# Supervised Analysis: PLS, OPLS and O2PLS in Metabolic Profiling

Dr Tim Ebbels

Computational and Systems Medicine

Imperial College

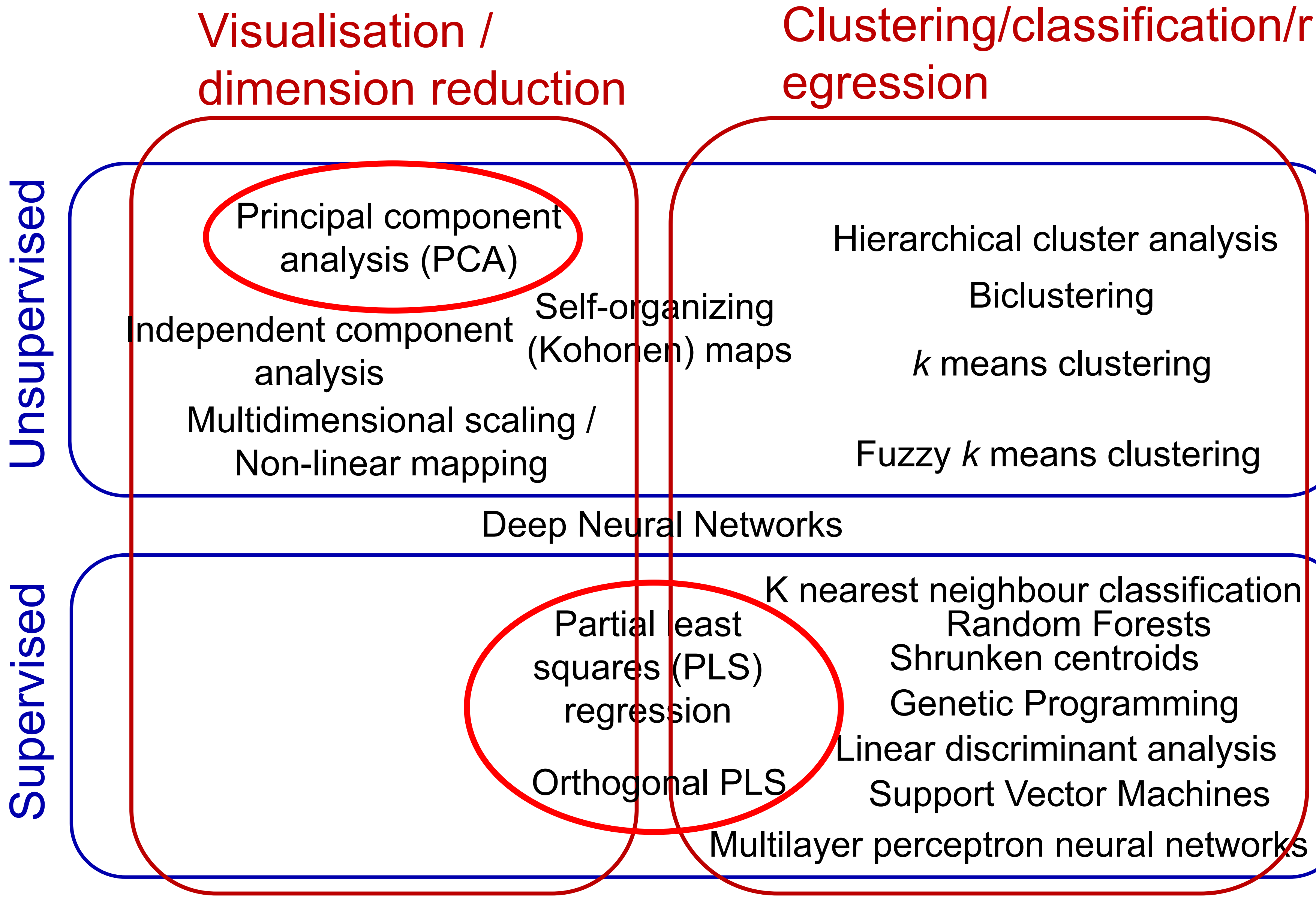
# Plan

- **Partial Least Squares**
  - Geometrical overview
  - PLS-DA
- **Orthogonal variation**
  - What is it? Why bother with it?
  - How to make use of it?
- **O-PLS model**
  - Geometrical view
  - Compare with PCA/PLS
  - Exploring orthogonal variation
- **O2-PLS model**
  - Conceptual overview
  - Comparison with other methods
  - O-PLS-DA & O2-PLS for multiblock exploration

# Plan

- **Partial Least Squares**
  - Geometrical Overview
  - PLS-DA
- **Orthogonal variation**
  - What is it? Why bother with it?
  - How to make use of it?
- **O-PLS model**
  - Geometrical view
  - Compare with PCA/PLS
  - Exploring orthogonal variation
- **O2-PLS model**
  - Conceptual overview
  - Comparison with other methods
  - O-PLS-DA & O2-PLS for multiblock exploration

# Multivariate analysis techniques



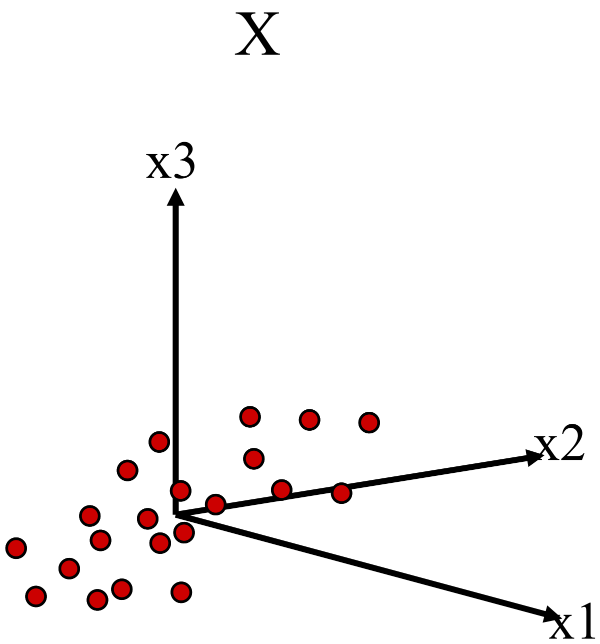
# Supervised methods

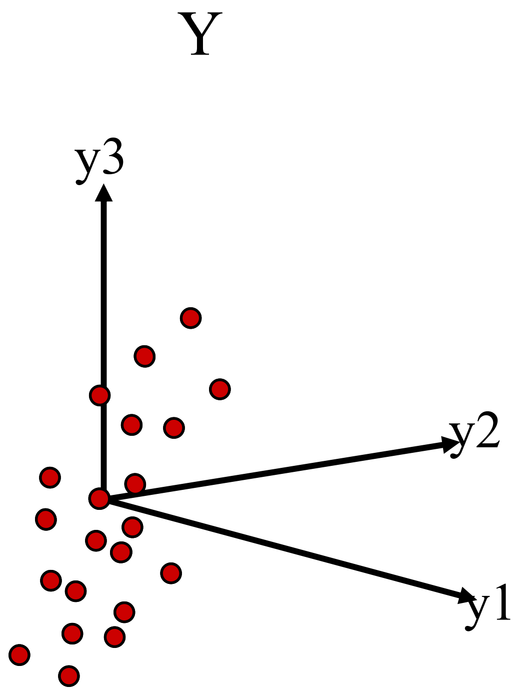
- *Unsupervised* = algorithm does not know true answer/output (e.g. class) – e.g. PCA
- *Supervised* = algorithm does know true answer
  - Attempt to find rule which predicts output (e.g. class) for given input (e.g. metabolic profile)
- Two cases
  - Classification: output is class label
  - Regression: output is continuous
- Usually denote input data as X and output as Y

# Partial Least Squares (PLS)

- Regression technique - relates X to Y
- Projection technique (linear like PCA)
- Often considered the 'regression extension of PCA'
- One of *many* multivariate regression techniques
  - OLS, PCR, RR, RRR…
- Good when
  - More variables than samples
  - Many highly correlated variables
  - More than one Y variable
  - Missing data
- Common conditions in metabolomics!

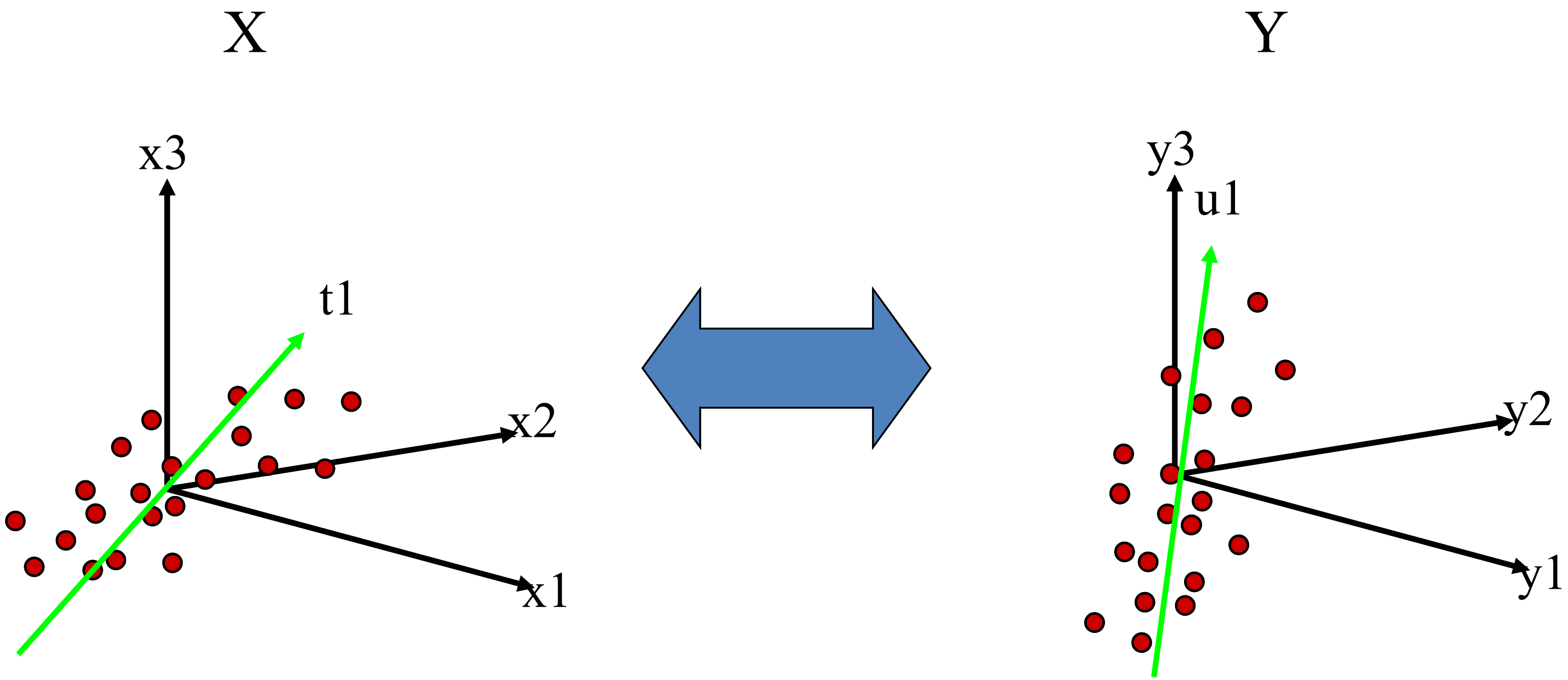
# PLS - Step by step (1)




Initially, we have two sets of *N* (mean centred) data points in the *X* and *Y* spaces

