Bayesian calculation: A graduated non convexity (GNC) algorithm and its application in linear and nonlinear inverse problems.

Ali MOHAMMAD-DJAFARI*
and
Jérôme IDIER, Mila NIKOLOVA, Hervé CARFANTAN

Laboratoire des Signaux et Systèmes
Supélec, Plateau de Moulon
91192 Gif-sur-Yvette, FRANCE.

* 1997-1998: Visiting Associate Professor
Dept. of Electrical Eng., Univ. of Notre Dame
Notre Dame, IN 46556, USA.

Summary

1. Inverse problems
2. Bayesian estimation framework
3. MAP estimation for some special signal and image models
4. GNC algorithm
5. Linear problems with compound markov models
6. Nonlinear inverse scattering (diffraction tomography) problem
7. Conclusions, Discussions and Questions

Inverse problems

$$\mathcal{H} \quad (\quad \boldsymbol{y}, \qquad \boldsymbol{x}, \qquad \boldsymbol{z}, \qquad \boldsymbol{e} \quad) \quad = \quad 0$$

model	measurement (Data)	unknown quantities	other unknowns	errors and noise

Explicit relation: $\boldsymbol{y} = \mathcal{H}(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{e})$

Output error model: $\boldsymbol{y} = \mathcal{H}(\boldsymbol{x}, \boldsymbol{z}) \diamond \boldsymbol{e}$

Additive error model: $\boldsymbol{y} = \mathcal{H}(\boldsymbol{x}, \boldsymbol{z}) + \boldsymbol{e}$

Relation between $\boldsymbol{x}$ and $\boldsymbol{z}$: $\begin{cases} \boldsymbol{y} = \mathcal{H}_1(\boldsymbol{x}, \boldsymbol{z}) + \boldsymbol{e} \\ \mathcal{H}_2(\boldsymbol{x}, \boldsymbol{z}) = 0 \end{cases}$

Nonlinear+ additive errors: $\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x}) + \boldsymbol{e}$ or $y_i = h_i(\boldsymbol{x}) + e_i$

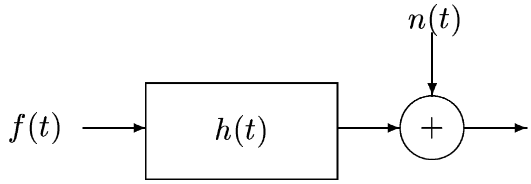
Linear + additive errors: $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{e}$ or $y_i = \sum_j h_{ij} x_j + e_i$

Examples of Linear inverse problems

$$g(\boldsymbol{s}_i) = \int f(\boldsymbol{r})\, h(\boldsymbol{r}, \boldsymbol{s}_i)\, \mathrm{d}\boldsymbol{r} + n(\boldsymbol{s}_i), \quad i = 1, \cdots, M$$

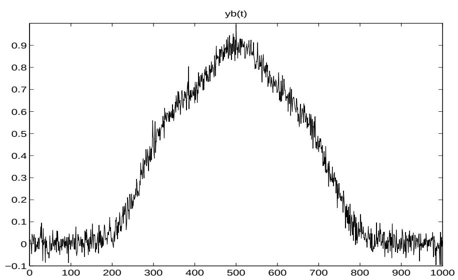
- Noise removing & Interpolation: $g(\boldsymbol{r}_i) = f(\boldsymbol{r}) + n(\boldsymbol{r}_i)$
- Deconvolution: $g(\boldsymbol{r}_i) = \int f(\boldsymbol{r}')\, h(\boldsymbol{r}_i - \boldsymbol{r}')\, \mathrm{d}\boldsymbol{r}' + n(\boldsymbol{r}_i)$
- Image restoration:
  $g(x_i, y_j) = \iint f(x', y')\, h(x_i - x', y_j - y')\, \mathrm{d}x'\, \mathrm{d}y' + n(x_i, y_j)$
- Image reconstruction:
  $g(r_i, \phi_j) = \iint f(x, y)\, \delta(r_i - x\cos\phi - y\sin\phi)\, \mathrm{d}x\, \mathrm{d}y + n(r_i, \phi_j)$
- Fourier synthesis:
  $g(\Omega_i, \phi_j) = \iint f(x, y) \exp\left[j(x\Omega_i \cos\phi_j + y\Omega_i \sin\phi_j\right]\, \mathrm{d}x\, \mathrm{d}y + n(\Omega_i, \phi_j)$

Deconvolution



$g(t) = h(t) * f(t) + n(t)$

Observation model



?

