Alpha Models

O. Croissant
Natixis,Fixed Income Dpt

October 17, 2008

Abstract

We present the alpha models and the technology associated with the pricing of vanilla and a few exotics

0.1 Purpose

This document explains how the alpha model technology helps solving the problem of mixing smiles of different underlyings to compute option on spreads, digitals associated with these options that pays float and other exotics.We present also the calibrating methodology that derives from this technology

0.2 Introduction

The alpha model technology grew from sevral attempts to compute spreadoption prices in models where the underlyings are modelled by smiled processes. We mean by smiled positive processes, processes such that ordinary options on underlyings modeled by such processes exhibit a non constant lognormal smile.

Different possibilities have been tried:

- Sabr processes linked by a gaussian copula
- Sabr processes linked by a student copula
- Sabr processes linked by a BiSabr copula
- Heston processes projected inside a wishart process.

Let's review the different alternatives , pointing out the strength and weakness of each.

0.2.1 Sabr processes linked by a gaussian copula

The copula is parametrized by one number: the gaussian correlation that is measured between the two underlying when they are mapped to a normal variable. It is different from an instantaneous correlation that you can model from their processes because it it measured at a certain maturity. So the gaussian correlation shoud be provided for every maturity. It is not a process. The correlation is also different from a historical correlation that you can measure

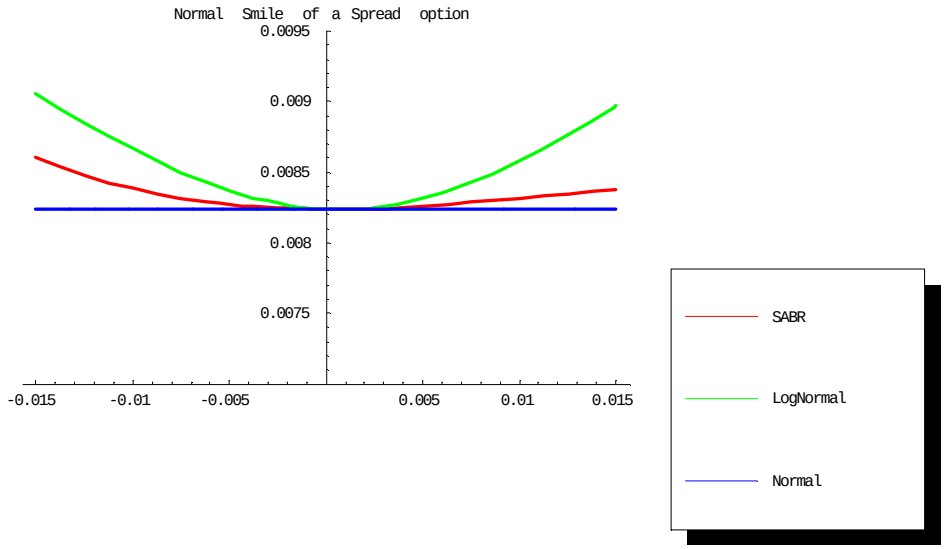

Figure 1:

through a monte carlo experiment for exemple.Therefore it has to be calculated for each maturity.

In terms of modeling power , The fact that SABR variable are somwhere between Lognormal and normal variable when it comes to tail indexes, implies that tha convexity of the the gaussian copula smiles also are between the those two smiles, as illustrated by the preceding figure. Also, we have no control on the convexity and the skew of the spreadoption smile.

Vanilla spreadoption can be expressed as numerical double integrals using the gaussian bidimensional kernel and 2 SABR quantile functions. It is better to use the following formulas (one dimension numerical integrals):

$$Call = E[(S_1 - S_1 - K)^+] = Digital(S_1) - Digital(S_2) - K \times Digital(K) \quad (1)$$

with

$$\begin{aligned} Digital(S_1) &= E[S_1 1_{S_1 - S_1 - K \geq 1}] = & (2) \\ & \int_{-\infty}^{\infty} S_1 dS_1 \partial_1 C(F_1(S_1), F_2(S_1 - K))) \frac{dF_1}{dS_1}(S_1) & (3) \end{aligned}$$

and

$$Digital(S_2) = E[S_2 1_{S_1 - S_1 - K \geq 1}] = \quad (4)$$

$$\int_{-\infty}^{\infty} S_2 dS_2 (1 - \partial_2 C(F_1(S_2 + K), F_2(S_2))) \frac{dF_2}{dS_2}(S_2) \quad (5)$$

and

$$Digital(1) = E[1_{S_1 - S_1 - K \geq 1}] = \quad (6)$$

$$\int_{-\infty}^{\infty} dS_2 (1 - \partial_2 C(F_1(S_2 + K), F_2(S_2))) \frac{dF_2}{dS_2}(S_2) \quad (7)$$

where

$$\partial_1 C(X_1, X_2)) = N[\frac{N^{-1}(X_2) - \rho N^{-1}(X_1)}{\sqrt{1 - \rho^2}}] \quad (8)$$

and

$$\partial_2 C(X_1, X_2)) = \partial_1 C(X_2, X_1)) = N[\frac{N^{-1}(X_1) - \rho N^{-1}(X_2)}{\sqrt{1 - \rho^2}}] \quad (9)$$

0.2.2 - Sabr processes linked by a student copula

The student copula is defined by two parameters :

-the correlation parameter ρ

-the degre parameter ν

by varying the degre , we can introduce more fat tails, as the following picture show us:

the Vanilla calls are computed using the double numerical integral:

$$Call = \int \int dY_1 dY_2 f(S_1^{-1}(D_\nu(Y_1)), S_2^{-1}(D_\nu(Y_2)) d(Y_1, Y_2, \rho, \nu) \quad (10)$$

where S_1^{-1}andS_2^{-1} are the two quantile functions associated with the two SABR underlyings and the cumulative student distribution function is given by

$$D_\nu(X) = \frac{1}{2}(1 + \text{Sign}(X) Beta_R(\frac{\nu}{X^2 + \nu}, 1, \frac{\nu}{2}, \frac{1}{2}) \quad (11)$$

where the regularised beta function is defined by:

$$Beta_R(x, y, a, b) = \frac{Beta(x, y, a, b)}{Beta(a, b)} \quad (12)$$

and the incomplete beta function is defined by

$$Beta(x, y, a, b) = \int_x^y t^{a-1}(1 - t)^{b-1} dt \quad (13)$$

and the usual beta function given by

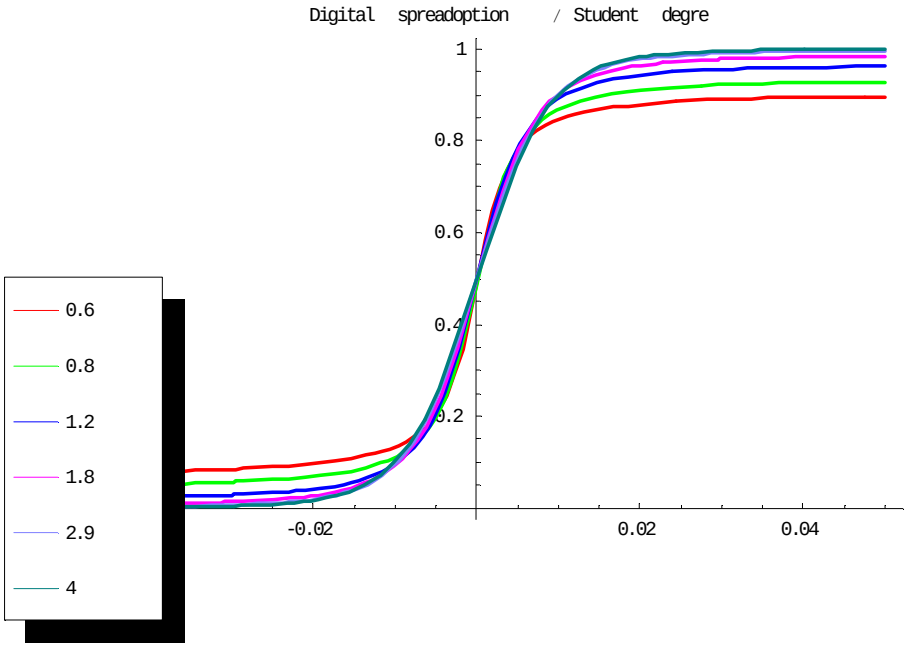
Digital spreadoption / Student degre
1
0.8
0.6
0.4
0.2
-0.02
0.02
0.04
0.6
0.8
1.2
1.8
2.9
4

Figure 2:

$$Beta(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \tag{14}$$

the joint bidimensional student density being an elliptic copula is defined by:

$$d(x,y,\rho,\nu) = \frac{\Gamma(\frac{\nu+2}{2})}{\Gamma(\frac{\nu}{2})\pi\nu\sqrt{1-\rho^2}}(1+\frac{x^2+y^2-2\rho xy}{\nu(1-\rho^2)})^{-\frac{\nu+2}{2}} \tag{15}$$

the problem of student copula is that, if it give us some flexibility to module the convexity of the correlation smile, it does not give us a way to control the slope of the correlation smile. This is bad because the slope is an order higher than the convexity.

