BEADS : filtrage asymétrique de ligne de base (tendance) et débruitage pour des signaux positifs avec parcimonie des dérivées

Séminaire ICube

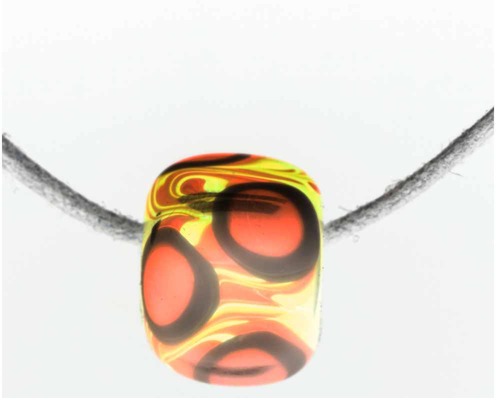
L. DUVAL, A. PIRAYRE
IFP Energies nouvelles
1 et 4 av. de Bois-Préau, 92852 Rueil-Malmaison - France

X. NING, I. W. SELESNICK
Polytechnic School of Engineering
New York University

19 juin 2015

The fast way

- Question: where is the string behind the bead?
- Smoothness, sparsity, asymmetry

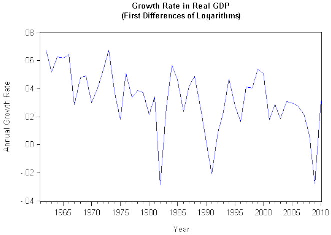


Outline

Background on background

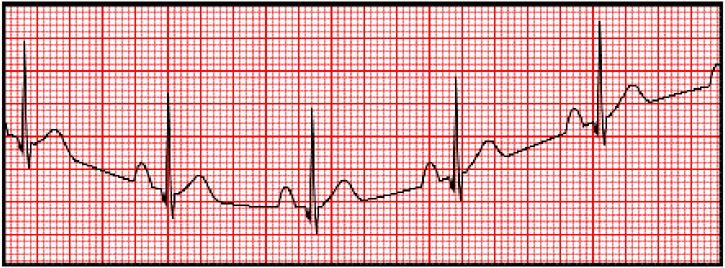
- Background affects quantitative evaluation/comparison
- In some domains: (instrumental) bias, (seasonal) trend
- In analytical chemistry: drift, continuum, wander, *baseline*
- Rare cases of parametric modeling

Background on background



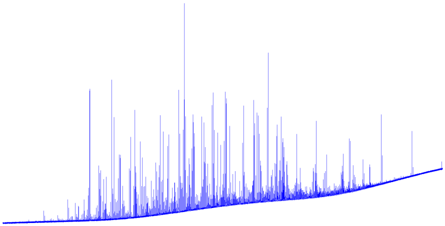
- Background affects quantitative evaluation/comparison
- In some domains: (instrumental) bias, (seasonal) trend
- In analytical chemistry: drift, continuum, wander, *baseline*
- Rare cases of parametric modeling

Background on background



- Background affects quantitative evaluation/comparison
- In some domains: (instrumental) bias, (seasonal) trend
- In analytical chemistry: drift, continuum, wander, *baseline*
- Rare cases of parametric modeling

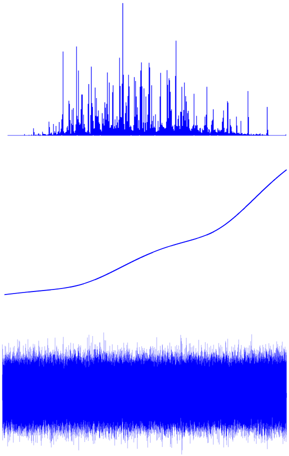
Background on background



- Background affects quantitative evaluation/comparison
- In some domains: (instrumental) bias, (seasonal) trend
- In analytical chemistry: drift, continuum, wander, *baseline*
- Rare cases of parametric modeling

Background on background

For analytical chemistry data:



Notations

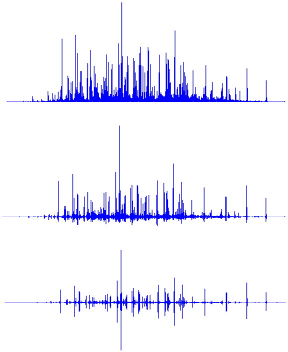
Morphological decomposition:
$\mathbf{y} = \mathbf{x} + \mathbf{f} + \mathbf{w}, \quad (\mathbf{y}, \mathbf{x}, \mathbf{f}, \mathbf{w}) \in (\mathbb{R}^N)^4.$