$\Longrightarrow$

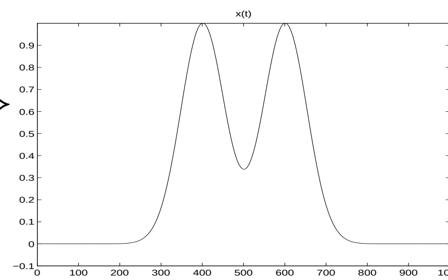
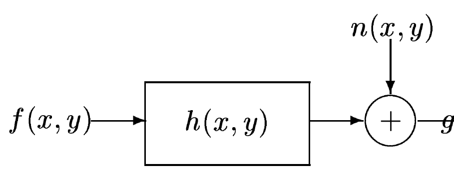
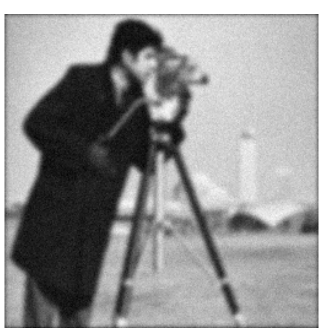


Image restoration



$g(x,y) = f(x,y) * h(x,y) + n(x,y)$

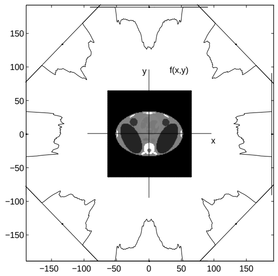
Observation model



?

$\Longrightarrow$

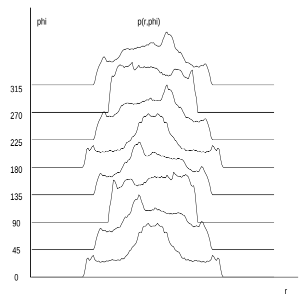
X ray computed tomography



$$f(x,y) \longrightarrow \boxed{\mathrm{TR}} \longrightarrow p(r,\phi)$$

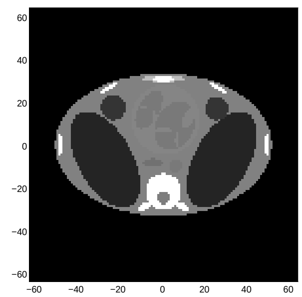
$$p(r,\phi) = \int_{L_{c,\phi}} f(x,y)\ \mathrm{d}l$$

$$p(r,\phi) = \iint_D f(x,y)\,\delta(r - x\cos\phi - y\sin\phi)\,\mathrm{d}x\,\mathrm{d}y$$



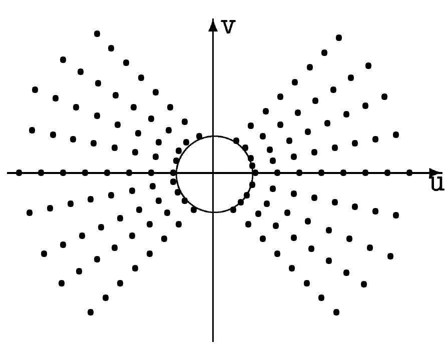
?

$\Longrightarrow$



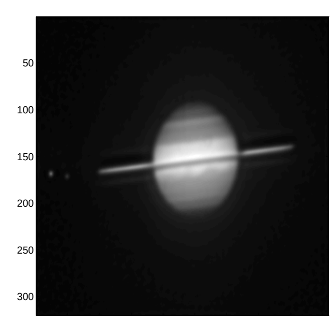
Fourier synthesis in radioastronomy imaging

$$g(u,v) = \iint_D f(x,y)\ \exp\left[-j(ux+vy)\right]\ \mathrm{d}x\,\mathrm{d}y + b(u,v),$$

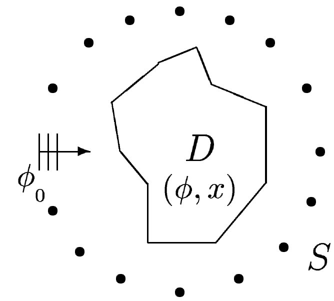


?

$\Longrightarrow$



Inverse scattering and diffraction tomography



Data: diffracted field $\phi_d(\boldsymbol{r}_i)$

Unknown quantity: $x(\boldsymbol{r}) = k_0^2(n^2(\boldsymbol{r}) - 1)$

Intermediate unknown: $\phi(\boldsymbol{r})$

$$
\begin{aligned}
\phi_d(\boldsymbol{r}_i) &= \iint_D G_m(\boldsymbol{r}_i, \boldsymbol{r}')\phi(\boldsymbol{r}')x(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}', \quad \boldsymbol{r}_i \in S \\
\phi(\boldsymbol{r}) &= \phi_0(\boldsymbol{r}) + \iint_D G_o(\boldsymbol{r}, \boldsymbol{r}')\phi(\boldsymbol{r}')x(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}', \quad \boldsymbol{r} \in D
\end{aligned}
$$

