



Third edition 2026 in Fréjus



MOFA: Multi-omic Factor Analysis

Jimmy Vandel, Arnaud Gloaguen

Method | 20 June 2018 |  OPEN ACCESS

 TRANSPARENT PROCESS

Multi-Omics Factor Analysis—a framework for unsupervised integration of multi-omics data sets

Ricard Argelaguet , Britta Velten , Damien Arnol , Sascha Dietrich , Thorsten Zenz ,
John C Marioni , Florian Buettner  ✉, Wolfgang Huber  ✉, Oliver Stegle  ✉

[Author Information](#)

Molecular Systems Biology (2018) 14: e8124 | <https://doi.org/10.15252/msb.20178124>

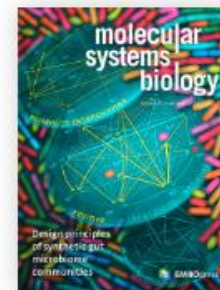
Method | [Open access](#) | [Published: 11 May 2020](#)

MOFA+: a statistical framework for comprehensive integration of multi-modal single-cell data

[Ricard Argelaguet](#) ✉, [Damien Arnol](#), [Danila Bredikhin](#), [Yonatan Deloro](#), [Britta Velten](#), [John C. Marioni](#) ✉ & [Oliver Stegle](#) ✉

Genome Biology **21**, Article number: 111 (2020) | [Cite this article](#)

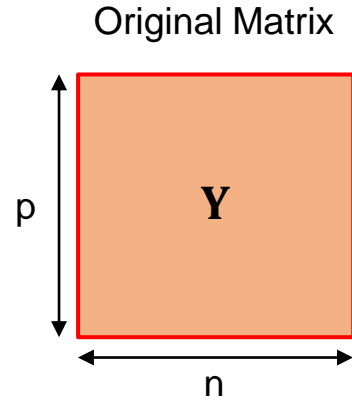
44k Accesses | 237 Citations | 118 Altmetric | [Metrics](#)

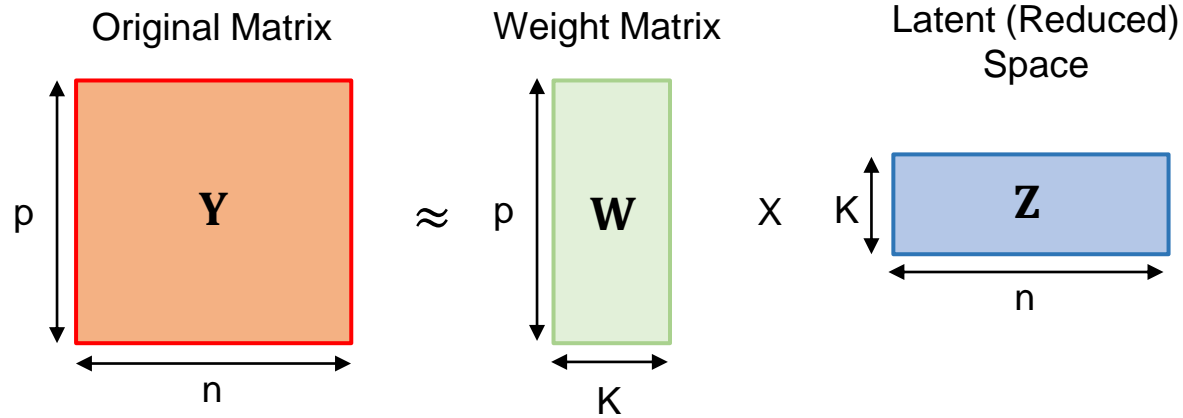


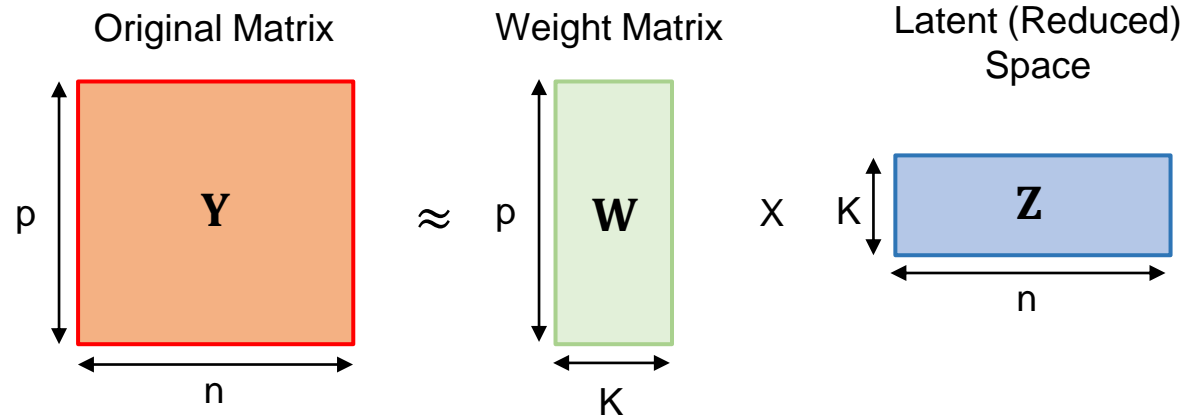
[About the cover](#)

Volume 14
Issue 6
1 June 2018

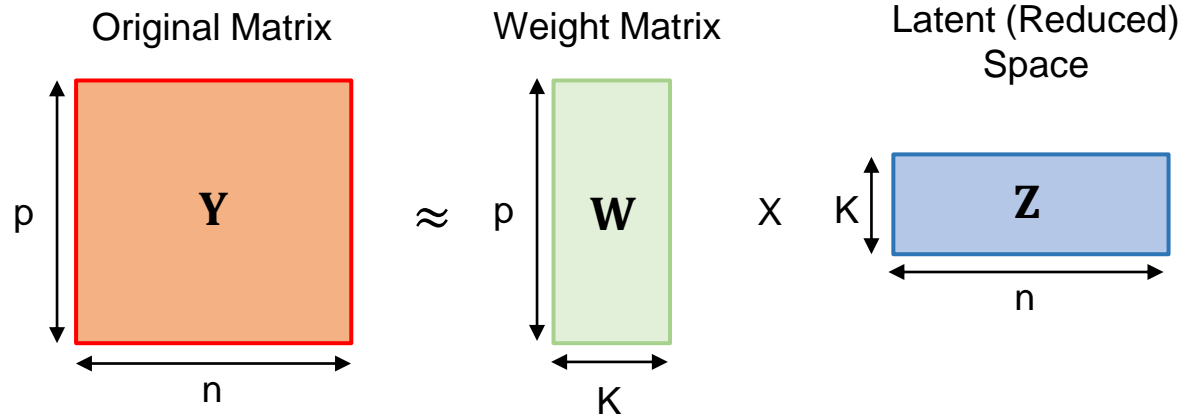
[IN THIS ISSUE](#)

