



**HAL**  
open science

## **A model for dissipation: cascade SDE with Markov regime-switching and Dirichlet prior**

Didier Bernard, Richard Emilion, Srikanth K. Iyer, Adaté Tossa

### ► **To cite this version:**

Didier Bernard, Richard Emilion, Srikanth K. Iyer, Adaté Tossa. A model for dissipation: cascade SDE with Markov regime-switching and Dirichlet prior. 2008. hal-00286131

**HAL Id: hal-00286131**

**<https://hal.science/hal-00286131>**

Preprint submitted on 17 Jun 2008

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A model for dissipation: cascade SDE with Markov regime-switching and Dirichlet prior

Didier BERNARD <sup>a</sup>, Richard EMILION <sup>b</sup>, Srikanth K. IYER <sup>c</sup>, Adaté TOSSA <sup>d</sup>

<sup>a</sup>*LPAT, Université Antilles-Guyane, B.P. 250 - 97157 Pointe-à-Pitre, Guadeloupe*

<sup>b</sup>*MAPMO, Université d'Orléans, B.P 6759 - 45067 Orléans Cedex 2, France*

<sup>c</sup>*Department of Mathematics, Indian Institute of Science, Bangalore 560 012 India*

<sup>d</sup>*CEREMADE, Université Paris Dauphine, 75016 Paris, France*

---

## Abstract

Cascade Stochastic Differential Equation (SDE), a continuous time model for energy dissipation in turbulence, is a generalization of the Yaglom discrete cascade model. We extend this SDE to a model in random environment by assuming that its two parameters are switched by a continuous time Markov chain whose states represent the states of the environment. Moreover, a Dirichlet process is placed as a prior on the space of sample paths of this chain. We propose a Bayesian estimation method of this model which is tested both on simulated data and on real data of wind speed measured at the entrance of the mangrove ecosystem in Guadeloupe.

*Key words:* Cascade model; Dirichlet process; dissipation; Mangrove; Markov regime switching; random environment; Stochastic Differential equation.

## 1. Introduction

In atmospheric turbulent flows, whirlwinds of various sizes can be observed. These whirlwinds whose sizes range from a largest scale  $L$  to a smallest scale  $l_0$  determine energy transfer amplitudes. Micrometeorological measures confirm important variations of the energy *dissipation*, putting in evidence an *intermittency* phenomenon (see Fig. 1 below).

Turbulent flows are characterized by their high Reynolds number  $R = vL/\nu$  where  $v$  is the fluid mean velocity and  $\nu$  the fluid viscosity. At small scales  $l_0 = \nu/v$  the flow is laminar (that is stable with smooth dependence on the flow data). At this length scale, *dissipation* effects dominate and energy is rapidly converted into heat. At length scales comparable to  $L$ , the flow is dominated by system dependent features, e.g., large vortices in the flow pattern. Thus neither the small nor the large length scales should be considered to be stochastic. For length scales in the range  $l_0 \leq l \leq L$ , called the *inertial range*, Kolmogorov hypothesis states that turbulence is independent on the Reynolds number  $R$  and is isotropic. In this range, energy is transferred from larger vortices to smaller ones, a process referred to as *cascades*. This is the range of length scales to which stochastic modelling applies. Kolmogorov and Obukhov established that

$$\langle \Delta_l v^2 \rangle = \langle |V(x+l) - V(x)|^2 \rangle \approx \langle \varepsilon^{3/2} \rangle l^{3/2}$$

where the *dissipation parameter*  $\varepsilon$  is assumed to be a random variable with a log-normal distribution ([15], [17]),  $\langle . \rangle$  denoting expectation. Taking into account experimental studies which showed that the dissipation field has a long-range power law correlations ([9], [19]), Yaglom [26] proposed a discrete multiplicative random cascade model with long-range correlations and small scale log-normal statistics for  $\varepsilon$ . Yaglom cascades are described using a formalism involving infinitely divisible probability measures. This has leaded F. Schmitt and D. Marsan [20] to prove that the limit of discrete cascade model defines a *stochastic dissipation process*  $\varepsilon$ , with a continuum of scale levels, which verifies a stochastic differential equation (SDE), called here *cascade SDE*, which depends on an *intermittency* parameter  $\mu$  and a *scale* parameter  $\lambda$ . Actually the singularity process  $\gamma_\lambda(t) = \log(\varepsilon_\lambda(t))$  satisfies the SDE

$$\gamma_\lambda(t) = -\frac{\mu}{2} \ln \lambda + \mu^{1/2} \int_{t+1-\lambda}^t (t+1-u)^{-1/2} dB(u). \quad (1)$$

---

*Email addresses:* [didier.bernard@univ-ag.fr](mailto:didier.bernard@univ-ag.fr) (Didier BERNARD), [richard.emilion@univ-orleans.fr](mailto:richard.emilion@univ-orleans.fr) (Richard EMILION), [skiyer@math.iisc.ernet.in](mailto:skiyer@math.iisc.ernet.in) (Srikanth K. IYER), [adate.tossa@ceremade.dauphine.fr](mailto:adate.tossa@ceremade.dauphine.fr) (Adaté TOSSA).

The aim of this paper is to propose an extension of this cascade SDE to a model in random environment by defining a dissipation process subject to *regime* changes. More precisely we assume that the parameters  $\mu$  and  $\lambda$  are driven by a continuous time Markov chain  $(X_t)_{t \geq 0}$  that takes values in a set  $S = \{1, 2, \dots, M\}$  representing the environment regimes. Further, the new idea of this paper, which at our knowledge has never been considered before in turbulence domain, is to make the model more flexible by putting on the space of sample paths of the Markov chain, a Dirichlet prior  $\mathcal{D}(\alpha H)$  centered around a basic probability measure  $H$ . The rationale for using such a model with regime changes came from observing the variations of  $\mu$  and  $\lambda$  estimated on sliding windows of wind speed data measured at the entrance of the mangrove ecosystem in Guadeloupe island [1], where the intermittency phenomenon allows an equilibrium between the wind turbulent flow and the vegetal cover (see [4], [7], and [25]). Our model, which is a complex *mixture* of cascade SDEs, better encompasses such variations.

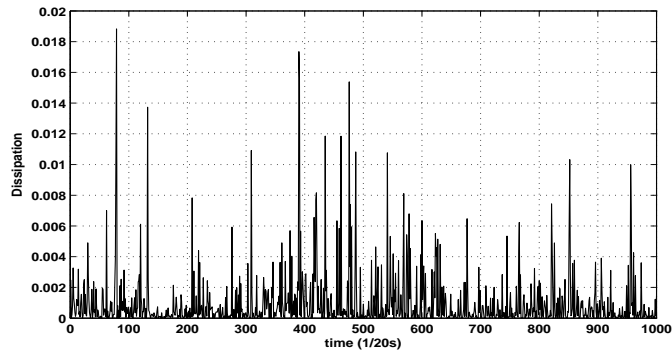


Fig. 1. Intermittency of the dissipation  $\varepsilon(t)$  over the Guadeloupe mangrove ecosystem

The paper is organized as follows. Section 2 reviews the cascade models: the discrete multiplicative cascade model and the continuous multiplicative cascades one. This last one, in the log-normal case, leads to the cascade SDE upon which hinges our model. Section 3 introduces the Cascade SDE with Markov switching regimes. After reviewing the Dirichlet process and its stick-breaking representation in section 4, section 5 is concerned with the complete model specification, its Bayesian estimation procedure through computations of posterior distributions as done in Ishwaran and Zarepour [10], Ishwaran and James [11]. Section 6 presents numerical results on both simulated and real data and we conclude in the last section by giving some perspectives on further research.

