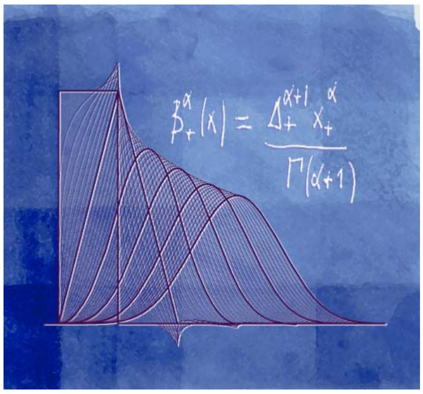
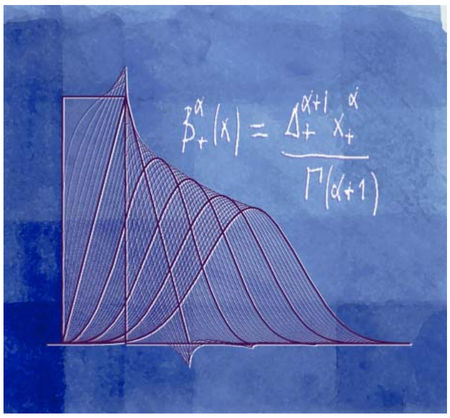
Sparse modeling and the resolution of inverse problems in biomedical imaging

Michael Unser
Biomedical Imaging Group
EPFL, Lausanne, Switzerland



Plenary talk, *IEEE Int. Symp. Biomedical Imaging* (ISBI'15), 16-19 April, 2015, New York, USA

IEEE International Symposium on Biomedical Imaging: Macro to Nano
7-10 July 2002
Washington, DC, USA

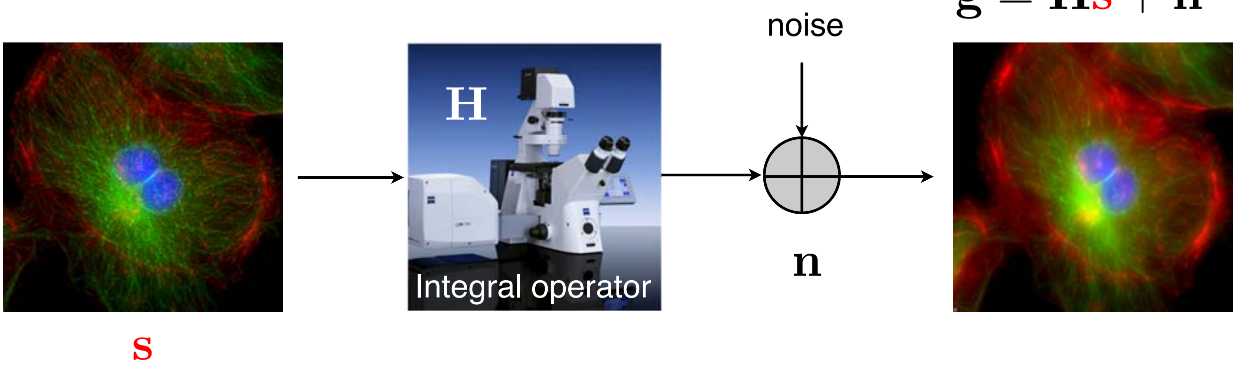


Logo design: Annette Unser

Variational formulation of image reconstruction

- Linear forward model

$$\mathbf{g} = \mathbf{H}\mathbf{s} + \mathbf{n}$$



Ill-posed inverse problem: recover $\mathbf{s}$ from noisy measurements $\mathbf{g}$

- Reconstruction as an optimization problem

$$\mathbf{s}^{\star} = \operatorname{argmin} \underbrace{\|\mathbf{g} - \mathbf{H}\mathbf{s}\|_2^2}_{\text{data consistency}} + \underbrace{\lambda \mathcal{R}(\mathbf{s})}_{\text{regularization}}$$

Classical reconstruction = linear algorithm

$$\mathbf{s}^{\star} = \operatorname{argmin} \underbrace{\|\mathbf{g} - \mathbf{H}\mathbf{s}\|_2^2}_{\text{data consistency}} + \underbrace{\lambda \mathcal{R}(\mathbf{s})}_{\text{regularization}}$$

- Quadratic regularization (Tikhonov)

  $\mathcal{R}(\mathbf{s}) = \|\mathbf{L}\mathbf{s}\|^2$

  Formal linear solution: $\mathbf{s} = (\mathbf{H}^T\mathbf{H} + \lambda \mathbf{L}^T\mathbf{L})^{-1}\mathbf{H}^T\mathbf{g} = \mathbf{R}_\lambda \cdot \mathbf{g}$

$\Updownarrow$ $\mathbf{L} = \mathbf{C}_s^{-1/2}$: Whitening filter

- Statistical formulation under Gaussian hypothesis

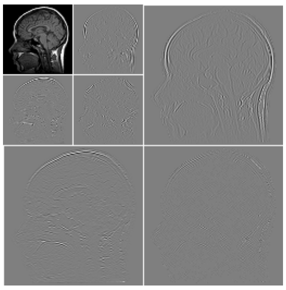
  Wiener (LMMSE) solution = Gauss MMSE = Gauss MAP

$$\mathbf{s}_{\mathrm{MAP}} = \arg\min_{\mathbf{s}} \underbrace{\frac{1}{\sigma^2}\|\mathbf{g} - \mathbf{H}\mathbf{s}\|_2^2}_{\text{Data Log likelihood}} + \underbrace{\|\mathbf{C}_s^{-1/2}\mathbf{s}\|_2^2}_{\text{Gaussian prior likelihood}}$$

Signal covariance: $\mathbf{C}_s = \mathbb{E}\{\mathbf{s} \cdot \mathbf{s}^T\}$

Current trend: non-linear algorithms (l_1 optimization)

$$\mathbf{s}^\star = \arg\min \underbrace{\|\mathbf{y} - \mathbf{H}\mathbf{s}\|_2^2}_{\text{data consistency}} + \underbrace{\lambda\mathcal{R}(\mathbf{s})}_{\text{regularization}}$$



- Wavelet-domain regularization

  Wavelet expansion: $\mathbf{s} = \mathbf{W}\mathbf{v}$ (typically, sparse)

  Wavelet-domain sparsity-constraint: $\mathcal{R}(\mathbf{s}) = \|\mathbf{v}\|_{\ell_1}$ with $\mathbf{v} = \mathbf{W}^{-1}\mathbf{s}$

  (Nowak et al., Daubechies et al. 2004)

- ℓ_1 regularization (Total variation=TV) (Rudin-Osher, 1992)

  $\mathcal{R}(\mathbf{s}) = \|\mathbf{L}\mathbf{s}\|_{\ell_1}$ with $\mathbf{L}$: gradient

- Compressed sensing/sampling (Candes-Romberg-Tao; Donoho, 2006)

Key research questions (for biomedical imaging)

1. Formulation of ill-posed reconstruction problem

   Statistical modeling (beyond Gaussian)
   supporting ***non-linear reconstruction*** *schemes*
   (including CS)

   Sparse stochastic processes

2. Efficient implementation for large-scale imaging problem

   ADMM = smart chaining of simple modules

3. Future trends and open issues

OUTLINE

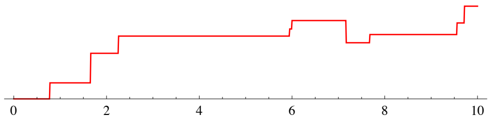
- Variational formulation of inverse problems ✔
- **Statistical modeling**
  Introduction to sparse stochastic processes
  - Generalized innovation model
  - Statistical characterization of signal
- **Algorithm design**
  Reconstruction of biomedical images
  - Discretization of inverse problem
  - Generic MAP estimator (iterative reconstruction algorithm)
  - Applications

Deconvolution microscopy
Computed tomography
Cryo-electron tomography
Differential phase-contrast tomography

An introduction to sparse stochastic processes

Random spline: archetype of sparse signal

non-uniform spline of degree 0



$$\mathrm{D}s(t) = \sum_n a_n \delta(t - t_n) = w(t)$$

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

- Anti-derivative operators

  Shift-invariant solution: $\mathrm{D}^{-1}\varphi(t) = (\mathbb{1}_+ * \varphi)(t) = \int_{-\infty}^{t} \varphi(\tau)\mathrm{d}\tau$

