

Méthodes modernes d'analyse de données en biophysique analytique: Résolution des problèmes inverses en RMN DOSY et MS

Afef Cherni

Directeur de thèse: Marc-André Delsuc

Co-encadrante: Émilie Chouzenoux

Strasbourg, le 20 septembre 2018



Plan

- Introduction
- Analyse des données RMN de type DOSY
- Déconvolution des massifs isotopiques en MS
- Conclusions & Perspectives



Mesure
 $T(\text{objet})$

(+ perturbations)

Objet d'intérêt

Appareil de mesure
(T)

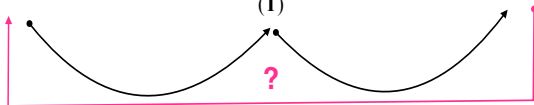


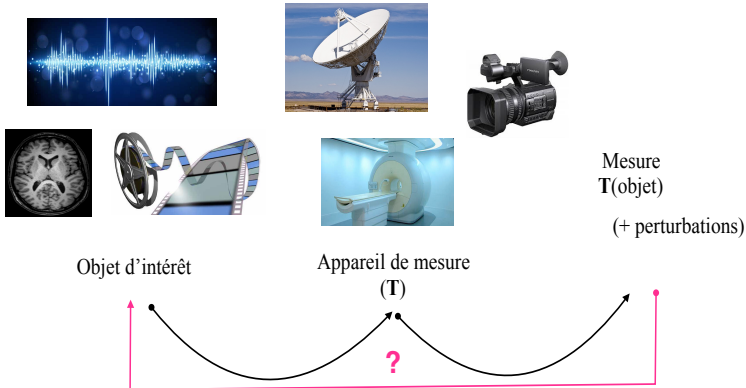


Objet d'intérêt

Appareil de mesure

(T)





- Objet d'intérêt : $\mathbf{x} \in \mathbb{R}^N$
- Mesure : $\mathbf{y} \in \mathbb{R}^M$
- Perturbations : $\mathbf{b} \in \mathbb{R}^M$

$$\mathbf{y} = T(\mathbf{x}) + \mathbf{b} \quad (1)$$

- Trouver x à partir de y : Problème inverse [**Hadamard, 1902**] :
 - bien posé : Existence - Unicité - Continuité
 - mal posé



- Trouver \mathbf{x} à partir de \mathbf{y} : Problème inverse [Hadamard, 1902] :
 - bien posé : Existence - Unicité - Continuité
 - mal posé
- ▶ Si $M > N$: Méthodes des moindres carrés :

$$\underset{\mathbf{x} \in \mathbb{R}^N}{\text{minimiser}} \quad \|\mathbf{T}(\mathbf{x}) - \mathbf{y}\|^2 \quad (2)$$

- Trouver \mathbf{x} à partir de \mathbf{y} : Problème inverse [Hadamard, 1902] :
 - bien posé : Existence - Unicité - Continuité
 - mal posé

- ▶ Si $M > N$: Méthodes des moindres carrées :

$$\underset{\mathbf{x} \in \mathbb{R}^N}{\text{minimiser}} \quad \|\mathbf{T}(\mathbf{x}) - \mathbf{y}\|^2 \quad (2)$$

- ▶ Si $M = N$: Inversion de l'opérateur de mesure :

$$\hat{\mathbf{x}} = \mathbf{T}^{-1}(\mathbf{y}) \quad (3)$$



- Trouver \mathbf{x} à partir de \mathbf{y} : Problème inverse [Hadamard, 1902] :
 - bien posé : Existence - Unicité - Continuité
 - mal posé

- ▶ Si $M > N$: Méthodes des moindres carrées :

$$\underset{\mathbf{x} \in \mathbb{R}^N}{\text{minimiser}} \quad \|\mathbf{T}(\mathbf{x}) - \mathbf{y}\|^2 \quad (2)$$

- ▶ Si $M = N$: Inversion de l'opérateur de mesure :

$$\hat{\mathbf{x}} = \mathbf{T}^{-1}(\mathbf{y}) \quad (3)$$

- ▶ Si $M < N$: Ajout d'une information *a priori* f :

- ◊ Approche sous contraintes :

$$\underset{\mathbf{x} \in \mathbb{R}^N}{\text{minimiser}} \quad f(\mathbf{x}) \quad \text{tel que } \|\mathbf{T}(\mathbf{x}) - \mathbf{y}\| \leq \eta \quad (4)$$

- ◊ Approche de régularisation :

$$\underset{\mathbf{x} \in \mathbb{R}^N}{\text{minimiser}} \quad \|\mathbf{T}(\mathbf{x}) - \mathbf{y}\|^2 + \lambda f(\mathbf{x}) \quad (5)$$

- Trouver \mathbf{x} à partir de \mathbf{y} : Problème inverse [Hadamard, 1902] :
 - bien posé : Existence - Unicité - Continuité
 - mal posé

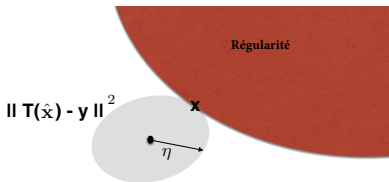
- ▶ Si $M > N$: Méthodes des moindres carrés :

$$\underset{\mathbf{x} \in \mathbb{R}^N}{\text{minimiser}} \quad \|\mathbf{T}(\mathbf{x}) - \mathbf{y}\|^2 \quad (2)$$

- ▶ Si $M = N$: Inversion de l'opérateur de mesure :

$$\hat{\mathbf{x}} = \mathbf{T}^{-1}(\mathbf{y}) \quad (3)$$

- ▶ Si $M < N$: Ajout d'une information *à priori* f :



Exemples :

- Entropie maximale
- Positivité
- Énergie (Tikhonov)
- etc.

- Trouver \mathbf{x} à partir de \mathbf{y} : Problème inverse [Hadamard, 1902] :
 - bien posé : Existence - Unicité - Continuité
 - mal posé

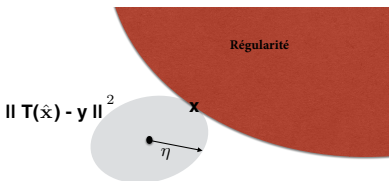
- ▶ Si $M > N$: Méthodes des moindres carrés :

$$\underset{\mathbf{x} \in \mathbb{R}^N}{\text{minimiser}} \quad \|\mathbf{T}(\mathbf{x}) - \mathbf{y}\|^2 \quad (2)$$

- ▶ Si $M = N$: Inversion de l'opérateur de mesure :

$$\hat{\mathbf{x}} = \mathbf{T}^{-1}(\mathbf{y}) \quad (3)$$

- ▶ Si $M < N$: Ajout d'une information *à priori* f :



Exemples :

- Entropie maximale
- Positivité
- Énergie (Tikhonov)
- etc.

Objectif

Résoudre les problèmes inverses mal-posés en :

- Résonance magnétique nucléaire de type DOSY
- Spectrométrie de masse

Objectif

Résoudre les problèmes inverses mal-posés en :

- Résonance magnétique nucléaire de type DOSY
- Spectrométrie de masse

Notations

- scalaire : $x \in \mathbb{R}$
- vecteur : $\mathbf{x} = (x_1, x_2, \dots, x_N) \in \mathbb{R}^N$
- matrice : $\mathbf{A} \in \mathbb{R}^{M \times N}$ (M lignes, N colonnes)

Première partie :
Analyse des données RMN de type DOSY

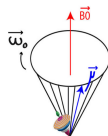
RMN de type DOSY

▶ RMN : Résonance Magnétique Nucléaire

- une technique qui permet d'étudier les interactions chimiques, déterminer les déplacement chimiques, etc.

○ Mouvement de procession de Larmor

- \vec{B}_0 : Champ magnétique.
- $\vec{\gamma}$: Moment du spin.
- $\vec{\omega}_0 = \gamma \vec{B}_0$: Fréquence de Larmor.



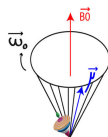
RMN de type DOSY

► RMN : Résonance Magnétique Nucléaire

- une technique qui permet d'étudier les interactions chimiques, déterminer les déplacement chimiques, etc.

○ Mouvement de procession de Larmor

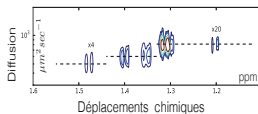
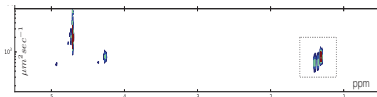
- \vec{B}_0 : Champ magnétique.
- $\vec{\gamma}$: Moment du spin.
- $\vec{\omega}_0 = \gamma \vec{B}_0$: Fréquence de Larmor.



► DOSY : Diffusion Ordered SpectroscopY

- une expérience RMN qui mesure les coefficients de diffusion

- Représentation 2D
- Déplacement chimique (axe horizontal)
- Coefficients de diffusion (axe vertical)



Problématique

- Signal mesuré **monodisperse** :

$$I(q) = I_0 \exp(-D\Delta q^2)$$

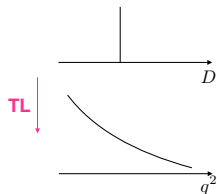
- I : Intensité mesurée.
- Δ : Temps de diffusion.
- $q = \gamma\delta g$: Phase de dispersion.