# PLS - Step by step (2)

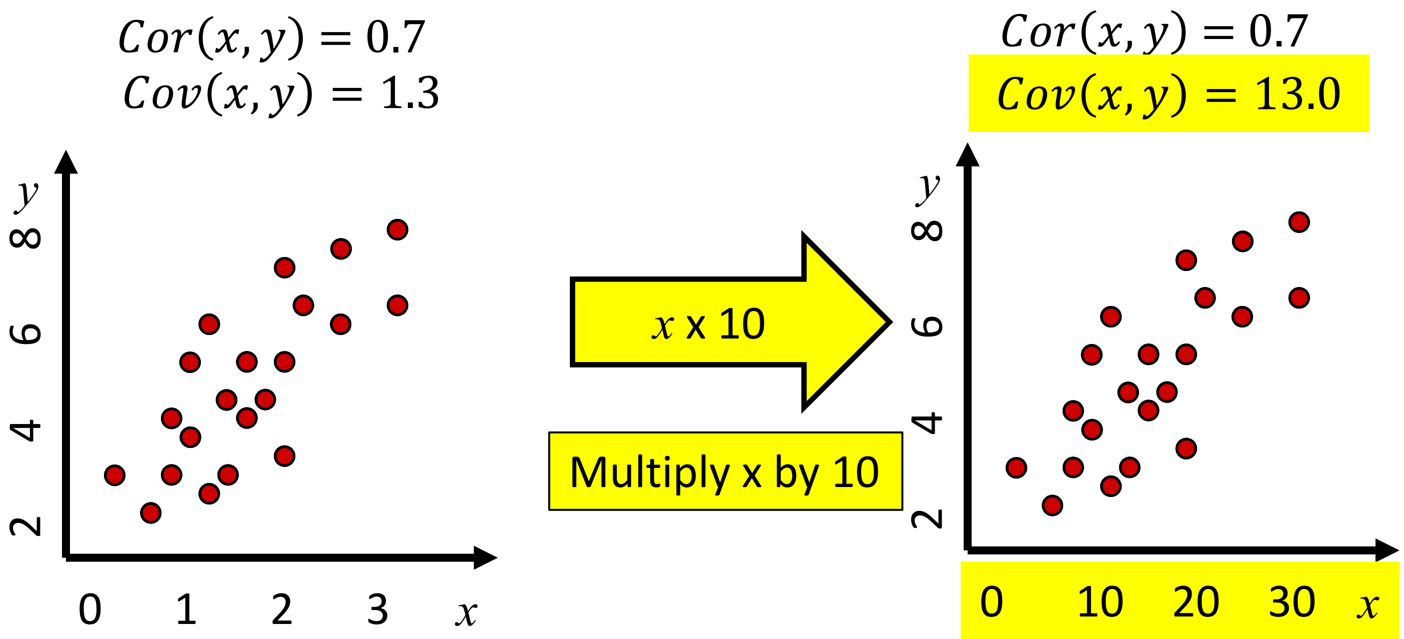



First PLS component: maximises *covariance* between X scores (t) and Y scores (u)

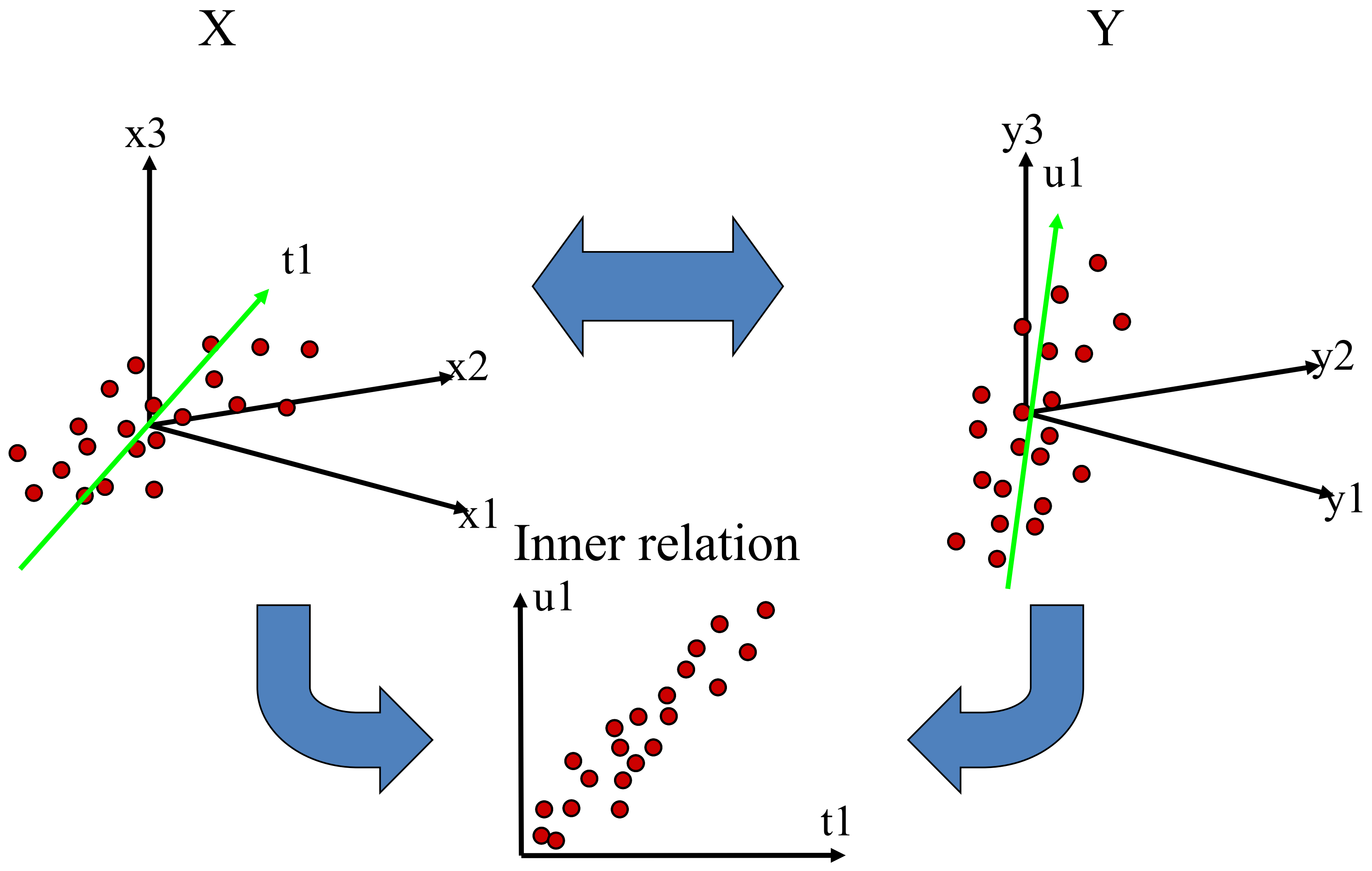
- Good summary of X space (X variance)
- Good representation of relationship between X & Y (X-Y correlation)

# Covariance and Correlation

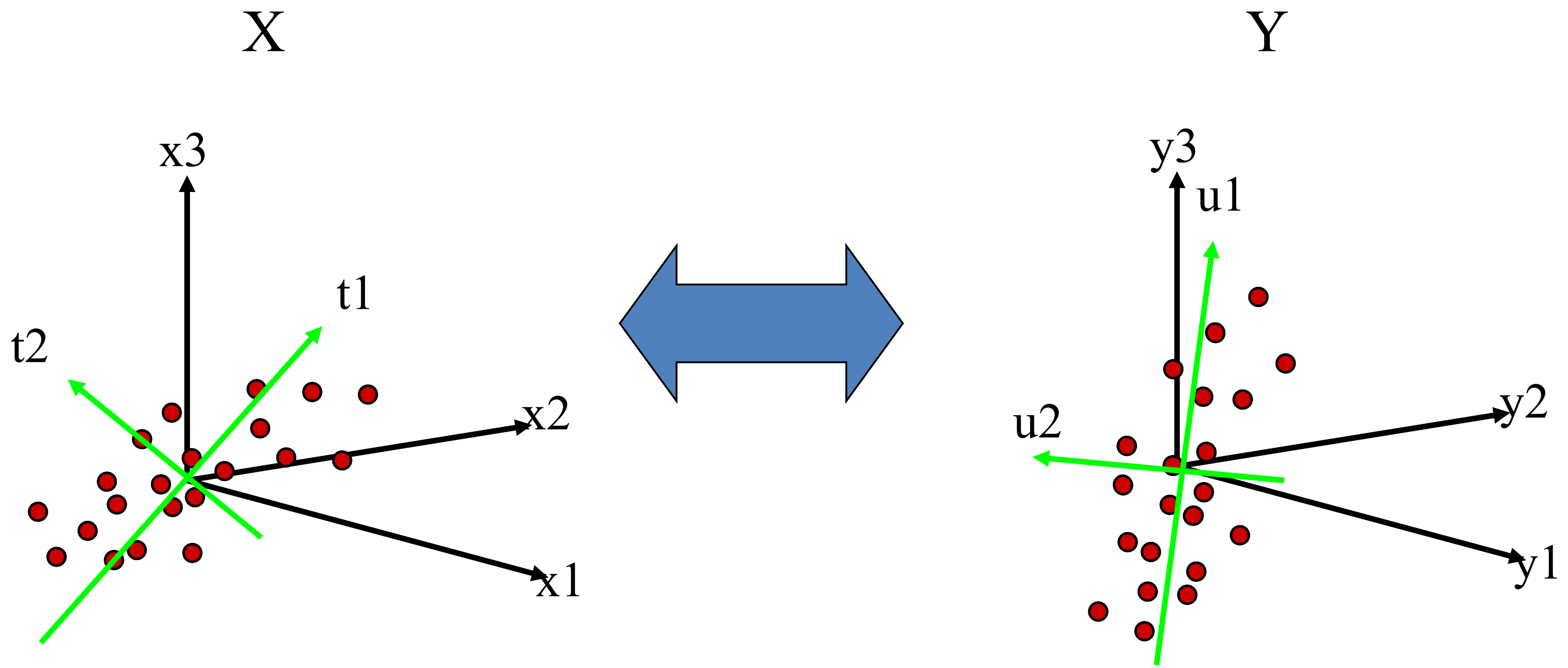
$$Cov(x,y) = Cor(x,y) \times \sigma_x \sigma_y$$



# PLS - Step by step (3)



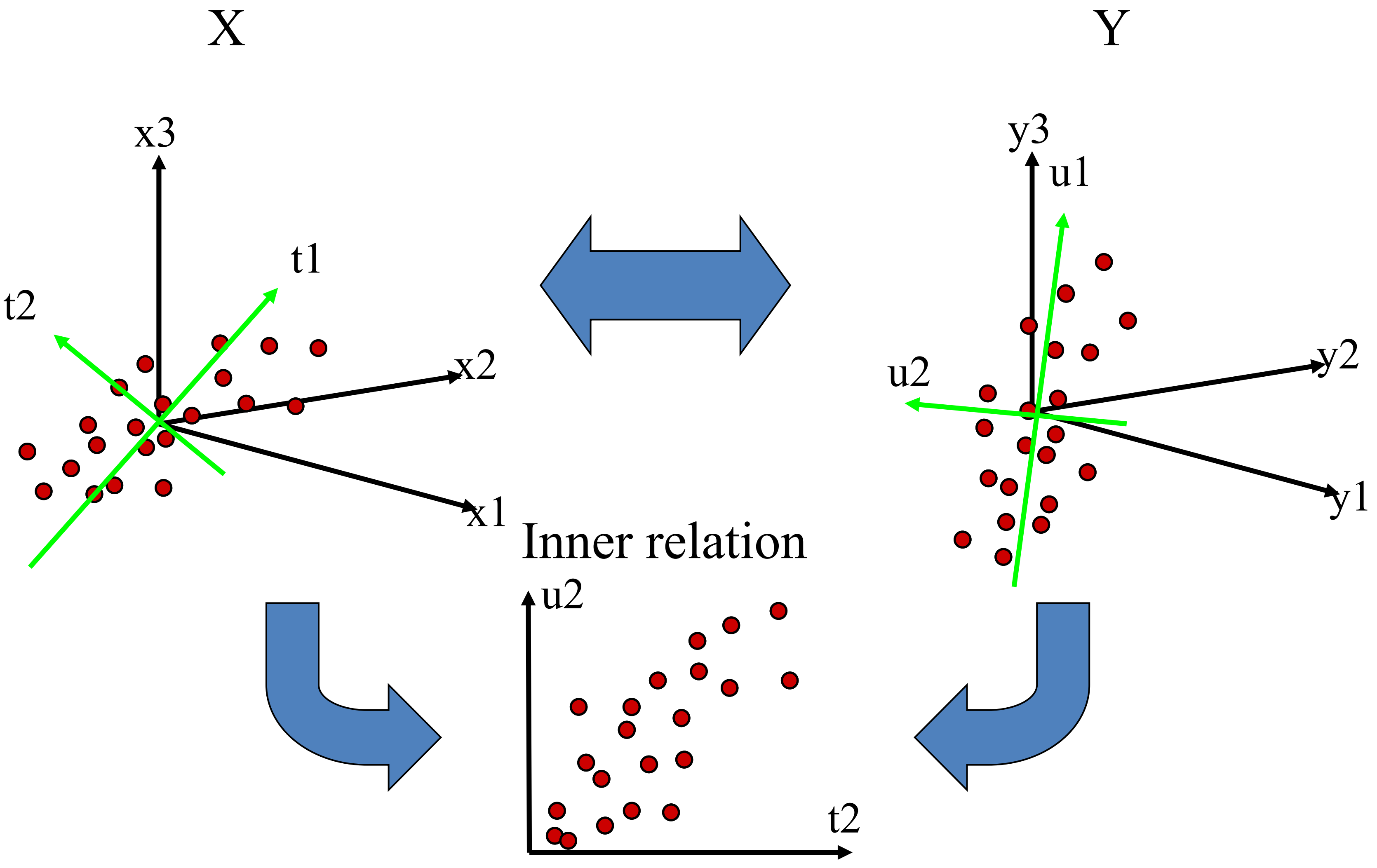
# PLS - Step by step (4)



