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

\[ g = Hs + n \]

Ill-posed inverse problem: recover \( s \) from noisy measurements \( g \)

- Reconstruction as an optimization problem

\[
\begin{align*}
s^* &= \arg\min_s \left( \|g - Hs\|_2^2 + \lambda \mathcal{R}(s) \right) \\
\text{data consistency} &+ \\text{regularization}
\end{align*}
\]

Classical reconstruction = linear algorithm

- Quadratic regularization (Tikhonov)

\[ \mathcal{R}(s) = \|Ls\|^2 \]

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

\[ \Leftrightarrow \quad L = C_s^{-1/2} : \text{Whitening filter} \]

- Statistical formulation under Gaussian hypothesis

Wiener (LMMSE) solution = Gauss MMSE = Gauss MAP

\[
\begin{align*}
s_{\text{MAP}} &= \arg\min_s \left( \frac{1}{\sigma^2} \|g - Hs\|_2^2 + \frac{1}{2} \|C_s^{-1/2} s\|_2^2 \right) \\
\text{Data Log likelihood} &+ \\text{Gaussian prior likelihood}
\end{align*}
\]

Signal covariance: \( C_s = \mathbb{E}\{s \cdot s^T\} \)
Current trend: non-linear algorithms ($l_1$ optimization)

\[ s^* = \text{argmin} \bigg\| y - Hs \bigg\|_2^2 + \lambda R(s) \]

- Wavelet-domain regularization
  Wavelet expansion: \( s = Wv \) (typically, sparse)
  Wavelet-domain sparsity-constraint: \( R(s) = \|v\|_{l_1} \) with \( v = W^{-1}s \)
  (Nowak et al., Daubechies et al. 2004)

- $l_1$ regularization (Total variation=TV) (Rudin-Osher, 1992)
  \( R(s) = \|Ls\|_{l_1} \) with \( L \): 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

\[ Ds(t) = \sum_n a_n \delta(t - t_n) = w(t) \]

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

- Anti-derivative operators
  - Shift-invariant solution: \( D^{-1}_1 \varphi(t) = (\mathbb{1} \ast \varphi)(t) = \int_{-\infty}^{t} \varphi(\tau) d\tau \)
  - Scale-invariant solution: \( D^{-1}_0 \varphi(t) = \int_{0}^{t} \varphi(\tau) d\tau \)

Compound Poisson process

- Stochastic differential equation
  \[ Ds(t) = w(t) \]
  with boundary condition \( s(0) = 0 \)

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

- Formal solution
  \[ s(t) = D^{-1}w(t) = \sum_n a_n D^{-1}\{\delta(\cdot - t_n)\}(t) \]
  \[ = \sum_n a_n \mathbb{1}_+(t - t_n) \]
**Lévy processes: all admissible brands of innovations**

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

\[ Ds = w \quad \text{(perfect decoupling!)} \]

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)

**innovation process**

\[ X = \langle \varphi, w \rangle \]

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

(Unser et al, IEEE-IT 2014)

\[ Y = \langle \varphi, s \rangle = \langle \varphi, L^{-1}w \rangle = \langle L^{-1}\varphi, w \rangle \]

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 $p_{id}(x)$ is **infinitely divisible** (id) iff., for any $N \in \mathbb{Z}^+$, there exist i.i.d. random variables $X_1, \ldots, X_N$ such that $X \overset{d}{=} X_1 + \cdots + X_N$.

- **Rectangular test function**
  
  $$X = \langle w, \text{rect} \rangle = \left\langle \underbrace{\vdots \vdots \vdots}_{n} \right\rangle$$

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

  $$\hat{p}_{id}(\omega) = e^{f(\omega)} = \int_{\mathbb{R}} p_{id}(x)e^{i\omega x}dx$$

  - **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**

- **Statistical description of white Lévy noise $w$ (innovation)**

  - **Graphical representation**
    - 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
      
      $$f_\varphi(\omega) = \log \hat{p}_X(\omega) = \int_{\mathbb{R}^d} f(\omega \varphi(x))dx$$
Probability laws of sparse processes are id

