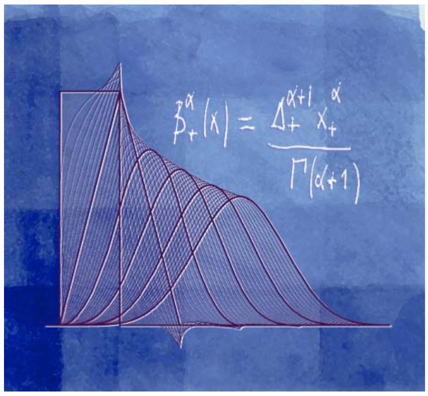
Hybrid sparse stochastic processes and the resolution of linear inverse problems

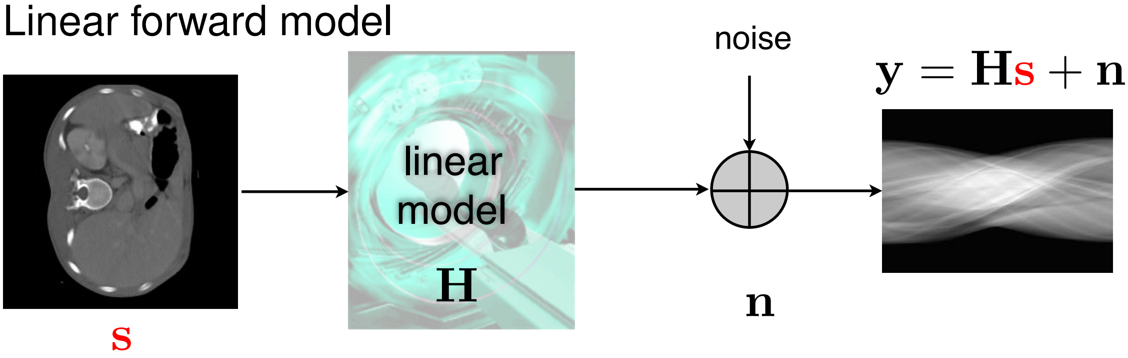
Michael Unser
Biomedical Imaging Group
EPFL, Lausanne, Switzerland



Statistical Models for Shape and Imaging, March 11-15, 2019, Institut Poincaré, Paris

Variational-MAP formulation of inverse problem

- Linear forward model



- Reconstruction as an optimization problem

$$\mathbf{s}_{\mathrm{rec}} = \arg\min \underbrace{\|\mathbf{y} - \mathbf{H}\mathbf{s}\|_2^2}_{\text{data consistency}} + \underbrace{\lambda\|\mathbf{L}\mathbf{s}\|_p^p}_{\text{regularization}}, \quad p = 1, 2$$

$$-\log \mathrm{Prob}(\mathbf{s})$$: prior likelihood

An introduction to sparse stochastic processes

Random spline: archetype of sparse signal

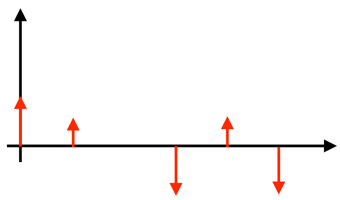
Random weights $\{a_n\}$ i.i.d. and random knots $\{t_n\}$ (Poisson with rate λ)

- Stochastic differential equation

$$\mathrm{D}s(t) = w(t)$$

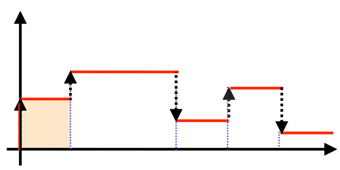
with boundary condition $s(0) = 0$

Innovation: $$w(t) = \sum_n a_n \delta(t - t_n)$$



- Formal solution = Compound Poisson process

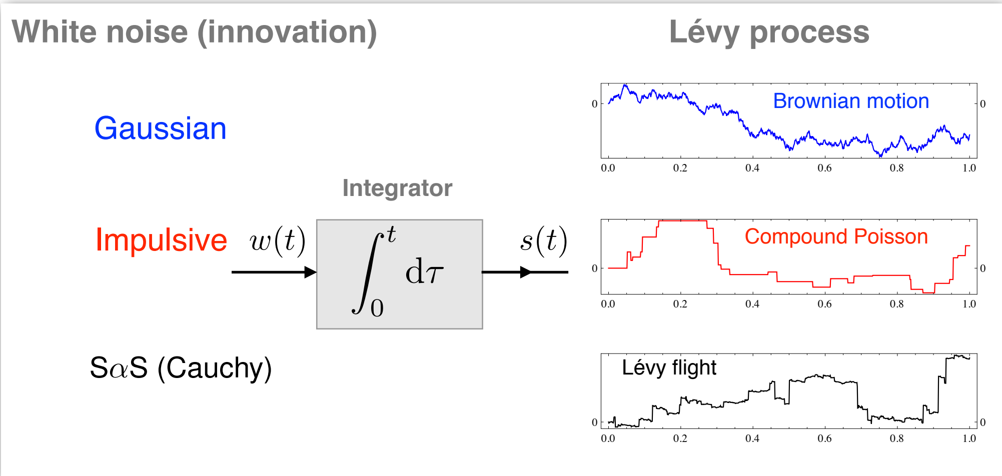
$$\begin{aligned} s(t) = \mathrm{D}^{-1}w(t) &= \sum_n a_n \mathrm{D}^{-1}\{\delta(\cdot - t_n)\}(t) \\ &= b_1 + \sum_n a_n \mathbb{1}_+(t - t_n) \end{aligned}$$



Lévy processes: all admissible brands of innovations

Generalized innovations : white Lévy noise with $\mathbb{E}\{w(t)w(t')\} = \sigma_w^2 \delta(t - t')$

$\mathrm{D}s = w$ (perfect decoupling!) $\in \mathcal{S}'(\mathbb{R}^d)$



(Wiener 1923)

(Paul Lévy circa 1930)

Generalized innovation model

Theoretical framework: Gelfand's theory of **generalized stochastic processes**

Generic test function $\varphi \in \mathcal{S}(\mathbb{R}^d)$ plays the role of index variable

Solution of SDE (general operator) $\mathrm{L}s = w$

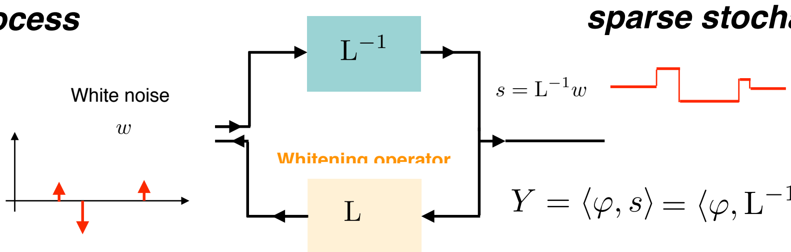
2

innovation process

sparse stochastic process

1

$X = \langle \varphi, w \rangle$



3

$Y = \langle \varphi, s \rangle = \langle \varphi, \mathrm{L}^{-1}w \rangle = \langle \mathrm{L}^{-1*}\varphi, w \rangle$

4

Approximate decoupling

Proper definition of ***continuous-domain*** white noise

(Unser et *al*, *IEEE-IT* 2014)

Regularization operator vs. wavelet analysis

Main feature: inherent sparsity
(few significant coefficients)

Description of sparse stochastic process

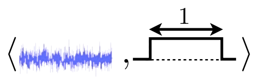
- Specification of spatial dependencies

Whitening operator L $\quad\Rightarrow\quad s = \mathrm{L}^{-1}w \qquad \in \mathcal{S}'(\mathbb{R}^d)$

- Specification of innovation (sparsity behavior)

Canonical observation through a rectangular window

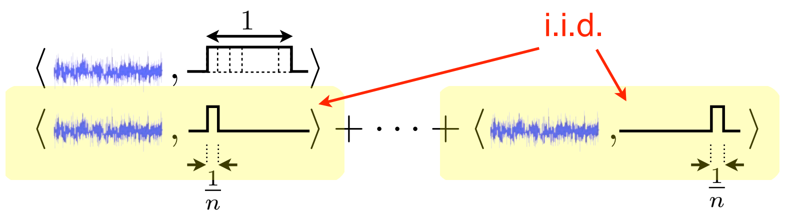
$X_{\mathrm{id}} = \langle w, \mathrm{rect}\rangle \;=$



w = white noise $\quad\Rightarrow\quad X_{\mathrm{id}} = \langle w, \mathrm{rect}\rangle$ is **infinitely divisible**
with **canonical Lévy exponent** $f(\omega) = \log \mathbb{E}\{\mathrm{e}^{\mathrm{j}\omega X_{\mathrm{id}}}\}$.

> **Definition**: A random variable X with generic pdf $p_{\mathrm{id}}(x)$ is **infinitely divisible** (id) iff., for any $N \in \mathbb{Z}^+$, there exist i.i.d. random variables $X_1, \dots, X_N$ such that $X \stackrel{\mathrm{d}}{=} X_1 + \dots + X_N$.