  Scale-invariant solution: $\mathrm{D}_0^{-1}\varphi(t) = \int_0^t \varphi(\tau)\mathrm{d}\tau$

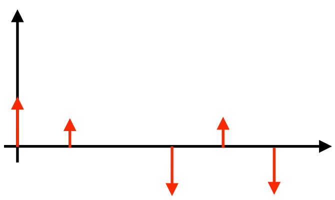
Compound Poisson process

- Stochastic differential equation

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

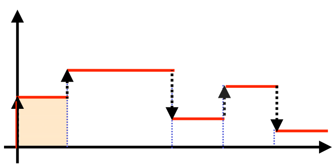
  with boundary condition $s(0) = 0$

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



- Formal solution

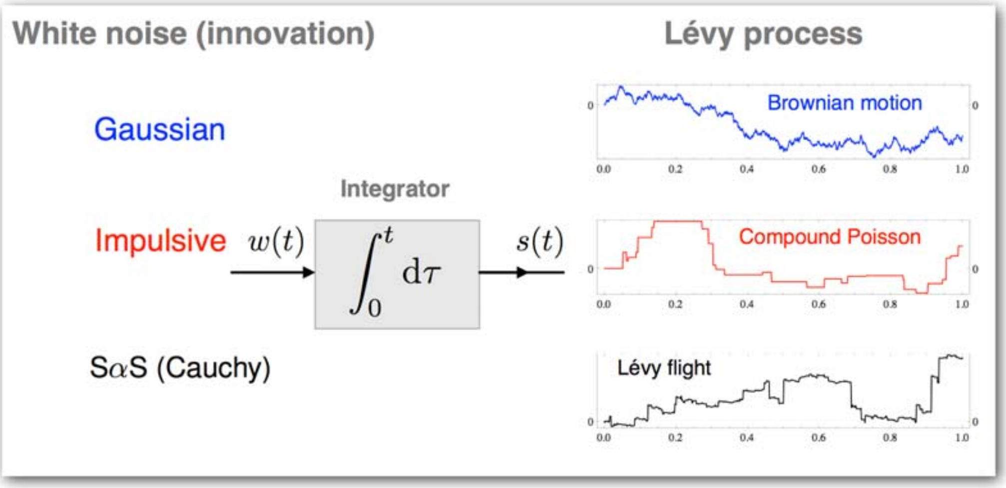
$$s(t) = \mathrm{D}^{-1}w(t) = \sum_n a_n \mathrm{D}^{-1}\{\delta(\cdot - t_n)\}(t)$$
$$= \sum_n a_n \mathbb{1}_+(t - t_n)$$



Lévy processes: all admissible brands of innovations

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

$$\mathrm{D}s = w \qquad \text{(perfect decoupling!)}$$



(Wiener 1923)

(Paul Lévy circa 1930)

Generalized innovation model

Theoretical framework: Gelfand's theory of generalized stochastic processes

Generic test function $\varphi \in \mathcal{S}$ plays the role of index variable

Solution of SDE (general operator)

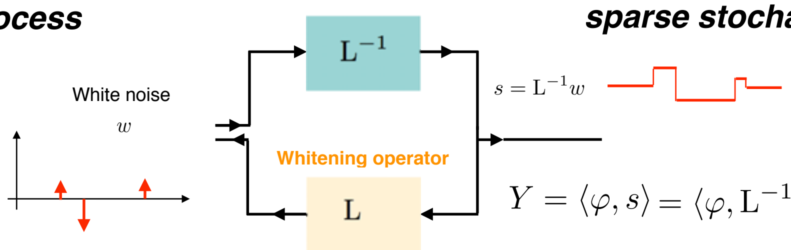
2

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

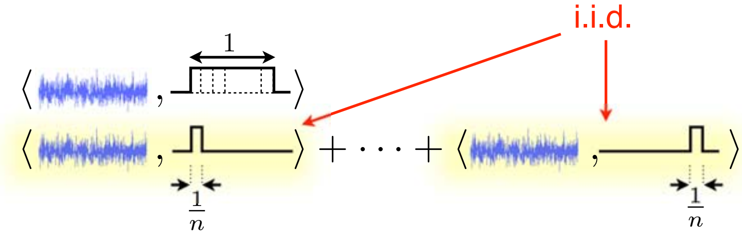
Main feature: inherent sparsity
(few significant coefficients)

Infinite divisibility and Lévy exponents

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 $X \overset{\mathrm{d}}{=} X_1 + \dots + X_N$.

- Rectangular test function

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



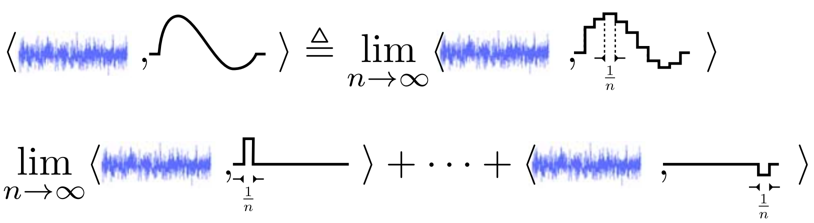
Proposition
The random variable $X = \langle w, \mathrm{rect}\rangle$ where w is a generalized innovation process is infinitely divisible. It is uniquely characterized by its **Lévy exponent** $f(\omega) = \log \hat{p}_{\mathrm{id}}(\omega)$.

$$\widehat{p}_{\mathrm{id}}(\omega) = \mathrm{e}^{f(\omega)} = \int_{\mathbb{R}} p_{\mathrm{id}}(x)\mathrm{e}^{\mathrm{j}\omega x}\mathrm{d}x$$

Bottom line: There is a one-to-one correspondence between Lévy exponents and infinitely divisible distributions and, by extension, innovation processes.

Probability laws of innovations are infinite divisible

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



- Statistical description of white Lévy noise w (innovation)
  - Characterized by canonical (p-admissible) Lévy exponent $f(\omega)$
  - Generic observation: $X = \langle \varphi, w\rangle$ with $\varphi \in L_p(\mathbb{R}^d)$
  - X is ***infinitely divisible*** with (modified) Lévy exponent

$$f_\varphi(\omega) = \log \widehat{p}_X(\omega) = \int_{\mathbb{R}^d} f\big(\omega\varphi(\boldsymbol{x})\big)\mathrm{d}\boldsymbol{x}$$

⇒ 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 Lévy exponent $f_\phi(\omega) = \int_{\mathbb{R}^d} f\big(\omega\phi(\boldsymbol{x})\big)\mathrm{d}\boldsymbol{x}$

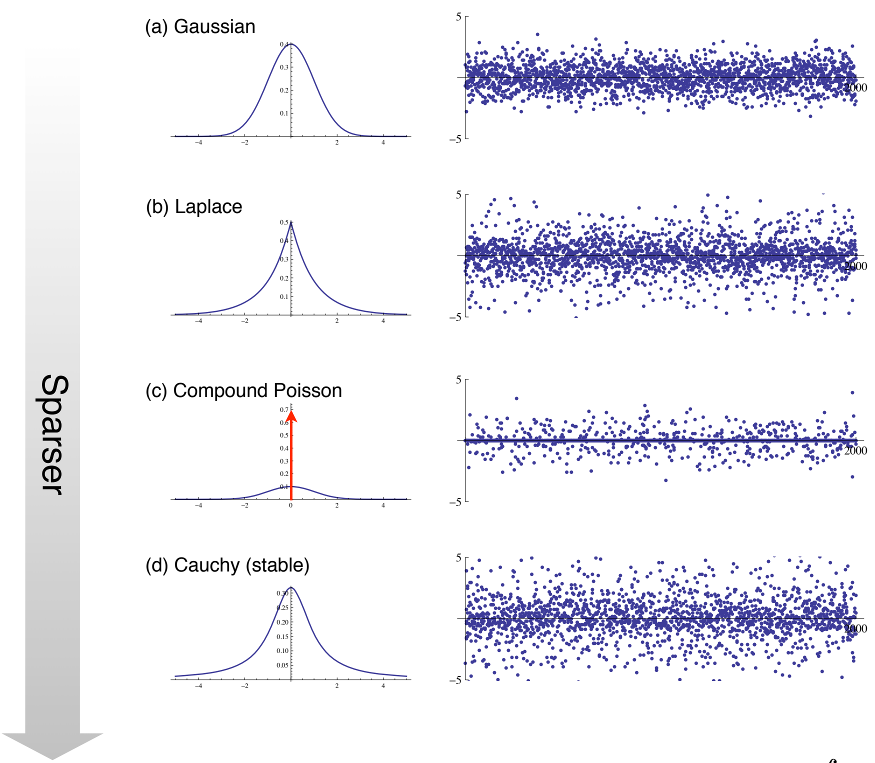
$$\Rightarrow \quad p_Y(y) = \mathcal{F}^{-1}\{\mathrm{e}^{f_\phi(\omega)}\}(y) = \int_{\mathbb{R}} \mathrm{e}^{f_\phi(\omega) - \mathrm{j}\omega y} \frac{\mathrm{d}\omega}{2\pi}$$