Problématique

- Signal mesuré **monodisperse** :

$$I(q) = I_0 \exp(-D\Delta q^2)$$

→ profil **parcimonieux**



Problématique

- Signal mesuré **monodisperse** :

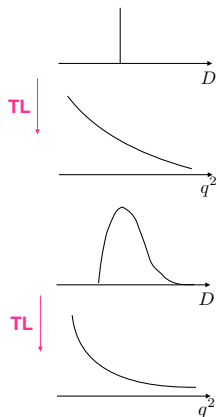
$$I(q) = I_0 \exp(-D\Delta q^2)$$

→ profil **parcimonieux**

- Signal mesuré **polydisperse** :

$$I(q) = \int_{D_{\min}}^{D_{\max}} x(D) \exp(-D\Delta q^2) dD$$

→ profil **étalé**



Problématique

- Signal mesuré **monodisperse** :

$$I(q) = I_0 \exp(-D\Delta q^2)$$

- profil **parcimonieux**

- Signal mesuré **polydisperse** :

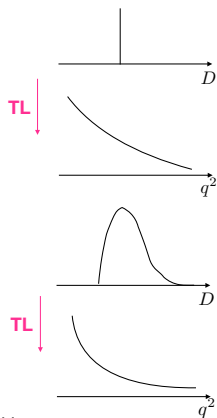
$$I(q) = \int_{D_{\min}}^{D_{\max}} x(D) \exp(-D\Delta q^2) dD$$

- profil **étalé**

► Modélisation discrète

$$y = Hx + b$$

- $x \in \mathbb{R}^N$: signal cible
- $H = (H_{i,j})_{1 \leq i \leq M, 1 \leq j \leq N}$: matrice de mesure
- $H_{i,j} = \exp(-D_j \Delta q_i^2)$
- $y \in \mathbb{R}^M$: mesure
- $b \in \mathbb{R}^M$: bruit



Problématique

- Signal mesuré **monodisperse** :

$$I(q) = I_0 \exp(-D\Delta q^2)$$

→ profil **parcimonieux**

- Signal mesuré **polydisperse** :

$$I(q) = \int_{D_{\min}}^{D_{\max}} x(D) \exp(-D\Delta q^2) dD$$

→ profil **étalé**

► Modélisation discrète

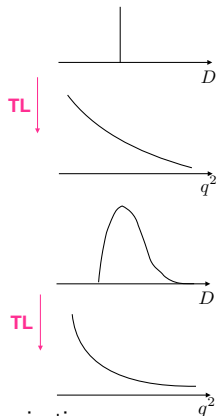
$$y = Hx + b$$

✗ $M \ll N$: sous-détermination

✗ Opérateur de Laplace : mauvais conditionnement

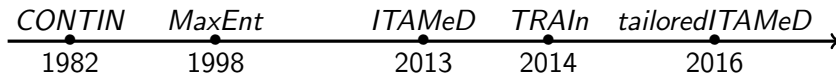
✗ Présence de bruit : instabilité

→ Problème **inverse mal-posé**



État de l'art

► Méthodes proposées



État de l'art

► Méthodes proposées



CONTIN : CONTINuous diffusion coefficient [[Provencher, 1982](#)]

- Principe : Méthode des moindres carrés avec contrainte de positivité

$$\underset{\mathbf{x} \in \mathbb{R}^N}{\text{minimiser}} \quad \|\mathbf{H}\mathbf{x} - \mathbf{y}\|^2 \quad \text{tel que} \quad \mathbf{x} \geq 0$$

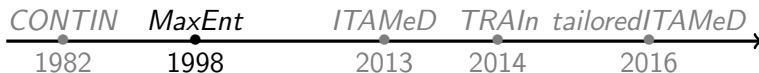
- Rappel :

$$(\forall \mathbf{x} = (x_1, x_2, \dots, x_N) \in \mathbb{R}^N) \quad \mathbf{x} \geq 0 \Leftrightarrow x_i \geq 0$$

- Langage : Fortran

État de l'art

► Méthodes proposées



MaxEnt : Maximum **EN**tropy [Delsuc and Malliavin, 1998]

- Principe : Régularisation avec l'entropie maximale

$$\underset{\mathbf{x} \in \mathbb{R}^N}{\text{minimiser}} \quad \frac{1}{2} \|\mathbf{H}\mathbf{x} - \mathbf{y}\|^2 + \lambda \text{ent}(\mathbf{x})$$

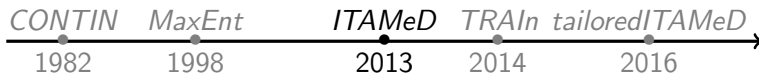
- **Rappel** : Maximiser l'entropie \Leftrightarrow Minimiser - l'entropie :

$$(\forall \mathbf{x} \in \mathbb{R}^N) \quad \text{ent}(\mathbf{x}) = \sum_{i=1}^N \text{ent}(x_i) \quad \text{avec} \quad \text{ent}(x_i) = \begin{cases} x_i \log x_i & \text{si } x_i > 0 \\ 0 & \text{si } x_i = 0 \\ +\infty & \text{sinon} \end{cases}$$

- Langage : Fortran

État de l'art

► Méthodes proposées



ITAMeD : Iterative Thresholding Algorithm for Multi-exponentiel Decay [[Urbanczyk et al., 2013](#)]

- Principe : Régularisation avec la norme ℓ_1

$$\underset{\mathbf{x} \in \mathbb{R}^N}{\text{minimiser}} \quad \frac{1}{2} \|\mathbf{H}\mathbf{x} - \mathbf{y}\|^2 + \lambda \ell_1(\mathbf{x})$$

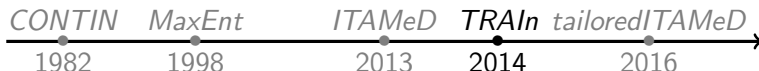
- Rappel :

$$(\forall \mathbf{x} \in \mathbb{R}^N) \quad \ell_1(\mathbf{x}) = \sum_{i=1}^N |x_i|$$

- Langage : MATLAB

État de l'art

► Méthodes proposées



TRAIIn : Trust-Region **A**lgorithm for **I**nversion of Molecular Diffusion NMR [Xu and Zhang, 2014]

- Principe : Minimisation sous contrainte adaptative

$$\underset{\mathbf{x} \in \mathbb{R}^N}{\text{minimiser}} \quad \frac{1}{2} \|\mathbf{H}\mathbf{x}^{(k)} - \mathbf{y}\|^2 \quad \text{tel que} \quad \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|_2^2 \leq r^{(k)}$$

(r_k désigne le rayon de la région de confiance à la $k^{\text{ème}}$ itération)

- Langage : MATLAB

État de l'art

► Méthodes proposées



tailored ITAMeD : Iterative Thresholding Algorithm for Multi-exponentiel Decay [Urbańczyk et al., 2016]

- Principe : Régularisation avec la norme ℓ_p^p

$$\underset{x \in \mathbb{R}^N}{\text{minimiser}} \quad \frac{1}{2} \| \mathbf{H}x - y \|^2 + \lambda \ell_p^p(x)$$

- Rappel :

$$(\forall x \in \mathbb{R}^N) \quad \ell_p(x) = \left(\sum_{i=1}^N |x_i|^p \right)^{\frac{1}{p}}$$

- Langage : MATLAB

Comparaison

Algorithmes	Complexité de calcul	Robustesse % bruit	Profil		Stratégie d'optimisation	À priori physique	Preuve de convergence
			monodisperse	polydisperse			
CONTIN	✗	✗	=	=	programmation quadratique	Positivité	✗
MaxEnt	✗	✓	=	✓	régularisation adaptative	Entropie	✓
ITAMeD	=	=	✓	✗	seuillage itératif	Parcimonie	✓
TRAIN	✓	=	=	✓	régularisation itérative	non explicite	✗
tailored ITAMeD	✗	=	✓	✓	moindre carrée pondérée	Entropie/Parcimonie	✗

Comparaison

Algorithmes	Complexité de calcul	Robustesse % bruit	Profil		Stratégie d'optimisation	À priori physique	Preuve de convergence
			monodisperse	polydisperse			
CONTIN	✗	✗	=	=	programmation quadratique	Positivité	✗
MaxEnt	✗	✓	=	✓	régularisation adaptative	Entropie	✓
ITAMeD	=	=	✓	✗	seuillage itératif	Parcimonie	✓
TRAln	✓	=	=	✓	régularisation itérative	non explicite	✗
tailored ITAMeD	✗	=	✓	✓	moindre carrée pondérée	Entropie/Parcimonie	✗

✗ Aucune méthode n'est satisfaisante !

Objectif

Développer une approche **rapide** et **robuste** avec un **modèle à priori adapté** et des **garanties de convergence solides**, qui permet de résoudre le problème inverse en DOSY **quelque soit le type de données** : monodisperse et polydisperse.

