Stochastic Evolution of Distributions - Applications to CDS indices

Guillaume Bernis, Nicolas Brunel, Antoine Kornprobst, Simone Scotti

UNIVERSITÉ PARIS 1
PANTHÉON SORBONNE

Documents de Travail du
Centre d'Economie de la Sorbonne

CES Working Papers

Stochastic Evolution of Distributions – Applications to CDS indices

Guillaume BERNIS, Nicolas BRUNEL,
Antoine KORNPROBST, Simone SCOTTI

2017.07

Maison des Sciences Économiques, 106-112 boulevard de L'Hôpital, 75647 Paris Cedex 13
http://centredeconomiesorbonne.univ-paris1.fr/
ISSN : 1955-611X

Stochastic Evolution of Distributions - Applications to CDS indices

Guillaume Bernis[*], Nicolas Brunel[†], Antoine Kornprobst[‡] and Simone Scotti[§]

December 5, 2016

Abstract: We use mixture of percentile functions to model credit spread evolution, which allows to obtain a flexible description of credit indices and their components at the same time. We show regularity results in order to extend mixture percentile to the dynamic case. We characterise the stochastic differential equation of the flow of cumulative distribution function and we link it with the ordered list of the components of the credit index. The main application is to introduce a functional version of Bollinger bands. The crossing of bands by the spread is associated with a trading signal. Finally, we show the richness of the signals produced by functional Bollinger bands compared with standard one with a practical example.

1 Introduction

The modelling of both market indices and their components is an open question in finance. Notably, there are many relevant papers (e.g. [12], [13]) focusing on the global evolution of a market and its underlying assets. However, this analysis is generally applies to stock assets but cannot be extended directly to other classes and mainly to credit risk market. The aim of this paper is, first, to fill this gap

[*]Natixis Asset Management, Fixed Income. eMail: guillaume.bernis@am.natixis.com. The opinions and views expressed in this document are those of the authors and do not necessarily reflect those of Natixis Asset Management.

[†]Laboratoire de Mathématiques et Modélisation d'Evry - UMR CNRS 8071. Université d'Evry Val d'Essonne and ENSIIE

[‡]Université Paris 1 Panthéon Sorbonne, Labex RéFi

[§]LPMA, Université Paris Diderot. eMail: scotti@math.univ-paris-diderot.fr

proposing a model to describe the evolution of a credit default swap (CDS) index and their components in a parsimonious but flexible way, second, to produce trading signals based on this model. The trading signals are an extension of the widely used Bollinger bands (See [17]) to infinite dimensional, functional indicators.

The guiding idea of this paper is that the distribution of the fair spreads of the components of the index can be used to forecast the evolution of the index itself: in many situation, the deformation of the distribution anticipates the widening or tightening of the index fair spread. The problem of optimal investment and efficient trading signals has been widely studied in literature. Moreover, the techniques applied span a large class of mathematical tools, for instance by restricting to recent years (but without any claim to being exhaustive) optimal switching [25], optimal investment with trend detection [8], [5] and [6], adding jumps and over/under-reaction to information [7] or using techniques from neural network [10]. The significant innovation of the present paper is to deal with the whole components and not the only index, using a diffusion process with values in a functional space.

The iTraxx Europe and CDX Investment Grade indices are liquid CDS indices composed of the 125 most liquid single name CDS referencing, respectively, the European and the US investment grade credit issuers. These are the typical indices on which our method can be applied, although it is not limited to them and can be adapted to cash credit indices. The CDS index is quoted by its fair spread and treated as a CDS itself (and not a pool of single name CDS). Some details about CDS indices can be found in section 2.1 in [5]. Due to technical effects, the fair spread of the iTraxx is not the average fair spread of the pool of single name CDS, but the basis is generally small, which can lead us to look at the distribution of the underlying spreads in order to analyse the future spread evolution of the index.

These indices are equally weighted with standardized maturities and have a large number of components (namely, 125). This leads us to consider the list of the underlying single name CDS as a probability distribution on $\mathbb{R}^+$. Several references on the question of deterministic or stochastic evolution of densities can be found across various fields of mathematics. The scope of applications encompasses statistics, biology, physics and finance. Bellomo and Pistone [4] study the action of an abstract dynamical system on a probability density. From a statistical point of view, some authors analyse the density as the outcome of a Dirichlet law: Ferguson [11] or Shao [20]. This question is also at the core of the optimal transport, see for instance Villani [23], Alfonsi et al [2] and the references therein. In this vein, Bass [3] studies the deformation of a parameter-based family of densities. Another approach is based on the mixture of percentiles. This method introduced by Sillitto [21] has been used, more recently, by Karvanen [16] and Gouriéroux and Jasiak [14] to fit distributions of stock returns. This is the approach that we will adopt in this paper.

Let us consider some integrable probability distribution on the positive half-line, represented by its percentile function. Assume that we have a breakdown the percentile as a sum of percentile functions of distributions on the positive half-line. Then, we can construct deformations of the initial distribution by considering weighted sums of the percentile function with positive coefficients. Then, we replace the mixing coefficients by correlated diffusion processes.

This method construct a process with values in a space of probability distributions, represented by their percentiles: hence a random measure on the positive half-line. We show that, under mild assumptions on the underlying diffusion, regularity results hold for the distribution-valued process. First, the process is continuous in time with respect to the expected Wasserstein distance, see Bass [3], which is a natural metric in this context. Second, the stochastic differential equation which drives the cumulated distribution function is explicit. We also provide results on the average and the variance of the cumulated distribution function.

Using this framework, at each date, we can calculate a confidence interval for the percentile at a given (short) time horizon. We show that the upper and lower boundaries of this interval are also percentile functions. We study trading signals triggered when the realised percentile function crosses either the upper or the lower band of this confidence interval. This method can be seen as a extension of the Bollinger bands, where the bands are not only functions of the time, but functions of the time and of the level of the percentile. Hence, it defines functional lower and upper bands. We also analyse the link between the crossing of one of the bands and the measure of risk according to the second order stochastic dominance.

In the first section, we introduce the mixture of percentiles method in a static setting and analyse its properties. Topological aspects as well as differentiability of the cumulated distribution functions are studied. Insights on the notion of stochastic dominance, in relation with the possible deformations of the initial distribution, are given. Then, in section 3, we use Markov diffusion processes as mixing coefficients and provide the dynamics of the distribution. Thus, we obtain a model where both percentile and cumulative distribution functions have some explicit diffusion equation, and satisfies some regularity results. Last section is dedicated to a case study for CDS indices.

2 Decomposition of probability distributions

This section is dedicated to the analysis of the mixture of percentiles, as defined in Sillitto [21], in a static context. We prove regularity results that will be useful for the dynamic case, detailed in section 3. After setting our framework, we present the mixture method and study its properties. Then, we derive results on the derivatives of the cumulative distribution function. Finally, we investigate the deformation of the distribution in term of stochastic dominance.

First, let us define the setting under which we will work throughout the paper. Consider the measured space $\big([0,+\infty), \mathcal{B}([0,+\infty)), dx\big)$, where $\mathcal{B}\big([0,+\infty)\big)$ is the Borelian sigma-field over $[0,+\infty)$ and dx the Lebesgue measure. Let f be a probability distribution function (hereafter p.d.f.) on this space, F the cumulative distribution function (c.d.f.) of f, and q the percentile function, i.e. the inverse[1] of F. We denote by $\mathcal{D}$ the set of all probability distributions on $\big([0,+\infty), \mathcal{B}([0,+\infty)), dx\big)$ and by (f, F, q) an element of $\mathcal{D}$. Where there is no ambiguity, we will indicate only one element among the p.d.f, the c.d.f or the percentile function. The next straightforward Lemma highlights some properties of $\mathcal{D}$.

Lemma 1 (Properties of percentile function) *The set $\mathcal{D}$ is a convex cone i.e. $\forall (q_1, q_2) \in \mathcal{D}^2$ and $(\lambda_1, \lambda_2) \in (\mathbb{R}_*^+)^2$, $\lambda_1 q_1 + \lambda_2 q_2 \in \mathcal{D}$. Moreover, it can be endowed with a partial order relation on the set $\mathcal{D}$, equivalent to the stochastic dominance of first order, as defined by Quirk and Saposnik [19], see also Levy [18]: $q_1 \succeq q_2$ if, for any $\varepsilon \in [0,1)$, $q_1(\varepsilon) \geq q_2(\varepsilon)$, which is equivalent to, $\forall\, x \geq 0$, $F_1(x) \leq F_2(x)$, in terms of c.d.f. For any $q_1 \succeq q_2$ and q, $q_1 + q \succeq q_2$.*

Throughout the paper, we consider the set $\mathcal{D}_0 \subset \mathcal{D}$ of probability distributions satisfying the following assumption:

Assumption 1 *A probability distribution $(f, F, q) \in \mathcal{D}$ is said to belong to $\mathcal{D}_0$ if:*

1. *Its p.d.f. is positive, almost everywhere (a.e.).*
2. *Its c.d.f. is a diffeomorphism from $[0,+\infty)$ onto $[0,1)$ and is twice continuously derivable with right-derivatives at 0. Moreover, $F(0) = q(0) = 0$.*
3. *Its probability distribution has finite first two moments.*

[1] The notion of generalized inverse should be used at this stage, but, hereafter, we will restrict our attention to distributions for which the c.d.f is actually invertible.

