Diffusion bridges and affine processes

Habilitációs cikkgyűjtemény

Dr. Barczy Mátyás

Debreceni Egyetem, Informatikai Kar
Alkalmazott Matematika és Valószínűségszámítás Tanszék

Debreceni Egyetem
Természettudományi Kar

Debrecen
2015

Contents

Introduction (in Hungarian)

A habilitációs cikkgyűjtemény a szerző nyolc dolgozatát tartalmazza, melyek PhD disszertációjának megvédése (2006) után készültek. Témájuk szerint e munkák két nagy csoportba oszthatóak: diffúziós folyamatokból származtatott hidak konstrukciója, tulajdonságai, speciális hidak vizsgálata; illetve affin folyamatok stacionaritása, ergodicitása és paraméterbecslése. A fentieknek megfelelő struktúrában, két rész, négy-négy fejezetében, az eredmények angol nyelven, bizonyítások nélkül kerülnek bemutatásra. A cikkeket magukat (ahol a bizonyítások is megtalálhatóak) mellékletként csatoljuk.

Az első részbeli első fejezet a Barczy és Kern [**18**] cikk összefoglalása: többdimenziós idő-inhomogén lineáris diffúziós folyamatokból származtatunk hídfolyamatokat csak az alapul vett folyamat átmenetvalószínűségeit felhasználva, megadva a származtatott hídfolyamatok ún. integrál- és anticipatív reprezentációját is. Megmutatjuk, hogy az integrálreprezentáció egy alkalmas sztochasztikus differenciálegyenlet erős megoldása is egyben, belátjuk továbbá, hogy a származtatott hídfolyamatok végesdimenziós eloszlásai tekinthetők alkalmas feltételes eloszlásoknak is. Eredményeinket külön megfogalmazzuk az egydimenziós esetben, speciálisan egydimenziós Ornstein-Uhlenbeck hidakkal is foglalkozunk.

Az első részbeli második fejezet a Barczy és Pap [**28**] cikk összefoglalása: egy sztochasztikus differenciálegyenlet egyértelmű erős megoldásaként adott idő-inhomogén diffúziós folyamat esetén, a drift- és diffúziós együtthatókra vonatkozó bizonyos feltételek mellett, a folyamat bizonyos funkcionáljainak együttes Laplace transzformáltjára származtatunk explicit képletet. Vizsgálatainkat az motiválja, hogy a szóbanforgó diffúziós folyamat drift együtthatójában szereplő paraméter maximum likelihood becslésében a szóbanforgó funkcionálok szerepelnek. A fenti eredmények alkalmazásaként megmutatjuk, hogy a szóbanforgó maximum likelihood becslés aszimptotikusan normális. Az ún. α-Wiener hidak (skálázott Wiener hidak) esetén specializáljuk mind a Laplace transzformáltra, mind a paraméterbecslésre vonatkozó eredményeinket. Megjegyezzük, hogy $\alpha = 1$ esetén egy α-Wiener híd nem más, mint a szokásos Wiener híd, az $\alpha = 0$ esetben pedig egy standard Wiener folyamat.

Az első részbeli harmadik fejezet a Barczy és Iglói [**14**] cikk összefoglalása: az α-Wiener hidak súlyozott és súlyozatlan Karhunen–Loève sorfejtésével foglalkozunk. Alkalmazásként megadjuk ezen hidak L^2-normanégyzetének Laplace transzformáltját és eloszlásfüggvényét, vizsgálva ez utóbbi aszimptotikus viselkedését is (nagy- és kiseltérések).

Az első részbeli negyedik fejezet a Barczy et al. [**21**] cikk összefoglalása: ún. operátor-skálázott Wiener hidakat vezetünk be, egy mátrix skálázási faktorral módosítva egy többdimenziós Wiener híd differenciálegyenletének drift együtthatóját. A skálázó mátrix sajátértékeinek segítségével egy elegendő feltételt származtatunk arra vonatkozóan, hogy a fentiek szerint módosított sztochasztikus differenciálegyenlet erős megoldása valóban hídfolyamat legyen. A hídfolyamat aszimptotikus viselkedésével is foglalkozunk, és röviden tárgyaljuk azt is, hogy a skálázási mátrix egyértelműen meghatározza-e a hídfolyamat eloszlását.

Nem kerültek be a cikkgyűjteménybe, de az első rész témájához kapcsolódnak a Barczy és Pap [**25**], [**27**], és a Barczy és Kern [**17**], [**19**], [**20**] cikkek. A Barczy és Pap [**25**] cikk hídfolyamatok általános konstrukciójáról, a Barczy és Pap [**27**] cikk az egydimenziós α-Wiener hidak alaptulajdonságairól, többek között pályatulajdonságairól szól. A cikkgyűjteményben szereplő Barczy et al. [**21**] cikk a Barczy és Pap [**27**] cikk általánosításának tekinthető. A Barczy és Kern [**17**] cikk az α-Wiener hidak azon általánosításával foglalkozik, mikor is a konstans α paramétert egy alkalmas függvénnyel helyettesítjük.

A második részbeli ötödik fejezet a Barczy et al. [**11**] cikk összefoglalása: affin folyamatokra vonatkozó skálázási tételekkel, illetve egy kritikus kétfaktoros affin folyamat paraméterei legkisebb négyzetes- és feltételes legkisebb négyzetes becslésének aszimptotikájával foglalkozunk.

A második részbeli hatodik fejezet a Barczy et al. [**12**] cikk összefoglalása: egy kétfaktoros szubkritikus affin folyamat esetén vizsgáljuk az egyértelmű stacionárius eloszlás létezését és az ergodicitás témakörét. A tekintett affin folyamat első koordinátája az ún. α-gyök folyamat, ahol $\alpha \in (1, 2]$. Megmutatjuk, hogy tetszőleges $\alpha \in (1, 2]$ esetén egyértelműen létezik stacionárius eloszlás, továbbá, az $\alpha = 2$ esetben az affin folyamat ergodicitását is bizonyítjuk.

A második részbeli hetedik fejezet a Barczy et al. [**13**] cikk összefoglalása: a Barczy et al. [**12**] cikkben vizsgált kétfaktoros szubkritikus diffúziós ($\alpha = 2$) affin folyamat esetén megvizsgáljuk a drift együtthatóban szereplő paraméterek folytonos idejű mintára támaszkodó maximum likelihood-, ill. feltételes legkisebb négyzetes becslésének aszimptotikus viselkedését. A szóbanforgó becslések erős konzisztenciáját és aszimptotikus normalitását bizonyítjuk.

A második részbeli nyolcadik fejezet a Barczy és Pap [**30**] cikk összefoglalása: a pénzügyi matematikában sokat használt Heston folyamat esetén a log-ár folyamatra vonatkozó folytonos idejű megfigyelés alapján viszgáljuk a modell drift együtthatójában szereplő paraméterek maximum likelihood becslésének aszimptotikus viselkedését. Három esetet különböztetünk meg: szubkritikus (ergodikus), kritikus és szuperkritikus. Megmutatjuk, hogy a szóbanforgó maximum likelihood becslés aszimptotikusan normális a szubkritikus esetben, ellentétben a kritikus és szuperkritikus esetekkel, ez utóbbi két esetben is megadva a határeloszlást.

A cikkgyűjteményben nem szereplő cikkek közül a második rész témájához kapcsolódik a Barczy et al. [**31**] cikk, melyben a második részbeli nyolcadik fejezetben vizsgált Heston folyamat esetén a drift együttható paramétereinek diszkrét idejű mintára vonatkozó feltételes legkisebb négyzetes becslésének aszimptotikus tulajdonságaival foglalkozunk a szubkritikus esetben. Tágabb értelemben, a folytonos idejű, folytonos állapotterű bevándorlásos elágazó folyamatokkal, mint "egy-faktoros" affin folyamatokkal foglalkozó Barczy et al. [**22**], [**23**], [**29**] és [**24**] cikkek is kapcsolódnak a második rész témájához.

A habilitációs cikkgyűjteményben szereplő cikkek, illetve a fentiekben említett egyéb cikkek egyike sem kapcsolódik a szerző PhD értekezéséhez, ahhoz képest új irányoknak tekinthetők.

Part 1

Diffusion bridges

Introduction and summary

This part is based on the articles Barczy and Kern [**18**], Barczy and Pap [**28**], Barczy and Iglói [**14**] and Barczy et al. [**21**].

In Barczy and Kern [**18**], we derive bridges from general multidimensional linear non time-homogeneous processes using only the transition densities of the original process giving their integral representations (in terms of a standard Wiener process) and so-called anticipative representations. We derive a stochastic differential equation satisfied by the integral representation and we prove a usual conditioning property for general multidimensional linear process bridges. We specialize our results for the one-dimensional case; especially, we study one-dimensional Ornstein-Uhlenbeck bridges.

In Barczy and Pap [**28**], we consider a time inhomogeneous diffusion process given by a pathwise unique strong solution of a stochastic differential equation, and assuming some conditions between the drift and diffusion coefficients, we derive an explicit formula for the joint Laplace transform of some functionals of the process in question. Our motivation for investigating these functionals is that the maximum likelihood estimator of a parameter in the drift part of the diffusion process can be expressed in terms of these functionals. As an application, we prove asymptotic normality of the maximum likelihood estimator in question. To give an example, we study so-called α-Wiener bridges (scaled Wiener bridges) and maximum likelihood estimation of the parameter α. Note that in case of $\alpha = 1$, this process is the usual Wiener bridge, while in case of $\alpha = 0$ it is a standard Wiener process.

In Barczy and Iglói [**14**], we study weighted and unweighted Karhunen–Loève expansions of an α-Wiener bridge. As applications, we calculate the Laplace transform and the distribution function of the L^2-norm square of an α-Wiener bridge studying also its asymptotic behavior (large and small deviations).

In Barczy et al. [**21**], we introduce operator scaled Wiener bridges by incorporating a matrix scaling in the drift part of the stochastic differential equation of a multidimensional Wiener bridge. A sufficient condition for the bridge property of the solution of this stochastic differential equation is derived in terms of the eigenvalues of the scaling matrix. We analyze the asymptotic behavior of the bridges and briefly discuss the question whether the scaling matrix determines uniquely the law of the corresponding bridge.

CHAPTER 1

Representations of multidimensional linear process bridges

Co-author: Peter Kern

1.1. Introduction

In this paper we deal with deriving bridges from general multidimensional linear processes giving their integral representations (in terms of a standard Wiener process) and their so-called anticipative representations. Our results are also specialized for the one-dimensional case. A bridge process is a stochastic process that is pinned to some fixed point at a future time point. Important examples are provided by Wiener bridges, Bessel bridges and general Markovian bridges, which have been extensively studied and find numerous applications. See, for example, Karlin and Taylor [**101**, Chapter 15], Fitzsimmons, Pitman and Yor [**71**], Privault and Zambrini [**140**], Delyon and Hu [**54**], Gasbarra, Sottinen and Valkeila [**76**], Goldys and Maslowski [**79**], Chaumont and Uribe Bravo [**44**] and Baudoin and Nguyen-Ngoc [**33**]. Recently, Hoyle, Hughston and Macrina [**87**] studied the so-called Lévy random bridges, that are Lévy processes conditioned to have a prespecified marginal law at the endpoint of the bridge (see also the Ph.D. dissertation of Hoyle [**86**]). Bichard [**37**] considered the so-called bridged Wiener sheets, that are Wiener sheets which are forced to take some values along specified curves. Very recently, Campi et al. [**43**] studied the so-called dynamic Markov bridges, i.e., given a Markovian Brownian martingale Z, they built a process U which is a martingale in its own filtration and satisfies $U_1 = Z_1$.

In what follows first we give a motivation for our multidimensional results by presenting different representations of the one-dimensional Ornstein-Uhlenbeck bridges, and then we briefly summarize the structure of the paper.

Motivation: representations of one-dimensional Ornstein-Uhlenbeck bridges

Let $(B_t)_{t\geqslant 0}$ be a standard Wiener process and for $q \neq 0$, $\sigma \neq 0$ let us consider the stochastic differential equation (SDE)

$$\begin{cases} \mathrm{d}Z_t = q\, Z_t \,\mathrm{d}t + \sigma \,\mathrm{d}B_t, & t \geqslant 0, \\ Z_0 = 0. \end{cases} \tag{1.1.1}$$

It is known that there exists a strong solution of this SDE, namely

$$Z_t = \sigma \int_0^t \mathrm{e}^{q(t-s)} \,\mathrm{d}B_s, \qquad t \geqslant 0, \tag{1.1.2}$$

and strong uniqueness for the SDE (1.1.1) holds. The process $(Z_t)_{t\geqslant 0}$ is called a one-dimensional Ornstein-Uhlenbeck process (OU-process). It is a time-homogeneous Gauss-Markov process with transition densities

$$p_t^Z(x,y) = \frac{1}{\sqrt{2\pi\sigma^2\kappa_q(t)}} \exp\left\{ -\frac{(y-\mathrm{e}^{qt}x)^2}{2\sigma^2\kappa_q(t)} \right\}, \qquad t > 0, \quad x, y \in \mathbb{R}, \tag{1.1.3}$$

where we set

$$\kappa_q(t) := \frac{\mathrm{e}^{2qt}-1}{2q} = \frac{\mathrm{e}^{qt}}{q}\sinh(qt), \qquad t \geqslant 0. \tag{1.1.4}$$

For $a, b \in \mathbb{R}$ and $T > 0$, by an Ornstein-Uhlenbeck bridge from a to b over the time interval $[0,T]$ derived from Z we understand a Markov process $(U_t)_{t\in[0,T]}$ with initial distribution $P(U_0 = a) = 1$, with $P(U_T = b) = 1$ and with transition densities

$$p^U_{s,t}(x,y) = \frac{p^Z_{t-s}(x,y)\,p^Z_{T-t}(y,b)}{p^Z_{T-s}(x,b)}, \quad x,y\in\mathbb{R}, \quad 0\leqslant s<t<T. \tag{1.1.5}$$

We also note that U_t converges almost surely to b as $t \uparrow T$, see, e.g., Fitzsimmons, Pitman and Yor [**71**, Proposition 1]. For the construction of bridges derived from a general time-homogeneous Markov process by using only its transition densities, see, e.g., Barczy and Pap [**25**] and Chaumont and Uribe Bravo [**44**]. Standard calculations yield that for $x, y \in \mathbb{R}$ and $0 \leqslant s < t < T$,

$$\frac{p^Z_{t-s}(x,y)\,p^Z_{T-t}(y,0)}{p^Z_{T-s}(x,0)} = \frac{1}{\sqrt{2\pi\sigma(s,t)}}\exp\left\{-\frac{\left(y-\frac{\sinh(q(T-t))}{\sinh(q(T-s))}\,x\right)^2}{2\sigma(s,t)}\right\}, \tag{1.1.6}$$

which is a Gauss density (as a function of y) with mean $\frac{\sinh(q(T-t))}{\sinh(q(T-s))}\,x$ and variance $\sigma(s,t)$, where for all $0 \leqslant s \leqslant t < T$,

$$\sigma(s,t) := \sigma^2\,\frac{\kappa_q(T-t)\kappa_q(t-s)}{\kappa_q(T-s)} = \frac{\sigma^2}{q}\,\frac{\sinh(q(T-t))\sinh(q(t-s))}{\sinh(q(T-s))}. \tag{1.1.7}$$

Note that if $\sigma = 0$ then for any $q \in \mathbb{R}$ the unique (deterministic) solution of (1.1.1) is $Z_t = 0$ for all $t \geqslant 0$ (which coincides with its own bridge from 0 to 0). On the other hand, if $q = 0$ and $\sigma \neq 0$, the unique strong solution of the SDE (1.1.1) is the Wiener process $Z_t = \sigma B_t$, $t \geqslant 0$, and it is well known that the Wiener bridge $(\widetilde{U}_t)_{t\in[0,T]}$ from 0 to 0 over $[0,T]$ derived from $Z = \sigma B$ admits the (stochastic) integral representation

$$\widetilde{U}_t = \sigma\int_0^t \frac{T-t}{T-s}\,\mathrm{d}B_s, \qquad t\in[0,T), \tag{1.1.8}$$

see, e.g., Section 5.6.B in Karatzas and Shreve [**100**]. Moreover, one can easily verify that $(\widetilde{U}_t)_{t\in[0,T]}$ is a Markov process with transition densities

(1.1.9)

$$p^{\widetilde{U}}_{s,t}(x,y) = \frac{1}{\sqrt{2\pi\widetilde{\sigma}(s,t)}}\exp\left\{-\frac{\left(y-\frac{T-t}{T-s}\,x\right)^2}{2\widetilde{\sigma}(s,t)}\right\}, \quad x,y\in\mathbb{R}, \quad 0\leqslant s<t<T,$$

where $\widetilde{\sigma}(s,t) := \sigma^2\frac{(T-t)(t-s)}{T-s}$ for all $0 \leqslant s < t < T$, and that (1.1.5) is satisfied with $b = 0$, U being replaced by $\widetilde{U}$ and

$$p^Z_t(x,y) = \frac{1}{\sqrt{2\pi t\sigma^2}}\exp\left\{-\frac{(y-x)^2}{2t\sigma^2}\right\}, \quad x,y\in\mathbb{R}, \quad t>0.$$

Comparing (1.1.6) with (1.1.9), it is quite reasonable that an integral representation for the Ornstein-Uhlenbeck bridge from 0 to 0 over $[0,T]$ derived from the process Z given by the SDE (1.1.1) should have the form

$$U_t = \sigma\int_0^t \frac{\sinh(q(T-t))}{\sinh(q(T-s))}\,\mathrm{d}B_s, \qquad t\in[0,T),$$

and in fact this is made precise in the sequel (see Remark 1.3.9). We will further consider general multivariate linear process bridges.

Besides the integral representation (1.1.8) of the Wiener bridge $(\widetilde{U}_t)_{t\in[0,T]}$ from 0 to 0 over $[0,T]$, one can find two equivalent representations in the literature. These are given in Section 5.6.B of Section 5.6.B in Karatzas and Shreve [**100**], namely,

$$\begin{cases} \mathrm{d}\widetilde{U}_t = -\frac{1}{T-t}\widetilde{U}_t\,\mathrm{d}t + \mathrm{d}B_t, \qquad t\in[0,T), \\ \widetilde{U}_0 = 0, \end{cases} \tag{1.1.10}$$

and

$$\widehat{U_t} = B_t - \frac{t}{T}B_T, \qquad t\in[0,T]. \tag{1.1.11}$$

The representation (1.1.8) with $\sigma = 1$ is just a strong solution of the SDE (1.1.10). So, the equations (1.1.8) with $\sigma = 1$ and (1.1.10) define the same process $(\widetilde{U}_t)_{t\in[0,T]}$. However, the equation (1.1.11) does not define the same process as the equations (1.1.8) with $\sigma = 1$ and (1.1.10). The equality between representations (1.1.8) with $\sigma = 1$, (1.1.10) and (1.1.11) is only an equality in law, i.e., they determine the same probability measure on $(C([0,T]),\mathcal{B}(C([0,T])))$, where $C([0,T])$ denotes the set of all real-valued continuous functions on $[0,T]$ and $\mathcal{B}(C([0,T]))$ is the Borel σ-algebra on it. The fact that the processes $\widetilde{U}$ and $\widehat{U}$ are different follows from the fact that the process $\widetilde{U}$ is adapted to the filtration generated by B, while the process $\widehat{U}$ is not. Indeed, to construct $\widehat{U}$ we need the random variable B_T. One can call (1.1.11) a non-adapted, anticipative representation of a Wiener bridge. The attribute anticipative indicates that for the definition of $\widehat{U_t}$ we use the random variable B_T, where the time point T is after the time point t.

A similar anticipative representation of an Ornstein-Uhlenbeck bridge derived from the SDE (1.1.1) can be found on page 378 in Donati-Martin [**60**] and in Lemma 1 in Papież and Sandison [**137**]. Donati-Martin gave an anticipative representation of an Ornstein-Uhlenbeck bridge from $a = 0$ to $b = 0$ derived from the SDE (1.1.1) with $q < 0$ and $\sigma = 1$, while Papież and Sandison formulated their lemma in case of arbitrary starting point a and ending point b, but only for special values of q and σ. Note that the proof in [**137**] also valid for all $q \neq 0$ and $\sigma \neq 0$ (see our Remark 1.3.7).

Moreover, concerning the relationship between a Wiener process and a Wiener bridge, by Problem 5.6.13 in Karatzas and Shreve [**100**], if $T > 0$ is fixed and $(B_t)_{t\geqslant 0}$ is a standard Wiener process (starting from 0), then for all $n\in\mathbb{N}$, $0 < t_1 < \ldots < t_n < T$, the conditional distribution of $(B_{t_1},\ldots,B_{t_n})$ given $B_T = 0$ coincides with the distribution of $(\widetilde{U}_{t_1},\ldots,\widetilde{U}_{t_n})$, where $\widetilde{U}$ is given by (1.1.8) with $\sigma = 1$ or by (1.1.10).

Finally, we note that the transition densities $p^{U}_{s,t}(x,y)$, $x,y\in\mathbb{R}$, $0\leqslant s<t<T$, of the process bridge $(U_t)_{t\in[0,T]}$ can be derived by using Doob's h-transform (see Doob [**61**]). In Section 1.2 we briefly study this approach for general multivariate linear process bridges.

Structure of the paper

In Section 1.2 we derive multidimensional linear process bridges from a multidimensional linear non time-homogeneous process Z given by the SDE (1.2.1) by using only the transition densities of Z, see Theorem 1.2.2 and Definition 1.2.4. We also give an integral and a so-called anticipative representation of the derived bridge, see formulae (1.2.11) and (1.2.14), respectively. We derive an SDE satisfied by this integral representation, see Theorem 1.2.5, and in Proposition 1.2.8 we prove a usual conditioning property for general multidimensional linear process bridges. In Remark 1.2.9 we point out that the integral representation and anticipative representation of the bridge are quite different. To shed more light on the different behavior of the different bridge representations, in a companion paper we examined sample path deviations of the Wiener process and the Ornstein-Uhlenbeck process from its bridges, see Barczy and Kern [**19**]. In Remark 1.2.10 we study that the SDE derived for the integral representation can be considered as a consequence of Proposition 3 in Delyon and Hu [**54**]. We use the expression 'can be considered' since the definition of bridges given in Delyon and Hu [**54**] and in the present paper are different. We have a different approach coming from the possibility that in our special case we are able to explicitly calculate the transition densities of the bridge from which we deduce an integral representation and finally end up with the same SDE of Proposition 3 in Delyon and Hu [**54**] such that this integral representation is a strong solution of the above mentioned SDE. We also note that the SDE of Proposition 3 in Delyon and Hu [**54**] contains the solution of a deterministic differential equation which solution always remains abstract, while in our special case we have an explicit solution via evolution matrices (see Section 1.2). Concerning anticipative representations of process bridges see also Delyon and Hu [**54**, Theorem 2] and the recent paper of Gasbarra, Sottinen and Valkeila [**76**].

In Section 1.3 we formulate our multidimensional results in case of dimension one which includes also the study of usual Ornstein-Uhlenbeck bridges. We note that not all of the results are immediate consequences of the multidimensional ones and in case of dimension one we can give an illuminating explanation for the anticipative representation motivated by Lemma 1 in Papież and Sandison [**137**], see Remark 1.3.7.

The Appendix contains a supplement for our assumption on Kalman type matrices (introduced in Section 1.2), two auxiliary lemmata on matrix identities and a modification of an appropriate strong law of large numbers for continuous square integrable multidimensional martingales needed to prove almost sure continuity of our process bridges at the endpoint T in Section 1.2.

1.2. Multidimensional linear process bridges

Let $\mathbb{N}$, $\mathbb{R}$ and $\mathbb{R}_+$ denote the set of positive integers, real numbers and non-negative real numbers, respectively. For all $n, m \in \mathbb{N}$, let $\mathbb{R}^{n\times m}$ and I_n denote the set of $n \times m$ matrices with real entries and the $n \times n$ identity matrix, respectively.

For all $d, p \in \mathbb{N}$, let us consider a general d-dimensional linear process given by the linear SDE

$$\mathrm{d}\mathbf{Z}_t = \big(Q(t)\mathbf{Z}_t + \mathbf{r}(t)\big)\,\mathrm{d}t + \Sigma(t)\,\mathrm{d}\mathbf{B}_t, \qquad t \geqslant 0, \tag{1.2.1}$$

with continuous functions $Q : \mathbb{R}_+ \to \mathbb{R}^{d\times d}$, $\Sigma : \mathbb{R}_+ \to \mathbb{R}^{d\times p}$ and $\mathbf{r} : \mathbb{R}_+ \to \mathbb{R}^d$, where $(\mathbf{B}_t)_{t\geqslant 0}$ is a p-dimensional standard Wiener process on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, P)$ satisfying the usual conditions (the filtration being constructed by the help of $\mathbf{B}$), i.e., $(\Omega, \mathcal{F}, P)$ is complete, $(\mathcal{F}_t)_{t\geqslant 0}$ is right continuous, $\mathcal{F}_0$ contains all the P-null sets in $\mathcal{F}$ and $\mathcal{F}_\infty = \mathcal{F}$, where $\mathcal{F}_\infty := \sigma\left(\bigcup_{t\geqslant 0}\mathcal{F}_t\right)$, see, e.g., Karatzas and Shreve [**100**, Section 5.2.A]. It is known that there exists a strong solution of the SDE (1.2.1), namely

$$\mathbf{Z}_t = \Phi(t)\left[\mathbf{Z}_0 + \int_0^t \Phi^{-1}(s)\mathbf{r}(s)\,\mathrm{d}s + \int_0^t \Phi^{-1}(s)\Sigma(s)\,\mathrm{d}\mathbf{B}_s\right], \qquad t \geqslant 0, \tag{1.2.2}$$

where $\mathbf{Z}_0$ is independent of the Wiener process $(\mathbf{B}_t)_{t\geqslant 0}$, Φ is a solution to the deterministic matrix differential equation $\Phi'(t) = Q(t)\Phi(t)$, $t \geqslant 0$, with $\Phi(0) = I_d$, and strong uniqueness for the SDE (1.2.1) holds, see, e.g., Karatzas and Shreve [**100**, Section 5.6]. The unique solution of the above matrix differential equation can be given as $\Phi(t) = E(t,0)$, $t \geqslant 0$, in terms of the evolution matrices (also known as state transition matrices)

$$E(t,s) = I_d + \int_s^t Q(t_1)\,\mathrm{d}t_1 + \sum_{k=2}^{\infty} \int_s^t \int_s^{t_1} \cdots \int_s^{t_{k-1}} Q(t_1)\cdots Q(t_k)\,\mathrm{d}t_k \mathrm{d}t_{k-1}\cdots \mathrm{d}t_1$$

for $s,t \geqslant 0$. Indeed, by Theorem 1.8.2 in Conti [**47**], the general d-dimensional solution $\mathbf{y}(t) : \mathbb{R}_+ \to \mathbb{R}^d$ of $\mathbf{y}'(t) = Q(t)\mathbf{y}(t)$, $t \geqslant 0$, is represented by $\mathbf{y}(t) = E(t,s)\mathbf{y}(s)$ for all $s,t \geqslant 0$, which shows that $\Phi(t) = E(t,0)$, $t \geqslant 0$. Note that, since Q is continuous, there exists an $L > 0$ such that $\|Q(u)\| \leqslant L$ for all $u \in [\min(s,t), \max(s,t)]$, $s,t \geqslant 0$ (with some fixed matrix norm $\|.\|$ on $\mathbb{R}^{d\times d}$), and hence one can easily verify that $\|E(t,s)\| \leqslant \mathrm{e}^{L|t-s|}$. Note also that if $Q(t) = Q \in \mathbb{R}^{d\times d}$, $t \geqslant 0$, is constant then $E(t,s) = \mathrm{e}^{(t-s)Q}$ for $t,s \geqslant 0$, and hence $\Phi(t) = \mathrm{e}^{tQ}$, $t \geqslant 0$.

We will make frequent use of the following properties of evolution matrices stated as equations (1.9.2) and (1.9.3) in Conti [**47**]. For all $r,s,t \geqslant 0$ we have

$$E(t,s)E(s,r) = E(t,r), \tag{1.2.3}$$

$$E(t,t) = I_d, \qquad E(t,s)^{-1} = E(s,t), \tag{1.2.4}$$

$$\partial_1 E(t,s) = Q(t)E(t,s), \qquad \partial_2 E(t,s) = -E(t,s)Q(s). \tag{1.2.5}$$

The unique strong solution of the SDE (1.2.1) can now be written as

$$\mathbf{Z}_t = E(t,0)\mathbf{Z}_0 + \int_0^t E(t,s)\mathbf{r}(s)\,\mathrm{d}s + \int_0^t E(t,s)\Sigma(s)\,\mathrm{d}\mathbf{B}_s, \qquad t \geqslant 0.$$

Here and in what follows we assume that $\mathbf{Z}_0$ has a Gauss distribution independent of the Wiener process $(\mathbf{B}_t)_{t\geqslant 0}$. Then we may define the filtration $(\mathcal{F}_t)_{t\geqslant 0}$ such that $\sigma\{\mathbf{Z}_0, \mathbf{B}_s : 0 \leqslant s \leqslant t\} \subset \mathcal{F}_t$ for all $t \geqslant 0$, see, e.g., Karatzas and Shreve [**100**, Section 5.2.A].

We will call the process $(\mathbf{Z}_t)_{t\geqslant 0}$ a d-dimensional linear process.

One can easily derive that for $0 \leqslant s \leqslant t$ we have

$$\mathbf{Z}_t = E(t,s)\mathbf{Z}_s + \int_s^t E(t,u)\mathbf{r}(u)\,\mathrm{d}u + \int_s^t E(t,u)\Sigma(u)\,\mathrm{d}\mathbf{B}_u. \tag{1.2.6}$$

Hence, given $\mathbf{Z}_s = \mathbf{x}$, the distribution of $\mathbf{Z}_t$ does not depend on $(\mathbf{Z}_u)_{u\in[0,s)}$ and thus $(\mathbf{Z}_t)_{t\geqslant 0}$ is a Gauss-Markov process (see, e.g., Karatzas and Shreve [**100**, Problem 5.6.2]). For any $0 \leqslant s \leqslant t$ and $\mathbf{x} \in \mathbb{R}^d$ let us define

$$\mathbf{m}^+_{\mathbf{x}}(s,t) := \mathbf{x} + \int_s^t E(s,u)\mathbf{r}(u)\,\mathrm{d}u \quad \text{and} \quad \mathbf{m}^-_{\mathbf{x}}(s,t) := \mathbf{x} - \int_s^t E(t,u)\mathbf{r}(u)\,\mathrm{d}u.$$

Then for any $\mathbf{x} \in \mathbb{R}^d$ and $0 \leqslant s < t$ the conditional distribution of $\mathbf{Z}_t$ given $\mathbf{Z}_s = \mathbf{x}$ is Gauss with mean

$$\mathbf{m}_{\mathbf{x}}(s,t) := E(t,s)\mathbf{m}^+_{\mathbf{x}}(s,t) = E(t,s)\mathbf{x} + \int_s^t E(t,u)\mathbf{r}(u)\mathrm{d}u,$$

and with covariance matrix of Kalman type (see Kalman [**99**])

$$\kappa(s,t) := \int_s^t E(t,u)\Sigma(u)\Sigma(u)^\top E(t,u)^\top\,\mathrm{d}u.$$

The matrices $\kappa(s,t)$ are symmetric and positive semi-definite for all $0 \leqslant s < t$, and in what follows we put the following assumption:

$$\kappa(s,t) \text{ is positive definite for all } 0 \leqslant s < t. \tag{1.2.7}$$

From control theory of linear systems we owe sufficient conditions for positive definiteness of the Kalman matrices (see, e.g., Theorems 7.7.1 - 7.7.3 in Conti [**47**]) which we present in the Appendix, see Proposition 1.4.1.

Hence the transition densities of the Gauss-Markov process $(\mathbf{Z}_t)_{t\geqslant 0}$ read as

(1.2.8)
$$p^{\mathbf{Z}}_{s,t}(\mathbf{x},\mathbf{y}) = \frac{1}{\sqrt{(2\pi)^d \det\kappa(s,t)}} \exp\left\{-\frac{1}{2}\langle \kappa(s,t)^{-1}(\mathbf{y}-\mathbf{m}_{\mathbf{x}}(s,t)), \mathbf{y}-\mathbf{m}_{\mathbf{x}}(s,t)\rangle\right\}$$

for all $0 \leqslant s < t$ and $\mathbf{x},\mathbf{y}\in\mathbb{R}^d$. Our aim is to derive a process bridge from $\mathbf{Z}$, namely, we will consider a bridge from $\mathbf{a}$ to $\mathbf{b}$ over the time interval $[0,T]$, where $\mathbf{a},\mathbf{b}\in\mathbb{R}^d$ and $T>0$. Generalizing the formula (2.7) in Fitzsimmons, Pitman and Yor [**71**] to multidimensional non time-homogeneous Markov processes, for fixed $T>0$ we are looking for a Markov process $(\mathbf{U}_t)_{t\in[0,T]}$ with initial distribution $P(\mathbf{U}_0=\mathbf{a})=1$ and with transition densities

$$p^{\mathbf{U}}_{s,t}(\mathbf{x},\mathbf{y}) = \frac{p^{\mathbf{Z}}_{s,t}(\mathbf{x},\mathbf{y})\, p^{\mathbf{Z}}_{t,T}(\mathbf{y},\mathbf{b})}{p^{\mathbf{Z}}_{s,T}(\mathbf{x},\mathbf{b})}, \quad \mathbf{x},\mathbf{y}\in\mathbb{R}^d, \quad 0\leqslant s<t<T, \tag{1.2.9}$$

provided that such a process exists. To properly speak of $(\mathbf{U}_t)_{t\in[0,T]}$ as a process bridge, we shall study the limit behavior of $\mathbf{U}_t$ as $t\uparrow T$, namely, we shall show that $\mathbf{U}_t\to\mathbf{b}=:\mathbf{U}_T$ almost surely and also in L^2 as $t\uparrow T$ (see Theorem 1.2.2).

Our approach can also be viewed in the context of Doob's h-transform (see Doob [**61**]) as follows. For bounded Borel-measurable functions $f:\mathbb{R}_+\times\mathbb{R}^d\to\mathbb{R}$ one can define a family of operators $(P_{s,t})_{0\leqslant s<t}$ by

$$P_{s,t}f(s,\mathbf{x}) := \int_{\mathbb{R}^d} f(t,\mathbf{y})\, p^{\mathbf{Z}}_{s,t}(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{y}$$

for $0\leqslant s<t$ and $\mathbf{x}\in\mathbb{R}^d$. Then

$$\begin{aligned} |P_{s,t}f(s,\mathbf{x})| &\leqslant \int_{\mathbb{R}^d} |f(t,\mathbf{y})|\, p^{\mathbf{Z}}_{s,t}(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{y} \leqslant \sup_{\mathbf{y}\in\mathbb{R}^d}|f(t,\mathbf{y})|<\infty,\\ P_{s,t}f(s,\mathbf{Z}_s) &= \mathbb{E}(f(t,\mathbf{Z}_t)\,|\,\mathbf{Z}_s) \qquad P\text{-a.s.,}\end{aligned}$$

and the family $(P_{s,t})_{0\leqslant s<t}$ forms a hemigroup of transition operators for the Markov process $\mathbf{Z}$. Indeed, for $0\leqslant s<r<t$ and $\mathbf{x}\in\mathbb{R}^d$ we observe

$$\begin{aligned} P_{s,r}P_{r,t}f(s,\mathbf{x}) &= \int_{\mathbb{R}^d} P_{r,t}f(r,\mathbf{y})\, p^{\mathbf{Z}}_{s,r}(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{y}\\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(t,\mathbf{z})\, p^{\mathbf{Z}}_{r,t}(\mathbf{y},\mathbf{z})\,\mathrm{d}\mathbf{z}\, p^{\mathbf{Z}}_{s,r}(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{y}\\ &= \int_{\mathbb{R}^d} f(t,\mathbf{z}) \int_{\mathbb{R}^d} p^{\mathbf{Z}}_{s,r}(\mathbf{x},\mathbf{y})\, p^{\mathbf{Z}}_{r,t}(\mathbf{y},\mathbf{z})\,\mathrm{d}\mathbf{y}\,\mathrm{d}\mathbf{z}\\ &= \int_{\mathbb{R}^d} f(t,\mathbf{z})\, p^{\mathbf{Z}}_{s,t}(\mathbf{x},\mathbf{z})\,\mathrm{d}\mathbf{z} = P_{s,t}f(s,\mathbf{x}).\end{aligned}$$

For fixed $T>0$ and $\mathbf{b}\in\mathbb{R}^d$ we now define the function

$$h:[0,T)\times\mathbb{R}^d\to\mathbb{R}_+ \quad \text{by} \quad h(t,\mathbf{x}) = p^{\mathbf{Z}}_{t,T}(\mathbf{x},\mathbf{b}), \qquad t\in[0,T),\ \ \mathbf{x}\in\mathbb{R}^d.$$

By (1.2.8), h is positive and bounded on $[0,t]\times\mathbb{R}^d$ for every $0<t<T$. Indeed, (1.2.7) yields that

$$\inf_{s\in[0,t]} \det\kappa(s,T)>0, \qquad t\in[0,T),$$

and hence

$$\sup_{(s,x)\in[0,t]\times\mathbb{R}^d} |h(s,x)| \leqslant \left((2\pi)^d \inf_{s\in[0,t]}\det\kappa(s,T)\right)^{-1/2} < \infty, \qquad t\in[0,T).$$

This yields that $P_{s,t}h(s,\mathbf{x})$ is defined for all $0 \leqslant s < t < T$ and $\mathbf{x} \in \mathbb{R}^d$, although it can happen that h is not bounded on $[0,T) \times \mathbb{R}^d$ (as it is in the case of $\mathbf{Z}$ being a one-dimensional standard Wiener process). Then h is space-time harmonic for the Markov process $\mathbf{Z}$ in the sense that

$$\begin{aligned} P_{s,t}h(s,\mathbf{x}) &= \int_{\mathbb{R}^d} h(t,\mathbf{y})\, p^{\mathbf{Z}}_{s,t}(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{y} = \int_{\mathbb{R}^d} p^{\mathbf{Z}}_{s,t}(\mathbf{x},\mathbf{y})\, p^{\mathbf{Z}}_{t,T}(\mathbf{y},\mathbf{b})\,\mathrm{d}\mathbf{y} \\ &= p^{\mathbf{Z}}_{s,T}(\mathbf{x},\mathbf{b}) = h(s,\mathbf{x}) \end{aligned}$$

for $0 \leqslant s < t < T$ and $\mathbf{x} \in \mathbb{R}^d$. Now a generalization of Doob's h-transform approach (see Doob [**61**] gives a new operator hemigroup

$$\widetilde{P}_{s,t}f = \frac{1}{h}\,P_{s,t}(hf), \quad 0 \leqslant s < t < T$$

with

$$\begin{aligned} \widetilde{P}_{s,t}f(s,\mathbf{x}) &= \frac{1}{h(s,\mathbf{x})}\,P_{s,t}(hf)(s,\mathbf{x}) = \frac{1}{h(s,\mathbf{x})}\int_{\mathbb{R}^d} h(t,\mathbf{y})f(t,\mathbf{y})\, p^{\mathbf{Z}}_{s,t}(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{y} \\ &= \int_{\mathbb{R}^d} f(t,\mathbf{y})\,\frac{p^{\mathbf{Z}}_{s,t}(\mathbf{x},\mathbf{y})\, p^{\mathbf{Z}}_{t,T}(\mathbf{y},\mathbf{b})}{p^{\mathbf{Z}}_{s,T}(\mathbf{x},\mathbf{b})}\,\mathrm{d}\mathbf{y} = \int_{\mathbb{R}^d} f(t,\mathbf{y})\, p^{\mathbf{U}}_{s,t}(\mathbf{x},\mathbf{y})\,\mathrm{d}\mathbf{y}, \end{aligned}$$

where $f : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ is a bounded Borel-measurable function and $\mathbf{x} \in \mathbb{R}^d$, i.e., the transition operators $(\widetilde{P}_{s,t})_{0\leqslant s<t<T}$ belong to a new Markov process $(\mathbf{U}_t)_{0\leqslant t<T}$, the desired process bridge, with transition densities $(p^{\mathbf{U}}_{s,t})_{0\leqslant s<t<T}$ given by (1.2.9).

For $T > 0$, $0 \leqslant s < t < T$ and $\mathbf{a},\mathbf{b} \in \mathbb{R}^d$, let us define

$$\begin{aligned} \Gamma(s,t) &:= E(s,t)\kappa(s,t) = \int_s^t E(s,u)\Sigma(u)\Sigma(u)^\top E(t,u)^\top\,\mathrm{d}u, \\ \Sigma(s,t) &:= \Gamma(t,T)\Gamma(s,T)^{-1}\Gamma(s,t), \end{aligned}$$

and

$$\mathbf{n}_{\mathbf{a},\mathbf{b}}(s,t) := \Gamma(t,T)\Gamma(s,T)^{-1}\mathbf{m}^{+}_{\mathbf{a}}(s,t) + \Gamma(s,t)^\top\big(\Gamma(s,T)^\top\big)^{-1}\mathbf{m}^{-}_{\mathbf{b}}(t,T). \tag{1.2.10}$$

The next result is about the existence of a Markov process $(\mathbf{U}_t)_{t\in[0,T]}$ with initial distribution $P(\mathbf{U}_0 = \mathbf{a}) = 1$ and with transition densities $p^{\mathbf{U}}_{s,t}$ given in (1.2.9) such that $\mathbf{U}_t \to \mathbf{b} =: \mathbf{U}_T$ almost surely and also in L^2 as $t \uparrow T$. First we present an auxiliary lemma.

Lemma 1.2.1. *Let us suppose that condition* (1.2.7) *holds. Let* $\mathbf{b} \in \mathbb{R}^d$ *and* $T > 0$ *be fixed. Then for all* $0 \leqslant s < t < T$ *and* $\mathbf{x},\mathbf{y} \in \mathbb{R}^d$ *we have*

$$\begin{aligned} &\frac{p^{\mathbf{Z}}_{s,t}(\mathbf{x},\mathbf{y})\, p^{\mathbf{Z}}_{t,T}(\mathbf{y},\mathbf{b})}{p^{\mathbf{Z}}_{s,T}(\mathbf{x},\mathbf{b})} \\ &\quad = \frac{1}{\sqrt{(2\pi)^d \det\Sigma(s,t)}}\,\exp\left\{-\frac{1}{2}\Big\langle \Sigma(s,t)^{-1}\big(\mathbf{y} - \mathbf{n}_{\mathbf{x},\mathbf{b}}(s,t)\big), \mathbf{y} - \mathbf{n}_{\mathbf{x},\mathbf{b}}(s,t)\Big\rangle\right\}, \end{aligned}$$

which is a Gauss density (in $\mathbf{y}$*) with mean vector* $\mathbf{n}_{\mathbf{x},\mathbf{b}}(s,t)$ *and with covariance matrix* $\Sigma(s,t)$*.*

Theorem 1.2.2. *Let us suppose that condition* (1.2.7) *holds. For fixed* $\mathbf{a},\mathbf{b} \in \mathbb{R}^d$ *and* $T > 0$*, let the process* $(\mathbf{U}_t)_{t\in[0,T)}$ *be given by*

$$\mathbf{U}_t := \mathbf{n}_{\mathbf{a},\mathbf{b}}(0,t) + \Gamma(t,T)\int_0^t \Gamma(u,T)^{-1}\Sigma(u)\,\mathrm{d}\mathbf{B}_u, \quad t \in [0,T). \tag{1.2.11}$$

Then for any $t \in [0,T)$ *the distribution of* $\mathbf{U}_t$ *is Gauss with mean* $\mathbf{n}_{\mathbf{a},\mathbf{b}}(0,t)$ *and covariance matrix* $\Sigma(0,t)$*. Especially,* $\mathbf{U}_t \to \mathbf{b}$ *almost surely (and hence in probability) and in* L^2 *as* $t \uparrow T$*. Hence the process* $(\mathbf{U}_t)_{t\in[0,T)}$ *can be extended to an* *almost surely (and hence stochastically) and* L^2*-continuous process* $(\mathbf{U}_t)_{t\in[0,T]}$ *with* $\mathbf{U}_0 = \mathbf{a}$ *and* $\mathbf{U}_T = \mathbf{b}$. *Moreover,* $(\mathbf{U}_t)_{t\in[0,T]}$ *is a Gauss-Markov process and for any* $\mathbf{x} \in \mathbb{R}^d$ *and* $0 \leqslant s < t < T$ *the transition density* $\mathbb{R}^d \ni \mathbf{y} \mapsto p^{\mathbf{U}}_{s,t}(\mathbf{x},\mathbf{y})$ *of* $\mathbf{U}_t$ *given* $\mathbf{U}_s = \mathbf{x}$ *is given by*

$$p^{\mathbf{U}}_{s,t}(\mathbf{x},\mathbf{y}) = \frac{1}{\sqrt{(2\pi)^d \det \Sigma(s,t)}} \exp\left\{-\frac{1}{2}\left\langle \Sigma(s,t)^{-1}\big(\mathbf{y}-\mathbf{n}_{\mathbf{x},\mathbf{b}}(s,t)\big), \mathbf{y}-\mathbf{n}_{\mathbf{x},\mathbf{b}}(s,t)\right\rangle\right\},$$

which coincides with the density given in Lemma 1.2.1.

Next we formulate an auxiliary result which is helpful for proving almost sure continuity of the linear process bridge at the endpoint T.

Lemma 1.2.3. *Let us assume that condition* (1.2.7) *holds. Let* $T \in (0,\infty)$ *be fixed and let* $(\mathbf{B}_t)_{t\geqslant 0}$ *be an p-dimensional standard Wiener process on a filtered probability space* $(\Omega, \mathcal{A}, (\mathcal{A}_t)_{t\in[0,T)}, P)$ *satisfying the usual conditions, constructed by the help of the standard Wiener process* $\mathbf{B}$ *(see, e.g., Karatzas and Shreve* [**100**, Section 5.2.A]*). The process* $(\mathbf{S}_t)_{t\in[0,T]}$ *defined by*

$$\mathbf{S}_t := \begin{cases} \Gamma(t,T)\int_0^t \Gamma(u,T)^{-1}\Sigma(u)\,\mathrm{d}\mathbf{B}_u & \text{if } t\in[0,T),\\ \mathbf{0} & \text{if } t = T,\end{cases} \tag{1.2.12}$$

is a centered Gauss process with almost surely continuous paths.

Definition 1.2.4. Let $(\mathbf{Z}_t)_{t\geqslant 0}$ be the d-dimensional linear process given by the SDE (1.2.1) with an initial Gauss random variable $\mathbf{Z}_0$ independent of $(\mathbf{B}_t)_{t\geqslant 0}$ and let us assume that condition (1.2.7) holds. For fixed $\mathbf{a},\mathbf{b}\in\mathbb{R}^d$ and $T>0$, the process $(\mathbf{U}_t)_{t\in[0,T]}$ defined in Theorem 1.2.2 is called a linear process bridge from $\mathbf{a}$ to $\mathbf{b}$ over $[0,T]$ derived from $\mathbf{Z}$. More generally, we call any almost surely continuous (Gauss) process on the time interval $[0,T]$ having the same finite-dimensional distributions as $(\mathbf{U}_t)_{t\in[0,T]}$ a multidimensional linear process bridge from $\mathbf{a}$ to $\mathbf{b}$ over $[0,T]$ derived from $\mathbf{Z}$.

Note that Definition 1.2.4 can be reformulated alternatively in a way that by a bridge from $\mathbf{a}$ to $\mathbf{b}$ over $[0,T]$ derived from $\mathbf{Z}$ we mean any almost surely continuous Gauss-Markov process $(\mathbf{U}_t)_{t\in[0,T]}$ with $\mathbf{U}_0=\mathbf{a}$, $\mathbf{U}_T=\mathbf{b}$ and with transition densities $(p^{\mathbf{U}}_{s,t})_{0\leqslant s<t<T}$ satisfying (1.2.9). Note also that the law of $(\mathbf{U}_t)_{t\in[0,T]}$ on $\big(C([0,T]),\mathcal{B}(C([0,T]))\big)$ is uniquely determined.

Formula (1.2.11) can be considered as an integral representation of the linear process bridge $\mathbf{U}$.

In the next theorem we present an SDE satisfied by the linear process bridge $\mathbf{U}$.

Theorem 1.2.5. *Let us suppose that condition* (1.2.7) *holds. The process* $(\mathbf{U}_t)_{t\in[0,T)}$ *defined by* (1.2.11) *is a strong solution of the linear SDE*

$$\begin{aligned}\mathrm{d}\mathbf{U}_t =&\Big[\big(Q(t)-\Sigma(t)\Sigma(t)^\top E(T,t)^\top\Gamma(t,T)^{-1}\big)\mathbf{U}_t\\ &+\Sigma(t)\Sigma(t)^\top\big(\Gamma(t,T)^\top\big)^{-1}\mathbf{m}^-_{\mathbf{b}}(t,T)+\mathbf{r}(t)\Big]\mathrm{d}t+\Sigma(t)\,\mathrm{d}\mathbf{B}_t\end{aligned} \tag{1.2.13}$$

for $t\in[0,T)$ *and with initial condition* $\mathbf{U}_0=\mathbf{a}$, *and strong uniqueness for the SDE* (1.2.13) *holds.*

Now we turn to give alternative representations of the bridge. The next theorem is about the existence of a so-called anticipative representation of the bridge which is a weak solution to the bridge SDE (1.2.13).

THEOREM 1.2.6. *Let* $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$ *and* $T > 0$ *be fixed. Let* $(\mathbf{Z}_t)_{t\geqslant 0}$ *be the linear process given by the SDE* (1.2.1) *with initial condition* $\mathbf{Z}_0 = \mathbf{0}$ *and let us assume that condition* (1.2.7) *holds. Then the process* $(\mathbf{Y}_t)_{t\in[0,T]}$ *given by*

$$\mathbf{Y}_t := \Gamma(t,T)\Gamma(0,T)^{-1}\mathbf{a} + \mathbf{Z}_t - \Gamma(0,t)^\top \big(\Gamma(0,T)^\top\big)^{-1}(\mathbf{Z}_T - \mathbf{b}), \quad t \in [0,T], \tag{1.2.14}$$

coincides in law the linear process bridge $(\mathbf{U}_t)_{t\in[0,T]}$ *from* $\mathbf{a}$ *to* $\mathbf{b}$ *over* $[0,T]$ *derived from* $\mathbf{Z}$.

Next we present the following result on the covariance structure of the linear process $\mathbf{Z}$ and its bridge $\mathbf{U}$ (given in Definition 1.2.4). We use this lemma in the proofs of Theorem 1.2.6 and Proposition 1.2.8.

LEMMA 1.2.7. *For fixed* $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$ *and* $T > 0$, *let* $(\mathbf{Z}_t)_{t\geqslant 0}$ *be the* d*-dimensional linear process given by the SDE* (1.2.1) *with a Gauss initial random vector* $\mathbf{Z}_0$ *independent of the underlying Wiener process* $(\mathbf{B}_t)_{t\geqslant 0}$. *Let us suppose that condition* (1.2.7) *holds and let* $(\mathbf{U}_t)_{t\in[0,T]}$ *be the linear process bridge from* $\mathbf{a}$ *to* $\mathbf{b}$ *over* $[0,T]$ *derived from* $\mathbf{Z}$ *(given by Theorem 1.2.2 and Definition 1.2.4). Then for* $0 \leqslant s \leqslant t$ *the covariance matrices of* $\mathbf{Z}$ *and* $\mathbf{U}$ *are given by*

(a) $\quad \operatorname{Cov}(\mathbf{Z}_s, \mathbf{Z}_t) = \operatorname{Cov}(\mathbf{Z}_t, \mathbf{Z}_s)^\top = \big(E(t,0)\Gamma(0,s)\big)^\top,$

(b) $\quad \operatorname{Cov}(\mathbf{U}_s, \mathbf{U}_t) = \operatorname{Cov}(\mathbf{U}_t, \mathbf{U}_s)^\top = \big(\Gamma(t,T)\Gamma(0,T)^{-1}\Gamma(0,s)\big)^\top.$

Next we present a usual conditioning property for multidimensional linear processes.

PROPOSITION 1.2.8. *Let* $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$ *and* $T > 0$ *be fixed. Let* $(\mathbf{Z}_t)_{t\geqslant 0}$ *be the* d*-dimensional linear process given by the SDE* (1.2.1) *with initial condition* $\mathbf{Z}_0 = \mathbf{a}$ *and let us assume that condition* (1.2.7) *holds. Let* $n \in \mathbb{N}$ *and* $0 < t_1 < t_2 < \ldots < t_n < T$. *Then the conditional distribution of* $(\mathbf{Z}_{t_1}^\top, \ldots, \mathbf{Z}_{t_n}^\top)^\top$ *given* $\mathbf{Z}_T = \mathbf{b}$ *coincides with the distribution of* $(\mathbf{U}_{t_1}^\top, \ldots, \mathbf{U}_{t_n}^\top)^\top$, *where* $(\mathbf{U}_t)_{t\in[0,T]}$ *is the linear process bridge from* $\mathbf{a}$ *to* $\mathbf{b}$ *over* $[0,T]$ *derived from* $(\mathbf{Z}_t)_{t\geqslant 0}$.

One can also realize that in case of time-homogeneity Proposition 1.2.8 is a consequence of Proposition 1 in Fitzsimmons, Pitman and Yor [**71**]. For more details, see our ArXiv preprint [**16**, page 10].

The next remark shows that the integral and anticipative representation of the bridge are quite different. To shed more light on the different behavior of various bridge representations, in a companion paper Barczy and Kern [**19**] we examined sample path deviations of the Wiener process and the Ornstein-Uhlenbeck process from its bridges.

REMARK 1.2.9. Note that the process $(\mathbf{Y}_t)_{t\in[0,T]}$ defined in (1.2.14) is only a weak solution of the SDE (1.2.13), since in contrast to the integral representation $(\mathbf{U}_t)_{t\in[0,T]}$ it is not adapted to the filtration $(\mathcal{F}_t)_{t\geqslant 0}$ of the underlying Wiener process $\mathbf{B}$. This can be easily seen by the definition of $\mathbf{Y}_t$ which requires the knowledge of $\mathbf{Z}_T$ at any time point $t \in (0,T)$. Nevertheless we have $\mathbf{Y}_t$ and $\mathbf{Z}_T$ are independent for any $t \in [0,T]$, since by part (a) of Lemma 1.2.7,

$$\begin{aligned}
\operatorname{Cov}(\mathbf{Y}_t, \mathbf{Z}_T) &= \operatorname{Cov}(\mathbf{Z}_t, \mathbf{Z}_T) - \Gamma(0,t)^\top \big(\Gamma(0,T)^\top\big)^{-1} \operatorname{Cov}(\mathbf{Z}_T, \mathbf{Z}_T) \\
&= \Gamma(0,t)^\top E(T,0)^\top - \Gamma(0,t)^\top \big(\Gamma(0,T)^\top\big)^{-1} \Gamma(0,T)^\top E(T,0)^\top \\
&= 0 \in \mathbb{R}^{d\times d},
\end{aligned}$$

and the random vector $(\mathbf{Y}_t^\top, \mathbf{Z}_T^\top)^\top$ has a Gauss distribution. □

In the next remark we compare the SDE (1.2.13) derived for the integral representation (1.2.11) and the anticipative representation (1.2.14) of the bridge $\mathbf{U}$ with the corresponding results of Delyon and Hu [**54**].

Remark 1.2.10. One may realize that Proposition 3 in Delyon and Hu [**54**] and our Theorem 1.2.5, and Theorem 2 in Delyon and Hu [**54**] and our Theorem 1.2.6 in principle are the same. For a more detailed comparison, see Remarks 2.2 and 2.3 in our ArXiv preprint [**16**]. Here first we only note that the definition of a bridge in Delyon and Hu [**54**] is different from our definition: they define a bridge as in Qian and Zheng [**141**], Lyons and Zheng [**122**], i.e., via Radon-Nikodym derivatives. We also note that the results of Qian and Zheng [**141**] and Lyons and Zheng [**122**] are valid for time-homogeneous diffusions, while Delyon and Hu [**54**] consider time inhomogeneous diffusions. Further, Qian and Zheng [**141**] refer to their Section 2.1 on conditional processes as a set of folklore facts for which they could not find a reference. We also remark that the SDE of Proposition 3 in Delyon and Hu [**54**] contains the solution of a deterministic differential equation which solution always remains abstract, while in our special case we have an explicit solution via evolution matrices. Moreover, also the proofs of Proposition 3 in Delyon and Hu [**54**] and of our Theorem 1.2.5 are different. □

1.3. One-dimensional linear process bridges

Let us consider a general one-dimensional linear process given by the linear SDE

$$\mathrm{d}Z_t = \big(q(t)\,Z_t + r(t)\big)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}B_t, \qquad t \geqslant 0, \tag{1.3.1}$$

with continuous functions $q : \mathbb{R}_+ \to \mathbb{R}$, $\sigma : \mathbb{R}_+ \to \mathbb{R}$ and $r : \mathbb{R}_+ \to \mathbb{R}$, where $(B_t)_{t\geqslant 0}$ is a standard Wiener process. By Section 5.6 in Karatzas and Shreve [**100**], it is known that there exists a strong solution of the SDE (1.3.1), namely

$$Z_t = \mathrm{e}^{\bar{q}(t)}\left(Z_0 + \int_0^t \mathrm{e}^{-\bar{q}(s)} r(s)\,\mathrm{d}s + \int_0^t \mathrm{e}^{-\bar{q}(s)}\sigma(s)\,\mathrm{d}B_s\right), \qquad t \geqslant 0, \tag{1.3.2}$$

with $\bar{q}(t) := \int_0^t q(u)\,\mathrm{d}u$, $t \geqslant 0$, and strong uniqueness for the SDE (1.3.1) holds. In what follows, we assume that Z_0 has a Gauss distribution independent of $(B_t)_{t\geqslant 0}$. We call the process $(Z_t)_{t\geqslant 0}$ a one-dimensional linear process. One can easily derive that for $0 \leqslant s < t$ we have

$$Z_t = \mathrm{e}^{\bar{q}(t)-\bar{q}(s)} Z_s + \int_s^t \mathrm{e}^{\bar{q}(t)-\bar{q}(u)} r(u)\,\mathrm{d}u + \int_s^t \mathrm{e}^{\bar{q}(t)-\bar{q}(u)}\sigma(u)\,\mathrm{d}B_u. \tag{1.3.3}$$

Hence, given $Z_s = x$, the distribution of Z_t does not depend on $(Z_r)_{r\in[0,s)}$ which yields that $(Z_t)_{t\geqslant 0}$ is a Markov process. Moreover, for any $x \in \mathbb{R}$ and $0 \leqslant s < t$ the conditional distribution of Z_t given $Z_s = x$ is Gauss with mean

$$m_x(s,t) := \mathrm{e}^{\bar{q}(t)-\bar{q}(s)} x + \int_s^t \mathrm{e}^{\bar{q}(t)-\bar{q}(u)} r(u)\,\mathrm{d}u,$$

and with variance

$$\gamma(s,t) := \int_s^t \mathrm{e}^{2(\bar{q}(t)-\bar{q}(u))}\sigma^2(u)\,\mathrm{d}u < \infty.$$

In what follows we put the following assumption

$$\sigma(t) \neq 0 \qquad \text{for all } \ t \geqslant 0. \tag{1.3.4}$$

We also note that one may weaken condition (1.3.4) to the following one: for every nonempty open interval $I \subseteq [0,\infty)$ there exist $t \in I$ such that $\sigma(t) \neq 0$. This yields that the variance $\gamma(s,t)$ is positive for all $0 \leqslant s < t$ (which corresponds to condition (1.2.7) in dimension one). Hence $(Z_t)_{t\geqslant 0}$ is a Gauss-Markov process with transition densities

(1.3.5)
$$p^Z_{s,t}(x,y) = \frac{1}{\sqrt{2\pi\gamma(s,t)}} \exp\left\{ -\frac{(y - m_x(s,t))^2}{2\gamma(s,t)} \right\}, \qquad 0 \leqslant s < t, \quad x, y \in \mathbb{R}.$$

For all $a, b \in \mathbb{R}$ and $0 \leqslant s \leqslant t < T$, let us introduce the notations

$$n_{a,b}(s,t) := \frac{\gamma(s,t)}{\gamma(s,T)} \mathrm{e}^{\bar{q}(T)-\bar{q}(t)} \left(b - \int_t^T \mathrm{e}^{\bar{q}(T)-\bar{q}(u)} r(u) \,\mathrm{d}u \right) + \frac{\gamma(t,T)}{\gamma(s,T)} m_a(s,t),$$

and

$$\sigma(s,t) := \frac{\gamma(s,t)\,\gamma(t,T)}{\gamma(s,T)}. \tag{1.3.6}$$

Theorem 1.2.2 has the following consequence.

THEOREM 1.3.1. *Let us suppose that condition* (1.3.4) *holds. For fixed* $a, b \in \mathbb{R}$ *and* $T > 0$, *let the process* $(U_t)_{t\in[0,T)}$ *be given by*

$$U_t := n_{a,b}(0,t) + \int_0^t \frac{\gamma(t,T)}{\gamma(s,T)} \mathrm{e}^{\bar{q}(t)-\bar{q}(s)} \sigma(s) \,\mathrm{d}B_s, \quad t \in [0,T). \tag{1.3.7}$$

Then for any $t \in [0,T)$ *the distribution of* U_t *is Gauss with mean* $n_{a,b}(0,t)$ *and with variance* $\sigma(0,t)$. *Especially,* $U_t \to b$ *almost surely (and hence in probability) and in* L^2 *as* $t \uparrow T$. *Hence the process* $(U_t)_{t\in[0,T)}$ *can be extended to an almost surely (and hence stochastically) and* L^2*-continuous process* $(U_t)_{t\in[0,T]}$ *with* $U_0 = a$ *and* $U_T = b$. *Moreover,* $(U_t)_{t\in[0,T]}$ *is a Gauss-Markov process and for any* $x \in \mathbb{R}$ *and* $0 \leqslant s < t < T$ *the transition density* $\mathbb{R} \ni y \mapsto p^U_{s,t}(x,y)$ *of* U_t *given* $U_s = x$ *is given by*

$$p^U_{s,t}(x,y) = \frac{1}{\sqrt{2\pi\sigma(s,t)}} \exp\left\{ -\frac{(y - n_{x,b}(s,t))^2}{2\sigma(s,t)} \right\}, \qquad y \in \mathbb{R}.$$

For completeness we formulate the definition of a one-dimensional linear process bridge, which definition is a special case of the multidimensional one (see Definition 1.2.4).

DEFINITION 1.3.2. Let $(Z_t)_{t\geqslant 0}$ be the one-dimensional linear process given by the SDE (1.3.1) with an initial Gauss random variable Z_0 independent of $(B_t)_{t\geqslant 0}$ and let us assume that condition (1.3.4) holds. For fixed $a, b \in \mathbb{R}$ and $T > 0$, the process $(U_t)_{t\in[0,T]}$ defined in Theorem 1.3.1 is called a linear process bridge from a to b over $[0,T]$ derived from Z. More generally, we call any almost surely continuous (Gauss) process on the time interval $[0,T]$ having the same finite-dimensional distributions as $(U_t)_{t\in[0,T]}$ a linear process bridge from a to b over $[0,T]$ derived from Z.

Theorem 1.2.5 has the following consequence.

THEOREM 1.3.3. *Let us suppose that condition* (1.3.4) *holds. The process* $(U_t)_{t\in[0,T)}$ *defined by* (1.3.7) *is a unique strong solution of the linear SDE*

$$\begin{aligned} \mathrm{d}U_t &= \left(q(t) - \frac{\mathrm{e}^{2(\bar{q}(T)-\bar{q}(t))}}{\gamma(t,T)} \sigma^2(t) \right) U_t \,\mathrm{d}t \\ &\quad + \left(r(t) + \frac{\mathrm{e}^{\bar{q}(T)-\bar{q}(t)}}{\gamma(t,T)} \left(b - \int_t^T \mathrm{e}^{\bar{q}(T)-\bar{q}(u)} r(u)\,\mathrm{d}u \right) \sigma^2(t) \right) \mathrm{d}t \\ &\quad + \sigma(t)\,\mathrm{d}B_t \end{aligned} \tag{1.3.8}$$

for $t \in [0,T)$ *and with initial condition* $U_0 = a,$ *and strong uniqueness for the SDE* (1.3.8) *holds.*

As a consequence of Theorem 1.2.6 we give an anticipative representation of the linear process bridge introduced in Theorem 1.3.1 and Definition 1.3.2.