- $\mathbf{y}$: observation
- $\mathbf{x}$: clean series of peaks
- $\mathbf{f}$: baseline
- $\mathbf{w}$: noise

Assumption: in the absence of peaks, the baseline can be approximately recovered from a noise-corrupted observation by low-pass filtering

- $\hat{\mathbf{f}} = \mathbf{L}(\mathbf{y} - \hat{\mathbf{x}})$ ($\mathbf{L}$: low-pass filter)
- formulated as $\|\mathbf{y} - \hat{\mathbf{s}}\|_2^2 = \|\mathbf{H}(\mathbf{y} - \hat{\mathbf{x}})\|_2^2$
- $\mathbf{H} = \mathbf{I} - \mathbf{L}$: high-pass filter

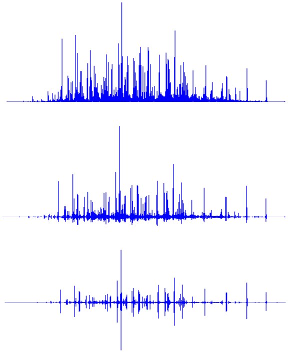
Compound sparse derivative modeling



An estimate $\hat{\mathbf{x}}$ can be obtained (with $\mathbf{D}_i$ diff. operators) via:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \left\{ F(\mathbf{x}) = \frac{1}{2} \|\mathbf{H}(\mathbf{y} - \mathbf{x})\|_2^2 + \sum_{i=0}^{M} \lambda_i R_i \left(\mathbf{D}_i \mathbf{x} \right) \right\}.$$

Compound sparse derivative modeling



Examples of (smooth) sparsity promoting functions for R_i

- $\phi_i^A = |x|$
- $\phi_i^B = \sqrt{|x|^2 + \epsilon}$
- $\phi_i^C = |x| - \epsilon \log(|x| + \epsilon)$

Compound sparse derivative modeling

Take the positivity of chromatogram peaks into account:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \Big\{ F(\mathbf{x}) = \frac{1}{2}\|\mathbf{H}(\mathbf{y}-\mathbf{x})\|_2^2 + \lambda_0 \sum_{n=0}^{N-1} \theta_\epsilon(x_n; r) + \sum_{i=1}^{M} \lambda_i \sum_{n=0}^{N_i-1} \phi\left([\mathbf{D}_i \mathbf{x}]_n\right) \Big\}.$$

Start from:

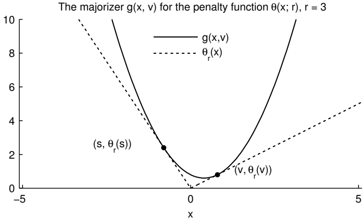
$$\theta(x;\, r) = \begin{cases} x, & x \geqslant 0 \\ -rx, & x < 0 \end{cases}$$

Compound sparse derivative modeling

Take the positivity of chromatogram peaks into account:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \Big\{ F(\mathbf{x}) = \frac{1}{2}\|\mathbf{H}(\mathbf{y}-\mathbf{x})\|_2^2 + \lambda_0 \sum_{n=0}^{N-1} \theta_\epsilon(x_n; r) + \sum_{i=1}^{M} \lambda_i \sum_{n=0}^{N_i-1} \phi\left([\mathbf{D}_i\mathbf{x}]_n\right) \Big\}.$$

and majorize it

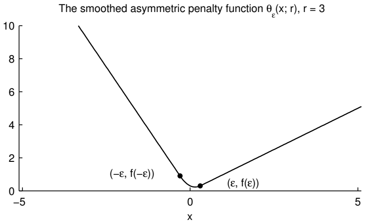


Compound sparse derivative modeling

Take the positivity of chromatogram peaks into account:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \Big\{ F(\mathbf{x}) = \frac{1}{2}\|\mathbf{H}(\mathbf{y}-\mathbf{x})\|_2^2 + \lambda_0 \sum_{n=0}^{N-1} \theta_\epsilon(x_n; r) + \sum_{i=1}^{M} \lambda_i \sum_{n=0}^{N_i-1} \phi\left([\mathbf{D}_i\mathbf{x}]_n\right) \Big\}.$$

then smooth it:



