World Scientific
www.worldscientific.com

MOST-LIKELY-PATH IN ASIAN OPTION PRICING UNDER LOCAL VOLATILITY MODELS

LOUIS-PIERRE ARGUIN

Department of Mathematics, Baruch College, CUNY
1 Bernard Baruch Way, New York, NY 10010, USA
louis-pierre.arguin@baruch.cuny.edu

NIEN-LIN LIU

BKC Research Organization of Social Sciences, Ritsumeikan University
Noji-higashi 1-1-1, Kusatsu, Shiga 525-8577, Japan
nienlin1126@gmail.com

TAI-HO WANG*

Department of Mathematics, Baruch College, CUNY
1 Bernard Baruch Way, New York, NY 10010, USA
tai-ho.wang@baruch.cuny.edu

This paper addresses the problem of approximating the price of options on discrete and continuous arithmetic averages of the underlying, i.e. discretely and continuously monitored Asian options, in local volatility models. A "path-integral"-type expression for option prices is obtained using a Brownian bridge representation for the transition density between consecutive sampling times and a Laplace asymptotic formula. In the limit where the sampling time window approaches zero, the option price is found to be approximated by a constrained variational problem on paths in time-price space. We refer to the optimizing path as the *most-likely path* (MLP). An approximation for the implied normal volatility follows accordingly. The small-time asymptotics and the existence of the MLP are also rigorously recovered using large deviation theory.

Keywords: Asian option pricing; asymptotic expansion; exotic option; large deviation theory; most-likely-path.

1. Introduction

Asian options, also known as average price options, are among the most liquidly traded exotic options in commodities such as agriculture, energy, and fixed incomes

---

*Corresponding author.

markets. Nowadays average price options represent a high percentage of options on oil; some are directly on the futures contracts of oil, while others on spreads between two types of oil futures. Asian options are also commonly used as a risk management vehicle for, owing to its averaging feature, (a) the underlying average price is more difficult to manipulate; (b) the average price is less sensitive to abrupt shocks; and (c) such options are cheaper than similar vanilla options.

There is a rich literature on the problem of pricing Asian options, in part due to the difficulty of finding analytical solutions even for simpler cases like the Black–Scholes model. We mention only the following few and refer the interested readers to the references therein. To our knowledge, Kemna & Vorst (1990) was the first published work tackling the problem of Asian option pricing in Black–Scholes model. As numerically pricing Asian options by Monte Carlo simulations is concerned, Kemna & Vorst (1990) also introduced a variance reduction technique by using the price of geometric average option, whose analytic form is readily obtained, as a control variate. See also Curran (1994) for further analysis and extension to portfolio options on the technique of conditioning. Further development and improvement of Monte Carlo schemes since then were followed up by Broadie & Glasserman (1996) (for a direct method of estimating the Greeks of an Asian option by Monte Carlo simulation, see Sec. 4.2, p. 275), Vazquez-Abad & Dufresne (1998) (combining control variate and change of measure/likelihood ratio), and Glasserman *et al.* (1999) (importance sampling), etc. Attempts to find closed or semi-closed form expressions for Asian option pricing in Black–Scholes models first appeared in the seminal work of Geman & Yor (1993). Among other interesting results in Geman & Yor (1993), the Laplace transform of the Asian option price with respect to time to maturity is derived and has been known as the celebrated Geman–Yor formula. An extension of the Geman–Yor approach to a jump diffusion model can be found in Cai & Kou (2012). However, numerical inversion of the Geman–Yor formula was shown slow and needed to be handled with care, see for example the discussions in Dufresne (2000) and Fu *et al.* (1999). Analytical approximation of the risk neutral density for the average price, which in turn yields an approximation of the Asian option, under Black–Scholes model dates back to the work of Turnbull & Wakeman (1991) and Ritchken *et al.* (1993). Both papers applied the Edgeworth expansion to the density of discretely monitored average price around lognormal distribution and obtained Black–Scholes type formula (up to correction terms) for the price of Asian option. Such analytical approximations are more appealing in practice than the Monte Carlo solutions because explicit expressions for the Greeks are readily accessible. Finally, PDE method is pursued by Rogers & Shi (1995) and Večeř (2001) (see also Večeř (2014) for a more recent development) after an ingeniously chosen change of variable.

Literatures on the pricing of Asian options under more general dynamics for the underlying security such as local volatility or stochastic volatility models, as opposed to those in the Black–Scholes model, are comparatively little. Less

ambitious approaches to the pricing of Asian options include arbitrage free bounds as in Albrecher *et al.* (2008) and Rogers & Shi (1995), and also approximative and asymptotic solutions such as Dassios & Nagaradjasarma (2006), Foschi *et al.* (2013), and Pagliarani *et al.* (2017). Approximations resorting to asymptotic expansions are mostly Itô–Taylor type expansion based as originated from the work of Kunitomo & Takahashi (1992), see also Cai *et al.* (2014) for a more recent development along this line. Despite being straightforward but tedious, such expansions usually require calculations up to third or fourth order in order to achieve satisfactory accuracy. Finally, though not directly related to the current paper, Asian option pricing under stochastic volatility models is discussed in Fouque & Han (2003) and Shiraya & Takahashi (2012).

In the current paper, we address the problem of approximating the price of options on the discrete arithmetic average, and its continuous-time limit, with the underlying following a local volatility model. For the discretely monitored Asian option, we assume that the average is over a set of equally spaced discrete time samples before expiry. The application of the Brownian bridge representation for the transition density (see Theorem 2.1 below) obtained in Wang & Gatheral (2005) between consecutive sampling time points leads to a "path-integral" type expression for the Asian option price, see (2.12). A direct application of a Laplace asymptotic formula (in this case high-dimensional, see Lemma 2.1) yields an approximation of the option price (see Theorem 2.2). In the limit where the sampling time window approaches zero, the leading order term (in small time to expiry) can be expressed as a constrained variational problem of finding an optimal path, referred to as the *most-likely-path* (MLP), in the time-price space. An approximation of the continuously monitored Asian option price is obtained once the variational problem is solved, see Definition 2.1 and Theorem 2.3. The MLP approximation coincides with a rigorous derivation of the leading order based on a recent extension of the Freidlin–Wentzell theorem for a large class of models, see Theorem 3.1. As for implied volatility, we opt to use the Bachelier model as benchmark rather than the Black–Scholes model because of the lack of analytical expression for Asian options in the Black–Scholes model. Such defined implied volatility in the European option case is sometimes referred to as the *implied normal volatility* in practice. By comparing corresponding expansions from the benchmark Bachelier model and from the local volatility model, we obtain, as the main result of the paper, the lowest order approximation of the implied normal volatility for Asian option in Theorem 2.5.

The paper is organized as follows. Section 2 lays out the model and provides derivations of the Laplace type approximation for discretely monitored Asian calls and the most-likely-path approximations for continuously monitored Asian call options. Section 3 is devoted to a rigorous derivation of the asymptotic behavior obtained in Sec. 2 based on large deviation theory. It was drawn to our attention that a rigorous proof of the most-likely-path approximation for the price of continuously monitored Asian option by the theory of large deviation has been done

independently in Pirjol & Zhu (2016). Finally, we conclude by presenting numerical tests of the most-likely-path approximation.

Throughout the text, $(W_t, t \geq 0)$ denotes the standard Brownian motion defined on the filtered probability space $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ satisfying the usual conditions. Dot will always refer to the partial derivative with respect to the time variable and prime to the space variable s.

2. Asian Option Pricing in Local Volatility Models

In this section, we derive asymptotic expansions of the prices of both discretely and continuously monitored Asian calls in local volatility models. First, the most-likely-path approximation in the discrete case is obtained in Theorem 2.2 using a Brownian bridge representation for the transition density obtained in Wang & Gatheral (2005) (see Theorem 2.1 below) and a high-dimensional Laplace asymptotic formula (Lemma 2.1). Second, an expression for the most-likely-path approximation in the continuous case is derived in Theorem 2.3 by taking a formal limit of the approximation in the discrete case. This provides a heuristic derivation for the leading order of the most-likely-path approximation for continuously monitored Asian calls.

We assume that the underlying asset S follows the local volatility model

$$dS_t = S_t \sigma_\ell(S_t, t) dW_t = a(S_t, t) dW_t \quad S_0 = s_0 > 0. \tag{2.1}$$

We will suppose that the diffusion function $a(s, t)$ is strictly positive (except possibly at $s = 0$ where it could be 0), and grows *at most linearly* in s: there exists $C > 0$ such that

$$0 \leq a(s, t) \leq C(1 + |s|) \quad \text{for all } t \in [0, T] \text{ and for all } s \in \mathbb{R}; \tag{2.2}$$

and is *locally Lipschitz*: for every $R > 0$, there exists C_R such that for all $t \in [0, T]$ and for every $x, y \in \mathbb{R}$ with $|s|, |s'| < R$

$$|a(s, t) - a(s', t)| \leq C_R |s - s'|. \tag{2.3}$$

These assumptions are sufficient to imply the existence and uniqueness of a strong solution to the SDE (2.1). Such solutions will also satisfy a *large deviation principle* based on the work in Chiarini & Fischer (2014) as explained in Sec. 3.

Let $p(T, s_T | t, s_t)$, $t < T$, be the transition probability density from (t, s_t) to (T, s_T) for the local volatility model (2.1). Consider the Lamperti transformation from s to x

$$x = \varphi(s, t) = \int_{s_0}^{s} \frac{d\xi}{a(\xi, t)} \quad \text{for } s > 0. \tag{2.4}$$

Since a is assumed locally Lipschitz and is strictly positive except at $s = 0$, the transformation is well-defined for all $s > 0$, except possibly at $s = 0$. The following representation for the transition density p is derived in Wang & Gatheral (2005).

Theorem 2.1. *Let $S = (S_t, t \geq 0)$ be the diffusion process given by (2.1). Define the function h by $h(x,t) = \varphi_t(s,t) - a_s(s,t)/2$, with $s = \varphi^{-1}(x,t)$, where φ is the Lamperti transformation (2.4) and subindices refer to corresponding partial derivatives. Then the transition density p of S from (t, s_t) to (T, s_T) has the representation:*

$$p(T, s_T | t, s_t) = \frac{g(T - t, \varphi(s_T, T) - \varphi(s_t, t))}{a(s_T, T)} \tilde{\mathbb{E}}_{\varphi(s_t, t), \varphi(s_T, T)}$$

$$\times \left[e^{\int_t^T h(X_s, s) dX_s - \frac{1}{2} \int_t^T h^2(X_s, s) ds} \right], \tag{2.5}$$

where g denotes the centered Gaussian density with variance t : $g(t, \xi) = \exp(-\xi^2/2t)/\sqrt{2\pi t}$ and $\tilde{\mathbb{E}}_{x,y}[\cdot]$ is the expectation under the Brownian bridge measure from x to y.

