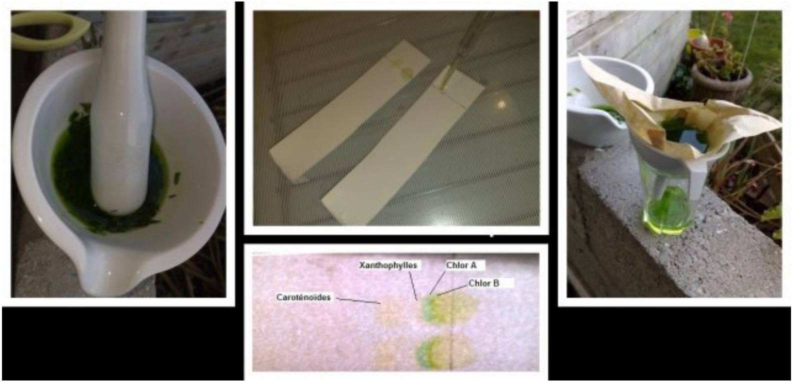
Pics, ligne de base, bruit : séparation ternaire de sources assistée (BEADS : positivité, parcimonie), spectres chimiques & miscellanées

L. DUVAL, A. PIRAYRE
IFP Energies nouvelles
X. NING, I. W. SELESNICK
Polytechnic School of Engineering, New York University

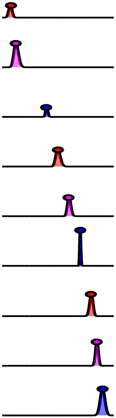
23 mars 2018

Old peaks cast long shadows



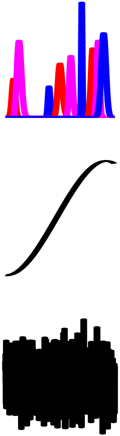
Chromatography: the traditional 2D way.

Old peaks cast long shadows



Chromatography: individual 1D peaks for single compounds

Old peaks cast long shadows

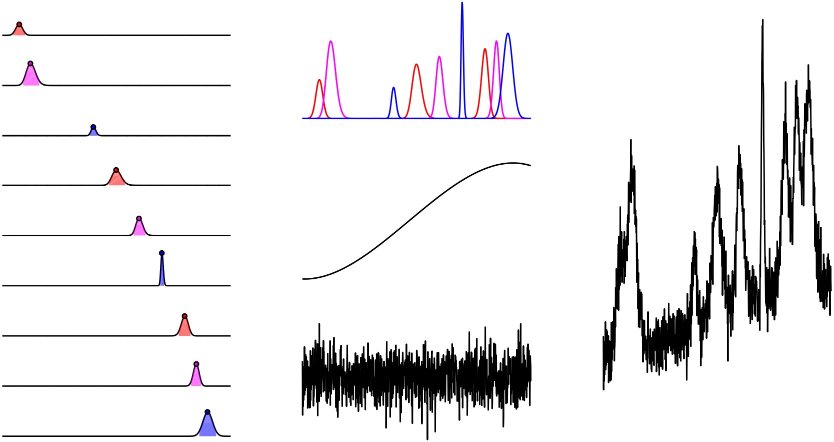


Chromatography: ternary sources separated

Old peaks cast long shadows

Chromatography: observed signal

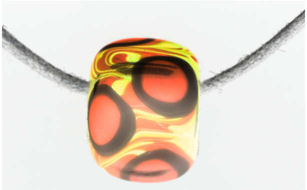
Old peaks cast long shadows



Chromatography: wrapping it up

The quick version

- *Issue*: how to accurately & repeatably quantize peaks?
  - avoiding separate baseline and noise removal
- *Question*: where is the string behind the bead?
  - without too accurate models for: peak, noise, baseline



- *Answer*: use main measurement properties + optimization
  - sparsity+symmetry, stationarity, smoothness
- BEADS: **B**aseline **E**stimation **A**nd **D**enoising w/ **S**parsity
  - other properties + optimization for further processing (BARCHAN)

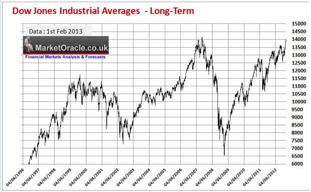
Outline

Background on background

Image processing: varying illumination

- Background affects quantitative evaluation/comparison
- In other domains: (instrumental) bias, (seasonal) trend
- In analytical chemistry: drift, continuum, wander, *baseline*
- Very rare cases of parametric modeling (piecewise linear, polynomial, spline)

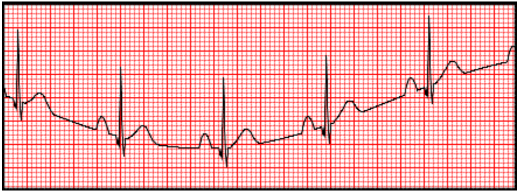
Background on background



Econometrics: trends and seasonality

- Background affects quantitative evaluation/comparison
- In other domains: (instrumental) bias, (seasonal) trend
- In analytical chemistry: drift, continuum, wander, *baseline*
- Very rare cases of parametric modeling (piecewise linear, polynomial, spline)

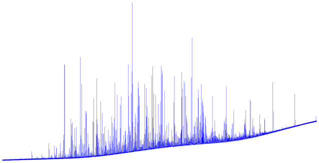
Background on background



Biomedical: ECG isoelectric line or baseline wander

- Background affects quantitative evaluation/comparison
- In other domains: (instrumental) bias, (seasonal) trend
- In analytical chemistry: drift, continuum, wander, *baseline*
- Very rare cases of parametric modeling (piecewise linear, polynomial, spline)

Background on background



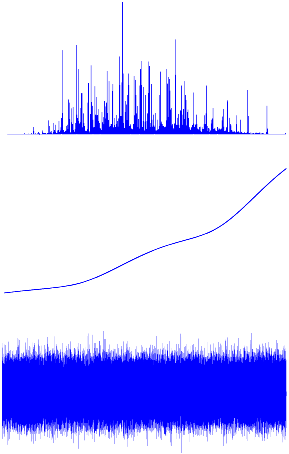
Gas chromatography: baseline

- Background affects quantitative evaluation/comparison
- In other domains: (instrumental) bias, (seasonal) trend
- In analytical chemistry: drift, continuum, wander, *baseline*
- Very rare cases of parametric modeling (piecewise linear, polynomial, spline)

Background on background

Analytical chemistry, biological data

- Signal separation into three main morphological components



Notations and assumptions

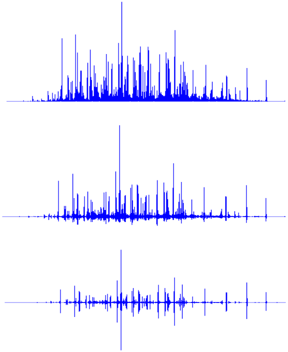
Morphological decomposition: $\mathbf{y} = \mathbf{x} + \mathbf{f} + \mathbf{w}$, signals in $\mathbb{R}^N$

- $\mathbf{y}$: observation (spectrum, analytical data)
- $\mathbf{x}$: clean series of peaks (no baseline, no noise)
- $\mathbf{f}$: baseline
- $\mathbf{w}$: noise

Assumption: without peaks, the baseline can be (approx.) recovered from noise-corrupted data by low-pass filtering

- $\hat{\mathbf{f}} = \mathbf{L}(\mathbf{y} - \hat{\mathbf{x}})$: $\mathbf{L}$: low-pass filter; $\mathbf{H} = \mathbf{I} - \mathbf{L}$: high-pass filter
- formulated as $\|\mathbf{y} - \hat{\mathbf{x}} - \hat{\mathbf{f}}\|_2^2 = \|\mathbf{H}(\mathbf{y} - \hat{\mathbf{x}})\|_2^2$
- Going further with $\mathbf{D}_i$: differentiation operators

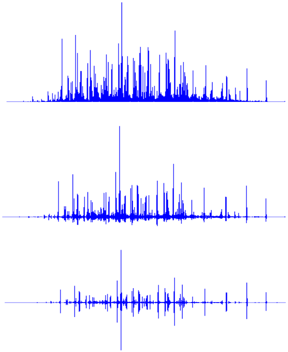
Compound sparse derivative modeling



An estimate $\hat{\mathbf{x}}$ can be obtained 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

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

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:

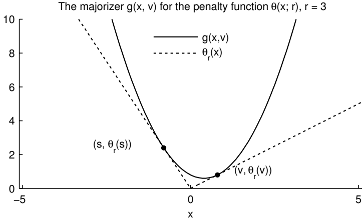
$$\theta(x;\, r) = \begin{cases} x, & x \geqslant 0 \\ -rx, & x < 0 \end{cases}$$

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

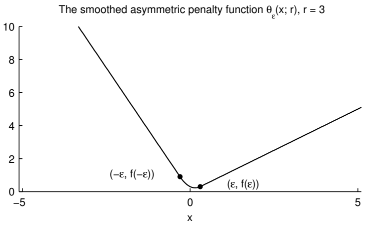


Compound sparse derivative modeling

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

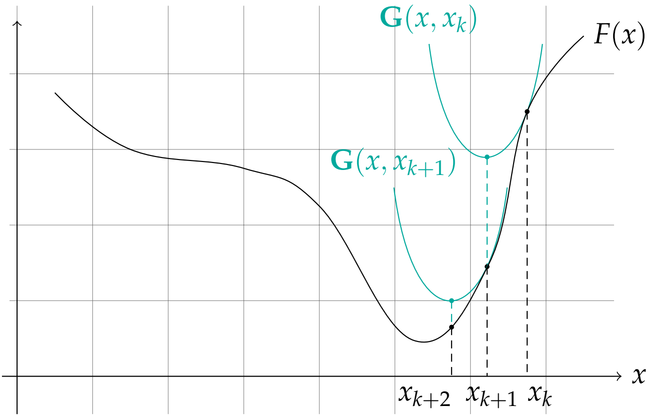
Take the positivity of chromatogram peaks into account:

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

then majorize it:

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

Overall principle for Majoration-Minimization



MM principles.

BEADS Algorithm (short)

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

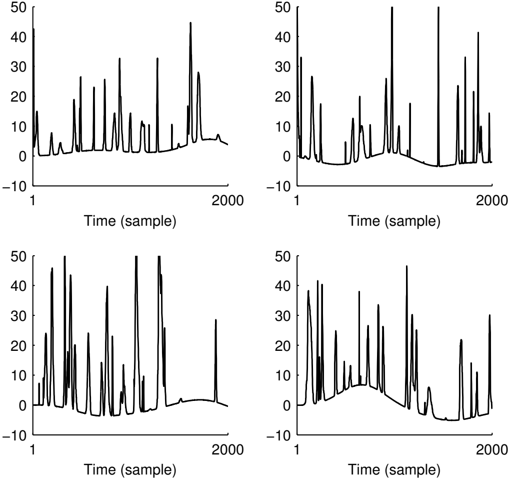
3. $$[\boldsymbol{\Lambda}_i]_{n,n} = \frac{\phi'([\mathbf{D}_i\mathbf{x}]_n)}{[\mathbf{D}_i\mathbf{x}]_n}, \quad i = 0, \ldots, M,$$
4. $$\mathbf{M} = \sum_{i=0}^{M} \lambda_i \mathbf{D}_i^\mathsf{T} \boldsymbol{\Lambda}_i \mathbf{D}_i$$
5. $$\mathbf{Q} = \mathbf{B}^\mathsf{T}\mathbf{B} + \mathbf{A}^\mathsf{T}\mathbf{M}\mathbf{A}$$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

Until converged

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

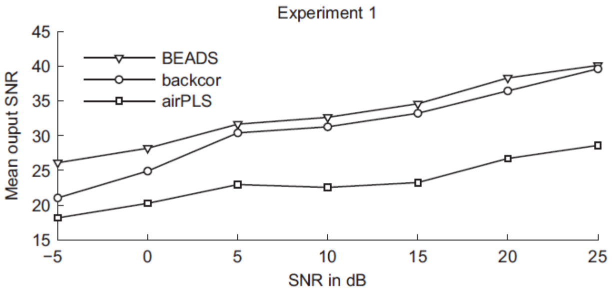
Output: $\mathbf{x}$, $\mathbf{f}$

Evaluation 1



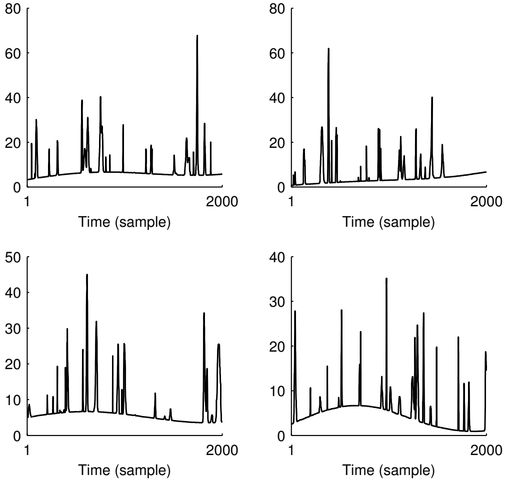
Simulated chromatograms w/ polynomial+sine baseline

Evaluation 1 with Gaussian noise



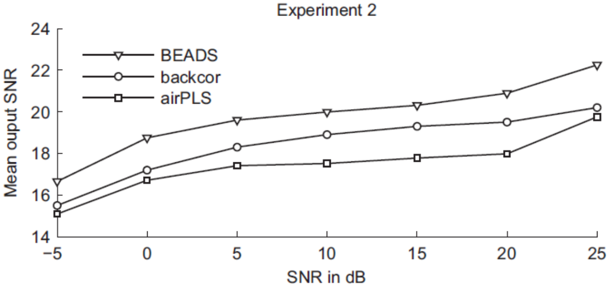
	0 dB		10 dB		20 dB	
	Mean	Std	Mean	Std	Mean	Std
BEADS	28.1	8.52	32.64	8.02	38.33	6.74
backcor	24.91	9.75	31.27	8.33	36.47	6.53
airLPS	20.26	9.65	22.54	10.15	26.71	7.76

Evaluation 2



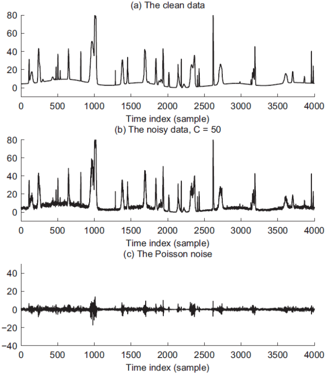
Simulated chromatograms w/ limited power spectrum noise

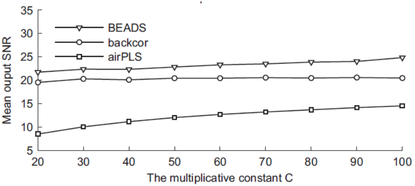
Evaluation 2 with Gaussian noise



	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

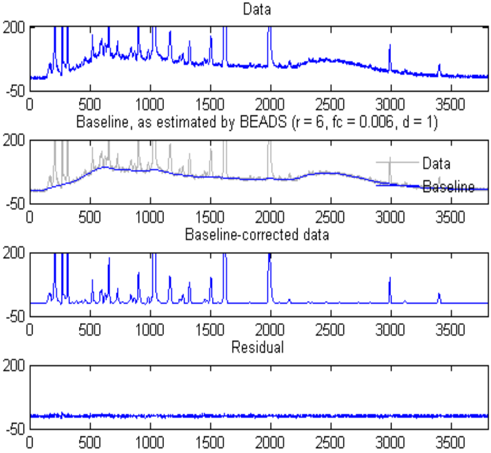
Evaluation 3 with Poisson noise





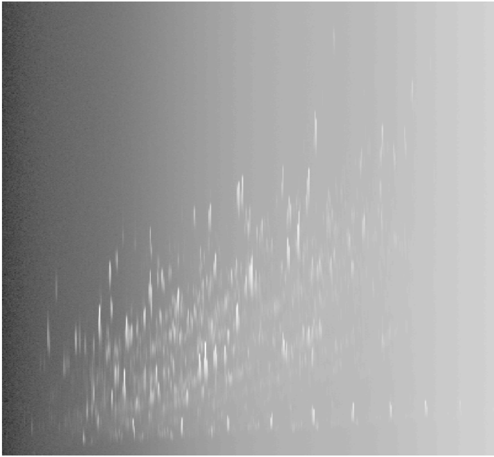
Simulated chromatograms w/ Poisson noise

Results: mono-dimensional chromatography (data 1)



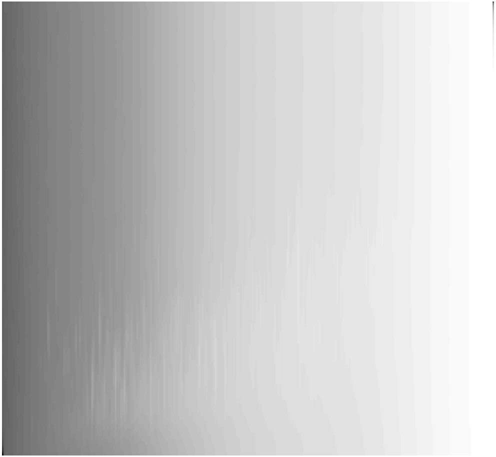
Original, superimposed, clean, noise

Results: two-dimensional chromatography (data 2)



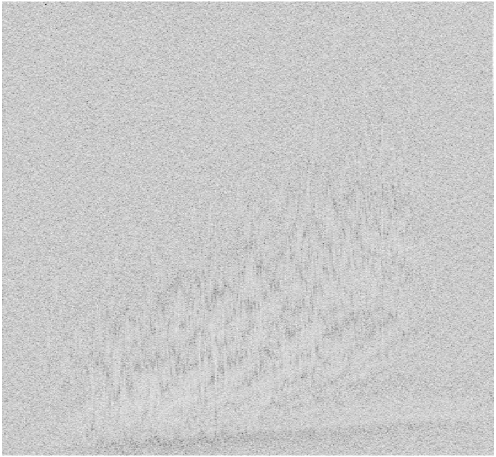
Original data

Results: two-dimensional chromatography (data 2)



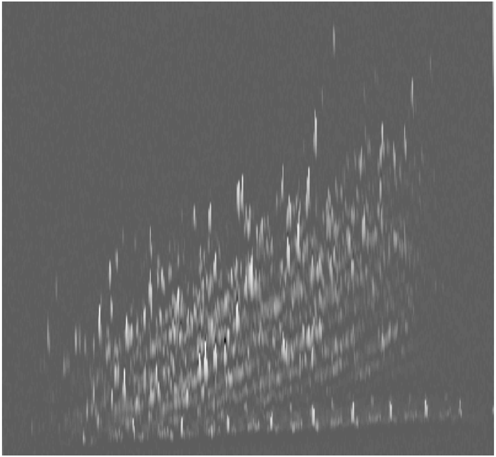
2D background (estimated)

Results: two-dimensional chromatography (data 2)



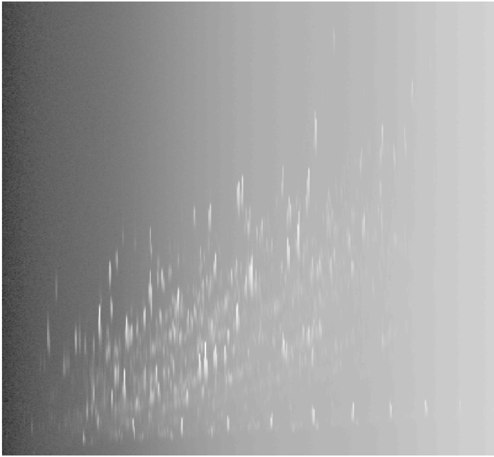
Noise (estimated)

Results: two-dimensional chromatography (data 2)



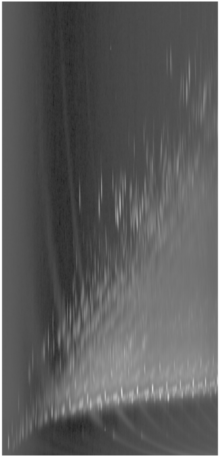
BEADS corrected data

Results: two-dimensional chromatography (data 2)



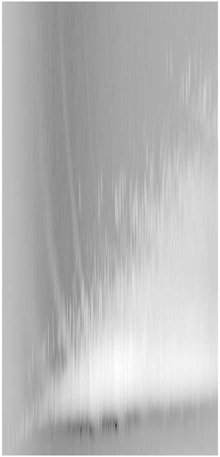
Original data (again!)

Results: two-dimensional chromatography (data 3)



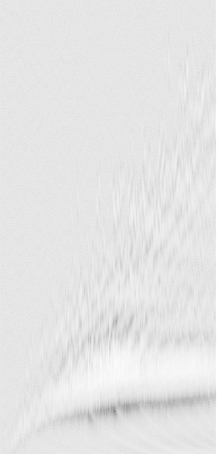
Original data

Results: two-dimensional chromatography (data 3)



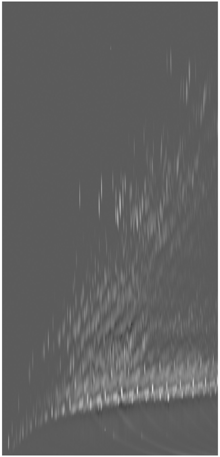
2D background (estimated)

Results: two-dimensional chromatography (data 3)



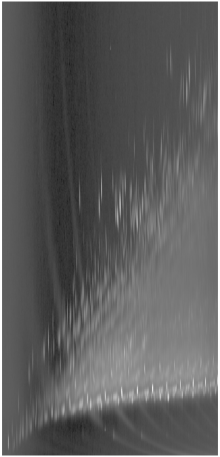
Noise (estimated)

Results: two-dimensional chromatography (data 3)



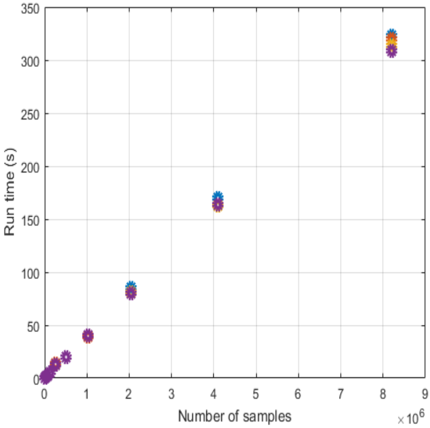
BEADS corrected data

Results: two-dimensional chromatography (data 3)



Original data (again!)

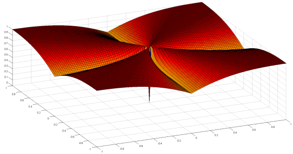
Results: computing scalability



Linear cost per sample (almost)

Ongoing work

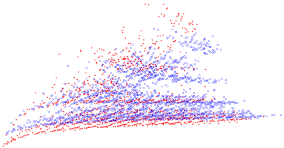
- Tests on analytical chemistry data: NIR, NMR, XPS
- Novel filtering: improved Savitzky-Golay filters
- Novel deconvolution: sparse & positive with norm ratios



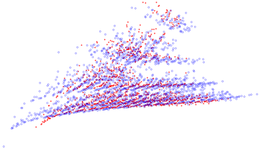
SOOT: Non-convex ℓ_0 count index approximation

- Novel metrics: errors related to peak quantities
- Baseline and noise use: uncertainty, trace products
- **2D chromatography comparisons: BARCHAN warping**
- **Improved usability: parameter estimation**

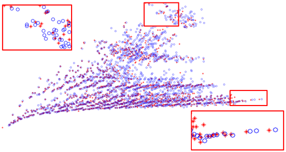
BARCHAN: 2D chromatography warping



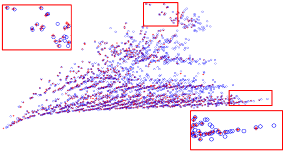
No transformation



Rigid



Non-rigid



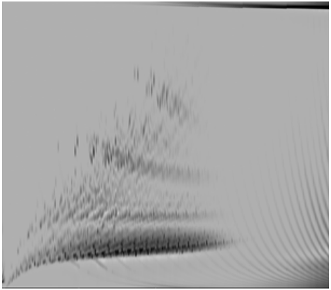
BARCHAN

Semi-rigid morphing of two different 2D chromatograms.

BARCHAN: 2D chromatography warping

Ingredients of a GMM plus EM optimization:

- Point sets $X = \{X_1, \dots, X_N\}$ and $Y = \{Y_1, \dots, Y_M\}$
- $p(X_n) = \frac{w}{N} + \sum_{m=1}^{M} \frac{1-w}{2M\pi\sigma^2} \exp\left(-\frac{\|X_n - T(Y_m)\|^2}{2\sigma^2}\right)$
- $\min_{\sigma,W,s,t} E = E_1(\sigma, W, s, t) + \frac{\lambda}{2} \operatorname{Tr}(W^\top G W)$

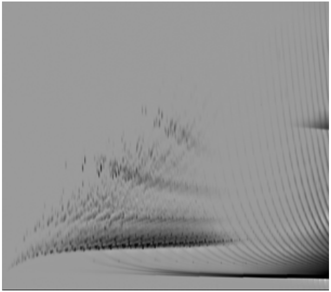


Calculated deformation of a 2D chromatogram with BARCHAN.

BARCHAN: 2D chromatography warping

Ingredients of a GMM plus EM optimization:

- Point sets $X = \{X_1, \ldots, X_N\}$ and $Y = \{Y_1, \ldots, Y_M\}$
- $p(X_n) = \frac{w}{N} + \sum_{m=1}^{M} \frac{1-w}{2M\pi\sigma^2} \exp\left(-\frac{\|X_n - T(Y_m)\|^2}{2\sigma^2}\right)$
- $\min_{\sigma,W,s,t} E = E_1(\sigma, W, s, t) + \frac{\lambda}{2} \mathrm{Tr}(W^\top G W)$

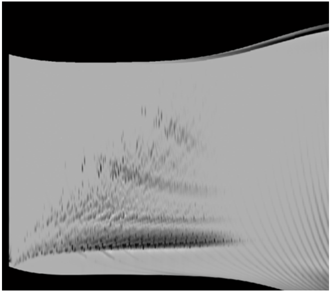


Calculated deformation of a 2D chromatogram with BARCHAN.

BARCHAN: 2D chromatography warping

Ingredients of a GMM plus EM optimization:

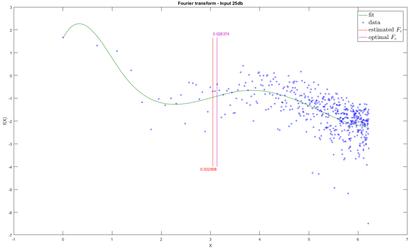
- Point sets $X = \{X_1, \ldots, X_N\}$ and $Y = \{Y_1, \ldots, Y_M\}$
- $p(X_n) = \frac{w}{N} + \sum_{m=1}^{M} \frac{1-w}{2M\pi\sigma^2} \exp\left(-\frac{\|X_n - T(Y_m)\|^2}{2\sigma^2}\right)$
- $\min_{\sigma,W,s,t} E = E_1(\sigma, W, s, t) + \frac{\lambda}{2}\operatorname{Tr}(W^\top G W)$

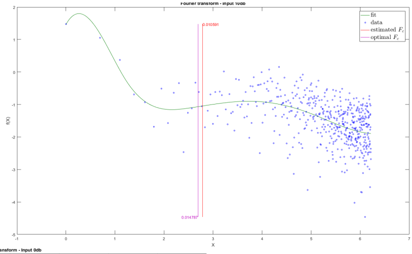


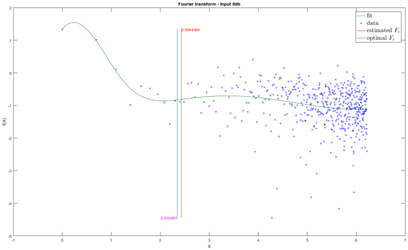
Calculated deformation of a 2D chromatogram with BARCHAN.

Improved usability: parameter estimation

- Cut-off frequency estimation

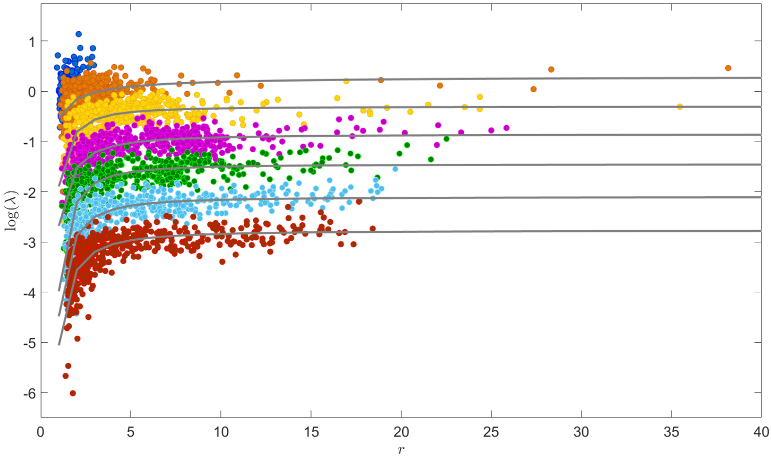






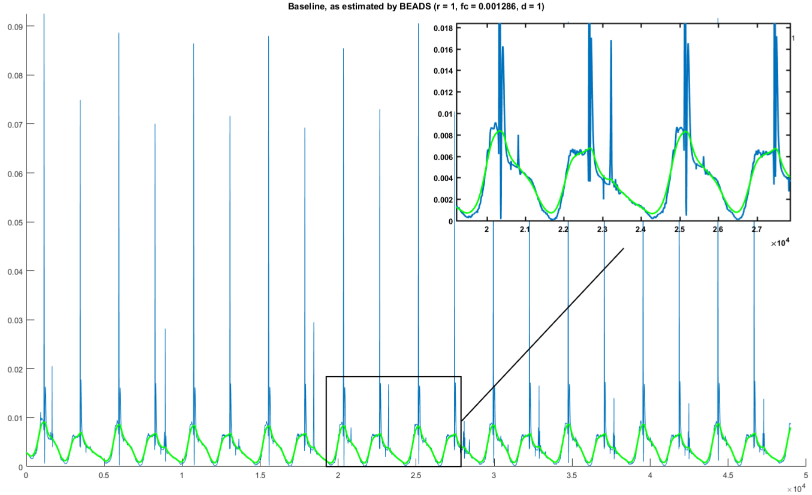
Improved usability: parameter estimation

- Noise, asymmetry (r) and regularization (λ)



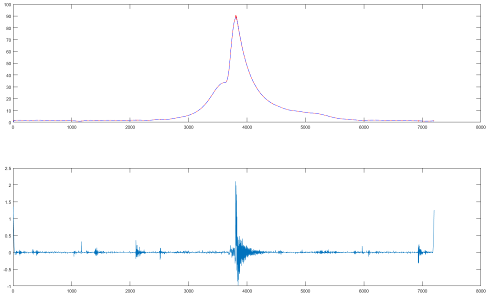
Extended applications

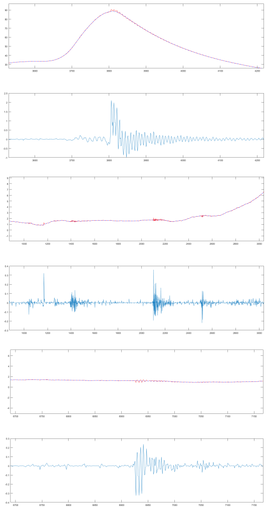
- Lidar application



Extended applications

- Engine knocking application





Other known uses

- A fairly generic model (sparsity, positivity/negativity), reused by other authors
  - gas chromatography: mono-dimensional and comprehensive/two-dimensional
  - Raman spectra: biological and biomedical
  - MUSE (Multi Unit Spectroscopic Explorer): astronomical hyperspectral galaxy spectrum
  - X-ray absorption spectroscopy (XAS), X-ray diffraction (XRD), and combined XAS/XRD
  - high-resolution mass spectrometry
  - postprandial Plasma Glucose (PPG), multichannel electroencephalogram (EEG) and single-channel electrocardiogram (ECG)
  - arabic characters

Other known uses

- A fairly generic model (sparsity, positivity/negativity), reused by other authors
  - gas chromatography: mono-dimensional and comprehensive/two-dimensional
  - Raman spectra: biological and biomedical
  - MUSE (Multi Unit Spectroscopic Explorer): astronomical hyperspectral galaxy spectrum
  - X-ray absorption spectroscopy (XAS), X-ray diffraction (XRD), and combined XAS/XRD
  - high-resolution mass spectrometry
  - postprandial Plasma Glucose (PPG), multichannel electroencephalogram (EEG) and single-channel electrocardiogram (ECG)
  - arabic characters

Conclusions

- Joint baseline/background and noise estimation
  - Interaction between "separative science" and "source separation"
  - Little "hard" modeling
  - Easy to tune, scalable
  - Codes in Matlab, R and C++[1]

- A wide range of applications to unveil

[1] http://www.laurent-duval.eu/siva-beads-baseline-background-removal-filtering-sparsity.html

A little more: additional references

C. Vendeuvre, F. Bertoncini, L. Duval, J.-L. Duplan, D. Thiébaut, and M.-C. Hennion.
Comparison of conventional gas chromatography and comprehensive two-dimensional gas chromatography for the detailed analysis of petrochemical samples, .
J. Chrom. A, 2004, http://dx.doi.org/10.1016/j.chroma.2004.05.071

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, 2005, http://dx.doi.org/10.1016/j.chroma.2005.05.106

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., 2007, http://dx.doi.org/10.2516/ogst:2007004

X. Ning, I. W. Selesnick, and L. Duval.
Chromatogram baseline estimation and denoising using sparsity (BEADS).
Chemometr. Intell. Lab. Syst., 2014, http://dx.doi.org/10.1016/j.chemolab.2014.09.014

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., 2015, http://dx.doi.org/10.1109/LSP.2014.2362861

C. Couprie, L. Duval, M. Moreaud, S. Hénon, M. Tebib, V. Souchon.
BARCHAN: Blob Alignment for Robust CHromatographic ANalysis.
Journal of Chromatography A., 2017, http://dx.doi.org/10.1016/j.chroma.2017.01.003

L. Duval, A. Pirayre and I. W. Selesnick.
Peaks, baseline and noise separation.
Chapter in preparation for *Source Separation in Physical-Chemical Sensing*, 2018.

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

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

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

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

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

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

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:

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

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

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