

# Analyse comparative d’algorithmes de restauration en architecture dépliée pour des signaux chromatographiques parcimonieux

Mouna GHARBI<sup>1</sup> Silvia VILLA<sup>2</sup> Emilie CHOUZENOUX<sup>3</sup> Jean-Christophe PESQUET<sup>3</sup> Laurent DUVAL<sup>4</sup>

<sup>1</sup>University of Graz, Department of Mathematics and Scientific Computing, Graz, Austria

<sup>2</sup>MaLGa, DIMA, Università degli Studi di Genova, Via Dodecaneso 35, 16146 Genova, Italy

<sup>3</sup>CVN, CentraleSupélec, Inria Saclay, University Paris Saclay, France

<sup>4</sup>IFP Energies nouvelles, Rueil-Malmaison, France

**Résumé** – La restauration de données à partir d’observations dégradées, sous hypothèse de parcimonie, est un champ d’étude actif. Aux traditionnelles méthodes itératives d’optimisation s’ajoutent aujourd’hui l’usage de techniques d’apprentissage profond. Le développement des méthodes dépliées bénéficie des avantages de ces deux familles. Nous réalisons une étude comparative de trois architectures sur des bases de données paramétrées de signaux en chromatographie, mettant en évidence la performance de ces approches, notamment en employant des métriques adaptées à la caractérisation de signaux de pics physico-chimiques.

**Abstract** – Data restoration from degraded observations, of sparsity hypotheses, is an active field of study. Traditional iterative optimization methods are now complemented by deep learning techniques. The development of unfolded methods benefits from both families. We carry out a comparative study of three architectures on parameterized chromatographic signal databases, highlighting the performance of these approaches, especially when employing metrics adapted to physico-chemical peak signal characterization.

## 1 Introduction

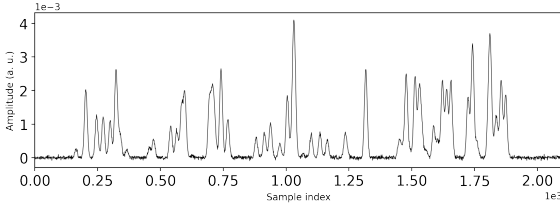


FIGURE 1 : Observation bruitée  $\mathbf{z}$  d’une somme de pics  $\mathbf{p}$ .

Ce travail<sup>1</sup> s’intéresse à une classe de signaux parcimonieux communément décrits comme formés d’une somme de pics relativement étroits ( $\mathbf{p}$ ), potentiellement dégradés par un noyau de lissage et du bruit ( $\mathbf{e}$ ), résultant en une observation  $\mathbf{z}$ . Ce modèle est courant en chimie analytique, notamment pour le traitement de données de type chromatographique, représentées dans la figure 1. On considère connue une forme paramétrique  $\pi$  de pic, illustré en figure 2. Le problème inverse à résoudre est l’estimation d’un train d’impulsions parcimonieux  $\mathbf{s}$  (impulsion représentée par une barre noire verticale, figure 2). La prise en compte de la réponse globale du système d’acquisition se fait par un opérateur linéaire  $\mathbf{G}$ , par exemple un lissage ou opérateur de flou (14), une transformée de Fourier (15) ou de Laplace (6). Il en résulte :

$$\mathbf{z} = \mathbf{G}\mathbf{p} + \mathbf{e} = \mathbf{G}(\pi * \mathbf{s}) + \mathbf{e}. \quad (1)$$

<sup>1</sup>Projets supports : EU H2020-MSCA-ETN TraDE-OPT-861137; ERC Starting Grant MAJORIS ERC-2019-STG850925; AFOSR FA8655-22-1-7034; ERC Consolidator Grant SLING 819789; EU H2020-MSCA-RISE NoMADS-777826.

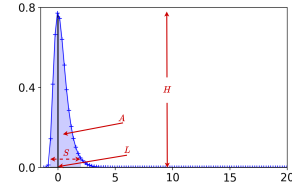


FIGURE 2 : Caractérisation d’un pic : hauteur ( $H$ ), aire ( $A$ ), localisation ( $L$ ). Support ( $S$ ) et aire colorée (bleu) déterminés en fonction d’un seuil sur  $H$  (ici  $H/20$ ).