## 2. Cascade SDE

### 2.1. Discrete multiplicative Cascades

Random multiplicative cascade models have been introduced by Novikov and Stewart [16] and by Yaglom [26] in 1966 as a simple way to describe stochastic transfer of energy along the inertial range. Yaglom model can be seen as a weighed random tree depending on a family of positive random variables (r.v.) ( $W_{p,i}$ ). There are  $n$  levels on the path from the root to any leaf node and each internal node of the tree has  $\lambda_1 > 1$  children nodes. The root node is associated with  $W_{1,1}$  and with the external scale number  $L = l_0 \lambda_1^n$  where  $l_0$  is the dissipation scale number. Moreover to the edge  $i$ , ( $i = 1, \dots, \lambda_1^p$ ) in layer  $p$  is associated with a random variable  $W_{p,i}$ , and a scale number  $L/\lambda_1^p = l_0 \lambda_1^{n-p}$ . Besides, for two different levels  $p$  and  $p'$ , the r.v.  $W_{p,\cdot}$  and  $W_{p',\cdot}$  are independent and identically distributed (i.i.d.) with a common distribution  $W$  that satisfies the conservative property  $\langle W \rangle = 1$ . Hence each node is associated with the product of the  $W_{p,i}$ 's which are on the path from the root to this node. In particular, a leaf node  $x$  is associated with a r.v.

$$\varepsilon(x) = \prod_{p=1}^n W_{p,x(p)} \quad (2)$$

where  $x(p)$  is the index of the level  $p$  edge on the path from the root to node  $x$ .

Similar constructions are done for random Polya trees [18] or in Kraft construction of a continuous random distribution [14]. Since, the  $W_{p,x(p)}$  are assumed independent by construction, the moments of order  $q > 0$  of  $\varepsilon$  are given by:

$$\begin{aligned} \langle (\varepsilon(x))^q \rangle &= \left\langle \left( \prod_{i=1}^n W_{p,x(p)} \right)^q \right\rangle \\ &= \prod_{i=1}^n \langle (W_{p,x(p)})^q \rangle \\ &= \langle W^q \rangle^n . \end{aligned}$$

Moreover, the correlation properties of  $\varepsilon$  are derived in Schmitt [20] and we reproduce them here for completeness:

$$\begin{aligned}
\langle \varepsilon(x)\varepsilon(x+r) \rangle &= \langle \prod_{p=1}^n W_{p,x(p)} \prod_{p'=1}^n W_{p',(x+r)(p')} \rangle \\
&= \prod_{p=1}^{n-m} \langle W_{p,x(p)}^2 \rangle \prod_{p=n-m+1}^n \langle W_{p,x(p)} \rangle \prod_{p'=n-p+1}^n \langle W_{p',(x+r)(p')} \rangle \\
&= \langle W^2 \rangle^{n-m} \langle W \rangle^{2m}
\end{aligned}$$

where  $m$  is such that  $\lambda_1^m \approx r$ . Introducing the total scale ratio  $\lambda = L/l_0 = \lambda_1^n$ , we get

$$\langle \varepsilon_\lambda^q \rangle = \lambda^{K(q)} \quad (3)$$

$$\langle \varepsilon_\lambda(x)\varepsilon_\lambda(x+r) \rangle \approx \lambda^{K(2)} r^{-K(2)} \quad (4)$$

where  $K(q) = \log_{\lambda_1} \langle W^q \rangle$  and  $K(2) = \mu$ .

## 2.2. Continuous multiplicative Cascades

### 2.2.1. Definition

Continuous cascades are obtained by letting the depth  $n \rightarrow +\infty$  while the total scale ratio  $\lambda = L/l_0 = \lambda_1^n$  is kept fixed (but large) so that  $\lambda_1 = \lambda^{1/n} \rightarrow 1^+$  (see [21], [23], [24]). In the previous section, cascades have been presented in space. When considering cascades developing in time,  $\varepsilon_\lambda$  may be seen as a stochastic process and equation (1) writes

$$\langle \varepsilon_\lambda(t)\varepsilon_\lambda(t+\tau) \rangle \approx \lambda^{K(2)} \tau^{-K(2)},$$

showing that  $(\varepsilon_\lambda(t))_t$  has long-range correlations for time increment  $\tau$  with parameter  $\mu = K(2)$  for fixed  $\lambda$ .

### 2.2.2. The stochastic equation

As shown by F.G. Schmitt [20], if all the  $W$ 's are log-normal, that is

$$W = \exp(\sqrt{\mu \ln 2} g_0 - \frac{\mu}{2} \ln 2),$$

where  $g_0$  is a standard Gaussian, we obtain a dissipation stochastic process:

$$\varepsilon_\lambda(t) = \lambda^{-\frac{\mu}{2}} \exp\left(\mu^{\frac{1}{2}} \int_{t+1-\lambda}^t (t+1-u)^{-1/2} dB(u)\right), \quad (5)$$

and its generator, called *singularity* process:

$$\gamma_\lambda(t) = \ln \varepsilon_\lambda(t) = -\frac{\mu}{2} \ln \lambda + \mu^{1/2} \int_{t+1-\lambda}^t (t+1-u)^{-1/2} dB(u). \quad (6)$$

This last process  $\gamma_\lambda(t)$  has a stochastic drift:

$$d\gamma_\lambda(t) = -\frac{\mu^{1/2}}{2} \left( \int_{t+1-\lambda}^t (t+1-u)^{-3/2} dB(u) \right) dt + \mu^{1/2} dB(t) - \lambda^{-1/2} dB(t+1-\lambda).$$

For large scale ratios  $\lambda \gg 1$  the term  $\lambda^{-1/2} dB(t+1-\lambda)$  is negligible, so

$$d\gamma(t) \approx \mu^{1/2} \left( dB(t) - \frac{1}{2} \int_{t+1-\lambda}^t (t+1-u)^{-3/2} dB(u) dt \right). \quad (7)$$

### 3. Cascade SDE with Markov regime-switching

In order to take in account environment changes, we assume that the above parameters  $\lambda$  and  $\mu$  are themselves stochastic processes, namely governed by a continuous-time Markov chain. This idea is well-known in mathematical finance when modelling regime switching markets with stochastic volatility (see e.g. [3], [8]).

Consider the following cascade stochastic equation

$$\gamma(t) = -\frac{\mu_{X(t)}}{2} \log(\lambda_{X(t)}) + \mu_{X(t)}^{1/2} \int_{t+1-\lambda_{X(t)}}^t (t+1-u)^{-1/2} dB(u), \quad (8)$$

where  $(B_t)$  is a standard Brownian motion,  $X = (X(t))_{t \geq 0}$  is a continuous time Markov chain taking values in a finite set  $S = \{1, 2, \dots, M\}$  that represents different *regimes* of the environment.

$\mu = (\mu_1, \dots, \mu_M)$  and  $\lambda = (\lambda_1, \dots, \lambda_M)$  are random vectors such that:

$$\mu_i \stackrel{i.i.d.}{\sim} \Gamma_1 \quad \text{for } i = 1, \dots, M,$$

$$\lambda_i \stackrel{i.i.d.}{\sim} \Gamma_2 \quad \text{for } i = 1, \dots, M,$$

$\Gamma_1$  and  $\Gamma_2$  being arbitrarily fixed distributions.

The meaning of this setting is that during regime  $i$  the intermittency parameter and the scale parameter are constant and equal to  $\mu_i$  and  $\lambda_i$  respectively. If, at a given time  $t$ , the state of the chain is  $i$ , that is

if  $X(t) = i$ , ( $i \in S$ ), then the environment is in regime  $i$  so that  $\mu_{X(t)} = \mu_i$  and  $\lambda_{X(t)} = \lambda_i$ .

Observe that the distribution of  $\gamma$  appears as a mixture of distributions of cascade SDE. The chain spends an exponential amount of time with parameter  $\beta_i$  in state  $i \in S$ , and then jumps to state  $j \in S, j \neq i$  with probability  $p_{ij}$ . The transition rate matrix  $Q = (q_{ij})_{i,j \in S}$  is such that off-diagonal entries  $q_{ij}$  are nonnegative reals values, and the diagonal element  $q_{ii}$  is constrained to be  $q_{ii} = -\sum_{j \neq i} q_{ij}$ , so that the raw sums of rows of  $Q$  are zero. We take  $\beta_i = -q_{ii} > 0$  for all  $i \in S$ , and  $p_{ij} = -q_{ij}/q_{ii}$  for  $i \neq j$ .

#### 4. Dirichlet process and Stick-breaking

As introduced in (Emilion, Faires, Iyer, 2008) [5], we complete the above model by placing a Dirichlet process (DP) as a prior on the path space of the Markov chain  $X$ . This will lead us to a Bayesian estimation of the model. In this section we review some properties of the DP that will be needed. Let  $(\Omega, \mathcal{B}, \mathcal{P})$  be a fixed probability space and let  $\mathcal{V}$  be a Polish space, that is a complete separable metric space. Let  $\mathbf{P}(\mathcal{V})$  denote the set of all probability measures defined on  $\mathcal{V}$ . It is well-know that if  $\mathcal{V}$  is a Polish space, so is  $\mathbf{P}(\mathcal{V})$ .

