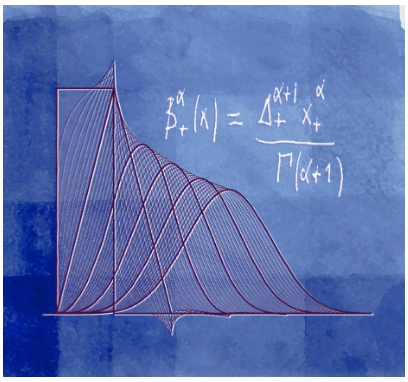
ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Stochastic models for sparse and piecewise-smooth signals

Michael Unser
Biomedical Imaging Group
EPFL, Lausanne, Switzerland



Sparse representations and efficient sensing of data, Dagstuhl, Feb, 2011

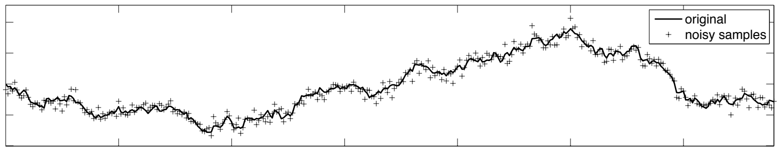
Motivation: Beyond Wiener filtering

Simple denoising experiment

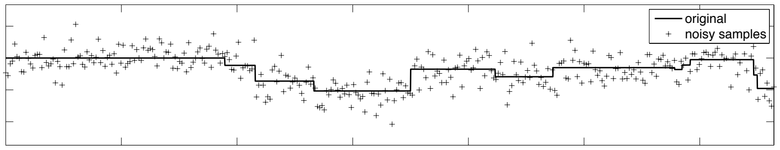
- Measurement model

  $g[k] = s(k) + n[k]$

  - $s(x)$: Continuously-defined process
  - $n[k]$: Discrete white Gaussian noise



Wiener process (Gaussian)



Compound Poisson process (Sparse)

- Controlled experiment
  - Matched 2nd-order statistics (correlation function)
  - Generalized spectrum $\sim \frac{1}{\omega}$

Three dominant paradigms

- Wiener solution (best linear estimator) = Smoothing spline

$$\begin{aligned}\tilde{s}_{\mathrm{spline}}(x) &= \arg\min_{s(x)} \left\{ \sum_{k\in\mathbb{Z}} |g[k]-s(k)|^2 + \mu \int_{\mathbb{R}} |\mathrm{D}s(x)|^2 \mathrm{d}x \right\} \\ &= \arg\min_{s(x)} \left\{ \|g-s\|_{\ell_2}^2 + \mu \|\mathrm{D}s\|_{L_2(\mathbb{R})}^2 \right\}\end{aligned}$$

Theoretical result: MMSE solution = piecewise-linear smoothing spline (Blu-U., 2005)

- Wavelet solution = sparse signal recovery

$$\tilde{s}_{\mathrm{wave}}(x) = \arg\min_{s(x)} \left\{ \|g-s\|_{\ell_2}^2 + \sum_i \mu_i \|w_i\|_{\ell_p}^p \right\}$$

with $\quad w_i[k] = \left\langle s, 2^{i/2}\psi(x/2^i - k)\right\rangle_{L_2(\mathbb{R})}$

Theoretical result: Simple wavelet shrinkage algorithm (Chambolle et al., 2005)

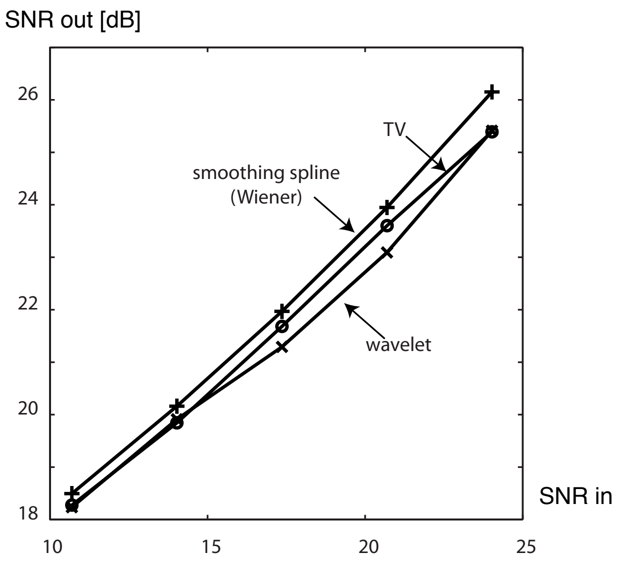
Three dominant paradigms (cont'd)

- Total variation = non-quadratic regularization

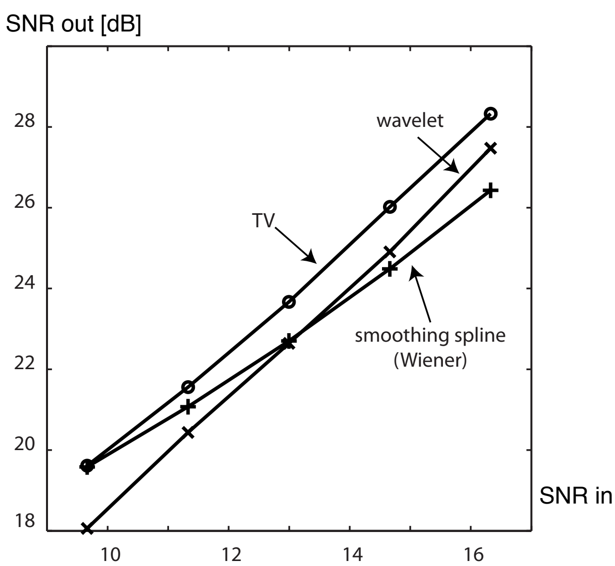
$$\tilde{s}_{\mathrm{TV}}(x) = \arg\min_{s(x)} \left\{ \|g - s\|_{\ell_2}^2 + \mu \, \mathrm{TV}(s) \right\}$$

Theoretical result [Mammen, *Annals of Statistics*, 1997]
Piecewise-constant spline with adaptive knots is a global minimizer

Denoising results



Wiener process (Gaussian)

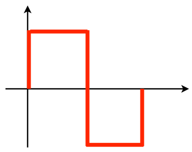


Poisson process (Sparse)

Commonalities

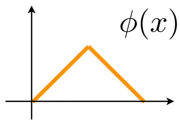
- Central role of derivative operator
  - Quadratic spline energy: $\|\mathrm{D}s\|^2_{L_2(\mathbb{R})}$
  - TV as an L_1 norm: $s \in W_1^1 \quad \Leftrightarrow \quad \mathrm{TV}(s) = \|\mathrm{D}s\|_{L_1(\mathbb{R})}$
  - Wavelet as a smoothed derivative: $\psi_{\mathrm{Haar}}(x) = \mathrm{D}\phi(x)$