La résolution du problème inverse consiste en l’obtention d’une estimation de  $\mathbf{p}$  (ainsi que de  $\mathbf{s}$  si possible), connaissant  $\mathbf{z}$  et  $\mathbf{G}$ , avec des hypothèses sur le bruit  $\mathbf{e}$ , et bien entendu en promouvant des solutions parcimonieuses. Pour que les résultats restent interprétables et utiles pour les chimistes, il est important de considérer des indicateurs de qualité complémentaires aux quantités usuelles. Dans le modèle paramétrique considéré, l’estimation des facteurs de forme (notamment : hauteur ( $H$ ), aire ( $A$ ) et localisation ( $L$ )) pour chaque pic est également importante, car ces derniers permettent d’accéder à des informations sur la nature chimique du mélange (type de composé, concentration de chaque espèce du mélange, etc.).

On distingue des approches (i) basées sur des modèles, employant des algorithmes d’optimisation itératifs et (ii) d’apprentissage (profond), basées sur des architectures neuronales. Chacune a des lacunes. Un paradigme récent, dit dépliement ou déroulement profond (*deep unfolding/unrolling*) (11), tente de les combler. Ce travail prolonge (10), et s’intéresse :

- au développement d’un simulateur réaliste de bases de données, aux paramètres ajustables (section 2) ;

- à l'analyse comparative de trois architectures dépliées basées sur des modèles, pour la résolution de problèmes inverses en chimie analytique et à leur formulation (sections 3 et 4);
- à leur évaluation, incluant des métriques HAL (hauteur-aire-localisation) inspirées de la chimie, plus adaptées à cette classe de signaux parcimonieux (section 5).

Les codes proposés sont disponibles sous forme d'une boîte à outils écrite en Pytorch pour satisfaire à des exigences de reproductibilité.

## 2 Modèle de données et évaluation

L'équation 1 nous conduit à simuler un modèle additif basé sur la forme de pic asymétrique et paramétrique à noyau de Fraser-Suzuki (8, p. 104) :  $\forall x > m - \frac{\sigma_f}{a}$ ,

$$\pi(x) \propto \exp \left( -\frac{1}{2a^2} \log \left( 1 + a \frac{(x-m)}{\sigma_f} \right)^2 \right). \quad (2)$$

Des pics gaussiens s'obtiennent quand  $a \rightarrow 0$ . La discrétisation des signaux induit les paramètres suivants :

- signal impulsif  $\mathbf{s}$  : nombre d'échantillons  $N = 2000$ ; nombre d'impulsions distinctes  $P$ , exprimé relativement à la taille du signal par  $P/N$ ; distance minimale séparant deux pics  $d_{\min}$ ; intensités des pics tirées aléatoirement (valeur absolue de la réalisation d'une loi normale de moyenne 0 et de variance 1);
- noyau de Fraser-Suzuki  $\pi$  : largeur de pic  $\sigma_f > 0$ ; coefficient d'asymétrie  $a > 0$ ;
- sources d'altération : bruit additif gaussien de moyenne nulle et d'écart-type  $\sigma_e > 0$ ; opérateur de flou gaussien de largeur  $\sigma_G = 1$  (fixé pour cette étude).

Pour tester l'influence des paramètres ( $N, P, d_{\min}, \sigma_f, a, \sigma_e$ ), nous proposons sept bases de données (table 1) dont la complexité relative est résumée en table 2.

Notre objectif est ensuite de comparer la performance des différents algorithmes. Cependant, il est connu que les métriques "quadratiques" standard se prêtent mal aux signaux parcimonieux, comportant une information entre-pics "inutile". Nous allons supposer que chaque pic possède (dans la vérité-terrain) un support effectif  $\mathcal{S}_j$  connu, sur lequel ses caractéristiques HAL (hauteur-aire-localisation) peuvent être évaluées et comparées. Étant donné le signal original  $\mathbf{s} * \pi = \mathbf{p} \in \mathbb{R}^N$  et sa restauration  $\hat{\mathbf{p}} \in \mathbb{R}^N$ , nous calculons dans un premier temps trois métriques usuelles : erreur quadratique moyenne (MSE), rapport signal/bruit (SNR, exprimé en décibels) et sa version tronquée (TSNR) sur le support. Cette dernière correspond à un rapport signal/bruit évalué uniquement sur l'union des supports de tous les pics, c'est-à-dire uniquement sur le signal d'intérêt physico-chimique,  $\mathcal{S} = \bigcup_{1 \leq j \leq P} \mathcal{S}_j$  :

$$\text{MSE}(\mathbf{p}, \hat{\mathbf{p}}) = \frac{1}{n} \|\mathbf{p} - \hat{\mathbf{p}}\|^2, \quad (3)$$

