MULTI-ASSET EMPIRICAL MARTINGALE PRICE ESTIMATORS FOR FINANCIAL DERIVATIVES

Shih-Feng Huang and Guan-Chih Ciou

National University of Kaohsiung, Taiwan

Supplementary Material

We conducted several simulation scenarios under multivariate Brownian motion, multivariate GARCH models with normal and double exponential innovations, multivariate versions of Merton (1976)'s jump-diffusion models and multivariate versions of Hull and White (1987)'s and Heston (1993)'s stochastic volatility models to investigate the efficiency of the multi-asset EMS and EPMS price estimators. The performance of the asymptotic distribution of the multi-asset EPMS price estimator was also examined under various simulation cases. Detailed proofs were also given in this supplement.

S1 Simulation study

S1.1 Multivariate geometric Brownian motion model

Let the underlying assets $(S_{1,t},\ldots,S_{n,t})$ satisfy the following multivariate geometric Brownian motion model:

$$dS_{i,t}/S_{i,t} = \mu_i dt + \sigma_i dW_{i,t}, \tag{S1.1}$$

where μ_i is the expected return, σ_i is the volatility and $W_{i,t}$ is the Brownian motion of the ith underlying asset, $i = 1, \ldots, n$. Moreover, we assume that $W_{i,t}$, $i = 1, \ldots, n$, are correlated with a time-homogeneous correlation matrix $\boldsymbol{\rho} = (\rho_{i,j})$, where $\rho_{i,j}$ denotes the correlation of $W_{i,t}$ and $W_{j,t}$ for $i, j = 1, \ldots, n$.

The parameters for computing the prices of the two multi-asset options with (S1.1) are set to be $K = 100$, $S_{i,0}/K = 0.97$,1.00 and 1.03 for the cases of T =30, 270 days (1 year = 365 days), $r = 0.1$ (annualized), $\mu_i = 0.15$ (annualized), $\sigma_i = 0.2$ (annualized) for $i = 1, \ldots, n$, $\rho_{j,k} = 0.5$ for $j, k = 1, \ldots, n$ and $j \neq k$, and $m = 10^4$ sample paths, where $n = 3$ for maximum call options and $n = 10$ for geometric average put options. Moreover, we also report the results of a deeper in-the-money (ITM) case for each option when $T = 270$ days, that is, $S_{i,0}/K = 1.10$ for the maximum call options and $S_{i,0}/K = 0.90$ for the geometric average put options. For model (S1.1), the change of measure Λ_T is derived by the multiple dimensional Girsanov Theorem (Shreve, 2004, Theorem 5.4.1) and has the following representation:

$$\Lambda_T = \exp\{-L\Sigma^{-1}M\},$$

where $L = (\mu_1 - r, \ldots, \mu_n - r)$, $M = (0.5(\mu_1 - r)T + \sigma_1 W_{1,T}, \ldots, 0.5(\mu_n - r)T + \sigma_n W_{n,T})^\top$ and $\Sigma = (\sigma_{ij})$ with $\sigma_{ij} = \rho_{ij}\sigma_i\sigma_j$, for $i, j = 1, \ldots, n$.

Under model (S1.1) the theoretical values of the maximum call and geometric average put options can be obtained from their closed-form representations in equation (16) of Huang and Yu (2012) and in Example 3.2 of Huang and Guo (2009), respectively. Table 1 presents the MSE ratios MR^{Q} and MR^{P}, which indicate that the EMS and EPMS perform better than the MCQ and MCP under the Q and P measures, respectively, especially for ITM options. For financial risk management and the construction of portfolios, we also care about how to price a vector of options whose payoffs are correlated because they depend on common underlying assets. In Table 1, for each pair of (K, T) and fixed model parameters, the geometric average put option and the maximal call option can be viewed as a 2-dimensional vector of options. The price estimates of these two options are obtained simultaneously by using the same generated underlying asset prices. According to the numerical results given in Table 1, the proposed price estimators can also help to improve the efficiency when pricing a vector of options.

To evaluate the performance of Theorem 2, we compute the coverage rates under various scenarios on the basis of 1,000 replications, where the coverage rate is defined by the percentage of times in which the theoretical value falls into the confidence interval given in (4.1). Simulation results

Table 1: The MSE ratios MR^{Q} and MR^{P}, and the coverage rates of the multi-asset EPMS price estimator by using the asymptotic distribution given in Theorem 2 for pricing the maximum call options and geometric average put options under the multivariate geometric Brownian motion model defined in (S1.1).

Maximum Call Option ($n = 3$)	$T = 30$ days			$T = 270$ days			
S_0/K	0.97	1.00	1.03	0.97	1.00	1.03	1.10
MR^{Q}	3.17	5.28	6.36	5.49	6.06	6.32	6.77
MR^{P}	2.69	4.21	4.82	2.85	3.06	2.99	2.83
25% cov. rate	0.251	0.266	0.239	0.241	0.247	0.259	0.259
50% cov. rate	0.476	0.512	0.519	0.497	0.510	0.494	0.486
75% cov. rate	0.746	0.762	0.774	0.757	0.752	0.738	0.759
95% cov. rate	0.945	0.947	0.951	0.956	0.950	0.953	0.954

Geometric Put Option ($n = 10$)	$T = 30$ days			$T = 270$ days			
S_0/K	0.97	1.00	1.03	0.90	0.97	1.00	1.03
MR^{Q}	6.65	5.16	5.26	4.87	5.28	5.42	4.86
MR^{P}	2.93	3.85	3.62	2.40	2.21	2.05	1.78
25% cov. rate	0.234	0.245	0.275	0.252	0.234	0.254	0.254
50% cov. rate	0.487	0.502	0.533	0.504	0.490	0.483	0.483
75% cov. rate	0.761	0.749	0.740	0.760	0.752	0.731	0.731
95% cov. rate	0.957	0.943	0.947	0.957	0.950	0.951	0.951

presented in Table 1 indicate that the coverage rates of the multi-asset EPMS price estimator are close to their nominal values. As a result, the asymptotic distribution given in Theorem 2 is a satisfactory approximation to the finite-sample distribution of the EPMS price estimator.

As mentioned in Duan and Simonato (1998) and Huang and Tu (2014) the EMS and EPMS can be combined with two popular variance reduc-

Table 2: The price, MR^{Q}, $\mathrm{MR}^{\mathrm{Q}*}$, $P^{(MCQ)}$, $P^{(EMS)}$ and $P^{(GHS)}$ for Asian options in the Black-Scholes model.

		T = 1 year				T = 3 years			
σ		K=40	45	50	55	40	45	50	55
0.1	Price	10.80	6.06	1.92	0.20	12.08	7.83	4.06	1.55
	MR^{Q}	375.7	291.7	6.5	1.3	74.6	32.9	14.3	4.2
	$\mathrm{MR}^{\mathrm{Q}*}$	37.6	11.1	5.9	20.3	19.3	8.6	6.6	9.4
	$P^{(MCQ)}$	0.95	0.94	0.62	0.12	0.86	0.83	0.65	0.35
	$P^{(EMS)}$	0.95	0.94	0.63	0.12	0.86	0.83	0.65	0.35
	$P^{(GHS)}$	0.95	0.95	0.85	0.69	0.86	0.85	0.80	0.71
0.3	Price	11.09	7.15	4.17	2.21	13.08	9.94	7.37	5.37
	MR^{Q}	67.0	6.9	5.2	3.2	18.8	14.4	6.8	6.3
	$\mathrm{MR}^{\mathrm{Q}*}$	12.0	4.3	8.0	10.3	7.8	8.8	9.0	10.0
	$P^{(MCQ)}$	0.86	0.70	0.49	0.30	0.68	0.57	0.46	0.35
	$P^{(EMS)}$	0.86	0.70	0.49	0.30	0.68	0.57	0.46	0.35
	$P^{(GHS)}$	0.93	0.88	0.82	0.76	0.82	0.79	0.75	0.72

tion techniques, antithetic and control variate simulations. For multi-asset derivatives, one can also combine the proposed method with antithetic and control variate techniques to improve the efficiency of the MC estimator. Another popular and frequently used variance reduction technique is importance sampling. Glasserman, Heidelberger and Shahabuddin (1999) proposed an importance sampling method based on a change of drift for pricing path-dependent options. In the following, we consider the pricing of an arithmetic Asian option on a single underlying asset under standard Black-Scholes assumptions Black and Scholes (1973), which is a special case of model (S1.1) with $n = 1$. The risk-neutral model is described by model (2.2). The payoff of an arithmetic Asian option with strike K is given by $f(S_{1,1}, \ldots, S_{1,T}) = \max(M^{-1} \sum_{i=1}^{M} S_{1,t_i} - K, 0)$, where $0 = t_0 < t_1 < \cdots < t_M$ and $t_i - t_{i-1}$, for $i = 1, \ldots, M$, are equidistance. We compare the estimation performance of Glasserman, Heidelberger and Shahabuddin (1999)'s importance sampling price estimator, denoted by GHS, with the EMS. Table 2 presents MR^{Q} and $\mathrm{MR}^{\mathrm{Q}*}$, which denotes the MSE ratio of the MCQ to the GHS, with $S_0 = 50$, $K = 40, 45, 50, 55$, $r = 0.05$, $\sigma = 0.1, 0.3$, $T = 1, 3$, $M = 16$ and 10^4 runs. Table 2 also reports the probabilities of $\mathrm{ITM}_T = \{S_{1,t_i}, i = 1, \ldots, M : M^{-1} \sum_{i=1}^{M} S_{1,t_i} > K\}$, where S_{1,t_i} are separately generated from the MCQ, EMS and GHS methods, and the corresponding probabilities are denoted by $P^{(MCQ)}$, $P^{(EMS)}$ and $P^{(GHS)}$, respectively. As shown in Table 2, $P^{(MCQ)}$ and $P^{(EMS)}$ are very close and decrease as K increases for Asian call options. By using the technique of changing drift, $P^{(GHS)}$ are greater than $P^{(MCQ)}$ and $P^{(EMS)}$ in the considered scenarios, especially when K is large. This property of the GHS method helps to generate more samples with positive payoffs than the MC and EMS methods. From the numerical results, this property also leads to more remarkable variance reduction than the EMS for out-of-the-money (OTM) options. On the other hand, when the option is ITM, the absolute difference between $P^{(GHS)}$ and $P^{(EMS)}$ decreases as K decreases. From this perspective, the superiority of generating samples with positive payoffs by the GHS over the EMS is insignificant for ITM options. Numerical results in Table 2 indicate that the magnitude of reducing variation by the EMS scheme is greater than the GHS for ITM Asian call options.

S1.2 Multivariate GARCH model

Let the daily prices of the underlying assets, $(S_{1,t}, \ldots, S_{n,t})$, satisfy the following multivariate GARCH(1,1) model:

$$\begin{cases} \log S_{i,t} &= \log S_{i,t-1} + r_d + \lambda_i \sigma_{i,t} - 0.5\sigma_{i,t}^2 + \sigma_{i,t}\epsilon_{i,t}, \\ \sigma_{i,t}^2 &= \beta_{i0} + \beta_{i1}\sigma_{i,t-1}^2 + \beta_{i2}\sigma_{i,t-1}^2\epsilon_{i,t-1}^2, \end{cases} \tag{S1.2}$$

for $i = 1, \ldots, n$, where r_d denotes the daily risk-free interest rate, $(\epsilon_{1,t}, \ldots, \epsilon_{n,t})$ are assumed to follow a multivariate normal distribution with mean $\mathbf{0} \in \mathbb{R}^n$ and a covariance matrix Σ, denoted by $N(\mathbf{0}, \Sigma)$. In particular, $\Sigma = (\rho_{j,k})$ is positive definite and $\rho_{j,k} = 1$ if $j = k$, where $j, k = 1, \ldots, n$. Moreover, β_{i0}, β_{i1}, β_{i2} are nonnegative, $\beta_{i1} + \beta_{i2} < 1$, and λ_i represents the unit risk premium for $i = 1, \ldots, n$. We denote model (S1.2) with normal innovations by GARCH-N. In particular, if $n = 1$, then model (S1.2) reduces to the GARCH model considered in Duan (1995).

The parameters for computing the prices of the two multi-asset options with (S1.2) are set to be $\sigma_{i,1}^2 = \beta_{i0}(1 - \beta_{i1} - \beta_{i2})^{-1}$, $\beta_{i0} = 10^{-5}$, $\beta_{i1} =$

0.7, $\beta_{i2} = 0.2$, $\lambda_i = 0.01$ for $i = 1, \ldots, n$, $\rho_{j,k} = 0.5$ for $j \neq k$, where $j, k = 1, \ldots, n$, and others are the same as in Table 1. Moreover, let $R_{i,t} = \log(S_{i,t}/S_{i,t-1})$ and $\mu_{i,t} = r_d + \lambda_i \sigma_{i,t} - 0.5\sigma_{i,t}^2$ for $i = 1, \ldots, n$ and $t = 1, \ldots, T$. For model (S1.2), the change of measure Λ_t is derived by the multivariate Esscher transform (Kijima, 2006) and has the following representation:

$$\Lambda_t = \prod_{k=1}^{t} \exp(\ -L_k \Sigma_k^{-1} M_k \ ), \tag{S1.3}$$

where $L_k = (\lambda_1 \sigma_{1,k}, \ldots, \lambda_n \sigma_{n,k})$, $M_k = (0.5\lambda_1\sigma_{1,k} + R_{1,k} - \mu_{1,k}, \ldots, 0.5\lambda_n\sigma_{n,k} + R_{n,k} - \mu_{n,k})^\top$ and $\Sigma_k = (\sigma_{ijk})$ with $\sigma_{ijk} = \rho_{i,j}\sigma_{i,k}\sigma_{j,k}$, for $i, j = 1, \ldots, n$. Details of the derivation of (S1.3) are presented in Section S2.4.