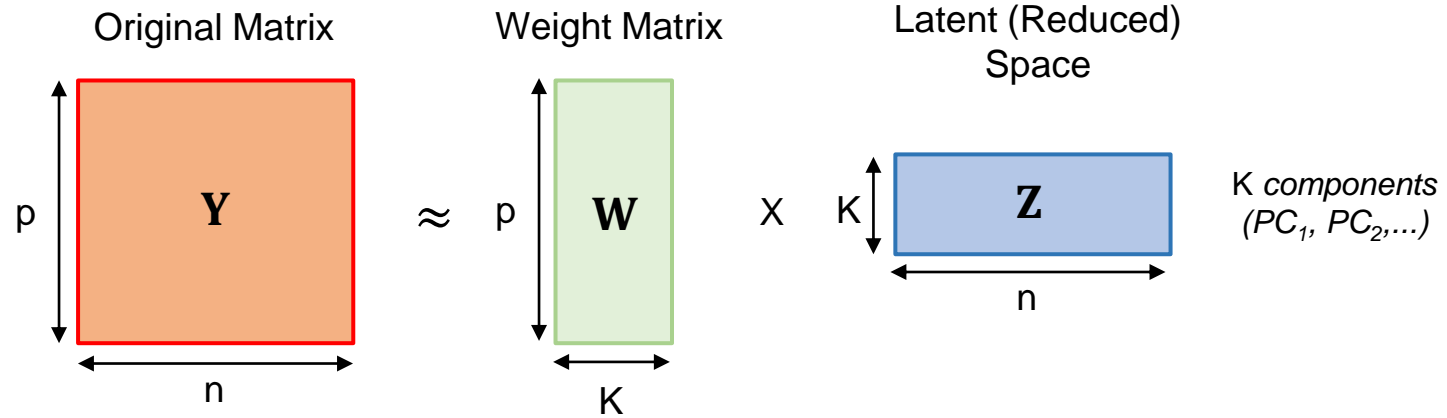


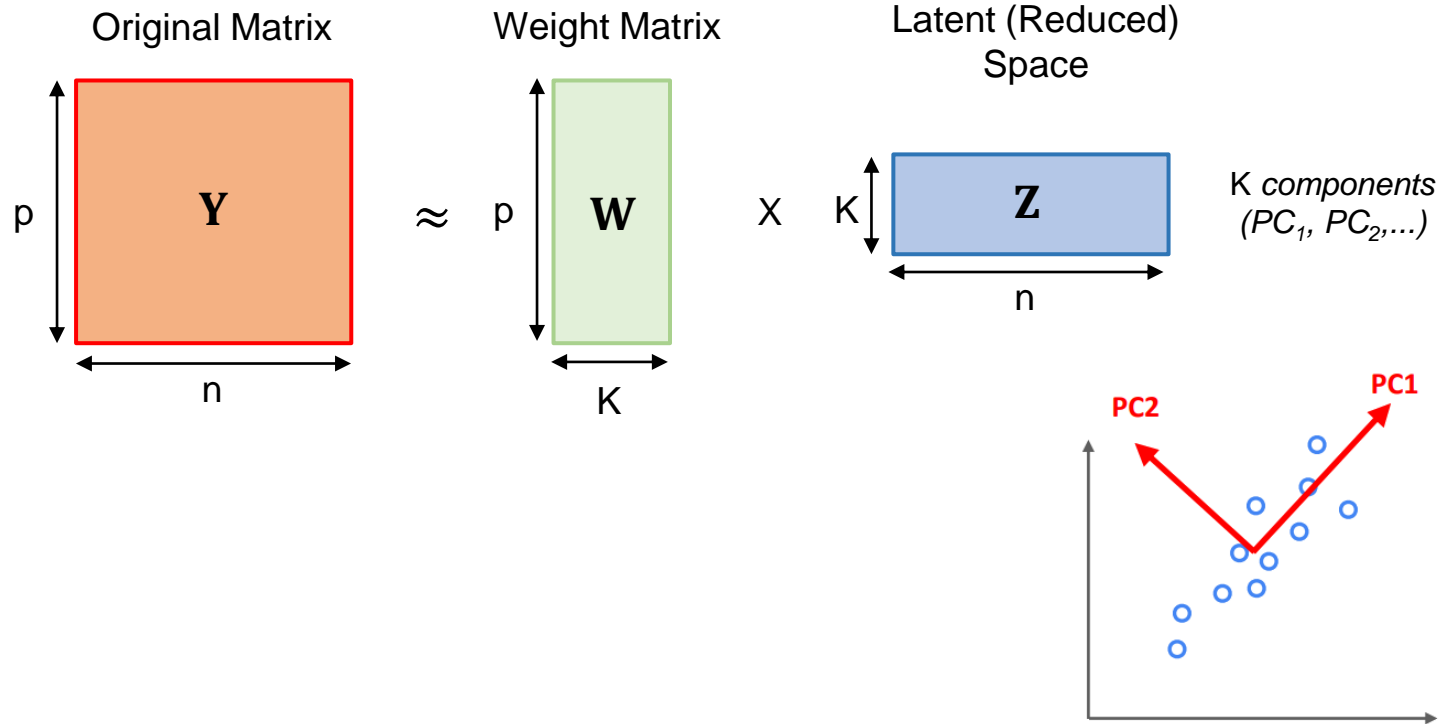


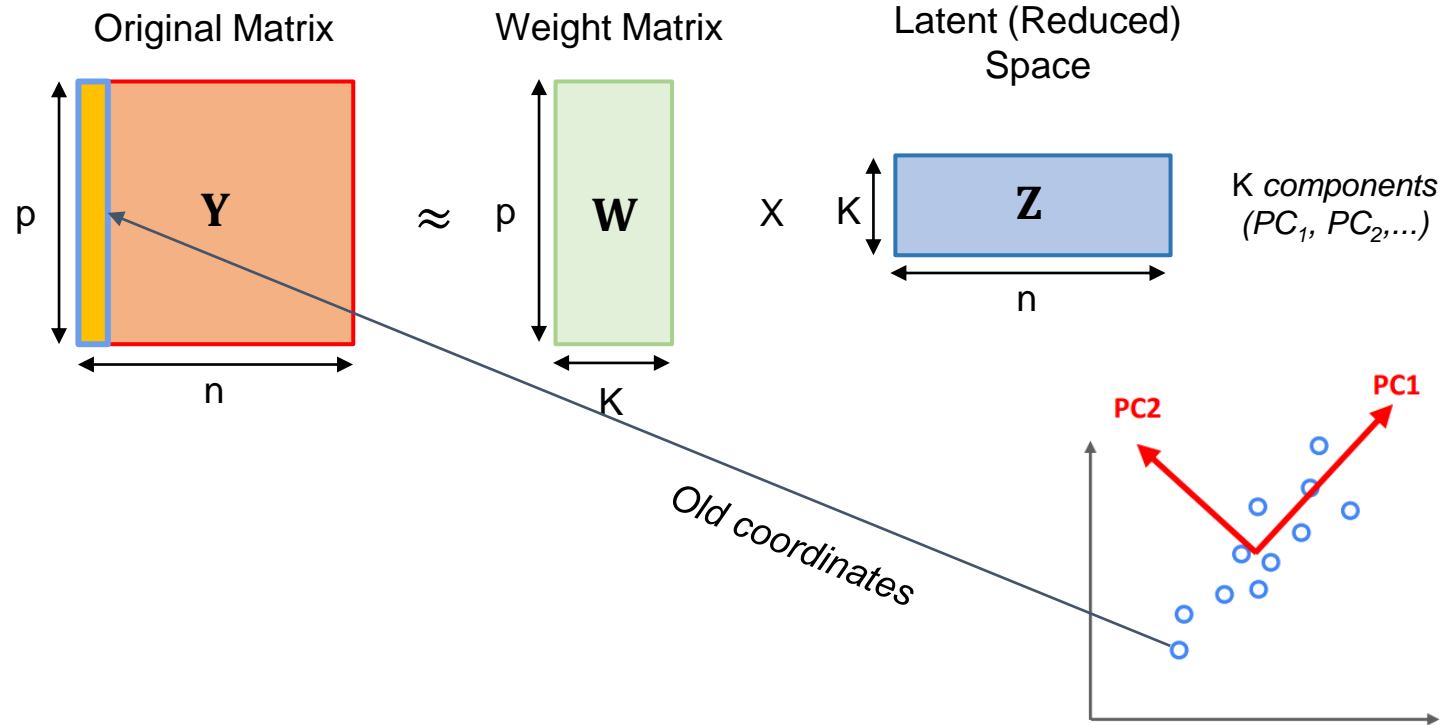


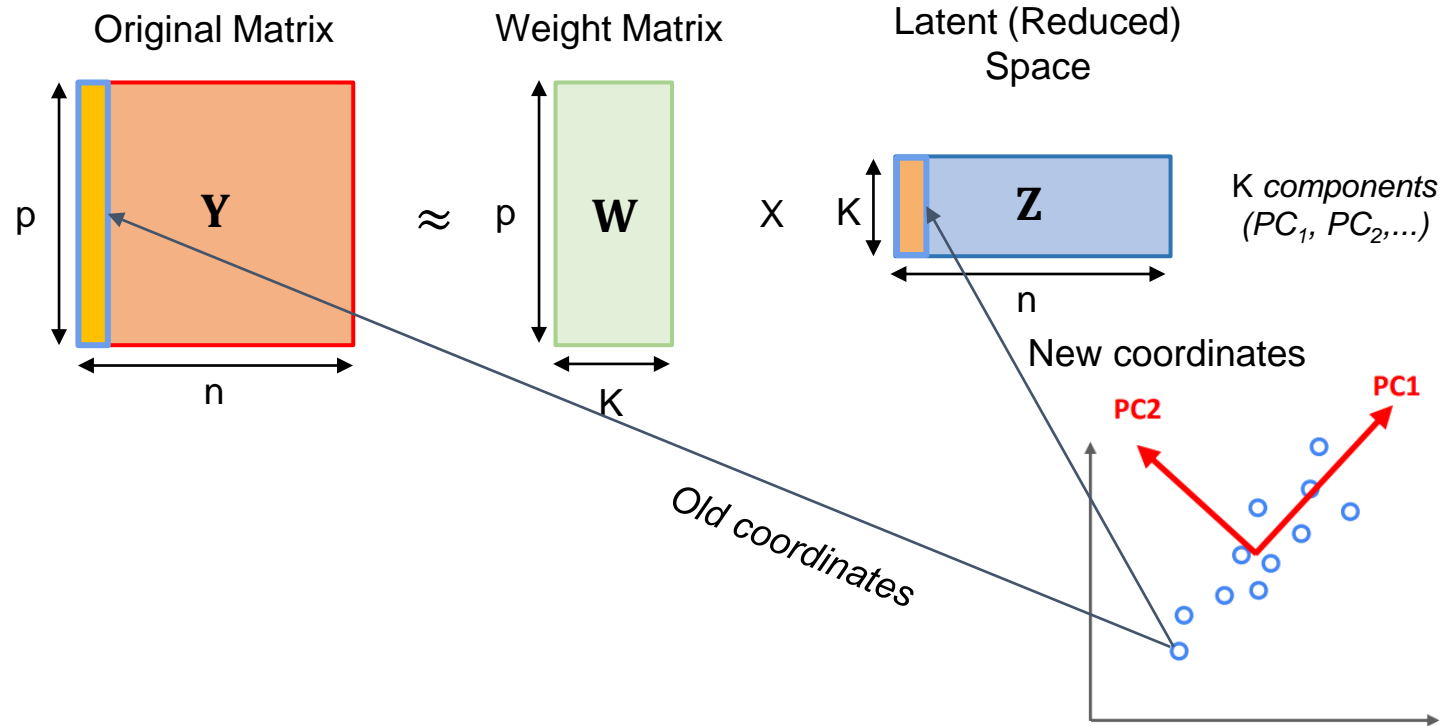
→ Interests: Dimension Reduction + Noise Reduction

