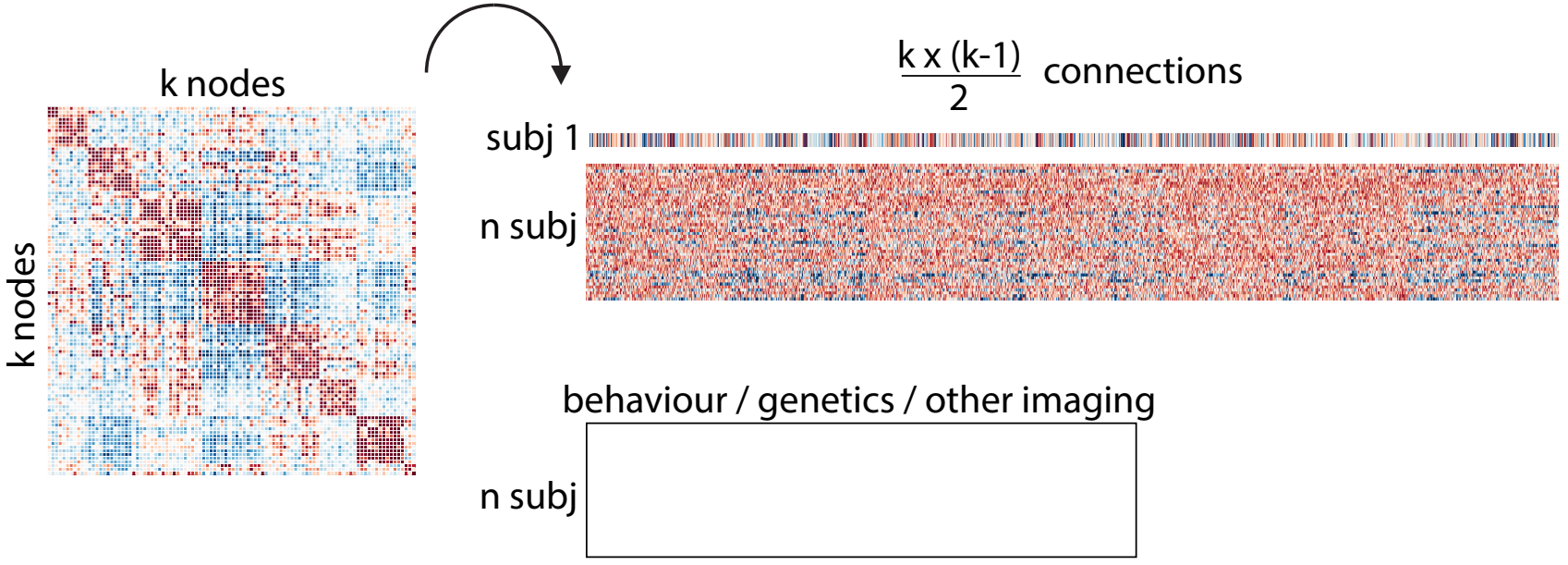
Associative statistics

Bratislav Mišić

NEUR 602 Week 4

October 6^{th} 2017

Why multivariate statistics?

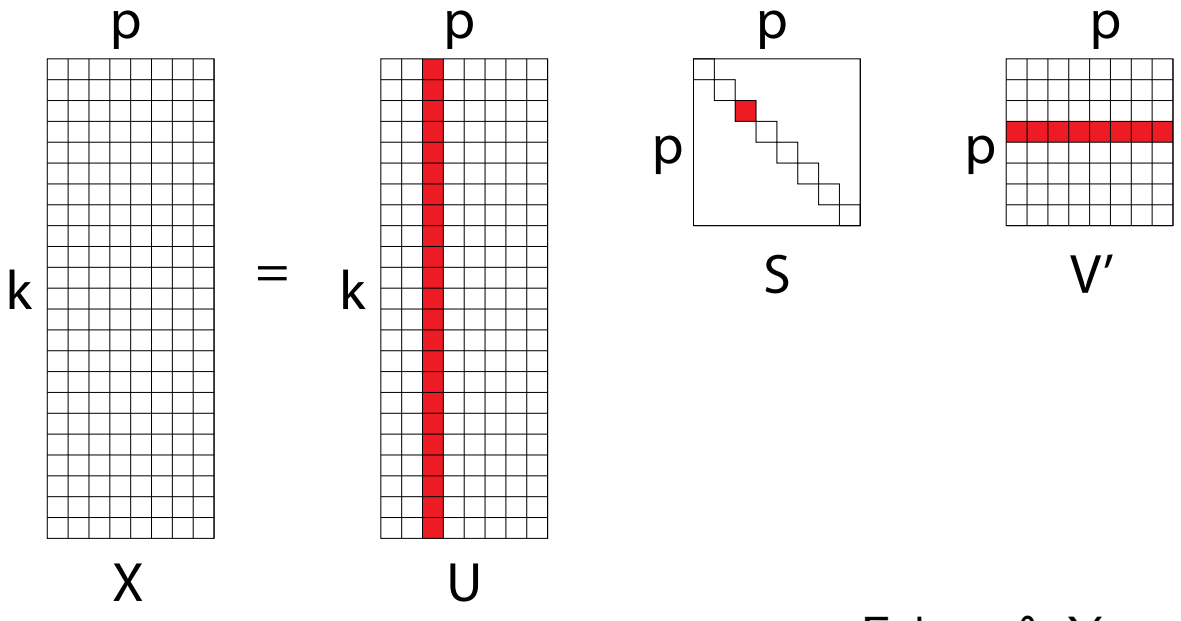


1. how to operationalize network property?
2. how to deal with with more variables than observations?
3. how to relate multiple data sets to one another?

Singular value decomposition

Spectral decomposition:
$\text{EIG}(\mathbf{X}'\mathbf{X}) = \mathbf{U}\mathbf{\Lambda}\mathbf{U}'$
$\text{EIG}(\mathbf{X}\mathbf{X}') = \mathbf{V}\mathbf{\Lambda}\mathbf{V}'$

Singular value decomposition:
$\text{SVD}(\mathbf{X}) = \mathbf{U}\mathbf{S}\mathbf{V}'$



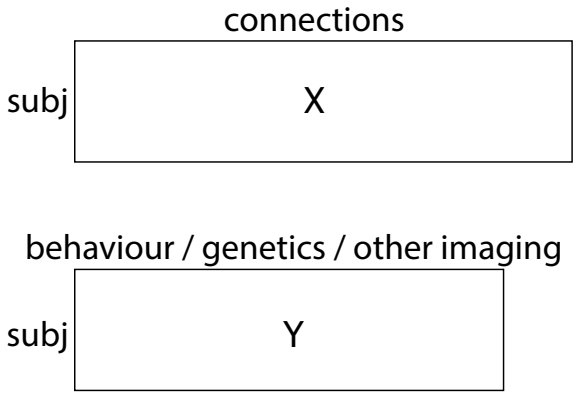
Eckart & Young (1936) *Psychometrika*

A family of techniques

PCA: SVD($\mathbf{X}$)

PLS: SVD($\mathbf{X}'\mathbf{Y}$)

CCA: SVD($(\mathbf{X}'\mathbf{X}')^{-1/2}(\mathbf{X}'\mathbf{Y})(\mathbf{Y}'\mathbf{Y})^{-1/2}$)