$X = \langle w, \mathrm{rect}\rangle \;=$



⇒ Probability laws of sparse processes are id

- Analysis: go back to **innovation process**: $w = \mathrm{L}s$
  - Generic random observation: $X = \langle \varphi, w\rangle$ with $\varphi \in \mathcal{S}(\mathbb{R}^d)$ or $\varphi \in L_p(\mathbb{R}^d)$ (by extension)
  - Linear functional: $Y = \langle \psi, s\rangle = \langle \psi, \mathrm{L}^{-1}w\rangle = \langle \overbrace{\mathrm{L}^{-1*}\psi}, \; w\rangle$

If $\phi = \mathrm{L}^{-1*}\psi \in L_p(\mathbb{R}^d)$ then $Y = \langle \psi, s\rangle = \langle \phi, w\rangle$ is ***infinitely divisible***
with (modified) **Lévy exponent** $f_\phi(\omega) = \int_{\mathbb{R}^d} f\big(\omega\phi(\boldsymbol{x})\big)\mathrm{d}\boldsymbol{x}$

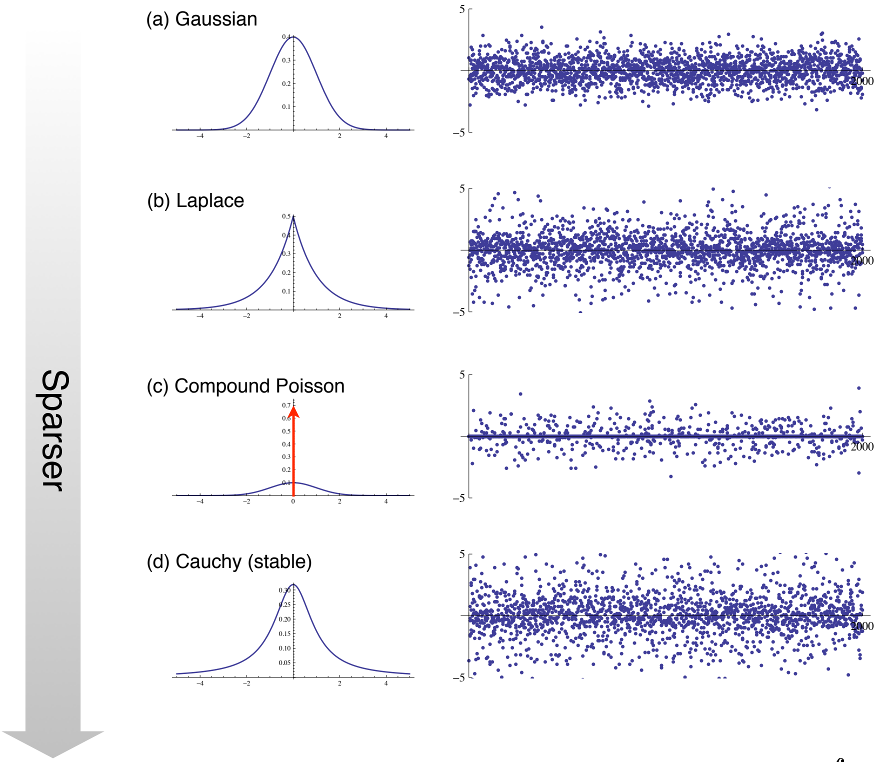
$$\Rightarrow \quad p_Y(y) = \mathcal{F}^{-1}\{\mathrm{e}^{f_\phi(\omega)}\}(y) = \int_{\mathbb{R}} \mathrm{e}^{f_\phi(\omega) - \mathrm{j}\omega y}\frac{\mathrm{d}\omega}{2\pi}$$

= explicit form of pdf

Unser and Tafti — An Introduction to Sparse Stochastic Processes — CAMBRIDGE

Examples of infinitely divisible laws

$p_{\rm id}(x)$



$$p_{\rm Gauss}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$$

$$p_{\rm Laplace}(x) = \frac{\lambda}{2} e^{-\lambda|x|}$$

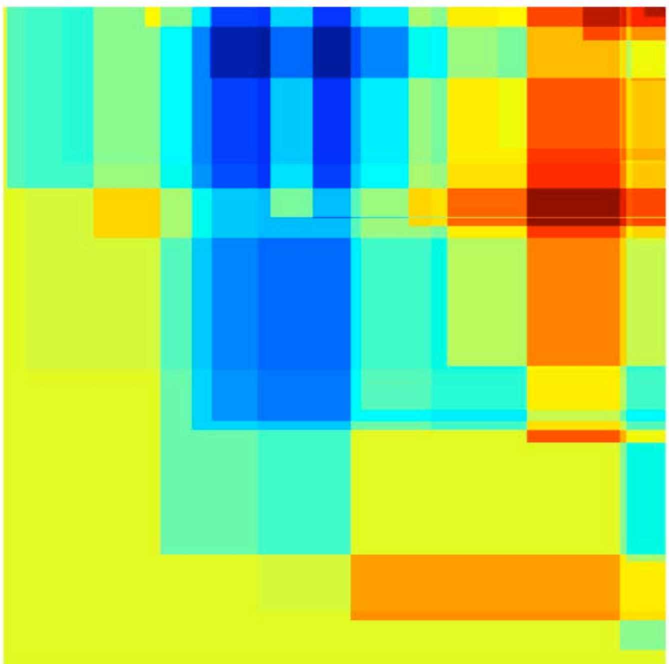
$$p_{\rm Poisson}(x) = \mathcal{F}^{-1}\{e^{\lambda(\hat{p}_A(\omega)-1)}\}$$

$$p_{\rm Cauchy}(x) = \frac{1}{\pi\,(x^2+1)}$$

Characteristic function: $\widehat{p}_{\rm id}(\omega) = \int_{\mathbb{R}} p_{\rm id}(x) e^{j\omega x} dx = e^{f(\omega)}$

Aesthetic sparse signal: the Mondrian process

$$\mathrm{L} = \mathrm{D}_x \mathrm{D}_y \quad \xleftrightarrow{\mathcal{F}} \quad (j\omega_x)(j\omega_y)$$



$\lambda = 30$

High-level properties of SSP

- **Infinite divisible probability laws**: broadest class of distributions preserved through linear transformation.
- **Explicit calculations**: Analytical determination of transform-domain statistics (including, joint pdfs).
- **Unifying framework**: includes all traditional families of stochastic processes (ARMA, fBm), as well as their non-Gaussian generalizations.
- **Sparsifying transforms / ICA**: SSP are (approximately) decoupled in a matched operator-like wavelet basis. (Pad-U., *IEEE-SP* 2015)
- N**-term approximation properties**: SSP are truly “sparse” as described by their inclusion in (weighted) Besov spaces. (Fageot et al., *ACHA* 2015)

Unser and Tafti | An Introduction to Sparse Stochastic Processes | Cambridge

OUTLINE

- Non-Gaussian statistical modeling ✔
  - Sparse stochastic processes
- Hybrid sparse stochastic processes
- Iterative image reconstruction (MAP formulation)

Hybrid stochastic processes

- Hybrid sparse processes: $s_{\mathrm{hyb}} = s_1 + \cdots + s_I$
- Description of independent components
  - Innovation model: $\mathrm{L}_i s_i = w_i \quad \Rightarrow \quad s_i = \mathrm{L}_i^{-1} w_i$
  - Whitening operator: $\mathrm{L}_i : \mathcal{S}'(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$
  - Lévy exponent of white noise: $f_i : \mathbb{R} \to \mathbb{C}$
  - Characteristic functionals $\quad \widehat{\mathscr{P}}_{w_i}, \widehat{\mathscr{P}}_{s_i} : \mathcal{S}(\mathbb{R}^d) \to \mathbb{C}$

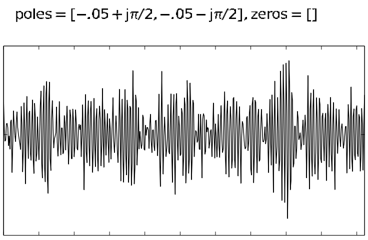
$$\widehat{\mathscr{P}}_{w_i}(\varphi) \triangleq \mathbb{E}\{e^{\mathrm{j}\langle w,\varphi\rangle}\} = \exp\left(\int_{\mathbb{R}^d} f_i(\varphi(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}\right)$$

$$\widehat{\mathscr{P}}_{s_i}(\varphi) = \mathbb{E}\{e^{\mathrm{j}\langle s_i,\varphi\rangle}\} = \mathbb{E}\{e^{\mathrm{j}\langle \mathrm{L}_i^{-1} w_i,\varphi\rangle}\} = \mathbb{E}\{e^{\mathrm{j}\langle w_i,\mathrm{L}_i^{-1*}\varphi\rangle}\} = \widehat{\mathscr{P}}_{w_i}(\mathrm{L}_i^{-1*}\varphi)$$