$\psi(x) = \mathrm{L}^*\phi(x)$



$\mathrm{L}^* = \frac{\mathrm{d}}{\mathrm{d}x}$

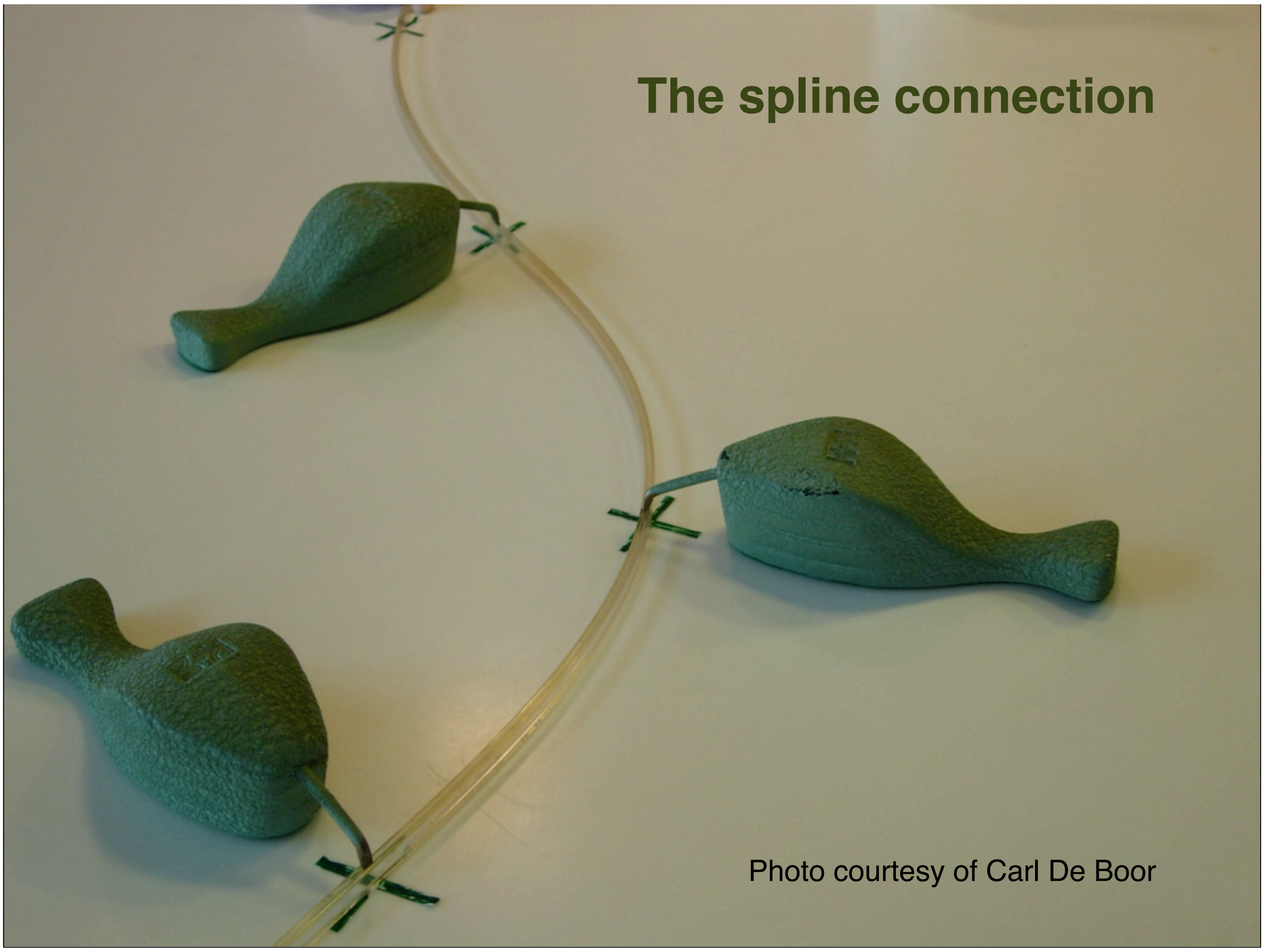
$\phi(x)$



$\Rightarrow \langle f, \psi(\cdot - x_0)\rangle = \mathrm{L}(f * \phi^*)(x_0) = -\frac{\mathrm{d}}{\mathrm{d}x}\,(f*$  $)\,(x_0)$

OUTLINE

- Motivation: beyond Wiener filtering ✔
- The spline connection
  - *L*-splines and signals with finite rate of innovation
  - Green functions as elementary building blocks
- Sparse stochastic processes
  - Innovation model
  - Gelfand's theory of generalized stochastic processes
  - Distributional solution of operator equations
- Imposing (TSR) invariance
  - Scale-invariant operators and their inverses
  - Fractal random processes (Gaussian vs. sparse)
  - Wavelet analysis of sparse processes
- B-spline conquest of fractal/sparse processes
  - Fractional B-splines
  - "Stationarization" and supression of long-range dependencies
  - On the statistical optimality of TV



Splines: signals with finite rate of innovation

$\mathrm{L}\{\cdot\}$: differential operator

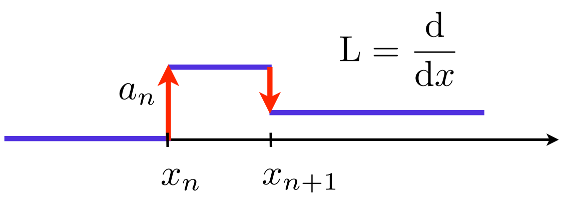
$\delta(x)$: Dirac distribution

Definition

The function $s(x)$ is a (non-uniform) L-spline with knots $\{x_n\}$ iff.

$$\mathrm{L}\{s\}(x) = \sum_{n=1}^{N} a_n \delta(x - x_n)$$

Spline theory: (Schultz-Varga, 1967)

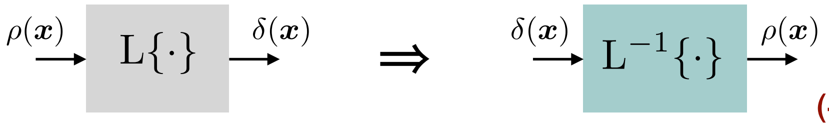


- FIR signal processing: Innovation variables $(2N)$ (Vetterli et al., 2002)
  - Location of singularities (knots) : $\{x_n\}_{n=1,\ldots,N}$
  - Strength of singularities (linear weights): $\{a_n\}_{n=1,\ldots,N}$

Splines and Green's functions

Definition

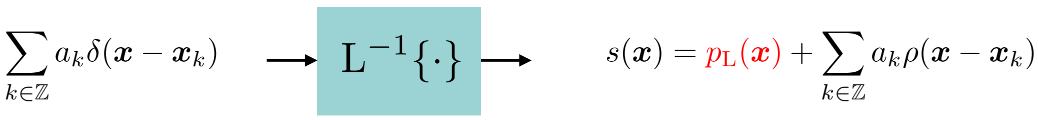
$\rho(\boldsymbol{x})$ is a Green function of the shift-invariant operator L iff $\mathrm{L}\{\rho\} = \delta$



(+ null-space component?)

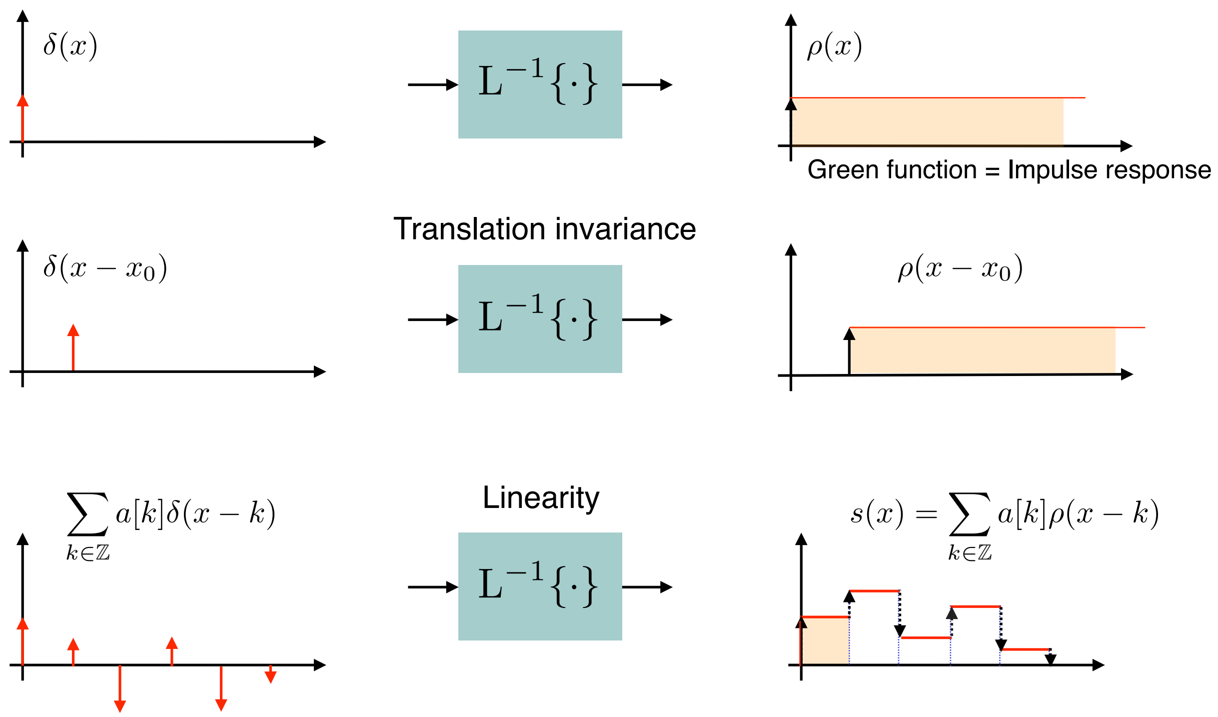
- General (non-uniform) L-spline: $\mathrm{L}\{s\}(\boldsymbol{x}) = \sum_{k\in\mathbb{Z}} a_k \delta(\boldsymbol{x} - \boldsymbol{x}_k)$