**Definition** : A random distribution (RD) on  $\mathcal{V}$  is a measurable map from  $\Omega$  to  $\mathbf{P}(\mathcal{V})$ .

Note that if  $X : \Omega \longrightarrow \mathbf{P}(\mathcal{V})$  is a RD then its distribution  $\mathcal{P}_X$  is a probability measure on  $\mathbf{P}(\mathcal{V})$ .

##### 4.1. Finite dimensional Dirichlet distribution

If  $\mathcal{V}$  is a finite set with cardinality  $l$  then  $\mathbf{P}(\mathcal{V})$  can be identified to the simplex

$$\Delta_l = \{y = (y_1, \dots, y_l), y_j \geq 0, \sum_{j=1}^l y_j = 1\}$$

and any RD to a random vector

$$X = (X_1, \dots, X_l) : (\Omega, \mathcal{P}) \rightarrow \mathbb{R}_+^l, \quad \sum_{k=1}^l X_k = 1.$$

**Definition** : Let  $\alpha = (\alpha_1, \dots, \alpha_l)$ , with  $\alpha_1 \geq 0, \dots, \alpha_l \geq 0$ , and let  $Z_1, \dots, Z_l$  be  $l$  independent real random variables with gamma distributions  $\Gamma(\alpha_1, 1), \dots, \Gamma(\alpha_l, 1)$  respectively. The Dirichlet distribution with pa-



parameter  $\alpha$ , denoted  $D(\alpha_1, \dots, \alpha_l)$  or  $D_\alpha$ , is defined as the distribution of the random vector  $(\frac{Z_1}{Z}, \dots, \frac{Z_l}{Z})$  where  $Z = Z_1 + \dots + Z_l$ .

Recall that the density of a Gamma random variable is

$$\Gamma(a, b)(x) = \frac{1}{\Gamma(a)} b^a e^{-bx} x^{a-1} I_{(x>0)}.$$

Let

$$\nabla_l = \{p = (p_1, \dots, p_{l-1}), p_j \geq 0, \sum_{j=1}^{l-1} p_j \leq 1\}.$$

It can be seen that if  $X = (X_1, \dots, X_l) \sim D(\alpha_1, \dots, \alpha_l)$  then  $X = (X_1, \dots, X_{l-1})$  has a probability density function

$$\pi(p_1, \dots, p_{l-1}) = \frac{\Gamma\left(\sum_{i=1}^l \alpha_i\right)}{\prod_{i=1}^l \Gamma(\alpha_i)} \prod_{i=1}^{l-1} p_i^{\alpha_i-1} \left(1 - \sum_{j=1}^{l-1} p_j\right)^{\alpha_l-1} \mathbb{I}_{p \in \nabla_l}$$

with respect to the Lebesgue measure in  $\mathbb{R}^{l-1}$ .

**Definition** (Ferguson 1973): Let  $H$  be a probability measure on  $\mathcal{V}$  and  $\alpha$  a positive number. A random distribution  $X : \Omega \rightarrow \mathbf{P}(\mathcal{V})$  is a Dirichlet process with scaling parameter  $\alpha$  and base distribution  $H$ , shortly  $X \sim \mathcal{D}(\alpha H)$ , if for any integer  $k \geq 2$  and any measurable partition  $B_1, \dots, B_k$  of  $\mathcal{V}$ , the joint distribution of the random vector  $(X(B_1), \dots, X(B_k))$  is a Dirichlet distribution with parameters  $(\alpha H(B_1), \dots, \alpha H(B_k))$ .

Shortly,

$$(X(B_1), \dots, X(B_k)) \sim D(\alpha H(B_1), \dots, \alpha H(B_k)) \quad (9)$$

It can be shown that  $\mathcal{D}(\alpha H)$  is centered around  $H$ , that is, for any measurable set  $B \in \mathcal{B}$ ,

$$E(X(B)) = H(B)$$

and

$$Var(X(B)) = \frac{H(B)(1-H(B))}{1+\alpha}.$$

Probability measures drawn from a Dirichlet process are discrete with probability one (Ferguson 1973).

This property is made explicit by the following stick-breaking construction due to Sethuraman (1994).

#### 4.2. Stick-breaking representation of Dirichlet process

A constructive approach to the Dirichlet process is the so called stick-breaking scheme. Let  $H$  be a probability distribution on  $\mathcal{V}$  and let  $N \geq 2$  be an integer. Let  $V_k \stackrel{i.i.d.}{\sim} \text{Beta}(a_k, b_k)$   $k = 1 \dots N - 1$  with shape parameters  $a_k, b_k > 0$  and let  $V_N = 1$ . Let  $p_1 \dots p_N$  be defined by

$$p_1 = V_1, \quad p_k = (1 - V_1)(1 - V_2) \dots (1 - V_{k-1})V_k, \quad k \geq 2. \quad (10)$$

Note that the  $p_k$ 's sum to 1. Let  $Z_k \stackrel{i.i.d.}{\sim} H$  ( $k = 1, \dots, N$ ) be independent of  $(p_k)_{k=1 \dots N}$ . The random measure  $\mathcal{P}_N$  defined by

$$\mathcal{P}_N = \sum_{k=1}^N p_k \delta_{Z_k}(\cdot) \quad (11)$$

is said to be a *stick-breaking* construction. The Ferguson Dirichlet process  $\mathcal{D}(\alpha H)$  is the best known example of an infinite stick-breaking prior. This is made explicit by the following proposition due to Sethuraman (1994):

**Proposition 4.1**

$$\mathcal{P}_N(\cdot) = \sum_{k=1}^N p_k \delta_{Z_k}(\cdot) \xrightarrow{a.s.} \mathcal{D}(\alpha H).$$

This proposition yields an efficient approximation of a Dirichlet process that is very useful in Bayesian nonparametrics statistics. The following lemma will be crucial in our estimation procedure.

**Lemma 1** *Let  $X : \Omega \rightarrow \{1, \dots, N\}$  be a r.v. with conditional distribution  $\Pr(X \in \cdot | P) = \sum_{k=1}^N p_k \delta_k(\cdot)$ , where  $P$  is defined by the stick-breaking construction (10). Then the conditional distribution of  $P$  given  $X$  is also defined by (10), where now, the  $V_k$ 's are independent  $\text{Beta}(a^*, b^*)$  r.v.'s with*

$$a^* = a_k + I_{\{X=k\}} \quad (12)$$

$$b^* = b + \sum_{j=k+1}^N I_{\{X=j\}}. \quad (13)$$

### 5. Complete model specification and estimation

In addition to the conditions precised in section 3 after equation (8)

$$\gamma(t) = -\frac{\mu_{X(t)}}{2} \log(\lambda_{X(t)}) + \mu_{X(t)}^{1/2} \int_{t+1-\lambda_{X(t)}}^t (t+1-u)^{-1/2} dB(u),$$

we place a Dirichlet prior on the path space  $D((0, \infty), S)$  of the process  $X$ , that is

$$\left\{ \begin{array}{l} (X|P, \alpha) \sim P \\ (P|\alpha) \sim \mathcal{D}(\alpha H) \\ (\alpha) \sim \text{Gamma}(\eta_1, \eta_2) \end{array} \right. \quad (14)$$

where  $H$  is the distribution of a specific Markov chain determined by a transition rate matrix  $Q$  and an initial distribution  $\pi_0$ . Here  $\pi_0$  will be the uniform distribution  $\pi_0 = (1/M, \dots, 1/M)$  and  $Q = (q_{ij})$  with  $q_{ij} = \beta_i p_{ij}$   $i \neq j$  where  $\beta_i = \beta > 0$ ,  $p_{ij} = 1/(M-1)$  and  $q_{ii} = -\sum_{j \neq i} q_{ij}$ .