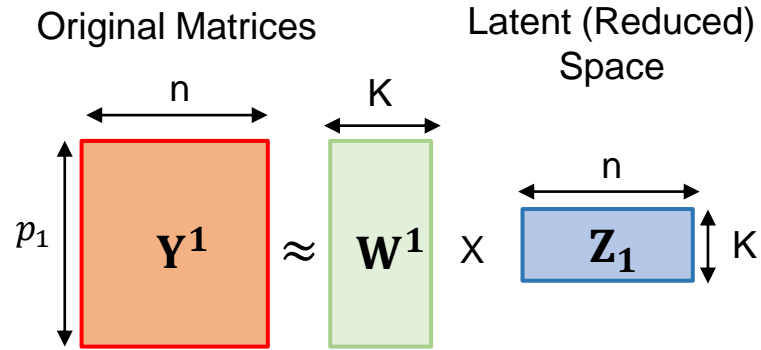


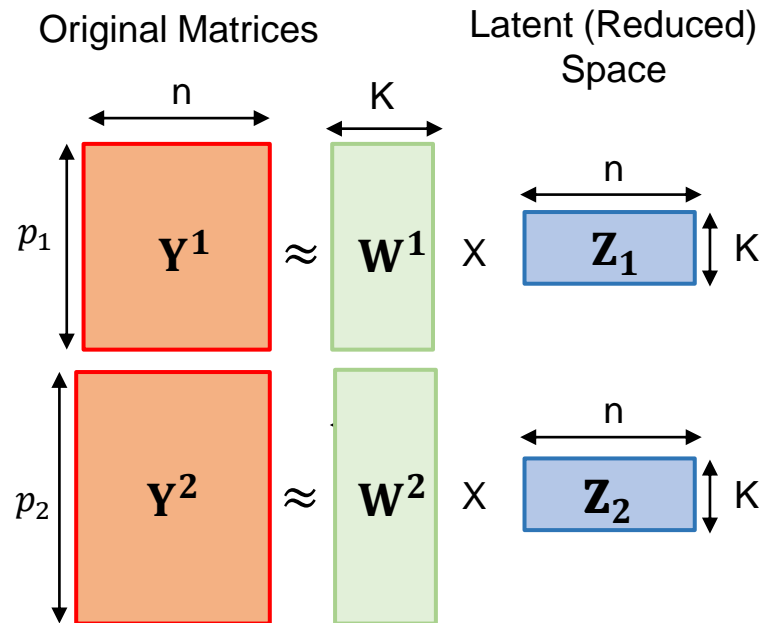


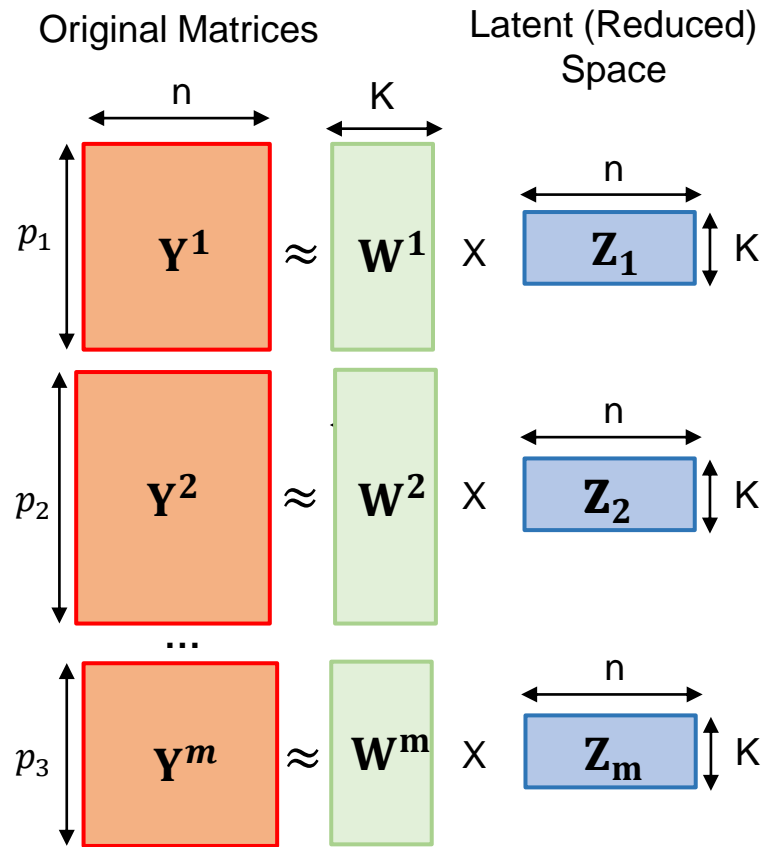


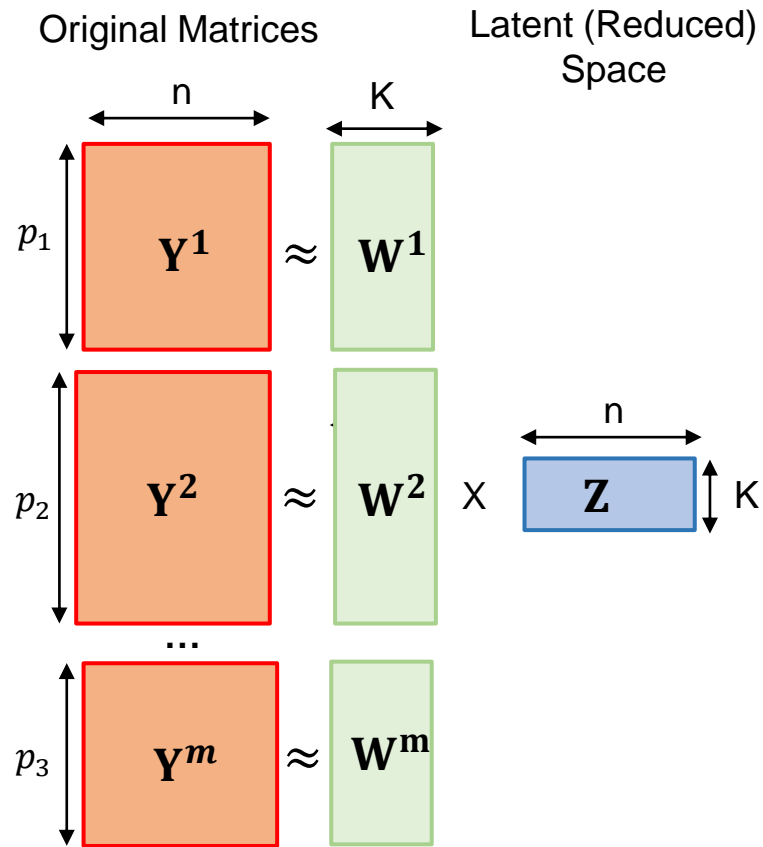


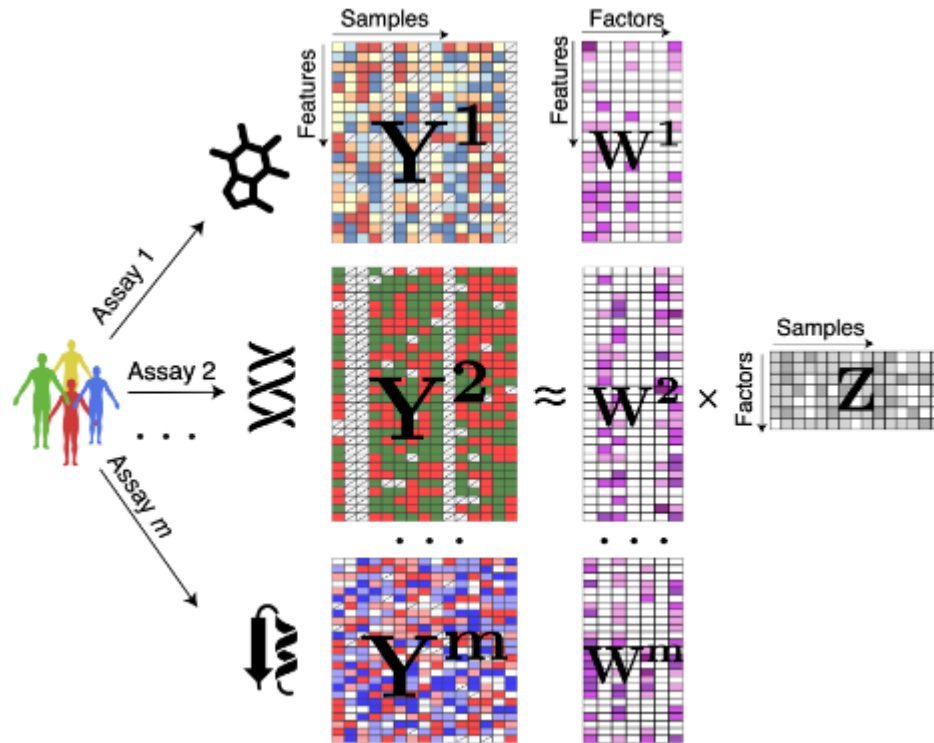


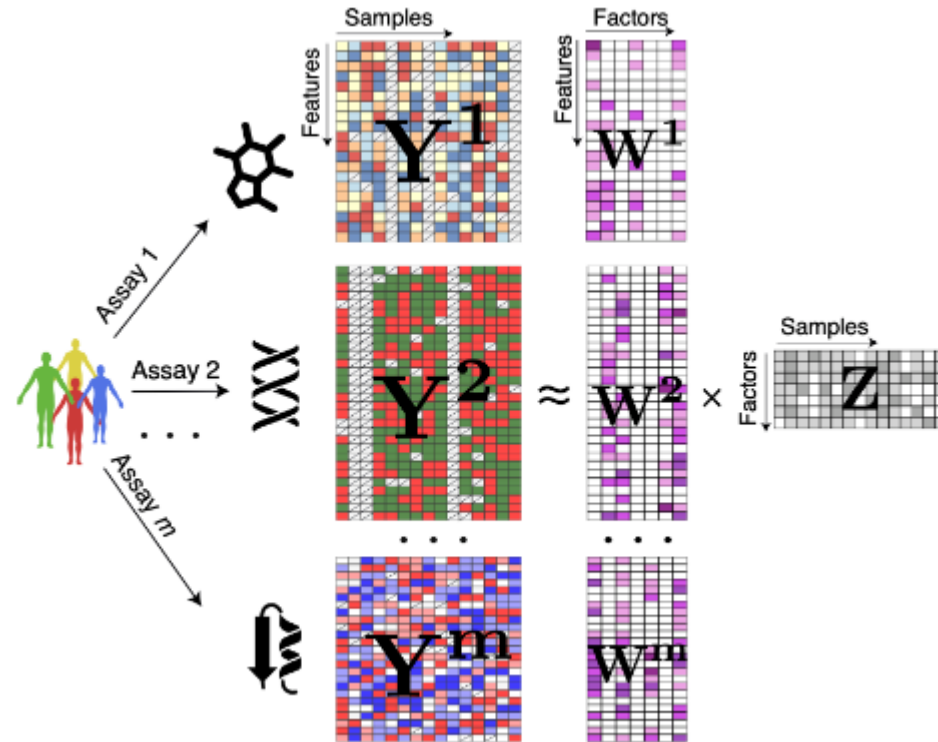




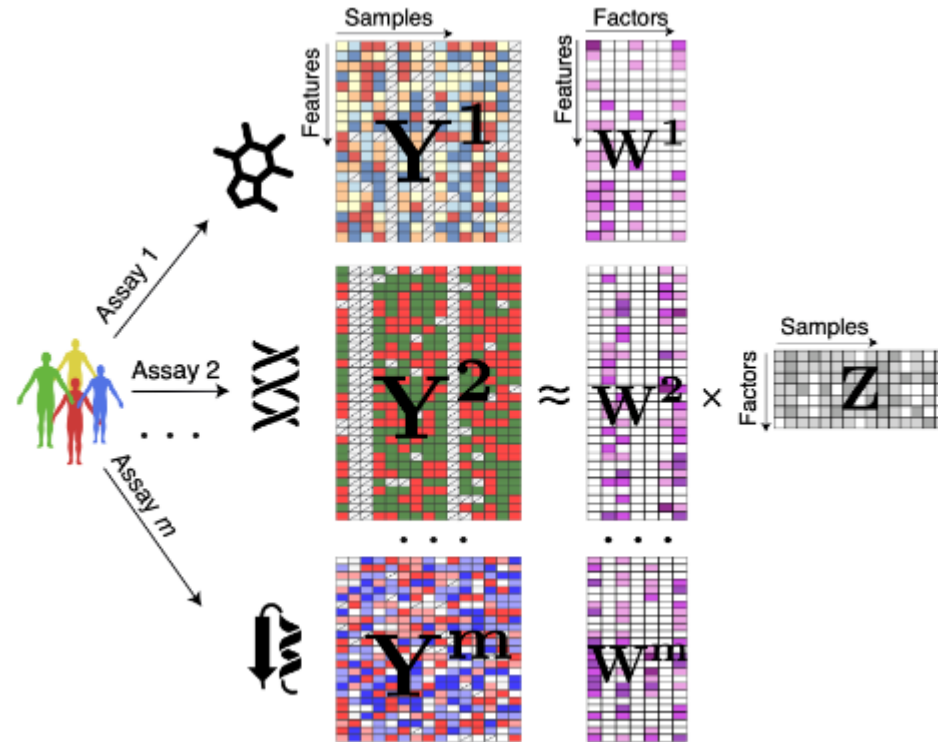






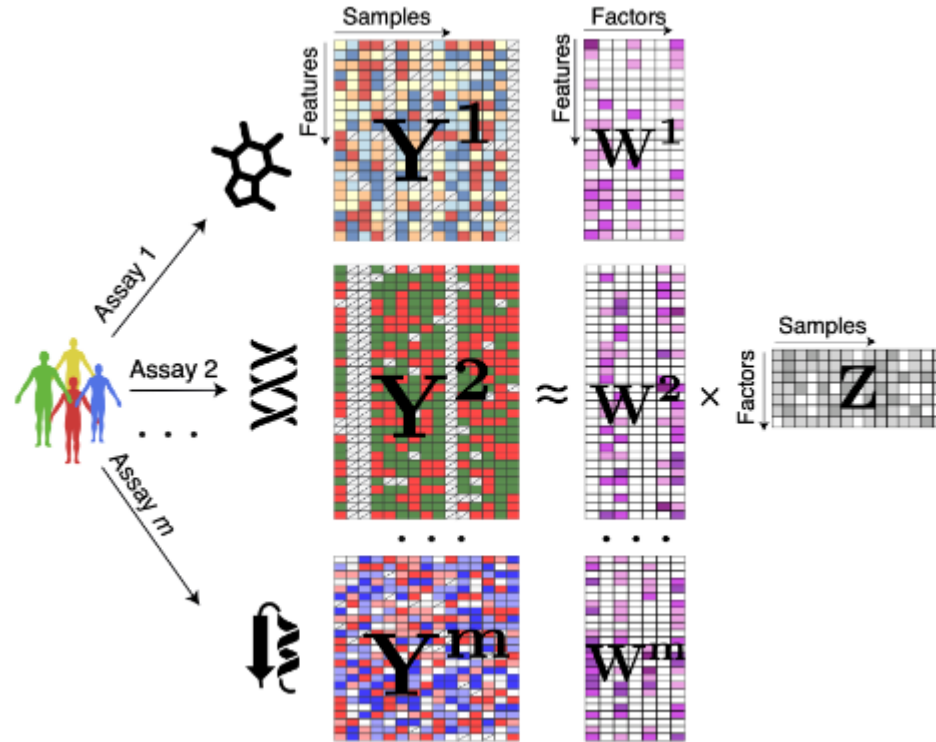


$$Y^m = W^m Z + \epsilon^m$$



$$Y^m = W^m Z + \epsilon^m$$

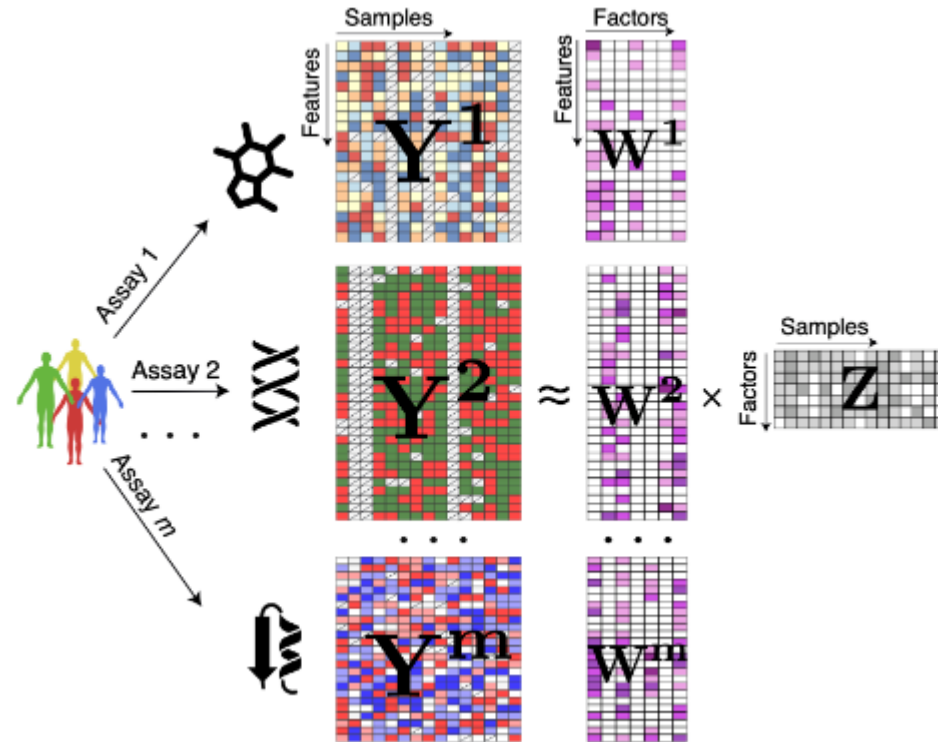
Main differences with PCA :



$$Y^m = W^m Z + \epsilon^m$$

Main differences with PCA :

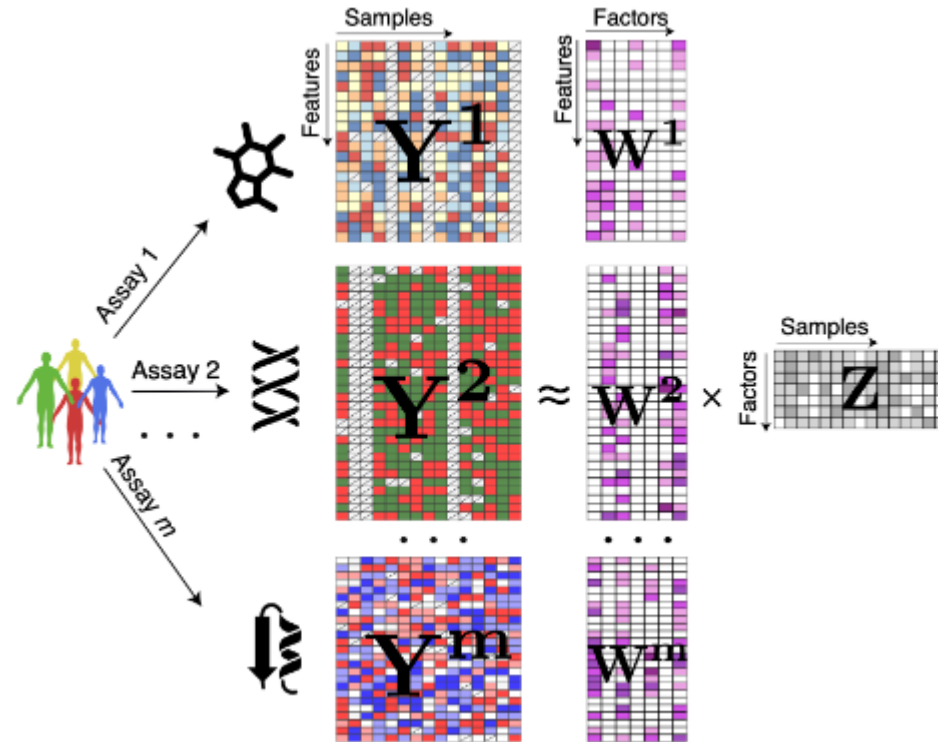