0.2.3 - Sabr processes linked by a BiSabr copula

The underlying processes are driven by 4 brownian motions the following way:

$$\begin{array}{c} dS_1 = \alpha_1 S_1^{\beta_1} dW_1 \\ dS_2 = \alpha_2 S_2^{\beta_2} dW_2 \\ d\alpha_1 = \nu_1 \alpha_1 dW_3 \\ d\alpha_2 = \nu_2 \alpha_2 dW_4 \end{array} \tag{16}$$

with the 2 correlations that determines (pour part) the skew of each processes

$$\begin{array}{c} dW_1.dW_3 = \rho_{13}dt \\ dW_2.dW_4 = \rho_{24}dt \end{array} \tag{17}$$

and the additional 4 correlations that determines the copula

$$\begin{array}{c} dW_1.dW_2 = \rho_{12}dt \\ dW_3.dW_4 = \rho_{34}dt \\ dW_1.dW_4 = \rho_{14}dt \\ dW_2.dW_3 = \rho_{23}dt \end{array} \tag{18}$$

The role of each these copula-correlations are the following:
ρ_{12} determine the general level of implicit volatility for ATM spreadoption.
$\rho_{14}$and ρ_{23}acts on the slope of the correlation smile .
ρ_{34} condition slightly the convexity of the correlation smile

The bisabr option formula structure follows it line of demonstration. First by singular perturbation we drop 2 dimensions,
so we determine the coefficients $A_{s_1s_1}, A_{\alpha_{11}s_1}, A_{\alpha_{11}\alpha_{11}}$and $A_{\alpha_{11}}$
such that the solution of the following differential equation :

$$-\partial_t P = \frac{1}{2}\ A_{ss}\frac{\partial^2 P}{\partial s^2} + A_{\alpha_{11}s}\frac{\partial^2 P}{\partial \alpha_1 1}\partial s_1 + A_{\alpha_{11}\alpha_{11}}\frac{\partial^2 P}{\partial {\alpha_{11}}^2} + A_{\alpha_{11}}\frac{\partial P}{\partial \alpha_{11}} \tag{19}$$

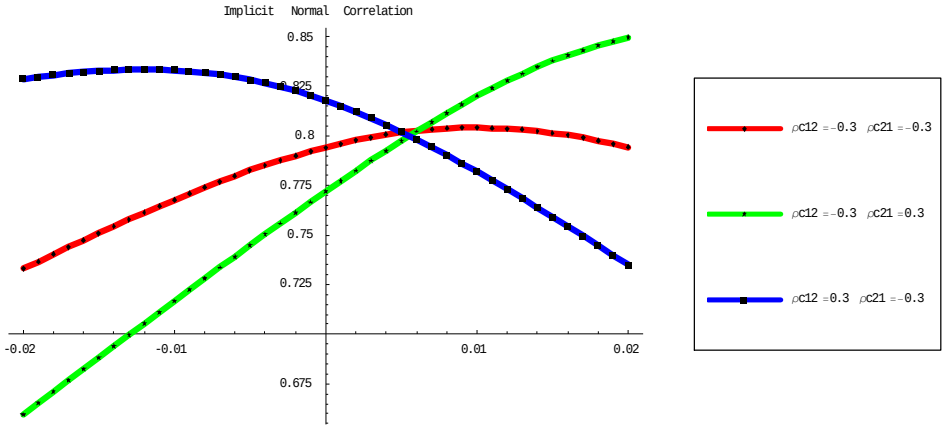

Figure 3:

with initial conditions

$$P[0] = (\frac{1}{2}\epsilon^2 {\alpha_1}^2 (K+b)^{2\beta_1} + \frac{1}{2}\epsilon^2 {\alpha_2}^2 b^{2\beta_2} - \epsilon^2 \rho_s \alpha_1 \alpha_2 (-K+b)^{\beta_1} b^{\beta_2}) \delta[s-K] \quad (20)$$

is the cal with strike K

$$V(t_{\text{tex}}, S_1, S_2, \alpha_1, \alpha_2) = (S_1 - S_2 - K)^+ + \int_0^{t_{\text{tex}}} P\left[t, s+b, b, \alpha_1, \alpha_2\right] dT \quad (21)$$

where we reparametrize the coordinates by :

$$\begin{aligned} s &= S_1 - S_2; && (22)\\ b &= S_2; && (23)\\ \alpha_{11} &= \sqrt{{\alpha_1}^2 (K+b)^{2\beta_1} + {\alpha_2}^2 b^{2\beta_2} - 2\rho_s \alpha_1 \alpha_2 (K+b)^{\beta_1} b^{\beta_2}} && (24) \end{aligned}$$

This bidimensional equation is then handled by the techniques described in Patrick Hagan Documents, by doing a further singular perturbation and reducing the problem to a local volatility problem which is solved using the fact that the associated implied volatility is the harmonique average of the local volatility of the problem.

Strength of the BiSABR approch. The BiSABR Copula is parametrized by very intuitive parameters and the phenomenology is correct.

Weakness of the BiSABR approach The formula is not perfectible, like the SABR formula, because the order reasonning which at the basis of the singular perturbation is not extensible.Iti is a blackbox with a validity domain.

If it is very sensible for short to medium term option, Long term options (20-30 years) begin to show calibration problems.

obtaining the bidimensional formula was a kind of tour de force .

We do not know how to address 3 dimensional issues (like two indexes spread-option paying a third index)

0.2.4 - Heston processes projected inside a Wishart process.

Here we use a multidimensional generalisation of the CIR process which at the basis of the Heston proccess.

And we use the matrix value to make it Multi Asset.

More precisely the Volatility sector is modeled by a matrix process V:

$$dV = (\beta Q^T Q + MV + VM)dt + \sqrt{V}dWQ + Q^T dW^T \sqrt{V} \tag{25}$$

with Q, M are matrix parameters, dW is a matrix brownian motion

The Asset Sector is modeled by a vector Process

$$dS_i = Trace[A_i(\sqrt{V}dWR + dZ\sqrt{I_n - RR^T})] \tag{26}$$

Where the matrix parametre R has a spectral radius <1

and Trace is the usual trace and the matrix Parameter A_i is specific to each asset

β is a positive number that induce convexity in the asset smile and correlation smile as well.

Strength of the Wishart approch. The Whishart approach is the maximaly flexible class of process that generalizes the Heston process.It Allows to have an arbitrary number of factors that influences the assets prices while inthe same time it offers correlations between these factors with some constraints. Its tractability rely on the existence of fourier transforms for the transition probabilities.

When it comes to pricing spreadoptions, this implies a bidimensional fourier transform of expression that are solution of multidimentional riccati equation and that can be represented as exponential of matrices, therfore efficiently computable.

Weakness of the Wishart approach The computation of the bidimensional inverse fourier transform can be very tricky, because of carracteristics like rapid oscillations and infinite branches that for monodimnentional case can be handled by numerical adaptative techniques, butr for the bidimensional cases are more delicate.

The main problem of wishart based process is the gyndinkin domain. In a nutshell we can say that the market asset need more convexity than what is representable in the gyndinkin domain, therefore safely computable

0.2.5 Alpha processes .

Alpha processes are just a particular case of whishart based processes, in wich we forget about the correlation between the different factors of volatility.

So they are of the type:

$$Z_t = f(X_t) \tag{27}$$

Where f is any differentiable function and X_t is a p dimension stochastic process

$$dX_i = (\sum_{j=1}^{p} \beta_{i,j} X_j + \sum_{k=1}^{n} \gamma_{i,j} V_k)dt + \sum_{k=0}^{m} \alpha_{i,k} \sqrt{V_k} dW_k(t) \tag{28}$$

and

$$\begin{array}{c} dV_k = \lambda_k(1 - V_k)dt + \nu_k \sqrt{V_k} d\bar{W}_k(t) \\ dW_k(t).d\bar{W}_k(t) = \rho_k dt \end{array} \tag{29}$$

where $V_k(0) = 1$ and $1 \leq i \leq m$

and $\nu_0 = 0$: So the case $k = 0$ represents the residual for which we do not have information.

But it comes also with a set of approximating assumptions and a few analytical results that make them unique.

The assumption is that the behaviour of the second order chaotic expansion is representative of the process. Then projection of the process on one dimension affine process is performed. This step can be avoided for the assets itself. it unavoidable for the spreadoptions

Strength of the Alpha approch. Universality: Thanks to the projection mecanism, any model can be integrated in the alpha methodology.Of course starting from pur Alpha Assets give the precise modeling.

The Exact change of numeraire formula, gives access to aplicability in building a full Interest rate model and hybrid FX-FixedIncome models as it will be seen later.

Modularity. Pricing rely on projection mechanism that can be improved or changed on purpose without affecting the general framework.The equivalence principle allows to use algebra of second order chaos expansions .

Robustness As soon the right representant of the process is determined, the pricing is performed by a Heston formula or a normal heston formula , which give us a control on problems tha can happen when strikes, maturity or parameters are pushed beyond usual values

Weakness of the Alpha approach For too long maturity or too extrem situations, the approximation may not be valid. Third order and higher order chaos may become important. Meaning for exemple that the reduction of a smile to its ATM behaviour may not be possible because a change of convexity of the smile or the farness of the strike we are considering.

1 Pricings in the Alpha Models

1.1 Presentation of the Alpha Methodology

The alpha methodology follows the Principal Component Analysis metaphor, applied to the smile, seen as representing a distribution.

Like the PCA it offers a approximation based view of nature, where we can add as many factor we want to better describe the set of prices we are looking at.

But instead of being restricted to only one measure of the distribution (variance) we are looking at the whole distribution. Therefore it offers a perturbative approach of the joint distribution of asset as well.Therefore we can also develop a perturbative approach of the copula .

To generate the smile we use a set a CIR that are independant of each other, but correlated to one and only one of the drivers of teh asset.