Equivalently, if H is an antiderivative of h with respect to x, namely, $\partial_x H(x,t) = h(x,t)$, for all x and t, then

$$p(T, s_T | t, s_t) = \frac{g(T - t, \varphi(s_T, T) - \varphi(s_t, t))}{a(s_T, T)} e^{H(\varphi(s_T, T), T) - H(\varphi(s_t, t), t)}$$

$$\times \tilde{\mathbb{E}}_{\varphi(s_t, t), \varphi(s_T, T)} \left[e^{-\frac{1}{2} \int_t^T h^2(X_s, s) + h_x(X_s, s) + 2H_t(X_s, s) ds} \right]. \tag{2.6}$$

For notational simplicity, hereafter we shall denote the expectation term in (2.5) as

$$\Psi(s_t, s_T) = \tilde{\mathbb{E}}_{\varphi(s_t, t), \varphi(s_T, T)} \left[e^{\int_t^T h(X_s, s) dX_s - \frac{1}{2} \int_t^T h^2(X_s, s) ds} \right]. \tag{2.7}$$

With this notation, we also have for the term in (2.6)

$$\Psi(s_t, s_T) = e^{H(\varphi(s_T, T), T) - H(\varphi(s_t, t), t)}$$

$$\times \tilde{\mathbb{E}}_{\varphi(s_t, t), \varphi(s_T, T)} \left[e^{-\frac{1}{2} \int_t^T h^2(X_s, s) + h_x(X_s, s) + 2H_t(X_s, s) ds} \right]. \tag{2.8}$$

2.1. *Small time asymptotic for discretely monitored Asian call*

Assume an (arithmetic) Asian call is sampled discretely at the time points $t_1 < t_2 < \cdots < t_n$ with $t_0 = 0$ and $t_n = T$. In other words, the payoff of such an Asian call is

$$\left(\frac{1}{n} \sum_{i=1}^n S_{t_i} - K \right)^+, \tag{2.9}$$

where the time interval between the sampling points is assumed equal, i.e. $t_i - t_{i-1} = \Delta t = T/n$ for $i = 1, \ldots, n$. By using the Brownian bridge representation (2.5) or (2.6), the joint density for $S_{t_1}, \ldots, S_{t_n}$ can be written as

$$p(t_1, s_{t_1} | t_0, s_{t_0}) p(t_2, s_{t_2} | t_1, s_{t_1}) \cdots p(t_n, s_{t_n} | t_{n-1}, s_{t_{n-1}}) \tag{2.10}$$

$$= \prod_{i=1}^n g(\Delta t, \varphi(s_{t_i}, t_i) - \varphi(s_{t_{i-1}}, t_{i-1})) \frac{\Psi(s_{t_i}, s_{t_{i-1}})}{a(s_{t_i}, t_i)}. \tag{2.11}$$

In this notation, the price $C = C(s_0, 0; K, T)$ of a discretely monitored Asian call struck at K can be written as

$$C = \mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n} S_{t_i} - K\right)^+\right]$$

$$= \iint \left(\frac{1}{n}\sum_{i=1}^{n} s_{t_i} - K\right)^+ \prod_{i=1}^{n} p(t_i, s_{t_i}|t_{i-1}, s_{t_{i-1}}) ds_{t_1} \cdots ds_{t_n}$$

$$= \frac{1}{(2\pi\Delta t)^{\frac{n}{2}}} \iint \left(\frac{1}{n}\sum_{i=1}^{n} s_{t_i} - K\right)^+ e^{-\frac{D(\mathbf{s},\mathbf{t})}{\Delta t}} W(\mathbf{s},\mathbf{t}) d\mathbf{s}, \qquad (2.12)$$

where

$$D(\mathbf{s},\mathbf{t}) = \frac{1}{2}\sum_{i=1}^{n} |\varphi(s_{t_i}, t_i) - \varphi(s_{t_{i-1}}, t_{i-1})|^2, \quad W(\mathbf{s},\mathbf{t}) = \prod_{i=1}^{n} \frac{\Psi(s_{t_{i-1}}, s_{t_i})}{a(s_{t_i}, t_i)}, \qquad (2.13)$$

for $\mathbf{t} = (t_1, \ldots, t_n)$, $\mathbf{s} = (s_{t_1}, \ldots, s_{t_n})$ and $d\mathbf{s} = ds_{t_1} \cdots ds_{t_n}$.

Hence, the evaluation of the Asian call price C becomes the computation of the multidimensional integral (2.12). The following Laplace asymptotic formula will be applied to estimate the n-dimensional integral in (2.12) for small Δt.

Lemma 2.1 (Laplace Asymptotic Formula). *Let R be a closed set in $\mathbb{R}^n$ with nonempty and smooth boundary ∂R. Suppose θ is a continuous function in R and attains its minimum uniquely at $x^* \in \partial R$ and, given any $\epsilon > 0$, there exists $\delta > 0$ such that $\theta(x) \geq \theta(x^*) + \delta$ for all $x \in R\backslash B_\epsilon(x^*)$, where $B_\epsilon(x^*) = \{x : |x - x^*| < \epsilon\}$ is the open ball of radius ϵ centered at x^*. Assume that f is integrable in R, i.e. $\int_R |f(x)|dx < \infty$ and that f vanishes identically in R^c and on the boundary ∂R but the inward normal directional derivative of f at x^* is nonzero. Then we have the asymptotic expansion as $\tau \to 0^+$*

$$\int_R e^{-\frac{\theta(x)}{\tau}} f(x)dx = \frac{(2\pi)^{\frac{n-1}{2}} \tau^{\frac{n+3}{2}} e^{-\frac{\theta(x^*)}{\tau}}}{\sqrt{\det \partial_{\mathbf{t}}^2 \theta(x^*)}|\nabla\theta(x^*)|} \left[\frac{\nabla f(x^*) \cdot \nabla\theta(x^*)}{|\nabla\theta(x^*)|^2}\right.$$

$$\left. + \frac{1}{2}\mathrm{tr}\{\partial_{\mathbf{t}}^2 f(x^*)[\partial_{\mathbf{t}}^2 \theta(x^*)]^{-1}\} + \mathcal{O}(\tau)\right], \qquad (2.14)$$

where $\partial_{\mathbf{t}}^2 f(x^*)$ and $\partial_{\mathbf{t}}^2 \theta(x^*)$ are the Hessian matrices of f and θ respectively in the tangential direction to R at x^*.

The proof of the lemma is standard and straightforward. See for instance Sec. 8.3 in Bleinstein & Handelsman (1986).

Now we apply the Laplace asymptotic formula (2.14) to the multidimensional integral (2.12) by taking $\theta = D(\mathbf{s},\mathbf{t})$, $f = (\frac{1}{n}\sum_{i=1}^{n} s_{t_i} - K)W(\mathbf{s},\mathbf{t})$, and R as the half space $R := \{\mathbf{s} : 1/n\sum_{i=1}^{n} s_{t_i} \geq K\}$. The crucial step in applying the Laplace asymptotic formula (2.14) is the determination of the minimum point of

D in the half space R, which boils down to solving the constrained optimization problem:

$$\min_{\mathbf{s}} \frac{1}{2} \sum_{i=1}^{n} |\varphi(s_{t_i}, t_i) - \varphi(s_{t_{i-1}}, t_{i-1})|^2 \tag{2.15}$$

subject to

$$\frac{1}{n} \sum_{i=1}^{n} s_{t_i} \geq K. \tag{2.16}$$

Remark 2.1. We shall assume $s_0 < K$ in the following for if $s_0 \geq K$, the value of the constrained optimization problem (2.15):(2.16) is 0 since one can simply take $s_{t_1} = s_{t_2} = \cdots = s_{t_n} = s_0$. Then $s_1 + \cdots + s_n = ns_0 \geq nK$ and $\varphi(s_{t_i}, t_i) = \varphi(s_{t_0}, t_i) = 0$ for all $1 \leq i \leq n$. Thus the objective function in (2.15) attains its global minimum 0.

Remark 2.2. We show in Appendix that, for Δt small enough, the objective function in (2.15) is in fact convex, which in turn implies that the minimizer, if there exists any, is unique since the constraint is a linear inequality.

We summarize the result for the price of a discretely monitored Asian call in Theorem 2.2 whose proof in the time homogeneous case is simply a direct application of the Laplace asymptotic formula (2.14) since the functions D and W are independent of $\mathbf{t}$. We remark that, modulo the exponential term, the result suggests the next order term is of order $3/2$ in Δt, which coincides with the order in the case for European options, see for example Theorem 2.3 in Gatheral *et al.* (2012).

Theorem 2.2 (Discrete Monitored Asian Option). *The price* $C = C(s_0, 0; K, T)$ *of a discretely monitored Asian call struck at* K *with* $K > s_0$ *and expiry time* T *has the following asymptotic expansion as* $T \to 0^+$, *for fixed* n,

$$C = \mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n} S_{t_i} - K\right)^+\right]$$

$$= \iint_{\{\mathbf{s}: \frac{1}{n}\sum_{i=1}^{n} s_{t_i} \geq K\}} \left(\frac{1}{n}\sum_{i=1}^{n} s_{t_i} - K\right) e^{-\frac{D(\mathbf{s},\mathbf{t})}{\Delta t}} W(\mathbf{s}, \mathbf{t}) d\mathbf{s}$$

$$= \frac{\Delta t^{\frac{3}{2}}}{\sqrt{2\pi}} \frac{e^{-\frac{D(\mathbf{s}^*,\mathbf{t})}{\Delta t}}}{|\nabla D(\mathbf{s}^*,\mathbf{t})|} \times \left[\frac{\nabla W(\mathbf{s}^*,\mathbf{t}) \cdot \nabla D(\mathbf{s}^*,\mathbf{t})}{\sqrt{\det \partial_{\mathbf{t}}^2 D(\mathbf{s}^*,\mathbf{t})}|\nabla D(\mathbf{s}^*,\mathbf{t})|^2}\right.$$

$$\left. + \frac{1}{2}\text{tr}\{\partial_{\mathbf{t}}^2 W(\mathbf{s}^*,\mathbf{t})[\partial_{\mathbf{t}}^2 D(\mathbf{s}^*,\mathbf{t})]^{-1}\} + \mathcal{O}(\Delta t)\right], \tag{2.17}$$

where $\mathbf{s}^* = (s_{t_1}^*, \ldots, s_{t_n}^*)$ *is the minimizer of the minimization problem* (2.15) *subject to the constraint* $1/n \sum_{i=1}^{n} s_{t_i} \geq K$.

2.2. *Continuously monitored Asian call and the most-likely-path*

The approximate price of an Asian call obtained in Theorem 2.2 is subject to solving a high-dimensional constrained optimization problem which is daunting in general. However, in the limit as Δt approaches zero, the optimization problem converges to a variational problem to which, in certain cases such as Black–Scholes and CIR, the associated Euler–Lagrange equations have closed form solution. The heuristic computation is given in this section. The rigorous derivation for the leading order using large deviation is done in Sec. 3.

Let $\{0 = t_0 < t_1 < \cdots < t_n = T\}$ be a partition of the interval $[0, T]$ with $t_i - t_{i-1} = \Delta t := T/n$, for $i = 1, \ldots, n$. Then, the price of a continuously monitored Asian call can be written as the limit of the prices of discretely monitored Asian calls as $n \to \infty$. Precisely,

$$\mathbb{E}\left[\left(\frac{1}{T}\int_0^T S_t dt - K\right)^+\right] = \lim_{n \to \infty} \mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^n S_{t_i} - K\right)^+\right] \qquad (2.18)$$

by applying the Lebesgue dominated convergence theorem. Hence, to the lowest order, it is natural to approximate the price of a continuously monitored Asian call by taking the limit of the approximate price of discretely monitored Asian call in (2.17) as $\Delta t \to 0$. To be specific, rewrite the logarithm of (2.17) as

$$\log C = -\frac{D(\mathbf{s}^*, \mathbf{t})}{\Delta t} - \frac{1}{2}\log(2\pi) + \frac{3}{2}\log(\Delta t) - \log|\nabla D(\mathbf{s}^*, \mathbf{t})|$$

$$+ \log\left[\frac{\nabla W(\mathbf{s}^*, \mathbf{t}) \cdot \nabla D(\mathbf{s}^*, \mathbf{t})}{\sqrt{\det \partial_{\mathbf{t}}^2 D(\mathbf{s}^*, \mathbf{t})}|\nabla D(\mathbf{s}^*, \mathbf{t})|^2}\right.$$

$$\left.+ \frac{1}{2}\text{tr}\{\partial_{\mathbf{t}}^2 W(\mathbf{s}^*, \mathbf{t})[\partial_{\mathbf{t}}^2 D(\mathbf{s}^*, \mathbf{t})]^{-1}\} + \mathcal{O}(\Delta t)\right], \qquad (2.19)$$

where we recall that $\mathbf{s}^*$ is the minimizer of (2.15). Note that the first term in the last expression is dominant as $\Delta t \to 0$.

