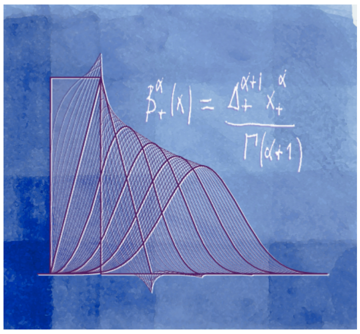
ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

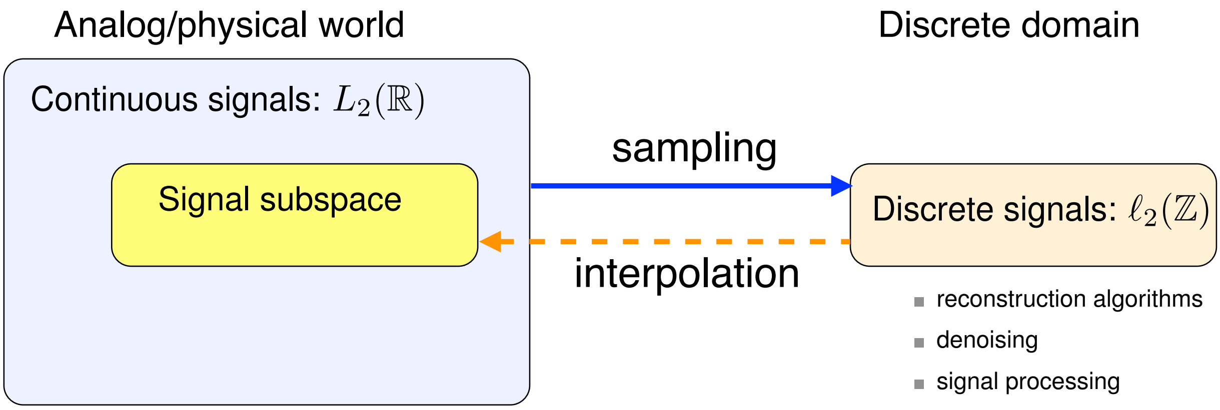
Sampling: 60 years after Shannon

Michael Unser
Biomedical Imaging Group
EPFL, Lausanne
Switzerland



Plenary talk, DSP2009, Santorini, Greece, July 2009

SAMPLING: 50 years after Shannon

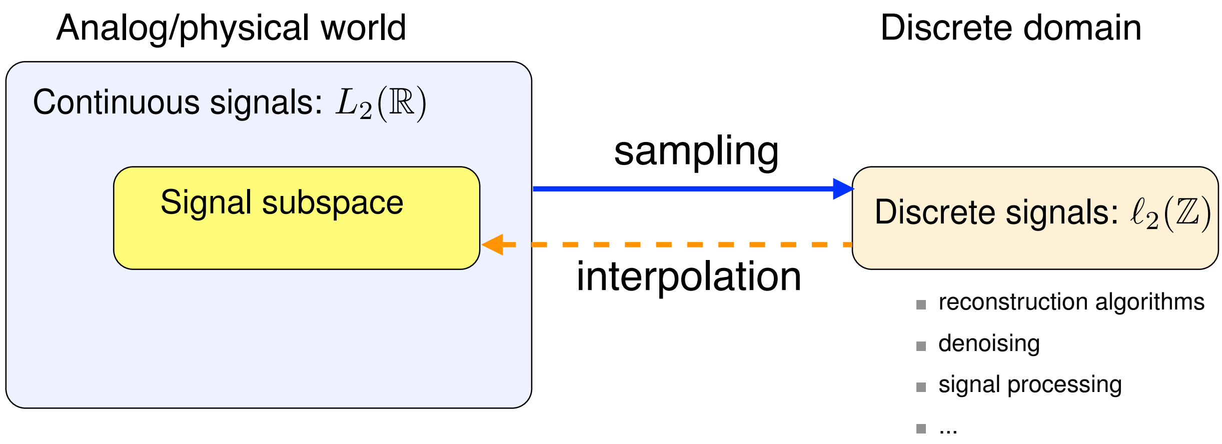


- reconstruction algorithms
- denoising
- signal processing
- ...

Universal

- Shannon's sampling theory (40's)
- Generalized sampling using splines (70's and 90's)
- Consistent sampling (noise-free scenario) (90's)

SAMPLING: 60 years after Shannon



Universal

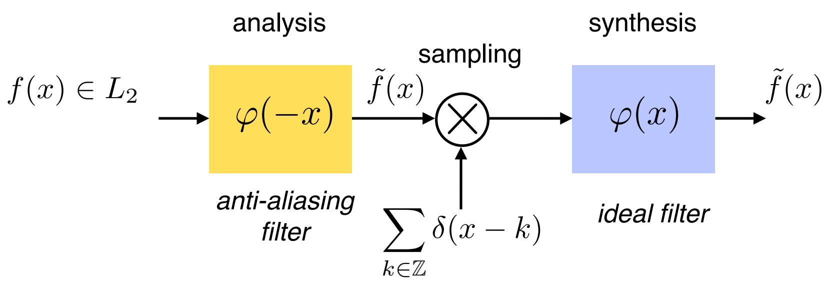
- Shannon's sampling theory (40's)
- Generalized sampling using splines (70's and 90's)
- Consistent sampling (noise-free scenario) (90's)

Constrained
(prior knowledge)

- Regularized sampling (for noisy data) (00's)
- Sampling with sparsity constraints ...

Shannon's sampling reinterpreted

- Generating function: $\varphi(x) = \mathrm{sinc}(x)$
- Subspace of bandlimited functions: $V(\varphi) = \mathrm{span}\{\varphi(x-k)\}_{k\in\mathbb{Z}}$



- Analysis: $\tilde{f}(k) = \langle \mathrm{sinc}(x-k), f(x)\rangle$
- Synthesis: $\tilde{f}(x) = \sum_{k\in Z} \tilde{f}(k)\, \mathrm{sinc}(x-k)$
- Orthogonal basis: $\langle \mathrm{sinc}(x-k), \mathrm{sinc}(x-l)\rangle = \delta_{k-l}$

→ Orthogonal projection operator !