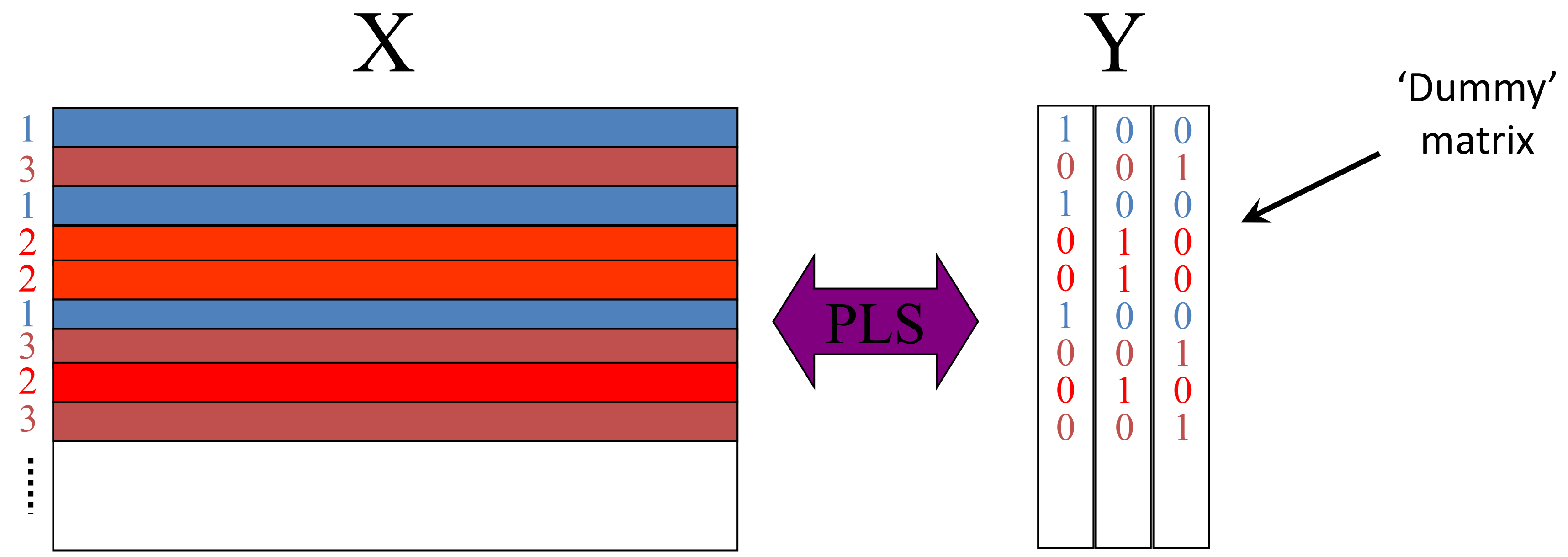

Subsequent components:

- orthogonal (uncorrelated) to previous components and
- continue to maximise the covariance between X & Y scores
    - good description of X space & relationship between X & Y

# PLS - Step by step (5)



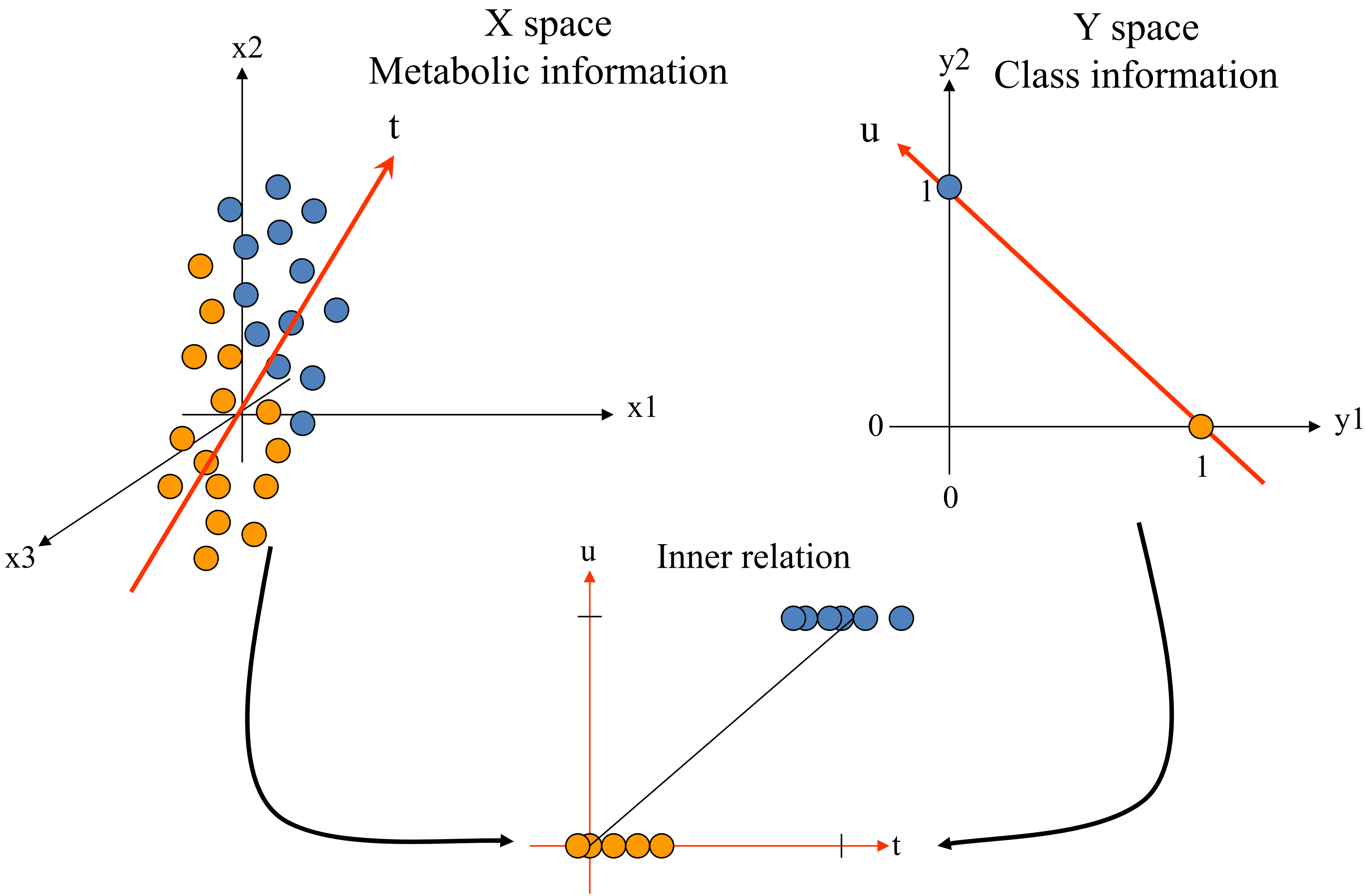
# PLS-DA schematic




- PLS-DA models separation between classes
- Y = 'dummy matrix' – gives class membership
- Q: Why can't Y just be a single column of class numbers 1,2,3 etc?

