Estimation statistique des paramètres pour les processus de Cox-Ingersoll-Ross et de Heston

Marie Du Roy de Chaumaray

Estimation statistique des paramètres pour les processus de Cox-Ingersoll-Ross et de Heston

THÈSE

présentée et soutenue publiquement le 2 Décembre 2016

pour l'obtention du grade de

Docteur de l'Université de Bordeaux

(en Mathématiques appliquées)

par

Marie du Roy de Chaumaray

Composition du jury

Président :	Marguerite Zani	Université d'Orléans
Rapporteurs :	Matyas Barczy	Université de Debrecen, Hongrie
	Antoine Jacquier	Imperial College London, Angleterre
Examinateur :	Mathieu Rosenbaum	Ecole Polytechnique
Directeurs :	Bernard Bercu	Université de Bordeaux
	Adrien Richou	Université de Bordeaux

Institut de Mathématiques de Bordeaux — UMR 5251

Mis en page avec la classe thesul.

Remerciements

En tout premier lieu, je tiens à remercier Adrien et Bernard pour leur implication, leur disponibilité et leur sympathie. J'ai pris beaucoup de plaisir à travailler avec eux ces trois dernières années.

Je remercie Antoine Jacquier d'avoir bien voulu rapporter cette thèse avec autant d'attention. Je tiens aussi à remercier Matyas Barczy pour l'intérêt qu'il a porté à mon travail. Ses nombreux commentaires ont permis d'améliorer très largement ce manuscrit. Je remercie également Marguerite Zani et Mathieu Rosenbaum d'avoir accepté de faire partie de mon jury. C'est un honneur pour moi.

Mon aventure bordelaise n'aurait pas été la même sans les quatre générations de doctorants que j'ai eu la chance de rencontrer à l'IMB. Si, cette année, ce sont Elsa, Fabien, Jonathan, Nicolas, Nikola, Thomas et Sami qui égayent mon quotidien, je n'oublie pas pour autant les anciens : Alice, Aurel, Bruno, Camille, Delphine, Giovanni, JB (et Mélanie), Jocelyn, Louis (et Joanna), Maarten, Marc, Pierre, Romain, Samuel, Toni et Zoé. Je les remercie pour le ski, les week-end arcachonnais, garorock et nos innombrables mais non moins mémorables *party Hard* à l'Apollo, au Rose, rue Charles Domercq ou ailleurs. Je remercie aussi les très anciens : Frédéric, Philippe, Romain et Vassili, qui m'ont encouragée (et ils ont bien fait !) à travailler avec Bernard. Par ailleurs, je tiens à remercier l'équipe de natation synchronisée de l'université de Bordeaux, grâce à qui j'ai eu la chance de remporter deux années de suite les championnats de France universitaires. En particulier, je remercie ma chère Leïla de m'avoir fait passer de si bons mercredis et jeudis après-midi de 2012 à 2014.

Si je remonte un peu dans le passé, je me dois de remercier mes deux professeurs de mathématiques de classe prépa, Roger Mansuy et Denis Monasse, ainsi que mes camarades de l'époque, Aline, Benjamin, Lydie, Malix, Margaret, Quentin et Stein, sans qui je n'aurais certainement pas intégré l'ENS de Rennes.

C'est d'ailleurs à Rennes qu'est né mon goût pour les probabilités, grâce, en particulier, aux cours de Benoit Cadre, Mihai Gradinaru et Florent Malrieu. J'en profite pour remercier Mihai Gradinaru pour sa bienveillance et pour m'avoir initiée aux équations différentielles stochastiques en TER. Si j'en suis là aujourd'hui, il n'y est pas pour rien. Je ne pourrais clore ce paragraphe rennais sans remercier vivement les promotions maths2008 et maths2009, et en particulier Arnaud, Basile, Blandine, Charlotte, Florent, Florian, Gautier, Geoffrey, Kévin, Ophélie, Pierre et Simon.

Pour continuer avec les bretons, je voudrais remercier ma belle-famille. C'est toujours avec un grand plaisir que je les retrouve pour les vacances. Mais en plus d'une famille, j'ai aussi la chance d'avoir des amis par alliance : Benoît et Aurélie, Fanchon et Hélène, Karl et Patty, sans oublier Ewen, Gee et Luce, qui étaient là bien avant la thèse et le seront encore longtemps après.

Je vais m'envoler maintenant vers le sud-est pour remercier ma famille et, en particulier, mes parents, d'avoir toujours eu confiance en mes choix et pour leur soutien sans faille. J'en profite pour faire un gros clin d'oeil à Caro, qui a grandi à mes côtés, et sans qui rien n'aurait été pareil. Je tiens aussi à remercier Sylvie et Hervé Gras pour les nombreux documents mathématiques qu'ils m'ont prêtés depuis mon départ de Nîmes et pour leurs réserves, qui m'ont souvent amenée à bien réfléchir à mes différents projets pour ne rien regretter. Je remercie enfin Valérie Lacognata sans qui je ne serais pas celle que je suis. Elle m'a appris la danse mais aussi "la vie".

Je ne peux terminer sans une petite pensée pour Alan et Gireg, les deux hommes de ma vie, et pour la petite puce qui sera là dans trois mois.

Sommaire

Chapitre 1

Introduction générale

Tous les résultats de cette thèse s'articulent autour de l'estimation des paramètres de certains processus de diffusion utilisés en particulier en finance pour modéliser les prix d'actifs financiers ou leur volatilité. Il s'agit dans un premier temps du processus CIR, pour Cox, Ingersoll et Ross qui furent les premiers à l'utiliser dans un contexte économique pour modéliser des taux d'intérets à court terme [22]. Ce processus est aussi appelé processus d'Ornstein-Uhlenbeck radial carré ou encore diffusion de Feller [34]. Dans un second temps, on étudie le processus de Heston, qui correspond à un processus de diffusion de type Black-Scholes dont la volatilité est générée par un processus CIR.

Sommaire

1.1 Les processus

1.1.1 Le processus CIR

Le processus CIR est la solution forte de l'équation différentielle stochastique suivante :

$$\mathrm{d}X_t = (a + bX_t)\mathrm{d}t + 2\sqrt{X_t}\,\mathrm{d}B_t \tag{1.1}$$

où $x_0 \geqslant 0$ est le point de départ, $a \in \mathbb{R}^+$ est le paramètre de dimension, $b \in \mathbb{R}$ est le coefficient de dérive et (B_t) est un mouvement Brownien standard. Certains auteurs introduisent un paramètre supplémentaire σ dans l'équation (1.1), qui devient alors

$$\mathrm{d}X_t = (a + bX_t)\mathrm{d}t + \sigma\sqrt{X_t}\,\mathrm{d}B_t.$$

Si le paramètre σ est connu, le changement de temps $s = \sigma^2 t/4$ permet de se ramener au cas $\sigma = 2$. On a donc choisi de se placer dans cette configuration.

Outre ses applications en finance, le processus CIR est aussi utilisé en biologie des populations. Il peut en effet être vu comme un processus de branchement continu avec immigration, c'est-à-dire comme diffusion limite de processus de branchement avec immigration à espace d'états discret, voir [5] et [55] pour plus de détails sur ces processus. Le paramètre a correspond au taux d'immigration tandis que b représente l'intensité de branchement.

Le comportement du processus CIR a été très largement étudié et dépend des valeurs des paramètres a et b. En particulier, on connait la loi de X_t. On trouvera dans [58] un calcul de la densité de $-bX_t/(1-e^{bt})$ via la transformée de Laplace du couple $(X_t, \int_0^t X_s\,\mathrm{d}s)$. Elle est donnée pour tout $x > 0$, par

$$f_t(x) = \psi_t \exp\left(-x/2\right) x^{(a-2)/4} I_{(a-2)/2}\left(\sqrt{x\xi_t}\right)$$

où ξ_t et ψ_t sont deux constantes respectivement données par

$$\xi_t = -\frac{x_0 b}{e^{-bt} - 1} \qquad \text{et} \qquad \psi_t = \frac{e^{-\xi_t/2}}{2\xi_t^{(a-2)/4}},$$

et I_ν est la fonction de Bessel modifiée de première espèce (voir [44] ou [59] par exemple pour plus de détails sur cette fonction spéciale) définie, pour tout $x \in \mathbb{R}$ et tout $\nu \in \mathbb{R}$, par

$$I_\nu(x) = \sum_{k=0}^{+\infty} \frac{1}{k!\,\Gamma(k+\nu+1)} \left(\frac{x}{2}\right)^{2k+\nu}.$$

Il s'agit d'une loi du chi-deux décentrée à a degrés de liberté et de paramètre de décentrage ξ_t. On note au passage que la transformée de Laplace de X_t vérifie, pour tout $\lambda \geqslant 0$,

$$\mathbb{E}\left[e^{-\lambda X_t}\right] = \left(1 - \frac{2\lambda}{b}(1-e^{bt})\right)^{-a/2} \exp\left(\frac{-x_0\lambda e^{bt}}{1 - \frac{2\lambda}{b}(1-e^{bt})}\right).$$

Par ailleurs, X_t est un processus affine régulier, dont le générateur infinitésimal est donné par l'opérateur de différentiation de Laguerre $\mathcal{L}$, qui est défini sur $\mathcal{C}_c^2(\mathbb{R}^+, \mathbb{R})$ par

$$\mathcal{L} = 2x\frac{\mathrm{d}^2}{\mathrm{d}x^2} + (a+bx)\frac{\mathrm{d}}{\mathrm{d}x}.$$

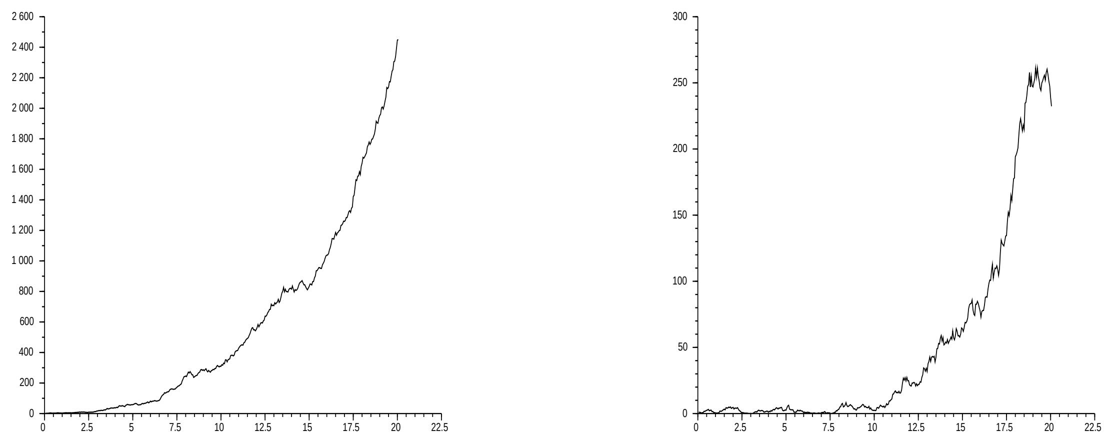
Pour la définition, la caractérisation et les applications en mathématiques financières des processus affines régulier, on pourra consulter [8] ou [29].

On note aussi que le processus CIR peut être relié à d'autres types de processus. Ainsi, pour a entier plus grand que 2, le processus CIR est la norme euclidienne au carré d'un processus d'Ornstein-Uhlenbeck a-dimensionnel de paramètre b. Par ailleurs, lorsque b vaut 0, (X_t) est un carré de Bessel de dimension a (voir [66]). Et plus généralement, si on note $BESQ_{x_0}^a(t)$ un carré de Bessel de dimension a et de point de départ x_0, alors X_t a la même loi que $e^{bt}BESQ_{x_0}^a\left(-\frac{1}{b}(e^{-bt}-1)\right)$.

On observe une trichotomie bien connue sur le processus CIR, suivant les valeurs du paramètre b.

• $b > 0$: cas supercritique

Le processus explose exponentiellement à la vitesse $\exp(bt)$. Cependant, si $0 \leqslant a < 2$, il y a un probabilité non nulle pour que le processus s'annule et même s'éteigne (dans le cas où $a = 0$). On a tracé dans la figure 1.1 deux trajectoires du processus partant de $x_0 = 1$ dans le cas supercritique pour deux valeurs différentes de a. Sur celle de droite, on a $a < 2$ et le processus s'annule plusieurs fois.



Cas $a = 5$ Cas $a = 1$

FIGURE 1.1 – Deux trajectoires pour $b = 0.2$

• $b = 0$: cas critique

Si $a = 0$, le processus s'éteint presque sûrement, mais le temps d'extinction n'est pas intégrable. Sinon, le processus croît linéairement « en moyenne ». On trace une trajectoire de (X_t) pour $a = 5$, partant de $x_0 = 1$.

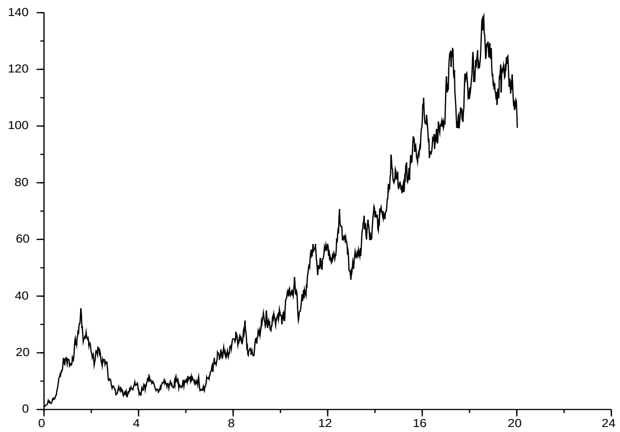


FIGURE 1.2 – Une trajectoire pour $b = 0$ et $a = 5$.

• $b < 0$: cas sous-critique

Si $a = 0$, le processus s'éteint presque sûrement. Par contre, si $a > 0$, l'immigration empêche l'extinction et le processus est même géométriquement ergodique. Ainsi, quand T tend vers l'infini, X_T converge en distribution vers X qui suit une loi Gamma $(a/2, -b/2)$, dont la densité de probabilité est donnée par

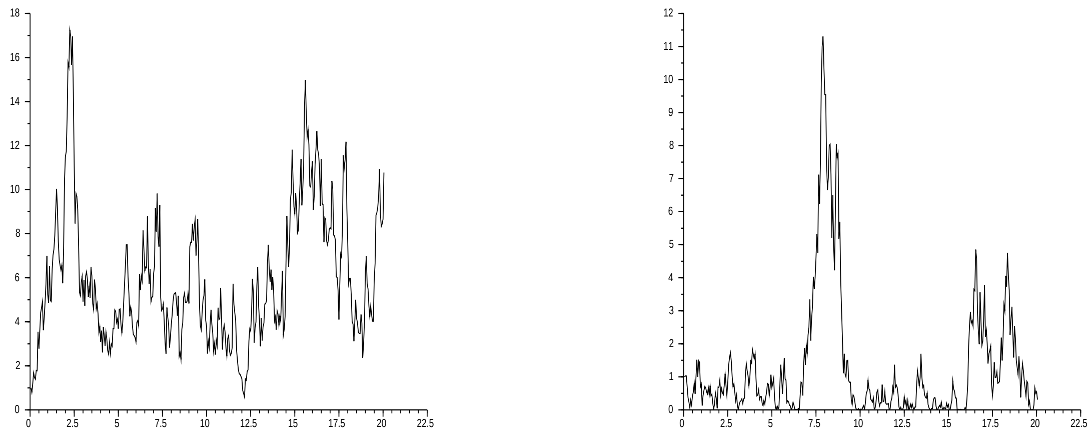
$$f(x) = (\Gamma(a/2))^{-1} \, (-b/2)^{a/2} \, x^{a/2-1} e^{xb/2} \mathbb{1}_{x>0}.$$

Lorsque $0 < a < 2$, le processus s'annule fréquemment, comme on peut le constater sur la figure 1.3. Mais il ne s'annule pas dans le cas où $a \geqslant 2$. De plus, pour $a > 2$, X^{-1} est intégrable et son espérance vaut

$$\mathbb{E}[X^{-1}] = -\frac{b}{2} \frac{\Gamma(a/2-1)}{\Gamma(a/2)}.$$

Ce résultat joue un rôle important dans l'étude du comportement asymptotique des estimateurs que l'on étudie par la suite.

On remarque que, dans cette trichotomie, le comportement en zéro est déterminé par la valeur de a. En effet, si $a = 0$, zéro est absorbant donc si le processus s'annule, il s'éteint. Si $a \geqslant 2$, zéro est inaccessible, le processus reste strictement positif. Et enfin, si $0 < a < 2$, le processus peut s'annuler, mais il est alors instantanément réfléchi. Dans toute la suite, on supposera que le paramètre a est strictement positif.



Cas $a = 5$ Cas $a = 1$

FIGURE 1.3 – Deux trajectoires pour $b = -1$

1.1.2 Le processus de Heston

Le processus de Heston [47] est la solution de l'équation différentielle stochastique suivante :

$$\begin{cases} \mathrm{d}X_t = (a + bX_t)\,\mathrm{d}t + 2\sqrt{X_t}\,\mathrm{d}B_t \\ \mathrm{d}Y_t = (c + dX_t)\,\mathrm{d}t + 2\sqrt{X_t}\left(\rho\,\mathrm{d}B_t + \sqrt{1-\rho^2}\,\mathrm{d}W_t\right) \end{cases} \tag{1.2}$$

avec $a > 0$, $(b, c, d) \in \mathbb{R}^3$ et $\rho \in]-1, 1[$, où (B_t, W_t) est un mouvement Brownien standard de dimension 2 et le point de départ vérifie $(x_0, y_0) \in \mathbb{R}^+ \times \mathbb{R}$. Ce processus est utilisé en finance comme généralisation du modèle de Black-Scholes. En effet, si Y_t désigne le logarithme du prix d'un certain actif financier, X_t représente sa volatilité, c'est-à-dire sa variance instantanée. Celle-ci n'est plus supposée constante comme dans le modèle de Black-Scholes mais est définie par un processus stochastique. On parle dans ce cas de modèle à volatilité stochastique. Le livre de Gatheral [39] fait un panorama des différents modèles à volatilité stochastique utilisés en finance. Dans le modèle de Heston, la volatilité est un processus CIR.

Le processus de Heston est lui aussi un processus affine régulier (voir la Proposition 2.1 de [11]). Son générateur infinitésimal $\mathcal{A}$ est défini sur $\mathcal{C}_c^2(\mathbb{R}^+ \times \mathbb{R}, \mathbb{R})$ par :

$$\mathcal{A} = (a + bx)\frac{\partial}{\partial x} + (c + dx)\frac{\partial}{\partial y} + 2x\left(\frac{\partial^2}{\partial x^2} + 2\rho\frac{\partial^2}{\partial x \partial y} + \frac{\partial^2}{\partial y^2}\right).$$

Bien qu'on ne puisse pas donner la loi du couple (X_t, Y_t) explicitement, on trouve un calcul de son espérance dans [48] ou encore dans [11] avec une méthode différente.

Proposition 1.1.1. *Soit* (X_t, Y_t) *solution de* (1.2). *Pour tout* $t > 0$, *son espérance vérifie*

$$\begin{pmatrix} \mathbb{E}[X_t] \\ \mathbb{E}[Y_t] \end{pmatrix} = \begin{pmatrix} e^{bt} & 0 \\ d\int_0^t e^{bs}\,\mathrm{d}s & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + \begin{pmatrix} \int_0^t e^{bs}\,\mathrm{d}s & 0 \\ d\int_0^t \int_0^s e^{bu}\,\mathrm{d}u\mathrm{d}s & t \end{pmatrix} \begin{pmatrix} a \\ c \end{pmatrix}. \tag{1.3}$$

On peut en déduire le résultat asymptotique suivant :

Corollaire 1.1.1. *Soit* (X_t, Y_t) *solution de* (1.2). *On a les convergences ponctuelles suivantes, quand* t *tend vers l'infini :*

- *si b<0 :*
$$\mathbb{E}[X_t] \to -\frac{a}{b} \qquad et \qquad t^{-1}\,\mathbb{E}[Y_t] \to c - \frac{ad}{b},$$
- *si b=0 :*
$$t^{-1}\,\mathbb{E}[X_t] \to a \qquad et \qquad t^{-2}\,\mathbb{E}[Y_t] \to \frac{ad}{2},$$
- *si b>0 :*
$$e^{-bt}\,\mathbb{E}[X_t] \to x_0 + \frac{a}{b} \qquad et \qquad e^{-bt}\,\mathbb{E}[Y_t] \to \frac{d}{b}y_0 + \frac{da}{b^2}.$$

Les résultats concernant le processus CIR (X_t) ont été montrés auparavant dans [22] et plus récemment dans [54]. On a choisi de les mentionner ici pour une présentation plus synthétique.

1.2 Estimation des paramètres du processus CIR

Une fois qu'un modèle a été proposé pour ses bonnes propriétés de modélisation en finance ou en biologie, il est nécessaire de le calibrer. Autrement dit, il faut trouver des valeurs pour les paramètres ajustant bien le modèle aux observations, afin que celui-ci puisse être utilisé ensuite pour faire de la prédiction par exemple. Dans toute cette partie, on cherchera donc à estimer les paramètres du processus à partir de l'observation d'une trajectoire de (X_t) sur l'intervalle de temps $[0, T]$. Pour ce faire, on va considérer l'estimateur du maximum de vraisemblance (EMV) du couple de paramètres (a, b) et on s'intéressera en particulier à son comportement asymptotique.

1.2.1 Estimateur du maximum de vraisemblance

Dans le cas du processus CIR, on trouve les premiers résultats concernant l'EMV $(\widehat{a}_T, \widehat{b}_T)$ de (a, b) et son comportement asymptotique en temps long dans les papiers d'Overbeck [65] et de Fournié et Talay [37]. On commence par introduire quelques notations. Soit τ_0 le temps d'arrêt défini par

$$\tau_0 = \inf\left\{T > 0 \,\middle|\, \int_0^T X_t^{-1}\,\mathrm{d}t = \infty\right\}. \tag{1.4}$$

On note $\mathbb{P}_{a,b}$ la mesure de probabilité associées au processus CIR de paramètres (a, b) . Par une application du théorème de Girsanov [57], on obtient facilement pour le processus CIR que la log-vraisemblance vaut, pour tout $T < \tau_0$,

$$\mathcal{L}_T(a,b) = \log \frac{\mathrm{d}\mathbb{P}_{a,b}}{\mathrm{d}\mathbb{P}_{0,0}} = \frac{1}{4}\int_0^T \frac{a+bX_t}{X_t}\,\mathrm{d}X_t - \frac{1}{8}\int_0^T \frac{(a+bX_t)^2}{X_t}\,\mathrm{d}t. \tag{1.5}$$

Ainsi, pour tout $a > 0$, l'EMV $(\widehat{a}_T, \widehat{b}_T)$ est donné, pour tout $T < \tau_0$, par les valeurs de a et b minimisant $\mathcal{L}_T(a, b)$:

$$\begin{pmatrix} \widehat{a}_T \\ \widehat{b}_T \end{pmatrix} = \Gamma_T^{-1} U_T \tag{1.6}$$

où

$$U_T = \left(\int_0^T X_t^{-1}\,\mathrm{d}X_t, \int_0^T \mathrm{d}X_t \right)^{\intercal} \quad \text{et} \quad \Gamma_T = \begin{pmatrix} \int_0^T X_t^{-1}\,\mathrm{d}t & T \\ T & \int_0^T X_t\,\mathrm{d}t \end{pmatrix}.$$

Remarque 1.2.1. *Si on suppose que l'on connait l'un des deux paramètres, l'estimateur du maximum de vraisemblance pour l'autre paramètre en est grandement simplifié. Ainsi, si b est connu, l'estimateur du maximum de vraisemblance de a est donné par*

$$\bar{a}_T = \frac{\int_0^T X_t^{-1}\,\mathrm{d}X_t - bT}{\int_0^T X_t^{-1}\,\mathrm{d}t} \tag{1.7}$$

tandis que si a est connu, l'EMV pour b vérifie

$$\bar{b}_T = \frac{X_T - x_0 - aT}{\int_0^T X_t\,\mathrm{d}t}. \tag{1.8}$$

On trouve dans [13] un résumé du comportement asymptotique de ces deux estimateurs (convergence en probabilité et en loi) selon les valeurs des paramètres. Les résultats de grandes déviations et déviations modérées que l'on présentera respectivement dans les sections 1.2.4 et 1.2.5 se placent dans ce cadre. Un des objectifs de cette thèse est de les étendre au cas où les deux paramètres a et b sont estimés simultanément.