Nouvelle approche

► Principe : Minimisation sous contrainte

$$\underset{\mathbf{x} \in \mathbb{R}^N}{\text{minimiser}} \quad \Psi(\mathbf{x}) \quad \text{tel que} \quad \|\mathbf{H}\mathbf{x} - \mathbf{y}\| \leq \eta$$

avec $\eta = \sigma\sqrt{M}$ et σ (supposé connu) : l'écart type du bruit.

Nouvelle approche

► Principe : Minimisation sous contrainte

$$\underset{\mathbf{x} \in \mathbb{R}^N}{\text{minimiser}} \quad \Psi(\mathbf{x}) \quad \text{tel que} \quad \|\mathbf{H}\mathbf{x} - \mathbf{y}\| \leq \eta$$

avec $\eta = \sigma\sqrt{M}$ et σ (supposé connu) : l'écart type du bruit.

► Régularisation : hybride

$$(\forall \mathbf{x} \in \mathbb{R}^N, \alpha \geq 0, \beta \geq 0) \quad \Psi(\mathbf{x}) = \underbrace{\alpha \Psi_1(\mathbf{x})}_{\text{Entropie}} + \underbrace{\beta \Psi_2(\mathbf{x})}_{\text{Parcimonie}}$$

Nouvelle approche

► Principe : Minimisation sous contrainte

$$\underset{\mathbf{x} \in \mathbb{R}^N}{\text{minimiser}} \quad \Psi(\mathbf{x}) \quad \text{tel que} \quad \|\mathbf{H}\mathbf{x} - \mathbf{y}\| \leq \eta$$

avec $\eta = \sigma\sqrt{M}$ et σ (supposé connu) : l'écart type du bruit.

► Régularisation : hybride

$$(\forall \mathbf{x} \in \mathbb{R}^N, \alpha \geq 0, \beta \geq 0) \quad \Psi(\mathbf{x}) = \underbrace{\alpha \Psi_1(\mathbf{x})}_{\text{Entropie}} + \underbrace{\beta \Psi_2(\mathbf{x})}_{\text{Parcimonie}}$$

- Si $\alpha = 0$: L'algorithme ITAMeD
- Si $\beta = 0$: L'algorithme MaxEnt

Nouvelle approche

► Entropie

$$\Psi_1: \mathbb{R}^N \rightarrow]-\infty, +\infty]$$

$$\mathbf{x} \rightarrow \sum_{i=1}^N \psi_1(x_i)$$

Critère d'entropie ψ_1	
$x \log(x) + \iota_{[0,+\infty)}(x)$	<i>Shannon</i>
$-\log(x) + \iota_{[0,+\infty)}(x)$	<i>Burg</i>

Nouvelle approche

► Entropie

$$\Psi_1: \mathbb{R}^N \rightarrow]-\infty, +\infty]$$

$$\mathbf{x} \rightarrow \sum_{i=1}^N \psi_1(x_i)$$

Critère d'entropie ψ_1	
$x \log(x) + \iota_{[0,+\infty)}(x)$	<i>Shannon</i>
$-\log(x) + \iota_{[0,+\infty)}(x)$	<i>Burg</i>

► Parcimonie

$$\Psi_2: \mathbb{R}^N \rightarrow \mathbb{R}$$

$$\mathbf{x} \rightarrow \sum_{i=1}^N \psi_2(x_i)$$

Critère de parcimonie ψ_2	
$ x ^0$	l_0
$ x $	l_1
$\log(\delta + x), \delta > 0$	<i>log-sum</i>
$\log(\delta + x^2), \delta > 0$	<i>Cauchy</i>

Nouvelle approche

► Entropie

$$\Psi_1: \mathbb{R}^N \rightarrow]-\infty, +\infty]$$

$$\mathbf{x} \rightarrow \sum_{i=1}^N \psi_1(x_i)$$

Critère d'entropie ψ_1	
$x \log(x) + \iota_{[0,+\infty)}(x)$	<i>Shannon</i>
$-\log(x) + \iota_{[0,+\infty)}(x)$	<i>Burg</i>

► Parcimonie

$$\Psi_2: \mathbb{R}^N \rightarrow \mathbb{R}$$

$$\mathbf{x} \rightarrow \sum_{i=1}^N \psi_2(x_i)$$

Critère de parcimonie ψ_2	
$ x ^0$	l_0
$ x $	l_1
$\log(\delta + x), \delta > 0$	<i>log-sum</i>
$\log(\delta + x^2), \delta > 0$	<i>Cauchy</i>

minimiser $\Psi(\mathbf{x})$ tel que $\|\mathbf{H}\mathbf{x} - \mathbf{y}\| \leq \eta$
 $\mathbf{x} \in \mathbb{R}^N$

Opérateur proximal

Définition (<http://proximity-operator.net/>)

Soit $\Psi : \mathbb{R} \rightarrow]-\infty, +\infty]$ une fonction propre semi-continue inférieurement (lsc). L'opérateur proximal de Ψ est défini comme suit [*Hiriart-Urruty and Lemaréchal, 1993, Bauschke and Combettes, 2011*] :

$$\begin{aligned} \text{prox}_{\Psi} : \mathbb{R}^N &\mapsto \mathbb{R}^N \\ \mathbf{x} &\mapsto \underset{\mathbf{y} \in \mathbb{R}^N}{\text{Argmin}} \left(\Psi(\mathbf{y}) + \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|^2 \right) \end{aligned}$$

Opérateur proximal

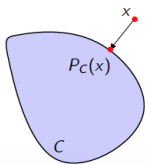
Définition

Soit $\Psi : \mathbb{R} \rightarrow]-\infty, +\infty]$ une fonction propre semi-continue inférieurement (lsc). L'opérateur proximal de Ψ est défini comme suit [*Hiriart-Urruty and Lemaréchal, 1993, Bauschke and Combettes, 2011*] :

$$\text{prox}_{\Psi} : \mathbb{R}^N \mapsto \mathbb{R}^N$$

$$x \mapsto \underset{y \in \mathbb{R}^N}{\text{Argmin}} \left(\Psi(y) + \frac{1}{2} \|y - x\|^2 \right)$$

- Exemple : Soit C un espace fermé non vide.



$$\iota(x) : x \rightarrow \begin{cases} 0 & \text{si } x \in C \\ +\infty & \text{sinon} \end{cases}$$

$$(\forall x \in \mathbb{R}^N) \quad \text{prox}_{\iota_C}(x) = \underset{y \in C}{\text{argmin}} \frac{1}{2} \|x - y\|^2 = P_C(x)$$

Opérateur proximal

Définition

Soit $\Psi : \mathbb{R} \rightarrow]-\infty, +\infty]$ une fonction propre semi-continue inférieurement (lsc). L'opérateur proximal de Ψ est défini comme suit [*Hiriart-Urruty and Lemaréchal, 1993, Bauschke and Combettes, 2011*] :

$$\begin{aligned} \text{prox}_{\Psi} : \mathbb{R}^N &\mapsto \mathbb{R}^N \\ \mathbf{x} &\mapsto \underset{\mathbf{y} \in \mathbb{R}^N}{\text{Argmin}} \left(\Psi(\mathbf{y}) + \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|^2 \right) \end{aligned}$$

Propriété de séparabilité

Pour tout $i \in \{1, \dots, N\}$ et pour $\mathbf{x} = (x_1, \dots, x_N)$:

$$\text{prox}_{\Psi}(\mathbf{x}) = (p(x_i))_{1 \leq i \leq N}$$

$$\text{avec } p(x_i) = \text{prox}_{\alpha\psi_1 + \beta\psi_2}(x_i)$$

Opérateurs proximaux pour une régularisation hybride (1/2)

► Cas de l'entropie de Shannon (ψ_1)

ψ_2	$\text{prox}_{\alpha\psi_1+\beta\psi_2}(x) \quad / \quad x \in \mathbb{R}$
$\beta = 0$	$\alpha\mathcal{W}((1/\alpha) \exp((x/\alpha) - 1))$
ℓ_1	$\alpha\mathcal{W}((1/\alpha) \exp((x - \beta)/\alpha - 1))$
ℓ_0	$\begin{cases} p & \text{si } \beta < \bar{\beta} \\ \{0, p\} & \text{si } \beta = \bar{\beta} \\ 0 & \text{sinon} \end{cases}$ <p>tel que $p = \alpha\mathcal{W}((1/\alpha) \exp((1/\alpha) - 1))$ et $\bar{\beta} = (1/2)p^2 + \alpha p \in]0, +\infty[$</p>
<i>log-sum</i>	$\underset{p \in]0, +\infty[\text{ t.q. } \varphi(p)=0}{\text{Argmin}} \quad \left((1/2)(x - p)^2 + \psi(p) \right)$ <p>avec $\varphi(p) = p^2 + (\delta - x + \alpha)p + \alpha(\delta + p) \log(p) + \delta(\alpha - x) + \beta$</p>
<i>Cauchy</i>	$\underset{p \in]0, +\infty[\text{ t.q. } \varphi(p)=0}{\text{Argmin}} \quad \left((1/2)(x - p)^2 + \psi(p) \right)$ <p>avec $\varphi(p) = p^3 + (\alpha - x)p^2 + (\delta + 2\beta)p + \alpha(\delta + p^2) \log(p) + \delta(\alpha - x)$</p>

\mathcal{W} désigne la fonction de Lambert [Corless et al., 1996] : l'inverse de $f : z \rightarrow ze^z$.