$$\text{SNR}(\mathbf{p}, \hat{\mathbf{p}}) = 20 \log_{10} \left( \frac{\|\mathbf{p}\|}{\|\mathbf{p} - \hat{\mathbf{p}}\|} \right), \quad (4)$$

$$\text{TSNR}(\mathbf{p}, \hat{\mathbf{p}}) = 20 \log_{10} \left( \frac{\sum_i |\mathbf{p}^{(i)}|^2}{\sum_i |\mathbf{p}^{(i)} - \hat{\mathbf{p}}^{(i)}|^2} \right), \forall i \in \mathcal{S}. \quad (5)$$

Si l'introduction du TSNR améliore l'évaluation des performances pour des signaux parcimonieux, cette mesure est complétée par des indicateurs de morphologie des pics, utiles en chimie. Nous employons trois quantités pour chaque pic : hauteur ( $H$ ), aire sous le pic ( $A$ ), localisation ( $L$ ), à l'aide de l'information sur le support essentiel  $\mathcal{S}_j$ . Elles sont illustrées sur leur figure 2. Les grandeurs  $H$  et  $A$  peuvent être reliées à la prévalence d'une entité chimique donnée, tandis que  $L$  peut servir à la caractérisation ou la détection. Les valeurs vraies issues du modèle de pic  $\bar{H}_j$  et  $\bar{L}_j$  sont obtenues directement pour chaque pic convolué. Nous faisons l'hypothèse que le support essentiel  $\mathcal{S}_j$  est défini par les indices auxquels l'amplitude du signal est supérieure à une certaine fraction ( $0 < \vartheta < 1$ ) de la hauteur  $\bar{H}_j$ . Ici, on prend  $\vartheta = \frac{1}{20}$ . L'aire  $\bar{A}_j$  est calculée par intégration discrète (méthode des trapèzes) des échantillons supportés par  $\mathcal{S}_j$ . En utilisant cette information comme oracle, des quantités analogues sont obtenues sur le signal restauré  $\hat{\mathbf{p}}$  :  $\hat{L}_j, \hat{H}_j$  et  $\hat{A}_j$ .

Ces indicateurs sont résumés dans des vecteurs  $\bar{\mathbf{H}} = (\bar{H}_j)_{1 \leq j \leq P}$  et  $\bar{\mathbf{L}} = (\bar{L}_j)_{1 \leq j \leq P}$  de longueur  $P$ , en définissant de manière analogue  $\bar{\mathbf{A}}, \bar{\mathbf{L}}, \hat{\mathbf{A}}$  et  $\hat{\mathbf{L}}$ . Ils permettent d'obtenir des graphes de dispersion comme en figure 3. Leur écart plus ou moins prononcé à la linéarité permettant de comparer subjectivement les performances. Cette analyse peut être affinée en distinguant les pics suivant leur taux de recouvrement. En cyan et jaune sont représentés respectivement les pics présentant moins (plus) de 30% de recouvrement. On voit que les points jaunes sont souvent situés au-dessus de la diagonale. Ceci est probablement lié à l'emprunt d'information entre deux pics adjacents, mal séparés, provoquant une surestimation des hauteurs. Ces observations motivent la définition d'une mesure objective normalisée d'écart des hauteurs :

$$\text{NMAE}(\bar{\mathbf{H}}, \hat{\mathbf{H}}) = \frac{\sum_{j=1}^P |\bar{H}_j - \hat{H}_j|}{\sum_{j=1}^P |\bar{H}_j|}. \quad (6)$$

Des mesures similaires sont définies pour l'aire  $\text{NMAE}(\bar{\mathbf{A}}, \hat{\mathbf{A}})$  et la localisation  $\text{NMAE}(\bar{\mathbf{L}}, \hat{\mathbf{L}})$ .

