A Yosida's parametrix approach to Varadhan's estimates for a degenerate diffusion under the weak Hörmander condition

(Article begins on next page)

18 April 2024

A Yosida's parametrix approach to Varadhan's estimates for a degenerate diffusion under the weak Hörmander condition

Stefano Pagliarani[a,*], Sergio Polidoro[b]

[a]*Dipartimento di Matematica, Università di Bologna, Bologna, Italy.*
[b]*Dipartimento di Scienze Fisiche, Informariche e Matematiche, Università di Modena e Reggio Emilia, Modena, Italy.*

Abstract

We adapt and extend Yosida's parametrix method, originally introduced for the construction of the fundamental solution to a parabolic operator on a Riemannian manifold, to derive Varadhan-type asymptotic estimates for the transition density of a degenerate diffusion under the weak Hörmander condition. This diffusion process, widely studied by Yor in a series of papers, finds direct application in the study of a class of path-dependent financial derivatives known as Asian options. We obtain the Varadhan formula

$$\frac{-2 \log p(t,x;T,y)}{\Psi(t,x;T,y)} \to 1, \qquad \text{as} \quad T - t \to 0^+,$$

where p denotes the transition density and Ψ denotes the optimal cost function of a deterministic control problem associated to the diffusion. We provide a partial proof of this formula, and present numerical evidence to support the validity of an intermediate inequality that is required to complete the proof. We also derive an asymptotic expansion of the cost function Ψ, expressed in terms of elementary functions, which is useful in order to design efficient approximation formulas for the transition density.

Keywords: weak Hörmander condition, asymptotic estimates, parametrix, hypoelliptic diffusion, Asian options
2000 MSC: 60H10, 35A17, 35B40

1. Introduction

The celebrated result obtained by S.R.S Varadhan in [1] can be stated as

$$\frac{-2 \log p(t,x;T,y)}{\Psi(t,x;T,y)} \to 1, \qquad \text{as} \quad T - t \to 0^+, \tag{1.1}$$

where $p(t,x;T,y)$ denotes the transition density of a diffusion process generated by a uniformly elliptic differential operator with Hölder continuous coefficients, and $\Psi(t,x;T,y)$ is the value function, or cost function, of a deterministic control problem associated to the diffusion coefficients. In the framework considered by Varadhan, the value function $\Psi(t,x;T,y)$ agrees with $\frac{d^2(x,y)}{T-t}$, where d denotes the Riemannian distance induced by the diffusion coefficients. This result, which is a cornerstone in the study of the small-time asymptotic behavior of diffusion processes, was generalized by Leandre in [2], [3] to cover the case of hypoelliptic generator satisfying the so-called strong Hörmander condition. In this setting, the Riemannian

*Corresponding author
Email addresses:* `stefano.pagliarani9@unibo.it` (Stefano Pagliarani), `sergio.polidoro@unimore.it` (Sergio Polidoro)

distance needs to be replaced by the (sub-Riemannian) Carnot-Charathéodory distance. However, in the weak hypoelliptic case, i.e. when the generator satisfies the so-called weak Hörmander condition, (1.1) is only known to hold in particular cases, such as constant diffusion and linear drift. In this paper, we move one step forward toward closing this gap in the literature by showing that (1.1) holds true for a two-dimensional degenerate diffusion with variable coefficients, which is weakly hypoelliptic in the aforementioned sense. In this setting, the function $\Psi(t, x; T, y)$ is not related to any sub-Riemannian distance.

We consider the system of Itô stochastic differential equations (SDEs)

$$\begin{cases} dX_s^1 = \sigma X_s^1 dW_s, & X_t^1 = x_1 \\ dX_s^2 = X_s^1 ds, & X_t^2 = x_2 \end{cases}, \tag{1.2}$$

where σ is a positive constant. The study of this equation is motivated by the financial problem of pricing arithmetically averaged Asian options. Indeed, the process $\left(X_s^1\right)_{s \geq t}$ is a geometric Brownian motion starting from x_1 at $s = t$, and describes the evolution of the price of some financial asset in the Black-Scholes setting, while $\left(X_s^2\right)_{s \geq t}$ is the time-integral of $\left(X_s^1\right)_{s \geq t}$ and appears whenever the payoff of the option depends on the past average of the asset price. We refer to Section 1.3 below for a more detailed discussion of the related applications.

Although the explicit expression of the solution to (1.2) is well known, namely

$$\begin{cases} X_s^1 = x_1 e^{\sigma(W_s - W_t) - \frac{\sigma^2}{2}(s-t)} \\ X_s^2 = x_2 + \int_t^s X_\tau^1 d\tau \end{cases}, \qquad s > t, \tag{1.3}$$

the analytical study of its density is more involved. We recall that Yor provided us with the following closed-form expression for the joint density of the process $X = (X^1, X^2)$ with $\sigma = 1$ starting from $(1, 0)$ at time 0:

$$p(s, y_1, y_2) = \frac{e^{\frac{\pi^2}{2s}}}{\pi y_2^2 \sqrt{2\pi s}} \exp\left(-\frac{1 + y_1^2}{2y_2}\right) q\left(s, \frac{y_1}{y_2}\right), \tag{1.4}$$

for $s, y_1, y_2 > 0$, where

$$q(s, \eta) = \int_0^\infty e^{-\frac{\xi^2}{2s}} e^{-\eta \cosh(\xi)} \sinh(\xi) \sin\left(\frac{\pi\xi}{s}\right) d\xi.$$

This expression was first obtained in [4]. We also refer to [5, Theorem 4.1] for a more compact statement and a concise proof. Note that the density p is trivially null for non-positive values of y_1 and y_2, as the process X^1 is strictly positive. Furthermore, an elementary change of variable (see (2.3)-(2.4) below) allows to write the transition density $p(t, x_1, x_2; s, y_1, y_2)$ of the process $X = (X^1, X^2)$ for a general $\sigma > 0$ starting from (x_1, x_2) at time t, which is constantly null if $y_1 \leq 0$ or $y_2 \leq x_2$.

Although (1.4) provides a semi-closed-form expression for the density, the latter is notoriously hard to handle from the numerical and analytical point of view. On the one hand, the numerical integration of the density as written in (1.4) is not accurate when the time s is close to 0, and this results in relevant errors when computing the price of financial options; we refer to Section 1.3 for more details and appropriate references about this problem. On the other hand, it is difficult to extract relevant analytical information, such as asymptotic properties of the density for extreme values of (y_1, y_2) and small values of time s. We undertake the study of these issues by following a different approach, namely we study the transition density as the fundamental solution of the associated backward Kolmogorov ultra-parabolic operator

$$\mathcal{L} = \partial_t + x_1 \partial_{x_2} + \frac{\sigma^2 x_1^2}{2} \partial_{x_1 x_1}, \qquad (t, x_1, x_2) \in \mathbb{R} \times D, \tag{1.5}$$

where we set $D := \mathbb{R}^+ \times \mathbb{R}$, with $\mathbb{R}^+ :=]0, +\infty[$. The operator $\mathcal{L}$ is hypoelliptic in that it can be written in Hörmander form as $\mathcal{L} = Y - \frac{\sigma}{2} Z + \frac{1}{2} Z^2$ with the vector fields

$$Y = \partial_t + x_1 \partial_{x_2}, \qquad Z = \sigma x_1 \partial_{x_1} \tag{1.6}$$

satisfying the Hörmander condition (see Section 1.1). Note that the commutator $[Z, Y]$ is necessary in order to span the space $\mathbb{R}^3$. In the Probability community this is often referred to as *weak Hörmander condition*.

In [6] the authors showed that the transition density of (X^1, X^2) coincides with the fundamental solution $p = p(t, x; T, y)$ for $\mathcal{L}$, and derived upper and lower bounds in terms of the optimal cost function Ψ, defined as the solution of the control problem given by:

$$\Psi(t, x; T, y) = \min_{\omega} \int_t^T |\omega(s)|^2 ds, \tag{1.7}$$

where the minimum is taken over all controls $\omega \in L^2([t, T])$ such that the problem

$$\begin{cases} \dot{\gamma}_1(s) = \sigma \omega(s) \gamma_1(s) \\ \dot{\gamma}_2(s) = \gamma_1(s) \end{cases} \quad t < s < T, \qquad \text{and} \qquad \begin{cases} \gamma_1(t) = x_1, \ \gamma_1(T) = y_1 \\ \gamma_2(t) = x_2, \ \gamma_2(T) = y_2 \end{cases} \tag{1.8}$$

admits a solution. Such a control exists if and only if $x_2 < y_2$. As usual in control theory, we agree to set $\Psi(z; w) := +\infty$ whenever $x_2 > y_2$, in accordance with the fact that $p(t, x; T, y)$ is null.

Roughly speaking, the main results in [6] are an explicit representation for the cost function Ψ and the following upper-lower bounds:

$$\frac{c}{\sigma^2 y_1^2 (T - t)^2} e^{-\frac{C}{2} \Psi(z; w)} \leq p(z; w) \leq \frac{C}{\sigma^2 y_1^2 (T - t)^2} e^{-\frac{c}{2} \Psi(z; w)}$$

for some positive constants $C \gg 1$ and $c \ll 1$, and for $z = (t, x), w = (T, y)$ such that $t < T$ and $x_2 < y_2$. We sharpen such bounds by proving that the exponents are asymptotically equivalent to $\Psi(z; w)/2$ as $T - t$ tends to zero, namely

$$\log p(z; w) \sim -\frac{\Psi(z; w)}{2}, \qquad \text{as} \quad T - t \to 0^+, \tag{1.9}$$

uniformly w.r.t. x, y ranging over compact subsets of D, where we agree to let

$$f \sim g \qquad \text{if} \quad \frac{f}{g} \to 1.$$

For clarity we state two separate results for the upper and lower bounds, respectively. We anticipate that one step of the proof of such estimates relies on some bounds, precisely the Key Inequalities 4.5 below, which we prove numerically in Section 5. We refer to Section 1.2 for a preliminary discussion about the interpretation of such inequalities.

Theorem 1.1 (Upper bound). *If the Key Inequalities 4.5 hold, then for any $\tau > 0$ there exists a positive constant $C > 0$, only dependent on τ and σ, such that*

$$p(z; w) \leq C \frac{\sqrt{\mathbf{h}(z; w)} + \mathbf{h}(z; w)}{\sigma^2 y_1^2 (T - t)^2} e^{-\frac{1}{2} \Psi(z; w)}, \tag{1.10}$$

3

for any $z = (t, x_1, x_2), w = (T, y_1, y_2) \in \mathbb{R} \times D$ such that $T - \tau < t < T$ and $x_2 < y_2$, where

$$\mathbf{h}(z; w) := \frac{(T - t)\sqrt{x_1 y_1}}{y_2 - x_2}, \tag{1.11}$$

and Ψ is the cost function defined as the solution of the control problem above, whose explicit representation is given in (2.7).

Theorem 1.2 (Lower bound). *If the Key Inequalities 4.5 hold, then for any $\varepsilon, \kappa > 0$ there exists $\tau > 0$, only dependent on ε, κ and σ, such that*

$$p(z; w) \geq (1 - \varepsilon) \frac{\mathbf{h}(z; w)}{\sigma^2 y_1^2 (T - t)^2} e^{-\frac{1}{2}\Psi(z; w)}, \tag{1.12}$$

for any $z = (t, x_1, x_2), w = (T, y_1, y_2) \in \mathbb{R} \times D$ such that $T - \tau < t < T$, $x_2 < y_2$ and $\frac{1}{\kappa} \leq \frac{\sqrt{x_1 y_1}}{y_2 - x_2} < \kappa$.

The results above are obtained by adapting to the strictly hypoelliptic setting the Yosida's parametrix method ([7]) for the construction of the fundamental solution to parabolic-type operators. Specifically, we use the cost function Ψ to define a parametrix function. We believe that this method might be of separate interest and we refer to Section 1.2 for further details and related references.

The interest in the results above is duplex. From the theoretical point of view, they provide Varadhan-type estimates for a degenerate operator satisfying a weak Hörmander condition. A detailed discussion about this aspect is deferred to Section 1.1 below. Secondly, knowing the sharp choice of the constant in the exponent paves the way for developing numerically tractable asymptotic expansions of the transition density: we discuss this possible application in Section 1.3.

With regard to the latter point, the computational aspects of the cost function Ψ are also important. In [6] the Authors solved the control problem (1.7)-(1.8) and provided an explicit representation of Ψ, which we report in Section 2. Tough exact, such expression involves the inverse of hyperbolic trigonometric functions. In Section 3 we derive an expansion for the the cost function Ψ whose terms only contain elementary functions, namely

$$\begin{aligned}
\Psi(z; \zeta) = \frac{4}{\sigma^2(T - t)} &\left[\mathbf{h}(z, w) \left(\sqrt{\frac{y_1}{x_1}} + \sqrt{\frac{x_1}{y_1}} - 2 \right) \right. \\
&\left. + \sum_{n=2}^{\infty} a_n \mathbf{1}_{[1,\infty[}(\mathbf{h}(z, w)) \frac{(\mathbf{h}(z, w) - 1)^n}{(\mathbf{h}(z, w))^{n-1}} + b_n \mathbf{1}_{]0,1[}(\mathbf{h}(z, w)) \frac{(-\log \mathbf{h}(z, w))^n}{(1 - \log \mathbf{h}(z, w))^{n-2}} \right],
\end{aligned} \tag{1.13}$$

with $\mathbf{h}(z, w)$ as defined in (1.11), and where the coefficients a_n, b_n can be computed recursively (see Proposition 3.3). A remarkable feature of such expansion is that it seems both point-wise and asymptotically convergent, in the sense that it approximates the asymptotic behavior of $\Psi(z; \zeta)$ as $\mathbf{h}(z, w)$ approaches 0 and $+\infty$. In Section 3 we managed to prove point-wise convergence for $\mathbf{h}(z; \zeta) > 0$ and asymptotic convergence as $\mathbf{h}(z, w) \to +\infty$.

After Sections 1.1, 1.2 and 1.3 below, the remainder of the paper unfolds as follows. Section 2 contains some preliminaries about the cost function Ψ and the fundamental solution of $\mathcal{L}$. In Section 3 we derive the representation (1.13) and provide a partial proof of convergence. In Section 4 we prove Theorems 1.1 and 1.2 by extending Yosida's parametrix method. Section 5 contains the numerical evidence supporting the validity of the Key Inequalities 4.5. Appendix A and Appendix B contain, respectively, a topological lemma needed to prove the results of Section 3 and the proof of Lemma 4.1 appearing in Section 4.

1.1. Varadhan-type estimates

In order to firm our results within the existing literature about fundamental solution, and density, estimates, let us consider the non-divergence-form second-order differential operator

$$\mathcal{H} = \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(t,x) \partial_{x_i x_j} + \sum_{i=1}^{n} \mu_i(t,x) \partial_{x_i} + \partial_t, \qquad (t,x) \in \mathbb{R} \times \tilde{D}, \tag{1.14}$$

with $\tilde{D}$ being a domain of $\mathbb{R}^n$. Clearly, the operator $\mathcal{L}$ in (1.5) is a particular instance of (1.14). From the probabilistic stand-point, under suitable assumptions, $\mathcal{H}$ is the infinitesimal generator of a solution to the following Itô SDE:

$$dX_s = \mu(t, X_s) ds + \sigma(t, X_s) dW_s, \qquad X_t = x, \tag{1.15}$$

with W being a n-dimensional Brownian motion, and σ a $n \times n$ matrix such that $\sigma \sigma^\top = a = (a_{i,j})_{i,j=1,\cdots,n}$. Also, the transition density of X, hereafter denoted by $p = p(t,x;T,y)$, coincides with the fundamental solution of $\mathcal{H}$.

In the uniformly parabolic case, a fairly general classical result (see [8]) states that

$$\frac{c}{(T-t)^{n/2}} e^{-\frac{C|y-x|^2}{2(T-t)}} \leq p(t,x;T,y) \leq \frac{C}{(T-t)^{n/2}} e^{-\frac{c|y-x|^2}{2(T-t)}}, \qquad t < T, \quad x,y \in \mathbb{R}^n, \tag{1.16}$$

where c, C are positive constants independent of $(t,x), (T,y)$. These estimates can be proved under boundedness and Hölder regularity assumptions on the coefficients of $\mathcal{H}$, and assuming a uniform ellipticity condition on the second-order coefficients a_{ij}. More precisely, the upper-bound above can be obtained by means of the so-called *parametrix method*. With the same method it is possible to prove the lower bound in (1.16), but only locally in space: the global version can be achieved with the method of the Harnack chains, introduced by Aronson in [9], which is based on the Harnack inequality. Though very general, the bounds in (1.16) are not sharp enough to characterize the asymptotic behavior for small times of the fundamental solution p away from the pole. In the groundbreaking paper [1] Varadhan proved, for $\mu \equiv 0$ and $a(t,x) = a(x)$, that

$$\lim_{T-t \to 0^+} 2(T-t) \log p(t,x;T,y) = -d^2(x,y), \tag{1.17}$$

uniformly with respect to x, y over a compact subset of $\mathbb{R}^n$, where d represents the geodesic distance with respect to the Riemann metric $a^{-1}(x) dx_i dx_j$, namely

$$d(x,y) = \inf_{\gamma: \gamma(0)=x, \, \gamma(1)=y} \int_0^1 \sqrt{\left\langle a^{-1}\big(\gamma(s)\big)\dot{\gamma}(s), \dot{\gamma}(s)\right\rangle} \, ds. \tag{1.18}$$

The assumptions in [1] are, again, uniform ellipticity and Hölder continuity for the second-order coefficients. Note that (1.17) is more precise than (1.16), in that it yields the exact asymptotic behavior of the logarithm of p for small times. In the subsequent paper [10], building upon (1.17), Varadhan established a large-deviation principle for elliptic diffusions of the form (1.15). The cornerstone contributions [1]-[10] initiated a whole stream of literature, dealing with asymptotic expansions of the transition densities (or heat-kernels in the context of PDEs) of diffusion processes on Riemannian manifolds. Such expansions aim at characterizing the full asymptotic behavior of $p(t,x;T,y)$, under various assumptions on the underlying geometry, by adding up additional terms to the leading one given by (1.17). For instance, following the WKB (Wentzel-Kramers-Brillouin) method, one seeks representations of the type

$$p(t,x;T,y) = \frac{\exp\left(-\frac{d^2(x,y)}{2(T-t)}\right)}{(T-t)^{n/2}} \Big(\alpha_0(x,y) + (T-t)\alpha_1(x,y) + (T-t)^2 \alpha_2(x,y) + \cdots\Big). \tag{1.19}$$

5

We refer, for instance, to [11] and [12] for some relevant contributions to this field, which mainly developed in the years 1970s and 1980s. We also mention the important contribution of Freidlin and Wentzell (e.g. [13]) to the subject of large deviations for diffusion processes. In the last two decades, these asymptotic techniques were employed for the study of volatility models in mathematical finance (see [14] and the references therein).

