Small-$\triangle$-optimal range-based estimation for diffusions

Hartmuth Henkel * Markus Reiß†

October 3, 2010

Abstract: This paper extends the theory of small-Δ-optimality for ordinary martingale estimating functions that are constructed by means of equidistant observations to a situation where the maxima and the minima of the observation intervals are added to the sample. The sampling frequency is denoted with Δ. Second order expansions of the expression $\mathbb{E}_x[g(H_\Delta, L_\Delta, X_\Delta)]$ with respect to $\sqrt{\Delta}$ are used to establish lower bounds of the variance for different classes of estimating functions. The case of Ornstein-Uhlenbeck processes is studied in detail.

KEY WORDS: Martingale estimating functions; Range–based estimation; Small–Δ–optimality; Joint distribution of the running maximum, the running minimum and the terminal value of diffusions.

MSC: 62M05, 60J60

*corresponding author, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, hartmuth.henkel@gmx.de

†Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, mreiss@math.hu-berlin.de

1 Introduction

Range-based parameter estimation methods are developed for diffusion processes. Existing inference methods for parameterized stochastic differential equations almost exclusively rely on (equidistant) discrete observations of the underlying process, whereas in real world applications sometimes additional information – such as the maximum or the minimum on disjoint intervals – is available. Financial assets for example are usually traded at an extremely high frequency and hence the aforementioned statistics can easily be determined. It is not trivial to incorporate the maxima and minima into the analysis, though. For the basic case of Brownian motion with drift several range-based estimators for the diffusion coefficient have already been proposed in the literature: for example the *Parkinson estimator*, see [15], the *Rogers and Satchell estimator*, see [16] and the *Garman-Klass estimator*, see [5]. These estimators have in common that they are moment-type estimators that rely on elementary properties of Brownian motion. Magdon-Ismail and Atiya [14] made use of an explicit representation of the joint density of the running maximum, the running minimum and the terminal value, that was originally derived by Dominé [3], to simulate the maximum likelihood estimator (MLE) and compared it to the aforementioned moment estimators.

When it comes to models other than Brownian motion, a range-based ansatz is particularly challenging because for more general processes the joint distributions is usually not known explicitly. Indeed only few properties of the required joint distribution or the joint density for diffusions are known at all. It is noteworthy that, apart from Dominé's result, a closed form expansion of the joint density can also be calculated for the Ornstein-Uhlenbeck process, see Sweet and Hardin [20]. In our analysis we will consider a diffusion model defined by

$$dX_t = \mu(X_t;\theta)dt + \sigma(X_t;\theta)dB_t, \quad X_0 = U,\ t \geq 0, \tag{1}$$

where B denotes the Brownian motion of $\mathbb{R}$ and U denotes the initial distribution. Moreover, $\mu : \mathbb{R} \times \Theta \to \mathbb{R}$ and $\sigma : \mathbb{R} \times \Theta \to \mathbb{R}_+$ are supposed to be sufficiently smooth functions parameterized by $\theta \in \Theta \subset \mathbb{R}$. Although our upcoming results can easily be proved for multi-dimensional parameters, for the ease of presentation, we will stick to the one-dimensional case throughout this paper.

It is straightforward to generalize existing results for martingale estimating functions (MEF) constructed from equidistant observations of the process X proposed by the authors Hansen and Scheinkmann [7], Kessler [13] and especially Sørensen [18]. One obtains results about consistency and asymptotic normality for estimators inferred from MEFs that are constructed from the sample

$$\left(\bar{H}_{(i\Delta)}, \bar{L}_{(i\Delta)}, X_{i\Delta}\right)_{i=1,\dots,n} = \left(\sup_{(i-1)\Delta \leq t \leq i\Delta} X_t, \inf_{(i-1)\Delta \leq t \leq i\Delta} X_t, X_{i\Delta}\right)_{i=1,\dots,n}, \tag{2}$$

on equidistant intervals $((i-1)\Delta, i\Delta]$, $\Delta > 0$, as the number of observations $n \in \mathbb{N}$ tends to infinity – see Section 2. However, these findings are highly theoretical for the calculation of the asymptotic variance still requires the knowledge of the joint moments. Of course, in a diffusion model the joint densities can always be approximated by means of numerical methods for PDEs with boundary conditions or simply by extensive simulations of the underlying diffusion. But both methods are quite computer intensive. This is why further simplifications are desirable. The means of choice in our analysis will be a small-Δ-approach similar to the one of Jacobsen, see [9] and [10]. For an overview of the existing methods concerning martingale estimating functions and the relation to small-Δ-optimality, we suggest [1].

It turns out that for certain diffusion processes it is possible to derive a second order expansion of the expression $\mathbb{E}_x[g(H_\Delta, L_\Delta, X_\Delta)]$ with respect to $\sqrt{\Delta}$, where

$$(H_\Delta, L_\Delta, X_\Delta) = \left(\sup_{0 \le t \le \Delta} X_t, \inf_{0 \le t \le \Delta} X_t, X_\Delta \right), \tag{3}$$

and $g : \mathbb{R}^3 \to \mathbb{R}$ must be a sufficiently smooth function that does not grow too fast. This expansion, which entirely relies on elementary estimates, is sufficient to establish the asymptotic theory for martingale estimating functions that are based on a fixed number of observations of the triplet (3) when the parameter Δ tends to 0. In particular, this proceeding helps us to avoid the complexity issues that arise when dealing with the MLE or even standard MEFs. Nevertheless our results imply that in the limit the simplified small-Δ-optimal estimators experience no loss in efficiency. It is remarkable that for parameter-dependent diffusion coefficients $\sigma(\cdot;\theta)$ the range-based approach yields a significant improvement, up to 80% for the mean squared error. The drift estimation $\mu(\cdot;\theta)$, however, does not benefit from the generalization and small-Δ-optimal generalized martingale estimating functions reduce to the ordinary case.

This paper is structured as follows: In Section 2 we describe the model we intend to work with and quote some theoretical foundations that we are going to use in the subsequent sections. In Section 3 and Section 4 we present the main small-Δ-optimality results for generalized martingale estimating functions. A simulation study in Section 5 corroborates our theoretical findings. Lastly, in Section 6 we concisely prove the technical results fundamental to the analysis of the Sections 3 and 4.

2 Theoretical Foundations for generalized small-Δ-optimal MEFs

2.1 Martingale Estimating Functions for Diffusion Processes

Let us consider a process of the form (1). Formally, the diffusion X is defined on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t)$ with U $\mathcal{F}_0$-measurable. We assume that, for any $\theta \in \Theta$ and any probability measure ν on $\mathbb{R}$, there is a probability measure $\mathbb{P}_{\nu,\theta}$ on $(\Omega, \mathcal{F})$, with respect to which the σ-algebra $\mathcal{F}_0$ and the Brownian motion B are independent and such that, for the prescribed θ-value, the equation (1) has a unique strong solution with

ν being the distribution of U. Accordingly, for $X_0 = x \in \mathbb{R}$, i.e. if ν equals the Dirac measure δ_x, we will denote the corresponding Markov measure with $\mathbb{P}_{x,\theta}$.

First we specify some regularity assumptions for the coefficients μ and σ.

Assumption 2.1.1. *For all $\theta \in \Theta$, $\mu(x;\theta)$ is continuous in x and, for each $x \in \mathbb{R}$, $\mu(x;\theta)$ is continuously differentiable in θ. The function $\sigma(x;\theta)$ is supposed to be continuously differentiable in (x,θ) and to be uniformly bounded away from 0 in (x,θ).*

The theory of generalized martingale estimating functions extends the concept of the maximum likelihood estimator (MLE). The MLE solves the score equation $U_n(\theta) = 0$, where

$$U_n(\theta) = \sum_{i=1}^{n} \partial_\theta \log f\big(\Delta_i, X_{(i-1)\Delta}, \bar{H}_{(i\Delta)}, \bar{L}_{(i\Delta)}, X_{i\Delta}; \theta\big). \tag{4}$$

Here, $(h,l,y) \mapsto f(\Delta, x, h, l, y; \theta)$ denotes the joint density of the vector (3) conditional on $X_0 = x$ with $\bar{H}_{(i\Delta)}$ and $\bar{L}_{(i\Delta)}$, $i = 1,\dots,n$, defined as in (2). From the Markov property it can be inferred that, on mild regularity assumptions, $U_n(\theta)$ is a martingale with respect to the filtration $\mathcal{F}_n = \sigma(X_\Delta, \dots, X_{n\Delta})$.

The idea behind MEFs is to replace $U_n(\theta)$ by another martingale

$$G_n(\theta) = \sum_{i=1}^{n} g\big(\Delta_i, X_{(i-1)\Delta}, \bar{H}_{(i\Delta)}, \bar{L}_{(i\Delta)}, X_{i\Delta}; \theta\big), \tag{5}$$

where

$$g(\Delta, x, h, l, y; \theta) = \sum_{j=1}^{J} a_j(\Delta, x; \theta) k_j(\Delta, x, h, l, y; \theta), \tag{6}$$

and the real valued functions k_j, $j = 1,\dots,J$, must satisfy

$$\mathbb{E}_{x,\theta}\Big[k_j(\Delta, X_0, H_\Delta, L_\Delta, X_\Delta; \theta)\Big] = 0, \quad \forall x \in \mathbb{R}, \tag{7}$$

in order to make $G_n(\theta)$ a martingale. The associated estimator $\hat{\theta}_n$ is given by any root of $G_n(\theta)$. We will usually refer to estimating functions of the type (6) as *generalized martingale estimating functions*, whereas functions that are constructed from the equidistant observations $(X_{i\Delta},\, i = 1,\dots,n)$ alone will be called *ordinary martingale estimating functions.*

In generalization of the ordinary case (Theorem 3.6 in [19]), Theorem 4.2.1.7 in [8] states that for regular functions g and on the additional assumption of ergodicity an estimator $\hat{\theta}_n$ for θ, inferred from a generalized martingale estimating function (5), is consistent and asymptotically normally distributed as the number of observations $n \to \infty$; more specifically under the true measure $\mathbb{P}_{\nu,\theta_0}$ we have

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N\left(0, \frac{v(\theta_0)}{\xi^2(\theta_0)}\right), \quad n \to \infty. \tag{8}$$

In this context, ν denotes the invariant measure and

$$v(\theta_0) = \mathbb{E}_{\nu,\theta_0}\big[g(\Delta, X_0, H_\Delta, L_\Delta, X_\Delta; \theta_0)^2\big], \tag{9}$$
$$\xi(\theta_0) = \mathbb{E}_{\nu,\theta_0}\big[\partial_\theta g(\Delta, X_0, H_\Delta, L_\Delta, X_\Delta; \theta_0)\big]. \tag{10}$$

In the framework of (ordinary or generalized) martingale estimating functions, one chooses optimal weights $(a_1^*, \dots, a_J^*)$ according to the optimality criteria of Godambe and Heyde [6] such that the asymptotic variance

$$\mathrm{Var}_{\nu,\Delta,\theta}(g) := \frac{v(\theta)}{\xi^2(\theta)}, \tag{11}$$

is minimized for a fixed $\Delta > 0$. By contrast, for the discussion of small-Δ-optimality, one considers $\mathrm{Var}_{\nu,\Delta,\theta}$ for a fixed sample size and for $\Delta \to 0$. We will show that

$$\mathrm{Var}_{\Delta,\nu,\theta}(g) = \frac{1}{\Delta}\mathcal{V}_{-1,\theta}(g) + \mathcal{V}_{0,\theta}(g) + o\,(1)\,, \quad \Delta \to 0, \tag{12}$$

and we will call an estimating function g small-Δ-optimal within a class of estimating functions $\mathcal{G}_\theta$ if the first term in this expansion, $\mathcal{V}_{-1}$ or $\mathcal{V}_0$, is minimized over $\mathcal{G}_\theta$. Depending on the structure of the diffusion model, we will distinguish two cases in our analysis:

(1) minimizing $\mathcal{V}_{-1}$: $\sigma(\cdot;\theta) = \sigma(\cdot)$ does not depend on θ,

(2) minimizing $\mathcal{V}_0$ with $\mathcal{V}_{-1} \equiv 0$: $\sigma(\cdot;\theta)$ depends on θ.

Consequently, for both the drift and the volatility estimation, in the limit a universal lower bound for the asymptotic variance can be obtained. This implies that, for small values of Δ, an estimator obtained from a small-Δ-optimal estimating function is in practice as good as the maximum likelihood estimator. To sum up, small-Δ-optimality is a global optimality criterion and, although small-Δ-optimality refers explicitly to the limit $\Delta \to 0$, for any fixed $\Delta > 0$ the estimator is still $\sqrt{n}$-consistent and asymptotically Gaussian for $n \to \infty$. Of course, there is no guarantee that the aforementioned estimators for fixed Δ are Godambe and Heyde optimal, but for Δ not too large, they should still behave well.

2.2 Notation and assumptions

Assumption 2.2.1.
(i) We assume that, if the coefficient σ depends on θ, then

$$\sigma(x;\theta) = r(\theta) \cdot \sigma(x) \quad \forall \theta \in \Theta,\ \forall x \in \mathbb{R}, \tag{13}$$

where $r : \Theta \to \mathbb{R}_+$ *is a once and* $\sigma : \mathbb{R} \to \mathbb{R}_+$ *is a twice continuously differentiable function.*

(ii) For each $\theta \in \Theta$, *the function* $\mu^Y(\cdot;\theta) : \mathbb{R} \to \mathbb{R}$ *defined by the mapping* $y \mapsto \mu^Y(y;\theta) = (\mu/\sigma - \sigma'/2)(y;\theta)$, *satisfies a global Lipschitz condition. Moreover, for all* $x \in \mathbb{R}$, $\mu^Y(x;\theta)$ *is continuously differentiable in* θ.