Opérateurs proximaux pour une régularisation hybride (2/2)

► Cas de l'entropie de Burg (ψ_1)

ψ_2	$\text{prox}_{\alpha\psi_1+\beta\psi_2}(x)$
$\beta = 0$	$(x + \sqrt{x^2 + 4\alpha})/2$
l_1	$(x - \beta + \sqrt{(\beta - x)^2 + 4\alpha})/2$
l_0	$(x + \sqrt{x^2 + 4\alpha})/2$
<i>log-sum</i>	<p>Argmin $\left(\frac{1}{2}(x - p)^2 + \psi(p)\right)$ $p \in]0, +\infty[\text{ t.q. } \varphi(p) = 0$</p> <p>avec $\varphi(p) = p^3 + (\delta - x)p^2 + p(\beta - \delta x - \alpha) - \delta\alpha$</p>
<i>Cauchy</i>	<p>Argmin $\left(\frac{1}{2}(x - p)^2 + \psi(p)\right)$ $p \in]0, +\infty[\text{ t.q. } \varphi(p) = 0$</p> <p>avec $\varphi(p) = p^4 - xp^3 + (\delta + 2\beta - \alpha)p^2 - \delta xp - \delta\alpha$</p>

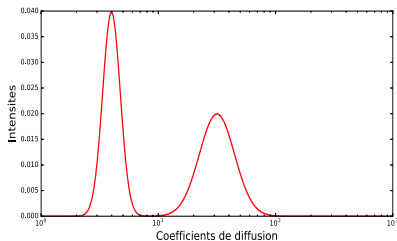
Signaux synthétiques

- ▶ Approches testées : Régularisations hybrides **convexes** :

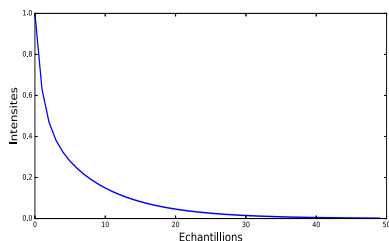
$$\alpha \text{ Entropie de Shannon} + \beta l_1$$

$$\alpha \text{ Entropie Burg} + \beta l_1$$

- ▶ $M = 50, N = 200$
- ▶ $\alpha + \beta = 1, \alpha \in [0, 1]$
- ▶ Algorithme utilisé : **PPXA+** (approche proximale)
- ▶ Langage : **Python**



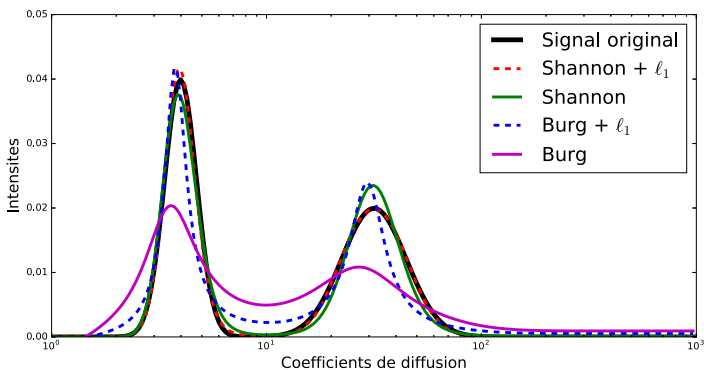
Signal simulé



Mesure

Résultats numériques (1/2)

- Résultats avec $\sigma = 10^{-5}$ (niveau de bruit).



Reconstruction du signal avec différentes fonctions de régularisation

Résultats numériques (2/2)

- ▶ Résultats avec différents niveaux de bruit

σ	Shannon	Shannon + ℓ_1	Burg	Burg + ℓ_1
10^{-2}	12.45	13.16	12.92	12.92
10^{-3}	18.16	20.86	12.11	13.44
10^{-4}	20.87	25.95	12.03	15.53

Qualité de reconstruction en dB pour différents choix de la fonction de régularisation ($\frac{\|X_{sim}\|}{\|X_{sim} - X_{calc}\|}$ en dB).

↪ Choix optimal : **Entropie de Shannon + ℓ_1**

Publication 1

Afef Cherni, Emilie Chouzenoux, Marc-André Delsuc, **Proximity for a class of hybrid sparsity + entropy prior. Application to DOSY NMR signal reconstruction**, In Proceedings of the ISIVC, Tunis-Tunisia, November 2016.

Publication 1

Afef Cherni, Emilie Chouzenoux, Marc-André Delsuc, **Proximity for a class of hybrid sparsity + entropy prior. Application to DOSY NMR signal reconstruction**, In Proceedings of the ISIVC, Tunis-Tunisia, November 2016.

► **Régularisation hybride :**

$$\Psi(\mathbf{x}) = \alpha \text{ent}_{\text{Shannon}}(\mathbf{x}) + (1 - \alpha)\ell_1(\mathbf{x})$$

► **PALMA** : **P**roximal **A**lgorithm for ℓ_1 combined with **MA**xent prior.

$$(\forall \mathbf{x} \in \mathbb{R}^N) \quad \text{prox}_{\Psi}(\mathbf{x}) = (p(x_i))_{1 \leq i \leq N}$$

avec :

$$p(x_i) = \begin{cases} \frac{\alpha}{a} \mathcal{W} \left[\frac{a}{\alpha} \exp\left(\frac{ax_i - a(1-\alpha)}{\alpha}\right) + \log(a) - 1 \right] & \text{si } \alpha \in]0, 1] \\ \text{sign}(x_i) \max(|x_i| - 1, 0) & \text{si } \alpha = 0 \end{cases}$$

Comparaison de PALMA avec d'autres algorithmes

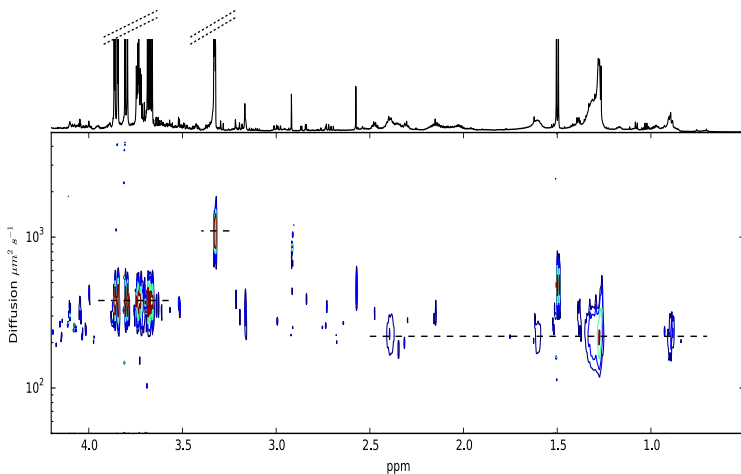
- Signal simulé : polydisperse, $N = 256$, $M = 64$.

Algorithme	niveau du bruit		
	0.1%	0.01%	0.001%
ITAMeD	18.65	29.04	29.40
tailored ITAMeD	25.26	36.69	37.08
TRAI _n	28.63	26.53	19.47
PALMA avec $\alpha = 0.01$	28.75	41.69	53.25

Qualité de reconstruction du signal avec différents algorithmes en variant le niveau de bruit.

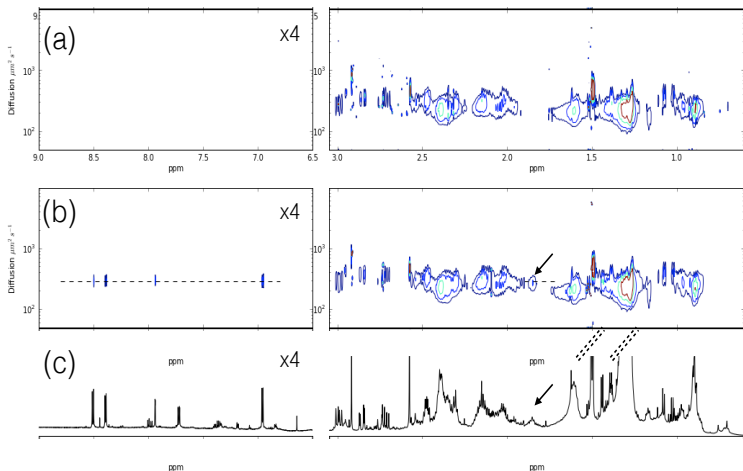
→ PALMA : algorithme robuste par rapport au niveau de bruit.

Application de PALMA (1/2)



Application PALMA sur des signaux d'extrait d'algues. La ligne horizontale indique les chaînes d'acides gras à $220 \mu\text{m}^2 \text{s}^{-1}$, le glycérol à $380 \mu\text{m}^2 \text{s}^{-1}$ et le méthanol à $1100 \mu\text{m}^2 \text{s}^{-1}$

Application de PALMA (2/2)



Comparaison de deux expériences DOSY : extrait d'algues : a) résultat de l'expérience initiale, b) même expérience avec 0.16 mg mL^{-1} de chloroquine, c) le spectre 1D de l'extrait d'algue avec la chloroquine

Publication 2

Afel Cherni, Emilie Chouzenoux, Marc-André Delsuc, **Improved algorithm for DOSY signal processing**, Analyst, Vol. 142, No. 5, pages 772-779, 2016.