To determine the limit as $\Delta t \to 0$ of the dominant term, we bring Δt back to the objective function (2.15) as

$$\frac{1}{2\Delta t}\sum_{k=1}^n \left|\varphi(s_{t_i}, t_i) - \varphi(s_{t_{i-1}}, t_{i-1})\right|^2. \qquad (2.20)$$

Note that, since

$$\varphi(s_{t_i}, t_i) - \varphi(s_{t_{i-1}}, t_{i-1})$$

$$= \varphi_s(s_{t_{i-1}}, t_{i-1})\Delta s_{t_i} + \varphi_t(s_{t_{i-1}}, t_{i-1})\Delta t + o(\Delta s_{t_i}, \Delta t), \qquad (2.21)$$

where $\Delta s_{t_i} = s_{t_i} - s_{t_{i-1}}$, we have

$$\lim_{\Delta t \to 0} \frac{1}{2\Delta t} \sum_{k=1}^{n} |\varphi(s_{t_i}, t_i) - \varphi(s_{t_{i-1}}, t_{i-1})|^2$$

$$= \lim_{\Delta t \to 0} \frac{1}{2\Delta t} \sum_{k=1}^{n} |\varphi_s(s_{t_{i-1}}, t_{i-1}) \Delta s_{t_i}|^2 + o((\Delta s_{t_i})^2, \Delta t)$$

$$= \lim_{\Delta t \to 0} \frac{1}{2} \sum_{k=1}^{n} \left| \frac{\frac{\Delta s_i}{\Delta t}}{a(s_{i-1}, t_{i-1})} \right|^2 \Delta t$$

$$= \frac{1}{2} \int_0^T \left[\frac{\dot{s}(t)}{a(s(t), t)} \right]^2 dt. \tag{2.22}$$

Therefore, in the limit as Δt approaches zero, the optimization problem (2.15) becomes the following variational problem[a] on the space of paths $s : t \mapsto s(t)$

$$\min_{s: \ t \mapsto s(t)} \frac{1}{2} \int_0^T \left[\frac{\dot{s}(t)}{a(s(t), t)} \right]^2 dt \tag{2.23}$$

subject to

$$\frac{1}{T} \int_0^T s(t) dt = K, \quad s(0) = s_0. \tag{2.24}$$

Definition 2.1. The optimal path of the variational problem (2.23):(2.24) is referred to as the *most-likely-path* (MLP) for the Asian call struck at K.

In view of the above, one expects that, at the heuristic level and for the leading order, the logarithm of the price C of an out-of-the-money Asian call is approximately given by the solution to the constrained variational problem (2.23):(2.24), which is equivalent to determining the most-likely-path. To determine the MLP, the Euler–Lagrange equation associated with the constrained variational problem (2.23):(2.24) along with proper boundary conditions is derived as follows.

Lemma 2.2. *The optimal path of the constrained variational problem* (2.23):(2.24) *satisfies the Euler–Lagrange equation*

$$\frac{d}{dt} \left(\frac{\dot{s}}{a} \right) - \frac{a_t}{a^2} \dot{s} + \frac{\lambda}{T} a = 0 \tag{2.25}$$

with boundary conditions

$$s(0) = s_0, \quad \dot{s}(T) = 0, \tag{2.26}$$

where λ is chosen such that $1/T \int_0^T s(t) dt = K$.

[a]An equivalent formulation of the same problem in the Black–Scholes case was obtained in Cibelli *et al.* (2016) [see (1.4)] for the analysis of fundamental solution.

Proof. We first rewrite the constrained variational problem (2.23):(2.24) in the Lagrangian form

$$L(s, \lambda) = \frac{1}{2} \int_0^T \left[\frac{\dot{s}(t)}{a(s(t), t)} \right]^2 dt - \lambda \left(\frac{1}{T} \int_0^T s(t) dt - K \right), \quad (2.27)$$

where λ is the Lagrange multiplier. Let $f : [0, T] \mapsto \mathbb{R}$ be a perturbation around the optimal path $s(t)$ with $f(0) = 0$. The first order criterion of optimality yields

$$0 = \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} L(s + \epsilon f, \lambda) \quad (2.28)$$

$$= \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \frac{1}{2} \int_0^T \left[\frac{\dot{s} + \epsilon \dot{f}}{a(s + \epsilon f, t)} \right]^2 dt - \lambda \left(\frac{1}{T} \int_0^T \{s + \epsilon f\} dt - K \right) \quad (2.29)$$

$$= \int_0^T \left[\frac{\dot{s}}{a(s, t)} \right] \left[\frac{\dot{f}}{a(s, t)} - \frac{a_s(s, t) \dot{s} f}{a^2(s, t)} \right] dt - \frac{\lambda}{T} \int_0^T f \, dt \quad (2.30)$$

$$= - \int_0^T \left\{ \frac{a_s(s, t)}{a^3(s, t)} \dot{s}^2 + \frac{\lambda}{T} \right\} f \, dt + \int_0^T \frac{\dot{s} \dot{f}}{a^2(s, t)} dt. \quad (2.31)$$

By applying integration by parts to the second integral and noting that $f(0) = 0$, the last equality becomes

$$0 = - \int_0^T \left\{ \frac{a_s(s, t)}{a^3(s, t)} \dot{s}^2 + \frac{\lambda}{T} \right\} f \, dt - \int_0^T \frac{d}{dt} \left[\frac{\dot{s}}{a^2(s, t)} \right] f \, dt + \frac{\dot{s}(T)}{a^2(s(T), T)} f(T). \quad (2.32)$$

Finally, since f is arbitrary and $a(s, t) > 0$, we obtain the Euler–Lagrange equation

$$\frac{a_s(s, t)}{a^3(s, t)} \dot{s}^2 + \frac{\lambda}{T} + \frac{d}{dt} \left[\frac{\dot{s}}{a^2(s, t)} \right] = 0, \quad (2.33)$$

which simplifies to

$$\frac{d}{dt} \left(\frac{\dot{s}}{a} \right) - \frac{a_t}{a^2} \dot{s} + \frac{\lambda}{T} a = 0 \quad (2.34)$$

with boundary conditions $s(0) = s_0$, $\dot{s}(T) = 0$. □

We summarize the final result in the following theorem. A rigorous proof based on large deviation theory is postponed to Theorem 3.1 in Sec. 3.

Theorem 2.3 (Log Price of a Continuously Monitored Asian Call). *The price $C(S_t, t; K, T)$ at time t of a continuously monitored out-of-the-money Asian call struck at $K > s_0$ and expiry time T, is approximately equal to*

$$\log C(S_t, t; K, T) = -\frac{1}{2} \int_t^T \left| \frac{\dot{s}(\tau)}{a(s(\tau), \tau)} \right|^2 d\tau + o(T - t)^{-1}, \quad (2.35)$$

where $s : t \mapsto s(t)$ is the solution to the constrained variational problem (2.23):(2.24).

We conclude the section by deriving closed form expressions for the most-likely-path in the Bachelier, the Black–Scholes, and the Cox–Ingersoll–Ross (CIR) models by solving their corresponding boundary value problems (2.25):(2.26). However, for more general cases we will have to resort to an iteration scheme for numerical computations of the most-likely-path. See Sec. 4 for a numerical scheme.

Eaxmple 2.1 (Bachelier Model). In the Bachelier model, $a(s,t) = \sigma$, a constant. The Euler–Lagrange equation (2.25) reduces to

$$\frac{d}{dt}\left(\frac{\dot{s}}{\sigma}\right) + \frac{\lambda}{T}\sigma = 0, \tag{2.36}$$

whose general solution is $s(t) = -\lambda\sigma^2 t^2/(2T) + c_1 t + c_0$. From the boundary conditions (2.26) together with $1/T \int_0^T s(t)dt = K$, we conclude that $\lambda = 3(K - s_0)/(\sigma^2 T^2)$, $c_1 = 3(K - s_0)/T$, and $c_0 = s_0$. Therefore, the most likely path for Asian call in the Bachelier model is a downward parabola in the t-s plane given by

$$s(t) = -\frac{3(K - s_0)}{2}\left(\frac{t}{T}\right)^2 + 3(K - s_0)\frac{t}{T} + s_0. \tag{2.37}$$

We remark that the most-likely-path in this case does not depend on σ. Also, (2.35) for the Bachelier model becomes

$$-\frac{1}{2}\int_0^T \left[\frac{\dot{s}(t)}{a(s(t),t)}\right]^2 dt = -\frac{3(K - s_0)^2}{2\sigma^2 T}. \tag{2.38}$$

Thus, it captures the exponential decay of an out-of-the-money Asian call in the Bachelier model, see (2.64) below.

Eaxmple 2.2 (Black–Scholes Model). In the Black–Scholes model, $a(s,t) = \sigma s$ with σ being a constant. The Euler–Lagrange equation (2.25) reads

$$-\frac{\ddot{s}}{s^2} + \frac{\dot{s}^2}{s^3} = \lambda\sigma^2 \tag{2.39}$$

to which the general solutions are

$$s(t) = \frac{c_1^2}{2\lambda\sigma^2}\left[1 - \tanh^2\left(\frac{c_1}{2}[c_2 - t]\right)\right] \tag{2.40}$$

and

$$s(t) = -\frac{c_1^2}{2\lambda\sigma^2}\left[1 + \tan^2\left(\frac{c_1}{2}[c_2 - t]\right)\right], \tag{2.41}$$

where c_1 and c_2 are (to be determined) constants. The condition $\dot{s}(T) = 0$ implies that $c_2 = T$ in either case and the conditions $s(0) = s_0$, $\int_0^T s(t)dt = TK$ imply

respectively that

$$\frac{c_1^2}{\lambda\sigma^2(1+\cosh(c_1 T))} = s_0, \quad c_1 \tanh\left(\frac{c_1 T}{2}\right) = \lambda\sigma^2 K T \qquad (2.42)$$

and

$$\frac{-c_1^2}{\lambda\sigma^2(1+\cos(c_1 T))} = s_0, \quad -c_1 \tan\left(\frac{c_1 T}{2}\right) = \lambda\sigma^2 K T. \qquad (2.43)$$

By dividing the two equations in (2.42) and rearranging terms, c_1 is given by the solution to the equation

$$\frac{c_1 T}{\sinh(c_1 T)} = \frac{s_0}{K} \quad \text{if } s_0 < K \qquad (2.44)$$

or $c_1 = f^{-1}(s_0/K)/T$, where $f(x) = x/\sinh x$. On the other hand, from (2.43) we obtain that

$$\frac{\sin(c_1 T)}{c_1 T} = \frac{K}{s_0} \quad \text{if } K > s_0 \qquad (2.45)$$

or $c_1 = g^{-1}(K/s_0)/T$, where $g(x) = \sin x/x$ for $x \in [0,\pi]$.

Hence, the most-likely-path $s(t)$ for $0 \le t \le T$ in the Black–Scholes model after simplification becomes

$$s(t) = \frac{s_0 \cosh^2\left(\dfrac{c_1 T}{2}\right)}{\cosh^2\left(\dfrac{c_1(T-t)}{2}\right)}, \qquad (2.46)$$

where $c_1 = f^{-1}(s_0/K)/T$ for $s_0 < K$; whereas for $s_0 > K$

$$s(t) = \frac{s_0 \cos^2\left(\dfrac{c_1 T}{2}\right)}{\cos^2\left(\dfrac{c_1(T-t)}{2}\right)} \qquad (2.47)$$

and $c_1 = g^{-1}(K/s_0)/T$. Moreover, for $s_0 < K$,

$$\frac{1}{2}\int_0^T \left|\frac{\dot{s}(t)}{\sigma s(t)}\right|^2 dt = \frac{f^{-1}\left(\dfrac{s_0}{K}\right)}{\sigma^2 T}\left[\frac{1}{2}f^{-1}\left(\frac{s_0}{K}\right) - \tanh\left(\frac{1}{2}f^{-1}\left(\frac{s_0}{K}\right)\right)\right], \qquad (2.48)$$

and for $s_0 > K$

$$\frac{1}{2}\int_0^T \left|\frac{\dot{s}(t)}{\sigma s(t)}\right|^2 dt = \frac{g^{-1}\left(\dfrac{K}{s_0}\right)}{\sigma^2 T}\left[\tan\left(\frac{1}{2}g^{-1}\left(\frac{K}{s_0}\right)\right) - \frac{1}{2}g^{-1}\left(\frac{K}{s_0}\right)\right]. \qquad (2.49)$$

Eaxmple 2.3 (Cox–Ingersoll–Ross Model). For the CIR model, $a(s,t) = \sigma\sqrt{s}$, where σ is a constant. The Euler–Lagrange equation (2.25) becomes

$$\frac{d}{dt}\left(\frac{\dot{s}(t)}{\sqrt{s(t)}}\right) + 2\phi\sqrt{s(t)} = 2\frac{d^2}{dt^2}(\sqrt{s(t)}) + 2\phi\sqrt{s(t)} = 0 \qquad (2.50)$$

for some constant ϕ with general solution given by

$$\sqrt{s(t)} = c_2 \cos(\sqrt{\phi}(c_1 - t)) \tag{2.51}$$

if $\phi > 0$ and

$$\sqrt{s(t)} = c_2 \cosh(\sqrt{-\phi}(c_1 - t)) \tag{2.52}$$

if $\phi < 0$. The boundary conditions (2.26) imply respectively that

$$c_1 = T \quad \text{and} \quad c_2 = \frac{\sqrt{s_0}}{\cos(\sqrt{\phi}T)} \quad \text{or} \quad \frac{\sqrt{s_0}}{\cosh(\sqrt{-\phi}T)}. \tag{2.53}$$

