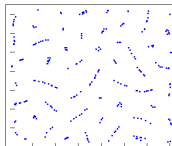
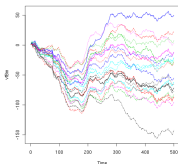


# Statistical Inference results for fractional processes and spatial Gibbs point processes



Jean-François Coeurjolly

Laboratoire Jean Kuntzmann (LJK) et Grenoble Images Parole Signal et Automatique  
(Gipsa-lab), Grenoble University



LABORATOIRE  
JEAN KUNTZMANN  
MATHÉMATIQUES APPLIQUÉES - INFORMATIQUE



gipsa-lab

# Research interests - Chronological aspect



## *I. Fractional Processes:*

- PhD Thesis (98-00): inference for fBm, optimality.
  - In LJK (01-): robust estimation, functional estimation, finite sample size aspect
  - In Gipsa-lab (09-11): multivariate aspect, application to FMRI.
-

# Research interests - Chronological aspect



## *I. Fractional Processes:*

- PhD Thesis (98-00): inference for fBm, optimality.
- In LJK (01-): robust estimation, functional estimation, finite sample size aspect
- In Gipsa-lab (09-11): multivariate aspect, application to FMRI.

## *II. Spatial Gibbs point processes (06-..):*

- started with my assistant teacher position at UPMF, to the contact of E. Bertin, JM Billiot and R. Drouilhet.
- In LJK (06-): parametric estimation, goodness-of-fit-tests for alternative models to Poisson point processes.
- In Gipsa-lab (11-): spatio-temporal p.p. on a sphere (sunspots)



# Research interests - Chronological aspect



## *I. Fractional Processes:*

- PhD Thesis (98-00): inference for fBm, optimality.
- In LJK (01-): robust estimation, functional estimation, finite sample size aspect
- In Gipsa-lab (09-11): multivariate aspect, application to FMRI.

## *II. Spatial Gibbs point processes (06-..):*

- started with my assistant teacher position at UPMF, to the contact of E. Bertin, JM Billiot and R. Drouilhet.
- In LJK (06-): parametric estimation, goodness-of-fit-tests for alternative models to Poisson point processes.
- In Gipsa-lab (11-): spatio-temporal p.p. on a sphere (sunspots)



## *III. Around applied statistics (01-..)*

- teaching statistics: experimental approach for visualizing complex concepts.
- since 2001: collaborations with economists, psychologists,...
- Co-supervision of two Phd students (06,10)
  - dimension reduction and information theory
  - multi-state models and epidemiology

# Research interests - Chronological aspect



## *I. Fractional Processes:*

- PhD Thesis (98-00): inference for fBm, optimality.
- In LJK (01-): robust estimation, functional estimation, finite sample size aspect
- In Gipsa-lab (09-11): multivariate aspect, application to FMRI.

## *II. Spatial Gibbs point processes (06-..):*

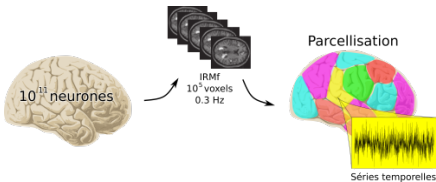
- started with my assistant teacher position at UPMF, to the contact of E. Bertin, JM Billiot and R. Drouilhet.
- In LJK (06-): parametric estimation, goodness-of-fit-tests for alternative models to Poisson point processes.
- In Gipsa-lab (11-): spatio-temporal p.p. on a sphere (sunspots)



## *III. Around applied statistics (01-..)*

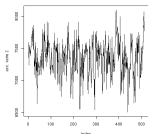
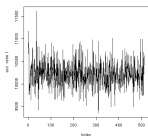
- teaching statistics: experimental approach for visualizing complex concepts.
- since 2001: collaborations with economists, psychologists,...
- Co-supervision of two Phd students (06,10)
  - dimension reduction and information theory
  - multi-state models and epidemiology

# Motivations (1): Analysis of FMRI for coma patients



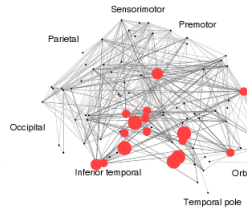
## Scientific questions

- Detect level of activation through regularity of signals.
- Find connections between regions of interest.

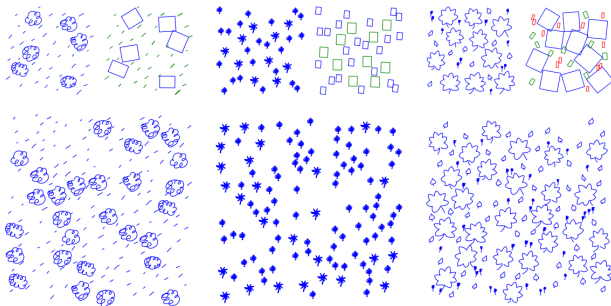


Objective: develop models and methodologies to deal with characteristics of FMRI time series

- 1 long memory, fractal, self-similar properties.
- 2 non stationary artefacts, outliers, additive noise.
- 3 strong spatial dependence.



# Motivations (2): Arrangement of geometrical objects



(Hurtut, Landes, Thollot, Gousseau, Drouilhet, C.'09.)

## Computer Graphics Problem

simulation at large scale of a pattern drawn in a small window by an artist.

## Statistical Problem

learn the complex spatial structure of the point pattern given by the centers of the objects (independence or interaction? regular or clustered distribution?).

# Common objective: develop methodologies and obtain asymptotic results in particular

---

## I. Fractional processes

## II. Spatial point processes

---

Data

Continuous stoch.  
discr. at times  $i = 0, \dots, n - 1$   
 $X = (X(0), \dots, X(n - 1))$

Set of locations  
observed in  $\Lambda_n \subset \mathbb{R}^d$   
 $\{x_1, \dots, x_n\}$

---

Asymptotic

$n \rightarrow +\infty$

$\Lambda_n \rightarrow \mathbb{R}^d$

---



# Common objective: develop methodologies and obtain asymptotic results in particular

## I. Fractional processes

## II. Spatial point processes

Data

Continuous stoch.  
discr. at times  $i = 0, \dots, n - 1$   
 $X = (X(0), \dots, X(n - 1))$

Set of locations  
observed in  $\Lambda_n \subset \mathbb{R}^d$   
 $\{x_1, \dots, x_n\}$

Asymptotic

$n \rightarrow +\infty$

$\Lambda_n \rightarrow \mathbb{R}^d$

Main result

$\widehat{\theta} \xrightarrow{a.s.} \theta^*$  and  $\sqrt{v_n}(\widehat{\theta} - \theta^*) \xrightarrow{d} \mathcal{N}(0, \Sigma(\theta^*))$   
where

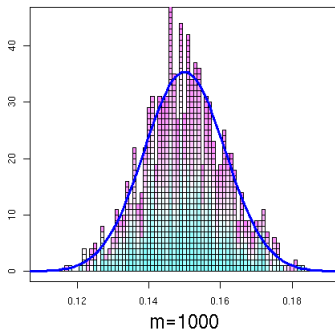
$\sqrt{v_n} =$

$\sqrt{n}$

$\sqrt{|\Lambda_n|}$

# To avoid a technical presentation ...

... in the following  
such a result could  
be summarized by  
the following  
figure ...



# To avoid a technical presentation ...

... or by following  
one ...