- Analysis: go back to innovation process: \( w = Ls \)
  - Generic random observation: \( X = \langle \varphi, w \rangle \) with \( \varphi \in S(\mathbb{R}^d) \) or \( \varphi \in L_p(\mathbb{R}^d) \) (by extension)
  - Linear functional: \( Y = \langle \psi, s \rangle = \langle \psi, L^{-1}w \rangle = \langle L^{-1}\ast \psi, w \rangle \)
    
    If \( \phi = L^{-1}\ast \psi \in L_p(\mathbb{R}^d) \) then \( Y = \langle \psi, s \rangle = \langle \phi, w \rangle \) is infinitely divisible
    
    with Lévy exponent \( f_\phi(\omega) = \int_{\mathbb{R}^d} f(\omega \phi(x)) \, dx \)
    
    \[ p_Y(y) = \mathcal{F}^{-1}\{e^{f_\phi(\omega)}\}(y) = \int_{\mathbb{R}} e^{f_\phi(\omega)j\omega y} \frac{d\omega}{2\pi} \]

= explicit form of pdf

Examples of infinitely divisible laws

\[ p_{id}(x) \]

(a) Gaussian

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

(b) Laplace

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

(c) Compound Poisson

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

(d) Cauchy (stable)

\[ p_{Cauchy}(x) = \frac{1}{\pi (x^2 + 1)} \]

Characteristic function:

\[ \hat{p}_{id}(\omega) = \int_{\mathbb{R}} p_{id}(x)e^{j\omega x} \, dx = e^{f(\omega)} \]
Examples of id noise distributions

\[ p_{id}(x) \]

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

(a) Gaussian

\[ f(\omega) = -\frac{\sigma^2}{2} |\omega|^2 \]

(b) Laplace

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

(c) Compound Poisson

\[ f(\omega) = \lambda \int_{\mathbb{R}} (e^{jx\omega} - 1)p(x)dx \]

(d) Cauchy (stable)

\[ f(\omega) = -s_0 |\omega| \]

Complete mathematical characterization:

\[ \mathcal{P}_w(\varphi) = \exp \left( \int_{\mathbb{R}} f(\varphi(x)) \, dx \right) \]

Generation of self-similar processes:

\[ s = L^{-1} \mathcal{W} \]

\[ L \xrightarrow{\mathcal{F}} (j\omega)^{H+\frac{1}{2}} \xrightarrow{\text{L}^{-1}} \text{fractional integrator} \]

Fractional Brownian motion (Mandelbrot, 1968)

Sparse (generalized Poisson) (U.-Tafti, IEEE-SP 2010)
Aesthetic sparse signal: the Mondrian process

\[ L = D_x D_y \quad \leftrightarrow \quad (j\omega_x)(j\omega_y) \]

\( \lambda = 30 \)

Scale- and rotation-invariant processes

Stochastic partial differential equation:

\[ (-\Delta)^{H+\frac{1}{2}} s(x) = w(x) \]

Gaussian

\[
\begin{align*}
H=0.5 & \quad H=0.75 & \quad H=1.25 & \quad H=1.75
\end{align*}
\]

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: \( s(r) = \sum_{k \in \Omega} s[k] \beta_k(r) \quad \leftrightarrow \quad s = (s[k])_{k \in \Omega} \)

- Innovation model

\[
\begin{align*}
Ls & = w \\
\int s & = L^{-1}w
\end{align*}
\]

Discretization \( u = Ls \) (matrix notation)

\( p_U \) is part of infinitely divisible family

- Physical model: image formation and acquisition

\[
y_m = \int \mathcal{S}_1(x) \eta_m(x) dx + n[m] = \langle s_1, \eta_m \rangle + n[m], \quad (m = 1, \ldots, M)
\]

\[
y = y_0 + n = Hs + n
\]

\( n \): i.i.d. noise with pdf \( p_N \)

\[
[H]_{m,k} = \langle \eta_m, \beta_k \rangle = \int \eta_m(r) \beta_k(r) dr: \quad (M \times K) \text{ system matrix}
\]