- ▶ **Nouvel algorithme PALMA** : robuste, fiable, et précis.
- ▶ **Serveur Web** : <http://palma.labo.igbmc.fr>.

PALMA for DOSY Analysis

Home Data analysis Methods How-To Contact us

Welcome

This Web service allows to analyse your DOSY experiment in simple, efficient and easy way.
 This site is associated to the work:

Improved algorithm for DOSY signal processing.
 Afel Cherni, Emilie Chouzenoux, and Marc-André Delsuc
 Analyst (2017) DOI: 10.1039/C6AN01902A
 It introduces the MacEnt / F_2 hybrid regularization used in this approach, as well as the PALMA algorithm to solve the associated inverse problem.
 Visit the [Methods](#) page for details.

This web site allows to use the PALMA algorithm to process you DOSY experiments.

To process a data-set

Use the "Data Analysis" page to send your job, you just have to indicate your e-mail address and the type of molecules you're studying.

Terms of use

By using this service, you agree to acknowledge this program when presenting results using this program. In particular, you should always cite the reference above.
 This server is provided as an aid for the analysis of DOSY. As it is, it might help you, and it is our goal to provide you with the best possible results. However, it is not meant to provide an "official" answer, and many possible errors are possible. **Please, use it at your own risk!** In particular we do not take any special care to protect you data for being viewed or copied by unwanted people. Please do not submit here critical data.

Conclusion 1

✓ Nouveaux opérateurs proximaux traitant la combinaison entre l'entropie et la parcimonie

[Cherni et al., 2016].

✓ Nouvelle approche pour résoudre le problème de reconstruction des signaux DOSY quelque soit le type des données

[Cherni et al., 2017].

✓ Nouveau serveur web gratuit, disponible en ligne pour toute analyse DOSY

<https://palma.labo.igbmc.fr>.

Deuxième partie :
Déconvolution des massifs isotopiques en MS

MS

- ▶ **Spectrométrie de Masse :**
 - Une méthode de mesure des rapports masse/charge
 - 1919, **Joseph John Thomson.**

MS

▶ **Spectrométrie de Masse :**

- Une méthode de mesure des rapports masse/charge

- 1919, **Joseph John Thomson.**

● **Utilité :**

- Pharmacie & Clinique / Métabolomique / Sécurité alimentaire /
Environnement.

MS

▶ Spectrométrie de Masse :

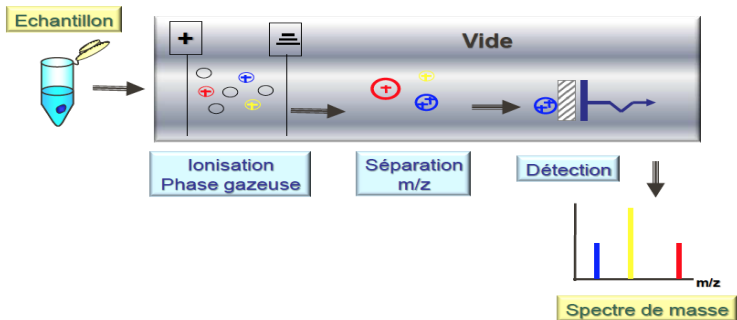
- Une méthode de mesure des rapports masse/charge

- 1919, **Joseph John Thomson**.

● Utilité :

- Pharmacie & Clinique / Métabolimique / Sécurité alimentaire / Environnement.

● Principe :



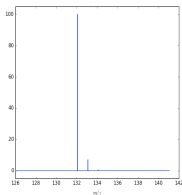
Problématique (1/3)

- ▶ **Application d'intérêt** : L'analyse protéomique.
Déterminer la masse monoisotopique des **protéines**.

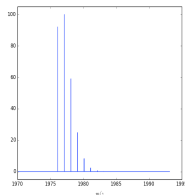
Rappel : Protéines

- Grosses molécules formées de longues chaînes d'acides aminés.
- Formule générique : $C_{N_C} H_{N_H} O_{N_O} N_{N_N} S_{N_S}$
- Massifs isotopiques : différentes probabilités d'apparition.

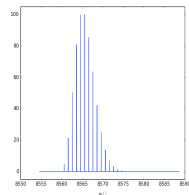
▶ Exemple :



Acide aminé



Peptide



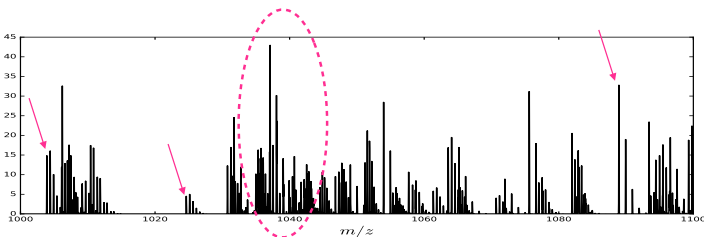
Protéine Ubiquitine

Problématique (1/3)

- **Application d'intérêt** : L'analyse protéomique.
Déterminer la masse monoisotopique des **protéines**.

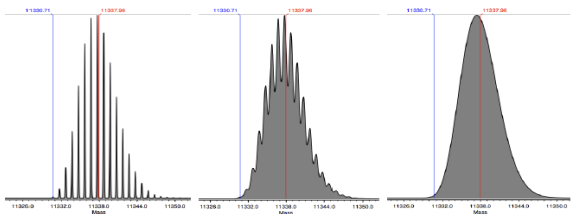
Rappel : Protéines

- Grosses molécules formées de longues chaînes d'acides aminés.
- Formule générique : $C_{N_C} H_{N_H} O_{N_O} N_{N_N} S_{N_S}$
- Massifs isotopiques : différentes probabilités d'apparition.



Problématique (2/3)

- ✗ Massif isotopique \rightsquigarrow plusieurs pics moléculaires.
- ✗ Massif isotopique multi-chargé \rightsquigarrow plusieurs profils superposés.
- ✗ Basse résolution \rightsquigarrow pics non détectables



- ✗ Niveau de bruit élevé
- ✗ Spectre de masse de grande dimension ($> 500k$)

Objectif

Déterminer les masses monoisotopiques à partir d'un spectre de masse mesuré avec ces conditions, pour identifier les protéines.

Problématique (3/3)

- ▶ Modèle Averagine [Senko et al., 1994] :
une modélisation des protéines qui permet de déterminer la
distribution la plus adéquate à une masse donnée :

$$\begin{aligned} \mathcal{A} : \mathbb{R}_+ \times \mathbb{N}^* &\rightarrow \mathbb{R}_+ \\ (m, z) &\rightarrow \mathbf{d}(m, z) \end{aligned} \quad (4)$$

Problématique (3/3)

- ▶ Modèle Averagine [Senko et al., 1994] :
une modélisation des protéines qui permet de déterminer la distribution la plus adéquate à une masse donnée :

$$\begin{aligned} \mathcal{A} : \mathbb{R}_+ \times \mathbb{N}^* &\rightarrow \mathbb{R}_+ \\ (m, z) &\rightarrow \mathbf{d}(m, z) \end{aligned} \quad (4)$$

- Rappel :

Problème de DOSY

Trouver \mathbf{x} à partir de \mathbf{y} :

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{b}$$

Modèle \mathbf{H} connu (TL)

Problème en MS

Trouver \mathbf{x} à partir de \mathbf{y} :

- ✗ Pas de modèle de mesure explicite
- ✗ Modèle d'averagine \mathcal{A} non linéaire
- ✗ Grande dimension des mesures
- ✗ Bruit intense

↪ Problème inverse complexe à résoudre

État de l'art

Méthodes proposées

- Approche Bayésienne par entropie maximale [Ferrige et al., 1991]
 - ✗ Lente convergence
- Technique du peak-picking
 - Une approche par référence
 - ✗ Non applicable avec un spectre de basse résolution
- L'algorithme NITPICK [Renard et al., 2008]
 - Débruitage avec régularisation ℓ_1
 - ✗ Technique de peak picking
- Approche NNLS [Slawski et al., 2012]
 - Débruitage sous contrainte de positivité
 - ✗ Technique de peak picking

✗ Aucune méthode n'est satisfaisante !