Assumption 2.2.1 *(i)* allows to consider the Lamperti transform of X given by

$$Y_t = \sigma(X_0;\theta) \int_{y_0}^{X_t} \frac{1}{\sigma(u;\theta)} du = \sigma(X_0) \int_{y_0}^{X_t} \frac{1}{\sigma(u)} du, \tag{14}$$

for some value $y_0 \in \mathbb{R}$, independently from the parameter θ. Let F denote the primitive $\int_{y_0}^{\cdot} 1/\sigma(u)du$ of $1/\sigma(\cdot)$. Without loss of generality, let us assume that the starting point of integration y_0 in (14) satisfies

$$y_0 = F^{-1}\left(F(X_0) - \frac{X_0}{\sigma(X_0)} \right). \tag{15}$$

Then, the Lamperti transform Y of the process X starting in $X_0 = x$ starts in x as well and it satisfies the following stochastic differential equation

$$dY_t = \sigma(X_0;\theta) \left(\frac{\mu(Y_t;\theta)}{\sigma(Y_t;\theta)} - \frac{1}{2} \frac{\partial}{\partial y} \sigma(y;\theta) \Big|_{y=Y_t} \right) dt + \sigma(X_0;\theta) dB_t, \quad t \geq 0,\ Y_0 = x. \tag{16}$$

Moreover, we define

$$H_t^Y = \sup_{0 \leq s \leq t} Y_s, \quad \text{and} \quad L_t^Y = \inf_{0 \leq s \leq t} Y_s, \quad t \geq 0. \tag{17}$$

For a sufficiently smooth function $g : \mathbb{R}^4 \to \mathbb{R}$ and for a multi-index $\alpha = (\alpha_1, \ldots, \alpha_4) \in \mathbb{N}_0^4$ we set $|\alpha| = \alpha_1 + \ldots + \alpha_4$ and

$$g_\alpha(x_1, \ldots, x_4) = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}} g(x_1, \ldots, x_4). \tag{18}$$

Let $g \in \mathcal{C}^{0,2,2,4}(\mathbb{R}^4, \mathbb{R})$ such that g all of its partial derivatives satisfy a polynomial growth condition. One of our key tools in the upcoming analysis is the following expansion:

$$\mathbb{E}_{x,\theta}\big[g\big(Y_0, H_\Delta^Y, L_\Delta^Y, Y_\Delta\big)\big] = g(\mathbf{x}) + \sqrt{\Delta}\mathcal{A}_\theta^{(\frac{1}{2})} g(\mathbf{x}) + \Delta \mathcal{A}_\theta^{(1)} g(\mathbf{x}) + O(\Delta^{3/2}), \tag{19}$$

where $\mathbf{x} = (x, x, x, x)$ and the operators $\mathcal{A}_\theta^{(1/2)}$ and $\mathcal{A}_\theta^{(1)}$ are given by

$$\mathcal{A}_\theta^{(\frac{1}{2})} g(\mathbf{x}) = \sigma(x;\theta) \frac{2}{\sqrt{2\pi}} g_{0,1,0,0}(\mathbf{x}) - \sigma(x;\theta) \frac{2}{\sqrt{2\pi}} g_{0,0,1,0}(\mathbf{x}) \tag{20}$$

and

$$
\begin{aligned}
\mathcal{A}_\theta^{(1)} g(\mathbf{x}) &= g_{0,1,0,0}(\mathbf{x})\frac{1}{2}\Big(\mu(x;\theta) - \frac{1}{2}\sigma(x;\theta)\frac{\partial}{\partial x}\sigma(x;\theta)\Big) + g_{0,2,0,0}(\mathbf{x})\frac{1}{2}\sigma^2(x;\theta) \\
&+ \ g_{0,0,1,0}(\mathbf{x})\frac{1}{2}\Big(\mu(x;\theta) - \frac{1}{2}\sigma(x;\theta)\frac{\partial}{\partial x}\sigma(x;\theta)\Big) + g_{0,0,2,0}(\mathbf{x})\frac{1}{2}\sigma^2(x;\theta) \\
&+ \ (1-2\log 2)g_{0,1,1,0}(\mathbf{x})\sigma^2(x;\theta) + \frac{1}{2}g_{0,1,0,1}(\mathbf{x})\sigma^2(x;\theta) + \frac{1}{2}g_{0,0,1,1}(\mathbf{x})\sigma^2(x;\theta) \\
&+ \ g_{0,0,0,1}(\mathbf{x})\left(\mu(x;\theta) - \frac{1}{2}\sigma(x;\theta)\frac{\partial}{\partial x}\sigma(x;\theta)\right) + \frac{1}{2}g_{0,0,0,2}(\mathbf{x})\sigma^2(x;\theta). \qquad (21)
\end{aligned}
$$

This formula is derived in Chapter 5 of [8], some intuition is given in Section 6.1 below. The expansion can only be proved for diffusions with constant diffusion coefficient. This is why Assumption 2.2.1 *(i)* and the Lamperti transform (14), respectively, are required. Likewise, Assumption 2.2.1*(ii)* is necessary since, for (19) to hold, μ^Y must satisfy a global Lipschitz condition with respect to the state variable x.

In its previous from, (19) is a result for functions $g : \mathbb{R}^4 \to \mathbb{R}$ that are independent of the time variable Δ. However, in our analysis we wish to consider flows $\mathcal{G}_\theta = (g_{\Delta,\theta})_{\Delta\geq 0,\theta\in\Theta}$, i.e. families of functions $g_{\Delta,\theta} : \mathbb{R}^4 \to \mathbb{R}$, parameterized by $(\Delta,\theta) \in \mathbb{R}_+ \times \Theta$. Formula (19) indicates that different square root terms in the time variable Δ are involved when one deals with running maxima and minima of diffusions. To come to grips with this additional difficulty, we introduce a particular notation: for $g \in \mathcal{G}_\theta$ and for $n \in \mathbb{N}$, we set

$$
\tilde{g}_{0,\theta}^{(n)}(x,h,l,y) = \lim_{s\to 0}\frac{1}{\sqrt{s^n}}\left(g_{s,\theta}(x,h,l,y) - \sum_{j=0}^{n-1} s^{j/2}\tilde{g}_{0,\theta}^{(j)}(x,h,l,y)\right), \qquad (22)
$$

where $\tilde{g}_{0,\theta}^{(0)} = g_{0,\theta}$ and provided that the limits exists. We will sometimes simply write $\tilde{g}_{0,\theta}$ for $\tilde{g}_{0,\theta}^{(1)}$ and $\tilde{\tilde{g}}_{0,\theta}$ instead of $\tilde{g}_{0,\theta}^{(2)}$. In essence, our strategy will be to apply (19) to the functions $g_{0,\theta}, \tilde{g}_{0,\theta}, \tilde{\tilde{g}}_{0,\theta}$ in order to obtain the necessary expansions of the estimating functions.

We now formulate precise conditions about the classes of estimating functions $\mathcal{G}_\theta$.

Assumption 2.2.2. *The class of flows $\mathcal{G}_\theta$ consists of functions of the form* (6)*, for which $\tilde{g}_{0,\theta}$ and $\tilde{g}_{0,\theta}^{(2)}$ exist and such that, for all $(x,h,l,y) \in \mathbb{R}^4$, the following expansion holds*

$$
\begin{aligned}
&g_{\Delta,\theta}(x,h,l,y) \\
&\quad = g_{0,\theta}(x,h,l,y) + \sqrt{\Delta}\,\tilde{g}_{0,\theta}(x,h,l,y) + \Delta\,\tilde{g}_{0,\theta}^{(2)}(x,h,l,y) + O\big(\Delta^{3/2};\theta,x,h,l,y\big). \qquad (23)
\end{aligned}
$$

The notation $O\big(\Delta^{3/2};\theta,x,h,l,y\big)$ means that the remainder term belongs to $O\big(\Delta^{3/2}\big)$ for fixed $\big(\theta,x,h,l,y\big)$ and has polynomial growth in the variables x, h, l and y. We assume that, for all $\theta \in \Theta$ and $\Delta \geq 0$, the function $g_{\Delta,\theta}(x,h,l,y)$ is continuous in x and 3 times continuously differentiable with respect to each variable h, l and y.

In the sequel, we will – among other things – be interested in estimating functions that consist of linear and quadratic terms. This approach can be considered as a generalization of the method of moments, for we will use the following set of functions:

$$\begin{aligned}
\kappa_h(\Delta, x, h, y; \theta) &= h - \mathbb{E}_{x,\theta}[H_\Delta], \\
\kappa_l(\Delta, x, h, y; \theta) &= l - \mathbb{E}_{x,\theta}[L_\Delta], \\
\kappa_y(\Delta, x, h, y; \theta) &= y - \mathbb{E}_{x,\theta}[Y_\Delta], \\
\kappa_{hh}(\Delta, x, h, y; \theta) &= \big\{h - \mathbb{E}_{x,\theta}[H_\Delta]\big\}^2 - \mathrm{Cov}_{x,\theta}[\, H_\Delta, H_\Delta\,], \\
\kappa_{ll}(\Delta, x, h, y; \theta) &= \big\{l - \mathbb{E}_{x,\theta}[L_\Delta]\big\}^2 - \mathrm{Cov}_{x,\theta}[\, L_\Delta, L_\Delta\,], \\
\kappa_{yy}(\Delta, x, h, y; \theta) &= \big\{y - \mathbb{E}_{x,\theta}[Y_\Delta]\big\}^2 - \mathrm{Cov}_{x,\theta}[\, Y_\Delta, Y_\Delta\,], \\
\kappa_{hl}(\Delta, x, h, y; \theta) &= \big\{h - \mathbb{E}_{x,\theta}[H_\Delta]\big\}\big\{l - \mathbb{E}_{x,\theta}[L_\Delta]\big\} - \mathrm{Cov}_{x,\theta}[\, H_\Delta, L_\Delta\,], \\
\kappa_{hy}(\Delta, x, h, y; \theta) &= \big\{h - \mathbb{E}_{x,\theta}[H_\Delta]\big\}\big\{y - \mathbb{E}_{x,\theta}[Y_\Delta]\big\} - \mathrm{Cov}_{x,\theta}[\, H_\Delta, Y_\Delta\,], \\
\kappa_{ly}(\Delta, x, h, y; \theta) &= \big\{l - \mathbb{E}_{x,\theta}[L_\Delta]\big\}\big\{y - \mathbb{E}_{x,\theta}[Y_\Delta]\big\} - \mathrm{Cov}_{x,\theta}[\, L_\Delta, Y_\Delta\,]. \qquad (24)
\end{aligned}$$

3 Linear estimators for the drift

3.1 Ordinary vs. generalized linear MEFs for the drift

First, by the statement of Proposition 6.2.3,

$$\begin{aligned}
&\mathbb{E}_{\nu,\theta}\left[\frac{\partial}{\partial\theta} g_{\Delta,\theta}\big(Y_0, H_\Delta^Y, L_\Delta^Y, Y_\Delta\big)\right] \\
&= \sqrt{\Delta}\mathbb{E}_{\nu,\theta}[Z_{lin}(Y_0)U_{lin}^{(1/2)}(Y_0)^T] + \Delta\mathbb{E}_{\nu,\theta}[Z_{lin}(Y_0)U_{lin}^{(1)}(Y_0)^T] + o(\Delta) \\
&= \mathbb{E}_{\nu,\theta}[Z_{lin}U_{lin}^T(\Delta, Y_0)] + o(\Delta), \qquad (25)
\end{aligned}$$

where the vector Z_{lin}, $x \in \mathbb{R}$, is defined via

$$Z_{lin}(x) = \left(\frac{\partial}{\partial h}\, g_{0,\theta}(x,h,x,x)\Big|_{h=x}, \frac{\partial}{\partial l}\, g_{0,\theta}(x,x,l,x)\Big|_{l=x}, \frac{\partial}{\partial y}\, g_{0,\theta}(x,x,x,y)\Big|_{y=x}\right). \qquad (26)$$

and where we set

$$U_{lin}(\Delta, x) = \sqrt{\Delta}U_{lin}^{(1/2)}(x) + \Delta U_{lin}^{(1)}(x), \;\; x \in \mathbb{R}, \qquad (27)$$

with the vector

$$U_{lin}^{(1/2)}(x) = \Big(U_{lin,1}^{(1/2)}(x), U_{lin,2}^{(1/2)}(x), U_{lin,3}^{(1/2)}(x)\Big) = \frac{\partial}{\partial\theta}\sigma(x;\theta)\frac{2}{\sqrt{2\pi}}\Big(1, -1, 0\Big) \qquad (28)$$

and the vector $U_{lin}^{(1)}(x)$, whose entries are defined by

$$U_{lin,1}^{(1)}(x) = U_{lin,2}^{(1)}(x) = \frac{1}{2}U_{lin,3}^{(1)}(x) = \frac{1}{2}\left(\frac{\partial}{\partial\theta}\mu(x;\theta) - \frac{1}{2}\frac{\partial}{\partial\theta}\left\{\sigma(x;\theta)\frac{\partial}{\partial x}\sigma(x;\theta)\right\}\right). \qquad (29)$$

Finally, note that $U_{lin}^{(1/2)}$ is zero if the diffusion coefficient of the underlying process X given by (1) does not depend on θ. Secondly, according to Proposition 6.2.5,

$$\mathbb{E}_{\nu,\theta}\Big[g^2_{\Delta,\theta}\big(Y_0, H^Y_\Delta, L^Y_\Delta, Y_\Delta\big)\Big] = \Delta \mathbb{E}_{\nu,\theta}[Z_{lin}(Y_0) S^{-1}_{lin}(Y_0) Z_{lin}(Y_0)^T] + o(\Delta), \tag{30}$$

where we used the additional notation

$$S^{-1}_{lin}(x) = \sigma^2(x;\theta) \begin{pmatrix} 1-\frac{2}{\pi} & \frac{2}{\pi}+(1-2\log 2) & \frac{1}{2} \\ \frac{2}{\pi}+(1-2\log 2) & 1-\frac{2}{\pi} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix}, \; x \in \mathbb{R}. \tag{31}$$

The matrix S^{-1}_{lin} is positive definite for all $x \in \mathbb{R}$. This follows since (up to the factor $\sigma^2(x;\theta)$) S^{-1}_{lin} equals the covariance matrix

$$\operatorname{Cov}\Big[\big(\sup_{0\le t\le 1} B_t, \inf_{0\le t\le 1} B_t, B_1\big)^T\Big]. \tag{32}$$

The formula

$$\mathbb{E}\left[\left(\sup_{0\le t\le 1} B_t\right)\left(\inf_{0\le t\le 1} B_t\right)\right] = 1 - 2\log 2 \tag{33}$$

was proved by Rogers [17]. The remaining moments can be calculated directly since by Lévy's formula the joint density of the running maximum and the terminal value on an interval is explicitly known in the Brownian case, see e.g. (106) below. But the relation between S^{-1}_{lin} and (32) is already clear from the fact that we approximate the moments of the diffusion by its expansion with respect to $\sqrt{\Delta}$. As we will briefly explain in Section 6, usually the first coefficients in the expansion coincide with the respective moments of the triplet $\big(\sup_{0\le t\le 1} B_t, \inf_{0\le t\le 1} B_t, B_1\big)$.

The above expansions enable us to state our first result.

Theorem 3.1.1. *Let X be a diffusion process whose Lamperti transform Y satisfies Assumption 2.2.1 and suppose we are given a class of flows $\mathcal{G}_\theta$ that satisfies Assumption 2.2.2.*
(i) If both coefficients μ and σ of the underlying process X depend on θ, then for all $g \in \mathcal{G}_\theta$,

$$\operatorname{Var}_{\Delta,\nu,\theta}(g) = \mathcal{V}_{0,\theta}(g,\Delta) + o\,(1)\,, \quad \Delta \to 0, \tag{34}$$

where

$$\begin{aligned} \mathcal{V}_{0,\theta}(g,\Delta) &\ge \Delta\Big(\mathbb{E}_{\nu,\theta}\left[U_{lin}(\Delta, Y_0)\, S_{lin}(Y_0)\, U_{lin}(\Delta,Y_0)^T\right]\Big)^{-1} \\ &= \Delta\left(\mathbb{E}_{\nu,\theta}\left[\Delta\left(\Delta U^{(1)}_{lin}(Y_0)^2 + \frac{\pi U^{(1/2)}_{lin}(Y_0)^2}{\pi\log 2 - 2}\right)\right]\right)^{-1}. \end{aligned} \tag{35}$$

In (35) *equality holds, and consequently* g *is small-*Δ*-optimal if there is a scalar* $K \in \mathbb{R}$*, possibly depending on* θ*, such that*

$$Z_{lin} = K\, U_{lin}(\Delta, \cdot) S_{lin}. \tag{36}$$

(ii) If σ *does not depend on* θ*, then for all* $g \in \mathcal{G}_\theta$*,*

$$\mathrm{Var}_{\Delta,\nu,\theta}(g) = \frac{1}{\Delta}\mathcal{V}_{-1,\theta}(g) + o\left(\frac{1}{\Delta}\right), \quad \Delta \to 0, \tag{37}$$

where

$$\mathcal{V}_{-1,\theta}(g) \geq \left(\mathbb{E}_{\nu,\theta}\left[U_{lin}^{(1)}(Y_0)\, S_{lin}(Y_0)\, U_{lin}^{(1)}(Y_0)^T\right]\right)^{-1} = \left(\mathbb{E}_{\nu,\theta}\left[\frac{\frac{\partial}{\partial\theta}\mu(Y_0,\theta)}{\sigma^2(Y_0)}\right]\right)^{-1}. \tag{38}$$

In this case an estimating function $g \in \mathcal{G}_\theta$ *is small-*Δ*-optimal if there is a constant* K *possibly depending on* θ *such that*

$$Z_{lin} = K\, U_{lin}^{(1)} S_{lin} = K\left(0,\, 0,\, \frac{\frac{\partial}{\partial\theta}\mu(\cdot,\theta)}{\sigma^2(\cdot)}\right). \tag{39}$$

Proof. Lemma 2 in [9] states that for integrable random variables UZ^T, $ZS^{-1}Z^T$ with $U \in \mathbb{R}^{1\times b}, Z \in \mathbb{R}^{1\times b}, S \in \mathbb{R}^{b\times b}$, $b \in \mathbb{N}$, and such that $\mathbb{E}(USU^T)$ and $\mathbb{E}(UZ^T)$ are non-zero, the following inequality is satisfied:

$$\mathbb{E}[ZS^{-1}Z^T] \cdot \left(\mathbb{E}[UZ^T]\right)^{-2} \geq (\mathbb{E}[USU^T])^{-1}. \tag{40}$$

Moreover, Jacobsen shows that in (40) there is equality if for some non-random, non-zero value $K \in \mathbb{R}$, $U = KZS^{-1}$, equivalently if $Z = KUS$.