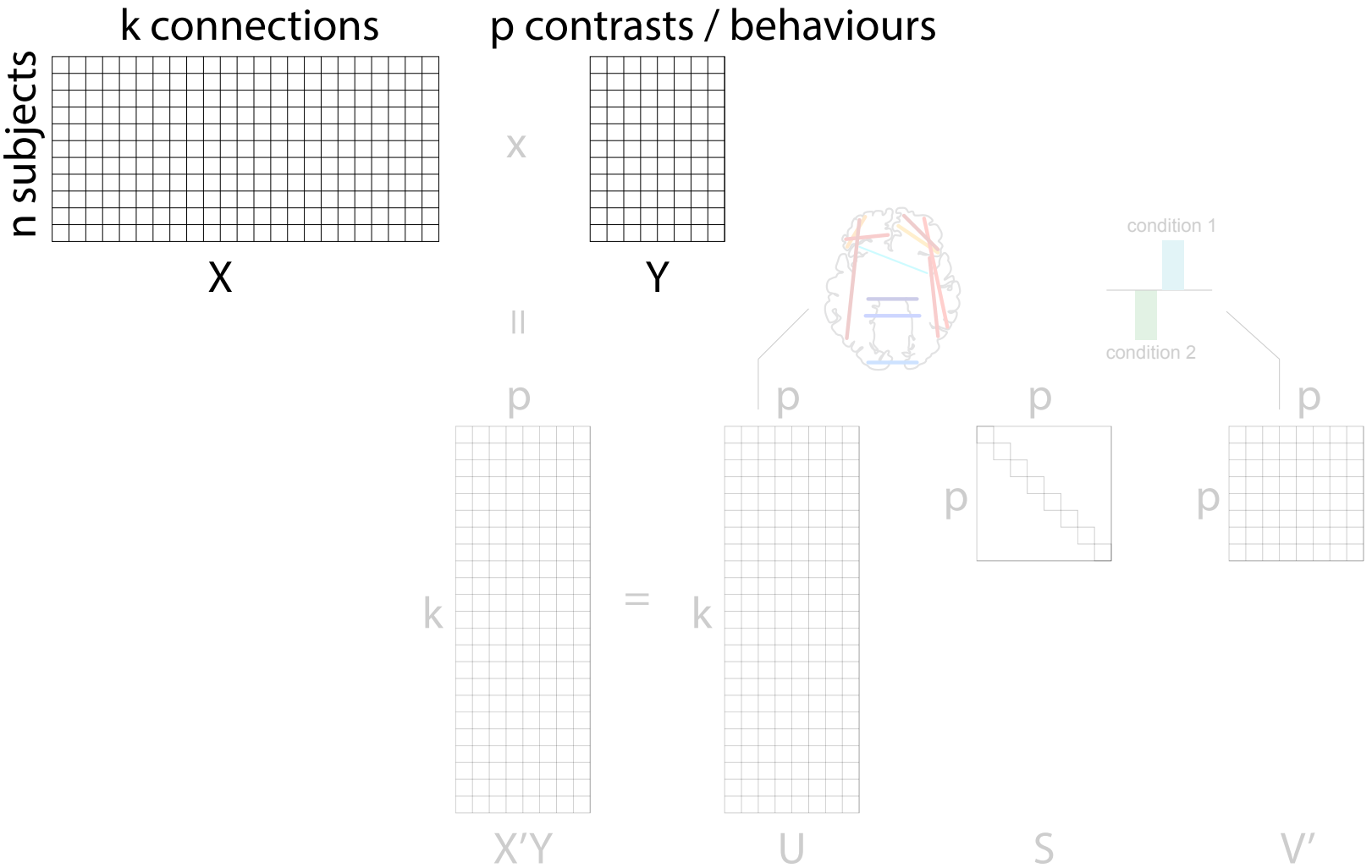


Worsley et al. (1997) *NeuroImage*

Tijl et al. (2005) *Handbook of Geometric Computing - Springer*

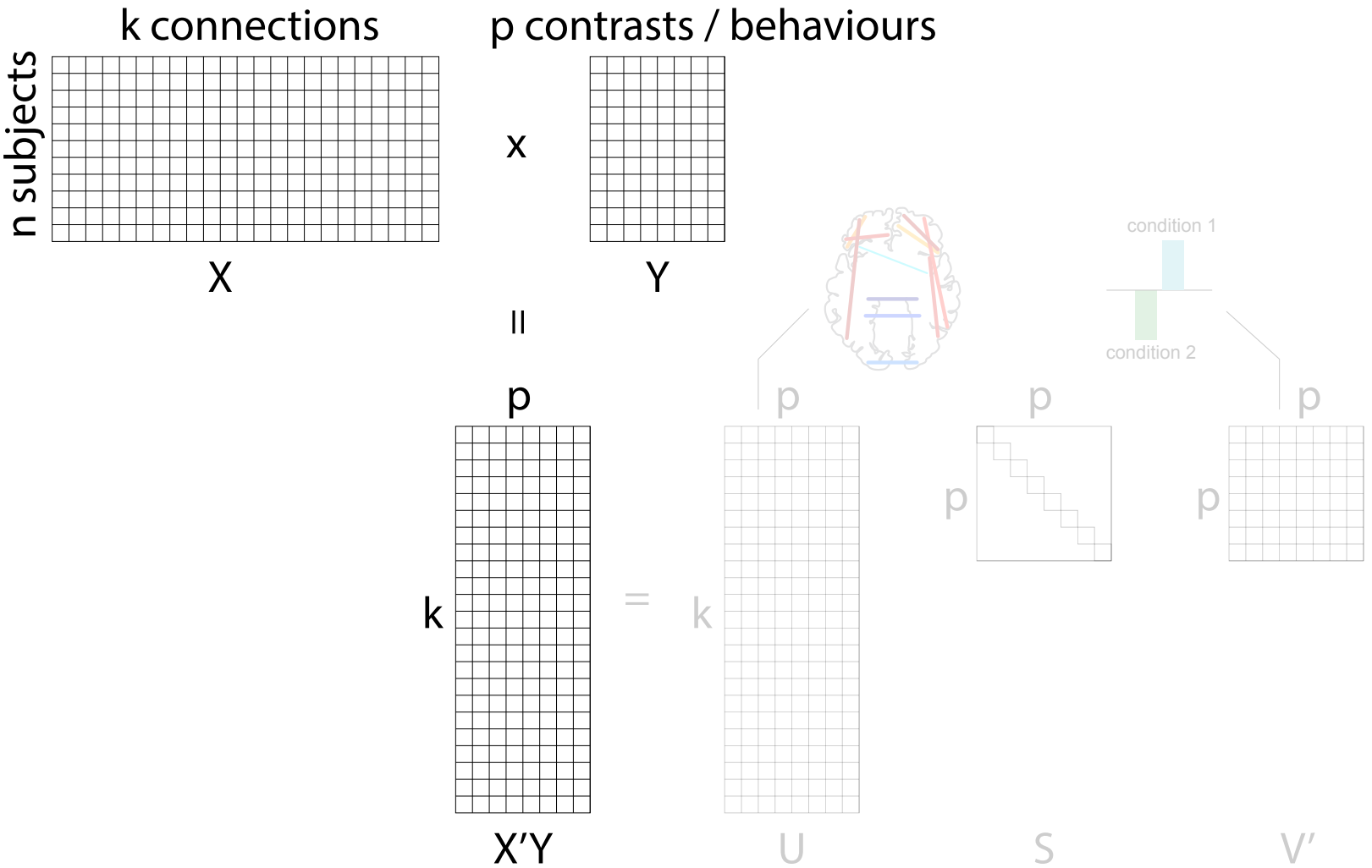
McIntosh & Mišić (2013) *Annu Rev Psychol*

Partial least squares (PLS)



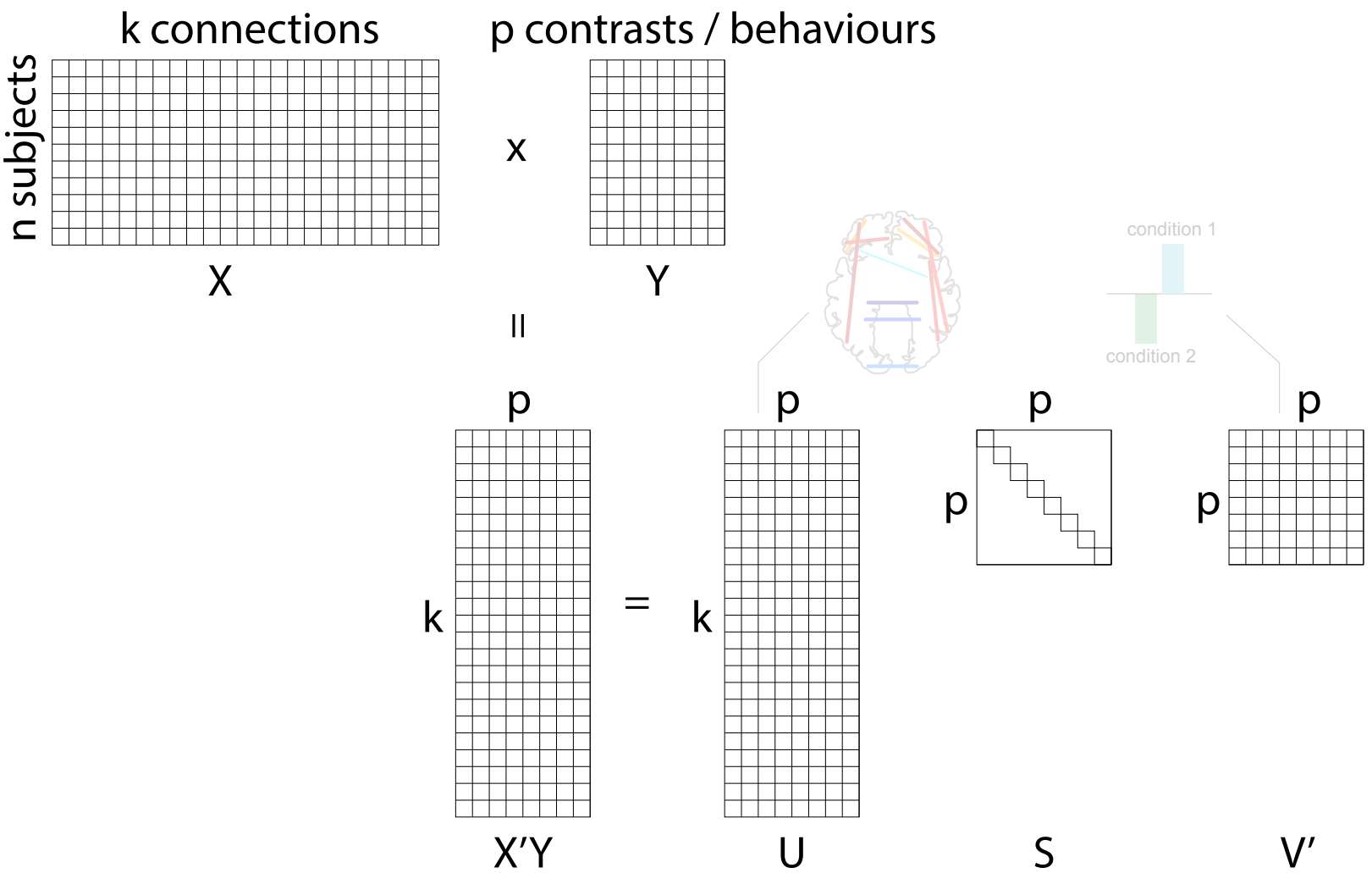
Wold et al. (1982, 1984) *Mathematics and statistics in chemistry*

Partial least squares (PLS)



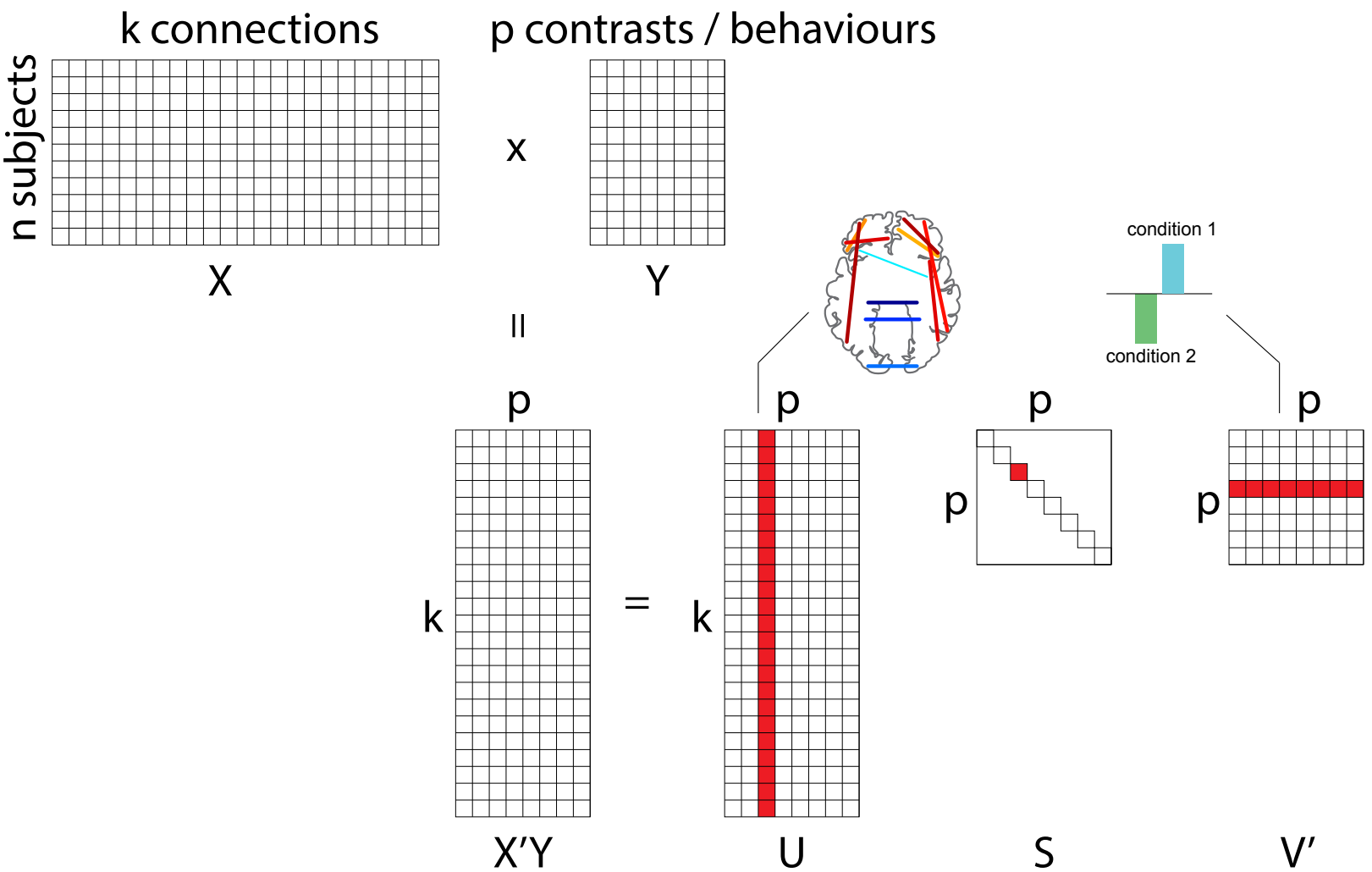
Wold et al. (1982, 1984) *Mathematics and statistics in chemistry*

