

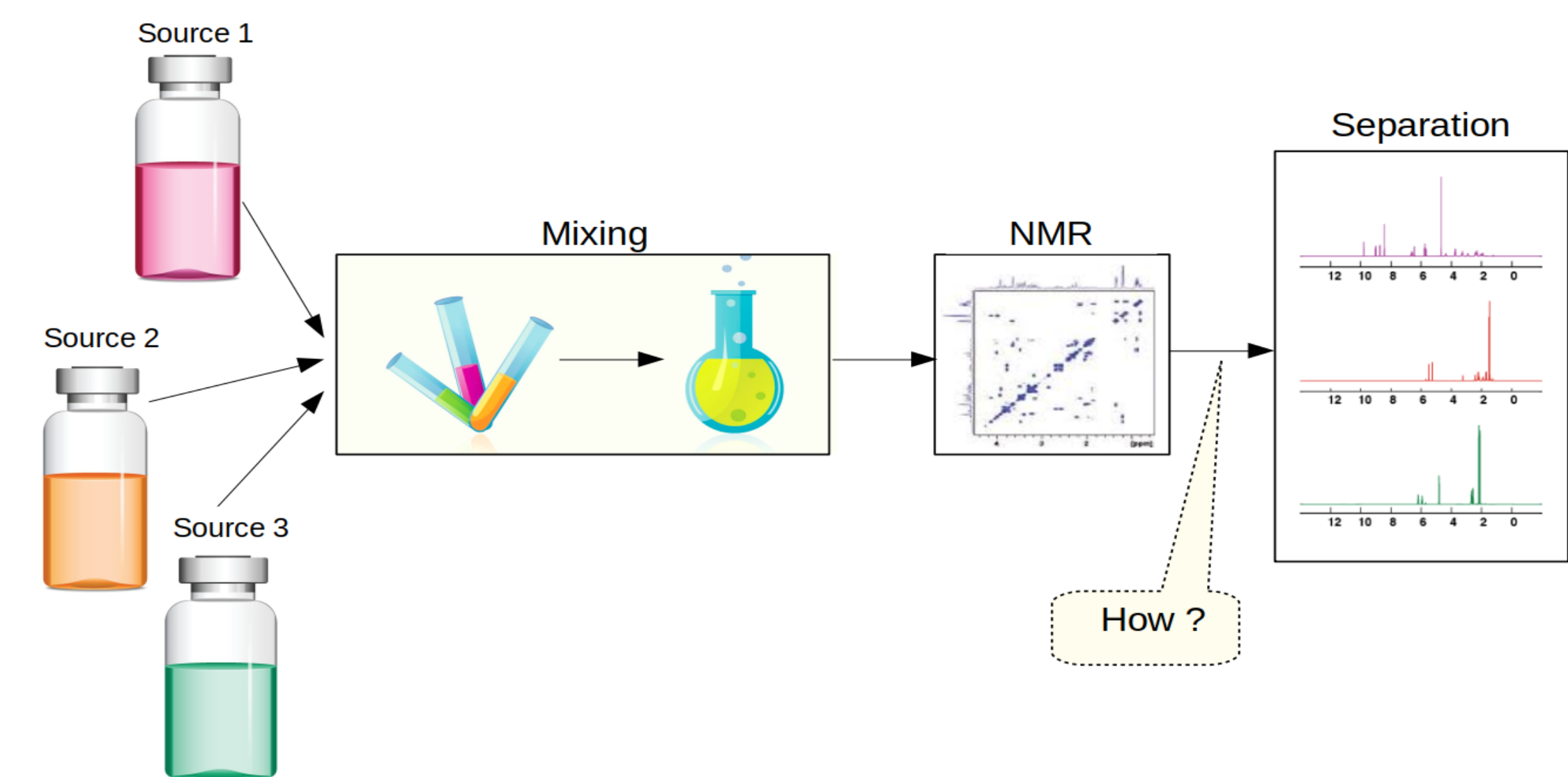
NMF-BASED SPARSE UNMIXING OF COMPLEX MIXTURES

Afef CHERNI*, Elena PIERSANTI*, Caroline CHAUX*

* Aix-Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France

* Aix-Marseille Univ, CNRS, Centrale Marseille, iSm2, Marseille, France.

