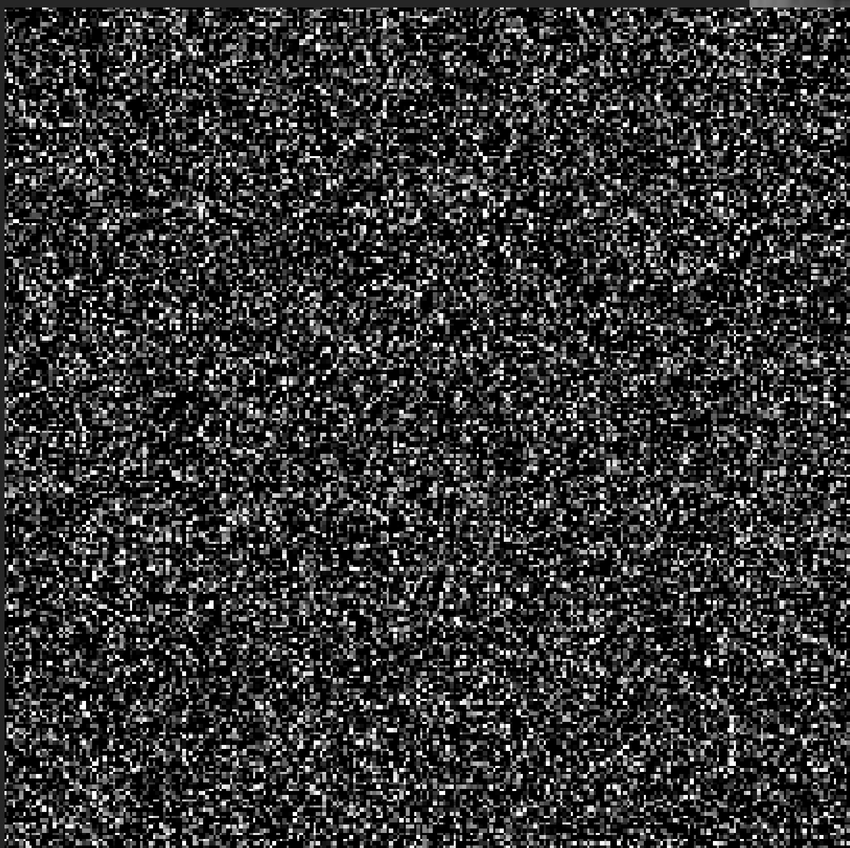
A Bayesian Approach to Tensor Decompositions

Clara Menzen, Manon Kok, Kim Batselier

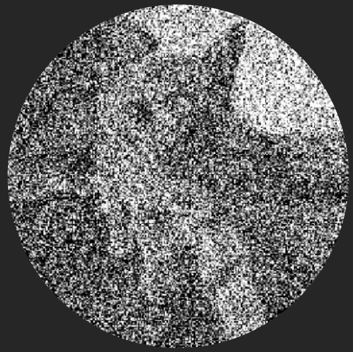
May, 21st, 2021

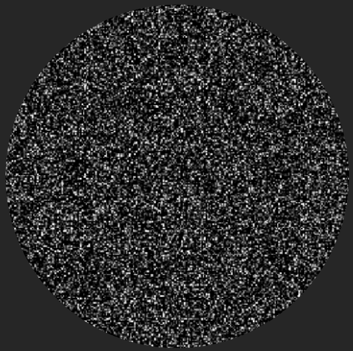
TU Delft

Noise



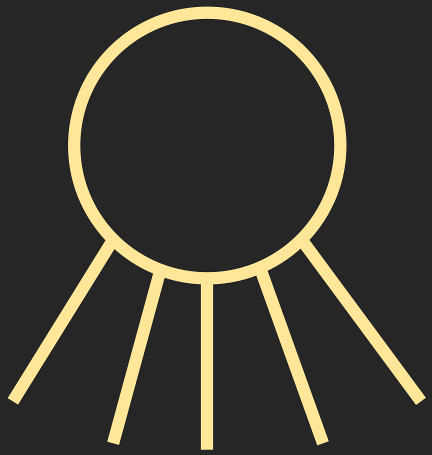
Low-rank



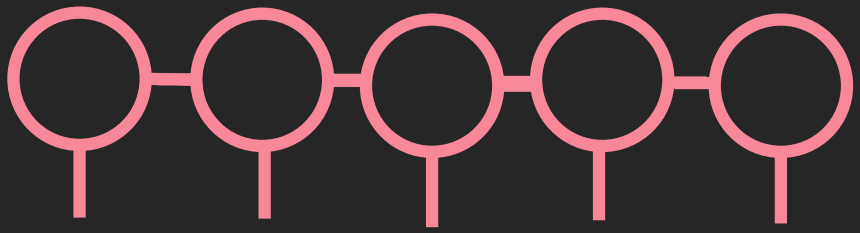


$$\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{Y}}_{\text{lr}} + \boldsymbol{\mathcal{E}}$$