Posterior probability distribution

\[
p_{S|Y}(s|y) = \frac{p_{Y|S}(y|s)p_S(s)}{p_Y(y)} = \frac{p_N(y - Hs)p_S(s)}{p_Y(y)}
\]

\[
= \frac{1}{Z} p_N(y - Hs)p_S(s)
\]

\[
u = Ls \quad \Rightarrow \quad p_S(s) \propto p_U(Ls) \approx \prod_{k \in \Omega} p_U([Ls]_k)
\]

- Additive white Gaussian noise scenario (AWGN)

\[
p_{S|Y}(s|y) \propto \exp \left( -\frac{\|y - Hs\|^2}{2\sigma^2} \right) \prod_{k \in \Omega} p_U([Ls]_k)
\]

... and then take the log and maximize ...
General form of MAP estimator

\[ s_{\text{MAP}} = \arg \min (\frac{1}{2} \| y - Hs \|_2^2 + \sigma^2 \sum_n \Phi_U([Ls]_n)) \]

- 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: \( p_U(x) = \frac{1}{B(r, \frac{1}{2})} \left( \frac{1}{x^2 + 1} \right)^{r+\frac{1}{2}} \) \( \Rightarrow \Phi_U(x) = (r + \frac{1}{2}) \log(1 + x^2) + C_3 \)

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

Proximal operator: pointwise denoiser

\[ \text{prox}_{\Phi_U}(y; \sigma^2) = \arg \min_{u \in \mathbb{R}} \frac{1}{2} \| y - u \|_2^2 + \sigma^2 \Phi_U(u) \]

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

- linear attenuation
- soft-threshold
- shrinkage function

\( \ell_2 \) minimization
\( \ell_1 \) minimization
\( \approx \ell_p \) relaxation for \( p \rightarrow 0 \)
Maximum a posteriori (MAP) estimation

- Constrained optimization formulation

Auxiliary innovation variable: \( u = Ls \)

\[
s_{\text{MAP}} = \arg \min_{s \in \mathbb{R}^k} \left( \frac{1}{2} \| y - Hs \|_2^2 + \sigma^2 \sum_n \Phi_U([u]_n) \right) \text{ subject to } u = Ls
\]

- Augmented Lagrangian method

Quadratic penalty term: \( \frac{\mu}{2} \| Ls - u \|_2^2 \)

Lagrange multiplier vector: \( \alpha \)

\[
\mathcal{L}_A(s, u, \alpha) = \frac{1}{2} \| g - Hs \|_2^2 + \sigma^2 \sum_n \Phi_U([u]_n) + \alpha^T(Ls - u) + \frac{\mu}{2} \| Ls - u \|_2^2
\]

Alternating direction method of multipliers (ADMM)

\[
\mathcal{L}_A(s, u, \alpha) = \frac{1}{2} \| g - Hs \|_2^2 + \sigma^2 \sum_n \Phi_U([u]_n) + \alpha^T(Ls - u) + \frac{\mu}{2} \| Ls - u \|_2^2
\]

Sequential minimization

\[
s^{k+1} \leftarrow \arg \min_{s \in \mathbb{R}^N} \mathcal{L}_A(s, u^k, \alpha^k)
\]

\[
\alpha^{k+1} = \alpha^k + \mu(Ls^{k+1} - u^k)
\]

\[
u^{k+1} \leftarrow \arg \min_{u \in \mathbb{R}^N} \mathcal{L}_A(s^{k+1}, u, \alpha^{k+1})
\]

Linear inverse problem: 

\[
s^{k+1} = (H^T H + \mu L^T L)^{-1} (H^T y + z^{k+1})
\]

with \( z^{k+1} = L^T (\mu u^k - \alpha^k) \)

Nonlinear denoising: 

\[
u^{k+1} = \text{prox}_{\Phi_U} (Ls^{k+1} + \frac{1}{\mu} \alpha^{k+1}; \sigma^2)
\]

- Proximal operator tailored to stochastic model

\[
\text{prox}_{\Phi_U} (y; \lambda) = \arg \min_u \frac{1}{2} | y - u |^2 + \lambda \Phi_U(u)
\]
Deconvolution of fluorescence micrographs