Compound sparse derivative modeling

Take the positivity of chromatogram peaks into account:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \Big\{ F(\mathbf{x}) = \frac{1}{2}\|\mathbf{H}(\mathbf{y}-\mathbf{x})\|_2^2 + \lambda_0 \sum_{n=0}^{N-1} \theta_\epsilon(x_n; r) + \sum_{i=1}^{M} \lambda_i \sum_{n=0}^{N_i-1} \phi\left([\mathbf{D}_i\mathbf{x}]_n\right) \Big\}.$$

then majorize it:

$$g_0(x, v) = \begin{cases} \frac{1+r}{4|v|}x^2 + \frac{1-r}{2}x + |v|\frac{1+r}{4}, & |v| > \epsilon \\ \frac{1+r}{4\epsilon}x^2 + \frac{1-r}{2}x + \epsilon\frac{1+r}{4}, & |v| \leqslant \epsilon. \end{cases}$$

BEADS Algorithm

We now have a majorizer for F

$$G(\mathbf{x},\mathbf{v}) = \frac{1}{2}\|\mathbf{H}(\mathbf{y}-\mathbf{x})\|_2^2 + \lambda_0 \mathbf{x}^\mathsf{T}[\mathbf{\Gamma}(\mathbf{v})]\mathbf{x} + \lambda_0 \mathbf{b}^\mathsf{T}\mathbf{x} + \sum_{i=1}^{M}\left[\frac{\lambda_i}{2}(\mathbf{D}_i\mathbf{x})^\mathsf{T}\left[\Lambda(\mathbf{D}_i\mathbf{v})\right](\mathbf{D}_i\mathbf{x})\right] + c(\mathbf{v}).$$

Minimizing $G(\mathbf{x},\mathbf{v})$ with respect to $\mathbf{x}$ yields

$$\mathbf{x} = \left[\mathbf{H}^\mathsf{T}\mathbf{H} + 2\lambda_0\,\mathbf{\Gamma}(\mathbf{v}) + \sum_{i=1}^{M}\lambda_i \mathbf{D}_i^\mathsf{T}\left[\Lambda(\mathbf{D}_i\mathbf{v})\right]\mathbf{D}_i\right]^{-1}\left(\mathbf{H}^\mathsf{T}\mathbf{H}\mathbf{y} - \lambda_0\mathbf{b}\right).$$

with notations

$$c(\mathbf{v}) = \sum_n \left[\phi(v_n) - \frac{v_n}{2}\phi'(v_n)\right].$$

BEADS Algorithm

We now have a majorizer for F

$$G(\mathbf{x},\mathbf{v}) = \frac{1}{2}\|\mathbf{H}(\mathbf{y}-\mathbf{x})\|_2^2 + \lambda_0 \mathbf{x}^\mathsf{T}[\boldsymbol{\Gamma}(\mathbf{v})]\mathbf{x} + \lambda_0 \mathbf{b}^\mathsf{T}\mathbf{x} + \sum_{i=1}^{M}\left[\frac{\lambda_i}{2}(\mathbf{D}_i\mathbf{x})^\mathsf{T}\left[\Lambda(\mathbf{D}_i\mathbf{v})\right](\mathbf{D}_i\mathbf{x})\right] + c(\mathbf{v}).$$

Minimizing $G(\mathbf{x},\mathbf{v})$ with respect to $\mathbf{x}$ yields

$$\mathbf{x} = \left[\mathbf{H}^\mathsf{T}\mathbf{H} + 2\lambda_0\,\boldsymbol{\Gamma}(\mathbf{v}) + \sum_{i=1}^{M}\lambda_i \mathbf{D}_i^\mathsf{T}\left[\Lambda(\mathbf{D}_i\mathbf{v})\right]\mathbf{D}_i\right]^{-1}\left(\mathbf{H}^\mathsf{T}\mathbf{H}\mathbf{y} - \lambda_0\mathbf{b}\right).$$

with notations

$$[\boldsymbol{\Gamma}(\mathbf{v})]_{n,n} = \begin{cases} \frac{1+r}{4|v_n|}, & |v_n| \geqslant \epsilon \\ \frac{1+r}{4\epsilon}, & |v_n| \leqslant \epsilon \end{cases}$$

