Dear Dr. Liebling,

I am pleased to inform you that you were selected to receive the 2004 Research Award of the Swiss Society of Biomedical Engineering for your thesis work "On Fresnelets, interference fringes, and digital holography". The award will be presented during the general assembly of the SSBE, September 3, Zurich, Switzerland.

Please let us know if
1) you will be present to receive the award,
2) you would be willing to give a 10 minutes presentation of the work during the general assembly.

The award comes with a cash prize of 1000.- CHF. Would you please send your banking information to the treasurer of the SSBE, Uli Diermann (Email: uli.diermann@bfh.ch), so that he can transfer the cash prize to your account?

I congratulate you on your achievement.

With best regards,
Michael Unser, Professor
Chairman of the SSBE Award Committee

cc: Ralph Mueller, president of the SSBE; Uli Diermann, treasurer

Dr. Michael Liebling
Biological Imaging Center
California Inst. of Technology
Mail Code 139-74
Pasadena, CA 91125, USA

BIOMEDICAL IMAGING GROUP (BIG)
LABORATOIRE D’IMAGERIE BIOMEDICALE
EPFL LIB
Bât. BM 4.127
CH 1015 Lausanne, Switzerland

Téléphone : + 41 21 693 51 85
Fax : + 41 21 693 37 01
E-mail : Michael.Unser@epfl.ch
Site web : http://bigwww.epfl.ch

June 2017

Paper print: 21st. June 2017

Lausanne, August 19, 2004

Dear Dr. Liebling,

I am pleased to inform you that you were selected to receive the 2004 Research Award of the Swiss Society of Biomedical Engineering for your thesis work "On Fresnelets, interference fringes, and digital holography". The award will be presented during the general assembly of the SSBE, September 3, Zurich, Switzerland.

Please let us know if
1) you will be present to receive the award,
2) you would be willing to give a 10 minutes presentation of the work during the general assembly.

The award comes with a cash prize of 1000.- CHF. Would you please send your banking information to the treasurer of the SSBE, Uli Diermann (Email: uli.diermann@bfh.ch), so that he can transfer the cash prize to your account?

I congratulate you on your achievement.

With best regards,
Michael Unser, Professor
Chairman of the SSBE Award Committee

cc: Ralph Mueller, president of the SSBE; Uli Diermann, treasurer

Dr. Michael Liebling
Biological Imaging Center
California Inst. of Technology
Mail Code 139-74
Pasadena, CA 91125, USA

BIOMEDICAL IMAGING GROUP (BIG)
LABORATOIRE D’IMAGERIE BIOMEDICALE
EPFL LIB
Bât. BM 4.127
CH 1015 Lausanne, Switzerland

Téléphone : + 41 21 693 51 85
Fax : + 41 21 693 37 01
E-mail : Michael.Unser@epfl.ch
Site web : http://bigwww.epfl.ch

June 2017

Paper print: 21st. June 2017

Lausanne, August 19, 2004

Plenary talk, Int. Conf. Acoust. Speech Sig. Proc. (ICASSP), March 20-25, 2016, Shanghai, China

Sparsity and inverse problems:
Think analog, act digital

Michael Unser
Biomedical Imaging Group
EPFL, Lausanne, Switzerland

Pareto distribution

\[
\Psi_{\alpha}(x) = \frac{x^\alpha}{\Gamma(\alpha+1)}
\]

Plenary talk, Int. Conf. Acoust. Speech Sig. Proc. (ICASSP), March 20-25, 2016, Shanghai, China
OUTLINE

- Brief history of inverse problems in imaging
  - Classical linear reconstruction methods
  - The sparsity (r)evolution
- Think analog: the spline connection
  - gTV and new optimality results
- Think analog & statistical
  - Introduction to sparse stochastic processes
- Act digital: Algorithm design
  - Reconstruction of biomedical images

Specific examples:
- Deconvolution microscopy
- Computed tomography
- Differential phase-contrast tomography