In practice, process (8) is sampled at time  $t_1, \dots, t_n$ , providing a finite random vector  $\gamma = (\gamma_1, \dots, \gamma_n)$  of observations, where  $\gamma_i$  is the value recorded at time  $t_i$ . It will be proved in the next section (Prop. 5.1 below) that the conditional distribution of the vector  $\gamma$  is Gaussian so that we arrive at the following hierarchical model that we want estimate from the observed data:

$$\left\{ \begin{array}{l} (\gamma|X, \mu, \lambda, P) \sim N_n(m(X), \Sigma) \\ \mu_1, \dots, \mu_M \stackrel{iid}{\sim} \Gamma_1 \\ \lambda_1, \dots, \lambda_M \stackrel{iid}{\sim} \Gamma_2 \\ (X|P, \alpha) \sim P \\ (P|\alpha) \sim \mathcal{D}(\alpha H) \\ (\alpha) \sim \text{Gamma}(\eta_1, \eta_2) \end{array} \right. \quad (15)$$

### 5.1. Conditional distributions

The estimation procedure of the various parameters of the model hinges on Gibbs sampling scheme which implementation requires the computation of the following conditional distributions:

$$(\gamma|X, \mu, \lambda, p)$$

$$(\mu|\gamma, \lambda, X)$$

$$(\lambda|\gamma, \mu, X)$$

$$(X|p, \alpha)$$

$$(p|\alpha).$$

For simplicity,  $\gamma(t_i)$  and  $X(t_i)$  will be denoted below by  $\gamma_i$  and  $X(i)$ , respectively.

### 5.1.1. Conditional for $\gamma$

#### Proposition 5.1

$$(\gamma|X, \mu, \lambda, p) \sim \mathcal{N}_n(m, \Sigma)$$

with

$$m = \left( -\frac{1}{2}\mu_{X(1)} \log(\lambda_{X(1)}), \dots, \frac{1}{2}\mu_{X(n)} \log(\lambda_{X(n)}) \right) \quad \text{and} \quad \Sigma = (\sigma_{st})_{s,t=1,\dots,n}$$

where

$$\sigma_{st} = (\mu_{X(s)}\mu_{X(t)})^{1/2} COV\left( \int_{s+1-\lambda_{X(s)}}^s (s+1-u)^{-1/2} dB(u), \int_{t+1-\lambda_{X(t)}}^t (t+1-u)^{-1/2} dB(u) \right).$$

The coefficients  $\sigma_{st}$  can be given explicitly using a computation of F. G. Schmitt ([20] pages 89-90).

#### Proposition 5.2

Let  $s$  and  $t$  be integers, let  $w = \min(s, t)$ ,  $a = \max(s+1-\lambda_{X(s)}, t+1-\lambda_{X(t)})$  and  $\tau = |(s-t)|$ . Then

$$\sigma_{st} = 2(\mu_{X(s)}\mu_{X(t)})^{1/2} \log \left( \frac{\sqrt{w+1-a} + \sqrt{w+1-a+\tau}}{1 + \sqrt{1+\tau}} \right).$$

If  $j$  is a state reached by the Markov chain, let  $t_{1j}, \dots, t_{n_jj}$  be the times at which  $j$  is reached and let

$$\gamma_j = (\gamma_{t_{1j}}, \dots, \gamma_{t_{n_jj}}) \tag{16}$$

**Corollary 2**

$$\gamma_j | \mu, \lambda, X \sim \mathcal{N}_{n_j} \left( -\frac{\mu_j}{2} \log(\lambda_j) \underbrace{(1, 1, \dots, 1)}_{n_j \text{ times}}, \mu_j \Sigma_j \right),$$

where  $\Sigma_j = (\sigma_j(k\ell))$ , is a  $n_j \times n_j$  matrix with

$$\sigma_j(k\ell) = \mu_j \text{COV} \left( \int_{t_k+1-\lambda_j}^{t_k} (t_k + 1 - u)^{-1/2} dB(u), \int_{t_\ell+1-\lambda_j}^{t_\ell} (t_\ell + 1 - v)^{-1/2} dB(v) \right).$$

In the following sections, the probability densities of  $\mathcal{N}_n$  and  $\mathcal{N}_{n_j}$  will be denoted by  $f_n$  and  $f_{n_j}$  respectively, that is

$$f_n(\gamma, \mu, \lambda) = \frac{\exp \left( -\frac{1}{2} (\gamma - m) \Sigma^{-1} (\gamma - m)^T \right)}{(2\pi)^{n/2} (\det(\mu \Sigma))^{1/2}}, \quad (17)$$

$$f_{n_j}(\gamma_j, \mu_j, \lambda_j) = \frac{\exp \left( -\frac{1}{2} (\gamma_j - m_j) (\mu_j \Sigma_j)^{-1} (\gamma_j - m_j)^T \right)}{(2\pi)^{n_j/2} (\det(\mu_j \Sigma_j))^{1/2}}.$$

where  $m_j = -\frac{\mu_j}{2} \log(\lambda_j) \underbrace{(1, 1, \dots, 1)}_{n_j \text{ times}}$

5.1.2. *Conditional for X.*

**Proposition 5.3** *Let N be a positive integer.*

$$\left( X | p = \sum_{i=1}^N p_i \delta_{X_i}, \gamma, \mu, \lambda \right) \propto \sum_{i=1}^N p_i^* \delta_{X_i}$$

where  $p_i^* = p_i f_n(\gamma, \mu, \lambda)$ ,  $f_n$  being defined in (17).

5.1.3. *Conditional for p*

**Proposition 5.4**

$$(p_1 | X) \sim V_1^*, \quad (p_k | X) \sim V_k^* \prod_{i=1}^{k-1} (1 - V_i^*) \quad k = 2, 3, \dots, N-1 \quad (18)$$

where  $V_k^* \stackrel{\text{ind}}{\sim} \text{Beta}(a_k^*, b_k^*)$  for  $k = 1, \dots, N$  and  $a_k^*, b_k^*$  are given by Lemma 1.

**Proof**

By proposition (5.3),  $(X | p) \propto \sum_{i=1}^N p_i^* \delta_{X_i}$ . The result then follows from Lemma 1.

**Remark 3** *The fact stated in this proposition is that the current value of parameter  $a_k$  increases by 1 each time the path  $k$  is chosen. What this does is that as we repeat the process of generating from the*

conditional within each iteration many times, the Beta distributions will gradually concentrate on paths that best explain the data.

#### 5.1.4. Conditional for $\alpha$

##### Proposition 5.5

$$(\alpha|p) \sim \text{Gamma} \left( N + \eta_1 - 1, \eta_2 - \sum_{i=1}^{N-1} \log(1 - V_i^*) \right),$$

where the  $V_i^*$  are like those obtained in the conditional of  $p$ .

#### 5.1.5. Conditional for $\mu$

Let  $j \in S$  be a state of  $X$  and  $\pi_1$  be the prior for  $\mu_j$ .

**Proposition 5.6** *If for all  $t \in \{1, \dots, n\}$   $X(t) \neq j$  then*

$$(\mu_j|\gamma, \lambda_j, X, \alpha) \sim \pi_1$$

otherwise

$$(\mu_j|\gamma, \lambda_j, X, \alpha) \sim \Gamma_1 \propto \mathcal{N}_{n_j}(m_j, \mu_j \Sigma_j) \cdot \pi_1$$

where  $m_j = -\frac{\mu_j}{2} \log(\lambda_j) \underbrace{(1, 1, \dots, 1)}_{n_j \text{ times}}$ .

#### 5.1.6. Conditional for $\lambda$

Let  $j \in S$  be a state of  $X$  and let  $\pi_2$  be the prior for  $\lambda_j$ . Similarly to the case of  $\mu$ , if for all  $t \in \{1, \dots, n\}$   $X(t) \neq j$  then

$$(\lambda_j|\gamma, \mu_j, X, \alpha) \sim \pi_2$$

otherwise, for each of the other values of  $j \in S$ ,

$$(\lambda_j|\gamma, \mu_j, X, \alpha) \sim \Gamma_2 \propto (\gamma|\lambda_j, \mu_j, X, \alpha) \pi_2. \tag{19}$$

## 5.2. Estimation procedure

Our estimation procedure is described as follows. Observe that steps (i), (iv), (v), (viii) and (ix) are new with regard to Ishwaran-Zarepour (2000) [10] and Ishwaran-James (2002) [11] procedures.