In the late 1980s these results were generalized to the hypoelliptic setting, under the so-called *strong Hörmander condition*. To explain these generalizations, it is useful to write the operator $\mathcal{H}$ above (with time-independent coefficients) in Hörmander form, namely

$$\mathcal{H} = \partial_t + Z_0 + \frac{1}{2} \sum_{j=1}^{m} Z_j^2, \qquad m \le n, \tag{1.20}$$

where $Z = (Z_0, Z_1, \ldots, Z_m)$ is a system of vector fields defined on a domain $\tilde{D} \subset \mathbb{R}^n$ satisfying the strong Hörmander's condition

$$\mathrm{Lie}(Z_1, \ldots, Z_m)(x) = \mathbb{R}^n, \qquad x \in \tilde{D}, \tag{1.21}$$

where $\mathrm{Lie}(Z_1, \ldots, Z_m)$ is the Lie algebra generated by the vector fields $Z_1, \ldots, Z_m$, namely the vector space spanned by $Z_1, \ldots, Z_m$ and their commutators.

Remark 1.3. The operator $\mathcal{H}$, in the form (1.20), can still be seen as the generator of X in (1.15), with W being now an m-dimensional Brownian motion, and σ a $n \times m$ matrix whose j-th column coincides with the vector field Z_j, namely

$$Z_j = \sum_{i=1}^{n} \sigma_{ij} \partial_{x_i}, \qquad j = 1, \cdots, m.$$

Note that the vector field Z_0 depends on both the drift coefficients μ and σ: precisely, Z_0 contains some terms that cancel out the first-order part stemming from $\frac{1}{2} \sum_{j=1}^{m} Z_j^2$ in (1.20), namely

$$Z_0 = \sum_{i=1}^{n} \mu_i(t,x) \partial_{x_i} - \frac{1}{2} \sum_{j=1}^{m} \sum_{i,k=1}^{n} \sigma_{ij} (\partial_{x_i} \sigma_{kj}) \partial_{x_k}.$$

Under the assumption of the vector fields Z_j being smoth, Leandre ([2], [3]) proved that (1.17) remains valid in this setting, with $d(x,t)$ being the Carnot-Caratheodory *sub-Riemannian* metric induced by the vector fields $Z_1, \ldots, Z_m$. In [15] and [16] also (1.19) was extended to this setting. A contribution in this direction was also given in [17].

In order to discuss extensions to strictly hypoelliptic settings (when (1.21) fails), it is crucial the following

Remark 1.4. Both in the Riemannian and sub-Riemannian case, (1.17) can also be written in the form of (1.9), where Ψ represents here the solution of the optimal control problem given by (1.7), with the minimum taken over all controls $\omega \in L^2([t,T])$ such that the problem

$$\begin{cases} \dot{\gamma}(s) = \omega(s) \sigma(\gamma(s)), & t < s < T, \\ \gamma(t) = x, \ \gamma(T) = y \end{cases} \tag{1.22}$$

admits a solution. Indeed, it is a standard result (see for instance [10, Lemma 2.2] for the elliptic case) that

$$\frac{d^2(x,y)}{T-t} = \Psi(t,x;T,y). \tag{1.23}$$

Under the so-called *weak Hörmander* condition, namely

$$\text{Lie}(\partial_t + Z_0, Z_1, \ldots, Z_m)(x) = \mathbb{R}^{n+1}, \qquad x \in \tilde{D},$$

the hypoellipticity of the operator $\mathcal{H}$ is preserved, but the second-order vector fields $Z_1, \ldots, Z_m$ and their commutators are no longer enough to span the space $\mathbb{R}^n$. Therefore, there is no sub-Riemannian metric on $\mathbb{R}^n$ and an estimate in the form of (1.17) can be no longer achieved. The cost function Ψ, however, remains well defined as the solution of the same control problem (1.7) where

$$\begin{cases} \dot{\gamma}(s) = \mu\big(\gamma(s)\big) + \omega(s)\sigma\big(\gamma(s)\big), & t < s < T, \\ \gamma(t) = x, \ \gamma(T) = y \end{cases} \tag{1.24}$$

replaces (1.22). Therefore, in light of Remark 1.4, (1.9) appears as the natural generalization of (1.17) to strictly hypoelliptic settings. In this sense, Theorems 1.1 and 1.2 above, which lead to (1.9), can be viewed as Varadhan-type estimates for the degenerate parabolic operator $\mathcal{L}$. As already mentioned above, (1.9) sharpens the estimates proved in [6] for the fundamental solution of $\mathcal{L}$. Furthermore, we are not aware of other Varadhan-type formulas in the context of hypoelliptic operators under the weak Hörmander condition, except for the case of constant diffusion and linear drift coefficients (see [18] and [19] among others). In this setting, we also mention the on-diagonal asymptotic estimates in [20].

Note that the drift coefficient μ in (1.24), which is not controlled, is needed to ensure the existence of a path connecting x and y. Also, the fact that (1.23) does not hold in general is evident as $\Psi(t, x; T, y)$ can exhibit different rates of explosion, as $T - t \to 0^+$, depending on the choice of (x, y). This is a well-known phenomenom, which reflects the different time-scales of the single components of the underlying diffusion. A stylized example is given by the stochastic Langevin equation

$$\begin{cases} dX_s^1 = \sigma dW_s, & X_t^1 = x_1 \\ dX_s^2 = X_s^1 ds, & X_t^2 = x_2 \end{cases} \tag{1.25}$$

with σ positive constant, whose generator is given by

$$\mathcal{H} = \partial_t + x_1 \partial_{x_2} + \frac{a}{2} \partial_{x_1 x_1}, \qquad a = \sigma^2. \tag{1.26}$$

Its fundamental solution is given exactly by

$$p_\sigma(t, x; T, y) = \frac{1}{2\pi\sqrt{\det \mathbf{C}(\sigma, T - t)}} \exp\left(-\frac{1}{2}\Psi_\sigma(t, x; T, y)\right), \tag{1.27}$$

$$\mathbf{C}(\sigma, s) := \sigma^2 \begin{pmatrix} s & -\frac{s^2}{2} \\ -\frac{s^2}{2} & \frac{s^3}{3} \end{pmatrix}, \tag{1.28}$$

where

$$\Psi_\sigma(t, x; T, y) = \left\langle \mathbf{C}^{-1}(\sigma, T - t)\big(x_1 - y_1, x_2 - y_2 + (T - t)y_1\big), \big(x_1 - y_1, x_2 - y_2 + (T - t)y_1\big) \right\rangle \tag{1.29}$$

$$= \frac{3\big(2(y_2 - x_2) - (T - t)(x_1 + y_1)\big)^2}{\sigma^2(T - t)^3} + \frac{(y_1 - x_1)^2}{\sigma^2(T - t)}$$

is the cost function of the associated control problem (see for instance [19, Example 9.53]).

1.2. Yosida's parametrix

In order to obtain the bounds in Theorems 1.1 and 1.2, we adapt and extend the method introduced by Yosida in his seminal paper [7], where he outlined a geometrical variation of the classical *Levi's parametrix method* for the construction of the fundamental solution of a parabolic-type operator $\mathcal{H}$ of type (1.14) on a Riemannian manifold. Such method, which was published 14 years before Varadhan's paper [1], already contained the idea that an expansion like (1.19) should hold, and in particular that the logarithm of the fundamental solution p should asymptotically behave like in (1.17). In spite of this, Yosida's method did not become a standard tool in the study of the asymptotic properties of transition densities. We suspect this is partially due the fact that his construction was mainly heuristic: no precise assumptions on the coefficients were given to ensure convergence, nor upper/lower bounds leading to (1.17) were proved in [7]. In [21], Gatheral et al. brought Yosida's method to the attention of the mathematical finance community emphasizing its versatility from the computational point of view, with a particular focus on its ability to deal with time-dependent coefficients. In this paper, inspired by [21], we push Yosida's method one step further and adapt it to the strictly hypoelliptic setting. In Section 4 we perform the construction of the fundamental solution p for the operator $\mathcal{L}$ in (1.5), and prove the estimates (1.10)-(1.12) which lead to (1.9). Although we treat here a special case, we claim that the principles of this approach can be employed in general for a wider class of degenerate parabolic operators under the weak Hörmander condition, to establish full expansions analogous to (1.19). The study of general assumptions on the coefficients, possibly time-dependent, under which this method yields rigorous results is subject of ongoing investigation.

Consider the parabolic operator $\mathcal{H}$ in (1.14), and assume for simplicity $\mu \equiv 0$ and $a_{ij}(t,x) = a(x)$. In the Yosida's parametrix, the key difference with respect to Levi's original method lies in the choice of the *parametrix function* (the starting point of the iterative construction), which is dependent on the geodesic distance d induced by the second-order coefficients of the differential operator. Precisely, in Yosida's method the kernel that defines the parametrix function is set as

$$\frac{(T-t)^{-n/2}}{\sqrt{(2\pi)^n \det a(y)}} \exp\left(-\frac{1}{2}\frac{d^2(x,y)}{T-t}\right), \tag{1.30}$$

where d is the Riemannian distance (1.18) induced by the coefficients a_{ij}. By opposite, in Levi's method the parametrix function is the fundamental solution of the constant coefficients operator obtained from $\mathcal{H}$ by freezing the coefficients at y. In other words, the distance d in (1.30) is replaced by the geodesic distance taken with respect to the constant metric tensor $a^{-1}(y)dx_i dx_j$, namely the kernel is

$$\frac{(T-t)^{-n/2}}{\sqrt{(2\pi)^n \det a(y)}} \exp\left(-\frac{1}{2}\frac{d_y^2(x,y)}{T-t}\right), \qquad \text{where} \quad d_y^2(x,y) = \langle a^{-1}(y)(y-x), y-x \rangle.$$

The idea is that the choice (1.30) provides us with sharp asymptotic estimates of the fundamental solution for small times, while preserving the correct behavior of the kernel near the singularity (the starting point of the stochastic process). As already mentioned in the previous subsection, the second-order coefficients of the operator $\mathcal{L}$ do not induce a distance on $\mathbb{R}^2$. However, proceeding in analogy with (1.30)-(1.23), our idea is to consider the following kernel for the parametrix function:

$$H_1(t,x;T,y) := \frac{1}{2\pi\sqrt{\det \mathbf{C}(\sigma y_1, T-t)}} \exp\left(-\frac{1}{2}\Psi(t,x;T,y)\right), \tag{1.31}$$

where the matrix $\mathbf{C}$ is set as in (1.28) in order to yield the right behavior near the singularity. In particular, the choice for $\mathbf{C}$ can be explained as follows. First note that the integral curve of the vector field Y in (1.6),

8

starting at $z = (t, x) \in \mathbb{R} \times D$, is given by

$$e^{\delta Y} z = (t + \delta, x_1, x_2 + \delta x_1),$$

which yields (just set $\omega \equiv 0$ in (1.8))

$$\Psi\big(t, y_1, y_2 - (T - t)y_1; T, y_1, y_2\big) = 0, \qquad T - t > 0, \ y \in D.$$

As it turns out, expanding now $\Psi(t, \cdot; T, y)$ around $(y_1, y_2 - (T - t)y_1)$, at second order, one obtains

$$\Psi(t, x; T, y) = \Psi_{\sigma y_1}(t, x; T, y) + o\big(|x_1 - y_1|^2 + |x_2 - y_2 + (T - t)y_1|^2\big), \qquad \text{as } (x_1, x_2) \to (y_1, y_2 - (T - t)y_1),$$

where $\Psi_{\sigma y_1}(t, \cdot; T, y)$ is the quadratic form in (1.29). Therefore, the term

$$1/\sqrt{\det \mathbf{C}(\sigma y_1, T - t)}$$

in (1.31) ensures a Gaussian behavior for H_1 near the diagonal. In particular, one can show (see Proposition 4.4) that $H(t, x; T, y) \to \delta_x$ as $t \to T^-$. Note that, replacing Ψ with $\Psi_{\sigma y_1}$ in (1.31), the function H_1 becomes the fundamental solution of the operator $\mathcal{H}$ in (1.26) with $a = \sigma^2 y_1^2$. Such a kernel can be seen as the counterpart of Levi's parametrix kernel for degenerate (1.26)-like operators with variable coefficients, and allows to prove (see [22], [23], [24]) Gaussian bounds for the fundamental solution when a is a positive uniformly Hölder-continuous function, which is bonded and bounded away from zero.

The parametrix method is an iterative procedure to constructs the fundamental solution taking a parametrix function as a leading term, and then representing the remainder as a series whose terms can be recursively determined via successive approximations based on Duhamel's principle. A key element to prove the convergence of the series consists in estimating the convolution of the parametrix function with itself. In the original Levi's method for parabolic operators, the parametrix is a Gaussian function and thus one can rely on the Chapman-Kolmogorov equation.

In our case, the bounds we need in order to complete the parametrix construction are those in the Key Inequalities 4.5, which are estimates that look roughly like

$$\int_{\mathbb{R}^+} \int_{x_2}^{y_2} H_1(t, x; s, \xi) H_1(s, \xi; T, y) d\xi \leq C_{T-t} H_1(t, x; T, y), \qquad s \in]t, T[,$$

uniformly in $x, y \in]0, +\infty[\times \mathbb{R}$ with $x_2 < y_2$. The fact that this bound holds locally in x, y is clear as H_1 tends to a Dirac delta for small times. At the moment we only have numerical evidence (reported in Section 5) for the fact that it holds uniformly. The techniques used in the Riemannian and sub-Riemannian case to prove the Chapman-Kolmogorov identity seem to fail here. Therefore, although the numerical evidence reported in Section 5 strongly supports the validity of such estimates, the problem of providing a rigorous proof remains an open problem.

1.3. Application to arithmetic Asian options

The Itô process in (1.3) finds a direct application in the problem of pricing and hedging of a class of path-dependent financial derivatives known as arithmetic Asian options. The first component X^1, which is a geometric Brownian motion, denotes the evolution price of a risky asset in the well-known Black-Scholes model under the risk-neutral probability measure. For simplicity, it is assumed here zero risk-free interest rate so that the risk-neutral dynamics of the asset is exactly given by (1.2), namely X^1 is an exponential

Brownian martingale. The process $s \mapsto \frac{1}{s}X_s^2$ represents instead the continuous arithmetic time-average of the risky asset X^1. An arithmetic average Asian option is a random variable of the form

$$\varphi\left(X_T^1, \frac{1}{T}X_T^2\right),$$

where φ is a payoff function that determines the value of the financial claim at a given maturity $T > 0$. A variety of payoff functions can be considered, some popular choices being for example

$$\varphi(x_1, x_2) = (x_1 - x_2)^+, \qquad\qquad \text{(floating-strike Call)}$$
$$\varphi(x_1, x_2) = (x_2 - K)^+, \quad K > 0. \qquad\qquad \text{(fixed-strike Call)}$$

Within the theory of continuous-time arbitrage pricing, the no-arbitrage price at time $t < T$ of the such derivatives, given the initial values $x = (x_1, x_2)$ for $X = (X^1, X^2)$, are given by the risk-neutral evaluation

$$\mathbb{E}_{t,x}\left[\varphi\left(X_T^1, X_T^2/T\right)\right] = \int_{\mathbb{R}_{>0}} \int_{x_2}^{+\infty} p(t, x_1, x_2; T, y_1, y_2)\varphi\left(y_1, y_2/T\right) dy_2 dy_1. \tag{1.32}$$

While the expected value above determines the price of the claim, its derivatives with respect to the variables t, x, T and the parameter σ, also known as sensitivities, are related to the hedging strategies of such claim. In the evaluation of (1.32) one encounters computational difficulties due the involved expression of the transition density p. Intuitively, this is largely due to the fact that the problem is somehow *ill-posed*, meaning that X^2 is an arithmetic average of a geometric Brownian motion. For instance, the integral representation (1.4) given by Yor is of limited practical use in the numerical computation of (1.32).

In the last decades, the pursue of the efficient methods to approximate the integral in (1.32) has challenged many authors, of whom we give here an incomplete account. In [25], Geman and Yor gave an explicit representation of Asian option prices in terms of the Laplace transform of hypergeometric functions. However, several authors (see for instance [26] and [27]) pointed out the difficulty of pricing Asian options with short maturities or small volatilities using the analytical method in [25]. This is also a disadvantage of the Laguerre expansion proposed in [28]. In [29] a contour integral approach was employed to improve the accuracy in the case of low volatilities, though at a higher computational cost. In [30] the problem was tackled using the spectral theory of singular Sturm-Liouville operators, yielding a series formula that gives very accurate results. However, again for low volatilities, the convergence might be slow and the method becomes computationally expensive. We also mention the Monte Carlo approach, which was taken by a number of authors to price efficiently Asian options under the Black-Scholes model (see [31] among others). In the Black-Scholes model and for special homogeneous payoff functions, it is possible to reduce the study of Asian options to a PDE with only one state variable. We refer to this approach as to PDE reduction, which was followed in [32], [33] and [34] among other works.

The results of this paper pave the road for deriving sharp asymptotic expansions for prices and sensitivities through an expansion of the transition density p of the form

$$p(t, x; T, y) = H_1(t, x; T, y)\mathbf{u}(t, x; T, y)\Big(1 + (T - t)\alpha_1(t, x; T, y) + (T - t)^2\alpha_2(t, x; T, y) + \cdots\Big), \tag{1.33}$$

with H_1 as in (1.31), $\mathbf{u}$ as determined in Proposition 4.3, and where the further correction terms $\alpha_1, \alpha_2, \cdots$ have to be determined through recursive procedure as in [7]. In order to obtain approximations of (1.32) that are computationally tractable, the expansion (1.13) of Ψ in terms of elementary functions, which we derive in Section 3, plays an essential role as it allows to avoid the numerical inversion of the hyperbolic

10

trigonometric functions that appear in the closed-form representation (2.7). The expansion (1.33) can be seen as a refinement of the asymptotic expansions derived in [35], [36], which are based on a perturbation procedure that approximates, at leading order, the Black-Scholes dynamics (1.2) with some Langevin dynamics like in (1.25). The result of this procedure is an expansion for p of the form (1.33) whose leading term is a Gaussian transition density as in (1.27). A similar approach was taken in [37] to obtain analytical approximations by means of Malliavin calculus.

In general, analytical approaches based on asymptotic expansions exhibit several advantages in that they provide approximations in closed form, which are fast to compute and display an explicit dependence on the parameters. Other asymptotic methods for Asian options were studied in [38] and [39] by Malliavin calculus techniques. Within this stream of literature we also mention the results in [40], where sharp asymptotic expansions for *fixed* and *floating-strike* Asian options are derived. The latter approach is related to ours, as it exploits Varadhan's principle of large deviation ([10]), together with a contraction principle to deal with the arithmetic average.