Every CIR is defined by an equation :

$$\begin{array}{c} dV_k = \lambda_k(1 - V_k)dt + \nu_k\sqrt{V_k(t)}d\bar{W}_k(t) \\ dW_k(t).d\bar{W}_k(t) = \rho_k dt \end{array} \tag{30}$$

so it is defined by two numbers :

$$\begin{array}{c} \nu_i > 0 \\ -1 < \rho_k < 1 \end{array} \tag{31}$$

these two numbers can be amalgamated into a single complex number called the color of the noise $\sqrt{V_k(t)}d\bar{W}_k(t)$:

$$C_k = \nu_k(e^{i\pi\rho_k}) \tag{32}$$

λ_k do not have the same statut and is assumed constant across the all framework for now.

$$\lambda_k = \lambda \tag{33}$$

1.1.1 Alpha Models for Assets in the delta neutral measure

With no surprise, all asset are assumed to be a martingale of the following form

$$\frac{dS_i}{S_i} = \sum_{k=0}^{m} \alpha_{i,k} dN_{k,C_k} \tag{34}$$

where each noise $dN_k$is defined by

$$dN_{k,C_k}(t) = \sqrt{V_k(t)}dW_k(t) \tag{35}$$

Each V_k implementing its color C_k.

This decomposition is called the spectral sequence associated with the asset S_i and is noted

$$\frac{dS_i}{S_i} \bowtie \{\alpha_{i,k}, C_k\}_{0\leq k\leq n} \tag{36}$$

It is costumary to associate the dimension 0 with a special color=0. It represents the residual with no specific structure in a PCA analysis.

1.1.2 Alpha Models for spreads in the delta neutral measure

We model $s_{i,j} = S_i - S_j$

$$ds_{i,j}(t) = \sum_{k=0}^{m} \alpha_{i,j,k} dN_{k,C_k}(t) \tag{37}$$

the projection procedure that will be detailed later aims to determine $\alpha_{i,j,k}$ in fonction of $\alpha_{i,k}$ and $\alpha_{j,k}$

The sequence spectral is defined by:

$$ds_{i,j}(t) \bowtie \{\alpha_{i,j,k}, C_k\}_{0\leq k\leq n} \tag{38}$$

1.1.3 Alpha Models for Assets in the measure induced by a new numeraire

In many cases, we will have to change of numeraire, The Alpha framework is robust to change of measure as we show here:

let

$$\frac{dX}{X} = (\sum_{k=0}^{m} \gamma_k V_k)dt + \sum_{k=0}^{m} \alpha_k \sqrt{V_k} dW_k(t) \tag{39}$$

We extend the spectral sequence, by including the coeficient of the drift that maintain it insided affine framework

$$\frac{dX}{X} = \bowtie \{\gamma_k, \alpha_k, C_k\}_{0\leq k\leq n} \tag{40}$$

or

$$\frac{dX}{X} = (\sum_{k=0}^{m} \gamma_k V_k)dt + \sum_{k=0}^{m} \alpha_k \sqrt{V_k} dW_k(t) \tag{41}$$

Taking the log :

$$dLog(X) = (\sum_{k=0}^{m} \gamma_k - \frac{\alpha_k^2}{2} V_k)dt + \sum_{k=0}^{m} \alpha_k \sqrt{V_k} dW_k(t) \tag{42}$$

$$dLog(S_i) = (\sum_{k=0}^{m} -\frac{\alpha_{i,k}^2}{2} V_k)dt + \sum_{k=0}^{m} \alpha_{i,k} \sqrt{V_k} dW_k(t) \tag{43}$$

we compute

$$dLog(\frac{S_i}{X}) = (\sum_{k=0}^{m} (\frac{\alpha_k^2}{2} - \frac{\alpha_{i,k}^2}{2} - \gamma_k) V_k)dt + \sum_{k=0}^{m} (\alpha_{i,k} - \alpha_k) \sqrt{V_k} dW_k(t) \tag{44}$$

$$\frac{d\frac{S_i}{X}}{\frac{S_i}{X}} = (\sum_{k=0}^{m} (\alpha_k^2 - \alpha_{i,k}\alpha_k - \gamma_k) V_k)dt + \sum_{k=0}^{m} (\alpha_{i,k} - \alpha_k) \sqrt{V_k} dW_k(t) \tag{45}$$

we want the drift under the change of measure :

$$dW_k = d\hat{W}_k + \Lambda_k dt \tag{46}$$

be 0,
so

$$\sum_{k=0}^{m} (\alpha_k^2 - \alpha_{i,k}\alpha_k - \gamma_k) V_k + \sum_{k=0}^{m} (\alpha_{i,k} - \alpha_k) \sqrt{V_k} \Lambda_k = 0 \tag{47}$$

that can localize in

$$(\alpha_k^2 - \alpha_{i,k}\alpha_k - \gamma_k) \sqrt{V_k} + (\alpha_{i,k} - \alpha_k) \Lambda_k = 0 \tag{48}$$

Generally, this can be implemented as

$$\Lambda_k = (\alpha_k - \frac{\gamma_k}{(\alpha_k - \alpha_{i,k})}) \sqrt{V_k} \tag{49}$$

so we get the dynamic under this new numeraire as:

$$\frac{dS_i}{S_i} = (\sum_{k=0}^{m} \alpha_{i,k} (\alpha_k - \frac{\gamma_k}{(\alpha_k - \alpha_{i,k})}) V_k)dt + \sum_{k=0}^{m} \alpha_{i,k} d\hat{W}_k \sqrt{V_k} \tag{50}$$

which is associated with

$$\frac{dS_i}{S_i} \bowtie \{\alpha_{i,k} (\alpha_k - \frac{\gamma_k}{(\alpha_k - \alpha_{i,k})}), \alpha_{i,k}, C_k\}_{0 \leq k \leq n} \tag{51}$$

we create an operation on the spectral sequences:

$$\{\alpha_{i,k}, C_k\}_{0 \leq k \leq n} \circledast \{\gamma_k, \alpha_k, C_k\}_{0 \leq k \leq n} = \{\alpha_{i,k} (\alpha_k - \frac{\gamma_k}{(\alpha_k - \alpha_{i,k})}), \alpha_{i,k}, C_k\}_{0 \leq k \leq n} \tag{52}$$

we can check that it works for the Bilognormal case:
for the change of Asset1 under the Asset1 numeraire:

$$\{\sigma_1, 0\} \circledast \{0, \sigma_1, 0\} = \{\sigma_1^2, \sigma_1, 0\} \tag{53}$$

for the change of Asset1 under the Asset2 numeraire:

$$\{\sigma_1, 0\} \circledast \{\{0, \rho\sigma_2, 0\}, \{0, \sqrt{1-\rho^2}\sigma_2, 0\}\} = \{\rho\sigma_1\sigma_2, \sigma_1, 0\} \tag{54}$$

where by convention $\{\{\sigma_1, 0\}, \{0, 0\}\}$ is noted $\{\sigma_1, 0\}$

1.2 The Alpha spreadoption projection technique

1.2.1 Second order Simple Gaussian Heston Development

Let a simple gaussian process

$$dS \bowtie \{\alpha, C\} \tag{55}$$

using Ito, we can write the first order integral developement

$$\begin{array}{c} S(t) = S_0 + \int_0^t \alpha\sqrt{V(u)}dW(u) \\ \sqrt{V(t)} = 1 + \lambda\int_0^t \frac{(1-V(u))}{2\sqrt{V(u)}}du + \nu\int_0^t \frac{\rho dW(u)+\sqrt{1-\rho^2}d\hat{W}(u)}{2\sqrt{V(u)}} \end{array} \tag{56}$$

by reinjecting the second equation inside the first one , we get

$$S(t) = S_0 + \int_0^t \alpha(1 + \lambda\int_0^h \frac{(1-V(u))}{2\sqrt{V(u)}}du + \tag{57}$$

$$\nu\int_0^h \frac{\rho dW(u) + \sqrt{1-\rho^2}d\hat{W}(u)}{2\sqrt{V(u)}})dW(h) \tag{58}$$

so by reinjecting iteratively the expression of $\sqrt{V(t)}$ and taking into account that du is of order 1 and dW is of order $1/2$, we can get approximations to order n in t

the second order approximation can be written as:

$$S(t) = S_0 + \int_0^t \alpha dW(h) + \frac{\alpha\nu}{2}\rho\int_0^t\int_0^h dW(u)dW(h) + \tag{59}$$

$$\frac{\alpha\nu}{2}\sqrt{1-\rho^2}\int_0^t\int_0^h d\hat{W}(u)dW(h) \tag{60}$$

where we see the first order approximation

$$\Delta^{(1)}S(t) = \alpha\int_0^t dW(h) \tag{61}$$

At first order the Base process is the brownian motion $W(t)$

$$B^{(1,1)}(t) = \int_0^t dW(h) \tag{62}$$

and the second order:

$$\Delta^{(2)}S(t) = \frac{\alpha\nu}{2}\rho\int_0^t\int_0^h dW(u)dW(h) + \frac{\alpha\nu}{2}\sqrt{1-\rho^2}\int_0^t\int_0^h d\hat{W}(u)dW(h) \tag{63}$$

So the two second order Base processes are :

$$\begin{array}{c} B^{(2,1)}(t) = \int_0^t\int_0^h dW(u)dW(h) \\ B^{(2,2)}(t) = \int_0^t\int_0^h d\hat{W}(u)dW(h) \end{array} \tag{64}$$

and we have the first correction after the gaussian approximation is::

$$\Delta^{(2)}S(t) = \frac{\alpha\nu}{2}\rho B^{(2,1)}(t) + \frac{\alpha\nu}{2}\sqrt{1-\rho^2}B^{(2,2)}(t) \tag{65}$$

where it is known that

$$\begin{array}{c} E[B^{(2,1)}(t)] = E[B^{(2,2)}(t)] = E[B^{(2,1)}(t)B^{(2,2)}(t)] = \\ E[B^{(2,1)}(t)B^{(1,1)}(t)] = E[B^{(2,2)}(t)B^{(1,1)}(t)] = 0 \\ E[B^{(1,1)}(t)^2] = t \\ E[B^{(2,1)}(t)^2] = E[B^{(2,2)}(t)^2] = t^2 \end{array} \tag{66}$$

the distribution of $B^{(2,1)}(t)$ is a Chi2, the distribution of $B^{(2,2)}(t)$ is not known

The linearity of the development implies that

$$\begin{array}{c} E[S(t)B^{(1,1)}(t)] = \alpha t \\ E[S(t)B^{(2,1)}(t)] = \frac{\alpha\nu}{2}\rho t^2 \\ E[S(t)B^{(2,1)}(t)] = \frac{\alpha\nu}{2}\sqrt{1-\rho^2}t^2 \end{array} \tag{67}$$

We are also interested in what we call statistics or observables of the processes. We mean by that quantities that are not dependent on the dynamics for their definition, but that can be used for infering the parameters of any raisonable dynamics.