1.2.2 Convergence presque sûre

Le comportement asymptotique de l'EMV dépend des valeurs des paramètres. On distingue comme attendu une trichotomie, selon les valeurs du paramètre b.

• $b > 0$: cas supercritique

Overbeck [65] a montré que l'EMV $\widehat{b}_T$ de b est fortement consistant, c'est-à-dire qu'il converge presque sûrement vers b lorsque T tend vers l'infini :

$$\widehat{b}_T \to b \quad \text{p.s.}$$

Par contre, *il n'existe pas d'estimateur consistant pour a.* Ceci peut s'expliquer par le fait que le processus croît exponentiellement à la vitesse $\exp(bT)$ tandis que la dépendance en a est seulement linéaire, ce qui rend ce paramètre indiscernable dès que T est suffisamment grand.

• $b < 0$: cas sous-critique

Lorsque $a \geqslant 2$, l'EMV $(\widehat{a}_T, \widehat{b}_T)$ du couple (a, b) est fortement consistant : quand T tend vers l'infini,

$$(\widehat{a}_T, \widehat{b}_T) \to (a, b) \quad \text{p.s.}$$

C'est une conséquence directe de l'ergodicité du processus et de la loi des grands nombres pour les martingales [65].

Par contre, lorsque $a < 2$, le processus visite fréquemment zéro et le temps d'arrêt τ_0 défini par (1.4) est fini presque sûrement. On a une connaissance parfaite de a en temps fini (c'est une fonction de $(X_T)_{T<\tau_0}$). Overbeck [65] propose un autre couple d'estimateurs fortement consistant :

$$\mathbb{1}_{T<\tau_0} \begin{pmatrix} \widehat{a}_T \\ \widehat{b}_T \end{pmatrix} + \mathbb{1}_{\tau_0 \leqslant T} \begin{pmatrix} \lim\limits_{t\uparrow\tau_0} S_t \Sigma_t^{-1} \\ \left(\int_0^T X_s \, \mathrm{d}s\right)^{-1} \left(X_T - T \lim\limits_{t\uparrow\tau_0} S_t \Sigma_t^{-1}\right) \end{pmatrix} \tag{1.9}$$

où $S_t = \int_0^t X_s^{-1} \, \mathrm{d}X_s$ et $\Sigma_t = \int_0^t X_s^{-1} \, \mathrm{d}s$. Cet estimateur naturel n'est pas très facile à manipuler en pratique puisqu'il fait intervenir un temps d'arrêt. On propose dans le Chapitre 4 un autre couple d'estimateurs fortement consistants plus simple à implémenter.

• $b = 0$: cas critique

Pour $a \geqslant 2$, l'EMV $(\widehat{a}_T, \widehat{b}_T)$ du couple (a, b) est faiblement consistant : quand T tend vers l'infini,

$$(\widehat{a}_T, \widehat{b}_T) \xrightarrow{\mathbb{P}} (a, b).$$

Si $a < 2$, comme dans le cas sous-critique, on construit un nouveau couple d'estimateurs à l'aide du temps d'arrêt τ_0, donné par (1.9). Ce couple est faiblement consistant quand T tend vers l'infini.

1.2.3 Normalité asymptotique

Tous les résultats de cette partie sont issus de [14] et [65]. On trouve dans le théorème 3 de [65] des calculs de distributions asymptotiques dans chacun des cas avec des échelles aléatoires, que l'on choisit de ne pas mentionner ici.

• $b < 0$: cas sous-critique

Lorsque $a > 2$, on montre facilement à l'aide du théorème de la limite centrale (TLC) pour les martingales que, pour T tendant vers l'infini,

$$\sqrt{T} \begin{pmatrix} \widehat{a}_T - a \\ \widehat{b}_T - b \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 4\Sigma^{-1}) \quad \text{où} \quad \Sigma = \begin{pmatrix} \frac{-b}{a-2} & 1 \\ 1 & -\frac{a}{b} \end{pmatrix}. \tag{1.10}$$

Pour $a = 2$, on peut montrer la convergence en loi suivante :

$$\begin{pmatrix} T(\widehat{a}_T - a) \\ \sqrt{T}(\widehat{b}_T - b) \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} -b/\tau_1 \\ \sqrt{-2b}\, G \end{pmatrix}$$

où G est une gaussienne centrée réduite indépendante de τ_1, lui-même défini par

$$\tau_1 = \inf \{t > 0 \,|B_t = -b/2\}\,.$$

• $b = 0$ **: cas critique**

On note $(\mathcal{X}_t)$ la solution de (1.1) partant de $x_0 = 0$ et τ_2 le temps d'atteinte défini par

$$\tau_2 = \inf \{t > 0 \,|B_t = 1/2\}\,.$$

Si $a = 2$, l'EMV vérifie la convergence en distribution suivante :

$$\begin{pmatrix} \log T(\widehat{a}_T - a) \\ T(\widehat{b}_T - b) \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} \tau_2^{-1} \\ \frac{a-\mathcal{X}_1}{\int_0^1 \mathcal{X}_t \,\mathrm{d}t} \end{pmatrix},$$

tandis que si $a > 2$, on a

$$\begin{pmatrix} \sqrt{\log T}(\widehat{a}_T - a) \\ T(\widehat{b}_T - b) \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} \sqrt{2(a-2)}\,G \\ \frac{a-\mathcal{X}_1}{\int_0^1 \mathcal{X}_t \,\mathrm{d}t}, \end{pmatrix}$$

où G est une gaussienne centrée réduite indépendante de $(\mathcal{X}_1, \int_0^1 \mathcal{X}_t \,\mathrm{d}t)$.

• $b > 0$ **: cas supercritique**

L'estimateur $\widehat{a}_T$ n'est pas consistant, mais on peut donner un résultat de convergence en loi pour $\widehat{b}_T$:

$$\left(\int_0^T X_t \,\mathrm{d}t\right)^{1/2} (\widehat{b}_T - b) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 4).$$

1.2.4 Grandes déviations

Une fois que la normalité asymptotique d'une suite de variables aléatoires a été établie, il est naturel de se questionner sur l'existence d'un principe de grandes déviations. Pour une introduction très complète aux grandes déviations, le livre de Dembo et Zeitouni [25] est une excellente référence. On commence par rappeler des résultats classiques de la théorie des grandes déviations que l'on utilisera par la suite, puis on présente les principes de grandes déviations obtenus par Zani [73] pour les estimateurs $\bar{a}_T$ et $\bar{b}_T$ définis respectivement par (1.7) et (1.8).

Définition 1.2.1 (Principe de Grandes Déviations)**.** Soit (v_T) une suite réelle positive croissant vers plus l'infini avec T. Une suite (Z_T) de vecteurs aléatoires de $\mathbb{R}^d$ satisfait un *principe de grandes déviations* de vitesse v_T et de *fonction de taux* I si I est une fonction de $\mathbb{R}^d$ dans $[0, +\infty]$ semi-continue inférieurement (i.e. ses ensembles de niveau sont fermés) telle que (Z_T) vérifie les deux propriétés suivantes :

◇ borne supérieure : pour tout fermé F de $\mathbb{R}^d$

$$\limsup_{T \to +\infty} v_T^{-1} \log \mathbb{P}\,(Z_T \in F) \leqslant -\inf_{z \in F} I\,(z),$$

◇ borne inférieure : pour tout ouvert G de $\mathbb{R}^d$

$$\liminf_{T\to+\infty} v_T^{-1} \log \mathbb{P}\left(Z_T \in G\right) \geqslant - \inf_{z\in G} I\left(z\right).$$

De plus, si les ensembles de niveau de la fonction de taux I sont compacts, on dit que I est une *bonne* fonction de taux.

On introduit quelques notations avant d'énoncer un théorème fondamental qui étend le théorème de Cramer (reliant PGD et transformée de Laplace) au cadre non indépendant. Cette généralisation est due aux travaux successifs de Gärtner [46] et Ellis [32].

Soit (Z_T) une suite de vecteurs aléatoires de $\mathbb{R}^d$. La fonction génératrice des moments Λ_T de Z_T est définie, pour tout $\lambda \in \mathbb{R}^d$, par

$$\Lambda_T\left(\lambda\right) = \log \mathbb{E}\left[e^{\langle \lambda, Z_T\rangle}\right],$$

où $\langle .,.\rangle$ désigne le produit scalaire sur $\mathbb{R}^d$. On suppose que, pour tout $\lambda \in \mathbb{R}^d$, la limite

$$\Lambda\left(\lambda\right) = \lim_{T\to\infty} v_T^{-1}\, \Lambda_T\left(v_T \lambda\right)$$

existe comme un nombre réel étendu. On note $\mathcal{D}_\Lambda$ son domaine de finitude

$$\mathcal{D}_\Lambda = \left\{\lambda \in \mathbb{R}^d | \Lambda(\lambda) < +\infty\right\}.$$

On remarque qu'en tant que limite de fonction convexe, Λ est elle-même une fonction convexe. On appelle Λ^* la transformée de Fenchel-Legendre de Λ, définie pour tout $z \in \mathbb{R}^d$ par

$$\Lambda^*(z) = \sup_{\lambda\in\mathbb{R}^d} \left\{\langle z, \lambda\rangle - \Lambda\left(\lambda\right)\right\}. \tag{1.11}$$

Définition 1.2.2. Soit $z \in \mathbb{R}^d$. z est appelé *point exposé* de Λ^* s'il existe $\lambda \in \mathbb{R}^d$ tel que, pour tout $y \in \mathbb{R}^d \backslash \{z\}$,

$$\langle \lambda, z\rangle - \Lambda^*\left(z\right) > \langle \lambda, y\rangle - \Lambda^*\left(y\right).$$

Un tel λ est alors appelé *hyperplan d'exposition.*

Définition 1.2.3. On dit que Λ est *escarpée* si l'intérieur de $\mathcal{D}_\Lambda$ est non vide, si elle est différentiable sur l'intérieur de $\mathcal{D}_\Lambda$ et si la norme de son gradient tend vers l'infini pour toute suite de points de l'intérieur de $\mathcal{D}_\Lambda$ qui converge vers un point du bord.

Théorème 1.2.1 (Théorème de Gärtner-Ellis)**.** *Si la limite Λ existe et si l'origine appartient à l'intérieur de son domaine de finitude $\mathcal{D}_\Lambda$, alors la famille (Z_T) satisfait un PGD faible :*

◇ *pour tout fermé F de $\mathbb{R}^d$*

$$\limsup_{T\to+\infty} v_T^{-1} \log \mathbb{P}\left(Z_T \in F\right) \leqslant - \inf_{z\in F} \Lambda^*\left(z\right),$$

- *pour tout ouvert G de $\mathbb{R}^d$*

$$\liminf_{T\to+\infty} v_T^{-1} \log \mathbb{P}\left(Z_T \in G\right) \geqslant -\inf_{z\in G\cap\mathcal{F}} \Lambda^*\left(z\right),$$

  où Λ^* est donnée par (1.11) et $\mathcal{F}$ est l'ensemble des points exposés de Λ^* dont l'hyperplan d'exposition est inclus dans l'intérieur de $\mathcal{D}_\Lambda$.

De plus, si Λ est semi-continue inférieurement et escarpée, alors la famille (Z_T) satisfait un PGD de vitesse v_T et de bonne fonction de taux Λ^*.

Remarque 1.2.2. *Dans la preuve de la borne inférieure, on veut effectuer, autour de chaque point z du domaine de finitude de Λ, un changement de probabilité de la forme* $\exp\left(-v_T\left\langle z,\lambda_z\right\rangle - v_T\Lambda(\lambda_z)\right)$ *où λ_z vérifie $\nabla\Lambda(\lambda_z) = z$. Le fait que la fonction Λ soit escarpée permet de justifier qu'un tel λ_z existe pour tout point du domaine de finitude de Λ et non pas uniquement sur l'ensemble $\mathcal{F}$ des points exposés de Λ^*.*

Remarque 1.2.3. *Il existe de nombreux cas où un PGD est satisfait bien que la fonction Λ ne soit pas escarpée.*

On donne dans le lemme suivant un outil très pratique qui permet de "transporter" des PGD à la façon de la δ-méthode pour le TLC.

Lemme 1.2.1 (Principe de contraction)**.** *Soit (Z_T) une suite de vecteurs aléatoires de $\mathbb{R}^d$ qui satisfait un PGD de vitesse v_T et de bonne fonction de taux I et soit $g : \mathbb{R}^d \to \mathbb{R}^p$ une fonction continue sur $\mathcal{D}_I = \left\{x \in \mathbb{R}^d | I(x) < +\infty\right\}$. Alors, la suite $(g(Z_T))$ satisfait un PGD de vitesse v_T et de bonne fonction de taux J définie pour tout $y \in \mathbb{R}^p$ par*

$$J\left(y\right) = \inf_{\{x\in\mathcal{D}_I | g(x)=y\}} I\left(x\right),$$

où l'infimum sur l'ensemble vide est supposé infini.

On revient maintenant à l'estimation des paramètres a et b du processus CIR. On se place dans le cadre de la remarque 1.2.1 : on estime chacun des paramètres successivement en supposant que l'autre est connu. Dans [73], Zani établit un PGD pour $\bar{a}_T$ en supposant que $a > 2$ et que $b \leqslant 0$ est connu. Pour $b < 0$, elle obtient un PGD pour $\bar{b}_T$ lorsque a est connu et strictement positif, ainsi que des grandes déviations précises. On notera que dans le cas particulier où $a = 1$, les résultats de [73] découlent de ceux prouvés auparavant dans [19] pour le processus d'Ornstein-Uhlenbeck.

Théorème 1.2.2. *Soit $b \leqslant 0$ donné et soit $a > 2$ le paramètre à estimer.*

- *Si $b = 0$, la suite $(\bar{a}_T)$ satisfait un PGD de vitesse $\log T$ et de bonne fonction de taux I_a^0 donnée pour tout $\alpha \in \mathbb{R}$ par*

$$I_a^0(\alpha) = \begin{cases} \dfrac{(\alpha-a)^2}{8(\alpha-2)} & \text{si } \alpha \geqslant \alpha_1, \\ 2 - \dfrac{\alpha}{2} + \dfrac{1}{2}\sqrt{(\alpha-a)^2 - 8(a-2)} & \text{si } \alpha < \alpha_1, \end{cases} \tag{1.12}$$

  où $\alpha_1 = \frac{1}{3}\left(-a + 4 + 2\sqrt{(a-2)^2 + 2a}\right)$.

◇ *Si $b < 0$, la suite $(\bar{a}_T)$ satisfait un PGD de vitesse T et de bonne fonction de taux I_a qui vérifie, pour tout $\alpha \in \mathbb{R}$,*

$$I_a(\alpha) = -bI_a^0(\alpha)$$

où I_a^0 est donnée par (1.12).

Remarque 1.2.4. *Le Théorème 3.1 de [73] donne en fait un PGD pour $\nu = (a-2)/2$. On se ramène facilement au résultat énoncé ci-dessus en utilisant le principe de contraction (voir Lemme 1.2.1). On peut aussi noter l'oubli du facteur $-b$ lorsque $b < 0$.*

Théorème 1.2.3. *Soit $a > 0$ donné et soit $b < 0$ le paramètre à estimer. La suite $(\bar{b}_T)$ satisfait un PGD de vitesse T et de bonne fonction de taux I_b donnée pour tout $\beta \in \mathbb{R}$ par*

$$I_b(\beta) = \begin{cases} -\dfrac{a\beta}{8}\left(1 - \dfrac{b}{\beta}\right)^2 & \text{si } \beta \leqslant \frac{b}{3}, \\ \dfrac{a}{2}\left(2\beta - b\right) & \text{si } \beta \geqslant \frac{b}{3}. \end{cases} \tag{1.13}$$

On termine cette partie en donnant les résultats de grandes déviations précises au premier ordre obtenus pour $(\bar{b}_T)$ dans [73].

Théorème 1.2.4. *Soit $a > 0$ connu et soit $b < 0$ le paramètre à estimer. Lorsque T tend vers l'infini, la suite $(\bar{b}_T)$ vérifie :*

◇ *Pour tout $\beta > \frac{b}{3}$,*

$$\mathbb{P}(\bar{b}_T \geqslant \beta) = T^{a/2-1} \exp\left[-TI_b(\beta) + K_1(\beta)\right] K_2(\beta) \left(1 + o(1)\right)$$

où I_b est donnée par (1.13),

$$K_1(\beta) = \frac{a}{2}\log\left(\frac{(2\beta - b)(3\beta - b)}{2(\beta - b)}\right) \quad \text{et} \quad K_2(\beta) = \frac{2(2\beta - b)}{\Gamma(a/2)(3\beta - b)(\beta - b)}.$$

◇ *Pour $\beta = \frac{b}{3}$,*

$$\mathbb{P}(\bar{b}_T \geqslant \frac{b}{3}) = T^{(a-2)/4} \exp\left[-TI_b(b/3)\right] (-b)^{(a-2)/4}\, C_a \left(1 + o(1)\right)$$

où $C_a = 2^{-a/2-1}\, 3^{(2-a)/4}\, a^{(a-2)/4} \left(\Gamma\left((a+2)/4\right)\right)^{-1}$.

◇ *Pour tout $b < \beta < \frac{b}{3}$,*

$$\mathbb{P}(\bar{b}_T \geqslant \beta) = T^{-1/2} \frac{\exp\left[-TI_b(\beta) + K_3(\beta)\right]}{\sqrt{2\pi a}\sigma(\beta)\phi(\beta)} \left(1 + o(1)\right)$$

où $\sigma^2(\beta) = -\beta^{-1}$,

$$\phi(\beta) = \frac{\beta^2 - b^2}{4\beta} \quad \text{et} \quad K_3(\beta) = -\frac{a}{2}\log\left(\frac{(\beta + b)(3\beta - b)}{4\beta^2}\right).$$

◇ *Pour tout $\beta < b$, avec les notations précédentes,*

$$\mathbb{P}(\bar{b}_T \leqslant \beta) = T^{-1/2} \frac{\exp\left[-TI_b(\beta) + K_3(\beta)\right]}{\sqrt{2\pi a}\sigma(\beta)(-\phi(\beta))} \left(1 + o(1)\right).$$

1.2.5 Déviations modérées

Dans certaines situations, la fonction de taux du PGD est difficile à évaluer en pratique et peut ne pas être explicite. Il arrive aussi qu'un PGD soit compliqué à déterminer voire qu'il n'existe pas. On peut alors établir un résultat intermédiaire : le principe de déviations modérées (PDM). On trouvera dans [71] d'autres intérêts des PDM, comme leur lien avec les lois du logarithme itéré ou leurs applications à la detection de rupture.

Définition 1.2.4. Soient (v_T) et (λ_T) deux suites réelles positives croissant vers l'infini avec T et telles que $\lambda_T = o(v_T)$. On dit qu'une suite (Z_T) de vecteurs aléatoires de $\mathbb{R}^d$ satisfait un *principe de déviations modérées* de vitesse λ_T et de fonction de taux I, si la suite $(\sqrt{v_T/\lambda_T}\, Z_T)$ satisfait un PGD de vitesse λ_T et de fonction de taux I.

Ce résultat donne, certes, des convergences exponentielles moins rapides qu'un PGD, mais il a le mérite d'exister et d'avoir une fonction de taux souvent simple à manipuler, quadratique en général du fait de sa proximité avec le TLC (c'est le cas par exemple des résultats de déviations modérées que l'on obtient dans le Chapitre 3 de cette thèse).

On trouve dans [38] des résultats de déviations modérées pour l'EMV $\bar{a}_T$ de a connaissant b et pour l'EMV $\bar{b}_T$ de b connaissant a. Soient (λ_T) et (γ_T) deux suites croissantes positives telles que pour T allant vers l'infini

$$\lambda_T \to +\infty\ ,\ \gamma_T \to +\infty\ ,\ \frac{\lambda_T}{T} \to 0 \quad \text{et} \quad \frac{\gamma_T}{\log T} \to 0.$$

Théorème 1.2.5. *Soit $b \leqslant 0$ donné et soit $a > 2$ le paramètre à estimer.*

- *Si $b < 0$, la suite $\sqrt{\frac{T}{\lambda_T}}(\bar{a}_T - a)$ satisfait un PGD de vitesse λ_T et de bonne fonction de taux J_a, donnée pour tout $\alpha \in \mathbb{R}$, par*

$$J_a(\alpha) = -\frac{\alpha^2}{b(a-2)}.$$

- *Si $b = 0$, la suite $\sqrt{\frac{\log T}{\gamma_T}}(\bar{a}_T - a)$ satisfait un PGD de vitesse γ_T et de bonne fonction de taux $\widetilde{J}_a$ donnée, pour tout $\alpha \in \mathbb{R}$, par*

$$\widetilde{J}_a(\alpha) = \frac{\alpha^2}{2(a-2)}.$$

Théorème 1.2.6. *Soit $a > 0$ connu et soit $b < 0$ le paramètre à estimer. La suite $\sqrt{\frac{T}{\lambda_T}}(\bar{b}_T - b)$ satisfait un PGD de vitesse λ_T et de bonne fonction de taux J_b donnée, pour tout $\beta \in \mathbb{R}$, par*

$$J_b(\beta) = -\frac{\beta^2}{2ab}.$$

1.3 Estimation des paramètres du processus de Heston

On souhaite estimer les paramètres du processus donné par (1.2) à partir de l'observation d'une trajectoire de (X_t) et de (Y_t) sur l'intervalle de temps $[0, T]$. On considère l'estimateur du maximum de vraisemblance (EMV) du quadruplet de paramètres $\theta = (a, b, c, d)$. On n'estime pas le coefficient de corrélation ρ, puisqu'on peut l'obtenir à partir des deux trajectoires grâce à la variation croisée de (X_t) et (Y_t), qui vérifie, pour tout $T > 0$:

$$\langle X, Y\rangle_T = 4\rho \int_0^T X_t \, \mathrm{d}t.$$

1.3.1 Estimateur du maximum de vraisemblance

L'EMV $\widehat{\theta}_T$ du quadruplet θ a été étudié très récemment par Barczy et Pap dans [11]. On résume ici les résultats qu'ils ont obtenus.

Soit τ_0 le temps d'arrêt défini par (1.4) et $\mathbb{P}_{a,b}^{c,d}$ la mesure de probabilité associée au processus de Heston de paramètres $\theta = (a, b, c, d)$. On trouve dans [11] que la log-vraisemblance $\mathcal{L}_T(\theta)$ vérifie, pour tout $T < \tau_0$,

$$\begin{aligned}(1-\rho^2)\mathcal{L}_T(\theta) = & \int_0^T ((a+bX_t) - \rho(c+dX_t)) \frac{\mathrm{d}X_t}{4X_t} + \int_0^T ((c+dX_t) - \rho(a+bX_t)) \frac{\mathrm{d}Y_t}{4X_t} \\ & - \int_0^T \frac{1}{8X_t} \left((a+bX_t)^2 - 2\rho(a+bX_t)(c+dX_t) + (c+dX_t)^2\right) \mathrm{d}t.\end{aligned}$$

L'EMV $\widehat{\theta}_T$ est donc donné, pour tout $T < \tau_0$, par

$$\widehat{\theta}_T = \begin{pmatrix} \Gamma_T^{-1} & 0 \\ 0 & \Gamma_T^{-1} \end{pmatrix} \begin{pmatrix} U_T \\ V_T \end{pmatrix} \tag{1.14}$$

où $V_T = \left(\int_0^T X_t^{-1} \, \mathrm{d}Y_t, \int_0^T \mathrm{d}Y_t\right)^{\intercal}$, et U_T et Γ_T sont définies après l'équation (1.6). On remarque que les deux premiers estimateurs coincident avec l'EMV des paramètres (a, b) du processus CIR obtenu à partir de l'observation d'une trajectoire de (X_t) uniquement donné par (1.6). On note aussi que l'expression des estimateurs ne fait pas intervenir la valeur de ρ.

1.3.2 Convergence presque sûre

Comme pour le processus CIR, la trichotomie selon les valeurs du paramètre b va intervenir. Dans toute cette partie et la suivante, on ne met pas de condition sur les valeurs de c et d, qui sont donc réels quelconques.

• $b > 0$: cas supercritique

L'EMV $\widehat{b}_T$ de b est fortement consistant, c'est-à-dire qu' il converge presque sûrement vers b lorsque T tend vers l'infini :

$$\widehat{b}_T \to b \quad \text{p.s.}$$

L'EMV $\widehat{d}_T$ de d est faiblement consistant, il converge en probabilité vers d quand T tend vers l'infini :

$$\widehat{d}_T \xrightarrow{\mathbb{P}} d.$$

Par contre, $\widehat{a}_T$ et $\widehat{c}_T$ ne sont pas consistants.