Table 3 reports the MSE ratios MR^{Q} and MR^{P}. The coverage rates obtained from the asymptotic distribution of the multi-asset EPMS price estimator are also presented. The phenomena shown in Table 3 for the multivariate GARCH-N model are similar to those in Table 1 for the multivariate geometric Brownian motion model. Under the multivariate GARCH-N framework, both the EMS and EPMS price estimators are more efficient than their MC counterparts. In addition, the asymptotic distribution given in Theorem 2 also provides a good approximation in our scenarios.

Next, we demonstrate a particular example to show that the explicit representation of a risk-neutral GARCH model is not convenient to obtain and thus the EMS is not feasible. Suppose that we use the Esscher trans-

Table 3: The MSE ratios MR^{Q} and MR^{P}, and the coverage rates of the multi-asset EPMS price estimator by using the asymptotic distribution given in Theorem 2 for the maximum call options and geometric average put options under the multivariate GARCH-N model.

Maximum Call Option ($n = 3$)	$T = 30$ days			$T = 270$ days			
S_0/K	0.97	1.00	1.03	0.97	1.00	1.03	1.10
MR^{Q}	2.72	3.81	6.06	5.03	5.78	5.88	6.49
MR^{P}	2.43	3.31	5.00	3.06	3.44	3.40	3.45
25% cov. rate	0.266	0.245	0.246	0.237	0.251	0.271	0.237
50% cov. rate	0.525	0.489	0.500	0.502	0.483	0.531	0.496
75% cov. rate	0.761	0.747	0.739	0.767	0.740	0.750	0.754
95% cov. rate	0.953	0.957	0.948	0.952	0.944	0.957	0.955

Geometric Put Option ($n = 10$)	$T = 30$ days			$T = 270$ days			
S_0/K	0.97	1.00	1.03	0.90	0.97	1.00	1.03
MR^{Q}	7.93	2.95	1.61	4.69	2.36	1.97	1.72
MR^{P}	10.03	3.38	1.71	8.49	3.44	2.56	2.18
25% cov. rate	0.255	0.257	0.241	0.228	0.270	0.267	0.255
50% cov. rate	0.500	0.507	0.510	0.511	0.512	0.502	0.504
75% cov. rate	0.741	0.756	0.745	0.737	0.750	0.743	0.766
95% cov. rate	0.953	0.959	0.950	0.949	0.955	0.946	0.953

form (Gerber and Shiu, 1994) to define the change of measure process for the GARCH model in 1-dimensional case. Let $g_t(\cdot)$ be the conditional density function of the log return $R_t = \log(S_t/S_{t-1})$ given $\mathcal{F}_{t-1}$ and $M_t(\delta_t) = E_{t-1}\{\exp(\delta_t R_t)\}$ be the corresponding conditional moment generating function (mgf). The Esscher transform aims to find a risk-neutral density from the family $g_t^*(x) = \exp(\delta_t x)g_t(x)/M_t(\delta_t)$ by choosing a δ_t such that the dis-

Table 4: The MSE ratios MR^{P} for pricing the maximum call options and geometric average put options and the coverage rates of the multi-asset EPMS price estimator by using the asymptotic distribution given in Theorem 2 under the multivariate GARCH-DE model.

Maximum Call Option ($n = 3$)	$T = 30$ days			$T = 270$ days			
S_0/K	0.97	1.00	1.03	0.97	1.00	1.03	1.10
MR^{P}	1.81	1.94	1.85	1.65	1.75	1.82	1.59
25% cov. rate	0.255	0.252	0.253	0.255	0.273	0.257	0.253
50% cov. rate	0.495	0.505	0.507	0.511	0.508	0.533	0.492
75% cov. rate	0.748	0.760	0.743	0.746	0.764	0.768	0.755
95% cov. rate	0.954	0.944	0.947	0.955	0.956	0.956	0.947

Geometric Put Option ($n = 10$)	$T = 30$ days			$T = 270$ days			
S_0/K	0.97	1.00	1.03	0.90	0.97	1.00	1.03
MR^{P}	39.64	2.95	1.09	27.11	2.86	1.77	1.24
25% cov. rate	0.227	0.243	0.220	0.265	0.222	0.228	0.254
50% cov. rate	0.475	0.504	0.488	0.511	0.476	0.482	0.496
75% cov. rate	0.727	0.753	0.738	0.753	0.739	0.744	0.742
95% cov. rate	0.943	0.944	0.943	0.955	0.950	0.947	0.943

counted stock prices being a martingale, that is, $E^*_{t-1}\{\exp(-r_d)S_t\} = S_{t-1}$ or $M^*_t(1) = \exp(r_d)$, where $E^*_{t-1}(\cdot)$ denotes to compute the conditional expectation under g^*_t and $M^*_t(z) = M_t(\delta_t + z)/M_t(\delta_t)$ is the conditional mgf of R_t under g^*_t. In the GARCH-N model, we have $M_t(\delta_t) = \exp\{(r_d + \lambda\sigma_t - 0.5\sigma_t^2)\delta_t + 0.5\sigma_t^2\delta_t^2\}$. By choosing $\delta_t = -\lambda/\sigma_t$ and after some straightforward calculation, we have $M^*_t(z) = \exp\{(r_d - 0.5\sigma_t^2)z + 0.5\sigma_t^2 z^2\}$, which is the mgf of a normal random variable with mean $r_d - 0.5\sigma_t^2$ and variance σ_t^2. In the other words, the risk-neutral density obtained by the Esscher transform is still normally distributed and thus the corresponding risk-neutral GARCH-N model can be represented accordingly. However, if we replace the normal assumption of the innovation $\epsilon_{i,t}$ in model (S1.2) by a standardized double exponential distribution, which has fatter tails than $N(0,1)$, and denoted this model by GARCH-DE, the task becomes difficult. In the GARCH-DE model, we have $M_t(\delta_t) = \exp(\mu_t \delta_t)/(1 - 0.5\sigma_t^2\delta_t^2)$ for $|\delta_t| < 2^{1/2}\sigma_t^{-1}$, where $\mu_t = r_d + \lambda\sigma_t - 0.5\sigma_t^2$. By choosing $\delta_t = (-\sigma_t + \{a_t\sigma_t^2 + 2(a_t - 1)^2\}^{1/2})/\{\sigma_t(1 - a_t)\}$, where $a_t = \exp(\lambda\sigma_t - 0.5\sigma_t^2)$, we have $E^*_{t-1}\{\exp(-r_d)S_t\} = S_{t-1}$ and $M_t^*(z) = \exp(z\mu_t)(1 - 0.5\delta_t^2\sigma_t^2)/\{1 - 0.5(\delta_t + z)^2\sigma_t^2\}$, which is not an mgf of a standard double-exponential random variable anymore. In addition, it is not convenient to derive the distribution from this $M_t^*(z)$. Therefore, the risk-neutral model for the GARCH-DE case is not convenient to obtain.

For the multivariate GARCH-DE case, we assume that the processes of the underlying assets are independent. Hence, the change of measure process derived by the Esscher transform for the GARCH-DE model can be obtained directly by Proposition 3.3 of Huang (2014):

$$\Lambda_t = \prod_{k=1}^{t}\prod_{i=1}^{n}[1 - 0.5(\delta^*_{i,k}\sigma_{i,k})^2]\exp\{\delta^*_{i,k}(R_{i,k} - r_d - \lambda_i\sigma_{i,k} + 0.5\sigma^2_{i,k})\},$$

where $\delta^*_{i,k} = (-\sigma_{i,k} + \{a_{i,k}\sigma^2_{i,k} + 2(a_{i,k} - 1)^2\}^{1/2})/\{\sigma_{i,k}(1 - a_{i,k})\}$ and $a_{i,k} = \exp(\lambda_i\sigma_{i,k} - 0.5\sigma^2_{i,k})$. Since the EMS procedure is not feasible for the

GARCH-DE model under the framework of Esscher transform, only the results of the EPMS and MCP price estimators are presented in Table 4. Numerical results indicate that the EPMS still has better performance than the MCP and the asymptotic distribution of the EPMS still provides a reliable approximation in the GARCH-DE model.

S1.3 Multivariate jump-diffusion model

In this section, we consider applying the proposed price estimators to multivariate versions of Merton (1976)'s jump-diffusion models for investigating the effects of the jump frequency and jump size on the pricing performance. Let the underlying assets $(S_{1,t}, \ldots, S_{n,t})$ satisfy the following multivariate jump-diffusion model under the P measure:

$$dS_{i,t}/S_{i,t} = (\mu_i - \lambda_i k)dt + \sigma_i dW_{i,t} + (Y_{i,t} - 1)dN_{i,t}, \tag{S1.4}$$

where μ_i, σ_i and $W_{i,t}$ are defined the same as in model (S1.1), $N_{i,t}$, $i = 1, \ldots, n$, are independent Poisson processes with intensity λ_i, $i = 1, \ldots, n$, respectively, $\log Y_{i,t}$, $i = 1, \ldots, n$, are i.i.d. $N(\alpha, \theta^2)$, $k \equiv E(Y_{i,t} - 1) = \exp(\alpha + 0.5\theta^2) - 1$, and $W_{i,t}$, $N_{i,t}$ and $Y_{i,t}$ are independent for each $i = 1, \ldots, n$. For computing the price of a geometric average put option under model (S1.4), we first have the following diffusion process for $\overline{S}_T =$

Table 5: The MSE ratios MR^{Q} and MR^{P} for pricing maximum call options and geometric average put options, and the coverage rates of the multi-asset EPMS price estimator by using the asymptotic distribution given in Theorem 2 under the multivariate jump-diffusion model (S1.4).

Maximum Call Option ($n = 3$)		$T = 30$ days			$T = 270$ days			
	S_0/K	0.97	1.00	1.03	0.97	1.00	1.03	1.10
$(\lambda, \theta) = (1{,}0.05)$	MR^{Q}	2.42	3.55	4.63	4.00	5.17	4.95	5.13
	MR^{P}	2.15	3.07	3.49	2.15	2.82	2.43	2.59
	25% cov. rate	0.249	0.238	0.276	0.234	0.250	0.248	0.233
	50% cov. rate	0.507	0.488	0.529	0.501	0.512	0.478	0.465
	75% cov. rate	0.761	0.748	0.779	0.763	0.755	0.749	0.727
	95% cov. rate	0.943	0.938	0.952	0.946	0.946	0.949	0.959
$(\lambda, \theta) = (1{,}0.2)$	MR^{Q}	1.74	2.22	2.64	3.07	3.30	3.52	3.91
	MR^{P}	1.72	2.10	2.36	2.44	2.49	2.76	2.86
	25% cov. rate	0.247	0.252	0.260	0.238	0.237	0.256	0.255
	50% cov. rate	0.493	0.513	0.508	0.507	0.485	0.490	0.493
	75% cov. rate	0.746	0.757	0.773	0.767	0.736	0.738	0.742
	95% cov. rate	0.951	0.941	0.956	0.950	0.946	0.940	0.949
$(\lambda, \theta) = (3{,}0.05)$	MR^{Q}	2.54	2.97	3.88	3.56	4.07	4.34	4.54
	MR^{P}	2.27	2.63	3.04	2.30	2.31	2.36	2.74
	25% cov. rate	0.261	0.242	0.251	0.260	0.242	0.248	0.244
	50% cov. rate	0.498	0.488	0.493	0.507	0.486	0.507	0.513
	75% cov. rate	0.751	0.759	0.740	0.758	0.732	0.747	0.761
	95% cov. rate	0.953	0.956	0.947	0.941	0.941	0.956	0.951
$(\lambda, \theta) = (3{,}0.2)$	MR^{Q}	1.83	2.13	2.64	3.16	3.30	3.47	3.94
	MR^{P}	1.81	2.17	2.45	2.90	3.09	3.14	3.29
	25% cov. rate	0.254	0.245	0.257	0.258	0.249	0.248	0.264
	50% cov. rate	0.503	0.503	0.535	0.504	0.488	0.478	0.506
	75% cov. rate	0.767	0.735	0.765	0.741	0.751	0.773	0.739
	95% cov. rate	0.947	0.957	0.953	0.956	0.953	0.963	0.952