- (i) Choose a large integer  $N$  and generate  $N$  paths  $X_1, \dots, X_N$  of the continuous time Markov chain with distribution  $H$ .
- (ii) Draw  $\alpha$  from  $Gamma(\eta_1, \eta_2)$  and draw  $p_1, \dots, p_N$  according to (10) with  $a_k = 1$  and  $b_k = \alpha$ .
- (iii) Draw  $\lambda = (\lambda_1, \dots, \lambda_M)$  and  $\mu = (\mu_1, \dots, \mu_M)$  from their priors  $\pi_1, \pi_2$  respectively.
- (iv) Draw one of the paths  $X_1, \dots, X_N$  with probability  $p_1, \dots, p_N$ , respectively.
  - Iterate over the following steps (v) through (ix):
  - (v) - Compute  $\sigma_{ij} = COV(\gamma_i, \gamma_j)$  from Proposition 5.2
    - Define  $p_j^* \propto p_j f_n(\gamma|m, \Sigma)$ , using Proposition 5.1
    - Execute (iv) with  $p_j$  replaced by  $p_j^*$ .
  - (vi) - Define  $a_k^*$  and  $b_k^*$  using (12) and (13) where  $X$  is the index of the chosen path.
    - Compute  $p_1 = V_1^*$ , and  $p_k = (1 - V_1^*) \cdots (1 - V_{k-1}^*) V_k^*$ ,  $k = 2, 3, \dots, N - 1$  where  $V_k^* \stackrel{ind}{\sim} Beta(a_k^*, b_k^*)$ .
  - (vii) Draw  $\alpha$  from  $(\alpha|p) \sim Gamma\left(N + \eta_1 - 1, \eta_2 - \sum_{i=1}^{N-1} \log(1 - V_i^*)\right)$ .
- (viii) Given  $\gamma, \lambda$  and a chosen path  $X$ , for each state  $j \in S$ ,
  - If  $X(t) \neq j$  for all  $t \in \{1, \dots, n\}$ , then draw  $\mu_j$  from the prior  $\pi_1$ .
  - otherwise, determine the times  $t_{1j}, \dots, t_{n_j, j}$  at which the Markov chain takes the value  $j$  and compute  $\sigma_j(k\ell)$  from Proposition 5.2.
  - draw  $\mu_j$  from the conditional distribution of  $\mu_j$  given by Proposition 5.6.
- (ix) - If  $X(t) \neq j$  for all  $t \in \{1, \dots, n\}$ , then draw  $\lambda_j$  from the prior  $\pi_2$ .
  - otherwise, determine the times  $t_{1j}, \dots, t_{n_j, j}$  at which the Markov chain takes the value  $j$  and compute  $\sigma_j(k\ell)$  from Proposition 5.2.
  - draw  $\lambda_j$  from the conditional distribution of  $\lambda_j$  given by (19)

**Remark 4** Since  $n$  will be very large for the application that we have in mind ( $n \approx 72,000$ ), it may be

hard to compute the above since we have to invert the matrix  $\Sigma$ . Hence breaking the data set into smaller pieces will help. However we keep the size large enough to estimate the largest significant correlation.

**Remark 5** Another way to simulate  $\lambda$  is to generate a path of the Brownian motion  $B$ . For all times  $t_{1j}, \dots, t_{n_jj}$  for which  $X$  takes value  $j$ , solve equation (8) in  $\lambda_j$  numerically, using a discretization of  $[t_j + 1 - \lambda_j, t_j]$ . This gives values  $\lambda_{1j}, \dots, \lambda_{n_jj}$ . Use the values obtained this way for a large number of Brownian paths to compute the conditional for  $\lambda_j$  and draw value of  $\lambda_j$  from this conditional.

The algorithm below summarizes the estimation procedure. In practice it will be implemented using the Gibbs sampling technique.

(i) **Initialization**

- Let  $\gamma$  be the vector of observations
- Choose the hyperameters  $\eta_1$ ,  $\eta_2$  and  $N$
- Generate  $\alpha$  from Gamma( $\eta_1, \eta_2$ )
- Generate  $N$  paths of the Markov chain
- Draw  $p = p_1 \dots p_N$  from stick-breaking( $\alpha, N$ )
- Choose one of the  $N$  paths according to  $p$
- Generate  $\mu$  and  $\lambda$  from their priors

(ii) **Iterations**

- Compute  $f_n(\gamma, \mu, \lambda)$  and update  $p_k$ ,  $k$  being the index of the chosen path.
- Choose one of the  $N$  paths according to  $p$
- For each state  $j$ , draw  $\mu_j$ , draw  $\lambda_j$
- Draw  $\alpha$
- Draw a new  $p$

### 5.3. Truncation error bound

Let  $\gamma(t)$  be defined as in (8) and let  $\gamma = (\gamma_1, \dots, \gamma_n)$  be a  $n$  sample from the process  $\gamma(t)$ . Let  $m_N(\gamma)$  and  $m_\infty(\gamma)$  denote the marginal density of  $\gamma$  subject to  $\mathcal{P}_N$  and  $D(\alpha H)$  respectively. Using a result in Ishwaran 2002 [11] it can be shown that



$$\int_{\mathbb{R}^n} |m_N(\gamma) - m_\infty(\gamma)| d\gamma \leq 4\exp(-(N-1)/\alpha).$$

This result provides an error bound for the truncated Dirichlet process and shows that the sample size  $n$  has no effect on the bound. The adequacy of the truncation then depends on  $N$  and  $\alpha$ . Of course the value of  $\alpha$  changes during the different iterations of our Gibbs sampler; but, since the bound decreases exponentially, even for a fairly large value  $\alpha = 3$  for example, a truncation of  $N = 30$  leads to an error bound of  $25 \cdot 10^{-5}$ . For the computations in the next section we have chosen a value of  $N = 50$ .

## 6. Numerical results

### 6.1. Simulated data

The present subsection aims at testing the reliability of the model. We perform numerical simulations of the stochastic process  $\gamma(t)$ . We consider a model with five regimes. The associated Markov chain then has five states and is defined by the following uniform transition rate matrix

$$Q = \begin{pmatrix} -1 & 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & -1 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & -1 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & -1 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 & -1 \end{pmatrix}$$

and the initial distribution  $\pi_o = [0.20 \ 0.20 \ 0.20 \ 0.20 \ 0.20]$ . We first run 1,500 burn-in iterations and at the end of each of these iterations we estimate from the obtained path, the transition probabilities and the average time the chain spends in each regime. We get:

$$P = \begin{pmatrix} 0 & 0.2 & 0.7 & 0 & 0.1 \\ 0.5 & 0 & 0.1 & 0.26 & 0.14 \\ 0.6 & 0.18 & 0 & 0.17 & 0.05 \\ 0.55 & 0.13 & 0.1 & 0 & 0.22 \\ 0.08 & 0.43 & 0.16 & 0.33 & 0 \end{pmatrix}$$

and

	Regime 1	Regime 2	Regime 3	Regime 4	Regime 5
$\beta_i$	0.05	0.1	0.02	0.2	0.041

where  $\beta_i$  is the parameter of the exponential distribution that gives the time spent by the chain in state  $i$ . This leads to a new transition rate matrix

$$Q_O = \begin{pmatrix} -0.05 & 0.01 & 0.035 & 0 & 0.005 \\ 0.05 & -0.1 & 0.01 & 0.026 & 0.014 \\ 0.012 & 0.0036 & -0.02 & 0.0034 & 0.001 \\ 0.11 & 0.026 & 0.02 & -0.2 & 0.044 \\ 0.00328 & 0.01763 & 0.00656 & 0.01353 & -0.041 \end{pmatrix}$$

We choose the prior of  $\mu$  and  $\lambda$  to be independent truncated Gaussian distributions, and simulate the parameters  $\mu = (\mu_1, \dots, \mu_5)$  and  $\lambda = (\lambda_1, \dots, \lambda_5)$ , corresponding to the five regimes. We also simulate a path of length  $n = 600$  of the Markov chain. Using  $\mu$ ,  $\lambda$  and the Markov chain, we simulate a sample path  $\gamma$  of the stochastic process  $\gamma(t)$ .

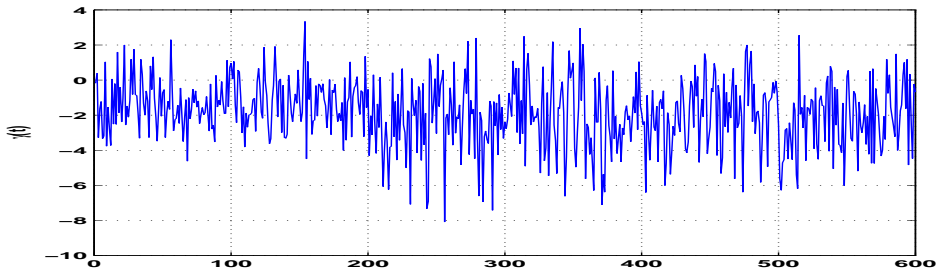


Fig. 2. A sample path of the stochastic process  $\gamma(t)$

Taking the data  $\gamma$  as input, we estimate the parameters of the model through the algorithm presented in Section 5. For the purpose, we performed 500 separate Gibbs sampling runs of length 1,000 each one. The values of the last iteration of each run were retained, leading to the following statistics.

