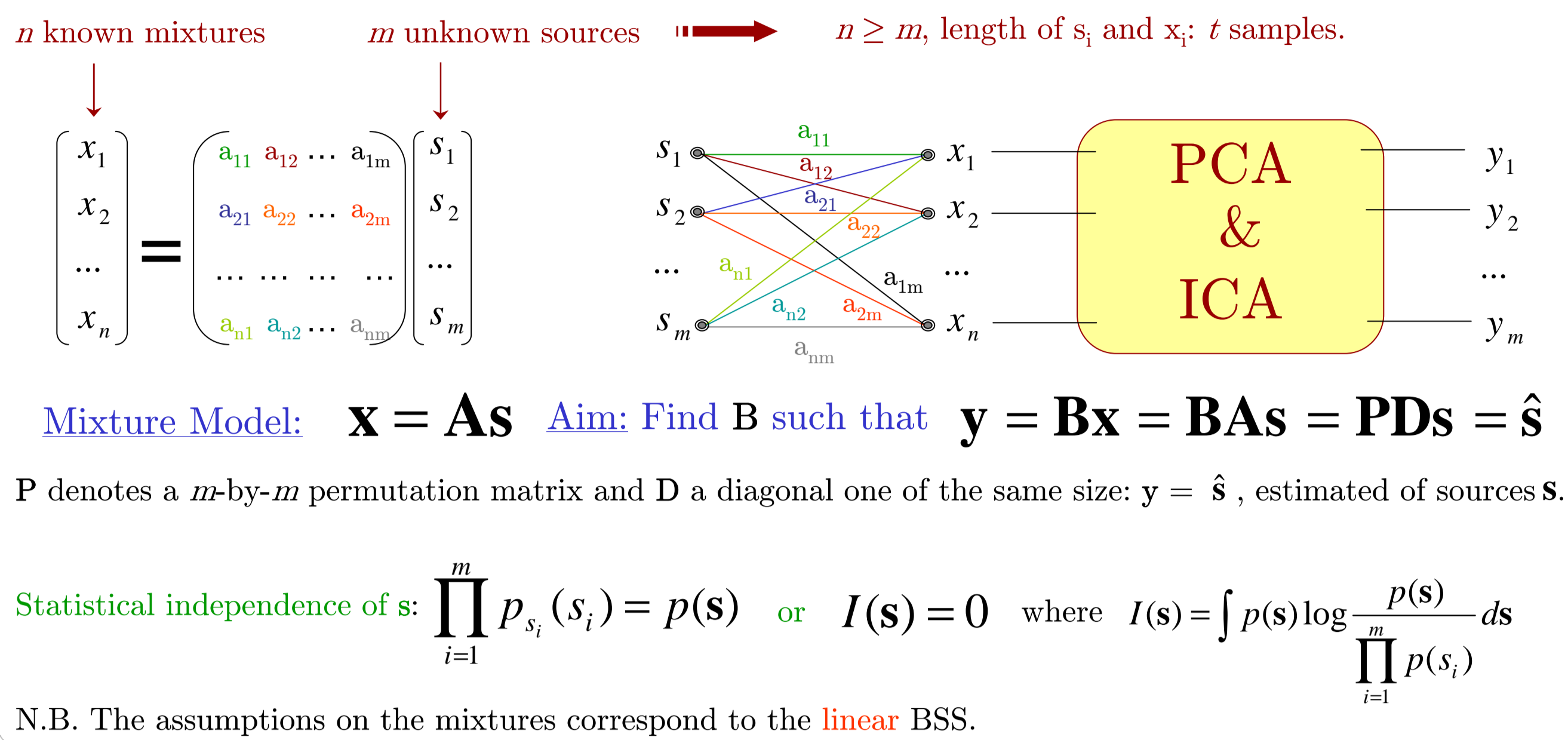


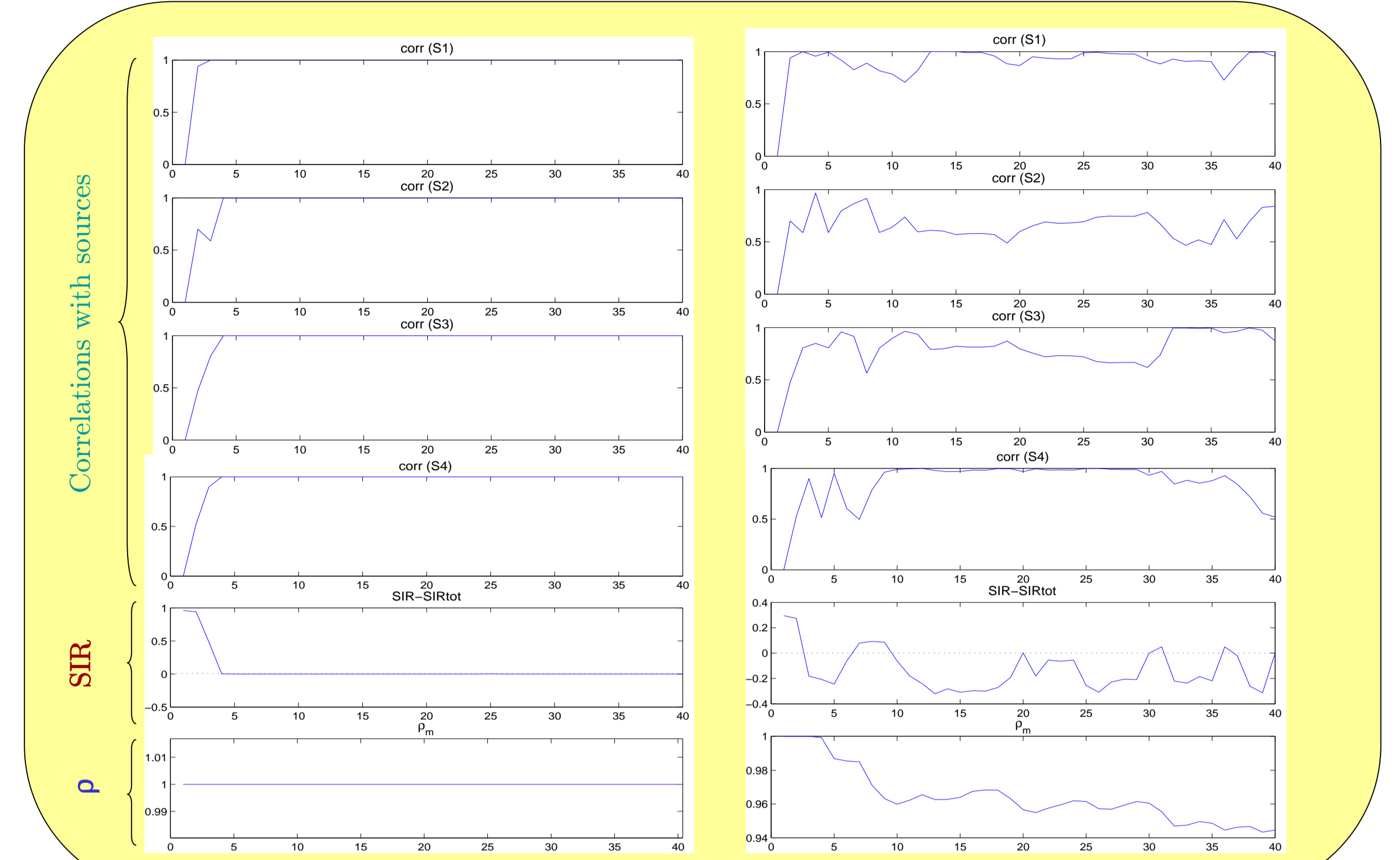
Improving Independent Component Analysis Performances by Variable Selection

F. Vrins, J. A. Lee, M. Verleysen (UCL – Machine Learning Group)
 V. Vigneron, C. Jutten (INPG – Images and Signals laboratory)
 {vrins,lee,verleysen}@dice.ucl.ac.be, {vvigne,jutten}@lis.inpg.fr