We compute:

from

$$\begin{array}{c} dS(t) = \alpha\sqrt{V(u)}dW(u) \\ dV(t) = \lambda(1-V(u))dt + \nu(\rho dW(u) + \sqrt{1-\rho^2}d\hat{W}(u)) \end{array} \tag{68}$$

we deduce

$$Var = \langle dS, dS\rangle = \alpha^2 V \tag{69}$$

$$dVar = \alpha^2 dV = \alpha^2\nu(\rho dW(u) + \sqrt{1-\rho^2}d\hat{W}(u)) \tag{70}$$

then the observables are:

$$\frac{\langle dVar, dVar\rangle}{dt} = \alpha^4\nu^2 \tag{71}$$

$$\frac{\langle dVar, dS\rangle}{dt} = \alpha^3\nu\rho \tag{72}$$

$$\frac{\langle dS, dS\rangle}{dt} = \alpha^2 \tag{73}$$

1.2.2 Second order Spreadoption Development

We assume two assets following

$$\begin{array}{c} \frac{dS_1}{S_1} \bowtie \{\alpha_{1,k}, C_{1,i}\}_{0\le k\le n} \\ \frac{dS_2}{S_2} \bowtie \{\alpha_{2,k}, C_{2,i}\}_{0\le k\le n} \end{array} \tag{74}$$

so we have:

$$\begin{array}{c} S_i(t) = S_{i,0} + \int_0^t \sum \alpha_{i,k} S_i(u)\sqrt{V_k(u)}dW_k(u) \\ \sqrt{V_k(t)} = 1 + \nu_k \int_0^t \frac{\rho_k dW_k(u) + \sqrt{1-\rho_k^2}d\hat{W}_k(u)}{2\sqrt{V_k(u)}} + \lambda\int_0^t \frac{(1-V_k(u))}{2\sqrt{V_k(u)}}du \end{array} \tag{75}$$

from which we deduce by reinjection:

$$S_i(t) = S_{i,0} + \int_0^t \sum_k \alpha_{i,k}(S_{i,0} + \tag{76}$$

$$\int_0^u \sum_m \alpha_{i,m} S_i(h)\sqrt{V_m(h)}dW_m(h)). \tag{77}$$

$$(1 + \nu_k \int_0^u \frac{\rho dW_k(l) + \sqrt{1-\rho_k^2}d\hat{W}_k(l)}{2\sqrt{V_k(l)}})dW_k(l) \tag{78}$$

then we expand keeping only the second order in dW

$$S_i(t) = S_{i,0} + \int_0^t \sum_k \alpha_{i,k} S_{i,0} dW_k(u) + \tag{79}$$

$$\int_0^t \sum_k \alpha_{i,k} S_{i,0}\nu_k \int_0^u \frac{\rho dW_k(h) + \sqrt{1-\rho_k^2}d\hat{W}_k(h)}{2\sqrt{V_k(h)}})dW_k(u) \tag{80}$$

$$+ \int_0^t \sum_k \alpha_{i,k} S_{i,0} \int_0^u \sum_m \alpha_{i,m} S_i(h)\sqrt{V_m(h)}dW_m(h))dW_k(u) \tag{81}$$

so we get the approximation:

$$S_i(t) = S_{i,0} + \int_0^t \sum_k \alpha_{i,k} S_{i,0} dW_k(u) + \quad (82)$$

$$\sum_k \frac{\alpha_{i,k} S_{i,0} \nu_k}{2\sqrt{V_{k,0}}} \int_0^t \int_0^u (\rho_k dW_k(h) + \sqrt{1-\rho_k^2} d\hat{W}_k(h)) dW_k(u) \quad (83)$$

$$+ \sum_k \sum_m \alpha_{i,m} \alpha_{i,k} S_{i,0}^2 \sqrt{V_{m,0}} \int_0^t \int_0^u dW_m(h)) dW_k(u) \quad (84)$$

that we can rename

$$S_i(t) = S_{i,0} + \sum_k \alpha_{i,k} S_{i,0} B_k^{(1)}(t) + \quad (85)$$

$$\sum_k \frac{\alpha_{i,k} S_{i,0} \nu_k}{2} (\rho_k B_k^{(2)}(t) + \sqrt{1-\rho^2} B_k^{(2,1)}(t)) \quad (86)$$

$$+ \sum_k \sum_m \alpha_{i,m} \alpha_{i,k} S_{i,0}^2 B_{k,m}^{(2)}(t) \quad (87)$$

where all $B_k^{(1)}(t), B_k^{(2)}(t) B_k^{(2,1)}(t)), B_{k,m}^{(2)}(t)$ are orthogonals
So now we can compute the second order spreadoption development:

$$S_1(t) - S_2(t) = S_{1,0} - S_{2,0} + \sum_k (\alpha_{1,k} S_{1,0} - \alpha_{2,k} S_{2,0}) B_k^{(1)}(t) + \quad (88)$$

$$\sum_k \frac{(\alpha_{1,k} S_{1,0} \nu_k - \alpha_{2,k} S_{2,0} \nu_k)}{2} (\rho_k B_k^{(2)}(t) + \quad (89)$$

$$\sqrt{1-\rho_k^2} B_k^{(2,1)}(t)) + \quad (90)$$

$$\sum_k \sum_m (\alpha_{1,m} \alpha_{1,k} S_{1,0}^2 - \alpha_{2,m} \alpha_{2,k} S_{2,0}^2) B_{k,m}^{(2)}(t) \quad (91)$$

If we look at the intantaneous variance of $S_1(t) - S_2(t)$, it follow a process for wich we can write the Ito process:

$$V_{1,2} = (dS_1 - dS_2).(dS_1 - dS_2) = \quad (92)$$

$$\sum_k (\alpha_{1,k} S_1(u) - \alpha_{2,k} S_2(u))^2 V_k(u) \quad (93)$$

its process is

$$dV_{1,2} = \sum_k (\alpha_{1,k} S_1 - \alpha_{2,k} S_2)^2 dV_k + \quad (94)$$

$$2 \sum_k (\alpha_{1,k} S_1 - \alpha_{2,k} S_2)(\alpha_{1,k} dS_1 - \alpha_{2,k} dS_2) + (...)dt \quad (95)$$

So we rewrite

$$dV_{1,2} = \sum_k (\alpha_{1,k}S_1 - \alpha_{2,k}S_2)^2 \nu_k \sqrt{V_k} d\hat{W}_k + \quad (96)$$

$$2\sum_{k,m} (\alpha_{1,k}S_1 - \alpha_{2,k}S_2)(S_1\alpha_{1,k}\alpha_{1,m} - S_2\alpha_{2,k}\alpha_{2,m})\sqrt{V_m} dW_m) \quad (97)$$

so we can write its integrated process at the first order. We ar not going to be interested by a second order because we do not look at instaneous derivatives or smiles. the only use of this development is to scaler product with S_1 and itself in order to get enough intrinseq information to be able to infer the parameters of the model:

$$V_{1,2}(t) = V_{1,2,0} + \sum_k \nu_k (\alpha_{1,k}S_{1,0} - \alpha_{2,k}S_{2,0})^2 V_k((\rho_k B_k^{(1)}(t) + \quad (98)$$

$$\sqrt{1-\rho^2}\hat{B}_k^{(1)}(t))) + \quad (99)$$

$$2\sum_{k,m} (\alpha_{1,k}S_1 - \alpha_{2,k}S_2)(S_1\alpha_{1,k}\alpha_{1,m} - \quad (100)$$

$$S_2\alpha_{2,k}\alpha_{2,m})V_m B_m^{(1)}(t) \quad (101)$$

$$V_{1,2}(t) = V_{1,2,0} + \sum_k (\nu_k (\alpha_{1,k}S_{1,0} - \alpha_{2,k}S_{2,0})^2 \rho_k + \quad (102)$$

$$2\sum_m (\alpha_{1,m}S_1 - \alpha_{2,m}S_2)(S_1\alpha_{1,k}\alpha_{1,m} - \quad (103)$$

$$S_2\alpha_{2,k}\alpha_{2,m}))V_k B_k^{(1)}(t) + \quad (104)$$

$$\sum_k \nu_k (\alpha_{1,k}S_{1,0} - \alpha_{2,k}S_{2,0})^2 V_k(\sqrt{1-\rho^2}\hat{B}_k^{(1)}(t)) \quad (105)$$

where we used independant brownian motions $B_k^{(1)}(t)$ and $\hat{B}_k^{(1)}(t)$.

where we observe that the correlation act now at the first order of the process instead of the second order for the spread itself.

$$S_{12}(t) = S_{1,0} - S_{2,0} + \sum_k (\alpha_{1,k}S_{1,0} - \alpha_{2,k}S_{2,0})V_k B_k^{(1)}(t) \quad (106)$$