(i) Recall the definition of $\mathrm{Var}_{\Delta,\nu,\theta}(g)$ in (11). Applying inequality (40) to the quotient of the first term in the expansion (30) and the first terms of (25) squared yields

$$\frac{\Delta\mathbb{E}_{\nu,\theta}[Z_{lin}(Y_0)S_{lin}^{-1}(Y_0)Z_{lin}(Y_0)^T]}{\left(\mathbb{E}_{\nu,\theta}[Z_{lin}U_{lin}^T(\Delta, Y_0)]\right)^2} \geq \frac{1}{\mathbb{E}_{\nu,\theta}[U_{lin}(\Delta, Y_0)\, S_{lin}^{-1}(Y_0)\, U_{lin}^T(\Delta, Y_0)]}, \tag{41}$$

which is precisely (35). According to Jacobsen's result, in formula (41) equality holds if there is a constant K such that $Z_{lin} = K\, U_{lin}(\Delta, \cdot)S_{lin}$. Hence, the statement about the small-Δ-optimality becomes obvious.

(ii) If the diffusion coefficient σ does not depend on the parameter θ, the vector $U_{lin}^{(1/2)}$ vanishes identically and $U_{lin}^{(1)}(x)$ becomes

$$U_{lin}^{(1)}(x) = (U_{lin,1}^{(1)}(x), U_{lin,2}^{(1)}(x), U_{lin,3}^{(1)}(x)) = \frac{\partial}{\partial\theta}\mu(x;\theta)\left(\frac{1}{2}, \frac{1}{2}, 1\right), \quad x \in \mathbb{R}. \tag{42}$$

We consequently find the following lower bound for the first term $\mathcal{V}_{-1,\theta}(g)$ of the variance:

$$\mathcal{V}_{-1,\theta}(g) = \frac{\mathbb{E}_{\nu,\theta}[Z_{lin}(Y_0)\, S_{lin}^{-1}(Y_0)\, Z_{lin}(Y_0)^T]}{\Big(\mathbb{E}_{\nu,\theta}[Z_{lin}(Y_0) U_{lin}^{(1)}(Y_0)^T]\Big)^2} \geq \Big(\mathbb{E}_{\nu,\theta}\Big[U_{lin}^{(1)}(Y_0)\, S_{lin}(Y_0)\, U_{lin}^{(1)}(Y_0)^T\Big]\Big)^{-1}$$

$$= \left(\mathbb{E}_{\nu,\theta}\left[\frac{\frac{\partial}{\partial\theta}\mu(Y_0,\theta)}{\sigma^2(Y_0)}\right]\right)^{-1}. \tag{43}$$

Equality holds if there is a constant K such that $Z_{lin} = K\, U_{lin}^{(1)} S_{lin}$. Thus, the statement about the small-Δ-optimality is obvious. □

Let us briefly outline the situation for the linear estimating function, i.e. for the function g_{lin} which is defined by

$$g_{lin}(\Delta, x, h, l, y; \theta) = \sum_{j \in M_{lin}} a_j^{lin}(\Delta, x; \theta)\kappa_j(\Delta, x, h, l, y; \theta), \tag{44}$$

where the index set is given by $M_{lin} = \{h, l, y\}$ and the respective functions $\kappa_h(\cdot)$, $\kappa_l(\cdot)$, $\kappa_y(\cdot)$ are described in (24). If both μ and σ depend on θ the lower bound (35) is not very handy. Thus, the interesting case is the case where σ is independent of θ. By the above statements, in this situation, (44) is small-Δ-optimal if the weights $(a_h^{lin}, a_l^{lin}, a_y^{lin})$ satisfy

$$a_h^{lin}(\Delta, x; \theta) = 0, \quad a_l^{lin}(\Delta, x; \theta) = 0, \quad a_y^{lin}(\Delta, x; \theta) = \frac{\partial}{\partial\theta}\mu(x,\theta)\Big/\sigma(x)^2. \tag{45}$$

3.2 Assessment of the results for linear MEFs

Let us focus on the case where σ does not depend on θ. If (44) in turn does not depend on h and l, that means if we neglect the observations H_Δ and L_Δ in our analyis, the small-Δ-asymptotic lower bound of the variance $\mathrm{Var}_{\Delta,\nu,\theta}(g)$ is given by

$$\frac{1}{\Delta} \cdot \left(\mathbb{E}_{\nu,\theta}\left[\frac{\frac{\partial}{\partial\theta}\mu(\cdot,\theta)}{\sigma^2(\cdot)}\right]\right)^{-1}. \tag{46}$$

Jacobsen already encountered this result when dealing with ordinary MEFs. Especially, compare Theorem 1 (i) in [9]. Formula (46) implies that the variance of an estimating function, inferred from the equidistant discrete sample $X_0, \ldots, X_T$, with n fixed and $T = n\Delta$, explodes as $\Delta \to 0$. Also, incorporating the maximum and the minimum over the observation intervals does not lower the variance – formula (38) shows that the first term in the expansion of $\mathrm{Var}_{\Delta,\nu,\theta}(g)$ remains the same.

An explanation for this phenomenon can be obtained by the following deliberations: if $t > 0$ is small and if X starts in x we have $\int_0^t \mu(X_s)ds + \sigma B_t \approx \mu(x)t + \sigma B_t$. Therefore, on small intervals, Itô-diffusions are very well approximated by a Brownian motion with drift. Without loss of generality let us consider the example ($X_t = \mu t + B_t$, $0 \le t \le 1$) with $\mu \in \mathbb{R}$. From Girsanov's Theorem we infer that the Log-likelihood is

$$\log L(\mu) = \mu X_1 - \frac{1}{2}\mu^2. \tag{47}$$

Thus a sufficient statistic for the parameter μ is already given by X_1. In a nutshell, in a Brownian model, all the information about the drift is contained in one single point, namely the endpoint X_1 of the trajectory (X_t, $0 \le t \le 1$).

4 Estimators for the diffusion coefficient

4.1 A special quadratic estimator for the diffusion coefficient

First we want to consider the case where the estimating function g is independent of the minimum variable l. In Proposition 6.2.5 we will prove the following approximation:

$$\mathbb{E}_{\nu,\theta}\Big[g^2_{\Delta,\theta}\big(Y_0, H^Y_\Delta, Y_\Delta\big)\Big] = \Delta \mathbb{E}_{\nu,\theta}\Big[\mathcal{A}^S g_{0,\theta}(Y_0)\Big] + O\big(\Delta^{3/2}\big), \tag{48}$$

where the operator $\mathcal{A}^S$ is given by (131) below. If

$$\frac{\partial}{\partial h} g_{0,\theta}(x,h,x)\Big|_{h=x} = 0 \quad \text{and} \quad \frac{\partial}{\partial y} g_{0,\theta}(x,x,y)\Big|_{y=x} = 0, \tag{49}$$

a further expansion of $\mathbb{E}_{\nu,\theta}\Big[g^2_{\Delta,\theta}\big(Y_0, H^Y_\Delta, Y_\Delta\big)\Big]$ is required, since in this case the expression $\mathbb{E}_{\nu,\theta}\Big[\mathcal{A}^S g_{0,\theta}(Y_0)\Big]$ on the right hand side of (48) identically equals 0. This becomes evident by the fact that, basically, (49) means that we deal with estimating functions consisting of terms that are at least quadratic. Consequently, the respective expansion of (48) w.r.t. $\sqrt{\Delta}$ starts with a term proportional to Δ^2 in this particular situation. This, in turn, follows from the estimate

$$\begin{aligned} &\mathbb{E}_{x,\theta}\left[\left(\sup_{0\le t\le\Delta} Y_t - x\right)^k (Y_\Delta - x)^l\right] \\ &= \sigma(x;\theta)^{k+l}\mathbb{E}\left[\left(\sup_{0\le t\le\Delta} B_t\right)^k \cdot B^l_\Delta\right] + O(\Delta^{(k+l+1)/2}) \\ &= \sigma(x;\theta)^{k+l} \cdot O(\Delta^{(k+l)/2}) + O(\Delta^{(k+l+1)/2}), \end{aligned} \tag{50}$$

where $k, l \in \mathbb{N}$. Formula (50) is a consequence of Doob's inequality and it also allows to calculate the coefficients belonging to Δ^2. According to Proposition 6.2.6 we have

$$\mathbb{E}_{\nu,\theta}\Big[g^2_{\Delta,\theta}\big(Y_0, H^Y_\Delta, Y_\Delta\big)\Big] = \Delta^2 \mathbb{E}_{\nu,\theta}[Z_{qua}(Y_0) S^{-1}_{qua}(Y_0) Z_{qua}(Y_0)^T] + O(\Delta^{5/2}), \tag{51}$$

where for $x \in \mathbb{R}$ we set

$$\begin{aligned} Z_{qua}(x) &= (Z_{qua,1}(x), Z_{qua,2}(x), Z_{qua,3}(x)) \\ &= \left(\frac{\partial^2}{\partial h^2} g_{0,\theta}(x,h,x)\Big|_{h=x}, \frac{\partial^2}{\partial y^2} g_{0,\theta}(x,x,y)\Big|_{y=x}, \frac{\partial}{\partial h}\frac{\partial}{\partial y} g_{0,\theta}(x,h,y)\Big|_{y=x} \right), \end{aligned} \tag{52}$$

and

$$S_{qua}^{-1}(x) = \sigma^4(x;\theta) \begin{pmatrix} \frac{1}{2} - \frac{4}{\pi^2} & \frac{1}{2}\left(\frac{1}{2} - \frac{2}{3\pi}\right) & \frac{1}{2}\left(\frac{7}{4} - \frac{4}{\pi}\right) \\ \frac{1}{2}\left(\frac{1}{2} - \frac{2}{3\pi}\right) & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2}\left(\frac{7}{4} - \frac{4}{\pi}\right) & \frac{1}{2} & \frac{7}{4} - \frac{10}{3\pi} \end{pmatrix}. \tag{53}$$

Note that analogously to formula (32) it can be shown that

$$S_{qua}^{-1}(x) = \sigma^4(x;\theta)\,\mathrm{Cov}\left[\left(\left(\sup_{0\le t\le 1} B_t - \frac{2}{\pi} \right)^2, \left(\inf_{0\le t\le 1} B_t + \frac{2}{\pi} \right)^2, B_1^2 \right)^T \right], \tag{54}$$

and, hence, S_{qua}^{-1} is a covariance matrix. For more details we make reference to Section 6, but we emphasize that the proof of Proposition 6.2.6 does not work any more for martingale estimating functions that have both linear and quadratic terms, since in this case condition (49) is violated. To capture the particularities of such a model, a sophisticated expansion of $\mathbb{E}_{x,\theta}\left[g_{\Delta,\theta}^2(Y_0, H_\Delta^Y, L_\Delta^Y, Y_\Delta) \right]$ with respect to $\sqrt{\Delta}$ is necessary. To exemplify, let us consider the expression $\gamma(\Delta) = H_\Delta^Y + Y_\Delta^2$. Obviously, $\mathbb{E}_{0,\theta}[\gamma(\Delta)^2] = \mathbb{E}_{0,\theta}[(H_\Delta^Y)^2] + 2\mathbb{E}_{0,\theta}[Y_\Delta^2 H_\Delta^Y] + \ldots = O(\Delta) + O(\Delta^{3/2}) + \ldots$. According to (50), the first term in the expansion of $\mathbb{E}_{0,\theta}[Y_\Delta^2 H_\Delta^Y]$ is of the order $\Delta^{3/2}$. Also, we are able to determine the coefficient that belongs to Δ in the expansion of $\mathbb{E}_{0,\theta}[(H_\Delta^Y)^2]$, but not the coefficient that belongs to $\Delta^{3/2}$.

The result concerning the expansion of the derivative with respect to θ remains unaffected by (49). Let us define the vector

$$\begin{aligned} U_{qua}(x) &= \big(U_{qua,1}(x), U_{qua,2}(x), U_{qua,3}(x)\big) \\ &= \frac{1}{2}\frac{\partial}{\partial\theta}\sigma(x;\theta)^2 \left(\left(\frac{2}{\pi} - 1 \right), -1, -1 \right), \ x \in \mathbb{R}, \end{aligned} \tag{55}$$

then we are able to write

$$\mathbb{E}_{\nu,\theta}\left[\frac{\partial}{\partial\theta} g_{\Delta,\theta}\big(Y_0, H_\Delta^Y, Y_\Delta\big) \right] = \Delta \mathbb{E}_{\nu,\theta}[Z_{qua}(Y_0) U_{qua}(Y_0)^T] + O(\Delta^{3/2}), \tag{56}$$

compare Proposition 6.2.6 in Section 6. Overall, we find the following expansion of the variance of an estimator derived from a strictly quadratic generalized martingale estimating function g.