THEOREM 1.3.4. *Let* $(Z_t)_{t\geqslant 0}$ *be a linear process given by the SDE* (1.3.1) *with initial condition* $Z_0 = 0$ *and let us suppose that condition* (1.3.4) *holds. Then the process* $(Y_t)_{t\in[0,T]}$ *given by*

$$Y_t := a\frac{\widetilde{R}(t,T)}{\widetilde{R}(0,T)} + Z_t - \frac{\widetilde{R}(0,t)}{\widetilde{R}(0,T)}(Z_T - b), \quad t \in [0,T],$$

equals in law the linear process bridge $(U_t)_{t\in[0,T]}$ *from* a *to* b *over* $[0,T]$ *derived from the process* $Z,$ *where*

$$\widetilde{R}(s,t) := \gamma(s,t)\mathrm{e}^{\bar{q}(s)-\bar{q}(t)}, \qquad 0 \leqslant s \leqslant t \leqslant T.$$

Moreover,

$$\widetilde{R}(s,t) = \mathrm{e}^{\bar{q}(s)-\bar{q}(t)}R(t,t) - \mathrm{e}^{\bar{q}(t)-\bar{q}(s)}R(s,s), \qquad 0 \leqslant s \leqslant t \leqslant T,$$

where R *denotes the covariance function of* $Z,$ *and*

$$\begin{aligned} Y_t = a &\left(\mathrm{e}^{\bar{q}(t)} - \mathrm{e}^{2\bar{q}(T)-\bar{q}(t)}\frac{\gamma(0,t)}{\gamma(0,T)}\right) + b\mathrm{e}^{\bar{q}(T)-\bar{q}(t)}\frac{\gamma(0,t)}{\gamma(0,T)} \\ &+ \left(Z_t - \mathrm{e}^{\bar{q}(T)-\bar{q}(t)}\frac{\gamma(0,t)}{\gamma(0,T)}Z_T\right), \qquad t \in [0,T]. \end{aligned} \tag{1.3.9}$$

We remark that the process $(Y_t)_{t\in[0,T]}$ in Theorem 1.3.4 can be written also in the form

$$Y_t = a\left(\mathrm{e}^{\bar{q}(t)} - \mathrm{e}^{\bar{q}(T)}\frac{R(t,T)}{R(T,T)}\right) + b\frac{R(t,T)}{R(T,T)} + \left(Z_t - \frac{R(t,T)}{R(T,T)}Z_T\right), \quad t \in [0,T].$$

We also note that in Remark 1.3.7 we will give an illuminating explanation for the representation (1.3.9).

As a consequence of Proposition 1.2.8 now we present a usual conditioning property for one-dimensional linear processes.

PROPOSITION 1.3.5. *Let* $a, b \in \mathbb{R}$ *and* $T > 0$ *be fixed. Let* $(Z_t)_{t\geqslant 0}$ *be the one-dimensional linear process given by the SDE* (1.3.1) *with initial condition* $Z_0 = a$ *and let us assume that condition* (1.3.4) *holds. Let* $n \in \mathbb{N}$ *and* $0 < t_1 < t_2 < \ldots < t_n < T.$ *Then the conditional distribution of* $(Z_{t_1}, \ldots, Z_{t_n})$ *given* $Z_T = b$ *coincides with the distribution of* $(U_{t_1}, \ldots, U_{t_n}),$ *where* $(U_t)_{t\in[0,T]}$ *is the linear process bridge from* a *to* b *over* $[0,T]$ *derived from* $(Z_t)_{t\geqslant 0}$.

Next we give an illuminating explanation for the representation (1.3.9) in Theorem 1.3.4 (see Remark 1.3.7), but preparatory we present a generalization of Lamperti transformation (see, e.g., Karlin and Taylor [**101**, page 218]) for one-dimensional linear processes. This generalization may be known, but we were not able to find any reference.

PROPOSITION 1.3.6. *Let* $(B_t^*)_{t\geqslant 0}$ *be a standard Wiener process starting from* 0 *and*

$$Z_t^* := m_0(0,t) + \mathrm{e}^{\bar{q}(t)}B^*(\mathrm{e}^{-2\bar{q}(t)}\gamma(0,t)), \quad t \geqslant 0.$$

Then $(Z_t^*)_{t\geqslant 0}$ *is a weak solution of the SDE* (1.3.1) *with initial condition* $Z_0^* = 0.$

REMARK 1.3.7. Using Proposition 1.3.6 one can give an illuminating explanation for the representation (1.3.9) in Theorem 1.3.4. By Problem 5.6.14 in Karatzas and Shreve [**100**], the process $(\widehat{U}_t)_{t\in[0,T]}$ defined by

$$\widehat{U}_t := a\frac{T-t}{T} + b\frac{t}{T} + \left(\widehat{B}_t - \frac{t}{T}\widehat{B}_T\right), \quad t\in[0,T],$$

equals in law the Wiener bridge from a to b over $[0,T]$, where $(\widehat{B}_t)_{t\geqslant 0}$ is a standard Wiener process, i.e., $(\widehat{U}_t)_{t\in[0,T]}$ is the anticipative representation of the Wiener bridge from a to b over $[0,T]$. Motivated by Lemma 1 in Papież and Sandison [**137**] and Proposition 1.3.6, first we will do the time change $[0,T]\ni t\mapsto \mathrm{e}^{-2\bar{q}(t)}\gamma(0,t)$, the rescaling with coefficient $\mathrm{e}^{\bar{q}(t)}$, and then the translation with $m_0(0,t)$ for the process $(\widehat{U}_t)_{t\in[0,T]}$. Namely, we consider the process

$$\begin{aligned}U_t^* := m_0(0,t) + \mathrm{e}^{\bar{q}(t)}\Bigg(&a\frac{\mathrm{e}^{-2\bar{q}(T)}\gamma(0,T)-\mathrm{e}^{-2\bar{q}(t)}\gamma(0,t)}{\mathrm{e}^{-2\bar{q}(T)}\gamma(0,T)} + b\frac{\mathrm{e}^{-2\bar{q}(t)}\gamma(0,t)}{\mathrm{e}^{-2\bar{q}(T)}\gamma(0,T)}\\
&+\widehat{B}(\mathrm{e}^{-2\bar{q}(t)}\gamma(0,t)) - \frac{\mathrm{e}^{-2\bar{q}(t)}\gamma(0,t)}{\mathrm{e}^{-2\bar{q}(T)}\gamma(0,T)}\widehat{B}(\mathrm{e}^{-2\bar{q}(T)}\gamma(0,T))\Bigg)\end{aligned}$$

for all $t\in[0,T]$. Then for all $t\in[0,T]$ we have

$$\begin{aligned}U_t^* &= m_0(0,t) + a\left(\mathrm{e}^{\bar{q}(t)} - \mathrm{e}^{2\bar{q}(T)-\bar{q}(t)}\frac{\gamma(0,t)}{\gamma(0,T)}\right) + b\mathrm{e}^{\bar{q}(T)}\mathrm{e}^{\bar{q}(T)-\bar{q}(t)}\frac{\gamma(0,t)}{\gamma(0,T)}\\
&\quad + \mathrm{e}^{\bar{q}(t)}\widehat{B}(\mathrm{e}^{-2\bar{q}(t)}\gamma(0,t)) - \mathrm{e}^{\bar{q}(T)-\bar{q}(t)}\frac{\gamma(0,t)}{\gamma(0,T)}\mathrm{e}^{\bar{q}(T)}\widehat{B}(\mathrm{e}^{-2\bar{q}(T)}\gamma(0,T))\\
&= a\left(\mathrm{e}^{\bar{q}(t)} - \mathrm{e}^{2\bar{q}(T)-\bar{q}(t)}\frac{\gamma(0,t)}{\gamma(0,T)}\right) + \big(\mathrm{e}^{\bar{q}(T)}b + m_0(0,T)\big)\mathrm{e}^{\bar{q}(T)-\bar{q}(t)}\frac{\gamma(0,t)}{\gamma(0,T)}\\
&\quad + \left(Z_t^* - \mathrm{e}^{\bar{q}(T)-\bar{q}(t)}\frac{\gamma(0,t)}{\gamma(0,T)}Z_T^*\right),\end{aligned}$$

where, using Proposition 1.3.6, $(Z_t^*)_{t\geqslant 0}$ equals in law the one-dimensional linear process given by the SDE (1.3.1) with initial condition $Z_0=0$. By Theorem 1.3.4, the process $(U_t^*)_{t\in[0,T]}$ equals in law the one-dimensional linear process bridge $(U_t)_{t\in[0,T]}$ from a to $\mathrm{e}^{\bar{q}(T)}b+m_0(0,T)$ over $[0,T]$ derived from Z given by the SDE (1.3.1) with initial condition $Z_0=0$. Roughly speaking, we have to apply the same time change, rescaling and translation to the anticipative representation of the Wiener bridge from a to b over $[0,T]$ in order to get the linear process bridge from a to $\mathrm{e}^{\bar{q}(T)}b+m_0(0,T)$ over $[0,T]$ (derived from Z given by the SDE (1.3.1) with initial condition $Z_0=0$) what we apply to a Wiener process in order to get the linear process Z.

Especially, concerning Wiener bridges and Ornstein-Uhlenbeck bridges, we have to apply the same time change and rescaling to the anticipative representation of the Wiener bridge from a to b over $[0,T]$ in order to get the Ornstein-Uhlenbeck bridge from a to $\mathrm{e}^{qT}b$ over $[0,T]$ (derived from Z given by the SDE (1.1.1)) what we apply to a Wiener process in order to get the Ornstein-Uhlenbeck process Z. We note that the original definition of an Ornstein-Uhlenbeck bridge of Papież and Sandison is different from ours, they define the bridge as a probability measure on the space of continuous functions $f:[0,T]\to\mathbb{R}$ such that $f(0)=a$ and $f(T)=\mathrm{e}^{qT}b$. □

Next we formulate special cases of the presented one-dimensional results.

REMARK 1.3.8. Note that in case of $q(t)=q\neq 0$, $t\geqslant 0$, and $\sigma(t)=\sigma\neq 0$, $t\geqslant 0$, (for any continuous deterministic forcing term r) the variance $\sigma(s,t)$ defined by (1.3.6) gives back (1.1.7). □

Theorem 1.3.1 has the following consequence.

REMARK 1.3.9. Note that in case of $q(t) = q = 0$, $\sigma(t) = \sigma \neq 0$, $r(t) = 0$, $t \geqslant 0$, and $a = 0 = b$, we recover the Wiener bridge $(\widetilde{U}_t)_{t\in[0,T]}$ from 0 to 0 stated in (1.1.8). Moreover, in case of $q(t) = q \neq 0$, $\sigma(t) = \sigma \neq 0$, and $r(t) = 0$, $t \geqslant 0$, the linear process bridge (Ornstein-Uhlenbeck bridge) $(U_t)_{t\in[0,T]}$ from a to b over $[0,T]$ defined in (1.3.7) has the form

(1.3.10)
$$U_t = a\,\frac{\sinh(q(T-t))}{\sinh(qT)} + b\,\frac{\sinh(qt)}{\sinh(qT)} + \sigma\int_0^t \frac{\sinh(q(T-t))}{\sinh(q(T-s))}\,\mathrm{d}B_s, \qquad t\in[0,T),$$

and admits transition densities

$$p^U_{s,t}(x,y) = \frac{1}{\sqrt{2\pi\sigma(s,t)}}\exp\left\{-\frac{\left(y - \frac{\sinh(q(t-s))}{\sinh(q(T-s))}\,b - \frac{\sinh(q(T-t))}{\sinh(q(T-s))}\,x\right)^2}{2\sigma(s,t)}\right\}$$

for all $0\leqslant s<t<T$ and $x,y\in\mathbb{R}$, where $\sigma(s,t)$ is given by (1.1.7). □

Theorem 1.3.3 has the following consequence.

REMARK 1.3.10. Note that in case of $q(t) = q \neq 0$, $\sigma(t) = \sigma \neq 0$ and $r(t) = 0$, $t\geqslant 0$, the SDE (1.3.8) has the form

(1.3.11)
$$\begin{cases} \mathrm{d}U_t = q\left(-\coth(q(T-t))\,U_t + \frac{b}{\sinh(q(T-t))}\right)\mathrm{d}t + \sigma\,\mathrm{d}B_t, & t\in[0,T),\\ U_0 = a. \end{cases}$$

Note also that both the SDE (1.3.11) and the integral representation (1.3.10) are invariant under a change of sign for the parameter q. Hence the Ornstein-Uhlenbeck bridges derived from the SDE (1.1.1) with q and $-q$ are (almost surely) pathwise identical. □

Theorem 1.3.4 has the following consequence.

REMARK 1.3.11. We consider a special case of Theorem 1.3.4, namely, let us suppose that $r(t) = 0$, $t\geqslant 0$, and that there exist real numbers $q\neq 0$ and $\sigma\neq 0$ such that $q(t) = q$, $t\geqslant 0$, and $\sigma(t) = \sigma$, $t\geqslant 0$. Then $\bar{q}(t) = qt$, $t\geqslant 0$, and

(1.3.12)
$$\begin{aligned}\widetilde{R}(s,t) &= \gamma(s,t)\mathrm{e}^{\bar{q}(s)-\bar{q}(t)} = \sigma^2\mathrm{e}^{q(s-t)}\int_s^t \mathrm{e}^{2q(t-u)}\,\mathrm{d}u = \sigma^2\mathrm{e}^{q(s-t)}\frac{1}{2q}(\mathrm{e}^{2q(t-s)}-1)\\ &= \frac{\sigma^2}{2q}(\mathrm{e}^{q(t-s)} - \mathrm{e}^{-q(t-s)}) = \frac{\sigma^2}{q}\sinh(q(t-s)), \quad 0\leqslant s\leqslant t\leqslant T,\end{aligned}$$

and

$$R(s,t) = \operatorname{Cov}(Z_s,Z_t) = \frac{\sigma^2}{2q}\mathrm{e}^{q(s+t)}(1-\mathrm{e}^{-2qs}) = \frac{\sigma^2}{q}\mathrm{e}^{qt}\sinh(qs), \quad 0\leqslant s\leqslant t.$$

An easy calculation shows that for all $t\in[0,T]$,

$$\begin{aligned}\frac{\widetilde{R}(0,t)}{\widetilde{R}(0,T)} &= \mathrm{e}^{q(T-t)}\,\frac{R(t,t)}{R(T,T)} = \frac{R(t,T)}{R(T,T)},\\ \frac{\widetilde{R}(t,T)}{\widetilde{R}(0,T)} &= \mathrm{e}^{qt} - \mathrm{e}^{2qT-qt}\,\frac{R(t,t)}{R(T,T)} = \frac{\sigma^2}{2q}\,\frac{\mathrm{e}^{qt+2qT}(1-\mathrm{e}^{-2qT}) - \mathrm{e}^{2qT+qt}(1-\mathrm{e}^{-2qt})}{R(T,T)}\\ &= \frac{\sigma^2}{2q}\,\frac{\mathrm{e}^{2qT-qt}-\mathrm{e}^{qt}}{R(T,T)} = \frac{R(T-t,T)}{R(T,T)}.\end{aligned}$$

Hence the process $(Y_t)_{t\in[0,T]}$ introduced in Theorem 1.3.4 (with our special choices of q, r and σ) has the form

$$Y_t = a\frac{R(T-t,T)}{R(T,T)} + b\frac{R(t,T)}{R(T,T)} + \left(Z_t - \frac{R(t,T)}{R(T,T)}Z_T\right), \quad t\in[0,T].$$

Moreover, by (1.3.12),

$$Y_t = a\frac{\sinh(q(T-t))}{\sinh(qT)} + b\frac{\sinh(qt)}{\sinh(qT)} + \left(Z_t - \frac{\sinh(qt)}{\sinh(qT)}Z_T\right), \quad t\in[0,T]. \tag{1.3.13}$$

Finally, we remark that in case of $q(t) = q \neq 0$, $\sigma(t) = \sigma \neq 0$, $t \geqslant 0$ and $r(t) = 0$, $t \geqslant 0$ with the special choices $q = -\sqrt{k\gamma}/2$ and $\sigma = k/4$, where $k > 0$ and $\gamma > 0$, our Theorem 1.3.4 is the same as Lemma 1 in Papież and Sandison [**137**]. □

For the comparison of Propositions 4 and 9 in Gasbarra, Sottinen and Valkeila [**76**] and our Theorems 1.3.4 and 1.3.1 (anticipative and integral representation of the bridge in case of dimension one), respectively, see our ArXiv preprint [**16**, Remark 3.6].

1.4. Appendix

First we give sufficient conditions for positive definiteness of the Kalman matrices introduced in Section 1.2, see, e.g., Theorems 7.7.1-7.7.3 in Conti [**47**].

PROPOSITION 1.4.1. *Let $0 \leqslant s < t$ be given. Then $\kappa(s,t)$ is positive definite if one of the following conditions is satisfied:*

(a) there exists $t_0 \in (s,t)$ such that $\Sigma(t_0)$ has full rank d (for which $p \geqslant d$ is required).

(b) there exist $t_0 \in (s,t)$, an open neighborhood I_0 around t_0 and some $k \in \mathbb{N}$ such that $\Sigma \in \mathcal{C}^{(k)}_{d\times p}(I_0)$, $Q \in \mathcal{C}^{(k-1)}_{d\times d}(I_0)$ and the controllability matrix $\big[\Sigma(t_0), \Delta\Sigma(t_0), \dots, \Delta^k\Sigma(t_0)\big]$ has full rank d, where Δ is the operator $\Delta\Sigma(t) = \Sigma'(t) - Q(t)\Sigma(t)$, $t \in I_0$ and for all $n, m \in \mathbb{N}$, $\mathcal{C}^{(k)}_{n\times m}(I_0)$ denotes the set of k-times continuously differentiable functions on I_0 with values in $\mathbb{R}^{n\times m}$ (for which $(k+1)p \geqslant d$ is required).

(c) there exist $t_0 \in (s,t)$, an open neighborhood I_0 around t_0 and some $k \in \mathbb{N}$ such that $\Sigma \in \mathcal{C}^{(\infty)}_{d\times p}(I_0)$, $Q \in \mathcal{C}^{(\infty)}_{d\times d}(I_0)$ and the controllability matrix $\big[\Sigma(t_0), \Delta\Sigma(t_0), \dots, \Delta^k\Sigma(t_0)\big]$ has full rank d, where $\mathcal{C}^{(\infty)}_{n\times m}(I_0)$ denotes the set of infinitely differentiable functions on I_0 with values in $\mathbb{R}^{n\times m}$ (for which $(k+1)p \geqslant d$ is required).

Next we present two auxiliary lemmata on matrix identities which are frequently used in the proofs.

LEMMA 1.4.2. *Let us suppose that condition* (1.2.7) *holds. For fixed $T > 0$ and all $0 \leqslant s < t < T$,*

$$\Sigma(s,t)^{-1} = \kappa(s,t)^{-1} + E(T,t)^\top \kappa(t,T)^{-1} E(T,t). \tag{1.4.1}$$

Especially, $\Sigma(s,t)$ is symmetric and positive definite for all $0 \leqslant s < t < T$.

LEMMA 1.4.3. *Let us suppose that condition* (1.2.7) *holds. For fixed $T > 0$ and all $0 \leqslant s < t < T$ we have*

(a) $\kappa(t,T)^{-1} - \kappa(s,T)^{-1} = \Gamma(s,T)^{-1}\Gamma(s,t)\big(\Gamma(t,T)^\top\big)^{-1}$,

(b) $\Sigma(s,t) = \Gamma(t,T)\displaystyle\int_s^t \Gamma(u,T)^{-1}\Sigma(u)\Sigma(u)^\top\big(\Gamma(u,T)^\top\big)^{-1}\,\mathrm{d}u\,\Gamma(t,T)^\top$,

(c) $E(t,0) - \Gamma(0,t)^\top\big(\Gamma(0,T)^\top\big)^{-1}E(T,0) = \Gamma(t,T)\Gamma(0,T)^{-1}$,

(d) $\Gamma(s,T)^\top E(t,s)^\top - E(T,s)\Gamma(s,t) = \Gamma(t,T)^\top$.

For proving almost sure continuity of the linear process bridge at the endpoint T, we recall a strong law of large numbers for continuous square integrable multivariate martingales with deterministic quadratic variation process due to Dzhaparidze and Spreij [**64**, Corollary 2]; see also Koval' [**110**, Corollary 1]. We note that the above mentioned citations are about continuous square integrable martingales with time interval $[0,\infty)$, but they are also valid for continuous square integrable martingales with time interval $[0,T)$, $T \in (0,\infty)$, with appropriate modifications in the conditions, see as follows (the proof of Koval' [**110**, Corollary 1] can be easily formulated for the time interval $[0,T)$, $T \in (0,\infty)$).

THEOREM 1.4.4. *Let $T \in (0,\infty]$ be fixed and let $(\Omega,\mathcal{G},(\mathcal{G}_t)_{t\in[0,T)},P)$ be a filtered probability space satisfying the usual conditions, i.e., $(\Omega,\mathcal{G},P)$ is complete, the filtration $(\mathcal{G}_t)_{t\in[0,T)}$ is right continuous, $\mathcal{G}_0$ contains all the P-null sets in $\mathcal{G}$ and $\mathcal{G}_{T-} = \mathcal{G}$, where $\mathcal{G}_{T-} := \sigma\left(\bigcup_{t\in[0,T)}\mathcal{G}_t\right)$. Let $(\mathbf{M}_t)_{t\in[0,T)}$ be an $\mathbb{R}^d$-valued continuous square integrable martingale with respect to the filtration $(\mathcal{G}_t)_{t\in[0,T)}$ such that $P(\mathbf{M}_0 = \mathbf{0}) = 1$. (The square integrability means that $\mathbb{E}(m_t^{(i)})^2 < \infty$, $t \in [0,T)$, $i = 1,\ldots,d$, where $\mathbf{M}_t := (m_t^{(1)},\ldots,m_t^{(d)})$, $t \in [0,T)$.) Further, we assume that the quadratic variation process $(\langle\mathbf{M}\rangle_t)_{t\in[0,T)}$ is deterministic (which yields that $(\langle\mathbf{M}\rangle_t)_{i,j} = \mathbb{E}(m_t^{(i)}m_t^{(j)})$, $t\in[0,T)$, $i,j = 1,\ldots,d$). If there exists some $t_0 \in [0,T)$ such that $\langle\mathbf{M}\rangle_{t_0}$ is positive definite and $\lim_{t\uparrow T}\langle\mathbf{M}\rangle_t^{-1} = 0 \in \mathbb{R}^{d\times d}$, then $P(\lim_{t\uparrow T}\langle\mathbf{M}\rangle_t^{-1}\mathbf{M}_t = \mathbf{0}) = 1$.*

CHAPTER 2

Explicit formulas for Laplace transforms of certain functionals of some time inhomogeneous diffusions

Co-author: Gyula Pap

2.1. Introduction

Several contributions have already been appeared containing explicit formulae for Laplace transforms of functionals of diffusion processes, see, e.g., Borodin and Salminen [**39**], Liptser and Shiryaev [**118**, Sections 7.7 and 17.3], Arató [**7**], Yor [**157**], Deheuvels and Martynov [**57**], Deheuvels, Peccati and Yor [**58**], Mansuy [**125**], Albanese and Lawi [**4**], Kleptsyna and Le Breton [**106**], [**107**], Hurd and Kuznetsov [**92**] and Gao, Hannig, Lee and Torcaso [**75**] (the latter one is about the Laplace transform of the squared L^2-norm of some Gauss processes). These formulae play an important role in theory of parameter estimation. Most of the literature concern time homogeneous diffusion processes.

To describe our aims, let us start with the usual Ornstein–Uhlenbeck process $(Z_t^{(\alpha)})_{t\geqslant 0}$ given by the stochastic differential equation (SDE)

$$\begin{cases} \mathrm{d}Z_t^{(\alpha)} = \alpha\, Z_t^{(\alpha)}\,\mathrm{d}t + \mathrm{d}B_t, \qquad t \geqslant 0,\\ Z_0^{(\alpha)} = 0, \end{cases}$$

where $\alpha \in \mathbb{R}$ and $(B_t)_{t\geqslant 0}$ is a standard Wiener process. An explicit formula is available for the Laplace transform of the random variable $\int_0^t (Z_s^{(\alpha)})^2\,\mathrm{d}s$, $t \geqslant 0$, namely, for all $t \geqslant 0$ and $\mu > 0$,

$$\begin{aligned} &\mathbb{E}\exp\left\{-\mu\int_0^t (Z_s^{(\alpha)})^2\,\mathrm{d}s\right\} \\ &\quad= \left(\frac{\mathrm{e}^{-\alpha t}\sqrt{\alpha^2+2\mu}}{\sqrt{\alpha^2+2\mu}\cosh(t\sqrt{\alpha^2+2\mu}) - \alpha\sinh(t\sqrt{\alpha^2+2\mu})}\right)^{\frac{1}{2}}, \end{aligned} \tag{2.1.1}$$

see, e.g., Liptser and Shiryaev [**118**, Lemma 17.3] or Gao, Hannig, Lee and Torcaso [**75**, Theorem 4].

Kleptsyna and Le Breton [**106**, Proposition 3.2] presented an extension of the above mentioned result for fractional Ornstein–Uhlenbeck type processes.

In case of a time homogeneous diffusion process $(H_t)_{t\geqslant 0}$, Albanese and Lawi [**4**] and Hurd and Kuznetsov [**92**] recently addressed the question whether it is possible to compute the Laplace transform

$$\mathbb{E}\left[\mathrm{e}^{-\int_0^t \phi(H_s)\,\mathrm{d}s} q(H_t)\right], \qquad t > 0,$$

in an analitically closed form, where $\phi, q : \mathbb{R} \to \mathbb{R}$ are Borel measurable functions. These papers provided a number of interesting cases when the Laplace transform can be evaluated in terms of special functions, such as hypergeometric functions. Their methods are based on probabilistic arguments involving Girsanov theorem, and alternatively on partial differential equations involving Feynman–Kac formula.

As new results, in case of some time inhomogeneous diffusion processes, we will derive an explicit formula for the joint Laplace transform of certain functionals of these processes using the ideas of Florens-Landais and Pham [**72**, Lemma 4.1], and see also Liptser and Shiryaev [**118**, Lemma 17.3]. Let $T \in (0, \infty]$ be fixed. Let $b : [0, T) \to \mathbb{R}$ and $\sigma : [0, T) \to \mathbb{R}$ be continuously differentiable functions. Suppose that $\sigma(t) > 0$ for all $t \in [0, T)$, and $b(t) \neq 0$ for all $t \in [0, T)$ (and hence $b(t) > 0$ for all $t \in [0, T)$ or $b(t) < 0$ for all $t \in [0, T)$). For all $\alpha \in \mathbb{R}$, consider the process $(X_t^{(\alpha)})_{t \in [0,T)}$ given by the SDE

$$\begin{cases} \mathrm{d}X_t^{(\alpha)} = \alpha\, b(t) X_t^{(\alpha)}\, \mathrm{d}t + \sigma(t)\, \mathrm{d}B_t, \qquad t \in [0, T), \\ X_0^{(\alpha)} = 0. \end{cases} \tag{2.1.2}$$

The SDE (2.1.2) is a special case of Hull–White (or extended Vasicek) model, see, e.g., Bishwal [**38**, page 3]. Assuming

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{b(t)}{\sigma(t)^2}\right) = -2K \frac{b(t)^2}{\sigma(t)^2}, \qquad t \in [0, T), \tag{2.1.3}$$

with some $K \in \mathbb{R}$, we derive an explicit formula for the joint Laplace transform of

$$\int_0^t \frac{b(s)^2}{\sigma(s)^2} (X_s^{(\alpha)})^2\, \mathrm{d}s \qquad \text{and} \qquad (X_t^{(\alpha)})^2 \tag{2.1.4}$$

for all $t \in [0, T)$ and for all $\alpha \in \mathbb{R}$, see Theorem 2.2.2.

We note that, using Lemma 11.6 in Liptser and Shiryaev [**118**], not assuming condition (2.1.3), one can derive the following formula for the Laplace transform of $\int_0^t \frac{b(s)^2}{\sigma(s)^2} (X_s^{(\alpha)})^2\, \mathrm{d}s$,

$$\mathbb{E} \exp\left\{-\mu \int_0^t \frac{b(s)^2}{\sigma(s)^2} (X_s^{(\alpha)})^2\, \mathrm{d}s\right\} = \exp\left\{\int_0^t \sigma(s)^2 \gamma_t(s)\, \mathrm{d}s\right\}, \qquad \mu > 0,$$

for all $t \in [0, T)$, where $\gamma_t : [0, t] \to \mathbb{R}$ is the unique solution of the Riccati differential equation

$$\begin{cases} \frac{\mathrm{d}\gamma_t}{\mathrm{d}s}(s) = 2\mu \frac{b(s)^2}{\sigma(s)^2} - 2\alpha b(s) \gamma_t(s) - \sigma(s)^2 \gamma_t(s)^2, \qquad s \in [0, t], \\ \gamma_t(t) = 0. \end{cases} \tag{2.1.5}$$

As a special case of our formula for the joint Laplace transform of (2.1.4), under the assumption (2.1.3), we have an explicit formula for the Laplace transform of $\int_0^t \frac{b(s)^2}{\sigma(s)^2} (X_s^{(\alpha)})^2\, \mathrm{d}s$, $t \in [0, T)$, see Theorem 2.2.2 with $\nu = 0$. We suspect that, under the assumption (2.1.3), the Riccati differential equation (2.1.5) may be solved explicitly.

We note that Deheuvels and Martynov [**57**] considered weighted Brownian motions $W_\gamma(t) := t^\gamma W_t$, $t \in (0, 1]$, with $W_\gamma(0) := 0$, and weighted Brownian bridges $B_\gamma(t) := t^\gamma W_t - t^{\gamma+1} W_1$, $t \in (0, 1]$, with $B_\gamma(t) := 0$, and with exponent $\gamma > -1$, where $(W_t)_{t \geqslant 0}$ is a standard Wiener process, and they explicitly calculated the Laplace transforms of the quadratic functionals $\int_0^1 W_\gamma(s)^2\, \mathrm{d}s$ and $\int_0^1 B_\gamma(s)^2\, \mathrm{d}s$ by means of Karhunen–Loève expansions. Deheuvels, Peccati and Yor [**58**] derived similar results for weighted Brownian sheets and bivariate weighted Brownian bridges. Motivated by Theorems 1.3 and 1.4 in Deheuvels and Martynov [**57**] and Theorem 4.1 in Deheuvels, Peccati and Yor [**58**], we conjecture that our explicit formula in Theorem 2.2.2 for the joint Laplace transform of (2.1.4) may be expressed as an infinite product containing the eigenvalues of the integral operator associated with the covariance function of $(X_t^{(\alpha)})_{t \in [0,T)}$. Assumption (2.1.3) may play a crucial role in the calculation of these eigenvalues and also for deriving a (weighted) Karhunen–Loève expansion for $(X_t^{(\alpha)})_{t\in[0,T)}$. Once a (weighted) Karhunen–Loève expansion is available for $(X_t^{(\alpha)})_{t\in[0,T)}$, one may derive the Laplace transform of $\int_0^t \frac{b(s)^2}{\sigma(s)^2}(X_s^{(\alpha)})^2\,\mathrm{d}s$, $t\in[0,T)$, as an infinite product. We note that this approach can be carried through in the special case of a so-called α-Wiener bridge with $\alpha = 1/2$ (introduced and discussed later on). Finally, we also remark that Gao, Hannig, Lee and Torcaso [**75**] used the same approach via Karhunen–Loève expansions for calculating the Laplace transform of the squared L^2-norm of some Gauss processes such as Ornstein-Uhlenbeck processes, time-changed Wiener bridges and integrated Wiener processes.

In Remark 2.2.4 we give a third possible explanation for the role of the assumption (2.1.3).

The random variables in (2.1.4) appear in the maximum likelihood estimator (MLE) $\widehat{\alpha}_t$ of α based on an observation $(X_s^{(\alpha)})_{s\in[0,\,t]}$. This is the reason why it is useful to calculate their joint Laplace transform explicitly. For a more detailed discussion, see Sections 2.3 and 2.4.

It is known that, under some conditions on b and σ (but without assumption (2.1.3)), the distribution of the MLE $\widehat{\alpha}_t$ of α normalized by Fisher information can converge to the standard normal distribution, to the Cauchy distribution or to the distribution of $c\int_0^1 W_s\,\mathrm{d}W_s \,/\, \int_0^1 (W_s)^2\,\mathrm{d}s$, where $(W_s)_{s\in[0,1]}$ is a standard Wiener process, and $c = 1/\sqrt{2}$ or $c = -1/\sqrt{2}$, see Luschgy [**120**, Section 4.2] and Barczy and Pap [**26**]. As an application of the joint Laplace transform of (2.1.4), under the conditions $\int_0^T \sigma(s)^2\,\mathrm{d}s < \infty$ and

$$b(t) = \frac{\sigma(t)^2}{-2K\int_t^T \sigma(s)^2\,\mathrm{d}s}, \qquad t\in[0,T), \tag{2.1.6}$$

with some $K \neq 0$ (note that in this case condition (2.1.3) is satisfied), we give an alternative proof for

$$\sqrt{I_\alpha(t)}\Big(\widehat{\alpha}_t - \alpha\Big) \xrightarrow{\mathcal{L}} \begin{cases} \mathcal{N}(0,1) & \text{if } \operatorname{sign}(\alpha-K) = \operatorname{sign}(K),\\ -\frac{\operatorname{sign}(K)}{\sqrt{2}}\,\frac{\int_0^1 W_s\,\mathrm{d}W_s}{\int_0^1 (W_s)^2\,\mathrm{d}s} & \text{if } \alpha = K, \end{cases}$$

as $t\uparrow T$, where $I_\alpha(t)$ denotes the Fisher information for α contained in the observation $(X_s^{(\alpha)})_{s\in[0,\,t]}$, $(W_s)_{s\in[0,1]}$ is a standard Wiener process and $\xrightarrow{\mathcal{L}}$ denotes convergence in distribution, see Theorem 2.3.6. In fact, in case of $\alpha = K$, for all $t\in(0,T)$,

$$\sqrt{I_K(t)}\,(\widehat{\alpha}_t - K) \stackrel{\mathcal{L}}{=} -\frac{\operatorname{sign}(K)}{2\sqrt{2}}\,\frac{(W_1)^2-1}{\int_0^1 (W_s)^2\,\mathrm{d}s} = -\frac{\operatorname{sign}(K)}{\sqrt{2}}\,\frac{\int_0^1 W_s\,\mathrm{d}W_s}{\int_0^1 (W_s)^2\,\mathrm{d}s},$$

where $\stackrel{\mathcal{L}}{=}$ denotes equality in distribution, see Theorem 2.3.6. We note that in case of $\operatorname{sign}(\alpha-K) = -\operatorname{sign}(K)$, one can prove $\sqrt{I_\alpha(t)}\,(\widehat{\alpha}_t-\alpha) \xrightarrow{\mathcal{L}} \zeta$ as $t\uparrow T$, where ζ is a random variable with standard Cauchy distribution, see, e.g., Luschgy [**120**, Section 4.2] or Barczy and Pap [**26**]. The proof in this case is based on a martingale limit theorem, and we do not know whether one can find a proof using the explicit form of the joint Laplace transform of (2.1.4).

By Barczy and Pap [**26**, Corollaries 9 and 11], under the conditions (2.1.6) and $\int_0^T \sigma(s)^2\,\mathrm{d}s < \infty$, we have for all $\alpha \neq K$, $K \neq 0$, the MLE $\widehat{\alpha}_t$ of α is asymptotically normal with an appropriate *random* normalizing factor, see also Remark 2.3.10. In case of $\alpha = K$, $K \neq 0$, under the above conditions, we determine the distribution of this randomly normalized MLE using the joint Laplace transform of (2.1.4), see Theorem 2.3.9. As a by-product of this result, giving a counterexample, we show that Remark 1.47 in Prakasa Rao [**142**] contains a mistake, see Remark 2.3.11.

Using the explicit form of the Laplace transform we also prove strong consistency of the MLE of α for all $\alpha \in \mathbb{R}$, see Theorem 2.3.12.

As an example, for all $\alpha \in \mathbb{R}$ and $T \in (0, \infty)$, we study the process $(X_t^{(\alpha)})_{t\in[0,T)}$ given by the SDE

$$\begin{cases} \mathrm{d}X_t^{(\alpha)} = -\frac{\alpha}{T-t} X_t^{(\alpha)}\,\mathrm{d}t + \mathrm{d}B_t, \qquad t \in [0,T), \\ \;X_0^{(\alpha)} = 0. \end{cases} \tag{2.1.7}$$

In case of $\alpha > 0$, this process is known as an α-Wiener bridge, and in case of $\alpha = 1$, this is the usual Wiener bridge. As a special case of the explicit form of the joint Laplace transform of (2.1.4), we obtain the joint Laplace transform of $\int_0^t \frac{(X_u^{(\alpha)})^2}{(T-u)^2}\,\mathrm{d}u$ and $(X_t^{(\alpha)})^2$ for all $t \in [0,T)$, see Theorem 2.4.1. As a special case of this latter formula we get the Laplace transform of $\int_0^t \frac{(B_u)^2}{(T-u)^2}\,\mathrm{d}u$, $t \in [0,T)$, which was first calculated by Mansuy [**125**, Proposition 5], see Remark 2.2.8. Finally, we remark that in case of $\alpha > 0$ unweighted and weighted Karhunen–Loève expansions are available for the α-Wiener bridge $(X_t^{(\alpha)})_{t\in[0,T)}$ on $[0,T]$ and $[0,S]$ with $0 < S < T$, respectively, see Barczy and Iglói [**14**]. Further, using the weighted Karhunen–Loève expansion, one can also get the Laplace transform of $\int_0^t \frac{(X_s^{(1/2)})^2}{(T-s)^2}\,\mathrm{d}s$, $t \in [0,T)$, see Barczy and Iglói [**14**, Proposition 3.1], i.e., in the special case of an α-Wiener bridge with $\alpha = 1/2$ the approach using Karhunen–Loève expansions mentioned earlier can be carried through.

2.2. Laplace transform

Let $T \in (0,\infty]$ be fixed. Let $b : [0,T) \to \mathbb{R}$ and $\sigma : [0,T) \to \mathbb{R}$ be continuously differentiable functions. Suppose that $\sigma(t) > 0$ for all $t \in [0,T)$, and $b(t) \neq 0$ for all $t \in [0,T)$ (and hence $b(t) > 0$ for all $t \in [0,T)$ or $b(t) < 0$ for all $t \in [0,T)$). For all $\alpha \in \mathbb{R}$, consider the SDE (2.1.2). Note that the drift and diffusion coefficients of the SDE (2.1.2) satisfy the local Lipschitz condition and the linear growth condition (see, e.g., Jacod and Shiryaev [**95**, Theorem 2.32, Chapter III]). By Jacod and Shiryaev [**95**, Theorem 2.32, Chapter III], the SDE (2.1.2) has a unique strong solution

$$X_t^{(\alpha)} = \int_0^t \sigma(s) \exp\left\{\alpha \int_s^t b(u)\,\mathrm{d}u\right\} \mathrm{d}B_s, \qquad t \in [0,T). \tag{2.2.1}$$

Note that $(X_t^{(\alpha)})_{t\in[0,T)}$ has continuous sample paths by the definition of strong solution, see, e.g., Jacod and Shiryaev [**95**, Definition 2.24, Chapter III]. For all $\alpha \in \mathbb{R}$ and $t \in (0,T)$, let $\mathbb{P}_{X^{(\alpha)},t}$ denote the distribution of the process $(X_s^{(\alpha)})_{s\in[0,t]}$ on $\big(C([0,t]), \mathcal{B}(C([0,t]))\big)$, where $C([0,t])$ and $\mathcal{B}(C([0,t]))$ denote the set of all continuous real valued functions defined on $[0,t]$ and the Borel σ-field on $C([0,t])$, respectively. The measures $\mathbb{P}_{X^{(\alpha)},t}$ and $\mathbb{P}_{X^{(\beta)},t}$ are equivalent for all α, $\beta \in \mathbb{R}$ and for all $t \in (0,T)$, and

$$\begin{aligned} &\frac{\mathrm{d}\,\mathbb{P}_{X^{(\alpha)},t}}{\mathrm{d}\,\mathbb{P}_{X^{(\beta)},t}}\Big(X^{(\beta)}\big|_{[0,t]}\Big) \\ &\quad = \exp\left\{(\alpha-\beta)\int_0^t \frac{b(s)}{\sigma(s)^2} X_s^{(\beta)}\,\mathrm{d}X_s^{(\beta)} - \frac{\alpha^2-\beta^2}{2}\int_0^t \frac{b(s)^2}{\sigma(s)^2}(X_s^{(\beta)})^2\,\mathrm{d}s\right\}, \end{aligned} \tag{2.2.2}$$

see, e.g., Liptser and Shiryaev [**117**, Theorem 7.19]. Note also that for all $s \in [0,T)$, $X_s^{(\alpha)}$ is normally distributed with mean 0 and with variance

$$V(s;\alpha) := \mathbb{E}\left(X_s^{(\alpha)}\right)^2 = \int_0^s \sigma(u)^2 \exp\left\{2\alpha \int_u^s b(v)\,\mathrm{d}v\right\} \mathrm{d}u, \qquad s \in [0,T), \tag{2.2.3}$$

and then, by the conditions on b and σ, $V(s;\alpha) > 0$ for all $s \in (0,T)$.

The next lemma is about the solutions of the differential equation (DE) (2.1.3).

Lemma 2.2.1. *Let $T \in (0,\infty]$ be fixed and let $b : [0,T) \to \mathbb{R} \setminus \{0\}$ and $\sigma : [0,T) \to (0,\infty)$ be continuously differentiable functions. The DE (2.1.3) leads to a Bernoulli type DE having solutions*

$$b(t) = \frac{\sigma(t)^2}{2\left(K \int_0^t \sigma(s)^2\,\mathrm{d}s + C\right)}, \qquad t \in [0,T), \tag{2.2.4}$$

where $C \in \mathbb{R}$ is such that the denominator $K \int_0^t \sigma(s)^2\,\mathrm{d}s + C \neq 0$ for all $t \in [0,T)$.

Now we derive an explicit formula for the joint Laplace transform of $\left(X_t^{(\alpha)}\right)^2$ and $\int_0^t \frac{b(s)^2}{\sigma(s)^2}\left(X_s^{(\alpha)}\right)^2 \mathrm{d}s$ for all $t \in [0,T)$ under the assumption (2.2.4) on b and σ. We use the same technique (sometimes called Novikov's method, see, e.g., Arató [**7**]) as in the proof of Lemma 4.1 in Florens-Landais and Pham [**72**] or see also the proof of Lemma 17.3 in Liptser and Shiryaev [**118**].

Theorem 2.2.2. *Let $(X_t^{(\alpha)})_{t\in[0,T)}$ be the process given by the SDE (2.1.2) where b is given by (2.2.4). Then for all $\mu > 0$, $\nu \geqslant 0$, and $t \in [0,T)$, we have*

$$\mathbb{E}\exp\left\{-\mu \int_0^t \frac{b(s)^2}{\sigma(s)^2}(X_s^{(\alpha)})^2\,\mathrm{d}s - \nu[X_t^{(\alpha)}]^2\right\}$$

$$= \frac{B_{K,C}(t)^{\frac{K-\alpha}{4}}}{\sqrt{\cosh\left(\frac{\sqrt{2\mu+(\alpha-K)^2}}{2}\ln(B_{K,C}(t))\right) - \frac{\alpha-K-4\nu\left(K\int_0^t \sigma(s)^2\,\mathrm{d}s + C\right)}{\sqrt{2\mu+(\alpha-K)^2}}\sinh\left(\frac{\sqrt{2\mu+(\alpha-K)^2}}{2}\ln(B_{K,C}(t))\right)}},$$

where

$$B_{K,C}(t) := \begin{cases} \left(1 + \frac{K}{C}\int_0^t \sigma(s)^2\,\mathrm{d}s\right)^{\frac{1}{K}} & \text{if } K \neq 0, \\ \exp\left\{\frac{1}{C}\int_0^t \sigma(s)^2\,\mathrm{d}s\right\} & \text{if } K = 0, \end{cases} \qquad t \in [0,T).$$

For the proof of Theorem 2.2.2 we need two lemmas. The first one can be considered as a preliminary version of Theorem 2.2.2, the second one is about the variance of $X_t^{(\alpha)}$.

Lemma 2.2.3. *Let $(X_t^{(\alpha)})_{t\in[0,T)}$ be the process given by the SDE (2.1.2). If assumption (2.1.3) is satisfied with some $K \in \mathbb{R}$ and if $\operatorname{sign}(b) = \pm\mathbf{1}_{[0,T)}$, then for all $\mu > 0$, $\nu \geqslant 0$ and $t \in [0,T)$, we have*

$$\begin{aligned}&\mathbb{E}\exp\left\{-\mu \int_0^t \frac{b(s)^2}{\sigma(s)^2}(X_s^{(\alpha)})^2\,\mathrm{d}s - \nu[X_t^{(\alpha)}]^2\right\} \\ &\qquad = \left(\frac{\exp\left\{-A^{\pm}_{\mu,\alpha,K}\int_0^t b(s)\,\mathrm{d}s\right\}}{1 + \left(2\nu - A^{\pm}_{\mu,\alpha,K}\frac{b(t)}{\sigma(t)^2}\right)V(t;\alpha - A^{\pm}_{\mu,\alpha,K})}\right)^{\frac{1}{2}},\end{aligned} \tag{2.2.5}$$

where $A^{\pm}_{\mu,\alpha,K} := \alpha - K \mp \sqrt{2\mu + (\alpha-K)^2}$.

Remark 2.2.4. Note that in Lemma 2.2.3 we do not use the explicit solutions of the DE (2.1.3) given in Lemma 2.2.1, since we wanted to demonstrate the role of condition (2.1.3) in the proof of Theorem 2.2.3. By this condition, the process $\int_0^t \left[\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{b(s)}{\sigma(s)^2}\right)\right](X_s^{(\beta)})^2\,\mathrm{d}s$, $t \in [0,T)$, has the form $-2K\int_0^t \frac{b(s)^2}{\sigma(s)^2}(X_s^{(\beta)})^2\,\mathrm{d}s$, $t \in$

$[0,T)$, and hence $\int_0^t \frac{b(s)}{\sigma(s)^2} X_s^{(\beta)} \, \mathrm{d}X_s^{(\beta)}$ can be expressed in terms of only the random variables $\big(X_t^{(\beta)}\big)^2$ and $\int_0^t \frac{b(s)^2}{\sigma(s)^2} (X_s^{(\beta)})^2 \, \mathrm{d}s$. As a consequence, in the calculation of $\Psi_t(\alpha,\mu,\nu)$ in the proof of Theorem 2.2.3, by the special choice of β, one can get rid of the stochastic integral $\int_0^t \frac{b(s)^2}{\sigma(s)^2} (X_s^{(\beta)})^2 \, \mathrm{d}s$. □

In the next lemma we calculate explicitly the variance $V(t;\alpha)$ of $X_t^{(\alpha)}$ for all $t \in [0,T)$.

Lemma 2.2.5. *Let* $(X_t^{(\alpha)})_{t\in[0,T)}$ *be the process given by the SDE* (2.1.2)*, where* b *is given by* (2.2.4)*. Then*

$$V(t;\alpha) = \begin{cases} \frac{C}{\alpha-K}\left(B_{K,C}(t)^\alpha - B_{K,C}(t)^K\right) & \text{if } \alpha \neq K, \\ C B_{K,C}(t)^K \ln(B_{K,C}(t)) & \text{if } \alpha = K, \end{cases}$$

where $B_{K,C}(t)$, $t\in[0,T)$, *is defined in Theorem 2.2.2.*

Remark 2.2.6. Note that formula (2.2.5) in Lemma 2.2.3 for the joint Laplace transform of (2.1.4) depends on the sign of the function $\operatorname{sign}(b)$, but in Theorem 2.2.2 it turned out that the sign is indifferent. We also remark that the case $b(t) < 0$, $t\in[0,T)$, can be traced back to the case $b(t) > 0$, $t\in[0,T)$, using the same arguments that are written for the case $b(t) < 0$, $t\in[0,T)$, at the end of the proof of Lemma 2.2.5. The point is that the formulae in Theorem 2.2.2 are invariant under the replacement of (α, b, K, C) with $(-\alpha, -b, -K, -C)$. □

In the next two remarks we consider special cases of Theorem 2.2.2.

Remark 2.2.7. As a special case of Theorem 2.2.2, one can get back formula (2.1.1) due to Liptser and Shiryaev [**118**, Lemma 17.3], and also the well-known Cameron–Martin formula for a standard Wiener process. Namely, let $T := \infty$, $b(t) := 1$, $t \geqslant 0$, and $\sigma(t) := 1$, $t\geqslant 0$. Let us consider the process $(X_t^{(\alpha)})_{t\in[0,T)}$ given by the SDE (2.1.2), which is the usual Ornstein–Uhlenbeck process starting from 0. Clearly, $\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{b(t)}{\sigma(t)^2}\right) = 0$, $t > 0$, and hence Theorem 2.2.2 with $\nu = 0$, $K = 0$ and with $C = \frac{1}{2}$ implies (2.1.1). With $\alpha = 0$, we get back the Cameron–Martin formula for a standard Wiener process,

$$\mathbb{E}\exp\left\{-\mu\int_0^t (B_u)^2 \, \mathrm{d}u\right\} = \frac{1}{\sqrt{\cosh(t\sqrt{2\mu})}}, \qquad t\geqslant 0, \quad \mu > 0,$$

see, e.g., Liptser and Shiryaev [**117**, formula (7.147)]. □

Remark 2.2.8. Let $T \in (0,\infty)$, $b(t) := -\frac{1}{T-t}$, $t\in[0,T)$, and $\sigma(t) := 1$, $t\in[0,T)$. Let us consider the process $(X_t^{(\alpha)})_{t\in[0,T)}$ given by the SDE (2.1.2). Hence condition (2.2.4) is satisfied with $K := \frac{1}{2}$ and $C := -\frac{T}{2}$, and clearly, $B_{K,C}(t) = (1-t/T)^2$, $t\in[0,T)$. Then Theorem 2.2.2 with $\nu = 0$ and $\alpha = 0$ implies that for all $\mu > 0$ and $t \in [0,T)$,

$$\mathbb{E}\exp\left\{-\frac{\mu}{2}\int_0^t \frac{(B_u)^2}{(T-u)^2}\,\mathrm{d}u\right\}$$
$$= \frac{\left(1-\frac{t}{T}\right)^{\frac{1}{4}}}{\sqrt{\cosh\left(\ln\left(1-\frac{t}{T}\right)\sqrt{\mu+\frac{1}{4}}\right) + \frac{1}{2\sqrt{\mu+\frac{1}{4}}}\sinh\left(\ln\left(1-\frac{t}{T}\right)\sqrt{\mu+\frac{1}{4}}\right)}}.$$

An easy calculation shows that for all $\mu > 0$ and $t \in [0, T)$,

$$\mathbb{E} \exp\left\{-\frac{\mu}{2}\int_0^t \frac{(B_u)^2}{(T-u)^2}\,\mathrm{d}u\right\} = \frac{\left(\frac{T-t}{T}\right)^{\frac{1+\sqrt{4\mu+1}}{4}}}{\sqrt{1-\frac{1+\sqrt{4\mu+1}}{2\sqrt{4\mu+1}}\left(1-\left(1-\frac{t}{T}\right)^{\sqrt{4\mu+1}}\right)}}.$$

This is the corrected formula of Proposition 5 in Mansuy [**125**], which contains a misprint. □

2.3. Maximum likelihood estimation via Laplace transform

As a special case of (2.2.2), the measures $\mathbb{P}_{X^{(\alpha)},\,t}$ and $\mathbb{P}_{X^{(0)},\,t}$ are equivalent for all $\alpha \in \mathbb{R}$ and for all $t \in (0, T)$, and

$$\frac{\mathrm{d}\,\mathbb{P}_{X^{(\alpha)},\,t}}{\mathrm{d}\,\mathbb{P}_{X^{(0)},\,t}}\left(X^{(\alpha)}\big|_{[0,t]}\right) = \exp\left\{\alpha\int_0^t \frac{b(s)}{\sigma(s)^2}X_s^{(\alpha)}\,\mathrm{d}X_s^{(\alpha)} - \frac{\alpha^2}{2}\int_0^t \frac{b(s)^2}{\sigma(s)^2}(X_s^{(\alpha)})^2\,\mathrm{d}s\right\}.$$

Here $\mathbb{P}_{X^{(0)},\,t}$ is nothing else but the Wiener measure on $\big(C([0,t]), \mathcal{B}(C([0,t]))\big)$.

For all $t \in (0, T)$, the maximum likelihood estimator $\widehat{\alpha}_t$ of the parameter α based on the observation $(X_s^{(\alpha)})_{s\in[0,\,t]}$ is defined by

$$\widehat{\alpha}_t := \arg\max_{\alpha\in\mathbb{R}} \ln\left(\frac{\mathrm{d}\,\mathbb{P}_{X^{(\alpha)},\,t}}{\mathrm{d}\,\mathbb{P}_{X^{(0)},\,t}}\left(X^{(\alpha)}\big|_{[0,t]}\right)\right).$$

The following lemma due to Barczy and Pap [**26**, Lemma 1] guarantees the existence of a unique MLE of α.

LEMMA 2.3.1. *For all $\alpha \in \mathbb{R}$ and $t \in (0, T)$, we have*

$$\mathbb{P}\left(\int_0^t \frac{b(s)^2}{\sigma(s)^2}(X_s^{(\alpha)})^2\,\mathrm{d}s > 0\right) = 1.$$

By Lemma 2.3.1, for all $t \in (0, T)$, there exists a unique maximum likelihood estimator $\widehat{\alpha}_t$ of the parameter α based on the observation $(X_s^{(\alpha)})_{s\in[0,\,t]}$ given by

$$\widehat{\alpha}_t = \frac{\int_0^t \frac{b(s)}{\sigma(s)^2}X_s^{(\alpha)}\,\mathrm{d}X_s^{(\alpha)}}{\int_0^t \frac{b(s)^2}{\sigma(s)^2}(X_s^{(\alpha)})^2\,\mathrm{d}s}, \qquad t \in (0, T).$$

To be more precise, by Lemma 2.3.1, for all $t \in (0, T)$, the MLE $\widehat{\alpha}_t$ exists $\mathbb{P}$-almost surely. Using the SDE (2.1.2) we obtain

$$\widehat{\alpha}_t - \alpha = \frac{\int_0^t \frac{b(s)}{\sigma(s)}X_s^{(\alpha)}\,\mathrm{d}B_s}{\int_0^t \frac{b(s)^2}{\sigma(s)^2}(X_s^{(\alpha)})^2\,\mathrm{d}s}, \qquad t \in (0, T). \tag{2.3.1}$$

For all $t \in (0, T)$, the Fisher information for α contained in the observation $(X_s^{(\alpha)})_{s\in[0,\,t]}$, is defined by

$$I_\alpha(t) := \mathbb{E}\left(\frac{\partial}{\partial\alpha}\ln\left(\frac{\mathrm{d}\,\mathbb{P}_{X^{(\alpha)},\,t}}{\mathrm{d}\,\mathbb{P}_{X^{(0)},\,t}}\left(X^{(\alpha)}\big|_{[0,\,t]}\right)\right)\right)^2 = \int_0^t \frac{b(s)^2}{\sigma(s)^2}\,\mathbb{E}\left(X_s^{(\alpha)}\right)^2\mathrm{d}s,$$

where the last equality follows by the SDE (2.1.2) and Karatzas and Shreve [**100**, Proposition 3.2.10]. Note that, by the conditions on b and σ, $I_\alpha : (0, T) \to (0, \infty)$ is an increasing function. Now we calculate the Fisher information $I_\alpha(t)$, $t \in (0, T)$, explicitly.

LEMMA 2.3.2. *Let* $(X_t^{(\alpha)})_{t\in[0,T)}$ *be the process given by the SDE* (2.1.2), *where* b *is given by* (2.2.4). *Then for all* $t \in (0, T)$,

$$I_\alpha(t) = \begin{cases} \frac{1}{4(\alpha-K)^2}(B_{K,C}(t)^{\alpha-K} - 1) - \frac{1}{4(\alpha-K)} \ln(B_{K,C}(t)) & \text{if } \alpha \neq K, \\ \frac{1}{8}(\ln(B_{K,C}(t)))^2 & \text{if } \alpha = K, \end{cases}$$

where $B_{K,C}(t)$, $t \in [0, T)$, *is defined in Theorem 2.2.2.*

Later on we intend to prove limit theorems for the MLE $\widehat{\alpha}_t$ of α normalized by Fisher information $I_\alpha(t)$. For proving these limit theorems, condition $\lim_{t\uparrow T} I_\alpha(t) = \infty$ plays a crucial role. In what follows we examine under what additional conditions on b and σ, $\lim_{t\uparrow T} I_\alpha(t) = \infty$ is satisfied.

LEMMA 2.3.3. *Let* $(X_t^{(\alpha)})_{t\in[0,T)}$ *be the process given by the SDE* (2.1.2), *where* b *is given by* (2.2.4). *In case of* $K \neq 0$,

$$\lim_{t\uparrow T} I_\alpha(t) = \infty \qquad \Longleftrightarrow \qquad \lim_{t\uparrow T} \int_0^t \sigma(u)^2 \,\mathrm{d}u = \begin{cases} \infty & \text{if } \frac{C}{K} > 0, \\ -\frac{C}{K} & \text{if } \frac{C}{K} < 0. \end{cases}$$

In case of $K = 0$, *we have*

$$\lim_{t\uparrow T} I_\alpha(t) = \infty \qquad \Longleftrightarrow \qquad \lim_{t\uparrow T} \int_0^t \sigma(u)^2 \,\mathrm{d}u = \infty.$$

Note that if the function $b : [0, T) \to \mathbb{R} \setminus \{0\}$ is given by (2.2.4) and if we suppose also that $K \neq 0$, $\frac{C}{K} < 0$, then, by Lemma 2.3.3, we have

$$C = -K \lim_{t\uparrow T} \int_0^t \sigma(u)^2 \,\mathrm{d}u =: -K \int_0^T \sigma(u)^2 \,\mathrm{d}u \in \mathbb{R} \setminus \{0\}, \tag{2.3.2}$$

and hence

$$b(t) = \frac{\sigma(t)^2}{2\left(K \int_0^t \sigma(u)^2 \,\mathrm{d}u - K \int_0^T \sigma(u)^2 \,\mathrm{d}u\right)} = \frac{\sigma(t)^2}{-2K \int_t^T \sigma(u)^2 \,\mathrm{d}u}, \qquad t \in [0, T),$$

which is nothing else but the form (2.1.6) of b. Moreover, by Lemma 2.3.3, we have $\lim_{t\uparrow T} I_\alpha(t) = \infty$ holds in this case.

In all what follows we will suppose that the function b is given by (2.1.6) with some $K \neq 0$, where $\int_0^T \sigma(u)^2 \,\mathrm{d}u < \infty$, and in this case, as an application of the explicit form of the joint Laplace transform of (2.1.4), we will give a complete description of the asymptotic behavior of the MLE $\widehat{\alpha}_t$ of α as $t \uparrow T$. In the other cases (for which $\lim_{t\uparrow T} I_\alpha(t) = \infty$) the asymptotic behavior of the MLE $\widehat{\alpha}_t$ as $t \uparrow T$ may be worked out using the same arguments as follows, but we do not consider these cases.

For our later purposes, we examine the asymptotic behavior of $I_\alpha(t)$ as $t \uparrow T$.

LEMMA 2.3.4. *Let* $(X_t^{(\alpha)})_{t\in[0,T)}$ *be the process given by the SDE* (2.1.2), *where* b *is given by* (2.1.6) *with some* $K \neq 0$ *and we suppose that* $\int_0^T \sigma(s)^2 \,\mathrm{d}s < \infty$. *Then in case of* $\operatorname{sign}(\alpha - K) = -\operatorname{sign}(K)$,

$$\lim_{t\uparrow T} \frac{I_\alpha(t)}{\frac{1}{4(K-\alpha)^2}\left(\frac{\int_0^T \sigma(s)^2 \,\mathrm{d}s}{\int_t^T \sigma(s)^2 \,\mathrm{d}s}\right)^{\frac{K-\alpha}{K}}} = 1,$$

in case of $\alpha = K$,

$$\lim_{t\uparrow T} \frac{I_\alpha(t)}{\frac{1}{8K^2}\left(\ln\left(\int_t^T \sigma(s)^2 \,\mathrm{d}s\right)\right)^2} = 1,$$

and in case of $\operatorname{sign}(\alpha - K) = \operatorname{sign}(K)$,

$$\lim_{t\uparrow T} \frac{I_\alpha(t)}{\frac{1}{4K(K-\alpha)} \ln\left(\int_t^T \sigma(s)^2 \,\mathrm{d}s\right)} = 1.$$

The next lemma is about the asymptotic behavior of the Laplace transform of the denominator in (2.3.1).

Lemma 2.3.5. *Let* $(X_t^{(\alpha)})_{t\in[0,T)}$ *be the process given by the SDE* (2.1.2)*, where* b *is given by* (2.1.6) *with some* $K \neq 0$ *and we suppose that* $\int_0^T \sigma(s)^2 \,\mathrm{d}s < \infty$. *Then*

(2.3.3)

$$\frac{1}{I_\alpha(t)} \int_0^t \frac{b(u)^2}{\sigma(u)^2} (X_u^{(\alpha)})^2 \,\mathrm{d}u \xrightarrow{\mathcal{L}} \begin{cases} (W_1)^2 & \text{if } \operatorname{sign}(\alpha-K) = -\operatorname{sign}(K), \\ 2\int_0^1 (W_s)^2 \,\mathrm{d}s & \text{if } \alpha = K, \\ 1 & \text{if } \operatorname{sign}(\alpha-K) = \operatorname{sign}(K), \end{cases}$$

as $t \uparrow T$, *where* $(W_s)_{s\in[0,1]}$ *is a standard Wiener process. In fact, in case of* $\alpha = K$, *for all* $t \in (0,T)$,

$$\frac{1}{I_K(t)} \int_0^t \frac{b(u)^2}{\sigma(u)^2} (X_u^{(K)})^2 \,\mathrm{d}u \stackrel{\mathcal{L}}{=} 2\int_0^1 (W_s)^2 \,\mathrm{d}s, \qquad t \in (0,T). \tag{2.3.4}$$

Theorem 2.3.6. *Let* $(X_t^{(\alpha)})_{t\in[0,T)}$ *be the process given by the SDE* (2.1.2)*, where* b *is given by* (2.1.6) *with some* $K \neq 0$ *and we suppose that* $\int_0^T \sigma(s)^2 \,\mathrm{d}s < \infty$. *Then*

$$\sqrt{I_\alpha(t)}\,(\widehat{\alpha}_t - \alpha) \xrightarrow{\mathcal{L}} \begin{cases} \mathcal{N}(0,1) & \text{if } \operatorname{sign}(\alpha-K) = \operatorname{sign}(K), \\ -\frac{\operatorname{sign}(K)}{\sqrt{2}} \frac{\int_0^1 W_s \,\mathrm{d}W_s}{\int_0^1 (W_s)^2 \,\mathrm{d}s} & \text{if } \alpha = K, \end{cases}$$

as $t \uparrow T$, *where* $(W_s)_{s\in[0,1]}$ *is a standard Wiener process. In fact, in case of* $\alpha = K$, *for all* $t \in (0,T)$,

$$\sqrt{I_K(t)}\,(\widehat{\alpha}_t - K) \stackrel{\mathcal{L}}{=} -\frac{\operatorname{sign}(K)}{2\sqrt{2}} \frac{(W_1)^2 - 1}{\int_0^1 (W_s)^2 \,\mathrm{d}s} = -\frac{\operatorname{sign}(K)}{\sqrt{2}} \frac{\int_0^1 W_s \,\mathrm{d}W_s}{\int_0^1 (W_s)^2 \,\mathrm{d}s}. \tag{2.3.5}$$

Remark 2.3.7. We note that Theorem 2.3.6 can be derived from our more general results, namely, from Barczy and Pap [**26**, Theorems 5 and 10]. We also remark that using these results one can also weaken the conditions on b and σ in Theorem 2.3.6. □

Remark 2.3.8. In case of $\operatorname{sign}(\alpha - K) = -\operatorname{sign}(K)$, under the conditions of Theorem 2.3.6, one can prove that

$$\sqrt{I_\alpha(t)}\,(\widehat{\alpha}_t - \alpha) \xrightarrow{\mathcal{L}} \zeta \qquad \text{as } t \uparrow T,$$

where ζ is a standard Cauchy distributed random variable, see, e.g., Luschgy [**120**, Section 4.2] or Barczy and Pap [**26**]. The proof in this case is based on a martingale limit theorem, and we do not know whether one can find a proof using the explicit form of the joint Laplace transform of (2.1.4). Lemma 2.3.5 implies only

$$\frac{1}{I_\alpha(t)} \int_0^t \frac{b(u)^2}{\sigma(u)^2} (X_u^{(\alpha)})^2 \,\mathrm{d}u \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)^2 \qquad \text{as } t \uparrow T. \tag{2.3.6}$$

However, using a martingale limit theorem, one can prove that the convergence in (2.3.6) holds almost surely (with some appropriate random variable ξ^2 as the limit). To be able to use Theorem 4 in Barczy and Pap [**26**], we need convergence in probability in (2.3.6). Hence the question is whether we can improve the convergence in distribution in (2.3.6) to convergence in probability using only the explicit form of the joint Laplace transform of (2.1.4). We do not know if one can find such a technique. □

The next theorem is about the (asymptotic) behavior of the MLE of $\alpha = K$, $K \neq 0$ using an appropriate *random* normalizing factor.