So we get the following observables:

$$\frac{\langle dV_{1,2}, dV_{1,2}\rangle}{dt} = \sum_k (\nu_k (\alpha_{1,k}S_{1,0} - \alpha_{2,k}S_{2,0})^2 \rho_k + \quad (107)$$

$$2\sum_m (\alpha_{1,m}S_1 - \alpha_{2,m}S_2)(S_1\alpha_{1,k}\alpha_{1,m} - \quad (108)$$

$$S_2\alpha_{2,k}\alpha_{2,m}))^2 V_k^2 + \quad (109)$$

$$\sum_k (\nu_k (\alpha_{1,k}S_{1,0} - \alpha_{2,k}S_{2,0})^2)^2 V_k^2 (1-\rho^2) \quad (110)$$

$$\frac{\langle dV_{1,2}, dS_{1,2}\rangle}{dt} = \sum_k (\nu_k(\alpha_{1,k}S_{1,0} - \alpha_{2,k}S_{2,0})^2\rho_k + \tag{111}$$

$$2\sum_m (\alpha_{1,m}S_1 - \alpha_{2,m}S_2)(S_1\alpha_{1,k}\alpha_{1,m} - \tag{112}$$

$$S_2\alpha_{2,k}\alpha_{2,m})) \tag{113}$$

$$(\alpha_{1,k}S_{1,0} - \alpha_{2,k}S_{2,0})V_k \tag{114}$$

$$\frac{\langle dS_{1,2}, dS_{1,2}\rangle}{dt} = \sum_k (\alpha_{1,k}S_{1,0} - \alpha_{2,k}S_{2,0})^2 \tag{115}$$

1.2.3 Development around the bilog

let look at a model that we call the bilog model tangent to the Alpha model: it is defined by :

$$\tilde{S}_i(t) = S_{i,0} + \int_0^t \sum_k \alpha_{i,k}\tilde{S}_i dW_k(u) \tag{116}$$

It give us the observables when we have $\nu_i -> 0$:

$$\frac{\langle d\tilde{V}_{1,2}, d\tilde{V}_{1,2}\rangle}{dt} = 2\sum_m ((\alpha_{1,m}S_1 - \alpha_{2,m}S_2)(S_1\alpha_{1,k}\alpha_{1,m} - \tag{117}$$

$$S_2\alpha_{2,k}\alpha_{2,m}))^2 V_k^2 \tag{118}$$

$$\frac{\langle d\tilde{V}_{1,2}, d\tilde{S}_{1,2}\rangle}{dt} = 2\sum_{k,m} (\alpha_{1,m}S_1 - \alpha_{2,m}S_2)(S_1\alpha_{1,k}\alpha_{1,m} - \tag{119}$$

$$S_2\alpha_{2,k}\alpha_{2,m})(\alpha_{1,k}S_{1,0} - \alpha_{2,k}S_{2,0})\tilde{V}_k \tag{120}$$

$$\frac{\langle d\tilde{S}_{1,2}, d\tilde{S}_{1,2}\rangle}{dt} = \sum_k (\alpha_{1,k}S_{1,0} - \alpha_{2,k}S_{2,0})^2 \tag{121}$$

We want to use the tangent bilog as a control variable. therfore we need to define the relative observables:

$$\frac{\langle dS_{1,2}, dS_{1,2}\rangle^R}{dt} = \frac{\langle dS_{1,2}, dS_{1,2}\rangle}{dt} - \frac{\langle d\tilde{S}_{1,2}, d\tilde{S}_{1,2}\rangle}{dt} \tag{122}$$

we notice here that $\frac{\langle dS_{1,2}, dS_{1,2}\rangle^R}{dt} = 0$;

$$\frac{\langle dV_{1,2}, dS_{1,2}\rangle^R}{dt} = \frac{\langle dV_{1,2}, dS_{1,2}\rangle}{dt} - \frac{\langle d\tilde{V}_{1,2}, d\tilde{S}_{1,2}\rangle}{dt} \tag{123}$$

$$\frac{\langle dV_{1,2}, dV_{1,2}\rangle^R}{dt} = \frac{\langle dV_{1,2}, dV_{1,2}\rangle}{dt} - \frac{\langle d\tilde{V}_{1,2}, d\tilde{V}_{1,2}\rangle}{dt} \tag{124}$$

It is obvious that they tend to 0 when $\nu_i -> 0$:

We want to use the Gaussian Heston process as a proxy for the spread $S_{1,2}$ therefore we need the absolute observable values (V_{0H}, ν_H, ρ_H) that defines the bilog.

They are obtained by an optimisation process that minimizes:

$$\sum_{n=1}^{N}(Bilog(S_1, S_2, \sigma_1, \sigma_2, \rho, K_n, T) - GHeston(S_1 - S_2, K_n, T, V_{0H}, \nu_H, \rho_H))^2 w_n \tag{125}$$

Where the mean reversion speed is equal to λ and the longterm variance is equal to V_{0H}

Then the effective observables are defined by :

$$\frac{\langle dS_{1,2}, dS_{1,2}\rangle^{eff}}{dt} = V_{0H}^2 \tag{126}$$

$$\frac{\langle dV_{1,2}, dS_{1,2}\rangle^{eff}}{dt} = V_{0H}^3 \nu_H \rho_H + \frac{\langle dV_{1,2}, dS_{1,2}\rangle^{R}}{dt} \tag{127}$$

$$\frac{\langle dV_{1,2}, dV_{1,2}\rangle^{eff}}{dt} = V_{0H}^4 \nu_H^2 + \frac{\langle dV_{1,2}, dV_{1,2}\rangle^{R}}{dt} \tag{128}$$

We now transform back the synthetic observables into the gaussian Heston parameters by

$$V_{0eff} = V_{0H} \tag{129}$$

$$\rho_{eff} = \frac{1}{V_{0H}} \frac{\frac{\langle dV_{1,2}, dS_{1,2}\rangle^{eff}}{dt}}{\sqrt{\frac{\langle dV_{1,2}, dV_{1,2}\rangle^{eff}}{dt}}} \tag{130}$$

$$\nu_{eff} = \frac{1}{V_{0H}^2} \sqrt{\frac{\langle dV_{1,2}, dV_{1,2}\rangle^{eff}}{dt}} \tag{131}$$

Of course, we need to keep raisonable values for the ρ_{eff}, ν_{eff} parameters by

$$\rho_{final} = Min[Max[\rho_{eff}, -\rho_{extrem}], \rho_{extrem}] \tag{132}$$

$$\nu_{final} = Max[\nu_{eff}, 0] \tag{133}$$

where $\rho_{extrem} = 0.95$ in order to keep the numerical algorithm still working.It is worth mentioning that the limiting condition on ν_{eff} comes from the fact that it may happen that $\frac{\langle dV_{1,2}, dV_{1,2}\rangle^{R}}{dt} < 0$.

2 Calibrations

Use of the alpha freedom degres

CIR1	CIR2	CIR3	CIR4	CIR5	CIR6	CIR7	CIR8
-0.487131072	-0.05988321	-0.338267131	-0.491737439	-0.15034582	-0.967333522	0.999	-0.982239154
0.969986087	1.65782433	6.979499344	0.782815677	0.00100239	0.001188125	0.010587713	0.008969456
1	1	1	1	0	0.000320669	3.15254E-06	0.000903218

Figure 4:

PCA	Forward	alpha1	alpha2	alpha3	alpha4
asset1	0.048290765	0.17857584	-0.056406384	-0.026813545	-0.006500946

Figure 5:

definition We call monodimensional alpha equivalent model of a given alpha model $\frac{dS}{S} \bowtie \{\alpha_k, C_k\}_{0\leq k\leq n}$ the model defined by : $\frac{d\tilde{S}}{\tilde{S}} = \bowtie \{\sqrt{\sum_{k=1}^{n}\alpha_k^2}, C\}$ and C is defined by $\{\rho, \nu\}$ where

$$\frac{\langle dVar, dVar\rangle}{dt} = S^4\sum_k(\nu_k(\alpha_k)^2\rho_k + 2\sum_m(\alpha_m^2)\alpha_k)V_k^2 + S^4\sum_k(\nu_k(\alpha_k)^2V_k^2(1-\rho_k^2) \tag{134}$$

$$\frac{\langle dVar, dS\rangle}{dt} = S^3\sum_k(\nu_k(\alpha_k)^2\rho_k + 2\alpha_k\sum_m\alpha_m^2)V_k \tag{135}$$

$$\frac{\langle dS, dS\rangle}{dt} = S^2\sum_k(\alpha_k)^2 \tag{136}$$

and

$$V_0 = \sqrt{Var} = \sqrt{\frac{\langle dS, dS\rangle}{dt}} \tag{137}$$

$$\rho = \frac{1}{\sqrt{\frac{\langle dS,dS\rangle}{dt}}}\frac{\frac{\langle dVar,dS\rangle}{dt}}{\sqrt{\frac{\langle dVar,dVar\rangle}{dt}}} \tag{138}$$

$$\nu = \frac{1}{\frac{\langle dS,dS\rangle}{dt}}\sqrt{\frac{\langle dVar, dVar\rangle}{dt}} \tag{139}$$

We take a typical alpha situation :
with the alphas:

We compute the Smile for vanilla call at different maturities:
We observe that the equivalent lognormal heston has a smile very close to its equivalent lognormal smile.