• $b < 0$: cas sous-critique

Lorsque $a > 2$, l'EMV $\widehat{\theta}_T$ du quadruplet θ est fortement consistant : quand T tend vers l'infini,

$$\widehat{\theta}_T \to \theta \quad \text{p.s.}$$

Dans le cas $a = 2$, il est faiblement consistant et la convergence ci-dessus est vérifiée en probabilité.

• $b = 0$: cas critique

Pour $a > 2$, l'EMV $\widehat{\theta}_T$ est faiblement consistant : quand T tend vers l'infini,

$$\widehat{\theta}_T \xrightarrow{\mathbb{P}} \theta.$$

Remarque 1.3.1. *Pour $a > 2$, on retrouve les résultats de consistance obtenus pour l'EMV du CIR dans la section précédente. Les autres cas ne sont que partiellement traités et ne permettent pas de faire des comparaisons.*

1.3.3 Normalité asymptotique

• $b < 0$: cas sous-critique

Lorsque $a > 2$, comme dans le cas du processus CIR, on montre facilement à l'aide du Théorème de la Limite Centrale (TLC) pour les martingales que, pour T tendant vers l'infini,

$$\sqrt{T}\left(\widehat{\theta}_T - \theta\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 4R \otimes \Sigma^{-1})$$

où $\otimes$ désigne le produit de Kronecker, Σ est définie dans (1.10) et R est la matrice donnée par

$$R = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}. \tag{1.15}$$

• $b = 0$: cas critique

On note $(\mathcal{X}_t, \mathcal{Y}_t)$ la solution de (1.2) partant de $(x_0, y_0) = (0, 0)$. Si $a > 2$, on a la convergence en distribution suivante :

$$\begin{pmatrix} \sqrt{\log T}(\widehat{a}_T - a) \\ \sqrt{\log T}(\widehat{c}_T - c) \\ T\widehat{b}_T \\ T(\widehat{d}_T - d) \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} 2\sqrt{a-2}\, R^{1/2}\, G \\ \frac{a - \mathcal{X}_1}{\int_0^1 \mathcal{X}_t \,\mathrm{d}t} \\ \frac{c - \mathcal{Y}_1}{\int_0^1 \mathcal{X}_t \,\mathrm{d}t}, \end{pmatrix}$$

où R est la matrice donnée dans (1.15) et G est une gaussienne bidimensionnelle centrée réduite indépendante de $(\mathcal{X}_1, \int_0^1 \mathcal{X}_t \,\mathrm{d}t, \mathcal{Y}_1)$.

• $b > 0$ **: cas supercritique**

Les estimateurs $\widehat{a}_T$ et $\widehat{c}_T$ n'étant pas consistants, on donne un résulat de convergence en loi pour $(\widehat{b}_T, \widehat{d}_T)$ uniquement. Pour tout $a \geqslant 2$,

$$\left(\int_0^T X_t \,\mathrm{d}t\right)^{1/2} \begin{pmatrix} \widehat{b}_T - b \\ \widehat{d}_T - d \end{pmatrix} \xrightarrow{\mathcal{L}} R^{1/2}\, G,$$

où R est donnée dans (1.15) et G est une gaussienne bidimensionnelle centrée réduite.

1.4 Aperçu des principaux résultats de la thèse

Les trois chapitres qui suivent présentent les résultats obtenus au cours de la thèse, issus respectivement de [27], [28] et [26].

1.4.1 Grandes déviations pour le processus CIR

Dans un premier temps, on étend les résultats de grandes déviations de Zani [73] au cas où les deux paramètres du processus CIR sont estimés simultanément (on ne suppose plus que l'un des paramètres est connu pour estimer l'autre), ce qui, on l'a vu, complique grandement l'expression des estimateurs. On établit ainsi dans le Chapitre 2 un principe de grandes déviations pour le couple $(\widehat{a}_T, \widehat{b}_T)$, dans le cas super-critique $(b < 0)$ pour $a > 2$. La méthode utilisée pour obtenir les résultats de grandes déviations est différente de celle de Zani et repose uniquement sur des outils standards de grandes déviations : le théorème de Gärtner-Ellis et le principe de contraction, que l'on a rappelés ci-dessus. En effet, on remarque que les estimateurs se réécrivent, en utilisant la formule d'Itô pour $\log X_T$ ainsi que l'équation (1.1), sous la forme

$$\widehat{a}_T = \frac{S_T\,(2\,\Sigma_T + L_T) - \frac{X_T - x_0}{T}}{V_T} \quad \text{et} \quad \widehat{b}_T = \frac{(\frac{X_T - x_0}{T} - 2)\,\Sigma_T - L_T}{V_T}$$

où le dénominateur $V_T = S_T\,\Sigma_T - 1$ avec

$$S_T = \frac{1}{T}\int_0^T X_t \,\mathrm{d}t \quad \text{et} \quad \Sigma_T = \frac{1}{T}\int_0^T \frac{1}{X_t}\,\mathrm{d}t,$$

et

$$L_T = \frac{\log X_T - \log x_0}{T}.$$

Les estimateurs s'expriment donc en fonction d'un quadruplet de fonctionnelles du processus : $(X_T/T, S_T, \Sigma_T, L_T)$. Une idée naturelle est d'établir un PGD pour ce quadruplet grâce au théorème de Gärtner-Ellis, puis d'en déduire un PGD pour les estimateurs en utilisant le principe de contraction. On montre que la log-Laplace pour ce quadruplet, convenablement renormalisée, admet une limite qui n'est pas escarpée. Pour contourner ce problème, on s'inspire des idées de [17] pour l'étude du PGD de l'EMV des deux paramètres d'un processus d'Ornstein-Uhlenbeck classique avec shift et on considère plutôt le quadruplet $\mathcal{Q}_T = \left(\sqrt{X_T/T}, S_T, \Sigma_T, \mathcal{L}_T\right)$, où

$$\mathcal{L}_T = -\sqrt{\frac{-\log X_T}{T}}\mathbf{1}_{X_T<1} + \frac{\log X_T}{T}\mathbf{1}_{X_T\geqslant 1}.$$

On obtient le résultat suivant

Proposition 1.4.1. *Soit $\Lambda_T(\lambda,\mu,\nu,\gamma)$ la log-Laplace renormalisée du quadruplet $\mathcal{Q}_T$ donnée sur $\mathbb{R}^4$ par*

$$\Lambda_T(\lambda,\mu,\nu,\gamma) = \frac{1}{T}\log\left(\mathbb{E}\left[\exp\left(\lambda\sqrt{T}\sqrt{X_T} + \gamma\, T\mathcal{L}_T + \mu\int_0^T X_t\,\mathrm{d}t + \nu\int_0^T \frac{1}{X_t}\,\mathrm{d}t\right)\right]\right).$$

Soit Λ sa limite ponctuelle quand T tend vers l'infini.

Pour tous $\lambda,\gamma \in \mathbb{R}$, $\mu < \frac{b^2}{8}$ et $\nu < \frac{(a-2)^2}{8}$,

$$\Lambda(\lambda,\mu,\nu,\gamma) = \begin{cases} -\dfrac{d}{2}(1+f) - \dfrac{ab}{4} + \dfrac{\lambda^2}{d-b} & \text{si } \lambda > 0 \text{ et } \gamma \geqslant 0 \\ & \quad \text{ou si } \gamma < 0, \lambda > 0 \text{ et } \frac{\gamma^2}{\lambda^2} < \frac{2f+a+2}{d-b}, \\ -\dfrac{d}{2}(1+f) - \dfrac{ab}{4} + \dfrac{\gamma^2}{2f+a+2} & \text{si } \lambda \leqslant 0 \text{ et } \gamma < 0 \\ & \quad \text{ou si } \gamma < 0, \lambda > 0 \text{ et } \frac{\gamma^2}{\lambda^2} \geqslant \frac{2f+a+2}{d-b}, \\ -\dfrac{d}{2}(1+f) - \dfrac{ab}{4} & \text{si } \lambda \leqslant 0 \text{ et } \gamma \geqslant 0, \end{cases}$$

où $d = \sqrt{b^2 - 8\mu}$ et $f = \frac{1}{2}\sqrt{(a-2)^2 - 8\nu}$.

La fonction Λ est escarpée. Le théorème de Gärtner-Ellis assure alors l'existence d'un PGD de vitesse T pour le quadruplet $\mathcal{Q}_T$, ainsi que pour le couple d'estimateurs $(\widehat{a}_T, \widehat{b}_T)$ par une application du principe de contraction. Cependant, comme il est impossible d'effectuer explicitement les différentes optimisations intervenant dans le calcul de la fonction de taux de ce PGD, on est amené à introduire deux couples d'estimateurs simplifiés $(\widetilde{a}_T, \widetilde{b}_T)$ et $(\breve{a}_T, \breve{b}_T)$ construits à partir de l'EMV en supprimant alternativement les termes en L_T et en X_T/T (qui tendent presque sûrement vers zéro quand T tend vers l'infini). Ces deux couples sont définis comme suit :

$$\widetilde{a}_T = \frac{2\,S_T\,\Sigma_T - \frac{X_T}{T}}{V_T} \quad \text{et} \quad \widetilde{b}_T = \frac{(\frac{X_T}{T} - 2)\,\Sigma_T}{V_T},$$

et

$$\breve{a}_T = \frac{S_T\,(2\,\Sigma_T + L_T)}{V_T} \quad \text{et} \quad \breve{b}_T = \frac{-2\,\Sigma_T - L_T}{V_T}.$$

Ces deux couples d'estimateurs sont fortement consistants et on montre qu'ils vérifient le même théorème de la limite centrale que l'EMV. Des principes de grandes déviations sont établis pour ces deux couples d'estimateurs simplifiés. Pour le premier couple, on obtient le résultat suivant.

Théorème 1.4.1. *Le couple* $(\widetilde{a}_T, \widetilde{b}_T)$ *satisfait un PGD de vitesse* T *et de bonne fonction de taux* $J_{a,b}$ *donnée sur* $\mathbb{R}^2$ *par*

$$J_{a,b}(\alpha,\beta) = \begin{cases} \dfrac{(a-2)^2\beta}{8(2-\alpha)}\left(1+\dfrac{(2-\alpha)b}{\beta(a-2)}\right)^2 + 2\beta - b & \text{si } \alpha > 2, \frac{b}{3} \leqslant \beta < 0 \\ & \text{ou si } \alpha < 2, \beta > 0, \\ \dfrac{(a-2)^2\beta}{8(2-\alpha)}\left(1+\dfrac{(2-\alpha)b}{\beta(a-2)}\right)^2 - \dfrac{\beta}{4}\left(1-\dfrac{b}{\beta}\right)^2 & \text{si } \alpha > 2, \beta \leqslant \frac{b}{3}, \\ -b & \text{si } (\alpha,\beta) = (2,0), \\ +\infty & \text{sinon.} \end{cases}$$

En utilisant de nouveau le principe de contraction, on en déduit un PGD pour chacun des estimateurs :

Corollaire 1.4.1. *La suite* $(\widetilde{a}_T)$ *satisfait un PGD de vitesse* T *et bonne fonction de taux*

$$J_a(\alpha) = \begin{cases} \dfrac{b}{4}\left(a - 6 - \sqrt{(a-2)^2 + 16(2-\alpha)}\right) & \text{si } \alpha \leqslant \ell_a, \\ \dfrac{b}{4}\left(a - \sqrt{\alpha\left(\dfrac{(a-2)^2}{\alpha-2} + 2\right)}\right) & \text{si } \alpha \geqslant \ell_a, \end{cases}$$

où $\ell_a = \frac{10}{9} + \frac{1}{9}\sqrt{64 + 9(a-2)^2}$.

Corollaire 1.4.2. *La suite* $(\widetilde{b}_T)$ *satisfait un PGD de vitesse* T *et de bonne fonction de taux*

$$J_b(\beta) = \begin{cases} -\dfrac{\beta}{4}\left(1-\dfrac{b}{\beta}\right)^2 & \text{si } \beta \leqslant \frac{b}{3}, \\ 2\beta - b & \text{si } \beta \geqslant \frac{b}{3}. \end{cases}$$

De la même manière, on obtient les trois résultats suivants pour le second couple d'estimateurs.

Théorème 1.4.2. *Le couple* $(\breve{a}_T, \breve{b}_T)$ *satisfait un PGD de vitesse* T *et de bonne fonction de taux* $K_{a,b}$ *définie sur* $\mathbb{R}^2$ *par*

$$K_{a,b}(\alpha,\beta) = \begin{cases} \dfrac{a}{4}(b-\beta) - \dfrac{\alpha}{8\beta}\left(b^2-\beta^2\right) - \dfrac{\beta}{\alpha}\left(\sqrt{2}+\sqrt{C_\alpha}\right)^2 & \text{si } \beta < 0, 0 < \alpha \leqslant \alpha_a \\ & \text{ou si } \beta > 0, \alpha < 0, \\ \dfrac{a}{4}(b-\beta) - \dfrac{\alpha}{8\beta}\left(b^2-\beta^2\right) - \dfrac{\beta\,(a-\alpha)^2}{8\,(\alpha-2)} & \text{si } \beta < 0, \alpha \geqslant \alpha_a, \\ -\dfrac{b}{4}\left(4 - a + \sqrt{a^2+16}\right) & \text{si } (\alpha,\beta) = (0,0), \\ +\infty & \text{sinon,} \end{cases}$$

où $C_\alpha = \frac{1}{8}(a-\alpha)^2 + 2 - \alpha$ *et* $\alpha_a = -\frac{2}{3}\left(\frac{a}{2} - 2 - \sqrt{a^2 - 2a + 4}\right)$.

Corollaire 1.4.3. *La suite $(\breve{a}_T)$ satisfait un PGD de vitesse T et de bonne fonction de taux*

$$K_a(\alpha) = \begin{cases} -\dfrac{b}{4}\left(4 - a + \sqrt{a^2+16}\right) & \text{si } \alpha = 0, \\ K_{a,b}(\alpha, \beta_b) & \text{si } \alpha < \alpha_a, \alpha \neq 0, \\ \dfrac{b}{4}\left(a - \sqrt{\alpha\left(\dfrac{(a-2)^2}{\alpha-2} + 2\right)}\right) & \text{si } \alpha \geqslant \alpha_a \end{cases}$$

où $\beta_b = b\alpha\left(16\sqrt{2C_\alpha} + a^2 - 8\alpha + 32\right)^{-1/2}$ et α_a, C_α sont définis dans le théorème précédent.

Corollaire 1.4.4. *La suite $(\breve{b}_T)$ satisfait un PGD de vitesse T et de bonne fonction de taux*

$$K_b(\beta) = \inf\left\{K_{a,b}(\alpha,\beta) \ / \ \alpha \in \mathbb{R}\right\}.$$

En particulier, on a $K_b(0) = K_{a,b}(0,0) = -\frac{b}{4}\left(4 - a + \sqrt{a^2+16}\right)$.

Enfin, en mélangeant les deux couples d'estimateurs précédents, on construit un troisième couple d'estimateurs exponentiellement équivalent à l'EMV, qui partage donc le même PGD. On est capable de calculer la fonction de taux associée au PGD de ce troisième couple d'estimateurs. On obtient ainsi le résultat suivant, qui répond à la question initiale :

Théorème 1.4.3. *Le couple $(\widehat{a}_T, \widehat{b}_T)$ satisfait un PGD de vitesse T et de bonne fonction de taux $I_{a,b}$ donnée sur $\mathbb{R}^2$ par*

$$I_{a,b}(\alpha,\beta) = \min\left(J_{a,b}(\alpha,\beta), K_{a,b}(\alpha,\beta)\right).$$

On en déduit immédiatemment un principe de grandes déviations pour chacun des estimateurs $\widehat{a}_T$ et $\widehat{b}_T$ grâce au principe de contraction.

Corollaire 1.4.5. *La suite $(\widehat{a}_T)$ satisfait un PGD de vitesse T et de bonne fonction de taux*

$$I_a(\alpha) = \min\left(J_a(\alpha), K_a(\alpha)\right).$$

Corollaire 1.4.6. *La suite $(\widehat{b}_T)$ satisfait un PGD de vitesse T et de bonne fonction de taux*

$$I_b(\beta) = \min\left(J_b(\beta), K_b(\beta)\right).$$

Par ailleurs, on peut retrouver les principes de grandes déviations obtenus dans [73] par notre méthode (voir section 2.6, qui est un ajout par rapport à l'article original [27]).

1.4.2 Déviations modérées pour les processus CIR et Heston

Dans un second temps, on étend les résultats de déviations modérées obtenus dans [38] au cas où les paramètres sont estimés simultanément, toujours pour $b < 0$ et $a > 2$. On ne considère plus uniquement le processus CIR et on établit un principe de déviations modérées pour l'EMV $\widehat{\theta}_T$ du quadruplet de paramètres du processus de Heston ainsi que pour l'EMV $(\widehat{a}_T, \widehat{b}_T)$ du couple de paramètres du CIR.

Soit $(\lambda_T)_T$ une suite croissante positive telle que pour T allant vers l'infini

$$\lambda_T \to +\infty \quad \text{et} \quad \frac{\lambda_T}{T} \to 0.$$

On obtient dans le Chapitre 3 les deux résultats suivants :

Théorème 1.4.4. *La suite* $\left(\sqrt{\frac{T}{\lambda_T}}(\widehat{a}_T - a, \widehat{b}_T - b)\right)$ *satisfait un PGD de vitesse* λ_T *et de bonne fonction de taux* $I_{a,b}$ *donnée pour tout* $(\alpha, \beta) \in \mathbb{R}^2$ *par*

$$I_{a,b}(\alpha, \beta) = -\frac{b}{8(a-2)}\alpha^2 - \frac{a}{8b}\beta^2 + \frac{\alpha\beta}{4}. \tag{1.16}$$

Théorème 1.4.5. *La suite* $\left(\sqrt{\frac{T}{\lambda_T}}(\widehat{\theta}_T - \theta)\right)$ *satisfait un PGD de vitesse* λ_T *et de bonne fonction de taux* I_θ *donnée pour tout* $(\alpha, \beta, \gamma, \delta) \in \mathbb{R}^4$ *par*

$$I_\theta(\alpha, \beta, \gamma, \delta) = \left(1 - \rho^2\right)^{-1} \left(I_{a,b}\left(\alpha, \beta\right) + I_{a,b}\left(\gamma, \delta\right) + \rho\, J\left(\alpha, \beta, \gamma, \delta\right)\right).$$

où $I_{a,b}$ *est définie par* (1.16) *et*

$$J\left(\alpha, \beta, \gamma, \delta\right) = -\alpha\delta + \frac{b}{a-2}\alpha\gamma - \beta\gamma + \frac{a}{b}\beta\delta.$$

La démonstration repose sur le théorème de Gärtner-Ellis et sur des résultats élémentaires de la théorie des martingales à temps continu. En effet, en utilisant (1.2), on peut réécrire les estimateurs (1.14) sous la forme suivante

$$\widehat{\theta}_T = \theta + 2\begin{pmatrix} \langle M \rangle_T^{-1} & 0 \\ 0 & \langle M \rangle_T^{-1} \end{pmatrix} \begin{pmatrix} M_T \\ N_T \end{pmatrix},$$

où M_T et N_T sont des martingales à temps continu données respectivement par

$$M_T = \left(\int_0^T X_t^{-1/2}\,\mathrm{d}B_t, \int_0^T X_t^{1/2}\,\mathrm{d}B_t\right)^{\intercal} \quad \text{et} \quad N_T = \left(\int_0^T X_t^{-1/2}\,\mathrm{d}\widetilde{B}_t, \int_0^T X_t^{1/2}\,\mathrm{d}\widetilde{B}_t\right)^{\intercal}$$

avec $\mathrm{d}\widetilde{B}_t = \rho\,\mathrm{d}B_t + \sqrt{1-\rho^2}\,\mathrm{d}W_t$, et $\langle M \rangle_T$ est le processus croissant de M_T qui vérifie

$$\langle M \rangle_T = T\begin{pmatrix} \Sigma_T & 1 \\ 1 & S_T \end{pmatrix}$$

en adoptant la notation $S_T = T^{-1}\int_0^T X_t\,\mathrm{d}t$ et $\Sigma_T = T^{-1}\int_0^T X_t^{-1}\,\mathrm{d}t$. Puis, on utilise la décomposition suivante :

$$\sqrt{T\lambda_T^{-1}}\left(\widehat{\theta}_T - \theta\right) = 2T\left(I_2 \otimes \langle M \rangle_T^{-1}\right)\left(\lambda_T T\right)^{-1/2} \mathcal{M}_T,$$

où $\mathcal{M}_T = (M_T, N_T)^{\intercal}$. On montre l'existence d'un PGD de vitesse λ_T pour la martingale quadridimensionnelle $\left(\lambda_T T\right)^{-1/2} \mathcal{M}_T$, qui repose sur la convergence exponentielle de son processus croissant, et donc sur l'existence d'un PGD pour le couple (S_T, Σ_T).

1.4.3 Estimation des paramètres quand le processus CIR s'annule

Pour terminer, on s'intéresse, toujours dans le cas sous-critique, au problème de l'estimation des paramètres pour $a < 2$. En effet, dans cette situation le temps d'arrêt τ_0 donné par (1.4), jusqu'auquel l'EMV est défini, est fini presque sûrement. Dans le cas du processus CIR, Overbeck [65] propose un couple d'estimateurs fortement consistant (voir (1.9)) mais qui est difficilement utilisable en pratique, à cause de sa construction à l'aide d'un temps d'arrêt. Pourtant, il existe de nombreuses situations, en mathématiques financières par exemple, où le paramètre a est strictement plus petit que 2 et le processus (X_t) s'annule fréquemment. On trouvera des exemples de calibrations aboutissant à cette situation dans [30], table 1, et dans [33], table 1. On propose au Chapitre 4 un quadruplet d'estimateurs, pour les paramètres du processus de Heston, fortement consistant et asymptotiquement normal. Ce nouveau quadruplet d'estimateurs est construit à l'aide d'une pondération permettant d'éviter l'explosion du dénominateur lorsque le processus s'annule, comme cela a été fait auparavant dans le cas des processus de branchement à temps discrets par Wei et Winnicki [70].

Soit $C_T = X_T + c$ où c est une constante positive. Notre nouveau couple d'estimateurs des moindres carrés pondérés [1] est donné comme suit

$$\breve{\theta}_T = \begin{pmatrix} \Gamma_{c,T}^{-1} & 0 \\ 0 & \Gamma_{c,T}^{-1} \end{pmatrix} \begin{pmatrix} U_{c,T} \\ V_{c,T} \end{pmatrix} \tag{1.17}$$

où $U_{c,T} = \left(\int_0^T \frac{1}{C_t} \, \mathrm{d}X_t, \int_0^T \frac{X_t}{C_t} \, \mathrm{d}X_t \right)^\intercal$, $V_{c,T} = \left(\int_0^T \frac{1}{C_t} \, \mathrm{d}Y_t, \int_0^T \frac{X_t}{C_t} \, \mathrm{d}Y_t \right)^\intercal$ et

$$\Gamma_{c,T} = \begin{pmatrix} \int_0^T \frac{1}{C_t} \, \mathrm{d}t & \int_0^T \frac{X_t}{C_t} \, \mathrm{d}t \\ \int_0^T \frac{X_t}{C_t} \, \mathrm{d}t & \int_0^T \frac{X_t^2}{C_t} \, \mathrm{d}t \end{pmatrix}.$$

On ne se restreint pas au cas particulier où $c = 1$ afin d'améliorer éventuellement la variance des estimateurs. On remarque que dans le cas particulier où $c = 0$, ce nouveau quadruplet d'estimateurs coincide avec l'EMV. On montre les deux résultats suivants :

Théorème 1.4.6. *Soient $a > 0$ et $b < 0$. Alors, l'estimateur $\breve{\theta}_T$ est fortement consistant : pour T qui tend vers l'infini,*

$$\breve{\theta}_T \xrightarrow{p.s.} \theta. \tag{1.18}$$

On rappelle que, lorsque T tend vers l'infini, X_T converge en loi vers X qui suit une loi Gamma $\Gamma(a/2, -b/2)$. On note C la limite en loi de $X_T + c$, quand T tend vers l'infini.