Geometric Put Option ($n = 10$)		$T = 30$ days			$T = 270$ days			
	S_0/K	0.97	1.00	1.03	0.90	0.97	1.00	1.03
$(\lambda, \theta) = (1,0.05)$	MR^{Q}	8.70	3.65	1.76	7.43	3.55	2.86	2.15
	MR^{P}	12.95	4.70	2.06	20.06	7.23	5.06	3.47
	25% cov. rate	0.264	0.247	0.223	0.242	0.250	0.252	0.245
	50% cov. rate	0.504	0.510	0.478	0.519	0.505	0.498	0.503
	75% cov. rate	0.766	0.774	0.743	0.764	0.751	0.739	0.763
	95% cov. rate	0.952	0.958	0.948	0.953	0.956	0.953	0.948
$(\lambda, \theta) = (1,0.2)$	MR^{Q}	7.12	2.61	1.68	6.54	3.51	3.03	2.22
	MR^{P}	10.00	5.55	2.12	17.53	7.33	5.15	3.84
	25% cov. rate	0.258	0.230	0.243	0.257	0.260	0.254	0.243
	50% cov. rate	0.480	0.488	0.498	0.500	0.508	0.497	0.502
	75% cov. rate	0.749	0.745	0.755	0.724	0.752	0.749	0.747
	95% cov. rate	0.940	0.951	0.942	0.952	0.952	0.945	0.954
$(\lambda, \theta) = (3,0.05)$	MR^{Q}	8.48	3.47	1.78	7.24	3.55	2.64	2.30
	MR^{P}	11.95	4.59	2.23	21.09	7.10	4.67	3.75
	25% cov. rate	0.254	0.274	0.255	0.226	0.243	0.273	0.254
	50% cov. rate	0.520	0.510	0.529	0.485	0.470	0.524	0.491
	75% cov. rate	0.762	0.764	0.762	0.731	0.739	0.759	0.742
	95% cov. rate	0.944	0.956	0.953	0.952	0.950	0.956	0.944
$(\lambda, \theta) = (3,0.2)$	MR^{Q}	5.99	3.02	1.92	5.34	3.42	3.12	2.47
	MR^{P}	7.36	3.85	2.10	17.83	7.30	5.90	4.99
	25% cov. rate	0.231	0.243	0.245	0.255	0.227	0.240	0.251
	50% cov. rate	0.505	0.506	0.507	0.527	0.496	0.503	0.479
	75% cov. rate	0.748	0.746	0.758	0.759	0.747	0.738	0.745
	95% cov. rate	0.937	0.956	0.939	0.966	0.954	0.945	0.961

$(S_{1,T} \cdots S_{n,T})^{1/n}$ under the P measure:

$$d \log \overline{S}_t = \Big(\overline{\mu} - \frac{1}{2n} \sum_{i=1}^{n} \sigma_i^2 - \overline{\lambda} k\Big) dt + \tilde{\sigma} d\widetilde{W}_t + \frac{1}{n} \log Y_t d\widetilde{N}_t, \tag{S1.5}$$

where $\overline{\mu} = n^{-1} \sum_{i=1}^{n} \mu_i$, $\overline{\lambda} = n^{-1} \sum_{i=1}^{n} \lambda_i$, $\tilde{\sigma} = n^{-2} \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{i,j} \sigma_i \sigma_j$, $\widetilde{W}_t$ is a Brownian motion, $\widetilde{N}_t$ is a Poisson process with intensity rate $\tilde{\lambda} =$

$\lambda_1 + \cdots + \lambda_n$, $\log Y_t$ is $N(\alpha, \theta_2)$, and $\widetilde{W}_t$, $\widetilde{N}_t$ and Y_t are independent. The risk-neutral counterpart of model (S1.5) is

$$d \log \overline{S}_t = \Big(r - \frac{1}{2n}\sum_{i=1}^{n} \sigma_i^2 - \overline{\lambda} k\Big) dt + \tilde{\sigma} d\widetilde{W}_t^* + \frac{1}{n} \log Y_t d\widetilde{N}_t, \qquad \text{(S1.6)}$$

where $\widetilde{W}_t^* = \widetilde{W}_t + (\overline{\mu} - r)t$ is a Brownian motion under the Q measure derived from the change of measure process $\Lambda_t = E_t(\Lambda_T)$, in which Λ_T is defined the same as in model (S1.1). In addition, Λ_t is still independent of $\widetilde{N}_t$ and Y_t, which have the same distributions under the Q measure as in the P measure. By using model (S1.6), the price of a geometric average put option has the following closed-form representation:

$$\begin{aligned} V_0 &= e^{-(r-\tilde{r})T} \sum_{i=0}^{\infty} \frac{e^{-\tilde{\lambda}^* T} (\tilde{\lambda}^* T)^i}{i!} \\ &\quad \text{BSPut}\Big(\overline{S}_0, K, r - \tilde{\lambda} k + \frac{i \log(1+k)}{T}, \sqrt{\tilde{\sigma}^2 + \frac{i\theta^2}{T}}, T\Big), \end{aligned} \qquad \text{(S1.7)}$$

where $\tilde{\lambda}^* = \tilde{\lambda}(1+k)$, $\tilde{r} = r - 0.5 \sum_{i=1}^{n} \sigma_i^2/n + 0.5\tilde{\sigma}^2 - \overline{\lambda} k + \tilde{\lambda}\tilde{k}$ with $\tilde{k} = \exp(\alpha/d + 0.5\theta^2/d^2) - 1$, and $\text{BSPut}(S_0, K, r, \sigma, T)$ denotes the value of a European put option under the Black-Scholes model with initial stock price S_0, risk-free interest rate r and instantaneous volatility σ.

The parameters for computing the prices of the two multi-asset options with (S1.4) are set to be the same as in Table 1 except for $r = 0.04$ (annualized), $\mu_i = 0.10$ (annualized), $\lambda_i = 1, 3$ for $i = 1, \ldots, n$, $\alpha = 0.06$ and $\theta = 0.05, 0.2$. The parameters in the jump part are set the same as in Table

1 of Zhu, Ye and Zhou (2015). Table 5 presents the MSE ratios MR^{Q} and MR^{P} for geometric average put options, where the true values are obtained by (S1.7). Since the derivation of the pricing formula of maximum call options under model (S1.4) is beyond the scope of this study, the expected maximum call option prices are replaced by using the MC method with 10^5 simulations for computing MR^{Q} and MR^{P} in Table 5. Numerical results in Table 5 indicate that the EMS and EPMS price estimators are capable of improving the efficiency of the MC method. The improvement is significant when applying the EMS to maximum call options with less jump frequency and smaller variance of the jump size, and applying the EMS and EPMS to ITM geometric put options.

S1.4 Multivariate stochastic volatility model

In this section, we consider applying the proposed price estimators to multivariate versions of Hull and White (1987)'s and Heston (1993)'s stochastic volatility models for investigating the effects of stochastic volatilities on the pricing performance. Let the underlying assets $(S_{1,t}, \ldots, S_{n,t})$ satisfy the following multivariate stochastic volatility model under the P measure (Hull and White, 1987):

$$\begin{cases} dS_{i,t}/S_{i,t} &= \mu_i dt + \sqrt{V_{i,t}} dW_{i,t}^{(1)} \\ dV_{i,t}/V_{i,t} &= \kappa_i(\sigma_i^* - \sqrt{V_{i,t}})dt + \sigma_i dW_{i,t}^{(2)} \end{cases} \tag{S1.8}$$

Table 6: The MSE ratios MR^{Q} and MR^{P} for pricing maximum call options and geometric average put options and the coverage rates of the multi-asset EPMS price estimator by using the asymptotic distribution given in Theorem 2 under the multivariate versions of Hull and White (1987)'s stochastic volatility model (S1.8).

Maximum Call Option ($n = 3$)		$T = 30$ days			$T = 270$ days			
	S_0/K	0.97	1.00	1.03	0.97	1.00	1.03	1.10
$(\kappa, \sigma) = (10,0.1)$	MR^{Q}	2.19	2.30	2.56	2.66	2.93	2.23	3.00
	MR^{P}	1.92	2.01	2.10	1.91	2.04	2.03	2.08
	25% cov. rate	0.249	0.258	0.243	0.249	0.254	0.259	0.261
	50% cov. rate	0.528	0.508	0.501	0.531	0.499	0.494	0.497
	75% cov. rate	0.761	0.737	0.748	0.775	0.758	0.747	0.753
	95% cov. rate	0.944	0.942	0.950	0.956	0.946	0.946	0.946
$(\kappa, \sigma) = (5,0.1)$	MR^{Q}	2.17	2.44	2.89	2.74	2.69	3.06	3.16
	MR^{P}	1.92	2.10	2.32	2.05	1.93	2.17	2.11
	25% cov. rate	0.238	0.286	0.257	0.245	0.261	0.264	0.245
	50% cov. rate	0.495	0.526	0.527	0.509	0.485	0.531	0.514
	75% cov. rate	0.763	0.770	0.769	0.749	0.752	0.767	0.751
	95% cov. rate	0.952	0.958	0.966	0.946	0.945	0.959	0.952
$(\kappa, \sigma) = (10,1)$	MR^{Q}	2.17	2.25	2.29	2.10	2.06	2.17	2.44
	MR^{P}	1.81	1.86	1.82	1.51	1.42	1.62	1.50
	25% cov. rate	0.266	0.229	0.257	0.261	0.254	0.231	0.252
	50% cov. rate	0.513	0.492	0.523	0.517	0.483	0.483	0.504
	75% cov. rate	0.759	0.746	0.761	0.726	0.731	0.745	0.750
	95% cov. rate	0.957	0.943	0.949	0.940	0.949	0.955	0.955
$(\kappa, \sigma) = (5,1)$	MR^{Q}	2.02	2.33	2.66	1.97	2.11	2.27	2.25
	MR^{P}	1.70	1.94	2.09	1.17	1.36	1.36	1.46
	25% cov. rate	0.248	0.231	0.268	0.268	0.262	0.236	0.253
	50% cov. rate	0.507	0.494	0.506	0.509	0.510	0.496	0.484
	75% cov. rate	0.751	0.731	0.763	0.758	0.772	0.753	0.747
	95% cov. rate	0.962	0.949	0.961	0.959	0.964	0.957	0.952

Geometric Put Option ($n = 10$)		$T = 30$ days			$T = 270$ days			
	S_0/K	0.97	1.00	1.03	0.90	0.97	1.00	1.03
$(\kappa, \sigma) = (10{,}0.1)$	MR^{Q}	13.92	3.11	1.29	4.56	2.42	1.86	1.54
	MR^{P}	25.89	4.09	1.43	10.46	5.09	3.09	2.18
	25% cov. rate	0.268	0.273	0.258	0.267	0.244	0.249	0.253
	50% cov. rate	0.496	0.531	0.512	0.510	0.478	0.503	0.505
	75% cov. rate	0.748	0.769	0.741	0.754	0.738	0.753	0.768
	95% cov. rate	0.953	0.956	0.942	0.944	0.953	0.948	0.954
$(\kappa, \sigma) = (5{,}0.1)$	MR^{Q}	12.89	3.07	1.35	4.59	2.46	1.93	1.48
	MR^{P}	24.21	4.00	1.52	15.12	4.86	2.98	2.06
	25% cov. rate	0.252	0.252	0.248	0.269	0.256	0.247	0.252
	50% cov. rate	0.488	0.508	0.494	0.500	0.496	0.498	0.526
	75% cov. rate	0.742	0.742	0.747	0.739	0.758	0.747	0.760
	95% cov. rate	0.941	0.954	0.948	0.953	0.957	0.947	0.952
$(\kappa, \sigma) = (10{,}1)$	MR^{Q}	14.71	3.17	1.37	5.13	2.29	1.87	1.49
	MR^{P}	23.34	4.08	1.48	7.50	3.61	2.69	1.92
	25% cov. rate	0.246	0.233	0.244	0.231	0.246	0.262	0.237
	50% cov. rate	0.514	0.500	0.480	0.475	0.504	0.510	0.487
	75% cov. rate	0.773	0.734	0.744	0.750	0.764	0.749	0.736
	95% cov. rate	0.944	0.954	0.952	0.957	0.943	0.950	0.937
$(\kappa, \sigma) = (5{,}1)$	MR^{Q}	14.76	2.92	1.35	4.44	2.26	2.02	1.46
	MR^{P}	26.17	3.80	1.45	4.59	2.82	2.25	1.76
	25% cov. rate	0.261	0.247	0.236	0.235	0.228	0.254	0.246
	50% cov. rate	0.487	0.493	0.499	0.497	0.479	0.496	0.514
	75% cov. rate	0.733	0.728	0.756	0.776	0.746	0.743	0.752
	95% cov. rate	0.943	0.947	0.943	0.953	0.957	0.941	0.950

where $\{S_{i,t}, t \geq 0\}$ and $\{V_{i,t}, t \geq 0\}$ are the price and volatility processes, respectively, for the ith underlying asset, $W_{i,t}^{(1)}$ and $W_{i,t}^{(2)}$ are correlated Brownian motions processes with correlation ρ_i, $\{V_{i,t}, t \geq 0\}$ is a mean reverting process with rate of reversion κ_i for $i = 1, \ldots, n$. Following Hull and White (1987)'s derivation, a risk-neutral counterpart of model (S1.8) is

$$\begin{cases} dS_{i,t}/S_{i,t} &= rdt + \sqrt{V_{i,t}}d\widetilde{W}_{i,t}^{(1)} \\ dV_{i,t}/V_{i,t} &= \kappa_i(\sigma_i^* - \sqrt{V_{i,t}})dt + \sigma_i d\widetilde{W}_{i,t}^{(2)} \end{cases} \tag{S1.9}$$

where $\widetilde{W}_{i,t}^{(1)} = W_{i,t}^{(1)} + t\eta_{i,t}$ with $\eta_{i,t} = (\mu_i - r)/\sqrt{V_{i,t}}$, $\widetilde{W}_{i,t}^{(2)} = W_{i,t}^{(2)}$, and $d\widetilde{W}_{i,t}^{(1)} d\widetilde{W}_{i,t}^{(2)} = \rho_i dt$ for each $i = 1, \ldots, n$.