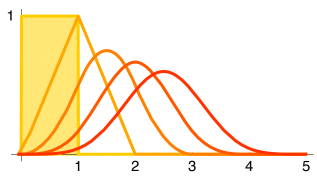
Fundamental sampling questions

- Q1: Are there alternative choices of representations?
  ANSWER: Yes, of course!
  Specification of reconstruction space
- Q2: How good is the representation/signal reconstruction?
  ANSWER: Approximation theory
  Rate of decay of the error as sampling step goes to zero
- Q3: How should we formulate the reconstruction problem?
  - Noise-free: consistent (but, not exact) reconstruction
  - Noisy data: regularized sampling
    smoothness and/or sparsity constraints
- Q4: Can we design fast/efficient algorithms?
- Q5: Can we specify optimal reconstruction spaces/solutions?
  ANSWER: Yes, under specific conditions
- Q6: Should we redesign the whole system?
  Compressive sensing ...

Part 1: Sampling theory and splines

- More general generating function

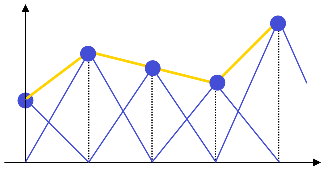
$$\mathrm{sinc}(x) \quad \rightarrow \quad \varphi(x) \quad \rightarrow \quad \beta^n(x)$$ (polynomial B-spline of degree n)



- Justifications for using (B-)splines
  - Ease of use: short, piecewise-polynomial basis functions
  - Generality: progressive transition from piecewise-constant ($n = 0$) to bandlimitted ($n \rightarrow \infty$)
  - Improved performance: best cost/quality tradeoff
  - Optimal from a number of perspectives
    - Approximation theory: shortest basis functions for a given order of approximation
    - Link with differential operators (Green functions)
    - Variational properties
    - Minimum Mean Square Error estimators for certain classes of stochastic processes
    - Fundamental role in wavelet theory

PRELIMINARIES

- Function and sequence spaces
- Shift-invariant subspaces
- Splines and operators



Continuous-domain signals

Mathematical representation: a function of the continuous variable $x \in \mathbb{R}$

- Lebesgue's space of finite-energy functions
  - $L_2(\mathbb{R}) = \left\{ f(x), x \in \mathbb{R} : \int_{x \in \mathbb{R}} |f(x)|^2 \mathrm{d}x < +\infty \right\}$
  - L_2-inner product: $\langle f, g \rangle = \int_{x \in \mathbb{R}} f(x) g^*(x) \mathrm{d}x$
  - L_2-norm: $\|f\|_{L_2} = \left(\int_{x \in \mathbb{R}} |f(x)|^2 \mathrm{d}x \right)^{1/2} = \sqrt{\langle f, f \rangle}$
- Fourier transform
  - Integral definition: $\hat{f}(\omega) = \int_{x \in \mathbb{R}} f(x) e^{-j\omega x} \mathrm{d}x$
  - Parseval relation: $\|f\|_{L_2}^2 = \frac{1}{2\pi} \int_{\omega \in \mathbb{R}} |\hat{f}(\omega)|^2 \mathrm{d}\omega$

Discrete-domain signals

Mathematical representation: a sequence indexed by the discrete variable $k \in \mathbb{Z}$

- Space of finite-energy sequences
  - $\ell_2(\mathbb{Z}) = \left\{ a[k], k \in \mathbb{Z} : \sum_{k \in \mathbb{Z}} |a[k]|^2 < +\infty \right\}$
  - ℓ_2-norm: $\|a\|_{\ell_2} = \left(\sum_{k \in \mathbb{Z}} |a[k]|^2 \right)^{1/2}$

- Discrete-time Fourier transform
  - z-transform: $A(z) = \sum_{k \in \mathbb{Z}} a[k] z^{-k}$
  - Fourier transform: $A(e^{j\omega}) = \sum_{k \in \mathbb{Z}} a[k] e^{-j\omega k}$

Shift-invariant spaces

Integer-shift-invariant subspace associated with a generating function φ (e.g., B-spline):

$$V(\varphi) = \left\{ f(x) = \sum_{k \in \mathbb{Z}} c[k] \varphi(x-k) : c \in \ell_2(\mathbb{Z}) \right\}$$

Generating function: $\varphi(x) \quad \stackrel{\mathcal{F}}{\longleftrightarrow} \quad \hat{\varphi}(\omega) = \int_{x \in \mathbb{R}} \varphi(x) e^{-j\omega x} \mathrm{d}x$

- Autocorrelation (or Gram) sequence

$$a_\varphi[k] \triangleq \langle \varphi(\cdot), \varphi(\cdot - k) \rangle \quad \stackrel{\mathcal{F}}{\longleftrightarrow} \quad A_\varphi(e^{j\omega}) = \sum_{n \in \mathbb{Z}} |\hat{\varphi}(\omega + 2\pi n)|^2$$

- Riesz-basis condition

Positive-definite Gram sequence: $0 < A^2 \le A_\varphi(e^{j\omega}) \le B^2 < +\infty$

$$\Updownarrow$$

$$A \cdot \|c\|_{\ell_2} \le \underbrace{\left\| \sum_{k \in \mathbb{Z}} c[k] \varphi(x-k) \right\|_{L_2}}_{\|f\|_{L_2}} \le B \cdot \|c\|_{\ell_2}$$

Orthonormal basis $\Leftrightarrow$ $a_\varphi[k] = \delta_k$ $\Leftrightarrow$ $A_\varphi(e^{j\omega}) = 1$ $\Leftrightarrow$ $\|c\|_{\ell_2} = \|f\|_{L_2}$ (Parseval)

Example of reconstruction spaces

- Piecewise-constant functions

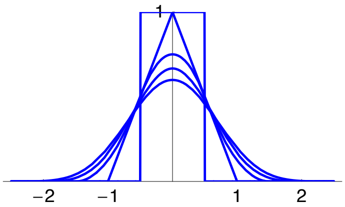
$\varphi(x) = \mathrm{rect}(x) = \beta^0(x)$ $\quad a_\varphi[k] = \delta_k \quad \Leftrightarrow \quad$ the basis is orthonormal

- bandlimited functions

$\varphi(x) = \mathrm{sinc}(x)$ $\quad \sum_{n\in\mathbb{Z}} |\hat{\varphi}(\omega + 2\pi n)|^2 = 1 \quad \Leftrightarrow \quad$ the basis is orthonormal

- Polynomial splines of degree n

$$\varphi(x) = \beta^n(x) = (\underbrace{\beta^0 * \beta^0 \cdots * \beta^0}_{(n+1) \text{ times}})(x)$$



Autocorrelation sequence: $\quad a_{\beta^n}[k] = (\beta^n * \beta^n)(x)|_{x=k} = \beta^{2n+1}(k)$

Proposition. The B-spline of degree n, $\beta^n(x)$, generates a Riesz basis with lower and upper Riesz bounds $A = \inf_\omega\{A_{\beta^n}(e^{j\omega})\} \geq \left(\frac{2}{\pi}\right)^{n+1}$ and $B = \sup_\omega\{A_{\beta^n}(e^{j\omega})\} = 1$.