Théorème 1.4.7. *On suppose que $a > 0$ et $b < 0$. Alors, pour T qui tend vers l'infini, l'estimateur $\breve{\theta}_T$ satisfait le Théorème de la limite centrale suivant*

$$\sqrt{T} \left(\breve{\theta}_T - \theta \right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, 4\Lambda\right), \tag{1.19}$$

où la variance asymptotique Λ est donnée comme une matrice par blocs

$$\Lambda = \begin{pmatrix} ALA & \rho ALA \\ \rho ALA & ALA \end{pmatrix}, \tag{1.20}$$

1. Les notations sont différentes de celles de l'article original [26] et donc du Chapitre 4.

avec les matrices A et L respectivement données par

$$A = \left(\mathbb{E}\left[C\right]\mathbb{E}\left[1/C\right] - 1\right)^{-1}\begin{pmatrix} \mathbb{E}\left[X^2/C\right] & -\mathbb{E}\left[X/C\right] \\ -\mathbb{E}\left[X/C\right] & \mathbb{E}\left[1/C\right] \end{pmatrix}$$

et

$$L = \begin{pmatrix} \mathbb{E}\left[X/C^2\right] & \mathbb{E}\left[X^2/C^2\right] \\ \mathbb{E}\left[X^2/C^2\right] & \mathbb{E}\left[X^3/C^2\right] \end{pmatrix}.$$

De ce quadruplet d'estimateurs pour le processus de Heston, on déduit un couple d'estimateurs $(\breve{a}_T, \breve{b}_T)$ pour les paramètres du processus CIR, en ne considérant que les deux premiéres coordonnées :

$$\begin{pmatrix} \breve{a}_T \\ \breve{b}_T \end{pmatrix} = \Gamma_{c,T}^{-1} U_{c,T}$$

où $\Gamma_{c,T}$ et $U_{c,T}$ sont définis après (1.17). On montre que ce couple d'estimateurs est fortement consistant et asymptotiquement normal :

Corollaire 1.4.7. *Soient $a > 0$ et $b < 0$. Alors, l'estimateur $(\breve{a}_T, \breve{b}_T)$ du couple de paramètres (a, b) est fortement consistant quand T tend vers l'infini,*

$$\begin{pmatrix} \breve{a}_T \\ \breve{b}_T \end{pmatrix} \xrightarrow{p.s.} \begin{pmatrix} a \\ b \end{pmatrix}.$$

De plus, il satisfait le théorème de la limite centrale

$$\sqrt{T}\begin{pmatrix} \breve{a}_T - a \\ \breve{b}_T - b \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, 4ALA\right).$$

Chapitre 2

Grandes déviations pour l'EMV des paramètres du processus CIR

Résumé. On établit un principe de grandes déviations pour l'estimateur du maximum de vraisemblance du couple des paramètres du processus CIR. On se restreint au cas le plus simple où le paramètre de dimension est tel que $a > 2$ et la dérive vérifie $b < 0$. La nouveauté réside dans le fait que l'on estime les deux paramètres simultanément.

Abstract. We establish large deviation principles for the couple of the maximum likelihood estimators of dimensional and drift coefficients in the CIR process. We focus our attention to the most tractable situation where the dimensional parameter $a > 2$ and the drift parameter $b < 0$. In contrast to the previous literature, we state large deviation principles when both dimensional and drift coefficients are estimated simultaneously.

Sommaire

2.1 Introduction

The generalized squared radial Ornstein-Uhlenbeck process, also known as the Cox-Ingersoll-Ross process, is the strong solution of the stochastic differential equation

$$\mathrm{d}X_t = (a + bX_t)\mathrm{d}t + 2\sqrt{X_t}\,\mathrm{d}B_t \tag{2.1}$$

where the initial state $X_0 = x \geqslant 0$, the dimensional parameter $a > 0$, the drift coefficient $b \in \mathbb{R}$ and (B_t) is a standard Brownian motion. The behaviour of the process has been widely investigated and depends on the values of both coefficients a and b. We shall restrict ourself to the most tractable situation where $a > 2$ and $b < 0$. In this case, the process is ergodic and never reaches zero.

We estimate parameters a and b at the same time using a trajectory of the process over the time interval $[0, T]$. The maximum likelihood estimators (MLE) of a and b are given by:

$$\widehat{a}_T = \frac{\int_0^T X_t\,\mathrm{d}t \int_0^T \frac{1}{X_t}\,\mathrm{d}X_t - T(X_T - x)}{\int_0^T X_t\,\mathrm{d}t \int_0^T \frac{1}{X_t}\,\mathrm{d}t - T^2} \quad \text{and} \quad \widehat{b}_T = \frac{(X_T - x)\int_0^T \frac{1}{X_t}\,\mathrm{d}t - T\int_0^T \frac{1}{X_t}\,\mathrm{d}X_t}{\int_0^T X_t\,\mathrm{d}t \int_0^T \frac{1}{X_t}\,\mathrm{d}t - T^2}. \tag{2.2}$$

Overbeck [65] has shown that $\widehat{a}_T$ and $\widehat{b}_T$ both converge almost surely to a and b. In addition, he has proven that

$$\sqrt{T}\begin{pmatrix} \widehat{a}_T - a \\ \widehat{b}_T - b \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 4C^{-1}) \quad \text{where} \quad C = \begin{pmatrix} \frac{-b}{a-2} & 1 \\ 1 & -\frac{a}{b} \end{pmatrix}.$$

Moderate deviation results for $\widehat{a}_T$ and $\widehat{b}_T$ are achieved in [38]. In addition, Zani [73] established large deviation principles (LDP) for the MLE of a assuming b known and, conversely, for the MLE of b assuming a known. Our goal is to extend her results to the case where both parameters are estimated simultaneously. Our method is also different and we explain how we have simplified her approach at the beginning of Section 2.2 and Section 4, using a new strategy introduced by Bercu and Richou in [17] for the study of the Ornstein-Uhlenbeck process with shift.

The paper is organised as follows. Section 2 is devoted to an LDP for the couple $(\widehat{a}_T, \widehat{b}_T)$, which is obtained via LDPs for two other couples of estimators constructed on the MLE. Before we prove those results, which is respectively the aim of Sections 5 and 6, we investigate in Section 3 LDPs for some useful functionals of the process and compute in Section 4 the normalized cumulant generating function of a given quadruplet, which is a keystone for every LDP we establish in this paper. Technical proofs are postponed to Appendix A to E.

2.2 Main results

We start by rewriting the estimators $\widehat{a}_T$ and $\widehat{b}_T$ in such a way that they are much easier to handle. We need to suppose the starting point $x > 0$ to apply the well-known

Itô's formula to $\log X_T$. We obtain that

$$\int_0^T \frac{1}{X_t}\,\mathrm{d}X_t = \log X_T - \log x + 2\int_0^T \frac{1}{X_t}\,\mathrm{d}t \tag{2.3}$$

which leads to

$$\widehat{a}_T = \frac{S_T\,(2\,\Sigma_T + L_T) - \frac{X_T - x}{T}}{V_T} \quad \text{and} \quad \widehat{b}_T = \frac{(\frac{X_T-x}{T} - 2)\,\Sigma_T - L_T}{V_T}$$

where the denominator $V_T = S_T\,\Sigma_T - 1$ with

$$S_T = \frac{1}{T}\int_0^T X_t\,\mathrm{d}t \quad \text{and} \quad \Sigma_T = \frac{1}{T}\int_0^T \frac{1}{X_t}\,\mathrm{d}t,$$

and

$$L_T = \frac{\log X_T - \log x}{T}.$$

For the remaining of the paper, we remove the parts involving the starting point x. It does not change the large deviation results because both couple of estimators, with and without x and $\log x$, are exponentially equivalent so that they share the same LDP. Thus, we now suppose that the estimators are respectively given by

$$\widehat{a}_T = \frac{S_T\,(2\,\Sigma_T + L_T) - \frac{X_T}{T}}{V_T} \quad \text{and} \quad \widehat{b}_T = \frac{(\frac{X_T}{T} - 2)\,\Sigma_T - L_T}{V_T} \tag{2.4}$$

where $L_T = \log X_T/T$.

As the rate function of the LDP for the MLE will turn out to be not directly computable, we first consider two couples of simplified estimators constructed from $(\widehat{a}_T, \widehat{b}_T)$ using the fact that L_T and X_T/T both tend to zero almost surely for T going to infinity. For those estimators LDPs are more straightforward and will finally be involved in the computation of the rate function of the MLE $(\widehat{a}_T, \widehat{b}_T)$. All the LDPs established in this paper are satisfied with speed T.

2.2.1 Simplified estimators

A first strategy to propose simplified estimators of a and b is to remove the logarithmic term L_T in the expression of $\widehat{a}_T$ and $\widehat{b}_T$ given by (2.4). This way, we obtain a new couple $(\widetilde{a}_T, \widetilde{b}_T)$ defined by

$$\widetilde{a}_T = \frac{2\,S_T\,\Sigma_T - \frac{X_T}{T}}{V_T} \quad \text{and} \quad \widetilde{b}_T = \frac{(\frac{X_T}{T} - 2)\,\Sigma_T}{V_T}. \tag{2.5}$$

It is clear that $\widetilde{a}_T$ and $\widetilde{b}_T$ converge almost surely to a and b. Moreover, we also have the same Central Limit Theorem (CLT)

$$\sqrt{T}\begin{pmatrix}\widetilde{a}_T - a \\ \widetilde{b}_T - b\end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 4C^{-1}).$$

The proof of this result can be found in appendix A. We also state an LDP for the couple $(\widetilde{a}_T, \widetilde{b}_T)$ assuming both parameters a and b unknown.

Theorem 2.2.1. *The couple $(\widetilde{a}_T, \widetilde{b}_T)$ satisfies an LDP with good rate function*

$$J_{a,b}(\alpha,\beta) = \begin{cases} \dfrac{(a-2)^2\beta}{8(2-\alpha)}\left(1+\dfrac{(2-\alpha)b}{\beta(a-2)}\right)^2 + 2\beta - b & \text{if } \alpha > 2, \frac{b}{3} \leqslant \beta < 0 \\ & \text{or if } \alpha < 2, \beta > 0, \\ \dfrac{(a-2)^2\beta}{8(2-\alpha)}\left(1+\dfrac{(2-\alpha)b}{\beta(a-2)}\right)^2 - \dfrac{\beta}{4}\left(1-\dfrac{b}{\beta}\right)^2 & \text{if } \alpha > 2, \beta \leqslant \frac{b}{3}, \\ -b & \text{if } (\alpha,\beta) = (2,0), \\ +\infty & \text{otherwise.} \end{cases}$$

Proof. The proofs of this theorem and the two following corollaries are postponed to Section 5. □

We give the shape of this rate function in Figure 2.1 below in the particular case $(a, b) = (4, -1)$ and over $[3, 5] \times [-4, -0.5]$. One can notice that the rate function reaches zero at point $(4, -1)$.

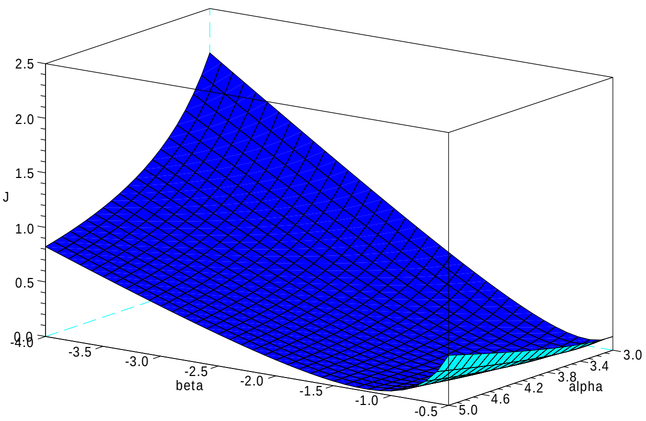


Figure 2.1 – Rate function for the couple of simplified estimators $(\widetilde{a}_T, \widetilde{b}_T)$

LDPs for each estimator $\widetilde{a}_T$ and $\widetilde{b}_T$ immediately follow from the contraction principle (see Theorem 4.2.1 of [25] and the following Remarks) which is recalled here for completeness.

Lemma 2.2.1 (Contraction Principle)**.** *Let $(Z_T)_T$ be a sequence of random variables of $\mathbb{R}^d$ satisfying an LDP with good rate function I and $g : \mathbb{R}^d \to \mathbb{R}^n$ be a continuous function over $\mathcal{D}_I = \left\{x \in \mathbb{R}^d | I(x) < +\infty\right\}$. The sequence $(g(Z_T))_T$ satisfies an LDP with good rate function J defined for all $y \in \mathbb{R}^n$ by*

$$J(y) = \inf_{\{x \in \mathcal{D}_I | g(x) = y\}} I(x).$$

Corollary 2.2.1. *The sequence* $(\widetilde{a}_T)$ *satisfies an LDP with good rate function*

$$J_a(\alpha) = \begin{cases} \dfrac{b}{4}\left(a - 6 - \sqrt{(a-2)^2 + 16(2-\alpha)}\right) & \text{if } \alpha \leqslant \ell_a, \\ \dfrac{b}{4}\left(a - \sqrt{\alpha\left(\dfrac{(a-2)^2}{\alpha - 2} + 2\right)}\right) & \text{if } \alpha \geqslant \ell_a, \end{cases}$$

with $\ell_a = \frac{10}{9} + \frac{1}{9}\sqrt{64 + 9(a-2)^2}$.

Corollary 2.2.2. *The sequence* $(\widetilde{b}_T)$ *satisfies an LDP with good rate function*

$$J_b(\beta) = \begin{cases} -\dfrac{\beta}{4}\left(1 - \dfrac{b}{\beta}\right)^2 & \text{if } \beta \leqslant \frac{b}{3}, \\ 2\beta - b & \text{if } \beta \geqslant \frac{b}{3}. \end{cases}$$

Remark 2.2.1. *This rate function is the same than the one obtained by Zani [73] for the MLE of* b *assuming* a *known.*

Figure 2.3 displays in blue the rate functions J_a and J_b in the particular case where $(a, b) = (4, -1)$.

A second strategy to propose simplified estimators of a and b is to remove the term X_T/T in the expression of $\widehat{a}_T$ and $\widehat{b}_T$ given by (2.4). Then, we obtain a new couple $(\breve{a}_T, \breve{b}_T)$ defined by

$$\breve{a}_T = \frac{S_T\,(2\,\Sigma_T + L_T)}{V_T} \qquad \text{and} \qquad \breve{b}_T = \frac{-2\,\Sigma_T - L_T}{V_T}. \tag{2.6}$$

As previously, $\breve{a}_T$ and $\breve{b}_T$ converge almost surely to a and b, and

$$\sqrt{T}\begin{pmatrix} \breve{a}_T - a \\ \breve{b}_T - b \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 4C^{-1}).$$

The proof of this result is given in appendix A. Again, we establish an LDP for the couple $(\breve{a}_T, \breve{b}_T)$, and deduce as corollaries LDPs for both estimators, assuming a and b unknown.

Theorem 2.2.2. *The couple* $(\breve{a}_T, \breve{b}_T)$ *satisfies an LDP with good rate function*

$$K_{a,b}(\alpha,\beta) = \begin{cases} \dfrac{a}{4}(b-\beta) - \dfrac{\alpha}{8\beta}\left(b^2 - \beta^2\right) - \dfrac{\beta}{\alpha}\left(\sqrt{2} + \sqrt{C_\alpha}\right)^2 & \text{if } \beta < 0, 0 < \alpha \leqslant \alpha_a \\ & \text{or if } \beta > 0, \alpha < 0, \\ \dfrac{a}{4}(b-\beta) - \dfrac{\alpha}{8\beta}\left(b^2 - \beta^2\right) - \dfrac{\beta\,(a-\alpha)^2}{8\,(\alpha-2)} & \text{if } \beta < 0, \alpha \geqslant \alpha_a, \\ -\dfrac{b}{4}\left(4 - a + \sqrt{a^2 + 16}\right) & \text{if } (\alpha,\beta) = (0,0), \\ +\infty & \text{otherwise,} \end{cases}$$

where $C_\alpha = \frac{1}{8}\,(a-\alpha)^2 + 2 - \alpha$ *and* $\alpha_a = -\frac{2}{3}\left(\frac{a}{2} - 2 - \sqrt{a^2 - 2a + 4}\right)$.

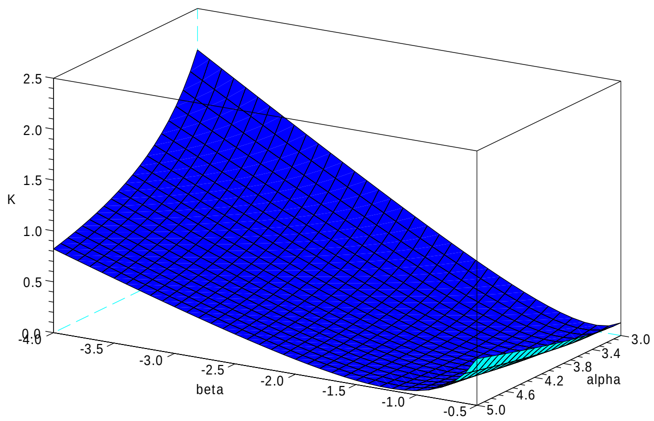


Figure 2.2 – Rate function for the couple of simplified estimators $(\breve{a}_T, \breve{b}_T)$.

One can observe that the rate functions $J_{a,b}$ and $K_{a,b}$ are equal over some domain of $\mathbb{R}^2$. It is possible to see it on Figure 2.2 below which is quite similar to the previous one and displays the rate function $K_{a,b}$ in the particular case where $(a, b) = (4, -1)$.

Corollary 2.2.3. *The sequence $(\breve{a}_T)$ satisfies an LDP with good rate function*

$$K_a(\alpha) = \begin{cases} -\dfrac{b}{4}\left(4 - a + \sqrt{a^2+16}\right) & \text{if } \alpha = 0, \\ K_{a,b}(\alpha, \beta_b) & \text{if } \alpha < \alpha_a, \alpha \neq 0, \\ \dfrac{b}{4}\left(a - \sqrt{\alpha\left(\dfrac{(a-2)^2}{\alpha - 2} + 2\right)}\right) & \text{if } \alpha \geqslant \alpha_a \end{cases}$$

with $\beta_b = b\alpha \left(16\sqrt{2C_\alpha} + a^2 - 8\alpha + 32\right)^{-1/2}$ and α_a, C_α are defined in Theorem 2.2.2.

Corollary 2.2.4. *The sequence $(\breve{b}_T)$ satisfies an LDP with good rate function*

$$K_b(\beta) = \inf\Big\{K_{a,b}(\alpha, \beta) \ / \ \alpha \in \mathbb{R}\Big\}.$$

In particular, $K_b(0) = K_{a,b}(0,0) = -\frac{b}{4}\left(4 - a + \sqrt{a^2+16}\right)$.

Figure 2.3 displays in green the rate functions K_a and K_b in the particular case where $(a, b) = (4, -1)$.

2.2.2 Large deviation results for the MLE

The next theorem gives a large deviation principle for the MLE $(\widehat{a}_T, \widehat{b}_T)$ of the couple (a, b). In contrast with the previous literature, we consider both parameters unknown and estimate them simultaneously. We also simplified the approach of the previous literature as our proofs only rely on the Gärtner-Ellis theorem and do not need, for example, accurate time-depending changes of probability.

Theorem 2.2.3. *The couple* $(\widehat{a}_T, \widehat{b}_T)$ *satisfies an LDP with good rate function* $I_{a,b}$ *given over* $\mathbb{R}^2$ *by*

$$I_{a,b}(\alpha,\beta) = \min\left(J_{a,b}(\alpha,\beta), K_{a,b}(\alpha,\beta)\right).$$

Proof. Section 6 is devoted to the proof of this result. □

Making use of the contraction principle (see Lemma 2.2.1), once again, we obtain straightforwardly the two following corollaries.

Corollary 2.2.5. *The sequence* $(\widehat{a}_T)$ *satisfies an LDP with good rate function*

$$I_a(\alpha) = \min\left(J_a(\alpha), K_a(\alpha)\right).$$

Corollary 2.2.6. *The sequence* $(\widehat{b}_T)$ *satisfies an LDP with good rate function*

$$I_b(\beta) = \min\left(J_b(\beta), K_b(\beta)\right).$$

Figure 2.3 displays in red the rate functions I_a and I_b in the particular case where $(a, b) = (4, -1)$.

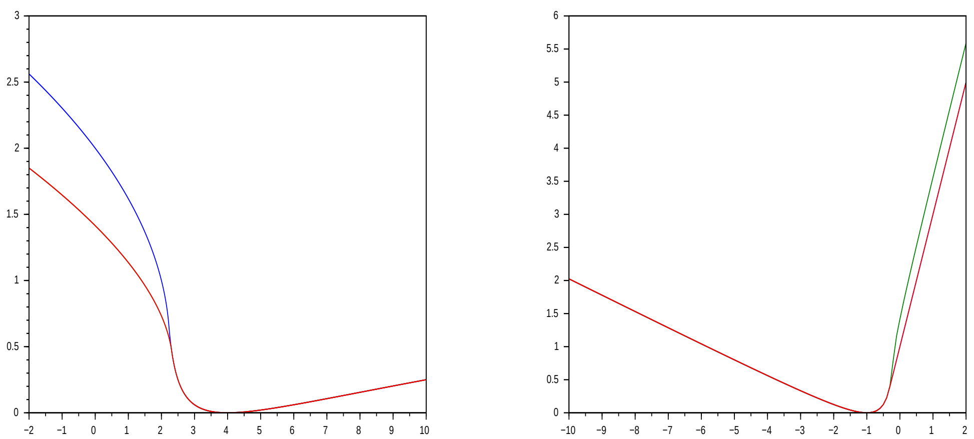


Figure 2.3 – Rate functions for dimensional and drift parameters.

Remark 2.2.2. *Both couples of simplified estimators perform better than the MLE in terms of large deviations, as their rate functions are always greater.*

2.3 Some results about the process

The aim of this section is to establish LDPs with speed T for S_T, Σ_T and V_T, which will be involved in the proof of the main theorem.

Lemma 2.3.1. *The couple* (S_T, Σ_T) *satisfies an LDP with good rate function*

$$I(x,y) = \begin{cases} \dfrac{y}{2(xy-1)} + \dfrac{b^2}{8}x + \dfrac{(a-2)^2}{8}y + \dfrac{ab}{4} & \text{if } x>0, y>0 \text{ and } xy-1>0 \\ +\infty & \text{otherwise.} \end{cases}$$

Proof. See appendix B. □

The following result can be proven either directly with the same method or using the previous lemma together with the contraction principle recalled in Lemma 2.2.1.

Theorem 2.3.1. *The sequence* (S_T) *satisfies an LDP with good rate function*

$$I(x) = \begin{cases} \frac{(a+bx)^2}{8x} & \text{if } x > 0 \\ +\infty & \text{if } x \leqslant 0. \end{cases}$$

In addition, the sequence (Σ_T) *satisfies an LDP with good rate function*

$$J(x) = \begin{cases} \frac{\left((a-2)x+b\right)^2}{8x} & \text{if } x > 0 \\ +\infty & \text{if } x \leqslant 0. \end{cases}$$

It is now easy to establish an LDP for V_T. We recall that $V_T = h\left(S_T, \Sigma_T\right)$ where h is the function defined on $\mathbb{R}^2$ by $h(x,y) = xy - 1$.

