

Implied and Local Correlations from Spread Options

René Carmona* and Youhong Sun †

Bendheim Center for Finance

Department of Operations Research and Financial Engineering,
Princeton University, Princeton, NJ 08544

August 18, 2011

Abstract

The multivariate lognormal model is a basic pricing model for derivatives with multiple underlying processes, for example, spread options. However, the market observation of implied correlation skew exemplifies how inaccurate the constant correlation assumption in the multivariate lognormal model can be. In this paper, we study alternative modeling approaches that generate implied correlation skews while at the same time maintain practical tractability. First, we propose a multiscale stochastic volatility model, and derive asymptotic formulas for option valuation and implied correlation. The model is a two-dimensional extension of the multiscale stochastic volatility model proposed in [20] which was tested on single underlying options. To obtain option valuations, we only need to calibrate a set of special parameters, and we propose a calibration procedure using option prices on individual underlying assets. From our simulated results, the multiscale stochastic volatility model generates implied correlation skews, and the asymptotic formulas are easy and fast to implement. However, in the multiscale stochastic volatility model, the stochastic volatilities introduce non-tradable sources of risk, and the market is no longer complete. Alternatively, we propose a local correlation model, which assumes the instantaneous correlation to be a deterministic function of time and the underlying prices. This model can be viewed as a two-dimensional extension of Dupire's local volatility model. The local correlation approach preserves the completeness of the market and low dimensionality of uncertainty.

Key words. spread option, implied correlation, local correlation

AMS subject classifications. 60H10, 91G20, 91G80

1 Introduction

Spread options are the simplest example of multi-underlying derivatives. It is a contract written on the difference between two underlying interests. If we denote the values of the two underlying interests at time t by X_t and Y_t , the payoff of a spread option at maturity T is $(X_T - M \times Y_T - K)^+$, where M and K are the ratio and strike of the option, and the function x^+ is defined as $x^+ = \max(x, 0)$. In other words, the buyer of the spread option has the right to buy the spread $X_T - M \times Y_T$ at the

*email: rcarmona@princeton.edu

†email: youhongs@princeton.edu

prespecified price K at maturity T . In a Markovian set-up, the price C at time t of the spread option with date of maturity T is given by the risk-neutral conditional expectation

$$C(x, y, t; T, M, K) = e^{-r(T-t)} \mathbf{E}^Q[(X_T - M \times Y_T - K)^+ | X_t = x, Y_t = y]. \quad (1)$$

More generally, if we denote the payoff at maturity by $h(X_T, Y_T)$, then the price of any European payoff multi-asset option is

$$C(x, y, t; T) = e^{-r(T-t)} \mathbf{E}^Q[h(X_T, Y_T) | X_t = x, Y_t = y]. \quad (2)$$

Spread options can be used for speculation, risk hedging, or even physical asset valuation. They are traded in many markets, such as the commodity markets, the currency and foreign exchange markets and the fixed-income markets, sometimes on exchanges, but most often as over-the-counter transactions. For example, the spark spread option is a spread between natural gas and electric power. The underlying is

$$S_t = P_t - H_{eff} G_t, \quad (3)$$

where P_t and G_t denote futures prices of power and gas, and H_{eff} is the heat rate, or efficiency ratio, of a typical gas fired power plant. It provides a proxy for the cost of converting natural gas into electricit. In this paper, the analysis will mainly focus on the spark spread option, but the methods can be easily used for other spreads.

The multivariate lognormal model, as a generalization of the Samuelson [36] lognormal model, is a basic model for spread option valuation. The model assumes that the risk-neutral dynamics of X_t and Y_t are given by geometric Brownian motions

$$\begin{cases} dX_t = X_t[r dt + \sigma_1 dW_X(t)] \\ dY_t = Y_t[r dt + \sigma_2 dW_Y(t)] \end{cases} \quad (4)$$

where W_X and W_Y are standard Brownian motions with correlation ρ .

The price of a spread option with strike $K = 0$, maturity T and payoff $(X_T - Y_T)^+$ under model (4) is given by Margrabe's formula [38]. For non-zero strikes, there is no closed form valuation formula under the multivariate lognormal model (4), and one needs to rely on approximations and numerical procedures. For example, Kirk suggested an approximation formula in [24] which provides a good approximation of spread option prices when the strike K is not far from zero. When the strike is large, Bjerksund and Stensland extended Kirk's idea and propose a more accurate, though more involved, closed form approximation for spread option price in [6]. Carmona and Durrleman [9] [10] developed sophisticated price bounds for spread option prices which produced accurate estimates to the true option value. Dempster and Hong [15] proposed an efficient numerical method using the fast Fourier transform [12].

The multivariate lognormal model (4) assumes constant correlation between the Brownian motions which drive the two underlying interests. However, past research results provide evidence that correlations change over time [8] [28], and there is a large correlation risk premium [16]. Return correlations increase when prices decrease, and often peak during financial crises. Roll [35] provides an analysis of the 1987 crash and Jorion [23] analyzes the Long Term Capital Management (LTCM) crisis, highlighting the role of correlation changes during these periods of crises.

In the spread option market, the observation of implied correlation skew exemplifies the inaccuracy of the constant correlation assumption. Implied correlations are defined similarly to implied volatilities, as the levels of correlation we need to input in a basic pricer in order to recover the market spread option prices. The basic pricing model is usually the multivariate lognormal model (4), and

the pricer can be chosen to be given by one of the approximation methods reviewed earlier, e.g. Kirk's formula [24]. Other parameters of the basic pricer, such as the volatilities of the two underlying assets, are fixed at pre-set levels which fit market conditions. The implied correlation is defined as a function of the ratio M and the strike K and as such, can be visualized as a two-dimensional surface. Practitioners observe implied correlation skews when both M and K change, especially M which for the spark spread options can be interpreted as the converting efficiency ratio of a power plant. Blakey and Scheule perform a nonparametric analysis of the implied correlation skew in [7]. The observation of an implied correlation skew indicates that tail distributions are not captured properly in the multivariate lognormal framework. Other models are needed in order to value spread options consistently with the implied correlation skew.

The fact that the multivariate lognormal model is inappropriate has been noted by many researchers. Instead of using a constant correlation, an alternative approach is to extract correlation information from historical data and market prices. When pricing new spreads or other multi-underlying options, the correlation in the multivariate lognormal model is modified according to the historical calibration and new prices are generated. For example, Cont and Deguest [13] suggested a statistical approach to extract an implied correlation matrix from index option prices. Cotter [14] proposed to calculate implied correlation from value at risk (VaR). Shevchenko [37] derived implied correlation formulas for foreign exchange rates. However, calibration from market prices can be extremely unstable in pricing and hedging. The idea that past observations provide a reasonable representation of the future is problematic.

Another approach is to rely on new models for the underlying assets. Ma [26] [27], Fonseca, Grasselli and Tebaldi [19] and others proposed stochastic correlation models for option pricing. The authors of [19] argue that their model generates implied correlation skew for the "best-of" basket option. Fengler, Herwatz, Menn and Werner [18] showed that their model explains the index volatility skew. Alexander and Scourse [4] proposed a bivariate normal mixture model for the underlying assets at maturity. Benth and Kettler [5] modeled the two underlying asset prices with bivariate non-Gaussian autoregressive processes, and assumed the innovations are correlated by a non-symmetric copula. However, these models either fail to explain the implied correlation skew for spread options, or do not assume evolutionary dynamics of the two underlying assets. Without dynamics, these models can not be used to price path-depend spread derivatives.

In this paper, we develop a two factor multiscale stochastic volatility model, and use asymptotic methods to approximate the spread option valuations and calculate the implied correlation. Our model is a plain generalization of a model introduced by Fouque, Papanicolaou, Sircar and Solna in [20], and tested on single underlying options. Our stochastic volatility model allows more flexibility than the multivariate lognormal model, and the asymptotic formulas we derive guarantee fast implementation and tractability. We show this model generates the desired implied correlation structure, and illustrate how to control the implied correlation structure by calibrating the parameters.

In practice, spread options are not as liquid as options on individual underlying assets. If we calibrate our model to spread option prices and market implied correlation, the result will be unstable and involve too much noise. We propose an alternative approach and discuss how to calibrate the model to the prices of options on the two underlying assets. As a result, the implied volatilities of the two underlying assets and the implied correlation of the spread options are "coupled". This approach guarantees that the valuation of spread options are consistent with market information for individual options.

Our two factor multiscale stochastic volatility model enriches the multivariate lognormal model to create an implied correlation skew. However, in the multiscale stochastic volatility model, the stochastic volatilities introduce non-tradable source of risk, and the market is no longer complete.

Motivated by Dupire’s local volatility approach [17], we propose in the second part of the paper, a local correlation theory for spread option modeling. We assume that the instantaneous correlation is a deterministic local correlation function of time and the underlying prices. The local correlation function is calibrated to derivatives and transformations of the spread option prices. The local correlation approach preserves the completeness of the market and low dimensionality of uncertainty.

We close the introduction with an outline of the paper. Section 2 introduces the two factor multiscale stochastic volatility model, and derives an asymptotic formula for spread option valuation. The formula provides approximations for the prices of European options with payoffs depending upon two underlying assets. A theoretical result is proved regarding the accuracy of the formula. A delta hedging experiment is performed, and comparisons with the multivariate lognormal model are provided. In Section 3, we derive an asymptotic formula for the implied correlation. We analyze the implied correlation surface generated by the multiscale stochastic volatility model and study its dependence upon and sensitivity with respect to the input parameters. Section 3 also discusses the details of the *coupling smiles* procedure, which is the calibration of implied correlations and spread option prices to the prices of options on the two underlying assets. Numerical tests are performed with market power and natural gas data. The use of the calibrated model is demonstrated on an example of the real option approach to power plant valuation.

Section 4 introduces the local correlation theory, and derives formulas of the local correlation function involving transformations of the spread option price surface. The calibrated local correlation and the underlying processes are used to price exotic and path-dependent derivatives. A numerical example is provided in which the calibrated model is used to price a capped variance swap on the spread.

2 Two Factor Multiscale Stochastic Volatility Model

In this section, we introduce a two factor multiscale stochastic volatility model for the two underlying processes, and derive an asymptotic pricing formula for spread options. We also provide theoretical and numerical evidence for the accuracy of the formula.

2.1 Model Setup

First, we present the various component of a two factor multiscale stochastic volatility model. We explain the rationale behind the model, and analyze how the dependence between the two underlying assets is incorporated in the model. We also discuss the change of measure from the physical measure to a risk neutral pricing measure. The results will be used to price European derivatives written on the two underlying assets.

2.1.1 Model under the Physical Measure

We assume that the dynamics of the two underlying prices X_t and Y_t are given by the stochastic differential equations

$$\begin{aligned} dX_t &= \mu_1 X_t dt + \sigma_t^X X_t dW_t^{(X)}, \\ dY_t &= \mu_2 Y_t dt + \sigma_t^Y Y_t dW_t^{(Y)}, \end{aligned} \tag{5}$$

where σ_t^X and σ_t^Y are adapted stochastic processes which will be described below. When σ_t^X and σ_t^Y are constants, the model reduces to the classic multivariate lognormal model reviewed in Section 1.

$W_t^{(X)}$ and $W_t^{(Y)}$ are standard Brownian motions, and their correlation is given by

$$d \langle W_t^{(X)}, W_t^{(Y)} \rangle = \rho dt, \quad (6)$$

where $|\rho| < 1$. As we discussed in Section 1, the two underlying assets for the spread options are expected to be highly correlated. Natural gas is one of the major fuels used for power generation, so the price of natural gas is bound to have a strong impact on the price of electric power. To model the correlation between the two underlying processes, we assume the stochastic volatilities σ_t^X and σ_t^Y are driven by common factors. Motivated by [20], we choose factors Z_t and V_t to capture fast and slow scale volatilities respectively.

Fast scale volatility factor

The first factor Z_t driving the volatility is assumed to be fast mean-reverting. This is motivated by empirical observations that energy prices have random spikes as a result of supply shortage or demand-supply imbalance. After a rapid increase, the price usually comes back to a regular level in a short period of time. This behavior motivates the introduction of the first volatility factor as a fast mean-reverting diffusion process. It is chosen to be a CIR process.

$$dZ_t = \frac{1}{\epsilon}(m - Z_t)dt + \frac{\nu\sqrt{2}}{\sqrt{\epsilon}}\sqrt{Z_t}dW_t^{(Z)}. \quad (7)$$

The mean reverting rate is $1/\epsilon$ where $\epsilon > 0$ is a small parameter which corresponds to the fast time scale of the process. As in the original work for Fouque, Papanicolaou and Sircar, we choose the parametrization of the dynamics of Z so that the invariant distribution is independent of ϵ . It is a Gamma distribution with shape parameter m/ν^2 and scale parameter ν^2 . $W_t^{(Z)}$ is a standard Brownian motion correlated with the Brownian motions $W_t^{(X)}$ and $W_t^{(Y)}$.

Slow scale volatility factor

For some assets, the fast mean-reverting process is not enough to explain some long term trends in the volatilities. For this purpose, we need a second slow-varying factor V_t whose dynamics are chosen as

$$dV_t = \delta c(V_t)dt + \sqrt{\delta}g(V_t)dW_t^{(V)}, \quad (8)$$

where δ is a small parameter. The motivation for the choice of such an equation is that V_t is a time change $t \rightarrow \delta t$ of a regular diffusion process

$$d\tilde{V}_t = c(\tilde{V}_t)dt + g(\tilde{V}_t)dW_t. \quad (9)$$

We assume that the functions c and g are smooth and at most linearly growing at infinity. $W_t^{(V)}$ is a standard Brownian motion correlated with the other Brownian motions in the model.

Volatility processes

We assume that the volatility processes σ_t^X and σ_t^Y are of the form:

$$\sigma_t^X = f(Z_t)f_1(V_t) \quad \text{and} \quad \sigma_t^Y = f(Z_t)f_2(V_t), \quad (10)$$

where the functions f , f_1 and f_2 are positive, bounded and bounded away from zero. f_1 and f_2 are assumed to be smooth and f to be integrable with respect to the invariant distribution of Z_t .

To summarize, our two factor multi-scale stochastic volatility model is defined by the following stochastic differential system for the underlying price processes X_t, Y_t and the volatility factors Z_t, V_t

$$\begin{cases} dX_t &= \mu_1 X_t dt + X_t f(Z_t) f_1(V_t) dW_t^{(X)}, \\ dY_t &= \mu_2 Y_t dt + Y_t f(Z_t) f_2(V_t) dW_t^{(Y)}, \\ dZ_t &= \frac{1}{\epsilon} (m - Z_t) dt + \frac{\nu \sqrt{2}}{\sqrt{\epsilon}} \sqrt{Z_t} dW_t^{(Z)}, \\ dV_t &= \delta c(V_t) dt + \sqrt{\delta} g(V_t) dW_t^{(V)}. \end{cases} \quad (11)$$

Moreover, the 4-dimensional Brownian motion $\mathbf{W}_t = (W_t^{(X)}, W_t^{(Y)}, W_t^{(Z)}, W_t^{(V)})$ can be written as a linear transformation of a standard 4-dimensional Brownian motion \mathbf{W}_t^0 with independent components:

$$\mathbf{W}_t = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \rho & \sqrt{1-\rho^2} & 0 & 0 \\ \rho_{11} & \widetilde{\rho}_{21} & \sqrt{1-\rho_{11}^2-\widetilde{\rho}_{21}^2} & 0 \\ \rho_{12} & \widetilde{\rho}_{22} & \widetilde{\rho}_0 & \sqrt{1-\rho_{12}^2-\widetilde{\rho}_{22}^2-\widetilde{\rho}_0^2} \end{pmatrix} \mathbf{W}_t^0. \quad (12)$$

The constants satisfy $|\rho| < 1, |\rho_{11}| < 1, |\rho_{12}| < 1$. Also, we assume $\rho_{11}^2 + \widetilde{\rho}_{21}^2 < 1$ and $\rho_{12}^2 + \widetilde{\rho}_{22}^2 + \widetilde{\rho}_0^2 < 1$. The correlation between $W^{(Y)}$ and $W^{(Z)}$ is given by $\rho_{21} = \rho \rho_{11} + \widetilde{\rho}_{21} \sqrt{1-\rho^2}$. The correlation between $W^{(Y)}$ and $W^{(V)}$ is given by $\rho_{22} = \rho \rho_{12} + \widetilde{\rho}_{22} \sqrt{1-\rho^2}$. The correlation between $W^{(V)}$ and $W^{(Z)}$ is given by $\rho_0 = \rho_{11} \rho_{12} + \widetilde{\rho}_{21} \widetilde{\rho}_{22} + \widetilde{\rho}_0 \sqrt{1-\rho_{11}^2-\widetilde{\rho}_{21}^2}$.