Partial least squares (PLS)



Wold et al. (1982, 1984) *Mathematics and statistics in chemistry*

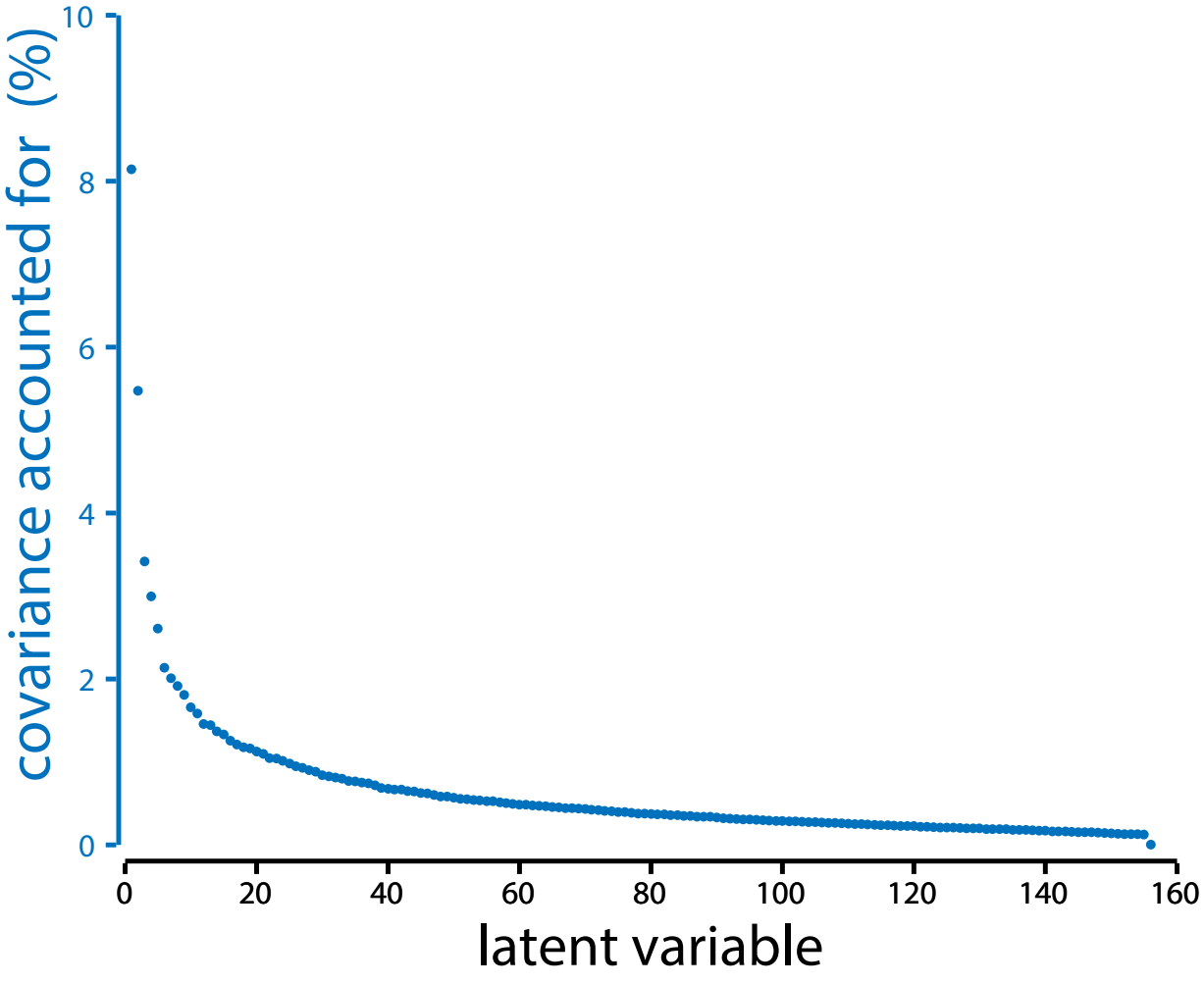
Partial least squares (PLS)



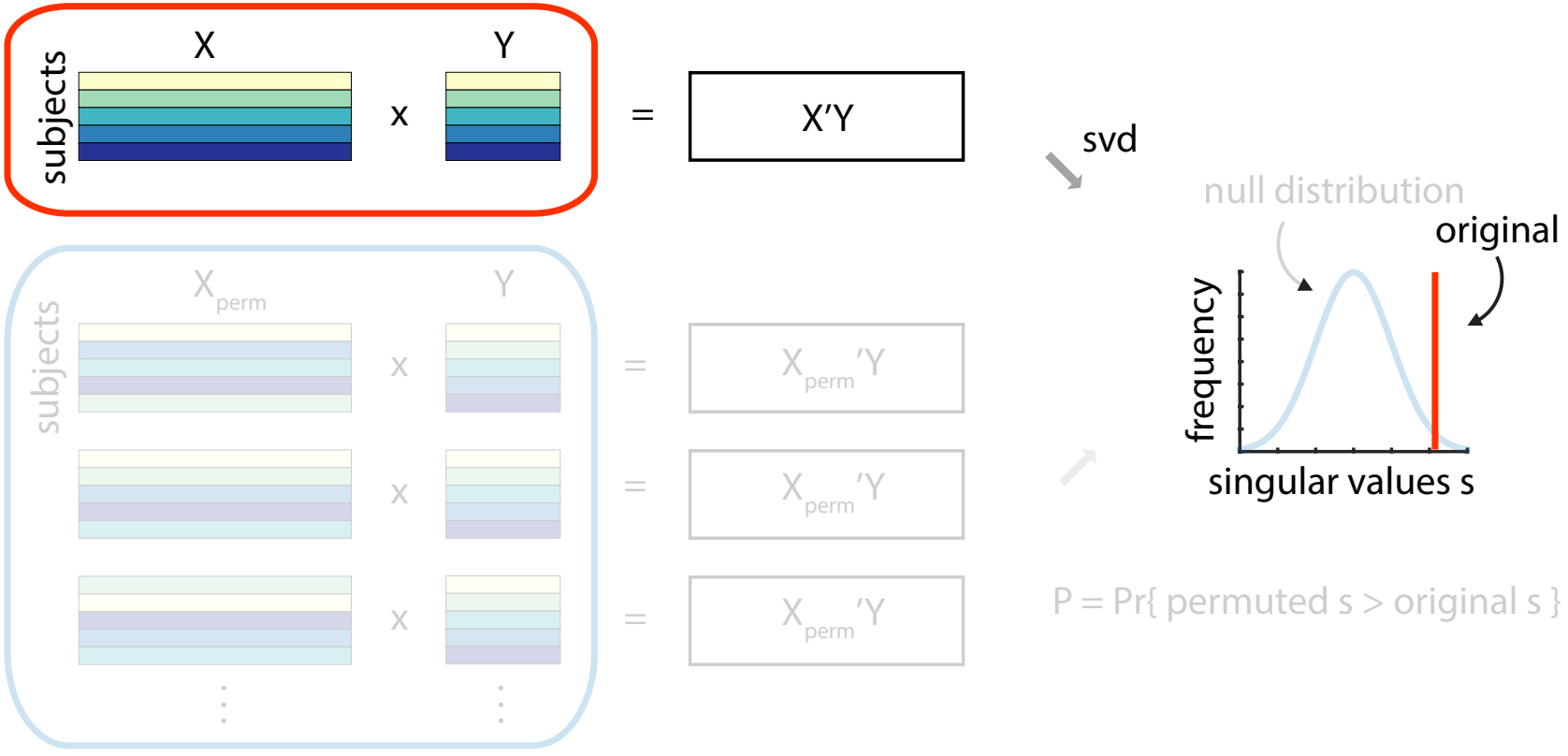
Wold et al. (1982, 1984) *Mathematics and statistics in chemistry*

How many components to retain?

$$\% \text{ covariance} = s_i^2 / \sum_j s_j^2$$



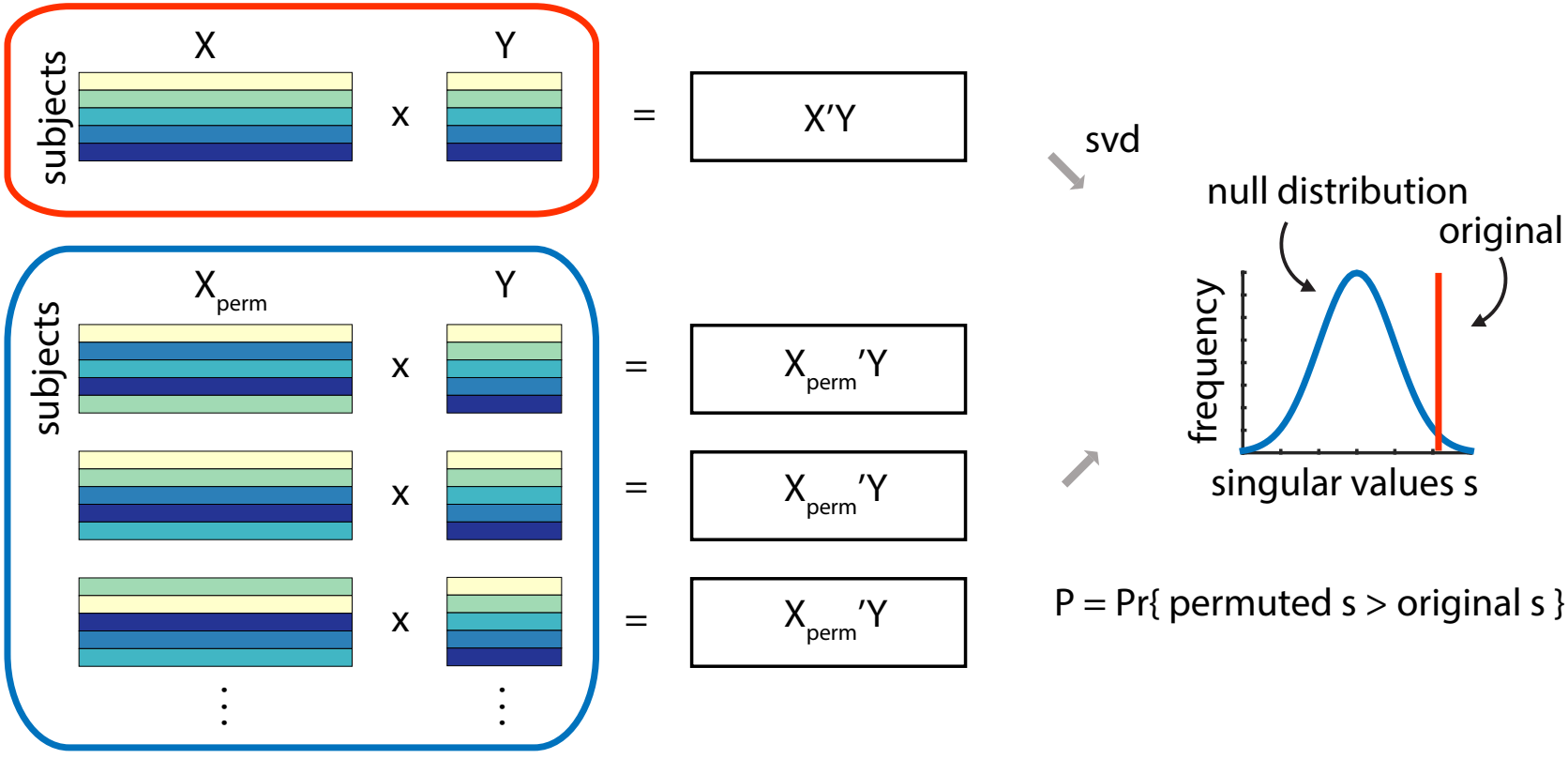
Statistical significance: permutation tests