Inverse problems in bio-imaging

- Linear forward model
  \[ y = Hs + n \]

- The easy scenario
  - Inverse problem is well posed
  \[ s \approx H^{-1}y \]
- Backprojection (poor man’s solution)

Basic limitations

1) Inherent noise amplification
2) Difficulty to invert \( H \) (too large or non-square)
3) All interesting inverse problems are ill-posed
**Linear inverse problems (20th century theory)**

- Dealing with *ill-posed problems*: Tikhonov regularization

  \[ \mathcal{R}(s) = \|Ls\|_2^2 : \text{regularization (or smoothness) functional} \]

  \[ L : \text{regularization operator (i.e., Gradient)} \]

  \[ \min_s \|y - Hs\|_2^2 \quad \text{subject to} \quad \mathcal{R}(s) \leq C_0 \]

- Equivalent variational problem

  \[ s^* = \arg \min_s \underbrace{\|y - Hs\|_2^2}_{\text{data consistency}} + \underbrace{\lambda\|Ls\|_2^2}_{\text{regularization}} \]

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

  Interpretation: “filtered” backprojection

**Statistical formulation (20th century)**

- Linear measurement model: \[ y = Hs + n \]

  \( n \): additive white Gaussian noise (i. i. d.)

  \( s \): realization of Gaussian process with zero-mean and covariance matrix \( \mathbb{E}\{s \cdot s^T\} = C_s \)

- Wiener (LMMSE) solution = Gauss MMSE = Gauss MAP

  \[ s_{MAP} = \arg \min_s \frac{1}{\sigma^2} \|y - Hs\|^2_2 + \|C_s^{-1/2} s\|^2_2 \]

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

- Quadratic regularization (Tikhonov)

  \[ s_{Tik} = \arg \min_s \left( \|y - Hs\|^2_2 + \lambda \mathcal{R}(s) \right) \quad \text{with} \quad \mathcal{R}(s) = \|Ls\|^2_2 \]

  **Linear solution**: \[ s = (H^T H + \lambda L^T L)^{-1} H^T y = R_\lambda \cdot y \]
**Linear inverse problems:** The sparsity (r)evolution

(20th Century) $p = 2 \rightarrow 1$ (21st Century)

$s_{\text{rec}} = \arg \min_s (\| y - Hs \|_2^2 + \lambda R(s))$

- **Non-quadratic regularization regularization**
  $R(s) = \| Ls \|_2^2 \rightarrow \| Ls \|_p^p \rightarrow \| Ls \|_1$

- **Total variation** (Rudin-Osher, 1992)
  $R(s) = \| Ls \|_1$ with $L$: gradient

- **Wavelet-domain regularization** (Figuereido et al., Daubechies et al. 2004)
  $v = W^{-1}s$: wavelet expansion of $s$ (typically, sparse)
  $R(s) = \| v \|_1$

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

---

**Inverse problems in imaging:** Current status

- **Higher reconstruction quality**: Sparsity-promoting schemes almost systematically outperform the classical linear reconstruction methods in MRI, x-ray tomography, deconvolution microscopy, etc... (Lustig et al. 2007)

- **Increased complexity**: Resolution of linear inverse problems using $\ell_1$ regularization requires more sophisticated algorithms (iterative and non-linear); efficient solutions (FISTA, ADMM) have emerged during the past decade. (Chambolle 2004; Figueiredo 2004; Beck-Teboule 2009; Boyd 2011)

- The paradigm is supported by the theory of **compressed sensing** (Candes-Romberg-Tao; Donoho, 2006)

- **Outstanding research issues**
  - Beyond $\ell_1$ and TV: Connection with **statistical modeling & learning**
  - Beyond matrix algebra: **Continuous-domain** formulation
  - Guarantees of **optimality** (either deterministic or statistical)
Sparsity: Continuous-domain formulation