2.1.2 Model under a Risk-neutral Measure

In order to price derivatives, we need to change measure from the physical measure \mathbb{P} to a risk-neutral pricing measure \mathbb{Q} . Our market model is incomplete, and the choice of the risk-neutral pricing measure is fully determined by the market price of volatility risk, which is a function that reflects the risk preferences of agents in the market.

In the multiscale stochastic volatility model, we introduce a combined market price of volatility risk defined by $\Lambda(z, v)$ and $\Gamma(z, v)$ which we assume to be smooth and bounded functions of z and v . The corresponding risk-neutral measure \mathbb{Q} is defined by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} \Big|_{\mathcal{F}_t} = \exp \left(- \int_0^t \mathbf{a}_s d\mathbf{W}_t^0 - \frac{1}{2} \int_0^t \mathbf{a}'_s \mathbf{a}_s ds \right), \quad (13)$$

where

$$\begin{aligned}
 \mathbf{a}_t &= \begin{pmatrix} \frac{1}{\sqrt{1-\rho^2}} \left(\frac{\frac{\mu_1-r}{f(Z_t)f_1(V_t)}}{\frac{\mu_2-r}{f(Z_t)f_2(V_t)} - \rho \frac{\mu_1-r}{f(Z_t)f_1(V_t)}} \right) \\ \gamma(Z_t, V_t) \\ \xi(Z_t, V_t) \end{pmatrix}, \\
 \gamma(z, v) &= \frac{1}{\sqrt{1-\rho_{11}^2 - \tilde{\rho}_{21}^2}} \left\{ \Lambda(z, v) - \rho_{11} \frac{\mu_1-r}{f(z)f_1(v)} \right. \\
 &\quad \left. - \frac{\tilde{\rho}_{21}}{\sqrt{1-\rho^2}} \left[\frac{\mu_2-r}{f(z)f_2(v)} - \rho \frac{\mu_1-r}{f(z)f_1(v)} \right] \right\}, \\
 \xi(z, v) &= \frac{1}{\sqrt{1-\rho_{12}^2 - \tilde{\rho}_{22}^2 - \tilde{\rho}_0^2}} \left\{ \Gamma(z, v) - \rho_{12} \frac{\mu_1-r}{f(z)f_1(v)} \right. \\
 &\quad \left. - \frac{\tilde{\rho}_{22}}{\sqrt{1-\rho^2}} \left[\frac{\mu_2-r}{f(z)f_2(v)} - \rho \frac{\mu_1-r}{f(z)f_1(v)} \right] - \tilde{\rho}_0 \gamma(z, v) \right\}.
 \end{aligned} \tag{14}$$

Notice that \mathbf{a}_t is well-defined since we assume f , f_1 and f_2 to be positive, bounded and bounded away from zero. The functions $\gamma(z, v)$ and $\xi(z, v)$ in the change of measure are determined by the market price of volatility risk functions $\Lambda(z, v)$ and $\Gamma(z, v)$. By Girsanov theorem [29], the process \mathbf{W}_t^{0*} defined as

$$\mathbf{W}_t^{0*} = \mathbf{W}_t^0 + \int_0^t \mathbf{a}_s ds \tag{15}$$

is a standard 4-dimensional Brownian motion under the new measure \mathbb{Q} . The model under \mathbb{Q} can be written as

$$\begin{cases} dX_t &= rX_t dt + X_t f(Z_t) f_1(V_t) dW_t^{(X)*}, \\ dY_t &= rY_t dt + Y_t f(Z_t) f_2(V_t) dW_t^{(Y)*}, \\ dZ_t &= \left[\frac{1}{\epsilon} (m - Z_t) - \frac{\nu\sqrt{2}}{\sqrt{\epsilon}} \sqrt{Z_t} \Lambda(Z_t, V_t) \right] dt + \frac{\nu\sqrt{2}}{\sqrt{\epsilon}} \sqrt{Z_t} dW_t^{(Z)*}, \\ dV_t &= [\delta c(V_t) - \sqrt{\delta} g(V_t) \Gamma(Z_t, V_t)] dt + \sqrt{\delta} g(V_t) dW_t^{(V)*}, \end{cases} \tag{16}$$

where the relationship between \mathbf{W}_t^{0*} and $\mathbf{W}_t^* = (W_t^{(X)*}, W_t^{(Y)*}, W_t^{(Z)*}, W_t^{(V)*})$ is the same as in (12).

For any European option with maturity T and payoff function $h(x, y)$, the option price at time $t < T$ can be calculated by risk-neutral expectation

$$C^{\epsilon, \delta}(x, y, z, v, t) = e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}}[h(X_T, Y_T) | X_t = x, Y_t = y, Z_t = z, V_t = v]. \tag{17}$$

The option price depends on the values (x, y, z, v) of the four components of the model at time t , and also the present time t . It also depends on the parameters of the model $(\epsilon, \delta, \rho, \rho_{11}, \rho_{12}, \rho_{21}, \rho_{22}, \rho_0)$, the risk-free rate r and the functions $(f, f_1, f_2, c, g, \Lambda, \Gamma)$. Calibration of these functions is extremely complicated and unstable. We will show in our asymptotic analysis that the option price only depends upon a few special parameters. These special parameters are explicit functions of $(\epsilon, \delta, f, f_1, f_2, c, g, \Lambda, \Gamma)$, which can be calibrated to market data. Details are given in Section 3.

2.2 Asymptotic Analysis of Spread Option Prices

In this part, we derive spread option price asymptotics based on the scheme introduced in [20]. The final result is given in (68) and (71).

Applying the multi-dimensional Feynman-Kac formula, the spread option price $C^{\epsilon, \delta}$ solves the following partial differential equation

$$\begin{aligned}
& \frac{1}{2}x^2 f^2(z) f_1^2(v) C_{xx} + \frac{1}{2}y^2 f^2(z) f_2^2(v) C_{yy} + \frac{\nu^2}{\epsilon} z C_{zz} + \frac{1}{2} \delta g^2(v) C_{vv} \\
& + \rho xy f^2(z) f_1(v) f_2(v) C_{xy} + \rho_{11} \frac{\nu \sqrt{2}}{\sqrt{\epsilon}} x \sqrt{z} f(z) f_1(v) C_{xz} + \rho_{21} \frac{\nu \sqrt{2}}{\sqrt{\epsilon}} y \sqrt{z} f(z) f_2(v) C_{yz} \\
& + \rho_{12} x \sqrt{\delta} g(v) f(z) f_1(v) C_{xv} + \rho_{22} y \sqrt{\delta} g(v) f(z) f_2(v) C_{yv} + \rho_0 \sqrt{\frac{\delta}{\epsilon}} \nu \sqrt{2} \sqrt{z} g(v) C_{zv} \\
& + \left[\frac{1}{\epsilon} (m - z) - \frac{\nu \sqrt{2}}{\sqrt{\epsilon}} \sqrt{z} \Lambda(z, v) \right] C_z + [\delta c(v) - \sqrt{\delta} g(v) \Gamma(z, v)] C_v + r x C_x + r y C_y \\
& - r C + C_t = 0
\end{aligned} \tag{18}$$

with the terminal condition

$$C^{\epsilon, \delta}(x, y, z, v, T) = h(x, y). \tag{19}$$

For now, we only assume the payoff function $h(x, y)$ to be smooth and growing at most linearly so that the option price and its derivatives exist. We denote by $\mathcal{L}^{\epsilon, \delta} C^{\epsilon, \delta}$ the left hand side of the above equation. The operator $\mathcal{L}^{\epsilon, \delta}$ can be split into different powers of the small parameters (ϵ, δ) . The decomposition is

$$\mathcal{L}^{\epsilon, \delta} = \frac{1}{\epsilon} \mathcal{L}_0 + \frac{1}{\sqrt{\epsilon}} \mathcal{L}_1 + \mathcal{L}_2 + \sqrt{\delta} \mathcal{M}_1 + \delta \mathcal{M}_2 + \sqrt{\frac{\delta}{\epsilon}} \mathcal{M}_3, \tag{20}$$

where the operators appearing in the right hand side are defined by:

$$\begin{aligned}
\mathcal{L}_0 &= (m - z) \frac{\partial}{\partial z} + \nu^2 z \frac{\partial^2}{\partial z^2}, \\
\mathcal{L}_1 &= \nu \sqrt{2} [\rho_{11} x \sqrt{z} f(z) f_1(v) \frac{\partial^2}{\partial x \partial z} + \rho_{21} y \sqrt{z} f(z) f_2(v) \frac{\partial^2}{\partial y \partial z} - \sqrt{z} \Lambda(z, v) \frac{\partial}{\partial z}], \\
\mathcal{L}_2 &= \frac{\partial}{\partial t} + \frac{1}{2} x^2 f^2(z) f_1^2(v) \frac{\partial^2}{\partial x^2} + \frac{1}{2} y^2 f^2(z) f_2^2(v) \frac{\partial^2}{\partial y^2} \\
& + \rho xy f^2(z) f_1(v) f_2(v) \frac{\partial^2}{\partial x \partial y} + r(x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} - 1), \\
\mathcal{M}_1 &= -g(v) \Gamma(z, v) \frac{\partial}{\partial v} + \rho_{12} x g(v) f(z) f_1(v) \frac{\partial^2}{\partial x \partial v} + \rho_{22} y g(v) f(z) f_2(v) \frac{\partial^2}{\partial y \partial v}, \\
\mathcal{M}_2 &= c(v) \frac{\partial}{\partial v} + \frac{1}{2} g^2(v) \frac{\partial^2}{\partial v^2}, \\
\mathcal{M}_3 &= \nu \sqrt{2} \rho_0 \sqrt{z} g(v) \frac{\partial^2}{\partial z \partial v}.
\end{aligned} \tag{21}$$

In the following discussion, we consider an expansion of $C^{\epsilon, \delta}$ in powers of $\sqrt{\delta}$:

$$C^{\epsilon, \delta} = C_0^\epsilon + \sqrt{\delta} D_1^\epsilon + \delta D_2^\epsilon + \dots \tag{22}$$

For the purpose of identification, we plug in the expansion (22) into (18) and separate terms of different orders of $\sqrt{\delta}$. This gives the following equations for C_0^ϵ

$$\begin{aligned}
& \left(\frac{1}{\epsilon} \mathcal{L}_0 + \frac{1}{\sqrt{\epsilon}} \mathcal{L}_1 + \mathcal{L}_2 \right) C_0^\epsilon = 0, \\
& C_0^\epsilon(x, y, z, v, T) = h(x, y),
\end{aligned} \tag{23}$$

and for D_1^ϵ

$$\left(\frac{1}{\epsilon}\mathcal{L}_0 + \frac{1}{\sqrt{\epsilon}}\mathcal{L}_1 + \mathcal{L}_2\right) D_1^\epsilon = -\left(\mathcal{M}_1 + \sqrt{\frac{1}{\epsilon}}\mathcal{M}_3\right) C_0^\epsilon, \quad (24)$$

$$D_1^\epsilon(x, y, z, v, T) = 0.$$

C_0^ϵ and D_1^ϵ in the approximation (22) are the solution of (23) and (24). In the following discussion, we expand C_0^ϵ and D_1^ϵ in powers of $\sqrt{\epsilon}$, and calculate the coefficients of the different terms of the expansions.

2.2.1 Asymptotics for C_0^ϵ

To solve for the asymptotics of C_0^ϵ , we consider an expansion of C_0^ϵ in powers of $\sqrt{\epsilon}$.

$$C_0^\epsilon = C_0 + \sqrt{\epsilon}C_1 + \epsilon C_2 + \epsilon^{\frac{3}{2}}C_3 + \dots. \quad (25)$$

We plug (25) into (23) and separate the terms of different orders. The equations resulting from the identification of the first few leading terms are

$$\mathcal{L}_0 C_0 = 0, \quad (26)$$

$$\mathcal{L}_1 C_0 + \mathcal{L}_0 C_1 = 0, \quad (27)$$

$$\mathcal{L}_2 C_0 + \mathcal{L}_1 C_1 + \mathcal{L}_0 C_2 = 0, \quad (28)$$

$$\mathcal{L}_2 C_1 + \mathcal{L}_1 C_2 + \mathcal{L}_0 C_3 = 0, \quad (29)$$

with the terminal conditions

$$C_0(x, y, z, v, T) = h(x, y), \quad (30)$$

$$C_1(x, y, z, v, T) = 0. \quad (31)$$

Solution for C_0

Equations (26) is an ordinary differential equation in z . Every term in (27) has derivative with respect to z . Since the terminal conditions for C_0 and C_1 are independent of z , we derive that the solutions of (26) and (27) are constant in z , that is $C_0 = C_0(x, y, v, t)$ and $C_1 = C_1(x, y, v, t)$. As a result, we have $\mathcal{L}_1 C_1 = 0$ and (28) becomes

$$\mathcal{L}_2 C_0 + \mathcal{L}_0 C_2 = 0. \quad (32)$$

We define the bracket notation here to stand for integration with respect to the invariant distribution of the CIR process with infinitesimal generator \mathcal{L}_0 . In other words, it is the integration with respect to the Gamma distribution $G(m/\nu^2, \nu)$ with shape parameter m/ν^2 and scale parameter ν^2 . For any function integrable with respect to this distribution, the bracket notation is defined as

$$\langle g(z) \rangle = \int g(z)G(dz). \quad (33)$$

We treat (32) as a Poisson equation for \mathcal{L}_0 so that $\mathcal{L}_0 C_2 = -\mathcal{L}_2 C_0$. In order to have a solution, $\mathcal{L}_2 C_0$ must be in the kernel of the operator \mathcal{L}_0 , or equivalently, $\langle \mathcal{L}_2 C_0 \rangle = 0$. Recall that C_0 is independent

of z , and \mathcal{L}_2 does not contain partial derivatives in z , we can write $\langle \mathcal{L}_2 C_0 \rangle$ as $\langle \mathcal{L}_2 \rangle C_0$, where the operator $\langle \mathcal{L}_2 \rangle$ is defined as

$$\begin{aligned} \langle \mathcal{L}_2 \rangle = & \frac{\partial}{\partial t} + \frac{1}{2} f_1^2(v) x^2 \langle f^2(z) \rangle \frac{\partial^2}{\partial^2 x} + \frac{1}{2} f_2^2(v) y^2 \langle f^2(z) \rangle \frac{\partial^2}{\partial^2 y} \\ & + \rho f_1(v) f_2(v) x y \langle f^2(z) \rangle \frac{\partial^2}{\partial x \partial y} + r \left(x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} - 1 \right). \end{aligned} \quad (34)$$

This is the operator for the multivariate lognormal model (4) with the volatilities $\sqrt{\langle f^2(z) \rangle} f_1(v)$ and $\sqrt{\langle f^2(z) \rangle} f_2(v)$, and correlation ρ .

In the following discussion, we define in the same way the "bracket operation" of operators which do not have partial differentials with respect to z .