2. Preliminaries

Throughout the paper we will use the following

Notation 2.1. Let $x = (x_1, x_2), y = (y_1, y_2) \in D$, we write

$$x \prec y$$

if $x_2 < y_2$. Furthermore, given $z = (t, x), w = (T, y) \in \mathbb{R} \times D$, we write

$$z \prec w$$

if $x \prec y$ and $t < T$.

Definition 2.2. A fundamental solution for the operator $\mathcal{L}$ is a function $p(z, w)$ defined for any $z, w \in \mathbb{R} \times D$ with $z \prec w$ such that:

i) for any $w \in \mathbb{R} \times D$, the function $p(\cdot; w)$ solves

$$\mathcal{L}u(z) = 0, \qquad z \in \mathbb{R} \times D, \ z \prec w, \tag{2.1}$$

in the classical sense;

ii) for any $z = (T, x) \in \mathbb{R} \times D$, we have that $p(t, x; T, \cdot) \to \delta_z$ as $t \to T^-$ in the following sense:

$$\lim_{\substack{(t,x') \to z \\ t < T}} \int_{x' \prec y} p(t, x'; T, y)\varphi(y)dy = \varphi(x), \qquad \varphi \in C_b(D).$$

Remark 2.3. Given a fundamental solution $p = p(z, w)$ for $\mathcal{L}$, one can consider its zero extension to the whole space $D \times D$ minus the diagonal $\{z = w\}$. This makes sense if we interpret the function $p(t, x; T, \cdot)$ as the transition density of the process X_T starting at time t from the point x. Indeed the second component X^2 of the process is strictly increasing due to the fact that the first component X^1 is a geometric Brownian motion, and as such it strictly positive.

In [41] it was first observed that $\mathcal{L}$ is left-invariant with respect to the non-commutative group law

$$z \circ w = (t + T, x_1 y_1, x_2 + y_2 x_1), \qquad z = (t, x), w = (T, y) \in \mathbb{R} \times D.$$

Precisely:

$$\mathcal{L}u(z) = 0 \Leftrightarrow \mathcal{L}u^w(z) = 0, \qquad \text{with } u^w(z) := u(w \circ z). \tag{2.2}$$

As it turns out, $\mathbb{G} = (\mathbb{R} \times D, \circ)$ is a Lie Group with identity and inverse given by

$$\mathbf{1}_{\mathbb{G}} = (0, 1, 0), \qquad z^{-1} = \left(-t, x_1^{-1}, -x_2 x_1^{-1} \right),$$

respectively. By the left-invariance of $\mathcal{L}$, it is straightforward to see that

$$y_1^2 p(z; w) = p(w^{-1} \circ z; \mathbf{1}_{\mathbb{G}}) = p\left(t - T, \frac{x_1}{y_1}, \frac{x_2 - y_2}{y_1}; \mathbf{1}_{\mathbb{G}} \right), \tag{2.3}$$

for any $z = (t, x), w = (T, y) \in \mathbb{R} \times D$ with $z \prec w$. Also, denoting by $\bar{p}$ the fundamental solution when $\sigma = 1$, it is easy to check that

$$p(z; \mathbf{1}_{\mathbb{G}}) = \sigma^2 \bar{p}(\sigma^2 t, x_1, \sigma^2 x_2; \mathbf{1}_{\mathbb{G}}), \tag{2.4}$$

for any $z = (t, x) \in \mathbb{R} \times D$ with $z \prec \mathbf{1}_{\mathbb{G}}$.

The following result was proved in [6].

Theorem 2.4. *The operator $\mathcal{L}$ has a unique fundamental solution, which is positive, smooth (C^∞), and for which the following upper/lower bounds hold: for any arbitrary $\varepsilon \in]0, 1[$ and $T_0 > 0$, there exist two positive constants $c_\varepsilon^-, C_\varepsilon^+$ depending on ε and T_0, and two positive universal constants C^-, c^+ such that*

$$\frac{c_\varepsilon^-}{\sigma^2 y_1^2 (T - t)^2} \exp\left(-C^- \Psi\left(t + \varepsilon(T - t), x_1, x_2 + y_1 \varepsilon(T - t); T, y_1, y_2 \right) \right) \leq$$

$$\Gamma(t, x_1, x_2; T, y_1, y_2) \leq$$

$$\frac{C_\varepsilon^+}{\sigma^2 y_1^2 (T - t)^2} \exp\left(-c^+ \Psi\left(t - \varepsilon, x_1, x_2 - y_1 \varepsilon; T, y_1, y_2 \right) \right),$$

for every $(t, x_1, x_2), (T, y_1, y_2) \in \mathbb{R}^+ \times D$ such that $(t + \varepsilon(T - t), x_1, x_2 + y_1 \varepsilon(T - t)) \prec (T, y_1, y_2)$ and $T - t < T_0$.

We recall here the explicit representation for the cost function Ψ, as it was computed in [6]. For given $z = (t, x), w = (T, y) \in \mathbb{R} \times D$ with $z \prec w$, set

$$E = E_{z,w} := \frac{4}{(T - t)^2} g^{-1}\left(\frac{1}{\mathbf{h}(z; w)} \right), \tag{2.5}$$

$$\varsigma = \varsigma_{z,w} := \text{sgn}\left(E_{z,w} + \frac{\pi^2}{(T - t)^2} \right) = \text{sgn}\left(4g^{-1}\left(\frac{1}{\mathbf{h}(z; w)} \right) + \pi^2 \right),$$

with

$$g(r) = \frac{\sinh(\sqrt{r})}{\sqrt{r}} = \begin{cases} \frac{\sinh(\sqrt{r})}{\sqrt{r}}, & r > 0 \\ 1, & r = 0 \\ \frac{\sin(\sqrt{-r})}{\sqrt{-r}}, & -\pi^2 < r < 0 \end{cases} \tag{2.6}$$

and $\mathbf{h}$ as defined in (1.11). Sometimes, here and throughout the paper, the notation E will be preferred to $E_{z,w}$ when the dependence on (z, w) is clear from the context. The optimal cost is then given by

$$\sigma^2 \Psi(z; w) = E_{z,w}(T - t) + \frac{4(x_1 + y_1)}{y_2 - x_2} - 4\varsigma_{z,w}\sqrt{E + \frac{4x_1 y_1}{(y_2 - x_2)^2}}, \tag{2.7}$$

or equivalently

$$\Psi(z; w) = \frac{4}{\sigma^2(T - t)}\left[g^{-1}\left(\frac{1}{\mathbf{h}(z; w)}\right) + \mathbf{h}(z; w)\left(\sqrt{\frac{x_1}{y_1}} + \sqrt{\frac{y_1}{x_1}}\right) - 2\varsigma_{z,w}\sqrt{g^{-1}\left(\frac{1}{\mathbf{h}(z; w)}\right) + \mathbf{h}^2(z; w)}\right].$$

Remark 2.5. It is straightforward to check that

$$\mathbf{h}(z; w) = \mathbf{h}(w^{-1} \circ z; \mathbf{1}_{\mathbb{G}}) = \mathbf{h}\left(t - T, \frac{x_1}{y_1}, \frac{x_2 - y_2}{y_1}; 0, 1, 0\right) \tag{2.8}$$

$$= \mathbf{h}\left(0, 1, 0; T - t, \frac{y_1}{x_1}, \frac{y_2 - x_2}{x_1}\right) = \mathbf{h}(\mathbf{1}_{\mathbb{G}}; z^{-1} \circ w),$$

and thus

$$\Psi(z, w) = \Psi(w^{-1} \circ z; \mathbf{1}_{\mathbb{G}}) = \Psi\left(t - T, \frac{x_1}{y_1}, \frac{x_2 - y_2}{y_1}; 0, 1, 0\right) \tag{2.9}$$

$$= \Psi\left(0, 1, 0; T - t, \frac{y_1}{x_1}, \frac{y_2 - x_2}{x_1}\right) = \Psi(\mathbf{1}_{\mathbb{G}}; z^{-1} \circ w).$$

This, together with (2.3) and (2.4), implies that the estimates of Theorems 1.1 and 1.2 only need to be proved for $w = \mathbf{1}_{\mathbb{G}}$ and $\sigma = 1$.

3. A novel representation for the cost function

In this section we present a convergent expansion for the cost function Ψ, which allows to write the latter in terms of elementary functions, and preserves its asymptotic properties as $\mathbf{h}(z; w)$ tends to zero and infinity. We start off by re-writing the cost function as

$$\Psi(z, w) = \frac{4}{\sigma^2(T - t)}\left[\mathbf{h}(z, w)\left(\sqrt{\frac{y_1}{x_1}} + \sqrt{\frac{x_1}{y_1}} - 2\right) + G\big(\mathbf{h}(z, w)\big)\right], \tag{3.1}$$

where we set

$$G(\eta) := 2\eta - 2\text{sgn}\big(4g^{-1}(\eta^{-1}) + \pi^2\big)\sqrt{\eta^2 + g^{-1}(\eta^{-1})} + g^{-1}(\eta^{-1}) \tag{3.2}$$

$$= 2\eta - 2\left(\eta^{-1} - \frac{2}{\pi}\right)\sqrt{\big(\eta^2 + g^{-1}(\eta^{-1})\big)\left(\eta^{-1} - \frac{2}{\pi}\right)^{-2}} + g^{-1}\left(\eta^{-1}\right), \qquad \eta > 0.$$

The last equality above stems from the following relation

$$\text{sgn}\big(4g^{-1}(\eta^{-1}) + \pi^2\big) = \begin{cases} 1, & \text{if } \eta < \frac{\pi}{2} \\ 0, & \text{if } \eta = \frac{\pi}{2} \\ -1, & \text{if } \eta > \frac{\pi}{2} \end{cases}. \tag{3.3}$$

The representation (3.1) for the cost function is particularly meaningful as the function G is strictly convex on $\mathbb{R}^+$ and has global minimum
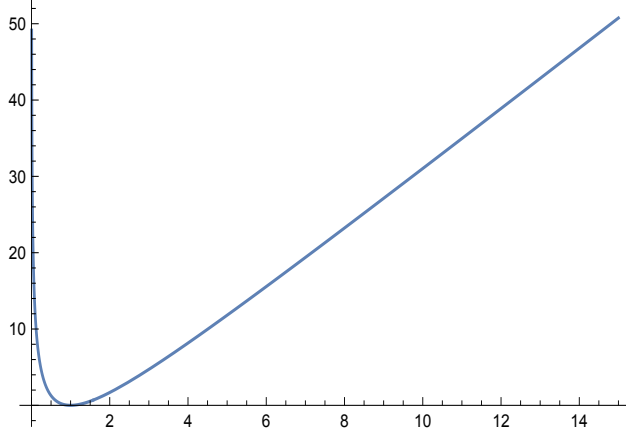
$$\min G = G(1) = 0.$$

Figure 1: Plot of the function G in (3.2).

In particular, $G(\eta)$ is strictly increasing for $\eta > 1$ and strictly decreasing for $\eta \in]0, 1[$ (see Figure 3). Therefore, the function Ψ appears as a strictly increasing function of the non-negative quantities

$$\mathbf{h}(z, w), \qquad \sqrt{\frac{y_1}{x_1}} + \sqrt{\frac{x_1}{y_1}} - 2,$$

with zero minimum at

$$\left(\mathbf{h}(z, w), \sqrt{\frac{y_1}{x_1}} + \sqrt{\frac{x_1}{y_1}} - 2 \right) = (1, 0). \tag{3.4}$$

Remark 3.1. Solving (3.4) for $w = (T, y)$, we obtain

$$\begin{cases} y_1 = x_1, \\ y_2 = x_2 + (T - t)x_1, \end{cases}$$

which is coherent with the control problem (1.7)-(1.8). Indeed, a solution to (1.8) for $y = (y_1, y_2)$ as above is given by $\omega \equiv 0$, which clearly yields $\Psi(t, x; T, y) = 0$.

The computation of the function G in (3.2) requires the numerical inversion of trigonometric functions. This makes the use of Ψ numerically challenging, especially in regards to the computation of numerical integrals of the type

$$\int_{x \prec y} e^{-\frac{\Psi(t, x; T, y)}{2}} f(y) dy,$$

which are in turn related to the evaluation of arithmetic Asian options (see (1.32)). For this reason, we look for an expansion of the function G that makes the numerical computation of Ψ more manageable. We would also like such expansion to preserve the asymptotic properties of G as η tends to 0^+ and $+\infty$.

We introduce the following

Definition 3.2. For any $N \in \mathbb{N}$ with $N \geq 2$, we define the *N-th order expansion of* Ψ as

$$\Psi_N(z, w) := \frac{4}{\sigma^2(T - t)} \left[\mathbf{h}(z, w) \left(\sqrt{\frac{y_1}{x_1}} + \sqrt{\frac{x_1}{y_1}} - 2 \right) + \sum_{n=2}^{N} G_n\big(\mathbf{h}(z, w)\big) \right], \qquad N \geq 2,$$

where

$$G_n(\eta) := a_n \mathbf{1}_{]1, \infty[}(\eta) \frac{(\eta - 1)^n}{\eta^{n-1}} + b_n \mathbf{1}_{]0, 1[}(\eta) \frac{(-\log \eta)^n}{(1 - \log \eta)^{n-2}}, \tag{3.5}$$

14

and where the coefficients a_n, b_n are recursively determined by solving the equations

$$\frac{d^n}{d\eta^n}G(\eta)\Big|_{\eta=1} = \frac{d^n}{d\eta^n}\sum_{k=2}^{n}G_k(\eta)\Big|_{\eta=1}, \qquad n \geq 1. \tag{3.6}$$

Before focusing on the convergence results and the asymptotic properties of the approximation Ψ_N, we discuss some aspects related to its computation. We first point out that the functions G_n are written in terms of elementary functions. Therefore, the numerical evaluation of Ψ_N is straightforward save computing the coefficients a_n, b_n in (3.5). In regard to this matter, the characterization of the coefficients in Proposition 3.3 below comes in handy. For the sake of completeness, the first 38 coefficients a_n, b_n are reported in Table 1.

Proposition 3.3. *For any $n \in \mathbb{N}$ with $n \geq 2$, the coefficients a_n, b_n in (3.5) are equal to*

$$a_n = (-1)^n\left(\beta_n + \beta_{n-1} - \frac{2}{n!}\sum_{h=0}^{n}\frac{(-1)^h(2h)!}{(1-2h)(h!)4^h}B_{n,h}\big(c_1, 2!c_2, \cdots, (n-h+1)!c_{n-h+1}\big)\right), \tag{3.7}$$

$$b_n = d_{n-2} - 2d_{n-1} + d_n + 2(e_{n-2} - 2e_{n-1} + e_n - f_{n-2} + 2f_{n-1} - f_n), \tag{3.8}$$

with

$$c_n = \beta_n + 2\beta_{n-1} + \beta_{n-2},$$

$$d_n = \frac{1}{n!}\sum_{h=1}^{n}L(n,h)\sum_{k=1}^{h}\beta_k k!\begin{Bmatrix}h\\k\end{Bmatrix}, \tag{3.9}$$

$$e_n = \frac{1}{n!}\sum_{h=1}^{n}(-1)^h L(n,h), \qquad \tilde{e}_n = \frac{1}{n!}\sum_{h=1}^{n}(-2)^h L(n,h), \tag{3.10}$$

$$f_n = \frac{1}{n!}\sum_{h=1}^{n}\frac{(-1)^h(2h)!}{(1-2h)(h!)4^h}B_{n,h}\big(d_1 + \tilde{e}_1, 2!(d_2 + \tilde{e}_2), \cdots, (n-h+1)!(d_{n-h+1} + \tilde{e}_{n-h+1})\big),$$

and

$$\beta_{-1} = \beta_0 = 0, \qquad \beta_1 = 6, \qquad \beta_k = -\frac{6^k}{k!}\sum_{h=1}^{k-1}h!b_h B_{k,h}\left(\frac{1}{3!}, \frac{2!}{5!}, \cdots, \frac{(k-h+1)!}{[2(k-h+1)+1]!}\right), \; k \geq 2. \tag{3.11}$$

Above, the functions $B_{k,h}$ represent the exponential Bell polynomials, $L(n,h)$ denote the Lah numbers, and $\begin{Bmatrix}h\\k\end{Bmatrix}$ the Stirling numbers of the second kind.

Proof. We first prove (3.7). By the change of variable $\xi := \frac{\eta-1}{\eta}$, we obtain that (3.6) is equivalent to

$$\frac{1}{n!}\frac{d^n}{d\xi^n}F(\xi)\Big|_{\xi=0} = \frac{1}{n!}\frac{d^n}{d\xi^n}\sum_{k=2}^{n}a_k\xi^k\Big|_{\xi=0} = a_n, \qquad n \geq 2, \tag{3.12}$$

where

$$F(\xi) := (1-\xi)G\left(\frac{1}{1-\xi}\right) = 2 - 2\sqrt{1 + (1-\xi)^2 g^{-1}(1-\xi)} + (1-\xi)g^{-1}(1-\xi), \qquad |\xi| << 1.$$

We now go on to compute the power series of F at $\xi = 0$. Denoting by $(\beta_n)_{n\in\mathbb{N}_0}$ the coefficients of the powers series of g^{-1} in 1, a direct application of Faà di Bruno formula yields

$$\beta_0 = 0,$$

15

n	a_n	b_n
2	3	3
3	0.6	-0.6
4	0.205714	-0.444286
5	0.0891429	-0.322
6	0.044144	-0.227237
7	0.0238419	-0.15493
8	0.0136883	-0.100762
9	0.00822413	-0.0610761
10	0.00511786	-0.0328045
11	0.00327527	-0.0133935
12	0.00214449	-0.00074011
13	0.00143104	0.00686845
14	0.0009704	0.0108089
15	0.000667139	0.0121717
16	0.00046414	0.0118058
17	0.000326287	0.0103592
18	0.000231492	0.00831401
19	0.000165581	0.00601859
20	0.000119303	0.00371439
21	0.0000865255	0.00155955
22	0.0000631266	-0.000351322
23	0.0000463045	-0.00197063
24	0.0000341329	-0.00328434
25	0.0000252745	-0.00430119
26	0.000018793	-0.00504427
27	0.0000140273	-0.00554463
28	0.0000105073	-0.0058366
29	$7.89662 * 10^-6$	-0.00595453
30	$5.95282 * 10^-6$	-0.00593069
31	$4.50037 * 10^-6$	-0.00579403
32	$3.41143 * 10^-6$	-0.00556962
33	$2.59248 * 10^-6$	-0.00527856
34	$1.97478 * 10^-6$	-0.00493825
35	$1.5076 * 10^-6$	-0.00456279
36	$1.15335 * 10^-6$	-0.00416351
37	$8.84092 * 10^-7$	-0.00374953
38	$6.78961 * 10^-7$	-0.00332826

Table 1: Values of the coefficients a_n, b_n up to $n = 38$.

$$\beta_1 = \big(g'(0)\big)^{-1},$$

$$n!\beta_n = -\big(g'(0)\big)^{-n}\sum_{h=1}^{n-1}h!\beta_h B_{n,h}\big(g'(0), g''(0), \cdots, g^{(n-h+1)}(0)\big), \qquad n \geq 2,$$

where $B_{n,h}$ represent the exponential Bell polynomials. Note now that, by (3.22), we obtain

$$g^{(n)}(0) = \frac{n!}{(2n+1)!}, \qquad n \in \mathbb{N}_0,$$

which yields (3.11). Thus we obtain

$$(1-\xi)g^{-1}(1-\xi) = \sum_{n=1}^{\infty}(-1)^n(\beta_n + \beta_{n-1})\xi^n,$$

$$(1-\xi)^2 g^{-1}(1-\xi) = \sum_{n=1}^{\infty}(\beta_n + 2\beta_{n-1} + \beta_{n-2})\xi^n,$$

for any ξ close to 0. Eventually, applying once more Faà di Bruno formula, together with

$$\sqrt{1+x} = 1 + \sum_{n=1}^{\infty}\frac{(-1)^n(2n)!}{(1-2n)(n!)^2 4^n}x^n, \qquad |x| << 1, \tag{3.13}$$

yield (3.12) with a_n as given by (3.7).