BEADS Algorithm

We now have a majorizer for F

$$G(\mathbf{x}, \mathbf{v}) = \frac{1}{2}\|\mathbf{H}(\mathbf{y}-\mathbf{x})\|_2^2 + \lambda_0 \mathbf{x}^\mathsf{T}[\boldsymbol{\Gamma}(\mathbf{v})]\mathbf{x} + \lambda_0 \mathbf{b}^\mathsf{T}\mathbf{x} + \sum_{i=1}^{M}\left[\frac{\lambda_i}{2}(\mathbf{D}_i\mathbf{x})^\mathsf{T}\left[\Lambda(\mathbf{D}_i\mathbf{v})\right](\mathbf{D}_i\mathbf{x})\right] + c(\mathbf{v}).$$

Minimizing $G(\mathbf{x}, \mathbf{v})$ with respect to $\mathbf{x}$ yields

$$\mathbf{x} = \left[\mathbf{H}^\mathsf{T}\mathbf{H} + 2\lambda_0\,\boldsymbol{\Gamma}(\mathbf{v}) + \sum_{i=1}^{M}\lambda_i \mathbf{D}_i^\mathsf{T}\left[\Lambda(\mathbf{D}_i\mathbf{v})\right]\mathbf{D}_i\right]^{-1}\left(\mathbf{H}^\mathsf{T}\mathbf{H}\mathbf{y} - \lambda_0\mathbf{b}\right).$$

with notations

$$[\Lambda(\mathbf{v})]_{n,n} = \frac{\phi'(v_n)}{v_n}$$

BEADS Algorithm

We now have a majorizer for F

$$G(\mathbf{x}, \mathbf{v}) = \frac{1}{2}\|\mathbf{H}(\mathbf{y}-\mathbf{x})\|_2^2 + \lambda_0 \mathbf{x}^\mathsf{T}[\mathbf{\Gamma}(\mathbf{v})]\mathbf{x} + \lambda_0 \mathbf{b}^\mathsf{T}\mathbf{x} + \sum_{i=1}^{M}\left[\frac{\lambda_i}{2}(\mathbf{D}_i\mathbf{x})^\mathsf{T}\left[\Lambda(\mathbf{D}_i\mathbf{v})\right](\mathbf{D}_i\mathbf{x})\right] + c(\mathbf{v}).$$

Minimizing $G(\mathbf{x}, \mathbf{v})$ with respect to $\mathbf{x}$ yields

$$\mathbf{x} = \left[\mathbf{H}^\mathsf{T}\mathbf{H} + 2\lambda_0\, \mathbf{\Gamma}(\mathbf{v}) + \sum_{i=1}^{M}\lambda_i \mathbf{D}_i^\mathsf{T}\left[\Lambda(\mathbf{D}_i\mathbf{v})\right]\mathbf{D}_i\right]^{-1}\left(\mathbf{H}^\mathsf{T}\mathbf{H}\mathbf{y} - \lambda_0\mathbf{b}\right).$$

with notations

$$[\mathbf{b}]_n = \frac{1-r}{2}$$

BEADS Algorithm

Writing filter $\mathbf{H} = \mathbf{A}^{-1}\mathbf{B} \approx \mathbf{B}\mathbf{A}^{-1}$ (banded matrices) we have

$$\mathbf{x} = \mathbf{A}\mathbf{Q}^{-1}\left(\mathbf{B}^{\mathsf{T}}\mathbf{B}\mathbf{A}^{-1}\mathbf{y} - \lambda_0\mathbf{A}^{\mathsf{T}}\mathbf{b}\right)$$

where $\mathbf{Q}$ is the banded matrix,

$$\mathbf{Q} = \mathbf{B}^{\mathsf{T}}\mathbf{B} + \mathbf{A}^{\mathsf{T}}\mathbf{M}\mathbf{A},$$

and $\mathbf{M}$ is the banded matrix,

$$\mathbf{M} = 2\lambda_0\mathbf{\Gamma}(\mathbf{v}) + \sum_{i=1}^{M}\lambda_i\mathbf{D}_i^{\mathsf{T}}\left[\Lambda(\mathbf{D}_i\mathbf{v})\right]\mathbf{D}_i.$$