Theorem 2.3.2. *The sequence* (V_T) *verifies an LDP with good rate function*

$$K(x) = \begin{cases} -\dfrac{b}{4}\sqrt{(x+1)\left((a-2)^2 + \dfrac{4}{x}\right)} + \dfrac{ab}{4} & \text{if } x > 0 \\ +\infty & \text{if } x \leqslant 0. \end{cases}$$

Proof. It follows immediately from Lemma 2.3.1 together with the contraction principle (see Lemma 2.2.1). It only remains to explicitly evaluate the rate function K given, for all real z, by

$$K(z) = \inf_{\{(x,y) | z = xy-1\}} I(x,y)$$

where I is defined in Lemma 2.3.1. □

2.4 Cumulant generating function for the quadruplet

To establish LDPs for the estimators $\left(\widehat{a}_T, \widehat{b}_T\right)$, we need to compute the normalized cumulant generating function of the quadruplet $(X_T/T, S_T, \Sigma_T, L_T)$. However, this does not lead to a steep function (see [25] for the definition), which is a necessary condition to apply Gärtner-Ellis theorem . In contrast with the previous literature, we will not search another method to obtain large deviation results. Following the strategy of [17], the idea to overcome this difficulty is to consider instead the quadruplet $\mathcal{Q}_T = \left(\sqrt{X_T/T}, S_T, \Sigma_T, \mathcal{L}_T\right)$, where

$$\mathcal{L}_T = -\sqrt{\frac{-\log X_T}{T}}\mathbf{1}_{X_T<1} + \frac{\log X_T}{T}\mathbf{1}_{X_T \geqslant 1}. \tag{2.7}$$

Proposition 2.4.1. *Let $\Lambda_T(\lambda,\mu,\nu,\gamma)$ be the normalized cumulant generating function of the quadruplet $\mathcal{Q}_T$ given over $\mathbb{R}^4$ by*

$$\Lambda_T(\lambda,\mu,\nu,\gamma)=\frac{1}{T}\log\left(\mathbb{E}\left[\exp\left(\lambda\sqrt{T}\sqrt{X_T}+\gamma T\mathcal{L}_T+\mu\int_0^T X_t\,\mathrm{d}t+\nu\int_0^T\frac{1}{X_t}\,\mathrm{d}t\right)\right]\right).$$

Denote by Λ its pointwise limit as T tends to $+\infty$. For all $\lambda,\gamma\in\mathbb{R}$, $\mu<\frac{b^2}{8}$ and $\nu<\frac{(a-2)^2}{8}$,

$$\Lambda(\lambda,\mu,\nu,\gamma)=\begin{cases}-\dfrac{d}{2}(1+f)-\dfrac{ab}{4}+\dfrac{\lambda^2}{d-b} & \text{if } \lambda>0 \text{ and } \gamma\geqslant 0\\ & \quad\text{or if } \gamma<0,\lambda>0 \text{ and } \frac{\gamma^2}{\lambda^2}<\frac{2f+a+2}{d-b},\\ -\dfrac{d}{2}(1+f)-\dfrac{ab}{4}+\dfrac{\gamma^2}{2f+a+2} & \text{if } \lambda\leqslant 0 \text{ and } \gamma<0\\ & \quad\text{or if } \gamma<0,\lambda>0 \text{ and } \frac{\gamma^2}{\lambda^2}\geqslant\frac{2f+a+2}{d-b},\\ -\dfrac{d}{2}(1+f)-\dfrac{ab}{4} & \text{if } \lambda\leqslant 0 \text{ and } \gamma\geqslant 0,\end{cases}$$

where $d=\sqrt{b^2-8\mu}$ and $f=\dfrac{1}{2}\sqrt{(a-2)^2-8\nu}$.

Lemma 2.4.1. *The function Λ is steep.*

Proof. Λ is differentiable over its domain $\mathcal{D}_\Lambda=\mathbb{R}\times]-\infty,\frac{b^2}{8}[\times]-\infty,\frac{(a-2)^2}{8}[\times\mathbb{R}$ and its gradient is given by

$$\nabla\Lambda=\begin{pmatrix}\frac{2\lambda}{d-b}\mathbf{1}_{\Delta_1}\\ \frac{2(1+f)}{d}+\frac{4\lambda^2}{d(d-b)^2}\mathbf{1}_{\Delta_1}\\ \frac{d}{2f}+\frac{2\gamma^2}{f(2f+a+2)^2}\mathbf{1}_{\Delta_2}\\ \frac{2\gamma}{2f+a+2}\mathbf{1}_{\Delta_2}\end{pmatrix}, \tag{2.8}$$

where $\Delta_1=\Big\{(\lambda,\mu,\nu,\gamma)\in\mathcal{D}_\Lambda\ /\ \lambda>0 \text{ and } \gamma\geqslant 0 \text{ or } \gamma<0,\lambda>0 \text{ and } \frac{\gamma^2}{\lambda^2}<\frac{2f+a+2}{d-b}\Big\}$ and $\Delta_2=\Big\{(\lambda,\mu,\nu,\gamma)\in\mathcal{D}_\Lambda\ /\ \lambda\leqslant 0 \text{ and } \gamma<0 \text{ or } \gamma<0,\lambda>0 \text{ and } \frac{\gamma^2}{\lambda^2}\geqslant\frac{2f+a+2}{d-b}\Big\}$. We easily obtain that the norm of (2.8) goes to infinity for any sequence in the interior of $\mathcal{D}_\Lambda$ converging to a boundary point. □

Proof of Proposition 2.4.1. We want to find the limit of $\Lambda_T(\lambda,\mu,\nu,\gamma)$ as $T\to+\infty$. It follows from Girsanov's theorem (see [57]) that for all $\lambda,\gamma\in\mathbb{R}$, $\mu<\frac{b^2}{8}$ and $\nu<\frac{(a-2)^2}{8}$,

$$\begin{aligned}\Lambda_T(\lambda,\mu,\nu,\gamma)\ &=\ \frac{c-a}{4T}\log x+\frac{d-b}{4T}x+(dc-ab)/4\\ &\quad+\frac{1}{T}\log\mathbb{E}_{c,d}\left[X_T^{\frac{a-c}{4}}\exp\left(\frac{b-d}{4}X_T+\lambda\sqrt{T}\sqrt{X_T}+\gamma T\mathcal{L}_T\right)\right]\end{aligned} \tag{2.9}$$

where we changed parameters a and b to new parameters c and d, respectively given by $c=2+\sqrt{(a-2)^2-8\nu}$ and $d=\sqrt{b^2-8\mu}$. Besides, we know that X_T follows a noncentral chi-squared distribution and its transition density $p_{c,d}$ is given for any $T \in \mathbb{R}_*^+$, $x > 0$ and $y \in \mathbb{R}_*^+$ by

$$p_{c,d}(T,x,y) = \frac{\beta_T}{2\sqrt{x}} \left(\frac{y}{x}\right)^{\frac{f}{2}} \exp\left(-\frac{d}{4}\left(cT + (x+y)\coth\left(\frac{dT}{2}\right) + x - y\right)\right) I_f\left(\beta_T \sqrt{y}\right)$$

with $f = (c-2)/2$ and $\beta_T = d\sqrt{x}/(2\sinh(dT/2))$, I_f being the modified Bessel function of the first kind. Thus, equation (2.9) becomes

$$\begin{aligned}\Lambda_T(\lambda,\mu,\nu,\gamma) &= \frac{c-a}{4T}\log x + \frac{d-b}{4T}x + (dc-ab)/4 \\ &\quad + \frac{1}{T}\log \int_0^\infty y^{\frac{a-c}{4}} e^{\frac{b-d}{4}y + \lambda\sqrt{Ty} + \gamma T l(T,y)} p_{c,d}(T,x,y)\,\mathrm{d}y\end{aligned}$$

where $l(T,y) = -\sqrt{\frac{-\log y}{T}}\,\mathbf{1}_{y<1} + \frac{\log y}{T}\,\mathbf{1}_{y\geqslant 1}$. Replacing $p_{c,d}$ by its expression and taking out of the integral all the terms that do not depend on y, we obtain that

$$\Lambda_T(\lambda,\mu,\nu,\gamma) = \frac{1}{T}\left(\log \mathcal{J}_T + \log\left(\beta_T\sqrt{x}/2\right) - \frac{1}{4}\left(abT + dx\coth\left(dT/2\right) + bx + a\log x\right)\right)$$

where

$$\mathcal{J}_T = \int_0^\infty e^{\lambda\sqrt{Ty} + \gamma T l(T,y) - \frac{y}{4}(d\coth(dT/2)-b)} y^{\frac{a-2}{4}} I_f\left(\beta_T\sqrt{y}\right)\,\mathrm{d}y. \tag{2.10}$$

However, as soon as T tends to infinity, $\coth(dT/2)$ goes to 1, which implies that

$$\lim_{T\to+\infty} -\frac{1}{4T}\left(abT + d\coth\left(\frac{dT}{2}\right) + b\right) = -\frac{ab}{4}. \tag{2.11}$$

On the other hand,

$$\frac{1}{T}\log\left(\sinh\left(\frac{dT}{2}\right)\right) = \frac{1}{T}\frac{dT}{2} + \frac{1}{T}\log\left(\frac{(1-e^{-dT})}{2}\right)$$

which clearly leads to

$$\lim_{T\to+\infty} \frac{1}{T}\log\left(\frac{\beta_T}{2}\right) = -\frac{d}{2}. \tag{2.12}$$

We have to establish the asymptotic behaviour of $\frac{1}{T}\log \mathcal{J}_T$. We split $\mathcal{J}_T$ into two terms: $\mathcal{J}_T = H_T + K_T$ where

$$H_T = \int_0^1 e^{\lambda\sqrt{Ty} - \gamma\sqrt{-T\log y} - \alpha_T y} y^{\frac{a-2}{4}} I_f\left(\beta_T\sqrt{y}\right)\,\mathrm{d}y, \tag{2.13}$$

$$K_T = \int_1^\infty e^{\lambda\sqrt{Ty} - \alpha_T y} y^{\gamma + \frac{a-2}{4}} I_f\left(\beta_T\sqrt{y}\right)\,\mathrm{d}y \tag{2.14}$$

with $\alpha_T = (d\coth(dT/2) - b)/4$.

We need the four following lemmas, whose proofs are postponed to Appendix C.

Lemma 2.4.2. *For all $\gamma < 0$ and $\lambda \in \mathbb{R}$, one can find the following bounds for H_T as T goes to infinity.*

$$H_T \leqslant \frac{2^{2-f}}{\Gamma(f+1)} \sqrt{\pi} g^{-3/2} |\gamma| \sqrt{T}\, e^{|\lambda| + \alpha_T/T + \beta_T/\sqrt{T}} \beta_T^f \exp\left(\frac{\gamma^2 T}{4g}\right)$$

and

$$H_T \geqslant \frac{2^{-1-f}}{\Gamma(f+1)} \sqrt{\pi} g^{-3/2} |\gamma| \sqrt{T}\, e^{-|\lambda| - \alpha_T/T} \beta_T^f \exp\left(\frac{\gamma^2 T}{4g}\right),$$

where $g = \frac{2f+a+2}{4}$.

Lemma 2.4.3. *For all $\gamma \geqslant 0$ and $\lambda \in \mathbb{R}$, bounds for H_T are given by*

$$H_T \leqslant \frac{(\beta_T)^f}{\Gamma(f+1)2^f} \exp\left(|\lambda| \sqrt{T} + \beta_T\right)$$

and

$$H_T \geqslant \frac{(\beta_T)^f \ \varepsilon_T}{2^f\, \Gamma\left(1+f\right)} \exp\left(-\gamma \sqrt{T} \sqrt{\log T} - g \log T\right),$$

where $\varepsilon_T = e^{-\alpha_T} \left(\varepsilon + \frac{e^{-|\lambda|}}{g}\left(1 - \frac{\gamma\sqrt{T}}{2g\sqrt{\log T} + \gamma\sqrt{T}}\right)\right)$ and $\varepsilon = \exp\left(\lambda \mathbf{1}_{\lambda \geqslant 0} + \lambda \sqrt{T} \mathbf{1}_{\lambda < 0}\right)$.

Lemma 2.4.4. *For all $\lambda \leqslant 0$ and $\gamma \in \mathbb{R}$ and T tending to infinity, $K_T = O\left((\beta_T)^f\right)$. Moreover if $\gamma \geqslant 0$, we have the following lower bound*

$$K_T \geqslant \frac{2^{1-f} \left(\beta_T\right)^f}{\Gamma(1+f)} \frac{e^{-\alpha_T}}{2\alpha_T} \frac{1}{1 - \frac{\lambda\sqrt{T}}{2\alpha_T}} \exp\left(\lambda\sqrt{T}\right).$$

Lemma 2.4.5. *For all $\lambda > 0$ and $\gamma \in \mathbb{R}$, K_T is bounded as follows for T going to infinity,*

$$K_T \leqslant \frac{2^{2-f} \sqrt{2\pi}\, (\beta_T)^f}{\Gamma(1+f)} \left(\beta_T + \lambda\sqrt{T}\right)^{2\gamma + 2g - 1} \exp\left(\frac{(\lambda\sqrt{T} + \beta_T)^2}{4\alpha_T}\right)$$

and

$$K_T \geqslant 2^{1-f} \sqrt{\frac{\pi}{d-b}} \frac{(\beta_T)^f}{\Gamma(1+f)}\, m_{\gamma,\lambda,T} \exp\left(\frac{\lambda^2 T}{4\alpha_T}\right).$$

where $m_{\gamma,\lambda,T} = \min\left\{\left(\frac{\lambda\sqrt{T}}{\alpha_T}\right)^{2(\gamma+g-1)}; \frac{\lambda\sqrt{T}}{\alpha_T}\right\}$.

It is clear with those lemmas that the asymptotic behaviour of $\mathcal{J}_T$ depends on the sign of λ and γ:

• For all $\lambda > 0$ and $\gamma \geqslant 0$: we directly deduce from Lemmas 2.4.5 and 2.4.3 that, for T large enough,

$$H_T + K_T \leqslant \frac{2^{2-f} \sqrt{2\pi}\, (\beta_T)^f}{\Gamma(1+f)} \left(\beta_T + \lambda\sqrt{T}\right)^{2\gamma+2g-1} \left(e^{\frac{(\lambda\sqrt{T}+\beta_T)^2}{4\alpha_T}} + e^{|\lambda|\sqrt{T} + \beta_T}\right)$$

and thus

$$\overline{\lim_{T\to\infty}} \frac{1}{T} \log \mathcal{J}_T \leqslant f \lim_{T\to\infty} \frac{1}{T} \log \beta_T + \lim_{T\to\infty} \frac{1}{T} \frac{(\lambda\sqrt{T} + \beta_T)^2}{4\alpha_T} = -f\,\frac{d}{2} + \frac{\lambda^2}{d-b}. \tag{2.15}$$

We show alike by using the lower bounds of Lemmas 2.4.5 and 2.4.3 that

$$\underline{\lim_{T\to\infty}} \frac{1}{T} \log \mathcal{J}_T \geqslant -f\,\frac{d}{2} + \frac{\lambda^2}{d-b} \tag{2.16}$$

and we finally obtain

$$\lim_{T\to\infty} \frac{1}{T} \log \mathcal{J}_T = -f\,\frac{d}{2} + \frac{\lambda^2}{d-b}. \tag{2.17}$$

• For all $\lambda \leqslant 0$ and $\gamma < 0$: With Lemma 2.4.4, we know that $K_T = O\left((\beta_T)^f\right)$. Thus, $H_T + K_T = H_T + O\left((\beta_T)^f\right)$. Lemma 2.4.2 gives bounds for H_T, which lead to

$$\overline{\lim_{T\to\infty}} \frac{1}{T} \log \mathcal{J}_T \leqslant \lim_{T\to\infty} \frac{1}{T} \log \left((\beta_T)^f\, h_{T,\lambda,\gamma} \exp\left(\frac{\gamma^2 T}{4g}\right)\right) + \lim_{T\to\infty} \frac{1}{T} \log \left(1 + C h_{T,\lambda,\gamma}^{-1} e^{-\frac{\gamma^2 T}{4g}}\right)$$

where C is some positive constant and

$$h_{T,\lambda,\gamma} = \frac{2^{2-f}}{\Gamma(f+1)} \sqrt{\pi} g^{-3/2} |\gamma| \sqrt{T}\, e^{|\lambda| + \alpha_T/T + \beta_T/\sqrt{T}}.$$

Using the fact that $h_{T,\lambda,\gamma}^{-1} e^{-\frac{\gamma^2 T}{4g}}$ tends to zero as T goes to infinity, we obtain

$$\overline{\lim_{T\to\infty}} \frac{1}{T} \log \mathcal{J}_T \leqslant -f\,\frac{d}{2} + \frac{\gamma^2}{4g}. \tag{2.18}$$

We obtain the same lower bound by using the lower bound in Lemma 2.4.2.
• For all $\lambda \leqslant 0$ and $\gamma \geqslant 0$: Lemma 2.4.3 gives $H_T = O\left((\beta_T)^f \exp\left(|\lambda|\sqrt{T}\right)\right)$ and by Lemma 2.4.4, we know that $K_T = O\left((\beta_T)^f\right)$. Consequently,

$$\overline{\lim_{T\to\infty}} \frac{1}{T} \log \mathcal{J}_T \leqslant f \lim_{T\to\infty} \frac{1}{T} \log \beta_T = -f\,\frac{d}{2}. \tag{2.19}$$

And the lower bounds given in Lemmas 2.4.3 and 2.4.4 lead to

$$\underline{\lim_{T\to\infty}} \frac{1}{T} \log \mathcal{J}_T \geqslant f \lim_{T\to\infty} \frac{1}{T} \log \beta_T = -f\,\frac{d}{2}. \tag{2.20}$$

• For all $\lambda > 0$ and $\gamma < 0$: using Lemmas 2.4.5 and 2.4.2, we obtain that

$$H_T + K_T \leqslant C\,(\beta_T)^f \left(\overline{h}_{T,\lambda} \exp\left(\frac{\gamma^2 T}{4g}\right) + \left(\beta_T + \lambda\sqrt{T}\right)^{2\gamma+2g-1} \exp\left(\frac{(\lambda\sqrt{T}+\beta_T)^2}{4\alpha_T}\right)\right)$$

where $\overline{h}_{T,\lambda} = \sqrt{T}\, e^{|\lambda|+\alpha_T/T+\beta_T/\sqrt{T}}$ and C is some positive constant. Thus

$$\overline{\lim_{T\to\infty}} \frac{1}{T} \log \mathcal{J}_T \leqslant \lim_{T\to\infty} \frac{1}{T} \log \left((\beta_T)^f \, e^{T\times \max\left(\frac{\gamma^2}{4g};\frac{\lambda^2}{d-b}\right)} \right) = -f\frac{d}{2} + \max\left(\frac{\gamma^2}{4g};\frac{\lambda^2}{d-b}\right).$$

We also show that

$$H_T + K_T \geqslant C\,(\beta_T)^f \left(\underline{h}_{T,\lambda} \exp\left(\frac{\gamma^2 T}{4g}\right) + m_{\gamma,\lambda,T}\, \exp\left(\frac{\lambda^2 T}{4\alpha_T}\right)\right)$$

where C is still some positive constant and $\underline{h}_{T,\lambda} = \sqrt{T}\, e^{-|\lambda|-\alpha_T/T}$. It leads to

$$\underline{\lim_{T\to\infty}} \frac{1}{T} \log \mathcal{J}_T \geqslant -f\frac{d}{2} + \max\left(\frac{\gamma^2}{4g};\frac{\lambda^2}{d-b}\right). \tag{2.21}$$

□

2.5 Proofs of the LDPs for the couples of simplified estimators

2.5.1 Proof of Theorem 2.2.1

We will now establish an LDP for the first couple of simplified estimators. We notice that $(\widetilde{a}_T, \widetilde{b}_T) = f(\sqrt{X_T/T}, S_T, \Sigma_T)$ where f is the function defined on $\{(x,y,z) \in \mathbb{R}^3 | yz \neq 1\}$ by

$$f(x,y,z) = \left(\frac{2zy - x^2}{yz-1}, \frac{(x^2-2)z}{yz-1}\right).$$

Thus, we first compute an LDP for the triplet $\left(\sqrt{X_T/T}, S_T, \Sigma_T\right)$ and then apply the contraction principle to the obtained rate function.

Lemma 2.5.1. *The sequence* $\left\{\left(\sqrt{X_T/T}, S_T, \Sigma_T\right)\right\}$ *satisfies an LDP with good rate function:*

$$I(x,y,z) = \begin{cases} \dfrac{ab}{4} + \dfrac{b^2}{8}y + \dfrac{(a-2)^2}{8}z - \dfrac{b}{4}x^2 + \dfrac{(x^2+2)^2 z}{8(yz-1)} & \textit{if } x \geqslant 0, y, z, yz-1 > 0 \\ +\infty & \textit{otherwise.} \end{cases}$$

Proof. See appendix D. □

As f is continuous over $\mathcal{D}_I = \{(x,y,z) \in \mathbb{R}^3 | I(x,y,z) < +\infty\}$, we deduce from Lemma 2.5.1 together with the contraction principle that $(\widetilde{a}_T, \widetilde{b}_T)$ satisfies an LDP with good rate function $J_{a,b}$ given by

$$J_{a,b}(\alpha,\beta) = \inf_{\mathcal{D}_I} \{ I(x,y,z) \,|\, f(x,y,z) = (\alpha,\beta)\} \tag{2.22}$$

which reduces to

$$J_{a,b}(\alpha,\beta) = \inf_{\mathcal{D}_I} \left\{ \frac{ab}{4} + \frac{b^2}{8}y + \frac{(a-2)^2}{8}z - \frac{b}{4}x^2 + \frac{(x^2+2)^2 z}{8(yz-1)} \,|f(x,y,z) = (\alpha,\beta) \right\}$$

where the infimum over the empty set is equal to the infinity. One easily see that $J_{a,b}(\alpha,\beta) = +\infty$ as soon as $\alpha = 2$ and $\beta \neq 0$ or $\beta = 0$ and $\alpha \neq 2$, since we take the infimum over the empty set. For the remaining particular case $(\alpha,\beta) = (2,0)$, as $z > 0$ on $\mathcal{D}_I$, the only way of satisfying $f(x,y,z) = (2,0)$ is to take $x^2 = 2$. Therefore

$$J_{a,b}(2,0) = \inf_{y>0,z>0,yz-1>0} \left\{ \frac{ab}{4} + \frac{b^2}{8}y + \frac{(a-2)^2}{8}z - \frac{b}{2} + \frac{2z}{yz-1} \right\} \tag{2.23}$$

The stationnary points are given by

$$y_0 = -\frac{a+2}{b} \qquad \text{and} \qquad z_0 = -\frac{b}{a-2}$$

which leads to

$$J_{a,b}(2,0) = -b \tag{2.24}$$

Otherwise, the condition $f(x,y,z) = (\alpha,\beta)$ implies that

$$z = \frac{\beta}{2-\alpha}, \quad yz - 1 = \frac{\beta y + \alpha - 2}{2-\alpha} \quad \text{and} \quad x^2 = \beta y + \alpha.$$

Thus, if $\frac{\beta}{2-\alpha} < 0$ then $J_{a,b}(\alpha,\beta) = +\infty$, otherwise

$$J_{a,b}(\alpha,\beta) = \inf_{\mathcal{D}_y} \left\{ \frac{ab}{4} + \frac{(a-2)^2}{8}\frac{\beta}{2-\alpha} - \frac{b}{4}\alpha + (b^2 - 2b\beta)\frac{y}{8} + \frac{\beta}{8}\frac{(\beta y + \alpha + 2)^2}{\beta y + \alpha - 2} \right\}$$

where $\mathcal{D}_y = \{y > \frac{2-\alpha}{\beta} | \beta y + \alpha \geqslant 0\}$. We set on $\mathcal{D}_y$,

$$g(y) := \frac{ab}{4} + \frac{(a-2)^2}{8}\frac{\beta}{2-\alpha} - \frac{b}{4}\alpha + (b^2 - 2b\beta)\frac{y}{8} + \frac{\beta}{8}\frac{(\beta y + \alpha + 2)^2}{\beta y + \alpha - 2}. \tag{2.25}$$

Its derivative vanishes at point

$$y_0 = \frac{2-\alpha}{\beta} - \frac{4}{b-\beta}$$

if $\beta \neq b$ and at $y_0 = \frac{2-\alpha}{\beta}$ if $\beta = b$. Depending on the values of β and α, the infimum will be reached either at the critical point y_0 or at the boundary of the domain $\mathcal{D}_y$.

• **For $\beta < 0$:** Only the case $\alpha > 2$ has not been considered yet. The condition $\beta y + \alpha \geqslant 0$ implies that $y \leqslant -\frac{\alpha}{\beta}$. Therefore,

$$J_{a,b}(\alpha,\beta) = \inf_{\frac{2-\alpha}{\beta} < y \leqslant -\frac{\alpha}{\beta}} g(y).$$

For $\beta \leqslant b$, y_0 is not inside the domain over which we take the infimum. Moreover g tends to infinity when y tends to $\frac{2-\alpha}{\beta}$, so necessarily

$$J_{a,b}(\alpha,\beta) = g\left(-\frac{\alpha}{\beta}\right) = \frac{(a-2)^2\beta}{8(2-\alpha)}\left(1+\frac{(2-\alpha)b}{\beta(a-2)}\right)^2 - \frac{\beta}{4}\left(1-\frac{b}{\beta}\right)^2.$$

For $\beta > b$, the condition $\frac{2-\alpha}{\beta} - \frac{4}{b-\beta} > \frac{2-\alpha}{\beta}$ is always satisfied and $\frac{2-\alpha}{\beta} - \frac{4}{b-\beta} \leqslant \frac{-\alpha}{\beta}$ if and only if $\beta \geqslant \frac{b}{3}$. Consequently, if $b < \beta \leqslant \frac{b}{3}$, the derivative does not vanish on the domain and we find the same value of $J_{a,b}$ as above, while if $\frac{b}{3} < \beta < 0$, we get

$$J_{a,b}(\alpha,\beta) = g\left(\frac{2-\alpha}{\beta} - \frac{4}{b-\beta}\right) = \frac{(a-2)^2\beta}{8(2-\alpha)}\left(1+\frac{(2-\alpha)b}{\beta(a-2)}\right)^2 + 2\beta - b.$$

• **For $\boldsymbol{\beta} > 0$:** the condition $\beta y + \alpha \geqslant 0$ becomes $y \geqslant -\frac{\alpha}{\beta}$ which is smaller than $\frac{2-\alpha}{\beta}$. Consequently,

$$J_{a,b}(\alpha,\beta) = \inf_{\frac{2-\alpha}{\beta}<y} g(y).$$

The derivative is equal to zero for $y = \frac{2-\alpha}{\beta} - \frac{4}{b-\beta}$ which is always greater than $\frac{2-\alpha}{\beta}$ so inside the domain. We get

$$J_{a,b}(\alpha,\beta) = g\left(\frac{2-\alpha}{\beta} - \frac{4}{b-\beta}\right) = \frac{(a-2)^2\beta}{8(2-\alpha)}\left(1+\frac{(2-\alpha)b}{\beta(a-2)}\right)^2 + 2\beta - b.$$

□

2.5.2 Proofs of Corollaries 2.2.2 and 2.2.1

Using the contraction principle again, we deduce from theorem 2.2.1 LDPs for both estimators. We begin with $\widetilde{b}_T$ because the calculations are really straightforward.