... an estimate  $\hat{\theta}$   
with nice properties

# Part I. Inference for fractional processes

1. Fractional models: FBM, MBM and MFBM.
2. Identification of models via filtering techniques.
3. Perspectives.

## 8 Papers (Int. Journ.)

Journ. Stat. Soft.'00  
 Stat. Inf. Stoch. Proc. '01  
 Stat. Prob. Lett.'01,'08  
 Bernoulli'05  
 Ann. Stat.'08  
 Stat. Surv.'10  
 Sém. Cong. SMF'10

**+ 2 Preprints subm.**



## Collaborators

S. Achard  
 P.O. Amblard  
 J. Istas  
 F. Lavancier  
 A. Philippe

# Fractional Brownian motion

Definition: unique centered Gaussian process with stationary and self-similar increments with parameters  $(H, \sigma) \in (0, 1) \times \mathbb{R}^+$ , with covariance function

$$E[X(s)X(t)] = \frac{\sigma^2}{2}(|s|^{2H} + |t|^{2H} - |t - s|^{2H})$$

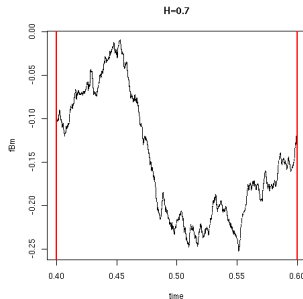
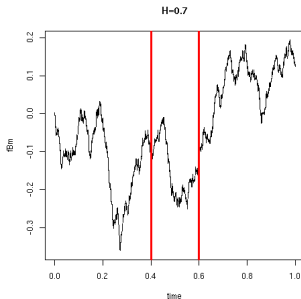
---

# Fractional Brownian motion

Definition: unique centered Gaussian process with stationary and self-similar increments with parameters  $(H, \sigma) \in (0, 1) \times \mathbb{R}^+$ , with covariance function

$$E[X(s)X(t)] = \frac{\sigma^2}{2} (|s|^{2H} + |t|^{2H} - |t - s|^{2H})$$

Properties: (1) Self-similarity:  $X(\lambda t) \stackrel{f.d.}{=} \lambda^H X(t)$

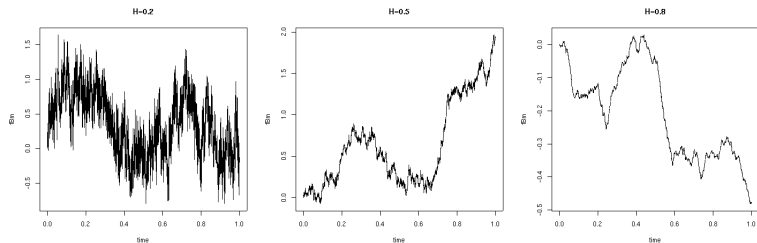


# Fractional Brownian motion

Definition: unique centered Gaussian process with stationary and self-similar increments with parameters  $(H, \sigma) \in (0, 1) \times \mathbb{R}^+$ , with covariance function

$$E[X(s)X(t)] = \frac{\sigma^2}{2} (|s|^{2H} + |t|^{2H} - |t - s|^{2H})$$

Properties: (2) Hausdorff dimension:  $2 - H$

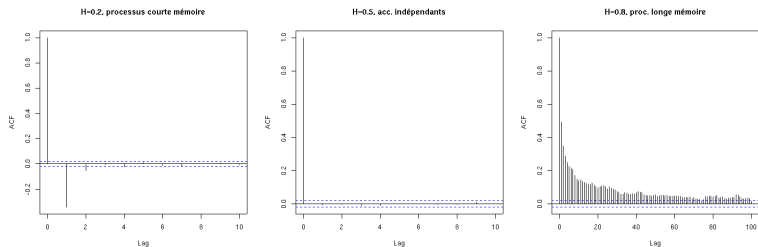


# Fractional Brownian motion

Definition: unique centered Gaussian process with stationary and self-similar increments with parameters  $(H, \sigma) \in (0, 1) \times \mathbb{R}^+$ , with covariance function

$$E[X(s)X(t)] = \frac{\sigma^2}{2} (|s|^{2H} + |t|^{2H} - |t-s|^{2H})$$

Properties: (3) let  $\gamma(\cdot)$  be the ACF  $\gamma(k) = \mathcal{O}(|k|^{2H-2})$





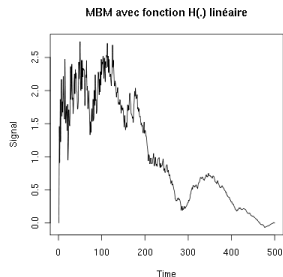
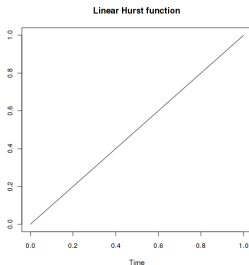
# Multifractional Brownian motion (MBM)

- Extension of the FBM allowing the regularity to vary with time [PLV'94,BCI'98].
- Intuitive definition: replace parameters by function of time (in stochastic representation integrals).

$$E[X(s)X(t)] = f_{s,t} (|s|^{H_{s,t}} + |t|^{H_{s,t}} - |t-s|^{H_{s,t}}),$$

where  $H_{s,t} = H(s) + H(t)$  and  $f_{s,t} = f(H(s), H(t), \sigma(s), \sigma(t))$ .

- **locally** self-similar and stationary increments.



# Multivariate fractional Brownian motion (MFBM)

- Recent Extension [LPS'10, DP'10] of the FBM to the multivariate case (could be used for modelling FMRI time series)
- Centered Gaussian process with stationary and self-similar increments

$$(X_1(\lambda t), \dots, X_p(\lambda t)) \stackrel{f.d.}{=} (\lambda^{H_1} X_1(t), \dots, \lambda^{H_p} X_p(t))$$

- The cross-covariances are necessarily written as (if  $H_i + H_j \neq 1$ )

[LPS'10] [ACLP'10]

$$E[X_i(s)X_j(t)] = \frac{\sigma_i \sigma_j}{2} \left\{ (\rho_{i,j} + \eta_{i,j} \text{sign}(s)) |s|^{H_i+H_j} + (\rho_{i,j} - \eta_{i,j} \text{sign}(t)) |t|^{H_i+H_j} - (\rho_{i,j} - \eta_{i,j} \text{sign}(t-s)) |t-s|^{H_i+H_j} \right\},$$

where  $H_i \in (0, 1)$ ,  $\sigma_i > 0$ ,  $\rho_{i,j} = \rho_{j,i} = E[X_i(1)X_j(1)] \in (-1, 1)$  and  $\eta_{i,j} = -\eta_{j,i} \in \mathbb{R}$ .

# A few properties of the MFBM

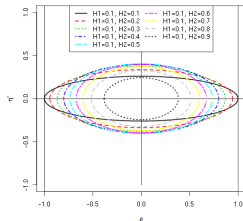
[ACLP'10,CAA'10]

- The cross-covariance (spectrum, wavelet transform, . . .) “behaves” like a FGN with parameter  $\frac{H_i+H_j}{2}$ . For example,  $\gamma_{i,j}(k) = \mathcal{O}(|k|^{H_i+H_j-2})$ .