Formal integration



Example of spline synthesis

$\mathrm{L} = \frac{\mathrm{d}}{\mathrm{d}x} = \mathrm{D} \quad \Rightarrow \quad \mathrm{L}^{-1}$: integrator



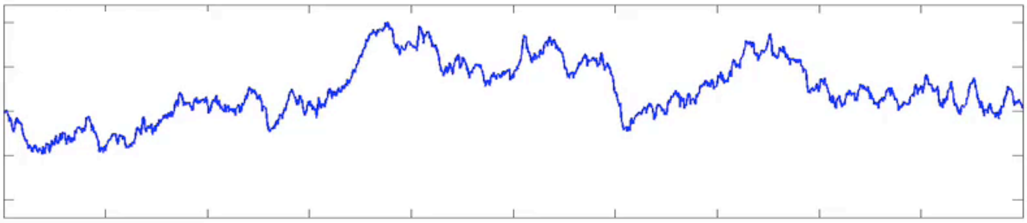
Sparse stochastic processes

Brownian motion vs. spline synthesis

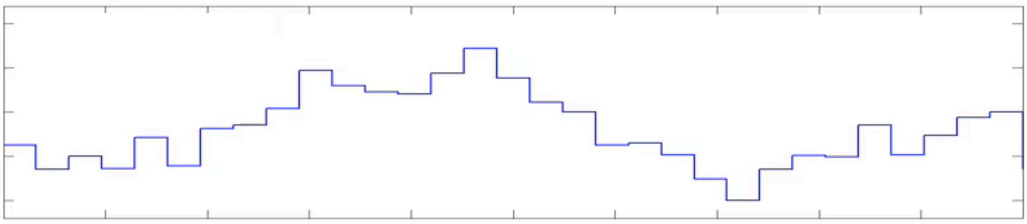
$\mathrm{L} = \frac{\mathrm{d}}{\mathrm{d}x} \quad \Rightarrow \quad \mathrm{L}^{-1}$: integrator

white noise
or
stream of Diracs

$\longrightarrow \boxed{\mathrm{L}^{-1}\{\cdot\}} \longrightarrow$



Brownian motion



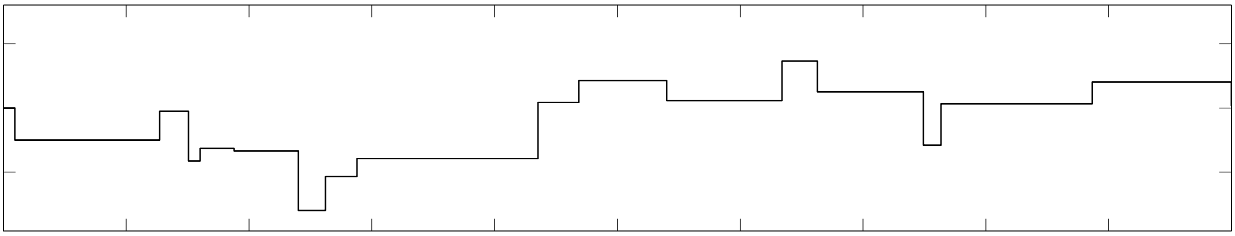
Cardinal spline (Schoenberg, 1946)

Compound Poisson process (sparse)

$\mathrm{L} = \frac{\mathrm{d}}{\mathrm{d}x} \quad \Rightarrow \quad \mathrm{L}^{-1}$: integrator

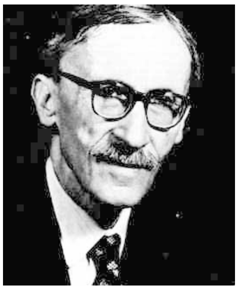
$$r(x) = \sum_k a_k \delta(x - x_k) \longrightarrow \boxed{\mathrm{L}^{-1}\{\cdot\}} \longrightarrow s(x)$$

random stream of Diracs | **Compound Poisson process**



Random jumps with rate λ (Poisson point process)

Jump size distribution: $a \backsim p(a)$



(Paul Lévy, 1934)

Generalized stochastic processes

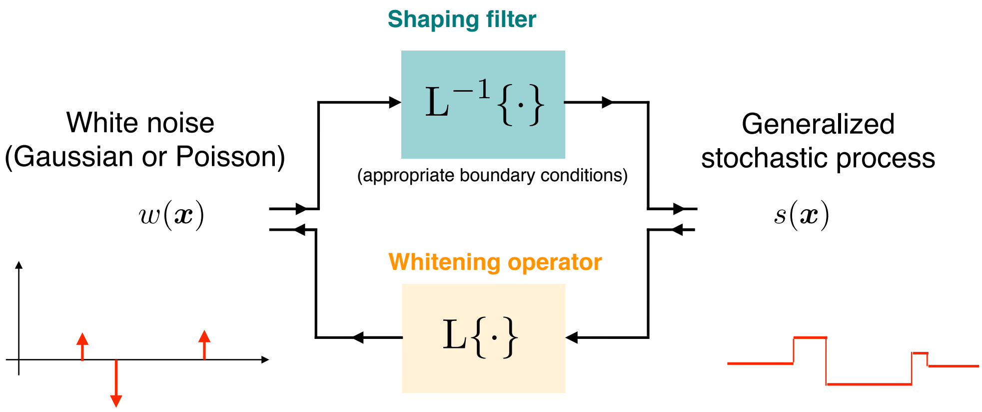
Splines are in direct correspondence with stochastic processes (stationary or fractals) that are solution of the same partial differential equation, but with a random driving term.

Defining operator equation: $\mathrm{L}\{s(\cdot)\}(\boldsymbol{x}) = r(\boldsymbol{x})$

- Specific driving terms
  - $r(\boldsymbol{x}) = \delta(\boldsymbol{x}) \quad \Rightarrow \quad s(\boldsymbol{x}) = \mathrm{L}^{-1}\{\delta\}(\boldsymbol{x})$: Green function
  - $r(\boldsymbol{x}) = \sum_{\boldsymbol{k} \in \mathbb{Z}^d} a[\boldsymbol{k}]\delta(\boldsymbol{x} - \boldsymbol{k}) \quad \Rightarrow \quad s(\boldsymbol{x})$: Cardinal L-spline
  - $r(\boldsymbol{x})$: white noise $\quad \Rightarrow \quad s(\boldsymbol{x})$: generalized stochastic process

⚠ non-empty null space of L, boundary conditions

References: stationary proc. (U.-Blu, *IEEE-SP* 2006), fractals (Blu-U., *IEEE-SP* 2007), sparse processes (U.-Tafti, *IEEE-SP* in press)



- What is white noise ?
  - **The problem**: Continuous-domain white noise does not have a pointwise interpretation.
  - **Standard stochastic calculus**. Statisticians circumvent the difficulty by working with *random measures* ($\mathrm{d}W(x) = w(x)\mathrm{d}x$) and *stochastic integrals*; i.e, $s(x) = \int_{\mathbb{R}} \rho(x, x')\mathrm{d}W(x')$ where $\rho(x, x')$ is a suitable kernel.
  - **Innovation model**. The white noise interpretation is more appealing for engineers (cf. Papoulis), but is requires a proper distributional formulation and operator calculus.

Gelfand's distributional framework

- Formal specification of Brownian motion (Wiener process)

$$\frac{\mathrm{d}}{\mathrm{d}t}s(t) = w(t) \qquad \Rightarrow \qquad s(t) = W(t) = \int_0^t w(\tau)\mathrm{d}\tau \quad ???$$

- Distributional interpretation

$$\mathrm{L}s = w \quad \Leftrightarrow \quad \forall \varphi \in \mathcal{S}, \quad \langle \varphi, \mathrm{L}s \rangle = \langle \varphi, w \rangle$$

$$s = \mathrm{L}^{-1}w \quad \Leftrightarrow \quad \forall \varphi \in \mathcal{S}, \quad \langle \varphi, s \rangle = \langle \varphi, \mathrm{L}^{-1}w \rangle$$

- Concept
  - $r(\varphi) = \langle \varphi, s \rangle$ is a conventional scalar random variable whose probablity density function (PDF) $p_{r(\varphi)}(r)$ can be specified
  - Getting back point values: $s(t_0) = \langle \delta(\cdot - t_0), s \rangle$

Characteristic form

- s is a generalized random process over $\mathcal{S}'$ (Schwartz's space of tempered distributions)
- $r_0 = \langle \varphi_0, s \rangle$ is a conventional random variable with characteristic function

  $\hat{p}_{r_0}(\omega) = \int_{\mathbb{R}} p_{r_0}(r) e^{-j\omega r} \mathrm{d}r = \mathcal{F}\{p_{r_0}\}(\omega) = E\{e^{-j\langle s, \omega\varphi_0 \rangle}\}.$
- Idea: Make φ_0 generic and use it as a space/frequency domain index variable...