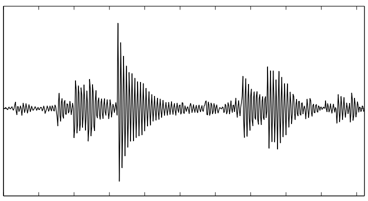
$$\widehat{\mathscr{P}}_{s_{\mathrm{hyd}}}(\varphi) = \prod_{i=1}^{I} \widehat{\mathscr{P}}_{s_i}(\varphi) = \exp\left(\int_{\mathbb{R}^d} \sum_{i=1}^{I} f_i(\mathrm{L}_i^{-1*}\varphi(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}\right)$$

Hybrid processes: Audio signal modeling

- Sparse, bandpass processes $\qquad \mathrm{L} = \frac{\mathrm{d}^n}{\mathrm{d}t^n} + a_{n-1}\frac{\mathrm{d}^{n-1}}{\mathrm{d}t^{n-1}} + \cdots + a_1 \frac{\mathrm{d}}{\mathrm{d}t} + a_0 \mathrm{I}$



(a) Gaussian


(b) Alpha stable α=1.2

- Hybrid sparse processes: $s_{\mathrm{hyb}} = s_1 + \cdots + s_I$

$$\widehat{\mathscr{P}}_{s_{\mathrm{hyd}}}(\varphi) = \prod_{i=1}^{I} \widehat{\mathscr{P}}_{s_i}(\varphi) = \exp\left(\int_{\mathbb{R}} \sum_{i=1}^{I} f_i(\mathrm{L}_i^{-1*}\varphi(t))\mathrm{d}t\right)$$

Gaussian (Am) $\qquad$ generalized Lévy (Am, SαS)

Discretization of linear inverse problem

$$s(\boldsymbol{r}) = \sum_{\boldsymbol{k}\in\Omega} s[\boldsymbol{k}]\beta_{\boldsymbol{k}}(\boldsymbol{r}) \qquad\qquad s(\boldsymbol{r}) = \sum_{i=1}^{I} s_i(\boldsymbol{r})\text{: hybrid sparse process}$$

Partial signals and innovation vectors: $\mathbf{s}_i = \big(s_i[\boldsymbol{k}]\big)_{\boldsymbol{k}\in\Omega}$ and $\mathbf{u}_i = \big(u[\boldsymbol{k}]\big)_{\boldsymbol{k}\in\Omega}$ of dimension K

- Discrete innovation model

$$\boxed{\mathbf{u}_i = \mathbf{L}_i\mathbf{s}_i} \qquad\qquad \mathbf{L}_i\text{: } (K\times K) \text{ matrix representation of } \mathrm{L}_{i,\mathrm{d}}$$

- Measurement model (image formation)

$$y_m = \int_{\mathbb{R}^d} s(\boldsymbol{r})\eta_m(\boldsymbol{r})\mathrm{d}\boldsymbol{r} + n[m] = \langle s, \eta_m\rangle + n[m], \quad (m = 1,\dots,M)$$

η_m: sampling/imaging function (mth detector)

$n[\cdot]$: additive i.i.d. noise with pdf p_N

$$\boxed{\mathbf{y} = \mathbf{y}_0 + \mathbf{n} = \mathbf{H}(\mathbf{s}_1 + \cdots + \mathbf{s}_I) + \mathbf{n}}$$

$(M\times K)$ system matrix : $\qquad [\mathbf{H}]_{m,\boldsymbol{k}} = \langle \eta_m, \beta_{\boldsymbol{k}}\rangle = \int_{\mathbb{R}^d} \eta_m(\boldsymbol{r})\beta_{\boldsymbol{k}}(\boldsymbol{r})\mathrm{d}\boldsymbol{r}$

Posterior probability distribution

$$\begin{aligned} p_{S|Y}(\mathbf{s}_1,\dots,\mathbf{s}_I|\mathbf{y}) &= \frac{p_{Y|S}(\mathbf{y}|\mathbf{s}_1,\dots,\mathbf{s}_I)\, p_S(\mathbf{s}_1,\dots,\mathbf{s}_I)}{p_Y(\mathbf{y})} \qquad \text{(Bayes' rule)}\\ &= \frac{p_N\big(\mathbf{y} - \mathbf{H}(\mathbf{s}_1+\cdots+\mathbf{s}_I)\big)\, p_S(\mathbf{s}_1,\dots,\mathbf{s}_I)}{p_Y(\mathbf{y})}\\ &= \frac{1}{Z} p_N\big(\mathbf{y} - \mathbf{H}(\mathbf{s}_1+\cdots+\mathbf{s}_I)\big)\, p_S(\mathbf{s}_1,\dots,\mathbf{s}_I) \end{aligned}$$

Independence of partial signals: $\quad p_S(\mathbf{s}_1,\dots,\mathbf{s}_I) = \prod_{i=1}^{I} p_S(\mathbf{s}_i)$

Innovation model: $\quad \mathbf{u}_i = \mathbf{L}_i\mathbf{s}_i \quad \Rightarrow \quad p_S(\mathbf{s}_i) \propto p_{U_i}(\mathbf{L}_i\mathbf{s}_i)$

$$\begin{aligned} p_{S|Y}(\mathbf{s}_1,\dots,\mathbf{s}_I|\mathbf{y}) &\propto p_N\big(\mathbf{y} - \mathbf{H}(\mathbf{s}_1+\cdots+\mathbf{s}_I)\big) \prod_{i=1}^{I} p_{U_i}(\mathbf{u}_i) \qquad \text{(decoupling simplication)}\\ &\approx p_N\big(\mathbf{y} - \mathbf{H}(\mathbf{s}_1+\cdots+\mathbf{s}_I)\big) \prod_{\boldsymbol{k}\in\Omega}\prod_{i=1}^{I} p_{U_i}\big([\mathbf{L}_i\mathbf{s}_i]_{\boldsymbol{k}}\big) \end{aligned}$$

Statement of MAP reconstruction problem

- Hypotheses

  $\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n}$ where $\mathbf{n}$ AWGN with variance σ^2

  $\mathbf{s} = \mathbf{s}_1 + \cdots + \mathbf{s}_I$ (sum of independent sparse processes)

  $\mathbf{L}_i \mathbf{s}_i = \mathbf{u}_i$: i.i.d. with pdf p_{U_i} and id potential function $\Phi_{U_i}(x) \triangleq -\log p_{U_i}(x)$

- Additive white Gaussian noise scenario (AWGN)

$$p_{S|Y}(\mathbf{s}_1, \dots, \mathbf{s}_I | \mathbf{y}) \propto \exp\left(-\frac{\|\mathbf{y} - \mathbf{H}(\mathbf{s}_1 + \cdots + \mathbf{s}_I)\|^2}{2\sigma^2}\right) \prod_{\boldsymbol{k}\in\Omega} \prod_{i=1}^{I} p_{U_i}([\mathbf{L}_i \mathbf{s}_i]_{\boldsymbol{k}})$$

- Maximum a posteriori (MAP) estimator

$$\arg \min_{\mathbf{s}_1,\dots,\mathbf{s}_I \in \mathbb{R}^K} \left(\frac{1}{2\sigma^2} \big\|\mathbf{y} - \mathbf{H}(\mathbf{s}_1 + \cdots + \mathbf{s}_I)\big\|_2^2 + \sum_{\boldsymbol{k}\in\Omega} \Phi_{U_1}([\mathbf{L}_1 \mathbf{s}_1]_{\boldsymbol{k}}) + \cdots + \sum_{\boldsymbol{k}\in\Omega} \Phi_{U_I}([\mathbf{L}_I \mathbf{s}_I]_{\boldsymbol{k}}) \right)$$

Hybrid model: reformulation

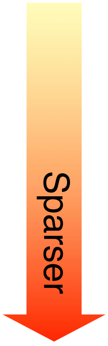
$$\arg \min_{\mathbf{s}_1,\dots,\mathbf{s}_I \in \mathbb{R}^K} \left(\frac{1}{2\sigma^2} \big\|\mathbf{y} - \mathbf{H}(\mathbf{s}_1 + \cdots + \mathbf{s}_I)\big\|_2^2 + \sum_{\boldsymbol{k}\in\Omega} \Phi_{U_1}([\mathbf{L}_1 \mathbf{s}_1]_{\boldsymbol{k}}) + \cdots + \sum_{\boldsymbol{k}\in\Omega} \Phi_{U_I}([\mathbf{L}_I \mathbf{s}_I]_{\boldsymbol{k}}) \right)$$