Thus

$$s(t) = s_0 \left\{ \frac{\cos(\sqrt{\phi}(T-t))}{\cos(\sqrt{\phi}T)} \right\}^2 \quad \text{or} \quad s(t) = s_0 \left\{ \frac{\cosh(\sqrt{-\phi}(T-t))}{\cosh(\sqrt{-\phi}T)} \right\}^2. \tag{2.54}$$

The parameter ϕ is determined by the solution to the equation

$$\frac{1}{T} \int_0^T s(t)\, dt = K = \frac{s_0}{2\sqrt{\phi}T} [\tan(\sqrt{\phi}T) + \sqrt{\phi}T \sec^2(\sqrt{\phi}T)] \tag{2.55}$$

if $s_0 < K$ and by

$$\frac{K}{s_0} = \frac{1}{2\sqrt{-\phi}T} [\tanh(\sqrt{-\phi}T) + \sqrt{-\phi}T \operatorname{sech}^2(\sqrt{-\phi}T)] \tag{2.56}$$

if $s_0 > K$. Finally, we have, subject to the determination of ϕ, that

$$\int_0^T \left| \frac{\dot{s}(t)}{\sigma \sqrt{s(t)}} \right|^2 dt = \frac{s_0 \sqrt{\phi}}{\sigma^2} [2\sqrt{\phi}T - \sin(2\sqrt{\phi}T)] \sec^2(\sqrt{\phi}T) \tag{2.57}$$

for $s_0 < K$ and

$$\int_0^T \left| \frac{\dot{s}(t)}{\sigma \sqrt{s(t)}} \right|^2 dt = \frac{s_0 \sqrt{-\phi}}{\sigma^2} [-2\sqrt{-\phi}T + \sinh(2\sqrt{-\phi}T)] \operatorname{sech}^2(\sqrt{-\phi}T) \tag{2.58}$$

if $s_0 > K$.

2.3. *Implied normal volatility for Asian options*

As far as implied volatility is concerned, we opt to use the Bachelier model as benchmark rather than the conventional Black–Scholes model because of the lack of analytical expression for Asian options in the Black–Scholes model. Such defined implied volatility in the European counterpart is sometimes used and referred to as the *implied normal volatility* in practice. For European calls, this approximation is good whenever $\sigma\sqrt{T}$ is small, at least for at-the-money options, see for example Schachermayer & Teichmann (2007). We expect the same should hold for Asian options, though the lack of analyticity makes it hard to check. In particular, note that the approximation might be problematic for small strike prices as the ratio S/K will be large.

Recall that, by straightforward calculations, under the Bachelier model

$$dS_t = \sigma_b dW_t, \tag{2.59}$$

where σ_b is a constant, the price of a continuously monitored Asian call option struck at K with expiry T has the closed form expression

$$C_b(K, T, \sigma_b) = \frac{\sigma_b \sqrt{T}}{\sqrt{6\pi}} e^{-\frac{3(s_0 - K)^2}{2\sigma_b^2 T}} + (s_0 - K)N\left(\frac{\sqrt{3}(s_0 - K)}{\sigma_b \sqrt{T}}\right) \tag{2.60}$$

since the average price $1/T \int_0^T S_t dt$ is normally distributed with mean s_0 and variance $\sigma_b^2 T/3$. In (2.60), $N(\cdot)$ denotes the cumulative distribution function of standard normal distribution. On the other hand, for a discretely monitored Asian call struck at K with expiry T in the Bachelier model, since $S_{t_1} + \cdots + S_{t_n}$ is normally distributed with

$$\mathbb{E}[S_{t_1} + \cdots + S_{t_n}] = ns_0, \tag{2.61}$$

$$\text{var}[S_{t_1} + \cdots + S_{t_n}] = \sigma_b^2 T \frac{(n+1)(2n+1)}{6}, \tag{2.62}$$

its price C_b^d is given by

$$C_b^d(K, T, \sigma_b) = \frac{\sigma_b \sqrt{T}}{\sqrt{2A_n \pi}} e^{-\frac{A_n(s_0 - K)^2}{2\sigma_b^2 T}} + (s_0 - K)N\left(\frac{\sqrt{A_n}(s_0 - K)}{\sigma_b \sqrt{T}}\right), \tag{2.63}$$

where $A_n := 6n^2/(n+1)(2n+1)$. Apparently, $A_n \to 3$ as $n \to \infty$.

Regarding the small time expansion, we remark that by using the asymptotic expansion $N(x) = N'(x)[-1/x + 1/x^3 + \mathcal{O}(1/x^5)]$ as $x \to -\infty$, we have

$$C_b(s_t, K) = \frac{\sigma_b \sqrt{T-t}}{\sqrt{6\pi}} e^{-\frac{3(s_t - K)^2}{2\sigma_b^2 (T-t)}} + (s_t - K)N\left(\frac{\sqrt{3}(s_t - K)}{\sigma_b \sqrt{T-t}}\right)$$

$$= e^{-\frac{3(s_t - K)^2}{2\sigma_b^2 (T-t)}} \left[\frac{\sigma_b^3 (T-t)^{\frac{3}{2}}}{3\sqrt{6\pi}(K - s_0)^2} + \mathcal{O}(T-t)^{\frac{5}{2}}\right] \tag{2.64}$$

as $t \to T^-$. Similarly, in the discrete case as $t \to T$,

$$C_b^d(s_t, K) = e^{-\frac{A_n(s_t - K)^2}{2\sigma_b^2 (T-t)}} \left[\frac{\sigma_b^3 (T-t)^{\frac{3}{2}}}{A_n \sqrt{2A_n \pi}(K - s_0)^2} + \mathcal{O}(T-t)^{\frac{5}{2}}\right]. \tag{2.65}$$

We remark that once the small time asymptotic for the price of an out-of-the-money call is established on the model side, it is a common practice to derive the small time asymptotic of implied volatility thereby. To that end, the following expansion for the implied normal volatility in terms of the call price given in Grunspan (2011) will be helpful.

Proposition 2.1. *For a fixed strike out-of-the-money call with time to expiry T, let $C = C(T)$ be the price of the European call regarded as a function of T. Then*

as time to expiry $T \to 0$, the implied normal volatility σ_N has the asymptotic

$$\sigma_N^2 T = \frac{(s_0 - K)^2}{2(\log s_0 - \log C(T))} + o(\log C(T)), \qquad (2.66)$$

as $T \to 0$.

2.3.1. *Implied normal volatility for discretely monitored Asian option*

The implied normal volatility for a discretely monitored Asian call is defined by solving the following equation for σ_b

$$C^d(s_t, K, T) = \frac{\sigma_b \sqrt{T}}{\sqrt{2A_n \pi}} e^{-\frac{A_n(s_0 - K)^2}{2\sigma_b^2 T}} + (s_0 - K)N\left(\frac{\sqrt{A_n}(s_0 - K)}{\sigma_b \sqrt{T}}\right), \qquad (2.67)$$

where $A_n := 6n^2/(n+1)(2n+1)$ and C^d is the price of a discretely monitored Asian call obtained by the model or from the market. Obviously, among other parameters, such defined σ_b depends on K and T.

To derive an asymptotic expansion for the implied normal volatility σ_b defined in (2.67) in small time, the idea, as in Gatheral *et al.* (2012), is to compare the corresponding terms in the expansions on both side of (2.67). The lowest order term is thus obtained by matching the exponential terms on both side of (2.67). Precisely, recall the small time asymptotic of the price of a discrete Asian call from Theorem 2.2

$$\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n} S_{t_i} - K\right)^+\right] = \frac{\Delta t^{\frac{3}{2}}}{\sqrt{2\pi}} \frac{e^{-\frac{D(\mathbf{s}^*, \mathbf{t})}{\Delta t}}}{|\nabla D(\mathbf{s}^*, \mathbf{t})|} \times \left[\frac{\nabla W(\mathbf{s}^*, \mathbf{t}) \cdot \nabla D(\mathbf{s}^*, \mathbf{t})}{\sqrt{\det \partial_{\mathbf{t}}^2 D(\mathbf{s}^*, \mathbf{t})}|\nabla D(\mathbf{s}^*, \mathbf{t})|^2}\right.$$

$$\left. + \frac{1}{2}\mathrm{tr}\{\partial_{\mathbf{t}}^2 W(\mathbf{s}^*, \mathbf{t})[\partial_{\mathbf{t}}^2 D(\mathbf{s}^*, \mathbf{t})]^{-1}\} + \mathcal{O}(\Delta t)\right],$$

where $\Delta t = T/n$. By matching the exponential terms in the above expression and in (2.65), we have

$$e^{-\frac{A_n(s_t - K)^2}{2\sigma_b^2 T}} = e^{-\frac{D(\mathbf{s}^*, \mathbf{t})}{\Delta t}}$$

$$\Rightarrow \frac{A_n(s_t - K)^2}{2\sigma_b^2 T} = \frac{D(\mathbf{s}^*, \mathbf{t})}{\Delta t} = \frac{nD(\mathbf{s}^*, \mathbf{t})}{T}$$

$$\Rightarrow \frac{1}{\sigma_b^2} = \frac{2nD(\mathbf{s}^*, \mathbf{t})}{A_n(s_t - K)^2} = \frac{n}{A_n(s_t - K)^2}\sum_{k=1}^{n}|\varphi(s_{t_i}^*, t_i) - \varphi(s_{t_{i-1}}^*, t_{i-1})|^2.$$

$$(2.68)$$

Hence, the lowest (zeroth) order approximation of the implied normal volatility is given by

$$\sigma_b = \sqrt{A_n}|s_t - K|\left(n\sum_{k=1}^{n}|\varphi(s_{t_i}^*, t_i) - \varphi(s_{t_{i-1}}^*, t_{i-1})|^2\right)^{-\frac{1}{2}} + o(1), \qquad (2.69)$$

where recall that $(s_{t_1}^*, \ldots, s_{t_n}^*)$ is the solution to the n-dimensional constrained optimization problem (2.15):(2.16). We summarize the result in the following theorem.

Theorem 2.4 (Implied Normal Volatility Asymptotic for Discrete Asian Call). *For a discretely monitored out-of-the-money Asian call struck at K, i.e. $s_0 < K$, in which the underlying is driven by the local volatility model (2.1), the implied normal volatility σ_b^d defined as in (2.67) has the asymptotic expansion as $T \to 0^+$*

$$\sigma_b^d(K, T) = \sigma_{b,0}^d(K) + \mathcal{O}(T), \tag{2.70}$$

where

$$\sigma_{b,0}^d(K) = \sqrt{A_n}|s_t - K| \left(n \sum_{k=1}^n |\varphi(s_{t_i}^*, t_i) - \varphi(s_{t_{i-1}}^*, t_{i-1})|^2 \right)^{-\frac{1}{2}}, \tag{2.71}$$

with $A_n = 6n^2/(n+1)(2n+1)$ and $(s_{t_1}^*, \ldots, s_{t_n}^*)$ the solution to the constrained optimization problem (2.15):(2.16).

2.3.2. *Implied normal volatility for continuously monitored Asian option*

For continuously monitored Asian calls, the implied normal volatility is thus defined by solving the following equation for σ_b

$$C(s_t, K, T) = \frac{\sigma_b \sqrt{T-t}}{\sqrt{6\pi}} e^{-\frac{3(s_t - K)^2}{2\sigma_b^2(T-t)}} + (s_t - K)N\left(\frac{\sqrt{3}(s_t - K)}{\sigma_b \sqrt{T-t}} \right), \tag{2.72}$$

where C is the price of an Asian call obtained by the model or from the market. Notice that the right-hand side of (2.72) can be regarded as the price of a European call option in the Bachelier model but with one third of σ_b^2 as the variance (volatility squared) parameter. In other words, in the Bachelier world, the price of an Asian call is equal to the price of its European counterpart but with only one third of variance. By combining the asymptotics (2.35) for the call price and (2.66) for the implied normal volatility, a small time asymptotic for the implied normal volatility of an out-of-the-money Asian call is established. Alternatively, we may also obtain the same approximation by straightforwardly taking the limit as $n \to \infty$ of $\sigma_{b,0}^d$ in Theorem 2.4. We summarize the result in the following theorem.