We can extend the previous Lemma 1 to show:

Proposition 1 *The set $\mathcal{D}_0$ is a convex cone.*

Proof: Let $q_1, q_2 \in \mathcal{D}_0$, since $\mathcal{D}_0 \subset \mathcal{D}$, applying the results of Lemma 1 we have that $q := \lambda_1 q_1 + \lambda_2 q_2 \in \mathcal{D}$ for all $\lambda_1, \lambda_2 \in \mathbb{R}_*^+$. Then we need to show the three properties in Assumption 1:

1. The mapping q is strictly increasing, with $q(0) = 0$, and tends to $+\infty$ in 1^-. Thus, it defines a bijection from $[0, 1)$ onto $[0, +\infty)$. It is, therefore, invertible, with an inverse denoted by F and strictly increasing. This implies that the p.d.f is well defined and positive a.e.

2. By implied function theorem, the derivative of F exists. Moreover, as q_1 and q_2 have derivatives of order up to two, so q has the same property and F too.

3. The existence of the moments of first and second orders stems from the following identity, obtained by the variable change $x = q(\varepsilon)$, i.e. $\varepsilon = F(x)$: for $k = 1, 2$
$$\int_0^\infty x^k f(x)dx = \int_0^1 (q(\varepsilon))^k d\varepsilon$$
Thus, if (f, F, q) has a second order moment so has $\left(\frac{1}{\lambda} f\left(\frac{\cdot}{\lambda}\right),\ F\left(\frac{\cdot}{\lambda}\right),\ \lambda q\right)$, for any $\lambda > 0$: $\int_0^1 (\lambda q(\varepsilon))^k d\varepsilon = \lambda^k \int_0^1 (q(\varepsilon))^k d\varepsilon$, $k \in \{1, 2\}$. Therefore, it is sufficient to show that, for $q_1, q_2 \in \mathcal{D}_0$, $q_1 + q_2$ and $(q_1 + q_2)^2$ are integrable. The integrability of $q_1 + q_2$ is a consequence of the linearity of the integral. In order to show the integrability of $(q_1 + q_2)^2 = q_1^2 + q_2^2 + 2q_1q_2$, it is sufficient to show that q_1q_2 is integrable. This is a consequence of the Cauchy-Schwarz inequality. $\square$

In the following, we endow the space $\mathcal{D}_0$ with the second-order Wasserstein distance, see for instance Vallander [22], which is defined, for any $(q_1, q_2) \in \mathcal{D}_0$ by

$$\mathcal{W}_2(q_1, q_2) := \sqrt{\int_0^1 \left[q_1(\varepsilon) - q_2(\varepsilon)\right]^2 d\varepsilon}$$

This distance is well suited to the analysis of deformation of probability distributions, and can be adapted in a dynamic setting. See, for instance, Bass [3] and Alfonsi et al. [2].

2.1 Mixture of percentile functions

In this section, following Sillitto [21], we introduce the method to construct various percentiles functions from an initial one. We provide results on the regularity of the functions obtained by this method. In particular, we prove that they belong to $\mathcal{D}_0$. In this section, we set out $(f, F, q) \in \mathcal{D}_0$.

Definition 1 (Basis of percentile function associated to (f, F, q)) *A n-uple of mappings $\psi := (\psi_i)_{1 \leq i \leq n}$, is called a n-basis of percentile functions associated to (f, F, q) if*

1. *The family is linearly independent;*
2. *For all $i \in \{1, \dots, n\}$ ψ_i is non-decreasing, taking non-negative values, twice continuously derivable on $[0, 1)$ and*
3. *The total sum of the percentile functions ψ_i reconstruct the percentile function q, i.e. $\sum_{i=1}^n \psi_i = q$.*

The set of all n-basis of percentile function associated to (f, F, q) will be denoted by $\mathcal{P}_n(q)$.

The following Lemma summarizes some properties of the basis representation:

Lemma 2 *Let ψ be a n-basis of percentile functions associated to (f, F, q) satisfying Assumption 1. Then, we have*

1. **Initial value**: *For all $i \in \{1, \dots, n\}$ $\psi_i(0) = 0$*
2. **Existence of a diverging term**: *There exists $i^* \in \{1, \dots, n\}$ such that $\lim\limits_{\varepsilon \to 1^-} \psi_{i^*}(\varepsilon) = \infty$.*

Proof: The first property is a direct consequence of $q(0) = 0$ and the fact that the functions ψ_i take non-negative values. If the second property does not hold, the limit of q, as ε goes to $1-$, would be finite. Hence a contradiction with point 1 in Assumption 1. $\square$

Now, let us introduce a method to generate a large class of probability distribution based on a basis of percentile functions.

Definition 2 (Mixture Method) *Let $y := (y_i)_{1 \le i \le n} \in (\mathbb{R}_*^+)^n$, and $\psi \in \mathcal{P}_n(q)$. The mixing of ψ with coefficients y is defined by*

$$\forall \varepsilon \in [0,1),\ \ q(\varepsilon, y) = q(x, y_1, \ldots, y_n) := \langle y, \psi(\varepsilon) \rangle = \sum_{i=1}^{n} y_i \psi_i(\varepsilon). \tag{1}$$

where $\langle \cdot, \cdot \rangle$ stands for the Euclidean scalar product on $\mathbb{R}^n$.

Note that $q(\varepsilon) = \langle e, \psi(\varepsilon) \rangle = q(\varepsilon, e)$, where $e := (1, \ldots, 1)^T \in \mathbb{R}^n$. In order to maintain consistent notations throughout the paper, the first variable ε will always denote the level of probability used in the percentile function and x and element of $[0, +\infty)$, i.e. the variable of the p.d.f. and c.d.f., whereas the element y of $(\mathbb{R}_*^+)^n$ will represent a vector of mixing coefficients. The function $\varepsilon \mapsto q(\varepsilon, y)$ will be denoted, for short, $q(\cdot, y)$. Accordingly, we denote by $f(\cdot, y)$ the p.d.f. and by $F(\cdot, y)$ the c.d.f. associated to $q(\cdot, y)$. Both functions are defined on $[0, +\infty)$. It is clear that the mixing method defines an element of $\mathcal{D}$. As we will see in Corollary 1, it also defines an element of $\mathcal{D}_0$, this is a direct consequence of Proposition 1.

Corollary 1 *Let y and ψ as in Definition 2. Then, the p.d.f associated to $q(\cdot, y)$, given by Equation (1), satisfies Assumption 1. Moreover, if $y_{min} = \min\{y_i | 1 \le i \le n\}$, we have, for $1 \le k \le 2$,*

$$y_{min}^k \int_0^{+\infty} x^k f(x) dx \le \int_0^{+\infty} x^k f(x, y) dx \le \langle y, e \rangle^k \int_0^{+\infty} x^k f(x) dx \tag{2}$$

Proof: The first part of the proof is a direct consequence of Lemma 2. For the integrability condition, let us write $\forall \varepsilon \in [0,1),\ \ q(\varepsilon) y_{min} \le q(\varepsilon, y) \le q(\varepsilon) \times \langle y, e \rangle$.

Therefore, we have boundaries on $q(\cdot, y)$ for the first order stochastic dominance. This can be translated into the inverse order for integral of any increasing mapping, see Levy [18]. Taking $x \mapsto x^k$, $k \in \{1, 2\}$ yields Equation (2). □

Another result on the influence of the mixture method on the distribution is given in Proposition 2. To this purpose, let us denote by $|\cdot|$ the Euclidean norm on $\mathbb{R}^n$, and by $||\cdot||_2$ the L^2-norm on $\mathcal{D}_0$.

Proposition 2 *The mapping $y \mapsto f(\cdot, y)$, from $(\mathbb{R}_*^+)^n$ into $\mathcal{D}_0$, endowed with the Wasserstein distance $\mathcal{W}_2$, is Lipschitz with coefficient $||q||_2$.*

Proof: For any $y, z \in (\mathbb{R}_*^+)^n$, the Wasserstein distance satisfies the following inequality, which is a consequence of Schwarz inequality in $\mathbb{R}^n$:

$$\mathcal{W}_2\big(q(\cdot, y), q(\cdot, z)\big) = \left[\int_0^1 < y - z, \psi(\varepsilon) >^2 d\varepsilon \right]^{\frac{1}{2}} \le \left[\sum_{i=1}^{n} \int_0^1 |\psi(\varepsilon)|^2 \times |y - z|^2 d\varepsilon \right]^{\frac{1}{2}}.$$

But, we have $|\psi(\cdot)| \leq q(\cdot)$ due to the non-negativity of ψ, and, thus, the following inequality holds: $\mathcal{W}_2(q(\cdot,y),q(\cdot,z)) \leq ||q||_2 \times |y-z|$. $\square$

With these result in hand, we can investigate the topological properties of the set of all possible mixtures given a basis of percentile functions.