- Augmented formulation of the problem
  - **Unmixed** signal: $\mathbf{s} = (\mathbf{s}_1, \dots, \mathbf{s}_I) \in \mathbb{R}^{K \times I}$
  - Augmented **innovation**: $\mathbf{u} = (\mathbf{u}_1, \dots, \mathbf{u}_I) \in \mathbb{R}^{K \times I}$
  - Augmented system matrix: $\mathbf{H}_{\mathrm{aug}} = [\mathbf{H} \ \cdots \ \mathbf{H}] \in \mathbb{R}^{M \times KI}$
  - Augmented whitening operator: $\mathbf{L}_{\mathrm{aug}} = \mathrm{diag}(\mathbf{L}_1, \dots, \mathbf{L}_I) \in \mathbb{R}^{KI \times KI}$

$$\arg \min_{\mathbf{s} \in \mathbb{R}^{IK}} \left(\frac{1}{2\sigma^2} \big\|\mathbf{y} - \mathbf{H}_{\mathrm{aug}} \mathbf{s}\big\|_2^2 + \sum_n \Phi_{U_n}([\mathbf{u}]_n) \right) \quad \text{s.t.} \quad \mathbf{u} = \mathbf{L}_{\mathrm{aug}} \mathbf{s}$$

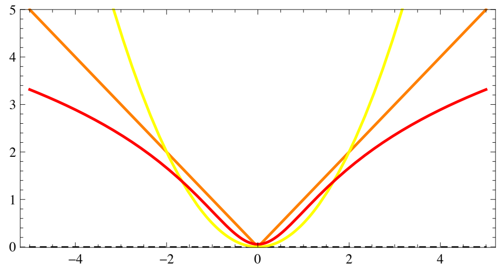
General form of MAP estimator (hybrid)

$$\mathbf{s}_{\mathrm{MAP}} = \operatorname{argmin}\left(\frac{1}{2}\|\mathbf{y}-\mathbf{H}_{\mathrm{aug}}\mathbf{s}\|_2^2 + \sigma^2 \sum_n \Phi_{U_n}([\mathbf{L}_{\mathrm{aug}}\mathbf{s}]_n)\right)$$



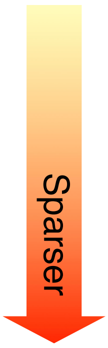
- Gaussian: $p_U(x) = \frac{1}{\sqrt{2\pi}\sigma_0} e^{-x^2/(2\sigma_0^2)}$ $\quad\Rightarrow\quad$ $\Phi_U(x) = \frac{1}{2\sigma_0^2}x^2 + C_1$
- Laplace: $p_U(x) = \frac{\lambda}{2} e^{-\lambda|x|}$ $\quad\Rightarrow\quad$ $\Phi_U(x) = \lambda|x| + C_2$
- Student: $p_U(x) = \dfrac{1}{B\left(r, \frac{1}{2}\right)} \left(\dfrac{1}{x^2+1}\right)^{r+\frac{1}{2}}$ $\quad\Rightarrow\quad$ $\Phi_U(x) = \left(r + \dfrac{1}{2}\right)\log(1+x^2) + C_3$

Potential: $\Phi_U(x) = -\log p_U(x)$



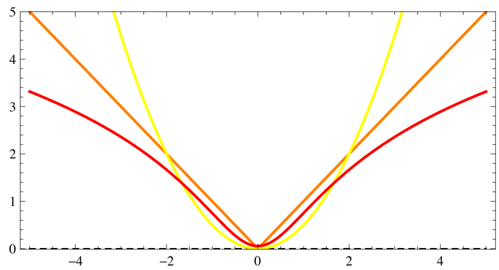
General form of MAP estimator (standard)

$$\mathbf{s}_{\mathrm{MAP}} = \operatorname{argmin}\left(\tfrac{1}{2}\|\mathbf{y}-\mathbf{H}\mathbf{s}\|_2^2 + \sigma^2 \textstyle\sum_n \Phi_U([\mathbf{L}\mathbf{s}]_n)\right)$$



- Gaussian: $p_U(x) = \frac{1}{\sqrt{2\pi}\sigma_0} e^{-x^2/(2\sigma_0^2)}$ $\quad\Rightarrow\quad$ $\Phi_U(x) = \frac{1}{2\sigma_0^2}x^2 + C_1$
- Laplace: $p_U(x) = \frac{\lambda}{2} e^{-\lambda|x|}$ $\quad\Rightarrow\quad$ $\Phi_U(x) = \lambda|x| + C_2$
- Student: $p_U(x) = \dfrac{1}{B\left(r, \frac{1}{2}\right)} \left(\dfrac{1}{x^2+1}\right)^{r+\frac{1}{2}}$ $\quad\Rightarrow\quad$ $\Phi_U(x) = \left(r + \dfrac{1}{2}\right)\log(1+x^2) + C_3$

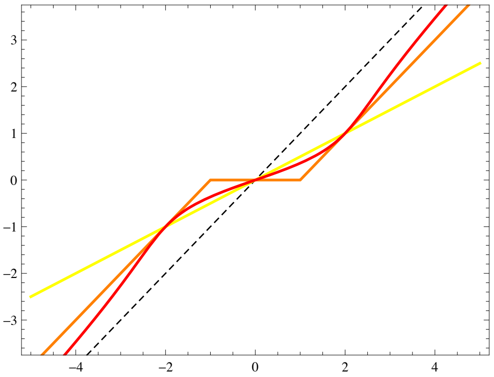
Potential: $\Phi_U(x) = -\log p_U(x)$



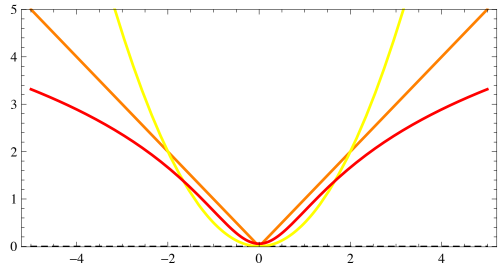
Proximal operator: pointwise denoiser

$$\mathrm{prox}_{\Phi_U}(y;\sigma^2) = \arg\min_{u\in\mathbb{R}} \frac{1}{2}|y-u|^2 + \sigma^2\Phi_U(u)$$

$\tilde{u} = \mathrm{prox}_{\Phi_U}(y;1)$



$\sigma^2\Phi_U(u)$



- linear attenuation — ℓ_2 minimization
- soft-threshold — ℓ_1 minimization
- shrinkage function $\approx$ ℓ_p relaxation for $p \to 0$

Maximum a posteriori (MAP) estimation

- Constrained optimization formulation

Auxiliary **innovation** variable: $\mathbf{u} = \mathbf{L}\mathbf{s}$

$$\mathbf{s}_{\mathrm{MAP}} = \arg\min_{\mathbf{s}\in\mathbb{R}^K} \left(\frac{1}{2}\|\mathbf{y}-\mathbf{H}\mathbf{s}\|_2^2 + \sigma^2 \sum_n \Phi_U\big([\mathbf{u}]_n\big) \right) \text{ subject to } \mathbf{u} = \mathbf{L}\mathbf{s}$$

- Augmented Lagrangian method

Quadratic penalty term: $\frac{\mu}{2}\|\mathbf{L}\mathbf{s}-\mathbf{u}\|_2^2$

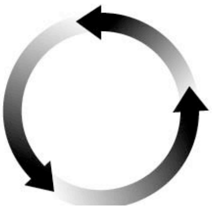
Lagrange multipler vector: $\boldsymbol{\alpha}$

$$\mathcal{L}_{\mathcal{A}}(\mathbf{s},\mathbf{u},\boldsymbol{\alpha}) = \frac{1}{2}\|\mathbf{y}-\mathbf{H}\mathbf{s}\|_2^2 + \sigma^2 \sum_n \Phi_U([\mathbf{u}]_n) + \boldsymbol{\alpha}^T(\mathbf{L}\mathbf{s}-\mathbf{u}) + \frac{\mu}{2}\|\mathbf{L}\mathbf{s}-\mathbf{u}\|_2^2$$

Alternating direction method of multipliers (ADMM)

$$\mathcal{L}_{\mathcal{A}}(\mathbf{s}, \mathbf{u}, \boldsymbol{\alpha}) = \frac{1}{2}\|\mathbf{y} - \mathbf{H}\mathbf{s}\|_2^2 + \sigma^2 \sum_n \Phi_U([\mathbf{u}]_n) + \boldsymbol{\alpha}^T(\mathbf{L}\mathbf{s} - \mathbf{u}) + \frac{\mu}{2}\|\mathbf{L}\mathbf{s} - \mathbf{u}\|_2^2$$

Sequential minimization



$$\mathbf{s}^{k+1} \leftarrow \arg\min_{\mathbf{s}\in\mathbb{R}^N} \mathcal{L}_{\mathcal{A}}(\mathbf{s}, \mathbf{u}^k, \boldsymbol{\alpha}^k)$$

$$\boldsymbol{\alpha}^{k+1} = \boldsymbol{\alpha}^k + \mu\big(\mathbf{L}\mathbf{s}^{k+1} - \mathbf{u}^k\big)$$

$$\mathbf{u}^{k+1} \leftarrow \arg\min_{\mathbf{u}\in\mathbb{R}^N} \mathcal{L}_{\mathcal{A}}(\mathbf{s}^{k+1}, \mathbf{u}, \boldsymbol{\alpha}^{k+1})$$

