



HAL
open science

Coupling Index and Stocks

Benjamin Jourdain, Mohamed Sbai

► **To cite this version:**

| Benjamin Jourdain, Mohamed Sbai. Coupling Index and Stocks. 2008. hal-00350652v1

HAL Id: hal-00350652

<https://hal.science/hal-00350652v1>

Preprint submitted on 7 Jan 2009 (v1), last revised 14 Nov 2009 (v2)

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.

Coupling Index and stocks

Benjamin Jourdain¹ and Mohamed Sbai¹

Abstract

In this paper, we are interested in continuous time models in which the index level induces some feedback on the dynamics of its composing stocks. More precisely, we propose a model in which the log-returns of each stock may be decomposed into a systemic part proportional to the log-returns of the index plus an idiosyncratic part. We show that, when the number of stocks in the index is large, this model may be approximated by a local volatility model for the index and a stochastic volatility model for each stock with volatility driven by the index. We address calibration of both the limit and the original models.

Introduction

From the early eighties, when trading on stock index was introduced, quantitative finance faced the problem of efficiently pricing and hedging index options along with their underlying components. Many advances have been made for single stock modeling and a variety of solutions to escape from the very restrictive Black & Scholes model has been deeply investigated (such as local volatility models, models with jumps or stochastic volatility models). However, when the number of underlyings is large, index option pricing, or more generally basket option pricing, remains a challenge unless one simply assumes constantly correlated dynamics for the stocks. The problem then is the impossibility of fitting both the stocks and the index smiles.

We try to address this issue by making the dynamics of the stocks depend on the index. The natural fact that the volatility of the index is related to the volatilities of its underlying components has already been accounted for in the works of Avellaneda *et al.* [1] and Lee *et al.* [12]. In the first paper, the authors use a large deviation asymptotics to reconstruct the local volatility of the index from the local volatilities of the stocks. They express this dependence in terms of implied volatilities using the results of Berestecky *et al.* ([4],[3]). In the second paper, the authors reconstruct the Gram-Charlier expansion of the probability density of the index from the stocks using a moments-matching technique. Both papers consider local volatility models for the stocks and a constant correlation matrix but the generalization to stochastic volatility models or to varying correlation coefficients is not straightforward.

Another point of view is to say that the volatility of a composing stock should be related to the index level, or say to the volatility of the index, in some way. This is not astonishing since the index represents the move of the market and reflects the view of the investors on the state of the economy. Moreover, it

¹Université Paris-Est, CERMICS, Projet MathFi ENPC-INRIA-UMLV. This research benefited from the support of the "Chair Risques Financiers", Fondation du Risque.

Postal address : 6-8 av. Blaise Pascal, Cité Descartes, Champs-sur-Marne, 77455 Marne-la-Vallée Cedex 2.

E-mails : jourdain@cermics.enpc.fr and sbai@cermics.enpc.fr

is coherent with equilibrium economic models like CAPM. Following this idea, we propose a new modeling framework in which the volatility of the index and the volatilities of the stocks are related. We show that, when the number of underlying stocks tends to infinity, our model reduces to a local volatility model for the index and to a stochastic volatility model for the stocks where the stochastic volatility depends on the index level. This asymptotics is reasonable since the number of underlying stocks is usually large. As a consequence, the correlation matrix between the stocks in our model is not constant but stochastic and we show that it is coherent with empirical studies. Finally, we address calibration issues and we show that it is possible, within our framework, to fit both index and stocks smiles. The method we introduce is based on the simulation of SDEs nonlinear in the sense of McKean, and non-parametric estimation of conditional expectations.

This paper is organized as follows. In Section 1, we specify our model for the index and its composing stocks and in Section 2 we study the limiting model when the number of underlying stocks goes to infinity. Section 3 is devoted to calibration issues. Numerical results are presented in Section 4 and the conclusion is given in Section 5.

Acknowledgements: We thank Lorenzo Bergomi, Julien Guyon and all the equity quantitative research team of Societe Generale CIB for numerous fruitful discussions and for providing us with the market data.

1 Model Specification

An index is a collection of stocks that reflects the performance of a whole stock market or a specific sector of a market. It is valued as a weighted sum of the value of its underlying components. More precisely, if I_t^M stands for the value at time t of an index composed of M underlyings, then

$$I_t^M = \sum_{j=1}^M w_j S_t^{j,M}, \quad (1)$$

where $S_t^{j,M}$ is the value of the stock j at time t and the weightings $(w_j)_{j=1\dots M}$ are given constants².

Unless otherwise stated, we always work under the risk-neutral probability measure. In order to account for the influence of the index on its underlying components, we specify the following stochastic differential equations for the stocks

$$\forall j \in \{1, \dots, M\}, \quad \frac{dS_t^{j,M}}{S_t^{j,M}} = (r - \delta_j)dt + \beta_j \sigma(t, I_t^M)dB_t + \eta_j(t, S_t^{j,M})dW_t^j \quad (2)$$

where

- r is the short interest rate.
- $\delta_j \in [0, \infty[$ is the continuous dividend rate of the stock j .
- β_j is the usual beta coefficient of the stock j that quantifies the sensitivity of the stock returns to the index returns (see the seminal paper of Sharpe [16]). It is defined as $\frac{Cov(r_j, r_I)}{Var(r_I)}$ where r_j (respectively r_I) is the rate of return of the stock j (respectively of the index).

²In most cases, the weightings are either proportional to stock prices or to market capitalization (stock price \times number of shares outstanding) and they are periodically updated but, as usually assumed, we suppose that, up to maturities of the options considered, they do not evolve in time.

- $(B_t)_{t \in [0, T]}, (W_t^1)_{t \in [0, T]}, \dots, (W_t^M)_{t \in [0, T]}$ are independent Brownian motions.
- The coefficients $\sigma, \eta_1, \dots, \eta_M$ satisfy the usual Lipschitz and growth assumptions that ensure existence and strong uniqueness of the solutions (see for example Theorem 5.2.9 of [11]) :

$$(\mathcal{H}1) \quad \exists K \text{ such that } \forall (t, s_1, s_2) \in [0, T] \times \mathbb{R}^M \times \mathbb{R}^M,$$

$$\begin{aligned} \sum_{j=1}^M \left| s_1^j \sigma \left(t, \sum_{k=1}^M w_k s_1^k \right) \right| + \left| s_1^j \eta_j(t, s_1^j) \right| &\leq K (1 + |s_1|) \\ \sum_{j=1}^M \left| s_1^j \sigma \left(t, \sum_{k=1}^M w_k s_1^k \right) - s_2^j \sigma \left(t, \sum_{k=1}^M w_k s_2^k \right) \right| &\leq K |s_1 - s_2| \\ \sum_{j=1}^M \left| s_1^j \eta_j(t, s_1^j) - s_2^j \eta_j(t, s_2^j) \right| &\leq K |s_1 - s_2| \end{aligned}$$

As a consequence, the index satisfies the following stochastic differential equation :

$$dI_t^M = rI_t^M dt - \left(\sum_{j=1}^M \delta_j w_j S_t^{j, M} \right) dt + \left(\sum_{j=1}^M \beta_j w_j S_t^{j, M} \right) \sigma(t, I_t^M) dB_t + \sum_{j=1}^M w_j S_t^{j, M} \eta_j(t, S_t^{j, M}) dW_t^j \quad (3)$$

Before going any further, let us make some preliminary remarks on this framework.

- We have M coupled stochastic differential equations. The dynamics of a given stock depends on all the other stocks composing the index through the volatility term $\sigma(t, I_t^M)$.
- Accounting for the dividends is not relevant for all types of indices. Indeed, for many performance-based indices (such as the German DAX index) dividends and other events are rolled into the final value of the index.
- The cross-correlations between stocks are not constant but stochastic :

$$\rho_{ij} = \frac{\beta_i \beta_j \sigma^2(t, I_t^M)}{\sqrt{\beta_i^2 \sigma^2(t, I_t^M) + \eta_i^2(t, S_t^{i, M})} \sqrt{\beta_j^2 \sigma^2(t, I_t^M) + \eta_j^2(t, S_t^{j, M})}}$$

Note that they depend not only on the stocks but also on the index.

In a recent paper, Cizeau *et al.* [8] show that it is possible to capture the essential features of stocks cross-correlations, in particular in extreme market conditions, by a simple non-Gaussian one factor model. The authors successfully compare different empirical measures of correlation with the prediction of the following model :

$$r_j(t) = \beta_j r_I(t) + \epsilon_j(t) \quad (4)$$

where $r_j(t) = \frac{S_t^j}{S_{t-1}^j} - 1$ is the daily return of stock j , $r_I(t)$ is the daily return of the market and the residuals $\epsilon_j(t)$ are independent random variables following a fat-tailed distribution³.

Our model is in line with (4). Indeed, since the beta coefficients are usually narrowly distributed around 1, the factor $\sum_{j=1}^M \beta_j w_j S_t^{j, M}$ of $\sigma(t, I_t^M)$ in (3) is close to I_t^M . Moreover, in the next section we show that, for a large number of underlying stocks, one can neglect the term $\sum_{j=1}^M w_j S_t^j \eta_j(t, S_t^j) dW_t^j$ in the dynamics of the index. Hence, if we denote by r_j the log-return of the stock j and by r_{I^M} the log-return of the index, both on a daily basis, we will have

$$r_j = \beta_j r_{I^M} + \eta_j \Delta W^j + \text{drift},$$

³The authors have chosen a Student distribution in their numerical experiments.

where ΔW^j is an independent Gaussian noise. Consequently, in our model too, the return of a stock is decomposed into a systemic part driven by the index, which represents the market, and a residual part.

2 Asymptotics for a large number of underlying stocks

The number of underlying components of an index is usually large⁴. It is then meaningful to let M tend to infinity. Since the Brownian motions $(W^j)_{j=1\dots M}$ are independent, one can expect that their contribution to the dynamics governing the index is not significant and drop the corresponding terms in the stochastic differential equation (3) which will drastically simplify the model. The aim of this section is to quantify the error we commit by doing so.

To be specific, consider the limit candidate $(I_t)_{t \in [0, T]}$ solution of the following SDE :

$$\begin{cases} dI_t &= (r - \delta)I_t dt + \beta I_t \sigma(t, I_t) dB_t \\ I_0 &= I_0^M \end{cases} \quad (5)$$

with δ and β two constant parameters that will be discussed later.