Independent Component Analysis



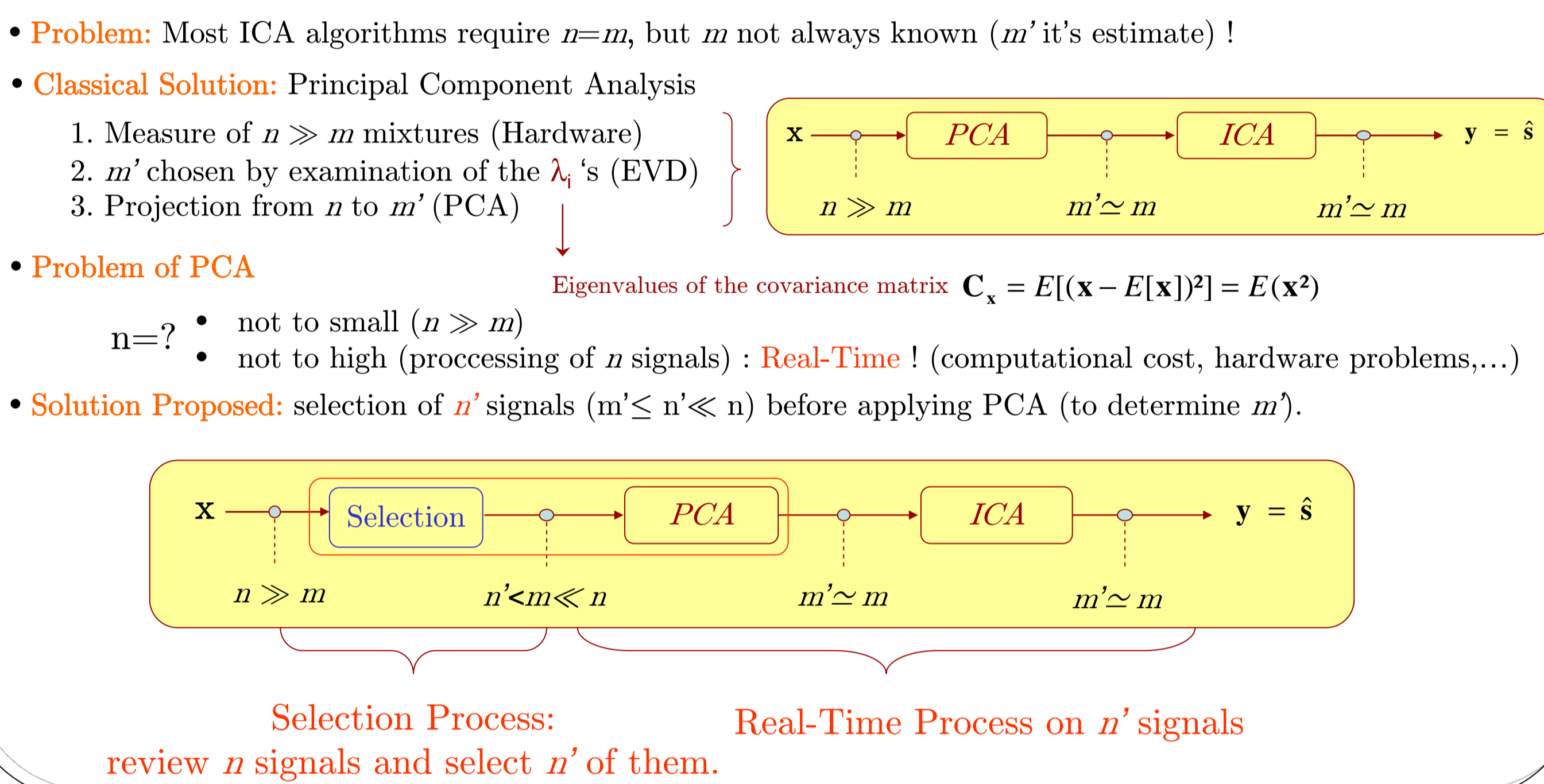
Impact of the number of mixtures if $m' \neq m$