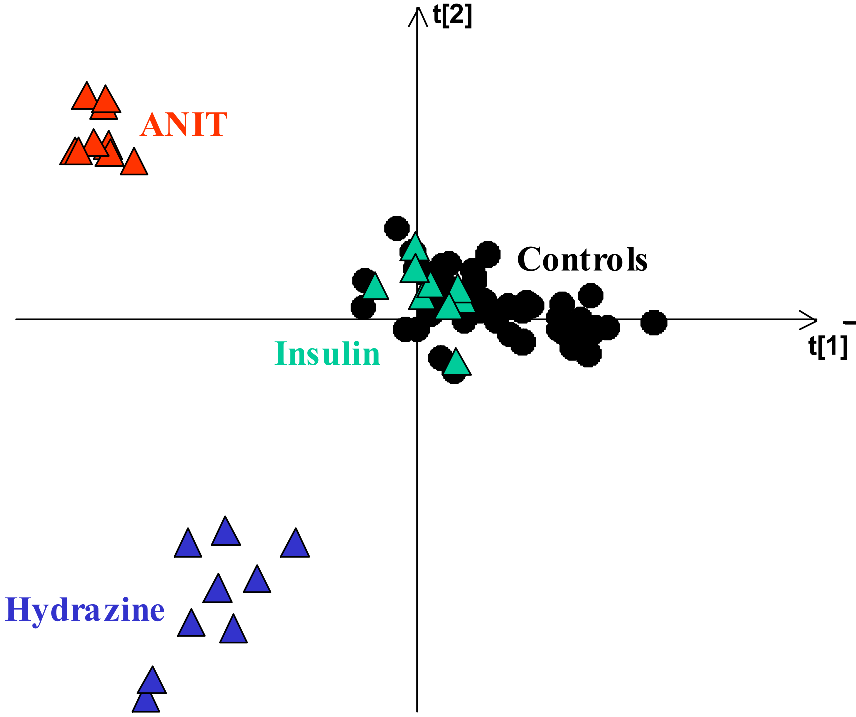
Credit: Henrik Antti / Umea University

# PLS-DA - geometrical view

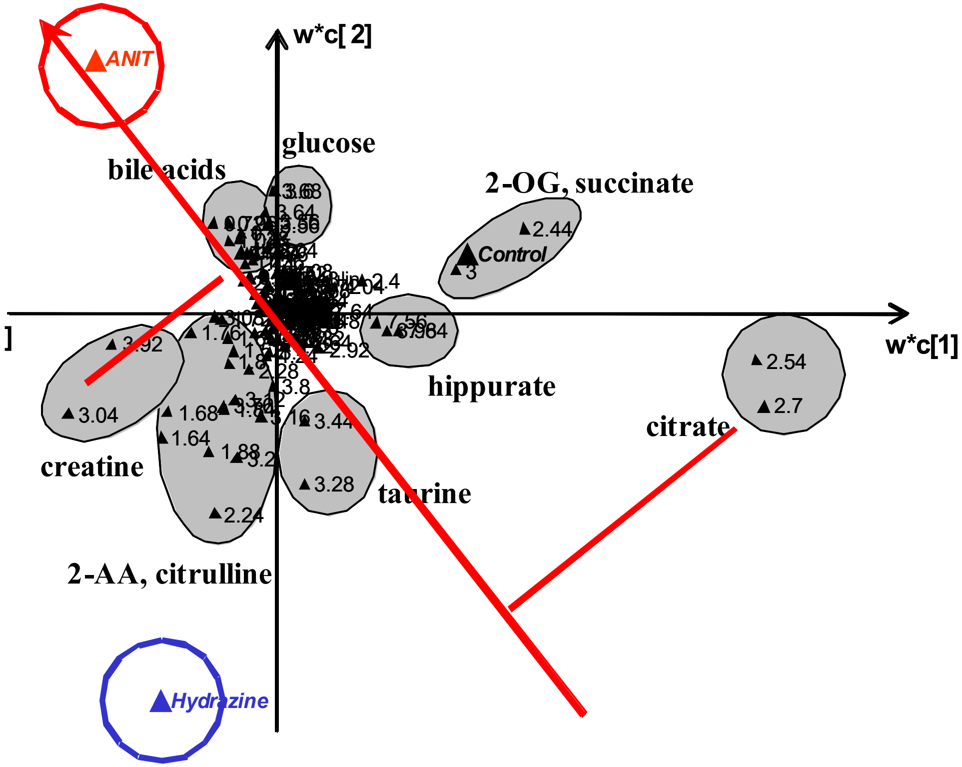


# PLS-DA Interpretation

*PLS-DA 'scores'*




*PLS-DA 'loadings'*




*Scores*

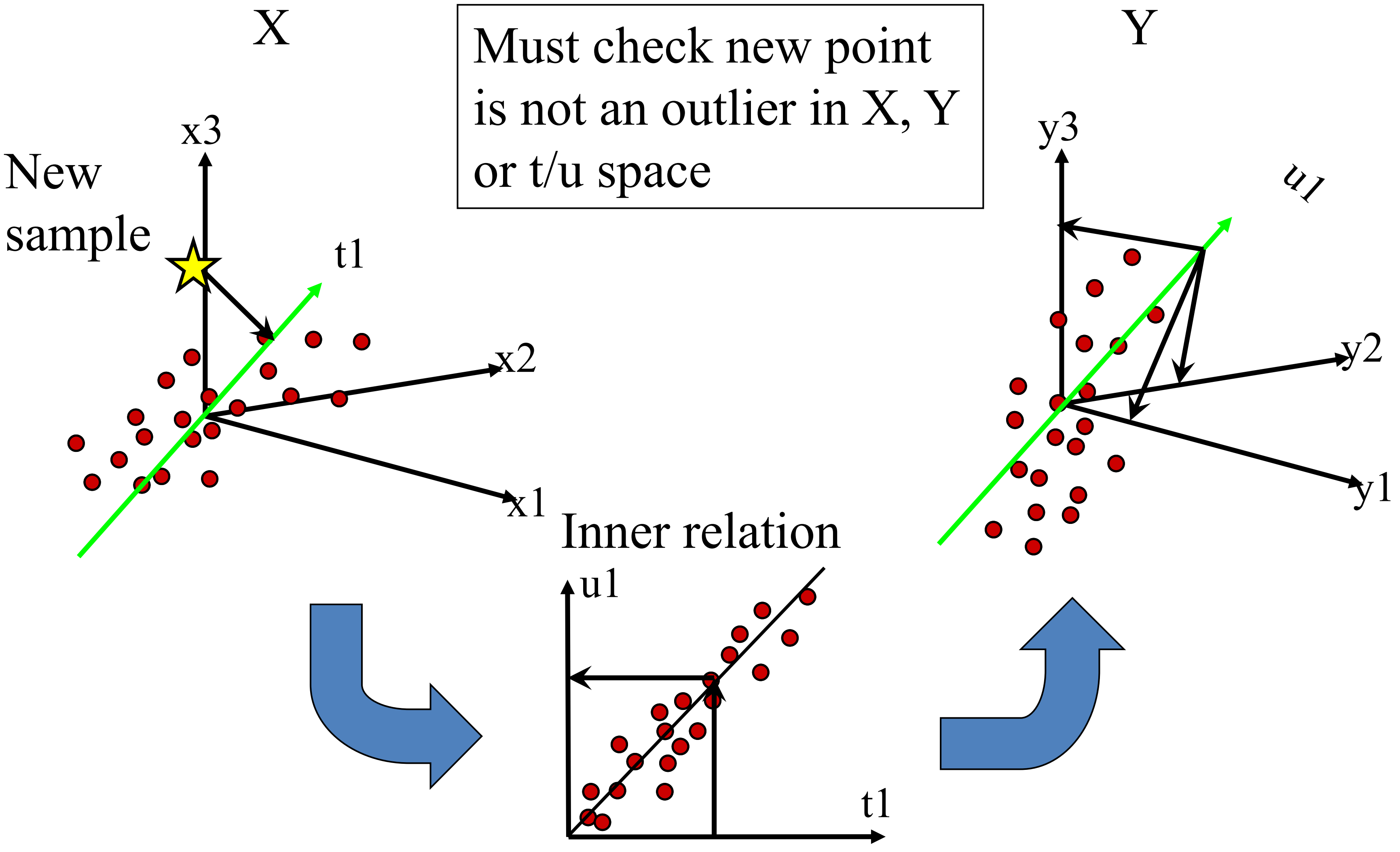
- information about class separation
- ANIT, Hydrazine → effect
- Insulin → no effect

*Loadings*

- information about variables responsible for class separation.
- ANIT c.f. Controls: citrate ↓, bile acids ↑

Credit: Consortium for Metabonomic Toxicology / Imperial College

# PLS - Prediction



# PLS & PLS-DA - Summary

- PLS – regression method
  - models relationship between X & Y
- PLS components maximise *covariance* between scores in X & Y spaces
- PLS-DA – classification with PLS
  - Y = dummy matrix, gives class info

# Plan

- **Partial Least Squares**
  - Geometrical overview
  - PLS-DA
- **Orthogonal variation**
  - What is it? Why bother with it?
  - How to make use of it?
- **O-PLS model**
  - Geometrical view
  - Compare with PCA/PLS
  - Exploring orthogonal variation
- **O2-PLS model**
  - Conceptual overview
  - Comparison with other methods
  - O-PLS-DA & O2-PLS for multiblock exploration

# What is orthogonal variation?

- Orthogonal variation: Systematic variation in one block which is *not* linearly related to the other block(s)

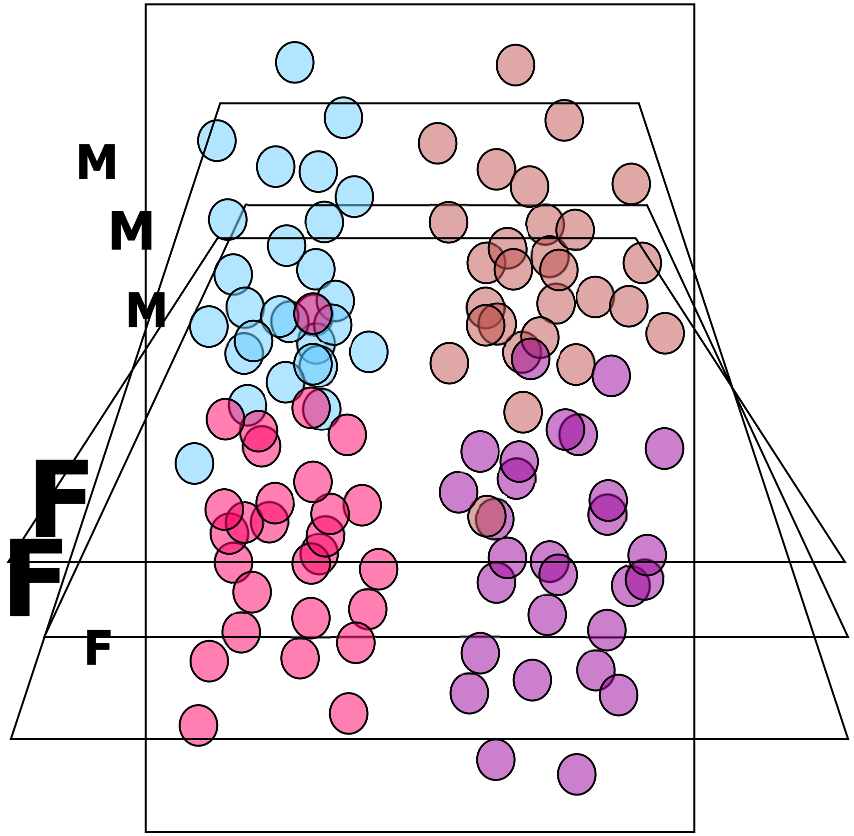
- Not all systematic variation in X is related to Y

- The 'O'-methods, OPLS and O2PLS, are able to divide the systematic variation in two parts:
  - What in X is correlated to Y – 'predictive' or shared variation
  - What in X is not correlated to Y – orthogonal variation
  - What in Y is not correlated to X – orthogonal variation

- Orthogonal variation is important information for the total understanding of the system

# What is orthogonal variation?

- Effect of interest often masked by other unwanted variation
- Orthogonal methods can rotate the projection to focus on effect of interest
- Here we want to focus on **control vs treated** but **gender** is the bigger influence on X
- OPLS rotates the PLS model so that the first OPLS component shows the between class difference




Control vs Treated

Credit: Erik Johansson / Umetrics

# Why bother with orthogonal variation?

## Improve interpretation

- Interpret 'predictive' and orthogonal effects separately

## Useful when effects cannot be controlled

- Human studies, epidemiology
- Environmental studies
- Observational studies
- Orthogonal effects akin to confounders or covariates in conventional regression
  - But confounders not necessarily orthogonal

## What about controlled experiments?

- Do not expect orthogonal variation?
- But if you find it, you will discover how to improve the experiment!

# O-PLS and O2-PLS

- Regression problem
  PLS and OPLS are unidirectional, i.e., X → Y

- Integration problem
  O2PLS is bi-directional, i.e., X ↔ Y

- Differences in preferred terminology
  OPLS: ’Predictive’ & Orthogonal variabilities
  O2PLS: Joint & Unique variabilities

# Plan

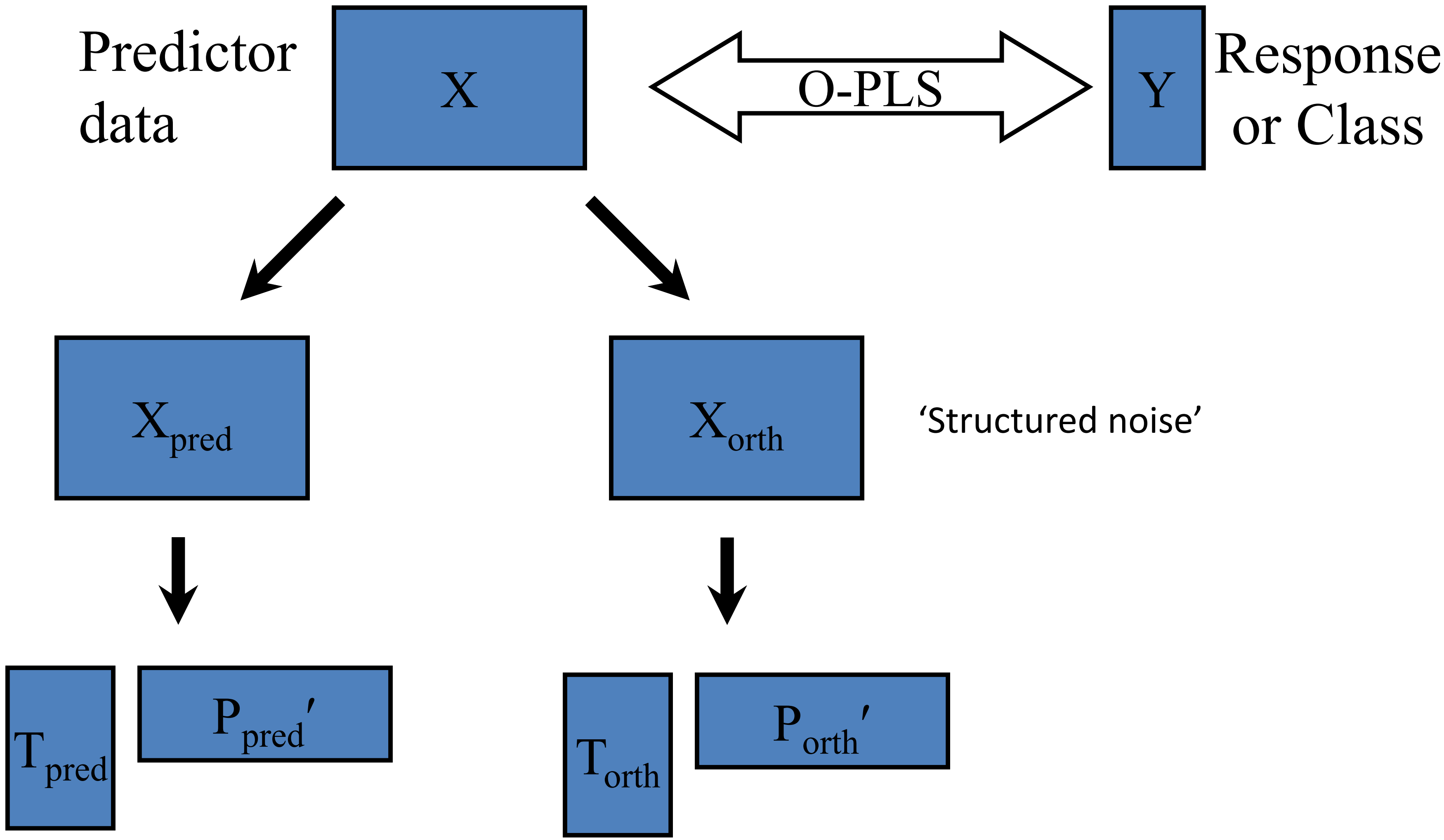
- **Partial Least Squares**
  - Geometrical overview
  - PLS-DA
- **Orthogonal variation**
  - What is it? Why bother with it?
  - How to make use of it?
- **O-PLS model**
  - Geometrical view
  - Compare with PCA/PLS
  - Exploring orthogonal variation
- **O2-PLS model**
  - Conceptual overview
  - Comparison with other methods
  - O-PLS-DA & O2-PLS for multiblock exploration

# Orthogonal PLS (O-PLS)

- Divide variation in to *predictive* and *orthogonal* components
- Same number of components & prediction ability as ordinary PLS, but…
- *Interpretation* is improved
- E.g. 6 component, 2 class PLS-DA model
  - What variables important for separating classes?
  - Interpret 6 sets of weights
- 1 pred + 5 orth O-PLS-DA model
  - Interpret *one* set of weights

Trygg, J. and Wold, S. (2002) *J. Chemometrics*, **16**, 119-128.