To prove (3.8) we use an analogous argument. By the change of variable $\xi := \frac{-\log\eta}{1-\log\eta}$, we obtain that (3.6) is equivalent to

$$\frac{1}{n!}\frac{d^n}{d\xi^n}F(\xi)\Big|_{\xi=0} = \frac{1}{n!}\frac{d^n}{d\xi^n}\sum_{k=2}^{n}b_k\xi^k\Big|_{\xi=0} = b_n, \qquad n \geq 2, \tag{3.14}$$

where

$$F(\xi) := (1-\xi)^2 G\big(e^{\frac{\xi}{\xi-1}}\big) = (1-\xi)^2\Big(2e^{\frac{\xi}{\xi-1}} - 2\sqrt{e^{\frac{2\xi}{\xi-1}} + g^{-1}\big(e^{\frac{\xi}{1-\xi}}\big)} + g^{-1}\big(e^{\frac{\xi}{1-\xi}}\big)\Big), \qquad |\xi| << 1.$$

By applying now Faà di Bruno formula we obtain

$$e^{\frac{\xi}{\xi-1}} - 1 = \sum_{n=1}^{\infty}\frac{\xi^n}{n!}\sum_{h=1}^{n}B_{n,h}\big(-1, -2!, \cdots, -(n-h+1)!\big) = \sum_{n=1}^{\infty}\frac{\xi^n}{n!}\sum_{h=1}^{n}(-1)^h\binom{n-1}{h-1}\frac{n!}{h!} = \sum_{n=1}^{\infty}e_n\xi^n,$$

$$e^{\frac{2\xi}{\xi-1}} - 1 = \sum_{n=1}^{\infty}\frac{\xi^n}{n!}\sum_{h=1}^{n}B_{n,h}\big(-2, -2\cdot 2!, \cdots, -2\cdot(n-h+1)!\big) = \sum_{n=1}^{\infty}\tilde{e}_n\xi^n,$$

$$g^{-1}\big(e^{\frac{\xi}{1-\xi}}\big) = \sum_{n=1}^{\infty}\frac{\xi^n}{n!}\sum_{h=1}^{n}B_{n,h}\big(1, 2!, \cdots, (n-h+1)!\big)\sum_{k=1}^{h}k!\beta_k B_{n,k}(1, \cdots, 1) = \sum_{n=1}^{\infty}d_n\xi^n,$$

for any ξ close to 0, where $d_n, e_n, \tilde{e}_n$ are as in (3.9)-(3.10). Eventually, applying again Faà di Bruno formula together with (3.13) yield (3.14) with b_n as in (3.8). $\qquad\square$

We now address both point-wise and asymptotic convergence of Ψ_N to Ψ. Concerning the former one, the desired result would be

$$\sum_{n=2}^{N}G_n(\eta) \to G(\eta) \qquad \text{as } N \to +\infty, \tag{3.15}$$

for any $\eta > 0$, which would in turn imply

$$\Psi_N(z, w) \to \Psi(z, w) \qquad \text{as } N \to \infty,$$

17

for any $z, w \in \mathbb{R} \times D$ with $z \prec w$. However, unfortunately, we were able to provide a rigorous proof of (3.15) only for $\eta > 1$ (see Theorem 3.4 below). Despite of this, we point out that strong numerical evidence suggests that (3.15) is satisfied also for $\eta \in]0, 1[$. To support this claim, in Figure 2 we compare the plot of the truncated series $\sum_{n=2}^{N} G_n$ with that of G, for different values of N.
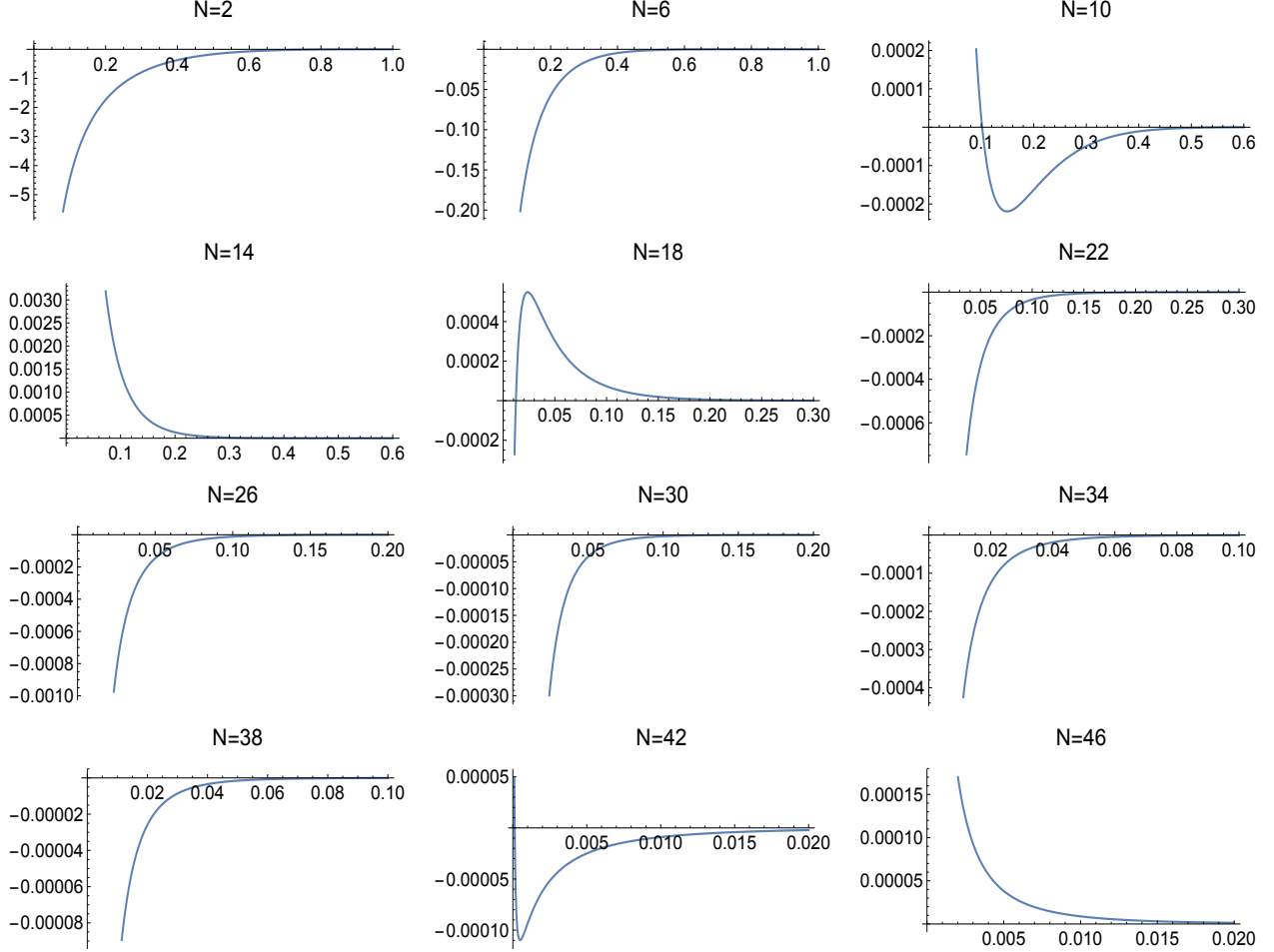
Figure 2: Plot of the difference $G(\eta) - \sum_{n=2}^{N} G_n(\eta)$, with G_n as in Definition 3.2, for values of η close to 0.

Theorem 3.4. *For any $\eta > 1$ the limit (3.15) holds true.*

The proof of Theorem 3.4 is deferred until Section 3.1.

We now turn our attention to the asymptotic convergence. By definition of g we can easily obtain

$$g^{-1}(h) \sim \log^2 h \qquad \text{as } h \to +\infty, \tag{3.16}$$

where, as already stated in the introduction, we adopted the notation

$$f \sim g \Leftrightarrow \frac{f}{g} \longrightarrow 1.$$

The definition of G together with (3.16) then yield

$$G(\eta) \sim 4\eta \qquad \text{as } \eta \to +\infty, \tag{3.17}$$

18

$$G(\eta) \sim \log^2 \eta \qquad \text{as } \eta \to 0^+.$$

The idea behind the approximation Ψ_N is to expand the function G by means of suitable basis functions so as to obtain

$$\sum_{n=2}^{N} G_n(\eta) \sim \eta \sum_{n=2}^{N} a_n \qquad \text{as } \eta \to \infty,$$

$$\sum_{n=2}^{N} G_n(\eta) \sim (\log^2 \eta) \sum_{n=2}^{N} b_n \qquad \text{as } \eta \to 0^+. \qquad (3.18)$$

Indeed, the latter easily stem from

$$\frac{(\eta - 1)^n}{\eta^{n-1}} \sim \eta \qquad \text{as } \eta \to +\infty,$$

$$\frac{(-\log \eta)^n}{(1 - \log \eta)^{n-2}} \sim \log^2 \eta \qquad \text{as } \eta \to 0^+.$$

Ideally, we would like to show that

$$\lim_{N \to \infty} \sum_{n=2}^{N} a_n = 4, \qquad (3.19)$$

$$\lim_{N \to \infty} \sum_{n=2}^{N} b_n = 1, \qquad (3.20)$$

which means the asymptotic behavior of $\sum_{n=2}^{N} G_n(\eta)$ converges to the asymptotic behavior of $G(\eta)$ as η tends to 0^+ and $+\infty$. Note that this property is in general not granted for convergent expansions. We start by considering the case $\eta > 1$. In Corollary 3.5 below, we show that (3.19) is satisfied provided that the coefficients a_n are all positive. Unfortunately, we were not able to rigorously prove the positiveness of the coefficients a_n, thought the numerical values reported in Table 1 strongly support this conjecture.

Corollary 3.5. *Under the assumption that the coefficients $(a_n)_{n \geq 2}$ are all positive, the limit (3.19) holds true.*

Proof. Let $\bar{a} := \lim_{N \to \infty} \sum_{n=2}^{N} a_n$ and assume $\bar{a} < 4$. By definition (3.5), and by (3.17), there exists $\bar{\eta} > 1$ such that

$$\sum_{n=2}^{N} G_n(\bar{\eta}) < \bar{a}\bar{\eta} < \frac{\bar{a}+4}{2}\bar{\eta} < G(\bar{\eta}), \qquad N \geq 2.$$

This violates Theorem 3.4 and yields a contradiction. Therefore, we have $\bar{a} \geq 4$.

Assume now that $\bar{a} > 4$. Then there exits $\bar{N}$ such that

$$\sum_{n=2}^{\bar{N}} a_n > \frac{\bar{a}+4}{2}.$$

On the other hand, the positiveness of the coefficients a_n, together with (3.17) and (3.18), implies that there exists $\bar{\eta} > 1$ such that

$$\sum_{n=2}^{N} G_n(\bar{\eta}) > \sum_{n=2}^{\bar{N}} G_n(\bar{\eta}) > \frac{\bar{a}+4}{2}\bar{\eta} > G(\bar{\eta}), \qquad N > \bar{N},$$

which again violates Theorem 3.4 and yields a contradiction. This proves that $\bar{a} \leq 4$ and concludes the proof.

$\square$

For $\eta \in]0, 1[$ the situation is less clear. On the one hand, the values in Table 1 suggests that the sign of the coefficients b_n oscillates as n grows, and thus we cannot rely on the positiveness (or negativeness) of the coefficients b_n to infer (3.20). On the other hand, the plot in Figure 3 shows that $\sum_{n=2}^{N} b_n$ is close to 1 for (not too) large values of N. Unfortunately, the numerical complexity of the representation (3.8) makes it very difficult to compute the summation above for larger values of N. Therefore, the question whether (3.20) is actually true remains open.
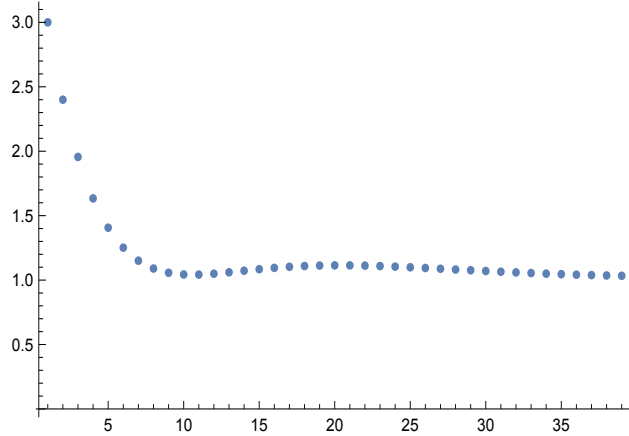

Figure 3: Plot of the sum $\sum_{n=2}^{N} b_n$ for N up to 40.

3.1. Proof of Theorem 3.4

This section is devoted to the proof of Theorem 3.4, which is preceded by some preliminary results. Hereafter, for any $z \in \mathbb{C}$ and $\delta > 0$, we denote by $B_\delta(z)$ the complex ball centered at z with radius δ.

Lemma 3.6. *The function g defined by (2.6) is holomorphic on $\mathbb{C}$, and*

$$g'(z) \neq 0, \qquad z \in B_{\pi^2}(0). \tag{3.21}$$

Furthermore, the restriction of g to $\partial B_{\pi^2}(0)$ is a homeomorphism, and its image surrounds $B_1(1)$.

Proof. Owing to the Taylor series expansion of $\sinh z$ around $z = 0$, we directly obtain

$$g(z) = \sum_{n=0}^{\infty} \frac{z^n}{(2n+1)!}, \qquad r \in \mathbb{C}, \tag{3.22}$$

which shows that g is holomorphic on $\mathbb{C}$.

We now prove that $g' \neq 0$ on $B_{\pi^2}(0)$. Differentiating (2.6) we obtain

$$g'(z) = \frac{f(\sqrt{z})}{2z^{3/2}}, \qquad z \in \mathbb{C},$$

where

$$f(z) := z \cosh z - \sinh z = \sum_{n=1}^{\infty} \frac{2n}{(2n+1)!} z^{2n+1}. \tag{3.23}$$

Therefore, owing to the fact that $g'(0) = 1/6$ (by (3.22)), and to $f(\bar{z}) = \overline{f(z)}$, we have that (3.21) is equivalent to

$$f(z) \neq 0, \qquad z \in \tilde{B}_\pi(0) \setminus \{0\}, \tag{3.24}$$

20

where

$$\tilde{B}_\pi(0) := \{x + iy \in B_\pi(0) : x, y \geq 0\}.$$

We first observe, that

$$f(x) \neq 0, \qquad x \in \mathbb{R}, \ x \neq 0, \tag{3.25}$$

which follows directly from the series representation in (3.23). Finally, a direct computation shows that

$$|f(z)|^2 = \frac{1}{2}\Big((x^2 + y^2 - 1)\cos(2y) + (x^2 + y^2 + 1)\cosh(2x) - 2x\sinh(2x) - 2y\sin(2y) \Big)$$
$$> |f(x)|^2$$

for any $z = x + iy \in \tilde{B}_\pi(0)$ with $y > 0$. This, together with (3.25), proves (3.24).

We now prove that $g|_{\partial B_{\pi^2}(0)}$ is a homeomorphism: it is enough to show that it is injective. Owing to

$$\overline{g(z)} = g(\bar{z}), \qquad z \in \mathbb{C}, \tag{3.26}$$

the latter property can be proved by checking that the function $\theta \mapsto |\sinh(\pi e^{i\theta})|$ is strictly decreasing on $[0, \pi/2]$, and that the imaginary part of $e^{-i\theta}\sinh(\pi e^{i\theta})$ is strictly positive for $\theta \in]0, \pi/2[$. We omit the details for brevity.

Finally, we prove that the image of $g|_{\partial B_{\pi^2}(0)}$ surrounds $B_1(1)$. Owing again to (3.26), it is enough to prove that

$$\left| \frac{\sinh(\pi e^{i\theta})}{\pi e^{i\theta}} - 1 \right| \geq 1, \qquad \theta \in [0, \pi/2]. \tag{3.27}$$

We have

$$\left| \frac{\sinh(\pi e^{i\theta})}{\pi e^{i\theta}} - 1 \right| = \frac{1}{\pi}\left| \sum_{n=1}^{\infty} \frac{(\pi e^{i\theta})^{2n+1}}{(2n+1)!} \right| \geq \frac{1}{\pi}|I_1(\theta) - I_2|,$$

with

$$I_1(\theta) = \left| \sum_{n=1}^{3} \frac{(\pi e^{i\theta})^{2n+1}}{(2n+1)!} \right|, \qquad I_2 = \sum_{n=4}^{\infty} \frac{\pi^{2n+1}}{(2n+1)!}.$$

Now, it is not difficult to prove that there exists $\varsigma > 0$ such that

$$I_1(\theta) > \varsigma + \pi > I_2 + \pi, \qquad \theta \in [0, \pi/2],$$

which proves (3.27). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

We are now in the position to prove Theorem 3.4.