THEOREM 2.3.9. *Let $(X_t^{(K)})_{t\in[0,T)}$ be the process given by the SDE (2.1.2), where b is given by (2.1.6) with some $K \neq 0$ and we suppose that $\int_0^T \sigma(s)^2 \,\mathrm{d}s < \infty$. Then for all $t \in (0,T)$,*

$$\left(\int_0^t \frac{b(u)^2}{\sigma(u)^2}(X_u^{(K)})^2 \,\mathrm{d}u\right)^{\frac{1}{2}} (\widehat{\alpha}_t - K)$$
$$\stackrel{\mathcal{L}}{=} -\operatorname{sign}(K)\frac{\int_0^1 W_u \,\mathrm{d}W_u}{\left(\int_0^1 (W_u)^2 \,\mathrm{d}u\right)^{\frac{1}{2}}} = -\frac{\operatorname{sign}(K)}{2}\frac{(W_1)^2 - 1}{\left(\int_0^1 (W_u)^2 \,\mathrm{d}u\right)^{\frac{1}{2}}}.$$

REMARK 2.3.10. We note that, by Barczy and Pap [**26**, Corollaries 9 and 11], under the conditions $\int_0^T \sigma(s)^2 \,\mathrm{d}s < \infty$ and (2.1.6), we have for all $\alpha \neq K$, $K \neq 0$, the MLE of α is asymptotically normal with a corresponding *random* normalizing factor, namely, for all $\alpha \neq K$, $K \neq 0$,

$$\left(\int_0^t \frac{b(u)^2}{\sigma(u)^2}(X_u^{(\alpha)})^2 \,\mathrm{d}u\right)^{\frac{1}{2}} (\widehat{\alpha}_t - \alpha) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1) \qquad \text{as} \quad t \uparrow T.$$

□

As a consequence of Theorem 2.3.9, giving an illuminating counterexample, we show that Remark 1.47 in Prakasa Rao [**142**] contains a mistake.

REMARK 2.3.11. By giving a counterexample, we show that condition (1.5.26) in Remark 1.47 in Prakasa Rao [**142**] is not enough to assure (1.5.35) in Prakasa Rao [**142**]. By (2.3.1), we have for all $\alpha \in \mathbb{R}$ and $t \in (0,T)$,

$$\left(\int_0^t \frac{b(u)^2}{\sigma(u)^2}(X_u^{(\alpha)})^2 \,\mathrm{d}u\right)^{\frac{1}{2}} (\widehat{\alpha}_t - \alpha) = \frac{\frac{1}{\sqrt{I_\alpha(t)}} \int_0^t \frac{b(u)}{\sigma(u)} X_u^{(\alpha)} \,\mathrm{d}B_u}{\left(\frac{1}{I_\alpha(t)} \int_0^t \frac{b(u)^2}{\sigma(u)^2}(X_u^{(\alpha)})^2 \,\mathrm{d}u\right)^{1/2}}. \tag{2.3.7}$$

By Lemma 2.3.5 (under its conditions), we have

$$\frac{1}{I_K(t)} \int_0^t \frac{b(u)^2}{\sigma(u)^2}(X_u^{(K)})^2 \,\mathrm{d}u \stackrel{\mathcal{L}}{=} 2\int_0^1 (W_u)^2 \,\mathrm{d}u, \qquad t \in (0,T).$$

Hence if Remark 1.47 in Prakasa Rao [**142**] were true, then we would have

$$\left(\frac{1}{\sqrt{I_K(t)}} \int_0^t \frac{b(s)}{\sigma(s)} X_s^{(K)} \,\mathrm{d}B_s, \frac{1}{I_K(t)} \int_0^t \frac{b(s)^2}{\sigma(s)^2}(X_s^{(K)})^2 \,\mathrm{d}s\right)$$
$$\xrightarrow{\mathcal{L}} \left(\left(2\int_0^1 (W_u)^2 \,\mathrm{d}u\right)^{\frac{1}{2}} \xi\,,\ 2\int_0^1 (W_u)^2 \,\mathrm{d}u\right) \quad \text{as} \quad t \uparrow T,$$

where ξ is a standard normally distributed random variable independent of $\int_0^1 (W_u)^2 \,\mathrm{d}u$. By (2.3.7) and continuous mapping theorem, we would have

$$\left(\int_0^t \frac{b(u)^2}{\sigma(u)^2}(X_u^{(K)})^2 \,\mathrm{d}u\right)^{\frac{1}{2}} (\widehat{\alpha}_t - K) \xrightarrow{\mathcal{L}} \frac{\left(2\int_0^1 (W_u)^2 \,\mathrm{d}u\right)^{\frac{1}{2}} \xi}{\left(2\int_0^1 (W_u)^2 \,\mathrm{d}u\right)^{\frac{1}{2}}} = \xi \qquad \text{as} \quad t \uparrow T,$$

which is a contradiction, since, by Theorem 2.3.9, the limit distribution is

$$-\frac{\operatorname{sign}(K)}{2}\frac{(W_1)^2-1}{\left(\int_0^1 (W_u)^2\,\mathrm{d}u\right)^{\frac{1}{2}}}.$$

Note that this limit distribution can not be a standard normal distribution, see, e.g., Feigin [**67**, Section 2]. Indeed, in case of $K < 0$,

$$\mathbb{P}\left(-\frac{\operatorname{sign}(K)}{2}\frac{(W_1)^2-1}{\left(\int_0^1 (W_u)^2\,\mathrm{d}u\right)^{\frac{1}{2}}} > 0\right) = \mathbb{P}((W_1)^2 > 1) = 2(1-\Phi(1)),$$

which is not equal to $\mathbb{P}(\mathcal{N}(0,1) > 0) = \frac{1}{2}$. In case of $K > 0$, we can arrive at a contradiction similarly. $\square$

The next theorem is about the strong consistency of the MLE of α.

Theorem 2.3.12. *Let $(X_t^{(\alpha)})_{t\in[0,T)}$ be the process given by the SDE* (2.1.2), *where b is given by* (2.1.6) *with some $K \neq 0$ and we suppose that $\int_0^T \sigma(s)^2\,\mathrm{d}s < \infty$. Then the maximum likelihood estimator of α is strongly consistent, i.e., for all $\alpha \in \mathbb{R}$,*

$$\mathbb{P}\left(\lim_{t\uparrow T}\widehat{\alpha}_t = \alpha\right) = 1.$$

Finally, we note that in this section we studied the MLE $\widehat{\alpha}_t$ of α based on a *continuous* observation $(X_s^{(\alpha)})_{s\in[0,t]}$ using the results on Laplace transforms presented in Section 2.2. However, a continuous observation of a diffusion process is only a mathematical idealization, in practice the observation is always discrete. Hence one can pose the question whether our results on the MLE of α based on continuous observations give some information also for discrete observations. Parameter estimation for discretely observed diffusion processes has been studied by many authors, for a detailed discussion and references see, e.g., Bishwal [**38**]. For discrete observations, one possible approach is to try to find a good approximation of the MLE of α based on continuous observations (for example, Itô type approximation for the stochastic integral in the numerator of (2.3.1) and usual rectangular approximation for the ordinary integral in the denumerator of (2.3.1)). In this paper we do not consider this question.

2.4. α-Wiener bridge

For $T \in (0,\infty)$ and $\alpha \in \mathbb{R}$, let $(X_t^{(\alpha)})_{t\in[0,T)}$ be the process given by the SDE (2.1.7). To our knowledge, these kinds of processes in the case of $\alpha > 0$ have been first considered by Brennan and Schwartz [**41**], and see also Mansuy [**125**]. In Brennan and Schwartz [**41**] the SDE (2.1.7) is used to model the arbitrage profit associated with a given futures contract in the absence of transaction costs. By (2.2.1), the unique strong solution of the SDE (2.1.7) is

$$X_t^{(\alpha)} = \int_0^t \left(\frac{T-t}{T-s}\right)^{\alpha}\mathrm{d}B_s, \qquad t \in [0,T).$$

Theorem 2.2.2 has the following consequence on the joint Laplace transform of $\int_0^t \frac{(X_u^{(\alpha)})^2}{(T-u)^2}\,\mathrm{d}u$ and $(X_t^{(\alpha)})^2$.

THEOREM 2.4.1. *Let* $(X_t^{(\alpha)})_{t\in[0,T)}$ *be the process given by the SDE* (2.1.7). *For all* $\mu > 0$, $\nu \geqslant 0$ *and* $t \in [0,T)$, *we have*

$$\mathbb{E}\exp\left\{-\mu\int_0^t \frac{(X_u^{(\alpha)})^2}{(T-u)^2}\,\mathrm{d}u - \nu[X_t^{(\alpha)}]^2\right\}$$
$$= \frac{\left(1-\frac{t}{T}\right)^{(1-2\alpha)/4}}{\sqrt{\cosh\left(\frac{\sqrt{8\mu+(2\alpha-1)^2}}{2}\ln\left(1-\frac{t}{T}\right)\right) + \frac{1-2\alpha-4\nu(T-t)}{\sqrt{8\mu+(2\alpha-1)^2}}\sinh\left(\frac{\sqrt{8\mu+(2\alpha-1)^2}}{2}\ln\left(1-\frac{t}{T}\right)\right)}}.$$

Theorem 2.3.6 has the following consequence on the asymptotic behavior of the maximum likelihood estimator $\widehat{\alpha}_t$ of α as $t \uparrow T$.

THEOREM 2.4.2. *Let* $(X_t^{(\alpha)})_{t\in[0,T)}$ *be the process given by the SDE* (2.1.7). *For each* $\alpha > \frac{1}{2}$, *the maximum likelihood estimator* $\widehat{\alpha}_t$ *of* α *is asymptotically normal, namely, for each* $\alpha > \frac{1}{2}$,

$$\sqrt{I_\alpha(t)}(\widehat{\alpha}_t - \alpha) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1) \qquad \text{as } t \uparrow T.$$

If $\alpha = \frac{1}{2}$, *then the distribution of* $\sqrt{I_{1/2}(t)}\left(\widehat{\alpha}_t - \frac{1}{2}\right)$ *is the same for all* $t \in (0,T)$, *namely,*

$$\sqrt{I_{1/2}(t)}\left(\widehat{\alpha}_t - \frac{1}{2}\right) \stackrel{\mathcal{L}}{=} -\frac{1}{2\sqrt{2}}\,\frac{(W_1)^2-1}{\int_0^1 (W_s)^2\,\mathrm{d}s} = -\frac{1}{\sqrt{2}}\,\frac{\int_0^1 W_s\,\mathrm{d}W_s}{\int_0^1 (W_s)^2\,\mathrm{d}s},$$

where $(W_s)_{s\in[0,1]}$ *is a standard Wiener process.*

The following remark is about the asymptotic behavior of the MLE of α in case of $\alpha < \frac{1}{2}$. We note that up to our knowledge this case can not be handled using only Laplace transforms.

REMARK 2.4.3. If $\alpha < \frac{1}{2}$, then

$$\sqrt{I_\alpha(t)}\left(\widehat{\alpha}_t - \alpha\right) \xrightarrow{\mathcal{L}} \zeta \qquad \text{as } t \uparrow T,$$

where ζ is a standard Cauchy distributed random variable, see, e.g., Luschgy [**120**, Section 4.2] or Barczy and Pap [**26**]. □

Theorem 2.3.9 has the following consequence on the (asymptotic) behavior of the MLE of $\alpha = 1/2$ using a *random* normalization.

THEOREM 2.4.4. *Let* $(X_t^{(\alpha)})_{t\in[0,T)}$ *be the process given by the SDE* (2.1.7). *For all* $t \in (0,T)$, *we have*

$$\left(\int_0^t \frac{(X_u^{(1/2)})^2}{(T-u)^2}\,\mathrm{d}u\right)^{1/2}\left(\widehat{\alpha}_t - \frac{1}{2}\right) \stackrel{\mathcal{L}}{=} -\frac{\int_0^1 W_s\,\mathrm{d}W_s}{\left(\int_0^1 (W_s)^2\,\mathrm{d}s\right)^{1/2}} = -\frac{1}{2}\,\frac{(W_1)^2-1}{\left(\int_0^1 (W_s)^2\,\mathrm{d}s\right)^{1/2}}.$$

Finally, we note that Es-Sebaiy and Nourdin [**65**] studied the parameter estimation for so-called α-fractional bridges which are given by the SDE (2.1.7) replacing the standard Wiener process B by a fractional Wiener process.

CHAPTER 3

Karhunen-Loève expansions of alpha-Wiener bridges

Co-author: Endre Iglói

3.1. Introduction

There are few stochastic processes of interest, even among Gaussian ones, for which the Karhunen–Loève (KL) expansion is explicitly known. Some examples are those of the Wiener process, the Ornstein–Uhlenbeck process and the Wiener bridge, see, e.g., Ash and Gardner [**8**, Example 1.4.4], Papoulis [**138**, Problem 12.7], Liu and Ulukus [**119**, Section III], Corlay and Pagès [**48**, Section 5.4 B] and Deheuvels [**55**, Remark 1.1]. Recently, there is a renewed interest in this field: some KL expansions were provided for weighted Wiener processes and weighted Wiener bridges with weighting function having the form t^β (these expansions make use of Bessel functions), see Deheuvels and Martynov [**57**]. The most recent results on this topic are those of Deheuvels, Peccati and Yor [**58**], Deheuvels [**55**], [**56**], Luschgy and Pagès [**121**], Nazarov and Nikitin [**132**] and Nazarov and Pusev [**133**] (the latter two ones are about exact small deviation asymptotics for weighted L^2-norm of some Gaussian processes).

Let $0 < S < T < \infty$ and $0 < \alpha < \infty$ be arbitrarily fixed and let $(B_t)_{t\geqslant 0}$ be a standard Wiener process on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Let us consider the stochastic differential equation (SDE)

$$\begin{cases} \mathrm{d}X_t^{(\alpha)} = -\dfrac{\alpha}{T-t}\,X_t^{(\alpha)}\mathrm{d}t + \mathrm{d}B_t, & t \in [0, S], \\ X_0^{(\alpha)} = 0. \end{cases}$$

The drift and diffusion terms satisfy the usual Lipschitz and linear growth conditions, so, by Øksendal [**134**, Theorem 5.2.1] or Jacod and Shiryaev [**95**, Chapter III, Theorem 2.32], the above SDE has a strong solution which is pathwise unique (i.e., it has a unique strong solution). Since $S \in (0, T)$ is chosen arbitrarily, we obtain by successive extension that also the SDE

$$\begin{cases} \mathrm{d}X_t^{(\alpha)} = -\dfrac{\alpha}{T-t}\,X_t^{(\alpha)}\mathrm{d}t + \mathrm{d}B_t, & t \in [0, T), \\ X_0^{(\alpha)} = 0, \end{cases} \tag{3.1.1}$$

has a unique strong solution. Namely, it is

$$X_t^{(\alpha)} = \int_0^t \left(\frac{T-t}{T-s}\right)^{\alpha} \mathrm{d}B_s, \qquad t \in [0, T), \tag{3.1.2}$$

as it can be checked by Itô's formula. The process $(X_t^{(\alpha)})_{t\in[0,T)}$ given by (3.1.2) is called an α-Wiener bridge (from 0 to 0 on the time interval $[0, T]$). To our knowledge, these kind of processes have been first considered by Brennan and Schwartz [**41**], and see also Mansuy [**125**]. In Brennan and Schwartz [**41**] α-Wiener bridges are used to model the arbitrage profit associated with a given futures contract in the absence of transaction costs. Sondermann, Trede and Wilfling [**147**] and Trede and

Wilfling [**149**] use the SDE (3.1.1) to describe the fundamental component of an exchange rate process and they call the process $X^{(\alpha)}$ as a scaled Brownian bridge. The essence of these models is that the coefficient of $X_t^{(\alpha)}$ in the drift term in (3.1.1) represents some kind of mean reversion, a stabilizing force that keeps pulling the process towards its mean (zero in this reduced form), and the absolute value of this force is increasing proportionally to the inverse of the remaining time $T-t$, with the rate constant α.

This process has been also studied by Barczy and Pap [**27**], [**28**, Section 4] from several points of view, e.g., singularity of probability measures induced by the process $X^{(\alpha)}$ with different values of α, sample path properties, Laplace transforms of certain functionals of $X^{(\alpha)}$ and maximum likelihood estimation of α. The process $(X_t^{(\alpha)})_{t\in[0,T)}$ is Gaussian and for all $t\in[0,T)$, $\mathbb{E}\,X_t^{(\alpha)}=0$ and the covariance function of $X^{(\alpha)}$ given in Barczy and Pap [**27**, Lemma 2.1] is

$$\begin{aligned} R^{(\alpha)}(s,t) &:= \operatorname{Cov}\left(X_s^{(\alpha)},X_t^{(\alpha)}\right)\\ &= \begin{cases} \frac{(T-s)^\alpha(T-t)^\alpha}{1-2\alpha}\left(T^{1-2\alpha}-(T-(s\wedge t))^{1-2\alpha}\right) & \text{if } \alpha\neq\frac{1}{2},\\ \sqrt{(T-s)(T-t)}\,\ln\left(\frac{T}{T-(s\wedge t)}\right) & \text{if } \alpha=\frac{1}{2}, \end{cases} \end{aligned} \tag{3.1.3}$$

for all $s,t\in[0,T)$, where $s\wedge t:=\min(s,t)$. By Barczy and Pap [**27**, Lemma 3.1], the α-Wiener bridge$(X_t^{(\alpha)})_{t\in[0,T)}$ has an almost surely continuous extension $(X_t^{(\alpha)})_{t\in[0,T]}$ to the time interval $[0,T]$ such that $X_T^{(\alpha)}=0$ with probability one. The possibility of such an extension is based on that the parameter α is positive and on a strong law of large numbers for square integrable local martingales. We note here also that (3.1.1–3.1.2) continue to hold for $\alpha\leqslant 0$ as well. However, there does not exist an almost surely continuous extension of the process $(X_t^{(\alpha)})_{t\in[0,T)}$ onto $[0,T]$ which would take some constant at time T with probability one (i.e., which would be a bridge), and this is why the range of the parameter α is restricted to positive values. Indeed, for $\alpha=0$ we obtain the Wiener process, and in case of $\alpha<0$ the second moment of the solution $X_t^{(\alpha)}$ given by (3.1.2) converges to infinity, as (3.1.3) (with $s=t$) shows. Hence the assumption of the existence of an almost surely continuous extension to $[0,T]$ such that this extension takes some constant at time T with probability one (i.e., we have a bridge) would result in a contradiction. We note that another proof of the impossibility of such an extension in the case of $\alpha<0$ can be found in Barczy and Pap [**27**, Remark 3.5]. Finally, we remark that Mansuy [**125**, Proposition 4] studied the question whether it is possible to derive the α-Wiener bridge from a (single) Gaussian process by taking a bridge.

Next we check that the α-Wiener bridge$(X_t^{(\alpha)})_{t\in[0,T]}$ is L^2-continuous. By Theorem 1.3.4 in Ash and Gardner [**8**], it is enough to show that the covariance function $R^{(\alpha)}$ can be extended continuously onto $[0,T]^2:=[0,T]\times[0,T]$ such that this extension (which will be also denoted by $R^{(\alpha)}$) is zero on the set $\{(s,T):s\in[0,T]\}\cup\{(T,t):t\in[0,T]\}$. This follows by

$$\lim_{(s,t)\to(s_0,T)} R^{(\alpha)}(s,t)=\lim_{(s,t)\to(T,t_0)} R^{(\alpha)}(s,t)=0,\qquad s_0,t_0\in[0,T]. \tag{3.1.4}$$

Indeed, if $\alpha\neq 1/2$ and $s_0<T$, then

$$\lim_{(s,t)\to(s_0,T)} R^{(\alpha)}(s,t)=\frac{(T-s_0)^\alpha}{1-2\alpha}(T^{1-2\alpha}-(T-s_0)^{1-2\alpha})\lim_{t\uparrow T}(T-t)^\alpha=0.$$

If $0<\alpha<1/2$ and $s_0=T$, then

$$\lim_{(s,t)\to(s_0,T)} R^{(\alpha)}(s,t)=\frac{T^{1-2\alpha}}{1-2\alpha}\lim_{(s,t)\uparrow(T,T)}(T-s)^\alpha(T-t)^\alpha=0.$$

If $\alpha > 1/2$ and $s_0 = T$, then

$$\begin{aligned}\lim_{(s,t)\to(s_0,T)} R^{(\alpha)}(s,t) &= \frac{1}{2\alpha-1}\lim_{(s,t)\uparrow(T,T)} \frac{(T-s)^\alpha(T-t)^\alpha}{(T-(s\wedge t))^{2\alpha-1}}\\ &\leqslant \frac{1}{2\alpha-1}\lim_{(s,t)\uparrow(T,T)}(T-(s\wedge t)) = 0.\end{aligned}$$

If $\alpha = 1/2$ and $s_0 < T$, then

$$\lim_{(s,t)\to(s_0,T)} R^{(\alpha)}(s,t) = \sqrt{T-s_0}\ln\left(\frac{T}{T-s_0}\right)\lim_{t\uparrow T}\sqrt{T-t} = 0.$$

If $\alpha = 1/2$ and $s_0 = T$, then

$$\begin{aligned}\lim_{(s,t)\to(s_0,T)} R^{(\alpha)}(s,t) &= \lim_{(s,t)\uparrow(T,T)} \sqrt{(T-s)(T-t)}\ln\left(\frac{T}{T-(s\wedge t)}\right)\\ &\leqslant \lim_{(s,t)\uparrow(T,T)}(T-(s\wedge t))\ln\left(\frac{T}{T-(s\wedge t)}\right) = 0.\end{aligned}$$

We also have $R^{(\alpha)} \in L^2([0,T]^2)$. So, the integral operator associated to the kernel function $R^{(\alpha)}$, i.e., the operator $A_{R^{(\alpha)}} : L^2[0,T] \to L^2[0,T]$,

$$(3.1.5)\qquad (A_{R^{(\alpha)}}(\phi))(t) := \int_0^T R^{(\alpha)}(t,s)\phi(s)\,\mathrm{d}s,\quad t\in[0,T],\qquad \phi\in L^2[0,T],$$

is of the Hilbert–Schmidt type, thus $(X_t^{(\alpha)})_{t\in[0,T]}$ has a Karhunen–Loève (KL) expansion based on $[0,T]$:

$$(3.1.6)\qquad X_t^{(\alpha)} = \sum_{k=1}^{\infty}\sqrt{\lambda_k^{(\alpha)}}\,\xi_k e_k^{(\alpha)}(t),\quad t\in[0,T],$$

where $\xi_k,\ k\in\mathbb{N}$, are independent standard normally distributed random variables, $\lambda_k^{(\alpha)},\ k\in\mathbb{N}$, are the non-zero eigenvalues of the integral operator $A_{R^{(\alpha)}}$ and $e_k^{(\alpha)}(t),\ t\in[0,T],\ k\in\mathbb{N}$, are the corresponding normed eigenfunctions, which are pairwise orthogonal in $L^2[0,T]$, see, e.g., Ash and Gardner [**8**, Theorem 1.4.1]. Observe that (3.1.6) has infinitely many terms. Indeed, if it had a finite number of terms, i.e., if there were only a finite number of eigenfunctions, say N, then by the help of (3.1.1) (considering it as an integral equation) we would obtain that the Wiener process $(B_t)_{t\in[0,T]}$ is concentrated in an N-dimensional subspace of $L^2[0,T]$, and so even of $C[0,T]$, with probability one. This results in a contradiction, since the integral operator associated to the covariance function (as a kernel function) of a standard Wiener process has infinitely many eigenvalues and eigenfunctions. We also note that the normed eigenfunctions are unique only up to sign. The series in (3.1.6) converges in $L^2(\Omega,\mathcal{A},P)$ to $X_t^{(\alpha)}$, uniformly on $[0,T]$, i.e.,

$$\sup_{t\in[0,T]}\mathbb{E}\left(\left|X_t^{(\alpha)} - \sum_{k=1}^{n}\sqrt{\lambda_k^{(\alpha)}}\,\xi_k e_k^{(\alpha)}(t)\right|^2\right)\to 0\qquad\text{as } n\to\infty.$$

Moreover, since $R^{(\alpha)}$ is continuous on $[0,T]^2$, the eigenfunctions corresponding to non-zero eigenvalues are also continuous on $[0,T]$, see, e.g. Ash and Gardner [**8**, p. 38] (this will be important in the proof of Theorem 3.2.1, too). Since the terms on the right-hand side of (3.1.6) are independent normally distributed random variables and $(X_t^{(\alpha)})_{t\in[0,T]}$ has continuous sample paths with probability one, the series converges even uniformly on $[0,T]$ with probability one (see, e.g., Adler [**1**, Theorem 3.8]).

The rest of the paper is organized as follows. In Section 3.2 we make the KL representation (3.1.6) of the α-Wiener bridge $(X_t^{(\alpha)})_{t\in[0,T]}$ as explicit as it is possible, see Theorem 3.2.1. We also consider two special cases of the KL representation (3.1.6), first we study the case $\alpha\downarrow 0$ (standard Wiener process) and then the case $\alpha=1$ (Wiener bridge), see Remark 3.2.3 and Remark 3.2.4, respectively. In Remark 3.2.5 we present a so-called space-time transformed Wiener process representation of the α-Wiener bridge. Using this representation and a result of Gutiérrez and Valderrama [**80**, Theorem 1], we derive a weighted KL representation of the α-Wiener bridge$(X_t^{(\alpha)})_{t\in[0,T]}$ based on $[0,S]$, where $0<S<T$, see Theorem 3.2.6. We also consider two special cases of this weighted KL representation, first we study the case $\alpha\downarrow 0$ and then the case $\alpha=1$, see Remark 3.2.7 and Remark 3.2.8, respectively. Further, we give an infinite series representation of $\int_0^S (X_u^{(1/2)})^2/(T-u)^2\,\mathrm{d}u$, where $0<S<T$, see Remark 3.2.9. Section 3.3 is devoted to the applications. In Proposition 3.3.1 we determine the Laplace transform of the $L^2[0,T]$-norm square of $(X_t^{(\alpha)})_{t\in[0,T]}$ and of the $L^2[0,S]$-norm square of $(X_t^{(1/2)}/(T-t))_{t\in[0,S]}$, where $0<S<T$. In Corollary 3.3.2 we give a new probabilistic proof for the well-known result of Rayleigh, namely, for the sum of the square of the reciprocals of the positive zeros of Bessel functions of the first kind (with order greater than $-1/2$ in our case). Based on the Smirnov formula (see, e.g., Smirnov [**146**, formula (97)]) we write the survival function of the $L^2[0,T]$-norm square of $(X_t^{(\alpha)})_{t\in[0,T]}$ in an infinite series form, see Proposition 3.3.3. We also consider two special cases of Proposition 3.3.3, the case $\alpha\downarrow 0$ and the case $\alpha=1$, see Remark 3.3.4 and Remark 3.3.5. In Corollary 3.3.7, based on a result of Zolotarev [**159**], we study large deviation probabilities for the $L^2[0,T]$-norm square of the α-Wiener bridge. Finally, based on a result of Li [**114**, Theorem 2], we describe the behavior at zero of the distribution function (small deviation probabilities) of the $L^2[0,T]$-norm square of the α-Wiener bridge, see Corollay 3.3.8. In the appendix we list some important properties of Bessel functions of the first and second kind, respectively.

We remark that our results for α-Wiener bridges may have some generalizations for random fields. Namely, for all $S>0$, $T>0$ and $\alpha\geqslant 0$, $\beta\geqslant 0$, one can consider a zero-mean Gaussian random field $(X_{s,t}^{(\alpha,\beta)})_{(s,t)\in[0,S]\times[0,T]}$ with the covariances

$$\mathbb{E}(X_{s_1,t_1}^{(\alpha,\beta)}X_{s_2,t_2}^{(\alpha,\beta)})=R^{(\alpha)}(s_1,s_2)R^{(\beta)}(t_1,t_2),\qquad (s_1,t_1),(s_2,t_2)\in[0,S]\times[0,T].$$

Such a random field exists, since $(X_s^{(\alpha)}X_t^{(\beta)})_{(s,t)\in[0,S]\times[0,T]}$ admits the above covariances, where $X^{(\alpha)}$ and $X^{(\beta)}$ are independent, and Kolmogorov's consistency theorem comes into play. This class of Gaussian processes may deserve more attention since it would generalize some well-known limit processes in mathematical statistics such as the Kiefer process (known also a tied down Brownian sheet), see, e.g., Csörgő and Révész [**50**, Section 1.15]. Indeed, with $S=1$, $T=\infty$, $\alpha=1$ and $\beta=0$ the process $X^{(\alpha,\beta)}$ is nothing else but the Kiefer process having covariance function $(s_1\wedge s_2-s_1s_2)(t_1\wedge t_2)$, $(s_1,t_1),(s_2,t_2)\in[0,1]\times[0,\infty)$.

In all what follows $\mathbb{N}$, $\mathbb{Z}_+$ and $\mathbb{Z}$ denote the set of natural numbers, nonnegative integers and integers, respectively.

3.2. Karhunen–Loève expansions of α-Wiener bridges

First we recall the notion of Bessel functions of the first kind which plays a key role in the KL expansions we will obtain. They can be defined as

$$J_\nu(x):=\sum_{k=0}^{\infty}\frac{(-1)^k}{k!\,\Gamma(k+\nu+1)}\left(\frac{x}{2}\right)^{2k+\nu},\quad x\in(0,\infty),\ \nu\in\mathbb{R},$$

where $\Gamma(z)$ for $z < 0$, $z \notin \mathbb{Z}$, is defined by a recursive application of the rule $\Gamma(z) = \Gamma(z+1)/z$, $z < 0$, $z \notin \mathbb{Z}$, and we use the convention that $1/\Gamma(-k) := 0$, $k \in \mathbb{Z}_+$, yielding that the first n terms in the series of $J_\nu(x)$ vanish if $\nu = -n$, $n \in \mathbb{N}$, see, e.g., Watson [**152**, pp. 40, 64].

In all what follows we will put $\nu := \alpha - 1/2$, where $\alpha > 0$. Next we present our main theorem.

THEOREM 3.2.1. *Let $\alpha > 0$, $\nu := \alpha - 1/2$, and $z_k^{(\nu)}$, $k \in \mathbb{N}$, be the (positive) zeros of J_ν. Then in the KL expansion (3.1.6) of the α-Wiener bridge $(X_t^{(\alpha)})_{t\in[0,T]}$ the eigenvalues and the corresponding normed eigenfunctions are*

$$\lambda_k^{(\alpha)} = \frac{T^2}{(z_k^{(\nu)})^2}, \quad k \in \mathbb{N}, \tag{3.2.1}$$

$$\begin{aligned} e_k^{(\alpha)}(t) &= \sqrt{\frac{2}{T}\left(1 - \frac{t}{T}\right)}\, \frac{J_\nu\big(z_k^{(\nu)}(1 - t/T)\big)}{\big|J_{\nu+1}\big(z_k^{(\nu)}\big)\big|} \\ &= \sqrt{\frac{2}{T}\left(1 - \frac{t}{T}\right)}\, \frac{J_\nu\big(z_k^{(\nu)}(1 - t/T)\big)}{\big|J_{\nu-1}\big(z_k^{(\nu)}\big)\big|}, \quad t \in [0, T], \end{aligned}$$

where we take the continuous extension of $e_k^{(\alpha)}$ at $t = T$ for $-1/2 < \nu < 0$, i.e., $e_k^{(\alpha)}(T) = 0$ for $\alpha < 1/2$ (see part (i) of Proposition 3.4.1).

In the next remark we study the question whether 0 is an eigenvalue of the integral operator $A_{R^{(\alpha)}}$ or not.

REMARK 3.2.2. We note that 0 is not an eigenvalue of the integral operator $A_{R^{(\alpha)}}$. Indeed, on the contrary let us suppose that 0 is an eigenvalue of $A_{R^{(\alpha)}}$. We may assume without loss of generality that $T = 1$ (see the end of the proof of Theorem 3.2.1). Then there exists a function $e : [0, 1] \to \mathbb{R}$ which is not 0 almost everywhere and

$$\int_0^1 R^{(\alpha)}(t, s) e^{(\alpha)}(s)\, \mathrm{d}s = 0, \qquad t \in [0, 1].$$

First let us suppose that $\alpha \neq 1/2$. By the proof of Theorem 3.2.1, we have

$$\int_0^1 (1 - s)^\alpha e(s)\, \mathrm{d}s - \int_0^t (1 - s)^{1-\alpha} e(s)\, \mathrm{d}s - (1 - t)^{1-2\alpha} \int_t^1 (1 - s)^\alpha e(s)\, \mathrm{d}s = 0$$

for all $t \in (0, 1)$, and differentiating with respect to t,

$$\begin{aligned} -(1 - t)^{1-\alpha} e(t) + (1 - 2\alpha)(1 - t)^{-2\alpha} &\int_t^1 (1 - s)^\alpha e(s)\, \mathrm{d}s \\ &+ (1 - t)^{1-2\alpha}(1 - t)^\alpha e(t) = 0, \end{aligned}$$

or equivalently

$$\int_t^1 (1 - s)^\alpha e(s)\, \mathrm{d}s = 0, \qquad t \in (0, 1).$$

Differentiating again with respect to t one can derive $e(t) = 0$, $t \in (0, 1)$, which leads us to a contradiction. The case $\alpha = 1/2$ can be handled in a similar way. □

In the next remark we will study the convergence of the coefficients of the random variables in the terms on the right-hand side of the KL representation (3.1.6) as $\alpha \downarrow 0$.

REMARK 3.2.3. First we recall that in the case of $\alpha = 0$ the process $(X_t^{(0)})_{t\in[0,T)}$ is the standard Wiener process on $[0,T]$. If $\alpha \downarrow 0$, then the left-hand side of the KL representation (3.1.6) converges in $L^2(\Omega, \mathcal{A}, P)$ to the standard Wiener process $(B_t)_{t\in[0,T]}$, uniformly in $t \in [0,S]$ on every interval $[0,S] \subset [0,T)$. Indeed, for all $t \in [0,S]$,

$$\begin{aligned}\mathbb{E}(X_t^{(\alpha)} - B_t)^2 &= \mathbb{E}\left(\int_0^t \left(\left(\frac{T-t}{T-s}\right)^\alpha - 1\right) \mathrm{d}B_s\right)^2 \\ &= \int_0^t \left(\left(\frac{T-t}{T-s}\right)^\alpha - 1\right)^2 \mathrm{d}s \leqslant S\left(\left(\frac{T-S}{T}\right)^\alpha - 1\right)^2 \to 0 \quad \text{as } \alpha \downarrow 0,\end{aligned}$$

where the last inequality follows by $(T-S)/T \leqslant (T-t)/(T-s) \leqslant 1$, $0 \leqslant s \leqslant t \leqslant S < T$. Hence

$$\sup_{t\in[0,S]} \mathbb{E}(X_t^{(\alpha)} - B_t)^2 \to 0 \quad \text{as } \alpha \downarrow 0.$$

Hereafter we show that the coefficients of the random variables in the terms on the right-hand side of (3.1.6) also converge uniformly in $t \in [0,S]$ to those of the corresponding terms of the KL expansion of $(B_t)_{t\in[0,T]}$, based on $[0,T]$. For the latter KL representation see, e.g., Papoulis [**138**, Example 12.10] (which unfortunately contains a misprint). Indeed, exploiting the fact that the eigenfunctions are unique only up to sign, the KL expansion of $(B_t)_{t\in[0,T]}$, based on $[0,T]$, can be written in the form

$$B_t = \sum_{k=1}^\infty \eta_k\, (-1)^{k-1}\sqrt{2T}\,\frac{\sin\big((k-1/2)\pi t/T\big)}{(k-1/2)\pi} \tag{3.2.2}$$

for all $t \in [0,T]$, where $\eta_k,\ k \in \mathbb{N}$, are independent, standard normally distributed random variables. Moreover, using Theorem 3.2.1, parts (ii) and (vii) of Proposition 3.4.1 and that therefore $z_k^{(-1/2)} = (k-1/2)\pi,\ k \in \mathbb{N}$, we obtain

$$\begin{aligned}\lim_{\alpha\downarrow 0}\left(\sqrt{\lambda_k^{(\alpha)}}\, e_k^{(\alpha)}(t)\right) &= \lim_{\alpha\downarrow 0}\left(\frac{T}{z_k^{(\nu)}}\sqrt{\frac{2}{T}\left(1-\frac{t}{T}\right)}\,\frac{J_\nu\big(z_k^{(\nu)}(1-t/T)\big)}{\big|J_{\nu+1}\big(z_k^{(\nu)}\big)\big|}\right) \\ &= \frac{T}{z_k^{(-1/2)}}\sqrt{\frac{2}{T}\left(1-\frac{t}{T}\right)}\,\frac{J_{-1/2}\big(z_k^{(-1/2)}(1-t/T)\big)}{\big|J_{1/2}\big(z_k^{(-1/2)}\big)\big|} \\ &= \sqrt{2T}\,\frac{\cos\big((k-1/2)\pi(1-t/T)\big)}{(k-1/2)\pi} \\ &= (-1)^{k-1}\sqrt{2T}\,\frac{\sin\big((k-1/2)\pi t/T\big)}{(k-1/2)\pi}\end{aligned}$$

for each $k \in \mathbb{N}$. Further, the convergence is uniform in $t \in [0,S]$, since $\lim_{\alpha\downarrow 0} z_k^{(\nu)} = \lim_{\nu\downarrow -1/2} z_k^{(\nu)} = z_k^{(-1/2)} > 0$, and the function

$$\left[z_k^{(-1/2)}\left(1-\frac{S}{T}\right) - \varepsilon,\, z_k^{(-1/2)} + \varepsilon\right] \times \left[-\frac{1}{2}, -\frac{1}{2}+\varepsilon\right] \ni (x,\nu) \mapsto J_\nu(x)$$

is uniformly continuous (where $\varepsilon > 0$ is sufficiently small), since, by part (ii) of Proposition 3.4.1, $(0,\infty) \times (-1,\infty) \ni (x,\nu) \mapsto J_\nu(x)$ is an analytic and hence continuous function, $\left[z_k^{(-1/2)}\left(1-\frac{S}{T}\right) - \varepsilon, z_k^{(-1/2)} + \varepsilon\right] \times \left[-\frac{1}{2}, -\frac{1}{2}+\varepsilon\right]$ is a compact set and a continuous function on a compact set is uniformly continuous. □

In the next remark we consider the special case $\alpha = 1$ in Theorem 3.2.1.

REMARK 3.2.4. If $\alpha = 1$, i.e., $\nu = 1/2$, then by Theorem 3.2.1, part (vii) of Proposition 3.4.1 and that therefore $z_k^{(1/2)} = k\pi$, $k \in \mathbb{N}$, we obtain that the eigenvalue–normed eigenfunction pairs are $\lambda_k^{(1)} = T^2/(k\pi)^2$ and

$$\begin{aligned}
e_k^{(1)}(t) &= \sqrt{\frac{2}{T}\left(1-\frac{t}{T}\right)}\,\frac{J_{1/2}\big(k\pi(1-t/T)\big)}{\big|J_{-1/2}(k\pi)\big|}\\
&= \sqrt{\frac{2}{T}\left(1-\frac{t}{T}\right)}\,\frac{\sqrt{2T/\big(k\pi^2(T-t)\big)}\,\sin\big(k\pi(1-t/T)\big)}{\sqrt{2/\big(k\pi^2\big)}\,\big|\cos(k\pi)\big|}\\
&= (-1)^k\sqrt{\frac{2}{T}}\sin\left(\frac{k\pi t}{T}\right), \qquad t\in[0,T],\ k\in\mathbb{N}.
\end{aligned}$$

Further,

$$X_t^{(1)} = \sqrt{2T}\sum_{k=1}^{\infty}\xi_k(-1)^k\frac{\sin\big(k\pi t/T\big)}{k\pi} \stackrel{\mathcal{L}}{=} \sqrt{2T}\sum_{k=1}^{\infty}\xi_k\frac{\sin\big(k\pi t/T\big)}{k\pi}$$

for all $t \in [0,T]$, where $\stackrel{\mathcal{L}}{=}$ denotes equality in distribution. As we expected, this is the KL expansion of the Wiener bridge, see, e.g., Deheuvels [**55**, Remark 1.1] or Gutiérrez and Valderrama [**80**, formula (10)]. □

In the next remark we present a space-time transformed Wiener process representation of the α-Wiener bridge, needed further on. The idea comes from the similar representation of the Wiener bridge, see, e.g., Csörgő and Révész [**50**, Proposition 1.4.2], and from Barczy and Pap [**27**, proof of Lemma 3.1]. For historical fidelity we note that our representation (for the α-Wiener bridge) is an analogue of formula (20) in Brennan and Schwartz [**41**].

REMARK 3.2.5. Let $(W_u)_{u\geqslant 0}$ be a standard Wiener process,

$$\tau_T^{(\alpha)}(t) := \int_0^t \frac{1}{(T-s)^{2\alpha}}\,\mathrm{d}s = \frac{R^{(\alpha)}(t,t)}{(T-t)^{2\alpha}}, \qquad t\in[0,T) \tag{3.2.3}$$

and

$$Z_t^{(\alpha)} := (T-t)^{\alpha} W_{\tau_T^{(\alpha)}(t)}, \quad t\in[0,T). \tag{3.2.4}$$

Since $\tau_T^{(\alpha)}(0) = 0$ and $\tau_T^{(\alpha)}$ is strictly increasing and continuous, $(Z_t^{(\alpha)})_{t\in[0,T)}$ can be called a space-time transformed Wiener process. One can see at once that this is a Gaussian process with almost surely continuous sample paths, zero mean and the covariance function (3.1.3), therefore the process $(Z_t^{(\alpha)})_{t\in[0,T)}$ is a weak solution of the SDE (3.1.1). Since the SDE (3.1.1) has a strong solution which is pathwise unique, we get $(Z_t^{(\alpha)})_{t\in[0,T)}$ is an α-Wiener bridge (there exists some appropriate standard Wiener process for which (3.1.2) holds). □

In the following we deal with the weighted KL expansion of the α-Wiener bridge. The series expansion which we call the weighted KL expansion of a space-time transformed centered process with continuous covariance function was introduced by Gutiérrez and Valderrama [**80**]. Let $S \in (0,T)$ and $\mu_T^{(\alpha)}$ be a (weight) measure defined on (the Borel sets of) $[0,S]$ by the help of the space-time transform in (3.2.4) as

$$\mathrm{d}\mu_T^{(\alpha)}(s) := (T-s)^{-2\alpha}\mathrm{d}\tau_T^{(\alpha)}(s) = (T-s)^{-4\alpha}\mathrm{d}s$$

and let us denote by $L^2\big([0,S],\mu_T^{(\alpha)}\big)$ the Hilbert space of measurable functions on $[0,S]$, which are square integrable with respect to $\mu_T^{(\alpha)}$. Furthermore, let

$$W_u = \sum_{k=1}^{\infty} \sqrt{\kappa_k^{(\alpha)}}\,\xi_k d_k^{(\alpha)}(u), \quad u \in [0,\tau_T^{(\alpha)}(S)], \tag{3.2.5}$$

be the (unweighted) KL expansion of the standard Wiener process $(W_u)_{u\in[0,\tau_T^{(\alpha)}(S)]}$, based on $[0,\tau_T^{(\alpha)}(S)]$, i.e., $(\kappa_k^{(\alpha)}, d_k^{(\alpha)}),\ k\in\mathbb{N}$, are the eigenvalue–normed eigenfunction pairs of the integral operator associated to the covariance function of the standard Wiener process (for explicit formulae see (3.2.9) and (3.2.10) later on) and $\xi_k,\ k\in\mathbb{N}$, are independent standard normally distributed random variables. Finally, let

$$f_k^{(\alpha)}(t) := (T-t)^\alpha d_k^{(\alpha)}(\tau_T^{(\alpha)}(t)), \quad t\in[0,S],\ k\in\mathbb{N}, \tag{3.2.6}$$

i.e., we apply the same time change and rescaling to the normed eigenfunction $d_k^{(\alpha)}$ in order to define $f_k^{(\alpha)}$ what we apply to a standard Wiener process in order to get an α-Wiener bridge, see (3.2.4). Using (3.2.4) and Gutiérrez and Valderrama [**80**, Theorem 1], the weighted KL expansion of the α-Wiener bridge $(X_t^{(\alpha)})_{t\in[0,T]}$, based on $[0,S]$, with respect to the weight measure $\mu_T^{(\alpha)}$ is

$$X_t^{(\alpha)} = \sum_{k=1}^{\infty} \sqrt{\kappa_k^{(\alpha)}}\,\xi_k f_k^{(\alpha)}(t), \quad t\in[0,S]. \tag{3.2.7}$$

It also follows that the properties of the weighted KL expansion (3.2.7) and the weighted normed eigenfunctions (3.2.6) are completely analogous to those of (3.2.5) and the unweighted normed eigenfunctions therein. The difference is in the measure with respect to which we integrate. Namely, in the weighted case we integrate with respect to a Lebesgue–Stieltjes measure and there is the $L^2\big([0,S],\mu_T^{(\alpha)}\big)$ space in the background, instead of the Lebesgue measure and the $L^2([0,\tau_T^{(\alpha)}(S)])$ space in the unweighted case. So, the series in (3.2.7) is convergent in $L^2(\Omega,\mathcal{A},P)$ uniformly in $t\in[0,S]$ and

$$\begin{aligned} &\int_0^S R^{(\alpha)}(t,s) f_k^{(\alpha)}(s)\,\mathrm{d}\mu_T^{(\alpha)}(s) = \kappa_k^{(\alpha)} f_k^{(\alpha)}(t), \qquad t\in[0,S],\ k\in\mathbb{N},\\ &\int_0^S (f_k^{(\alpha)}(s))^2\,\mathrm{d}\mu_T^{(\alpha)}(s) = 1, \quad \int_0^S f_k^{(\alpha)}(s) f_\ell^{(\alpha)}(s)\,\mathrm{d}\mu_T^{(\alpha)}(s) = 0, \quad k\neq\ell,\ k,\ell\in\mathbb{N}. \end{aligned} \tag{3.2.8}$$

By Papoulis [**138**, Example 12.10] (which unfortunately contains a misprint), we have

$$\kappa_k^{(\alpha)} = \left(\frac{\tau_T^{(\alpha)}(S)}{(k-1/2)\pi}\right)^2, \quad k\in\mathbb{N}, \tag{3.2.9}$$

$$d_k^{(\alpha)}(u) = \sqrt{\frac{2}{\tau_T^{(\alpha)}(S)}}\,\sin\left(\left(k-\frac{1}{2}\right)\pi\frac{u}{\tau_T^{(\alpha)}(S)}\right), \quad u\in[0,\tau_T^{(\alpha)}(S)],\ k\in\mathbb{N}, \tag{3.2.10}$$

and then using (3.2.6–3.2.7) we obtain the following theorem.

THEOREM 3.2.6. *In the weighted KL expansion* (3.2.7) *of the α-Wiener bridge the weighted eigenvalues $\kappa_k^{(\alpha)}$, $k\in\mathbb{N}$, are given by* (3.2.9) *and the corresponding* *weighted normed eigenfunctions*

(3.2.11)

$$f_k^{(\alpha)}(t) = \sqrt{\frac{2}{\tau_T^{(\alpha)}(S)}}\,(T-t)^\alpha \sin\left(\left(k-\frac{1}{2}\right)\pi\,\frac{\tau_T^{(\alpha)}(t)}{\tau_T^{(\alpha)}(S)}\right), \quad t\in[0,S],\ \ k\in\mathbb{N}.$$

Hence

$$X_t^{(\alpha)} = \sum_{k=1}^{\infty} \sqrt{2\tau_T^{(\alpha)}(S)}(T-t)^\alpha \frac{\sin\left(\left(k-\frac{1}{2}\right)\pi\frac{\tau_T^{(\alpha)}(t)}{\tau_T^{(\alpha)}(S)}\right)}{(k-1/2)\pi}\xi_k, \qquad t\in[0,S].$$

In the next remark we will study the convergence of the coefficients of the random variables in the terms on the right-hand side of (3.2.7) as $\alpha\downarrow 0$.

Remark 3.2.7. If $\alpha\downarrow 0$, then the left-hand side of the weighted KL representation (3.2.7) converges in $L^2(\Omega,\mathcal{A},P)$ to B_t uniformly in $t\in[0,S]$, see the beginning of Remark 3.2.3. Hereafter we show that the coefficients of the random variables in the terms on the right-hand side of (3.2.7) (given by the help of (3.2.3), (3.2.9) and (3.2.11)) converge uniformly in $t\in[0,S]$ to the coefficients of the corresponding terms of the (unweighted) KL expansion of the standard Wiener process, based on $[0,S]$. Indeed, we have

$$\lim_{\alpha\downarrow 0}\sqrt{2\tau_T^{(\alpha)}(S)}\,(T-t)^\alpha = \sqrt{2S},$$

uniformly in $t\in[0,S]$ (the uniform convergence follows by mean value theorem), and

$$\lim_{\alpha\downarrow 0}\frac{\tau_T^{(\alpha)}(t)}{\tau_T^{(\alpha)}(S)} = \frac{t}{S}\,,$$

also uniformly in $t\in[0,S]$. Hence for all $k\in\mathbb{N}$,

$$\begin{aligned}\lim_{\alpha\downarrow 0}\sqrt{\kappa_k^{(\alpha)}}\,f_k^{(\alpha)}(t) &= \lim_{\alpha\downarrow 0}\sqrt{2\tau_T^{(\alpha)}(S)}\,(T-t)^\alpha\frac{\sin\big((k-1/2)\pi\tau_T^{(\alpha)}(t)/\tau_T^{(\alpha)}(S)\big)}{(k-1/2)\pi}\\ &= \sqrt{2S}\,\frac{\sin\big((k-1/2)\pi t/S\big)}{(k-1/2)\pi},\end{aligned}$$

uniformly in $t\in[0,S]$. Indeed,

$$\left|\sin\left(\left(k-\frac{1}{2}\right)\pi\frac{\tau_T^{(\alpha)}(t)}{\tau_T^{(\alpha)}(S)}\right) - \sin\left(\left(k-\frac{1}{2}\right)\pi\frac{t}{S}\right)\right| \leqslant \left(k-\frac{1}{2}\right)\pi\left|\frac{\tau_T^{(\alpha)}(t)}{\tau_T^{(\alpha)}(S)} - \frac{t}{S}\right|.$$

□

In the next remark we consider the special case $\alpha=1$ in Theorem 3.2.6.

Remark 3.2.8. If $\alpha=1$, then $\tau_T^{(\alpha)}$ (given by (3.2.3)) takes the form $\tau_T^{(1)}(t) = t/(T(T-t))$, $t\in[0,T)$, so (3.2.9) takes the form

$$\kappa_k^{(1)} = \left(\frac{S}{T(T-S)}\right)^2\frac{1}{(k-1/2)^2\pi^2}, \quad k\in\mathbb{N},$$

and (3.2.11) becomes

$$f_k^{(1)}(t) = \sqrt{\frac{2T(T-S)}{S}}\,(T-t)\sin\left(\left(k-\frac{1}{2}\right)\pi\,\frac{t(T-S)}{S(T-t)}\right), \quad t\in[0,S],\ \ k\in\mathbb{N}.$$

Particularly, for $T=1$ and $S=1/2$ we reobtain the weighted KL expansion

$$X_t^{(1)} = \sqrt{2}\,(1-t)\sum_{k=1}^{\infty}\xi_k\frac{\sin\big((k-1/2)\pi t/(1-t)\big)}{(k-1/2)\pi}, \quad t\in[0,1/2],$$

given by Gutiérrez and Valderrama [**80**, formula (12)]. □

In the next remark we formulate a corollary of Theorem 3.2.6 in the case of $\alpha = 1/2$.

REMARK 3.2.9. For all $0 < S < T$, we have

$$\int_0^S \frac{(X_u^{(1/2)})^2}{(T-u)^2}\,\mathrm{d}u = \left(\ln\left(\frac{T}{T-S}\right)\right)^2 \sum_{k=1}^\infty \frac{1}{(k-1/2)^2\pi^2}\xi_k^2, \tag{3.2.12}$$

where ξ_k, $k \in \mathbb{N}$, are independent standard normally distributed random variables. Indeed, by Theorem 3.2.6 and the Parseval identity in $L^2([0,S], \mu_T^{(1/2)})$, we get

$$\int_0^S \frac{(X_u^{(1/2)})^2}{(T-u)^2}\,\mathrm{d}u = \sum_{k=1}^\infty \kappa_k^{(1/2)}\xi_k^2 = \left(\ln\left(\frac{T}{T-S}\right)\right)^2 \sum_{k=1}^\infty \frac{1}{(k-1/2)^2\pi^2}\xi_k^2,$$

where the last equality follows by

$$\tau_T^{(1/2)}(t) = \int_0^t \frac{1}{T-u}\,\mathrm{d}u = \ln\left(\frac{T}{T-t}\right), \qquad t \in [0,T).$$

□

3.3. Applications

In this section we present some applications of the KL expansion (3.1.6) given in Theorem 3.2.1. First we calculate the Laplace transform of the $L^2[0,T]$-norm square of $(X_t^{(\alpha)})_{t\in[0,T]}$.

PROPOSITION 3.3.1. *Let $T > 0$, $\alpha > 0$ and $\nu := \alpha - 1/2$. Then*

$$\mathbb{E}\exp\left\{-c\int_0^T (X_u^{(\alpha)})^2\,\mathrm{d}u\right\} = \prod_{k=1}^\infty \frac{1}{\sqrt{1+2cT^2/(z_k^{(\nu)})^2}}, \qquad c \geqslant 0, \tag{3.3.1}$$

$$\mathbb{E}\exp\left\{-c\int_0^1 (X_u^{(0)})^2\,\mathrm{d}u\right\} = \frac{1}{\sqrt{\cosh(\sqrt{2c})}}, \qquad c \geqslant 0, \tag{3.3.2}$$

$$\mathbb{E}\exp\left\{-c\int_0^1 (X_u^{(1)})^2\,\mathrm{d}u\right\} = \sqrt{\frac{\sqrt{2c}}{\sinh(\sqrt{2c})}}, \qquad c > 0. \tag{3.3.3}$$

Further, for all $0 < S < T$,

$$\mathbb{E}\exp\left\{-c\int_0^S \frac{(X_u^{(1/2)})^2}{(T-u)^2}\,\mathrm{d}u\right\} = \frac{1}{\sqrt{\cosh\left(\sqrt{2c}\ln\left(\frac{T}{T-S}\right)\right)}}, \qquad c \geqslant 0. \tag{3.3.4}$$

We remark that a corresponding version of (3.3.4) for general α-Wiener bridges can be proved by a different technique, see Barczy and Pap [**28**, Theorem 4.1].

Next we give a simple probabilistic proof for the sum of the square of the reciprocals of the positive zeros of J_ν with $\nu > -1/2$. For $\nu > -1$ this is a well-known result due to Rayleigh, see, e.g., Watson [**152**, Section 15.51, p. 502]. We note that Yor [**156**, (11.47)–(11.49)] and Deheuvels and Martynov [**57**, Corollary 1.3] also gave probabilistic proofs of Rayleigh's results; we show that the proof of Deheuvels and Martynov can be carried through starting from the Karhunen–Loève expansion of the α-Wiener bridge as well.

COROLLARY 3.3.2. *Let $\alpha > 0$, $\nu := \alpha - 1/2$ and $z_k^{(\nu)}$, $k \in \mathbb{N}$, be the positive zeros of J_ν. Then*

$$\sum_{k=1}^{\infty} \frac{1}{(z_k^{(\nu)})^2} = \frac{1}{4(\nu+1)}.$$

A consequence of (3.3.1) is the following form of the survival function (complementary distribution function) of the $L^2[0,T]$-norm square of the α-Wiener bridge.

PROPOSITION 3.3.3. *Let $\alpha > 0$ and $\nu := \alpha - 1/2$. Then*

(3.3.5)

$$\mathbb{P}\left(\int_0^T (X_t^{(\alpha)})^2 \,\mathrm{d}t > x\right) = \frac{2^{1-\nu/2}}{\pi\sqrt{\Gamma(\nu+1)}} \sum_{k=1}^{\infty} (-1)^{k+1} \int_{z_{2k-1}^{(\nu)}}^{z_{2k}^{(\nu)}} u^{\nu/2-1} \frac{\mathrm{e}^{-xu^2/(2T^2)}}{\sqrt{|J_\nu(u)|}} \,\mathrm{d}u$$

for all $x > 0$.

In the next remark we check that the formula for the survival function of the $L^2[0,1]$-norm square of a standard Wiener process (see, e.g., Deheuvels and Martynov [**57**, formula (1.50)]) can be derived by taking the limit of (3.3.5) with $T = 1$ as $\alpha \downarrow 0$.

REMARK 3.3.4. According to Deheuvels and Martynov [**57**, formula (1.50)] the survival function of the $L^2[0,1]$-norm square of a standard Wiener process is

$$\mathbb{P}\left(\int_0^1 B_t^2 \,\mathrm{d}t > x\right) = \frac{2}{\pi} \sum_{k=1}^{\infty} (-1)^{k+1} \int_{(2k-3/2)\pi}^{(2k-1/2)\pi} \frac{\mathrm{e}^{-xu^2/2}}{u\sqrt{-\cos(u)}} \,\mathrm{d}u \tag{3.3.6}$$

for all $x > 0$. The right-hand side of (3.3.6) is continuous in $x \in (0, \infty)$, which can be derived using Lebesgue's dominated convergence theorem. Indeed,

$$\begin{aligned}
\int_{(2k-3/2)\pi}^{(2k-1/2)\pi} \frac{\mathrm{e}^{-xu^2/2}}{u\sqrt{-\cos(u)}} \,\mathrm{d}u &\leqslant \frac{\mathrm{e}^{-x(2k-3/2)^2\pi^2/2}}{(2k-3/2)\pi} \int_{(2k-3/2)\pi}^{(2k-1/2)\pi} \frac{1}{\sqrt{-\cos(u)}} \,\mathrm{d}u \\
&= \frac{\mathrm{e}^{-x(2k-3/2)^2\pi^2/2}}{(2k-3/2)\pi} \int_{\pi/2}^{3\pi/2} \frac{1}{\sqrt{-\cos(u)}} \,\mathrm{d}u, \quad k \in \mathbb{N},
\end{aligned}$$

and, by D'Alembert's criterion, for all $x > 0$,

$$\sum_{k=1}^{\infty} \frac{\mathrm{e}^{-x(2k-3/2)^2\pi^2/2}}{(2k-3/2)\pi} < \infty.$$

Then the left-hand side of (3.3.6) is also continuous in $x \in (0, \infty)$. Using the continuity of probability and that the $L^2[0,1]$-norm square of a standard Wiener process takes the value zero with probability 0, we have that the left-hand side of (3.3.6) is continuous in $x = 0$ too, and

$$\lim_{x\downarrow 0} \mathbb{P}\left(\int_0^1 B_t^2 \,\mathrm{d}t > x\right) = \mathbb{P}\left(\int_0^1 B_t^2 \,\mathrm{d}t > 0\right) = 1.$$

Hence $\mathbb{P}\left(\int_0^1 B_t^2 \mathrm{d}t > x\right)$ is continuous at every $x \in \mathbb{R}$. Using that $\int_0^1 (X_t^{(\alpha)})^2 \mathrm{d}t$ converges in distribution to $\int_0^1 B_t^2 \mathrm{d}t$ as $\alpha \downarrow 0$ (which was verified in the proof of Proposition 3.3.1), we get

$$\lim_{\alpha\downarrow 0} \mathbb{P}\left(\int_0^1 (X_t^{(\alpha)})^2 \,\mathrm{d}t > x\right) = \mathbb{P}\left(\int_0^1 B_t^2 \,\mathrm{d}t > x\right)$$

for all $x \in \mathbb{R}$. Therefore the right-hand side of (3.3.5) must also converge to the right-hand side of (3.3.6) for every $x > 0$, i.e., the survival function of the $L^2[0,1]$-norm square of a standard Wiener process is the limit of the survival function of the $L^2[0,1]$-norm square of the α-Wiener bridge (on the time interval $[0,1]$) as $\alpha \downarrow 0$. □

In the next remark we consider the Proposition 3.3.3 with the special choices $\alpha = 1$ and $T = 1$.

Remark 3.3.5. With the special choices $\alpha = 1$, i.e., $\nu = 1/2$ and $T = 1$ in Proposition 3.3.3 we have for all $x > 0$,

$$\mathbb{P}\left(\int_0^1 \big(X_t^{(1)}\big)^2 \,\mathrm{d}t > x\right) = \frac{2^{1-1/4}}{\pi\sqrt{\Gamma(3/2)}} \sum_{k=1}^{\infty} (-1)^{k+1} \int_{z_{2k-1}^{(1/2)}}^{z_{2k}^{(1/2)}} u^{1/4-1} \frac{\mathrm{e}^{-xu^2/2}}{\sqrt{\big|J_{1/2}(u)\big|}} \,\mathrm{d}u$$

$$= \frac{2^{3/4}}{\pi\sqrt{\sqrt{\pi}/2}} \sum_{k=1}^{\infty} (-1)^{k+1} \int_{(2k-1)\pi}^{2k\pi} u^{-3/4} \frac{\mathrm{e}^{-xu^2/2}}{\sqrt{-\sqrt{2}\,\sin(u)/\sqrt{\pi u}}} \,\mathrm{d}u$$

$$= \frac{2}{\pi} \sum_{k=1}^{\infty} (-1)^{k+1} \int_{(2k-1)\pi}^{2k\pi} \frac{\mathrm{e}^{-xu^2/2}}{\sqrt{-u\sin(u)}} \,\mathrm{d}u\,,$$

where we used part (vii) of Proposition 3.4.1. We reobtained the survival function of the $L^2[0,1]$-norm square of the Wiener bridge, see Deheuvels and Martynov [**57**, formula (1.51)]. □

Remark 3.3.6. We note that Deheuvels and Martynov [**57**, formula (1.43)] gave an expression for the survival function of the $L^2[0,1]$-norm square of a particularly weighted time transformed Wiener bridge, namely $(t^{1/2-\nu}X^{(1)}_{t^{2\nu}})_{t\in[0,1]}$, where $(X_t^{(1)})_{t\in[0,1]}$ is a Wiener bridge on the time interval $[0,1]$ and $\nu > 0$. We can notice that the only difference between that formula and our formula (3.3.5) is the denominator of the fraction in the argument of the exponential function, namely instead of 4ν we have $2T^2$. This means that the distribution of the $L^2[0,1]$-norm square of the above mentioned particularly weighted time transformed Wiener bridge is the same as the distribution of the $L^2[0,T]$-norm square of an appropriate α-Wiener bridge. Namely, in case of $\alpha > 1/2$, i.e., $\nu > 0$, the random variables

$$\int_0^1 t^{1-2\nu}(X^{(1)}_{t^{2\nu}})^2\,\mathrm{d}t = \int_0^1 t^{2(1-\alpha)}(X^{(1)}_{t^{2\alpha-1}})^2\,\mathrm{d}t \qquad \text{and} \qquad \int_0^{\sqrt{2\alpha-1}} (X_t^{(\alpha)})^2\,\mathrm{d}t$$

have the same distribution, where $(X_t^{(\alpha)})_{t\in[0,\sqrt{2\alpha-1}]}$ is an α-Wiener bridge on the time interval $[0,\sqrt{2\alpha-1}]$. □

Zolotarev [**159**, formula (6)] gives the tail behaviour of the distribution function of $\sum_{k=1}^{\infty}\lambda_k\xi_k^2$, where ξ_k, $k\in\mathbb{N}$, are independent standard normally distributed random variables and $(\lambda_k)_{k\in\mathbb{N}}$ is a sequence of positive real numbers such that $\lambda_1 > \lambda_2 > \cdots > 0$ and $\sum_{k=1}^{\infty}\lambda_k < \infty$ (see also Hwang [**83**, Theorem 1] or Deheuvels and Martynov [**57**, Lemma 1.1 and Remark 1.2]). This result can be directly applied together with Theorem 3.2.1 to obtain the following corollary about the large deviation probabilities for the $L^2[0,T]$-norm square of the α-Wiener bridge.