- Characteristic form (Kolmogorov 1935; Gelfand, 1955)

$$Z_s(\varphi) = E\{e^{-j\langle s, \varphi \rangle}\} = \int_{\mathbb{R}} e^{-jr(\varphi)} p_{r(\varphi)}(r) \mathrm{d}r \quad \text{with} \quad r(\varphi) = \langle \varphi, s \rangle$$

  - Infinite-dimensional generalization of the characteristic function (since φ is generic)
  - Uniquely defines the process while condensing all the statistical information
  - Example:

    $$Z_s\big(\omega_1 \delta(\cdot - x_1) + \omega_2 \delta(\cdot - x_2)\big) = E\{e^{j(\omega_1 s_1 + \omega_2 s_2)}\} = \mathcal{F}\{p(s_1, s_2)\}(\omega_1, \omega_2)$$

    with $s_1 = s(x_1)$ and $s_2 = s(x_2)$

Complete functional characterization

- Properties of the characteristic form
  - The characteristic form $Z_s(\varphi)$ of a generalized stochastic process over $\mathcal{S}'$ is a continuous, positive-define functional on $\mathcal{S}$ such that $Z_s(0) = 1$.

- Bochner-Minlos theorem (Minlos, 1963)
  - Let $Z(\varphi)$ be a continuous, positive-definite functional on $\mathcal{S}$ such that $Z(0) = 1$. Then Z uniquely defines a generalized stochastic process whose characteristic form is $Z(\varphi)$.

    Moreover, one has the guarantee that all the finite dimensional probabilities densities that can be derived from $Z(\varphi)$ by setting $\varphi = \omega_1\varphi_1 \cdots + \omega_N \varphi_N$ are mutually compatible.

- Abstract formulation
  - Theory of measures on topological vector spaces
  - Generalized stochastic process defined by a probability measure $\mu(s)$ on $\mathcal{S}'$

    $$Z_s(\varphi) = \int_{\mathcal{S}'} e^{-j\langle s, \varphi \rangle} \mathrm{d}\mu(s)$$

Characteristic forms of white noises

- White noise process: w
- Scalar observation (random variable): $r = r(\varphi) = \langle w, \varphi \rangle$
- Characteristic function of r: $\hat{p}_r(\omega) = E\{e^{-j\omega\langle w,\varphi\rangle}\} = Z_w(\omega\varphi)$

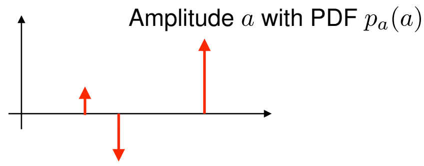
- White Gaussian noise

$$Z_w(\varphi) = \exp\left(-\frac{1}{2}\|\varphi\|_{L_2}^2\right) \quad \Rightarrow \quad \hat{p}_r(\omega) = e^{-\frac{1}{2}\omega^2\sigma^2} \quad \text{with} \quad \sigma^2 = \|\varphi\|_{L_2}^2$$

- Poisson noise

Amplitude distribution $p_a(a)$, Poisson rate λ

$$Z_w(\varphi) = \exp\left(\lambda \int_{\mathbb{R}^d}\int_{\mathbb{R}} (e^{ja\varphi(\boldsymbol{x})} - 1) p_a(a) \mathrm{d}a \, \mathrm{d}\boldsymbol{x}\right)$$



- Correlation form

$$E\{\langle w, \varphi_1\rangle \cdot \langle w, \varphi_2\rangle\} = \sigma_0^2 \, \langle \varphi_1, \varphi_2 \rangle \quad \Rightarrow \quad R_w(\boldsymbol{x}) = \sigma_0^2 \, \delta(\boldsymbol{x})$$

Distributional solution of operator equation

$$s = \mathrm{L}^{-1} w \quad \Leftrightarrow \quad \forall \varphi \in \mathcal{S}, \quad \langle \varphi, s\rangle = \langle \varphi, \mathrm{L}^{-1} w\rangle = \langle \mathrm{L}^{-1*}\varphi, w\rangle$$

- Technical aspect: functional analysis

Find an acceptable inverse of L such that the adjoint operator L^{-1*} is well-defined over Schwartz's class of test functions

Ideally: $\mathrm{L}^{-1*}\varphi \in \mathcal{S}$

or at least $\|\mathrm{L}^{-1*}\varphi\|_{L_p} < +\infty$ (continuity)

Example: L_p-stable version of anti-derivative

$$\mathrm{D}^{-1*}\varphi(x) \quad \stackrel{\mathcal{F}}{\longleftrightarrow} \quad \frac{\hat{\varphi}(\omega) - \hat{\varphi}(0)}{-j\omega}$$

Characteristic forms of generalized processes

$$s = \mathrm{L}^{-1} w \quad \Leftrightarrow \quad \forall \varphi \in \mathcal{S}, \quad \langle \varphi, s \rangle = \langle \varphi, \mathrm{L}^{-1} w \rangle = \langle \mathrm{L}^{-1*} \varphi, w \rangle$$

- Generalized Gaussian process

Classical spectral power density

Whitening operator L

$$Z_s(\varphi) = Z_w(\mathrm{L}^{-1*}\varphi) = \exp\left(-\frac{1}{2}\|\mathrm{L}^{-1*}\varphi\|_{L_2}^2\right) = \exp\left(-\frac{1}{2}\int_{\mathbb{R}^d} |\hat{\varphi}(\boldsymbol{\omega})|^2 \frac{1}{|\hat{L}(\boldsymbol{\omega})|^2} \frac{\mathrm{d}\boldsymbol{\omega}}{(2\pi)^d}\right)$$

- Generalized Poisson process (sparse)

Amplitude PDF $p_a(a)$, Poisson rate λ, whitening operator L

$$Z_s(\varphi) = Z_w(\mathrm{L}^{-1*}\varphi) = \exp\left(\lambda \int_{\mathbb{R}^d} \int_{\mathbb{R}} (e^{ja\mathrm{L}^{-1*}\varphi(\boldsymbol{x})} - 1)\, p_a(a) \mathrm{d}a \, \mathrm{d}\boldsymbol{x}\right)$$

(U. & Tafti, *IEEE-SP*, in press)

- Correlation form

$$E\{\langle s, \varphi_1 \rangle \cdot \langle s, \varphi_2 \rangle\} = \sigma_0^2 \, \langle \mathrm{L}^{-1}\mathrm{L}^{-1*}\varphi_1, \varphi_2 \rangle \quad \Rightarrow \quad R_s(\boldsymbol{x}) = \rho_{\mathrm{L}}(\boldsymbol{x}) * \rho_{\mathrm{L}}(-\boldsymbol{x})$$

Concrete example: (f)Brownian motion

$$\mathrm{D}s = w \qquad\qquad \mathrm{D}^{\gamma} s = w$$

$$s = \mathrm{D}_0^{-1} w \quad \Leftrightarrow \quad \forall \varphi \in \mathcal{S}, \quad \langle \varphi, s \rangle = \langle \mathrm{D}_0^{-1*}\varphi, w \rangle$$

L_2-stable anti-derivative: $\mathrm{D}_0^{-1*}\varphi(x) = \int_{\mathbb{R}} \frac{\hat{\varphi}(\omega) - \hat{\varphi}(0)}{-j\omega} e^{j\omega x} \frac{\mathrm{d}\omega}{2\pi}$