Overview



• **Aim:** Given the measurements $\mathbf{X} \in \mathbb{R}^{M \times L}$, find the unknown sources $\mathbf{S} \in \mathbb{R}^{N \times L}$ and the unknown mixing matrix $\mathbf{A} \in \mathbb{R}^{M \times N}$ related to \mathbf{X} through

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{N} \approx \mathbf{A}\mathbf{S} \quad (1)$$

• **Application:** Nuclear Magnetic Resonance (NMR) spectroscopy \mathbf{A} represents the concentrations in the mixtures:

☞ positive constraint.

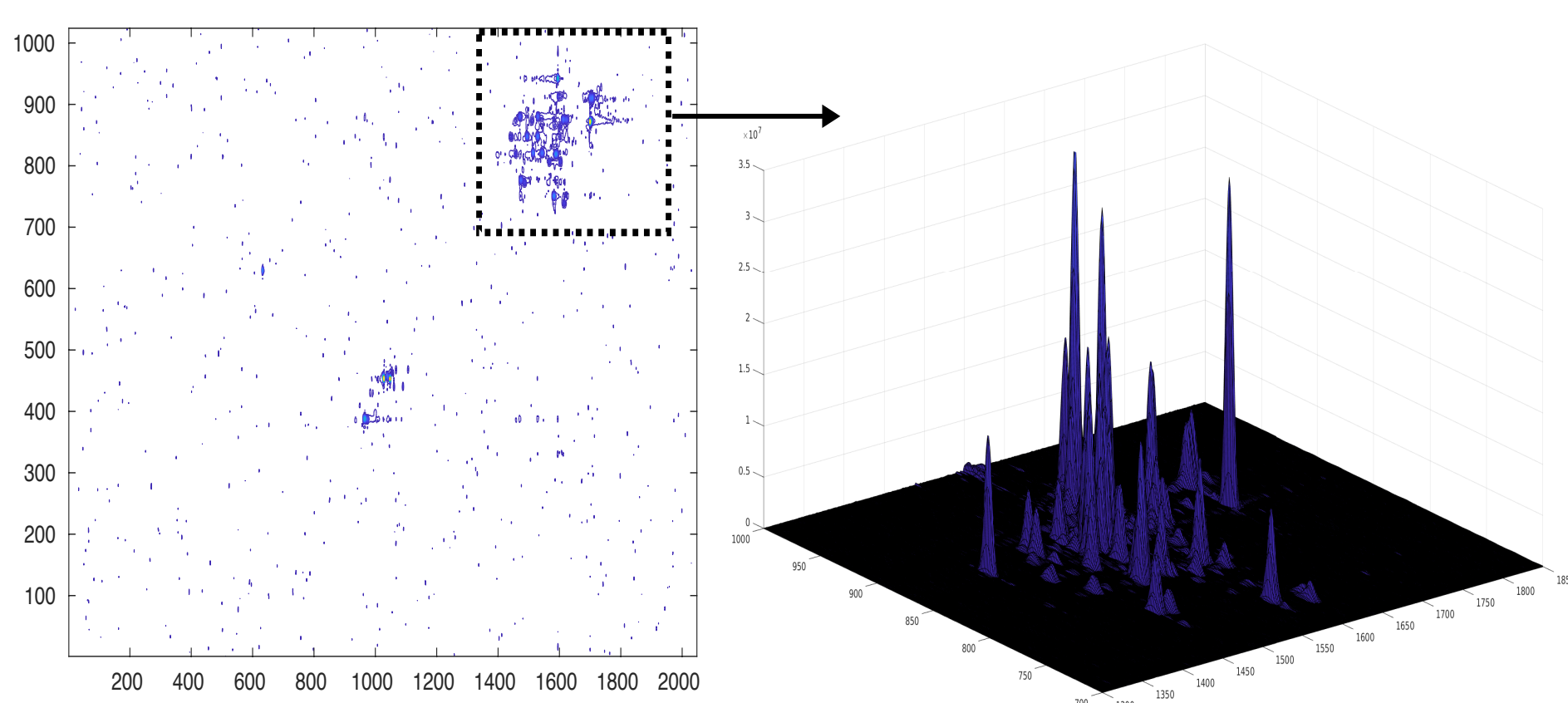
\mathbf{S} represents source coefficients:

☞ positive and sparse constraints.

• **Existing solutions:** BSS techniques (NMF, JADE, ICA, etc.).