Edgington (1965, 1969) *J Psychol*

McIntosh et al. (1996, 2004) *NeuroImage*

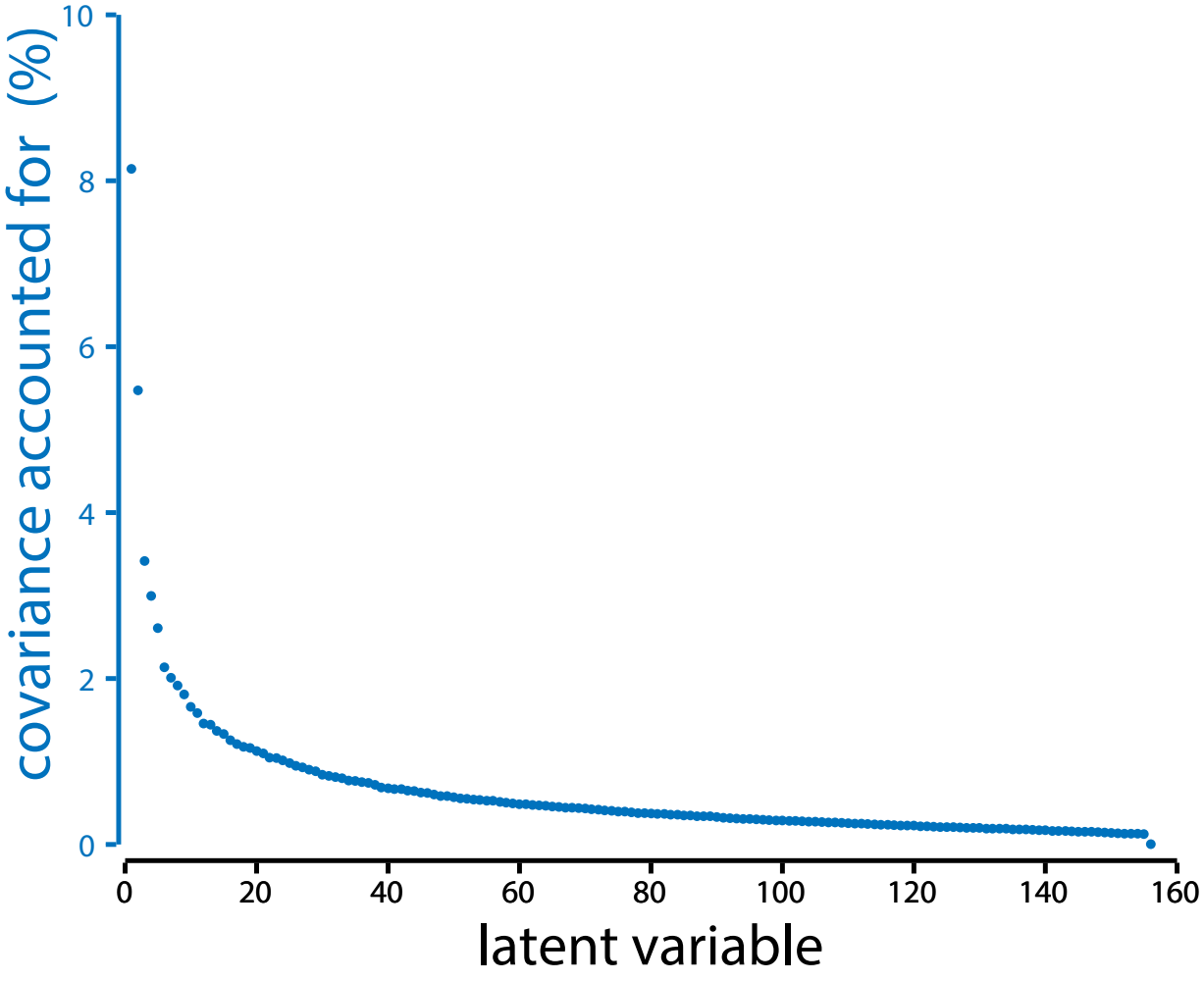
Statistical significance: permutation tests