Cardinal *L*-splines

$\mathrm{L}\{\cdot\}$: differential operator (translation-invariant)

$\delta(x)$: Dirac distribution

Definition

The continuous-domain function $s(x)$ is called a cardinal L-spline iff.

$$\mathrm{L}\{s\}(x) = \sum_{k\in\mathbb{Z}} a[k]\delta(x-k)$$

- Location of singularities = spline knots (integers)
- Generalization: includes polynomial splines as particular case ($\mathrm{L} = \frac{\mathrm{d}^N}{\mathrm{d}x^N}$)

Example: piecewise-constant splines

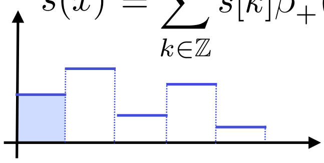
- Spline-defining operators

Continuous-domain derivative: $\mathrm{D} = \frac{\mathrm{d}}{\mathrm{d}x} \longleftrightarrow j\omega$

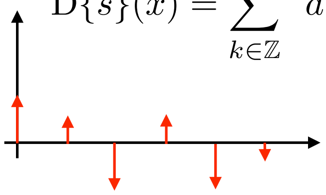
Discrete derivative: $\Delta_+\{\cdot\} \longleftrightarrow 1 - e^{-j\omega}$

- Piecewise-constant or D-spline

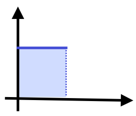
$$s(x) = \sum_{k\in\mathbb{Z}} s[k]\beta_+^0(x-k)$$



$$\mathrm{D}\{s\}(x) = \sum_{k\in\mathbb{Z}} \overbrace{a[k]}^{\Delta_+ s(k)}\,\delta(x-k)$$



- B-spline function

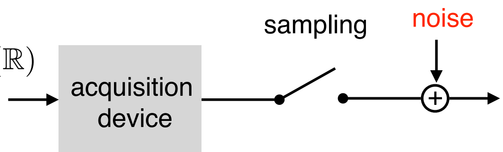


$$\beta_+^0(x) = \Delta_+\mathrm{D}^{-1}\{\delta\}(x) \longleftrightarrow \frac{1-e^{-j\omega}}{j\omega}$$

Basic sampling problem

- Sampling system

$f(x) \in L_2(\mathbb{R})$



Discrete measurements:

$$g[k] = (h * f)(x)|_{x=k} + n[k]$$

Reconstruction algorithm

Constraints (prior knowledge)

signal coefficients

$\{c[k]\}_{k\in\mathbb{Z}}$

Riesz-basis property

Continuous-domain reconstruction

$$\tilde{f}(x) = \sum_{k\in\mathbb{Z}} c[k]\varphi(x-k)$$

Goal: Specify a set of constraints, a reconstruction space and a reconstruction algorithm so that $\tilde{f}(x)$ is a good approximation of $f(x)$

VARIATIONAL RECONSTRUCTION

- Regularized interpolation
- Generalized smoothing splines
- Optimal reconstruction space
- Splines and total variation

Regularized interpolation (Ideal sampler)

Given the noisy data $g[k] = f(k) + n[k]$, obtain an estimation $\tilde{f}$ of f that is

1. (piecewise-) smooth to reduce the effect of noise (regularization)
2. consistent with the given data (data fidelity)

- Variational formulation

$$f_\lambda = \arg \min_{f \in V(\varphi)} J(f, g; \lambda),$$

$$J(f, g; \lambda) = \underbrace{\sum_{k \in \mathbb{Z}} |g[k] - f(k)|^2}_{\text{Data Fidelity Term}} + \lambda \underbrace{\int_{\mathbb{R}} \Phi(|\mathrm{L}\{f\}(x)|) \, \mathrm{d}x}_{\text{Regularization}}$$

- L: Differential operator used to quantify lack of smoothness; e.g., $\mathrm{D} = \frac{\mathrm{d}}{\mathrm{d}x}$ or D^2
- $\Phi(\cdot)$: Increasing potential function used to penalize non-smooth solutions (e.g., $\Phi(u) = |u|^2$)
- $\lambda \geq 0$: Regularization parameter to strike a balance between “smoothing” and “consistency”

Regularized fit: Smoothing splines

Theorem: The solution (among all functions) of the smoothing spline problem

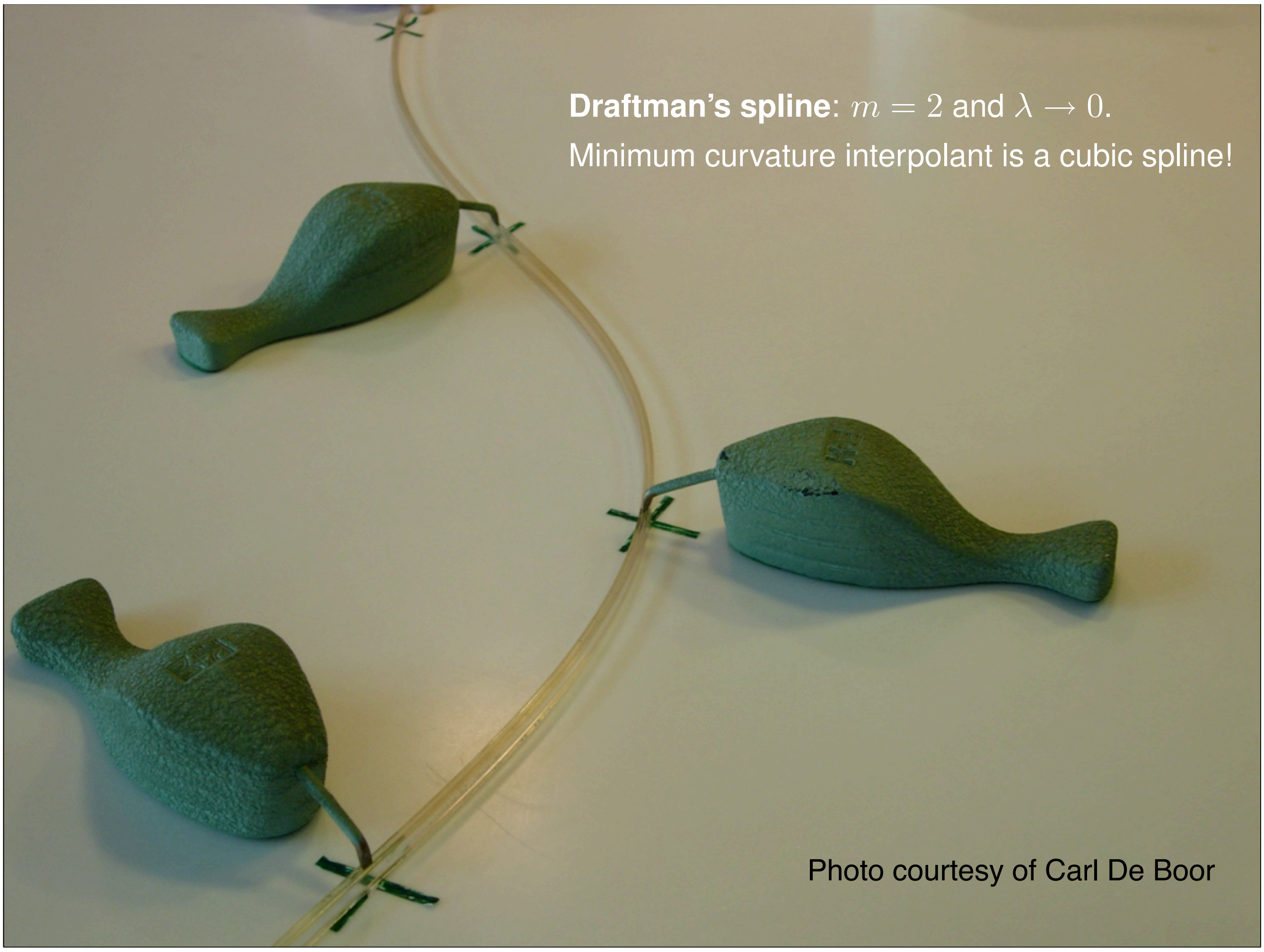
$$\min_{f(x)} \left\{ \sum_{k\in\mathbb{Z}} |g[k]-f(k)|^2 + \lambda \int_{-\infty}^{+\infty} |\mathrm{D}^m f(x)|^2 dx \right\}$$