= explicit form of pdf

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

Examples of infinitely divisible laws

$p_{\text{id}}(x)$



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

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

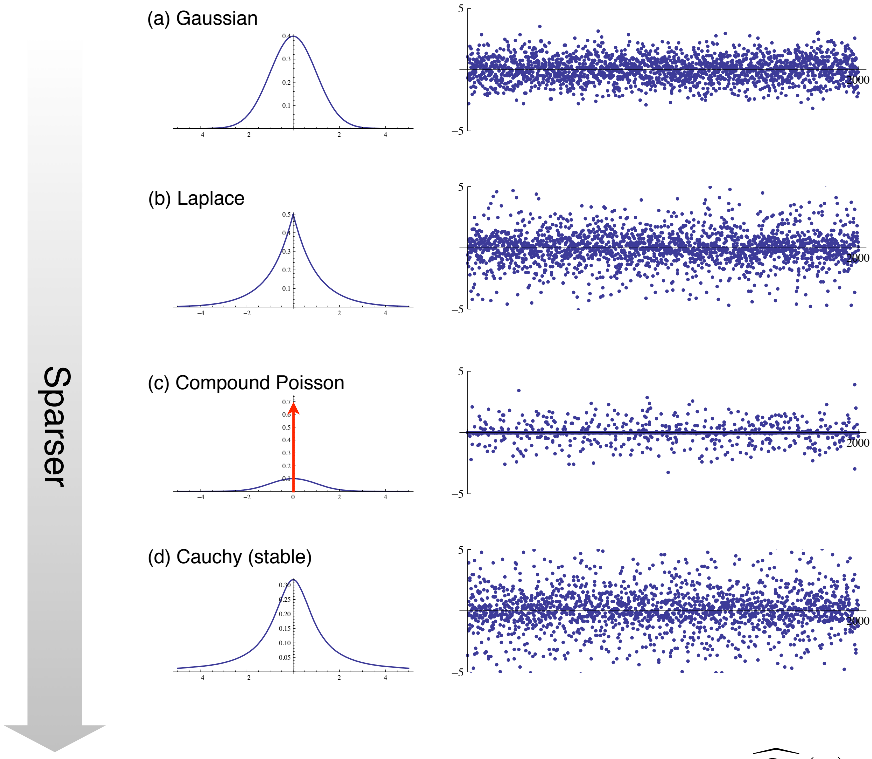
$$p_{\text{Poisson}}(x) = \mathcal{F}^{-1}\{\mathrm{e}^{\lambda(\hat{p}_A(\omega) - 1)}\}$$

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

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

Examples of id noise distributions

$p_{\text{id}}(x)$ Observations: $X_n = \langle w, \text{rect}(\cdot - n)\rangle$



$f(\omega) = -\frac{\sigma_0^2}{2}|\omega|^2$

$f(\omega) = \log\left(\frac{1}{1+\omega^2}\right)$

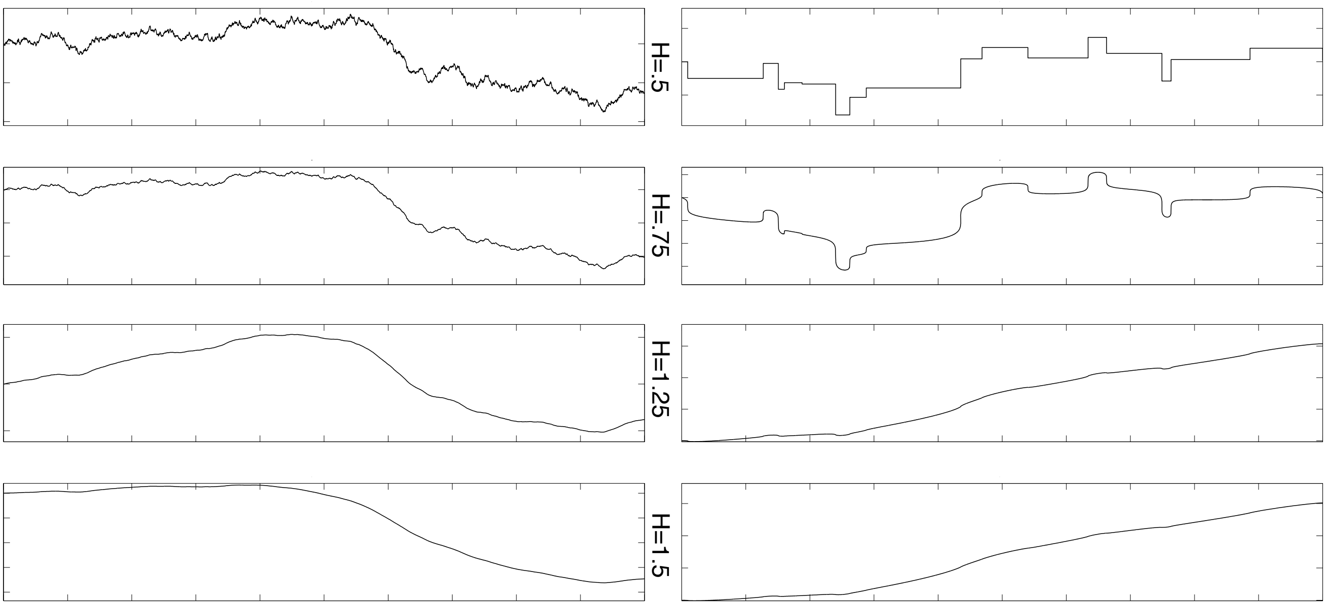
$f(\omega) = \lambda \int_{\mathbb{R}} (e^{jx\omega} - 1)p(x)\mathrm{d}x$

$f(\omega) = -s_0|\omega|$

Complete mathematical characterization: $\widehat{\mathscr{P}_w}(\varphi) = \exp\left(\int_{\mathbb{R}^d} f\left(\varphi(\boldsymbol{x})\right)\mathrm{d}\boldsymbol{x}\right)$

Generation of self-similar processes: $s = \mathrm{L}^{-1}w$

$\mathrm{L} \quad \overset{\mathcal{F}}{\longleftrightarrow} \quad (\mathrm{j}\omega)^{H+\frac{1}{2}} \quad \Rightarrow \quad \mathrm{L}^{-1}$: fractional integrator

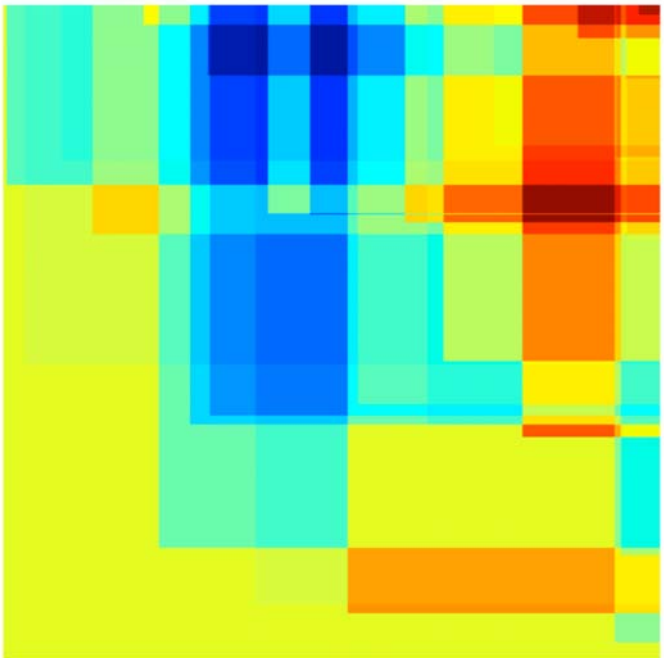


Gaussian **Sparse (generalized Poisson)**

Fractional Brownian motion (Mandelbrot, 1968) (U.-Tafti, *IEEE-SP* 2010)