# A few properties of the MFBM

[ACLP'10,CAA'10]

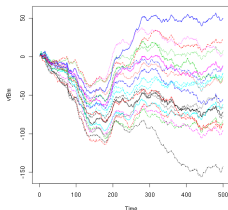
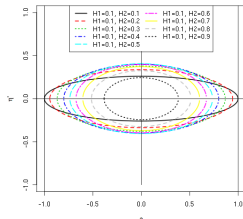
- The cross-covariance (spectrum, wavelet transform, ...) “behaves” like a FGN with parameter  $\frac{H_i+H_j}{2}$ . For example,  $\gamma_{i,j}(k) = \mathcal{O}(|k|^{H_i+H_j-2})$ .
- Existence conditions:  $E[X_i(s)X_j(t)]$  is a covariance function iff the matrix with entries  $\Gamma(H_{i,j} + 1) \left( \rho_{i,j} \sin\left(\frac{\pi}{2}H_{i,j}\right) - \mathbf{i} \eta_{i,j} \cos\left(\frac{\pi}{2}H_{i,j}\right) \right)$ , where  $H_{i,j} = H_i + H_j$ , is Hermitian.
- $p = 2$ , set of  $(\rho, \eta)$  described by an ellipse.



# A few properties of the MFBM

[ACLP'10,CAA'10]

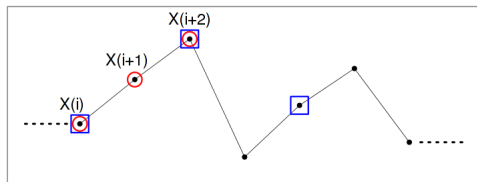
- The cross-covariance (spectrum, wavelet transform, ...) “behaves” like a FGN with parameter  $\frac{H_i+H_j}{2}$ . For example,  $\gamma_{i,j}(k) = \mathcal{O}(|k|^{H_i+H_j-2})$ .
- Existence conditions:  $E[X_i(s)X_j(t)]$  is a covariance function iff the matrix with entries  $\Gamma(H_{i,j} + 1) \left( \rho_{i,j} \sin\left(\frac{\pi}{2}H_{i,j}\right) - \mathbf{i} \eta_{i,j} \cos\left(\frac{\pi}{2}H_{i,j}\right) \right)$ , where  $H_{i,j} = H_i + H_j$ , is Hermitian.
- $p = 2$ , set of  $(\rho, \eta)$  described by an ellipse.



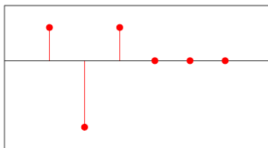
- Exact simulation can be done using general embedding circulant matrix method [wc97].
- Example:  $p = 20$ ,  $H_j \in [0.7, 0.8]$ ,  $\rho_{i,j} = 0.7$ ,  $\eta_{i,j} = 0$ .

# Identification using filtering techniques: a few notation

- $X = (X(0), \dots, X(n-1))$  sample of a fBm.
- $a$  filter with  $p$  zero moments, i.e.  $\sum_q a_q q^r = 0, r = 0, \dots, p-1$  (ex: increments, Daubechies, ...)
- $a^m$  filter  $a$  dilated  $m$  times
- $X^{a^m}$ : vector  $X$  filtered with  $a^m$ .

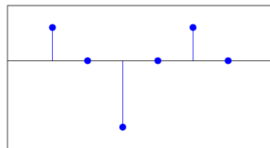


$$a = (1, -2, 1)$$



$$X^a(i) = X(i) - 2X(i+1) + X(i+2)$$

$$a^2 = (1, 0, -2, 0, 1)$$



$$X^{a^2}(i) = X(i) - 2X(i+2) + X(i+4)$$

# Applications to the estimation of a fBm

## Idea

- Self-similarity property of  $X^{\alpha^m}$ :  $E[X^{\alpha^m}(i)^2] = m^{2H} \underbrace{\sigma^2 E[X^\alpha(i)^2]}_{:=\gamma}$
- Let  $S^2(m)$  be the sample variance of  $X^{\alpha^m}$

$$\log(S^2(m)) \simeq 2H \log m + \log(\gamma)$$

## Remarks

- $H$  can be estimated by a simple linear regression.
- $\widehat{H}$  independent of scaling factor, quick procedure.
- Interest of filtering:  $\gamma^{\alpha^m}(k) = O(|k|^{2H-2p})$

# Applications to the estimation of a fBm

## Idea

- Self-similarity property of  $X^{\alpha^m}$ :  $E[X^{\alpha^m}(i)^2] = m^{2H} \underbrace{\sigma^2 E[X^\alpha(i)^2]}_{:=\gamma}$
- Let  $S^2(m)$  be the sample variance of  $X^{\alpha^m}$

$$\log(S^2(m)) \simeq 2H \log m + \log(\gamma)$$

## Remarks

- $H$  can be estimated by a simple linear regression.
- $\widehat{H}$  independent of scaling factor, quick procedure.
- Interest of filtering:  $\gamma^{\alpha^m}(k) = O(|k|^{2H-2p})$

⇒ standard results on subordinated Gaussian stationary sequences [T<sup>79</sup>,BM<sup>83</sup>] ensure nice asymptotic properties [C<sup>01</sup>]





# Applications to the estimation of a fBm

## Idea

- Self-similarity property of  $X^{a^m}$ :  $E[X^{a^m}(i)^2] = m^{2H} \underbrace{\sigma^2 E[X^a(i)^2]}_{:=\gamma}$
- Let  $S^2(m)$  be the sample variance of  $X^{a^m}$

$$\log(S^2(m)) \simeq 2H \log m + \log(\gamma)$$

## Remarks

- $H$  can be estimated by a simple linear regression.
- $\widehat{H}$  independent of scaling factor, quick procedure.
- Interest of filtering:  $\gamma^{a^m}(k) = O(|k|^{2H-2p})$

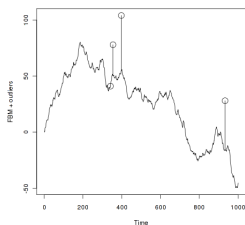
⇒ standard results on subordinated Gaussian stationary sequences [T'79,BM'83] ensure nice asymptotic properties [C'01]



- Same rate of convergence as the Cramèr-Rao-Bounds [C'01].
- The method can be "localized" to estimate the regularity function  $H(\cdot)$  of a MBM [C'05].

# Robustness to outliers

- Previous Procedure sensitive to **outliers**.

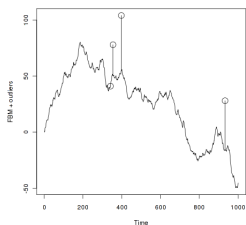


(parameters: 500 samples,  $n = 1000$ ,  $H = .8$ ,  $\sigma^2 = 1$ ,  $SNR = -30$  Db)

Method	Emp. Bias (Std Dev.)
No outliers	0.801 (0.021)
With outliers	0.329 (0.162)

# Robustness to outliers

- Previous Procedure sensitive to **outliers**.