	Regime1	Regime 2	Regime 3	Regime 4	Regime 5
Simulated Values	0.19	0.33	0.36	0.41	0.45
Estimated Values	0.22	0.31	0.38	0.42	0.46
95% confidence Interval	[0.189, 0.25]	[0.285, 0.335]	[0.35, 0.40]	[0.405, 0.43]	[0.44, 0.48]

Table 1: Summary statistics of  $\mu$ .

	Regime1	Regime 2	Regime 3	Regime 4	Regime 5
Simulated Values	1067	997	1234	1743	1408
Estimated Values	1070	996	1234	1742	1410
95% confidence Interval	[1064, 1075]	[992, 1000]	[1228, 1240]	[1734, 1749]	[1403, 1416]

Table 2: Summary statistics of  $\lambda$ .

At each run of the Gibbs sampling, we also choose the path with maximum likelihood. At the end of the procedure, we chose among the 500 paths retained, the one with the highest likelihood. It has the following characteristics:

	Regime1	Regime 2	Regime 3	Regime 4	Regime 5
$\mu$	0.22	0.31	0.38	0.43	0.46
$\lambda$	1070	996	1234	1742	1410

Table 3: Regime characteristics of the process  $\gamma(t)$

## 6.2. Wind speed data

This section is aimed at testing the model on real data. We consider a dataset collected at the entrance of the mangrove ecosystem in Guadeloupe island. Wind velocity was recorded at a frequency of 20Hz by its

3D components  $v_x, v_y$  and  $v_z$ . As we are interested by the longitudinal velocity, only the component  $v_x$  and  $v_y$  are considered. Our observation time interval is one hour, providing a 2D series of length  $n = 72,000$ .

Let  $u = \frac{1}{\sqrt{(v_x)^2 + (v_y)^2}}(\bar{v}_x, \bar{v}_y)$  be the mean longitudinal velocity vector and  $w = \frac{1}{\sqrt{(v_x)^2 + (v_y)^2}}(-\bar{v}_y, \bar{v}_x)$ . Let  $(S_1, S_2)$  be the new coordinates of  $(v_x, v_y)$  in the basis  $(u, w)$  and let  $S = \sqrt{S_1^2 + S_2^2}$  be the wind modulus. Computing the energy dissipation series  $\varepsilon(t) = (S(t+1) - S(t))^2 / (1/20)$ , the aim is to fit our model (15) to the series  $\gamma(t) = \log(S(t))$ .

Estimates of  $\mu$  and  $\lambda$ , using energy spectrum and equation (3) on sliding windows of 60 seconds length, show that these parameters remain quite constant for random durations of time and then jump to another value. Regime changes can be observed in Fig. 2 below as mentioned in the Introduction section.

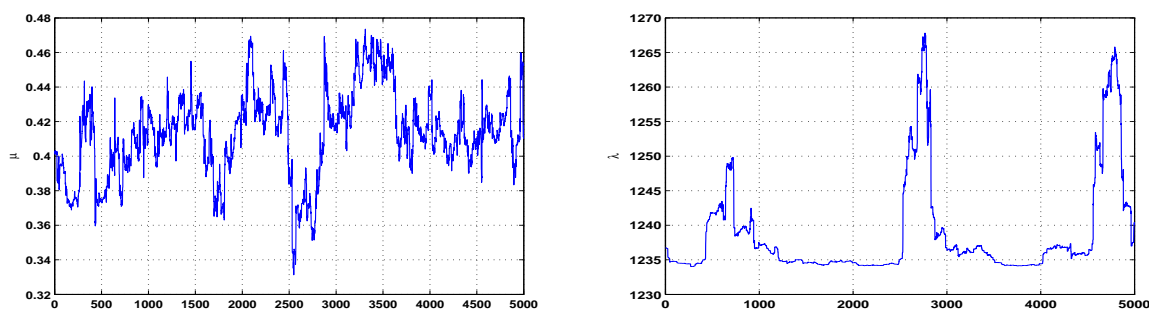


Fig. 3. Regime changes for  $\mu$  (left) and  $\lambda$  (right).

Considering the histogram of the values of  $\mu$  (resp. of  $\lambda$ ) over the above sliding windows (see Fig. 3 below), a truncated Gaussian (resp. a *Gamma*) distribution is taken as initial prior for  $\mu$  (resp. for  $\lambda$ ).

The results of our algorithm for 500 Gibbs sampling runs of length 1,000 each after 3,000 burn-in

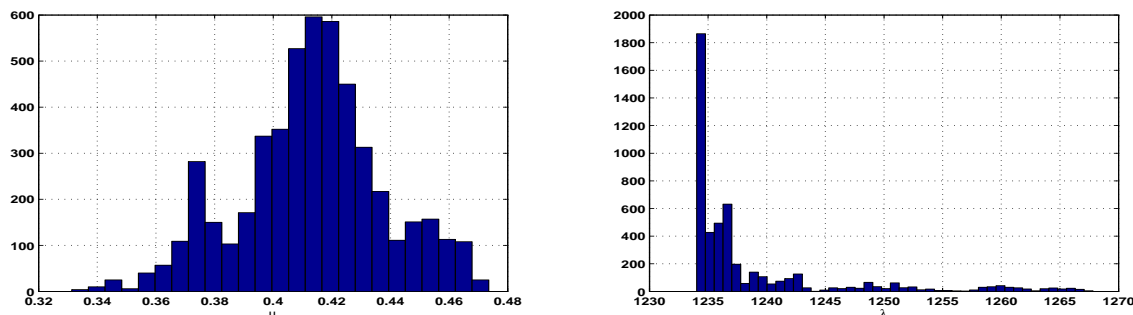


Fig. 4. Histograms of  $\mu$  (left) and  $\lambda$  (right)

iterations, are displayed in the following table. We first ran the algorithm with  $M = 5$  states but some values of the parameters were not significantly different so that we took  $M = 4$  states. We also evaluated

the likelihood of the paths selected at each of the 1,000 iterations of a Gibbs sampling run and we chose the path with maximum likelihood. Among these paths, the one with the highest likelihood has the following characteristics:

	Regime 1	Regime 2	Regime 3	Regime 4
$\mu$	0.33	0.37	0.41	0.47
$\lambda$	1248	1262	1236	1280
% of occupation	11.9%	26.8%	47%	14.3%

Table 4: Characteristics of the path with highest likelihood.

Regimes	2	1	3	1	3	4	3	2	3	4	1	3	4
Duration	58	24	112	44	23	34	53	103	73	36	4	21	16

Table 5: Sequence of regimes in the highest likelihood path.

## 7. Conclusion

We have proposed a new model for dissipation: cascade SDE with Markov regime switching to represent the environment randomness, and Dirichlet prior on the path space of the continuous time Markov chain to make the model more flexible. It can be seen that this model is a complex mixture hierarchical model. The numerical results obtained lead us to think that such mixture model better fit to the real world than usual SDE models. The proposed Bayesian algorithmic method, whose key idea is the simulation of paths, can be extended to many other situations as soon as posterior distributions can be computed or simulated and priors used cautiously. A nice topic for further research may consist in replacing the continuous time Markov chain by a diffusion process. This requires a deeper study of the behaviour of the parameters.

