Most-likely-path in Asian option pricing under local volatility models

Louis-Pierre Arguin
Department of Mathematics, Baruch College and Graduate Center, CUNY
1 Bernard Baruch Way, New York, NY10010
e-mail: louis-pierre.arguin@baruch.cuny.edu

Nien-Lin Liu
Department of Mathematical Sciences, Ritsumeikan University
Noji-higashi 1-1-1, Kusatsu, Shiga 525-8577, Japan
e-mail: nienlin1126@gmail.com

Tai-Ho Wang
Department of Mathematics, Baruch College, CUNY
1 Bernard Baruch Way, New York, NY10010
e-mail: tai-ho.wang@baruch.cuny.edu

Current Version: June 9, 2017

Abstract

This article addresses the problem of approximating the price of options on discrete and continuous arithmetic average 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). Approximation for the implied normal volatility follows accordingly. The small-time asymptotics and the existence of the MLP are also recovered rigorously using large deviation theory.

Keywords: Asian option pricing, Asymptotic expansion, Exotic option, Large deviation theory, Most likely path
MOST-LIKELY-PATH IN ASIAN OPTION PRICING UNDER LOCAL VOLATILITY MODELS

LOUIS-PIERRE ARGUIN, NIEN-LIN LIU, AND TAI-HO WANG

ABSTRACT. This article addresses the problem of approximating the price of options on discrete and continuous arithmetic average 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). Approximation for the implied normal volatility follows accordingly. The small-time asymptotics and the existence of the MLP are also recovered rigorously using large deviation theory.

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 markets. Nowadays average price options represent a high percentage of options on oil; some are directly on oil prices, while others on spreads between two types of oil. 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 and Vorst [19] 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 and Vorst [19] 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 [9] 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 and Glasserman [6] (for a direct method of estimating the Greeks of an Asian option by Monte Carlo simulation, see Section 4.2 on P.275), Vazquez-Abad and Dufresne [26] (combining control variate and change of measure/likelihood ratio), and Glasserman, Heidelberger, and Shahabuddin [17] (importance sampling), etc. Attempts to closed or semi-closed form expressions for Asian option pricing in Black-Scholes models first appeared in the seminal work of Geman and Yor [16]. Among other interesting results in Geman and Yor [16], 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. However, numerical inversion of the Geman-Yor formula was shown slow and needed to handle with care, see for example the discussions in Dufresne [11] and Fu, Madan, and Wang [13]. 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 and Wakeman [25] and Ritchken, Sankarasubramanian, and Vijh [22]. 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 and Shi [23] and Verčer [27] (see also Verčer [28] 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, Mayer, and Schoutens [1] and Rogers and Shi [23], and also approximative and asymptotic solutions. Approximations resorting to asymptotic expansions are mostly Itô-Taylor type expansion based as originated from the work of Kunitomo and Takahashi [20], see also Cai, Li, and Shi [7] 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 accuracies. Finally, though not directly related to the current paper, Asian option pricing under stochastic volatility models is discussed in Fouque and Han [12] and Shiraya and Takahashi [24].
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 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 and Gatheral [29] between consecutive sampling time points leads to a “path-integral” type expression for the Asian option price, see (2.7). A direct application of a Laplace asymptotic formula (in this case high-dimensional and nonstandard, 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 Section 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 by Pirjol and Zhu in [21]. 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.
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 prices for 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 and Gatheral [29] (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

$$dS_t = S_t \sigma_t(S_t, t)dW_t = a(S_t, t)dW_t \quad S_0 = s_0. \quad (2.1)$$

We will suppose that the diffusion function $a(x, t)$ grows at most linearly in $x$: there exists $C > 0$ such that

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

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 $|x|, |y| < R$

$$|a(x, t) - a(y, t)| \leq C_R|x - y|. \quad (2.3)$$

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 by Chiarini and Fischer [8] as explained in Section 3. For simplicity, we will also assume that the diffusion function $a(S_t, t)$ is non-zero for almost all paths of the diffusion. This assumption is not technically necessary but simplifies the expression of the large deviation principle.

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 $x$

$$x = \varphi(s, t) = \int_{s_0}^{s} \frac{d\xi}{a(\xi, t)}. \quad (2.4)$$
The following representation for the transition density \( p \) is derived in Wang and Gatheral [29].

**Theorem 2.1.** Let \( S = (S_t, t \geq 0) \) be the diffusion process given by \( \Phi_1 \). Define the function \( h \) by \( h(x,t) = \varphi_t(s,t) - \frac{a(s,t)}{2} \), with \( s = \varphi^{-1}(x,t) \), where \( \varphi \) is the Lamperti transformation \( \Phi_4 \) 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:

\[
p(T, s_T | t, s_t) = \frac{g(T - t, \varphi(s_T, T) - \varphi(s_t, t))}{a(s_T, T)} \mathbb{E}_{\varphi(s_t, t), \varphi(s_T, T)} \left[ e^{\int_t^T h(X_s, s) dX_s - \frac{1}{2} \int_t^T h^2(X_s, s) ds} \right]
\]

(2.5)

where \( g \) denotes the centered Gaussian density with variance \( t \): \( g(t, \xi) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{\xi^2}{2t}} \) and \( \mathbb{E}_{x,y}[\cdot] \) 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, \( \frac{\partial}{\partial x} H(x,t) = h(x,t) \), for all \( x \) and \( t \), then

\[
p(T, s_T | t, s_t) = \frac{g(T - t, \varphi(s_T, T) - \varphi(s_t, t))}{a(s_T, T)} e^{H(\varphi(s_T, T), T) - H(\varphi(s_t, t), t)} \times \\
\mathbb{E}_{\varphi(s_t, t), \varphi(s_T, T)} \left[ e^{-\frac{1}{2} \int_t^T h^2(X_s, s) + h_s(X_s, s) + 2H_t(X_s, s) ds} \right].
\]

(2.6)

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

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

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

\[
\Psi(s_T, s_T) = e^{H(\varphi(s_T, T), T) - H(\varphi(s_t, t), t)} \times \\
\mathbb{E}_{\varphi(s_t, t), \varphi(s_T, T)} \left[ e^{-\frac{1}{2} \int_t^T h^2(X_s, s) + h_s(X_s, s) + 2H_t(X_s, s) ds} \right].
\]

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

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

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

\[
p(t_1, s_{t_1} | t_0, s_{t_0}) \cdots p(t_n, s_{t_n} | t_{n-1}, s_{t_{n-1}})
\]

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

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

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

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

where

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

$t = (t_1, \cdots, t_n), s = (s_{t_1}, \cdots, s_{t_n})$ and $ds = ds_{t_1} \cdots ds_{t_n}$.

Hence, the evaluation of the Asian call price $C$ becomes the computation of the multidimensional integral (2.7). The following Laplace asymptotic formula will be applied to estimate the $n$-dimensional integral in (2.7) for small $\Delta t$.

**Lemma 2.1.** *(Laplace asymptotic formula)*

Let $R$ be a closed set in $\mathbb{R}^n$ with nonempty and smooth boundary $\partial R$. Suppose $\theta$ 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 $x \in R \setminus B_\epsilon(x^*)$, where $B_\epsilon(x^*) = \{x : |x - x^*| < \epsilon\}$ is the open ball of radius $\epsilon$ 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 $\partial R$ but the inward normal directional derivative of $f$ at $x^*$ is nonzero. Then we have the asymptotic expansion as $\tau \to 0^+$

$$\int_{R} e^{-\frac{\theta(x)}{\tau}} f(x) dx = \frac{(2\pi)^{\frac{n}{2}} \tau^{\frac{n-1}{2}} e^{-\frac{\theta(x^*)}{\tau}}}{\sqrt{\det \partial^2 \theta(x^*) \nabla \theta(x^*)}} \left[ \frac{\nabla f(x^*) \cdot \nabla \theta(x^*)}{|\nabla \theta(x^*)|^2} + \frac{1}{2} \text{tr} \left\{ \partial^2 f(x^*) \left[ \partial^2 \theta(x^*) \right]^{-1} \right\} + \mathcal{O}(\tau) \right],$$

where $\partial^2 f(x^*)$ and $\partial^2 \theta(x^*)$ are the Hessian matrices of $f$ and $\theta$ respectively in the tangential direction to $R$ at $x^*$.