(parameters: 500 samples,  $n = 1000$ ,  $H = .8$ ,  $\sigma^2 = 1$ ,  $SNR = -30$  Db)

Method	Emp. Bias (Std Dev.)
No outliers	0.801 (0.021)
With outliers	0.329 (0.162)

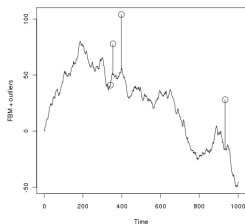
- **Idea:** replace  $S^2(m)$  by the empirical median of  $(X^{a^m})^2$ . Indeed,

$$\text{Med}(X^{a^m}(i)^2) = m^{2H} \underbrace{\gamma \text{Med}(\chi_1^2)}_{:=\gamma'} \Rightarrow \log \widehat{\text{Med}}((X^{a^m})^2) \simeq 2H \log(m) + \log(\gamma')$$

- methodology can be extended to trimmed-means,...

# Robustness to outliers

- Previous Procedure sensitive to **outliers**.



(parameters: 500 samples,  $n = 1000$ ,  $H = .8$ ,  $\sigma^2 = 1$ ,  $SNR = -30$  Db)

Method	Emp. Bias (Std Dev.)
No outliers	0.801 (0.021)
With outliers	0.329 (0.162)
<b>Median based procedure</b>	<b>0.798 (0.047)</b>

- **Idea:** replace  $S^2(m)$  by the empirical median of  $(X^{a^m})^2$ . Indeed,

$$\text{Med}(X^{a^m}(i)^2) = m^{2H} \underbrace{\gamma \text{Med}(\chi_1^2)}_{:=\gamma'} \Rightarrow \log \widehat{\text{Med}}((X^{a^m})^2) \simeq 2H \log(m) + \log(\gamma')$$

- methodology can be extended to trimmed-means,...

## Robustness to outliers (2)

- Nature of quantiles  $\Rightarrow$  results are more complicated to establish.
- **Key-ingredient:** (nice) Bahadur representation of sample quantiles for subordinated Gaussian stationary sequences.

## Robustness to outliers (2)

- Nature of quantiles  $\Rightarrow$  results are more complicated to establish.
- **Key-ingredient:** (nice) Bahadur representation of sample quantiles for subordinated Gaussian stationary sequences.

$g(Y) := (g(Y(1)), \dots, g(Y(n)))$  where  $Y$  is a Gaussian centered stationary sequence with variance 1, let  $p \in (0, 1)$  and  $\widehat{\xi}(p)$  be the emp. quantile of order  $p$

$$\widehat{\xi}(p) - \xi(p) = \frac{\widehat{F}(\xi(p)) - F(\xi(p))}{f(\xi(p))} + r_n,$$

where  $r_n \rightarrow 0$  (in a certain sense). For example, if  $g(Y)$  is short memory,

$$r_n = \begin{cases} O_{a.s.}(n^{-3/4} \log(n)^{3/4}) & \text{[C'08a] under } C^2\text{-type assumption on } g \\ O_P(n^{-1/2}) & \text{[C'08b] under mild assumption on } g \end{cases}$$

- Based on this representation ...

## Robustness to outliers (2)

- Nature of quantiles  $\Rightarrow$  results are more complicated to establish.
- **Key-ingredient:** (nice) Bahadur representation of sample quantiles for subordinated Gaussian stationary sequences.

$g(Y) := (g(Y(1)), \dots, g(Y(n)))$  where  $Y$  is a Gaussian centered stationary sequence with variance 1, let  $p \in (0, 1)$  and  $\widehat{\xi}(p)$  be the emp. quantile of order  $p$

$$\widehat{\xi}(p) - \xi(p) = \frac{\widehat{F}(\xi(p)) - F(\xi(p))}{f(\xi(p))} + r_n,$$

where  $r_n \rightarrow 0$  (in a certain sense). For example, if  $g(Y)$  is short memory,

$$r_n = \begin{cases} O_{a.s.}(n^{-3/4} \log(n)^{3/4}) & \text{[C'08a] under } C^2\text{-type assumption on } g \\ O_P(n^{-1/2}) & \text{[C'08b] under mild assumption on } g \end{cases}$$

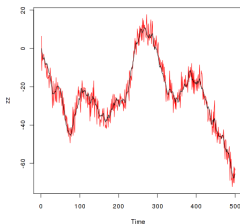


- Based on this representation ... [C'08a]

# Robustness to additive noise

- Standard procedure sensitive to an additive noise.

Assume  $X(i) = B_H(i) + \sigma_v^2 Z(i)$ , where  $\sigma_v > 0$  and  $Z(i)$  i.i.d  $\mathcal{N}(0, 1)$



(parameters  $n = 1000, H = .8, \sigma^2 = 1, SNR = 0$  Db)

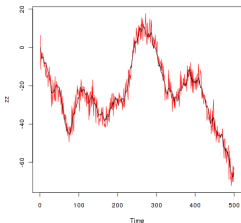
Method	Emp. Bias (Std Dev.)
No noise	0.801 (0.021)
With noise	0.543 (0.047)



# Robustness to additive noise

- Standard procedure sensitive to an additive noise.

Assume  $X(i) = B_H(i) + \sigma_v^2 Z(i)$ , where  $\sigma_v > 0$  and  $Z(i)$  i.i.d  $\mathcal{N}(0, 1)$



(parameters  $n = 1000, H = .8, \sigma^2 = 1, SNR = 0$  Db)

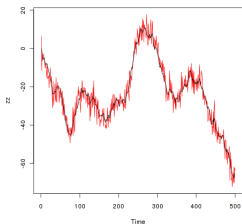
Method	Emp. Bias (Std Dev.)
No noise	0.801 (0.021)
With noise	0.543 (0.047)

- Why?  $E[X^{a^m}(i)^2] = m^{2H} \gamma + \sigma_v^2 \sum_q (a_q)^2$
- Idea to correct:  $E[X^{a^{2m}}(i)^2] - E[X^{a^m}(i)^2] = m^{2H} \underbrace{(2^{2H} - 1)}_{:=\gamma'} \gamma$   
 $\Rightarrow \log |S^2(2m) - S^2(m)| \simeq 2H \log(m) + \log \gamma'$

# Robustness to additive noise

- Standard procedure sensitive to an additive noise.

Assume  $X(i) = B_H(i) + \sigma_v^2 Z(i)$ , where  $\sigma_v > 0$  and  $Z(i)$  i.i.d  $\mathcal{N}(0, 1)$



(parameters: 500 samples,  $n = 1000$ ,  $H = .8$ ,  $\sigma^2 = 1$ ,  $SNR = 0$  Db)

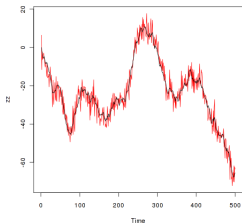
Method	Emp. Bias (Std Dev.)
No noise	0.801 (0.021)
With noise	0.543 (0.047)
<b>Diff. based procedure</b>	<b>0.807 (0.077)</b>

- Why?  $E[X^{a^m}(i)^2] = m^{2H} \gamma + \sigma_v^2 \sum_q (a_q)^2$
- Idea to correct:  $E[X^{a^{2m}}(i)^2] - E[X^{a^m}(i)^2] = m^{2H} \overbrace{(2^{2H} - 1)}{:=\gamma} \gamma$   
 $\Rightarrow \log |S^2(2m) - S^2(m)| \simeq 2H \log(m) + \log \gamma'$

# Robustness to additive noise

- Standard procedure sensitive to an additive noise.