Lemma 3 *Let $\psi_i \in \mathcal{D}_0$, for $i \in \{1,\ldots,n\}$ and assume that the family $\psi := (\psi)_{1\leq i\leq n}$ is linearly independent. Let $A := \left\{q(\cdot,y) \mid y \in (\mathbb{R}_*^+)^n\right\}$ be the set of mixtures associated to the family $(\psi_i)_{1\leq i\leq n}$. Then, the closure of A for $\mathcal{W}_2$ is the convex cone spanned by the ψ_i, $i \in \{1,\ldots,n\}$ and an element of the closure is either 0 or an element of $\mathcal{D}_0$.*

Proof: Let $(q_k)_{k\geq 0} \in A^{\mathbb{N}}$ converging toward q^*. For each q_k, there exists $y_k \in (\mathbb{R}_*^+)^n$. On the one hand, a converging sequence is also a Cauchy sequence, hence, for any $k \geq 0$, $\mathcal{W}_2(q_k,q_{k+p})$ tends to 0 as p goes to infinity. On the other hand, $\mathcal{W}_2^2(q_k,q_{k+p}) = (y_{k+p}-y_k)^T\Psi(y_{k+p}-y_k)$ where Ψ is the $n\times n$-matrix, the elements of which are given by $\Psi_{k,i} := \int_0^1 \psi_k(\varepsilon)\psi_i(\varepsilon)d\varepsilon$ for $(i,k) \in \{1,\ldots,n\}^2$. As the family ψ is free, matrix Ψ is definite positive: it is the matrix of the quadratic form associated to the L^2-norm, with respect to the basis ψ. Thus, it implies that $(y_{k+p}-y_k)$ tends to 0 (for any norm on $\mathbb{R}^n$) as p tends to infinity. As $\mathbb{R}^n$ is complete, the sequence $(y_k)_{k\geq 0}$ is converging towards $y \in (\mathbb{R}^+)^n$. It remains to show that $q^*(\cdot) = q(\cdot,y)$, which is a direct consequence of the triangular inequality for $\mathcal{W}_2$. Hence, any limit of a sequence of elements of A can be written as $q(\cdot,y)$ with $y \in (\mathbb{R}^+)^n$. This is exactly the convex cone spanned by the ψ_i for $i \in \{1,\ldots,n\}$. $\square$

A direct consequence of the previous lemma is the following:

Corollary 2 *Let the assumptions of Lemma 3 prevail. Set, for $U \subseteq (\mathbb{R}_*^+)^n$, $B := \{q(\cdot,y) \mid y \in U\}$. Then, if U is closed (respectively, compact) the set B is closed (respectively, compact) for $\mathcal{W}_2$.*

We now turn our attention to investigate the form of the derivatives of the c.d.f. that will be used to define the dynamics of the c.d.f. in Section 3.

2.2 Derivatives of the c.d.f

Our purpose is to express the derivatives of the c.d.f. $F(\cdot,y)$, as given in Definition 2, in terms of y, ψ and F. These results will be at the core of the expression of the stochastic evolution of the c.d.f. in Section 3. Indeed, in order to apply Itō calculus, we need to obtain the expression of the first and second order derivatives.

Corollary 1 provides the differentiability of F. In order to clarify the notations of partial derivatives F, we set, $\dot{F}_x(x,y) := \frac{\partial F}{\partial x}(x,y)$, the partial derivative with respect to x and for $y = (y_i)_{1\le i\le n}$, $\dot{F}_i(x,y) := \frac{\partial F}{\partial y_i}(x,y)$, the partial derivative with respect to the i^{th} component of the vector y, with $i \in \{1,\ldots,n\}$. Accordingly, we set $\ddot{F}_{x,j}(x)$, with $j \in \{1,\ldots,n\}$ and $\ddot{F}_{i,j}(x)$, with $(i,j) \in \{1,\ldots,n\}^2$, the second order derivatives. Concerning functions on the real line, such as the ψ_i, $i \in \{1,\ldots,n\}$, the successive derivatives will be denoted by ψ_i' and ψ_i''.

Proposition 3 *We have the two following derivatives.*

$$\dot{F}_x(x,y) = \frac{1}{\langle y, \psi'(F(x,y))\rangle} \tag{3}$$

$$\dot{F}_i(x,y) = -\frac{\psi_i(F(x,y))}{\Big\langle y,\ \psi'(F(x,y))\Big\rangle} \tag{4}$$

Proof: As seen in the proof of Proposition 1, the mapping $F(\cdot,y)$ is the inverse of $q(\cdot,y)$ and this relation writes, for any $\varepsilon \in [0,1)$,

$$F(q(\varepsilon,y),y) = \varepsilon \tag{5}$$

By derivation of (5) with respect to y_i, $i \in \{1,\ldots,n\}$, we find out

$$\psi_i(\varepsilon)\dot{F}_x(q(\varepsilon,y),y) + \dot{F}_i(q(\varepsilon,y),y) = 0 \tag{6}$$

By derivation of (5) with respect to ε, we find out

$$\langle y, \psi'(\varepsilon)\rangle\, \dot{F}_x(q(\varepsilon,y),y) = 1 \tag{7}$$

Substituting (7) in (6) yields, for any $i \in \{1,\ldots,n\}$,

$$\dot{F}_i(q(\varepsilon,y),y) = \frac{-\psi_i(\varepsilon)}{\langle y, \psi'(\varepsilon)\rangle} \tag{8}$$

We recall that, if $x = q(\varepsilon,y)$, then $\varepsilon = F(x,y)$. Thus, Equation (8) gives us a formulation of the first order derivatives of F with respect to y_i, $i \in \{1,\ldots,n\}$, as a function of $F(x,y)$, ψ and y. $\square$

Now, let us turn to the second order derivatives with similar arguments used in Proposition 3.

Proposition 4 *We have the following second derivatives.*

$$\ddot{F}_{xx}(x,y) = -\frac{\langle y, \psi''(F(x,y))\rangle}{\left(\langle y,\, \psi'(F(x,y))\rangle\right)^3} \tag{9}$$

$$\ddot{F}_{x,i}(x,y) = \psi_i\left((F(x,y)\right)\frac{\langle y, \psi''(F(x,y))\rangle}{\left(\langle y,\, \psi'(F(x,y))\rangle\right)^3} - \frac{\psi_i'\left((F(x,y)\right)}{\left(\langle y,\, \psi'(F(x,y))\rangle\right)^2} \tag{10}$$

$$\begin{aligned}\ddot{F}_{i,j}(x,y) &= \frac{\psi_i'\left((F(x,y)\right)\psi_j\left((F(x,y)\right) + \psi_j'\left((F(x,y)\right)\psi_i\left((F(x,y)\right)}{\left(\langle y,\, \psi'(F(x,y))\rangle\right)^2} \\ &\quad -\psi_i\left((F(x,y)\right)\psi_j\left((F(x,y)\right)\frac{\langle y, \psi''\left((F(x,y)\right)\rangle}{\left(\langle y,\, \psi'(F(x,y))\rangle\right)^3}\end{aligned} \tag{11}$$

Now, let us state a result that will prove itself useful to describe the form of the volatility in Section 3.

Lemma 4 *For any $y \in (\mathbb{R}_*^+)^n$, we have the following limit*

$$\lim_{x\to+\infty} \dot{F}_i(x,y) = 0$$

for $i \in \{1,\dots,n\}$ and this convergence is uniform with respect to y on any compact set.

Proof: The existence of such a c.d.f and its differentiability has been proved in Proposition 1. As q is increasing in y_i, $i \in \{1,\dots,n\}$, F is decreasing in this variable. This implies that the convergence of $F(x, y_1,\dots,y_n)$ towards 1, when x tends to $+\infty$, is uniform in y_i, on interval of the form $(0, \overline{y}_i]$. Without loss of generality, let us consider the case $i = 1$ for some given $(y_2,\dots,y_n)$. For any $y_1 \in (\eta, \overline{y}_1 - \eta]$, and for any $0 < |h| \le \eta$, set

$$g(x,h) := \frac{1}{h}\Big[F(x, y_1 + h, y_2, \dots, y_n) - F(x, y_1, \dots, y_n)\Big]$$

The function g converges to 0 when x tends to $+\infty$, uniformly in h. It also converges to $\dot{F}_{i+1}(x, y_1,\dots,y_n)$ when h tends to 0. Therefore, we can permute the limits over x and h to obtain the result. $\square$

Before turning to the study of dynamic distributions, i.e. stochastic processes with values in $\mathcal{D}_0$, let us analyse the effect of the mixture of percentiles on the risk of the underlying distribution.