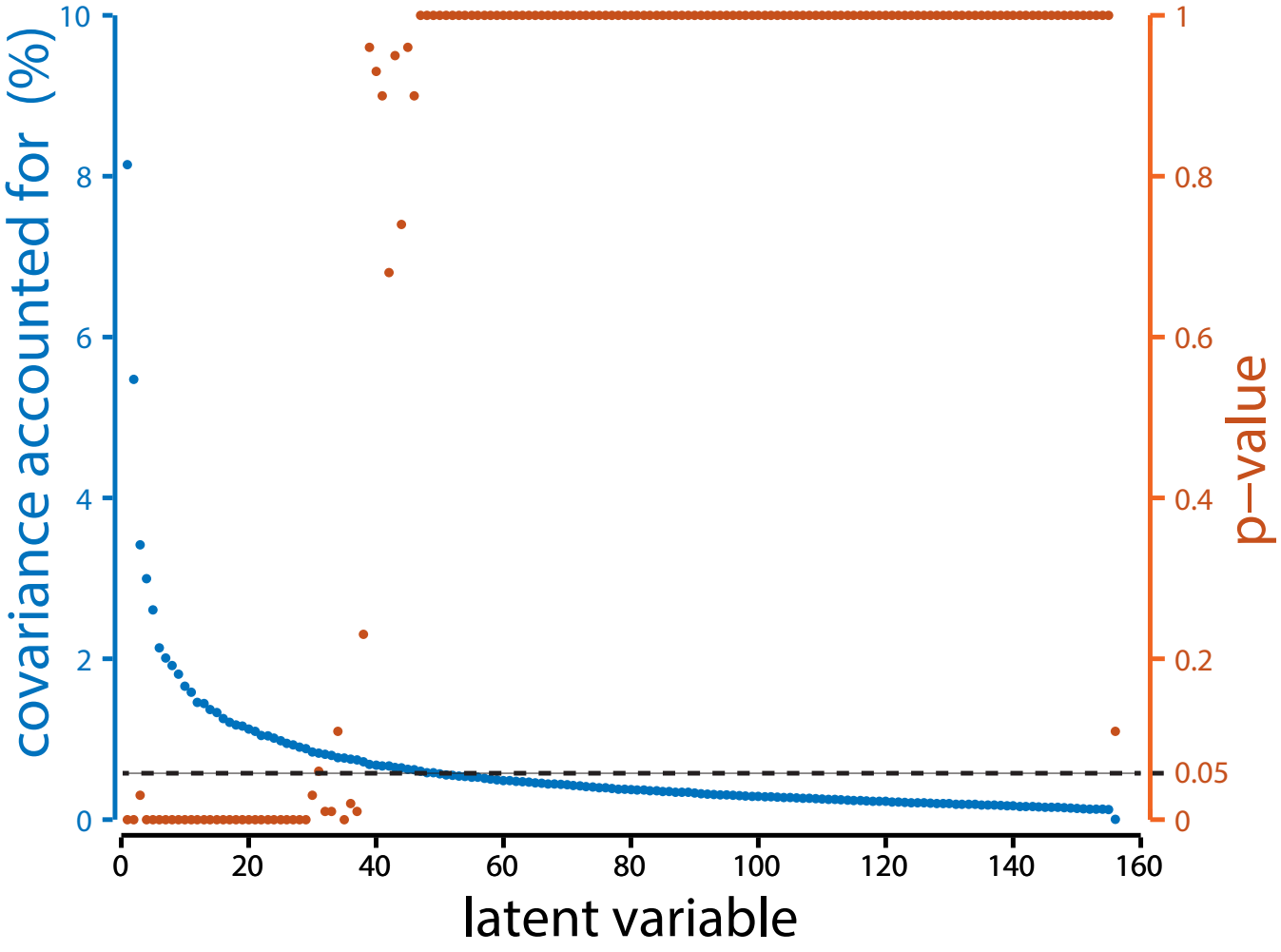
Edgington (1965, 1969) *J Psychol*

McIntosh et al. (1996, 2004) *NeuroImage*

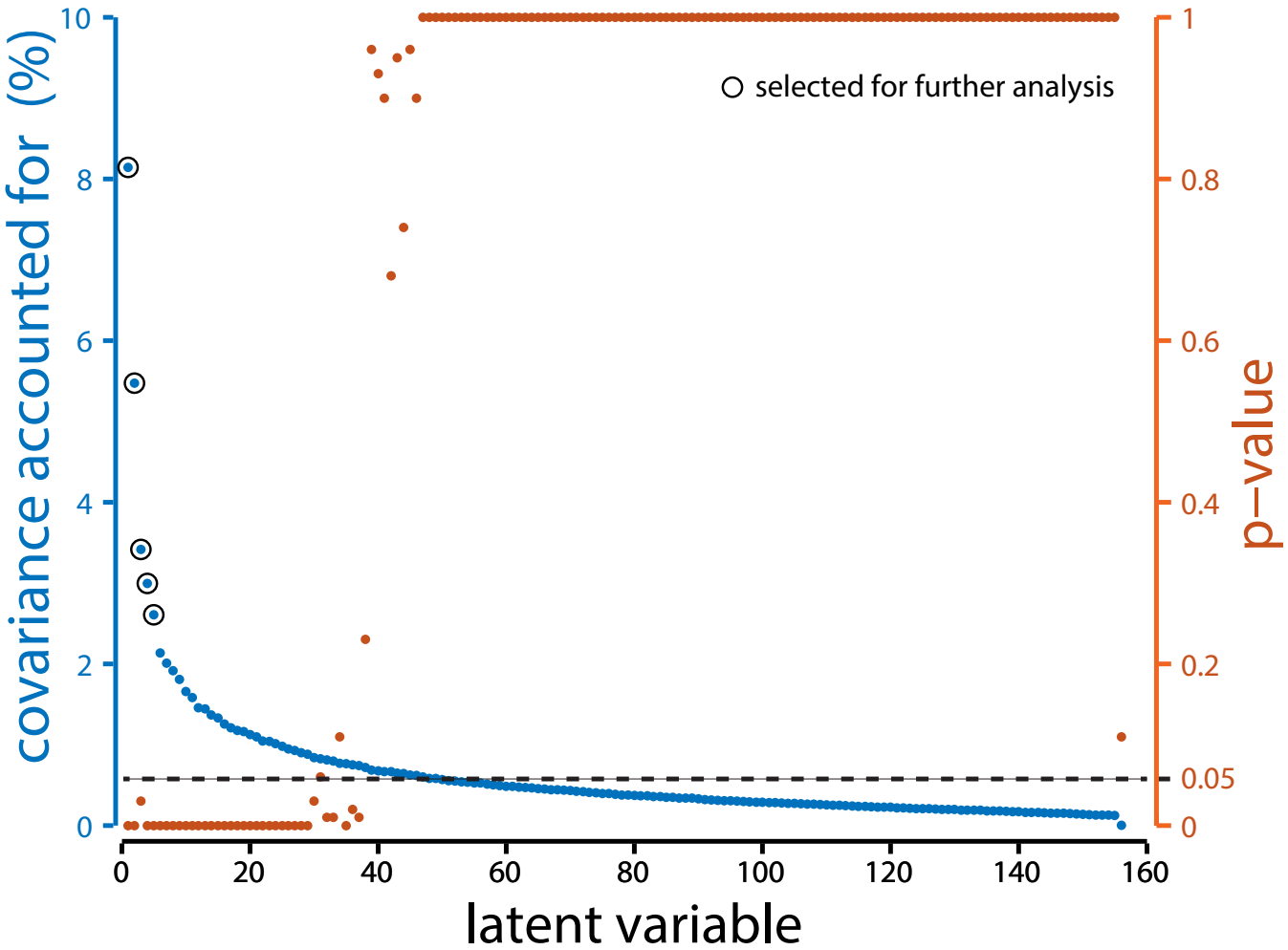
How many components to retain?



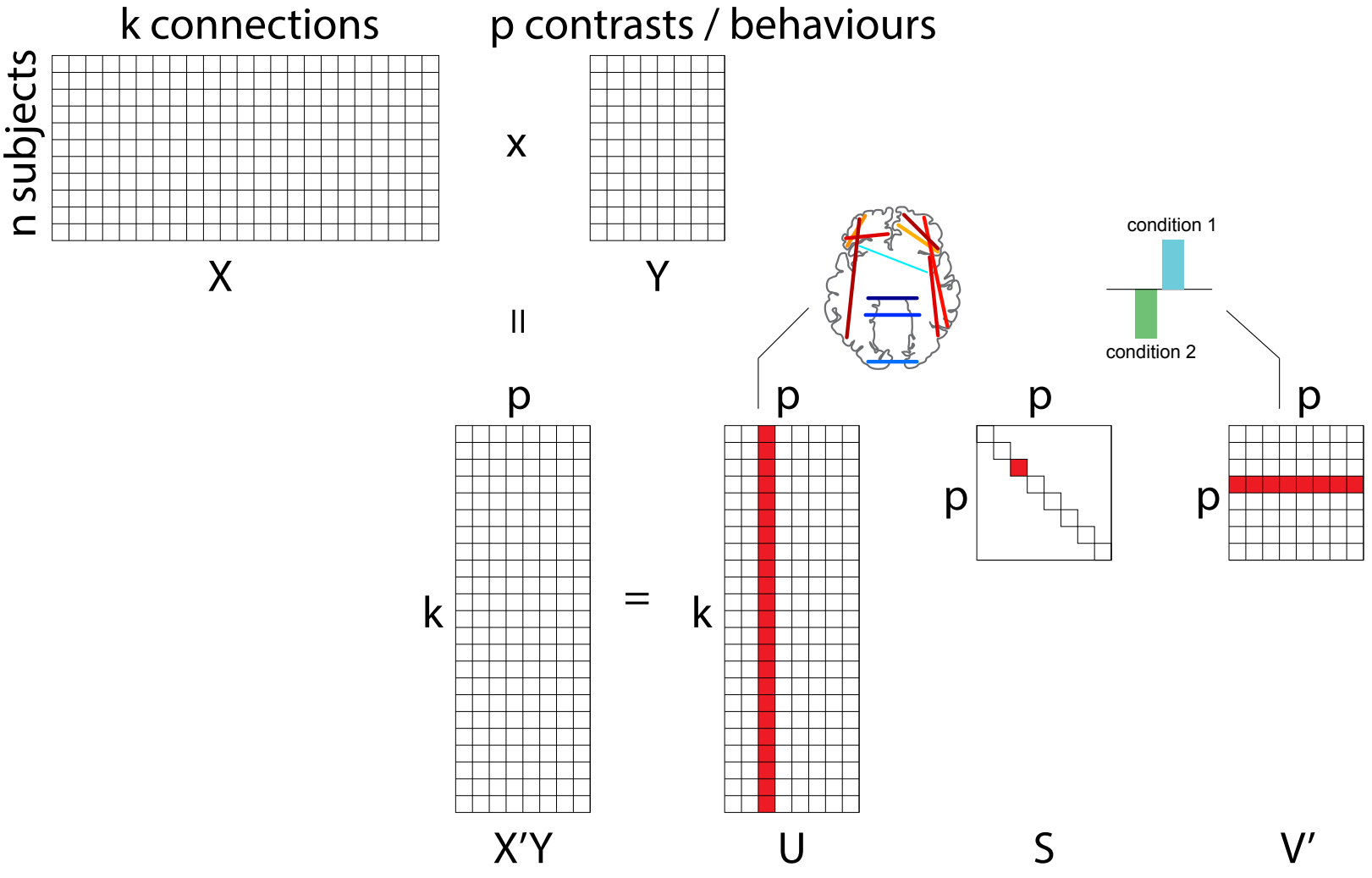
How many components to retain?



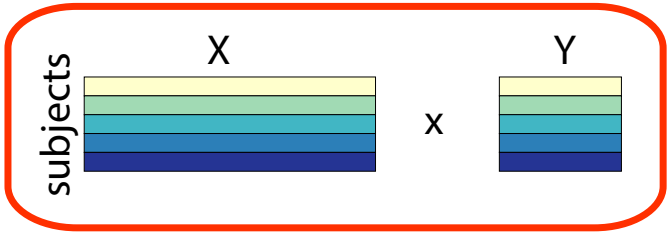
How many components to retain?



Which connections to focus on?