Proof of Theorem 3.4. By the change of variable $\xi := \frac{\eta-1}{\eta}$, it suffices to prove

$$F(\xi) := (1 - \xi)G\Big(\frac{1}{1-\xi}\Big) = \sum_{n=2}^{\infty} a_n \xi^n, \qquad \xi \in]-1, 1[. \tag{3.28}$$

We prove (3.28) by showing that F is holomorphic on $B_1(0)$. Explicitly, we have

$$F(\xi) := (1 - \xi)\left(\frac{2}{1-\xi} - 2(1 - \xi - 2/\pi)\sqrt{\frac{(1-\xi)^{-2} + g^{-1}(1-\xi)}{(1 - \xi - 2/\pi)^2}} + g^{-1}(1-\xi) \right).$$

Lemma 3.6 combined with Lemma Appendix A.1 imply that $g|_{B_{\pi^2}(0)}$ is a biholomorphism on its image. In particular, the inverse g^{-1} is well defined and holomorphic on $B_1(1)$. Setting now

$$f(z) := z^{-2} + g^{-1}(z), \qquad z \in B_1(1),$$

it is straightforward to check that

$$f\left(\frac{2}{\pi}\right) = 0 = f'\left(\frac{2}{\pi}\right).$$

The first equality above follows directly from (3.3), whereas the second stems from (3.3) together with

$$f'(2/\pi) = -2z^{-3} + \left.\frac{2g^{-1}(z)}{\sqrt{1 + zg^{-1}(z) - z}}\right|_{z=2/\pi}, \qquad z \in \mathbb{R}.$$

It follows that

$$z \mapsto \frac{f(z)}{(z - 2/\pi)^2}$$

is holomorphic on $B_1(1)$. it is also possible to observe that $f(B_1(1))$ does not contain non-negative real numbers, which finally yields holomorphicity of F on $B_1(0)$.

$\square$

4. Yosida's parametrix construction

For the sake of clarity, before dwelling on the details of our extension of Yosida's parametrix method, we outline the general parametrix construction of the fundamental solution p, for a general choice of the parametrix function.

4.1. General parametrix construction

Consider a given function $H(z; w)$, hereafter referred to as *parametrix function*, or simply *parametrix*, which enjoys suitable regularity and boundedness properties, together with the Dirac delta property ii) in Definition 2.2, and such that

$$H(t, x_1, x_2; T, y_1, y_2) \to 0 \qquad \text{as } y_2 \to x_2^+. \tag{4.1}$$

Following the standard approach, we look for $p = p(z; w)$ of the form

$$p(z, w) = H(z, w) + \int_{z \prec \zeta \prec w} H(z; \zeta)\Phi(\zeta; w)d\zeta, \qquad z \prec w. \tag{4.2}$$

By formally applying Point i) in Definition 2.2, we obtain

$$0 = \mathcal{L}p(z, w) = \mathcal{L}H(z, w) + \mathcal{L}\int_{z \prec \zeta \prec w} H(z; \zeta)\Phi(\zeta; w)d\zeta$$

(owing to the Dirac delta property $H(t, x; t, \xi) = \delta_x$ and to (4.1))

$$= \mathcal{L}H(z, w) + \int_{z \prec \zeta \prec w} \mathcal{L}H(z; \zeta)\,\Phi(\zeta; w)d\zeta - \Phi(z, w),$$

which yields

$$\Phi(z, w) = \mathcal{L}H(z, w) + \int_{z \prec \zeta \prec w} \mathcal{L}H(z; \zeta)\,\Phi(\zeta; w)d\zeta, \qquad z \prec w. \tag{4.3}$$

Now by Picard iteration we look for a solution to (4.3) of the form

$$\Phi(z;\zeta) = \sum_{n=0}^{\infty} K_n(z;\zeta), \qquad z \prec w, \tag{4.4}$$

where the functions K_n are defined through the recursion

$$K_0(z;w) = \mathcal{L}H(z;w), \qquad K_n(z;w) = \int_{z \prec \zeta \prec w} \mathcal{L}H(z;\zeta)K_{n-1}(\zeta,w)d\zeta, \qquad n \geq 1. \tag{4.5}$$

In order for the construction above to be formalized, one needs to find a parametrix H satisfying suitable estimates so that:

a) the series in (4.4) is convergent;

b) the properties i) and ii) of Definition 2.2 can be verified.

Furthermore, as a by-product of (4.2) and of the estimates for H and Φ, one can prove bounds for p and its derivatives. In our case, thanks to a specific choice of H, we will be able to prove the estimates in Theorem 1.1 and Theorem 1.2.

4.2. Generalized Yosida's parametrix

Following Yosida [7], we introduce our parametrix H in two steps. We first define a *pre-parametrix* H_1 as a function having the expected asymptotic behavior both away from the pole, and close to the pole. As we will see, the function H_1 is not suitable to define the integrals appearing in (4.5). Hence, in the second step, we introduce a correction term $\mathbf{u}$ that makes the functions K_n's well defined if we use $H := \mathbf{u}H_1$ as a parametrix.

Hereafter throughout the whole section we will set $\sigma = 1$ without loss of generality (see Remark 2.5). For any $z = (t,x), w = (T,y) \in \mathbb{R} \times D$ with $z \prec w$, define the *pre-parametrix* H_1 as in (1.31). Explicitly,

$$H_1(z;w) = \frac{\sqrt{12}}{2\pi(T-t)^2 y_1^2} e^{-\frac{1}{2}\Psi(z;w)}. \tag{4.6}$$

The idea behind this definition is the following. On one hand, the cost function Ψ at the exponential is supposed to provide the exact asymptotic behavior of the fundamental solution away from the pole. On the other hand, the square root of the determinant of $\mathbf{C}$ in the denominator yields the same singular behavior near the pole as the fundamental solution of the *frozen* operator

$$\tilde{\mathcal{L}}_y = \partial_t + x_1 \partial_{x_2} + \frac{y_1^2}{2} \partial_{x_1 x_1}, \qquad (t, x_1, x_2) \in \mathbb{R}^3.$$

In practice, this specific choice for the denominator ensures that $H_1(z,w)$ enjoys the Dirac Delta property ii) in Definition 2.2. The exact statement for this is in Proposition 4.4 below.

Now, if we were to choose H_1 as a parametrix, the first step towards proving the convergence of the series (4.4) would be to bound $\mathcal{L}H_1$ from above. To do so, one first computes

$$\partial_t H_1 = \left(\frac{2}{T-t} - \frac{1}{2}\partial_t \Psi\right) H_1,$$

$$\partial_{x_1} H_1 = -\frac{1}{2}\left(\partial_{x_1}\Psi\right) H_1,$$

23

$$\partial_{x_2} H_1 = -\frac{1}{2}\big(\partial_{x_2}\Psi\big)H_1,$$

$$\partial_{x_1 x_1} H_1 = \left(\frac{1}{4}\big(\partial_{x_1}\Psi\big)^2 - \frac{1}{2}\partial_{x_1 x_1}\Psi\right)H_1,$$

and obtains

$$\mathcal{L}H_1(z;w) = \left(\left(\frac{x_1}{2}\partial_{x_1}\Psi(z;w)\right)^2 - Y\Psi(z;w) + f(z,w)\right)H_1(z;w), \tag{4.7}$$

with

$$f(t,x;T,y) := \frac{2}{T-t} - \frac{x_1^2 \partial_{x_1 x_1}\Psi(t,x;T,y)}{4}. \tag{4.8}$$

At first glance, a twofold problem appears:

- the function $f(z,w)$ is singular of order $\frac{1}{T-t}$;

- the functions

$$\left(\frac{x_1}{2}\partial_{x_1}\Psi(z;w)\right)^2 \quad \text{and} \quad Y\Psi(z;w)$$

are singular of order $\frac{1}{(T-t)^2}$.

In light of the space-time convolutions that define the terms K_n in (4.5), the presence of these singular terms seems to undermine the construction using H_1. However, the next lemma shows that the singular terms of order $\frac{1}{(T-t)^2}$ cancel each other out.

Lemma 4.1 (HJB equation). *For any $w = (T,y) \in \mathbb{R} \times D$, the function $\Psi(\cdot,\cdot;T,y)$ satisfies*

$$Y\Psi(z;w) = \left(\frac{x_1}{2}\partial_{x_1}\Psi(z;w)\right)^2, \qquad z = (t,x) \in \mathbb{R} \times D,\ z \prec w. \tag{4.9}$$

Furthermore, the optimal control ω for the problem (1.7)-(1.8) *satisfies*

$$\omega(s) = -\frac{\gamma_1(s)}{2}\partial_{\gamma_1}\Psi\big(s,\gamma(s);w\big), \qquad s \in [t,T]. \tag{4.10}$$

The proof is deferred until Appendix B.

By applying (4.9) to (4.7) we obtain

$$\mathcal{L}H_1(z;w) = f(z,w)H_1(z;w). \tag{4.11}$$

In order to get rid of the singularity of order $\frac{1}{T-t}$, we define the *parametrix H* as

$$H(z;w) := H_1(z;w)\mathbf{u}(z;w), \tag{4.12}$$

where $\mathbf{u}$ is a regular function, suitably chosen so as to obtain a uniform bound for $\mathcal{L}H(z,w)$. By (4.11) we obtain

$$\mathcal{L}H(z;w) = \Big(f(z,w)\mathbf{u}(z;w) - g(z,w)\partial_{x_1}\mathbf{u}(z;w) + \mathcal{L}\mathbf{u}(z;w)\Big)H_1(z;w)$$

$$= \left(\big(Y - g(z;w)\partial_{x_1} + f(z,w)\big)\mathbf{u}(z;w) + \frac{x_1^2}{2}\partial_{x_1 x_1}\mathbf{u}(z;w)\right)H_1(z;w), \tag{4.13}$$

with f as given in (4.8) and

$$g(z;w) := \frac{x_1^2}{2}\partial_{x_1}\Psi(z;w).$$

The idea is then to find $\mathbf{u}$ that verifies

$$\big(Y - g(z;w)\partial_{x_1}\big)\mathbf{u}(z;w) + f(z;w)\mathbf{u}(z;w) = 0, \qquad z \prec w, \tag{4.14}$$

and such that

$$\frac{x_1^2}{2}\partial_{x_1 x_1}\mathbf{u}(z;w)$$

is bounded with respect to $T - t$. We also impose that that the initial condition

$$\mathbf{u}(w;w) = 1 \tag{4.15}$$

is verified in some sense. Intuitively, the latter is required in order for the Dirac Delta property, as $(T-t) \to 0^+$, to be transferred from H_1 to H. We seek a solution to (4.14)-(4.15) by employing the method of the characteristic curves. We have the following crucial

Lemma 4.2. *For any $w = (T, y) \in \mathbb{R} \times D$, the integral curves of the vector field*

$$z \mapsto Y - g(z, w)\partial_{x_1}, \qquad z = (t, x) \in \mathbb{R} \times D, \ z \prec w,$$

are the optimal curves $\big(s, \gamma(s)\big)$ for the optimal control problem (1.7)-(1.8) with $\gamma(T) = y$.

Proof. By (4.10), we have

$$\begin{cases} \dot\gamma_1(s) = -g\big(s, \gamma(s); w\big) \\ \dot\gamma_2(s) = \gamma_1(s) \end{cases}, \qquad t < s < T,$$

which completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Applying now Lemma 4.2, together with the method of the characteristic curves, we find that the solution to (4.14)-(4.15) has to be

$$\begin{aligned} \mathbf{u}(z;w) &:= \exp\left(\int_t^T f\big(s, \gamma(s); T, y\big)ds\right) \\ &= \exp\left(\int_t^T \left[\frac{2}{T - s} - \frac{\gamma_1^2(s)\partial_{\gamma_1\gamma_1}\Psi\big(s, \gamma(s); w\big)}{4}\right]ds\right), \qquad z \prec w, \end{aligned} \tag{4.16}$$

where γ denotes the optimal trajectory of the control problem (1.7)-(1.8). In the next section we will prove that such $\mathbf{u}$ is well defined, and that the iterative construction described in Section 4.1 converges to the fundamental solution p of $\mathcal{L}$, with the parametrix function H given by (4.12)-(4.6)-(4.16).

4.3. Convergence of the Picard series and proofs of Theorems 1.1 and 1.2

We prove Theorems 1.1 and 1.2 by estimating the series (4.4) and the integral in (4.2) with H as defined in the previous section, and by proving that p as defined by (4.2) is indeed the fundamental solution of $\mathcal{L}$.

We start with the following two propositions, whose proofs are deferred until Sections 4.3 and 4.4.

Proposition 4.3. *The function $\mathbf{u}$ given by (4.16) is well defined and we have*

$$\mathbf{u}(z;w) = v\left(4\,g^{-1}\Big(\frac{1}{\mathbf{h}(z;w)}\Big)\right), \tag{4.17}$$

25

with

$$v(\eta) = \begin{cases} \dfrac{|\eta|}{2\sqrt{3\sqrt{\eta}\sinh(\sqrt{\eta}) - 6\cosh(\sqrt{\eta}) + 6}}, & \eta \in \,]-4\pi^2, +\infty[\,\backslash\{0\} \\ 1, & \eta = 0 \end{cases}. \tag{4.18}$$

Furthermore, the function $\mathbf{u}(\cdot\,;\zeta)$ *is smooth and solves* (4.14).

Finally there exists a universal constant $\kappa > 0$ *such that*

$$\kappa^{-1}\mathbf{1}_{]0,1]}\big(\mathbf{h}(z;w)\big)\mathbf{h}(z;w) \le |\mathbf{u}(z;w)| \le \kappa\big(\sqrt{\mathbf{h}} + \mathbf{h}\big)(z;w), \tag{4.19}$$

$$\big|x_1^2\partial_{x_1 x_1}\mathbf{u}(z;w)\big| \le \kappa\,\sqrt{\mathbf{h}(z;w)}, \tag{4.20}$$

for any $z = (t,x), w = (T,y) \in \mathbb{R} \times D$ *with* $z \prec w$.

The next statement is crucial in order to check that parametrix H enjoys the Dirac delta property ii) in Definition 2.2.

Proposition 4.4. *For any function* $\tilde{v} \in C(\mathbb{R}^+)$ *bounded by a power function, such that* $\tilde{v}(1) = 1$, *we have*

$$\lim_{T \to 0^+} \int_{(1,0) \prec \xi} \tilde{v}\big(\mathbf{h}(\mathbf{1}_{\mathbb{G}};T,\xi)\big)H_1(\mathbf{1}_{\mathbb{G}};T,\xi)\varphi(\xi)d\xi = \varphi(1,0), \qquad \varphi \in C_b(D). \tag{4.21}$$

In order to carry on with the convergence analysis it is essential to provide bounds for the integrals in (4.5) and (4.2). In light of estimates (4.20)-(4.19), such bounds are consequences of the following

Key Inequalities 4.5. *Let* $\mathbf{h}$ *be as defined in* (4.20) *and* $\tilde{g}$ *be the function*

$$\tilde{g}(h) = \sqrt{h} + h.$$

For any $\tau > 0$, *there exists a constant* $C_\tau > 0$ *such that*

$$\int_{x \prec \xi \prec (1,0)} \sqrt{\mathbf{h}(z;s,\xi)}H_1(z;s,\xi)\sqrt{\mathbf{h}(s,\xi;\mathbf{1}_{\mathbb{G}})}H_1(s,\xi;\mathbf{1}_{\mathbb{G}})d\xi \le C_\tau\,\sqrt{\mathbf{h}(z;\mathbf{1}_{\mathbb{G}})}H_1(z,\mathbf{1}_{\mathbb{G}}), \qquad s \in\,]t,0[, \tag{4.22}$$

$$\int_{x \prec \xi \prec (1,0)} \tilde{g}\big(\mathbf{h}(z;s,\xi)\big)H_1(z;s,\xi)\tilde{g}\big(\mathbf{h}(s,\xi;\mathbf{1}_{\mathbb{G}})\big)H_1(s,\xi;\mathbf{1}_{\mathbb{G}})d\xi \le C_\tau\,\tilde{g}\big(\mathbf{h}(z;\mathbf{1}_{\mathbb{G}})\big)H_1(z,\mathbf{1}_{\mathbb{G}}), \qquad s \in\,]t,0[, \tag{4.23}$$

for any $z = (t,x) \in \mathbb{R} \times D$ *such that* $z \prec \mathbf{1}_{\mathbb{G}}$ *and* $t > -\tau$.

As already pointed out in the introduction, at the current stage we were not able to provide a proof for the bounds in Key Inequalities 4.5. In Section 4.5 we collect a considerable amount of numerical evidence in favor of the claim that these estimates hold true. This makes us comfortable in conjecturing their validity. However, given that a rigorous proof is currently missing, the Key Inequalities 4.5 are part of the hypotheses of Theorems 1.1 and 1.2.

Theorem 4.6. *Assume that Key Inequalities 4.5 hold true. Then the series in* (4.4) *converges, and for any* $\tau > 0$ *there exists a positive constant* $C > 0$, *only dependent on* τ, *such that*

$$\left| \int_{z \prec \zeta \prec w} H(z;\zeta)\Phi(\zeta;w)d\zeta \right| \le C(T-t)\big(\sqrt{\mathbf{h}(z;w)} + \mathbf{h}(z;w)\big)H_1(z,w), \tag{4.24}$$

for any $z = (t,x), \zeta = (T,w) \in \mathbb{R} \times D$ *with* $z \prec w$ *and* $T - t < \tau$.

Furthermore, the function p *given by* (4.2) *is the fundamental solution of* $\mathcal{L}$ *with* $\sigma = 1$.

Before proving Theorem 4.6, we prove Theorems 1.1 and 1.2, which are straightforward consequences of Theorem 4.6.

Proof of Theorem 1.1. In light of Remark 2.5, it is not restrictive to assume $\sigma = 1$. Thus the bound (1.10) stems from Theorem 4.6, in particular by applying the estimates (4.19)-(4.24) to (4.2). $\qquad\square$

Proof of Theorem 1.2. In light of Remark 2.5, it is not restrictive to assume $\sigma = 1$ and $w = \mathbf{1}_{\mathbb{G}}$. To ease notation we remove the explicit dependence on $\mathbf{1}_{\mathbb{G}}$ in the functions below.

