THE SMILE CALIBRATION PROBLEM SOLVED

JULIEN GUYON AND PIERRE HENRY-LABORDÈRE

Abstract. Following previous work on calibration of multi-factor local stochastic volatility models to market smiles [7], we show how to calibrate *exactly* any such models. Our approach, based on McKean's particle method, extends to hybrid models, for which we provide a Malliavin representation of the effective local volatility. We illustrate the efficiency of our algorithm on hybrid local stochastic volatility models.

Introduction

The calibration of stochastic volatility/hybrid models to market smiles is a long-standing issue in quantitative finance. Partial answers have been given: for low-dimensional factor models such as old-fashioned one-factor local stochastic volatility models (in short LSV models) or hybrid Dupire local volatility model with a one factor interest rate model, this calibration can be achieved by solving a two-dimensional non-linear Fokker-Planck PDE [10, 11]. For multi-factor stochastic models such as variance swap curve models [4], Libor Market Models with stochastic volatility such as SABR-LMM model or hybrid Dupire local volatility appearing in the pricing of PRD derivatives [12], approximate solutions have been suggested based on heat kernel perturbation expansions [6], time-averaging methods and the so-called Markovian projection techniques [12, 7].

In this paper, we introduce the particle algorithm. This Monte-Carlo-type method allows to calibrate *exactly* any LSV/hybrid model to market smiles. This method relies on the fact that the dynamics of the calibrated model is a *non-linear* McKean stochastic differential equation (SDE). Here non-linear means that the volatility depends on the marginal distribution of the process. As a consequence, this SDE is associated to a non-linear Fokker-Planck PDE. The particle method consists in considering this equation as the large N limit of a N-dimensional linear Fokker-Planck PDE that can be simulated efficiently with a Monte-Carlo algorithm.

The paper is organized as follows. We first exhibit examples of non-linear McKean SDEs arising in the calibration of LSV/hybrid models to market smiles. Then, we present the particle algorithm, including important implementation details. Eventually, we illustrate the efficiency of our algorithm on various models commonly used by practitioners.

Local stochastic volatility model

A local stochastic volatility (LSV) model is defined by the following SDE for the T-forward f_t in the forward measure $\mathbb{P}^T$

$$df_t = f_t \sigma(t, f_t) a_t \, dW_t \tag{1}$$

where a_t is a (possibly multi-factor) stochastic process. It can be seen as an extension to the Dupire local volatility (LV) model, or as an extension to the stochastic volatility (SV) model. In the SV model, one handles only a finite number of parameters (volatility-of-volatility, spot/volatility correlations, etc.). As a consequence, one is not able to perfectly calibrate the implied volatility surface. In order to be able to calibrate exactly market smiles, one decorates the volatility of the forward with a local volatility function $\sigma(t,f)$. By definition, the effective local volatility is [5]

$$\sigma_{\mathrm{loc}}(t,f)^2 = \sigma(t,f)^2 \mathbb{E}^{\mathbb{P}^T}[a_t^2 | f_t = f]$$

From [5], this model is exactly calibrated to market smiles if and only if this effective volatility equals the square of the Dupire local volatility $\sigma_{\mathrm{Dup}}(t,f)^2$. SDE (1), once the requirement that market marginals have to be calibrated exactly has been taken into account, can be rewritten as

$$df_t = f_t \frac{\sigma_{\mathrm{Dup}}(t,f_t)}{\sqrt{\mathbb{E}^{\mathbb{P}^T}[a_t^2|f_t]}} a_t dW_t \tag{2}$$

The local volatility function depends on the joint pdf $p(t,f,a)$ of (f_t, a_t):

$$\sigma(t,f,p) = \sigma_{\mathrm{Dup}}(t,f) \sqrt{\frac{\int p(t,f,a')da'}{\int a'^2 p(t,f,a')da'}} \tag{3}$$

This is an example of McKean SDEs.

McKean SDEs. A McKean equation for an n-dimensional process X is a SDE in which the drift and volatility depend not only on the current value X_t of the process, but also on the probability distribution $\mathbb{P}_t$ of X_t:

$$dX_t = b(t, X_t, \mathbb{P}_t)\ dt + \sigma(t, X_t, \mathbb{P}_t) \cdot dW_t, \qquad \mathbb{P}_t = \mathrm{Law}\,(X_t) \tag{4}$$

where W_t is a d-dimensional Brownian motion. In [13], uniqueness and existence are proved for equation (4) if the drift and volatility coefficients are Lipschitz-continuous functions of $\mathbb{P}_t$, with respect to the so-called Wasserstein metric. The probability density function $p(t,\cdot)$ of X_t is solution to the Fokker-Planck PDE:

$$\begin{aligned} &-\partial_t p - \sum_{i=1}^n \partial_i \left(b^i(t,x,\mathbb{P}_t) p(t,x) \right) + \frac{1}{2} \sum_{i,j=1}^n \partial_{ij} \left(\sum_{k=1}^d \sigma_k^i(t,x,\mathbb{P}_t) \sigma_k^j(t,x,\mathbb{P}_t) p(t,x) \right) = 0 \\ &\lim_{t \to 0} p(t,x) = \delta(x - X_0) \end{aligned} \tag{5}$$

It is non-linear because $b^i(t,x,\mathbb{P}_t)$ and $\sigma_k^i(t,x,\mathbb{P}_t)$ depend on the unknown p.

Existence of the LSV model. In (3), the Lipschitz condition is not satisfied. Hence a uniqueness and existence result for Equation (2) is not at all obvious. In particular, given a set of stochastic volatility parameters, it is not clear at all whether a LSV model exists for a given arbitrary arbitrage-free implied volatility surface: some smiles may not be attainable by the model. However, a partial result exists: in [1] it is shown that the calibration problem for a LSV model is well posed but only (a) until some maturity T^*, (b) if the volatility-of-volatility is small enough, and (c) in the case of suitably regularized initial conditions - hence the result does not apply to Equation (5) because of the initial Dirac mass. Our numerical experiments show that the calibration does not work for large enough volatility-of-volatility, whatever the algorithm used: PDE, particle method or Markovian projection method [7]. This may come from numerical issues, or from non-existence of a solution. The problem of deriving the set of stochastic volatility parameters for which the LSV model does exist for a given market smile, is very challenging and open (see an illustration in the numerical experiments section).

Hybrid Models

A hybrid LSV model is defined in a risk-neutral measure $\mathbb{P}$ by

$$\frac{dS_t}{S_t} = r_t dt + \sigma(t, S_t) a_t dW_t$$

where the short-term rate r_t and the stochastic volatility a_t are Itô processes. For the sake of simplicity, we assume at this stage that there are no dividends. We explain how to include (discrete) dividends in our calibration procedure in the appendix. This model is exactly calibrated to the market smile if and only if (see Proof 1)

$$\sigma(T, K)^2 \mathbb{E}^{\mathbb{P}^T}[a_T^2 | S_T = K] = \sigma_{\mathrm{Dup}}(T, K)^2 - P_{0T} \frac{\mathbb{E}^{\mathbb{P}^T}[\left(r_T - r_T^0\right) 1_{S_T > K}]}{\frac{1}{2} K \partial_K^2 \mathcal{C}(T, K)} \tag{6}$$

with $r_T^0 = \mathbb{E}^{\mathbb{P}^T}[r_T] = -\partial_T \ln P_{0T}$ and

$$\sigma_{\mathrm{Dup}}(T, K)^2 = \frac{\partial_T \mathcal{C}(T, K) + r_T^0 K \partial_K \mathcal{C}(T, K)}{\frac{1}{2} K^2 \partial_K^2 \mathcal{C}(T, K)}$$

Here $\mathbb{P}^T$ denotes the T-forward measure, $\mathcal{C}(T, K)$ the market fair value of a vanilla option with strike K and maturity T, and P_{tT} the time-t value of the bond of maturity T. Hence the dynamics of the calibrated hybrid LSV model reads as the following non-linear McKean diffusion for the forward $f_t = S_t / P_{t\bar{T}}$ in the forward measure $\mathbb{P}^{\bar{T}}$, where $\bar{T}$ denotes the last maturity date for which we want to calibrate the market smile:

$$\frac{df_t}{f_t} = \sigma\left(t, P_{t\bar{T}} f_t, \mathbb{P}_t^{\bar{T}}\right) a_t \, dW_t^{\bar{T}} - \sigma_{\mathrm{P}}^{\bar{T}}(t).dB_t^{\bar{T}}$$

where[1]

$$\sigma\left(t, K, \mathbb{P}_t^{\bar{T}}\right)^2 = \left(\sigma_{\mathrm{Dup}}(t, K)^2 - P_{0\bar{T}} \frac{\mathbb{E}^{\mathbb{P}^{\bar{T}}}[P_{t\bar{T}}^{-1}\left(r_t - r_t^0\right) 1_{S_t > K}]}{\frac{1}{2} K \partial_K^2 \mathcal{C}(t, K)}\right) \frac{\mathbb{E}^{\mathbb{P}^{\bar{T}}}[P_{t\bar{T}}^{-1} | S_t = K]}{\mathbb{E}^{\mathbb{P}^{\bar{T}}}[P_{t\bar{T}}^{-1} a_t^2 | S_t = K]} \tag{7}$$

$\sigma_{\mathrm{P}}^{\bar{T}}(t)$ is the volatility of the bond $P_{t\bar{T}}$, and $B_t^{\bar{T}}$ the (possibly multidimensional) $\mathbb{P}^{\bar{T}}$-Brownian motion that drives the interest rate curve.