is a cardinal polynomial spline of degree $2m-1$. Morever, its B-spline coefficients can be obtained by suitable recursif filtering of the input samples $g[k]$.

[Schoenberg, 1973; U., 1992]

Polynomial spline reconstruction: $f_\lambda(x) = \sum_{k\in\mathbb{Z}} c[k]\beta^n(x-k)$

Discrete, noisy input: $g[k] = f(k) + n[k]$ → **Smoothing spline filter** → $c[k] = (h_\lambda * g)[k]$



Generalized smoothing spline

L: Spline-defining differential operator

Theorem: The solution (among all functions) of the smoothing spline problem

$$\min_{f(x)} \left\{ \sum_{k\in\mathbb{Z}} |g[k] - f(k)|^2 + \lambda \int_{-\infty}^{+\infty} |\mathrm{L}f(x)|^2 \mathrm{d}x \right\}$$

is a cardinal $\mathrm{L}^*\mathrm{L}$ spline. The solution can calculated as

$$f_\lambda(x) = \sum_{k\in\mathbb{Z}} (h_\lambda * g)[k] \varphi_{\mathrm{L}}(x-k)$$

where φ_{L} is an "optimal" B-spline generator and h_λ a corresponding digital reconstruction filter parametrized by λ.

[U.-Blu, IEEE-SP, 2005]

Variational reconstruction: optimal discretization

Definition: φ_{L} is an **optimal generator** with respect to L iff

- it generates a shift-invariant Riesz basis $\{\varphi_{\mathrm{L}}(x-k)\}_{k\in\mathbb{Z}}$
- φ_{L} is a cardinal $\mathrm{L}^*\mathrm{L}$-spline; i.e., there exists a sequence $q[k]$ s.t.

  $$\mathrm{L}^*\mathrm{L}\{\varphi_{\mathrm{L}}\}(x) = \sum_{k\in\mathbb{Z}} q[k]\delta(x-k).$$

Short support: φ_{L} can be chosen of size $2N$ where N is the order of the operator

- Optimal digital reconstruction filter

$$H_\lambda(z) = \frac{1}{B_{\mathrm{L}}(z) + \lambda Q(z)} \qquad \text{with} \qquad B_{\mathrm{L}}(z) = \sum_{k\in\mathbb{Z}} \varphi_{\mathrm{L}}(k) z^{-k}$$

Stochastic optimality of splines

- Stationary processes
  - A smoothing spline estimator provides the MMSE estimation of a continuously-defined signal $f(\boldsymbol{x})$ given its noisy samples iff L is the whitening operator of the process and $\lambda = \frac{\sigma^2}{\sigma_0^2}$ [Unser-Blu, 2005].
  - Advantages: the spline machinery often yields a most efficient implementation: shortest basis functions (B-splines) together with recursive algorithms (especially in 1D).

- Fractal processes
  - Fractional Brownian motion (fBm) is a self-similar process of great interest for the modeling of natural signals and images. fBms are non-stationary, meaning that the Wiener formalism is not applicable (their power spectrum is not defined !).
  - Yet, using a distributional formalism (Gelfand's theory of generalized stochastic processes), it can be shown that these are whitened by fractional derivatives.
  - The optimal MSE estimate of a fBm with Hurst exponent H is a fractional smoothing spline of order $\gamma = 2H + 1$: $\hat{L}(\omega) = (j\omega)^{\gamma/2}$ [Blu-Unser, 2007].
  - Special case: the MMSE estimate of the Wiener process (Brownian motion) is a linear spline ($\gamma = 2$).

Generalization: non-quadratic data term

- General cost function with quadratic regularization

$$J(f,g) = J_{\text{data}}(f,g) + \lambda \|\mathrm{L}f\|^2_{L_2(\mathbb{R}^d)}$$

$J_{\text{data}}(f,g)$: arbitrary, but depends on the input data $g[k]$ and the samples $\{f(k)\}_{k\in\mathbb{Z}}$ only

Theorem. If φ_{L} is optimum with respect to L and a solution exists, then the optimum reconstruction over ALL continuously-defined functions f is such that

$$\min_f J(f,g) = \min_{f\in V(\varphi_{\mathrm{L}})} J(f,g).$$

Hence, there is an optimal solution of the form $\sum_{k\in\mathbb{Z}} c[k]\varphi_{\mathrm{L}}(x-k)$ that can be found by DISCRETE optimization.