Corollary 3.3.7. *Let $\alpha > 0$, $\nu := \alpha - 1/2$ and $z_k^{(\nu)}$, $k\in\mathbb{N}$, be the positive zeros of J_ν. Then*

(3.3.7)
$$\mathbb{P}\Big(\int_0^T (X_t^{(\alpha)})^2\,\mathrm{d}t > x\Big) = \big(1+o(1)\big)\frac{2^{1-\nu/2}T\big(z_1^{(\nu)}\big)^{(\nu-3)/2}}{\sqrt{\pi\,\Gamma(\nu+1)J_{\nu+1}\big(z_1^{(\nu)}\big)}}\; x^{-1/2}\,\mathrm{e}^{-(z_1^{(\nu)})^2x/(2T^2)}$$

$$= \big(1+o(1)\big)\sqrt{\frac{2}{\pi}}\,\frac{T}{z_1^{(\nu)}}\prod_{k=2}^{\infty}\left(1-\frac{(z_1^{(\nu)})^2}{(z_k^{(\nu)})^2}\right)^{-1/2} x^{-1/2}\,\mathrm{e}^{-(z_1^{(\nu)})^2x/(2T^2)} \tag{3.3.8}$$

as $x\to\infty$.

The next corollary describes the behaviour at zero of the distribution function (small deviation probabilities) of the $L^2[0,T]$-norm square of the α-Wiener bridge.

COROLLARY 3.3.8. *Let* $\alpha > 0$ *and* $\nu := \alpha - 1/2$. *Then there exists some constant* $c > 0$ *such that*

$$\mathbb{P}\left(\int_0^T (X_t^{(\alpha)})^2 \,\mathrm{d}t < \varepsilon\right) = \big(c + o(1)\big)\varepsilon^{1/4-\nu/2}\mathrm{e}^{-T^2/(8\varepsilon)} \tag{3.3.9}$$

as $\varepsilon \downarrow 0$.

REMARK 3.3.9. In case of $\alpha \geqslant 1/2$, Corollary 3.3.8 can be improved by which we mean that the constant c can be explicitly given. Namely, by Nazarov [**131**, Lemma 3.2], if ξ_k, $k \in \mathbb{N}$, are independent standard normally distributed random variables, then for all $\nu \geqslant 0$,

$$\mathbb{P}\left(\sum_{n=1}^{\infty} \frac{\xi_n^2}{(z_n^{(\nu)})^2} \leqslant \varepsilon^2\right) \sim \sqrt{\frac{\sqrt{\pi}}{2^{\nu-1/2}\Gamma(1+\nu)}} \frac{\sqrt{2}\varepsilon_1^{1/2-\nu}}{\sqrt{\pi \mathcal{D}_1}} \exp\left\{-\frac{\mathcal{D}_1}{2\varepsilon_1^2}\right\},$$

as $\varepsilon \to 0$, where

$$\varepsilon_1 = \varepsilon\sqrt{2\sin(\pi/2)} = \sqrt{2}\varepsilon \qquad \text{and} \qquad \mathcal{D}_1 = \frac{1}{2\sin(\pi/2)} = 1/2.$$

Hence

$$\mathbb{P}\left(\sum_{n=1}^{\infty} \frac{\xi_n^2}{(z_n^{(\nu)})^2} \leqslant \varepsilon^2\right) \sim \frac{2^{\frac{3}{2}-\nu}\pi^{-1/4}}{\sqrt{\Gamma(1+\nu)}}\varepsilon^{1/2-\nu}\mathrm{e}^{-\frac{1}{8\varepsilon^2}}, \qquad \text{as } \varepsilon \downarrow 0.$$

Then for all $T > 0$ and $\varepsilon > 0$ we have

$$\mathbb{P}\left(T^2\sum_{n=1}^{\infty} \frac{\xi_n^2}{(z_n^{(\nu)})^2} \leqslant \varepsilon\right) \sim \frac{2^{\frac{3}{2}-\nu}\pi^{-1/4}}{\sqrt{\Gamma(1+\nu)}}\frac{\varepsilon^{\frac{1}{4}-\frac{\nu}{2}}}{T^{1/2-\nu}}\mathrm{e}^{-\frac{T^2}{8\varepsilon}}, \qquad \text{as } \varepsilon \downarrow 0.$$

By the proof of Corollary 3.3.8, this yields that in case of $\alpha \geqslant 1/2$ the constant c in Corollary 3.3.8 takes the following form

$$c = \frac{2^{\frac{3}{2}-\nu}\pi^{-1/4}}{\sqrt{\Gamma(1+\nu)}T^{1/2-\nu}}.$$

The reason for restricting ourselves to the case $\alpha \geqslant 1/2$ is that Lemma 3.2 in Nazarov [**131**] is valid for $\nu \geqslant 0$, while in Corollary 3.3.8 we have $\nu = \alpha - 1/2$, $\alpha > 0$. □

3.4. Appendix: Some properties of Bessel functions

In the next proposition we list some properties of the Bessel functions J_ν of the first kind (introduced in Section 3.2).

PROPOSITION 3.4.1.
- (i) *For all* $\nu \in \mathbb{R}$, J_ν *is continuous on* $(0,\infty)$, *and in case of* $\nu \geqslant 0$ *it can be continuously extended to* $[0,\infty)$ *by* $J_\nu(0) := 0$ *if* $\nu > 0$ *and by* $J_\nu(0) := 1$ *if* $\nu = 0$. *However, in case of* $-1/2 < \nu < 0$ *we have* $\lim_{x\downarrow 0} J_\nu(x) = \infty$, *and* $\lim_{x\downarrow 0}(\sqrt{x}J_\nu(x)) = 0$ *holds for all* $\nu > -1/2$.
- (ii) *By Watson* [**152**, p. 44], $(0,\infty)\times(-1,\infty) \ni (x,\nu) \mapsto J_\nu(x)$ *is an analytic function of both variables.*
- (iii) *By the Bessel–Lommel theorem, see, e.g., Watson* [**152**, pp. 478, 482], *if* $\nu > -1$, *then* J_ν *has infinitely many positive real zeros with multiplicities one. We denote them by* $z_k^{(\nu)}$, $k \in \mathbb{N}$, *where we assume*

$$0 < z_1^{(\nu)} < z_2^{(\nu)} < \cdots < z_k^{(\nu)} < z_{k+1}^{(\nu)} < \cdots .$$

By Watson [**152**, p. 508] *or Korenev* [**109**, p. 96], *for fixed* $k \in \mathbb{N}$, $z_k^{(\nu)}$ *is a strictly increasing and continuous function of* $\nu \in (-1, \infty)$.

(iv) *By Korenev* [**109**, p. 96], *if* $\nu > -1$, *then*

$$z_k^{(\nu)} = \left(k + \frac{1}{2}\left(\nu - \frac{1}{2}\right)\right)\pi + O\left(\frac{1}{k}\right) \qquad \text{as } k \to \infty.$$

(v) *If* $\nu > -1$, *then* J_ν *can be written by the help of its zeros* $z_k^{(\nu)}$, $k \in \mathbb{N}$, *as*

$$J_\nu(x) = \frac{\left(\frac{x}{2}\right)^\nu}{\Gamma(\nu+1)} \prod_{k=1}^{\infty} \left(1 - \frac{x^2}{(z_k^{(\nu)})^2}\right), \quad x \in (0, \infty).$$

This is the so called Euler formula, see Watson [**152**, p. 498].

(vi) *By Bowman* [**40**, p. 107], *for* $\nu > -1$ *and for all zeros* $z_k^{(\nu)}$, $k \in \mathbb{N}$, *of* J_ν, *it holds that*

$$\int_0^{z_k^{(\nu)}} x J_\nu^2(x)\,\mathrm{d}x = \frac{\left(z_k^{(\nu)}\right)^2}{2} J_{\nu+1}^2\left(z_k^{(\nu)}\right) = \frac{\left(z_k^{(\nu)}\right)^2}{2} J_{\nu-1}^2\left(z_k^{(\nu)}\right).$$

(vii) *The Bessel functions* $J_{1/2}$ *and* $J_{-1/2}$ *(of the first kind) can be expressed in closed forms by elementary functions:*

$$J_{1/2}(x) = \sqrt{\frac{2}{\pi x}} \sin(x) \qquad \text{and} \qquad J_{-1/2}(x) = \sqrt{\frac{2}{\pi x}} \cos(x), \qquad x \in (0, \infty),$$

see, e.g., Watson [**152**, pp. 54, 55].

CHAPTER 4

Operator scaled Wiener bridges

Co-authors: Peter Kern and Vincent Krause

4.1. Introduction

This paper deals with a multidimensional generalization of the so-called α-Wiener bridges also known as scaled Wiener bridges. For fixed $T > 0$ and given matrices $A \in \mathbb{R}^{d\times d}$ and $\Sigma \in \mathbb{R}^{d\times m}$, a d-dimensional process $(X_t)_{t\in[0,T)}$ is given by the SDE

$$\mathrm{d}X_t = -\frac{1}{T-t} A X_t \,\mathrm{d}t + \Sigma \,\mathrm{d}B_t, \quad t \in [0,T), \tag{4.1.1}$$

with initial condition $X_0 = 0 \in \mathbb{R}^d$, where $(B_t)_{t\in[0,T)}$ is an m-dimensional standard Wiener process defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T)}, P)$ with the completion $(\mathcal{F}_t)_{t\in[0,T)}$ of the canonical filtration of $(B_t)_{t\in[0,T)}$. Note that in case $m = d$ and if A and Σ are both the $d \times d$ identity matrix, then the process $(X)_{t\in[0,T)}$ is nothing else but the usual d-dimensional Wiener bridge over $[0,T]$.

To our knowledge, in case of dimension $d = 1$, these kinds of processes have been first considered by Brennan and Schwartz [**41**]; see also Mansuy [**125**]. In Brennan and Schwartz [**41**] α-Wiener bridges, where $A = \alpha \in \mathbb{R}^{1\times 1}$ with $\alpha > 0$, are used to model the arbitrage profit associated with a given futures contract in the absence of transaction costs. This model is also meaningful in a multidimensional context when a finite number of contracts is considered with possible dependencies between the contracts. Operator scaled Wiener bridges offer a tool for modeling the arbitrage profit in this multidimensional setting.

Sondermann, Trede and Wilfling [**147**] and Trede and Wilfling [**149**] used α-Wiener bridges with $\alpha > 0$ to describe the fundamental component of an exchange rate process and they call the process a scaled Brownian bridge. The essence of these models is that the coefficient $-\alpha/(T-t)$ of X_t in the drift term in (4.1.1) represents some kind of mean reversion, a stabilizing force that keeps pulling the process towards its mean 0, and the absolute value of this force is increasing proportionally to the inverse of the remaining time $T-t$, with the constant rate α. This model is used in [**149**] to analyze the exchange rate of the Greek drachma to the Euro before the Greek EMU entrance on 1 January 2001 with a priorly fixed conversion rate. Trede and Wilfling [**149**] observe an increase in interventions towards the fixed conversion rate, well described by an α-Wiener bridge plus deterministic drift with MLE-estimator $\widehat{\alpha} = 1.24$. If more than two countries join the EMU at the same time, most recently Cyprus and Malta on 1 January 2008, operator scaled Wiener bridges may offer a useful tool to analyze interventions for all the exchange rates, commonly. In this context the replacement of a constant rate α by some scaling matrix A is meaningful, since the economies of EU countries are tightly linked and thus interventions are likely to be strongly dependent on each other.

The SDE (4.1.1) with initial condition $X_0 = 0$ has a unique strong solution $(X_t)_{t\in[0,T)}$ given by the d-dimensional integral representation

$$X_t = \int_0^t \left(\frac{T-t}{T-s}\right)^A \Sigma \, \mathrm{d}B_s \quad \text{for } t \in [0,T), \tag{4.1.2}$$

where r^A is defined by the exponential operator

$$r^A = \mathrm{e}^{A \log r} = \sum_{k=0}^{\infty} \frac{(\log r)^k}{k!} A^k \quad \text{for } r > 0. \tag{4.1.3}$$

The validity of (4.1.2) can be easily checked using Itô's formula and properties of the exponential operator. Indeed,

$$\begin{aligned}
\mathrm{d}X_t &= \left(\left(\frac{\mathrm{d}}{\mathrm{d}t}(T-t)^A \right) \int_0^t (T-s)^{-A} \Sigma \, \mathrm{d}B_s \right) \mathrm{d}t + (T-t)^A (T-t)^{-A} \Sigma \, \mathrm{d}B_t \\
&= \left(\left(-A(T-t)^{A-I_d} \right) \int_0^t (T-s)^{-A} \Sigma \, \mathrm{d}B_s \right) \mathrm{d}t + \Sigma \, \mathrm{d}B_t \\
&= -\frac{1}{T-t} A X_t \, \mathrm{d}t + \Sigma \, \mathrm{d}B_t, \quad t \in [0,T),
\end{aligned}$$

where I_d denotes the $d \times d$ identity matrix. Further, by Section 5.6 in Karatzas and Shreve [**100**], strong uniqueness holds for the SDE (4.1.1). Note also that $(X_t)_{t\in[0,T)}$ is a Gauss process with almost surely continuous sample paths, see, e.g., Problem 5.6.2 in Karatzas and Shreve [**100**]. Later on, we will frequently assume that Σ has rank d (and consequently $m \geqslant d$), but the assumption will always be stated explicitly. Note that this is only a minor restriction, since otherwise the d-dimensional Gaussian driving process $(\Sigma B_t)_{t\in[0,T)}$ in (4.1.2) has linearly dependent coordinates.

The paper is organized as follows. In Section 4.2 we recall a spectral decomposition of the matrix A and of the process X, respectively. We further present a result on the growth behavior of the exponential operator t^A near the origin, and we also recall a strong law of large numbers and a law of the iterated logarithm valid for the martingale $((T-t)^{-A} X_t)_{t\in[0,T)}$. In Section 4.3, in order to properly speak of a process bridge, we derive some sufficient conditions on A and Σ such that X_t converges to the origin almost surely as $t \uparrow T$, see Theorem 4.3.4. Provided that the conditions of Theorem 4.3.4 hold, we will call the process $(X_t)_{t\in[0,T]}$ an *operator scaled Wiener bridge* associated to the matrices A and Σ over the time interval $[0,T]$. By giving an example, we point out that if the conditions of Theorem 4.3.4 do not hold, then in general one cannot expect that X_t converges to some deterministic d-dimensional vector almost surely as $t \uparrow T$. Section 4.4 is devoted to study the asymptotic behavior of the sample paths of operator scaled Wiener bridges as $t \uparrow T$. Finally, in Section 4.5 we address the question of uniqueness of bridges. By giving examples, we point out that there exist matrices A, $\widetilde{A} \in \mathbb{R}^{d\times d}$ and $\Sigma \in \mathbb{R}^{d\times m}$ such that the laws of the bridges associated to the matrices A and Σ, and $\widetilde{A}$ and Σ coincide, but $A \neq \widetilde{A}$. We also formulate a partial result on the uniqueness of bridges in terms of the spectrum of A, see Proposition 4.5.2.

4.2. Preliminaries

4.2.1. Spectral decomposition. Factor the minimal polynomial f of A into $f(\lambda) = f_1(\lambda) \cdots f_p(\lambda)$, $\lambda \in \mathsf{C}$, with $p \leqslant d$ such that every root of f_j has real part a_j, where $a_1 < \cdots < a_p$ denote the distinct real parts of the eigenvalues of A. Note that $f, f_1, \ldots, f_p$ are polynomials with real coefficients. According to the primary decomposition theorem of linear algebra we can decompose $\mathbb{R}^d$ into a direct sum $\mathbb{R}^d = V_1 \oplus \cdots \oplus V_p$, where each $V_j := \mathrm{Ker}(f_j(A))$ is an A-invariant subspace. Let us denote the dimension of V_j by d_j, $j = 1, \ldots, p$. Now, in an appropriate basis, say $\{b_i^{(j)} : i = 1, \ldots, d_j,\, j = 1, \ldots, p\}$, A can be represented as a block-diagonal matrix $A = A_1 \oplus \cdots \oplus A_p$, where every eigenvalue of A_j has real part a_j. For this reason, we will call each matrix A_j *real spectrally simple*, i.e., all its eigenvalues have the same real part. We can further choose a unique inner product $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^d$ such that the basis $\{b_i^{(j)} : i = 1, \ldots, d_j,\, j = 1, \ldots, p\}$ is orthonormal, and consequently, the subspaces V_j, $1 \leqslant j \leqslant p$, are mutually orthogonal. For $x = x_1 + \cdots + x_p$ with $x_j \in V_j$, $j = 1, \ldots, p$, let $\pi_j(x)$ be the coordinates of x_j with respect to the basis $\{b_i^{(j)} : i = 1, \ldots, d_j\}$ of V_j. Then $\pi_j : \mathbb{R}^d \to \mathbb{R}^{d_j}$ is a linear projection mapping. To conclude, for every $x \in \mathbb{R}^d$ there exist unique $x_j \in V_j$, $j = 1, \ldots, p$, such that $x = x_1 + \cdots + x_p = (\pi_1(x), \ldots, \pi_p(x))$ and $t^A x = (t^{A_1}\pi_1(x), \ldots, t^{A_p}\pi_p(x))$ for all $t > 0$. This later fact is a consequence of $t^A = t^{A_1} \oplus \cdots \oplus t^{A_p}$ which can be easily checked using (4.1.3). Moreover, for our multidimensional process we have $X_t = (X_t^{[1]}, \ldots, X_t^{[p]})$, where $(X_t^{[j]} = \pi_j(X_t))_{t \in [0,T)}$ is again of the same structure (4.1.2) which will be shown below in Lemma 4.2.1. Thus it suffices to show that for each component $X_t^{[j]} \to 0 \in \mathbb{R}^{d_j}$ almost surely to deduce $X_t \to 0 \in \mathbb{R}^d$ almost surely as $t \uparrow T$.

LEMMA 4.2.1. *For every $j = 1, \ldots, p$, the j-th spectral component of $(X_t)_{t \in [0,T)}$ can almost surely be represented as*

$$X_t^{[j]} = \int_0^t \Big(\frac{T-t}{T-s}\Big)^{A_j} \Sigma_j \, \mathrm{d}B_s \qquad \text{for } t \in [0,T), \tag{4.2.1}$$

where $\Sigma_j \in \mathbb{R}^{d_j \times m}$ is given by $\pi_j(\Sigma y) = \Sigma_j y$ for $y \in \mathbb{R}^m$.

Note that, by Lemma 4.2.1, $(X_t^{[j]})_{t \in [0,T)}$ structurally has the same integral representation (4.1.2) but with real spectrally simple exponent A_j whose eigenvalues all have the same real part a_j. Concluding, we only need to consider real spectrally simple exponents A to decide whether $X_t \to 0$ almost surely as $t \uparrow T$ or not.

We will need the following result on the growth behavior of the exponential operator t^{A_j} near the origin $t = 0$ for $j = 1, \ldots, p$. For a matrix $Q \in \mathbb{R}^{d_j \times d_j}$, now we choose the associated matrix norm

$$\|Q\| := \sup\big\{\|Qy\| : \|y\| = 1,\, y \in \mathbb{R}^{d_j}\big\}$$

with respect to the standard Euclidean norm $\|y\|$ for $y \in \mathbb{R}^{d_j}$.

LEMMA 4.2.2. *For every $j = 1, \ldots, p$ and every $\varepsilon > 0$, there exists a constant $K \in (0, \infty)$ such that for all $0 < t \leqslant T$ we have*

$$\|t^{A_j}\| \leqslant K\, t^{a_j - \varepsilon} \quad \text{and} \quad \|t^{-A_j}\| \leqslant K\, t^{-(a_j + \varepsilon)}.$$

We note that in Lemma 4.2.2 one can use any matrix norm on $\mathbb{R}^{d_j \times d_j}$ (since any two matrix norms on $\mathbb{R}^{d_j \times d_j}$ are equivalent).

4.2.2. SLLN and LIL for martingales on $[0,T)$**.** Recall the integral representation (4.1.2) of the solution $(X_t)_{t \in [0,T)}$ of (4.1.1) with $X_0 = 0$. We may write

$$X_t = (T-t)^A M_t \quad \text{with} \quad M_t := \int_0^t (T-s)^{-A} \Sigma \, \mathrm{d}B_s, \quad t \in [0,T). \tag{4.2.2}$$

Here $(M_t)_{t \in [0,T)}$ is a continuous square-integrable martingale whose i-th coordinate $(M_t^{(i)})_{t \in [0,T)}$ has quadratic variation process given by

$$\langle M^{(i)} \rangle_t = \int_0^t \big\|e_i^\top (T-s)^{-A} \Sigma\big\|^2 \, \mathrm{d}s, \quad t \in [0,T), \tag{4.2.3}$$

for every $i = 1, \dots, d$, where $\{e_1, \dots, e_d\}$ denotes the canonical basis of $\mathbb{R}^d$. Note that $(\langle M^{(i)} \rangle_t)_{t\in[0,T)}$ is a continuous deterministic function.

We call the attention that from now on the superscripts in curved brackets denote coordinates rather than spectral components denoted by superscripts with squared brackets as in Section 4.2.1.

Usually, the strong law of large numbers for martingales is formulated as a limit theorem as $t \to \infty$. In our case we need to consider the limiting behavior as $t \uparrow T$. Due to the strictly increasing and continuous time change $t(s) = (2T/\pi) \arctan s$, $s \geqslant 0$ (which is a bijection between $[0, \infty)$ and $[0, T)$), we get that $(\widetilde{M}_s := M_{t(s)})_{s\geqslant 0}$ is a continuous square-integrable martingale with respect to the filtration $(\widetilde{\mathcal{F}}_s := \mathcal{F}_{t(s)})_{s\geqslant 0}$ and we can easily adopt the following well-known versions of the strong law of large numbers for continuous square-integrable martingales.

LEMMA 4.2.3. *If* $\lim_{t\uparrow T} \langle M^{(i)} \rangle_t < \infty$ *for every* $i = 1, \dots, d$, *then*

$$P\left(\lim_{t\uparrow T} M_t \text{ exists}\right) = 1.$$

For the proof we refer to Proposition 4.1.26 together with Proposition 5.1.8 in [**143**].

LEMMA 4.2.4. *Let* $f : [x_0, \infty) \to (0, \infty)$ *be an increasing function, where* $x_0 > 0$ *such that* $\int_{x_0}^{\infty} f(x)^{-2} \,\mathrm{d}x < \infty$. *If* $\lim_{t\uparrow T} \langle M^{(i)} \rangle_t = \infty$ *for some* $i \in \{1, \dots, d\}$, *then*

$$P\left(\lim_{t\uparrow T} \frac{M_t^{(i)}}{f(\langle M^{(i)} \rangle_t)} = 0\right) = 1.$$

For the proof we refer to Exercise 5.1.16 in [**143**] or to Theorem 2.3 in [**27**].

Next we present a law of the iterated logarithm for $(M_t)_{t\in[0,T)}$.

LEMMA 4.2.5. *If* $P(\lim_{t\uparrow T} \langle M^{(i)} \rangle_t = \infty) = 1$ *for some* $i \in \{1, \dots, d\}$, *then*

$$\begin{aligned} P&\left(\limsup_{t\uparrow T} \frac{M_t^{(i)}}{\sqrt{2\langle M^{(i)} \rangle_t \ln(\ln\langle M^{(i)} \rangle_t)}} = 1\right) \\ &= P\left(\liminf_{t\uparrow T} \frac{M_t^{(i)}}{\sqrt{2\langle M^{(i)} \rangle_t \ln(\ln\langle M^{(i)} \rangle_t)}} = -1\right) = 1. \end{aligned}$$

Lemma 4.2.5 follows by Exercise 1.15 in Chapter V of Revuz and Yor [**143**].

4.3. Bridge property

Let $\operatorname{ReSpec}(A) := \{\operatorname{Re}\lambda : \lambda \in \operatorname{Spec}(A)\}$ be the collection of distinct real parts of the eigenvalues of the matrix A, where $\operatorname{Spec}(A)$ denotes the set of eigenvalues of A. If there exists $\lambda \in \operatorname{Spec}(A)$ with $\operatorname{Re}\lambda \leqslant 0$, then the process $(X_t)_{t\in[0,T)}$ defined by (4.1.2) with initial condition $X_0 = 0 \in \mathbb{R}^d$ does not fulfill that X_t converges to some deterministic d-dimensional vector almost surely as $t \uparrow T$ in general. This fact is known for the one-dimensional situation $d = 1$ from Remark 3.5 in [**27**]. To give an explicit multidimensional example, we consider a $d \times d$ matrix A having only purely imaginary eigenvalues.

EXAMPLE 4.3.1. Let $\Sigma = I_d$ be the $d \times d$ identity matrix and let $A \in \mathbb{R}^{d\times d}$ be a skew symmetric matrix, i.e. $A^\top = -A$. Then all the non-zero eigenvalues of A are purely imaginary and r^A is an orthogonal matrix for every $r > 0$. Due to the invariance of the incremental distributions of a standard Wiener process with respect to orthogonal transformations, one can easily derive that the distributions of X_t and B_t coincide for every $t \in [0, T)$. Hence X_t converges in distribution to

B_T as $t \uparrow T$, which shows that it cannot hold that X_t converges almost surely to some deterministic d-dimensional vector as $t \uparrow T$.

Our next result is about the limit behavior of the quadratic variation processes $\langle M^{(i)} \rangle_t$ as $t \uparrow T$ for $i = 1, \ldots, d$.

LEMMA 4.3.2. *If* $\operatorname{ReSpec}(A) \subseteq (0, 1/2)$*, then for all* $i = 1, \ldots, d$*, the quadratic variation process* $(\langle M^{(i)} \rangle_t)_{t \in [0,T)}$ *is bounded. If* $\operatorname{ReSpec}(A) \subseteq (1/2, \infty)$ *and* Σ *has full rank* d *(and consequently* $m \geqslant d$*), then* $\lim_{t \uparrow T} \langle M^{(i)} \rangle_t = \infty$ *for all* $i = 1, \ldots, d$.

REMARK 4.3.3. We conjecture that $\lim_{t \uparrow T} \langle M^{(i)} \rangle_t = \infty$ for every $i \in \{1, \ldots, d\}$ in case $\operatorname{ReSpec}(A) = \{1/2\}$. However, we cannot address a precise argument. Note that in dimension 1 this holds; see the proof of Lemma 3.1 in Barczy and Pap [**27**]. Fortunately, for proving the bridge property of $(X_t)_{t \in [0,T)}$ we do not need any information about the limit behavior of the quadratic variation process in case A has eigenvalues with real part all equal to $\frac{1}{2}$, see the proof of Theorem 4.3.4 below. $\square$

Now we are ready to formulate our main result.

THEOREM 4.3.4. *Let us suppose that* Σ *has full rank* d *(and consequently* $m \geqslant d$*). If* $\operatorname{ReSpec}(A) \subseteq (0, \infty)$*, then the process*

$$\widehat{X}_t := \begin{cases} \displaystyle\int_0^t \left(\frac{T-t}{T-s} \right)^A \Sigma \, \mathrm{d}B_s & \text{if } t \in [0, T), \\ 0 & \text{if } t = T \end{cases} \tag{4.3.1}$$

is a centered Gauss process with almost surely continuous sample paths.

REMARK 4.3.5. Note that the condition $\operatorname{ReSpec}(A) \subseteq (0, \infty)$ is equivalent to $t^A \to 0 \in \mathbb{R}^{d \times d}$ as $t \downarrow 0$. We call the attention that the condition that Σ has full rank d in Theorem 4.3.4 is needed only for the case $\operatorname{ReSpec}(A) \cap [1/2, \infty) \neq \emptyset$; see the proof given below. Moreover, as mentioned in the Introduction, the assumption that Σ has full rank d is only a minor restriction to the generality of Theorem 4.3.4. $\square$

4.4. Asymptotic behavior of the bridge

In this section we study asymptotic behavior of the sample paths of the operator scaled Wiener bridge $(X_t)_{t \in [0,T)}$ given by (4.1.1) with initial condition $X_0 = 0$.

Our first result is a partial generalization of Theorem 3.4 in Barczy and Pap [**27**].

PROPOSITION 4.4.1. *If* $\operatorname{ReSpec}(A) \subseteq (0, 1/2)$*, then*

$$P\Big(\lim_{t \uparrow T} (T-t)^{-A} X_t = M_T \Big) = 1, \tag{4.4.1}$$

where M_T *is a* d*-dimensional normally distributed random variable. Consequently, for all* $\widetilde{A} \in \mathbb{R}^{d \times d}$ *with* $A\widetilde{A} = \widetilde{A}A$*, we have*

$$P\Big(\lim_{t \uparrow T} (T-t)^{-\widetilde{A}} X_t = 0 \Big) = 1 \quad \text{if} \quad \operatorname{ReSpec}(A - \widetilde{A}) \subseteq (0, \infty), \tag{4.4.2}$$

$$P\Big(\lim_{t \uparrow T} \|(T-t)^{-\widetilde{A}} X_t\| = \infty \Big) = 1 \ \text{if} \ \operatorname{ReSpec}(A - \widetilde{A}) \subseteq (-\infty, 0). \tag{4.4.3}$$

Recall the spectral decomposition of the process $(X_t)_{t \in [0,T)}$, see Lemma 4.2.1. For the spectral components, one can get the following precise asymptotic result.

THEOREM 4.4.2. *If* $\operatorname{ReSpec}(A) \subseteq (0,\infty)$ *and* Σ *has full rank* d *(and consequently* $m \geqslant d$*), then for all* $\varepsilon > 0$,

$$P\Big(\lim_{t\uparrow T}(T-t)^{-\min(a_j,1/2)+\varepsilon}\|X_t^{[j]}\| = 0\Big) = 1, \tag{4.4.4}$$

$$P\Big(\limsup_{t\uparrow T}(T-t)^{-\min(a_j,1/2)-\varepsilon}\|X_t^{[j]}\| = \infty\Big) = 1, \tag{4.4.5}$$

where $a_1 < \cdots < a_p$ *denote the distinct real parts of the eigenvalues of* A *and* $(X_t^{[j]})_{t\in[0,T)}$, $j = 1,\ldots,p$, *are the corresponding spectral components of* $(X_t)_{t\in[0,T)}$, *see Section 4.2.1. Further, if* $\operatorname{ReSpec}(A) \subseteq (0,1/2)$, *then* (4.4.5) *can be strengthened to*

$$P\Big(\lim_{t\uparrow T}(T-t)^{-a_j-\varepsilon}\|X_t^{[j]}\| = \infty\Big) = 1. \tag{4.4.6}$$

4.5. Uniqueness in law of operator scaled Wiener bridges

For $A, \widetilde{A} \in \mathbb{R}^{d\times d}$ and $\Sigma \in \mathbb{R}^{d\times m}$, $\widetilde{\Sigma} \in \mathbb{R}^{d\times \widetilde{m}}$, let the processes $(X_t)_{t\in[0,T)}$ and $(Y_t)_{t\in[0,T)}$ be given by the SDEs

$$\mathrm{d}X_t = -\frac{1}{T-t}A\,X_t\,\mathrm{d}t + \Sigma\,\mathrm{d}B_t, \quad t \in [0,T),$$
$$\mathrm{d}Y_t = -\frac{1}{T-t}\widetilde{A}\,Y_t\,\mathrm{d}t + \widetilde{\Sigma}\,\mathrm{d}\widetilde{B}_t, \quad t \in [0,T),$$

with initial conditions $X_0 = 0$ and $Y_0 = 0$, where $(B_t)_{t\geqslant 0}$ and $(\widetilde{B}_t)_{t\geqslant 0}$ are m-, respectively $\widetilde{m}$-dimensional standard Wiener process. Assume that $(X_t)_{t\in[0,T)}$ and $(Y_t)_{t\in[0,T)}$ generate the same law on the space of real-valued continuous functions defined on $[0,T)$. Since $(X_t)_{t\in[0,T)}$ and $(Y_t)_{t\in[0,T)}$ are centered Gauss processes, their laws coincide if and only if their covariance functions coincide. Let $(U(t) := \mathbb{E}(X_tX_t^\top))_{t\in[0,T)}$ and $(V(t) := \mathbb{E}(Y_tY_t^\top))_{t\in[0,T)}$ be the corresponding covariance functions. Then, by Problem 5.6.1 in [**100**], we have

$$U'(t) = -\frac{1}{T-t}\,AU(t) - U(t)A^\top\frac{1}{T-t} + \Sigma\Sigma^\top, \quad t \in [0,T), \tag{4.5.1}$$

$$V'(t) = -\frac{1}{T-t}\,\widetilde{A}V(t) - V(t)\widetilde{A}^\top\frac{1}{T-t} + \widetilde{\Sigma}\widetilde{\Sigma}^\top, \quad t \in [0,T). \tag{4.5.2}$$

Since $U(t) = V(t)$ for all $t \in [0,T)$, we get

$$-\frac{1}{T-t}\,AU(t) - U(t)A^\top\frac{1}{T-t} + \Sigma\Sigma^\top = -\frac{1}{T-t}\,\widetilde{A}U(t) - U(t)\widetilde{A}^\top\frac{1}{T-t} + \widetilde{\Sigma}\widetilde{\Sigma}^\top$$

for $t \in [0,T)$. Since $U(0) = 0 \in \mathbb{R}^{d\times d}$, we have $\Sigma\Sigma^\top = \widetilde{\Sigma}\widetilde{\Sigma}^\top$, and hence

$$(A-\widetilde{A})U(t) = -U(t)(A-\widetilde{A})^\top, \quad t \in [0,T).$$

Unfortunately, this does not imply that $A = \widetilde{A}$ in general. Before we construct counterexamples, we will give the solutions of the $\mathbb{R}^{d\times d}$-valued differential equations (4.5.1) and (4.5.2) with initial condition $U(0) = 0$ and $V(0) = 0$, respectively. By Section 5.6.A in [**100**], one easily calculates that

$$U(t) = \int_0^t \left(\frac{T-t}{T-s}\right)^A \Sigma\Sigma^\top \left(\frac{T-t}{T-s}\right)^{A^\top} \mathrm{d}s \tag{4.5.3}$$

for every $t \in [0,T)$. Analogously, using also that $\Sigma\Sigma^\top = \widetilde{\Sigma}\widetilde{\Sigma}^\top$, we have

$$V(t) = \int_0^t \left(\frac{T-t}{T-s}\right)^{\widetilde{A}} \Sigma\Sigma^\top \left(\frac{T-t}{T-s}\right)^{\widetilde{A}^\top} \mathrm{d}s, \quad t \in [0,T). \tag{4.5.4}$$

Next we give examples for bridges associated to the matrices A and Σ, and $\widetilde{A}$ and Σ, respectively, such that their laws on the space of real-valued continuous functions on $[0,T)$ coincide, but $A \neq \widetilde{A}$.

EXAMPLE 4.5.1. Let $A \in \mathbb{R}^{d\times d}$ be a normal matrix, i.e. $AA^\top = A^\top A$. Choose $\widetilde{A} = A^\top$ and let $\Sigma = I_d = \widetilde{\Sigma}$, then for every $r > 0$ we have

$$r^A \Sigma\Sigma^\top r^{A^\top} = r^{A+A^\top} = r^{\widetilde{A}+\widetilde{A}^\top} = r^{\widetilde{A}}\Sigma\Sigma^\top r^{\widetilde{A}^\top}. \tag{4.5.5}$$

By (4.5.3) and (4.5.4) it follows that $U(t) = V(t)$ for all $t \in [0,T)$. Using Theorem 4.3.4, the bridges associated to the matrices A and Σ, and $\widetilde{A}$ and $\widetilde{\Sigma}$ coincide, but $A \neq \widetilde{A}$. Note also that here the eigenvalues of A and $\widetilde{A} = A^\top$ coincide.

We further wish to give an example, where the eigenvalues of A and $\widetilde{A}$ do not coincide, but still $U(t) = V(t)$ holds for all $t \in [0,T)$. Choose the normal matrices

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \quad \text{and} \quad \widetilde{A} = I_2$$

together with $\Sigma = I_2 = \widetilde{\Sigma}$, then due to $A + A^\top = 2I_2 = \widetilde{A} + \widetilde{A}^\top$ again (4.5.5) holds for every $r > 0$, which yields that $U(t) = V(t)$ for all $t \in [0,T)$ as above. Note that now the eigenvalues $1+i$ and $1-i$ of A do not coincide with those of $\widetilde{A} = I_2$, but the real parts of the eigenvalues do, including their multiplicity.

To conclude, we formulate a partial result on the uniqueness of the scaling matrix.

PROPOSITION 4.5.2. *Let A, $\widetilde{A} \in \mathbb{R}^{d\times d}$ and $\Sigma \in \mathbb{R}^{d\times m}$, $\widetilde{\Sigma} \in \mathbb{R}^{d\times \widetilde{m}}$ be such that* $\operatorname{ReSpec}(A) \subseteq (0,1/2)$, $\operatorname{ReSpec}(\widetilde{A}) \subseteq (0,1/2)$ *and Σ, $\widetilde{\Sigma}$ have full rank d (and consequently $m \geqslant d$ and $\widetilde{m} \geqslant d$). If the bridges associated to the matrices A and Σ, and $\widetilde{A}$ and $\widetilde{\Sigma}$ induce the same law on the space of real-valued continuous functions on $[0,T)$, then* $\operatorname{ReSpec}(A) = \operatorname{ReSpec}(\widetilde{A})$.

REMARK 4.5.3. We conjecture that Proposition 4.5.2 also holds in the situation $\operatorname{ReSpec}(A) \subseteq (0,\infty)$, $\operatorname{ReSpec}(\widetilde{A}) \subseteq (0,\infty)$ but we were not able to give a rigorous proof. $\square$

Part 2

Two-factor affine processes

Introduction and summary

This part is based on the articles Barczy et al. [**11**], [**12**], [**13**] and Barczy and Pap [**30**].

In Barczy et al. [**11**], first we provide a simple set of sufficient conditions for the weak convergence of scaled affine processes with state space $\mathbb{R}_+ \times \mathbb{R}^d$. We specialize our result to one-dimensional continuous state and continuous time branching processes with immigration as well. As an application, we study the asymptotic behavior of least squares estimators of some parameters of a two-dimensional critical affine diffusion process.

In Barczy et al. [**12**], we study the existence of a unique stationary distribution and ergodicity for a two-dimensional (subcritical) affine process. The first coordinate is supposed to be a so-called α-root process with $\alpha \in (1, 2]$. The existence of a unique stationary distribution for the affine process is proved in case of $\alpha \in (1, 2]$; further, in case of $\alpha = 2$, the ergodicity is also shown.

In Barczy et al. [**13**], for a subcritical diffusion $(\alpha = 2)$ affine two-factor model, we study the asymptotic properties of the maximum likelihood and least squares estimators of some appearing parameters based on continuous time observations. We prove strong consistency and asymptotic normality of the estimators in question.

In Barczy and Pap [**30**], we study asymptotic properties of maximum likelihood estimators for Heston models based on continuous time observations of the log-price process. We distinguish three cases: subcritical (also called ergodic), critical and supercritical. In the subcritical case, asymptotic normality is proved for all the parameters, while in the critical and supercritical cases, non-standard asymptotic behavior is described.

CHAPTER 5

On parameter estimation for critical affine processes

Co-authors: Leif Doering, Zenghu Li and Gyula Pap

5.1. Introduction

In recent years quickly growing interest in pricing of credit-risky securities (e.g., defaultable bonds) has been seen in the mathematical finance literature. One of the basic models (for applications see for instance Chen and Joslin [**45**]) is the following two-dimensional affine diffusion process:

$$(5.1.1) \qquad \begin{cases} \mathrm{d}Y_t = (a - bY_t)\,\mathrm{d}t + \sqrt{Y_t}\,\mathrm{d}W_t, \\ \mathrm{d}X_t = (m - \theta X_t)\,\mathrm{d}t + \sqrt{Y_t}\,\mathrm{d}B_t, \end{cases} \qquad t \geqslant 0,$$

where a, b, θ and m are real parameters such that $a > 0$ and B and W are independent standard Wiener processes. Note that Y is a Cox-Ingersol-Ross (CIR) process. For practical use, it is important to estimate the appearing parameters from some discretely observed real data set. In the case of the one-dimensional CIR process, the parameter estimation of a and b goes back to Overbeck and Rydén [**135**], Overbeck [**136**], and see also the very recent papers of Ben Alaya and Kebaier [**34, 35**]. For asymptotic results on discrete time critical branching processes with immigration, one may refer to Wei and Winnicki [**153**] and [**154**].

The process (Y, X) given by (5.1.1) is a very special affine process. The set of affine processes contains a large class of important Markov processes such as continuous state branching processes and Orstein-Uhlenbeck processes. Further, a lot of models in financial mathematics are also special affine processes such as the Heston model [**84**], the model due to Barndorff-Nielsen and Shephard [**32**] or the model due to Carr and Wu [**42**]. A precise mathematical formulation and complete characterization of regular affine processes are due to Duffie et al. [**63**]. Later several authors have contributed to the study of properties of general affine processes: to name a few, Andersen and Piterbarg [**5**] (moment explosions in stochastic volatility models), Dawson and Li [**53**] (jump-type SDE representation for two-dimensional affine processes), Filipović and Mayerhofer [**69**] (applications to the pricing of bond and stock options), Glasserman and Kim [**78**] (the range of finite exponential moments and the convergence to stationarity in affine diffusion models), Jena et al. [**97**] (long-term and blow-up behaviors of exponential moments in multidimensional affine diffusions), Keller-Ressel et al. [**104, 105**] (stochastically continuous, time-homogeneous affine processes with state space $\mathbb{R}_+^n \times \mathbb{R}^d$ or more general ones are regular). We also refer to the overview articles Cuchiero et al. [**51**] and Friz and Keller-Ressel [**73**].

To the best knowledge of the authors the parameter estimation problem for multidimensional affine processes has not been tackled so far. Since affine processes are being used in financial mathematics very frequently, the question of parameter estimation for them is of high importance. Our aim is to start the discussion with a simple non-trivial example: the two-dimensional affine diffusion process given by (5.1.1).

In Section 5.2 we recall some notations, the definition of affine processes and some of their basic properties, and then a simple set of sufficient conditions for the weak convergence of scaled affine processes is presented. Roughly speaking, given a family of affine processes $(Y^{(\theta)}(t), X^{(\theta)}(t))_{t\geqslant 0}$, $\theta > 0$, such that the corresponding admissible parameters converge in an appropriate way (see Theorem 5.2.9), the scaled process $\big(\theta^{-1}Y^{(\theta)}(\theta t), \theta^{-1}X^{(\theta)}(\theta t)\big)_{t\geqslant 0}$ converge weakly towards an affine diffusion process as $\theta \to \infty$. We specialize our result for one-dimensional continuous state branching processes with immigration which generalizes Theorem 2.3 in Huang et al. [**91**]. The scaling Theorem 5.2.9 is proved for quite general affine processes since it might have applications elsewhere later on. In Section 5.3 the scaling Theorem 5.2.9 is applied to study the asymptotic behavior of least squares and conditional least squares estimators of some parameters of a critical two-dimensional affine diffusion process given by (5.1.1), see Theorems 5.4.1, 5.5.1 and 5.6.2.

5.2. A scaling theorem for affine processes

Let $\mathbb{N}$, $\mathbb{Z}_+$, $\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{R}_-$, $\mathbb{R}_{++}$, and C denote the sets of positive integers, non-negative integers, real numbers, non-negative real numbers, non-positive real numbers, positive real numbers and complex numbers, respectively. For $x, y \in \mathbb{R}$, we will use the notations $x \wedge y := \min(x, y)$ and $x \vee y := \max(x, y)$. For $x, y \in \mathsf{C}^k$, $k \in \mathbb{N}$, we write $\langle x, y\rangle := \sum_{i=1}^k x_i y_i$ (notice that this is not the scalar product on C^k, however for $x \in \mathsf{C}^k$ and $y \in \mathbb{R}^k$, $\langle x, y\rangle$ coincides with the usual scalar product of x and y). By $\|x\|$ and $\|A\|$ we denote the Euclidean norm of a vector $x \in \mathbb{R}^p$ and the induced matrix norm of a matrix $A \in \mathbb{R}^{p\times p}$, respectively. Further, let $U := \{z_1 + iz_2 : z_1 \in \mathbb{R}_-, z_2 \in \mathbb{R}\} \times (i\mathbb{R}^d)$. By $C^2_c(\mathbb{R}_+ \times \mathbb{R}^d)$ $(C^\infty_c(\mathbb{R}_+ \times \mathbb{R}^d))$ we denote the set of twice (infinitely) continuously differentiable complex-valued functions on $\mathbb{R}_+ \times \mathbb{R}^d$ with compact support, where $d \in \mathbb{N}$. The set of càdlàg functions from $\mathbb{R}_+$ to $\mathbb{R}_+ \times \mathbb{R}^d$ will be denoted by $\mathsf{D}(\mathbb{R}_+, \mathbb{R}_+ \times \mathbb{R}^d)$. For a bounded function $g : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^p$, let $\|g\|_\infty := \sup_{x\in\mathbb{R}_+\times\mathbb{R}^d} \|g(x)\|$. Convergence in distribution, in probability and almost sure convergence will be denoted by $\stackrel{\mathcal{L}}{\longrightarrow}$, $\stackrel{\mathbb{P}}{\longrightarrow}$ and $\stackrel{\text{a.s.}}{\longrightarrow}$, respectively.

Next we briefly recall the definition of affine processes with state space $\mathbb{R}_+ \times \mathbb{R}^d$ based on Duffie et al. [**63**].

Definition 5.2.1. A transition semigroup $(P_t)_{t\in\mathbb{R}_+}$ with state space $\mathbb{R}_+ \times \mathbb{R}^d$ is called a (general) affine semigroup if its characteristic function has the representation

$$\int_{\mathbb{R}_+\times\mathbb{R}^d} \mathrm{e}^{\langle u,\xi\rangle} P_t(x, \mathrm{d}\xi) = \mathrm{e}^{\langle x,\psi(t,u)\rangle + \phi(t,u)} \tag{5.2.1}$$

for $x \in \mathbb{R}_+ \times \mathbb{R}^d$, $u \in U$ and $t \in \mathbb{R}_+$, where $\psi(t, \cdot) = (\psi_1(t, \cdot), \psi_2(t, \cdot)) \in \mathsf{C} \times \mathsf{C}^d$ is a continuous C^{1+d}-valued function on U and $\phi(t, \cdot)$ is a continuous C-valued function on U satisfying $\phi(t, 0) = 0$. The affine semigroup $(P_t)_{t\in\mathbb{R}_+}$ defined by (5.2.1) is called regular if it is stochastically continuous (equivalently, for all $u \in U$, the functions $\mathbb{R}_+ \ni t \mapsto \Psi(t, u)$ and $\mathbb{R}_+ \ni t \mapsto \phi(t, u)$ are continuous) and $\partial_1\psi(0, u)$ and $\partial_1\phi(0, u)$ exist for all $u \in U$ and are continuous at $u = 0$ (where $\partial_1\psi$ and $\partial_1\phi$ denote the partial derivatives of ψ and ϕ, respectively, with respect to the first variable).

Remark 5.2.2. We call the attention that Duffie et al. [**63**] in their Definition 2.1 assume only that Equation (5.2.1) hold for $x \in \mathbb{R}_+ \times \mathbb{R}^d$, $u \in \partial U = i\mathbb{R}^{1+d}$,

$t \in \mathbb{R}_+$, i.e., instead of $u \in U$ they only require that u should be an element of the boundary ∂U of U. However, by Proposition 6.4 in Duffie et al. [**63**], one can formulate the definition of a regular affine process as we did. Note also that this kind of definition was already given by Dawson and Li [**53**, Definitions 2.1 and 3.3]. Finally, we remark that every stochastically continuous affine semigroup is regular due to Keller-Ressel et al. [**104**, Theorem 5.1]. □

Definition 5.2.3. A set of parameters $(a, \alpha, b, \beta, m, \mu)$ is called admissible if

(i) $a = (a_{i,j})_{i,j=1}^{1+d} \in \mathbb{R}^{(1+d)\times(1+d)}$ is a symmetric positive semidefinite matrix with $a_{1,1} = 0$ (hence $a_{1,k} = a_{k,1} = 0$ for all $k \in \{2, \dots, 1+d\}$),
(ii) $\alpha = (\alpha_{i,j})_{i,j=1}^{1+d} \in \mathbb{R}^{(1+d)\times(1+d)}$ is a symmetric positive semidefinite matrix,
(iii) $b = (b_i)_{i=1}^{1+d} \in \mathbb{R}_+ \times \mathbb{R}^d$,
(iv) $\beta = (\beta_{i,j})_{i,j=1}^{1+d} \in \mathbb{R}^{(1+d)\times(1+d)}$ with $\beta_{1,j} = 0$ for all $j \in \{2, \dots, 1+d\}$,
(v) $m(\mathrm{d}\xi) = m(\mathrm{d}\xi_1, \mathrm{d}\xi_2)$ is a σ-finite measure on $\mathbb{R}_+ \times \mathbb{R}^d$ supported by $(\mathbb{R}_+ \times \mathbb{R}^d) \setminus \{(0,0)\}$ such that
$$\int_{\mathbb{R}_+\times\mathbb{R}^d} \left[\xi_1 + (\|\xi_2\| \wedge \|\xi_2\|^2)\right] m(\mathrm{d}\xi) < \infty,$$
(vi) $\mu(\mathrm{d}\xi) = \mu(\mathrm{d}\xi_1, \mathrm{d}\xi_2)$ is a σ-finite measure on $\mathbb{R}_+ \times \mathbb{R}^d$ supported by $(\mathbb{R}_+ \times \mathbb{R}^d) \setminus \{(0,0)\}$ such that
$$\int_{\mathbb{R}_+\times\mathbb{R}^d} \|\xi\| \wedge \|\xi\|^2 \mu(\mathrm{d}\xi) < \infty.$$

Remark 5.2.4. Note that our Definition 5.2.3 of the set of admissible parameters is not so general as Definition 2.6 in Duffie et al. [**63**]. Firstly, the set of admissible parameters is defined only for affine process with state space $\mathbb{R}_+ \times \mathbb{R}^d$, while Duffie et al. [**63**] consider affine processes with state space $\mathbb{R}_+^n \times \mathbb{R}^d$. We restrict ourselves to this special case, since our scaling Theorem 5.2.9 is valid only in this case. Secondly, our conditions (v) and (vi) of Definition 5.2.3 are stronger than that of (2.10) and (2.11) of Definition 2.6 in Duffie et al. [**63**]. Thirdly, according to our definition, a set of admissible parameters does not contain parameters corresponding to killing, while in Definition 2.6 in Duffie et al. [**63**] such parameters are included. Our definition of admissible parameters can be considered as a $(1+d)$-dimensional version of Definition 6.1 in Dawson and Li [**53**]. The reason for this definition is to have a more pleasant form of the infinitesimal generator of an affine process compared to that of Duffie et al. [**63**, formula (2.12)]. For more details, see Remark 5.2.6. □

Theorem 5.2.5. *(Duffie et al.* [**63**, Theorem 2.7]*) Let $(a, \alpha, b, \beta, m, \mu)$ be a set of admissible parameters. Then there exists a unique regular affine semigroup $(P_t)_{t\in\mathbb{R}_+}$ with infinitesimal generator*

$$\begin{aligned}(\mathcal{A}f)(x) = &\sum_{i,j=1}^{1+d} (a_{i,j} + \alpha_{i,j}x_1) f''_{i,j}(x) + \langle f'(x), b + \beta x\rangle \\ &+ \int_{\mathbb{R}_+\times\mathbb{R}^d} (f(x+\xi) - f(x) - \langle f'_{(2)}(x), \xi_2\rangle)\, m(\mathrm{d}\xi) \\ &+ \int_{\mathbb{R}_+\times\mathbb{R}^d} (f(x+\xi) - f(x) - \langle f'(x), \xi\rangle) x_1\, \mu(\mathrm{d}\xi)\end{aligned} \tag{5.2.2}$$

for $x = (x_1, x_2) \in \mathbb{R}_+ \times \mathbb{R}^d$ and $f \in C_c^2(\mathbb{R}_+ \times \mathbb{R}^d)$, where f'_i, $i \in \{1, \dots, 1+d\}$, and $f''_{i,j}$, $i, j \in \{1, \dots, 1+d\}$, denote the first and second order partial derivatives of f with respect to its i-th and i-th and j-th variables, and $f'(x) :=$

$(f_1'(x),\dots,f_{1+d}'(x))^\top$, $f_{(2)}'(x) := (f_2'(x),\dots,f_{1+d}'(x))^\top$. *Further,* $\mathcal{C}_c^\infty(\mathbb{R}_+\times\mathbb{R}^d)$ *is a core of* $\mathcal{A}$.

REMARK 5.2.6. Note that the form of the infinitesimal generator $\mathcal{A}$ in Theorem 5.2.5 is slightly different from the one given in (2.12) in Duffie et al. [**63**]. Our formula (5.2.2) is in the spirit of Dawson and Li [**53**, formula (6.5)]. On the one hand, the point is that under the conditions (v) and (vi) of Definition 5.2.3, one can rewrite (2.12) in Duffie et al. [**63**] into the form (5.2.2), by changing the 2-nd, ..., $(1+d)$-th coordinates of $b\in\mathbb{R}_+\times\mathbb{R}^d$ and the first column of $\beta\in\mathbb{R}^{(1+d)\times(1+d)}$, respectively, in appropriate ways (see Remark 5.6.4). To see this, it is enough to check that the integrals in (5.2.2) are well-defined (i.e., elements of C) under the conditions (v) and (vi) of Definition 5.2.3. For further details, see also Remark 5.6.4. On the other hand, the killing rate (see page 995 in Duffie et al. [**63**]) of the affine semigroup $(P_t)_{t\in\mathbb{R}_+}$ in Theorem 5.2.5 is identically zero. This also implies that the affine processes that we will consider later on will have lifetime infinity. $\Box$

REMARK 5.2.7. In dimension 2 (i.e., if $d=1$), by Theorem 6.2 in Dawson and Li [**53**] and Theorem 2.7 in Duffie et al. [**63**] (see also Theorem 5.2.5), for an infinitesimal generator $\mathcal{A}$ given by (5.2.2) with $d=1$ one can construct a two-dimensional system of jump type SDEs of which there exists a pathwise unique strong solution $(Y(t),X(t))_{t\in\mathbb{R}_+}$ which is a regular affine Markov process with the given infinitesimal generator $\mathcal{A}$. $\Box$

The next lemma is simple but very useful.

LEMMA 5.2.8. *Let* $(Z(t))_{t\in\mathbb{R}_+}$ *be a time-homogeneous Markov process with state space* $\mathbb{R}_+\times\mathbb{R}^d$ *and let us denote its infinitesimal generator by* $\mathcal{A}_Z$. *Suppose that* $\mathcal{C}_c^2(\mathbb{R}_+\times\mathbb{R}^d)$ *is a subset of the domain of* $\mathcal{A}_Z$. *Then for all* $\theta\in\mathbb{R}_{++}$, *the time-homogeneous Markov process* $(Z_\theta(t))_{t\in\mathbb{R}_+} := (\theta^{-1}Z(\theta t))_{t\in\mathbb{R}_+}$ *has infinitesimal generator*

$$(\mathcal{A}_{Z_\theta}f)(x) = \theta(\mathcal{A}_Z f_\theta)(\theta x), \qquad x\in\mathbb{R}_+\times\mathbb{R}^d, \quad f\in C_c^2(\mathbb{R}_+\times\mathbb{R}^d),$$

where $f_\theta(x) := f(\theta^{-1}x)$, $x\in\mathbb{R}_+\times\mathbb{R}^d$.

THEOREM 5.2.9. *For all* $\theta\in\mathbb{R}_{++}$, *let* $(Y^{(\theta)}(t),X^{(\theta)}(t))_{t\in\mathbb{R}_+}$ *be a* $(1+d)$-*dimensional affine process with state space* $\mathbb{R}_+\times\mathbb{R}^d$ *and with admissible parameters* $(a^{(\theta)},\alpha^{(\theta)},b^{(\theta)},\beta^{(\theta)},m,\mu)$ *such that additionally*

$$\int_{\mathbb{R}_+\times\mathbb{R}^d}\|\xi\|\,m(\mathrm{d}\xi)<\infty \qquad \text{and} \qquad \int_{\mathbb{R}_+\times\mathbb{R}^d}\|\xi\|^2\,\mu(\mathrm{d}\xi)<\infty. \tag{5.2.3}$$

If there exist $a,\alpha,\beta\in\mathbb{R}^{(1+d)\times(1+d)}$, $b\in\mathbb{R}_+\times\mathbb{R}^d$, *and a random vector* $(Y(0),X(0))$ *with values in* $\mathbb{R}_+\times\mathbb{R}^d$ *such that*

$$\theta^{-1}a^{(\theta)}\to a, \qquad \alpha^{(\theta)}\to\alpha, \qquad b^{(\theta)}\to b, \qquad \theta\beta^{(\theta)}\to\beta,$$
$$\theta^{-1}(Y^{(\theta)}(0),X^{(\theta)}(0)) \xrightarrow{\mathcal{L}} (Y(0),X(0))$$

as $\theta\to\infty$, *then*

$$\left(Y_\theta^{(\theta)}(t),X_\theta^{(\theta)}(t)\right)_{t\in\mathbb{R}_+} = \left(\theta^{-1}Y^{(\theta)}(\theta t),\theta^{-1}X^{(\theta)}(\theta t)\right)_{t\in\mathbb{R}_+} \xrightarrow{\mathcal{L}} (Y(t),X(t))_{t\in\mathbb{R}_+}$$

in $\mathsf{D}(\mathbb{R}_+,\mathbb{R}_+\times\mathbb{R}^d)$ *as* $\theta\to\infty$, *where* $(Y(t),X(t))_{t\in\mathbb{R}_+}$ *is a* $(1+d)$-*dimensional affine process with state space* $\mathbb{R}_+\times\mathbb{R}^d$ *and with the set of admissible parameters* $(a,\widetilde{\alpha},\widetilde{b},\beta,0,0)$, *where*

$$\widetilde{\alpha} := \alpha + \frac{1}{2}\int_{\mathbb{R}_+\times\mathbb{R}^d}\xi\xi^\top\,\mu(\mathrm{d}\xi),$$

and $\widetilde{b} = (\widetilde{b}_i)_{i=1}^{1+d}$ *with* $\widetilde{b}_i := b_i$ *for* $i \in \{2, \dots, 1+d\}$ *and*

$$\widetilde{b}_1 := b_1 + \int_{\mathbb{R}_+ \times \mathbb{R}^d} \xi_1 \, m(\mathrm{d}\xi).$$

Remark 5.2.10. (i) Note that the limit process $(Y(t), X(t))_{t \in \mathbb{R}_+}$ in Theorem 5.2.9 has continuous sample paths almost surely. However, this is not a big surprise, since in condition (5.2.3) of Theorem 5.2.9 we require finite second moment for the measure μ.
(ii) Note also that the matrix $\widetilde{\alpha} \in \mathbb{R}^{(1+d)\times(1+d)}$ given in Theorem 5.2.9 is symmetric and positive semidefinite, since α is symmetric and positive semidefinite, and for all $z \in \mathbb{R}^{1+d}$,

$$\left\langle \int_{\mathbb{R}_+ \times \mathbb{R}^d} \xi\xi^\top \mu(\mathrm{d}\xi) z, z \right\rangle = \int_{\mathbb{R}_+ \times \mathbb{R}^d} (z^\top \xi)^2 \, \mu(\mathrm{d}\xi) \geqslant 0.$$

$\square$

Remark 5.2.11. By giving an example, we shed some light on why we consider only $(1+d)$-dimensional affine processes with state space $\mathbb{R}_+ \times \mathbb{R}^d$ in Theorem 5.2.9 instead of $(n+d)$-dimensional ones with state space $\mathbb{R}_+^n \times \mathbb{R}^d$, $n \in \mathbb{N}$. Let $(Y_t)_{t \in \mathbb{R}_+}$ be a two-dimensional continuous state branching process with infinitesimal generator

$$(\mathcal{A}_Y f)(y) = \sum_{i=1}^{2} y_i \int_{\mathbb{R}_+^2 \setminus \{0\}} \Big(f(y+u) - f(y) - f_i'(y) u_i \Big) p_i(\mathrm{d}u),$$

for $f \in \mathcal{C}_c^2(\mathbb{R}_+^2)$ and $y = (y_1, y_2) \in \mathbb{R}_+^2$, where p_i, $i = 1, 2$, are σ-finite measures on $\mathbb{R}_+^2 \setminus \{0\}$ such that

(5.2.4)
$$\int_{\mathbb{R}_+^2 \setminus \{0\}} (u_1 + \|u\|^2) p_2(\mathrm{d}u) < \infty \qquad \text{and} \qquad \int_{\mathbb{R}_+^2 \setminus \{0\}} (u_2 + \|u\|^2) p_1(\mathrm{d}u) < \infty,$$

see, e.g., Duffie et al. [**63**, Theorem 2.7]. Note that Y can be considered as a two-dimensional affine process with state space $\mathbb{R}_+^2$ (formally with $d = 0$). Then, by a simple modification of Lemma 5.2.8, for all $\theta > 0$, $f \in \mathcal{C}_c^2(\mathbb{R}_+^2)$ and $y = (y_1, y_2) \in \mathbb{R}_+^2$,

$$\begin{aligned}
(\mathcal{A}_{Y_\theta} f)(y) &= \theta (\mathcal{A}_Y f_\theta)(\theta y) \\
&= \theta \sum_{i=1}^{2} \theta y_i \int_{\mathbb{R}_+^2 \setminus \{0\}} \Big(f(\theta^{-1}(\theta y + u)) - f(\theta^{-1}\theta y) - \theta^{-1} f_i'(\theta^{-1}\theta y) u_i \Big) p_i(\mathrm{d}u) \\
&= \theta^2 \sum_{i=1}^{2} y_i \int_{\mathbb{R}_+^2 \setminus \{0\}} \Big(f(y + \theta^{-1}u) - f(y) - \langle f'(y), \theta^{-1}u \rangle \Big) p_i(\mathrm{d}u) \\
&\quad + \theta \sum_{i=1}^{2} y_i f_{3-i}'(y) \int_{\mathbb{R}_+^2 \setminus \{0\}} u_{3-i} \, p_i(\mathrm{d}u),
\end{aligned}$$

where the last equality follows by (5.2.4). Supposing that f is real-valued, by Taylor's theorem,

$$\begin{aligned}
f(y + \theta^{-1}u) - f(y) - \langle f'(y), \theta^{-1}u \rangle &= \frac{1}{2} \langle f''(y + \tau\theta^{-1}u)\theta^{-1}u, \theta^{-1}u \rangle \\
&= \frac{\theta^{-2}}{2} \langle f''(y + \tau\theta^{-1}u)u, u \rangle
\end{aligned}$$

with some $\tau = \tau(u, y) \in [0, 1]$. Hence, similarly to the proof of (2.7) in Barczy et al. [**11**], we get

$$\begin{aligned}
\lim_{\theta\to\infty} \theta^2 \sum_{i=1}^2 y_i \int_{\mathbb{R}_+^2\setminus\{0\}} \Big(f(y+\theta^{-1}u) - f(y) - \langle f'(y), \theta^{-1}u\rangle \Big) p_i(\mathrm{d}u) \\
= \frac{1}{2} \sum_{i=1}^2 y_i \int_{\mathbb{R}_+^2\setminus\{0\}} \langle f''(y)u, u\rangle p_i(\mathrm{d}u)
\end{aligned}$$

for real-valued $f \in \mathcal{C}_c^2(\mathbb{R}_+^2)$ and $y = (y_1, y_2) \in \mathbb{R}_+^2$. However, $(\mathcal{A}_{Y_\theta} f)(y)$ does not converge as $\theta \to \infty$ provided that

$$\sum_{i=1}^2 y_i f'_{3-i}(y) \int_{\mathbb{R}_+^2\setminus\{0\}} u_{3-i}\, p_i(\mathrm{d}u) \neq 0.$$

We also note that this phenomena is somewhat similar to that of Remark 2.1 in Ma [**123**]. □

In the next remark we formulate some special cases of Theorem 5.2.9.