In the following theorem, we give an upper bound for the L^{2p} -distance between $(I_t^M)_{t \in [0, T]}$ and $(I_t)_{t \in [0, T]}$ under mild assumption on the volatility coefficients :

Theorem 1 — *Let $p \in \mathbb{N}^*$. Under assumption (H1) and if the following assumptions on the volatility coefficients hold,*

$$(H2) \quad \exists K_b \text{ such that } \forall (t, s) \in [0, T] \times \mathbb{R}_+, \quad |\sigma(t, s)| + |\eta_j(t, s)| \leq K_b.$$

$$(H3) \quad \exists K_\sigma \text{ such that } \forall (t, s_1, s_2) \in [0, T] \times \mathbb{R}_+ \times \mathbb{R}_+, \quad |s_1 \sigma(t, s_1) - s_2 \sigma(t, s_2)| \leq K_\sigma |s_1 - s_2|.$$

then

$$\mathbb{E} \left(\sup_{0 \leq t \leq T} |I_t^M - I_t|^{2p} \right) \leq C_T \left(\left(\sum_{j=1}^M w_j^2 \right)^p + \left(\sum_{j=1}^M w_j |\beta_j - \beta| \right)^{2p} + \left(\sum_{j=1}^M w_j |\delta_j - \delta| \right)^{2p} \right)$$

where

$$C_T = 8^{2p-1} T^p (T^p + K_p K_b^{2p}) C_p \exp \left(4^{2p-1} T (2^{2p-1} K_p T^{p-1} (\beta K_\sigma)^{2p} + (2T)^{2p-1} \delta^{2p} + r^{2p} T^{2p-1}) \right)$$

and

$$C_p = \max_{1 \leq j \leq M} |S_0^{j, M}|^{2p} \exp \left(\left(2r + (2p-1) (\max_{j \geq 1} \beta_j^2 + 1) K_b^2 \right) pT \right).$$

The next theorem states that, under an additional assumption on the volatility coefficients, the L^{2p} -distance between a stock $(S_t^{j, M})_{t \in [0, T]}$ and the solution of the SDE obtained by replacing I^M by I

$$\frac{dS_t^j}{S_t^j} = (r - \delta_j) dt + \beta_j \sigma(t, I_t) dB_t + \eta_j(t, S_t^j) dW_t^j, \quad S_0^j = S_0^{j, M}$$

is controlled by the L^{2p} -distance between I^M and I :

⁴500 stocks for the S&P 500 index, 100 stocks for the FTSE 100 index, 40 stocks for the CAC40 index, etc.

Theorem 2 — Let $p \in \mathbb{N}^*$. Under the assumptions of Theorem 1 and if

$$\begin{aligned} (\mathcal{H4}) \quad & \exists K_\eta \text{ such that } \forall (t, s_1, s_2) \in [0, T] \times \mathbb{R}_+ \times \mathbb{R}_+, \quad |s_1 \eta(t, s_1) - s_2 \eta(t, s_2)| \leq K_\eta |s_1 - s_2|. \\ & \exists K_{Lip} \text{ such that } \forall (t, s_1, s_2) \in [0, T] \times \mathbb{R}_+ \times \mathbb{R}_+, \quad |\sigma(t, s_1) - \sigma(t, s_2)| \leq K_{Lip} |s_1 - s_2|. \end{aligned}$$

Then, $\forall j \in \{1, \dots, M\}$,

$$\mathbb{E} \left(\sup_{0 \leq t \leq T} |S_t^{j,M} - S_t^j|^{2p} \right) \leq \tilde{C}_T^j \left(\left(\sum_{j=1}^M w_j^2 \right)^p + \left(\sum_{j=1}^M w_j |\beta_j - \beta| \right)^{2p} + \left(\sum_{j=1}^M w_j |\delta_j - \delta| \right)^{2p} \right)$$

where

$$\tilde{C}_T^j = 6^{2p-1} K_p T^p \beta_j^{2p} C_{2p}^{\frac{1}{2}} K_{Lip}^{2p} e^{3^{2p-1} ((r-\delta_j)^{2p} T^{2p-1} + K_p T^{p-1} K_\eta^{2p} + 2^{2p-1} K_p T^{p-1} \beta_j^{2p} K_b^{2p}) T}.$$

Moreover, for $\bar{I}_t^M = \sum_{j=1}^M w_j S_t^j$, one has

$$\mathbb{E} \left(\sup_{0 \leq t \leq T} |I_t^M - \bar{I}_t^M|^{2p} \right) \leq \tilde{C}_T \left(\sum_{j=1}^M w_j \right)^{2p} \left(\left(\sum_{j=1}^M w_j^2 \right)^p + \left(\sum_{j=1}^M w_j |\beta_j - \beta| \right)^{2p} + \left(\sum_{j=1}^M w_j |\delta_j - \delta| \right)^{2p} \right)$$

where $\tilde{C}_T = \max_{1 \leq j \leq M} \tilde{C}_T^j$.

The proof for these two theorems can be found in the appendix. Note that, Theorems 1 and 2 yield that I^M is also close to I . In the following corollary, we make explicit the dependence of the coefficients on M and we consider the limit $M \rightarrow \infty$:

Corollary 3 — Under the assumptions of Theorems 1 and 2 and if

$$(\mathcal{H5}) \quad \text{there exists a constant } A \text{ such that } \max_{j \geq 1} \left((S_0^{j,M})^2 + (\beta_j^M)^2 + (\delta_j^M)^2 \right) \leq A,$$

$$(\mathcal{H6}) \quad P_w^M = \sum_{j=1}^M (w_j^M)^2 \xrightarrow{M \rightarrow \infty} 0,$$

$$(\mathcal{H7}) \quad P_\beta^M = \sum_{j=1}^M w_j^M |\beta_j^M - \beta| \xrightarrow{M \rightarrow \infty} 0,$$

$$(\mathcal{H8}) \quad P_\delta^M = \sum_{j=1}^M w_j^M |\delta_j^M - \delta| \xrightarrow{M \rightarrow \infty} 0,$$

then one has

$$\mathbb{E} \left(\sup_{0 \leq t \leq T} |I_t^M - I_t|^2 \right) \xrightarrow{M \rightarrow \infty} 0$$

and

$$\forall j \in \{1, \dots, M\}, \quad \mathbb{E} \left(\sup_{0 \leq t \leq T} |S_t^{j,M} - S_t^j|^2 \right) \xrightarrow{M \rightarrow \infty} 0.$$

If, in addition, $\sup_M \sum_{j=1}^M w_j^M < \infty$ then

$$\mathbb{E} \left(\sup_{0 \leq t \leq T} |I_t^M - \bar{I}_t^M|^2 \right) \xrightarrow{M \rightarrow \infty} 0.$$

Let us briefly comment on these additional assumptions :

- Assumption ($\mathcal{H}5$) is a technical assumption that prevents the constants C_T and \tilde{C}_T appearing in the Theorems 1 and 2 from depending on M . It says that the initial stock levels, the beta coefficients and the dividend yields are uniformly bounded which is not restrictive.
- Assumption ($\mathcal{H}6$) sets a condition on the weightings $(w_j^M)_{j=1 \dots M}$. For example, uniform weights do satisfy this condition :

$$\sum_{j=1}^M \frac{1}{M^2} = \frac{1}{M} \xrightarrow{M \rightarrow \infty} 0$$

In Table 1, we compute the quantity P_w^M for the Eurostoxx index and find that it is indeed very small (of the order $\frac{1}{M}$).

- Assumptions ($\mathcal{H}7$) and ($\mathcal{H}8$) are similar. They express the fact that the distance between $(\beta_j^M)_{j=1 \dots M}$ and β and the distance between $(\delta_j^M)_{j=1 \dots M}$ and δ tends to 0 when M tends to infinity. More importantly, they give us a means of determining the parameters β and δ :

$$\frac{\sum_{j=1}^M w_j^M |\beta_j^M - \beta|}{\sum_{i=1}^M w_i^M} = \mathbb{E} |Y_\beta - \beta| \quad \text{and} \quad \frac{\sum_{j=1}^M w_j^M |\delta_j^M - \delta|}{\sum_{i=1}^M w_i^M} = \mathbb{E} |Y_\delta - \delta|$$

where Y_β and Y_δ are discrete random variables having the following probability distributions :

$$\forall j \in \{1, \dots, M\}, \quad \mathbb{P}(Y_\beta = \beta_j) = \frac{w_j^M}{\sum_{i=1}^M w_i^M} \quad \text{and} \quad \mathbb{P}(Y_\delta = \delta_j^M) = \frac{w_j^M}{\sum_{i=1}^M w_i^M}.$$

Consequently, the optimal choice of the parameters is the median⁵ of Y_β for β and the median of Y_δ for δ . Nevertheless, one does not actually have the choice for the coefficient β . Indeed, recall that by definition of the beta coefficients :

$$\beta_j^M := \frac{\text{Cov}(r_j, r_I)}{\text{Var}(r_I)} = \frac{\beta_j \beta \sigma^2}{\beta^2 \sigma^2} = \frac{\beta_j}{\beta},$$

so one should take $\beta = 1$. In Table 1, we see that the optimal choice of β is very close to 1 and that the quantities of interest, $(P_{\beta_{opt}}^M)^2$ and $(P_{\beta=1}^M)^2$ are also very close to each other.

⁵The median of a real random variable X is any real number m satisfying :

$$\mathbb{P}(X \leq m) \geq \frac{1}{2} \quad \text{and} \quad \mathbb{P}(X \geq m) \geq \frac{1}{2}.$$

It has the property of minimizing the L^1 -distance to X : $m = \arg \min_{x \in \mathbb{R}} \mathbb{E}|X - x|$.

P_w^M	β_{opt}	$(P_{\beta_{opt}}^M)^2$	$(P_{\beta=1}^M)^2$
0.026	0.975	0.0173	0.0174

Table 1: Computation of P_w^M, β_{opt} and $(P_{\beta_{opt}}^M)^2$ for the Eurostoxx index at December 21, 2007. The beta coefficients are estimated on a two year history.

Simplified model

To sum up, we have shown that, under mild assumptions, when the number of underlying stocks is large, the original model may be approximated by the following dynamics

$$\forall j \in \{1, \dots, M\}, \quad \frac{dS_t^j}{S_t^j} = (r - \delta_j)dt + \beta_j \sigma(t, I_t)dB_t + \eta_j(t, S_t^j)dW_t^j \quad (6)$$

$$\frac{dI_t}{I_t} = (r - \delta_I)dt + \sigma(t, I_t)dB_t.$$

Interestingly, we end up with a local volatility model for the index and, for each stock, a stochastic volatility model decomposed into a systemic part driven by the index level and an intrinsic part. Note that this simplified model is not valid for options written on the index together with all its composing stocks since the index is no longer an exact, but an approximate, weighted sum of the stocks. In this case, one should consider the reconstructed index $\bar{I}_t^M = \sum_{j=1}^M w_j S_t^j$ or use the original model.

The fact remains that the simplified model can be used for options written on the stocks or on the index or even on the index together with few stocks.

3 Model calibration

Calibration, which is how to determine the model parameters in order to fit market prices at best, is of paramount importance in practice. In the following, we try to tackle this issue for both our simplified and original model :

3.1 Simplified model

$$\forall j \in \{1, \dots, M\}, \quad \frac{dS_t^j}{S_t^j} = (r - \delta_j)dt + \beta_j \sigma(t, I_t)dB_t + \eta_j(t, S_t^j)dW_t^j$$

$$\frac{dI_t}{I_t} = (r - \delta_I)dt + \sigma(t, I_t)dB_t$$
(7)

The short interest rate and the dividend yields can be extracted from the market. The calibration of the local volatility σ to fit index option prices is a classic problem. What seems to be the market practice is to do a best-fit of a chosen parametric form and match it to the available market prices. This is an important feature of our model : even though the index is reconstructed from the stocks, its calibration remains comparatively easy. Actually our model gives an advantage to the fit of index option prices in comparison with options written on the stocks, which is in line with the market since index options are usually very liquid in comparison with individual stock options.

