A Bayesian framework for the multifractal analysis of images using data augmentation and a Whittle approximation

S. Combrexelle[1,*], H. Wendt[1], Y. Altmann[2], J.-Y. Tourneret[1], S. McLaughlin[2], P. Abry[3]

[1] IRIT - ENSEEIHT, Toulouse, France
[2] Heriot-Watt University, Edinburgh, Scotland
[3] Ecole Normale Supérieure, Lyon, France
* Supported by the Direction Générale de l'Armement (DGA)

ICASSP 2016, Shanghai

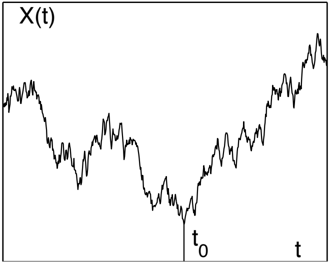
Introduction

- Multifractal analysis (MFA)
  - a widely used tool in signal/image processing
  - applications in various fields (texture analysis, . . .)
  - challenging estimation for images of small sizes
- Recent work
  - statistical estimation procedure based on a Bayesian framework
  - excellent estimation performance
  - high computational cost/ designed only for univariate data
- Contribution
  - more efficient estimation algorithm
  - suitable for extension to the analysis of collections of data

Characterization by local regularity

- Local regularity of $X(t)$ at t_0

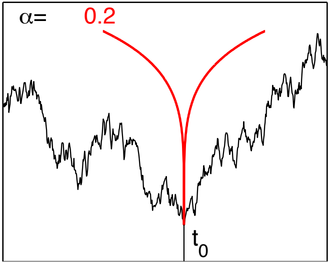
$$X \in C^{\alpha}(t_0) \quad \text{if} \quad \exists K, \alpha > 0;\ P_{t_0}(t);\ deg(P_{t_0}) < \alpha :$$
$$|X(t) - P_{t_0}(t)| < K|t - t_0|^{\alpha}$$



Characterization by local regularity

- Local regularity of $X(t)$ at t_0

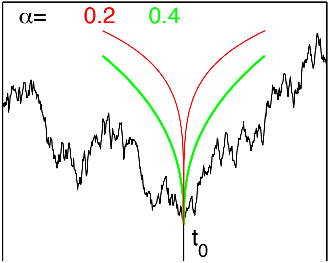
$$X \in C^{\alpha}(t_0) \quad \text{if} \quad \exists K, \alpha > 0;\ P_{t_0}(t);\ deg(P_{t_0}) < \alpha :$$
$$|X(t) - P_{t_0}(t)| < K|t - t_0|^{\alpha}$$



Characterization by local regularity

- Local regularity of $X(t)$ at t_0

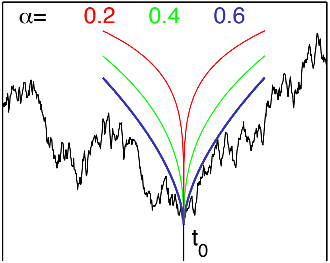
$$X \in C^{\alpha}(t_0) \quad \text{if} \quad \exists K, \alpha > 0;\ P_{t_0}(t);\ deg(P_{t_0}) < \alpha :$$
$$|X(t) - P_{t_0}(t)| < K|t - t_0|^{\alpha}$$



Characterization by local regularity

- Local regularity of $X(t)$ at t_0

$$X \in C^{\alpha}(t_0) \quad \text{if} \quad \exists K, \alpha > 0;\ P_{t_0}(t);\ deg(P_{t_0}) < \alpha :$$
$$|X(t) - P_{t_0}(t)| < K|t - t_0|^{\alpha}$$



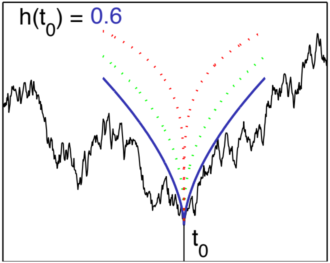
Characterization by local regularity

- Local regularity of $X(t)$ at t_0

$$X \in C^{\alpha}(t_0) \quad \text{if} \quad \exists K, \alpha > 0;\ P_{t_0}(t);\ deg(P_{t_0}) < \alpha :$$
$$|X(t) - P_{t_0}(t)| < K|t - t_0|^{\alpha}$$

- Hölder exponent:

$$h(t_0) = \sup_{\alpha}\{\alpha :\ X \in C^{\alpha}(t_0)\}$$



Characterization by local regularity

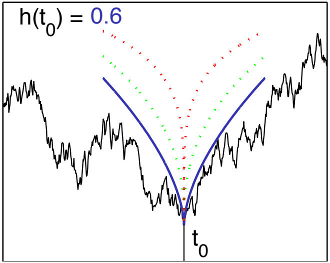
- Local regularity of $X(t)$ at t_0

$$X \in C^{\alpha}(t_0) \quad \text{if} \quad \exists K, \alpha > 0;\ P_{t_0}(t);\ deg(P_{t_0}) < \alpha :$$
$$|X(t) - P_{t_0}(t)| < K|t - t_0|^{\alpha}$$

- Hölder exponent:

$$h(t_0) = \sup_{\alpha}\{\alpha :\ X \in C^{\alpha}(t_0)\}$$

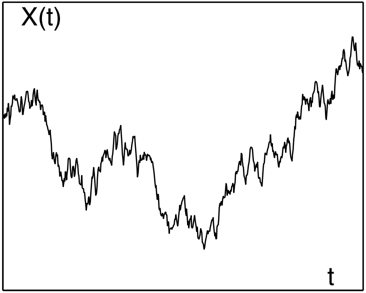
$h(t_0) \to 1 \Rightarrow$ smooth, very regular
$h(t_0) \to 0 \Rightarrow$ rough, very irregular