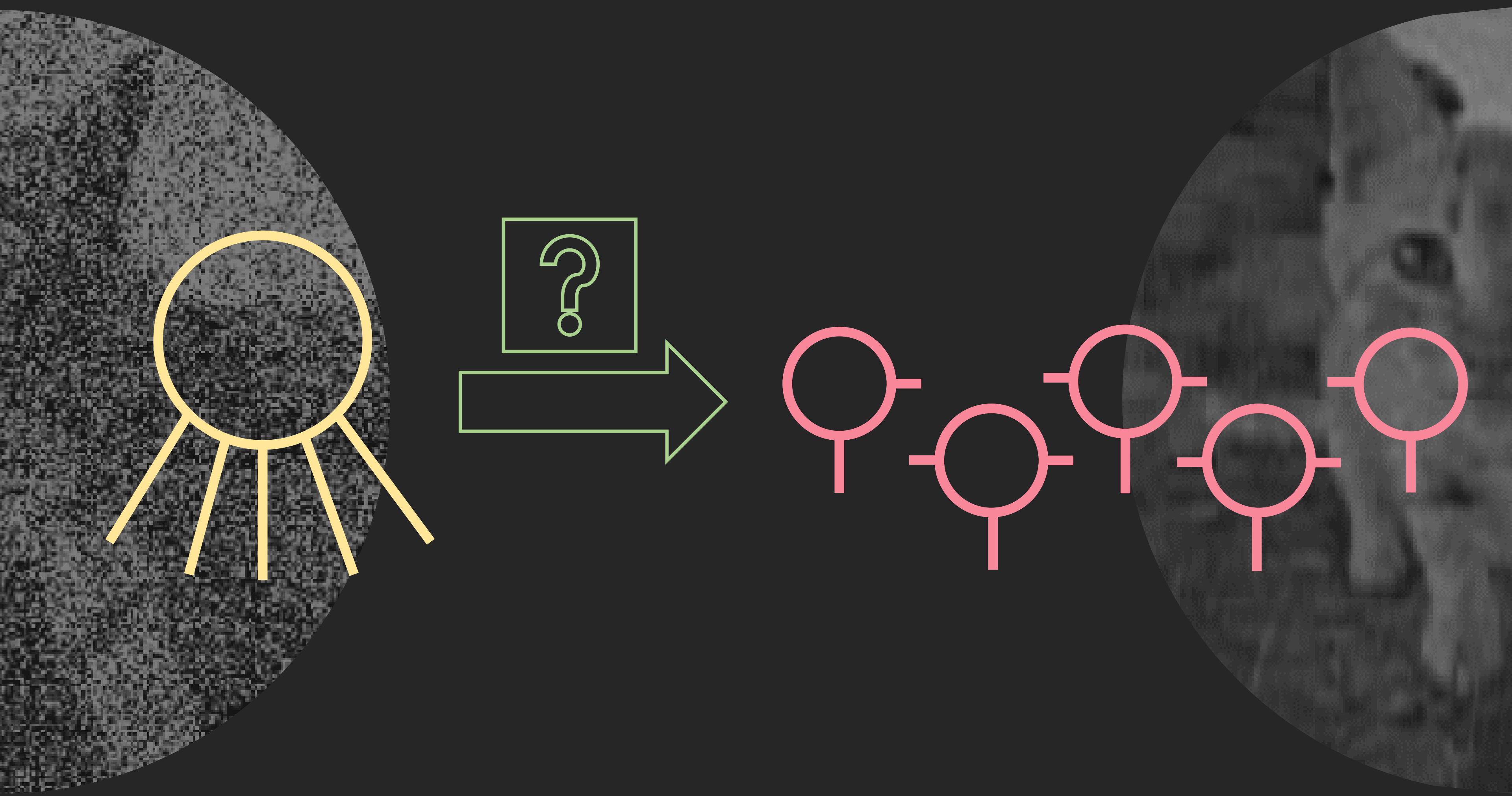
$$\text{vec}(\boldsymbol{\mathcal{E}}) \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$$