## APPENDIX: PROOFS

### Proof of Lemma 1

Note that  $Pr(X = j|p) = p_j = (1 - V_1) \cdots (1 - V_{j-1})V_j$ . Thus, if  $A_1, \dots, A_n$  are measurable subsets of  $[0,1]$ , and if  $\pi$  is the joint distribution of the  $V_i$ 's, we have,

$$\begin{aligned}
 Pr(V_1 \in A_1, \dots, V_N \in A_N, X = j) &= \int \prod_{k=1}^N I_{\{V_k \in A_k\}} Pr(X = j|V_1, \dots, V_N) \pi(dV_1, \dots, dV_N) \\
 &= \int \prod_{k=1}^N I_{\{V_k \in A_k\}} (1 - V_1) \cdots (1 - V_{j-1}) V_j \pi(dV_1, \dots, dV_N) \\
 &= \int I_{\{x_k \in A_k, k=1, \dots, N\}} (1 - x_1) \cdots (1 - x_{j-1}) x_j \prod_{k=1}^N x_k^{a_k-1} (1 - x_k)^{b_k-1} (dx_1, \dots, dx_N) \\
 &= \int I_{\{x_k \in A_k, k=1, \dots, N\}} x_j^{a_j} (1 - x_j)^{b_j-1} \prod_{k=1}^{j-1} x_k^{a_k-1} (1 - x_k)^{b_k} \prod_{k=j+1}^N x_k^{a_k-1} (1 - x_k)^{b_k-1} (dx_1, \dots, dx_N).
 \end{aligned}$$

This implies that

$$\begin{aligned}
 V_j | X = j &\sim \text{Beta}(a_j + 1, b_j) \\
 V_k | X = j &\sim \text{Beta}(a_k, b_k + 1) \quad \text{for } 0 < k < j \\
 V_k | X = j &\sim \text{Beta}(a_k, b_k) \quad \text{for } k > j
 \end{aligned}$$

as summarized in the Lemma.

### Proof of Proposition 5.1

According to definition (8), each component  $\gamma_i$  of  $\gamma$  is a Gaussian r.v. with mean

$$m_i(X) = -\frac{1}{2} \mu_{X(i)} \log(\lambda_{X(i)})$$

since

$$\left\langle \int_{i+1-\lambda_{X(i)}}^i (i+1-u)^{-1/2} dB(u) \right\rangle = 0.$$

Therefore

$$\langle \gamma \rangle = m(X) = \left( -\frac{1}{2} \mu_{X(1)} \log(\lambda_{X(1)}), \dots, \frac{1}{2} \mu_{X(n)} \log(\lambda_{X(n)}) \right).$$

Now let  $Z = \sum_{j=1}^J \alpha_j \gamma_j$  be a linear combination of components of  $\gamma$ .

$$Z = \sum_{j=1}^J \alpha_j \left( -\frac{\mu_{X(t_j)}}{2} \log(\lambda_{X(t_j)}) + (\mu_{X(t_j)})^{1/2} \int_{t_j+1-\lambda_{X(t_j)}}^{t_j} (t_j + 1 - u)^{-1/2} dB(u) \right)$$

which can be written in the form

$$Z = A_J + \sum_{j=1}^J \alpha_j (\mu_{X(t_j)})^{1/2} \int (t_j + 1 - u)^{-1/2} I_{[t_j+1-\lambda_{X(t_j)}, t_j]} dB(u)$$

that is

$$Z = A_J + \int (t_j + 1 - u)^{-1/2} B_J I_{[t_j+1-\lambda_{X(t_j)}, t_j]} dB(u),$$

showing that  $Z$  is a Gaussian r.v. It follows that  $\gamma$  is a Gaussian random vector.

Moreover

$$\begin{aligned} \sigma_{st} &= \text{Cov}(\gamma_s(X), \gamma_t(X)) = \langle \gamma_s(X) - m_s(X), \gamma_t(X) - m_t(X) \rangle \\ &= \left\langle (\mu_{X(s)})^{1/2} \int_{s+1-\lambda_{X(s)}}^s (s+1-u)^{-1/2} dB(u), (\mu_{X(t)})^{1/2} \int_{t+1-\lambda_{X(t)}}^t (t+1-v)^{-1/2} dB(v) \right\rangle \\ &= (\mu_{X(s)} \mu_{X(t)})^{1/2} \left\langle \int_{s+1-\lambda_{X(s)}}^s (s+1-u)^{-1/2} dB(u), \int_{t+1-\lambda_{X(t)}}^t (t+1-v)^{-1/2} dB(v) \right\rangle. \quad \square \end{aligned}$$

### Proof of Proposition 5.2

The covariance matrix coefficients  $\sigma_{st} = \text{COV}(\gamma_s, \gamma_t)$  involve two Gaussian stochastic integrals:

$$\sigma_{st} = (\mu_{X(s)} \mu_{X(t)})^{1/2} \left\langle \int_{s+1-\lambda_{X(s)}}^s (s+1-u)^{-1/2} dB(u), \int_{t+1-\lambda_{X(t)}}^t (t+1-u)^{-1/2} dB(u) \right\rangle.$$

Recall that Gaussian stochastic integrals are zero mean Gaussian r.v.s that have the property

$$\left\langle \int_{E_1} F(x) dB(x) \int_{E_2} G(x) dB(x) \right\rangle = \int_{E_1 \cap E_2} F(x) G(x) dx.$$

So, if

$$I = \int_b^a F(x) dB(x)$$

then

$$\sigma_I^2 = \int_b^a F^2(x) dx.$$

Suppose that  $s < t$ , then  $w = s$  and  $t = s + \tau$ . Let

$$K = \left\langle \int_{s+1-\lambda_{X(s)}}^s (s+1-u)^{-1/2} dB(u), \int_{t+1-\lambda_{X(t)}}^t (t+1-u)^{-1/2} dB(u) \right\rangle.$$

It follows that

$$K = \int_a^s (s+1-u)^{-1/2} (t+1-u)^{-1/2} du.$$

As in Schmitt [20] we see that

$$\begin{aligned} K &= \int_a^s \frac{du}{\sqrt{(s+1-u)(t+1-u)}} \\ &= \int_a^s \frac{du}{\sqrt{(s+1-u)(s+1-u+\tau)}} \\ &= \int_1^{1+s-a} \frac{dx}{\sqrt{x(x+\tau)}} \\ &= 2 \log \left( \frac{\sqrt{s+1-a} + \sqrt{s+1-a+\tau}}{1 + \sqrt{1+\tau}} \right) \end{aligned}$$

where we have made the variable change  $x = s+1-u$  and used the identity

$$\int \frac{dx}{\sqrt{x(x+\tau)}} = 2 \log(\sqrt{x} + \sqrt{x+\tau}).$$

Therefore

$$\sigma_{st} = 2(\mu_{X(s)}\mu_{X(t)})^{1/2} \log \left( \frac{\sqrt{w+1-a} + \sqrt{w+1-a+\tau}}{1 + \sqrt{1+\tau}} \right). \quad \square$$

### Proof of Proposition 5.3

We know that the conditional density of  $X$  is

$$(X = x | p = y, \gamma = g, \mu = m, \lambda = l) = \frac{(X, p, \gamma, \mu, \lambda)(x, y, g, m, l)}{\int (X, p, \gamma, \mu, \lambda)(x, y, g, m, l) dP(x)}.$$



But,

$$\begin{aligned}
(X, p, \gamma, \mu, \lambda)(x, y, g, m, l) &= (\gamma = g | X = x, p = y, \mu = m, \lambda = l)(X = x, p = y, \mu = m, \lambda = l) \\
&= (\gamma = g | X = x, p = y, \mu = m, \lambda = l)(X = x | p = y, \mu = m, \lambda = l) \\
&\quad \times (p = y, \mu = m, \lambda = l) \\
&= f_n(g, m, l)(X = x | p = y, \mu = m, \lambda = l)(p = y, \mu = m, \lambda = l).
\end{aligned}$$

Since

$$(X = x | p, \mu, \lambda) = \sum_{i=1}^N p_i \delta_{X_i}(\{x\}),$$

and since the distribution of  $(p, \mu, \lambda)$  does not depend on  $X$ , we have

$$(X, p, \gamma, \mu, \lambda)(x, y, g, m, l) = f_n(g, m, l) \sum_{i=1}^N p_i \delta_{X_i}(\{x\}).$$

Using that  $f(x) \delta_{X_i}(\{x\}) = f(X_i) \delta_{X_i}(\{x\})$ , for any function  $f$ , we get

$$\begin{aligned}
(X, p, \gamma, \mu, \lambda)(x, p, g, m, l) &= \sum_{i=1}^N f_n(g, m, l) p_i \delta_{X_i}(\{x\}) \\
&= \sum_{i=1}^N p_i^* \delta_{X_i}(\{x\}) \\
(X | p, \gamma, \mu, \lambda) &\propto \sum_{i=1}^N p_i^* \delta_{X_i}.
\end{aligned}$$