Note: similar optimality results apply for the non-ideal sampling problem

[Ramani-U., IEEE-IP, 2008]

Splines and total variation

- Variational formulation with TV-type regularization

$$\tilde{f} = \arg \min_{f \in L_2(\mathbb{R})} J(f, g),$$

$$J(f,g) = \underbrace{\sum_{k\in\mathbb{Z}} |g[k] - f(k)|^2}_{\text{Data Fidelity Term}} + \lambda \underbrace{\int_{\mathbb{R}} |\mathrm{D}^n\{f\}(x)|^{1}\, \mathrm{d}x}_{\mathrm{TV}\{\mathrm{D}^{n-1}f\}}$$

Theorem: The above optimization problem admits a solution that is a ***non-uniform*** spline of degree $n-1$ with ***adaptive*** knots.

[Mammen, Van de Geer, *Annals of Statistics*, 1997]

More complex algorithm (current topic of research)

Part 2: From smoothness to sparsity

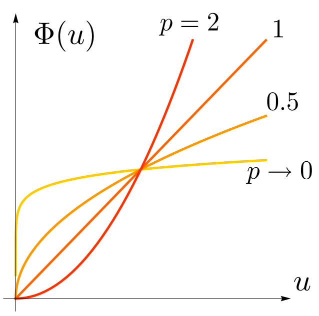
- Choices of regularization functionals
  - Aim: Penalize non-smooth (or highly oscillation) solutions
  - Limitation of quadratic regularization: over-penalizes sharp signal transitions

Signal domain

$\|\mathrm{L}f\|_2^2$ (Sobolev-type norm) → $\|\mathrm{L}f\|_1$

e.g., $\|\mathrm{D}f\|_1 = \mathrm{TV}\{f\}$ (total variation)

Wavelet domain



$\|\mathcal{W}f\|_1 \sim \|f\|_{B_1^1(L_1(\mathbb{R}))}$ (Besov norm)

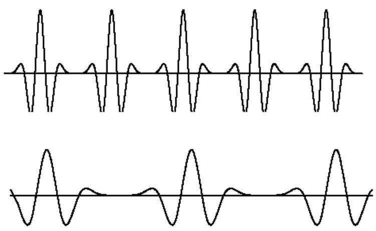
↓ **Compressive sensing theory**

$\|\mathcal{W}f\|_0$ Sparsity index (non-convex)

SAMPLING AND SPARSITY

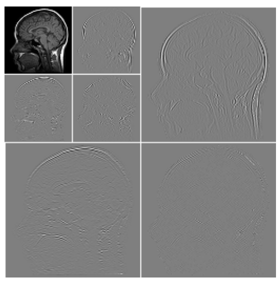
- Wavelets yield sparse representations
- Theory of compressive sensing
- Wavelet-regularized solution of general linear inverse problems
- Biomedical imaging examples
  - 3D deconvolution
  - Parallel MRI

Wavelet bases of L_2



- Family of wavelet templates (basis functions)

$$\psi_{i,k}(x) = 2^{-i/2}\psi\left(\frac{x - 2^i k}{2^i}\right)$$

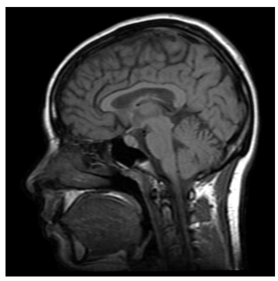


- Orthogonal wavelet basis

$$\langle \psi_{i,k}, \psi_{j,l} \rangle = \delta_{i-j,k-l} \quad \Leftrightarrow \quad \mathbf{W}^{-1} = \mathbf{W}^T$$

Analysis: $w_i[k] = \langle f, \psi_{i,k} \rangle$ (wavelet coefficients)

Reconstruction: $\forall f(x) \in L_2(\mathbb{R}), \quad f(x) = \sum_{i\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} w_i[k]\ \psi_{i,k}(x)$



- Vector/matrix notation

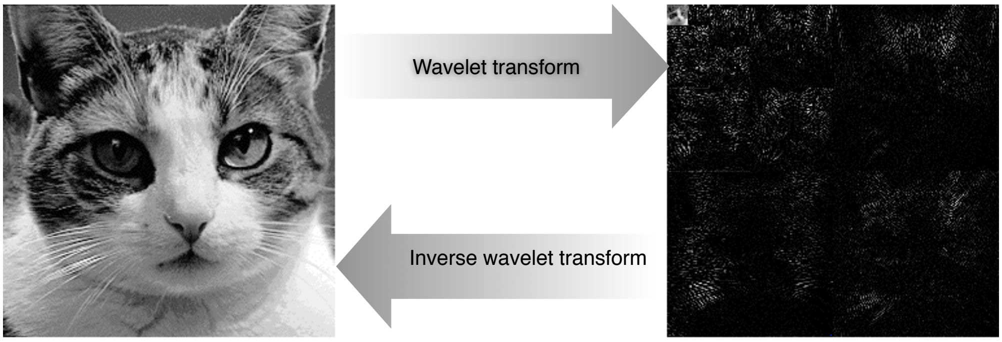
Discrete signal: $\mathbf{f} = (\cdots, c[0], c[1], c[2], \cdots)$

Wavelet coefficients: $\mathbf{w} = (\cdots, w_1[0], w_1[1], \cdots, w_2[0], \cdots)$

Analysis formula: $\mathbf{w} = \mathbf{W}^T\mathbf{f}$

Synthesis formula: $\mathbf{f} = \mathbf{W}\mathbf{w} = \sum_k w_k \boldsymbol{\psi}_k$

Wavelets yield sparse decompositions



66.4 dB

0.00%

Discarding “small coefficients”

Theory of compressive sensing

- Generalized sampling setting (after discretization)
  - Linear inverse problem: $\mathbf{u} = \mathbf{H}\mathbf{f} + \mathbf{n}$
  - Sparse representation of signal: $\mathbf{f} = \mathbf{W}^T\mathbf{v}$ with $\|\mathbf{v}\|_0 = K \ll N_v$
  - $N_u \times N_v$ system matrix : $\mathbf{A} = \mathbf{H}\mathbf{W}^T$

- Formulation of ill-posed recovery problem when $2K < N_u \ll N_v$

(P0) $\min_{\mathbf{v}} \|\mathbf{u} - \mathbf{A}\mathbf{v}\|_2^2$ subject to $\|\mathbf{v}\|_0 \leq K$

- Theoretical result

Under suitable conditions on $\mathbf{A}$ (e.g., restricted isometry), the solution is unique and the recovery problem (P0) is equivalent to:

(P1) $\min_{\mathbf{v}} \|\mathbf{u} - \mathbf{A}\mathbf{v}\|_2^2$ subject to $\|\mathbf{v}\|_1 \leq C_1$

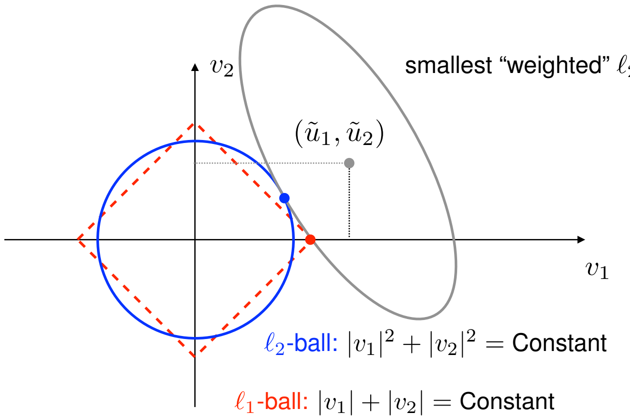
[Donoho et al., 2005
Candès-Tao, 2006, ...]