In discrete form:

$$
\begin{cases} \boldsymbol{\phi}_d = \boldsymbol{G}_M \boldsymbol{X} \boldsymbol{\phi} \\ \boldsymbol{\phi} = \boldsymbol{\phi}_0 + \boldsymbol{G}_O \boldsymbol{X} \boldsymbol{\phi} \end{cases} \longrightarrow \begin{cases} \boldsymbol{\phi}_d = \boldsymbol{G}_M \boldsymbol{X}(\boldsymbol{I} - \boldsymbol{G}_O \boldsymbol{X})^{-1}\boldsymbol{\phi}_0, \quad \boldsymbol{X} = \mathrm{diag}(\boldsymbol{x}) \\ \boldsymbol{\phi}_d = \boldsymbol{H}(\boldsymbol{x}) \text{ with } \boldsymbol{H}(\boldsymbol{x}) = \boldsymbol{G}_M \boldsymbol{X}(\boldsymbol{I} - \boldsymbol{G}_O \boldsymbol{X})^{-1}\boldsymbol{\phi}_0 \end{cases}
$$

Bayesian estimation approach

- Data observation and noise models: $\longrightarrow p(\boldsymbol{y}|\boldsymbol{x};\boldsymbol{\beta})$
- Prior information on $\boldsymbol{x}$: $\longrightarrow p(\boldsymbol{x}|\boldsymbol{\alpha})$
- Bayes:

$$
\pi(\boldsymbol{x}|\boldsymbol{y};\boldsymbol{\beta},\boldsymbol{\alpha}) = \frac{p(\boldsymbol{y}|\boldsymbol{x};\boldsymbol{\beta})\,p(\boldsymbol{x}|\boldsymbol{\alpha})}{m(\boldsymbol{y};\boldsymbol{\beta},\boldsymbol{\alpha})}
$$

$$
\text{where} \qquad m(\boldsymbol{y};\boldsymbol{\beta},\boldsymbol{\alpha}) = \int p(\boldsymbol{y}|\boldsymbol{x};\boldsymbol{\beta})\,p(\boldsymbol{x}|\boldsymbol{\alpha})\,\mathrm{d}\boldsymbol{x}
$$

- Inference or estimation rule: cost function $c(\boldsymbol{x}, \widehat{\boldsymbol{x}})$

$$
\widehat{\boldsymbol{x}}(\boldsymbol{y};\boldsymbol{\beta},\boldsymbol{\alpha}) = \arg\min_{\boldsymbol{z}} \left\{ \int c(\boldsymbol{x},\boldsymbol{z})\pi(\boldsymbol{x}|\boldsymbol{y};\boldsymbol{\beta},\boldsymbol{\alpha})\,\mathrm{d}\boldsymbol{x} \right\}
$$

Example:

Maximum A Posteriori (MAP):

$$
\widehat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \{-\ln \pi(\boldsymbol{x}|\boldsymbol{y};\boldsymbol{\beta},\boldsymbol{\alpha})\} = \arg\min_{\boldsymbol{x}} \{-\ln p(\boldsymbol{y}|\boldsymbol{x};\boldsymbol{\beta}) - \ln p(\boldsymbol{x}|\boldsymbol{\alpha})\}
$$

Point Estimators:

- Maximum *a posteriori* (MAP):

$$c(\boldsymbol{x}, \widehat{\boldsymbol{x}}) = 1 - \delta(\boldsymbol{x} - \widehat{\boldsymbol{x}}) \longrightarrow \widehat{\boldsymbol{x}} = \arg\max_{\boldsymbol{x}} \{\pi(\boldsymbol{x}|\boldsymbol{y}; \boldsymbol{\alpha}, \boldsymbol{\beta})\}$$

- Posterior mean (PM):

$$c(\boldsymbol{x}, \widehat{\boldsymbol{x}}) = [\boldsymbol{x} - \widehat{\boldsymbol{x}}]^t \boldsymbol{Q} [\boldsymbol{x} - \widehat{\boldsymbol{x}}]^t \longrightarrow \widehat{\boldsymbol{x}} = \mathrm{E}_{x|y}\{\boldsymbol{x}\} = \int \boldsymbol{x}\, \pi(\boldsymbol{x}|\boldsymbol{y}; \boldsymbol{\alpha}, \boldsymbol{\beta})\, \mathrm{d}\boldsymbol{x}$$

- Marginal MAP (MMAP):

$$c(\boldsymbol{x}, \widehat{\boldsymbol{x}}) = \prod_i 1 - \delta(x_i - \widehat{x}_i) \longrightarrow \widehat{x}_i = \arg\max_{x_i} \{\pi_i(x_i|\boldsymbol{y}; \boldsymbol{\alpha}, \boldsymbol{\beta})\},$$

where

$$\pi_i(x_i|\boldsymbol{y}; \boldsymbol{\alpha}, \boldsymbol{\beta}) = \int p_{x|y}(\boldsymbol{x}|\boldsymbol{y}; \boldsymbol{\alpha})\, \mathrm{d}x_1 \cdots \mathrm{d}x_{i-1}\, .\, \mathrm{d}x_{i+1} \cdots \mathrm{d}x_n$$