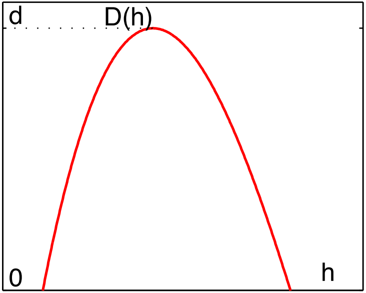


Multifractal spectrum

- Multifractal spectrum $D(h)$:
  - geometric description of fluctuations of the local regularity
  - Haussdorf dimension of the sets $\{t_i | h(t_i) = h\}$

$$D(h) = \dim_H\{t :\ h(t) = h\}$$

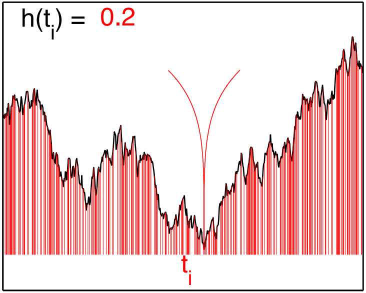


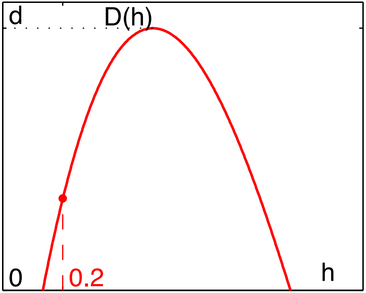


Multifractal spectrum

- Multifractal spectrum $D(h)$:
  - geometric description of fluctuations of the local regularity
  - Haussdorf dimension of the sets $\{t_i | h(t_i) = h\}$

$$D(h) = \dim_H\{t : \; h(t) = h\}$$

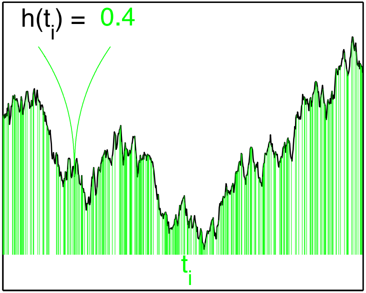


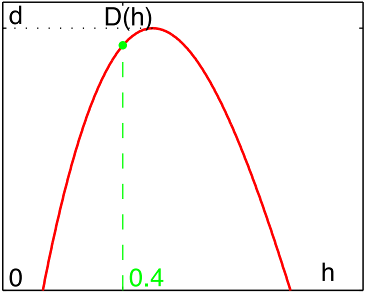


Multifractal spectrum

- Multifractal spectrum $D(h)$:
  - geometric description of fluctuations of the local regularity
  - Haussdorf dimension of the sets $\{t_i | h(t_i) = h\}$

$$D(h) = \dim_H \{t : \ h(t) = h\}$$

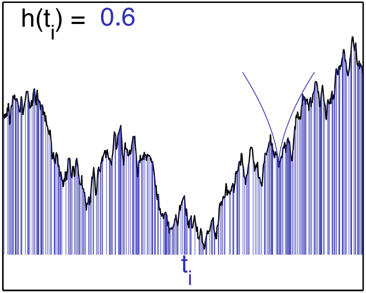


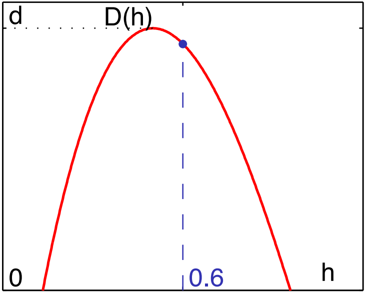


Multifractal spectrum

- Multifractal spectrum $D(h)$:
  - geometric description of fluctuations of the local regularity
  - Haussdorf dimension of the sets $\{t_i | h(t_i) = h\}$

$$D(h) = \dim_H \{t : \; h(t) = h\}$$



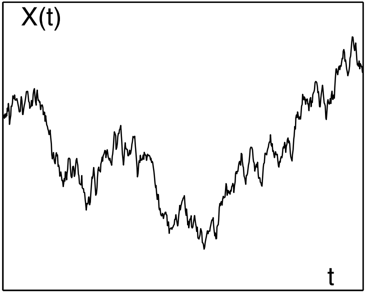


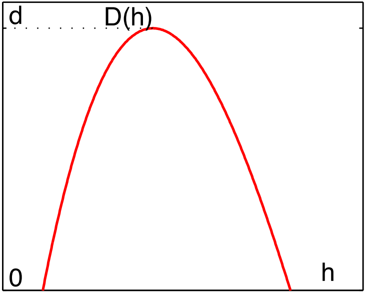
Multifractal spectrum

- Multifractal spectrum $D(h)$:
  - geometric description of fluctuations of the local regularity
  - Haussdorf dimension of the sets $\{t_i | h(t_i) = h\}$

$$D(h) = \dim_H\{t : \; h(t) = h\}$$

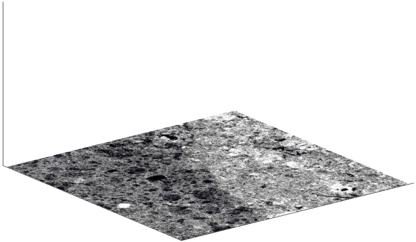
- In practice $\rightarrow$ multifractal formalism [Parisi85]





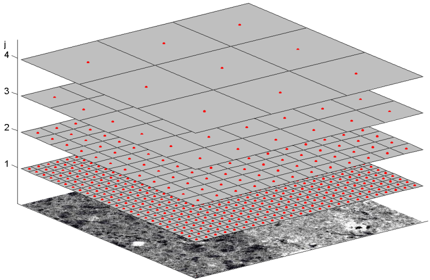
Multifractal formalism

- Dyadic wavelet transform: image **X**



Multifractal formalism

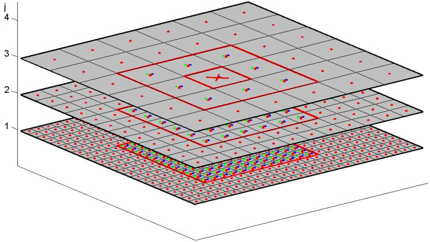
- Dyadic wavelet transform: image $\mathbf{X} \to \{d^{(m)}(j,\cdot,\cdot)\}_{m=1,2,3}$