2.3 Mixture and stochastic order

Let us consider a basis of percentiles $(\psi_i)_{1\leq i\leq n}$. Let $\left(y^{(1)}\right)$ and $\left(y^{(2)}\right)$ be in $(\mathbb{R}_*^+)^n$. If, for any $i \in \{1,\dots,n\}$, $y_i^{(1)} \geq y_i^{(2)}$, then $q\left(\cdot, y^{(1)}\right) \succeq q\left(\cdot, y^{(2)}\right)$, i.e. the distribution associated to $y^{(1)}$ dominates the distribution associated to $y^{(2)}$ for the first order stochastic dominance. Now, let us investigate the case of the second order stochastic dominance and answer the question: can we characterize a mixture which dominates the initial distribution with respect to this order? First, let us recall that $(f_1, F_1, q_1) \in \mathcal{D}_0$ dominates (f_2, F_2, q_2) for the second order stochastic dominance if, for any $x \geq 0$,

$$\int_0^x F_1(x)dx \leq \int_0^x F_2(x)dx,$$

with a strict inequality for at least one value of x.

Lemma 5 *Let $q \in \mathcal{D}_0$ and $\psi \in \mathcal{P}_n(q)$. There exists $h \in \mathbb{R}^n$ and $\widehat{\delta} > 0$, such that, for any $\delta \in \left(0, \widehat{\delta}\right]$, $q(\cdot, e+\delta h) \succeq q(\cdot)$, if, and only if,*

$$\forall \varepsilon \in [0,1], \quad \sum_{i=1}^n h_i \int_0^\varepsilon \psi_i(u)du \geq 0, \tag{12}$$

and the inequality is strict for one ε.

Proof: If $q(\cdot, e+\delta h) \succeq q(\cdot)$ for any $\delta \in \left(0, \widehat{\delta}\right]$, we have

$$\forall x \geq 0, \quad \int_0^x \Big[F(u, e+\delta h) - F(u)\Big] du \leq 0\,.$$

If we divide the two members of the previous inequality by $\delta > 0$, and let it goes to 0, we obtain

$$\forall x \geq 0, \quad \int_0^x \sum_{i=1}^n \dot{F}_i(u,e) h_i \, du \leq 0.$$

By Equations (3) and (4), it yields

$$\forall x \geq 0, \; \sum_{i=1}^n h_i \int_0^x \psi_i(F(u,e)) f(u,e) du = \sum_{i=1}^n h_i \int_0^{F(x)} \psi_i(v) dv \geq 0.$$

Hence, the sufficient condition. For the necessary condition, we write a second order Taylor expansion: for any $u \in [0,x]$, there exists $\theta(u) \in [0,\delta]$, such that

$$\begin{aligned}\frac{1}{\delta}\int_0^x [F(u,e+\delta h) - F(u)]\,du = &\int_0^x \sum_{i=1}^n \dot{F}_i(u,e)h_i\,du \\ &+ \delta \int_0^x \sum_{i,j} \ddot{F}_{i,i}(u,e+\theta(u)h)h_i h_j\,du\end{aligned}$$

As F is twice continuously derivable, its second order derivatives are uniformly bounded on $[0,1]\times[e,e+\delta h]$. Therefore, there exists $K(h)>0$ such that

$$\frac{1}{\delta}\int_0^x [F(u,e+\delta h) - F(u)]\,du \leq \int_0^x \sum_{i=1}^n \dot{F}_i(u,e)h_i du + \delta\,K(h)$$

This implies that, for δ small enough $\int_0^x [F(u,e+\delta h) - F(u)]\,du \leq 0$ and the inequality is strict if $x \neq 0$. □

According, to Lemma 5, a deviation from the initial distribution q, with increasing second order stochastic dominance, can decrease some mixture coefficients as long as condition (12) prevails. Let us consider an example:

Example 1 *Let us consider a basis of two log-normal percentiles, i.e.* $\psi_i(\varepsilon) = e^{\sigma_i \Phi^{-1}(\varepsilon)}$, $i \in \{1,\ldots,n\}$, *where* Φ *is the c.d.f. of the standard normal law,* $0 < \sigma_1 < \sigma_2$. *We have* $\int_0^\varepsilon \psi_i(u)du = \Phi\left(\Phi^{-1}(\varepsilon) - \sigma_i\right) \times e^{\frac{\sigma_i^2}{2}}$. *Assuming* $h_2 < 0 < h_1$, *condition (12) writes, for any* $\varepsilon \in (0,1)$,

$$\frac{h_1}{|h_2|} \geq \frac{\Phi\left(\Phi^{-1}(\varepsilon) - \sigma_2\right)}{\Phi\left(\Phi^{-1}(\varepsilon) - \sigma_1\right)}\, e^{\frac{\sigma_2^2-\sigma_1^2}{2}}$$

The right hand side of the inequality is uniformly bounded by $e^{\frac{\sigma_2^2-\sigma_1^2}{2}}$ *on* $[0,1)$ *since the numerator and denominator are equivalent in* 0^+. *Hence, the existence of* h_1 *and* h_2. *Besides, set* $h_1 = 1$ *and* $h_2 = -\exp\left(\frac{(\sigma_1^2-\sigma_2^2)}{2}\right)$. *Condition* (12) *is satisfied and the distributions* $f(\cdot,e+h)$ *and* $f(\cdot)$ *have the same expectation. In this case, the two percentile functions do cross each over. The modified percentile function,* $q(\cdot,e+h)$ *is above* $q(\cdot)$ *on* $[0,\varepsilon_0)$ *and below* $q(\cdot)$ *on* $(\varepsilon_0,1)$, *for a unique* $\varepsilon_0 > 0$. *It means that, although* $q(\cdot,e+h)$ *has the same expectation than* $q(\cdot)$, *it puts more weight on both low and large values: it is a so called mean-preserving spread, see Levy* [18].

3 Dynamic distributions

We now turn to the dynamic extension of the previous analysis in order to model the evolution of distributions.

Consider some $f \in \mathcal{D}_0$, with c.d.f F and percentile function q. We also set-out some $\psi \in \mathcal{P}_n(q)$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space. The operator $\mathbb{E}[\cdot]$ denotes the expectation under $\mathbb{P}$. We consider a n-dimensional Brownian motion $\{W(t) := (W_i(t))_{1\leq i\leq n}\}_{t\geq 0}$, centred, with reduced volatilities, but a non-degenerated correlation matrix $C = [c_{i,j}]_{1\leq i,j\leq n}$ on $(\Omega, \mathcal{F}, \mathbb{P})$. The natural filtration of W will be denoted by $\mathbb{F} := \{\mathcal{F}_t\}_{t\geq 0}$. We shall deal with a $\mathbb{F}$-adapted diffusion process, denoted by $\left\{Y(t) := (Y_i(t))_{1\leq i\leq n}\right\}_{t\geq 0}$, with values in $(\mathbb{R}_*^+)^n$. More precisely, set, for $i \in \{1, \dots, n\}$, $Y_i(0) = 1$ and

$$dY_i(t) = \mu_i(Y_i)dt + \sigma_i(Y_i)dW_i(t)$$

where μ_i and σ_i are Borelian mappings. We will work in the remaining of the paper under the following assumption.

Assumption 2 *For any $t \geq 0$, the process $\{Y(t)\}_{t\geq 0}$ is a Markov diffusion, square integrable, with positive values $\mathbb{P}$-a.s. Moreover, there exists $K > 0$ such that, for any $y > 0$ and $i \in \{1, \dots, n\}$, we have $\mu_i(y)^2 + \sigma_i(y)^2 \leq K \times (1 + y^2)$.*

Let us provide two examples of stochastic processes satisfying Assumption 2.

Example 2 (Log-normal diffusion) *Set, for any $i \in \{1, \dots, n\}$, $\sigma_i(x) = \overline{\sigma}_i \times x$, with $\overline{\sigma}_i > 0$ and $\mu_i(x) = \overline{\mu}_i \times x$. We have*

$$Y_i(t) = \exp\left[\left(\overline{\mu}_i - \frac{\overline{\sigma}_i^2}{2}\right)t + \overline{\sigma}_i W_i(t)\right]$$

The process Y_i is positive and square integrable.

Example 3 (Jacobi process) *Let $0 < m < \overline{\mu}_i < M$, $\overline{\sigma}_i > 0$, $\lambda_i > 0$, for any $i \in \{1, \dots, n\}$, with*

$$\frac{\overline{\sigma}_i^2}{2\lambda_i} \leq \frac{\overline{\mu}_i - m}{M - m} \leq 1 - \frac{\overline{\sigma}_i^2}{2\lambda_i}$$

We set out

$$dY_i(t) = \lambda_i(\overline{\mu}_i - Y_i(t))dt + \overline{\sigma}_i\sqrt{(M - Y_i(t))(Y_i(t) - m)}\, dW_i(t)$$

The process Y_i is a Jacobi process with values in (m, M),hence positive. It is also square integrable. See, e.g., Delbaen and Shirakawa [9] or Ackerer et al. [1]. The condition on drift and volatility functions in Assumption 2 is clearly satisfied because the drift is linear and the volatility bounded.

Let us define the distribution-valued process.