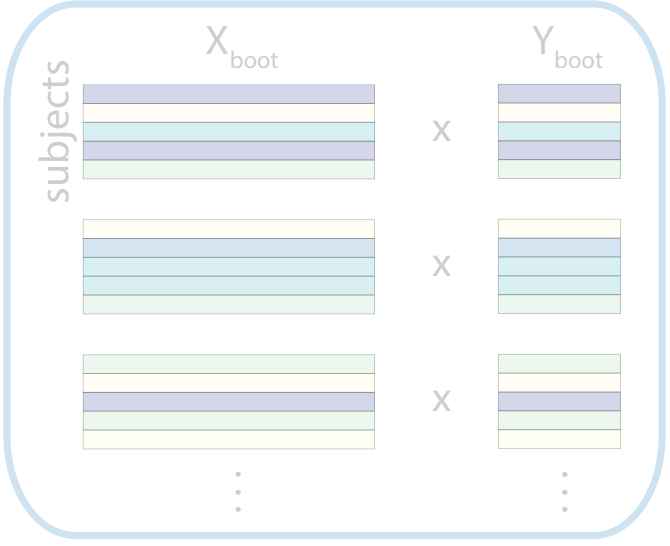
Reliability: bootstrapping



=

X'Y

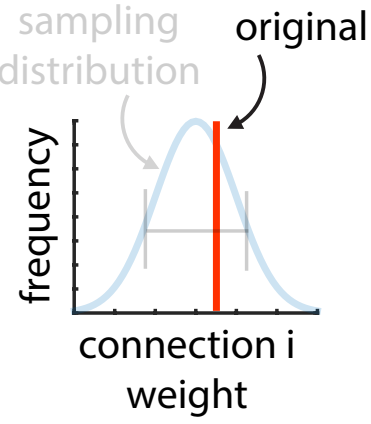
svd



= $X_{boot}'Y_{boot}$

= $X_{boot}'Y_{boot}$

= $X_{boot}'Y_{boot}$

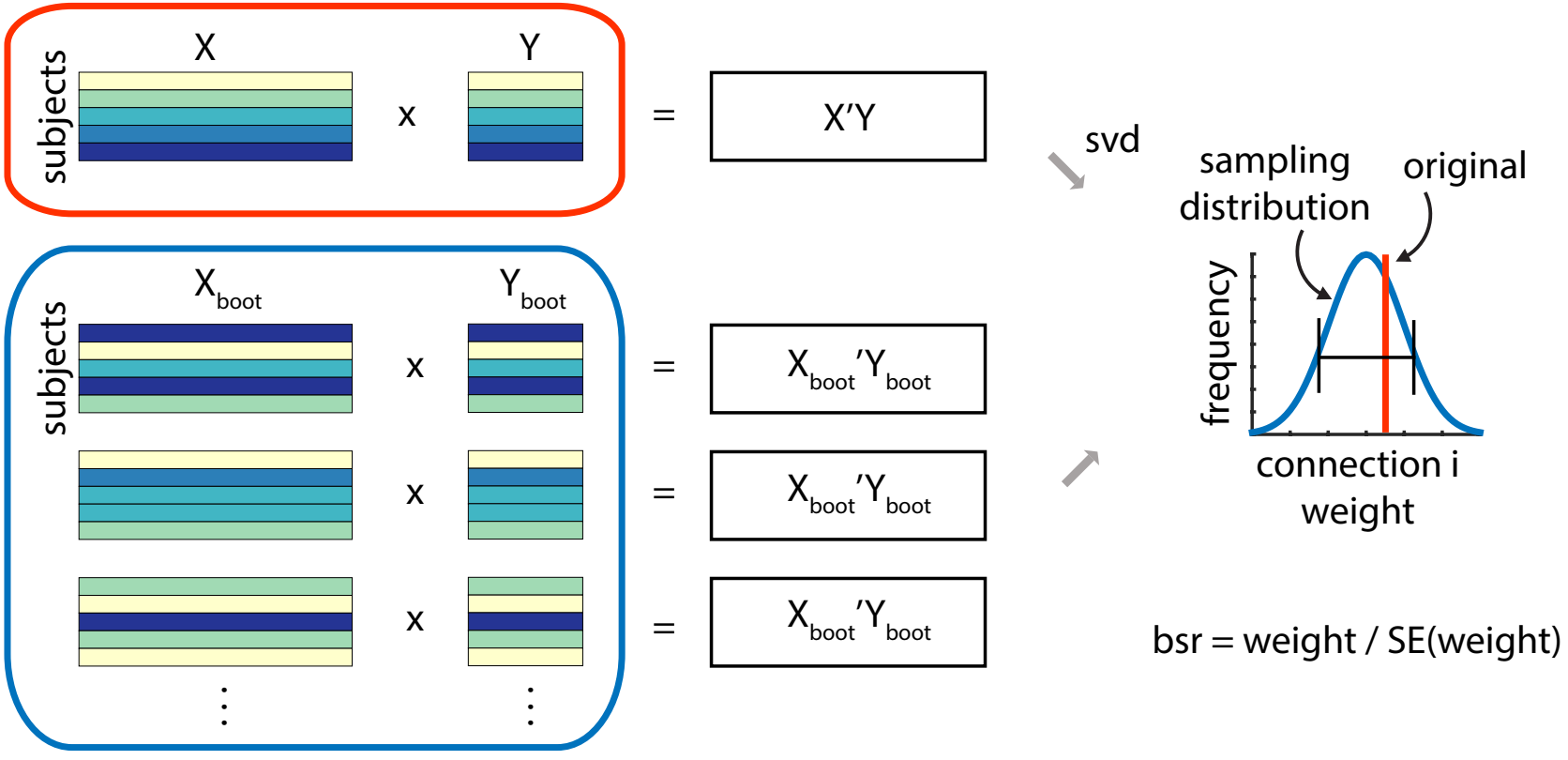


bsr = weight / SE(weight)

Efron & Tibshirani (1986) *Stat Sci*

McIntosh et al. (1996, 2004) *NeuroImage*

Reliability: bootstrapping

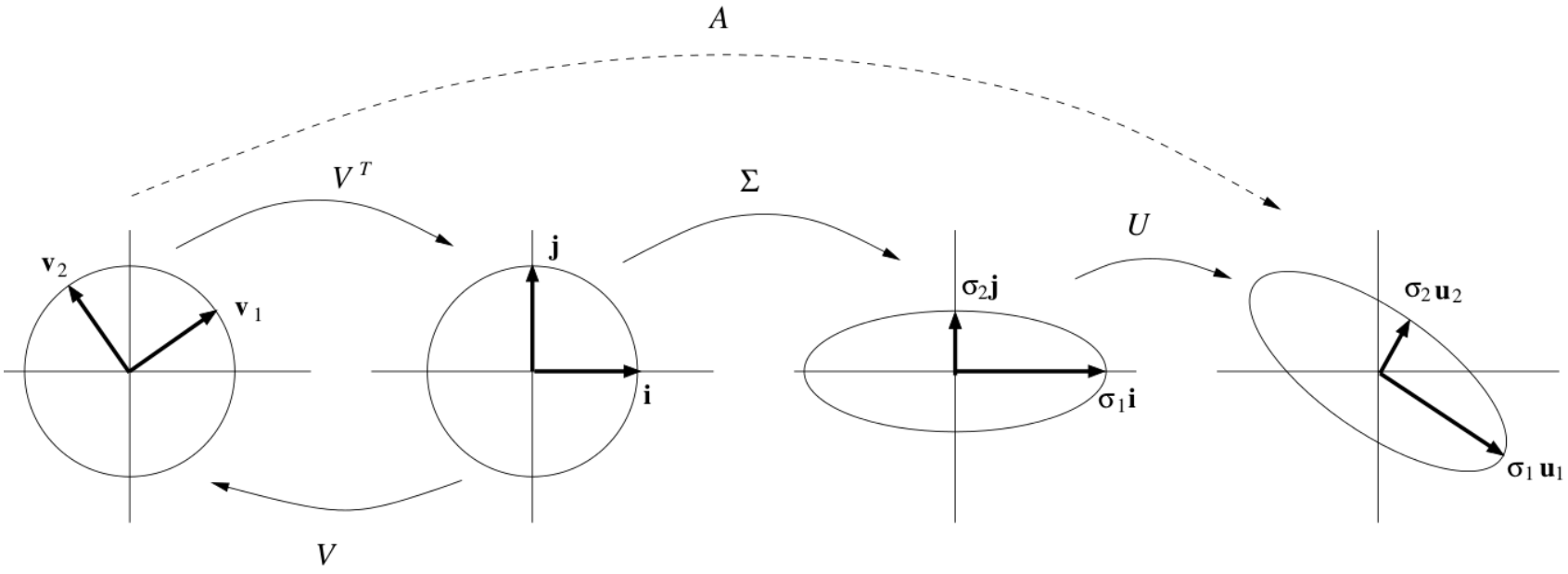


Efron & Tibshirani (1986) *Stat Sci*

McIntosh et al. (1996, 2004) *NeuroImage*

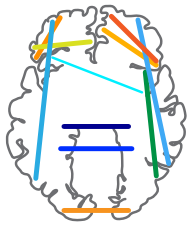
Matching randomized components

$$\mathbf{X}'\mathbf{Y} = \mathbf{USV}' \longleftrightarrow \mathbf{X}'_{\text{boot}}\mathbf{Y} = \mathbf{U}_{\text{boot}}\mathbf{S}_{\text{boot}}\mathbf{V}'_{\text{boot}}$$