The calibration of the beta coefficients is more tedious. Indeed, estimation based on historical data can be unsuitable for our model when the historical beta is much larger than the implied one : in this case, since the slope of the local volatility of the index is usually steeper than the one of the stock, the systemic part of the volatility of the stock in our model can be larger than the local volatility of the stock.

To be specific, thanks to the usual formula relating the stochastic volatility to the local volatility (for the theoretical result, see the paper of Gyöngy [10]), one can express the local variance of the stock as

$$v_{loc}(t, K) = \eta^2(t, K) + \beta^2 \mathbb{E}(\sigma^2(t, I_t) | S_t = K).$$
(8)

We see that when $\beta_{hist}^2 \mathbb{E}(\sigma^2(t, I_t) | S_t = K)$ becomes larger than $v_{loc}(t, K)$, the local volatility given by our model is larger than the true local volatility of the stock. The right way to handle the estimation of the beta coefficient is then to compute an implied beta calibrated to the options market. Unfortunately, there is no option product that permits us to do this reasonably⁶ and one should take a beta coefficient lower than the historical beta whenever the preceding problem is encountered and a beta coefficient higher than the historical one whenever it is possible, such that the following rule of thumb is observed :

$$\sum_{j=1}^M w_j \beta_j \simeq 1.$$

In Figure 1, we have plotted both the local volatility of the stock, the local volatility of the index, the systemic part of the volatility of the stock $\beta_{hist} \sigma(T, I_T)$ and $\beta_{hist} \mathbb{E}(\sigma(T, I_T) | S_T = K)$ when η is set to zero (which intuitively gives the lowest local volatility that one can obtain in our model) for a maturity $T = 1$ year. We considered three components of the Eurostoxx : AXA, ALCATEL and CARREFOUR at December 21, 2007. We made this choice deliberately in order to point out the extreme situations one can face :

⁶One financial product that can lead to an easy calibration of the beta coefficient should revolve around the correlation between an index and one of its composing stocks. This is not the case for the most liquid correlation swaps which are sensitive to an average correlation between all the stocks.

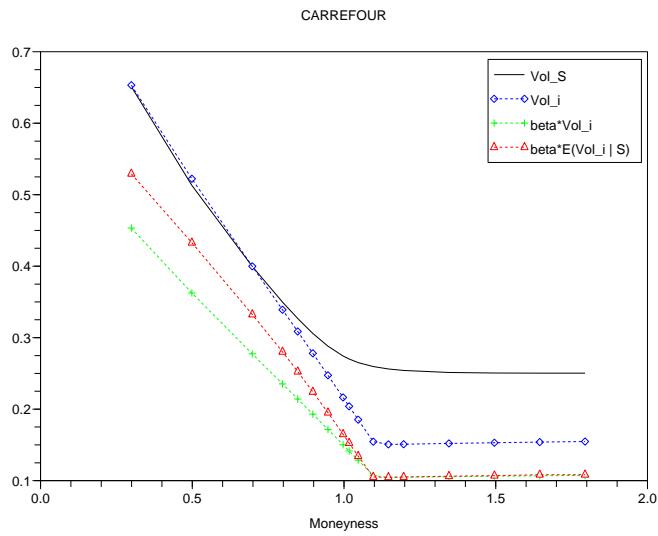
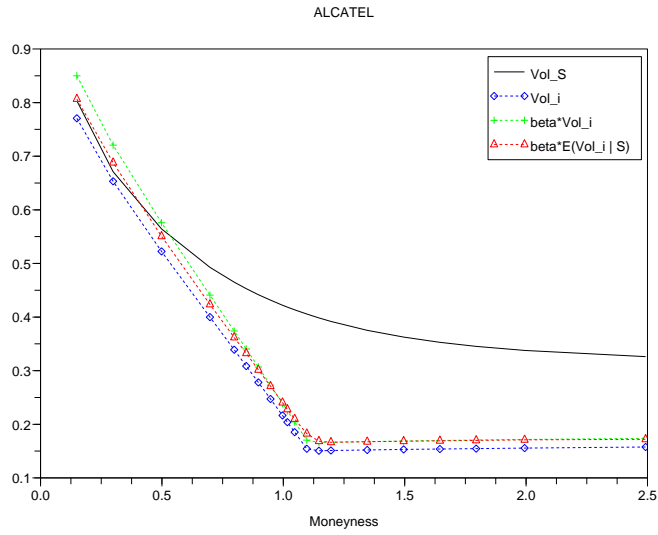
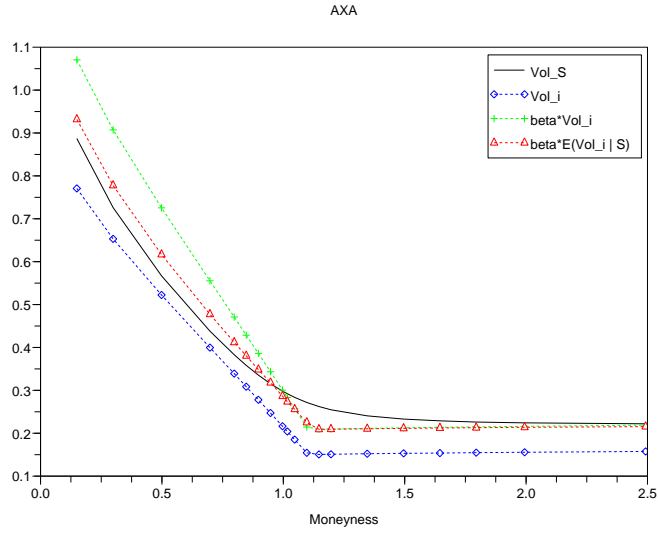


Figure 1: Local volatilities of AXA, ALCATEL and CARREFOUR together with $\sigma(T, I_T^{Eurostoxx})$, $\beta_{hist}\sigma(T, I_T^{Eurostoxx})$ and $\beta_{hist}\mathbb{E}(\sigma(T, I_T^{Eurostoxx})|S_T = K)$ when η is set to zero.

- AXA is an example of a stock with a high beta coefficient ($\beta = 1.4$).
- CARREFOUR is an example of a stock with a low beta coefficient ($\beta = 0.7$).
- ALCATEL is an example of a stock with a high volatility level but with a low smile effect ($\beta = 1.1$).

The local volatilities are obtained from a parametric function of the forward moneyness achieving a best-fit to market smile data. The x -axis represents the moneyness, that is the strike over the spot ($\frac{K}{S_0}$ for a the stock and $\frac{K}{I_0}$ for the index). Clearly, we can deduce that the market is choosing a *beta* coefficient for both AXA and ALCATEL that is lower than the historical one whereas, for CARREFOUR, one can plug the historical beta, or even a larger one, in (7) and still be able to calibrate the model.

Finally, the remaining parameters that have to be calibrated to fit option prices are the volatility coefficients η_1, \dots, η_M . From now on, we omit the index j to simplify the notations and we consider the issue of calibrating the volatility coefficient η for a given stock.

From equation (8), one gets

$$\eta(t, K) = \sqrt{v_{loc}(t, K) - \beta^2 \mathbb{E}(\sigma^2(t, I_t) | S_t = K)}. \quad (9)$$

As previously mentioned, v_{loc} can be determined with the best-fit of a parametric form to the stock market smile but determining the conditional expectation is a more challenging task. Note that, since the law of (S_t, I_t) depends on η , so does the conditional expectation and therefore it is difficult to get an estimation of it or to simulate a stochastic differential equation that gives the same vanilla prices as those given by the market. In order to address this issue, we suggest two different simulation based approaches. The first one is based on non-parametric estimation of the conditional expectation and the second one on parametric estimation.

3.1.1 Estimation of the conditional expectation

The idea behind the following techniques is to circumvent the difficulty of calibrating the volatility coefficient η . Indeed, if we plug the formula (8) in the dynamics of the stock, we obtain a stochastic differential equation that is nonlinear in the sense of McKean :

$$\begin{aligned} \frac{dS_t}{S_t} &= (r - \delta)dt + \beta \sigma(t, I_t)dB_t + \sqrt{v_{loc}(t, S_t) - \beta^2 \mathbb{E}(\sigma^2(t, I_t) | S_t)}dW_t \\ \frac{dI_t}{I_t} &= (r - \delta_I)dt + \sigma(t, I_t)dB_t \end{aligned} \quad (10)$$

For an introduction to the topics of nonlinear stochastic differential equations and propagation of chaos, we refer to the lecture notes of Sznitman [17] and Méléard [14]. In our case, the nonlinearity appears in the diffusion coefficient through the conditional expectation term. This makes the natural question of existence and uniqueness of a solution very difficult to handle. The case of a drift involving a conditional expectation has only been handled recently even for constant diffusion coefficient (see Talay and Vaillant [18] and Dermoune [9]). Meanwhile, it is possible to simulate such a stochastic differential equation by means of a system of N interacting paths using either a non-parametric estimation of the conditional expectation or regression techniques. The advantage of the regression approach over the non-parametric estimation is that it also yields a smooth approximation of the function $\mathbb{E}(\sigma^2(t, I_t) | S_t = s)$ whereas, with a non-parametric method, one has to interpolate the estimated function and to carefully tune the window parameter to obtain a smooth approximation.

3.1.1a Non-parametric estimation

Non-parametric estimators of the conditional expectation, and more generally non-parametric density estimators, have been widely studied in the literature. We will focus on kernel estimators of the Nadaraya-Watson type (see [19] and [15]) : given N observations $(S_t^i, I_t^i)_{i=1\dots N}$ of (S_t, I_t) , we consider the kernel conditional expectation estimator of $\mathbb{E}(\sigma^2(t, I_t) | S_t = s)$ given by

$$\frac{\sum_{i=1}^N \sigma^2(t, I_t^i) K\left(\frac{s - S_t^i}{h_N}\right)}{\sum_{i=1}^N K\left(\frac{s - S_t^i}{h_N}\right)}$$

where K is a non-negative kernel such that $\int_{\mathbb{R}} K(x)dx = 1$ and h_N is a smoothing parameter which tends to zero as $N \rightarrow +\infty$. This leads to the following system with N interacting particles : $\forall 1 \leq i \leq N$,

$$\begin{cases} \frac{dS_t^{i,N}}{S_t^{i,N}} = (r - \delta)dt + \beta \sigma(t, I_t^i)dB_t^i + \sqrt{v_{loc}(t, S_t^{i,N}) - \beta^2 \frac{\sum_{j=1}^N \sigma^2(t, I_t^j) K\left(\frac{S_t^{i,N} - S_t^{j,N}}{h_N}\right)}{\sum_{j=1}^N K\left(\frac{S_t^{i,N} - S_t^{j,N}}{h_N}\right)}} dW_t^i, S_0^{i,N} = S_0 \\ \frac{dI_t^i}{I_t^i} = (r - \delta_I)dt + \sigma(t, I_t^i)dB_t^i, I_0^i = I_0 \end{cases}$$

where $(B^i, W^i)_{i \geq 1}$ is a sequence of independent two-dimensional Brownian motions. This $2N$ -dimensional SDE may be discretized using the Euler scheme :

Let $0 = t_0 < \dots < t_M = T$ be a subdivision with step $\frac{T}{M}$ of $[0, T]$. For each $k \in \{0, \dots, M-1\}$, $\forall 1 \leq i \leq N$,

$$\begin{cases} \bar{S}_{t_{k+1}}^{i,N} = \bar{S}_{t_k}^{i,N} \left((r - \delta) \frac{T}{M} + \beta \sigma(t_k, \bar{I}_{t_k}^i) \sqrt{\frac{T}{M}} G_{i,k}^1 + \sqrt{v_{loc}(t_k, \bar{S}_{t_k}^{i,N}) - \beta^2 \frac{\sum_{j=1}^N \sigma^2(t_k, \bar{I}_{t_k}^j) K\left(\frac{\bar{S}_{t_k}^{i,N} - \bar{S}_{t_k}^{j,N}}{h_N}\right)}{\sum_{j=1}^N K\left(\frac{\bar{S}_{t_k}^{i,N} - \bar{S}_{t_k}^{j,N}}{h_N}\right)}} \sqrt{\frac{T}{M}} G_{i,k}^2 \right) \\ \bar{I}_{t_{k+1}}^i = \bar{I}_{t_k}^i \left((r - \delta_I) \frac{T}{M} + \sigma(t_k, \bar{I}_{t_k}^i) \sqrt{\frac{T}{M}} G_{i,k}^1 \right) \end{cases}$$

where $(G_{i,k}^1)_{1 \leq i \leq N, 0 \leq k \leq M-1}$ and $(G_{i,k}^2)_{1 \leq i \leq N, 0 \leq k \leq M-1}$ are independent centered and reduced Gaussian random variables.