Definition 3 *We set, for any $\varepsilon \in [0,1)$, and $t \geq 0$, $\widetilde{q}(t,\varepsilon) = q(\varepsilon, Y(t))$, with $q(\cdot, y)$ defined by Equation (1). The p.d.f, respectively, c.d.f., associated to $\widetilde{q}(t,\cdot)$ is denoted by $\widetilde{f}(t,\cdot)$, respectively $\widetilde{F}(t,\cdot)$.*

The following lemma provides some boundaries on the expected Wasserstein distance between the distributions at time s and t, with $0 \leq t \leq T$. In particular, it provides the continuity with respect to time, in terms on the expected Wasserstein distance, which is a natural tool to control the stochastic evolution of probability densities. The reader can refer to Alfonsi et al. [2], for applications in the convergence of Euler schemes.

Lemma 6 *For any $T \geq 0$, with K as in Assumption 2, there exist a constant, $C_{n,T,K}$ such that, for any $0 \leq s < t \leq T$,*

$$\mathbb{E}\left[\mathcal{W}_2\big(\widetilde{q}(t,\varepsilon), \widetilde{q}(s,\varepsilon)\big)\right] \leq ||q||_2 C_{n,T,K}\sqrt{1+n}\sqrt{t-s}$$

Proof: We use Jensen inequality and Proposition 2 to state

$$\mathbb{E}\left[\mathcal{W}_2\big(\widetilde{q}(t,\varepsilon), \widetilde{q}(s,\varepsilon)\big)\right]^2 \leq \mathbb{E}\left[\mathcal{W}_2^2\big(\widetilde{q}(t,\varepsilon), \widetilde{q}(s,\varepsilon)\big)\right] \leq ||q||_2^2 \times |Y(t) - Y(s)|^2$$

Then, by Problem 3.15, p. 306, in Karatzas and Shreve [15], under Assumption 2, we have the existence of a constant $L_{n,K,T}$ such that $\mathbb{E}\left[|Y(t) - Y(s)|^2\right] \leq L_{K,T}\big(1 + |Y(0)|^2\big)(t-s)$. But $|Y(0)|^2 = n$, which yield the result. □

Now, let us turn to the explicit dynamics of the c.d.f.

Proposition 5 *For any $x > 0$, the process $\big(\widetilde{F}(t,x), Y(t)\big)_{t\geq 0}$ is a Markov diffusion, with values on $[0,1] \times (\mathbb{R}_*^+)^n$, with $\widetilde{F}(t,0) = 0$. The dynamics of $\big(\widetilde{F}(t,x)\big)_{t\geq 0}$ is given by*

$$d\widetilde{F}(t,x) = \sum_{i=1}^{n} A_i(x, Y(t))dt + \sum_{i=1}^{n} B_i(x, Y(t))dW_i(t) \tag{13}$$

where the mappings A_i *and* B_i *from* $(\mathbb{R}_*^+)^{n+1}$ *into* $\mathbb{R}$ *are given by*

$$\begin{aligned} B_i(x,y) &= \frac{-\psi_i(F(x,y))}{\langle y, \psi'(F(x,y))\rangle}\sigma_i(y_i) \\ A_i(x,y) &= \frac{-\psi_i(F(x,y))}{\langle y, \psi'(F(x,y))\rangle}\mu_i(y_i) + \frac{1}{2}(\sigma_i(y_i))^2 V_{i,i}(F(x,y),y) \\ &\quad +\sigma_i(y_i)\sum_{j>i}^{n} c_{i,j}\sigma_j(y_j) y_j V_{i,j}(F(x,y),y) \end{aligned}$$

with, for $F \in [0,1]$ *and* $y \in (\mathbb{R}_*^+)^n$,

$$V_{i,j}(F,y) := \left[\frac{\psi_i'(F)\psi_j(F) + \psi_j'(F)\psi_i(F)}{\langle y, \psi'(F)\rangle)^2} - \psi_i(F)\psi_j(F)\frac{\langle y, \psi''(F)\rangle}{(\langle y, \psi'(F)\rangle)^3}\right] \tag{14}$$

Proof: By applying Itō calculus, we have

$$\begin{aligned} d\widetilde{F}(t,x) = &\sum_{i=1}^{n} \dot{F}_i\big(x, Y(t)\big)\left[\mu_i(Y_i(t))dt + \sigma_i(Y_i(t))dW_i(t)\right] \\ &+ \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \ddot{F}_{i,j}\big(x, Y(t)\big)\sigma_i(Y_i(t))\sigma_j(Y_j(t))c_{i,j}dt \end{aligned} \tag{15}$$

Using Equations (8) and (9), we obtain, for $i \in \{1,\ldots,n\}$, the mappings A_i and B_i. $\square$

The fact that $x \mapsto \widetilde{F}(x,t)$ is a c.d.f has some implication on its volatility (as a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$), in particular, when x goes to to infinity. This property is declined into two results, depending on the type of behaviour of the volatility of Y.

Corollary 3 *Assume that*

(i) For $i \in \{1,\ldots,n\}$, $\sigma_i(x) \le \overline{\sigma}_i x$

(ii) There exists $L > 0$ *such that, for all* $x \in \mathbb{R}^+$ *and* $y \in (\mathbb{R}_*^+)^n$, $0 \le xf(x,y) \le L$

Then, for any $T > 0$ *and* $i \in \{1,\ldots,n\}$, $\lim_{x\to+\infty} \mathbb{E}\left[\int_0^T B_i^2\big(x, Y(s)\big)ds\right] = 0.$

Proof: By Problem 315 p. 306 in Karatzas and Shreve [15], we know that, under Assumption 2, the random variable $U^*_{i,T} := \max_{0\leq s\leq T} Y_i^2(s)$ is integrable. We write

$$\mathbb{E}\left[\int_0^T B_i^2\big(x, Y(s)\big)ds\right] \leq \overline{\sigma}_i^2 N^2\, \mathbb{E}\left[\mathbb{I}_{\{U^*_{i,T}\leq N\}} \int_0^T \dot{F}_i^2(x, Y(s))ds\right] + \overline{\sigma}_i^2 L^2 T\, \mathbb{P}\left[U^*_{i,T} \geq N\right]$$

Now, set $\eta > 0$ and choose N such that $\mathbb{P}\left[U^*_{i,T} \geq N\right] \leq \frac{\eta}{2\overline{\sigma}_i^2 T L^2}$. As a consequence of Lemma 4, there exists X_N such that, for all $x \geq X_N$, for any $y \in ([0,N])^n$,

$$\left|\dot{F}_i(x, Y(s))\right| \leq \frac{\eta}{2\overline{\sigma}_i^2 T}$$

Hence, $\mathbb{E}\left[\int_0^T B_i^2\big(x, Y(s)\big)ds\right] \leq \eta$, for $x \geq X_N$. □

Remark 1 *Point (ii) of Lemma 3 is satisfied, for instance, for log-normal distributions as can be seen in the proof of Proposition 6.*

Corollary 4 *Assume that, for $i \in \{1,\ldots,n\}$, the process Y_i takes its values in the finite interval (m, M). Then, for any $T > 0$ and $i \in \{1,\ldots,n\}$,*

$$\lim_{x\to+\infty} \mathbb{E}\left[\int_0^T B_i^2\big(x, Y(s)\big)ds\right] = 0$$

Proof: By Assumption 2, we have the straightforward inequality

$$\mathbb{E}\left[\int_0^T B_i^2\big(x, Y(s)\big)ds\right] \leq K(1+M^2)T\left[\sup_{y\in[m,M]^n}\{\dot{F}_i(x,y)\}\right]^2$$

As a consequence of Lemma 4, the right-hand side converges to 0 as x goes to infinity. □

We can provide a result on the regularity of the expected value of $\widetilde{F}(t,\cdot)$ and its link with the expected value of $\widetilde{f}(t,\cdot)$. For technical reasons, we will work with a basis of log-normal percentile functions. Indeed, in this setting, we can show that the derivation with respect to x and the integration with respect to $\mathbb{P}$ do permute.

Proposition 6 *Assume that, for all $i \in \{1,\ldots,n\}$, $\psi_i : \varepsilon \mapsto e^{\gamma_i \Phi^{-1}(\varepsilon)}$, with $\gamma_i > 0$, with the notations set out in Example 1. Assume that $[\langle Y(t), e\rangle]^{-1}$ is square-integrable. Then, $M(t,\cdot) := \mathbb{E}\{F(\cdot, Y(t))\}$ is the c.d.f of an element of $\mathcal{D}_0$, with p.d.f $m(t,\cdot) = \mathbb{E}\{f(\cdot, Y(t))\}$.*

Proof: First of all, it is clear that $M(t,\cdot)$ is increasing and $M(t,0) = 0$. By Lebesgue monotone convergence theorem, we also have $M(t,x) \to 1$, when $x \to +\infty$. Therefore, it defines an element of $\mathcal{D}$. In order to show that is is, actually, in $\mathcal{D}_0$, we need to show that the derivatives of $F(\cdot, Y(t))$ of order 1 and 2 are uniformly bounded. Let us start with some facts deduced from the log-normal distribution. Set $\phi := \Phi'$. We have

$$\psi_i'(\varepsilon) = \frac{\gamma_i \psi_i(\varepsilon)}{\phi\left(\Phi^{-1}(\varepsilon)\right)} \quad \text{and} \quad \psi_i''(\varepsilon) = \frac{\gamma_i^2 \psi_i(\varepsilon) + \gamma_i \psi_i(\varepsilon)\Phi^{-1}(\varepsilon)}{\left[\phi\left(\Phi^{-1}(\varepsilon)\right)\right]^2}$$

We also notice that $q(\frac{1}{2}, y) = \langle y, e\rangle$. We define $\gamma_{min} := \min\{\gamma_i \mid 1 \leq i \leq n\}$ and $\gamma_{max} := \max\{\gamma_i \mid 1 \leq i \leq n\}$ and set out, for any $z > 0$,

$$K(x,z) := \begin{cases} \Phi\left(\frac{1}{\gamma_{min}} \ln\left(\frac{x}{z}\right)\right) & \text{if } x \geq z \\ \Phi\left(\frac{1}{\gamma_{max}} \ln\left(\frac{x}{z}\right)\right) & \text{if } x < z \end{cases}$$

We obtain the following inequality $0 \leq F(x,y) \leq K(x, \langle y, e\rangle)$. This shows that, basically, the c.d.f. $F(\cdot, y)$ is below the log-normal c.d.f with the smallest volatility if $F(x,y) \geq \frac{1}{2}$ and below the log-normal c.d.f with the largest volatility if $F(x,y) < \frac{1}{2}$.

Equations (7) and (9) and the calculations above yield the following inequality:

$$0 \leq f(x,y) \leq \frac{\phi\left(\Phi^{-1}(F(x,y))\right)}{\gamma_{min} x}$$

Using the boundary $K(x, \langle y, e\rangle)$, we obtain

$$0 \leq f(x,y) \leq \frac{1}{\gamma_{min} x} \begin{cases} \phi\left(\frac{1}{\gamma_{min}} \ln\left(\frac{x}{\langle y,e\rangle}\right)\right) & \text{if } x \geq \langle y, e\rangle \\ \phi\left(\frac{1}{\gamma_{max}} \ln\left(\frac{x}{\langle y,e\rangle}\right)\right) & \text{if } x < \langle y, e\rangle \end{cases}$$

In the left hand side we recognize the log-normal densities with mean $\ln(\langle y, e\rangle)$ and volatility γ_{min} and γ_{max}, respectively. By using the mode of these densities, we obtain the following inequality:

$$0 \leq f(x,y) \leq H(y) := \frac{e^{\frac{\gamma_{max}^2}{2}}}{\gamma_{min}\sqrt{2\pi}\langle y, e\rangle}$$

Hence, $f(x, Y(t)) = \dot{F}_x(x, Y(t))$ is uniformly bounded in x by some $\mathbb{P}$-integrable random variable $H(Y(t))$. Hence, $M(t,\cdot)$ is derivable and its derivative is $m(t,\cdot)$.

We can go one step further and compute the second order derivative. For $x < \langle y, e\rangle$, we have, for some constant $c_1 > 0$,

$$|\ddot{F}_{xx}(x,y)| \leq \frac{1}{\gamma_{min}^3}\phi\left(\frac{1}{\gamma_{max}}\ln\left(\frac{x}{\langle y,e\rangle}\right)\right) \times \frac{\gamma_{max}x - x\ln\left(\frac{x}{\langle y,e\rangle}\right)}{x^3}$$

The maximum of the function in right hand side is achieved at

$$\langle y,e\rangle \exp\left[-\gamma_{max}^2 + \frac{\gamma_{max}}{2} - \sqrt{\gamma_{max}^4 + \frac{3}{2}\gamma_{max}^3 + \frac{5}{4}\gamma_{max}^2}\right] < \langle y,e\rangle$$

At this point, the maximum is of the form $v(\gamma)(\langle y,e\rangle)^{-2}$, hence $\mathbb{P}$-integrable, by assumption. The same argument applies to the case $x > \langle y,e\rangle$. This provides the second order derivative as the expectation of $\ddot{F}_{xx}(x, Y(t))$. $\square$

Now, we can characterize the form of the probability density of $\widetilde{q}(t,\cdot)$ and $\widetilde{F}(t,\cdot)$. Let us introduce the following replacing function $\mathcal{R} : \mathbb{R}^n \times \{1,\ldots,n\} \times \mathbb{R} \to \mathbb{R}^n$: for any $r \in \mathbb{R}^n$, $i \in \{1,\ldots,n\}$, and $u \in \mathbb{R}$, $\mathcal{R}(r;i,u) = z$ where $z_j = r_j\mathbb{I}_{j\neq i} + u\mathbb{I}_{j=i}$. That is the replacing function $\mathcal{R}$ transform the original vector r in the one in which coordinate i has been replaced by u.

Proposition 7 *Let $\mathcal{K}_t : (\mathbb{R}_*^+)^n \to \mathbb{R}^+$ be the density $Y(t)$. Let $\eta(t,\varepsilon,x)$ and $\rho(t,\varepsilon,x)$ denote, respectively, the derivative of the percentile and the p.d.f., at fixed time t. Assume also that there exists $i^\star \in \{1,\ldots,n\}$ such that $\psi_{i^\star} \in \mathcal{D}_0$.*

Then, the derivative function satisfies, for any $x \geq 0$ and $\varepsilon > 0$,

$$\eta(t,\varepsilon,x) = \int_{(\mathbb{R}_*^+)^{n-1}} \mathcal{K}_t\left[\mathcal{R}\left(z; i^\star, \frac{x - \sum_{i\neq i^\star}\psi_i(\varepsilon)z_i}{\psi_{i^\star}(\varepsilon)}\right)\right]\prod_{i\neq i^\star} dz_i$$

Proof: The result is obtained by a change of variable $z_{i^\star} = \langle \psi(\varepsilon), y\rangle$, $z_i = y_i$, $i \neq i^\star$, and the fact that $\psi_{i^\star}(\varepsilon) > 0$ as soon as $\varepsilon > 0$. $\square$

In order to construct the functional Bollinger bands, we need to be able to define a confidence interval for $F(t,x)$, seen as a random variable on $(\Omega,\mathcal{F},\mathbb{P})$. For this purpose, we use the following functions:

Definition 4 *For any $\eta \in (0,1)$, $t > 0$ and $x > 0$, $H_\eta(t,x)$ is the solution (in $H \in [0,1)$) of $\mathbb{P}\left[\widetilde{F}(t,x) \leq H\right] = \eta$. For any $\eta \in (0,1)$, $t > 0$ and $\varepsilon \in (0,1)$, $I_\eta(t,x)$ is the solution in $I \geq 0$ of $\mathbb{P}\left[\widetilde{q}(t,\varepsilon) \leq I\right] = \eta$*

The functions $x \mapsto H_\eta(t,x)$ and $\varepsilon \mapsto I_{1-\eta}(t,\varepsilon)$ are, respectively, a (non random) c.d.f and a (non-random) percentile function, associated to the same probability distribution. This is, basically, what is proved in next proposition. These functions can be used to define the confidence interval on $\widetilde{F}(t,x)$ or $\widetilde{q}(t,\varepsilon)$, respectively.

Proposition 8 *Let the assumption of Theorem 7 prevails. Assume that Y takes its values in a open (possibly not bounded) subset of $(\mathbb{R}_*^+)^n$, with unattainable boundaries, and, for every non empty ball B in this subset $\mathbb{P}[Y \in B] > 0$. Then, the mapping $\varepsilon \mapsto I_\eta(t,\varepsilon)$ is derivable, increasing, with $\lim_{\varepsilon\to 0+} I_\eta(t,\varepsilon) = 0$ and $\lim_{\varepsilon\to 1-} I_\eta(t,\varepsilon) = +\infty$*

Proof: First, let us denote by $\mathcal{K}_t(y)$ the density of $Y(t)$. Set, for any $\varepsilon \in [0,1)$, $h \geq 0$,

$$Q(\varepsilon,h) := \mathbb{P}\left[q(\varepsilon,Y(t)) \leq h\right] = \int_{y\in(\mathbb{R}_*^+)^d} \mathbb{I}_{\{q(\varepsilon,y)\leq h\}}\mathcal{K}_t(y)dy$$

First, if $\varepsilon > 0$, the mapping $h \mapsto Q(\varepsilon,h)$ is increasing, with $Q(\varepsilon,0) = 0$ and $\lim_{h\to+\infty} Q(\varepsilon,h) = 1$. The fact that $Q(\varepsilon,0) = 0$ is clear, because $q(\varepsilon,y) > 0$ if and only if $\varepsilon > 0$. It is increasing because, if $0 \leq h_1 < h_2$, $\{q(\varepsilon,y) \leq h_1\} \neq \{q(\varepsilon,y) \leq h_2\}$. Indeed, as at least one of the ψ_i is not equal to 0 (because $q(\varepsilon) > 0$), it is always possible to increase $q(\varepsilon,y)$ by increasing the corresponding component of y. Therefore, there is an open ball included in the second interval and not in the first. By assumption of this proposition, it is given a positive weight by $K(t,\cdot)$. The limit when h goes to infinity is a consequence of the Beppo-Levi monotone convergence theorem. Besides, the mapping $(\varepsilon,h) \mapsto Q(\varepsilon,h)$ is also derivable. This is a consequence of theorem 7.