## 3 Travaux analogues

Nous mentionnons ici quelques travaux en lien avec cette étude, avec des applications en chimie analytique, en commençant par les méthodes d'optimisation itérative. Le débruitage de spectres Raman emploie une régularisation en variation totale (20). Une approche bayésienne de débruitage et déconvolution d'images spectroscopiques en infrarouge à transformée de Fourier est proposée dans (17). La reconstruction de données spectroscopiques est obtenue en incorporant des contraintes de positivité, par régularisation en maximum d'entropie (4) ou par méthode de point intérieur primal-dual dans (5), plus

TABLE 1 : Bases de données : parcimonie relative ( $P/N$ ), distance minimale entre impulsions ( $d_{\min}$ ), largeur de  $\pi$  ( $\sigma_f$ ) et facteur d'asymétrie ( $a$ ), écart-type du bruit ( $\sigma_e$ ).

| Param.     | D0    | D1   | D2    | D3    | D4    | D5   | D6   |
|------------|-------|------|-------|-------|-------|------|------|
| $P/N$      | 1.5 % | 3 %  | 4.5 % | 1.5 % | 1.5 % | 3 %  | 3 %  |
| $d_{\min}$ | 5     | 3    | 1     | 5     | 5     | 3    | 3    |
| $\sigma_f$ | 0.5   | 0.5  | 0.5   | 0.5   | 0.5   | 0.5  | 0.5  |
| $a$        | 0.2   | 0.2  | 0.2   | 0.4   | 0.6   | 0.2  | 0.2  |
| $\sigma_e$ | 0.02  | 0.02 | 0.02  | 0.02  | 0.02  | 0.04 | 0.06 |

TABLE 2 : Variations paramétriques (cf. table 1) sur la parcimonie, l'asymétrie et le bruit, avec trois niveaux de difficulté.

| Difficulté                     | Faible        | Moyenne     | Haute         |
|--------------------------------|---------------|-------------|---------------|
| Parcimonie ( $P/N, d_{\min}$ ) | D0 (1.5 %, 5) | D1 (3 %, 3) | D2 (4.5 %, 1) |
| Asymétrie ( $a$ )              | D0 (0.2)      | D3 (0.4)    | D4 (0.6)      |
| Bruit ( $\sigma_e$ )           | D1 (0.02)     | D5 (0.04)   | D6 (0.06)     |

récemment dans (3) par méthode de majoration-minimisation. La restauration de données chromatographiques par moindres carrés alternés est proposée dans (19). Les méthodes d'apprentissage profond jouent un rôle croissant en chimie, avec des revues génériques (16; 2; 7). En chromatographie, notons la classification de profils d'élution ou la détection automatisée de pics (18; 12). Le recours à des algorithmes dépliés en chimie n'est pas encore très développé. Nous mentionnons notamment (1; 21; 13; 9).

## 4 Algorithmes

Nous considérons trois architectures dépliées, basées sur les algorithmes itératifs : primal-dual proximal, ISTA et semi-quadratique (HQ). Nous rappelons uniquement les problèmes de minimisation, et renvoyons par manque d'espace à (10) pour les détails :

- primal-dual déplié (U-PD) :  $\rho > 0$  est un hyper-paramètre lié au niveau de bruit.

$$\hat{x} \in \arg \min_{x \in \mathbb{R}^n} \|x\|_1 \quad \text{s.t.} \quad \|Hx - z\|_2 \leq \rho. \quad (7)$$

- seuillage doux itératif (ISTA) déplié (U-ISTA) :  $\chi$  est un hyper-paramètre couplant les deux termes.

$$\hat{x} \in \arg \min_{x \in \mathbb{R}^n} \left( F(x) = \frac{1}{2} \|Hx - z\|_2^2 + \chi \|x\|_1 \right). \quad (8)$$

- Semi-quadratique déplié (U-HQ) :

$$\hat{x} \in \arg \min_{x \in \mathbb{R}^n} \left( F(x) = \frac{1}{2} \|Hx - z\|_2^2 + \Psi(x) \right), \quad (9)$$

avec  $\Psi(x) = \sum_{i=1}^n \psi(x_i)$  et

$$(\forall t \in \mathbb{R}) \quad \psi(t) = \lambda_1 \delta_1 \left( |t| - \delta_1 \log \left( \frac{|t|}{\delta_1} + 1 \right) \right) + \lambda_2 \frac{\delta_2^2}{2} \log \left( 1 + \frac{t^2}{\delta_2^2} \right). \quad (10)$$

## 5 Résultats

Nous comparons ces trois algorithmes dépliés pour un nombre de couches fixé à  $K = 4$ . La table 3 rassemble les métriques

MSE, SNR, TSNR et HAL. **U-HQ** présente de manière évidente des résultats supérieurs à ceux de **U-PD** et **U-ISTA**. Notons que cette performance se répercute en coût de calcul, le temps moyen d'exécution par jeu de données étant respectivement de 0,09, 0,009 et 0,003 secondes.