Approche par dictionnaire

Pour une mesure formée de P molécules ayant différentes masses monoisotopiques $m_p^{\text{iso}} \in (0, +\infty)$, de charges $z_p \in \mathbb{N}^*$ et d'abondance $a_p \in (0, +\infty)$:

$$\mathbf{y} = \sum_{p=1}^P a_p \mathbf{d}(m_p^{\text{iso}}, z_p) + \mathbf{b} \quad (5)$$

- $\mathbf{y} \in \mathbb{R}^M$: spectre MS mesuré
- $\mathbf{d}(m_p^{\text{iso}}, z_p)$: distribution de masse associée à la masse monoisotopique m_p^{iso} et la charge z_p
- \mathbf{b} : bruit

Approche par dictionnaire

Pour une mesure formée de P molécules ayant différentes masses monoisotopiques $m_p^{\text{iso}} \in (0, +\infty)$, de charges $z_p \in \mathbb{N}^*$ et d'abondance $a_p \in (0, +\infty)$:

$$\mathbf{y} = \sum_{p=1}^P a_p \mathbf{d}(m_p^{\text{iso}}, z_p) + \mathbf{b} \quad (5)$$

- $\mathbf{y} \in \mathbb{R}^M$: spectre MS mesuré
- $\mathbf{d}(m_p^{\text{iso}}, z_p)$: distribution de masse associée à la masse monoisotopique m_p^{iso} et la charge z_p
- \mathbf{b} : bruit

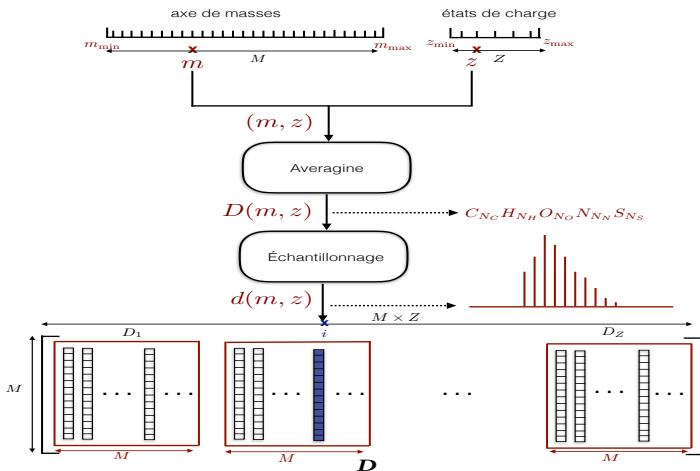
Pour M masses et Z charges, on définit une grille de taille $T = MZ$, et on construit le dictionnaire $\mathbf{D} \in \mathbb{R}^{M \times T}$:

$$\mathbf{D} = [\mathbf{D}_1, \mathbf{D}_2, \dots, \mathbf{D}_Z] \quad (6)$$

- \mathbf{D}_ℓ : Dictionnaire associé à la charge d'indice ℓ
- La i -ème colonne de \mathbf{D}_ℓ est la distribution de masse déterminée avec le modèle **averagine** à la i ème position sur la grille.

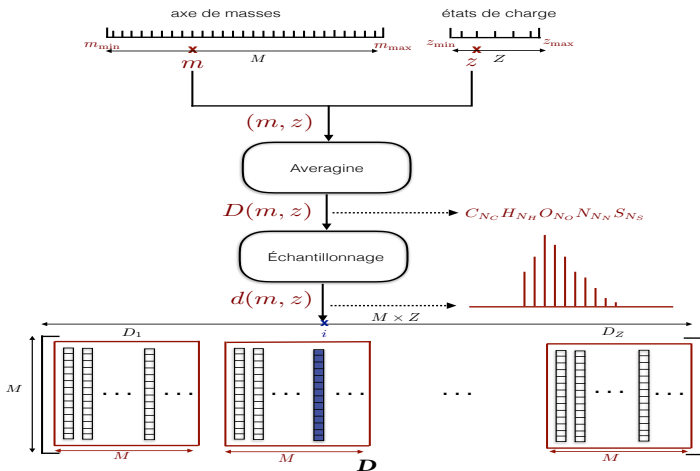
Approche par dictionnaire

► Construction du dictionnaire



Approche par dictionnaire

► Construction du dictionnaire



► Problème inverse mal-posé : $y = Dx + b$

Stratégie d'optimisation

$$\underset{\mathbf{x} \in \mathbb{R}^N}{\text{minimiser}} \quad \Psi(\mathbf{x}) \quad \text{tel que} \quad \|\mathbf{D}\mathbf{x} - \mathbf{y}\| \leq \eta \quad (5)$$

- $\Psi : \mathbb{R}^T \mapsto]-\infty, +\infty]$ fonction de régularisation utilisée pour renforcer la parcimonie et la positivité du signal cible
- $\eta > 0$ paramètre lié au niveau de bruit

Choix adoptés

- $\Psi = \ell_1$: critère de parcimonie convexe
- Algorithme Primal Dual : Aucune inversion matricielle & Convergence garantie [[Chambolle and Pock, 2011](#)]
- Langage : Python

Approximation circulante

► Difficultés

- ✗ Grande dimension du dictionnaire $T = MZ$
- ↪ complexité de calcul
- ↪ large mémoire de stockage

► Hypothèse

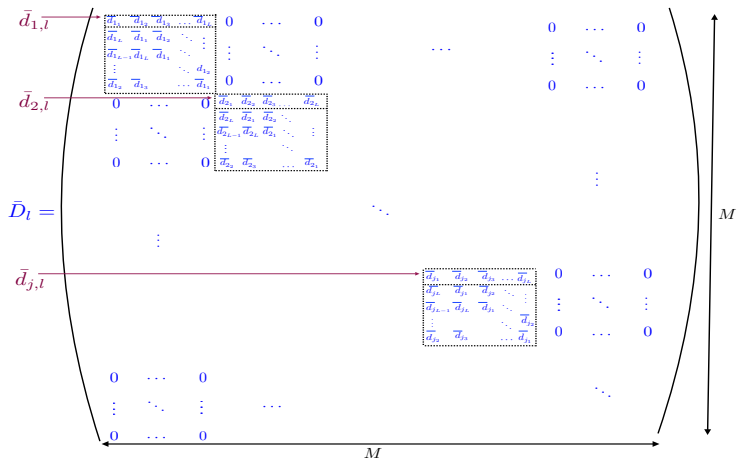
- Le profil isotopique est localement stable
- ↪ Décomposer l'axe de masse en fenêtres de taille L

► Nouveau dictionnaire

$$\bar{\mathbf{D}} = [\bar{\mathbf{D}}_1, \bar{\mathbf{D}}_2, \dots, \bar{\mathbf{D}}_Z] \quad (6)$$

$$\bar{\mathbf{D}}_\ell = \mathbf{B} \text{Diag} \left(\left[\text{Circ}(\bar{\mathbf{d}}_{s,\ell}) \right]_{1 \leq s \leq M/L} \right) \quad (7)$$

Approximation circulante

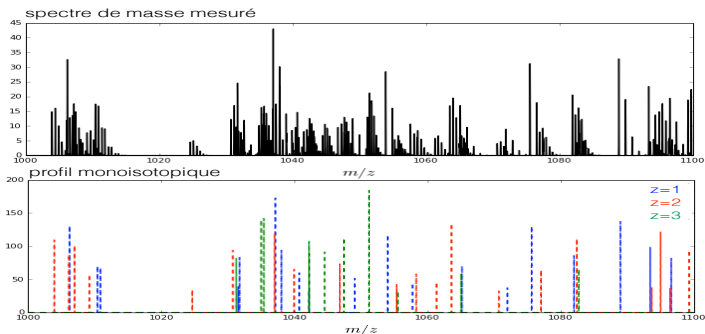


- Mémoire de stockage **réduite**
- Simplicité de calcul avec la Transformée de Fourier

Résultats de simulation (1/3)

Signal simulé A

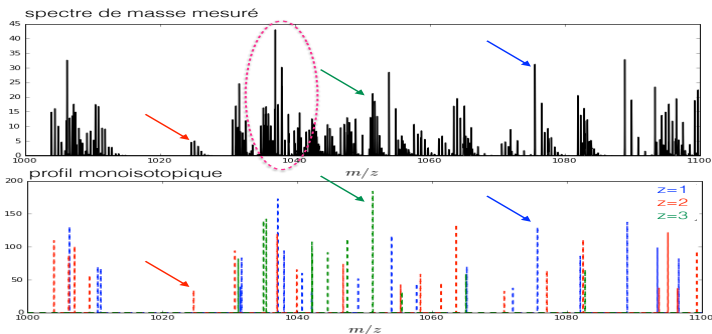
- $M = 5000$, $P = 50$ protéines
- $Z = 3$, $z_{\min} = 1$, $z_{\max} = 3$
- Bruit gaussien, $\sigma = 10^{-2}$
- Axe de masse = $[1000, 1100]$ Daltons
- **D** de dimension (5000, 15000) : 572 MB.



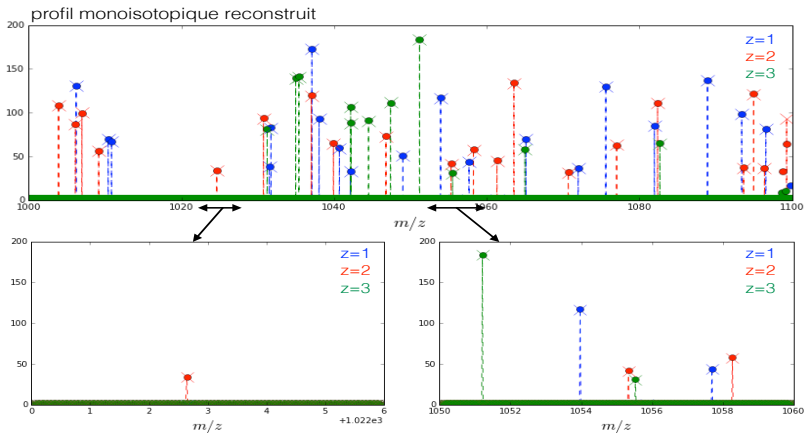
Résultats de simulation (1/3)

Signal simulé A

- $M = 5000$, $P = 50$ protéines
- $Z = 3$, $z_{\min} = 1$, $z_{\max} = 3$
- Bruit gaussien, $\sigma = 10^{-2}$
- Axe de masse = $[1000, 1100]$ Daltons
- D de dimension $(5000, 15000)$: 572 MB.