Case of linear inverse problems

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{e}$$

- Hypothesis on noise: $\boldsymbol{e} \sim \mathcal{N}(\boldsymbol{0}, 1/\beta \boldsymbol{I}) \longrightarrow \boldsymbol{y}|\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{H}\boldsymbol{x}, 1/\beta \boldsymbol{I})$

$$p(\boldsymbol{y}|\boldsymbol{x}, \beta) \propto \exp\left[-\frac{1}{2}\beta \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|^2\right]$$

- Hypothesis on $\boldsymbol{x}$: $\quad \boldsymbol{x} \sim \mathcal{N}(\boldsymbol{x}_0, 1/\alpha \boldsymbol{P}_0)$

$$p(\boldsymbol{x}|\alpha) \propto \exp\left[-\frac{1}{2}\alpha [\boldsymbol{x} - \boldsymbol{x}_0]^t \boldsymbol{P}_0^{-1} [\boldsymbol{x} - \boldsymbol{x}_0]\right]$$

- Posterior $\quad \pi(\boldsymbol{x}|\boldsymbol{y}) \sim \mathcal{N}(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{P}})$ with

$$\widehat{\boldsymbol{x}} = \widehat{\boldsymbol{P}} \boldsymbol{H}^t (\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}_0), \quad \widehat{\boldsymbol{P}} = \left(\boldsymbol{H}^t \boldsymbol{H} + \lambda \boldsymbol{P}_0^{-1}\right)^{-1}, \quad \lambda = \frac{\alpha}{\beta}$$

- MAP estimate:

$$\widehat{\boldsymbol{x}} = \arg\max_{\boldsymbol{x}} \{\pi(\boldsymbol{x}|\boldsymbol{y})\} = \arg\min_{\boldsymbol{x}} \{J(\boldsymbol{x}) = Q(\boldsymbol{x}) + \lambda \Phi(\boldsymbol{x})\} \text{ with}$$

$$Q(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|^2, \quad \Phi(\boldsymbol{x}) = \boldsymbol{x}^t \boldsymbol{P}_0^{-1} \boldsymbol{x} = \|\boldsymbol{D}\boldsymbol{x}\|^2, \quad \lambda = \frac{\alpha}{\beta}$$

$$J(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|^2 + \lambda \Phi_\alpha(\boldsymbol{x})$$

• Gamma law hypothesis for $\boldsymbol{x}$:

$$p(x_j) \propto (x_j/m_j)^\alpha \exp\left[-x_j/m_j\right] \longrightarrow \quad \Phi_\alpha(\boldsymbol{x}) = \alpha \sum_j \ln \frac{x_j}{m_j} + \frac{x_j}{m_j}$$

• Beta law hypothesis for $\boldsymbol{x}$:

$$p(x_j) \propto x_j^\alpha (1-x_j)^{1-\alpha} \longrightarrow \quad \Phi_\alpha(\boldsymbol{x}) = \alpha \sum_j \ln x_j + (1-\alpha) \sum_j \ln(1-x_j)$$

• Generalized Gaussian laws for $\boldsymbol{x}$:

$$p(x_j) \propto \exp\left[-\alpha |x_j|^p\right], \quad 1 < p < 2 \longrightarrow \quad \Phi_\alpha(\boldsymbol{x}) = \alpha \sum_j |x_j|^p$$

• Markovian models for $\boldsymbol{x}$:

$$p(x_j|\boldsymbol{x}) \propto \exp\left[-\alpha \sum_{i \in N_j} \phi(x_j, x_i)\right] \longrightarrow \quad \Phi_\alpha(\boldsymbol{x}) = \alpha \sum_j \sum_{i \in N_j} \phi(x_j, x_i)$$

General MAP estimate:

$$\widehat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \{J(\boldsymbol{x})\} \quad \text{with} \quad J(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|^2 + \lambda \Phi_\alpha(\boldsymbol{x})$$

• Gaussian laws: $\Phi_\alpha(\boldsymbol{x})$ quadratic
$\longrightarrow J(\boldsymbol{x})$ quadratic $\longrightarrow \widehat{\boldsymbol{x}}$ linear function of $\boldsymbol{y}$

• Entropic laws: $\Phi_\alpha(\boldsymbol{x})$ convex and separable: $\Phi(\boldsymbol{x}) = \sum_j \phi_\alpha(x_j)$
with $\quad \phi_\alpha(x_j) = \left\{\alpha x_j^p, \quad \alpha(x_j \ln x_j - x_j), \quad \alpha(\ln x_j - x_j)\right\}$
$\longrightarrow J(\boldsymbol{x})$ convex $\longrightarrow$ Nonlinear estimators but easy to calculate

• Markovian models: $\Phi_\alpha(\boldsymbol{x}) = \sum_j \sum_{i \in N_j} \phi_\alpha(x_j - x_i)$ with $\phi_\alpha(t) =$

$$\left\{\begin{array}{ll} |t|^2 & \text{if } |t| < \alpha, \\ \alpha^2 & \text{else,} \end{array}\right., \left\{\begin{array}{ll} t^2 & \text{if } |t| < \alpha, \\ 2\alpha t - \alpha^2 & \text{else,} \end{array}\right., \frac{\alpha^2 t^2}{1+t^2}, \quad \log\cosh(t/\alpha)$$