BEADS Algorithm

Using previous equations, the MM iteration takes the form:

$$\mathbf{M}^{(k)} = 2\lambda_0 \mathbf{\Gamma}(\mathbf{x}^{(k)}) + \sum_{i=1}^{M} \lambda_i \mathbf{D}_i^{\mathsf{T}} \left[\Lambda(\mathbf{D}_i \mathbf{x}^{(k)})\right] \mathbf{D}_i.$$

$$\mathbf{Q}^{(k)} = \mathbf{B}^{\mathsf{T}}\mathbf{B} + \mathbf{A}^{\mathsf{T}}\mathbf{M}^{(k)}\mathbf{A}$$

$$\mathbf{x}^{(k+1)} = \mathbf{A}[\mathbf{Q}^{(k)}]^{-1} \left(\mathbf{B}^{\mathsf{T}}\mathbf{B}\mathbf{A}^{-1}\mathbf{y} - \lambda_0 \mathbf{A}^{\mathsf{T}}\mathbf{b}\right)$$

BEADS Algorithm

Input: $\mathbf{y}$, $\mathbf{A}$, $\mathbf{B}$, λ_i, $i = 0, \ldots, M$

1. $\mathbf{b} = \mathbf{B}^{\mathsf{T}}\mathbf{B}\mathbf{A}^{-1}\mathbf{y}$
2. $\mathbf{x} = \mathbf{y}$ (Initialization)

Repeat

3. $$[\mathbf{\Lambda}_i]_{n,n} = \frac{\phi'([\mathbf{D}_i\mathbf{x}]_n)}{[\mathbf{D}_i\mathbf{x}]_n}, \quad i = 0, \ldots, M,$$
4. $$\mathbf{M} = \sum_{i=0}^{M} \lambda_i \mathbf{D}_i^{\mathsf{T}} \mathbf{\Lambda}_i \mathbf{D}_i$$
5. $$\mathbf{Q} = \mathbf{B}^{\mathsf{T}}\mathbf{B} + \mathbf{A}^{\mathsf{T}}\mathbf{M}\mathbf{A}$$
6. $$\mathbf{x} = \mathbf{A}\mathbf{Q}^{-1}\mathbf{b}$$

Until converged

8. $\mathbf{f} = \mathbf{y} - \mathbf{x} - \mathbf{B}\mathbf{A}^{-1}(\mathbf{y} - \mathbf{x})$

Output: $\mathbf{x}$, $\mathbf{f}$

Evaluation 1

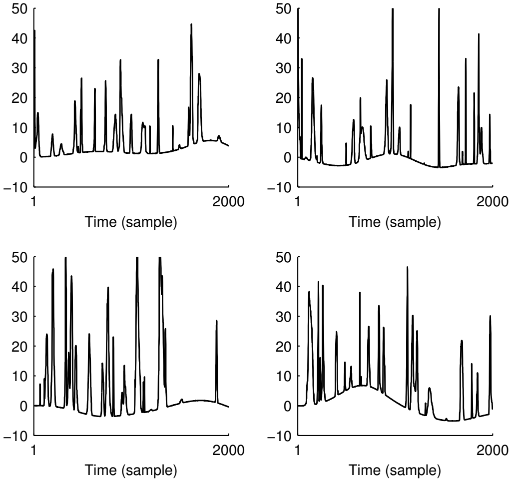
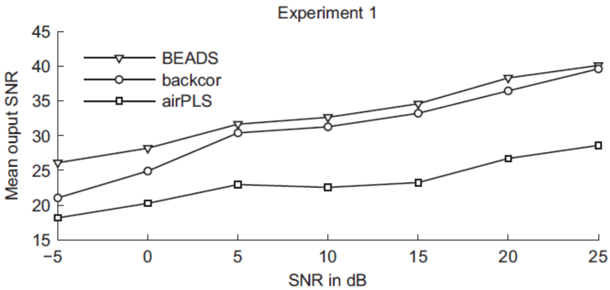


Figure : Simulated chromatograms w/ polynomial+sine baseline.