Aesthetic sparse signal: the Mondrian process

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

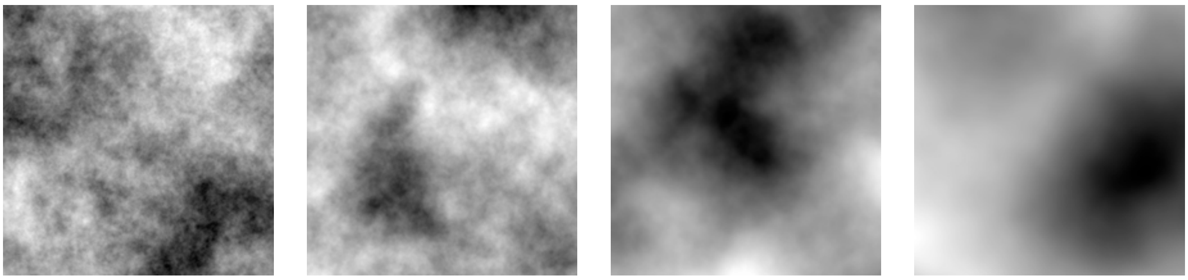


$\lambda = 30$

Scale- and rotation-invariant processes

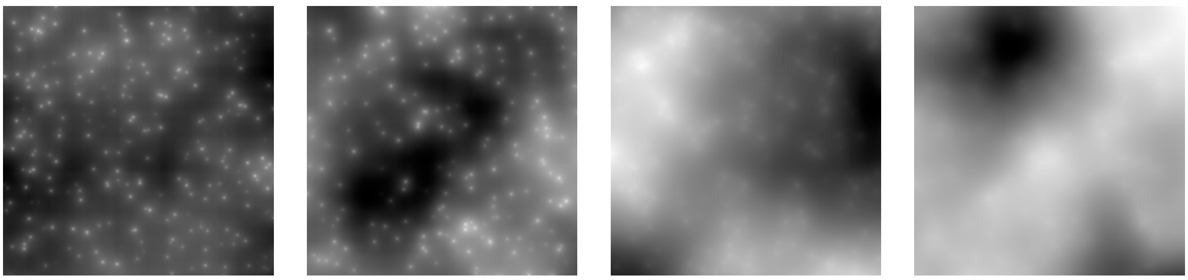
Stochastic partial differential equation : $(-\Delta)^{\frac{H+1}{2}} s(\boldsymbol{x}) = w(\boldsymbol{x})$

Gaussian



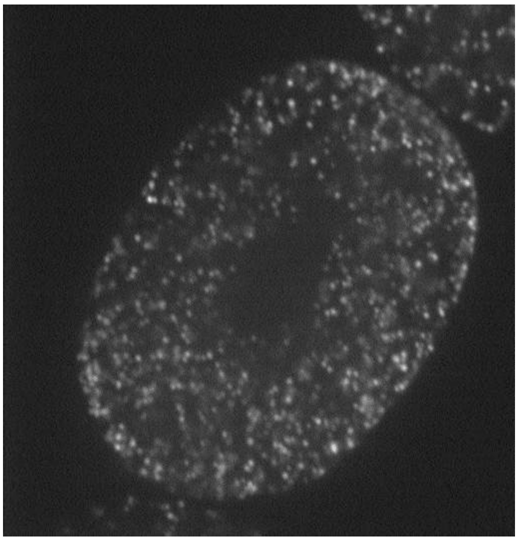
H=.5 H=.75 H=1.25 H=1.75

Sparse (generalized Poisson)



(U.-Tafti, *IEEE-SP* 2010)

Powers of ten: from astronomy to biology



RECONSTRUCTION OF BIOMEDICAL IMAGES

- Discretization of reconstruction problem
- Signal reconstruction algorithm (MAP)
- Examples of image reconstruction
  - Deconvolution microscopy
  - X-ray tomography
  - Cryo-electron tomography
  - Phase contrast tomography

Discretization of reconstruction problem

Spline-like reconstruction model: $s(\boldsymbol{r}) = \sum_{\boldsymbol{k}\in\Omega} s[\boldsymbol{k}]\beta_{\boldsymbol{k}}(\boldsymbol{r}) \quad \longleftrightarrow \quad \mathbf{s} = (s[\boldsymbol{k}])_{\boldsymbol{k}\in\Omega}$

- Innovation model

$$\begin{aligned} \mathrm{L}s &= w \\ s &= \mathrm{L}^{-1}w \end{aligned}$$

Discretization →

$\mathbf{u} = \mathbf{L}\mathbf{s}$ (matrix notation)

p_U is part of **infinitely divisible** family

- Physical model: image formation and acquisition

$$y_m = \int_{\mathbb{R}^d} s_1(\boldsymbol{x})\eta_m(\boldsymbol{x})\mathrm{d}\boldsymbol{x} + n[m] = \langle s_1, \eta_m\rangle + n[m], \quad (m = 1, \dots, M)$$

$$\mathbf{y} = \mathbf{y}_0 + \mathbf{n} = \mathbf{H}\mathbf{s} + \mathbf{n}$$

$\mathbf{n}$: i.i.d. noise with pdf p_N

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

Posterior probability distribution

$$\begin{aligned} p_{S|Y}(\mathbf{s}|\mathbf{y}) &= \frac{p_{Y|S}(\mathbf{y}|\mathbf{s})p_S(\mathbf{s})}{p_Y(\mathbf{y})} = \frac{p_N(\mathbf{y} - \mathbf{H}\mathbf{s})p_S(\mathbf{s})}{p_Y(\mathbf{y})} \\ &= \frac{1}{Z}p_N(\mathbf{y} - \mathbf{H}\mathbf{s})p_S(\mathbf{s}) \end{aligned}$$

(Bayes' rule)

$$\mathbf{u} = \mathbf{L}\mathbf{s} \quad \Rightarrow \quad p_S(\mathbf{s}) \propto p_U(\mathbf{L}\mathbf{s}) \approx \prod_{\boldsymbol{k}\in\Omega} p_U([\mathbf{L}\mathbf{s}]_{\boldsymbol{k}})$$

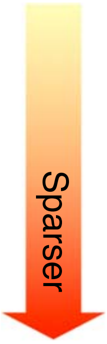
- Additive white Gaussian noise scenario (AWGN)

$$p_{S|Y}(\mathbf{s}|\mathbf{y}) \propto \exp\left(-\frac{\|\mathbf{y} - \mathbf{H}\mathbf{s}\|^2}{2\sigma^2}\right) \prod_{\boldsymbol{k}\in\Omega} p_U([\mathbf{L}\mathbf{s}]_{\boldsymbol{k}})$$

... and then take the log and maximize ...

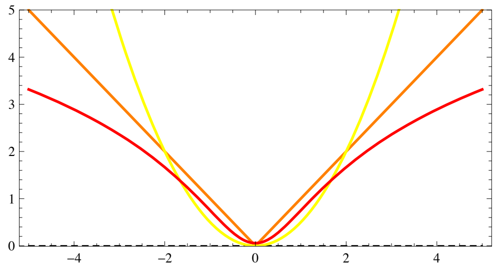
General form of MAP estimator

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



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

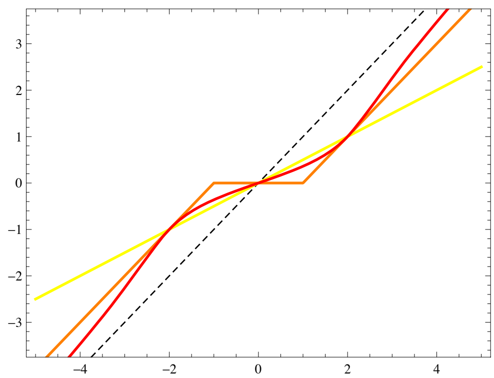
Potential: $\Phi_U(x) = -\log p_U(x)$



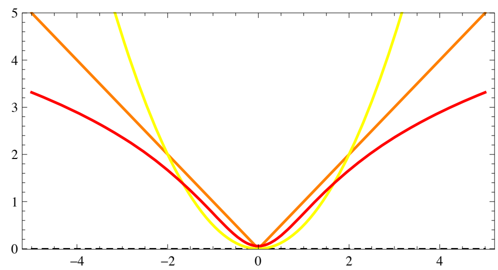
Proximal operator: pointwise denoiser

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

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



$\sigma^2 \Phi_U(u)$



- 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: $\mathbf{u} = \mathbf{L}\mathbf{s}$

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

- Augmented Lagrangian method

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

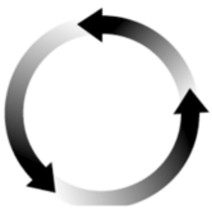
Lagrange multiplier vector: $\boldsymbol{\alpha}$

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

Alternating direction method of multipliers (ADMM)

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

Sequential minimization



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

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

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

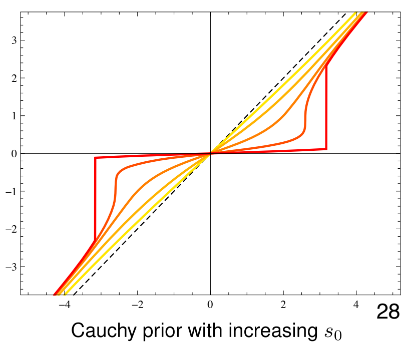
Linear inverse problem: $\mathbf{s}^{k+1} = \left(\mathbf{H}^T\mathbf{H} + \mu\mathbf{L}^T\mathbf{L}\right)^{-1}\left(\mathbf{H}^T\mathbf{y} + \mathbf{z}^{k+1}\right)$

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

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

- Proximal operator taylored to stochastic model

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

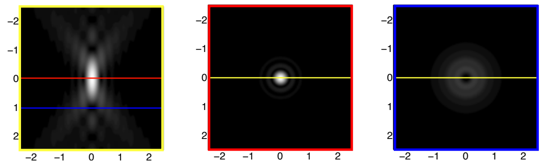


Cauchy prior with increasing s_0

Deconvolution of fluorescence micrographs

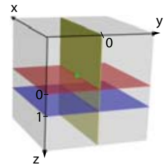
- Physical model of a diffraction-limited microscope

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



3-D point spread function (PSF)

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



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

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: numerical set-up

- Discretization

$\omega_0 \leq \pi$ and representation in (separable) sinc basis $\{\mathrm{sinc}(\boldsymbol{x}-\boldsymbol{k})\}_{\boldsymbol{k}\in\mathbb{Z}^d}$

Analysis functions: $\eta_{\boldsymbol{m}}(x,y,z) = h_{\mathrm{3D}}(x-m_1, y-m_2, z-m_3)$

$$\begin{aligned}[\mathbf{H}]_{\boldsymbol{m},\boldsymbol{k}} &= \langle \eta_{\boldsymbol{m}}, \mathrm{sinc}(\cdot - \boldsymbol{k})\rangle \\ &= \langle h_{\mathrm{3D}}(\cdot - \boldsymbol{m}), \mathrm{sinc}(\cdot - \boldsymbol{k})\rangle \\ &= \big(\mathrm{sinc} * h_{\mathrm{3D}}\big)(\boldsymbol{m}-\boldsymbol{k}) = h_{\mathrm{3D}}(\boldsymbol{m}-\boldsymbol{k}).\end{aligned}$$

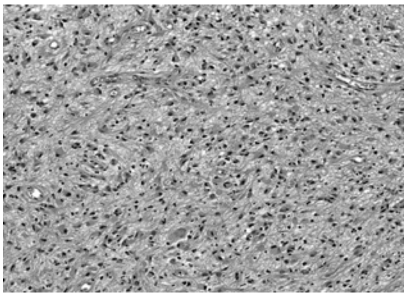
$\mathbf{H}$ and $\mathbf{L}$: convolution matrices diagonalized by discrete Fourier transform

- Linear step of ADMM algorithm implemented using the FFT

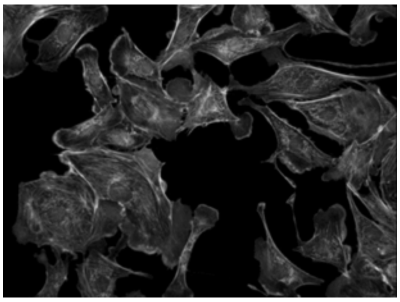
$$\mathbf{s}^{k+1} = \left(\mathbf{H}^T\mathbf{H} + \mu \mathbf{L}^T\mathbf{L}\right)^{-1} \left(\mathbf{H}^T\mathbf{y} + \mathbf{z}^{k+1}\right)$$

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

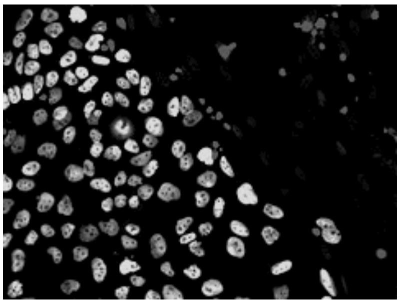
2D deconvolution experiment



Astrocytes cells



bovine pulmonary artery cells



human embryonic stem cells

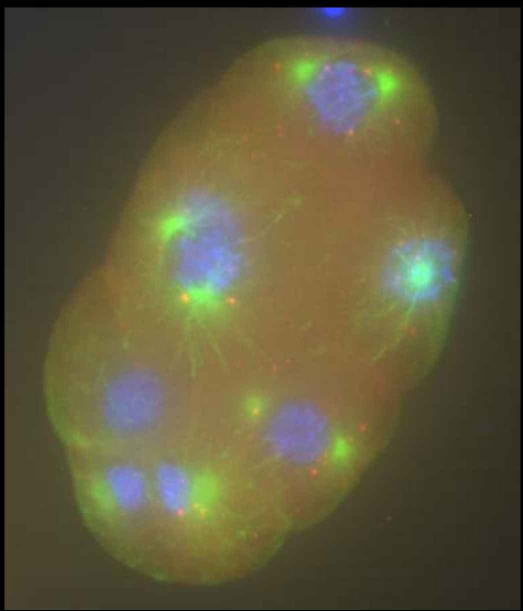
L : gradient

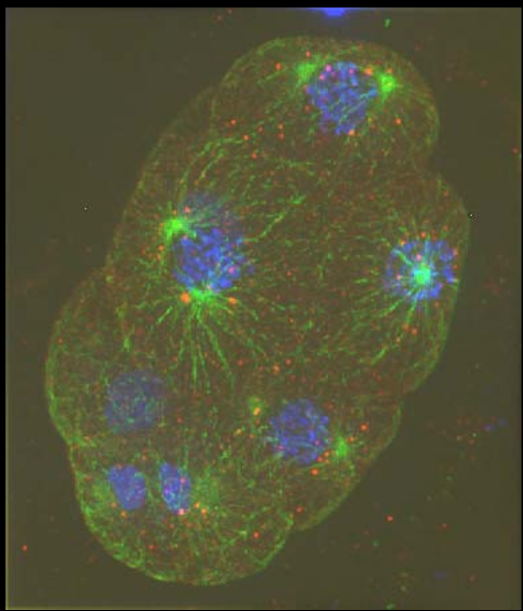
Optimized parameters

Deconvolution results in dB

	Gaussian Estimator	Laplace Estimator	Student's Estimator
Astrocytes cells	**12.18**	10.48	10.52
Pulmonary cells	16.90	**19.04**	18.34
Stem cells	15.81	20.19	**20.50**

3D deconvolution with sparsity constraints





Maximum intensity projections of $384 \times 448 \times 260$ image stacks;

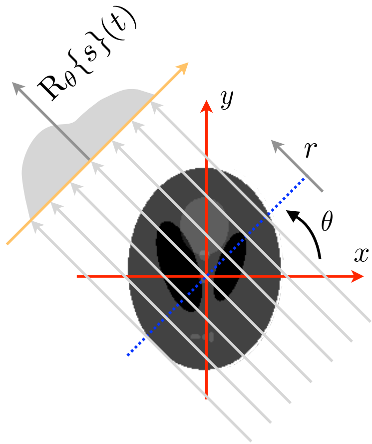
Leica DM 5500 widefield epifluorescence microscope with a $63\times$ oil-immersion objective;

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

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

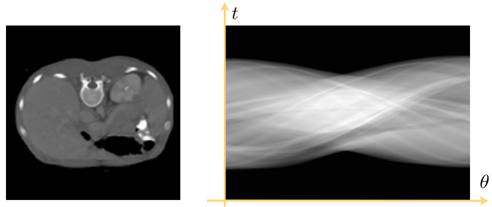
Computed tomography (straight rays)

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



- Radon transform (line integrals)

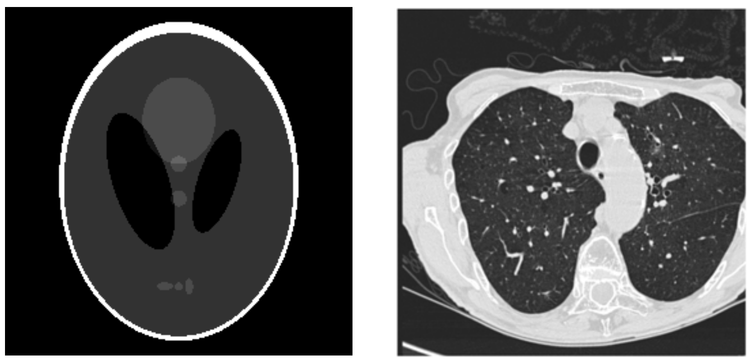
$$R_\theta\{s(\boldsymbol{x})\}(t) = \int_{\mathbb{R}} s(t\boldsymbol{\theta} + r\boldsymbol{\theta}^{\perp})\mathrm{d}r$$
$$= \int_{\mathbb{R}^2} s(\boldsymbol{x})\delta(t - \langle \boldsymbol{x}, \boldsymbol{\theta}\rangle)\mathrm{d}\boldsymbol{x}$$



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.

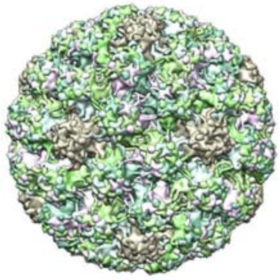
Table 10.4 Reconstruction results of X-ray computed tomography using different estimators.

	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

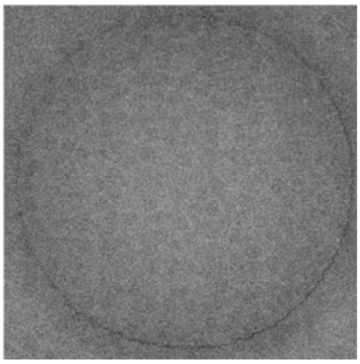
$\mathbf{L}$: discrete gradient

EM: Single particle analysis

Bovine papillomavirus



Cryo-electron micrograph

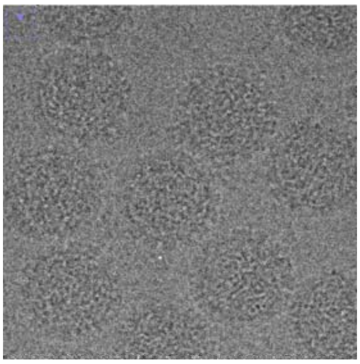


Number of pixels: $256 \times 256 \times 256$
Resolution: $2.474\ \mathring{A}$
Number of particles: 800
Type of symmetry: i1 (60 fold symmetry)

CENTRO NACIONAL DE BIOTECNOLOGIA

mipp

C.-O. Sorzano



noisy projection of identical particles, with unknown orientations

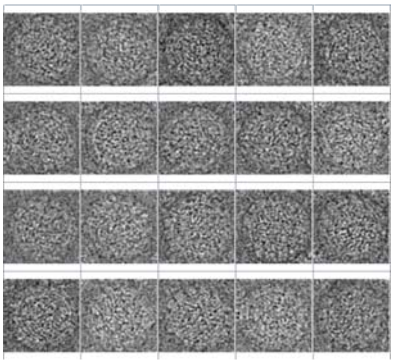
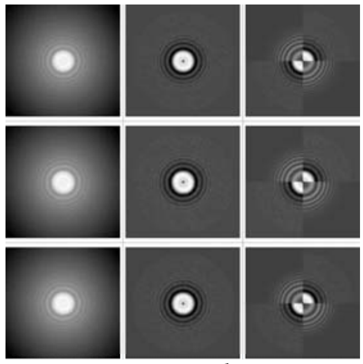


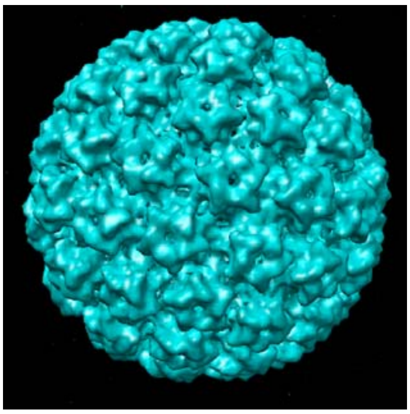
image alignment and classification



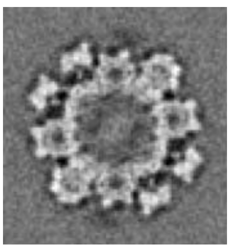
CTF estimation and correction

Image reconstruction (real data)

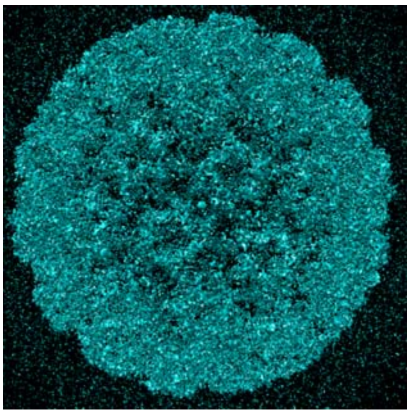
Standard Fourier-based reconstruction

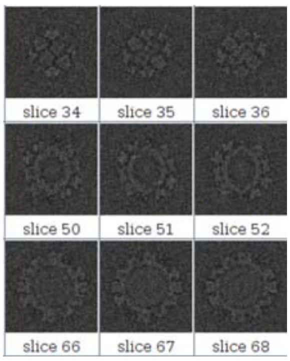


$6.185\ \mathring{A}$

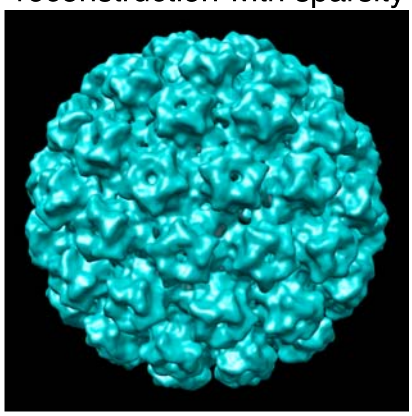


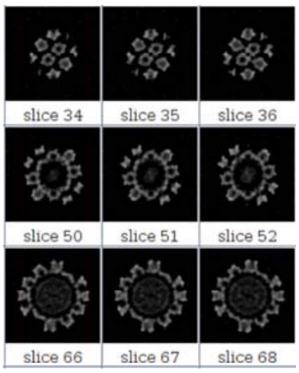
High-resolution Fourier-based reconstruction





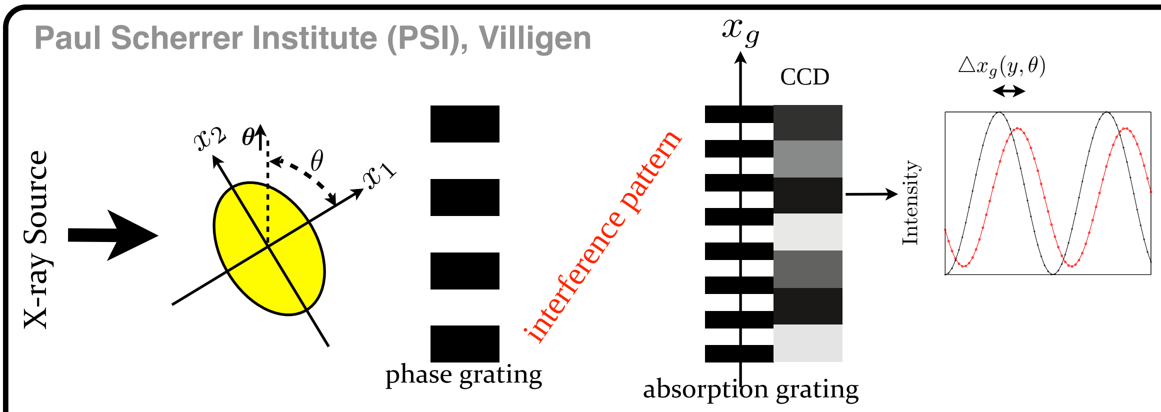
High-resolution reconstruction with sparsity