The parameters for computing the prices of the two multi-asset options with (S1.8) are set to be the same as in Table 1. For simplicity, we consider to compute the option prices under the case of independent underlying assets and further set $\sigma_i = 0.1, 1$, $\sigma_i^* = 0.30$ and $\kappa_i = 5, 10$ for $i = 1, \ldots, n$. In addition, the correlation, ρ_i, of $W_{i,t}^{(1)}$ and $W_{i,t}^{(2)}$ is assumed to be -0.7 and initial value of $V_{i,t}$ is assumed to be $V_{i,0} = (\sigma_i^*)^2$ for each $i = 1, \ldots, n$. Therefore, the change of measure process can be represented by

$$\Lambda_t = \prod_{i=1}^{n} \exp\Big\{ -\frac{1}{1-\rho_i^2}\Big(\frac{1}{2}\int_0^t \eta_{i,s}^2 ds + \int_0^t \eta_{i,s} dW_{i,s}^{(1)} - \rho_i \int_0^t \eta_{i,s} dW_{i,s}^{(2)}\Big)\Big\}.$$

By using a daily discretized approximation, model (S1.8) is used to compute the MCP and EPMS price estimates, and model (S1.9) is used to compute the MCQ and EMS price estimates. Table 6 presents the MSE ratios MR^{Q} and MR^{P} for maximum call options with $n = 3$ and for geometric average put options with $n = 10$. Numerical results in Table 6 indicate that the

EMS and EPMS price estimators are capable of improving the efficiency of the MC method, especially for ITM options.

Another widely discussed stochastic volatility model in the literature is Heston (1993)'s model. Let the underlying assets $(S_{1,t},\dots,S_{n,t})$ satisfy the following multivariate stochastic volatility model under the P measure:

$$\begin{cases} dS_{i,t} &= \mu_i S_{i,t}dt + \sqrt{V_{i,t}}S_{i,t}dW_{i,t}^{(1)} \\ dV_{i,t} &= \kappa_i(\theta_i - V_{i,t})dt + \sigma_i\sqrt{V_{i,t}}dW_{i,t}^{(2)} \end{cases} \tag{S1.10}$$

where $\{S_{i,t}, t \geq 0\}$ and $\{V_{i,t}, t \geq 0\}$ are the price and volatility processes, respectively, for the ith underlying asset, $W_{i,t}^{(1)}$ and $W_{i,t}^{(2)}$ are defined the same as in model (S1.8), $\{V_{i,t}, t \geq 0\}$ is a square root mean reverting process proposed by Cox, Ingersoll and Ross (1985) with long run mean θ_i, rate of reversion κ_i, and the volatility of volatility σ_i for each $i = 1,\dots,n$. Following Heston (1993)'s derivation, a risk-neutral counterpart of model (S1.10) is

$$\begin{cases} dS_{i,t} &= rS_{i,t}dt + \sqrt{V_{i,t}}S_{i,t}d\widetilde{W}_{i,t}^{(1)} \\ dV_{i,t} &= \kappa_i^*(\theta_i^* - V_{i,t})dt + \sigma_i\sqrt{V_{i,t}}d\widetilde{W}_{i,t}^{(2)} \end{cases} \tag{S1.11}$$

where $\widetilde{W}_{i,t}^{(1)} = W_{i,t}^{(1)} + t\eta_{i,t}$ with $\eta_{i,t} = (\mu_i - r)/\sqrt{V_{i,t}}$, $\widetilde{W}_{i,t}^{(2)} = W_{i,t}^{(2)} + t\xi_{i,t}$ with $\xi_{i,t} = \lambda_i\sqrt{V_{i,t}}/\sigma_i$, which is called the market price of volatility risk and λ_i is a constant, $d\widetilde{W}_{i,t}^{(1)}d\widetilde{W}_{i,t}^{(2)} = \rho_i dt$, $\kappa_i^* = \kappa + \lambda$ and $\theta_i^* = \kappa_i\theta_i/\kappa_i^*$ for each $i = 1,\dots,n$.

Table 7: The MSE ratios MR^{Q} and MR^{P} for pricing maximum call options and geometric average put options and the coverage rates of the multi-asset EPMS price estimator by using the asymptotic distribution given in Theorem 2 under the multivariate versions of Heston (1993)'s stochastic volatility model (S1.10).

Maximum Call Option ($n = 3$)		T = 30 days			T = 270 days			
	S_0/K	0.97	1.00	1.03	0.97	1.00	1.03	1.10
$(\lambda, \kappa^*) = (0.01,2)$	MR^{Q}	2.29	2.36	2.65	2.87	2.62	2.84	2.53
	MR^{P}	1.95	2.04	2.20	2.08	1.95	2.08	2.10
	25% cov. rate	0.257	0.227	0.262	0.234	0.246	0.235	0.254
	50% cov. rate	0.508	0.485	0.511	0.472	0.496	0.471	0.511
	75% cov. rate	0.746	0.752	0.767	0.728	0.761	0.740	0.754
	95% cov. rate	0.940	0.955	0.949	0.954	0.962	0.959	0.955
$(\lambda, \kappa^*) = (0.05,2)$	MR^{Q}	2.13	2.37	2.85	2.60	2.82	2.72	2.66
	MR^{P}	1.88	2.09	2.43	2.68	3.06	3.05	3.22
	25% cov. rate	0.240	0.237	0.231	0.263	0.262	0.240	0.239
	50% cov. rate	0.483	0.511	0.510	0.506	0.513	0.494	0.501
	75% cov. rate	0.764	0.754	0.759	0.741	0.769	0.735	0.738
	95% cov. rate	0.952	0.955	0.958	0.934	0.946	0.947	0.955
$(\lambda, \kappa^*) = (0.01,4)$	MR^{Q}	2.24	2.42	2.48	2.60	2.54	3.05	2.86
	MR^{P}	1.95	2.01	2.05	1.94	1.94	2.14	2.38
	25% cov. rate	0.238	0.251	0.237	0.250	0.220	0.235	0.259
	50% cov. rate	0.475	0.528	0.488	0.519	0.486	0.490	0.515
	75% cov. rate	0.756	0.739	0.739	0.734	0.763	0.744	0.750
	95% cov. rate	0.955	0.954	0.947	0.944	0.943	0.952	0.954
$(\lambda, \kappa^*) = (0.05,4)$	MR^{Q}	2.15	2.42	2.49	2.49	2.74	3.04	2.82
	MR^{P}	1.95	2.21	2.14	2.49	2.91	3.33	3.60
	25% cov. rate	0.252	0.256	0.270	0.238	0.240	0.254	0.248
	50% cov. rate	0.499	0.502	0.507	0.513	0.472	0.498	0.479
	75% cov. rate	0.751	0.745	0.765	0.764	0.731	0.747	0.766
	95% cov. rate	0.938	0.953	0.951	0.950	0.951	0.957	0.962

Geometric Put Option ($n = 10$)		$T = 30$ days			$T = 270$ days			
	S_0/K	0.97	1.00	1.03	0.90	0.97	1.00	1.03
$(\lambda, \kappa^*) = (0.01,2)$	MR^{Q}	15.42	3.26	1.33	4.07	2.48	1.92	1.38
	MR^{P}	28.80	4.29	1.49	15.48	4.71	2.99	1.94
	25% cov. rate	0.241	0.268	0.240	0.251	0.254	0.271	0.257
	50% cov. rate	0.507	0.520	0.509	0.486	0.493	0.500	0.500
	75% cov. rate	0.750	0.744	0.766	0.728	0.756	0.749	0.753
	95% cov. rate	0.950	0.951	0.944	0.948	0.947	0.955	0.952
$(\lambda, \kappa^*) = (0.05,2)$	MR^{Q}	13.99	2.97	1.36	4.58	2.38	1.77	1.40
	MR^{P}	24.83	4.16	1.51	19.74	5.70	3.34	2.17
	25% cov. rate	0.239	0.251	0.259	0.250	0.246	0.227	0.243
	50% cov. rate	0.491	0.488	0.511	0.493	0.473	0.479	0.477
	75% cov. rate	0.732	0.732	0.731	0.751	0.737	0.737	0.733
	95% cov. rate	0.948	0.949	0.950	0.948	0.950	0.940	0.935
$(\lambda, \kappa^*) = (0.01,4)$	MR^{Q}	14.75	3.00	1.39	4.80	2.18	2.03	1.44
	MR^{P}	25.54	3.89	1.51	12.42	4.48	3.18	2.10
	25% cov. rate	0.269	0.249	0.251	0.248	0.281	0.275	0.247
	50% cov. rate	0.526	0.494	0.502	0.479	0.506	0.502	0.504
	75% cov. rate	0.753	0.749	0.748	0.721	0.775	0.747	0.751
	95% cov. rate	0.944	0.951	0.951	0.948	0.965	0.957	0.957
$(\lambda, \kappa^*) = (0.05,4)$	MR^{Q}	14.37	2.91	1.31	4.49	2.39	1.81	1.48
	MR^{P}	27.18	3.93	1.49	18.46	5.32	3.13	2.22
	25% cov. rate	0.265	0.245	0.259	0.251	0.259	0.244	0.256
	50% cov. rate	0.516	0.512	0.522	0.497	0.502	0.482	0.488
	75% cov. rate	0.761	0.731	0.778	0.749	0.736	0.728	0.744
	95% cov. rate	0.946	0.946	0.961	0.949	0.944	0.932	0.946

The parameters for computing the prices of the two multi-asset options with (S1.10) are set to be the same as in Table 1. For simplicity, we consider to compute the option prices under the case of independent underlying assets and further set $\sigma_i = 0.1$, $\theta_i = 0.09$, $\kappa_i^* = \kappa^* = 2, 4$ and $\lambda_i = \lambda =$

$0.01, 0.05$ for $i = 1, \ldots, n$. In addition, the correlation, ρ_i, of $W_{i,t}^{(1)}$ and $W_{i,t}^{(2)}$ is assumed to be -0.7 and initial value of $V_{i,t}$ is assumed to be $V_{i,0} = \theta_i = 0.09$ for each $i = 1, \ldots, n$. The change of measure process can be represented by

$$\begin{aligned}\Lambda_t &= \prod_{i=1}^{n} \exp\Big\{ -\frac{1}{1-\rho_i^2}\Big(\frac{1}{2}\int_0^t (\eta_{i,s}^2 + \xi_{i,s}^2 - 2\rho_i\eta_{i,s}\xi_{i,s})ds \\ &\qquad + \int_0^t (\eta_{i,s} - \rho_i\xi_{i,s})dW_{i,s}^{(1)} + \int_0^t (\xi_{i,s} - \rho_i\eta_{i,s})dW_{i,s}^{(2)}\Big)\Big\}.\end{aligned}$$

By using a daily discretized approximation, model (S1.10) is used to compute the MCP and EPMS price estimates, and model (S1.11) is used to compute the MCQ and EMS price estimates. Table 7 presents the values of MR^{Q} and MR^{P} for maximum call options with $n = 3$ and for geometric average put options with $n = 10$. Numerical results indicate that the EMS and EPMS have satisfactory performance, especially when an option is ITM.

S2 Theoretical proofs

S2.1 Proof of Example 1

Proof. In the following, we show that

$$f(S_{1,T}, S_{2,T}) = \max(K - \sqrt{S_{1,T}S_{2,T}}, 0)$$

is Lipschitz continuous on $A_1 = \{(S_{1,T}, S_{2,T}) : (S_{1,T}S_{2,T})^{1/2} \le K$ for $S_{1,T} \ge \eta$ and $S_{2,T} \ge \eta\}$, where $0 < \eta < K^{1/2}$. If (S_1, S_2) and (P_1, P_2) are two different points in A_1, then we have

$$\begin{aligned}
|f(S_1,S_2) - f(P_1,P_2)| &\le \left|\sqrt{P_1}(\sqrt{P_2}-\sqrt{S_2}) \times \frac{\sqrt{P_2}+\sqrt{S_2}}{\sqrt{P_2}+\sqrt{S_2}}\right| \\
&\quad + \left|\sqrt{S_2}(\sqrt{P_1}-\sqrt{S_1})\frac{\sqrt{P_1}+\sqrt{S_1}}{\sqrt{P_1}+\sqrt{S_1}}\right| \\
&\le \frac{K}{2\eta}\Big(|P_2 - S_2| + |P_1 - S_1|\Big) \\
&\le \frac{K}{\sqrt{2}\eta}\sqrt{(P_1-S_1)^2 + (P_2-S_2)^2} \\
&= \frac{K}{\sqrt{2}\eta}\|(S_1,S_2) - (P_1,P_2)\|,
\end{aligned}$$

where the second inequality holds because $(P_1P_2)^{1/2} \le K$, $(S_1S_2)^{1/2} \le K$ and P_1, P_2, S_1, $S_2 \ge \eta$, and the last inequality holds by the Cauchy-Schwarz inequality. Hence, the desired result holds. □

S2.2 Proof of Theorem 1

Proof. In the multi-asset EPMS procedure, we define

$$\Lambda^*_{j,t} = \hat{\Lambda}_{j,t}/\bar{\Lambda}_{m,t} \quad \text{and} \quad S^*_{i,j,t} = S_{i,0}e^{rt}\hat{S}_{i,j,t}/\bar{S}^*_{i,m,t}$$

where $\hat{\Lambda}_{j,t} = \Lambda_t(\hat{S}_{1,j,u}, \ldots, \hat{S}_{n,j,u}, u = 1, \ldots, t)$, $\bar{\Lambda}_{m,t} = m^{-1}\sum_{j=1}^m \hat{\Lambda}_{j,t}$ and $\bar{S}^*_{i,m,t} = m^{-1}\sum_{j=1}^m \hat{S}_{i,j,t}\Lambda^*_{j,t}$. Furthermore, let

$$a_i(m,t) = S^*_{i,j,t}/\hat{S}_{i,j,t} = S_{i,0}e^{rt}/\bar{S}^*_{i,m,t},$$

which is independent of paths $j = 1, \ldots, m$ for any fixed $i \in \{1, \ldots, n\}$ and $t \in \{1, \ldots, T\}$. Let

$$b(m,t) = \Lambda^*_{j,t}/\hat{\Lambda}_{j,t} = 1/\bar{\Lambda}_{m,t},$$

which depends only on $t = 1, \ldots, T$.