Evaluation 1 with Gaussian noise



	0 dB		10 dB		20 dB	
	Mean	Std	Mean	Std	Mean	Std
BEADS	28.1	8.52	32.64	8.02	38.33	6.74
backcor	24.91	9.75	31.27	8.33	36.47	6.53
airLPS	20.26	9.65	22.54	10.15	26.71	7.76

Evaluation 2

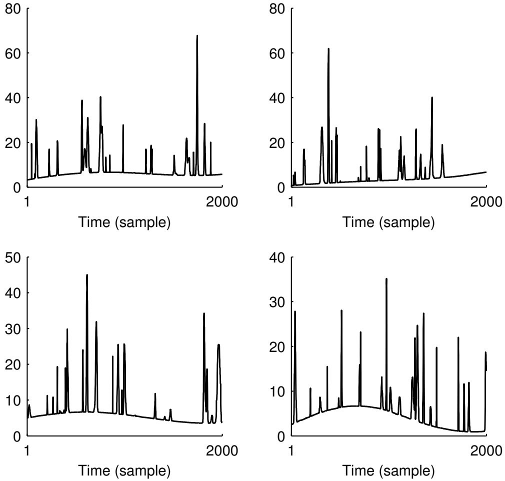
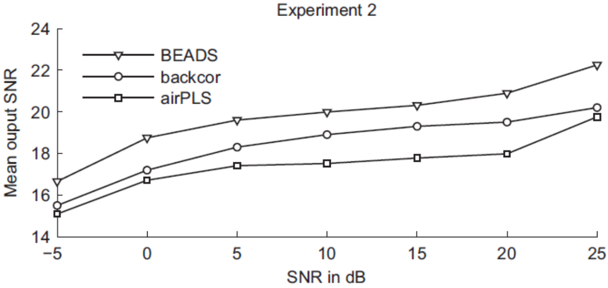


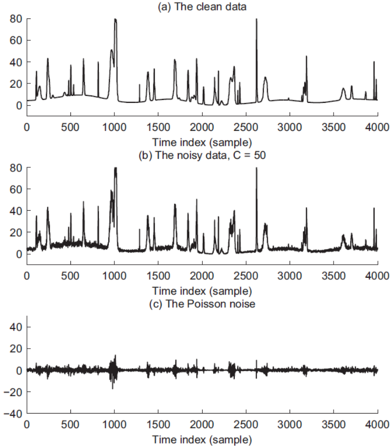
Figure : Simulated chromatograms w/ limited power spectrum noise.

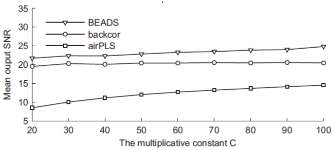
Evaluation 2 with Gaussian noise



	0 dB		10 dB		20 dB	
	Mean	Std	Mean	Std	Mean	Std
BEADS	18.75	3.71	19.99	3.17	20.89	3.32
backcor	17.20	4.57	18.93	3.74	19.54	3.18
airLPS	16.71	4.80	17.52	5.54	17.98	4.82

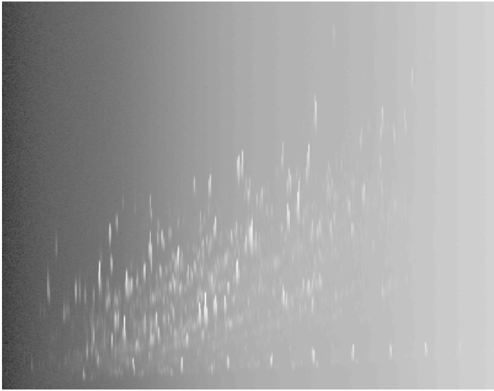
Evaluation 3 with Poisson noise





Two-dimensional chromatography data 1

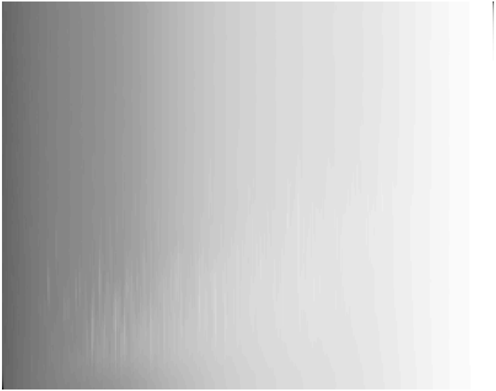
Hyphenated, two-dimensional gas chromatography data