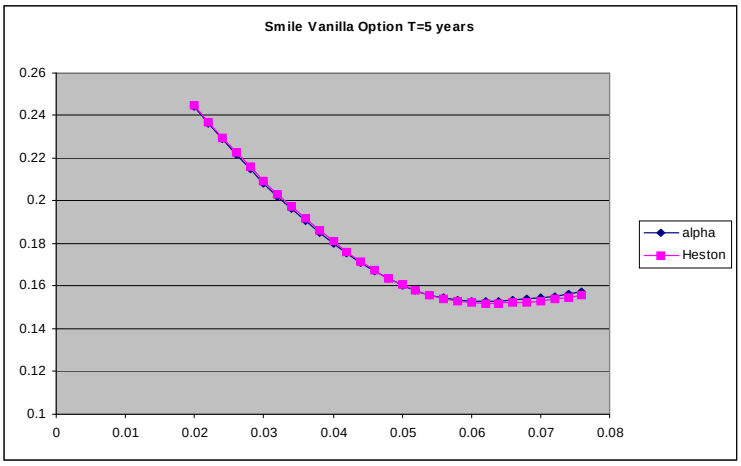

Figure 6:

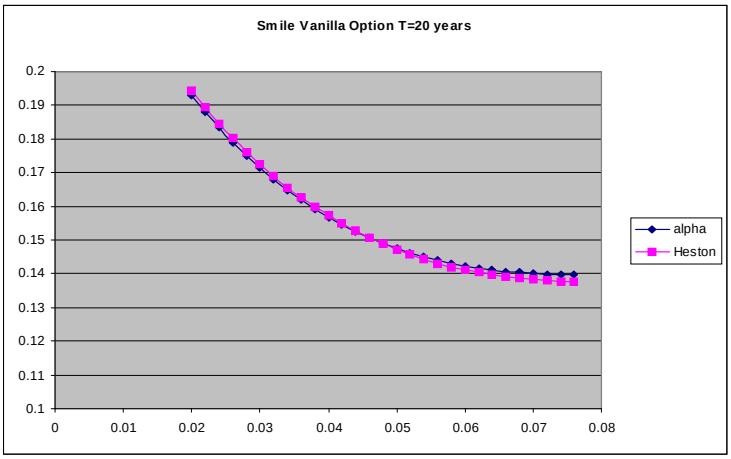

Figure 7:

This fact is a general situation that we observe for lognormal alpha models and normal alpha models (with normal volatility of course). This has two consequences.

1) The introduction of additional volatility factor do not increase the richness of the model. it do not make the smile more flexible.Specifically If we consider that one factor allow us to adjust the smile ATM trough its slope and convexity ATM, THe introduction of more factor do not allow us to adjust the smile far from the money while keeping its behaviour ATM.

2) The decomposition of the smile of the equivalent monodimensional Heston model into the whole spectra of volatility driver of the initial alpha model has to be understood as a "smiled PCA" . It is this point of view that is going to be exploited.

2.0.4 Calibration through var, varvar and covar observables

Calibrating a p-Asset Alpha model is equivalent to determine the set

$\{\alpha_{i,k}, C_{i,k}\}_{0\leq k\leq n, 1\leq i\leq p}$ such that

-The lognormal smiles of underlyings are as close as possible from a set of given option.?

-The normal smiles of spreadoptions are as close as possible from a set of given spreadoption.

Calibrating the Alpha model through pricing of the underlyings and spreadoptions is something we could have tried. But only with little hope of success. The reason is the large number of parameters to calibrate , the time necessary to perform all the necessary pricing and the difficulty for a gradient based algorithm to perform in a context where the pricing are done with a truncated summation due to the inverse fourier transform of the pricing algorithm.

The idea here is to bypass the Fourier transform in the calibrating process .

Thanks the result of the preceding paragraph, we assimilate a smile to a set of data:

-wether it is normal or lognormal

-An ATM point

-3 observables: $\{\frac{\langle dS,dS\rangle}{dt}, \frac{\langle d(\frac{\langle dS,dS\rangle}{dt}),d(\frac{\langle dS,dS\rangle}{dt})\rangle}{dt}, \frac{\langle d(\frac{\langle dS,dS\rangle}{dt}),dS\rangle}{dt}\}$ that we call from now on : $\{Var, CoVar, Varvar\}$

These observables have to be thought as being the roots of Heston or Normal heston pricing model parameters.

So the calibration is a two step process:

- first we prepare the global calibrationthrough a complete set of pre-calibrations.

The Heston parameters representing the underlyings

The Gaussian Heston parameters representing the spreadoption.

- second, we minimize the objective function that is built the following way:

We write $P = \{\alpha_{i,k}, C_{i,k}\}_{0 \leq k \leq n, 1 \leq i \leq p}$

$$\begin{aligned} Obtf_{Under}[P] &= \sum_{k \in Undlngs} \{C_{Under,Var,k}(Var_{U,P,k} - Var_{U,k})^2 + & (140)\\ & C_{Under,CoVar,k}(CoVar_{U,P,k} - CoVar_{U,k})^2 + & (141)\\ & C_{Under,Varvar,k}(Varvar_{U,P,k} - Varvar_{U,k})^2\} & (142) \end{aligned}$$

$$\begin{aligned} Obtf_{Spreadopt}[P] &= \sum_{k \in Spreadopts} \{C_{Spreadopt,Var,k}(Var_{S,P,k} - Var_{S,k})^2 & (143)\\ & C_{Spreadopt,CoVar,k}(CoVar_{S,P,k} - CoVar_{S,k})^2 + & (144)\\ & C_{Spreadopt,Varvar,k}(Varvar_{S,P,k} - Varvar_{S,k})^2\} & (145) \end{aligned}$$

So the calibration consists of minimizing the objective:

$$Obtf[P] = Obtf_{Under}[P] + Obtf_{Spreadopt}[P] \quad (146)$$

The constants $\{C_{Under,Var,k}, C_{Under,CoVar,k}, C_{Under,Varvar,k},$
$C_{Spreadopt,Var,k}, C_{Spreadopt,CoVar,k}, C_{Spreadopt,Varvar,k}\}$ are to be tuned to make the calibration effective.

A good starting point for tuning these constant is:

$$\begin{aligned} C_{Under,Var,k} &= \frac{1}{Var_{U,k}^2} & (147)\\ C_{Under,CoVar,k} &= Min\{\frac{1}{CoVar_{U,k}^2}, \frac{0.01}{Var_{U,k}Varvar_{U,k}}\} & (148)\\ C_{Under,Varvar,k} &= \frac{1}{Varvar_{U,k}^2} & (149)\\ C_{Spreadopt,Var,k} &= \frac{1}{Var_{S,k}^2} & (150)\\ C_{Spreadopt,CoVar,k} &= Min\{\frac{1}{CoVar_{S,k}^2}, \frac{0.01}{Var_{S,k}Varvar_{S,k}}\} & (151)\\ C_{Spreadopt,Varvar,k} &= \frac{1}{Varvar_{S,k}^2} & (152) \end{aligned}$$

Where the conventions are:

$Var_{U,k}$ means the Observable Variance computed on the calibrated parameters to the market prices of the the underlying k.

$Var_{S,k}$ means the Observable Variance computed on the calibrated parameters to the market prices of the the spreadoption k.

and the equivalent notations for the CoVar and Varvar observables.

These way to control the optimiser with a choice of the dimension on which we optimise allow us to experiment the "smiled PCA" perception of the Alpha calibration.

2.1 Computation of digital spreadoptions that pay float

Let assume that we want to compute the following option

$$opt_2 = E[S_2 1_{S_2 - S_1 - K \geq 0}] \tag{153}$$

by a change of numeraire its price is equal to:

$$opt_2 = E[S_2] E^{Q_{S_2}}[1_{S_2 - S_1 - K \geq 0}] \tag{154}$$

We know that the price of an ordinary digital is minus the derivative of the associated spreadoption

$$opt_2 = -E[S_2] \frac{\partial}{\partial K} E^{Q_{S_2}}[(S_2 - S_1 - K)^+] \tag{155}$$

We assume that the pricing of the spreadoption $(S_2 - S_1 - K)^+$in the measure Q_{S_2} is given again by a gaussian heston pricing associated with parameters $\{V_{0,D_1}, \rho_{D_1}, \nu_{D_1}\}$. Therefore the value of the digital that pay float is given by :

$$opt_2 = -E[S_2] \frac{\partial}{\partial K} Heston[V_{0,D_1}, \rho_{D_1}, \nu_{D_1}, K] \tag{156}$$

Where the Heston price is computed taking into account that the forward should be:

$$F_{21} = S_2 e^{\sigma_2^2 T} - S_1 e^{\sigma_1 \sigma_2 \rho T} \tag{157}$$

Now there is another digital that pay float which is:

$$opt_1 = E[S_1 1_{S_2 - S_1 - K \geq 0}] \tag{158}$$

and we can reexpress it using the duality as :

$$opt_1 = S_1 - E[S_1 1_{S_2 - S_1 - K \leq 0}] \tag{159}$$

that we can reexpress again as

$$opt_1 = S_1 - E[S_1 1_{S_1 - S_2 - (-K) \geq 0}] \tag{160}$$

now we see that we have to compute a digital which pays float (the first float of the spread) but with the role of $S_1$and S_2 exchanged and for a strike which is (-K).