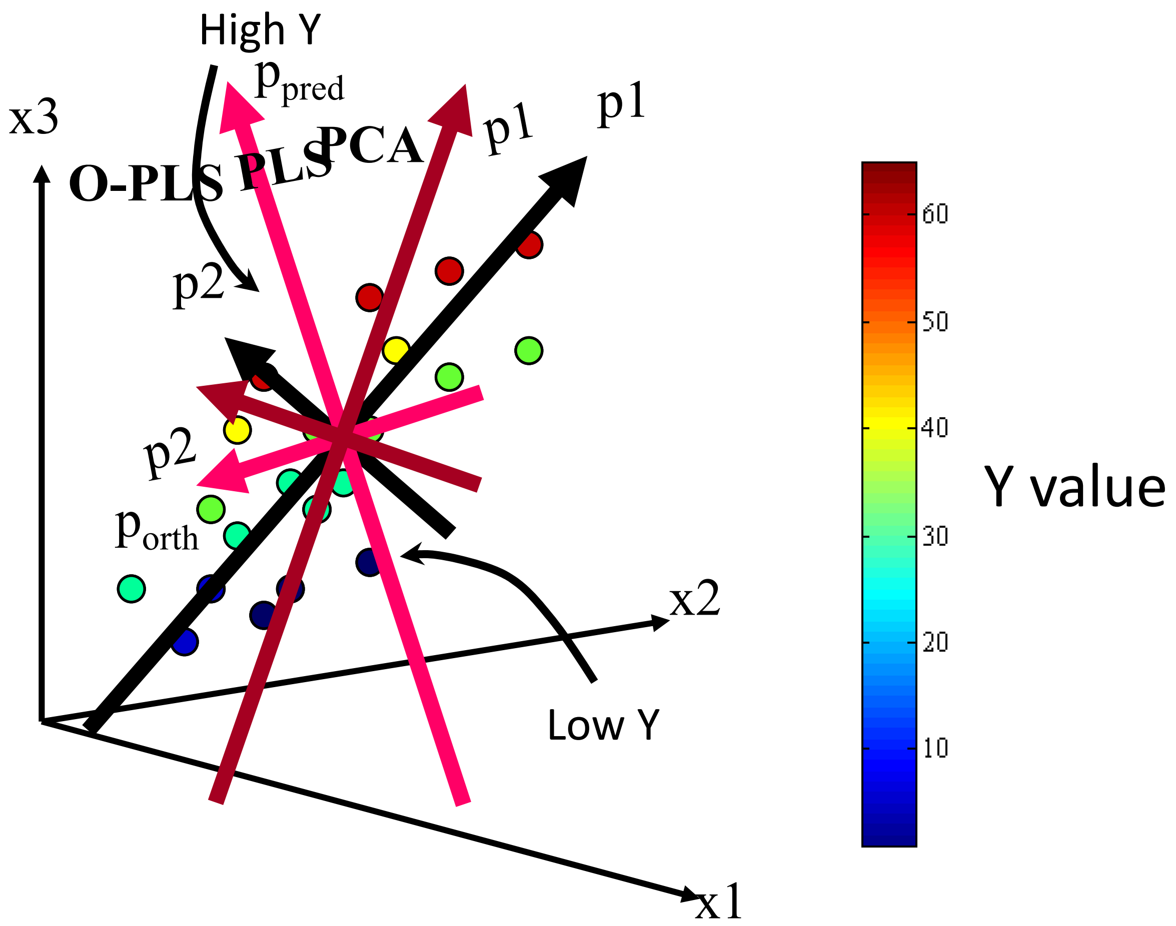
# O-PLS model structure



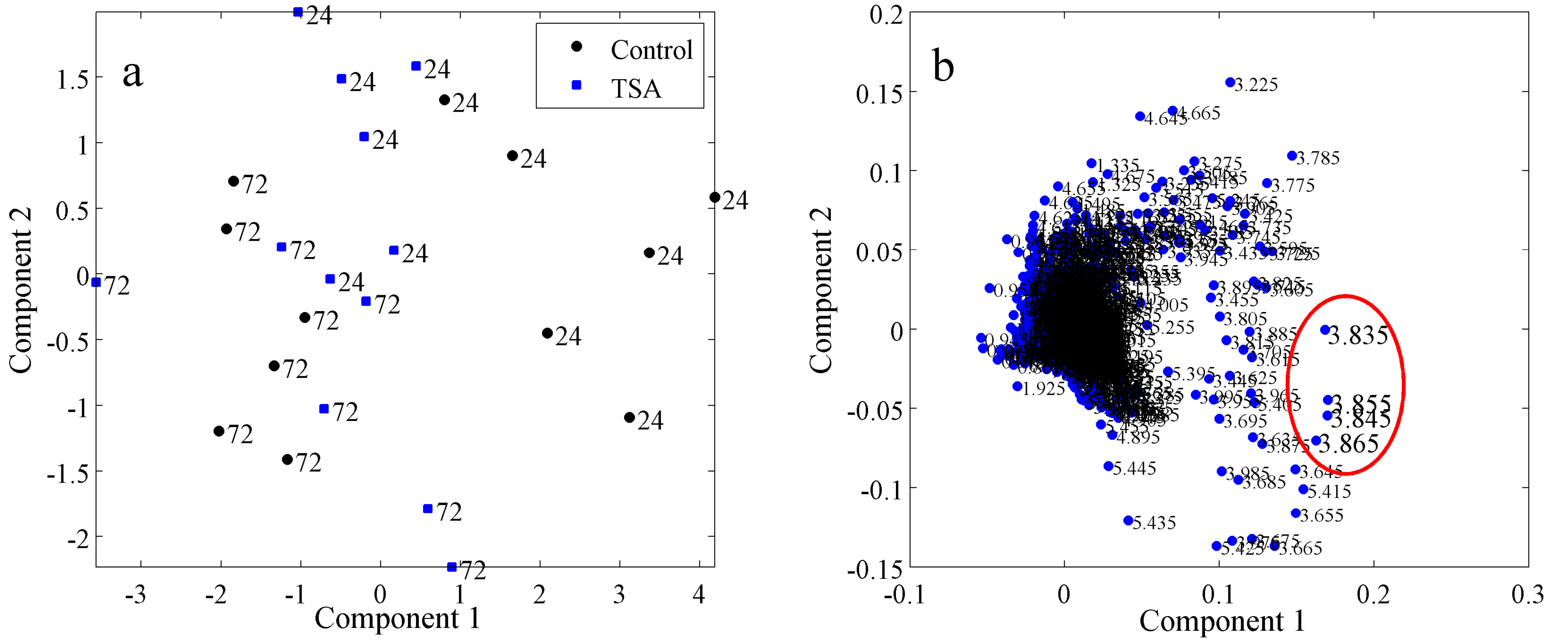
# O-PLS – geometrical view

**X space**

Points colour-coded according to Y value



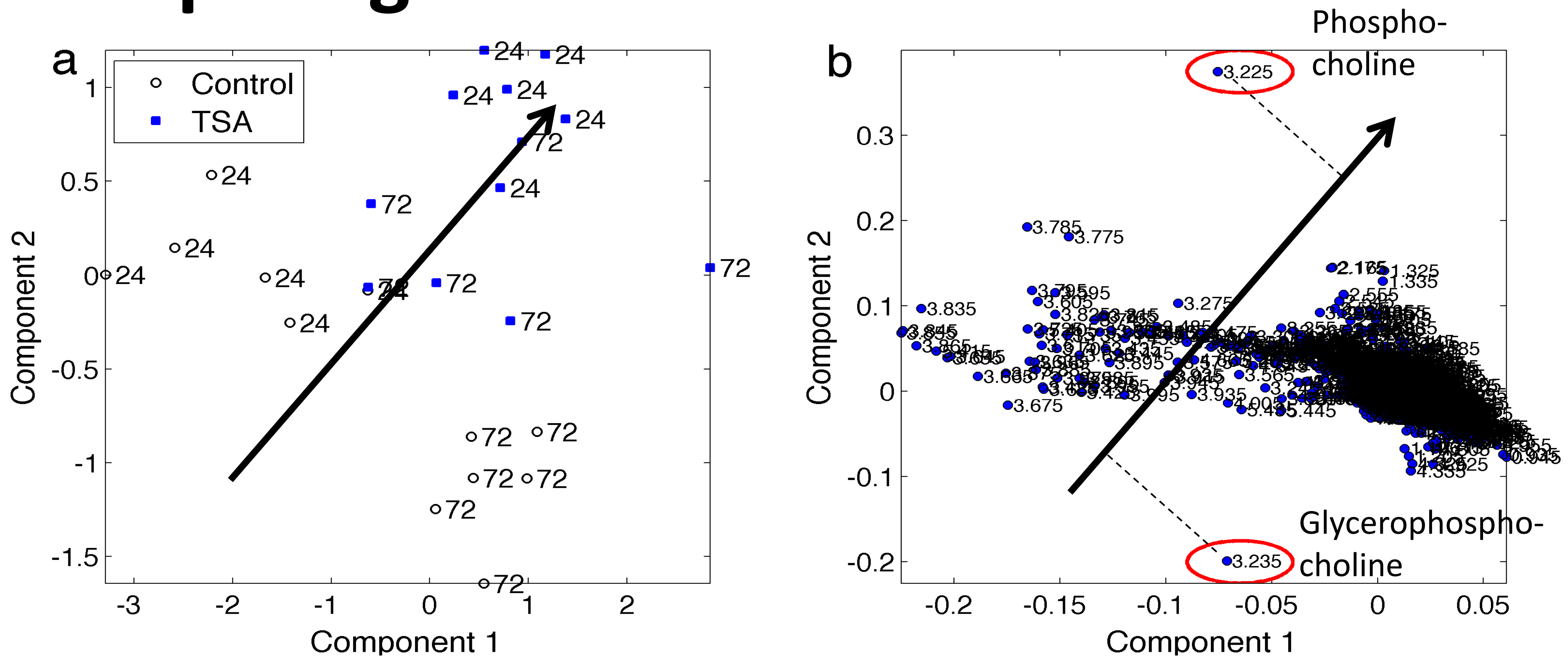
# Comparing models - PCA




- 1-D $^{1}$H NMR metabolic profiles of rat hepatocytes treated with trichostatin-A (TSA) & control.
- Some separation visible in scores – mostly time related