REMARK 5.2.12. (i) If $(Y(t), X(t))_{t\in\mathbb{R}_+}$ is a $(1 + d)$-dimensional affine process on $\mathbb{R}_+ \times \mathbb{R}^d$ with admissible parameters $(a, \alpha, b, 0, m, \mu)$ such that condition (5.2.3) is satisfied, then the conditions of Theorem 5.2.9 are satisfied for $(Y^{(\theta)}(t), X^{(\theta)}(t))_{t\in\mathbb{R}_+} := (Y(t), X(t))_{t\in\mathbb{R}_+}$, $\theta \in \mathbb{R}_{++}$, and hence

$$\big(\theta^{-1}Y(\theta t), \theta^{-1}X(\theta t)\big)_{t\in\mathbb{R}_+} \xrightarrow{\mathcal{L}} (\mathcal{Y}(t), \mathcal{X}(t))_{t\in\mathbb{R}_+} \qquad \text{as } \theta \to \infty$$

in $\mathsf{D}(\mathbb{R}_+, \mathbb{R}_+ \times \mathbb{R}^d)$, where $(\mathcal{Y}(t), \mathcal{X}(t))_{t\in\mathbb{R}_+}$ is a $(1+d)$-dimensional affine process on $\mathbb{R}_+ \times \mathbb{R}^d$ with admissible parameters $(0, \widetilde{\alpha}, \widetilde{b}, 0, 0, 0)$, where $\widetilde{\alpha}$ and $\widetilde{b}$ are given in Theorem 5.2.9.

(ii) If $(Y(t), X(t))_{t\in\mathbb{R}_+}$ is a $(1 + d)$-dimensional affine process on $\mathbb{R}_+ \times \mathbb{R}^d$ with $(Y(0), X(0)) = (0, 0)$ and with admissible parameters $(0, \alpha, b, 0, 0, 0)$, then

$$\big(\theta^{-1}Y(\theta t), \theta^{-1}X(\theta t)\big)_{t\in\mathbb{R}_+} \stackrel{\mathcal{L}}{=} (Y(t), X(t))_{t\in\mathbb{R}_+} \qquad \text{for all } \theta \in \mathbb{R}_{++},$$

where $\stackrel{\mathcal{L}}{=}$ denotes equality in distribution. Indeed, by Proposition 1.6 on page 161 in Ethier and Kurtz [**66**], it is enough to check that the semigroups (on the Banach space of bounded Borel measurable functions on $\mathbb{R}_+ \times \mathbb{R}^d$) corresponding to the processes in question coincide. By the definition of a core, this follows from the equality of the infinitesimal generators of the processes in question on the core $\mathcal{C}_c^\infty(\mathbb{R}_+ \times \mathbb{R}^d)$, which has been shown in the proof of Theorem 5.2.9. □

Next we present a corollary of Theorem 5.2.9 which states weak convergence of appropriately normalized one-dimensional continuous state branching processes with immigration. Our corollary generalizes Theorem 2.3 in Huang et al. [**91**] in the sense that we do not have to suppose that $\int_1^\infty \xi^2\, m(\mathrm{d}\xi) < \infty$, only that $\int_1^\infty \xi\, m(\mathrm{d}\xi) < \infty$ (with the notations of Huang et al. [**91**]), and our proof defers from that of Huang et al. [**91**].

COROLLARY 5.2.13. *For all $\theta \in \mathbb{R}_{++}$, let $(Y^{(\theta)}(t))_{t\in\mathbb{R}_+}$ be a one-dimensional continuous state branching process with immigration on $\mathbb{R}_+$ with branching mechanism*

$$R^{(\theta)}(z) := \beta^{(\theta)}z + \alpha^{(\theta)}z^2 + \int_{\mathbb{R}_+} (\mathrm{e}^{-zu} - 1 + zu)\, p(\mathrm{d}u), \qquad z \in \mathbb{R}_+,$$

and with immigration mechanism

$$F^{(\theta)}(z) := b^{(\theta)}z + \int_{\mathbb{R}_+} (1 - \mathrm{e}^{-zu})\, n(\mathrm{d}u), \qquad z \in \mathbb{R}_+,$$

where $\alpha^{(\theta)} \geqslant 0$, $b^{(\theta)} \geqslant 0$, $\beta^{(\theta)} \in \mathbb{R}$ *and* n *and* p *are measures on* $(0, \infty)$ *such that*

$$\int_{\mathbb{R}_+} u \, n(\mathrm{d}u) < \infty \qquad \text{and} \qquad \int_{\mathbb{R}_+} u^2 \, p(\mathrm{d}u) < \infty.$$

Let $\alpha, b, \beta \in \mathbb{R}$, *and let* $(Y(t))_{t \in \mathbb{R}_+}$ *be a one-dimensional continuous state branching process with immigration on* $\mathbb{R}_+$ *with branching mechanism*

$$R(z) := -\beta z + \left(\alpha + \frac{1}{2} \int_{\mathbb{R}_+} u^2 \, p(\mathrm{d}u) \right) z^2, \qquad z \in \mathbb{R}_+,$$

and with immigration mechanism

$$F(z) := \left(b + \int_{\mathbb{R}_+} u \, n(\mathrm{d}u) \right) z, \qquad z \in \mathbb{R}_+.$$

If

$$\lim_{\theta \to \infty} \alpha^{(\theta)} = \alpha, \qquad \lim_{\theta \to \infty} b^{(\theta)} = b, \qquad \lim_{\theta \to \infty} \theta \beta^{(\theta)} = \beta, \qquad Y^{(\theta)}(0) \xrightarrow{\mathcal{L}} Y(0)$$

as $\theta \to \infty$, *then*

$$\left(\theta^{-1} Y^{(\theta)}(\theta t) \right)_{t \in \mathbb{R}_+} \xrightarrow{\mathcal{L}} (Y(t))_{t \in \mathbb{R}_+} \qquad \text{as } \theta \to \infty$$

in $\mathsf{D}(\mathbb{R}_+, \mathbb{R}_+)$.

5.3. A two-dimensional affine diffusion process

From this section, continuous time stochastic processes will be written as $(\xi_t)_{t \in \mathbb{R}_+}$ instead of $(\xi(t))_{t \in \mathbb{R}_+}$. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions, i.e., $(\Omega, \mathcal{F}, \mathbb{P})$ is complete, the filtration $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ is right-continuous and $\mathcal{F}_0$ contains all the $\mathbb{P}$-null sets in $\mathcal{F}$. Let $(W_t)_{t \in \mathbb{R}_+}$ and $(B_t)_{t \in \mathbb{R}_+}$ be independent standard $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$-Wiener processes. Let us consider the following two-dimensional diffusion process given by the SDE

$$\text{(5.3.1)} \qquad \begin{cases} \mathrm{d}Y_t = (a - bY_t) \, \mathrm{d}t + \sqrt{Y_t} \, \mathrm{d}W_t, \\ \mathrm{d}X_t = (m - \theta X_t) \, \mathrm{d}t + \sqrt{Y_t} \, \mathrm{d}B_t, \end{cases} \qquad t \in \mathbb{R}_+,$$

where $a \in \mathbb{R}_{++}$ and $b, \theta, m \in \mathbb{R}$.

The next proposition is about the existence and uniqueness of a strong solution of the SDE (5.3.1).

Proposition 5.3.1. *Let* (η, ζ) *be a random vector independent of* $(W_t, B_t)_{t \in \mathbb{R}_+}$ *satisfying* $\mathbb{P}(\eta \geqslant 0) = 1$. *Then, for all* $a \in \mathbb{R}_{++}$ *and* $b, m, \theta \in \mathbb{R}$, *there is a (pathwise) unique strong solution* $(Y_t, X_t)_{t \in \mathbb{R}_+}$ *of the SDE* (5.3.1) *such that* $\mathbb{P}((Y_0, X_0) = (\eta, \zeta)) = 1$ *and* $\mathbb{P}(Y_t \geqslant 0 \text{ for all } t \in \mathbb{R}_+) = 1$. *Further, for all* $0 \leqslant s \leqslant t$,

$$\text{(5.3.2)} \qquad Y_t = \mathrm{e}^{-b(t-s)} \left(Y_s + a \int_s^t \mathrm{e}^{-b(s-u)} \, \mathrm{d}u + \int_s^t \mathrm{e}^{-b(s-u)} \sqrt{Y_u} \, \mathrm{d}W_u \right),$$

and

$$\text{(5.3.3)} \qquad X_t = \mathrm{e}^{-\theta(t-s)} \left(X_s + m \int_s^t \mathrm{e}^{-\theta(s-u)} \, \mathrm{d}u + \int_s^t \mathrm{e}^{-\theta(s-u)} \sqrt{Y_u} \, \mathrm{d}B_u \right).$$

Note that it is the assumption $a \in \mathbb{R}_{++}$ that ensures $\mathbb{P}(Y_t \geqslant 0, \ \forall \, t \in \mathbb{R}_+) = 1$.

Next we present a result about the first moment of $(Y_t, X_t)_{t \in \mathbb{R}_+}$.

Proposition 5.3.2. *Let* $(Y_t, X_t)_{t\in\mathbb{R}_+}$ *be a strong solution of the SDE* (5.3.1) *satisfying* $\mathbb{P}(Y_0 \geqslant 0) = 1$, $\mathbb{E}(Y_0) < \infty$, *and* $\mathbb{E}(|X_0|) < \infty$. *Then*

$$\begin{bmatrix}\mathbb{E}(Y_t)\\ \mathbb{E}(X_t)\end{bmatrix} = \begin{bmatrix}\mathrm{e}^{-bt} & 0\\ 0 & \mathrm{e}^{-\theta t}\end{bmatrix}\begin{bmatrix}\mathbb{E}(Y_0)\\ \mathbb{E}(X_0)\end{bmatrix} + \begin{bmatrix}\int_0^t \mathrm{e}^{-bs}\,\mathrm{d}s & 0\\ 0 & \int_0^t \mathrm{e}^{-\theta s}\,\mathrm{d}s\end{bmatrix}\begin{bmatrix}a\\ m\end{bmatrix}, \qquad t \in \mathbb{R}_+,$$

Next we show that the process $(Y_t, X_t)_{t\in\mathbb{R}_+}$ given by the SDE (5.3.1) is an affine process.

Proposition 5.3.3. *Let* $(Y_t, X_t)_{t\in\mathbb{R}_+}$ *be a strong solution of the SDE* (5.3.1) *satisfying* $\mathbb{P}(Y_0 \geqslant 0) = 1$. *Then* $(Y_t, X_t)_{t\in\mathbb{R}_+}$ *is an affine process with infinitesimal generator*

$$(\mathcal{A}_{(Y,X)}f)(x) = (a - bx_1)f_1'(x) + (m - \theta x_2)f_2'(x) + \frac{1}{2}x_1(f_{1,1}''(x) + f_{2,2}''(x)) \tag{5.3.4}$$

for $x = (x_1, x_2) \in \mathbb{R}_+ \times \mathbb{R}$ *and* $f \in \mathcal{C}_c^2(\mathbb{R}_+ \times \mathbb{R})$.

By Proposition 5.3.3, the process $(Y_t, X_t)_{t\in\mathbb{R}_+}$ given by (5.3.1) is a two-dimensional affine process with admissible parameters

$$\left(\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}, \frac{1}{2}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, \begin{bmatrix}a\\ m\end{bmatrix}, \begin{bmatrix}-b & 0\\ 0 & -\theta\end{bmatrix}, 0, 0\right).$$

In what follows we define and study criticality of the affine process given by the SDE (5.3.1).

Definition 5.3.4. Let $(Y_t, X_t)_{t\in\mathbb{R}_+}$ be an affine diffusion process given by the SDE (5.3.1) satisfying $\mathbb{P}(Y_0 \geqslant 0) = 1$. We call $(Y_t, X_t)_{t\in\mathbb{R}_+}$ subcritical, critical or supercritical if the spectral radius of the matrix

$$\begin{pmatrix}\mathrm{e}^{-bt} & 0\\ 0 & \mathrm{e}^{-\theta t}\end{pmatrix}$$

is less than 1, equal to 1 or greater than 1, respectively.

Note that, since the spectral radius of the matrix given in Definition 5.3.4 is $\max(\mathrm{e}^{-bt}, \mathrm{e}^{-\theta t})$, the affine process given in Definition 5.3.4 is

$$\begin{aligned} &\text{subcritical} && \text{if } b > 0 \text{ and } \theta > 0,\\ &\text{critical} && \text{if } b = 0,\ \theta \geqslant 0 \text{ or } b \geqslant 0,\ \theta = 0,\\ &\text{supercritical} && \text{if } b < 0 \text{ or } \theta < 0. \end{aligned}$$

Definition 5.3.4 of criticality is in accordance with the corresponding definition for one-dimensional continuous state branching processes, see, e.g., Li [**115**, page 58].

In this section we will always suppose that

$$\textbf{Condition (C):}\quad \begin{aligned} &(b,\theta) = (0,0),\ \mathbb{P}(Y_0 \geqslant 0) = 1,\\ &\mathbb{E}(Y_0) < \infty, \text{ and } \mathbb{E}(X_0^2) < \infty. \end{aligned}$$

For some explanations why we study only this special case, see Remarks 5.4.2, 5.4.3 and 5.5.2. In the next sections under Condition (C) we will study asymptotic behaviour of least squares estimator of θ and (θ, m), respectively. Before doing so we recall some critical models both in discrete and continuous time.

In general, parameter estimation for critical models has a long history. A common feature of the estimators for parameters of critical models is that one may prove weak limit theorems for them by using norming factors that are usually different from the norming factors for the subcritical and supercritical models. Further, it may happen that one has to use different norming factors for two different critical cases.

We recall some discrete time critical models. If $(\xi_k)_{k\in\mathbb{Z}_+}$ is an AR(1) process, i.e., $\xi_k = \varrho\xi_{k-1} + \zeta_k$, $k \in \mathbb{N}$, with $\xi_0 = 0$ and an i.i.d. sequence $(\zeta_k)_{k\in\mathbb{N}}$ having mean 0 and positive variance, then the (ordinary) least squares estimator of the so-called stability parameter ϱ based on the sample $\xi_1, \dots, \xi_n$ takes the form

$$\widetilde{\varrho}_n = \frac{\sum_{k=1}^n \xi_{k-1}\xi_k}{\sum_{k=1}^n \xi_k^2}, \qquad n \in \mathbb{N},$$

see, e.g., Hamilton [**81**, 17.4.2]. In the critical case, i.e., when $\varrho = 1$, by Hamilton [**81**, 17.4.7],

$$n(\widetilde{\varrho}_n - 1) \xrightarrow{\mathcal{L}} \frac{\int_0^1 \mathcal{W}_t \,\mathrm{d}\mathcal{W}_t}{\int_0^1 \mathcal{W}_t^2 \,\mathrm{d}t} \qquad \text{as } n \to \infty,$$

where $(\mathcal{W}_t)_{t\in\mathbb{R}_+}$ is a standard Wiener process and $\xrightarrow{\mathcal{L}}$ denotes convergence in distribution. Here $n(\widetilde{\varrho}_n - 1)$ is known as the Dickey-Fuller statistics. We emphasize that the asymptotic behaviour of $\widetilde{\varrho}_n$ is completely different in the subcritical $(|\rho| < 1)$ and supercritical $(|\rho| > 1)$ cases, where it is asymptotically normal and asymptotically Cauchy, respectively, see, e.g., Mann and Wald [**124**], Anderson [**6**] and White [**155**].

For continuous time critical models, we recall that Huang et al. [**91**, Theorem 2.4] studied asymptotic behaviour of weighted conditional least squares estimator of the drift parameters for discretely observed continuous time critical branching processes with immigration given by

$$\begin{aligned}
\widetilde{Y}_t = \widetilde{Y}_0 &+ \int_0^t (a + b\widetilde{Y}_s)\,\mathrm{d}s + \sigma\int_0^t \sqrt{\widetilde{Y}_s}\,\mathrm{d}\mathcal{W}_s + \int_0^t \int_{[0,\infty)} \xi\,\mathcal{N}_0(\mathrm{d}s, \mathrm{d}\xi) \\
&+ \int_0^t \int_0^{\widetilde{Y}_{s-}} \int_{[0,\infty)} \xi\,(\mathcal{N}_1(\mathrm{d}s, \mathrm{d}u, \mathrm{d}\xi) - \mathrm{d}s\,\mathrm{d}u\,p(\mathrm{d}\xi)), \qquad t \in \mathbb{R}_+,
\end{aligned}$$

where $\widetilde{Y}_0 \geqslant 0$, $a \geqslant 0$, $b \in \mathbb{R}$, $\sigma \geqslant 0$, $\mathcal{W}$ is a standard Wiener process, $\mathcal{N}_0(\mathrm{d}s, \mathrm{d}\xi)$ is a Poisson random measure on $(0,\infty)\times[0,\infty)$ with intensity $\mathrm{d}s\,n(\mathrm{d}\xi)$, $\mathcal{N}_1(\mathrm{d}s, \mathrm{d}u, \mathrm{d}\xi)$ is a Poisson random measure on $(0,\infty)\times(0,\infty)\times[0,\infty)$ with intensity $\mathrm{d}s\,\mathrm{d}u\,p(\mathrm{d}\xi)$ such that the σ-finite measures n and p are supported by $(0,\infty)$ and

$$\int_0^\infty \xi\,n(\mathrm{d}\xi) + \int_0^\infty \xi \wedge \xi^2\,p(\mathrm{d}\xi) < \infty.$$

Our technique differs from that of Huang et al. [**91**] and for completeness we note that the limit distribution and some parts of the proof of their Theorem 2.4 suffer from some misprints. Furthermore, Hu and Long [**90**] studied the problem of parameter estimation for critical mean-reverting α-stable motions

$$\mathrm{d}\widetilde{X}_t = (m - \theta\widetilde{X}_t)\,\mathrm{d}t + \mathrm{d}Z_t, \qquad t \in \mathbb{R}_+,$$

where Z is an α-stable Lévy motion with $\alpha \in (0,2))$ observed at discrete instants. A least squares estimator is obtained and its asymptotics is discussed in the singular case $(m, \theta) = (0, 0)$. We note that the forms of the limit distributions of least squares estimators for critical two-dimensional affine diffusion processes in our Theorems 5.4.1 and 5.5.1 are the same as that of the limit distributions in Theorems 3.2 and 4.1 in Hu and Long [**90**], respectively. We also recall that Hu and Long [**89**] considered the problem of parameter estimation not only for critical mean-reverting α-stable motions, but also for some subcritical ones $(m = 0$ and $\theta > 0)$ by proving limit theorems for the least squares estimators that are completely different from the ones in the critical case. Huang et al. [**91**] investigated the asymptotic behaviour of weighted conditional least squares estimator of the drift parameters not only for critical continuous time branching processes with immigration, but also for subcritical and supercritical ones.

Using our scaling Theorem 5.2.9 we can only handle a special critical affine diffusion model given by (5.1.1) (for a more detailed discussion, see Remark 5.4.3). The other critical and non-critical cases are under investigation but different techniques are needed.

From this section, we will study least squares and conditional least squares estimation for the SDE (5.3.1).

5.4. Least squares estimator of θ when m is known

The least squares estimator (LSE) of θ based on the observations X_i, $i = 0, 1, \ldots, n$, can be obtained by solving the extremum problem

$$\widetilde{\theta}_n^{\mathrm{LSE}} := \arg\min_{\theta\in\mathbb{R}} \sum_{i=1}^n (X_i - X_{i-1} - (m - \theta X_{i-1}))^2.$$

This definition of LSE of θ can be considered as the counterpart of the one given in Hu and Long [**89**, formula (1.2)] for generalized Ornstein-Uhlenbeck processes driven by α-stable motions, see also Hu and Long [**90**, formulas (3.1) and (4.1)]. For a mathematical motivation of the definition of the LSE of θ, see later on Remark 5.6.1. With the notation $f(\theta) := \sum_{i=1}^n (X_i - X_{i-1} - (m - \theta X_{i-1}))^2$, $\theta \in \mathbb{R}$, the equation $f'(\theta) = 0$ takes the form:

$$2\sum_{i=1}^n (X_i - X_{i-1} - (m - \theta X_{i-1}))X_{i-1} = 0.$$

Hence

$$\left(\sum_{i=1}^n X_{i-1}^2\right)\theta = -\sum_{i=1}^n (X_i - X_{i-1} - m)X_{i-1},$$

i.e.,

$$\begin{aligned} \widetilde{\theta}_n^{\mathrm{LSE}} &= -\frac{\sum_{i=1}^n (X_i - X_{i-1} - m)X_{i-1}}{\sum_{i=1}^n X_{i-1}^2} \\ &= -\frac{\sum_{i=1}^n (X_i - X_{i-1})X_{i-1} - \left(\sum_{i=1}^n X_{i-1}\right) m}{\sum_{i=1}^n X_{i-1}^2} \end{aligned} \tag{5.4.1}$$

provided that $\sum_{i=1}^n X_{i-1}^2 > 0$. Since $f''(\theta) = 2\sum_{i=1}^n X_{i-1}^2$, $\theta \in \mathbb{R}$, we have $\widetilde{\theta}_n^{\mathrm{LSE}}$ is indeed the solution of the extremum problem provided that $\sum_{i=1}^n X_{i-1}^2 > 0$.

THEOREM 5.4.1. *Let us assume that Condition (C) holds. Then, for all* $n \geqslant 2$, $\mathbb{P}(\sum_{i=1}^n X_{i-1}^2 > 0) = 1$, *and there exists a unique LSE* $\widetilde{\theta}_n^{\mathrm{LSE}}$ *which has the form given in* (5.4.1). *Further,*

$$n\widetilde{\theta}_n^{\mathrm{LSE}} \xrightarrow{\mathcal{L}} -\frac{\int_0^1 \mathcal{X}_t \,\mathrm{d}\mathcal{X}_t - m\int_0^1 \mathcal{X}_t \,\mathrm{d}t}{\int_0^1 \mathcal{X}_t^2 \,\mathrm{d}t} \qquad \text{as } n \to \infty, \tag{5.4.2}$$

where $(\mathcal{X}_t)_{t\in\mathbb{R}_+}$ *is the second coordinate of a two-dimensional affine process* $(\mathcal{Y}_t, \mathcal{X}_t)_{t\in\mathbb{R}_+}$ *given by the unique strong solution of the SDE*

$$\begin{cases} \mathrm{d}\mathcal{Y}_t = a\,\mathrm{d}t + \sqrt{\mathcal{Y}_t}\,\mathrm{d}\mathcal{W}_t, \\ \mathrm{d}\mathcal{X}_t = m\,\mathrm{d}t + \sqrt{\mathcal{Y}_t}\,\mathrm{d}\mathcal{B}_t, \end{cases} \qquad t \in \mathbb{R}_+, \tag{5.4.3}$$

with initial value $(\mathcal{Y}_0, \mathcal{X}_0) = (0, 0)$, *where* $(\mathcal{W}_t)_{t\in\mathbb{R}_+}$ *and* $(\mathcal{B}_t)_{t\in\mathbb{R}_+}$ *are independent standard Wiener processes.*

REMARK 5.4.2. (i) The limit distributions in Theorem 5.4.1 have the same forms as those of the limit distributions in Theorem 3.2 in Hu and Long [**90**].
(ii) The limit distribution of $n\widetilde{\theta}_n^{\mathrm{LSE}}$ as $n \to \infty$ in Theorem 5.4.1 can be written also in the form

$$-\frac{\int_0^1 \mathcal{X}_t \,\mathrm{d}(\mathcal{X}_t - mt)}{\int_0^1 \mathcal{X}_t^2 \,\mathrm{d}t} = -\frac{\int_0^1 \mathcal{X}_t \sqrt{\mathcal{Y}_t} \,\mathrm{d}\mathcal{B}_t}{\int_0^1 \mathcal{X}_t^2 \,\mathrm{d}t}.$$

(iii) By Proposition 5.3.3, the affine process $(\mathcal{Y}_t, \mathcal{X}_t)_{t\in\mathbb{R}_+}$ given in Theorem 5.4.1 has infinitesimal generator

$$(\mathcal{A}_{(\mathcal{Y},\mathcal{X})} f)(x) = \frac{1}{2} x_1 f_{1,1}''(x) + \frac{1}{2} x_1 f_{2,2}''(x) + a f_1'(x) + m f_2'(x)$$

where $x = (x_1, x_2) \in \mathbb{R}_+ \times \mathbb{R}$ and $f \in \mathcal{C}_c^2(\mathbb{R}_+ \times \mathbb{R})$.
(iv) Under the Condition (C), by Theorem 5.4.1 and Slutsky's lemma, we get $\widetilde{\theta}_n^{\mathrm{LSE}}$ converges stochastically to the parameter $\theta = 0$ as $n \to \infty$. □

REMARK 5.4.3. If the affine diffusion process given by the SDE (5.3.1) is critical but $(b, \theta) \neq (0, 0)$ (i.e., $b = 0$, $\theta > 0$ or $b > 0$, $\theta = 0$), then the asymptotic behaviour of the LSE $\widetilde{\theta}_n^{\mathrm{LSE}}$ cannot be studied using Theorem 5.2.9 since its condition $\lim_{\theta\to\infty} \theta\beta^{(\theta)} = \beta$ is not satisfied. □

5.5. Least squares estimator of (θ, m)

The LSE of (θ, m) based on the observations X_i, $i = 0, 1, \ldots, n$, can be obtained by solving the extremum problem

$$(\widehat{\theta}_n^{\mathrm{LSE}}, \widehat{m}_n^{\mathrm{LSE}}) := \operatorname*{arg\,min}_{(\theta,m)\in\mathbb{R}^2} \sum_{i=1}^n (X_i - X_{i-1} - (m - \theta X_{i-1}))^2.$$

We need to solve the following system of equations with respect to (θ, m):

$$2\sum_{i=1}^n (X_i - X_{i-1} - (m - \theta X_{i-1}))X_{i-1} = 0,$$
$$2\sum_{i=1}^n (X_i - X_{i-1} - (m - \theta X_{i-1})) = 0,$$

which can be written also in the form

$$\begin{bmatrix} \sum_{i=1}^n X_{i-1}^2 & -\sum_{i=1}^n X_{i-1} \\ -\sum_{i=1}^n X_{i-1} & n \end{bmatrix} \begin{bmatrix} \theta \\ m \end{bmatrix} = \begin{bmatrix} -\sum_{i=1}^n (X_i - X_{i-1})X_{i-1} \\ \sum_{i=1}^n (X_i - X_{i-1}) \end{bmatrix}.$$

Then one can check that

$$\widehat{\theta}_n^{\mathrm{LSE}} = -\frac{n\sum_{i=1}^n (X_i - X_{i-1})X_{i-1} - \sum_{i=1}^n X_{i-1} \sum_{i=1}^n (X_i - X_{i-1})}{n\sum_{i=1}^n X_{i-1}^2 - \left(\sum_{i=1}^n X_{i-1}\right)^2}, \tag{5.5.1}$$

and

$$\widehat{m}_n^{\mathrm{LSE}} = \frac{\sum_{i=1}^n X_{i-1}^2 \sum_{i=1}^n (X_i - X_{i-1}) - \sum_{i=1}^n X_{i-1} \sum_{i=1}^n (X_i - X_{i-1})X_{i-1}}{n\sum_{i=1}^n X_{i-1}^2 - \left(\sum_{i=1}^n X_{i-1}\right)^2} \tag{5.5.2}$$

provided that $n\sum_{i=1}^n X_{i-1}^2 - \left(\sum_{i=1}^n X_{i-1}\right)^2 > 0$. Since the matrix

$$\begin{bmatrix} 2\sum_{i=1}^n X_{i-1}^2 & -2\sum_{i=1}^n X_{i-1} \\ -2\sum_{i=1}^n X_{i-1} & 2n \end{bmatrix}$$

which consists of the second order partial derivatives of the function $\mathbb{R}^2 \ni (\theta, m) \mapsto \sum_{i=1}^n (X_i - X_{i-1} - (m - \theta X_{i-1}))^2$ is positive definite provided that $n\sum_{i=1}^n X_{i-1}^2 -$

$\big(\sum_{i=1}^n X_{i-1}\big)^2 > 0$, we have $(\widehat{\theta}_n^{\mathrm{LSE}}, \widehat{m}_n^{\mathrm{LSE}})$ is indeed the solution of the extremum problem provided that $n\sum_{i=1}^n X_{i-1}^2 - \big(\sum_{i=1}^n X_{i-1}\big)^2 > 0$.

Theorem 5.5.1. *Let us assume that Condition (C) holds. Then*

$$\mathbb{P}\left(n\sum_{i=1}^n X_{i-1}^2 - \left(\sum_{i=1}^n X_{i-1}\right)^2 > 0\right) = 1 \qquad \text{for all } n \geqslant 2, \tag{5.5.3}$$

and there exists a unique LSE $(\widehat{\theta}_n^{\mathrm{LSE}}, \widehat{m}_n^{\mathrm{LSE}})$ *which has the form given in* (5.5.1) *and* (5.5.2). *Further,*

$$n\widehat{\theta}_n^{\mathrm{LSE}} \xrightarrow{\mathcal{L}} -\frac{\int_0^1 \mathcal{X}_t \,\mathrm{d}\mathcal{X}_t - \mathcal{X}_1\int_0^1 \mathcal{X}_t\,\mathrm{d}t}{\int_0^1 \mathcal{X}_t^2\,\mathrm{d}t - \left(\int_0^1 \mathcal{X}_t\,\mathrm{d}t\right)^2} \qquad \text{as } n\to\infty,$$

and

$$\widehat{m}_n^{\mathrm{LSE}} \xrightarrow{\mathcal{L}} \frac{\mathcal{X}_1\int_0^1 \mathcal{X}_t^2\,\mathrm{d}t - \int_0^1 \mathcal{X}_t\,\mathrm{d}t\int_0^1 \mathcal{X}_t\,\mathrm{d}\mathcal{X}_t}{\int_0^1 \mathcal{X}_t^2\,\mathrm{d}t - \left(\int_0^1 \mathcal{X}_t\,\mathrm{d}t\right)^2} \qquad \text{as } n\to\infty,$$

where $(\mathcal{X}_t)_{t\in\mathbb{R}_+}$ *is the second coordinate of a two-dimensional affine process* $(\mathcal{Y}_t,\mathcal{X}_t)_{t\in\mathbb{R}_+}$ *given by the unique strong solution of the SDE*

$$\begin{cases} \mathrm{d}\mathcal{Y}_t = a\,\mathrm{d}t + \sqrt{\mathcal{Y}_t}\,\mathrm{d}\mathcal{W}_t, \\ \mathrm{d}\mathcal{X}_t = m\,\mathrm{d}t + \sqrt{\mathcal{Y}_t}\,\mathrm{d}\mathcal{B}_t, \end{cases} \qquad t\in\mathbb{R}_+,$$

with initial value $(\mathcal{Y}_0,\mathcal{X}_0) = (0,0)$, *where* $(\mathcal{W}_t)_{t\in\mathbb{R}_+}$ *and* $(\mathcal{B}_t)_{t\in\mathbb{R}_+}$ *are independent standard Wiener processes.*

Remark 5.5.2. (i) The limit distributions in Theorem 5.5.1 have the same forms as those of the limit distributions in Theorem 4.1 in Hu and Long [**90**].
(ii) By Proposition 5.3.3, the affine process $(\mathcal{Y}_t,\mathcal{X}_t)_{t\in\mathbb{R}_+}$ given in Theorem 5.5.1 has infinitesimal generator

$$(\mathcal{A}_{(\mathcal{Y},\mathcal{X})}f)(x) = \frac{1}{2}x_1 f_{1,1}''(x) + \frac{1}{2}x_1 f_{2,2}''(x) + af_1'(x) + mf_2'(x),$$

where $x=(x_1,x_2)\in\mathbb{R}_+\times\mathbb{R}$ and $f\in\mathcal{C}_c^2(\mathbb{R}_+\times\mathbb{R})$.
(iii) Under the Condition (C), by Theorem 5.5.1 and Slutsky's lemma, we get $\widehat{\theta}_n^{\mathrm{LSE}}$ converges stochastically to the parameter $\theta=0$ as $n\to\infty$, and one can show that $\widehat{m}_n^{\mathrm{LSE}}$ does not converge stochastically to the parameter m as $n\to\infty$, see Remark 5.6.4. □

5.6. Conditional least squares estimator of (θ, m)

For all $t\in\mathbb{R}_+$, let $\mathcal{F}_t^{(Y,X)}$ be the σ-algebra generated by $(Y_s,X_s)_{s\in[0,t]}$. The conditional least squares estimator (CLSE) of (θ,m) based on the observations X_i, $i=0,1,\ldots,n$, can be obtained by solving the extremum problem

$$(\widehat{\theta}_n^{\mathrm{CLSE}}, \widehat{m}_n^{\mathrm{CLSE}}) := \operatorname*{arg\,min}_{(\theta,m)\in\mathbb{R}^2} \sum_{i=1}^n \Big(X_i - \mathbb{E}\big(X_i \,|\, (Y_j,X_j),\, j=0,\ldots,i-1\big)\Big)^2.$$

By (5.3.3), for all $(y_0,x_0)\in\mathbb{R}_+\times\mathbb{R}$, we have

$$\mathbb{E}\big(X_t \,|\, (Y_0,X_0)=(y_0,x_0)\big) = \mathrm{e}^{-\theta t}x_0 + m\int_0^t \mathrm{e}^{-\theta(t-u)}\,\mathrm{d}u, \qquad t\geqslant 0,$$

where we used that $\left(\int_0^t \mathrm{e}^{\theta u}\sqrt{Y_u}\,\mathrm{d}B_u\right)_{t\in\mathbb{R}_+}$ is a martingale (which follows by the proof of Proposition 5.3.2). Using that $(Y_t, X_t)_{t\in\mathbb{R}_+}$ is a time-homogeneous Markov process, we have

$$\mathbb{E}(X_t \,|\, \mathcal{F}_s^{(Y,X)}) = \mathbb{E}(X_t \,|\, (Y_s, X_s)) = \mathrm{e}^{-\theta(t-s)}X_s + m\int_s^t \mathrm{e}^{-\theta(t-u)}\,\mathrm{d}u$$

for $0 \leqslant s \leqslant t$. Then

$$\begin{aligned}
X_i - \mathbb{E}\left(X_i \,|\, (Y_j, X_j),\, j = 0,\dots,i-1\right) &= X_i - \mathbb{E}(X_i \,|\, \mathcal{F}_{i-1}^{(Y,X)}) \\
&= X_i - \mathrm{e}^{-\theta}X_{i-1} - m\int_{i-1}^{i} \mathrm{e}^{-\theta(i-u)}\,\mathrm{d}u = X_i - \mathrm{e}^{-\theta}X_{i-1} - m\int_0^1 \mathrm{e}^{-\theta v}\,\mathrm{d}v \\
&= X_i - \gamma X_{i-1} - \delta, \qquad i = 1,\dots,n,
\end{aligned}$$

where

$$\gamma := \mathrm{e}^{-\theta} \qquad \text{and} \qquad \delta := m\int_0^1 \mathrm{e}^{-\theta v}\,\mathrm{d}v = \begin{cases} m\frac{1-\mathrm{e}^{-\theta}}{\theta} & \text{if } \theta \neq 0, \\ m & \text{if } \theta = 0. \end{cases}$$

The CLSE $(\widehat{\gamma}_n^{\mathrm{CLSE}}, \widehat{\delta}_n^{\mathrm{CLSE}})$ of (γ, δ) on $\mathbb{R}^2$ based on the observations X_i, $i = 0, 1, \dots, n$, can be obtained by solving the extremum problem

$$(\widehat{\gamma}_n^{\mathrm{CLSE}}, \widehat{\delta}_n^{\mathrm{CLSE}}) := \operatorname*{arg\,min}_{(\gamma,\delta)\in\mathbb{R}^2} \sum_{i=1}^n (X_i - \gamma X_{i-1} - \delta)^2. \tag{5.6.1}$$

For the extremum problem (5.6.1), we need to solve the following system of equations with respect to (γ, δ):

$$\begin{aligned}
2\sum_{i=1}^n (X_i - \gamma X_{i-1} - \delta)X_{i-1} &= 0, \\
2\sum_{i=1}^n (X_i - \gamma X_{i-1} - \delta) &= 0,
\end{aligned}$$

which can be written also in the form

$$\begin{bmatrix} \sum_{i=1}^n X_{i-1}^2 & \sum_{i=1}^n X_{i-1} \\ \sum_{i=1}^n X_{i-1} & n \end{bmatrix} \begin{bmatrix} \gamma \\ \delta \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n X_{i-1}X_i \\ \sum_{i=1}^n X_i \end{bmatrix}.$$

Then

$$\widehat{\gamma}_n^{\mathrm{CLSE}} = \frac{n\sum_{i=1}^n X_{i-1}X_i - \sum_{i=1}^n X_{i-1}\sum_{i=1}^n X_i}{n\sum_{i=1}^n X_{i-1}^2 - \left(\sum_{i=1}^n X_{i-1}\right)^2}, \tag{5.6.2}$$

and

$$\widehat{\delta}_n^{\mathrm{CLSE}} = \frac{\sum_{i=1}^n X_{i-1}^2\sum_{i=1}^n X_i - \sum_{i=1}^n X_{i-1}\sum_{i=1}^n X_{i-1}X_i}{n\sum_{i=1}^n X_{i-1}^2 - \left(\sum_{i=1}^n X_{i-1}\right)^2}, \tag{5.6.3}$$

provided that $n\sum_{i=1}^n X_{i-1}^2 - \left(\sum_{i=1}^n X_{i-1}\right)^2 \neq 0$. Since the matrix

$$\begin{bmatrix} 2\sum_{i=1}^n X_{i-1}^2 & 2\sum_{i=1}^n X_{i-1} \\ 2\sum_{i=1}^n X_{i-1} & 2n \end{bmatrix}$$

consisting of the second order partial derivatives of the function $\mathbb{R}^2 \ni (\gamma, \delta) \mapsto \sum_{i=1}^n (X_i - \gamma X_{i-1} - \delta)^2$ is positive definite provided that

$$n\sum_{i=1}^n X_{i-1}^2 - \left(\sum_{i=1}^n X_{i-1}\right)^2 > 0,$$

we have $(\widehat{\gamma}_n^{\mathrm{CLSE}}, \widehat{\delta}_n^{\mathrm{CLSE}})$ is indeed the solution of the extremum problem provided that $n\sum_{i=1}^n X_{i-1}^2 - \left(\sum_{i=1}^n X_{i-1}\right)^2 > 0$. By the same technique as in the proof of Theorem 5.5.1, we get

$$n(\widehat{\gamma}_n^{\mathrm{CLSE}} - 1) \xrightarrow{\mathcal{L}} \frac{\int_0^1 \mathcal{X}(t)\,\mathrm{d}\mathcal{X}(t) - \mathcal{X}(1)\int_0^1 \mathcal{X}(t)\,\mathrm{d}t}{\int_0^1 \mathcal{X}(t)^2\,\mathrm{d}t - \left(\int_0^1 \mathcal{X}(t)\,\mathrm{d}t\right)^2} \qquad \text{as } n\to\infty,$$

and

$$\widehat{\delta}_n^{\mathrm{CLSE}} \xrightarrow{\mathcal{L}} \frac{\mathcal{X}(1)\int_0^1 \mathcal{X}(t)^2\,\mathrm{d}t - \int_0^1 \mathcal{X}(t)\,\mathrm{d}t\int_0^1 \mathcal{X}(t)\,\mathrm{d}\mathcal{X}(t)}{\int_0^1 \mathcal{X}(t)^2\,\mathrm{d}t - \left(\int_0^1 \mathcal{X}(t)\,\mathrm{d}t\right)^2} \qquad \text{as } n\to\infty.$$

By Slutsky's lemma, we also have $\widehat{\gamma}_n^{\mathrm{CLSE}} \xrightarrow{\mathbb{P}} 1$ as $n\to\infty$. Hence with probability converging to one we can introduce the estimators $\widehat{\theta}_n^{\mathrm{CLSE}}$ and $\widehat{m}_n^{\mathrm{CLSE}}$ in a way that

$$\begin{aligned}\widehat{\gamma}_n^{\mathrm{CLSE}} &= \mathrm{e}^{-\widehat{\theta}_n^{\mathrm{CLSE}}},\\ \widehat{\delta}_n^{\mathrm{CLSE}} &= \widehat{m}_n^{\mathrm{CLSE}} \int_0^1 \mathrm{e}^{-\widehat{\theta}_n^{\mathrm{CLSE}} v}\,\mathrm{d}v.\end{aligned} \tag{5.6.4}$$

Since the function $A:\mathbb{R}^2\to\mathbb{R}_+\times\mathbb{R}$,

$$\mathbb{R}^2 \ni (\theta', m') \mapsto A(\theta', m') := \begin{bmatrix}\gamma'\\ \delta'\end{bmatrix} =: \begin{bmatrix}\mathrm{e}^{-\theta'}\\ m'\int_0^1 \mathrm{e}^{-\theta' v}\,\mathrm{d}v\end{bmatrix} \in \mathbb{R}_+\times\mathbb{R}$$

is bijective and measurable, and for all $n\in\mathbb{N}$, $(x_0, x_1, \dots, x_n)\in\mathbb{R}^{n+1}$ and $(\gamma', \delta')\in\mathbb{R}_+\times\mathbb{R}$,

$$\begin{aligned}\sum_{i=1}^n (x_i - \gamma' x_{i-1} - \delta')^2 &= \sum_{i=1}^n \left(x_i - \begin{bmatrix}\gamma'\\ \delta'\end{bmatrix}^\top \begin{bmatrix}x_{i-1}\\ 1\end{bmatrix}\right)^2\\ &= \sum_{i=1}^n \left(x_i - (A(\theta', m'))^\top \begin{bmatrix}x_{i-1}\\ 1\end{bmatrix}\right)^2,\end{aligned}$$

we have there is a bijection between the set of CLSEs of the parameters (θ, m) on $\mathbb{R}^2$ and the set of CLSEs of the parameters $A(\theta, m)$ on $\mathbb{R}_+\times\mathbb{R}$. Consequently, with probability converging to one, we have $(\widehat{\theta}_n^{\mathrm{CLSE}}, \widehat{m}_n^{\mathrm{CLSE}})$ is a CLSE of (θ, m).

Remark 5.6.1. Using the definition of CLSE of (θ, m) we give a mathematical motivation of the definition of the LSE $\widetilde{\theta}_n$ of θ introduced in Section 5.4. Note that if $\theta = 0$, then

$$X_i - \mathbb{E}(X_i \mid \mathcal{F}_{i-1}^{(Y,X)}) = X_i - X_{i-1} - m, \quad i = 1, \dots, n.$$

If $\theta\neq 0$, then, by Taylor's theorem, $1-\mathrm{e}^{-\theta} = \mathrm{e}^{-\tau\theta}\theta$ with some $\tau = \tau(\theta)\in[0,1]$, and hence

$$\begin{aligned}X_i - \mathbb{E}(X_i \mid \mathcal{F}_{i-1}^{(Y,X)}) &= X_i - \mathrm{e}^{-\theta} X_{i-1} - m\int_0^1 \mathrm{e}^{-\theta v}\,\mathrm{d}v\\ &= X_i - X_{i-1} + \mathrm{e}^{-\tau\theta}\theta X_{i-1} - m\mathrm{e}^{-\tau\theta}\end{aligned}$$

for $i = 1, \dots, n-1$. Hence for small values of θ one can approximate $X_i - \mathbb{E}(X_i \mid \mathcal{F}_{i-1}^{(Y,X)})$ by $X_i - X_{i-1} + \theta X_{i-1} - m = X_i - X_{i-1} - (m - \theta X_{i-1})$, $i = 1, \dots, n$. Based on this, for small values of θ, in the definition of the LSE of θ, the sum $\sum_{i=1}^n (X_i - X_{i-1} - (m - \theta X_{i-1}))^2$ can be considered as an approximation of the sum $\sum_{i=1}^n (X_i - \mathbb{E}(X_i \mid \mathcal{F}_{i-1}^{(Y,X)}))^2$ in the definition of the CLSE of (θ, m). □

THEOREM 5.6.2. *Let us assume that Condition (C) holds. Then*

$$\mathbb{P}\left(n \sum_{i=1}^{n} X_{i-1}^2 - \left(\sum_{i=1}^{n} X_{i-1} \right)^2 > 0 \right) = 1 \qquad \text{for all } n \geqslant 2, \tag{5.6.5}$$

and there exists a unique CLSE $(\widehat{\theta}_n^{\mathrm{CLSE}}, \widehat{m}_n^{\mathrm{CLSE}})$ *which has the form given in* (5.6.4). *Further,*

$$n\widehat{\theta}_n^{\mathrm{CLSE}} \xrightarrow{\mathcal{L}} -\frac{\int_0^1 \mathcal{X}_t \,\mathrm{d}\mathcal{X}_t - \mathcal{X}_1 \int_0^1 \mathcal{X}_t \,\mathrm{d}t}{\int_0^1 \mathcal{X}_t^2 \,\mathrm{d}t - \left(\int_0^1 \mathcal{X}_t \,\mathrm{d}t\right)^2} \qquad \text{as } n \to \infty, \tag{5.6.6}$$

and

$$\widehat{m}_n^{\mathrm{CLSE}} \xrightarrow{\mathcal{L}} \frac{\mathcal{X}_1 \int_0^1 \mathcal{X}_t^2 \,\mathrm{d}t - \int_0^1 \mathcal{X}_t \,\mathrm{d}t \int_0^1 \mathcal{X}_t \,\mathrm{d}\mathcal{X}_t}{\int_0^1 \mathcal{X}_t^2 \,\mathrm{d}t - \left(\int_0^1 \mathcal{X}_t \,\mathrm{d}t\right)^2} \qquad \text{as } n \to \infty, \tag{5.6.7}$$

where $(\mathcal{X}_t)_{t \in \mathbb{R}_+}$ *is the second coordinate of a two-dimensional affine process* $(\mathcal{Y}_t, \mathcal{X}_t)_{t \in \mathbb{R}_+}$ *given by the unique strong solution of the SDE*

$$\begin{cases} \mathrm{d}\mathcal{Y}_t = a \,\mathrm{d}t + \sqrt{\mathcal{Y}_t} \,\mathrm{d}\mathcal{W}_t, \\ \mathrm{d}\mathcal{X}_t = m \,\mathrm{d}t + \sqrt{\mathcal{Y}_t} \,\mathrm{d}\mathcal{B}_t, \end{cases} \qquad t \in \mathbb{R}_+,$$

with initial value $(\mathcal{Y}_0, \mathcal{X}_0) = (0, 0)$, *where* $(\mathcal{W}_t)_{t \in \mathbb{R}_+}$ *and* $(\mathcal{B}_t)_{t \in \mathbb{R}_+}$ *are independent standard Wiener processes.*

REMARK 5.6.3. (i) We do not consider the CLSE of θ supposing that m is known since the corresponding extremum problem is rather complicated, and from statistical point of view it has less importance.
(ii) Under the Condition (C), by Theorem 5.6.2 and Slutsky's lemma, we get $\widehat{\theta}_n^{\mathrm{CLSE}}$ converges stochastically to the parameter $\theta = 0$ as $n \to \infty$, and one can show that $\widehat{m}_n^{\mathrm{CLSE}}$ does not converge stochastically to the parameter m as $n \to \infty$, see Remark 5.6.4. □

REMARK 5.6.4. Barczy et al. [**11**] contains two additional appendices (Appendices A and B) as well, where we check that the integrals in (5.2.2) are well-defined, i.e., elements of C, under the conditions (v) and (vi) of Definition 5.2.3 (Appendix A), and we show that $\widehat{m}_n^{\mathrm{LSE}}$ and $\widehat{m}_n^{\mathrm{CLSE}}$ do not converge stochastically to the parameter m as $n \to \infty$, respectively (Appendix B). □

CHAPTER 6

Stationarity and ergodicity for an affine two-factor model

Co-authors: Leif Doering, Zenghu Li and Gyula Pap

6.1. Introduction

We consider the following two-dimensional affine process (affine two-factor model)

$$\begin{cases} \mathrm{d}Y_t = (a - bY_t)\,\mathrm{d}t + \sqrt[\alpha]{Y_{t-}}\,\mathrm{d}L_t, & t \geqslant 0, \\ \mathrm{d}X_t = (m - \theta X_t)\,\mathrm{d}t + \sqrt{Y_t}\,\mathrm{d}B_t, & t \geqslant 0, \end{cases} \tag{6.1.1}$$

where $a > 0$, $b, \theta, m \in \mathbb{R}$, $\alpha \in (1, 2]$, $(L_t)_{t\geqslant 0}$ is a spectrally positive α-stable Lévy process with Lévy measure $C_\alpha z^{-1-\alpha}\mathbf{1}_{\{z>0\}}$ with $C_\alpha := (\alpha\Gamma(-\alpha))^{-1}$ (where Γ denotes the Gamma function) in case $\alpha \in (1, 2)$, a standard Wiener process in case $\alpha = 2$, and $(B_t)_{t\geqslant 0}$ is an independent standard Wiener process. Note that in case of $\alpha = 2$, due to the almost sure continuity of the sample paths of a standard Wiener process, instead of $\sqrt{Y_{t-}}$ one can write $\sqrt{Y_t}$ in the first SDE of (6.1.1), and Y is the so-called Cox-Ingersol-Ross (CIR) process; while in case of $\alpha \in (1, 2)$, Y is called the α-root process. Note also that the process $(Y_t)_{t\geqslant 0}$ given by the first SDE of (6.1.1) is a continuous state branching process with immigration with branching mechanism $bz + \frac{1}{\alpha}z^\alpha$, $z \geqslant 0$, and with immigration mechanism az, $z \geqslant 0$ (for more details, see the proof of Theorem 6.3.1, part (i)). Chen and Joslin [**45**] have found several applications of the model (6.1.1) with $\alpha = 2$ in financial mathematics, see their equations (25) and (26).

The process (Y, X) given by (6.1.1) is a special affine process. The set of affine processes contains a large class of important Markov processes such as continuous state branching processes and Orstein-Uhlenbeck processes. Further, a lot of models in financial mathematics are affine such as the Heston model [**84**], the model of Barndorff-Nielsen and Shephard [**32**] or the model due to Carr and Wu [**42**]. A precise mathematical formulation and a complete characterization of regular affine processes are due to Duffie et al. [**63**]. Later several authors have contributed to the theory of general affine processes: to name a few, Andersen and Piterbarg [**5**], Dawson and Li [**53**], Filipović and Mayerhofer [**69**], Glasserman and Kim [**78**], Jena et al. [**97**] and Keller-Ressel et al. [**104**].

This article is devoted to study the existence of a unique stationary distribution and ergodicity of the affine process given by the SDE (6.1.1). These kinds of results are important on their own rights, further they can be used for studying parameter estimation for the given model. For the existing results on ergodicity of affine processes, see the beginning of Section 6.3.

Next we give a brief overview of the structure of the paper. Section 6.2 is devoted to a preliminary discussion of the existence and uniqueness of a strong solution of the SDE (6.1.1) by proving also that this solution is indeed an affine process, see, Theorem 6.2.2. In Section 6.3 we prove the existence of a unique stationary distribution for the affine process given by (6.1.1) in both cases $\alpha \in (1, 2)$

and $\alpha = 2$, provided that $a > 0$, $b > 0$ and $\theta > 0$, see, Theorem 6.3.1. In Section 6.4, in case of $\alpha = 2$, we prove ergodicity of the process in question provided that $a > 0$, $b > 0$ and $\theta > 0$, and we also show that the unique stationary distribution of the process is absolutely continuous, has finite (mixed) moments of any order by calculating some moments explicitly, too, see Theorems 6.4.1 and 6.4.2, respectively.

6.2. The affine two-factor model

Let $\mathbb{N}$, $\mathbb{Z}_+$, $\mathbb{R}$ and $\mathbb{R}_+$ denote the sets of positive integers, non-negative integers, real numbers and non-negative real numbers, respectively. By $\|x\|$ and $\|A\|$ we denote the Euclidean norm of a vector $x \in \mathbb{R}^m$ and the induced matrix norm $\|A\| = \sup\{\|Ax\| : x \in \mathbb{R}^m,\ \|x\| = 1\}$ of a matrix $A \in \mathbb{R}^{n\times m}$, respectively. By $\mathcal{C}^2(\mathbb{R}_+ \times \mathbb{R}, \mathbb{R})$, $\mathcal{C}^2_c(\mathbb{R}_+ \times \mathbb{R}, \mathbb{R})$ and $\mathcal{C}^\infty_c(\mathbb{R}_+ \times \mathbb{R}, \mathbb{R})$, we denote the set of twice continuously differentiable real-valued functions on $\mathbb{R}_+ \times \mathbb{R}$, the set of twice continuously differentiable real-valued functions on $\mathbb{R}_+ \times \mathbb{R}$ with compact support and the set of infinitely differentiable real-valued functions on $\mathbb{R}_+ \times \mathbb{R}$ with compact support, respectively. Convergence in distribution will denoted by $\stackrel{\mathcal{L}}{\longrightarrow}$.

Let $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P}\big)$ be a filtered probability space satisfying the usual conditions, i.e., $(\Omega, \mathcal{F}, \mathbb{P})$ is complete, the filtration $(\mathcal{F}_t)_{t\geqslant 0}$ is right-continuous and $\mathcal{F}_0$ contains all the $\mathbb{P}$-null sets in $\mathcal{F}$. Let $(B_t)_{t\geqslant 0}$ be a standard $(\mathcal{F}_t)_{t\geqslant 0}$-Wiener process and $(L_t)_{t\geqslant 0}$ be a spectrally positive $(\mathcal{F}_t)_{t\geqslant 0}$-stable process with index $\alpha \in (1, 2]$. We assume that B and L are independent. If $\alpha = 2$, we understand that L is a standard $(\mathcal{F}_t)_{t\geqslant 0}$-Wiener process. If $\alpha \in (1, 2)$, we understand that L is a $(\mathcal{F}_t)_{t\geqslant 0}$-Lévy process with Lévy-Khintchine formula

$$\mathbb{E}(\mathrm{e}^{iuL_1}) = \exp\left\{\int_0^\infty (\mathrm{e}^{iuz} - 1 - iuz) C_\alpha z^{-1-\alpha}\,\mathrm{d}z\right\} = \exp\left\{\frac{1}{\alpha}(-iu)^\alpha\right\}, \quad u \in \mathbb{R},$$

where $C_\alpha = (\alpha\Gamma(-\alpha))^{-1}$. Recall that in case of $\alpha \in (1, 2)$ the Lévy-Itô representation of L takes the form

$$L_t = \int_{(0,t]}\int_{(0,\infty)} z\widetilde{N}(\mathrm{d}s, \mathrm{d}z), \quad t \geqslant 0,$$

where $\widetilde{N}(\mathrm{d}s, \mathrm{d}z)$ is a compensated Poisson random measure on $(0, \infty)^2$ with intensity measure $C_\alpha z^{-1-\alpha}\mathbf{1}_{\{z>0\}}\,\mathrm{d}s\,\mathrm{d}z$.

Remark 6.2.1. We shed some light on the definition of the stochastic integral with respect to the spectrally positive α-stable process L in the first SDE of (6.1.1) in case of $\alpha \in (1, 2)$. By Jacod and Shiryaev [**95**, Corollary II.4.19], L is a semimartingale so that Theorems I.4.31 and I.4.40 in Jacod and Shiryaev [**95**] describe the classes of processes which are integrable with respect to L. A more accessible integrability criteria is due to Kallenberg [**98**, Theorem 3.1]. Roughly speaking, a predictable process V is locally integrable with respect to L (i.e., the stochastic integral $\int_0^t V_s\,\mathrm{d}L_s$ exists for all $t \geqslant 0$) if and only if $\int_0^t |V_s|^\alpha\,\mathrm{d}s < \infty$ almost surely for all $t \geqslant 0$. For the construction of stochastic integrals with respect to symmetric α-stable processes, see also Rosinski and Woyczynski [**144**, Theorem 2.1]. Another possible way is to consider the stochastic integral with respect to L as a stochastic integral with respect to a certain compensated Poisson random measure, see the last equality on page 230 in Li [**115**]. □

The next theorem is about the existence and uniqueness of a strong solution of the SDE (6.1.1).

THEOREM 6.2.2. *Let* (η_0, ζ_0) *be a random vector independent of* $(L_t, B_t)_{t\geqslant 0}$ *satisfying* $\mathbb{P}(\eta_0 \geqslant 0) = 1$. *Then for all* $a > 0$, $b, m, \theta \in \mathbb{R}$ *and* $\alpha \in (1, 2]$, *there is a (pathwise) unique strong solution* $(Y_t, X_t)_{t\geqslant 0}$ *of the SDE* (6.1.1) *such that* $\mathbb{P}((Y_0, X_0) = (\eta_0, \zeta_0)) = 1$ *and* $\mathbb{P}(Y_t \geqslant 0,\ \forall\, t \geqslant 0) = 1$. *Further, we have*

$$Y_t = \mathrm{e}^{-b(t-s)} \left(Y_s + a \int_s^t \mathrm{e}^{-b(s-u)}\, \mathrm{d}u + \int_s^t \mathrm{e}^{-b(s-u)} \sqrt[\alpha]{Y_{u-}}\, \mathrm{d}L_u \right) \tag{6.2.1}$$

for $0 \leqslant s \leqslant t$, *and*

$$X_t = \mathrm{e}^{-\theta(t-s)} \left(X_s + m \int_s^t \mathrm{e}^{-\theta(s-u)}\, \mathrm{d}u + \int_s^t \mathrm{e}^{-\theta(s-u)} \sqrt{Y_u}\, \mathrm{d}B_u \right) \tag{6.2.2}$$

for $0 \leqslant s \leqslant t$. *Moreover,* $(Y_t, X_t)_{t\geqslant 0}$ *is a regular affine process with infinitesimal generator*

$$\begin{aligned}(\mathcal{A}f)(y, x) &= (a - by) f_1'(y, x) + (m - \theta x) f_2'(y, x) + \frac{1}{2} y f_{2,2}''(y, x) \\ &\quad + y \int_0^\infty \Big(f(y + z, x) - f(y, x) - z f_1'(y, x) \Big) C_\alpha z^{-1-\alpha}\, \mathrm{d}z\end{aligned} \tag{6.2.3}$$

in case of $\alpha \in (1, 2)$, *and*

$$(\mathcal{A}f)(y, x) = (a - by) f_1'(y, x) + (m - \theta x) f_2'(y, x) + \frac{1}{2} y (f_{1,1}''(y, x) + f_{2,2}''(y, x)) \tag{6.2.4}$$

in case of $\alpha = 2$, *where* $(y, x) \in \mathbb{R}_+ \times \mathbb{R}$, $f \in \mathcal{C}_c^2(\mathbb{R}_+ \times \mathbb{R}, \mathbb{R})$, *and* f_i', $i = 1, 2$, *and* $f_{i,j}''$, $i, j \in \{1, 2\}$, *denote the first and second order partial derivatives of* f *with respect to its* i*-th and* i*-th and* j*-th variables.*

REMARK 6.2.3. Note that in Theorem 6.2.2 it is the assumption $a > 0$ which ensures $\mathbb{P}(Y_t \geqslant 0,\ \forall\, t \geqslant 0) = 1$. □

6.3. Stationarity

The study of existence of stationary distributions for affine processes in general is currently under active research.

In the special case of continuous state branching processes with immigration the question of existence of a unique stationary distribution has been well-studied, see Li [**115**, Theorem 3.20 and Corollary 3.21] or Keller-Ressel and Mijatović [**103**, Theorem 2.6].

Glasserman and Kim [**78**, Theorem 2.4] proved existence of a unique stationary distribution for the process

$$\begin{cases} \mathrm{d}Y_t = (a - bY_t)\, \mathrm{d}t + \sqrt{Y_t}\, \mathrm{d}L_t, & t \geqslant 0, \\ \mathrm{d}X_t = -\theta X_t\, \mathrm{d}t + \sqrt{1 + \sigma Y_t}\, \mathrm{d}B_t, & t \geqslant 0, \end{cases} \tag{6.3.1}$$

where $a > 0, b > 0, \theta > 0, \sigma \geqslant 0$ and L and B are independent standard Wiener processes.

The following result states the existence of a unique stationary distribution of the affine process given by the SDE (6.1.1) for both cases $\alpha \in (1, 2)$ and $\alpha = 2$.

THEOREM 6.3.1. *Let us consider the two-dimensional affine model* (6.1.1) *with* $a > 0$, $b > 0$, $m \in \mathbb{R}$, $\theta > 0$, *and with a random initial value* (η_0, ζ_0) *independent of* $(L_t, B_t)_{t\geqslant 0}$ *satisfying* $\mathbb{P}(\eta_0 \geqslant 0) = 1$. *Then*

(i) $(Y_t, X_t) \xrightarrow{\mathcal{L}} (Y_\infty, X_\infty)$ *as* $t \to \infty$, *and the distribution of* (Y_∞, X_∞) *is given by*

$$\mathbb{E}\left(\mathrm{e}^{-\lambda_1 Y_\infty + i\lambda_2 X_\infty}\right) = \exp\left\{ -a \int_0^\infty v_s(\lambda_1, \lambda_2)\, \mathrm{d}s + i \frac{m}{\theta} \lambda_2 \right\} \tag{6.3.2}$$

for $(\lambda_1, \lambda_2) \in \mathbb{R}_+ \times \mathbb{R}$, *where* $v_t(\lambda_1, \lambda_2)$, $t \geqslant 0$, *is the unique non-negative solution of the (deterministic) differential equation*

$$\begin{cases} \frac{\partial v_t}{\partial t}(\lambda_1, \lambda_2) = -bv_t(\lambda_1, \lambda_2) - \frac{1}{\alpha}(v_t(\lambda_1, \lambda_2))^\alpha + \frac{1}{2}\mathrm{e}^{-2\theta t}\lambda_2^2, & t \geqslant 0, \\ v_0(\lambda_1, \lambda_2) = \lambda_1. \end{cases} \tag{6.3.3}$$

(ii) *supposing that the random initial value* (η_0, ζ_0) *has the same distribution as* (Y_∞, X_∞) *given in part (i), we have* $(Y_t, X_t)_{t\geqslant 0}$ *is strictly stationary.*

6.4. Ergodicity

Such as the existence of a unique stationary distribution, the question of ergodicity for an affine process is also in the focus of current investigations.

Recently, Sandrić [**145**] has proved ergodicity of so called stable-like processes using the same technique that we applied. Further, the ergodicity of the so-called α-root process with $\alpha \in (1, 2]$ (see, the first SDE of (6.1.1)) and some statistical applications were given in Li and Ma [**116**].

The following result states the ergodicity of the affine diffusion process given by the SDE (6.1.1) with $\alpha = 2$.

THEOREM 6.4.1. *Let us consider the two-dimensional affine diffusion model* (6.1.1) *with* $\alpha = 2$, $a > 0$, $b > 0$, $m \in \mathbb{R}$, $\theta > 0$, *and with a random initial value* (η_0, ζ_0) *independent of* $(L_t, B_t)_{t\geqslant 0}$ *satisfying* $\mathbb{P}(\eta_0 \geqslant 0) = 1$. *Then, for all Borel measurable functions* $f : \mathbb{R}^2 \to \mathbb{R}$ *such that* $\mathbb{E}\,|f(Y_\infty, X_\infty)| < \infty$, *we have*

$$\mathbb{P}\left(\lim_{T\to\infty} \frac{1}{T}\int_0^T f(Y_s, X_s)\,\mathrm{d}s = \mathbb{E}\,f(Y_\infty, X_\infty)\right) = 1, \tag{6.4.1}$$

where the distribution of (Y_∞, X_∞) *is given by* (6.3.2) *and* (6.3.3) *with* $\alpha = 2$.

In the next theorem we collected several facts about the limiting random variable (Y_∞, X_∞) given by (6.3.2) and (6.3.3) with $\alpha = 2$.

THEOREM 6.4.2. *The random variable* (Y_∞, X_∞) *given by* (6.3.2) *and* (6.3.3) *with* $\alpha = 2$ *is absolutely continuous, the Laplace transform of* Y_∞ *takes the form*

$$\mathbb{E}(\mathrm{e}^{-\lambda_1 Y_\infty}) = \left(1 + \frac{\lambda_1}{2b}\right)^{-2a}, \qquad \lambda_1 \in \mathbb{R}_+, \tag{6.4.2}$$

yielding that Y_∞ *has Gamma distribution with parameters* $2a$ *and* $2b$. *Further, all the (mixed) moments of* (Y_∞, X_∞) *of any order are finite, i.e., we have* $\mathbb{E}(Y_\infty^n |X_\infty|^p) < \infty$ *for all* $n, p \in \mathbb{Z}_+$, *and especially,*

$$\mathbb{E}(Y_\infty) = \frac{a}{b}, \qquad \mathbb{E}(X_\infty) = \frac{m}{\theta},$$

$$\mathbb{E}(Y_\infty^2) = \frac{a(2a+1)}{2b^2}, \qquad \mathbb{E}(Y_\infty X_\infty) = \frac{ma}{\theta b}, \qquad \mathbb{E}(X_\infty^2) = \frac{a\theta + 2bm^2}{2b\theta^2},$$

$$\mathbb{E}(Y_\infty X_\infty^2) = \frac{a}{(b+2\theta)2b^2\theta^2}\big(\theta(ab + 2a\theta + \theta) + 2m^2 b(2\theta + b)\big).$$

CHAPTER 7

Parameter estimation for a subcritical affine two-factor model

Co-authors: Leif Doering, Zenghu Li and Gyula Pap

7.1. Introduction

We consider the following two-dimensional affine process (affine two factor model)

$$
\begin{cases} \mathrm{d}Y_t = (a - bY_t)\,\mathrm{d}t + \sqrt{Y_t}\,\mathrm{d}L_t, \\ \mathrm{d}X_t = (m - \theta X_t)\,\mathrm{d}t + \sqrt{Y_t}\,\mathrm{d}B_t, \end{cases} \qquad t \geqslant 0, \tag{7.1.1}
$$

where $a > 0$, $b, m, \theta \in \mathbb{R}$, and $(L_t)_{t\geqslant 0}$ and $(B_t)_{t\geqslant 0}$ are independent standard Wiener processes. Note that the process $(Y_t)_{t\geqslant 0}$ given by the first SDE of (7.1.1) is the so-called Cox–Ingersol–Ross (CIR) process which is a continuous state branching process with branching mechanism $bz + z^2/2$, $z \geqslant 0$, and with immigration mechanism az, $z \geqslant 0$. Chen and Joslin [**45**] applied (7.1.1) for modelling quantitative impact of stochastic recovery on the pricing of defaultable bonds, see their equations (25) and (26).

The process (Y, X) given by (7.1.1) is a special affine diffusion process. The set of affine processes contains a large class of important Markov processes such as continuous state branching processes and Ornstein–Uhlenbeck processes. Further, a lot of models in financial mathematics are affine such as the Heston model [**84**], the model of Barndorff-Nielsen and Shephard [**32**] or the model due to Carr and Wu [**42**]. A precise mathematical formulation and a complete characterization of regular affine processes are due to Duffie et al. [**63**]. These processes are widely used in financial mathematics due to their computational tractability, see Gatheral [**77**].

This article is devoted to estimate the parameters a, b, m and θ from some continuously observed real data set $(Y_t, X_t)_{t\in[0,T]}$, where $T > 0$. To the best knowledge of the authors the parameter estimation problem for multidimensional affine processes has not been tackled so far. Since affine processes are frequently used in financial mathematics, the question of parameter estimation for them is of high importance. In Barczy et al. [**11**] we started the discussion with a simple non-trivial two-dimensional affine diffusion process given by (7.1.1) in the so called critical case: $b \geqslant 0, \theta = 0$ or $b = 0, \theta \geqslant 0$ (for the definition of criticality, see Section 7.2). In the special critical case $b = 0, \theta = 0$ we described the asymptotic behavior of least squares estimator (LSE) of (m, θ) from some discretely observed low frequency real data set $X_0, X_1, \ldots, X_n$ as $n \to \infty$. The description of the asymptotic behavior of the LSE of (m, θ) in the other critical cases $b = 0, \theta > 0$ or $b > 0, \theta = 0$ remained opened. In this paper we deal with the same model (7.1.1) but in the so-called subcritical (ergodic) case: $b > 0, \theta > 0$, and we consider the maximum likelihood estimator (MLE) of (a, b, m, θ) using some continuously observed real data set $(Y_t, X_t)_{t\in[0,T]}$, where $T > 0$, and the LSE of (m, θ) using some continuously observed real data set $(X_t)_{t\in[0,T]}$, where $T > 0$. For studying the asymptotic behaviour of the MLE and LSE in the subcritical (ergodic) case, one first needs to examine the question of existence of a unique stationary distribution and ergodicity for the model given by (7.1.1). In a companion paper Barczy et al. [**12**] we solved this problem, see also Theorem 7.2.5. Further, in a more general setup by replacing the CIR process $(Y_t)_{t\geqslant 0}$ in the first SDE of (7.1.1) by a so-called α-root process (stable CIR process) with $\alpha \in (1,2)$, the existence of a unique stationary distribution for the corresponding model was proved in Barczy et al. [**12**].

In general, parameter estimation for subcritical (also called ergodic) models has a long history, see, e.g., the monographs of Liptser and Shiryaev [**118**, Chapter 17], Kutoyants [**112**], Bishwal [**38**] and the papers of Klimko and Nelson [**108**] and Sørensen [**148**]. In case of the one-dimensional CIR process Y, the parameter estimation of a and b goes back to Overbeck and Rydén [**135**] (conditional LSE), Overbeck [**136**] (MLE), and see also Bishwal [**38**, Example 7.6] and the very recent papers of Ben Alaya and Kebaier [**34**], [**35**] (MLE). In Ben Alaya and Kebaier [**34**], [**35**] one can find a systematic study of the asymptotic behavior of the quadruplet $\big(\log(Y_t), Y_t, \int_0^t Y_s \,\mathrm{d}s, \int_0^t 1/Y_s \,\mathrm{d}s\big)$ as $t \to \infty$. Finally, we note that Li and Ma [**116**] started to investigate the asymptotic behaviour of the (weighted) conditional LSE of the drift parameters for a CIR model driven by a stable noise (they call it a stable CIR model) from some discretely observed low frequency real data set. To give another example besides the one-dimensional CIR process, we mention a model that is somewhat related to (7.1.1) and parameter estimation of the appearing parameters based on continuous time observations has been considered. It is the so-called Ornstein–Uhlenbeck process driven by α-stable Lévy motions, i.e., $\mathrm{d}U_t = (m - \theta U_t)\,\mathrm{d}t + \mathrm{d}Z_t$, $t \geqslant 0$, where $\theta > 0, m \neq 0$, and $(Z_t)_{t\geqslant 0}$ is an α-stable Lévy motion with $\alpha \in (1,2)$. For this model Hu and Long investigated the question of parameter estimation, see [**88**], [**89**] and [**90**].

It would be possible to calculate the discretized version of the estimators presented in this paper using the same procedure as in Ben Alaya and Kebaier [**34**, Section 4] valid for discrete time observations of high frequency. However, it is out of the scope of the present paper.

We give a brief overview of the structure of the paper. Section 7.2 is devoted to a preliminary discussion of the existence and uniqueness of a strong solution of the SDE (7.1.1), we make a classification of the model (see Definition 7.2.4), we also recall our results in Barczy et al. [**12**] on the existence of a unique stationary distribution and ergodicity for the affine process given by SDE (7.1.1), see Theorem 7.2.5. Further, we recall some limit theorems for continuous local martingales that will be used later on for studying the asymptotic behaviour of the MLE of (a, b, m, θ) and the LSE of (m, θ), respectively. In Sections 7.3–7.8 we study the asymptotic behavior of the MLE of (a, b, m, θ) and LSE of (m, θ) proving that the estimators are strongly consistent and asymptotically normal under appropriate conditions on the parameters. We call the attention that for the MLE of (a, b, m, θ) we require a continuous time observation $(Y_t, X_t)_{t\in[0,T]}$ of the process (Y, X), but for the LSE of (m, θ) we need a continuous time observation $(X_t)_{t\in[0,T]}$ only for the process X. We note that in the critical case we obtained a different limit behaviour for the LSE of (m, θ), see Barczy et al. [**11**, Theorem 3.2].

A common generalization of the model (7.1.1) and the well-known Heston model [**84**] is a general affine diffusion two factor model

$$
\begin{cases}
\mathrm{d}Y_t = (a - bY_t)\,\mathrm{d}t + \sigma_1\sqrt{Y_t}\,\mathrm{d}L_t, \\
\mathrm{d}X_t = (m - \kappa Y_t - \theta X_t)\,\mathrm{d}t + \sigma_1\sqrt{Y_t}\,(\varrho L_t + \sqrt{1-\varrho^2}\mathrm{d}B_t),
\end{cases} \qquad t \geqslant 0, \tag{7.1.2}
$$

where $a, \sigma_1, \sigma_2 > 0$, $b, m, \kappa, \theta \in \mathbb{R}$, $\varrho \in (-1, 1)$, and $(L_t)_{t \geqslant 0}$ and $(B_t)_{t \geqslant 0}$ are independent standard Wiener processes. One does not need to estimate the parameters σ_1, σ_2 and ϱ, since these parameters could —in principle, at least— be determined (rather than estimated) using an arbitrarily short continuous time observation of (Y, X), see Remark 2.5 in Barczy and Pap [**30**]. For studying the parameter estimation of a, b, m, κ and θ in the subcritical case, one needs to investigate ergodicity properties of the model (7.1.2). For the submodel (7.1.1), this has been proved in Barczy et al. [**12**], see also Theorem 7.2.5. For the Heston model, ergodicity of the first coordinate process Y is sufficient for statistical purposes, see Barczy and Pap [**30**]; the existence of a unique stationary distribution and the ergodicity for Y has been proved by Cox et al. [**49**, Equation (20)] and Li and Ma [**116**, Theorem 2.6]. After showing appropriate ergodicity properties of the model (7.1.2), one could obtain the asymptotic behavior of the MLE and LSE of $(a, b, m, \kappa, \theta)$ with a similar method used in the present paper.

7.2. Preliminaires

Let $\mathbb{N}$, $\mathbb{Z}_+$, $\mathbb{R}$ and $\mathbb{R}_+$ denote the sets of positive integers, non-negative integers, real numbers and non-negative real numbers, respectively. By $\|x\|$ and $\|A\|$ we denote the Euclidean norm of a vector $x \in \mathbb{R}^m$ and the induced matrix norm $\|A\| = \sup\{\|Ax\| : x \in \mathbb{R}^m, \ \|x\| = 1\}$ of a matrix $A \in \mathbb{R}^{n \times m}$, respectively.

The next proposition is about the existence and uniqueness of a strong solution of the SDE (7.1.1) which follows from the theorem due to Yamada and Watanabe, further it clarifies that the process given by the SDE (7.1.1) belongs to the family of regular affine processes, see Barczy et al. [**12**, Theorem 2.2 with $\alpha = 2$].

Proposition 7.2.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and $(L_t, B_t)_{t \in \mathbb{R}_+}$ be a two-dimensional standard Wiener process. Let (η_0, ζ_0) be a random vector independent of $(L_t, B_t)_{t \geqslant 0}$ satisfying $\mathbb{P}(\eta_0 \geqslant 0) = 1$. Then, for all $a > 0$, and $b, m, \theta \in \mathbb{R}$, there is a (pathwise) unique strong solution $(Y_t, X_t)_{t \geqslant 0}$ of the SDE (7.1.1) such that $\mathbb{P}((Y_0, X_0) = (\eta_0, \zeta_0)) = 1$ and $\mathbb{P}(Y_t \geqslant 0, \ \forall\, t \geqslant 0) = 1$. Further, we have*

(7.2.1)
$$Y_t = \mathrm{e}^{-b(t-s)} \left(Y_s + a \int_s^t \mathrm{e}^{-b(s-u)} \, \mathrm{d}u + \int_s^t \mathrm{e}^{-b(s-u)} \sqrt{Y_u} \, \mathrm{d}L_u \right), \qquad 0 \leqslant s \leqslant t,$$

and

(7.2.2)
$$X_t = \mathrm{e}^{-\theta(t-s)} \left(X_s + m \int_s^t \mathrm{e}^{-\theta(s-u)} \, \mathrm{d}u + \int_s^t \mathrm{e}^{-\theta(s-u)} \sqrt{Y_u} \, \mathrm{d}B_u \right), \qquad 0 \leqslant s \leqslant t.$$

Moreover, $(Y_t, X_t)_{t \geqslant 0}$ is a regular affine process with infinitesimal generator

$$(\mathcal{A}f)(y, x) = (a - by) f_1'(y, x) + (m - \theta x) f_2'(y, x) + \frac{1}{2} y (f_{1,1}''(y, x) + f_{2,2}''(y, x)),$$

where $(y, x) \in \mathbb{R}_+ \times \mathbb{R}$, $f \in \mathcal{C}_c^2(\mathbb{R}_+ \times \mathbb{R}, \mathbb{R})$, f_i', $i = 1, 2$, and $f_{i,j}''$, $i, j \in \{1, 2\}$, denote the first and second order partial derivatives of f with respect to its i-th and i-th and j-th variables, respectively, and $\mathcal{C}_c^2(\mathbb{R}_+ \times \mathbb{R}, \mathbb{R})$ is the set of twice continuously differentiable real-valued functions defined on $\mathbb{R}_+ \times \mathbb{R}$ having compact support.

Remark 7.2.2. Note that in Proposition 7.2.1 the unique strong solution $(Y_t, X_t)_{t \geqslant 0}$ of the SDE (7.1.1) is adapted to the augmented filtration $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ corresponding to $(L_t, B_t)_{t \in \mathbb{R}_+}$ and (η_0, ζ_0), constructed as in Karatzas and Shreve [**100**, Section 5.2]. Note also that $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ satisfies the usual conditions, i.e., the filtration $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ is right-continuous and $\mathcal{F}_0$ contains all the $\mathbb{P}$-null sets in $\mathcal{F}$. Further, $(L_t)_{t\in\mathbb{R}_+}$ and $(B_t)_{t\in\mathbb{R}_+}$ are independent $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$-standard Wiener processes. In Proposition 7.2.1 it is the assumption $a > 0$ which ensures $\mathbb{P}(Y_t \geqslant 0,\ \forall\, t \geqslant 0) = 1$. □

In what follows we will make a classification of the affine processes given by the SDE (7.1.1). First we recall a result about the first moment of $(Y_t, X_t)_{t\in\mathbb{R}_+}$, see Proposition 3.2 in Barczy et al. [**11**].

Proposition 7.2.3. *Let $(Y_t, X_t)_{t\in\mathbb{R}_+}$ be an affine diffusion process given by the SDE (7.1.1) with a random initial value (η_0, ζ_0) independent of $(L_t, B_t)_{t\geqslant 0}$ such that $\mathbb{P}(\eta_0 \geqslant 0) = 1$, $\mathbb{E}(\eta_0) < \infty$ and $\mathbb{E}(|\zeta_0|) < \infty$. Then*

$$\begin{bmatrix}\mathbb{E}(Y_t)\\ \mathbb{E}(X_t)\end{bmatrix} = \begin{bmatrix}\mathrm{e}^{-bt} & 0\\ 0 & \mathrm{e}^{-\theta t}\end{bmatrix}\begin{bmatrix}\mathbb{E}(\eta_0)\\ \mathbb{E}(\zeta_0)\end{bmatrix} + \begin{bmatrix}\int_0^t \mathrm{e}^{-bs}\,\mathrm{d}s & 0\\ 0 & \int_0^t \mathrm{e}^{-\theta s}\,\mathrm{d}s\end{bmatrix}\begin{bmatrix}a\\ m\end{bmatrix}, \qquad t \in \mathbb{R}_+.$$

Proposition 7.2.3 shows that the asymptotic behavior of the first moment of $(Y_t, X_t)_{t\in\mathbb{R}_+}$ as $t \to \infty$ is determined by the spectral radius of the diagonal matrix $\operatorname{diag}(\mathrm{e}^{-bt}, \mathrm{e}^{-\theta t})$, which motivates our classification of the affine processes given by the SDE (7.1.1).

Definition 7.2.4. Let $(Y_t, X_t)_{t\in\mathbb{R}_+}$ be an affine diffusion process given by the SDE (7.1.1) with a random initial value (η_0, ζ_0) independent of $(L_t, B_t)_{t\geqslant 0}$ satisfying $\mathbb{P}(\eta_0 \geqslant 0) = 1$. We call $(Y_t, X_t)_{t\in\mathbb{R}_+}$ subcritical, critical or supercritical if the spectral radius of the matrix

$$\begin{bmatrix}\mathrm{e}^{-bt} & 0\\ 0 & \mathrm{e}^{-\theta t}\end{bmatrix}$$

is less than 1, equal to 1 or greater than 1, respectively.

Note that, since the spectral radius of the matrix given in Definition 7.2.4 is $\max(\mathrm{e}^{-bt}, \mathrm{e}^{-\theta t})$, the affine process given in Definition 7.2.4 is

$$\begin{aligned} &\text{subcritical} && \text{if } b > 0 \text{ and } \theta > 0,\\ &\text{critical} && \text{if } b = 0,\ \theta \geqslant 0 \text{ or } b \geqslant 0,\ \theta = 0,\\ &\text{supercritical} && \text{if } b < 0 \text{ or } \theta < 0.\end{aligned}$$

Further, under the conditions of Proposition 7.2.3, by an easy calculation, if $b > 0$ and $\theta > 0$, then

$$\lim_{t\to\infty}\begin{bmatrix}\mathbb{E}(Y_t)\\ \mathbb{E}(X_t)\end{bmatrix} = \begin{bmatrix}\frac{a}{b}\\ \frac{m}{\theta}\end{bmatrix},$$

if $b = 0$ and $\theta = 0$, then

$$\lim_{t\to\infty}\begin{bmatrix}\frac{1}{t}\,\mathbb{E}(Y_t)\\ \frac{1}{t}\,\mathbb{E}(X_t)\end{bmatrix} = \begin{bmatrix}a\\ m\end{bmatrix},$$

if $b = 0$ and $\theta > 0$, then

$$\lim_{t\to\infty}\begin{bmatrix}\frac{1}{t}\,\mathbb{E}(Y_t)\\ \mathbb{E}(X_t)\end{bmatrix} = \begin{bmatrix}a\\ \frac{m}{\theta}\end{bmatrix},$$

if $b > 0$ and $\theta = 0$, then

$$\lim_{t\to\infty}\begin{bmatrix}\mathbb{E}(Y_t)\\ \frac{1}{t}\,\mathbb{E}(X_t)\end{bmatrix} = \begin{bmatrix}\frac{a}{b}\\ m\end{bmatrix},$$

and if $b < 0$ and $\theta < 0$, then

$$\lim_{t\to\infty}\begin{bmatrix}\mathrm{e}^{bt}\,\mathbb{E}(Y_t)\\ \mathrm{e}^{\theta t}\,\mathbb{E}(X_t)\end{bmatrix} = \begin{bmatrix}\mathbb{E}(\eta_0) - \frac{a}{b}\\ \mathbb{E}(\zeta_0) - \frac{m}{\theta}\end{bmatrix}.$$

Remark also that Definition 7.2.4 of criticality is in accordance with the corresponding definition for one-dimensional continuous state branching processes, see, e.g., Li [**115**, page 58].

In the sequel $\stackrel{\mathbb{P}}{\longrightarrow}$ and $\stackrel{\mathcal{L}}{\longrightarrow}$ will denote convergence in probability and in distribution, respectively.

The following result states the existence of a unique stationary distribution and the ergodicity for the affine process given by the SDE (7.1.1), see Theorems 3.1 with $\alpha = 2$ and Theorem 4.2 in Barczy et al. [**12**].

THEOREM 7.2.5. *Let us consider the two-dimensional affine model* (7.1.1) *with* $a > 0$, $b > 0$, $m \in \mathbb{R}$, $\theta > 0$, *and with a random initial value* (η_0, ζ_0) *independent of* $(L_t, B_t)_{t\geqslant 0}$ *satisfying* $\mathbb{P}(\eta_0 \geqslant 0) = 1$. *Then*

(i) $(Y_t, X_t) \stackrel{\mathcal{L}}{\longrightarrow} (Y_\infty, X_\infty)$ *as* $t \to \infty$, *and the distribution of* (Y_∞, X_∞) *is given by*

$$\begin{aligned}
&\mathbb{E}\left(\mathrm{e}^{-\lambda_1 Y_\infty + i\lambda_2 X_\infty}\right) \\
&\qquad = \exp\left\{-a\int_0^\infty v_s(\lambda_1,\lambda_2)\,\mathrm{d}s + i\frac{m}{\theta}\lambda_2\right\}, \qquad (\lambda_1,\lambda_2) \in \mathbb{R}_+ \times \mathbb{R},
\end{aligned} \tag{7.2.3}$$

where $v_t(\lambda_1, \lambda_2)$, $t \geqslant 0$, *is the unique non-negative solution of the (deterministic) differential equation*

$$\begin{cases} \frac{\partial v_t}{\partial t}(\lambda_1,\lambda_2) = -bv_t(\lambda_1,\lambda_2) - \frac{1}{2}(v_t(\lambda_1,\lambda_2))^2 + \frac{1}{2}\mathrm{e}^{-2\theta t}\lambda_2^2, \qquad t \geqslant 0, \\ v_0(\lambda_1,\lambda_2) = \lambda_1. \end{cases} \tag{7.2.4}$$

(ii) *supposing that the random initial value* (η_0, ζ_0) *has the same distribution as* (Y_∞, X_∞) *given in part (i), we have* $(Y_t, X_t)_{t\geqslant 0}$ *is strictly stationary.*

(iii) *for all Borel measurable functions* $f : \mathbb{R}^2 \to \mathbb{R}$ *with* $\mathbb{E}\left(|f(Y_\infty, X_\infty)|\right) < \infty$, *we have*

$$\mathbb{P}\left(\lim_{T\to\infty} \frac{1}{T}\int_0^T f(Y_s, X_s)\,\mathrm{d}s = \mathbb{E}(f(Y_\infty, X_\infty))\right) = 1, \tag{7.2.5}$$

where the distribution of (Y_∞, X_∞) *is given by* (7.2.3) *and* (7.2.4).

Moreover, the random variable (Y_∞, X_∞) *is absolutely continuous, the Laplace transform of* Y_∞ *takes the form*

$$\mathbb{E}(\mathrm{e}^{-\lambda_1 Y_\infty}) = \left(1 + \frac{\lambda_1}{2b}\right)^{-2a}, \qquad \lambda_1 \in \mathbb{R}_+, \tag{7.2.6}$$

i.e., Y_∞ *has Gamma distribution with parameters* $2a$ *and* $2b$, *all the (mixed) moments of* (Y_∞, X_∞) *of any order are finite, i.e.,* $\mathbb{E}(Y_\infty^n |X_\infty|^p) < \infty$ *for all* $n, p \in \mathbb{Z}_+$, *and especially,*

$$\begin{aligned}
&\mathbb{E}(Y_\infty) = \frac{a}{b}, \qquad \mathbb{E}(X_\infty) = \frac{m}{\theta}, \\
&\mathbb{E}(Y_\infty^2) = \frac{a(2a+1)}{2b^2}, \qquad \mathbb{E}(Y_\infty X_\infty) = \frac{ma}{\theta b}, \qquad \mathbb{E}(X_\infty^2) = \frac{a\theta + 2bm^2}{2b\theta^2}, \\
&\mathbb{E}(Y_\infty X_\infty^2) = \frac{a}{(b+2\theta)2b^2\theta^2}\big(\theta(ab + 2a\theta + \theta) + 2m^2 b(2\theta + b)\big).
\end{aligned}$$

In all what follows we will suppose that we have continuous time observations for the process (Y, X), i.e., $(Y_t, X_t)_{t\in[0,T]}$ can be observed for some $T > 0$, and our aim is to deal with parameter estimation of (a, b, m, θ). We also deal with parameter estimation of θ provided that the parameter $m \in \mathbb{R}$ is supposed to be known.

Next we recall some limit theorems for continuous local martingales. We will use these limit theorems in the sequel for studying the asymptotic behaviour of different kinds of estimators for (a, b, m, θ). First we recall a strong law of large numbers for continuous local martingales, see, e.g., Liptser and Shiryaev [**118**, Lemma 17.4].

THEOREM 7.2.6. *Let* $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P}\big)$ *be a filtered probability space satisfying the usual conditions. Let* $(M_t)_{t\geqslant 0}$ *be a square-integrable continuous local martingale with respect to the filtration* $(\mathcal{F}_t)_{t\geqslant 0}$ *started from* 0. *Let* $(\xi_t)_{t\geqslant 0}$ *be a progressively measurable process such that*

$$\mathbb{P}\left(\int_0^t (\xi_u)^2 \,\mathrm{d}\langle M\rangle_u < \infty\right) = 1, \quad t \geqslant 0,$$

and

$$\mathbb{P}\left(\lim_{t\to\infty} \int_0^t (\xi_u)^2 \,\mathrm{d}\langle M\rangle_u = \infty\right) = 1, \tag{7.2.7}$$

where $(\langle M\rangle_t)_{t\geqslant 0}$ *denotes the quadratic variation process of* M. *Then*

$$\mathbb{P}\left(\lim_{t\to\infty} \frac{\int_0^t \xi_u \,\mathrm{d}M_u}{\int_0^t (\xi_u)^2 \,\mathrm{d}\langle M\rangle_u} = 0\right) = 1. \tag{7.2.8}$$

In case of $M_t = B_t$, $t \geqslant 0$, *where* $(B_t)_{t\geqslant 0}$ *is a standard Wiener process, the progressive measurability of* $(\xi_t)_{t\geqslant 0}$ *can be relaxed to measurability and adaptedness to the filtration* $(\mathcal{F}_t)_{t\geqslant 0}$.

The next theorem is about the asymptotic behaviour of continuous multivariate local martingales.

THEOREM 7.2.7. *(van Zanten [**158**, Theorem 4.1]) Let* $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P}\big)$ *be a filtered probability space satisfying the usual conditions. Let* $(M_t)_{t\geqslant 0}$ *be a* d*-dimensional square-integrable continuous local martingale with respect to the filtration* $(\mathcal{F}_t)_{t\geqslant 0}$ *started from* 0. *Suppose that there exists a function* $Q : [0,\infty) \to \mathbb{R}^{d\times d}$ *such that* $Q(t)$ *is a non-random, invertible matrix for all* $t \geqslant 0$, $\lim_{t\to\infty} \|Q(t)\| = 0$ *and*

$$Q(t)\langle M\rangle_t Q(t)^\top \xrightarrow{\mathbb{P}} \eta\eta^\top \quad \textit{as} \quad t\to\infty,$$

where η *is a* $d \times d$ *random matrix defined on* $\big(\Omega, \mathcal{F}, \mathbb{P}\big)$. *Then, for each* $\mathbb{R}^k$*-valued random variable* V *defined on* $\big(\Omega, \mathcal{F}, \mathbb{P}\big)$, *it holds that*

$$(Q(t)M_t, V) \xrightarrow{\mathcal{L}} (\eta Z, V) \qquad \textit{as} \qquad t\to\infty,$$

where Z *is a* d*-dimensional standard normally distributed random variable independent of* (η, V).

We note that Theorem 7.2.7 remains true if the function Q, instead of the interval $[0,\infty)$, is defined only on an interval $[t_0,\infty)$ with some $t_0 > 0$.

7.3. Existence and uniqueness of maximum likelihood estimator

We will denote by $\mathbb{P}_{(a,b,m,\theta)}$ the probability measure on the measurable space $(\mathcal{C}(\mathbb{R}_+, \mathbb{R}_+\times\mathbb{R}), \mathcal{B}(\mathcal{C}(\mathbb{R}_+, \mathbb{R}_+\times\mathbb{R})))$ induced by the process $(Y_t, X_t)_{t\geqslant 0}$ corresponding to the parameters (a, b, m, θ) and initial value (Y_0, X_0). Here $\mathcal{C}(\mathbb{R}_+, \mathbb{R}_+\times\mathbb{R})$ denotes the set of continuous $\mathbb{R}_+\times\mathbb{R}$-valued functions defined on $\mathbb{R}_+$, $\mathcal{B}(\mathcal{C}(\mathbb{R}_+, \mathbb{R}_+\times\mathbb{R}))$ is the Borel σ-algebra on it, and we suppose that the space $(\mathcal{C}(\mathbb{R}_+, \mathbb{R}_+\times\mathbb{R}), \mathcal{B}(\mathcal{C}(\mathbb{R}_+, \mathbb{R}_+\times\mathbb{R})))$ is endowed with the natural filtration $(\mathcal{A}_t)_{t\geqslant 0}$, given by $\mathcal{A}_t := \varphi_t^{-1}(\mathcal{B}(\mathcal{C}(\mathbb{R}_+, \mathbb{R}_+\times\mathbb{R})))$, where $\varphi_t : \mathcal{C}(\mathbb{R}_+, \mathbb{R}_+\times\mathbb{R}) \to$

$\mathcal{C}(\mathbb{R}_+, \mathbb{R}_+ \times \mathbb{R})$ is the mapping $\varphi_t(f)(s) := f(t \wedge s)$, $s \geqslant 0$. For all $T > 0$, let $\mathbb{P}_{(a,b,m,\theta),T} := \mathbb{P}_{(a,b,m,\theta)} |_{\mathcal{A}_T}$ be the restriction of $\mathbb{P}_{(a,b,m,\theta)}$ to $\mathcal{A}_T$.

LEMMA 7.3.1. *Let* $a \geqslant 1/2$, $b, m, \theta \in \mathbb{R}$, $T > 0$, *and suppose that* $\mathbb{P}(Y_0 > 0) = 1$. *Let* $\mathbb{P}_{(a,b,m,\theta)}$ *and* $\mathbb{P}_{(1,0,0,0)}$ *denote the probability measures induced by the unique strong solutions of the SDE* (7.1.1) *corresponding to the parameters* (a, b, m, θ) *and* $(1, 0, 0, 0)$ *with the same initial value* (Y_0, X_0), *respectively. Then* $\mathbb{P}_{(a,b,m,\theta),T}$ *and* $\mathbb{P}_{(1,0,0,0),T}$ *are absolutely continuous with respect to each other, and the Radon-Nykodim derivative of* $\mathbb{P}_{(a,b,m,\theta),T}$ *with respect to* $\mathbb{P}_{(1,0,0,0),T}$ *(so called likelihood ratio) takes the form*

$$
\begin{aligned}
L_T^{(a,b,m,\theta),(1,0,0,0)}((Y_s, X_s)_{s\in[0,T]}) = \exp\Bigg\{ \int_0^T \left(\frac{a - bY_s - 1}{Y_s} \,\mathrm{d}Y_s + \frac{m - \theta X_s}{Y_s} \,\mathrm{d}X_s \right) \\
- \frac{1}{2} \int_0^T \frac{(a - bY_s - 1)(a - bY_s + 1) + (m - \theta X_s)^2}{Y_s} \,\mathrm{d}s \Bigg\},
\end{aligned}
$$

where $(Y_t, X_t)_{t\geqslant 0}$ *denotes the unique strong solution of the SDE* (7.1.1) *corresponding to the parameters* (a, b, m, θ) *and the initial value* (Y_0, X_0).

By Lemma 7.3.1, under its conditions the log-likelihood function takes the form

$$
\begin{aligned}
\log L_T^{(a,b,m,\theta),(1,0,0,0)}&((Y_s, X_s)_{s\in[0,T]}) \\
&= (a - 1) \int_0^T \frac{1}{Y_s} \,\mathrm{d}Y_s - b(Y_T - Y_0) + m \int_0^T \frac{1}{Y_s} \,\mathrm{d}X_s \\
&\quad - \theta \int_0^T \frac{X_s}{Y_s} \,\mathrm{d}X_s - \frac{a^2 - 1}{2} \int_0^T \frac{1}{Y_s} \,\mathrm{d}s + abT - \frac{b^2}{2} \int_0^T Y_s \,\mathrm{d}s \\
&\quad - \frac{m^2}{2} \int_0^T \frac{1}{Y_s} \,\mathrm{d}s + m\theta \int_0^T \frac{X_s}{Y_s} \,\mathrm{d}s - \frac{\theta^2}{2} \int_0^T \frac{X_s^2}{Y_s} \,\mathrm{d}s \\
&=: f_T(a, b, m, \theta), \qquad T > 0.
\end{aligned}
$$

We remark that for all $T > 0$ and all initial values (Y_0, X_0), the probability measures $\mathbb{P}_{(a,b,m,\theta),T}$, $a \geqslant 1/2, b, m, \theta \in \mathbb{R}$, are absolutely continuous with respect to each other, and hence it does not matter which measure is taken as a reference measure for defining the MLE (we have chosen $\mathbb{P}_{(1,0,0,0),T}$). For more details, see, e.g., Liptser and Shiryaev [**117**, page 35]. Then the equation $\frac{\partial f_T}{\partial \theta}(a, b, m, \theta) = 0$ takes the form

$$
- \int_0^T \frac{X_s}{Y_s} \,\mathrm{d}X_s + m \int_0^T \frac{X_s}{Y_s} \,\mathrm{d}s - \theta \int_0^T \frac{X_s^2}{Y_s} \,\mathrm{d}s = 0.
$$

Moreover, the system of equations

$$
\begin{aligned}
&\frac{\partial f_T}{\partial a}(a, b, m, \theta) = 0, \qquad \frac{\partial f_T}{\partial b}(a, b, m, \theta) = 0, \\
&\frac{\partial f_T}{\partial m}(a, b, m, \theta) = 0, \qquad \frac{\partial f_T}{\partial \theta}(a, b, m, \theta) = 0,
\end{aligned}
$$

takes the form

$$
\begin{bmatrix}
\int_0^T \frac{1}{Y_s} \,\mathrm{d}s & -T & 0 & 0 \\
-T & \int_0^T Y_s \,\mathrm{d}s & 0 & 0 \\
0 & 0 & \int_0^T \frac{1}{Y_s} \,\mathrm{d}s & -\int_0^T \frac{X_s}{Y_s} \,\mathrm{d}s \\
0 & 0 & -\int_0^T \frac{X_s}{Y_s} \,\mathrm{d}s & \int_0^T \frac{X_s^2}{Y_s} \,\mathrm{d}s
\end{bmatrix}
\begin{bmatrix} a \\ b \\ m \\ \theta \end{bmatrix}
=
\begin{bmatrix}
\int_0^T \frac{1}{Y_s} \,\mathrm{d}Y_s \\
-(Y_T - Y_0) \\
\int_0^T \frac{1}{Y_s} \,\mathrm{d}X_s \\
-\int_0^T \frac{X_s}{Y_s} \,\mathrm{d}X_s
\end{bmatrix}.
$$

First, we suppose that $a \geqslant 1/2$, and $b \in \mathbb{R}$ and $m \in \mathbb{R}$ are known. By maximizing $\log L_T^{(a,b,m,\theta),(1,0,0,0)}$ in $\theta \in \mathbb{R}$, we get the MLE of θ based on the observations $(Y_t, X_t)_{t\in[0,T]}$,

$$\widetilde{\theta}_T^{\mathrm{MLE}} := \frac{-\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}X_s + m\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s}{\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s}, \qquad T>0, \tag{7.3.1}$$

provided that $\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s > 0$. Indeed,

$$\frac{\partial^2 f_T}{\partial\theta^2}(\theta, m) = -\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s < 0.$$

Using the SDE (7.1.1), one can also get

$$\widetilde{\theta}_T^{\mathrm{MLE}} - \theta = -\frac{\int_0^T \frac{X_s}{\sqrt{Y_s}}\,\mathrm{d}B_s}{\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s}, \qquad T>0, \tag{7.3.2}$$

provided that $\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s > 0$. Note that the estimator $\widetilde{\theta}_T^{\mathrm{MLE}}$ does not depend on the parameters $a \geqslant 1/2$ and $b\in\mathbb{R}$. In fact, if we maximize $\log L_T^{(a,b,m,\theta),(1,0,0,0)}$ in $(a,b,\theta)\in\mathbb{R}^3$, then we obtain the MLE of (a,b,θ) supposing that $m\in\mathbb{R}$ is known, and one can observe that the MLE of θ by this procedure coincides with $\widetilde{\theta}_T^{\mathrm{MLE}}$.

By maximizing $\log L_T^{(a,b,m,\theta),(1,0,0,0)}$ in $(a,b,m,\theta)\in\mathbb{R}^4$, the MLE of (a,b,m,θ) based on the observations $(Y_t, X_t)_{t\in[0,T]}$ takes the form

$$\widehat{a}_T^{\mathrm{MLE}} := \frac{\int_0^T Y_s\,\mathrm{d}s\int_0^T \frac{1}{Y_s}\,\mathrm{d}Y_s - T(Y_T - Y_0)}{\int_0^T Y_s\,\mathrm{d}s\int_0^T \frac{1}{Y_s}\,\mathrm{d}s - T^2}, \qquad T>0, \tag{7.3.3}$$

$$\widehat{b}_T^{\mathrm{MLE}} := \frac{T\int_0^T \frac{1}{Y_s}\,\mathrm{d}Y_s - (Y_T - Y_0)\int_0^T \frac{1}{Y_s}\,\mathrm{d}s}{\int_0^T Y_s\,\mathrm{d}s\int_0^T \frac{1}{Y_s}\,\mathrm{d}s - T^2}, \qquad T>0, \tag{7.3.4}$$

$$\widehat{m}_T^{\mathrm{MLE}} := \frac{\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s\int_0^T \frac{1}{Y_s}\,\mathrm{d}X_s - \int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}X_s}{\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s\int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2}, \qquad T>0, \tag{7.3.5}$$

$$\widehat{\theta}_T^{\mathrm{MLE}} := \frac{\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\int_0^T \frac{1}{Y_s}\,\mathrm{d}X_s - \int_0^T \frac{1}{Y_s}\,\mathrm{d}s\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}X_s}{\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s\int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2}, \qquad T>0, \tag{7.3.6}$$

provided that $\int_0^T Y_s\,\mathrm{d}s\int_0^T \frac{1}{Y_s}\,\mathrm{d}s - T^2 > 0$ and $\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s\int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2 > 0$. Indeed,

$$\begin{bmatrix} \frac{\partial^2 f_T}{\partial a^2}(a,b,m,\theta) & \frac{\partial^2 f_T}{\partial b\partial a}(a,b,m,\theta) \\ \frac{\partial^2 f_T}{\partial a\partial b}(a,b,m,\theta) & \frac{\partial^2 f_T}{\partial b^2}(a,b,m,\theta) \end{bmatrix} = \begin{bmatrix} -\int_0^T \frac{1}{Y_s}\,\mathrm{d}s & T \\ T & -\int_0^T Y_s\,\mathrm{d}s \end{bmatrix},$$

$$\begin{bmatrix} \frac{\partial^2 f_T}{\partial m^2}(a,b,m,\theta) & \frac{\partial^2 f_T}{\partial \theta\partial m}(a,b,m,\theta) \\ \frac{\partial^2 f_T}{\partial m\partial \theta}(a,b,m,\theta) & \frac{\partial^2 f_T}{\partial \theta^2}(a,b,m,\theta) \end{bmatrix} = \begin{bmatrix} -\int_0^T \frac{1}{Y_s}\,\mathrm{d}s & \int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s \\ \int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s & -\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \end{bmatrix},$$

and the positivity of $\int_0^T Y_s \,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - T^2$ and $\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2$ yield $\int_0^T \frac{1}{Y_s}\,\mathrm{d}s > 0$, respectively. Using the SDE (7.1.1) one can check that

$$\begin{bmatrix} \widehat{a}_T^{\mathrm{MLE}} - a \\ \widehat{b}_T^{\mathrm{MLE}} - b \end{bmatrix} = \begin{bmatrix} \int_0^T \frac{1}{Y_s}\,\mathrm{d}s & -T \\ -T & \int_0^T Y_s\,\mathrm{d}s \end{bmatrix}^{-1} \begin{bmatrix} \int_0^T \frac{1}{\sqrt{Y_s}}\,\mathrm{d}L_s \\ -\int_0^T \sqrt{Y_s}\,\mathrm{d}L_s \end{bmatrix},$$

$$\begin{bmatrix} \widehat{m}_T^{\mathrm{MLE}} - m \\ \widehat{\theta}_T^{\mathrm{MLE}} - \theta \end{bmatrix} = \begin{bmatrix} \int_0^T \frac{1}{Y_s}\,\mathrm{d}s & -\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s \\ -\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s & \int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \end{bmatrix}^{-1} \begin{bmatrix} \int_0^T \frac{1}{\sqrt{Y_s}}\,\mathrm{d}B_s \\ -\int_0^T \frac{X_s}{\sqrt{Y_s}}\,\mathrm{d}B_s \end{bmatrix},$$

and hence

$$\widehat{a}_T^{\mathrm{MLE}} - a = \frac{\int_0^T Y_s\,\mathrm{d}s \int_0^T \frac{1}{\sqrt{Y_s}}\,\mathrm{d}L_s - T\int_0^T \sqrt{Y_s}\,\mathrm{d}L_s}{\int_0^T Y_s\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - T^2}, \qquad T > 0, \tag{7.3.7}$$

$$\widehat{b}_T^{\mathrm{MLE}} - b = \frac{T\int_0^T \frac{1}{\sqrt{Y_s}}\,\mathrm{d}L_s - \int_0^T \frac{1}{Y_s}\,\mathrm{d}s \int_0^T \sqrt{Y_s}\,\mathrm{d}L_s}{\int_0^T Y_s\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - T^2}, \qquad T > 0, \tag{7.3.8}$$

$$\widehat{m}_T^{\mathrm{MLE}} - m = \frac{\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{\sqrt{Y_s}}\,\mathrm{d}B_s - \int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s \int_0^T \frac{X_s}{\sqrt{Y_s}}\,\mathrm{d}B_s}{\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2}, \qquad T > 0, \tag{7.3.9}$$

and

$$\widehat{\theta}_T^{\mathrm{MLE}} - \theta = \frac{\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{\sqrt{Y_s}}\,\mathrm{d}B_s - \int_0^T \frac{1}{Y_s}\,\mathrm{d}s \int_0^T \frac{X_s}{\sqrt{Y_s}}\,\mathrm{d}B_s}{\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2}, \qquad T > 0, \tag{7.3.10}$$

provided that $\int_0^T Y_s\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - T^2 > 0$ and $\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2 > 0$.

Remark 7.3.2. For the stochastic integrals $\int_0^T \frac{1}{Y_s}\,\mathrm{d}Y_s$, $\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}X_s$ and $\int_0^T \frac{1}{Y_s}\,\mathrm{d}X_s$ in (7.3.3), (7.3.4), (7.3.5) and (7.3.6), we have

$$\begin{aligned} &\sum_{i=1}^{\lfloor nT\rfloor} \frac{1}{Y_{\frac{i-1}{n}}}\big(Y_{\frac{i}{n}} - Y_{\frac{i-1}{n}}\big) \xrightarrow{\mathbb{P}} \int_0^T \frac{1}{Y_s}\,\mathrm{d}Y_s \qquad \text{as } n\to\infty,\\ &\sum_{i=1}^{\lfloor nT\rfloor} \frac{X_{\frac{i-1}{n}}}{Y_{\frac{i-1}{n}}}\big(X_{\frac{i}{n}} - X_{\frac{i-1}{n}}\big) \xrightarrow{\mathbb{P}} \int_0^T \frac{X_s}{Y_s}\,\mathrm{d}X_s \qquad \text{as } n\to\infty,\\ &\sum_{i=1}^{\lfloor nT\rfloor} \frac{1}{Y_{\frac{i-1}{n}}}\big(X_{\frac{i}{n}} - X_{\frac{i-1}{n}}\big) \xrightarrow{\mathbb{P}} \int_0^T \frac{1}{Y_s}\,\mathrm{d}X_s \qquad \text{as } n\to\infty, \end{aligned} \tag{7.3.11}$$

following from Proposition I.4.44 in Jacod and Shiryaev [**95**] with the Riemann sequence of deterministic subdivisions $\left(\frac{i}{n}\wedge T\right)_{i\in\mathbb{N}}$, $n\in\mathbb{N}$. Thus, there exist measurable functions $\Phi, \Psi, \Xi : C([0,T], \mathbb{R}_+\times\mathbb{R}) \to \mathbb{R}$ such that

$$\begin{aligned} \int_0^T \frac{1}{Y_s}\,\mathrm{d}Y_s &= \Phi((Y_s, X_s)_{s\in[0,T]}),\\ \int_0^T \frac{X_s}{Y_s}\,\mathrm{d}X_s &= \Psi((Y_s, X_s)_{s\in[0,T]}),\\ \int_0^T \frac{1}{Y_s}\,\mathrm{d}X_s &= \Xi((Y_s, X_s)_{s\in[0,T]}), \end{aligned}$$

since the convergences in (7.3.11) hold almost surely along suitable subsequences, the members of the sequences in (7.3.11) are measurable functions of $(Y_s, X_s)_{s\in[0,T]}$, and one can use Theorems 4.2.2 and 4.2.8 in Dudley [**62**]. Hence the right hand sides of (7.3.3), (7.3.4), (7.3.5) and (7.3.6) are measurable functions of $(Y_s, X_s)_{s\in[0,T]}$, i.e., they are statistics. □

The next lemma is about the existence of $\widetilde{\theta}_T^{\mathrm{MLE}}$ (supposing that $a \geqslant \frac{1}{2}$, $b \in \mathbb{R}$ and $m \in \mathbb{R}$ are known).

LEMMA 7.3.3. *If* $a \geqslant \frac{1}{2}$, $b, m, \theta \in \mathbb{R}$, *and* $\mathbb{P}(Y_0 > 0) = 1$, *then*

$$\mathbb{P}\left(\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \in (0,\infty)\right) = 1 \quad \text{for all } T > 0, \tag{7.3.12}$$

and hence there exists a unique MLE $\widetilde{\theta}_T^{\mathrm{MLE}}$ *which has the form given in* (7.3.1).

The next lemma is about the existence of $(\widehat{a}_T^{\mathrm{MLE}}, \widehat{b}_T^{\mathrm{MLE}}, \widehat{m}_T^{\mathrm{MLE}}, \widehat{\theta}_T^{\mathrm{MLE}})$.

LEMMA 7.3.4. *If* $a \geqslant \frac{1}{2}$, $b, m, \theta \in \mathbb{R}$, *and* $\mathbb{P}(Y_0 > 0) = 1$, *then*

$$\mathbb{P}\left(\int_0^T Y_s\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - T^2 \in (0,\infty)\right) = 1 \qquad \text{for all } T > 0, \tag{7.3.13}$$

$$\mathbb{P}\left(\int_0^T \frac{X_s^2}{Y_s}\,\mathrm{d}s \int_0^T \frac{1}{Y_s}\,\mathrm{d}s - \left(\int_0^T \frac{X_s}{Y_s}\,\mathrm{d}s\right)^2 \in (0,\infty)\right) = 1 \qquad \text{for all } T > 0, \tag{7.3.14}$$

and hence there exists a unique MLE $(\widehat{a}_T^{\mathrm{MLE}}, \widehat{b}_T^{\mathrm{MLE}}, \widehat{m}_T^{\mathrm{MLE}}, \widehat{\theta}_T^{\mathrm{MLE}})$ *which has the form given in* (7.3.3), (7.3.4) (7.3.5) *and* (7.3.6).

7.4. Existence and uniqueness of least squares estimator

Studying LSE for the model (7.1.1), the parameters $a > 0$ and $b \in \mathbb{R}$ will be not supposed to be known. However, we will not consider the LSEs of a and b, we will focus only on the LSEs of m and θ, since we would like use a continuous time observation only for the process X, and not for the process (Y, X), studying LSEs.

First we give a motivation for the LSE based on continuous time observations using the form of the LSE based on discrete time low frequency observations.

Let us suppose that $m \in \mathbb{R}$ is known ($a > 0$ and $b \in \mathbb{R}$ are not supposed to be known). The LSE of θ based on the discrete time observations X_i, $i = 0, 1, \ldots, n$, can be obtained by solving the following extremum problem

$$\widetilde{\theta}_n^{\mathrm{LSE,D}} := \operatorname*{arg\,min}_{\theta\in\mathbb{R}} \sum_{i=1}^n (X_i - X_{i-1} - (m - \theta X_{i-1}))^2.$$

Here in the notation $\widetilde{\theta}_n^{\mathrm{LSE,D}}$ the letter D refers to discrete time observations, and we note that X_0 denotes an observation for the second coordinate of the initial value of the process (Y, X). This definition of LSE of θ can be considered as the corresponding one given in Hu and Long [**89**, formula (1.2)] for generalized Ornstein-Uhlenbeck processes driven by α-stable motions, see also Hu and Long [**90**, formula (3.1)]. For a motivation of the LSE of θ based on the discrete observations X_i, $i \in \{0, 1, \ldots, n\}$, see Remark 3.4 in Barczy et al. [**11**]. Further, by Barczy et al. [**11**, formula (3.5)],

$$\widetilde{\theta}_n^{\mathrm{LSE,D}} = -\frac{\sum_{i=1}^n (X_i - X_{i-1})X_{i-1} - m\left(\sum_{i=1}^n X_{i-1}\right)}{\sum_{i=1}^n X_{i-1}^2} \tag{7.4.1}$$

provided that $\sum_{i=1}^n X_{i-1}^2 > 0$. Motivated by (7.4.1), the LSE of θ based on the continuous time observations $(X_t)_{t\in[0,T]}$ is defined by

$$\widetilde{\theta}_T^{\mathrm{LSE}} := -\frac{\int_0^T X_s \,\mathrm{d}X_s - m\int_0^T X_s \,\mathrm{d}s}{\int_0^T X_s^2 \,\mathrm{d}s}, \tag{7.4.2}$$

provided that $\int_0^T X_s^2 \,\mathrm{d}s > 0$, and using the SDE (7.1.1) we have

$$\widetilde{\theta}_T^{\mathrm{LSE}} - \theta = -\frac{\int_0^T X_s\sqrt{Y_s} \,\mathrm{d}B_s}{\int_0^T X_s^2 \,\mathrm{d}s}, \tag{7.4.3}$$

provided that $\int_0^T X_s^2 \,\mathrm{d}s > 0$.

Let us suppose that the parameters $a > 0$ and $b, m \in \mathbb{R}$ are not known. The LSE of (m,θ) based on the discrete time observations X_i, $i = 0, 1, \ldots, n$, can be obtained by solving the following extremum problem

$$(\widehat{m}_n^{\mathrm{LSE,D}}, \widehat{\theta}_n^{\mathrm{LSE,D}}) := \operatorname*{arg\,min}_{(\theta,m)\in\mathbb{R}^2} \sum_{i=1}^n (X_i - X_{i-1} - (m - \theta X_{i-1}))^2.$$

By Barczy et al. [**11**, formulas (3.27) and (3.28)],

$$\widehat{m}_n^{\mathrm{LSE,D}} = \frac{\sum_{i=1}^n X_{i-1}^2 \sum_{i=1}^n (X_i - X_{i-1}) - \sum_{i=1}^n X_{i-1}\sum_{i=1}^n (X_i - X_{i-1})X_{i-1}}{n\sum_{i=1}^n X_{i-1}^2 - \left(\sum_{i=1}^n X_{i-1}\right)^2} \tag{7.4.4}$$

and

$$\widehat{\theta}_n^{\mathrm{LSE,D}} = \frac{\sum_{i=1}^n X_{i-1} \sum_{i=1}^n (X_i - X_{i-1}) - n\sum_{i=1}^n (X_i - X_{i-1})X_{i-1}}{n\sum_{i=1}^n X_{i-1}^2 - \left(\sum_{i=1}^n X_{i-1}\right)^2}, \tag{7.4.5}$$

provided that $n\sum_{i=1}^n X_{i-1}^2 - \left(\sum_{i=1}^n X_{i-1}\right)^2 > 0$. Motivated by (7.4.4) and (7.4.5), the LSE of (m,θ) based on the continuous time observations $(X_t)_{t\in[0,T]}$ is defined by

$$\widehat{m}_T^{\mathrm{LSE}} := \frac{(X_T - X_0)\int_0^T X_s^2 \,\mathrm{d}s - \left(\int_0^T X_s \,\mathrm{d}s\right)\left(\int_0^T X_s \,\mathrm{d}X_s\right)}{T\int_0^T X_s^2 \,\mathrm{d}s - \left(\int_0^T X_s \,\mathrm{d}s\right)^2}, \tag{7.4.6}$$

$$\widehat{\theta}_T^{\mathrm{LSE}} := \frac{(X_T - X_0)\int_0^T X_s \,\mathrm{d}s - T\int_0^T X_s \,\mathrm{d}X_s}{T\int_0^T X_s^2 \,\mathrm{d}s - \left(\int_0^T X_s \,\mathrm{d}s\right)^2}, \tag{7.4.7}$$

provided that $T\int_0^T X_s^2 \,\mathrm{d}s - \left(\int_0^T X_s \,\mathrm{d}s\right)^2 > 0$. Note that, by Cauchy-Schwarz's inequality, $T\int_0^T X_s^2 \,\mathrm{d}s - \left(\int_0^T X_s \,\mathrm{d}s\right)^2 \geqslant 0$, and $T\int_0^T X_s^2 \,\mathrm{d}s - \left(\int_0^T X_s \,\mathrm{d}s\right)^2 > 0$ yields that $\int_0^T X_s^2 \,\mathrm{d}s > 0$. Then

$$\begin{bmatrix} \widehat{m}_T^{\mathrm{LSE}} \\ \widehat{\theta}_T^{\mathrm{LSE}} \end{bmatrix} = \begin{bmatrix} T & -\int_0^T X_s \,\mathrm{d}s \\ -\int_0^T X_s \,\mathrm{d}s & \int_0^T X_s^2 \,\mathrm{d}s \end{bmatrix}^{-1} \begin{bmatrix} X_T - X_0 \\ -\int_0^T X_s \,\mathrm{d}X_s \end{bmatrix},$$

and using the SDE (7.1.1) one can check that

$$\begin{bmatrix} \widehat{m}_T^{\mathrm{LSE}} - m \\ \widehat{\theta}_T^{\mathrm{LSE}} - \theta \end{bmatrix} = \begin{bmatrix} T & -\int_0^T X_s \,\mathrm{d}s \\ -\int_0^T X_s \,\mathrm{d}s & \int_0^T X_s^2 \,\mathrm{d}s \end{bmatrix}^{-1} \begin{bmatrix} \int_0^T \sqrt{Y_s} \,\mathrm{d}B_s \\ -\int_0^T X_s\sqrt{Y_s} \,\mathrm{d}B_s \end{bmatrix},$$

and hence

$$\widehat{m}_T^{\mathrm{LSE}} - m = \frac{-\int_0^T X_s \,\mathrm{d}s \int_0^T X_s \sqrt{Y_s}\,\mathrm{d}B_s + \int_0^T X_s^2 \,\mathrm{d}s \int_0^T \sqrt{Y_s}\,\mathrm{d}B_s}{T\int_0^T X_s^2\,\mathrm{d}s - \left(\int_0^T X_s\,\mathrm{d}s\right)^2}, \tag{7.4.8}$$

and

$$\widehat{\theta}_T^{\mathrm{LSE}} - \theta = \frac{-T\int_0^T X_s \sqrt{Y_s}\,\mathrm{d}B_s + \int_0^T X_s \,\mathrm{d}s \int_0^T \sqrt{Y_s}\,\mathrm{d}B_s}{T\int_0^T X_s^2\,\mathrm{d}s - \left(\int_0^T X_s\,\mathrm{d}s\right)^2}, \tag{7.4.9}$$

provided that $T\int_0^T X_s^2\,\mathrm{d}s - \left(\int_0^T X_s\,\mathrm{d}s\right)^2 > 0$.

Remark 7.4.1. For the stochastic integral $\int_0^T X_s\,\mathrm{d}X_s$ in (7.4.2), (7.4.6) and (7.4.7), we have

$$\sum_{i=1}^{\lfloor nT\rfloor} X_{\frac{i-1}{n}}\big(X_{\frac{i}{n}} - X_{\frac{i-1}{n}}\big) \xrightarrow{\mathbb{P}} \int_0^T X_s\,\mathrm{d}X_s \qquad \text{as } n\to\infty,$$

following from Proposition I.4.44 in Jacod and Shiryaev [**95**]. For more details, see Remark 7.3.2. □

The next lemma is about the existence of $\widetilde{\theta}_T^{\mathrm{LSE}}$ (supposing that $m\in\mathbb{R}$ is known, but $a>0$ and b are unknown).

Lemma 7.4.2. *If* $a>0$, $b,m,\theta\in\mathbb{R}$, *and* $\mathbb{P}(Y_0>0)=1$, *then*

$$\mathbb{P}\left(\int_0^T X_s^2\,\mathrm{d}s \in (0,\infty)\right) = 1 \quad \textit{for all } T>0, \tag{7.4.10}$$

and hence there exists a unique LSE $\widetilde{\theta}_T^{\mathrm{LSE}}$ *which has the form given in* (7.4.2).

The next lemma is about the existence of $(\widehat{m}_T^{\mathrm{LSE}}, \widehat{\theta}_T^{\mathrm{LSE}})$.

Lemma 7.4.3. *If* $a>0$, $b,m,\theta\in\mathbb{R}$, *and* $\mathbb{P}(Y_0>0)=1$, *then*

$$\mathbb{P}\left(T\int_0^T X_s^2\,\mathrm{d}s - \left(\int_0^T X_s\,\mathrm{d}s\right)^2 \in (0,\infty)\right) = 1 \quad \textit{for all } T>0, \tag{7.4.11}$$

and hence there exists a unique LSE $(\widehat{m}_T^{\mathrm{LSE}}, \widehat{\theta}_T^{\mathrm{LSE}})$ *which has the form given in* (7.4.6) *and* (7.4.7).

7.5. Consistency of maximum likelihood estimator

Theorem 7.5.1. *If* $a\geqslant\frac{1}{2}$, $b>0$, $m\in\mathbb{R}$, $\theta>0$, *and* $\mathbb{P}(Y_0>0)=1$, *then the MLE of* θ *is strongly consistent:* $\mathbb{P}\left(\lim_{T\to\infty}\widetilde{\theta}_T^{\mathrm{MLE}}=\theta\right)=1$.

Theorem 7.5.2. *If* $a>\frac{1}{2}$, $b>0$, $m\in\mathbb{R}$, $\theta>0$, *and* $\mathbb{P}(Y_0>0)=1$, *then the MLE of* (a,b,m,θ) *is strongly consistent:*

$$\mathbb{P}\left(\lim_{T\to\infty}(\widehat{a}_T^{\mathrm{MLE}}, \widehat{b}_T^{\mathrm{MLE}}, \widehat{m}_T^{\mathrm{MLE}}, \widehat{\theta}_T^{\mathrm{MLE}}) = (a,b,m,\theta)\right) = 1.$$

Remark 7.5.3. If $a=\frac{1}{2}$, $b>0$, $\theta>0$, $m\in\mathbb{R}$, and $\mathbb{P}(Y_0>0)=1$, then one should find another approach for studying the consistency behaviour of the MLE of (a,b,m,θ), since in this case

$$\mathbb{E}\left(\frac{1}{Y_\infty}\right) = \int_0^\infty \frac{2b\mathrm{e}^{-2bx}}{x}\,\mathrm{d}x = \infty,$$

and hence one cannot use part (iii) of Theorem 7.2.5. In this paper we renounce to consider it. □

7.6. Consistency of least squares estimator

THEOREM 7.6.1. *If* $a > 0$, $b > 0$, $m \in \mathbb{R}$, $\theta > 0$, *and* $\mathbb{P}(Y_0 > 0) = 1$, *then the LSE of* θ *is strongly consistent:* $\mathbb{P}\left(\lim_{T\to\infty} \widetilde{\theta}_T^{\mathrm{LSE}} = \theta\right) = 1$.

THEOREM 7.6.2. *If* $a > 0$, $b > 0$, $m \in \mathbb{R}$, $\theta > 0$, *and* $\mathbb{P}(Y_0 > 0) = 1$, *then the LSE of* (m, θ) *is strongly consistent:* $\mathbb{P}\left(\lim_{T\to\infty}(\widehat{m}_T^{\mathrm{LSE}}, \widehat{\theta}_T^{\mathrm{LSE}}) = (m, \theta)\right) = 1$.