Original data

Two-dimensional chromatography data 1

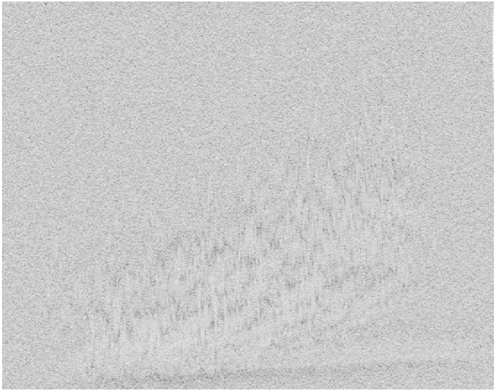
Hyphenated, two-dimensional gas chromatography data



2D background

Two-dimensional chromatography data 1

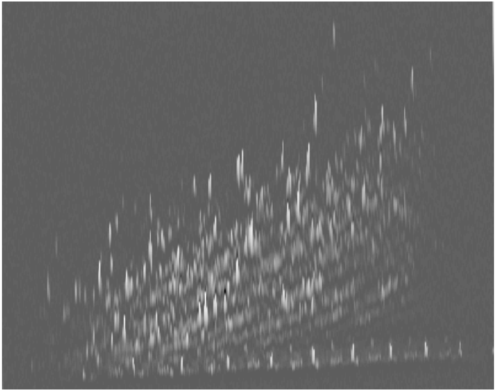
Hyphenated, two-dimensional gas chromatography data



Noise

Two-dimensional chromatography data 1

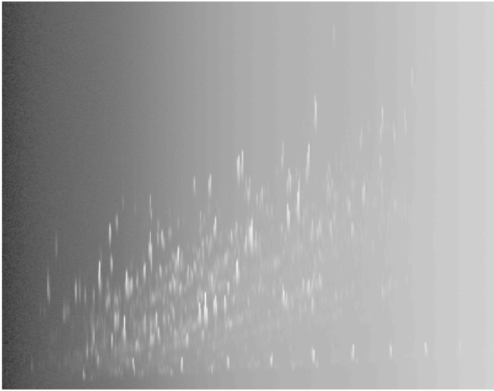
Hyphenated, two-dimensional gas chromatography data



Corrected

Two-dimensional chromatography data 1

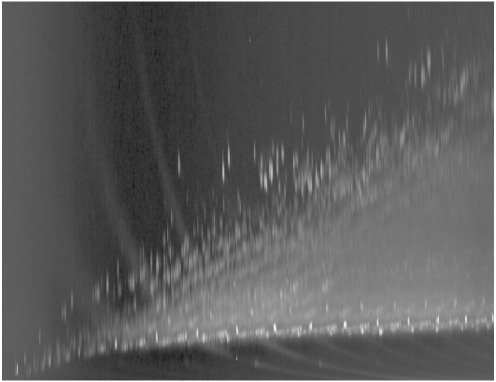
Hyphenated, two-dimensional gas chromatography data



Original data (again!)

Two-dimensional chromatography data 2

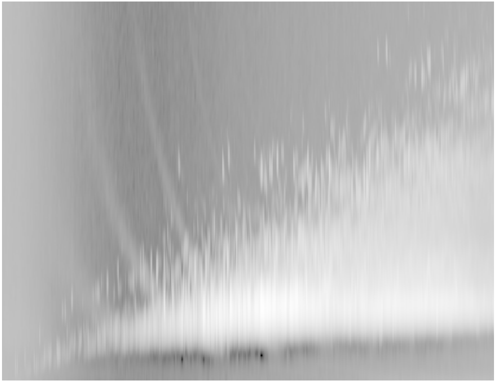
Hyphenated, two-dimensional gas chromatography data



Original data

Two-dimensional chromatography data 2

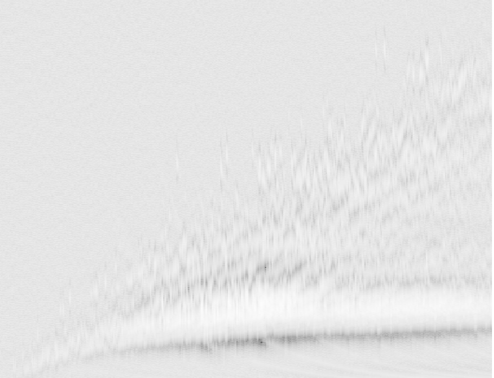
Hyphenated, two-dimensional gas chromatography data