- Compressed sensing (CS)
  - Generalized sampling and infinite-dimensional CS (Adcock-Hansen, 2011)
  - Xampling: CS of analog signals (Eldar, 2011)

- Splines and approximation theory
  - $L_1$ splines (Fisher-Jerome, 1975)
  - Locally-adaptive regression splines (Mammen-van de Geer, 1997)
  - Generalized TV (Steidl et al. 2005; Bredies et al. 2010)

- Statistical modeling
  - Sparse stochastic processes (Unser et al. 2011-2014)

Think analog: The spline connection

Photo courtesy of Carl De Boor
Splines are analog and intrinsically sparse

\[ L\{\cdot\} : \text{admissible differential operator} \]
\[ \delta(\cdot - x_0) : \text{Dirac impulse shifted by } x_0 \in \mathbb{R}^d \]

**Definition**

The function \( s : \mathbb{R}^d \to \mathbb{R} \) is a (non-uniform) \( L \)-spline with knots \((x_k)_{k=1}^K\) if

\[
L\{s\} = \sum_{k=1}^{K} a_k \delta(\cdot - x_k) = w_\delta : \text{spline’s innovation}
\]

Spline theory: (Schultz-Varga, 1967)

- **FRI signal processing:** Innovation variables \((2K)\) (Vetterli et al., 2002)
  
  - Location of singularities (knots): \(\{x_k\}_{k=1}^{K}\)
  
  - Strength of singularities (linear weights): \(\{a_k\}_{k=1}^{K}\)

---

**Spline synthesis: example**

\[
L = D = \frac{d}{dx} \quad \text{Null space: } \mathcal{N}_D = \text{span}\{p_1\}, \quad p_1(x) = 1
\]

\[
\rho_D(x) = D^{-1}\{\delta\}(x) = \mathbb{1}_+(x) : \text{Heaviside function}
\]

\[
w_\delta(x) = \sum_k a_k \delta(x - x_k)
\]

\[
s(x) = b_1 p_1(x) + \sum_k a_k \mathbb{1}_+(x - x_k)
\]
**Spline synthesis: generalization**

$L$: spline admissible operator (LSI)

\[ \rho_L(x) = L^{-1}\{\delta\} : \text{Green’s function of } L \]

Finite-dimensional null space: \( \mathcal{N}_L = \text{span}\{p_n\}_n^{N_0} \)

Spline’s innovation:

\[ w_\delta(x) = \sum_k a_k \delta(x - x_k) \]

\[ \Rightarrow \quad s(x) = \sum_k a_k \rho_L(x - x_k) + \sum_{n=1}^{N_0} b_n p_n(x) \]

Requires specification of boundary conditions

**New optimality result: universality of splines**

$L$: spline-admissible operator

\[ \mathcal{M}_L(\mathbb{R}^d) = \{ f : gTV(f) = \|L\{f\}\|_{TV} = \sup_{\|\varphi\|_\infty \leq 1} \langle L\{f\}, \varphi \rangle < \infty \} \]

Generalized total variation:

\[ gTV(f) = \|L\{f\}\|_{L_1} \text{ when } L\{f\} \in L_1(\mathbb{R}^d) \]

Linear measurement operator \( \mathcal{M}_L(\mathbb{R}) \to \mathbb{R}^M : f \mapsto z = H\{f\} \)

\[ \Leftrightarrow \quad z_m = \langle h_m, f \rangle \]

**Theorem**: The generic linear-inverse problem

\[ \min_{f \in \mathcal{M}_L(\mathbb{R}^d)} \left( \|y - H\{f\}\|_2^2 + \lambda \|L\{f\}\|_{TV} \right) \]

admits global solution(s) of the form

\[ f(x) = \sum_{k=1}^K a_k \rho_L(x - x_k) + \sum_{n=1}^{N_0} b_n p_n(x) \]

with \( K \leq M - N_0 \), which is a non-uniform \( L \)-spline with knots \( (x_k)_{k=1}^K \).