7.7. Asymptotic behaviour of maximum likelihood estimator

THEOREM 7.7.1. *If* $a > 1/2$, $b > 0$, $m \in \mathbb{R}$, $\theta > 0$, *and* $\mathbb{P}(Y_0 > 0) = 1$, *then the MLE of* θ *is asymptotically normal, i.e.,*

$$\sqrt{T}(\widetilde{\theta}_T^{\mathrm{MLE}} - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{1}{\mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right)}\right) \qquad \text{as } T \to \infty,$$

where $\mathbb{E}\left(X_\infty^2/Y_\infty\right)$ *is positive and finite.*

THEOREM 7.7.2. *If* $a > 1/2$, $b > 0$, $m \in \mathbb{R}$, $\theta > 0$, *and* $\mathbb{P}(Y_0 > 0) = 1$, *then the MLE of* (a, b, m, θ) *is asymptotically normal, i.e.,*

$$\sqrt{T}\begin{bmatrix} \widehat{a}_T^{\mathrm{MLE}} - a \\ \widehat{b}_T^{\mathrm{MLE}} - b \\ \widehat{m}_T^{\mathrm{MLE}} - m \\ \widehat{\theta}_T^{\mathrm{MLE}} - \theta \end{bmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}_4\left(0, \Sigma^{\mathrm{MLE}}\right) \qquad \text{as } T \to \infty,$$

where $\mathcal{N}_4\left(0, \Sigma^{\mathrm{MLE}}\right)$ *denotes a* 4*-dimensional normally distribution with mean vector* $0 \in \mathbb{R}^4$ *and with covariance matrix* $\Sigma^{\mathrm{MLE}} := \operatorname{diag}(\Sigma_1^{\mathrm{MLE}}, \Sigma_2^{\mathrm{MLE}})$ *with blockdiagonals given by*

$$\Sigma_1^{\mathrm{MLE}} := \frac{1}{\mathbb{E}\left(\frac{1}{Y_\infty}\right)\mathbb{E}(Y_\infty) - 1} D_1, \qquad D_1 := \begin{bmatrix} \mathbb{E}(Y_\infty) & 1 \\ 1 & \mathbb{E}\left(\frac{1}{Y_\infty}\right) \end{bmatrix},$$

$$\Sigma_2^{\mathrm{MLE}} := \frac{1}{\mathbb{E}\left(\frac{1}{Y_\infty}\right)\mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right) - \left(\mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right)\right)^2} D_2, \qquad D_2 := \begin{bmatrix} \mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right) & \mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) \\ \mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right) & \mathbb{E}\left(\frac{1}{Y_\infty}\right) \end{bmatrix}.$$

REMARK 7.7.3. The asymptotic variance $1/\mathbb{E}\left(X_\infty^2/Y_\infty\right)$ of $\widetilde{\theta}_T^{\mathrm{MLE}}$ in Theorem 7.7.1 is less than the asymptotic variance

$$\frac{\mathbb{E}\left(\frac{1}{Y_\infty}\right)}{\mathbb{E}\left(\frac{1}{Y_\infty}\right)\mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right) - \left(\mathbb{E}\left(\frac{X_\infty}{Y_\infty}\right)\right)^2}$$

of $\widehat{\theta}_T^{\mathrm{MLE}}$ in Theorem 7.7.2. This is in accordance with the fact that $\widetilde{\theta}_T^{\mathrm{MLE}}$ is the MLE of θ provided that the value of the parameter m is known, which gives extra information, so the MLE estimator of θ becomes better. □

7.8. Asymptotic behaviour of least squares estimator

THEOREM 7.8.1. *If* $a > 0$, $b > 0$, $m \in \mathbb{R}$, $\theta > 0$, *and* $\mathbb{P}(Y_0 > 0) = 1$, *then the LSE of* θ *is asymptotically normal, i.e.,*

$$\sqrt{T}(\widetilde{\theta}_T^{\mathrm{LSE}} - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\mathbb{E}(X_\infty^2 Y_\infty)}{(\mathbb{E}(X_\infty^2))^2}\right) \qquad \text{as } T \to \infty,$$

where $\mathbb{E}(X_\infty^2 Y_\infty)$ *and* $\mathbb{E}(X_\infty^2)$ *are given explicitly in Theorem 7.2.5.*

REMARK 7.8.2. The asymptotic variance $\mathbb{E}(X_\infty^2 Y_\infty)/(\mathbb{E}(X_\infty^2))^2$ of the LSE $\widetilde{\theta}_T^{\mathrm{LSE}}$ in Theorem 7.8.1 is greater than the asymptotic variance $1/\mathbb{E}\left(X_\infty^2/Y_\infty\right)$ of the MLE $\widetilde{\theta}_T^{\mathrm{MLE}}$ in Theorem 7.7.1, since, by Cauchy and Schwarz's inequality,

$$(\mathbb{E}(X_\infty^2))^2 = \left(\mathbb{E}\left(\frac{X_\infty}{\sqrt{Y_\infty}} X_\infty \sqrt{Y_\infty}\right)\right)^2 < \mathbb{E}\left(\frac{X_\infty^2}{Y_\infty}\right) \mathbb{E}\left(X_\infty^2 Y_\infty\right).$$

Note also that using the limit theorem for $\widetilde{\theta}_T^{\mathrm{LSE}}$ given in Theorem 7.8.1, one can not give asymptotic confidence intervals for θ, since the variance of the limit normal distribution depends on the unknown parameters a and b as well. $\Box$

THEOREM 7.8.3. *If* $a > 0$, $b > 0$, $m \in \mathbb{R}$, $\theta > 0$, *and* $\mathbb{P}(Y_0 > 0) = 1$, *then the LSE of* (m, θ) *is asymptotically normal, i.e.,*

$$\sqrt{T} \begin{bmatrix} \widehat{m}_T^{\mathrm{LSE}} - m \\ \widehat{\theta}_T^{\mathrm{LSE}} - \theta \end{bmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}_2\left(0, \Sigma^{\mathrm{LSE}}\right) \qquad \text{as } T \to \infty,$$

where $\mathcal{N}_2\left(0, \Sigma^{\mathrm{LSE}}\right)$ *denotes a two-dimensional normally distribution with mean vector* $0 \in \mathbb{R}^2$ *and with covariance matrix* $\Sigma^{\mathrm{LSE}} = (\Sigma_{i,j}^{\mathrm{LSE}})_{i,j=1}^2$, *where*

$$\Sigma_{1,1}^{\mathrm{LSE}} := \frac{(\mathbb{E}(X_\infty))^2\, \mathbb{E}(X_\infty^2 Y_\infty) - 2\,\mathbb{E}(X_\infty)\,\mathbb{E}(X_\infty^2)\,\mathbb{E}(X_\infty Y_\infty) + (\mathbb{E}(X_\infty^2))^2\,\mathbb{E}(Y_\infty)}{(\mathbb{E}(X_\infty^2) - (\mathbb{E}(X_\infty))^2)^2},$$

$$\begin{aligned} \Sigma_{1,2}^{\mathrm{LSE}} &= \Sigma_{2,1}^{\mathrm{LSE}} \\ &:= \frac{\mathbb{E}(X_\infty)\big(\mathbb{E}(X_\infty^2 Y_\infty) + \mathbb{E}(X_\infty^2)\,\mathbb{E}(Y_\infty)\big) - \mathbb{E}(X_\infty Y_\infty)\big(\mathbb{E}(X_\infty^2) + (\mathbb{E}(X_\infty))^2\big)}{(\mathbb{E}(X_\infty^2) - (\mathbb{E}(X_\infty))^2)^2}, \end{aligned}$$

$$\Sigma_{2,2}^{\mathrm{LSE}} := \frac{\mathbb{E}(X_\infty^2 Y_\infty) - 2\,\mathbb{E}(X_\infty)\,\mathbb{E}(X_\infty Y_\infty) + (\mathbb{E}(X_\infty))^2\,\mathbb{E}(Y_\infty)}{(\mathbb{E}(X_\infty^2) - (\mathbb{E}(X_\infty))^2)^2}.$$

REMARK 7.8.4. Using the explicit forms of the mixed moments given in (iii) of Theorem 7.2.5, one can check that the asymptotic variance $\mathbb{E}(X_\infty^2 Y_\infty)/(\mathbb{E}(X_\infty^2))^2$ of $\widetilde{\theta}_T^{\mathrm{LSE}}$ in Theorem 7.8.1 is less than the asymptotic variance $\Sigma_{1,1}^{\mathrm{LSE}}$ of $\widehat{\theta}_T^{\mathrm{LSE}}$ in Theorem 7.8.3. This can be interpreted similarly as in Remark 7.7.3. Note also that using the limit theorem for $(\widehat{m}_T^{\mathrm{LSE}}, \widehat{\theta}_T^{\mathrm{LSE}})$ given in Theorem 7.8.3, one can not give asymptotic confidence regions for (m, θ), since the variance matrix of the limit normal distribution depends on the unknown parameters a and b as well. $\Box$

7.9. Appendix: Radon-Nykodim derivatives for certain diffusions

We consider the SDEs

$$\mathrm{d}\xi_t = (A\xi_t + a)\,\mathrm{d}t + \sigma(\xi_t)\,\mathrm{d}W_t, \qquad t \in \mathbb{R}_+, \tag{7.9.1}$$

$$\mathrm{d}\eta_t = (B\eta_t + b)\,\mathrm{d}t + \sigma(\eta_t)\,\mathrm{d}W_t, \qquad t \in \mathbb{R}_+, \tag{7.9.2}$$

with the same initial values $\xi_0 = \eta_0$, where $A, B \in \mathbb{R}^{2\times 2}$, $a, b \in \mathbb{R}^2$, $\sigma : \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$ is a Borel measurable function, and $(W_t)_{t\in\mathbb{R}_+}$ is a two-dimensional standard Wiener process. Suppose that the SDEs (7.9.1) and (7.9.2) admit pathwise unique strong solutions. Let $\mathbb{P}_{(A,a)}$ and $\mathbb{P}_{(B,b)}$ denote the probability measures on the measurable space $(\mathcal{C}(\mathbb{R}_+, \mathbb{R}^2), \mathcal{B}(\mathcal{C}(\mathbb{R}_+, \mathbb{R}^2)))$ induced by the processes $(\xi_t)_{t\in\mathbb{R}_+}$ and $(\eta_t)_{t\in\mathbb{R}_+}$, respectively. Here $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^2)$ denotes the set of continuous $\mathbb{R}^2$-valued functions defined on $\mathbb{R}_+$, $\mathcal{B}(\mathcal{C}(\mathbb{R}_+, \mathbb{R}^2))$ is the Borel σ-algebra on it, and we suppose that the space $(\mathcal{C}(\mathbb{R}_+, \mathbb{R}^2), \mathcal{B}(\mathcal{C}(\mathbb{R}_+, \mathbb{R}^2)))$ is endowed with the natural filtration $(\mathcal{A}_t)_{t\in\mathbb{R}_+}$, given by $\mathcal{A}_t := \varphi_t^{-1}(\mathcal{B}(\mathcal{C}(\mathbb{R}_+, \mathbb{R}^2)))$, where $\varphi_t : \mathcal{C}(\mathbb{R}_+, \mathbb{R}^2) \to \mathcal{C}(\mathbb{R}_+, \mathbb{R}^2)$ is the mapping $\varphi_t(f)(s) := f(t \wedge s)$, $s \in \mathbb{R}_+$. For all $T > 0$, let

$\mathbb{P}_{(A,a),T} := \mathbb{P}_{(A,a)}\,|_{\mathcal{A}_T}$ and $\mathbb{P}_{(B,b),T} := \mathbb{P}_{(B,b)}\,|_{\mathcal{A}_T}$ be the restrictions of $\mathbb{P}_{(A,a)}$ and $\mathbb{P}_{(B,b)}$ to $\mathcal{A}_T$, respectively.

From the very general result in Section 7.6.4 of Liptser and Shiryaev [**117**], one can deduce the following lemma.

LEMMA 7.9.1. *Let* $A, B \in \mathbb{R}^{2\times 2}$, $a, b \in \mathbb{R}^2$, *and let* $\sigma : \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$ *be a Borel measurable function. Suppose that the SDEs* (7.9.1) *and* (7.9.2) *admit pathwise unique strong solutions. Let* $\mathbb{P}_{(A,a)}$ *and* $\mathbb{P}_{(B,b)}$ *denote the probability measures induced by the unique strong solutions of the SDEs* (7.9.1) *and* (7.9.2) *with the same initial value* $\xi_0 = \eta_0$, *respectively. Suppose that* $\mathbb{P}(\exists\,\sigma(\xi_t)^{-1}) = 1$ *and* $\mathbb{P}(\exists\,\sigma(\eta_t)^{-1}) = 1$ *for all* $t \in \mathbb{R}_+$, *and*

$$\begin{aligned}\mathbb{P}\Big(\int_0^t \Big[&(A\xi_s + a)^\top (\sigma(\xi_s)\sigma(\xi_s)^\top)^{-1}(A\xi_s + a)^\top \\ &+ (B\xi_s + b)^\top (\sigma(\xi_s)\sigma(\xi_s)^\top)^{-1}(B\xi_s + b)^\top\Big]\mathrm{d}s < \infty\Big) = 1\end{aligned}$$

and

$$\begin{aligned}\mathbb{P}\Big(\int_0^t \Big[&(A\eta_s + a)^\top (\sigma(\eta_s)\sigma(\eta_s)^\top)^{-1}(A\eta_s + a)^\top \\ &+ (B\eta_s + b)^\top (\sigma(\eta_s)\sigma(\eta_s)^\top)^{-1}(B\eta_s + b)^\top\Big]\mathrm{d}s < \infty\Big) = 1\end{aligned}$$

for all $t \in \mathbb{R}_+$. *Then for all* $T > 0$, *the probability measures* $\mathbb{P}_{(A,a),T}$ *and* $\mathbb{P}_{(B,b),T}$ *are absolutely continuous with respect to each other, and the Radon-Nykodim derivative of* $\mathbb{P}_{(A,a),T}$ *with respect to* $\mathbb{P}_{(B,b),T}$ *(so called likelihood ratio) takes the form*

$$\begin{aligned}L_T^{(A,a),(B,b)}((\xi_s)_{s\in[0,T]}) = \exp\Bigg\{ &\int_0^T (A\xi_s + a - B\xi_s - b)^\top (\sigma(\xi_s)\sigma(\xi_s)^\top)^{-1}\,\mathrm{d}\xi_s \\ -\frac{1}{2}\int_0^T (A\xi_s + a - B\xi_s - b)^\top &(\sigma(\xi_s)\sigma(\xi_s)^\top)^{-1}(A\xi_s + a + B\xi_s + b)^\top \mathrm{d}s\Bigg\}.\end{aligned}$$

We call the attention that conditions (4.110) and (4.111) are also required for Section 7.6.4 in Liptser and Shiryaev [**117**], but the Lipschitz condition (4.110) in Liptser and Shiryaev [**117**] does not hold in general for the SDEs (7.9.1) and (7.9.2). However, we can use formula (7.139) in Liptser and Shiryaev [**117**], since they use their conditions (4.110) and (4.111) only in order to ensure the SDE they consider in Section 7.6.4 has a pathwise unique strong solution (see, the proof of Theorem 7.19 in Liptser and Shiryaev [**117**]), which is supposed in Theorem 7.9.1.

CHAPTER 8

Asymptotic properties of maximum likelihood estimators for Heston models based on continuous time observations

Co-author: Gyula Pap

8.1. Introduction

Affine processes and especially the Heston model have been frequently applied in financial mathematics since they can be well-fitted to financial time series, and also due to their computational tractability. They are characterized by their characteristic function which is exponentially affine in the state variable. A precise mathematical formulation and a complete characterization of regular affine processes are due to Duffie et al. [**63**]. A very recent monograph of Baldeaux and Platen [**10**] gives a detailed survey on affine processes and their applications in financial mathematics.

Let us consider a Heston model

$$\begin{cases} \mathrm{d}Y_t = (a - bY_t)\,\mathrm{d}t + \sigma_1\sqrt{Y_t}\,\mathrm{d}W_t, \\ \mathrm{d}X_t = (\alpha - \beta Y_t)\,\mathrm{d}t + \sigma_2\sqrt{Y_t}\big(\varrho\,\mathrm{d}W_t + \sqrt{1-\varrho^2}\,\mathrm{d}B_t\big), \end{cases} \qquad t \geqslant 0, \tag{8.1.1}$$

where $a > 0$, $b, \alpha, \beta \in \mathbb{R}$, $\sigma_1 > 0$, $\sigma_2 > 0$, $\varrho \in (-1, 1)$ and $(W_t, B_t)_{t\geqslant 0}$ is a two-dimensional standard Wiener process. In this paper we study maximum likelihood estimator (MLE) of (a, b, α, β) based on continuous time observations $(X_t)_{t\in[0,T]}$ with $T > 0$, starting the process (Y, X) from some known non-random initial value $(y_0, x_0) \in (0,\infty) \times \mathbb{R}$. We do not suppose the process $(Y_t)_{t\in[0,T]}$ being observed, since it can be determined using the observations $(X_t)_{t\in[0,T]}$, see Remark 8.2.5. We do not estimate the parameters σ_1, σ_2 and ϱ, since these parameters could —in principle, at least— be determined (rather than estimated) using the observations $(X_t)_{t\in[0,T]}$, see Remark 8.2.6. Further, it will turn out that for the calculation of the MLE of (a, b, α, β), one does not need to know the values of the parameters $\sigma_1 > 0$, $\sigma_2 > 0$, and $\varrho \in (-1, 1)$, see (8.3.2). Note also that $(Y_t, X_t)_{t\geqslant 0}$ is a two-dimensional affine diffusion process with state space $[0, \infty) \times \mathbb{R}$, see Proposition 8.2.1. In the language of financial mathematics, provided that $\beta = \sigma_2^2/2$, one can interpret

$$S_t := \exp\left\{X_t - \alpha + \frac{\sigma_2^2}{2}t\right\}$$

as the asset price, $X_t - \alpha + \frac{\sigma_2^2}{2}t$ as the log-price (log-spot) and $\sigma_2\sqrt{Y_t}$ as the volatility of the asset price at time $t \geqslant 0$. Indeed, using (8.1.1), by an application of Itô's formula, if $\beta = \sigma_2^2/2$, then we have

$$\mathrm{d}S_t = (\alpha + \sigma_2^2/2)S_t\,\mathrm{d}t + \sigma_2\sqrt{Y_t}S_t\big(\varrho\,\mathrm{d}W_t + \sqrt{1-\varrho^2}\,\mathrm{d}B_t\big), \qquad t \geqslant 0,$$

which is Equation (19) in Heston [**84**]. The squared volatility process $(\sigma_2^2 Y_t)_{t\geqslant 0}$ is a continuous time continuous state branching process with immigration, also called Cox–Ingersoll–Ross (CIR) process, first studied by Feller [**68**].

Parameter estimation for continuous time models has a long history, see, e.g., the monographs of Liptser and Shiryaev [**118**, Chapter 17], Kutoyants [**112**] and Bishwal [**38**]. For estimating continuous time models used in finance, Phillips and Yu [**139**] gave an overview of maximum likelihood and Gaussian methods. Since the exact likelihood can be constructed only in special cases (e.g., geometric Brownian motion, Ornstein–Uhlenbeck process, CIR process and inverse square-root process), much attention has been devoted to the development of methods designed to approximate the likelihood.

Aït-Sahalia [**2**] provides closed-form expansions for the log-likelihood function of multivariate diffusions based on discrete time observations. He proved that, under some conditions, the approximate maximum likelihood exists almost surely, and the difference of the approximate and the true maximum likelihood converges in probability to 0 as the time interval separating observations tends to 0. The above mentioned closed-form expansions for the Heston model can be found in Aït-Sahalia and Kimmel [**3**, Appendix A.1]. We note that in Sørensen [**148**] one can find a brief and concise summary of the approach of Aït-Sahalia. In fact, Sørensen [**148**] gives a survey of estimation techniques for stationary and ergodic (one-dimensional) diffusion processes observed at discrete time points. Besides the above mentioned approach of Aït-Sahalia, she recalls estimating functions with special emphasis on martingale estimating functions and so-called simple estimating functions, together with Bayesian analysis of discretely observed diffusion processes.

Azencott and Gadhyan [**9**] considered another parametrization of the Heston model (8.1.1), and they investigated only the subcritical (also called ergodic) case, i.e., when $b > 0$ (see Definition 8.2.3). They developed an algorithm to estimate the parameters of the Heston model based on discrete time observations for the asset price and the volatility. They supposed that $\sigma_2 = 1$ and $\beta = 1/2$, and estimated the parameters σ_1 and ϱ as well. They assumed the time interval separating two consecutive observations also to be unknown and used MLE based on Euler and Milstein discretization schemes. They showed that parameter estimates derived from the Euler scheme using constrained optimization of the approximate MLE are strongly consistent. Note that we obtain results also on the asymptotic behavior of the MLE, and not only in the subcritical case.

Hurn et al. [**93**] developed a quasi-maximum likelihood procedure for estimating the parameters of multidimensional diffusions based on discrete time obervations by replacing the original transition density by a multivariate Gaussian density with first and second moments approximating the true moments of the unknown density. For affine drift and diffusion functions, these moments are exactly those of the true transitional density. As an example, the Heston stochastic volatility model has been analyzed in the subcritical case. However, they did not investigate consistency or asymptotic behavior of their estimators.

Recently, Varughese [**151**] has studied parameter estimation for time inhomogeneous multidimensional diffusion processes given by SDEs based on discrete time observations. The likelihood of a diffusion process in question sampled at discrete time points has been estimated by a so-called saddlepoint approximation. In general, the saddlepoint approximation is an algebraic expression based on a random variable's cumulant generation function. In cases where the first few moments of a random variable are known but the corresponding probability density is difficult to obtain, the saddlepoint approximation to the density can be calculated. The parameter estimates are taken to be the values that maximize this approximate likelihood, which may be estimated by a Markov Chain Monte Carlo (MCMC) procedure. However, the asymptotic properties of the estimators have not been studied.

As an example, the saddlepoint MCMC is used to fit a subcritical Heston model to the S&P 500 and the VIX indices over the period December 2009–November 2010.

In case of the one-dimensional CIR process Y, the parameter estimation of a and b goes back to Overbeck and Rydén [**135**] (conditional least squares estimator (LSE)), Overbeck [**136**] (MLE), and see also Bishwal [**38**, Example 7.6] and the very recent papers of Ben Alaya and Kebaier [**34**], [**35**] (MLE). We also note that Li and Ma [**116**] started to investigate the asymptotic behaviour of the (weighted) conditional LSE of the drift parameters for a CIR model driven by a stable noise (they call it a stable CIR model) from some discretely observed low frequency data set.

To the best knowledge of the authors the parameter estimation problem for multidimensional affine processes has not been tackled so far. Since affine processes are frequently used in financial mathematics, the question of parameter estimation for them needs to be well-investigated. In Barczy et al. [**11**] we started the discussion with a simple non-trivial two-dimensional affine diffusion process given by the SDE

$$\begin{cases} \mathrm{d}Y_t = (a - bY_t)\,\mathrm{d}t + \sqrt{Y_t}\,\mathrm{d}W_t, \\ \mathrm{d}X_t = (m - \theta X_t)\,\mathrm{d}t + \sqrt{Y_t}\,\mathrm{d}B_t, \end{cases} \qquad t \geqslant 0, \tag{8.1.2}$$

where $a > 0$, $b, m, \theta \in \mathbb{R}$, $(W_t, B_t)_{t\geqslant 0}$ is a two-dimensional standard Wiener process. Chen and Joslin [**45**] have found several applications of the model (8.1.2) in financial mathematics, see their equations (25) and (26). In the special critical case $b = 0$, $\theta = 0$ we described the asymptotic behavior of the LSE of (m, θ) based on discrete time observations $X_0, X_1, \ldots, X_n$ as $n \to \infty$. The description of the asymptotic behavior of the LSE of (m, θ) in the other critical cases $b = 0$, $\theta > 0$ or $b > 0$, $\theta = 0$ remained opened. In Barczy et al. [**13**] we dealt with the same model (8.1.2) but in the so-called subcritical (ergodic) case: $b > 0$, $\theta > 0$, and we considered the MLE of (a, b, m, θ) and the LSE of (m, θ) based on continuous time observations. To carry out the analysis in the subcritical case, we needed to examine the question of existence of a unique stationary distribution and ergodicity for the model given by (8.1.2). We solved this problem in a companion paper Barczy et al. [**12**].

Next, we summarize our results comparing with those of Overbeck [**136**] and Ben Alaya and Kebaier [**34**], [**35**], and give an overview of the structure of the paper. Section 8.2 is devoted to some preliminaries. We recall that the SDE (8.1.1) has a pathwise unique strong solution and show that it is a regular affine process, see Proposition 8.2.1. We describe the asymptotic behaviour of the first moment of $(Y_t, X_t)_{t\geqslant 0}$, and, based on it, we introduce a classification of Heston processes given by the SDE (8.1.1), see Proposition 8.2.2 and Definition 8.2.3. Namely, we call $(Y_t, X_t)_{t\geqslant 0}$ subcritical, critical or supercritical if $b > 0$, $b = 0$, or $b < 0$, respectively. We recall a result about existence of a unique stationary distribution and ergodicity for the process $(Y_t)_{t\geqslant 0}$ given by the first equation in (8.1.1) in the subcritical case, see Theorem 8.2.4. From Section 8.3 we will consider the Heston model (8.1.1) with a non-random initial value. In Section 8.3 we study the existence and uniqueness of the MLE of (a, b, α, β) by giving an explicit formula for this MLE as well. It turned out that the MLE of (a, b) based on the observations $(Y_t)_{t\in[0,T]}$ for the CIR process Y is the same as the MLE of (a, b) based on the observations $(X_t)_{t\in[0,T]}$ for the Heston process (Y, X) given by the SDE (8.1.1), see formula (8.3.2) and Overbeck [**136**, formula (2.2)] or Ben Alaya and Kebaier [**35**, Section 3.1].

In Section 8.4 we investigate consistency of MLE. For subcritical Heston models we prove that the MLE of (a, b, α, β) is strongly consistent whenever $a \in \big(\frac{\sigma_1^2}{2}, \infty\big)$ (which is an extension of strong consistency of the MLE of (a, b) proved by

Overbeck [**136**, Theorem 2 (ii)], see Remark 8.4.5), and weakly consistent whenever $a = \frac{\sigma_1^2}{2}$ (which is an extension of weak consistency of the MLE of (a, b) following from part 1 of Theorem 7 in Ben Alaya and Kebaier [**35**], see Remark 8.4.5), see Theorem 8.4.1. For critical Heston models with $a \in \big(\frac{\sigma_1^2}{2}, \infty\big)$, we obtain weak consistency of the MLE of (a, b, α, β) (as a consequence of Theorem 8.6.2), which is an extension of weak consistency of the MLE of (a, b) following from Theorem 6 in Ben Alaya and Kebaier [**35**], see Remark 8.4.6. For supercritical Heston models $a \in \big[\frac{\sigma_1^2}{2}, \infty\big)$, we get strong consistency of the MLE of b, see Theorem 8.4.4, and weak consistency of the MLE of β, see Theorem 8.7.1, and it turns out that the MLE of a and α is not even weakly consistent, see Corollary 8.7.3. This is an extension of Overbeck [**136**, Theorem 2, parts (i) and (v)], see Remark 8.4.7.

Sections 8.5, 8.6 and 8.7 are devoted to study asymptotic behaviour of the MLE of (a, b, α, β) for subcritical, critical and supercritical Heston models, respectively. In Section 8.5 we show that the MLE of (a, b, α, β) is asymptotically normal in the subcritical case with $a \in \big(\frac{\sigma_1^2}{2}, \infty\big)$, which is a generalization of the asymptotic normality of the MLE of (a, b) proved by Ben Alaya and Kebaier [**35**, Theorem 5], see Remark 8.5.2. We also show asymptotic normality with random scaling for the MLE of (a, b, α, β) generalizing the asymptotic normality with random scaling for the MLE of (a, b) due to Overbeck [**136**, Theorem 3 (iii)], see Remark 8.5.2. In Section 8.6 we describe the asymptotic behaviour of the MLE in the critical case with $a \in \big(\frac{\sigma_1^2}{2}, \infty\big)$ generalizing the second part of Theorem 6 in Ben Alaya and Kebaier [**35**], see Remark 8.6.3. It turns out that the MLE of a and α is asymptotically normal, but we have a different limit behaviour for the MLE of b and β, see Theorem 8.6.2. In Theorem 8.6.4 we incorporate random scaling for the MLE of (a, b, α, β) in case of critical Heston models generalizing part (ii) of Theorem 3 in Overbeck [**136**], see Remark 8.6.5. In Section 8.7 for supercritical Heston models with $a \in \big[\frac{\sigma_1^2}{2}, \infty\big)$, we prove that the MLE of a and α has a weak limit without any scaling (consequently, not weakly consistent, see Corollary 8.7.3), and the appropriately normalized MLE of b and β has a mixed normal limit distribution, which is a generalization of the second part of Theorem 3 (i) of Overbeck [**136**], see Remark 8.7.2. We also show asymptotic normality with random scaling for the MLE of (b, β) generalizing the asymptotic normality with random scaling for the MLE of b due to Overbeck [**136**, first part of Theorem 3 (i)], see Remark 8.7.2. In the Appendix we recall some limit theorems for continuous local martingales for studying asymptotic behaviour of the MLE of (a, b, α, β).

In the proofs, mainly for the critical and supercritical cases, we extensively used the following results of Ben Alaya and Kebaier [**34**, Propositions 3 and 4], [**35**, Theorems 4 and 6]: for $b > 0$ and $a = \frac{\sigma_1^2}{2}$, weak convergence of $\frac{1}{T^2}\int_0^T \frac{\mathrm{d}s}{Y_s}$ as $T \to \infty$; for $b = 0$ and $a > \frac{\sigma_1^2}{2}$, the explicit form of the moment generating function of the quadruplet $\big(\log Y_T, Y_T, \int_0^T Y_s \,\mathrm{d}s, \int_0^T \frac{\mathrm{d}s}{Y_s}\big)$, $T > 0$; for $b < 0$ and $a \geqslant \frac{\sigma_1^2}{2}$, a representation of the weak limit of $\big(\mathrm{e}^{bT} Y_T, \int_0^T \frac{\mathrm{d}s}{Y_s}\big)$ as $T \to \infty$. However, our results are not simple consequences of those of Ben Alaya and Kebaier, we will have to find appropriate decompositions of the derived MLEs and then to investigate the joint weak convergence of the components via continuity theorem.

In Barczy et al. [**31**], we study conditional least squares estimation for the drift parameters (a, b, α, β) of the Heston model (8.1.1) starting from some known non-random initial value $(y_0, x_0) \in [0, \infty) \times \mathbb{R}$ based on discrete time observations $(Y_i, X_i)_{i \in \{1, \ldots, n\}}$, and in the subcritical case we describe its asymptotic properties.

Finally, note that Benke and Pap [**36**] study local asymptotic properties of likelihood ratios of the Heston model (8.1.1) under the assumption $a \in \big(\frac{\sigma_1^2}{2}, \infty\big)$. Local asymptotic normality has been proved in the subcritical case and for the submodel when $b = 0$ and $\beta \in \mathbb{R}$ are known in the critical case. Moreover, local asymptotic mixed normality has been shown for the submodel when $a \in \big(\frac{\sigma_1^2}{2}, \infty\big)$ and $\alpha \in \mathbb{R}$ are known in the supercritical case. As a consequence, there exist asymptotic minimax bounds for arbitrary estimators in these models, the MLE (for the appropriate submodels in the critical and supercritical cases) attains this bound for bounded loss function, and the MLE is asymptotically efficient in Hájek's convolution theorem sense, see Benke and Pap [**36**].

8.2. Preliminaries

Let $\mathbb{N}$, $\mathbb{Z}_+$, $\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{R}_{++}$, $\mathbb{R}_-$ and $\mathbb{R}_{--}$ denote the sets of positive integers, non-negative integers, real numbers, non-negative real numbers, positive real numbers, non-positive real numbers and negative real numbers, respectively. For $x, y \in \mathbb{R}$, we will use the notations $x \wedge y := \min(x, y)$ and $x \vee y := \max(x, y)$. By $\|x\|$ and $\|A\|$, we denote the Euclidean norm of a vector $x \in \mathbb{R}^d$ and the induced matrix norm of a matrix $A \in \mathbb{R}^{d\times d}$, respectively. By $\boldsymbol{I}_d \in \mathbb{R}^{d\times d}$, we denote the d-dimensional unit matrix.

Let $\big(\Omega, \mathcal{F}, \mathbb{P}\big)$ be a probability space. By $C_c^2(\mathbb{R}_+ \times \mathbb{R}, \mathbb{R})$ and $C_c^\infty(\mathbb{R}_+ \times \mathbb{R}, \mathbb{R})$, we denote the set of twice continuously differentiable real-valued functions on $\mathbb{R}_+ \times \mathbb{R}$ with compact support, and the set of infinitely differentiable real-valued functions on $\mathbb{R}_+ \times \mathbb{R}$ with compact support, respectively.

The next proposition is about the existence and uniqueness of a strong solution of the SDE (8.1.1) stating also that (Y, X) is a regular affine process. Note that these statements for the first equation of (8.1.1) are well known.

Proposition 8.2.1. *Let (η_0, ζ_0) be a random vector independent of the process $(W_t, B_t)_{t\in\mathbb{R}_+}$ satisfying $\mathbb{P}(\eta_0 \in \mathbb{R}_+) = 1$. Then for all $a \in \mathbb{R}_{++}$, $b, \alpha, \beta \in \mathbb{R}$, $\sigma_1, \sigma_2 \in \mathbb{R}_{++}$, $\varrho \in (-1, 1)$, there is a (pathwise) unique strong solution $(Y_t, X_t)_{t\in\mathbb{R}_+}$ of the SDE (8.1.1) such that $\mathbb{P}((Y_0, X_0) = (\eta_0, \zeta_0)) = 1$ and $\mathbb{P}(Y_t \in \mathbb{R}_+$ for all $t \in \mathbb{R}_+) = 1$. Further, for all $s, t \in \mathbb{R}_+$ with $s \leqslant t$,*

$$\begin{cases} Y_t = \mathrm{e}^{-b(t-s)} \Big(Y_s + a \int_s^t \mathrm{e}^{-b(s-u)}\,\mathrm{d}u + \sigma_1 \int_s^t \mathrm{e}^{-b(s-u)} \sqrt{Y_u}\,\mathrm{d}W_u \Big), \\ X_t = X_s + \int_s^t (\alpha - \beta Y_u)\,\mathrm{d}u + \sigma_2 \int_s^t \sqrt{Y_u}\,(\varrho\,\mathrm{d}W_u + \sqrt{1-\varrho^2}\,\mathrm{d}B_u). \end{cases} \tag{8.2.1}$$

Moreover, $(Y_t, X_t)_{t\in\mathbb{R}_+}$ is a regular affine process with infinitesimal generator

$$\begin{aligned} (\mathcal{A}f)(y, x) &= (a - by) f_1'(y, x) + (\alpha - \beta y) f_2'(y, x) \\ &\qquad + \frac{1}{2} y \big(\sigma_1^2 f_{1,1}''(y, x) + 2\varrho\sigma_1\sigma_2 f_{1,2}''(y, x) + \sigma_2^2 f_{2,2}''(y, x)\big), \end{aligned} \tag{8.2.2}$$

where $(y, x) \in \mathbb{R}_+ \times \mathbb{R}$, $f \in C_c^2(\mathbb{R}_+ \times \mathbb{R}, \mathbb{R})$, and f_i' and $f_{i,j}''$, $i, j \in \{1, 2\}$, denote the first and second order partial derivatives of f with respect to its i-th, and i-th and j-th variables, respectively.

Next we present a result about the first moment of $(Y_t, X_t)_{t\in\mathbb{R}_+}$. We note that Hurn et al. [**93**, Equation (23)] derived the same formula for the expectation of (Y_t, X_t), $t \in \mathbb{R}_+$, by a different method. Note also that the formula for $\mathbb{E}(Y_t)$, $t \in \mathbb{R}_+$, is well known.

PROPOSITION 8.2.2. *Let* $(Y_t, X_t)_{t\in\mathbb{R}_+}$ *be the unique strong solution of the SDE* (8.1.1) *satisfying* $\mathbb{P}(Y_0 \in \mathbb{R}_+) = 1$ *and* $\mathbb{E}(Y_0) < \infty$, $\mathbb{E}(|X_0|) < \infty$. *Then*

$$\begin{bmatrix}\mathbb{E}(Y_t)\\ \mathbb{E}(X_t)\end{bmatrix} = \begin{bmatrix} \mathrm{e}^{-bt} & 0\\ -\beta\int_0^t \mathrm{e}^{-bu}\,\mathrm{d}u & 1\end{bmatrix}\begin{bmatrix}\mathbb{E}(Y_0)\\ \mathbb{E}(X_0)\end{bmatrix} + \begin{bmatrix}\int_0^t \mathrm{e}^{-bu}\,\mathrm{d}u & 0\\ -\beta\int_0^t\left(\int_0^u \mathrm{e}^{-bv}\,\mathrm{d}v\right)\mathrm{d}u & t\end{bmatrix}\begin{bmatrix}a\\ \alpha\end{bmatrix}$$

for $t \in \mathbb{R}_+$. *Consequently, if* $b \in \mathbb{R}_{++}$, *then*

$$\lim_{t\to\infty} \mathbb{E}(Y_t) = \frac{a}{b}, \qquad \lim_{t\to\infty} t^{-1}\,\mathbb{E}(X_t) = \alpha - \frac{\beta a}{b},$$

if $b = 0$, *then*

$$\lim_{t\to\infty} t^{-1}\,\mathbb{E}(Y_t) = a, \qquad \lim_{t\to\infty} t^{-2}\,\mathbb{E}(X_t) = -\frac{1}{2}\beta a,$$

if $b \in \mathbb{R}_{--}$, *then*

$$\lim_{t\to\infty} \mathrm{e}^{bt}\,\mathbb{E}(Y_t) = \mathbb{E}(Y_0) - \frac{a}{b}, \qquad \lim_{t\to\infty} \mathrm{e}^{bt}\,\mathbb{E}(X_t) = \frac{\beta}{b}\,\mathbb{E}(Y_0) - \frac{\beta a}{b^2}.$$

Based on the asymptotic behavior of the expectations $(\mathbb{E}(Y_t), \mathbb{E}(X_t))$ as $t \to \infty$, we introduce a classification of Heston processes given by the SDE (8.1.1).

DEFINITION 8.2.3. Let $(Y_t, X_t)_{t\in\mathbb{R}_+}$ be the unique strong solution of the SDE (8.1.1) satisfying $\mathbb{P}(Y_0 \in \mathbb{R}_+) = 1$. We call $(Y_t, X_t)_{t\in\mathbb{R}_+}$ subcritical, critical or supercritical if $b \in \mathbb{R}_{++}$, $b = 0$ or $b \in \mathbb{R}_{--}$, respectively.

In the sequel $\stackrel{\mathbb{P}}{\longrightarrow}$, $\stackrel{\mathcal{L}}{\longrightarrow}$ and $\stackrel{\text{a.s.}}{\longrightarrow}$ will denote convergence in probability, in distribution and almost surely, respectively.

The following result states the existence of a unique stationary distribution and the ergodicity for the process $(Y_t)_{t\in\mathbb{R}_+}$ given by the first equation in (8.1.1) in the subcritical case, see, e.g., Feller [**68**], Cox et al. [**49**, Equation (20)], Li and Ma [**116**, Theorem 2.6] or Theorem 3.1 with $\alpha = 2$ and Theorem 4.1 in Barczy et al. [**12**].

THEOREM 8.2.4. *Let* $a, b, \sigma_1 \in \mathbb{R}_{++}$. *Let* $(Y_t)_{t\in\mathbb{R}_+}$ *be the unique strong solution of the first equation of the SDE* (8.1.1) *satisfying* $\mathbb{P}(Y_0 \in \mathbb{R}_+) = 1$.

(i) *Then* $Y_t \stackrel{\mathcal{L}}{\longrightarrow} Y_\infty$ *as* $t \to \infty$, *and the distribution of* Y_∞ *is given by*

$$\mathbb{E}(\mathrm{e}^{-\lambda Y_\infty}) = \left(1 + \frac{\sigma_1^2}{2b}\lambda\right)^{-2a/\sigma_1^2}, \qquad \lambda \in \mathbb{R}_+, \tag{8.2.3}$$

i.e., Y_∞ *has Gamma distribution with parameters* $2a/\sigma_1^2$ *and* $2b/\sigma_1^2$, *hence*

$$\mathbb{E}(Y_\infty^\kappa) = \frac{\Gamma\left(\frac{2a}{\sigma_1^2} + \kappa\right)}{\left(\frac{2b}{\sigma_1^2}\right)^\kappa \Gamma\left(\frac{2a}{\sigma_1^2}\right)}, \qquad \kappa \in \left(-\frac{2a}{\sigma_1^2}, \infty\right).$$

Especially, $\mathbb{E}(Y_\infty) = \frac{a}{b}$. *Further, if* $a \in \left(\frac{\sigma_1^2}{2}, \infty\right)$, *then* $\mathbb{E}\left(\frac{1}{Y_\infty}\right) = \frac{2b}{2a-\sigma_1^2}$.

(ii) *Supposing that the random initial value* Y_0 *has the same distribution as* Y_∞, *the process* $(Y_t)_{t\in\mathbb{R}_+}$ *is strictly stationary.*

(iii) *For all Borel measurable functions* $f : \mathbb{R} \to \mathbb{R}$ *such that* $\mathbb{E}(|f(Y_\infty)|) < \infty$, *we have*

$$\frac{1}{T}\int_0^T f(Y_s)\,\mathrm{d}s \stackrel{\text{a.s.}}{\longrightarrow} \mathbb{E}(f(Y_\infty)) \qquad \text{as } T \to \infty. \tag{8.2.4}$$

In the next remark we explain why we suppose only that the process X is observed.

REMARK 8.2.5. If $a \in \mathbb{R}_{++}$, $b, \alpha, \beta \in \mathbb{R}$, $\sigma_1, \sigma_2 \in \mathbb{R}_{++}$, $\varrho \in (-1, 1)$, and $(Y_0, X_0) = (y_0, x_0) \in \mathbb{R}_{++} \times \mathbb{R}$, then, by the SDE (8.1.1),

$$\langle X \rangle_t = \sigma_2^2 \int_0^t Y_s \,\mathrm{d}s, \qquad t \in \mathbb{R}_+.$$

By Theorems I.4.47 a) and I.4.52 in Jacod and Shiryaev [**95**],

$$\sum_{i=1}^{\lfloor nt \rfloor} (X_{\frac{i}{n}} - X_{\frac{i-1}{n}})^2 \xrightarrow{\mathbb{P}} \langle X \rangle_t \qquad \text{as} \ \ n \to \infty, \quad t \in \mathbb{R}_+.$$

This convergence holds almost surely along a suitable subsequence, the members of this sequence are measurable functions of $(X_s)_{s \in [0,t]}$, hence, using Theorems 4.2.2 and 4.2.8 in Dudley [**62**], we obtain that $\langle X \rangle_t = \sigma_2^2 \int_0^t Y_s \,\mathrm{d}s$ is a measurable function of $(X_s)_{s \in [0,t]}$. Moreover,

$$\frac{\langle X \rangle_{t+h} - \langle X \rangle_t}{h} = \frac{\sigma_2^2}{h} \int_t^{t+h} Y_s \,\mathrm{d}s \xrightarrow{\text{a.s.}} \sigma_2^2 Y_t \qquad \text{as} \ \ h \to 0, \quad t \in \mathbb{R}_+, \tag{8.2.5}$$

since Y has almost surely continuous sample paths. In particular,

$$\frac{\langle X \rangle_h}{h y_0} = \frac{\sigma_2^2}{h y_0} \int_0^h Y_s \,\mathrm{d}s \xrightarrow{\text{a.s.}} \sigma_2^2 \frac{Y_0}{y_0} = \sigma_2^2 \qquad \text{as} \ \ h \to 0,$$

hence, for any fixed $T > 0$, σ_2^2 is a measurable function of $(X_s)_{s \in [0,T]}$, i.e., it can be determined from a sample $(X_s)_{s \in [0,T]}$ (provided that (Y, X) starts from some known non-random initial value $(y_0, x_0) \in (0, \infty) \times \mathbb{R}$). However, we also point out that this measurable function remains abstract. Consequently, by (8.2.5), for all $t \in [0, T]$, Y_t is a measurable function of $(X_s)_{s \in [0,T]}$, i.e., it can be determined from a sample $(X_s)_{s \in [0,T]}$ (provided that (Y, X) starts from some known non-random initial value $(y_0, x_0) \in (0, \infty) \times \mathbb{R}$). Finally, we note that the sample size T is fixed above, and it is enough to know any short sample $(X_s)_{s \in [0,T]}$ to carry out the above calculations. □

Next we give statistics for the parameters σ_1, σ_2 and ϱ using continuous time observations $(X_t)_{t \in [0,T]}$ with some $T > 0$ (provided that (Y, X) starts from some known non-random initial value $(y_0, x_0) \in (0, \infty) \times \mathbb{R}$). Due to this result we do not consider the estimation of these parameters, they are supposed to be known.

REMARK 8.2.6. If $a \in \mathbb{R}_{++}$, $b, \alpha, \beta \in \mathbb{R}$, $\sigma_1, \sigma_2 \in \mathbb{R}_{++}$, $\varrho \in (-1, 1)$, and $(Y_0, X_0) = (y_0, x_0) \in \mathbb{R}_{++} \times \mathbb{R}$, then for all $T > 0$,

$$\boldsymbol{S} = \frac{1}{\int_0^T Y_s \,\mathrm{d}s} \begin{bmatrix} \langle Y \rangle_T & \langle Y, X \rangle_T \\ \langle Y, X \rangle_T & \langle X \rangle_T \end{bmatrix} =: \widehat{\boldsymbol{S}}_T \qquad \text{almost surely,}$$

where $(\langle Y, X \rangle_t)_{t \in \mathbb{R}_+}$ denotes the quadratic cross-variation process of Y and X, since, by the SDE (8.1.1),

$$\langle Y \rangle_T = \sigma_1^2 \int_0^T Y_s \,\mathrm{d}s, \qquad \langle X \rangle_T = \sigma_2^2 \int_0^T Y_s \,\mathrm{d}s, \qquad \langle Y, X \rangle_T = \varrho \sigma_1 \sigma_2 \int_0^T Y_s \,\mathrm{d}s.$$

Here $\widehat{\boldsymbol{S}}_T$ is a statistic, i.e., there exists a measurable function $\Xi : C([0, T], \mathbb{R}) \to \mathbb{R}^{2 \times 2}$ such that $\widehat{\boldsymbol{S}}_T = \Xi((X_s)_{s \in [0,T]})$, where $C([0, T], \mathbb{R})$ denotes the space of continuous real-valued functions defined on $[0, T]$, since

$$\frac{1}{\frac{1}{n} \sum_{i=1}^{\lfloor nT \rfloor} Y_{\frac{i-1}{n}}} \sum_{i=1}^{\lfloor nT \rfloor} \begin{bmatrix} Y_{\frac{i}{n}} - Y_{\frac{i-1}{n}} \\ X_{\frac{i}{n}} - X_{\frac{i-1}{n}} \end{bmatrix} \begin{bmatrix} Y_{\frac{i}{n}} - Y_{\frac{i-1}{n}} \\ X_{\frac{i}{n}} - X_{\frac{i-1}{n}} \end{bmatrix}^\top \xrightarrow{\mathbb{P}} \widehat{\boldsymbol{S}}_T \qquad \text{as} \ \ n \to \infty, \tag{8.2.6}$$

where $\lfloor x\rfloor$ denotes the integer part of a real number $x \in \mathbb{R}$, the convergence in (8.2.6) holds almost surely along a suitable subsequence, by Remark 8.2.5, the members of the sequence in (8.2.6) are measurable functions of $(X_s)_{s\in[0,T]}$, and one can use Theorems 4.2.2 and 4.2.8 in Dudley [**62**]. Next we prove (8.2.6). By Theorems I.4.47 a) and I.4.52 in Jacod and Shiryaev [**95**],

$$\sum_{i=1}^{\lfloor nT\rfloor}(Y_{\frac{i}{n}}-Y_{\frac{i-1}{n}})^2 \xrightarrow{\mathbb{P}} \langle Y\rangle_T, \qquad \sum_{i=1}^{\lfloor nT\rfloor}(X_{\frac{i}{n}}-X_{\frac{i-1}{n}})^2 \xrightarrow{\mathbb{P}} \langle X\rangle_T,$$
$$\sum_{i=1}^{\lfloor nT\rfloor}(Y_{\frac{i}{n}}-Y_{\frac{i-1}{n}})(X_{\frac{i}{n}}-X_{\frac{i-1}{n}}) \xrightarrow{\mathbb{P}} \langle Y, X\rangle_T$$

as $n\to\infty$. Consequently,

$$\sum_{i=1}^{\lfloor nT\rfloor}\begin{bmatrix} Y_{\frac{i}{n}}-Y_{\frac{i-1}{n}}\\ X_{\frac{i}{n}}-X_{\frac{i-1}{n}}\end{bmatrix}\begin{bmatrix} Y_{\frac{i}{n}}-Y_{\frac{i-1}{n}}\\ X_{\frac{i}{n}}-X_{\frac{i-1}{n}}\end{bmatrix}^\top \xrightarrow{\mathbb{P}} \left(\int_0^T Y_s\,\mathrm{d}s\right)\widehat{\boldsymbol{S}}_T$$

as $n\to\infty$, see, e.g., van der Vaart [**150**, Theorem 2.7, part (vi)]. Moreover,

$$\frac{1}{n}\sum_{i=1}^{\lfloor nT\rfloor} Y_{\frac{i-1}{n}} \xrightarrow{\text{a.s.}} \int_0^T Y_s\,\mathrm{d}s \qquad \text{as } n\to\infty$$

since Y has almost surely continuous sample paths. Here $\mathbb{P}\big(\int_0^T Y_s\,\mathrm{d}s \in \mathbb{R}_{++}\big)=1$. Indeed, if $\omega\in\Omega$ is such that $[0,T]\ni s\mapsto Y_s(\omega)$ is continuous and $Y_t(\omega)\in\mathbb{R}_+$ for all $t\in\mathbb{R}_+$, then we have $\int_0^T Y_s(\omega)\,\mathrm{d}s = 0$ if and only if $Y_s(\omega)=0$ for all $s\in[0,T]$. Using the method of the proof of Theorem 3.1 in Barczy et. al [**11**], we get $\mathbb{P}(\int_0^T Y_s = 0) = 0$, as desired. Hence (8.2.6) follows by properties of convergence in probability. □

8.3. Existence and uniqueness of MLE

From this section, we will consider the Heston model (8.1.1) with a known non-random initial value $(y_0, x_0)\in\mathbb{R}_{++}\times\mathbb{R}$, and we equip $\big(\Omega,\mathcal{F},\mathbb{P}\big)$ with the augmented filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ corresponding to $(W_t, B_t)_{t\in\mathbb{R}_+}$, constructed as in Karatzas and Shreve [**100**, Section 5.2]. Note that $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ satisfies the usual conditions, i.e., the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ is right-continuous and $\mathcal{F}_0$ contains all the $\mathbb{P}$-null sets in $\mathcal{F}$.

Let $\mathbb{P}_{(Y,X)}$ denote the probability measure induced by $(Y_t,X_t)_{t\in\mathbb{R}_+}$ on the measurable space $(C(\mathbb{R}_+,\mathbb{R}_+\times\mathbb{R}),\mathcal{B}(C(\mathbb{R}_+,\mathbb{R}_+\times\mathbb{R})))$ endowed with the natural filtration $(\mathcal{G}_t)_{t\in\mathbb{R}_+}$, given by $\mathcal{G}_t := \varphi_t^{-1}(\mathcal{B}(C(\mathbb{R}_+,\mathbb{R}_+\times\mathbb{R})))$, $t\in\mathbb{R}_+$, where $\varphi_t : C(\mathbb{R}_+,\mathbb{R}_+\times\mathbb{R}) \to C(\mathbb{R}_+,\mathbb{R}_+\times\mathbb{R})$ is the mapping $\varphi_t(f)(s) := f(t\wedge s)$, $s,t\in\mathbb{R}_+$, $f\in C(\mathbb{R}_+,\mathbb{R}_+\times\mathbb{R})$. Here $C(\mathbb{R}_+,\mathbb{R}_+\times\mathbb{R})$ denotes the set of $\mathbb{R}_+\times\mathbb{R}$-valued continuous functions defined on $\mathbb{R}_+$, and $\mathcal{B}(C(\mathbb{R}_+,\mathbb{R}_+\times\mathbb{R}))$ is the Borel σ-algebra on it. Further, for all $T\in\mathbb{R}_{++}$, let $\mathbb{P}_{(Y,X),T} := \mathbb{P}_{(Y,X)}|_{\mathcal{G}_T}$ be the restriction of $\mathbb{P}_{(Y,X)}$ to $\mathcal{G}_T$.

LEMMA 8.3.1. *Let* $a\in\big[\frac{\sigma_1^2}{2},\infty\big)$, $b,\alpha,\beta\in\mathbb{R}$, $\sigma_1,\sigma_2\in\mathbb{R}_{++}$, *and* $\varrho\in(-1,1)$. *Let* $(Y_t,X_t)_{t\in\mathbb{R}_+}$ *and* $(\widetilde{Y}_t,\widetilde{X}_t)_{t\in\mathbb{R}_+}$ *be the unique strong solutions of the SDE* (8.1.1) *with initial values* $(y_0,x_0)\in\mathbb{R}_{++}\times\mathbb{R}$, $(\widetilde{y}_0,\widetilde{x}_0)\in\mathbb{R}_{++}\times\mathbb{R}$ *such that* $(y_0,x_0)=(\widetilde{y}_0,\widetilde{x}_0)$, *corresponding to the parameters* $(a,b,\alpha,\beta,\sigma_1,\sigma_2,\varrho)$ *and* $(\sigma_1^2,0,0,0,\sigma_1,\sigma_2,\varrho)$, *respectively. Then for all* $T\in\mathbb{R}_{++}$, *the measures* $\mathbb{P}_{(Y,X),T}$ *and* $\mathbb{P}_{(\widetilde{Y},\widetilde{X}),T}$ *are absolutely continuous with respect to each other, and the Radon–Nikodym derivative of* $\mathbb{P}_{(Y,X),T}$ *with respect to* $\mathbb{P}_{(\widetilde{Y},\widetilde{X}),T}$ *(the so called likelihood* *ratio) takes the form*

$$L_T^{(Y,X),(\widetilde{Y},\widetilde{X})}\big((Y_s,X_s)_{s\in[0,T]}\big) = \exp\Bigg\{\int_0^T \frac{1}{Y_s}\begin{bmatrix} a-bY_s-\sigma_1^2 \\ \alpha-\beta Y_s\end{bmatrix}^\top \boldsymbol{S}^{-1}\begin{bmatrix}\mathrm{d}Y_s\\ \mathrm{d}X_s\end{bmatrix} - \frac{1}{2}\int_0^T \frac{1}{Y_s}\begin{bmatrix} a-bY_s-\sigma_1^2 \\ \alpha-\beta Y_s\end{bmatrix}^\top \boldsymbol{S}^{-1}\begin{bmatrix} a-bY_s+\sigma_1^2 \\ \alpha-\beta Y_s\end{bmatrix}\mathrm{d}s\Bigg\},$$

where

$$\boldsymbol{S} := \begin{bmatrix}\sigma_1^2 & \varrho\sigma_1\sigma_2\\ \varrho\sigma_1\sigma_2 & \sigma_2^2\end{bmatrix}. \tag{8.3.1}$$

.

By Lemma 8.3.1, under its conditions the log-likelihood function satisfies

$$\begin{aligned}
&(1-\varrho^2)\log L_T^{(Y,X),(\widetilde{Y},\widetilde{X})}\big((Y_s,X_s)_{s\in[0,T]}\big)\\
&= \int_0^T \frac{1}{Y_s}\Bigg[\left(\frac{a-bY_s-\sigma_1^2}{\sigma_1^2} - \frac{\varrho(\alpha-\beta Y_s)}{\sigma_1\sigma_2}\right)\mathrm{d}Y_s\\
&\qquad + \left(-\frac{\varrho(a-bY_s-\sigma_1^2)}{\sigma_1\sigma_2} + \frac{\alpha-\beta Y_s}{\sigma_2^2}\right)\mathrm{d}X_s\Bigg]\\
&\quad - \frac{1}{2}\int_0^T \frac{1}{Y_s}\left[\frac{(a-bY_s)^2-\sigma_1^4}{\sigma_1^2} - \frac{2\varrho(a-bY_s)(\alpha-\beta Y_s)}{\sigma_1\sigma_2} + \frac{(\alpha-\beta Y_s)^2}{\sigma_2^2}\right]\mathrm{d}s\\
&= a\int_0^T\left(\frac{\mathrm{d}Y_s}{\sigma_1^2 Y_s} - \frac{\varrho\,\mathrm{d}X_s}{\sigma_1\sigma_2 Y_s}\right) + b\int_0^T\left(-\frac{\mathrm{d}Y_s}{\sigma_1^2} + \frac{\varrho\,\mathrm{d}X_s}{\sigma_1\sigma_2}\right)\\
&\quad + \alpha\int_0^T\left(-\frac{\varrho\,\mathrm{d}Y_s}{\sigma_1\sigma_2 Y_s} + \frac{\mathrm{d}X_s}{\sigma_2^2 Y_s}\right) + \beta\int_0^T\left(\frac{\varrho\,\mathrm{d}Y_s}{\sigma_1\sigma_2} - \frac{\mathrm{d}X_s}{\sigma_2^2}\right)\\
&\quad - \frac{1}{2}a^2\int_0^T\frac{\mathrm{d}s}{\sigma_1^2 Y_s} + ab\int_0^T\frac{\mathrm{d}s}{\sigma_1^2} - \frac{1}{2}b^2\int_0^T\frac{Y_s\,\mathrm{d}s}{\sigma_1^2} - \frac{1}{2}\alpha^2\int_0^T\frac{\mathrm{d}s}{\sigma_2^2 Y_s} + \alpha\beta\int_0^T\frac{\mathrm{d}s}{\sigma_2^2}\\
&\quad - \frac{1}{2}\beta^2\int_0^T\frac{Y_s\,\mathrm{d}s}{\sigma_2^2} + a\alpha\int_0^T\frac{\varrho\,\mathrm{d}s}{\sigma_1\sigma_2 Y_s} - (b\alpha+a\beta)\int_0^T\frac{\varrho\,\mathrm{d}s}{\sigma_1\sigma_2} + b\beta\int_0^T\frac{\varrho Y_s\,\mathrm{d}s}{\sigma_1\sigma_2}\\
&\quad - \int_0^T\frac{\mathrm{d}Y_s}{Y_s} + \int_0^T\frac{\varrho\sigma_1\,\mathrm{d}X_s}{\sigma_2 Y_s} + \frac{1}{2}\int_0^T\frac{\sigma_1^2\,\mathrm{d}s}{Y_s}\\
&= \boldsymbol{\theta}^\top\boldsymbol{d}_T - \frac{1}{2}\boldsymbol{\theta}^\top\boldsymbol{A}_T\,\boldsymbol{\theta} - \int_0^T\frac{\mathrm{d}Y_s}{Y_s} + \int_0^T\frac{\varrho\sigma_1\,\mathrm{d}X_s}{\sigma_2 Y_s} + \frac{1}{2}\int_0^T\frac{\sigma_1^2\,\mathrm{d}s}{Y_s},
\end{aligned}$$

where

$$\boldsymbol{\theta} := \begin{bmatrix} a\\ b\\ \alpha\\ \beta\end{bmatrix}, \qquad \boldsymbol{d}_T := \boldsymbol{d}_T^{(\sigma_1,\sigma_2,\varrho)}\big((Y_s,X_s)_{s\in[0,T]}\big) := \begin{bmatrix} \int_0^T\left(\frac{\mathrm{d}Y_s}{\sigma_1^2 Y_s} - \frac{\varrho\,\mathrm{d}X_s}{\sigma_1\sigma_2 Y_s}\right)\\ \int_0^T\left(-\frac{\mathrm{d}Y_s}{\sigma_1^2} + \frac{\varrho\,\mathrm{d}X_s}{\sigma_1\sigma_2}\right)\\ \int_0^T\left(-\frac{\varrho\,\mathrm{d}Y_s}{\sigma_1\sigma_2 Y_s} + \frac{\mathrm{d}X_s}{\sigma_2^2 Y_s}\right)\\ \int_0^T\left(\frac{\varrho\,\mathrm{d}Y_s}{\sigma_1\sigma_2} - \frac{\mathrm{d}X_s}{\sigma_2^2}\right)\end{bmatrix},$$

$$\begin{aligned}
\boldsymbol{A}_T &:= \boldsymbol{A}_T^{(\sigma_1,\sigma_2,\varrho)}\big((Y_s,X_s)_{s\in[0,T]}\big)\\
&:= \begin{bmatrix}
\int_0^T\frac{\mathrm{d}s}{\sigma_1^2 Y_s} & -\int_0^T\frac{\mathrm{d}s}{\sigma_1^2} & -\int_0^T\frac{\varrho\,\mathrm{d}s}{\sigma_1\sigma_2 Y_s} & \int_0^T\frac{\varrho\,\mathrm{d}s}{\sigma_1\sigma_2}\\
-\int_0^T\frac{\mathrm{d}s}{\sigma_1^2} & \int_0^T\frac{Y_s\,\mathrm{d}s}{\sigma_1^2} & \int_0^T\frac{\varrho\,\mathrm{d}s}{\sigma_1\sigma_2} & -\int_0^T\frac{\varrho Y_s\,\mathrm{d}s}{\sigma_1\sigma_2}\\
-\int_0^T\frac{\varrho\,\mathrm{d}s}{\sigma_1\sigma_2 Y_s} & \int_0^T\frac{\varrho\,\mathrm{d}s}{\sigma_1\sigma_2} & \int_0^T\frac{\mathrm{d}s}{\sigma_2^2 Y_s} & -\int_0^T\frac{\mathrm{d}s}{\sigma_2^2}\\
\int_0^T\frac{\varrho\,\mathrm{d}s}{\sigma_1\sigma_2} & -\int_0^T\frac{\varrho Y_s\,\mathrm{d}s}{\sigma_1\sigma_2} & -\int_0^T\frac{\mathrm{d}s}{\sigma_2^2} & \int_0^T\frac{Y_s\,\mathrm{d}s}{\sigma_2^2}
\end{bmatrix}.
\end{aligned}$$

If we fix $\sigma_1, \sigma_2 \in \mathbb{R}_{++}$, $\varrho \in (-1,1)$, the initial value $(y_0, x_0) \in \mathbb{R}_{++} \times \mathbb{R}$, and $T \in \mathbb{R}_{++}$, then the probability measures $\mathbb{P}_{(Y,X),T}$ induced by $(Y_t, X_t)_{t\in\mathbb{R}_+}$ corresponding to the parameters $(a, b, \alpha, \beta, \sigma_1, \sigma_2, \varrho)$, where $a \in \big[\frac{\sigma_1^2}{2}, \infty\big)$, $b, \alpha, \beta \in \mathbb{R}$, are absolutely continuous with respect to each other. Hence it does not matter which measure is taken as a reference measure for defining the MLE (we have chosen the measure corresponding to the parameters $(\sigma_1^2, 0, 0, 0, \sigma_1, \sigma_2, \varrho)$). For more details, see, e.g., Liptser and Shiryaev [**117**, page 35].

The random symmetric matrix $\boldsymbol{A}_T$ can be written as a Kronecker product of a deterministic symmetric matrix and a random symmetric matrix, namely,

$$\boldsymbol{A}_T = \begin{bmatrix} \frac{1}{\sigma_1^2} & -\frac{\varrho}{\sigma_1\sigma_2} \\ -\frac{\varrho}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2} \end{bmatrix} \otimes \begin{bmatrix} \int_0^T \frac{\mathrm{d}s}{Y_s} & -\int_0^T 1\,\mathrm{d}s \\ -\int_0^T 1\,\mathrm{d}s & \int_0^T Y_s\,\mathrm{d}s \end{bmatrix}.$$

The first matrix is strictly positive definite. The second matrix is strictly positive definite if and only if $\int_0^T Y_s\,\mathrm{d}s \int_0^T \frac{\mathrm{d}s}{Y_s} > T^2$. The eigenvalues of $\boldsymbol{A}_T$ coincides with the products of the eigenvalues of the two matrices in question (taking into account their multiplicities), hence the matrix $\boldsymbol{A}_T$ is strictly positive definite if and only if $\int_0^T Y_s\,\mathrm{d}s \int_0^T \frac{\mathrm{d}s}{Y_s} > T^2$, and in this case the inverse $\boldsymbol{A}_T^{-1}$ has the form (applying the identity $(\boldsymbol{A}\otimes\boldsymbol{B})^{-1} = \boldsymbol{A}^{-1}\otimes\boldsymbol{B}^{-1}$)

$$\begin{aligned}\boldsymbol{A}_T^{-1} &= \begin{bmatrix} \frac{1}{\sigma_1^2} & -\frac{\varrho}{\sigma_1\sigma_2} \\ -\frac{\varrho}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2} \end{bmatrix}^{-1} \otimes \begin{bmatrix} \int_0^T \frac{\mathrm{d}s}{Y_s} & -T \\ -T & \int_0^T Y_s\,\mathrm{d}s \end{bmatrix}^{-1} \\ &= \frac{\boldsymbol{S} \otimes \begin{bmatrix} \int_0^T Y_s\,\mathrm{d}s & T \\ T & \int_0^T \frac{\mathrm{d}s}{Y_s} \end{bmatrix}}{(1-\varrho^2)\left(\int_0^T Y_s\,\mathrm{d}s \int_0^T \frac{\mathrm{d}s}{Y_s} - T^2\right)}.\end{aligned}$$

Hence we have

$$\begin{aligned} 2(1-\varrho^2)\log L_T^{(Y,X),(\widetilde{Y},\widetilde{X})}\big((Y_s,X_s)_{s\in[0,T]}\big) &= -(\boldsymbol{\theta} - \boldsymbol{A}_T^{-1}\boldsymbol{d}_T)^\top \boldsymbol{A}_T(\boldsymbol{\theta} - \boldsymbol{A}_T^{-1}\boldsymbol{d}_T) + \boldsymbol{d}_T^\top\boldsymbol{A}_T^{-1}\boldsymbol{d}_T \\ &\quad - 2\int_0^T \frac{\mathrm{d}Y_s}{Y_s} + 2\int_0^T \frac{\varrho\sigma_1\,\mathrm{d}X_s}{\sigma_2 Y_s} + \int_0^T \frac{\sigma_1^2\,\mathrm{d}s}{Y_s}, \end{aligned}$$

provided that $\int_0^T Y_s\,\mathrm{d}s \int_0^T \frac{\mathrm{d}s}{Y_s} > T^2$. Recall that $\sigma_1, \sigma_2 \in \mathbb{R}_{++}$ and $\varrho \in (-1,1)$ are supposed to be known. Then maximizing $(1-\varrho^2)\log L_T^{(Y,X),(\widetilde{Y},\widetilde{X})}\big((Y_s,X_s)_{s\in[0,T]}\big)$ in $(a,b,\alpha,\beta) \in \mathbb{R}^4$ gives the MLE of (a,b,α,β) based on the observations $(X_t)_{t\in[0,T]}$ having the form

$$\widehat{\boldsymbol{\theta}}_T = \begin{bmatrix} \widehat{a}_T \\ \widehat{b}_T \\ \widehat{\alpha}_T \\ \widehat{\beta}_T \end{bmatrix} = \boldsymbol{A}_T^{-1}\boldsymbol{d}_T,$$

provided that $\int_0^T Y_s\,\mathrm{d}s \int_0^T \frac{\mathrm{d}s}{Y_s} > T^2$. The random vector $\boldsymbol{d}_T$ can be expressed as

$$\boldsymbol{d}_T = \begin{bmatrix} \frac{1}{\sigma_1^2} \\ -\frac{\varrho}{\sigma_1\sigma_2} \end{bmatrix} \otimes \begin{bmatrix} \int_0^T \frac{\mathrm{d}Y_s}{Y_s} \\ -\int_0^T \mathrm{d}Y_s \end{bmatrix} + \begin{bmatrix} -\frac{\varrho}{\sigma_1\sigma_2} \\ \frac{1}{\sigma_2^2} \end{bmatrix} \otimes \begin{bmatrix} \int_0^T \frac{\mathrm{d}X_s}{Y_s} \\ -\int_0^T \mathrm{d}X_s \end{bmatrix}.$$

Applying the identity $(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \otimes \boldsymbol{D}) = (\boldsymbol{AC}) \otimes (\boldsymbol{BD})$, we can calculate

$$
\begin{aligned}
&\left(\boldsymbol{S} \otimes \begin{bmatrix} \int_0^T Y_s \, \mathrm{d}s & T \\ T & \int_0^T \frac{\mathrm{d}s}{Y_s} \end{bmatrix}\right) \boldsymbol{d}_T \\
&= \left(\boldsymbol{S} \begin{bmatrix} \frac{1}{\sigma_1^2} \\ -\frac{\varrho}{\sigma_1\sigma_2} \end{bmatrix}\right) \otimes \left(\begin{bmatrix} \int_0^T Y_s \, \mathrm{d}s & T \\ T & \int_0^T \frac{\mathrm{d}s}{Y_s} \end{bmatrix} \begin{bmatrix} \int_0^T \frac{\mathrm{d}Y_s}{Y_s} \\ -\int_0^T \mathrm{d}Y_s \end{bmatrix}\right) \\
&\quad + \left(\boldsymbol{S} \begin{bmatrix} -\frac{\varrho}{\sigma_1\sigma_2} \\ \frac{1}{\sigma_2^2} \end{bmatrix}\right) \otimes \left(\begin{bmatrix} \int_0^T Y_s \, \mathrm{d}s & T \\ T & \int_0^T \frac{\mathrm{d}s}{Y_s} \end{bmatrix} \begin{bmatrix} \int_0^T \frac{\mathrm{d}X_s}{Y_s} \\ -\int_0^T \mathrm{d}X_s \end{bmatrix}\right) \\
&= \begin{bmatrix} 1-\varrho^2 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} \int_0^T Y_s \, \mathrm{d}s \int_0^T \frac{\mathrm{d}Y_s}{Y_s} - T(Y_T - y_0) \\ T \int_0^T \frac{\mathrm{d}Y_s}{Y_s} - (Y_T - y_0) \int_0^T \frac{\mathrm{d}s}{Y_s} \end{bmatrix} \\
&\quad + \begin{bmatrix} 0 \\ 1-\varrho^2 \end{bmatrix} \otimes \begin{bmatrix} \int_0^T Y_s \, \mathrm{d}s \int_0^T \frac{\mathrm{d}X_s}{Y_s} - T(X_T - x_0) \\ T \int_0^T \frac{\mathrm{d}X_s}{Y_s} - (X_T - x_0) \int_0^T \frac{\mathrm{d}s}{Y_s} \end{bmatrix}.
\end{aligned}
$$

Consequently, we obtain

$$
\begin{bmatrix} \widehat{a}_T \\ \widehat{b}_T \\ \widehat{\alpha}_T \\ \widehat{\beta}_T \end{bmatrix} = \frac{1}{\int_0^T Y_s \, \mathrm{d}s \int_0^T \frac{\mathrm{d}s}{Y_s} - T^2} \begin{bmatrix} \int_0^T Y_s \, \mathrm{d}s \int_0^T \frac{\mathrm{d}Y_s}{Y_s} - T(Y_T - y_0) \\ T \int_0^T \frac{\mathrm{d}Y_s}{Y_s} - (Y_T - y_0) \int_0^T \frac{\mathrm{d}s}{Y_s} \\ \int_0^T Y_s \, \mathrm{d}s \int_0^T \frac{\mathrm{d}X_s}{Y_s} - T(X_T - x_0) \\ T \int_0^T \frac{\mathrm{d}X_s}{Y_s} - (X_T - x_0) \int_0^T \frac{\mathrm{d}s}{Y_s} \end{bmatrix}, \tag{8.3.2}
$$

provided that $\int_0^T Y_s \, \mathrm{d}s \int_0^T \frac{\mathrm{d}s}{Y_s} > T^2$. In fact, it turned out that for the calculation of the MLE of (a, b, α, β), one does not need to know the values of the parameters $\sigma_1, \sigma_2 \in \mathbb{R}_{++}$ and $\varrho \in (-1, 1)$. Note that the MLE of (a, b) based on the observations $(X_t)_{t\in[0,T]}$ for the Heston model (Y, X) is the same as the MLE of (a, b) based on the observations $(Y_t)_{t\in[0,T]}$ for the CIR process Y, see, e.g., Overbeck [**136**, formula (2.2)] or Ben Alaya and Kebaier [**35**, Section 3.1].