Assume  $X(i) = B_H(i) + \sigma_v^2 Z(i)$ , where  $\sigma_v > 0$  and  $Z(i)$  i.i.d  $\mathcal{N}(0, 1)$



(parameters: 500 samples,  $n = 1000$ ,  $H = .8$ ,  $\sigma^2 = 1$ ,  $SNR = 0$  Db)

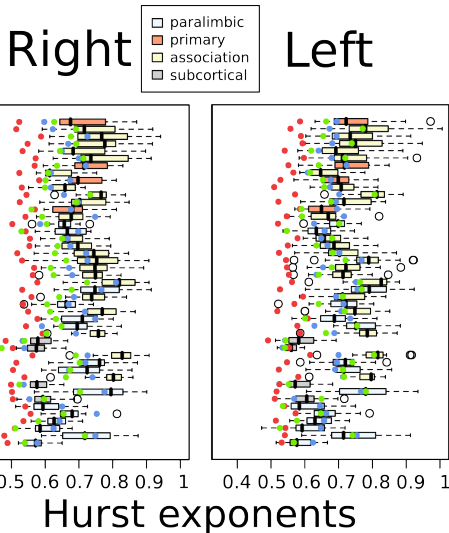
Method	Emp. Bias (Std Dev.)
No noise	0.801 (0.021)
With noise	0.543 (0.047)
<b>Diff. based procedure</b>	<b>0.807 (0.077)</b>

- Why?  $E[X^{a^m}(i)^2] = m^{2H} \gamma + \sigma_v^2 \sum_q (a_q)^2$
- Idea to correct:  $E[X^{a^{2m}}(i)^2] - E[X^{a^m}(i)^2] = m^{2H} \overbrace{(2^{2H} - 1)}{:=\gamma}$
- $\Rightarrow \log |S^2(2m) - S^2(m)| \simeq 2H \log(m) + \log \gamma'$
- [AC'10]: other type of contaminations (Brownian drift, outliers + noise, ...) can be undertaken using similar techniques; partial asymptotic results

Identification of fractional processes

# A first recent application to fMRI time series [AC et al'10]

- 20 healthy patients (boxplots)
- 3 coma patients (●, ●, ●).
- 90 regions of interest (left and right hemisphere)



# Some ongoing works and perspectives

- [BC<sup>10</sup>] Confidence interval for the Hurst parameter, based on concentration inequalities for subordinated Gaussian processes (alternative approach to asymptotic confidence intervals)
- Use of the Multivariate FBM for FMRI time series:
  - Joint estimation of the parameters  $H_i, \rho_{ij}, \dots$ , and asymptotic results [AC<sup>11</sup>].
  - Construction of connectivity graphs based on correlations, partial correlations, mutual information,...
- Develop models for FMRI time series with stimuli (cognitive or visual task,...)
- Exact simulation of multivariate Gaussian time series, multivariate multifractional processes, ...
- ...

# Part II. Inference for spatial Gibbs point processes

1. Definition of Gibbs models and examples
2. Estimation and validation of Gibbs models

## 3 Papers (Int. Journ.)

Elect. Journ. Stat.'08,'10

Scand. Journ. Stat.'10 (in rev.)

+ 1 Preprint subm.

## 1 Proceedings

NPAR'09



## Collaborators

(method. papers)

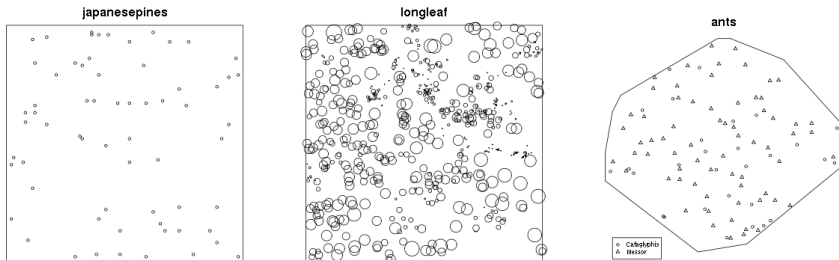
J.M. Billiot

D. Dereudre

R. Drouilhet

F. Lavancier

# Type of data and a few notation



- Data: configuration of marked points, i.e. set  $\varphi := \{x_1^{m_1}, \dots, x_n^{m_n}\}$ , where  $n := |\varphi|$  is not known in advance.
- $x_i^{m_i} = (x_i, m_i)$  is a marked point
- $x_i \in \mathbb{R}^d$ : location (tree, ant, ...),  $m_i \in \mathbb{M}$ : associated mark (diameter of a tree  $\mathbb{M} = [0, K]$ , ant species  $\mathbb{M} = \{1, 2, \dots\}$ ).
- $\varphi_\Lambda$  is the restriction of  $\varphi$  on  $\Lambda$ , that is  $\varphi \cap \Lambda$ .

# Stationary Gibbs models in $\mathbb{R}^d$ in one slide!

- Reference model = homogeneous Poisson point process with intensity  $\theta$ :  $\forall \Lambda, \Lambda' \in \mathbb{R}^d, \Lambda \cap \Lambda' = \emptyset$

$$|\Phi_\Lambda| \sim \mathcal{P}(\theta|\Lambda|), \quad |\Phi_\Lambda| \text{ and } |\Phi_{\Lambda'}| \text{ are independent.}$$

- Gibbs pp in  $\Lambda$** : admits a density with respect to a Ppp with unit rate of the form

$$f_\Lambda(\varphi; \theta) = Z_\Lambda(\theta)^{-1} e^{-V_\Lambda(\varphi; \theta)}$$

$$Z_\Lambda(\theta) = \sum_{k \geq 0} \frac{1}{k!} \int_\Lambda \dots \int_\Lambda e^{-V_\Lambda(\{x_1, \dots, x_k\}; \theta)} dx_1 \dots dx_k$$

intractable normalizing constant called partition function.

$$V_\Lambda(\varphi; \theta) = \text{measures the energy of a configuration of points } \varphi.$$



# Stationary Gibbs models in $\mathbb{R}^d$ in one slide!

- Reference model= homogeneous Poisson point process with intensity  $\theta$ :  $\forall \Lambda, \Lambda' \in \mathbb{R}^d, \Lambda \cap \Lambda' = \emptyset$

$$|\Phi_\Lambda| \sim \mathcal{P}(\theta|\Lambda|), \quad |\Phi_\Lambda| \text{ and } |\Phi_{\Lambda'}| \text{ are independent.}$$

- **Gibbs pp in  $\Lambda$** : admits a density with respect to a Ppp with unit rate of the form

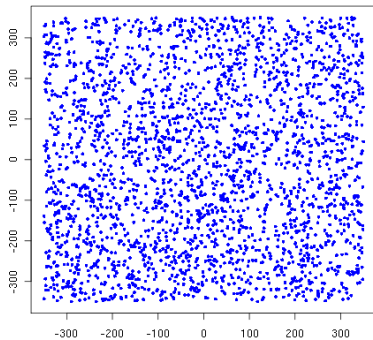
$$f_\Lambda(\varphi; \theta) = Z_\Lambda(\theta)^{-1} e^{-V_\Lambda(\varphi; \theta)}$$

- **Stationary Gibbs models in  $\mathbb{R}^d$** :
  - Defined via conditional specifications
  - **essential** for asymptotic properties; technical conditions on  $(V_\Lambda)_\Lambda$  for existence of stationary Gibbs measures: [R'69,BBD'94,D'08,DDG'10].
- definition of marked Gpp follows similar developments.