The proof of the lemma is standard and straightforward. See for instance Section 8.3 in Bleistein and Handelsman [1]. For reader’s convenience, we provide a proof in Appendix II.
Now we apply the Laplace asymptotic formula (2.8) to the multidimensional integral (2.7) by taking
\[ \theta = D(s, t), \quad f = \left( \frac{1}{n} \sum_{i=1}^{n} s_{t_i} - K \right) W(s, t), \]
and \( R = \{ s : \frac{1}{n} \sum_{i=1}^{n} s_{t_i} \geq K \} \). The crucial step in applying the Laplace asymptotic formula (2.8) is the determination of the minimum point of \( D \) in the half space \( R \), which boils down to solving the constrained optimization problem:
\[
\min_{s} \frac{1}{2} \sum_{i=1}^{n} \left| \varphi(s_{t_i}, t_i) - \varphi(s_{t_{i-1}}, t_{i-1}) \right|^2 \tag{2.9}
\]
subject to
\[
\frac{1}{n} \sum_{i=1}^{n} s_{t_i} \geq K. \tag{2.10}
\]

**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 (2.9);(2.10) 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 (2.9) attains its global minimum 0.

**Remark 2.2.** We show in Section 6 Appendix I that, for \( \Delta t \) small enough, the objective function in (2.9) 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 (2.8) since the functions \( D \) and \( W \) are independent of \( t \). We remark that, modulo the exponential term, the result suggests the next order term is of order \( \frac{3}{2} \) in \( \Delta t \), which coincides with the order in the case for European options, see for example Theorem 2.3 in Gatheral et al [14].

**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 \( n \),
\[
C = \mathbb{E} \left[ \left( \frac{1}{n} \sum_{i=1}^{n} S_{t_i} - K \right)^+ \right] = \int \int_{\{ s : \frac{1}{n} \sum_{i=1}^{n} s_{t_i} \geq K \}} \left( \frac{1}{n} \sum_{i=1}^{n} s_{t_i} - K \right) e^{-\frac{D(s,t)}{\Delta t}} W(s, t) ds
\]
where \( s^* = (s^*_1, \ldots, s^*_n) \) is the minimizer of the minimization problem (2.9) subject to the constraint \( \frac{1}{n} \sum_{i=1}^{n} s_i \geq K \).

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 \( \Delta t \) 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 Section 3.

Let \( \{0 = t_0 < t_1 < \cdots < t_n = T\} \) be a partition of the interval \([0, T]\) with \( t_i - t_{i-1} = \Delta t := \frac{T}{n} \), 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,

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

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 (2.11) as \( \Delta t \to 0 \). To be specific, rewrite the logarithm of (2.11) as

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

\[
+ \log \left( \frac{\nabla W(s^*, t) \cdot \nabla D(s^*, t)}{\sqrt{\det \partial_t^2 D(s^*, t)|\nabla D(s^*, t)|^2}} \right) + \frac{1}{2} \text{tr} \left\{ \partial_t^2 W(s^*, t) \left[ \partial_t^2 D(s^*, t) \right]^{-1} \right\} + O(\Delta t),
\]

where recall that \( s^* \) is the minimizer of (2.9). Notice 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 $\Delta t$ back to the objective function (2.9) as
\[
\frac{1}{2\Delta t} \sum_{k=1}^{n} \left| \varphi(s_{t_k}, t_k) - \varphi(s_{t_{k-1}}, t_{k-1}) \right|^2.
\]
Note that, since
\[
\varphi(s_{t_k}, t_k) - \varphi(s_{t_{k-1}}, t_{k-1}) = \varphi_s(s_{t_{k-1}}, t_{k-1}) \Delta s_{t_k} + \varphi_t(s_{t_{k-1}}, t_{k-1}) \Delta t + o(\Delta s_{t_k}, \Delta t),
\]
where $\Delta s_{t_k} = s_{t_k} - s_{t_{k-1}}$, we have
\[
\lim_{\Delta t \to 0} \frac{1}{2\Delta t} \sum_{k=1}^{n} \left| \varphi(s_{t_k}, t_k) - \varphi(s_{t_{k-1}}, t_{k-1}) \right|^2
\]
\[
= \lim_{\Delta t \to 0} \frac{1}{2\Delta t} \sum_{k=1}^{n} \left| \varphi_s(s_{t_{k-1}}, t_{k-1}) \Delta s_{t_k} \right|^2 + o((\Delta s_{t_k})^2, \Delta t)
\]
\[
= \lim_{\Delta t \to 0} \frac{1}{2} \sum_{k=1}^{n} \left| \frac{\Delta s_{t_k}}{\Delta t} a(s_{t_{k-1}}, t_{k-1}) \right|^2 \Delta t
\]
\[
= \frac{1}{2} \int_{0}^{T} \left[ \frac{\dot{s}(t)}{a(s(t), t)} \right]^2 dt.
\]
Therefore, in the limit as $\Delta t$ approaches zero, the optimization problem (2.9) becomes the following variational problem on the space of paths $s : t \mapsto s(t)$
\[
\min_{s : t \mapsto s(t)} \frac{1}{2} \int_{0}^{T} \left[ \frac{\dot{s}(t)}{a(s(t), t)} \right]^2 dt
\]
subject to
\[
\frac{1}{T} \int_{0}^{T} s(t) dt = K, \quad s(0) = s_0.
\]

Definition 2.1.
The optimal path of the variational problem (2.12):(2.13) is referred to as the most-likely-path (MLP) for 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-money Asian call is approximately given by the solution to the constrained variational problem (2.12):(2.13), which is equivalent to determining the most-likely-path. To determine the MLP, the Euler-Lagrange equation associated with the constrained variational problem (2.12):(2.13) along with proper boundary conditions is derived as follows.
Lemma 2.2. The optimal path of the constrained variational problem (2.12); (2.13) satisfies the Euler-Lagrange equation

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

(2.14)

with boundary conditions

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

(2.15)

where $\lambda$ is chosen such that $\frac{1}{T} \int_0^T s(t) dt = K$.

Proof. We first rewrite the constrained variational problem (2.12); (2.13) in the Lagrangian form

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

where $\lambda$ 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

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

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

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

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

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

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

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

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

which simplifies to

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

with boundary conditions

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

□
We summarize the final result in the following theorem. A rigorous proof based on large deviation theory is postponed to Theorem 3.1 in Section 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-money Asian call struck at \( K > s_0 \) and expiry time \( T \), is approximately equal to

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

(2.16)

where \( s : t \mapsto s(t) \) is the solution to the constrained variational problem (2.12):(2.13).

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 (2.14):(2.15). However, for more general cases we will have to resort to an iteration scheme for numerical computations of the most-likely-path. See Section 4 for a numerical scheme.