- Physical model of a diffraction-limited microscope

\[ g(x, y, z) = (h_{3D} \ast s)(x, y, z) \]

3-D point spread function (PSF)

\[ h_{3D}(x, y, z) = I_0 \left| p_\lambda \left( \frac{x}{M}, \frac{y}{M}, \frac{z}{M} \right) \right|^2 \]

\[ p_\lambda(x, y, z) = \int_{\mathbb{R}^2} P(\omega_1, \omega_2) \exp \left( j2\pi \frac{\omega_1^2 + \omega_2^2}{2\lambda f_0^2} \right) \exp \left( -j2\pi \frac{x\omega_1 + y\omega_2}{\lambda f_0} \right) d\omega_1 d\omega_2 \]

Optical parameters

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

Deconvolution: numerical set-up

- Discretization

\[ \omega_0 \leq \pi \] and representation in (separable) sinc basis \( \{\text{sinc}(x - k)\}_{k \in \mathbb{Z}^d} \)

Analysis functions:

\[ \eta_m(x, y, z) = h_{3D}(x - m_1, y - m_2, z - m_3) \]

\[ [H]_{m,k} = \langle \eta_m, \text{sinc}(\cdot - k) \rangle = \langle h_{3D}(\cdot - m), \text{sinc}(\cdot - k) \rangle = (\text{sinc} \ast h_{3D})(m - k) = h_{3D}(m - k). \]

\( H \) and \( L \): convolution matrices diagonalized by discrete Fourier transform

- Linear step of ADMM algorithm implemented using the FFT

\[ s^{k+1} = (H^T H + \mu L^T L)^{-1} (H^T y + z^{k+1}) \]

\[ z^{k+1} = L^T (\mu u^k - \alpha^k) \]
2D deconvolution experiment

Deconvolution results in dB

<table>
<thead>
<tr>
<th></th>
<th>Gaussian Estimator</th>
<th>Laplace Estimator</th>
<th>Student’s Estimator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Astrocytes cells</td>
<td>12.18</td>
<td>10.48</td>
<td>10.52</td>
</tr>
<tr>
<td>Pulmonary cells</td>
<td>16.90</td>
<td>19.04</td>
<td>18.34</td>
</tr>
<tr>
<td>Stem cells</td>
<td>15.81</td>
<td>20.19</td>
<td>20.50</td>
</tr>
</tbody>
</table>

L : gradient
Optimized parameters

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;

Computed tomography (straight rays)

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

- Radon transform (line integrals)

\[
R_\theta \{ s(x) \}(t) = \int_{\mathbb{R}} s(t\theta + r\theta^\perp) dr = \int_{\mathbb{R}^2} s(x) \delta(t - \langle x, \theta \rangle) dx
\]

Equivalent analysis functions: \( \eta_m(x) = \delta(t_m - \langle x, \theta_m \rangle) \)

---

Computed tomography reconstruction results

![sinogram](image)

**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.

<table>
<thead>
<tr>
<th></th>
<th>Directions</th>
<th>Gaussian</th>
<th>Laplace</th>
<th>Student’s</th>
</tr>
</thead>
<tbody>
<tr>
<td>SL Phantom</td>
<td>120</td>
<td>16.8</td>
<td>17.53</td>
<td>18.76</td>
</tr>
<tr>
<td>SL Phantom</td>
<td>180</td>
<td>18.13</td>
<td>18.75</td>
<td>20.34</td>
</tr>
<tr>
<td>Lung</td>
<td>180</td>
<td>22.49</td>
<td>21.52</td>
<td>21.45</td>
</tr>
<tr>
<td>Lung</td>
<td>360</td>
<td>24.38</td>
<td>22.47</td>
<td>22.37</td>
</tr>
</tbody>
</table>

*: discrete gradient
EM: Single particle analysis

Cryo-electron micrograph

C.-O. Sorzano

Bovine papillomavirus

Number of pixels: 256 × 256 × 256
Resolution: 2.47 Å
Number of particles: 800
Type of symmetry: ii (60 fold symmetry)