Sparsity and l_1-minimization

- Prototypical inverse problem

$$\min_{\mathbf{v}} \left\{ \|\mathbf{u} - \mathbf{A}\mathbf{v}\|_{\ell_2}^2 + \lambda \|\mathbf{v}\|_{\ell_2}^2 \right\} \quad \Leftrightarrow \quad \min_{\mathbf{v}} \|\mathbf{u} - \mathbf{A}\mathbf{v}\|_{\ell_2}^2 \text{ subject to } \|\mathbf{v}\|_{\ell_2} = C_2$$

$$\min_{\mathbf{v}} \left\{ \|\mathbf{u} - \mathbf{A}\mathbf{v}\|_{\ell_2}^2 + \lambda \|\mathbf{v}\|_{\ell_1} \right\} \quad \Leftrightarrow \quad \min_{\mathbf{v}} \|\mathbf{u} - \mathbf{A}\mathbf{v}\|_{\ell_2}^2 \text{ subject to } \|\mathbf{v}\|_{\ell_1} = C_1$$



Elliptical norm: $\|\mathbf{u} - \mathbf{A}\mathbf{v}\|_2^2 = (\mathbf{v} - \tilde{\mathbf{u}})^T \mathbf{A}^T \mathbf{A} (\mathbf{v} - \tilde{\mathbf{u}})$ with $\tilde{\mathbf{u}} = \mathbf{A}^{-1}\mathbf{u}$

Solving general linear inverse problems

- Space-domain measurement model

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

$\mathbf{H}$: system matrix (image formation)

$\mathbf{n}$: additive noise component

- Wavelet-regularized signal recovery
  - Wavelet expansion of signal: $\tilde{\mathbf{f}} = \mathbf{W}\tilde{\mathbf{w}}$
  - Data term: $\|\mathbf{g} - \mathbf{H}\tilde{\mathbf{f}}\|_2^2 = \|\mathbf{g} - \mathbf{H}\mathbf{W}\tilde{\mathbf{w}}\|_2^2$
  - Wavelet-domain sparsity constraint: $\|\tilde{\mathbf{w}}\|_{\ell_1} \leq C_1$

Convex optimization problem

$$\tilde{\mathbf{w}} = \arg\min_{\mathbf{w}} \left\{ \|\mathbf{g} - \mathbf{A}\mathbf{w}\|_2^2 + \lambda \|\mathbf{w}\|_{\ell_1} \right\} \quad \text{with } \mathbf{A} = \mathbf{H}\mathbf{W}$$

or

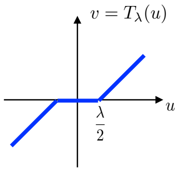
$$\tilde{\mathbf{f}} = \arg\min_{\mathbf{f}} \left\{ \|\mathbf{g} - \mathbf{H}\mathbf{f}\|_2^2 + \lambda \|\mathbf{W}^T \mathbf{f}\|_{\ell_1} \right\}$$

Alternating minimization: ISTA

- Convex cost functional: $J(\mathbf{f}) = \|\mathbf{g} - \mathbf{H}\mathbf{f}\|_2^2 + \lambda\|\mathbf{W}^T\mathbf{f}\|_1$
- Special cases
  - Classical least squares: $\lambda = 0 \quad \Rightarrow \quad \mathbf{f} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{g}$

    Landweber algorithm: $\mathbf{f}_{n+1} = \mathbf{f}_n + \gamma\mathbf{H}^T(\mathbf{g} - \mathbf{H}\mathbf{f}_n)$ (steepest descent)
  - Pure denoising: $\mathbf{H} = \mathbf{I} \quad \Rightarrow \quad \mathbf{f} = \mathbf{W}\, T_\lambda\{\mathbf{W}^T\mathbf{g}\}$ (Chambolle et al., *IEEE-IP* 1998)
- Iterative Shrinkage-Thresholding Algorithm (ISTA)

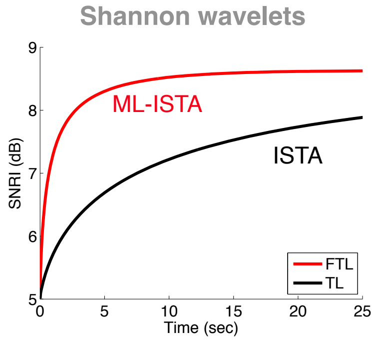
1. Initialization $(n \leftarrow 0), \mathbf{f}_0 = \mathbf{g}$ (Figueiredo, Nowak, *IEEE-IP* 2003)
2. Landweber update: $\mathbf{z} = \mathbf{f}_n + \gamma\mathbf{H}^T(\mathbf{g} - \mathbf{H}\mathbf{f}_n)$
3. Wavelet denoising: $\mathbf{w} = \mathbf{W}^T\mathbf{z}, \quad \tilde{\mathbf{w}} = T_{\gamma\lambda}\{\mathbf{w}\}$ (soft threshold)
4. Signal update: $\mathbf{f}_{n+1} \leftarrow \mathbf{W}\tilde{\mathbf{w}}$ and repeat from Step 2 until convergence

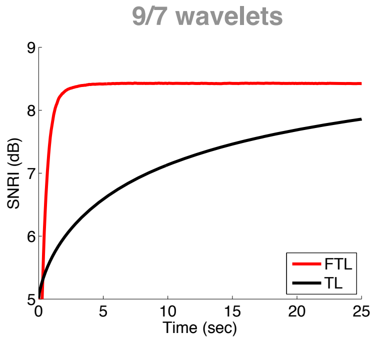


Proof of convergence: (Daubechies, Defrise, De Mol, 2004)

Fast multilevel wavelet-regularized deconvolution

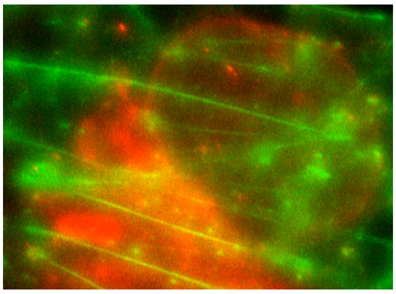
- Key features of multilevel wavelet deconvolution algorithm (ML-ISTA)
  - Acceleration by one order of magnitude with respect to state-of-the art algorithm (ISTA) (multigrid iteration strategy)
  - Applicable in 2D or 3D:
    first wavelet attempt for the deconvolution of 3D fluorescence micrographs
  - Works for any wavelet basis
  - Typically outperforms oracle Wiener solution (best linear algorithm)



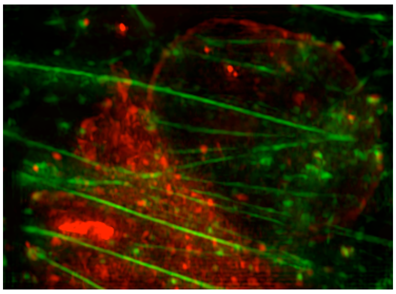


(Vonesch-Unser, *IEEE-IP*, 2009)

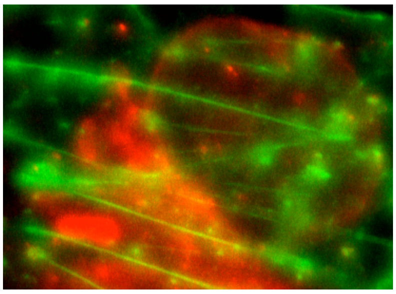
Deconvolution of 3D fluorescence micrographs



Widefield micrograph



ML-ISTA 5 iterations



ISTA 5 iterations

384×288×32 stack (maximum-intensity projections); sample: fibroblast cells;
staining: actine filaments in green (Phalloidin-Alexa488), vesicles and nucleus membrane in red (DiI);
objective: 63× plan-apochromat 1.4 NA oil-immersion;
diffraction-limited PSF model; initialization: measured data.

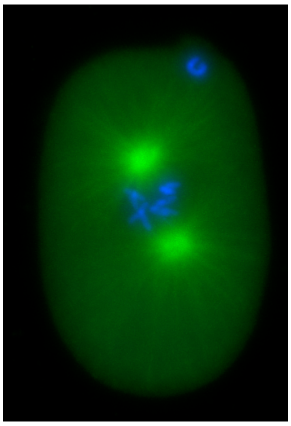
3D fluorescence microscopy experiment

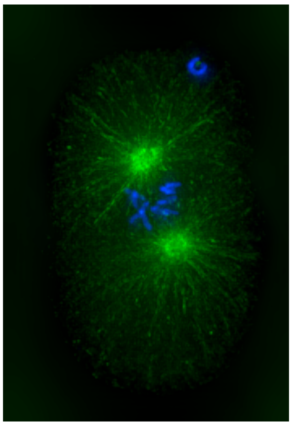
Input data (open pinhole)

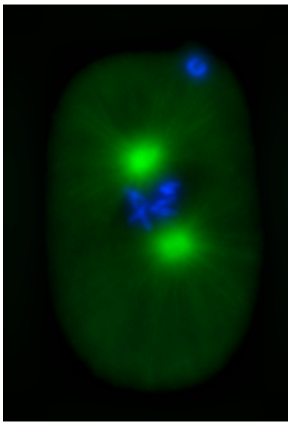
ML-ISTA 15 iterations

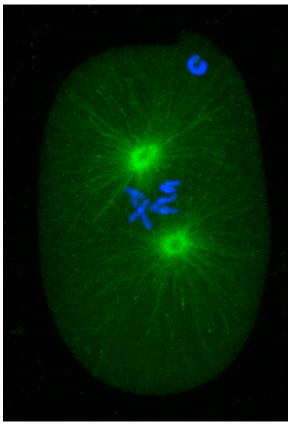
ISTA 15 iterations

Confocal reference









Maximum-intensity projections of 512×352×96 image stacks;
Zeiss LSM 510 confocal microscope with a 63× oil-immersion objective;
C. Elegans embryo labeled with Hoechst, Alexa488, Alexa568;
each channel processed separately; computed PSF based on diffraction-limited model;
separable orthonormalized linear-spline/Haar basis.

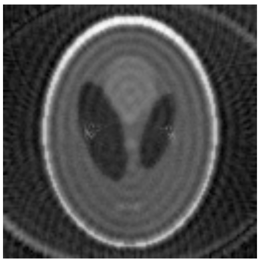
Preliminary results with parallel MRI

Simulated parallel MRI experiment

(M. Guerquin-Kern, BIG)

- Shepp-Logan brain phantom
- 4 coils, undersampled spiral acquisition, 15dB noise

Space



Backprojection

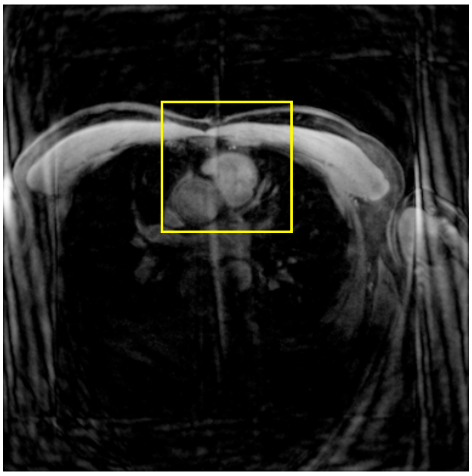
L_2 regularization (CG)

ℓ_1 wavelet regularization

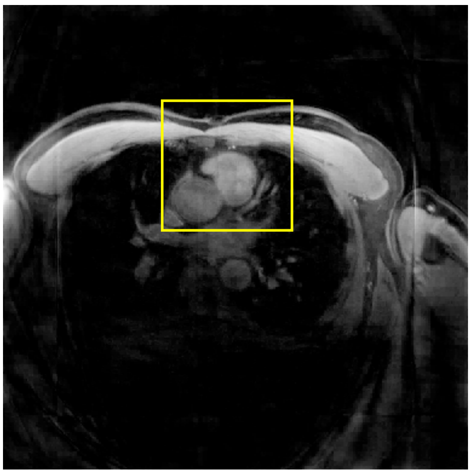
NCCBI collaboration with K. Prüssmann, ETHZ

Fresh try at ISMRM reconstruction challenge

L_2 regularization (Laplacian)



ℓ_1 wavelet regularization



Sampling-related problems and formulations

	Variational	MMSE	TV	Sparsity ℓ_1-norm
Ideal sampling	Optimal discretization and solution Smoothing spline	Optimal discretization and solution Hybrid Wiener filter	Optimal solution space Nonuniform spline	Exact solution (for ortho basis) Soft-threshold
Generalized sampling	Direct numerical solution Digital filtering	Gaussian MAP	Iterative TV deconvolution	Numerical optimization Multi-level, iterated, threshold
Linear inverse problems	Numerical, matrix-form solution CG (iterative)	Gaussian MAP	Iterative TV reconstruction	Numerical optimization Iterated thresholding

Level of complexity

CONCLUSION

- **Generalized sampling**
  - Unifying Hilbert-space formulation: Riesz basis, etc.
  - Approximation point of view: projection operators
  - Increased flexibility; closer to real-world systems
  - Generality: nonideal sampling, interpolation, etc...
- **Regularized sampling**
  - Regularization theory: smoothing splines
  - Stochastic formulation: hybrid form of Wiener filter
  - Non-linear techniques (e.g., TV)
- Quest for the "best" representation space
  - Optimal choice determined by regularization operator L
  - Spline-like representation; compactly-supported basis functions
  - Not bandlimited !

CONCLUSION (Cont'd)

- **Sampling with sparsity constraints**
  - Requires sparse signal representation (wavelets)
  - Theory of compressed sensing
  - Qualitatively equivalent to non-quadratic regularization (e.g. TV)
  - Challenge: Can we re-engineer the acquisition process in order to sample with fewer measurements?
- **Further research issues**
  - Fast algorithms for l_1-constrained signal reconstruction
  - CS: beyond toy problems
    real-word applications of the "compressed" part of theory
  - Strengthening the link with spline theory
  - Better sparsifying transforms of signal and images:
    tailored basis functions, rotation-invariance, ...

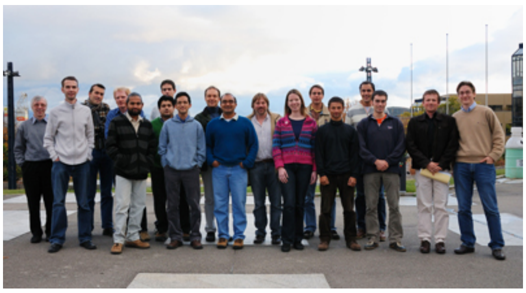
Acknowledgments

Many thanks to

- Prof. Thierry Blu
- Prof. Akram Aldroubi
- Dr. Philippe Thévenaz
- Dr. Sathish Ramani
- Dr. Cédric Vonesch
- Prof. D. Van De Ville
- Prof. Yonina Eldar

\+ many other researchers, and graduate students

EPFL's Biomedical Imaging Group



- Preprint and demos at: http://bigwww.epfl.ch/

Bibliography

- Sampling theory and splines
  - M. Unser, A. Aldroubi, “A General Sampling Theory for Nonideal Acquisition Devices,” *IEEE Trans Signal Processing*, vol. 42, no. 11, pp. 2915-2925, 1994.
  - M. Unser, “Sampling—50 Years After Shannon,” *Proc. IEEE*, vol. 88, no. 4, pp. 569-587, 2000.
  - M. Unser, “Splines: A Perfect Fit for Signal and Image Processing,” *IEEE Signal Processing Magazine*, vol. 16, no. 6, pp. 22-38, 1999.

- Regularized sampling
  - M. Unser, T. Blu, “Generalized Smoothing Splines and the Optimal Discretization of the Wiener Filter,” *IEEE Trans. Signal Processing*, vol. 53, no. 6, pp. 2146-2159, 2005.
  - Y.C. Eldar, M. Unser, “Nonideal Sampling and Interpolation from Noisy Observations in Shift-Invariant Spaces,” *IEEE Trans. Signal Processing*, vol. 54, no. 7, pp. 2636-2651, 2006.
  - T. Blu, M. Unser, “Self-Similarity: Part II—Optimal Estimation of Fractal Processes,” *IEEE Trans Signal Processing*, vol. 55, no. 4, pp. 1364-1378, 2007.
  - S. Ramani, D. Van De Ville, T. Blu, M. Unser, “Nonideal Sampling and Regularization Theory,” *IEEE Trans. Signal Processing*, vol. 56, no. 3, pp. 1055-1070, March 2008.
  - E. Mammen, S. van de Geer, “Locally adaptive regression splines,” *Annals of Statistics*, vol. 25, no. 1, pp. 387-413, 1997.

Bibliography (Cont’d)

- Compressive sampling
  - E. Candès, J. Romberg, T. Tao, “Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,”, *IEEE Trans. Information Theory*, vol. 52, no. 2, pp. 489 - 509, 2006.
  - D. Donoho, “Compressed sensing,” *IEEE Trans. Information Theory*, vol. 52, no. 4, pp. 1289 - 1306, 2006.
  - E. Candès and M. Wakin, “An introduction to compressive sampling,” *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21 - 30, 2008.
  - A.M. Bruckstein, D.L. Donoho, M. Elad,, “From Sparse Solutions of Systems of Equations to Sparse Modeling of Signals and Images,” *SIAM Review*, vol. 51, no. 1, pp. 34.81, 2009.

- ℓ_1-regularized signal reconstruction
  - L. Rudin, S.J. Osher, E. Fatemi, “Nonlinear total variation based noise removal algorithms,” *Physica D.*, vol. 60, pp. 259-268, 1992.
  - A. Chambolle, R.A. DeVore, N.-Y. Lee and B.J. Lucier, “Nonlinear Wavelet Image Processing: Variational Problems, Compression, and Noise Removal Through Wavelet Shrinkage,” *IEEE Trans. Image Processing*, vol. 7, no. 33, pp. 319-335, April 1998.
  - M. Figueiredo and R. Nowak, “An EM algorithm for wavelet-based image restoration,” *IEEE Trans. Image Processing*, vol. 12, no. 8, pp. 906–916, Aug. 2003.
  - I. Daubechies, M. Defrise, and C. De Mol, “An Iterative Thresholding Algorithm for Linear Inverse Problems with a Sparsity Constraint,” *Comm. Pure and Applied Mathematics*, vol. 57, no. 11, pp. 1413–1457, August 2004.
  - C. Vonesch, M. Unser, “A Fast Multilevel Algorithm for Wavelet-Regularized Image Restoration,” *IEEE Trans. Image Processing*, , vol. 18, no. 3, pp. 509-523, 2009.