All the following inequalities are meant for any $z = (t, x) \in]-1, 0[\times D$ such that

$$\frac{1}{\kappa} \leq \frac{\sqrt{x_1}}{-x_2} \leq \kappa. \tag{4.25}$$

By Theorem 4.6 we obtain

$$p(t, x) \geq H(t, x) - \left| \int_{(t,x)\prec\zeta\prec w} H(t, x; \zeta)\Phi(\zeta)d\zeta \right| \geq H(t, x) - C(-t)\big(\sqrt{\mathbf{h}} + \mathbf{h}\big)(t, x)H_1(t, x), \tag{4.26}$$

where C is a universal constant independent of any variable. Furthermore, by (4.25), we have

$$\mathbf{h}(t, x) \leq 1, \qquad t \in [-\kappa^{-1}, 0[,$$

and thus (4.26) together with the first equality in (4.19) yield

$$p(t, x) \geq \Big(\mathbf{h}(t, x) - C(-t)\big(\sqrt{\mathbf{h}} + \mathbf{h}\big)(t, x)\Big)H_1(t, x)$$
$$= \Big(1 - C(-t) - C\sqrt{-t}\Big(\frac{-x_2}{\sqrt{x_1}}\Big)^{1/2}\Big)\mathbf{h}(t, x)H_1(t, x)$$

(by the first inequality in (4.25))

$$\geq \big(1 - C(-t) - C\sqrt{-t}\sqrt{\kappa}\,\big)\mathbf{h}(t, x)H_1(t, x)$$

for any $t \in [-\kappa^{-1}, 0[$. This completes the proof. $\qquad\square$

We conclude the section with the proof of Theorem 4.6.

Proof of Theorem 4.6. We first prove convergence of the series in (4.4) and estimate (4.24). In light of (2.8)-(2.9), it is not restrictive to assume $w = \mathbf{1}_{\mathbb{G}}$. To ease notation we remove the explicit dependence on $\mathbf{1}_{\mathbb{G}}$ in the functions below. Furthermore, we will denote by $C_1, C_2, \dots$ any positive constant that depends at most on τ.

By Proposition 4.3, in particular by applying (4.14) to (4.13), we obtain

$$\mathcal{L}H(z, \zeta) = \frac{x_1^2}{2}\partial_{x_1 x_1}\mathbf{u}(z, \zeta)H_1(z, \zeta), \qquad z \prec \zeta. \tag{4.27}$$

Therefore, by (4.5) and employing (4.20), we obtain

$$|K_0(z)| \leq C_1\sqrt{\mathbf{h}(z)}H_1(z).$$

In general, by induction we can also prove

$$|K_n(z)| \leq C_2^n C_1^{n+1}\frac{(-t)^n}{n!}\sqrt{\mathbf{h}(z)}H_1(z) \tag{4.28}$$

27

for any $n \in \mathbb{N}_0$. Indeed, assuming (4.28) true, we have

$$|K_{n+1}(z)| \leq \frac{C_2^n C_1^{n+1}}{n!} \int_{z \prec \zeta \prec \mathbf{1}_\mathbb{G}} (-s)^n |\mathcal{L}H(z,\zeta)| \sqrt{\mathbf{h}(\zeta)} H_1(\zeta) d\zeta$$

(by (4.27)-(4.20))

$$\leq \frac{C_2^n C_1^{n+2}}{n!} \int_{z \prec \zeta \prec \mathbf{1}_\mathbb{G}} (-s)^n \sqrt{\mathbf{h}(z;\zeta)} H_1(z;\zeta) \sqrt{\mathbf{h}(\zeta)} H_1(\zeta) d\zeta$$

(by (4.22))

$$\leq \frac{C_2^{n+1} C_1^{n+2}}{n!} \int_t^0 (-s)^n ds \sqrt{\mathbf{h}(z)} H_1(z)$$

$$\leq C_2^{n+1} C_1^{n+2} \frac{(-t)^{n+1}}{(n+1)!} \sqrt{\mathbf{h}(z)} H_1(z).$$

Summing over n, we obtain that the series in (4.4) converges and that

$$|\Phi(z)| \leq C_1 \, e^{-C_1 C_2 t} \sqrt{\mathbf{h}(z)} H_1(z).$$

This, together with (4.19)-(4.23), yields (4.24).

We now go on to prove the second part of the statement, namely that p as defined by (4.2) is the fundamental solution of $\mathcal{L}$ with $\sigma = 1$.

For any $x' \in D$ and $t < s$, and for any $\varphi \in C_b(D)$, we have

$$\int_{x' \prec y} H(t, x'; T, y) \varphi(y) dy = \int_{x' \prec y} \mathbf{u}(t, x'; T, y) H_1(t, x'; T, y) \varphi(y) dy$$

(by (2.8)-(2.9))

$$= \int_{x' \prec y} \mathbf{u}\big(\mathbf{1}_\mathbb{G}; (t, x')^{-1} \circ (T, y)\big)(x_1')^{-2} H_1\big(\mathbf{1}_\mathbb{G}; (t, x')^{-1} \circ (T, y)\big) \varphi(y) dy$$

$$= \int_{(1,0) \prec \xi} \mathbf{u}(\mathbf{1}_\mathbb{G}; T - t, \xi) H_1(\mathbf{1}_\mathbb{G}; T - t, \xi) \varphi_{x'}(\xi) d\xi,$$

where we used the notation

$$\varphi_x(\xi) := \big(\xi_1 x_1^{-1}, x_2 + \xi_2 x_1^{-1}\big).$$

Therefore, employing Proposition 4.4, it is straightforward to show that

$$\lim_{\substack{(t,x') \to z \\ t < T}} \int_{x' \prec y} H(t, x'; T, y) \varphi(y) dy = \varphi(x) \tag{4.29}$$

for any $z = (T, x) \in \mathbb{R} \times D$ and $\varphi \in C_b(D)$. In analogous way, we can employ estimate (4.24) together with Theorem 4.4, to prove

$$\lim_{\substack{(t,x') \to z \\ t < T}} \int_{x' \prec y} \left(\int_{(t,x') \prec \zeta \prec (T,y)} H(t, x'; \zeta) \Phi(\zeta; T, y) d\zeta \right) \varphi(y) dy = 0$$

for any $z = (T, x) \in \mathbb{R} \times D$ and $\varphi \in C_b(D)$. This together with (4.29) imply that p satisfies property ii) in Definition 2.2.

28

We now prove property i) in Definition 2.2, namely that $p(\cdot, w)$ as defined by (4.2) satisfies (2.1). By (2.8)-(2.9) it is easy to show that

$$y_1^2 p(z; w) = p(w^{-1} \circ z; \mathbf{1}_{\mathbb{G}}).$$

In light of this, and of the left invariance of $\mathcal{L}$ (see (2.2)), it is not restrictive to set $w = \mathbf{1}_{\mathbb{G}}$. Once more, to ease notation we remove the explicit dependence on $\mathbf{1}_{\mathbb{G}}$ in the functions below. We need to show that

$$\mathcal{L}p(z) = 0, \qquad z = (t, x) \in \Omega, \tag{4.30}$$

with $\Omega := \mathbb{R}^- \times \mathbb{R}^+ \times \mathbb{R}^-$. Since the operator $\mathcal{L}$ is hypoelliptic, it is enough to show that p solves (4.30) in the sense of distributions, namely

$$\int_\Omega p(z) \tilde{\mathcal{L}} \phi(z) dz = 0, \qquad \phi \in C_0^\infty(\Omega), \tag{4.31}$$

where the operator $\tilde{\mathcal{L}}$ denotes the formal adjoint of $\mathcal{L}$, i.e.

$$\tilde{\mathcal{L}} \phi(z) = -\partial_t \phi(z) - x_1 \partial_{x_2} \phi(z) + \frac{1}{2} \partial_{x_1 x_1} \big(x_1^2 \phi(z) \big).$$

We first prove that, for any $\zeta = (s, \xi) \in \Omega$, we have

$$\int_{z \prec \zeta} H(z; \zeta) \tilde{\mathcal{L}} \phi(z) dz = \int_{z \prec \zeta} \phi(z) \mathcal{L} H(z; \zeta) dz - \phi(\zeta), \qquad \phi \in C_0^\infty(\Omega). \tag{4.32}$$

Note that, proceeding like we did above to prove (4.29) (we skip the details for brevity), we obtain

$$\lim_{t \to s^-} \int_{x \prec \eta} H(t, x; \zeta) \varphi(x) dx = \varphi(\eta), \qquad \varphi \in C_b(D). \tag{4.33}$$

Furthermore, since $\mathbf{h}(t, x_1, x_2; \zeta) \to +\infty$ as $x_2 \to \xi_2^-$, (3.1)-(3.17) together with (4.19) simply yield

$$H(t, x_1, x_2; \zeta) \to 0 \quad \text{as } x_2 \to \xi_2^-, \qquad t < s, \; x_1 \in \mathbb{R}^+. \tag{4.34}$$

For any $\delta > 0$ we obtain

$$\int_{-\infty}^{s-\delta} \int_{x \prec \xi} H(t, x; s, \xi) \tilde{\mathcal{L}} \phi(t, x) dx dt$$

(integrating by parts w.r.t. x and using the boundary condition (4.34))

$$= \int_{-\infty}^{s-\delta} \int_{x \prec \xi} \mathcal{L} H(t, x; s, \xi) \phi(t, x) dx dt + \int_{x \prec \xi} H(s - \delta, x; s, \xi) \phi(s - \delta, x) dx.$$

Passing to the limit as $\delta \to 0^+$, together with (4.33), yields (4.32).

We now go on to prove (4.31). Integrating by parts, we clearly obtain

$$\int_\Omega H(z; \mathbf{1}_{\mathbb{G}}) \tilde{\mathcal{L}} \phi(z) dz = \int_\Omega \phi(z) \mathcal{L} H(z; \mathbf{1}_{\mathbb{G}}) dz, \qquad \phi \in C_0^\infty(\Omega). \tag{4.35}$$

Furthermore,

$$\int_\Omega \left(\int_{z \prec \zeta \prec \mathbf{1}_{\mathbb{G}}} H(z; \zeta) \Phi(\zeta; \mathbf{1}_{\mathbb{G}}) d\zeta \right) \tilde{\mathcal{L}} \phi(z) dz$$

29

(by Fubini's Theorem)

$$= \int_\Omega \left(\int_{z \prec \zeta} H(z; \zeta) \tilde{\mathcal{L}} \phi(z) dz \right) \Phi(\zeta; \mathbf{1}_\mathbb{G}) d\zeta$$

(by (4.32))

$$= \int_\Omega \left(\int_{z \prec \zeta} \phi(z) \mathcal{L} H(z; \zeta) dz - \phi(\zeta) \right) \Phi(\zeta; \mathbf{1}_\mathbb{G}) d\zeta$$

(again by Fubini's Theorem)

$$= \int_\Omega \left(\int_{z \prec \zeta \prec \mathbf{1}_\mathbb{G}} \mathcal{L} H(z; \zeta) \, \Phi(\zeta; \mathbf{1}_\mathbb{G}) d\zeta \right) \phi(z) dz - \int_\Omega \phi(\zeta) \Phi(\zeta; \mathbf{1}_\mathbb{G}) d\zeta$$

(by construction Φ solves (4.3))

$$= - \int_\Omega \phi(z) \mathcal{L} H(z; \mathbf{1}_\mathbb{G}) dz$$

for any $\phi \in C_0^\infty(\Omega)$.

This, together with (4.35) and (4.2), proves (4.31) and completes the proof.

$\square$

4.4. Proof of Proposition 4.3

We start with the following

Lemma 4.7. *For any $z = (t, x), w = (T, y) \in \mathbb{R} \times D$ with $z \prec \zeta$, denoting by $\gamma = \gamma(s) = \big(\gamma_1(s), \gamma_2(s)\big)$ the optimal trajectory of the control problem (1.7)-(1.8), we have*

$$\gamma_1^2(s) \partial_{\gamma_1 \gamma_1} \Psi\big(s, \gamma(s); T, y\big) = \frac{h\left(\frac{(T-s)^2}{(T-t)^2} \, g^{-1}\big(\frac{1}{\mathbf{h}(z;w)} \big) \right)}{T - s}, \qquad s \in [t, T[,$$

with

$$h(\eta) = \begin{cases} 2\sqrt{\eta} \coth(\sqrt{\eta}) - \frac{2\eta}{1 - \sqrt{\eta} \coth(\sqrt{\eta})}, & \eta \in]-\pi^2, +\infty[\setminus \{0\} \\ 8, & z = 0 \end{cases} \tag{4.36}$$

Proof. In [6] it was shown that the optimal curve γ is given by

$$\gamma(s) = \begin{cases} \left(\frac{4y_1}{(k(T-s)+2)^2}, y_2 - \frac{2(T-s)y_1}{k(T-s)+2} \right), & \text{if } E = 0 \\[3mm] \left(\frac{E y_1}{\left(\sqrt{E} \cosh\left(\frac{T-s}{2} \sqrt{E} \right) + k \sinh\left(\frac{T-s}{2} \sqrt{E} \right) \right)^2}, y_2 - \frac{2 \sinh\left(\frac{T-s}{2} \sqrt{E} \right) y_1}{\sqrt{E} \cosh\left(\frac{T-s}{2} \sqrt{E} \right) + k \sinh\left(\frac{T-s}{2} \sqrt{E} \right)} \right), & \text{if } E > 0 \\[3mm] \left(\frac{-E y_1}{\left(\sqrt{-E} \cos\left(\frac{T-s}{2} \sqrt{-E} \right) + k \sin\left(\frac{T-s}{2} \sqrt{-E} \right) \right)^2}, y_2 - \frac{2 \sin\left(\frac{T-s}{2} \sqrt{-E} \right) y_1}{\sqrt{-E} \cos\left(\frac{T-s}{2} \sqrt{-E} \right) + k \sin\left(\frac{T-s}{2} \sqrt{-E} \right)} \right), & \text{if } E < 0 \end{cases} \tag{4.37}$$

where $E = E_{z,w}$ is as defined in (2.5)-(2.6), and where

$$k = k_{z,w} := \begin{cases} \frac{2y_1}{y_2 - x_2} - \sqrt{E + \frac{4x_1 y_1}{(y_2 - x_2)^2}}, & \text{if } E \geq -\frac{\pi^2}{(T-t)^2} \\[3mm] -\frac{2y_1}{y_2 - x_2} - \sqrt{E + \frac{4x_1 y_1}{(y_2 - x_2)^2}}, & \text{if } -\frac{4\pi^2}{(T-t)^2} < E < -\frac{\pi^2}{(T-t)^2} \end{cases}.$$

We also have

$$\frac{y_2 - \gamma_2(s)}{(T-s)\sqrt{\gamma_1(s)y_1}} = \frac{2\sinh\left(\frac{(T-s)\sqrt{E}}{2}\right)}{(T-s)\sqrt{E}},$$

which implies

$$g^{-1}\left(\frac{y_2 - \gamma_2(s)}{(T-s)\sqrt{\gamma_1(s)y_1}}\right) = \frac{(T-s)^2 E}{4}. \qquad (4.38)$$

The statement now follows from a direct computation of $\partial_{x_1 x_1}\Psi(t,x;T,y)$, together with (4.37) and (4.38). $\qquad\square$

We are now in the position to prove Theorem 4.3.

Proof of Theorem 4.3. Fix $z = (t,x), w = (T,y) \in \mathbb{R} \times D$ such that $z \prec w$, and denote as usual by $\gamma = \gamma(s)$ the optimal curve connecting z to w. By Lemma 4.7 together with (4.8) we obtain

$$f\big(s,\gamma(s);T,y\big) := \frac{2 - \frac{1}{4}h\left(\frac{(T-s)^2}{(T-t)^2}\,g^{-1}\left(\frac{1}{\mathbf{h}(z;w)}\right)\right)}{T-s}, \qquad s \in [t,T[.$$

A simple change of variables now yields

$$\int_t^T f\big(t,\gamma(s);T,y\big)ds = \int_0^1 \tilde{f}(\tau)d\tau, \qquad (4.39)$$

with

$$\tilde{f}(\tau) = \frac{2 - \frac{1}{4}h\left(\tau^2 g^{-1}\left(\frac{1}{\mathbf{h}(z;w)}\right)\right)}{\tau}$$

and h as in (4.36). By Taylor-expanding the function $h(z)$ around $z = 0$ we obtain

$$2 - \frac{1}{4}h(z) = -\frac{16}{15}z + O\big(z^{3/2}\big) \qquad \text{as } z \to 0.$$

Furthermore, it can be checked that

$$1 - \sqrt{\eta}\coth(\sqrt{\eta}) \neq 0, \qquad \eta \in\,]-\pi^2, +\infty[\,\backslash\{0\},$$

and thus $\tilde{f}$ is continuous on $[0,1]$. Therefore, the integral in (4.39) is well defined, and so is the function $\mathbf{u}$ as given by (4.16).

We now go on to prove that $\mathbf{u}$ can be represented as in (4.17)-(4.18), and that it is smooth. Set

$$a := 4g^{-1}\left(\frac{1}{\mathbf{h}(z;w)}\right).$$

If $\mathbf{h}(z;w) = 1$, then $a = 0$ and (4.17)-(4.18) is trivially satisfied. If $\mathbf{h}(z;w) \neq 1$, then $a \neq 0$ and it is easy to check that the function

$$\tilde{F}(\tau) := 2\log\tau - \frac{1}{2}\log\left(\sqrt{a}\tau\sinh(\sqrt{a}\tau) - 2\cosh(\sqrt{a}\tau) + 2\right)$$

is a primitive of $\tilde{f}$ on $]0,\infty[$. Thus we obtain

$$\int_0^1 \tilde{f}(\tau)d\tau = \tilde{F}(1) - \lim_{\tau \to 0^+}\tilde{F}(\tau) =$$

$$= -\frac{\log\left(\sqrt{a}\sinh(\sqrt{a}) - 2\cosh(\sqrt{a}) + 2\right)}{2} + \frac{1}{2}\log\left(\frac{a^2}{12}\right).$$

This yields

$$\mathbf{u}(z;\zeta) = \frac{|a|}{2\sqrt{3\sqrt{a}\sinh\left(\sqrt{a}\right) - 6\cosh\left(\sqrt{a}\right) + 6}}, \qquad a := 4g^{-1}\left(\frac{1}{\mathbf{h}(z;w)}\right),$$

which is (4.17)-(4.18). Furthermore, by Taylor expanding the functions sinh and cosh around zero we obtain

$$\frac{1}{v(\eta)} = \sqrt{1 + \sum_{n=1}^{\infty} \frac{12}{(2n+3)!}\frac{n+1}{n+2}\eta^n}, \qquad \eta \in]-4\pi^2, +\infty[.$$

This shows that v is smooth on $]-4\pi^2, +\infty[$, and since g^{-1} and $\mathbf{h}$ are smooth, then $\mathbf{u}$ is smooth on its existence domain.