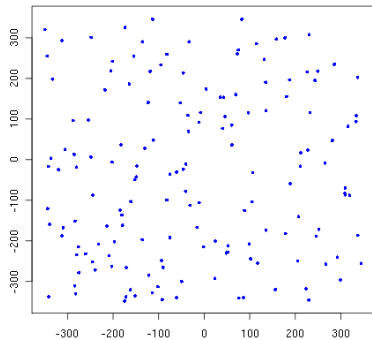
Poisson point process,  $\mathbb{M} = \{0\}$

$$V_{\Lambda}(\varphi; \theta) = \theta_1 |\varphi_{\Lambda}|$$

$\theta_1=5$



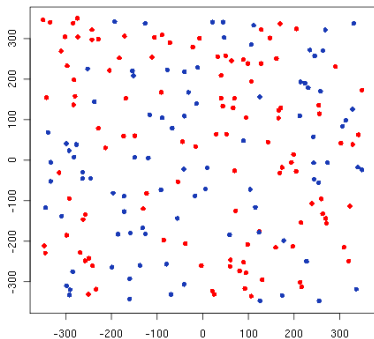
$\theta_1=8$



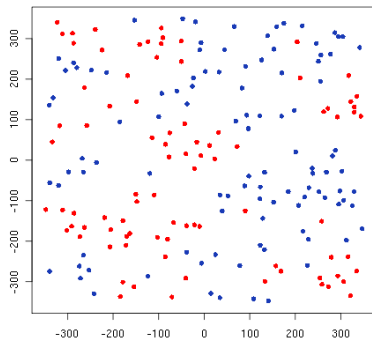
Strauss marked point process,  $\mathbb{M} = \{r, b\}$ 

$$V_{\Lambda}(\varphi; \theta) = \theta_{11} |\varphi_{\Lambda}^r| + \theta_{12} |\varphi_{\Lambda}^b| \\ + \theta_{21} \sum_{(x^r, y^r)} \mathbf{1}_{\|y^r - x^r\| < 10} + \theta_{22} \sum_{(x^b, y^b)} \mathbf{1}_{\|y^b - x^b\| < 10} + \theta_{23} \sum_{(x^b, y^r)} \mathbf{1}_{\|y^r - x^b\| < 10}$$

$$\theta_1 = (2, 2), \theta_2 = (2, 2, 2)$$



$$\theta_1 = (2, 2), \theta_2 = (6, 6, 10)$$



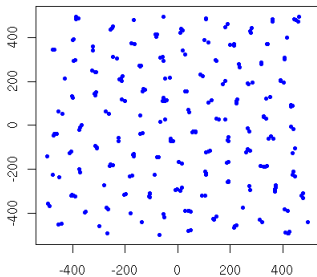
Multi-Strauss point process,  $\mathbb{M} = \{0\}$  on a planar structured graph

$$V_{\Lambda}(\varphi; \theta) = \theta_1 |\varphi_{\Lambda}| + \sum_{\{x,y\} \in G_2(\varphi)} \left\{ \theta_2 \mathbf{1}_{[\|y-x\| \leq 20]} + \theta_3 \mathbf{1}_{20 \leq \|y-x\| \leq 80} \right\}$$

$$\theta = (1, 2, 4)$$

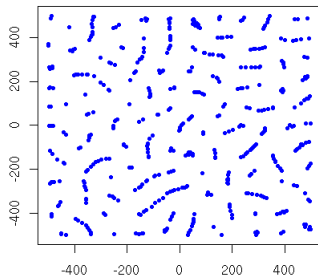
$$G_2(\varphi) = \mathcal{P}_2(\varphi)$$

number of points=266



$$G_2(\varphi) = Del_2(\varphi)$$

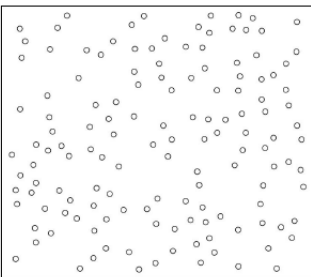
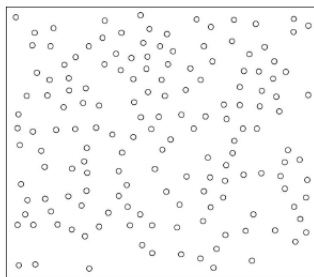
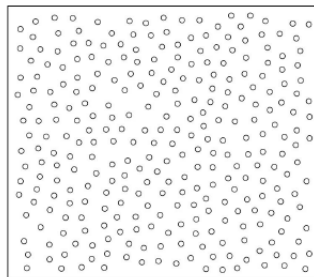
number of points=472



## Lennard-Jones model

$$V_{\Lambda}(\varphi; \theta) := \theta_1 |\varphi_{\Lambda}| + 4\theta_2 \sum_{\{x,y\} \in P_2(\varphi)} \left( \left( \frac{\theta_3}{\|y-x\|} \right)^{12} - \left( \frac{\theta_3}{\|y-x\|} \right)^6 \right)$$

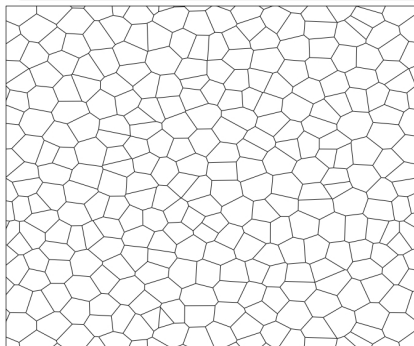
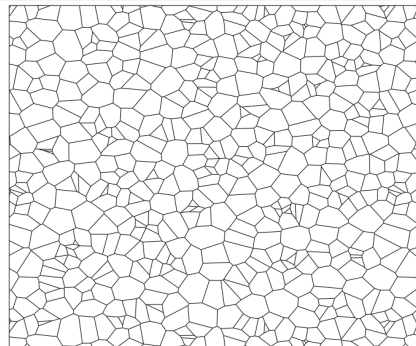
with  $\theta = (\theta_1, \theta_2, \theta_3) \in \mathbb{R} \times (\mathbb{R}^+)^2$ .


 $\theta_2 = 0$ 

 $\theta_2 = 0.1$ 

 $\theta_2 = 2$

## Gibbs Voronoi tessellation

$$V_{\Lambda}(\varphi) = \sum_{\substack{C \in \text{Vor}(\varphi) \\ C \cap \Lambda \neq \emptyset}} V_1(C) + \sum_{\substack{C, C' \in \text{Vor}(\varphi) \\ C \text{ and } C' \text{ are neighbors} \\ (C \cup C') \cap \Lambda \neq \emptyset}} V_2(C, C')$$

$V_1(C)$ : deals with the shape of the cell and  $V_2(C, C') = \theta d(\text{vol}(C), \text{vol}(C'))$ .


 $\theta > 0$ 

 $\theta < 0$

## Quermass model $\mathbb{M} = [0, \bar{R}]$

For finite configuration  $\varphi$

$$V(\varphi; \theta) = \theta_1 |\varphi| + \theta_2 \mathcal{P}(\Gamma) + \theta_3 \mathcal{A}(\Gamma) + \theta_4 \mathcal{E}(\Gamma) \quad \text{where } \Gamma = \bigcup_{(x,R) \in \varphi} \mathcal{B}(x, R)$$

where  $\mathcal{P}(\Gamma)$ ,  $\mathcal{A}(\Gamma)$  and  $\mathcal{E}(\Gamma)$  respectively denote the perimeter, the volume and the Euler-Poincaré characteristic of  $\Gamma$ .