Theorem 2.5 (Implied Normal Volatility Asymptotic for Continuous Asian Call). *For a continuously monitored out-of-the-money Asian call struck at K, i.e. $s_0 < K$, in which the underlying is driven by the local volatility model (2.1), the implied normal volatility σ_b defined in (2.72) has the asymptotic expansion as $T \to 0^+$*

$$\sigma_b(K, T) = \sigma_{b,0} + o(T), \tag{2.73}$$

where

$$\sigma_{b,0} = \left(\frac{T}{3(K-s_0)^2} \int_0^T \left[\frac{\dot{\tilde{s}}(t)}{a(\tilde{s}(t),t)} \right]^2 dt \right)^{-\frac{1}{2}}, \tag{2.74}$$

and $\tilde{s}(t)$ is the most-likely-path for Asian option determined by solving the variational problem (2.23):(2.24)*.*

Proof. Recall from (2.66) that the implied normal volatility σ of a European option has the asymptotic

$$\sigma^2 T = \frac{(s_0 - K)^2}{2(\log s_0 - \log C(T))} + o(\log C(T)). \tag{2.75}$$

Thus, by substituting $\log C(T)$ with (2.35) and using the fact that Asian variance equals one third of its European counterpart in Bachelier world, we obtain for implied normal volatility σ_b of an Asian call

$$\frac{\sigma_b^2 T}{3} = \frac{(s_0 - K)^2}{2(\log s_0 - \log C(T))} + o(\log C(T)) \tag{2.76}$$

$$= -\frac{(K-s_0)^2}{2\log C(T)} + o(\log C(T)) \tag{2.77}$$

$$= (K-s_0)^2 \left[\int_0^T \left[\frac{\dot{\tilde{s}}(t)}{a(\tilde{s}(t),t)} \right]^2 dt \right]^{-1} + o(T), \tag{2.78}$$

where we used (2.35) in the last equality. Finally, the result is obtained by rearranging terms. $\square$

Eaxmple 2.4 (Time-Dependent Bachelier Model). Finally, we consider a time-dependent Bachelier model in which

$$dS_t = \sigma\theta(t)dW_t. \tag{2.79}$$

Note that in this case $a(s,t) = \sigma\theta(t)$, the Euler–Lagrange equation (2.25) reduces to

$$\frac{d}{dt}\left(\frac{\dot{s}}{\sigma\theta(t)} \right) - \frac{\theta'(t)\dot{s}}{\sigma\theta^2(t)} + \frac{\lambda}{T}\sigma\theta(t) = 0, \tag{2.80}$$

$$\Rightarrow \frac{1}{\theta(t)}\frac{d}{dt}\left(\frac{\dot{s}}{\theta(t)} \right) - \frac{\theta'(t)}{\theta^2(t)}\frac{\dot{s}}{\theta(t)} + \frac{\lambda}{T}\sigma^2 = 0. \tag{2.81}$$

Integrating the last equation subject to the condition $\dot{s}(T) = 0$ gives

$$\frac{\dot{s}(t)}{\theta^2(t)} = \frac{\lambda}{T}\sigma^2(T-t) \tag{2.82}$$

for some constant λ set using the condition

$$\frac{1}{T}\int_0^T s(t)\,dt = K. \tag{2.83}$$

This gives

$$\lambda = \frac{T^2(K - s_0)}{\sigma^2 \int_0^T (T - u)^2 \theta^2(u) du}. \tag{2.84}$$

Also, from (2.82) one easily obtain

$$\int_0^T \left| \frac{\dot{s}(t)}{a(s, t)} \right|^2 dt = \int_0^T \left| \frac{\dot{s}(t)}{\sigma \theta(t)} \right|^2 dt = \frac{\lambda^2 \sigma^2}{T^2} \int_0^T (T - t)^2 \theta^2(t) \, dt$$

$$= \frac{T^2(K - s_0)^2}{\sigma^2 \int_0^T (T - t)^2 \theta^2(t) dt}. \tag{2.85}$$

This gives

$$\sigma_{b,0}^2 = \frac{3\sigma^2}{T^3} \int_0^T (T - t)^2 \theta^2(t) \, dt. \tag{2.86}$$

On the other hand, we have $\bar{S}_T = 1/T \int_0^T S_t dt$ where

$$S_t = s_0 + \sigma \int_0^t \theta(u) dW_u. \tag{2.87}$$

A straightforward computation then yields

$$\text{var}[\bar{S}_T] = \frac{\sigma^2}{T^2} \int_0^T (T - u)^2 \theta(u)^2 \, du. \tag{2.88}$$

So the *exact* Bachelier implied volatility in the time-dependent Bachelier case is given by

$$\sigma_b^2 = \frac{3\sigma^2}{T^3} \int_0^T (T - u)^2 \theta(u)^2 \, du. \tag{2.89}$$

Thus, the most-likely-path approximation is exact in the time-dependent Bachelier case.

2.4. *Approximation of Greeks*

The approximation of implied volatility in Theorem 2.5 is also applicable for approximations of Greeks. For example, we may calculate the delta as follows. For notational simplicity, denote by $v_b := \sigma_b^2 T$. Suppressing and holding the other parameters fixed, since the Bachelier implied volatility is defined through

$$C(s) = C_b(s, v_b), \tag{2.90}$$

the delta Δ satisfies

$$\Delta := \frac{\partial C}{\partial s} = \frac{\partial C_b}{\partial s} + \frac{\partial C_b}{\partial v_b} \frac{\partial v_b}{\partial s}, \tag{2.91}$$

where C_b is the function defined in (2.72). Note that by straightforward calculations we have

$$\frac{\partial C_b}{\partial s} = N\left(\frac{\sqrt{3}(s-K)}{\sqrt{v_b}}\right) \quad \text{and} \quad \frac{\partial C_b}{\partial v_b} = \frac{1}{2\sqrt{6\pi v_b}}e^{-\frac{3(s-K)^2}{2v_b}}. \tag{2.92}$$

Thus, with $v_b \approx v_{b,0} := \sigma_{b,0}^2 T$, it follows that

$$\Delta \approx N\left(\frac{\sqrt{3}(s-K)}{\sqrt{v_{b,0}}}\right) + \frac{1}{2\sqrt{6\pi v_{b,0}}}e^{-\frac{3(s-K)^2}{2v_{b,0}}}\frac{\partial v_{b,0}}{\partial s}. \tag{2.93}$$

The expressions on the right-hand side of (2.93) can be calculated easily except the last term which can be calculated as follows. Recall that

$$v_{b,0} = 3(K-s)^2\left[\int_0^T \left[\frac{\dot{\tilde{s}}(t)}{a(\tilde{s}(t),t)}\right]^2 dt\right]^{-1}, \tag{2.94}$$

we have

$$\frac{\partial v_{b,0}}{\partial s} = 6(s-K)\left[\int_0^T \left[\frac{\dot{\tilde{s}}(t)}{a(\tilde{s}(t),t)}\right]^2 dt\right]^{-1}$$

$$-3(K-s)^2\left[\int_0^T \left[\frac{\dot{\tilde{s}}(t)}{a(\tilde{s}(t),t)}\right]^2 dt\right]^{-2}\frac{\partial}{\partial s}\int_0^T \left[\frac{\dot{\tilde{s}}(t)}{a(\tilde{s}(t),t)}\right]^2 dt$$

$$= \frac{2v_{b,0}}{s-K} - \frac{v_{b,0}^2}{3(K-s)^2}\frac{\partial}{\partial s}\int_0^T \left[\frac{\dot{\tilde{s}}(t)}{a(\tilde{s}(t),t)}\right]^2 dt. \tag{2.95}$$

The integral in (2.95) in the general case needs to be evaluated numerically. Similarly, since the gamma Γ satisfies

$$\Gamma := \frac{\partial \Delta}{\partial s} = \frac{\partial^2 C_b}{\partial s^2} + 2\frac{\partial^2 C_b}{\partial s \partial v_b}\frac{\partial v_b}{\partial s} + \frac{\partial^2 C_b}{\partial v_b^2}\left(\frac{\partial v_b}{\partial s}\right)^2 + \frac{\partial C_b}{\partial v_b}\frac{\partial^2 v_b}{\partial s^2}, \tag{2.96}$$

an approximation of Γ is given by substituting v_b with $v_{b,0}$, subject to numerically evaluations of the derivatives $\partial_s v_{b,0}$ and $\partial_s^2 v_{b,0}$. However, we remark that in the Black–Scholes (Example 2.2) and the CIR (Example 2.3) models, since the $\sigma_{b,0}$'s have closed form expressions, closed expressions for Greeks are available. We refer the reader to Pirjol & Zhu (2018) for more detailed discussions on Greeks of Asian options in the Black–Scholes model.

3. Large Deviation Principle

In this section, we prove Theorem 3.1 which is a large deviation reformulation of Theorem 2.3 for continuously monitored options. Finer tools would be needed to go beyond the leading order of the asymptotic expansion in Theorem 2.2 suggested by the discrete approximation. However, in particular cases, these terms are given by a Girsanov change of measure. We will illustrate this using the Bachelier model

at the end of the section. A more refined expansion for more general processes does not appear tractable with current methods.

Theorem 3.1 (Large Deviation for the Log-Price of Continuous Monitored Asian Call). *The price $C(s_0, 0; K, T)$ at time $t = 0$ of a continuously monitored out-of-the-money Asian call struck at $K > s_0$ with expiry time T for the price process $(S_t, t \geq 0)$ of (2.1) admits the following expansion in T*

$$C(s_0, 0; K, T) = \exp\left\{-\frac{J(K)}{T} + o(T^{-1})\right\} \tag{3.1}$$

where

$$J(K) = \inf\left\{I(f) : f \in \mathcal{C}([0,T]) \text{ and } \int_0^T f(u) = K\right\}. \tag{3.2}$$

Moreover, if $f(t) = x + \int_0^t a(f(s), s)g(s)ds$ for some $g \in L^2([0,T])$ then

$$I(f) = \frac{1}{2}\int_0^T \left(\frac{\dot{f}(t)}{a(f(t), 0)}\right)^2 dt, \tag{3.3}$$

and $I(f)$ is ∞ otherwise.

The idea to prove the theorem is to treat $\varepsilon = T$ as a small parameter and expand around $\varepsilon = 0$. We do a simple time-change $t = uT$ and write for simplicity

$$S^\varepsilon := (S_u^\varepsilon, u \in [0,1]) = (S_{uT}, u \in [0,1]). \tag{3.4}$$

Note that with this notation:

$$\frac{1}{T}\int_0^T S_t dt = \int_0^1 S_u^\varepsilon \, du. \tag{3.5}$$

By the scaling property of Brownian motion, the process $(S_u^\varepsilon, u \in [0,1])$ satisfies the SDE

$$dS_u^\varepsilon = a_\varepsilon(S_u^\varepsilon, u)\sqrt{\varepsilon}dW_u, \tag{3.6}$$

where $a_\varepsilon(x, u) := a(x, u\varepsilon)$. Since $a(x, \cdot)$ is assumed to be continuous uniformly in x, we can write

$$a(x, u\varepsilon) = a(x, 0) + \mathcal{O}(\varepsilon). \tag{3.7}$$

A *large deviation principle* for the process means that there exists a *rate function* $I : \mathcal{C}([0,T]) \to [0, \infty]$ that is lower semi-continuous and has compact level sets such that for any Borel subset A of paths in $\mathcal{C}([0,1])$,

$$-\inf_{s \in \text{int}(A)} I(s) \leq \liminf_{\varepsilon \to 0} \varepsilon \log \mathbb{P}(S^\varepsilon \in A) \leq \limsup_{\varepsilon \to 0} \varepsilon \log \mathbb{P}(S^\varepsilon \in A) \leq -\inf_{s \in \text{cl}(A)} I(s), \tag{3.8}$$

where $\text{int}(A)$ denotes the interior of A and $\text{cl}(A)$ is its closure. Roughly speaking, a large deviation principle quantifies the probability of atypical path at the exponential scale with the help of the rate function. The proof of Theorem 3.1 is based on an extension of the Freidlin–Wentzell theorem, see e.g. Dembo & Zeitouni (1999). The

standard statement of the theorem holds for time-homogeneous drift and volatility under Lipschitz and boundedness assumptions. Here we will use a recent result of Chiarini & Fisher (2014). Note that the locally Lipschitz condition (2.3) is not satisfied in the CIR model. However, in this case, the volatility is time-homogeneous and a weaker assumption is needed for a large deviation to hold. This is the content of Theorem 4 in Chiarini & Fisher (2014).