By the Law of Large Numbers (LLN) and since $\{\Lambda_t\}$ and $\{e^{-rt}S_{i,t}\Lambda_t\}$ are both P-martingale processes, we have

$$\bar{\Lambda}_{m,t} = \frac{1}{m}\sum_{j=1}^{m}\hat{\Lambda}_{j,t} \to E_0(\Lambda_t) = \Lambda_0 = 1$$

and

$$e^{-rt}\bar{S}^*_{i,m,t} = \frac{1}{m}\sum_{j=1}^{m} e^{-rt}\hat{S}_{i,j,t}\hat{\Lambda}_{j,t}/\bar{\Lambda}_{m,t} \to E_0(e^{-rt}S_{i,t}\Lambda_t) = S_{i,0}$$

, almost surely (abbreviated as a.s.), as $m \to \infty$. Consequently,

$$a_i(m,t) \to 1 \quad \text{and} \quad b(m,t) \to 1, \tag{S2.1}$$

a.s. as $m \to \infty$, for $i = 1, \ldots, n$ and $t = 1, \ldots, T$.

Next, we show that

$$\frac{1}{m}\sum_{j=1}^{m} f(\mathbf{S}^*_{j,1}, \ldots, \mathbf{S}^*_{j,T})\Lambda^*_{j,T} \to E_0\{f(\mathbf{S}_1, \ldots, \mathbf{S}_T)\Lambda_T\},$$

a.s. as $m \to \infty$, where $\mathbf{S}^*_{j,t} = (S^*_{1,j,t}, \ldots, S^*_{n,j,t})$, $t = 1, \ldots, T$. Since the LLN guarantees that $m^{-1}\sum_{j=1}^{m} f(\hat{\mathbf{S}}_{j,1}, \ldots, \hat{\mathbf{S}}_{j,T})\hat{\Lambda}_{j,T} \to E_0\{f(\mathbf{S}_1, \ldots, \mathbf{S}_T)\Lambda_T\}$, a.s. as $m \to \infty$, it suffices to prove that

$$\frac{1}{m}\sum_{j=1}^{m} |f(\mathbf{S}^*_{j,1}, \ldots, \mathbf{S}^*_{j,T})\Lambda^*_{j,T} - f(\hat{\mathbf{S}}_{j,1}, \ldots, \hat{\mathbf{S}}_{j,T})\hat{\Lambda}_{j,T}| \to 0, \tag{S2.2}$$

a.s. as $m \to \infty$.

Before proving (S2.2), we define the following notations:

(a) For the jth sample path $(\hat{\mathbf{S}}_{j,1}, \ldots, \hat{\mathbf{S}}_{j,T})$, $j = 1, \ldots, m$, we define $\mathbf{x}_j$ to be a point in G such that $||\mathbf{x}_j - (\hat{\mathbf{S}}_{j,1}, \ldots, \hat{\mathbf{S}}_{j,T})|| \leq ||\mathbf{y} - (\hat{\mathbf{S}}_{j,1}, \ldots, \hat{\mathbf{S}}_{j,T})||$, for any $\mathbf{y} \in G$, where G is defined in (3.2). Let $B(\mathbf{x}_j, \delta)$ be an open ball with center $\mathbf{x}_j$ and radius $\delta > 0$. Furthermore, let $B_{j,\delta} = B(\mathbf{x}_j, \delta) \cap D^f$.

(b) Let

$$D_u = \{\|(\widehat{\mathbf{S}}_1, \ldots, \widehat{\mathbf{S}}_T)\|_1 \leq u\},$$

where $\|(\widehat{\mathbf{S}}_1, \ldots, \widehat{\mathbf{S}}_T)\|_1 = \sum_{i=1}^n \sum_{t=1}^T |\widehat{S}_{i,t}|$ is the ℓ_1-norm of $(\widehat{\mathbf{S}}_1, \ldots, \widehat{\mathbf{S}}_T)$, and let $A^*_{j,\delta,\ell} = A_\ell \cap B^{\mathrm{c}}_{j,\delta}$ for $\ell = 1, \ldots, k$, where $B^{\mathrm{c}}_{j,\delta}$ is the complement of $B_{j,\delta}$. Moreover, let $D_{j,\delta,u} = B_{j,\delta} \cap D_u$ and $\mathbf{A}_{j,\delta,u} = \bigcup_{\ell=1}^k A^*_{j,\delta,\ell} \cap D_u$.

By using the above notations, the left-hand-side of (S2.2) can be rewritten as

$$\begin{aligned}
&\frac{1}{m}\sum_{j=1}^m |f(\mathbf{S}^*_{j,1}, \ldots, \mathbf{S}^*_{j,T})\Lambda^*_{j,T} - f(\hat{\mathbf{S}}_{j,1}, \ldots, \hat{\mathbf{S}}_{j,T})\hat{\Lambda}_{j,T}| \\
= \;& \frac{1}{m}\sum_{j=1}^m |f(\mathbf{S}^*_{j,1}, \ldots, \mathbf{S}^*_{j,T})\Lambda^*_{j,T} - f(\hat{\mathbf{S}}_{j,1}, \ldots, \hat{\mathbf{S}}_{j,T})\hat{\Lambda}_{j,T}| \; I_{\mathbf{A}_{j,\delta,u}} \\
&+ \frac{1}{m}\sum_{j=1}^m |f(\mathbf{S}^*_{j,1}, \ldots, \mathbf{S}^*_{j,T})\Lambda^*_{j,T} - f(\hat{\mathbf{S}}_{j,1}, \ldots, \hat{\mathbf{S}}_{j,T})\hat{\Lambda}_{j,T}| \; I_{D_{j,\delta,u}} \\
&+ \frac{1}{m}\sum_{j=1}^m |f(\mathbf{S}^*_{j,1}, \ldots, \mathbf{S}^*_{j,T})\Lambda^*_{j,T} - f(\hat{\mathbf{S}}_{j,1}, \ldots, \hat{\mathbf{S}}_{j,T})\hat{\Lambda}_{j,T}| \; I_{D^{\mathrm{c}}_u}, \quad \text{(S2.3)}
\end{aligned}$$

for any $u > 0$ and $\delta > 0$, where I_D is an indicator function with value 1 when the event D occurs. In the following, we show that the three terms on the right-hand-side (rhs) of (S2.3) converge to 0, a.s. as $m \to \infty$.

(i) By using similar arguments used in the proof of Proposition 2 in Duan and Simonato (1998), if $(\hat{\mathbf{S}}_{j,1}, \ldots, \hat{\mathbf{S}}_{j,T}) \in A^*_{j,\delta,\ell} \subseteq A_\ell$ and the event $\mathbf{A}_{j,\delta,u}$ occurs, then there exists an integer M such that for all $m \geq M$, $(\mathbf{S}^*_{j,1}, \ldots, \mathbf{S}^*_{j,T}) \in A_\ell$. Consequently, the assumption (A2) yields that

$$\begin{aligned} &|f(\mathbf{S}^*_{j,1}, \ldots, \mathbf{S}^*_{j,T}) - f(\hat{\mathbf{S}}_{j,1}, \ldots, \hat{\mathbf{S}}_{j,T})| \\ &\leq c \sum_{i=1}^{n} \sum_{t=1}^{T} |S^*_{i,j,t} - \hat{S}_{i,j,t}|, \end{aligned} \tag{S2.4}$$

for any jth path and for some positive constant $c < \infty$. Hence, (S2.4) yields that

$$\begin{aligned} &\frac{1}{m} \sum_{j=1}^{m} |f(\mathbf{S}^*_{j,1}, \ldots, \mathbf{S}^*_{j,T}) \Lambda^*_{j,T} - f(\hat{\mathbf{S}}_{j,1}, \ldots, \hat{\mathbf{S}}_{j,T}) \hat{\Lambda}_{j,T}| \ I_{\mathbf{A}_{j,\delta,u}} \\ \leq \ & c\ (|b(m,T) - 1| + 1) \sum_{i=1}^{n} \sum_{t=1}^{T} \left[|a_i(m,t) - 1| \Big(\frac{1}{m} \sum_{j=1}^{m} \hat{S}_{i,j,t} \hat{\Lambda}_{j,T} \Big) \right] \\ & + |b(m,T) - 1| \left(\frac{1}{m} \sum_{j=1}^{m} |f(\hat{\mathbf{S}}_{j,1}, \ldots, \hat{\mathbf{S}}_{j,T})| \ \hat{\Lambda}_{j,T} \right). \end{aligned} \tag{S2.5}$$

In addition, by the LLN and (A4), we have

$$\frac{1}{m} \sum_{j=1}^{m} \hat{S}_{i,j,t} \hat{\Lambda}_{j,T} \to E(S_{i,t} \Lambda_T) = e^{rt} S_{i,0}, \ \text{a.s.},$$

and

$$\frac{1}{m}\sum_{j=1}^{m}|f(\hat{\mathbf{S}}_{j,1},\ldots,\hat{\mathbf{S}}_{j,T})|\hat{\Lambda}_{j,T}\to E^{Q}(|f(\hat{\mathbf{S}}_{1},\ldots,\hat{\mathbf{S}}_{T})|)<\infty,\ \text{a.s.},$$

which together with (S2.1) yield that the rhs of (S2.5) converges to 0, a.s. as $m\to\infty$.

(ii) Note that

$$\begin{aligned}
& \frac{1}{m}\sum_{j=1}^{m}|f(\mathbf{S}^{*}_{j,1},\ldots,\mathbf{S}^{*}_{j,T})\Lambda^{*}_{j,T}-f(\hat{\mathbf{S}}_{j,1},\ldots,\hat{\mathbf{S}}_{j,T})\hat{\Lambda}_{j,T}|\ I_{D_{j,\delta,u}}\\
\leq\ & \frac{1}{m}\sum_{j=1}^{m}\{|f(\mathbf{S}^{*}_{j,1},\ldots,\mathbf{S}^{*}_{j,T})|\ \Lambda^{*}_{j,T}+|f(\hat{\mathbf{S}}_{j,1},\ldots,\hat{\mathbf{S}}_{j,T})|\ \hat{\Lambda}_{j,T}\}\ I_{D_{j,\delta,u}}\\
\leq\ & \frac{c}{m}\sum_{j=1}^{m}\left\{\left[1+\left(\sum_{i=1}^{n}\sum_{t=1}^{T}S^{*}_{i,j,t}\right)^{q}\right]\frac{\Lambda^{*}_{j,T}}{\hat{\Lambda}_{j,T}}+(1+u^{q})\right\}\hat{\Lambda}_{j,T}I_{D_{j,\delta,u}}\\
\leq\ & c\left\{[1+u^{q}(\max_{i,t}\ a_i(m,t))^{q}]b(m,T)+(1+u^{q})\right\}\\
& \quad\left(\frac{1}{m}\sum_{j=1}^{m}\hat{\Lambda}_{j,T}I_{D_{j,\delta,u}}\right),
\end{aligned}$$

where the second inequality holds by the assumption (A3), $||\mathbf{x}||^{q}\leq||\mathbf{x}||_1^{q}$ for a vector $\mathbf{x}$ and $(\hat{\mathbf{S}}_{j,1},\ldots,\hat{\mathbf{S}}_{j,T})\in D_{j,\delta,u}\subset D_u$ for each $j=1,\ldots,m$. In addition, the LLN yields that $m^{-1}\sum_{j=1}^{m}\hat{\Lambda}_{j,T}I_{D_{j,\delta,u}}$ converges to $P^{Q}(D_{1,\delta,u})$, a.s., as $m\to\infty$, where P^{Q} is the probability under a Q measure. In addition, $P^{Q}(D_{1,\delta,u})\leq P^{Q}(B_{1,\delta})\leq K\delta^{nT}$ for some finite positive constant K, where the last inequality holds by the assumption (A5). By (S2.1) and since u and δ are arbitrary positive numbers, let $\delta = u^{-2q}$ and thereby the second term on the rhs of (S2.3) converges to 0, a.s. as $u \to \infty$ and $m \to \infty$.