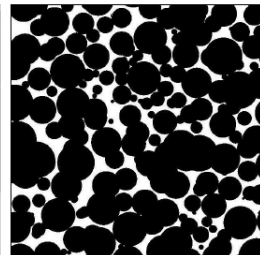
Simulation for a  $\mathcal{U}([0, 2])$  on the radius;  $\theta_1$  constant [MH'10]



$$(\theta_2, \theta_3, \theta_4) = (0, 0.2, 0)$$



$$(\theta_2, \theta_3, \theta_4) = (0, 0, 1)$$



# Examples with different characteristics

	Examples	Interaction on	Linearity of $V$ in $\theta$ ?	Local stability?	Hereditary?	Observed points?
1	(Multi-)Strauss Marked	$\mathcal{P}_2(\varphi)$	✓	✓	✓	✓
2	Structured Multi-Strauss	$Del_2(\varphi)$	✓	✓	✓	✓
3	Lennard-Jones	$\mathcal{P}_2(\varphi)$	✗	✗	✓	✓
4	Tesselation	Voronoi cells	✓	✓	✗	✓
5	Quermass	$f(\Gamma)$	✓	✓	✓	✗

- Local stability:  $V(x^m|\varphi; \theta) := V(\varphi \cup x^m) - V(\varphi) \geq -K$ .
- Hereditary:  $f(\varphi) > 0 \Rightarrow f(\psi) > 0 \quad \forall \varphi \subset \psi$ .
- Aim: asymptotic results encompassing all these examples.



# Maximum Likelihood estimation

	Examples	Interaction on	Linearity of $V$ in $\theta$ ?	Local stability?	Hereditary?	Observed points?
1	(Multi-)Strauss Marked	$\mathcal{P}_2(\varphi)$	✓	✓	✓	✓
2	Structured Multi-Strauss	$Del_2(\varphi)$	✓	✓	✓	✓
3	Lennard-Jones	$\mathcal{P}_2(\varphi)$	✗	✗	✓	✓
4	Tesselation	Voronoi cells	✓	✓	✗	✓
5	Quermass	$f(\Gamma)$	✓	✓	✓	✗

- Defined as:  $\text{Argmax}_{\theta} Z_{\Lambda_n}(\theta)^{-1} e^{-V(\varphi; \theta)}$
- theoretically unfeasible.
- practically (almost always) feasible with intensive MC simulations to approximate the partition function [M07].
- few theoretical results (general consistency?)

# Maximum Pseudo-Likelihood estimation

	Examples	Interaction on	Linearity of $V$ in $\theta$ ?	Local stability?	Hereditary?	Observed points?
1	(Multi-)Strauss Marked	$\mathcal{P}_2(\varphi)$	✓	✓	✓	✓
2	Structured Multi-Strauss	$Del_2(\varphi)$	✓	✓	✓	✓
3	Lennard-Jones	$\mathcal{P}_2(\varphi)$	✗	✗	✓	✓
4	Tessellation	Voronoi cells	✓	✓	✗	✓
5	Quermass	$f(\Gamma)$	✓	✓	✓	✗

- On the lattice [B<sup>+</sup>68]: consider the product of the conditional densities in each site conditionally on the other ones.
- Extended by [JM<sup>+</sup>91] for p.p.:

$$LPL_{\Lambda_n}(\varphi; \theta) = - \int_{\Lambda_n \times \mathbb{M}} e^{-V(x^m | \varphi; \theta)} dx^m - \sum_{x^m \in \varphi_{\Lambda_n}} V(x^m | \varphi \setminus x^m; \theta).$$

# Maximum Pseudo-Likelihood estimation

	Examples	Interaction on	Linearity of $V$ in $\theta$ ?	Local stability?	Hereditary?	Observed points?
1	(Multi-)Strauss Marked	$\mathcal{P}_2(\varphi)$	✓	✓	✓	✓
2	Structured Multi-Strauss	$Del_2(\varphi)$	✓	✓	✓	✓
3	Lennard-Jones	$\mathcal{P}_2(\varphi)$	✗	✗	✓	✓
4	Tessellation	Voronoi cells	✓	✓	✗	✓
5	Quermass	$f(\Gamma)$	✓	✓	✓	✗

- On the lattice [B'68]: consider the product of the conditional densities in each site conditionally on the other ones.
- Extended by [JM'91] for p.p.:

$$LPL_{\Lambda_n}(\varphi; \theta) = - \int_{\Lambda_n \times \mathbb{M}} e^{-V(x^m | \varphi; \theta)} dx^m - \sum_{x^m \in \varphi_{\Lambda_n}} V(x^m | \varphi \setminus x^m; \theta).$$

- ‘quick’ and computable estimator and asymptotic results (as  $\Lambda_n \rightarrow \mathbb{R}^d$ ) ...



are valid for a large class of models

[BCD'08]: 1,2, [CD'10] 3, [DL'09] 4

# Takacs-Fiksel method (1)

## Theorem (Georgii-Nguyen-Zessin)

For any  $h(\cdot, \cdot; \theta) : \mathbb{S} \times \Omega \rightarrow \mathbb{R}$ , for any  $\theta \in \Theta$ ,

$$E_{\theta^*} \left( \sum_{x^m \in \varphi} h(x^m, \varphi \setminus x^m; \theta) \right) = E_{\theta^*} \left( \int_{\mathbb{R}^d \times \mathbb{M}} h(x^m, \varphi; \theta) e^{-V(x^m | \varphi; \theta^*)} dx^m \right)$$

Define

$$I_{\Lambda}(\varphi; h, \theta) = \int_{\Lambda \times \mathbb{M}} h(x^m, \varphi; \theta) e^{-V(x^m | \varphi; \theta)} dx^m - \sum_{x^m \in \varphi_{\Lambda}} h(x^m, \varphi \setminus x^m; \theta).$$

Idea of TF method: ergodic theorem and GNZ formula  $\Rightarrow I_{\Lambda_n}(\varphi; \theta^*) \simeq 0$ .

Let us give  $K$  test functions  $h_k(\cdot, \cdot; \theta) : \mathbb{S} \times \Omega \rightarrow \mathbb{R}$  (for  $k = 1, \dots, K$ ).

$$\widehat{\theta}_{TF}(\varphi) := \arg \min_{\theta \in \Theta} \sum_{k=1}^K I_{\Lambda_n}(\varphi; h_k, \theta)^2,$$

# Takacs-Fiksel method (2)

$$\widehat{\theta}_{TF}(\varphi) := \arg \min_{\theta \in \Theta} \sum_{k=1}^K \left( \int_{\Lambda_n \times \mathbb{M}} h_k(x^m, \varphi; \theta) e^{-V(x^m | \varphi; \theta)} dx^m - \sum_{x^m \in \varphi_{\Lambda_n}} h_k(x^m, \varphi \setminus x^m; \theta) \right)^2$$

## Remarks and interest of the TF method:

- when  $\mathbf{h} = \mathbf{V}^{(1)}$ ,  $\widehat{\theta}_{TF}(\varphi) = \widehat{\theta}_{MPLE}(\varphi)$ .
- quick estimator: for example  $h_k(x^m, \varphi; \theta) := 1_{\mathcal{B}(0, r_k)}(\|x\|) e^{V(x^m | \varphi; \theta^*)}$ .
- [CDDL'10]: allows the identification of the Quermass model with a pertinent choice of test functions.
- [CDDL'10] Asymptotic results may be obtained (covering 1 → 5) but ...