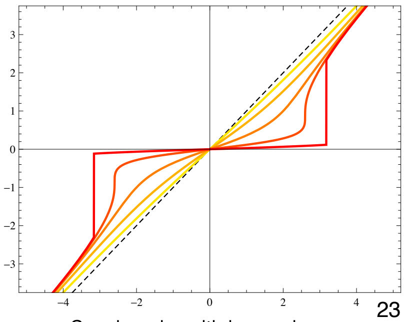
Linear inverse problem: $\mathbf{s}^{k+1} = \left(\mathbf{H}^T\mathbf{H} + \mu\mathbf{L}^T\mathbf{L}\right)^{-1}\left(\mathbf{H}^T\mathbf{y} + \mathbf{z}^{k+1}\right)$

with $\mathbf{z}^{k+1} = \mathbf{L}^T\left(\mu\mathbf{u}^k - \boldsymbol{\alpha}^k\right)$

Nonlinear denoising: $\mathbf{u}^{k+1} = \operatorname{prox}_{\Phi_U}\left(\mathbf{L}\mathbf{s}^{k+1} + \frac{1}{\mu}\boldsymbol{\alpha}^{k+1}; \frac{\sigma^2}{\mu}\right)$

- Proximal operator taylored to stochastic model

$$\operatorname{prox}_{\Phi_U}(y;\lambda) = \arg\min_u \frac{1}{2}|y-u|^2 + \lambda\Phi_U(u)$$

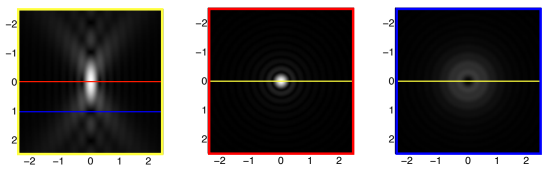


Cauchy prior with increasing s_0

Deconvolution of fluorescence micrographs

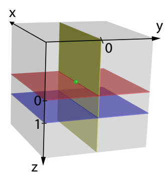
- Physical model of a diffraction-limited microscope

$$g(x,y,z) = (h_{3\mathrm{D}} * s)(x,y,z)$$



3-D point spread function (PSF)

$$h_{3\mathrm{D}}(x,y,z) = I_0\left|p_\lambda\left(\tfrac{x}{M}, \tfrac{y}{M}, \tfrac{z}{M^2}\right)\right|^2$$

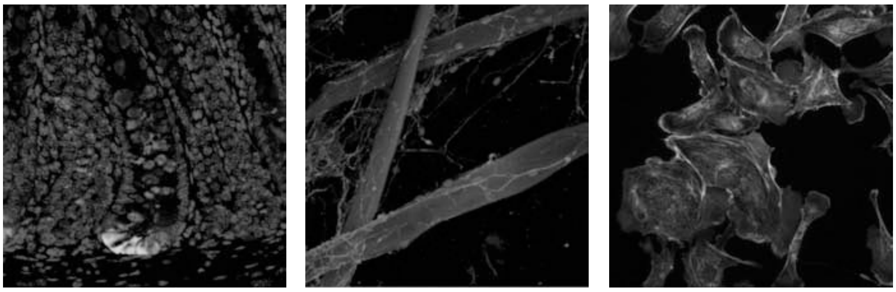


$$p_\lambda(x,y,z) = \int_{\mathbb{R}^2} P(\omega_1,\omega_2)\exp\left(\mathrm{j}2\pi z\frac{\omega_1^2+\omega_2^2}{2\lambda f_0^2}\right)\exp\left(-\mathrm{j}2\pi\frac{x\omega_1+y\omega_2}{\lambda f_0}\right)\mathrm{d}\omega_1\mathrm{d}\omega_2$$

Optical parameters

- λ: wavelength (emission)
- M: magnification factor
- f_0: focal length
- $P(\omega_1,\omega_2) = \mathbb{1}_{\|\boldsymbol{\omega}\|<R_0}$: pupil function
- $\mathrm{NA} = n\sin\theta = R_0/f_0$: numerical aperture

Deconvolution experiments



(a) (b) (c)

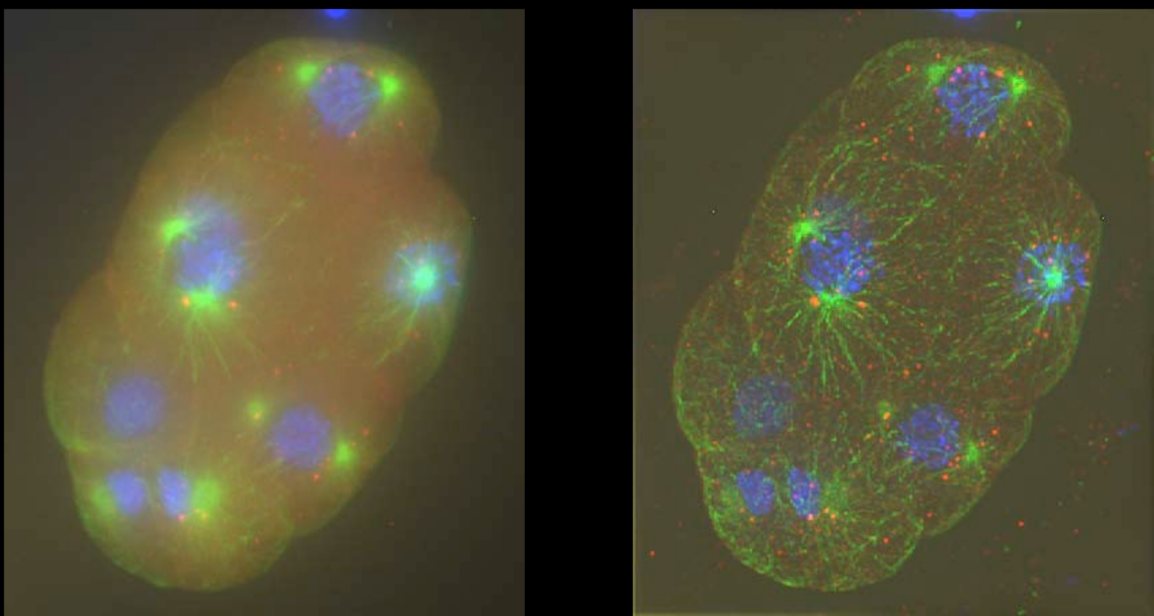
Figure 10.3 Images used in deconvolution experiments. (a) Stem cells surrounded by goblet cells. (b) Nerve cells growing around fibers. (c) Artery cells.

Table 10.2 Deconvolution performance of MAP estimators based on different prior distributions.

		Estimation performance (SNR in dB)		
	BSNR (dB)	Gaussian	Laplace	Student's
Stem cells	20	**14.43**	13.76	11.86
	30	**15.92**	15.77	13.15
	40	**18.11**	**18.11**	13.83
Nerve cells	20	13.86	**15.31**	14.01
	30	15.89	**18.18**	15.81
	40	18.58	**20.57**	16.92
Artery cells	20	14.86	**15.23**	13.48
	30	16.59	**17.21**	14.92
	40	18.68	**19.61**	15.94

$\mathbf{L}$: discrete gradient

3D deconvolution with sparsity constraints



Maximum intensity projections of 384×448×260 image stacks;

Leica DM 5500 widefield epifluorescence microscope with a 63× oil-immersion objective;

C. Elegans embryo labeled with Hoechst, Alexa488, Alexa568;

(Vonesch-U. *IEEE Trans. Im. Proc.* 2009)

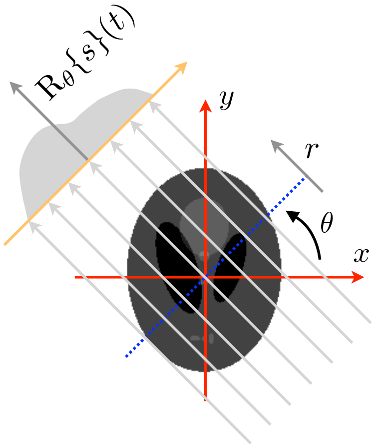
Computed tomography (straight rays)

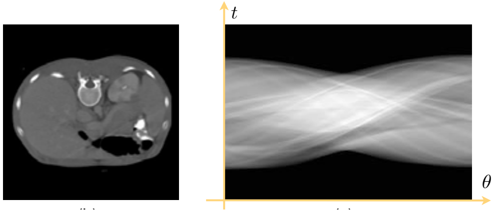
Projection geometry: $\boldsymbol{x} = t\boldsymbol{\theta} + r\boldsymbol{\theta}^{\perp}$ with $\boldsymbol{\theta} = (\cos\theta, \sin\theta)$

- Radon transform (line integrals)

$$\mathrm{R}_{\theta}\{s(\boldsymbol{x})\}(t) = \int_{\mathbb{R}} s(t\boldsymbol{\theta} + r\boldsymbol{\theta}^{\perp})\mathrm{d}r$$

$$= \int_{\mathbb{R}^2} s(\boldsymbol{x})\delta(t - \langle \boldsymbol{x}, \boldsymbol{\theta}\rangle)\mathrm{d}\boldsymbol{x}$$