Theorem 4.1.1. *Let X be a diffusion process whose Lamperti transform Y satisfies Assumption 2.2.1 and suppose we are given a class of flows $\mathcal{G}_\theta$ that satisfies Assumption 2.2.2. We assume that, for each $g \in \mathcal{G}_\theta$, the function*

$$(x, h, l, y) \longmapsto g_{\Delta,\theta}(x, h, l, y) \tag{57}$$

is independent of l, for all $\theta \in \Theta$ and for all $\Delta \geq 0$. Moreover, we assume that any $g \in \mathcal{G}_\theta$ satisfies (49)*. Let us assume that the coefficient μ of the underlying process X does not depend on θ. Then, for all $g \in \mathcal{G}_\theta$,*

$$\mathrm{Var}_{\Delta,\nu,\theta}(g) = \mathcal{V}_{0,\theta}(g) + o(1), \quad \Delta \to 0, \tag{58}$$

where the first term on the right hand side of (58) *is lower bounded by*

$$\begin{aligned}\mathcal{V}_{0,\theta}(g) &\geq \left(\mathbb{E}_{\nu,\theta}\left[U_{qua}(Y_0)\, S_{qua}(Y_0)\, U_{qua}(Y_0)^T\right]\right)^{-1} \\ &= \frac{r(\theta)^2}{r'(\theta)^2}\left(6 + \frac{96(28-9\pi)}{3\pi(56+3\pi)-608}\right)^{-1} \approx \frac{r(\theta)^2}{r'(\theta)^2}\cdot 0.33983,\end{aligned} \tag{59}$$

provided that $r'(\theta)$ does not vanish. Moreover, small-Δ-optimality holds for $g \in \mathcal{G}_\theta$ if there is a scalar $K \in \mathbb{R}$, possibly depending on θ, such that

$$Z_{qua} = K U_{qua} S_{qua}. \tag{60}$$

Proof. By definition of $\mathrm{Var}_{\Delta,\nu,\theta}(g)$ in (11) one obtains the result by dividing (51) by formula (56) squared, in combination with the estimates in Lemma 2 of [10]. □

Now, let g denote the particular quadratic martingale estimating function

$$g_{qua}(\Delta, x, h, l, y; \theta) = \sum_{j \in M_{qua}} a_j^{qua}(\Delta, x; \theta)\kappa_j(\Delta, x, h, l, y; \theta) \tag{61}$$

with the index set $M_{qua} = \{hh, hy, yy\}$. According to our analysis, the small-Δ-optimal weights for the function (61) can be chosen as

$$\begin{aligned}a_{hh}^{qua}(\Delta, x; \theta) = -\, a_{hy}^{qua}(\Delta, x; \theta) &= 6.2355 \big/ \sigma(x)^2, \\ a_{yy}^{qua}(\Delta, x; \theta) &= 2.3232 \big/ \sigma(x)^2.\end{aligned} \tag{62}$$

4.2 A more general quadratic estimator for the diffusion coefficient

Let us assume that g depends on all variables and that it satisfies

$$\frac{\partial}{\partial h} g_{0,\theta}(x, h, x, x)\Big|_{h=x} = \frac{\partial}{\partial l} g_{0,\theta}(x, x, l, x)\Big|_{l=x} = \frac{\partial}{\partial y} g_{0,\theta}(x, x, x, y)\Big|_{y=x} = 0. \tag{63}$$

We consider the vector $Z_{qua}^a(x)$ defined in formulae (157). In this case we are able to write

$$\mathbb{E}_{x,\theta}\Big[g_{\Delta,\theta}^2\big(Y_0, H_\Delta^Y, L_\Delta^Y, Y_\Delta\big)\Big] = \Delta^2\, Z_{qua}^a(x)\cdot\big(S_{qua}^a\big)^{-1}(x)\cdot\big(Z_{qua}^a(x)\big)^T + O\big(\Delta^{5/2}\big), \tag{64}$$

where a simulation of $\mathrm{Cov}[S^B_{H,L,X}]$, with $S^B_{H,L,X}$ defined by (159), shows that the matrix $\left(S^a_{qua}\right)^{-1}$ is approximately given by

$$\left(S^a_{qua}\right)^{-1}(x) = \sigma^4(x;\theta)\begin{pmatrix} 0.095 & 0.144 & 0.259 & 0.057 & 0.01 & 0.059 \\ 0.144 & 0.499 & 0.5 & 0.143 & 0.144 & 0.499 \\ 0.259 & 0.5 & 0.8 & 0.174 & 0.059 & 0.261 \\ 0.057 & 0.143 & 0.174 & 0.082 & 0.057 & 0.174 \\ 0.01 & 0.144 & 0.059 & 0.057 & 0.094 & 0.259 \\ 0.059 & 0.499 & 0.261 & 0.174 & 0.259 & 0.799 \end{pmatrix}, \quad x \in \mathbb{R}, \quad (65)$$

compare the statement of Proposition 6.2.7. The simulation can be avoided for there seems to be a closed form expansion of the moment generating function, see Garman and Klass [5]. Consequently, not only (33) but also the other relevant moments of the pair $(\sup_{0\leq t\leq 1} B_s, \inf_{0\leq t\leq 1} B_s)$ can be derived. However, we decided to simulate because of the missing proof and because of the alleged generating function's complexity. Our simulations included $5 \cdot 10^5$ independent trajectories of the standard Brownian motion on the interval $[0, 1]$ and each trajectory was computed with an accuracy of 10^6 steps.

Moreover, we define the vector $U^a_{qua}(x)$, $x \in \mathbb{R}$, by the entries

$$\begin{aligned} U^a_{qua,1}(x) = U^a_{qua,5}(x) &= \frac{1}{2}\left(\frac{2}{\pi} - 1\right)\frac{\partial}{\partial\theta}\sigma(x;\theta)^2, \\ U^a_{qua,2}(x) = U^a_{qua,3}(x) = U^a_{qua,6}(x) &= -\frac{1}{2}\frac{\partial}{\partial\theta}\sigma(x;\theta)^2, \\ U^a_{qua,4}(x) &= \left(-\frac{2}{\pi} - 1 + 2\log 2\right)\frac{\partial}{\partial\theta}\sigma(x;\theta)^2. \end{aligned} \quad (66)$$

According to Proposition 6.2.3 we are able to write

$$\mathbb{E}_{\nu,\theta}\left[\frac{\partial}{\partial\theta}g_{\Delta,\theta}\big(Y_0, H^Y_\Delta, L^Y_\Delta, Y_\Delta\big)\right] = \Delta\mathbb{E}_{\nu,\theta}[Z^a_{qua}(Y_0)U^a_{qua}(Y_0)^T] + O(\Delta^{3/2}). \quad (67)$$

With these expressions at hand we are able to determine the small-Δ-optimal lower bound and the small-Δ-optimal weights for the most general strictly quadratic estimating functions.

Theorem 4.2.1. *Suppose we are given a class of flows $\mathcal{G}_\theta$ that satisfies Assumption 2.2.2. We assume that any $g \in \mathcal{G}_\theta$ satisfies for all $x \in \mathbb{R}$ and for all $\theta \in \Theta$ the additional condition* (63)*. Moreover, let us assume that the coefficient μ of the underlying process X does not depend on θ. Then, for all $g \in \mathcal{G}_\theta$,*

$$\mathrm{Var}_{\Delta,\nu,\theta}(g) = \mathcal{V}^a_{0,\theta}(g) + o(1), \quad \Delta \to 0, \quad (68)$$

where the first term on the right hand side of (68) *is lower bounded by*

$$\begin{aligned} \mathcal{V}^a_{0,\theta}(g, \hat{\theta}) &\geq \left(\mathbb{E}_{\nu,\theta}\left[U^a_{qua}(Y_0)\, S^a_{qua}(Y_0)\, U^a_{qua}(Y_0)^T\right]\right)^{-1} \\ &\approx 0.1228 \cdot \frac{r(\theta)^2}{r'(\theta)^2}, \end{aligned} \quad (69)$$

provided that $r'(\theta)$ does not vanish. Moreover, small-Δ-optimality holds for $g \in \mathcal{G}_\theta$ if there is a scalar $K \in \mathbb{R}$, possibly depending on θ, such that

$$Z^a_{qua} = K U^a_{qua} S^a_{qua}. \tag{70}$$

Proof. Again, by definition of $\mathrm{Var}_{\Delta,\nu,\theta}(g)$ in (11), the expansion in formulae (64) and (67) which are proved in Proposition 6.2.3 and Proposition 6.2.7, respectively, in combination with the simulated matrix (65) and the inequality of Lemma 2 in [9] gives the assertion. □

We consider a concrete example. Let

$$g^a_{qua}(\Delta, x, h, y; \theta) = \sum_{j \in M^a_{qua}} a^{qua,a}_j(\Delta, x; \theta) \kappa_j(\Delta, x, h, y; \theta), \tag{71}$$

with the index-set $M^a_{qua} = \{hh, ll, yy, hy, ly, hl\}$. A set of small-$\Delta$-optimal weights is given by

$$\begin{aligned} a^{qua,a}_{hh}(\Delta, x; \theta) = a^{qua,a}_{ll}(\Delta, x; \theta) &= 5.436 \Big/ \sigma(x)^2, \\ a^{qua,a}_{yy}(\Delta, x; \theta) &= 11.056 \Big/ \sigma(x)^2, \\ a^{qua,a}_{hy}(\Delta, x; \theta) = a^{qua,a}_{ly}(\Delta, x; \theta) &= -16.304 \Big/ \sigma(x)^2, \\ a^{qua,a}_{hl}(\Delta, x; \theta) &= 21.756 \Big/ \sigma(x)^2. \end{aligned} \tag{72}$$

4.3 Assessment of the results for quadratic martingale estimating functions

One important fact to state is that the lower bounds we found in (59) and (69) are independent of the initial distribution. This is due to the fact that we imposed a special structure for $\sigma(\cdot;\theta) = r(\theta) \cdot \sigma(\cdot)$, see Assumption 2.2.1. The respective moments of the function $\sigma(\cdot)$ can be canceled out and we are left with a term that depends on $r(\theta)$ alone.

Furthermore, we stress that, if the martingale estimating function g does not depend on h and l, the lower bound for the first term $\mathcal{V}_{0,\theta}(g)$ in the expansion of the variance equals $0.5 \cdot r(\theta)^2 / r'(\theta)^2$. This result can easily be obtained as a special case of our analysis. Jacobsen already stated this fact in Theorem 1 (ii) of his paper [9]. A comparison with the formulae (59) and (69) shows that, in contrast to the case of linear MEFs, we benefit from using generalized martingale estimating functions.

4.4 Range based estimators for the diffusion coefficient

In this section we finally want to consider martingale estimating functions g that are independent of the variable y that corresponds to the end point Y_Δ. By the Propositions 6.2.3 and 6.2.5 we have the following expansions:

$$\mathbb{E}_{\nu,\theta}\left[\frac{\partial}{\partial\theta} g_{\Delta,\theta}(Y_0, H^Y_\Delta, L^Y_\Delta)\right] = \sqrt{\Delta}\,\mathbb{E}_{\nu,\theta}[Z_{range}(Y_0) U_{range}(Y_0)^T] + o(\sqrt{\Delta}) \tag{73}$$

and

$$\mathbb{E}_{\nu,\theta}\Big[g^2_{\Delta,\theta}\big(Y_0, H^Y_\Delta, L^Y_\Delta\big)\Big] = \Delta\mathbb{E}_{\nu,\theta}[Z_{range}(Y_0)S^{-1}_{range}(Y_0)Z_{range}(Y_0)^T] + o(\Delta), \qquad (74)$$

where

$$Z_{range}(x) = \left(\frac{\partial}{\partial h}g_{0,\theta}(x,h,x)\Big|_{h=x}, \frac{\partial}{\partial l}g_{0,\theta}(x,x,l)\Big|_{l=x}\right), \ x \in \mathbb{R}, \qquad (75)$$

and U_{range} is the vector

$$U_{range}(x) = \big(U^{(1/2)}_{lin,1}(x), U^{(1/2)}_{lin,2}(x)\big) = \left(\frac{\partial}{\partial\theta}\sigma(x;\theta)\frac{2}{\sqrt{2\pi}}, -\frac{\partial}{\partial\theta}\sigma(x;\theta)\frac{2}{\sqrt{2\pi}}\right), \ x \in \mathbb{R}. \ (76)$$

The matrix S^{-1}_{range} is defined by

$$S^{-1}_{range}(x) = \sigma^2(x;\theta)\begin{pmatrix} 1-\frac{2}{\pi} & \frac{2}{\pi}+(1-2\log 2) \\ \frac{2}{\pi}+(1-2\log 2) & 1-\frac{2}{\pi}\end{pmatrix}, \ x \in \mathbb{R}. \qquad (77)$$

Again, we have $S^{-1}_{range}(x) = \sigma^2(x;\theta)\,\mathrm{Cov}\Big[\big(\sup_{0\le t\le 1} B_s, \inf_{0\le t\le 1} B_s\big)^T\Big]$. We find an overall expansion of the variance of a range based estimating function g_{range} in the following theorem.

Theorem 4.4.1. *Suppose we are given a class of flows $\mathcal{G}_\theta$ that satisfies Assumption 2.2.2. We assume that $\mathcal{G}_\theta$ is such that, for each $g \in \mathcal{G}_\theta$, the function*

$$(x,h,l,y) \longmapsto g_{\Delta,\theta}(x,h,l,y) \qquad (78)$$

is independent of y, for all $\theta \in \Theta$ and for all $\Delta \ge 0$. If the coefficient σ of the underlying process X depends on θ, then, for all $g \in \mathcal{G}_\theta$,

$$\mathrm{Var}_{\Delta,\theta}(g) = \mathcal{V}_{0,\theta}(g) + o(1), \quad \Delta \to 0. \qquad (79)$$

The first term in the expansion is lower bounded by

$$\begin{aligned}\mathcal{V}_{0,\theta}(g) &\ge \big(\mathbb{E}_{\nu,\theta}\left[U_{range}(Y_0)\,S_{range}(Y_0)\,U_{range}(Y_0)^T\right]\big)^{-1} \\ &= \frac{r(\theta)^2}{r'(\theta)^2}\cdot\left(\frac{2}{\pi\log 2-2}\right)^{-1} \approx \frac{r(\theta)^2}{r'(\theta)^2}\cdot 0.088793, \end{aligned} \qquad (80)$$

provided that $r'(\theta)$ does not vanish. And finally, equality holds in (80) and $g \in \mathcal{G}_\theta$ is small-Δ-optimal if there is a scalar $K \in \mathbb{R}$, possibly depending on θ, such that

$$Z_{range} = KU_{range}S_{range}. \qquad (81)$$

Proof. Recall the definition of $\mathrm{Var}_{\Delta,\nu,\theta}[g,\hat\theta]$ in (11). If one divides (74) by (73) squared, the result follows by means of Lemma 2 in [9]. □

In order to prove Theorem 4.4.1 it is not necessary to work with the Lamperti transform of X and thus Assumption 2.2.1 is redundant. A first order expansion (with respect to $\sqrt{\Delta}$) of the expression

$$\mathbb{E}_{x,\theta}\left[\frac{\partial}{\partial\theta} g_{\Delta,\theta}\big(X_0, H_\Delta, L_\Delta\big)\right] \tag{82}$$

and a second order expansion of

$$\mathbb{E}_{x,\theta}\Big[g^2_{\Delta,\theta}\big(X_0, H_\Delta, L_\Delta\big)\Big], \tag{83}$$

are sufficient to determine the asymptotic lower bounds for the variance of strictly range based martingale estimating functions. Ceteris paribus, such expansions can be obtained for any diffusion process of the type (1). In the sequel, we will give a short heuristic why this is the case. For a sufficiently smooth function $\gamma : \mathbb{R} \to \mathbb{R}$ we have

$$\begin{aligned}\mathbb{E}_{x,\theta}\left[\gamma\left(\sup_{0\le s\le t} X_s\right)\right] &= \gamma(x) + \gamma'(x)\mathbb{E}_{x,\theta}\left[\left(\sup_{0\le s\le t} X_s - x\right)\right] \\ &\quad + \frac{1}{2}\gamma''(x)\mathbb{E}_{x,\theta}\left[\left(\sup_{0\le s\le t} X_s - x\right)^2\right] \\ &\quad + \frac{1}{6}\mathbb{E}_{x,\theta}\left[\gamma'''(\xi)\left(\sup_{0\le s\le t} X_s - x\right)^3\right],\end{aligned} \tag{84}$$

where ξ is between x and $\sup_{0\le s\le t} X_s$. If $\gamma'''(y)$ has polynomial growth, by Doob's inequality, it can be derived that

$$\mathbb{E}_{x,\theta}\left[\left(\sup_{0\le s\le t} X_s - x\right)^2\right] = O(\Delta), \text{ and } \mathbb{E}_{x,\theta}\left[\gamma(\xi)\left(\sup_{0\le s\le t} X_s - x\right)^3\right] = O(\Delta^{3/2}). \tag{85}$$

And, if one replaces Y with X, formula (50) remains valid. It might be a bit surprising that the estimates are identical for the process (1) and for the model based on the Lamperti transform (14), but this fact becomes obvious when taking into account that (14) is rescaled by $\sigma(X_0;\theta)$. By means of (50), it is always possible to derive a first order expansion of (82) and one can show that it has the same form as (73). Also compare Remark 6.2.4 below. On the other hand, the martingale condition (7) implies $g_{0,\theta}(x,x,x) \equiv 0$. It follows directly that the term associated with the first derivative in the expansion of (83) vanishes – recall that $(g^2)' = 2g\,g'$. It remains a second order expansion w.r.t. $\sqrt{\Delta}$, which formally coincides with the expansion (74). Again, this is due to the estimate (50).