3.1.1b Parametric estimation

Another approach to estimate conditional expectations is to use parametric estimators, or projection. This idea has also been widely used and studied previously (for example in finance, one can think of the Longstaff-Schwartz algorithm for pricing American options [13]). Noting that the conditional expectation is a projection operator on the space of square integrable random variables, one can approximate $\mathbb{E}(\sigma^2(t, I_t) | S_t = s)$ by the parametric estimator

$$\sum_{k=1}^K \alpha_k f_k(s)$$

where $(f_k)_{k=1\dots K}$ is a functional basis and $\alpha = (\alpha_k)_{k=1\dots K}$ is a vector of parameters estimated by least mean squares : given N observations $(S_t^i, I_t^i)_{i=1\dots N}$ of (S_t, I_t) , α minimizes $\sum_{i=1}^N \left(\sigma^2(t, I_t^i) - \sum_{k=1}^K \alpha_k f_k(S_t^i) \right)^2$.

3.1.2 Numerical results

3.1.2a A toy model

In the first numerical example, we suppose that the local volatility of the stock is constant and we try to reconstruct it by simulating the particle system of the non-parametric method presented above. We consider the Eurostoxx index and we determine its local volatility by fitting the market prices at December 21, 2007.

As described above, we can approximate the following SDE using a system of N interacting particles :

$$\begin{aligned}\frac{dS_t}{S_t} &= (r - \delta)dt + \beta \sigma(t, I_t)dB_t + \sqrt{v - \beta^2 \mathbb{E}(\sigma^2(t, I_t) | S_t)}dW_t \\ \frac{dI_t}{I_t} &= (r - \delta_I)dt + \sigma(t, I_t)dB_t\end{aligned}\tag{11}$$

Using these simulations to price European call options for different strikes, one should obtain the same results as a Black & Scholes model with volatility \sqrt{v} . In Figure 2, we plot the implied volatility obtained by independent simulations of $N = 5000$ paths and see that the implied volatilities obtained are indeed close to the exact volatility level. This example was generated with the following arbitrary set of parameters :

- $S_0 = 100$.
- $\beta = 0.7$.
- $r = 0.05$.
- $\delta = \delta_I = 0$.
- $\sqrt{v} = 0.3$.
- $T = 1$.
- Number of simulated paths : $N = 5000$.
- Number of time steps in the Euler scheme : $M = 20$.

In this example and for all the following numerical experiments, we use a Gaussian kernel : $K(u) = \frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}$. The smoothing parameter h_N is set to $N^{-\frac{1}{5}}$ which is the optimal bandwidth that one obtains when minimizing the asymptotic mean square error of the Nadaraya-Watson estimator under some regularity assumptions and assuming independence of the random variables involved (see for example Bosq [6]).

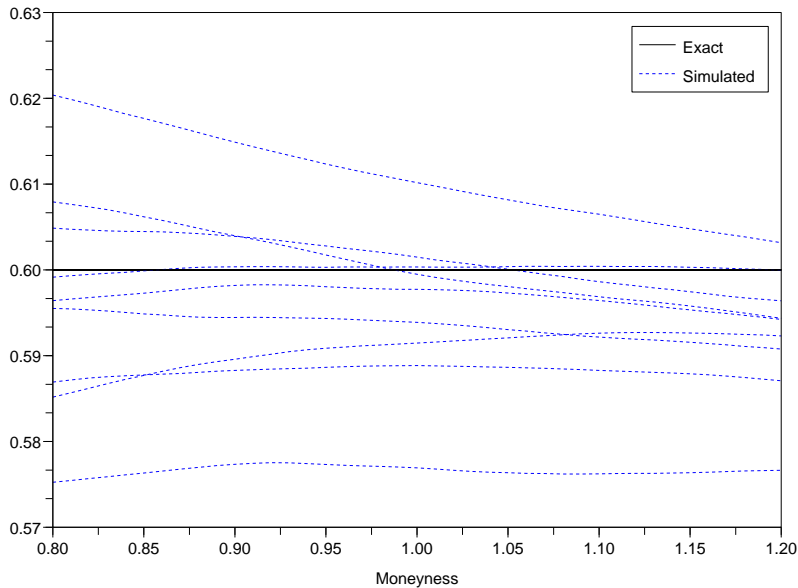


Figure 2: Implied volatility obtained for nine independent simulations with $N = 5000$ paths.

3.1.2b An example with real data

In the following, we test our model with real data. More precisely, given the local volatilities of the Eurostoxx index and of Carrefour at December 21, 2007, we simulate the particle system (10) by different methods for a one year maturity.

An acceleration technique

The simulation of the particle system is very time consuming : for each discretization step and for each stock particle, one has to make N computations which yield a global complexity of order $O(MN^2)$ where M is the number of time steps in the Euler scheme. Acceleration techniques are thus unavoidable. One possible method consists in reducing the number of interactions : instead of making N computations for each estimation of the conditional expectation, one can neglect interactions which involve particles which are far away from each other. When the kernel used is non increasing with the absolute value of its argument, the easiest way to implement this idea is to sort the particles at each step and, whenever a contribution of a particle is lower than some fixed threshold, to stop the estimation of the conditional expectation.

Of course, by doing this, we lose in precision for the same number of interacting particles, especially for deep in/out of the money strikes. But what we gain in terms of computation time is much more important : in Figure 3, we plot the implied volatility obtained by the naive method and the method with the above acceleration technique for the same number $N = 10000$ of particles. We take as threshold $\frac{1}{N}$ and set $h_N = N^{-\frac{1}{10}}$ for the bandwidth parameter⁷ and $M = 20$ for the number of time steps in the Euler scheme. The computation time, on a computer with a 2.8 Ghz Intel Pentium 4 processor, is of 52 minutes for the naive method and of 5 minutes for the accelerated one.

More importantly, we see that the implied volatility $\hat{\sigma}_{simul}$ obtained by simulations converges to the exact

⁷In order to smooth the estimation, one has to choose a bandwidth parameter that is greater than the theoretical optimal parameter $N^{-\frac{1}{5}}$.

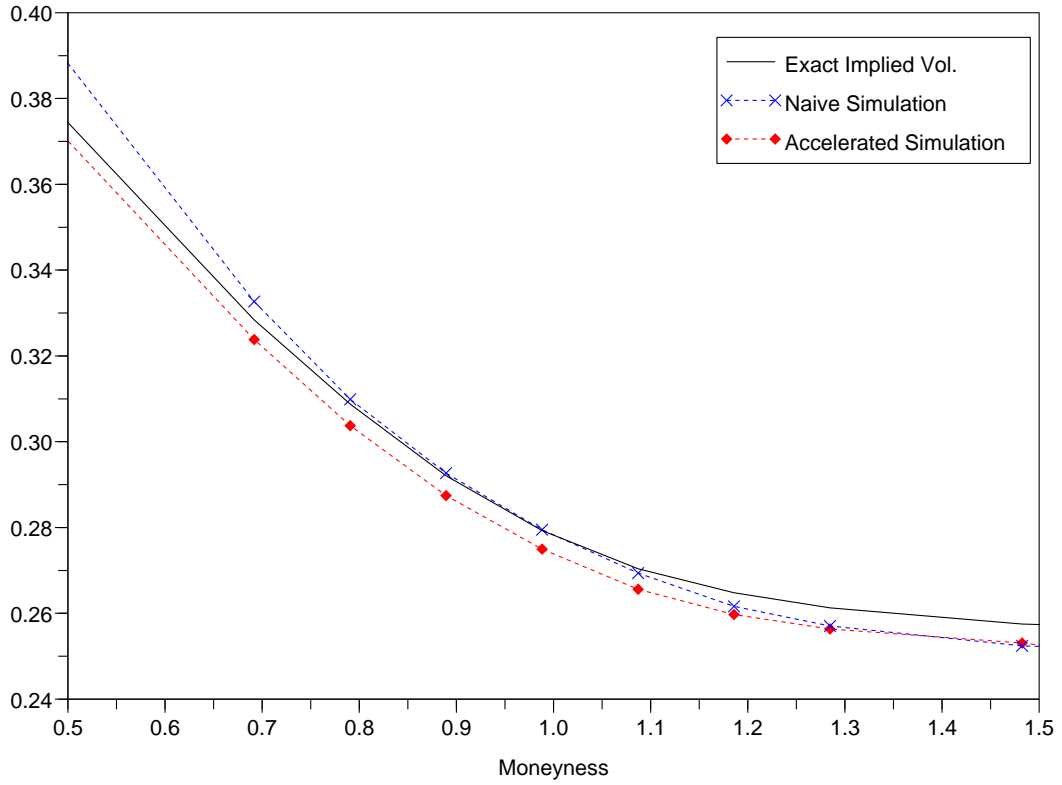


Figure 3: Comparison between the naive technique and the accelerated one.

volatility $\hat{\sigma}_{exact}$: see Figure 4 and Table 2. With a reasonable number of simulated paths, $N = 200000$, the error on the implied volatility remains clearly tolerable for practitioners (of the order of 10 bp) except for a deep in the money call ($K = 0.3S_0$) where it attains 195 bp.

Moneyness ($\frac{K}{S_0}$)	0.30	0.49	0.69	0.79	0.89	0.99	1.09	1.19	1.28	1.48	1.98
Error : $ \hat{\sigma}_{simul} - \hat{\sigma}_{exact} $	195	36	8	5	2	1	2	9	17	32	56

Table 2: Error (in bp) on the implied volatility with $N = 200000$ particles.