(U.-Fageot-Ward, ArXiv 2016)
OUTLINE

- Brief history of inverse problems in imaging ✔
- **Think analog & deterministic ✔**
  Optimality of splines for gTV
- **Think analog & statistical**
  Introduction to sparse stochastic processes
- **Act digital: Algorithm design**
  Reconstruction of biomedical images

Specific examples:

- Deconvolution microscopy
- Computed tomography
- Differential phase-contrast tomography
Random spline: archetype of sparse signal

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

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

- Anti-derivative operators
  - Shift-invariant solution: \( D^{-1} \varphi(t) = (\mathbb{1}_+ \ast \varphi)(t) = \int_{-\infty}^{t} \varphi(\tau) d\tau \)
  - Scale-invariant solution: \( D_0^{-1} \varphi(t) = \int_{0}^{t} \varphi(\tau) d\tau \)

Compound Poisson process

- Stochastic differential equation
  \[ Ds(t) = w(t) \]
  with boundary condition \( s(0) = 0 \)
  \[ \text{Innovation: } w(t) = \sum_n a_n \delta(t - t_n) \]

- 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 \( \mathbb{E}\{w(t)w(t')\} = \sigma_w^2 \delta(t - t') \)

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

White noise (innovation)  \hspace{1cm} \text{Lévy process}

- Gaussian
- Impulsive \( w(t) \int_0^t d\tau \)
- \( \text{S\&S (Cauchy)} \)

(Wiener 1923)

(Paul Lévy circa 1930)

Generalized innovation model

Theoretical framework: Gelfand’s theory of generalized stochastic processes

Generic test function \( \varphi \in S \) plays the role of index variable

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

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

Proper definition of \textit{continuous-domain} white noise

(Unser et al, IEEE-IT 2014)

Regularization operator vs. wavelet analysis

Main feature: inherent sparsity
(few significant coefficients)
Probability laws of innovations are infinite divisible

\( w \) is a generalized innovation process (or continuous-domain white noise) in \( \mathcal{S}'(\mathbb{R}^d) \) if

1. **Observability**: \( X = \langle \varphi, w \rangle \) is an ordinary random variable for any \( \varphi \in \mathcal{S}(\mathbb{R}^d) \).

2. **Stationarity**: \( X_{x_0} = \langle \varphi(\cdot - x_0), w \rangle \) is identically distributed for all \( x_0 \in \mathbb{R}^d \).

3. **Independent atoms**: \( X_1 = \langle \varphi_1, w \rangle \) and \( X_2 = \langle \varphi_2, w \rangle \) are independent whenever \( \varphi_1 \) and \( \varphi_2 \) have non-intersecting support.

**Theorem** (under mild technical conditions) ([Amini-U., IEEE-IT 2014])

\( w \) is an innovation process in \( \mathcal{S}'(\mathbb{R}^d) \)

\[ \Rightarrow \ X = \langle \varphi, w \rangle \text{ is well defined and infinitely divisible for any } \varphi \in L_p(\mathbb{R}^d) \]

**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

\[ X \overset{d}{=} X_1 + \cdots + X_N. \]

\[ X = \langle w, \text{rect} \rangle = \left\langle \frac{1}{n}, \underbrace{\hfill \hfill \hfill \hfill \hfill}_{n} \right\rangle \]

\[ = \left\langle \frac{1}{n}, \underbrace{\hfill \hfill \hfill \hfill}_{n} \right\rangle + \cdots + \left\langle \frac{1}{n}, \underbrace{\hfill \hfill \hfill \hfill}_{n} \right\rangle \]

Probability laws of innovations are **infinite divisible**

- **Canonical observation through a rectangular test function**

  \[ X_{id} = \langle w, \text{rect} \rangle = \left\langle \frac{1}{n}, \underbrace{\hfill \hfill \hfill \hfill}_{n} \right\rangle \]