Multifractal formalism

- Dyadic wavelet transform: image $\mathbf{X} \to \{d^{(m)}(j,\cdot,\cdot)\}_{m=1,2,3}$
- Wavelet leaders $\{\ell(j\cdot,\cdot)\}$ [Jaffard04]

$$\ell(j,k_1,k_2) = \sup_{m,\, \lambda' \subset 3\lambda_{j,k_1,k_2}} |d^{(m)}(\lambda')|$$



Multifractal formalism

- Dyadic wavelet transform: image $\mathbf{X} \rightarrow \{d^{(m)}(j,\cdot,\cdot)\}_{m=1,2,3}$
- Wavelet leaders $\{\ell(j\cdot,\cdot)\}$ [Jaffard04]

$$\ell(j,k_1,k_2) = \sup_{m,\,\lambda' \subset 3\lambda_{j,k_1,k_2}} |d^{(m)}(\lambda')|$$

- Polynomial expansion [Castaing93]

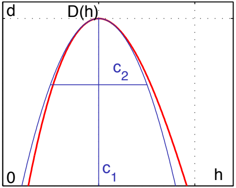
$$D(h) \approx 2 + \frac{c_2}{2!}\left(\frac{h-c_1}{c_2}\right)^2 - \frac{c_3}{3!}\left(\frac{h-c_1}{c_2}\right)^3 + \dots$$

$\rightarrow c_p$ tied to multiscale statistics of log-leaders $\{\ln \ell(j\cdot,\cdot)\}$

- Multifractality parameter c_2
  - tied to the variance of $\{\ln \ell(j\cdot,\cdot)\}$

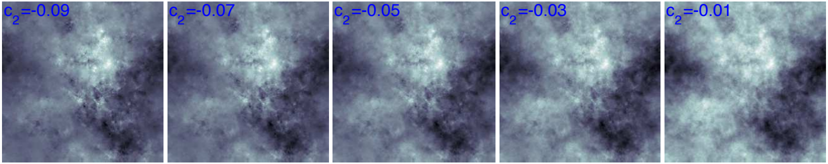
$$\boxed{\mathrm{Var}\left[\,\ln \ell(j,\cdot,\cdot)\,\right] = c_2^0 + c_2 \ln 2^j}$$

  - width of the multifractal spectrum $D(h)$
  - self-similar processes $\leftrightarrow$ multiplicative cascades



Estimation of the multifractality parameter

- Synthetic multiplicative cascades with different c_2



- Estimation of c_2
  - benchmark: linear regression-based estimation [Castaing93]

$$\boxed{\mathrm{Var}\,[\,\ln \ell(j,\cdot,\cdot)\,] = c_2^0 + c_2 \ln 2^j}$$

    X estimation performance

  - recently proposed Bayesian estimation [Combrexelle15]

    √ estimation performance

    X computational cost: Metropolis-Hasting (MH) moves

Space-domain statistical model of log-leaders

1. Marginal distribution of log-leaders approximated by Gaussian

$$\boxed{l(j,\cdot,\cdot) = \ln L(j,\cdot,\cdot) \ \sim\ \mathcal{N}\left(\cdot, c_2^0 + c_2 \ln 2^j\ \right)}$$

2. Intra-scale dependence captured by a radial covariance model

$$\boxed{\mathrm{Cov}[l(j,\boldsymbol{k}), l(j,\boldsymbol{k}+\boldsymbol{\Delta k})] \overset{\Delta r = |\boldsymbol{\Delta k}|}{\approx} \varrho_j(\Delta r;\boldsymbol{\theta}), \quad \boldsymbol{\theta} = (c_2, c_2^0)}$$

- Likelihood of centered log-leaders $\boldsymbol{l}_j$ stacked in $\boldsymbol{l} = [\boldsymbol{l}_{j_1}^T, ..., \boldsymbol{l}_{j_2}^T]^T$

  $\rightarrow$ scale-wise product of Gaussian likelihoods

$$p(\boldsymbol{l}|\boldsymbol{\theta}) \propto \prod_{j=j_1}^{j_2} |\boldsymbol{\Sigma}_{j,\boldsymbol{\theta}}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\boldsymbol{l}_j^T \boldsymbol{\Sigma}_{j,\boldsymbol{\theta}}^{-1} \boldsymbol{l}_j\right), \text{ with } \boldsymbol{\Sigma}_{j,\boldsymbol{\theta}} \text{ induced by } \varrho_j(\Delta r;\boldsymbol{\theta})$$

X *evaluation of $p(\boldsymbol{l}|\boldsymbol{\theta})$ numerically unstable*

[Combrexelle15]

Whittle approximation

- Evaluation of the Gaussian likelihood in the spectral domain

$$p_W(\boldsymbol{l}|\boldsymbol{\theta}) \propto \prod_{j=j_1}^{j_2} |\boldsymbol{\Gamma}_{j,\boldsymbol{\theta}}|^{-1} \exp\left(-\boldsymbol{y}_j^H \boldsymbol{\Gamma}_{j,\boldsymbol{\theta}}^{-1} \boldsymbol{y}_j\right)$$

  - $\boldsymbol{y}_j$ Fourier coefficients of $\boldsymbol{l}_j$
  - $\boldsymbol{\Gamma}_{j,\boldsymbol{\theta}}$ parametric spectral density associated with $\varrho_j(\Delta r; \boldsymbol{\theta})$