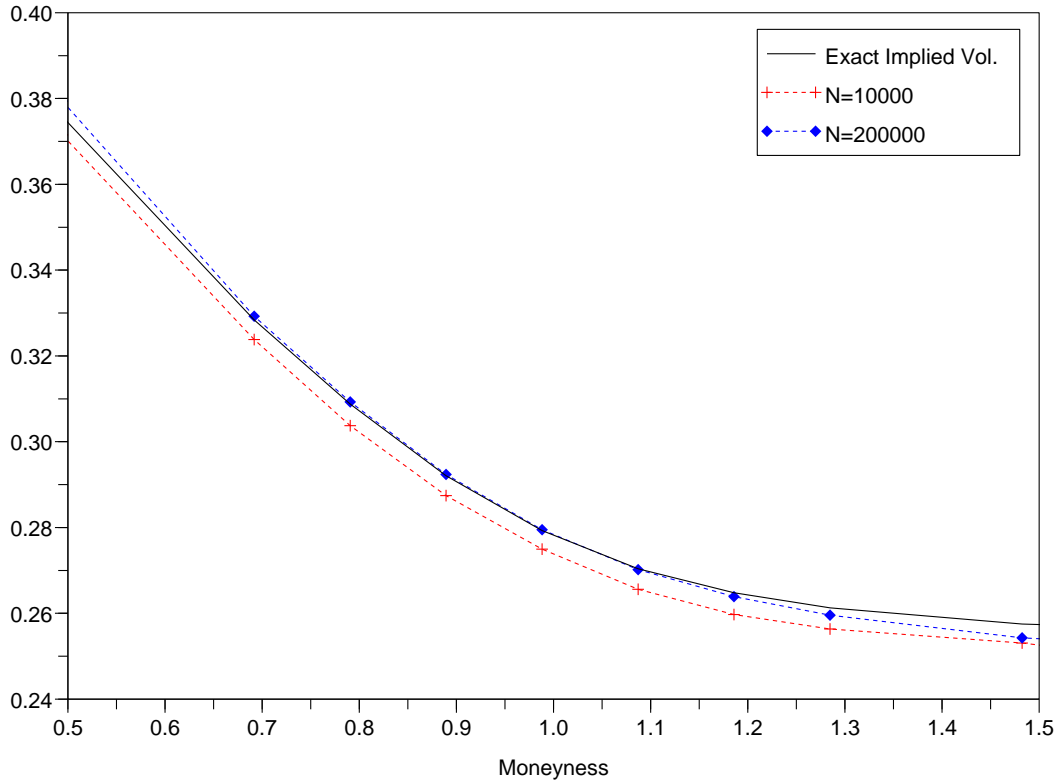


Figure 4: Convergence of the implied volatility obtained with non-parametric estimation.

Independent particles

Unlike the parametric method, non-parametric estimation of the conditional expectation gives the value of the intrinsic volatility η at the simulated points only. However, using an interpolation technique, one can first reconstruct η with N_1 dependent particles and then simulate the 2-dimensional stochastic differential equation with N_2 independent draws, N_2 being larger than N_1 . By doing so, we speed up the simulations but one has to choose carefully the size N_1 of the particle system in order to have a reasonable estimation of the intrinsic volatility and to tune the bandwidth parameter in order to smooth the estimation (our numerical tests were done with $N_1 = 1000$, $N_2 = 100000$ and $h_{N_1} = N_1^{-\frac{1}{10}}$). In Figures 5 and 6, we give the surfaces of both the local volatility and the intrinsic volatility of the stock. This latter is used to draw independent simulations of the index along with the stock and we see in Figure 7 that the implied volatility obtained is close to the right one, especially near the money.

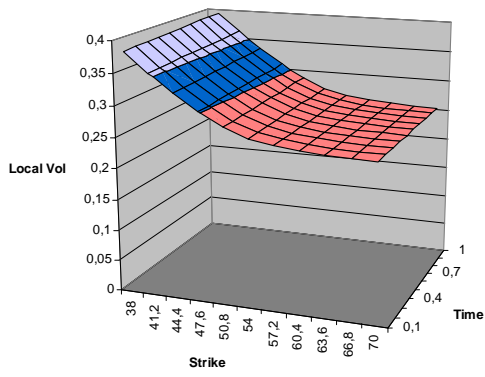


Figure 5: Local volatility surface of the stock.

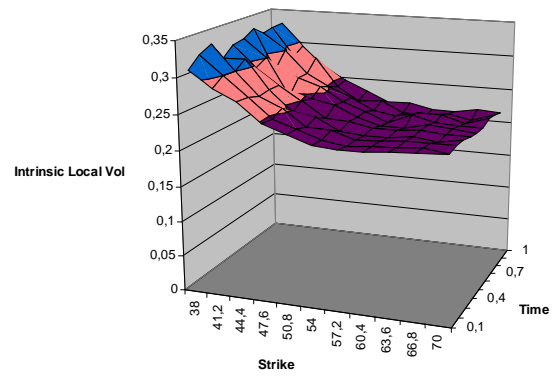


Figure 6: Intrinsic part of the stochastic volatility.

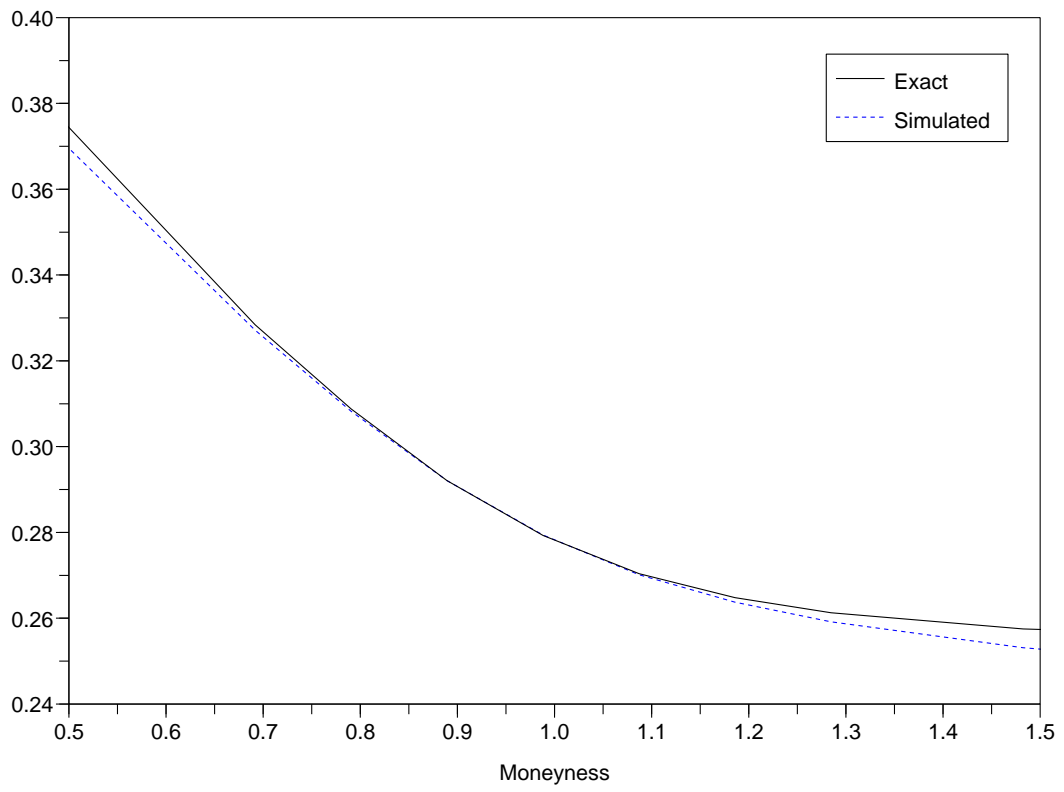


Figure 7: Simulated implied volatility with independent draws.

3.2 Original model

We now turn to the calibration of our original model :

$$\forall j \in \{1, \dots, M\}, \quad \frac{dS_t^{j,M}}{S_t^{j,M}} = (r - \delta_j)dt + \beta_j \sigma(t, I_t^M)dB_t + \eta_j(t, S_t^{j,M})dW_t^j \quad (12)$$

with $I_t^M = \sum_{i=1}^M w_i S_t^{i,M}$.

Obviously, it is rather complicated to have a perfect calibration for both index and stocks within this framework. Nevertheless, one can either

- take for σ the calibrated local volatility of the index and then calibrate the volatility coefficients η_j using an adaptation of the non-parametric method presented above in order to fit all the individual stock smiles at the same time. In this case, the index is not perfectly calibrated but, thanks to Theorem 1, one can expect the error to be small.

Or,

- take for σ and η_j the calibrated coefficients in the simplified model framework. Once again, the calibration is not perfect and this time for both index and individual stocks but Theorems 1 and 2 suggest that the calibration error will be negligible.

Hence, in comparison with the simplified model, we allow ourselves a slight error in the calibration but we guarantee the additivity constraint $I_t^M = \sum_{i=1}^M w_i S_t^{i,M}$. In what follows, we illustrate the effect of Theorems 1 and 2 and compare our models with a constant correlation model.

4 Illustration of Theorems 1 and 2 and comparison with a constant correlation model

The objective of this section is to compare index and individual stock smiles obtained with three different models : our original model (12), the simplified one (after letting $M \rightarrow \infty$) and a model with constant correlation coefficient. More precisely, we consider the following dynamics

1. The original model

$$\forall j \in \{1, \dots, M\}, \quad \frac{dS_t^{j,M}}{S_t^{j,M}} = rdt + \sigma(t, I_t^M)dB_t + \eta(t, S_t^{j,M})dW_t^j \quad (13)$$

$$\text{with } I_t^M = \sum_{i=1}^M w_i S_t^{i,M}.$$

2. The simplified model

$$\forall j \in \{1, \dots, M\}, \quad \frac{dS_t^j}{S_t^j} = rdt + \sigma(t, I_t)dB_t + \eta(t, S_t^j)dW_t^j \quad (14)$$

$$\frac{dI_t}{I_t} = rdt + \sigma(t, I_t)dB_t.$$

Where we can also compute the reconstructed index $\bar{I}_t^M = \sum_{i=1}^M w_i S_t^i$.

3. The "Market" model

$$\forall j \in \{1, \dots, M\}, \quad \frac{dS_t^j}{S_t^j} = rdt + \sqrt{v_{loc}(t, S_t^j)}d\widetilde{W}_t^j \quad (15)$$

with, $\forall i \neq j, d\langle \widetilde{W}^i, \widetilde{W}^j \rangle_t = \rho dt$.

We deliberately dropped the dividend yields and the beta coefficients in order to simplify the numerical experiment. For the volatility coefficient σ , we take as previously the calibrated local volatility of the Eurostoxx. We choose an arbitrary parametric form, function of the forward moneyness, for the volatility coefficient η and we evaluate v_{loc} such that the market model and the simplified model yield the same implied volatility for individual stocks. Indeed, it suffices to take

$$v_{loc}(t, s) = \eta^2(t, s) + \mathbb{E}(\sigma^2(t, I_t) | S_t = s)$$

where the conditional expectation can be approximated using the non-parametric method presented above.

Finally, we fix the correlation coefficient ρ such that the market model and the simplified one have the same ATM implied volatility for the index.