Proof of Corollary 2.2.2. From the contraction principle, we know that

$$J_b(\beta) = \inf_{\alpha\in\mathbb{R}} I_{a,b}(\alpha,\beta).$$

We have directly that $J_b(0) = J_{a,b}(2,0) = -b$ and that, for $\beta \neq 0$,

$$J_b(\beta) = J_{a,b}(2+\beta\,\frac{a-2}{b},\beta).$$

This leads to the result noticing that it is continuous at point zero. □

Proof of Corollary 2.2.1. With the contraction principle again, we have

$$J_a(\alpha) = \inf_{\beta\in\mathbb{R}} J_{a,b}(\alpha,\beta).$$

• **For $\boldsymbol{\alpha} = 2$:** We easily show that $J_a(2) = J_{a,b}(2,0) = -b$.
• **For $\boldsymbol{\alpha} < 2$:** Investigating for critical points, we obtain that

$$J_a(\alpha) = \inf_{\beta>0} \left\{ \frac{(a-2)^2\beta}{8(2-\alpha)} \left(1 + \frac{(2-\alpha)b}{\beta(a-2)}\right)^2 + 2\beta - b \right\} = J_{a,b}(\alpha, \beta_0),$$

where β_0 is the critical point given by

$$\beta_0 = -(2-\alpha)\, b \left((a-2)^2 + 16\,(2-\alpha)\right)^{-1/2}.$$

This straightforwardly leads to the announced result

$$J_a(\alpha) = \frac{b}{4}\left(a - 6 - \sqrt{(a-2)^2 + 16\,(2-\alpha)}\right).$$

• **For $\boldsymbol{\alpha} > 2$:** $J_{a,b}(\alpha,\beta) = \min(I_1, I_2)$ where

$$I_1 = \inf_{\beta \leqslant \frac{b}{3}} \left\{ \frac{(a-2)^2\beta}{8(2-\alpha)} \left(1 + \frac{(2-\alpha)b}{\beta(a-2)}\right)^2 - \frac{\beta}{4}\left(1 - \frac{b}{\beta}\right)^2 \right\} \tag{2.26}$$

and

$$I_2 = \inf_{\frac{b}{3} \leqslant \beta < 0} \left\{ \frac{(a-2)^2\beta}{8(2-\alpha)} \left(1 + \frac{(2-\alpha)b}{\beta(a-2)}\right)^2 + 2\beta - b \right\}. \tag{2.27}$$

For I_2, with the calculations of the second case, we already know that the derivative equals zero for β_0 satisfying

$$\beta_0^2 = \frac{(2-\alpha)^2\, b^2}{(a-2)^2 + 16\,(2-\alpha)}.$$

This is well defined if and only if $(a-2)^2 + 16\,(2-\alpha) > 0$. And, β_0 is in the domain if and only if $\beta_0 < 0$ and $\beta_0^2 \leqslant \frac{b^2}{9}$. All those conditions are fulfilled if and only if $9(2-\alpha)^2 \leqslant (a-2)^2 + 16(2-\alpha)$. As $\alpha > 2$, we obtain the condition

$$\alpha < \ell_a := \frac{10}{9} + \frac{1}{9}\sqrt{64 + 9(a-2)^2}$$

and $\beta_0 = -(2-\alpha)\, b \left((a-2)^2 + 16\,(2-\alpha)\right)^{-1/2}$. Thus, for $2 < \alpha < \ell_a$

$$I_2 = J_{a,b}(\alpha, \beta_0) = \frac{b}{4}\left(a - 6 - \sqrt{(a-2)^2 + 16(2-\alpha)}\right). \tag{2.28}$$

Otherwise, for $\alpha \geqslant \ell_a$, the derivative never vanishes on the domain and the minimum is reached at one of the boundaries. When β goes to zero, the function goes to infinity. Consequently,

$$I_2 = J_{a,b}(\alpha, \frac{b}{3}) = \frac{b}{3}\left(\frac{\left((a-2) + 3(2-\alpha)\right)^2}{8(2-\alpha)} - 1\right). \tag{2.29}$$

For I_1, the idea is similar. The derivative equals zero for β_1 satisfying

$$\beta_1^2 = \frac{\alpha\, b^2\,(\alpha-2)}{(a-2)^2 + 2\,(\alpha-2)}.$$

This time the domain is $\left\{\beta < \frac{b}{3}\right\}$. So β_1 is inside the domain if $\beta_1 < 0$ and $\beta_1^2 \geqslant \frac{b^2}{9}$. It leads us to

$$\beta_1 = b\sqrt{\alpha\,(\alpha-2)}\,\left((a-2)^2 + 2\,(\alpha-2)\right)^{-1/2}$$

with the condition $9\alpha(\alpha-2) > (a-2)^2 + 2(\alpha-2)$ on α which gives the same limit value ℓ_a. We get

$$I_1 = \begin{cases} J_{a,b}(\alpha, \frac{b}{3}) = \dfrac{b}{3}\left(\dfrac{\left((a-2)+3(2-\alpha)\right)^2}{8(2-\alpha)} - 1\right) & \text{if } 2 < \alpha < \ell_a \\ J_{a,b}(\alpha, \beta_1) = \dfrac{b}{4}\left(a - \sqrt{\alpha\left(\dfrac{(a-2)^2}{\alpha-2} + 2\right)}\right) & \text{if } \alpha \geqslant \ell_a. \end{cases} \tag{2.30}$$

We now come back to $J_{a,b}(\alpha,\beta)$. Combining (2.28), (2.29) and (2.30), we obtain

$$J_{a,b}(\alpha,\beta) = \min(I_1, I_2) = \begin{cases} \min\left(J_{a,b}(\alpha,\beta_0), J_{a,b}(\alpha,\frac{b}{3})\right) & \text{if } 2 < \alpha \leqslant \ell_a \\ \min\left(J_{a,b}(\alpha,\beta_1), J_{a,b}(\alpha,\frac{b}{3})\right) & \text{if } \alpha > \ell_a, \end{cases}$$

and it is easy to deduce that

$$J_{a,b}(\alpha,\beta) = \begin{cases} J_{a,b}(\alpha,\beta_0) & \text{if } 2 < \alpha \leqslant \ell_a \\ J_{a,b}(\alpha,\beta_1) & \text{if } \alpha > \ell_a. \end{cases} \tag{2.31}$$

This leads to the conclusion, noticing that it is continuous at the point $\alpha = 2$. $\square$

2.5.3 Proof of Theorem 2.2.2

We consider the second couple of simplified estimators defined by

$$\breve{a}_T = \frac{S_T\,(2\,\Sigma_T + L_T)}{V_T} \quad \text{and} \quad \breve{b}_T = \frac{-2\,\Sigma_T - L_T}{V_T}.$$

We notice that $\left(\breve{a}_T, \breve{b}_T\right) = h(S_T, \Sigma_T, \mathcal{L}_T)$, where h is the function defined on the subset $\{(y,z,t) \in \mathbb{R}^3 | yz - 1 \neq 0\}$ by

$$h(y,z,t) = \left(\frac{y\,(2z - t^2\mathbf{1}_{t\leqslant 0} + t\mathbf{1}_{t>0})}{yz-1}, \frac{t^2\mathbf{1}_{t\leqslant 0} - t\mathbf{1}_{t>0} - 2z}{yz-1}\right). \tag{2.32}$$

Once again, we start by computing an LDP for the triplet $(S_T, \Sigma_T, \mathcal{L}_T)$ and then we deduce an LDP for the couple of estimators applying the contraction principle to the obtained rate function.

Lemma 2.5.2. *The sequence $\{(S_T, \Sigma_T, \mathcal{L}_T)\}$ satisfies an LDP with good rate function*

$$\widetilde{I}(y,z,t) = \begin{cases} \frac{ab}{4} + \frac{b^2}{8}y + \frac{(a-2)^2}{8}z + \frac{a}{4}t^2 + \frac{4z(yt^2+1)+t^4y}{8(yz-1)} & \text{if } t \leqslant 0, y > 0, z > 0 \\ & \text{and } yz-1>0, \\ +\infty & \text{otherwise.} \end{cases}$$

Proof. The proof is postponed to Appendix E. □

As the function h given by (2.32) is continuous over the domain where the rate function $\widetilde{I}$ of Lemma 2.5.2 is finite, the contraction principle applied to $\widetilde{I}$ shows that $\left(\breve{a}_T, \breve{b}_T\right)$ satisfies an LDP with good rate function $K_{a,b}$ given by

$$K_{a,b}(\alpha,\beta) = \inf_{\mathcal{D}_{\widetilde{I}}} \left\{ \widetilde{I}(y,z,t) | h(y,z,t) = (\alpha,\beta) \right\}$$

which reduces to

$$K_{a,b}(\alpha,\beta) = \inf_{\mathcal{D}_{\alpha,\beta}} \left\{ \frac{ab}{4} + \frac{b^2}{8}y + \frac{(a-2)^2}{8}z + \frac{a}{4}t^2 + \frac{4z(yt^2+1)+t^4y}{8(yz-1)} \right\}$$

where $\mathcal{D}_{\alpha,\beta} = \{(y,z,t) \in \mathbb{R}^3 | y > 0, z > 0, yz-1 > 0, t \leqslant 0 \text{ and } h(y,z,t) = (\alpha,\beta)\}$ and the infimum over the empty set is equal to infinity. The condition $h(y,z,t) = (\alpha,\beta)$ implies that

$$\beta y = -\alpha \quad \text{and} \quad t^2 = (2-\alpha)z - \beta.$$

It gives us some additional conditions on the parameters. First of all, we notice that for α or β equal to zero, $\mathcal{D}_{\alpha,\beta}$ is not empty if and only if the other one is also zero. Thus, $K_{a,b}(\alpha,\beta) = +\infty$ over $\{0\} \times \mathbb{R}_*^+$ and $\mathbb{R}_*^+ \times \{0\}$. If $\alpha = \beta = 0$, necessarily $t^2 = 2z$ and we easily obtain the critical points leading to

$$K_{a,b}(0,0) = -\frac{b}{4}\left(4 - a + \sqrt{a^2+16}\right).$$

Moreover, for $\beta \neq 0$, as $y = -\frac{\alpha}{\beta}$, $\mathcal{D}_{\alpha,\beta}$ is empty as soon as α and β have the same sign. So $K_{a,b}(\alpha,\beta) = +\infty$ over $\mathbb{R}_*^+ \times \mathbb{R}_*^+$ and $\mathbb{R}_*^- \times \mathbb{R}_*^-$. Besides, both expressions will give us boundaries for z depending on the sign of α, β and $2-\alpha$, because t^2 must be positive and z must be greater than $\frac{1}{y}$. Assuming that all conditions are fulfilled, we derive from Lemma 2.5.2 that

$$\widetilde{I}\left(-\frac{\alpha}{\beta}, z, \sqrt{(2-\alpha)z-\beta}\right) = A_{\alpha,\beta} + C_\alpha z - \frac{2\beta z}{\alpha z + \beta}$$

where C_α and $A_{\alpha,\beta}$ do not depend on z, and are defined by

$$C_\alpha = \frac{1}{8}(a-\alpha)^2 + 2 - \alpha \quad \text{and} \quad A_{\alpha,\beta} = -\frac{\alpha}{\beta}\frac{b^2}{8} + \frac{ab}{4} - \frac{a\beta}{4} + \frac{\alpha\beta}{8}. \tag{2.33}$$

Thus,

$$K_{a,b}(\alpha,\beta) = \inf_{\mathcal{D}_z} \left\{ A_{\alpha,\beta} + C_\alpha z - \frac{2\beta z}{\alpha z + \beta} \right\} \tag{2.34}$$

where $\mathcal{D}_z = \left\{z > 0 | z > -\frac{\beta}{\alpha} \text{ and } (2-\alpha)z - \beta \geqslant 0\right\}$. Depending on the values of α and β, the infimum will be reached either at a critical point or at the boundary of the domain $\mathcal{D}_z$.

• **For $\boldsymbol{\alpha} < \mathbf{0}$:** Only remains the case $\beta > 0$. As $0 < \frac{\beta}{2-\alpha} \leqslant \frac{\beta}{-\alpha}$ the domain $\mathcal{D}_z$ reduces to $\mathcal{D}_z = \left\{z > -\frac{\beta}{\alpha}\right\}$. We look for critical points of $A_{\alpha,\beta} + C_\alpha z - \frac{2\beta z}{\alpha z+\beta}$ over this domain. We find that critical points z_0 satisfy

$$(\alpha z_0 + \beta)^2 = \frac{2\beta^2}{C_\alpha}$$

We notice that for α negative C_α is always positive. So the only critical point that remains in the domain is $z_0 = -\frac{\beta}{\alpha}\left(1 + \sqrt{\frac{2}{C_\alpha}}\right)$. As the function tends to infinity on the boundaries of $\mathcal{D}_z$, it actually reaches the infimum we were looking for at this critical point z_0. Replacing it into (2.34), we find

$$\begin{aligned} K_{a,b}(\alpha,\beta) &= A_{\alpha,\beta} + C_\alpha z_0 - \frac{2\beta z_0}{\alpha z_0 + \beta} \\ &= \frac{a}{4}(b-\beta) - \frac{\alpha}{8\beta}\left(b^2 - \beta^2\right) - \frac{\beta}{\alpha}\left(\sqrt{2} + \sqrt{C_\alpha}\right)^2. \end{aligned} \tag{2.35}$$

• **For $\mathbf{0} < \boldsymbol{\alpha} \leqslant \mathbf{2}$:** As $\beta < 0$, the condition $(2-\alpha)z - \beta \geqslant 0$ is always verified. Thus $\mathcal{D}_z = \left\{z > -\frac{\beta}{\alpha}\right\}$. We obtain the same critical point than in the firsta case and the infimum is still given by formula (2.35).

• **For $\boldsymbol{\alpha} > \mathbf{2}$.** The case $\beta > 0$ has already been seen. We consider $\beta < 0$. We investigate the critical points of $A_{\alpha,\beta} + C_\alpha z - \frac{2\beta z}{\alpha z+\beta}$ over the domain $\mathcal{D}_z$, given in this case by $\mathcal{D}_z = \left\{-\frac{\beta}{\alpha} < z \leqslant \frac{\beta}{2-\alpha}\right\}$. We need to distinguish cases depending on the sign of C_α.

If $\alpha < a + 4 - 2\sqrt{a}$ (which is greater than 2 because $a > 2$), C_α is positive and we find the same critical points than in the first case. The condition $z > -\frac{\beta}{\alpha}$ is still only verified by

$$z_0 = -\frac{\beta}{\alpha}\left(1 + \sqrt{\frac{2}{C_\alpha}}\right).$$

But, the condition $z_0 \leqslant \frac{\beta}{2-\alpha}$ is not satisfied for all α in $[2, a + 4 - 2\sqrt{a}[$. Indeed,

$$z_0 > \frac{\beta}{2-\alpha} \text{ if and only if } \frac{2}{C_\alpha}(\alpha-2)^2 > 4 \tag{2.36}$$

which leads to the following condition on α:

$$3\alpha^2 + 2(a-4)\alpha - (a^2+8) > 0.$$

One of the roots is negative. The other one is inside $[2, a + 4 - 2\sqrt{a}[$:

$$\alpha_a = -\frac{2}{3}\left(\frac{a}{2} - 2 - \sqrt{a^2 - 2a + 4}\right). \tag{2.37}$$

Thus, for $\alpha \leqslant \alpha_a$, the infimum is reached at z_0 and is given by (2.35), while for $\alpha > \alpha_a$, z_0 is not inside the domain $\mathcal{D}_z$ so the derivative does not vanish and the infimum is reached at one of the boundaries. We notice that for z tending to $-\frac{\beta}{\alpha}$, $A_{\alpha,\beta}+C_\alpha z-\frac{2\beta z}{\alpha z+\beta}$ tends to the infinity. Thus, the infimum is reached at the other boundary of the domain: $z_0=\frac{\beta}{2-\alpha}$. Replacing it into (2.34), we obtain

$$K_{a,b}(\alpha,\beta)=\frac{a}{4}(b-\beta)-\frac{\alpha}{8\beta}\left(b^2-\beta^2\right)-\frac{\beta\,(a-\alpha)^2}{8\,(\alpha-2)}. \tag{2.38}$$

If $\alpha\in[a+4-2\sqrt{a};a+4+2\sqrt{a}]$, then C_α is null or negative so the derivative cannot vanish and $K_{a,b}$ is given by (2.38).

If $\alpha>a+4+2\sqrt{a}$, then C_α is positive but as $\alpha>\alpha_a$ the critical point z_0 is greater than $\frac{\beta}{2-\alpha}$ so outside the domain $\mathcal{D}_z$. The infimum is reached at the boundary $\frac{\beta}{2-\alpha}$ of the domain and is still given by (2.38). □

2.5.4 Proofs of Corollaries 2.2.3 and 2.2.4

Proof of Corollary 2.2.4. The result is a direct application of the contraction principle to the rate function $K_{a,b}$ of Theorem 2.2.2. We did not obtain an explicit expression of the infimum. □

Proof of Corollary 2.2.3. With the contraction principle again, we know that $(\breve{a}_T)$ satisfies an LDP with good rate function

$$K_a(\alpha)=\inf_{\beta\in\mathbb{R}}K_{a,b}(\alpha,\beta).$$

- **For $\boldsymbol{\alpha}=\mathbf{0}$.** We easily show that $K_a(0)=K_{a,b}(0,0)$.
- **For $\boldsymbol{\alpha}<\mathbf{0}$.** We rewrite

$$K_a(\alpha)=\inf_{\beta>0}\left\{A_{\alpha,\beta}-\frac{\beta}{\alpha}\left(\sqrt{2}+\sqrt{C_\alpha}\right)^2\right\}.$$

The critical points β_b satisfy

$$\beta_b^2=\frac{\alpha^2b^2}{16\sqrt{2}\sqrt{C_\alpha}+a^2-8\alpha+32}, \tag{2.39}$$

which is clearly well defined for all $\alpha<0$. Using the fact that $\alpha<0$ and $b<0$ and as β_b must be positive, we obtain

$$\beta_b=b\alpha\left(16\sqrt{2}\sqrt{C_\alpha}+a^2-8\alpha+32\right)^{-1/2} \tag{2.40}$$

and

$$K_a(\alpha)=K_{a,b}(\alpha,\beta_b). \tag{2.41}$$

- **For $\mathbf{0}<\boldsymbol{\alpha}\leqslant\boldsymbol{\alpha_a}$.** This time

$$K_a(\alpha)=\inf_{\beta<0}\left\{A_{\alpha,\beta}-\frac{\beta}{\alpha}\left(\sqrt{2}+\sqrt{C_\alpha}\right)^2\right\}.$$

The critical points β_b are still given by (2.39). It is well defined for all $\alpha < \alpha_a$. Namely, $16\sqrt{2}\sqrt{C_\alpha}+a^2-8\alpha+32$ is a decreasing function on α over this domain and is positive at the point α_a. Indeed, we know from (2.36) and (2.37) that $\alpha_a > 2$ satisfies $(\alpha_a-2)^2 = 2C_{\alpha_a}$, which leads to $16\sqrt{2}\sqrt{C_{\alpha_a}} + a^2 - 8\alpha_a + 32 = 8\alpha_a + a^2 > 0$. This time $\alpha > 0$ and β_b must be negative but we obtain anyway the same critical point β_b given by (2.40) and the infimum K_a by (2.41).

• **For $\boldsymbol{\alpha > \alpha_a}$.** This time

$$K_a(\alpha) = \inf_{\beta<0}\left\{\frac{(a-2)^2\beta}{8(2-\alpha)}\left(1+\frac{(2-\alpha)b}{\beta(a-2)}\right)^2 - \frac{\beta}{4}\left(1-\frac{b}{\beta}\right)^2\right\} = J_a(\alpha).$$

Using the results of the third case of the proof of Corollary 2.2.1, we obtain the same critical point and the same infimum. □

2.6 Proof of Theorem 2.2.3

We now come back to the MLE $\left(\widehat{a}_T, \widehat{b}_T\right)$. As we did for the couples of simplified estimators, we first establish an LDP for the quadruplet $\mathcal{Q}_T$ and we deduce an LDP for the MLE via the contration principle.