- m views for m omic sources with a share the Z matrix between views



$$Y^m = W^m Z + \epsilon^m$$

Main differences with PCA :

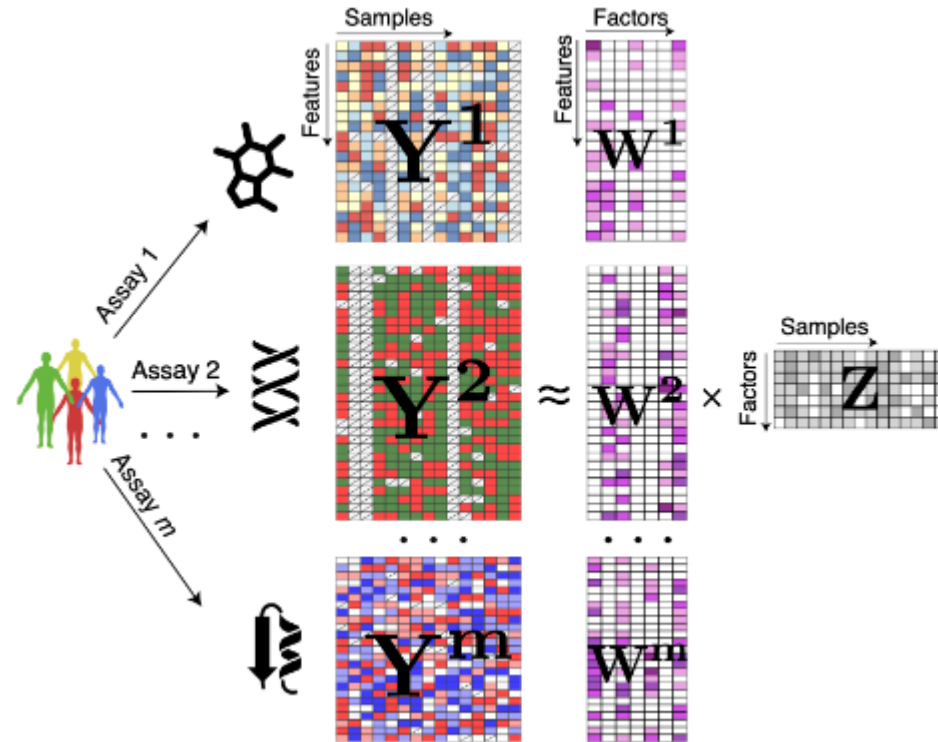
- m views for m omic sources with a share the Z matrix between views
→ extract common effects across omics



$$\mathbf{Y}^m = \mathbf{W}^m \mathbf{Z} + \epsilon^m$$

Main differences with PCA :

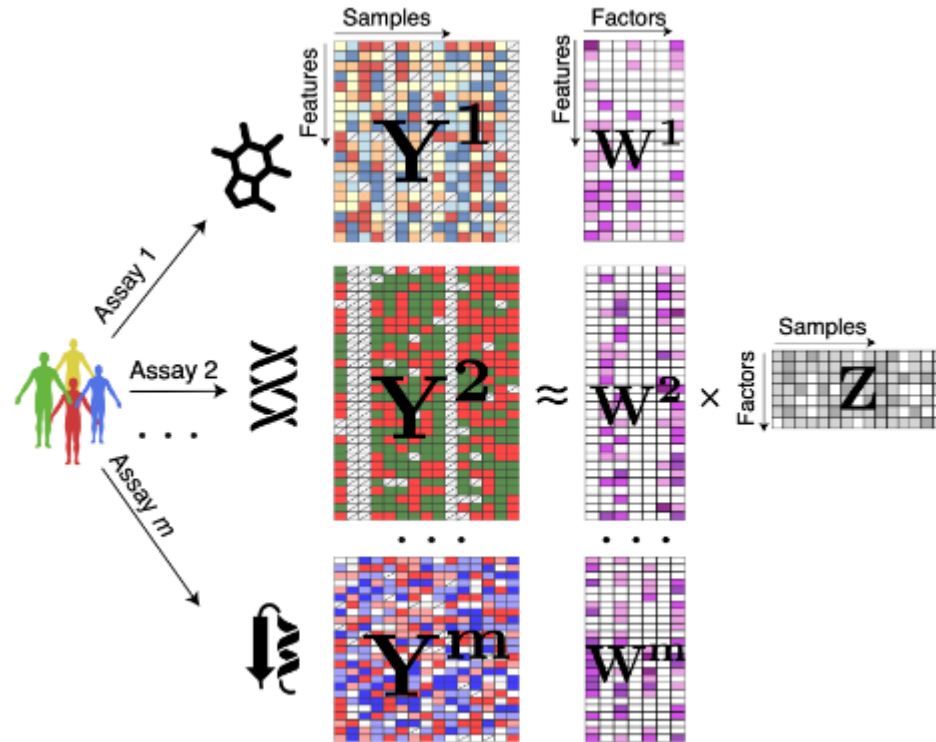
- m views for m omic sources with a share the \mathbf{Z} matrix between views
→ extract common effects across omics
- Bayesian Framework



$$\mathbf{Y}^m = \mathbf{W}^m \mathbf{Z} + \epsilon^m$$

Main differences with PCA :