  $\rightarrow$ closed-form expression via Hankel transform

$$\boldsymbol{\Gamma}_{j,\boldsymbol{\theta}} = c_2 \, \mathbf{F}_{1,j} + c_2^0 \, \mathbf{F}_{2,j} \quad \text{with} \quad \mathbf{F}_{i,j} = \operatorname{diag}(\mathbf{f}_{i,j})$$

[Combrexelle15]

Whittle approximation

- Evaluation of the Gaussian likelihood in the spectral domain

$$p_W(\boldsymbol{l}|\boldsymbol{\theta}) \propto \prod_{j=j_1}^{j_2} |\boldsymbol{\Gamma}_{j,\boldsymbol{\theta}}|^{-1} \exp\left(-\boldsymbol{y}_j^H \boldsymbol{\Gamma}_{j,\boldsymbol{\theta}}^{-1} \boldsymbol{y}_j\right)$$

  - $\boldsymbol{y}_j$ Fourier coefficients of $\boldsymbol{l}_j$
  - $\boldsymbol{\Gamma}_{j,\boldsymbol{\theta}}$ parametric spectral density associated with $\varrho_j(\Delta r; \boldsymbol{\theta})$

  $\rightarrow$ closed-form expression via Hankel transform

$$\boldsymbol{\Gamma}_{j,\boldsymbol{\theta}} = c_2\, \mathbf{F}_{1,j} + c_2^0\, \mathbf{F}_{2,j} \quad \text{with} \quad \mathbf{F}_{i,j} = \mathrm{diag}(\mathbf{f}_{i,j})$$

- Bayesian estimation of $\boldsymbol{\theta} = (c_2, c_2^0)$
  - space-domain likelihood (approximated) + common priors

  X *non-standard posterior distribution $\rightarrow$ acceptance/reject moves*

[Combrexelle15]

C1 - Fourier-domain statistical model

- Whittle approximation

$$p_W(\boldsymbol{l}|\boldsymbol{\theta}) \propto \prod_{j=j_1}^{j_2} |\boldsymbol{\Gamma}_{j,\boldsymbol{\theta}}|^{-1} \exp\left(-\boldsymbol{y}_j^H \boldsymbol{\Gamma}_{j,\boldsymbol{\theta}}^{-1} \boldsymbol{y}_j\right)$$

  - $\boldsymbol{y}_j$ Fourier coefficients of $\boldsymbol{l}_j$
  - $\boldsymbol{\Gamma}_{j,\boldsymbol{\theta}} = c_2\, \mathbf{F}_{1,j} + c_2^0\, \mathbf{F}_{2,j}$ parametric spectral density

$$\Updownarrow$$

- Generative model for $\boldsymbol{y} = [\boldsymbol{y}_{j_1}^T, ..., \boldsymbol{y}_{j_2}^T]^T$

$$p(\boldsymbol{y}|\theta) \propto |\boldsymbol{\Gamma}_{\boldsymbol{\theta}}|^{-1} \exp\left(-\boldsymbol{y}^H \boldsymbol{\Gamma}_{\boldsymbol{\theta}}^{-1} \boldsymbol{y}\right)$$

  - complex Gaussian model $\boldsymbol{y} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Gamma}_{\boldsymbol{\theta}})$
  - $\boldsymbol{\Gamma}_{\boldsymbol{\theta}} = c_2 \boldsymbol{F}_1 + c_2^0 \boldsymbol{F}_2$ and $\boldsymbol{F}_i = \mathrm{diag}(\boldsymbol{F}_{i,j_1}, \ldots, \boldsymbol{F}_{i,j_2})$

X *model not separable in* (c_2, c_2^0)

C1 - Fourier-domain statistical model

- Whittle approximation

$$p_W(\boldsymbol{l}|\boldsymbol{\theta}) \propto \prod_{j=j_1}^{j_2} |\boldsymbol{\Gamma}_{j,\boldsymbol{\theta}}|^{-1} \exp\left(-\boldsymbol{y}_j^H \boldsymbol{\Gamma}_{j,\boldsymbol{\theta}}^{-1} \boldsymbol{y}_j\right)$$

  - $\boldsymbol{y}_j$ Fourier coefficients of $\boldsymbol{l}_j$
  - $\boldsymbol{\Gamma}_{j,\boldsymbol{\theta}} = c_2\, \mathbf{F}_{1,j} + c_2^0\, \mathbf{F}_{2,j}$ parametric spectral density

$$\Updownarrow$$

- Generative model for $\boldsymbol{y} = [\boldsymbol{y}_{j_1}^T, ..., \boldsymbol{y}_{j_2}^T]^T$

$$p(\boldsymbol{y}|\boldsymbol{\theta}) \propto |\boldsymbol{\Gamma}_{\boldsymbol{\theta}}|^{-1} \exp\left(-\boldsymbol{y}^H \boldsymbol{\Gamma}_{\boldsymbol{\theta}}^{-1} \boldsymbol{y}\right)$$

  - complex Gaussian model $\boldsymbol{y} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Gamma}_{\boldsymbol{\theta}})$
  - $\boldsymbol{\Gamma}_{\boldsymbol{\theta}} = c_2 \boldsymbol{F}_1 + c_2^0 \boldsymbol{F}_2$ and $\boldsymbol{F}_i = \mathrm{diag}(\boldsymbol{F}_{i,j_1}, \ldots, \boldsymbol{F}_{i,j_2})$

X *model not separable in* (c_2, c_2^0)

C2 - Reparametrization

- Non-separable constraints on (c_2, c_2^0)

$$\boldsymbol{\theta} \in \mathcal{A} = \{(c_2, c_2^0) \in \mathbb{R}_\star^- \times \mathbb{R}_\star^+ | \boldsymbol{\Gamma}_{\boldsymbol{\theta}} = c_2 \mathbf{F}_1 + c_2^0 \mathbf{F}_2 \text{ positive-definite}\}$$