Résultats de simulation (2/3)



Résultat de reconstruction :

Signal original (--) Signal reconstruit avec \mathbf{D} (x) Signal reconstruit avec $\bar{\mathbf{D}}$ (●)

Résultats de simulation (3/3)

■ Effet du bruit ? Mémoire de stockage ?

σ	Approche par dictionnaire exact			Approximation circulante		
	RSB	Temps	Mémoire	RSB	Temps	Mémoire
1	16.18	303.33	572	15.57	127.85	0.53
0.1	35.73	206.84		35.43	44.48	
0.01	39.56	377.80		38.38	290.56	

RSB (en dB), temps de calcul (en s) et mémoire de stockage (en MB) pour la reconstruction du signal A pour différents niveaux de bruit. L'approximation circulante est testée pour $L = 10$.

- ▶ Bonne qualité de reconstruction quelque soit le niveau de bruit.
- ▶ L'approximation circulante est **plus rapide** que l'approche par dictionnaire exacte avec une **légère dégradation**, mais une **optimisation de mémoire** très importante.

Mesures réelles

■ Paramètres de mesure

- peptide EVEALEKKVAALESKVQALEKKVEALEHG-NH₂
- 3 μ M (C₁₄₀H₂₄₀N₃₈O₄₅)
- Forme trimère à partir de 50 mM de NH₄OAc
- BRUKER Solarix 15 T, instrument FT-ICR, source ESI
- $M = 8.130.981$ sur $[153.57, 4999.96]$ Daltons

■ Approche par dictionnaire exact ?

D de dimension ($M = 8.130.981, M = 8.130.981 \times z$) avec $z \geq 1$:
plusieurs TéraBytes, très large mémoire de stockage !

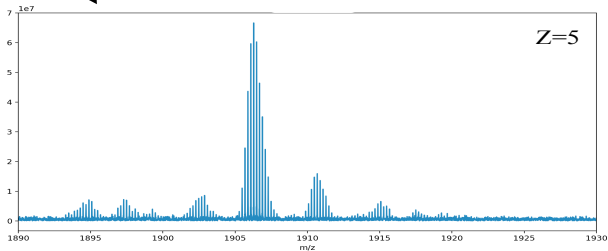
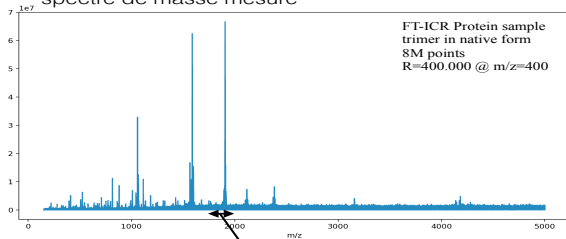
↪ Approximation circulante avec $L = 10$

■ Paramètres de l'algorithme

- σ : estimé à partir du signal mesuré
- nombre d'itérations : 2000

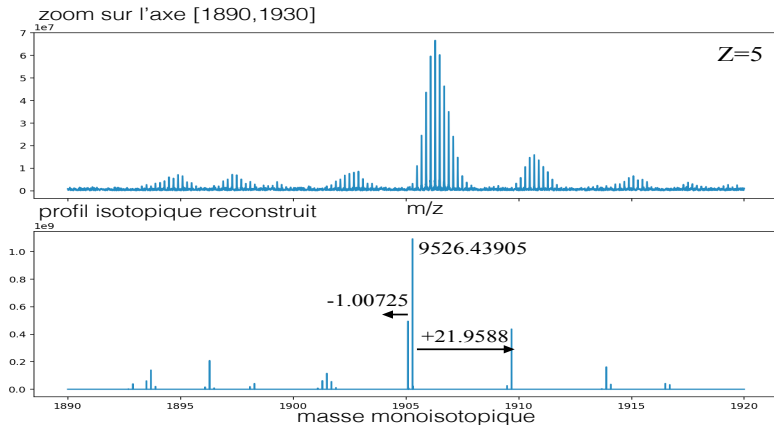
Résultats expérimentaux

spectre de masse mesuré



Spectre FT-ICR du peptide sur l'axe de masse [1890, 1930]

Résultats expérimentaux



Analyse du spectre FT-ICR du peptide (8 s avec un Mac Pro Intel Xeon)

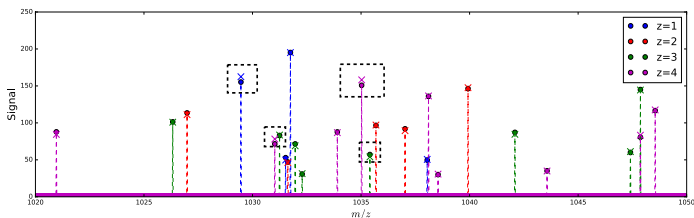
- Masse monoisotopique théorique du trimère : $m = 9526.337$ Daltons.
- Masse monoisotopique théorique de l'ajout du Sodium : $m = +21.982$ Daltons.
- Masse monoisotopique théorique de l'Hydrogène : $m = +1,007$ Daltons.

Publication 3

Afef Cherni, Emilie Chouzenoux, Marc-André Delsuc, **Fast dictionary based approach for mass spectrometry data analysis**, In Proceedings of the 43th IEEE ICASSP 2018, pp. x-x+5, Calgary-Alberta-Canada, April 2018.

- ▶ Nouvelle approche par dictionnaire qui reconstruit les pics monoisotopiques à partir d'un spectre MS.
- ▶ Nouvelle approximation du dictionnaire pour réduire la mémoire de stockage et le temps de calcul.

Observation



✗ Sous estimation d'amplitude liée à l'utilisation de l_1
 ↪ Ratio de normes l_1/l_2 [Repetti et al., 2015].

Nouvelle proposition

Fonction de régularité l_p/l_q plus générale dédiée à la restauration des signaux parcimonieux.

Nouvelle pénalité

Même formulation

$$\underset{\mathbf{x} \in \mathbb{R}_+^N}{\text{minimiser}} \quad \Psi(\mathbf{x}) \quad \text{tel que} \quad \|\mathbf{D}\mathbf{x} - \mathbf{y}\| \leq \eta \quad (6)$$

Nouvelle régularisation

$$(\forall \mathbf{x} \in \mathbb{R}^N) \quad \Psi(\mathbf{x}) = \frac{\ell_p(\mathbf{x})}{\ell_q(\mathbf{x})} \quad (7)$$

avec :

$$(\forall p \in]0, 2]) \quad \ell_p(\mathbf{x}) = \left(\sum_{n=1}^N |x_n|^p \right)^{1/p} \quad (8)$$

$$(\forall q \in [2, +\infty[) \quad \ell_q(\mathbf{x}) = \left(\sum_{n=1}^N |x_n|^q \right)^{1/q} \quad (9)$$

Nouvelle pénalité

Même formulation

$$\underset{\mathbf{x} \in \mathbb{R}_+^N}{\text{minimiser}} \quad \Psi(\mathbf{x}) \quad \text{tel que} \quad \|\mathbf{D}\mathbf{x} - \mathbf{y}\| \leq \eta \quad (6)$$

Nouvelle régularisation

$$(\forall \mathbf{x} \in \mathbb{R}^N) \quad \Psi(\mathbf{x}) = \frac{\ell_{p,\alpha}(\mathbf{x})}{\ell_{q,\eta}(\mathbf{x})} \quad (7)$$

tel que :

$$(\forall p \in]0, 2]) \quad \ell_{p,\alpha}(\mathbf{x}) = \left(\sum_{n=1}^N \left((x_n^2 + \alpha^2)^{p/2} - \alpha^p \right) \right)^{1/p} \quad (8)$$

$$(\forall q \in [2, +\infty[) \quad \ell_{q,\eta}(\mathbf{x}) = \left(\sum_{n=1}^N |x_n|^q + \eta^q \right)^{1/q} \quad (9)$$

Nouvelle pénalité

Même formulation

$$\underset{\mathbf{x} \in \mathbb{R}_+^N}{\text{minimiser}} \quad \Psi(\mathbf{x}) \quad \text{tel que} \quad \|\mathbf{D}\mathbf{x} - \mathbf{y}\| \leq \eta \quad (6)$$

Nouvelle régularisation

$$(\forall \mathbf{x} \in \mathbb{R}^N) \quad \Psi(\mathbf{x}) = \log \left(\frac{\ell_{p,\alpha}(\mathbf{x})}{\ell_{q,\eta}(\mathbf{x})} \right) \quad (7)$$

tel que :

$$(\forall p \in]0, 2]) \quad \ell_{p,\alpha}(\mathbf{x}) = \left(\sum_{n=1}^N \left((x_n^2 + \alpha^2)^{p/2} - \alpha^p \right) \right)^{1/p} \quad (8)$$

$$(\forall q \in [2, +\infty[) \quad \ell_{q,\eta}(\mathbf{x}) = \left(\sum_{n=1}^N |x_n|^q + \eta^q \right)^{1/q} \quad (9)$$