Theorem 3.2 (Theorem 2, Theorem 4 and Example 1 in Chiarini & Fisher (2014)). *The family of diffusions (3.6) satisfy a large deviation principle with rate function*

$$I(f) = \inf_{\{g \in L^2([0,1]): f(t) = x + \int_0^t a(f(s),0)g(s)ds\}} \frac{1}{2} \int_0^T |g(t)|^2 dt, \tag{3.9}$$

whenever the set $\{g \in L^2([0,1]) : f(t) = x + \int_0^t a(f(s),0)g(s)ds\}$ is nonempty and $I(f)$ is ∞ otherwise.

Formally, it is good to think of the set $\{g \in L^2([0,T]) : f(t) = x + \int_0^t a(f(s),0)g(s)ds\}$ as the set of "white-noise paths" $t \mapsto \dot{W}_t$. The map $g \mapsto f$ where f is the solution of $f(t) = x + \int_0^t \bar{a}(f(s),s)g(s)ds$ can then be thought of as the map sending an underlying Brownian path to the corresponding diffusion path. In the case where the volatility is nonzero for the path f, the map can be inverted and the rate function reduces to the simplest case

$$I(f) = \frac{1}{2} \int_0^1 \left(\frac{\dot{f}(t)}{a(f(t),0)} \right)^2 dt. \tag{3.10}$$

This is certainly the case for geometric Brownian motion and the CIR model. The reader is referred to Theorem 1 in Chiarini & Fisher (2014) and Proposition 3.11 in Baldi & Caramellino (2011) for general sufficient conditions for (3.10) to hold.

Proof. [Proof of Theorem 3.1] Observe that for any random variable X and $K > 0$, we have the identity

$$\mathbb{E}[(X - K)^+] = \int_K^\infty \mathbb{P}(X > x)dx. \tag{3.11}$$

For simplicity, write T for the functional $T : \mathcal{C}([0,1]) \to \mathbb{R}$ with $T(f) = \int_0^1 f(u)du$. Note that T is continuous on $\mathcal{C}([0,1])$ equipped with the topology of uniform convergence. By the contraction principle [see e.g. Dembo & Zeitouni (1999)] and Theorem 3.2, the family of random variables $(T(S^\varepsilon), \varepsilon > 0)$ satisfies a large deviation principle with rate function $J : \mathbb{R} \to [0,\infty]$ with

$$J(y) = \inf\{I(f) : T(f) = y, f \in \mathcal{C}([0,1])\}, \quad y \in \mathbb{R}. \tag{3.12}$$

By (3.11), we have

$$\log \mathbb{E}[(T(S^\varepsilon) - K)^+] = \log \int_K^\infty \mathbb{P}(T(S^\varepsilon) > x)dx. \tag{3.13}$$

On one hand, we have for any $M > K$

$$\log \int_K^\infty \mathbb{P}(T(S^\varepsilon) > x)\,dx = \left(\log \int_K^M \mathbb{P}(T(S^\varepsilon) > x)\,dx\right)$$

$$+ \log\left(1 + \frac{\log \int_M^\infty \mathbb{P}(T(S^\varepsilon) > x)\,dx}{\log \int_K^M \mathbb{P}(T(S^\varepsilon) > x)\,dx}\right). \tag{3.14}$$

For M large enough (possibly dependent on ε), the term in the second parenthesis is smaller than 2. Pick $M = \varepsilon^{-1}$. For ε small enough we thus have

$$\log \int_K^\infty \mathbb{P}(T(S^\varepsilon) > x)\,dx \le \log 2 + \log(\varepsilon^{-1} - K) + \log \mathbb{P}(T(S^\varepsilon) > K). \tag{3.15}$$

This proves that

$$\limsup_{\varepsilon \to 0} \varepsilon \log \int_K^\infty \mathbb{P}(T(S^\varepsilon) > x)\,dx \le \lim_{\varepsilon \to 0} \varepsilon \log \mathbb{P}(T(S^\varepsilon) > K) = -J(K), \tag{3.16}$$

since $(T(S^\varepsilon), \varepsilon > 0)$ satisfies a large deviation with rate function J.

On the other hand, for $\delta > 0$

$$\int_K^\infty \mathbb{P}(T(S^\varepsilon) > x)\,dx \ge \int_K^{K+\delta} \mathbb{P}(T(S^\varepsilon) > x)\,dx \ge \delta \mathbb{P}(T(S^\varepsilon) > K + \delta). \tag{3.17}$$

Therefore, for any $\delta > 0$, the above with the use of (3.8) becomes

$$\liminf_{\varepsilon \to 0} \varepsilon \log \int_K^\infty \mathbb{P}(T(S^\varepsilon) > x)\,dx \ge -J(K + \delta). \tag{3.18}$$

It remains to show that $\lim_{\delta \to 0} J(K+\delta) = J(K)$. First, notice that $\liminf_{\delta \to 0} J(K + \delta) \ge J(K)$ since J is lower semi-continuous (it is a rate function). So it suffices to show that $\limsup_{\delta \to 0} J(K + \delta) \le J(K)$. By definition, $J(K + \delta) = \inf\{I(x) : x \in \mathcal{C}([0,1])$ and $\int_0^1 x(u)du = K + \delta\}$. Pick a sequence $(y_n) \in \mathcal{C}([0,1])$ such that $I(y_n) \to J(K)$ and $\int_0^1 y_n(u)du = K$. By definition of the infimum, this sequence can be picked such that $I(y_n) < J(K) - 1/n$. Pick z a differentiable function on $[0,1]$ such that $\int z = 1$ and $z(0) = 0$. Then $\int(y_n(u) + \delta z(u))du = K + \delta$. Moreover

$$J(K + \delta) < I(y_n + \delta z). \tag{3.19}$$

It is easy to check that for fixed n, $\lim_{\delta \to 0} I(y_n + \delta z) = I(y_n)$. Therefore

$$\limsup_{\delta \to 0} J(K + \delta) < I(y_n) < J(K) - \frac{1}{n}. \tag{3.20}$$

Since n is arbitrary, this proves the claim. □

The asymptotic expansion (2.35) suggests that the lower order corrections to (3.1) are much smaller than the dominant term in T^{-1}. We expect more precisely that

$$\mathbb{E}\left[\left(\frac{1}{T}\int_0^T S_u\,du - K\right)^+\right] = \exp\left\{-\frac{J(K)}{T} + \frac{3}{2}\log T + \mathcal{O}(1)\right\}. \qquad (3.21)$$

This can be verified rigorously for the Bachelier model (Example 1) using a Girsanov change of measure designed to tilt towards the *most-likely path* minimizing the rate function J for a given K. The correction is expected to be of the same order for other models as well as for other types of options.

For the Bachelier model, the diffusion on $[0,1]$ is $dS_u^\varepsilon = \sqrt{\varepsilon}\,dW_u$ and the most-likely path (2.37) when written as a path on $[0,1]$ is

$$s(u) = -\frac{3(K-s_0)}{2}u^2 + 3(K-s_0)u + s_0, \quad u \in [0,1]. \qquad (3.22)$$

By (3.11), the price of the call option becomes

$$\mathbb{E}\left[\left(\int_0^1 S_u^\varepsilon\,du - K\right)^+\right] = \int_K^\infty \mathbb{P}\left(\int_0^1 S_u^\varepsilon\,du > x\right)dx. \qquad (3.23)$$

Write $\mathbb{Q}$ for the measure with $d\mathbb{Q}/d\mathbb{P} = \exp(\varepsilon^{-1/2}\int_0^1 \dot{s}(u)dW_u - \frac{\varepsilon^{-1}}{2}\int_0^1 \dot{s}(u)^2 du)$. Under $\mathbb{Q}$, the process $(W_u, u \in [0,1])$ is a Brownian motion with drift $\varepsilon^{-1/2}\dot{s}(u)$. With this notation, (3.23) becomes

$$e^{-\frac{\varepsilon^{-1}}{2}\int_0^1 \dot{s}(u)^2 du}\int_K^\infty \mathbb{E}_{\mathbb{Q}}[e^{-\varepsilon^{-1/2}\int_0^1 \dot{s}(u)d\widetilde{W}_u}1_{\{\int_0^1 \varepsilon^{1/2}\widetilde{W}_u du+K>x\}}]dx \qquad (3.24)$$

where $(\widetilde{W}_u, u \in [0,1])$ is a standard Brownian motion under $\mathbb{Q}$ and we use the fact that $\int_0^1 s(u)du = K$ by definition of the most likely path. By doing the change of variable $y = x - K$, this reduces to

$$\mathbb{E}\left[\left(\int_0^1 S_u^\varepsilon\,du - K\right)^+\right] = e^{-\frac{\varepsilon^{-1}}{2}\int_0^1 \dot{s}(u)^2 du}$$

$$\times \int_0^\infty \mathbb{E}_{\mathbb{Q}}[e^{-\varepsilon^{-1/2}\int_0^1 \dot{s}(u)d\widetilde{W}_u}1_{\{\int_0^1 \varepsilon^{1/2}\widetilde{W}_u du>x\}}]dx. \qquad (3.25)$$

The first term is $e^{-\varepsilon^{-1}J(K)}$ and gives the first order. To evaluate the second term, it is convenient to first integrate x to get

$$\mathbb{E}_{\mathbb{Q}}\left[e^{-\varepsilon^{-1/2}\int_0^1 \dot{s}(u)d\widetilde{W}_u}\left(\int_0^1 \varepsilon^{1/2}\widetilde{W}_u\,du\right)^+\right]. \qquad (3.26)$$

Note that $\dot{s}(u) = 3(K-s_0)(1-u)$, therefore

$$\int_0^1 \dot{s}(u)d\widetilde{W}_u = 3(K-s_0)\int_0^1 \widetilde{W}_u\,du. \qquad (3.27)$$

Write X for the random variable $\int_0^1 \widetilde{W}_u du$ which is Gaussian with mean 0 and variance $1/3$. We have

$$\varepsilon^{1/2} \mathbb{E}_{\mathbb{Q}} \left[X^+ e^{-\varepsilon^{-1/2} 3(K-s_0)X} \right] = \varepsilon^{3/2} \int_0^\infty y e^{-3(K-s_0)y} \frac{e^{-\frac{3\varepsilon y^2}{2}}}{\sqrt{\frac{2\pi}{3}}} dy. \tag{3.28}$$

The integral is of order 1. We conclude that

$$\mathbb{E} \left[\left(\frac{1}{T} \int_0^T S_u \, du - K \right)^+ \right] = \exp \left(-T^{-1} J(K) + \frac{3}{2} \log T + \mathcal{O}(1) \right). \tag{3.29}$$

4. Numerical Tests

From Theorem 2.5, we have the following approximate formula for Asian Bachelier implied volatility:

$$\sigma_{b,0} = \left(\frac{T}{3(K-s_0)^2} \int_0^T \left[\frac{\dot{\tilde{s}}(t)}{a(\tilde{s}(t),t)} \right]^2 dt \right)^{-\frac{1}{2}}, \tag{4.1}$$

where $\tilde{s}(t)$ is the most-likely-path for an Asian option determined by solving the variational problem

$$\min_{s:t \mapsto s(t)} \frac{1}{2} \int_0^T \left[\frac{\dot{s}(t)}{a(s(t),t)} \right]^2 dt \tag{4.2}$$

subject to

$$\frac{1}{T} \int_0^T s(t) dt = K, \quad s(0) = s_0. \tag{4.3}$$

We now proceed to test numerically the approximate implied volatility formula (4.1) for various definitions of the local volatility function $a(s,t) = s\, \sigma_\ell(s,t)$ in the local volatility model (2.1). In each case, to evaluate (4.1), we need to compute the most-likely-path $\tilde{s}(t)$. To this end, we exploit the following iteration scheme.