(iii) By (A3), we have

$$\begin{aligned}
&\frac{1}{m}\sum_{j=1}^{m} |f(\mathbf{S}^*_{j,1},\ldots,\mathbf{S}^*_{j,T})\ \Lambda^*_{j,T} - f(\hat{\mathbf{S}}_{j,1},\ldots,\hat{\mathbf{S}}_{j,T})\ \hat{\Lambda}_{j,T}|\ I_{D_u^{\mathrm{c}}} \\
\le\ & c\,[1+b(m,T)]\Big(\frac{1}{m}\sum_{j=1}^{m}\hat{\Lambda}_{j,T} I_{D_u^{\mathrm{c}}}\Big) \\
&+c\,\{1+(\max_j\ S_n(j,T))^q b(m,T)\} \\
&\ \Big(\frac{1}{m}\sum_{j=1}^{m}\|(\hat{\mathbf{S}}_{j,1},\ldots,\hat{\mathbf{S}}_{j,T})\|^q\hat{\Lambda}_{j,T} I_{D_u^{\mathrm{c}}}\Big),
\end{aligned}$$

where $S_n(j,T) = \|(\mathbf{S}^*_{j,1},\ldots,\mathbf{S}^*_{j,T})\|/\|(\hat{\mathbf{S}}_{j,1},\ldots,\hat{\mathbf{S}}_{j,T})\|$, which converges to 1, a.s. as $m \to \infty$ by (S2.1). In addition, the LLN ensures that

$$\frac{1}{m}\sum_{j=1}^{m}\hat{\Lambda}_{j,T} I_{D_u^{\mathrm{c}}} \to P^Q(D_u^{\mathrm{c}})$$

and

$$\frac{1}{m}\sum_{j=1}^{m}\|(\hat{\mathbf{S}}_{j,1},\ldots,\hat{\mathbf{S}}_{j,T})\|^q\hat{\Lambda}_{j,T} I_{D_u^{\mathrm{c}}} \to E^Q[\|(\hat{\mathbf{S}}_1,\ldots,\hat{\mathbf{S}}_T)\|^q I_{D_u^{\mathrm{c}}}],$$

a.s. as $m \to \infty$. Since (A5) ensures that $E^Q[\|(\hat{\mathbf{S}}_1,\ldots,\hat{\mathbf{S}}_T)\|^q I_{D_u^{\mathrm{c}}}]$ and $P^Q(D_u^{\mathrm{c}})$ converge to 0 as $u \to \infty$, thus the third term on the rhs of (S2.3) converges to 0, a.s. as $m \to \infty$ and $u \to \infty$.

According to (S2.3) and (i)-(iii), let $u = m$ and thereby the desired result holds. $\square$

S2.3 Proof of Theorem 2

Proof. Throughout this proof, we use $|\mathbf{a}|$ to denote $(|a_1|, \ldots, |a_n|)$, use $\mathbf{a}^{-1}$ to denote $(a_1^{-1}, \ldots, a_n^{-1})$, and use $\mathbf{a} \leq \mathbf{b}$ to denote that $a_i \leq b_i$ for each $i = 1, \ldots, n$, where $\mathbf{a} = (a_1, \ldots, a_n)$ and $\mathbf{b} = (b_1, \ldots, b_n)$. Let ε_M, $M = 1, 2, \ldots$, be a sequence of decreasing positive numbers and satisfy $\lim_{M\to\infty} \varepsilon_M = 0$. By the assumption (A4'), for any ε_M, there exists a compact subset of D^f, denoted by K_M, such that

$$\left| C - e^{-rT} E[f(\mathbf{S}_T)\Lambda_T I_{K_M}] \right| < \frac{1}{2}\varepsilon_M, \tag{S2.6}$$

where $C = e^{-rT} E[f(\mathbf{S}_T)\Lambda_T]$ denotes the theoretical price of the derivative. Note that $K_1 \subset K_2 \subset \cdots$ and $\lim_{M\to\infty} K_M = D^f$.

Next, since $f(\mathbf{S}_T)$ is continuous on K_M, by Weierstrass's Approximation Theorem, for any given $\varepsilon_M > 0$ there exists a polynomial function $\varphi_M : K_M \to R$ such that

$$\sup_{\mathbf{S}_T \in K_M} |f(\mathbf{S}_T) - \varphi_M(\mathbf{S}_T)| < \frac{1}{2}\varepsilon_M. \tag{S2.7}$$

We refer the reader to Timan (1963, page 1) for details of (S2.7). Furthermore, let $G_M = \bigcup_{\mathbf{S}_T \in G} B(\mathbf{S}_T, \delta_M)$, where $B(\mathbf{S}_T, \delta_M)$ is an open ball centered at $\mathbf{S}_T$ with radius δ_M, which satisfies $\lim_{M\to\infty} \delta_M = 0$, and G is defined in (3.2). Since $K_M \backslash G_M$ is a finite union of disjoint compact sets, by (S2.7) and (Duistermaat and Kolk, 2004, page 667) we also can find a large enough integer M' such that for all $M \geq M'$

$$\sup_{\mathbf{S}_T \in K_M \backslash G_M} ||\nabla f(\mathbf{S}_T) - \nabla \varphi_M(\mathbf{S}_T)||_1 < \varepsilon_M, \tag{S2.8}$$

for any given $\varepsilon_M > 0$. Moreover, for any $\mathbf{S}_T \in K_M$ we have

$$\lim_{M \to \infty} ||\nabla \varphi_M(\mathbf{S}_T)||_1 \leq 2 \sum_{\ell \in H} ||\nabla f_\ell(\mathbf{S}_T)||_1, \tag{S2.9}$$

where $H \subseteq \{1, \ldots, k\}$ is the collection of the index of A_ℓ, whose closure contains $\mathbf{S}_T$. We prove (S2.9) in the following two steps. First, if $\mathbf{S}_T \notin G$, then there exist only one ℓ and an integer M'' such that for all $M \geq M''$ $\mathbf{S}_T \in A_\ell \cap (K_M \backslash G_M)$. That is, $H = \{\ell\}$ and $\mathbf{S}_T \in A_\ell^* \cap G_M^c$, where $A_{\ell_i}^*$ denotes the interior of A_{ℓ_i} and G_M^c denotes the complement of G_M. By (S2.8), we have

$$||\nabla \varphi_M(\mathbf{S}_T)||_1 \leq ||\nabla f_\ell(\mathbf{S}_T)||_1 + \varepsilon_M, \text{ for } \mathbf{S}_T \notin G. \tag{S2.10}$$

On the other hand, if $\mathbf{S}_T \in G$, then there could be more than one $\overline{A}_\ell$ containing $\mathbf{S}_T$. Let $H = \{\ell_1, \ldots, \ell_p\}$, where $\ell_1 < \cdots < \ell_p$ and $p \leq n$. That is, $\mathbf{S}_T \in \bigcap_{i=1}^p \overline{A}_{\ell_i}$. If $||\nabla \varphi_M(\mathbf{S}_T)||_1 = 0$, then (S2.9) holds. If $||\nabla \varphi_M(\mathbf{S}_T)||_1 > 0$, by using the continuity of $\nabla \phi_M$ on K_M, there exists an $\eta_M > 0$ such that $||\nabla \varphi_M(\mathbf{x})||_1 > 0$ and $||\nabla \varphi_M(\mathbf{S}_T)||_1 \leq 2||\nabla \varphi_M(\mathbf{x})||_1$ for all $\mathbf{x} \in B(\mathbf{S}_T, \eta_M)$. By (S2.10) and choosing an $\mathbf{x}_i$ from the intersection of $A_{\ell_i}^*$ and $B(\mathbf{S}_T, \eta_M)$, we have

$$||\nabla \varphi_M(\mathbf{x}_i)||_1 \leq ||\nabla f_\ell(\mathbf{x}_i)||_1 I_{A_{\ell_i}^* \cap B(\mathbf{S}_T, \eta_M)}(\mathbf{x}_i) + \varepsilon_M$$

and thereby

$$||\nabla\varphi_M(\mathbf{S}_T)||_1 \le 2\sum_{i=1}^{p} ||\nabla f_{\ell_i}(\mathbf{x}_i)||_1 I_{A^*_{\ell_i}\cap B(\mathbf{s}_T,\eta_M)}(\mathbf{x}_i) + 2p\varepsilon_M.$$

Moreover, by using the continuity of ∇f_{ℓ_i} assumed in (A2'), we have

$$||\nabla f_{\ell_i}(\mathbf{x}_i)||_1 I_{A^*_{\ell_i}\cap B(\mathbf{s}_T,\eta_M)}(\mathbf{x}_i) - ||\nabla f_{\ell_i}(\mathbf{S}_T)||_1 \to 0, \text{ as } \eta_M \to 0,$$

for $i = 1, \dots, p$, and thereby

$$||\nabla\varphi_M(\mathbf{S}_T)||_1 \le 2\sum_{i=1}^{p} ||\nabla f_{\ell_i}(\mathbf{S}_T)||_1 + \tilde{\varepsilon}_M, \text{ for } \mathbf{S}_T \in G, \tag{S2.11}$$

where $\tilde{\varepsilon}_M > 0$ and $\lim_{M\to\infty}\tilde{\varepsilon}_M = 0$. Consequently, (S2.10) and (S2.11) yield (S2.9).

By (S2.6), (S2.7) and the fact of $E(\Lambda_T) = 1$, we have

$$|C - e^{-rT}E[\varphi_M(\mathbf{S}_T)\Lambda_T]| < \varepsilon_M$$

and thereby

$$\begin{aligned}
& e^{rT}\sqrt{m}(C^{(m)}_{\text{EPMS}} - C^{(m)}_{\text{MC}}) \\
= \ & \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\lim_{M\to\infty}\{\varphi_M(\mathbf{S}^*_{j,T})\Lambda^*_{j,T} - \varphi_M(\mathbf{S}_{j,T})\Lambda_{j,T}\} \\
= \ & \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\lim_{M\to\infty}\{\varphi_M(\mathbf{S}^*_{j,T}) - \varphi_M(\hat{\mathbf{S}}_{j,T})\}\ \hat{\Lambda}_{j,T} \\
& +\frac{1}{\sqrt{m}}\sum_{j=1}^{m} f(\hat{\mathbf{S}}_{j,T})(\Lambda^*_{j,T} - \hat{\Lambda}_{j,T}) \\
& +\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\lim_{M\to\infty}\{\varphi_M(\mathbf{S}^*_{j,T}) - \varphi_M(\hat{\mathbf{S}}_{j,T})\} \times (\Lambda^*_{j,T} - \hat{\Lambda}_{j,T}).
\end{aligned} \tag{S2.12}$$

In the following, we deal with the three terms on the rhs of (S2.12) separately.

Note that for any fixed $\mathbf{s} > \mathbf{0}$, the function $\mathbf{v} \to \varphi_M(\mathbf{s} \circ \mathbf{v}^{-1})$ is continuous and differentiable for $\mathbf{v} \in (0, \infty)^n$. Since $\mathbf{S}_T$ is a positive random vector, we apply the multi-dimensional mean value theorem to $\varphi_M(\mathbf{s} \circ \mathbf{v}^{-1})$. That is, there exists a constant $d \in [0, 1]$ such that

$$
\begin{aligned}
& \varphi_M(\mathbf{S}^*_{j,T}) - \varphi_M(\hat{\mathbf{S}}_{j,T}) \\
= \; & \varphi_M(e^{rT}\mathbf{S}_0 \circ \hat{\mathbf{S}}_{j,T} \circ (\bar{\mathbf{S}}^*_{m,T})^{-1}) - \varphi_M(e^{rT}\mathbf{S}_0 \circ \hat{\mathbf{S}}_{j,T} \circ (e^{rT}\mathbf{S}_0)^{-1}) \\
= \; & -(\bar{\mathbf{S}}^*_{m,T} - e^{rT}\mathbf{S}_0)\Big\{\nabla\varphi_M(\mathbf{W}^{(M)}_{j,1}) \circ (\mathbf{W}^{(M)}_{j,2})^\top\Big\}, \qquad \text{(S2.13)}
\end{aligned}
$$

where $\bar{\mathbf{S}}^*_{m,T} = (\bar{S}^*_{1,m,T}, \ldots, \bar{S}^*_{n,m,T})$, in which $\bar{S}^*_{i,m,T} = m^{-1}\sum_{j=1}^m \hat{S}_{i,j,T}\Lambda^*_{j,T}$, $\mathbf{S}^*_{j,T} = e^{rT}\mathbf{S}_0 \circ \hat{\mathbf{S}}_{j,T} \circ (\bar{\mathbf{S}}^*_{m,T})^{-1}$, $\mathbf{W}^{(M)}_{j,k} = e^{rT}\mathbf{S}_0 \circ \hat{\mathbf{S}}_{j,T} \circ (\mathbf{V}^{(M)}_{j,m})^{-k}$, for $k = 1, 2$, and $\mathbf{V}^{(M)}_{j,m} = (1-d)e^{rT}\mathbf{S}_0 + d\overline{\mathbf{S}}^*_{m,T}$. In addition, by similar arguments used in (S2.13) and recalling that $\Lambda^*_{j,T} = \hat{\Lambda}_{j,T}/\bar{\Lambda}_{m,T}$, we also have

$$
\Lambda^*_{j,T} - \hat{\Lambda}_{j,T} = -\frac{\hat{\Lambda}_{j,T}}{(U^*_{j,m})^2}(\bar{\Lambda}_{m,T} - 1), \qquad \text{(S2.14)}
$$

where $U^*_{j,m} \in [\min(\bar{\Lambda}_{m,T}, 1), \max(\bar{\Lambda}_{m,T}, 1)]$. By substituting (S2.13) and (S2.14) into (S2.12), we have

$$
\text{(S2.12)} = \text{(I)} + \text{(II)} + \text{(III)}, \qquad \text{(S2.15)}
$$

where

$$
\begin{aligned}
\text{(I)} &= -(\bar{\mathbf{S}}^*_{m,T} - e^{rT}\mathbf{S}_0)\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\hat{\Lambda}_{j,T}\lim_{M\to\infty}\Big\{\nabla\varphi_M(\mathbf{W}^{(M)}_{j,1})\circ(\mathbf{W}^{(M)}_{j,2})^\top\Big\},\\
\text{(II)} &= -\frac{1}{\sqrt{m}}\sum_{j=1}^{m} f(\hat{\mathbf{S}}_{j,T})\left(\frac{\hat{\Lambda}_{j,T}}{(U^*_{j,m})^2}(\bar{\Lambda}_{m,T}-1)\right)\\
\text{(III)} &= \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\lim_{M\to\infty}\Big[\varphi_M(\mathbf{S}^*_{j,T}) - \varphi_M(\hat{\mathbf{S}}_{j,T})\Big](\Lambda^*_{j,T} - \hat{\Lambda}_{j,T}).
\end{aligned}
$$

Let $\Phi = e^{-rT}E[\Lambda_T\nabla f(\mathbf{S}_T)\circ(\mathbf{S}_T\circ\mathbf{S}_0^{-1})^\top]$. Lemmas 1-3, respectively, given below yield that

$$
\begin{aligned}
\text{(I)} &= -\sqrt{m}(\bar{\mathbf{S}}_{m,T} - e^{rT}\bar{\Lambda}_{m,T}\mathbf{S}_0)\Phi + o_p(1), && \text{(S2.16)}\\
\text{(II)} &= -\sqrt{m}(\bar{\Lambda}_{m,T}-1)E[f(\mathbf{S}_T)\Lambda_T] + o_p(1), && \text{(S2.17)}\\
\text{(III)} &= o_p(1), && \text{(S2.18)}
\end{aligned}
$$

where $\bar{\mathbf{S}}_{m,T} = (\bar{S}_{1,m,T},\ldots,\bar{S}_{n,m,T})$. Consequently, by (S2.15)-(S2.18), and let $\Psi = E(f(\mathbf{S}_T)\Lambda_T) - e^{rT}\mathbf{S}_0\Phi$, we have $C^{(m)}_{\text{MC}} - C^{(m)}_{\text{EPMS}} = e^{-rT}\{(\bar{\mathbf{S}}_{m,T} - e^{rT}\mathbf{S}_0)\Phi + (\bar{\Lambda}_{m,T}-1)\Psi\} + o_p(m^{-1/2})$. Hence, Theorem 2 (i) holds.