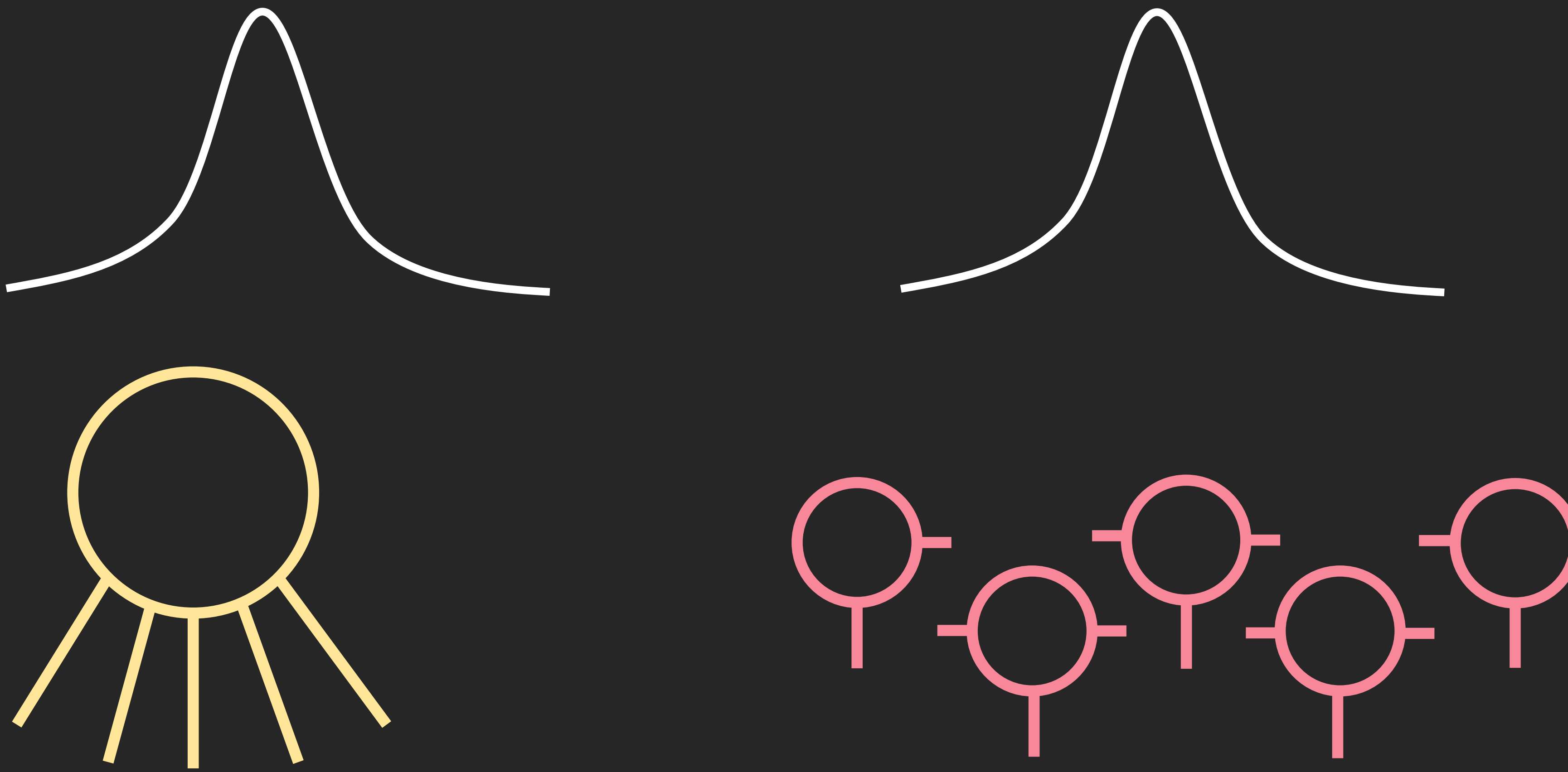


5-way tensor

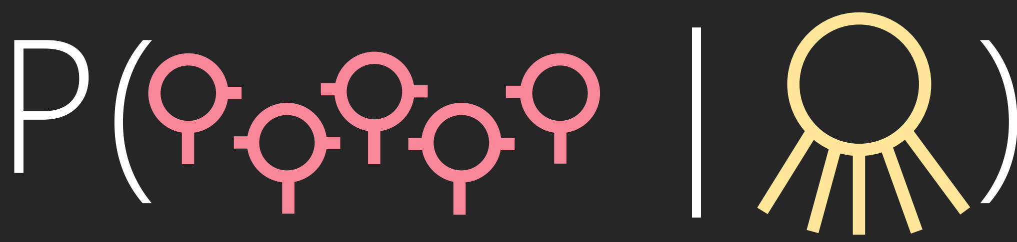


Tensor Decomposition



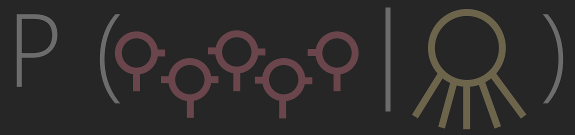


$$\mathrm{P}(\bullet) \sim \mathcal{N}(\mu, \Sigma)$$

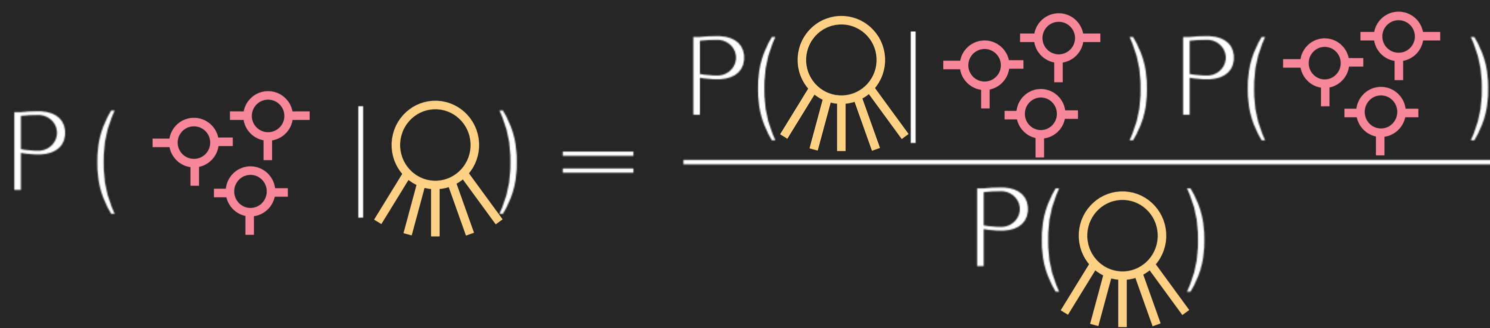

P(
|
)

Bayes rule

$$P\left(\{\mathbf{g}_i\} \mid \mathbf{y}\right)$$



Bayes rule



Model choices

Gaussian random variables

$$P\left(\mathbf{y} \mid \{\mathbf{g}_i\}\right) \sim \mathcal{N}\left(\mu_{\mathbf{y}}, \sigma^2 \mathbf{I}\right)$$

Prior for tensor decomposition components

$$P(\{\mathbf{g}_i\}) \sim \mathcal{N}\left(\begin{bmatrix} \mu_1^0 \\ \mu_2^0 \\ \vdots \\ \mu_N^0 \end{bmatrix}, \begin{bmatrix} \Sigma_1^0 & 0 & \dots & 0 \\ 0 & \Sigma_2^0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \Sigma_N^0 \end{bmatrix} \right)$$