If we can compute a set of parameters $\{V_{0,D_1}, \rho_{D_1}, \nu_{D_1}\}$. for the spread $S_2 - S_1$ in the measure Q_{S_2}, then the parameters for the spread $S_1 - S_2$ in the measure $Q_{S_2}$are the same except for a $\rho_{D_1} - > -\rho_{D_1}$.and the forward should be

$$F_{12} = S_1 e^{\sigma_1^2 T} - S_2 e^{\sigma_1 \sigma_2 \rho T} \tag{161}$$

If we assume that the parameters $\{V_{0,D_2}, \rho_{D_2}, \nu_{D_2}\}$.which are the parameters of the spread $S_1 - S_2$ in the measure Q_{S_1} are nevertheless very close, then

$$\{V_{0,D_2}, \rho_{D_2}, \nu_{D_2}\} = \{V_{0,D_1}, -\rho_{D_1}, \nu_{D_1}\} \tag{162}$$

Now there is an arbitrage relatioship that needs to hold:

$$opt_1 - opt_2 - Kopt_D = spreadoption \tag{163}$$

If we write it as functional on the K variable we get a function equality:

$$-E[S_2]\frac{\partial}{\partial K}Heston_{F_{21}}[V_{0,D_1}, \rho_{D_1}, \nu_{D_1}, K] + E[S_1](1 - \tag{164}$$

$$\frac{\partial}{\partial K}Heston_{F_{12}}[V_{0,D_1}, -\rho_{D_1}, \nu_{D_1}, K]) \tag{165}$$

$$= \quad spreadoption(K) + K\frac{\partial}{\partial K}spreadoption(K) \tag{166}$$

That in our framework we translate in an optimisation problem,
whose solution should give us the triplet$\{V_{0,D_1}, \rho_{D_1}, \nu_{D_1}\}$
we mean : the minimisation of

$$obj \quad = \quad \sum_{i=1}^{n} w_i\{spreadoption(K_i) + K\frac{\partial}{\partial K}spreadoption(K_i) \tag{167}$$

$$+E[S_2]\frac{\partial}{\partial K}Heston_{F_{21}}[V_{0,D_1}, \rho_{D_1}, \nu_{D_1}, K] - \tag{168}$$

$$E[S_1](1 - \frac{\partial}{\partial K}Heston_{F_{12}}[V_{0,D_1}, -\rho_{D_1}, \nu_{D_1}, K])\}^2 \tag{169}$$

where w_i as usual will be equal to the normal vega of the strike K_i to perform an optimisation on the smile

Sometime the optimisation do not perform a good job, a possible improvement is to look for two sets of parameters $\{V_{0,D_2}, \rho_{D_2}, \nu_{D_2}\}, \{V_{0,D_1}, \rho_{D_1}, \nu_{D_1}\}$ and enough strikes or additional conditions likes a digital market prices or stipulating that those two triplets being not too different from each other or from the optimum of the preceding simplied optimisation.

Exemples of option prices
$S_1 = 0.0507, S_2 = 0.0505, T = 10$
$\alpha_{11} = 0.07, \alpha_{12} = 0.06, \alpha_{13} = 0.03, \alpha_{14} = 0.02, \alpha_{15} = \sqrt{0.0005}$
$\alpha_{21} = 0.07, \alpha_{21} = 0.06, \alpha_{31} = 0.03, \alpha_{41} = 0.02, \alpha_{15} = -\sqrt{0.0005}$
$\nu_i = 2, \rho_i = -0.5, \nu_5 = 0$
the goal here is to get correlated assets with correlation=0.951456
thanks to the optimisation, the arbitrage is small
the two digitals that pays float:
the digital that pays 1

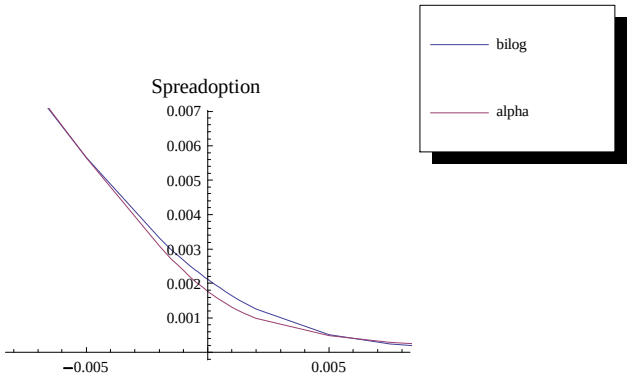

Figure 8:

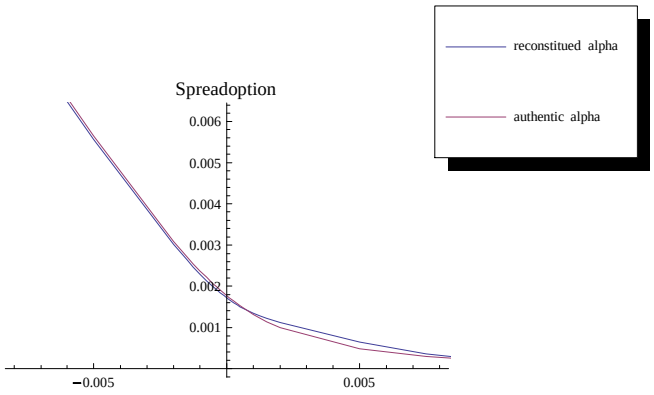

Figure 9:

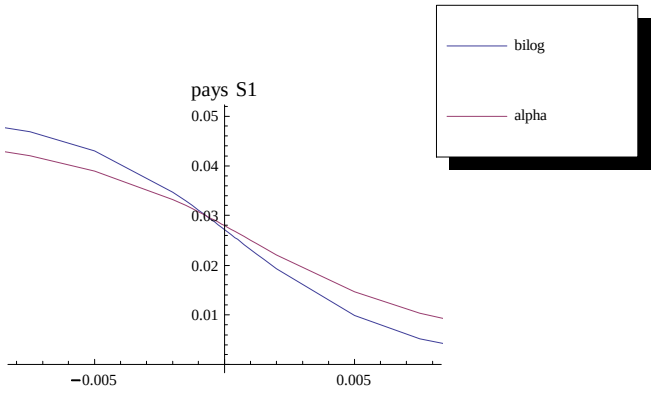

Figure 10:

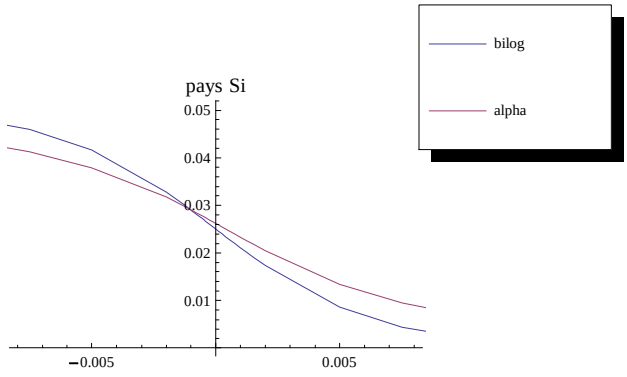

Figure 11:

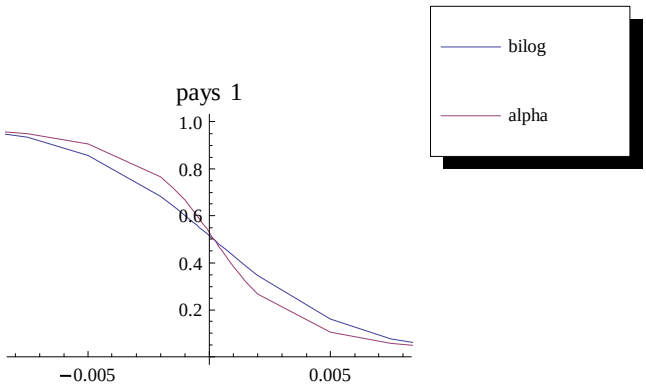

Figure 12:

3 Annexes

List of add-ins that implements the precedings functions,

The add-ins need a certain number of arguments. Some of them are obvious, others are technical, but they are needed to make the numerical algorithms working.

The add-ins compute the requested quantities plus additional quantities that can be used to make cross verification or further computations

3.0.1 ARM_CF_ALPHA_4Under_VanillaOption

Computes the vanilla options for $1 \leq i \leq 4$

$$\frac{dS_i}{S_i} = \sum_{k=0}^{4} \alpha_{i,k} d\hat{W}_k \sqrt{V_k} + d\hat{W}_{4+i}\sqrt{V_{4+i}} \tag{170}$$

The idea is to have up to 4 drivers and a residual that carries a smile (of course) . The residual is specific to the asset, but the 4 drivers are common to all assets.

The add-in computes

$$E[(S_i - K)^+] \tag{171}$$

3.0.2 ARM_CF_ALPHA_4Under_VanillaSpreadoption12

The main pricing function of this frame work: prices the spreadoptions associated with

The add-in computes

$$E[(S_2 - S_1 - K)^+] \tag{172}$$

the other add-in works the same way, they just transform the input to use the preceding add-in:

-ARM_CF_ALPHA_4Under_VanillaSpreadoption13
-ARM_CF_ALPHA_4Under_VanillaSpreadoption14
-ARM_CF_ALPHA_4Under_VanillaSpreadoption23
-ARM_CF_ALPHA_4Under_VanillaSpreadoption24
-ARM_CF_ALPHA_4Under_VanillaSpreadoption34

3.0.3 ARM_CF_ALPHA_4Under_DigitalSpreadoption12

Pricing function that compute all digitals associated with a spreadoption, it computes also the spreadoption .

The reason for that is that, the pricing compute a price which is arbitrage free, with respect to the all set of digitals

and the arbitrage:

$$E[(S_2 - S_1 - K)^+] = E[S_2 1_{S_2 - S_1 - K > 0}] - E[S_1 1_{S_2 - S_1 - K > 0}] - KE[1_{S_2 - S_1 - K > 0}] \tag{173}$$

3.0.4 ARM_CF_ALPHA_4Unders_40p_Smiles_Calibration

Calibrating function that computes the optimal $\alpha_{i,k}$ and (ρ_i, ν_i) that reproduce a given set of smiles for the underlyings : $E[(S_i - K_k)^+]$ and prices for the spreadoptions: $E[(S_j - S_i - K_k)^+]$

Arguments : the list of arguments is:

Forwards, a list of 4 forwards

Alphas, a list of 4X4 alphas . Each value is either an initial value for the optimisation, either a value that is not optimised.