From the discussion above, C_0 must solve the following partial differential equation:

$$\begin{aligned} \langle \mathcal{L}_2 \rangle C_0 &= 0, \\ C_0(x, y, v, T) &= h(x, y). \end{aligned} \quad (35)$$

In other word, C_0 is the price function of options with a multivariate lognormal underlying model (4) where the volatilities of the two underlying processes are depending on v , namely $\bar{\sigma}_1^2(v) = \langle f^2(z) \rangle f_1^2(v)$, $\bar{\sigma}_2^2(v) = \langle f^2(z) \rangle f_2^2(v)$, and the correlation is ρ . Again, the bracket notation stands for integration with respect to the invariant distribution of the CIR process with infinitesimal generator \mathcal{L}_0 . Notice that the option price depends on the initial value of v in this model.

$$C_0 = C_0(x, y, t; \bar{\sigma}_1(v), \bar{\sigma}_2(v), \rho). \quad (36)$$

Solution for C_1

We rewrite equation (29) as a Poisson equation for the operator \mathcal{L}_0 :

$$\mathcal{L}_0 C_3 = -\mathcal{L}_1 C_2 - \mathcal{L}_2 C_1. \quad (37)$$

Take the bracket operation on both sides of (37), and recall that C_1 is independent of z . We have

$$\langle \mathcal{L}_2 \rangle C_1 = -\langle \mathcal{L}_1 C_2 \rangle. \quad (38)$$

To solve for C_1 , we first need to solve for C_2 . From equation (32) with the fact $\langle \mathcal{L}_2 C_0 \rangle = 0$, we have

$$\mathcal{L}_0 C_2 = -\mathcal{L}_2 C_0 = -(\mathcal{L}_2 C_0 - \langle \mathcal{L}_2 C_0 \rangle). \quad (39)$$

Because \mathcal{L}_0 is invertible on the space orthogonal to the constant functions, and C_0 is independent of z , we have

$$C_2 = -\mathcal{L}_0^{-1}(\mathcal{L}_2 - \langle \mathcal{L}_2 \rangle) C_0. \quad (40)$$

Therefore, for C_1 , we choose

$$\begin{aligned} \langle \mathcal{L}_2 \rangle C_1 &= -\langle \mathcal{L}_1 C_2 \rangle \\ &= \langle \mathcal{L}_1 \mathcal{L}_0^{-1}(\mathcal{L}_2 - \langle \mathcal{L}_2 \rangle) \rangle C_0 = \mathcal{A} C_0. \end{aligned} \quad (41)$$

To calculate \mathcal{A} , let us assume that $\phi(z)$ solves the ordinary differential equation

$$\mathcal{L}_0\phi(z) = f^2(z) - \langle f^2(z) \rangle \quad (42)$$

with $\langle \phi(z) \rangle = 0$, so that $\mathcal{L}_0^{-1}(f^2(z) - \langle f^2(z) \rangle) = \phi(z)$. Then we have:

$$\mathcal{L}_0^{-1}(\mathcal{L}_2 - \langle \mathcal{L}_2 \rangle) = \phi(z) \left[\frac{1}{2}x^2 f_1^2(v) \frac{\partial^2}{\partial x^2} + \frac{1}{2}y^2 f_2^2(v) \frac{\partial^2}{\partial y^2} + \rho xy f_1(v) f_2(v) \frac{\partial^2}{\partial x \partial y} \right], \quad (43)$$

and consequently

$$\begin{aligned} \mathcal{A} = \nu\sqrt{2} \{ & \langle \sqrt{z} f(z) \phi'(z) \rangle (\rho_{11} f_1(v) x \frac{\partial}{\partial x} + \rho_{21} f_2(v) y \frac{\partial}{\partial y}) - \langle \sqrt{z} \Lambda(z, v) \phi'(z) \rangle \} \\ & \times \left[\frac{1}{2} f_1^2(v) x^2 \frac{\partial^2}{\partial x^2} + \frac{1}{2} f_2^2(v) y^2 \frac{\partial^2}{\partial y^2} + \rho f_1(v) f_2(v) xy \frac{\partial^2}{\partial x \partial y} \right]. \end{aligned} \quad (44)$$

From the discussion above, C_1 solves the equation

$$\langle \mathcal{L}_2 \rangle C_1 = \mathcal{A} C_0 \quad (45)$$

with the terminal condition

$$C_1(x, y, v, T) = 0. \quad (46)$$

Next, we show that the solution to (45) is given by

$$C_1 = -(T - t) \mathcal{A} C_0. \quad (47)$$

Clearly such a C_1 satisfies the terminal condition (46). To prove the fact that (47) solves (45), we use the property:

$$(x^k \frac{d^k}{dx^k})(x^l \frac{d^l}{dx^l}) = (x^l \frac{d^l}{dx^l})(x^k \frac{d^k}{dx^k}). \quad (48)$$

which holds for all integer k and l and can be proved by a simple induction argument after changing variable to $\xi = \log x$. Using (48), we see that the operators $\langle \mathcal{L}_2 \rangle$ and \mathcal{A} commute. Moreover

$$\langle \mathcal{L}_2 \rangle [-(T - t) \mathcal{A} C_0] = \mathcal{A} C_0 - (T - t) \mathcal{A} \langle \mathcal{L}_2 \rangle C_0 = \mathcal{A} C_0, \quad (49)$$

proving that (47) is the solution for C_1 .

Finally, the approximation of C_0^ϵ is given by:

$$C_0^\epsilon \approx C_0 + \sqrt{\epsilon} C_1 = C_0 - \sqrt{\epsilon} (T - t) \mathcal{A} C_0. \quad (50)$$

2.2.2 Asymptotics for D_1^ϵ

Similarly, we expand D_1^ϵ in powers of $\sqrt{\epsilon}$:

$$D_1^\epsilon = D_1 + \sqrt{\epsilon} D_2 + \epsilon D_3 + \epsilon^{\frac{3}{2}} D_4 + \dots \quad (51)$$

Inserting the expansion (51) into equation (24) and separating the terms of different orders, we get

$$\mathcal{L}_0 D_1 = 0, \quad (52)$$

$$\mathcal{L}_0 D_2 + \mathcal{L}_1 D_1 = -\mathcal{M}_3 C_0, \quad (53)$$

$$\mathcal{L}_0 D_3 + \mathcal{L}_1 D_2 + \mathcal{L}_2 D_1 = -\mathcal{M}_1 C_0 - \mathcal{M}_3 C_1, \quad (54)$$

$$\mathcal{L}_0 D_4 + \mathcal{L}_1 D_3 + \mathcal{L}_2 D_2 = -\mathcal{M}_1 C_1 - \mathcal{M}_3 C_2, \quad (55)$$

with the terminal conditions

$$D_1(x, y, z, v, T) = 0, \quad (56)$$

$$D_2(x, y, z, v, T) = 0. \quad (57)$$

As we argued for C_0 and C_1 , D_1 is independent of z from (52) and the zero terminal condition (56). In (53), since $\mathcal{M}_3 C_0 = 0$ and $\mathcal{L}_1 D_1 = 0$, we derive

$$\mathcal{L}_0 D_2 = 0. \quad (58)$$

As a result from the above equation and the zero terminal condition (57), D_2 is independent of z . In (54), we have $\mathcal{M}_3 C_1 = 0$ and $\mathcal{L}_1 D_2 = 0$, and (54) becomes

$$\mathcal{L}_0 D_3 + \mathcal{L}_2 D_1 = -\mathcal{M}_1 C_0. \quad (59)$$

Solution for D_1

We rewrite (59) as a Poisson equation of \mathcal{L}_0 and take the bracket operation on both sides. Since D_1 is independent of z , we derive the partial differential equation for D_1

$$\begin{aligned} \langle \mathcal{L}_2 \rangle D_1 &= -\langle \mathcal{M}_1 \rangle C_0, \\ D_1(x, y, v, T) &= 0. \end{aligned} \quad (60)$$

We claim that the solution D_1 of (60) is

$$D_1 = \frac{T-t}{2} \langle \mathcal{M}_1 \rangle C_0, \quad (61)$$

and we derive the proof as followed. Denote

$$\begin{aligned} \langle \mathcal{M}_1 \rangle &= \left[-g(v) \langle \Gamma(z, v) \rangle + \rho_{12} g(v) f_1(v) \langle f(z) \rangle x \frac{\partial}{\partial x} \right. \\ &\quad \left. + \rho_{22} g(v) f_2(v) \langle f(z) \rangle y \frac{\partial}{\partial y} \right] \frac{\partial}{\partial v} \\ &= M_1 \frac{\partial}{\partial v}, \end{aligned} \quad (62)$$

where

$$M_1 = -g(v) \langle \Gamma(z, v) \rangle + \rho_{12} g(v) f_1(v) \langle f(z) \rangle x \frac{\partial}{\partial x} + \rho_{22} g(v) f_2(v) \langle f(z) \rangle y \frac{\partial}{\partial y}. \quad (63)$$

Plug in (61) into (60) and we have

$$\langle \mathcal{L}_2 \rangle \frac{T-t}{2} \langle \mathcal{M}_1 \rangle C_0 = \langle \mathcal{L}_2 \rangle \left[\frac{T-t}{2} (M_1 \frac{\partial}{\partial v}) C_0 \right]. \quad (64)$$

As proved by Reiss and Wystup [33], for any European claims in the multivariate lognormal model given by (4) with $X_0 = x$ and $Y_0 = y$, gamma and vega follow the relationship

$$\begin{aligned} \sigma_1 \frac{\partial C}{\partial \sigma_1} &= \sigma_1^2 x^2 (T-t) \frac{\partial^2 C}{\partial x^2} + \rho \sigma_1 \sigma_2 x y (T-t) \frac{\partial^2 C}{\partial x \partial y}, \\ \sigma_2 \frac{\partial C}{\partial \sigma_2} &= \sigma_2^2 y^2 (T-t) \frac{\partial^2 C}{\partial y^2} + \rho \sigma_1 \sigma_2 x y (T-t) \frac{\partial^2 C}{\partial x \partial y}. \end{aligned} \quad (65)$$

Using formula (65) we get

$$\begin{aligned} \frac{\partial}{\partial v} C_0 &= \frac{\partial C_0}{\partial \sigma_1} \bar{\sigma}'_1(v) + \frac{\partial C_0}{\partial \sigma_2} \bar{\sigma}'_2(v) \\ &= (T-t) \left\{ \bar{\sigma}_1(v) \bar{\sigma}'_1(v) x^2 \frac{\partial^2 C_0}{\partial x^2} + \bar{\sigma}_2(v) \bar{\sigma}'_2(v) y^2 \frac{\partial^2 C_0}{\partial y^2} \right. \\ &\quad \left. + [\bar{\sigma}_1(v) \bar{\sigma}'_2(v) + \bar{\sigma}_2(v) \bar{\sigma}'_1(v)] xy \frac{\partial^2 C_0}{\partial x \partial y} \right\}. \end{aligned}$$

As a result, (64) equals

$$\begin{aligned} \langle \mathcal{L}_2 \rangle \frac{T-t}{2} \langle \mathcal{M}_1 \rangle C_0 &= \langle \mathcal{L}_2 \rangle \left\{ \frac{(T-t)^2}{2} M_1 [\bar{\sigma}_1(v) \bar{\sigma}'_1(v) x^2 \frac{\partial^2 C_0}{\partial x^2} + \bar{\sigma}_2(v) \bar{\sigma}'_2(v) y^2 \frac{\partial^2 C_0}{\partial y^2} \right. \\ &\quad \left. + (\bar{\sigma}_1(v) \bar{\sigma}'_2(v) + \bar{\sigma}_2(v) \bar{\sigma}'_1(v)) xy \frac{\partial^2 C_0}{\partial x \partial y} \right\} \\ &= - (T-t) M_1 [\bar{\sigma}_1(v) \bar{\sigma}'_1(v) x^2 \frac{\partial^2 C_0}{\partial x^2} + \bar{\sigma}_2(v) \bar{\sigma}'_2(v) y^2 \frac{\partial^2 C_0}{\partial y^2} \\ &\quad + (\bar{\sigma}_1(v) \bar{\sigma}'_2(v) + \bar{\sigma}_2(v) \bar{\sigma}'_1(v)) xy \frac{\partial^2 C_0}{\partial x \partial y}] \\ &\quad + \frac{(T-t)^2}{2} \langle \mathcal{L}_2 \rangle M_1 [\bar{\sigma}_1(v) \bar{\sigma}'_1(v) x^2 \frac{\partial^2 C_0}{\partial x^2} + \bar{\sigma}_2(v) \bar{\sigma}'_2(v) y^2 \frac{\partial^2 C_0}{\partial y^2} \\ &\quad + (\bar{\sigma}_1(v) \bar{\sigma}'_2(v) + \bar{\sigma}_2(v) \bar{\sigma}'_1(v)) xy \frac{\partial^2 C_0}{\partial x \partial y}]. \end{aligned} \tag{66}$$

Using again (48), we see that the operator $\langle \mathcal{L}_2 \rangle$ commutes with $x^k \partial^k / \partial x^k$. Also, since $\langle \mathcal{L}_2 \rangle C_0 = 0$, the second term after the last equal sign of (66) vanishes

$$\begin{aligned} \langle \mathcal{L}_2 \rangle \frac{T-t}{2} \langle \mathcal{M}_1 \rangle C_0 &= - (T-t) M_1 [\bar{\sigma}_1(v) \bar{\sigma}'_1(v) x^2 \frac{\partial^2 C_0}{\partial x^2} + \bar{\sigma}_2(v) \bar{\sigma}'_2(v) y^2 \frac{\partial^2 C_0}{\partial y^2} \\ &\quad + (\bar{\sigma}_1(v) \bar{\sigma}'_2(v) + \bar{\sigma}_2(v) \bar{\sigma}'_1(v)) xy \frac{\partial^2 C_0}{\partial x \partial y}] \\ &= - \langle \mathcal{M}_1 \rangle C_0. \end{aligned} \tag{67}$$

This proves that (61) is a solution for D_1 .

2.2.3 Option Price Approximation Formula

With the calculations above, the spread option price is approximated by:

$$\begin{aligned} C^{\epsilon, \delta} &\approx C_0 + \sqrt{\epsilon} C_1 + \sqrt{\delta} D_1 \\ &= C_0 + (T-t) [-\sqrt{\epsilon} \mathcal{A} + \frac{\sqrt{\delta}}{2} \langle \mathcal{M}_1 \rangle] C_0, \end{aligned} \tag{68}$$

where \mathcal{A} is defined by (44). To analyze the properties of this price approximation and its relationship to individual option prices, we introduce the following parameters. The first and second groups are

related to the volatility functions of X_t and Y_t :

$$\begin{aligned}
\bar{\sigma}_{1,2}(v) &= \sqrt{\langle f^2(z) \rangle f_{1,2}^2(v)} \\
P_{1,2}^{(0)}(v) &= \frac{\sqrt{\delta}}{2} g(v) \langle \Gamma(z, v) \rangle \bar{\sigma}_{1,2}(v) \bar{\sigma}'_{1,2}(v), \\
P_{1,2}^{(1)}(v) &= -\frac{\sqrt{\delta}}{2} \rho_{12,22} g(v) \langle f(z) \rangle f_{1,2}(v) \bar{\sigma}_{1,2}(v) \bar{\sigma}'_{1,2}(v), \\
P_{1,2}^{(2)}(v) &= -\frac{\sqrt{\epsilon}}{\sqrt{2}} \nu \langle \sqrt{z} \Lambda(z, v) \phi'(z) \rangle f_{1,2}^2(v), \\
P_{1,2}^{(3)}(v) &= \frac{\sqrt{\epsilon}}{\sqrt{2}} \nu \rho_{11,21} \langle \sqrt{z} f(z) \phi'(z) \rangle f_{1,2}^3(v).
\end{aligned} \tag{69}$$

The third group comes from the cross-terms:

$$\begin{aligned}
P^{(0)}(v) &= -\frac{\sqrt{\delta}}{2} \rho_{12} g(v) \langle f(z) \rangle f_1(v) \bar{\sigma}_2(v) \bar{\sigma}'_2(v), \\
P^{(1)}(v) &= -\frac{\sqrt{\delta}}{2} \rho_{22} g(v) \langle f(z) \rangle f_2(v) \bar{\sigma}_1(v) \bar{\sigma}'_1(v), \\
P^{(2)}(v) &= -\sqrt{2} \sqrt{\epsilon} \nu \rho \langle \sqrt{z} \Lambda(z, v) \phi'(z) \rangle f_1(v) f_2(v), \\
P^{(3)}(v) &= \frac{\sqrt{\epsilon}}{\sqrt{2}} \nu \rho_{11} \langle \sqrt{z} f(z) \phi'(z) \rangle f_1(v) f_2^2(v), \\
P^{(4)}(v) &= \frac{\sqrt{\epsilon}}{\sqrt{2}} \nu \rho_{21} \langle \sqrt{z} f(z) \phi'(z) \rangle f_2(v) f_1^2(v), \\
P^{(5)}(v) &= \sqrt{2} \sqrt{\epsilon} \nu \rho \rho_{11} \langle \sqrt{z} f(z) \phi'(z) \rangle f_2(v) f_1^2(v), \\
P^{(6)}(v) &= \sqrt{2} \sqrt{\epsilon} \nu \rho \rho_{21} \langle \sqrt{z} f(z) \phi'(z) \rangle f_1(v) f_2^2(v).
\end{aligned} \tag{70}$$