2D background

Two-dimensional chromatography data 2

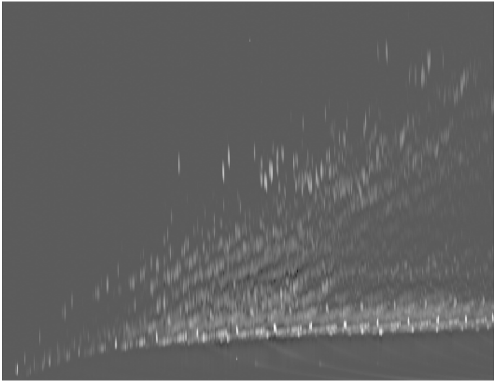
Hyphenated, two-dimensional gas chromatography data



Noise

Two-dimensional chromatography data 2

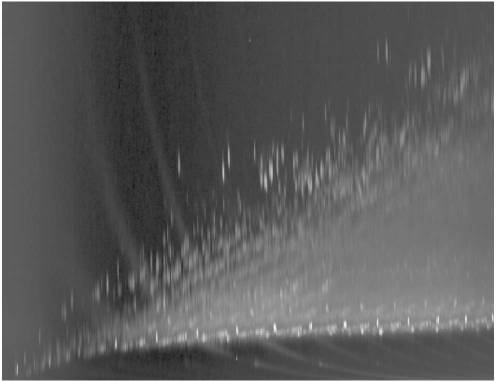
Hyphenated, two-dimensional gas chromatography data



Corrected

Two-dimensional chromatography data 2

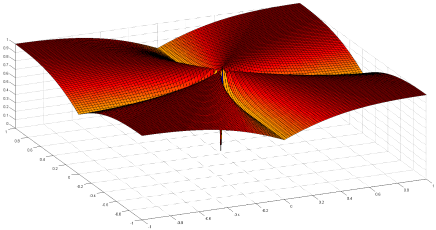
Hyphenated, two-dimensional gas chromatography data



Original data (again!)

Conclusions and work to come

- BEADS: Baseline Estimation And Denoising…Sparsely
- Asymmetric penalties with Majorization-Minimization
- Matlab toolbox: `http://lc.cx/beads`
- Important pre-processing for image alignment
- Further tests on other analytical chemistry signals for routine analysis
  - gas, liquid or ion chromatography; infrared, Raman, Nuclear Magnetic Resonance (NMR) spectroscopy; mass spectrometry
- Use closer to ℓ_0 sparse penalties: SOOT or smoothed ℓ_1/ℓ_2



More for free: additional references

V. Mazet, C. Carteret, D. Brie, J. Idier, and B. Humbert.
Background removal from spectra by designing and minimising a non-quadratic cost function.
Chemometr. Intell. Lab. Syst., 76(2):121–133, 2005.

C. Vendeuvre, R. Ruiz-Guerrero, F. Bertoncini, L. Duval, D. Thiébaut, and M.-C. Hennion.
Characterisation of middle-distillates by comprehensive two-dimensional gas chromatography (GC $\times$ GC): A powerful alternative for performing various standard analysis of middle-distillates.
J. Chrom. A, 1086(1-2):21–28, 2005.

C. Vendeuvre, R. Ruiz-Guerrero, F. Bertoncini, L. Duval, and D. Thiébaut.
Comprehensive two-dimensional gas chromatography for detailed characterisation of petroleum products.
Oil Gas Sci. Tech., 62(1):43–55, 2007.

X. Ning, I. W. Selesnick, and L. Duval.
Chromatogram baseline estimation and denoising using sparsity (BEADS).
Chemometr. Intell. Lab. Syst., 139:156–167, Dec. 2014.

A. Repetti, M. Q. Pham, L. Duval, E. Chouzenoux, and J.-C. Pesquet.
Euclid in a taxicab: Sparse blind deconvolution with smoothed ℓ_1/ℓ_2 regularization.
IEEE Signal Process. Lett., 22(5):539–543, May 2015.

C. Couprie, M. Moreaud, L. Duval, S. Henon and V. Souchon,
BARCHAN: Blob Alignment for Robust CHromatographic ANalysis.
J. Chrom. A, 2016...