**Example 2.1. (Bachelier model)**

In the Bachelier model, \( a(s, t) = \sigma \), a constant. The Euler-Lagrange equation (2.14) reduces to

\[
    \frac{d}{dt} \left( \frac{\dot{s}}{\sigma} \right) + \frac{\lambda}{T} \sigma = 0
\]

whose general solution is \( s(t) = -\frac{\lambda \sigma^2}{2T} t^2 + c_1 t + c_0. \) From the boundary conditions (2.15) together with \( \frac{1}{T} \int_0^T s(t) dt = K \), we conclude that \( \lambda = \frac{3(K-s_0)}{\sigma^2 T}, \) \( c_1 = \frac{3(K-s_0)}{T} \), 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

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

(2.17)

We remark that the most-likely-path in this case does not depend on \( \sigma \). Also, (2.16) for the Bachelier model becomes

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

Thus it captures the exponential decay of an out-of-money Asian call in Bachelier model, see (2.22) below.

**Example 2.2. (Black-Scholes model)**

In the Black-Scholes model, \( a(s, t) = \sigma s \) with \( \sigma \) being a constant. The Euler-Lagrange equation (2.14) reads

\[
    -\frac{\ddot{s}}{s^2} + \frac{\dot{s}^2}{s^3} = \lambda \sigma^2
\]
to which the general solution is
\[ s(t) = \frac{c_1^2}{2\lambda\sigma^2} \left[ 1 - \tanh^2 \left( \frac{c_1}{2} [c_2 - t] \right) \right], \]
where \( c_1 \) and \( c_2 \) are (to be determined) constants. The condition \( \dot{s}(T) = 0 \) imply that \( c_2 = T \) and the conditions \( s(0) = s_0, \int_0^T s(t)dt = TK \) become respectively
\[ \frac{c_1^2}{\lambda\sigma^2(1 + \cosh(c_1 T))} = s_0, \quad c_1 \tanh \left( \frac{c_1 T}{2} \right) = \lambda\sigma^2 KT. \]
By dividing the two equations and rearranging terms, \( c_1 \) is given by the solution to the equation
\[ \frac{c_1 T}{\sinh(c_1 T)} = \frac{s_0}{K}, \]
or \( c_1 = \frac{1}{T} f^{-1} \left( \frac{s_0}{K} \right) \) where \( f(x) = \frac{x}{\sinh x} \). Hence, the most-likely-path \( s(t) \) for \( 0 \leq t \leq T \) in the Black-Scholes model after simplification becomes
\[ s(t) = \frac{s_0 \cosh^2 \left( \frac{c_1 T}{2} \right)}{\cosh \left( \frac{c_1 (T - t)}{2} \right)}. \] (2.18)
Moreover
\[ \int_0^T \left| \frac{\dot{s}(t)}{\sigma s(t)} \right|^2 dt = \frac{f^{-1}(\frac{s_0}{K})}{\sigma^2 T} \left[ f^{-1}(\frac{s_0}{K}) - 2 \tanh \left( \frac{1}{2} f^{-1}(\frac{s_0}{K}) \right) \right]. \]

**Example 2.3. (Cox-Ingersoll-Ross model)**
For the CIR model, \( a(s, t) = \sigma \sqrt{s} \), where \( \sigma \) is a constant. The Euler-Lagrange equation (2.14) becomes
\[ \frac{d}{dt} \left( \frac{\dot{s}(t)}{\sqrt{s(t)}} \right) + 2 \phi^2 \sqrt{s(t)} = 2 \frac{d^2}{dt^2} \left( \sqrt{s(t)} \right) + 2 \phi^2 \sqrt{s(t)} = 0 \]
for some constant \( \phi \) with general solution given by
\[ \sqrt{s(t)} = c_2 \cos \left( \phi (c_1 - t) \right). \] (2.19)
The boundary conditions (2.15) imply respectively that
\[ c_1 = T \quad \text{and} \quad c_2 = \frac{s_0}{\cos \phi T}. \]
Thus
\[ s(t) = s_0 \left( \frac{\cos \phi (T - t)}{\cos \phi T} \right)^2. \]
The parameter \( \phi \) is determined by the solution to the equation
\[ \frac{1}{T} \int_0^T s(t) dt = K = \frac{s_0}{2 \phi T} \left[ \tan \phi T + \phi T \sec^2 \phi T \right]. \]
Finally, we have
\[
\int_0^T \left| \frac{\dot{s}(t)}{a(s,t)} \right|^2 dt = s_0 \phi \left[ 2 \phi T - \sin(2 \phi T) \right] \sec^2 \phi T
\]
subject to the determination of \( \phi \).

### 2.3. Implied normal volatility for Asian options

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 \textit{implied normal volatility} in practice.

Recall that, by straightforward calculations, under the Bachelier model
\[
dS_t = \sigma_b dW_t,
\]
where \( \sigma_b \) is a constant, the price of a continuously monitored Asian call option struck at \( K \) with expiry \( T \) has the closed form expression
\[
C_b(K, T, \sigma_b) = \frac{\sigma_b \sqrt{T}}{\sqrt{2 \pi}} e^{-\frac{A_n (s_0 - K)^2}{2\sigma_b^2 t}} + (s_0 - K)N \left( \frac{\sqrt{3}(s_0 - K)}{\sigma_b \sqrt{T}} \right),
\]  
(2.20)

since the average price \( \frac{1}{T} \int_0^T S_t dt \) is normally distributed with mean \( s_0 \) and variance \( \frac{\sigma^2 T}{3} \). In (2.20), \( N(\cdot) \) denotes the cumulative distribution function of standard normal distribution. On the other hand, for 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
\[
\mathbb{E}[S_{t_1} + \cdots + S_{t_n}] = n s_0,
\]
\[
\text{var}[S_{t_1} + \cdots + S_{t_n}] = \sigma^2 T \frac{(n + 1)(2n + 7)}{6},
\]
its price \( C_{bd} \) is given by
\[
C_{bd}(K, T, \sigma_b) = \frac{\sigma_b \sqrt{T}}{\sqrt{2 A_n \pi}} e^{-\frac{A_n (s_0 - K)^2}{2\sigma_b^2 t}} + (s_0 - K)N \left( \frac{\sqrt{A_n}(s_0 - K)}{\sigma_b \sqrt{T}} \right),
\]  
(2.21)

where \( A_n := \frac{(n+1)(2n+7)}{6n^2} \). Apparently, \( A_n \to 3 \) as \( n \to \infty \).

As small time expansion is concerned, we remark that by using the asymptotic expansion \( N(x) = N(x)[-\frac{1}{x} + \frac{1}{x^3} + O(x^{-5})] \) as \( x \to -\infty \), we have
\[
C_b(s_t, K) = \frac{\sigma_b \sqrt{T - t}}{\sqrt{6 \pi}} e^{-\frac{3(s_t - K)^2}{2\sigma_b^2 (T - t)}} + (s_t - K)N \left( \frac{\sqrt{3}(s_t - K)}{\sigma_b \sqrt{T - t}} \right)
\]
\[
= e^{-\frac{3(s_t - K)^2}{2\sigma_b^2 (T - t)}} \left[ \frac{\sigma_b^2 (T - t)^{\frac{3}{2}}}{3 \sqrt{6 \pi} (K - s_0)^2} + O(T - t)^{\frac{5}{2}} \right]
\]  
(2.22)
as \( t \to T^− \). Similarly, in the discrete case as \( t \to T \),
\[
C^d_b(s_t, K) = e^{-\frac{A_n(s_t-K)^2}{2\sigma^2_b(T-t)}} \left[ \frac{\sigma^b_3(T-t)^{\frac{3}{2}}}{A_n\sqrt{2A_n\pi}(K-s_0)^2} + \mathcal{O}(T-t)^{\frac{5}{2}} \right].
\] (2.23)

We remark that once the small time asymptotic for the price of an out-of-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 call price given in [18] will be helpful.