- Characteristic form of Wiener process

Stabilization ⇔ non-stationary behavior

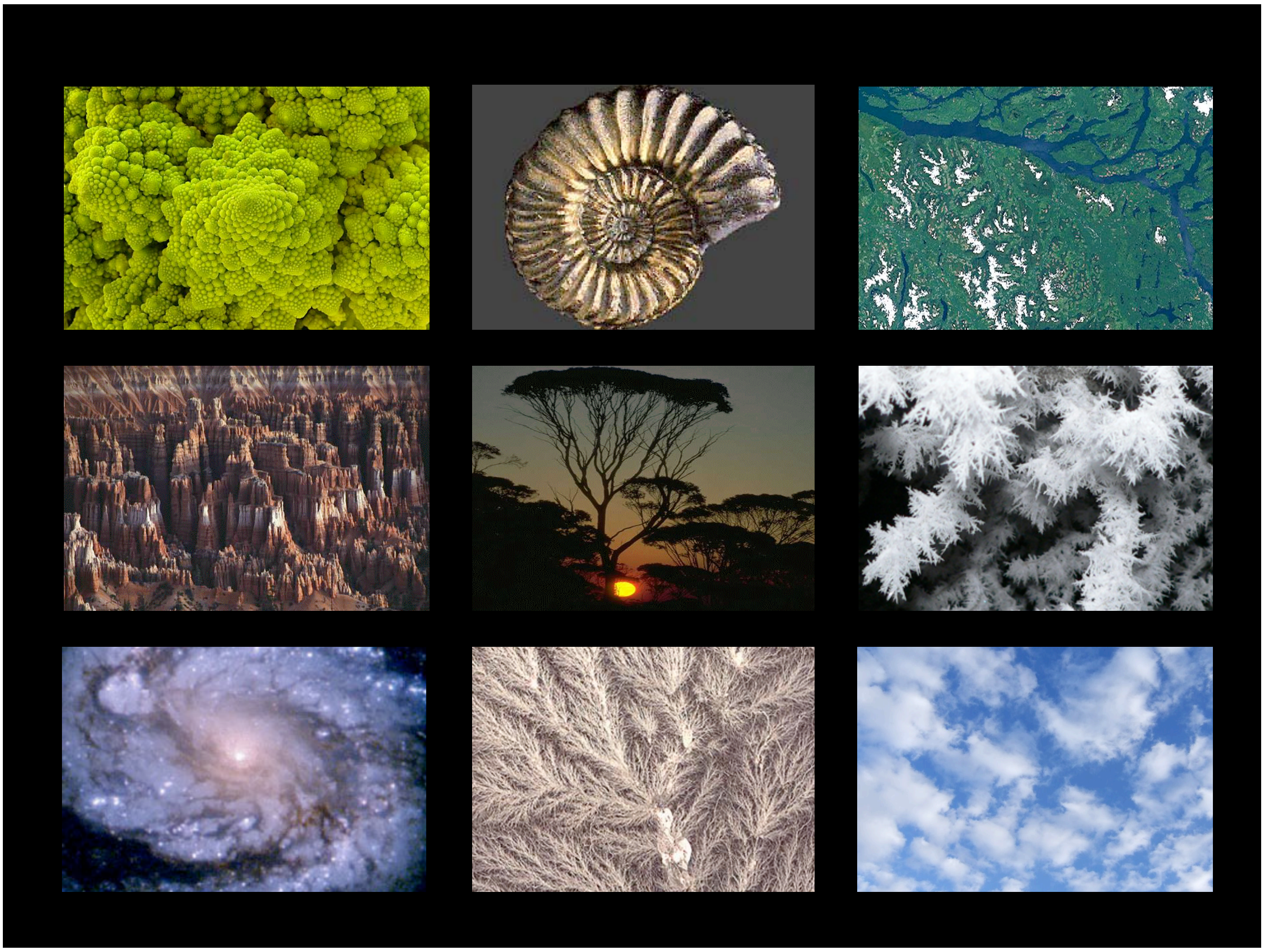
$$Z_s(\varphi) = \exp\left(-\frac{1}{2}\|\mathrm{D}_0^{-1*}\varphi\|_{L_2}^2\right) = \exp\left(-\frac{1}{2}\int_{\mathbb{R}} \left|\frac{\hat{\varphi}(\omega) - \hat{\varphi}(0)}{j\omega}\right|^2 \frac{\mathrm{d}\omega}{2\pi}\right)$$

(by Parseval)

- Characteristic form of fractional Brownian motion

$$Z_s(\varphi) = \exp\left(-\frac{1}{2}\int_{\mathbb{R}} \left|\frac{\hat{\varphi}(\omega) - \hat{\varphi}(0)}{|\omega|^{\gamma}}\right|^2 \frac{\mathrm{d}\omega}{2\pi}\right)$$

(Blu-U., *IEEE-SP* 2007)



IMPOSING SCALE INVARIANCE

- Scale-invariant operators
- Inverse operators (fractional integrators)
- Fractal random processes
  (Gaussian vs. Sparse)
- Wavelet analysis of fractal/sparse processes

Scale-invariant operators

Definition: An operator L is scale-invariant iff it commutes with dilation: i.e., $\forall s(\boldsymbol{x}), \mathrm{L}\{s(\cdot)\}(\boldsymbol{x}/a) = C_a\, \mathrm{L}\{s(\cdot/a)\}(\boldsymbol{x})$.

Theorem

The complete family of real scale-invariant 1D convolution operators is given by the fractional derivatives ∂_τ^γ, whose frequency response is

$$\hat{L}(\omega) = (-j\omega)^{\frac{\gamma}{2}-\tau}(j\omega)^{\frac{\gamma}{2}+\tau}$$

$\gamma \in \mathbb{R}^+$: order of the derivative (i.e., $|\hat{L}(\omega)| = |\omega|^\gamma$)

$\tau \in \mathbb{R}$: phase (or asymmetry)

(Unser & Blu, *IEEE-SP*, 2007)

Scale- and rotation-invariant operators

Definition: An operator L is scale- and rotation-invariant iff.

$$\forall s(\boldsymbol{x}),\ \mathrm{L}\{s(\cdot)\}(\mathbf{R}_\theta \boldsymbol{x}/a) = C_a \cdot \mathrm{L}\{s(\mathbf{R}_\theta \cdot /a)\}(\boldsymbol{x})$$

where $\mathbf{R}_\theta$ is an arbitrary $d \times d$ unitary matrix and C_a a constant

- **Invariance theorem**

The complete family of real, scale- and rotation-invariant convolution operators is given by the fractional Laplacians

$$(-\Delta)^{\frac{\gamma}{2}} \quad \stackrel{\mathcal{F}}{\longleftrightarrow} \quad \|\boldsymbol{\omega}\|^\gamma$$

- Invariant Green functions (a.k.a. RBF) (Duchon, 1979)

$$\rho(\boldsymbol{x}) = \begin{cases} \|\boldsymbol{x}\|^{\gamma-d}\log\|\boldsymbol{x}\|, & \text{if } \gamma-d \text{ is even} \\ \|\boldsymbol{x}\|^{\gamma-d}, & \text{otherwise} \end{cases}$$

Inverse operators: fractional calculus

L	$\rho = \mathrm{L}^{-1}\delta$	$\mathrm{L}^{-1}\varphi$	$\mathrm{L}^{-1*}\varphi$	$0 < \gamma < 1 + d/2$
D^γ	$\frac{x_+^{\gamma-1}}{\Gamma(\gamma)}$	$\int_{\mathbb{R}} \frac{e^{j\omega x}-1}{(j\omega)^\gamma}\hat{\varphi}(\omega)\frac{\mathrm{d}\omega}{2\pi}$	$\int_{\mathbb{R}} e^{j\omega x}\frac{\hat{\varphi}(\omega)-\hat{\varphi}(0)}{(-j\omega)^\gamma}\frac{\mathrm{d}\omega}{2\pi}$	
∂_τ^γ	$\frac{\lvert x\rvert^{\gamma-1}}{\Gamma(\gamma)}\left(A_{\gamma,\tau} + B_{\gamma,\tau}\mathrm{sign}(x)\right)$, $\gamma \notin \mathbb{N}$	$\int_{\mathbb{R}} \frac{e^{j\omega x}-1}{(-j\omega)^{\frac{\gamma}{2}-\tau}(j\omega)^{\frac{\gamma}{2}+\tau}}\hat{\varphi}(\omega)\frac{\mathrm{d}\omega}{2\pi}$	$\int_{\mathbb{R}} e^{j\omega x}\frac{\hat{\varphi}(\omega)-\hat{\varphi}(0)}{(j\omega)^{\frac{\gamma}{2}-\tau}(-j\omega)^{\frac{\gamma}{2}+\tau}}\frac{\mathrm{d}\omega}{2\pi}$	
$(-\Delta)^{\frac{\gamma}{2}}$	$C_\gamma\lVert\boldsymbol{x}\rVert^{\gamma-d}, \gamma - d \notin 2\mathbb{N}$	$\int_{\mathbb{R}^d} \frac{e^{j\langle\boldsymbol{x},\boldsymbol{\omega}\rangle}-1}{\lVert\boldsymbol{\omega}\rVert^\gamma}\hat{\varphi}(\boldsymbol{\omega})\frac{\mathrm{d}\boldsymbol{\omega}}{(2\pi)^d}$	$\int_{\mathbb{R}^d} e^{j\langle\boldsymbol{x},\boldsymbol{\omega}\rangle}\frac{\hat{\varphi}(\boldsymbol{\omega})-\hat{\varphi}(\boldsymbol{0})}{\lVert\boldsymbol{\omega}\rVert^\gamma}\frac{\mathrm{d}\boldsymbol{\omega}}{(2\pi)^d}$	