- Design of a linear diffeomorphism ψ
  1. maps joint constraints into independent positivity constraints

$$\begin{aligned} \psi : \mathcal{A} &\to \mathbb{R}_\star^{+2} \\ : \boldsymbol{\theta} &\mapsto \psi(\boldsymbol{\theta}) \triangleq \tilde{\boldsymbol{\theta}} \end{aligned}$$

  2. leads to a more convenient likelihood

$$p(\boldsymbol{y}|\tilde{\boldsymbol{\theta}}) \propto |\boldsymbol{\Gamma}_{\tilde{\boldsymbol{\theta}}}|^{-1} \exp\left(-\boldsymbol{y}^H \boldsymbol{\Gamma}_{\tilde{\boldsymbol{\theta}}}^{-1} \boldsymbol{y}\right) \quad \text{with}$$

$$\text{for } \tilde{\boldsymbol{\theta}} \in \mathbb{R}_\star^{+2} \begin{cases} \boldsymbol{\Gamma}_{\tilde{\boldsymbol{\theta}}} = \tilde{\theta}_1 \tilde{\boldsymbol{F}}_1 + \tilde{\theta}_2 \tilde{\boldsymbol{F}}_2 & \text{positive-definite} \\ \tilde{\theta}_1 \tilde{\boldsymbol{F}}_1 \text{ and } \tilde{\theta}_2 \tilde{\boldsymbol{F}}_2 & \text{positive-definite} \end{cases}$$

$\rightarrow$ *likelihood separable via data augmentation*

C2 - Reparametrization

- Non-separable constraints on (c_2, c_2^0)

$$\boldsymbol{\theta} \in \mathcal{A} = \{(c_2, c_2^0) \in \mathbb{R}_\star^- \times \mathbb{R}_\star^+ | \boldsymbol{\Gamma}_{\boldsymbol{\theta}} = c_2 \mathbf{F}_1 + c_2^0 \mathbf{F}_2 \text{ positive-definite}\}$$

- Design of a linear diffeomorphism ψ

  1. maps joint constraints into independent positivity constraints

$$\psi : \mathcal{A} \to \mathbb{R}_\star^{+2}$$
$$: \boldsymbol{\theta} \mapsto \psi(\boldsymbol{\theta}) \triangleq \tilde{\boldsymbol{\theta}}$$

  2. leads to a more convenient likelihood

$$p(\boldsymbol{y}|\tilde{\theta}) \propto |\boldsymbol{\Gamma}_{\tilde{\theta}}|^{-1} \exp\left(-\boldsymbol{y}^H \boldsymbol{\Gamma}_{\tilde{\theta}}^{-1} \boldsymbol{y}\right) \quad \text{with}$$

$$\text{for } \tilde{\theta} \in \mathbb{R}_\star^{+2} \begin{cases} \boldsymbol{\Gamma}_{\tilde{\theta}} = \tilde{\theta}_1 \tilde{\boldsymbol{F}}_1 + \tilde{\theta}_2 \tilde{\boldsymbol{F}}_2 & \text{positive-definite} \\ \tilde{\theta}_1 \tilde{\boldsymbol{F}}_1 \text{ and } \tilde{\theta}_2 \tilde{\boldsymbol{F}}_2 & \text{positive-definite} \end{cases}$$

$\rightarrow$ *likelihood separable via data augmentation*

C2 - Reparametrization

- Non-separable constraints on (c_2, c_2^0)

$$\boldsymbol{\theta} \in \mathcal{A} = \{(c_2, c_2^0) \in \mathbb{R}_\star^- \times \mathbb{R}_\star^+ | \boldsymbol{\Gamma}_{\boldsymbol{\theta}} = c_2 \mathbf{F}_1 + c_2^0 \mathbf{F}_2 \text{ positive-definite}\}$$

- Design of a linear diffeomorphism ψ

1. maps joint constraints into independent positivity constraints

$$\psi : \mathcal{A} \to \mathbb{R}_\star^{+2}$$
$$: \boldsymbol{\theta} \mapsto \psi(\boldsymbol{\theta}) \triangleq \tilde{\boldsymbol{\theta}}$$

2. leads to a more convenient likelihood

$$p(\boldsymbol{y}|\tilde{\boldsymbol{\theta}}) \propto |\boldsymbol{\Gamma}_{\tilde{\boldsymbol{\theta}}}|^{-1} \exp\left(-\boldsymbol{y}^H \boldsymbol{\Gamma}_{\tilde{\boldsymbol{\theta}}}^{-1} \boldsymbol{y}\right) \quad \text{with}$$

$$\text{for } \tilde{\boldsymbol{\theta}} \in \mathbb{R}_\star^{+2} \begin{cases} \boldsymbol{\Gamma}_{\tilde{\boldsymbol{\theta}}} = \tilde{\theta}_1 \tilde{\boldsymbol{F}}_1 + \tilde{\theta}_2 \tilde{\boldsymbol{F}}_2 & \text{positive-definite} \\ \tilde{\theta}_1 \tilde{\boldsymbol{F}}_1 \text{ and } \tilde{\theta}_2 \tilde{\boldsymbol{F}}_2 & \text{positive-definite} \end{cases}$$

$\rightarrow$ *likelihood separable via data augmentation*

C3 - Data augmentation

- Definition of an augmented model

$$\begin{cases} \boldsymbol{y}|\boldsymbol{\mu},\tilde{\theta}_2 \sim \mathcal{CN}(\boldsymbol{\mu},\tilde{\theta}_2\tilde{\boldsymbol{F}}_2) & \text{observed data} \\ \boldsymbol{\mu}|\tilde{\theta}_1 \sim \mathcal{CN}(\boldsymbol{0},\tilde{\theta}_1\tilde{\boldsymbol{F}}_1) & \text{hidden mean} \end{cases}$$

with augmented likelihood

$$p(\boldsymbol{y},\boldsymbol{\mu}|\tilde{\boldsymbol{\theta}}) \propto \tilde{\theta}_2{}^{-N_Y} \exp\left(-\frac{1}{\tilde{\theta}_2}(\boldsymbol{y}-\boldsymbol{\mu})^H \tilde{\boldsymbol{F}}_2^{-1}(\boldsymbol{y}-\boldsymbol{\mu})\right) \times \tilde{\theta}_1{}^{-N_Y} \exp\left(-\frac{1}{\tilde{\theta}_1}\boldsymbol{\mu}^H \tilde{\boldsymbol{F}}_1^{-1}\boldsymbol{\mu}\right)$$