**Proposition 2.1.** For a fixed strike out-of-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 \( \sigma_N^2T \) has the asymptotic
\[
\sigma_N^2T = \frac{(s_0-K)^2}{2(\log s_0 - \log C(T))} + o(\log C(T)).
\] (2.24)
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 \( \sigma_b \)
\[
C^d(s_t, K, T) = \frac{\sigma_b\sqrt{T}}{\sqrt{2A_n\pi}} e^{-\frac{A_n(s_t-K)^2}{2\sigma^2_bT}} + (s_0-K)N\left( \frac{\sqrt{A_n}(s_0-K)}{\sigma_b\sqrt{T}} \right),
\] (2.25)
where \( A_n := \frac{(n+1)(2n+7)}{6n^2} \) and \( C^d \) is price of a discretely monitored Asian call obtained by the model or from the market. Obviously, among other parameters, such defined \( \sigma_b \) depends on \( K \) and \( T \).

To derive an asymptotic expansion for the implied normal volatility \( \sigma_b \) defined in (2.25) in small time, the idea, as in Gatheral et al [13], is to compare the corresponding terms in the expansions on both side of (2.25). The lowest order term is thus obtained by matching the exponential terms on both side of (2.25). Precisely, recall the small time asymptotic of the price of a discrete Asian call from Theorem 2.2
\[
E \left[ \left( \frac{1}{n} \sum_{i=1}^{n} S_{t_i} - K \right)^+ \right] = \frac{\Delta t^{\frac{3}{2}}}{\sqrt{2\pi}} e^{-\frac{D(s^*, t)^2}{\Delta t}} \left[ \frac{\nabla W(s^*, t) \cdot \nabla D(s^*, t)}{\sqrt{\det \partial_t^2 D(s^*, t) \nabla D(s^*, t)^2}} \right] + \mathcal{O}(\Delta t),
\]
where $\Delta t = \frac{T}{n}$. By matching the exponential terms in the above expression and in (2.23), we have

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

which implies

$$1 = \frac{2nD(s^*, t)}{A_n(s_t-K)^2} = \frac{n}{A_n(s_t-K)^2} \sum_{k=1}^{n} \left| \varphi(s^*_t, t_i) - \varphi(s^*_{i-1}, t_{i-1}) \right|^2.$$

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

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

where recall that $(s^*_t, \cdots, s^*_n)$ is the solution to the $n$-dimensional constrained optimization problem (2.9)-(2.10). 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-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 $\sigma^d_b$ defined as in (2.25) has the asymptotic expansion as $T \to 0^+$

$$\sigma^d_b(K,T) = \sigma^d_{b,0}(K) + O(T),$$

where

$$\sigma^d_{b,0}(K) = \sqrt{A_n|s_t-K|} \left( n \sum_{k=1}^{n} \left| \varphi(s^*_t, t_i) - \varphi(s^*_{i-1}, t_{i-1}) \right|^2 \right)^{-\frac{1}{2}},$$

with $A_n = \frac{(n+1)(2n+7)}{6n^2}$ and $(s^*_t, \cdots, s^*_n)$ the solution to the constrained optimization problem (2.9)-(2.10).

**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 $\sigma_b$

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

where $C$ is price of an Asian call obtained by the model or from the market. Notice that the right hand side of (2.26) can be regarded as the price of a European call option in Bachelier model but with one third of $\sigma^2_b$ 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 (2.16) for call price and (2.24) for implied normal volatility, a small time asymptotic for the implied normal volatility of an out-of-money Asian call is established. Alternative, we may also obtain the same approximation by straightforwardly taking the limit as \( n \to \infty \) of \( \sigma_{b,0} \) 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-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 \( \sigma_b \) defined in (2.26) has the asymptotic expansion as \( T \to 0^+ \)

\[
\sigma_b(K, T) = \sigma_{b,0} + o(T),
\]

where

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

and \( \hat{s}(t) \) is the most-likely-path for Asian option determined by solving the variational problem (2.12) : (2.13).

**Proof.** Recall from (2.24) that the implied normal volatility \( \sigma \) of a European option has the asymptotic

\[
\sigma^2 T = \frac{(s_0 - K)^2}{2(\log s_0 - \log C(T))} + o(\log C(T)).
\]

Thus, by substituting \( \log C(T) \) with (2.16) and using the fact that Asian variance equals one third of its European counterpart in Bachelier world, we obtain for implied normal volatility \( \sigma_b \) of an Asian call

\[
\frac{\sigma_b^2 T}{3} = \frac{(s_0 - K)^2}{2(\log s_0 - \log C(T))} + o(\log C(T))
\]

\[
= -\frac{(K-s_0)^2}{2\log C(T)} + o(\log C(T))
\]

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

where we used (2.16) in the last equality. Finally, the result is obtained by rearranging terms. \( \square \)

**Example 2.4. (Time-dependent Bachelier model)**
Finally, we consider a time-dependent Bachelier model in which
\[ dS_t = \sigma \theta(t) \, dW_t. \]

Note that in this case \( a(s, t) = \sigma \theta(t) \), the Euler-Lagrange equation (2.14) reduces to
\[ \frac{d}{dt} \left( \frac{\dot{s}}{\sigma \theta(t)} \right) - \frac{\theta'(t)}{\theta^2(t)} \frac{\dot{s}}{\theta(t)} + \frac{\lambda}{T} \sigma \theta(t) = 0 \]
\[ \rightarrow \frac{1}{\theta(t)} \frac{d}{dt} \left( \frac{\dot{s}}{\theta(t)} \right) - \frac{\theta'(t)}{\theta^2(t)} \frac{\dot{s}}{\theta(t)} + \frac{\lambda}{T} \sigma^2 = 0 \]
Integrating the last equation subject to the condition \( \dot{s}(T) = 0 \) gives
\[ \frac{\dot{s}(t)}{\theta^2(t)} = \frac{\lambda}{T} \sigma^2 (T - t) \] (2.27)
for some constant \( \lambda \) set using the condition
\[ \frac{1}{T} \int_0^T s(t) \, dt = K. \]

This gives
\[ \lambda = \frac{T^2 (K - s_0)}{\sigma^2 \int_0^T (T - u)^2 \theta^2(u) \, du}. \]

Also, from (2.27) one easily obtain
\[ \int_0^T \left| \frac{\dot{s}(t)}{a(s, t)} \right|^2 dt = \int_0^T \left| \frac{\dot{s}(t)}{\sigma \theta(t)} \right|^2 dt = \frac{\lambda^2 \sigma^2}{T^2} \int_0^T (T - t)^2 \theta^2(t) \, dt = \frac{T^2 (K - s_0)^2}{\sigma^2 \int_0^T (T - t)^2 \theta^2(t) \, dt}. \]
This gives
\[ \sigma^2_{b,0} = \frac{3 \sigma^2}{T^3} \int_0^T (T - t)^2 \theta^2(t) \, dt. \] (2.28)

On the other hand, we have \( \bar{S}_T = \frac{1}{T} \int_0^T S_t \, dt \) where
\[ S_t = s_0 + \sigma \int_0^t \theta(u) \, dW_u. \]
A straightforward computation then yields
\[ \text{var} \left[ \bar{S}_T \right] = \frac{\sigma^2}{T^2} \int_0^T (T - u)^2 \theta(u)^2 \, du \]
so the exact Bachelier implied volatility in the time-dependent Bachelier case is given by
\[ \sigma^2_b = \frac{3 \sigma^2}{T^3} \int_0^T (T - u)^2 \theta(u)^2 \, du. \]
Thus, the most-likely-path approximation is exact in the time-dependent Bachelier case.
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 (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-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 $T$

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

(3.1)

where

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

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

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

(3.2)

and $I(f)$ is $\infty$ 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

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

(3.3)

Note that with this notation:

$$\frac{1}{T} \int_0^T S_t dt = \int_0^1 S^\varepsilon_u du.$$

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

$$dS^\varepsilon_u = a_\varepsilon(S^\varepsilon_u, u) \sqrt{\varepsilon}dW_u,$$

(3.4)

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

$$a(x, u\varepsilon) = a(x, 0) + O(\varepsilon).$$
A large deviation principle for the process means that there exists a rate function $I : C([0, T]) \rightarrow [0, \infty]$ that is lower semi-continuous and has compact level sets such that for any Borel subset $A$ of paths in $C([0, 1])$,

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

(3.5)

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. [10]. 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 and Fischer [8]. 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 [8].