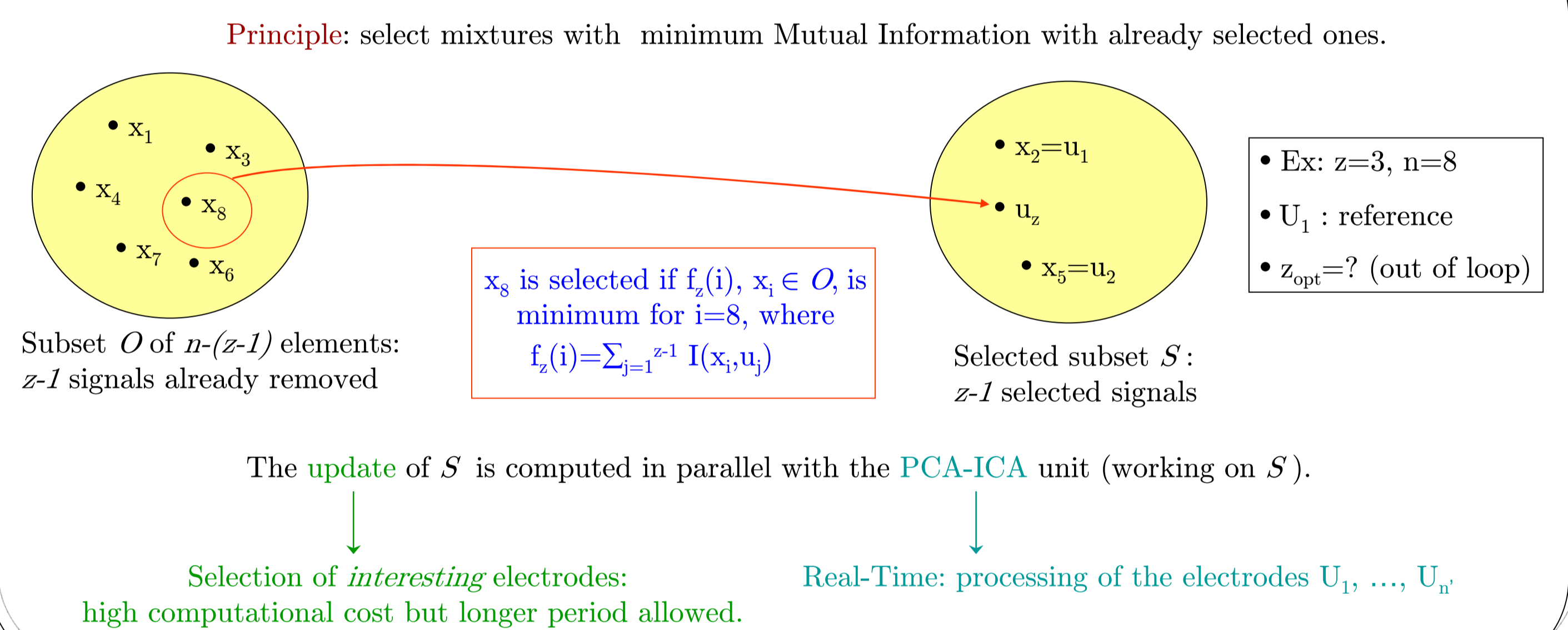
More than 4 mixtures are projected onto a 4th-dimensional subspace ($m'=m=4$).

More than 3 mixtures are projected onto a 3th-dimensional subspace ($m'=3, m=4$).

Problem: The Number of Observations n



Variable Selection algorithm using Mutual Information



Preprocessing to ICA

- Whitening**
 - Centering data: $\mathbf{x} = \mathbf{x} - E[\mathbf{x}]$ ($\rightarrow E[\mathbf{x}] = 0$), size(\mathbf{x}) = $n \times s$
 - Diagonalize $\mathbf{C}_x = E[(\mathbf{x} - E[\mathbf{x}])^2] = E(\mathbf{x}^2)$ by an eigenvalues decomposition (EVD): $\mathbf{C}_x = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$
 - $\mathbf{y} = \sqrt{\mathbf{\Lambda}^{-1}}\mathbf{V}^T\mathbf{x}$ $\mathbf{C}_y = E[\mathbf{y}^2] = \sqrt{\mathbf{\Lambda}^{-1}}\mathbf{V}^T\mathbf{C}_x\mathbf{V}\sqrt{\mathbf{\Lambda}^{-1}} = \mathbf{I}$ **\mathbf{y} are decorrelated!**
- Principal Component Analysis**

Dimension reduction from n to m' : $\mathbf{y} = \sqrt{\mathbf{\Lambda}_{(1:m',1:m')}^{-1}}\mathbf{V}_{1:m'}^T\mathbf{x}$

Projection and Separation Quality Criteria

- Projection (Principal Component Analysis)**

EVD(\mathbf{C}_x) $\rightarrow \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$, $\lambda_i (\geq 0) \propto$ variance along the axe \mathbf{V}_i

Projection from n to m' -subspace: keeping the m' (here 4) first $\lambda_i =$ keeping ρ % of the total variance :

$$\rho = \frac{\sum_{i=1}^{m'} \lambda_i}{\sum_{i=1}^n \lambda_i}$$
- Separation (Independent Component Analysis)**