L'intérêt de la constitution de bases de données paramétrables est d'analyser plus précisément l'impact d'un facteur en particulier. Si l'on regarde la parcimonie (ligne 1, table 2), on observe pour les jeux de données D0, D1 et D2 que l'accroissement du nombre d'impulsions (d'amplitude non nulle) dégrade la qualité de reconstruction. Pour les trois méthodes, la mesure NMAE augmente pour **H**, **A** et **L**. Les méthodes dépliées ont des difficultés croissantes à estimer correctement hauteur, amplitude et localisation, à mesure que les signaux sont plus denses en pics (moins parcimonieux). On remarque également que les métriques globales SNR et TSNR ne sont pas des indicateurs fiables pour ce phénomène, car leurs valeurs sont assez similaires pour les trois ensembles de données.

La figure 3 détaille l'évolution des hauteurs prédites. La comparaison par colonnes confirme les bonnes performances de **U-HQ**, et indique que **U-ISTA** et **U-PD** souffrent apparemment du biais  $\ell_1$  typique, avec une tendance à sous-estimer une portion importante des hauteurs de pics.

## 6 Conclusions

Nous avons proposé une analyse comparative de trois algorithmes de restauration en architecture dépliée : primal-dual, ISTA (seuillage doux itératif) et semi-quadratique. L'application choisie est le traitement de signaux chromatographiques parcimonieux. Cette analyse se base sur la réalisation d'une boîte à outils permettant de générer différents ensembles de données chromatographiques paramétrées. Les résultats expérimentaux soulignent l'aspect pratique des architectures déroulées face à des données de complexité graduelle, leur robustesse, leur flexibilité et leur efficacité. Il s'agit d'un domaine prometteur puisqu'il peut permettre aux chimistes de traiter leurs spectres obtenus rapidement et efficacement à l'aide d'outils d'optimisation sophistiqués. Les codes Python sont disponibles<sup>2</sup>.

<sup>2</sup><https://github.com/GHARBIMouna/Unrolled-Half-Quadratic/>

TABLE 3 : Comparaison des méthodes dépliées **U-HQ**, **U-ISTA**, **U-PD** à  $K = 4$  couches : moyenne (écart-type) des MSE, SNR et TSNR entre  $\mathbf{p}$  et  $\hat{\mathbf{p}}$  ; moyenne (écart-type) des NMAE pour les hauteurs ( $\bar{\mathbf{H}}, \hat{\mathbf{H}}$ ), aires ( $\bar{\mathbf{A}}, \hat{\mathbf{A}}$ ) et localisations ( $\bar{\mathbf{L}}, \hat{\mathbf{L}}$ ). Meilleures performances en gras.