The mapping $Q(\varepsilon,\cdot)$ defines a (continuous) bijection from $[0,+\infty)$ onto $[0,1)$. Hence, $\eta \mapsto I_\eta(t,\varepsilon)$ is well defined and continuous with respect to both variable. It is increasing because $q(\cdot,y)$ is increasing. Set $Q_{max}(h) := \mathbb{P}(\langle Y(t),e\rangle \leq h)$ and $Q_{min}(h) := \mathbb{P}\big[\min_i\{Y_i(t)\} \leq h\big]$. From the inequality $\min_i\{Y_i(t)\}\times q(\varepsilon) \leq q(\varepsilon,Y(t)) \leq \langle Y(t),e\rangle \times q(\varepsilon)$, we deduce that

$$Q_{max}\left(\frac{I_\eta(t,\varepsilon)}{q(\varepsilon)}\right) \leq \eta \leq Q_{min}\left(\frac{I_\eta(t,\varepsilon)}{q(\varepsilon)}\right)$$

The first inequality yields the limit when ε goes to 0^+ and the second the limit when ε goes to 1^-. By definition of $F(x,y)$, we have $q(\varepsilon,y) \leq h \Rightarrow \varepsilon \leq F(h,y)$. Hence, the inverse of $I_\eta(t,\cdot)$ is $H_{1-\delta}(t,\cdot)$. $\square$

At this stage, we can define the functional Bollinger bands by

Definition 5 *Set $\eta \in \left(0, \frac{1}{2}\right)$, the functional Bollinger bands, for the percentile, at horizon t are defined by the mappings $\varepsilon \mapsto I_\eta(t, \varepsilon)$ (the lower band) and $\varepsilon \mapsto I_{1-\eta}(t, \varepsilon)$ (the upper band). The functional Bollinger bands, for the c.d.f., at horizon t are defined by the mappings $x \mapsto H_\eta(t, x)$ (the lower band) and $x \mapsto H_{1-\eta}(t, x)$ (the upper band). These mappings are elements of $\mathcal{D}$, as proved in Proposition 8.*

An interesting property of this approach is that the lower and upper bands for the percentile (respectively, the c.d.f) are percentiles (resp., c.d.f). Moreover, the percentile upper band is the inverse of the c.d.f lower band (and symmetrically for the other bands). Hence, the percentile $\widetilde{F}(t, \cdot)$ can be compared to the lower and upper bands $H_\eta(t, \cdot)$ and $H_{1-\eta}(t, \cdot)$ through the concept of second order stochastic dominance. If $\widetilde{F}(t, \cdot)$ dominates $H_\eta(t, \cdot)$ for the second order stochastic dominance, it implies that the distribution at time t involves more risk than what was expected at time 0, with a confidence level of $2\,\eta$. This situation can occur even if the averages of the two distributions are the same. Such a situation could be a good trading signal. This is what we want to illustrate in the following subsection.

4 Application to Credit Indices

In this section, we consider the evolution of the spreads of all the 125 components of a the iTraxx Europe (with maturity 5 years). Although this index is rolled every 6 months, we will consider - for a purely illustrative purpose - a fixed composition, corresponding to one given series of this index.

Given these samples, we can calculate, at each date t_j, with $t_1 = 0 < \cdots < t_m$, the empirical percentiles function, denoted by $\overline{q}(t_j, \cdot)$.

Consider n elements of $\mathcal{D}$, represented by their percentile functions $v := (v_k)_{1 \le k \le n}$, assuming that one of those is in $\mathcal{D}_0$. We also assume that these elements are linearly independent. As a typical example, we consider $v_k(\varepsilon) := \exp(\gamma_k \Phi^{-1}(\varepsilon))$, where Φ is the c.d.f of the standard normal law and the γ_k are positive, two by two distinct real numbers. In this case, each v_k is the percentile of a log-normal law.

A each date t_j, we will perform a constrained regression of $\overline{q}(t_j, \cdot)$ on v, using the Wasserstein distance. A short calculation shows that it amounts solving the following (constrained) quadratic program in $\mathbb{R}^n$:

$$(\mathcal{R})_j \ : \left\{ \begin{array}{l} \min_{z \in \mathbb{R}^n} \left[z^T \cdot \Psi \cdot z - 2Q(j)^T \cdot z \right] \\ \text{s.t. } \forall k \in \{1, \ldots, p\} \ z_k \geq \delta > 0 \end{array} \right.$$

where Ψ is defined in the proof of Lemma 3, and $Q(j)$ is a vector of $\mathbb{R}^n$ with $Q_k(j) = \int_0^1 \psi_k(\varepsilon)\overline{q}(t_j,\varepsilon)d\varepsilon$. As in Lemma 3, the matrix is invertible because the family of log-normal percentiles is free as soon as the volatilities are distinct.

In many practical cases that we will analyse in this section, the constraints of program $(\mathcal{R})_j$ are not binding and this program amounts to a classic least-square method.

Let us denote by $\widehat{z}^{(j)}$ the solution of program $(\mathcal{R})_j$, and set $\widehat{Y}_k(t_j) := \widehat{z}_k^{(j)} \times (\widehat{z}_k^{(0)})^{-1}$ and $\psi_k(\cdot) = \widehat{z}_k^{(0)} \times v_k(\cdot)$, $k \in \{1,\ldots,n\}$. By definition, ψ is an element of $\mathcal{P}_p(\widehat{q}_0)$ where $\widehat{q}_0 := \langle \widehat{z}^{(0)}, v(\cdot)\rangle \in \mathcal{D}_0$. Hence, we are in the framework developed above, and we have obtained a sample path for Y, on the basis of which we can calibrate the parameters defined in Section 3. For each date t_j, we have $\widehat{q}_j(\cdot) = \langle v(\cdot), \widehat{z}^{(j)}\rangle$.

Once the parameters of the diffusion Y are calibrated, according to Definition 3, we have the full dynamics of $\widetilde{q}(t,\cdot)$.

In the following examples, we shall develop our method based on the function Bollinger bands, defined in 5. Let us emphasize the analogy with the standard, one dimensional case. For a single valued process, in the trading strategies area, the Bollinger bands method consists, basically, to look at a confidence interval on the price of a financial instrument, based on historical trailing volatility and average. Hence, implicitly a Gaussian case. Trading signals are triggered when the price crosses the bands. See Kaufman [17], for definition and use in trading strategies, or Bernis and Scotti [6] for applications to credit indices in the context of non-linear filtering. In our setting, the equivalent of crossing the upper (respectively, lower) band will be the case where the current distribution $F(t,\cdot)$ dominates the upper (resp., lower) functional band $H_\eta(t,\cdot)$ (resp. $H_{1-\eta}(t,\cdot)$) for the second order stochastic dominance. More precisely, we denote by $F(t_j,\cdot)$ the c.d.f. at time t_j, stemming from the fit of the empirical percentiles at this date. Given the calibration of the diffusion parameters for Y on the sample $0 \le t_i \le t_m$, we calculate at time t_m, the lower and upper functional bands, at horizon t_{m+l}, $H_\eta(t_{m+l}-t_m,\cdot)$, $H_{1-\eta}(t_{m+l}-t_m,\cdot)$, with $l>0$ some fixed horizon. Then, we can compare $\overline{F}(t_{m+l}-t_m,\cdot)$ to the bands. This formulation can be transposed to the lower and upper functional bands on the percentiles.

As an example, we display in Figure 1 the upper and lower bands on July the 13^{th} 2015 (t_{m+l}), as well as the c.d.f at this date. We take $n=2$, $\gamma_1 = 25\%$ and $\gamma_2 = 85\%$. The dynamics of Y_1, Y_2 is assumed to be log-normal, with no drift. Calibration on market data (1 year) yields $\overline{\sigma}_1 = 29\%$, $\overline{\sigma_2} = 80\%$ and a correlation of some 23%. The lower and upper bands are computed using a normal approximation over the last 5 business days, according to formula (13). It can be interesting to observe that the drift has a second order effect in this example. At this date, the c.d.f dominates the upper band, for the SOSD, which means that the distribution is significantly less risky than expected. In the same time, the average spread on July the 13^{th} 2015 (t_m) is around 75 bps (a rather high level over the last weeks) and is about 68 bps at t_{m+l}. We can reasonably expect that the index average spread will keep on tightening over the next few days: 5 business days later, it is around 63 bps.