(U.-Tafti, 2011)

Theorem (Generalized Riesz potentials)

Unique left-inverse of $\mathrm{L}^* = (-\Delta)^{\frac{\gamma}{2}}$ that is L_p-stable and scale-invariant:

$$\mathrm{I}_p^\gamma\varphi(\boldsymbol{x}) = \int_{\mathbb{R}^d} e^{j\langle\boldsymbol{x},\boldsymbol{\omega}\rangle}\frac{\hat{\varphi}(\boldsymbol{\omega}) - \sum_{|\boldsymbol{k}|=0}^{\lfloor\gamma-d+\frac{d}{p}\rfloor}\hat{\varphi}^{(\boldsymbol{k})}(\boldsymbol{0})\frac{\boldsymbol{\omega}^{\boldsymbol{k}}}{\boldsymbol{k}!}}{\lVert\boldsymbol{\omega}\rVert^\gamma}\frac{\mathrm{d}\boldsymbol{\omega}}{(2\pi)^d} = \mathrm{L}^{-1*}\varphi(\boldsymbol{x})$$

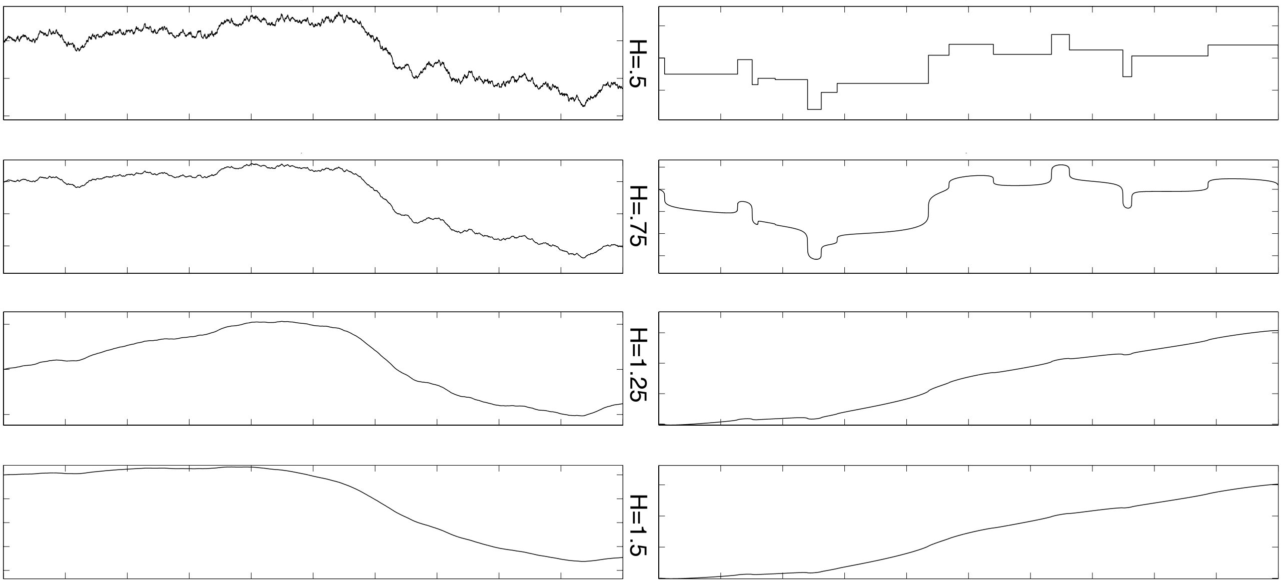
$$\forall\varphi \in \mathcal{S}(\mathbb{R}^d), \qquad \lVert\mathrm{I}_p^\gamma\varphi\rVert_{L_p(\mathbb{R}^d)} < C\cdot\lVert\varphi\rVert_{L_p(\mathbb{R}^d)}$$

for $\gamma \in \mathbb{R}^+\backslash\mathbb{Z}^+$ and $1 \leq p \leq +\infty$.

(Sun-U., 2011)

Self-similar processes (TS-invariant)

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



Gaussian

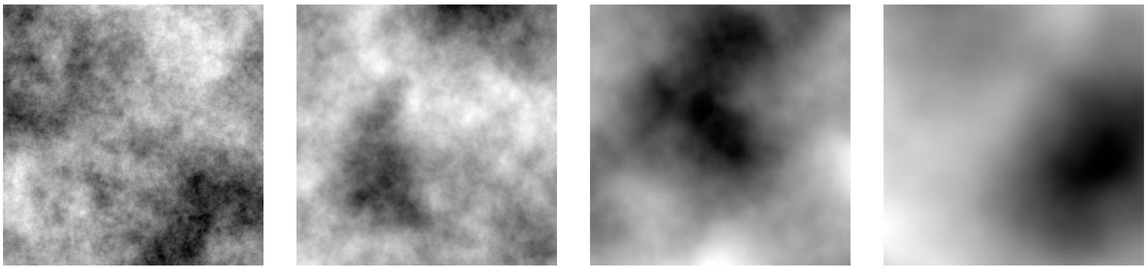
Sparse (generalized Poisson)

Fractional Brownian motion (Mandelbrot, 1968)

Scale- and rotation-invariant processes

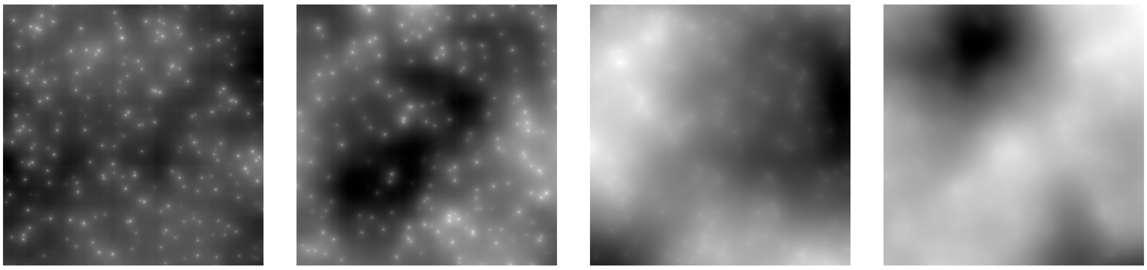
Stochastic partial differential equation : $(-\Delta)^{\frac{H+1}{2}} s(\boldsymbol{x}) = w(\boldsymbol{x})$

Gaussian

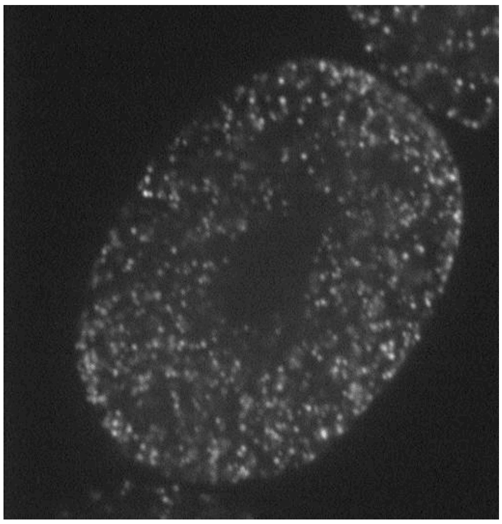


H=.5 H=.75 H=1.25 H=1.75

Sparse (generalized Poisson)