To sum up, we are able to work with any diffusion model (1) that satisfies the remaining assumptions of Section 2 and we find the lower bound

$$\begin{aligned}\mathcal{V}_{0,\theta}(g) &\ge \big(\mathbb{E}_{\nu,\theta}\left[U_{range}(Y_0)\, S_{range}(Y_0)\, U_{range}(Y_0)^T\right]\big)^{-1} \\ &= \left(\frac{2}{\pi\log 2 - 2}\right)^{-1}\left(\mathbb{E}_{\nu,\theta}\left[\frac{\big(\frac{\partial}{\partial\theta}\sigma(Y_0;\theta)\big)^2}{\sigma(Y_0;\theta)^2}\right]\right)^{-1}.\end{aligned} \tag{86}$$

To underpin its importance, let us state this result in a corollary.

Corollary 4.4.2. *We assume that we are given a stochastic differential equation of the type* (1)*, whose coefficients* $\mu(\cdot;\theta)$ *and* $\sigma(\cdot;\theta)$ *satisfy 2.1.1, but not necessarily Assumption 2.2.1. Let* $\mathcal{G}_\theta$ *be a class of flows that satisfies Assumption 2.2.2. We assume that* $\mathcal{G}_\theta$ *is such that, for each* $g \in \mathcal{G}_\theta$*, the function*

$$(x, h, l, y) \longmapsto g_{\Delta,\theta}(x, h, l, y) \tag{87}$$

is independent of y *for all* $\theta \in \Theta$ *and for all* $\Delta \geq 0$*. Set* $v(\theta) = \mathbb{E}_{\nu,\theta}[g^2_{\Delta,\theta}(X_0, H_\Delta, L_\Delta)]$ *and* $\xi(\theta) = \mathbb{E}_{\nu,\theta}[\partial_\theta g_{\Delta,\theta}(X_0, H_\Delta, L_\Delta)]$*. If the coefficient* σ *of the process* X *depends on* θ*, then the variance* $\mathrm{Var}_{\Delta,\nu,\theta}(g) = v(\theta)/\xi(\theta)^2$ *satisfies, for all* $g \in \mathcal{G}_\theta$*, the expansion*

$$\mathrm{Var}_{\Delta,\theta}(g) = \mathcal{V}_{0,\theta}(g) + o(1), \quad \Delta \to 0, \tag{88}$$

where the first term $\mathcal{V}_{0,\theta}(g)$ *satisfies the inequality* (86) *provided that* $r'(\theta)$ *does not vanish. Finally, equality holds in* (86) *and* $g \in \mathcal{G}_\theta$ *is small-*Δ*-optimal if there is a scalar* $K \in \mathbb{R}$*, possibly depending on* θ*, such that*

$$Z_{range} = K U_{range} S_{range}. \tag{89}$$

Proof. According to our above discussions, the proof follows from formulae (73) and (74) in combination with Lemma 2 in [9]. □

We end our analysis of the range based case by examining the concrete estimating function

$$g(\Delta, x, h, l, y; \theta) = \sum_{j \in M_{range}} a_j^{range}(\Delta, x; \theta) \kappa_j(\Delta, x, h, l, y; \theta), \tag{90}$$

with the index-set $M_{range} = \{h, l\}$. Small-Δ-optimal weights for (90) are clearly given by

$$a_h^{range}(\Delta, x; \theta) = \frac{\partial}{\partial\theta}\sigma(x;\theta) \quad \text{and} \quad a_l^{range}(\Delta, x; \theta) = -\frac{\partial}{\partial\theta}\sigma(x;\theta). \tag{91}$$

4.5 Assessment of the results for range based MEFs

The factor $\left(\frac{2}{\pi \log 2 - 2}\right)^{-1} \approx 0.088793$ appearing in the asymptotic lower bounds of the variance for the class of range based estimating functions we displayed in formula (80) and formula (86) coincides with the variance of the expression

$$\sqrt{\frac{\pi}{2}} \frac{(H_\Delta^B - L_\Delta^B)}{\Delta}, \tag{92}$$

which is an unbiased estimator for the diffusion coefficient in the Brownian model without drift. As above H_Δ^B and L_Δ^B stand for $H_\Delta^B = \sup_{0\le t\le\Delta} B_t$ and $L_\Delta^B = \inf_{0\le t\le\Delta} B_t$, respectively. Consequently, we infer that

$$\mathrm{Var}\left[\sqrt{\frac{2}{\pi}}\frac{(H_\Delta^B - L_\Delta^B)}{\Delta}\right] = \frac{\pi}{2}\,\frac{\mathbb{E}[(H_\Delta^B)^2] + \mathbb{E}[(L_\Delta^B)^2] - 2\mathbb{E}\left[H_\Delta^B \cdot L_\Delta^B\right]}{\Delta} - 1 = \frac{\pi \log 2}{2} - 1. \tag{93}$$

A comparison of formula (80), which states that the first term in the expansion of the variance is lower bounded by $0.088793 \cdot r(\theta)^2/r'(\theta)^2$, with the lower bound for the ordinary martingale estimating function, which is given by $0.5 \cdot r(\theta)^2/r'(\theta)^2$, shows that there is a gain in efficiency of about 82 %. The asymptotic lower bound for the variance in the strictly range-based model is even lower than the one for the variance of the generalized quadratic martingale estimating function, which is given by $0.1228 \cdot r(\theta)^2/r'(\theta)^2$, see formula (69) and also the discussion in Paragraph 4.3. This result is particularly interesting because it shows that in our model the linear estimator constructed from the ranges $\left(\bar{H}_{(i\Delta)} - \bar{L}_{(i\Delta)}\right)_{i=1,\ldots,n}$ is more efficient than the quadratic estimator inferred from the samples $\left((\bar{H}_{(i\Delta)} - \bar{L}_{(i\Delta)})^2\right)_{i=1,\ldots,n}$.

Small-Δ-optimality of strictly range-based martingale estimating functions cannot only be stated in a model where the underlying process has the structure of a properly rescaled Lamperti transform, but for fairly general diffusions as well – see Corollary 4.4.2. A comparison of this refined result with the corresponding result for ordinary martingale estimating functions is not possible in general. Both lower bounds depend on $\sigma(\cdot;\theta)$ in a different way. Compare formula (86) and the respective formula in Theorem 1 (ii) in [9].

5 Case study

We consider an Ornstein-Uhlenbeck process defined by the following stochastic differential equation

$$dX_t = -\mu X_t dt + \sigma dB_t, \quad X_0 = x,\ t \ge 0. \tag{94}$$

Simulations were performed for for $\mu \equiv 1$ and $\sigma \equiv 1$ with the values $\Delta = 0.5,\ 0.1,\ 0.01$ based on $n + 1 = 501$ observations. For each value of Δ, 1000 trajectories were created according to an Euler-scheme such that there are at least 1000 simulated values of X in each observation interval $((i-1)\Delta, i\Delta]$, for $i = 1, \ldots, 500$ and for every value of $\Delta = 0.5,\ 0.1,\ 0.01$.

We will only consider estimators of the diffusion parameter $\sigma = 1$, where we assume that the parameter $\mu = 1$ is known. We do not conduct a simulation study for the drift parameter μ since the generalized martingale estimating functions coincides with the ordinary martingale estimating function in this case, see Section 3. A simulation study for

Estimation method	Δ	mean	std. dev.	max	min
Ordinary MEF: $\hat{\sigma}^2_{ord}$	0.5	0.392	0.027	0.478	0.319
	0.1	0.854	0.054	1.032	0.700
	0.01	0.983	0.062	1.203	0.807
Generalized MEF: $\hat{\sigma}^2_{qua}$	0.5	0.68	0.03	0.781	0.594
	0.1	0.917	0.047	1.072	0.798
	0.01	0.992	0.054	1.19	0.853
Gen. MEF – all quad. terms: $(\hat{\sigma}^a_{qua})^2$	0.5	0.555	0.063	1.00	0.372
	0.1	0.908	0.038	1.028	0.788
	0.01	0.989	0.04	1.114	0.862
Range Based MEF: $\hat{\sigma}^2_{range}$	0.5	0.914	0.024	0.997	0.844
	0.1	0.974	0.026	1.068	0.902
	0.01	0.988	0.026	1.08	0.921

Table 1: Estimators of the diffusion parameter σ^2 from 1000 trajectories

the drift estimation with ordinary martingale estimating functions can be found in [11]. The mean value and the standard deviation of the different estimators are given in Table 1. The columns labeled "min" and "max" indicate the range of the 1000 estimators we calculated.

In the sequel, we display the different types of martingale estimating functions for the diffusion coefficient that were used to generate Table 1. First, we compare the estimator inferred from the ordinary martingale estimating function to the one inferred from a generalized martingale estimating function. Then, we compare the ordinary estimator to a range based estimator.

5.1 Ordinary vs. Generalized Martingale Estimating Functions

We assume that $\mu \equiv 1$ is known and we compare different estimators of σ^2. We compare the ordinary martingale estimating function

$$g_{ord}(\Delta, x, h, l, y; \theta) = a_{ord}(\Delta, x; \theta)\kappa_{yy}(\Delta, x, h, l, y; \theta) \tag{95}$$

with the generalized MEFs

$$g_{qua}(\Delta, x, h, l, y; \theta) = \sum_{j \in M_{qua}}^{3} a_j^{qua}(\Delta, x; \theta)\kappa_j(\Delta, x, h, l, y; \theta), \tag{96}$$

and

$$g^a_{qua}(\Delta, x, h, l, y; \theta) = \sum_{j \in M^a_{qua}} a_j^{qua,a}(\Delta, x; \theta)\kappa_j(\Delta, x, h, l, y; \theta), \tag{97}$$

respectively, where the functions κ_j are the functions described in (24). The relevant index sets are defined by $M_{qua} = \{hh, hy, yy\}$ and $M^a_{qua}\{hh, ll, yy, hl, hy, ly\}$.

We do not only replace the weights by their small-Δ-optimal analogue, but we approximate the complete martingale estimating functions (95)–(97) by their second order expansion with respect to $\sqrt{\Delta}$. The concrete expansions we take into account are

$$\begin{aligned}
&\big(y - \mathbb{E}_{x,\theta}[X_\Delta]\big)^2 - \mathrm{Var}_{x,\theta}[X_\Delta] \\
&\quad = y^2 - 2y(x + \mu(x)\Delta) + 2(x^2 + 2x\mu(x)\Delta) - (x^2 + 2x\mu(x)\Delta + \sigma^2\Delta) + O(\Delta^2), \\
&\big(h - \mathbb{E}_{x,\theta}[H_\Delta]\big)^2 - \mathrm{Var}_{x,\theta}[H_\Delta] \\
&\quad = h^2 - 2h\left(x + \sqrt{\frac{2}{\pi}}\sigma\sqrt{\Delta} + \frac{1}{2}\mu(x)\Delta\right) \\
&\qquad + 2\left(x^2 + \frac{2}{\pi}\sigma^2\Delta + 2x\sqrt{\frac{2}{\pi}}\sigma\sqrt{\Delta} + x\mu(x)\Delta\right) \\
&\qquad - \left(x^2 + 2x\sqrt{\frac{2}{\pi}}\sigma\sqrt{\Delta} + x\mu(x)\Delta + \sigma^2\Delta\right) + O(\Delta^{3/2})
\end{aligned} \tag{98}$$

and

$$\begin{aligned}
&\big(h - \mathbb{E}_{x,\theta}[H_\Delta]\big)\big(y - \mathbb{E}_{x,\theta}[X_\Delta]\big) - \mathrm{Cov}_{x,\theta}[H_\Delta, X_\Delta] \\
&\quad = hy - y\left(x + \sqrt{\frac{2}{\pi}}\sigma\sqrt{\Delta} + \frac{1}{2}\mu(x)\Delta\right) - 2h(x + \mu(x)\Delta) \\
&\qquad + 2\left(x^2 + x\sqrt{\frac{2}{\pi}}\sigma\sqrt{\Delta} + \frac{3}{2}\mu(x)\Delta\right) \\
&\qquad - \left(x^2 + x\sqrt{\frac{2}{\pi}}\sigma\sqrt{\Delta} + \frac{3}{2}\mu(x)\Delta + \frac{1}{2}\sigma^2\Delta\right) + O(\Delta^{3/2}).
\end{aligned} \tag{99}$$

For the estimating functions including the random variable L_Δ one obtains similar expression. The small-Δ-optimal weights for the generalized martingale estimating function g_{qua} and g^a_{qua} can be chosen according to (62) and to (72), respectively.

Comparing the lines for $\Delta = 0.01$ in Table 1, we see that the estimator $\hat{\sigma}^2_{qua}$ inferred from the generalized martingale estimating function (96) has slightly smaller bias and significantly smaller standard deviation than the estimator $\hat{\sigma}^2_{ord}$ inferred from the ordinary martingale estimating function (95). Let us compare the mean squared errors of $\hat{\sigma}^2_{qua}$ and $\hat{\sigma}^2_{ord}$ for $\Delta = 0.01$. The quotient of both quantities is

$$\frac{\mathbb{E}[(\hat{\sigma}^2_{qua} - \sigma^2)^2]}{\mathbb{E}[(\hat{\sigma}^2_{ord} - \sigma^2)^2]} = 0.833752. \tag{100}$$

The standard deviation of the estimator $\left(\hat{\sigma}^a_{qua}\right)^2$, inferred from the martingale estimating function (97) that consists of all quadratic terms, is even smaller. We obtain the quotients

$$\frac{\mathbb{E}[(\left(\hat{\sigma}^a_{qua}\right)^2 - \sigma^2)^2]}{\mathbb{E}[(\hat{\sigma}^2_{ord} - \sigma^2)^2]} = 0.414096 \tag{101}$$

and

$$\frac{\mathbb{E}[(\left(\hat{\sigma}^a_{qua}\right)^2 - \sigma^2)^2]}{\mathbb{E}[(\hat{\sigma}^2_{qua} - \sigma^2)^2]} = 0.496667. \tag{102}$$

Evidently, both generalized martingale estimating functions are superior to the ordinary martingale estimating function, even though we do not exactly discover the gain in efficiency that was predicted by our theoretical results – see Paragraph 4 above. This mismatch between theory and practice might be due to the discretization of the trajectories. Consequently, the suprema are underestimated and the infima are overestimated. Another source of inaccuracy is the simulation of the matrix (65) and the weights (72). As a result, the lower bound on the right hand side of (69) might be too low. Alternatively, with more accurately calculated weights $a_1^{qua,a}, \ldots, a_6^{qua,a}$ in (72) the standard deviation in Table 1 for the estimator derived from the generalized martingale estimating function g^a_{qua} might be smaller.

5.2 Ordinary vs. Range Based Martingale Estimating Functions

Again, we assume that $\mu \equiv 1$ is known. Here, we compare the ordinary quadratic martingale estimating function g_{ord} given by (95) to the range based martingale estimating function

$$g_{range}(\Delta, x, h, l, y; \theta) = \sum_{j \in M_{range}} a_j^{range}(\Delta, x; \theta)\kappa_j(\Delta, x, h, l, y; \theta), \tag{103}$$

where the functions κ_j are defined by (24) and the index set is given by $M_{range} = \{h, l\}$. Again, the range-based martingale estimating function is approximated by its second order expansion with respect to $\sqrt{\Delta}$, cf. (6.181) and (6.182) in [8]. Moreover, the small-Δ-optimal weights can be chosen according to (91). The results for the range based estimators for the true parameter $\sigma^2 = 1$ for different values of Δ are displayed in the last block of Table 1. We compare them to the estimators inferred from the ordinary martingale estimating function. As we expected, for small values of Δ, the range based estimator has smaller variance than the estimator inferred from the ordinary quadratic estimating function $g_{qua,ord}$. It is also superior to the estimator inferred from the generalized quadratic martingale estimating function $g_{qua,gen}$. Concretely, a comparison of the biases and the standard deviations for $\Delta = 0.01$ in Table 1 shows that

$$\frac{\mathbb{E}[(\hat{\sigma}^2_{range} - \sigma^2)^2]}{\mathbb{E}[(\hat{\sigma}^2_{ord} - \sigma^2)^2]} = 0.20778, \tag{104}$$

where $\hat{\sigma}_{range}$ denotes the estimator inferred from (103). This means that the mean squared error for the range-based model is about 80 % lower than the mean squared error for the ordinary estimating function. This almost corresponds to the theoretical values we discovered in Section 4.