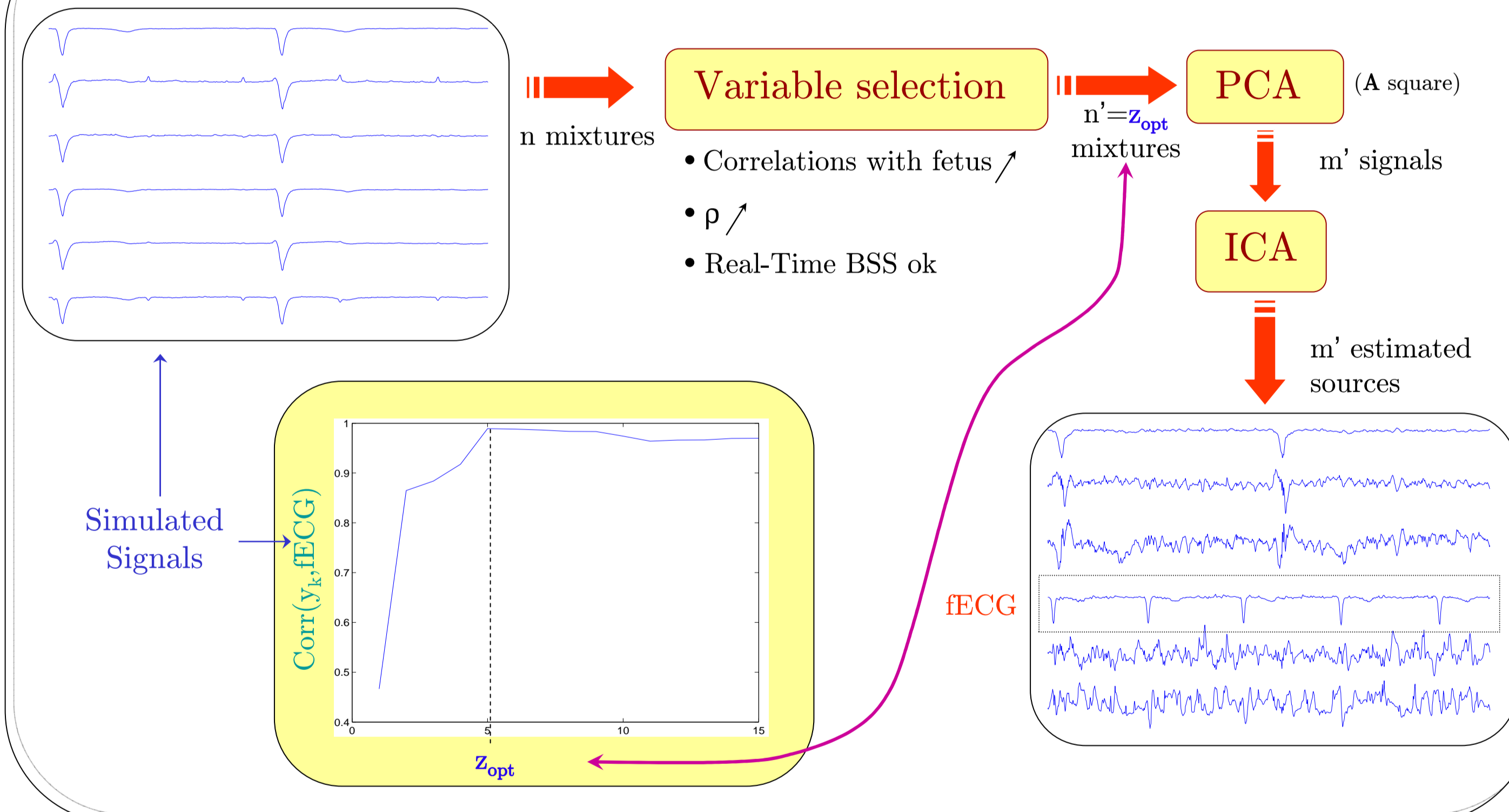
Correlation with sources (not 'blind')

$$\text{Corr}(s_j, x_i) = \frac{E[(s_j - E[s_j])(x_i - E[x_i])]}{\sqrt{E[(s_j - E[s_j])^2]E[(x_i - E[x_i])^2]}}$$