• **Difficulties:**

- 2D NMR spectra present a high level of sparsity with a spectral overlap and poor resolution.
- Scale and order indeterminacies [3].



Practical implementation

• **Algorithm:** Block-Coordinate Variable Metric Forward-Backward

Inputs: \mathbf{X} , ≥ 0 , $\epsilon > 0$

Initialization: \mathbf{A}_0 , \mathbf{S}_0

For $k = 0, 1, \dots$

$$\begin{cases} \mathbf{A}_{k+1} = \underset{\mathbf{A}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{X} - \mathbf{A}\mathbf{S}_k\|_F^2 + \Psi_{\mathbf{A}}(\mathbf{A}) \\ \mathbf{S}_{k+1} = \underset{\mathbf{S}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{X} - \mathbf{A}_{k+1}\mathbf{S}\|_F^2 + \Psi_{\mathbf{S}}(\mathbf{S}) \\ \text{If } \|\mathbf{A}_{k+1}\mathbf{S}_{k+1} - \mathbf{X}\|_F^2 \leq \epsilon : \\ \quad \hat{\mathbf{A}} = \mathbf{A}_{k+1} \text{ and } \hat{\mathbf{S}} = \mathbf{S}_{k+1} \end{cases}$$

Outputs: $\hat{\mathbf{A}}$, $\hat{\mathbf{S}}$

• **Proximity operator**

For every $\Psi : \mathbb{R} \rightarrow]-\infty, +\infty]$, lower semi continuous (lsc) and proper function, the proximity operator of Ψ relative to the metric induced by a symmetric positive and definite matrix $\mathbf{P} \in \mathbb{R}^{L \times L}$ is defined as

$$\operatorname{prox}_{\Psi, \mathbf{P}} : \mathbf{Z} \rightarrow \underset{\mathbf{Y} \in \mathbb{R}^L}{\operatorname{argmin}} \left(\Psi(\mathbf{Y}) + \frac{1}{2} \|\mathbf{Y} - \mathbf{Z}\|_{\mathbf{P}}^2 \right)$$

$\Psi_{\mathbf{S}}$	Proximity operator
ℓ_1 (Eq. 4)	Computed [4]
Ψ_1 (Eq. 5)	Computed [2]
Ψ_2 (Eq. 6)	Approximated algorithm [5]

Optimization tools

$$\underset{\mathbf{A}, \mathbf{S}}{\operatorname{minimize}} \frac{1}{2} \|\mathbf{X} - \mathbf{A}\mathbf{S}\|_F^2 + \lambda_{\mathbf{A}} \Psi_{\mathbf{A}}(\mathbf{A}, \mathbf{S}) + \lambda_{\mathbf{S}} \Psi_{\mathbf{S}}(\mathbf{A}, \mathbf{S}) \quad (2)$$

• $\Psi_{\mathbf{A}}$: encodes the nonnegativity of \mathbf{A} .

$$(\forall \mathbf{Z} \in \mathbb{R}^L) \quad \iota_+ : z_i \rightarrow \begin{cases} 0 & \text{if } z_i \geq 0, \\ +\infty & \text{elsewhere.} \end{cases} \quad (3)$$

• $\Psi_{\mathbf{S}}$: encodes the nonnegativity and the sparsity of \mathbf{S} .