Differential phase-contrast tomography

Paul Scherrer Institute (PSI), Villigen



(Pfeiffer, *Nature* 2006)

Mathematical model

$$g(t,\theta) = \frac{\partial}{\partial t} \mathrm{R}_\theta\{f\}(t)$$

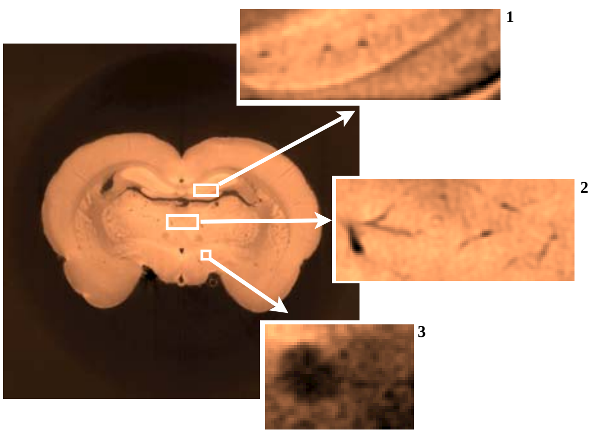
$$\mathbf{g} = \mathbf{H}\ \mathbf{s}$$

$$[\mathbf{H}]_{(i,j),\mathbf{k}} = \frac{\partial}{\partial t} \mathrm{P}_{\theta_j} \beta_{\mathbf{k}}(t_j)$$

Experimental results

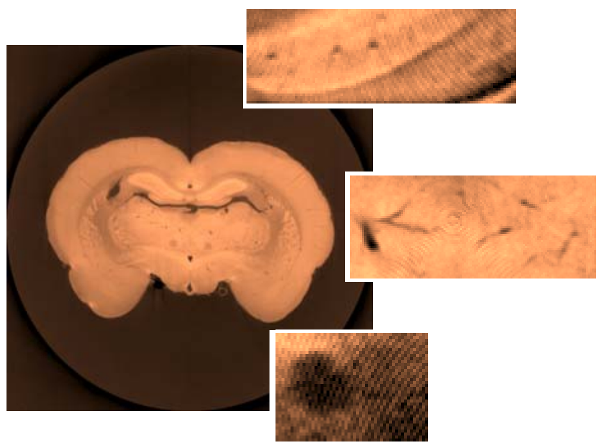
Rat brain reconstruction with 721 projections

ADMM-PCG (TV)



FBP

(Pfeiffer et al. *Nucl. Inst. Meth. Phys. Res.* 2007)



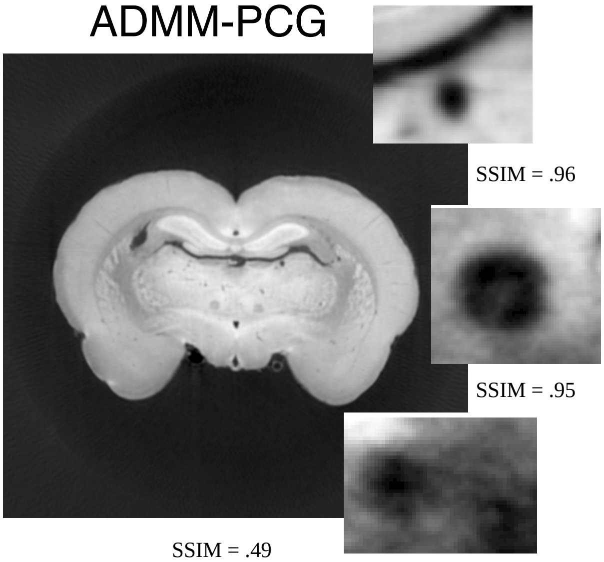
$\mathbf{L}$: discrete gradient, $\Phi(u) = \lambda|u|$

Collaboration: Prof. Marco Stampanoni, TOMCAT PSI / ETHZ

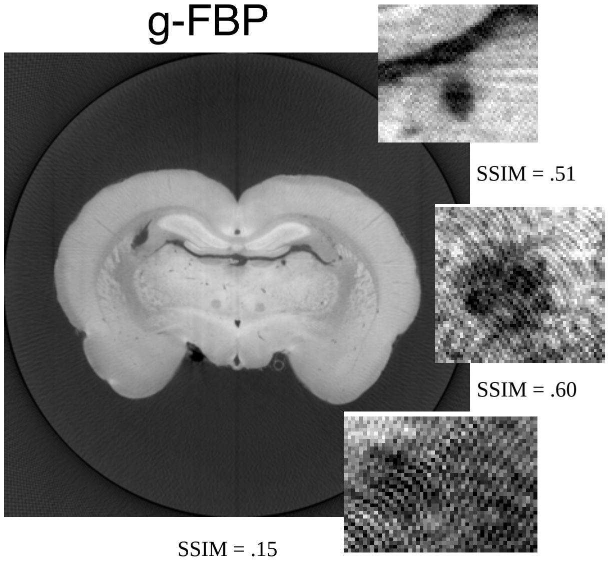
Reducing the numbers of views

Rat brain reconstruction with 181 projections

ADMM-PCG



g-FBP

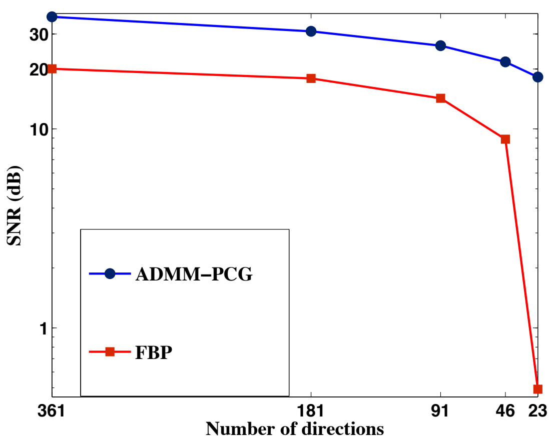


Collaboration: Prof. Marco Stampanoni, TOMCAT PSI / ETHZ

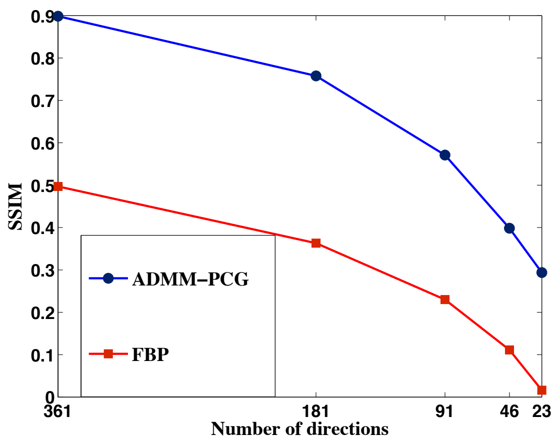
(Nichian et al. *Optics Express* 2013)

Performance evaluation

Goldstandard: high-quality iterative reconstruction with 721 views



(a)



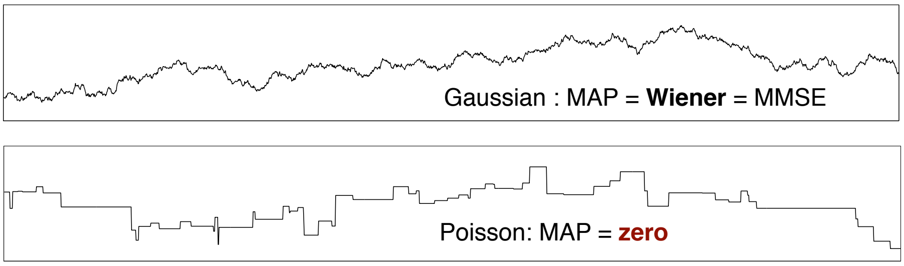
(b)

⇒ Reduction of acquisition time by a factor 10 (or more) ?

CAN WE GO BEYOND MAP ESTIMATION ?