Credit: CARCINOGENOMICS / Ellis *et al.* J. Prot. Res. 2009

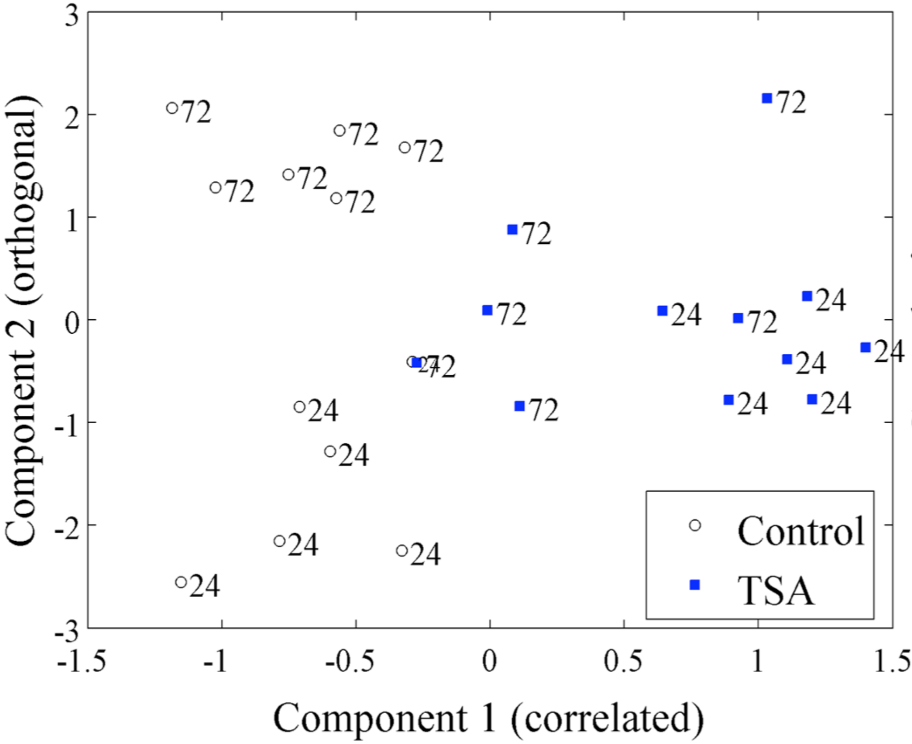
# Comparing models – PLS-DA

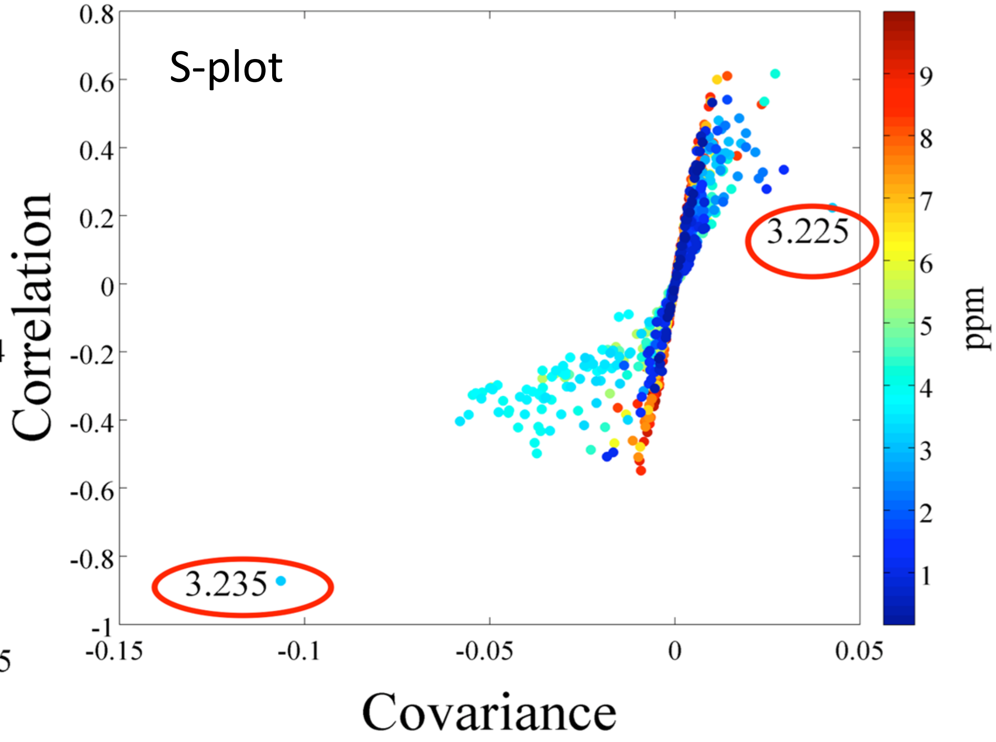



- Scores – better separation
  - But still some time related substructure
- Loadings – two clear variables related to TSA treatment

Credit: CARCINOGENOMICS / Ellis *et al.* J. Prot. Res. 2009

# Comparing models – O-PLS-DA

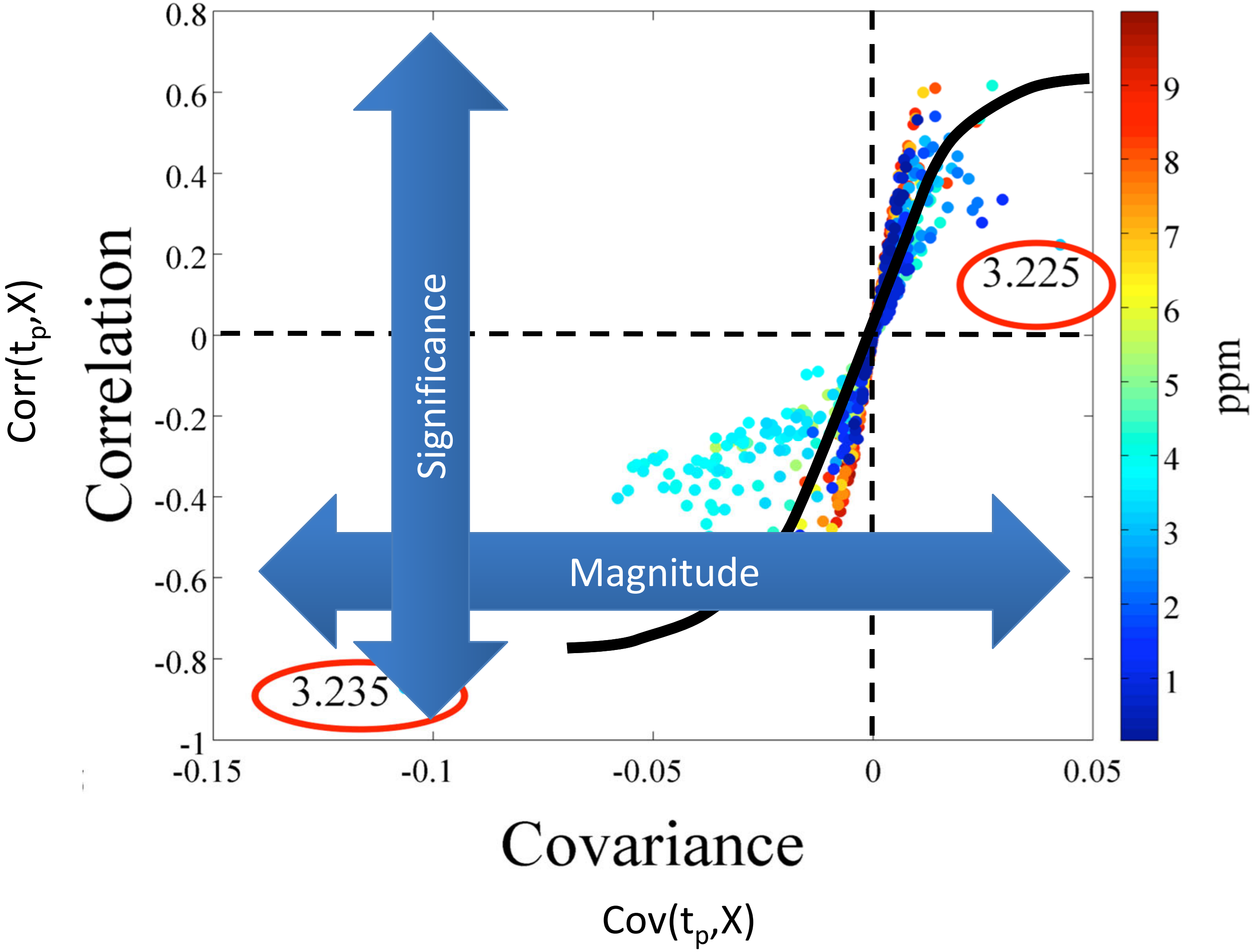



- Discrimination in ‘predictive’ component only
- Interpretation using ‘S-plot’:
  - Covariance: magnitude of change
  - Correlation: ‘reliability’ of change

Credit: CARCINOGENOMICS / Ellis *et al.* J. Prot. Res. 2009

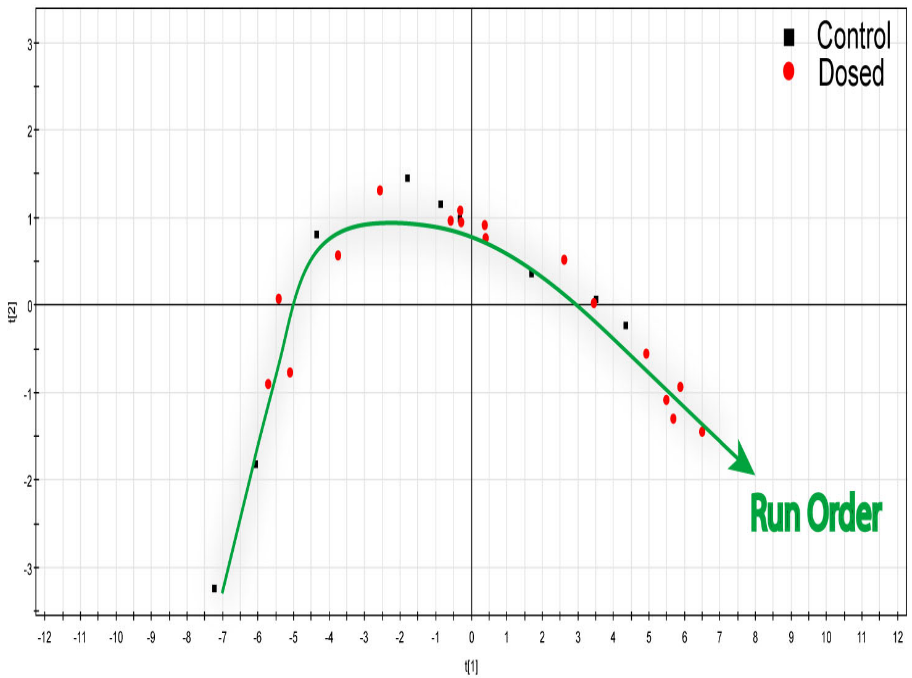
# S-Plot



Corr($t_p$,X)
Correlation
Significance
Magnitude
3.225
3.235
ppm
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-0.15
-0.1
-0.05
0
0.05
1
2
3
4
5
6
7
8
9
Covariance
Cov($t_p$,X)

# What about the orthogonal variation?

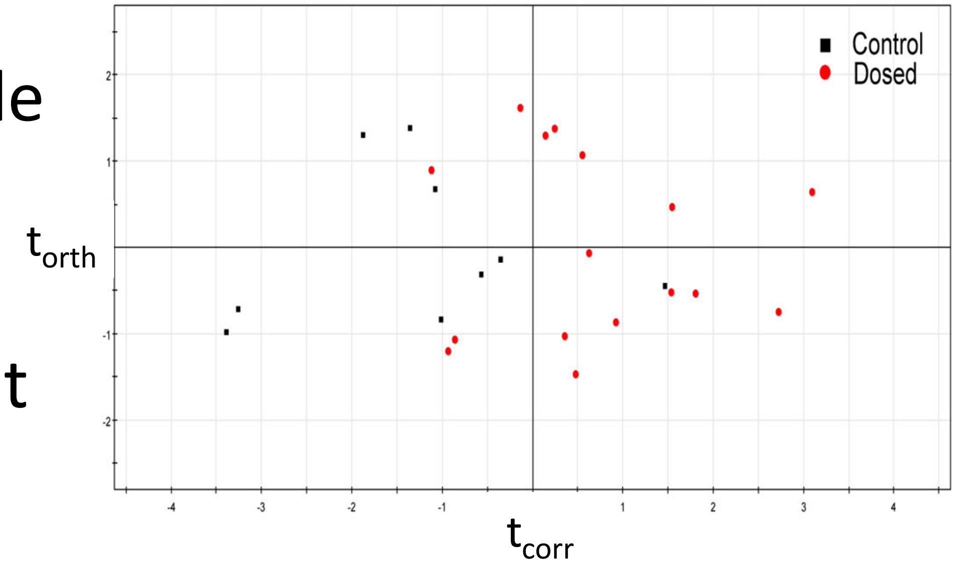
- Rats dosed with paracetamol
- Liver extracts profiled by NMR
- PCA - strong trend with run order
- Impossible to separate control & dosed classes