- m views for m omic sources with a share the \mathbf{Z} matrix between views
→ extract common effects across omics
- Bayesian Framework
- No orthogonality imposed for \mathbf{W}^m or \mathbf{Z}



$$\mathbf{Y}^m = \mathbf{W}^m \mathbf{Z} + \epsilon^m$$

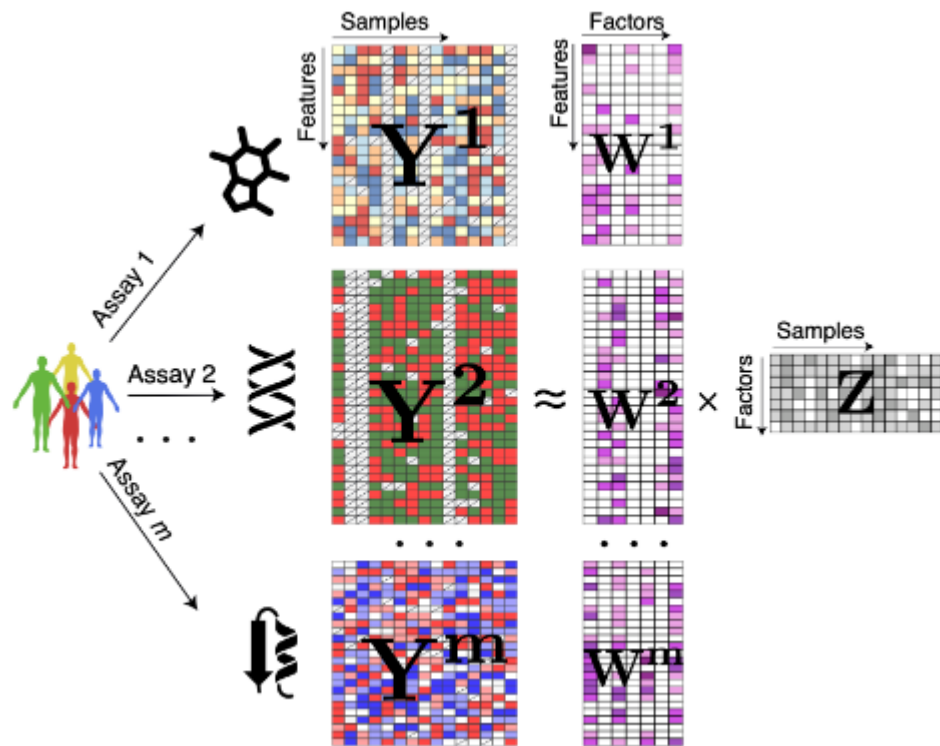
Main differences with PCA :

- m views for m omic sources with a share the \mathbf{Z} matrix between views
→ extract common effects across omics
- Bayesian Framework
- No orthogonality imposed for \mathbf{W}^m or \mathbf{Z}
- Sparsity on \mathbf{W}^m :

$$\mathbf{Y}^m = \mathbf{W}^m \mathbf{Z} + \epsilon^m$$

Main differences with PCA :

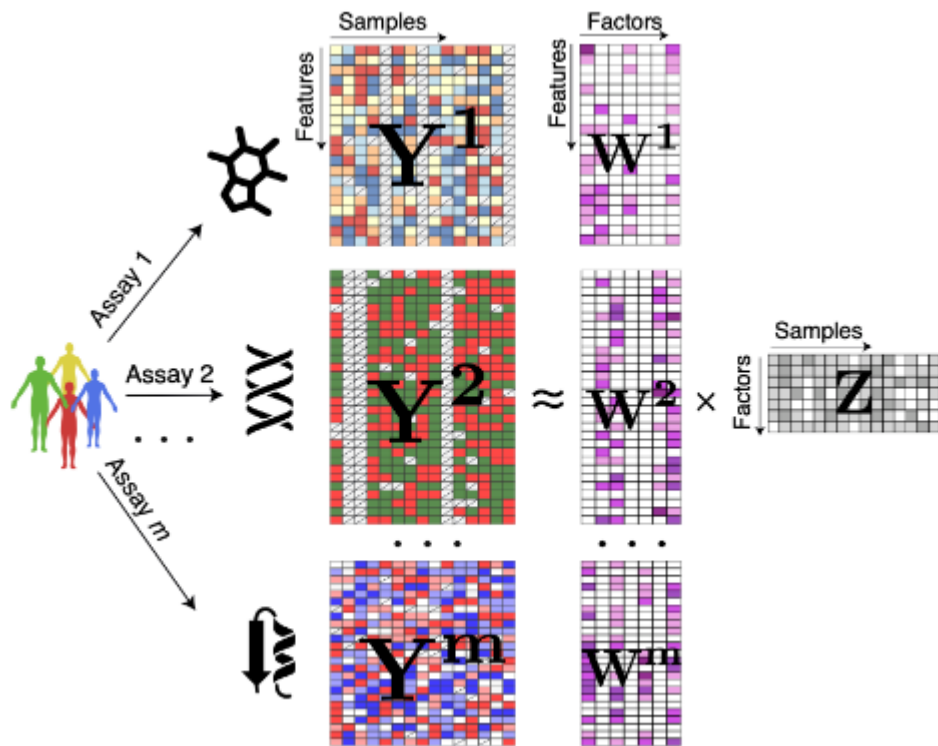
- m views for m omic sources with a share the \mathbf{Z} matrix between views
→ extract common effects across omics
- Bayesian Framework
- No orthogonality imposed for \mathbf{W}^m or \mathbf{Z}
- Sparsity on \mathbf{W}^m :
→ sparse biological phenomenon

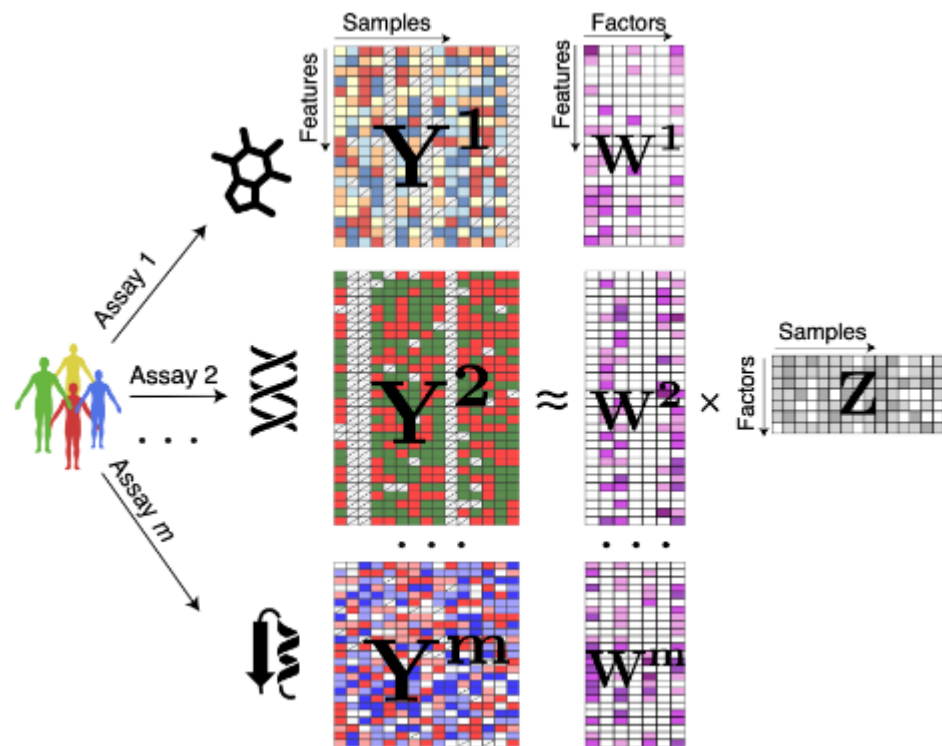


$$\mathbf{Y}^m = \mathbf{W}^m \mathbf{Z} + \epsilon^m$$

Main differences with PCA :

- m views for m omic sources with a share the \mathbf{Z} matrix between views
→ extract common effects across omics
- Bayesian Framework
- No orthogonality imposed for \mathbf{W}^m or \mathbf{Z}
- Sparsity on \mathbf{W}^m :
→ sparse biological phenomenon
- $\mathbf{Y}^m / \epsilon^m$ can follow different distributions :

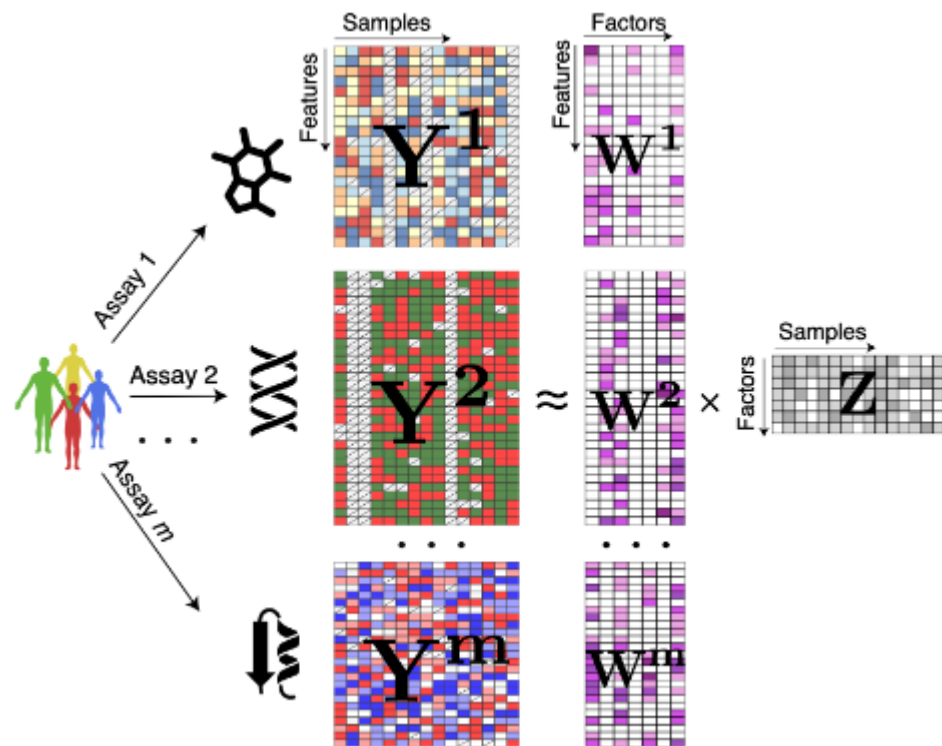




$$\mathbf{Y}^m = \mathbf{W}^m \mathbf{Z} + \epsilon^m$$

Main differences with PCA :

- m views for m omic sources with a share the \mathbf{Z} matrix between views
→ extract common effects across omics
- Bayesian Framework
- No orthogonality imposed for \mathbf{W}^m or \mathbf{Z}
- Sparsity on \mathbf{W}^m :
→ sparse biological phenomenon
- $\mathbf{Y}^m / \epsilon^m$ can follow different distributions :
 - Gaussian (continuous)



$$\mathbf{Y}^m = \mathbf{W}^m \mathbf{Z} + \epsilon^m$$

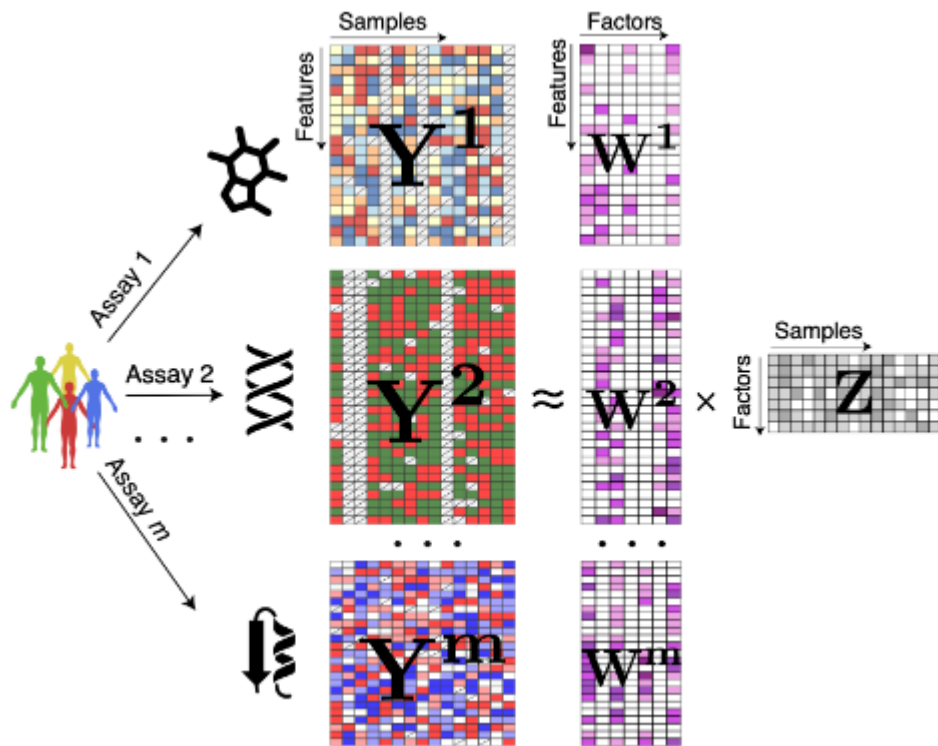
Main differences with PCA :

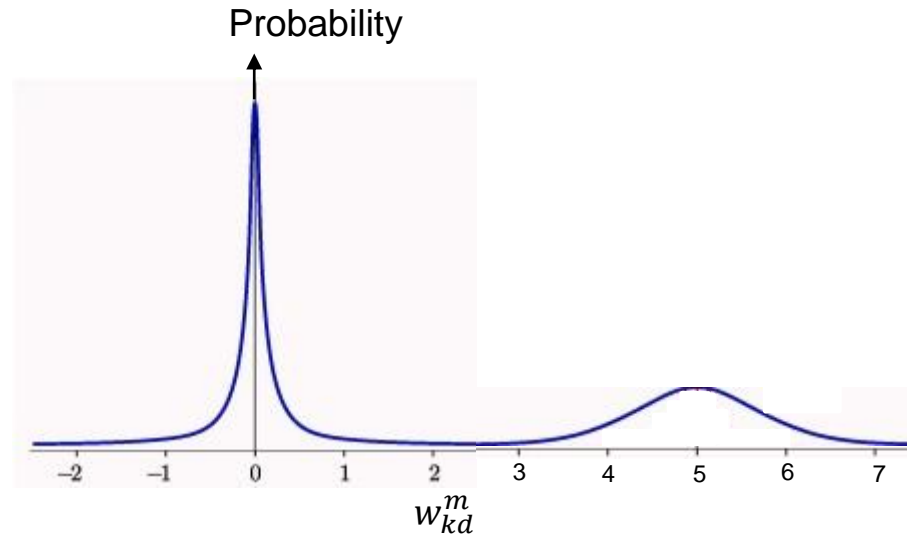
- m views for m omic sources with a share the \mathbf{Z} matrix between views
→ extract common effects across omics
- Bayesian Framework
- No orthogonality imposed for \mathbf{W}^m or \mathbf{Z}
- Sparsity on \mathbf{W}^m :
→ sparse biological phenomenon
- $\mathbf{Y}^m / \epsilon^m$ can follow different distributions :
 - Gaussian (continuous)
 - Poisson (natural)

$$\mathbf{Y}^m = \mathbf{W}^m \mathbf{Z} + \epsilon^m$$

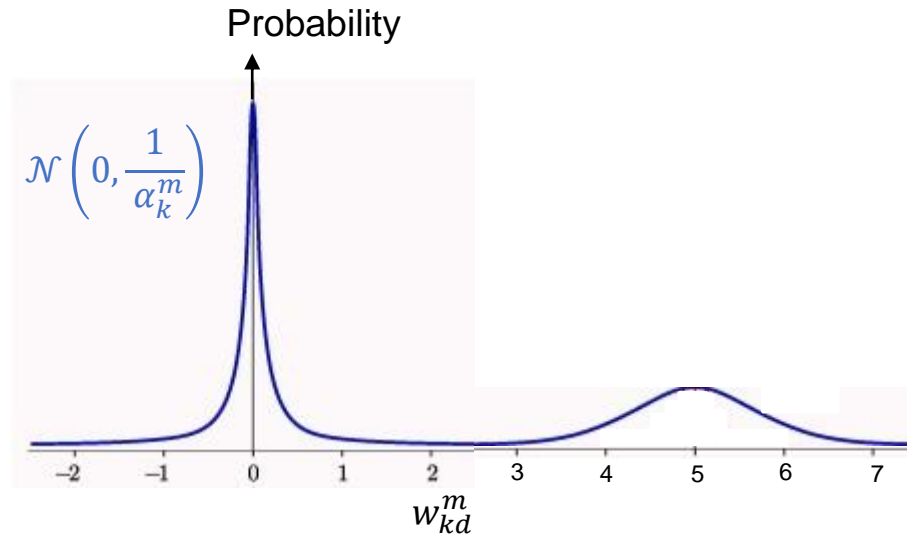
Main differences with PCA :

- m views for m omic sources with a share the \mathbf{Z} matrix between views
→ extract common effects across omics
- Bayesian Framework
- No orthogonality imposed for \mathbf{W}^m or \mathbf{Z}
- Sparsity on \mathbf{W}^m :
→ sparse biological phenomenon
- $\mathbf{Y}^m / \epsilon^m$ can follow different distributions :
 - Gaussian (continuous)
 - Poisson (natural)
 - Bernoulli (binary)