Lemma 4.1. *The most-likely path $\tilde{s}(t)$ satisfies the recursive formula*

$$\tilde{s}(t) = s_0 + \frac{I(t)}{\bar{I}(T)}[K - s_0], \tag{4.4}$$

where

$$I(t) = \int_0^t \int_r^T a(\tilde{s}(r),r) a(\tilde{s}(u),u) e^{-\int_t^u \frac{a_t(\tilde{s}(v),v)}{a(\tilde{s}(v),v)} dv} \, du dr,$$

$$\bar{I}(T) = \frac{1}{T} \int_0^T I(u) \, du.$$

Proof. From Lemma 2.2, $\tilde{s}(t)$ satisfies the Euler–Lagrange equation (2.25) which we reiterate in the following for convenience

$$\frac{d}{dt}\left(\frac{\dot{s}}{a}\right) - \frac{a_t}{a^2}\dot{s} + \frac{\lambda}{T}a = 0 \tag{4.5}$$

with boundary conditions

$$s(0) = s_0, \quad \dot{s}(T) = 0, \tag{4.6}$$

where λ is chosen such that $1/T \int_0^T s(t)dt = K$.

For ease of notation, define $\tilde{a}(t) := a(\tilde{s}(t), t)$ and $f(s,t) = a_t/a = \partial_t \log a$. Also, $\tilde{f}(t) = f(\tilde{s}(t), t)$. Applying the integrating factor $\exp(-\int_0^t \tilde{f}(u)du)$ and integrating (4.5) with the boundary condition $\dot{s}(T) = 0$ gives

$$-e^{-\int_0^t \tilde{f}(v)dv}\frac{\dot{s}(t)}{\tilde{a}(t)} = -\frac{\lambda}{T}\int_t^T \tilde{a}(u)e^{-\int_0^u \tilde{f}(v)dv}\,du. \tag{4.7}$$

It follows that

$$\frac{\dot{s}(t)}{\tilde{a}(t)} = \frac{\lambda}{T}\int_t^T \tilde{a}(u)e^{-\int_t^u \tilde{f}(v)dv}\,du. \tag{4.8}$$

Rearranging and integrating again gives

$$s(t) - s_0 = \frac{\lambda}{T}I(t), \tag{4.9}$$

where $I(t) = \int_0^t \int_r^T \tilde{a}(r)\tilde{a}(u)\exp(-\int_t^u \tilde{f}(v)dv)\,du\,dr$. Now apply the boundary condition $1/T\int_0^T s(t)dt = K$ to get

$$K - s_0 = \frac{\lambda}{T}\bar{I}(T) \tag{4.10}$$

and the result follows. $\qquad\square$

Lemma 4.1 leads to an efficient fixed-point algorithm for solving for the most-likely-path. The natural choice of first guess is the Bachelier most-likely-path (2.37):

$$s(t) = s_0 + 3(K - s_0)\frac{t}{T} - \frac{3(K - s_0)}{2}\left(\frac{t}{T}\right)^2. \tag{4.11}$$

The resulting algorithm typically converges sufficiently after three or four iterations.

4.1. *A time-dependent CIR model*

We consider the model

$$dS_t = e^{-\lambda t}\sigma\sqrt{S_t}\,dW_t \tag{4.12}$$

with $S_0 = 1$, $\sigma = 0.2$ and $\lambda = 1$. Thus

$$a(s,t) = e^{-\lambda t}\sigma\sqrt{s}. \tag{4.13}$$

Though we computed a quasi-closed-form for the most-likely-path for the time-homogeneous case in Example 2.3, in this time-inhomogeneous case, we choose to
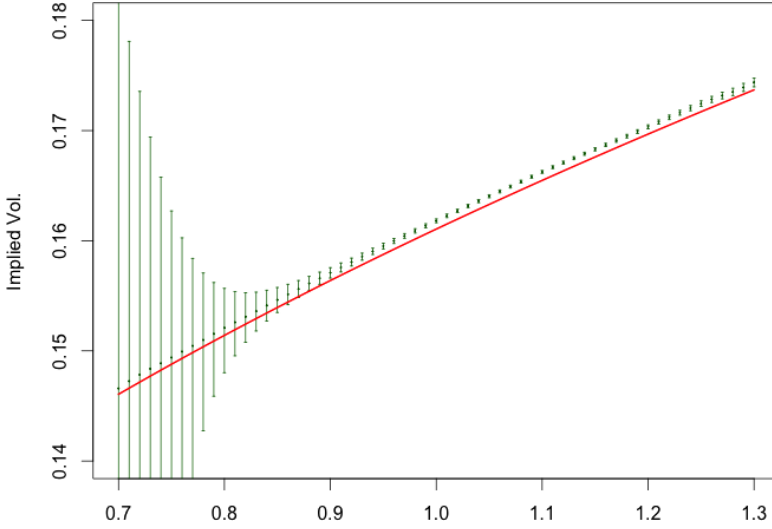
Fig. 1. The 1-year Bachelier Asian implied volatility smile corresponding to the time-dependent CIR local volatility function (4.13). The blue dotted line is from Monte Carlo simulation with error bars; the red solid line is the approximation $\sigma_{b,0}$.

compare our approximate implied volatility formula (4.1) evaluated using the fixed-point iteration algorithm against Monte Carlo simulations generated with 1000 time steps and 2 million sample paths. The results are shown in Fig. 1. We remark that the most-likely-path approximation in this example slightly underestimates the normal implied volatility for Asian option inferred from simulation.

4.2. *Time-dependent quadratic local volatility*

Next we consider the following quadratic local volatility model:

$$dS_t = e^{-\lambda t}\sigma \left[1 + \psi(S_t - 1) + \frac{\gamma}{2}(S_t - 1)^2\right] dW_t \qquad (4.14)$$

with $\sigma = 0.2$, $\psi = -0.5$, $\gamma = 0.1$, and $\lambda = 1$. We remark that though in this example the function a grows quadratically to infinity as $|x| \to \infty$ which violates the linear growth condition (2.2) required for the theoretical argument, we did the numerical experiment for testing the applicability of the most-likely-path methodology.

Though a closed-form solution for European options with these parameters is given in Andersen (2001), we again resort to Monte Carlo simulation to estimate the value of Asian options in this model. Likewise, simulations are generated with 1000 time steps and 2 million sample paths. The results are shown in Fig. 2. Similarly, the most-likely-path approximation in this example also underestimates the normal implied volatility for Asian option inferred from simulation.
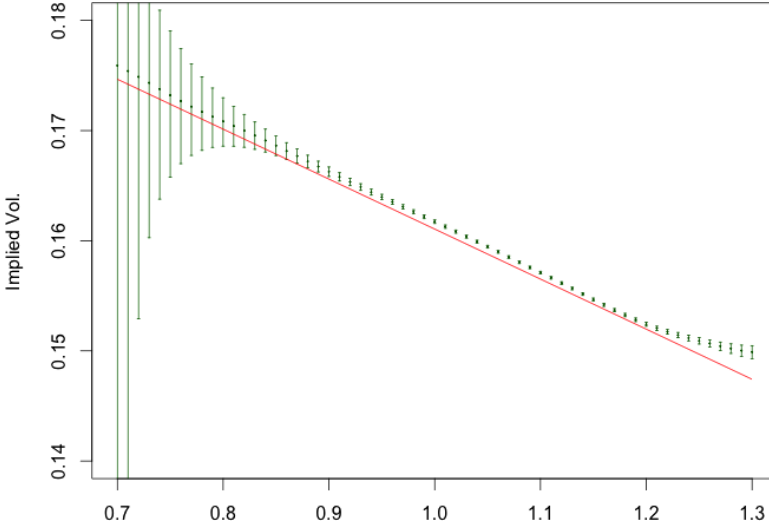
Fig. 2. The 1-year Bachelier Asian implied volatility smile corresponding to the time-dependent quadratic local volatility function (4.14). The blue dotted line is from Monte Carlo simulation with error bars; the red solid line is the approximation $\sigma_{b,0}$.

4.3. *Benchmark scenarios in Black–Scholes and CIR models*

In this subsection, for the Black–Scholes and CIR models, we compare the most-likely-path approximation to a few existing approximations with the benchmark scenarios proposed in Geman & Eydeland (1995) and Fu *et al.* (1999) that were commonly used in the literatures on Asian option pricing, see for instance Dassios & Nagaradjasarma (2006), Dufresne (2000), Foschi *et al.* (2013), Linetsky (2004), Večeř & Xu (2002).

With the approximation of Bachelier implied volatility $\sigma_{b,0}$ given in Theorem 2.5, we may calculate the approximate price of the Asian call struck at K and expired at T by

$$e^{-rT}\left\{\sqrt{\frac{v}{2\pi}}e^{-\frac{(A-K)^2}{2v}} + (A-K)N\left(\frac{A-K}{\sqrt{v}}\right)\right\} \tag{4.15}$$

with

$$A = \frac{S_0(e^{\mu T}-1)}{\mu T}, \quad v = \frac{\sigma_{b,0}^2}{\mu^2 T^2}\left(\frac{3 - 4e^{\mu T} + e^{2\mu T}}{2\mu} + T\right) \tag{4.16}$$

and $\mu = r - q$. Note that (4.15) is indeed the price of the Asian call struck at K and expired at T under the Bachelier model

$$dS_t = (r-q)S_t dt + \sigma dW_t \tag{4.17}$$

with risk free rate r and dividend rate q, both assumed constant. We remark that, since many of the benchmark scenarios are of ATM, the ATM approximate implied

volatility is obtained by taking the limit of $\sigma_{b,0}$ as K approaches S_0. Explicitly, the limits in the Black–Scholes and CIR cases are given by

$$\lim_{K \to S_0} \sigma_{b,0} = \begin{cases} \sigma S_0 & \text{for the Black–Scholes model,} \\ \sigma \sqrt{S_0} & \text{for the CIR model.} \end{cases} \tag{4.18}$$

Table 1 exhibits the numerical results for the asymptotic approximation for the Asian options obtained from (4.15) for the scenarios considered in Fu *et al.* (1999). Based on the results in Linetsky (2004), the relative discrepancies of the approximate prices are less than 1.5% in all the seven benchmark scenarios and within 1% for options expiring in a year. Table 2 shows the numerical tests for Asian option pricing in the CIR model for the seven scenarios proposed in Dassios & Nagaradjasarma (2006). Data is quoted from Table 5 in Foschi *et al.* (2013). Based on the results in Foschi *et al.* (2013), the relative discrepancies of the approximate prices are less than 1% in all the seven benchmark scenarios and within 0.6% for options expiring in a year. We may thus arguably conclude that the short expiry approximation results of the current paper provide a reasonable approximation for Asian option prices in these models.

Table 1. Prices of Asian calls in the Black–Scholes model in the benchmark scenarios. The last four columns correspond to the approximate price from (4.15) (ALW), the third order approximation from Foschi *et al.* (2013) (FPP3), the precise evaluation in Linetsky (2004) (Linetsky), and the relative discrepancy of ALW to Linetsky.

Case	S_0	K	r	σ	T	ALW	FPP3	Linetsky	rel. discrp.
1	2	2	0.02	0.1	1	0.056042	0.055986	0.055986	0.10%
2	2	2	0.18	0.3	1	0.219607	0.218387	0.218387	0.56%
3	2	2	0.0125	0.25	2	0.172939	0.172267	0.172269	0.39%
4	1.9	2	0.05	0.5	1	0.195034	0.193164	0.193174	0.96%
5	2	2	0.05	0.5	1	0.248277	0.246406	0.246416	0.73%
6	2.1	2	0.05	0.5	1	0.308029	0.306210	0.306220	0.59%
7	2	2	0.05	0.5	2	0.355167	0.350040	0.350095	1.45%

Table 2. Prices of Asian calls in the CIR model in the benchmark scenarios considered in Dassios & Nagaradjasarma (2006). The last four columns correspond to the approximate price from (4.15) (ALW), from Dassios & Nagaradjasarma (2006) (DN), the third order approximation from Foschi *et al.* (2013) (FPP3), and the relative discrepancy of ALW to FPP3.