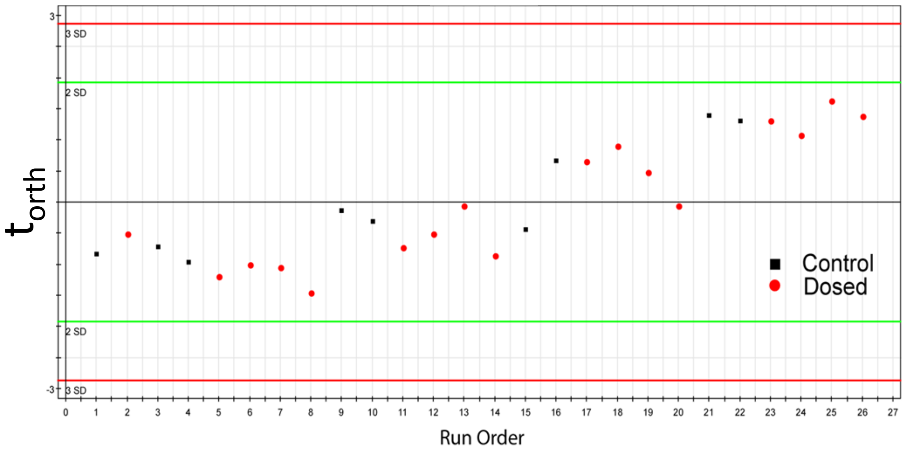



PCA

# Examining the orthogonal variation

- Separation now visible (though marginal)
- Interpret loadings on predictive component
- Orthogonal component shows trend with run order - removed



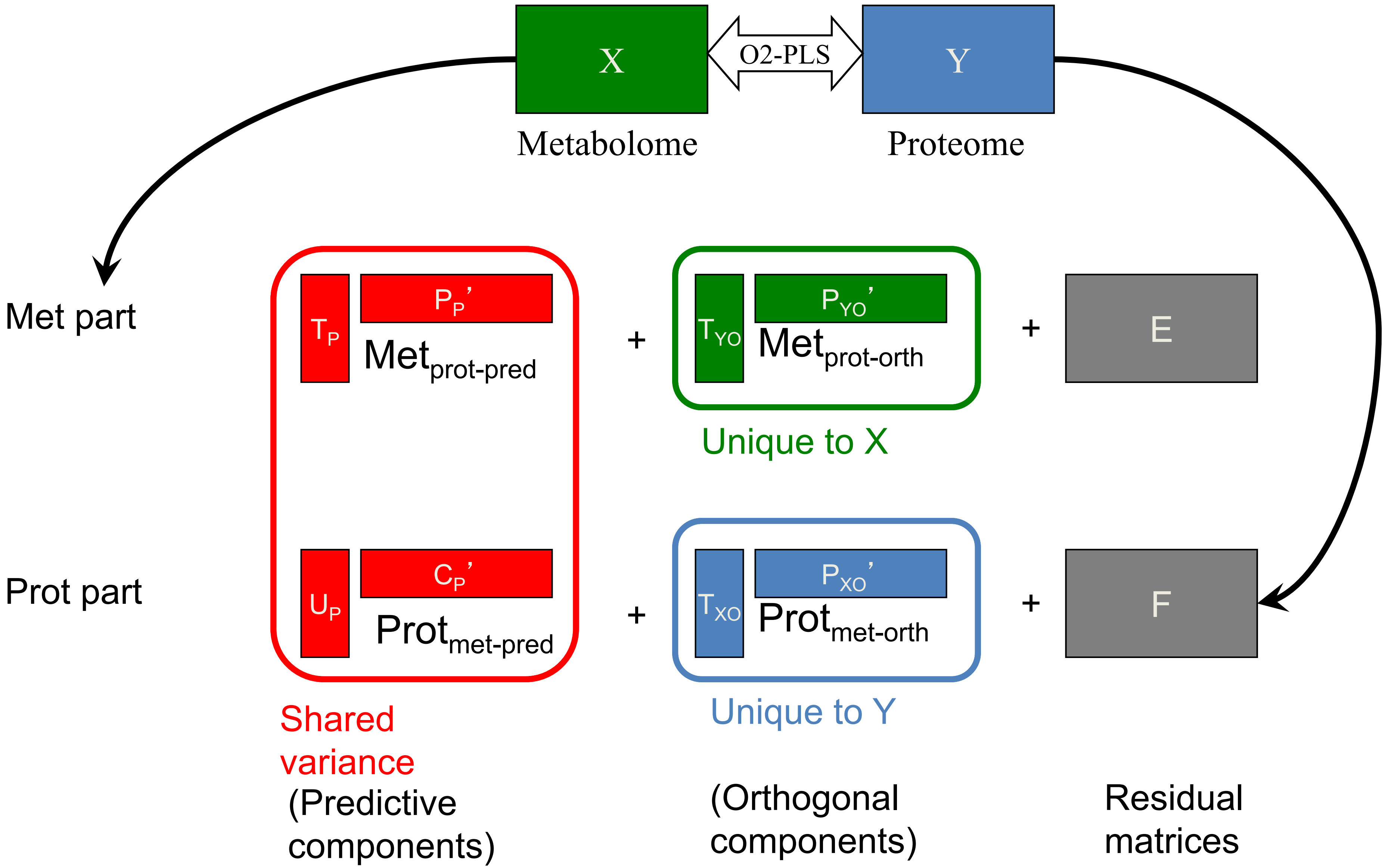
# Plan

- **Partial Least Squares**
  - Geometrical overview
  - PLS-DA
- **Orthogonal variation**
  - What is it? Why bother with it?
  - How to make use of it?
- **O-PLS model**
  - Geometrical view
  - Compare with PCA/PLS
  - Exploring orthogonal variation
- **O2-PLS model**
  - Conceptual overview
  - Comparison with other methods
  - O-PLS-DA & O2-PLS for multiblock exploration

# O2PLS

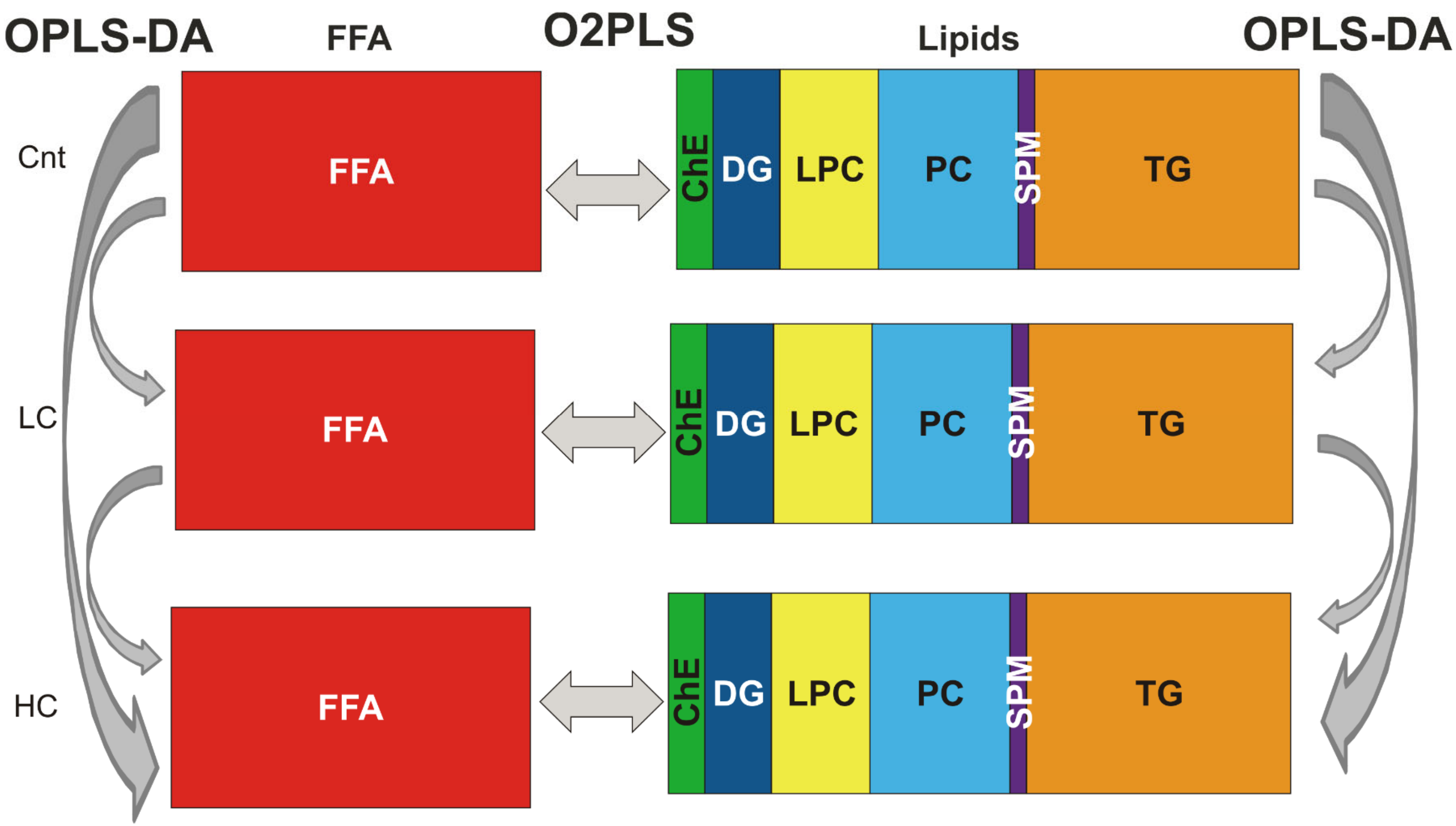
- Objective: *integration* of two data blocks (X & Y)
- What information overlaps between the two data blocks?
- What information is unique to a specific data blocks?
- Three types of components:
  - Model of joint variation
  - Model of variation unique to X
  - Model of variation unique to Y
- Improved interpretability compared to PLS

# O2-PLS model structure



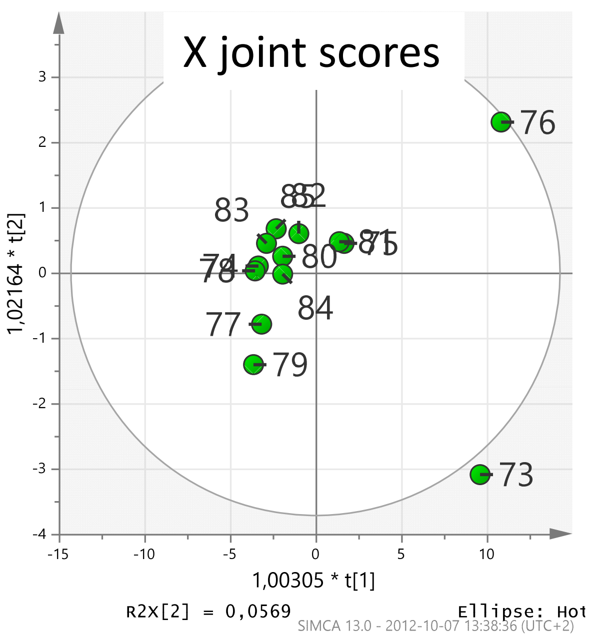
# O-PLS-DA & O2-PLS for multiblock data

- O-PLS-DA: Same variables. What is difference between sample blocks (classes)?
- O2-PLS: Same samples. What is difference between variable blocks?
- Example: integration of Free Fatty Acid (FFA) and Lipidomic data

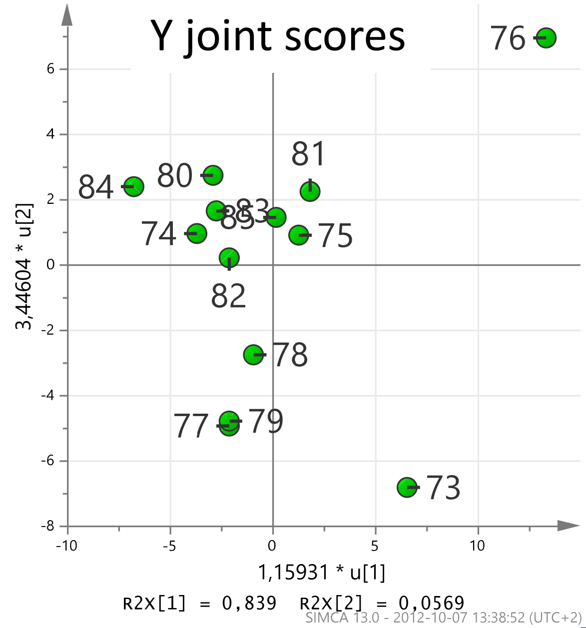



Kirwan, G.M. et al. Ana.l Chem., 2012. **84**(16): p. 7064-71.