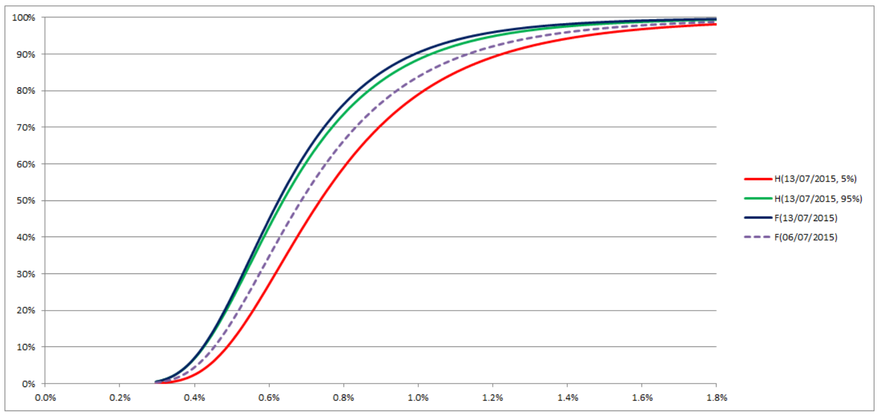

Figure 1: **Lower (5%) and upper (95%) bands, c.d.f on July the 13^{th} 2015 and July the 6^{th} 2015.**

We propose, for a deeper understanding of this example, to have a look at the form of the volatility functions $B_i(\cdot, Y(t_m))$, $1 \leq i \leq 2$. The two mappings are represented in Figure 2, in the context of Figure 1. The volatility function $B_1(\cdot, Y(t_m))$ is smaller than $B_2(\cdot, Y(t_m))$ for average and large values of x. This is due to the fact that the mapping ψ_2, associated to a large value of γ_2, mainly controls the extreme percentiles values. This basis function requires more volatility stemming from Y_2, in order to fit the percentiles in case of turmoil. In these periods, the extreme percentiles tend to increase sharply, showing some decorrelation from the lower percentiles. The same effect is captured by B_2, which remains larger than B_1 for large values of x. However, as given by Corollary 3, both functions tends to 0 as x goes to $+\infty$.

It may be interesting to investigate a criterion less restrictive than SOSD. For instance, as displayed in Figure 3, on February the 1^{st} 2016 the c.d.f began to cross the lower band (even if not dominating it for SOSD). The average spread at t_m is close to its level at t_{m+l}: respectively, 111 and 112 bps. However, the band crossing detects the increase of the risk in the index distribution: 5 days later the average spread is around 130 bps. The c.d.f. dominates the lower band in terms of SOSD shortly after this date, but when it occurs most part of the spread widening has already occurred.

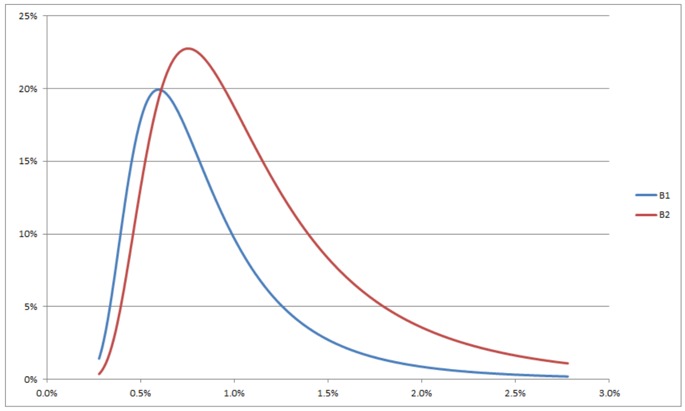

Figure 2: **Volatility functions** $B_i(\cdot, Y(t_m))$**,** $1 \leq i \leq 2$, where t_m is July the 13^{th} 2015.

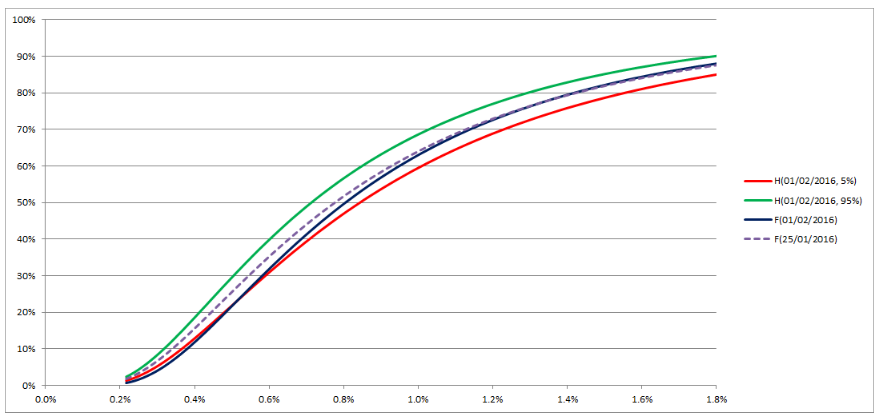

Figure 3: **Lower** (5%) **and upper** (95%) **bands, c.d.f on July the** 13^{th} **2015 and July the** 6^{th} **2015**.

References

[1] Ackerer, D., Filipović, D. and Pulido, S. (2016) The Jacobi stochastic volatility model *Swiss Finance Institute Research Paper* pp. 16-35.

[2] Alfonsi, A., Jourdain, B. and Kohatsu-Higa, A. (2014) Pathwise Optimal Transport Bounds between a One-Dimensional Diffusion and its Euler Scheme *The Annals of Applied Probability*, vol. 24(3), pp. 1049-1080.

[3] Bass, P. (2013) Optimal transportation with infinitely many marginals. *J. Funct. Anal.*, vol. 264(4), pp.947-963.

[4] Bellomo, N. and Pistone, G. (1980), Time-Evolution of the Probability Density Under the Action of a Deterministic Dynamical System. *Journal of Mathematical Analysis and Applications*, vol. 77, pp. 215-224.

[5] Bernis, G., Carassus, L., Docq, G., and Scotti, S. (2015), Optimal credit allocation under regime uncertainty with sensitivity analysis. *International Journal of Theoretical and Applied Finance*, 18(01), 1550002.

[6] Bernis, G. and Scotti, S. (2015), Trend Detection under Erroneous Observations: Application to Quantitative Finance Strategies. *Journal of Investment Strategies*, vol. 4(3), pp. 83-104.

[7] Callegaro, G.; Gaigi, M.; Scotti, S. and Sgarra C. (2016), Optimal Investment in Markets with Over and Under-Reaction to Information, *Mathematics and Financial Economics*, forthcoming.

[8] Dai, M.; Yang, Z.; Zhang, Q. and Zhu Q.J. (2016), Optimal Trend Following Trading Rules, *Mathematics of Operations Research*, vol. 41-2, pp. 626-642.

[9] Delbaen, F. and Shirakawa, H. (2002) An Interest Rate Model with Upper and Lower Bounds, *Asia-Pacific Financial Markets*, vol. 9, pp. 191-209.

[10] Di Persio, L. and Honchar, O. (2016), Artificial Neural Networks Architectures for Stock Price Prediction: Comparisons and Applications, working paper.

[11] Ferguson, T.S. (1973), A Bayesian analysis of some non-parametric problems. *Annals of Statistics*, vol. 1, pp. 209-230.

[12] Fernholz, E.R. (2001) Equity portfolios generated by ranked market weights. *Finance and Stochastics*. 5, pp. 469-486.

[13] Fernholz, E.R.; Kardaras C. and Karatzas I. (2005), Diversity and arbitrage in models of equity markets. *Finance and Stochastics*, volume 9, pp. 1-27.

[14] Gouriéroux, C. and Jasiak, J. (2008) Dynamic quantile models. *Journal of Econometrics*, vol. 147, pp. 198-205.

[15] Karatzas, I. and Shreve, S.E. (1991) Brownian Motion and Stochastic Calculus. 2nd ed. Springer-Verlag, Berlin.

[16] Karvanen, J. (2006) Estimation of quantile mixtures via L-moments and trimmed L- moments. *Computational Statistics and Data Analysis*, vol. 51, pp. 947-959.

[17] Kaufman, P.J. (2004): *New Trading Systems and Methods.*Fourth Edition, John Wiley & Sons, Inc.

[18] Levy, H. (1992), Stochastic dominance and expected utility: survey and analysis. *Management Science*, vol. 38, pp. 555–593.

[19] Quirk J.P. and Saposnik, R. (1962). Admissibility and measurable utility functions. *Review of Economic Studies*, vol. 29, pp. 140-146.

[20] Shao, J. (2011) A new probability measure-valued stochastic process with Ferguson-Dirichlet process as reversible measure. Electronic Journal Of Probability, vol. 16, pp. 271-292.

[21] Sillitto, G. (1969) Derivation of approximants to the inverse distribution function of a continuous univariate population from the order statistics of the sample. *Biometrika*, vol. 56, pp. 641-650.

[22] Vallander, S.S. (1973), Calculation of the Wasserstein distance between probability distributions on the line. *Teor. Veroyatnost. i Primenen.*, vol. 18(4), pp. 824-827.

[23] Villani, C. (2003), Topics in Optimal Transportation. *Graduate Studies in Mathematics.*, vol. 58, American Society, Providence.

[24] Wagner, N. (2008) Credit Risk, Models, Derivatives and Management. Chapman & Hall Financial Mathematics Series.

[25] Zervos, M.; Johnson, T.C. and Alazemi, F. (2013) Buy-low and sell-high investment strategies. *Mathematical Finance*, vol. 23, pp. 560-578.