Then the option price approximation can be rewritten as

$$\begin{aligned}
C^{\epsilon, \delta} &= C_0 - (T - t) \left\{ \frac{1}{\bar{\sigma}_1(v)} \left[P_1^{(0)} \frac{\partial}{\partial \sigma_1} + P_1^{(1)} x \frac{\partial^2}{\partial x \partial \sigma_1} + P_1^{(1)} y \frac{\partial^2}{\partial y \partial \sigma_1} \right] \right. \\
&+ \frac{1}{\bar{\sigma}_2(v)} \left[P_2^{(0)} \frac{\partial}{\partial \sigma_2} + P_2^{(1)} y \frac{\partial^2}{\partial y \partial \sigma_2} + P_2^{(0)} x \frac{\partial^2}{\partial x \partial \sigma_2} \right] \\
&+ \left[P_1^{(2)} x^2 \frac{\partial^2}{\partial x^2} + P_1^{(3)} x \frac{\partial}{\partial x} x^2 \frac{\partial^2}{\partial x^2} \right. \\
&+ P_2^{(2)} y^2 \frac{\partial^2}{\partial y^2} + P_2^{(3)} y \frac{\partial}{\partial y} y^2 \frac{\partial^2}{\partial y^2} \\
&+ P^{(3)} x \frac{\partial}{\partial x} y^2 \frac{\partial^2}{\partial y^2} + P^{(4)} y \frac{\partial}{\partial y} x^2 \frac{\partial^2}{\partial x^2} \\
&\left. + P^{(2)} xy \frac{\partial^2}{\partial x \partial y} + P^{(5)} x \frac{\partial}{\partial x} xy \frac{\partial^2}{\partial x \partial y} + P^{(6)} y \frac{\partial}{\partial y} xy \frac{\partial^2}{\partial x \partial y} \right] \Big\} C_0.
\end{aligned} \tag{71}$$

As we will show later, the parameters in (69) are calibrated to the marginal option prices. The parameters in (70) can be calculated from (69) and the parameter ρ . They satisfy the equations

$$\begin{aligned} P^{(0)} &= P_1^{(1)} \frac{P_2^{(0)}}{P_1^{(0)}}, & P^{(1)} &= P_2^{(1)} \frac{P_1^{(0)}}{P_2^{(0)}}, & P^{(2)} &= 2\rho\sqrt{P_1^{(2)}P_2^{(2)}}, \\ P^{(3)} &= P_1^{(3)} \left[\frac{\bar{\sigma}_2}{\bar{\sigma}_1}\right]^2, & P^{(4)} &= P_2^{(3)} \left[\frac{\bar{\sigma}_1}{\bar{\sigma}_2}\right]^2, & P^{(5)} &= 2\rho P_1^{(3)} \left[\frac{\bar{\sigma}_2}{\bar{\sigma}_1}\right], & P^{(6)} &= 2\rho P_2^{(3)} \left[\frac{\bar{\sigma}_1}{\bar{\sigma}_2}\right]. \end{aligned} \quad (72)$$

As a result, the asymptotic formula for $C^{\epsilon,\delta}$ depends only on the parameters related to the volatility functions of X_t and Y_t : $(\bar{\sigma}_1, P_1^{(0)}, P_1^{(1)}, P_1^{(2)}, P_1^{(3)})$, $(\bar{\sigma}_2, P_2^{(0)}, P_2^{(1)}, P_2^{(2)}, P_2^{(3)})$ and the Brownian motion correlation ρ . When we calibrate the multiscale stochastic volatility model, instead of calibrating the functions $(f, f_1, f_2, c, g, \Lambda, \Gamma)$ and the model parameters $(\epsilon, \delta, \rho, \rho_{11}, \rho_{12}, \rho_{21}, \rho_{22}, \rho_0)$, we only need to determine the seven $P^{(j)}$ parameters and ρ . This greatly reduces the complexity of the calibration and increases the efficiency of the model.

2.3 Accuracy of the Option Price Approximation

So far we have derived an asymptotic pricing formula for European options. Next, we study the accuracy of this approximation in the case of smooth and bounded option payoffs.

Theorem 2.1. *We assume that the following properties hold:*

1. *there exists a unique solution (X, Y, Z, V) for the system of stochastic differential equations (11) under the physical measure and fixed ϵ, δ ;*
2. *there exists a unique solution (X, Y, Z, V) for the system of stochastic differential equations (16) under the risk-neutral measure and fixed ϵ, δ ;*
3. *the market prices of volatility risk Λ and Γ are bounded;*
4. *the parameters satisfy $m > \nu^2$. ;*
5. *the functions $c(v)$ and $g(v)$ are smooth and at most linearly growing at infinity;*
6. *$\phi(z)$, $\varphi(z)$, $\psi(z, v)$, $\lambda(z)$ and $\zeta(z, v)$ as unique solutions of the following ordinary differential equations are at most polynomially growing in (z, v)*

$$\begin{aligned} \mathcal{L}_0\phi(z) &= f^2(z) - \langle f^2(z) \rangle, \\ \langle \phi(z) \rangle &= 0, \end{aligned} \quad (73)$$

$$\begin{aligned} \mathcal{L}_0\varphi(z) &= f(z) - \langle f(z) \rangle, \\ \langle \varphi(z) \rangle &= 0, \end{aligned} \quad (74)$$

$$\begin{aligned} \mathcal{L}_0\psi(z, v) &= \Gamma(z, v) - \langle \Gamma(z, v) \rangle, \\ \langle \psi(z, v) \rangle &= 0, \end{aligned} \quad (75)$$

$$\begin{aligned} \mathcal{L}_0\lambda(z) &= \sqrt{z}f(z)\phi'(z) - \langle \sqrt{z}f(z)\phi'(z) \rangle, \\ \langle \lambda(z) \rangle &= 0, \end{aligned} \quad (76)$$

$$\begin{aligned} \mathcal{L}_0\zeta(z, v) &= \sqrt{z}\Lambda(z, v)\phi'(z) - \langle \sqrt{z}\Lambda(z, v)\phi'(z) \rangle, \\ \langle \zeta(z, v) \rangle &= 0; \end{aligned} \quad (77)$$

7. the volatility functions f , f_1 and f_2 are positive, bounded and bounded away from zero. The payoff $h(x, y)$ and its derivatives are smooth and bounded such that the option price C_0 under the multivariate lognormal model

$$C_0 = C_0(x, y, v, t; \bar{\sigma}_1(v), \bar{\sigma}_2(v), \rho), \quad (78)$$

where

$$\bar{\sigma}_1^2(v) = \langle f^2(z) \rangle f_1^2(v), \quad \bar{\sigma}_2^2(v) = \langle f^2(z) \rangle f_2^2(v), \quad (79)$$

and all the derivatives $x^k y^j \frac{\partial^{j+k+t}}{\partial x^k \partial y^j \partial v^t} C_0$ for different t are uniformly bounded in (x, y, v) ;

8. $\epsilon \leq 1$ and $\delta \leq 1$.

Then for fixed (x, y, z, v, t) , there exists a constant $c > 0$ such that

$$|C^{\epsilon, \delta} - \tilde{C}^{\epsilon, \delta}| \leq c(\epsilon + \delta + \sqrt{\epsilon\delta}). \quad (80)$$

The proof of Theorem 2.1 follows the same line as the corresponding result proven in [20], so we shall omit the details here. Theorem 2.1 quantifies the accuracy of the price approximation when the payoff function $h(x, y)$ is bounded and smooth. For the piecewise smooth payoffs (call option on spreads for example), we will illustrate the accuracy of the pricing formula with a delta hedging experiment in the next section.

2.4 Delta Hedging Experiment

Delta hedging is a procedure that is widely used by derivative investors to reduce a portfolio's exposure to underlying price movements. In option trading, the deltas are the sensitivity of the value of an option to changes in the values of the underlying assets. The investors calculate the portfolio's deltas and then add an offsetting position in the underlying assets to make the portfolio's delta zero, or delta neutral. Delta hedging experiments are efficient ways to test the accuracy of option pricing formulas. The idea is to simulate sample paths of the underlying prices, and calculate the delta-hedged portfolios along each path. The option price and hedging ratio is calculated with the option pricing formula. If the formula is accurate, the profit and loss from trading the delta-hedged portfolio along each path should be small and centered around zero.

In the first simulation, we assume that the historical dynamics of the underlying asset prices are given by the following specific form of our multiscale stochastic volatility model:

$$\begin{aligned} dX_t &= \mu_1 X_t dt + X_t \sqrt{Z_t V_t} dW_t^{(X)*}, \\ dY_t &= \mu_2 Y_t dt + Y_t \sqrt{Z_t V_t} dW_t^{(Y)*}, \\ dZ_t &= \frac{1}{\epsilon} (m - Z_t) dt + \frac{\nu \sqrt{2}}{\sqrt{\epsilon}} \sqrt{Z_t} dW_t^{(Z)}, \\ dV_t &= \sqrt{\delta} V_t dW_t^{(V)*}. \end{aligned} \quad (81)$$

The slow scale stochastic volatility V_t is chosen to be a geometric Brownian motion. This choice is motivated by the popularity of the SABR model [30]. The volatility function is the square root function, which is the same as in the Heston model [22]. Correlations between Brownian motions are defined in (12).

Numerical values of the parameters are chosen as $x = 50$, $y = 5$, $\mu_1 = 0.1$, $\mu_2 = 0.1$, $z = 0.5$, $v = 0.5$, $\epsilon = 1$, $\delta = 1$, $m = 0.5$, $\nu = 0.1$, $\rho = 0.5$, $\rho_{11} = 0.3$, $\rho_{12} = 0.3$, $\rho_{21} = 0.2$, $\rho_{22} = 0.2$, $\rho_0 = 0.1$, $M = 10$, $K = 0$, $T = 1$, $r = 0$.

We use the model above to generate 2000 sample paths of the process X_t and Y_t of the underlying assets. For each path, we sell one spread option. The portfolio is delta-hedged with both the multiscale stochastic volatility model and the multivariate lognormal model. The pricing formula for the multiscale stochastic volatility model is given in (68). The multivariate lognormal model is priced by Kirk’s formula [24]. The final profit and loss (P&L for short) of each path is collected for the two models.

Figure 1 gives the histograms of the P&L for the two pricing models. When hedged by the multiscale stochastic volatility model, the distribution of the P&L is slim and centered around zero compared with the multivariate lognormal model.

To quantify the difference between the two models, we calculate the the P&L ratio of the two portfolios defined as the ratio of the payoff over the option price. The mean and variance of the the P&L ratios are listed in Table 1. The the P&L ratio is on average 2.4% when hedged by the multiscale stochastic volatility model, compared with 18.3% for the multivariate lognormal model. The statistics prove again the accuracy of the formula (68). However, the risk is not completely eliminated. This is because under a stochastic volatility model, a delta-hedged portfolio is still exposed to volatility risk and other types of risk.

Table 1: Profit and loss ratio statistics for the simulation experiment with the true dynamics of the multiscale stochastic volatility model

	Mean	Variance
Multiscale Stochastic Volatility Model	0.0240	0.1416
Multivariate Lognormal model	0.1834	0.1431

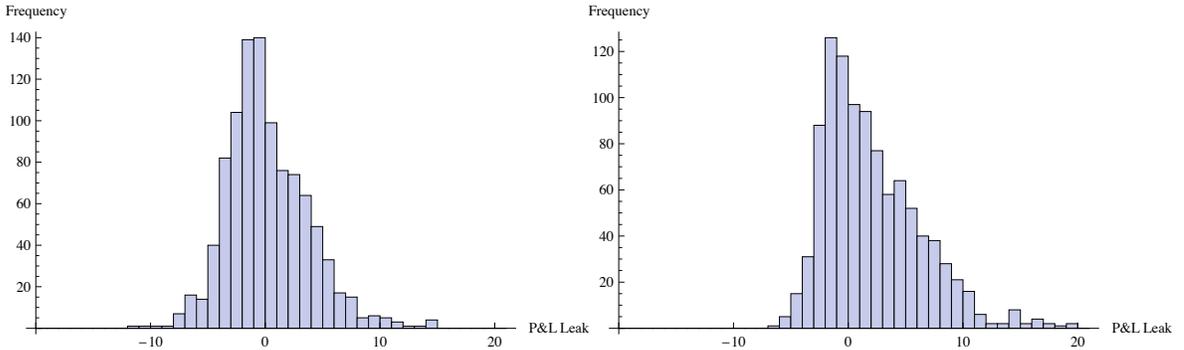


Figure 1: the P&L of delta hedging using ratios computed with the multiscale stochastic volatility model (left pane) and the multivariate lognormal model (right pane) when the true dynamics of the underlying assets are given by the multiscale stochastic volatility model

For the sake of fairness, we also report the results of a second simulation in which the historical

dynamics are assumed to be the multivariate lognormal model (4) with parameters $x = 50$, $y = 5$, $\mu_1 = 0.1$, $\mu_2 = 0.1$, $\sigma_1 = 0.5$, $\sigma_2 = 0.5$, $\rho = 0.5$, $M = 10$, $K = 0$, $T = 1$, $r = 0$.

Again, we use Monte Carlo simulation to generate 2000 sample paths of X_t and Y_t . For each path, we sell one spread option. The portfolio is delta-hedged with both the multiscale stochastic volatility model and the multivariate lognormal model. The final the P&L of each path is collected for the two models. The histograms of the the P&L are given by Figure 2. As expected, when hedged by the multivariate lognormal model, the profit and loss ratio mean is reduced to 2.1%. However, when hedged by a multiscale stochastic volatility model, the ratio mean is only 7.8%. This result is satisfactory compared with the 18.3% average profit and loss when hedging with the multivariate lognormal model in the first simulation.

Our numerical experiment shows the accuracy of the asymptotic formula (68) for option valuation under the multiscale stochastic volatility model. The delta-hedging performance of the multiscale stochastic volatility model is consistent when the underlying processes are given by different models.

Table 2: the P&L ratio statistics for the simulation experiment with the true dynamics of the multivariate lognormal model

	Mean	Variance
Multiscale Stochastic Volatility Model	0.07770	0.1051
Multivariate Lognormal model	0.0209	0.0527

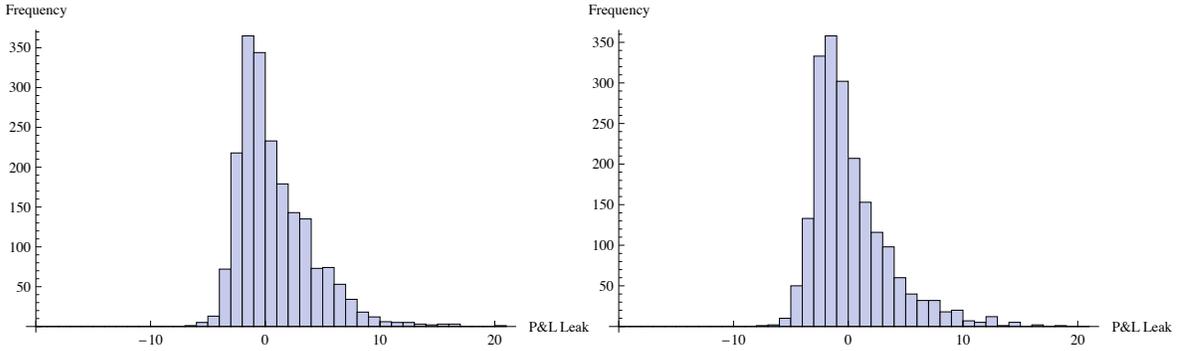


Figure 2: the P&L of delta hedging using ratios computed with the multiscale stochastic volatility model (left pane) and the multivariate lognormal model (right pane) when the true dynamics of the underlying assets are given by the multivariate lognormal model

3 Implied Correlation and Calibration

In Section 2, we discussed the setup of our two factor multiscale stochastic volatility model and derived an asymptotic pricing formula for European options. In this section, we derive a corresponding asymptotic formula for the implied correlation of the spread options. We also discuss the calibration of the model to real data.

3.1 Asymptotics for Implied Correlation

Recall that the implied correlation of a spread option is the level of correlation we need to put into a basic pricer to generate the market price of the option. We assume the option to be the spread option with payoff $h(x, y) = (x - My - K)^+$. We also assume the basic pricer to be any spread option pricer under the multivariate lognormal model (4), which is also the function C_0 in our approximation. The implied correlation is the value $I^{\epsilon, \delta}$ such that

$$C^{\epsilon, \delta} = C_0(x, y, t; \bar{\sigma}_1(v), \bar{\sigma}_2(v), I^{\epsilon, \delta}). \quad (82)$$

When the strike K is zero, the spread option price C_0 is given by Margrabe's formula [38], and C_0 is monotonically decreasing in the correlation $I^{\epsilon, \delta}$. As a result, a unique implied correlation can be extracted from the above equation.

When the strike is not zero, there is no closed form spread option valuation formula under the multivariate lognormal model. The monotonicity depends on the choice of the approximation pricer. In general, we expect that the spread option pricer remains monotonically decreasing in the correlation. Kirk's formula [24] is a case in point. The intuition is that when the correlation is large, the two underlying asset prices tend to move in the same direction, the spread is lower and thus the value of the option decreases. As a result, a unique implied correlation $I^{\epsilon, \delta}$ can be solved from (82).