- Virtues of the augmented likelihood

√ recovers the original likelihood by marginalization

$$p(\boldsymbol{y}|\tilde{\theta}) = \int p(\boldsymbol{y},\boldsymbol{\mu}|\tilde{\theta})d\boldsymbol{\mu}$$

√ *separable in* $(\tilde{\theta}_1,\tilde{\theta}_2)$

√ *conjugate to inverse-gamma priors*

C3 - Data augmentation

- Definition of an augmented model

$$\begin{cases} \boldsymbol{y}|\boldsymbol{\mu}, \tilde{\theta}_2 \sim \mathcal{CN}(\boldsymbol{\mu}, \tilde{\theta}_2 \tilde{\boldsymbol{F}}_2) & \text{observed data} \\ \boldsymbol{\mu}|\tilde{\theta}_1 \sim \mathcal{CN}(\boldsymbol{0}, \tilde{\theta}_1 \tilde{\boldsymbol{F}}_1) & \text{hidden mean} \end{cases}$$

with augmented likelihood

$$p(\boldsymbol{y}, \boldsymbol{\mu}|\tilde{\boldsymbol{\theta}}) \propto \tilde{\theta}_2{}^{-N_Y} \exp\left(-\frac{1}{\tilde{\theta}_2}(\boldsymbol{y}-\boldsymbol{\mu})^H \tilde{\boldsymbol{F}}_2^{-1}(\boldsymbol{y}-\boldsymbol{\mu})\right) \times \tilde{\theta}_1{}^{-N_Y} \exp\left(-\frac{1}{\tilde{\theta}_1}\boldsymbol{\mu}^H \tilde{\boldsymbol{F}}_1^{-1}\boldsymbol{\mu}\right)$$

- Virtues of the augmented likelihood

√ recovers the original likelihood by marginalization

$$p(\boldsymbol{y}|\tilde{\boldsymbol{\theta}}) = \int p(\boldsymbol{y}, \boldsymbol{\mu}|\tilde{\boldsymbol{\theta}}) d\boldsymbol{\mu}$$

√ *separable in* $(\tilde{\theta}_1, \tilde{\theta}_2)$

√ *conjugate to inverse-gamma priors*

Bayesian model

- Bayesian model
  - augmented likelihood $p(\boldsymbol{y}, \boldsymbol{\mu} | \tilde{\boldsymbol{\theta}})$
  - independent inverse-gamma priors

$$\pi(\tilde{\theta}_i) \sim \mathcal{IG}(\alpha, \beta), \quad \alpha, \beta \ll 1$$

  - Bayes' theorem

$$p(\tilde{\boldsymbol{\theta}}, \boldsymbol{\mu} | \boldsymbol{y}) \propto p(\boldsymbol{y}, \boldsymbol{\mu} | \tilde{\boldsymbol{\theta}}) \times \prod_i \pi(\tilde{\theta}_i)$$

- Bayesian estimator
  - minimum mean squared error

$$\tilde{\boldsymbol{\theta}}^{\mathrm{MMSE}} = \mathbb{E}[\tilde{\boldsymbol{\theta}} | \boldsymbol{y}]$$

  - untractable integration over the posterior

    $\rightarrow$ computation via a Markov chain Monte Carlo algorithm

Bayesian model

- Bayesian model
  - augmented likelihood $p(\boldsymbol{y}, \boldsymbol{\mu} | \tilde{\boldsymbol{\theta}})$
  - independent inverse-gamma priors

$$\pi(\tilde{\theta}_i) \sim \mathcal{IG}(\alpha, \beta), \quad \alpha, \beta \ll 1$$

  - Bayes' theorem

$$p(\tilde{\boldsymbol{\theta}}, \boldsymbol{\mu} | \boldsymbol{y}) \propto p(\boldsymbol{y}, \boldsymbol{\mu} | \tilde{\boldsymbol{\theta}}) \times \prod_i \pi(\tilde{\theta}_i)$$

- Bayesian estimator
  - minimum mean squared error

$$\tilde{\boldsymbol{\theta}}^{\mathrm{MMSE}} = \mathbb{E}[\tilde{\boldsymbol{\theta}} | \boldsymbol{y}]$$

  - untractable integration over the posterior

    $\rightarrow$ computation via a Markov chain Monte Carlo algorithm

Markov chain Monte Carlo algorithm

- Generation of a collection of samples $(\tilde{\boldsymbol{\theta}}^{(k)}, \boldsymbol{\mu}^{(k)})_{k=1}^{N_{mc}}$ asymptotically distributed according to the posterior $p(\tilde{\boldsymbol{\theta}}, \boldsymbol{\mu} | \boldsymbol{y})$
- Approximation of the MMSE

$$\tilde{\boldsymbol{\theta}}^{\mathrm{MMSE}} \approx \frac{1}{N_{\mathrm{mc}} - N_{\mathrm{bi}}} \sum_{k=N_{\mathrm{bi}+1}}^{N_{\mathrm{mc}}} \tilde{\boldsymbol{\theta}}^{(k)}$$