Next, we derive the result of Theorem 2 (ii). By Theorem 2 (i), we have

$$
\begin{aligned}
&\sqrt{m}(C^{(m)}_{\text{EPMS}} - C)\\
&= \sqrt{m}(C^{(m)}_{\text{MC}} - C) - \sqrt{m}e^{-rT}\{(\bar{\mathbf{S}}_{m,T} - e^{rT}\mathbf{S}_0)\Phi + (\bar{\Lambda}_{m,T}-1)\Psi\} + o_p(1)\\
&= \frac{e^{-rT}}{\sqrt{m}}\sum_{j=1}^{m}\{[f(\hat{\mathbf{S}}_{j,T})\hat{\Lambda}_{j,T} - E(f(\mathbf{S}_T)\Lambda_T)] - [\hat{\Lambda}_{j,T}\hat{\mathbf{S}}_{j,T} - E(\Lambda_T\mathbf{S}_T)]\Phi\\
&\qquad -[\Lambda_{j,T} - E(\Lambda_T)]\Psi\} + o_p(1)
\end{aligned}
$$

By the central limit theorem (CLT), we have

$$\sqrt{m}(C^{(m)}_{\text{EPMS}} - C) \xrightarrow{\mathcal{L}} N(0, V), \text{ as } m \to \infty,$$

where

$$\begin{aligned} V &= e^{-2rT}\Big\{\text{Var}(f(\mathbf{S}_T)\Lambda_T) + \Phi^\top \text{Var}(\Lambda_T \mathbf{S}_T)\Phi + \Psi^2 \text{Var}(\Lambda_T) \\ &\quad -2\Big[\Phi^\top \text{Cov}(f(\mathbf{S}_T)\Lambda_T, \Lambda_T \mathbf{S}_T) + \Psi \text{Cov}(f(\mathbf{S}_T)\Lambda_T, \Lambda_T) \\ &\quad -\Psi\Phi^\top \text{Cov}(\Lambda_T, \Lambda_T \mathbf{S}_T)\Big]\Big\}. \end{aligned}$$

□

Lemma 1. *By using the same assumptions as in Theorem 2 (i), (S2.16) holds.*

Proof. First, we rewrite (I) as

$$\text{(I)} = -\sqrt{m}(\bar{\mathbf{S}}^*_{m,T} - e^{rT}\mathbf{S}_0)\left(\Phi - \mathbf{Y}_{m,T}\right), \tag{S2.19}$$

where

$$\begin{aligned} \mathbf{Y}_{m,T} &= E[e^{-rT}\Lambda_T \nabla f(\mathbf{S_T}) \circ (\mathbf{S}_T \circ \mathbf{S}_0^{-1})^\top] \\ &\quad -\frac{1}{m}\sum_{j=1}^m \hat{\Lambda}_{j,T} \lim_{M\to\infty}\Big\{\nabla\varphi_M(\mathbf{W}^{(M)}_{j,1}) \circ (\mathbf{W}^{(M)}_{j,2})^\top\Big\} \\ &= \frac{1}{m}\sum_{j=1}^m e^{-rT}\hat{\Lambda}_{j,T}\nabla f(\hat{\mathbf{S}}_{j,T}) \circ (\hat{\mathbf{S}}_{j,T} \circ \mathbf{S}_0^{-1})^\top \\ &\quad -\frac{1}{m}\sum_{j=1}^m \hat{\Lambda}_{j,T} \lim_{M\to\infty}\Big\{\nabla\varphi_M(\mathbf{W}^{(M)}_{j,1}) \circ (\mathbf{W}^{(M)}_{j,2})^\top\Big\} + \mathbf{o_p(1)} \\ &= \mathbf{Y}^{(1)}_{m,T} + \mathbf{Y}^{(2)}_{m,T} + \mathbf{o_p(1)}, \end{aligned} \tag{S2.20}$$

in which the 2nd equality holds by the LLN and $\mathbf{o_p}(\mathbf{1})$ stands for a random vector whose components are $o_p(1)$,

$$\mathbf{Y}^{(1)}_{m,T} = \frac{1}{m}\sum_{j=1}^{m} e^{-rT}\hat{\Lambda}_{j,T}\left(\hat{\mathbf{S}}_{j,T}\circ\mathbf{S}_0^{-1}\right)^{\top}\circ\left[\nabla f(\hat{\mathbf{S}}_{j,T}) - \lim_{M\to\infty}\nabla\varphi_M(\mathbf{W}^{(M)}_{j,1})\right]$$

and

$$\begin{aligned}\mathbf{Y}^{(2)}_{m,T} &= \frac{1}{m}\sum_{j=1}^{m}\left(e^{rT}\hat{\Lambda}_{j,T}\mathbf{S}_0\circ\hat{\mathbf{S}}_{j,T}\right)^{\top} \\ &\qquad\circ\Big\{\lim_{M\to\infty}\nabla\varphi_M(\mathbf{W}^{(M)}_{j,1})\circ\left(e^{-2rT}\mathbf{S}_0^{-2}-(\mathbf{V}^{(M)}_{j,m})^{-2}\right)^{\top}\Big\}.\end{aligned}$$

Recall that

$$\bar{\mathbf{S}}_{m,T} = \frac{1}{m}\sum_{j=1}^{m}\hat{\Lambda}_{j,T}\hat{\mathbf{S}}_{j,T} \quad\text{and}\quad \bar{\mathbf{S}}^{*}_{m,T} = \frac{1}{m}\sum_{j=1}^{m}\Lambda^{*}_{j,T}\hat{\mathbf{S}}_{j,T} = \bar{\Lambda}^{-1}_{m,T}\bar{\mathbf{S}}_{m,T}$$

for $i=1,\ldots,n$ since $\Lambda^{*}_{j,T}=\hat{\Lambda}_{j,T}/\bar{\Lambda}_{m,T}$. We then have

$$\begin{aligned}&\sqrt{m}(\bar{\mathbf{S}}^{*}_{m,T}-e^{rT}\mathbf{S}_0)\\ &= \sqrt{m}\bar{\Lambda}^{-1}_{m,T}\{(\bar{\mathbf{S}}_{m,T}-e^{rT}\mathbf{S}_0)-e^{rT}\mathbf{S}_0(\bar{\Lambda}_{m,T}-1)\}. \qquad \text{(S2.21)}\end{aligned}$$

By the CLT, $\sqrt{m}(\bar{S}_{i,m,T}-S_{i,0}e^{rT})$ and $\sqrt{m}S_{i,0}e^{rT}(\bar{\Lambda}_{m,T}-1)$ converge weakly to proper normal random variables for $i=1,\ldots,n$. In addition, since $\bar{\Lambda}_{m,T}\to 1$, a.s., thus, (S2.21) can be rewritten as

$$\sqrt{m}(\bar{\mathbf{S}}^{*}_{m,T}-e^{rT}\mathbf{S}_0) = \sqrt{m}(\bar{\mathbf{S}}_{m,T}-\bar{\Lambda}_{m,T}e^{rT}\mathbf{S}_0)+\mathbf{o_p}(\mathbf{1}). \qquad \text{(S2.22)}$$

By (S2.19) and (S2.22), we have

$$(\mathrm{I}) = -\sqrt{m}(\bar{\mathbf{S}}_{m,T}-\bar{\Lambda}_{m,T}e^{rT}\mathbf{S}_0)(\Phi-\mathbf{Y}_{m,T})+\mathbf{o_p}(\mathbf{1}).$$

To obtain (S2.16), it remains to show that the $\mathbf{Y}^{(1)}_{m,T}$ and $\mathbf{Y}^{(2)}_{m,T}$ defined in (S2.20) satisfy

$$||\mathbf{Y}^{(1)}_{m,T}||_1 + ||\mathbf{Y}^{(2)}_{m,T}||_1 = o_p(1). \tag{S2.23}$$

Let $\boldsymbol{\ell}_{j,m} := \lim_{M\to\infty} \mathbf{W}^{(M)}_{j,1}$ and we have

$$\begin{aligned}
& ||\mathbf{Y}^{(1)}_{m,T}||_1 \\
\leq \;& \frac{1}{m}\sum_{j=1}^{m} e^{-rT}\hat{\Lambda}_{j,T}||\hat{\mathbf{S}}_{j,T}\circ \mathbf{S}_0^{-1}||_1 \times ||\nabla f(\hat{\mathbf{S}}_{j,T}) - \lim_{M\to\infty}\nabla\varphi_M(\mathbf{W}^{(M)}_{j,1})||_1 \\
\leq \;& \frac{1}{m}\sum_{j=1}^{m} e^{-rT}\hat{\Lambda}_{j,T}||\hat{\mathbf{S}}_{j,T}\circ \mathbf{S}_0^{-1}||_1 \times ||\nabla f(\hat{\mathbf{S}}_{j,T}) - \nabla f(\boldsymbol{\ell}_{j,m})||_1 \; I_{\{\boldsymbol{\ell}_{j,m}\notin G\}} \\
& + \frac{1}{m}\sum_{j=1}^{m} e^{-rT}\hat{\Lambda}_{j,T}||\hat{\mathbf{S}}_{j,T}\circ \mathbf{S}_0^{-1}||_1 \times \\
& \qquad \Big(||\nabla f(\hat{\mathbf{S}}_{j,T})||_1 + 2\sum_{\ell\in H}||\nabla f_\ell(\ell_{j,m})||_1\Big) \; I_{\{\boldsymbol{\ell}_{j,m}\in G\}} \\
:= \;& B_1 + B_2,
\end{aligned} \tag{S2.24}$$

where the second inequality hold by (S2.8) and (S2.9). Since $\lim_{m\to\infty}\mathbf{V}^{(M)}_{j,m} = e^{rT}\mathbf{S}_0$, a.s., for all j and M, we have

$$B_1 = o_p(1), \text{ as } m\to\infty, \tag{S2.25}$$

and thus $\lim_{m\to\infty}\boldsymbol{\ell}_{j,m} = \hat{\mathbf{S}}_{j,T}$, a.s.. In addition, (A3') yields that

$$\begin{aligned}
\lim_{m\to\infty} E|B_2| \;&\leq\; c\, E\Big(\Lambda_T(1+\|\mathbf{S}_T\|^q)\times ||\mathbf{S}_T\circ \mathbf{S}_0^{-1}||_1 \; I_{\{\mathbf{S}_T\in G\}}\Big) \\
&\leq\; c\, ||\mathbf{S}_0^{-1}||\times E\Big(\Lambda_T(1+\|\mathbf{S}_T\|^q)\times ||\mathbf{S}_T|| \; I_{\{\mathbf{S}_T\in G\}}\Big) \;=\; 0,
\end{aligned}$$

where the second inequality is due to the Cauchy-Schwarz inequality, and the equality holds by (A5') and the fact that the volume of the boundary set G defined in (3.2) is zero. Consequently,

$$B_2 = o_p(1), \text{ as } m \to \infty. \tag{S2.26}$$

Moreover, (S2.9) and (A3') yield that

$$\begin{aligned} ||\mathbf{Y}^{(2)}_{m,T}||_1 &\le c\ e^{rT}\ ||\mathbf{q}_{m,T}||_1 \times ||\mathbf{S}_0||_1 \times \Big|\Big|\frac{1}{m}\sum_{j=1}^m \Big(1 + \|\mathbf{W}^{(M)}_{j,1}\|^q\Big)\ \hat{\Lambda}_{j,T}\hat{\mathbf{S}}_{j,T}\Big|\Big|_1 \\ &= o_p(1), \end{aligned} \tag{S2.27}$$

as $m \to \infty$, where

$$\mathbf{q}_{m,T} = \sup_{\mathbf{x}=(1-d)e^{rT}\mathbf{S}_0 + d\bar{\mathbf{S}}^*_{m,T},\ 0\le d\le 1} |e^{-2rT}\mathbf{S}_0^{-2} - \mathbf{x}^{-2}|.$$