# Takacs-Fiksel method (2)

$$\widehat{\theta}_{TF}(\varphi) := \arg \min_{\theta \in \Theta} \sum_{k=1}^K \left( \int_{\Lambda_n \times \mathbb{M}} h_k(x^m, \varphi; \theta) e^{-V(x^m | \varphi; \theta)} dx^m - \sum_{x^m \in \varphi_{\Lambda_n}} h_k(x^m, \varphi \setminus x^m; \theta) \right)^2$$

## Remarks and interest of the TF method:

- when  $\mathbf{h} = \mathbf{V}^{(1)}$ ,  $\widehat{\theta}_{TF}(\varphi) = \widehat{\theta}_{MPLE}(\varphi)$ .
- quick estimator: for example  $h_k(x^m, \varphi; \theta) := 1_{\mathcal{B}(0, r_k)}(\|x\|) e^{V(x^m | \varphi; \theta^*)}$ .
- [CDDL'10]: allows the identification of the Quermass model with a pertinent choice of test functions.
- [CDDL'10] Asymptotic results may be obtained (covering 1 → 5) but ...



• The contrast limit  
is not necessarily concave!



# Validation through residuals

Strategy proposed by [BTMH'05,BMP'08]:

- ① Let us give a parametric model with true parameter  $\theta^*$
- ② compute an estimate  $\widehat{\theta}$  of  $\theta^*$  (MLE,MPLE,TF,...).
- ③ Define the  $h$ -residuals,  $R_{\Lambda_n}(\varphi; h) = I_{\Lambda_n}(\varphi; h, \widehat{\theta})$  as a diagnostic tool

$$R_{\Lambda_n}(\varphi; h) = \int_{\Lambda \times \mathbb{M}} h(x^m, \varphi; \widehat{\theta}) e^{-V(x^m | \varphi; \widehat{\theta})} dx^m - \sum_{x^m \in \varphi_{\Lambda}} h(x^m, \varphi \setminus x^m; \widehat{\theta}).$$

Indeed, if the model is true (i.e.  $\Phi \sim P_{\theta^*}$ ), then one may expect that

$$R_{\Lambda_n}(\varphi; h) \simeq I_{\Lambda_n}(\varphi; h, \theta^*) \simeq 0 \text{ by the GNZ formula.}$$

# Validation through residuals

Strategy proposed by [BTMH'05,BMP'08]:

- ① Let us give a parametric model with true parameter  $\theta^*$
- ② compute an estimate  $\widehat{\theta}$  of  $\theta^*$  (MLE,MPLE,TF,...).
- ③ Define the  $h$ -residuals,  $R_{\Lambda_n}(\varphi; h) = I_{\Lambda_n}(\varphi; h, \widehat{\theta})$  as a diagnostic tool

$$R_{\Lambda_n}(\varphi; h) = \int_{\Lambda \times \mathbb{M}} h(x^m, \varphi; \widehat{\theta}) e^{-V(x^m | \varphi; \widehat{\theta})} dx^m - \sum_{x^m \in \varphi_{\Lambda}} h(x^m, \varphi \setminus x^m; \widehat{\theta}).$$

Indeed, if the model is true (i.e.  $\Phi \sim P_{\theta^*}$ ), then one may expect that

$$R_{\Lambda_n}(\varphi; h) \simeq I_{\Lambda_n}(\varphi; h, \theta^*) \simeq 0 \text{ by the GNZ formula.}$$

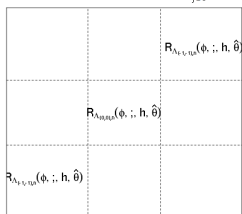
[CL'10]  $\Rightarrow$  investigate  $h$ -residuals for constructing several families of (asymptotic) gof tests



# Family of gof test based on quadrats

Fix some test function  $h$  (ex:  $h = 1$ ) and split  $\Lambda_n = \cup_{j \in \mathcal{J}} \Lambda_j$

$\hat{\theta}$  is based on  $\phi_{\Lambda_n}$  and  $\Lambda_n = \cup_{j \in \mathcal{J}} \Lambda_j$



- ① compute the  $h$ -residuals in each  $\Lambda_j$
- ② construct  $\|\mathbf{R}_1\|^2$  where

$$\mathbf{R}_1 := \left( R_{\Lambda_j}(\varphi, h, \widehat{\theta}_n) \right)_{j \in \mathcal{J}}.$$

- Based on a CLT for triangular arrays for spatial Markov random fields [\[CL10\]](#):

$$|\Lambda_{0,n}|^{-1} \widehat{\lambda}^{-1} \|\mathbf{R}_1(\Phi; h) - \overline{\mathbf{R}}_1(\Phi; h)\|^2 \xrightarrow{d} \chi_{|\mathcal{J}|-1}^2,$$

where  $\overline{\mathbf{R}}_1(\Phi; h) = \frac{1}{|\mathcal{J}|} \sum_j R_{\Lambda_j}(\varphi, h, \widehat{\theta}_n)$  and where  $\widehat{\lambda} = \widehat{Var}(I_\Delta(h, \Phi; \theta^*))$ .

- When  $h = 1 \Rightarrow$  generalization of the **dispersion test** for Poisson point processes.

## Further Work

- Simpler expressions for variance estimators (MPLE, Takacs-Fiksel) and variance of residuals.
- Simulation study to assess the efficiency and power of the proposed gof tests.
- Asymptotic results for MLE for general Gibbs points processes.
- Avoiding the finite range assumption in the CLT, results for non-stationary point processes.
- Nonparametric estimation of a pairwise interaction function (interaction on  $\mathcal{P}_2(\varphi)$ ,  $Del_2(\varphi), \dots$ )
- Current IXXI project (11-): spatio-temporal point process modelling of sunspots
- ...

# Summary

- Contributions
  - in modelling and inference for dependent data (I. 1D fractional processes, II. spatial point processes)
  - in applied statistics (III)
- **13** papers in international journals (published/accepted/in rev.):  
[**8** (I) + **3** (II) + **2** (III)]
- **3** Arxiv preprints.
- **22** (inter)national conferences/proceedings/techn. rep. (not submitted)
- **2** Phd co-supervisions.
- **7** grants/contracts (2 current: ANR InfoNetComaBrain / IXXI Project Sunspots).
- Development of procedures in R libraries (I, III): `SimEstFBM`, `dvfbm`, `asymptTest`

## Current IXXI project: bridge between Parts I and II?

- IXXI Project Sunspots (P.O. Amblard, N. Le Bihan, J. Lilenstein)
- Goal: statistical modelling of sunspots by a spatio-temporal point process on the sphere

Daily sunspots observations:

March 2001 → April 2001

<http://sohowww.nascom.nasa.gov/>

Thank you for your attention

# Characteristics of the data

- cyclostationary LRD time series of the sunspot number.
- Complex spatial structure and complex evolution with time.
- Spherical data (Geodesic Gaussian distribution on the circle, sphere [CLB'11])