- Gibbs sampler
  - iterative sampling according to conditional distributions

$$\tilde{\theta}_1 | \boldsymbol{y}, \boldsymbol{\mu}, \tilde{\theta}_2 \sim \mathcal{IG}(N_Y, \boldsymbol{\mu}^H \tilde{\boldsymbol{F}}_1^{-1} \boldsymbol{\mu})$$

$$\tilde{\theta}_2 | \boldsymbol{y}, \boldsymbol{\mu}, \tilde{\theta}_1 \sim \mathcal{IG}(N_Y, (\boldsymbol{y} - \boldsymbol{\mu})^H \tilde{\boldsymbol{F}}_2^{-1} (\boldsymbol{y} - \boldsymbol{\mu}))$$

$$\boldsymbol{\mu} | \boldsymbol{y}, \tilde{\boldsymbol{\theta}} \sim \mathcal{CN}\Big((\tilde{\theta}_1 \tilde{\boldsymbol{F}}_1 \boldsymbol{\Gamma}_{\tilde{\boldsymbol{\theta}}}^{-1}) \boldsymbol{y},\ ((\tilde{\theta}_1 \tilde{\boldsymbol{F}}_1)^{-1} + (\tilde{\theta}_2 \tilde{\boldsymbol{F}}_2)^{-1})^{-1} \Big)$$

√ *standard distributions → no Metropolis-Hasting (MH) moves*

Markov chain Monte Carlo algorithm

- Generation of a collection of samples $(\tilde{\boldsymbol{\theta}}^{(k)}, \boldsymbol{\mu}^{(k)})_{k=1}^{N_{mc}}$ asymptotically distributed according to the posterior $p(\tilde{\boldsymbol{\theta}}, \boldsymbol{\mu}|\boldsymbol{y})$
- Approximation of the MMSE

$$\tilde{\boldsymbol{\theta}}^{\mathrm{MMSE}} \approx \frac{1}{N_{\mathrm{mc}} - N_{\mathrm{bi}}} \sum_{k=N_{\mathrm{bi}+1}}^{N_{\mathrm{mc}}} \tilde{\boldsymbol{\theta}}^{(k)}$$

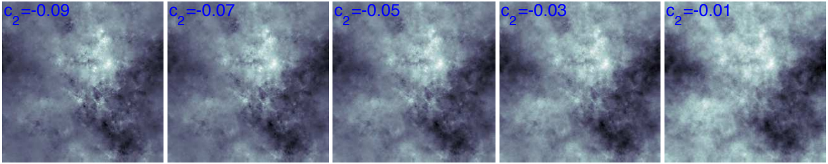
- Gibbs sampler
  - iterative sampling according to conditional distributions

$$\tilde{\theta}_1|\boldsymbol{y}, \boldsymbol{\mu}, \tilde{\theta}_2 \sim \mathcal{IG}\left(N_Y, \boldsymbol{\mu}^H \tilde{\boldsymbol{F}}_1^{-1} \boldsymbol{\mu}\right)$$
$$\tilde{\theta}_2|\boldsymbol{y}, \boldsymbol{\mu}, \tilde{\theta}_1 \sim \mathcal{IG}\left(N_Y, (\boldsymbol{y} - \boldsymbol{\mu})^H \tilde{\boldsymbol{F}}_2^{-1} (\boldsymbol{y} - \boldsymbol{\mu})\right)$$
$$\boldsymbol{\mu}|\boldsymbol{y}, \tilde{\boldsymbol{\theta}} \sim \mathcal{CN}\left(\left(\tilde{\theta}_1 \tilde{\boldsymbol{F}}_1 \boldsymbol{\Gamma}_{\tilde{\boldsymbol{\theta}}}^{-1}\right)\boldsymbol{y},\ \left((\tilde{\theta}_1 \tilde{\boldsymbol{F}}_1)^{-1} + (\tilde{\theta}_2 \tilde{\boldsymbol{F}}_2)^{-1}\right)^{-1}\right)$$

$\surd$ *standard distributions $\rightarrow$ no Metropolis-Hasting (MH) moves*

Numerical simulations

- Synthetic multiplicative cascades: 2D Multifractal Random Walk
  - non-Gaussian process
  - multifractal properties $\sim$ Mandelbrot's multiplicative cascades

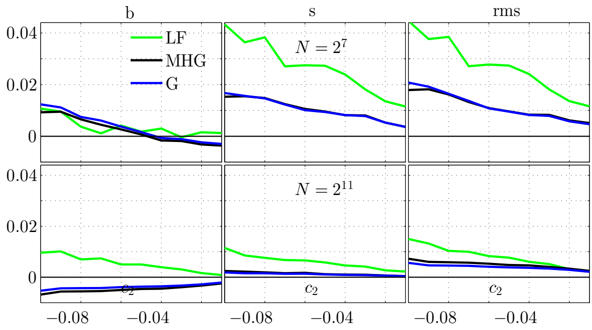


- Estimation setup
  - several image sizes $N \in \{2^7, 2^8, \ldots, 2^{11}\}$
  - Daubechies' mother wavelet $j_1 \in \{1, 2\}$, $j_2 = \ln_2 N - 4$
- Performance assessment

$$\mathrm{b} = \widehat{\mathbb{E}}[\hat{c}_2] - c_2, \quad \mathrm{s} = \sqrt{\widehat{\mathrm{Var}}[\hat{c}_2]}, \quad \mathrm{rms} = \sqrt{\mathrm{b}^2 + \mathrm{s}^2}$$

$\rightarrow$ computed over 100 independent realizations

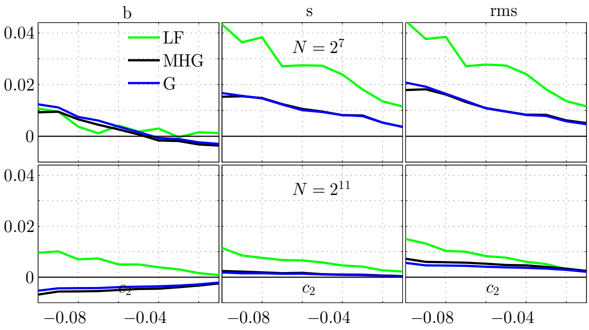
Results: estimation performance



LF	linear regression
MHG	space-domain model + Whittle approximation + uniform prior → Gibbs/MH [Combrexelle15]
G	Fourier-domain model + data augmentation + conjugate prior → Gibbs