It would be interesting to know if this effect carries over to martingale estimating function constructed with triplets of observations $(H_\Delta, L_\Delta, X_\Delta)$. We conjecture that the lower bound of the variance of such an estimating function is even smaller than the one of the range-based martingale estimating function obtained from the pair (H_Δ, L_Δ).

5.3 Annotations

In our simulation study we did not only replace the optimal weights $a_1(\Delta, x; \theta)$, $a_2(\Delta, x; \theta)$ and $a_3(\Delta, x; \theta)$ by the respective small-Δ-optimal weights, but we also approximated the expectations

$$\mathbb{E}_{x,\theta}[H_\Delta], \quad \mathbb{E}_{x,\theta}[L_\Delta] \quad \text{and} \quad \mathbb{E}_{x,\theta}[X_\Delta], \tag{105}$$

and the respective covariances by their second order approximations with respect to $\sqrt{\Delta}$. As we saw above, the resulting estimators were biased. This is due to the fact that the martingale property of the estimating functions is destroyed by the additional approximation. In order to analyse the theoretical behavior of such two-fold approximated MEFs requires to consider the simultaneous asymptotics $\Delta_n \to 0$ and $n\Delta_n \to \infty$ for $n \to \infty$. For ordinary estimating functions, consistency of the resulting estimators is proved in the article of Florens-Zmirou [4] in this scenario. Moreover, asymptotic normality can be stated on the further condition $n\Delta_n^3 \to 0$. Related results were also proved by Yoshida [21]. And finally, Kessler [12] used higher order expansions of the moments of the transition distribution to obtain estimators that are asymptotically normal, even when Δ_n tends more slowly to zero as n tends to ∞. Analogous results for our generalized martingale estimating functions remain to be studied.

6 Technical Results

6.1 Foundations

We give a brief heuristic of how to prove formula (19). Let $x \in \mathbb{R}$, $\Delta > 0$, and, for the time being, let us assume that μ and σ are real valued constants. Moreover, let Z denote the Brownian motion with drift $Z_t = \mu t + \sigma B_t$, $0 \leq t \leq \Delta$. The joint density of the process $(\sup_{0 \leq t \leq \Delta} Z_t, Z_\Delta)$, starting in (x, x), is given by

$$f_{(\mu,\sigma)}(\Delta, x, h, y) = \frac{2(2h - x - y)}{\sqrt{2\pi\Delta^3}\sigma^3} \exp\left(\frac{\mu}{\sigma^2}(y - x) - \frac{(2h - x - y)^2}{2\Delta\sigma^2} - \frac{\mu^2}{2\sigma^2}\Delta\right). \tag{106}$$

See e.g. [2]. ¿From this density it can be inferred that, for a sufficiently smooth function $g : \mathbb{R} \to \mathbb{R}$ that does not grow too fast,

$$\mathbb{E}_x \left[g \left(\sup_{0 \le t \le \Delta} Z_t \right) \right] = g(x) + \sigma \sqrt{\frac{2\Delta}{\pi}} \frac{\partial}{\partial x} g(x) + \Delta \left(\frac{1}{2} \mu \frac{\partial}{\partial x} g(x) + \frac{1}{2} \sigma^2 \frac{\partial^2}{\partial x^2} g(x) \right) + O(\Delta^3). \tag{107}$$

Now, let us consider the process Y defined by (16). From the estimate

$$\begin{aligned} &\mathbb{E}_{x,\theta} \left[\sup_{0 \le t \le \Delta} \Big| Y_t - \sigma(x;\theta) \Big\{ \mu^Y(x) t + B_t \Big\} \Big|^2 \right] \\ &\le \sigma(x;\theta)^2 \, \mathbb{E}_{x,\theta} \left[\sup_{0 \le t \le \Delta} \Big| \int_0^t \{ \mu^Y(X_s) - \mu^Y(x) \} ds \Big|^2 \right] \\ &\le \text{ const. } \sigma(x;\theta)^2 \Delta^3, \end{aligned} \tag{108}$$

which is a direct consequence of the Lipschitz property of μ^Y (recall Assumption 2.2.1) and Cauchy-Schwarz' inequality, in combination with (107) it follows easily that

$$\begin{aligned} &\mathbb{E}_x \left[g \left(\sup_{0 \le t \le \Delta} Y_t \right) \right] \\ &= x + \sigma(x;\theta) \sqrt{\frac{2\Delta}{\pi}} \frac{\partial}{\partial x} g(x) \\ &+ \Delta \left(\frac{1}{2} \left\{ \mu(x;\theta) - \frac{1}{2} \sigma(x;\theta) \frac{\partial}{\partial x} \sigma(x;\theta) \right\} \frac{\partial}{\partial x} g(x) + \frac{1}{2} \sigma^2(x;\theta) \frac{\partial^2}{\partial x^2} g(x) \right) \\ &+ O(\Delta^{3/2}). \end{aligned} \tag{109}$$

From this expansion, one can easily infer equation (19).

6.2 Expansions

The aim of the present paragraph is to find formulae analogous to (46) and (47) in Proposition 5 in [9] for generalized martingale estimating functions. Concretely, we derive expansion for both the expressions (9) and (10) with respect to $\sqrt{\Delta}$ in different classes of estimating functions and for the process Y. We start with a simple statement that will turn out to be crucial in the sequel.

Proposition 6.2.1. *Let Y denote the process defined by the stochastic differential equation* (16) *and let $\mathcal{G}_\theta$ be a class of flows that satisfies Assumption 2.2.2. Then, for $g \in \mathcal{G}_\theta$, we have*

$$\begin{aligned} 0 &= \mathbb{E}_{x,\theta} \big[g_{\Delta,\theta} \big(Y_0, H_t^Y, L_t^Y, Y_t \big) \big] \\ &= g_{0,\theta}(\mathbf{x}) + \sqrt{\Delta} \mathcal{A}_\theta^{(\frac{1}{2})} g_{0,\theta}(\mathbf{x}) + \sqrt{\Delta} \tilde{g}_{0,\theta}(\mathbf{x}) + \Delta \mathcal{A}_\theta^{(\frac{1}{2})} \tilde{g}_{0,\theta}(\mathbf{x}) \\ &\quad + \Delta \mathcal{A}_\theta^{(1)} g_{0,\theta}(\mathbf{x}) + \Delta \tilde{\tilde{g}}_{0,\theta}(\mathbf{x}) + O(\Delta^{3/2}), \end{aligned} \tag{110}$$

where the operators $\mathcal{A}_\theta^{(\frac{1}{2})}$ *and* $\mathcal{A}_\theta^{(1)}$ *are given by* (20) *and* (21)*, respectively.*

Proof. By Assumption 2.2.2 the function g has an expansion with respect to $\sqrt{\Delta}$ of the following form

$$\begin{aligned} & g_{\Delta,\theta}(x,h,l,y) \\ & = g_{0,\theta}(x,h,l,y) + \sqrt{\Delta}\tilde{g}_{0,\theta}(x,h,l,y) + \Delta\tilde{\tilde{g}}_{0,\theta}(x,h,l,y) + O\big(\Delta^{3/2};\theta,x,h,l,y\big). \end{aligned} \tag{111}$$

In view of the fact that the rest term $O\big(\Delta^{3/2};\, \theta,x,h,l,y\big)$ behaves like a polynomial in h, l and y, the result can be inferred by an expansion of each of the terms

$$\mathbb{E}_x\Big[\tilde{g}_{0,\theta}^{(k)}\big(Y_0,H_t^Y,L_t^Y,Y_t\big)\Big], \quad k=0,1,2. \tag{112}$$

A comparison with Corollary 5.2.3.6 in [8] yields the result. □

Before we proceed, let us introduce the following notation

$$\mathbf{D}_h\, g_{0,\theta}(\mathbf{x}) = \frac{\partial}{\partial h} g_{0,\theta}(x,h,x,x)\Big|_{h=x} \quad \text{and} \quad \mathbf{D}_{hh}\, g_{0,\theta}(\mathbf{x}) = \frac{\partial^2}{\partial h^2} g_{0,\theta}(x,h,x,x)\Big|_{h=x}. \tag{113}$$

The expressions $\mathbf{D}_l\, g_{0,\theta}(\mathbf{x})$, $\mathbf{D}_y\, g_{0,\theta}(\mathbf{x})$ and the terms $\mathbf{D}_{hl}\, g_{0,\theta}(\mathbf{x})$, $\mathbf{D}_{hy}\, g_{0,\theta}(\mathbf{x})$, ... etc. are defined in an analogous way.

An immediate consequence of the previous proposition is displayed in the following corollary.

Corollary 6.2.2. *Let the assumptions of Proposition 6.2.1 be satisfied. Then*

$$\begin{aligned} 0 &= \mathbb{E}_{x,\theta}\big[g_{\Delta,\theta}\big(Y_0,H_t^Y,L_t^Y,Y_t\big)\big] \\ &= g_{0,\theta}(\mathbf{x}) + \sqrt{\Delta}\mathcal{A}_\theta^{(\frac{1}{2})} g_{0,\theta}(\mathbf{x}) + \sqrt{\Delta}\tilde{g}_{0,\theta}(\mathbf{x}) - \Delta\left(\mathcal{A}_\theta^{(\frac{1}{2})}\right)^2 \tilde{g}_{0,\theta}(\mathbf{x}) \\ &\quad + \Delta\mathcal{A}_\theta^{(1)} g_{0,\theta}(\mathbf{x}) + \Delta\tilde{\tilde{g}}_{0,\theta}(\mathbf{x}) + O(\Delta^{3/2}). \end{aligned} \tag{114}$$

where the operator $\left(\mathcal{A}_\theta^{(\frac{1}{2})}\right)^2$ *is given by*

$$\left(\mathcal{A}_\theta^{(\frac{1}{2})}\right)^2 g_{0,\theta}(\mathbf{x}) = \sigma(x;\theta)^2\frac{2}{\pi}\mathbf{D}_{hh}\, g_{0,\theta}(\mathbf{x}) + \sigma(x;\theta)^2\frac{2}{\pi}\mathbf{D}_{ll}\, g_{0,\theta}(\mathbf{x}) - \sigma(x;\theta)^2\frac{4}{\pi}\mathbf{D}_{hl}\, g_{0,\theta}(\mathbf{x})\,. \tag{115}$$

Proof. Formula (110) shows that, by letting $\Delta \to 0$, $g_{0,\theta}(\mathbf{x}) = 0$. Moreover, one infers that

$$\mathcal{A}_\theta^{(\frac{1}{2})} g_{0,\theta}(\mathbf{x}) + \tilde{g}_{0,\theta}(\mathbf{x}) = 0, \tag{116}$$

$$\mathcal{A}_\theta^{(\frac{1}{2})}\tilde{g}_{0,\theta}(\mathbf{x}) + \mathcal{A}_\theta^{(1)} g_{0,\theta}(\mathbf{x}) + \tilde{\tilde{g}}_{0,\theta}(\mathbf{x}) = 0. \tag{117}$$

Note that equation (117) is equivalent to

$$-\left(\mathcal{A}_\theta^{(\frac{1}{2})}\right)^2 g_{0,\theta}(\mathbf{x}) + \mathcal{A}_\theta^{(1)} g_{0,\theta}(\mathbf{x}) + \tilde{\tilde{g}}_{0,\theta}(\mathbf{x}) = 0, \tag{118}$$

where one can easily see that the operator $\left(\mathcal{A}_\theta^{(\frac{1}{2})}\right)^2$ coincides with the operator (115). This follows directly by means of equation (116). The result now follows directly from formulae (116) and (118). □

We are now in a position to derive our first expansion. The next proposition states the analogue of formula (46) in [9] for generalized martingale estimating functions.

Proposition 6.2.3. *Let the assumptions of Proposition 6.2.1 be satisfied. Then*

$$\begin{aligned}\mathbb{E}_{x,\theta}\left[\frac{\partial}{\partial\theta} g_{\Delta,\theta}\big(Y_0, H_\Delta^Y, L_\Delta^Y, Y_\Delta\big)\right] = &- \sqrt{\Delta}\,\dot{\mathcal{A}}_\theta^{(\frac{1}{2})} g_{0,\theta}(\mathbf{x}) + \Delta \dot{\left(\mathcal{A}_\theta^{(\frac{1}{2})}\right)^2} g_{0,\theta}(\mathbf{x}) \\ &- \Delta\dot{\mathcal{A}}_\theta^{(1)} g_{0,\theta}(\mathbf{x}) + O\big(\Delta^{3/2}\big),\end{aligned} \tag{119}$$

where the operators $\dot{\mathcal{A}}_\theta^{(\frac{1}{2})}$ *and* $\dot{\left(\mathcal{A}_\theta^{(\frac{1}{2})}\right)^2}$ *are given by*

$$\dot{\mathcal{A}}_\theta^{(\frac{1}{2})} g_{0,\theta}(\mathbf{x}) = \frac{\partial}{\partial\theta}\sigma(x;\theta)\frac{2}{\sqrt{2\pi}}\mathbf{D}_h\, g_{0,\theta}(\mathbf{x}) - \frac{\partial}{\partial\theta}\sigma(x;\theta)\frac{2}{\sqrt{2\pi}}\mathbf{D}_l\, g_{0,\theta}(\mathbf{x}), \tag{120}$$

and

$$\begin{aligned}\dot{\left(\mathcal{A}_\theta^{(\frac{1}{2})}\right)^2} g_{0,\theta}(\mathbf{x}) =& \frac{1}{2}\frac{\partial}{\partial\theta}\sigma(x;\theta)^2\frac{2}{\pi}\mathbf{D}_{hh}\, g_{0,\theta}(\mathbf{x}) + \frac{1}{2}\frac{\partial}{\partial\theta}\sigma(x;\theta)^2\frac{2}{\pi}\mathbf{D}_{ll}\, g_{0,\theta}(\mathbf{x}) \\ &- \frac{\partial}{\partial\theta}\sigma(x;\theta)^2\frac{2}{\pi}\mathbf{D}_{hl}\, g_{0,\theta}(\mathbf{x}),\end{aligned} \tag{121}$$

respectively. Finally, $\dot{\mathcal{A}}^{(1)}$ *is given by*

$$\begin{aligned}\dot{\mathcal{A}}_\theta^{(1)} g_{0,\theta}(\mathbf{x}) =& \frac{1}{2}\left(\frac{\partial}{\partial\theta}\mu(x;\theta) - \frac{1}{2}\frac{\partial}{\partial\theta}\left\{\sigma(x;\theta)\frac{\partial}{\partial x}\sigma(x;\theta)\right\}\right)\mathbf{D}_h\, g_{0,\theta}(\mathbf{x}) \\ &+ \frac{1}{2}\left(\frac{\partial}{\partial\theta}\mu(x;\theta) - \frac{1}{2}\frac{\partial}{\partial\theta}\left\{\sigma(x;\theta)\frac{\partial}{\partial x}\sigma(x;\theta)\right\}\right)\mathbf{D}_l\, g_{0,\theta}(\mathbf{x}) \\ &+ \frac{1}{2}\frac{\partial}{\partial\theta}\sigma^2(x;\theta)\mathbf{D}_{hh}\, g_{0,\theta}(\mathbf{x}) + \frac{1}{2}\frac{\partial}{\partial\theta}\sigma^2(x;\theta)\mathbf{D}_{ll}\, g_{0,\theta}(\mathbf{x}) \\ &+ (1 - 2\log 2)\frac{\partial}{\partial\theta}\sigma^2(x;\theta)\mathbf{D}_{hl}\, g_{0,\theta}(\mathbf{x}) \\ &+ \frac{1}{2}\frac{\partial}{\partial\theta}\sigma^2(x;\theta)\mathbf{D}_{hy}\, g_{0,\theta}(\mathbf{x}) + \frac{1}{2}\frac{\partial}{\partial\theta}\sigma^2(x;\theta)\mathbf{D}_{ly}\, g_{0,\theta}(\mathbf{x}) \\ &+ \left(\frac{\partial}{\partial\theta}\mu(x;\theta) - \frac{1}{2}\frac{\partial}{\partial\theta}\left\{\sigma(x;\theta)\frac{\partial}{\partial x}\sigma(x;\theta)\right\}\right)\mathbf{D}_y\, g_{0,\theta}(\mathbf{x}) \\ &+ \frac{1}{2}\frac{\partial}{\partial\theta}\sigma(x;\theta)^2\mathbf{D}_{yy}\, g_{0,\theta}(\mathbf{x})\,.\end{aligned} \tag{122}$$