**Theorem 3.2** (Theorem 2, Theorem 4 and Example 1 in [8]). The family of diffusions (3.4) satisfy a large deviation principle with rate function

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

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

Formally, it is good to think of the set $\{g \in L^2([0,1]) : f(t) = x + \int_0^t a(f(s), 0)g(s)ds\}$ as the set of “white-noise paths” $t \mapsto W_t$. The map $g \mapsto f$ where $f$ is the solution of $f(t) = x + \int_0^t a(f(s), 0)g(s)ds$ 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 non-zero for the path $f$, the map can be inverted and the rate function reduces to the simplest case

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

(3.6)

This is certainly the case for geometric Brownian motion and the CIR model. The reader is referred to Theorem 1 in [8] and Proposition 3.11 in [3] for general sufficient conditions for (3.6) to hold.
Proof of Theorem 3.1. Observe that for any random variable \( X \) and \( K > 0 \), we have the identity
\[
\mathbb{E}[(X - K)^+] = \int_K^\infty \mathbb{P}(X > x) \, dx.
\] (3.7)

For simplicity, write \( T \) for the functional \( T : C([0, 1]) \to \mathbb{R} \) with \( T(f) = \int_0^1 f(u) \, du \). Note that \( T \) is continuous on \( C([0, 1]) \) equipped with the topology of uniform convergence. By the contraction principle (see e.g. [10]) 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
\[
J(y) = \inf \{ I(f) : T(f) = y, f \in C([0, 1]) \} , y \in \mathbb{R} .
\]

By (3.7), we have
\[
\log \mathbb{E}[(T(S^\varepsilon) - K)^+] = \log \int_K^\infty \mathbb{P}(T(S^\varepsilon) > x) \, dx.
\]

On one hand, we have for any \( M > K \)
\[
\log \int_K^\infty \mathbb{P}(T(S^\varepsilon) > x) \, dx = \left( \log \int_K^M \mathbb{P}(T(S^\varepsilon) > x) \, dx \right) + \log \left( 1 + \frac{\log \int_M^\infty \mathbb{P}(T(S^\varepsilon) > x) \, dx}{\log \int_K^M \mathbb{P}(T(S^\varepsilon) > x) \, dx} \right).
\]

For \( M \) large enough (possibly dependent on \( \varepsilon \)), the term in the second parenthesis is smaller than 2. Pick \( M = \varepsilon^{-8} \). For \( \varepsilon \) small enough we thus have
\[
\log \int_K^\infty \mathbb{P}(T(S^\varepsilon) > x) \, dx \leq \log 2 + \log(\varepsilon^{-1} - K) + \log \mathbb{P}(T(S^\varepsilon) > K).
\]

This proves that
\[
\limsup_{\varepsilon \to 0} \varepsilon \log \int_K^\infty \mathbb{P}(T(S^\varepsilon) > x) \, dx \leq \lim_{\varepsilon \to 0} \varepsilon \log \mathbb{P}(T(S^\varepsilon) > K) = -J(K),
\]

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

On the other hand, for \( \delta > 0 \)
\[
\int_K^\infty \mathbb{P}(T(S^\varepsilon) > x) \, dx \geq \int_K^{K+\delta} \mathbb{P}(T(S^\varepsilon) > x) \, dx \geq \delta \mathbb{P}(T(S^\varepsilon) > K + \delta).
\]

Therefore, for any \( \delta > 0 \), the above with the use of (3.5) becomes
\[
\liminf_{\varepsilon \to 0} \varepsilon \log \int_K^\infty \mathbb{P}(T(S^\varepsilon) > x) \, dx \geq -J(K + \delta).
\]

It remains to show that \( \lim_{\delta \to 0} J(K + \delta) = J(K) \). First, notice that \( \liminf_{\delta \to 0} J(K + \delta) \geq J(K) \) since \( J \) is lower semi-continuous (it is a rate function). So it suffices to show that \( \limsup_{\delta \to 0} J(K + \delta) \leq J(K) \). By definition, \( J(K + \delta) = \inf \{ I(x) : x \in C([0, 1]) \text{ and } \int_0^1 x(u) \, du = K + \delta \} \). Pick a sequence \( (y_n) \in C([0, 1]) \) 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

\[
J(K + \delta) < I(y_n + \delta z).
\]

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

\[
\limsup_{\delta \to 0} J(K + \delta) < I(y_n) < J(K) - 1/n.
\]

Since \( n \) is arbitrary, this proves the claim. \( \square \)

The asymptotic expansion (2.16) suggests that the lower order corrections to (3.1) are much smaller than the dominant term in \( T^{-1} \). We expect more precisely that

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

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 (2.17) when written as a path on \([0, 1]\) is

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

By (3.7), the price of the call option becomes

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

Write \( \mathbb{Q} \) for the measure with \( \frac{d\mathbb{Q}}{d\mathbb{P}} = \exp \left( \varepsilon^{-1/2} \int_0^1 \dot{s}(u) dW_u - \frac{\varepsilon^{-1/2}}{2} \int_0^1 \dot{s}(u)^2 du \right) \). 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, (3.9) becomes

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

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

\[
\mathbb{E} \left[ \left( \int_0^1 S_u^\varepsilon du - K \right)^+ \right] = e^{-\frac{\varepsilon^{-1}}{2} \int_0^1 \dot{s}(u)^2 du} \int_0^\infty \mathbb{E}_\mathbb{Q} \left[ e^{-\varepsilon^{-1/2} \int_0^1 \dot{s}(u) d\widetilde{W}_u} 1_{\{\int_0^1 \varepsilon^{1/2} \dot{s}(u) d\widetilde{W}_u > x\}} \right] dx.
\]
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 to get
\[ E_Q \left[ e^{-\varepsilon^{-1/2} \int_0^1 \dot{s}(u) d\tilde{W}_u \left( \int_0^1 \varepsilon^{1/2} \tilde{W}_u \, du \right)^2} \right]. \]

Note that $\dot{s}(u) = 3(K - s_0)(1 - u)$, therefore
\[ \int_0^1 \dot{s}(u) d\tilde{W}_u = 3(K - s_0) \int_0^1 \tilde{W}_u du. \]

Write $X$ for the random variable $\int_0^1 \tilde{W}_u du$ which is Gaussian with mean 0 and variance $1/3$. We have
\[ \varepsilon^{1/2} E_Q \left[ X^+ e^{-\varepsilon^{-1/2} 3(K - s_0) X} \right] = \varepsilon^{3/2} \int_0^\infty y e^{-3(K - s_0)y} \frac{e^{-\frac{3y^2}{2}}}{\sqrt{2\pi}/3} dy. \]

The integral is of order 1. We conclude that
\[ E \left[ \left( \frac{1}{T} \int_0^T S_u \, du - K \right)^+ \right] = \exp \left( -T^{-1} J(K) + \frac{3}{2} \log T + O(1) \right). \]

4. Numerical tests

From Theorem 2.5, we have the following approximate formula for Asian Bachelier implied volatility:
\[ \sigma_{b,0} = \left( \frac{T}{3(K - s_0)^2} \int_0^T \left[ \frac{\dot{s}(t)}{a(s(t), t)} \right]^2 dt \right)^{-\frac{1}{2}}, \quad (4.1) \]

where $\tilde{s}(t)$ is the most-likely-path for an Asian option determined by solving the variational problem
\[ \min_{s: \int_0^T s(t) dt = K, \ s(0) = s_0} \frac{1}{T} \int_0^T \left[ \frac{\dot{s}(t)}{a(s(t), t)} \right]^2 dt \]

subject to
\[ \frac{1}{T} \int_0^T s(t) dt = K, \ s(0) = s_0. \]