In the next remark we point out that the MLE (8.3.2) of (a, b, α, β) can be approximated using discrete time observations for X, which can be reassuring for practical applications, where data in continuous record is not available.

REMARK 8.3.2. For the stochastic integrals $\int_0^T \frac{\mathrm{d}X_s}{Y_s}$ and $\int_0^T \frac{\mathrm{d}Y_s}{Y_s}$ in (8.3.2), we have

$$
\sum_{i=1}^{\lfloor nT \rfloor} \frac{X_{\frac{i}{n}} - X_{\frac{i-1}{n}}}{Y_{\frac{i-1}{n}}} \xrightarrow{\mathbb{P}} \int_0^T \frac{\mathrm{d}X_s}{Y_s} \qquad \text{and} \qquad \sum_{i=1}^{\lfloor nT \rfloor} \frac{Y_{\frac{i}{n}} - Y_{\frac{i-1}{n}}}{Y_{\frac{i-1}{n}}} \xrightarrow{\mathbb{P}} \int_0^T \frac{\mathrm{d}Y_s}{Y_s} \tag{8.3.3}
$$

as $n \to \infty$, following from Proposition I.4.44 in Jacod and Shiryaev [**95**] with the Riemann sequence of deterministic subdivisions $\left(\frac{i}{n} \wedge T\right)_{i\in\mathbb{N}}$, $n \in \mathbb{N}$. Thus, there exist measurable functions $\Phi, \Psi : C([0,T], \mathbb{R}) \to \mathbb{R}$ such that $\int_0^T \frac{\mathrm{d}X_s}{Y_s} = \Phi((X_s)_{s\in[0,T]})$ and $\int_0^T \frac{\mathrm{d}Y_s}{Y_s} = \Psi((X_s)_{s\in[0,T]})$, since the convergences in (8.3.3) hold almost surely along suitable subsequences, by Remark 8.2.5, the members of both sequences in (8.3.3) are measurable functions of $(X_s)_{s\in[0,T]}$, and one can use Theorems 4.2.2 and 4.2.8 in Dudley [**62**]. Moreover, since Y has continuous sample paths almost surely,

$$
\frac{1}{n} \sum_{i=1}^{\lfloor nT \rfloor} Y_{\frac{i-1}{n}} \xrightarrow{\text{a.s.}} \int_0^T Y_s \, \mathrm{d}s \qquad \text{and} \qquad \frac{1}{n} \sum_{i=1}^{\lfloor nT \rfloor} \frac{1}{Y_{\frac{i-1}{n}}} \xrightarrow{\text{a.s.}} \int_0^T \frac{\mathrm{d}s}{Y_s}
$$

as $n \to \infty$, hence the right hand side of (8.3.2) is a measurable function of $(X_s)_{s\in[0,T]}$, i.e., it is a statistic. Further, one can define a sequence $(\widehat{\boldsymbol{\theta}}_{T,n})_{n\in\mathbb{N}}$ of estimators of $\boldsymbol{\theta} = (a, b, \alpha, \beta)^\top$ based only on the discrete time observations $(Y_{\frac{i}{n}}, X_{\frac{i}{n}})_{i\in\{1,\ldots,\lfloor nT\rfloor\}}$ such that $\widehat{\boldsymbol{\theta}}_{T,n} \xrightarrow{\mathbb{P}} \widehat{\boldsymbol{\theta}}_T$ as $n \to \infty$. This is also called infill asymptotics. This phenomenon is similar to the approximate MLE, used by Aït-Sahalia [**2**], as discussed in the Introduction. $\square$

Using the SDE (8.1.1) one can check that

$$\begin{bmatrix} \widehat{a}_T - a \\ \widehat{b}_T - b \\ \widehat{\alpha}_T - \alpha \\ \widehat{\beta}_T - \beta \end{bmatrix} = \frac{\begin{bmatrix} \int_0^T Y_s \,\mathrm{d}s \int_0^T \frac{\mathrm{d}Y_s}{Y_s} - T(Y_T - y_0) - a \int_0^T Y_s \,\mathrm{d}s \int_0^T \frac{\mathrm{d}s}{Y_s} + aT^2 \\ T \int_0^T \frac{\mathrm{d}Y_s}{Y_s} - (Y_T - y_0) \int_0^T \frac{\mathrm{d}s}{Y_s} - b \int_0^T Y_s \,\mathrm{d}s \int_0^T \frac{\mathrm{d}s}{Y_s} + bT^2 \\ \int_0^T Y_s \,\mathrm{d}s \int_0^T \frac{\mathrm{d}X_s}{Y_s} - T(X_T - x_0) - \alpha \int_0^T Y_s \,\mathrm{d}s \int_0^T \frac{\mathrm{d}s}{Y_s} + \alpha T^2 \\ T \int_0^T \frac{\mathrm{d}X_s}{Y_s} - (X_T - x_0) \int_0^T \frac{\mathrm{d}s}{Y_s} - \beta \int_0^T Y_s \,\mathrm{d}s \int_0^T \frac{\mathrm{d}s}{Y_s} + \beta T^2 \end{bmatrix}}{\int_0^T Y_s \,\mathrm{d}s \int_0^T \frac{\mathrm{d}s}{Y_s} - T^2}$$

$$(8.3.4) \qquad = \frac{1}{\int_0^T Y_s \,\mathrm{d}s \int_0^T \frac{\mathrm{d}s}{Y_s} - T^2} \begin{bmatrix} \sigma_1 \int_0^T Y_s \,\mathrm{d}s \int_0^T \frac{\mathrm{d}W_s}{\sqrt{Y_s}} - \sigma_1 T \int_0^T \sqrt{Y_s} \,\mathrm{d}W_s \\ \sigma_1 T \int_0^T \frac{\mathrm{d}W_s}{\sqrt{Y_s}} - \sigma_1 \int_0^T \frac{\mathrm{d}s}{Y_s} \int_0^T \sqrt{Y_s} \,\mathrm{d}W_s \\ \sigma_2 \int_0^T Y_s \,\mathrm{d}s \int_0^T \frac{\mathrm{d}\widetilde{W}_s}{\sqrt{Y_s}} - \sigma_2 T \int_0^T \sqrt{Y_s} \,\mathrm{d}\widetilde{W}_s \\ \sigma_2 T \int_0^T \frac{\mathrm{d}\widetilde{W}_s}{\sqrt{Y_s}} - \sigma_2 \int_0^T \frac{\mathrm{d}s}{Y_s} \int_0^T \sqrt{Y_s} \,\mathrm{d}\widetilde{W}_s \end{bmatrix},$$

provided that $\int_0^T Y_s \,\mathrm{d}s \int_0^T \frac{\mathrm{d}s}{Y_s} > T^2$, where the process

$$\widetilde{W}_s := \varrho W_s + \sqrt{1 - \varrho^2} B_s, \qquad s \in \mathbb{R}_+,$$

is a standard Wiener process.

The next lemma is about the existence of $\big(\widehat{a}_T, \widehat{b}_T, \widehat{\alpha}_T, \widehat{\beta}_T\big)$.

LEMMA 8.3.3. *If* $a \in \big[\frac{\sigma_1^2}{2}, \infty\big)$, $b \in \mathbb{R}$, $\sigma_1 \in \mathbb{R}_{++}$, *and* $Y_0 = y_0 \in \mathbb{R}_{++}$, *then*

$$(8.3.5) \qquad \mathbb{P}\left(\int_0^T Y_s \,\mathrm{d}s \int_0^T \frac{1}{Y_s} \,\mathrm{d}s > T^2\right) = 1 \qquad \text{for all } T \in \mathbb{R}_{++},$$

and hence, supposing also that $\alpha, \beta \in \mathbb{R}$, $\sigma_2 \in \mathbb{R}_{++}$, $\varrho \in (-1, 1)$, *and* $X_0 = x_0 \in \mathbb{R}$, *there exists a unique MLE* $\big(\widehat{a}_T, \widehat{b}_T, \widehat{\alpha}_T, \widehat{\beta}_T\big)$ *for all* $T \in \mathbb{R}_{++}$.

8.4. Consistency of MLE

First we consider the case of subcritical Heston models, i.e., when $b \in \mathbb{R}_{++}$.

THEOREM 8.4.1. *If* $b \in \mathbb{R}_{++}$, $\alpha, \beta \in \mathbb{R}$, $\sigma_1, \sigma_2 \in \mathbb{R}_{++}$, $\varrho \in (-1, 1)$, *and* $(Y_0, X_0) = (y_0, x_0) \in \mathbb{R}_{++} \times \mathbb{R}$, *then the MLE of* (a, b, α, β) *is strongly consistent, i.e.,* $\big(\widehat{a}_T, \widehat{b}_T, \widehat{\alpha}_T, \widehat{\beta}_T\big) \xrightarrow{\text{a.s.}} (a, b, \alpha, \beta)$ *as* $T \to \infty$, *whenever* $a \in \left(\frac{\sigma_1^2}{2}, \infty\right)$, *and it is weakly consistent, i.e.,* $\big(\widehat{a}_T, \widehat{b}_T, \widehat{\alpha}_T, \widehat{\beta}_T\big) \xrightarrow{\mathbb{P}} (a, b, \alpha, \beta)$ *as* $T \to \infty$, *whenever* $a = \frac{\sigma_1^2}{2}$.

In order to handle supercritical Heston models, i.e., when $b \in \mathbb{R}_{--}$, we need the following integral version of the Toeplitz Lemma, due to Dietz and Kutoyants [**59**].

LEMMA 8.4.2. *Let* $\{\varphi_T : T \in \mathbb{R}_+\}$ *be a family of probability measures on* $\mathbb{R}_+$ *such that* $\varphi_T([0, T]) = 1$ *for all* $T \in \mathbb{R}_+$, *and* $\lim_{T\to\infty} \varphi_T([0, K]) = 0$ *for all*

$K \in \mathbb{R}_{++}$. *Then for every bounded and measurable function* $f : \mathbb{R}_+ \to \mathbb{R}$ *for which the limit* $f(\infty) := \lim_{t\to\infty} f(t)$ *exists, we have*

$$\lim_{T\to\infty} \int_0^\infty f(t)\, \varphi_T(\mathrm{d}t) = f(\infty).$$

As a special case, we have the following integral version of the Kronecker Lemma, see Küchler and Sørensen [**113**, Lemma B.3.2].

LEMMA 8.4.3. *Let* $a : \mathbb{R}_+ \to \mathbb{R}_+$ *be a measurable function. Put* $b(T) := \int_0^T a(t)\,\mathrm{d}t$, $T \in \mathbb{R}_+$. *Suppose that* $\lim_{T\to\infty} b(T) = \infty$. *Then for every bounded and measurable function* $f : \mathbb{R}_+ \to \mathbb{R}$ *for which the limit* $f(\infty) := \lim_{t\to\infty} f(t)$ *exists, we have*

$$\lim_{T\to\infty} \frac{1}{b(T)} \int_0^T a(t) f(t)\,\mathrm{d}t = f(\infty).$$

The next theorem states strong consistency of the MLE of b in the supercritical case. Overbeck [**136**, Theorem 2, part (i)] contains this result for CIR processes with a slightly incomplete proof.

THEOREM 8.4.4. *If* $a \in \left[\frac{\sigma_1^2}{2}, \infty\right)$, $b \in \mathbb{R}_{--}$, $\alpha, \beta \in \mathbb{R}$, $\sigma_1, \sigma_2 \in \mathbb{R}_{++}$, $\varrho \in (-1, 1)$, *and* $(Y_0, X_0) = (y_0, x_0) \in \mathbb{R}_{++} \times \mathbb{R}$, *then the MLE of* b *is strongly consistent, i.e.,* $\widehat{b}_T \xrightarrow{\text{a.s.}} b$ *as* $T \to \infty$.

REMARK 8.4.5. For subcritical (i.e., $b \in \mathbb{R}_{++}$) CIR models with $a \in \big(\frac{\sigma_1^2}{2}, \infty\big)$, Overbeck [**136**, Theorem 2, part (ii)] proved strong consistency of the MLE of (a, b). For subcritical (i.e., $b \in \mathbb{R}_{++}$) CIR models with $a = \frac{\sigma_1^2}{2}$, weak consistency of the MLE of (a, b) follows from part 1 of Theorem 7 in Ben Alaya and Kebaier [**35**]. □

REMARK 8.4.6. For critical (i.e., $b = 0$) CIR models with $a \in \big[\frac{\sigma_1^2}{2}, \infty\big)$, weak consistency of the MLE of (a, b) follows from Theorem 2 (iii) in Overbeck [**136**] or Theorem 6 in Ben Alaya and Kebaier [**35**]. For critical Heston models with $a \in (\frac{\sigma_1^2}{2}, \infty)$, weak consistency of the MLE of (a, b, α, β) is a consequence of Theorem 8.6.2. □

REMARK 8.4.7. For supercritical (i.e., $b \in \mathbb{R}_{--}$) CIR models with $a \in \big[\frac{\sigma_1^2}{2}, \infty\big)$, Overbeck [**136**, Theorem 2, parts (i) and (v)] proved that the MLE of b is strongly consistent, however, there is no strongly consistent estimator of a. See also Ben Alaya and Kebaier [**35**, Theorem 7, part 2]. For supercritical Heston models with $a \in \big[\frac{\sigma_1^2}{2}, \infty\big)$, it will turn out that the MLE of a and α is not even weakly consistent, but the MLE of β is weakly consistent, see Theorem 8.7.1. □

8.5. Asymptotic behaviour of MLE: subcritical case

We consider subcritical Heston models, i.e., when $b \in \mathbb{R}_{++}$.

THEOREM 8.5.1. *If* $a \in \left(\frac{\sigma_1^2}{2}, \infty\right)$, $b \in \mathbb{R}_{++}$, $\alpha, \beta \in \mathbb{R}$, $\sigma_1, \sigma_2 \in \mathbb{R}_{++}$, $\varrho \in (-1, 1)$, *and* $(Y_0, X_0) = (y_0, x_0) \in \mathbb{R}_{++} \times \mathbb{R}$, *then the MLE of* (a, b, α, β) *is asymptotically normal, i.e.,*

$$\sqrt{T} \begin{bmatrix} \widehat{a}_T - a \\ \widehat{b}_T - b \\ \widehat{\alpha}_T - \alpha \\ \widehat{\beta}_T - \beta \end{bmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}_4\left(\mathbf{0}, \boldsymbol{S} \otimes \begin{bmatrix} \frac{2b}{2a-\sigma_1^2} & -1 \\ -1 & \frac{a}{b} \end{bmatrix}^{-1}\right) \qquad \text{as } T \to \infty, \tag{8.5.1}$$

where $\boldsymbol{S}$ *is defined in* (8.3.1).

With a random scaling, we have

$$\frac{1}{\left(\int_0^T \frac{\mathrm{d}s}{Y_s}\right)^{1/2}} \left(\boldsymbol{I}_2 \otimes \begin{bmatrix} \int_0^T \frac{\mathrm{d}s}{Y_s} & -T \\ 0 & \left(\int_0^T Y_s \, \mathrm{d}s \int_0^T \frac{\mathrm{d}s}{Y_s} - T^2\right)^{1/2} \end{bmatrix} \right) \begin{bmatrix} \widehat{a}_T - a \\ \widehat{b}_T - b \\ \widehat{\alpha}_T - \alpha \\ \widehat{\beta}_T - \beta \end{bmatrix} \tag{8.5.2}$$

$$\stackrel{\mathcal{L}}{\longrightarrow} \mathcal{N}_4 \left(\boldsymbol{0}, \boldsymbol{S} \otimes \boldsymbol{I}_2\right) \qquad \text{as } T \to \infty.$$

REMARK 8.5.2. For subcritical (i.e., $b \in \mathbb{R}_{++}$) CIR models, for the MLE of (a, b), Ben Alaya and Kebaier [**35**, Theorems 5 and 7] proved asymptotic normality whenever $a \in \left(\frac{\sigma_1^2}{2}, \infty\right)$, and derived a limit theorem with a non-normal limit distribution whenever $a = \frac{\sigma_1^2}{2}$. For subcritical (i.e., $b \in \mathbb{R}_{++}$) CIR models, for the MLE of (a, b), with random scaling, Overbeck [**136**, Theorem 3 (iii)] showed asymptotic normality. □

8.6. Asymptotic behaviour of MLE: critical case

We consider critical Heston models, i.e., when $b = 0$. First we present an auxiliary lemma.

LEMMA 8.6.1. *The mapping* $C(\mathbb{R}_+, \mathbb{R}) \ni f \mapsto \left(\int_0^t f(u) \, \mathrm{d}u\right)_{t \in \mathbb{R}_+} \in C(\mathbb{R}_+, \mathbb{R})$ *is continuous, hence measurable, where* $C(\mathbb{R}_+, \mathbb{R})$ *denotes the set of real-valued continuous functions defined on* $\mathbb{R}_+$.

The next result can be considered as a generalization of part 2 of Theorem 6 in Ben Alaya and Kebaier [**35**] for critical Heston models.

THEOREM 8.6.2. *If* $a \in \left(\frac{\sigma_1^2}{2}, \infty\right)$, $b = 0$, $\alpha, \beta \in \mathbb{R}$, $\sigma_1, \sigma_2 \in \mathbb{R}_{++}$, $\varrho \in (-1, 1)$ *and* $(Y_0, X_0) = (y_0, x_0) \in \mathbb{R}_{++} \times \mathbb{R}$, *then*

$$\begin{bmatrix} \sqrt{\log T}(\widehat{a}_T - a) \\ \sqrt{\log T}(\widehat{\alpha}_T - \alpha) \\ T\widehat{b}_T \\ T(\widehat{\beta}_T - \beta) \end{bmatrix} \stackrel{\mathcal{L}}{\longrightarrow} \begin{bmatrix} \left(a - \frac{\sigma_1^2}{2}\right)^{1/2} \boldsymbol{S}^{1/2} \boldsymbol{Z}_2 \\ \frac{a - \mathcal{Y}_1}{\int_0^1 \mathcal{Y}_s \, \mathrm{d}s} \\ \frac{\alpha - \mathcal{X}_1}{\int_0^1 \mathcal{Y}_s \, \mathrm{d}s} \end{bmatrix} \qquad \text{as } T \to \infty, \tag{8.6.1}$$

where $(\mathcal{Y}_t, \mathcal{X}_t)_{t \in \mathbb{R}_+}$ *is the unique strong solution of the SDE*

$$\begin{cases} \mathrm{d}\mathcal{Y}_t = a \, \mathrm{d}t + \sigma_1 \sqrt{\mathcal{Y}_t} \, \mathrm{d}\mathcal{W}_t, \\ \mathrm{d}\mathcal{X}_t = \alpha \, \mathrm{d}t + \sigma_2 \sqrt{\mathcal{Y}_t} \big(\varrho \, \mathrm{d}\mathcal{W}_t + \sqrt{1 - \varrho^2} \, \mathrm{d}\mathcal{B}_t\big), \end{cases} \qquad t \in \mathbb{R}_+, \tag{8.6.2}$$

with initial value $(\mathcal{Y}_0, \mathcal{X}_0) = (0, 0)$, *where* $(\mathcal{W}_t, \mathcal{B}_t)_{t \in \mathbb{R}_+}$ *is a two-dimensional standard Wiener process,* $\boldsymbol{Z}_2$ *is a two-dimensional standard normally distributed random vector independent of* $\left(\mathcal{Y}_1, \int_0^1 \mathcal{Y}_t \, \mathrm{d}t, \mathcal{X}_1\right)$, $\boldsymbol{S}$ *is defined in* (8.3.1), *and* $\boldsymbol{S}^{1/2}$ *denotes its uniquely determined symmetric, positive definite square root.*

REMARK 8.6.3. (i) As a consequence of Theorem 8.6.2 we get back the description of the asymptotic behavior of the MLE of (a, b) for the CIR process $(Y_t)_{t \in \mathbb{R}_+}$ in the critical case whenever $a \in \left(\frac{\sigma_1^2}{2}, \infty\right)$ proved by Ben Alaya and Kebaier [**35**, Theorem 6, part 2]. We note that Ben Alaya and Kebaier [**35**, Theorem 6, part 1] described the asymptotic behavior of the MLE of (a, b) in the critical case for the CIR process $(Y_t)_{t \in \mathbb{R}_+}$ with $a = \frac{\sigma_1^2}{2}$ as well.
(ii) Theorem 8.6.2 does not cover the case $a = \frac{\sigma_1^2}{2}$, we renounce to consider it.

(iii) Ben Alaya and Kebaier's proof of part 2 of their Theorem 6 relies on an explicit form of the moment generating-Laplace transform of the quadruplet

$$\left(\log Y_t, Y_t, \int_0^t Y_s \,\mathrm{d}s, \int_0^t \frac{\mathrm{d}s}{Y_s}\right), \qquad t \in \mathbb{R}_+.$$

Using this explicit form, they derived convergence

$$\begin{aligned}(8.6.3)\qquad &\left(\frac{\log Y_T - \log y_0 + \left(\frac{\sigma_1^2}{2} - a\right)\int_0^T \frac{\mathrm{d}s}{Y_s}}{\sqrt{\log T}}, \frac{Y_T}{T}, \frac{1}{T^2}\int_0^T Y_s \,\mathrm{d}s\right)\\ &\qquad \xrightarrow{\mathcal{L}} \left(\frac{\sigma_1}{\sqrt{a - \frac{\sigma_1^2}{2}}} Z_1, \mathcal{Y}_1, \int_0^1 \mathcal{Y}_s \,\mathrm{d}s\right) \qquad \text{as } T \to \infty,\end{aligned}$$

where Z_1 is a one-dimensional standard normally distributed random variable independent of $\left(\mathcal{Y}_1, \int_0^1 \mathcal{Y}_t \,\mathrm{d}t\right)$, which is a corner stone of the proof of our Theorem 8.6.2. □

The next theorem can be considered as a counterpart of Theorem 8.6.2 by incorporating random scaling.

Theorem 8.6.4. *If* $a \in \left(\frac{\sigma_1^2}{2}, \infty\right)$, $b = 0$, $\alpha, \beta \in \mathbb{R}$, $\sigma_1, \sigma_2 \in \mathbb{R}_{++}$, $\varrho \in (-1, 1)$ *and* $(Y_0, X_0) = (y_0, x_0) \in \mathbb{R}_{++} \times \mathbb{R}$, *then*

$$(8.6.4)\qquad \begin{bmatrix} \left(\int_0^T \frac{\mathrm{d}s}{Y_s}\right)^{1/2} (\widehat{a}_T - a) \\ \left(\int_0^T \frac{\mathrm{d}s}{Y_s}\right)^{1/2} (\widehat{\alpha}_T - \alpha) \\ \left(\int_0^T Y_s \,\mathrm{d}s\right)^{1/2} \widehat{b}_T \\ \left(\int_0^T Y_s \,\mathrm{d}s\right)^{1/2} (\widehat{\beta}_T - \beta) \end{bmatrix} \xrightarrow{\mathcal{L}} \begin{bmatrix} \boldsymbol{S}^{1/2} \boldsymbol{Z}_2 \\ \frac{a - \mathcal{Y}_1}{\left(\int_0^1 \mathcal{Y}_s \,\mathrm{d}s\right)^{1/2}} \\ \frac{\alpha - \mathcal{X}_1}{\left(\int_0^1 \mathcal{Y}_s \,\mathrm{d}s\right)^{1/2}} \end{bmatrix} \qquad \text{as } T \to \infty,$$

where $(\mathcal{Y}_t, \mathcal{X}_t)_{t \in \mathbb{R}_+}$ *is the unique strong solution of the SDE* (8.6.2) *with initial value* $(\mathcal{Y}_0, \mathcal{X}_0) = (0, 0)$, $\boldsymbol{Z}_2$ *is a two-dimensional standard normally distributed random vector independent of* $\left(\mathcal{Y}_1, \int_0^1 \mathcal{Y}_t \,\mathrm{d}t, \mathcal{X}_1\right)$, *and* $\boldsymbol{S}$ *is defined in* (8.3.1).

Remark 8.6.5. For a critical (i.e., $b = 0$) CIR models with $a \in \left(\frac{\sigma_1^2}{2}, \infty\right)$, using random scaling, Overbeck [**136**, Theorem 3, part (ii)] has already described the asymptotic behaviour of $\widehat{a}_T$ and $\widehat{b}_T$ separately, but he did not consider their joint asymptotic behaviour. □

8.7. Asymptotic behaviour of MLE: supercritical case

We consider supercritical Heston models, i.e., when $b \in \mathbb{R}_{--}$.

Theorem 8.7.1. *If* $a \in \left[\frac{\sigma_1^2}{2}, \infty\right)$, $b \in \mathbb{R}_{--}$, $\alpha, \beta \in \mathbb{R}$, $\sigma_1, \sigma_2 \in \mathbb{R}_{++}$, $\varrho \in (-1, 1)$, *and* $(Y_0, X_0) = (y_0, x_0) \in \mathbb{R}_{++} \times \mathbb{R}$, *then*

$$(8.7.1)\qquad \begin{bmatrix} \widehat{a}_T - a \\ \widehat{\alpha}_T - \alpha \\ \mathrm{e}^{-bT/2}(\widehat{b}_T - b) \\ \mathrm{e}^{-bT/2}(\widehat{\beta}_T - \beta) \end{bmatrix} \xrightarrow{\mathcal{L}} \begin{bmatrix} \widetilde{\mathcal{V}} \\ \varrho \frac{\sigma_2}{\sigma_1} \widetilde{\mathcal{V}} + \sigma_2 \sqrt{1 - \varrho^2} \left(\int_0^{-1/b} \widetilde{\mathcal{Y}}_u \,\mathrm{d}u\right)^{-1/2} Z_1 \\ \left(-\frac{\widetilde{\mathcal{Y}}_{-1/b}}{b}\right)^{-1/2} \boldsymbol{S}^{1/2} \boldsymbol{Z}_2 \end{bmatrix}$$

as $T \to \infty$, *where* $(\widetilde{\mathcal{Y}}_t)_{t \in \mathbb{R}_+}$ *is a CIR process given by the SDE*

$$\mathrm{d}\widetilde{\mathcal{Y}}_t = a \mathrm{d}t + \sigma_1 \sqrt{\widetilde{\mathcal{Y}}_t} \,\mathrm{d}\mathcal{W}_t, \qquad t \in \mathbb{R}_+,$$

with initial value $\widetilde{\mathcal{Y}}_0 = y_0$, *where* $(\mathcal{W}_t)_{t\in\mathbb{R}_+}$ *is a standard Wiener process,*

$$\widetilde{\mathcal{V}} := \frac{\log \widetilde{\mathcal{Y}}_{-1/b} - \log y_0}{\int_0^{-1/b} \widetilde{\mathcal{Y}}_u \, \mathrm{d}u} + \frac{\sigma_1^2}{2} - a,$$

Z_1 *is a one-dimensional standard normally distributed random variable,* $\boldsymbol{Z}_2$ *is a two-dimensional standard normally distributed random vector such that* Z_1, $\boldsymbol{Z}_2$ *and* $(\widetilde{\mathcal{Y}}_{-1/b}, \int_0^{-1/b} \widetilde{\mathcal{Y}}_u \, \mathrm{d}u)$ *are independent, and* $\boldsymbol{S}$ *is defined in* (8.3.1).

With a random scaling, we have

$$\begin{bmatrix} \widehat{a}_T - a \\ \widehat{\alpha}_T - \alpha \\ \left(\int_0^T Y_s \, \mathrm{d}s\right)^{1/2} (\widehat{b}_T - b) \\ \left(\int_0^T Y_s \, \mathrm{d}s\right)^{1/2} (\widehat{\beta}_T - \beta) \end{bmatrix} \xrightarrow{\mathcal{L}} \begin{bmatrix} \widetilde{\mathcal{V}} \\ \varrho \frac{\sigma_2}{\sigma_1} \widetilde{\mathcal{V}} + \sigma_2 \sqrt{1-\varrho^2} \left(\int_0^{-1/b} \widetilde{\mathcal{Y}}_u \, \mathrm{d}u\right)^{-1/2} Z_1 \\ \boldsymbol{S}^{1/2} \boldsymbol{Z}_2 \end{bmatrix}$$

as $T \to \infty$.

Remark 8.7.2. Overbeck [**136**, Theorem 3] has already derived the asymptotic behaviour of $\widehat{b}_T$ with non-random and random scaling for supercritical CIR processes. We also note that Ben Alaya and Kebaier [**34**, Theorem 1, Case 3] described the asymptotic behavior of the MLE of b for supercritical CIR processes supposing that $a \in \mathbb{R}_{++}$ is known. It turns out that in this case the limit distribution is different from that we have in (8.7.1). □

Corollary 8.7.3. *Under the conditions of Theorem 8.7.1, the MLEs of* b *and* β *are weakly consistent, however, the MLEs of* a *and* α *are not weakly consistent. (Recall also that earlier it turned out that the MLE of* b *is in fact strongly consistent, see Theorem 8.4.4.)*

8.8. Appendix: Limit theorems for continuous local martingales

In what follows we recall some limit theorems for continuous local martingales. We use these limit theorems for studying the asymptotic behaviour of the MLE of (a, b, α, β). First we recall a strong law of large numbers for continuous local martingales.

Theorem 8.8.1. *(Liptser and Shiryaev* [**118**, Lemma 17.4]*) Let us suppose that* $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\big)$ *is a filtered probability space satisfying the usual conditions. Let* $(M_t)_{t\in\mathbb{R}_+}$ *be a square-integrable continuous local martingale with respect to the filtration* $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ *such that* $\mathbb{P}(M_0 = 0) = 1$. *Let* $(\xi_t)_{t\in\mathbb{R}_+}$ *be a progressively measurable process such that*

$$\mathbb{P}\left(\int_0^t \xi_u^2 \, \mathrm{d}\langle M\rangle_u < \infty\right) = 1, \qquad t \in \mathbb{R}_+,$$

and

$$\int_0^t \xi_u^2 \, \mathrm{d}\langle M\rangle_u \xrightarrow{\text{a.s.}} \infty \qquad \text{as } t \to \infty, \tag{8.8.1}$$

where $(\langle M\rangle_t)_{t\in\mathbb{R}_+}$ *denotes the quadratic variation process of* M. *Then*

$$\frac{\int_0^t \xi_u \, \mathrm{d}M_u}{\int_0^t \xi_u^2 \, \mathrm{d}\langle M\rangle_u} \xrightarrow{\text{a.s.}} 0 \qquad \text{as } t \to \infty. \tag{8.8.2}$$

If $(M_t)_{t\in\mathbb{R}_+}$ *is a standard Wiener process, the progressive measurability of* $(\xi_t)_{t\in\mathbb{R}_+}$ *can be relaxed to measurability and adaptedness to the filtration* $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$.

The next theorem is about the asymptotic behaviour of continuous multivariate local martingales, see van Zanten [**158**, Theorem 4.1].

Theorem 8.8.2. *(van Zanten* [**158**, Theorem 4.1]*) Let* $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\big)$ *be a filtered probability space satisfying the usual conditions. Let* $(\boldsymbol{M}_t)_{t\in\mathbb{R}_+}$ *be a d-dimensional square-integrable continuous local martingale with respect to the filtration* $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ *such that* $\mathbb{P}(\boldsymbol{M}_0 = \boldsymbol{0}) = 1$. *Suppose that there exists a function* $\boldsymbol{Q} : \mathbb{R}_+ \to \mathbb{R}^{d\times d}$ *such that* $\boldsymbol{Q}(t)$ *is an invertible (non-random) matrix for all* $t \in \mathbb{R}_+$, $\lim_{t\to\infty} \|\boldsymbol{Q}(t)\| = 0$ *and*

$$\boldsymbol{Q}(t)\langle \boldsymbol{M}\rangle_t\, \boldsymbol{Q}(t)^\top \xrightarrow{\mathbb{P}} \boldsymbol{\eta}\boldsymbol{\eta}^\top \qquad \text{as } t \to \infty,$$

where $\boldsymbol{\eta}$ *is a* $d \times d$ *random matrix. Then, for each* $\mathbb{R}^k$*-valued random vector* $\boldsymbol{v}$ *defined on* $(\Omega, \mathcal{F}, \mathbb{P})$, *we have*

$$(\boldsymbol{Q}(t)\boldsymbol{M}_t, \boldsymbol{v}) \xrightarrow{\mathcal{L}} (\boldsymbol{\eta}\boldsymbol{Z}, \boldsymbol{v}) \qquad \text{as } t \to \infty,$$

where $\boldsymbol{Z}$ *is a d-dimensional standard normally distributed random vector independent of* $(\boldsymbol{\eta}, \boldsymbol{v})$.

We note that Theorem 8.8.2 remains true if the function $\boldsymbol{Q}$ is defined only on an interval $[t_0, \infty)$ with some $t_0 \in \mathbb{R}_{++}$.

To derive consequences of Theorem 8.8.2 one can use the following lemma which is a multidimensional version of Lemma 3 due to Kátai and Mogyoródi [**102**], see Barczy and Pap [**26**, Lemma 3].

Lemma 8.8.3. *Let* $(\boldsymbol{U}_t)_{t\in\mathbb{R}_+}$ *be a k-dimensional stochastic process such that* $\boldsymbol{U}_t$ *converges in distribution as* $t \to \infty$. *Let* $(\boldsymbol{V}_t)_{t\in\mathbb{R}_+}$ *be an* ℓ*-dimensional stochastic process such that* $\boldsymbol{V}_t \xrightarrow{\mathbb{P}} \boldsymbol{V}$ *as* $t \to \infty$, *where* $\boldsymbol{V}$ *is an* ℓ*-dimensional random vector. If* $g : \mathbb{R}^k \times \mathbb{R}^\ell \to \mathbb{R}^d$ *is a continuous function, then*

$$g(\boldsymbol{U}_t, \boldsymbol{V}_t) - g(\boldsymbol{U}_t, \boldsymbol{V}) \xrightarrow{\mathbb{P}} \boldsymbol{0} \qquad \text{as } t \to \infty.$$

Bibliography

[1] Adler, R. J. (1990). *An Introduction to Continuity, Extrema, and Related Topics for General Gaussian Processes.* Lecture Notes–Monograph Series 12, Institute of Mathematical Statistics, Hayward, CA.

[2] Aït-Sahalia, Y. (2008). Closed-form likelihood expansions for multivariate diffusions. *The Annals of Statististics* 36(2) 906–937.

[3] Aït-Sahalia, Y. and Kimmel, R. (2007). Maximum likelihood estimation of stochastic volatility models. *Journal of Financial Economics* 83 413–452.

[4] Albanese, C. and Lawi, S. (2005). Laplace transforms for integrals of Markov processes. *Markov Processes and Related Fields* 11(4) 677–724.

[5] Andersen, L. B. G. and Piterbarg, V. V. (2007). Moment explosions in stochastic volatility models. *Finance and Stochastics* 11(1) 29–50.

[6] Anderson, T. W. (1959). On asymptotic distributions of estimates of parameters of stochastic difference equations. *The Annals of Mathematical Statistics* 30 676–687.

[7] Arató, M. (1982). *Linear Stochastic Systems With Constant Coefficients.* A Statistical Approach: Lecture Notes in Control and Information Sciences (45). Springer.

[8] Ash, R. B. and Gardner, M. F. (1975). *Topics in Stochastic Processes.* Academic Press, New York.

[9] Azencott, R. and Gadhyan, Y. (2009). Accurate parameter estimation for coupled stochastic dynamics. *Discrete and Continuous Dynamical Systems, Series A.* Dynamical Systems, Differential Equations and Applications. 7th AIMS Conference, Arlington, Texas, USA, suppl., 44–53.

[10] Baldeaux, J. and Platen, E. (2013). *Functionals of multidimensional diffusions with applications to finance.* Bocconi & Springer Series 5, Springer, Cham, Bocconi University Press, Milano.

[11] Barczy, M., Doering, L., Li, Z. and Pap, G. (2013). On parameter estimation for critical affine processes. *Electronic Journal of Statistics* 7 647–696.

[12] Barczy, M., Doering, L., Li, Z. and Pap, G. (2014). Stationarity and ergodicity for an affine two factor model. *Advances in Applied Probability* 46(3) 878–898.

[13] Barczy, M., Doering, L., Li, Z. and Pap, G. (2014). Parameter estimation for a subcritical affine two factor model. *Journal of Statistical Planning and Inference* 151-152 37–59.

[14] Barczy, M. and Iglói, E. (2011). Karhunen-Loève expansions of alpha-Wiener bridges. *Central European Journal of Mathematics* 9(1) 65–84.

[15] Barczy, M., Ispány, M. and Pap, G. (2010). Asymptotic behavior of CLS estimators of autoregressive parameter for nonprimitive unstable INAR(2) models. Available on the ArXiv: `http://arxiv.org/abs/1006.4641`

[16] Barczy, M. and Kern, P. (2010). Representations of multidimensional linear process bridges. Available on the ArXiv: `http://arxiv.org/abs/1011.0067`. Extended version of [**18**].

[17] Barczy, M. and Kern, P. (2011). General alpha-Wiener bridges. *Communications on Stochastic Analysis* 5(3) 585–608.

[18] Barczy, M. and Kern, P. (2013). Representations of multidimensional linear process bridges. *Random Operators and Stochastic Equations* 21(2) 159–189.

[19] Barczy, M. and Kern, P. (2013). Sample path deviations of the Wiener and the Ornstein-Uhlenbeck process from its bridges. *Brazilian Journal of Probability and Statistics* 27(4) 437–466.

[20] Barczy, M. and Kern, P. (2015+). A link between Bougerol's identity and a formula due to Donati-Martin, Matsumoto and Yor. To appear in *Séminaire de Probabilités.* Available on the ArXiv: `http://arxiv.org/abs/1403.5785`

[21] Barczy, M., Kern, P. and Krause, V. (2015). Operator scaled Wiener bridges. *ESAIM: Probability and Statistics* 19 100–114.

[22] Barczy, M., Li, Z. and Pap, G. (2015). Stochastic differential equation with jumps for multi-type continuous state and continuous time branching processes with immigration. *ALEA. Latin American Journal of Probability and Mathematical Statistics* 12(1) 129–169.

[23] Barczy, M., Li, Z. and Pap, G. (2015+). Moment formulas for multi-type continuous state and continuous time branching processes with immigration. To appear in *Journal of Theoretical Probability.* DOI: 10.1007/s10959-015-0605-0

[24] Barczy, M., Körmendi, K. and Pap, G. (2015). Statistical inference for 2-type doubly symmetric critical irreducible continuous state and continuous time branching processes with immigration. *Journal of Multivariate Analysis* 139 92–123.

[25] Barczy, M. and Pap, G. (2005). Connection between deriving bridges and radial parts from multidimensional Ornstein-Uhlenbeck processes. *Periodica Mathematica Hungarica* 50(1-2) 47–60.

[26] Barczy, M. and Pap, G. (2010). Asymptotic behavior of maximum likelihood estimator for time inhomogeneous diffusion processes. *Journal of Statistical Planning and Inference* 140(6) 1576–1593.

[27] Barczy, M. and Pap, G. (2010). Alpha-Wiener bridges: singularity of induced measures and sample path properties. *Stochastic Analysis and Applications* 28(3) 447–466.

[28] Barczy, M. and Pap, G. (2011). Explicit formulas for Laplace transforms of certain functionals of some time inhomogeneous diffusions. *Journal of Mathematical Analysis and Applications* 380(2) 405–424.

[29] Barczy, M. and Pap, G. (2015+). Asymptotic behavior of critical irreducible multi-type continuous state and continuous time branching processes with immigration. To appear in *Stochastics and Dynamics.* DOI: 10.1142/S0219493716500088

[30] Barczy, M. and Pap, G. (2015+). Asymptotic properties of maximum likelihood estimators for Heston models based on continuous time observations. To appear in *Statistics.* DOI: 10.1080/02331888.2015.1044991

[31] Barczy, M., Pap, G. and Szabó, T. T. (2014). Parameter estimation for subcritical Heston models based on discrete time observations.
Available on the ArXiv: `http://arxiv.org/abs/1403.0527`

[32] Barndorff-Nielsen, O. and Shephard, N. (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society, Series B* 63 167–241.

[33] Baudoin, F. and Nguyen-Ngoc, L. (2004). The financial value of a weak information on a financial market. *Finance and Stochastics* 8 415–435.

[34] Ben Alaya, M. and Kebaier, M. (2012). Parameter estimation for the square root diffusions: ergodic and nonergodic cases. *Stochastic Models* 28(4) 609–634.

[35] Ben Alaya, M. and Kebaier, M. (2013). Asymptotic behavior of the maximum likelihood estimator for ergodic and nonergodic square-root diffusions. *Stochastic Analysis and Applications* 31(4) 552–573.

[36] Benke, J. and Pap, G. (2015). Local asymptotic quadraticity of statistical experiments connected with a Heston model. Available on the ArXiv: `http://arxiv.org/abs/1501.03664`

[37] Bichard, J. (2009). Some examples of the spatial evolution of two-parameter processes with non-adapted initial conditions, PhD thesis. *University of Warwick,* Warwick Research Archive Project.

[38] Bishwal, J. P. N. (2008). *Parameter Estimation in Stochastic Differential Equations.* Springer-Verlag Berlin Heidelberg.

[39] Borodin, A. N. and Salminen, P. (2002). *Handbook of Brownian Motion – Facts and Formulae, 2nd edition.* Birkhäuser.

[40] Bowman, F. (1958). *Introduction to Bessel Functions.* Dover, New York.

[41] Brennan, M. J. and Schwartz, E. S. (1990). Arbitrage in stock index futures. *The Journal of Business* 63(1) S7–S31.

[42] Carr, P. and Wu, L. (2003). The finite moment log stable process and option pricing. *The Journal of Finance* 58 753–777.

[43] Campi, L., Çetin, U. and Danilova, A. (2011). Dynamic Markov bridges motivated by models of insider trading. *Stochastic Processes and their Applications* 121 534–567.

[44] Chaumont, L. and Uribe Bravo, G. (2011). Markovian bridges: weak continuity and pathwise constructions. *The Annals of Probability* 39(2) 609–647.

[45] Chen, H. and Joslin, S. (2012). Generalized transform analysis of affine processes and applications in finance. *Review of Financial Studies* 25(7) 2225–2256.

[46] Chung, K. L. (1982). *Lectures from Markov Processes to Brownian Motion.* Springer-Verlag, New York–Berlin.

[47] Conti, R. (1976). *Linear Differential Equations and Control.* Academic Press, London.

[48] Corlay, S. and Pagès, G. (2015). Functional quantization-based stratified sampling methods. *Monte Carlo Methods and Applications* 21(1) 1–32.

[49] Cox, J. C., Ingersoll, J. E. and Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica* 53(2) 385–407.

[50] Csörgő, M. and Révész, P. (1981). *Strong Approximations in Probability and Statistics.* Academic Press, New York.

[51] Cuchiero, C., Filipović, D. and Teichmann, J. (2010). Affine Models. Contribution to *Encyclopedia of Quantitative Finance.* John Wiley & Sons, Ltd.

[52] Cuchiero, C., Keller-Ressel, M. and Teichmann, J. (2012). Polynomial processes and their applications to mathematical finance. *Finance and Stochastic* 16(4) 711–740.

[53] Dawson, D. A. and Li, Z. (2006). Skew convolution semigroups and affine processes. *The Annals of Probability* 34(3) 1103–1142.

[54] Delyon, B. and Hu, Y. (2006). Simulation of conditioned diffusion and application to parameter estimation. *Stochastic Processes and their Applications* 116 1660–1675.

[55] Deheuvels, P. (2006). Karhunen–Loève expansions of mean-centered Wiener processes. *High Dimensional Probability. IMS Lecture Notes–Monograph Series* 51 62–76.

[56] Deheuvels, P. (2007). A Karhunen–Loève expansion of a mean-centered Brownian bridge. *Statistics & Probability Letters* 77 1190–1200.

[57] Deheuvels, P. and Martynov, G. (2003). Karhunen–Loève expansions for weighted Wiener processes and Brownian bridges via Bessel functions. In: *Progress in Probability* vol. 55, pp. 57–93, Birkhäuser Verlag, Basel.

[58] Deheuvels, P., Peccati, G. and Yor, M. (2006). On quadratic functionals of the Brownian sheet and related processes. *Stochastic Processes and their Applications* 116(3) 493–538.

[59] Dietz, H. M. and Kutoyants, Yu. A. (1997). A class of minimum-distance estimators for diffusion processes with ergodic properties. *Statistics and Decisions* 15 211–227.

[60] Donati-Martin, C. (1990). Le probléme de Buffon-Synge pour une corde. *Advances in Applied Probability* 22 375–395.

[61] Doob, J. L. (1957). Conditional Brownian motion and the boundary limits of harmonic functions. *Bulletin de la Société Mathématique de France* 85 431–458.

[62] Dudley, R. M. (1989). *Real Analysis and Probability.* Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, California.

[63] Duffie, D., Filipović, D. and Schachermayer, W. (2003). Affine processes and applications in finance. *Annals of Applied Probability* 13 984–1053.

[64] Dzhaparidze, K. and Spreij, P. (1993). The strong law of large numbers for martingales with deterministic quadratic variation. *Stochastics and Stochastics Reports* 42 53–65.

[65] Es-Sebaiy, K. and Nourdin, I. (2013). Parameter estimation for alpha-fractional bridges. *Malliavin Calculus and Stochastic Analysis, A Festschrift in Honor of David Nualart, Series: Springer Proceedings in Mathematics & Statistics,* Vol. 34, Viens, F.; Feng, J.; Hu, Y.; Nualart, E. (Eds.), XII, 588 p.

[66] Ethier, S. N. and Kurtz, T. G. (1986). *Markov Processes: Characterization and Convergence.* Wiley.

[67] Feigin, P. D. (1979). Some comments concerning a curious singularity. *Journal of Applied Probability* 16(2) 440–444.

[68] Feller, W. (1951). Two Singular Diffusion Problems. *Annals of Mathematics* 54(1) 173–182.

[69] Filipović, D. and Mayerhofer, E. (2009). Affine diffusion processes: theory and applications. In Albrecher, Hansjörg (ed.) et al., Advanced Financial Modelling, *Radon Series on Computational and Applied Mathematics* 8 125–164, Berlin, Walter de Gruyter.

[70] Filipović, D., Mayerhofer, E. and Schneider, P. (2013). Density approximations for multivariate affine jump-diffusion processes. *Journal of Econometrics* 176(2) 93–111.

[71] Fitzsimmons, P., Pitman, J. and Yor, Y. (1992). Markovian bridges: construction, Palm interpretation, and splicing, in: E. Çinlar et al. (Eds.), Seminar on Stochastic Processes. *Progress in Probability* 33 Birkhäuser, Boston, 1992, pp. 101–134.

[72] Florens-Landais, D. and Pham, H. (1999). Large deviations in estimation of an Ornstein-Uhlenbeck model. *Journal of Applied Probability* 36(1) 60–70.

[73] Friz, P. K. and Keller-Ressel, M. (2010). Moment Explosions in Stochastic Volatility Models. Contribution to *Encyclopedia of Quantitative Finance.* John Wiley & Sons, Ltd.

[74] Fu, Z. and Li, Z. (2010). Stochastic equations of non-negative processes with jumps. *Stochastic Processes and their Applications* 120 306–330.

[75] Gao, F., Hannig, J., Lee, T-Y. and Torcaso, F. (2003). Laplace transforms via Hadamard factorization. *Electronic Journal of Probability* 8 no. 13, 20 pp.

[76] Gasbarra, D., Sottinen, T. and Valkeila, E. (2007). Gaussian bridges, in: F.E. Benth et al. (Eds.), Abel Symposia 2, Stochastic Analysis and Applications, *Proceedings of the Second*

Abel Symposium, Oslo, July 29 - August 4, 2005, held in honor of Kiyosi Itô, Springer, pp. 361–383.

[77] Gatheral, J. (2006). *The volatility surface : a practitioner's guide.* Jon Wiley & Sons, Inc., Hoboken, New Jersey.

[78] Glasserman, P. and Kim, K-K. (2010). Moment explosions and stationary distributions in affine diffusion models. *Mathematical Finance* 20(1) 1–33.

[79] Goldys, B. and Maslowski, M. (2008). The Ornstein Uhlenbeck bridge and applications to Markov semigroups. *Stochastic Processes and their Applications* 118 1738–1767.

[80] Gutiérrez, R. and Valderrama, M. J. (1987). On the Karhunen-Loève expansion for transformed processes. *Trabajos de Estadistica* 2 81–90.

[81] Hamilton, J. D. (1994). *Time Series Analysis.* Princeton University Press, Princeton.

[82] Hartman, P. (2002). *Ordinary Differential Equations,* 2nd ed. Society for Industrial and Applied Mathematics, Philadelphia.

[83] Hwang, C. R. (1980). Gaussian measure of large balls in a Hilbert space. *Proceedings of the American Mathematical Society* 78(1) 107–110.

[84] Heston, S. (1993). A closed-form solution for options with stochastic volatilities with applications to bond and currency options. *The Review of Financial Studies* 6 327–343.

[85] Horn, R. A. and Johnson, Ch. R. (1985). *Matrix Analysis.* Cambridge University Press, Cambridge.

[86] Hoyle, E. (2010). Information-based models for finance and insurance, PhD thesis. *Imperial College London.* Available on the ArXiv: `http://arxiv.org/abs/1010.0829`.

[87] Hoyle, E., Hughston, L. P. and Macrina, A. (2011). Lévy random bridges and the modelling of financial information. *Stochastic Proceses and their Applications* 121 856–884.

[88] Hu, Y. and Long, H. (2007). Parameter estimation for Ornstein–Uhlenbeck processes driven by α-stable Lévy motions. *Communications on Stochastic Analysis* 1(2) 175–192.

[89] Hu, Y. and Long, H. (2009). Least squares estimator for Ornstein–Uhlenbeck processes driven by α-stable motions. *Stochastic Processes and their Applications* 119(8) 2465–2480.

[90] Hu, Y. and Long, H. (2009). On the singularity of least squares estimator for mean-reverting α-stable motions. *Acta Mathematica Scientia* 29B(3) 599–608.

[91] Huang, J., Ma, C. and Zhu, C. (2011). Estimation for discretely observed continuous state branching processes with immigration. *Statistics and Probability Letters* 81 1104–1111.

[92] Hurd, T. R. and Kuznetsov, A. (2008). Explicit formulas for Laplace transforms of stochastic integrals. *Markov Processes and Related Fields* 14(2) 277–290.

[93] Hurn, A. S., Lindsay, K. A. and McClelland, A. J. (2013). A quasi-maximum likelihood method for estimating the parameters of multivariate diffusions. *Journal of Econometrics* 172 106–126.

[94] Ikeda, N. and Watanabe, S. (1981). *Stochastic Differential Equations and Diffusion Processes.* North-Holland Publishing Company.

[95] Jacod, J. and Shiryaev, A. N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. Springer-Verlag, Berlin.

[96] Jeanblanc, M., Yor, M. and Chesney, M. (2009). *Mathematical Methods for Financial Markets.* Springer-Verlag London Limited.

[97] Jena, R. P., Kim, K.-K. and Xing, H. (2012). Long-term and blow-up behaviors of exponential moments in multi-dimensional affine diffusions. *Stochastic Processes and their Applications* 122(8) 2961–2993.

[98] Kallenberg, O. (1992). Some time change representations of stable integrals, via predictable transformations of local martingales. *Stochastic Processes and their Applications* 40 199–223.

[99] Kalman, R. E. (1960). Contributions to the theory of optimal control. *Boletín de la Sociedad Matemática Mexicana* 5 102–119.

[100] Karatzas, I. and Shreve, S. E. (1991). *Brownian Motion and Stochastic Calculus, 2nd edition.* Springer-Verlag.

[101] Karlin, S. and Taylor, M. (1981). *A Second Course in Stochastic Processes.* Academic Press, New York.

[102] Kátai, I. and Mogyoródi, J. (1967). Some remarks concerning the stable sequences of random variables. *Publicationes Mathematicae Debrecen* 14 227–238.

[103] Keller-Ressel, M. and Mijatović, A. (2012). On the limit distributions of continuous-state branching processes with immigration. *Stochastic Processes and their Applications* 122 2329–2345.

[104] Keller-Ressel, M., Schachermayer, W. and Teichmann, J. (2011). Affine processes are regular. *Probabaility Theory and Related Fields* 151 591–611.

[105] Keller-Ressel, M., Schachermayer, W. and Teichmann, J. (2013). Regularity of affine processes on general state spaces. *Electronic Journal of Probability* 18 Article 43 1–17.

[106] Kleptsyna, M. L. and Le Breton, A. (2002). Statistical analysis of the fractional Ornstein–Uhlenbeck type process. *Statistical Inference for Stochastic Processes* 5(3) 229–248.

[107] Kleptsyna, M. L. and Le Breton, A. (2002). A Cameron-Martin type formula for general Gaussian processes – a filtering approach. *Stochastics and Stochastics Reports* 72(3-4) 229–250.

[108] Klimko, L. A. and Nelson, P. I. (1978). On conditional least squares estimation for stochastic processes. *The Annals of Statistics* 6(3) 629–642.

[109] Korenev, B. G. (2002). *Bessel Functions and their Applications.* Taylor & Francis, London.

[110] Koval', V. A. (2001). On the strong law of large numbers for multivariate martingales with continuous time. *Ukrainian Mathematical Journal* 53(9) 1554–1560.

[111] Kuo, H.-H. (2006). *Introduction to Stochastic Integration.* Universitext, Springer, New York.

[112] Kutoyants, Yu A. (2004). *Statistical Inference for Ergodic Diffusion Processes.* Springer-Verlag London Limited.

[113] Küchler, U. and Sørensen, M. (1997). *Exponential families of stochastic processes,* Springer-Verlag, New York.

[114] Li, W. V. and Ma, C. (1992). Comparison results for the lower tail of Gaussian seminorms. *Journal of Theoretical Probability* 5 1–31.

[115] Li, Z. (2011). *Measure-Valued Branching Markov Processes.* Springer-Verlag, Heidelberg.

[116] Li, Z. and Ma, C. (2013). Asymptotic properties of estimators in a stable Cox-Ingersoll-Ross model. *Stochastic Processes and their Applications* 125(8) 3196–3233.

[117] Liptser, R. S. and Shiryaev, A. N. (2001). *Statistics of Random Processes I. Applications, 2nd edition.* Springer-Verlag, Berlin, Heidelberg.

[118] Liptser, R. S. and Shiryaev, A. N. (2001). *Statistics of Random Processes II. Applications, 2nd edition.* Springer-Verlag, Berlin, Heidelberg.

[119] Liu, N. and Ulukus, S. (2006). Optimal Distortion-Power Tradeoffs in Sensor Networks: Gauss-Markov Random Processes. *2006. ICC '06. IEEE International Conference on Communications* 4 1543–1548.

[120] Luschgy, H. (1992). Local asymptotic mixed normality for semimartingale experiments. *Probability Theory and Related Fields* 92(2) 151–176.

[121] Luschgy, H. and Pagès, G. (2009). Expansions for Gaussian processes and Parseval frames. *Electronic Journal of Probability* 14 1198–1221.

[122] Lyons, T. J. and Zheng, W. A. (1990). On conditional diffusions. *Proceedings of the Royal Society of Edinburgh. Section A. Mathematics* 115A 243–255.

[123] Ma, C. (2009). A limit theorem of two-type Galton–Watson branching processes with immigration. *Statistics and Probability Letters* 79 1710–1716.

[124] Mann, H. B. and Wald, A. (1943). On the statistical treatment of linear stochastic difference equations. *Econometrica* 11(3/4) 173–220.

[125] Mansuy, R. (2004). On a one-parameter generalization of the Brownian bridge and associated quadratic functionals. *Journal of Theoretical Probability* 17(4) 1021–1029.

[126] Meerschaert, M. M. and Scheffler, H. P. (2001). *Limit Distributions for Sums of Independent Random Vectors.* Wiley, New York.

[127] Meyn, Sean P. and Tweedie, R. L. (1992). Stability of Markovian processes. I: Criteria for Discrete-Time Chains. *Advances in Applied Probability* 24(3) 542–574.

[128] Meyn, Sean P. and Tweedie, R. L. (1993). Stability of Markovian processes. II: Continuous-time processes and sampled chains. *Advances in Applied Probability* 25(3) 487–517.

[129] Meyn, Sean P. and Tweedie, R. L. (1993). Stability of Markovian processes. III: Foster-Lyapunov criteria for continuous-time processes. *Advances in Applied Probability* 25(3) 518–548.

[130] Meyn, Sean P. and Tweedie, R. L. (2009). *Markov Chain and Stochastic Stability, 2nd edition.* Cambridge University Press, Cambridge.

[131] Nazarov, A. I. (2003). On the sharp constant in the small ball asymptotics of some Gaussian processes under L_2-norm. *Journal of Mathematical Sciences* 117(3) 4185–4210.

[132] Nazarov, A. I. and Nikitin, Ya. Yu. (2004). Exact L_2-small ball behavior of integrated Gaussian processes and spectral asymptotics of boundary value problems. *Probability Theory and Related Fields* 129 469–494.

[133] Nazarov, A. I. and Pusev, R. S. (2009). Exact small ball asymptotics in weighted L_2-norm for some Gaussian processes. *Journal of Mathematical Sciences* 163(4) 409–429.

[134] Øksendal, B. (2003). *Stochastic Differential Equations, 6th edition.* Springer, Berlin, Heidelberg, New York.

[135] Overbeck, L. and Rydén, T. (1997). Estimation in the Cox-Ingersoll-Ross model. *Econometric Theory* 13(3) 430–461.

[136] Overbeck, L. (1998). Estimation for continuous branching processes. *Scandinavian Journal of Statistics* 25(1) 111–126.

[137] Papież, L. S. and Sandison, G. A. (1990). A diffusion model with loss of particles. *Advances in Applied Probability* 22 533–547.

[138] Papoulis, A. (1991). *Probability, Random Variables and Stochastic Processes* McGraw-Hill Inc., New York.

[139] Phillips, P. C..B. and Yu, J. (2009). Maximum likelihood and Gaussian estimation of continuous time models in finance. In: Andersen, Torben G., Davis, Richard A., Kreiß, Jens-Peter, Mikosch, Thomas (eds.), Handbook of financial time series, pp. 497–530, Berlin, Springer.

[140] Privault, N. and Zambrini, J-C. (2004). Markovian bridges and reversible diffusion processes with jumps. *Annales de l'Institut Henri Poincaré Probabilités et Statistiques* 40 599–633.

[141] Qian, Z. and Zheng, W. (2004). A representation formula for transition probability densities of diffusions and applications. *Stochastic Processes and their Applications* 111 57–76.

[142] Prakasa Rao, B. L. S. (1999). *Semimartingales and their Statistical Inference, 3rd ed.* Chapman & Hall/CRC.

[143] Revuz, D. and Yor, M. (1999). *Continuous martingales and Brownian motion*, 3rd ed. Springer-Verlag Berlin Heidelberg.

[144] Rosinski, J. and Woyczynski, W. A. (1986). On Itô stochastic integration with respect to p-stable motion: inner clock, integrability of sample paths, double and multiple integrals. *The Annals of Probability* 14(1) 271–286.

[145] Sandrić, N. (2013). Long-time behaviour of stable-like processes. *Stochastic Processes and their Applications* 123(4) 1276–1300.

[146] Smirnov, N. V. (1937). On the distribution of Mises' ω^2-test. *Matematicheskii Sbornik* 2 973–994. (In Russian.)

[147] Sondermann, D., Trede, M. and Wilfling, B. (2011). Estimating the degree of interventionist policies in the run-up to EMU. *Applied Economics* 43(2) 207–218.

[148] Sørensen, H. (2004). Parametric inference for diffusion processes observed at discrete points in time: a survey. *International Statistical Review* 72(3) 337–354.

[149] Trede, M. and Wilfling, B. (2007). Estimating exchange rate dynamics with diffusion processes: an application to Greek EMU data. *Empirical Economics* 33(1) 23–39.

[150] van der Vaart, A. W. (1998). *Asymptotic Statistics.* Cambridge University Press, Cambridge.

[151] Varughese, M. M. (2013). Parameter estimation for multivariate diffusion systems. *Computational Statistics and Data Analysis* 57 417–428.

[152] Watson, G. N. (1944). *A Treatise on the Theory of Bessel Functions, 2nd edition.* Cambridge Univ. Press, England.

[153] Wei, C. Z. and Winnicki, J. (1989). Some asymptotic properties of branching processes with immigration. *Stochastic Processes and their Applications* 31 261–282.

[154] Wei, C. Z. and Winnicki, J. (1990). Estimation of the means in the branching processes with immigration. *The Annals of Mathematical Statistics* 18(4) 1757–1773.

[155] White, J. S. (1958). The limiting distribution of the serial correlation coefficient in the explosive case. *The Annals of Mathematical Statistics* 29 1188–1197.

[156] Yor, M. (1997). *Some Aspects of Brownian Motion, Part II: Some Recent Martingale Problems.* Lectures in Mathematics, ETH Zürich, Birkhäuser, Basel.

[157] Yor, M. (2001). *Exponential Functionals of Brownian Motion and Related Processes.* Springer-Verlag, Berlin.

[158] van Zanten, H. (2000). A multivariate central limit theorem for continuous local martingales. *Statistics & Probability Letters* 50(3) 229–235.

[159] Zolotarev, V. M. (1961). Concerning a certain probability problem. *Theory of Probability and its Applications* 6 201–204.