|    |        | MSE                                                                        | SNR                 | TSNR                | NMAE( $\bar{\mathbf{H}}, \hat{\mathbf{H}}$ ) | NMAE( $\bar{\mathbf{A}}, \hat{\mathbf{A}}$ ) | NMAE( $\bar{\mathbf{L}}, \hat{\mathbf{L}}$ )                               |
|----|--------|----------------------------------------------------------------------------|---------------------|---------------------|----------------------------------------------|----------------------------------------------|----------------------------------------------------------------------------|
| D0 | U-HQ   | <b><math>4.70 \times 10^{-4}</math> (<math>1.15 \times 10^{-4}</math>)</b> | <b>19.60 (0.95)</b> | <b>19.99 (0.98)</b> | <b>0.09 (0.04)</b>                           | <b>0.012 (0.003)</b>                         | <b><math>3.12 \times 10^{-6}</math> (<math>1.45 \times 10^{-6}</math>)</b> |
|    | U-PD   | $3.68 \times 10^{-3}$ ( $1.60 \times 10^{-3}$ )                            | 10.88 (1.30)        | 10.91 (1.32)        | 0.23 (0.05)                                  | 0.107 (0.025)                                | $8.04 \times 10^{-6}$ ( $1.77 \times 10^{-6}$ )                            |
|    | U-ISTA | $2.99 \times 10^{-3}$ ( $1.03 \times 10^{-3}$ )                            | 11.68 (0.99)        | 11.71 (0.99)        | 0.23 (0.04)                                  | 0.110 (0.020)                                | $8.02 \times 10^{-6}$ ( $1.70 \times 10^{-6}$ )                            |
| D1 | U-HQ   | <b><math>1.06 \times 10^{-3}</math> (<math>2.28 \times 10^{-4}</math>)</b> | <b>19.36 (0.68)</b> | <b>19.60 (0.71)</b> | <b>0.22 (0.05)</b>                           | <b>0.014 (0.002)</b>                         | <b><math>3.81 \times 10^{-6}</math> (<math>9.46 \times 10^{-7}</math>)</b> |
|    | U-PD   | $9.97 \times 10^{-3}$ ( $2.51 \times 10^{-3}$ )                            | 9.67 (0.71)         | 9.70 (0.71)         | 0.36 (0.04)                                  | 0.107 (0.016)                                | $6.33 \times 10^{-6}$ ( $8.69 \times 10^{-7}$ )                            |
|    | U-ISTA | $8.02 \times 10^{-3}$ ( $2.06 \times 10^{-3}$ )                            | 10.61 (0.73)        | 10.63 (0.73)        | 0.35 (0.04)                                  | 0.109 (0.019)                                | $6.12 \times 10^{-6}$ ( $8.59 \times 10^{-7}$ )                            |
| D2 | U-HQ   | <b><math>1.69 \times 10^{-3}</math> (<math>3.37 \times 10^{-4}</math>)</b> | <b>19.56 (0.76)</b> | <b>19.74 (0.78)</b> | <b>0.45 (0.08)</b>                           | <b>0.015 (0.002)</b>                         | <b><math>5.17 \times 10^{-6}</math> (<math>9.55 \times 10^{-7}</math>)</b> |
|    | U-PD   | $1.72 \times 10^{-2}$ ( $3.25 \times 10^{-3}$ )                            | 9.46 (0.54)         | 9.49 (0.54)         | 0.50 (0.06)                                  | 0.101 (0.012)                                | $5.83 \times 10^{-6}$ ( $7.23 \times 10^{-7}$ )                            |
|    | U-ISTA | $1.38 \times 10^{-2}$ ( $2.60 \times 10^{-3}$ )                            | 10.43 (0.55)        | 10.44 (0.55)        | 0.50 (0.06)                                  | 0.102 (0.012)                                | $5.77 \times 10^{-6}$ ( $7.51 \times 10^{-7}$ )                            |
| D3 | U-HQ   | <b><math>8.74 \times 10^{-4}</math> (<math>2.13 \times 10^{-4}</math>)</b> | <b>16.92 (0.74)</b> | <b>17.30 (0.77)</b> | <b>0.10 (0.05)</b>                           | <b>0.015 (0.002)</b>                         | <b><math>3.68 \times 10^{-6}</math> (<math>1.74 \times 10^{-6}</math>)</b> |
|    | U-PD   | $4.35 \times 10^{-3}$ ( $1.37 \times 10^{-3}$ )                            | 10.04 (0.85)        | 10.08 (0.86)        | 0.26 (0.05)                                  | 0.107 (0.020)                                | $9.02 \times 10^{-6}$ ( $1.74 \times 10^{-6}$ )                            |
|    | U-ISTA | $3.70 \times 10^{-3}$ ( $1.10 \times 10^{-3}$ )                            | 10.71 (0.65)        | 10.75 (0.66)        | 0.25 (0.04)                                  | 0.104 (0.027)                                | $8.73 \times 10^{-6}$ ( $1.72 \times 10^{-6}$ )                            |
| D4 | U-HQ   | <b><math>1.36 \times 10^{-3}</math> (<math>2.84 \times 10^{-4}</math>)</b> | <b>14.89 (0.57)</b> | <b>15.11 (0.57)</b> | <b>0.13 (0.05)</b>                           | <b>0.012 (0.002)</b>                         | <b><math>4.49 \times 10^{-6}</math> (<math>1.88 \times 10^{-6}</math>)</b> |
|    | U-PD   | $5.01 \times 10^{-3}$ ( $1.36 \times 10^{-3}$ )                            | 9.28 (0.51)         | 9.32 (0.52)         | 0.27 (0.04)                                  | 0.10 (0.02)                                  | $9.56 \times 10^{-6}$ ( $1.65 \times 10^{-6}$ )                            |
|    | U-ISTA | $4.48 \times 10^{-3}$ ( $1.19 \times 10^{-3}$ )                            | 9.76 (0.55)         | 9.83 (0.55)         | 0.27 (0.04)                                  | 0.083 (0.011)                                | $9.26 \times 10^{-6}$ ( $1.64 \times 10^{-6}$ )                            |
| D5 | U-HQ   | <b><math>1.51 \times 10^{-3}</math> (<math>3.28 \times 10^{-4}</math>)</b> | <b>17.77 (0.75)</b> | <b>18.02 (0.78)</b> | <b>0.24 (0.06)</b>                           | <b>0.018 (0.003)</b>                         | <b><math>4.19 \times 10^{-6}</math> (<math>1.05 \times 10^{-6}</math>)</b> |
|    | U-PD   | $9.96 \times 10^{-3}$ ( $2.34 \times 10^{-3}$ )                            | 9.60 (0.65)         | 9.64 (0.66)         | 0.37 (0.05)                                  | 0.108 (0.015)                                | $6.47 \times 10^{-6}$ ( $9.10 \times 10^{-7}$ )                            |
|    | U-ISTA | $7.88 \times 10^{-3}$ ( $1.76 \times 10^{-3}$ )                            | 10.61 (0.62)        | 10.62 (0.63)        | 0.36 (0.05)                                  | 0.106 (0.014)                                | $6.30 \times 10^{-6}$ ( $9.10 \times 10^{-7}$ )                            |
| D6 | U-HQ   | <b><math>1.90 \times 10^{-3}</math> (<math>4.00 \times 10^{-4}</math>)</b> | <b>16.79 (0.76)</b> | <b>17.11 (0.80)</b> | <b>0.24 (0.05)</b>                           | <b>0.021 (0.003)</b>                         | <b><math>4.18 \times 10^{-6}</math> (<math>9.78 \times 10^{-7}</math>)</b> |
|    | U-PD   | $9.92 \times 10^{-3}$ ( $2.32 \times 10^{-3}$ )                            | 9.65 (0.66)         | 9.68 (0.66)         | 0.36 (0.04)                                  | 0.108 (0.016)                                | $6.27 \times 10^{-6}$ ( $8.70 \times 10^{-7}$ )                            |
|    | U-ISTA | $7.85 \times 10^{-3}$ ( $1.71 \times 10^{-3}$ )                            | 10.65 (0.62)        | 10.67 (0.62)        | 0.35 (0.04)                                  | 0.107 (0.014)                                | $6.10 \times 10^{-6}$ ( $8.59 \times 10^{-7}$ )                            |