We now proceed to test numerically the approximate implied volatility formula (4.1) for various definitions of the local volatility function $a(s, t) = s \sigma(t, s)$ 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
\[ \tilde{s}(t) = s_0 + \frac{I(t)}{I(T)} [K - s_0] \]
where
\[ I(t) = \int_0^t \int_r^T a(\tilde{s}(r), r) a(\tilde{s}(u), u)e^{-\int_u^r \frac{a_t(\tilde{s}(v), v)}{a(\tilde{s}(v), v)} dv} du dr, \]
\[ \bar{I}(T) = \frac{1}{T} \int_0^T I(u) du. \]

Proof. From Lemma 2.2, \( \tilde{s}(t) \) satisfies the Euler-Lagrange equation (2.14) which we reiterate in the following for convenience
\[ \frac{d}{dt} \left( \frac{\dot{s}}{a} \right) - \frac{a_t}{a^2} \dot{s} + \frac{\lambda}{T} a = 0 \quad (4.2) \]
with boundary conditions
\[ s(0) = s_0, \quad \dot{s}(T) = 0, \]
where \( \lambda \) is chosen such that \( \frac{1}{T} \int_0^T s(t) dt = K. \)

For ease of notation, define \( \tilde{a}(t) := a(\tilde{s}(t), t) \) and \( f(s, t) = \frac{a_t}{a} = \partial_t \log a. \) Also, \( \tilde{f}(t) = f(\tilde{s}(t), t). \) Applying the integrating factor \( e^{-\int_0^t \tilde{f}(u) du} \) and integrating (4.2) with the boundary condition \( \dot{s}(T) = 0 \) gives
\[ -e^{-\int_0^t \tilde{f}(v) dv} \dot{s}(t) = -\frac{\lambda}{T} \int_t^T \tilde{a}(u) e^{-\int_u^t \tilde{f}(v) dv} du. \]
It follows that
\[ \frac{\dot{s}(t)}{\tilde{a}(t)} = \frac{\lambda}{T} \int_t^T \tilde{a}(u) e^{-\int_u^t \tilde{f}(v) dv} du. \]
Rearranging and integrating again gives
\[ s(t) - s_0 = \frac{\lambda}{T} I(t), \]
where \( I(t) = \int_0^t \int_r^T \tilde{a}(r) \tilde{a}(u) e^{-\int_u^r \tilde{f}(v) dv} du dr. \) Now apply the boundary condition \( \frac{1}{T} \int_0^T s(t) dt = K \) to get
\[ K - s_0 = \frac{\lambda}{T} \bar{I}(T) \]
and the result follows.

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 (2.17):
\[ s(t) = s_0 + 3(K - s_0) \frac{t}{T} - \frac{3(K - s_0)}{2} \left( \frac{t}{T} \right)^2. \]
The resulting algorithm typically converges sufficiently after three or four iterations.
4.1. **A time-dependent CIR model.** We consider the model

\[ dS_t = e^{-\lambda t} \sigma \sqrt{S_t} dW_t \]

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

\[ a(s, t) = e^{-\lambda t} \sigma \sqrt{s}. \] (4.3)

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 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 Figure 4.1. We remark that the most-likely-path approximation in this example slightly underestimates the normal implied volatility for Asian option inferred from simulation.

![Figure 4.1](image-url)  

**Figure 4.1.** The 1-year Bachelier Asian implied volatility smile corresponding to the time-dependent CIR local volatility function (4.3). The blue dotted line is from Monte Carlo simulation with error bars; the red solid line is the approximation \( \sigma_{b,0} \).

4.2. **Time-dependent quadratic local volatility.** Next we consider the following quadratic local volatility model:

\[ dS_t = e^{-\lambda t} \sigma \left[ 1 + \psi(S_t - 1) + \frac{\gamma}{2} (S_t - 1)^2 \right] dW_t \] (4.4)
with $\sigma = 0.2$, $\psi = -0.5$, $\gamma = 0.1$, and $\lambda = 1$.

Though a closed-form solution for European options with these parameters is given in [2], 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 Figure 4.2. Similarly, the most-likely-path approximation in this example also underestimates the normal implied volatility for Asian option inferred from simulation.

![Figure 4.2. The 1-year Bachelier Asian implied volatility smile corresponding to the time-dependent quadratic local volatility function (4.4). The blue dotted line is from Monte Carlo simulation with error bars; the red solid line is the approximation $\sigma_{b,0}$.](image)

5. Conclusion

We derive 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 is 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 show that the most-likely-path approximation underestimates. It can possibly be improved by including 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 limit of corresponding terms from the discretely
monitored case may yield tractable expressions for numerical evaluations. The calcula-
tions of higher order terms are considerably more involved and henceforth are left
to further work.

ACKNOWLEDGEMENT

We would like to express our gratitude to Jim Gatheral for numerical results and
Dan Pirjol for sending us an early version of their paper. L.-P. A. is supported by NSF
CAREER 1653602, NSF grant DMS-1513441, and a Eugene M. Lang Junior Faculty
Research Fellowship. N.-L. L. is supported by JSPS KAKENHI Grant Numbers
25780213.

6. Appendix I - Convexity of the constrained optimization problem

We analyze the convexity of the constrained optimization problem  \((2.9):(2.10)\) in
this appendix. Recall the Lamperti transformation 
\[ \varphi(s,t) := \int_{s_0}^s \frac{d\xi}{a(\xi,t)} \]
and the objec-
tive function
\[ D(s,t) = \frac{1}{2} \sum_{i=1}^n (\varphi(s_i,t_i) - \varphi(s_{i-1},t_{i-1}))^2. \]

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
\[
\frac{\partial^2 D}{\partial s_i^2} = \frac{2}{a_i^2(s_i)} + \frac{a'_i(s_i)}{a_i^2(s_i)} \left[ \varphi_{i-1}(s_{i-1}) + \varphi_{i+1}(s_{i+1}) - 2\varphi_i(s_i) \right];
\]
\[
\frac{\partial^2 D}{\partial s_n^2} = \frac{1}{a_n^2(s_n)} - \frac{a'_n(s_n)}{a_n^2(s_n)} \left[ \varphi_{n-1}(s_{n-1}) - \varphi_{n-1}(s_{n-1}) \right];
\]
\[
\frac{\partial^2 D}{\partial s_i \partial s_j} = \frac{-1}{a_i(s_i)a_j(s_j)}, \quad \text{if } |i - j| = 1;
\]
\[
\frac{\partial^2 D}{\partial s_i \partial s_j} = 0, \quad \text{if } |i - j| \geq 2.
\]

We decompose the Hessian matrix \(H = \left[ \frac{\partial^2 D}{\partial s_i \partial s_j} \right]\) as \(H = H^1 + H^2\), where \(H^1\) is the
symmetric tridiagonal matrix with diagonal entries given by
\[
H^1_{kk} = \frac{2}{a_k^2(s_k)} \quad \text{for } k = 1, \ldots, n - 1, \quad H^1_{nn} = \frac{1}{a_n^2(s_n)}
\]
and off-diagonal entries by

\[ H_{k,k+1}^1 = \frac{-1}{a_k(s_k)a_{k+1}(s_{k+1})} \text{ for } k = 1, \ldots, n - 1. \]

\( H^2 \) is a diagonal matrix with diagonal entries given by

\[ H^2_{kk} = \frac{a'_k(s_k)}{a^2_k(s_k)}[\varphi_{k-1}(s_{k-1}) + \varphi_{k+1}(s_{k+1}) - 2\varphi_k(s_k)] \text{ for } k = 1, \ldots, n - 1, \]

\[ H^2_{nn} = -\frac{a'_n(s_n)}{a^2_n(s_n)}[\varphi_{n-1}(s_{n-1}) - \varphi_{n-1}(s_{n-1})]. \]

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

\[
\begin{pmatrix}
\hat{H}^1 & c^T \\
c & H^1_{nn}
\end{pmatrix},
\]

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 \((k+1)\prod_{j=1}^{k} \frac{1}{a^2_j(s_j)} > 0\), for \( k = 1, \ldots, n - 1 \).