Our strategy is to expand $I^{\epsilon, \delta}$ in powers of ϵ and δ :

$$I^{\epsilon, \delta} = I_0 + \sqrt{\epsilon}I_1 + \sqrt{\delta}I_2 + \dots. \quad (83)$$

Inserting (83) into equation (82), we develop the resulting expression using Taylor's formula

$$\begin{aligned} C^{\epsilon, \delta} &= C_0(x, y, t; I_0 + \sqrt{\epsilon}I_1 + \sqrt{\delta}I_2) \\ &= C_0(x, y, t; I_0) + (\sqrt{\epsilon}I_1 + \sqrt{\delta}I_2) \frac{\partial C_0}{\partial \rho}. \end{aligned} \quad (84)$$

Comparing with equation (71) to identify the various terms, we find $I_0 = \rho$ and

$$\begin{aligned} I_1 \frac{\partial C_0}{\partial \rho} &= -(T-t) \left\{ \frac{1}{\bar{\sigma}_1(v)} \left[P_1^{(0)} \frac{\partial}{\partial \sigma_1} + P_1^{(1)} x \frac{\partial^2}{\partial x \partial \sigma_1} + P_1^{(1)} y \frac{\partial^2}{\partial y \partial \sigma_1} \right] \right. \\ &\quad \left. + \frac{1}{\bar{\sigma}_2(v)} \left[P_2^{(0)} \frac{\partial}{\partial \sigma_2} + P_2^{(1)} y \frac{\partial^2}{\partial y \partial \sigma_2} + P_2^{(0)} x \frac{\partial^2}{\partial x \partial \sigma_2} \right] \right\} C_0, \\ I_2 \frac{\partial C_0}{\partial \rho} &= -(T-t) \left[P_1^{(2)} x^2 \frac{\partial^2}{\partial x^2} + P_1^{(3)} x \frac{\partial}{\partial x} x^2 \frac{\partial^2}{\partial x^2} \right. \\ &\quad \left. + P_2^{(2)} y^2 \frac{\partial^2}{\partial y^2} + P_2^{(3)} y \frac{\partial}{\partial y} y^2 \frac{\partial^2}{\partial y^2} + P_3^{(3)} x \frac{\partial}{\partial x} y^2 \frac{\partial^2}{\partial y^2} + P_4^{(4)} y \frac{\partial}{\partial y} x^2 \frac{\partial^2}{\partial x^2} \right. \\ &\quad \left. + P^{(2)} xy \frac{\partial^2}{\partial x \partial y} + P^{(5)} x \frac{\partial}{\partial x} xy \frac{\partial^2}{\partial x \partial y} + P^{(6)} y \frac{\partial}{\partial y} xy \frac{\partial^2}{\partial x \partial y} \right] C_0. \end{aligned} \quad (85)$$

We choose to approximate the implied correlation by

$$I^{\epsilon, \delta} \approx \rho + \sqrt{\epsilon}I_1 + \sqrt{\delta}I_2. \quad (86)$$

For any given analytic formula for C_0 , the implied correlation approximation (86) can be calculated explicitly. In most cases, formulas giving the partial derivatives are involved. However, when $K = 0$

and C_0 becomes Margrabe's formula, the results are relatively simple. The equations below give I_1 and I_2 in the implied correlation approximation (86) under Margrabe's formula.

$$\begin{aligned}
 I_1 &= -\frac{T-t}{\sigma_M^2} \left[\frac{\rho\bar{\sigma}_2 - \bar{\sigma}_1}{\bar{\sigma}_1\bar{\sigma}_2} (P_1^{(0)} + \frac{1}{2}P_1^{(1)} + \frac{1}{2}P^{(1)}) + \frac{\rho\bar{\sigma}_1 - \bar{\sigma}_2}{\bar{\sigma}_1\bar{\sigma}_2} (P_2^{(0)} + \frac{1}{2}P_2^{(1)} + \frac{1}{2}P^{(0)}) \right] \\
 &\quad + \left[\frac{\rho\bar{\sigma}_2 - \bar{\sigma}_1}{\bar{\sigma}_1\bar{\sigma}_2} (P_1^{(1)} - P^{(1)}) + \frac{\rho\bar{\sigma}_1 - \bar{\sigma}_2}{\bar{\sigma}_1\bar{\sigma}_2} (P^{(0)} - P_2^{(1)}) \right] \ln\left(\frac{x}{My}\right) \\
 I_2 &= -\frac{1}{\bar{\sigma}_1\bar{\sigma}_2\sigma_M^2} (-P_1^{(2)} - P_2^{(2)} + P^{(2)} - \frac{1}{2}P_1^{(3)} - \frac{1}{2}P_2^{(3)} + \frac{1}{2}P^{(5)} + \frac{1}{2}P^{(6)} - \frac{1}{2}P^{(3)} - \frac{1}{2}P^{(4)}) \\
 &\quad - \frac{1}{\bar{\sigma}_1\bar{\sigma}_2\sigma_M^2} (P_1^{(3)} - P_2^{(3)} - P^{(5)} + P^{(6)} + P^{(3)} - \frac{1}{2}P^{(4)}) \frac{\ln(\frac{x}{My})}{T-t},
 \end{aligned} \tag{87}$$

With (87), the implied correlation asymptotics in (86) can be simplified as

$$I^{\epsilon, \delta} = \rho + b^\epsilon + a^\epsilon \frac{\ln(\frac{x}{My})}{T} + a^\delta \ln\left(\frac{x}{My}\right) + b^\delta T, \tag{88}$$

where

$$\begin{aligned}
 a^\epsilon &= \frac{1}{\bar{\sigma}_1\bar{\sigma}_2} \left[\frac{P_1^{(3)}}{\bar{\sigma}_1^2} - \frac{P_2^{(3)}}{\bar{\sigma}_2^2} \right], \\
 b^\epsilon &= -\frac{\bar{\sigma}_0^2}{2\bar{\sigma}_1\bar{\sigma}_2} \left[\left(\frac{P_2^{(3)}}{\bar{\sigma}_2^2} + \frac{P_1^{(3)}}{\bar{\sigma}_1^2} \right) + \left(\frac{P_2^{(2)}}{\bar{\sigma}_2^2} + \frac{P_1^{(2)}}{\bar{\sigma}_1^2} \right) \right], \\
 a^\delta &= \frac{1}{\bar{\sigma}_0^2\bar{\sigma}_1\bar{\sigma}_2} \left[\frac{\rho\bar{\sigma}_2 - \bar{\sigma}_1}{\bar{\sigma}_1} (P_1^{(1)} - P_2^{(1)} \frac{P_1^{(0)}}{P_2^{(0)}}) + \frac{\rho\bar{\sigma}_1 - \bar{\sigma}_2}{\bar{\sigma}_2} (P_1^{(1)} \frac{P_1^{(0)}}{P_2^{(0)}} - P_2^{(1)}) \right], \\
 b^\delta &= -\frac{1}{\bar{\sigma}_1\bar{\sigma}_2} \left[\frac{\rho\bar{\sigma}_2 - \bar{\sigma}_1}{\bar{\sigma}_1} (P_1^{(0)} + \frac{1}{2}P_1^{(1)} + \frac{1}{2}P_2^{(1)} \frac{P_1^{(0)}}{P_2^{(0)}}) + \frac{\rho\bar{\sigma}_1 - \bar{\sigma}_2}{\bar{\sigma}_2} (P_2^{(0)} + \frac{1}{2}P_1^{(1)} \frac{P_1^{(0)}}{P_2^{(0)}} + \frac{1}{2}P_2^{(1)}) \right].
 \end{aligned} \tag{89}$$

As we will show later, $(\bar{\sigma}_1, P_1^{(0)}, P_1^{(1)}, P_1^{(2)}, P_1^{(3)})$ and $(\bar{\sigma}_2, P_2^{(0)}, P_2^{(1)}, P_2^{(2)}, P_2^{(3)})$ are calibrated to individual implied volatility curves. ρ is the additional input for the joint model that allows extra flexibility in the implied correlation curve. The coefficients $(a^\epsilon, b^\epsilon, a^\delta, b^\delta)$ are functions of the parameters of the individual implied volatilities and the Brownian motions correlation ρ .

3.2 Coupling Implied Correlation and Implied Volatilities

Spread option trading volume is usually small while individual options are much more liquid. For example, in the case of the spark spread option between electric power and natural gas, the volume of trades reported by financial data providers is rather small and it is practically impossible to calibrate implied correlations to market spread option data. Alternatively, natural gas options and power options are much more liquid. To fully utilize market information, we calibrate part of the implied correlation coefficients to the individual option data of the two underlying assets. In this way, the valuation of spread options are consistent with market information for individual options.

Assume that the two underlying assets X_t and Y_t follow the multiscale stochastic volatility model (16). Then each individual underlying asset follows the single asset multiscale stochastic volatility model discussed in [20]. When fitting individual implied volatility curves, the parameters $(\bar{\sigma}_1, P_1^{(0)}, P_1^{(1)}, P_1^{(2)}, P_1^{(3)})$

and $(\bar{\sigma}_2, P_2^{(0)}, P_2^{(1)}, P_2^{(2)}, P_2^{(3)})$ can be calibrated. Take X_t as an example. If viewed individually, X_t follows the risk-neutral model

$$\begin{cases} dX_t &= rX_t dt + X_t f(Z_t) f_1(V_t) dW_t^{(X)*}, \\ dZ_t &= [\frac{1}{\epsilon}(m - Z_t) - \frac{\nu\sqrt{2}}{\sqrt{\epsilon}}\sqrt{Z_t}\Lambda(Z_t, V_t)]dt + \frac{\nu\sqrt{2}}{\sqrt{\epsilon}}\sqrt{Z_t}dW_t^{(Z)*}, \\ dV_t &= [\delta c(V_t) - \sqrt{\delta}g(V_t)\Gamma(Z_t, V_t)]dt + \sqrt{\delta}g(V_t)dW_t^{(V)*}. \end{cases} \quad (90)$$

The price of European call options written on this asset are given by the conditional expectations

$$C^{X,\epsilon,\delta}(x, z, v, t; T, M, K) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[(X_T - K)^+ | X_t = x, Z_t = z, V_t = v]. \quad (91)$$

According to [20], the option price can be approximated by the following asymptotic formula

$$\begin{aligned} C^{X,\epsilon,\delta} &\approx C_0 - (T-t)\left\{\frac{1}{\bar{\sigma}_1(v)}\left[P_1^{(0)}(v)\frac{\partial}{\partial\sigma} + P_1^{(1)}(v)x\frac{\partial^2}{\partial x\partial\sigma}\right] \right. \\ &\quad \left. + [P_1^{(2)}(v)x^2\frac{\partial^2}{\partial x^2} + P_1^{(3)}(v)x\frac{\partial}{\partial x}(x^2\frac{\partial^2}{\partial x^2})]\right\}C_0, \end{aligned} \quad (92)$$

where the parameters $(\bar{\sigma}_1, P_1^{(0)}, P_1^{(1)}, P_1^{(2)}, P_1^{(3)})$ are defined in (69). The implied volatility is then defined as the value I^X such that

$$C_{BS}(x, t; T, K, I^X) = C^{X,\epsilon,\delta}. \quad (93)$$

I^X can be approximated by the following asymptotic expansion

$$\begin{aligned} I^X &\approx \bar{\sigma}_1(v) + \left[-\frac{P_1^{(2)}(v)}{\bar{\sigma}_1(v)} + \frac{P_1^{(3)}(v)}{\bar{\sigma}_1^3(v)}\left(r - \frac{\bar{\sigma}_1^2(v)}{2}\right)\right] \\ &\quad + \left[-\frac{P_1^{(3)}(v)}{\bar{\sigma}_1^3(v)}\right]\frac{\ln(\frac{K}{x})}{T-t} + \left[-\frac{P_1^{(1)}(v)}{\bar{\sigma}_1^3(v)}\right]\ln\left(\frac{K}{x}\right) \\ &\quad + \left[-\frac{P_1^{(0)}(v)}{\bar{\sigma}_1(v)} + \frac{P_1^{(1)}(v)}{\bar{\sigma}_1^3(v)}\left(r - \frac{\bar{\sigma}_1^2(v)}{2}\right)\right](T-t). \end{aligned} \quad (94)$$

Since we take forward prices as underlying, we have $r = 0$. Denote

$$\begin{aligned} a_1^\epsilon &= -\frac{P_1^{(3)}(v)}{\bar{\sigma}_1^3(v)}, & b_1^\epsilon &= -\frac{P_1^{(2)}(v)}{\bar{\sigma}_1(v)} - \frac{1}{2}\frac{P_1^{(3)}(v)}{\bar{\sigma}_1(v)}, \\ a_1^\delta &= -\frac{P_1^{(1)}(v)}{\bar{\sigma}_1^3(v)}, & b_1^\delta &= -\frac{P_1^{(0)}(v)}{\bar{\sigma}_1(v)} - \frac{1}{2}\frac{P_1^{(1)}(v)}{\bar{\sigma}_1(v)}. \end{aligned} \quad (95)$$

Then the implied volatility is approximated by

$$I^X \approx \bar{\sigma}_1 + b_1^\epsilon + a_1^\epsilon\frac{\ln(\frac{K}{x})}{T-t} + a_1^\delta\ln\left(\frac{K}{x}\right) + b_1^\delta(T-t). \quad (96)$$

Similarly, for the options written on Y_t , the implied volatility of Y_t can be approximated by

$$I^Y \approx \bar{\sigma}_2 + b_2^\epsilon + a_2^\epsilon\frac{\ln(\frac{K}{x})}{T-t} + a_2^\delta\ln\left(\frac{K}{x}\right) + b_2^\delta(T-t). \quad (97)$$

where

$$\begin{aligned} a_2^\epsilon &= -\frac{P_2^{(3)}(v)}{\bar{\sigma}_2^3(v)}, & b_2^\epsilon &= -\frac{P_2^{(2)}(v)}{\bar{\sigma}_2(v)} - \frac{1}{2} \frac{P_2^{(3)}(v)}{\bar{\sigma}_2(v)}, \\ a_2^\delta &= -\frac{P_2^{(1)}(v)}{\bar{\sigma}_2^3(v)}, & b_2^\delta &= -\frac{P_2^{(0)}(v)}{\bar{\sigma}_2(v)} - \frac{1}{2} \frac{P_2^{(1)}(v)}{\bar{\sigma}_2(v)}, \end{aligned} \quad (98)$$

and $(\bar{\sigma}_2, P_2^{(0)}, P_2^{(1)}, P_2^{(2)}, P_2^{(3)})$ are defined in (69).

In practice, the implied volatilities of X and Y are fitted to the option prices of the two individual assets to obtain the coefficients $(\bar{\sigma}_1, P_1^{(0)}, P_1^{(1)}, P_1^{(2)}, P_1^{(3)})$ and $(\bar{\sigma}_2, P_2^{(0)}, P_2^{(1)}, P_2^{(2)}, P_2^{(3)})$. Then, we introduce ρ as an additional input for extra flexibility in the implied correlation curve. The calibrated coefficients are used to calculate the implied correlation of the multi-asset option following (86) and (85) from which we obtain option prices. Next, we illustrate the calibration results using power and natural gas option data and discuss in more details the calibration procedure.

3.3 Implied Volatility and Implied Correlation Calibration

In this section, we discuss the calibration of formula (96) and (97) to market implied volatility data, and the derivation of the implied correlation curve from the calibrated parameters.

The calibration to implied volatility data follows the method proposed in [20]. For a set of time-to-maturity T_i , we have implied volatilities $I^X(T_i, K_{ij})$ and $I^Y(T_i, K_{ij})$ for different strikes K_{ij} . We estimate the parameters $(a_1^\epsilon, b_1^\epsilon, a_1^\delta, b_1^\delta)$ and $(a_2^\epsilon, b_2^\epsilon, a_2^\delta, b_2^\delta)$ in (96) and (97) with the following steps (use the calibration for I^X as an example)

1. Approximate $\bar{\sigma}_1$ by the average implied volatility value.
2. For fixed time-to-maturity T_i , regress the implied volatilities for different strikes $(K_{ij})_j$ on the corresponding log-moneyness to maturity ratio

$$(LMMR)_{ij} = \frac{\ln(K_{ij}/x)}{T_i}$$

following a linear model $I^X(T_i, K_{ij}) = \beta(T_i) + \alpha(T_i)(LMMR)_{ij}$

3. Regress the coefficients on the time-to-maturity T_i following linear models

$$\begin{aligned} \alpha(T_i) &= a_1^\epsilon + a_1^\delta T_i, \\ \beta(T_i) &= \bar{\sigma}_1 + b_1^\epsilon + b_1^\delta T_i, \end{aligned}$$

and get estimates of $(a_1^\epsilon, b_1^\epsilon, a_1^\delta, b_1^\delta)$.