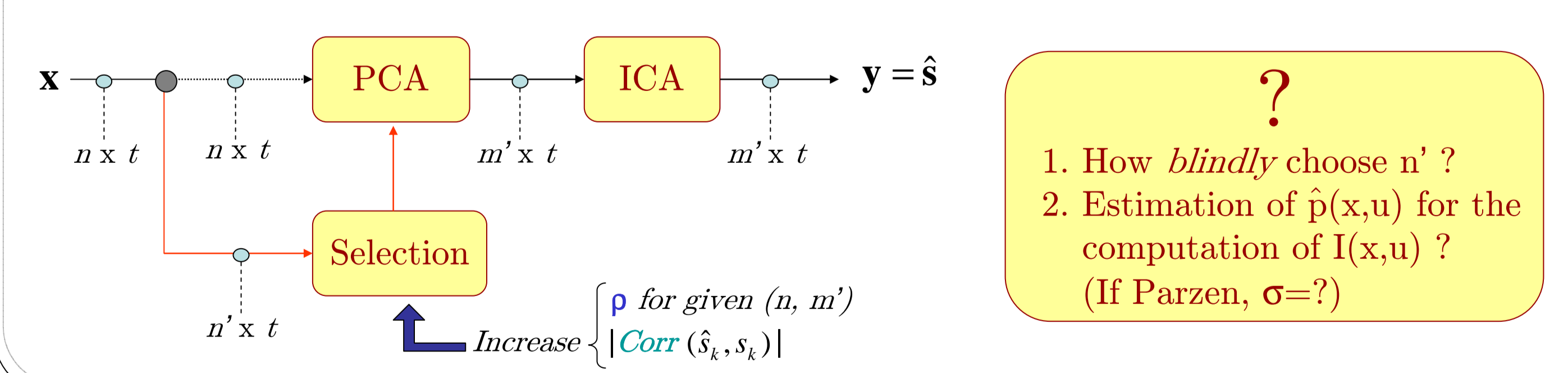
Signal Interference Ratio

$$\text{SIR} = \frac{1}{s} \sum_{i,j=1}^s \frac{|BA_{ij}|}{\max_{1 \leq k \leq s} |BA_{ik}|} - 1$$

Application to the fECG extraction



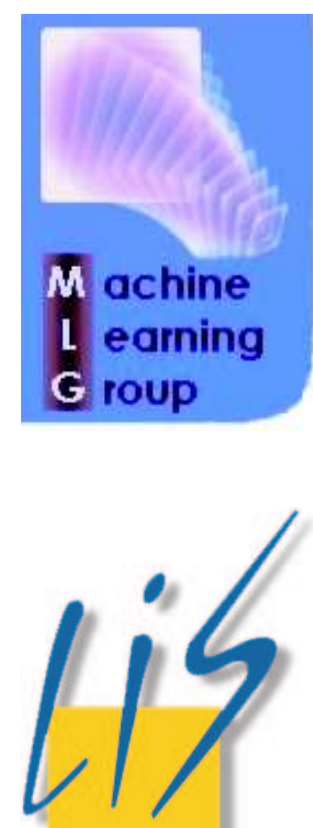
Conclusion



Institutions



- Université Catholique de Louvain (UCL)
 Machine Learning Group (DICE)
 Bât. Maxwell, 3 Place du Levant, 1348 Louvain-la-Neuve, Belgium
- Institut National Polytechnique de Grenoble (INPG)
 Laboratoire des Signaux et des Images (LIS)
 46 Avenue Félix Viallet, 38031 Grenoble Cedex, France



Main References

- L. De Lathauwer, B. De Moor and J. Vandewalle, Fetal Electrocardiogram by Blind Subspace Separation, IEEE Trans. On Biomedical Engineering, vol. 47, num. 5, pp. 567-572, 2000.
- J. V. Stone and J. Porril, Undercomplete Independent Component Analysis for Signal Separation and Dimension Reduction, Techn. Rep., Psychology Department, Sheffield University, Sheffield, S10 2UR, England, oct. 1997.
- I. T. Jolliffe, Principal Component Analysis, New York, Springer-Verlag, 1986.
- A. Hyvarinen, J. Karhunen and E. Oja, Independent Component Analysis, New York, Wiley-Interscience, 2000.
- P. Comon, Independent Component Analysis, A New Concept?, Signal Processing, v. 36, n. 3, pp. 287-314, 1994.
- T. M. Cover and J. A. Thomas, Elements of Information Theory, John Wiley and sons, 1991.
- J. Cardoso and A. Sououmiac, Blind Beamforming for Non-Gaussian Signals, in IEEE proceedings F, vol. 140, num. 6, pp. 362-370, 1993.