$\Phi_\alpha(\boldsymbol{x})$ non convex $\longrightarrow$ Nonlinear estimator and
need of global optimization

Open problems

- How to choose $p(\boldsymbol{x}|\boldsymbol{\alpha})$
  In the case of markovian models how to choose the potential functions
  $\longrightarrow$ Scale invariance (Djafari93, Brette *et al.* 93, 94)
- How to determine the hyperparameters $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\beta})$
  $\longrightarrow$ GMAP, MAPM, EM, SEM, ... (Djafari 95, Idier 94)
- Which estimator: MAP, PM ou MMAP
  $\longrightarrow$ Depending more on the calculation cost
- How to calculate effectively the solution $\widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\alpha}, \boldsymbol{\beta})$
  $\longrightarrow$ Optimization algorithmes:
  – Simulated annealing,
  – Deterministic relaxation (GNC) (Nikolova, Idier, Djafari 95)

Graduated non convexity (GNC)

Main idea:

MAP solution:

$$\widehat{\boldsymbol{x}} = \arg\max_{\boldsymbol{x}} \{\pi(\boldsymbol{x}|\boldsymbol{y})\} = \arg\min_{\boldsymbol{x}} \{J(\boldsymbol{x})\}$$

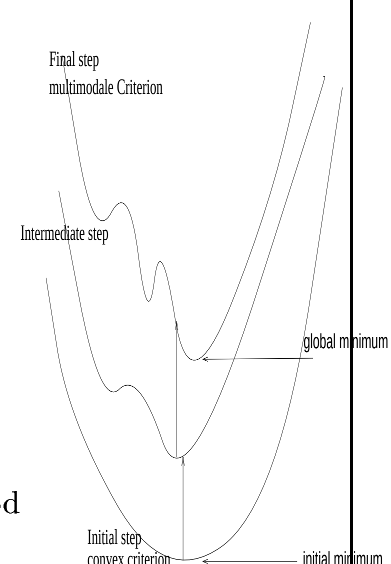
$J(\boldsymbol{x})$ multimodal

- Define a set of criteria $J_c(\boldsymbol{x})$ such that:

– $J_{c_0}(\boldsymbol{x})$ be convex

– $\forall \boldsymbol{x} \quad \lim_{c \longrightarrow c_\infty} J_c(\boldsymbol{x}) = J(\boldsymbol{x})$

- Minimize $J_{c_0}(\boldsymbol{x})$ to obtain $\boldsymbol{x}_0$
- For $c = \{c_1, \cdots, c_\infty\}$
  minimize localy $J_c(\boldsymbol{x})$ in the neighborhood of the precedent solution.



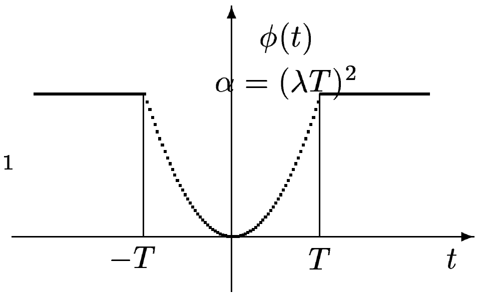
Linear inverse problems with piecewise Gaussian priors

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}$$

$$J(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 + \Omega(\boldsymbol{x})$$

with

$$\Omega(\boldsymbol{x}) = \textstyle\sum_j \phi(t_j), \quad t_j = x_j - x_{j-1}$$



$$\phi(t) = \begin{cases} (\lambda t)^2 & \text{if} \quad |t| < T \\ \alpha = (\lambda T)^2 & \text{if} \quad |t| > T \end{cases}$$

$$J_c(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 + \Omega_c(\boldsymbol{x}) \quad \text{with} \quad \Omega_c(\boldsymbol{x}) = \sum_j \phi_c(t_j)$$

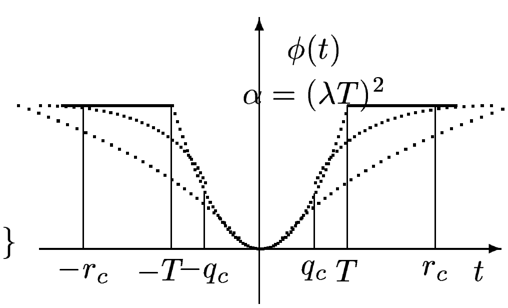
Originally developed for noise filtering and segmentation $\boldsymbol{A} = \boldsymbol{I}$ by Balke & Zisserman 1987

$$J_c(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 + \Omega_c(\boldsymbol{x}) \quad \text{with} \quad \Omega_c(\boldsymbol{x}) = \sum_j \phi_c(t_j)$$

$$\phi_c(t) = \begin{cases} (\lambda t)^2 & \text{if} \quad |t| < q_c \\ \alpha - \frac{1}{2}c(|t| - r_c)^2 & \text{if} \quad q_c > |t| > r_c \\ \alpha = (\lambda T)^2 & \text{if} \quad |t| > r_c \end{cases}$$

$$\begin{cases} q_c = T(1 + 2\lambda^2/c)^{-1/2} \\ r_c = T(1 + 2\lambda^2/c)^{1/2} \end{cases}$$

$$\widehat{\boldsymbol{x}}_{c_k} = \arg \min_{\boldsymbol{x} \in V(\widehat{\boldsymbol{x}}_{c_{k-1}})} \{J_{c_k}(\boldsymbol{x})\}$$