The implied volatilities for the index and for an individual stocks obtained by the three models are plotted in Figures 8 and 9. We also give the difference in basis points between the implied volatilities obtained with the simplified model and the original one in Tables 3, 4 and 5. The parameters we use in our numerical experiment are the following :

- $S_0^1 = \dots = S_0^M = 53$.
- M , I_0 and the weights w_1, \dots, w_M : the same as of the Eurostoxx index at December 21, 2007.
- $r = 0.045$.
- Maturity $T = 1$ year.
- Number of time steps: 10.
- Number of simulated paths : 100000.

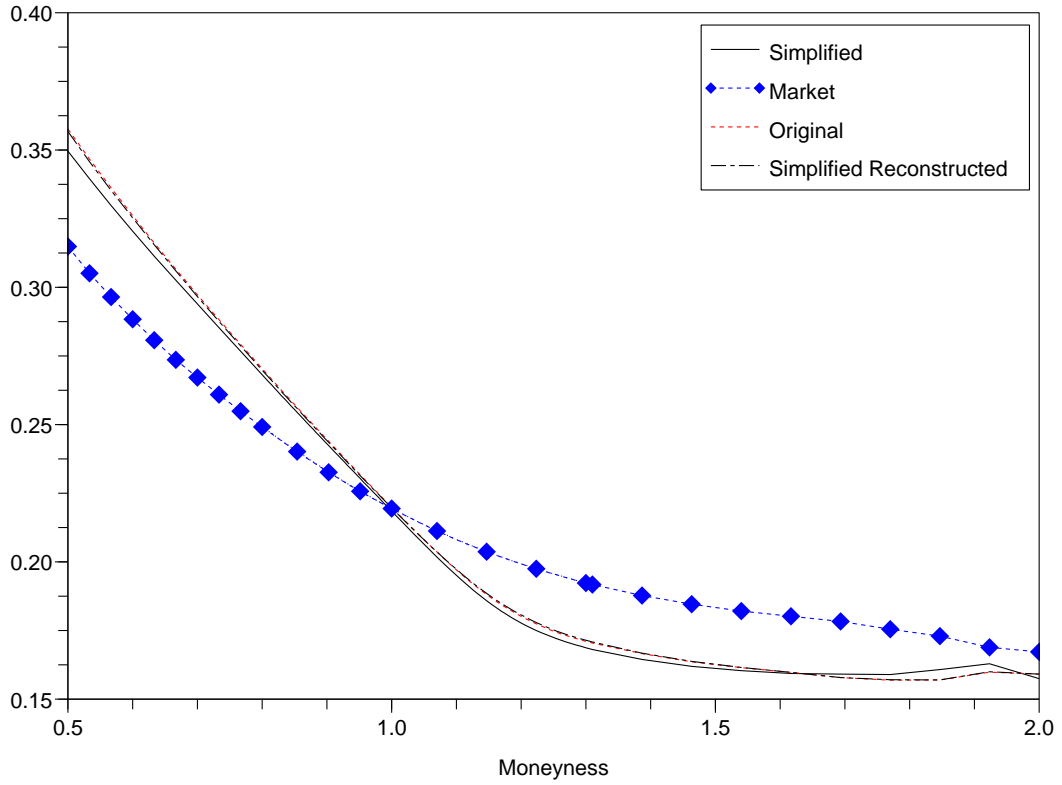


Figure 8: Implied volatility of the index.

Moneyness ($\frac{K}{I_0}$)	0.5	0.8	0.9	0.95	1	1.05	1.1	1.2	1.3	1.55	1.85	2
$ \hat{\sigma}_{simplified} - \hat{\sigma}_{original} $	81	22	16	14	14	17	20	24	24	11	38	17

Table 3: Difference (in bp) between the index implied volatility obtained with the simplified model and the one obtained with the original model.

Moneyness ($\frac{K}{I_0}$)	0.5	0.8	0.9	0.95	1	1.05	1.1	1.2	1.3	1.55	1.85	2
$ \hat{\sigma}_{reconstruct} - \hat{\sigma}_{original} $	10	5	4	3	2	1	2	5	4	1	0	0

Table 4: Difference (in bp) between the implied volatility of the reconstructed index \bar{I}^M in the simplified model and the index implied volatility obtained with the original model.

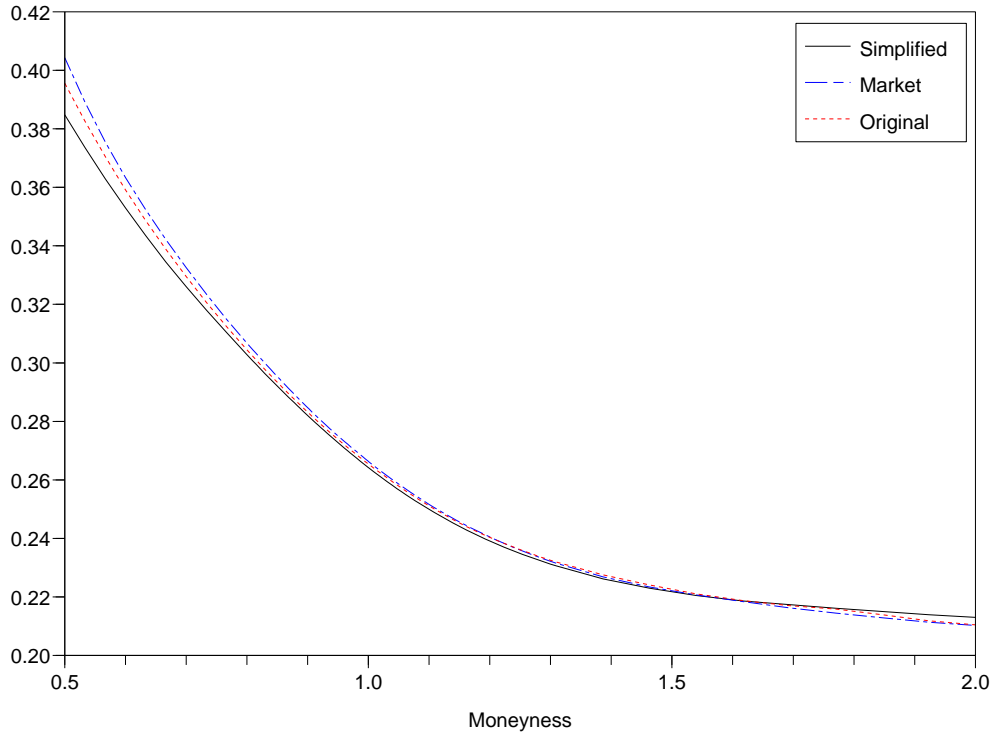


Figure 9: Implied volatility of an individual stock.

Moneyness ($\frac{K}{S_0}$)	0.5	0.8	0.9	0.95	1	1.05	1.1	1.2	1.3	1.55	1.85	2
$ \hat{\sigma}_{simplified} - \hat{\sigma}_{original} $	81	22	16	14	14	17	20	24	24	11	38	17

Table 5: Difference (in bp) between an individual stock implied volatility obtained with the simplified model and the one obtained with the original model.

As suggested by Theorems 1 and 2, we see that the original model and the simplified one yield implied volatility curves that are very close to each other, both for the index and for individual stocks. The difference in basis points between the implied volatilities is reasonable, especially between the reconstructed index implied volatility of the simplified model and the index implied volatility of the original model.

Concerning the market model, by construction we have the same implied volatility of an individual stock as for the simplified model but the implied volatility of the index obtained is far from the right one, especially the slope of the smile out-of-the-money. This phenomenon is well known in practice (see [2],[5] or [7]) : the implied volatility smile of an index is much steeper than the implied volatility smile of an individual stock, hence the market model of constantly correlated stocks is unable to retrieve the shape of the index smile. More sophisticated dependence structure between stocks is needed. Our modeling framework circumvents this difficulty since we force the index to have the correct volatility smile while the individual stocks can still be properly calibrated.

4.1 Application: Pricing of a worst-of option

Apart from handling both the index and its composing stocks, our models are also relevant for the widespread financial products that are sensitive to correlation in the equity world, such as rainbow options.

One example of such products is the worst-of performance option whose payout is referenced to the worst performer in a basket of shares. For a basket of M shares, the payoff of a call with strike K and maturity T writes $\left(\min_{1 \leq i \leq M} \frac{S_T^i}{S_0^i} - K \right)_+$. Our objective is to compare the prices obtained by our model to the prices obtained by the market model of constantly correlated stocks. The parameters of the numerical experiment are the same as previously and we set the correlation coefficient ρ such that all the models exhibit the same ATM implied volatility for the index.

The result, as can be seen in Figure 10, is that our prices are always lower than the market model price, especially in the money. Hence, a model with constant correlation coefficient, calibrated in order to fit at the money prices, will always overestimate the risks of worst-of options. Note that the prices obtained with the original model and the simplified one are barely distinguishable from each other.

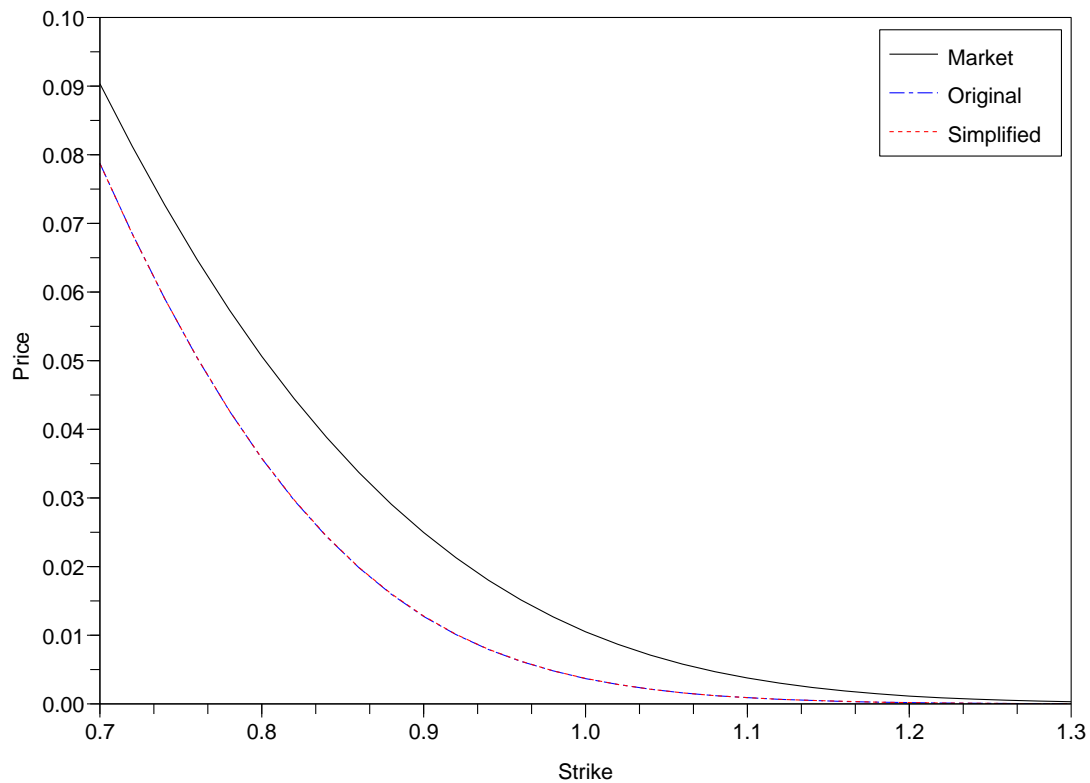


Figure 10: Worst-of price.