2.6.1 Existence of an LDP

Lemma 2.6.1. *The quadruplet $\mathcal{Q}_T = (\sqrt{X_T/T}, S_T, \Sigma_T, \mathcal{L}_T)$ satisfies an LDP with good rate function Λ^* given by*

$$\Lambda^*(x,y,z,t) = +\infty \text{ for } x<0, t>0, y \leqslant 0, z \leqslant 0 \text{ or } yz-1 \leqslant 0. \tag{2.42}$$

and, otherwise,

$$\Lambda^*(x,y,z,t) = \sup_{\mathcal{D}_{d,f}} h(d,f) \tag{2.43}$$

where $\mathcal{D}_{d,f} = \{d>0, f>0\}$ and, with $\varphi(f) = 2f+a+2$,

$$h(d,f) = \frac{1}{4}\left(t\sqrt{\varphi(f)} - x\sqrt{d-b}\right)^2 + y\,\frac{b^2-d^2}{8} + \frac{(a-2)^2-4f^2}{8}\,z + \frac{d}{2}(1+f) + \frac{ab}{4}.$$

Proof. Using Gärtner-Ellis theorem, we have to compute the Fenchel-Legendre transform Λ^* of the cumulant generating function Λ defined in Proposition 2.4.1:

$$\Lambda^*(x,y,z,t) = \sup_{\mathcal{D}}\{x\lambda + y\mu + z\nu + t\gamma - \Lambda(x,y,z,t)\} \tag{2.44}$$

where $\mathcal{D} = \left\{\lambda\in\mathbb{R},\ \gamma\in\mathbb{R},\ \mu<\frac{b^2}{8},\ \nu<\frac{(a-2)^2}{8}\right\}$. We show with the same arguments than for the other LDP proofs that

$$\Lambda^*(x,y,z,t) = +\infty \text{ for } x<0, t>0, y \leqslant 0, z \leqslant 0 \text{ or } yz-1 \leqslant 0. \tag{2.45}$$

Besides, for $x \geqslant 0$, the part involving λ in the function we want to optimize is always negative for $\lambda \leqslant 0$ and sometimes positive for $\lambda > 0$. Thus the supremum is necessarily reached for some $\lambda > 0$. With the same argument for $t \leqslant 0$, we show that we only have to consider $\gamma < 0$. Replacing μ and ν by their expression in d and f, the domain $\mathcal{D}$ over which we optimize reduces to $\mathcal{D} = \{\lambda > 0, \gamma < 0, d > 0, f > 0\}$. Replacing Λ by its value leads to

$$\Lambda^*(x, y, z, t) = \max(S_1, S_2),$$

where

$$S_1 = \sup_{\mathcal{D}\cap\left\{\frac{\gamma^2}{\lambda^2}\leqslant\frac{\varphi(f)}{d-b}\right\}} \left\{x\lambda + y\frac{b^2-d^2}{8} + z\frac{(a-2)^2-4f^2}{8} + t\gamma + \frac{d}{2}(1+f) + \frac{ab}{4} - \frac{\lambda^2}{d-b}\right\}$$

and

$$S_2 = \sup_{\mathcal{D}\cap\left\{\frac{\gamma^2}{\lambda^2}\geqslant\frac{\varphi(f)}{d-b}\right\}} \left\{x\lambda + y\frac{b^2-d^2}{8} + z\frac{(a-2)^2-4f^2}{8} + t\gamma + \frac{d}{2}(1+f) + \frac{ab}{4} - \frac{\gamma^2}{\varphi(f)}\right\}.$$

We first consider S_1. The domain over which we take the supremum is given by

$$\mathcal{D} \cap \left\{\frac{\gamma^2}{\lambda^2} \leqslant \frac{\varphi(f)}{d-b}\right\} = \left\{\lambda < 0, d > 0, f > 0, 0 > \gamma \geqslant -\sqrt{\frac{\varphi(f)}{d-b}}\lambda\right\}.$$

Over this domain, as $t \leqslant 0$, $0 \leqslant t\gamma \leqslant -t\sqrt{\frac{\varphi(f)}{d-b}}\lambda$, so that the supremum of $t\gamma$ is equal to $-t\lambda\sqrt{\frac{\varphi(f)}{d-b}}$. Thus, if we set $\mathcal{D}_1 = \{\lambda < 0, d > 0, f > 0\}$,

$$S_1 = \sup_{\mathcal{D}_1} \left\{x\lambda + y\frac{b^2-d^2}{8} + z\frac{(a-2)^2-4f^2}{8} - t\lambda\sqrt{\frac{\varphi(f)}{d-b}} + \frac{d}{2}(1+f) + \frac{ab}{4} - \frac{\lambda^2}{d-b}\right\}.$$

The supremum over λ is easy to compute. Indeed, the function is concave on λ and the critical point is given by

$$\lambda = \frac{d-b}{2}\left(x - t\sqrt{\frac{\varphi(f)}{d-b}}\right).$$

Finally, with $\mathcal{D}_{d,f} = \{d > 0, f > 0\}$, we obtain

$$S_1 = \sup_{\mathcal{D}_{d,f}} \left\{\frac{1}{4}\left(t\sqrt{\varphi(f)} - x\sqrt{d-b}\right)^2 + y\frac{b^2-d^2}{8} + \frac{(a-2)^2-4f^2}{8}z + \frac{d}{2}(1+f) + \frac{ab}{4}\right\}.$$

We do the same thing with S_2, computing first the supremum over λ and then over γ. We obtain $S_1 = S_2$, so that

$$\Lambda^*(x, y, z, t) = S_1 = \sup_{\mathcal{D}_{d,f}} h(d, f) \tag{2.46}$$

where

$$h(d, f) = \frac{1}{4}\left(t\sqrt{\varphi(f)} - x\sqrt{d-b}\right)^2 + y\,\frac{b^2-d^2}{8} + \frac{(a-2)^2-4f^2}{8}z + \frac{d}{2}(1+f) + \frac{ab}{4}.$$

□

Remark 2.6.1. *This supremum is not explicitly computable but, as the function h is concave, it is reached for some (d^*, f^*) and this gives the rate function of the LDP satisfied by the quadruplet $\mathcal{Q}_T$.*

Lemma 2.6.2. *The couple $(\widehat{a}_T, \widehat{b}_T)$ satisfies an LDP with good rate function $I_{a,b}$ given over $\mathbb{R}^2$ by*

$$I_{a,b}(\alpha,\beta) = \begin{cases} K_{a,b}(0,0) & \text{if } (\alpha,\beta) = (0,0), \\ J_{a,b}(2,0) & \text{if } (\alpha,\beta) = (2,0), \\ \inf\limits_{\mathcal{D}_{x,t}} \sup\limits_{\mathcal{D}_{d,f}} H(x,t,d,f) & \text{if } (\alpha,\beta) \in \mathcal{D}_1 \cup \mathcal{D}_2 \cup \mathcal{D}_3, \\ +\infty & \text{otherwise.} \end{cases}$$

where $\mathcal{D}_1 = \mathbb{R}^- \times \mathbb{R}^+_*$, $\mathcal{D}_2 =]0,2[\times\mathbb{R}$, $\mathcal{D}_3 = [2,+\infty[\times\mathbb{R}^-_*$, $\mathcal{D}_{d,f} = \{d > 0, f > 0\}$,

$$\mathcal{D}_{x,t} = \left\{ x \geqslant 0, t \leqslant 0 | \frac{x^2-\alpha}{\beta}\frac{t^2+\beta}{2-\alpha} > 1 \right\}$$

and

$$\begin{aligned} H(x,t,d,f) &= \frac{1}{4}\left(t\sqrt{2f+a+2} - x\sqrt{d-b}\right)^2 + \frac{b^2-d^2}{8}\frac{x^2-\alpha}{\beta} \\ &+ \frac{(a-2)^2-4f^2}{8}\frac{t^2+\beta}{2-\alpha} + \frac{d}{2}(1+f) + \frac{ab}{4}. \end{aligned}$$

Proof. $(\widehat{a}_T, \widehat{b}_T) = g(\mathcal{Q}_T)$ where g is the function defined on $\{(x,y,z,t) \in \mathbb{R}^4 | yz - 1 \neq 0\}$ by

$$g(x,y,z,t) = \left(\frac{y\left(2z - t^2\mathbf{1}_{t\leqslant 0} + t\mathbf{1}_{t>0}\right) - x^2}{yz-1}, \frac{t^2\mathbf{1}_{t\leqslant 0} - t\mathbf{1}_{t>0} + (x^2-2)z}{yz-1} \right).$$

As g is continuous over the domain $\mathcal{D}_{\Lambda^*}$ where the rate function Λ^* of Lemma 2.6.1 is finite, the contraction principle applies and give us that the couple $(\widehat{a}_T, \widehat{b}_T)$ satisfies an LDP with good rate function $I_{a,b}$ given by

$$I_{a,b}(\alpha,\beta) = \inf_{\mathcal{D}_{\Lambda^*}} \left\{ \Lambda^*(x,y,z,t) | g(x,y,z,t) = (\alpha,\beta) \right\}. \tag{2.47}$$

The condition $g(x,y,z,t) = (\alpha,\beta)$ gives us a link between x and y and one between t and z:

$$\beta y = x^2 - \alpha \quad \text{and} \quad (2-\alpha) z = t^2 + \beta. \tag{2.48}$$

We first notice that if $\alpha \leqslant 0$ and $\beta < 0$ then y is negative and for all x, z, t, $\Lambda^*(x,y,z,t) = +\infty$ such as $I_{a,b}(\alpha,\beta)$. Similarly, if $\beta \geqslant 0$ and $\alpha > 2$, z is negative and $\Lambda^*(x,y,z,t) = +\infty$ for all x, y, t, then $I_{a,b}(\alpha,\beta) = +\infty$. If $\beta = 0$ and $\alpha < 0$, the first condition in (2.48) leads to x^2 negative and if $\alpha = 2$ and $\beta > 0$ the second condition gives t^2 negative. So, in both cases, we get $I_{a,b}(\alpha,\beta) = +\infty$. We now focus on the values of α and β for which $I_{a,b}$ is not clearly infinite. We first consider the two remaining limit cases : $(0,0)$ and $(2,0)$. If $\alpha = \beta = 0$ then the first condition of (2.48) gives $x^2 = \alpha = 0$ so that

$$I_{a,b}(0,0) = K_{a,b}(0,0). \tag{2.49}$$

Similarly, if $\alpha = 2$ and $\beta = 0$, the second condition implies that $t^2 = (2-\alpha)z = 0$ and consequently

$$I_{a,b}(2,0) = J_{a,b}(2,0). \tag{2.50}$$

For all remaining values of (α, β), we define the function

$$\begin{aligned} H(x,t,d,f) &= \frac{1}{4}\left(t\sqrt{2f+a+2} - x\sqrt{d-b}\right)^2 + \frac{b^2-d^2}{8}\frac{x^2-\alpha}{\beta} \\ &+ \frac{(a-2)^2-4f^2}{8}\frac{t^2+\beta}{2-\alpha} + \frac{d}{2}(1+f) + \frac{ab}{4} \end{aligned} \tag{2.51}$$

and obtain the announced result:

$$I_{a,b}(\alpha,\beta) = \inf_{\mathcal{D}_{x,t}} \sup_{\mathcal{D}_{d,f}} H(x,t,d,f), \tag{2.52}$$

where $\mathcal{D}_{x,t} = \left\{x \geqslant 0, t \leqslant 0 | \frac{x^2-\alpha}{\beta}\frac{t^2+\beta}{2-\alpha} > 1\right\}$ and $\mathcal{D}_{d,f} = \{d > 0, f > 0\}$. $\square$

We were not able to compute $I_{a,b}$ explicitly at this stage. It is the aim of the next subsection.

2.6.2 Evaluating the rate function $I_{a,b}$

Our goal is to show that

$$I_{a,b}(\alpha,\beta) = \min\left(J_{a,b}(\alpha,\beta), K_{a,b}(\alpha,\beta)\right) \tag{2.53}$$

where $J_{a,b}$ and $K_{a,b}$ are the rate functions for the two couples of simplified estimators (see Theorems 2.2.1 and 2.2.2) and $I_{a,b}$ is given by Lemma 2.6.2. We notice that

$$K_{a,b}(\alpha,\beta) = \inf_{t\leqslant 0}\sup_{\mathcal{D}_{d,f}} H(0,t,d,f) \quad \text{and} \quad J_{a,b}(\alpha,\beta) = \inf_{x\geqslant 0}\sup_{\mathcal{D}_{d,f}} H(x,0,d,f).$$

Thus it easily follows that

$$I_{a,b}(\alpha,\beta) \leqslant \min\left(J_{a,b}(\alpha,\beta), K_{a,b}(\alpha,\beta)\right). \tag{2.54}$$

So, we just have to show the inequality in the other side. We denote $\widehat{\theta}_T = \left(\widehat{a}_T, \widehat{b}_T\right)$ and $\overline{\theta}_T = \left(\widetilde{a}_T, \widetilde{b}_T\right)\mathbf{1}_{X_T \geqslant 1} + \left(\breve{a}_T, \breve{b}_T\right)\mathbf{1}_{X_T<1}$.

Lemma 2.6.3. *The estimators $\overline{\theta}_T$ and $\widehat{\theta}_T$ are exponentially equivalent, which means that for all $\varepsilon > 0$,*

$$\limsup_{T\to+\infty}\frac{1}{T}\log\mathbb{P}\left(\| \widehat{\theta}_T - \overline{\theta}_T \| > \varepsilon\right) = -\infty.$$

In particular, as the sequence $(\widehat{\theta}_T)$ satisfies an LDP with good rate function $I_{a,b}$, then the same LDP holds true for $(\overline{\theta}_T)$.

Proof. From the definition of each estimator, we get that

$$\widehat{a}_T - (\widetilde{a}_T \mathbf{1}_{X_T \geqslant 1} + \breve{a}_T \mathbf{1}_{X_T<1}) = \frac{S_T\, L_T \mathbf{1}_{X_T \geqslant 1} - \frac{X_T}{T}\mathbf{1}_{X_T<1}}{V_T}$$

and

$$\widehat{b}_T - \left(\widetilde{b}_T \mathbf{1}_{X_T \geqslant 1} + \breve{b}_T \mathbf{1}_{X_T<1}\right) = \frac{\frac{X_T}{T}\Sigma_T \mathbf{1}_{X_T<1} - L_T \mathbf{1}_{X_T \geqslant 1}}{V_T}.$$

Thus, for all $\varepsilon > 0$,

$$\mathbb{P}\left(\| \widehat{\theta}_T - \overline{\theta}_T \| > \varepsilon\right) \leqslant P_T^\varepsilon + Q_T^\varepsilon + p_T^\varepsilon + q_T^\varepsilon$$

where $P_T^\varepsilon = \mathbb{P}\left(\left|\frac{S_T\, L_T \mathbf{1}_{X_T \geqslant 1}}{V_T}\right| \geqslant \frac{\varepsilon}{2\sqrt{2}}\right)$, $Q_T^\varepsilon = \mathbb{P}\left(\left|\frac{\frac{X_T}{T}\mathbf{1}_{X_T<1}}{V_T}\right| \geqslant \frac{\varepsilon}{2\sqrt{2}}\right)$, $p_T^\varepsilon = \mathbb{P}\left(\left|\frac{\frac{X_T}{T}\Sigma_T \mathbf{1}_{X_T<1}}{V_T}\right| \geqslant \frac{\varepsilon}{2\sqrt{2}}\right)$ and $q_T^\varepsilon = \mathbb{P}\left(\left|\frac{L_T \mathbf{1}_{X_T \geqslant 1}}{V_T}\right| \geqslant \frac{\varepsilon}{2\sqrt{2}}\right)$. For all $\eta > 0$, we have the following upper bounds:

$$\begin{aligned} P_T^\varepsilon &\leqslant \mathbb{P}\left(|S_T| \geqslant \frac{\varepsilon}{2\eta\sqrt{2}}\right) + \mathbb{P}\left(\frac{|L_T \mathbf{1}_{X_T \geqslant 1}|}{V_T} \geqslant \eta\right) \\ &\leqslant \mathbb{P}\left(S_T \geqslant \frac{\varepsilon}{2\eta\sqrt{2}}\right) + \mathbb{P}\left(L_T \mathbf{1}_{X_T \geqslant 1} \geqslant \eta^2\right) + \mathbb{P}\left(V_T \leqslant \eta\right), \end{aligned} \tag{2.55}$$

$$p_T^\varepsilon \leqslant \mathbb{P}\left(\Sigma_T \geqslant \frac{\varepsilon}{2\eta\sqrt{2}}\right) + \mathbb{P}\left(\frac{X_T}{T}\mathbf{1}_{X_T<1} \geqslant \eta^2\right) + \mathbb{P}\left(V_T \leqslant \eta\right), \tag{2.56}$$

$$q_T^\varepsilon \leqslant \mathbb{P}\left(L_T \mathbf{1}_{X_T \geqslant 1} \geqslant \frac{\varepsilon\eta}{2\sqrt{2}}\right) + \mathbb{P}\left(V_T \leqslant \eta\right), \tag{2.57}$$

and

$$Q_T^\varepsilon \leqslant \mathbb{P}\left(\frac{X_T}{T}\mathbf{1}_{X_T<1} \geqslant \frac{\varepsilon\eta}{2\sqrt{2}}\right) + \mathbb{P}\left(V_T \leqslant \eta\right). \tag{2.58}$$

First of all, using Theorem 2.3.1, we easily obtain that for all $c > -\frac{a}{b}$,

$$\lim_{T\to+\infty} \frac{1}{T} \log \mathbb{P}\left(S_T \geqslant c\right) = -I(c) \tag{2.59}$$

and for all $c > -\frac{b}{a-2}$,

$$\lim_{T\to+\infty} \frac{1}{T} \log \mathbb{P}\left(\Sigma_T \geqslant c\right) = -J(c) \tag{2.60}$$

where I and J are given in Theorem 2.3.1. Likewise, we deduce from Theorem 2.3.2 that for any $c > 0$ small enough

$$\lim_{T\to+\infty} \frac{1}{T} \log \mathbb{P}\left(V_T \leqslant c\right) = -K(c). \tag{2.61}$$

We now consider the parts involving L_T. For all $c > 0$ and $\lambda > 0$:

$$\begin{aligned}\mathbb{P}\left(L_T \mathbf{1}_{X_T \geqslant 1} \geqslant c\right) &= \mathbb{P}\left(\log X_T \mathbf{1}_{X_T \geqslant 1} \geqslant cT\right) \\ &\leqslant \mathbb{E}\left[e^{\lambda \log X_T}\right] e^{-\lambda c T}.\end{aligned}$$

Hence

$$\frac{1}{T}\log \mathbb{P}\left(L_T \mathbf{1}_{X_T \geqslant 1} \geqslant c\right) \leqslant -\lambda c + \frac{1}{T}\log\left(\mathbb{E}\left[X_T^{\lambda}\right]\right).$$

Asymptotic properties of the moments of the process X_T as T tends to infinity can be found in Proposition 3 of [14], and give that the second term tends to zero for T going to infinity. Thus, for any $\lambda > 0$ and $c > 0$, we have the following upper bound

$$\limsup_{T\to+\infty} \frac{1}{T}\log \mathbb{P}\left(L_T \mathbf{1}_{X_T \geqslant 1} \geqslant c\right) \leqslant -\lambda c.$$

Consequently, letting λ go to infinity, we obtain that for all $c > 0$,

$$\limsup_{T\to+\infty} \frac{1}{T}\log \mathbb{P}\left(L_T \mathbf{1}_{X_T \geqslant 1} \geqslant c\right) = -\infty. \tag{2.62}$$

Finally, we consider the terms involving $\frac{X_T}{T}$. For all $c > 0$ and $\lambda > 0$:

$$\begin{aligned}\mathbb{P}\left(\frac{X_T}{T}\mathbf{1}_{X_T < 1} \geqslant c\right) &\leqslant \mathbb{E}\left[e^{\lambda X_T \mathbf{1}_{X_T<1}}\right] e^{-\lambda c T} \\ &\leqslant e^{\lambda - \lambda c T}.\end{aligned}$$

Hence

$$\frac{1}{T}\log \mathbb{P}\left(\frac{X_T}{T}\mathbf{1}_{X_T < 1} \geqslant c\right) \leqslant -\lambda c + \frac{\lambda}{T}.$$

Thus, for any $\lambda > 0$ and $c > 0$, we have the following upper bound

$$\limsup_{T\to+\infty} \frac{1}{T}\log \mathbb{P}\left(\frac{X_T}{T}\mathbf{1}_{X_T < 1} \geqslant c\right) \leqslant -\lambda c.$$

Consequently, letting λ go to infinity, we obtain that for all $c > 0$,

$$\limsup_{T\to+\infty} \frac{1}{T}\log \mathbb{P}\left(\frac{X_T}{T}\mathbf{1}_{X_T < 1} \geqslant c\right) = -\infty. \tag{2.63}$$

Consequently, combining the limits (2.59) to (2.63) , we are able to compute the asymptotic behaviour of the bounds (2.55) to (2.58) and we show that for all $\varepsilon > 0$ and all $\eta > 0$ small enough,

$$\limsup_{T\to+\infty} \frac{1}{T}\log \mathbb{P}\left(\| \widehat{\theta}_T - \overline{\theta}_T \| > \varepsilon\right) \leqslant -M_{\varepsilon,\eta},$$

where $M_{\varepsilon,\eta} = \min\left\{I\left(\frac{\varepsilon}{2\eta\sqrt{2}}\right), J\left(\frac{\varepsilon}{2\eta\sqrt{2}}\right), K(\eta)\right\}$. Each term in this minimum tends to infinity as η goes to zero, so that $M_{\varepsilon,\eta}$ itself tends to infinity. This gives the announced result. □

Proof of Theorem 2.2.3. We have already shown in Lemma 2.6.2 that $I_{a,b}(2,0) = J_{a,b}(2,0)$ and $I_{a,b}(0,0) = K_{a,b}(0,0)$ and that, except at this two points, $I_{a,b}$ is infinite over $\mathbb{R}^- \times \mathbb{R}^-$ and over $[2,+\infty[\times \mathbb{R}^+$. We also know by (2.54) that

$$I_{a,b}(\alpha,\beta) \leqslant \min\left(J_{a,b}(\alpha,\beta); K_{a,b}(\alpha,\beta)\right),$$

so we still need to establish the other inequality over the remaining domain. In the sequel, we show that, for all compact subsets $C \subset \mathbb{R}^2$,

$$\limsup_{T\to+\infty} \frac{1}{T} \log \mathbb{P}\left(\overline{\theta}_T \in C\right) \leqslant -\inf_{(\alpha,\beta)\in C} \min\{J_{a,b}(\alpha,\beta), K_{a,b}(\alpha,\beta)\}. \tag{2.64}$$

It is sufficient to consider compact subsets of $\mathbb{R}^2$ instead of closed ones, as we already know that the sequence $\overline{\theta}_T$ satisfies an LDP with good rate function $I_{a,b}$ and $\mathbb{R}^2$ is locally compact so that the family $\left(\mathbb{P}\left(\overline{\theta}_T \in \bullet\right)\right)_T$ is exponentially tight (see Lemma 1.2.18 and Exercise 1.2.19 of [25]). This will prove the announced result as, by Lemma 2.6.3, the sequences $\left(\overline{\theta}_T\right)_T$ and $\left(\widehat{\theta}_T\right)_T$ share the same LDP.

First of all, we notice that $\overline{\theta}_T = g(\mathcal{Q}_T)$ where g is the function defined over the subset $\{(x,y,z,t) \in \mathbb{R}^4 | yz - 1 \neq 0\}$ by

$$g(x,y,z,t) = \left(\frac{y\left(2z - t^2\mathbf{1}_{t<0}\right) - x^2\mathbf{1}_{t\geqslant 0}}{yz-1}, \frac{t^2\mathbf{1}_{t<0} + (x^2\mathbf{1}_{t\geqslant 0} - 2)z}{yz-1}\right), \tag{2.65}$$

and the quadruplet $\mathcal{Q}_T$ satisfies an LDP with good rate function Λ^* given by Lemma 2.6.1. As the function g given by (2.65) is not continuous, we cannot apply directly the contraction principle. However,

$$\limsup_{T\to+\infty} \frac{1}{T} \log \mathbb{P}\left(\overline{\theta}_T \in C\right) \leqslant -\inf_{\overline{g^{-1}(C)}} \Lambda^*. \tag{2.66}$$

We need to describe the subset $\overline{g^{-1}(C)}$. A quadruplet (x,y,z,t) of $\mathbb{R}^4$ belongs to $\overline{g^{-1}(C)}$ if and only if there exists a sequence $(x_n,y_n,z_n,t_n)_n$ and a sequence $(\alpha_n,\beta_n)_n \in C$ such that, as n tends to infinity,

$$(x_n,y_n,z_n,t_n) \to (x,y,z,t) \tag{2.67}$$

and for all n

$$g\left(x_n,y_n,z_n,t_n\right) = (\alpha_n,\beta_n) \tag{2.68}$$

As C is a compact subset, up to a subsequence, there exists $(\alpha,\beta) \in C$ such that (α_n,β_n) converges to (α,β) as n goes to infinity. Moreover, (2.68) is equivalent to the following conditions for all n:

$$\beta_n y_n = x_n^2 \mathbf{1}_{t_n\geqslant 0} - \alpha_n$$

and

$$\left(2-\alpha_n\right) z_n = t_n^2 \mathbf{1}_{t_n<0} + \beta_n.$$

Up to a subsequence again, both indicator functions converge toward 1 or 0. Thus, letting n go to infinity, we obtain conditions on (x, y, z, t) which lead to

$$\overline{g^{-1}(C)} = \bigcup_{(\alpha,\beta)\in C} \mathcal{D}^+_{\alpha,\beta} \cup \mathcal{D}^-_{\alpha,\beta}$$

where

$$\mathcal{D}^+_{\alpha,\beta} = \left\{(x, y, z, t) \in \mathbb{R}^3 \times \mathbb{R}^+ | \beta y = x^2 - \alpha \text{ and } (2-\alpha) z = \beta\right\}$$

and

$$\mathcal{D}^-_{\alpha,\beta} = \left\{(x, y, z, t) \in \mathbb{R}^3 \times \mathbb{R}^- | \beta y = -\alpha \text{ and } (2-\alpha) z = t^2 + \beta\right\}.$$

Thus, (2.66) becomes

$$\limsup_{T\to+\infty} \frac{1}{T} \log \mathbb{P}\left(\overline{\theta}_T \in C\right) \leqslant - \inf_{(\alpha,\beta)\in C} \min\left\{\inf_{\mathcal{D}^+_{\alpha,\beta}} \Lambda^* ; \inf_{\mathcal{D}^-_{\alpha,\beta}} \Lambda^*\right\}. \tag{2.69}$$