Image reconstruction (real data)

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

6.185 Å

noisy projection of identical particles, with unknown orientations
image alignment and classification
CTF estimation and correction
Differential phase-contrast tomography

Mathematical model

\[ g(t, \theta) = \frac{\partial}{\partial t} R_{\theta}\{f\}(t) \]

\[ g = H s \]

\[ [H]_{(i,j),k} = \frac{\partial}{\partial t} P_{\theta j} \beta_k(t_j) \]

Experimental results

Rat brain reconstruction with 721 projections

ADMM-PCG (TV)  
FBP  

Collaboration: Prof. Marco Stampanoni, TOMCAT PSI / ETHZ
Reducing the numbers of views

Rat brain reconstruction with 181 projections

ADMM-PCG

SSIM = .49

SSIM = .96

SSIM = .95

g-FBP

SSIM = .15

SSIM = .51

SSIM = .60

Collaboration: Prof. Marco Stampanoni, TOMCAT PSI / ETHZ

(Nichian et al. Optics Express 2013)

Performance evaluation

Goldstandard: high-quality iterative reconstruction with 721 views

⇒ Reduction of acquisition time by a factor 10 (or more)?
CAN WE GO BEYOND MAP ESTIMATION?

- A detailed investigation of simpler denoising problems

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 \( y = x + n \)

\[ x_{\text{MMSE}}(y) = \mathbb{E}\{X|Y = y\} = \int_{\mathbb{R}} x \cdot p_{X|Y}(x|y) \, dx \]

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

Stein’s formula for AWGN

\[ x_{\text{MMSE}}(y) = y - \sigma^2 \Phi_Y'(y) \]

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

\[ g_\sigma: \text{Gaussian pdf (zero mean with variance } \sigma^2) \]
Iterative wavelet-based denoising: MAP → MMSE

Consistent Cycle Spinning (CCS) (Kamilov, IEEE-SPL 2012)

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

**Algorithm 3: CCS denoising** solves Problem (11.41) where $A$ is a tight frame matrix

**input:** $y, s^0 \in \mathbb{R}^N, \tau, \mu \in \mathbb{R}^+$

**set:** $k = 0, A^0 = 0, u = Ay$

**repeat**

$z^{k+1} = \text{prox}_{\frac{1}{1+\mu} \left( u + \mu A s^k + \lambda^k \right)}$

$s^{k+1} = A^\dagger (z^{k+1} - \frac{1}{\mu} A^k)$

$\lambda^{k+1} = \lambda^k - \mu (z^{k+1} - As^k)$

$k = k + 1$

until stopping criterion

return $s = s^k$

- CCS constraint: $z = As$ with $\|z\|_2^2 = M \|s\|_2^2$ (enforces energy conservation)

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

(Kazerouni, IEEE-SPL 2013) $\Rightarrow$ Iterative MMSE denoising

---

**Comparison of wavelet denoising strategies**

**Compound Poisson**

<table>
<thead>
<tr>
<th>Method</th>
<th>Δ SNR [dB]</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCS-MMSE</td>
<td>8</td>
</tr>
<tr>
<td>frame-MMSE</td>
<td>6</td>
</tr>
<tr>
<td>ortho-MMSE</td>
<td>4</td>
</tr>
<tr>
<td>ortho-ST</td>
<td>2</td>
</tr>
</tbody>
</table>

**Levy Flight**

<table>
<thead>
<tr>
<th>Method</th>
<th>Δ SNR [dB]</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCS-MMSE</td>
<td>4</td>
</tr>
<tr>
<td>CCS-MAP</td>
<td>3</td>
</tr>
<tr>
<td>frame-MMSE</td>
<td>2</td>
</tr>
<tr>
<td>frame-MAP</td>
<td>1</td>
</tr>
<tr>
<td>ortho-MMSE</td>
<td>0</td>
</tr>
<tr>
<td>ortho-MAP</td>
<td>-1</td>
</tr>
</tbody>
</table>

**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

- Algorithms and imaging applications

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

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

Cover design: Annette Unser

http://www.sparseprocesses.org