Parameter estimation for non-stationary Fisher-Snedecor diffusion

A.M. Kulik[a], N.N. Leonenko[b,*], I. Papić[c], N. Šuvak[c]

[a] *Wroclaw University of Science and Technology, Faculty of Pure and Applied Mathematics, Wybrzeże Wyspiańskiego Str. 27, 50-370 Wroclaw, Poland and Institute of Mathematics, Ukrainian National Academy of Sciences, 01601 Tereshchenkivska str. 3, Kyiv, Ukraine*
[b] *School of Mathematics, Cardiff University, Senghennydd Road, Cardiff CF244AG, UK*
[c] *Department of Mathematics, J.J. Strossmayer University of Osijek, Trg Ljudevita Gaja 6, HR-31000 Osijek, Croatia*

Abstract

The problem of parameter estimation for the non-stationary ergodic diffusion with Fisher-Snedecor invariant distribution, to be called Fisher-Snedecor diffusion, is considered. We propose generalized method of moments (GMM) estimator of unknown parameter, based on continuous-time observations, and prove its consistency and asymptotic normality. The explicit form of the asymptotic covariance matrix in asymptotic normality framework is calculated according to the new iterative technique based on evolutionary equations for the point-wise covariations. The results are illustrated in a simulation study covering various starting distributions and parameter values.

Keywords: Fisher-Snedecor diffusion, generalized method of moments (GMM), P-consistency, asymptotic normality, iterative technique for the calculation of the asymptotic covariance matrix
2000 MSC: 33C05, 33C47, 35P10, 60G10, 60J60, 62M05, 62M15

1. Introduction

Fisher-Snedecor diffusion (FSD) $X = (X_t,\, t \geq 0)$ is defined as the solution of the non-linear stochastic differential equation (SDE)

$$dX_t = -\theta\,(X_t - \kappa)\; dt + \sqrt{2\theta X_t \left(\frac{X_t}{\beta/2-1} + \frac{\kappa}{\alpha/2}\right)}\, dW_t, \quad t \geq 0 \tag{1.1}$$

with values in $(0,\infty)$, where $(W_t,\, t \geq 0)$ is the standard Brownian motion. From the SDE (1.1) we see that infinitesimal parameter of FSD are polynomials - the drift coefficient is linear and the squared-diffusion coefficient is quadratic, depending on parameters ensuring that its leading coefficient and its discriminant are positive. According to general results from Genon-Catalot et al. (2000), for $\alpha, \beta > 2$ the diffusion X satisfying (1.1) is ergodic. Under the particular choice $\kappa = \beta/(\beta-2)$, the unique invariant distribution is the well-known Fisher-Snedecor distribution $\mathcal{FS}(\alpha,\beta)$ with shape parameters (degrees of freedom) α and β and the PDF given by (2.7).

FSD belongs to the class of diffusion processes with invariant distributions from the Pearson family of continuous distributions, introduced by K. Pearson (1914). The study of such processes started in the 1930's by A. Kolmogorov (see Shiryayev (1992)) and therefore processes from this family are often called Kolmogorov-Pearson (KP) diffusions. For more detailed discussion on KP diffusions we refer to papers Forman and Sørensen (2008), Shaw and Munir (2009) and Avram et al. (2013a). Together with the reciprocal gamma and the Student diffusions, FSD forms the class of the so-called heavy-tailed KP diffusions. Statistical inference for heavy-tailed KP diffusions, relying on the GMM estimation and spectral representation of the diffusion transition density (eigendifferential expansion in Itô and McKean (1974)), is developed in recent papers Leonenko and Šuvak (2010a), Leonenko and Šuvak (2010b) and Avram et al. (2011), where the stationary version of the respective diffusions are observed. Statistical inference of non-stationary FSD was motivated mostly by theoretical reasons, to complete the recent theoretical results developed for stationary case in Avram et al. (2011), Avram et al. (2012) and Avram et al. (2013b).

*Corresponding author
Email addresses: kulik.alex.m@gmail.com (A.M. Kulik), LeonenkoN@cardiff.ac.uk (N.N. Leonenko), ipapic@mathos.hr (I. Papić), nsuvak@mathos.hr (N. Šuvak)

The theory of parametric inference for both stationary and non-stationary diffusions is a well studied field. Some of the classical references from this area are e.g. Kutoyants (2004), Bishwal (2007) and Kessler et al. (2012). The approach presented in Sørensen (2012), based on optimal martingale estimating functions, is applicable in estimation problems regarding stationary KP diffusions - some elements of this theory are already applied in Forman and Sørensen (2008). Furthermore, martingale estimating functions methodology for non-stationary diffusions is developed and thoroughly studied from theoretical and practical point of view in e.g. Kessler (1997), Kessler and Sørensen (1999) and Kessler (2000). The theory of simulation approaches are well covered by a classical book Kloeden and Platen (2011), more practical book Iacus (2008) and e.g. paper Kessler and Paredes (2002) devoted to computational properties of martingale estimating functions. However, in this paper we focus on GMM estimators. Beside their simplicity and appealing statistical properties such as consistency and asymptotic normality (with explicitly known asymptotic covariance matrix), additional reason why we focus on GMM estimators is that in the analysis of their asymptotic properties the continuous part of the spectrum of the corresponding diffusion infinitesimal generator is not neglected. The concise overview of spectral properties and spectral representation of transition density of stationary FSD, based on results from Avram et al. (2013b) and with the obvious impact of the continuous part of the spectrum of the corresponding infinitesimal generator, is given in Appendix A. We find this approach quite important since the impact of the continuous part of the spectrum and its negligence in most of the statistical problems regarding diffusions having such a structure of the spectrum still isn't properly investigated (by the best knowledge of the authors).

In this paper we consider the problem of parameter estimation of FSD X in the non-stationary setting, i.e. with the arbitrary distribution of the initial value X_0. Results of analysis of asymptotic properties of parameter estimators are obtained by application of the law of large numbers and the central limit theorem for additive functionals for the FSD from Kulik and Leonenko (2013) and relying on the study of explicit quantitative rates for the convergence rate of respective finite-dimensional distributions to that of the stationary FSD and for the β-mixing coefficient. The techniques in Kulik and Leonenko (2013) rely on the general theory developed for (possibly non-symmetric and non-stationary) Markov processes, with the significant novelty based on the so-called Lyapunov-type condition. The natural idea behind this approach is the extension of results in the stationary setting to the non-stationary setting using the bounds for the deviation between the stationary and non-stationary versions of the FSD. The technique is based on the notion of an (exponential) ϕ-coupling, introduced in Kulik (2011) as a tool for studying convergence rates of L_p-semigroups generated by a Markov process. Similar results for the reciprocal gamma and the Student diffusions were considered in Abourashchi and Veretennikov (2009) and Abourashchi and Veretennikov (2010), respectively.

This paper is a natural extension of results from Kulik and Leonenko (2013), where it is proved that the empirical moments of the FSD X are P-consistent, asymptotically normal and, under some additional conditions on the initial distribution of X, asymptotically unbiased. Relying on these results, it has been shown that the GMM estimator of parameter $(\alpha, \beta, \kappa, \theta)$ of FSD, given either the discrete-time or the continuous-time observations, are also P-consistent and asymptotically normal. We would like to emphasize that in the GMM estimation moments of negative order have also been used, yielding a simpler estimator than the one based on moments of positive order only. In this paper we develop an iterative procedure for calculation of the asymptotic covariances in the asymptotic normality framework, where the only drawback is experienced in the calculation of some asymptotic covariances related to discrete-time empirical moments of negative order. However, the simulation study in Section 4 revealed the deficiency of the proposed GMM estimators when diffusion parameters are estimated from observations from short time-intervals, in which the diffusion still didn't reach its asymptotic/stationary regime or is not in this regime long enough. The detected problem is illustrated and discussed in Remark 4.1.

The paper is organized as follows. After Introduction, in Section 2 we give some important preliminaries on FSD and references to certain results that are crucial for developing this paper. In Section 3 we focus to parameter estimation of FSD in non-stationary setting, based on continuous observations in $[0, T]$. In Subsections 3.1 and 3.2 we formulate results and give examples that are crucial for proving P-consistency and asymptotic normality of GMM estimator of parameters of non-stationary FSD as $T \to \infty$. Main result on asymptotic properties of the GMM estimator of parameters of FSD is formulated in Subsection 3.3 and proved in 5.5. All the proofs are postponed to the separate Section 5, while Section 4 contains results of the simulation study based on discrete observations. Although the focus of the paper is on the continuous observations, in order to provide the simulation study we established the connection between estimators and their asymptotic properties based on continuous and discrete observations and clarified it in Remark 5.1. Some results from previous researches of FSD are, for completeness of the exposition, given in Appendix A and Appendix B.

2. Non-stationary Fisher-Snedecor diffusion

From the governing SDE (1.1) of the FSD $X = (X_t,\, t \geq 0)$ we see that its infinitesimal parameters, i.e. the drift coefficient $a(x)$ and the diffusion coefficient $\sigma(x)$, are respectively given by

$$a(x) = -\theta\,(x-\kappa)\,, \qquad \sigma(x) = \sqrt{2\theta x\left(\frac{x}{\beta/2-1}+\frac{\kappa}{\alpha/2}\right)}. \tag{2.1}$$

Here we assume that

$$\theta > 0, \quad \kappa > 0, \quad \beta > 4, \quad \alpha > 2, \tag{2.2}$$

where the restrictions imposed on values of parameters α and β ensure ergodicity of the diffusion and existence of the second moment of its invariant distribution, respectively (see e.g. Forman and Sørensen (2008); Genon-Catalot et al. (2000)). If the invariant distribution has finite variance, the autocorrelation function is given by

$$\rho(t) = \mathrm{Corr}(X_{s+t}, X_s) = e^{-\theta t}, \quad t \geq 0, \quad s \geq 0 \tag{2.3}$$

(see Bibby et al. Bibby et al. (2005), Theorem 2.3.(iii)). Therefore, parameter $\theta > 0$ is usually called the autocorrelation parameter.

For $x \in (0,\infty)$,

$$\mathfrak{s}(x) = \exp\left(-\int_1^x \frac{2a(u)}{\sigma^2(u)}\,du\right) = Cx^{-\alpha/2}\left(x+\frac{\kappa(\beta-2)}{\alpha}\right)^{\alpha/2+\beta/2-1} \tag{2.4}$$

defines the scale density of FSD, where C is a constant which could be expressed explicitly. From assumptions on parameter values (2.2) it follows that

$$\int_x^\infty \mathfrak{s}(y)\,dy = \infty, \qquad \int_0^x \mathfrak{s}(y)\,dy = \infty, \quad x \in (0,\infty).$$

Therefore, for $\alpha > 2$ both 0 and ∞ are unattainable points of the state space for the diffusion X, i.e. the random time moment $T_{0,\infty}$ is a.s. infinite for any positive initial condition X_0, see e.g. Karlin and Taylor (1981), Chapter 18.6. This means that (1.1) uniquely determines a time-homogeneous strong Markov process X with the state space $\mathbb{X} = (0,\infty)$, considered here as the locally compact metric space with the metric $d(x,y) = |x-y| + |x^{-1}-y^{-1}|$.

Generally, for any $\kappa > 0$ the unique invariant distribution is given by the PDF

$$\mathfrak{p}(x) = \frac{1}{xB(\alpha/2,\beta/2)}\left(\frac{\alpha x}{\alpha x+\varrho}\right)^{\alpha/2}\left(\frac{\varrho}{\alpha x+\varrho}\right)^{\beta/2}\mathrm{I}_{(0,\infty)}(x), \tag{2.5}$$

with $\varrho = (\beta-2)\kappa$ and with the moments of order $\upsilon \in \left[-\frac{\alpha}{2}, \frac{\beta}{2}\right)$ of the following form:

$$m_\upsilon = \int_0^\infty x^\upsilon \mathfrak{p}(x)\,dx = \left(\frac{\varrho}{\alpha}\right)^\upsilon \frac{\Gamma(\alpha/2+\upsilon)\Gamma(\beta/2-\upsilon)}{\Gamma(\alpha/2)\Gamma(\beta/2)}. \tag{2.6}$$

If $\kappa = \beta/(\beta-2)$, i.e. $\varrho = \beta$, the unique invariant distribution is the classical Fisher-Snedecor distribution $\mathcal{FS}(\alpha,\beta)$ with degrees of freedom α and β and with the PDF

$$\mathfrak{fs}(x) = \frac{1}{xB(\alpha/2,\beta/2)}\left(\frac{\alpha x}{\alpha x+\beta}\right)^{\alpha/2}\left(\frac{\beta}{\alpha x+\beta}\right)^{\beta/2}\mathrm{I}_{(0,\infty)}(x). \tag{2.7}$$

The relation between the PDFs (2.5) and (2.7) is

$$\mathfrak{p}(x) = \frac{\beta}{\varrho}\,\mathfrak{fs}\left(\frac{\beta}{\varrho}x\right), \tag{2.8}$$

clearly stating that the invariant PDF for X is the image of $\mathcal{FS}(\alpha,\beta)$ under the linear transformation $x \mapsto (\varrho/\beta)x$. It is not hard to see that the process $(\beta/\varrho)X$ satisfies SDE (1.1) with $\kappa = \beta/(\beta-2)$, and therefore the general FSD defined by (1.1) could be represented as

$$X = \frac{\varrho}{\beta}U, \tag{2.9}$$

where the so-called canonical FSD U satisfies the SDE

$$dU_t = -\theta\left(U_t - \frac{\beta}{\beta-2}\right)dt + \sqrt{\frac{4\theta U_t}{\beta-2}\left(U_t + \frac{\beta}{\alpha}\right)}\,dW_t, \quad t \geq 0. \tag{2.10}$$

In this paper we focus on estimation of parameter $(\alpha, \beta, \kappa, \theta)$ of the FSD (1.1). Estimation and the analysis of asymptotic properties of GMM estimator in stationary setting ($X_0 \sim \mathcal{FS}(\alpha, \beta)$) are performed in Avram et al. (2011). The results of this paper rely on the law of large numbers (LLN) and the central limit theorem (CLT) for α-mixing sequences. The analogous of these theorems, as well as an extensive and detailed study of other properties of the FSD in non-stationary setting are given in the recent paper Kulik and Leonenko (2013). The proofs of the LLN and the CLT for additive functionals of the non-stationary FSD in both discrete-time and continuous-time setting in Kulik and Leonenko (2013) are based on coupling, ergodicity and mixing properties. For completeness, we give a short overview of these results in Appendix B.

3. Parameter estimation of the non-stationary FSD

In this section we apply Theorems from Appendix B.2 to verify that GMM estimator of parameter $(\alpha, \beta, \kappa, \theta)$ of the FSD (1.1) is P-consistent and asymptotically normal. In Subsections 3.1 and 3.2 we derive techniques for calculating the point-wise covariations of the estimator of the specific empirical moments, further used for GMM estimation of parameter $(\alpha, \beta, \kappa, \theta)$ in Subsection 3.3.

Let us introduce the notation

$$X^{st} = (X_t^{st},\, t \geq 0)$$

for the stationary version of the FSD ($X_0^{st} \sim \mathcal{FS}\,(\alpha, \beta)$) and its moments and mixed moments:

$$m_\upsilon = E(X_s^{st})^\upsilon, \qquad m_{\upsilon,\chi}(t) = E(X_s^{st})^\upsilon (X_{s+t}^{st})^\chi.$$

Next we introduce the notation for empirical moments based on continuous-time observations ($X_t,\, t \in [0, T]$) and discrete-time observations $(X_1, \ldots, X_n)$ from the non-stationary FSD $X = (X_t,\, t \geq 0)$, supposing certain restrictions on parameter values in order to ensure existence of the corresponding theoretical moments in particular situations:

$$\overline{m}_{\upsilon,\chi,c}(t) = \frac{1}{T}\int_0^T (X_s)^\upsilon (X_{t+s})^\chi\,ds, \qquad \overline{m}_{\upsilon,\chi,d}(t) = \frac{1}{n}\sum_{s=1}^{n} (X_s)^\upsilon (X_{t+s})^\chi, \quad t > 0. \tag{3.1}$$

Usual empirical moments

$$\overline{m}_{\upsilon,c} = \frac{1}{T}\int_0^T (X_s)^\upsilon\,ds, \quad \overline{m}_{\upsilon,d} = \frac{1}{n}\sum_{s=1}^{n}(X_s)^\upsilon \tag{3.2}$$

are equal to the empirical mixed moments with $\chi = 0$, and empirical covariances

$$\overline{R}_c(t) = \frac{1}{T}\int_0^T X_s X_{t+s}\,ds - \left(\frac{1}{T}\int_0^T X_s\,ds\right)^2, \quad \overline{R}_d(t) = \frac{1}{n}\sum_{s=1}^{n} X_s X_{t+s} - \left(\frac{1}{n}\sum_{s=1}^{n} X_s\right)^2, \tag{3.3}$$

can be written as

$$\overline{R}_c(t) = \overline{m}_{1,1,c}(t) - \left(\overline{m}_{1,c}\right)^2, \quad \overline{R}_d(t) = \overline{m}_{1,1,d}(t) - \left(\overline{m}_{1,d}\right)^2. \tag{3.4}$$

Since non-stationary FSD after some time reaches its steady state, it is well known that for any initial distribution empirical moments given in (3.1) are estimators of theoretical moments m_υ and $m_{\upsilon,\chi}(t)$, for each fixed $t > 0$ in the latter case. This presents the basis for the GMM estimation of parameter $(\alpha, \beta, \kappa, \theta)$ of the non-stationary FSD.

In Subsection 3.1 we introduce evolutionary equations approach, a useful tool for calculating asymptotic covariances of moment estimators (3.2) and (3.4), while the technically sound Subsection 3.2 gives the technique for calculating the specific covariances when the evolutionary approach is not applicable.

3.1. Evolutionary equations for the point-wise covariations

In the following list of propositions and examples we use the following basic notation:

$$f_\upsilon(x) := x^\upsilon, \qquad E_{\upsilon,g}(t) := E f_\upsilon(X_t^{st}) g(X_0^{st}), \qquad \langle f \rangle := \int_{(0,\infty)} f(x)\, \mathfrak{p}(x) dx.$$

Proposition 3.1. *For all $\upsilon \in \mathbb{N}$ and non-negative functions g satisfiying*

$$\langle g \rangle < +\infty, \quad \langle f_\upsilon g \rangle < +\infty, \tag{3.5}$$

$$C_{\upsilon,g}(t) := \operatorname{Cov}\Big(f_\upsilon\big(X_t^{st}\big), g\big(X_0^{st}\big)\Big) \tag{3.6}$$

satisfies identity

$$C_{\upsilon,g}(t) = e^{-\frac{\upsilon\theta(\beta-2\upsilon)}{\beta-2}t} C_{\upsilon,g}(0) + \frac{\upsilon\theta\varrho(\alpha+2\upsilon-2)}{\alpha(\beta-2)} \int_0^t e^{\frac{\upsilon\theta(\beta-2\upsilon)}{\beta-2}(s-t)} C_{\upsilon-1,g}(s)\, ds. \tag{3.7}$$

In particular, note that $C_{0,g}(t) = 0$, for all $t \geq 0$.

Example 3.1.

(a) According to the Proposition 3.1, for $\upsilon \in \mathbb{N}$ and non-negative function g satisfying (3.5), point-wise covariances (3.6) can be calculated iteratively from (3.7). In particular,

$$C_{1,g}(t) = e^{-\theta t} C_{1,g}(0), \tag{3.8}$$

$$C_{2,g}(t) = e^{-2\theta\frac{\beta-4}{\beta-2}t} C_{2,g}(0) + \frac{2(\alpha+2)\varrho}{\alpha(\beta-6)} \left(e^{-\theta t} - e^{-2\theta\frac{\beta-4}{\beta-2}t}\right) C_{1,g}(0). \tag{3.9}$$

Note that the condition for (3.9) to hold true is $\alpha > 2, \beta > 4$ and (3.5), since

$$e^{-\theta t} - e^{-2\theta\frac{\beta-4}{\beta-2}t} = e^{-\theta t}\left[1 - e^{-\theta\frac{\beta-6}{\beta-2}t}\right],$$

compensates the term $(\beta - 6)$ in denominator.

(b) From (2.6) we obtain

$$\langle f_{-1} \rangle = \frac{\alpha\beta}{(\alpha-2)\varrho}, \quad \langle f_1 \rangle = \frac{\varrho}{\beta-2}, \quad \langle f_2 \rangle = \frac{(\alpha+2)\varrho^2}{\alpha(\beta-2)(\beta-4)}.$$

Since

$$\langle f_{\upsilon-1} \rangle = \frac{(\beta-2\upsilon)\alpha}{(\alpha+2\upsilon-2)\varrho} \langle f_\upsilon \rangle,$$

the above identities, after introducing the notation $C_{\upsilon,\chi}(t)$ for $C_{\upsilon,f_\chi}(t)$, can be written in the following form:

$$C_{1,\chi}(t) = \left[\langle f_{\chi+1} \rangle - \frac{\varrho}{\beta-2}\langle f_\chi \rangle\right] e^{-\theta t} = \frac{2\chi(\alpha+\beta-2)\varrho}{\alpha(\beta-2)(\beta-2\chi-2)} \langle f_\chi \rangle e^{-\theta t},$$

$$\begin{aligned}
C_{2,\chi}(t) = &\left(\left[\langle f_{\chi+2} \rangle - \frac{(\alpha+2)\varrho^2}{\alpha(\beta-2)(\beta-4)}\langle f_\chi \rangle\right] - \frac{2(\alpha+2)\varrho}{\alpha(\beta-6)}\left[\langle f_{\chi+1} \rangle - \frac{\varrho}{\beta-2}\langle f_\chi \rangle\right]\right) e^{-2\theta\frac{\beta-4}{\beta-2}t} \\
&+ \frac{2(\alpha+2)\varrho}{\alpha(\beta-6)}\left[\langle f_{\chi+1} \rangle - \frac{\varrho}{\beta-2}\langle f_\chi \rangle\right] e^{-\theta t} \\
= &\left(\langle f_{\chi+2} \rangle - \frac{2(\alpha+2)\varrho}{\alpha(\beta-6)}\langle f_{\chi+1} \rangle + \frac{(\alpha+2)\varrho^2}{\alpha(\beta-4)(\beta-6)}\langle f_\chi \rangle\right) e^{-2\theta\frac{\beta-4}{\beta-2}t} \\
&+ \frac{4\chi(\alpha+2)(\alpha+\beta-2)\varrho^2}{\alpha^2(\beta-2)(\beta-6)(\beta-2\chi-2)} \langle f_\chi \rangle e^{-\theta t}.
\end{aligned}$$

(c) This part of the example contains the list of the covariances $C_{\upsilon,\chi}(t)$ that can be calculated according to the formulae from Example 3.1 (a) and (b). The items in the list are numerated as $I_{\upsilon,\chi}$. Each item is provided by the conditions on parameters α and β, required for the respective identity to hold true.

$I_{1,1}$ $(\alpha > 2, \beta > 4)$

$$C_{1,1}(t) = \frac{2(\alpha+\beta-2)\varrho^2}{\alpha(\beta-2)^2(\beta-4)} e^{-\theta t}.$$

$I_{1,2}$ $(\alpha > 2, \beta > 6)$

$$C_{1,2}(t) = \frac{4(\alpha+2)(\alpha+\beta-2)\varrho^3}{\alpha^2(\beta-2)^2(\beta-4)(\beta-6)} e^{-\theta t}.$$

$I_{1,-1}$ $(\alpha > 2, \beta > 2)$

$$C_{1,-1}(t) = -\frac{2(\alpha+\beta-2)}{(\beta-2)(\alpha-2)} e^{-\theta t}.$$

$I_{2,2}$ $(\alpha > 2, \beta > 8)$

$$\begin{aligned} C_{2,2}(t) = & \left[\frac{(\alpha+6)(\alpha+4)(\alpha+2)\varrho^4}{\alpha^3(\beta-2)(\beta-4)(\beta-6)(\beta-8)} - \frac{2(\alpha+4)(\alpha+2)^2\varrho^4}{\alpha^3(\beta-2)(\beta-4)(\beta-6)^2} \right. \\ & \left. + \frac{(\alpha+2)^2\varrho^4}{\alpha^2(\beta-2)(\beta-4)^2(\beta-6)} \right] e^{-2\theta \frac{\beta-4}{\beta-2} t} + \frac{8(\alpha+2)^2(\alpha+\beta-2)\varrho^4}{\alpha^3(\beta-2)^2(\beta-4)(\beta-6)^2} e^{-\theta t} = \\ = & \frac{8(\alpha+2)(\alpha+\beta-2)(\alpha+\beta-4)\varrho^4}{\alpha^3(\beta-2)(\beta-4)^2(\beta-6)^2(\beta-8)} e^{-2\theta \frac{\beta-4}{\beta-2} t} + \frac{8(\alpha+2)^2(\alpha+\beta-2)\varrho^4}{\alpha^3(\beta-2)^2(\beta-4)(\beta-6)^2} e^{-\theta t}. \end{aligned}$$

$I_{2,-1}$ $(\alpha > 2, \beta > 4)$

$$\begin{aligned} C_{2,-1}(t) = & \left[\frac{\varrho}{\beta-2} - \frac{2(\alpha+2)\varrho}{\alpha(\beta-6)} + \frac{(\alpha+2)\beta\varrho}{(\alpha-2)(\beta-4)(\beta-6)} \right] e^{-2\theta \frac{\beta-4}{\beta-2} t} \\ & - \frac{4(\alpha+2)(\alpha+\beta-2)\varrho}{\alpha(\alpha-2)(\beta-2)(\beta-6)} e^{-\theta t} = \\ = & \frac{8(\alpha+\beta-2)(\alpha+\beta-4)\varrho}{\alpha(\alpha-2)(\beta-2)(\beta-4)(\beta-6)} e^{-2\theta \frac{\beta-4}{\beta-2} t} - \frac{4(\alpha+2)(\alpha+\beta-2)\varrho}{\alpha(\alpha-2)(\beta-2)(\beta-6)} e^{-\theta t}. \end{aligned}$$

Remark 3.1. The relations for various covariances in Example 3.1 were deduced for $t \geq 0$. Clearly,

$$C_{\upsilon,\chi}(t) = C_{\chi,\upsilon}(-t), \quad t \in \mathbb{R}.$$

Observe in addition that X, like any ergodic one-dimensional diffusion, is time reversible in the sense that the process $X^{st}(-t), t \in \mathbb{R}$, has the same distribution with X^{st}. This gives the identities

$$C_{\upsilon,\chi}(t) = C_{\chi,\upsilon}(-t) = C_{\upsilon,\chi}(|t|), \quad t \in \mathbb{R}.$$

Hence, the above list provides explicit expressions for the point-wise covariations

$$C_{\upsilon,\chi}(t) = \operatorname{Cov}\left(\left(X_t^{st}\right)^{\upsilon}, \left(X_0^{st}\right)^{\chi} \right)$$

for all $t \in \mathbb{R}$ and every pair

$$(\upsilon, \chi), \ \upsilon, \chi \in \{-1, 1, 2\}$$

with the only exception

$$(\upsilon, \chi) = (-1, -1).$$

Proposition 3.2. *Covariances* $C_\upsilon(t,s) = \operatorname{Cov}\left(X_t^{st} X_s^{st}, \left(X_0^{st}\right)^{\upsilon} \right)$ *for* $\upsilon \in \{0, 1, 2\}$ *are given by following expressions:*

- *for* $t \geq s \geq 0:$

$$C_\upsilon(t,s) = e^{-\theta(t-s)} C_{2,\upsilon}(s) + \left[1 - e^{-\theta(t-s)}\right] C_{1,\upsilon}(s) \langle f_1 \rangle,$$

- *for* $t \geq 0 > s:$

$$\begin{aligned} C_\upsilon(t,s) &= e^{-\theta t}\left[C_{1,\upsilon+1}(s) - C_{1,\upsilon}(s)\langle f_1 \rangle + C_{1,\upsilon}(0)\langle f_1 \rangle \right] + C_{1,\upsilon}(s)\langle f_1 \rangle - C_{1,1}(t-s)\langle f_\upsilon \rangle \\ &= e^{-\theta t + \theta s} C_{1,\upsilon+1}(0) + \left[e^{-\theta t} + e^{\theta s} - e^{-\theta t + \theta s} \right] C_{1,\upsilon}(0)\langle f_1 \rangle - e^{-\theta t + \theta s} C_{1,1}(0)\langle f_\upsilon \rangle. \end{aligned}$$

Remark 3.2. Note that $C_\upsilon(t,s) = C_\upsilon(s,t)$ and $C_\upsilon(t,s) = C_\upsilon(-t,-s)$. Hence, the two cases considered above ($t \geq s \geq 0$ and $t \geq 0 > s$) cover all possibilities.

Example 3.2. This example contains the list of covariances that can be calculated using the formulae from the Proposition 3.2. The items in the list are numerated as II_υ and provided by the respective conditions on the parameters α, β.

II_{-1} $(\alpha > 2, \beta > 4)$

$$
\begin{aligned}
C_{-1}(t,s) = {} & \frac{8(\alpha+\beta-2)(\alpha+\beta-4)\varrho}{\alpha(\alpha-2)(\beta-2)(\beta-4)(\beta-6)} e^{-\theta(t-s)-2\theta\frac{\beta-4}{\beta-2}s} - \frac{4(\alpha+2)(\alpha+\beta-2)\varrho}{\alpha(\alpha-2)(\beta-2)(\beta-6)} e^{-\theta t} \\
& - \frac{2(\alpha+\beta-2)\varrho}{(\beta-2)^2(\alpha-2)} \Big[e^{-\theta s} - e^{-\theta t}\Big], \quad t \geq s \geq 0,
\end{aligned}
$$

$$
C_{-1}(t,s) = -\Big[e^{-\theta t} + e^{\theta s} - e^{-\theta t+\theta s}\Big] \frac{2(\alpha+\beta-2)\varrho}{(\beta-2)^2(\alpha-2)} - \frac{2\beta(\alpha+\beta-2)\varrho}{(\alpha-2)(\beta-2)^2(\beta-4)} e^{-\theta t+\theta s}, \quad t \geq 0 > s.
$$

II_1 $(\alpha > 2, \beta > 6)$

$$
C_1(t,s) = \frac{4(\alpha+2)(\alpha+\beta-2)\varrho^3}{\alpha^2(\beta-2)^2(\beta-4)(\beta-6)} e^{-\theta t} + \frac{2(\alpha+\beta-2)\varrho^3}{\alpha(\beta-2)^3(\beta-4)} \Big[e^{-\theta s} - e^{-\theta t}\Big], \quad t \geq s \geq 0,
$$

$$
C_1(t,s) = \frac{4(\alpha+2)(\alpha+\beta-2)\varrho^3}{\alpha^2(\beta-2)^2(\beta-4)(\beta-6)} e^{-\theta t+\theta s} + \frac{2(\alpha+\beta-2)\varrho^3}{\alpha(\beta-2)^3(\beta-4)} \Big[e^{-\theta t} + e^{\theta s} - 2e^{-\theta t+\theta s}\Big], \quad t \geq 0 > s.
$$

II_2 $(\alpha > 2, \beta > 8)$

$$
\begin{aligned}
C_2(t,s) = {} & \frac{8(\alpha+2)(\alpha+\beta-2)(\alpha+\beta-4)\varrho^4}{\alpha^3(\beta-2)(\beta-4)^2(\beta-6)^2(\beta-8)} e^{-\theta(t-s)-2\theta\frac{\beta-4}{\beta-2}s} + \frac{8(\alpha+2)^2(\alpha+\beta-2)\varrho^4}{\alpha^3(\beta-2)^2(\beta-4)(\beta-6)^2} e^{-\theta t} \\
& + \frac{4(\alpha+2)(\alpha+\beta-2)\varrho^4}{\alpha^2(\beta-2)^3(\beta-4)(\beta-6)} \Big[e^{-\theta s} - e^{-\theta t}\Big], \quad t \geq s \geq 0,
\end{aligned}
$$

$$
\begin{aligned}
C_2(t,s) = {} & \frac{6(\alpha+2)(\alpha+4)(\alpha+\beta-2)\varrho^4}{\alpha^3(\beta-2)^2(\beta-4)(\beta-6)(\beta-8)} e^{-\theta t+\theta s} \\
& + \frac{4(\alpha+2)(\alpha+\beta-2)\varrho^4}{\alpha^2(\beta-2)^3(\beta-4)(\beta-6)} \Big[e^{-\theta t} + e^{\theta s} - e^{-\theta t+\theta s}\Big] - \frac{2(\alpha+2)(\alpha+\beta-2)\varrho^4}{\alpha^2(\beta-2)^3(\beta-4)^2} e^{-\theta t+\theta s}, \quad t \geq 0 > s.
\end{aligned}
$$

Proposition 3.3. *Covariances* $C(t_1,\dots,t_4) = \mathrm{Cov}\left(X^{st}_{t_1}X^{st}_{t_2},\ X^{st}_{t_3}X^{st}_{t_4}\right)$, *where* $\{t_1,\dots,t_4\} = \{t_{[1]},\dots,t_{[4]}\}$ *and* $t_{[1]} \leq \dots t_{[4]}$ *can be explicitly determined in terms of the functions* $C_{\upsilon,\chi}(t)$ *and* $C_\upsilon(t,s)$. *In particular,*

$$
\begin{aligned}
C(t_1,\dots,t_4) = {} & \mathrm{Cov}\left(X^{st}_{t_{[1]}},\ X^{st}_{t_{[2]}}X^{st}_{t_{[3]}}X^{st}_{t_{[4]}}\right) - C_{1,1}(t_2-t_1)C_{1,1}(t_4-t_3) + \langle f_1\rangle \mathrm{Cov}\left(X^{st}_{t_{[2]}},\ X^{st}_{t_{[3]}}X^{st}_{t_{[4]}}\right) \\
& + \langle f_1\rangle^2 \Big[C_{1,1}(t_{[4]}-t_{[3]}) - C_{1,1}(t_2-t_1) - C_{1,1}(t_4-t_3)\Big],
\end{aligned}
$$

where

$$
\mathrm{Cov}\left(X^{st}_{t_{[2]}},\ X^{st}_{t_{[3]}}X^{st}_{t_{[4]}}\right) = e^{-\theta(t_{[3]}-t_{[2]})}\Big[C_{2,1}(t_{[4]}-t_{[3]}) - \langle f_1\rangle C_{1,1}(t_{[4]}-t_{[3]}) + \langle f_1\rangle C_{1,1}(0)\Big]
$$

and

$$
\begin{aligned}
\mathrm{Cov}\left(X^{st}_{t_{[1]}},\ X^{st}_{t_{[2]}}X^{st}_{t_{[3]}}X^{st}_{t_{[4]}}\right) = {} & e^{-\theta(t_{[2]}-t_{[1]})}\Big[C_2(t_{[4]}-t_{[2]}, t_{[3]}-t_{[2]}) \\
& - \langle f_1\rangle C_1(t_{[4]}-t_{[2]}, t_{[3]}-t_{[2]}) + C_{1,1}(0)\big(C_{1,1}(t_{[4]}-t_{[3]}) + \langle f_1\rangle^2\big)\Big].
\end{aligned}
$$

In the next section we introduce the technique for calculating the specific covariances when the previous evolutionary approach is not applicable (e.g. $C_{-1,-1}(t)$).

3.2. Integral formula for the continuous-time asymptotic covariances

In this section we consider continuous-time asymptotic covariances

$$\sigma_{f,g} = \int_{-\infty}^{\infty} \operatorname{Cov}\left(f(X_t^{st}), g(X_0^{st})\right) dt := \lim_{T\to\infty} \int_{-T}^{T} \operatorname{Cov}\left(f(X_t^{st}), g(X_0^{st})\right) dt \tag{3.10}$$

for functions $f, g\colon (0,\infty) \to \mathbb{R}$ satisfying some growth conditions. As an extension of the Subsection 3.1, here we further develop the approach for their explicit calculations, needed for determining the explicit form of the asymptotic covariance matrix in the light of the CLT from Appendix B. Calculations from the previous sections can be used to derive the (mutual) asymptotic covariances for empirical moments $\overline{m}_\upsilon, \upsilon \in \{-1, 1, 2\}$, and empirical covariances $C_{\upsilon,\chi}(t)$ in the continuous-time setting in all possible cases except the one where $\upsilon = \chi = -1$. In this section we provide the tool which is applicable in particular in this exceptional case.

Below we denote by π the invariant distribution of the FSD X and by $\mathbb{X}$ we denote its state space.

Theorem 3.1.

1. *Let*

$$|f(x)| \le C\left(x^{-\gamma_1} + x^{\delta_1}\right), \quad |g(x)| \le C\left(x^{-\gamma_2} + x^{\delta_2}\right), \quad x \in \mathbb{X}, \tag{3.11}$$

with some constants $C, \gamma_1, \gamma_2, \delta_1, \delta_2$. Assume that $\gamma_i, \delta_i, i = 1, 2$ satisfy

$$\gamma_i < \frac{\alpha}{2}, \quad \delta_i < \frac{\beta}{2}, \tag{3.12}$$

and

$$\gamma_1 + \gamma_2 < \frac{\alpha}{2} + 1, \quad \delta_1 + \delta_2 < \frac{\beta}{2}. \tag{3.13}$$

Then the limit (3.10) is well defined.

2. *Assume, in addition, that f, g are centered in the sense that*

$$\int_{\mathbb{X}} f\, d\pi = \int_{\mathbb{X}} g\, d\pi = 0,$$

and $\alpha \notin \{2(m+1), m \in \mathbb{N}\}$. Then

$$\begin{aligned}\sigma_{f,g} = \frac{(\beta-2)}{\theta B(\alpha/2, \beta/2)} \int_0^1 &\left(\int_0^u f\left(\frac{\varrho(1-v)}{\alpha v}\right) v^{\beta/2-1}(1-v)^{\alpha/2-1}\, dv\right) \times \\ &\times \left(\int_0^u g\left(\frac{\varrho(1-v)}{\alpha v}\right) v^{\beta/2-1}(1-v)^{\alpha/2-1}\, dv\right) u^{-\beta/2-1}(1-u)^{-\alpha/2}\, du,\end{aligned} \tag{3.14}$$

where $B(\cdot,\cdot)$ is the standard beta function.

The following examples illustrate the use of the last theorem.

Example 3.3. Let $f(x) = g(x) = x^{-1} - \frac{\alpha\beta}{(\alpha-2)\varrho} = f_{-1}(x) - \langle f_{-1}\rangle$. Conditions of the Theorem 3.1 are satisfied with $\gamma_{1,2} = 1$ and $\delta > 0$ small enough because $\alpha > 2, \beta > 2$. The function f is centered and hence we can calculate the covariance $\sigma_{f,f}$ according to (3.14). For the function F defined by (5.23) we have

$$\begin{aligned}F(u) &= \int_0^u \left(\frac{\alpha}{\varrho} v^{\beta/2}(1-v)^{\alpha/2-2} - \frac{\alpha\beta}{(\alpha-2)\varrho} v^{\beta/2-1}(1-v)^{\alpha/2-1}\right) dv \\ &= \frac{\alpha\beta}{\varrho}\left(\frac{1}{\beta} B(\beta/2+1, \alpha/2-1; u) - \frac{1}{\alpha-2} B(\beta/2, \alpha/2; u)\right) \\ &= \frac{\alpha\beta}{\varrho}\Bigg(\frac{u^{\beta/2+1}}{\beta(\beta/2+1)}\, {}_2F_1(\beta/2+1, 2-\alpha/2; \beta/2+2; u) \\ &\qquad - \frac{u^{\beta/2}}{(\alpha-2)(\beta/2)}\, {}_2F_1(\beta/2, 1-\alpha/2; \beta/2+1; u)\Bigg), \quad u \in (0,1).\end{aligned}$$

Here we have used first

$$B(a, b; z) = \int_0^z t^{a-1}(1-t)^{b-1}\, dt, \quad z \in (0,1),$$

under the assumption $a > 0, b > 0$, and then the definition of the incomplete beta function $B(a,b;z) = \frac{z^a}{a}\,{}_2F_1(a, 1-b; a+1; z)$, where ${}_2F_1$ is the hypergeometric function defined by

$${}_2F_1(a,b;c,z) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \frac{z^n}{n!}, \quad |z| < 1,$$

and where $(k)_n$, $k \in \{a, b, c\}$, is the Pochhammer symbol defined by

$$(k)_n = \begin{cases} 1 & , \quad n = 0 \\ k(k+1)\cdots(k+n-1) & , \quad n > 0 \end{cases}.$$

Then according to the recursive relation (for various relation on hypergeometric functions we refer e.g. to Abramowitz and Stegun (1967), Section 15)

$${}_2F_1(a,b;c;z) + \frac{bz}{c}\,{}_2F_1(a+1,b+1;c+1;z) = {}_2F_1(a+1,b;c;z), \quad |z| < 1, \quad c \neq -1,$$

with $a = \beta/2$, $b = 1 - \alpha/2$, $c = \beta/2 + 1$ we have

$$\begin{aligned} F(u) &= -\frac{\alpha u^{\beta/2}}{\varrho(\alpha/2-1)} \left(\frac{u(1-\alpha/2)}{(\beta/2+1)}\,{}_2F_1(\beta/2+1, 2-\alpha/2; \beta/2+2; u) \right. \\ &\qquad \left. + \,{}_2F_1(\beta/2, 1-\alpha/2; \beta/2+1; u) \right) \\ &= -\frac{\alpha u^{\beta/2}}{\varrho(\alpha/2-1)}\,{}_2F_1(\beta/2+1, 1-\alpha/2; \beta/2+1; u). \end{aligned}$$

According to the symmetry of the Gauss hypergeometric function in its first two parameters as well as the simple relation ${}_2F_1(a,b;b;z) = (1-z)^{-a}$, we obtain

$$F(u) = -\frac{\alpha}{\varrho(\alpha/2-1)} u^{\beta/2} (1-u)^{\alpha/2-1}.$$

Then

$$\begin{aligned} \sigma_{f,f} &= \frac{(\beta-2)}{\theta B(\alpha/2, \beta/2)} \frac{\alpha^2}{\varrho^2(\alpha/2-1)^2} \int_0^1 u^{2(\beta/2)-\beta/2-1} (1-u)^{2(\alpha/2-1)-\alpha/2}\, du \\ &= \frac{\alpha^2(\beta-2) B(\alpha/2-1, \beta/2)}{\theta\varrho^2(\alpha/2-1)^2 B(\alpha/2, \beta/2)} = \frac{4\alpha^2(\beta-2)(\alpha+\beta-2)}{\theta\varrho^2(\alpha-2)^3}. \end{aligned}$$

Example 3.4. Similar calculation with $g(x) = x - \frac{\varrho}{\beta-2}$ gives

$$G(u) = \frac{\varrho}{\alpha(\beta/2-1)} u^{\beta/2-1} (1-u)^{\alpha/2}, \quad u \in (0,1)$$

and

$$\sigma_{f,g} = -\frac{4(\alpha+\beta-2)}{\theta(\beta-2)(\alpha-2)},$$

$$\sigma_{g,g} = \frac{4\varrho^2(\alpha+\beta-2)}{\theta\alpha(\beta-2)^2(\beta-4)}.$$

Note that the last two formulae can also be obtained by integration of respective expressions for $C_{1,-1}$ and $C_{1,1}$ from Subsection 3.1.

3.3. Parameter estimation in continuous-time setting

The results of this section are a direct application of the LLN and the CLT from Appendix B, i.e. Theorem 3.4 from Kulik and Leonenko (2013), to the GMM estimator $(\widehat{\alpha}_c, \widehat{\beta}_c, \widehat{\kappa}_c, \widehat{\theta}_c)$ of the parameter $(\alpha, \beta, \kappa, \theta)$ of the non-stationary FSD in continuous-time setting, i.e. on fixed time interval $[0, T]$. In particular, the main novelty of this section is the application of the techniques developed in 3.1 and 3.2 for calculation of the covariance matrix in asymptotic normality framework (we assume that all covariance matrices are non-degenerate).

In Theorem 3.2 we use the following GMM estimators:

$$\widehat{\alpha}_c = \frac{2(\overline{m}_{-1,c}\overline{m}_{1,c}\overline{m}_{2,c} - \overline{m}_{1,c}^2)}{\overline{m}_{-1,c}\overline{m}_{1,c}\overline{m}_{2,c} - 2\overline{m}_{2,c} + \overline{m}_{1,c}^2}, \quad \widehat{\beta}_c = \frac{4\overline{m}_{-1,c}(\overline{m}_{1,c}^2 - \overline{m}_{2,c})}{2\overline{m}_{-1,c}\overline{m}_{1,c}^2 - \overline{m}_{-1,c}\overline{m}_{2,c} - \overline{m}_{1,c}},$$
$$\widehat{\kappa}_c = \overline{m}_{1,c}, \qquad \widehat{\theta}_c(t) = -\frac{1}{t}\log\left(\frac{\overline{m}_{1,1,c}(t) - \overline{m}_{1,c}^2}{\overline{m}_{2,c} - \overline{m}_{1,c}^2}\right). \tag{3.15}$$

Empirical moments $\overline{m}_{-1,c}$, $\overline{m}_{1,c}$, $\overline{m}_{2,c}$ and $\overline{m}_{1,1,c}$ of the invariant distribution (2.5) are given at the beginning of the Section 3. Now we state the main result of this section.

Theorem 3.2. *Let $\alpha > 2, \beta > 8$. Then, for an arbitrary initial distribution of the FSD, the estimator $(\widehat{\alpha}_c, \widehat{\beta}_c, \widehat{\kappa}_c, \widehat{\theta}_c)$ given by (3.15) has the following properties:*

(i) P-consistency, i.e. for every $t > 0$

$$(\widehat{\alpha}_c, \widehat{\beta}_c, \widehat{\kappa}_c, \widehat{\theta}_c(t)) \xrightarrow{P} (\alpha, \beta, \kappa, \theta), \quad T \to \infty$$

(ii) asymptotic normality, i.e. for every $t > 0$

$$(\widehat{\boldsymbol{\Sigma}}_c)^{-1/2}\sqrt{T}\,(\widehat{\alpha}_c - \alpha, \widehat{\beta}_c - \beta, \widehat{\kappa}_c - \kappa, \widehat{\theta}_c(t) - \theta) \xrightarrow{D} \mathcal{N}(\mathbf{0}, \mathbf{I}), \quad T \to \infty,$$

where $\widehat{\boldsymbol{\Sigma}}_c$ is the asymptotic covariance matrix depending on P-consistent estimate of the parameter $(\alpha, \beta, \kappa, \theta)$ and $\mathbf{I}$ is the (4×4) identity matrix.

Remark 3.3. P-consistency and asymptotic normality of the estimator of parameter $(\alpha, \beta, \kappa, \theta)$ in the discrete-time setting is proved by the means of the Theorem 3.2 in Remark 5.1 by simple use of the Chebyshev inequality, Slutsky theorem and delta method.

Example 3.5. Let us consider a special parametrization of the FSD where $\varrho = \beta$, or equivalently $\kappa = \frac{\beta}{\beta-2}$ and parameter θ is known (the so-called canonical parametrization). In this special case

$$m_{-1} = \frac{\alpha}{\alpha-2}, \quad m_1 = \frac{\beta}{\beta-2}, \quad m_2 = \frac{\beta^2(\alpha+2)}{\alpha(\beta-2)(\beta-4)}.$$

We introduce the following estimators of unknown parameter (α, β):

1. estimator $(\widehat{\alpha}, \widehat{\beta})$ depending on empirical moments $\overline{m}_{1,c}$ and $\overline{m}_{2,c}$:

$$\widehat{\alpha}_c = \frac{2\,\overline{m}_{1,c}^2}{\overline{m}_{2,c}(2 - \overline{m}_{1,c}) - \overline{m}_{1,c}^2}, \quad \widehat{\beta}_c = \frac{2\overline{m}_{1,c}}{\overline{m}_{1,c} - 1}, \tag{3.16}$$

2. estimator $(\widehat{\alpha}, \widehat{\beta})$ depending on empirical moments $\overline{m}_{-1,c}$ and $\overline{m}_{1,c}$:

$$\widehat{\alpha}_c = \frac{2\overline{m}_{-1,c}}{\overline{m}_{-1,c} - 1}, \quad \widehat{\beta}_c = \frac{2\overline{m}_{1,c}}{\overline{m}_{1,c} - 1}. \tag{3.17}$$

Estimator (3.17) is clearly analytically simpler, with $\widehat{\alpha}_c$ depending only on $\overline{m}_{-1,c}$ and $\widehat{\beta}_c$ depending only on $\overline{m}_{1,c}$. Also, the covariance matrix $\Sigma(\alpha, \beta) = \Sigma$ related to the estimator $(\overline{m}_{-1,c}, \overline{m}_{1,c})$ is simpler than the covariance matrix related to the estimator $(\overline{m}_{1,c}, \overline{m}_{2,c})$:

1. elements of the covariance matrix related to the estimator $(\overline{m}_{-1,c}, \overline{m}_{1,c})$:

$$\Sigma_{11} = \frac{\alpha^2(\alpha-2)(\beta-2)(\alpha+\beta-2)}{\theta\beta^2},$$
$$\Sigma_{12} = -\frac{(\alpha-2)(\beta-2)(\alpha+\beta-2)}{\theta} = \Sigma_{21}, \tag{3.18}$$
$$\Sigma_{22} = \frac{(\beta-2)^2\beta^2(\alpha+\beta-2)}{\theta\alpha(\beta-4)}.$$

2. elements of the covariance matrix related to the estimator $(\overline{m}_{1,c}, \overline{m}_{2,c})$:

$$\Sigma_{11} = \frac{(\beta-2)^2\beta^2(\alpha+\beta-2)}{\theta\alpha(\beta-4)},$$
$$\Sigma_{12} = -\frac{(\alpha+2)(\beta-8)(\beta-2)^2\beta(\alpha+\beta-2)}{(\beta-6)(\beta-4)^2\theta} = \Sigma_{12},$$
$$\Sigma_{22} = \frac{\alpha(\alpha+2)(\beta-2)^2(\alpha+\beta-2)\left(\alpha\left(\beta^3-24\beta^2+194\beta-520\right)+2\left(\beta^3-23\beta^2+184\beta-496\right)\right)}{(\beta-8)(\beta-6)^2(\beta-4)^3\theta}. \tag{3.19}$$

To illustrate the difference, for parameters $(\alpha, \beta) = (10, 20)$, assuming $\theta = 1$, we compute the matrix Σ for both estimators:

1. estimator $(\overline{m}_{-1,c}, \overline{m}_{1,c})$:

$$\begin{bmatrix} 1008 & -4032 \\ -4032 & 22680 \end{bmatrix},$$

2. estimator $(\overline{m}_{1,c}, \overline{m}_{2,c})$:

$$\begin{bmatrix} 22680 & -7290 \\ -7290 & 2437.23 \end{bmatrix}.$$

Estimators (3.17) and (3.16) are transformations of $(\overline{m}_{-1,c}, \overline{m}_{1,c})$ and $(\overline{m}_{1,c}, \overline{m}_{2,c})$, respectively. Note that the range of true values for the parameter (α, β) required for the estimator (3.17) to be asymptotically normal is wider than the respective range for (3.16): $\alpha > 2, \beta > 4$ against $\alpha > 2, \beta > 8$.

4. Simulation results

In this section we present results of the simulation study based on 10000 sample paths of the non-stationary FSD on interval $[0, T]$ with time-step $\Delta t = 1/252$, where 252 is taken to be the conventional number of working days in one year. The time-horizon is taken to be $T = 30000$, while some other options, explicitly $T \in \{100, 500, 1000, 10000\}$, are considered in the Remark 4.1. Sample paths are simulated in *Rcpp* environment in statistical software *R*, by using the Milstein approximation scheme (for details of the computational procedure we refer e.g. to Iacus (2008)). Simulation is performed in two different scenarios, depending on values of parameters $\alpha > 2$, $\beta > 8$, $\kappa > 0$ and $\theta > 0$. In each scenario three possible starting distributions are observed: uniform distribution on $(0, 1)$, normal distribution with mean 10 and variance 1 and Fisher-Snedecor distribution with shape parameters both equal to 10.

Mean values of estimates of moments m_{-1}, m_1, m_2 and $m_{1,1}(k)$ for $k = 1$ and $k = 5$ (see expression (2.6))of the invar iant distribution (2.5), based on the corresponding estimators given at the beginning of the Section 3, are given in Table 1. Mean values, as well as RMSEs (in %) and REs (in %) of estimates of parameters κ, α, β and θ, based on estimators (3.15), are given in Tables 2 and 3, respectively.

Estimation of moments and mixed moments					
				EX_tX_{t+k}	
	$E(X_t)^{-1}$	EX_t	$E(X_t)^2$	$k=1$	$k=5$
Scenario 1 $(\alpha = 10, \beta = 10, \kappa = 5, \theta = 0.05)$					
Theoretical values	0.3125	5	40	39.2684	36.682
$X_0 \sim \mathcal{U}(0,1)$	0.3131	4.9863	39.4061	38.6835	36.1345
$X_0 \sim \mathcal{N}(10,1)$	0.3125	4.9948	39.5151	38.7888	36.2281
$X_0 \sim \mathcal{FS}(10,10)$	0.3129	4.9877	39.4231	38.6999	36.1491
Scenario 2 $(\alpha = 5, \beta = 11, \kappa = 7, \theta = 0.025)$					
Theoretical values	0.291	7	88.2	87.2321	83.5939
$X_0 \sim \mathcal{U}(0,1)$	0.2919	6.9777	86.5939	85.6379	82.0476
$X_0 \sim \mathcal{N}(10,1)$	0.2907	6.9936	86.9163	85.9558	82.3479
$X_0 \sim \mathcal{FS}(10,10)$	0.2914	6.9755	86.5311	85.5757	81.9892

Table 1: Mean values of estimated moments and mixed moments of FSD.

Estimation of parameters κ, α and β									
	κ	**RMSE (%)**	**RE (%)**	α	**RMSE (%)**	**RE (%)**	β	**RMSE (%)**	**RE (%)**
Scenario 1									
Theoretical values	$\kappa = 5$			$\alpha = 10$			$\beta = 10$		
$X_0 \sim \mathcal{U}(0,1)$	4.9863	0.2738	-0.2731	9.6714	3.3972	-3.2856	10.7455	6.9375	7.4547
$X_0 \sim \mathcal{N}(10,1)$	4.9948	0.1049	-0.1048	9.6655	3.4606	-3.3448	10.7568	7.0359	7.5684
$X_0 \sim \mathcal{FS}(10,10)$	4.9877	0.2461	-0.2455	9.6903	3.1958	-3.0968	10.7427	6.9135	7.4269
Scenario 2									
Theoretical values	$\kappa = 7$			$\alpha = 5$			$\beta = 11$		
$X_0 \sim \mathcal{U}(0,1)$	6.9777	0.3192	-0.3182	4.9094	1.8460	-1.8126	12.4899	11.9292	13.5449
$X_0 \sim \mathcal{N}(10,1)$	6.9936	0.0916	-0.0915	4.9237	1.5496	-1.5259	12.4721	11.8033	13.3829
$X_0 \sim \mathcal{F}(10,10)$	6.9755	0.3508	-0.3496	4.9299	1.4216	-1.4016	12.4437	11.6016	13.1243

Table 2: Mean values, RMSE and RE for estimates of parameters α, β and κ of FSD.

Estimation of parameter θ						
	$k = 1$	**RMSE (%)**	**RE (%)**	$k = 5$	**RMSE (%)**	**RE (%)**
Scenario 1 (true values: $\alpha = 10, \beta = 10, \kappa = 5, \theta = 0.05$**)**						
$X_0 \sim \mathcal{U}(0,1)$	0.0515	3.0736	3.1711	0.0516	3.1349	3.2363
$X_0 \sim \mathcal{N}(10,1)$	0.0517	3.3954	3.5147	0.0518	3.4620	3.5862
$X_0 \sim \mathcal{FS}(10,10)$	0.0516	3.1398	3.2415	0.0517	3.2093	3.3157
Scenario 2 (true values: $\alpha = 5, \beta = 11, \kappa = 7, \theta = 0.025$**)**						
$X_0 \sim \mathcal{U}(0,1)$	0.0261	4.1839	4.3666	0.0261	4.2685	4.4588
$X_0 \sim \mathcal{N}(10,1)$	0.0261	4.3328	4.5290	0.0262	4.4425	4.6489
$X_0 \sim \mathcal{FS}(10,10)$	0.0261	4.1771	4.3592	0.0261	4.2357	4.4231

Table 3: Mean values, RMSE and RE for estimation of autocorrelation parameter θ of FSD.

From the theoretical point of view, GMM based estimates obtained here could be valuable as the starting values in some more advanced and numerically more complex estimation procedures.

Empirical variances $\widetilde{\sigma}^2_\alpha$, $\widetilde{\sigma}^2_\beta$, $\widetilde{\sigma}^2_\kappa$ and $\widetilde{\sigma}^2_\theta$ and asymptotic covariance matrices calculated for means of estimated values of parameters α, β, κ and θ given in Tables 2 and 3, in scenario 1 ($\alpha = 10, \beta = 10, \kappa = 5, \theta = 0.05$ for lag $k = 5$), for all three observed starting distributions are of the following form:

- $X_0 \sim \mathcal{U}(0,1)$
  empirical variances: $\widetilde{\sigma}^2_\alpha = 0.973843$, $\widetilde{\sigma}^2_\beta = 1.98796$, $\widetilde{\sigma}^2_\kappa = 0.01846974$, $\widetilde{\sigma}^2_\theta = 0.00001857$,

$$\widehat{\mathbf{\Sigma}}_c = \begin{bmatrix} 3.0257281683 & -4.073098390 & 0.0360270976 & 0.0007536235 \\ -4.0730983895 & 6.089502030 & -0.1092301826 & -0.0022849033 \\ 0.0360270976 & -0.109230183 & 0.0181306898 & 0.0003792621 \\ 0.0007536235 & -0.002284903 & 0.0003792621 & 0.0013140138 \end{bmatrix}, \tag{4.1}$$

- $X_0 \sim \mathcal{N}(10,1)$
  empirical variances $\widetilde{\sigma}^2_\alpha = 0.9373594$, $\widetilde{\sigma}^2_\beta = 1.928097$, $\widetilde{\sigma}^2_\kappa = 0.01912949$, $\widetilde{\theta}^2_\beta = 0.00001861$,

$$\widehat{\mathbf{\Sigma}}_c = \begin{bmatrix} 2.9908388099 & -4.041027043 & 0.0357738599 & 0.0007505208 \\ -4.0410270435 & 6.066711470 & -0.1090047438 & -0.0022868745 \\ 0.0357738599 & -0.109004744 & 0.0181164481 & 0.0003800756 \\ 0.0007505208 & -0.002286874 & 0.0003800756 & 0.0013130525 \end{bmatrix}, \tag{4.2}$$

- $X_0 \sim \mathcal{FS}(10,10)$
  empirical variances: $\widetilde{\sigma}^2_\alpha = 1.000106$, $\widetilde{\sigma}^2_\beta = 1.999009$, $\widetilde{\sigma}^2_\kappa = 0.01879328$, $\widetilde{\sigma}^2_\theta = 0.00001888$,

$$\widehat{\mathbf{\Sigma}}_c = \begin{bmatrix} 3.0531077176 & -4.089963619 & 0.0361923961 & 0.0007580354 \\ -4.0899636190 & 6.083323154 & -0.1091180825 & -0.0022854351 \\ 0.0361923961 & -0.109118083 & 0.0181147595 & 0.0003794065 \\ 0.0007580354 & -0.002285435 & 0.0003794065 & 0.0013152270 \end{bmatrix}. \tag{4.3}$$

The presented results testify that in the proposed scenarios, based on the simulation of 10000 independent sample paths consisting of $30000 \cdot 252$ points each, we are close to the real parameter values and that the empirical and the asymptotic variances of proposed GMM estimator are comparable (in a sense that they are of the approximately the same order of magnitude).

Remark 4.1. Non-stationary FSD (1.1) reaches its invariant (steady-state) distribution (2.5) very fast (e.g. for $\theta = 0.05$ approximately for $T = 100$), so the results of the simulation study presented in Tables 1, 2 and 3, based on 10000 sample paths on time-interval $[0, 30000]$, resemble the stationary setting. However, the observed sample paths are simulated from the non-stationary process and the results in the Table 4, for time-horizon $T \in \{100, 500, 1000, 10000\}$, show that estimation based on shorter time-intervals, in which the diffusion is not in its invariant distribution for very long time, are not on satisfactory level.

Estimation of parameters κ, α, β and θ								
Theoretical values	$\kappa = 5$		$\alpha = 10$		$\beta = 10$		$\theta = 0.05$, $k = 1$	
$T = 100$								
	κ	**RMSE (%)**	α	**RMSE (%)**	β	**RMSE (%)**	θ	**RMSE (%)**
$X_0 \sim \mathcal{U}(0,1)$	4.0838	22.4348	6.5395	52.9168	9.0412	10.6043	0.0966	48.2241
$X_0 \sim \mathcal{FS}(10,10)$	4.2578	17.4322	8.7197	14.6824	35.6906	71.9814	0.1038	51.8333
$T = 500$								
	κ	**RMSE (%)**	α	**RMSE (%)**	β	**RMSE (%)**	θ	**RMSE (%)**
$X_0 \sim \mathcal{U}(0,1)$	4.8278	3.5665	7.6774	30.2523	18.6641	46.4211	0.0665	24.8539
$X_0 \sim \mathcal{FS}(10,10)$	4.8311	3.4952	8.0709	23.9016	68.3760	85.3749	0.0678	26.2916
$T = 1000$								
	κ	**RMSE (%)**	α	**RMSE (%)**	β	**RMSE (%)**	θ	**RMSE (%)**
$X_0 \sim \mathcal{U}(0,1)$	4.9218	1.5889	8.3667	19.5214	56.8924	82.4229	0.0607	17.6082
$X_0 \sim \mathcal{FS}(10,10)$	4.9203	1.6205	8.4318	18.5985	26.7961	62.6811	0.0614	18.5682
$T = 10000$								
	κ	**RMSE (%)**	α	**RMSE (%)**	β	**RMSE (%)**	θ	**RMSE (%)**
$X_0 \sim \mathcal{U}(0,1)$	4.9889	0.2229	9.9086	0.9221	11.3669	12.0259	0.0522	4.2672
$X_0 \sim \mathcal{FS}(10,10)$	4.9960	0.0794	11.6923	14.4736	11.3174	11.6407	0.0521	4.1174

Table 4: Mean values and RMSE for estimates of parameters α, β, κ and θ of FSD.

This is surely a drawback of the proposed GMM estimator for non-stationary FSD parameters. However, since α, β and κ are present both in the defining SDE and in the density of its invariant distribution, it is hard to expect to obtain estimations arbitrarily close to the real values of parameters based on the sample of trajectories that didn't reach or are not long enough in the stationary setting. This deficiency is also reflected in the asymptotical performance of GMM estimator $(\hat{\alpha}, \hat{\beta}, \hat{\kappa}, \hat{\theta})$. In particular, there is a discordance between empirical variances $\widetilde{\sigma}^2_\alpha$ and $\widetilde{\sigma}^2_\beta$ with the corresponding asymptotic covariances of estimators for α and β, given in the estimated covariance matirces (4.1), (4.2) and (4.3).

Remark 4.2. From Table 3 it could be seen that the estimates of the autocorrelation parameter θ are stable with respect to small lags (1 and 5), for which the autocorrelation function contain the most information about the dependence structure. However, the practical question "which lag t to use for estimation of θ?" remains open and due to further research. As an alternative, θ could be estimated by the least-squares method

$$\widetilde{\theta} = \min_{\theta > 0} \sum_{t=1}^{T-1} (\rho(t;\theta) - \widehat{\rho}(t))^2, \tag{4.4}$$

where $\rho(t;\theta) = e^{-\theta t}$ is the theoretical autocorrelation function and $\widehat{\rho}(t) = (\overline{m}_{1,1,c}(t) - \overline{m}^2_{1,c})/(\overline{m}_{2,c} - \overline{m}^2_{1,c})$ is its empirical counterpart. However, it is well known that due to the non-linearity this estimator is not consistent (see e.g. Ivanov (1997) or Malinvaud (1970)). The similar procedure was used in part II.A of Forman (2007).

5. Proofs

5.1. Proof of the Proposition 3.1

According to Proposition 3.1 in Kulik and Leonenko (2013), for every $\upsilon \in (-\alpha/2+1, \beta/2)$ the function $f_\upsilon(x) = x^\upsilon$ belongs to the domain of the extended generator $\mathcal{A}$ of the FSD X, and

$$\begin{aligned}\mathcal{A}f_\upsilon(x) &= -\theta\left(x - \frac{\varrho}{\beta-2}\right)\upsilon x^{\upsilon-1} + \frac{2\theta}{\beta-2}\left(x + \frac{\varrho}{\alpha}\right)\upsilon(\upsilon-1)x^{\upsilon-1}\\ &= -\frac{\upsilon\theta(\beta-2\upsilon)}{\beta-2}f_\upsilon(x) + \frac{\upsilon\theta\varrho(\alpha+2\upsilon-2)}{\alpha(\beta-2)}f_{\upsilon-1}(x),\end{aligned}$$

with $\varrho = (\beta-2)\kappa$. This, by definition, means that for every $x \in (0,\infty)$ the process

$$f(X_t) - \int_0^t \mathcal{A}f(X_s)\,ds$$

is a P_x-martingale with respect to the natural filtration for the FSD X. Then for every bounded function g we have

$$Ef_\upsilon(X_t^{st})g(X_0^{st}) = Ef_\upsilon(X_0^{st})g(X_0^{st}) + \int_0^t E\mathcal{A}f_\upsilon(X_s^{st})g(X_0^{st})\,ds.$$

From the relations above we have

$$E_{\upsilon,g}(t) = \langle f_\upsilon g\rangle - \frac{\upsilon\theta(\beta-2\upsilon)}{\beta-2}\int_0^t E_{\upsilon,g}(s)\,ds + \frac{\upsilon\theta\varrho(\alpha+2\upsilon-2)}{\alpha(\beta-2)}\int_0^t E_{\upsilon-1,g}(s)\,ds.$$

Considering this identity as an equation of $E_{\upsilon,g}(t)$, by solving it we obtain

$$E_{\upsilon,g}(t) = e^{-\frac{\upsilon\theta(\beta-2\upsilon)}{\beta-2}t}\langle f_\upsilon g\rangle + \frac{\upsilon\theta\varrho(\alpha+2\upsilon-2)}{\alpha(\beta-2)}\int_0^t e^{\frac{\upsilon\theta(\beta-2\upsilon)}{\beta-2}(s-t)}E_{\upsilon-1,g}(s)\,ds. \tag{5.1}$$

When $\upsilon \in \mathbb{N}$, relation (5.1) allows expressing $E_{\upsilon,g}(t)$ in the terms of $\langle g\rangle, \langle f_1 g\rangle, \dots, \langle f_\upsilon g\rangle$. Indeed, $E_{0,g}(t) \equiv \langle g\rangle$. Then using (5.1) iteratively with $\upsilon = 1,\dots,\upsilon$ gives the required expression for $E_{\upsilon,g}(t)$. Note that by the monotone convergence theorem the same relations hold true for a non-negative g such that

$$\langle g\rangle < +\infty, \quad \langle f_\upsilon g\rangle < +\infty. \tag{5.2}$$

Finally, from (2.6) we see that

$$\langle f_{\upsilon-1}\rangle = \frac{(\beta-2\upsilon)\alpha}{(\alpha+2\upsilon-2)\varrho}\langle f_\upsilon\rangle, \tag{5.3}$$

and therefore

$$\langle f_\upsilon\rangle\langle g\rangle = e^{-\frac{\upsilon\theta(\beta-2\upsilon)}{\beta-2}t}\langle f_\upsilon\rangle\langle g\rangle + \frac{\upsilon\theta\varrho(\alpha+2\upsilon-2)}{\alpha(\beta-2)}\left(\int_0^t e^{\frac{\upsilon\theta(\beta-2\upsilon)}{\beta-2}(s-t)}\,ds\right)\langle f_{\upsilon-1}\rangle\langle g\rangle.$$

Combined with (5.1), this gives the relation

$$C_{\upsilon,g}(t) = e^{-\frac{\upsilon\theta(\beta-2\upsilon)}{\beta-2}t}C_{\upsilon,g}(0) + \frac{\upsilon\theta\varrho(\alpha+2\upsilon-2)}{\alpha(\beta-2)}\int_0^t e^{\frac{\upsilon\theta(\beta-2\upsilon)}{\beta-2}(s-t)}C_{\upsilon-1,g}(s)\,ds.$$

In particular, for $\upsilon = 0$, $f_\upsilon(x) = 1$ and therefore

$$C_{0,g}(t) = \mathrm{Cov}\left(1, g\big(X_0^{st}\big)\right) = 0,$$

for all $t \geq 0$ and all non-negative functions g such that $\langle g\rangle < \infty$.

5.2. Proof of the Proposition 3.2

First, consider the case $t \geq s \geq 0$. Write

$$\begin{aligned} C_\upsilon(t,s) &= EX_t^{st}X_s^{st}\left(X_0^{st}\right)^\upsilon - \left(EX_t^{st}X_s^{st}\right)\langle f_\upsilon\rangle \\ &= \mathrm{Cov}\left(X_t^{st}, X_s^{st}\left(X_0^{st}\right)^\upsilon\right) + \left(EX_s^{st}\left(X_0^{st}\right)^\upsilon\right)\langle f_1\rangle - \left(EX_{t-s}^{st}X_0^{st}\right)\langle f_\upsilon\rangle \\ &= \mathrm{Cov}\left(X_t^{st}, X_s^{st}\left(X_0^{st}\right)^\upsilon\right) + C_{1,\upsilon}(s)\langle f_1\rangle - C_{1,1}(t-s)\langle f_\upsilon\rangle. \end{aligned} \tag{5.4}$$

The same argument that leads to (3.8) yields

$$\mathrm{Cov}\left(X_t^{st}, X_s^{st}\left(X_0^{st}\right)^\upsilon\right) = e^{-\theta(t-s)}\mathrm{Cov}\left(X_s^{st}, X_s^{st}\left(X_0^{st}\right)^\upsilon\right) \tag{5.5}$$

under the assumption

$$\langle f_\upsilon\rangle < \infty, \quad \langle f_{\upsilon+2}\rangle < \infty.$$

From (5.4) we get

$$\mathrm{Cov}\left(X_s^{st}, X_s^{st}\left(X_0^{st}\right)^\upsilon\right) = C_\upsilon(s,s) - C_{1,\upsilon}(s)\langle f_1\rangle + C_{1,1}(0)\langle f_\upsilon\rangle.$$

Clearly, $C_\upsilon(s,s) = C_{2,\upsilon}(s)$. Finally, we obtain

$$C_\upsilon(t,s) = e^{-\theta(t-s)}C_{2,\upsilon}(s) + \left[1 - e^{-\theta(t-s)}\right]C_{1,\upsilon}(s)\langle f_1\rangle,$$

where we take into account that

$$e^{-\theta(t-s)}C_{1,1}(0) = C_{1,1}(t-s),$$

see (3.8).

When $t \geq 0, s < 0$ relation (5.4), still holds true, but instead of (5.5) we have

$$\mathrm{Cov}\left(X_t^{st}, X_s^{st}\left(X_0^{st}\right)^\upsilon\right) = e^{-\theta t}\mathrm{Cov}\left(X_0^{st}, X_s^{st}\left(X_0^{st}\right)^\upsilon\right), \tag{5.6}$$

where

$$\begin{aligned} \mathrm{Cov}\left(X_0^{st}, X_s^{st}\left(X_0^{st}\right)^\upsilon\right) &= EX_s^{st}\left(X_0^{st}\right)^{\upsilon+1} - EX_s^{st}\left(X_0^{st}\right)^\upsilon\langle f_1\rangle \\ &= C_{1,\upsilon+1}(s) - C_{1,\upsilon}(s)\langle f_1\rangle + \langle f_1\rangle\langle f_{\upsilon+1}\rangle - \langle f_\upsilon\rangle\langle f_1\rangle^2 \\ &= C_{1,\upsilon+1}(s) - C_{1,\upsilon}(s)\langle f_1\rangle + C_{1,\upsilon}(0)\langle f_1\rangle. \end{aligned}$$

Then, using (3.8) we obtain

$$\begin{aligned} C_\upsilon(t,s) &= e^{-\theta t}\left[C_{1,\upsilon+1}(s) - C_{1,\upsilon}(s)\langle f_1\rangle + C_{1,\upsilon}(0)\langle f_1\rangle\right] + C_{1,\upsilon}(s)\langle f_1\rangle - C_{1,1}(t-s)\langle f_\upsilon\rangle \\ &= e^{-\theta t+\theta s}C_{1,\upsilon+1}(0) + \left[e^{-\theta t} + e^{\theta s} - e^{-\theta t+\theta s}\right]C_{1,\upsilon}(0)\langle f_1\rangle - e^{-\theta t+\theta s}C_{1,1}(0)\langle f_\upsilon\rangle. \end{aligned}$$

5.3. Proof of the Proposition 3.3

$$\begin{aligned} C(t_1,\ldots,t_4) = \mathrm{Cov}\left(X_{t_1}^{st}X_{t_2}^{st},\ X_{t_3}^{st}X_{t_4}^{st}\right) &= EX_{t_1}^{st}X_{t_2}^{st}X_{t_3}^{st}X_{t_4}^{st} - \left(C_{1,1}(t_2-t_1) + \langle f_1\rangle^2\right)\left(C_{1,1}(t_4-t_3) + \langle f_1\rangle^2\right) \\ &= EX_{t_{[1]}}^{st}X_{t_{[2]}}^{st}X_{t_{[3]}}^{st}X_{t_{[4]}}^{st} - \left(C_{1,1}(t_2-t_1) + \langle f_1\rangle^2\right)\left(C_{1,1}(t_4-t_3) + \langle f_1\rangle^2\right). \end{aligned}$$

An explicit calculation gives

$$EX_{t_{[1]}}^{st}X_{t_{[2]}}^{st}X_{t_{[3]}}^{st}X_{t_{[4]}}^{st} = \mathrm{Cov}\left(X_{t_{[1]}}^{st},\ X_{t_{[2]}}^{st}X_{t_{[3]}}^{st}X_{t_{[4]}}^{st}\right) + \langle f_1\rangle\mathrm{Cov}\left(X_{t_{[2]}}^{st},\ X_{t_{[3]}}^{st}X_{t_{[4]}}^{st}\right) + \langle f_1\rangle^2\mathrm{Cov}\left(X_{t_{[3]}}^{st}, X_{t_{[4]}}^{st}\right) + \langle f_1\rangle^4,$$

and therefore

$$\begin{aligned} C(t_1,\ldots,t_4) &= \mathrm{Cov}\left(X_{t_{[1]}}^{st},\ X_{t_{[2]}}^{st}X_{t_{[3]}}^{st}X_{t_{[4]}}^{st}\right) - C_{1,1}(t_2-t_1)C_{1,1}(t_4-t_3) + \langle f_1\rangle\mathrm{Cov}\left(X_{t_{[2]}}^{st},\ X_{t_{[3]}}^{st}X_{t_{[4]}}^{st}\right) \\ &\quad + \langle f_1\rangle^2\left[C_{1,1}(t_{[4]}-t_{[3]}) - C_{1,1}(t_2-t_1) - C_{1,1}(t_4-t_3)\right]. \end{aligned}$$

The same argument that leads to (3.8) yields

$$\mathrm{Cov}\left(X^{st}_{t_{[2]}},\ X^{st}_{t_{[3]}}X^{st}_{t_{[4]}}\right) = e^{-\theta(t_{[3]}-t_{[2]})}\mathrm{Cov}\left(X^{st}_{t_{[3]}},\ X^{st}_{t_{[3]}}X^{st}_{t_{[4]}}\right),$$

which after straightforward transformations gives

$$\mathrm{Cov}\left(X^{st}_{t_{[2]}},\ X^{st}_{t_{[3]}}X^{st}_{t_{[4]}}\right) = e^{-\theta(t_{[3]}-t_{[2]})}\Big[C_{2,1}(t_{[4]}-t_{[3]}) - \langle f_1\rangle C_{1,1}(t_{[4]}-t_{[3]}) + \langle f_1\rangle C_{1,1}(0)\Big].$$

Similarly,

$$\begin{aligned}\mathrm{Cov}\left(X^{st}_{t_{[1]}},\ X^{st}_{t_{[2]}}X^{st}_{t_{[3]}}X^{st}_{t_{[4]}}\right) = e^{-\theta(t_{[2]}-t_{[1]})}\Big[&C_2(t_{[4]}-t_{[2]}, t_{[3]}-t_{[2]})\\ &- \langle f_1\rangle C_1(t_{[4]}-t_{[2]}, t_{[3]}-t_{[2]}) + C_{1,1}(0)\big(C_{1,1}(t_{[4]}-t_{[3]}) + \langle f_1\rangle^2\big)\Big].\end{aligned}$$

Hence $C(t_1,\dots,t_4)$ can be expressed explicitly in the terms of the functions $C_{v,\chi}(t)$ and $C_v(t,s)$.

5.4. Proof of the Theorem 3.1

Statement 1. We restrict ourselves to the case of centered f,g. If f,g are supported on some segment $[u,v]\subset(0,\infty)$, then $f,g\in L_2(\pi)$ and (3.10) holds true according to L_2-bounds for α-mixing processes (see e.g. Hall and Heyde (1980)) since the stationary Fisher-Snedecor diffusion is α-mixing with an exponentially decaying rate (see Avram et al. (2011)). The following approximation procedure, similar to the one from the proof of Theorem 3.4 in Kulik and Leonenko (2013), extends (3.10) to the general case.

Let $\phi_{1,2}$ be defined as in the Theorem Appendix B.1 with $\gamma_{1,2},\delta_{1,2}$ satisfying (3.11). According to the Proposition 3.2. from Kulik and Leonenko (2013), there exist non-negative functions $\psi_{1,2}$ from the domain of the extended generator $\mathcal{A}$, satisfying

$$\mathcal{A}\psi_{1,2} \le c_{1,2}\psi_{1,2} + C_{1,2}\mathrm{I}_{[u,v]}(x)$$

with positive constants $c_{1,2}$ and $C_{1,2}$, and such that

$$\mathcal{A}\psi_{1,2} \le c'_{1,2}\psi_{1,2}^{1+\varepsilon} + C'_{1,2}$$

with positive constants $c'_{1,2}$ and $C'_{1,2}$ from Theorem 3.2 in Kulik and Leonenko (2013). By the condition (3.13), respective parameters $\gamma'_{1,2},\delta'_{1,2}$ in the construction of $\psi_{1,2}$ can be chosen in such a way that

$$\gamma'_1+\gamma_2<\frac{\alpha}{2},\quad \gamma'_2+\gamma_1<\frac{\alpha}{2},\quad \delta'_1+\delta_2<\frac{\beta}{2},\quad \delta'_2+\delta_1<\frac{\beta}{2}. \tag{5.7}$$

For centered f,g we have

$$\int_{-T}^{T}\mathrm{Cov}\left(f(X^{st}_t), g(X^{st}_0)\right)dt = \int_0^T\left(\int_{\mathbb{X}} gT_tf\,d\pi\right)dt + \int_0^T\left(\int_{\mathbb{X}} fT_tg\,d\pi\right)dt = \varrho_T(f,g)+\varrho_T(g,f),$$

where

$$\varrho_T(f,g) = \int_{\mathbb{X}} g(x)\left(\int_{\mathbb{X}} f(y)\left(\int_0^T\Big((\delta_x)_t(dy)-\pi(dy)\Big)dt\right)\right)\pi(dx),$$

$$\mu_t(dy) = \int_{(0,\infty)} P_t(x,dy)\mu(dx).$$

Here we used the standard notation

$$T_tf(x) = \int\limits_{\mathbb{X}} f(y)P_t(x,dy),\quad P_t(x,\cdot) = (\delta_x)_t.$$

Then we have by Theorem 3.2 in Kulik and Leonenko (2013)

$$|\varrho_T(f,g)| \le C_2\|f\|_{\phi_1}\|g\|_{\phi_2}C^2\int_{\mathbb{X}}\phi_2(x)\psi_1(x)\pi(dx),$$

where $\|f\|_{\phi_1} = \sup\limits_{x=(x_1,\dots,x_r)}\frac{|f(x)|}{\sum_{i=1}^r\phi_1(x_i)}$, analogously for $\|g\|_{\phi_2}$. According to the procedure in Section 5.6 in Kulik and Leonenko (2013), pp. 22-23, one can construct centered compactly supported functions $f_n, g_n, n\ge 1$, such that

$$\|f-f_n\|_{\phi_1}\to 0,\quad \|g-g_n\|_{\phi_2}\to 0,\quad n\to\infty, \tag{5.8}$$

$$f_n(x) = f(x)\prod_{i=1}^{r} \mathrm{I}_{\{x_i \geq 1/n\}},$$

and analogously for functions g_n.

Note that (5.7) provides that the function $\phi_2\psi_1$ is integrable with respect to the invariant density π. For any $T > 0$ we have the following upper bound for $|\varrho_T(f,g) - \varrho_T(f_n,g_n)|$:

$$\begin{aligned} |\varrho_T(f,g) - \varrho_T(f_n,g_n)| &\leq |\varrho_T(f,g) - \varrho_T(f,g_n)| + |\varrho_T(f,g_n) - \varrho_T(f_n,g_n)| \\ &\leq C_2\Big(\|f-f_n\|_{\phi_1}\|g_n\|_{\phi_2} + \|f\|_{\phi_1}\|g-g_n\|_{\phi_2}\Big)\int_{\mathbb{X}} \phi_2(x)\psi_1(x)\pi(dx). \end{aligned} \tag{5.9}$$

Combined with similar estimate for $\varrho_T(g,f)$, this yields that the limit (3.10) exists and

$$\sigma_{f,g} = \lim_{n\to\infty} \sigma_{f_n,g_n}. \tag{5.10}$$

Statement 2. Consider first the case when f, g are compactly supported. We have

$$\sigma_{f,g} = \int_0^\infty \int_{\mathbb{X}} \Big(f(x)T_t g(x) + g(x)T_t f(x)\Big)\,\pi(dx)dt.$$

Because f, g are compactly supported, we can write $|f| \leq C_f\phi, |g| \leq C_g\phi$, where ϕ is defined as in the Theorem Appendix B.1 with γ, δ satisfying

$$\gamma < \frac{\alpha}{2} - 1, \quad \delta < \frac{\beta}{2}. \tag{5.11}$$

Then the Theorem Appendix B.1 (Theorem 3.1 in Kulik and Leonenko (2013)) with $\mu = \delta_x$ yields

$$|T_t f(x)| \leq Ce^{-ct}\phi(x), \quad |T_t g(x)| \leq Ce^{-ct}\phi(x).$$

Then according to the Lebesgue's dominated convergence theorem we obtain

$$\begin{aligned} \sigma_{f,g} &= \lim_{s\to 0+} \int_0^\infty \int_{\mathbb{X}} e^{-st}\Big(f(x)T_t g(x) + g(x)T_t f(x)\Big)\,\pi(dx)dt \\ &= \lim_{s\to 0+} \int_{\mathbb{X}} \Big(f(x)R_s g(x) + g(x)R_s f(x)\Big)\,\pi(dx) \\ &= \lim_{s\to 0+} \int_{\mathbb{X}}\int_{\mathbb{X}} G_s(x,y)\Big(f(x)g(y) + f(y)g(x)\Big)\,\pi(dx)\,dy, \end{aligned} \tag{5.12}$$

where $R_s = \int_0^\infty e^{-st}T_t\,dt$ is the resolvent operator, and G_s is the respective resolvent kernel (see e.g. Karlin and Taylor (1981)). The explicit formula for G_s is

$$G_s(x,y) = C_s f_1(x\wedge y, s) f_4(x\vee y, s)\,\mathfrak{p}(y), \tag{5.13}$$

with

$$f_1(x,s) = {}_2F_1\left(z_{+,s}, z_{-,s}; \frac{\alpha}{2}; -\frac{\alpha}{\varrho}x\right),$$

$$f_4(x,s) = \left(\frac{\alpha}{\varrho}x\right)^{-z_{+,s}} {}_2F_1\left(z_{+,s}, u_{+,s}; 1+2\Delta_s; -\frac{\varrho}{\alpha x}\right).$$

being the classical Gauss hypergeometric functions with parameters

$$\Delta_s = \sqrt{\frac{\beta^2}{16} + \frac{s(\beta-2)}{2\theta}}, \quad z_{\pm,s} = -\frac{\beta}{4} \pm \Delta_s, \quad u_{\pm,s} = 1 - \frac{\alpha}{2} + z_{\pm,s}$$

related to continuous part of the spectrum of the infinitesimal generator of FSD and involved in the spectral representation of its transition density (see Appendix A). For more information we refer to Avram et al. (2013b). Straightforward calculation then gives that

$$z_{+,s} \to 0+, \quad C_s \sim \frac{(\beta-2)}{\theta\beta}\Gamma(z_{+,s}), \quad s \to 0+.$$

All the above and the following relations regarding the hypergeometric functions can be found (with much more details) in e.g. Nikiforov and Uvarov (1988) or Luke (1969). According to the relation

$$\begin{aligned} {}_2F_1\Big(a,b;c;z\Big) =& \frac{\Gamma(c)\Gamma(b-a)}{\Gamma(b)\Gamma(c-a)}(-z)^{-a}{}_2F_1\Big(a,a-c+1;a-b+1;\frac{1}{z}\Big) \\ &+ \frac{\Gamma(c)\Gamma(a-b)}{\Gamma(a)\Gamma(c-b)}(-z)^{-b}{}_2F_1\Big(b,b-c+1;b-a+1;\frac{1}{z}\Big), \quad |\arg(-z)| < \pi \end{aligned} \tag{5.14}$$

with

$$a = z_{+,s}, \quad b = u_{+,s} = 1 - \frac{\alpha}{2} + z_{+,s}, \quad c = 1 + 2\Delta_s = 1 + \frac{\beta}{2} + 2z_{+,s},$$

we obtain

$$f_4(x,s) = \frac{\Gamma(1+\beta/2+2z_{+,s})\Gamma(1-\alpha/2)}{\Gamma(1-\alpha/2+z_{+,s})\Gamma(1+\beta/2+z_{+,s})} f_1(x,s) + \frac{\Gamma(1+\beta/2+2z_{+,s})\Gamma(\alpha/2-1)}{\Gamma(z_{+,s})\Gamma(\alpha/2+\beta/2+z_{+,s})} f_2(x,s)$$

with

$$f_2(x,s) = \Big(\frac{\alpha}{\varrho}x\Big)^{1-\alpha/2}{}_2F_1\Big(u_{+,s}, u_{-,s}; 2-\frac{\alpha}{2}; -\frac{\alpha}{\varrho}x\Big).$$

This leads to the representation

$$G_s(x,y) = \Big[C_s^1 f_1(x\wedge y,s) f_1(x\vee y,s) + C_s^2 f_1(x\wedge y,s) f_2(x\vee y,s)\Big]\,\mathfrak{p}(y) \tag{5.15}$$

with

$$C_s^1 \sim \frac{(\beta-2)}{\theta\beta}\Gamma(z_{+,s}), \quad C_s^2 \to \frac{(\beta-2)}{\theta\beta}\frac{\Gamma(1+\beta/2)\Gamma(\alpha/2-1)}{\Gamma(\alpha/2+\beta/2)}, \quad s\to 0+.$$

Because $f_1(x\wedge y,s)f_1(x\vee y,s) = f_1(x,s)f_1(y,s)$, we have by (5.12) that $\sigma_{f,g} = \sigma^1_{f,g} + \sigma^2_{f,g}$ with

$$\sigma^1_{f,g} = \frac{2(\beta-2)}{\theta\beta}\lim_{s\to 0+}\Gamma(z_{+,s})\left(\int_0^\infty f(x)f_1(x,s)\,\mathfrak{p}(x)\,dx\right)\left(\int_0^\infty g(x)f_1(x,s)\,\mathfrak{p}(x)\,dx\right), \tag{5.16}$$

$$\begin{aligned} \sigma^2_{f,g} =& \frac{(\beta-2)}{\theta\beta}\frac{\Gamma(1+\beta/2)\Gamma(\alpha/2-1)}{\Gamma(\alpha/2+\beta/2)} \\ &\times \lim_{s\to 0+}\int_0^\infty\int_0^\infty \Big(f(x)g(y)+f(y)g(x)\Big) f_1(x\wedge y,s) f_2(x\vee y,s)\,\mathfrak{p}(x)\,\mathfrak{p}(y)\,dxdy. \end{aligned} \tag{5.17}$$

Existence of the limits in (5.16), (5.17) is provided by the following lemma.

Lemma 5.1. *Let $[u,v]\subset(0,\infty)$ be an arbitrary segment. Then*

(i) uniformly for $x\in[u,v]$,

$$f_2(x,s)\to\chi(x) = \Big(\frac{\alpha}{\varrho}x\Big)^{1-\alpha/2}{}_2F_1\Big(1-\frac{\alpha}{2}, 1-\frac{\alpha}{2}-\frac{\beta}{2}; 2-\frac{\alpha}{2}; -\frac{\alpha}{\varrho}x\Big), \quad s\to 0+;$$

(ii) uniformly for $x\in[u,v]$,

$$\frac{f_1(x,s)-1}{z_{+,s}} \to \Psi(x), \quad s\to 0+,$$

where Ψ is some locally bounded function.

Proof. When $[u,v]$ is located close to 0 the required statements follow directly from the definition of the Gauss hypergeometric function for the complex-valued z with $|z|<1$. If the segment $[u,v]$ is located near other two singular points of the Gauss hypergeometric function, i.e. near 1 and ∞, we must use the relation providing the analytic continuation of this function to the respective parts of the complex plane.

When $[u,v]$ is located close to ∞, one can use the functional relation (5.14) with $z'' = 1/z$ and then the definition of the Gauss hypergeometric function.

At last, when $[u,v]$ is located near ϱ/α, the arguments of functions f_1 and f_2 are located near (-1) and so $|1-(\alpha/\varrho)x| = |1-z|$ is located near zero, and therefore the relation

$$\begin{aligned} {}_2F_1\Big(a,b;c;1-z\Big) =& \frac{\Gamma(c)\Gamma(c-b-a)}{\Gamma(c-a)\Gamma(c-b)}{}_2F_1\Big(a,b;a+b-c+1;z\Big) + \\ &+ z^{c-a-b}\frac{\Gamma(c)\Gamma(a+b-c)}{\Gamma(a)\Gamma(b)}{}_2F_1\Big(c-a,c-b;c-a-b+1;z\Big), \quad |\arg(z)|<\pi \end{aligned} \tag{5.18}$$

yields the proofs of the statements of the Lemma. Specifically, when $x = \varrho/\alpha$, i.e. the arguments of functions f_1 and f_2 are both (-1), according to Slater (1966), page 5, function f_2 is well-defined, while f_1 is well-defined when $\alpha/2 - 2\Delta_s > 0$ so we restrict our considerations according to this condition. □

Now, by the means provided by the Lemma 5.1 we can continue with the proof of the *Statement 2* from the Theorem 3.1. Lemma 5.1 provides the following explicit formulae for covariances $\sigma^{1,2}_{f,g}$:

$$\sigma^1_{f,g} = 0,$$

$$\sigma^2_{f,g} = \frac{(\beta-2)}{\theta\beta}\frac{\Gamma(1+\beta/2)\Gamma(\alpha/2-1)}{\Gamma(\alpha/2+\beta/2)}\int_0^\infty\int_0^\infty \Big(f(x)g(y)+f(y)g(x)\Big)\chi(x\vee y)\,\mathfrak{p}(x)\,\mathfrak{p}(y)\,dxdy.$$

The second relation follows from Lemma 5.1 directly because f and g are compactly supported. To verify the first relation, we note that statement (ii) of Lemma 5.1 provides

$$\begin{aligned}\int_0^\infty f(x)f_1(x,s)\,\mathfrak{s}(x)\,dx &= \int_0^\infty f(x)\,\mathfrak{p}(x)\,dx + \int_0^\infty f(x)(f_1(x,s)-1)\,\mathfrak{p}(x)\,dx\\ &\sim z_{+,s}\int_0^\infty f(x)\Psi(x)\,\mathfrak{p}(x)\,dx, \quad s\to 0+,\end{aligned}$$

where the first term vanishes because f is centered. Together with the same relation for g this yields the required identity, since

$$\lim_{s\to 0+} z^2_{+,s}\Gamma(z_{+,s}) = 0.$$

Hence, after the transformations $\Gamma(\beta/2+1) = (\beta/2)\Gamma(\beta/2), \Gamma(\alpha/2) = (\alpha/2-1)\Gamma(\alpha/2-1)$ we get

$$\begin{aligned}\sigma_{f,g} = &\frac{(\beta-2)}{\theta(\alpha-2)B(\alpha/2,\beta/2)}\int_0^\infty\int_0^\infty \Big(f(x)g(y)+f(y)g(x)\Big)\\ &\times\left(\frac{\chi(x\vee y)}{xy}\right)\left(\frac{\alpha x}{\alpha x+\varrho}\right)^{\alpha/2}\left(\frac{\varrho}{\alpha x+\varrho}\right)^{\beta/2}\left(\frac{\alpha y}{\alpha y+\varrho}\right)^{\alpha/2}\left(\frac{\varrho}{\alpha y+\varrho}\right)^{\beta/2} dxdy.\end{aligned} \tag{5.19}$$

Now changing the variables

$$u = \frac{\varrho}{\alpha x+\varrho} \Leftrightarrow x = \frac{\varrho(1-u)}{\alpha u},$$

$$v = \frac{\varrho}{\alpha y+\varrho} \Leftrightarrow y = \frac{\varrho(1-v)}{\alpha v},$$

and noting that

$$\frac{\alpha}{\sigma}(x\vee y) = \left(\frac{1}{u}-1\right)\vee\left(\frac{1}{v}-1\right) = \frac{1}{u\wedge v}-1,$$

transforms the integral on the right hand side of (5.19) as follows:

$$\int_0^1\int_0^1 \Big(\hat f(u)\hat g(v)+\hat f(v)\hat g(u)\Big)Q(u\wedge v)u^{\beta/2-1}(1-u)^{\alpha/2-1}v^{\beta/2-1}(1-v)^{\alpha/2-1}\,dudv, \tag{5.20}$$

with

$$\hat f(z) = f\left(\frac{\varrho(1-z)}{\alpha z}\right),\quad \hat g(z) = g\left(\frac{\varrho(1-z)}{\alpha z}\right),\quad z\in(0,1),$$

$$Q(z) = \left(\frac{1}{z}-1\right)^{1-\alpha/2} {}_2F_1\Big(1-\frac{\alpha}{2}, 1-\frac{\alpha}{2}-\frac{\beta}{2}; 2-\frac{\alpha}{2}; 1-\frac{1}{z}\Big),\quad z\in(0,1).$$

To simplify the expression for $Q(z), z\in(0,1)$, we use the relation

$$\begin{aligned}{}_2F_1(a,b;c;z) =&(1-z)^{-a}\frac{\Gamma(c)\Gamma(b-a)}{\Gamma(c-a)\Gamma(b)}{}_2F_1\left(a, c-b; 1+a-b; \frac{1}{1-z}\right)\\ &+(1-z)^{-b}\frac{\Gamma(c)\Gamma(a-b)}{\Gamma(c-b)\Gamma(a)}{}_2F_1\left(c-a, b; 1-a+b; \frac{1}{1-z}\right)\end{aligned} \tag{5.21}$$

valid for $|\arg(-z)| < \pi, |\arg(1-z)| < \pi$ with

$$a = 1-\frac{\alpha}{2},\quad b = 1-\frac{\alpha}{2}-\frac{\beta}{2},\quad c = 2-\frac{\alpha}{2},$$

and $z' = 1 - 1/z$ (note that for $z \in (0,1)$ one has $\arg(-z') = \arg(1-z') = 0$, therefore (5.21) is applicable). Since $1/(1-z') = z$, (5.21) provides

$$\begin{aligned} Q(z) &= \left(\frac{1}{z}-1\right)^{1-\alpha/2} z^{1-\alpha/2} \frac{\Gamma(2-\alpha/2)\Gamma(-\beta/2)}{\Gamma(1)\Gamma(1-\alpha/2-\beta/2)} {}_2F_1\left(1-\frac{\alpha}{2}, 1+\frac{\beta}{2}; 1+\frac{\beta}{2}; z\right) \\ &+ \left(\frac{1}{z}-1\right)^{1-\alpha/2} z^{1-\alpha/2-\beta/2} \frac{\Gamma(2-\alpha/2)\Gamma(\beta/2)}{\Gamma(1+\beta/2)\Gamma(1-\alpha/2)} {}_2F_1\left(1, 1-\frac{\alpha}{2}-\frac{\beta}{2}; 1-\frac{\beta}{2}; z\right) \\ &=: Q_1(z) + Q_2(z). \end{aligned}$$

We proceed with further transformations separately for Q_1 and Q_2. From the symmetry of the Gauss hypergeometric functions in its first two parameters, i.e. from ${}_2F_1(a,b;c;z) = {}_2F_1(b,a;c;z)$, it follows that

$$Q_1(z) = \left(\frac{1}{z}-1\right)^{1-\alpha/2} z^{1-\alpha/2} \frac{\Gamma(2-\alpha/2)\Gamma(-\beta/2)}{\Gamma(1)\Gamma(1-\alpha/2-\beta/2)} (1-z)^{-1+\alpha/2} = \frac{\Gamma(2-\alpha/2)\Gamma(-\beta/2)}{\Gamma(1-\alpha/2-\beta/2)}.$$

Since f and g are supposed to be centered, we have

$$\int_0^1 \hat{f}(u) u^{\beta/2-1}(1-u)^{\alpha/2-1}\, du = \int_0^1 \hat{g}(u) u^{\beta/2-1}(1-u)^{\alpha/2-1}\, du = 0$$

providing that this part of Q is negligible in the integral (5.20). Therefore,

$$\begin{aligned} &\int_0^1 \int_0^1 \Big(\hat{f}(u)\hat{g}(v) + \hat{f}(v)\hat{g}(u)\Big) Q(u \wedge v) u^{\beta/2-1}(1-u)^{\alpha/2-1} v^{\beta/2-1}(1-v)^{\alpha/2-1}\, dudv \\ &\quad = \frac{2\Gamma(2-\alpha/2)\Gamma(-\beta/2)}{\Gamma(1-\alpha/2-\beta/2)} \left(\int_0^1 \hat{f}(u) u^{\beta/2-1}(1-u)^{\alpha/2-1}\, du\right) \\ &\qquad \times \left(\int_0^1 \hat{g}(u) u^{\beta/2-1}(1-u)^{\alpha/2-1}\, du\right) = 0. \end{aligned}$$

Next, using the relation ${}_2F_1(a,b;c;z) = (1-z)^{c-a-b} {}_2F_1(c-a, c-b; c; z)$ we obtain

$$\begin{aligned} Q_2(z) &= \left(\frac{1}{z}-1\right)^{1-\alpha/2} z^{1-\alpha/2-\beta/2} (1-z)^{\alpha/2-1} \frac{\Gamma(2-\frac{\alpha}{2})\Gamma(\frac{\beta}{2})}{\Gamma(1+\frac{\beta}{2})\Gamma(1-\frac{\alpha}{2})} {}_2F_1\left(-\frac{\beta}{2}, \frac{\alpha}{2}; 1-\frac{\beta}{2}; z\right) \\ &= \frac{1-\alpha/2}{\beta/2} z^{-\beta/2} {}_2F_1\left(-\frac{\beta}{2}, \frac{\alpha}{2}; 1-\frac{\beta}{2}; z\right) = (\alpha/2-1) B\left(-\frac{\beta}{2}, 1-\frac{\alpha}{2}; z\right), \end{aligned}$$

where $B(a,b;z) = \frac{z^a}{a} {}_2F_1(a, 1-b; a+1; z)$ is the incomplete Beta function $B(a,b;z)$ (see e.g. Nikiforov and Uvarov (1988)). Summarizing all the above, we get the formula for covariances

$$\begin{aligned} \sigma_{f,g} &= \frac{(\beta-2)}{2\theta B(\alpha/2, \beta/2)} \int_0^1 \int_0^1 \Big(\hat{f}(u)\hat{g}(v) + \hat{f}(v)\hat{g}(u)\Big) \\ &\quad \times B\left(-\frac{\beta}{2}, 1-\frac{\alpha}{2}; u \wedge v\right) u^{\beta/2-1}(1-u)^{\alpha/2-1} v^{\beta/2-1}(1-v)^{\alpha/2-1}\, dudv, \end{aligned} \tag{5.22}$$

valid for any centered f, g having a compact support in $(0, \infty)$. Denote

$$F(u) = \int_0^u f\left(\frac{\varrho(1-v)}{\alpha v}\right) v^{\beta/2-1}(1-v)^{\alpha/2-1}\, dv, \quad G(u) = \int_0^u g\left(\frac{\varrho(1-v)}{\alpha v}\right) v^{\beta/2-1}(1-v)^{\alpha/2-1}\, dv, \tag{5.23}$$

and note that F, G vanish at some neighborhoods of the points 0 and 1 since f, g are centered and compactly supported. Then, using the integration-by-parts formula, we can write the integral in (5.22) as follows:

$$\int_0^1\int_0^1 \Big(F'(u)G'(v)+F'(u)G'(v)\Big)B\left(-\frac{\beta}{2},1-\frac{\alpha}{2};u\wedge v\right)\,dudv$$
$$=2\int_0^1 F'(u)\left(\int_0^u G'(v)B\left(-\frac{\beta}{2},1-\frac{\alpha}{2};v\right)\,dv\right)\,du$$
$$+\,2\int_0^1 G'(u)\left(\int_0^u F'(v)B\left(-\frac{\beta}{2},1-\frac{\alpha}{2};v\right)\,dv\right)\,du$$
$$=-2\int_0^1 \Big(F(u)G'(u)+F'(u)G(u)\Big)B\left(-\frac{\beta}{2},1-\frac{\alpha}{2};v\right)\,du$$
$$=2\int_0^1 F(u)G(u)\left[B\left(-\frac{\beta}{2},1-\frac{\alpha}{2};v\right)\right]'\,du=2\int_0^1 F(u)G(u)u^{-\beta/2-1}(1-u)^{-\alpha/2}\,du,$$

where, in the last identity, we have used the relation $\frac{d}{dz}B\left(a,b;z\right)=z^{a-1}(1-z)^{b-1}$ for the derivative of the incomplete Beta function. From this representation of the integral from (5.22) we finally obtain (3.14), under the additional assumption that f,g are compactly supported.

To remove this assumption, we use an approximation procedure, similar to the one from the proof of Theorem 3.4 in Kulik and Leonenko (2013). Let $\phi_{1,2}$ be defined as in the Theorem Appendix B.1 with $\gamma_{1,2},\delta_{1,2}$ from the condition (3.11). Consider sequences of centered compactly supported functions $f_n,g_n,n\geq 1$ such that (5.8) holds true. Then respective functions F_n,G_n converge to F,G point-wise, and there exist constants $\varepsilon\in(0,1)$ and $C>0$ such that

$$|F_n(u)|\leq Cu^{\beta/2-\delta_1},\quad |G_n(u)|\leq Cu^{\beta/2-\delta_2},\quad u\in(0,\varepsilon),\quad n\geq 1,$$

for u close enough to 0, and

$$|F_n(u)|\leq C(1-u)^{\alpha/2-\gamma_1},\quad |G_n(u)|\leq C(1-u)^{\alpha/2-\gamma_2},\quad u\in(0,1-\varepsilon),\quad n\geq 1.$$

Under condition (3.13) this provides

$$\int_0^1 F_n(u)G_n(u)u^{-\beta/2-1}(1-u)^{-\alpha/2}\,du\to\int_0^1 F(u)G(u)u^{-\beta/2-1}(1-u)^{-\alpha/2}\,du,\quad n\to\infty,$$

which together with (5.10) completes the proof.

5.5. Proof of the Theorem 3.2

(i) *P-consistency*

From the general expression (2.6) for moments of the FS distribution with the density (2.5), it follows that in particular moments of order −1, 1 and 2 are given by

$$m_{-1}=\frac{\alpha\beta}{(\alpha-2)(\beta-2)\kappa},\quad m_1=\kappa,\quad m_2=\frac{(\alpha+2)(\beta-2)\kappa^2}{\alpha(\beta-4)},\tag{5.24}$$

while

$$m_{1,1}(t)=\frac{2\kappa^2(\alpha+\beta-2)}{\alpha(\beta-4)}e^{-\theta t}+\kappa^2.$$

The assumptions $\alpha>2$ and $\beta>8$ in the statement of this theorem ensure that $\{-1,1,2\}\subset(-\alpha/4-1/2,\beta/4)$. Now from the LLN for the FSD with the arbitrary initial distribution (Theorem 3.4 from Kulik and Leonenko (2013), part 1, see also Appendix B), it follows that for every $t>0$

$$(\overline{m}_{-1,c},\overline{m}_{1,c},\overline{m}_{2,c},\overline{m}_{1,1,c}(t))\xrightarrow{P}(m_{-1},m_1,m_2,m_{1,1}(t)),\tag{5.25}$$

i.e. empirical moments are P-consistent estimators of the corresponding theoretical moments.

Furthermore, let $g\colon\mathbb{R}^4\to\mathbb{R}^4$ be a function defined by

$$g(x,y,z,w)=(g_1(x,y,z,w),\,g_2(x,y,z,w),\,g_3(x,y,z,w),\,g_4(x,y,z,w)),\tag{5.26}$$

where

$$g_1(x,y,z,w) = \frac{2(xyz-y^2)}{xyz-2z+y^2}, \qquad g_2(x,y,z,w) = \frac{4x(y^2-z)}{2xy^2-xz-y},$$
$$g_3(x,y,z,w) = y, \qquad g_4(x,y,z,w) = -\frac{1}{t}\log\left(\frac{w-y^2}{z-y^2}\right).$$

Clearly, for estimator (3.15), $(\widehat{\alpha}_c, \widehat{\beta}_c, \widehat{\kappa}_c, \widehat{\theta}_c(t)) = g(\overline{m}_{-1,c}, \overline{m}_{1,c}, \overline{m}_{2,c}, \overline{m}_{1,1,c}(t))$ and $(\alpha,\beta,\kappa,\theta) = g(m_{-1}, m_1, m_2, m_{1,1}(t))$.

The function g is well defined and smooth in the neighbourhood of the point $(m_{-1}, m_1, m_2, m_{1,1}(t))$, and therefore according to the continuity mapping theorem applied to the function g and result (5.25) it follows that for every $t > 0$

$$g(\overline{m}_{-1,c}, \overline{m}_{1,c}, \overline{m}_{2,c}, \overline{m}_{1,1,c}(t)) \xrightarrow{P} g(m_{-1}, m_1, m_2, m_{1,1}(t)),$$

i.e.

$$(\widehat{\alpha}_c, \widehat{\beta}_c, \widehat{\kappa}_c, \widehat{\theta}_c(t)) \xrightarrow{P} (\alpha,\beta,\kappa,\theta). \tag{5.27}$$

(ii) *Asymptotic normality with explicitly calculated asymptotic covariance matrix*

According to the CLT for the FSD with the arbitrary initial distribution (Theorem 3.4 from Kulik and Leonenko (2013), part 2, see also Appendix B), it follows that for the function $f\colon \mathbb{R}^2 \to \mathbb{R}^4$ defined by

$$f(x,y) = (x^{-1}, x, x^2, xy),$$

we have

$$\frac{1}{\sqrt{T}}\int_0^T \Big(f\big(X_s^{st}, X_{s+t}^{st}\big) - E\big[f\big(X_0^{st}, X_t^{st}\big)\big]\Big)\, ds \xrightarrow{D} \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}), \quad T\to\infty,$$

where

$$\mathbf{\Sigma}_{i,j} = \int_{-\infty}^{\infty} \mathrm{Cov}\left(f_i\left(X_s^{st}, X_{s+t}^{st}\right), f_j\left(X_0^{st}, X_t^{st}\right)\right) ds, \quad i,j \in \{1,2,3,4\}.$$

Hence,

$$\sqrt{T}\left(\overline{m}_{-1,c} - E[\overline{m}_{-1,c}], \overline{m}_{1,c} - E[\overline{m}_{1,c}], \overline{m}_{2,c} - E[\overline{m}_{2,c}], \overline{m}_{1,1,c}(t) - E[\overline{m}_{1,1c}(t)]\right) \xrightarrow{D} \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}), \quad T\to\infty$$

where

$$E[\overline{m}_{-1,c}] = E[1/X_0^{st}] = \frac{\alpha\beta}{(\alpha-2)(\beta-2)\kappa}, \qquad E[\overline{m}_{1,c}] = E[X_0^{st}] = \kappa,$$

$$E[\overline{m}_{2,c}] = E\left[\left(X_0^{st}\right)^2\right] = \frac{(\alpha+2)(\beta-2)\kappa^2}{\alpha(\beta-4)}, \qquad E[\overline{m}_{1,1,c}] = E[X_0^{st}X_t^{st}] = \kappa^2 + \frac{2\kappa^2(\alpha+\beta-2)}{\alpha(\beta-4)}e^{-\theta t}.$$

Therefore, $(\overline{m}_{-1,c}, \overline{m}_{1,c}, \overline{m}_{2,c}, \overline{m}_{1,1,c}(t))$ is asymptotically normal estimator of theoretical moments $(m_{-1}, m_1, m_2, m_{1,1}(t))$ of FSD invariant distribution. In the following part of the proof we provide the explicit expressions for the elements of the asymptotic covariance matrix $\mathbf{\Sigma} = \mathbf{\Sigma}(\alpha,\beta,\kappa,\theta)$, calculated by applying the evolutionary equations approach detailed in Section 3.1 together with the result of Theorem 3.1 (more precisely, Example 3.3). Covariances $\mathbf{\Sigma}_{i,j} = (\mathbf{\Sigma}(\alpha,\beta,\kappa,\theta))_{i,j}$ are given by following expressions:

$$\mathbf{\Sigma}_{11} = \frac{4\alpha^2(\alpha+\beta-2)}{\theta\kappa^2(\alpha-2)^3(\beta-2)}, \tag{5.28}$$

$$\mathbf{\Sigma}_{12} = -\frac{4(\alpha+\beta-2)}{\theta(\alpha-2)(\beta-2)} = \mathbf{\Sigma}_{21},$$

$$\mathbf{\Sigma}_{13} = -\frac{8\kappa(\alpha+\beta-2)(\alpha\beta+\beta-3\alpha-4)}{\theta\alpha(\alpha-2)(\beta-4)^2} = \mathbf{\Sigma}_{31},$$

$$\mathbf{\Sigma}_{14} = -\frac{4\kappa(\alpha+\beta-2)}{\theta(\alpha-2)(\beta-2)} - \frac{4\kappa(\alpha+\beta-2)\left(\alpha\left(\beta^2-2\beta-4\right)+2\left(\beta^2-6\beta+8\right)\right)}{\theta\alpha(\alpha-2)(\beta-2)(\beta-4)^2}e^{-\theta t} = \mathbf{\Sigma}_{41},$$

$$\mathbf{\Sigma}_{22} = \frac{4\kappa^2(\alpha+\beta-2)}{\theta\alpha(\beta-4)},$$

$$\mathbf{\Sigma}_{23} = \frac{8\kappa^3(\alpha+2)(\beta-2)(\alpha+\beta-2)}{\theta\alpha^2(\beta-4)(\beta-6)} = \mathbf{\Sigma}_{32},$$

$$\mathbf{\Sigma}_{24} = \frac{4\kappa^3(\alpha+\beta-2)}{\theta\alpha(\beta-4)} + \frac{4\kappa^3(\alpha+\beta-2)\left(\alpha\left(\beta+2\right)+4\left(\beta-2\right)\right)}{\theta\alpha^2(\beta-4)(\beta-6)}e^{-\theta t} = \mathbf{\Sigma}_{42},$$

$$\mathbf{\Sigma}_{33} = \frac{8\kappa^4(\alpha+2)(\beta-2)^2(\alpha+\beta-2)(\alpha(\beta-6)(2\beta-7)+(\beta-4)(5\beta-22))}{\theta\alpha^3(\beta-4)^3(\beta-6)(\beta-8)},$$

$$\begin{aligned}\mathbf{\Sigma}_{34} = {} & \frac{8\kappa^4(\alpha+2)(\beta-2)(\alpha+\beta-2)}{\theta\alpha^2(\beta-4)(\beta-6)} + \\ & \frac{8\kappa^4(\alpha+2)(\beta-2)(\alpha+\beta-2)\left(\alpha\left(\beta^3-7\beta^2+44\right)+5\beta^3-52\beta^2+172\beta-176\right)}{\theta\alpha^3(\beta-4)^3(\beta-6)(\beta-8)}e^{-\theta t},\end{aligned}$$

$$\begin{aligned}\mathbf{\Sigma}_{44} = {} & \frac{4\kappa^4(\alpha+\beta-2)(\alpha(4\beta-15)+\beta-2)}{\theta\alpha^2(\beta-4)^2} + \frac{12\kappa^4(\alpha+2)(\beta-2)(\alpha+\beta-2)}{\theta\alpha^2(\beta-4)(\beta-6)}e^{\theta s} + \\ & \frac{4\kappa^4(\alpha+2)(\beta-2)^4(\alpha+\beta-2)(\alpha+\beta-4)}{\theta\alpha^3(\beta-4)^2(\beta-6)^2(\beta-8)}e^{-2\theta\frac{(\beta-4)}{\beta-2}s} - \frac{8\kappa^4(\alpha+\beta-2)^2 s}{\alpha^2(\beta-4)^2}e^{-2\theta s} + \\ & \frac{4\kappa^4(\alpha+\beta-2)\left(16(\beta-2)^2+\alpha^2\left(-\left(3\beta^2-56\beta+2(\beta-10)(\beta-6)\theta s+164\right)\right)+2\alpha(\beta-2)(5\beta+2(\beta-6)\theta s+2)\right)}{\theta\alpha^3(\beta-4)(\beta-6)^2}e^{-\theta s} \\ & -\frac{4\kappa^4(\alpha+\beta-2)\left(\alpha^2(\beta((\beta-6)(\beta-1)\beta+72)-176)+\alpha(\beta((\beta-6)(\beta-1)\beta+72)-176)(\beta-2)+2(\beta-4)(\beta-2)^4\right)}{\theta\alpha^3(\beta-4)^3(\beta-6)(\beta-8)}e^{-2\theta s}.\end{aligned}$$

Asymptotic normality of the estimator $(\widehat{\alpha}_c, \widehat{\beta}_c, \widehat{\kappa}_c, \widehat{\theta}_c(t))$ now follows by applying the multivariate delta method (see e.g. Serfling (1980), Theorem 3.3.A.):

$$\sqrt{T}(\widehat{\alpha}_c-\alpha, \widehat{\beta}_c-\beta, \widehat{\kappa}_c-\kappa, \widehat{\theta}_c(t)-\theta) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \mathbf{\Sigma}_c\right), \quad T\to\infty,$$

i.e.

$$(\mathbf{\Sigma}_c)^{-1/2}\sqrt{T}(\widehat{\alpha}_c-\alpha, \widehat{\beta}_c-\beta, \widehat{\kappa}_c-\kappa, \widehat{\theta}_c(t)-\theta) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \mathbf{I}\right), \quad T\to\infty, \tag{5.29}$$

where $\mathbf{\Sigma}_c = D\mathbf{\Sigma}D^\tau$ and where the elements of the (4×4) matrix

$$D = D(\alpha,\beta,\kappa,\theta) = \left[\frac{\partial g_i}{\partial x_j}\right]_{(x_1,x_2,x_3,x_4)=(m_{-1},m_1,m_2,m_{1,1}(t))},$$

for function g from (5.26) and for $i\in\{1,2,3,4\}$, are given by following expressions:

$$d_{11} = -\frac{\kappa(\alpha-2)^2(\alpha+2)(\beta-2)}{8(\alpha+\beta-2)}, \quad d_{12} = -\frac{\alpha(\alpha^2-4)(3\beta-8)}{8\kappa(\alpha+\beta-2)}, \quad d_{13} = \frac{(\alpha-2)\alpha^2(\beta-4)^2}{8\kappa^2(\beta-2)(\alpha+\beta-2)},$$

$$d_{14} = d_{24} = d_{31} = d_{33} = d_{34} = d_{41} = 0, \quad d_{32} = 1,$$

$$d_{21} = \frac{\kappa(\alpha-2)^2(\beta-4)(\beta-2)^2}{8\alpha(\alpha+\beta-2)}, \quad d_{22} = \frac{(3\alpha+2)(\beta-4)(\beta-2)\beta}{8\kappa(\alpha+\beta-2)}, \quad d_{23} = -\frac{\alpha(\beta-4)^2\beta}{8\kappa^2(\alpha+\beta-2)},$$

$$d_{42} = \frac{\alpha(\beta-4)}{\kappa t(\alpha+\beta-2)}(e^{\theta t}-1), \quad d_{43} = \frac{\alpha(\beta-4)}{2\kappa^2 t(\alpha+\beta-2)}, \quad d_{44} = -\frac{\alpha(\beta-4)}{2\kappa^2 t(\alpha+\beta-2)}e^{\theta t}. \tag{5.30}$$

Now the explicit covariances from the asymptotic covariance matrix $\mathbf{\Sigma}_c = D\mathbf{\Sigma}D^\tau$ can easily be calculated by matrix multiplication. Since matrices $\mathbf{\Sigma} = \mathbf{\Sigma}(\alpha,\beta,\kappa,\theta)$ and $D = D(\alpha,\beta,\kappa,\theta)$ depend on unknown parameters α, β, κ and θ, so does the matrix $\mathbf{\Sigma}_c$. However, since $(\widehat{\alpha}_c, \widehat{\beta}_c, \widehat{\kappa}_c, \widehat{\theta}_c(t))$ is for each $t>0$ P-consistent estimator of the parameter $(\alpha,\beta,\kappa,\theta)$ and since

$$\left[D(\widehat{\alpha}_c, \widehat{\beta}_c, \widehat{\kappa}_c, \widehat{\theta}_c(t))\,\mathbf{\Sigma}(\widehat{\alpha}_c, \widehat{\beta}_c, \widehat{\kappa}_c, \widehat{\theta}_c(t))\,D(\widehat{\alpha}_c, \widehat{\beta}_c, \widehat{\kappa}_c, \widehat{\theta}_c(t))^\tau\right]^{-1/2} \times$$

$$\times\left[D(\alpha,\beta,\kappa,\theta)\,\mathbf{\Sigma}(\alpha,\beta,\kappa,\theta)\,D(\alpha,\beta,\kappa,\theta)^\tau\right]^{1/2} =$$

$$(\widehat{\mathbf{\Sigma}}_c)^{-1/2}(\mathbf{\Sigma}_c)^{1/2} \xrightarrow{P} \mathbf{I}, \tag{5.31}$$

according to (5.29), (5.31) and the Slutsky theorem (see Serfling (1980), Theorem 1.5.4.1) we finally obtain that for every $t > 0$

$$(\widehat{\mathbf{\Sigma}}_c)^{-1/2}\sqrt{T}\,(\widehat{\alpha}_c - \alpha, \widehat{\beta}_c - \beta, \widehat{\kappa}_c - \kappa, \widehat{\theta}_c(t) - \theta) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \mathbf{I}\right), \quad T \to \infty, \tag{5.32}$$

where $\widehat{\mathbf{\Sigma}}_c = \widehat{D}\widehat{\mathbf{\Sigma}}\widehat{D}^{\tau}$ depends on consistent estimate of parameter $(\alpha, \beta, \kappa, \theta)$.

Remark 5.1. If $(X_1, \ldots, X_T)$ are discrete observations from the non-stationary FSD, then $\overline{m}_{-1,d}$, $\overline{m}_{1,d}$, $\overline{m}_{2,d}$ and for every fixed $t > 0$ $m_{1,1,d}(t)$ given by (3.2) and (3.1) are, according to the Theorem 3.3 from Kulik and Leonenko (2013) (see also Appendix B), consistent estimators of respective moments. Therefore, for each $t > 0$

$$(\overline{m}_{-1,d}, \overline{m}_{1,d}, \overline{m}_{2,d}, \overline{m}_{1,1,d}(t)) \xrightarrow{P} \left(E[1/X_t^{st}], E[X_t^{st}], E[(X_t^{st})^2], E[X_0^{st}X_t^{st}]\right), \quad T \to \infty. \tag{5.33}$$

To verify the asymptotic normality of the estimator $(\overline{m}_{-1,d}, \overline{m}_{1,d}, \overline{m}_{2,d}, \overline{m}_{1,1,d}(t))$ by using the asymptotic normality of the continuous-time setting estimator from Theorem 3.2, rewrite

$$\sqrt{T}\,(\overline{m}_{-1,d}, \overline{m}_{1,d}, \overline{m}_{2,d}, \overline{m}_{1,1,d}(t))$$

as follows:

$$\begin{aligned}\sqrt{T}\,(\overline{m}_{-1,d}, \overline{m}_{1,d}, \overline{m}_{2,d}, \overline{m}_{1,1,d}(t)) + \sqrt{T}\,(\overline{m}_{-1,c}, \overline{m}_{1,c}, \overline{m}_{2,c}, \overline{m}_{1,1,c}(t)) - \sqrt{T}\,(\overline{m}_{-1,c}, \overline{m}_{1,c}, \overline{m}_{2,c}, \overline{m}_{1,1,c}(t)) \;=\\ \sqrt{T}\,(\overline{m}_{-1,c}, \overline{m}_{1,c}, \overline{m}_{2,c}, \overline{m}_{1,1,d}(t)) + \sqrt{T}\,(\overline{m}_{-1,d} - \overline{m}_{-1,c}, \overline{m}_{1,d} - \overline{m}_{1,c}, \overline{m}_{2,d} - \overline{m}_{2,c}, \overline{m}_{1,1,d} - \overline{m}_{1,1,c})\,.\end{aligned} \tag{5.34}$$

Note that in the expression (5.34) we have the following types of convergence:

- according to the part (ii) of the Theorem 3.2 for every $t > 0$

$$\sqrt{T}\,(\overline{m}_{-1,c} - E[\overline{m}_{-1,c}], \overline{m}_{1,c} - E[\overline{m}_{1,c}], \overline{m}_{2,c} - E[\overline{m}_{2,c}], \overline{m}_{1,1,c}(t) - E[\overline{m}_{1,1c}(t)]) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}), \quad T \to \infty, \tag{5.35}$$

  where $\mathbf{\Sigma}$, depending on unknown parameters, is given by (5.28)

- by direct computation we see that for every $t > 0$

$$E\left[\sqrt{T}(\overline{m}_{-1,d} - \overline{m}_{-1,c}, \overline{m}_{1,d} - \overline{m}_{1,c}, \overline{m}_{2,d} - \overline{m}_{2,c}, \overline{m}_{1,1,d}(t) - \overline{m}_{1,1,c}(t))\right] = \mathbf{0}$$

  and that

$$\mathrm{Var}\left(\sqrt{T}(\overline{m}_{-1,d} - \overline{m}_{-1,c}, \overline{m}_{1,d} - \overline{m}_{1,c}, \overline{m}_{2,d} - \overline{m}_{2,c}, \overline{m}_{1,1,d}(t) - \overline{m}_{1,1,c}(t))\right) =$$

$$\frac{1}{T}\,\mathrm{Var}\left(\sum_{i=1}^{T}\frac{1}{X_i^{st}} - \int_0^T \frac{1}{X_s^{st}}\,ds, \sum_{i=1}^{T} X_i^{st} - \int_0^T X_s^{st}\,ds, \sum_{i=1}^{T}(X_i^{st})^2 - \int_0^T (X_s^{st})^2\,ds, \sum_{i=1}^{T} X_0^{st}X_i^{st} - \int_0^T X_0^{st}X_s^{st}\,ds\right) \to \mathbf{0},$$

  as $T \to \infty$. Therefore, according to the Chebyshev inequality, it follows that for every $t > 0$

$$\sqrt{T}\,(\overline{m}_{-1,d} - \overline{m}_{-1,c}, \overline{m}_{1,d} - \overline{m}_{1,c}, \overline{m}_{2,d} - \overline{m}_{2,c}, \overline{m}_{1,1,d}(t) - \overline{m}_{1,1,c}(t)) \xrightarrow{P} \mathbf{0}, \quad T \to \infty. \tag{5.36}$$

According to the Slutsky theorem (see Serfling (1980), Theorem 1.5.4.1.), from expressions (5.34), (5.35) and (5.36) it follows that for every $t > 0$

$$\sqrt{T}\,(\overline{m}_{-1,d}, \overline{m}_{1,d}, \overline{m}_{2,d}, \overline{m}_{1,1,d}(t)) \xrightarrow{d} \mathcal{N}\left((E[\overline{m}_{-1,c}], E[\overline{m}_{1,c}], E[\overline{m}_{2,c}], E[\overline{m}_{1,1,c}]), \mathbf{\Sigma}\right), \quad T \to \infty.$$

Since $E[\overline{m}_{-1,c}] = E[1/X_t^{st}]$, $E[\overline{m}_{1,c}] = E[X_t^{st}]$, $E[\overline{m}_{2,c}] = E[(X_t^{st})^2]$, $E[\overline{m}_{1,1,c}] = E[X_0^{st}X_t^{st}]$, it follows that for every $t > 0$

$$(\mathbf{\Sigma})^{-1/2}\sqrt{T}\,\left(\overline{m}_{-1,d} - E[1/X_t^{st}], \overline{m}_{1,d} - E[X_t^{st}], \overline{m}_{2,d} - E[(X_t^{st})^2], \overline{m}_{1,1,d} - E[X_0^{st}X_t^{st}]\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \mathbf{I}\right), \quad T \to \infty. \tag{5.37}$$

Therefore, from the expressions (5.35) and (5.37) it follows that $(\overline{m}_{-1,c}, \overline{m}_{1,c}, \overline{m}_{2,c}, \overline{m}_{1,1,c})$ and $(\overline{m}_{-1,d}, \overline{m}_{1,d}, \overline{m}_{2,d}, \overline{m}_{1,1,d})$ have the same asymptotic distributions.
Let

$$\begin{aligned} \widehat{\alpha}_d &= \frac{2(\overline{m}_{-1,d}\overline{m}_{1,d}\overline{m}_{2,d} - \overline{m}_{1,d}^2)}{\overline{m}_{-1,d}\overline{m}_{1,d}\overline{m}_{2,d} - 2\overline{m}_{2,d} + \overline{m}_{1,d}^2}, & \widehat{\beta}_d &= \frac{4\overline{m}_{-1,d}(\overline{m}_{1,d}^2 - \overline{m}_{2,d})}{2\overline{m}_{-1,d}\overline{m}_{1,d}^2 - \overline{m}_{-1,d}\overline{m}_{2,d} - \overline{m}_{1,d}}, \\ \widehat{\kappa}_d &= \overline{m}_{1,d}, & \widehat{\theta}_d(t) &= -\frac{1}{t}\log\left(\frac{\overline{m}_{1,1,d}(t) - \overline{m}_{1,d}^2}{\overline{m}_{2,d} - \overline{m}_{1,d}^2}\right) \end{aligned} \tag{5.38}$$

be the discrete time estimators of the unknown parameters.
Finally, by the same procedure like in the proof of the Theorem 3.2, part (ii), it follows that for every $t > 0$

$$(\widehat{D}\widehat{\mathbf{\Sigma}}\widehat{D}^\tau)^{-1/2}\sqrt{T}(\widehat{\alpha}_d - \alpha, \widehat{\beta}_d - \beta, \widehat{\kappa}_d - \kappa, \widehat{\theta}_d(t) - \theta) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{I}), \quad T \to \infty, \tag{5.39}$$

where $\widehat{D} = D(\widehat{\alpha}_d, \widehat{\beta}_d, \widehat{\kappa}_d, \widehat{\theta}_d(t))$ and $\widehat{\mathbf{\Sigma}} = \mathbf{\Sigma}(\widehat{\alpha}_d, \widehat{\beta}_d, \widehat{\kappa}_d, \widehat{\theta}_d(t))$ are matrices as in the Theorem 3.2 depending on consistent estimates of parameters.

Acknowledgements

N.N. Leonenko was supported in particular by Cardiff Incoming Visiting Fellowship Scheme and International Collaboration Seedcorn Fund, Cardiff Data Innovation Research Institute Seed Corn Funding, Australian Research Council's Discovery Projects funding scheme (project number DP160101366), and by projects MTM2012-32674 and MTM2015-71839-P (co-funded with Federal funds), of the DGI, MINECO, Spain.

N. Šuvak and I. Papić were supported by the scientific project UNIOS-ZUP 2018-31 funded by the J. J. Strossmayer University of Osijek.

At last, authors would like to thank Danijel Grahovac (Department of Mathematics, J.J. Strossmayer University of Osijek) for many useful discussions regarding the tail-index estimation.

Appendix A. Spectral representation of transition density of Fisher-Snedecor diffusion

For deriving the closed-form results in the framework of statistical analysis, e.g. calculation of asymptotic covariances in explicit form, the explicit expression for diffusion transition density

$$p(x,t) = p(x; x_0, t) = \frac{d}{dx}P(X_t \le x \mid X_0 = x_0), \quad x > 0, \quad t \ge 0,$$

can be extremely useful.

For canonical FSD such representation is given in terms of the spectrum of the corresponding infinitesimal generator and it is thoroughly studied in Avram et al. (2013b). In the general case of the FSD (2.9) satisfying the SDE (1.1) and having the invariant density (2.8), infinitesimal generator is defined as follows:

$$(\mathcal{G}f)(x) = \frac{2\theta}{\beta - 2}\,x\left(x + \frac{\varrho}{\alpha}\right)f''(x) - \theta\left(x - \frac{\varrho}{\beta - 2}\right)f'(x), \quad x > 0. \tag{A.1}$$

The domain of the operator $\mathcal{G}$ is the space of functions

$$D(\mathcal{G}) = \left\{f \in L^2((0,\infty), \mathfrak{p}(x)) \cap C^2((0,\infty)) : \mathcal{G}f \in L^2((0,\infty), \mathfrak{p}(x)), \lim_{x\to 0}\frac{f'(x)}{\mathfrak{s}(x)} = \lim_{x\to\infty}\frac{f'(x)}{\mathfrak{s}(x)} = 0\right\},$$

where $\mathfrak{s}(x)$ is the scale density given by (2.4) with $\kappa = \varrho/(\beta - 2)$.

As for the canonical FSD, the spectrum of the operator $(-\mathcal{G})$ consists of two disjoint parts: the discrete spectrum and the essential spectrum (see Avram et al. (2013b), Subsection 4.3). The discrete spectrum of the operator $(-\mathcal{G})$ is the finite set $\sigma_d(-\mathcal{G}) = \{\lambda_n,\, n = 0, \ldots, \lfloor\beta/4\rfloor\}$, where the eigenvalues λ_n are given by

$$\lambda_n = \frac{\theta}{\beta - 2}\,n(\beta - 2n), \quad \theta > 0, \quad \beta > 2, \quad n = 0, \ldots, \left\lfloor\frac{\beta}{4}\right\rfloor, \tag{A.2}$$

and the corresponding eigenfunctions are orthogonal Fisher-Snedecor polynomials given by the Rodrigues formula

$$P_n(x) = K_n\widetilde{P}_n(x) = K_n x^{1-\frac{\alpha}{2}}\left(x + \frac{\varrho}{\alpha}\right)^{\frac{\alpha+\beta}{2}}\frac{d^n}{dx^n}\left\{x^{\frac{\alpha}{2}+n-1}\left(x + \frac{\varrho}{\alpha}\right)^{n-\frac{\alpha+\beta}{2}}\right\}, \tag{A.3}$$

where $\widetilde{P}_n(x)$ are non-normalized polynomials and the normalization constant K_n can be expresses explicitly. The essential spectrum of the operator $(-\mathcal{G})$ is $\sigma_{ess}(-\mathcal{G}) = [\Lambda, \infty)$, where

$$\Lambda = \frac{\theta\beta^2}{8(\beta-2)}, \quad \theta > 0, \quad \beta > 2.$$

Moreover, operator $(-\mathcal{G})$ has the absolutely continuous spectrum of multiplicity one in (Λ, ∞), i.e. $\sigma_{ac}(-\mathcal{G}) \subseteq (\Lambda, \infty) \subset \sigma_{ess}(-\mathcal{G})$, whose elements could be parameterized by

$$\lambda = \Lambda + \frac{2\theta k^2}{\beta-2} = \frac{2\theta}{\beta-2}\left(\frac{\beta^2}{16} + k^2\right), \quad \theta > 0, \quad \beta > 2, \quad k > 0, \tag{A.4}$$

and where Λ is the cutoff between the absolutely continuous spectrum and the discrete spectrum. According to Borodin and Salminen (2002), in the general case for spectral representation of the transition density two linearly independent solutions of the SL equation, one of which is strictly increasing while the other one is strictly decreasing, are crucial. Such solutions in the case of the FSD are

$$f_1(x) = f_1(x, -\lambda) \;\; = {}_2F_1\left(z_{+,\lambda}, z_{-,\lambda}; \frac{\alpha}{2}; -\frac{\alpha}{\varrho}x\right), \tag{A.5}$$

$$f_4(x) = f_4(x, -\lambda) = \left(\frac{\alpha}{\varrho}x\right)^{-z_{+,\lambda}} {}_2F_1\left(z_{+,\lambda}, u_{+,\lambda}; 1 + 2\Delta_\lambda; -\frac{\varrho}{\alpha x}\right), \tag{A.6}$$

where $\lambda > \Lambda$ is the spectral parameter,

$$\Delta_\lambda = \sqrt{\frac{\beta^2}{16} - \frac{\lambda(\beta-2)}{2\theta}}, \quad z_{\pm,\lambda} = -\frac{\beta}{4} \pm \Delta_\lambda, \quad u_{\pm,\lambda} = 1 - \frac{\alpha}{2} + z_{\pm,\lambda} \tag{A.7}$$

and ${}_2F_1(a, b; c; \cdot)$ is the Gauss hypergeometric function (see e.g. Nikiforov and Uvarov (1988) or Luke (1969)). Due to the procedure of the analytic continuation of the function ${}_2F_1(a, b; c; \cdot)$, solutions $f_1(x, -\lambda)$ and $f_4(x, -\lambda)$ are well defined on the whole state space of the FSD. Spectral representation of transition density $p(x; x_0, t)$ is given in Theorem Appendix A.1. The proof can be conducted analogously as in the canonical case for which we refer to Avram et al. (2013b), Theorem 4.1.

Theorem Appendix A.1. *Spectral representation of the transition density of the FSD with the PDF* (2.8) *with parameters $\alpha > 2$, $\alpha \notin \{2(m+1), m \in \mathbb{N}\}$, $\beta > 2$, $\varrho > 0$ and $\theta > 0$ is of the form*

$$p(x; x_0, t) = p_d(x; x_0, t) + p_c(x; x_0, t). \tag{A.8}$$

The discrete part of the spectral representation

$$p_d(x; x_0, t) = \mathfrak{p}(x) \sum_{n=0}^{\lfloor \frac{\beta}{4} \rfloor} e^{-\lambda_n t} P_n(x_0)\, P_n(x) \tag{A.9}$$

is given in terms of the eigenvalues λ_n given by (A.2) *and the normalized Fisher-Snedecor polynomials $P_n(\cdot)$ given by* (A.3)*. The continuous part of the spectral representation*

$$p_c(x; x_0, t) = \mathfrak{p}(x)\, \frac{1}{\pi} \int_{\Lambda}^{\infty} e^{-\lambda t}\, k(\lambda) \times \tag{A.10}$$

$$\times \left| \frac{B^{\frac{1}{2}}\left(\frac{\alpha}{2}, \frac{\beta}{2}\right) \Gamma\left(-\frac{\beta}{4} + ik(\lambda)\right) \Gamma\left(\frac{\alpha}{2} + \frac{\beta}{4} + ik(\lambda)\right)}{\Gamma\left(\frac{\alpha}{2}\right) \Gamma\left(1 + 2ik(\lambda)\right)} \right|^2 f_1(x_0, -\lambda) f_1(x, -\lambda)\, d\lambda$$

is given in terms of the elements λ of the absolutely continuous spectrum of the operator $(-\mathcal{G})$ given by (A.4)*, solution $f_1(\cdot, -\lambda)$ of the Sturm-Liouville equation $(\mathcal{G}f)(x) = -\lambda f(x)$ for $\lambda > \Lambda$ given by* (A.5) *and parameter $k(\lambda) = -i\Delta_\lambda$, where Δ_λ is given in* (A.7)*.*

Furthermore, the explicit expression for the corresponding two-dimensional density is given by following expression:

$$p(x, y, t) = \frac{\partial^2}{\partial x \partial y} P(X_{s+t} \le x, X_s \le y) = \mathfrak{p}(y)\, p(x; y, t) = \mathfrak{p}(y)\, \left(p_d(x; y, t) + p_c(x; y, t)\right), \tag{A.11}$$

where $p_d(x; y, t)$ is given by (A.9) and $p_c(x; y, t)$ is given by (A.10). Representation (A.11) of two-dimensional density of the FSD can be used in calculation of explicit form of expectation $E[X_s^m X_t^n]$, $s, t \in (0, \infty)$, which is very useful for calculating the explicit expressions of asymptotic covariances of parameter estimator in asymptotic normality framework (see Avram et al. (2011)).

Appendix B. Important results on non-stationary Fisher-Snedecor diffusion

Appendix B.1. Coupling, ergodicity, and β-mixing

This section collects the results on ergodic behavior of the FSD. A traditional tool for proving the ergodicity of a Markov process X is the coupling construction. A coupling for a pair of processes U and V is any two-component process $Z = (Z^{(1)}, Z^{(2)})$ such that $Z^{(1)}$ has the same distribution as U and $Z^{(2)}$ has the same distribution as V. According to this terminology, for a Markov process X and every pair of probability distributions $\mu, \nu \in \mathcal{P}$, where $\mathcal{P}$ is the family of probability distributions on the Borel σ-algebra on the diffusion state space $\mathbb{X}$, we consider two versions $X^{(\mu)}$ and $X^{(\nu)}$ of the process X with the initial distributions μ and ν, respectively. Any two-component process $Z = (Z^{(1)}, Z^{(2)})$ which is a coupling for $X^{(\mu)}$ and $X^{(\nu)}$ is called (μ, ν)-coupling for the process X. According to Kulik and Leonenko (2013), the Markov process X admits an exponential ϕ-coupling if there exists an invariant measure π for this process and positive constants C and c such that, for every $\mu \in \mathcal{P}$, there exists a (μ, π)-coupling $Z = (Z^{(1)}, Z^{(2)})$ such that

$$E\Big[\phi(Z_t^1) + \phi(Z_t^2)\Big]\mathbb{1}_{Z_t^1 \neq Z_t^2} \leq Ce^{-ct}\int_{\mathbb{X}} \phi \, d\mu, \quad t \geq 0. \tag{B.1}$$

In Kulik (2011) an exponential ϕ-coupling is introduced, and it was demonstrated that it is a convenient tool for studying convergence rates of L_p-semigroups, generated by a Markov process, and spectral properties of respective generators. In Kulik and Leonenko (2013) it is shown that this notion is also efficient for proving LLN and CLT for the FSD in non-stationary setting.

Next we provide the definition of the well-known β-mixing coefficient, also known as complete regularity or Kolmogorov's coefficient. Generally, β-mixing coefficient of the process X is defined as

$$\beta^{\mu}(t) = \sup_{s\geq 0} E_{\mu} \sup_{B \in \mathcal{F}^X_{\geq t+s}} |P_{\mu}(B|\mathcal{F}^X_s) - P_{\mu}(B)|, \quad \mu \in \mathcal{P}, \quad t \geq 0, \tag{B.2}$$

where $\mathcal{F}^X_{\geq r}$ for a given $r \geq 0$ denotes the σ-algebra generated by the process X at times $v \geq r$.

The state-dependent β-mixing coefficient is defined by

$$\beta_x(t) = \sup_{s\geq 0} E_x \sup_{B \in \mathcal{F}^X_{\geq t+s}} |P_x(B|\mathcal{F}^X_s) - P_x(B)|, \quad x \in \mathbb{X}, \quad t \geq 0, \tag{B.3}$$

where the initial distribution of X is the degenerate distribution $\mu = \delta_x$.

The stationary β-mixing coefficient is defined by

$$\beta(t) = \sup_{s\geq 0} E_{\pi} \sup_{B \in \mathcal{F}^X_{\geq t+s}} |P_{\pi}(B|\mathcal{F}^X_s) - P_{\pi}(B)|, \quad x \in \mathbb{X}, \quad t \geq 0, \tag{B.4}$$

where π denotes the (unique) invariant distribution of the process X. For more information about various types of mixing coefficients see e.g. Bradley (2005).

Finally, results concerning the ϕ-coupling and β-mixing for the non-stationary FSD are stated in the Theorem Appendix B.1. For the proof we refer to Kulik and Leonenko (2013), Theorem 3.1.

Theorem Appendix B.1. *Let the function ϕ be defined as $\phi = \phi_{\diamondsuit} + \phi_{\blacklozenge}$, where $\phi \geq 1$, $\phi_{\diamondsuit}, \phi_{\blacklozenge} \in C^2(0,\infty)$, $\phi_{\diamondsuit} = 0$ on $[2,\infty)$, $\phi_{\blacklozenge} = 0$ on $(0,1]$, $\phi_{\diamondsuit}(x) = x^{-\gamma}$ for x small enough and $\phi_{\blacklozenge}(x) = x^{\delta}$ for x large enough with non-negative γ and δ satisfying $\gamma < \dfrac{\alpha}{2} - 1$ and $\delta < \dfrac{\beta}{2}$. Then the following statements hold true.*

1. *FSD admits an exponential ϕ-coupling.*
2. *Finite-dimensional distributions of the FSD admit the following convergence rate in the weighted total variation norm with the weight ϕ: for any $m \geq 1$ and $0 \leq t_1 < \ldots < t_m$*

$$\|\mu_{t+t_1,\ldots,t+t_m} - \pi_{t_1,\ldots,t_m}\|_{\phi,var} \leq mCe^{-ct}\int_{\mathbb{X}} \phi \, d\mu, \quad \mu \in \mathcal{P}, \quad t \geq 0. \tag{B.5}$$

Here $\mu_{t_1,\ldots,t_m}$, $0 \leq t_1 < \ldots < t_m$, $m \geq 1$, denotes finite-dimensional distributions of the respective diffusion with the initial distribution μ, while $\pi_{t_1,\ldots,t_m}$ denotes the corresponding finite-dimensional invariant distribution. Constants C and c are the same as in the bound (B.1) in the definition of an exponential ϕ-coupling.

3. *FSD admits the following bound for the β-mixing coefficient:*

$$\beta^{\mu}(t) \le C' e^{-ct} \int_{\mathbb{X}} \phi \, d\mu, \quad \mu \in \mathcal{P}, \quad t \ge 0. \tag{B.6}$$

Here the constant c is the same as in the bound (B.1), and C' is a positive constant which can be given explicitly (see Kulik and Leonenko (2013), relation (5.15)).

Furthermore, from (B.6) and Corollary 3.1 from Kulik and Leonenko (2013), the following bounds for the β-mixing coefficients can be obtained:

- bound for the state-dependent β-mixing coefficient:

$$\beta_x(t) \le C' e^{-ct} \phi(x), \quad x \in \mathbb{X}, \quad t \ge 0 \tag{B.7}$$

- bound for the stationary β-mixing coefficient:

$$\beta(t) \le C'' e^{-ct}, \quad t \ge 0, \quad C'' := C' \int_{\mathbb{X}} \phi \, d\pi < +\infty. \tag{B.8}$$

Appendix B.2. Limit theorems for additive functionals for random samples from the Fisher-Snedecor diffusion

Here we state the LLN and CLT for additive functionals of the FSD X, separately for the discrete-time and the continuous-time observations. For the proofs we refer to the recent paper Kulik and Leonenko (2013), Theorems 3.3 and 3.4. For clarity of the exposition, we introduce the notation $X^{st} = (X_t^{st}, t \in (-\infty, \infty))$ for the stationary version of the FSD X, by which we understand the strictly stationary process such that for every $m \ge 1$ and $t_1 < \ldots < t_m$ the distribution of the random vector $X_{t_1}^{st}, \ldots, X_{t_m}^{st}$ is $\pi_{0,t_2-t_1,\ldots,t_m-t_1}$ (time-shift invariance of the finite-dimensional distributions). Heuristically, X^{st} is a solution of the SDE (1.1) defined on the whole time axis and starting at $(-\infty)$ from the invariant distribution π.

Theorem Appendix B.2. *(Discrete-time case)*
Let, for some $r, k \ge 1$, a vector-valued function

$$f = (f_1, \ldots, f_k) \colon \mathbb{X}^r \to \mathbb{R}^k$$

be such that for any $i = 1, \ldots, k$ for some γ_i, δ_i such that $\gamma_i < (\alpha/2) - 1$ and $\delta_i < \beta/2$

$$|f_i(x)| \le C \sum_{j=1}^{r} \left(x_j^{-\gamma_i} + x_j^{\delta_i} \right), \quad x = (x_1, \ldots, x_r) \tag{B.9}$$

with some constant C. Then the following statements hold true.

1. *Law of large numbers*
   For arbitrary initial distribution μ of X and arbitrary $t_1, \ldots, t_r \ge 0$,

$$\frac{1}{n} \sum_{l=1}^{n} f\big(X_{t_1+l}, \ldots, X_{t_r+l}\big) \xrightarrow{P} a_f, \tag{B.10}$$

   with the asymptotic mean vector

$$a_f = E f\big(X_{t_1}^{st}, \ldots, X_{t_r}^{st}\big).$$

   If, in addition, the initial distribution is such that for some positive ε

$$\int_{\mathbb{X}} \big(x^{-\gamma_i - \varepsilon} + x^{\delta_i + \varepsilon}\big) \, \mu(dx) < \infty, \quad i = 1, \ldots, k, \tag{B.11}$$

   then (B.10) holds true in the mean sense.
2. *Central limit theorem*
   Assume in addition that there exists $\varepsilon > 0$ such that

$$E \left\| f\big(X_{t_1}^{st}, \ldots, X_{t_r}^{st}\big) \right\|^{2+\varepsilon} < \infty. \tag{B.12}$$

Then

$$\frac{1}{\sqrt{n}}\sum_{l=1}^{n}\Big(f\big(X_{t_1+l},\dots,X_{t_r+l}\big)-a_f\Big)\Rightarrow\mathcal{N}(0,\mathbf{\Sigma}),\tag{B.13}$$

where the components of the asymptotic covariance matrix $\mathbf{\Sigma}$ *are given as follows:*

$$(\mathbf{\Sigma})_{i,j}=\sum_{l=-\infty}^{\infty}\operatorname{Cov}\Big(f_i\big(X^{st}_{t_1+l},\dots,X^{st}_{t_r+l}\big),f_j\big(X^{st}_{t_1},\dots,X^{st}_{t_r}\big)\big)\Big),\quad i,j=1,\dots,k.$$

Theorem Appendix B.3. *(Continuous-time case)*
Let the components of a vector-valued function $f\colon\mathbb{X}^r\to\mathbb{R}^k$ *satisfy (B.9) with* γ_i,δ_i *satisfying* $\gamma_i<\alpha/2$ *and* $\delta_i<\beta/2$ *for every* $i=1,\dots,k$*. Then the following statements hold true.*

1. *Law of large numbers*
   For arbitrary initial distribution μ *of* X

$$\frac{1}{T}\int_0^T f\big(X_{t_1+s},\dots,X_{t_r+s}\big)\,ds\xrightarrow{P}a_f.\tag{B.14}$$

   If, in addition, the initial distribution is such that for some positive ε

$$\int_{\mathbb{X}}\Big(x^{-(\gamma_i-1)\vee 0-\varepsilon}+x^{\delta_i+\varepsilon}\Big)\,\mu(dx)<\infty,\quad i=1,\dots,k,\tag{B.15}$$

   then (B.14) holds true in the mean sense.
2. *Central limit theorem*
   Assume in addition that

$$\gamma_i<\frac{\alpha}{4}+\frac{1}{2},\quad\delta_i<\frac{\beta}{4},\quad i=1,\dots,k.\tag{B.16}$$

   Then for arbitrary initial distribution μ *of* X

$$\frac{1}{\sqrt{T}}\int_0^T\Big(f\big(X_{t_1+s},\dots,X_{t_r+s}\big)-a_f\Big)\,ds\Rightarrow\mathcal{N}(0,\mathbf{\Sigma}),\tag{B.17}$$

   where the components of the asymptotic covariance matrix $\mathbf{\Sigma}$ *are given as follows:*

$$(\mathbf{\Sigma})_{i,j}=\int_{-\infty}^{\infty}\operatorname{Cov}\Big(f_i\big(X^{st}_{t_1+s},\dots,X^{st}_{t_r+s}\big),f_j\big(X^{st}_{t_1},\dots,X^{st}_{t_r}\big)\big)\Big)\,ds,\quad i,j=1,\dots,k.\tag{B.18}$$

Statements of Theorems Appendix B.2 and Appendix B.3 clearly show that the technique for calculation of asymptotic covariances $(\mathbf{\Sigma})_{i,j}$ in discrete and continuous-time setting rely on properties of the stationary FSD X^{st}. Therefore, we refer to Appendix A, where we give a short overview of the most important probabilistic properties of the stationary FSD in the canonical case.

Bibliography

Abourashchi, N., Veretennikov, A. Y., 2009. On exponential mixing bounds and convergence rate for reciprocal gamma diffusion processes. Mathematical Communications 14 (2), 331–339.

Abourashchi, N., Veretennikov, A. Y., 2010. On exponential bounds for mixing and the rate of convergence for Student processes. Theory of Probability and Mathematical Statistics 81, 1–13.

Abramowitz, M., Stegun, I. A., 1967. Handbook of mathematical functions: with formulas, graphs, and mathematical tables/edited by Milton Abramowitz and Irene A. Stegun.

Avram, F., Leonenko, N., Šuvak, N., 2011. Parameter estimation for Fisher Snedecor diffusion. Statistics 45 (1), 27–42.

Avram, F., Leonenko, N., Šuvak, N., 2012. Hypothesis testing for Fisher Snedecor diffusion. Journal of Statistical Planning and Inference 142 (8), 2308–2321.

Avram, F., Leonenko, N., Šuvak, N., 2013a. On spectral analysis of heavy-tailed Kolmogorov-Pearson diffusions. Markov Processes and Related Fields 19 (2), 249–298.

Avram, F., Leonenko, N., Šuvak, N., 2013b. Spectral representation of transition density of Fisher Snedecor diffusion. Stochastics 85 (2), 346–369.

Bibby, M., Skovgaard, M., Sørensen, M., 2005. Diffusion-type models with given marginal distribution and autocorrelation function. Bernoulli 11 (2), 191–220.

Bishwal, J., 2007. Parameter Estimation in Stochastic Differential Equations. Lecture Notes in Mathematics. Springer Berlin Heidelberg.

Borodin, A., Salminen, P., 2002. Handbook of Brownian Motion - Facts and Formulae. Birkhäuser Basel.

Bradley, R., 2005. Basic properties of strong mixing conditions. A survey and some open questions. Probability Surveys 2, 107–144.

Forman, J., Sørensen, M., 2008. The Pearson diffusions: A class of statistically tractable diffusion processes. Scandinavian Journal of Statistics 35 (3), 438–465.

Forman, J. L., 2007. Statistical inference from diffusion driven models. Ph.D. thesis, Department of Applied Mathematics and Statistics, Faculty of Science, University of Copenhagen.

Genon-Catalot, V., Jeantheau, T., Larédo, C., 2000. Stochastic volatility models as hidden Markov models and statistical applications. Bernoulli 6 (6), 1051–1079.

Hall, P., Heyde, C., 1980. Martingale limit theory and its application. Academic Press.

Iacus, S., 2008. Simulation and Inference for Stochastic Differential Equations: With R Examples. Springer-Verlag, New York.

Itô, K., McKean, H., 1974. Diffusion Processes and their Sample Paths. Springer, New York, Heidelberg, Berlin.

Ivanov, A., 1997. Asymptotic Theory of Nonlinear regression. Kluwer Academic Publisher, Dordrecht.

Karlin, S., Taylor, H., 1981. A Second Course in Stochastic Processes. Academic Press.

Kessler, M., 1997. Estimation of an ergodic diffusion from discrete observations. Scandinavian Journal of Statistics 24 (2), 211–229.

Kessler, M., 2000. Simple and explicit estimating functions for a discretely observed diffusion process. Scandinavian Journal of Statistics 27 (1), 65–82.

Kessler, M., Lindner, A., Sørensen, M., 2012. Statistical Methods for Stochastic Differential Equations. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. CRC Press.

Kessler, M., Paredes, S., 2002. Computational aspects related to martingale estimating functions for a discretely observed diffusion. Scandinavian Journal of Statistics 29 (3), 425–440.

Kessler, M., Sørensen, M., 04 1999. Estimating equations based on eigenfunctions for a discretely observed diffusion process. Bernoulli 5 (2), 299–314.

Kloeden, P., Platen, E., 2011. Numerical Solution of Stochastic Differential Equations. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg.

Kulik, A., 2011. Asymptotic and spectral properties of exponentially ϕ-ergodic Markov processes. Stochastic Processes and their Applications 121 (5), 1044–1075.

Kulik, A., Leonenko, N., 2013. Ergodicity and mixing bounds for the Fisher-Snedecor diffusion. Bernoulli 19 (5b), 2294–2329.

Kutoyants, Y., 2004. Statistical Inference for Ergodic Diffusion Processes. Springer Series in Statistics. Springer.

Leonenko, N., Šuvak, N., 2010a. Statistical inference for reciprocal gamma diffusion process. Journal of Statistical Planning and Inference 140 (1), 30 – 51.

Leonenko, N., Šuvak, N., 2010b. Statistical inference for Student diffusion process. Stochastic Analysis abd Applictions 28 (6), 972–1002.

Luke, Y., 1969. The Special Functions And Their Approximations. Vol. 1. Academic Press.

Malinvaud, E., 1970. The consistency of nonlinear regressions. Ann. Math. Statist. 41 (3), 956–969.

Nikiforov, A., Uvarov, V., 1988. Special functions of mathematical physics: a unified introduction with applications. Birkhäuser.

Pearson, K., 1914. Tables for Statisticians and Biometricians, Part I. Biometrics Laboratory, University College London.

Serfling, R., 1980. Approximation Theorems of Mathematical Statistics. John Wiley & Sons, New York.

Shaw, W. T., Munir, A., 2009. Dependency without copulas or ellipticity. The European Journal of Finance 15 (7-8), 661–674.

Shiryayev, A., 1992. Selected Works of A. N. Kolmogorov: Volume II Probability Theory and Mathematical Statistics. Springer Netherlands, Dordrecht, Ch. On Analytical Methods In Probability Theory, pp. 62–108.

Slater, L., 1966. Generalized Hypergeometric Functions. Cambridge University Press.

Sørensen, M., 2012. Estimating functions for diffusion-type processes. In: Kessler, M., Lindner, A., Sørensen, M. (Eds.), Statistical Methods for Stochastic Differential Equations. Chapman and Hall/CRC Press, Boca Raton, Florida, pp. 1–107.