Proof. Clearly, the following expansion holds

$$\begin{aligned}
&\mathbb{E}_{x,\theta}\left[\frac{\partial}{\partial\theta}g_{\Delta,\theta}\big(Y_0,H_\Delta^Y,L_\Delta^Y,Y_\Delta\big)\right]\\
&=\Bigg\{\frac{\partial}{\partial\theta}g_{0,\theta}+\sqrt{\Delta}\mathcal{A}_\theta^{(\frac{1}{2})}\frac{\partial}{\partial\theta}g_{0,\theta}+\sqrt{\Delta}\frac{\partial}{\partial\theta}\tilde{g}_{0,\theta}\\
&\qquad+\Delta\mathcal{A}^{(\frac{1}{2})}\frac{\partial}{\partial\theta}\tilde{g}_{0,\theta}+\Delta\mathcal{A}_\theta^{(1)}\frac{\partial}{\partial\theta}g_{0,\theta}+\Delta\frac{\partial}{\partial\theta}\tilde{\tilde{g}}_{0,\theta}\Bigg\}(\mathbf{x})+O(\Delta^{3/2}).
\end{aligned}\tag{123}$$

Now, $\frac{\partial}{\partial\theta}g_{0,\theta}(\mathbf{x})\equiv 0$ and

$$0=\frac{\partial}{\partial\theta}\Big(\mathcal{A}_\theta^{(\frac{1}{2})}g_{0,\theta}(\mathbf{x})+\tilde{g}_{0,\theta}(\mathbf{x})\Big)=\Big(\mathcal{A}_\theta^{(\frac{1}{2})}\frac{\partial}{\partial\theta}g_{0,\theta}(\mathbf{x})+\frac{\partial}{\partial\theta}\tilde{g}_{0,\theta}(\mathbf{x})\Big)+\dot{\mathcal{A}}_\theta^{(\frac{1}{2})}g_{0,\theta}(\mathbf{x}),\tag{124}$$

where $\dot{\mathcal{A}}_\theta^{(\frac{1}{2})}$ is the operator (120). Analogously, we find that

$$\begin{aligned}
0=&\frac{\partial}{\partial\theta}\left\{-\left(\mathcal{A}_\theta^{(\frac{1}{2})}\right)^2g_{0,\theta}(\mathbf{x})+\mathcal{A}_\theta^{(1)}g_{0,\theta}(\mathbf{x})+\tilde{\tilde{g}}_{0,\theta}(\mathbf{x})\right\}\\
=&\left\{-\left(\mathcal{A}_\theta^{(\frac{1}{2})}\right)^2\frac{\partial}{\partial\theta}g_{0,\theta}(\mathbf{x})+\mathcal{A}_\theta^{(1)}\frac{\partial}{\partial\theta}g_{0,\theta}(\mathbf{x})+\frac{\partial}{\partial\theta}\tilde{\tilde{g}}_{0,\theta}(\mathbf{x})\right\}\\
&-\dot{\left(\mathcal{A}_\theta^{(\frac{1}{2})}\right)^2}g_{0,\theta}(\mathbf{x})+\dot{\mathcal{A}}_\theta^{(1)}g_{0,\theta}(\mathbf{x}),
\end{aligned}\tag{125}$$

where $\dot{\mathcal{A}}^{(1)}$ is given by (122). Since

$$-\dot{\left(\mathcal{A}_\theta^{(\frac{1}{2})}\right)^2}g_{0,\theta}(\mathbf{x})=\dot{\mathcal{A}}_\theta^{(\frac{1}{2})}\tilde{g}_{0,\theta}(\mathbf{x}),\tag{126}$$

inserting (124) and (125) into equation (123) yields the result. □

Remark 6.2.4. If the diffusion coefficient σ of the diffusion (1) does not depend on the parameter θ, the operator $\dot{\mathcal{A}}_\theta^{(\frac{1}{2})}$ vanishes, and we have the following result

$$\mathbb{E}_{x,\theta}\left[\frac{\partial}{\partial\theta}g_{\Delta,\theta}\big(Y_0,H_\Delta^Y,L_\Delta^Y,Y_\Delta\big)\right]=\Delta\dot{\mathcal{A}}_\theta^{(1)}g_{0,\theta}(\mathbf{x})+O\big(\Delta^{3/2}\big).\tag{127}$$

The operator $\dot{\mathcal{A}}^{(1)}$ also takes a simple form in this case, namely

$$\begin{aligned}
&\dot{\mathcal{A}}_\theta^{(1)}g_{0,\theta}(\mathbf{x})\\
&=\frac{1}{2}\frac{\partial}{\partial\theta}\mu(x;\theta)\mathbf{D}_h\,g_{0,\theta}(\mathbf{x})+\frac{1}{2}\frac{\partial}{\partial\theta}\mu(x;\theta)\mathbf{D}_l\,g_{0,\theta}(\mathbf{x})+\frac{\partial}{\partial\theta}\mu(x;\theta)\mathbf{D}_y\,g_{0,\theta}(\mathbf{x}).
\end{aligned}\tag{128}$$

A similar result can be stated if μ does not depend on θ. In any case, we have the first order expansion

$$\mathbb{E}_{x,\theta}\left[\frac{\partial}{\partial\theta}g_{\Delta,\theta}\big(Y_0,H_\Delta^Y,L_\Delta^Y,Y_\Delta\big)\right]=\sqrt{\Delta}\dot{\mathcal{A}}_\theta^{(\frac{1}{2})}g_{0,\theta}(\mathbf{x})+O\big(\Delta\big).\tag{129}$$

The next result states the analogue of formula (47) in [9].

Proposition 6.2.5. *Let the assumptions of Proposition 6.2.1 be satisfied. Then*

$$\mathbb{E}_{x,\theta}\Big[g^2_{\Delta,\theta}\big(Y_0, H^Y_\Delta, L^Y_\Delta, Y_\Delta\big)\Big] = \Delta\mathcal{A}^S g_{0,\theta}(\mathbf{x}) + O\big(\Delta^{3/2}\big), \tag{130}$$

where the operator $\mathcal{A}^S$ *is defined via*

$$\begin{aligned}\mathcal{A}^S g_{0,\theta}(\mathbf{x}) = &\left(1-\frac{2}{\pi}\right)(\mathbf{D}_h\, g_{0,\theta}(\mathbf{x}))^2\,\sigma(x;\theta)^2 + \left(1-\frac{2}{\pi}\right)\Big(\mathbf{D}_l\, g_{0,\theta}(\mathbf{x})\Big)^2\sigma(x;\theta)^2\\ &+ \left(\frac{4}{\pi}+2\,(1-2\log 2)\right)\Big(\mathbf{D}_h\, g_{0,\theta}(\mathbf{x})\mathbf{D}_l\, g_{0,\theta}(\mathbf{x})\Big)\sigma(x;\theta)^2\\ &+\Big(\mathbf{D}_h\, g_{0,\theta}(\mathbf{x})\mathbf{D}_y\, g_{0,\theta}(\mathbf{x})\Big)\sigma(x;\theta)^2 + \Big(\mathbf{D}_l\, g_{0,\theta}(\mathbf{x})\mathbf{D}_y\, g_{0,\theta}(\mathbf{x})\Big)\sigma(x;\theta)^2\\ &+\Big(\mathbf{D}_y\, g_{0,\theta}(\mathbf{x})\Big)^2\sigma(x;\theta)^2.\end{aligned} \tag{131}$$

Proof. Because $g_{0,\theta}(\mathbf{x}) = 0$, we have

$$\mathbf{D}_h\, g^2_{0,\theta}(\mathbf{x}) = 2g_{0,\theta}(\mathbf{x})\mathbf{D}_h\, g_{0,\theta}(\mathbf{x}) = 0. \tag{132}$$

And similarly, we have

$$\mathbf{D}_l\, g^2_{0,\theta}(\mathbf{x}) = \mathbf{D}_y\, g^2_{0,\theta}(\mathbf{x}) = 0. \tag{133}$$

It remains to show that

$$\widetilde{g^2_{0,\theta}}(\mathbf{x}) = 0. \tag{134}$$

But this is evident, since

$$\frac{\partial}{\partial s} g^2_{s,\theta}(\mathbf{x}) = 2g_{s,\theta}(\mathbf{x})\frac{\partial}{\partial s} g_{s,\theta}(\mathbf{x}), \tag{135}$$

which implies

$$\widetilde{g^2_{0,\theta}}(\mathbf{x}) = 2g_{0,\theta}(\mathbf{x})\tilde{g}_{0,\theta}(\mathbf{x}) = 0. \tag{136}$$

Therefore, the $\sqrt{\Delta}$-term in the expansion of $\mathbb{E}_{x,\theta}\big[g^2_{\Delta,\theta}\big]$ vanishes and the Δ-term becomes

$$\begin{aligned}&-\left(\mathcal{A}_\theta^{(\frac{1}{2})}\right)^2 g^2_{0,\theta}(\mathbf{x}) + \mathcal{A}_\theta^{(1)} g^2_{0,\theta}(\mathbf{x}) + \widetilde{\widetilde{g^2_{0,\theta}}}(\mathbf{x})\\ &= -\sigma(x;\theta)^2\frac{2}{\pi}\Big\{\mathbf{D}_{hh}\, g^2_{0,\theta}(\mathbf{x}) + \mathbf{D}_{ll}\, g^2_{0,\theta}(\mathbf{x})\Big\} + \Big\{\frac{4}{\pi}+1-2\log 2\Big\}\sigma(x;\theta)^2\mathbf{D}_{hl}\, g^2_{0,\theta}(\mathbf{x})\\ &\quad+\frac{1}{2}\sigma(x;\theta)^2\Big\{\mathbf{D}_{hh}\, g^2_{0,\theta}(\mathbf{x}) + \mathbf{D}_{ll}\, g^2_{0,\theta}(\mathbf{x}) + \mathbf{D}_{hy}\, g^2_{0,\theta}(\mathbf{x}) + \mathbf{D}_{ly}\, g^2_{0,\theta}(\mathbf{x}) + \mathbf{D}_{yy}\, g^2_{0,\theta}(\mathbf{x})\Big\}\\ &\quad+\widetilde{\widetilde{g^2_{0,\theta}}}(x,x,x,x).\end{aligned} \tag{137}$$

It remains to calculate the derivatives on the right hand side of (137). First,

$$\mathbf{D}_h\, g^2_{0,\theta}(\mathbf{x}) = 2g_{0,\theta}(\mathbf{x})\mathbf{D}_{hh}\, g_{0,\theta}(\mathbf{x}) + 2\left(\mathbf{D}_h\, g_{0,\theta}(\mathbf{x})\right)^2 = 2\left(\mathbf{D}_h\, g_{0,\theta}(\mathbf{x})\right)^2. \tag{138}$$

The last equality in the previous equation (138) follows because $g_{0,\theta}(\mathbf{x}) = 0$. Analogously, we find

$$\mathbf{D}_{ll}\, g^2_{0,\theta}(\mathbf{x}) = 2\left(\mathbf{D}_l\, g_{0,\theta}(\mathbf{x})\right)^2,\ \mathbf{D}_{yy}\, g^2_{0,\theta}(\mathbf{x}) = 2\left(\mathbf{D}_y\, g_{0,\theta}(\mathbf{x})\right)^2. \tag{139}$$

For the first cross-term, we have

$$\mathbf{D}_{hl}\, g^2_{0,\theta}(\mathbf{x}) = 2g_{0,\theta}(\mathbf{x})\mathbf{D}_{hl}\, g_{0,\theta}(\mathbf{x}) + 2\mathbf{D}_h g_{0,\theta}(\mathbf{x})\mathbf{D}_l g_{0,\theta}(\mathbf{x}) = 2\mathbf{D}_h g_{0,\theta}(\mathbf{x})\mathbf{D}_l g_{0,\theta}(\mathbf{x}). \tag{140}$$

And analogously, one obtains

$$\mathbf{D}_{hy}\, g^2_{0,\theta}(\mathbf{x}) = 2\mathbf{D}_h g_{0,\theta}(\mathbf{x})\mathbf{D}_y g_{0,\theta}(\mathbf{x}),\ \mathbf{D}_{ly}\, g^2_{0,\theta}(\mathbf{x}) = 2\mathbf{D}_l g_{0,\theta}(\mathbf{x})\mathbf{D}_y g_{0,\theta}(\mathbf{x}). \tag{141}$$

Finally, let us consider the term $\widetilde{\widetilde{g^2_{0,\theta}}}(\mathbf{x})$. By (23), we obtain the following equation

$$\begin{aligned} g^2_{\Delta,\theta}(x,h,l,y) &= \Big(g_{0,\theta}(x,h,l,y) + \sqrt{\Delta}\widetilde{g}_{0,\theta}(x,h,l,y) + \Delta\widetilde{\widetilde{g}}_{0,\theta}(x,h,l,y) + O(\Delta^{3/2})\Big)^2 \\ &= g^2_{0,\theta}(x,h,l,y) + 2\sqrt{\Delta}g_{0,\theta}(x,h,l,y)\widetilde{g}_{0,\theta}(x,h,l,y) \\ &\quad + \Delta\left\{\Big(\widetilde{g}_{0,\theta}(x,h,l,y)\Big)^2 + 2g_{0,\theta}(x,h,l,y)\right\} + O(\Delta^{3/2}). \end{aligned} \tag{142}$$

By the definitions of the square-root-derivatives in formula (22), the Δ-term in (142) corresponds to $\widetilde{\widetilde{g}}^2$, which shows that

$$\widetilde{\widetilde{g^2_{0,\theta}}}(\mathbf{x}) = \Big(\widetilde{g}_{0,\theta}(\mathbf{x})\Big)^2 + 2g_{0,\theta}(\mathbf{x}). \tag{143}$$

By the fact that $g_{0,\theta}(\mathbf{x}) = 0$ and by the definition of $\mathcal{A}^{(\frac{1}{2})}$, we obtain the final equation

$$\begin{aligned} \widetilde{\widetilde{g^2_{0,\theta}}}(\mathbf{x}) &= \Big(\widetilde{g}_{0,\theta}(\mathbf{x})\Big)^2 = \Big(-\mathcal{A}^{(\frac{1}{2})}g_{0,\theta}(\mathbf{x})\Big)^2 = \left(\sigma(x;\theta)\sqrt{\frac{2}{\pi}}\Big\{-\mathbf{D}_h\, g_{0,\theta}(\mathbf{x}) + \mathbf{D}_l\, g_{0,\theta}(\mathbf{x})\Big\}\right)^2 \\ &= \sigma(x;\theta)^2\frac{2}{\pi}\Big\{\big(\mathbf{D}_h\, g_{0,\theta}(\mathbf{x})\big)^2 + \big(\mathbf{D}_l\, g_{0,\theta}(\mathbf{x})\big)^2 - 2\mathbf{D}_h\, g_{0,\theta}(\mathbf{x})\mathbf{D}_l\, g_{0,\theta}(\mathbf{x})\Big\}. \end{aligned} \tag{144}$$

Inserting the above terms for the derivatives of $g^2_{0,\theta}$ and for $\widetilde{\widetilde{g^2_{0,\theta}}}$ into (137), we obtain the operator $\mathcal{A}^S$. This completes the proof of the proposition. □

The special form of our martingale estimating functions allows us to determine a fourth order expansion if condition (49) or condition (63), respectively, holds. For the sake of simplicity it will first be derived for the case, where the martingale estimating function g depends on the variables h, y and is independent of the minimum variable l.