Malliavin representation of the hybrid local volatility. As we explain in our section on numerical experiments, estimating an unconditional expectation such as $\mathbb{E}^{\mathbb{P}^T}[\left(r_T - r_T^0\right) 1_{S_T > K}]$ is more costly than estimating a conditional expectation such as $\mathbb{E}^{\mathbb{P}^T}[a_T^2 | S_T = K]$. It is thus highly desirable to get a representation of the hybrid contribution to local volatility, i.e., the last term in (6), as a conditional expectation given $S_T = K$. We now explain how this is achieved.

From the martingale representation theorem,

$$r_T - r_T^0 = \int_0^T \sigma_{\mathrm{r}}^T(s).dB_s^T$$

with $\sigma_{\mathrm{r}}^T(s)$ an adapted process. Note that, from Clark-Ocone's formula, $\sigma_{\mathrm{r}}^T(s) = \mathbb{E}_s^{\mathbb{P}^T}[D_s^{B^T} r_T]$ with $D_s^{B^T}$ the Malliavin derivative with respect to the Brownian motion B^T, and $\mathbb{E}_s$ the conditional expectation given $\mathcal{F}_s$, the natural filtration of all the Brownian motions used. The application of Clark-Ocone's formula to the process $1_{S_T > K}$, combined with Itô's isometry, gives

$$P_{0T} \mathbb{E}^{\mathbb{P}^T}[\left(r_T - r_T^0\right) 1_{S_T > K}] = \partial_K^2 \mathcal{C}(T, K) \int_0^T \mathbb{E}^{\mathbb{P}^T}[\sigma_{\mathrm{r}}^T(s).D_s^{B^T} S_T | S_T = K] \, ds$$

[1] We have the identity: $P_{0T} \mathbb{E}^{\mathbb{P}^T}[X_T] = P_{0\bar{T}} \mathbb{E}^{\mathbb{P}^{\bar{T}}}[P_{T\bar{T}}^{-1} X_T]$.

so that the hybrid contribution to local volatility reads (see [2] for a similar expression)

$$P_{0T}\frac{\mathbb{E}^{\mathbb{P}^T}\left[\left(r_T - r_T^0\right) 1_{S_T>K}\right]}{\frac{1}{2}K\partial_K^2\mathcal{C}(T,K)} = \frac{2}{K}\int_0^T \mathbb{E}^{\mathbb{P}^T}\left[\sigma_{\mathrm{r}}^T(s).D_s^{B^T} S_T | S_T = K\right] ds \tag{8}$$

For the sake of simplicity, let us assume that the short rate r_t follows a one-factor Itô diffusion

$$dr_t = \mu_{\mathrm{r}}(t,r_t)dt + \sigma_{\mathrm{r}}(t,r_t)dB_t$$

where $\mu_{\mathrm{r}}(t,r_t)$ and $\sigma_{\mathrm{r}}^T(t) = \sigma_{\mathrm{r}}(t,r_t)$ are deterministic functions of the time t and the short rate r_t, and B_t is a one-dimensional $\mathbb{P}$-Brownian motion with $d\langle B,W\rangle_t = \rho\, dt$. Then $\sigma_{\mathrm{P}}^T(t)$, the volatility of the bond P_{tT}, is also a deterministic function $\sigma_{\mathrm{P}}^T(t,r_t)$ of the time t and the short rate r_t. Moreover, we assume that the stochastic volatility is not correlated with the stochastic rate r_t. Both assumptions can be easily relaxed but at the cost of additional straightforward computations. By explicitly computing $D_s^{B^T} S_T$, (8) can then be written as (see Proof 2)

$$P_{0T}\frac{\mathbb{E}^{\mathbb{P}^T}\left[\left(r_T - r_T^0\right) 1_{S_T>K}\right]}{\frac{1}{2}K\partial_K^2\mathcal{C}(T,K)} = 2\mathbb{E}^{\mathbb{P}^T}\left[V_T\left(\rho U_T + \Theta_T^T\Xi_T^T - \Lambda_T^T\right)|S_T = K\right] \tag{9}$$

with

$$\frac{dV_t}{V_t} = S_t\partial_S\sigma(t,S_t)a_t\left(dW_t - a_t\sigma(t,S_t)dt\right), \quad V_0 = 1 \tag{10}$$

$$dU_t = \sigma_{\mathrm{r}}(t,r_t)a_t\frac{\sigma(t,S_t)}{V_t}dt, \quad U_0 = 0 \tag{11}$$

$$\frac{dR_t^T}{R_t^T} = \left(\partial_r\mu_{\mathrm{r}}(t,r_t) + \sigma_{\mathrm{r}}(t,r_t)\partial_r\sigma_{\mathrm{P}}^T(t,r_t)\right)dt + \partial_r\sigma_{\mathrm{r}}(t,r_t)dB_t, \quad R_0 = 1 \tag{12}$$

$$d\Theta_t^T = \frac{R_t^T}{V_t}\left(1 + \rho\partial_r\sigma_{\mathrm{P}}^T(t,r_t)\sigma(t,S_t)a_t\right)dt, \quad \Theta_0 = 0 \tag{13}$$

$$d\Xi_t^T = \frac{\sigma_{\mathrm{r}}(t,r_t)^2}{R_t^T}dt, \quad \Xi_0 = 0, \qquad d\Lambda_t^T = \Theta_t^T d\Xi_t^T, \quad \Lambda_0 = 0 \tag{14}$$

In the commonly used short rate models, such as the Ho-Lee and Hull-White models, the volatility of the bond is deterministic so that $\partial_r\sigma_{\mathrm{P}}^T(t,r_t) = 0$. Then the processes R_t, Θ_t, Ξ_t and Λ_t are independent of T, R_t is the tangent process of r_t, and the computation of (9), for all T, requires the simulation of only 4 processes f_t, r_t, V_t, R_t and 4 integrals U_t, Θ_t, Ξ_t, Λ_t.

In this case, we eventually obtain the following representation of the local volatility (7):

$$\sigma\left(t,K,\mathbb{P}_t^{\bar{T}}\right)^2 = \left(\sigma_{\mathrm{Dup}}(t,K)^2 - 2\frac{\mathbb{E}^{\mathbb{P}^{\bar{T}}}\left[P_{t\bar{T}}^{-1}V_t\left(\rho U_t + \Theta_t\Xi_t - \Lambda_t\right)|S_t = K\right]}{\mathbb{E}^{\mathbb{P}^{\bar{T}}}\left[P_{t\bar{T}}^{-1}|S_t = K\right]}\right)\frac{\mathbb{E}^{\mathbb{P}^{\bar{T}}}\left[P_{t\bar{T}}^{-1}|S_t = K\right]}{\mathbb{E}^{\mathbb{P}^{\bar{T}}}\left[P_{t\bar{T}}^{-1}a_t^2|S_t = K\right]} \tag{15}$$

where the dynamics for V_t and R_t is[2]

$$\frac{dV_t}{V_t} = S_t\partial_S\sigma(t,S_t)a_t\left(dW_t^{\bar{T}} + \left(\rho\sigma_{\mathrm{P}}^{\bar{T}}(t,r_t) - a_t\sigma(t,S_t)\right)dt\right), \quad V_0 = 1$$
$$\frac{dR_t}{R_t} = \left(\partial_r\mu_{\mathrm{r}}(t,r_t) + \partial_r\sigma_{\mathrm{r}}(t,r_t)\sigma_{\mathrm{P}}^{\bar{T}}(t,r_t)\right)dt + \partial_r\sigma_{\mathrm{r}}(t,r_t)dB_t^{\bar{T}}, \quad R_0 = 1$$

For general short-rate models, $\partial_r\sigma_{\mathrm{P}}^T(t,r_t)$ has no reason to vanish but, provided it is small, we can neglect it and use (15). We now look at the case of hybrid Libor Market Models.

[2] U_t, Θ_t, Ξ_t and Λ_t are drift terms and are not affected by the change of measure from $\mathbb{P}$ to $\mathbb{P}^{\bar{T}}$.