In addition, the last equality in (S2.27) holds by the LLN, (A5') and $\lim_{m\to\infty} ||\mathbf{q}_{m,T}||_1 = 0$, a.s., provided by $\lim_{m\to\infty} \bar{\mathbf{S}}^*_{m,T} = e^{rT}\mathbf{S}_0$. Therefore, by (S2.24)-(S2.27), (S2.23) holds. □

Lemma 2. *By using the same assumptions as in Theorem 2 (i), (S2.17) holds.*

Proof. By (S2.15),

$$\begin{aligned} \text{(II)} &= -\sqrt{m}(\bar{\Lambda}_{m,T} - 1)\Big(\frac{1}{m}\sum_{j=1}^m f(\hat{\mathbf{S}}_{j,T})\frac{\hat{\Lambda}_{j,T}}{(U^*_{j,m})^2}\Big) \\ &= -\sqrt{m}(\bar{\Lambda}_{m,T} - 1)E[f(\mathbf{S}_T)\Lambda_T] \\ &\quad +\sqrt{m}(\bar{\Lambda}_{m,T} - 1) \\ &\qquad \left(E[f(\mathbf{S}_T)\Lambda_T] - \frac{1}{m}\sum_{j=1}^m f(\hat{\mathbf{S}}_{j,T})\frac{\hat{\Lambda}_{j,T}}{(U^*_{j,m})^2}\right), \end{aligned} \tag{S2.28}$$

where $U^*_{j,m}$ is defined in (S2.14). By the CLT, $\sqrt{m}(\bar{\Lambda}_{m,T}-1)$ converges weakly to a normal random variable. By (S2.28) and using similar arguments in (S2.23)-(S2.27),

$$E[f(\mathbf{S}_T)\Lambda_T]-\frac{1}{m}\sum_{j=1}^{m} f(\hat{\mathbf{S}}_{j,T})\frac{\hat{\Lambda}_{j,T}}{(U^*_{j,m})^2}=o_p(1),$$

and thereby (S2.17) holds. □

Lemma 3. *By using the same assumptions as in Theorem 2 (i), (S2.18) holds.*

Proof. Note that

$$\begin{aligned}
\text{(III)} \;&:=\; \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\lim_{M\to\infty}\{\varphi_M(\mathbf{S}^*_{j,T})-\varphi_M(\hat{\mathbf{S}}_{j,T})\}\times(\Lambda^*_{j,T}-\hat{\Lambda}_{j,T})\\
&\leq\; \max_{j=1,\ldots,m}\left|\frac{\Lambda^*_{j,T}}{\hat{\Lambda}_{j,T}}-1\right|\\
&\quad\;\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\lim_{M\to\infty}\{\varphi_M(\mathbf{S}^*_{j,T})-\varphi_M(\hat{\mathbf{S}}_{j,T})\}\;\hat{\Lambda}_{j,T}. \qquad \text{(S2.29)}
\end{aligned}$$

By (S2.12), (S2.13), (S2.15) and (S2.16),

$$\begin{aligned}
&\frac{1}{\sqrt{m}}\sum_{j=1}^{m}\lim_{M\to\infty}\{\varphi_M(\mathbf{S}^*_{j,T})-\varphi_M(\hat{\mathbf{S}}_{j,T})\}\;\hat{\Lambda}_{j,T}\\
=\;& -\sqrt{m}(\bar{\mathbf{S}}_{m,T}-e^{rT}\bar{\Lambda}_{m,T}\mathbf{S}_0)\Phi+o_p(1). \qquad \text{(S2.30)}
\end{aligned}$$

Furthermore, by the CLT, $\sqrt{m}(\bar{S}_{i,m,T}-S_{i,0}e^{rT}\bar{\Lambda}_{m,T})$ converges weakly to a proper normal random variable for $i=1,\ldots,n$. Hence, by (S2.1), (S2.29) and (S2.30), (S2.18) holds. □

S2.4 Multi-asset Esscher transform for multivariate GARCH-N models

Let $R_{i,t} = \log(S_{i,t}/S_{i,t-1})$ and $\mu_{i,t} = r - 0.5\sigma_{i,t}^2 + \lambda_i \sigma_{i,t}$. From model (S1.2), we have $\mathbf{R}_t | \mathcal{F}_{t-1} \sim N(\boldsymbol{\mu}_t, \Sigma_t)$, where $\mathbf{R}_t = (R_{1,t}, \ldots, R_{n,t})^\top$, $\boldsymbol{\mu}_t = (\mu_{1,t}, \ldots, \mu_{n,t})^\top$, $\Sigma_t = (\sigma_{ijt})$, $\sigma_{ijt} = \rho_{i,j}\sigma_{i,t}\sigma_{j,t}$ and $\mathcal{F}_{t-1}$ denotes the set of information from time 0 up to time $t-1$. Let $f_{t-1}(\mathbf{R}_t)$ denote the conditional density function of $\mathbf{R}_t$ given $\mathcal{F}_{t-1}$ under the P measure and the corresponding conditional moment generating function (m.g.f.) is

$$M_{t-1}(\boldsymbol{\delta}) = \exp\left\{\boldsymbol{\delta}^\top \boldsymbol{\mu}_t + \frac{1}{2}\boldsymbol{\delta}^\top \Sigma_t \boldsymbol{\delta}\right\}, \tag{S2.31}$$

where $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_n)^\top$. By the Esscher transform, let

$$f_{t-1}^Q(\mathbf{R}_t) = \frac{e^{\mathbf{s}^\top \mathbf{R}_t}}{M_{t-1}(\mathbf{s})} f_{t-1}(\mathbf{R}_t) \tag{S2.32}$$

denote the conditional density function of $\mathbf{R}_t$ given $\mathcal{F}_{t-1}$ under a Q measure by introducing an $n \times 1$ vector $\mathbf{s}$. To choose a vector $\mathbf{s}$ in (S2.32) such that Q is a risk-neutral measure, the following martingale identity for each underlying asset has to be satisfied under the Q measure, that is, $E_{t-1}^Q(e^{R_{i,t}}) = e^r$ for $i = 1, \ldots, n$.

By (S2.32), the conditional m.g.f. of $\mathbf{R}_t$ given $\mathcal{F}_{t-1}$ under the Q measure can be represented as

$$M_{t-1}^Q(\boldsymbol{\delta}) = \frac{M_{t-1}(\boldsymbol{\delta} + \mathbf{s})}{M_{t-1}(\mathbf{s})}. \tag{S2.33}$$

To solve $\mathbf{s}$, first let $\mathbf{e}_i = (0, \ldots, 1, \ldots, 0)^\top$, whose elements are all zero except for the ith element being 1. Hence, we have $E^Q_{t-1}(e^{R_{i,t}}) = M^Q_{t-1}(\mathbf{e}_i)$ and thereby (S2.31) and (S2.33) yield

$$M^Q_{t-1}(\mathbf{e}_i) = \exp\left\{\mathbf{e}_i^\top \boldsymbol{\mu}_t + \frac{1}{2}\mathbf{e}_i^\top \Sigma_t \mathbf{e}_i + \mathbf{e}_i^\top \Sigma_t \mathbf{s}\right\}. \tag{S2.34}$$

By (S2.34), the identities of $M^Q_{t-1}(\mathbf{e}_i) = E^Q_{t-1}(e^{R_{i,t}}) = e^r$ is equivalent to $\mu_{i,t} + \frac{1}{2}\sigma^2_{i,t} + (\sigma_{i1t}, \ldots, \sigma_{int})\mathbf{s}= r$. Furthermore, since $\mu_{i,t} = r - 0.5\sigma^2_{i,t} + \lambda_i\sigma_{i,t}$, the last identity can be represented as $(\sigma_{i1t}, \ldots, \sigma_{int})\mathbf{s} = -\lambda_i\sigma_{i,t}$, which is a linear equation of $\mathbf{s}$, for $i = 1, \ldots, n$. By rewriting the n linear equations in matrix, we have

$$\mathbf{s} = -\Sigma_t^{-1}\mathrm{D}_{\boldsymbol{\lambda}}\boldsymbol{\sigma}_t, \tag{S2.35}$$

where $\mathrm{D}_{\boldsymbol{\lambda}}$ is a diagonal matrix with the ith diagonal component being λ_i and $\boldsymbol{\sigma}_t = (\sigma_{1,t}, \ldots, \sigma_{n,t})^\top$. By substituting ($S2.35$) into equation ($S2.33$), the conditional m.g.f. of $\mathbf{R}_t$ given $\mathcal{F}_{t-1}$ under the Q measure is

$$M^Q_{t-1}(\boldsymbol{\delta}) = \exp\left\{\boldsymbol{\delta}^\top(\boldsymbol{\mu}_t - \mathrm{D}_{\boldsymbol{\lambda}}\boldsymbol{\sigma}_t) + \frac{1}{2}\boldsymbol{\delta}^\top\Sigma_t\boldsymbol{\delta}\right\}$$

which is the m.g.f. of $N\left(\boldsymbol{\mu}_t - \mathrm{D}_{\boldsymbol{\lambda}}\boldsymbol{\sigma}_t, \Sigma_t\right)$. Therefore, we have $R_{i,t}|\mathcal{F}_{t-1} \sim N(r - 0.5\sigma^2_{i,t}, \sigma^2_{i,t})$. As a result, the risk-neutral counterpart of model (S1.2) derived by the multi-asset Esscher transform is

$$\begin{cases} R_{i,t} = r - \frac{1}{2}\sigma^2_{i,t} + \sigma_{i,t}\epsilon^*_{i,t} \\ \sigma^2_{i,t} = \beta_0 + \beta_1\sigma^2_{i,t-1} + \beta_2\sigma^2_{i,t-1}(\epsilon^*_{i,t-1} - \lambda_i)^2 \end{cases},$$

where $\epsilon^*_{i,t} = \epsilon_{i,t} + \lambda_i$ is $N(0,1)$ distributed under the Q measure.

In addition, the Radon-Nykodým derivative of the measure Q with respect to the measure P is

$$\Lambda_T = \frac{f^Q_{T-1}(\mathbf{R}_T)}{f_{T-1}(\mathbf{R}_T)} \times \frac{f^Q_{T-2}(\mathbf{R}_{T-1})}{f_{T-2}(\mathbf{R}_{T-1})} \times \ldots \times \frac{f^Q_0(\mathbf{R}_1)}{f_0(\mathbf{R}_1)} = \prod_{t=1}^{T} \Lambda_t,$$

where

$$\Lambda_t = \frac{f^Q_{t-1}(\mathbf{R}_t)}{f_{t-1}(\mathbf{R}_t)} = \exp\left\{-\boldsymbol{\sigma}_t^\top \mathrm{D}_{\boldsymbol{\lambda}} \Sigma_t^{-1} \Big(\mathbf{R}_t - \boldsymbol{\mu}_t + \frac{1}{2}\mathrm{D}_{\boldsymbol{\lambda}}\boldsymbol{\sigma}_t\Big)\right\},$$

and (S1.3) holds. $\square$

References

Black, F. and Scholes, M. (1973). Efficient minimum distance estimation with multiple rates of convergence. *J. Political Economy* **81**, 637-654.

Cox, J. C., Ingersoll, J. E. and Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica* **53**, 385-407.

Duan, J. C. (1995). The GARCH option pricing model. *Mathematical Finance* **43**, 13-32.

Duan, J. C. and Simonato, J. G. (1998). Empirical martingale simulation for asset prices. *Management Science* **44**, 1218-1233.

Duistermaat, J. J. and Kolk, J. A. C. (2004). *Multidimensional Real Analysis II, Integration.* Cambridge, New York.

Gerber, H. U. and Shiu, S. W. (1994). Option pricing by Esscher transform. *Transactions of Society of Actuaries* **46**, 99-191.

Glasserman, P., Heidelberger, P. and Shahabuddin, P. (1999). Asymptotically optimal importance samling and stratification for pricing path-dependent options. *Mathematical Finance* **9**, 117-152.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bonds and currency options. *The Review of Financial Studies* **6**, 327-343.

Huang, S. F. (2014). A modified empirical martingale simulation for financial derivative pricing. *Communications in Statistics - Theory and Methods* **43**, 328-342.

Huang, S. F. and Guo, M. H. (2009). Financial derivative valuation - A dynamic semiparametric approach. *Statistica Sinica* **19**, 1037-1054.

Huang, S. F. and Tu, Y. T. (2014). Asymptotic distribution of the EPMS estimator for financial derivatives pricing. *Computational Statistics and Data Analysis* **73**, 129-145.

Huang, S. F. and Yu, J. F. (2012). Hedging rainbow options in discrete time. *J. Chinese Statistical Association* **50**, 1-20.

Hull, J. and White, A. (1987). The pricing of options on assets with stochastic volatilities. *J. Finance* **42**, 281-300.

Kijima, M. (2006). A multivariate extension of equilibrium pricing transforms: The multivariate Esscher and Wang transforms for pricing financial and insurance risks. *ASTIN Bulletin* **36**, 269-283.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *J. Financial Economics* **3**, 125-144.

Shreve, S. E. (2004). *Stochastic Calculus for Finance II, Continuous-Time Models*. Springer, New York.

Timan, A. F. (1963). *International Series of Monographs in Pure and Applied Mathematics*. Macmillan Company, New York.

Zhu, H., Ye, F. and Zhou, E. (2015). Fast estimation of true bounds on Bernudan option prices under jump-diffusion processes. *Quantitative Finance* **15**, 1885-1900.