Assuming statistical independence

$$P(\{\mathbf{g}_i\}) = P(\mathbf{g}_1)P(\mathbf{g}_2)\dots P(\mathbf{g}_N)$$

Model choices

Gaussian random variables

Model choices

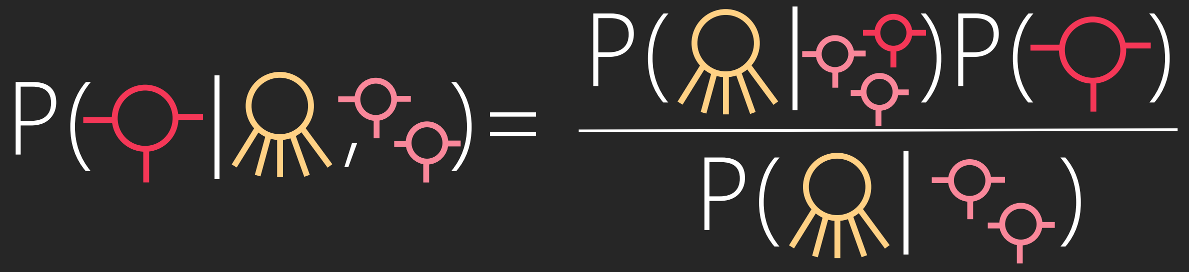
Gaussian random variables
Statistical independence

Model choices

Gaussian random variables
Statistical independence

Model choices

Gaussian random variables
Statistical independence
Block-coordinate descent

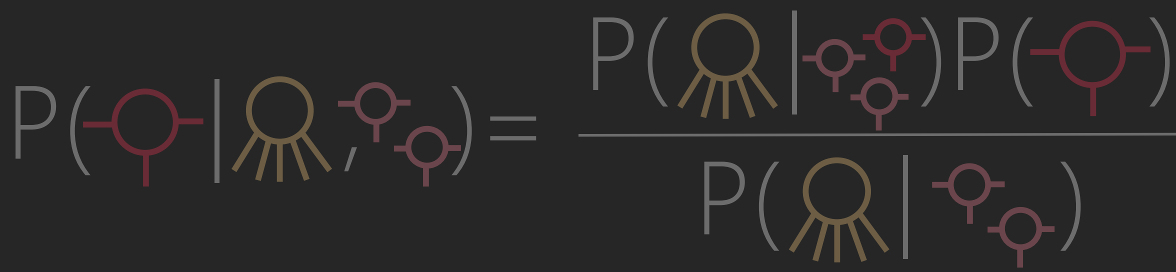


Model choices

Gaussian random variables
Statistical independence
Block-coordinate descent

$$P\left(\mathbf{y} \mid \{\mathbf{g}_i\}\right) \sim \mathcal{N}\left(\mu_{\mathbf{y}}, \sigma^2 \mathbf{I}\right)$$

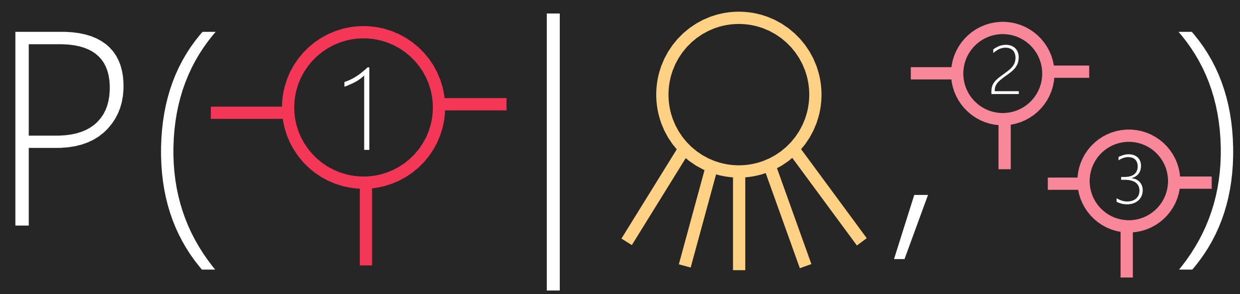
$$P\left(\mathbf{y} \mid \{\mathbf{g}_i\}\right) \sim \mathcal{N}\left(\mathbf{U}_{\setminus n} \mathbf{g}_n, \sigma^2 \mathbf{I}\right)$$



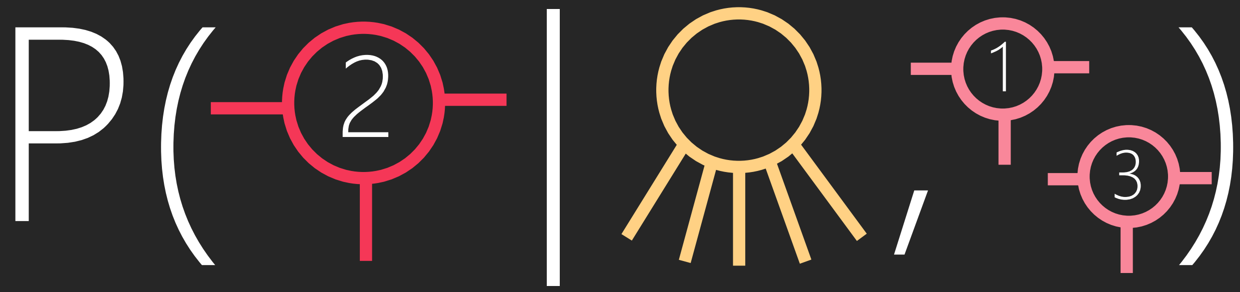
$$\mu_n^+ = \underbrace{\left[(\Sigma_n^0)^{-1} + \frac{\mathbf{U}_{\setminus n}^\top \mathbf{U}_{\setminus n}}{\sigma^2} \right]^{-1}}_{\Sigma_n^+} \left[\frac{\mathbf{U}_{\setminus n}^\top \mathbf{y}}{\sigma^2} + (\Sigma_n^0)^{-1} \mu_n^0 \right]$$

$$\Sigma_n^0 \to \infty: \quad \mu_n^+ = \left(\mathbf{U}_{\setminus n}^\top \mathbf{U}_{\setminus n} \right)^{-1} \mathbf{U}_{\setminus n}^\top \mathbf{y}$$

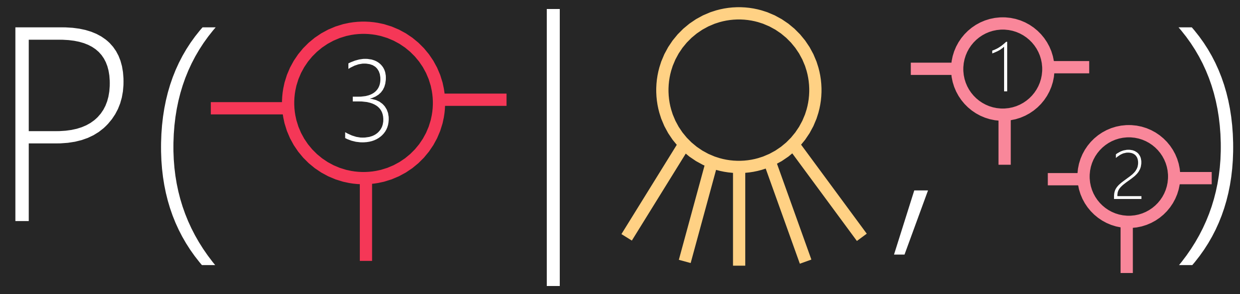
Alternating update

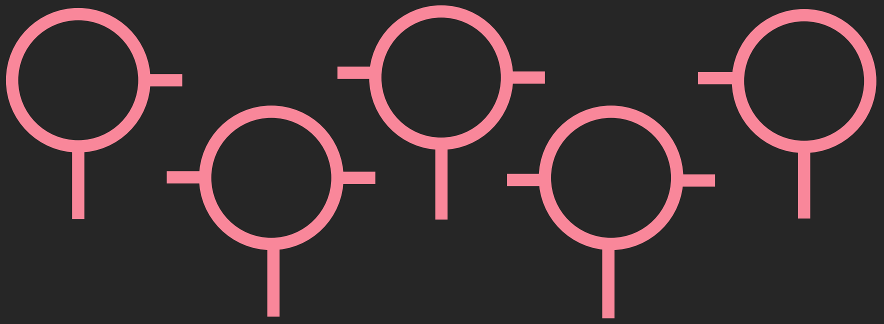


Alternating update

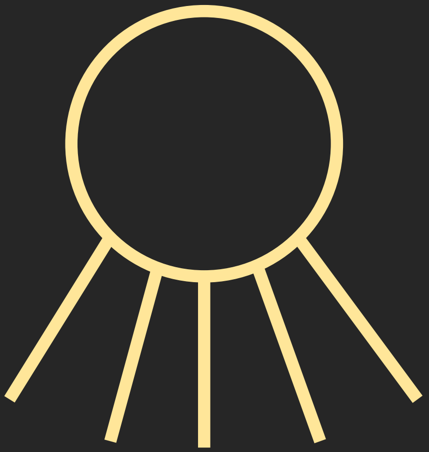


Alternating update





$$P\left(\mathbf{g}_i\right) \sim \mathcal{N}\left(\mu_i^+, \Sigma_i^+\right)$$



$$P(\ \hat{\mathbf{y}}\ ) \sim \mathcal{N}\left(\mu_{\hat{\mathbf{y}}}, \Sigma_{\hat{\mathbf{y}}}\right)$$

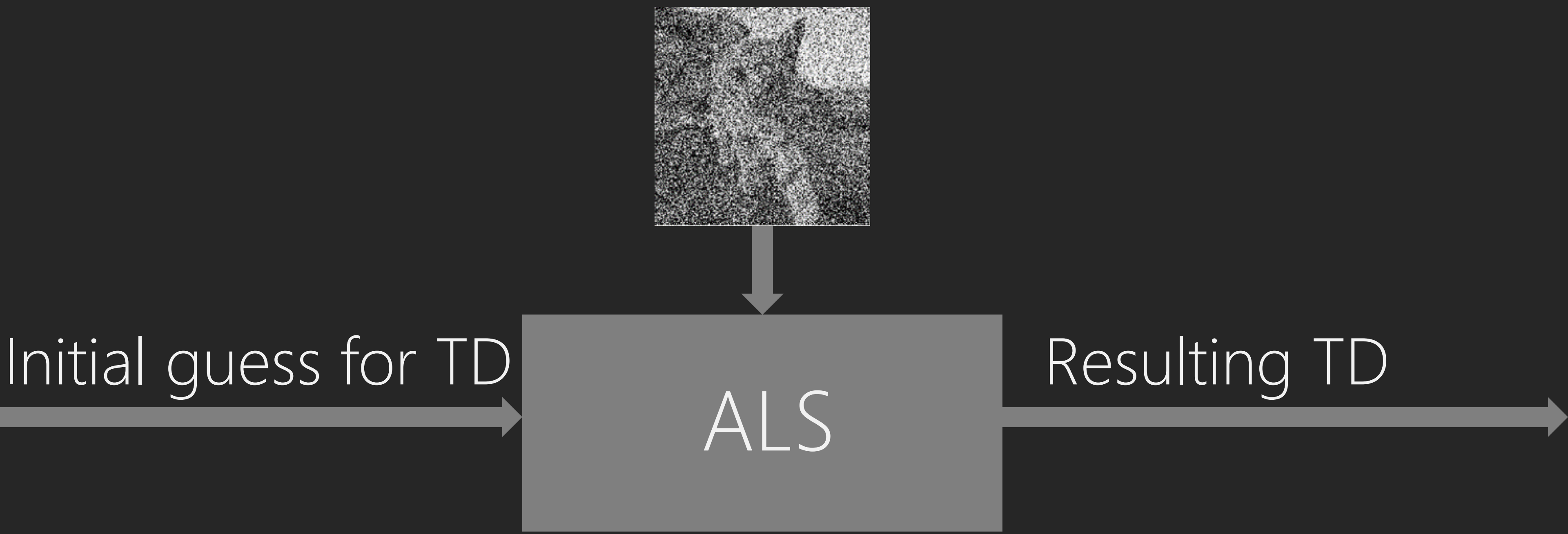
Recursive computation

$$\mathrm{P}(\mathbf{g}_i) \sim \mathcal{N}\left(\mu_i^0, \Sigma_i^0\right)$$

Bayesian ALS

$$\mathrm{P}(\mathbf{g}_i) \sim \mathcal{N}\left(\mu_i^+, \Sigma_i^+\right)$$

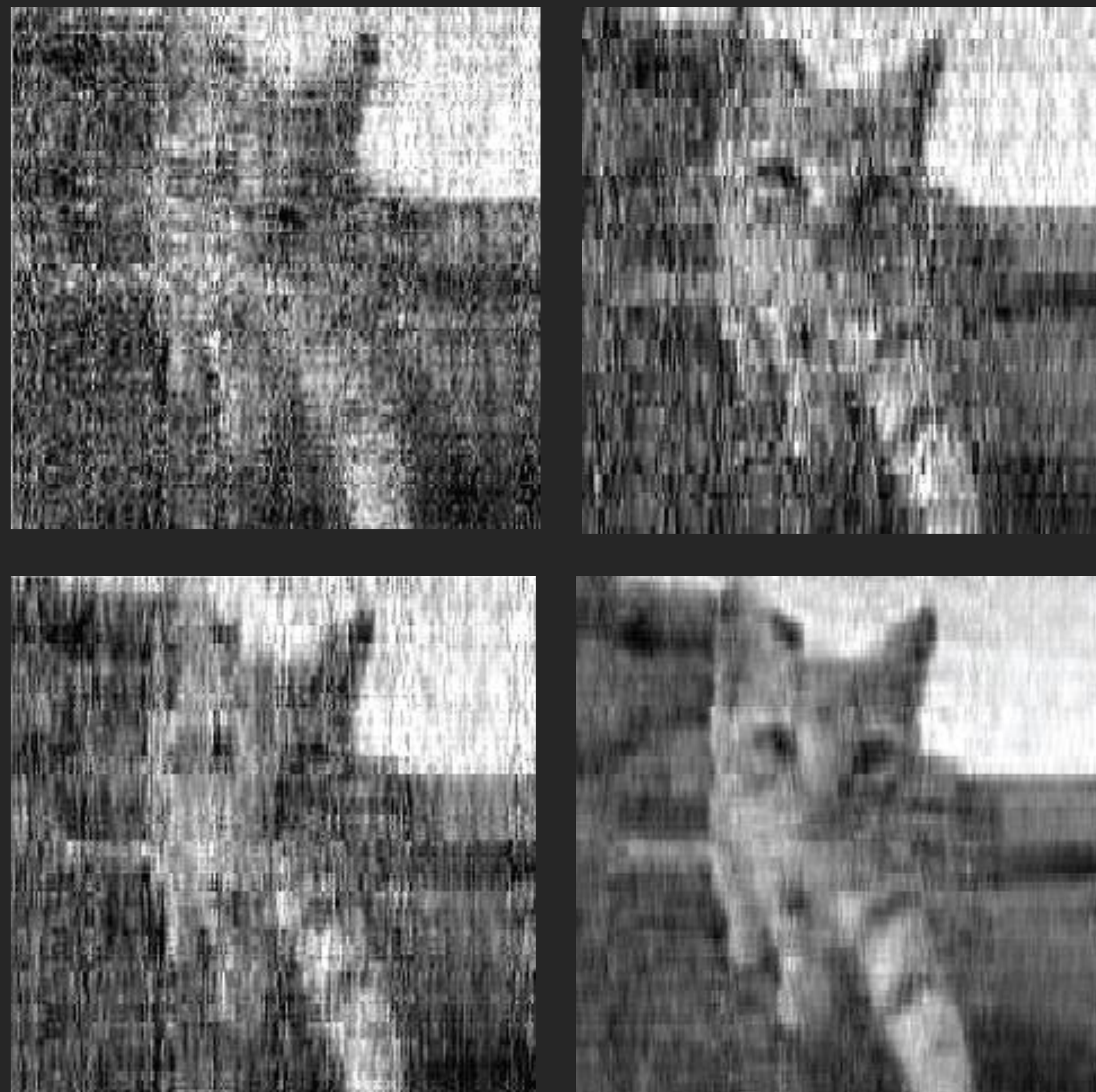
Recursive computation



1 noisy image | 10 noisy images

ALS

Bayesian ALS



Alternating linear scheme in a Bayesian framework for low-rank tensor approximation

Clara Menzen*, Manon Kok*, and Kim Batselier*

Abstract. Multiway data often naturally occurs in a tensorial format which can be approximately represented by a low-rank tensor decomposition. This is useful because complexity can be significantly reduced and the treatment of large-scale data sets can be facilitated. In this paper, we find a low-rank representation for a given tensor by solving a Bayesian inference problem. This is achieved by dividing the overall inference problem into sub-problems where we sequentially infer the posterior distribution of one tensor decomposition component at a time. This leads to a probabilistic interpretation of the well-known iterative algorithm alternating linear scheme (ALS). In this way, the consideration of measurement noise is enabled, as well as the incorporation of application-specific prior knowledge and the uncertainty quantification of the low-rank tensor estimate. To compute the low-rank tensor estimate from the posterior distributions of the tensor decomposition components, we present an algorithm that performs the unscented transform in tensor train format.

Key words. Low-rank approximation, alternating linear scheme, Bayesian inference, tensor decomposition, tensor train.

ArXiv:2012.11228