Extensions to general linear inverse problems
Initial convexity :

$$J_c(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 + \Omega_c(\boldsymbol{x})$$

- If $\boldsymbol{A}^t\boldsymbol{A}$ is full-rank $\longrightarrow \{J_c(\boldsymbol{x})\}$ has convex elements
  - μ_{min} smallest eigenvalue of $\boldsymbol{A}^t\boldsymbol{A}$
  - ν_{max} largest eigenvalue of $\boldsymbol{D}^t\boldsymbol{D}$
  - $J_c(\boldsymbol{x})$ is convex if $c < 2\mu_{min}/\nu_{max}$
- If $\boldsymbol{A}^t\boldsymbol{A}$ is singular
  (which is almost the case in Inverse problems)
  then $\{J_c(\boldsymbol{x})\}$ may not have any convex elements

$$J_c(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 + \Omega_{a,c}(\boldsymbol{x}) \quad \text{with} \quad \Omega_{a,c}(\boldsymbol{x}) = \sum_j \left[\phi_c(t_j) + a t_j^2\right]$$

Double relaxation scheme:

for fixed $c = c_0$ and for $a = a_0, \ldots, 0$ do:

$$\widehat{\boldsymbol{x}}_{a_k} = \arg \min_{\boldsymbol{x} \in V(\widehat{\boldsymbol{x}}_{a_{k-1}})} \{J_{a_k, c_0}(\boldsymbol{x})\}$$

for fixed $a = 0$ and for $c = c_0, \ldots, \infty$ do:

$$\widehat{\boldsymbol{x}}_{c_k} = \arg \min_{\boldsymbol{x} \in V(\widehat{\boldsymbol{x}}_{c_{k-1}})} \{J_{c_k, 0}(\boldsymbol{x})\}$$

Initial convexity is insured for $a_0 > \frac{c_0}{2}$.

Link with compound Markov modeling

$$(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{l}}) = \arg\max_{(\boldsymbol{x},\boldsymbol{l})} \{-\log p(\boldsymbol{x}, \boldsymbol{l}|\boldsymbol{y})\}$$

with the following prior laws:

$$\begin{aligned}
-\log p(\boldsymbol{y}|\boldsymbol{x}) &= \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 \\
-\log p(\boldsymbol{x}|\boldsymbol{l}) &= \sum_j \phi(t_j)(1 - l_j), \\
& \quad t_j = x_j - x_{j-1}, \quad \phi(t) = (\lambda t)^2 \\
-\log p(\boldsymbol{l}) &= \alpha \sum_j l_j
\end{aligned}$$

$$(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{l}}) = \arg\min_{(\boldsymbol{x},\boldsymbol{l})} \{J(\boldsymbol{x}, \boldsymbol{l})\}$$

with

$$J(\boldsymbol{x}, \boldsymbol{l}) = \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 + \sum_j \left[\lambda^2 (x_j - x_{j-1})^2 (1 - l_j) + \alpha l_j\right].$$

Line variables l_j are assumed non-interacting

The solution $(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{l}})$ is obtained equivalently by

$$\widehat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \left\{ J(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 + \sum_j \phi(x_j - x_{j-1}) \right\}$$

$$\widehat{l}_j = \begin{cases} 1 & \text{if} \quad |\widehat{x}_j - \widehat{x}_{j-1}| > T \\ 0 & \text{if} \quad |\widehat{x}_j - \widehat{x}_{j-1}| < T \end{cases}$$

with $\phi(t)$ the truncated quadratic function.

In 2D case:

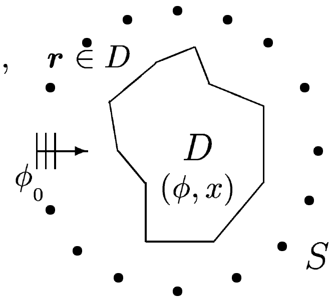
$$\sum_{ij} \phi(x_{i,j} - x_{i,j-1}) + \phi(x_{i,j} - x_{i-1,j})$$

Nonlinear inverse scattering

Diffraction tomography :

$$y(\boldsymbol{r}_i) = \iint_D G_m(\boldsymbol{r}_i, \boldsymbol{r}')\phi(\boldsymbol{r}')x(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}', \quad \boldsymbol{r}_i \in S$$

$$\phi(\boldsymbol{r}) = \phi_0(\boldsymbol{r}) + \iint_D G_o(\boldsymbol{r}, \boldsymbol{r}')\phi(\boldsymbol{r}')x(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}', \quad \boldsymbol{r} \in D$$