- MHG vs G
  → slight difference for the bias, vanishing for large sample sizes
- LF vs Bayesian (MHG/G)
  → significant reduction of the standard deviation
  → root mean square error divided by 4

Results: estimation performance



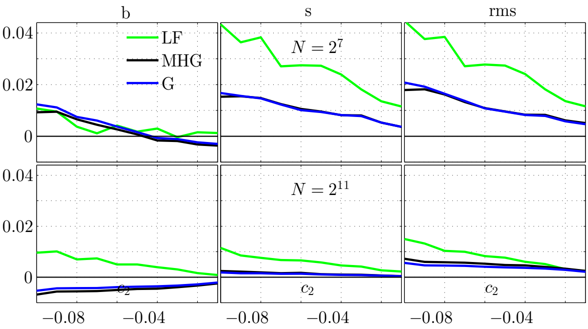
LF — linear regression

MHG — space-domain model
+ Whittle approximation
+ uniform prior
→ Gibbs/MH
[Combrexelle15]

G — Fourier-domain model
+ data augmentation
+ conjugate prior
→ Gibbs

- MHG vs G
  → slight difference for the bias, vanishing for large sample sizes
- LF vs Bayesian (MHG/G)
  → significant reduction of the standard deviation
  → root mean square error divided by 4

Results: estimation performance



LF — linear regression

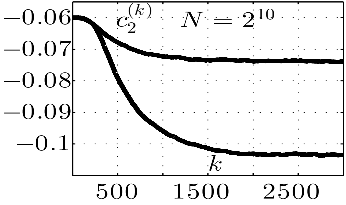
MHG — space-domain model
+ Whittle approximation
+ uniform prior
→ Gibbs/MH
[Combrexelle15]

G — Fourier-domain model
+ data augmentation
+ conjugate prior
→ Gibbs

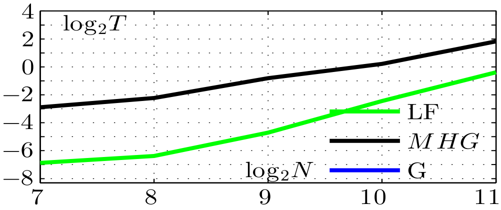
- MHG vs G
  → slight difference for the bias, vanishing for large sample sizes
- LF vs Bayesian (MHG/G)
  → significant reduction of the standard deviation
  → root mean square error divided by 4

Results: convergence and computational cost

Average Markov Chains



Computational time



- Convergence of Markov Chains

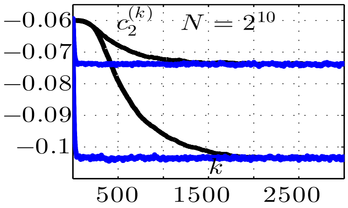
  $\rightarrow$ MHG long burn-in (tuning of the proposals) $\sim$ 3000 iterations

- Computational time $T =$ DWT + estimation algorithm

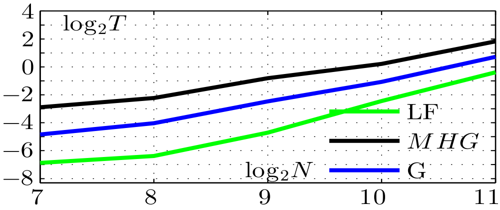
  $\rightarrow$ MHG $25 - 5$ times slower than LF

Results: convergence and computational cost

Average Markov Chains



Computational time



- Convergence of Markov Chains

  → MHG long burn-in (tuning of the proposals) ∼ 3000 iterations

  → G almost immediate convergence ∼ 600 iterations

- Computational time T = DWT + estimation algorithm

  → MHG $25 - 5$ times slower than LF

  → G $5 - 2$ times slower than LF

Conclusion and future work

- Conclusions
  - Bayesian estimation of c_2 for image texture
  - novel statistical model

$$\begin{cases} \text{generative model for Fourier coefficients (Whittle approximation)} \\ \text{separable model (reparametrization + data augmentation)} \end{cases}$$

  - efficient inference via a Gibbs sampler

  $\rightarrow$ excellent performance, competitive computational cost

- Work under investigation
  - MFA of multivariate data via the design of multivariate priors

Thanks for your attention

References

- [Parisi85] U. Frisch, and G. Parisi, *On the singularity structure of fully developed turbulence; appendix to Fully developped turbulence and intermittency, by U. Frisch*, in Proc. Int. Summer School Phys. Enrico Fermi, North-Holland, 1985, pp. 84-88
- [Jaffard04] S. Jaffard, *Wavelet techniques in multifractal analysis*, in Fractal Geometry and Applications: A Jubilee of Benoît Mandelbrot, Proc. Symp. Pure Math., M. Lapidus and M. van Frankenhuijsen, Eds. 2004, vol. 72(2), pp. 91-152, AMS
- [Castaing93] B. Castaing, Y. Gagne, and M. Marchand, *Log-similarity for turbulent flows?*, Physica D, vol. 68, no. 34, pp. 387-400, 1993
- [Combrexelle15] S. Combrexelle, H. Wendt, N. Dobigeon, J.-Y. Tourneret, S. McLaughlin, and P. Abry, *Bayesian Estimation of the Multifractality Parameter for Image Texture Using a Whittle Approximation*, IEEE T. Image Proces., vol. 24, no. 8, pp. 2540-2551, Aug. 2015
- [Combrexelle15] S. Combrexelle et al, *Bayesian Estimation of the Multifractality Parameter for Images Via a Closed-Form Whittle Likelihood*, Proc. 23rd European Signal Process. Conf. (EUSIPCO), Nice, France, 2015.