We perform the same procedure for I^Y , and get estimates of $(a_2^\epsilon, b_2^\epsilon, a_2^\delta, b_2^\delta)$. Using equations (95) and (98), we compute the coefficients $(P_1^{(0)}, P_1^{(1)}, P_1^{(2)}, P_1^{(3)})$ and $(P_2^{(0)}, P_2^{(1)}, P_2^{(2)}, P_2^{(3)})$.

We illustrate the fitting procedure using power and natural gas option data from the Global Insight database. The electric power data is from PJM electricity daily settlement prices. PJM Interconnection is a regional transmission organization that coordinates the movement of wholesale electricity in all or parts of 13 states in the North East of the US as well as the District of Columbia [1]. The natural gas data is from the Henry Hub (Sabine pipeline) natural gas daily settlement prices.

Henry Hub is the pricing point for natural gas futures contracts traded on the New York Mercantile Exchange (NYMEX). It is a point on the natural gas pipeline system in Erath, Louisiana. It is owned by Sabine Pipe Line LLC [2].

In the commodity market, the underlying asset of an option is not the commodity itself, but rather a futures contract for that commodity. For example, a November natural gas option will actually be an option for a November delivery natural gas futures contract. In this sense, the options are on futures instead of the physical commodity. Futures contracts with a definite maturity date imply physical delivery throughout the following month. Also, options have a date at which they mature and expire. For example, a \$7.00 November natural gas option is an option to buy or sell one November natural gas futures contract at \$7.00. The option can be exercised by the holder on any business day until mid-October at which time the option expires. Trading in most options will not be conducted during the futures contract delivery month.

By definition, the commodity options are more "American" than the standard European options which can only be exercised on the maturity date. However, if there is no dividend payment, the investors holding American-style call options usually do not have any reason to exercise the option before maturity. This is because the options have non-negative time value and are usually worth more unexercised. As a result, we can use the standard European option approach to treat the commodity call options.

Also, when we build implied volatility surfaces, we choose options with the same time-of-maturity, and different trading dates. We choose power and natural gas options maturing in August 2011, and traded in April 2010, May 2010, until April 2011. As we explained earlier, these options have the same underlying - the August 2011 futures contract. The option prices are more correlated than the options with the same trading date and different maturity time. Figure 3a and Figure 3b give the implied volatility surfaces calculated from these option prices.

For these two implied volatilities surfaces, we calibrated the coefficients $(P_1^{(0)}, P_1^{(1)}, P_1^{(2)}, P_1^{(3)})$ and $(P_2^{(0)}, P_2^{(1)}, P_2^{(2)}, P_2^{(3)})$ following the procedure discussed earlier. The fitted coefficients are listed in Table 3. The additional parameter ρ is chosen to be 0.3 for the model.

Table 3: Fitted multivariate stochastic volatility model coefficients

Coefficient from X and Y			
$\bar{\sigma}_1$	0.3174	$\bar{\sigma}_2$	0.3160
$P_1^{(0)}$	-0.0658	$P_2^{(0)}$	0.0740
$P_1^{(1)}$	-0.0034	$P_2^{(1)}$	0.0001
$P_1^{(2)}$	0.0943	$P_2^{(2)}$	-0.0585
$P_1^{(3)}$	0.0029	$P_2^{(3)}$	-0.0025

Figures 4a and 4b give the implied correlation surfaces calculated from the fitted coefficients. Figure 4a plots the computed implied correlation surface over the plane of log-moneyness $\ln(x/My)$ and time-to-maturity. The strike is fixed at 0 for this plot. The implied correlation is calculated by (88). For fixed maturity, the implied correlation is affine in the log-moneyness $\ln(x/My)$. When the maturity changes, the slope in the log-moneyness becomes more pronounced for short maturities compared with long maturities. The observation is consistent with the implied correlation formula (88), in which the coefficient of the log-moneyness is $a^\epsilon/T + a^\delta$. The coefficient increases in absolute value when T goes to zero. As a result, the slope becomes more pronounced for short maturity.

Figure 4b plots the computed correlation over the log-moneyness $\ln(x/My)$ and the strike K . Time-to-maturity is chosen as 1 year for this plot. When the strike is away from zero, the implied correlation have slight convexity or concavity in the log-moneyness $\ln(x/My)$, instead of being affine in the log-moneyness when $K = 0$.

From these results, the multiscale stochastic volatility model generates implied correlation skews. The slope and convexity of the skew are controlled by the parameters $(a^\epsilon, b^\epsilon, a^\delta, b^\delta)$, the level of the log-moneyness, the strike and the time-to-maturity. These results are consistent with market expectations.

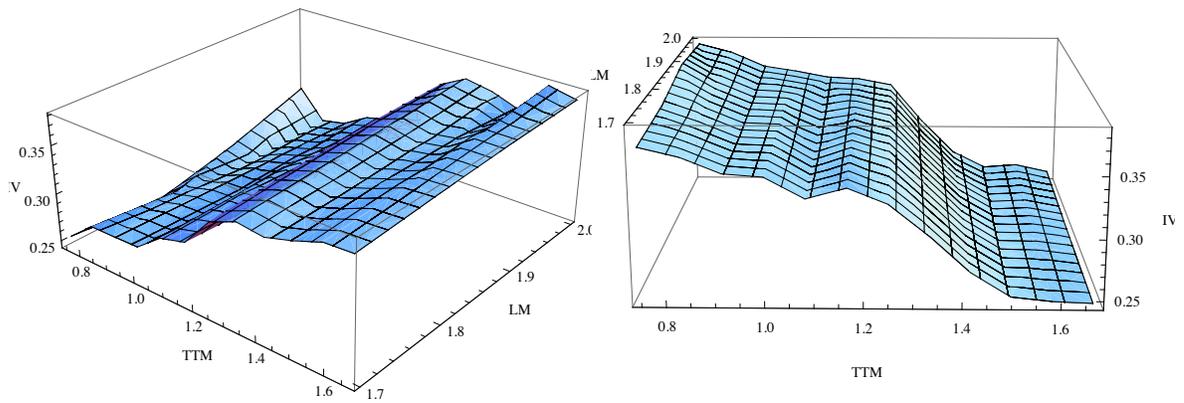


Figure 3: Implied volatility for options on electric power futures maturing in August 2011, and traded in April 2010, May 2010, until April 2011 (left pane) and for options on natural gas futures maturing in August 2011, and traded in April 2010, May 2010, until April 2011 (right pane)

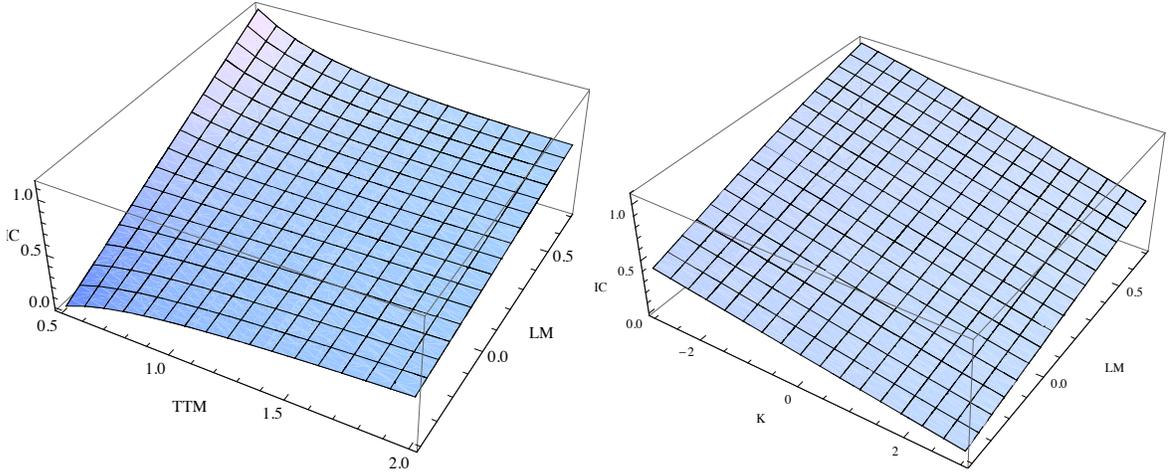


Figure 4: Implied correlation (IC) plotted over time-to-maturity (TTM) and log-moneyness (LM), with fitted parameters and strike $K = 0$ (left pane) and plotted over strike (K) and log-moneyness (LM), with fitted parameters and time-to-maturity $TTM = 1$ year (right pane)

3.4 Application to Power Plant Valuation

Managing power generation assets is a complicated task. Although for a regulated utility, the objective is to satisfy its customer demand and minimize cost, we restrict the focus of the energy service company in our example on profit maximization. Risk managers of power plants need a fast approximation tool to model the value of their assets, especially if they are the objects of tolling or leasing contracts. The tool is also useful for investors who need to evaluate the economic feasibility of investing in power generation assets.

The fuel is one of the primary costs in power generation. To capture the maximum spread between the profit from generating power and the underlying fuel costs, operators usually dispatch schedules that are optimized. In a simplified and idealized scenario, the operator of the power plant knows the daily fuel price and power price, and can choose to switch on or off the generation on a daily basis in response to market price movements in power and fuel.

Assume the power plant uses only one fuel type, say natural gas, then the daily operation of the plant over a period from time $t = 1$ to $t = T$ can be identified with the following payoff

$$C = \sum_{t=1}^T e^{-rt} L(t) \max\{P(t) - H(t)G(t) - X, 0\}, \quad (99)$$

where P refers to the price at which one unit of power can be sold, G to the price at which one unit of natural gas can be purchased, H is the energy efficiency heat rate of the power generation asset, X is the fixed costs of operations and management (O&M) of the plant, L is the daily load, and r the interest rate. Taking expectations on both sides of (99) shows that the expected P&L can be viewed as a strip of spread options.

From an implied correlation surface, one can calculate the spread option prices accurately. The advantage of the multiscale stochastic volatility model and the implied correlation approach is to provide calibration to the marginal option prices and the term structure.

For numerical purposes, we assume that the power plant can produce up to 1,000 Megawatt-hours of electricity everyday when it operates. The plant has a conversion heat rate of 9 Megawatt-hour/Million BTU. This is largely simplified for a real power plant, but it serves as an illustrative example. Assume the fixed cost is, and that the O&M are of the order of 500 USD everyday the plant operates, so that the daily P&L is

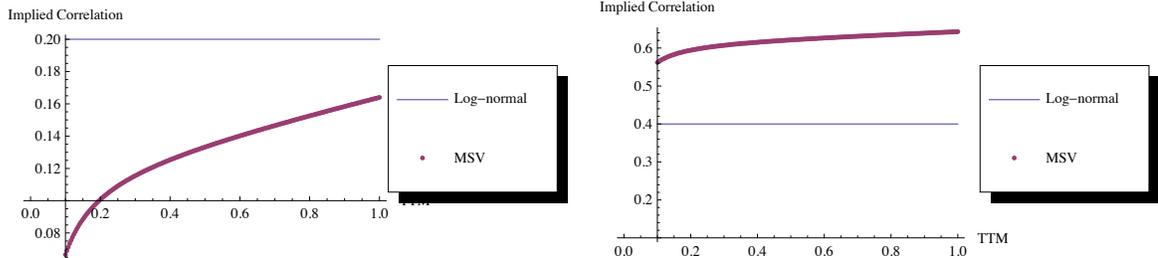
$$\max\{1000[P(t) - 9G(t)] - 500, 0\}. \tag{100}$$

The current power price is assumed to be 45 Dollar/Megawatt-hour, and the current natural gas price is assumed to be 5 Dollar/Million BTU. We assume the two underlying asset prices follow the multiscale stochastic volatility model with parameters and implied correlation surfaces calibrated in the previous section. The power plant valuation can be calculated as the price of a strip of spread options with payoff (99).

To analyze the effectiveness of the multiscale stochastic volatility approach, we compare the power plant valuation result with the traditional multivariate lognormal model. In the lognormal model, we assume that the volatilities $\bar{\sigma}_1$ and $\bar{\sigma}_2$ are calibrated as in the previous section, and we assume that the correlation is the Brownian motion correlation ρ . Recall that with this parametrization, the lognormal model becomes a special case of model (16), with Z_t and V_t being constant.

First, we compare the implied correlation values and spread option prices for different time-to-maturity under the two models. Figure 5c and Figure 5d shows the implied correlations and prices of spread options with time-to-maturity from 1 day to 1 year for different value of ρ . In these plots, the implied correlation curve generated by the multiscale stochastic volatility model has a term structure compared with the fixed correlation of the multivariate lognormal model. The implied correlation tends to increase with the time-to-maturity. In other words, the two underlying assets are de-correlated when the time-to-maturity goes to zero. This structure is consistent with market expectations. When the time-to-maturity is small, the two underlying assets are impacted by their own randomness, and the correlation becomes smaller.

The implied correlation curve generated by the multiscale stochastic volatility model can be either higher or lower than the correlation of the multivariate lognormal model. The relationship depends on the value of ρ . In Figure 5c, the implied correlation from the multiscale stochastic volatility model is lower when ρ is small. In Figure 5d, the multiscale stochastic volatility model implied correlation is higher when ρ is large. In practice, the correlation level of power and natural gas is impacted by many external factors, temperature being one of them, and it takes different value for different power markets. We gather that traders determine the value of ρ using price observation and their understanding of the market.



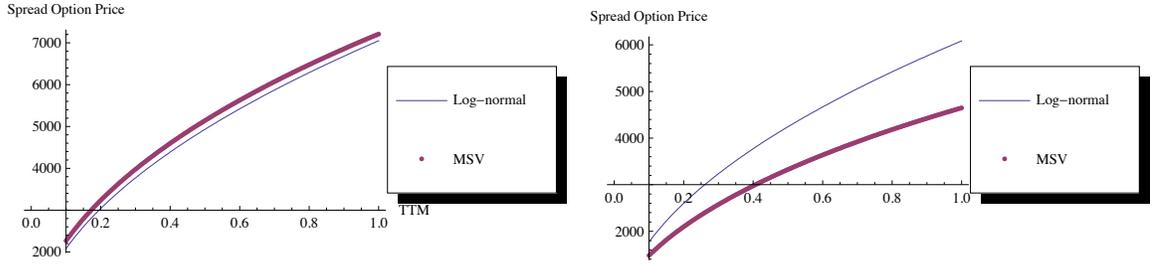


Figure 5: Implied correlations and spread option prices under the multiscale stochastic volatility model (MSV) and the multivariate lognormal model for different time-to-maturity. Parameter chosen as: $\rho = 0.2$ (left pane) and $\rho = 0.5$ (right pane)

Figure 6 displays the power plant valuation for different values of ρ . The higher ρ , the lower the power plant value. This is because the higher the correlation between natural gas and power, the more likely are periods of non-profitability.

As we discussed earlier, the multivariate lognormal model tends to overestimate the correlation between power and natural gas when the input ρ is small, and underestimate the correlation when the input ρ is large. This is confirmed by the power plant valuation plot. The valuation from the multiscale stochastic volatility model is higher than the valuation from the multivariate lognormal model when ρ is small, and lower when ρ is large.

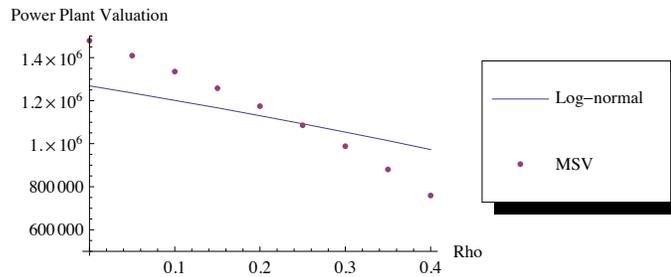


Figure 6: Power plant valuation for different values of ρ under the multiscale stochastic volatility model (MSV) and the multivariate lognormal model

The heat rate H is a very important characteristic of a power plant. The value of H was fixed at 9 for the pervious calculations. H can have different values depending on the efficiency of the power plant as well as other factors. The lower the heat rate, the more efficient the power plant, and the higher its value. For this example, Figure 7 exhibits valuations of the power plants with different values of the heat rate and the other parameters being held constant.

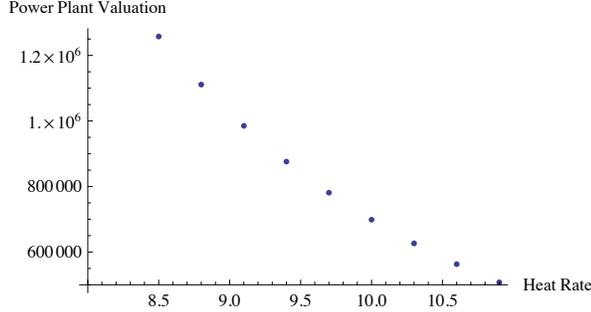


Figure 7: Power plant valuation for different heat rate under the multiscale stochastic volatility model

Remark 3.1. *A real power plant has so many optionalities that a simple strip of spread options cannot account for. In practice, the power market is divided into peak and off-peak hours. For example, if the operator runs the turbines overnight at a loss (low power prices) and still accumulates enough heat so he can ramp up the plant to full capacity in peak hours at low cost, he can still make money for the whole cycle. Conversely, if off-peak power prices are too low, the operator would turn the plant off and then incur the cost of a cold start to produce during peak hours.*

The rationale for valuing a power plant as a strip of spread option assumes the operators know the power and natural gas prices one day ahead. In reality, the optionality is better captured by an optimal switching algorithm like in [11].

4 Local Correlation

The multiscale stochastic volatility model introduced earlier enriches the multivariate lognormal model to create an implied correlation skew. However, the stochastic volatilities carry non-tradable sources of risk, and the market model is not complete. For options with a single underlying asset, Dupire [17] introduced a local volatility approach, which assumes that the instantaneous volatility is a deterministic function of time and the underlying prices. The local volatility approach extends the Samuelson model, while preserving completeness of the market and one-dimensionality of uncertainty.

Motivated by the success of Dupire’s approach, we establish a local correlation theory for spread option modeling. We assume that the instantaneous correlation is a deterministic function of time and the underlying prices, which is called the local correlation. We derive formulas for the local correlation function involving transformations of the spread option price surface. The calibrated local correlation model can be used to price exotic and path-dependent derivatives.

4.1 Local Correlation: Existence and Uniqueness

In [21] Gyongy gave the theoretical foundations underpinning Dupire’s approach. He gave sufficient conditions for an Itô process to have the same one dimensional marginal distributions as a Markov diffusion whose drift and volatility are given by deterministic functions of time and the process itself. The volatility and correlation functions are determined by conditional expectations. In the financial arena, this means that the two processes will provide the same European option prices, hence the relevance to Dupire’s approach.

4.2 Local Correlation Formulas

Dupire's local volatility result states that if the prices of European call options are given by a function $C(K, t)$ of strike K and time-of-maturity t which is smooth enough, and the underlying process follows a local volatility model

$$dS_t = S_t[r dt + \sigma(t, S_t)dW], \quad (101)$$

then the local volatility function $\sigma(t, x)$ is given by the following formula:

$$\sigma^2(t, K) = \frac{\frac{\partial C}{\partial t} + rK \frac{\partial C}{\partial K}}{\frac{1}{2}K^2 \frac{\partial^2 C}{\partial K^2}}. \quad (102)$$

In this section, we derive a similar relationship between the local correlation function and spread option prices. We assume that the instantaneous correlation is a function $\rho(t, x, y)$ of the two underlying asset prices x, y and the time t , and the prices of European call spread options with payoff $(X_t - M \times Y_t - K)^+$ are given by a function $C(M, K, t)$ of the strike K , ratio M and time-of-maturity t . We derive the solution for the local correlation function $\rho(t, x, y)$ with the price function $C(M, K, t)$ and other parameters of the underlying model.

Before we state the main result, we recall the definition of the Radon transform, and derive a few simple properties which we will need.

Definition 4.1. *Let $f(x, y)$ be an integrable continuous function defined on \mathbf{R}^2 . The Radon transform of f is the function Rf defined by the line integral*

$$Rf(M, K) = \int_{-\infty}^{\infty} f(K + My, y) dy.$$

The inverse Radon transform exists and is denoted by $R^{-1}f(x, y)$. The inversion formulas can be found in [3]. The first lemma yields the uniqueness of the inverse Radon transform among continuous functions. Its proof can be found in [39].

Lemma 4.1. *Assume that f is an integrable continuous function defined on \mathbf{R}^2 . If f satisfies $\int_l f ds = 0$ for all line l in \mathbf{R}^2 , then $f \equiv 0$.*

Lemma 4.2. *Let us assume that X, Y are random variables with a continuous joint density $f_{X,Y}$, denote $S = X - M \times Y$ where M is a constant, f_S be the density function of S , and $g(x, y)$ be a continuous function such that the Radon transform of $gf_{X,Y}$ exist. Then for every M and K , we have the following relationship*

$$R(gf_{X,Y})(M, K) = f_S(K) \mathbf{E}[g(X, Y) | S = K]. \quad (103)$$

Proof. The proof follows a straightforward calculation:

$$\begin{aligned} f_S(K) \mathbf{E}[g(X, Y) | S = K] &= f_S(K) \mathbf{E}[g(MY + K, Y) | S = K] \\ &= \int_{\mathbf{R}} g(My + K, y) \frac{f_{Y,S}(y, K)}{f_S(K)} f_S(K) dy \\ &= \int_{\mathbf{R}} g(My + K, y) f_{X,Y}(My + K, y) dy \\ &= R(gf_{X,Y})(M, K). \end{aligned} \quad (104)$$

□

For vanilla call options with one underlying asset, the payoff function is $(X_t - K)^+$ and the critical point is $X_t = K$. For spread options, the payoff function is $(X_t - M \times Y_t - K)^+$ and the *critical region* is the line $X_t - M \times Y_t - K = 0$.

Theorem 4.1. *Let $\{X_t, Y_t\}$ be the diffusion model given by*

$$dX_t = \sigma_X(t, X_t, Y_t)dW_X, \quad dY_t = \sigma_Y(t, X_t, Y_t)dW_Y, \quad d \langle W_X, W_Y \rangle = \rho_t dt, \quad (105)$$

where ρ_t is an adapted stochastic correlation process bounded between $+1$ and -1 . Let $\{\tilde{X}_t, \tilde{Y}_t\}$ be the diffusion model given by

$$d\tilde{X}_t = \sigma_X(t, \tilde{X}_t, \tilde{Y}_t)d\tilde{W}_X, \quad d\tilde{Y}_t = \sigma_Y(t, \tilde{X}_t, \tilde{Y}_t)d\tilde{W}_Y, \quad d \langle \tilde{W}_X, \tilde{W}_Y \rangle = \rho(t, \tilde{X}_t, \tilde{Y}_t)dt, \quad (106)$$

where $\rho(t, x, y)$ is a measurable function with values in $[-1, +1]$. We assume that the volatility functions $\sigma_X(t, x, y)$ and $\sigma_Y(t, x, y)$ are positive, continuous, bounded and bounded away from zero, and that the functions $\mathbf{E}(\rho_t | X_t = x, Y_t = y)$ and $\rho(t, x, y)$ are continuous. Also, we assume the joint densities f_{X_t, Y_t} of (X_t, Y_t) and $f_{\tilde{X}_t, \tilde{Y}_t}$ of $(\tilde{X}_t, \tilde{Y}_t)$ exist for all t and are continuous.

Define C and \tilde{C} by:

$$C(M, K, t) = \mathbf{E}[(X_t - M \times Y_t - K)^+], \quad \tilde{C}(M, K, t) = \mathbf{E}[(\tilde{X}_t - M \times \tilde{Y}_t - K)^+].$$

If $C(M, K, t) = \tilde{C}(M, K, t)$ for all M, K, t , then $\rho(t, x, y) = \mathbf{E}(\rho_t | X_t = x, Y_t = y)$ which can be expressed in terms of Radon transform and inverse Radon transform of $C(M, K, t)$, $\sigma_X(t, x, y)$ and $\sigma_Y(t, x, y)$.

Proof. In model (105), $S_t = X_t - M \times Y_t$ is a continuous martingale. By Tanaka's formula, we have

$$(S_t - K)^+ = (S_0 - K)^+ + \int_0^t I(S_t > K) dS_t + \frac{1}{2} L_t^K, \quad (107)$$

where L_t^K is the local time of S at K . See [25] Chapter 8.7 for details. For the spread option price $C(M, K, t)$, we have

$$C(M, K, t) = \mathbf{E}[(S_t - K)^+] = (S_0 - K)^+ + \frac{1}{2} \mathbf{E}(L_t^K), \quad (108)$$

and

$$\frac{\partial C}{\partial t} = \frac{1}{2} \frac{\partial}{\partial t} \mathbf{E}(L_t^K). \quad (109)$$

Next, we show

$$\frac{\partial}{\partial t} \mathbf{E}(L_t^K) = f_{S_t}(K) H_{S_t}(K), \quad (110)$$

where

$$H_{S_t}(K) = \mathbf{E}\left(\frac{d \langle S \rangle_t}{dt} \mid S_t = K\right), \quad (111)$$

and

$$\frac{d \langle S \rangle_t}{dt} = \sigma_X(t, X_t, Y_t)^2 + M^2 \sigma_Y(t, X_t, Y_t)^2 - 2M \sigma_X(t, X_t, Y_t) \sigma_Y(t, X_t, Y_t) \rho_t. \quad (112)$$

Using the occupation time formula for a positive measurable bounded function $g(x)$, we have

$$\int_{-\infty}^{\infty} L_t^K g(K) dK = \int_0^t g(S_t) d \langle S \rangle_t . \quad (113)$$

Taking expectations on both sides we get:

$$\begin{aligned} \int_{-\infty}^{\infty} \mathbf{E}(L_t^K) g(K) dK &= \mathbf{E} \int_0^t g(S_t) d \langle S \rangle_t \\ &= \int_0^t \mathbf{E}(g(S_t) \frac{d \langle S \rangle_t}{dt}) dt \\ &= \int_0^t \mathbf{E}(g(S_t) \mathbf{E}(\frac{d \langle S \rangle_t}{dt} | S_t)) dt \\ &= \int_0^t \int_{-\infty}^{\infty} g(K) H_{S_t}(K) f_{S_t}(K) dK dt \\ &= \int_{-\infty}^{\infty} g(K) \int_0^t H_{S_t}(K) f_{S_t}(K) dt dK . \end{aligned} \quad (114)$$

The order of the integrations can be changed since the integrands are positive and bounded. Since g is an arbitrary positive measurable bounded function, we have

$$\mathbf{E}(L_t^K) = \int_0^t H_{S_t}(K) f_{S_t}(K) dt , \quad (115)$$

for almost every K . Formula (110) follows by taking derivatives on both sides of the equation above and using the continuity of $f_{S_t} H_{S_t}$.

From now on, we assume that t fixed. Using formula (110), we have

$$\begin{aligned} \frac{\partial C}{\partial t} &= \frac{1}{2} \frac{\partial}{\partial t} \mathbf{E}(L_t^K) = \frac{1}{2} f_{S_t}(K) H_{S_t}(K) \\ &= \frac{1}{2} f_{S_t}(K) \mathbf{E}[\sigma_X(t, X_t, Y_t)^2 + M^2 \sigma_Y(t, X_t, Y_t)^2 \\ &\quad - 2M \sigma_X(t, X_t, Y_t) \sigma_Y(t, X_t, Y_t) \rho_t | X_t - M \times Y_t = K] . \end{aligned} \quad (116)$$

Since the σ -field $\sigma(X_t, Y_t)$ generated by X_t and Y_t contains the σ -field $\sigma(\alpha_1 X_t + \alpha_2 Y_t)$ generated by the random variable $\alpha_1 X_t + \alpha_2 Y_t$ for any constants α_1 and α_2 ,

$$\begin{aligned} \frac{\partial C}{\partial t} &= \frac{1}{2} f_{S_t}(K) \mathbf{E}\{ \mathbf{E}[\sigma_X^2(t, X_t, Y_t) + M^2 \sigma_Y^2(t, X_t, Y_t) \\ &\quad - 2M \sigma_X(t, X_t, Y_t) \sigma_Y(t, X_t, Y_t) \rho_t | X_t, Y_t] | X_t - M \times Y_t = K \} \\ &= \frac{1}{2} f_{S_t}(K) \mathbf{E}[\sigma_X^2(t, X_t, Y_t) + M^2 \sigma_Y^2(t, X_t, Y_t) \\ &\quad - 2M \sigma_X(t, X_t, Y_t) \sigma_Y(t, X_t, Y_t) \mathbf{E}[\rho_t | X_t, Y_t] | X_t - M \times Y_t = K] . \end{aligned} \quad (117)$$

Use Lemma 4.2, we have

$$\begin{aligned} \frac{\partial C}{\partial t} &= \frac{1}{2} R(\sigma_X^2(t, \cdot, \cdot) f_{X_t, Y_t}(\cdot, \cdot))(M, K) + \frac{1}{2} M^2 R(\sigma_Y^2(t, \cdot, \cdot) f_{X_t, Y_t}(\cdot, \cdot))(M, K) \\ &\quad - MR(\sigma_X(t, \cdot, \cdot) \sigma_Y(t, \cdot, \cdot) \mathbf{E}[\rho_t | X_t = x, Y_t = y] f_{X_t, Y_t}(\cdot, \cdot))(M, K) . \end{aligned} \quad (118)$$

The Radon transforms exist since the Radon transform of $f_{X_t, Y_t}(x, y)$ exist, and the volatility and correlation functions are bounded. Use again Lemma 4.2 with $g \equiv 1$, we have

$$\frac{\partial^2 C}{\partial K^2} = f_{S_t}(K) = R(f_{X_t, Y_t})(M, K). \quad (119)$$

The steps above apply to the model (106) as well. Without even using the tower property, we have

$$\begin{aligned} \frac{\partial \tilde{C}}{\partial t} &= \frac{1}{2} R(\sigma_X^2(t, \cdot, \cdot) f_{\tilde{X}_t, \tilde{Y}_t}(\cdot, \cdot))(M, K) + \frac{1}{2} M^2 R(\sigma_Y^2(t, \cdot, \cdot) f_{\tilde{X}_t, \tilde{Y}_t}(\cdot, \cdot))(M, K) \\ &\quad - MR(\sigma_X(t, \cdot, \cdot) \sigma_Y(t, \cdot, \cdot) \rho(t, \cdot, \cdot) f_{\tilde{X}_t, \tilde{Y}_t}(\cdot, \cdot))(M, K), \end{aligned} \quad (120)$$

and

$$\frac{\partial^2 \tilde{C}}{\partial K^2} = R(f_{\tilde{X}_t, \tilde{Y}_t})(M, K). \quad (121)$$

By $C = \tilde{C}$, the uniqueness of the inverse Radon transform, (119) and (121), we have

$$f_{X_t, Y_t} = f_{\tilde{X}_t, \tilde{Y}_t}. \quad (122)$$

Then, by (118) and (120), we have

$$\rho(t, x, y) = \mathbf{E}(\rho_t | X_t = x, Y_t = y). \quad (123)$$

Combining (120) and (121), we can solve for $\rho(t, x, y)$. \square

As an illustration, we give the local correlation formula for the Bachelier model. In the Bachelier model, in which the two underlying asset prices are assumed to be Brownian motions. It is a special case of model (106) with $\sigma_X(t, x, y) \equiv \sigma_1$ and $\sigma_Y(t, x, y) \equiv \sigma_2$. The Brownian motions for the two underlying assets are assumed to be correlated by a local correlation function $\rho(t, x, y)$. So

$$dX_t = \sigma_1 dW_X, \quad dY_t = \sigma_2 dW_Y, \quad d \langle W_X, W_Y \rangle = \rho(t, X_t, Y_t) dt. \quad (124)$$

Define $C(M, K, t)$ as the spread option price surface. Combining (120) and (121) with the inverse Radon transform R^{-1} , the local correlation function is given by the following formula:

$$\rho(t, x, y) = \frac{R^{-1}\left(\frac{\partial^2 C}{\partial K^2} \left(\frac{\sigma_1}{2M\sigma_2} + \frac{M\sigma_2}{2\sigma_1}\right) - \frac{\partial C}{\partial t} \frac{1}{M\sigma_1\sigma_2}\right)}{R^{-1}\left(\frac{\partial^2 C}{\partial K^2}\right)}. \quad (125)$$

This equation is the analog of Dupire's equation (102).

The local correlation formula is more involved when the functions $\sigma_X(t, x, y)$ and $\sigma_Y(t, x, y)$ are not constants. In those model, we need to compute the density function $f_{X_t, Y_t}(x, y)$ using

$$f_{X_t, Y_t} = R^{-1}\left(\frac{\partial^2 C}{\partial K^2}(\cdot, \cdot, t)\right), \quad (126)$$

and then calculate the value of

$$R(\sigma_X^2(t, \cdot, \cdot) f_{X_t, Y_t}(\cdot, \cdot)), \quad R(\sigma_Y^2(t, \cdot, \cdot) f_{X_t, Y_t}(\cdot, \cdot)). \quad (127)$$

Then we can invert the local correlation function by (120). The formula is more complicated, but it is still in closed-form.