CIRs, a list of 8X3 CIR (ρ, ν, V_0) parameters . Each value is either an initial value for the optimisation, either a value that is not optimised.

the V_i for i$\leq$4 should be put to 1 to avoid an overdertermination with the alphas that would make the optimisation non robust.

T, a double . Maturity of the optimisation

params, a list of 35 double.

the 6 first parameters concern the behaviour of the optimiser and other needed constants:

-LT : Ratio of the long term variance with the initial value of the stochastic variance for all CIR processes: $LT = \frac{V_\infty}{V_0}$

-lambda: mean reversion speed of the CIR processes, assumed to be the same for all processes. This is why the Alpha model is not used to model several maturities yet. There is no different time scale among the CIR processes.

-Algo

Specifies the optimisation alogithm which is used. Now there is only one:

= 0->the Nag routine:nag_opt_nlp_solve (e04wdc).

-Mode

Specifies different ways for the spreadoption pricing to work:

=0 -> uses the exact values of the observable to be transmitted to the Normal heston

Pricing function

=1 -> uses the Development around the Bilog as described in this document

linesearch

Specifies a number detween 0 and 1 that describes how the over shouting is performed in the

QP optimisation routine : for exemple 0.2.

Tol

Specifies the tolerance to be transmitted to the optimisation routine :

for exemple: 1 E-08

VU1
VU2
VU3
VU4

Specify the coefficients $C_{Under,Var,k}$

VarVU1
VarVU2
VarVU3
VarVU4

Specify the coefficients $C_{Under,Varvar,k}$

CovVU1
CovVU2
CovVU3
CovVU4

Specify the coefficients $C_{Under,CoVar,k}$

Spread12
Spread13
Spread14
Spread23
Spread24

Spread34
Specify the coefficients $C_{Spreadopt,Var,k}$
CovSpread12
CovSpread13
CovSpread14
CovSpread23
CovSpread24
CovSpread34
Specify the coefficients $C_{Spreadopt,CoVar,k}$
VarSpread12
VarSpread13
VarSpread14
VarSpread23
VarSpread24
VarSpread34
Specify the coefficients $C_{Spreadopt,Varvar,k}$

nbiter Specifiy the maximum number of global iteration of the QP optimisation routine .
Transmitted to Nag

opt1_vec,

opt2_vec,

opt3_vec,

opt4_vec, The Heston parameters $\{\rho, \nu, V_0\}$ associated with the Underlyings

spre12_vec,

spre13_vec,

spre14_vec,

spre23_vec,

spre24_vec,

spre34_vec, The Heston parameters $\{\rho, \nu, V_0\}$ associated with the Spread-options

Coefficients	
Nb Iteration	3000
LT	1
lambda	0.1
Algo	1
mode	-6
linesearch	0.2
Tol	1.00E-08
VU1	4.23E+03
VU2	4.23E+03
VU3	4.23E+03
VU4	4.23E+03
VarVU1	6.50E+06
VarVU2	6.50E+06
VarVU3	6.50E+06
VarVU4	6.50E+06
CovVU1	8.55E+05
CovVU2	8.55E+05
CovVU3	8.55E+05
CovVU4	8.55E+05
Spread12	1.00E+00
Spread13	0.00E+00
Spread14	1.00E+00
Spread23	0.00E+00
Spread24	1.00E+00
Spread34	0.00E+00
CovSpread12	0.00E+00
CovSpread13	1.00E+00
CovSpread14	0.00E+00
CovSpread23	0.00E+00
CovSpread24	0.00E+00
CovSpread34	0.00E+00
VarSpread12	0.00E+00
VarSpread13	1.00E+00
VarSpread14	0.00E+00
VarSpread23	0.00E+00
VarSpread24	0.00E+00
VarSpread34	0.00E+00

		yes/no	min	max	Initial
CIR1	rho1	1	-0.999	0.999	0
	kappa1	1	0.0001	5	1
	V1	0	0	2	1
CIR2	rho2	1	-0.999	0.999	0
	kappa2	1	0.0001	10	1
	V2	0	0	2	1
CIR3	rho3	1	-0.999	0.999	0
	kappa3	1	0.0001	10	1
	V3	0	0	2	1
CIR4	rho4	1	-0.999	0.999	0
	kappa4	1	0.0001	5	1
	V4	0	0	2	1
CIR5 (ass.1)	rho5	1	-0.999	0.999	0
	kappa5	1	0.0001	5	0.001
	V5	1	0	1	0.0001
CIR6(ass.2)	rho6	1	-0.999	0.999	0
	kappa6	1	0.0001	5	0.001
	V6	1	0	1	0.0001
CIR7(ass.3)	rho7	1	-0.999	0.999	0
	kappa7	1	0.0001	5	0.001
	V7	1	0	1	0.0001
CIR8(ass.4)	rho8	1	-0.999	0.999	0
	kappa8	1	0.0001	5	0.001
	V8	1	0	1	0.0001
U1	alpha11	1	0	1	0.001
	alpha12	1	-1	1	0
	alpha13	1	-1	1	0
	alpha14	1	-1	1	0
U2	alpha21	1	0	1	0.001
	alpha22	1	-1	1	0
	alpha23	1	-1	1	0
	alpha24	1	-1	1	0
U3	alpha31	1	0	1	0.001
	alpha32	1	-1	1	0
	alpha33	1	-1	1	0
	alpha34	1	-1	1	0
U4	alpha41	1	0	1	0.001
	alpha42	1	-1	1	0
	alpha43	1	-1	1	0
	alpha44	1	-1	1	0

Figure 13:

optim_vec, a vector of 40 integer $\in \{0, 1\}$ specifing if the associated dimension is to be optimised or not. the meaning of each number is given by the corresponding ligne in the figure 13.

Upper_Lim_vec, a vector of 40 double representing the upper limit of the associated dimension .the meaning of each number is given by the corresponding ligne in the figure 13.

Lower_Lim_vec a vector of 40 double representing the lower limit of the associated dimension. the meaning of each number is given by the corresponding ligne in the figure 13.

Exemple of input presented in tables :

Results: The result of optimisation are given first in the following order, followed by the value of the objective function at the optimum (the fit)

rho1	-0.412509979
kappa1	1.183580304
V1	1
rho2	-0.273142196
kappa2	1.964122138
V2	1
rho3	0
kappa3	1
V3	1
rho4	0
kappa4	1
V4	1
rho5	-0.279451445
kappa5	0.150507575
V5	0.005421222
rho6	0
kappa6	0.001
V6	0.0001
rho7	-0.940411184
kappa7	0.005154997
V7	0
rho8	0.999
kappa8	0.002294721
V8	0.003504209
alpha11	1.72E-01
alpha12	6.26E-03
alpha13	0.00E+00
alpha14	0.00E+00
alpha21	1.00E-03
alpha22	1.00E-03
alpha23	0.00E+00
alpha24	0.00E+00
alpha31	1.46E-01
alpha32	-4.77E-02
alpha33	0.00E+00
alpha34	0.00E+00
alpha41	1.49E-01
alpha42	-4.02E-02
alpha43	0.00E+00
alpha44	0.00E+00
Fit	4.2391E-04

Figure 14:

Components	U1	0.035007036
	U2	0.000102
	U3	0.023630254
	U4	0.027336516
	Spread12	7.78049E-05
	Spread13	2.01191E-05
	Spread14	2.60255E-05
	Spread23	6.29522E-05
	Spread24	6.15972E-05
	Spread34	8.45926E-06
	CovU1	-0.002707813
	CovU2	-1.02472E-09
	CovU3	-0.001464847
	CovU4	-0.001574338
	VarU1	0.001345765
	VarU2	1.05259E-10
	VarU3	0.00065865
	VarU4	0.000701679
	Sp12Cov	2.82181E-07
	Sp13Cov	3.5599E-08
	Sp14Cov	3.13424E-08
	Sp23Cov	-1.98073E-07
	Sp24Cov	-1.6559E-07
	Sp34Cov	8.20598E-10
	Sp12Var	4.73068E-09
	Sp13Var	8.03362E-10
	Sp14Var	7.08491E-10
	Sp23Var	3.63474E-09
	Sp24Var	2.63717E-09
	Sp34Var	5.18875E-13

Figure 15:

Then we have the value of each component of the objective function at the optimum. Because it is infact a multi-objective optimisation, it can be used to optimise the coefficients given in input. Typically the coefficients should be the inverse of the square of those. Unfortunatly we do not know them in advance. Optimally, we can reinject those as input of a second stage of optimisation and hope for a convergence after N steps. Practically, they do not change much, so for a class of problem, we determine them and resuse.

Finally we have the value of the derivative of the objective at optimum. These coefficients can be used to appreciate the sharpness of the optimum and also to know how critical are the boundaries, and if iti is important to relax those boundaries if the optimum which reached is not satisfactory.

Derivatives at optimum	rho1	-0.412509979
	kappa1	1.183580304
	V1	1
	rho2	-0.273142196
	kappa2	1.964122138
	V2	1
	rho3	0
	kappa3	1
	V3	1
	rho4	0
	kappa4	1
	V4	1
	rho5	-0.279451445
	kappa5	0.150507575
	V5	0.005421222
	rho6	0
	kappa6	0.001
	V6	0.0001
	rho7	-0.940411184
	kappa7	0.005154997
	V7	0
	rho8	0.999
	kappa8	0.002294721
	V8	0.003504209
	alpha11	0.171891228
	alpha12	0.006262532
	alpha13	0
	alpha14	0
	alpha21	0.001
	alpha22	0.001
	alpha23	0
	alpha24	0
	alpha31	0.146120177
	alpha32	-0.047740419
	alpha33	0
	alpha34	0
	alpha41	0.149062824
	alpha42	-0.040156967
	alpha43	0
	alpha44	0
	NbIteration	412
	NbFuncCall	-858993460

Figure 16: