OPTIMAL PARTIAL PROXY METHOD FOR COMPUTING GAMMAS OF FINANCIAL PRODUCTS WITH DISCONTINUOUS AND ANGULAR PAYOFFS

MARK S. JOSHI AND DAN ZHU

Abstract. We extend the limit optimal partial proxy method to compute second order sensitivities of financial products with discontinuous or angular payoffs by Monte Carlo simulation. The methodology is optimal in terms of minimizing the variance of likelihood ratios terms. Applications are presented for both equity options and interest rate products with discontinuous payoff structures. The first order optimal partial proxy method is also implemented to calculate the Hessians of insurance products with angular payoffs. Numerical results are presented which demonstrate the speed and efficacy of the method.

1. Introduction

Financial institutions have significant exposures to derivative products with complex structures. For example, insurance companies sell products with embedded options as part of their guarantee to the insured, banks use derivatives to protect themselves from interest-rate changes, and hedge funds use derivatives to manage investment risks or to speculate on the future direction of markets. Fluctuations in the underlying assets' prices will affect derivatives' prices directly, which may materially impact on the balance sheet of the financial institution. Therefore it is important for the financial institutions to manage the risks underlying their derivative exposures. In order to implement a hedging strategy, we need to know the Greeks, i.e. the price sensitivities of the derivative with respect to the parameters of interest. For a delta-hedging strategy, Deltas (the price sensitivity with respect to the underlying) indicate the number of units of the underlying securities to hold in the hedged position, and Gammas (the second derivative of the option price with respect to the underlying) are used to determine the optimal time interval required for rebalancing under transaction costs (Broadie and Glasserman, 1996).

To hedge derivative products with multiple underlying assets, we need to calculate the Gamma matrix, the Hessian (the square matrix of second-order partial derivatives of a function). If we write the price of an option as a function of θ and ψ

$$v(\theta, \psi) = \int g(x, \theta) f(x, \psi)\, dx, \tag{1.1}$$

where $g(x, \theta)$ is the discounted payoff function of the option, and $f(x, \psi)$ is the probability density function of the random variable X, the Hessian of v describes the local curvature of the price. Under a delta-hedged position, the first-order derivatives $\frac{\partial v}{\partial \theta}$ and $\frac{\partial v}{\partial \psi}$ are cancelled out; the Hessian of the price provides the local explanation of profit and loss movements of the delta-hedging strategy.

We can compute price sensitivities directly by differentiating the analytic formula of the option price, when it exists. Otherwise, Monte-Carlo simulation is often used. Traditionally, there are three methods of computing Greeks: finite differencing, the likelihood ratio, and the pathwise method (Glasserman, 2004). The three traditional methods as discussed by Glasserman, are subject to limitations:

- the finite differencing method produces biased estimators;

- the pathwise method produces unbiased estimators with the smallest variances among the three, but it has limited applications when the payoff of the product is discontinuous or angular;
- the likelihood ratio method also produces unbiased estimators, but the variance of the Greek estimator increases linearly with the number of random variables which the parameter of interest has influence upon. This method is not applicable when the density function of the random variable is singular.

The computation of first-order derivatives is problematic when the discounted payoff function is discontinuous. Pathwise discontinuities exclude the application of the pathwise method, and the use of the likelihood ratio method is likely to result in an estimator with a large variance.

To resolve the problem of first-order Greeks, Glasserman (2004) suggested approximating the discontinuity by a smooth function. Hong and Liu (2011) purposed a methodology to approximate the discontinuity by a kernel distribution supported in a small neighborhood; their method first decomposes the derivative into a combination of ordinary expectation and a derivative with respect to an auxiliary parameter. The ordinary expectation is then estimated using the sample mean, and the derivative part is estimated using kernel smoothing. However, the estimators produced by these two methods are biased. Measure-valued differentiation is used to estimate gradients by differentiating the measures (Heidergott and Leahu, 2010). The method requires sampling different versions of the random variables, which may lead to a considerable computational effort. Giles(2008) suggested the idea of vibrato Monte Carlo, which is a generalisation of the use of conditional expectation for payoff smoothing. This is a hybrid method for calculating sensitivities, applying the pathwise method to the path simulation and the likelihood ratio method to the payoff evaluation. Other solutions using Malliavin calculus have been proposed by Benhamou (2003). However, these methods suffer the same limitations as the likelihood ratio method.

Glasserman (2004) also suggested using the likelihood ratio near the discontinuity and the pathwise method away from it. Wang et al(2012) adopted this approach, by using a specific change of variable that targets discontinuities caused by indicator functions. Chan and Joshi (2012) took this idea to the extreme. They introduced a measure change on one standard uniform random variable each step, which eliminates the pathwise discontinuity. The variance of the likelihood ratio part of their first-order Greek estimator is minimal among all such measure changes with the requisite properties. Their method is named the optimal partial proxy scheme(OPP).

Estimating second-order derivatives is far more difficult than estimating first-order derivatives. The pathwise method is rarely applicable since it requires the first derivatives of the payoff function to be Lipschitz continuous everywhere and differentiable almost surely; the likelihood ratio method, on the other hand is often applicable if the density function is non-singular, but it results in even larger variances for second-order derivative estimators. Glasserman (2004) suggested using a combination of the pathwise and the likelihood ratio method to produce better estimators for Gammas than using the pure likelihood ratio method, i.e. the pathwise likelihood ratio method. This approach, however, still suffers from the same problem as the likelihood ratio method for first-order Greeks and it is not applicable when the payoff function is discontinuous.

Chan and Joshi(2012) showed that the OPP method is effective for computing first-order Greeks of financial products with discontinuous payoffs. The extension of this method to computing second-order Greeks such as Gammas and Cross-gammas is the main focus of this paper. The approach is to perform a change of variable to eliminate the limits' dependence on the parameter of interest, θ, so pathwise discontinuities of the price and the first order derivatives of it are both eliminated. The resulting discounted payoff function is twice differentiable almost everywhere, and has Lipschitz continuous first-order derivatives. Here, we shall say a function is $\hat{C}^2$ if it satisfies such conditions. We can then change the order of differentiation and integration to calculate the pathwise estimator of the Hessian. The change of variable is also optimal in terms of minimizing the variance of the likelihood ratio part of second-order derivatives. We call this method the second-order optimal partial proxy method for calculating Hessians(HOPP(2)).

Our simulation scheme is defined in terms of the discontinuities of the discounted payoff function. The change of variable is performed at every step on the first random uniform $v_{i,1}$; it is replaced by a function $U_i(\theta, v_{i,1})$ defined in terms of the critical values of $v_{i,1}$ for the discounted payoff function to cross the discontinuities. Therefore, we need to find these critical values of $v_{i,1}$ explicitly. For example, for a digital basket option, a strike crossing is determined by the average of a basket of stocks in the Black- Scholes model; the critical value is the value of the first uniform such that the average of the basket is equal to the strike. For such cases where we are unable to find closed form solutions of these critical values, we use the Newton-Raphson method with two steps to numerically approximate them. However, when the approximated critical values are used in the finite differencing HOPP(2) scheme, there is still a small probability for the bumped path and the unbumped path to finish on different sides of discontinuities. To remove such discontinuities, we use the unbumped path event for the bumped path. We show that the error produced is of order $||\theta - \theta_0||^4$, which is sufficient for the bias to vanish at the limit as $\theta \to \theta_0$ for the Hessian.

Derivative products with angular payoffs are different from derivative products with discontinuous payoffs. Their payoff functions are Lipschitz continuous everywhere and differentiable almost surely, so the pathwise method is applicable to calculate the first-order derivatives but not the second-order ones. We introduce a change of variable function which is affine in θ to remove the pathwise discontinuity for the first-order derivatives of the product. We call this method, the first-order optimal partial proxy method for calculating Hessians(HOPP(1)).

We apply the HOPP(2) and the HOPP(1) schemes to calculate the Hessian of both equity and interest rate derivatives. For high-dimensional products, the computational burden is another important issue to consider when calculating price sensitivities. The LIBOR Market Model of interest-rate derivatives is one such example, where multi-dimensional evolutions are involved in Monte-Carlo simulation (Brace et al, 1997). When the drift is calculated using the method introduced by Joshi(2003), the computational order of calculating the price under the LMM is $\mathcal{O}(nF)$ per step, where n is the number of underlying forward rates and F is the number of Brownian motions driving the evolution of the forward rates. For an interest-rate derivative with 30 underlying forward rates, the Hessian matrix has 465 different entries. It will take a substantial amount of time to compute the Hessian using the finite differencing approach. To reduce the computational complexity of these cases, Giles and Glasserman(2006) applied the adjoint and automatic differentiation method to calculating price sensitivities of financial products. Joshi and Yang (2010) introduced a methodology for computing the Hessian by decomposing the algorithm into elementary operations, then calculate the pathwise estimator of the Hessian using a backwards method. Here, HOPP(2) and HOPP(1) are combined with the algorithmic Hessian, to produce a simulation scheme for calculating the Hessian of derivative products with discontinuous or angular payoffs. The computational order for a path is $\mathcal{O}(nFMI)$ for calculating the Hessian of Target redemption notes, where M is the number of state variables and I is the number of observation dates up to and including the redemption.

In addition to derivative products, we also apply our simulation scheme to compute second-order derivatives of insurance contracts, since sensitivity analysis is one of the key risk management practices in pricing and reserving of insurance contracts. There is an extensive amount of insurance literature focusing on the first-order sensitivities. Insurance contracts often provide guaranteed minimum benefits in an event of death, ill, or withdrawal, these guarantees represent embedded derivatives in the liabilities that are often complex, path dependent options. A pathwise approach was introduced by Hobbs et al (2009) to provide unbiased sensitivity estimators of these types of products. Valuation of pension funds are subject to actuarial assumptions such as mortality rates, interest rates and levels of salary inflation, Joshi and Pitt (2010) applied the adjoint method to assess sensitivities of valuation results for pension funds to these key inputs. However, the Global

Financial Crisis that started in 2008 alerted the insurance industry to the crucial importance of higher order sensitivities in their hedging programs. Here, we apply the HOPP(1) scheme to calculate the Hessian of a portfolio insurance product, the North Guarantee, which has an angular payoff as a function of the underlying stocks.

The remaining sections of the report are organized as follows. The basic idea of HOPP(2) is presented in Section 2, and examples of the method applied to the Black-Scholes model are presented. In Section 3, we discuss the cases where the discontinuities are not explicitly computable. In Section 4, we apply the model to calculate the Hessian of financial products with angular payoffs using HOPP(1), numerical examples are shown for the vanilla put option and the North Guarantee. We then apply HOPP(2) to the LIBOR market model in Section 5, to calculate the second order Greeks of various vanilla and exotic style interest rate derivatives.

2. The Optimal Partial Proxy Method For Second-Order Greeks

2.1. **One-step and one-dimensional case.** In this section, we present the one-step and one-dimensional case because it is simple, yet illustrates many important features of HOPP(2). A simple discontinuous payoff can be expressed as

$$I_{\alpha_0 \le S \le \alpha_1} g(S),$$

where g is the twice-differentiable discounted payoff function of S.

The one-step and one-dimensional integral can be viewed as a one-dimensional Monte-Carlo simulation over $u \in U(0,1)$,

$$P(\theta) = \int_0^1 I_{[\alpha_0,\alpha_1]}\big(f(\theta,u)\big) g\big(f(\theta,u)\big)\, du. \tag{2.1}$$

Here f is the algorithm for turning a random uniform u and a parameter θ into the state variable S. We assume that there exist twice-differentiable functions $a_j(\theta)$, for $j = 0$, 1, such that

$$f(\theta,u) \in [\alpha_0, \alpha_1] \Leftrightarrow u \in \big[a_0(\theta), a_1(\theta)\big]. \tag{2.2}$$

The above integral becomes

$$F(\theta) = \int_{a_0(\theta)}^{a_1(\theta)} g\big(f(\theta,u)\big)\, du. \tag{2.3}$$

To remove the limits' dependence on θ for both F and F', we introduce a change of variable by replacing u with $u(v,\theta)$ in equation (2.3), where

$$u(v,\theta) = v + (\theta-\theta_0)\gamma(v) + \frac{1}{2}(\theta-\theta_0)^2\delta(v), \tag{2.4}$$

with conditions

(1) for F, we require $\gamma\big(a_j(\theta_0)\big) = a_j'(\theta_0)$ for $j = 0, 1$,
(2) for F', we require $\delta\big(a_j(\theta_0)\big) = a_j''(\theta_0)$ for $j = 0, 1$.

The change of variable function is quadratic in θ, since we consider only the first and second derivatives of the function. The conditions above guarantee the limits to not move with θ up to third order, i.e.

$$\frac{\partial u(a_j(\theta_0),\theta_0)}{\partial\theta} = a_j'(\theta_0),$$
$$\frac{\partial^2 u(a_j(\theta_0),\theta_0)}{\partial\theta^2} = a_j''(\theta_0).$$

We can now expanding the integral using Taylor's theorem up to third order

$$F(\theta) = \int_{a_0(\theta_0)}^{a_1(\theta_0)} g\Big(f\big(\theta, u(v,\theta)\big)\Big) \frac{\partial u(v,\theta)}{\partial v}\, dv. \tag{2.5}$$

The derivative evaluated at θ_0 is

$$F'(\theta_0) = \int_{a_0(\theta_0)}^{a_1(\theta_0)} g'\big(f(\theta_0, v)\big)\Upsilon(\theta_0, v) + g(f(\theta_0, v))\Omega(\theta_0, v)\, dv, \tag{2.6}$$

with

$$\Upsilon(\theta_0, v) = \frac{\partial}{\partial\theta}\Big[f\big(\theta, u(v,\theta)\big)\Big]_{|\theta=\theta_0} = \frac{\partial f(\theta_0, u)}{\partial\theta} + \frac{\partial f(\theta_0, u)}{\partial u}\gamma(v),$$

$$\Omega(\theta_0, v) = \frac{\partial^2 u(\theta_0, v)}{\partial\theta\partial v} = \gamma'(v).$$

The expression in equation (2.6) agrees with the one derived for first-order derivatives in the original OPP paper (Chan and Joshi, 2012). We further differentiate this expression to obtain the second-order derivative

$$\begin{aligned} F''(\theta_0) = \int_{a_0(\theta_0)}^{a_1(\theta_0)} & g''\Big(f\big(\theta_0, U(v,\theta_0)\big)\Big)\Upsilon^2 + g'\Big(f\big(\theta_0, U(v,\theta_0)\big)\Big)\eta(\theta_0, v) \\ & + g\Big(f\big(\theta_0, U(v,\theta_0)\big)\omega(\theta_0, v)\, dv, \end{aligned} \tag{2.7}$$

with

$$\omega(\theta_0, v) = \frac{\partial^3 U(v,\theta_0)}{\partial\theta^2\partial v} = \delta'(v).$$

and

$$\eta(\theta_0, v) = \lambda(\theta_0, v) + \upsilon(\theta_0, v),$$

where

$$\lambda(\theta_0, v) = \frac{\partial^2 f(\theta_0, u)}{\partial\theta^2} + 2\frac{\partial^2 f(\theta_0, u)}{\partial\theta\partial u}\gamma(v) + \frac{\partial^2 f(\theta_0, u)}{\partial u^2}(\gamma(v))^2 + \frac{\partial f\big(\theta_0, U(v,\theta_0)\big)}{\partial u}\delta(v),$$

$$\upsilon(\theta_0, v) = 2\frac{\partial f\big(\theta_0, U(v,\theta_0)\big)}{\partial\theta}\frac{\partial^2 u(\theta_0, v)}{\partial\theta\partial v} = 2\Upsilon(\theta_0, v)\gamma'(v),$$

$$\omega(\theta_0, v) = \frac{\partial^3 U(v,\theta_0)}{\partial\theta^2\partial v} = \delta'(v).$$

The term $\frac{\partial u(\theta_0, v)}{\partial v}$ is dropped from equations (2.6) and (2.7) since it is one at the limit.

The unity of the pathwise and the likelihood ratio methods is discussed by L'Ecuyer(1990), our approach can be viewed as a combination of the pathwise method and the likelihood ratio method. The change of variable performed is to shift some dependence of θ from the discounted payoff function to the probability density function, so we can apply the pathwise method to the payoff function away from the discontinuities and the likelihood ratio method near the discontinuities.

2.1.1. *Optimization.* The mild conditions on $\gamma(v)$ and $\delta(v)$ give us the flexibility to choose these functions in a way that minimizes the variance of the Monte-Carlo implementation. The pathwise method when applicable produces a lower variance than the likelihood ratio method; therefore it is preferable to use as much pathwise method as possible. The functions γ and δ can be chosen to minimize the variance of the likelihood ratio part, i.e. the variances of η and ω in equation (2.7). This is however dependent on the structure of the discounted payoff function, Chan and Joshi (2012) presented a special case where the continuous part of the payoff function is constant, and the optimal choice for γ is a piecewise linear function for computing first-order derivatives.

Here, we show that affine functions are also optimal for computing second-order derivatives of the expectation of piecewise constant functions. Under the HOPP(2) scheme, when the integrand is a function being constant between α_0 and α_1, and zero otherwise, the second derivative of the integral can be expressed as

$$F''(\theta_0) = \int_{a_0(\theta_0)}^{a_1(\theta_0)} g\Big(f\big(\theta_0, U(v, \theta_0)\big)\Big)\omega(\theta_0, v)dv.$$

Letting K be the constant value g takes between α_0 and α_1, the equation can be simplified to

$$F''(\theta_0) = K \int_{a_0(\theta_0)}^{a_1(\theta_0)} \delta'(v)dv.$$

Our objective is to choose $\delta(v)$ in such way so the variance of the simulated second-order derivative estimator is minimized, that is to minimize

$$\int_{a_0(\theta_0)}^{a_1(\theta_0)} \delta'(v)^2 dv.$$

Chan and Joshi (2012) showed that, given that δ' is continuously differentiable on $[a_0(\theta_0), a_1(\theta_0)]$ except at a finite number of points where it is continuous, this expression is minimized globally when δ is linear between $a_0(\theta_0)$ and $a_1(\theta_0)$. With given values of the function δ at $a_0(\theta_0)$ and $a_1(\theta_0)$, we have that the optimal δ is

$$\delta(v) = a_0''(\theta_0) + \frac{a_1''(\theta_0) - a_0''(\theta_0)}{a_1(\theta_0) - a_0(\theta_0)}\big(v - a_0(\theta_0)\big). \tag{2.8}$$

Here, we adopt the same linear function for γ as the first-order OPP scheme,

$$\gamma(v) = a_0'(\theta_0) + \frac{a_1'(\theta_0) - a_0'(\theta_0)}{a_1(\theta_0) - a_0(\theta_0)}\big(v - a_0(\theta_0)\big). \tag{2.9}$$

In general, we may have a discounted payoff function with several discontinuities. For m discontinuities, we assume there exist a set of twice-differentiable functions $a_j(\theta)$ for $j = 0, 1, \ldots, m+1$, such that

$$a_0(\theta) = 0 \text{ and } a_{m+1}(\theta) = 1,$$

and for $i = 0, 1, \ldots, m$,

$$a_i(\theta) < a_{i+1}(\theta),$$

where $a_i(\theta)$ is the ith strata function corresponding to the ith discontinuity of the payoff, a_i, as shown in equation (2.1.2). We assume the discounted payoff function is $\hat{C}^2$ on (a_i, a_{i+1}) and will be $\hat{C}^2$ extendible to the closed intervals. HOPP(2) is applicable to discounted payoff functions with multiple discontinuities, by having different δ and γ functions for each interval $(a_i(\theta_0), a_{i+1}(\theta_0)]$ as shown in equation (2.8) and (2.9).

2.1.2. *One-dimensional and one-step examples.* The first example we present here is the digital call option. The analytical formula for the Gamma is:

$$-\frac{e^{-rT}\phi(d_2)}{S_0^2\sigma\sqrt{T}}\Big(1 - \frac{d_2}{\sigma\sqrt{T}}\Big),$$

where ϕ is the standard normal probability density function, and

$$d_2 = \frac{\log(\frac{S_0}{K}) + (r - 0.5\sigma^2)T}{\sigma\sqrt{T}}.$$

The HOPP(2) method presented in this section produces an unbiased estimator of the Gamma. From equation (2.7), the HOPP(2) Gamma for a digital Call is given by

$$\mathbb{E}\big[I_{S_T > K}e^{-rT}\omega(S_0, v)\big],$$

where

$$\omega(S_0, v) = \delta'(v) = -\frac{\Phi''(d_2)(\frac{\partial d_2}{\partial S_0})^2 + \Phi'(d_2)\frac{\partial^2 d_2}{\partial S_0^2}}{\Phi(d_2)},$$

where Φ is the standard normal cumulative density function.

The HOPP(2) Gamma of a digital call option becomes

$$e^{-rT}\delta'(v)\mathbb{P}(S_T > K) = \Big[\Phi''(d_2)\Big(\frac{\partial d_2}{\partial S_0}\Big)^2 + \Phi'(d_2)\frac{\partial^2 d_2}{\partial S_0^2}\Big]e^{-rT},$$

this expression agrees with the analytical formula.

We present applications of the HOPP(2) to vanilla equity products, namely, the digital put option, the put option and the parabolic put option. Here, the payoff of a parabolic put option is defined as

$$\Big((K - S_T)_+\Big)^2$$

where K is the strike of the option and S_T is the stock price at the expiry. This hypothetical product has a discounted payoff function, which is $\hat{C}^2$ as a function of the underlying stock value. The pathwise method is applicable to this product to calculate the Gammas, which allows the comparison of the HOPP(2) and the pure pathwise method. For the Monte-Carlo simulation, we set $S_0 = 100$, $K = 100$, $\sigma = 0.2$ and $r = 0.05$ per annum. The expiry time is 0.2 years.

The standard errors generated by the HOPP(2) method are compared with the standard errors from the likelihood ratio, the pathwise likelihood ratio and the pure pathwise method based on a 20,000 paths sample, see Table 1 and 2.

Option	HOPP(2) Mean	HOPP(2) S.E.	LR S.E	PWLR S.E.	PW S.E.
Parabolic Delta	-5.5505	0.0788	0.1248		0.04
Parabolic Gamma	0.8221	0.0099	0.0326	0.012	0.0044
Put Delta	-0.4377	0.0044	0.0069		0.0033
Put Gamma	0.044	0.0004	0.0016	0.0004	
Digital Delta	-0.0439	0.0003	0.0005		
Digital Gamma	0.0008	2.61594E-06	1.77493E-05		

Table 1. Vanilla options: means and standard errors of the Delta and Gamma estimates by various methods with 20,000 paths samples

Option	LR/HOPP(2)	PWLR/HOPP(2)	PW/HOPP(2)
Parabolic Delta	1.585		0.5074
Parabolic Gamma	3.289	1.211	0.4406
Put Delta	1.583		0.7477
Put Gamma	4.015	1.083	
Digital Delta	1.397		
Digital Gamma	15.631		

Table 2. Vanilla options: standard errors of the Delta and Gamma estimates by various methods divided by standard errors by the HOPP(2) with 20,000 paths samples

We observe that the HOPP(2) method outperforms the likelihood ratio method and the pathwise likelihood ratio method. We perform further investigations on the performance of the four methods with different volatilities. We keep all the parameters the same as the above example, except the volatility. The volatility is varying from 0.1 to 0.5. The comparison of the standard errors of the four methods are illustrated in Figure 1.

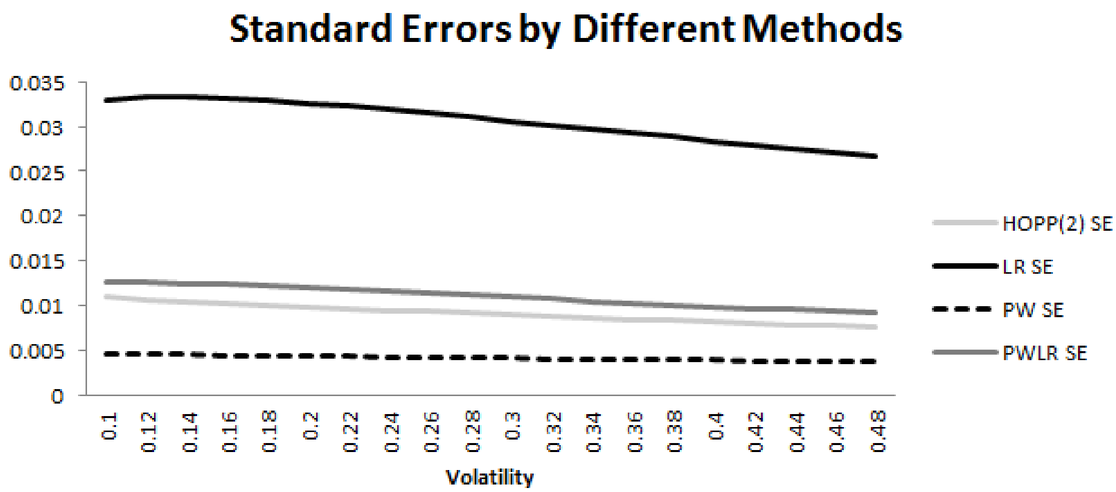

FIGURE 1. Parabolic Put Option: Comparison of standard errors of the Gamma calculated by various methods with varying volatility using 20,000 paths samples

2.2. **Extension to multiple step cases.** To calculate second-order derivatives of the price for products with path-dependent features by HOPP(2), we perform a change of variable at each step where there exist pathwise discontinuities. Here, we have a set of N time steps, $T_1, T_2, \ldots, T_N$, and $S_k \in \mathbb{R}$ is the value of the state variable observed at T_k. The change of variable performed on each step is to remove pathwise discontinuities of both the discounted payoff function and the second order derivative of it. We replace the simulated standard uniform v_i of step i, with the change of variable function $u_i(v_i, \theta)$ as shown in Section 2.1, which is:

- continuously twice differentiable as a function of θ,
- and piecewise differentiable as a function of v_i.

The resulting discounted payoff of the path is,

$$\hat{P}(\theta) = \prod_{i=1}^{N} \frac{\partial u_i(\theta, V_i)}{\partial v} P\Big(S\big(\theta, U(\theta, V)\big)\Big), \tag{2.10}$$

with the properties:

- $\hat{P}(\theta_0) = P(S(\theta_0, V))$,
- $\mathbb{E}[\hat{P}(\theta)] = \mathbb{E}[P(S(\theta, V))]$,
- $\hat{P}$ is $\hat{C}^2$ as a function of θ.

We differentiate the discounted payoff of the path in equation (2.1) to obtain the first-order Greek estimator under the HOPP(2) simulation scheme,

(2.2.11)

$$\hat{P}'(\theta_0) = \sum_{k=1}^{N} \frac{\partial P\Big(S\big(\theta_0, U(\theta_0, V)\big)\Big)}{\partial S_k} \prod_{i=1}^{k} \Big(\frac{\partial S_i(S_{i-1}(\theta_0), V_i)}{\partial S_{i-1}} + \frac{\partial S_i\big(S_{i-1}(\theta_0), u_i(\theta_0, V_i)\big)}{\partial u_i} \frac{\partial u_i(\theta_0, V_i)}{\partial \theta} \Big) \prod_{j=1}^{N} \frac{\partial u_j(\theta_0, V_j)}{\partial v}$$

$$+P\big(S(\theta_0,V)\big)\sum_{i=1}^{N}\frac{\partial^2 u_i(\theta_0,V_i)}{\partial v\partial\theta}\prod_{j\neq i}\frac{\partial u_j(\theta_0,V_j)}{\partial v}.$$

We further differentiate this expression with respect to θ to obtain the second derivative,

$$\hat{P}''(\theta_0)=\sum_{s=1}^{N}\sum_{k=1}^{N}\frac{\partial^2 P\Big(S\big(\theta_0,U(\theta_0,V)\big)\Big)}{\partial S_k\partial S_s}\prod_{i=1}^{k}\Upsilon_i(\theta_0,V_i)\prod_{i=1}^{s}\Upsilon_i(\theta_0,V_i) \tag{2.2.12}$$

$$+\sum_{k=1}^{N}\frac{\partial P\Big(S\big(\theta_0,U(\theta_0,V)\big)\Big)}{\partial S_k}\Big[\sum_{i=1}^{k}\lambda_i(\theta_0,V_i)\prod_{j\neq i}\Upsilon_j(\theta_0,V_j)+2\prod_{i=1}^{k}\Upsilon_i(\theta_0,V_i)\sum_{i=1}^{k}\Omega_i(\theta_0,V_i)\Big]$$

$$+P\big(S(\theta_0,V)\big)\Big[\sum_{i=1}^{N}\omega_j(\theta_0,V_j)+\sum_{i=1}^{N}\Omega_i(\theta_0,V_i)\sum_{j\neq i}\Omega_j(\theta_0,V_j)].$$

The functions Υ_i,λ_i,Ω_i and ω_i are as defined in section 3.1.

We apply HOPP(2) to calculate the Gamma of the discrete barrier option and the digital barrier option. For the Monte-Carlo simulation, we compute the Gammas and Deltas for the down-and-out put options with initial stock price and strike 100, and barrier 80. The volatility of the stock is 0.5 and the interest rate is 0.05 per annum. The expiry time is in 1 year with 10 observation dates equally-spaced across $[0,T]$. The discrete barrier option payoffs are path-dependent and discontinuous, as a consequence, the pathwise method is not applicable to these products.

The results from HOPP(2) and the likelihood ratio method converge. The standard errors and the ratio of standard errors are presented in Table 3 based on a sample of 500,000 paths.

Option	HOPP(2) Mean	LR Mean	HOPP(2)S.E.	LR S.E.	LR S.E./HOPP(2) S.E.
Digital Barrier Delta	0.002	0.002	1.45028E-05	2.33001E-05	1.607
Digital Barrier Gamma	-9.359E-05	-9.184E-05	9.52616E-07	1.86469E-06	1.957
Barrier Delta	0.027	0.027	0.000240	0.0003799	1.582
Barrier Gamma	-0.001	-0.001	1.4734E-05	3.00469E-05	2.039

Table 3. Down and Out Put options: Mean and standard errors of the Delta and Gamma estimates by various methods on samples with 500,000 paths.

One important drawback of the likelihood ratio method is its dependence on the volatility of the underlying stock. The next numerical experiment is based on 50,000 paths samples with the parameters the same as above except

- the number of observation dates is 90,
- and the volatility is varying from 0.05 to 0.5.

The ratio of standard errors of the Gamma estimated by HOPP(2), divided by the standard errors of the Gamma estimated by the likelihood ratio method is plotted in Figure 2. The ratio of standard error is 22.063 when the volatility is 0.05 and 2.823 when the volatility is 0.5.

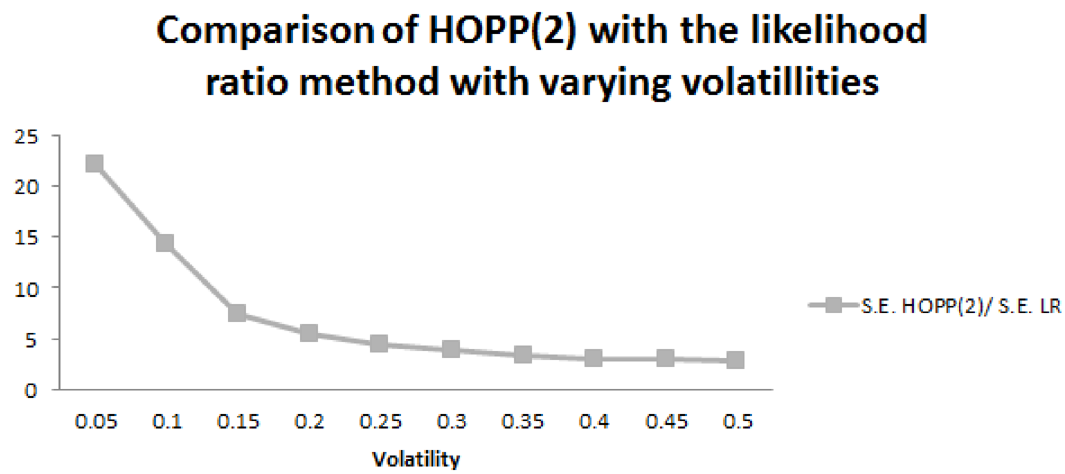

FIGURE 2. Ratio of the standard error computed by the likelihood ratio method divided by the standard error computed by HOPP(2) for changing volatilities with 50,000 paths samples.

2.3. **The multi-step and multi-dimensional case.**

2.3.1. *The generalization.* In our general framework, we allow multiple parameters of interest. The parameter θ in equation (2.1) is a vector of size m, $\theta \in \mathbb{R}^m$ lies within a small open set U_θ about the base point θ_0. We have a set of N time steps, $T_1, T_2, T_3, \ldots, T_N$, and multiple state variables $S_j \in \mathbb{R}^n$ at time T_j depending on the parameter θ and F standard uniforms driving the evolution on each step $V_j \in (0,1)^F$. We divide the evolution across each step into two phases. In the first phase, we perform the ordinary evolution of the state variables except the first random uniform. In the second phase, we perform the change of variable in a similar way to the one dimensional case to remove pathwise discontinuities of both the discounted payoff function and the first order derivative of it with respect to θ.

The resulting HOPP(2) discounted payoff $\hat{P}$ is

$$\hat{P}(\theta) = P\big(K_N(\theta)\big) \prod_{i=1}^{N} \frac{\partial U_i\big(K_j^*(\theta), V_{i,1}\big)}{\partial v}, \tag{2.3.1}$$

where P is the discounted payoff function, depending on the evolution of state-variables up to time T_N, and K_j is the composition of a sequence of mapping Q_i,

$$K_j(\theta) = Q_j \circ Q_{j-1} \circ Q_{j-2} \cdots \circ Q_1(\theta).$$

In this section, we introduce a general framework of constructing $\hat{P}$ under HOPP(2), which is applicable to most cases.

Financial products often have payoffs which are constant or linear away from their discontinuities, we want to make use of this regularity property in our HOPP(2) scheme. We first establish a class of functions which HOPP(2) is applicable to. The functions are defined in terms of strata sets.

Definition 1.
We shall say that disjoint sets $E_{j,k}$ with $j = 1, 2, \ldots, N$, and $k = 0, 1, \ldots, M_j$, stratify the state-space if $E_{j,k} \subset \mathbb{R}^n$ and

$$\bigcup_k E_{j,k} = \mathbb{R}^n.$$

Here M_j is the number of discontinuities at step j, the state variable $S_j \in E_{j,k}$ for some $k = 0$, $1, \ldots, M_j$. Our objective here is to apply the HOPP(2) scheme to a function which has regular properties away from the discontinuities, i.e. it is $\hat{C}^2$ if $S_j \in int(E_{j,k})$ and pathwise discontinuities of it may only occur if $S_j \in \partial(E_{j,k})$. Now, we define the class of discontinuous functions with these strata sets.

Definition 2.
Let $\mathcal{E}((E_{j,k}))$ denote the class of functions

$$\sum_a \alpha_a \beta_a$$

with $\beta_a(K_N)$ a $\hat{C}^2$ function and α_a a product of indicator functions of sets $E_{j,k}$.

This class includes a function which is equal to a general twice differentiable function if a sequence of strata E_{1,k_1} ,E_{2,k_2} , ..., E_{N,k_N} are all visited by the state variables, and zero otherwise. We also include all finite linear combinations of such functions. Now, we proceed to define the evolution under HOPP(2) across the steps.

Definition 3.
Let Q_j be the evolution of state variables over the time step $[T_{j-1}, T_j]$, for $j = 1, 2, \ldots, N$. The evolution of each step is decomposed again into two phases such that

$$Q_j = G_j \circ F_j,$$

with

$$F_j : \mathbb{R}^n_{S_{j-1}} \times U_\theta \times (0, 1)^{F-1} \to \mathbb{R}^n_{S_j^*},$$

and

$$G_j : \mathbb{R}^n_{S_j^*} \times U_\theta \times (0, 1) \to \mathbb{R}^n_{S_j}.$$

Since the composition of differentiable functions is differentiable, we shall say Q_j is C^2 as a function of S_{j-1} and θ if

- *F_j is C^2 with respect to both S_{j-1} and θ for fixed V_i^{F-1},*
- *G_j is C^2 as a function of (S_j^*, θ) for fixed $v_{i,1}$.*

Here, we assume we observe the standard uniforms $v_{j,2}$, $v_{j,3}$,..., $v_{j,F}$ in F_j before $v_{j,1}$, so the occurrence of event $E_{j,k}$ is then determined by $v_{j,1}$. Therefore, although the problem here is multi-dimensional, the change of variable is still one-dimensional, i.e, we only perform the change of variable on the first standard uniform of each step in G_j.

Following Chan and Joshi(2012), we assume that the evolution is adopted to the strata, that is there exists a sequence of twice differentiable functions $a_{j,i}\big(K_j^*(\theta)\big)$ with $j = 1, 2, \ldots, N$, and $i = 0, 1, \ldots, M_j + 1$ such that

$$a_{j,0} = 0 \text{ and } a_{j,M_j+1} = 1,$$

and

$$a_{j,i-1} < a_{j,i},$$

for i=1, 2,..., M_j+1, with the property that $G_j(S_j^*,\theta,v)\in E_{j,k}$ if and only if

$$a_{j,i-1}\Big(K_j^*(\theta)\Big) < v_{j,1} \le a_{j,i}\Big(K_j^*(\theta)\Big).$$

Unlike equation (2.4) where one parameter of interest needs to be considered, our change of variables in the general framework need to remove the limit dependence on m parameters of interest up to the third order. Here, we introduce a change of variable function $U_j\big(K_j^*(\theta),v_{j,1}\big)$ to replace $v_{j,1}$ at step j. It depends on the evolution of the state variable up to time T_{j-1} and the evolution F_j, i.e. $K_j^*=F_j\circ K_{j-1}$, with conditions that

- it is twice differentiable as a function of θ and piecewise differentiable as a function of $v_{j,1}$;
- it is bijective on $[0,1]$ for fixed $K_j^*(\theta)$;
- $U_j\big(K_j^*(\theta_0),v_{j,1}\big)=v_{j,1}$;
- $U_j\big(K_j^*(\theta_0),a_{j,i}(K_j^*(\theta_0))\big)=a_{j,i}(K_j^*(\theta_0))$ for i=0, 1, ..., M_j+1;
- for it to remove the limit dependence of $\hat{P}$ on all parameters, we require for k=1, 2, ..., m,
$$\frac{\partial}{\partial\theta_k}\Big[U_j\big(K_j^*(\theta),a_{j,i}\big)\Big]_{|\theta=\theta_0}=\frac{\partial}{\partial\theta_k}\Big[a_{j,i}\big(K_j^*(\theta)\big)\Big]_{|\theta=\theta_0},$$
for all $i=0,1,\dots,M_j+1$;
- for it to remove the limit dependence of $\hat{P}'$ on all parameters, we require for k=1, 2, ..., m and $s=1,2,\dots,m$,
$$\frac{\partial^2}{\partial\theta_k\partial\theta_s}\Big[U_j\big(K_j^*(\theta),a_{j,i}\big)\Big]_{|\theta=\theta_0}=\frac{\partial^2}{\partial\theta_k\partial\theta_s}\Big[a_{j,i}\big(K_j^*(\theta)\big)\Big]_{|\theta=\theta_0}$$
for all $i=0,1,\dots,M_j+1$.

There are many functions which satisfy the conditions listed above. One example is the unique continuous function which is linear on the intervals $(a_{j,i}(\theta_0),a_{j,i+1}(\theta_0)]$ with the above conditions satisfied,

$$U_j\big(K_j^*(\theta),v_{j,1}\big)=a_{j,i}(K_j^*(\theta))+\frac{a_{j,i+1}(K_j^*(\theta))-a_{j,i}(K_j^*(\theta))}{a_{j,i+1}(K_j^*(\theta_0))-a_{j,i}(K_j^*(\theta_0))}(v_{j,1}-a_{j,i}(K_j^*(\theta_0))), \tag{2.3.2}$$

if $a_{j,i}(\theta_0)<v_{j,1}\le a_{j,i+1}(\theta_0)$. By construction, the first and second-order partial derivatives of U_j with respect to θ satisfy the last two conditions above. As a result of such changes of variable, the strata do not move up to third order, i.e.

$$a_{j,i-1}\Big(K_j^*(\theta_0)\Big)<v_{j,0}<a_{j,i}\Big(K_j^*(\theta_0)\Big)\Leftrightarrow$$

$$a_{j,i-1}\Big(K_j^*(\theta)\Big)+\mathcal{O}\Big(||\theta-\theta_0||^3\Big)<U_j\big(K_j^*(\theta),v_{j,1}\big)<a_{j,i}\Big(K_j^*(\theta)\Big)+\mathcal{O}\Big(||\theta-\theta_0||^3\Big).$$

Another objective of HOPP(2) is to minimize the variance of the Monte-Carlo weight, i.e. to minimize the expression

$$\int_0^1\Big(\frac{\partial U_j\big(K_j^*(\theta),v\big)}{\partial v}\Big)^2dv. \tag{2.3.3}$$

The Monte-Carlo weight from the linear change of variable function in equation (2.3.2) is

$$\frac{a_{j,i+1}(K_j^*(\theta))-a_{j,i}(K_j^*(\theta))}{a_{j,i+1}(K_j^*(\theta_0))-a_{j,i}(K_j^*(\theta_0))},$$

if $a_{j,i}(\theta_0)<v_{j,1}\le a_{j,i+1}(\theta_0)$. Chan and Joshi (2012) showed that, this is optimal in terms of minimizing the expression (2.3.3).

The change of variable using equation (2.3.2) performed on each step has removed the limit dependence on θ of both $\hat{P}$ and $\hat{P}'$, so the small bump size does cause the bumped path and unbumped path to finish on different sides of discontinuities. Define $\theta_k(h)$ to be a vector of size m, which equals to zero except the kth coordinate, and the kth coordinate of it equals to h, for $h > 0$. The finite differencing HOPP(2) Gamma estimator of θ_k for $k = 1, 2, \ldots, m$, is

$$\frac{\hat{P}(\theta + \theta_k(h)) - 2\hat{P}(\theta) + \hat{P}(\theta - \theta_k(h))}{h^2}. \tag{2.3.4}$$

The finite differencing HOPP(2) Cross-Gamma estimator of θ_k and θ_s , for k=1, 2, …, m, and s=1, 2, …, $k-1$, $k+1$, …, m, is

$$\frac{\hat{P}(\theta + \theta_k(h)) - \hat{P}(\theta - \theta_k(h)) - \hat{P}(\theta - \theta_s(h)) + \hat{P}(\theta - \theta_s(h))}{4h^2}. \tag{2.3.5}$$

We are interested in the case as $h \to 0$, i.e. the pathwise estimator of the Hessian. As Glasserman(2004) pointed out, the main obstacle to the applicability of the pathwise method to calculate the Hessian is discontinuities in the first-order derivatives of the discounted payoff function, so we have made the first order derivatives of the discounted payoff function under HOPP(2) Lipschitz continuous. Our HOPP(2) scheme also satisfies other conditions of the pathwise method.

(1) The discounted payoff under HOPP(2) is twice differentiable almost surely since
- the Monte-Carlo weight $\prod_{i=1}^{N} \frac{\partial U_i\left(K_j^*(\theta), V_{i,1}\right)}{\partial v}$ and the evolution of state variables $K_N(\theta)$ are twice differentiable functions of θ provided that $a_{j,i}(\theta)$'s are twice differentiable functions of θ;
- the probability of the event that $\hat{P}$ is twice differentiable as a function of θ at θ_0 is 1.

(2) The finite differencing estimators of Gammas and Cross-Gammas as shown in equation (2.3.4) and (2.3.5) are uniformly integrable.

The discounted payoff function $\hat{P}$ is now $\hat{C}^2$, we can apply the pathwise method to $\hat{P}$ to calculate pathwise estimators of the Hessian. The HOPP(2) estimator of Cross-Gamma is given by

$$\mathbf{E}\left[\frac{\partial^2\left(P\left(K_N(\theta)\right)\prod_{i=1}^{N}\frac{\partial U_i\left(K_j(\theta), V_{i,0}\right)}{\partial v}\right)}{\partial\theta_k\partial\theta_s}\right]. \tag{2.3.6}$$

Here, the change of variable functions are as shown in equation (2.3.2).

2.3.2. *Automatic differentiation for higher dimensional applications.* Automatic differentiation differentiates functions expressed as computer programs. It is based on the fact that all computer programs can be decomposed into strings of simple operations; the derivatives with respect to the input variables are then computed by a chain-rule-based technique. It is only applicable when the program is continuous as a function of the parameters of interest.

To compute the Hessian of multi-dimensional path-dependent products simulated by the HOPP(2) algorithm, we adopt the methodology introduced by Joshi and Yang (2010) for computing Hessians of $\hat{C}^2$ functions to reduce the computational complexity. The speed of their algorithm depends on the number of state variables, M, which is typically greater than m, the number of parameters of interest. Their key result is that, given the Hessian H_g of a function, $g : \mathbb{R}^M \to \mathbb{R}$, if f is an elementary operation, i.e. a vector mapping which is the identity mapping in all coordinates except one and in that coordinate depends on only one or two coordinates, then the Hessian of $g \circ f$, $H_{g\circ f}$ will agree with H_g except in one or two rows and columns. The rows and columns of the Hessian contain M elements, so $H_{g\circ f}$ can be computed by overwriting H_g with $AM + B$ additional operations for some constants A and B. Therefore if we initialize the Hessian at the final payoff date and decompose the evolution of state variables into L operations, then we can compute the Hessian estimator of the path in $(AM + B)L$ additional operations. To resolve the problem of pathwise discontinuities of financial products, they used smooth functions to approximate the discontinuities so the resulting discounted payoff functions are continuously twice-differentiable. However, the Hessian estimator resulted by such approach is biased. Since the discounted payoff function under HOPP(2) is $\hat{C}^2$, the Hessian can be computed by applying the algorithmic Hessian method backwards to produce an unbiased estimator of the Hessian.

3. Inexplicit Critical Value Functions

Our simulation scheme is defined in terms of the discontinuities of the discounted payoff function. The change of variable function at each step is determined by the critical values of the first standard uniform for the discounted payoff function to pass the points of discontinuities, i.e. $a_{j,i}$'s in Section 2.3. So far, we assumed that these critical value functions $a_{j,i}$'s can be found explicitly. For cases where the closed-form solutions of $a_{j,i}$'s do not exist, numerical techniques such as the Newton-Raphson method need to be used to provide an approximation of those critical values.

For the first-order OPP, Chan and Joshi (2012) used the Newton-Raphson method with one step to approximate $a_{j,i}$. The bias resulted from using unbumped strata for the bumped path is of order $\mathcal{O}(||\theta-\theta_0||^2)$. The bias disappears at the limit after division by $||\theta-\theta_0||$. However, this is not sufficient for calculating second-order derivatives using HOPP(2), since the bias does not disappear after division by $||\theta-\theta_0||^2$.

3.1. **Using Newton Raphson's method with two steps to approximate the critical value functions.** We adopt the Newton-Raphson method with two steps to approximate critical value functions. Assume that there exists a function of state variables β_i at T_i, such that for $k = 0, 1, 2, \dots, M_i$

$$a_{i,k}(\theta) < v_{i,1} \le a_{i,k+1}(\theta) \Leftrightarrow D_{i,k}(\theta) < \beta_i\Big(G_i(S_i^*, \theta, v_{i,1})\Big) \le D_{i,k+1}(\theta).$$

We call β_i the proxy constraint function; $D_{i,k}$ is the discontinuity level corresponding to the kth discontinuity at step i of the discounted payoff function. For example, the proxy constraint function for the digital basket option is the average of the basket and the discontinuity level is the strike; the proxy constraint funtion determines a strike crossing of the discounted payoff function.

We perform the linearization using standard normals, to ensure that the estimated $\hat{a}_{i,j}$ lies in the range $(0,1)$. We also assume the function $\beta_i\Big(G\big(S_i^*, \theta, \Phi(Z_{i,1})\big)\Big)$ has which has a positive derivative as a function of $Z_{i,1}$, i.e. the first random standard normal variable in step i. The approximated critical values $\hat{a}_{i,j}(\theta, v_{i,1})$ of $a_{j,i}$ are calculated with the following procedures.

(1) *We approximate the proxy constraint function β_i about the base point $Z_{i,1} = \Phi^{-1}(v_{i,1})$, the simulated first random standard normal,*

$$\hat{\beta}_i^1\Big(G\big(S_i^*, \theta, \Phi(Z)\big)\Big) = \beta_i\Big(G\big(S_i^*, \theta, \Phi(Z_{i,1})\big)\Big) + \frac{\partial \beta_i\Big(G\big(S_i^*, \theta, \Phi(Z_{i,1})\big)\Big)}{\partial Z}(Z - Z_{i,1}). \tag{3.1.1}$$

(2) *We equate this linear approximation $\hat{\beta}_i^{(1)}\Big(G\big(S_i^*, \theta, \Phi(Z)\big)\Big)$ to the discontinuity level $D_{i,k}(\theta)$, and obtain the first estimate $\hat{Z}_{i,k}^1$ of the critical standard normal,*

$$\hat{Z}_{i,k}^{(1)}(\theta, v_{i,1}) = \left(D_{i,k}(\theta) - \beta_i\Big(G\big(S_i^*, \theta, \Phi(Z_{i,1})\big)\Big)\right)\left(\frac{\partial \beta_i\Big(G\big(S_i^*, \theta, \Phi(Z_{i,1})\big)\Big)}{\partial Z}\right)^{-1} + Z_{i,1}. \tag{3.1.2}$$

(3) *We apply this method again, to approximate the proxy constraint function β_i about this new value of $\hat{Z}^{(1)}_{i,k}$'s and then equate to the discontinuity level. We obtain the second approximation of the critical value,*

(3.1.3)

$$\hat{Z}^{(2)}_{i,k}(\theta, v_{i,1}) = \left(D_{i,k}(\theta) - \beta_i\Big(G\big(S_i^*, \theta, \Phi(Z^1_{i,k}(\theta, v_{i,1}))\big)\Big)\right)\left(\frac{\partial\beta_i\Big(G\big(S_i^*, \theta, \Phi(Z^1_{i,k}(\theta, v_{i,1}))\big)\Big)}{\partial Z}\right)^{-1} + Z^1_{i,k}(\theta, v_{i,1}).$$

(4) *The approximated strata function $\hat{a}_{i,j}(\theta, v_{i,1})$ of $a_{j,i}$ is given by*

$$\hat{a}_{i,j}(\theta, v_{i,1}) = \Phi[\hat{Z}^{(2)}_{i,k}(\theta, v_{i,1})]. \tag{3.1.4}$$

The functions $\hat{a}_{i,j}(\theta, v_{i,1})$ are twice-differentiable with respect to θ, provided the proxy constraint function β_i is twice-differentiable with respect to the parameter θ.

The change of variable function is

$$U_i(\theta, v_{i,1}) = \left(\frac{\hat{a}_{i,j+1}(\theta, v_{i,1}) - \hat{a}_{i,j}(\theta, v_{i,1})}{\hat{a}_{i,j+1}(\theta_0, v_{i,1}) - \hat{a}_{i,j}(\theta_0, v_{i,1})}\right)\Big(v_{i,1} - \hat{a}_{i,j}(\theta_0, v_{i,1})\Big) + \hat{a}_{i,j}(\theta, v_{i,1}), \tag{3.1.5}$$

when $\hat{a}_{i,j}(\theta_0, v_{i,1}) < v_{i,1} \leq \hat{a}_{i,j+1}(\theta_0, v_{i,1})$. The resulting Monte-Carlo weight is more complicated than the linear case, since the strata functions $\hat{a}_{i,j}(\theta, v_{i,1})$ are functions of both θ and the first random uniform. The Monte-Carlo weight is the derivative of the function $U_i(\theta, v_{i,1})$ with respect to $v_{i,1}$.

3.2. **The smallness of the bias arising from using the Newton-Raphson method.** The HOPP(2) simulation scheme has the following objectives:

(1) When $\theta = \theta_0$, we have the unbumped paths exactly the same as the bumped paths.
(2) The expectation of the discounted payoff from the HOPP(2) simulation scheme is equal to the expectation of the discounted payoff under the original scheme.
(3) The discounted payoff as a function of θ from the HOPP(2) scheme is $\hat{C}^2$, so we can apply the pathwise method to compute the Hessian. This resulting Hessian computed is equal to the Hessian of the discounted payoff under the original scheme.

Since the strata functions $\hat{a}_{i,j}$'s are approximations, we still have a small possibility of having the bumped and unbumped paths to finish on different sides of discontinuities. We remove this remaining discontinuity by using the unbumped strata $E_{j,k}$ for the bumped path. This creates a bias in the bumped path's expectation, we shall see that the bias is of order $\mathcal{O}\Big(||\theta - \theta_0||^4\Big)$.

We focus on the behaviour of the bumped path close to the critical points, this is because

- when $v_{i,1}$ equals the actual critical value $a_{i,j}$, the estimated critical value by Newton-Raphson's method with two steps is correct, i.e, $\hat{a}_{i,j} = a_{i,j}$;
- both $\hat{a}_{i,j}$ and $a_{i,j}$ are smooth functions of θ;
- the change of variable function is a smooth function between the critical points and will be extendible to the closed interval, and it is also the identity mapping in $v_{i,1}$ when $\theta = \theta_0$.

Therefore, the change of variable function will not map the bumped path across a critical point for θ lying in a small neighbourhood of θ_0 unless $v_{i,1}$ is in a small neighborhood of the true critical value.

We only show the one-step case with a single discontinuity, and v is just above or on a critical point since the case of below follows by symmetry. We assume $a_0(\theta_0) = 0$, and the support of the first random uniform is $[-\frac{1}{2}, \frac{1}{2}]$. We need to show that, the probability of the unbumped path is above the discontinuity while the bumped path actually finishes below the discontinuity is $\mathcal{O}(||\theta-\theta_0||^4)$, that is

$$\mathbb{P}[v>0, u(v,\theta)<a_0(\theta)] = \mathcal{O}\Big(||\theta-\theta_0||^4\Big). \tag{3.2.1}$$

We adopt the change of variable as equation (3.1.5)

$$u(v,\theta) = \hat{a}_0(\theta,v) + \frac{\hat{a}_1(\theta,v)-\hat{a}_0(\theta,v)}{\hat{a}_1(\theta_0,v)-\hat{a}_0(\theta_0,v)}\big(v-\hat{a}_0(\theta_0,v)\big).$$

We rewrite as

$$u(v,\theta) = v + [\hat{a}_0(\theta,v)-\hat{a}_0(\theta_0,v)] + \frac{(\hat{a}_1(\theta,v)-\hat{a}_0(\theta,v))-\hat{a}_1(\theta_0,v)-\hat{a}_0(\theta_0,v)}{\hat{a}_1(\theta_0,v)-\hat{a}_0(\theta_0,v)}(v-\hat{a}_0(\theta_0,v)).$$

We first consider the third term

- it is a smooth as a function of both θ and v;
- it vanishes at both $\theta=\theta_0$ and $v=0$.

We can then express this term as $v(\theta-\theta_0)^T\chi(v,\theta)$ by Taylor's Theorem, where χ is a smooth vector-valued function. So, we have

$$u(v,\theta)-a_0(\theta) = v + \big[\hat{a}_0(\theta,v)-\hat{a}_0(\theta_0,v)-a_0(\theta)\big] + v(\theta-\theta_0)^T\chi(v,\theta).$$

Now consider the second term

- it is a smooth function of both v and θ;
- when $\theta=\theta_0$ it vanishes;
- when $v=0$, we have $\hat{a}_0(\theta_0,0)=0$;
- when $v=0$, $\hat{a}_0(\theta,0)-a_0(\theta)$ vanishes to fourth order in $\theta-\theta_0$, since the Newton-Raphson method with a two-step approximation is used.

Applying Taylor's theorem, the expression becomes

$$\begin{aligned}u(v,\theta)-a_0(\theta) &= v + (\theta-\theta_0)^T v\varpi(v,\theta) + v(\theta-\theta_0)^T\chi(v,\theta) + \vartheta(v,\theta)\\ &= v + v(\theta-\theta_0)^T\kappa(v,\theta) + \vartheta(v,\theta),\end{aligned}$$

where $\kappa=\chi+\varpi$, and $\vartheta(v,\theta)=\mathcal{O}\Big((||\theta-\theta_0||)^4\Big)$ is a smooth function of both v and θ. Now κ is smooth, and in a small enough neighbourhood in θ_0, we have $v(\theta-\theta_0)^T\kappa > -v/2$, so on that set, our event is contained in

$$v/2+\vartheta(v,\theta)<0,\ v>0.$$

Since $\vartheta(v,\theta)=\mathcal{O}\Big(||\theta-\theta_0||^4\Big)$, the event is smaller than

$$v<C(||\theta-\theta_0||)^4,\ v>0,$$

for some constant C, which is $\mathcal{O}\Big(||\theta-\theta_0||^4\Big)$.

Under the finite differencing HOPP(2) scheme, the discounted payoff function is $\hat{C}^2$ for both the bumped and unbumped path, we conclude that the bias in the expectation of the bumped paths with linearization of proxy constraint functions is of order $\mathcal{O}\Big(||\theta-\theta_0||^4\Big)$. For our HOPP(2) scheme, i.e. $\theta\to\theta_0$, the bias disappears after division by $||\theta-\theta_0||^2$. Hence, the HOPP(2) estimator of Hessian is unbiased.

However, the measure change under each step determined by the approximated critical values is no longer fully optimal. Therefore before linearization of the proxy constraint function β_i, it is preferable to apply a simple mapping to β_i, so it is closer to linear then the measure change will be closer to optimal. For example, instead of linearizing β_i, it is sometimes better to linearize $\log(\beta_i)$.

3.3. **Application to the multi-dimensional Black-Scholes model. Notations and Model Setup**. Let $S_i(T_j)$ denote the price of stock for $i = 1, 2, ...n$ at time T_j for $j = 1, 2, ...N$. We consider

(1) a digital basket option, with payoff as follows
$$H\left(\sum_{i=1}^{n}\frac{1}{n}S_i(T) > K\right);$$

(2) a parabolic Asian basket option, with payoff as follows
$$\left(\Big(\frac{1}{N}\sum_{j=1}^{N}\sum_{i=1}^{n}\frac{1}{n}S_i(T_j) - K\Big)_+\right)^2,$$

where H is the Heaviside function.

For HOPP(2), we can adopt the reduced-factor log-Euler discretization scheme in the Monte-Carlo simulations. i.e. for an n asset model, we have F Brownian motions driving the evolution of the asset values where $F < n$. The evolution across the step j for each asset is
$$S_i(T_j) = S_i(T_{j-1})\exp\Big(r\triangle T + \sum_{f=1}^{F}\big(a_{i,f}(T_{j-1})Z_{j,f} - 0.5a_{i,f}(T_{j-1})^2\big)\Big).$$

Here, $Z_{j,f}$ is the fth standard normal random variable simulated at step j. We assume a constant volatility structure and compute a pseudo square root, $A = a_{i,j}$ of the covariance matrix using the spectral decomposition method, that gives us the components in order of significance. The first eigenvector is the one with the highest eigenvalue which is the dominant principal component of the covariance matrix (Lord-Pelsser, 2007).

For every step of the HOPP(2) scheme, in the first phase we evolve all the stocks values ordinarily except the first random uniform,
$$S_i^*(T_j) = S_i(T_{j-1})\exp\Big(r\triangle T - 0.5a_{i,1}(T_{j-1})^2 + \sum_{f=2}^{F}\big(a_{i,f}(T_{j-1})Z_{j,f} - 0.5a_{i,f}(T_{j-1})^2\big)\Big),$$

then in the second phase, the change of variable is performed as described in Section 3.1,
$$S_i(T_j) = S_i^*(T_j)\exp\left(a_{i,1}(T_{j-1})\Phi^{-1}\Big(U_j^*\big(K^*(\theta), v_{j,1}\big)\Big)\right),$$

for $i = 1, 2, \ldots, n$.

Since the likelihood ratio and the pathwise likelihood ratio methods are not applicable when the density function is singular, i.e. when $F < n$, the numerical experiments are conducted using a full-factor model to compare the performance of HOPP(2) with the other methods.

To show the convergence of the Hessian computed by HOPP(2), we apply it to calculate the Hessian of the parabolic Asian basket option. We present the hypothetical parabolic Asian basket option here, so we can compare the Hessians calculated by HOPP(2) and the pathwise method. We use a five-asset model with $\rho_{i,j} = 0.6$, $\sigma_i = 0.5$ for $i, j = 1, 2, \ldots, n$, and $r = 0.05$ per annum. The initial values of the stocks are 100 and the barrier is 100. The expiry time T is 0.5, with 15 observation dates equally space across $[0, T]$, the sample size is 20,000 paths for both HOPP(2) and the pathwise method. We present the results for only the first Gamma and Cross-Gamma in Table 4, since the product and the parameter inputs are symmetric for all stocks. The mean values calculated by the two methods converge; the pure pathwise method produces slightly smaller standard errors, but would not be applicable for non-parabolic payoffs.

We compare the HOPP(2) method with the pure likelihood ratio method, by comparing the mean and the standard errors of the Hessian calculated by the two methods on a digital basket option. For numerical experiment, we implement a five-asset basket with initial stocks equal to 88, the strike is 100. We set the volatility to be 0.1 and the correlation is 0.6. The interest rate is 0.05 and the maturity time is 0.7. The sample size is 20,000 paths for HOPP(2) and 6,000,000 paths for the likelihood ratio method, since the values calculated are of similar degree of precision in terms of the standard error, see Table 4.

	HOPP(2) Mean	HOPP(2) S.E.	PW Mean	PW.S.E
Parabolic Asian first Gamma	0.7896	0.0065	0.7886	0.0063
Parabolic Asian first Cross-Gamma	0.7744	0.0062	0.7733	0.0060
	HOPP(2) Mean	HOPP(2) S.E.	LR Mean	LR.S.E
Digital basket first Gamma	2.06×10^{-4}	4.916×10^{-6}	2.161×10^{-4}	6.534×10^{-6}
Digital basket first Cross-Gamma	2.055×10^{-4}	4.890×10^{-6}	1.98×10^{-4}	4.737×10^{-6}

Table 4. The Comparison of HOPP(2) with the pathwise method on a parabolic Asian basket option both with a 20,000 paths sample, and the comparison of HOPP(2) with a 20,000 paths sample with the likelihood ratio method with a 6,000,000 paths sample on a digital basket option

4. Application to Derivative Products with Angular Payoffs

For financial products with angular payoffs, both HOPP(1) and HOPP(2) are applicable to compute the second order derivatives. For illustration, first we apply the two methods to calculate the Gamma of a vanilla put option, and compare the standard errors and the computational times. HOPP(1) is then used to calculate the Hessian of a portfolio insurance product, the North Guarantee.

4.1. **Simple example: vanilla put option.** For HOPP(2), the change of variable function is

$$u^{(2)}(\theta, v) = v + (\theta - \theta_0)\gamma(v) + \frac{1}{2}(\theta - \theta_0)^2\delta(v).$$

The HOPP(2) estimator for the Gamma of a vanilla put option is

$$\exp(-rT)\Bigg(- 2\frac{S_T\sigma\sqrt{T}}{\phi(Z)S_0}\gamma(v) - \frac{S_T\sigma\sqrt{T}}{\phi(Z)^2}[\sigma\sqrt{T} - Z] - \frac{S_T\sigma\sqrt{T}}{\phi(Z)}\delta(v) - 2\Big[\frac{S_T}{S_0} + \frac{S_T\sigma\sqrt{T}}{\phi(Z)}\gamma(v)\Big]\Bigg) I_{S_T<K}$$
$$+ \exp(-rT)(K - S_T)_+\delta'(v),$$

where v is the simulated standard uniform, Z is the simulated standard normal random variable and ϕ is the standard normal density function. The functions γ and δ are defined as Section 3.1.

For HOPP(1), the change of variable function is

$$u^{(1)}(\theta, v) = v + (\theta - \theta_0)\gamma(v).$$

The HOPP(1) estimator for the Gamma has a simpler form

$$\exp(-rT)\Bigg(- 2\frac{S_T\sigma\sqrt{T}}{\phi(Z)S_0}\gamma(v) - \frac{S_T\sigma\sqrt{T}}{\phi(Z)^2}[\sigma\sqrt{T} - Z] - 2\Big[\frac{S_T}{S_0} + \frac{S_T\sigma\sqrt{T}}{\phi(Z)}\gamma(v)\Big]\Bigg) I_{S_T<K}.$$

The function γ is the same as the one in $u^{(2)}(\theta, v)$.

The two estimators are both unbiased, and converge to the analytic value. Since the HOPP(1) estimator of the Gamma does not involve the term with the payoff and the function δ, we can expect a smaller standard error and a shorter computational time. We perform a Monte-Carlo simulation to calculate the Gamma of an at-the-money put option to compare the two methods. We set $S_0 = K = 100$, $T = 0.2$, $\sigma = 0.2$ and $r = 0.05$ per annum. For a sample of 4,000,000 paths, the standard error is 2.45×10^{-5} calculated by HOPP(2) and 2.34×10^{-5} by HOPP(1).

The computation time is 0.239 seconds by HOPP(2) and 0.2 seconds by HOPP(1). Therefore, it appears that for financial products with angular payoffs, it is slightly more efficient to use HOPP(1) to calculate the Hessian.

4.2. **Multi-dimensional example: the North Guarantee.** The North Guarantee is an insurance product, which is dependent on a basket of assets. It provides investor a capital guarantee over their investments while still benefiting from any market growth in their selected investments. There are many portfolio insurance product available in the market with similar features, here we refer to the one provided by the insurer, AMP (AMP, 2009). An investment strategy(from two or three options) and the term of the contract (from 6, 8, 10 or 20 years) are chosen at the commencement, the investor cannot change the term or the investment strategy during the term of the contract, but can switch investment options within the chosen investment strategy. However, we do not model this option in our simulation exercises.

The payoff to the insured is defined in terms of the account value and the protected balance of the insured's account.

- The account value is the current market value of the underlying assets, after insurance premiums deducted as a percentage of the value monthly. It is subject to ordinary market fluctuations.
- The protected balance locks in any growth as a result of positive investment performance each year or every two years, provided that the account value exceeds the protected balance. For example, the initial market value of the basket is worth one million, at the first observation date after the commencement, the account value is the current value of the basket net of the insurance premiums
  - if the account value is more than one million, the protected value then locked in as the current account value,
  - else, the protected value is one million.

  The protected balance is not affected by negative market performances. At maturity, it is the maximum account value observed on the annual or biennial observation dates before the maturity.

Without the guarantee, only the account value at maturity is paid to the insured. With the guarantee, the terminal amount paid to the insured is the maximum of the protected balance and the account value observed at the maturity. Therefore the payoff to the insured is the difference between the two, i.e.

$$(\text{Protected Balance} - \text{Account Value})_+.$$

The discounted payoff function of this product as shown above is angular, so we apply HOPP(1) to calculate the Hessian of this product with respect to the initial asset values.

For numerical experiments, we implement a five-asset basket, with $19 + t_0$ years to the maturity (this is after its commencement). The observation dates are at t_0, $1+t_0$, $2+t_0$, $\ldots$, T. The account values are charged with insurance premiums 0.3% monthly in arrears, i.e. the account value at t_0 is $0.997^{12}\sum_{i=1}^{5} S_i(t_0)$. The premium is charged as a percentage of the account value, so we only evolve the assets over the observation dates. Since the investors are given the right to choose different investment options with different degrees of riskiness within each investment strategy, the volatilities are set to be 0.02, 0.06, 0.2, 0.3 and 0.4 per annum, the correlations are 0.6. The interest rate is 0.05 per annum.

We plot the sum of standard errors of the Hessian calculated by the pathwise likelihood ratio method and HOPP(1) with varying first observation date t_0 from 0.01 to 0.1 in Figure 3, the sample size is 20,000 paths. Here, we set the current stocks to be 80 and the original stocks to be 100. We observe the pathwise likelihood ratio method produces very large standard errors when the first observation date is small. The HOPP(1) method unlike the pathwise likelihood ratio method produces very stable standard errors as the first observation date changes. For the same set of t_0, the maximum sum of standard errors calculated by HOPP(1) is 0.01505 when $t_0 = 0.1$ compared to 0.44294 by the pathwise likelihood ratio method; the minimum is 0.01470 when $t_0 = 0.015$, compared to 1.13244 by the pathwise likelihood ratio method.

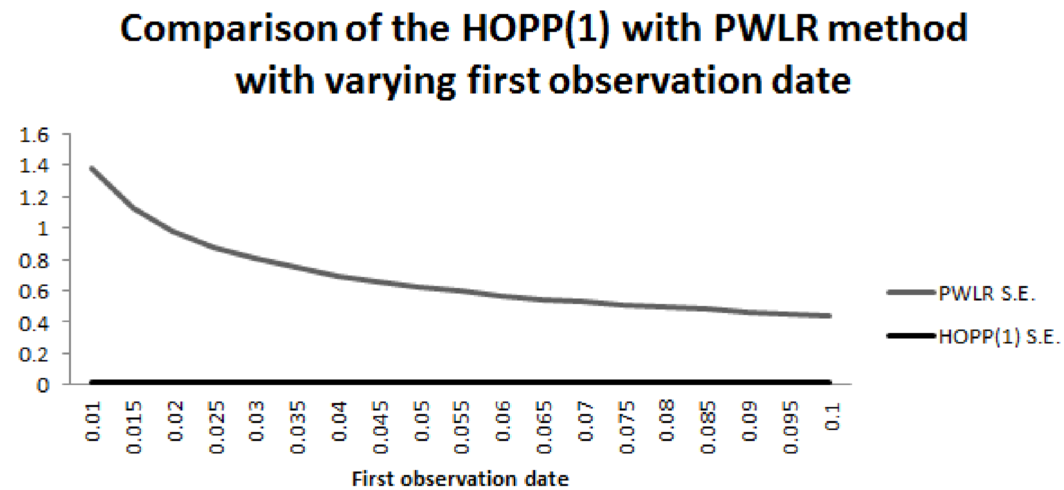

Figure 3. North Guarantee: The sum of standard errors of the Hessian calculated by the PWLR method and HOPP(1) with varying initial observation date, with 20,000 paths sample

We plot the sum of standard errors of the Hessian calculated by the pathwise likelihood ratio method and HOPP(1) with varying current stock price level from 100 to 50 in Figure 4, the sample size is 20,000 paths. Since when the current stock level is above the initial stock prices, the Gamma estimates are very small, we only consider the cases where the current stock is below the original stock level. Here the original stock prices are 100 and the first observation time is 0.1. The maximum sum of standard errors by HOPP(1) is 0.117088 when the stock level is 90, compared to 0.421071 by the pathwise likelihood ratio method; the minimum by HOPP(1) is 0.01463 when the stock level is 78, compared to 0.45681 by the pathwise likelihood ratio method.

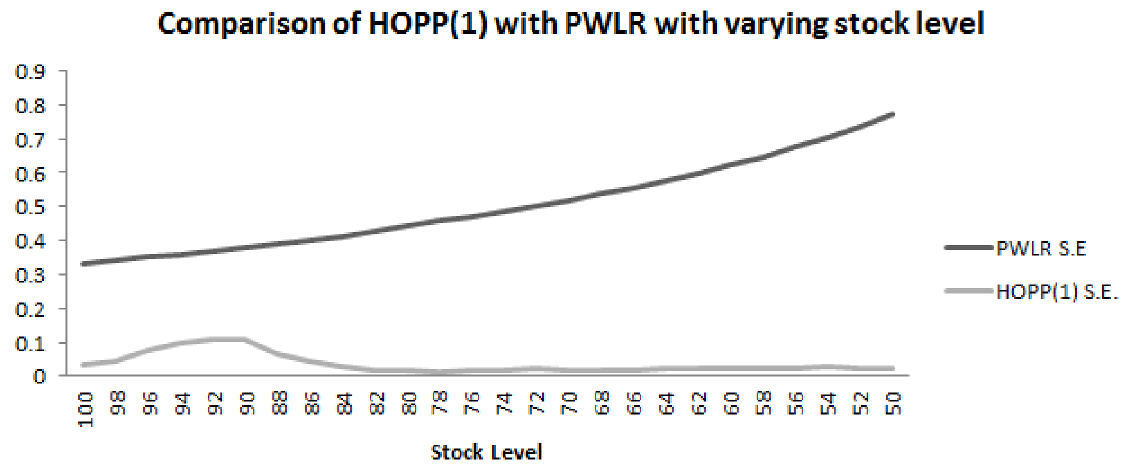

FIGURE 4. North Guarantee: The sum of standard errors of the Hessian calculated by the PWLR method and HOPP(1) with varying current stock values, with 20,000 paths sample

When the first observation date is small and the current stock level is far below the original stock level, the pathwise likelihood ratio method produces very large standard errors. We set the current stock price to be 55 and the first observation date to be 0.01. Since the pathwise likelihood ratio method produces also very large standard errors of Cross-Gammas when the volatilities are different, for this particular example, we use volatility 0.1 for all assets in order to compare the mean of the Hessian calculated by the two methods. We compare the results calculated by HOPP(1) on a sample with 20,000 paths and the results calculated by the pathwise likelihood ratio method on a sample with 14,000,000 paths, so the Hessian calculated by the two methods are similar precision, see Table 5

	HOPP(1) Mean	HOPP(1) S.E.	PWLR Mean	PWLR.S.E
North Guarantee first Gamma	2.0257×10^{-3}	2.594×10^{-4}	2.418×10^{-3}	2.972×10^{-4}
North Guarantee first Cross-Gamma	1.836×10^{-3}	2.412×10^{-4}	2.0187×10^{-3}	2.972×10^{-4}

TABLE 5. North Guarantee: The comparison of the Hessian calculated by HOPP(1) with a 20,000 paths sample and by the pathwise likelihood ratio method with a 14,000,000 paths sample

When the first observation date is long and the current stock level is close to the original stock level, the pathwise likelihood ratio method performs its best result. For the next numerical example, we set the first observation date to be 1 and the current stock equals to the original stock level at 100, see Table 8, 9 and 10 in Appendix One.

Unlike HOPP(1), the pathwise likelihood ratio method generates very large standard errors when

- the volatility is very small,
- the differences between the volatilities are large,
- the first observation time is small.

5. Application to the LIBOR Market Model

5.1. **Model description: the LIBOR market model.** In this section, we briefly summarize the LIBOR market model (Brace et al,1997). Given, the set of tenor dates

$$0 < T_0 < T_1 < T_2 \cdots < T_{n-1} < T_n,$$

where the time difference between two reset dates T_{i+1} and T_i is τ_i. Let $B_i(t)$ denote the value of zero-coupon bonds maturing at T_i, observed at time t. We define the LIBOR spanning over the period $[T_i, T_{i+1})$ at time $t < T_i$ as

$$f_i(t) = \frac{B_i(t) - B_{i+1}(t)}{\tau_i B_{i+1}(t)},$$

with dynamic

$$df_i(t) = u_i(f,t) f_i(t) dt + \sigma_i(t) f_i(t) dW_t^i, \tag{5.1}$$

where W_t is the F-dimensional standard Brownian motions. For $t > T_i$, we have $f_i(t) = f_i(T_i)$. Typically, the instantaneous volatility curve $\sigma_i(t)$ is chosen to be time-homogeneous and the correlations between the rates are assumed to be constant.

Define a function $\eta : [0, T_n) \to 0, 1, 2, ..., n-1$, to be the index of the next LIBOR reset date at time t. The LMM has a multi-dimensional setting, Monte-Carlo simulations are used to price and hedge exotic interest-rate derivatives. The modelling dynamics of f_i under the spot measure with the log-Euler scheme is

$$f_i(T_{j+1}) = f_i(T_j) e^{u_i(T_j) + \sum_{k=1}^{F} (a_{ik} Z_k - 0.5 a_{ik}^2)} \tag{5.2}$$

for $T_{j+1} \leq T_i$. Here a_{ik} is the ikth element of the pseudo square root of the covariance matrix calculated by the spectral decomposition and the drift is computed by the method introduced by Joshi(2003). Under the spot measure, the drift is

$$u_i(T_j) = \sum_{k=1}^{F} a_{ik} e_{ik}, \tag{5.3}$$

with

$$e_{ik} = \sum_{s=\eta(T_j)}^{i} \frac{\tau_s f_s}{1 + \tau_s f_s} a_{sk}. \tag{5.4}$$

The computational order of implementing the LMM is $O(nF)$ per step by this method. The drift in the model is state-dependent.

The LMM is widely used by practitioners and the academic community for valuing interest rate derivatives, especially exotic derivatives such as target redemption notes, swaps and Bermudan Swaptions, among many others. The key advantage of the model is that it models a set of forward rates which is directly observable in the market, rather than the short rate or the instantaneous forward rates. The model is also consistent with the market standard approach for pricing caps using Black's formula, which helps calibrating the model to current market prices of caplets and European swaptions. The LMM has developed from its standard lognormal basis to more sophisticated versions to produce more realistic market features, for example, stochastic volatilities are incorporated in the model to capture the market-observed volatility smile. The model has been studied extensively in the past; for more details of the model, we refer the reader to Brace (2007), Fries(2007) and Joshi(2011).

The computation of sensitivities under the LMM is especially complicated. The model has a multi-dimensional framework and the drift of each rate is also state-dependent. The payoffs of most interest rate products are either discontinuous, or angular. As a result, the pathwise method is not applicable to compute the second order derivatives of these products. The likelihood ratio method, on the other hand, is not applicable due to the use of reduced factor model, the joint density function at each step is singular. Even if the full factor model is used, the likelihood ratio method generally produces very large standard errors.

For interest-rate derivative products with angular payoffs, adjoint methods of algorithmic differentiation can be used to calculate the gradients. The general result is that the gradient can be computed with 4 times as many operations as the price (Griewank and Walther, 2008). Denson and Joshi(2009) extended the pathwise adjoint method for Deltas and Vegas to the displaced diffusion LMM, they also presented a way to improve the speed of the computation by working with the log-rates. However, these methods are not applicable to calculate the second-order derivatives if the payoff function is angular or discontinuous. Joshi and Yang (2010) applied the algorithmic Hessian approach to finding Gammas of interest rate derivatives using the LMM whilst adopting a smoothing technique to cope with the lack of twice differentiability of payoffs. HOPP(2) does apply to these cases and produces an unbiased estimator of the second-order derivatives, nonetheless has similar computational complexity to the algorithmic Hessian approach. In addition, it is also applicable when the probability density function is singular. Therefore HOPP(2) can be widely used to calculate second-order derivatives of interest rate products.

5.2. **LIBOR market model product specifications. Digital Caplet**: A digital caplet pays off 1 at T_1 when the value of the underlying forward rate f_0 is above the strike K at T_0, or zero otherwise. The payoff as observed at T_0 is

$$\frac{I_{f_0(T_0)>K}}{1+f_0(T_0)\tau_0}.$$

Caplet: A caplet pays off $(f_0(T_0)-K)\tau_0$ at T_1 when the value of the underlying forward rate f_0 is above the strike K at T_0, or zero otherwise. The payoff as observed at T_0 is

$$\frac{(f_0(T_0)-K)_+\tau_0}{1+f_0(T_0)\tau_0}.$$

For the Monte-Carlo simulation, we set $f_0 = K = 0.04$, $T_0 = 0.5$ and $T_1 = 1$. The volatility of the forward is 0.2. On a sample of 200,000 paths, the results are summarized in Table 6.

Option	Analytical Mean	HOPP(2) Mean	HOPP(2) S.E.	LR S.E	S.E Ratio
Caplet Delta	0.256	0.25589	0.00019	0.00031	1.591
Caplet Gamma	33.892	33.88921	0.022	0.103	4.700
Digital Caplet Delta	68.061	68.08076	0.036	0.050	1.390
Digital Caplet Gamma	-920.293	-919.91974	0.486	14.050	28.92

Table 6. Vanilla interest rate product: means and standard errors of the Delta and Gamma estimates by HOPP(2) and the likelihood ratio method on a 200,000 paths sample

5.2.1. *Target redemption note.* Here, we consider the definition of TARN in Piterbarg(2004), where the note knocks out when the total structured coupon reaches a target level R. The structured coupon of each time T_j is $C_j = (K - 2.0 f_j(T_j))_+$, paid at time T_{j+1}. With initial coupon C^*, it has payoff at time T_0

$$CF_0 = \frac{\Big(C^* + I_{C_0<R}C_0 + I_{C_0>R}R - f_0(T_0)\Big)\tau_0}{1+f_0(T_0)\tau_0},$$

the cashflow generated at time T_j is

$$CF_j = \frac{\Big((I_{(\sum_{k=0}^{j} C_k < R)} C_j + I_{(\sum_{k=0}^{j-1} C_k < R \sum_{k=0}^{j})}(R - \sum_{k=0}^{j-1} C_k) - f_j(T_j) I_{(\sum_{k=0}^{j-1} C_k < R)}\Big)\tau_j}{1 + f_j(T_j)\tau_j},$$

the cashflow at time T_{n-1} is

$$CF_{n-1} = \frac{I_{(\sum_{k=0}^{n-2} C_k < R)}(R - \sum_{k=0}^{n-2} C_k - f_{n-1}(T_{n-1}))\tau_{n-1}}{1 + f_{n-1}(T_{n-1})\tau_{n-1}}.$$

For the Monte-Carlo simulations, we implement a five-rate LIBOR model for the TARN to compare the results from the HOPP(2) and the likelihood ratio method. The likelihood ratio method is only applicable if the full-factor model is used, therefore it will be computationally intensive if a large number of forward rates is involved. The forward times are equally-spaced across $[0.5, 3]$. The initial forward rates increase linearly from $f_0 = 2.5\%$ to $f_4 = 4.5\%$. We set $K = 0.12$ and $R = 0.09$. The volatilities are 0.2 and the correlation ρ_{ij} is $0.5 + 0.5\exp(-0.2|T_i - T_j|)$. We compare the results of the Hessian calculated by HOPP(2) with a 200,000 paths sample, and the Hessian calculated by the likelihood ratio method on a sample of 20,000,000 paths, so the standard errors of the first two Gammas and the first Cross-Gamma by the two methods are of similar precision, see Table 11 and 12 in Appendix One.

We also compare the ratio of the sum of standard errors of the Hessian calculated by the likelihood ratio method, divided by the sum of standard errors of the Hessian calculated by HOPP(2) with the same sample size, 200,000 paths, refer to Table 13.

We analyze the performance of HOPP(2) and the likelihood ratio method, with different volatilities and different values of T_0. We plot the sum of the standard error of the Hessian calculated by the two methods on the same graph, the parameters are taken to be the same as the above example, except we vary the volatility from 0.1 to 0.5 in Figure 5 and T_0 from 0.05 to 0.5 in Figure 6. The sample sizes for these two experiments are 500,000 paths. The results are represented in the Figures 5 and 6.

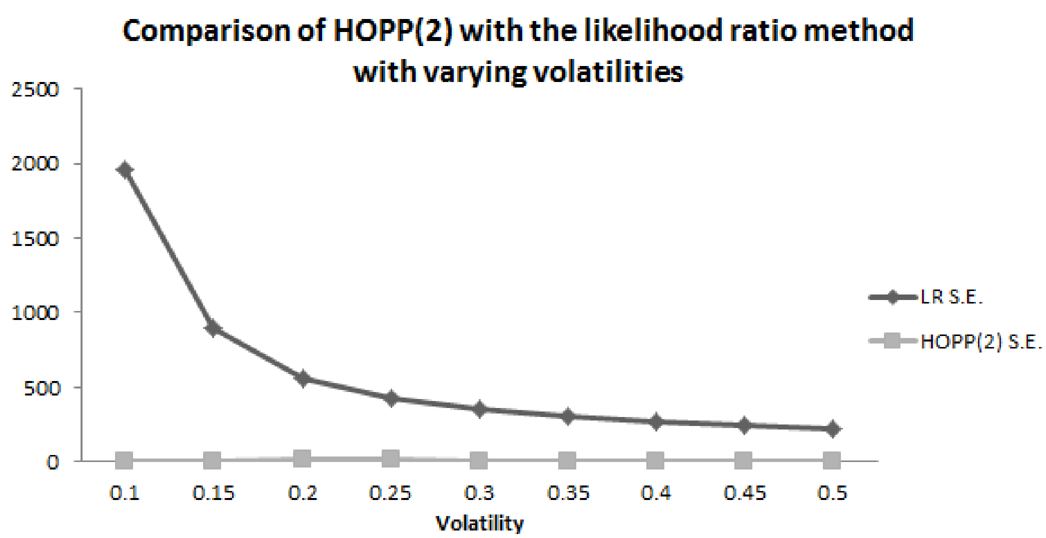

FIGURE 5. Hessian of five-rates TARN with varying volatility: the sum of standard errors of the Hessian calculated by HOPP(2) and the likelihood ratio method with 500,000 paths

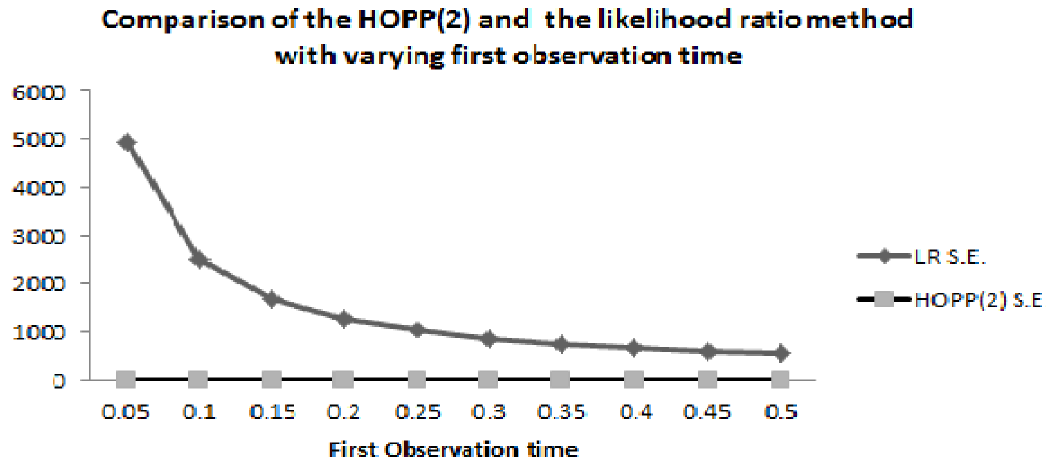

FIGURE 6. Hessian of five-rates TARN with varying first observation date: the sum of standard errors of the Hessian calculated by HOPP(2) and the likelihood ratio method with 500,000 paths

The HOPP(2) method always outperforms the likelihood ratio method with smaller standard errors. The standard errors of the Hessian calculated by HOPP(2) unlike the likelihood ratio method are very stable. The maximum sum of standard errors by HOPP(2) for the set of volatilities is 9.88 when $\sigma = 0.2$, compared to 555.89 by the likelihood ratio method; the minimum is 2.07 when $\sigma = 0.1$, compared to 1962.38 by the likelihood ratio method. The maximum sum of standard errors by HOPP(2) for the set of T_0 is 9.88 when $T_0 = 0.5$, compared to 555.89 by the likelihood ratio method; the minimum is 3.69 when $T_0 = 0.05$, compared to 4929.00 by the likelihood ratio method.

5.2.2. *Computational timing for TARN.* Finally, we investigate the computational timing of HOPP(2) for calculating Greeks. The algorithmic Hessian method by Joshi and Yang(2011) is able to compute all the Gammas with order $O(nFNM)$ with $M = n + F + 2$ if the order of computation for the price is $O(nFN)$ which is the case for the LMM.

For this experiment, we compute the price and the Greeks (the Hessian and the Deltas are computed simultaneously) with sample size of 50,000 paths. Since we are using n steps with n rates, the time taken for computing price should be with an order of n^2, and the Hessian should be n^3. The results are shown in Table 5. We also divide the times by n^2 and n^3 to compare the results.

n	5	10	15	20	25
Price Time	0.218	0.53	0.936	1.451	2.059
Additional HOPP(2) Greek Time	0.749	1.748	2.964	4.508	6.287
Price $\times n^-2$	0.00872	0.0053 7	0.00416	0.0036275	0.0032944
Greeks $\times n^-3$	0.005992	0.001748	0.000878222	0.0005635	0.000402368

TABLE 7. Time(in seconds) of pricing and calculating the Deltas and the Hessian simultaneously of a TARN

It is interesting to see the ratio of the additional HOPP(2) Greek times divided by the time of calculating the price is not increasing as the number of assets increases. This is possibly because we have used the same target rate for the TARNs with different number of forward rates in Table

7, so the TARN with 25 rates has a much higher possibility of early redemption than the one with 5 rates. The HOPP(2) method only initializes the Hessian at the redemption date, the computational order for Hessian is then $O(nFMI)$ for $I \leq N$, where I the number of observation dates up to and including the redemption date. The average value of I is likely to be similar for the five-rate TARN and the 25-rate TARN, since the same target rate is used.

6. Conclusion

We have presented an extension of the OPP method for computing second order Greeks of financial products. We introduced a measure change, which removes the pathwise discontinuity of both the discounted payoff function and its first order derivatives. The measure change is selected so that the variance of the likelihood ratio part is minimized. The simulated discounted payoff function under the new scheme is $\hat{C}^2$, the algorithmic Hessian method is used to calculate the pathwise Hessian of the discounted payoff under the new scheme. This method allows multiple discontinuities per step as well as for cases with non-linear proxy constraint functions. This approach is applicable for a wide range of products with discontinuous or angular payoffs, where the pure pathwise method is not applicable. Our numerical results suggest that the new method is significantly better than the pathwise-likelihood ratio and the pure likelihood ratio method in terms of reducing the standard error of the Hessian estimates, as well as being made more widely applicable.

Appendix One: Tables of Numerical Results

HOPP(1) vs PWLR	S_1	S_2	S_3	S_4	S_5
S_1	0.8069 vs 0.0207	0.8237 vs -0.3237	0.761 vs -0.3672	0.6349 vs -3.8794	0.525 vs 5.4115
S_2	0.8237 vs -0.3237	0.8705 vs 1.3621	0.7693 vs 1.1033	0.642 vs 0.6253	0.5222 vs -2.0246
S_3	0.761 vs -0.3672	0.7693 vs 1.1033	1.0118 vs 1.1948	0.504 vs 0.3869	0.3745 vs 1.7844
S_4	0.6349 vs -3.8794	0.642 vs 0.6253	0.504 vs 0.3869	1.0092 vs 1.4486	0.1535 vs 0.4795
S_5	0.525 vs 5.4115	0.5222 vs -2.0246	0.3745 vs 1.7844	0.1535 vs 0.4795	1.0664 vs 0.2248

Table 8. North-Guarantee: Mean×1000 of the Hessian calculated by the pathwise likelihood ratio method and by HOPP(1) using a 20,000 paths sample

HOPP(1) vs PWLR	S_1	S_2	S_3	S_4	S_5
S_1	0.1037 vs 0.9191	0.104 vs 0.8414	0.1177 vs 1.5667	0.1239 vs 3.0328	0.1009 vs 6.4483
S_2	0.104 vs 0.8414	0.1066 vs 0.283	0.1182 vs 0.522	0.1238vs 1.101	0.1003 vs 2.2495
S_3	0.1177 vs 1.5667	0.1182 vs 0.522	0.1631 vs 0.1596	0.1331vs 0.3033	0.1097 vs 0.6264
S_4	0.1239 vs 3.0328	0.1238 vs 1.101	0.1331 vs 0.3033	0.182vs 0.2171	0.1085 vs 0.5039
S_5	0.1009 vs 6.4483	0.1003 vs 2.2495	0.1097 vs 0.6264	0.1085 vs 0.5039	0.1703 vs 0.3053

Table 9. North-Guarantee: Standard Errors×1000 of the Hessian calculated by the pathwise likelihood ratio method and by HOPP(1) using a 20,000 paths sample

HOPP(1) vs PWLR	S_1	S_2	S_3	S_4	S_5
S_1	8.866	8.091	13.31	24.473	63.877
S_2	8.091	2.655	4.416	8.896	22.427
S_3	13.31	4.416	0.978	2.279	5.708
S_4	24.473	8.896	2.279	1.193	4.643
S_5	63.877	22.427	5.708	4.643	1.793

Table 10. North-Guarantee:Ratio of the standard error of the Hessian calculated by the pathwise likelihood ratio method divided by the standard errors of the Hessian calculated by HOPP(1) using a 20,000 paths sample

	f_1	f_2	f_3	f_4	f_5
f_1	-29.46 vs -29.36	-36.17 vs -34.89	-13.82 vs -16.48	-1.88 vs -2.5	-0.82 vs -0.48
f_2	-36.17 vs -34.89	-43 vs -41.71	-16.81 vs -16.58	-2.13 vs -1.27	-1.02 vs -2.23
f_3	-13.82 vs -16.48	-16.81 vs -16.58	-0.87 vs 1.16	-2.12 vs -2.29	-0.59 vs -0.19
f_4	-1.88 vs -2.5	-2.13 vs -1.27	-2.12 vs -2.29	1.22 vs -1.14	-0.12 vs 0.77
f_5	-0.82 vs -0.48	-1.02 vs -2.23	-0.59 vs -0.19	-0.12 vs0.77	0.01 vs -0.46

TABLE 11. TARN: Mean of the Hessian computed by HOPP(2) on 200,000 paths and by the likelihood ratio method on 200,000,000 paths

	f_1	f_2	f_3	f_4	f_5
f_1	1.88 vs 1.7	1.72 vs 1.64	1.01 vs 1.17	0.26 vs 1.02	0.03 vs 0.67
f_2	1.72 vs 1.64	2.84 vs 2.2	1.29 vs 1.48	0.34 vs1.16	0.04 vs 0.76
f_3	1.01 vs 1.17	1.29 vs 1.48	1.13 vs 1.61	0.27 vs1.11	0.03 vs 0.65
f_4	0.26 vs 1.02	0.34 vs 1.16	0.27 vs 1.11	0.16 vs1.23	0.01 vs 0.69
f_5	0.03 vs 0.67	0.04 vs 0.76	0.03 vs 0.65	0.01 vs 0.69	0.00 vs 0.52

TABLE 12. TARN: standard errors of the Hessian computed by HOPP(2) on 200,000 paths and by the likelihood ratio method on 200,000,000 paths

	f_1	f_2	f_3	f_4	f_5
f_1	28.31	30.06	35.95	121.57	690.56
f_2	30.06	24.53	36.36	106.65	585.63
f_3	35.95	36.36	44	124.86	703.82
f_4	121.57	106.65	124.86	232.28	1857.18
f_5	690.56	585.63	703.82	1857.18	66540.6

TABLE 13. TARN: standard errors of the Hessian computed by the likelihood ratio method over standard errors of the Hessian computed by HOPP(2) on 200,000 paths

References

[1] AMP. (2009). *North Investment Guarantee: Product Disclosure Statement* AMP, Australia, Retrieved 9th April 2014,<https:www.amp.com.au>.

[2] Benhamou E. (2003). *Optimal Malliavin Weighting Function for the Computation of the Greeks*. Mathematical Finance, 13(1), 37-53.

[3] Brace, A. (2002). *Engineering BGM*. Chapman and Hall.

[4] Brace, A., Gatarek, D., and Musiela, M. (1997). *The market model of interest-rate dynamics*. Mathematical Finance 7, 127-155.

[5] Broadie, M., and Glasserman, P. (1996). *Estimating security price derivatives using simulation*. Management science, 42(2), 269-285.

[6] Chan, J. H., and Joshi, M. S. (2012). *Optimal Limit Methods for Computing Sensitivities of Discontinuous Integrals Including Triggerable Derivative Securities*. Available at SSRN 2011690.

[7] Denson, N., and Joshi, M. S. (2009). *Flaming Logs*. Wilmott Journal Volume 1(5-6), 259-262.

[8] Fries, C. (2007). *Mathematical Finance: Theory, Modeling, Implementation*. Wiley.

[9] Giles, M. (2008). *Vibrato Monte Carlo sensitivities*. In Monte Carlo and Quasi-Monte Carlo Methods 2008. Springer, New York.

[10] Giles, M., and Glasserman, P. (2006). *Smoking adjoints: fast Monte Carlo Greeks*. RISK, 19(1), p88-92.

[11] Glasserman, P. (2004). *Monte Carlo Methods in Financial Engineering*. Springer, New York.

[12] Griewank, A., and Walther, B. (2008). *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*. SIAM

[13] Heidergott, B., Pflug, G., and Farenhorst-Yuan, T. (2010). *Gradient estimation for discrete-event systems by measure-valued differentiation*. ACM Transactions on Modeling and Computer Simulation (TOMACS), 20(1), 5.

[14] Hobbs, C., Krishnaraj, B., Liu, Y., and Musselman, J. (2009) *Calculation of variable annuity market sensitivities using a pathwise methodology*. Life & Pensions, September 2009.

[15] Joshi, M. S.(2003). *Rapid computation of drifts in a reduced factor LIBOR market model*. Wilmott Magazine, 5, 84-85.

[16] Joshi, M.S. (2011). *More mathematical finance*. Pilot Whale Press.

[17] Joshi, M. S., and Pitt, D. (2010). *Fast Sensitivity Computations for Monte Carlo Valuation of Pension Funds*. Astin Bulletin, 40(2). 655 - 667.

[18] Joshi, M. S., and Yang, C. (2011). *Algorithmic Hessians and the fast computation of cross-gamma risk*. IIE Transactions, 43(12),878-892.

[19] L'Ecuyer, P. (1990). *A unified view of the IPA, SF, and LR gradient estimation techniques*. Management Science,36(11) 1364-1383.

[20] Liu, G., and Hong, L. J. (2011). *Kernel estimation of the Greeks for options with discontinuous payoffs*. Operations research, 59(1), 96-108.

[21] R. Lord, A. A. J. Pelsser. (2007). *Level-Slope-Curvature - Fact or Artefact?*. Applied Mathematical Finance, 14(2), 105-130

[22] Piterbarg, V. V. (2004). *TARNs: models,valuation, risk sensitivities*. Wilmott Magazine, 14, 62-71.

[23] Wang, Y., Fu, M. C., and Marcus, S. I.(2012). *A new stochastic derivative estimator for discontinuous payoff functions with application to financial derivatives*. Operations Research, 60(2), 447-460.

Centre for Actuarial Studies, Department of Economics, University of Melbourne, Victoria 3010, Australia

E-mail address: `mark@markjoshi.com`

Centre for Actuarial Studies, Department of Economics, University of Melbourne, Victoria 3010, Australia

E-mail address: `d.zhu2@student.unimelb.edu.au`