➠ A detailed investigation of simpler denoising problem

Test case: Lévy processes



1. Can we compute the "best"= MMSE estimator ?

Yes, by using **belief propagation** (Kamilov et al., *IEEE-SP* 2013)

2. Can we compute it with an iterative MAP-type algorithm ?

Yes (with the help of **Haar wavelets**) by optimizing the **thresholding function**

Pointwise MMSE estimators for AWGN

- Minimum-mean-square-error (MMSE) estimator from $y = x + n$

$$x_{\mathrm{MMSE}}(y) = \mathbb{E}\{X|Y=y\} = \int_{\mathbb{R}} x \cdot p_{X|Y}(x|y)\mathrm{d}x$$

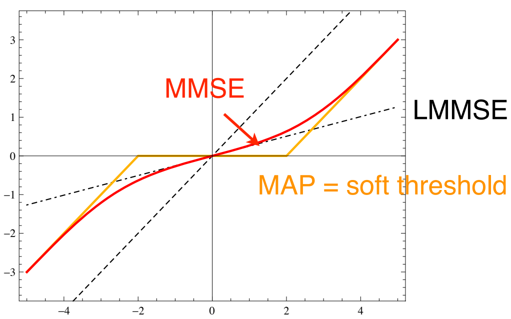
AWGN probability model $\implies$ $p_{Y|X}(y|x) = g_\sigma(y-x)$ and $p_Y = g_\sigma * p_X$

g_σ: Gaussian pdf (zero mean with variance σ^2)

- Stein's formula for AWGN

$$x_{\mathrm{MMSE}}(y) = y - \sigma^2 \Phi_Y'(y)$$

where $\Phi_Y'(y) = -\frac{\mathrm{d}}{\mathrm{d}y} \log p_Y(y) = -\frac{p_Y'(y)}{p_Y(y)}$



Laplacian

Iterative wavelet-based denoising: MAP →MMSE

Consistent Cycle Spinning (CCS)

(Kamilov, IEEE-SPL 2012)

CCS denoising: Solves $\min_{\mathbf{s}} \left\{\frac{1}{2}\|\mathbf{s}-\mathbf{y}\|_2^2 + \frac{\tau}{M}\Phi(\mathbf{A}\mathbf{s})\right\}$ where $\mathbf{A}$ is a tight frame

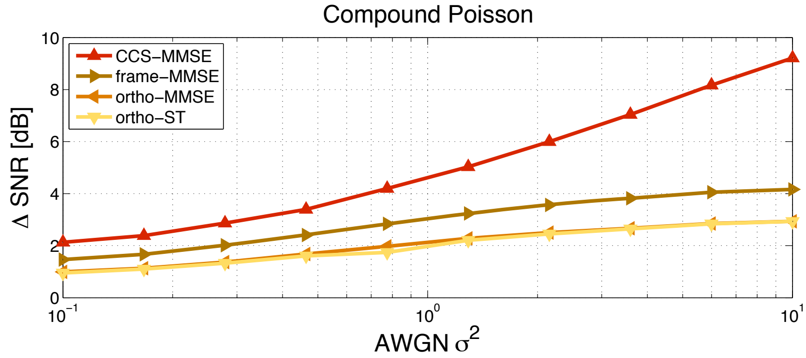
input: $\mathbf{y}, \mathbf{s}^0 \in \mathbb{R}^N, \tau, \mu \in \mathbb{R}^+$
set: $k = 0, \boldsymbol{\lambda}^0 = \mathbf{0}, \mathbf{u} = \mathbf{A}\mathbf{y}$;
repeat
$\mathbf{z}^{k+1} = \mathrm{prox}_\Phi\left(\frac{1}{1+\mu}\left(\mathbf{u} + \mu\mathbf{A}\mathbf{s}^k + \boldsymbol{\lambda}^k\right); \frac{\tau}{1+\mu}\right)$ or $\mathbf{z}^{k+1} = v_{\mathrm{MMSE}}\left(\tilde{\mathbf{z}}^{k+1}; \frac{\sigma^2}{1+\mu}\right)$
$\mathbf{s}^{k+1} = \mathbf{A}^\dagger\left(\mathbf{z}^{k+1} - \frac{1}{\mu}\boldsymbol{\lambda}^k\right)$
$\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k - \mu\left(\mathbf{z}^{k+1} - \mathbf{A}\mathbf{s}^{k+1}\right)$
$k = k + 1$
until stopping criterion
return $\mathbf{s} = \mathbf{s}^k$

- CCS constraint: $\mathbf{z} = \mathbf{A}\mathbf{s}$ with $\|\mathbf{z}\|^2 = M\|\mathbf{s}\|^2$ (enforces energy convervation)

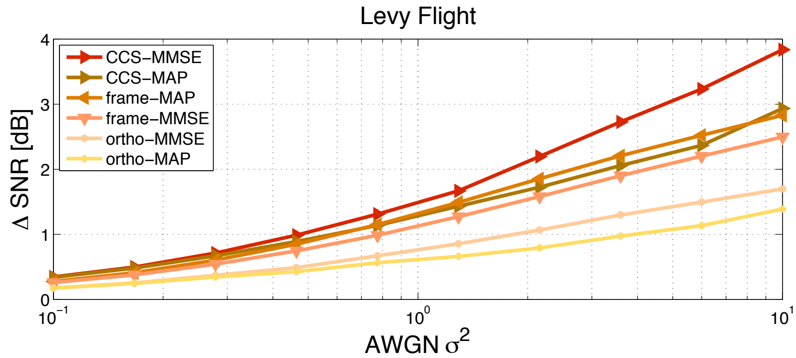
- Variation on a theme: substitute MAP shrinkage by MMSE shrinkage

(Kazerouni, IEEE-SPL 2013) ⇒ Iterative MMSE denoising

Comparison of wavelet denoising strategies



Key empirical finding: CCS MMSE denoising yields optimal solution !!!!



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 novel algorithms
  - Challenge: algorithms for solving large-scale problems in imaging:

    Cryo-electron tomography, diffraction tomography, dynamic MRI (3D + time), etc...
  - **Beyond MAP reconstruction**: MMSE with learning (self-tuning)

Conclusion (Cont’d)

The continuous-domain theory of sparse stochastic processes is compatible with both

20th century SP = **linear and Fourier-based algorithms**, and

21st century SP = **non-linear, sparsity-promoting, wavelet-based algorithms**

... but there are still many open questions ...

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.
  - A. Entezari, M. Nilchian, M. Unser, "A Box Spline Calculus for the Discretization of Computed Tomography Reconstruction Problems," *IEEE Trans. Medical Imaging*, vol. 31, no. 8, pp. 1532-1541, 2012.

Acknowledgments

Many thanks to

- Dr. Pouya Tafti
- Prof. Qiyu Sun
- Prof. Arash Amini
- Dr. Cédric Vonesch
- Dr. Matthieu Guerquin-Kern
- Emrah Bostan
- Ulugbek Kamilov
- Masih Nilchian

- **Members of EPFL's Biomedical Imaging Group**

erc

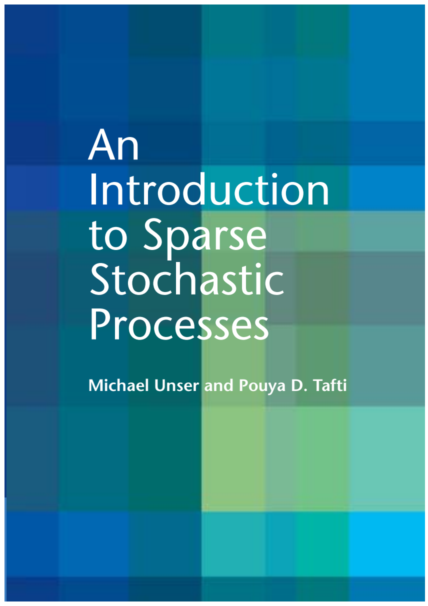
- Preprints and demos: http://bigwww.epfl.ch/

ebook: web preprint

Chapter by chapter

- Cover
- Introduction
- Road map to the monograph
- Mathematical context and background
- Continuous-domain innovation models
- Operators and their inverses
- Splines and wavelets
- Sparse stochastic processes
- Sparse representations
- Infinite divisibility and transform-domain statistics
- Sparse signal recovery
- Wavelet-domain methods

CAMBRIDGE UNIVERSITY PRESS



Cover design: Annette Unser

http://www.sparseprocesses.org