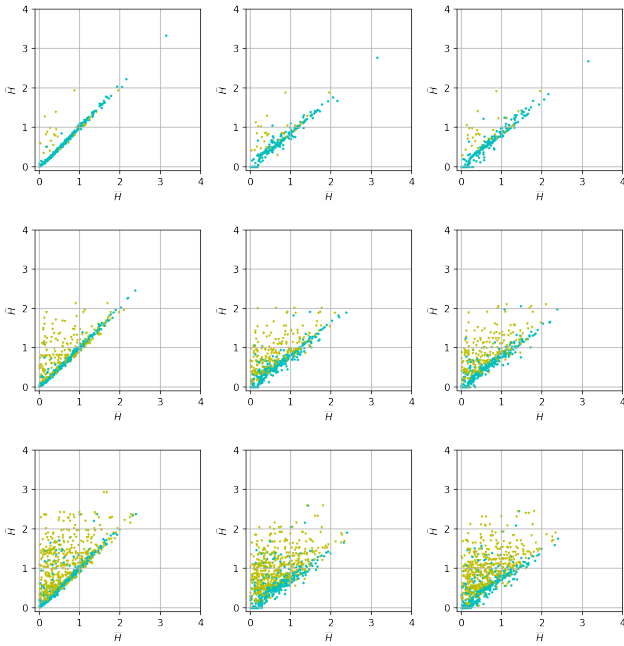


FIGURE 3 : Diagrammes de dispersion ( $\bar{\mathbf{H}}, \hat{\mathbf{H}}$ ) des hauteurs des pics, parcimonie variable pour D0, D1 et D2 (1e, 2e et 3e lignes), pour U-HQ, U-ISTA et U-PD de gauche à droite.