  \( w \) innovation process \( \iff \ X_{id} = \langle w, \text{rect} \rangle \) infinitely divisible

  with **canonical Lévy exponent**

  \[ f(\omega) = \log \tilde{p}_{id}(\omega) \]

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

  - **Generic observation**: \( X = \langle \varphi, w \rangle \) with \( \varphi \in L_p(\mathbb{R}^d) \)

  \[ X = \langle w, \varphi \rangle = \left\langle \frac{1}{n}, \underbrace{\hfill \hfill \hfill \hfill}_{n} \right\rangle \overset{\triangle}{=} \lim_{n \to \infty} \left\langle \frac{1}{n}, \underbrace{\hfill \hfill \hfill \hfill}_{n} \right\rangle \]

  \[ = \lim_{n \to \infty} \left\langle \frac{1}{n}, \underbrace{\hfill \hfill \hfill \hfill}_{n} \right\rangle + \cdots + \left\langle \frac{1}{n}, \underbrace{\hfill \hfill \hfill \hfill}_{n} \right\rangle \]

  - \( X \) is **infinitely divisible** with (modified) Lévy exponent

  \[ f_{\varphi}(\omega) = \log \tilde{p}_X(\omega) = \int_{\mathbb{R}^d} f(\omega \varphi(x)) \, dx \]
Examples of infinitely divisible laws

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

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

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

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

Characteristic function: \( \hat{p}_{id}(\omega) = \int_{\mathbb{R}} p_{id}(x) e^{i\omega x} dx = c^{\omega} \)
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

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

(d) Cauchy (stable)

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

Complete mathematical characterization: \( \mathcal{P}_w(\varphi) = \exp \left( \int_{\mathbb{R}^d} f(\varphi(x)) \, dx \right) \)

Aesthetic sparse signal: the Mondrian process

\[ L = D_x D_y \quad \overset{\mathcal{F}}{\leftrightarrow} \quad (j\omega_x)(j\omega_y) \]

\[ \lambda = 30 \]
Scale- and rotation-invariant processes

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

Gaussian

Sparse (generalized Poisson)

Powers of ten: from astronomy to biology
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*)

**OUTLINE**

- Brief history of inverse problems in imaging ✔
- **Think analog & deterministic ✔**
  - Optimality of splines for gTV
- **Think analog & statistical ✔**
  - Sparse stochastic processes
- **Act digital**: Reconstruction of biomedical images
  - Discretization of reconstruction problem
  - Signal reconstruction algorithm (MAP)
  - Examples of image reconstruction
Discretization of reconstruction problem

Spline-like reconstruction model: \( s(r) = \sum_{k \in \Omega} s[k] \beta_k(r) \) \( \leftrightarrow \) \( s = (s[k])_{k \in \Omega} \)

- Innovation model

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

\( p_U \) is part of infinitely divisible family

- Physical model: image formation and acquisition

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

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

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

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

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

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

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

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

Proximal operator: **pointwise denoiser**

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

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

- linear attenuation
- soft-threshold
- shrinkage function

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

- Constrained optimization formulation

Auxiliary innovation variable: \( u = Ls \)

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

- Augmented Lagrangian method

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

Lagrange multiplier vector: \( \mathbf{\alpha} \)

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

Alternating direction method of multipliers (ADMM)

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

Sequential minimization

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

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

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

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

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

Nonlinear denoising: \( \mathbf{u}^{k+1} = \text{prox}_{\Phi_U} (\mathbf{Ls}^{k+1} + \frac{1}{\mu} \mathbf{\alpha}^{k+1}; \frac{\sigma^2}{\mu}) \)

- Proximal operator tailored to stochastic model

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

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

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

Optical parameters

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

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.