The discretized version:

$\boldsymbol{y} = \boldsymbol{G}_m \boldsymbol{X} \boldsymbol{\phi} + \boldsymbol{b}$

$\boldsymbol{\phi} = \boldsymbol{\phi}_0 + \boldsymbol{G}_o \boldsymbol{X} \boldsymbol{\phi}$

$\boldsymbol{y} = \boldsymbol{G}_m \boldsymbol{X} \left(\boldsymbol{I} - \boldsymbol{G}_o \boldsymbol{X}\right)^{-1} \boldsymbol{\phi}_0$

$$\boldsymbol{y} = \boldsymbol{A}(\boldsymbol{x}) + \boldsymbol{b} \quad \text{with} \quad \boldsymbol{A}(\boldsymbol{x}) = \boldsymbol{G}_m \boldsymbol{X} \left(\boldsymbol{I} - \boldsymbol{G}_o \boldsymbol{X}\right)^{-1} \boldsymbol{\phi}_0$$

$$J(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{A}(\boldsymbol{x})\|^2 + \Omega(\boldsymbol{x})$$

$$\boldsymbol{A}(\boldsymbol{x}) = \boldsymbol{G}_m \boldsymbol{X} \left(\boldsymbol{I} - \boldsymbol{G}_o \boldsymbol{X}\right)^{-1} \boldsymbol{\phi}_0$$

Proposed GNC scheme:

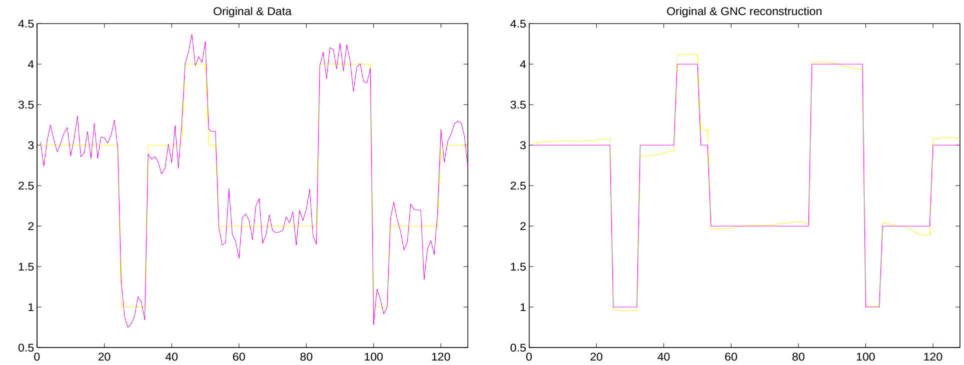
$$\boldsymbol{A}_{c_k}(\boldsymbol{x}) = \boldsymbol{G}_m \boldsymbol{X} (\boldsymbol{I} - c_k \boldsymbol{G}_o \boldsymbol{X})^{-1} \boldsymbol{\phi}_0$$

with $c_0 = 0$, and $\lim_{k \to \infty} c_k = 1$.

$c_0 = 0$:
$\boldsymbol{A}_0(\boldsymbol{x}) = \boldsymbol{G}_m \boldsymbol{X} \boldsymbol{\phi}_0 \longrightarrow \quad J_0(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{G}_m \boldsymbol{X} \boldsymbol{\phi}_0\|^2 + \Omega(\boldsymbol{x})$.
(Born approximation)

$c_k = 1$:
$\boldsymbol{A}_1(\boldsymbol{x}) = \boldsymbol{A}(\boldsymbol{x}) \longrightarrow \quad J_1(\boldsymbol{x}) = J(\boldsymbol{x})$

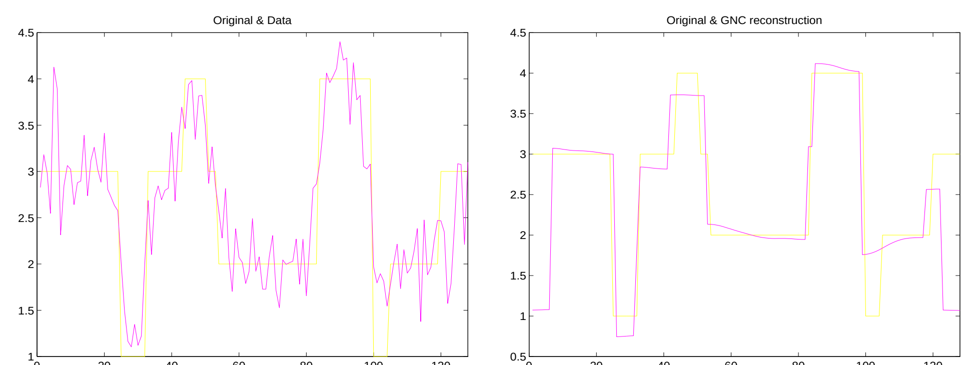
1-D noise filtering :



a b

a) Original and data, b) Piecewise Gauusian with GNC.

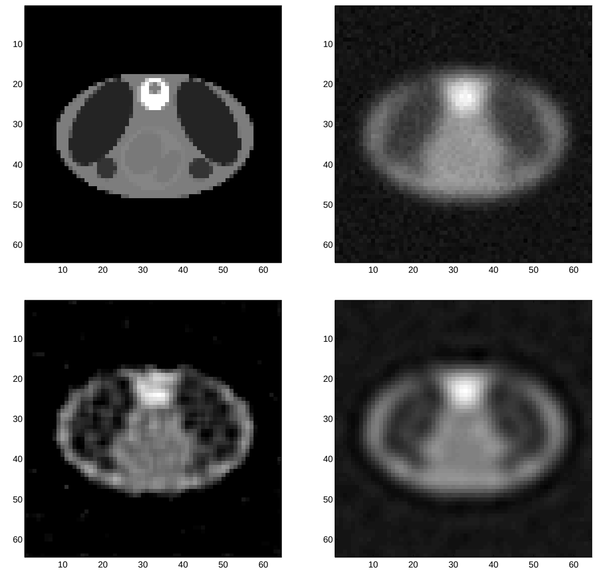
1-D signal deconvolution :



a b

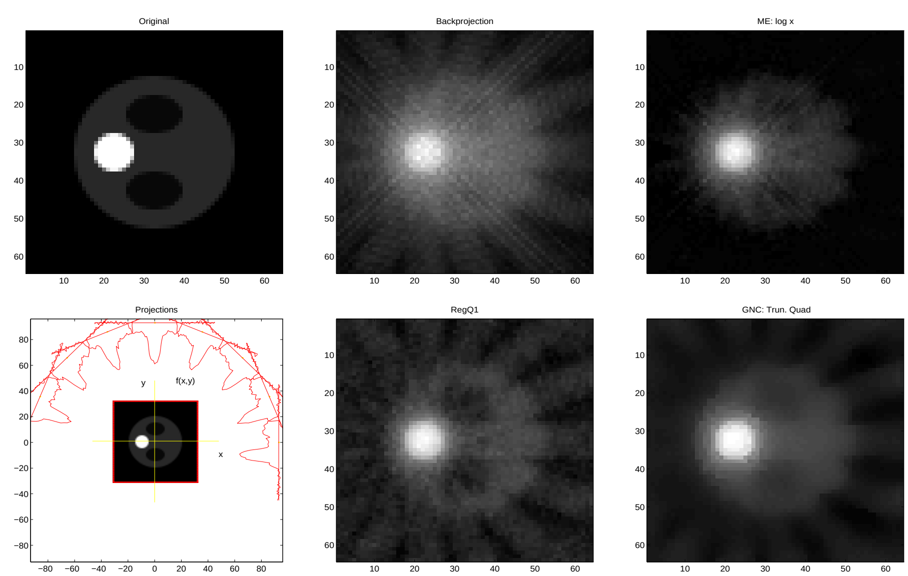
a) Original and data b) Piecewise Gauusian with GNC.

Image restoration :



a) original, b) data, c) Gaussian, d) Piecewise Gaussian and GNC

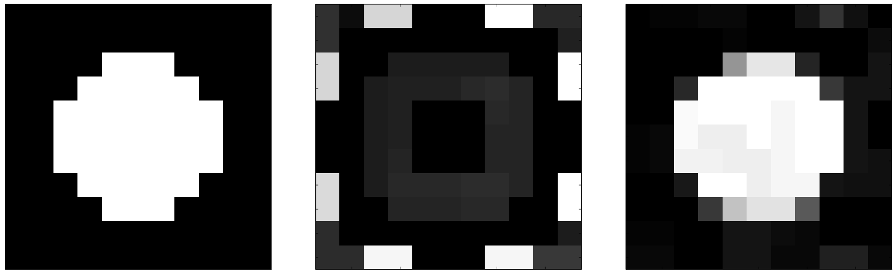
Image reconstruction in X-ray tomography :



a) original, c) Backprojection, e) Gamma prior,

b) projections (data), d) Gaussian, f) GNC reconstruction

Inverse scattering and diffraction tomography:



a b c

Diffraction tomography image reconstruction:
a) original,
b) linear Born approximation reconstruction,
c) GNC based reconstruction

Conclusions

- The Bayesian maximum *a posteriori* (MAP) estimates requires the minimization of a compound criterion.
- In many situations in inverse problems the criterion to be minimized is multimodal.
- The cost of the Simulated Annealing (SA) based techniques is in general huge for these problems.
- Graduated Non Convexity (GNC) has given satisfaction in two areas:

  - The linear inverse problems such as: noise filtering, deconvolution, image restoration, tomographic image reconstruction,..
  - A nonlinear inverse inverse scattering and diffraction tomography.