Compressed Sensing in Imaging Mass Spectrometry

Andreas Bartels[1]

Joint work with

Patrick Dülk[1], Dennis Trede[1,2], Theodore Alexandrov[1,2,3] and Peter Maaß[1,2]

1 Center for Industrial Mathematics, University of Bremen, Bremen, Germany
2 Steinbeis Innovation Center SCiLS (Scientific Computing in Life Sciences), Bremen, Germany
3 MALDI Imaging Lab, University of Bremen, Bremen, Germany

Copenhagen

26.-28. March 2014

Outline

Mass spectrometry (MS)

Technique of analytical chemistry that identifies the elemental composition of a chemical sample based on mass-to-charge ratio of charged particles.

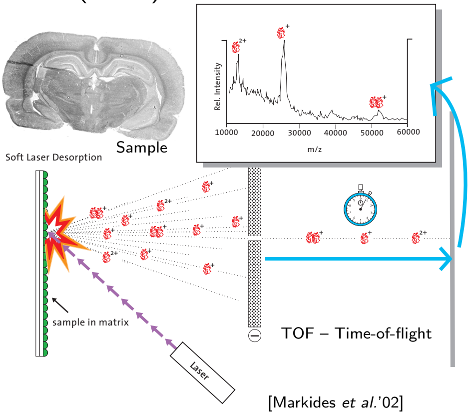
What is it used for?

- drug development
- detect/identify the use of drugs of abuse (dopings) in athlets
- identification of explosives and analysis of explosives in postblast residues (*puffer machine*)
- study the interaction of two (or more) bacterial cultures
- detection of disease biomarkers
- determination of proteins, peptides, metabolites
- and ...

MS methods

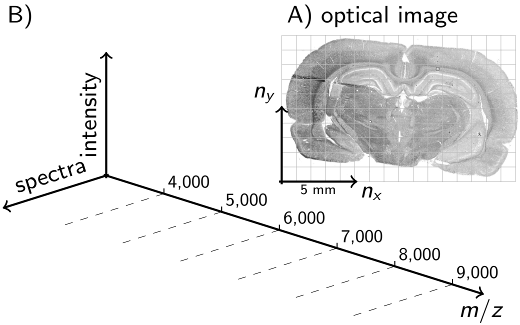
- Matrix-Assisted Laser Desorption/Ionization (MALDI)
- Secondary Ion Mass Spectrometry (SIMS)
- Desorption Electrospray Ionization (DESI)
- ...

1. sample is cut and mounted on glass slide
2. matrix solution is applied (acid crystalisation)
3. laser desorption of 'spots' (grid ∼ 20 µm – 200 µm)
4. computer aided analysis of m/z-slices



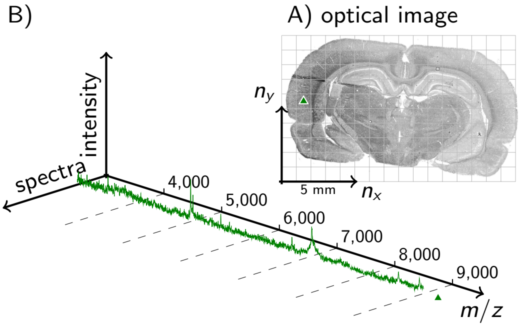
[Markides *et al.*'02]

Matrix-assisted laser desorption/ionization



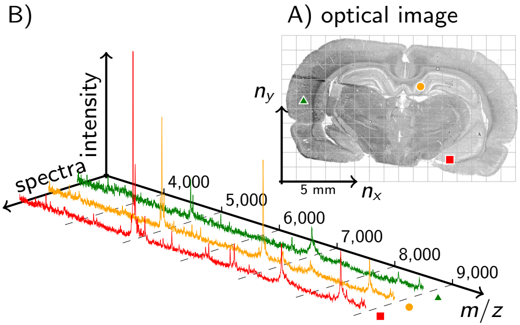
[Alexandrov *et al.*'11]

Matrix-assisted laser desorption/ionization



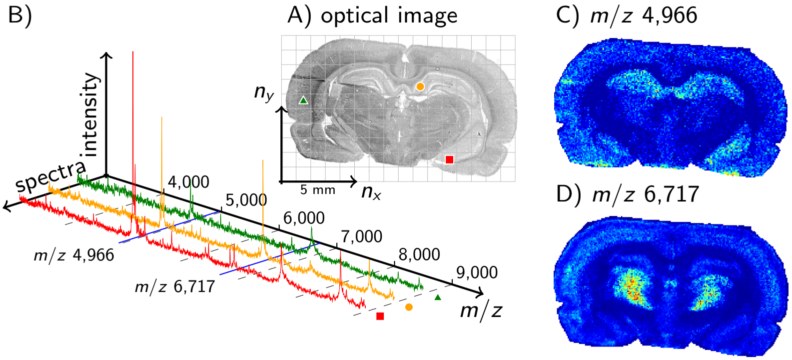
[Alexandrov *et al.*'11]