<table>
<thead>
<tr>
<th></th>
<th>BSNR (dB)</th>
<th>Estimation performance (SNR in dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Gaussian</td>
</tr>
<tr>
<td>Stem cells</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>14.43</td>
<td>13.76</td>
</tr>
<tr>
<td>30</td>
<td>15.92</td>
<td>15.77</td>
</tr>
<tr>
<td>40</td>
<td>18.11</td>
<td>18.11</td>
</tr>
<tr>
<td>Nerve cells</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>13.86</td>
<td>15.31</td>
</tr>
<tr>
<td>30</td>
<td>15.89</td>
<td>18.18</td>
</tr>
<tr>
<td>40</td>
<td>18.58</td>
<td>20.57</td>
</tr>
<tr>
<td>Artery cells</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>14.86</td>
<td>15.23</td>
</tr>
<tr>
<td>30</td>
<td>16.59</td>
<td>17.21</td>
</tr>
<tr>
<td>40</td>
<td>18.68</td>
<td>19.61</td>
</tr>
</tbody>
</table>

L: discrete gradient

37
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)

$$R_\theta \{s(x)\}(t) = \int_R s(t\theta + r\theta^\perp)dr$$

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

Sinogram

(appealable to tomographic phase microscopy with plane wave illumination)

Equivalent analysis functions: $\eta_m(x) = \delta(t_m - \langle x, \theta_m \rangle)$
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>Directions</th>
<th>SL Phantom 120</th>
<th>SL Phantom 180</th>
<th>Lung 180</th>
<th>Lung 360</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian</td>
<td>16.8</td>
<td>18.13</td>
<td>22.49</td>
<td>24.38</td>
</tr>
<tr>
<td>Laplace</td>
<td>17.53</td>
<td>18.75</td>
<td>21.52</td>
<td>22.47</td>
</tr>
<tr>
<td>Student's</td>
<td>18.76</td>
<td>20.34</td>
<td>21.45</td>
<td>22.37</td>
</tr>
</tbody>
</table>

For the reconstruction, we solve the quadratic minimization problem (10.21) iteratively by using 50 conjugate-gradient (inner) iterations. The reconstruction results are reported in Table 10.4.

We observe that the imposition of the strong level of sparsity brought by Student's priors is advantageous for the SL phantom. This is not overly surprising given that the SL phantom is an artificial construct composed of piecewise-constant regions (ellipses). For the realistic lung image (true CT), we find that the Gaussian solution outperforms the others. Similarly to the deconvolution and MRI problems, the present MAP estimators are in line with the Tikhonov-type [WLLL06] and TV [XQJ05] reconstructions used for X-ray CT.
SUMMARY: Sparsity in infinite dimensions

- Continuous-domain formulation
  - Linear measurement model
  - Linear signal model: PDE
  - $L$-splines = signals with “sparsest” innovation $\Rightarrow s = L^{-1}w$

- Deterministic optimality result
  - $gTV$ regularization: favors “sparse” innovations
  - Non-uniform $L$-splines: universal solutions of linear inverse problems

- Statistical model that supports sparsity
  - Statistical **decoupling**:
    - **Gaussian** vs. **sparse** innovations (Poisson, student, $S\alpha S$)
  - Unifying framework: “sparse stochastic processes” $s = L^{-1}w$
  - MAP enforces sparsity through non-quadratic regularization

Acknowledgments

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

- Dr. Pouya Tafti
- Prof. Arash Amini
- Dr. John-Paul Ward
- Julien Fageot
- Dr. Emrah Bostan
- Dr. Masih Nilchian
- Dr. Ulugbek Kamilov
- Dr. Cédric Vonesch
- ....

and collaborators ...

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

- Preprints and demos: [http://bigwww.epfl.ch/](http://bigwww.epfl.ch/)
Gaussian vs. Sparse

Fourier analysis

Splines

Wavelet analysis

Norbert Wiener

Isaac Schoenberg

Paul Lévy

References

- Theory of sparse stochastic processes

- $L_1$-optimality of splines for inverse problems

- Algorithms and imaging applications