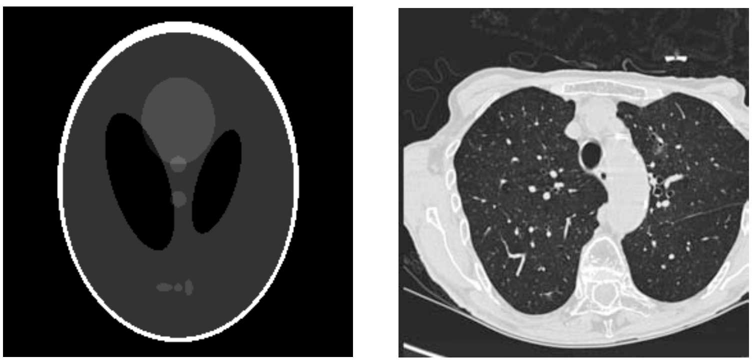




sinogram

Equivalent analysis functions: $\eta_m(\boldsymbol{x}) = \delta\big(t_m - \langle \boldsymbol{x}, \boldsymbol{\theta}_m\rangle\big)$

Computed tomography reconstruction results



(a) (b)

Figure 10.6 Images used in X-ray tomographic reconstruction experiments. (a) The Shepp-Logan (SL) phantom. (b) Cross section of a human lung.

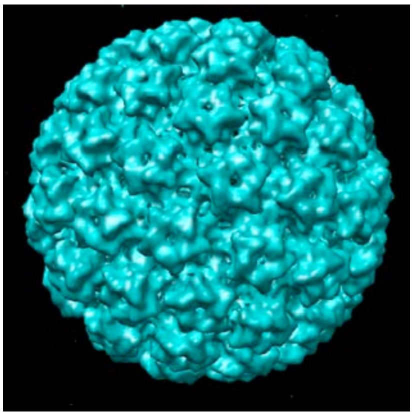
Table 10.4 Reconstruction results of X-ray computed tomography using different estimators.

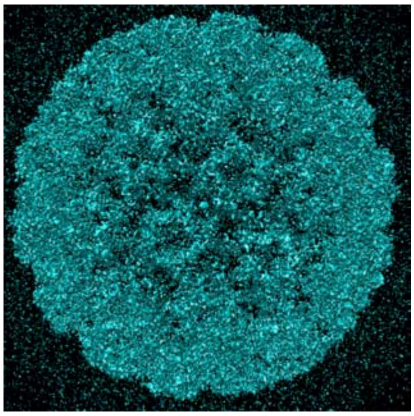
	Directions	Estimation performance (SNR in dB)		
		Gaussian	Laplace	Student's
SL Phantom	120	16.8	17.53	**18.76**
SL Phantom	180	18.13	18.75	**20.34**
Lung	180	**22.49**	21.52	21.45
Lung	360	**24.38**	22.47	22.37

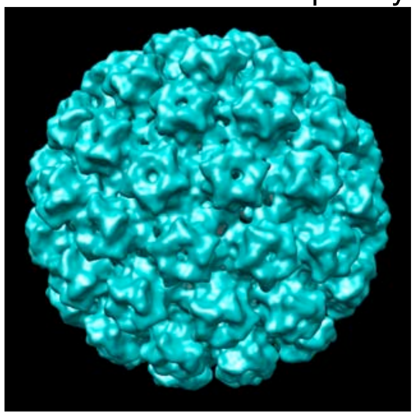
$\mathbf{L}$: discrete gradient

Cryo-electron tomography (real data)

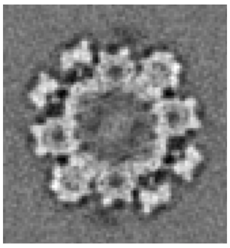
Standard Fourier-based reconstruction	High-resolution Fourier-based reconstruction	High-resolution reconstruction with sparsity

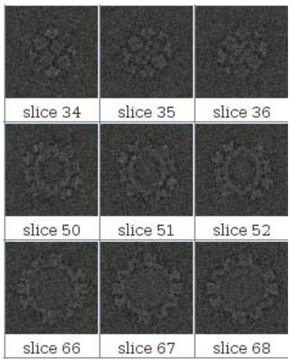


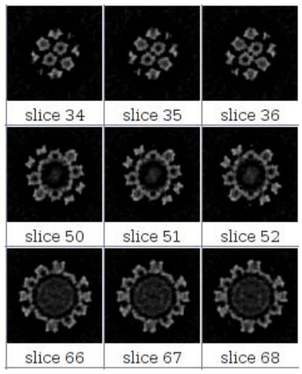




6.185 Å





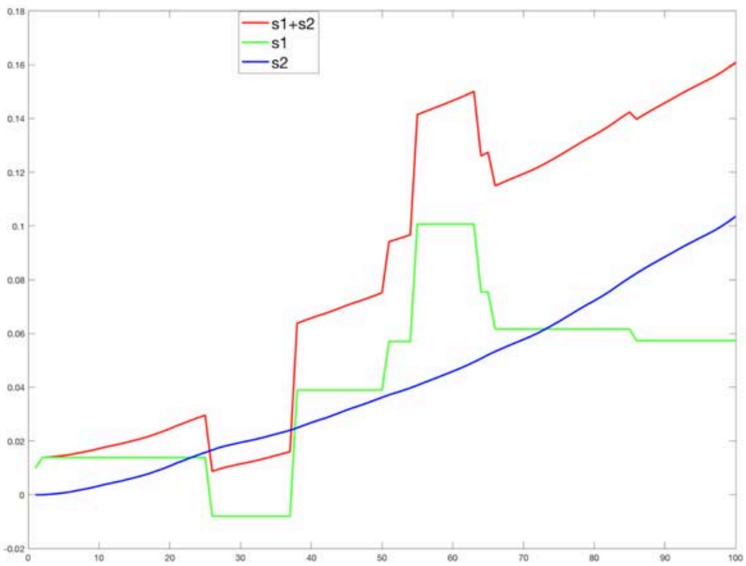


Hybrid Model

$$s = s_1 + s_2$$

$\mathrm{L}_1 = \mathrm{D}$, w_1: impulsive noise

$\mathrm{L}_2 = \mathrm{D}^2$, w_2: Gaussian noise



Fourier Sampling (1D MRI)

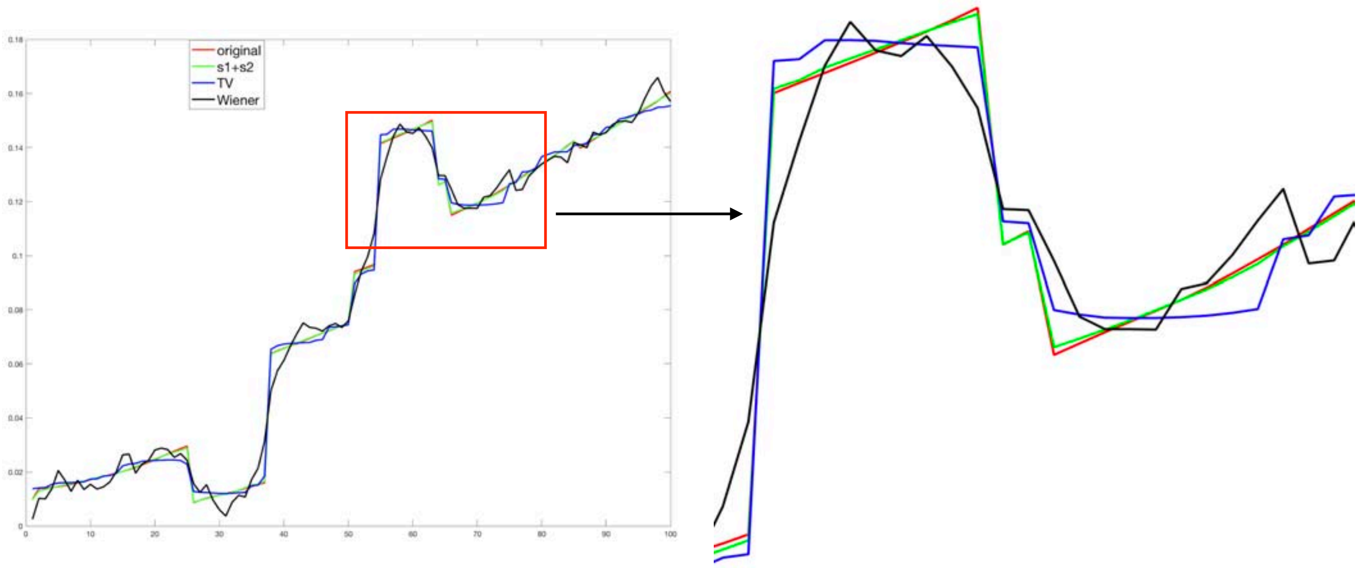
- Forward model: Quasi-random sampling in Fourier domain
- Denser sampling at low-frequencies
- Recovery methods:

Hybrid: $\min_{\mathbf{s}_1,\mathbf{s}_2} \|\mathbf{y} - \mathbf{H}(\mathbf{s}_1 + \mathbf{s}_2)\|_2^2 + \lambda_1 \|\mathbf{D}\mathbf{s}_1\|_1 + \lambda_2 \|\mathbf{D}^2\mathbf{s}_2\|_2^2$

TV: $\min_{\mathbf{s}} \|\mathbf{y} - \mathbf{H}\mathbf{s}\|_2^2 + \lambda \|\mathbf{D}\mathbf{s}\|_1$

Wiener $\min_{\mathbf{s}} \|\mathbf{y} - \mathbf{H}\mathbf{s}\|_2^2 + \lambda \|\mathbf{D}^2\mathbf{s}\|_2^2$

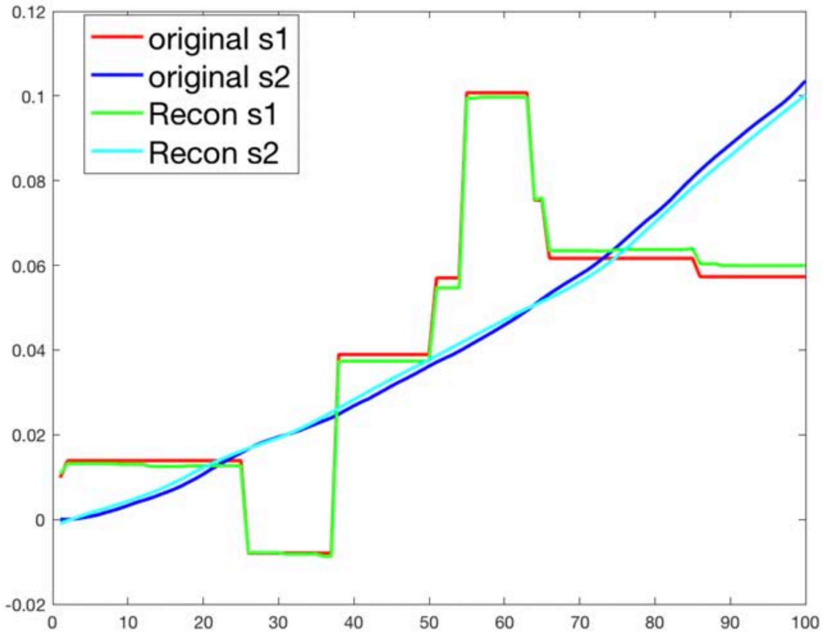
Example 1



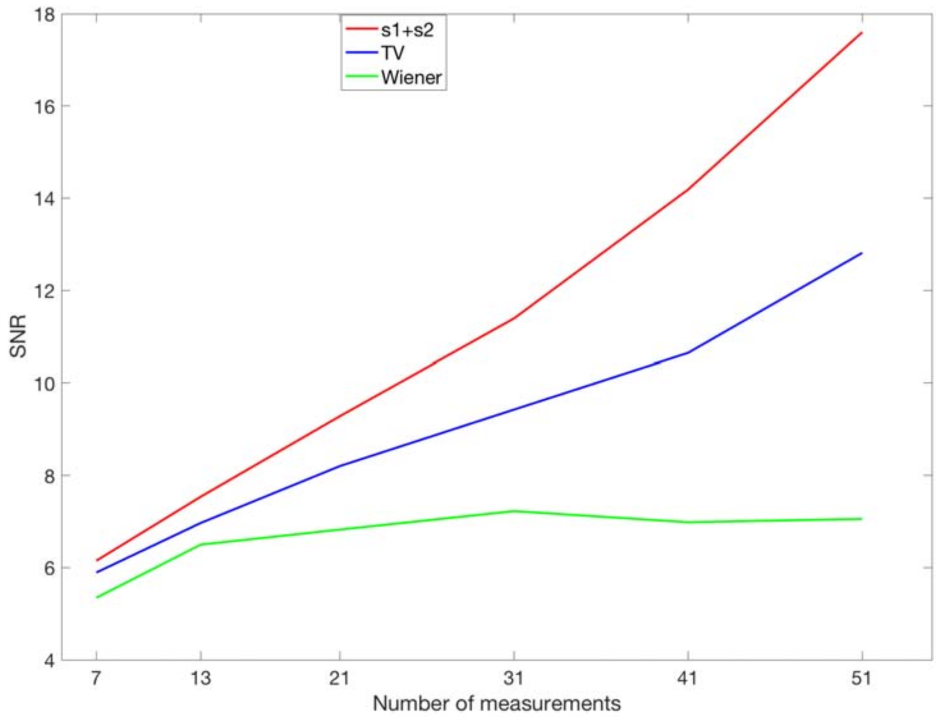
Number of measurements: M = 31

Number of basis functions: K=100

Example 1 - Unmixed components



Comparison

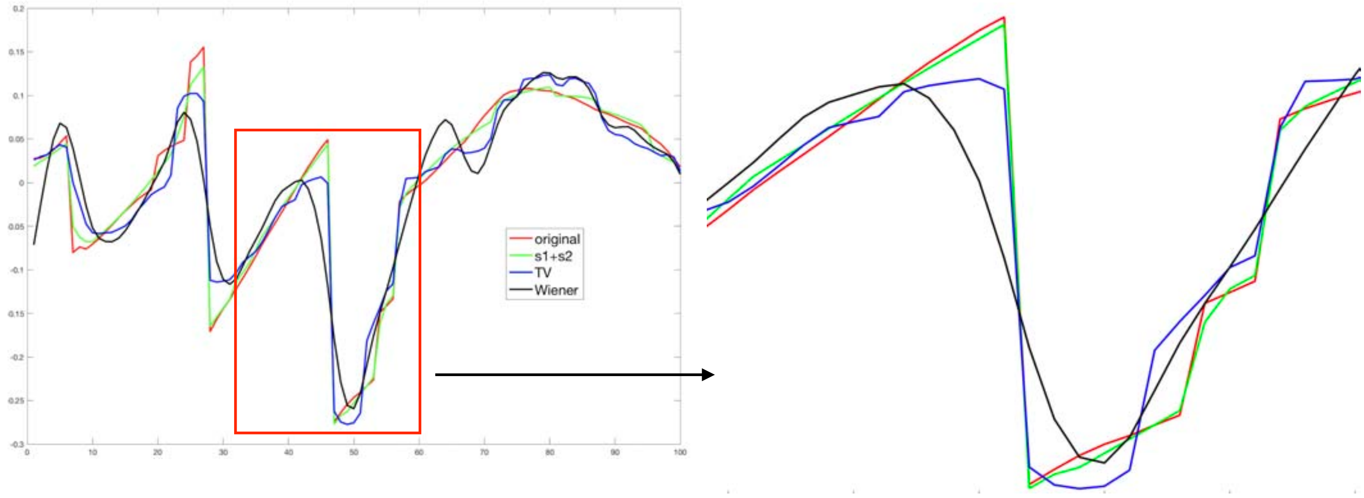


Average performance over 20 signals

Example 2

$s = s_1 + s_2$

$L_1 = D$, w_1: impulsive noise
$L_2 = D^2$, w_2: Gaussian noise

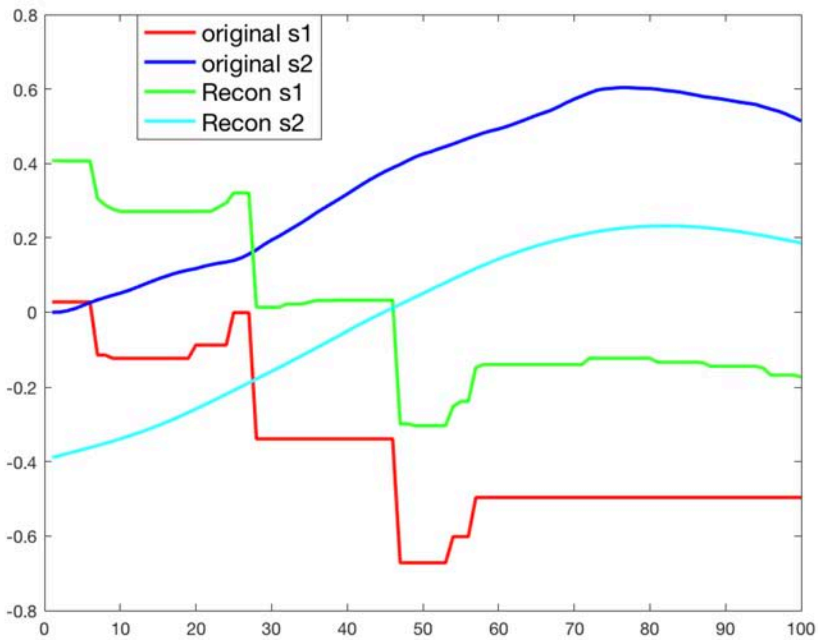


Number of measurements: M = 31

Number of basis functions: K=100

Example 2 - Unmixed components

$L_1 = D$, w_1: impulsive noise
$L_2 = D^2$, w_2: Gaussian noise



CONCLUSION

- Unifying continuous-domain stochastic model
  - Backward compatibility with classical Gaussian theory
  - Operator-based formulation: Lévy-driven SDEs or SPDEs
  - **Gaussian** vs. **sparse** (generalized Poisson, student, SαS)
- Regularization
  - Sparsification via “operator-like” behavior (whitening)
  - Specific family of id potential functions (typ., non-convex)
- Conceptual framework for sparse signal recovery
  - Principled approach for the development of algorithms
  - **Generalization**: hybrid models

    - link with sparse encoding, dictionary-based methods
- Challenges
  - Model identification / learning (self-tuning)

Acknowledgments

Many thanks to (former) members of EPFL’s Biomedical Imaging Group

- Dr. Pouya Tafti
- Prof. Arash Amini
- Dr. Julien Fageot
- Dr. Emrah Bostan
- Dr. Masih Nilchian
- Pakshal Bohra
- Thomas Debarre
- ....

and collaborators ...