5 Conclusion

In this paper, we have introduced a new model for describing the joint evolution of an index and its composing stocks. The idea behind our view is that an index is not only a weighted sum of stocks but can also be seen as a market factor that influences their dynamics. In order to have a more tractable model, we have studied the limit when the number of underlying stocks goes to infinity and we have shown that our model reduces to a local volatility model for the index and to a stochastic volatility model for each individual stock with volatility driven by the index. Unlike the existing models, we favor the fit of the index smile in comparison with the fit of the stock smiles which goes in accordance with the market since index options are usually more liquid than options on a given stock. We have discussed calibration issues and proposed a simulation-based technique for the calibration of the stock dynamics, which permits us to fit both index and stocks smiles. The numerical results obtained on real data for the Eurostoxx index are very encouraging, especially for accelerated techniques. We have also compared our models (before and after passing to the limit) to a market standard model consisting of local volatility models for the stocks which are constantly correlated and we have seen that it is not possible to retrieve the shape of the index smile. Finally, when considering the pricing of worst-of performance options, which are sensitive to the dependence structure between stocks, we have found that our prices are more aggressive than the prices obtained by the standard market model.

To sum up, we list some properties of our models depending on the options one wishes to handle in the Table below

Purpose	Simplified model	Original model
Options written on -few ($J \ll M$) stocks -the index.	-Simulation of a $(J + 1)$ -dimensional SDE : (I, S^1, \dots, S^J) . -Exact calibration of $(S^j)_{1 \leq j \leq J}$ and I possible.	-Simulation of an M -dimensional SDE : $(S^{1,M}, \dots, S^{M,M})$. -Exact calibration of $(S^{j,M})_{1 \leq j \leq J}$ possible. -Approximate calibration of I^M .
Options written on -all the stocks -the index.	-Simulation of an $(M + 1)$ -dimensional SDE : (I, S^1, \dots, S^M) . -Exact calibration of all the stocks possible. -Index value : $\bar{I}_t^M = \sum_{j=1}^M w_j S_t^j$. -Approximate calibration of \bar{I}^M .	-Simulation of an M -dimensional SDE : $(S^{1,M}, \dots, S^{M,M})$. -Exact calibration of all the stocks possible. -Approximate calibration of I^M .

Table 6: Which model to use and when.

References

- [1] M. Avellaneda, D. Boyer-Olson, J. Busca, and P. Friz. Reconstructing volatility. *Risk*, pages 87–91, October 2002.
- [2] G. Bakshi, N. Kapadia, and D. Madan. Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Review of Financial Studies*, 16:101–143, 2003.
- [3] H. Berestycki, J. Busca, and I. Florent. An inverse parabolic problem arising in finance. *C. R. Acad. Sci. Paris Sér. I Math.*, 331(12):965–969, 2000.
- [4] H. Berestycki, J. Busca, and I. Florent. Asymptotics and calibration of local volatility models. *Quantitative Finance*, 2(1):61–69, 2002.
- [5] N. P. B. Bollen and R. E. Whaley. Does net buying pressure affect the shape of implied volatility functions? *Journal of Finance*, 59(2):711–753, 04 2004.
- [6] D. Bosq. *Nonparametric statistics for stochastic processes*, volume 110 of *Lecture Notes in Statistics*. Springer-Verlag, New York, second edition, 1998. Estimation and prediction.
- [7] N. Branger and C. Schlag. Why is the index smile so steep? *Review of Finance*, 8(1):109–127, 2004.
- [8] P. Cizeau, M. Potters, and J-P. Bouchaud. Correlation structure of extreme stock returns. *Quantitative Finance*, 1(2):217–222, February 2001.
- [9] A. Dermoune. Propagation and conditional propagation of chaos for pressureless gas equations. *Probab. Theory Related Fields*, 126(4):459–476, 2003.
- [10] I. Gyöngy. Mimicking the one-dimensional marginal distributions of processes having an Itô differential. *Probab. Theory Relat. Fields*, 71(4):501–516, 1986.
- [11] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*. Springer-Verlag New-York, second edition, 1991.
- [12] P. Lee, L. Wang, and A. Kerim. Index volatility surface via moment-matching techniques. *Risk*, pages 85–89, December 2003.
- [13] FA Longstaff and ES Schwartz. Valuing American options by simulation: a simple least-squares approach. *Rev. Financ. Stud.*, 14(1):113–147, 2001.
- [14] S. Méléard. Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. In *Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995)*, volume 1627 of *Lecture Notes in Math.*, pages 42–95. Springer, Berlin, 1996.
- [15] E. A. Nadaraya. On estimating regression. *Theory of Probability and its Applications*, 9(1):141–142, 1964.
- [16] W. F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3):425–442, 1964.
- [17] A-S. Sznitman. Topics in propagation of chaos. In *École d’Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.
- [18] D. Talay and O. Vaillant. A stochastic particle method with random weights for the computation of statistical solutions of McKean-Vlasov equations. *Ann. Appl. Probab.*, 13(1):140–180, 2003.
- [19] G. S. Watson. Smooth regression analysis. *Sankhyā Ser. A*, 26:359–372, 1964.

Appendix

In order to prove the Theorems 1 and 2, we need the following technical estimation

Lemma 4 — Under assumption $(\mathcal{H}2)$, for all $p \geq 1$, one has

$$\forall j \in \{1, \dots, M\}, \quad \sup_{0 \leq t \leq T} \mathbb{E} \left(|S_t^{j,M}|^{2p} \right) \leq C_p \quad (16)$$

where $C_p = \max_{1 \leq j \leq M} |S_0^{j,M}|^{2p} \exp \left(\left(2r + (2p-1) \left(\max_{j \geq 1} \beta_j^2 + 1 \right) K_b^2 \right) pT \right)$.

Proof : By Itô's lemma one has

$$\begin{aligned} |S_t^{j,M}|^{2p} &= |S_0^{j,M}|^{2p} + \int_0^t |S_s^{j,M}|^{2p} \left((2p)(r - \delta_j) + p(2p-1)(\beta_j^2 \sigma^2(s, I_s^M) + \eta_j^2(s, S_s^{j,M})) \right) ds \\ &\quad + \int_0^t (2p) |S_s^{j,M}|^{2p} (\beta_j \sigma(s, I_s^M) dB_s + \eta_j(s, S_s^{j,M}) dW_s^j) \end{aligned}$$

In order to get rid of the stochastic integral, we use a localization technique : let ν_n be the stopping time defined for each $n \in \mathbb{N}$ by $\nu_n := \inf\{t \geq 0; |S_t^{j,M}| \geq n\}$. Then, using $(\mathcal{H}2)$, one has

$$\begin{aligned} \mathbb{E} \left(|S_{t \wedge \nu_n}^{j,M}|^{2p} \right) &= |S_0^{j,M}|^{2p} + \mathbb{E} \left(\int_0^{t \wedge \nu_n} |S_s^{j,M}|^{2p} \left((2p)(r - \delta_j) + p(2p-1)(\beta_j^2 \sigma^2(s, I_s^M) + \eta_j^2(s, S_s^{j,M})) \right) ds \right) \\ &\leq |S_0^{j,M}|^{2p} + \left((2p)(r - \delta_j) \mathbb{1}_{\{r - \delta_j \geq 0\}} + p(2p-1)(\beta_j^2 + 1) K_b^2 \right) \int_0^t \mathbb{E} \left(|S_{s \wedge \nu_n}^{j,M}|^{2p} \right) ds \end{aligned}$$

So, by Gronwall's lemma and the fact that the dividends are nonnegative,

$$\forall t \leq T, \mathbb{E} \left(|S_{t \wedge \nu_n}^{j,M}|^{2p} \right) \leq |S_0^{j,M}|^{2p} \exp \left(\left(2rp + p(2p-1)(\beta_j^2 + 1) K_b^2 \right) T \right) \quad (17)$$

Finally, Fatou's lemma permits us to conclude :

$$\sup_{0 \leq t \leq T} \mathbb{E} \left(|S_t^{j,M}|^{2p} \right) \leq |S_0^{j,M}|^{2p} \exp \left(\left(2rp + p(2p-1)(\beta_j^2 + 1) K_b^2 \right) T \right). \quad (18)$$

□

Proof of Theorem 1 : Using the SDEs (3) and (5), one has

$$\begin{aligned} |I_t^M - I_t|^{2p} &= \left| r \int_0^t (I_s^M - I_s) ds - \int_0^t \left(\sum_{j=1}^M \delta_j w_j S_s^{j,M} - \delta I_s \right) ds \right. \\ &\quad \left. + \int_0^t \left(\sum_{j=1}^M \beta_j w_j S_s^{j,M} \sigma(s, I_s^M) - \beta I_s \sigma(s, I_s) \right) dB_s + \sum_{j=1}^M w_j \int_0^t S_s^{j,M} \eta_j(s, S_s^{j,M}) dW_s^j \right|^{2p} \\ &\leq 4^{2p-1} \left(r^{2p} t^{2p-1} \int_0^t (I_s^M - I_s)^{2p} ds + t^{2p-1} \int_0^t \left(\sum_{j=1}^M \delta_j w_j S_s^{j,M} - \delta I_s \right)^{2p} ds \right. \\ &\quad \left. + \left| \int_0^t \left(\sum_{j=1}^M \beta_j w_j S_s^{j,M} \sigma(s, I_s^M) - \beta I_s \sigma(s, I_s) \right) dB_s \right|^{2p} + \left| \sum_{j=1}^M w_j \int_0^t S_s^{j,M} \eta_j(s, S_s^{j,M}) dW_s^j \right|^{2p} \right) \end{aligned}$$

Hence, using the Burkholder-Davis-Gundy inequality (see Karatzas and Shreve [11] p. 166), there exists a universal positive constant K_p such that

$$\mathbb{E} \left(\sup_{0 \leq t \leq T} |I_t^M - I_t|^{2p} \right) \leq 4^{2p-1} (a_M + b_M + c_M + d_M)$$

where

- $a_M = r^{2p} T^{2p-1} \int_0^T \mathbb{E}((I_s^M - I_s)^{2p}) ds$
- $b_M = T^{2p-1} \int_0^T \mathbb{E} \left(\left(\sum_{j=1}^M \delta_j w_j S_s^{j,M} - \delta I_s \right)^{2p} \right) ds$
- $c_M = K_p T^{p-1} \int_0^T \mathbb{E} \left(\left(\sum_{j=1}^M \beta_j w_j S_s^{j,M} \sigma(s, I_s^M) - \beta I_s \sigma(s, I_s) \right)^{2p} \right) ds$
- $d_M = K_p T^{p-1} \int_0^T \mathbb{E} \left(\left(\sum_{j=1}^M (w_j S_s^{j,M} \eta_j(s, S_s^{j,M}))^2 \right)^p \right) ds$

The term a_M is the easiest one to handle :

$$a_M \leq r^{2p} T^{2p-1} \int_0^T \mathbb{E} \left(\sup_{0 \leq u \leq s} |I_u^M - I_u|^{2p} \right) ds. \quad (19)$$

