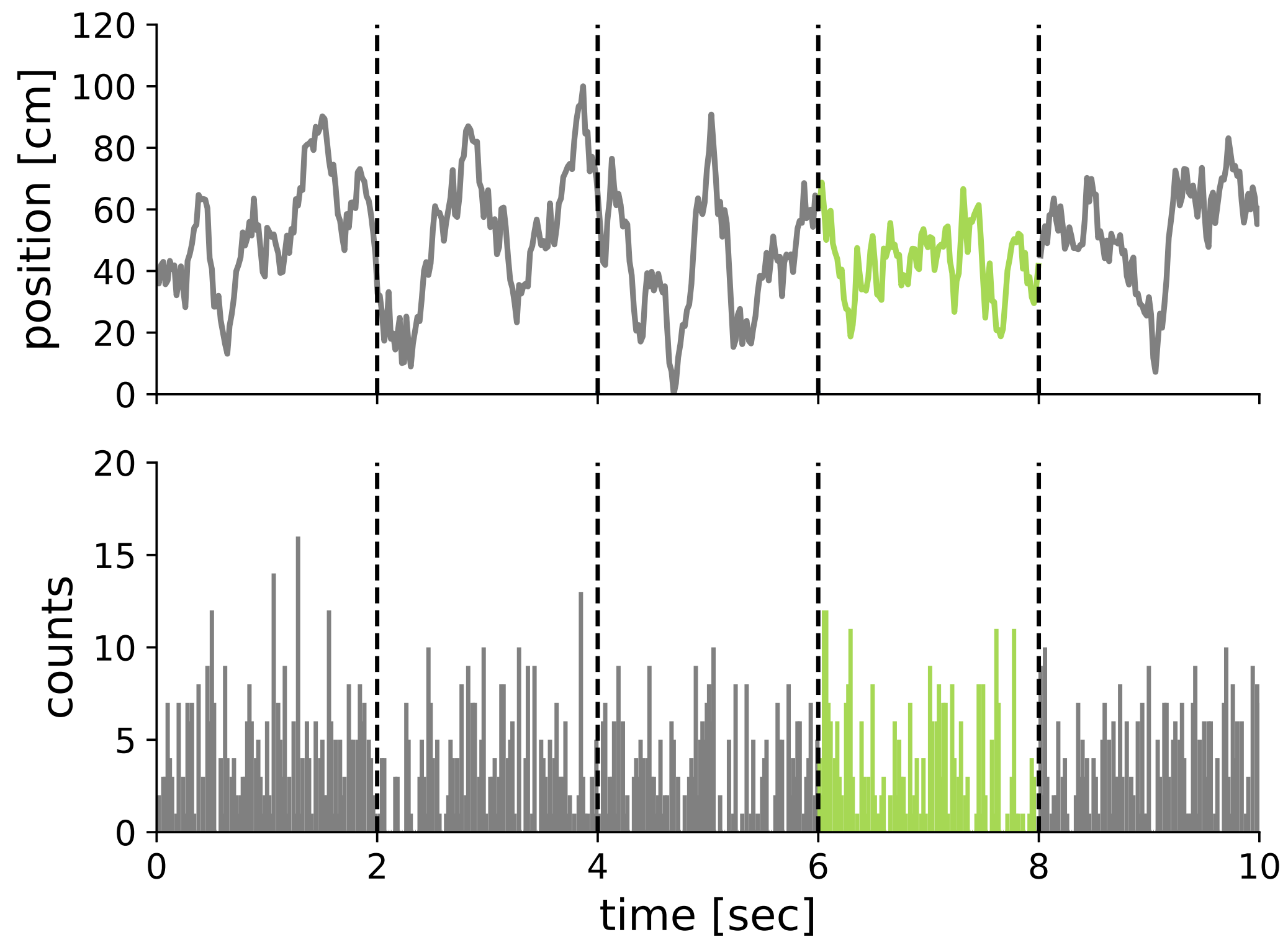
Example: K-Fold

Split 3



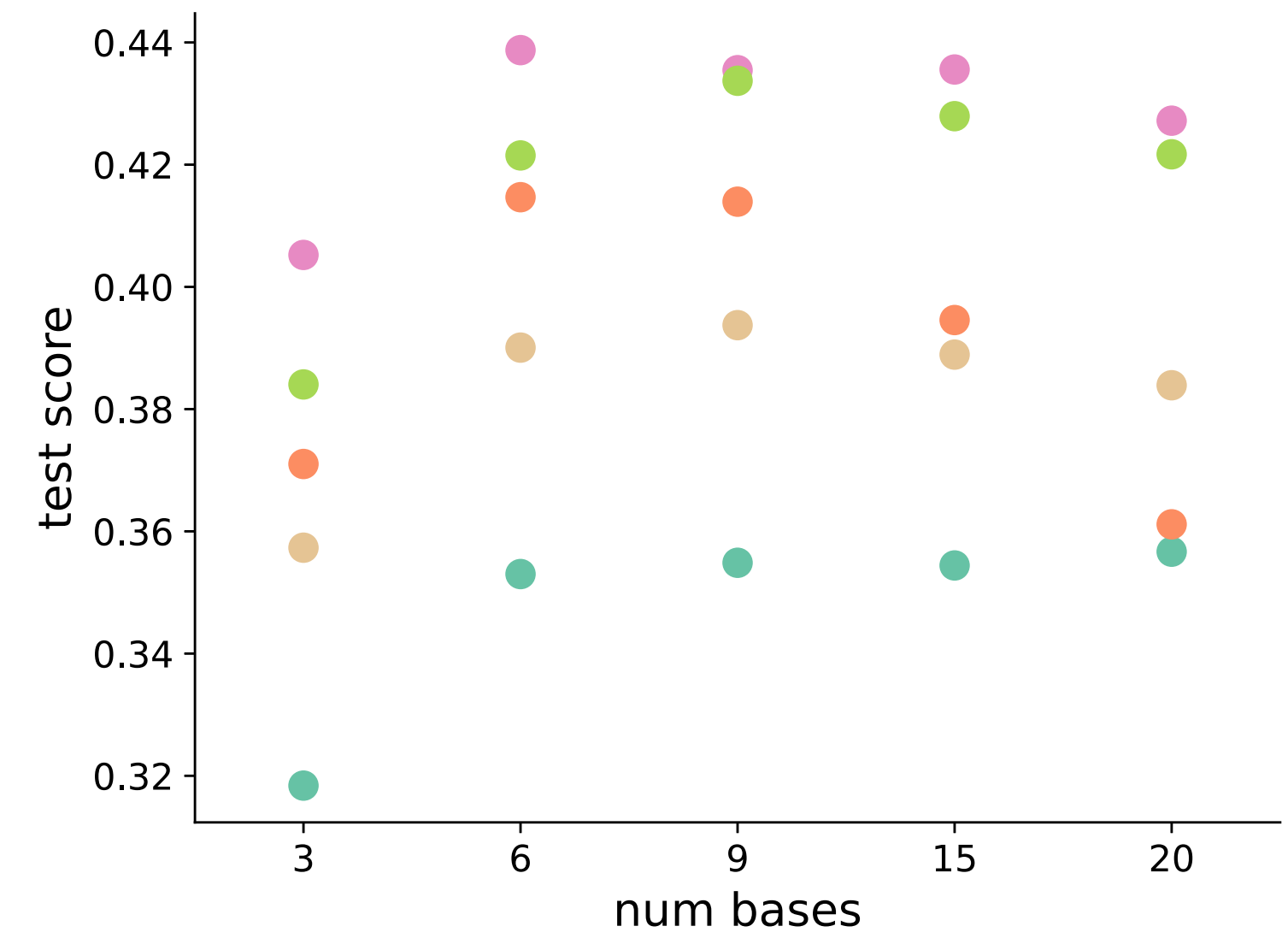
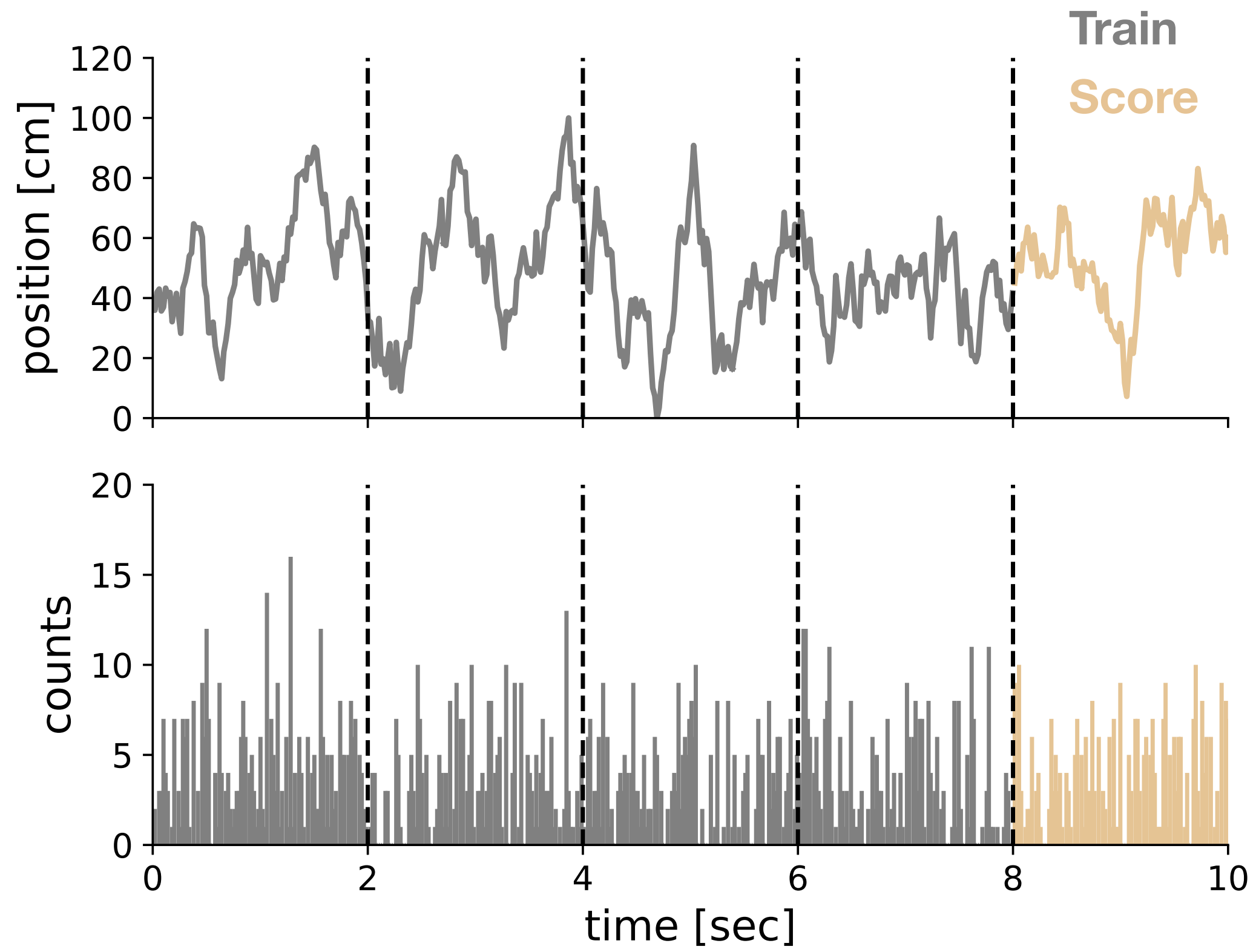
Example: K-Fold

Split 4

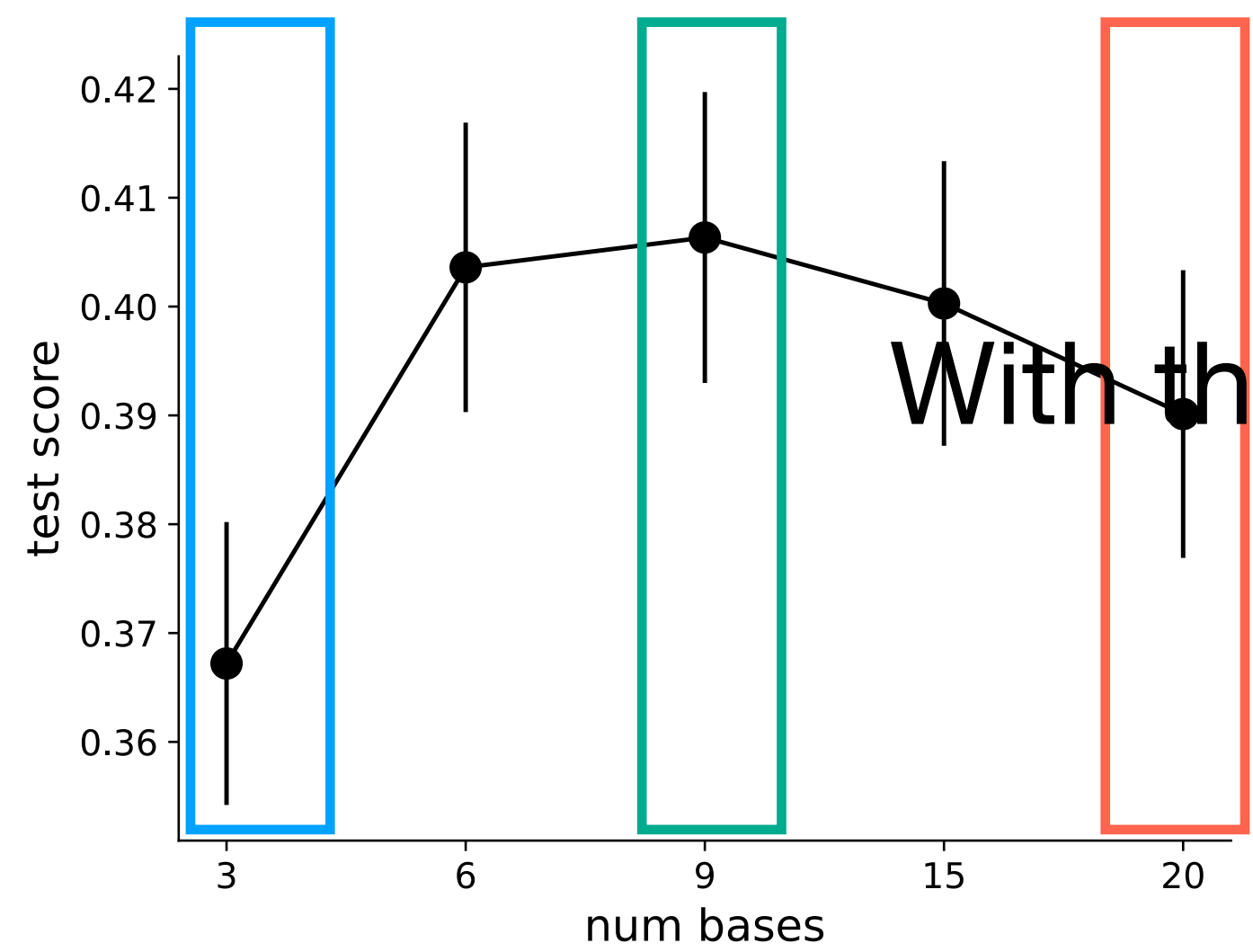


Example: K-Fold

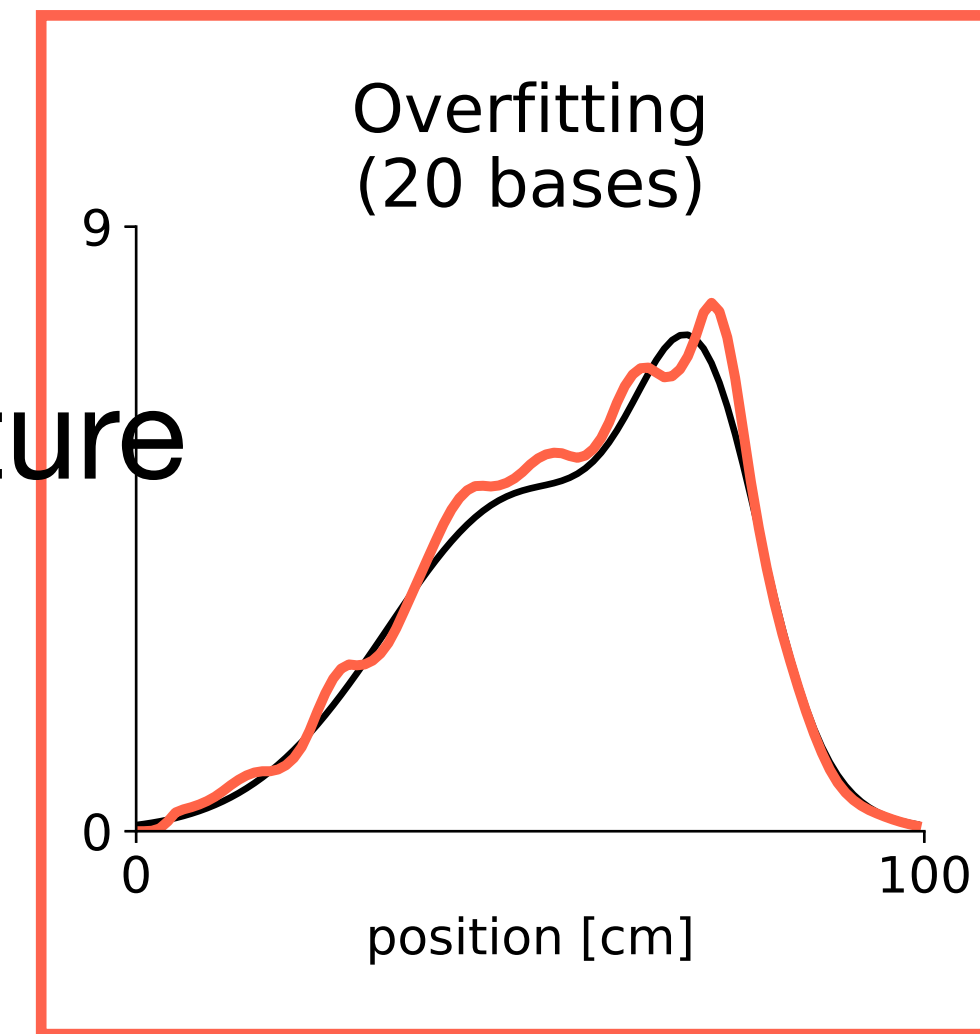
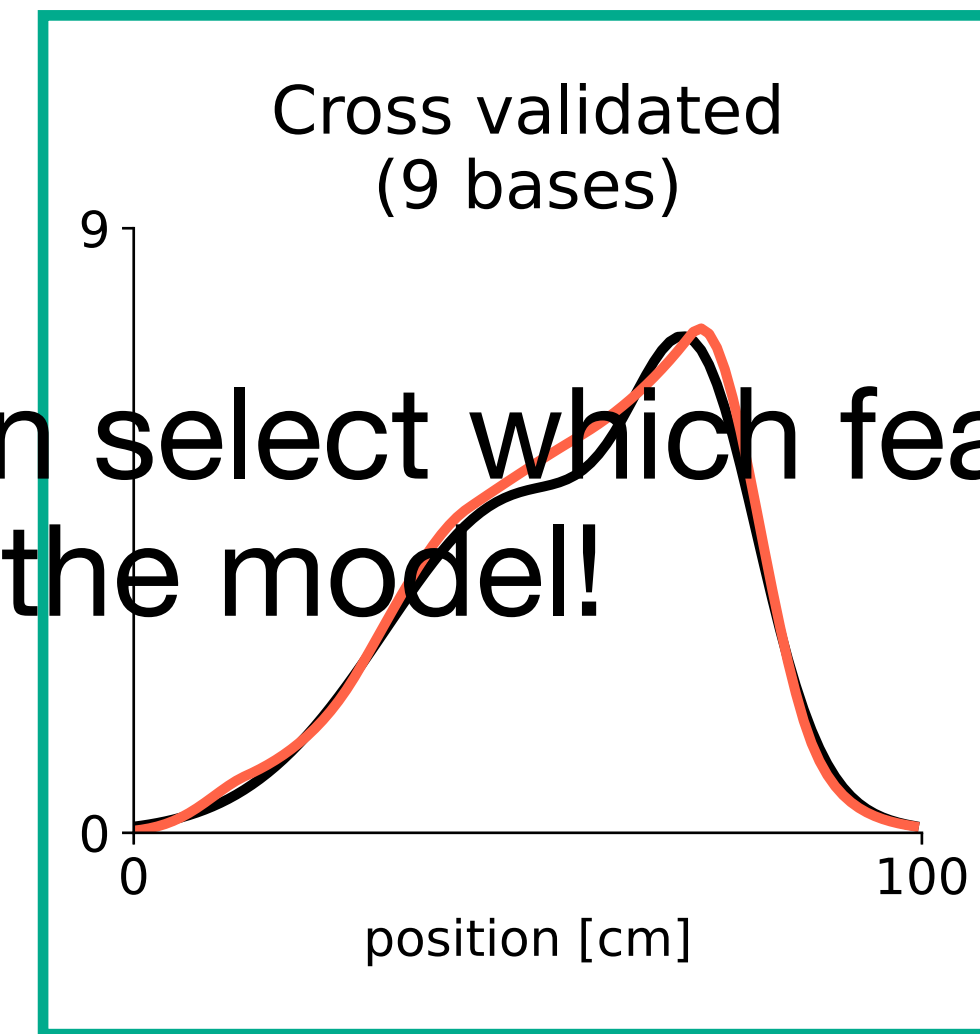
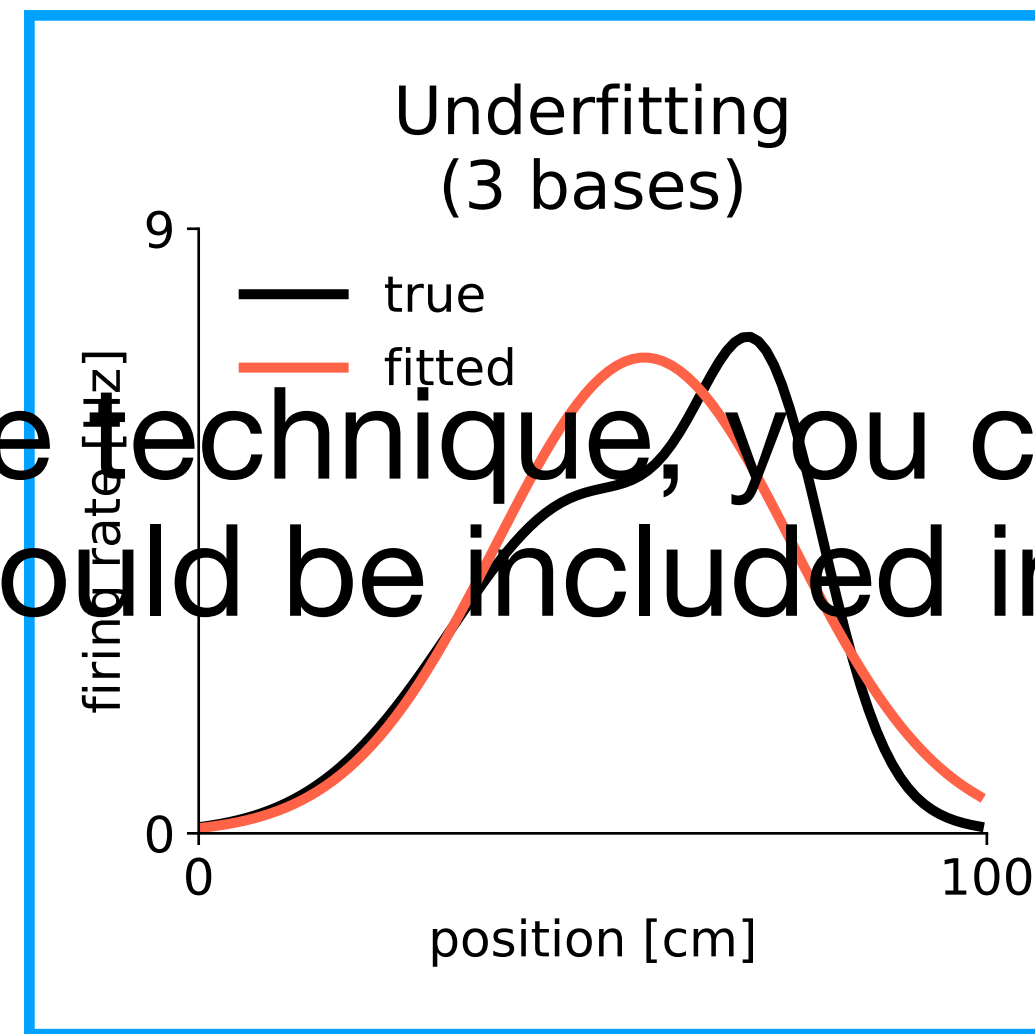
Split 5



Example: K-Fold



With the same technique, you can select which feature should be included in the model!



- Select the model with highest mean score.

Learn More

- [Scikit-Learn documentation](#)
- [Wikipedia](#)
- [NeMoS documentation](#)