Hybrid Libor Market Models. The risk-neutral measure $\mathbb{P}$ does not exist in Libor Market Models (LMM) as it is not possible to invest in an ultra-short Libor. The above discussion must then be refined. In the measure $\mathbb{P}^{T_{\bar{N}}}$ associated to the bond $P_{tT_{\bar{N}}}$ of maturity $T_{\bar{N}} = \bar{T}$, the forward $f_t = S_t/P_{t\bar{T}}$ is a local martingale

$$\frac{df_t}{f_t} = \sigma(t, S_t, \mathbb{P}_t^{\bar{T}}) a_t dW_t^{\bar{T}} - \sigma_{\mathrm{P}}^{\bar{T}}(t).dB_t^{\bar{T}}$$

with $\sigma_{\mathrm{P}}^{\bar{T}}(t)$ the volatility of the bond $P_{t\bar{T}}$. Note that because of the bond $P_{tT_{\beta(t)-1}}$, with $T_{\beta(t)-1}$ the nearest future Libor fixing date, the dynamics of the bond

$$P_{t\bar{T}} = P_{tT_{\beta(t)-1}} \prod_{i=\beta(t)}^{\bar{N}} \frac{1}{1+\tau_i L_i(t)}$$

(hence its volatility) is not defined in a LMM. Only the volatility of a ratio of bonds maturing at the Libor fixing dates, arising for example in a change of numéraire, is well-defined. In order to get the evolution of $P_{tT_{\beta(t)-1}}$, which cannot be deduced from our discrete forward rates, we need to model the instantaneous forward rate volatility. This is an additional arbitrary input to the discrete-tenor setting. As a convenient choice, we assume that the volatility of the short term instantaneous forward rate vanishes (zero ultra-short volatility). More specifically, we use a linear interpolation for the value of $P_{tT_{\beta(t)-1}}$

$$P_{tT_{\beta(t)-1}} = \frac{1}{1 + (T_{\beta(t)-1} - t) L_{\beta(t)-1}(T_{\beta(t)-2})}$$

Another possible choice, in which the volatility of the instantaneous forward rate does not vanish, is

$$P_{tT_{\beta(t)-1}} = \frac{1}{1 + (T_{\beta(t)-1} - t)\left(\theta L_{\beta(t)-1}(T_{\beta(t)-2}) + (1-\theta) L_{\beta(t)}(t)\right)}$$

Then, we get that the volatility $\sigma_{\mathrm{P}}^{\bar{T}}(t)$ (assuming here a one-dimensional LMM) is

$$\sigma_{\mathrm{P}}^{\bar{T}}(t) = -\sum_{i=\beta(t)}^{\bar{N}} \frac{\tau_i \sigma_i(t)}{1+\tau_i L_i(t)} \tag{16}$$

with $\sigma_i(t)$ the (normal) volatility of the Libor L_i. The instantaneous rate is

$$r_t = \frac{L_{\beta(t)-1}(T_{\beta(t)-2})}{1 + (T_{\beta(t)-1} - t) L_{\beta(t)-1}(T_{\beta(t)-2})} \tag{17}$$

This expression requires only the computation of the Libors L_i at each tenor date $\{T_{i-1}\}$.[3] Expression (8) can then be applied with $\sigma_{\mathrm{P}}^{\bar{T}}(t)$ given by (16) and

$$\sigma_{\mathrm{r}}^{T}(s) = \mathbb{E}_s^{\mathbb{P}^T}[D_s^{B^T} r_T] = \mathbb{E}_s^{\mathbb{P}^T}\left[\frac{D_s^{B^T} L_{\beta(T)-1}(T_{\beta(T)-2})}{\left(1 + (T_{\beta(T)-1} - T) L_{\beta(T)-1}(T_{\beta(T)-2})\right)^2}\right]$$

By explicitly computing $D_s^{B^T} S_T$, (8) can then be approximated by (see Proof 3, assuming here a one-dimensional LMM with a local volatility $\sigma_i(t) = \sigma_i(t, L_t^i)$)

$$P_{0T}\frac{\mathbb{E}^{\mathbb{P}^T}\left[\left(r_T - r_T^0\right) 1_{S_T > K}\right]}{\frac{1}{2} K \partial_K^2 \mathcal{C}(T,K)} = 2\mathbb{E}^{\mathbb{P}^T}\left[V_T \left(\rho U_T^{\beta(T)-1} + \Theta_T^{\beta(T)-1} \Xi_T^{\beta(T)-1} - \Lambda_T^{\beta(T)-1}\right) \Big| S_T = K\right] \tag{18}$$

[3]Note that in a Markov functional model, Libors $L_i(t) \equiv L_i(t, B_t^{\bar{T}})$ are functions of a one-dimensional $\mathbb{P}^{\bar{T}}$-Brownian motion $B_t^{\bar{T}}$ and are (usually) stored as splines at each tenor date.

with

$$\frac{dV_t}{V_t} = S_t \partial_S \sigma(t,S_t) a_t \left(dW_t - a_t \sigma(t,S_t)dt\right), \quad V_0 = 1$$

$$dU_t^{\beta(T)-1} = \sigma_{\beta(T)-1}(t) a_t \frac{\sigma(t,S_t)}{V_t} dt, \quad U_0^{\beta(T)-1} = 0$$

$$\frac{dR_t^{\beta(T)-1}}{R_t^{\beta(T)-1}} = \partial_{L^{\beta(T)-1}} \sigma_{\beta(T)-1}(t) dB_t^{T_{\beta(T)-1}}, \quad R_0^{\beta(T)-1} = 1$$

$$d\Theta_t^{\beta(T)-1} = \frac{R_t^{\beta(T)-1}}{V_t} dt, \quad \Theta_0^{\beta(T)-1} = 0$$

$$d\Xi_t^{\beta(T)-1} = \frac{(\sigma_t^{\beta(T)-1})^2}{R_t^{\beta(T)-1}} dt, \quad \Xi_0^{\beta(T)-1} = 0, \qquad d\Lambda_t^{\beta(T)-1} = \Theta_t^{\beta(T)-1} d\Xi_t^{\beta(T)-1}, \quad \Lambda_0^{\beta(T)-1} = 0$$

This expression requires the simulation of the processes f_t, L_t^i, V_t, R_t^i and the integrals U_t^i, Θ_t^i, Ξ_t^i, Λ_t^i.

Local correlation model

Let us consider an index $I_t = \sum_{i=1}^N \alpha_i S_t^i$ made of N weighted stocks, each modeled using a local volatility:

$$dS_t^i = S_t^i \sigma^i(t, S_t^i) dW_t^i, \qquad d\langle W^i, W^j \rangle_t = \rho_{ij}(t, S_t) dt$$

$\{W^i\}$ denotes a multi-dimensional Brownian motion with an instantaneous correlation function of the time and the N stock values $S_t = \left(S_t^1, \ldots, S_t^N\right)$. This model is calibrated to the basket smile if and only if[4]

$$I^2 \sigma_{\mathrm{Dup}}^I(t,I)^2 = \sum_{i,j=1}^N \alpha_i \alpha_j \mathbb{E}[\rho_{ij}(t,S_t) \sigma^i(t,S_t^i) \sigma^j(t,S_t^j) S_t^i S_t^j | I_t = I] \tag{19}$$

where σ_{Dup}^I denotes the Dupire local volatility of the index. Assume that the instantaneous correlation depends on the N stock values only through the index value [6]:

$$\rho_{ij}(t,I) = (1 - \lambda(t,I)) \rho_{ij}^{\mathrm{hist}} + \lambda(t,I), \qquad \lambda(t,I) \in [0,1]$$

The calibrated model then follows the McKean SDE

$$\begin{aligned} dS_t^i &= S_t^i \sigma^i(t,S_t^i) dW_t^i, \qquad d\langle W^i, W^j \rangle_t = \rho_{ij}(t, I_t) dt \\ \lambda(t,I) &= \frac{I^2 \sigma_{\mathrm{Dup}}^B(t,I)^2 - \sum_{i,j=1}^N \alpha_i \alpha_j \rho_{ij}^{\mathrm{hist}} \mathbb{E}[\sigma^i(t,S_t^i) \sigma^j(t,S_t^j) S_t^i S_t^j | I_t = I]}{\sum_{i,j=1}^N \alpha_i \alpha_j (1 - \rho_{ij}^{\mathrm{hist}}) \mathbb{E}[\sigma^i(t,S_t^i) \sigma^j(t,S_t^j) S_t^i S_t^j | I_t = I]} \end{aligned}$$

Although our choice of local correlation is convenient for calibration purposes, it is not at all obvious that this SDE is well-posed, i.e., that there exits a local correlation that matches the basket smile. In particular, the calibrated λ may not lie in $[0,1]$.

[4]Note that in [9], the author claims that at each point in time, in order to exclude arbitrage opportunities, we must have the stronger statement $I_t^2 \sigma_{\mathrm{Dup}}^I(t,I_t)^2 = \sum_{i,j=1}^N \alpha_i \alpha_j \rho_{ij}(t,S_t) \sigma^i(t,S_t^i) \sigma^j(t,S_t^j) S_t^i S_t^j$. This is incorrect as only the conditional expected value of the basket variance given the index value matters.

Particle method

Principle. The stochastic simulation of the McKean SDE (4) is very natural. It consists in replacing the law $\mathbb{P}_t$, which appears explicitly in the drift and diffusion coefficients, by its approximation given by the empirical distribution

$$\mathbb{P}_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}}$$

where the $X_t^{i,N}$ are solution to the $(\mathbb{R}^n)^N$-dimensional linear SDE

$$dX_t^{i,N} = b\left(t, X_t^{i,N}, \mathbb{P}_t^N\right)\, dt + \sigma\left(t, X_t^{i,N}, \mathbb{P}_t^N\right)\cdot dW_t^i, \qquad \text{Law}\left(X_0^{i,N}\right) = \mathbb{P}_0$$

$\{W_t^i\}_{i=1,\dots,N}$ are N independent d-dimensional Brownian motions; $\mathbb{P}_t^N$ is a random measure on $\mathbb{R}^n$. In the case of the McKean-Vlasov SDE where

$$\begin{aligned} b^i\left(t, x, \mathbb{P}_t\right) &= \int b^i(t,x,y)p(t,y)dy \;=\; \mathbb{E}\left[b^i(t,x,X_t)\right] \\ \sigma_j^i\left(t, x, \mathbb{P}_t\right) &= \int \sigma_j^i(t,x,y)p(t,y)dy \;=\; \mathbb{E}\left[\sigma_j^i(t,x,X_t)\right] \end{aligned}$$

$\{X_t^{i,N}\}_{i=1,\dots,N}$ are n-dimensional Itô processes given by

$$dX_t^{i,N} = \left(\int b(t, X_t^{i,N}, y)d\mathbb{P}_t^N(y)\right)dt + \left(\int \sigma(t, X_t^{i,N}, y)d\mathbb{P}_t^N(y)\right)\cdot dW_t^i$$

which is equivalent to

$$dX_t^{i,N} = \frac{1}{N}\sum_{j=1}^N b\left(t, X_t^{i,N}, X_t^{j,N}\right)dt + \frac{1}{N}\sum_{j=1}^N \sigma\left(t, X_t^{i,N}, X_t^{j,N}\right)dW_t^i$$

One can then show the *chaos propagation property* [13]: If at $t=0$, $X_0^{i,N}$ are independent particles then as $N\to\infty$, for any fixed $t>0$, the $X_t^{i,N}$ are asymptotically independent and their empirical measure $\mathbb{P}_t^N$ converges in distribution towards the true measure $\mathbb{P}_t$. This means that, in the space of probabilities on the space of probabilities, the distribution of the random measure $\mathbb{P}_t^N$ converges towards a Dirac mass at the deterministic measure $\mathbb{P}_t$. Practically, it means that for bounded continuous functions f:

$$\frac{1}{N}\sum_{i=1}^N f(X_t^{i,N}) \xrightarrow[N\to\infty]{\mathrm{L}^1} \int_{\mathbb{R}^d} f(x)p(t,x)dx$$

where $p(t,\cdot)$ is the fundamental solution of the non-linear Fokker-Planck PDE (5) - see [13].

Using an analogy with the mean-field approximation in statistical physics, the N processes $\{X_t^i\}_{i=1,\dots,N}$ can be seen as a system of N interacting (bosonic) particles. In the large N limit, the $(\mathbb{R}^n)^N$-dimensional (linear) Fokker-Planck PDE approximates the non-linear low-dimensional (n-dimensional) Fokker-Planck PDE (5).[5] Then, the resulting drift and diffusion coefficients of $X_t^{i,N}$ depend not only on the position of the particle $X_t^{i,N}$ but also on the interaction with the other $N-1$ particles.

[5] The inverse procedure is common in Physics. For example, the large N limit of N interacting (fermionic) particles modeled by a $3\times N$-dimensional Schrödinger equation can be described by a 3-dimensional non-linear PDE, the so-called Hartree-Fock PDE.

Local stochastic volatility model. In the LSV model, the approximated conditional expectation

$$\mathbb{E}^{\mathbb{P}_t^N}[a_t^2|f_t = f] = \frac{\int a'^2 p_N(t,f,a')da'}{\int p_N(t,f,a')da'} = \frac{\sum_{i=1}^N (a_t^{i,N})^2 \delta(f_t^{i,N} - f)}{\sum_{i=1}^N \delta(f_t^{i,N} - f)}$$

is not properly defined. Like in the Nadaraya-Watson regression, instead of the Dirac function $\delta(\cdot)$, we use a regularizing kernel $\delta_{t,N}(\cdot)$, define

$$\sigma_N(t,f) = \sigma_{\mathrm{Dup}}(t,f)\sqrt{\frac{\sum_{i=1}^N \delta_{t,N}\left(f_t^{i,N} - f\right)}{\sum_{i=1}^N \left(a_t^{i,N}\right)^2 \delta_{t,N}\left(f_t^{i,N} - f\right)}} \tag{20}$$

and simulate

$$df_t^{i,N} = f_t^{i,N}\sigma_N(t, f_t^{i,N})a_t^{i,N}dW_t^i \tag{21}$$

A similar algorithm has been used in [8] in the case of the joint calibration of smiles of a basket and its components.

Acceleration techniques. At first sight, (20)-(21) requires $O(N^2)$ operations at each discretization date: each calculation of $\sigma_N(t, f_t^{i,N})$ requires $O(N)$ operations, and there are N such local volatilities to compute. This naive method is too slow. Below, we present two crucial acceleration techniques that make our method efficient in practice.

First, computing $\sigma_N(t, f_t^{i,N})$ for all i is useless. One can save considerable time by computing $\sigma_N(t,f)$ for only a grid $G_{f,t}$ of values of f, of a size much smaller than N, say $N_{f,t}$, and then inter- and extrapolate. We use cubic splines, with a flat extrapolation, and $N_{f,t} = \max(N_f\sqrt{t}, N'_f)$; typical values are $N_f = 30$ and $N'_f = 15$. The range of the grid can be inferred from the prices of digital options: $\mathbb{E}[1_{f_t > \max G_{f,t}}] = \mathbb{E}[1_{f_t < \min G_{f,t}}] = \alpha$. In practice, we take $\alpha = 10^{-3}$.

Second, in the sums in (20), a large number of terms make a negligible contribution: we can disregard $f_t^{i,N}$ when it is far from f, say, when $\delta_{t,N}(f_t^{i,N} - f)$ is smaller than some threshold η. In practice, this requires sorting particles according to the spot value. The cost of sorting, $O(N \ln N)$, is more than compensated by the acceleration in the $N_{f,t}$ evaluations of (20).

Particle algorithm. Let $\{t_k\}$ denote a time discretization of $[0,T]$. The particle algorithm can now be described by the following steps:

(1) Initialize $k := 1$ and set $\sigma_N(t,f) = \frac{\sigma_{\mathrm{Dup}}(0,f)}{a_0}$ for all $t \in [t_0 = 0, t_1]$.
(2) Simulate the N processes $\{f_t^{i,N}, a_t^{i,N}\}_{i=1,\dots,N}$ from t_{k-1} to t_k according to (21).
(3) Sort the particles according to spot value. For $f \in G_{f,t_k}$, find the smallest index $\underline{i}(f)$ and the largest index $\bar{i}(f)$ for which $\delta_{t_k,N}\left(f_{t_k}^{i,N} - f\right) > \eta$, and compute the local volatility

$$\sigma_N(t_k,f) = \sigma_{\mathrm{Dup}}(t_k,f)\sqrt{\frac{\sum_{i=\underline{i}(f)}^{\bar{i}(f)} \delta_{t_k,N}\left(f_{t_k}^{i,N} - f\right)}{\sum_{i=\underline{i}(f)}^{\bar{i}(f)} \left(a_{t_k}^{i,N}\right)^2 \delta_{t_k,N}\left(f_{t_k}^{i,N} - f\right)}}$$

Interpolate the local volatility using cubic splines, and extrapolate flat outside the interval $[\min G_{f,t_k}, \max G_{f,t_k}]$. Set $\sigma_N(t,f) \equiv \sigma_N(t_k,f)$ for all $t \in [t_k, t_{k+1}]$.
(4) Set $k := k+1$. Iterate steps 2 and 3 up to the maturity date T.

Step 2 is easily parallelizable. Calibration and pricing can be achieved in the course of the same Monte-Carlo simulation. We only need to ensure that all spot observation dates needed in the calculation of the payout are included in the time discretization $\{t_k\}$. The price of an option is estimated as $\frac{1}{N}\sum_{i=1}^{N} H^{i,N}$ where $H^{i,N}$ is the discounted payout evaluated on the path of particle i.

Regularizing kernel. It is natural to take $\delta_{t,N}(x) = \frac{1}{h_{t,N}} K\left(\frac{x}{h_{t,N}}\right)$ where K is a fixed, symmetric kernel with a bandwidth $h_{t,N}$ that tends to zero as N grows to infinity. The exponential kernel $K(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-x^2/2\right)$ and the quartic kernel $K(x) = \frac{15}{16}\left(1-x^2\right)^2 1_{\{|x|\leq 1\}}$ are typical examples. We use the latter because it saves computational time. We take

$$h_{t,N} = \kappa f_0 \sigma_{\mathrm{VS},t} \sqrt{\max\left(t, t_{\min}\right)} N^{-\frac{1}{5}}$$

with $\sigma_{\mathrm{VS},t}$ the variance swap volatility at maturity t. The factor $N^{-\frac{1}{5}}$ comes from the minimization of the asymptotic mean integrated squared error of the Nadaraya-Watson estimator, which is the sum of two terms: bias and variance. The smaller the bandwidth, the smaller the bias, but the larger the variance. The critical bandwidth that minimizes the sum of bias and variance decreases as $N^{-\frac{1}{5}}$ for large N. Following Silverman's rule of thumb [14], the prefactor $\kappa f_0 \sigma_{\mathrm{VS},t}\sqrt{\max\left(t, t_{\min}\right)}$ is of order the standard deviation of the regressor f_t. The fine-tuning of κ is crucial. In practice, we take $\kappa \simeq 1.5$, $t_{\min} = 1/4$.

Hybrid local stochastic volatility model. In the case of hybrid LSV model, a particle is described by three processes (f_t, a_t, r_t). If we use representation (7) of the hybrid local volatility, we define

$$\sigma_N(t,S)^2 = \left(\sigma_{\mathrm{Dup}}(t,S)^2 - P_{0\bar{T}} \frac{\frac{1}{N}\sum_{i=1}^{N}\left(P_{t\bar{T}}^{i,N}\right)^{-1}\left(r_t^{i,N} - r_t^0\right) 1_{S_t^{i,N} > S}}{\frac{1}{2} K \partial_K^2 \mathcal{C}(t,K)}\right) \frac{\sum_{i=1}^{N}\left(P_{t\bar{T}}^{i,N}\right)^{-1} \delta_{t,N}\left(S_t^{i,N} - S\right)}{\sum_{i=1}^{N}\left(P_{t\bar{T}}^{i,N}\right)^{-1}\left(a_t^{i,N}\right)^2 \delta_{t,N}\left(S_t^{i,N} - S\right)} \tag{22}$$

and simulate

$$df_t^{i,N} = f_t^{i,N} \sigma_N(t, f_t^{i,N} P_{t\bar{T}}^{i,N}) a_t^{i,N} dW_t^i - f_t^{i,N} \sigma_{\mathrm{P}}^{\bar{T}}(t).dB_t^{\bar{T}}$$

Because of the unconditional expectation estimator $\frac{1}{N}\sum_{i=1}^{N}(P_{t\bar{T}}^{i,N})^{-1}(r_t^{i,N} - r_t^0)1_{S_t^{i,N}>S}$, and after sorting the particles, the $N_{S,t}$ evaluations of (22) cost approximately $N_{S,t}N/2$. For usual (small) values of the bandwidth $h_{t,N}$, this is more than the cost $N_{S,t}N_h$ of computing a conditional expectation estimator using the second acceleration technique described above, where N_h is the average number of particles which contribute to the estimator. If we use representation (15) of the hybrid local volatility, we need to add the Malliavin processes to the particle, which means more processes to simulate, but now only conditional expectations need to be estimated, which means that we can use the second acceleration technique, and this saves much computation time.

Numerical Tests

Ho-Lee/Dupire hybrid model. We consider a hybrid local volatility model ($a_t \equiv 1$) where the short rate follows a Ho-Lee model, for which the volatility $\sigma_{\mathrm{r}}(s) = \sigma_{\mathrm{r}}$ is a constant. A bond of maturity $\bar{T}$ is given by

$$P_{t\bar{T}} = \frac{P^{\text{mkt}}_{0\bar{T}}}{P^{\text{mkt}}_{0t}} e^{\frac{\sigma_{\text{r}}^2(\bar{T}-t)^2 t}{2} - \sigma_{\text{r}}(\bar{T}-t)B^{\bar{T}}_t}$$

with a volatility $\sigma^{\bar{T}}_{\text{P}}(t) = -\sigma_{\text{r}}(\bar{T}-t)$. From (15), the local volatility is

$$\sigma\left(t, K, \mathbb{P}^{\bar{T}}_t\right)^2 = \sigma_{\text{Dup}}(t,K)^2 - 2\rho\sigma_{\text{r}} \frac{\mathbb{E}^{\mathbb{P}^{\bar{T}}}[P^{-1}_{t\bar{T}} V_t U_t | S_t = K]}{\mathbb{E}^{\mathbb{P}^{\bar{T}}}[P^{-1}_{t\bar{T}} | S_t = K]} - 2\sigma_{\text{r}}^2 \frac{\mathbb{E}^{\mathbb{P}^{\bar{T}}}[P^{-1}_{t\bar{T}} V_t \left(t\Theta_t - \Lambda_t\right) | S_t = K]}{\mathbb{E}^{\mathbb{P}^{\bar{T}}}[P^{-1}_{t\bar{T}} | S_t = K]}$$

with

$$\begin{aligned} \frac{dV_t}{V_t} &= S_t \partial_S \sigma(t, S_t) \left(dW^{\bar{T}}_t + \left(\rho \sigma^{\bar{T}}_{\text{P}}(t) - \sigma(t, S_t)\right) dt \right) \ , \ V_0 = 1 \\ U_t &= \int_0^t \frac{\sigma(s, S_s)}{V_s} ds, \qquad \Theta_t = \int_0^t \frac{ds}{V_s}, \qquad \Lambda_t = \int_0^t \Theta_s ds \end{aligned}$$

As a sanity check, when σ_{Dup} depends only on the time t, we obtain the exact expression for $\sigma(\cdot)$ as expected:

$$\sigma(t)^2 = \sigma_{\text{Dup}}(t)^2 - 2\rho\sigma_{\text{r}} \int_0^t \sigma(s) ds - \sigma_{\text{r}}^2 t^2$$

Note that in [3], the local volatility is approximated by

$$\sigma(t,K)^2 \approx \sigma_{\text{Dup}}(t,K)^2 - 2\rho\sigma_{\text{r}} \int_0^t \sigma(s,K) ds$$

This equation in $\sigma(t,K)$ is then solved using a fixed-point method. Practitioners typically use such approximations for $\sigma(t,K)$ whose quality deteriorates significantly far out of the money or for long maturities. We emphasize that even in the simple case where σ_{Dup} depends only on the time t, the above approximation is not exact because of the missing term $\sigma_{\text{r}}^2 t^2$. Our algorithm achieves exact calibration in this case with a single particle!

We have checked the accuracy of our calibration procedure on the DAX market smile (30-May-11). We have chosen $\sigma_{\text{r}}(s) = 6.3$ bps per day (1% per year) and set the correlation between the stock and the rate to $\rho = 40\%$. The time discretization $\Delta = t_{k+1} - t_k$ has been set to $\Delta = 1/100$ and we have used $N = 2^{10}$ or $N = 2^{12}$ particles. After calibrating the model using the particle algorithm, we have computed vanilla smiles using a (quasi) Monte Carlo pricer with $N = 2^{15}$ paths and a timestep of 1/250. Figure 1 shows the implied volatility for the market smile (DAX, 30-May-11) and the hybrid LV model for maturities 4 years and 10 years. When we use the Malliavin representation, the computational time is around 4 seconds for maturities up to 10 years with $N = 2^{10}$ partciles (12 seconds with $N = 2^{12}$). Our algorithm definitively outperforms a (two-dimensional) PDE implementation and has already converged with $N = 2^{10}$ particles. Note that the calibration is also exact using Equation (22), i.e., with no use of the Malliavin representation, but with a larger computational time of 8 seconds with $N = 2^{10}$ particles (26 seconds with $N = 2^{12}$). As shown in Table 1, the absolute error in implied volatility is of a few basis points. For completeness, we have plotted the smile obtained from hybrid LV model without any calibration, i.e., $\sigma(t,K) = \sigma_{\text{Dup}}(t,K)$, to materialize the impact of the stochastic rate.

Strike	0.5	0.7	0.8	0.9	1	1.1	1.2	1.3	1.5	1.8
with Malliavin, time = 4s	14	10	10	9	8	7	6	5	3	1
without Malliavin, time = 8s	16	8	7	4	1	1	1	3	3	5

Table 1. DAX (30-May-11) Implied volatilities $T = 10$Y. Errors in bps using the particle method with $N = 2^{10}$ particles.

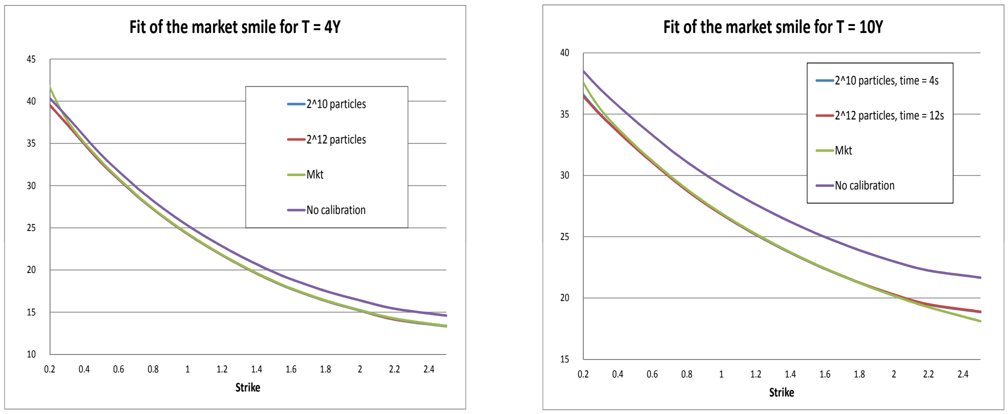


FIGURE 1. DAX (30-May-11) Implied volatilities $T = 4$Y, $T = 10$Y. Ho-Lee parameters: $\sigma_{\mathrm{r}}(s) = 6.3$ bps per day, $\rho = 40\%$. Computation time with $\Delta = 1/100$, $N = 2^{10}$ on a full 10Y implied volatility surface with a Intel Core Duo©, 3 Ghz, 3 GB of Ram: 4s.

Bergomi's local stochastic volatility model. As a next example, we consider Bergomi's LSV model [4, 7]:

$$\begin{aligned} df_t &= f_t \sigma(t, f_t) \sqrt{\xi_t^t} dW_t \\ \xi_t^T &= \xi_0^T f^T(t, x_t^T) \\ f^T(t, x) &= \exp(2\sigma x - 2\sigma^2 h(t, T)) \\ x_t^T &= \alpha_\theta \left((1-\theta) e^{-k_X (T-t)} X_t + \theta e^{-k_Y (T-t)} Y_t \right) \\ \alpha_\theta &= \left((1-\theta)^2 + \theta^2 + 2\rho_{XY} \theta (1-\theta) \right)^{-1/2} \\ dX_t &= -k_X X_t dt + dW_t^X \\ dY_t &= -k_Y Y_t dt + dW_t^Y \end{aligned}$$

where

$$\begin{aligned} h(t, T) &= (1-\theta)^2 e^{-2k_X(T-t)} \mathbb{E}\left[X_t^2\right] + \theta^2 e^{-2k_Y(T-t)} \mathbb{E}\left[Y_t^2\right] + 2\theta(1-\theta) e^{-(k_X+k_Y)(T-t)} \mathbb{E}\left[X_t Y_t\right] \\ \mathbb{E}\left[X_t^2\right] &= \frac{1 - e^{-2k_X t}}{2k_X}, \quad \mathbb{E}\left[Y_t^2\right] = \frac{1 - e^{-2k_Y t}}{2k_Y}, \quad \mathbb{E}\left[X_t Y_t\right] = \rho_{XY} \frac{1 - e^{-(k_X+k_Y)t}}{k_X + k_Y} \end{aligned}$$

This model, commonly used by practitioners, is a variance swap curve model which admits a two-dimensional Markovian representation. We have performed similar tests as in the previous section (see Figure 2). The Bergomi model parameters are $\sigma = 200\%$ (the volatility of an ultra-short volatility), $\theta = 22.65\%$, $k_X = 4$, $k_Y = 12.5\%$, $\rho_{XY} = 30\%$, $\rho_{\mathrm{SX}} = -50\%$, $\rho_{\mathrm{SY}} = -50\%$. The time discretization has been fixed to $\Delta = 1/100$ and we have used $N = 2^{10}$, $N = 2^{12}$ or $N = 2^{13}$ particles. Figure 2 shows the implied volatility for the market smile (DAX, 30-May-11) and the LSV model for maturities 4 years and 10 years. The computational time is 4 seconds for maturities up to 10 years with $N = 2^{10}$ particles (11 seconds with $N = 2^{12}$). This should be compared with the approximate calibration [7] with a computational time around 12 seconds. In order to illustrate

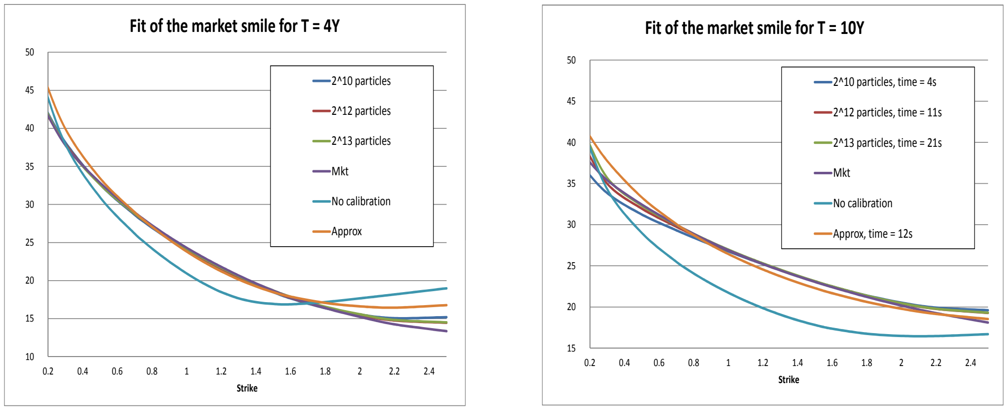


FIGURE 2. DAX (30-May-11) Implied volatilities $T = 4\text{Y}$, $T = 10\text{Y}$. Bergomi parameters: $\sigma = 200\%$, $\theta = 22.65\%$, $k_X = 4$, $k_Y = 12.5\%$, $\rho_{XY} = 30\%$, $\rho_{\text{SX}} = -50\%$, $\rho_{\text{SY}} = -50\%$.

that we have used stressed parameters to check the efficiency of our algorithm, we have plotted the smile produced by the naked SVM which significantly differs from the market smile.

Existence under question. As highlighted previously, the existence of LSV models for a given market smile is not at all obvious although this seems to be a common belief in the quant community. In order to illustrate this mathematical question, we have decided to check our algorithm with a volatility-of-volatility $\sigma = 350\%$ (see Figure 3). This large value of volatility-of-volatility is sometimes needed in order to generate typical levels of forward skew for indices, in a model where skew is generated by volatility-of-volatility only.

Our algorithm seems to converge with $N = 2^{13}$ particles around the money but the market smile is not calibrated. For the maturity $T = 4$ years, we have an error of around 61 bps at-the-money which increases to 245 bps for $K = 2$. This may come from numerical issues, or from non-existence of a solution. For comparison, we graph the result of an approximate calibration [7] which definitively breaks down for high levels of volatility-of-volatility.

Local Bergomi and Ho-Lee hybrid model. We now go one step further in complexity and consider a Bergomi LSV model with Ho-Lee stochastic rates. We should emphasize that since this model is driven by four Brownian motions, a calibration relying on a PDE solver is out of the picture. The Bergomi model parameters are those used in the previous section. Additionally, we have chosen $\sigma_{\text{r}}(s) = 6.3$ bps per day and set the correlation between the stock and the rate to $\rho = 40\%$ (see Figure 4). The computational time is 4 seconds for maturities up to 10 years with $N = 2^{10}$ particles (20 seconds with $N = 2^{12}$).

CONCLUSION

In this paper, we have explained how to calibrate *exactly* multi-factor hybrid local stochastic volatility models to market smiles using particle methods. We have also provided a Malliavin

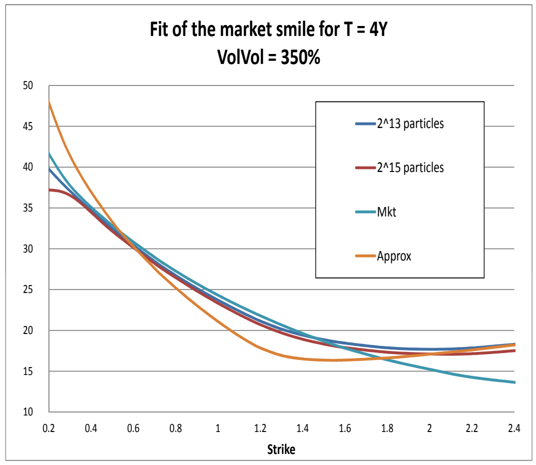
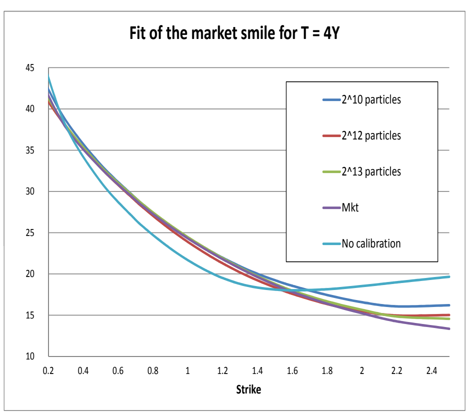


FIGURE 3. DAX (30-May-11) Implied volatilities $T = 4$Y. Bergomi parameters: $\sigma = \mathbf{350}\%$, $\theta = 22.65\%$, $k_X = 4$, $k_Y = 12.5\%$, $\rho_{XY} = 30\%$, $\rho_{\mathrm{SX}} = -50\%$, $\rho_{\mathrm{SY}} = -50\%$.



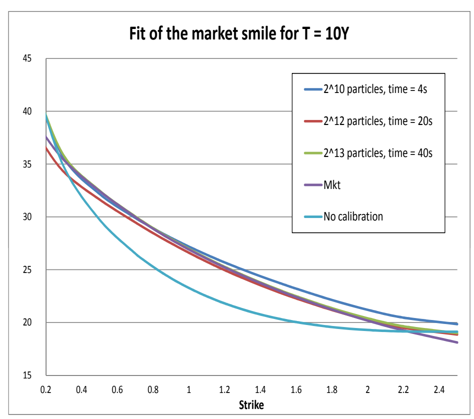


FIGURE 4. DAX (30-May-11) Implied volatilities $T = 4$Y, $T = 10$Y. Bergomi parameters: $\sigma = 200\%$, $\theta = 22.65\%$, $k_X = 4$, $k_Y = 12.5\%$, $\rho_{XY} = 30\%$, $\rho_{\mathrm{SX}} = -50\%$, $\rho_{\mathrm{SY}} = -50\%$. Ho-Lee parameters: $\sigma_{\mathrm{r}}(s) = 6.3$ bps per day, $\rho = 40\%$.

stochastic representation of the hybrid contribution to local volatility. The technique we proposed proves useful when models incorporating multiple volatility and interest rate risks are needed, typically for long-dated, forward skew-sensitive payoffs. Our algorithm represents, to the best of our knowledge, the first *exact* algorithm for the calibration of multi-factor hybrid local stochastic volatility models. Acceleration techniques make it efficient in practice. As highlighted in our numerical experiments, the computation time is excellent and even for low-dimensional (hybrid) LSV models, our algorithm outperforms PDE implementations.

The analysis of non-linear (kinetic) PDEs arising in statistical physics such as the McKean-Vlasov PDE and the Boltzmann equation has recently become more popular and has recently drawn attention in part thanks to the work of Fields medallist Prof. C. Villani. We hope that this work will initiate new research and attract the attention of practitioners to the world of non-linear SDEs.

Dr Julien Guyon and **Dr Pierre Henry-Labordère** are members of the Global Markets Quantitative Research team at Société Générale. They wish to thank their colleagues for useful discussions. Email: julien.guyon@sgcib.com, pierre.henry-labordere@sgcib.com. This paper is dedicated to the memory of Prof. P. Malliavin who pointed out to us the efficiency of the Clark-Ocone formula as we were working on this project.

Appendix

Proof 1. By applying Itô-Tanaka's formula on a discounted vanilla call payoff with maturity t and strike K, $\mathcal{P}_t \equiv D_{0t}(S_t - K)^+$, we have:

$$\begin{aligned} d\mathcal{P}_t &= -D_{0t}(S_t-K)^+ r_t dt + D_{0t} 1_{S_t>K} S_t \left(r_t dt + a_t \sigma(t,S_t) dW_t\right) \\ &+ \frac{1}{2} K^2 a_t^2 \sigma(t,K)^2 D_{0t} \delta(S_t - K) dt \end{aligned}$$

By taking the expectation $\mathbb{E}^{\mathbb{P}}[\cdot]$ on both sides of the above equation and by assuming that $M_t = \int_0^t D_{0t} 1_{S_t>K} a_s \sigma(s,S_s) dW_s$ is a true martingale, we get

$$\partial_t \mathcal{C}(t,K) = K\mathbb{E}^{\mathbb{P}}[D_{0t} 1_{S_t>K} r_t] + \frac{1}{2} K^2 \sigma(t,K)^2 \mathbb{E}^{\mathbb{P}}[D_{0t} a_t^2 \delta(S_t - K)]$$

Then, by doing a change of measure from $\mathbb{P}$ to $\mathbb{P}^t$ and using that $\partial_K^2 \mathcal{C}(t,K) = P_{0t} \mathbb{E}^{\mathbb{P}^t}[\delta(S_t - K)]$, we deduce our final result (6).

Proof 2. In the measure $\mathbb{P}^T$, we have

$$dS_t = \left(r_t S_t + \rho \sigma_P^T(t) \sigma(t,S_t) S_t a_t\right) dt + \sigma(t,S_t) S_t a_t \left(\rho dB_t^T + \sqrt{1-\rho^2} dZ_t^T\right)$$

with $d\langle B^T, Z^T\rangle_t = 0$. By applying the derivative $D_s^{B^T}$ on the above SDE, we obtain

$$\begin{aligned} dD_s^{B^T} S_t &= \left(D_s^{B^T} r_t \left(S_t + \rho \partial_r(\sigma_P^T(t)) \sigma(t,S_t) S_t a_t\right) + \left(r_t + \rho \sigma_P^T(t) \partial_S(\sigma(t,S_t) S_t) a_t\right) D_s^{B^T} S_t\right) dt \\ &+ \partial_S(S_t \sigma(t,S_t)) a_t D_s^{B^T} S_t \left(\rho dB_t^T + \sqrt{1-\rho^2} dZ_t^T\right) \end{aligned}$$

with $D_s^{B^T} S_s = \rho S_s \sigma(s,S_s) a_s$. We have used our hypothesis: $D_s^{B^T} a_t = 0$. Then the reader can easily check that the solution $D_s^{B^T} S_T$ can be written as a sum of adapted processes (see [6] for detailed similar computations)

$$D_s^{B^T} S_T = \left(\rho \sigma(s,S_s) S_s a_s - \frac{\sigma_{\mathrm{r}}^T(s)}{R_s^T} \theta_s^T\right) \frac{Y_T}{Y_s} + \frac{\sigma_{\mathrm{r}}^T(s)}{R_s^T} \theta_T^T \, , \; s \le T \tag{23}$$

with $D_s^{B^T} r_t = \sigma_r^T(s) \frac{R_t^T}{R_s^T} 1_{t\geq s}$, R_t^T defined by (12) and

$$\begin{aligned}
\frac{dY_t}{Y_t} &= \left(r_t + \rho\sigma_{\mathrm{P}}^T(t)\partial_S(\sigma(t,S_t)S_t)a_t\right)dt + \partial_S(\sigma(t,S_t)S_t)a_t dW_t^T \\
d\theta_t^T &= \theta_t^T \frac{dY_t}{Y_t} + S_t R_t^T dt + \rho S_t \sigma(t,S_t) a_t \partial_r(\sigma_{\mathrm{P}}^T) R_t^T dt
\end{aligned}$$

We have set $W^T = \rho B^T + \sqrt{1-\rho^2} Z^T$, $\Theta_t^T = \theta_t^T / Y_t$, $V_t = Y_t / S_t$ and $dU_t = \sigma_{\mathrm{r}}(t, r_t)\sigma(t,S_t)S_t a_t / V_t dt$. By plugging formula (23) into (8), we get our final expression (9).

Proof 3. Proceeding as above, we get

$$dD_s^{B^T} S_t = \left(S_t D_s^{B^T} r_t + \rho D_s^{B^T}(\sigma_P^T(t))\sigma(t,S_t)S_t a_t\right)dt + \frac{dY_t}{Y_t} D_s^{B^T} S_t$$

where r_t is given by (17) and

$$\sigma_{\mathrm{P}}^T(t) = -\sum_{i=\beta(t)}^{\beta(T)-1} \frac{\hat{\tau}_i \sigma_i}{1+\hat{\tau}_i L_i(t)}$$

with $\hat{\tau}_i = \tau_i,\ i = \beta(t), \ldots, \beta(T)-2$ and $\hat{\tau}_{\beta(T)-1} = T - T_{\beta(T)-2}$. Here, as an approximation at first order in the volatility, we have $D_s^{B^T} L_i(t) \simeq D_s^{B^{T_i}} L_i(t) = \sigma_i(s) \frac{R^i(t)}{R^i(s)} 1_{t\geq s}$ with R_t^i the tangent process to L_t^i. Then by neglecting the factor $1/\left(1 + (T_{\beta(t)-1} - t)L_{\beta(t)-1}(T_{\beta(t)-2})\right)$ in r_t, we obtain

$$\begin{aligned}
D_s^{B^T} r_t &\simeq \sigma_{\beta(t)-1}(s) \frac{R_{T_{\beta(t)-2}}^{\beta(t)-1}}{R_s^{\beta(t)-1}} 1_{T_{\beta(t)-2}\geq s} \\
\sigma_{\mathrm{r}}^T(s) &\simeq \sigma_{\beta(T)-1}(s) 1_{T_{\beta(T)-2}\geq s} \\
D_s^{B^T} \sigma_{\mathrm{P}}^T(t) &\simeq -\sum_{i=\beta(t)}^{\beta(T)-1} \partial_i \left(\frac{\hat{\tau}_i \sigma_i}{1+\hat{\tau}_i L_i(t)}\right) \sigma_i(s) \frac{R_t^i}{R_s^i} 1_{t\geq s}
\end{aligned}$$

The reader can easily check that the solution $D_s^{B^T} S_T$ can then be written as

$$D_s^{B^T} S_T = \left(\rho\sigma(s,S_s)S_s a_s - \frac{\sigma_{\beta(T)-1}(s)}{R_s^{\beta(T)-1}} \theta_s^{\beta(T)-1}\right)\frac{Y_T}{Y_s} + \frac{\sigma_{\beta(T)-1}(s)}{R_s^{\beta(T)-1}} \theta_T^{\beta(T)-1}\ ,\ s \leq T \tag{24}$$

with

$$d\theta_t^{\beta(t)-1} = \frac{dY_t}{Y_t} \theta_t^{\beta(t)-1} + S_t R_{T_{\beta(t)-2}}^{\beta(t)-1} dt$$

By plugging formula (24) into (8), we get our final expression (18).

Including (discrete) dividends. We assume that the spot process S_t jumps down by the dividend amount $D(t_i, S_{t_i^-})$ paid at time t_i and that between dividend dates $\{t_i\}$ it follows a (hybrid) LSV model under a risk-neutral measure $\mathbb{P}$

$$S_t = S_0 + \int_0^t r_s S_s ds + \int_0^t S_s \sigma(s,S_s) a_s dW_s - \sum_{t_i \leq t} D(t_i, S_{t_i^-})$$

This is in agreement with market prices only if

$$\mathcal{C}(t_i, K) = \mathbb{E}^{\mathbb{P}^{\mathrm{mkt}}}\left[D_{0t_i}\left(S_{t_i^-} - D(t_i, S_{t_i^-}) - K\right)^+\right] \tag{25}$$

with D_{st} the discount factor from t to s. Furthermore, we assume that dividends are part cash, part yield:

$$D(t,S) = \alpha(t)S_0 + \beta(t)S$$

In [7], we have explained how to build an implied volatility surface satisfying (25) by introducing a *continuous* local martingale f_t defined by $S_t = A(t)S_0 + B(t)f_t$. By following closely Proof 1, this LSV model is calibrated exactly to the market smile if and only if

$$\begin{aligned}\frac{1}{2}K^2\partial_K^2\mathcal{C}(T,K)\sigma(T,K)^2\mathbb{E}^{\mathbb{P}^T}[a_T^2|S_T=K] \\ = \partial_T\mathcal{C}(T,K) + r_T^0 K\partial_K\mathcal{C}(T,K) - \mathbb{E}^{\mathbb{P}}[D_{0T}\left(r_T - r_T^0\right)1_{S_T>K}]\end{aligned} \tag{26}$$

between two dividend dates and satisfies the following matching condition at a dividend date:

$$\mathcal{C}(t_i,K) = \mathbb{E}^{\mathbb{P}}\left[D_{0t_i}\left(S_{t_i^-} - D(t_i,S_{t_i^-}) - K\right)^+\right] \tag{27}$$

With a LV model with dividends and deterministic rate (i.e. $a_t \equiv 1$, $r_t \equiv r_t^0$), the necessary and sufficient condition is

$$\frac{1}{2}K^2\partial_K^2\mathcal{C}(T,K)\sigma_{\mathrm{Dup+Div}}(T,K)^2 = \partial_T\mathcal{C}(T,K) + r_T^0 K\partial_K\mathcal{C}(T,K) \tag{28}$$

with the same matching condition (27). Therefore by subtracting (28) from (26), we get that the LSV model is calibrated exactly to the market smile if and only if

$$\sigma(T,K)^2\mathbb{E}^{\mathbb{P}^T}[a_T^2|S_T=K] = \sigma_{\mathrm{Dup+Div}}(T,K)^2 - P_{0T}\frac{\mathbb{E}^{\mathbb{P}^T}[\left(r_T - r_T^0\right)1_{S_T>K}]}{\frac{1}{2}K\partial_K^2\mathcal{C}(T,K)}$$

between two dividend dates and satisfies (27) at dividend dates t_i, where we have used that the quantity $\mathbb{E}^{\mathbb{P}}[D_{0t_i}(S_{t_i^-} - D(t_i,S_{t_i^-}) - K)^+]$ is the same when computed in the LSV model or when computed in the LV model with deterministic rate, because both models are calibrated to the market smile. Finally, by following closely Proof 2, we obtain formula (9) with σ_{Dup} replaced by $\sigma_{\mathrm{Dup+Div}}$ and the extra processes (V_t,Θ_t) replaced by

$$\begin{aligned}V_t^{\mathrm{div}} &= V_t\prod_i\left(1+\alpha(t_i)\frac{S_0}{S_{t_i^-}}\right)1_{t_i\le t}\\ \Theta_t^{\mathrm{div}} &= \Theta_t\prod_i\frac{1}{1-\beta(t_i)}1_{t_i\le t}\end{aligned}$$

References

[1] Abergel, F., Tachet, R. : *A nonlinear partial integrodifferential equation from mathematical finance*, Discrete Cont. Dynamical Systems, Serie A, vol. 27, 3, pp 907-917, 2010

[2] Balland, P. : *Stoch-Vol Model with Interest Rate Volatility*, ICBI conference (2005).

[3] Benhamou, E., Gruz, A., Rivoira, A. : *Stochastic interest rates for local volatility hybrids models*, Wilmott Magazine, 2008.

[4] Bergomi, L. : *Smile dynamics III*, Risk Magazine, Oct. 2008.

[5] Dupire, B. : *A Unified Theory of Volatility*, In Derivatives Pricing: The Classic Collection, edited by Peter Carr, Risk publications.

[6] Henry-Labordère, P. : *Analysis, Geometry and Modeling in Finance: Advanced Methods in Option Pricing*, Chapman & Hall/CRC, Financial Mathematics Series (2008).

[7] Henry-Labordère, P. : *Calibration of Local Stochastic Volatility Models: A Monte-Carlo Approach*, Risk Magazine , Sept. (2009). Extended version: http://ssrn.com/abstract=1493306.

[8] Jourdain, B., Sbai, M. : *Coupling Index and Stocks*, Quantitative Finance (Oct. 2010).

[9] Langnau A : *A dynamic model for correlation risk*, Risk Magazine, April 2010, pages 74–78.

[10] Lipton, A. : *The Vol Smile Problem*, Risk Magazine, Feb. (2002).

[11] Madan, D., Qian, M.Q., Ren, Y. : *Calibrating and Pricing with Embedded Local Volatility Models*, Risk Magazine, Sept. 2007.
[12] Piterbarg, V. : *Smiling Hybrids*, Risk Magazine, May 2006.
[13] Sznitman, A.S. : *Topics in propagation of chaos*, Ecole d'été de probabilités de Saint-Flour XIX - 1989, volume 1464 of Lect. Notes in Math. Springer-Verlag (1991).
[14] Silverman, B.W. : *Density Estimation for Statistics and Data Analysis*, Chapman & Hall, New York, 1986.