Matrix-assisted laser desorption/ionization



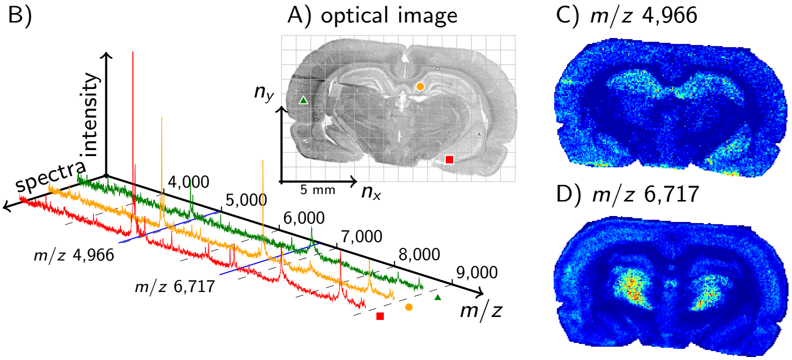
[Alexandrov *et al.*'11]

Matrix-assisted laser desorption/ionization



[Alexandrov *et al.*'11]

Matrix-assisted laser desorption/ionization



[Alexandrov *et al.*'11]

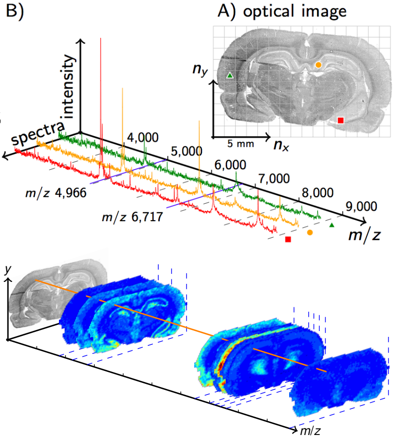
$\Longrightarrow$ IMS data: *Hyperspectral* data $X \in \mathbb{R}_+^{n_x \times n_y \times c}$ (m/z-spectra and -images)

The information disaster – data overflow

Data $X \in \mathbb{R}_+^{n_x \times n_y \times c}$ typically contains

- $n_x \cdot n_y = 10{,}000 - 100{,}000$ pixels
- $c = 10{,}000 - 100{,}000$ m/z-values
- $10^8 - 10^{10}$ values, altogether

Write $X \in \mathbb{R}_+^{n \times c}$, $n = n_x \cdot n_y$.



(General) Questions:

- ⇝ How to interpret the data?
- ⇝ What is the *main* information?
- ⇝ How to compress the data?
- ⇝ Where to compress the data?

Compression perspectives

Mass spectrometry data $X \in \mathbb{R}_+^{n \times c}$ is typically large!

- *Nonnegative matrix factorization*

$$X \approx MS,$$

where $M \in \mathbb{R}_+^{n \times \rho}$ and $S \in \mathbb{R}_+^{\rho \times c}$ with $\rho \ll \min\{n, c\}$.

$$\rightsquigarrow \quad \min_{M,S} \; \alpha\Theta_1(M) + \beta\Theta_2(S) \quad \text{s.t.} \quad \|X - MS\|_F \leq \varepsilon$$

M – pseudo m/z-images, S – pseudo spectra

- *Compressed Sensing*

$$Y = \Phi X \in \mathbb{R}_+^{m \times c},$$

where $\Phi \in \mathbb{R}_+^{m \times n}$, $m \ll n$.

$$\rightsquigarrow \quad \min_{X} \; \alpha\Theta_1(X) + \beta\Theta_2(X) \quad \text{s.t.} \quad \|Y - \Phi X\|_F \leq \varepsilon$$

Compressed Sensing in IMS

Problems:

- MALDI measurements require several hours in *time*
- Data *interpretation* on full data

Example: Rat brain dataset $\sim$ 5 hours

Idea: Make use of *compressed sensing* with the knowledge of

- sparse m/z-spectra (ℓ_1 minimization) and
- sparse m/z-images (TV minimization)

Compressed Sensing in IMS

Problems:

- MALDI measurements require several hours in *time*
- Data *interpretation* on full data

Example: Rat brain dataset ∼ 5 hours

Idea: Make use of *compressed sensing* with the knowledge of

- sparse m/z-spectra (ℓ_1 minimization) and
- sparse m/z-images (TV minimization)

⇝ A.B., P. Dülk, D. Trede, T. Alexandrov and P. Maaß, "Compressed Sensing in Imaging Mass Spectrometry", *Inverse Problems*, **29**(12), 125015 (24pp), 2013.

CS-IMS model - The data

IMS data is a hyperspectral data cube consisting of $n_x \cdot n_y$ (m/z-)spectra of length c (number of channels), whereas n_x and n_y are the number of pixels in each coordinate direction. Thus,

$$X \in \mathbb{R}_+^{n_x \times n_y \times c}.$$

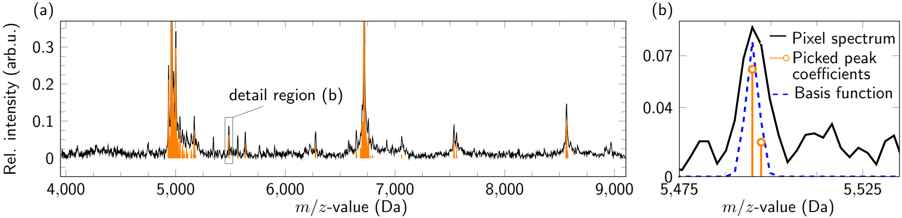
Concatenating each m/z-image as a vector the data X becomes

$$X \in \mathbb{R}_+^{n \times c},$$

where $n := n_x \cdot n_y$.

Each *column* corresponds to one m/z-image
Each *row* to one m/z-spectrum.

Sparsity in IMS data - m/z-spectra



Sparsity of the *spectra* in a basis $\Psi \in \mathbb{R}_+^{c \times c}$

- Only a few peaks arise with high intensities
- $\Longrightarrow$ Feature extraction via ℓ_1 minimization [Denis *et al.*'09]
- Shifted Gaussians: $\psi_k(x) = \frac{1}{\pi^{1/4}\sigma^{1/2}} \exp\left(-\frac{(x-k)^2}{2\sigma^2}\right)$, $k = 1, .., c$

Sparsity in IMS data - m/z-spectra

Let $X^T_{(k,\cdot)} \in \mathbb{R}^c_+$, $k = 1, \dots, n$, be the k-th row of $X \in \mathbb{R}^{n\times c}_+$, i.e. one spectrum. We assume the spectra to be *sparse* or *compressible* in a (known) basis $\Psi \in \mathbb{R}^{c\times c}_+$, i.e.

$$X^T_{(k,\cdot)} = \Psi\lambda, \quad \lambda \in \mathbb{R}^c_+ \tag{1}$$

at which $\|\lambda\|_0 \ll c$.
With coefficient matrix $\Lambda \in \mathbb{R}^{c\times n}_+$, Eq. (1) reads

$$X^T = \Psi\Lambda.$$

$\Longrightarrow$ Minimize *columns* $\Lambda_{(\cdot,k)}$ of Λ w.r.t. the ℓ_0 'norm', i.e.

$$\boxed{\|\Lambda_{(\cdot,k)}\|_0 \text{ for } k = 1, \dots, n.}$$

Sparsity in IMS data - m/z-spectra

Let $X^T_{(k,\cdot)} \in \mathbb{R}^c_+$, $k = 1, \dots, n$, be the k-th row of $X \in \mathbb{R}^{n \times c}_+$, i.e. one spectrum. We assume the spectra to be *sparse* or *compressible* in a (known) basis $\Psi \in \mathbb{R}^{c \times c}_+$, i.e.

$$X^T_{(k,\cdot)} = \Psi\lambda, \quad \lambda \in \mathbb{R}^c_+ \tag{1}$$

at which $\|\lambda\|_0 \ll c$.
With coefficient matrix $\Lambda \in \mathbb{R}^{c \times n}_+$, Eq. (1) reads

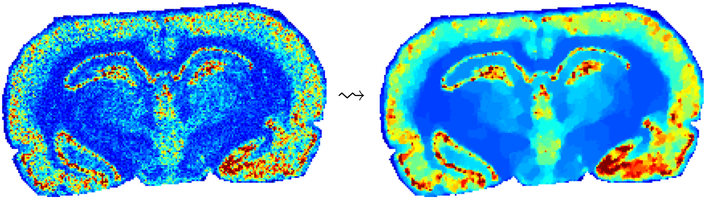
$$X^T = \Psi\Lambda.$$

$\Longrightarrow$ Minimize *columns* $\Lambda_{(\cdot,k)}$ of Λ w.r.t. the ℓ_1 'norm', i.e.

$$\boxed{\|\Lambda_{(\cdot,k)}\|_1 \text{ for } k = 1, \dots, n.}$$

Sparsity in IMS data - m/z-images

Sparsity of the m/z-*images*



[Alexandrov *et al.*'10]:
m/z-images of imaging mass spectrometry data usually

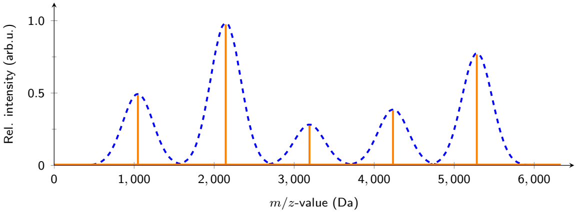
- inherent large variance of noise
- are piecewise constant

$\Longrightarrow$ Apply *TV* denoising on each m/z-image.

Sparsity in IMS data - m/z-images

Recall that $\Lambda \in \mathbb{R}_+^{c \times n}$ is the coefficient matrix in $X^T = \Psi\Lambda$.
$\Longrightarrow$ Minimize *rows* $\Lambda_{(k,\cdot)}$ of Λ w.r.t. the *TV* norm, i.e.

$$\boxed{\|\Lambda_{(j,\cdot)}\|_{TV} \text{ for } j = 1, \ldots, c.}$$



Instead of finding a reconstruction $\tilde{X}^T = \Psi\tilde{\Lambda}$, we aim to directly recover the *features* $\tilde{\Lambda}$.

The compressed sensing process

IMS data $X \in \mathbb{R}_+^{n \times c}$ *acquisition:* Ionizing the given sample on each of the n pixels on a predefined grid .

The compressed sensing process

IMS data $X \in \mathbb{R}_+^{n \times c}$ *acquisition:* Ionizing the given sample on each of the n pixels on a predefined grid $\rightsquigarrow$ *Compressed sensing*: $m \ll n$.

The compressed sensing process

IMS data $X \in \mathbb{R}_+^{n\times c}$ *acquisition:* Ionizing the given sample on each of the n pixels on a predefined grid $\rightsquigarrow$ *Compressed sensing*: $m \ll n$.

Take m measurements $y_i \in \mathbb{R}_+^c$, $i = 1,..,m$:

$$y_{ij} = \langle \varphi_i, X_{(\cdot,j)} \rangle, \quad j = 1,..,c, \quad \varphi_i \in \mathbb{R}_+^n,$$

φ_i from sub-gaussian distribution

Each y_i for $i = 1,..,m$ is a *measurement-mean spectrum* since it is calculated by the mean intensities on each channel:

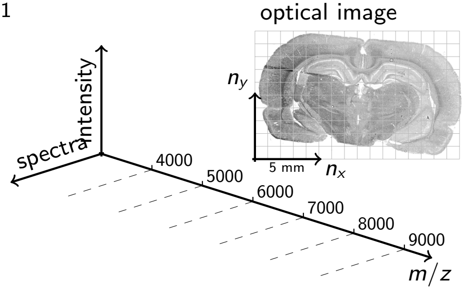
$$y_i^T = \varphi_i^T X = \sum_{k=1}^{n} \varphi_{ik} X_{(k,\cdot)},$$

$\rightsquigarrow$ y_i^T are *linear combinations* of the original spectra $X_{(k,\cdot)}$, $k = 1,..,n$.

CS-IMS model

CS in IMS data acquisition: Ionizing the given sample on randomly selected pixels on a predefined grid.

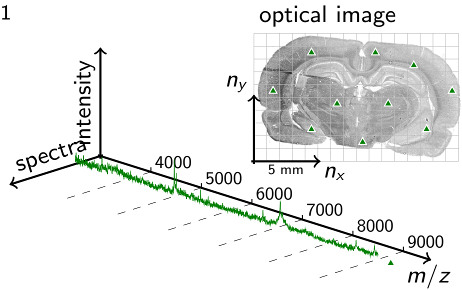
$$y_i{}^T = \varphi_i{}^T X = \sum_{k=1}^{n} \varphi_{ik} X_{(k,\cdot)}, \quad X^T = \Psi\Lambda.$$



CS-IMS model

CS in IMS data acquisition: Ionizing the given sample on randomly selected pixels on a predefined grid.

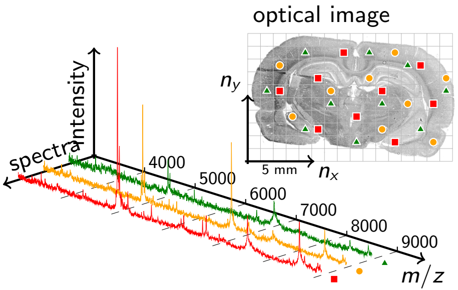
$$y_1{}^T = \varphi_1{}^T X = \sum_{k=1}^{n} \varphi_{1k} X_{(k,\cdot)}, \quad X^T = \Psi\Lambda.$$



CS-IMS model

CS in IMS data acquisition: Ionizing the given sample on randomly selected pixels on a predefined grid.

$$y_{2/3}^T = \varphi_{2/3}^T X = \sum_{k=1}^{n} \varphi_{2/3k} X_{(k,\cdot)}, \quad X^T = \Psi\Lambda.$$



CS-IMS model

CS in IMS data acquisition: Ionizing the given sample on randomly selected pixels on a predefined grid.

$$y_{2/3}^T = \varphi_{2/3}^T X = \sum_{k=1}^{n} \varphi_{2/3k} X_{(k,\cdot)}, \quad X^T = \Psi\Lambda.$$

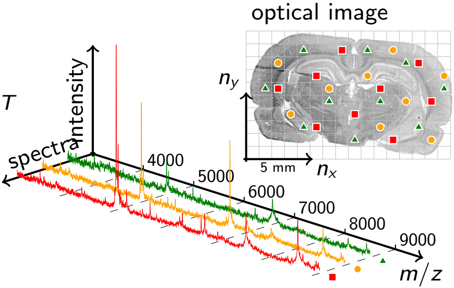
In matrix form this becomes

$$\mathbb{R}_+^{m\times c} \ni Y = \Phi X = \Phi\Lambda^T\Psi^T$$

$$Y = \Phi X + Z, \quad \|Z\|_F \leq \varepsilon$$

$\Psi \in \mathbb{R}_+^{c\times c}$ - Dictionary,

$\Lambda \in \mathbb{R}_+^{c\times n}$ - Coefficients



CS-IMS model

CS in IMS data acquisition: Ionizing the given sample on randomly selected pixels on a predefined grid.

$$y_{2/3}^T = \varphi_{2/3}^T X = \sum_{k=1}^{n} \varphi_{2/3k} X_{(k,\cdot)}, \quad X^T = \Psi\Lambda.$$

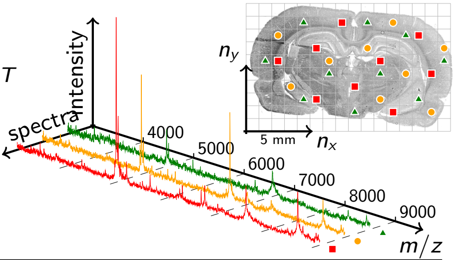
In matrix form this becomes

$$\mathbb{R}_+^{m\times c} \ni Y = \Phi X = \Phi\Lambda^T\Psi^T$$

$$Y = \Phi X + Z, \quad \|Z\|_F \leq \varepsilon$$

$\Psi \in \mathbb{R}_+^{c\times c}$ - Dictionary,
$\Lambda \in \mathbb{R}_+^{c\times n}$ - Coefficients

optical image



$$\min_{\Lambda\in\mathbb{R}^{c\times n}} \alpha \sum_{j=1}^{c} \|\Lambda_{(j,\cdot)}\|_{TV} + \beta\|\Lambda\|_1, \quad \text{s.t.} \quad \|Y - \Phi\Lambda^T\Psi^T\|_F \leq \varepsilon, \quad \Lambda \geq 0$$

Theorem: Robustness

A.B., P.Dülk, D.Trede, T.Alexandrov and P.Maaß, "CS in IMS", *Inverse Problems*, **29**(12), 125015 (24pp), 2013.

Let $\mathcal{M} : \mathbb{R}^{c\times n} \to \mathbb{R}^{4cm_1} \times \mathbb{R}^{m_2\times c}$ be the linear operator with components

$$\mathcal{M}(\Lambda) = \left(\mathcal{A}^0\Lambda_1, \mathcal{A}_0\Lambda_1, \mathcal{A}'^0\Lambda_1, \mathcal{A}'_0\Lambda_1, \dots, \mathcal{A}^0\Lambda_c, \mathcal{A}_0\Lambda_c, \mathcal{A}'^0\Lambda_c, \mathcal{A}'_0\Lambda_c, \Phi\Lambda^T\Psi^T\right).$$

If noisy measurements $Y = \mathcal{M}(\Lambda) + Z$ are observed with noise level $\|Z\|_F \leq \varepsilon$, then

$$\Lambda^\diamond = \operatorname*{argmin}_{W\in\mathbb{R}^{c\times n}} \|W\|_1 + \sum_{i=1}^{c} \|W_i\|_{TV} \quad \text{s.t.} \quad \|\mathcal{M}(W) - Y\|_F \leq \varepsilon,$$

satisfies both

$$\|\Lambda - \Lambda^\diamond\|_F + \sum_{i=1}^{c} \|\nabla\Lambda_i - \nabla\Lambda_i^\diamond\|_F \lesssim \frac{1}{\sqrt{K}} \left(\|\Lambda - \Lambda_{S_0}\|_1 + \sum_{i=1}^{c} \left\|\nabla\Lambda_i - (\nabla\Lambda_i)_{S_i}\right\|_1 \right) + \varepsilon,$$

and

$$\|\Lambda - \Lambda^\diamond\|_1 + \sum_{i=1}^{c} \|\Lambda_i - \Lambda_i^\diamond\|_{TV} \lesssim \|\Lambda - \Lambda_{S_0}\|_1 + \sum_{i=1}^{c} \left\|\nabla\Lambda_i - (\nabla\Lambda_i)_{S_i}\right\|_1 + \sqrt{K}\varepsilon.$$

Theorem: Robustness - The tools

1. RIP
   The linear operator $\mathcal{A} : \mathbb{R}^{n_x \times n_y} \to \mathbb{R}^{m \times p}$ has the *restricted isometry property* of order s and level $\delta \in (0,1)$ if

$$(1-\delta)\|X\|_F^2 \le \|\mathcal{A}(X)\|_F^2 \le (1+\delta)\|X\|_F^2$$

   for all s-sparse $X \in \mathbb{R}^{n_x \times n_y}$.

2. D-RIP (extends the RIP to matrices adapted to a dictionary)
   A linear operator $\mathcal{A} : \mathbb{R}^{n_x \times n_y} \to \mathbb{R}^{m \times p}$ has the *D-RIP* of order s and level $\delta^* \in (0,1)$, adapted to a dictionary D, if for all s-sparse $X \in \mathbb{R}^{n_x \times n_y}$ it holds

$$(1-\delta^*)\|DX\|_F^2 \le \|\mathcal{A}(DX)\|_F^2 \le (1+\delta^*)\|DX\|_F^2.$$

Theorem: Robustness - The tools

3. A-RIP

A matrix $D \in \mathbb{R}^{n_x \times n_x}$ satisfies the *asymmetric restricted isometry property* (A-RIP), if for all s-sparse $X \in \mathbb{R}^{n_x \times n_y}$ the following inequalities hold:

$$\mathcal{L}(D)\|X\|_F \leq \|DX\|_F \leq \mathcal{U}(D)\|X\|_F,$$

where $\mathcal{L}(D)$ and $\mathcal{U}(D)$ are the largest and the smallest constants for which the above inequalities hold. The restricted condition number of D is defined as

$$\xi(D) = \frac{\mathcal{U}(D)}{\mathcal{L}(D)} \leq \frac{\max_{\|X\|_F=1} \|DX\|_F}{\min_{\|X\|_F=1} \|DX\|_F} = \kappa(D).$$

Theorem: Robustness - The tools

$$\min_{\Lambda \in \mathbb{R}^{c \times n}} \sum_{j=1}^{c} \|\Lambda_{(j,\cdot)}\|_{TV} + \|\Lambda\|_1, \quad \text{s.t.} \quad \|Y - \underbrace{\Phi \Lambda^T \Psi^T}_{=:\mathcal{D}_{\Phi,\Psi}\Lambda}\|_F \leq \varepsilon, \quad \Lambda \geq 0$$

Argue that

- $\mathcal{D}_{\Phi,\Psi}$ fulfils the D-RIP
  Argument via Kronecker product and blockdiagonal RIP results [Eftekhari, A, *et al.*'12]
- Ψ fulfils the A-RIP
  Ψ will consist of shifted Gaussians
  $\rightsquigarrow$ Ψ is invertible, i.e. $\xi(\Psi)$ is bounded by $\kappa(\Psi)$
- $\mathcal{B} = [\mathcal{A}\ \mathcal{A}', \ldots, \mathcal{A}\ \mathcal{A}']$
  (operator with artificial gradient measurements)
  fulfils the RIP Argument similar as for $\mathcal{D}_{\Phi,\Psi}$.

Algorithm (PPXA)

We aim to solve the following problem

$$\min_{\Lambda \in \mathbb{R}^{c \times n}} \alpha \sum_{j=1}^{c} \|\Lambda_{(j,\cdot)}\|_{TV} + \beta \|\Lambda\|_1, \quad \text{s.t.} \quad \|Y - \Phi \Lambda^T \Psi^T\|_F \leq \varepsilon, \quad \Lambda \geq 0.$$

We use the *parallel proximal splitting algorithm* (PPXA) [Combettes&Pesquet'08] which solves problems of the kind:

$$\min_{x \in \mathcal{H}} \sum_{i=1}^{\ell} f_i(x),$$

where

- $\mathcal{H}$ is a Hilbert space and
- $(f_i)_{1 \leq i \leq \ell}$ are proper lower semicontinuous convex functions $f_i : \mathcal{H} \to]-\infty, +\infty]$

Algorithm (PPXA)

Here: $\mathcal{H} = \mathbb{R}^{c \times n}$, $\ell = 4$ and

$$f_1(\Lambda) = \alpha \sum_{i=1}^{c} \|\Lambda_i\|_{TV}, \qquad f_2(\Lambda) = \beta \|\Lambda\|_1,$$

$$f_3(\Lambda) = \iota_{\mathcal{B}_2^\varepsilon}(\Lambda), \qquad f_4(\Lambda) = \iota_{\mathcal{B}_+}(\Lambda).$$

$\iota_{\mathcal{C}}$ is the indicator function,

$$\iota_{\mathcal{C}}(\Lambda) = \begin{cases} 0 & \text{if } \Lambda \in \mathcal{C} \\ +\infty & \text{otherwise} \end{cases},$$

applied to the convex sets

$$\mathcal{B}_2^\varepsilon = \{A \in \mathbb{R}^{c \times n} : \|Y - \mathcal{D}_{\Phi,\Psi} A\|_F \leq \varepsilon\} \quad \text{(Fidelity constraint)},$$
$$\mathcal{B}_+ = \{A \in \mathbb{R}^{c \times n} : A \geq 0\} \quad \text{(Positive orthant)}.$$

The proximity operator. . .

. . . is defined as

$$\text{prox}_f : \mathcal{H} \to \mathcal{H}.$$

$\text{prox}_f(X)$ is the unique point in $\mathcal{H}$ that satisfies

$$\text{prox}_f(X) = \underset{Y \in \mathcal{H}}{\text{argmin}} \frac{1}{2}\|X - Y\|_F^2 + f(Y).$$

For f_1 (TV-norm): Via an implementation from [Beck'09].

For f_2 (1-norm): $\text{prox}_{\gamma\|\cdot\|_1}(Z) = \left(\max\left\{0, \left(1 - \frac{\gamma}{|Z_{i,j}|}\right)\right\} Z_{i,j} \right)_{\substack{1\le i\le c \\ 1\le j\le n}}$

For f_3 ($\mathcal{B}_2^\varepsilon$): Via a Douglas-Rachford splitting scheme [Fadili'09]

For f_4 ($\mathcal{B}_+$): $\text{prox}_{\gamma\iota_{\mathcal{B}_+(\cdot)}}(Z) = \left(\max\{0, Z_{i,j}\} \right)_{\substack{1\le i\le c \\ 1\le j\le n}}$

Parallel proximal splitting algorithm

Algorithm 1: PPXA

Input: $Y, \Psi, \Phi,\ \alpha, \beta, \varepsilon, \gamma > 0$

Initializations: $k = 0$; $\Lambda_0 = \Gamma_{1,0} = \Gamma_{2,0} = \Gamma_{3,0} = \Gamma_{4,0} \in \mathbb{R}^{c\times n}$

repeat

  for $j = 1 : 4$ **do**

    $P_{j,k} = \mathrm{prox}_{\gamma f_j}(\Gamma_{j,k})$

  $\Lambda_{k+1} = (P_{1,k} + P_{2,k} + P_{3,k} + P_{4,k})/4$

  for $j = 1 : 4$ **do**

    $\Gamma_{j,k+1} = \Gamma_{j,k} + 2\Theta_{k+1} - \Theta_k - P_{j,k}$

until *convergence*

Implementations given in the UNLocBoX [Perraudin'14].

Test-Setting - Rat brain dataset

Part of a rat brain dataset: $X \in \mathbb{R}_+^{n \times c}$, $n = 121 \cdot 202$, $c = 2,000$.

Assume mass spectra to be sparse in a basis Ψ consisting of shifted Gaussians [Denis *et al.*'09]

$$\psi_k(x) = \frac{1}{\pi^{1/4}\sigma^{1/2}} \exp\left(-\frac{(x-k)^2}{2\sigma^2}\right)$$

Choose std. deviation of $\Psi_k(x)$

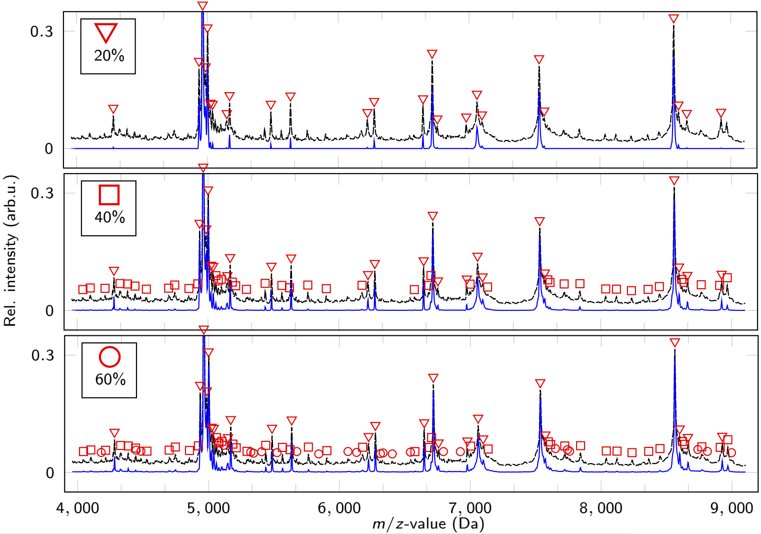
- consistent with the data and
- such that the condition $\kappa(\Psi)$ is small

$\rightsquigarrow \quad \sigma = 0.75.$

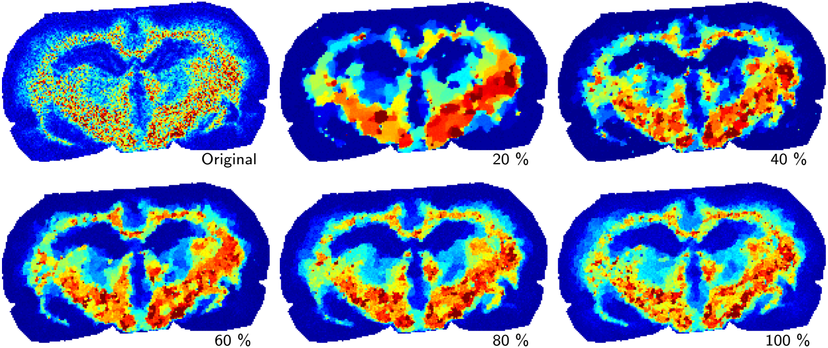
Elements of measurement matrix $\Phi \in \mathbb{R}^{r \times n}$ $(r \ll n)$ chosen at random from an i.i.d. Gaussian distribution.

Noise level $\varepsilon = 3.75 \times 10^3$. Parameters α, β set by hand.

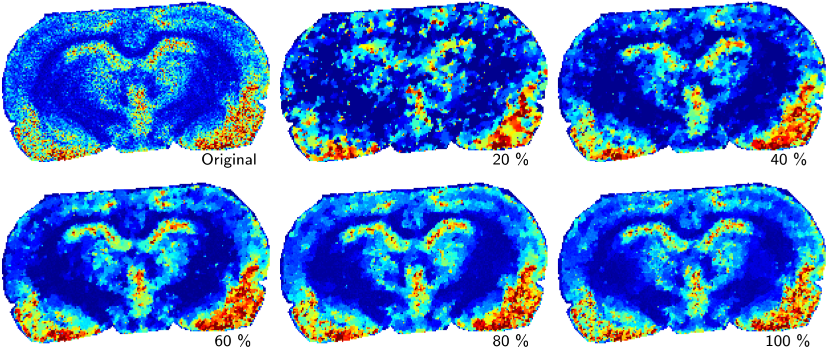
Reconstructed mean spectrum



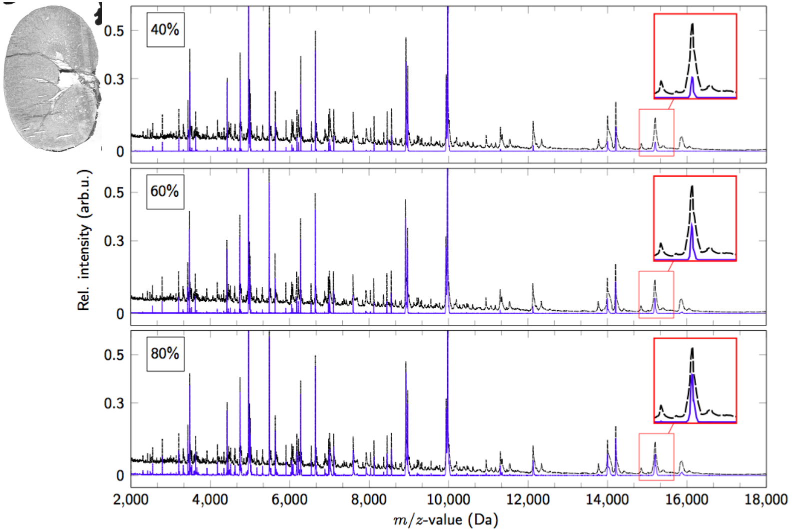
Reconstructed m/z-image



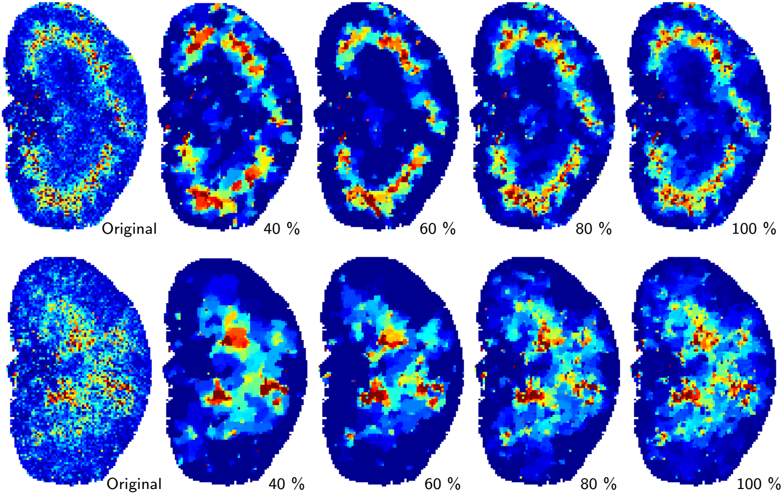
Reconstructed m/z-image - Cont.



Kidney - $X \in \mathbb{R}_+^{n \times c}$, $n = 113 \cdot 71$, $c = 10{,}000$, $\varepsilon = 7 \times 10^4$



Reconstructed m/z-image



Others?

Are there similar results for other MS systems?

Others?

Are there similar results for other MS systems?

⇝ Yes!

Gao, Y., Zhu, L., Norton, I., Agar, N. Y. R., Tannenbaum, A. *Reconstruction and feature selection for desorption electrospray ionization mass spectroscopy imagery* Proc. SPIE 9036, Medical Imaging 2014, March 12, 2014. (DESI)

From the abstract:

> *"...time it takes for imaging and data analysis becomes a critical factor. Therefore, [...] we utilize compressive sensing to perform the sparse sampling of the tissue, which halves the scanning time."*

Conclusions

- First model for compressed sensing in MALDI-IMS
- Reconstruction of whole dataset w.r.t. its features
- Robustness of reconstruction with respect to noise
  - combines ℓ_1 and TV
  - includes two sparsity aspects

Future Prospects

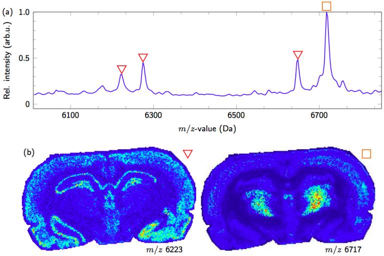
- How to chose regularization parameters α and β?
- Noise models (e.g. Poisson noise, etc.)
- Numerics (alternating minimization, surrogate functionals)
- *Include sparse representation*

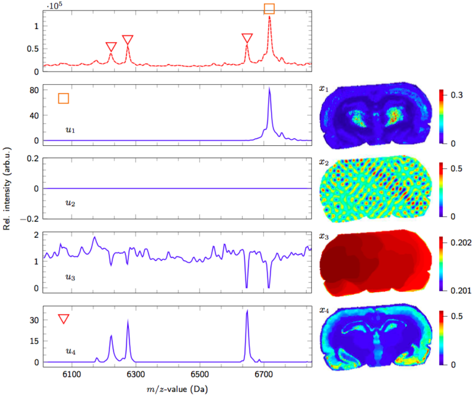
$$X \approx MS, \qquad M \in \mathbb{R}_+^{m\times\rho}, \qquad S \in \mathbb{R}_+^{\rho\times n}$$

M – Matrix with m/z-images, S – Matrix with *pseudo spectra*

Outlook - Nonnegative matrix factorization

$$\min_{M,S} \|X - MS\|_F^2 + \alpha \sum_{i=1}^{\rho} \|M_i\|_{TV} + \beta \|S\|_1$$
$$\text{s.t.} \quad M \geq 0, S \geq 0.$$





Pham, A.B., P.M. - *in progress*

Working group

Thank you!

bartels@math.uni-bremen.de

Bibliography

[1] Alexandrov, T. et al.: *Spatial segmentation of imaging mass spectrometry data with edge-preserving image denoising and clustering*, J Proteome Research, **6**, pp. 6535–6546, 2010

[2] Alexandrov, T. and Kobarg, J.H.: *Efficient spatial segmentation of large imaging mass spectrometry datasets with spatially aware clustering*, Bioinformatics, **27**, pp. i230–i238, 2011

[3] Denis, L., et al.: *Greedy Solution of Ill-Posed Problems: Error Bounds and Exact Inversion*, Inverse Problems, **25**:11, pp. 1–14, 2009

[4] Markides, K. and Gräslund, A.: *Mass spectrometry (MS) and nuclear magnetic resonance (NMR) applied to biological macromolecules*, Adv. inf. on the Nobel Prize in Chemistry 2002, The Royal Swedish Academy of Sciences

Bibliography - Continued

[5] Golbabaee, M. and Arberet, S. and Vandergheynst, P., *Distributed compressed sensing of hyperspectral images via blind source separation*, Presentation given at Asilomar conf. on signals, systems, and computers, Pacific Groove, CA, USA, November 7-10, 2010

[6] Golbabaee, M. and Arberet, S. and Vandergheynst, P., *Compressive source separation: Theory and Methods for hyperspectral imaging*, Submitted, 2012

[7] Combettes, P. L. and Pesquet, J.-C., *A proximal decomposition method for solving convex variational inverse problems*, Inverse Problems, **24**:065014, 2008

[8] Eftekhari, A, *et al.*: *The restricted isometry property for block diagonal matrices*, arXiv:1210.3395v2, 2014

Bibliography - Continued

[9] Beck, A. and Teboulle, M.: *Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems*, IEEE Trans. Image Proc. **18**: pp. 2419–2434

[10] Fadili, M. J. and Starck, J.-L..: *Monotone operator splitting for optimization problems in sparse recovery*, 16th IEEE Int. Conf. on Image Processing, pp. 1461–1464, 2009

[11] Bartels, A. *et al.*: *Compressed Sensing in Imaging Mass Spectrometry*, Inverse Problems, **29**(12), 125015 (24pp), 2013

[12] Perraudin, N. *et al.*: *UNLocBoX A Matlab convex optimization toolbox using proximal splitting methods*, arXiv:1402.0779 , 2014