Milan & Whittaker (1995) *J Roy Stat Soc C*

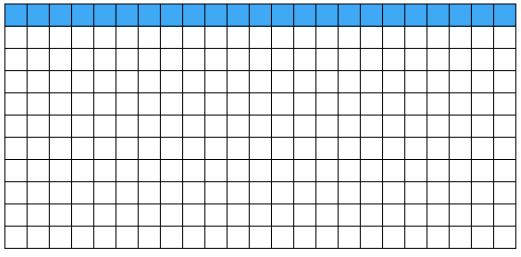
Individual participants



connection
strength

k connections

n subjects

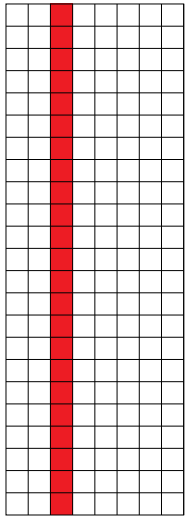


X

x

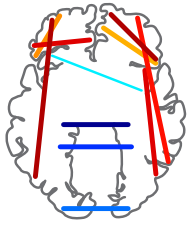
p

k



U

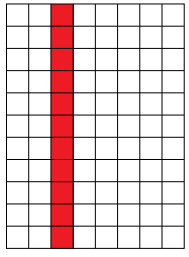
=



statistical
weight

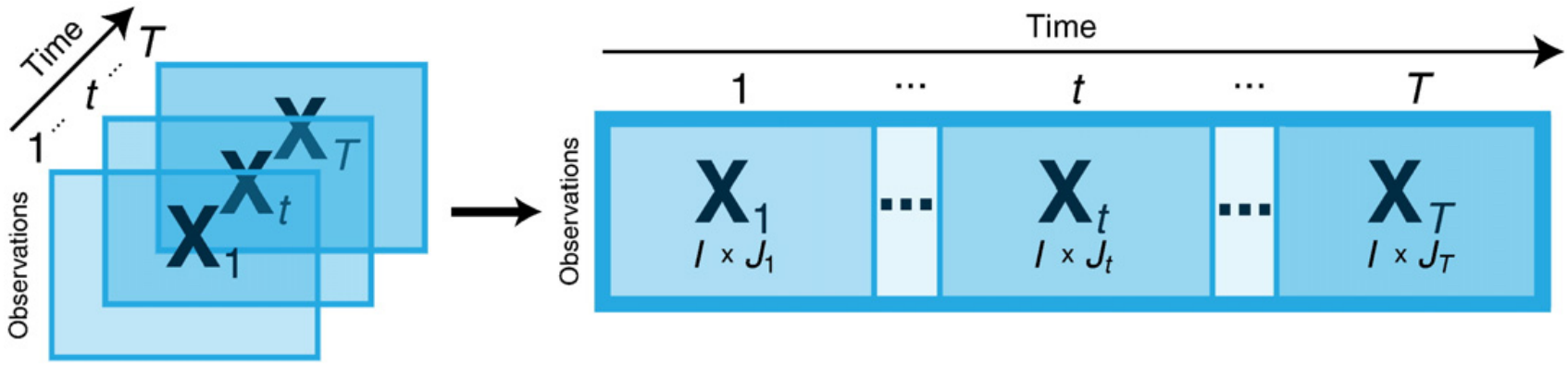
p

n



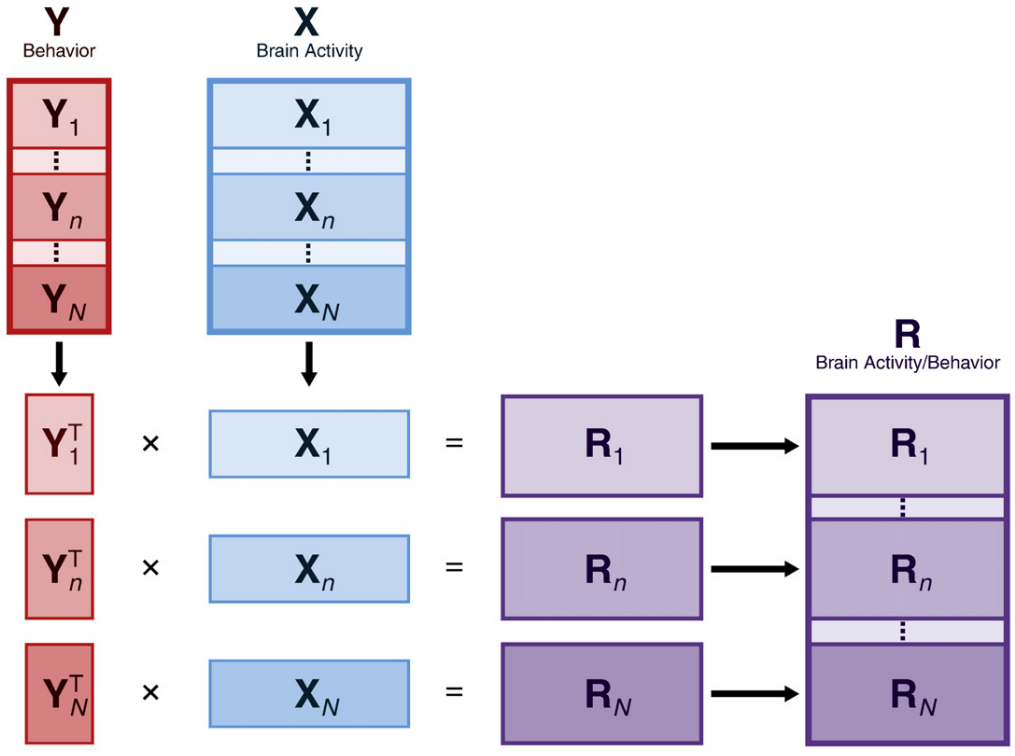
$F_i = XU_i$

Adding other dimensions



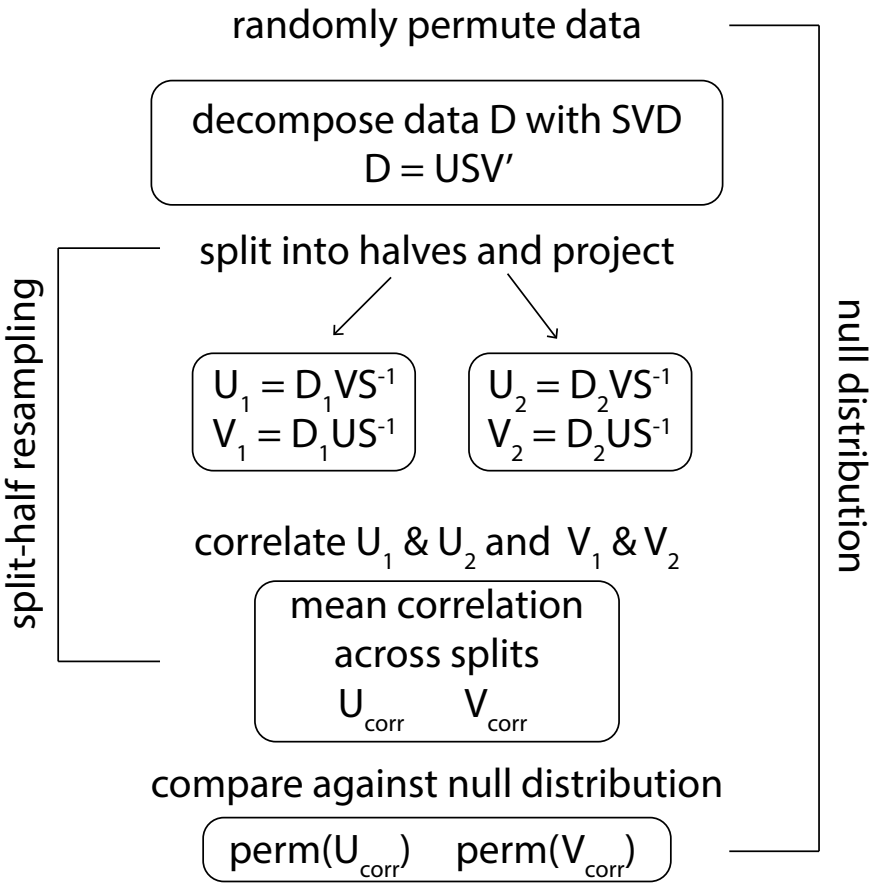
Krishnan et al. (2011) *NeuroImage*

Adding other dimensions



Krishnan et al. (2011) *NeuroImage*

Cross-validation: split-half resampling

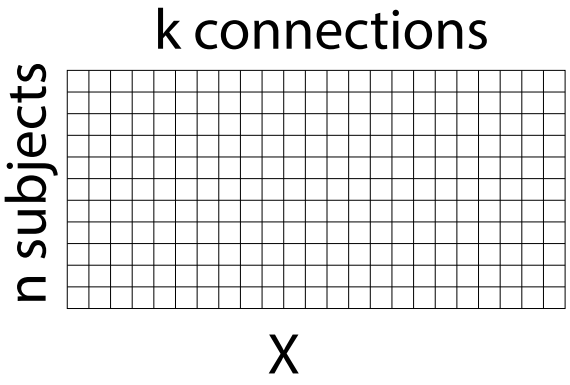


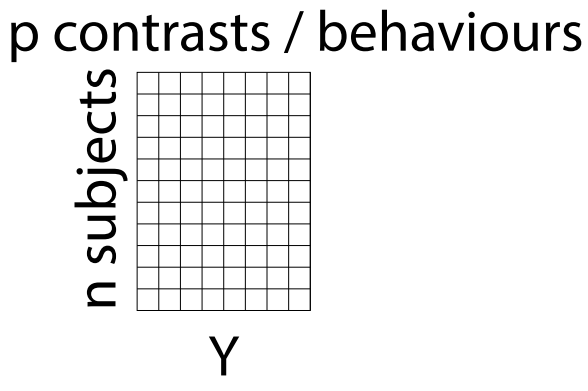
Strother et al. (2002) *NeuroImage*

Kovacevic et al. (2013) *New perspectives in Partial Least Squares* (Ed: Abdi)

Canonical correlation analysis (CCA)

$$\text{SVD}((\mathbf{X}'\mathbf{X}')^{-1/2}(\mathbf{X}'\mathbf{Y})(\mathbf{Y}'\mathbf{Y})^{-1/2})$$

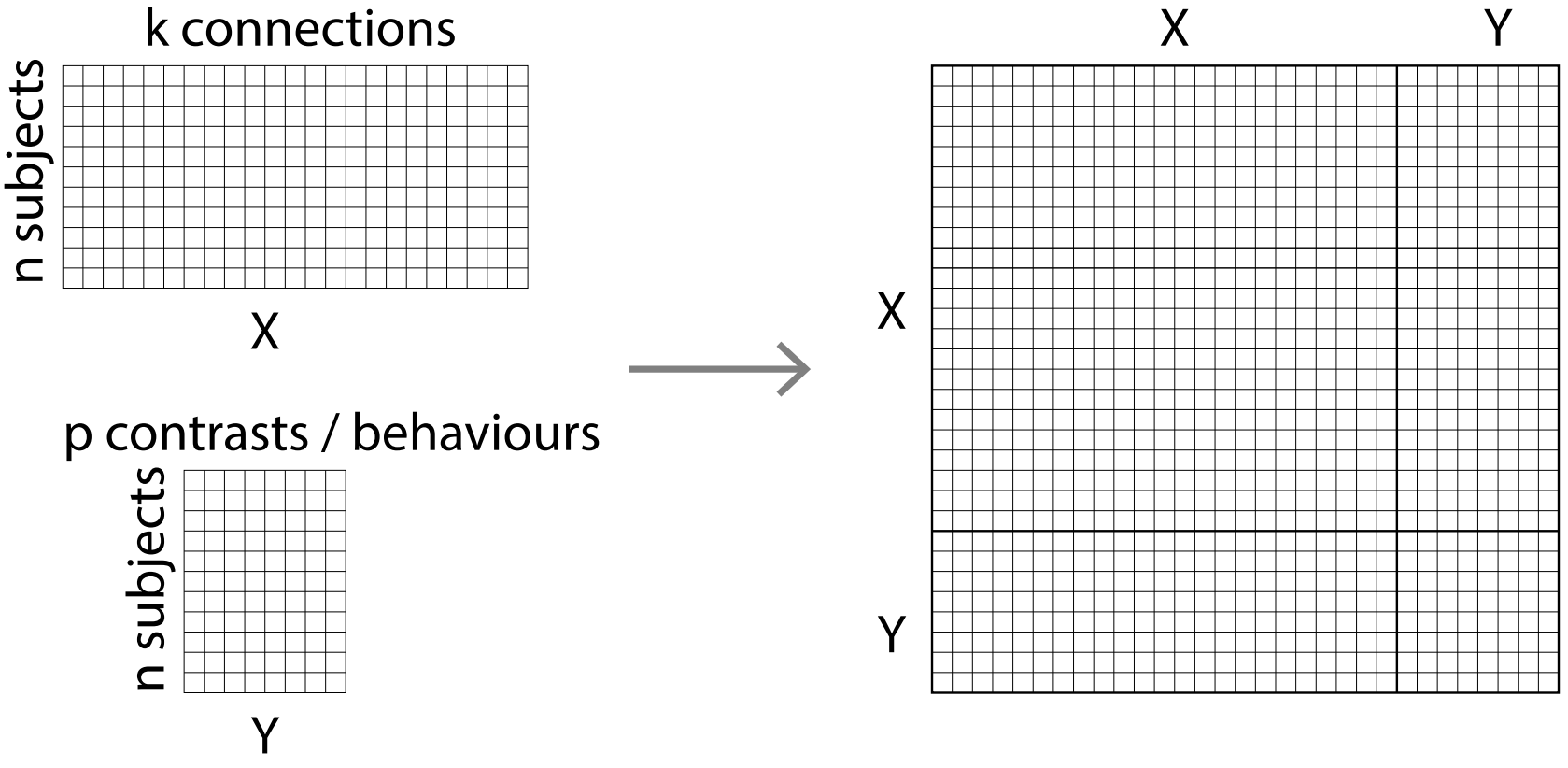




Hotelling (1936) *Biometrika*

Canonical correlation analysis (CCA)