- Prof. Demetri Psaltis
- Prof. Marco Stampanoni
- Prof. Carlos-Oscar Sorzano
- Dr. Arne Seitz
- ....

- Preprints and demos: http://bigwww.epfl.ch/

References

- Theory of sparse stochastic processes

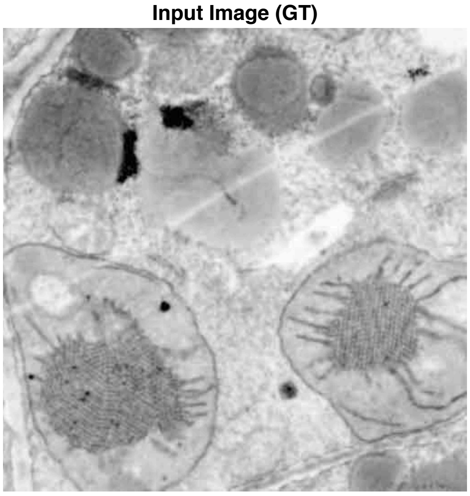
  - M. Unser and P. Tafti, ***An Introduction to Sparse Stochastic Processes***, Cambridge University Press, 2014; preprint, available at `http://www.sparseprocesses.org`.
  - M. Unser, P.D. Tafti, "Stochastic models for sparse and piecewise-smooth signals", *IEEE Trans. Signal Processing*, vol. 59, no. 3, pp. 989-1006, March 2011.
  - M. Unser, P. Tafti, and Q. Sun, "A unified formulation of Gaussian vs. sparse stochastic processes—Part I: Continuous-domain theory," *IEEE Trans. Information Theory*, vol. 60, no. 3, pp. 1945-1962, March 2014.

- Algorithms and imaging applications

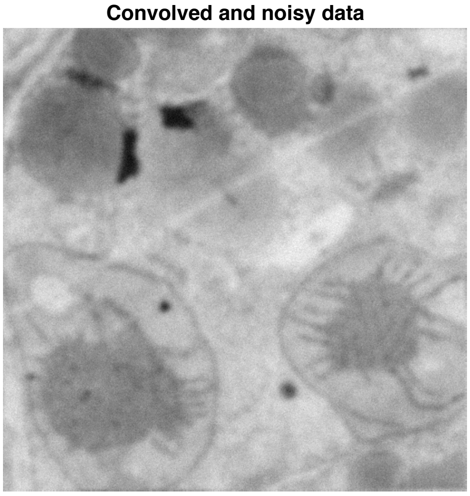
  - E. Bostan, U.S. Kamilov, M. Nilchian, M. Unser, "Sparse Stochastic Processes and Discretization of Linear Inverse Problems," *IEEE Trans. Image Processing*, vol. 22, no. 7, pp. 2699-2710, 2013.
  - C. Vonesch, M. Unser, "A Fast Multilevel Algorithm for Wavelet-Regularized Image Restoration," *IEEE Trans. Image Processing*, vol. 18, no. 3, pp. 509-523, March 2009.
  - M. Nilchian, C. Vonesch, S. Lefkimmiatis, P. Modregger, M. Stampanoni, M. Unser, "Constrained Regularized Reconstruction of X-Ray-DPCI Tomograms with Weighted-Norm," *Optics Express*, vol. 21, no. 26, pp. 32340-32348, 2013.
  - T. Debarre, J. Fageot , H. Gupta, M. Unser, "B-Spline-Based Exact Discretization of Continuous-Domain Inverse Problems With Generalized TV Regularization", *IEEE Trans. Information Theory*, in press.

Ground truth

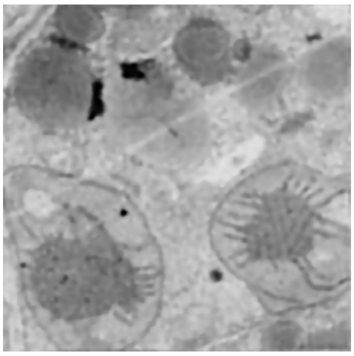
Input Image (GT)



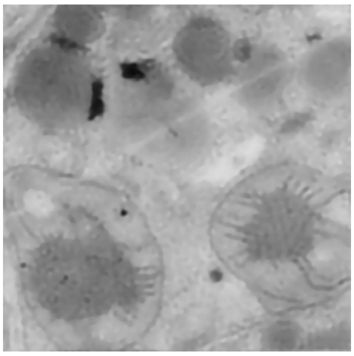
Convolved and noisy data



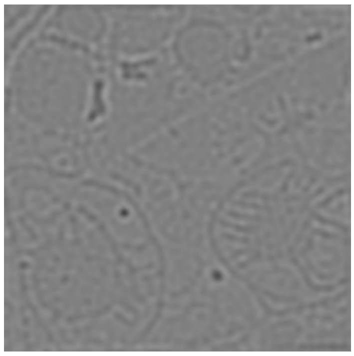
Hybrid



SNR = 27.96 dB



s_1 (Total variation)

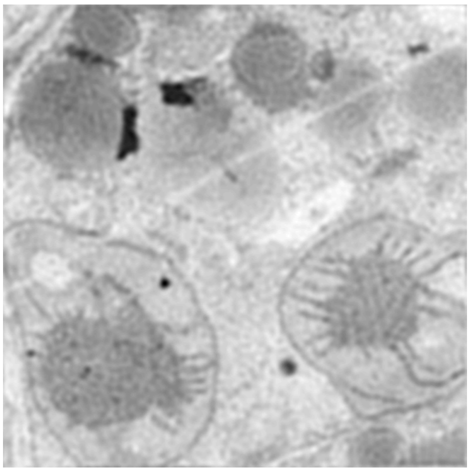


s_2 (Sobolev)

$$\underset{\boldsymbol{c}_1,\boldsymbol{c}_2\in\mathbb{R}^N}{\arg\min}\ \|\mathbf{H}(\boldsymbol{c}_1+\boldsymbol{c}_2)-\boldsymbol{y}\|_2^2+\lambda\Big(\alpha\|\nabla\boldsymbol{c}_1\|_1+(1-\alpha)\|(\boldsymbol{\Delta}+\beta\mathbf{I})\boldsymbol{c}_2\|_2^2\Big)$$

$$\lambda=10,\alpha=0.001,\beta=0.1$$

L2

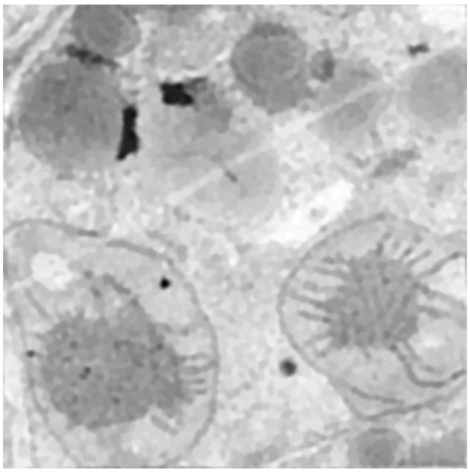


SNR = 27.87 dB

$$\underset{\boldsymbol{c}\in\mathbb{R}^N}{\arg\min}\ \|\mathbf{H}\boldsymbol{c}-\boldsymbol{y}\|_2^2+\lambda\|\boldsymbol{\Delta}\boldsymbol{c}\|_2^2$$

$$\lambda=0.2$$

TV



SNR = 27.97 dB

$$\arg\min_{\boldsymbol{c}\in\mathbb{R}^N} \|\mathbf{H}\boldsymbol{c} - \boldsymbol{y}\|_2^2 + \lambda\|\nabla\boldsymbol{c}\|_1$$

$$\lambda = 0.002$$