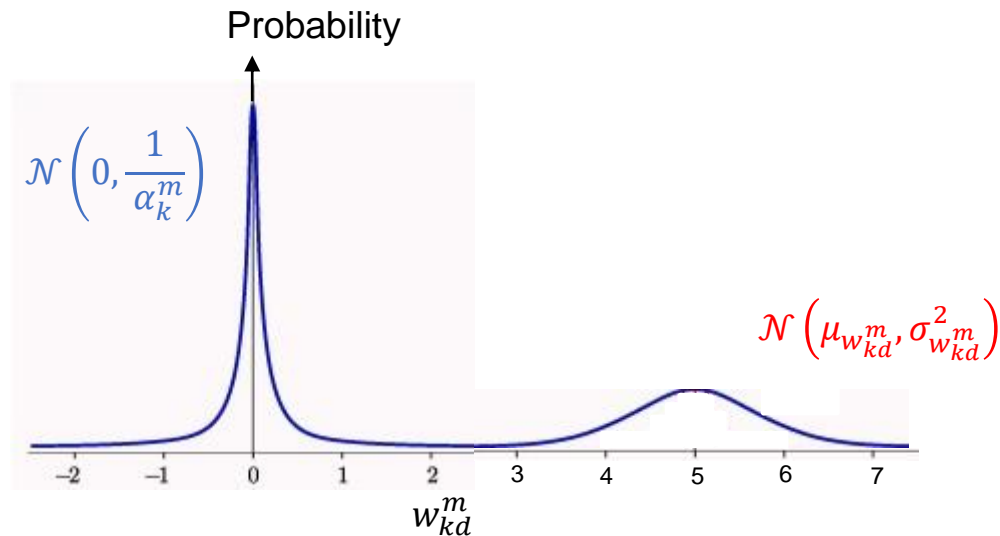




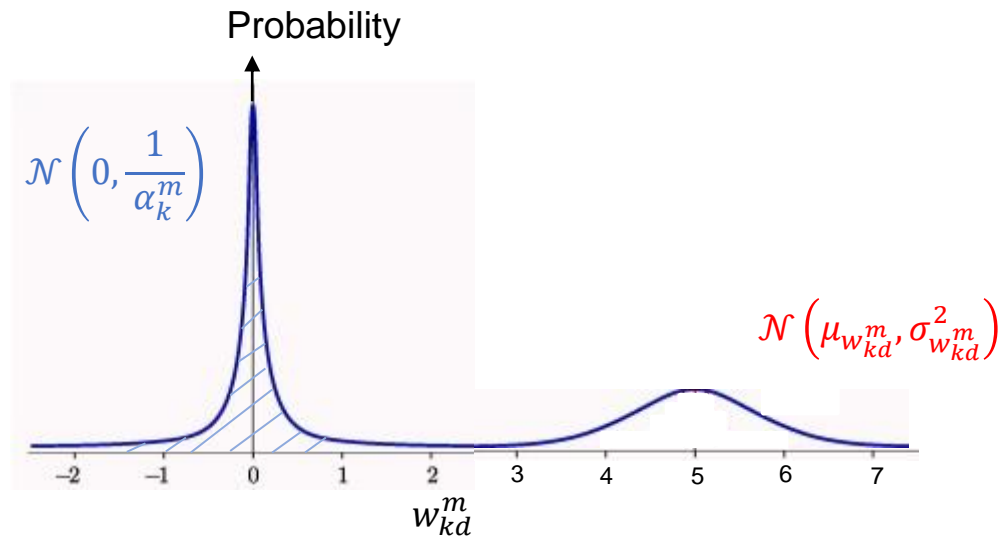
Posterior probability distribution



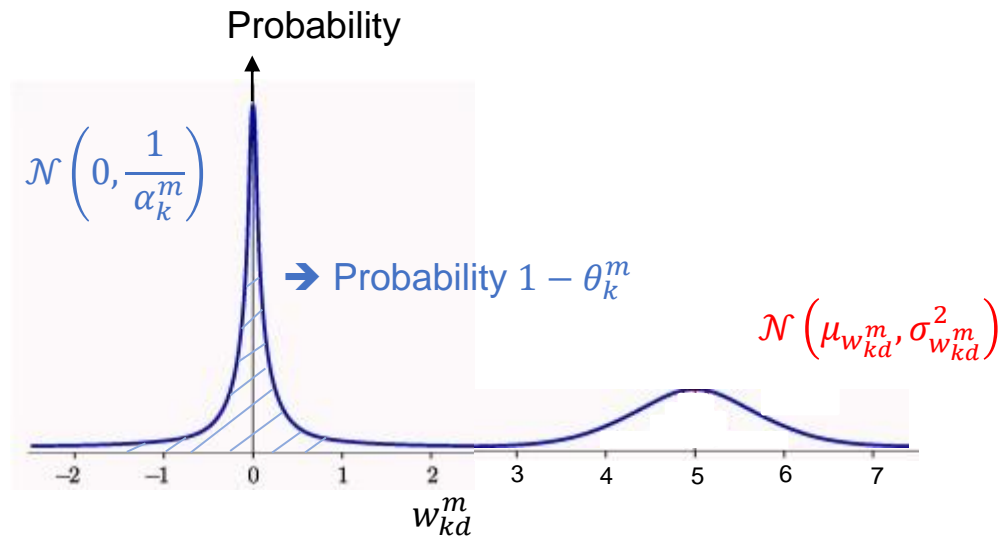
Posterior probability distribution



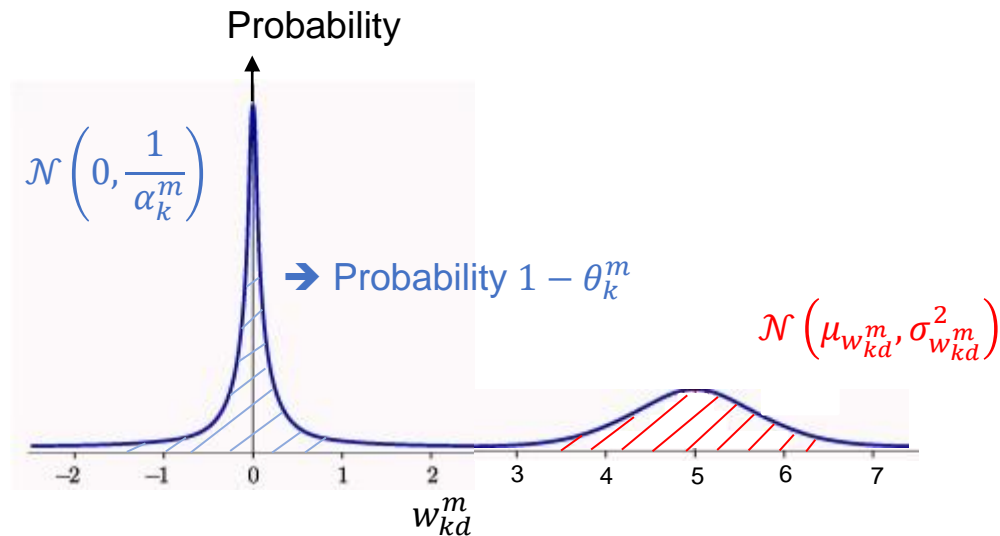
Posterior probability distribution



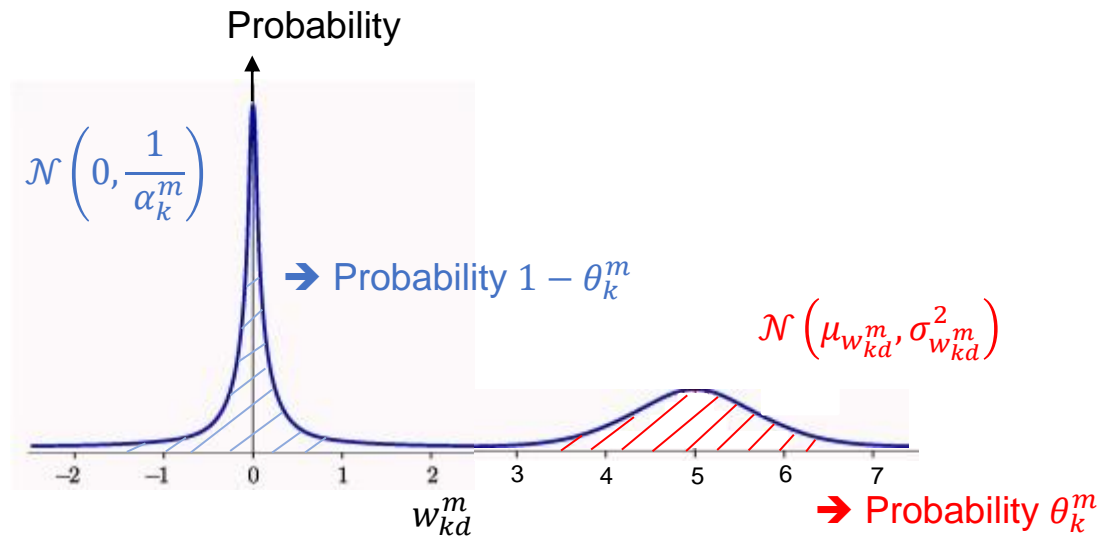
Posterior probability distribution



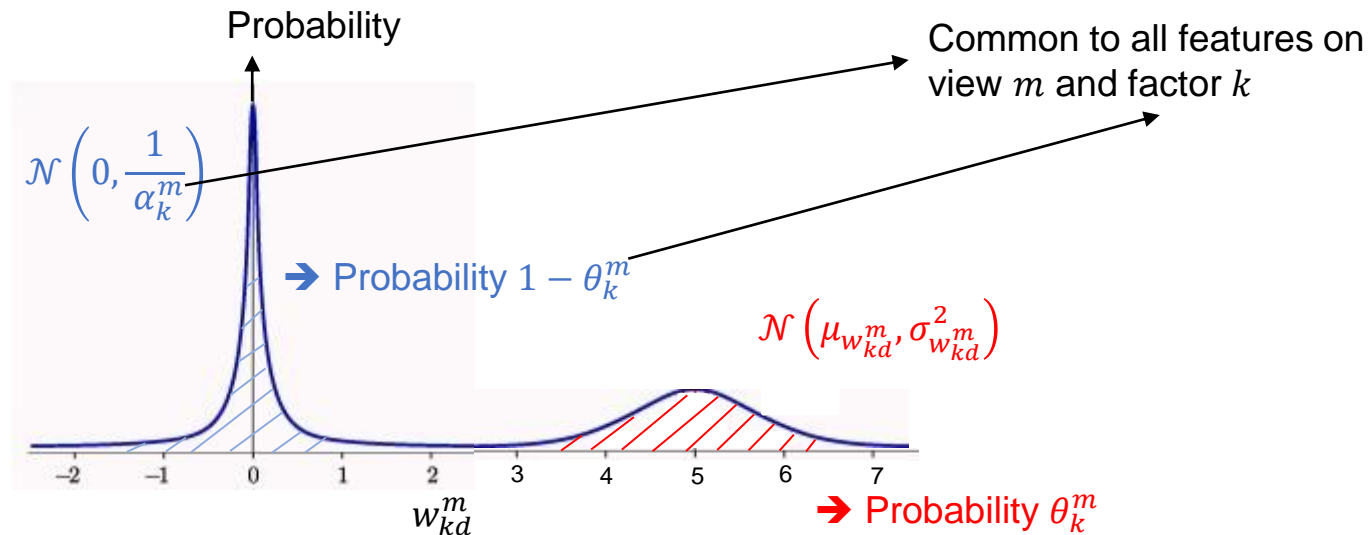
Posterior probability distribution



Posterior probability distribution



Posterior probability distribution



Posterior probability distribution



- Unsupervised, Mixed integration strategy, vertical integration



- Unsupervised, Mixed integration strategy, vertical integration
- Choice of **K** (number of factors)



- Unsupervised, Mixed integration strategy, vertical integration
- Choice of **K** (number of factors)
 - inactive factors can be removed through a user defined explained variance threshold during optimization



- Unsupervised, Mixed integration strategy, vertical integration
- Choice of **K** (number of factors)
→ inactive factors can be removed through a user defined explained variance threshold during optimization
- PCA initialization on concatenated views



- Unsupervised, Mixed integration strategy, vertical integration
- Choice of **K** (number of factors)
→ inactive factors can be removed through a user defined explained variance threshold during optimization
- PCA initialization on concatenated views
- Missing value



- Unsupervised, Mixed integration strategy, vertical integration
- Choice of **K** (number of factors)
 - inactive factors can be removed through a user defined explained variance threshold during optimization
- PCA initialization on concatenated views
- Missing value
 - no need for imputation, missing values are ignored in the model thanks to probabilistic modelling



- Unsupervised, Mixed integration strategy, vertical integration
- Choice of **K** (number of factors)
→ inactive factors can be removed through a user defined explained variance threshold during optimization
- PCA initialization on concatenated views
- Missing value
→ no need for imputation, missing values are ignored in the model thanks to probabilistic modelling
- Data pre-processing



- Unsupervised, Mixed integration strategy, vertical integration
- Choice of **K** (number of factors)
 - inactive factors can be removed through a user defined explained variance threshold during optimization
- PCA initialization on concatenated views
- Missing value
 - no need for imputation, missing values are ignored in the model thanks to probabilistic modelling
- Data pre-processing
 - use appropriate normalization to fit with available likelihoods



- Unsupervised, Mixed integration strategy, vertical integration
- Choice of **K** (number of factors)
 - inactive factors can be removed through a user defined explained variance threshold during optimization
- PCA initialization on concatenated views
- Missing value
 - no need for imputation, missing values are ignored in the model thanks to probabilistic modelling
- Data pre-processing
 - use appropriate normalization to fit with available likelihoods
 - e.g. for rna-seq: size factor normalization + VST*