We now prove that $\mathbf{u}(\cdot;\zeta)$ solves (4.14). We employ the method of the characteristics. Fix again $z = (t,x), w = (T,y) \in \mathbb{R} \times D$ such that $z \prec w$, and recall that $\gamma = \gamma(s)$ denotes the optimal curve connecting z to w. By (4.16) we have

$$\mathbf{u}\big(r,\gamma(r);w\big) = e^{\int_r^T f(s,\gamma(s);w)ds}, \qquad r \in [t,T[.$$

This yields

$$\frac{d}{dr}\mathbf{u}\big(r,\gamma(r);w\big) = -f\big(r,\gamma(r);w\big)\mathbf{u}\big(r,\gamma(r);w\big), \qquad r \in [t,T[. \tag{4.40}$$

On the other hand the smoothness of $\mathbf{u}$, together with Lemma 4.2, implies

$$\frac{d}{dr}\mathbf{u}\big(r,\gamma(r);w\big) = \big\langle(1,\dot{\gamma}(r)),\nabla\mathbf{u}\big(r,\gamma(r);w\big)\big\rangle = \big(\partial_t + \gamma_1(r)\partial_{x_2} - g(\gamma(r);w)\partial_{x_1}\big)\mathbf{u}\big(\gamma(r);w\big), \qquad r \in [t,T[. \tag{4.41}$$

Setting $r = t$ in (4.40) and (4.41) we obtain (4.14).

We now prove the bounds in (4.19). Set

$$r(\rho) := \frac{g^{-1}(\rho)}{\log^2\rho}, \qquad \rho >> 1,$$

and obtain

$$v\big(4g^{-1}(\rho)\big) = v\big(4r(\rho)\log^2\rho\big) = \rho^{-\sqrt{r(\rho)}}(\log\rho)^{\frac{3}{2}}R\big(\rho,r(\rho)\big), \tag{4.42}$$

with

$$R(\rho,r) := 2r\left(3\sqrt{r}\big(1 - \rho^{-4\sqrt{r}}\big) - \frac{3}{\log\rho}\big(1 + \rho^{-4\sqrt{r}}\big) + 6\frac{\rho^{-2\sqrt{r}}}{\log\rho}\right)^{-\frac{1}{2}}.$$

By (3.16), which is

$$r(\rho) \longrightarrow 1 \qquad \text{as } \rho \to +\infty,$$

we obtain

$$R\big(\rho,r(\rho)\big) \longrightarrow \frac{2}{\sqrt{3}} \qquad \text{as } \rho \to +\infty, \tag{4.43}$$

and

$$\rho^{-\sqrt{r(\rho)}+\frac{1}{2}}(\log\rho)^{\frac{3}{2}} \longrightarrow 0 \qquad \text{as } \rho \to +\infty. \tag{4.44}$$

Furthermore the definition of g yields

$$-\sqrt{g^{-1}(\rho)} + \log\rho = -\log\left(\sqrt{g^{-1}(\rho)} + \frac{e^{-\sqrt{g^{-1}(\rho)}}}{\rho}\right),$$

which in turn implies

$$\rho^{-\sqrt{r(\rho)}+1} = e^{-\sqrt{g^{-1}(\rho)}+\log\rho} = \frac{1}{\sqrt{g^{-1}(\rho)} + \frac{e^{-\sqrt{g^{-1}(\rho)}}}{\rho}}.$$

Applying again (3.16), the latter yields

$$\rho^{-\sqrt{r(\rho)}+1} \sim \frac{1}{\log \rho} \qquad \text{as } \rho \to +\infty$$

and thus

$$\rho^{-\sqrt{r(\rho)}+1}(\log \rho)^{\frac{3}{2}} \longrightarrow +\infty \qquad \text{as } \rho \to +\infty. \tag{4.45}$$

Now, plugging (4.43)-(4.44)-(4.45) into (4.42) yields

$$\rho^{-1} < v\big(4g^{-1}(\rho)\big) < \rho^{-1/2}, \qquad \rho \gg 1. \tag{4.46}$$

We now prove

$$v\big(4g^{-1}(\rho)\big) < \rho^{-1}, \qquad \rho \ll 1, \tag{4.47}$$

which is equivalent to

$$g(\eta)v(4\eta) < 1, \qquad 0 > \eta + \pi^2 \ll 1. \tag{4.48}$$

Noting that

$$g(\eta), \frac{1}{v(4\eta)} \longrightarrow 0 \qquad \text{as } \eta \to -\pi_+^2,$$

we study the limit of the derivatives. We find

$$\frac{d}{d\eta}g(\eta) = \frac{\sqrt{\eta}\cosh\sqrt{\eta} - \sinh\sqrt{\eta}}{2\eta^{3/2}} \longrightarrow \frac{1}{2\pi^2},$$

$$\frac{d}{d\eta}\frac{1}{v(4\eta)} = \frac{\sqrt{\frac{3}{2}}\left(-5\sqrt{\eta}\sinh\left(2\sqrt{\eta}\right) + 2(\eta+2)\cosh\left(2\sqrt{\eta}\right) - 4\right)}{4\eta\sqrt{\eta^2\left(\sqrt{\eta}\sinh\left(2\sqrt{\eta}\right) - \cosh\left(2\sqrt{\eta}\right) + 1\right)}} \longrightarrow +\infty,$$

as $\eta \to -\pi_+^2$. Therefore, applying L'Hôpital's rule yields

$$g(\eta)v(4\eta) \longrightarrow 0 \qquad \text{as } \eta \to -\pi_+^2,$$

which in turn implies (4.48) and thus (4.47). Finally, (4.46)-(4.47), together with the continuity of v and Weierstrass extreme value theorem, proves (4.19).

Eventually, analogous arguments allow to prove (4.20), using the representation

$$x_1^2 \partial_{x_1 x_1} \mathbf{u}(z;w) = \left(\rho\big(g^{-1}\big)'(\rho) + \rho^2\big(g^{-1}\big)''(\rho)\right)v'\big(4g^{-1}(\rho)\big) + \left(2\rho\big(g^{-1}\big)'(\rho)\right)^2 v''\big(4g^{-1}(\rho)\big), \qquad \rho = \frac{1}{\mathbf{h}(z;w)}.$$

The details are left to the reader for the sake of brevity. $\qquad\square$

4.5. Proof of Proposition 4.4

Proof of Proposition 4.4. For $T > 0$ we set

$$I(T) := \int_{(1,0)\prec\xi} g\big(\mathbf{h}(\mathbf{1}_{\mathbb{G}};T,\xi)\big)H_1(\mathbf{1}_{\mathbb{G}};T,\xi)\varphi(\xi)d\xi. \tag{4.49}$$

By (3.1)-(3.2), and by the change of variables

$$\begin{cases} \chi = \sqrt{\xi_1} \\ \eta = \mathbf{h}(\mathbf{1}_{\mathbb{G}};T,\xi) \end{cases} \qquad, \qquad (1,0) \prec \xi,$$

the integrals in (4.49) can be written as

$$I(T) = \frac{\sqrt{12}}{\pi T} \int_{\mathbb{R}^+ \times \mathbb{R}^+} \eta^{-2} \chi^{-2} g(\eta) \exp\left(-2 \frac{\eta(\chi + \chi^{-1} - 2) + G(\eta)}{T}\right) \bar{\varphi}(\chi, \eta) d\eta d\chi,$$

with

$$\bar{\varphi}(\chi, \eta) = \varphi\left(\chi^2, \frac{T\chi}{\eta}\right).$$

Employing now $G(1) = G'(1) = 0$ together with $G''(1) = 6$ (see (3.5)-(3.6) with $a_2 = 3$), we can write the 2nd order Taylor expansion

$$\eta(\chi + \chi^{-1} - 2) + G(\eta) = (\chi - 1)^2 + 3(\eta - 1)^2 + R(\chi, \eta),$$

where R is a continuous function such that

$$R(\chi, \eta) = o(\chi^2) + o(\eta^2), \qquad \text{as } (\chi, \eta) \to (1, 1).$$

Thus we have

$$I(T) = I_1(T) + I_2(T) + I_3(T),$$

with

$$I_1(T) = \frac{\sqrt{12}}{\pi T} \int_{([\frac{1}{2}, \frac{3}{2}])^2} \eta^{-2} \chi^{-2} g(\eta) \exp\left(-\frac{4(\chi-1)^2 + 12(\eta-1)^2}{2T}\right) \bar{\varphi}(\chi, \eta) d\eta d\chi,$$

$$I_2(T) = \frac{\sqrt{12}}{\pi T} \int_{([\frac{1}{2}, \frac{3}{2}])^2} \eta^{-2} \chi^{-2} g(\eta) \exp\left(-\frac{4(\chi-1)^2 + 12(\eta-1)^2}{2T}\right) \left(e^{-\frac{2R(\chi,\eta)}{T}} - 1\right) \bar{\varphi}(\chi, \eta) d\eta d\chi,$$

$$I_3(T) = \frac{\sqrt{12}}{\pi T} \int_{(\mathbb{R}^+)^2 \setminus ([\frac{1}{2}, \frac{3}{2}])^2} \eta^{-2} \chi^{-2} g(\eta) \exp\left(-2 \frac{\eta(\chi + \chi^{-1} - 2) + G(\eta)}{T}\right) \bar{\varphi}(\chi, \eta) d\eta d\chi.$$

We now study each term separately. Regarding I_1, it is enough to observe that the function

$$(\chi, \eta) \mapsto \frac{\sqrt{12}}{\pi T} \exp\left(-\frac{4(\chi-1)^2 + 12(\eta-1)^2}{2T}\right)$$

is a Gaussian probability density. Owing to the continuity of g and φ, and to the hypothesis $g(1) = 1$, it is standard to show that

$$\lim_{T \to 0^+} I_1(T) = \bar{\varphi}(1, 1) = \varphi(1, 0).$$

We now address I_2. By using

$$e^{-\frac{2R(\chi,\eta)}{T}} - 1 = o\big((\chi - 1)^2\big) + o\big((\eta - 1)^2\big), \qquad \text{as } (\chi, \eta) \to (1, 1),$$

it is again standard to prove that

$$\lim_{T \to 0^+} I_2(T) = 0.$$

We now employ

$$\frac{1}{T} \exp\left(-2 \frac{\eta(\chi + \chi^{-1} - 2) + G(\eta)}{T}\right) \longrightarrow 0 \quad \text{as } T \to 0^+, \qquad (\chi, \eta) \neq (1, 1),$$

together with

$$\eta(\chi + \chi^{-1} - 2) + G(\eta) \geq c > 0, \qquad (\chi, \eta) \in (\mathbb{R}^+)^2 \setminus ([1/2, 3/2])^2$$

and the fact that $\varphi \in C_b(D)$, to apply Lebesgue dominated convergence theorem so as to obtain

$$\lim_{T \to 0^+} I_3(T) = 0.$$

This proves (4.21) and concludes the proof. $\qquad \square$

5. Numerical evidences for Key Inequalities 4.5

In this section we collect numerical evidence in favor of the validity of the Key Inequalities 4.5. Unfortunately, so far, we were unable to provide a full mathematical proof. However, given the stability and the extensiveness of the numerical tests that we performed, we feel comfortable to conjecture that the Key Inequalities 4.5 are true statements. The completion of a rigorous proof seems challenging and remains an open problem, which we defer to further research.

From the numerical point of view, the two issues that one has to overcome in order to verify (4.22)-(4.23) are the following:

a) The cost function Ψ involves the inverse of a trigonometric function. Thus the complexity of the numerical inversion adds up to the complexity of the numerical integration. The result is a loss of precision and an increase of the computational time.

b) The integrals are to be computed on unbounded domains. This makes the numerical integration harder to manage.

To tackle Point a) above, we utilize the approximation of the cost function given by

$$\tilde{\Psi}_N(z,w) = \frac{4}{T-t}\left[\mathbf{h}(z,w)\left(\sqrt{\frac{y_1}{x_1}} + \sqrt{\frac{x_1}{y_1}} - 2\right) + \sum_{n=2}^{N} \tilde{G}_n\big(\mathbf{h}(z,w)\big)\right], \qquad N \geq 2,$$

where

$$\tilde{G}_n(\eta) := a_n \mathbf{1}_{]1,\infty[}(\eta)\frac{(\eta-1)^n}{\eta^{n-1}} + \tilde{b}_n \mathbf{1}_{]0,1[}(\eta)\frac{(-\log\eta)^n}{\left(1-\log\eta\right)^{n-1}}, \tag{5.1}$$

and where the coefficients $a_n, \tilde{b}_n$ are recursively determined by solving the equations

$$\frac{d^n}{d\eta^n}G(\eta)\Big|_{\eta=1} = \frac{d^n}{d\eta^n}\sum_{k=2}^{n}\tilde{G}_k(\eta)\Big|_{\eta=1}, \qquad n \geq 2. \tag{5.2}$$

The N-th order approximation $\tilde{\Psi}_N$ is clearly very close to the one in Definition 3.2. In fact, the functions $\tilde{G}_n$ coincide with the functions G_n on $]1,\infty[$, while on $]0,1[$ the two are only slightly different. The choice $\tilde{G}_n$ clearly does not preserve the asymptotic properties enjoyed by G_n described in Section 3. However, the advantage of using here $\tilde{\Psi}_N$ in place of Ψ_N is that the former approximates Ψ from below, namely

$$\tilde{\Psi}_N(z,w) \leq \Psi(z,w), \qquad z,w \in \mathbb{R} \times D, \ z \prec w. \tag{5.3}$$

Indeed, in Section 3 we have already showed that

$$\sum_{n=2}^{N} \tilde{G}_n(\eta) < G(\eta) \tag{5.4}$$

for any $\eta > 1$. This comes from the fact that the coefficients a_n are positive and G_n converges point-wise to G. On the other hand, the plots in Figure 4 provide numerical evidence of the fact that (5.4) is satisfied for $\eta \in]0,1[$ too. In particular, we test it for $N = 50$, which is the order used in the numerical tests below.

In order to address Point b) above, we make use of the Dirac Delta property of the function H_1 proved in Proposition 4.4, in order to reduce the integration domains to bounded sets. We have the following
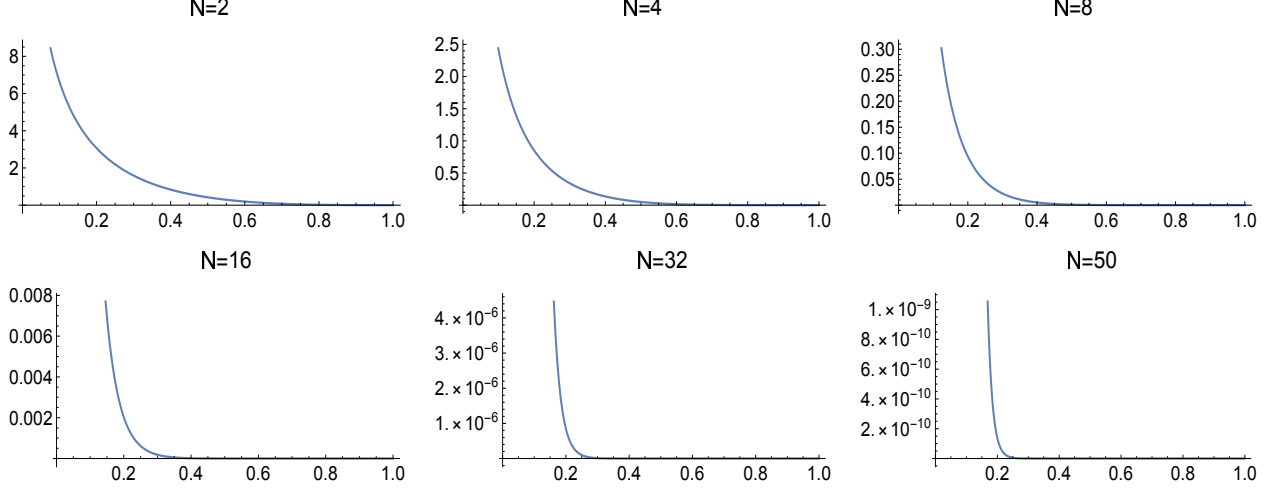
Figure 4: Plot of the difference $G(\eta) - \sum_{n=2}^{N} b_n \tilde{G}_n(\eta)$, with $\tilde{G}_n$ and $\tilde{b}_n$ as defined by (5.1)-(5.2).

Proposition 5.1. *The Key Inequalities 4.5 are satisfied if and only if there exists a constant $\kappa_\tau > 0$ such that*

$$\int_{x \prec \xi \prec (1,0)} \mathbf{1}_{D(t,s,x)}(\xi)\sqrt{\mathbf{h}(z;s,\xi)}H_1(z;s,\xi)\sqrt{\mathbf{h}(s,\xi;\mathbf{1}_\mathbb{G})}H_1(s,\xi;\mathbf{1}_\mathbb{G})d\xi \leq \kappa_\tau \sqrt{\mathbf{h}(z;\mathbf{1}_\mathbb{G})}H_1(z,\mathbf{1}_\mathbb{G}),$$

$$\int_{x \prec \xi \prec (1,0)} \mathbf{1}_{D'(t,s,x)}(\xi)\tilde{g}\big(\mathbf{h}(z;s,\xi)\big)H_1(z;s,\xi)\tilde{g}\big(\mathbf{h}(s,\xi;\mathbf{1}_\mathbb{G})\big)H_1(s,\xi;\mathbf{1}_\mathbb{G})d\xi \leq \kappa_\tau \tilde{g}\big(\mathbf{h}(z;\mathbf{1}_\mathbb{G})\big)H_1(z,\mathbf{1}_\mathbb{G}),$$

for any $s \in]t,0[$, and for any $z = (t,x) \in \mathbb{R} \times D$ such that $z \prec \mathbf{1}_\mathbb{G}$ and $t > -\tau$, where $D(x), D'(x) \subset D$ are the bounded domains

$$D(t,s,x) := \Big\{\xi \in D : \Big(\sqrt{\mathbf{h}(z;s,\xi)}H_1(z;s,\xi)\Big) \wedge \Big(\sqrt{\mathbf{h}(s,\xi;\mathbf{1}_\mathbb{G})}H_1(s,\xi;\mathbf{1}_\mathbb{G})\Big) > \sqrt{\mathbf{h}(z;\mathbf{1}_\mathbb{G})}H_1(z,\mathbf{1}_\mathbb{G})\Big\},$$

$$\tag{5.5}$$

$$D'(t,s,x) := \Big\{\xi \in D : \Big(\tilde{g}\big(\mathbf{h}(z;s,\xi)\big)H_1(z;s,\xi)\Big) \wedge \Big(\tilde{g}\big(\mathbf{h}(s,\xi;\mathbf{1}_\mathbb{G})\big)H_1(s,\xi;\mathbf{1}_\mathbb{G})\Big) > \tilde{g}\big(\mathbf{h}(z;\mathbf{1}_\mathbb{G})\big)H_1(z,\mathbf{1}_\mathbb{G})\Big\}.$$

Proof. It plainly follows from (5.5) and Proposition 4.4. □

5.1. Numerical tests and results

In force of (5.3) and of Proposition 5.1 above, in order to prove the Key Inequalities 4.5 it is enough to check that

$$I_1(t,s,x) := \frac{\int_{]x_2,0[\times]0,\bar{\xi}[} \sqrt{\mathbf{h}(z;s,\xi)}\tilde{H}_1(z;s,\xi)\sqrt{\mathbf{h}(s,\xi;\mathbf{1}_\mathbb{G})}\tilde{H}_1(s,\xi;\mathbf{1}_\mathbb{G})d\xi}{\sqrt{\mathbf{h}(z;\mathbf{1}_\mathbb{G})}H_1(z,\mathbf{1}_\mathbb{G})} \leq \kappa_\tau, \tag{5.6}$$

$$I_2(t,s,x) := \frac{\int_{]x_2,0[\times]0,\bar{\bar{\xi}}[} (\sqrt{\mathbf{h}}+\mathbf{h})(z;s,\xi)\tilde{H}_1(z;s,\xi)(\sqrt{\mathbf{h}}+\mathbf{h})(s,\xi;\mathbf{1}_\mathbb{G})\tilde{H}_1(s,\xi;\mathbf{1}_\mathbb{G})d\xi}{(\sqrt{\mathbf{h}}+\mathbf{h})(z;\mathbf{1}_\mathbb{G})H_1(z,\mathbf{1}_\mathbb{G})} \leq \kappa_\tau, \tag{5.7}$$

for any $z = (t,x) \in \mathbb{R} \times D$ such that $z \prec \mathbf{1}_\mathbb{G}$ and $t > -\tau$, where $\tilde{H}_1$ is defined as H_1 with Ψ replaced by $\tilde{\Psi}_N$ for a given $N > 2$, and where $\bar{\xi}, \bar{\bar{\xi}} > 0$ are such that

$$]x_2,0[\times]0,\bar{\xi}[\supset D(t,s,x) \cap \big(]x_2,0[\times\mathbb{R}^+\big), \qquad]x_2,0[\times]0,\bar{\bar{\xi}}[\supset D'(t,s,x) \cap \big(]x_2,0[\times\mathbb{R}^+\big).$$

The scope for turning the domains of integration into rectangles is facilitating the numerical integration.

n	t^*	s^*	x_1^*	x_2^*	$\log \chi^*$	$\log \eta^*$	$I_n(t^*, s^*, x^*)$
1	-0.81948	-0.54040	43.64581	-0.12497	1.88805	3.76863	1.74841
2	-0.78426	-0.67014	0.00002	-0.00605	-5.52146	-0.54472	2.48050

Table 2: The maxima $I_n(t^*, s^*, x^*) = \max_{t,s,x} I_n(t, s, x)$ and relative argmax, with t, s, x ranging over a batch counting 2×10^7 samples generated at random in accordance with (5.8)-(5.9).