### Proof of Proposition 5.5

By Connor and Mosimann [2], the probability density of  $p$  defined by equation (10) is

$$\left\{ \prod_{k=1}^{N-1} \frac{\Gamma(a_k, b_k)}{\Gamma(a_k) \Gamma(b_k)} \right\} p_1^{a_1-1} \dots p_{N-1}^{a_{N-1}-1} p_N^{b_{N-1}-1} \times (1 - P_1)^{b_1 - (a_2 + b_2)} \dots (1 - P_{N-2}^{b_{N-2} - (a_{N-1} + b_{N-1})}), \quad (20)$$

where  $P_k = p_1 + \dots + p_k$ . When  $a_k = 1$  and  $b_k = \alpha$ , using  $\Gamma(1 + \alpha) = \alpha \Gamma(\alpha)$ , we get that the conditional density of  $p$  given  $\alpha$  is

$$f(p|\alpha) \propto \alpha^{N-1} p_N^{\alpha-1} = \alpha^{N-1} e^{(\alpha-1) \log(p_N)}.$$

As  $f(\alpha|p) \propto f(p|\alpha) f(\alpha)$  and the prior for  $\alpha$  is  $Gamma(\eta_1, \eta_2)$ , we get

$$f(\alpha|p) \propto \alpha^{N-1+\eta_1} e^{-(\eta_2 - \log(p_N))\alpha}$$

So,  $(\alpha|p) \sim \text{Gamma}(N + \eta_1 - 1, \eta_2 - \log(p_N))$ , that is

$$(\alpha|p) \sim \text{Gamma}\left(N + \eta_1 - 1, \eta_2 - \sum_{i=1}^{N-1} \log(1 - V_i^*)\right).$$

### Proof of Proposition 5.6

Let  $t_{1j}, \dots, t_{n_jj}$  be the times at which the Markov chain takes value  $j$ .

We know from corollary (2) that,

$$(\gamma_{t_{1j}}, \dots, \gamma_{t_{n_jj}})|_{\mu, \lambda} \sim \mathcal{N}_{n_j}\left(-\frac{\mu_j}{2} \log(\lambda_j)(1, 1, \dots, 1), \mu_j \Sigma_j\right).$$

Moreover

$$(\gamma, \mu_j, \lambda_j, X, \alpha) = (\gamma|\mu_j, \lambda_j, X, \alpha) \otimes (\mu_j, \lambda_j, X, \alpha) = (\gamma|\mu_j, \lambda_j, X, \alpha) \otimes (\mu_j \otimes \lambda_j \otimes X \otimes \alpha)$$

since  $X, \mu_j, \lambda_j$  and  $\alpha$  are independent. It follows that

$$(\mu_j|\gamma, \lambda_j, X, \alpha) = \frac{(\gamma|\mu_j, \lambda_j, X, \alpha) \otimes \mu_j \otimes \lambda_j \otimes X \otimes \alpha}{\int (\gamma, \mu_j, \lambda_j, X, \alpha) dP(\mu_j)}.$$

As  $\lambda_j, X$  and  $\alpha$  do not depend on  $\mu_j$  we have

$$(\mu_j|\gamma, \lambda_j, X, \alpha) \propto (\gamma|\mu_j, \lambda_j, X, \alpha) \otimes \mu_j$$

That is

$$(\mu_j|\gamma, \lambda_j, X, \alpha) \sim \Gamma_1 \propto \mathcal{N}_{n_j}(m_j, \mu_j \Sigma_j). \pi_1. \quad \square$$

### References

- [1] BERNARD, D. and D'ALEXIS, C. (2006). Mesures atmosphériques à l'entrée de la mangrove en Guadeloupe. *Projet Atmo-Mangrove*, LPAT, Université Antilles-Guyane, Guadeloupe.
- [2] CONNOR, R. J. and MOSSIMANN, J. E. (1969). Concept of independence for proportions with a generalization of Dirichlet Distribution. *J. Am. Statist. Assoc.* **64** 194-206.
- [3] DI-MASI, G. B., KABANOV, M. Y. and RUNGALDIER, W. J. (1994). Mean-variance hedging of options on stocks with Markov volatility. *Theory Probab. App* **39** 173-174.

- [4] DUPONT, S. and BRUNET, Y. (2008). Edge flow and canopy structure : A large-Eddy simulation study. *Boundary-Layer Meteorol.* **126** 51-71.
- [5] EMILION R., FAIRES H. and IYER S. (2007). Dirichlet prior on path spaces with application to modeling asset prices. *Preprint*.  
<http://www.univ-orleans.fr/mapmo/membres/emilion/>
- [6] FERGUSON, T. F. (1973). A Bayseian analysis of some non parametric problems. *Ann. Stat.* **1** 209-230.
- [7] FOU DHIL, H., BRUNET, Y. and CALTAGIRONE J.P.(2005). A fine-scale k-e model for atmospheric flow over heterogeneous landscapes. *Environ. Fluid Mech.***5** 245-267.
- [8] DESHPANDE, A. and GHOSH, M. K. (2008). Risk minimizing option pricing in a regime switching market. *Stoch. Analysis and App* vol **28**, to appear.
- [9] GURVICH, A. S. and ZUBKOSKI, S. L. (1963). *Izv. Akad. Nauk. SSSR, Ser. Geofiz.* **1856**.
- [10] ISHWARAN, H. and ZAREPOUR, M. (2000) Markov Chain Monte Carlo in Approximate Dirichlet and Beta Two-Parameter Process Hierarchical Models, *Biometrike* **87**(2) 371-390.
- [11] ISHWARAN, H. and LANCELOT, F. J. (2002) Approximate Dirichlet process computing in finite normal mixtures: smoothing and prior information. *Journal of Computational and Graphical Statistics* **11**(3) 1-26.
- [12] ISHWARAN, H. and LANCELOT, F. J. (2003). Some futher developments for Stick-breaking priors: Finite and infinite clustering and clasification. *The Indian Journal of Statistics* **65**(3) 577-592.
- [13] KATUL, G. et al. (2006). An analysis of intermittency, scaling, and surface renewal in atmospheric surface layer turbulence. *Physica D* **215** 117-126.
- [14] KRAFT, C. H. (1964). A class of distribution function processes which have derivatives. *Journal of Applied Probability* **1** 384-388.
- [15] KOLMOGOROV, A. N. (1962). A refinement hypothesis concerning the localstructure of turbulence in a viscous incompressible fluid at high Reynolds number. *Journal of Fleuid Mech.* **13** 82-85.
- [16] NOVIKOV, E. A. and STEWART, R. W. (1964). *Izv. Geophys.* **3** 408.
- [17] OBUKHOV, A M. (1962). Some specific features of atmospheric turbulence. *J. Fluid Mech.* **13**(1) 7781
- [18] PADDOCK, S., RUGGERI, F., LAVINE M. and WEST, M. (2003). A randomised Polya tree models for nonparametric Bayesian inference. *Statistica Sinica* **13**(2) 443-460.
- [19] POND, S. and STEWART, R. W. 1965). *Izv. Akad. Nauk. SSSR, Ser. Fis. Atmosf. Okeana* **1** 914.
- [20] SCHMITT, F. G. (2003). A causal Stochastic multifractal stohastic equation and its statistical properties. *The European Physical Journal B* **34** 85-98.
- [21] SCHMITT, F. G., VANNITSEM, S. and BARBOSA, A. (1998). Modeling of rainfall time series using renewal processes and multifractals, *J. Geophys. Res.* **103** 23181-23193

- [22] SETHURAMAN, J. (1994). A constructive definition of Dirichlet priors. *Statistica sinica* **4** 639-650.
- [23] SHE, Z-S. and WAYEMIRE, E. C. (1995). Quantized Energy Cascade and Log-Poisson Statistics in Fully Developed Turbulence *Phys. Rev. Lett.* **74** 262-265
- [24] SCHERTZER D. and LOVEJOY, S. (1991). *In scaling, fractals and non-linear variability in geophysics*, edited by D. Schertzer, S. Lovejoy (Kluwer, Berlin) p.41.
- [25] SU, H-B., SHAW RH., PAW UKT., MOENG C-H. and SULLIVAN PP. (1998). Turbulent statistics of neutrally stratified flow within and above a sparse forest from large-eddy simulation and field observations. *Boundary-Layer Meteorol.* **88**363-397
- [26] YAGLOM, A. M. (1966). The influence of fluctuations in enegy dissipation on the shape of turbulent characteristics in the inertial interval. *Soviet Physics Doklady* **11** 26-29.