◦ classical choice: ℓ_1 norm

$$\ell_1(\mathbf{Z}) = \sum_i |z_i|. \quad (4)$$

◦ new proposition: entropy functions

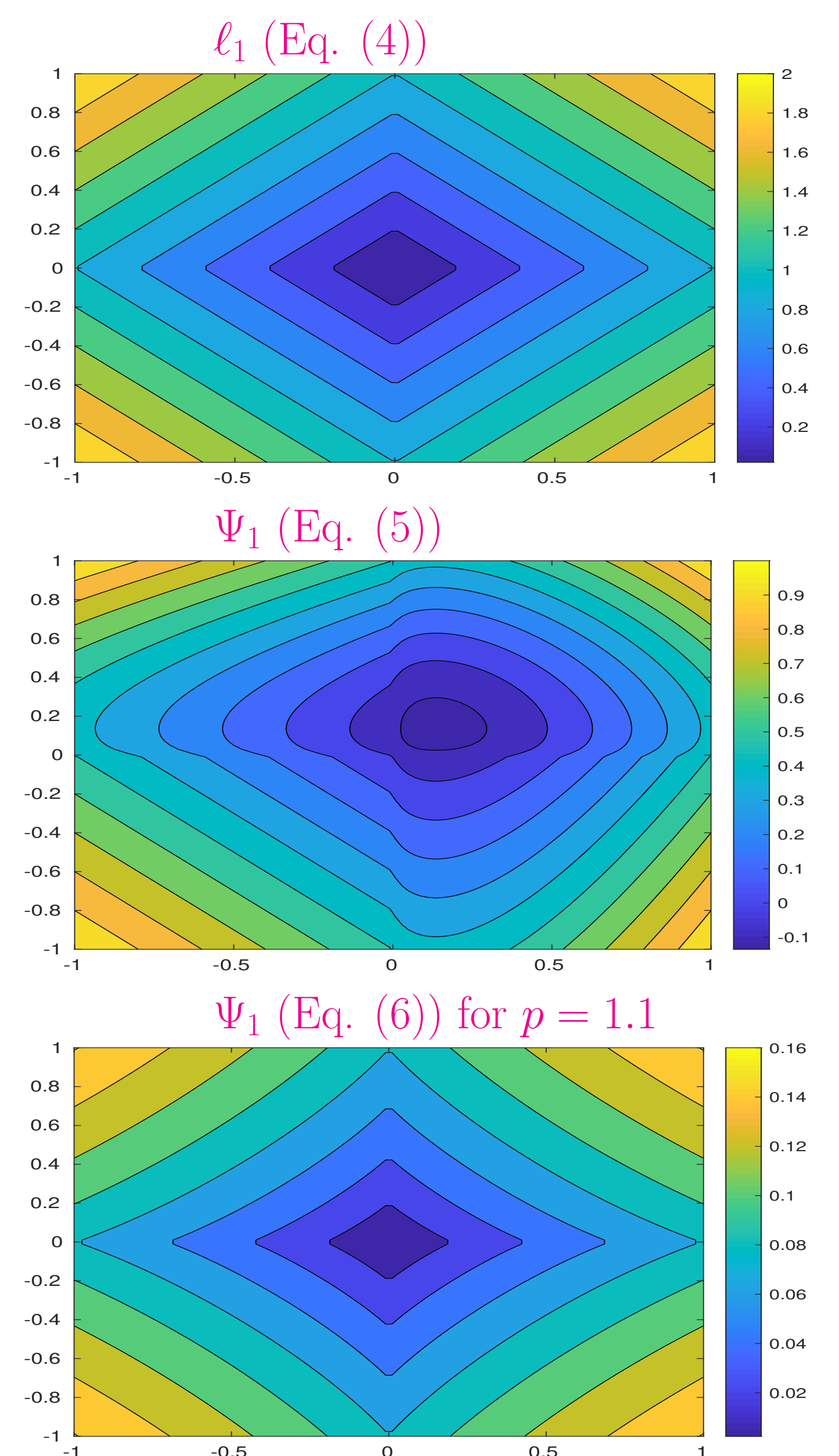
Shannon entropy + ℓ_1 norm [1, 2]:

$$\Psi_1(\mathbf{Z}) = \sum_i (z_i \log(z_i) + z_i + \iota_+(z_i)). \quad (5)$$

Generalized entropy function [5]:

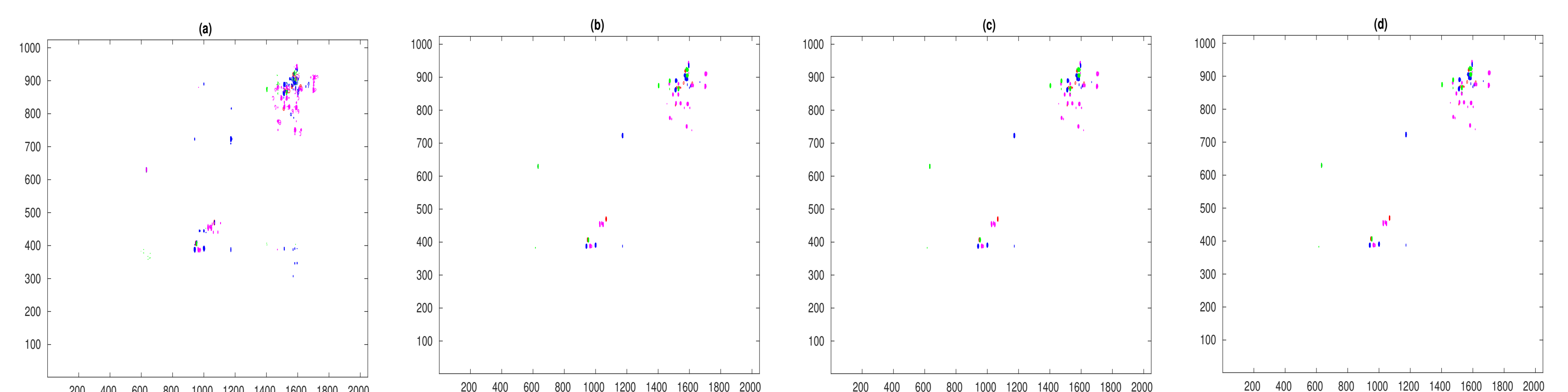
$$\Psi_2(\mathbf{Z}) = - \sum_i \frac{|z_i|^p}{\|\mathbf{Z}\|_p^p} \log \left(\frac{|z_i|^p}{\|\mathbf{Z}\|_p^p} \right). \quad (6)$$

• $\lambda_{\mathbf{A}}$, $\lambda_{\mathbf{S}}$: regularization parameters.



Results

- **2D HSQC data:** Limonene (red), Nerol (blue), α -Terpinolene (green), β -Caryophyllene (magenta)
- **Dimensions:** $N = 4$, $M = 5$, $L = (1024 \times 2048)$



Simulated data

Estimated sources using ℓ_1

Estimated sources using Ψ_1

Estimated sources using Ψ_2 ($p = 1.1$)

2D HSQC data: original (a) and estimated sources ((b)-(c)-(d))

		ℓ_1 (Eq. (4))	Ψ_1 (Eq. (5))	Ψ_2 (Eq. (6))
SIR	S1	40.10	34.11	34.04
	S2	42.98	44.58	50.23
	S3	30.72	31.58	33.18
	S4	35.87	29.03	28.13
SDR	S1	25.93	24.56	24.70
	S2	18.41	18.10	18.78
	S3	15.14	14.46	15.08
	S4	11.06	9.63	10.09
SAR	S1	26.10	25.07	25.25
	S2	18.42	18.11	18.79
	S3	15.26	14.55	15.15
	S4	11.08	9.68	10.17
Amari index		0.01	0.05	0.05

Quality performances using different regularization functions

• **Interpretations**

- ✓ The ℓ_1 , Ψ_1 and Ψ_2 ensure a good estimation of the 2D HSQC sources.
- ✓ ℓ_1 and Ψ_2 are the most efficient functions.
- ✓ Generalized entropy Ψ_2 seems to be an efficient function to estimate sparse signals.

• **Perspectives**

- ✓ Adapt the p value of the generalized entropy to better estimate 2D HSQC sources.
- ✓ Investigate on the combination of generalized entropy function Ψ_2 and ℓ_1 norm.
- ✓ Explore multiresolution transforms with new sparse NMF algorithms.

References

- [1] A. Cherni, E. Chouzenoux, and M.-A. Delsuc. Proximity operators for a class of hybrid sparsity+ entropy priors application to dosy NMR signal reconstruction. In *Int. Symp. on Image Video Communications*, pages 120–125, 2016.
- [2] A. Cherni, E. Chouzenoux, and M.-A. Delsuc. Palma, an improved algorithm for dosy signal processing. *Analyst*, 142(5):772–779, 2017.
- [3] A. Cherni, E. Piersanti, S. Anthoine, C. Chau, L. Shintu, M. Yemloul, and B. Torrèsani. Challenges in the decomposition of 2D NMR spectra of mixtures of small molecules. *Faraday Discussions*, 2019.
- [4] P.-L. Combettes and J.-C. Pesquet. Proximal splitting methods in signal processing. In *Fixed-point algorithms for inverse problems in science and engineering*, pages 185–212. Springer, 2011.
- [5] S. Huang and T.-D. Tran. Sparse signal recovery via generalized entropy functions minimization. *IEEE Trans. Signal Process.*, 67(5):1322–1337, March 2019.

SPARS 2019, Toulouse 1-4 July.