Next, using assumption $(\mathcal{H}2)$ for the first inequality, Hölder's inequality for the second and lemma 4 for the third, one gets

$$\begin{aligned} d_M &= K_p T^{p-1} \int_0^T \sum_{j_1=1}^M \cdots \sum_{j_p=1}^M \mathbb{E} \left(\prod_{k=1}^p w_{j_k}^2 (S_s^{j_k, M})^2 (\eta_{j_k}(s, S_s^{j_k, M}))^2 \right) ds \\ &\leq K_p K_b^{2p} T^{p-1} \int_0^T \sum_{j_1=1}^M \cdots \sum_{j_p=1}^M \left(\prod_{k=1}^p w_{j_k}^2 \right) \mathbb{E} \left(\prod_{k=1}^p (S_s^{j_k, M})^2 \right) ds \\ &\leq K_p K_b^{2p} T^{p-1} \int_0^T \sum_{j_1=1}^M \cdots \sum_{j_p=1}^M \prod_{k=1}^p w_{j_k}^2 \left(\mathbb{E} ((S_s^{j_k, M})^{2p}) \right)^{\frac{1}{p}} ds \\ &\leq K_p K_b^{2p} T^p C_p \left(\sum_{j=1}^M w_j^2 \right)^p \end{aligned} \quad (20)$$

The same arguments enable us to control the term b_M :

$$\begin{aligned}
b_M &= T^{2p-1} \int_0^T \mathbb{E} \left(\left(\sum_{j=1}^M \delta_j w_j S_s^{j,M} - \delta I_s \right)^{2p} \right) ds \\
&\leq (2T)^{2p-1} \left(\int_0^T \mathbb{E} \left(\left(\sum_{j=1}^M \delta_j w_j S_s^{j,M} - \delta I_s^M \right)^{2p} \right) ds + \mathbb{E} \left((\delta I_s^M - \delta I_s)^{2p} \right) \right) \\
&\leq (2T)^{2p-1} \int_0^T \mathbb{E} \left(\left(\sum_{j=1}^M (\delta_j - \delta) w_j S_s^{j,M} \right)^{2p} \right) ds + (2T)^{2p-1} \delta^{2p} \int_0^T \mathbb{E} \left(\sup_{0 \leq u \leq s} |I_u^M - I_u|^{2p} \right) ds \\
&\leq 2^{2p-1} T^{2p} C_p \left(\sum_{j=1}^M w_j |\delta_j - \delta| \right)^{2p} + (2T)^{2p-1} \delta^{2p} \int_0^T \mathbb{E} \left(\sup_{0 \leq u \leq s} |I_u^M - I_u|^{2p} \right) ds.
\end{aligned} \tag{21}$$

For the remaining term c_M , we will also need the Lipschitz assumption ($\mathcal{H}3$)

$$\begin{aligned}
c_M &= K_p T^{p-1} \int_0^T \mathbb{E} \left(\left(\sum_{j=1}^M \beta_j w_j S_s^{j,M} \sigma(s, I_s^M) - \beta I_s \sigma(s, I_s) \right)^{2p} \right) ds \\
&\leq 2^{2p-1} K_p T^{p-1} \left(\int_0^T \mathbb{E} \left(\left(\sum_{j=1}^M (\beta_j - \beta) w_j S_s^{j,M} \sigma(s, I_s^M) \right)^{2p} \right) ds + \mathbb{E} \left((\beta I_s^M \sigma(s, I_s^M) - \beta I_s \sigma(s, I_s))^{2p} \right) \right) \\
&\leq 2^{2p-1} K_p T^p K_b^{2p} C_p \left(\sum_{j=1}^M w_j |\beta_j - \beta| \right)^{2p} + 2^{2p-1} K_p T^{p-1} (\beta K_\sigma)^{2p} \int_0^T \mathbb{E} \left(\sup_{0 \leq u \leq s} |I_u^M - I_u|^{2p} \right) ds.
\end{aligned} \tag{22}$$

So, combining the inequalities (19), (20), (21) and (22), one obtains

$$\begin{aligned}
\mathbb{E} \left(\sup_{0 \leq t \leq T} |I_t^M - I_t|^{2p} \right) &\leq C_0 \left(\left(\sum_{j=1}^M w_j^2 \right)^p + \left(\sum_{j=1}^M w_j |\beta_j - \beta| \right)^{2p} + \left(\sum_{j=1}^M w_j |\delta_j - \delta| \right)^{2p} \right) \\
&\quad + C_1 \int_0^T \mathbb{E} \left(\sup_{0 \leq u \leq s} |I_u^M - I_u|^{2p} \right) ds
\end{aligned}$$

with $C_0 = 8^{2p-1} T^p (T^p + K_p K_b^{2p}) C_p$ and $C_1 = 4^{2p-1} (2^{2p-1} K_p T^{p-1} (\beta K_\sigma)^{2p} + (2T)^{2p-1} \delta^{2p} + r^{2p} T^{2p-1})$.

Finally, by means of Gronwall's lemma, we conclude that

$$\mathbb{E} \left(\sup_{0 \leq t \leq T} |I_t^M - I_t|^{2p} \right) \leq C_T \left(\left(\sum_{j=1}^M w_j^2 \right)^p + \left(\sum_{j=1}^M w_j |\beta_j - \beta| \right)^{2p} + \left(\sum_{j=1}^M w_j |\delta_j - \delta| \right)^{2p} \right)$$

where

$$C_T = C_0 e^{C_1 T}.$$

□

Proof of Theorem 2 : The proof is similar to the previous one :

$$\begin{aligned} |S_t^{j,M} - S_t^j|^{2p} &\leq 3^{2p-1} \left((r - \delta_j)^{2p} t^{2p-1} \int_0^t (S_s^{j,M} - S_s^j)^{2p} ds + \left| \int_0^t (S_s^{j,M} \eta_j(s, S_s^{j,M}) - S_s^j \eta_j(s, S_s^j)) dW_s^j \right|^{2p} \right. \\ &\quad \left. + \beta_j^{2p} \left| \int_0^t (S_s^{j,M} \sigma(s, I_s^M) - S_s^j \sigma(s, I_s)) dB_s \right|^{2p} \right) \end{aligned}$$

hence, using the Burkholder-Davis-Gundy inequality, there exists a constant K_p such that

$$\begin{aligned} \mathbb{E} \left(\sup_{0 \leq t \leq T} |S_t^{j,M} - S_t^j|^{2p} \right) &\leq 3^{2p-1} \left((r - \delta_j)^{2p} T^{2p-1} \int_0^T \mathbb{E} \left(\sup_{0 \leq u \leq s} |S_u^{j,M} - S_u^j|^2 \right) ds \right. \\ &\quad + K_p T^{p-1} \int_0^T \mathbb{E} \left((S_s^{j,M} \eta_j(s, S_s^{j,M}) - S_s^j \eta_j(s, S_s^j))^{2p} \right) ds \\ &\quad \left. + K_p T^{p-1} \beta_j^{2p} \int_0^T \mathbb{E} \left((S_s^{j,M} \sigma(s, I_s^M) - S_s^j \sigma(s, I_s))^{2p} \right) ds \right) \end{aligned}$$

Using assumption $(\mathcal{H}4)$, one gets

$$\int_0^T \mathbb{E} \left((S_s^{j,M} \eta_j(s, S_s^{j,M}) - S_s^j \eta_j(s, S_s^j))^{2p} \right) ds \leq K_\eta^{2p} \int_0^T \mathbb{E} \left(\sup_{0 \leq u \leq s} |S_u^{j,M} - S_u^j|^{2p} \right) ds.$$

Finally, by means of lemma 4 and assumptions $(\mathcal{H}2)$ and $(\mathcal{H}3)$,

$$\begin{aligned} \int_0^T \mathbb{E} \left((S_s^{j,M} \sigma(s, I_s^M) - S_s^j \sigma(s, I_s))^{2p} \right) ds &\leq 2^{2p-1} \int_0^T \mathbb{E} \left((S_s^{j,M})^{2p} (\sigma(s, I_s^M) - \sigma(s, I_s))^{2p} \right) ds \\ &\quad + 2^{2p-1} \int_0^T \mathbb{E} \left((\sigma(s, I_s))^{2p} (S_s^{j,M} - S_s^j)^{2p} \right) ds \\ &\leq 2^{2p-1} C_{2p}^{\frac{1}{2}} K_{Lip}^{2p} T \sqrt{\mathbb{E} \left(\sup_{0 \leq t \leq T} |I_t^M - I_t|^{4p} \right)} \\ &\quad + 2^{2p-1} K_b^{2p} \int_0^T \mathbb{E} \left(\sup_{0 \leq t \leq T} |S_t^{j,M} - S_t^j|^{2p} \right) ds \end{aligned}$$

We deduce using Gronwall's lemma :

$$\mathbb{E} \left(\sup_{0 \leq t \leq T} |S_t^{j,M} - S_t^j|^{2p} \right) \leq \tilde{C}_T^j \sqrt{\mathbb{E} \left(\sup_{0 \leq t \leq T} |I_t^M - I_t|^{4p} \right)}$$

where

$$\tilde{C}_T^j = 6^{2p-1} K_p T^p \beta_j^{2p} C_{2p}^{\frac{1}{2}} K_{Lip}^{2p} e^{3^{2p-1} ((r - \delta_j)^{2p} T^{2p-1} + K_p T^{p-1} K_\eta^{2p} + 2^{2p-1} K_p T^{p-1} \beta_j^{2p} K_b^{2p}) T}.$$

We conclude by Theorem 1 and the sublinearity of the square root function on \mathbb{R}_+ .

We now turn to the L^{2p} -distance between I^M and \bar{I}^M :

$$\begin{aligned} |I_t^M - \bar{I}_t^M|^{2p} &= \left| \sum_{j=1}^M w_j S_t^{j,M} - \sum_{j=1}^M w_j S_t^j \right|^{2p} \\ &\leq \left(\sum_{j=1}^M w_j |S_t^{j,M} - S_t^j| \right)^{2p} \\ &\leq \sum_{j_1=1}^M \cdots \sum_{j_{2p}=1}^M \prod_{k=1}^{2p} w_{j_k} |S_t^{j_k, M} - S_t^{j_k}| \end{aligned}$$

So, using Hölder inequality, one has

$$\begin{aligned}
\mathbb{E} \left(\sup_{0 \leq t \leq T} |I_t^M - \bar{I}_t^M|^{2p} \right) &\leq \sum_{j_1=1}^M \cdots \sum_{j_{2p}=1}^M \left(\prod_{k=1}^{2p} w_{j_k} \right) \prod_{k=1}^{2p} \left(\mathbb{E} \left(\sup_{0 \leq t \leq T} |S_t^{j_k, M} - S_t^{j_k}|^{2p} \right) \right)^{\frac{1}{2p}} \\
&\leq \left(\sum_{j=1}^M w_j \right)^{2p} \max_{1 \leq j \leq M} \tilde{C}_T^j \left(\left(\sum_{j=1}^M w_j^2 \right)^p + \right. \\
&\quad \left. \left(\sum_{j=1}^M w_j |\beta_j - \beta| \right)^{2p} + \left(\sum_{j=1}^M w_j |\delta_j - \delta| \right)^{2p} \right).
\end{aligned}$$

□