## Références

- [1] J. BOBIN, J. XU, A. de VISMES OTT et C. BOBIN : Learning to unmix from poisson measurements with application to  $\gamma$ -spectroscopy. In *Proceedings of the 8th Signal Processing with Adaptive Sparse Structured Representations (SPARS 2019)*, 2019.
- [2] D. CHEN, Z. WANG, D. GUO, V. OREKHOV et X. QU : Review and prospect : deep learning in Nuclear Magnetic Resonance spectroscopy. *Chemistry—A European Journal*, 26(46):10391–10401, 2020.
- [3] A. CHERNI, E. CHOUZENOUX, L. DUVAL et J. C. PESQUET : SPOQ  $\ell_p$ -over- $\ell_q$  regularization for sparse signal recovery applied to mass spectrometry. *IEEE Transactions on Signal Processing*, 68:6070–6084, 2020.
- [4] E. CHOUZENOUX, S. MOUSSAOUI, J. IDIER et F. MARIETTE : Efficient maximum entropy reconstruction of Nuclear Magnetic Resonance  $T_1 - T_2$  spectra. *IEEE Transactions on Signal Processing*, 58(12):6040–6051, 2010.
- [5] E. CHOUZENOUX, S. MOUSSAOUI, J. IDIER et F. MARIETTE : Primal-dual interior point optimization for a regularized reconstruction of NMR relaxation time distributions. In *Proceedings of the 38th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2013)*, pages 8747–8750, 2013.
- [6] J. CHRISTIAN, T. LEONARDO et S. MOUSSAOUI : *Source separation in physical-chemical sensing*. Wiley, 2024.
- [7] B. DEBUS, H. PARASTAR, P. HARRINGTON et D. KIRSANOV : Deep learning in analytical chemistry. *Trends in Analytical Chemistry*, 145:116459, 2021.
- [8] A. FELINGER : *Data analysis and signal processing in chromatography*. Elsevier, 1998.
- [9] M. GHARBI, E. CHOUZENOUX et J.-C. PESQUET : An unrolled half-quadratic approach for sparse signal recovery in spectroscopy. *Signal Process.*, 218:109369, mai 2024.
- [10] M. GHARBI, S. VILLA, E. CHOUZENOUX, J.-C. PESQUET et L. DUVAL : Unrolled deep networks for sparse signal restoration in analytical chemistry. In *IEEE Int. Workshop Mach. Learn. Signal Process.*, 2024.
- [11] J. HERSHEY, J. LE ROUX et F. WENINGER : Deep unfolding : Model-based inspiration of novel deep architectures. Rapport technique, Sep. 2014. <https://arxiv.org/abs/1409.2574>.
- [12] A. KENSERT, E. BOSTEN, G. COLLAERTS, K. EFTHYMIADIS, P. VAN BROECK, G. DESMET et D. CABOOTER : Convolutional neural network for automated peak detection in reversed-phase liquid chromatography. *J. Chrom. A*, 1672:463005, juin 2022.
- [13] C. KERVAZO, A. CHETOUTI et J. E. COHEN : Deep unrolling of the multiplicative updates algorithm for blind source separation, with application to hyperspectral unmixing. In *Proc. Eur. Sig. Image Proc. Conf.*, Aug. 2024.
- [14] H. LIU, T. ZHANG, L. YAN, H. FANG et Y. CHANG : A MAP-based algorithm for spectroscopic semi-blind deconvolution. *Analyst*, 137:3862–3873, 2012.
- [15] V. MARTIKYAN, C. BELUFFI, S.J. GLASER, M.A. DELSUC et D. SUGNY : Application of optimal control theory to Fourier transform ion cyclotron resonance. *Molecules*, 26(10), 2021.
- [16] A. C. MATER et M. L. COOTE : Deep learning in chemistry. *Journal of Chemical Information and Modeling*, 59(6):2545–2559, 2019.
- [17] T. H. NGUYEN, R. K. REDDY, M. J. WALSH, Matthew SCHULMERICH, G. POPESCU, M. N. DO et R. BHARGAVA : Denoising and deblurring of Fourier transform infrared spectroscopic imaging data. In *Proceedings of the 10th Computational Imaging Conference (COIMG 2012)*, volume 8296, pages 112–120. SPIE, 2012.
- [18] A. B. RISUM et R. BRO : Using deep learning to evaluate peaks in chromatographic data. *Talanta*, 204:255–260, nov. 2019.
- [19] J. SHI et S. E. REICHENBACH : Restoration for comprehensive two-dimensional gas chromatography. In *Proceedings of the 5th IEEE International Conference on Electro Information Technology (EIT 2005)*, pages 6–pp. IEEE, 2005.
- [20] P. TRIPATHI, N. KUMAR et A. H. SIDDIQI : De-noising Raman spectra using total variation de-noising with iterative clipping algorithm. *Computational Science and its Applications, Taylor and Francis Group*, pages 225–231, 2020.
- [21] Z. WANG, C. QIAN, D. GUO, H. SUN, R. LI, B. ZHAO et X. QU : One-dimensional deep low-rank and sparse network for accelerated MRI. *IEEE Transactions on Medical Imaging*, 42(1):79–90, 2022.