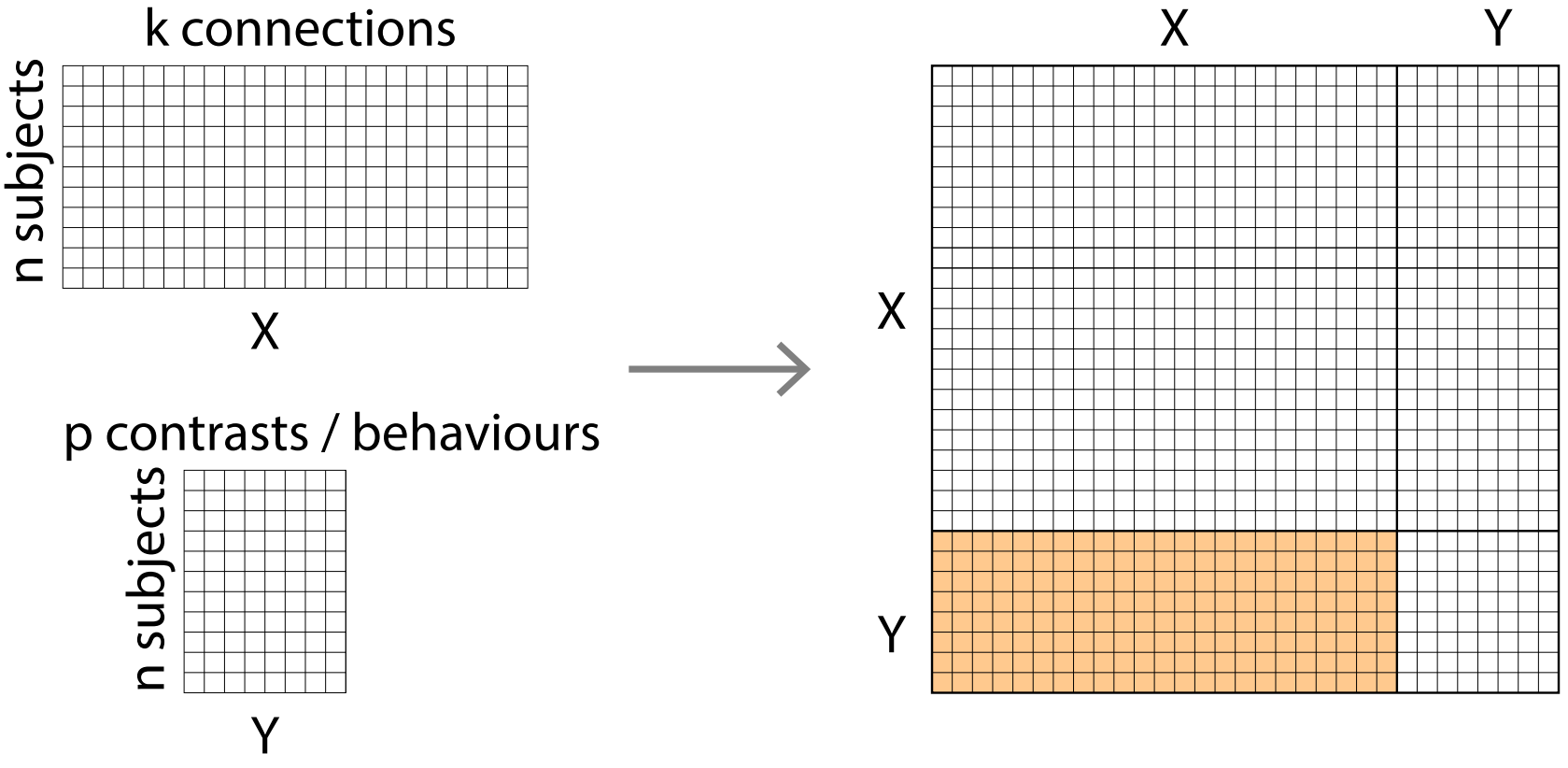
$$\text{SVD}((\mathbf{X}'\mathbf{X}')^{-1/2}(\mathbf{X}'\mathbf{Y})(\mathbf{Y}'\mathbf{Y})^{-1/2})$$



Hotelling (1936) *Biometrika*

Canonical correlation analysis (CCA)

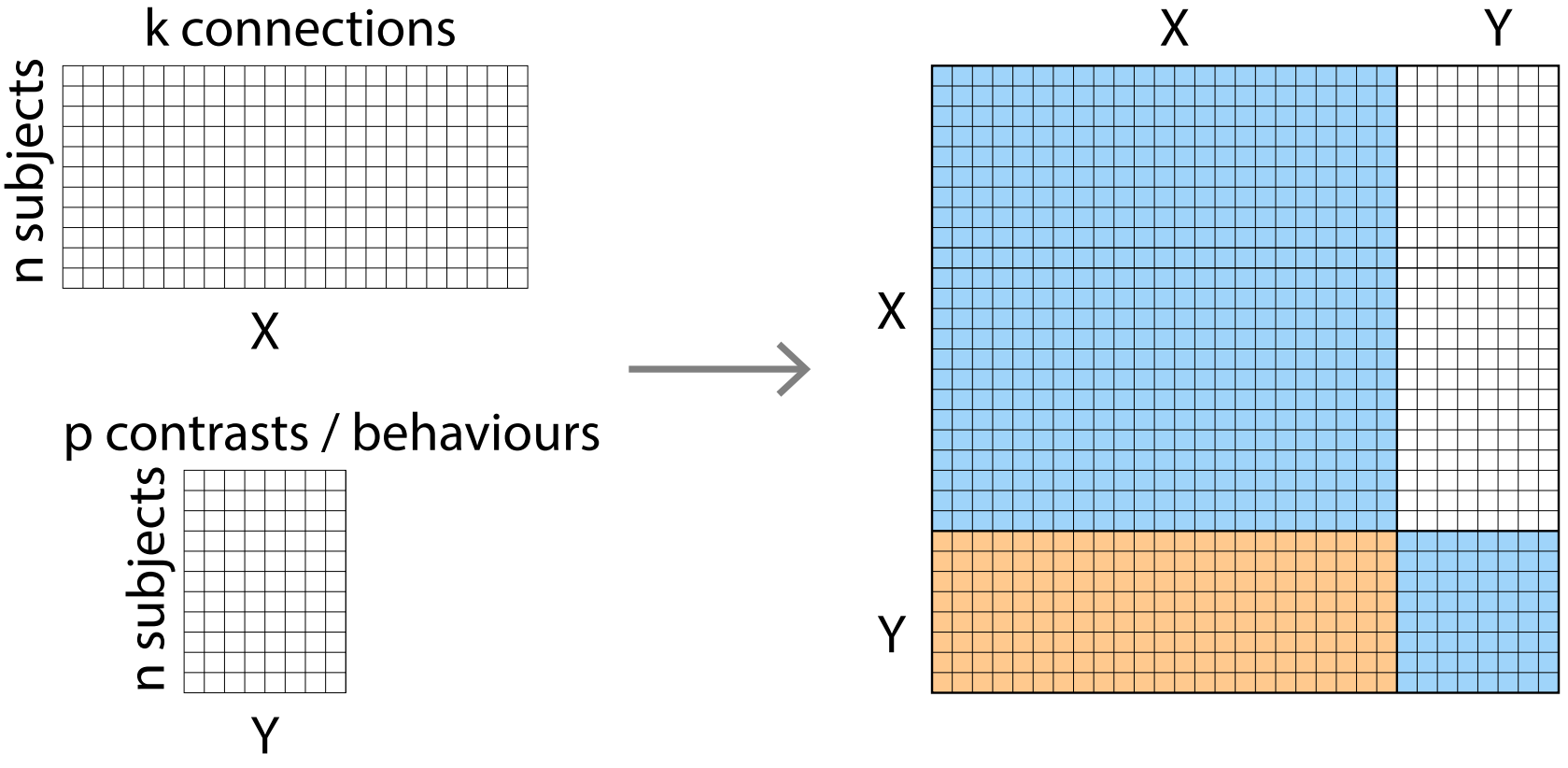
$$\text{SVD}((\mathbf{X}'\mathbf{X}')^{-1/2}(\mathbf{X}'\mathbf{Y})(\mathbf{Y}'\mathbf{Y})^{-1/2})$$



Hotelling (1936) *Biometrika*

Canonical correlation analysis (CCA)

$$\text{SVD}((\mathbf{X}'\mathbf{X}')^{-1/2}(\mathbf{X}'\mathbf{Y})(\mathbf{Y}'\mathbf{Y})^{-1/2})$$



Hotelling (1936) *Biometrika*

Linear discriminant analysis (LDA)

$$\text{SVD}(\mathbf{W}^{-1}\mathbf{B})$$

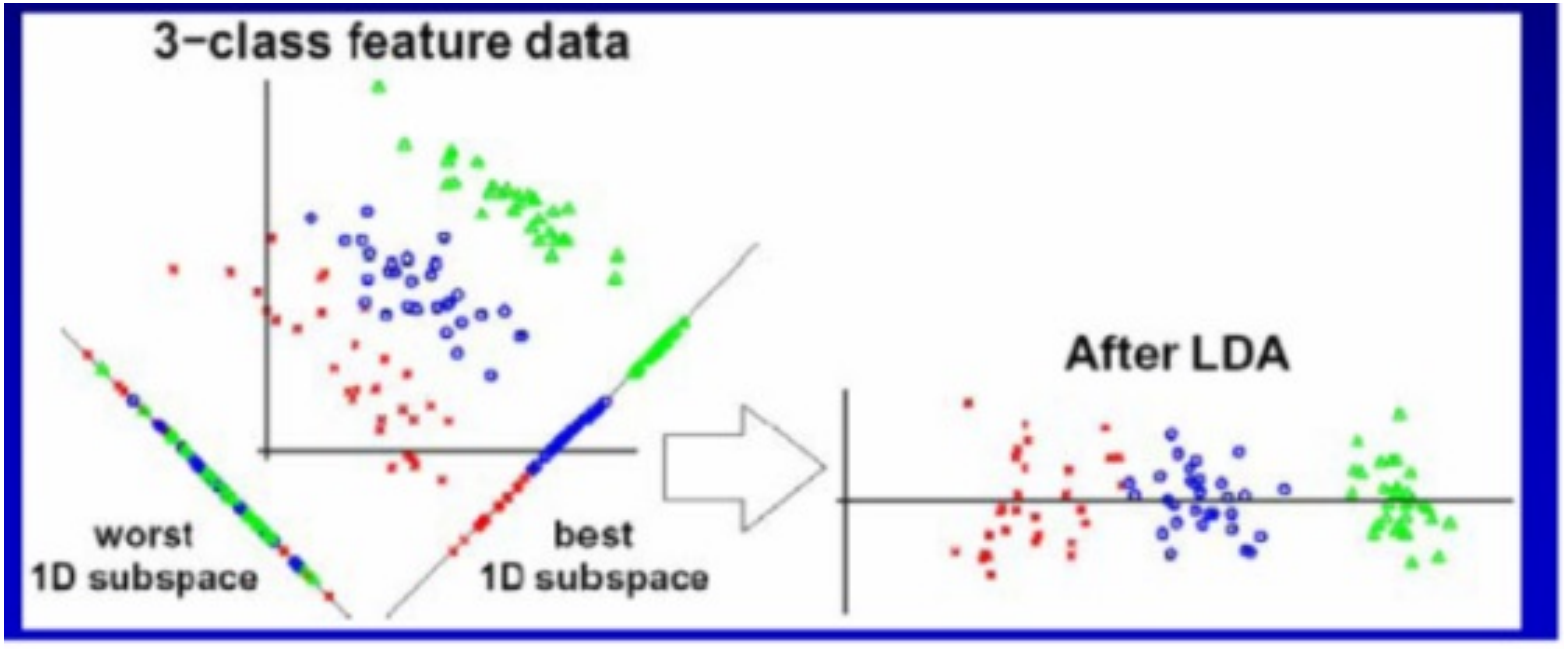


image: https://mlalgorithm.wordpress.com/

Friston et al. (1995) *NeuroImage*

Extensions

- sparse/regularized solutions, e.g.
  sCCA: Witten & Tibshirani (2009) *Stat Appl Genet Mol Biol*
  PLS-CA: Beaton et al. (2015) *Psychol Meth*
- extensions to 3+ data sets, e.g.
  PARAFAC: Bro (1997) *Chemometr Intell Lab*
  Multiway PLS: Wold et al. (1987) *J Chemometrics*
- nonlinear dependencies
- Bayesian implementations, e.g.
  IBFA: Virtanen et al. (2011) *ICML-II*
- prediction

Limitations and considerations

- overfitting
- linear
- unique partitioning of variance/covariance
- inference on individual variables

Summary

- multivariate models embody network property
- all techniques entail unique assumptions
- many linear multivariate techniques are related
- multivariate techniques are versatile