- Unsupervised, Mixed integration strategy, vertical integration
- Choice of **K** (number of factors)
 - inactive factors can be removed through a user defined explained variance threshold during optimization
- PCA initialization on concatenated views
- Missing value
 - no need for imputation, missing values are ignored in the model thanks to probabilistic modelling
- Data pre-processing
 - use appropriate normalization to fit with available likelihoods
 - e.g. for rna-seq: size factor normalization + VST*
 - (use Gaussian likelihood when possible (instead of Poisson or Bernoulli))



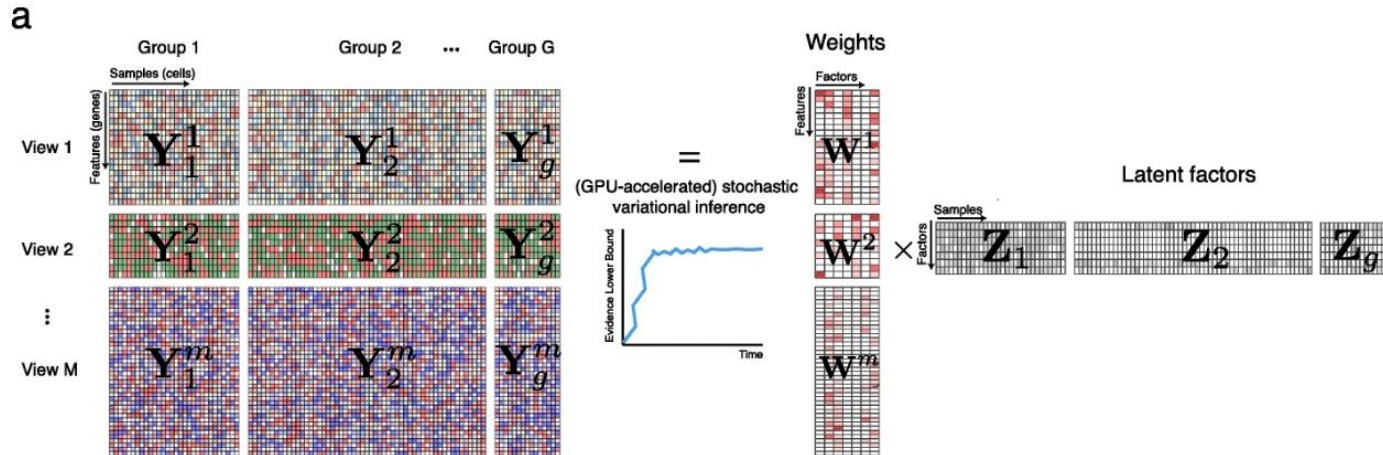
- Lack of function to predict the position of a new individual in the latent space
 - ➔ [Is Multi-Omics Transfer Learning \(MOTL\)](#) a solution ? ➔ JOBIM 2026 ?
 - ➔ Other strategy: [Using multiomic integration to improve blood biomarkers of major depressive disorder: a case-control study](#)
- Mainly linear relationships are captured
- Assumes independence between features in the model (and sample)
- Unbalanced modalities sensibility
 - ➔ consider feature selection procedure before MOFA
(e.g. *highly variable features*)
 - ❖ to limit view imbalance
 - ❖ to speedup the model



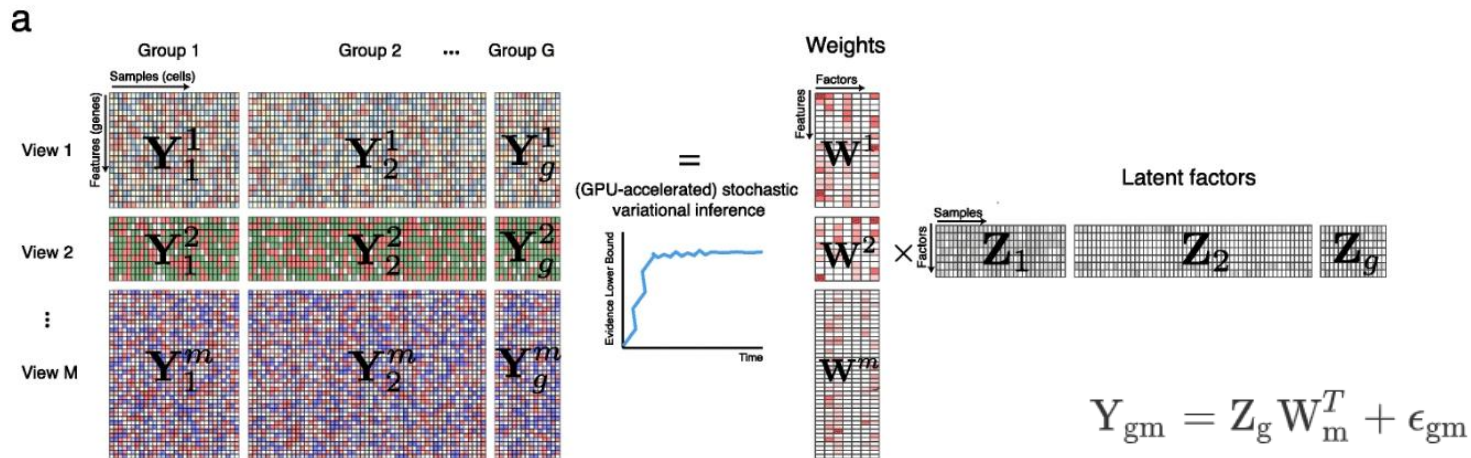


- Define sample groups (batch, experiments, conditions) → both Vertical/Horizontal integration

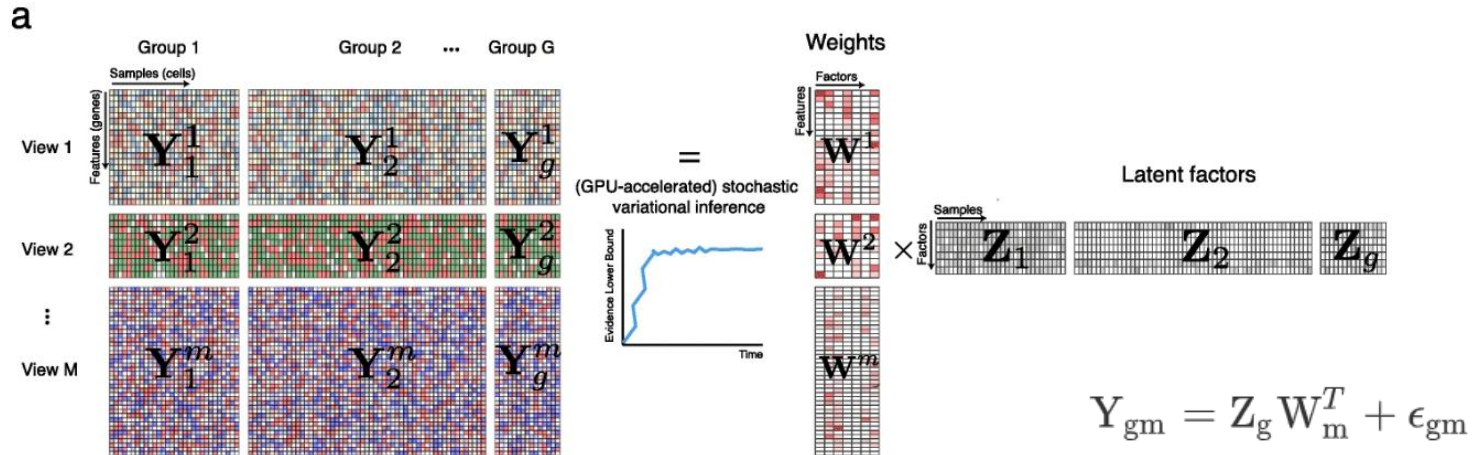
- Define sample groups (batch, experiments, conditions) → both Vertical/Horizontal integration



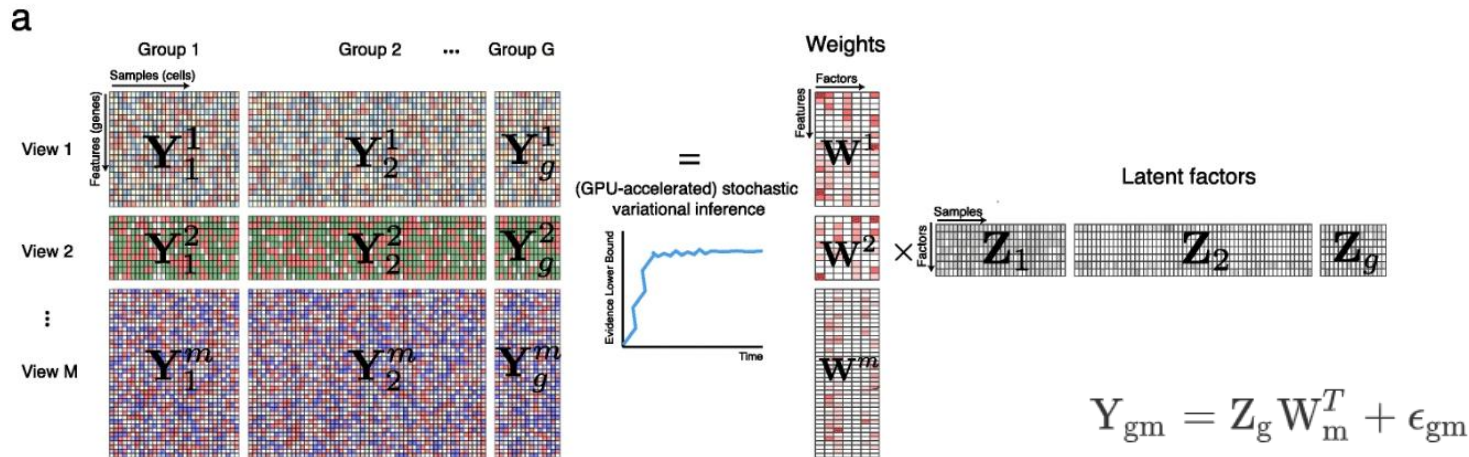
- Define sample groups (batch, experiments, conditions) → both Vertical/Horizontal integration



- Define sample groups (batch, experiments, conditions) → both Vertical/Horizontal integration
- Groups should not be defined as for a differential analysis (identify shared and exclusive source of variability in the data)



- Define sample groups (batch, experiments, conditions) → both Vertical/Horizontal integration
- Groups should not be defined as for a differential analysis (identify shared and exclusive source of variability in the data)
- Group-wise prior to the Z matrix



- Define sample groups (batch, experiments, conditions) → both Vertical/Horizontal integration
- Groups should not be defined as for a differential analysis (identify shared and exclusive source of variability in the data)
- Group-wise prior to the Z matrix
- Stochastic variational inference framework (GPU accelerated computations)