By applying the identity for determinant, if \( A \) is invertible,

\[
\det \begin{bmatrix} A & b^T \\ b & d \end{bmatrix} = \det(A) \times (d - bA^{-1}b^T),
\]

we calculate the determinant of \( H^1 \) as

\[
\det(H^1) = \det(\hat{H}^1) \times \left\{ \frac{1}{H^1_{nn} - c(\hat{H}^1)^{-1}c^T} \right\} = \det(\hat{H}^1)H^1_{nn} - (H^1_{n-1,n})^2 \det(\hat{H}^1) = \prod_{k=1}^{n} \frac{1}{a^2_k(s_k)} > 0,
\]

where \( \hat{H}^1 \) 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

\[
\varphi_{k-1}(s_{k-1}) + \varphi_{k+1}(s_{k+1}) - 2\varphi_k(s_k) = \varphi'_k(s_k)(\Delta s_{k+1} + \Delta s_k) + o(\Delta s_{k+1}, \Delta s_k, \Delta t)
\]

\[
\varphi_{n-1}(s_n) - \varphi_{n-1}(s_{n-1}) = \varphi'_{n-1}(s_{n-1})\Delta s_n + o(\Delta s_n),
\]

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 \( \Delta t \) is small. In this case, \( H \) is positive definite. Hence, the objective function \( D \) is convex.
7. Appendix II - Proof of the Laplace asymptotic formula

This appendix is devoted to the proof of Lemma 2.1, the Laplace asymptotic formula which is slightly different from the conventional one in that the minimum point is attained at a boundary point rather than an interior point. Nevertheless, the conventional technique applies in this case. We repeat the statement in the following for convenience.

**Lemma 7.1.** (Laplace asymptotic formula)

Let $R$ be a closed set in $\mathbb{R}^n$ with nonempty and smooth boundary $\partial R$. Suppose $\theta$ 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 $x \in R \setminus B_\epsilon(x^*)$, where $B_\epsilon(x^*) = \{ x : |x - x^*| < \epsilon \}$ is the open ball of radius $\epsilon$ 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 $\partial R$ but the inward normal directional derivative of $f$ at $x^*$ is nonzero.

Then we have the asymptotic expansion as $\tau \to 0^+$

$$
\int_R e^{-\theta(x)/\tau} f(x) dx = \left(\frac{2\pi}{\tau} \right)^{n-1} \frac{n+2}{n+3} e^{-\theta(x^*)/\tau} \frac{\nabla f(x^*) \cdot \nabla \theta(x^*)}{|\nabla \theta(x^*)|^2} + \frac{1}{2} \text{tr} \left( \partial^2 f(x^*) \left[ \partial^2 \theta(x^*) \right]^{-1} \right) + O(\tau),
$$

where $\partial^2 f(x^*)$ and $\partial^2 \theta(x^*)$ are the Hessian matrices of $f$ and $\theta$ respectively in the tangential direction to $R$ at $x^*$.

**Proof.** For any $\epsilon > 0$, we split the integral on the left side of (7.1) into two parts as

$$
\int_R e^{-\theta(x)/\tau} f(x) dx = \int_{R \cap B_\epsilon(x^*)} e^{-\theta(x)/\tau} f(x) dx + \int_{R \setminus B_\epsilon(x^*)} e^{-\theta(x)/\tau} f(x) dx.
$$

We treat the two terms on the right hand side of (7.2) individually. For the first term, since the integration region is restricted to a subset of the small ball $B_\epsilon(x^*)$, it can be reparametrized by $y = (y_1, \ldots, y_n)$ so that in the $y$-coordinates the set $\{ y : y^n = 0 \}$ corresponds to $\partial R$ and the vectors $\{ \partial_{y_1}, \ldots, \partial_{y^n} \}$ form a local orthonormal frame around $x^*$. For simplicity, we further assume that in the $y$-coordinates $x^*$ is located at the origin. Note that in the $y$-coordinates the vector $\partial_{y^n}$ is parallel to $\nabla \theta(x^*)$ and the inward normal vector of $R$ at $x^*$ as well. We shall use the conventions for the indices

$$
1 \leq \alpha, \beta, \gamma \cdots \leq n,
$$

$$
1 \leq i, j, k, \cdots \leq n - 1.
$$
Also, repeated indices will be summed up over their respective ranges.

Denote partial derivatives by subindices we have for \( y \in B_{\epsilon}(x^*) \)
\[
\theta(y) = \theta(0) + \theta_n(0)y^n + \frac{1}{2}\theta_{\alpha\beta}(0)y^\alpha y^\beta + \mathcal{O}(|y|^2),
\]
\[
f(y) = f_\alpha(0)y^\alpha + \frac{1}{2}f_{\alpha\beta}(0)y^\alpha y^\beta + \mathcal{O}(|y|^2)
\]
since \( \theta_1(0) = \cdots = \theta_{n-1}(0) = 0 \) for \( \theta \) attains its minimum at the boundary point \( x^* \). Hence, in the \( y \)-coordinates the first integral on the right hand side of (7.2) reads
\[
\int e^{-\Theta(x)} f(x)dx
\]
\[
= \int e^{-\frac{1}{2}(\theta(0) + \theta_n(0)y^n + \frac{1}{2}\theta_{\alpha\beta}(0)y^\alpha y^\beta + \mathcal{O}(|y|^2))} f_\alpha(0)y^\alpha + \frac{1}{2}f_{\alpha\beta}(0)y^\alpha y^\beta + \mathcal{O}(|y|^2)} \, dy d^n y,
\]
where \( dy = dy^1 \cdots dy^{n-1} \). Now do the change of variable
\[
y^n = \tau z^n, \quad y^i = \sqrt{\tau} z^i.
\]
The last integral becomes
\[
e^{-\frac{\Theta(0)}{\tau}} \tau^{n+1} \int_0^{\sqrt{\tau}} \int_{[-\sqrt{\tau}, \sqrt{\tau}]} e^{-\left(\theta_n(0)z^n + \frac{1}{2}\theta_{\alpha\beta}(0)z^\alpha z^\beta + o(\tau)\right)} \times
\]
\[
\left[ f_i(0)z^i\sqrt{\tau} + f_n(0)z^n\tau + \frac{1}{2}f_{ij}(0)z^i z^j \tau + o(\tau) \right] d\hat{z} dz^n.
\]
Define the integrals by
\[
I = \int_0^{\sqrt{\tau}} \int_{[-\sqrt{\tau}, \sqrt{\tau}]} e^{-\left(\theta_n(0)z^n + \frac{1}{2}\theta_{\alpha\beta}(0)z^\alpha z^\beta + o(\tau)\right)} f_i(0)z^i d\hat{z} dz^n,
\]
\[
II = \int_0^{\sqrt{\tau}} \int_{[-\sqrt{\tau}, \sqrt{\tau}]} e^{-\left(\theta_n(0)z^n + \frac{1}{2}\theta_{\alpha\beta}(0)z^\alpha z^\beta + o(\tau)\right)} f_n(0)z^n d\hat{z} dz^n,
\]
\[
III = \int_0^{\sqrt{\tau}} \int_{[-\sqrt{\tau}, \sqrt{\tau}]} e^{-\left(\theta_n(0)z^n + \frac{1}{2}\theta_{\alpha\beta}(0)z^\alpha z^\beta + o(\tau)\right)} \frac{1}{2}f_{ij}(0)z^i z^j d\hat{z} dz^n.
\]
As \( \tau \to 0^+ \), we calculate the each integral individually as follows.
\[
I = f_i(0) \int_0^{\sqrt{\tau}} e^{-\theta_n(0)z^n} d\hat{z} \times \int_{[-\sqrt{\tau}, \sqrt{\tau}]} e^{-\left(\frac{1}{2}\theta_{\alpha\beta}(0)z^\alpha z^\beta + o(\tau)\right)} z^i d\hat{z}
\]
\[
\to f_i(0) \int_0^{\infty} e^{-\theta_n(0)z^n} dz \times \int_{\mathbb{R}^{n-1}} e^{-\frac{1}{2}\theta_{\alpha\beta}(0)z^\alpha z^\beta} dz = 0.
\]
For $II$,