Nouvelle pénalité

Même formulation

$$\underset{\mathbf{x} \in \mathbb{R}_+^N}{\text{minimiser}} \quad \Psi(\mathbf{x}) \quad \text{tel que} \quad \|\mathbf{D}\mathbf{x} - \mathbf{y}\| \leq \eta \quad (6)$$

Nouvelle régularisation

$$(\forall \mathbf{x} \in \mathbb{R}^N) \quad \Psi(\mathbf{x}) = \log \left(\frac{(\ell_{p,\alpha}^p(\mathbf{x}) + \beta^p)^{1/p}}{\ell_{q,\eta}(\mathbf{x})} \right) \quad (7)$$

Smoothed p -Over- q norm (SPOQ)

$$(\forall p \in]0, 2]) \quad \ell_{p,\alpha}(\mathbf{x}) = \left(\sum_{n=1}^N \left((x_n^2 + \alpha^2)^{p/2} - \alpha^p \right) \right)^{1/p} \quad (8)$$

$$(\forall q \in [2, +\infty[) \quad \ell_{q,\eta}(\mathbf{x}) = \left(\sum_{n=1}^N |x_n|^q + \eta^q \right)^{1/q} \quad (9)$$

Implémentation

Principes

- ▶ Algorithme Forward-Backward à métrique variable.
- ▶ Ajustement local de la métrique, basé sur la stratégie de majoration-minimisation de Ψ dans une région de confiance.
- ▶ Garanties de convergence

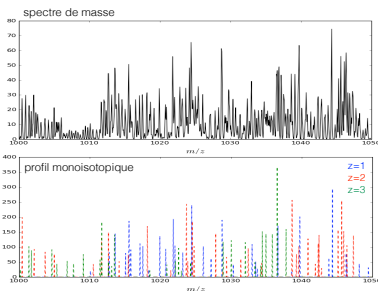
● Signal B :

$M = 1000$

$Z = 3$

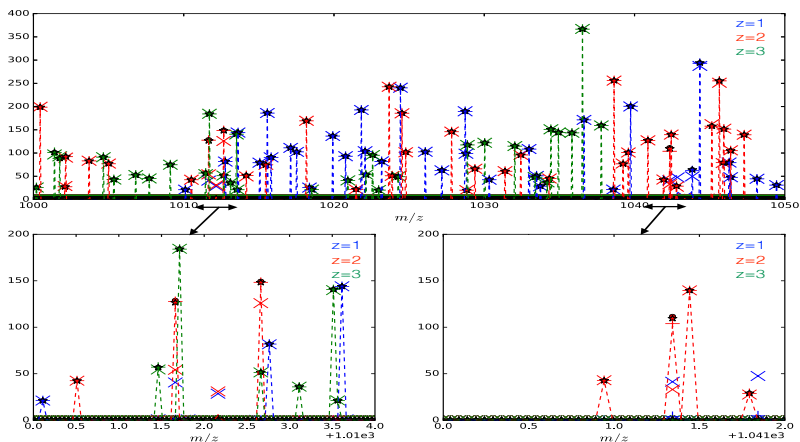
$P = 100$ protéines

axe de masse = $[1000, 1050]$



Résultats synthétiques

- Résultat de reconstruction du signal B avec la pénalité :
 l_1 en " × ", l_1/l_2 en " + ", $l_{1/4}/l_2$ en " * " et $l_{3/4}/l_3$ en " ● ".
 Le signal original est en " – "



Publication (en cours de finalisation)

Afef Cherni, Emilie Chouzenoux, Laurent Duval, Marc-André Delsuc, Jean-Christophe Pesquet, **SPOQ : Smoothed p -Over- q ℓ_p/ℓ_q norm ratio Regularization for Sparse Signal reconstruction.**

- ▶ Nouvelle régularisation SPOQ fondée sur une approximation lisse de ℓ_p/ℓ_q .
- ▶ Nouvel algorithme à métrique variable localement ajustée basé sur l'algorithme FB.
- ▶ Qualité de reconstruction des signaux parcimonieux en MS bien meilleure que le critère ℓ_1 .
- ✗ Complexité de calcul.

Conclusion 2

- ✓ Nouvelle approche par dictionnaire pour restaurer les signaux MS avec le critère ℓ_1 .
- ✓ Nouvelle approximation du dictionnaire qui réduit la complexité de calcul et les sources de stockage.
- ✓ Analyse des données MS de grande dimension
[Cherni et al., 2018].

- ✓ Nouvelle approche SPOQ qui assure l'estimation précise des signaux parcimonieux
[Article en cours].

Conclusion générale (1/2)

▶ RMN DOSY

- ✓ Nouvelle approche de régularisation hybride
- ✓ Algorithme PALMA + Serveur web
- ↪ 2 publications, et un serveur web

Conclusion générale (1/2)

▶ RMN DOSY

- ✓ Nouvelle approche de régularisation hybride
- ✓ Algorithme PALMA + Serveur web
- ~ 2 publications, et un serveur web

▶ Spectrométrie de masse

- ✓ Nouvelle approche par dictionnaire
- ✓ Nouvelle approximation circulante
- ~ 1 publication
- ✓ Nouvelle approche SPOQ
- ~ 1 publication en cours de finalisation.

Conclusion générale (2/2)

► Autres contributions :

✓ Nouvel opérateur proximal pour la régression logistique parcimonieuse

↪ 1 publication.

✓ Nouvel algorithme proximal alterné pour l'analyse spectrale des signaux sous échantillonnés

↪ 1 publication en cours de finalisation.

Perspectives

▶ RMN DOSY

- Optimisation non convexe (l_0)
- Déconvolution aveugle (nouvelle version de PALMA)

Perspectives

▶ RMN DOSY

- Optimisation non convexe (l_0)
- Déconvolution aveugle (nouvelle version de PALMA)

▶ Spectrométrie de masse

- Mieux adapter l'approche par dictionnaire au cas réel
- Estimer le modèle de convolution à partir des mesures
- Apprentissage par réseau de neurones des données MS
- Chercher le bon choix théorique de (p, q) pour l'approche SPOQ

Publications I

■ Articles acceptés

- **Afef Cherni**, Emilie Chouzenoux, Marc-André Delsuc, **Fast dictionary based approach for mass spectrometry data analysis**, In Proceedings of the 43th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2018), pp. x-x+5, Calgary-Alberta-Canada, 14-20 april 2018.
- Giovanni Chierchia, **Afef Cherni**, Emilie Chouzenoux, and Jean- Christophe Pesquet, **Approche de Douglas-Rachford aléatoire par blocs. Application à la régression logistique parcimonieuse**, In Actes du 26e colloque GRETSI (GRETSI 2017), Juan-les-Pins, 5-8 september 2017
- **Afef Cherni**, Emilie Chouzenoux, Marc-André Delsuc, **Improved algorithm for DOSY signal processing**, Analyst, Vol. 142, No. 5, pages 772-779, 2016.
- **Afef Cherni**, Emilie Chouzenoux, Marc-André Delsuc, **Proximity for a class of hybrid sparsity + entropy prior. Application to DOSY NMR signal reconstruction**, In Proceedings of the International Symposium on Signal, Image, Video and Communications (ISIVC 2016), Tunis-Tunisia, 21-23 november 2016.

■ Articles en cours de préparation

- **Afef Cherni**, Emilie Chouzenoux, Laurent Duval, Jean-Christophe Pesquet, **SPOQ : Smoothed p -Over- q ℓ_p/ℓ_q norm ratio Regularization for Sparse Signal reconstruction**.
- Lionel Chiron, **Afef Cherni**, Jean-Philippe Starck, Christian Rolando, Emilie Chouzenoux, Marc-André Delsuc, **Fast and robust analysis of non uniformly sampled 2D-FT-ICR-MS datasets**.

■ Logiciel diffusé :

- PALMA : Plateforme en Python, basée sur l'algorithme proximal PALMA, proposée pour l'analyse des données expérimentales réelles DOSY. Juin 2016. License Cecill-B <http://palma.labo.igbmc.fr/>

Remerciement

- Marc-André Delsuc, Émilie Chouzenoux
- Christian Rolando, Bruno Kieffer
- Jean-Christophe Pesquet, Laurent Duval, Giovanni Chierchia
- Équipe RMN de l'IGBMC, Strasbourg
- Équipe Signal et Communications de LIGM, Paris



Remerciement

- Marc-André Delsuc, Émilie Chouzenoux
- Christian Rolando, Bruno Kieffer
- Jean-Christophe Pesquet, Laurent Duval, Giovanni Chierchia
- Équipe RMN de l'IGBMC, Strasbourg
- Équipe Signal et Communications de LIGM, Paris



• **Merci pour votre attention** •

Méthodes modernes d'analyse de données en biophysique analytique : Résolution des problèmes inverses en RMN DOSY et MS

Afef Cherni

Directeur de thèse : Marc-André Delsuc

Co-encadrante : Émilie Chouzenoux

Strasbourg, le 20 septembre 2018