Proposition 6.2.6. *Let Y denote the process defined by the stochastic differential equation* (16)*. We assume that $\mathcal{G}_\theta$ is a class of flows that satisfies Assumption 2.2.2. Then for any $g \in \mathcal{G}_\theta$, $(h,l,y) \mapsto g_{\Delta,\theta}(x,h,l,y)$ that is independent of l for all $(\Delta,\theta,x) \in (\mathbb{R}_+,\Theta,\mathbb{R})$ and that satisfies the additional assumption* (49)*, the following expansion holds:*

$$\mathbb{E}_{x,\theta}\Big[g^2_{\Delta,\theta}\big(Y_0,H^Y_\Delta,Y_\Delta\big)\Big] = \Delta^2\mathcal{A}^S_2 g_{0,\theta}(\mathbf{x}) + O\big(\Delta^{5/2}\big), \tag{145}$$

where the operator $\mathcal{A}^S_2$ is defined via

$$\begin{aligned}\mathcal{A}^S_2 g_{0,\theta}(\mathbf{x}) = \sigma(x;\theta)^4 \ \Bigg\{&\left(\frac{1}{2}-\frac{4}{\pi^2}\right)\Big(\mathbf{D}_{hh}\, g_{0,\theta}(\mathbf{x})\Big)^2 + \frac{1}{2}\Big(\mathbf{D}_{yy}\, g_{0,\theta}(\mathbf{x})\Big)^2 \\ &+ \mathbf{D}_{hy}\, g_{0,\theta}(\mathbf{x})\mathbf{D}_{yy}\, g_{0,\theta}(\mathbf{x}) \ + \ \left(\frac{7}{4}-\frac{4}{\pi}\right)\mathbf{D}_{hy}\, g_{0,\theta}(\mathbf{x})\mathbf{D}_{hh}\, g_{0,\theta}(\mathbf{x}) \\ &+ \left(\frac{7}{4}-\frac{10}{3\pi}\right)\Big(\mathbf{D}_{hy}\, g_{0,\theta}(\mathbf{x})\Big)^2 \ + \ \left(\frac{1}{2}-\frac{2}{3\pi}\right)\mathbf{D}_{hh}\, g_{0,\theta}(\mathbf{x})\mathbf{D}_{yy}\, g_{0,\theta}(\mathbf{x}) \ \Bigg\}.\end{aligned} \tag{146}$$

Proof. For convenience, let us assume that the martingale estimating function depends on h alone. This means, we assume that $g_{\Delta,\theta}$ has the following form

$$g_{\Delta,\theta}(x,h,y) = a(\Delta,x;\theta)\left(\kappa\Big(h - F^H(\Delta,x;\theta)\Big) - \mathbb{E}_{x,\theta}\Big[\kappa\Big(H^Y_\Delta - F^H(\Delta,x;\theta)\Big)\Big]\right). \tag{147}$$

In order to satisfy Assumption 2.2.2 the function κ must be three times continuously differentiable and κ''' must have polynomial growth near infinity. Condition (49) is equivalent to $\kappa'(0) = 0$. We expand the expression $\kappa\Big(h - F^H(\Delta,x;\theta)\Big)$ around 0 in order to obtain

$$\kappa\Big(h - F^H(\Delta,x;\theta)\Big) = \kappa(0) + \frac{\kappa''(0)}{2}\Big(h - F^H(\Delta,x;\theta)\Big)^2 + \ \frac{k'''(\xi)}{6}\Big(h - F^H(\Delta,x;\theta)\Big)^3, \tag{148}$$

where ξ is between 0 and $h - F^H(\Delta,x;\theta)$. Due to the assumption that κ''' has polynomial growth, we find

$$\mathbb{E}_{x,\theta}\left[\kappa\Big(H^Y_\Delta - F^H(\Delta,x;\theta)\Big)\right] = \kappa(0) + \frac{1}{2}\kappa''(0)\mathbb{E}_{x,\theta}\left[\big\{H^Y_\Delta - F^H(\Delta,x;\theta)\big\}^2\right] + O(\Delta^{3/2}). \tag{149}$$

By the same argument, an expansion of $g^2_{\Delta,\theta}$ gives the estimate

$$\begin{aligned}&\mathbb{E}_{x,\theta}\Big[g^2_{\Delta,\theta}\big(Y_0,H^Y_\Delta,Y_\Delta\big)\Big] \\ &= \frac{\kappa''(0)}{4}a(\Delta,x;\theta)^2\mathbb{E}_{x,\theta}\left[\left\{\Big(H^Y_\Delta - F^H(\Delta,x;\theta)\Big)^2 - \mathbb{E}_{x,\theta}\Big[\Big(H^Y_\Delta - F^H(\Delta,x;\theta)\Big)^2\Big]\right\}^2\right] \\ &\quad + O(\Delta^{5/2}).\end{aligned} \tag{150}$$

Formula (50), which is concisely proved in Lemma 5.2.1.4 in [8], in combination with the scaling property of Brownian motion shows that

$$\mathbb{E}_{x,\theta}\left[\left\{\left(H^Y_\Delta - F^H(\Delta,x;\theta)\right)^2 - \mathbb{E}_{x,\theta}\left[\left(H_\Delta - F^H(\Delta,x;\theta)\right)^2\right]\right\}^2\right]$$
$$= \Delta^2\sigma(x;\theta)^4\mathbb{E}\Big[\left((H^B_1 - \mathbb{E}H^B_1)^2 - \mathrm{Var}(H^B_1)\right)^2\Big] + O(\Delta^{5/2}). \tag{151}$$

Here, for standard one-dimensional Brownian motion $(B_s,\ 0 \le s \le 1)$, H^B_1 denotes the random variable $H^B_1 = \sup_{0\le s\le 1} B_s$. Obviously,

$$\frac{\partial^2}{\partial h^2} g_{0,\theta}(x,h,x)\Big|_{h=x} = \mathbf{D}_{hh} g_{0,\theta}(\mathbf{x}) = a(\Delta,x;\theta)^2\kappa''(0), \tag{152}$$

and consequently it remains to state that

$$\mathbb{E}\big[((H^B_1 - \mathbb{E}H^B_1)^2 - \mathrm{Var}(H^B_1))^2\big] = \mathbb{E}\big[(H^B_1 - \mathbb{E}H^B_1)^4\big] - \mathrm{Var}(H^B_1)^2 = 2 - \frac{16}{\pi^2}. \tag{153}$$

Altogether, for the function $g_{\Delta,\theta}$ defined by formula (147), we have proved that

$$\mathbb{E}_{x,\theta}\Big[g^2_{\Delta,\theta}(Y_0, H^Y_\Delta, Y_\Delta)\Big] = \sigma(x;\theta)^4\ \left(\frac{1}{2} - \frac{4}{\pi^2}\right)\left(\mathbf{D}_{hh} g_{0,\theta}(\mathbf{x})\right)^2 + O(\Delta^{5/2}). \tag{154}$$

More general functions that depend on both variables h and y, are treated in the same way. One just has to consider all possible partial derivatives of $g_{\Delta,\theta}(x,h,y)$ with respect to h and y separately. The basic ideas behind the proof remain the same as above. We omit the details here. A list of the remaining relevant moments of (H^B_1, B_1) can be found in formulae (10.60)-(10.64) in [8]. These moments are easily calculated by means of the joint density of (H^B_1, B_1), which is displayed in formula (106). Altogether, this proves the result. □

For a class of flows that depends on all variables (h,l,y) a result similar to the one of Proposition 6.2.6 can be found, provided condition (63) is satisfied. We can state the following proposition.

Proposition 6.2.7. *Let Y denote the process defined by the stochastic differential equation* (16)*. We assume that $\mathcal{G}_\theta$ is a class of flows that satisfies Assumption 2.2.2. Then for any $g \in \mathcal{G}_\theta$, that satisfies the additional assumption* (63)*, the following expansion holds:*

$$\mathbb{E}_{x,\theta}\Big[g^2_{\Delta,\theta}(Y_0, H^Y_\Delta, L^Y_\Delta, Y_\Delta)\Big] = \Delta^2\mathcal{A}^{S,a}_2 g_{0,\theta}(\mathbf{x}) + O(\Delta^{5/2}), \tag{155}$$

where the operator $\mathcal{A}^{S,a}_2$ is defined via

$$\mathcal{A}^{S,a}_2 g_{0,\theta}(\mathbf{x}) = Z^a_{qua}(x)\cdot\left(S^a_{qua}\right)^{-1}(x)\cdot\left(Z^a_{qua}(x)\right)^T, \tag{156}$$

with

$$
\begin{aligned}
&Z^a_{qua}(x)\\
&= \big(Z^a_{qua,1}(x), Z^a_{qua,2}(x), Z^a_{qua,3}(x), Z^a_{qua,4}(x), Z^a_{qua,5}(x), Z^a_{qua,6}(x)\big)\\
&= \Big(\mathbf{D}_{hh}\, g_{0,\theta}(\mathbf{x}), \mathbf{D}_{yy}\, g_{0,\theta}(\mathbf{x}), \mathbf{D}_{hy}\, g_{0,\theta}(\mathbf{x}), \mathbf{D}_{hl}\, g_{0,\theta}(\mathbf{x}), \mathbf{D}_{ll}\, g_{0,\theta}(\mathbf{x}), \mathbf{D}_{ly}\, g_{0,\theta}(\mathbf{x})\Big). \qquad (157)
\end{aligned}
$$

Moreover, for $x \in \mathbb{R}$*, the matrix* $\big(S^a_{qua}\big)^{-1}(x)$ *is defined by*

$$
\big(S^a_{qua}\big)^{-1}(x) = \sigma^4(x;\theta) \cdot A \cdot \mathrm{Cov}[S^B_{H,L,X}] \cdot A. \qquad (158)
$$

Here, $A = \mathrm{diag}(1/2, 1/2, 1/2, 1, 1, 1)$ *and the entries of the column-vector* $S^B_{H,L,X}$ *are given by*

$$
\begin{aligned}
S^B_{H,L,X;1} &= \big(H^B_1 - \mathbb{E}H^B_1\big)^2 - \mathbb{E}\Big[\big(H^B_1 - \mathbb{E}H^B_1\big)^2\Big],\\
S^B_{H,L,X;2} &= B_1 - \mathbb{E}\Big[B_1^2\Big],\\
S^B_{H,L,X;3} &= \big(H^B_1 - \mathbb{E}H^B_1\big) \cdot B_1 - \mathbb{E}\Big[\big(H^B_1 - \mathbb{E}H^B_1\big) \cdot B_1\Big],\\
S^B_{H,L,X;4} &= \big(H^B_1 - \mathbb{E}H^B_1\big)\big(L^B_1 - \mathbb{E}L^B_1\big) - \mathbb{E}\Big[\big(H^B_1 - \mathbb{E}H^B_1\big)\big(L^B_1 - \mathbb{E}L^B_1\big)\Big],\\
S^B_{H,L,X;5} &= \big(L^B_1 - \mathbb{E}L^B_1\big)^2 - \mathbb{E}\Big[\big(L^B_1 - \mathbb{E}L^B_1\big)^2\Big],\\
S^B_{H,L,X;6} &= \big(L^B_1 - \mathbb{E}L^B_1\big) \cdot B_1 - \mathbb{E}\Big[\big(L^B_1 - \mathbb{E}L^B_1\big) \cdot B_1\Big], \qquad (159)
\end{aligned}
$$

with $(B_s,\, 0 \leq s \leq 1)$ *being standard Brownian motion of* $\mathbb{R}$ *and* $H^B_1 = \sup_{0\leq s\leq 1} B_s$*,* $L^B_1 = \inf_{0\leq s\leq 1} B_s$*.*

Proof. The proof follows the same procedure as the proof of Proposition 6.2.6. □

References

[1] B.M. Bibby, M. Jacobsen, and M. Sørensen. Estimating functions for discretely sampled diffusion-type models. *Preprint*, 2004. URL `http://www.math.ku.dk/~michael`.

[2] A. Borodin and P. Salminen. *Handbook of Brownian Motion - Facts and Formulae.* Birkhäuser, 2nd edition, 2002.

[3] M. Dominé. First passage time distribution of a Wiener process with drift concerning two elastic barriers. *Journal of Applied Probability*, 33:164–175, 1996.

[4] D. Florens-Zmirou. Approximate discrete time schemes for statistics of diffusion processes. *Statistics*, 20:547–557, 1989.

[5] M.B. Garman and M.J. Klass. On the estimation of security price volatility from historical data. *Journal of Business*, 53(1):67–78, 1980.

[6] V.P. Godambe and C.C. Heyde. Quasi likelihood and optimal estimation. *International Statistical Review*, 55:231–244, 1987.

[7] L.P. Hansen and J.A. Scheinkmann. Back to the future: generating moment implications for continuous-time markov-processes. *Econometrica*, 63:767–804, 1995.

[8] H. Henkel. *Range-based parameter estimation in diffusion models.* PhD Thesis, submitted at Humboldt-Universität zu Berlin, 2010.

[9] M. Jacobsen. Discretely observed diffusions: Classes of estimating functions and small-Δ-optimality. *Scandinavian Journal of Statistics*, 28:123–149, 2001.

[10] M. Jacobsen. Optimality and small-Δ-optimality of martingale estimating functions. *Bernoulli*, 8:643–668, 2002.

[11] M. Jacobsen. Small-Δ-optimal martingale estimating functions for discretely observed diffusions: a simulation study. *Preprint*, 2010.

[12] M. Kessler. Estimation of an ergodic diffusion from discrete observations. *Scand. J. Statist.*, 24:211–229, 1997.

[13] M. Kessler. Simple and explicit estimating functions for a discretely observed diffusion. *Scand. J. Statist.*, 27:65–82, 2000.

[14] Malik Magdon-Ismail and Amir F. Atiya. Volatility estimation using high, low and close data - a maximum likelihood approach. In *Computational Finance (CF2000), Proceedings*, 2000.

[15] M. Parkinson. The extreme value method for estimating the variance of the rate of return. *Journal of Business*, 53(1):61–65, January 1980. URL `http://ideas.repec.org/a/ucp/jnlbus/v53y1980i1p61-65.html`.

[16] L.C.G. Rogers and S.E. Satchell. Estimating variance form high, low and closing prices. *The Annals of Applied Probability*, 4:504–512, 1991.

[17] L.C.G. Rogers and L. Shepp. The correlation of the maxima of correlated Brownian motions. *Journal of Applied Probability*, 43:880–883, 2006.

[18] M. Sørensen. Estimating functions for discretely sampled diffusion-type models: A review. *IMS Lecture Notes - Monograph Series*, 32:305–325, 1997.

[19] M. Sørensen. On asymptotics of estimating functions. *Brazilian Journal of Probability and Statistics*, 13:111–136, 1999.

[20] A.L. Sweet and J.C. Hardin. Solutions for some diffusion processes with two barriers. *Journal of Applied Probability*, 7:423–431, 1970.

[21] N. Yoshida. Estimation for diffusion processes from discrete observations. *Jour. Mult. Anal.*, 41:220–242, 1990.