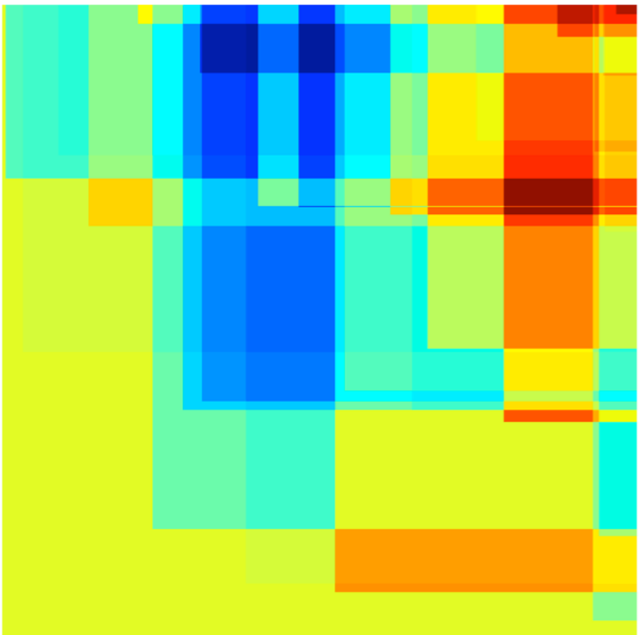


Powers of ten: from astronomy to biology

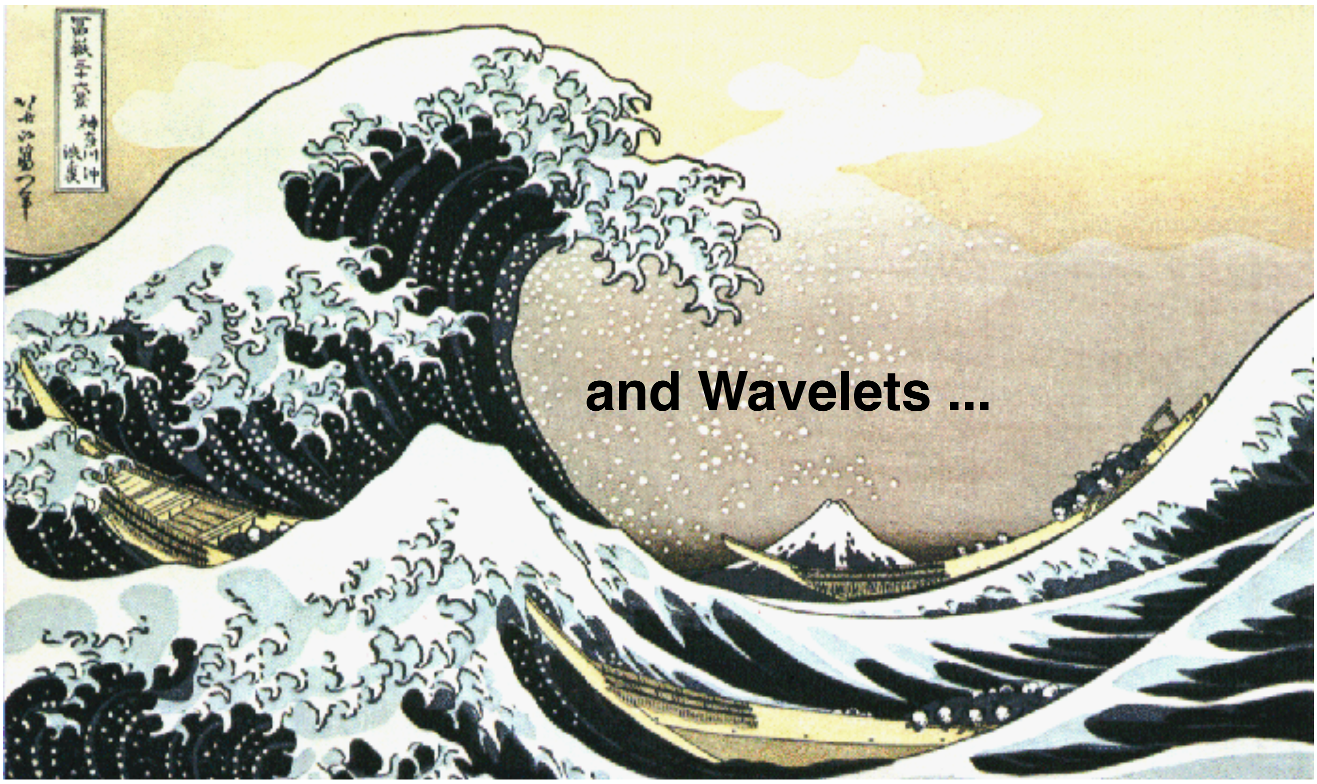


2D generalization: the Mondrian process

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

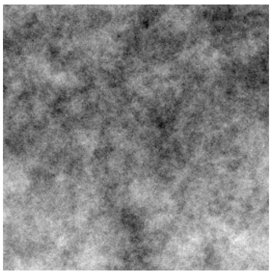


$\lambda = 30$



Wavelet analysis of self-similar processes

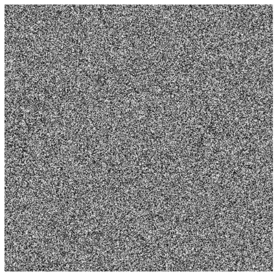
Generalized stochastic process



$(-\Delta)^{\frac{\gamma}{2}}$

(Gaussian or impulsive)

White noise



Multi-scale wavelet analysis

Approximate white noises

Laplacian-like wavelet: $\psi(\boldsymbol{x}) = (-\Delta)\varphi(\boldsymbol{x}) = (-\Delta)^{\frac{\gamma}{2}} \underbrace{(-\Delta)^{\frac{2-\gamma}{2}}\varphi(\boldsymbol{x})}_{\psi'(\boldsymbol{x})}$

1. Will approximately decorrelate fractal-like processes (fBm)
   (Quality of whitening depends on spectral characteristics of ψ')

2. Will yield sparse wavelet decomposition of generalized Poisson processes
   (Extent of sparsity depends on decay property of ψ')

Which signals are wavelet-sparse?

Conventional wavelets are optimal for piecewise-smooth signals (i.e., nonuniform splines) or 1/ω-like processes

- Generalized argument

$\tilde{\psi}_{i,k}(x) = 2^{-i/2}\mathrm{L}^*\phi_i(x - 2^i k)$

FRI or Poisson process

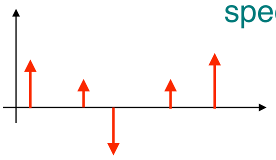
stochastic spline

sparse wavelet transform

$\sum_n a_n\delta(x - x_n)$ → $\mathrm{L}^{-1}\{\cdot\}$ → $f(x)$ → Operator-like wavelet transform → $w_i[k] = \langle f, \tilde{\psi}_{i,k}\rangle$

spectral shaping

multiscale "whitening"



Wavelet footprints: $\propto \sum_n a_n\phi_i(x_n - 2^i k)$

localized, lowpass kernel

Construction of piecewise-constant B-spline

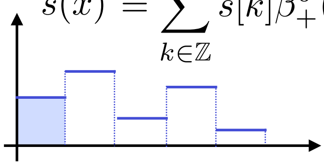
- Spline-defining operators

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

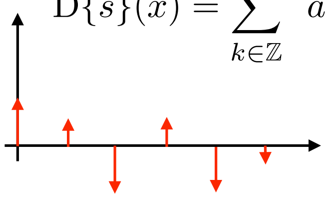
Discrete derivative: $\Delta_+\{\cdot\} \quad \longleftrightarrow \quad 1 - e^{-j\omega}$

- Cardinal D-spline

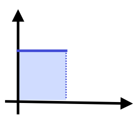
$$s(x) = \sum_{k\in\mathbb{Z}} s[k]\beta_+^0(x-k)$$



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



- B-spline function



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

Generalization: fractional B-splines

Derivative operator: $\mathrm{D} = \partial^1_{\frac{1}{2}} \quad \xleftrightarrow{\mathcal{F}} \quad j\omega$

Finite difference: $\Delta_+ \quad \xleftrightarrow{\mathcal{F}} \quad 1 - e^{-j\omega}$

Liouville's fractional derivative: $\mathrm{D}^\gamma = \partial^\gamma_{\gamma/2} \quad \xleftrightarrow{\mathcal{F}} \quad (j\omega)^\gamma$

Fractional finite differences: $\Delta^\gamma_+ \quad \xleftrightarrow{\mathcal{F}} \quad (1 - e^{-j\omega})^\gamma$