Case	S_0	K	r	σ	T	ALW	DN	FPP3	rel. discrp.
1	2	2	0.02	0.14	1	0.055591	0.0197	0.055562	0.05%
2	2	2	0.18	0.42	1	0.218521	0.2189	0.217874	0.30%
3	2	2	0.0125	0.35	2	0.171331	0.1725	0.170926	0.24%
4	1.9	2	0.05	0.69	1	0.191950	0.1902	0.190834	0.58%
5	2	2	0.05	0.72	1	0.252333	NA	0.251121	0.48%
6	2.1	2	0.05	0.72	1	0.309864	0.3098	0.308715	0.37%
7	2	2	0.05	0.71	2	0.356411	0.3339	0.353197	0.91%

5. Conclusion

We have derived a small time asymptotic of the price of discretely monitored Asian options up to first order. The most-likely-path approximation for continuously monitored Asian calls has been derived heuristically by taking the limit of the corresponding term from discretely monitored case and proved rigorously using the theory of large deviation. Numerical experiments in both time-dependent CIR and time-dependent quadratic local volatility models showed that there is room to improve the performance of the most-likely-path approximation. One possibility is to include higher order terms. Large deviation theory generically provides no insights beyond the term of exponential decay. On the other hand, it is conceivable that taking the limit of corresponding terms from the discretely monitored case may yield tractable expressions for numerical evaluations. The calculations of higher order terms are considerably more involved and henceforth were left to further work.

Appendix. Convexity of the Constrained Optimization Problem

We analyze the convexity of the constrained optimization problem (2.15):(2.16) in this Appendix. Recall the Lamperti transformation $\varphi(s,t) := \int_{s_0}^{s} 1/a(\xi,t)d\xi$ and the objective function D

$$D(\mathbf{s}, \mathbf{t}) = \frac{1}{2} \sum_{i=1}^{n} |\varphi(s_i, t_i) - \varphi(s_{i-1}, t_{i-1})|^2. \tag{A.1}$$

For notational simplicity, we shall write the function $a(\cdot, t_i)$ as $a_i(\cdot)$ and similarly $\varphi(\cdot, t_i)$ as $\varphi_i(\cdot)$. By straightforward calculations, the second partial derivatives of D are given by

$$\frac{\partial^2 D}{\partial s_i^2} = \frac{2}{a_i^2(s_i)} + \frac{a_i'(s_i)}{a_i^2(s_i)} [\varphi_{i-1}(s_{i-1}) + \varphi_{i+1}(s_{i+1}) - 2\varphi_i(s_i)]; \tag{A.2}$$

$$\frac{\partial^2 D}{\partial s_n^2} = \frac{1}{a_n^2(s_n)} - \frac{a_n'(s_n)}{a_n^2(s_n)} [\varphi_{n-1}(s_n) - \varphi_{n-1}(s_{n-1})]; \tag{A.3}$$

$$\frac{\partial^2 D}{\partial s_i \partial s_j} = \frac{-1}{a_i(s_i) a_j(s_j)}, \quad \text{if } |i - j| = 1; \tag{A.4}$$

$$\frac{\partial^2 D}{\partial s_i \partial s_j} = 0, \quad \text{if } |i - j| \geq 2. \tag{A.5}$$

We decompose the Hessian matrix $H = [\partial_{s_i} \partial_{s_j} D]$ as $H = H^1 + H^2$, where H^1 is the symmetric tridiagonal matrix with diagonal entries given by

$$H^1_{kk} = \frac{2}{a_k^2(s_k)} \quad \text{for } k = 1, \ldots, n - 1, \; H^1_{nn} = \frac{1}{a_n^2(s_n)} \tag{A.6}$$

and off-diagonal entries by

$$H^1_{k,k+1} = \frac{-1}{a_k(s_k) a_{k+1}(s_{k+1})} \quad \text{for } k = 1, \ldots, n - 1. \tag{A.7}$$

H^2 is a diagonal matrix with diagonal entries given by

$$H^2_{kk} = \frac{a_k'(s_k)}{a_k^2(s_k)} [\varphi_{k-1}(s_{k-1}) + \varphi_{k+1}(s_{k+1}) - 2\varphi_k(s_k)] \quad \text{for } k = 1, \ldots, n - 1, \tag{A.8}$$

$$H^2_{nn} = -\frac{a_n'(s_n)}{a_n^2(s_n)} [\varphi_{n-1}(s_n) - \varphi_{n-1}(s_{n-1})]. \tag{A.9}$$

We claim that H^1 is positively definite. Let $\hat{H}^1$ be the square submatrix of H^1 consisting of the first $n - 1$ rows and columns of H^1. We partition H^1 as

$$H^1 = \begin{bmatrix} \hat{H}^1 & c^T \\ c & H^1_{nn} \end{bmatrix}, \tag{A.10}$$

where $c = [0 \; \cdots \; 0 \; H^1_{n-1,n}]$ is an $(n - 1)$ row vector. Note that, by induction, one can show that the principle minors of $\hat{H}^1$ are $(k + 1) \prod_{j=1}^{k} a_j^{-2}(s_j) > 0$, for $k = 1, \ldots, n - 1$. By applying the identity for determinant, if A is invertible,

$$\det \begin{bmatrix} A & b^T \\ b & d \end{bmatrix} = \det(A) \times (d - bA^{-1}b^T), \tag{A.11}$$

we calculate the determinant of H^1 as

$$\det(H^1) = \det(\hat{H}^1) \times \{H^1_{nn} - c(\hat{H}^1)^{-1}c^T\} \tag{A.12}$$

$$= \det(\hat{H}^1) H^1_{nn} - (H^1_{n-1,n})^2 \det(\hat{\hat{H}}^1) \tag{A.13}$$

$$= \prod_{k=1}^{n} \frac{1}{a_k^2(s_k)} > 0, \tag{A.14}$$

where $\hat{\hat{H}}^1$ is the square submatrix of H^1 by deleting the last two rows and columns of H^1. Thus, by Sylvester's criterion H^1 is positive definite.

Finally, since

$$\varphi_{k-1}(s_{k-1}) + \varphi_{k+1}(s_{k+1}) - 2\varphi_k(s_k)$$
$$= \varphi'_k(s_k)(\Delta s_{k+1} + \Delta s_k) + o(\Delta s_{k+1}, \Delta s_k, \Delta t) \qquad (A.15)$$

$$\varphi_{n-1}(s_n) - \varphi_{n-1}(s_{n-1}) = \varphi'_{n-1}(s_{n-1})\Delta s_n + o(\Delta s_n), \qquad (A.16)$$

where $\Delta s_k = s_k - s_{k-1}$, the entries of H^2 are small if the s_i's are close to each other and Δt is small. In this case, H is positive definite. Hence, the objective function D is convex.

References

H. Albrecher, P. A. Mayer & W. Schoutens (2008) General lower bounds for arithmetic Asian option prices, *Applied Mathematical Finance* **15** (2), 123–149.

L. Andersen (2011) Option pricing with quadratic volatility: A revisit, *Finance and Stochastics* **15** (2), 191–219.

P. Baldi & L. Caramellino (2011) General Freidlin–Wentzell large deviations and positive diffusions, *Statistics and Probability Letters* **81** (8), 1218–1229.

N. Bleinstein & R. A. Handelsman (1986) *Asymptotic Expansions of Integrals*. Dover Publications.

P. P. Boyle & D. Emanuel (1980) Discretely adjusted option hedges, *Journal of Financial Economics* **8** (3), 259–282.

M. Broadie & P. Glasserman (1996) Estimating security price derivatives using simulation, *Management Science* **42**, 269–285.

N. Cai & S. G. Kou (2012) Pricing Asian options under a hyper-exponential jump diffusion model, *Operations Research* **60** (1), 64–77.

N. Cai, C. Li & C. Shi (2014) Closed-form expansions of discretely monitored Asian options in diffusion models, *Mathematics of Operations Research* **39** (3), 789–822.

A. Chiarini & M. Fisher (2014) On large deviations for small noise Itô processes, *Advances in Applied Probability* **46** (4), 1126–1147.

G. Cibelli, E. Polidoro & F. Rossi (2016) Sharp estimates for Geman–Yor processes and applications to arithmetic average Asian options, Preprint available in arXiv, 2016.

M. Curran (1994) Valuing Asian and portfolio options by conditioning on the geometric mean, *Management Science* **40** (21), 1705–1711.

A. Dassios & J. Nagaradjasarma (2006) The square-root process and Asian options, *Quantitative Finance* **6**, 337–347.

J. N. Dewynne & W. T. Shaw (2008) Differential equations and asymptotic solutions for arithmetic Asian options: 'Black–Scholes formulae' for Asian rate calls, *European Journal of Applied Mathematics* **19**, 353–391.

A. Dembo & O. Zeitouni (1998) *Large Deviations Techniques and Applications*, 2nd edn. Springer-Verlag.

D. Dufresne (2000) Laguerre series for Asian and other options, *Mathematical Finance* **10** (4), 407–428.

J.-P. Fouque & C. H. Han (2003) Pricing Asian options with stochastic volatility, *Quantitative Finance* **3** (5), 352–362.

M. C. Fu, D. Madan & T. Wang (1999) Pricing continuous Asian options: A comparison of Monte Carlo and Laplace transform inversion methods, *Journal of Computational Finance* **2** (2), 49–74.

P. Foschi, S. Pagliarani & A. Pascucci (2013) Approximations for Asian options in local volatility models, *Journal of Computational and Applied Mathematics* **237** (1), 442–459.

J. Gatheral, E. P. Hsu, P. Laurence, C. Ouyang & T.-H. Wang (2012) Asymptotics of implied volatility in local volatility models, *Mathematical Finance* **22** (4), 591–620.

J. Gatheral & T.-H. Wang (2012) The heat kernel most-likely-path approximation, *International Journal of Theoretical and Applied Finance* **15** (1), 1250001.

H. Geman & A. Eydeland (1995) Domino effect, *Risk Magazine* **8**, 65–67.

H. Geman & M. Yor (1993) Bessel processes, Asian options, and perpetuities, *Mathematical Finance* **3**, 349–375.

P. Glasserman, P. Heidelberger & P. Shahabuddin (1999) Asymptotically optimal importance sampling and stratification for pricing path-dependent options, *Mathematical Finance* **9** (2), 117–152.

C. Grunspan (2011) A note on the equivalence between the normal and lognormal implied volatility: A model free approach, Preprint, available on arXiv.

A. G. Z. Kemna & A. C. F. Vorst (1990) A pricing method for options based on average asset values, *Journal of Banking and Finance* **14**, 113–129.

N. Kunitomo & A. Takahashi (1992) Pricing average options, *Japan Financial Review* **14**, 1–20.

V. Linetsky (2004) Spectral expansions for Asian (average price) options, *Operations Research* **52** (6), 856–867.

S. Pagliarani, A. Pascucci & M. Pignotti (2017) Intrinsic expansions for averaged diffusion processes, *Stochastic Processes and Their Applications* **127** (8), 2560–2585.

D. Pirjol & L. Zhu (2016) Short maturity Asian options in local volatility models, *SIAM Journal of Financial Mathematics* **7** (1), 947–992.

D. Pirjol & L. Zhu (2018) Sensitivities of Asian options in the Black–Scholes model, *International Journal of Theoretical and Applied Finance* **21** (1), 1850008.

P. Ritchken, L. Sankarasubramanian & A. M. Vijh (1993) The valuation of path dependent contracts on the average, *Management Science* **39**, 1202–1213.

L. C. G. Rogers & Z. Shi (1995) The value of an Asian option, *Journal of Applied Probability* **32**, 1077–1088.

W. Schachermayer & J. Teichmann (2007) How close are the option pricing formulas of Bachelier and Black–Merton–Scholes? *Mathematical Finance* **18**, 155–170.

K. Shiraya & A. Takahashi (2012) Pricing average options on commodities, CIRJE Discussion Papers.

S. M. Turnbull & L. M. Wakeman (1991) A quick algorithm for pricing European average options, *The Journal of Financial and Quantitative Analysis* **26** (3), 377–389.

F. Vazquez-Abad & D. Dufresne (1998) Accelerated simulation for pricing Asian options, *Proceedings of the 1998 Winter Simulation Conference*, Piscataway, New Jersey. Institute of Electrical and Electronics Engineers, pp. 1493–1500.

J. Večeř (2001) A new PDE approach for pricing arithmetic average Asian options, *Journal of Computational Finance* **4** (4), 105–113.

J. Večeř (2014) Black–Scholes representation for Asian options, *Mathematical Finance* **24** (3), 598–626.

J. Večeř & M. Xu (2002) Unified Asian pricing, *Risk Magazine* **15**, 113–116.

T.-H. Wang & J. Gatheral (2015) Implied volatility from local volatility: A path integral approach, *Springer Proceedings in Mathematics & Statistics* **110**, Large Deviations and Asymptotic Methods in Finance.