In the next section, we perform a numerical test of local correlation calibration, and use the calibrated result to price a path-dependent derivative.

4.3 Numerical Example: Capped Variance Swap

With the local correlation function calibrated from spread option prices, we have complete knowledge of the dynamics of the underlying processes, and we can price exotic and path-dependent derivatives by Monte-Carlo methods. In this last subsection, we illustrate the calibration and computation procedures with a path-dependent derivative, a capped variance swap on a spread.

Capped variance swaps are derivatives on realized variance. In our example, we use the annualized variance of the spread $X_t - MY_t$ as the underlying asset. The realized variance is defined as

$$I(M; t_0, t_N) = \frac{AF}{N} \sum_{n=1}^N \left[\left(\frac{S_{t_n}}{S_{t_{n-1}}} - 1 \right)^2 \right], \quad (128)$$

where $S_t = X_t - MY_t$ is the spread between the underlying assets X_t and Y_t , AF is an annualization factor (typically $AF = 252$), N is the number of fixings during the period t_0 to t_N and $(S_{t_n}/S_{t_{n-1}} - 1)$ is the arithmetic rate of return. The realized variance is usually defined for the logarithm rate of return, but since the spread S_t can be negative, we use the arithmetic rate of return instead.

Capped variance swaps are swap contracts on the realized variance $I(M)$. The present value of the swap is given by

$$S(M, K, C) = \min(I(M), C) - K, \quad (129)$$

where K is the strike of the swap, and the cap C is chosen as $C = 2.5K$ in our example.

In this section, we perform a numerical experiment to test the accuracy of the local correlation calibration procedure. We assume that the underlying assets follow a stochastic correlation model, we simulate sample paths from the model, and compute the payoffs and prices of the spread options and capped variance swaps. The spread option prices are used to calibrate the corresponding local correlation model. Then we compute the prices of capped variance swaps under the calibrated local correlation model, and compare with the prices calculated under the original stochastic correlation model.

We assume that the historical dynamics of the two underlying asset prices are given by the following Bachelier model with stochastic correlation

$$dX_t = \sigma_1 dW_X, \quad dY_t = \sigma_2 dW_Y, \quad d \langle W_X, W_Y \rangle = \rho_t dt. \quad (130)$$

The stochastic correlation is an adapted process taking values between -1 and +1. As proposed in [27] and [34], we use the following Jacobi diffusion to model the bounded stochastic correlation ρ_t :

$$d\rho_t = \kappa(\rho_\infty - \rho_t)dt + \alpha\sqrt{(1 - \rho_t)(1 + \rho_t)}dW, \quad (131)$$

where W is a standard Brownian motion assumed to be independent of W_X and W_Y , and the constants satisfy $\kappa > 0$, $\alpha > 0$, $-1 < \rho_0 < 1$ and $-1 < \rho_\infty < 1$. According to [27], under the constraints

$$\frac{\alpha^2}{\kappa} - 1 < \rho_\infty < 1 - \frac{\alpha^2}{\kappa}, \quad (132)$$

the stochastic correlation process ρ_t does not exit the interval $[-1, +1]$.

Our goal is to find a local correlation model

$$d\tilde{X}_t = \sigma_1 d\tilde{W}_X, \quad d\tilde{Y}_t = \sigma_2 d\tilde{W}_Y, \quad d \langle \tilde{W}_X, \tilde{W}_Y \rangle = \rho(t, \tilde{X}_t, \tilde{Y}_t) dt \quad (133)$$

which generates the same spread option prices as model (130).

The Radon and inverse Radon transforms are widely used in image processing and tomography. Algorithms have already been implemented in existing numerical softwares. We use the PET package [31] of R [32] in our calculation. Our numerical experiment follows the steps:

- First, we discretize the time T into 500 steps, and simulate 10000 paths under model (130). The number of paths is limited to 10000 since we have to record all the simulated values along each path. In order to reduce simulation error, we use the underlying asset price as a control variate when we calculate the spread option price.
- Next, from the spread option price surface, we calibrate the local correlation function $\rho(t, x, y)$ in model (133) using formula (125) with the inverse Radon transform implemented in the PET package. We perform a projection pursuit regression of the calculated local correlation values on the underlying asset prices (x, y) and the time t .
- Last, we simulate 10000 paths under the local correlation model (133) with the fitted local correlation function. We calculate and compare the capped variance swaps prices from the sample paths simulated under models (130) and (133). We also study the variances of the Monte Carlo simulations in order to control their impacts.

The parameters in the experiment are chosen as $T = 2$, $X_0 = 40$, $Y_0 = 30$, $\sigma_1 = 5.0$, $\sigma_2 = 4.0$, $\rho_0 = 0.3$, $\rho_\infty = 0.3$, $\alpha = 1$, $\kappa = 10.6$. Figure 8a shows the simulated capped variance swap price under the original stochastic correlation model (130) for different strikes K and ratios M (the original price). Figure 8b shows the capped variance swaps price under the fitted local correlation model (133) (the fitted price). From these two plots, the fitted price surface has similar curvature as the original price surface.

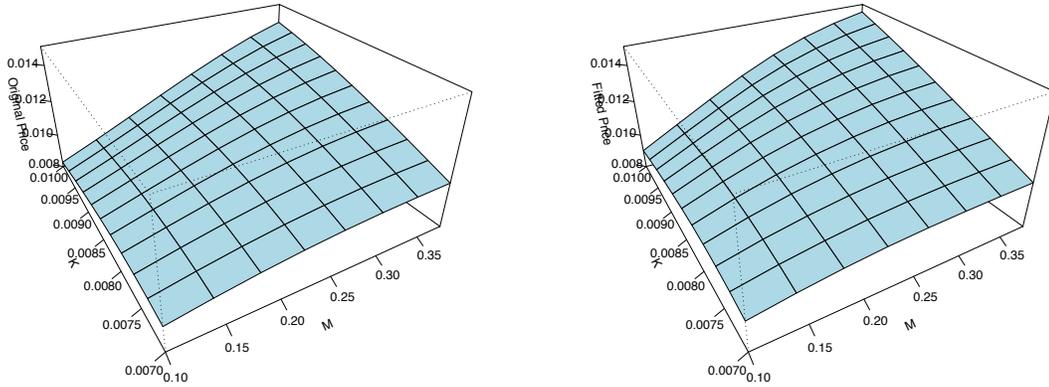


Figure 8: Capped variance swap prices plotted over the ratio M and the strike K , calculated from the original stochastic correlation model (130) (left pane) and from the fitted local correlation model (133) (right pane)

In order to study the errors of the simulations, we plot the standard deviations of the capped variance swap prices calculated from sample paths of models (130) and (133) for different ratios M and strikes K . The results are in Figures 9a and 9b. The pricing difference between the capped variance swap prices calculated from the two models mainly results from fitting errors instead of simulation errors.

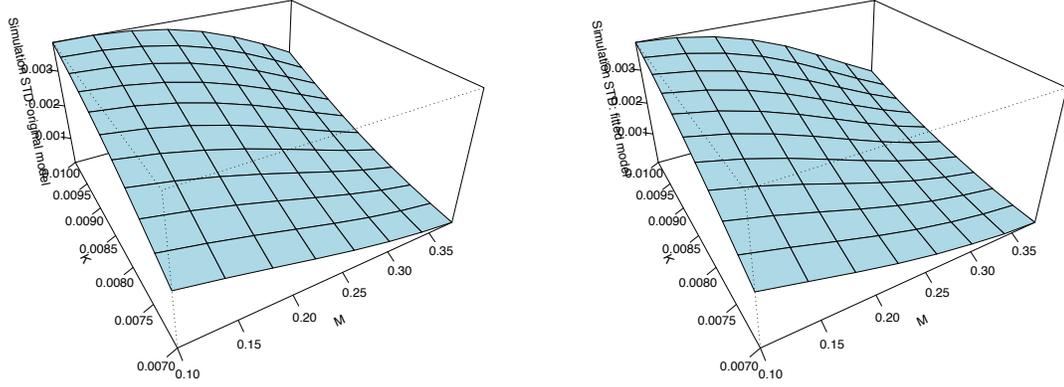


Figure 9: Standard deviation of the capped variance swap prices calculated from the stochastic correlation model (130) plotted over the ratio M and the strike K (left pane) and from the fitted local correlation model (133) (right pane)

In order to quantify the fitting error, we define the following pricing difference percentage and the square difference

$$\text{DIFF} = \frac{\text{Original Price} - \text{Fitted Price}}{\text{Original Price}}, \quad \text{SD} = \sqrt{\sum (\text{DIFF})^2}, \quad (134)$$

where "Original Price" stands for the capped variance swap prices calculated from the original stochastic correlation model (130), and "Fitted Price" stands for the prices computed from the fitted local correlation model (133).

In our experiment, the square difference is calculated as $\text{SD} = 0.0578$. Figure 10 shows the pricing difference percentage between the original price and fitted price. From these results, we can conclude that the local correlation fitting is satisfactory in general. However, the fitted local correlation model underprices the capped variance swap in some areas. The reason is probably due to some smoothing effect during the discretization of the inverse Radon transform when we calibrate the local correlation. The pricing discrepancy can be expected to reduce when the grids of the numerical inverse Radon transform become finer.

5 Conclusion

The main contributions of this paper are the two-dimensional extensions of popular asset price models and option pricing theories. These extensions are designed to provide joint models for the processes underlying spread options in order to generate implied correlation skews.

The first extension is based on a multiscale stochastic volatility model for two dimensional underlying processes. The dependence between the two underlying assets is incorporated in the stochastic volatility factors which are chosen to be the same for the two processes. We derived asymptotic formulas for option values and implied correlations. These formulas have two advantages: 1) Instead of fully

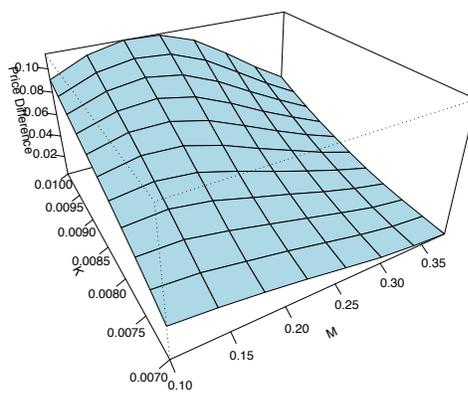


Figure 10: Pricing difference percentage between the capped variance swap prices simulated from the original stochastic correlation model (130) and the fitted local correlation model (133) plotted over the ratio M and the strike K

calibrating the whole model, we only need to calibrate a set of special parameters to obtain option prices. This improves the speed and accuracy of the calibration process. 2) With the asymptotic formula of the implied correlation, we can control the slope and curvature of the implied correlation skew by adjusting the parameters.

The second extension is a local correlation model which assumes the instantaneous correlation to be a deterministic function of time and the underlying prices. The local correlation approach preserves the completeness of the market and low dimensionality of uncertainty. The local correlation model can be viewed as a two-dimensional extension of Dupire's local volatility model. Its numerical implementation relies on the use of the Radon transform. As an example of the versatility of this approach, we computed prices of capped variance swaps on spreads.

Acknowledgments

Both authors were partially supported by NSF DMS-0806591

References

- [1] About PJM. www.pjm.com, 2011.
- [2] Henry Hub. en.wikipedia.org, 2011.
- [3] Radon Transform. en.wikipedia.org, 2011.
- [4] C. Alexander and A. Scourse. Bivariate normal mixture spread option valuation. *Quantitative Finance* 4(6), pages 637–648, 2004.

- [5] F.E. Benth and P.C. Kettler. Dynamic copula models for the spark spread. *Quantitative Finance* 11(3), 2010.
- [6] P. Bjerksund and G. Stensland. Closed form spread option valuation. *NHH Dept. of Finance & Management Science Discussion Paper No. 2006/20*, 2006.
- [7] J.M.S. Blakey and H. Scheule. A nonparametric analysis of implied correlation curve. 2006.
- [8] T. Bollerslev, R. Engle, and J. Woolridge. A capital asset pricing model with time-varying covariances. *Journal of Political Economy* 96, pages 116–131, 1988.
- [9] R. Carmona and V. Durrleman. Pricing and hedging spread options. *SIAM Review* 45, pages 627–685, 2003.
- [10] R. Carmona and V. Durrleman. Pricing and hedging spread options in a log-normal model. *Tech. report, Princeton University*, 2003.
- [11] R. Carmona and M. Ludkovski. Pricing asset scheduling flexibility using optimal switching. *Applied Mathematical Finance* 15(6), pages 405–447, 2008.
- [12] P. Carr and D. Madan. Option valuation using the fast fourier transform. *Journal of Computational Finance* 2(4), pages 61–73, 1999.
- [13] R. Cont and R. Deguest. Extracting implied correlation matrices from index option prices: a statistical approach. *finance-innovation.org*, 2008.
- [14] J. Cotter and F. Longin. Implied correlation from var. *MPRA Paper No. 3506, posted 07*, 2006.
- [15] M. A. H. Dempster and S. S. G. Hong. Spread option valuation and the fast fourier transform. *Technical Report WP 26/2000, The Judge Institute of Management Studies, University of Cambridge*, 2000.
- [16] J. Driessen, P. Maenhout, and G. Vilkov. Option-implied correlations and the price of correlation risk. *Working Paper INSEAD*, 2005.
- [17] B. Dupire. Pricing with a smile. *Risk* 7, pages 32–39, 1994.
- [18] M. Fengler, H. Herwartz, C. Menn, and C. Werner. Explaining the index skew via stochastic correlation models. Technical report, EEA-ESEM, 2008.
- [19] J. D. Fonseca. Option pricing when correlations are stochastic: an analytical framework. *ESILV and Zeliade Systems*, 2006.
- [20] J.P. Fouque, G. Papanicolaou, R. Sircar, and K. Solna. Multiscale stochastic volatility asymptotics. *SIAM J. Multiscale Modeling and Simulation* 2(1), pages 22–42, 2003.
- [21] I. Gyongy. Mimicking the one-dimensional marginal distributions of processes having an ito differential. *Probability Theory and Related Fields* 71, pages 501–516, 1986.
- [22] S. L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies* 6, pages 327–344, 1993.
- [23] P. Jorion. Risk management lessons from long-term capital management. *European Financial Management* 6, pages 277–300, 2000.

- [24] E. Kirk. Correlations in the energy markets. *Managing Energy Price Risk, Risk Publications and Enron, London*, pages 71–78, 1995.
- [25] F. Klebaner. *Introduction to stochastic calculus with applications*. Imperial College Press, 2005.
- [26] J. Ma. Pricing foreign equity options with stochastic correlation and volatility. *Annals of economics and finance 10-2*, pages 303–327, 2009.
- [27] J. Ma. A stochastic correlation model with mean reversion for pricing multi-asset options. *Asia-pacific finan markets 16*, pages 97–109, 2009.
- [28] T. Moskowitz. An analysis of covariance risk and pricing anomalies. *Review of Financial Studies 16*, pages 417–457, 2003.
- [29] B. Oksendal. *Stochastic Differential Equations*. Springer, 2005.
- [30] A.S. Lesniewski P.S. Hagan, D. Kumar and D.E. Woodward. Managing smile risk. *Wilmott Magazine*, pages 84–108, 2002.
- [31] Joern Schulz Peter Toft (pto@imm.dtu.dk) Jesper James Jensen (jjj@oedan.dk) Peter Philipsen (pap@imm.dtu.dk). *PET: Simulation and Reconstruction of PET Images*, 2010. R package version 0.4.9.
- [32] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2010. ISBN 3-900051-07-0.
- [33] O. Reiss and U. Wystup. Computing option price sensitivities using homogeneity and other tricks. *The Journal of Derivatives 9(2)*, pages 41–53, 2001.
- [34] Risk Latte. How do we model stochastic correlations. www.risklatte.com/quantlatte/42.php, 2009.
- [35] R. Roll. The international crash of october 1987. *Financial Analysts Journal 44*, pages 19–35, 1988.
- [36] P.A. Samuelson and R.C. Merton. A complete model of warrant pricing that maximizes utility. *Industrial Management Rvicut IO, 1746*, 1969.
- [37] P.V. Shevchenko. Implied correlation for pricing multi-fx options. *Quantitative Finance Papers 0904.4822*, 2009.
- [38] W.Margrabe. The value of an option to exchange one asset for another. *Journal of Finance 33*, pages 177–86, 1978.
- [39] L. Zalcman. Uniqueness and nonuniqueness for the radon transform. *Bull. London Math. Soc. 14(3)*, pages 241–245, 1982.