- Causal fractional B-splines

Discrete version of operator

Spline degree: $\alpha = \gamma - 1$

$$\frac{(1 - e^{-j\omega})^{\alpha+1}}{(j\omega)^{\alpha+1}} \quad \xrightarrow{\mathcal{F}^{-1}} \quad \beta^\alpha_+(x)$$

Continuous-domain operator: $\hat{L}(\omega)$

Causal fractional B-splines

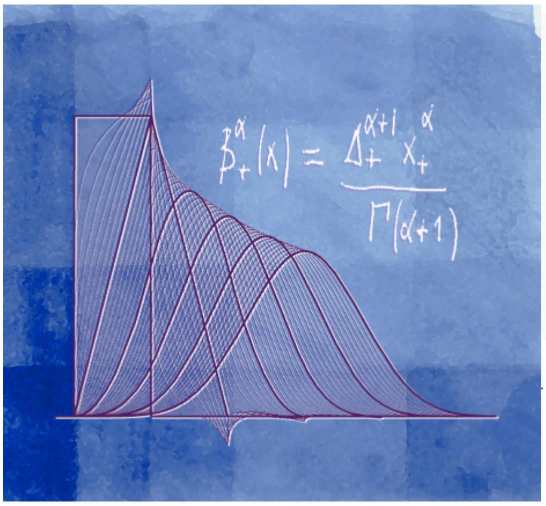
$$\mathrm{L} = \mathrm{D}^{\alpha+1} \xleftrightarrow{\mathcal{F}} (j\omega)^{\alpha+1}$$ (Liouville's fractional derivative)

$$\beta^0_+(x) = \Delta_+ x^0_+ \quad \xleftrightarrow{\mathcal{F}} \quad \frac{1 - e^{-j\omega}}{j\omega}$$

$$\vdots \qquad\qquad \vdots$$

$$\beta^\alpha_+(x) = \frac{\Delta^{\alpha+1}_+ x^\alpha_+}{\Gamma(\alpha+1)} \quad \xleftrightarrow{\mathcal{F}} \quad \left(\frac{1 - e^{-j\omega}}{j\omega}\right)^{\alpha+1}$$

One-sided power function: $x^\alpha_+ = \begin{cases} x^\alpha, & x \geq 0 \\ 0, & x < 0 \end{cases}$



(Unser & Blu, *SIAM Rev*, 2000)

$$\beta^\alpha_+(x) = \Delta^{\alpha+1} \mathrm{D}^{-(\alpha+1)}\{\delta\}(x)$$

Sparsifying fractal processes

- Fractional differential equation

$$\mathrm{D}^{\gamma} s = w$$
$$s = \mathrm{D}^{-\gamma} w$$

- Generalized increment process

$$\Delta_+^{\gamma} s = \Delta_+^{\gamma} \mathrm{D}^{-\gamma} w = \beta_+^{\gamma-1} * w$$

$$\langle \Delta_+^{\gamma} s, \varphi \rangle = \langle \beta_+^{\gamma-1} * w, \varphi \rangle = \langle w, \beta_-^{\gamma-1} * \varphi \rangle \quad \text{with} \quad \beta_-^{\gamma-1}(x) = \beta_+^{\gamma-1}(-x)$$

$$\Longrightarrow \quad Z_{\Delta_+^{\gamma} s}(\varphi) = Z_w(\beta_-^{\gamma} * \varphi)$$

- Statistical implications
  - $\Delta_+^{\gamma} s(x)$ is stationary with correlation function $(\beta_+^{\gamma-1} * \beta_-^{\gamma-1})(x)$
  - $\Delta_+^{\gamma} s(x_0)$ is sparse (mass distribution for 0) and nearly decorrelated
  - Integer case $\gamma = n$: $\Delta_+^{n} s(x_1)$ and $\Delta_+^{n} s(x_2)$ are independent if $|x_1 - x_2| \geq n$

Back to the Compound Poisson process

$$\mathrm{D}s = w \quad \Rightarrow \quad \Delta_+ s(x) = \beta_-^0 * w$$

- Statistical characterization of sampled increment process
  - The samples values $\{r(k) = \Delta_+ s(k)\}$ are independent, identically distributed
  - Explicit form of PDF: $p(r) = \mathcal{F}^{-1}\left\{\exp\left(\lambda \int_R (e^{ja\omega} - 1) p_a(a) \mathrm{d}a\right)\right\}(r)$

- Is there a MAP estimator that is equivalent to TV ? **Yes**
  - Choice of Lévy density: $\lambda \cdot p_a(a) = \dfrac{e^{-|a|}}{|a|}$
  - Relevant Fourier identity:

Laplace density

$$\exp \int_{\mathbb{R}} (e^{ja\omega} - 1) \frac{e^{-|a|}}{|a|} \mathrm{d}a = \frac{1}{1+\omega^2} = \mathcal{F}\left\{\frac{1}{2} e^{-|r|}\right\}(\omega)$$

  - Prior likelihood term: $R(s) = \sum_k -\log\big(p(\Delta_+ s(k)\big) \propto \sum_k |s(k) - s(k-1)|$

CONCLUSION

- Unifying operator-based paradigm
  - Backward compatibility with classical Gaussian theory
  - ***Operator*** identification based on ***invariance*** principles (TSR)
  - Characterization of ***stochastic processes*** (fractals)
    Gaussian vs. **sparse (generalized Poisson)**
  - Focus on non-stable PDEs ⇒ non-stationary, self-similar processes
- Wavelet analysis vs. regularization
  - Central role of B-spline
  - Sparsification via "operator-like" behavior
    - Multi-resolution: wavelets
    - Discrete-domain: finite-differences (generalized increments)
- Theoretical framework for sparse signal processing
  - New statistically-founded sparsity priors
  - Derivation of optimal estimators (MAP, MMSE)

Art of tempering Fourier-domain singularities ...

$$e^{jz} - 1 = jz + O(z^2)$$

- B-spline construction: $\beta^0_-(x) = \int_{\mathbb{R}} \frac{e^{j\omega} - 1}{j\omega} e^{j\omega x} \frac{d\omega}{2\pi}$

- Correct noise integrator: adjoint of L_p-stable inverse

$$\mathrm{D}^{-1} w(x) = \int_{\mathbb{R}} \hat{w}(\omega) \frac{e^{j\omega t} - 1}{j\omega} \frac{d\omega}{2\pi}$$

- Characteristic function(al) domain

$$\exp \int_{\mathbb{R}} (e^{ja\omega} - 1) \frac{e^{-|a|}}{|a|} da = \frac{1}{1+\omega^2} = \mathcal{F}\left\{\frac{1}{2} e^{-|r|}\right\}(\omega)$$

Epilogue: it is the unstable character of the underlying PDE that makes the processes interesting (e.g., piecewise smooth, long range-dependencies)

References

- M. Unser, T. Blu, "Fractional Splines and Wavelets," *SIAM Review*, vol. 42, no. 1, pp. 43-67, March 2000.
- M. Unser and T. Blu, "Generalized Smoothing Splines and the Optimal Discretization of the Wiener Filter," *IEEE Transactions on Signal Processing*, vol. 53, no. 6, pp. 2146–2159, June 2005.
- I. Khalidov and M. Unser, "From Differential Equations to the Construction of New Wavelet-Like Bases," *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1256–1267, April 2006.
- M. Unser and T. Blu, "Self-Similarity: Part I—Splines and Operators," *IEEE Transactions on Signal Processing*, vol. 55, no. 4, pp. 1352–1363, April 2007.
- T. Blu, M. Unser, "Self-Similarity: Part II—Optimal Estimation of Fractal Processes," *IEEE Trans. Signal Processing*, vol. 55, no. 4, pp. 1364-1378, April 2007.
- P.D. Tafti, D. Van De Ville, M. Unser, "Invariances, Laplacian-Like Wavelet Bases, and the Whitening of Fractal Processes," *IEEE Trans. Image Processing*, vol. 18, no. 4, pp. 689-702, April 2009.
- M. Unser, P.D. Tafti, "Stochastic models for sparse and piecewise-smooth signals", *IEEE Trans. Signal Processing*, in press.

Acknowledgments

Many thanks to

- Pouya Tafti
- Prof. Thierry Blu
- Prof. Dimitri Van De Ville
- Dr. Ildar Khalidov
- Arash Amini
- Annette Unser, Artist

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

erc

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