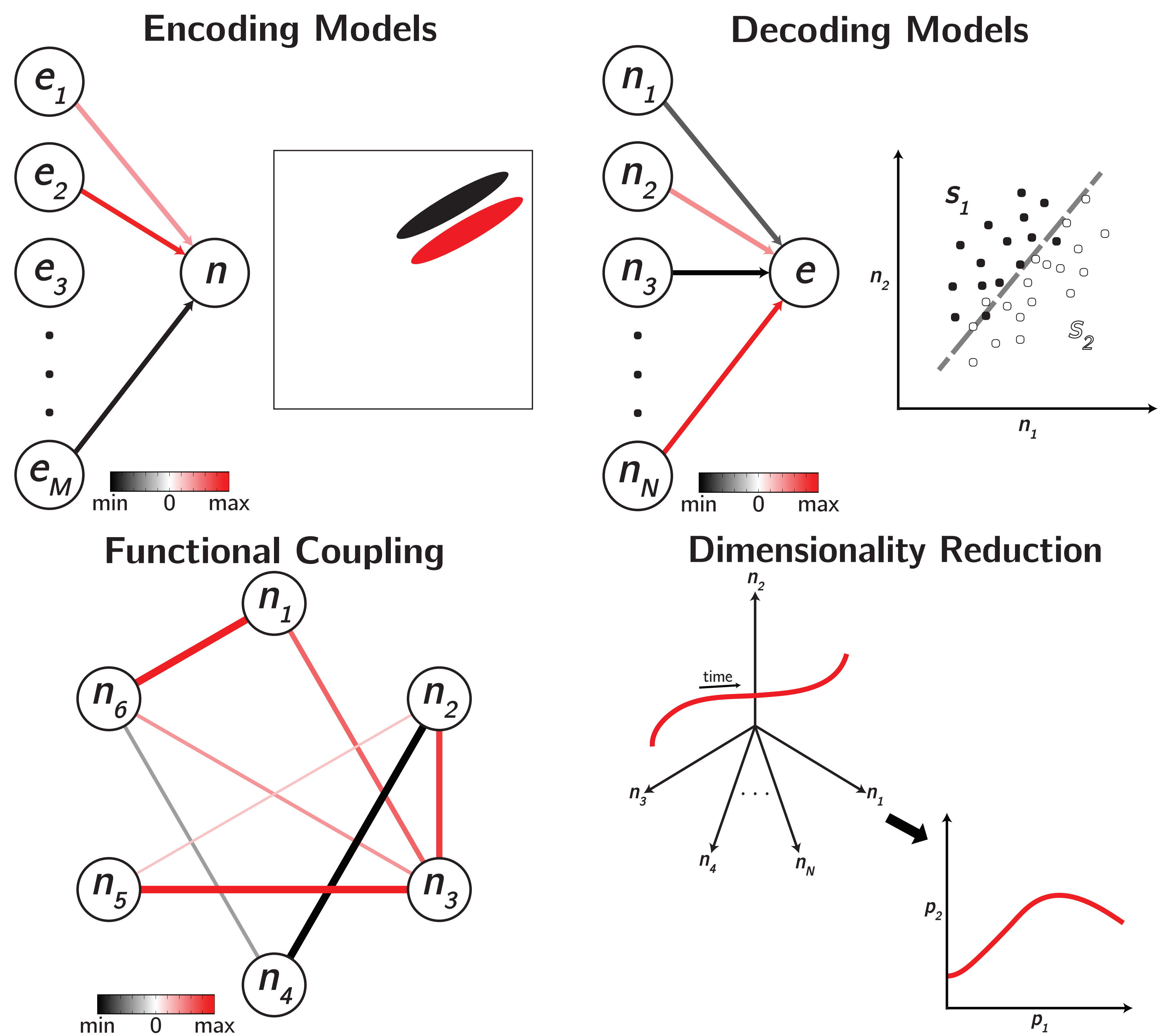
Sparse, Predictive, and Interpretable Functional Connectomics with Union of Intersections ($\text{UoI}_{\text{Lasso}}$)

Pratik S. Sachdeva[1,2],
Sharmodeep Bhattacharrya[3],
Kristofer E. Bouchard[1,2,4,5]

[1]Redwood Center for Theoretical Neuroscience, University of California, Berkeley
[2]Biological Sciences Division, Lawrence Berkeley National Laboratory
[3]Department of Statistics, Oregon State University
[4]Helen Wills Neuroscience Institute, University of California, Berkeley
[5]Computational Sciences Division, Lawrence Berkeley National Laboratory

Parametric Models in Neuroscience



Desired Properties of Statistical Models in (Neuro)science

property	*goal*	*metric*
predictive	ability to predict response variable(s) on new data	accuracy, log-likelihood, R^2, *etc.*
stable	returns the same values on multiple runs	variance
selective	only chooses features that influence the response variable(s)	false positives, false negatives, selection accuracy
accurate	values of estimated parameters are close to the "real" value	bias
scalable	capable of fitting to large dataset	time

Linear Model

features (*e.g.* activities of other neurons)

response variable (*e.g.* activity of a target neuron)

$$y = \sum_{i=1}^{p} \beta_i x_i + \epsilon$$

parameters to fit (*i.e.* which features are important, and how important they are)

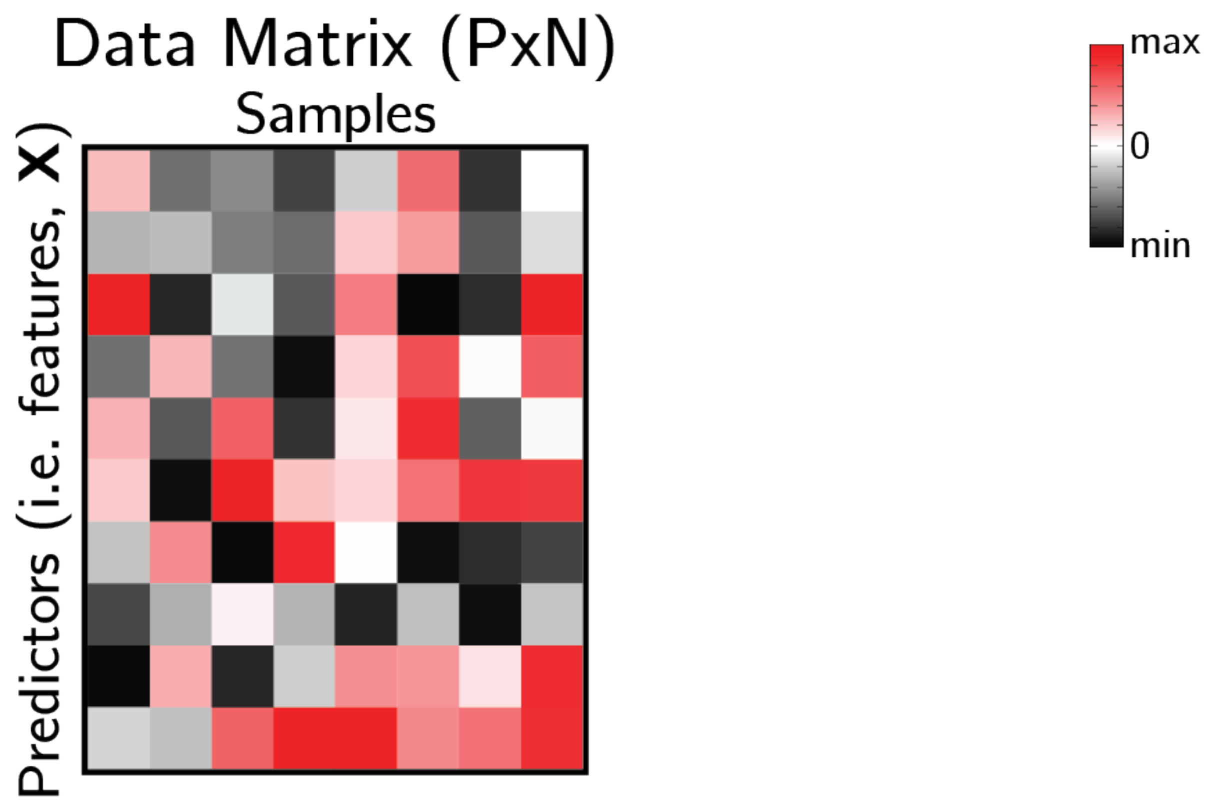
Lasso Penalty

$$\underset{\boldsymbol{\beta}}{\operatorname{argmin}} \; |\mathbf{y} - \mathbf{X}\boldsymbol{\beta}|_2^2 + \lambda|\boldsymbol{\beta}|_1$$

mean squared error (reconstruction)

weight penalty (enforces sparsity)

Feature Compression via Sparsity Regularization

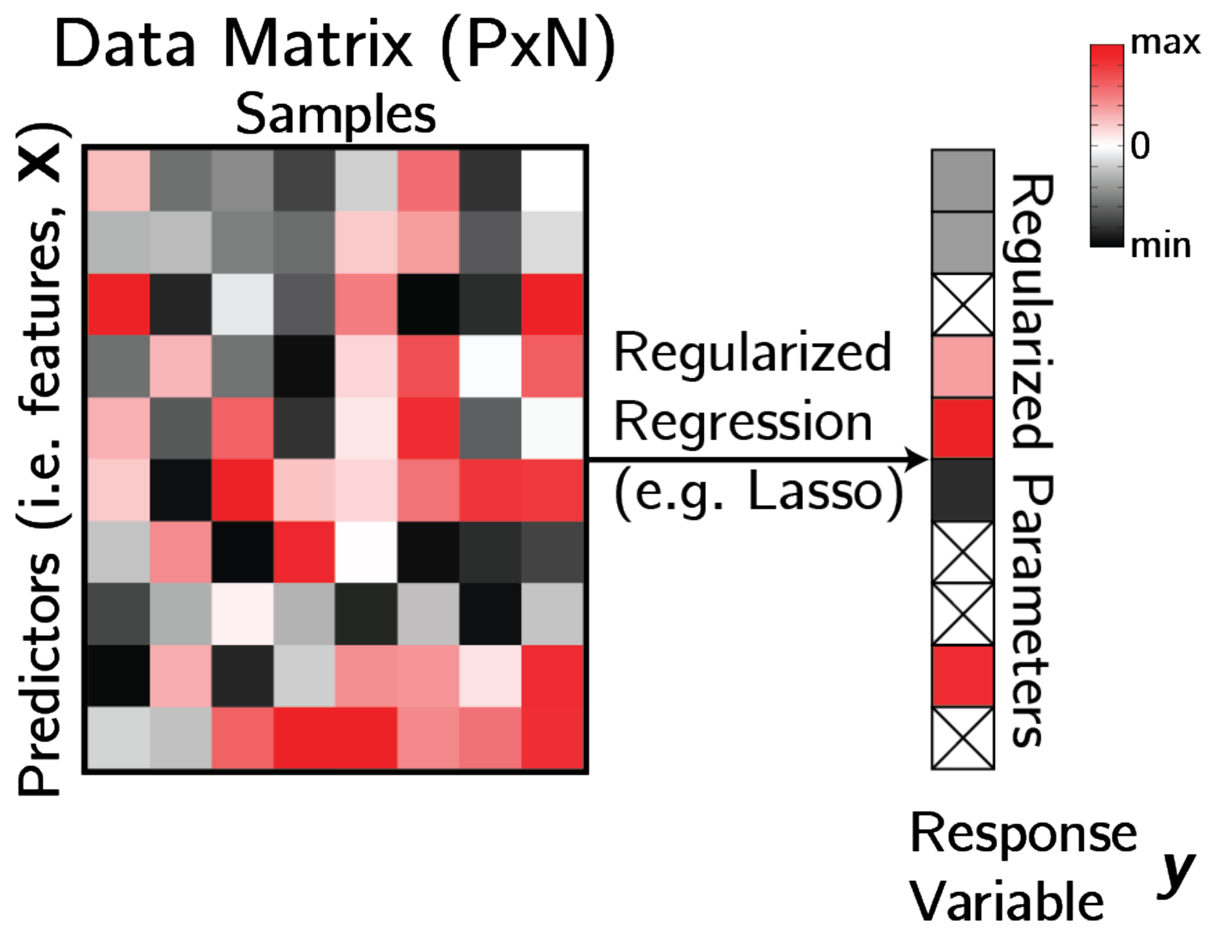


Response Variable ***y***

Breiman (1994)

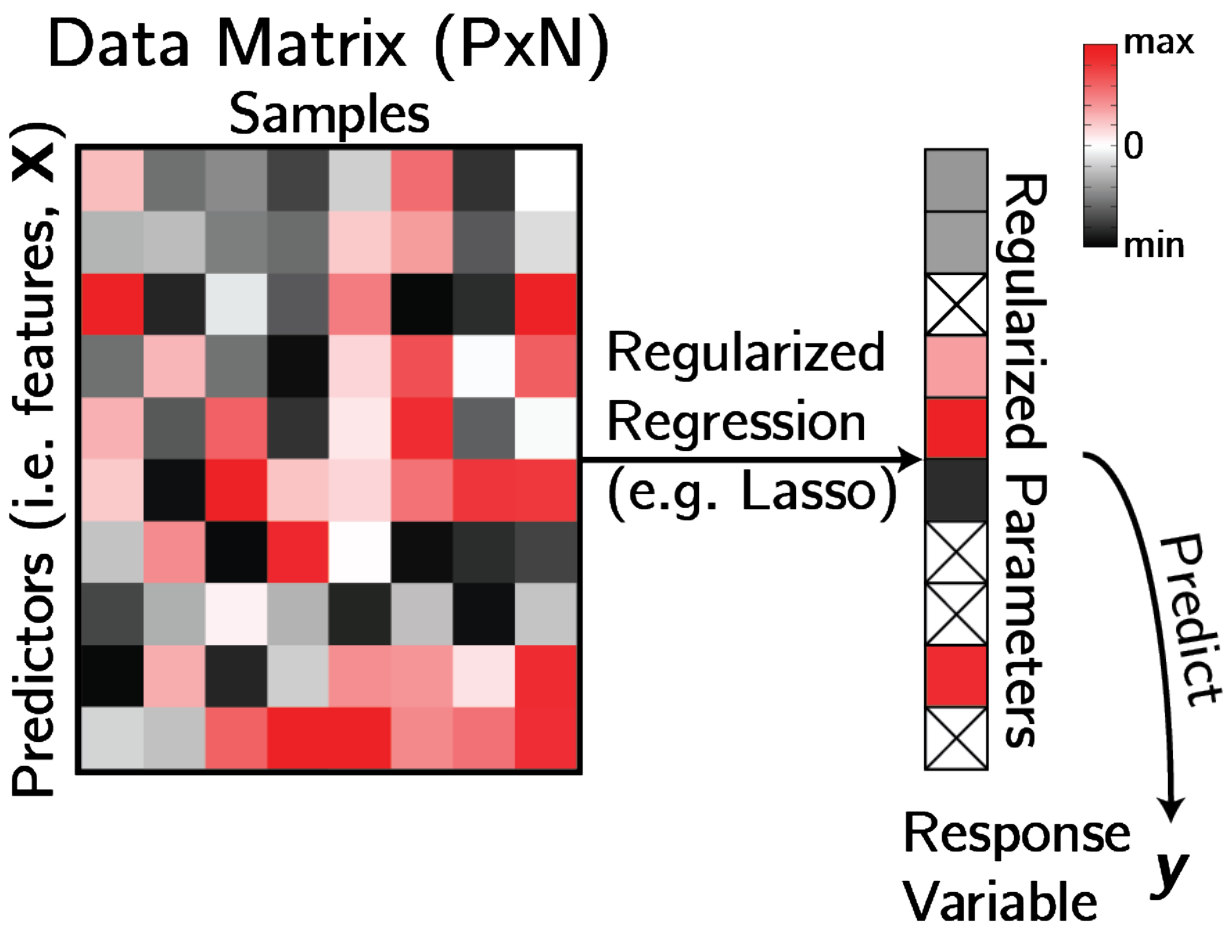
Tibshirani, J.R. Statist. Soc. B (1996)

Feature Compression via Sparsity Regularization



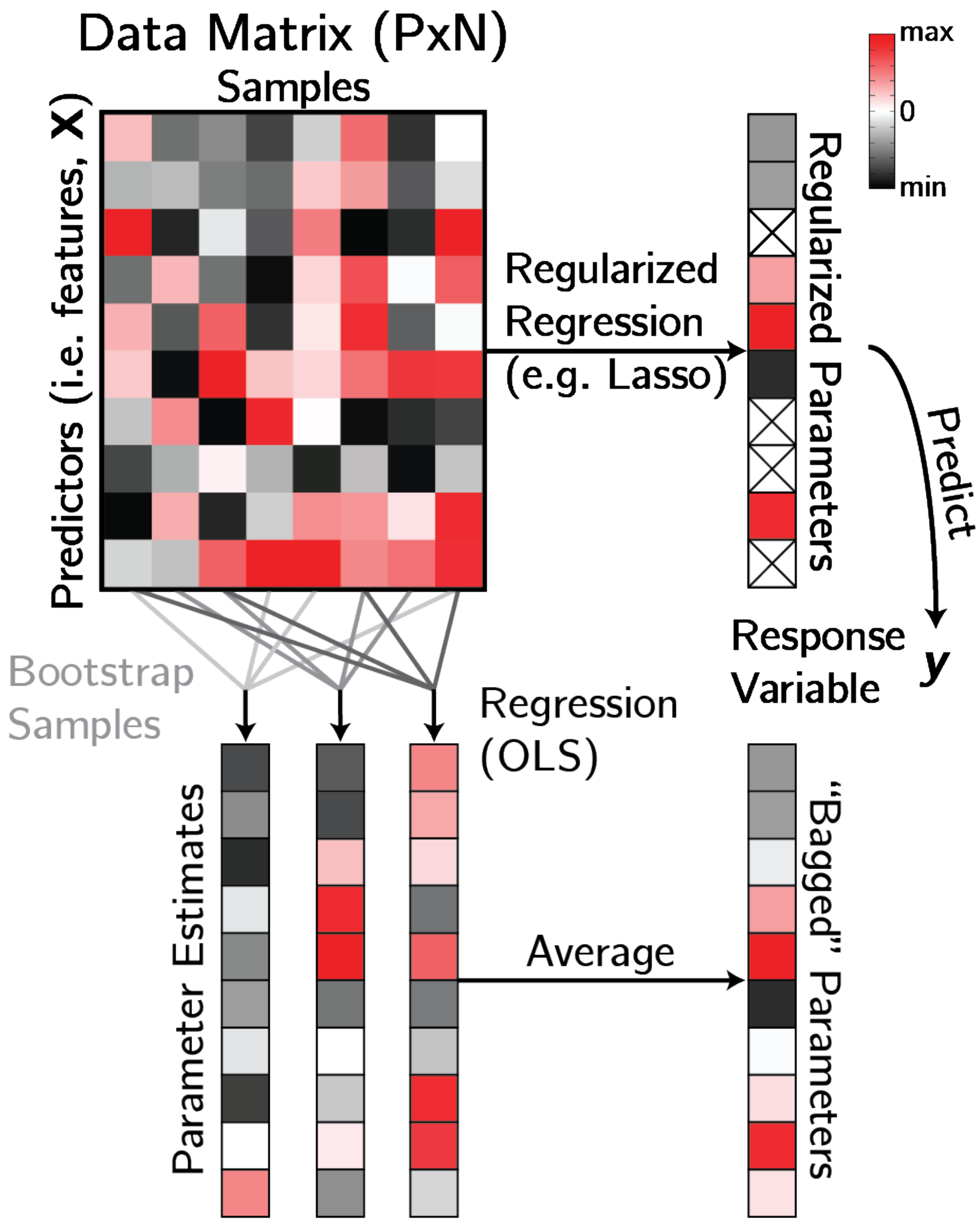
Breiman (1994)
Tibshirani, J.R. Statist. Soc. B (1996)

Feature Compression via Sparsity Regularization



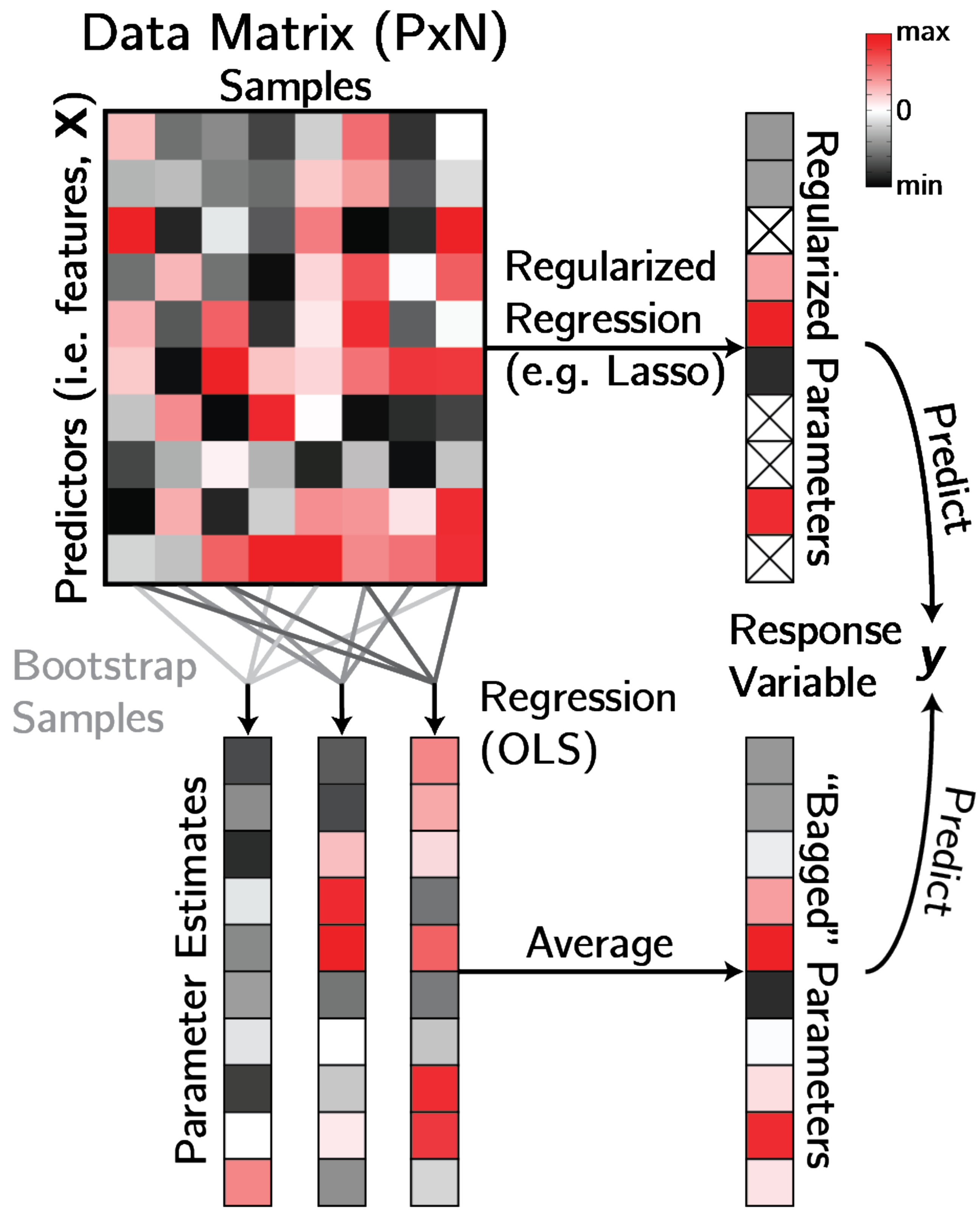
Breiman (1994)
Tibshirani, J.R. Statist. Soc. B (1996)

Feature Expansion via Ensemble Methods



Breiman (1994)
Tibshirani, J.R. Statist. Soc. B (1996)

Feature Expansion via Ensemble Methods



Breiman (1994)

Tibshirani, J.R. Statist. Soc. B (1996)

The Union of Intersections Framework

Intersection

Union

The Union of Intersections Framework

Intersection { **selection:** find features stable to perturbations
feature compression

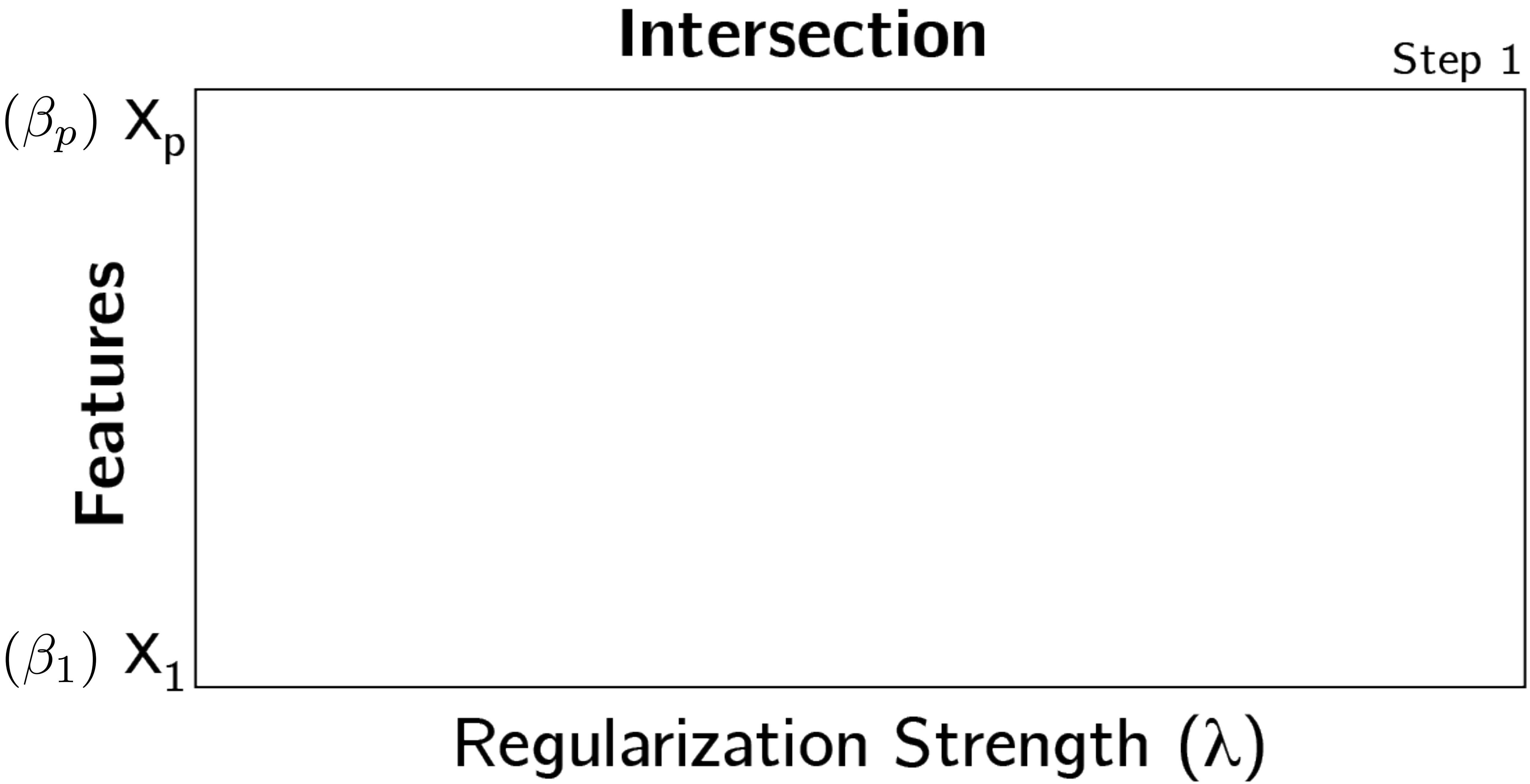
Union

The Union of Intersections Framework

Intersection {
selection: find features stable to perturbations
feature compression

Union {
estimation: combine the most predictive selection profiles
feature expansion

Intersection Module: Construct Selection Profiles



support

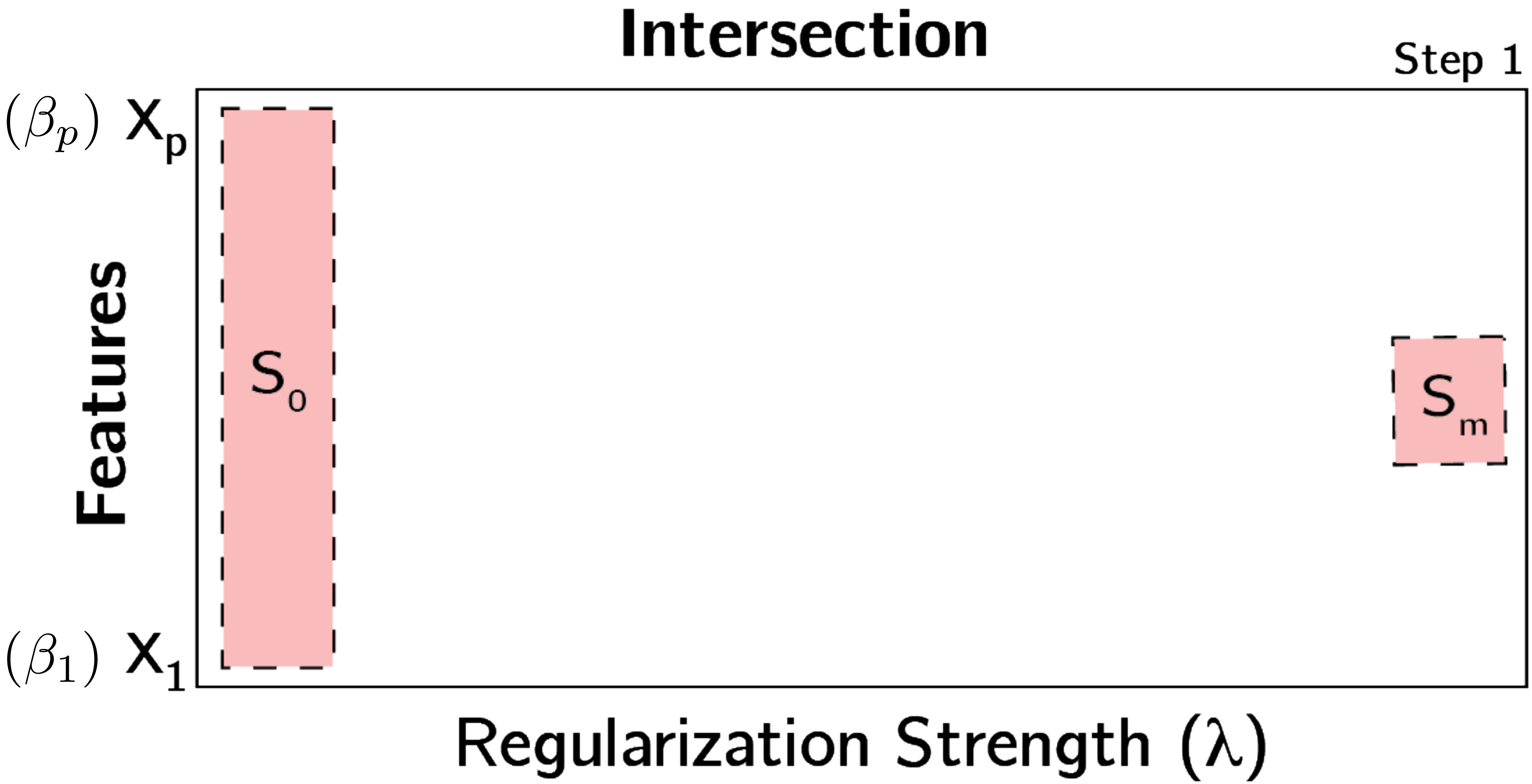
$$S_k = \{j : \beta_j \neq 0\} \text{ for } \lambda_k$$

total support

intersected support

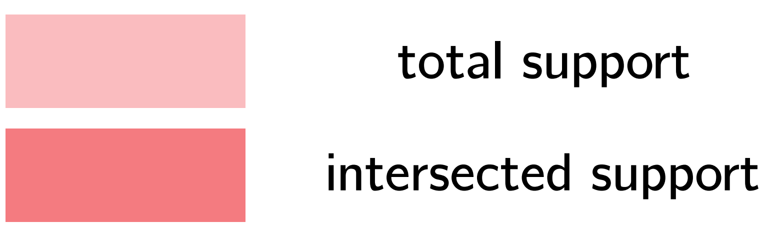
Bach, ICML (2008)

Intersection Module: Construct Selection Profiles



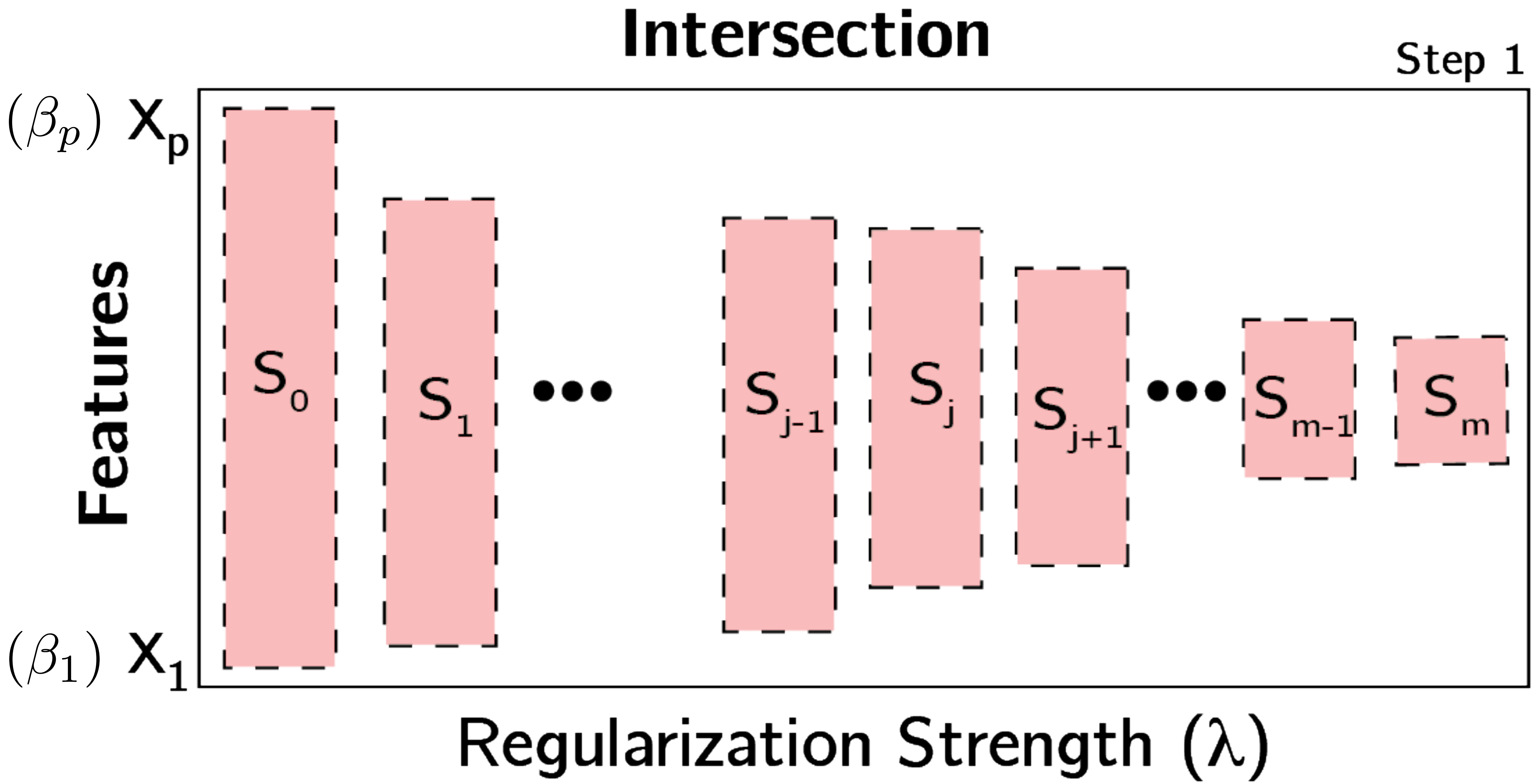
support

$S_k = \{j : \beta_j \neq 0\}$ for λ_k



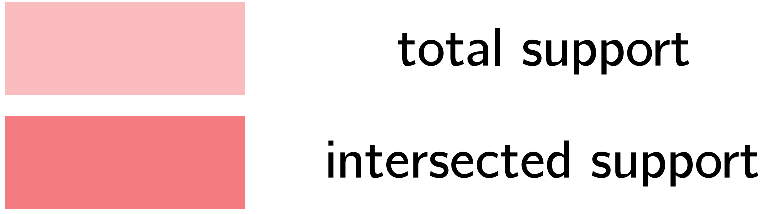
Bach, ICML (2008)

Intersection Module: Construct Selection Profiles



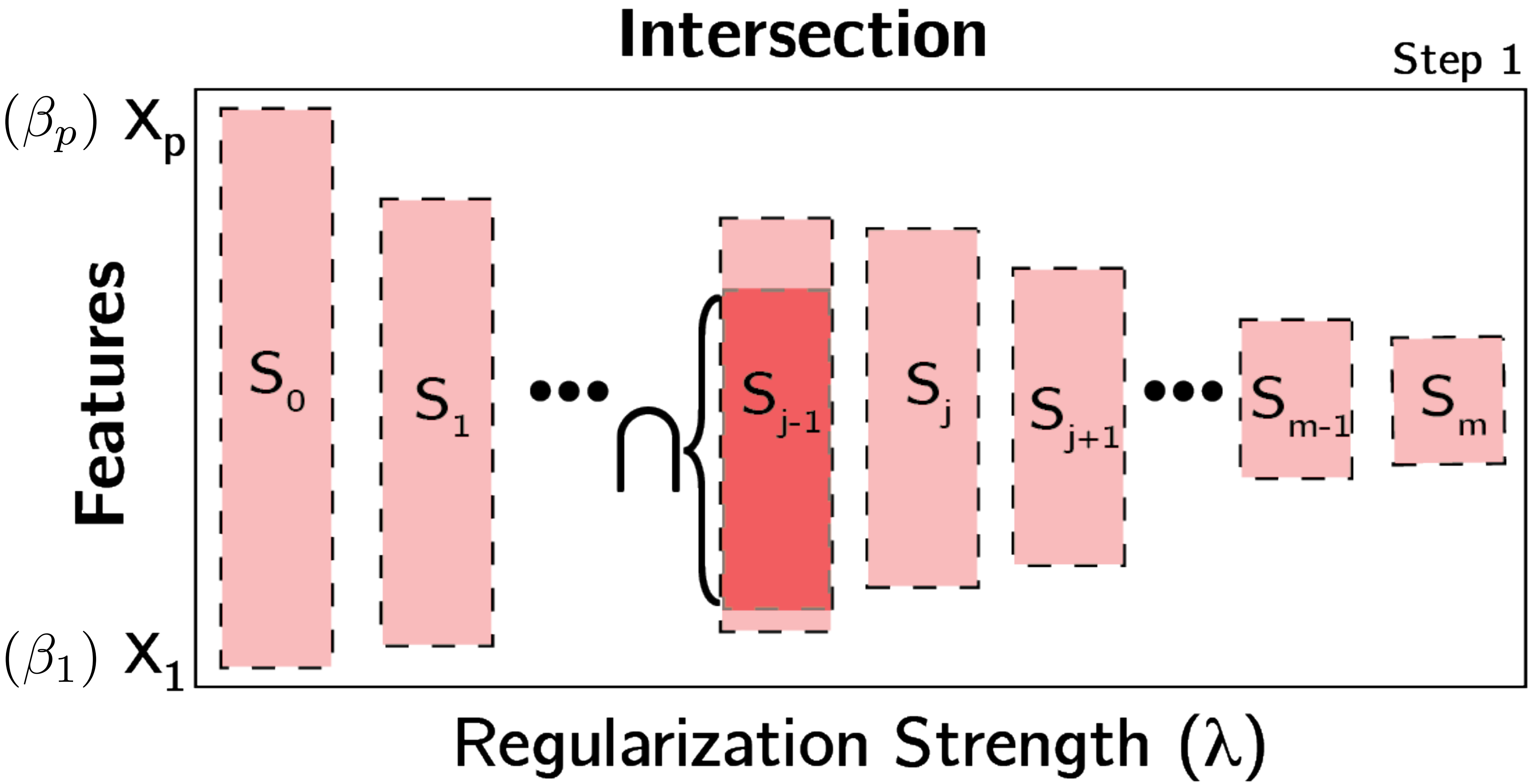
support

$S_k = \{j : \beta_j \neq 0\}$ for λ_k



Bach, ICML (2008)

Intersection Module: Construct Selection Profiles



support

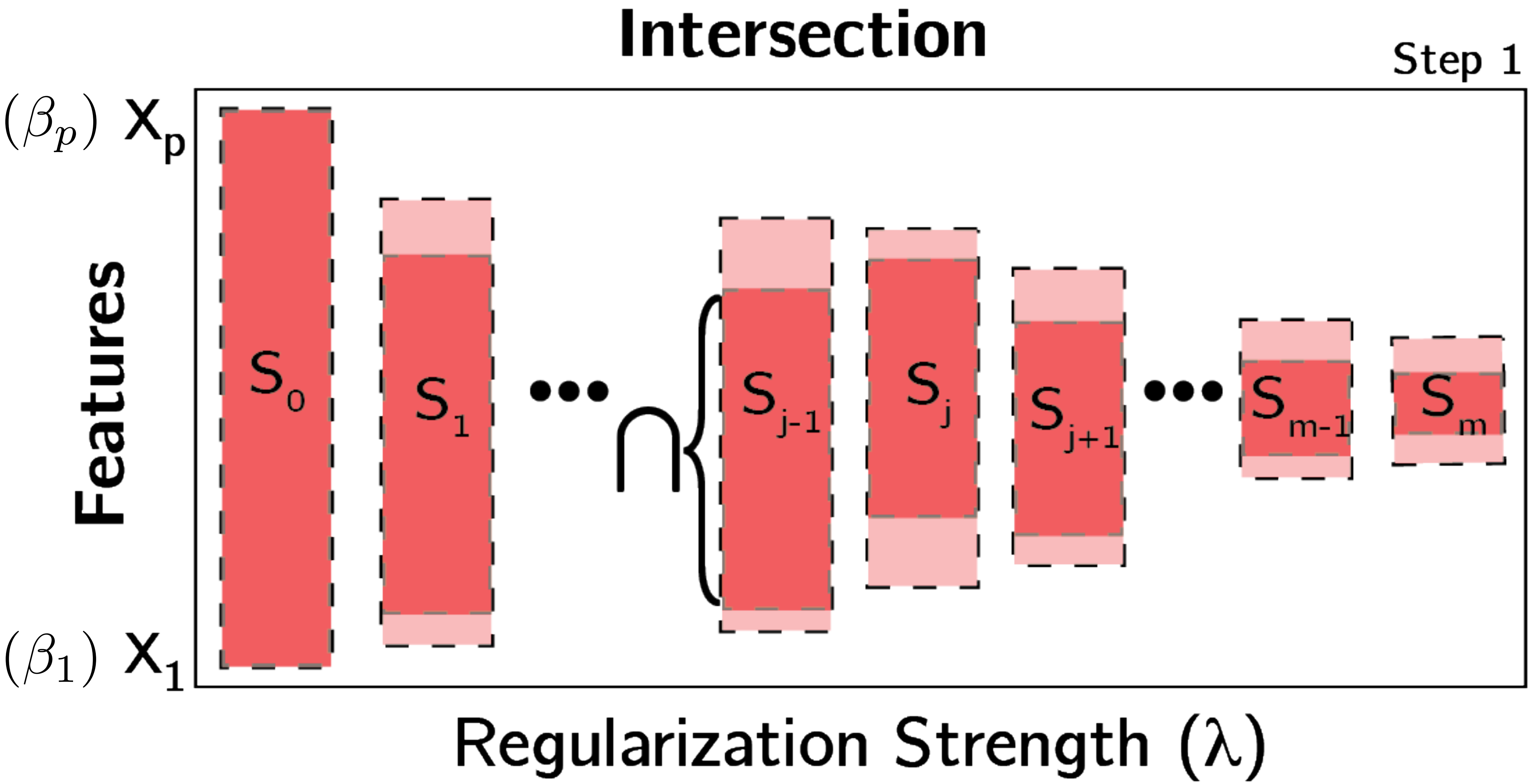
$$S_k = \{j : \beta_j \neq 0\} \text{ for } \lambda_k$$

total support

intersected support

Bach, ICML (2008)

Intersection Module: Construct Selection Profiles



support

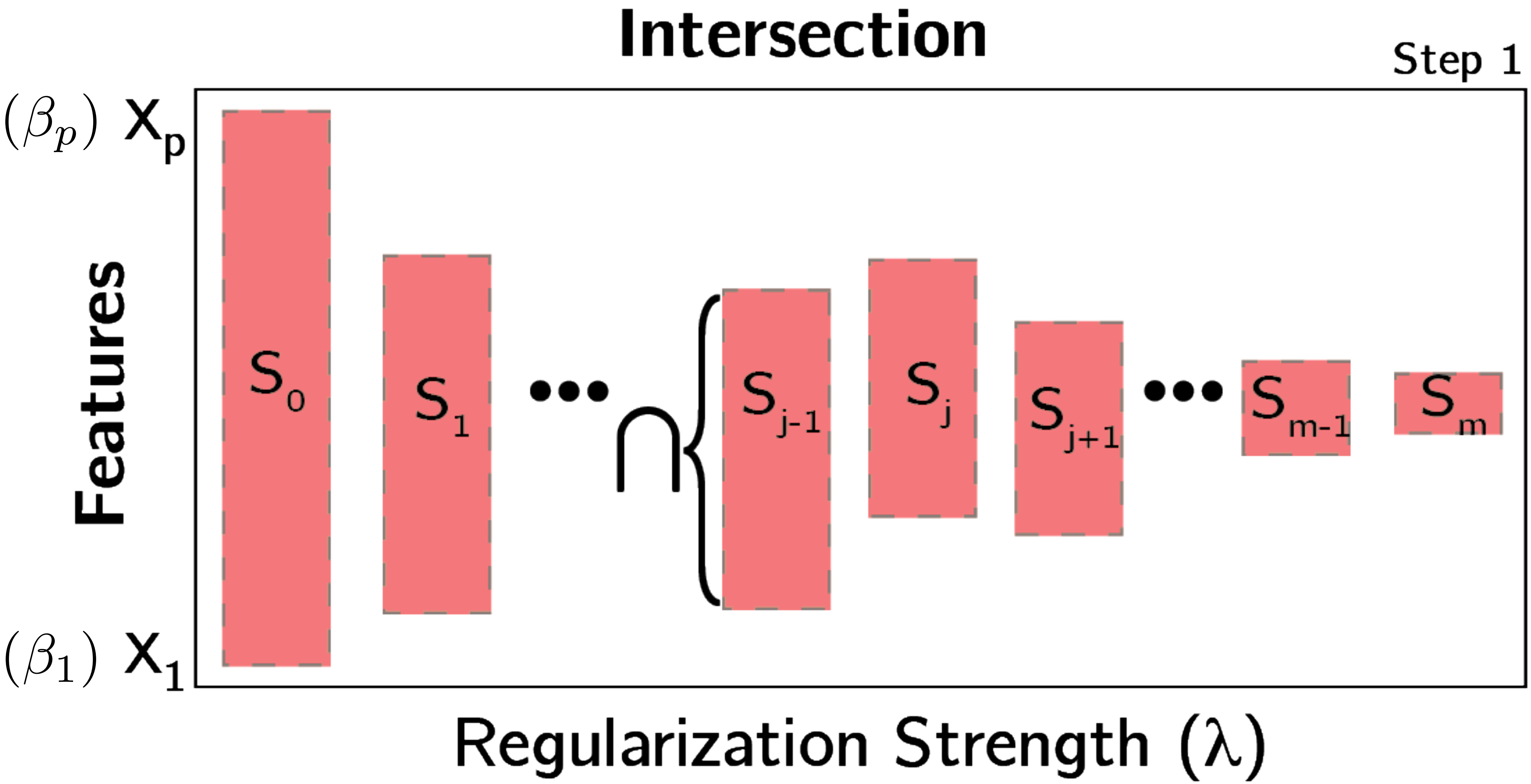
$$S_k = \{j : \beta_j \neq 0\} \text{ for } \lambda_k$$

total support

intersected support

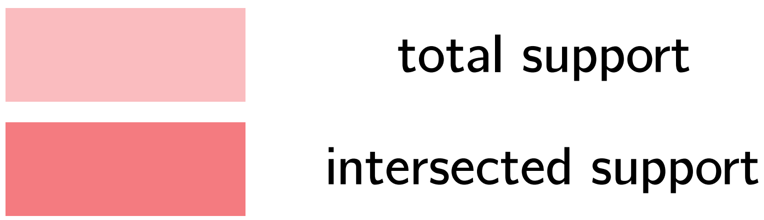
Bach, ICML (2008)

Intersection Module: Construct Selection Profiles



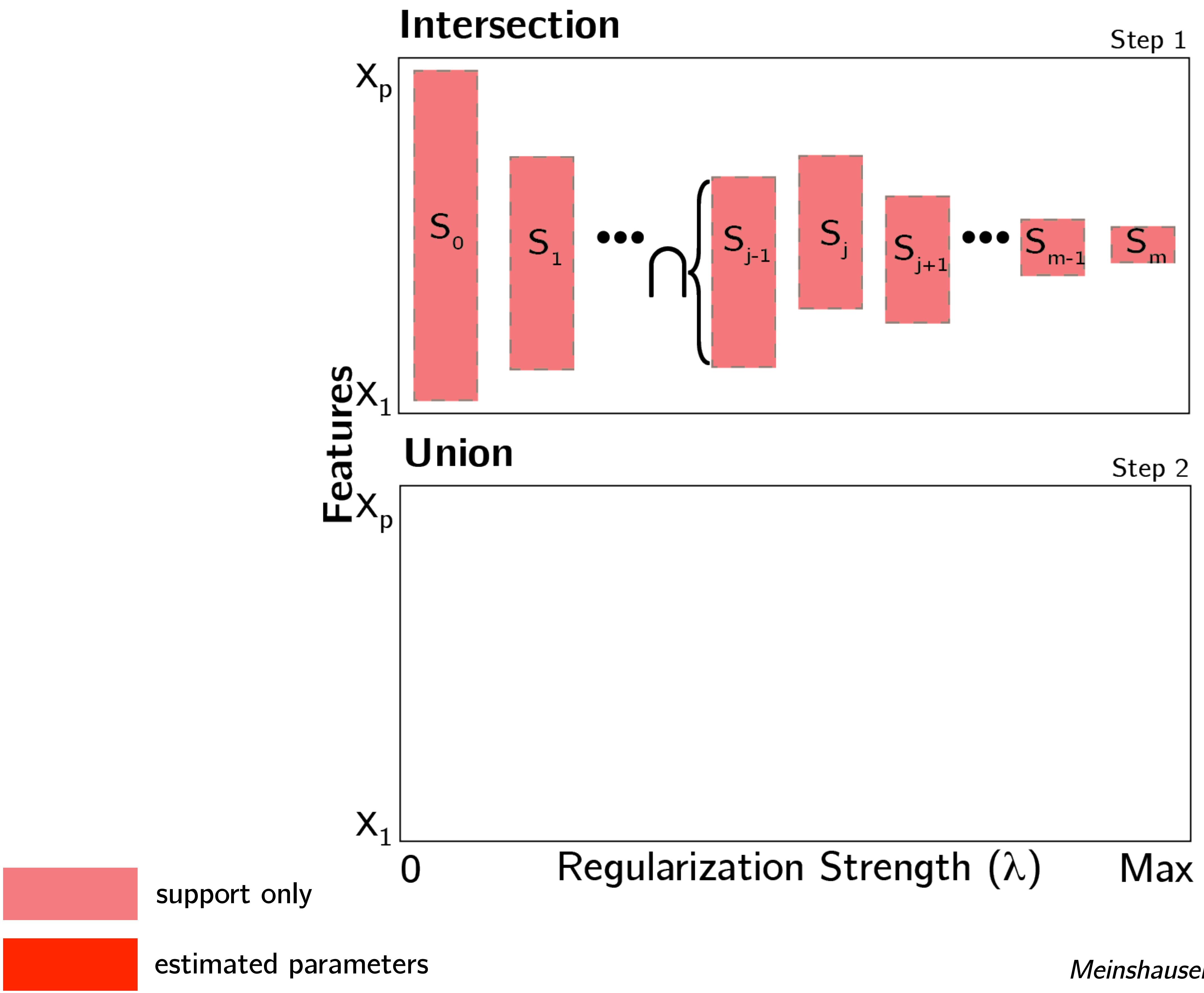
support

$$S_k = \{j : \beta_j \neq 0\} \text{ for } \lambda_k$$



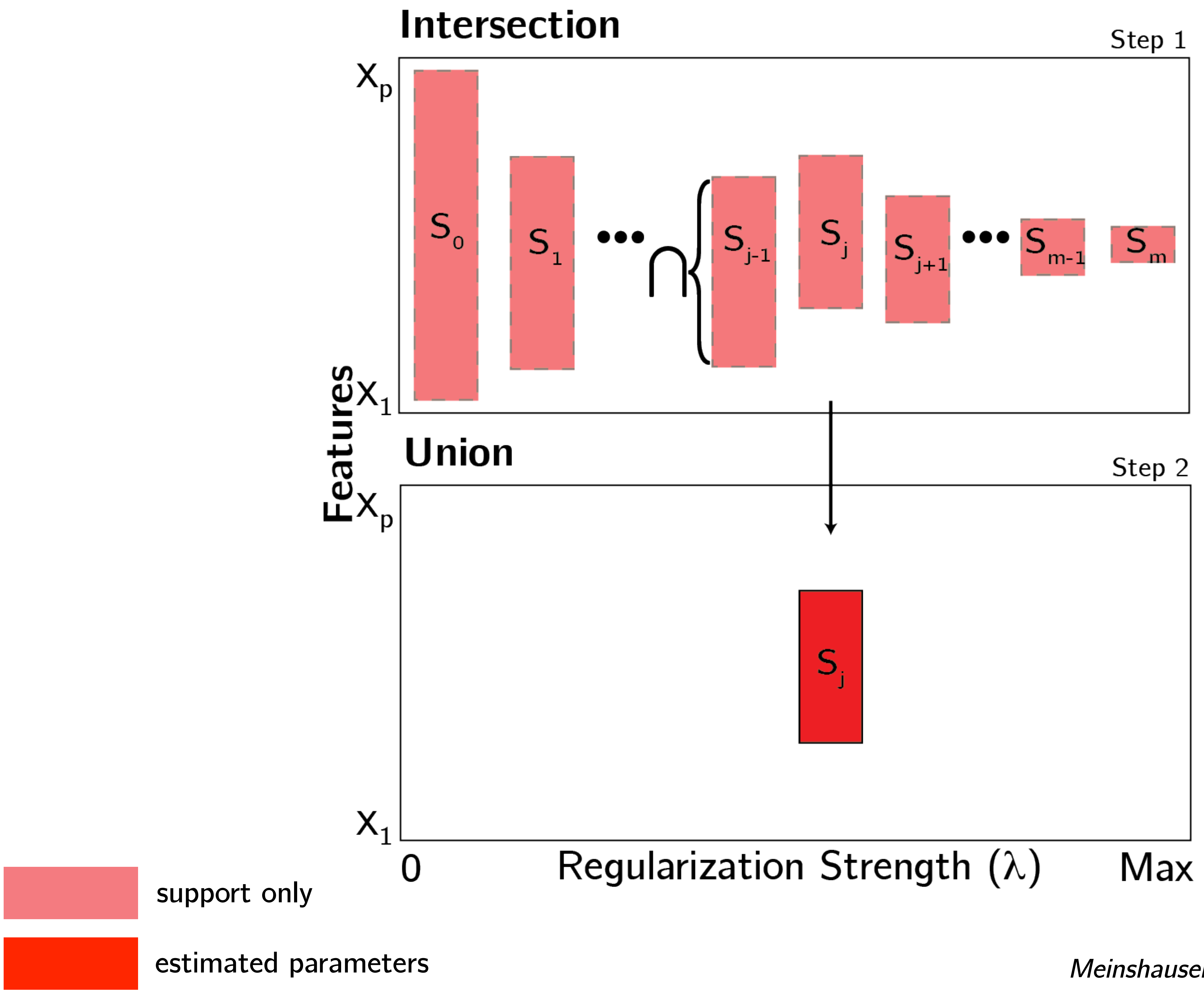
Bach, ICML (2008)

Estimation Module: Consolidate *Predictive* Selection Profiles



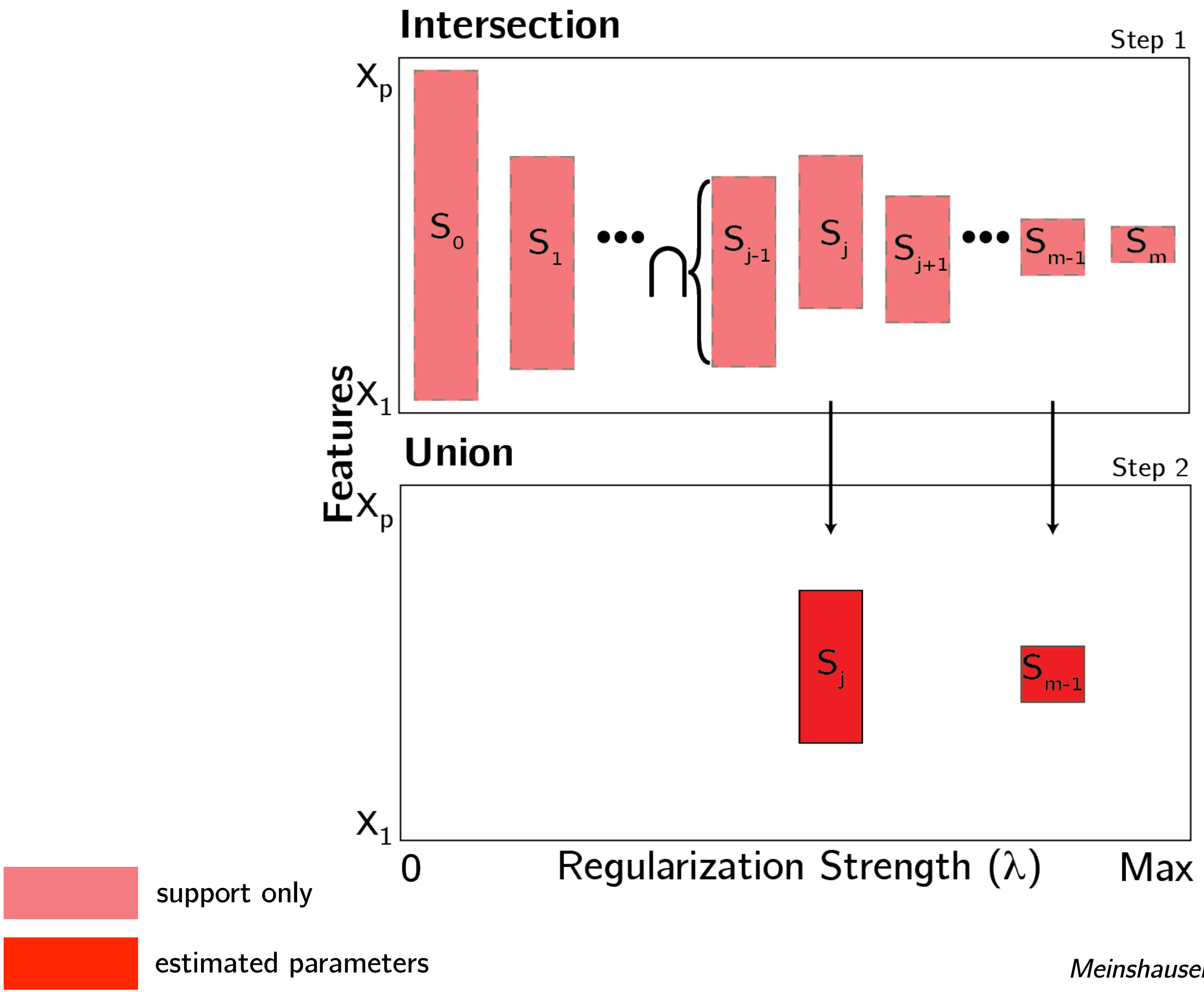
Breiman (1994)
Bach, ICML (2008)
Meinshausen & Bühlmann, RSS (2010)

Estimation Module: Consolidate *Predictive* Selection Profiles



Breiman (1994)
Bach, ICML (2008)
Meinshausen & Bühlmann, RSS (2010)

Estimation Module: Consolidate *Predictive* Selection Profiles

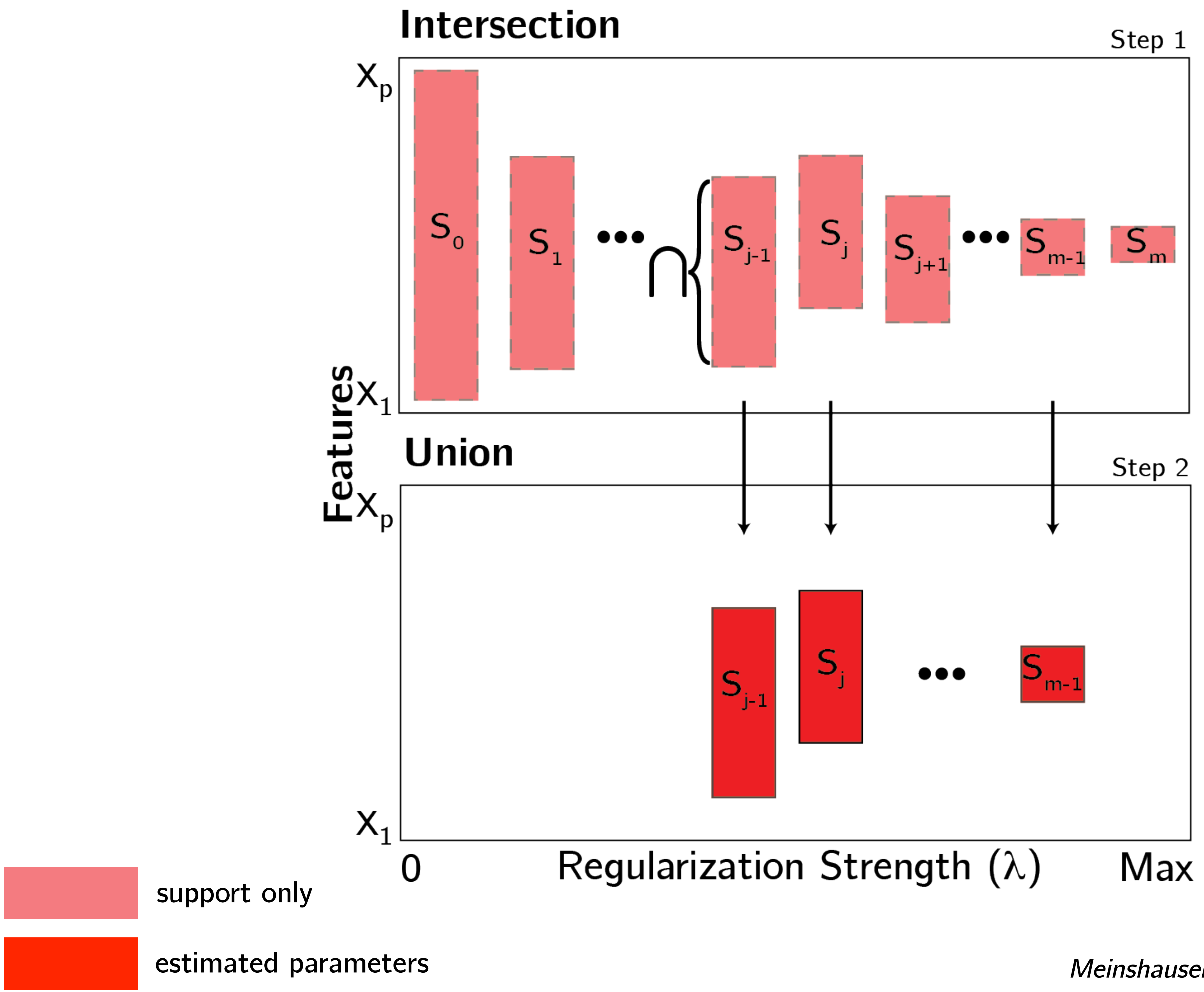


Breiman (1994)

Bach, ICML (2008)

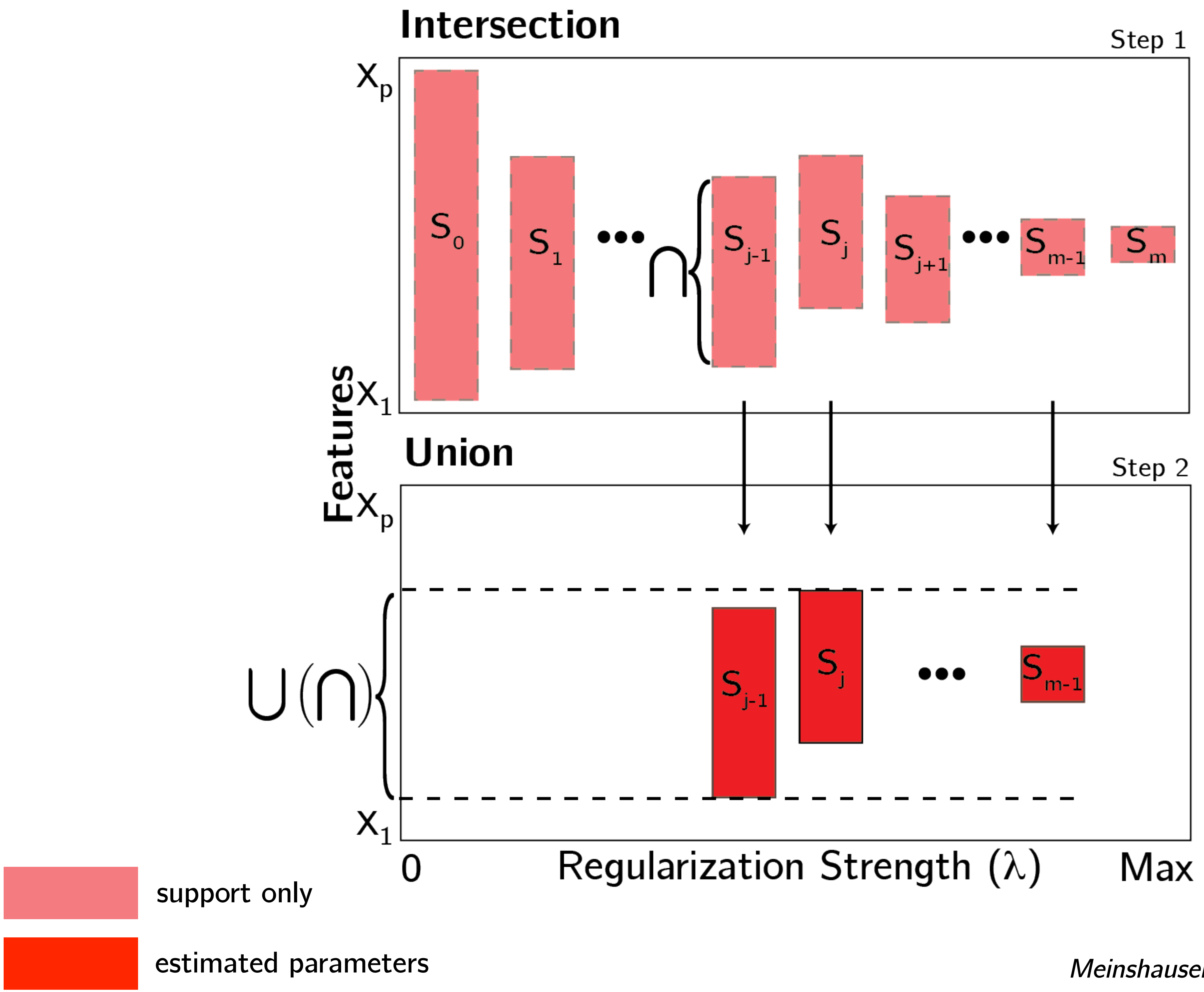
Meinshausen & Bühlmann, RSS (2010)

Estimation Module: Consolidate *Predictive* Selection Profiles



Breiman (1994)
Bach, ICML (2008)
Meinshausen & Bühlmann, RSS (2010)

Estimation Module: Consolidate *Predictive* Selection Profiles



Breiman (1994)
Bach, ICML (2008)
Meinshausen & Bühlmann, RSS (2010)

Results on Synthetic Data

Lasso | **UoI_{Lasso}**

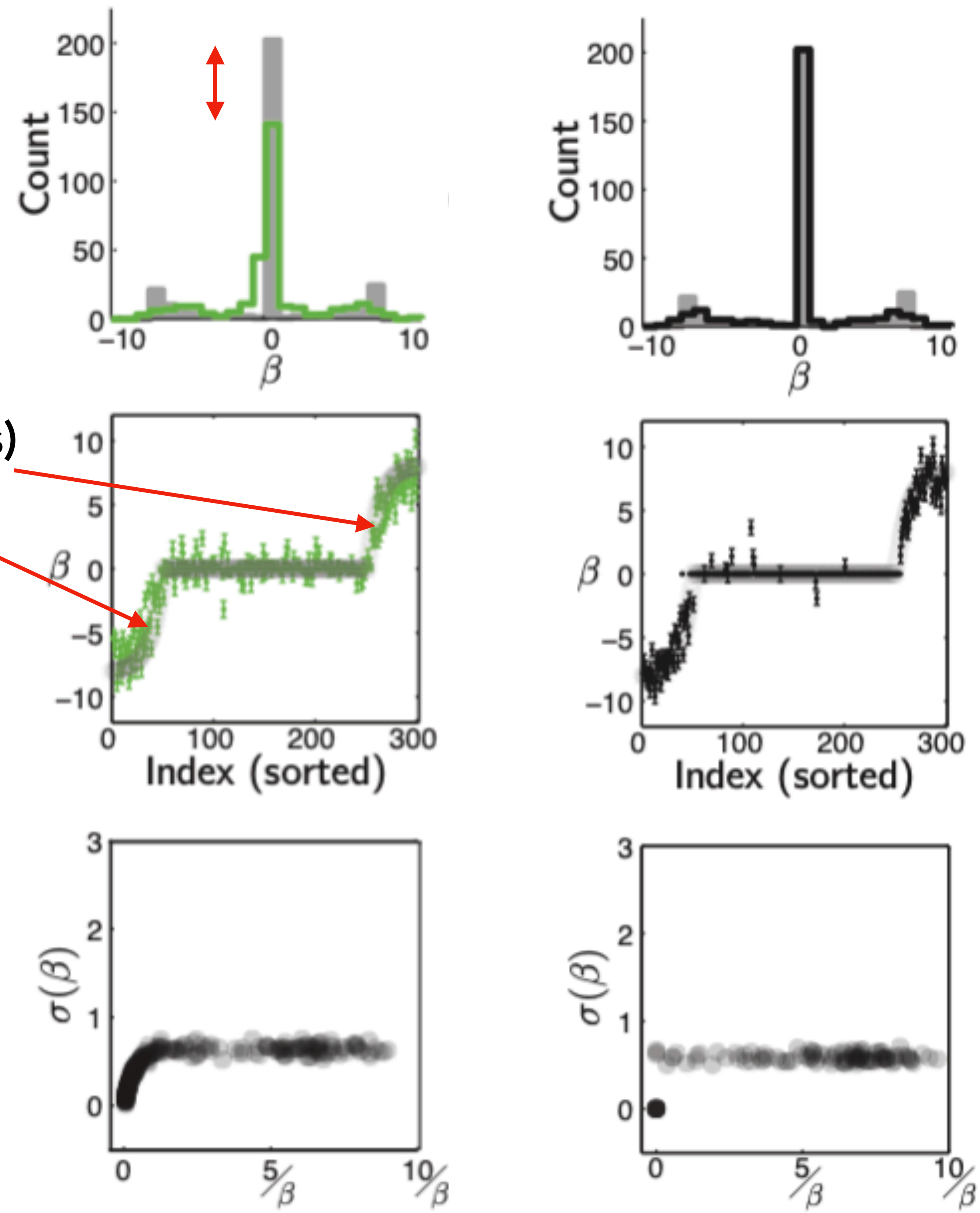
inaccurate selection

precise selection

shrinkage (bias)

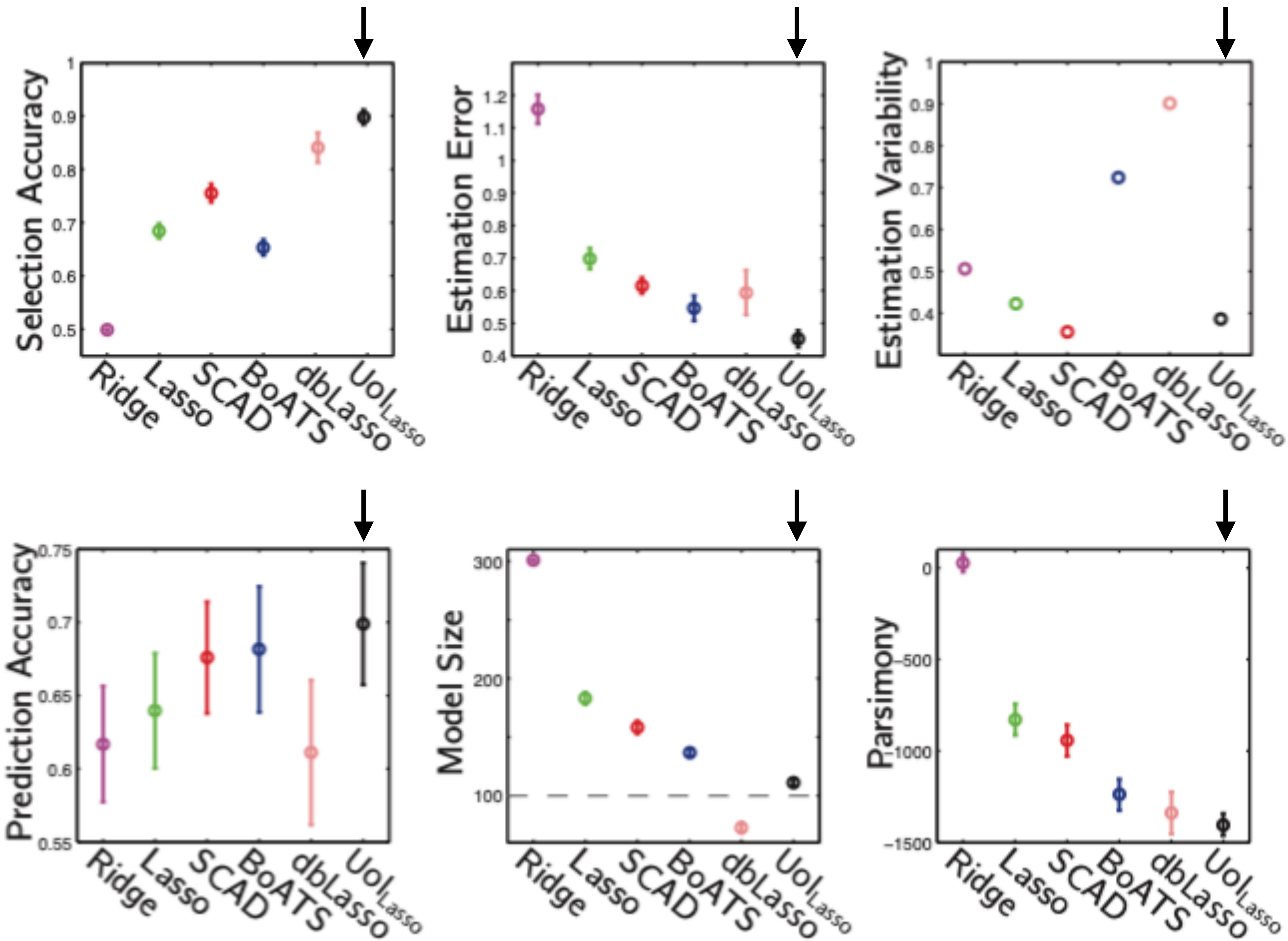
nearly unbiased

stable

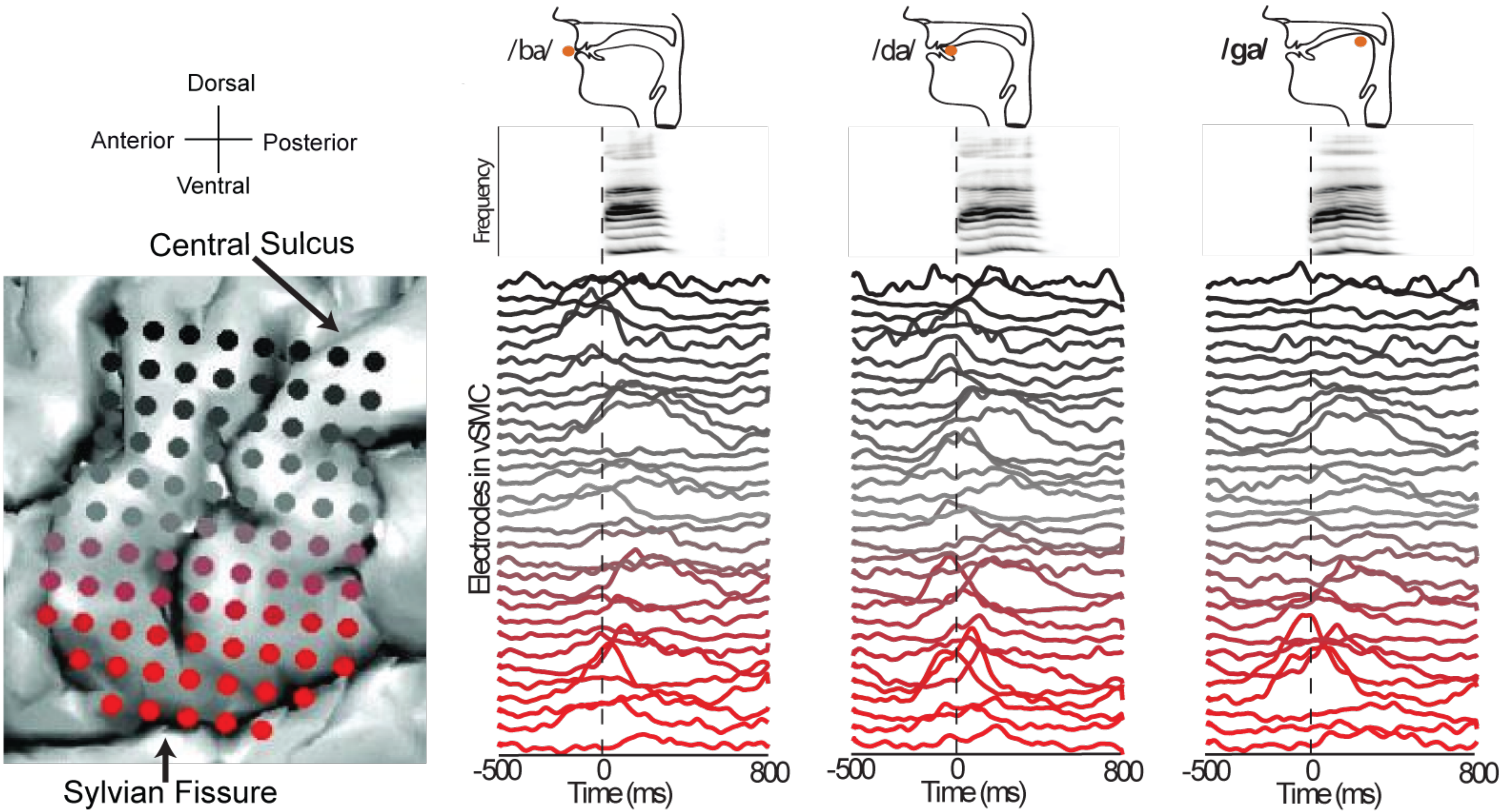


Linear Model
300 Features
100 Non-zero

Results on Synthetic Data



Electro-corticography Recordings from Human vSMC



Livezey et al., PLoS Comp. Bio (2019)

$\text{UoI}_{\text{Lasso}}$ Produces Sparse and Interpretable Functional Networks

$$n_i = \beta_{i0} + \sum_{j \neq i} \beta_{ij} n_j$$

$$n_1 = \beta_{10} + \beta_{11} n_2 + \cdots + \beta_{16} n_6$$

$$n_2 = \beta_{20} + \beta_{21} n_1 + \cdots + \beta_{26} n_6$$

$$\vdots$$

$$n_6 = \beta_{60} + \beta_{61} n_1 + \cdots + \beta_{65} n_5$$

/ba/

Xie & Huang, Ann. Stat., (2009)
Bouchard et al., NeurIPS (2017)

$\text{UoI}_{\text{Lasso}}$ Produces Sparse and Interpretable Functional Networks

$$n_i = \beta_{i0} + \sum_{j \neq i} \beta_{ij} n_j$$

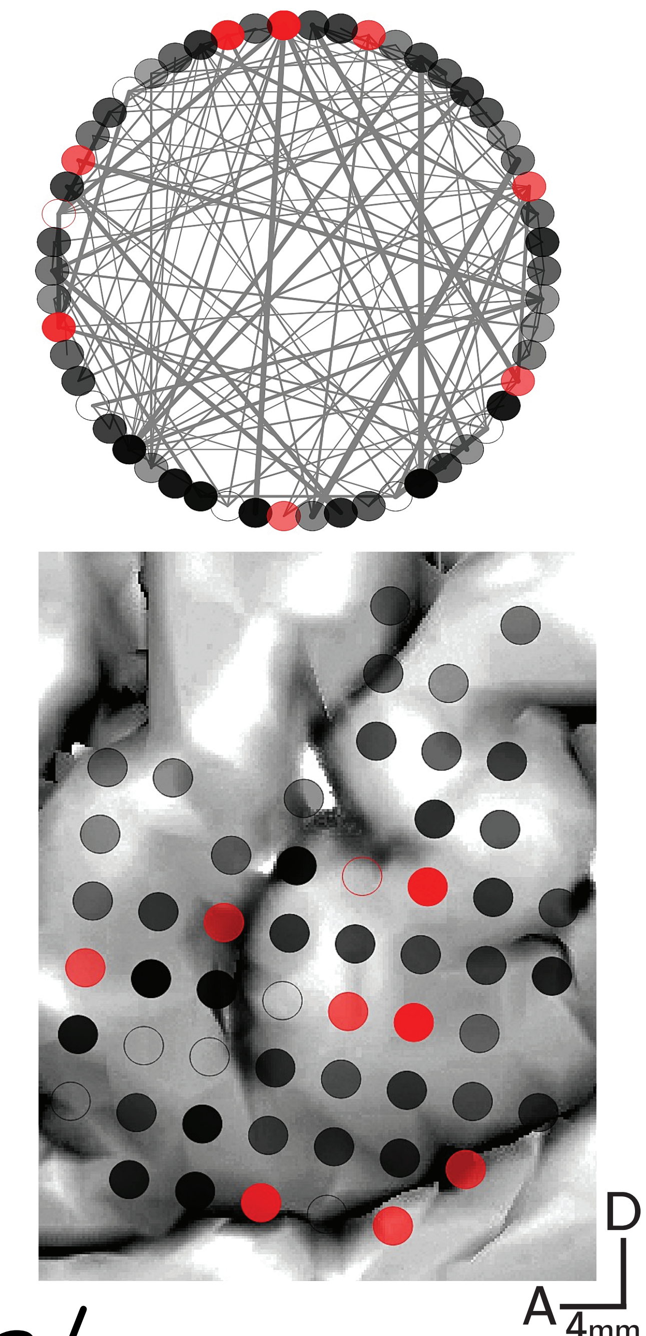
$$n_1 = \beta_{10} + \beta_{11} n_2 + \cdots + \beta_{16} n_6$$

$$n_2 = \beta_{20} + \beta_{21} n_1 + \cdots + \beta_{26} n_6$$

$$\vdots$$

$$n_6 = \beta_{60} + \beta_{61} n_1 + \cdots + \beta_{65} n_5$$

SCAD



/ba/

Xie & Huang, Ann. Stat., (2009)
Bouchard et al., NeurIPS (2017)

$\text{UoI}_{\text{Lasso}}$ Produces Sparse and Interpretable Functional Networks

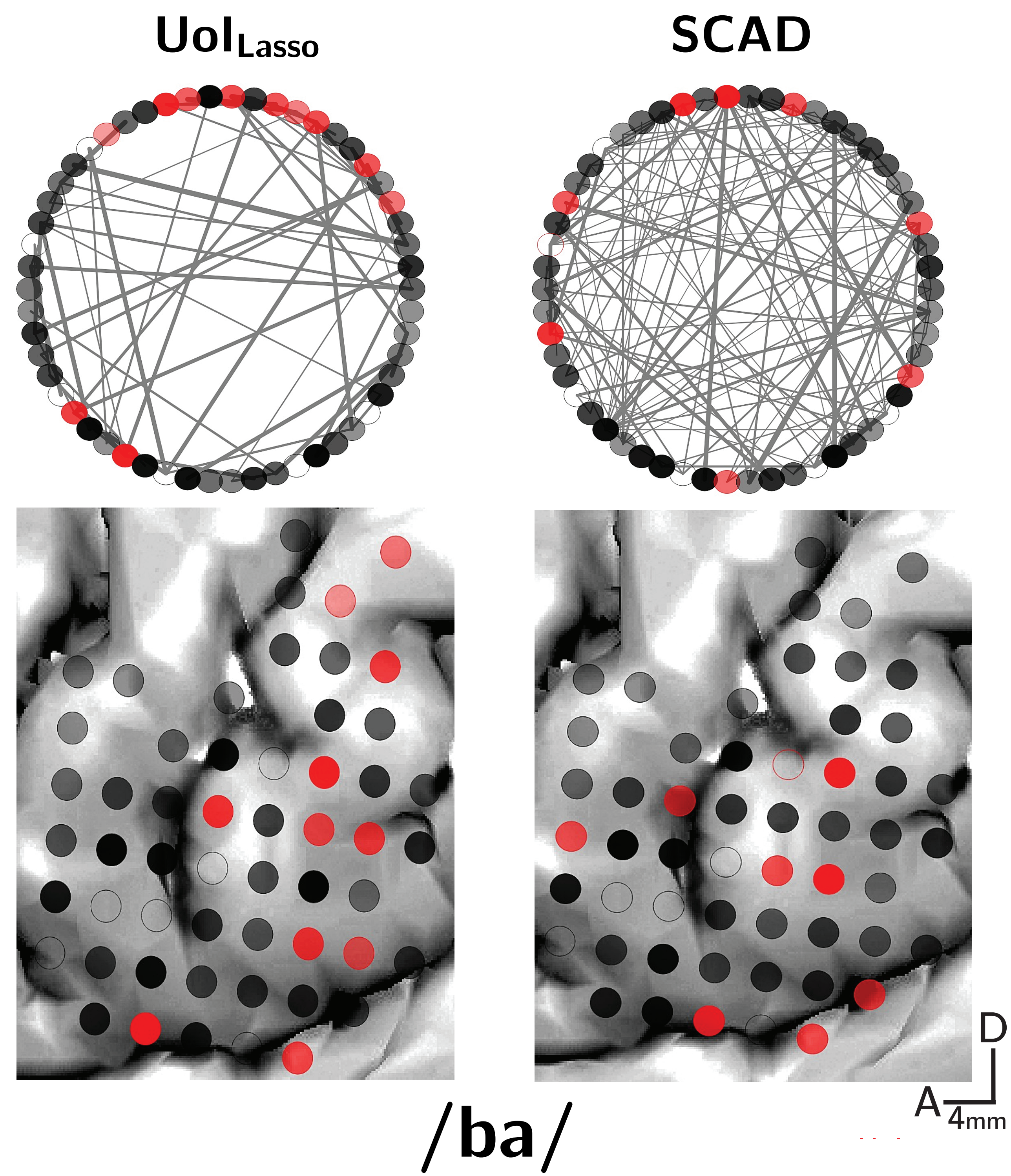
$$n_i = \beta_{i0} + \sum_{j \neq i} \beta_{ij} n_j$$

$$n_1 = \beta_{10} + \beta_{11} n_2 + \cdots + \beta_{16} n_6$$

$$n_2 = \beta_{20} + \beta_{21} n_1 + \cdots + \beta_{26} n_6$$

$$\vdots$$

$$n_6 = \beta_{60} + \beta_{61} n_1 + \cdots + \beta_{65} n_5$$



Xie & Huang, Ann. Stat., (2009)

Bouchard et al., NeurIPS (2017)

$\text{UoI}_{\text{Lasso}}$ Produces Sparse and Interpretable Functional Networks

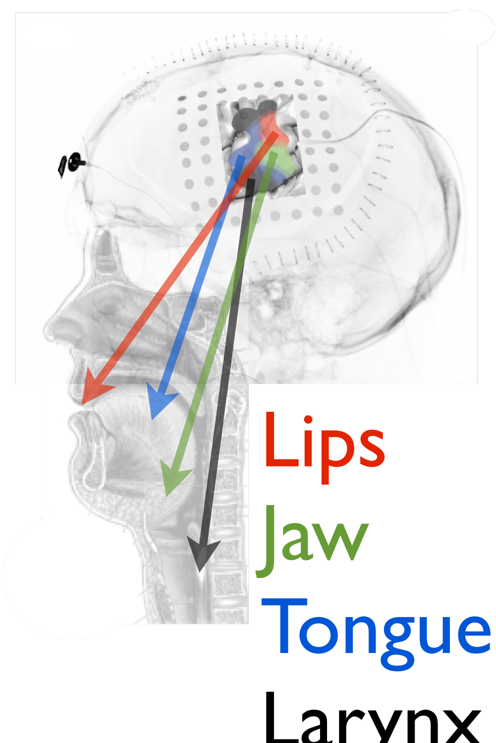
$$n_i = \beta_{i0} + \sum_{j \neq i} \beta_{ij} n_j$$

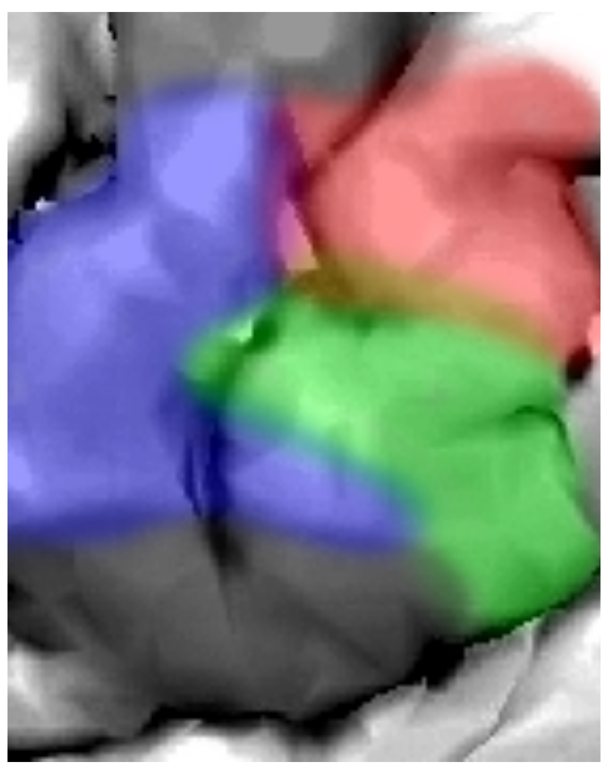
$$n_1 = \beta_{10} + \beta_{11} n_2 + \cdots + \beta_{16} n_6$$

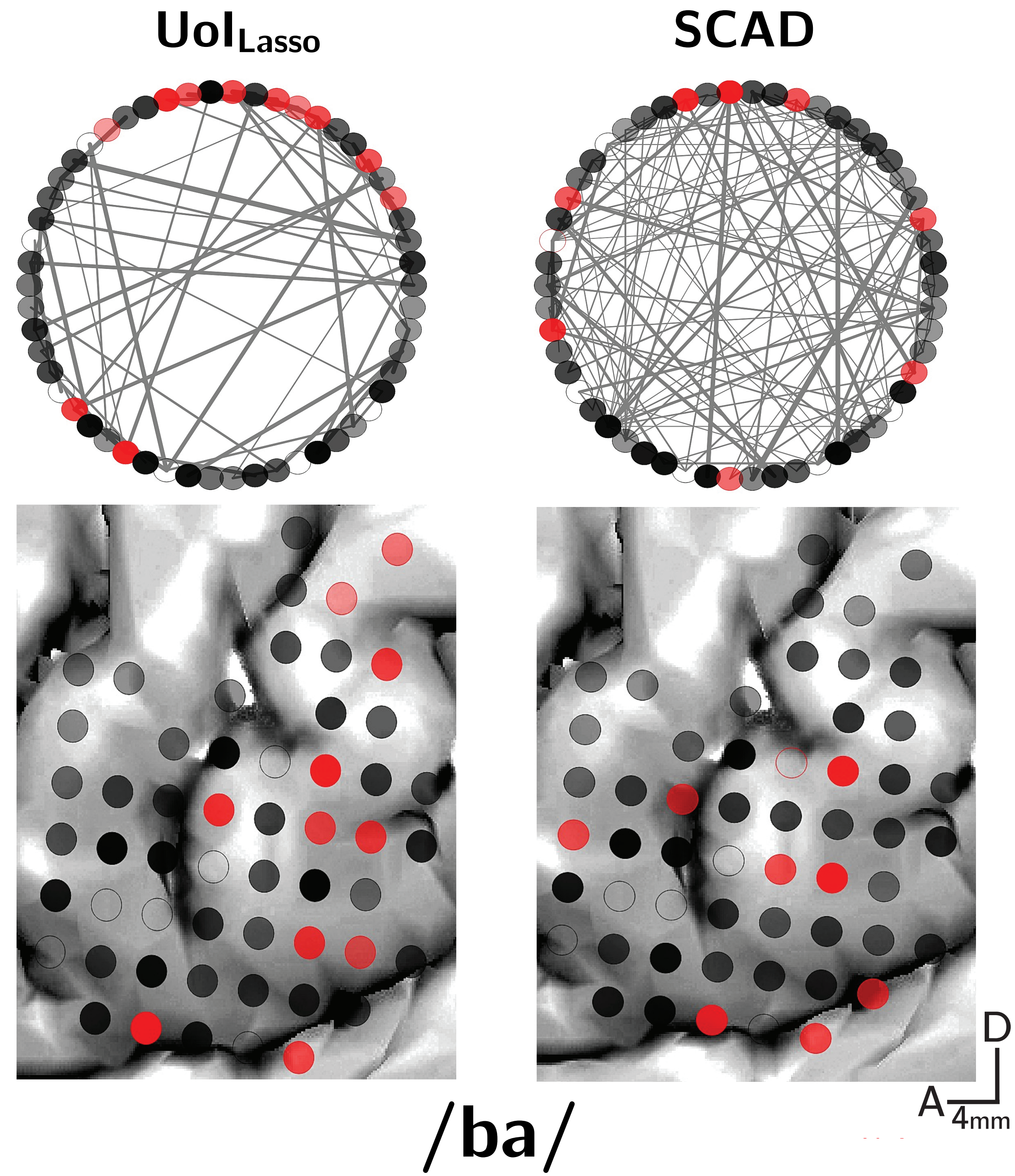
$$n_2 = \beta_{20} + \beta_{21} n_1 + \cdots + \beta_{26} n_6$$

$$\vdots$$

$$n_6 = \beta_{60} + \beta_{61} n_1 + \cdots + \beta_{65} n_5$$

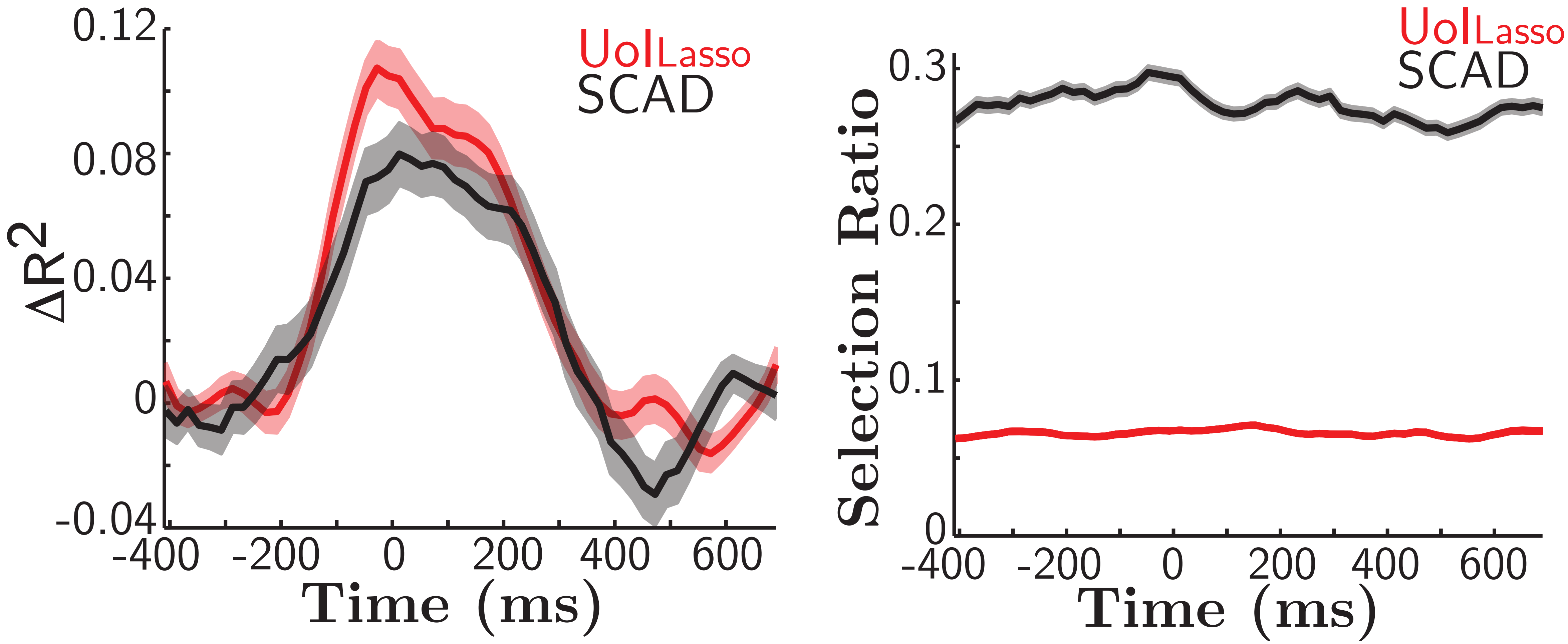




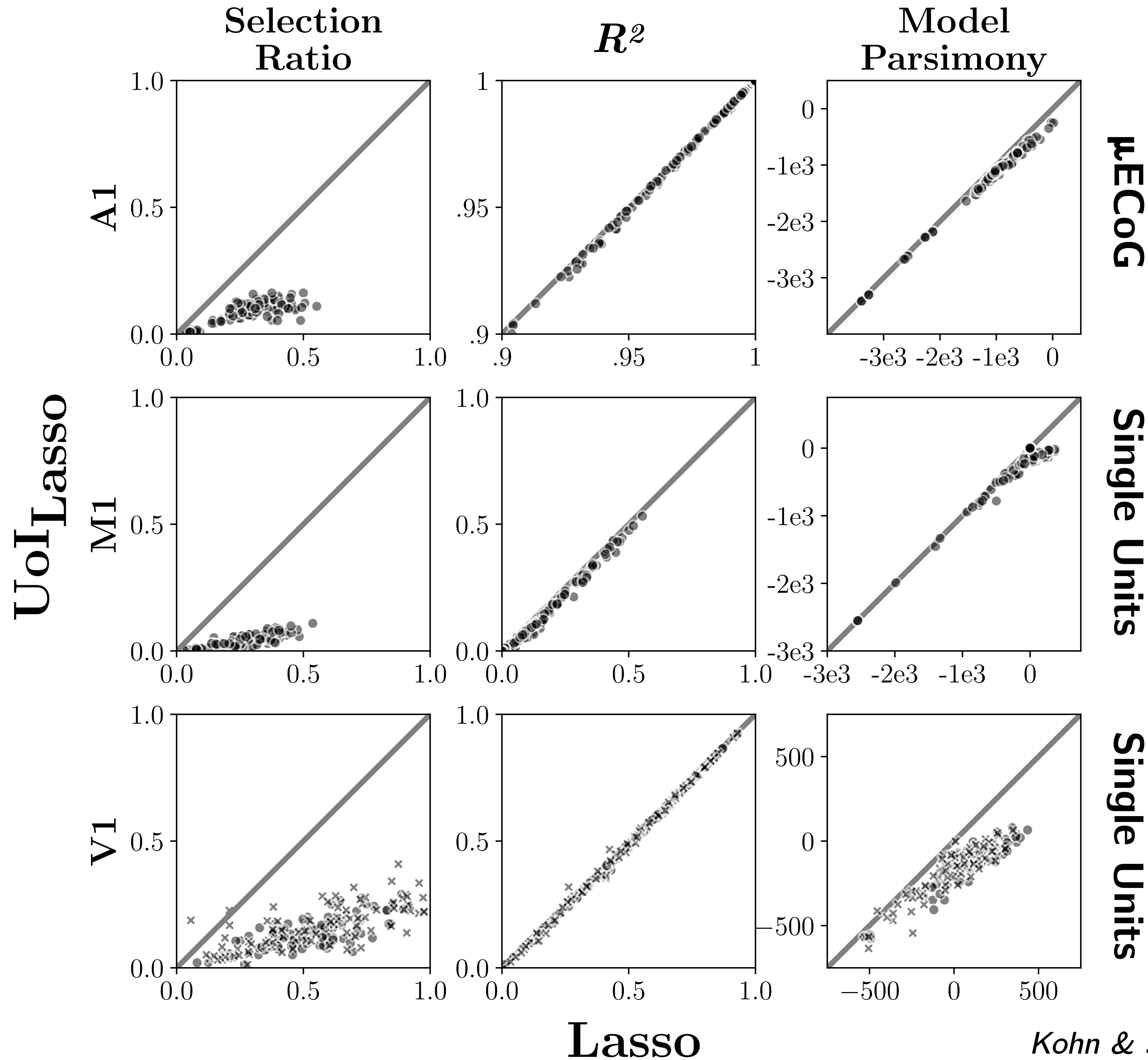


Xie & Huang, Ann. Stat., (2009)
Bouchard et al., NeurIPS (2017)

$\text{UoI}_{\text{Lasso}}$ Produces Sparse and Interpretable Functional Networks



Sparse and Predictive Functional Coupling Networks Across Brain Regions



Kohn & Smith, CRCNS.org (2016)
O'Doherty et al., Zenodo (2017)
Bouchard et al. (2019)

PyUoI: Union of Intersections in Python

github.com/BouchardLab/PyUoI

Neural Systems and Data Science Group

bouchardlab.lbl.gov

Collaborators

Center for Neural Engineering & Prostheses
University of California · Berkeley / San Francisco

Berkeley Lab

Berkeley
University of California
Lawrence Berkeley National Laboratory

UCSF Center for Integrative Neuroscience

Redwood Center for Theoretical Neuroscience

Helen Wills Neuroscience Institute
University of California, Berkeley

Funding

U.S. Department of Energy

Bill & Melinda Gates foundation

NIH National Institutes of Health
Turning Discovery Into Health

The Kavli Foundation

DARPA

Department of Defense United States of America

Interpretable Networks in Ventral Sensory-Motor Cortex

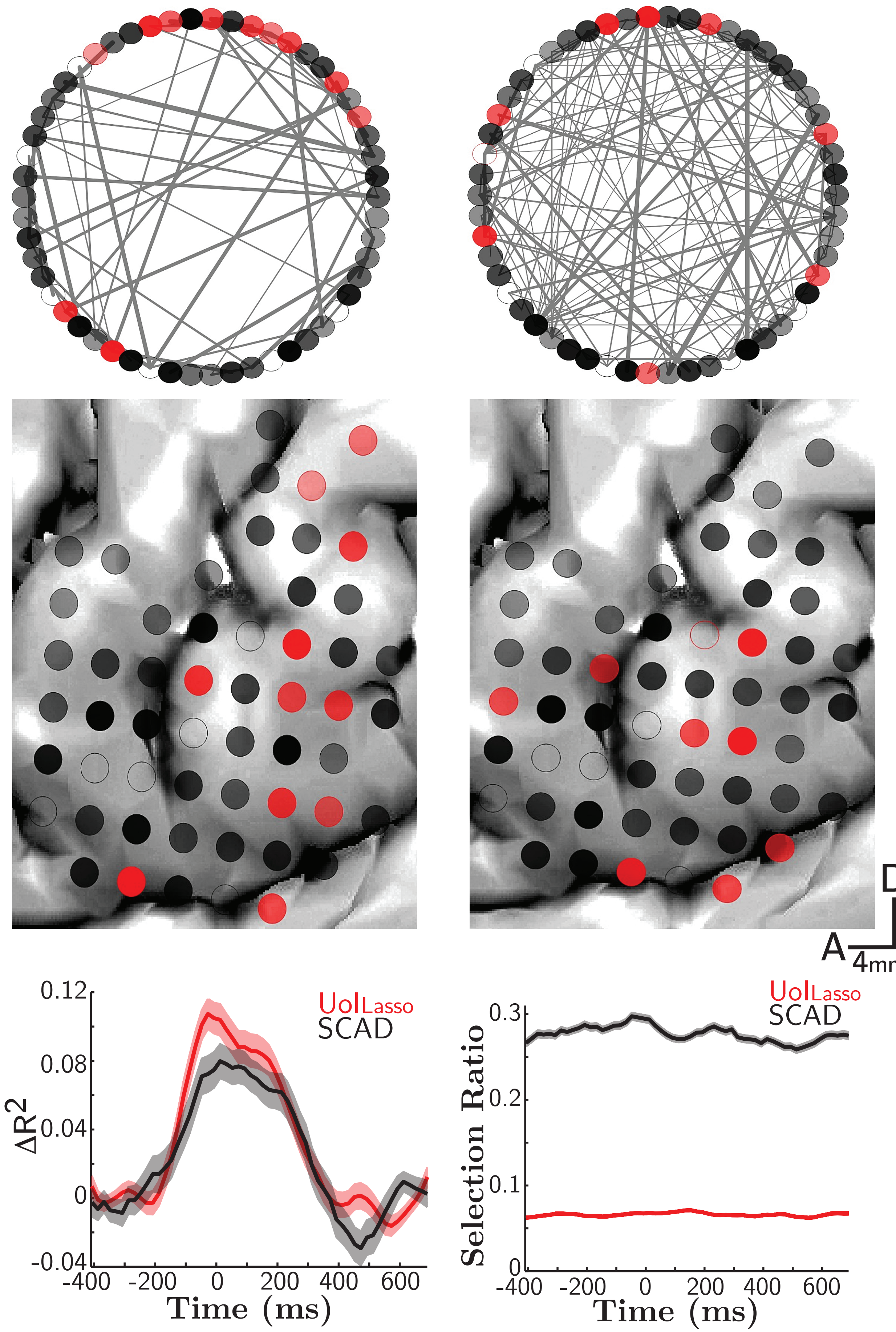
$$n_i = \beta_{i0} + \sum_{j \neq i} \beta_{ij} n_j$$

$$n_1 = \beta_{10} + \beta_{11} n_2 + \cdots + \beta_{16} n_6$$

$$n_2 = \beta_{20} + \beta_{21} n_1 + \cdots + \beta_{26} n_6$$

$$\vdots$$

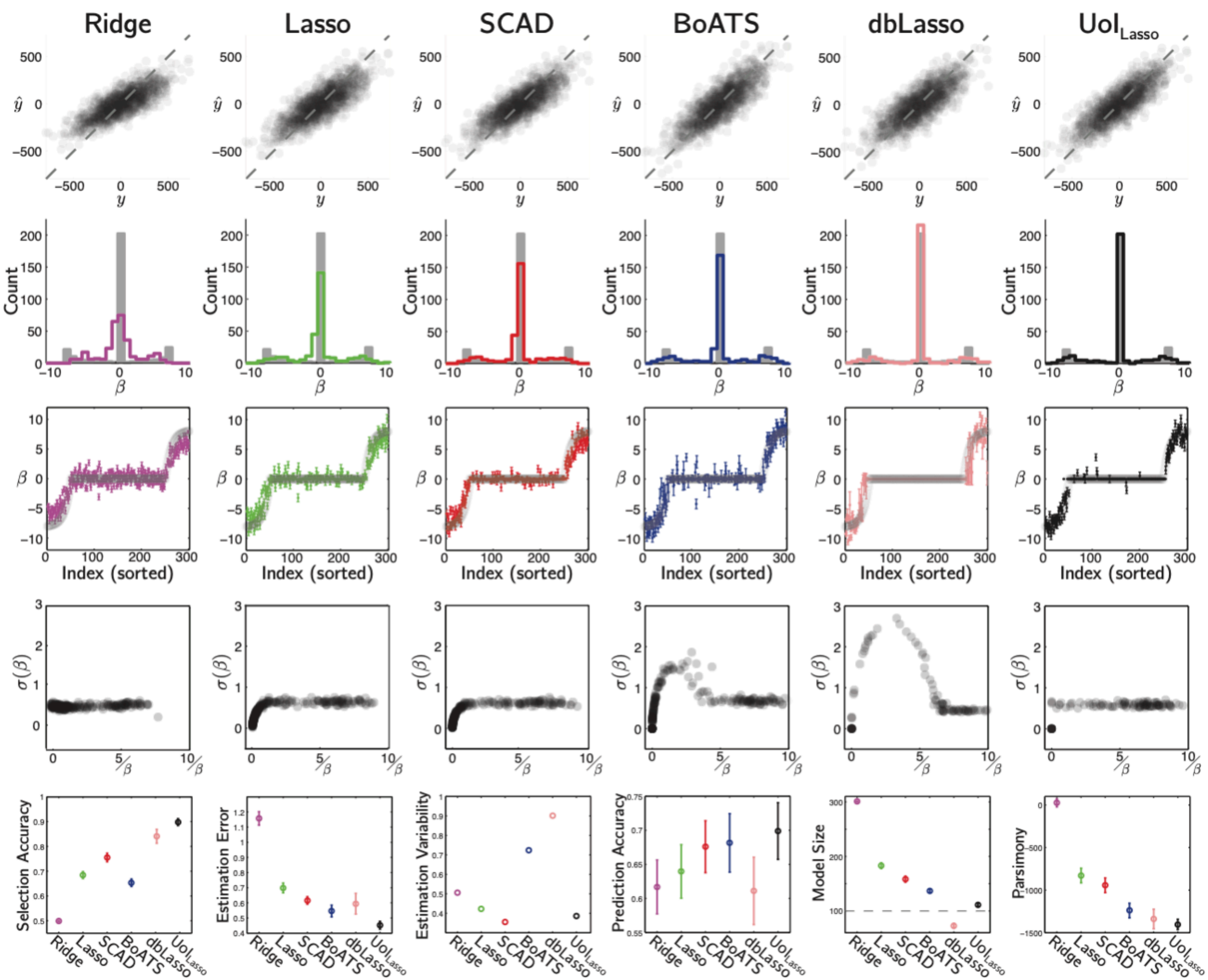
$$n_6 = \beta_{60} + \beta_{61} n_1 + \cdots + \beta_{65} n_5$$



Bouchard et al., NeurIPS (2017)

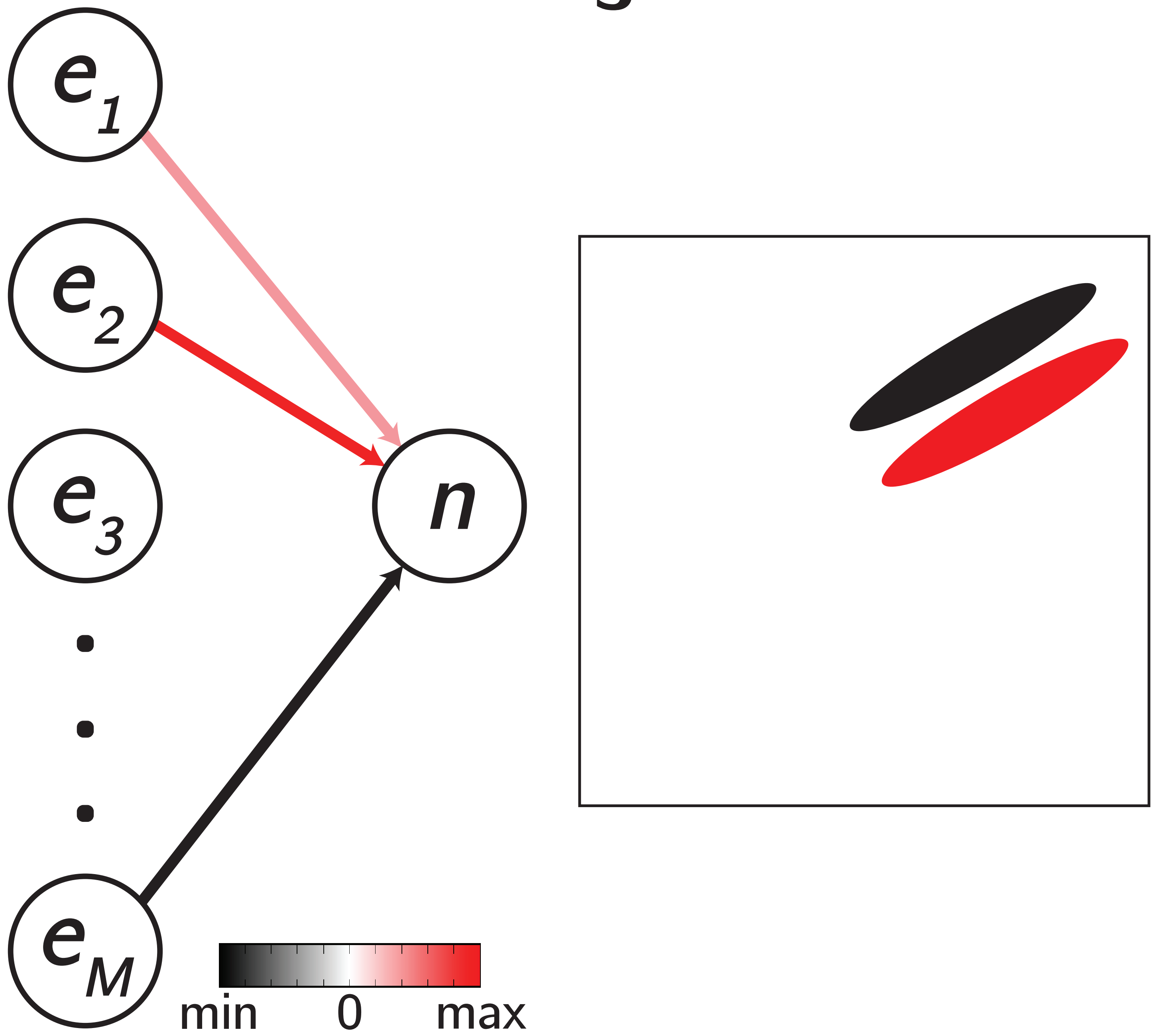
$\text{UoI}_{\text{Lasso}}$

Validation on Synthetic Data



Neuroscience Applications

Encoding Models

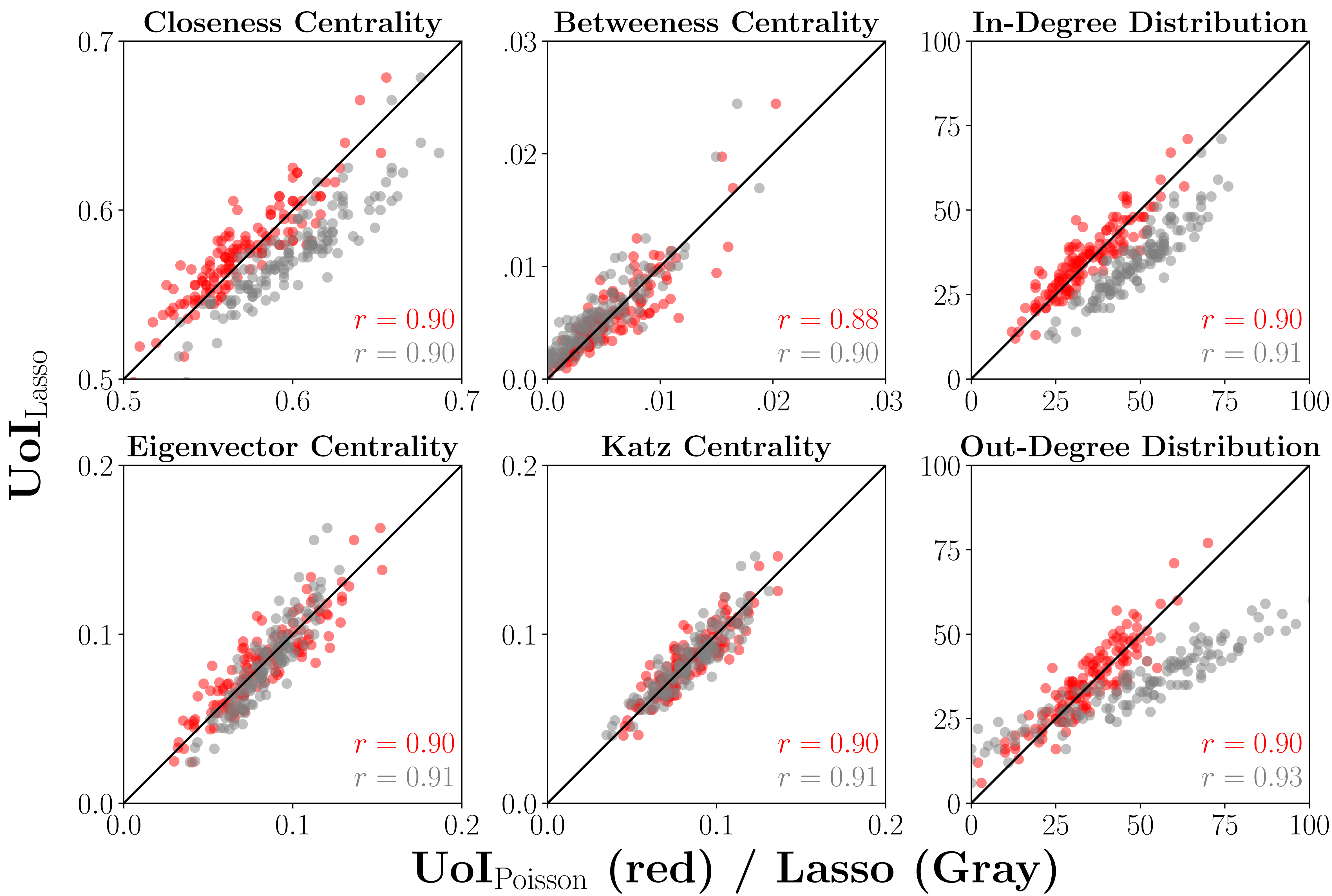


Applications:

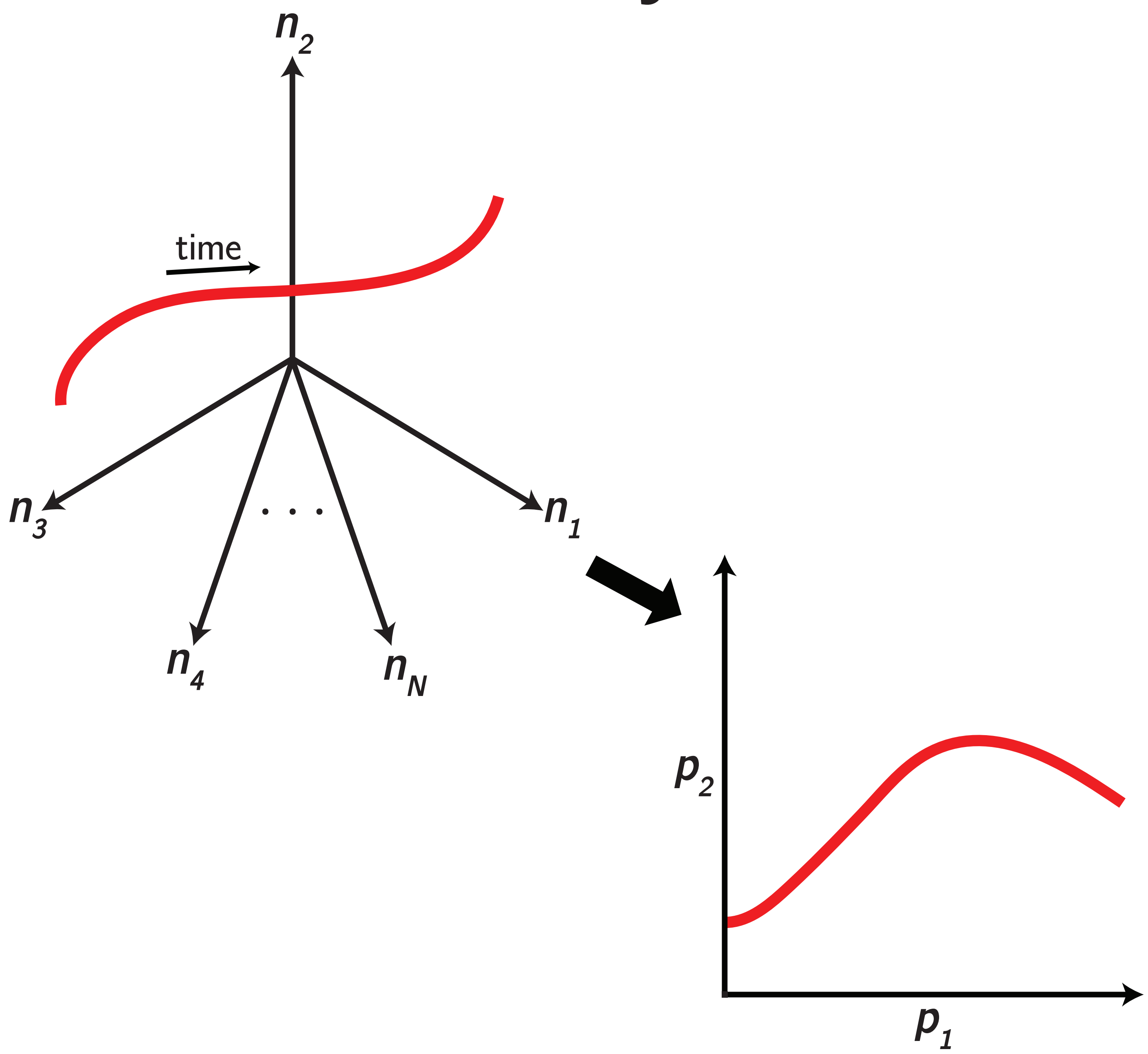
Retinal STRF

A1 Tuning Curves

$UoI_{Poisson}$ extracts similar networks as UoI_{Lasso}



Dimensionality Reduction



Column Subset Selection / CUR Decomposition

$$\mathbf{A} \approx \mathbf{CUR}$$

columns (C) | "carefully constructed" (U) | **rows** (R)

Mahoney, Drineas, PNAS (2009)

Column Subset Selection / CUR Decomposition

$$\mathbf{A} \approx \mathbf{CUR}$$

columns

rows

“carefully constructed”

Mahoney, Drineas, PNAS (2009)

Column Subset Selection / CUR Decomposition

$$\mathbf{A} \approx \mathbf{CUR}$$

columns

"carefully constructed"

rows

$\textsc{ColumnSelect}(A, k, c)$

Mahoney, Drineas, PNAS (2009)

UoI_{CSS}

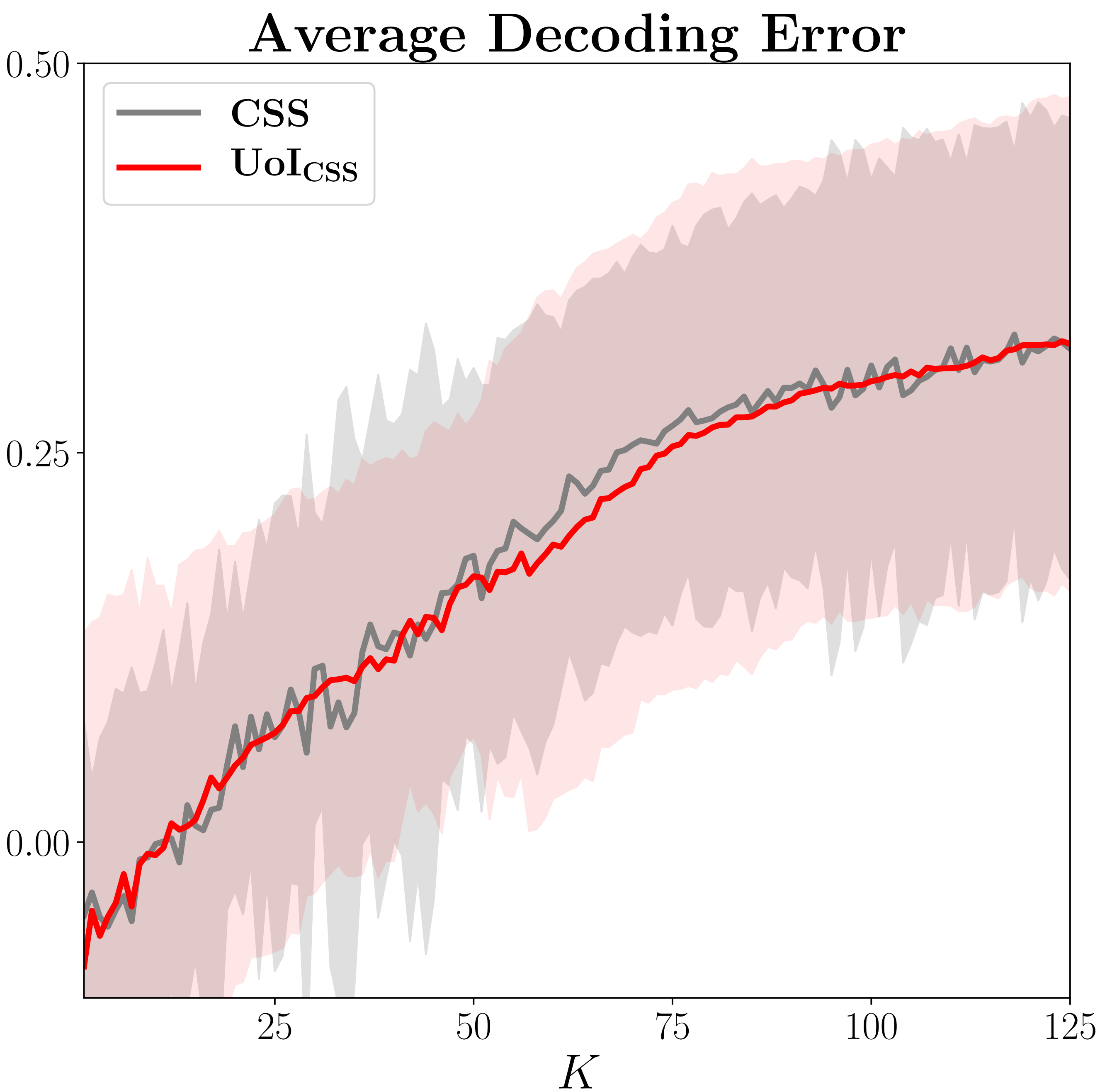
Algorithm 1 UoI_{CSS} (A, K, c, N_B).

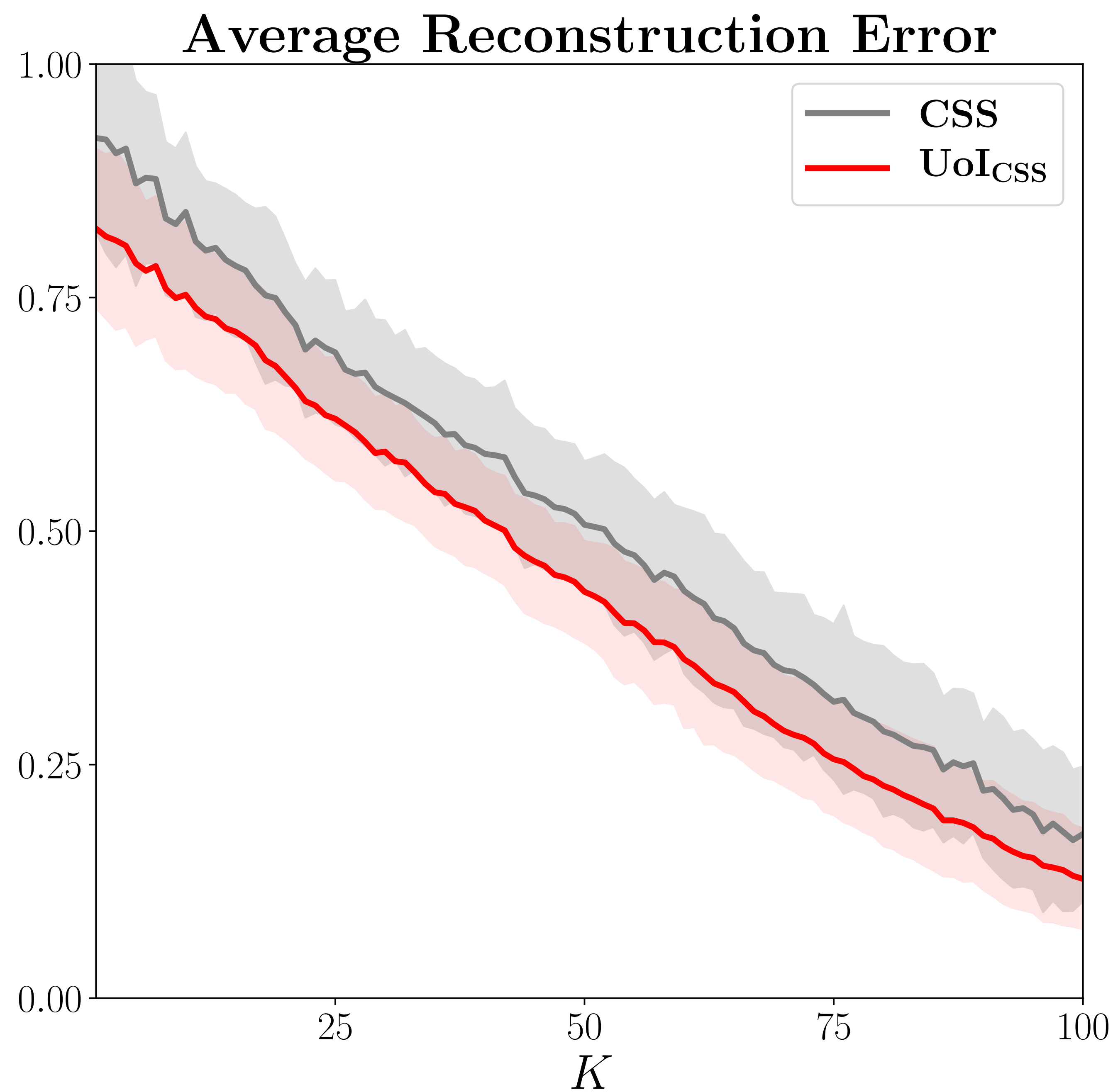
Input: A, $N \times M$ design matrix
$K = \{k_i\}_{i=1}^{N_k}$, the ranks of the SVD decomposition to union over
$c(k)$, the expected number of columns for each rank k

```
for j = 1 to N_B do                                      ▷ iterate over bootstraps
    Generate resample A^j of the data matrix A
    for k_i in K do                                      ▷ iterate over ranks
        C_i ← ColumnSelect(A^j, k_i, c(k_i))
C ← ⋃_{i=1}^{N_k} ( ⋂_{j=1}^{N_B} C_j )                  ▷ union of intersections
return C
```

$$C \leftarrow \bigcup_{i=1}^{N_k} \left(\bigcap_{j=1}^{N_B} C_j \right)$$

CSS in M1 Reaching Data





$$x, y = \beta_0 + \sum_{i=1}^{T} \beta_i N_i$$

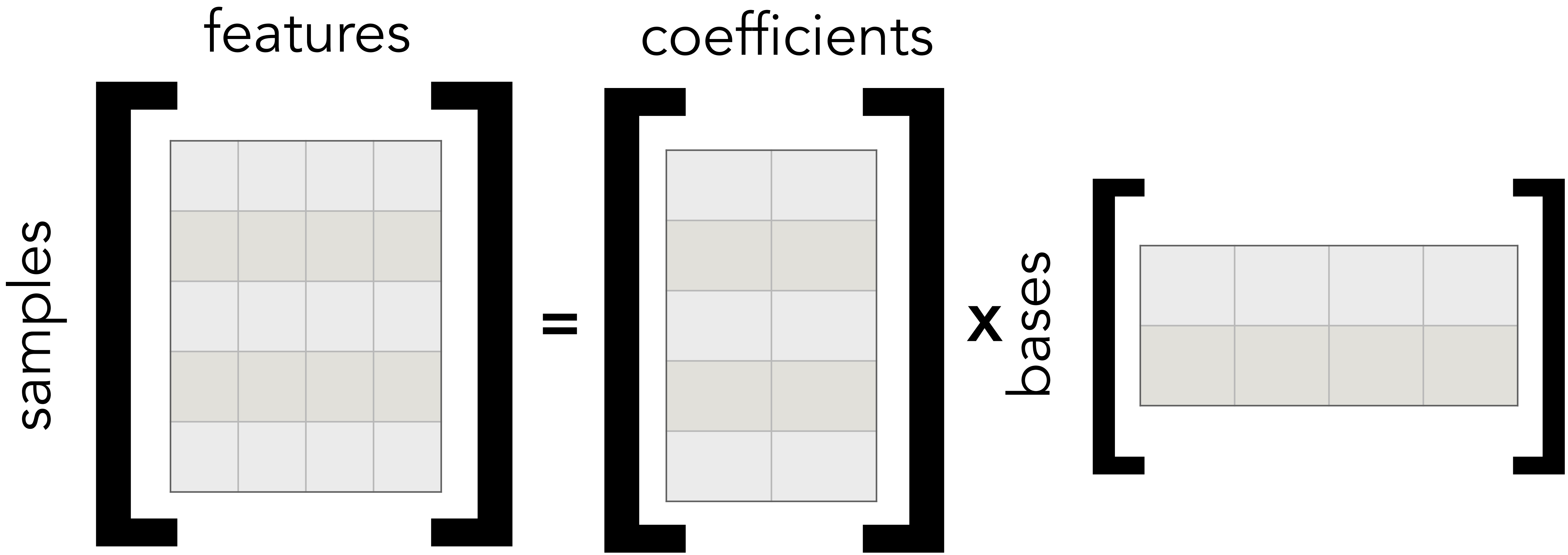
$$\text{sum}(A - A_c A_c^{\dagger} A)$$

Sabes Lab

UoI_{CSS}

Non-negative Matrix Factorization

$$\mathbf{A} \approx \mathbf{WH}$$



parts-based decomposition of the data

Ubaru et al., ICML-A (2017)

NMF in Auditory A1 Data