$$II = \frac{f_n(0)}{\theta_n^2(0)} \times \frac{(2\pi)^{n-1}}{\sqrt{\det(\theta_{kl}(0))}} \times e^{-\frac{\theta_n^2(0)\tau}{2}}$$

and $III$

$$III = \frac{f_{ij}(0)}{2} \times \frac{(2\pi)^{n-1}}{\sqrt{\det(\theta_{kl}(0))}} \times e^{-\frac{\theta_n^2(0)\tau}{2}}$$

where $\theta_{ij}(0)$ is the inverse matrix of $\theta_{ij}(0)$. Moreover, by applying the argument as in next paragraph, one can show that all the terms that are left out in each integral in the limit are exponentially small compared to the limiting term. Thus, they do not contribute to the asymptotic expansion.

On the other hand, concerning the second term

$$\left| \int_{R \setminus B(x^*)} e^{-\frac{\theta_n(0)}{\tau}} f(x) \, dx \right| \leq \int_{R \setminus B(x^*)} e^{-\frac{\theta_n(0)+\delta}{\tau}} |f(x)| \, dx \leq e^{-\frac{\delta}{\tau}} \int_{R} |f(x)| \, dx.$$

As a result the second term is exponentially small (at the rate $\delta$) as $\tau \to 0^+$ compared to the expansion (7.1) obtained for the first term, hence it does not contribute to the asymptotic expansion. Finally, we obtain the Laplace expansion (7.1) by rewriting the expressions for $II$ and $III$ in a coordinate independent form. $\square$

REFERENCES

[1] Albrecher, H., P.A. Mayer, and W. Schoutens, General lower bounds for arithmetic Asian option prices, Applied Mathematical Finance, 15(2), pp.123-149, 2008.
[2] Andersen, L, Option Pricing with Quadratic Volatility: A Revisit, Finance and Stochastics, 15(2), pp.191-219, 2011.
[3] Baldi, P. and Caramellino, L., General Freidlin-Wentzell large deviations and positive diffusions, Statistics and Probability Letters, 81(8), pp.1218-1229, 2011.
[4] Bleistein, N. and Handelsman, R.A., Asymptotic expansions of integrals, Dover Publications (1986).
[5] Boyle, P.P. and Emanuel, D., Discretely adjusted option hedges, Journal of Financial Economics, 8(3), pp.259-282, 1980.
[6] Broadie, M. and Glasserman, P., Estimating security price derivatives using simulation, Management Science, 42, pp.269-285, 1996.
[7] Cai, N., Li, C., and Shi, C., Closed-form expansions of discretely monitored Asian options in diffusion models, Mathematics of Operations Research, 39(3), pp.789-822, 2014.

[8] Chiarini, A. and Fisher, M., On large deviations for small noise Itô processes, Advances in Applied Probability, 46(4), pp.1126-1147, 2014.

[9] Currán, M., Valuing Asian and portfolio options by conditioning on the geometric mean, Management Science, 40(21), pp.1705-1711, 1994.

[10] Dembo, A. and Zeitouni, O., Large deviations techniques and applications, 2nd edition, Springer-Verlag (1998).

[11] Dufresne, D., Laguerre series for Asian and other options, Mathematical Finance, 10(4), pp.407-428, 2000.

[12] Fouque, J.-P. and Han, C.H., Pricing Asian options with stochastic volatility, Quantitative Finance, 3(5), pp.352-362, 2003.

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

[14] Gatheral, J., Hsu, E.P., Laurence, P., Ouyang, C., and Wang, T.-H., Asymptotics of implied volatility in local volatility models, Mathematical Finance, 22(4), pp.591-620, 2012.

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

[16] Geman, H. and Yor, M., Bessel processes, Asian options, and perpetuities, Mathematical Finance, 3, pp.349-375, 1993.

[17] Glasserman, P., Heidelberger, P., and Shahabuddin, P., Asymptotically optimal importance sampling and stratification for pricing path-dependent options, Mathematical Finance, 9(2), pp.117-152, 1999.

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

[19] Kemna, A.G.Z. and Vorst, A.C.F., A pricing method for options based on average asset values, Journal of Banking and Finance, 14, pp.113-129, 1990.

[20] Kunitomo, N. and Takahashi, A., Pricing average options, Japan Financial Review, 14, pp.1-20, 1992.

[21] Pirjol, D. and Zhu, L., Short maturity Asian options in local volatility models, SIAM Journal of Financial Mathematics, 7(1), pp.947-992, 2016.

[22] Ritchken, P., Sankarasubramanian, L., and VIJH, A.M., The valuation of path dependent contracts on the average, Management Science, 39, pp.1202-1213, 1993.

[23] Rogers, L.C.G. and Shi, Z., The value of an Asian option, Journal of Applied Probability, 32, pp.1077-1088, 1995.

[24] Shiraya, K. and Takahashi, A., Pricing average options on commodities, CIRJE Discussion Papers, 2012.

[25] Turnbull, S.M. and Wakeman, L.M., A quick algorithm for pricing European average options, The Journal of Financial and Quantitative Analysis, 26(3), pp.377-389, 1991.

[26] Vázquez-Abad, F., and Dufresne, D., 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, 1998.

[27] Vercér, J., A new PDE approach for pricing arithmetic average Asian options, Journal of Computational Finance, 4(4), pp.105-113, 2001.

[28] Vercér, J., Black-Scholes representation for Asian options, Mathematical Finance, 24(3), pp.598-626, 2014.

[29] Wang, T.-H. and Gatheral, J., Implied volatility from local volatility: a path integral approach, Springer Proceedings in Mathematics & Statistics, 110, Large Deviations and Asymptotic Methods in Finance, 2015.
LOUIS-PIERRE ARGUIN
DEPARTMENT OF MATHEMATICS
BARUCH COLLEGE, THE CITY UNIVERSITY OF NEW YORK
1 BERNARD BARUCH WAY, NEW YORK, NY10010

E-mail address: louis-pierre.arguin@baruch.cuny.edu

NIE-N-LIN LIU
DEPARTMENT OF MATHEMATICAL SCIENCES
RITSUMEIKAN UNIVERSITY
NOJI-HIGASHI 1-1-1, KUSATSU, SHIGA 525-8577, JAPAN

E-mail address: nienlin1126@gmail.com

TAI-HO WANG
DEPARTMENT OF MATHEMATICS
BARUCH COLLEGE, THE CITY UNIVERSITY OF NEW YORK
1 BERNARD BARUCH WAY, NEW YORK, NY10010

E-mail address: tai-ho.wang@baruch.cuny.edu