We provide numerical evidence of the fact that (5.6)-(5.7) are satisfied for $\tau = 1$, with $\kappa_\tau \approx 2.5$ and $N = 50$. We test the inequalities for multiple choices of the parameters s, t, x, generated at random according to the following distributions:

$$t \sim \text{Unif}_{[-1,0]}, \qquad s \sim \text{Unif}_{[t,0]}, \tag{5.8}$$

and

$$x_1 = \chi^2, \qquad x_2 = \frac{t\chi}{\eta}, \qquad \text{with } (\chi, \eta) \sim \text{LogNorm}_{0,2} \otimes \text{LogNorm}_{0,2}. \tag{5.9}$$

The distribution above is justified by the representation (3.1), which can be written here as

$$\Psi(t, x; \mathbf{1}_{\mathbb{G}}) = \frac{4}{-t}\left(\eta\left(\chi + \frac{1}{\chi} - 2\right) + G(\eta)\right), \qquad \eta = \mathbf{h}(t, x; \mathbf{1}_{\mathbb{G}}) = \frac{t\sqrt{x_1}}{x_2}, \ \chi = \sqrt{x_1},$$

and which shows that Ψ reaches its minimum at $(\chi, \eta) = (1, 1)$. We generated 2×10^7 vectors (s, t, x) and computed the integrals in the LHS of (5.6)-(5.7) using Monte Carlo integration. In Table 2 we report the maximum value of $I_1(t, s, x), I_2(t, s, x)$ and the corresponding arg max obtained for the batch of variables.

The results do not present evidence of the fact that the functions I_1 and I_2 are unbounded. Note that, due to the chosen distributions for (η, χ) and (t, s), the values of I_1 and I_2 were computed in a region that includes the extreme tails of the exponential kernel $e^{-\Psi(t,x;\mathbf{1}_{\mathbb{G}})/2}$, and hence in a region where the denominators in (5.6)-(5.7) can be very small. For instance, for $t = -1/2, s = -1/4$ and $\chi = \xi = e^8$ one has $e^{-\Psi(t,x;\mathbf{1}_{\mathbb{G}})/2} \approx 3.24 \times 10^{-126}$.

The numerical integration was performed using Wolfram Mathematica built-in numerical integration routine `NIntegrate` with the method `AdaptiveMonteCarlo`. The Mathematica notebook used to generate the results reported above can be found in the supplementary material.

Appendix A. A topological lemma

Let $B = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$ denote the unit open ball of $\mathbb{R}^2$, and let $f : \overline{B} \to \mathbb{R}^2$ be a continuous map such that $f|_{\partial B}$ is a homeomorphism. In accordance with the Jordan-Schoenflies separation theorem, in the sequel $f(\partial B)^*$ will denote the "inside" of $f(\partial B)$, meaning the bounded set of $\mathbb{R}^2$, whose border is $f(\partial B)$, homeomorphic to an open ball. We will also denote by $f(\partial B)_*$ the "outside" of $f(\partial B)$, namely the complementary of $f(\partial B) \cup f(\partial B)^*$, which is homeomorphic to the complementary of a closed ball.

Lemma Appendix A.1. *Let B be an open ball of $\mathbb{R}^2$, and let $f : \overline{B} \to \mathbb{R}^2$ be such that:*

(i) f is a local homeomorphism;

(ii) $f|_{\partial B}$ is a homeomorphism.

Then $f|_B$ is a homeomorphism between B and $f(\partial B)^*$.

Proof. We first prove that

$$f(B) = f(\partial B)^*. \tag{A.1}$$

Assume that $f(\partial B)_* \cap f(\overline{B}) \neq \emptyset$. Since $f(\overline{B})$ is a compact set, we also have $f(\partial B)_* \cap \left(f(\overline{B}) \right)^c \neq \emptyset$. Therefore, being $f(\partial B)_*$ an open connected set, we have $f(\partial B)_* \cap \partial \left(f(\overline{B}) \right) \neq \emptyset$. Let $y \in f(\partial B)_* \cap \partial \left(f(\overline{B}) \right)$, again by compactness of $f(\overline{B})$ there exists $x \in \overline{B}$ such that $f(x) = y$. Note that $x \in B$, as $y \in f(\partial B)_*$. However, by (i), there exists a neighborhood U of x such that $f(U)$ is a neighborhood of y, which contradicts the fact that $y \in \partial \left(f(\overline{B}) \right)$. We thus proved that $f(\overline{B}) \subset \overline{f(\partial B)^*}$.

An analogous argument shows that $f(\overline{B}) \supset \overline{f(\partial B)^*}$. Assume by contradiction that $f(\overline{B}) \cap f(\partial B)_* \neq \emptyset$. Then $f(\partial B)^* \cap \partial \left(f(\overline{B}) \right) \neq \emptyset$, and the conclusion follows. Thus,

$$f(\overline{B}) = \overline{f(\partial B)^*}.$$

Moreover, if there were $x \in B$ such that $f(x) \in f(\partial B)$, then (i) would be violated because $f(\partial B) = \partial f(B)$. This proves (A.1).

We now note that the compact subsets in the subspace topologies on B and $f(\partial B)^*$ are all the closed subsets of $\mathbb{R}^2$ contained in B and $f(\partial B)^*$, respectively. Therefore, $f|_B : B \to f(\partial B)^*$ is a proper function.

Now, in order to conclude the proof it is enough to apply Hadamard-Caccioppoli Theorem, a particular instance of which states that a local homeomorphism between two open and simply connected sets of $\mathbb{R}^n$ is a global homeomorphism if and only if it is a proper function. $\qquad\square$

Appendix B. Proof of Lemma 4.1

Proof of Lemma 4.1. First note that, by (2.9), it is enough to prove (4.9) for $w = \mathbf{1}_{\mathbb{G}}$. Now, we observe that it suffices to prove that

$$0 = \inf_{\omega \in \mathbb{R}} \left\{ \omega^2 + \left(\omega x_1 \partial_{x_1} + Y \right) \Psi(z, \mathbf{1}_{\mathbb{G}}) \right\}, \qquad z = (t, x) \in \mathbb{R} \times D, \ z \prec \mathbf{1}_{\mathbb{G}}. \tag{B.1}$$

Indeed, the function

$$\omega^2 + \left(\omega x_1 \partial_{x_1} + Y \right) \Psi(z, \mathbf{1}_{\mathbb{G}})$$

has a global minimum at $\omega = -\frac{x_1}{2} \partial_{x_1} \Psi(z, \mathbf{1}_{\mathbb{G}})$, which is

$$-\left(\frac{x_1}{2} \partial_{x_1} \Psi(z, \mathbf{1}_{\mathbb{G}}) \right)^2 + Y \Psi(z, \mathbf{1}_{\mathbb{G}}).$$

Consider now the following extended control problem. For any $z = (t, x) \in \mathbb{R} \times D$ with $z \prec \mathbf{1}_{\mathbb{G}}$, and for any $x_0 \in \mathbb{R}$, find

$$\bar{\Psi}(t, x_0, x_1, x_2) = \min_{\omega} \gamma_0(T),$$

where the minimum is taken over all the controls $\omega \in L^2([t, T])$ for which (1.8) with $w = \mathbf{1}_{\mathbb{G}}$ are satisfied, and

$$\dot{\gamma}_0(s) = \omega^2(s), \quad t < s < T, \qquad\qquad \gamma_0(t) = x_0.$$

It is trivial to see that ω is optimal for this problem if and only if it is optimal for the problem (1.7)-(1.8). Also, we have

$$\bar{\Psi}(t, x_0, x_1, x_2) = \Psi(t, x_1, x_2; \mathbf{1}_{\mathbb{G}}) + x_0.$$

Therefore, (B.1) is equivalent to

$$0 = \inf_{\omega \in \mathbb{R}} \left\{ \left(\omega^2 \partial_{x_0} \bar{\Psi}(t, x_0, x) + \omega x_1 \partial_{x_1} + Y \right) \bar{\Psi}(t, x_0, x) \right\}, \qquad z = (t, x) \in \mathbb{R} \times D, \ z \prec \mathbf{1}_{\mathbb{G}}, \quad x_0 \in \mathbb{R}.$$

Eventually, the latter holds true by Theorem IV-4.1 in [42], which applies to the our problem as $\Psi(\cdot, \mathbf{1}_{\mathbb{G}})$ is smooth on its domain and the optimal control ω is piecewise continuous. Theorem IV-4.1 in [42] also states (4.10), and this concludes the proof.

$\square$

References

[1] S. R. S. Varadhan, On the behavior of the fundamental solution of the heat equation with variable coefficients, Communications on Pure and Applied Mathematics 20 (2) (1967) 431–455.

[2] R. Léandre, Majoration en temps petit de la densité d'une diffusion dégénérée, Probability theory and related fields 74 (2) (1987) 289–294.

[3] R. Léandre, Minoration en temps petit de la densité d'une diffusion dégénérée, Journal of Functional Analysis 74 (2) (1987) 399–414.

[4] M. Yor, On some exponential functionals of Brownian motion, Adv. in Appl. Probab. 24 (3) (1992) 509–531. doi:10.2307/1427477.
URL https://doi.org/10.2307/1427477

[5] H. Matsumoto, M. Yor, et al., Exponential functionals of brownian motion, i: Probability laws at fixed time, Probability surveys 2 (2005) 312–347.

[6] G. Cibelli, S. Polidoro, F. Rossi, Sharp estimates for Geman–Yor processes and applications to arithmetic average asian options, Journal de Mathématiques Pures et Appliquées 129 (2019) 87–130.

[7] K. Yosida, On the fundamental solution of the parabolic equation in a riemannian space, Osaka Mathematical Journal 5 (1) (1953) 65–74.

[8] A. Friedman, Partial differential equations of parabolic type, Prentice-Hall Inc., Englewood Cliffs, N.J., 1964.

[9] D. G. Aronson, Bounds for the fundamental solution of a parabolic equation, Bulletin of the American Mathematical society 73 (6) (1967) 890–896.

[10] S. R. Varadhan, Diffusion processes in a small time interval, Communications on Pure and Applied Mathematics 20 (4) (1967) 659–685.

[11] S. A. Molchanov, Diffusion processes and riemannian geometry, Russian Mathematical Surveys 30 (1) (1975) 1.

[12] R. Azencott, Densité des diffusions en temps petit: développements asymptotiques. I, in: Seminar on probability, XVIII, Vol. 1059 of Lecture Notes in Math., Springer, Berlin, 1984, pp. 402–498. doi:10.1007/BFb0100057.
URL https://doi-org.ezproxy.unibo.it/10.1007/BFb0100057

[13] A. D. Ventsel', M. I. Freidlin, On small random perturbations of dynamical systems, Russian Mathematical Surveys 25 (1) (1970) 1–55.

[14] P. K. Friz, J. Gatheral, A. Gulisashvili, A. Jacquier, J. Teichmann, Large deviations and asymptotic methods in finance, Vol. 110, Springer, 2015.

[15] G. Ben Arous, Développement asymptotique du noyau de la chaleur hypoelliptique hors du cut-locus, in: Annales scientifiques de l'Ecole normale supérieure, Vol. 21, 1988, pp. 307–331.

[16] D. Barilari, U. Boscain, R. W. Neel, Heat kernel asymptotics on sub-riemannian manifolds with symmetries and applications to the bi-heisenberg group, in: Annales de la Faculté des sciences de Toulouse: Mathématiques, Vol. 28, 2019, pp. 707–732.

[17] S. De Marco, P. K. Friz, Local volatility, conditioned diffusions, and varadhan's formula, SIAM Journal on Financial Mathematics 9 (2) (2018) 835–874.

[18] A. Bellaïche, C. BELLAICHE, P. BOUGEROL, M. CHALEYAT-MAUREL, L. ELIE, J. GRANARA, géodésiques et diffusions en temps petit, Astérisque 84 (1981).

[19] A. Pascucci, PDE and martingale methods in option pricing, Springer Science & Business Media, 2011.

[20] D. Barilari, F. Boarotto, Kolmogorov–fokker–planck operators in dimension two: heat kernel and curvature, Journal of Evolution Equations 18 (3) (2018) 1115–1146.

[21] J. Gatheral, E. P. Hsu, P. Laurence, C. Ouyang, T.-H. Wang, Asymptotics of implied volatility in local volatility models, Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics 22 (4) (2012) 591–620.

[22] S. Polidoro, On a class of ultraparabolic operators of kolmogorov-fokker-planck type, Le Matematiche 49 (1) (1994) 53–105.

[23] M. D. Francesco, A. Pascucci, On a class of degenerate parabolic equations of kolmogorov type, Applied Mathematics Research eXpress 2005 (3) (2005) 77–116.

[24] F. Delarue, S. Menozzi, Density estimates for a random noise propagating through a chain of differential equations, J. Funct. Anal. 259 (6) (2010) 1577–1630. `doi:10.1016/j.jfa.2010.05.002`.
URL `https://doi.org/10.1016/j.jfa.2010.05.002`

[25] H. Geman, M. Yor, Quelques relations entre processus de Bessel, options asiatiques et fonctions confluentes hypergéométriques, C. R. Acad. Sci. Paris Sér. I Math. 314 (6) (1992) 471–474.

[26] M. Fu, D. Madan, T. Wang, Pricing continuous time Asian options: a comparison of Monte Carlo and Laplace transform inversion methods, J. Comput. Finance 2 (1998) 49–74.

[27] D. Dufresne, Asian and Basket asymptotics, Research paper n.100, University of Montreal (2002).

[28] D. Dufresne, Laguerre series for Asian and other options, Math. Finance 10 (4) (2000) 407–428. `doi:10.1111/1467-9965.00101`.
URL `http://dx.doi.org/10.1111/1467-9965.00101`

[29] W. T. Shaw, Pricing Asian options by contour integration, including asymptotic methods for low volatility, Working paper,
http://www.mth.kcl.ac.uk/∼shaww/web_page/papers/Mathematicafin.htm (2003).

[30] V. Linetsky, The spectral decomposition of the option value, International Journal of Theoretical and Applied Finance 7 (3) (2004) 337–384.

[31] P. Guasoni, S. Robertson, Optimal importance sampling with explicit formulas in continuous time, Finance Stoch. 12 (1) (2008) 1–19. doi:10.1007/s00780-007-0053-5.
URL http://dx.doi.org/10.1007/s00780-007-0053-5

[32] J. E. Ingersoll, Theory of Financial Decision Making, Blackwell, Oxford, 1987.

[33] J. N. Dewynne, W. T. Shaw, Differential equations and asymptotic solutions for arithmetic Asian options: 'Black-Scholes formulae' for Asian rate calls, European J. Appl. Math. 19 (4) (2008) 353–391. doi:10.1017/S095679250800750X.
URL http://dx.doi.org/10.1017/S095679250800750X

[34] N. C. Caister, J. G. O'Hara, K. S. Govinder, Solving the Asian option PDE using Lie symmetry methods, Int. J. Theor. Appl. Finance 13 (8) (2010) 1265–1277. doi:10.1142/S0219024910006194.
URL http://dx.doi.org/10.1142/S0219024910006194

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

[36] S. Pagliarani, A. Pascucci, M. Pignotti, Intrinsic expansions for averaged diffusion processes, Stochastic Processes and their Applications 127 (8) (2017) 2560–2585.

[37] E. Gobet, M. Miri, Weak approximation of averaged diffusion processes, Stochastic Processes and their Applications 124 (1) (2014) 475–504.

[38] K. Shiraya, A. Takahashi, Pricing average options on commodities, Journal of Futures Markets 31 (5) (2011) 407–439.

[39] K. Shiraya, A. Takahashi, M. Toda, Pricing Barrier and Average Options Under Stochastic Volatility Environment, SSRN eLibrary (2009).

[40] D. Pirjol, L. Zhu, Short maturity asian options in local volatility models, SIAM Journal on Financial Mathematics 7 (1) (2016) 947–992.

[41] L. Monti, A. Pascucci, Obstacle problem for arithmetic asian options, Comptes Rendus Mathematique 347 (23-24) (2009) 1443–1446.

[42] W. H. Fleming, R. W. Rishel, Deterministic and stochastic optimal control, Vol. 1, Springer Science & Business Media, 2012.