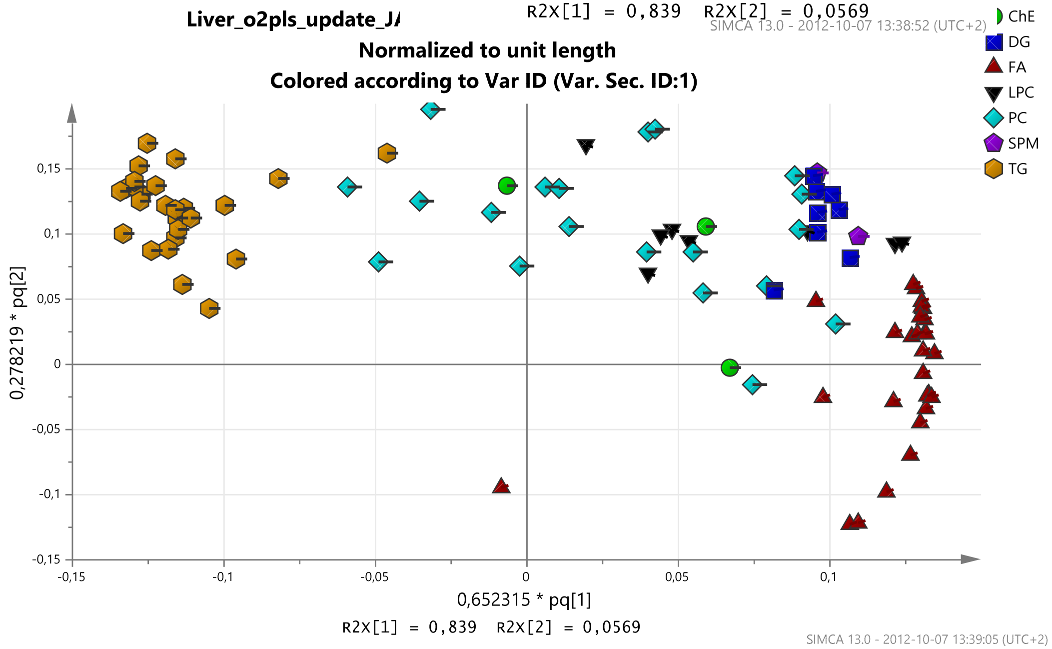
# O2-PLS Example: Joint Variation




- Joint scores give variation between samples shared by both blocks



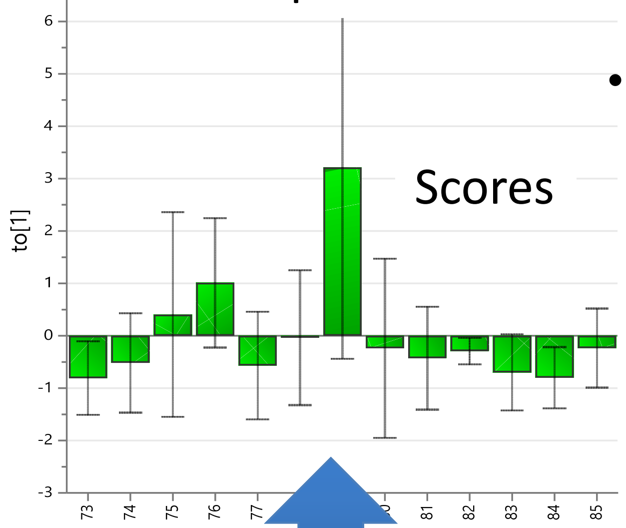

- Joint loadings help interpret joint variation in both X & Y
- Both X loadings (p) and Y loadings (q) visualised on same plot ('biplot')

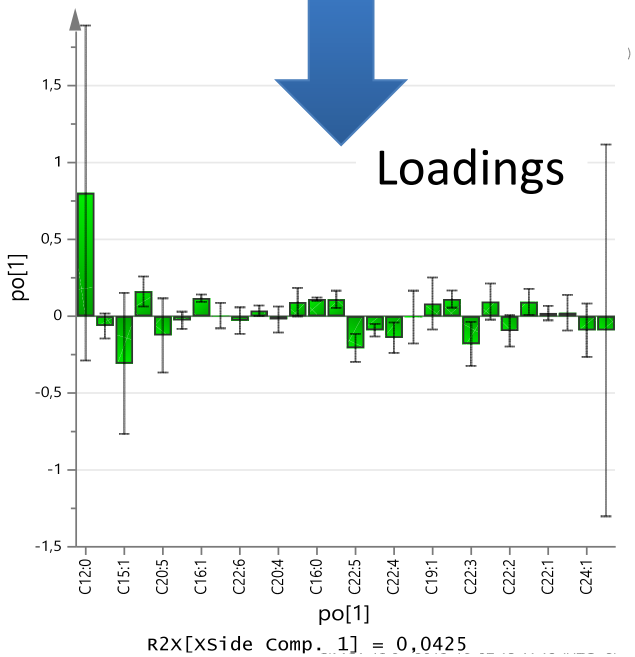



Erik Johansson / Umetrics

# O2-PLS Example: Unique Variation

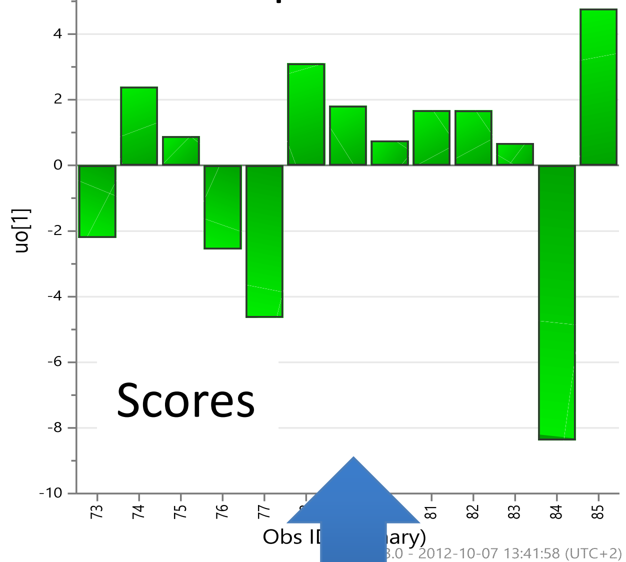
Unique to X

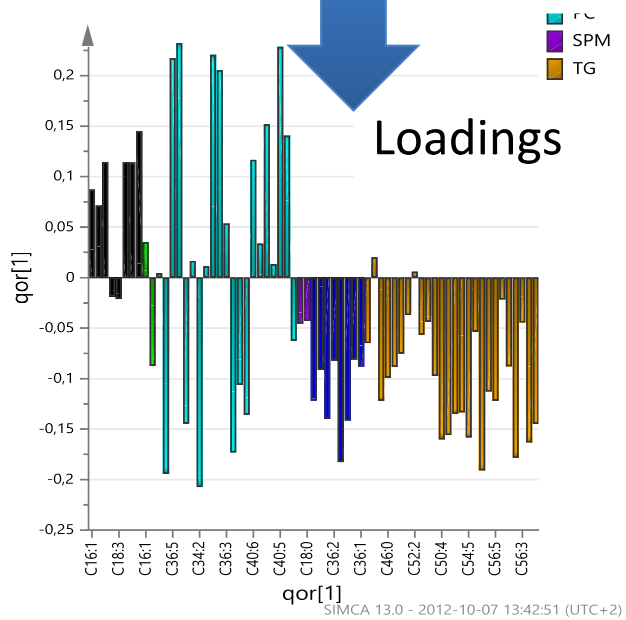



- Unique scores give variation between samples specific to each block
- Unique loadings help explain unique variation in each block

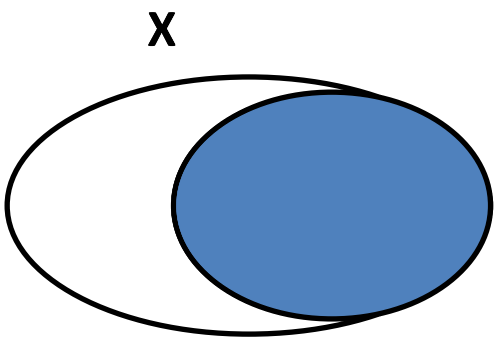
Unique to Y



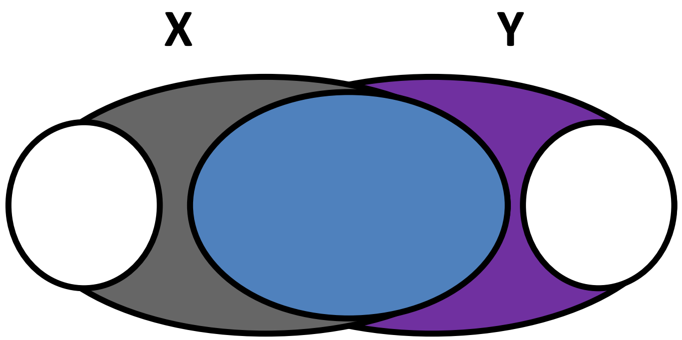
## PCA

All variation in one joint model
Remaining variation is seen as white noise



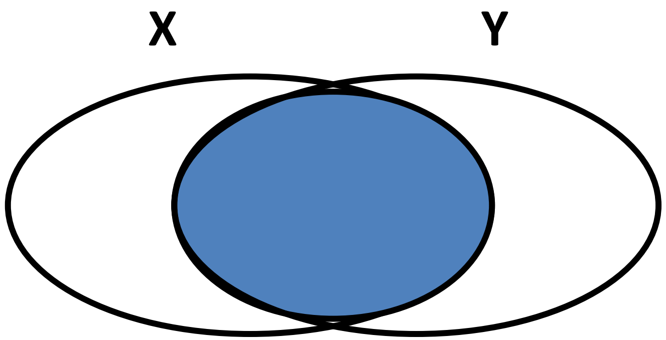

## O2PLS

Joint variation in one model in addition to a model of unique variation for each of X and Y.
Remaining variation is seen as noise



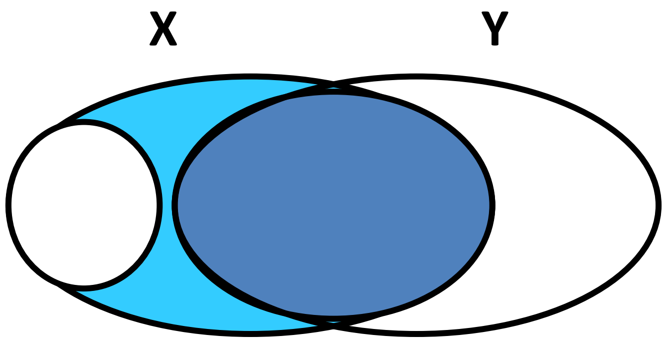

## PLS

Predicts Y from X
Remaining variation is seen as white noise




## OPLS

Predicts Y from X
Remaining variation is seen as white noise for Y
X is divided into orthogonal information and noise




Erik Johansson / Umetrics

# O-PLS, O2-PLS - summary

- Model variation in each block which is *orthogonal* to the other block(s)
- '*Predictive*' and *orthogonal* variation (O-PLS)
- *Shared* and *unique* variation (O2-PLS, On-PLS)
- Improve *interpretation* of models

# O2PLS & OPLS References


- Trygg, J. & S. Wold, J. Chemometrics 2002, 16, 283−293.
- Trygg, J. & S. Wold, J. Chemometrics, 2003. 17, 53-64.
- Wiklund, S. et al. Anal. Chem. 2008, 80, 115−122.
- Bylesjö M. et al. Plant J. 2007, 52, 1181−1191.
- Bylesjö M., J. Prot. Res. 2009, 8, 199−210.
- Rantalainen, M., et al. J Prot. Res, 2006, 5, 2642-55.
- Richards, S. E. et al. Chemom. Intell. Lab. Syst. 2010, 104, 121-131
- Löfstedt, T. and J. Trygg, 2011, 25, 441-455.