For all $\alpha \neq 2$ and $\beta \neq 0$, we easily rewrite $\mathcal{D}^+_{\alpha,\beta} = \left\{\left(x, \frac{x^2-\alpha}{\beta}, \frac{\beta}{2-\alpha}, t\right), x \in \mathbb{R}, t \geqslant 0\right\}$ and $\mathcal{D}^-_{\alpha,\beta} = \left\{\left(x, -\frac{\alpha}{\beta}, \frac{t^2+\beta}{2-\alpha}, t\right), x \in \mathbb{R}, t \leqslant 0\right\}$.
As the rate function Λ^* given by Lemma 2.6.1 is infinite for $t > 0$, the infimum over $\mathcal{D}^+_{\alpha,\beta}$ reduces to the infimum over $\left\{\left(x, \frac{x^2-\alpha}{\beta}, \frac{\beta}{2-\alpha}, 0\right), x \in \mathbb{R}\right\}$ which is equal to $J_{a,b}(\alpha, \beta)$. Over $\mathcal{D}^-_{\alpha,\beta}$, we know by Lemma 2.6.1 that

$$\Lambda^*\left(x, -\frac{\alpha}{\beta}, \frac{t^2+\beta}{2-\alpha}, t\right) = \sup_{\mathcal{D}_{d,f}} \left\{\frac{1}{4}\left(x\sqrt{d-b} - t\sqrt{\varphi(f)}\right)^2 + \psi_{d,f}(t)\right\}$$

where $\psi_{d,f}(t)$ does not depend on x. Thus,

$$\Lambda^*\left(x, -\frac{\alpha}{\beta}, \frac{t^2+\beta}{2-\alpha}, t\right) \geqslant \Lambda^*\left(0, -\frac{\alpha}{\beta}, \frac{t^2+\beta}{2-\alpha}, t\right)$$

and the infimum over $\mathcal{D}^-_{\alpha,\beta}$ is greater than the infimum over $\left\{\left(0, -\frac{\alpha}{\beta}, \frac{t^2+\beta}{2-\alpha}, t\right), t \leqslant 0\right\}$ which is equal to $K_{a,b}(\alpha, \beta)$.
We now need to investigate the cases $\alpha = 2$ and $\beta = 0$ before concluding. For $\beta = 0$ and $\alpha \notin]0, 2[$, we already know the value of $I_{a,b}$. If $\alpha \in]0, 2[$,

$$\mathcal{D}^+_{\alpha,0} = \left\{(x, y, z, t) \in \mathbb{R}^3 \times \mathbb{R}^+ | x^2 = \alpha \text{ and } z = 0\right\} \text{ and } \mathcal{D}^-_{\alpha,0} = \varnothing.$$

And with the argument than before we obtain that for all $\alpha \in]0, 2[$,

$$I_{a,b}(\alpha, 0) = J_{a,b}(\alpha, 0) = \min\left\{J_{a,b}(\alpha, 0); K_{a,b}(\alpha, 0)\right\}$$

as $K_{a,b}(\alpha, 0)$ is equal to infinity. Now, for $\alpha = 2$, we have already computed $I_{a,b}$ for all $\beta \geqslant 0$. For $\beta < 0$,

$$\mathcal{D}^+_{2,\beta} = \varnothing \text{ and } \mathcal{D}^-_{2,\beta} = \left\{(x, y, z, t) \in \mathbb{R}^3 \times \mathbb{R}^- | \beta y = -2 \text{ and } t^2 = -\beta\right\}.$$

We obtain that

$$\mathcal{D}_{2,\beta}^{-}=\left\{\left(x,-\frac{2}{\beta},z,t\right),x\in\mathbb{R},z\in\mathbb{R},t\leqslant 0\right\}$$

and with the same argument than before, the infimum over this subset is greater than the infimum over $\left\{\left(0,-\frac{2}{\beta},z,t\right),z\in\mathbb{R},t\leqslant 0\right\}$, which is equal to $K_{a,b}\left(2,\beta\right)$. Thus, as $J_{a,b}\left(2,\beta\right)$ is equal to infinity, we can conclude that

$$I_{a,b}\left(2,\beta\right)=\min\left\{J_{a,b}\left(2,\beta\right);K_{a,b}\left(2,\beta\right)\right\}.$$

□

Appendix A: Proofs of the CLT for the two couples of simplified estimators

The key to obtain those results is Slutsky's lemma. Indeed, we have

$$\sqrt{T}\begin{pmatrix}\widetilde{a}_T-a\\ \widetilde{b}_T-b\end{pmatrix}=\sqrt{T}\begin{pmatrix}\widehat{a}_T-a\\ \widehat{b}_T-b\end{pmatrix}+\sqrt{T}\begin{pmatrix}-\frac{S_TL_T}{V_T}\\ \frac{L_T}{V_T}\end{pmatrix}\tag{A.1}$$

where

$$\sqrt{T}\begin{pmatrix}\widehat{a}_T-a\\ \widehat{b}_T-b\end{pmatrix}\xrightarrow{\mathcal{L}}\mathcal{N}(0,4C^{-1})\text{ with }C=\begin{pmatrix}\frac{-b}{a-2}&1\\ 1&-\frac{a}{b}\end{pmatrix},$$

and we show that the second summand on the right-hand side of (A.1) converges to zero in probability. Namely, it is well known (see for instance Lemma 3 of [65]) that S_T converges almost surely to $-\frac{a}{b}$ and V_T to $\frac{2}{a-2}$. And for all $\varepsilon>0$, we have

$$\begin{array}{rcl}\mathbb{P}\left(\left|\frac{\log X_T}{\sqrt{T}}\right|\geqslant\varepsilon\right)&=&\mathbb{P}\left(\left|\log X_T\right|\geqslant\sqrt{T}\varepsilon\right)\\ &\leqslant&\mathbb{P}\left(\log X_T\geqslant\sqrt{T}\varepsilon\right)+\mathbb{P}\left(-\log X_T\geqslant\sqrt{T}\varepsilon\right)\\ &\leqslant&\mathbb{E}\left(X_T\right)e^{-\sqrt{T}\varepsilon}+\mathbb{E}\left(X_T^{-1}\right)e^{-\sqrt{T}\varepsilon}\\ &\xrightarrow{T\to+\infty}&0\end{array}$$

because $\mathbb{E}\left(X_T\right)$ converges to $\mathbb{E}\left(X_\infty\right)=-\frac{a}{b}$ (see [65] Lemma 3) and, as the parameter a is supposed greater than 2, we obtain from Proposition 3 in [14] that

$$\mathbb{E}\left(X_T^{-1}\right)\to-\frac{b}{2}\frac{\Gamma(a/2-1)}{\Gamma(a/2)}=-\frac{b}{a-2}=\mathbb{E}\left(X_\infty^{-1}\right).$$

This gives the announced convergence in probability to zero. Thus, with Slutsky's lemma, the simplified estimators $(\widetilde{a}_T,\widetilde{b}_T)$ satisfy the same asymptotic normality result than the MLE.

Similarly, for the second couple of simplified estimators, we have

$$\sqrt{T}\begin{pmatrix}\breve{a}_T-a\\ \breve{b}_T-b\end{pmatrix}=\sqrt{T}\begin{pmatrix}\widehat{a}_T-a\\ \widehat{b}_T-b\end{pmatrix}+\sqrt{T}\begin{pmatrix}\frac{X_T}{TV_T}\\ -\frac{X_T\Sigma_T}{TV_T}\end{pmatrix}.\tag{A.2}$$

As Σ_T converges almost surely to $-\frac{b}{a-2}$ we only have to show that $X_T/\sqrt{T}$ converges to zero in probability. For all $\epsilon > 0$,

$$\mathbb{P}\left(\left|\frac{X_T}{\sqrt{T}}\right| \geqslant \varepsilon\right) = \mathbb{P}\left(X_T \geqslant \sqrt{T}\varepsilon\right) \leqslant \frac{\mathbb{E}\left(X_T\right)}{\sqrt{T}\varepsilon} \xrightarrow{T\to+\infty} 0$$

with the same argument than before. Thus $(\breve{a}_T, \breve{b}_T)$ also satisfies the same asymptotic normality result. $\square$

Appendix B: proof of Lemma 2.3.1

We apply the Gärtner-Ellis theorem (see [25]). It is easy to deduce from Proposition 2.4.1 that the pointwise limit $\widetilde{\Lambda}$ of the normalized cumulant generating function $\widetilde{\Lambda}_T$ of the couple (S_T, Σ_T) is given by

$$\widetilde{\Lambda}(\mu,\nu) = -\frac{d}{2}(1+f) - \frac{ab}{4} \tag{B.1}$$

where $d = \sqrt{b^2 - 8\mu}$ and $f = \sqrt{\left(\frac{a}{2}-1\right)^2 - 2\nu}$. We easily get from Lemma 2.4.1 that the function $\widetilde{\Lambda}$ is steep. To obtain the rate function I, we just have to compute the Fenchel-Legendre transform of $\widetilde{\Lambda}$:

$$I(x,y) = \sup_{\mu<\frac{b^2}{8},\,\nu<\frac{(a-2)^2}{8}} \left\{x\mu + y\nu - \widetilde{\Lambda}(\mu,\nu)\right\}. \tag{B.2}$$

First we note that

$$\text{for } x \leqslant 0 \text{ or } y \leqslant 0,\ \ I(x,y) = +\infty, \tag{B.3}$$

because $\widetilde{\Lambda}(\mu,\nu)$ tends to $-\infty$ as ν or μ tends to $-\infty$. Only the case $x > 0$ and $y > 0$ remains to be studied.
We look for the critical points. If $xy - 1 \neq 0$, we obtain

$$\begin{pmatrix} d_0 \\ f_0 \end{pmatrix} = \frac{1}{xy-1}\begin{pmatrix} 2y \\ 1 \end{pmatrix}. \tag{B.4}$$

As d_0 and f_0 must both be positive, the solution is in the domain if and only if $xy-1 > 0$. It is easy to check that this critical point corresponds to a maximum of $\widetilde{\Lambda}$, as the principal minors are respectively given by

$$-\frac{8(1+f_0)}{d_0^3} < 0 \qquad \text{and} \qquad \frac{4}{d_0^2 f_0^3} > 0.$$

Using the fact that $\mu = \frac{b^2-d^2}{8}$ and $\nu = \frac{(\frac{a}{2}-1)^2 - f^2}{2}$ and replacing it into (B.2), we get

$$I(x,y) = \frac{y}{2(xy-1)} + \frac{b^2}{8}x + \frac{(a-2)^2}{8}y + \frac{ab}{4}. \tag{B.5}$$

To conclude, we need to examine the case $x > 0$, $y > 0$ and $xy - 1 < 0$. We already know that for ν or μ tending to $-\infty$, $-\widetilde{\Lambda}(\mu, \nu)$ tends to $+\infty$. However, as x and y are positive, we cannot conclude directly. But it is possible to find a direction in which $\widetilde{\Lambda}$ dominate the expression. Note that, for $-\nu$ and $-\mu$ large enough,

$$x\mu + y\nu - \widetilde{\Lambda}(\mu, \nu) \sim x\mu + y\nu + 2\sqrt{\mu\nu}.$$

Let $k > 0$ and $\nu = k\mu$. We just have to find a $k > 0$ that satisfies

$$k^2 y^2 + 2k(xy - 2) + x^2 < 0.$$

The above polynomial has two positive roots. Any k between this two roots will satisfy the condtion. For example, $k = (2 - xy)/y^2$ fits. We have found directions for which $x\mu + y\nu - \widetilde{\Lambda}(\mu, \nu)$ tends to $+\infty$, so that the supremum itself is equal to $+\infty$. And,

$$\text{for } x > 0, y > 0 \text{ such that } xy - 1 < 0, \quad I(x, y) = +\infty. \tag{B.6}$$

Combining (B.3), (B.5) and (B.6), we obtain the announced result. □

Appendix C: Proofs of Lemmas 2.4.2 to 2.4.5

The four following proofs rely on lower and upper bounds for the modified Bessel function of the first kind, given by formula (6.25) of [63]. More precisely, for all $z > 0$ and $\nu > -\frac{1}{2}$, we have:

$$1 < \left(\frac{2}{z}\right)^{\nu} \Gamma\left(\nu + 1\right) I_{\nu}\left(z\right) < e^{z}. \tag{C.1}$$

C.1. Proof of Lemma 2.4.2

It is easy to deduce the following upper and lower bounds from (C.1):

$$\frac{(\beta_T\sqrt{y})^f}{2^f\,\Gamma(f+1)} \leqslant I_f\left(\beta_T\sqrt{y}\right) \leqslant \frac{(\beta_T\sqrt{y})^f}{2^f\Gamma(f+1)} e^{\beta_T\sqrt{y}}. \tag{C.2}$$

Replacing it into the expression of H_T given by (2.13) leads to:

$$\frac{2^f\Gamma(1+f)}{\beta_T^f} H_T \geqslant \int_0^1 e^{\lambda\sqrt{Ty} - \gamma\sqrt{-T\log y} - \alpha_T y}\, y^{\frac{2f+a-2}{4}}\, \mathrm{d}y \tag{C.3}$$

and

$$\frac{2^f\Gamma(1+f)}{\beta_T^f} H_T \leqslant \int_0^1 e^{\beta_T\sqrt{y} + \lambda\sqrt{Ty} - \gamma\sqrt{-T\log y} - \alpha_T y}\, y^{\frac{2f+a-2}{4}}\, \mathrm{d}y. \tag{C.4}$$

We consider separately the integrals over $[0, \frac{1}{T}]$ and over $[\frac{1}{T}, 1]$. On the one hand,

$$\begin{aligned}
\int_{\frac{1}{T}}^{1} e^{\beta_T\sqrt{y}+\lambda\sqrt{Ty}-\gamma\sqrt{-T\log y}-\alpha_T y}\, y^{\frac{2f+a-2}{4}}\,\mathrm{d}y &\leqslant \int_{\frac{1}{T}}^{1} e^{\beta_T\sqrt{y}+\lambda\sqrt{Ty}-\gamma\sqrt{-T\log y}}\,\mathrm{d}y \\
&\leqslant 1 \times \sup_{[\frac{1}{T},1]}\left(e^{\beta_T\sqrt{y}+\lambda\sqrt{Ty}-\gamma\sqrt{-T\log y}}\right) \\
&\leqslant e^{\lambda\sqrt{T}+\beta_T-\gamma\sqrt{-T\log\frac{1}{T}}} \\
&= O\left(e^{(\lambda-\gamma)\sqrt{T}\sqrt{\log T}}\right).
\end{aligned}$$

On the other hand, with the same argument, we have that

$$\int_{\frac{1}{T}}^{1} e^{\lambda\sqrt{Ty}-\gamma\sqrt{-T\log y}-\alpha_T y}\, y^{\frac{2f+a-2}{4}}\,\mathrm{d}y = O\left(e^{(\lambda-\gamma)\sqrt{T}\sqrt{\log T}}\right).$$

Over $[0, \frac{1}{T}]$, $e^{\lambda\sqrt{Ty}}$, $e^{\beta_T\sqrt{y}}$ and $e^{-\alpha_T y}$ are bounded. So, as α_T and β_T are always positive, we have the following bounds

$$\int_{0}^{\frac{1}{T}} e^{\lambda\sqrt{Ty}-\gamma\sqrt{-T\log y}-\alpha_T y}\, y^{\frac{2f+a-2}{4}}\,\mathrm{d}y \geqslant e^{-\alpha_T/T-|\lambda|}\int_{0}^{\frac{1}{T}} e^{-\gamma\sqrt{-T\log y}}\, y^{\frac{2f+a-2}{4}}\,\mathrm{d}y,$$

$$\int_{0}^{\frac{1}{T}} e^{\beta_T\sqrt{y}+\lambda\sqrt{Ty}-\gamma\sqrt{-T\log y}-\alpha_T y}\, y^{\frac{2f+a-2}{4}}\,\mathrm{d}y \leqslant e^{|\lambda|+\alpha_T/T+\beta_T/\sqrt{T}}\int_{0}^{\frac{1}{T}} e^{-\gamma\sqrt{-T\log y}}\, y^{\frac{2f+a-2}{4}}\,\mathrm{d}y.$$

Using the change of variable given by $z = \sqrt{-\log y} + \frac{\gamma\sqrt{T}}{2g}$, where $g = \frac{2f+a+2}{4}$, we obtain:

$$\begin{aligned}
\int_{0}^{\frac{1}{T}} e^{-\gamma\sqrt{-T\log y}}\, y^{g-1}\,\mathrm{d}y &= 2e^{\frac{\gamma^2 T}{4g}}\int_{\sqrt{\log T}+\frac{\gamma\sqrt{T}}{2g}}^{+\infty} e^{-g\,z^2}\left(z-\frac{\gamma\sqrt{T}}{2g}\right)\mathrm{d}z \\
&= 2e^{\frac{\gamma^2 T}{4g}}\left(\int_{\sqrt{\log T}+\frac{\gamma\sqrt{T}}{2g}}^{+\infty} e^{-g\,z^2}\,z\,\mathrm{d}z - \frac{\gamma\sqrt{T}}{2g}\int_{\sqrt{\log T}+\frac{\gamma\sqrt{T}}{2g}}^{+\infty} e^{-g\,z^2}\,\mathrm{d}z\right).
\end{aligned}$$

By dominated convergence as $\gamma < 0$, the first integral tends to zero when T goes to infinity and the second one tends to the positive constant $\sqrt{\frac{\pi}{g}}$. This leads to the following bounds, for T large enough:

$$\int_{0}^{\frac{1}{T}} e^{-\gamma\sqrt{-T\log y}}\, y^{g-1}\,\mathrm{d}y \leqslant 2\sqrt{\frac{\pi}{g}}\,\frac{|\gamma|\sqrt{T}}{g}\,\exp\left(\frac{\gamma^2 T}{4g}\right)$$

$$\int_{0}^{\frac{1}{T}} e^{-\gamma\sqrt{-T\log y}}\, y^{g-1}\,\mathrm{d}y \geqslant \frac{1}{2}\sqrt{\frac{\pi}{g}}\,\frac{|\gamma|\sqrt{T}}{g}\,\exp\left(\frac{\gamma^2 T}{4g}\right).$$

Combined with the result over $[\frac{1}{T}, 1]$, it gives the announced result.

C.2. Proof of Lemma 2.4.3

The upper bound easily follows from (C.2). Actually, for all $\gamma > 0$ and $\lambda \in \mathbb{R}$, we obtain

$$\begin{aligned}
H_T &\leqslant \frac{\beta_T^f}{\Gamma(f+1)2^f}\int_0^1 e^{\left(\lambda\sqrt{T}+\beta_T\right)\sqrt{y}-\gamma\sqrt{-T\log y}}\, y^{\frac{2f+a-2}{4}}\,\mathrm{d}y \\
&\leqslant \frac{\beta_T^f}{\Gamma(f+1)2^f}\, e^{|\lambda|\sqrt{T}+\beta_T}.
\end{aligned} \tag{C.5}$$

Besides, using the lower bound of (C.2), we clearly have, for all $\gamma > 0$ and $\lambda \in \mathbb{R}$,

$$H_T \geqslant \frac{\beta_T^f \, e^{-\alpha_T}}{\Gamma(f+1)2^f} \int_0^1 e^{\lambda\sqrt{T}\sqrt{y}-\gamma\sqrt{-T\log y}} \, y^{\frac{2f+a-2}{4}} \, \mathrm{d}y. \tag{C.6}$$

To obtain the announced lower bound, we need to consider separately the integral over $[0, \frac{1}{T}]$ and over $[\frac{1}{T}, 1]$. On the one hand, the integral over $[\frac{1}{T}, 1]$ is easy to handle.

$$\int_{\frac{1}{T}}^1 e^{\lambda\sqrt{Ty}-\gamma\sqrt{-T\log y}} \, y^{\frac{2f+a-2}{4}} \, \mathrm{d}y \geqslant \varepsilon \, e^{-\gamma\sqrt{T}\sqrt{\log T}} \, T^{-\frac{2f+a-2}{4}}, \tag{C.7}$$

where $\varepsilon = \exp\left(\lambda \mathbf{1}_{\lambda \geqslant 0} + \lambda\sqrt{T}\mathbf{1}_{\lambda<0}\right)$. On the other hand,

$$\int_0^{\frac{1}{T}} e^{\lambda\sqrt{T}\sqrt{y}-\gamma\sqrt{-T\log y}} \, y^{\frac{2f+a-2}{4}} \, \mathrm{d}y \geqslant e^{-|\lambda|} \int_0^{\frac{1}{T}} e^{-\gamma\sqrt{-T\log y}} \, y^{\frac{2f+a-2}{4}} \, \mathrm{d}y$$

Using the variable change $z = \sqrt{-\log y} + \frac{\gamma\sqrt{T}}{2g}$, where we recall that $g = \frac{2f+a+2}{4}$, we obtain

$$\int_0^{\frac{1}{T}} e^{-\gamma\sqrt{-T\log y}} \, y^{\frac{2f+a-2}{4}} \, \mathrm{d}y = 2 \, e^{\frac{\gamma^2 T}{4g}} \int_{\sqrt{\log T}+\frac{\gamma\sqrt{T}}{2g}}^{+\infty} e^{-g\,z^2} \left(z - \frac{\gamma\sqrt{T}}{2g}\right) \mathrm{d}z.$$

Firstly, we clearly have

$$\int_{\sqrt{\log T}+\frac{\gamma\sqrt{T}}{2g}}^{+\infty} e^{-g\,z^2} \, z \, \mathrm{d}z = \frac{1}{2g} \exp\left(-g\left(\sqrt{\log T} + \frac{\gamma\sqrt{T}}{2g}\right)^2\right).$$

Besides, for the second part of the integral, we have

$$\begin{aligned}
\int_{\sqrt{\log T}+\frac{\gamma\sqrt{T}}{2g}}^{+\infty} e^{-g\,z^2} \, \mathrm{d}z &= \int_{\sqrt{\log T}+\frac{\gamma\sqrt{T}}{2g}}^{+\infty} e^{-g\,z^2} \, z \times \frac{1}{z} \mathrm{d}z \\
&\leqslant \frac{1}{\sqrt{\log T}+\frac{\gamma\sqrt{T}}{2g}} \int_{\sqrt{\log T}+\frac{\gamma\sqrt{T}}{2g}}^{+\infty} e^{-g\,z^2} \, z \mathrm{d}z \\
&= \frac{1}{2g\sqrt{\log T}+\gamma\sqrt{T}} \exp\left(-g\left(\sqrt{\log T} + \frac{\gamma\sqrt{T}}{2g}\right)^2\right).
\end{aligned}$$

Thus, for any positive γ, we have the following lower bound:

$$\int_0^{\frac{1}{T}} e^{-\gamma\sqrt{-T\log y}} \, y^{\frac{2f+a-2}{4}} \, \mathrm{d}y \geqslant \frac{1}{g}\left(1 - \frac{\gamma\sqrt{T}}{2g\sqrt{\log T}+\gamma\sqrt{T}}\right) \exp\left(-g\log T - \gamma\sqrt{T}\sqrt{\log T}\right).$$

Combined with (C.7), this leads to

$$H_T \geqslant \frac{(\beta_T)^f \; e^{-\alpha_T}}{2^f \, \Gamma(1+f)} \left(\varepsilon + \frac{e^{-|\lambda|}}{g}\left(1 - \frac{\gamma\sqrt{T}}{2g\sqrt{\log T}+\gamma\sqrt{T}}\right)\right) \exp\left(-\gamma\sqrt{T}\sqrt{\log T} - g\,\log T\right).$$

C.3. Proof of Lemma 2.4.4

Using the inequality (C.2) for the modified Bessel function I_f in (2.14), we obtain

$$K_T \leqslant \frac{(\beta_T)^f}{2^f\Gamma(1+f)} \int_1^{+\infty} e^{\beta_T\sqrt{y}+\lambda\sqrt{Ty}-\alpha_T y}\, y^{\gamma+\frac{2f+a-2}{4}}\, \mathrm{d}y \tag{C.8}$$

and

$$K_T \geqslant \frac{(\beta_T)^f}{2^f\Gamma(1+f)} \int_1^{+\infty} e^{\lambda\sqrt{Ty}-\alpha_T y}\, y^{\gamma+\frac{2f+a-2}{4}}\, \mathrm{d}y. \tag{C.9}$$

To go further, we need to consider the sign of the exponent $\gamma + \frac{2f+a-2}{4}$.

• If $\gamma + \frac{2f+a-2}{4} \leqslant 0$: using the fact that $y^{\gamma+\frac{2f+a-2}{4}} \leqslant 1$ and with the change of variable $u = \sqrt{y} - \frac{\lambda\sqrt{T}+\beta_T}{2\alpha_T}$, we obtain the following asymptotic behaviour for the upper bound of K_T,

$$\begin{aligned}
K_T &\leqslant \frac{(\beta_T)^f}{2^f\Gamma(1+f)} \int_1^{+\infty} e^{\left(\beta_T+\lambda\sqrt{T}\right)\sqrt{y}-\alpha_T y}\, \mathrm{d}y \\
&= \tfrac{2(\beta_T)^f}{2^f\Gamma(1+f)} e^{\frac{(\lambda\sqrt{T}+\beta_T)^2}{4\alpha_T}} \int_{1-\frac{\lambda\sqrt{T}+\beta_T}{2\alpha_T}}^{+\infty} e^{-\alpha_T u^2} \left(u + \frac{\lambda\sqrt{T}+\beta_T}{2\alpha_T}\right) \mathrm{d}u \\
&\leqslant \tfrac{2(\beta_T)^f}{2^f\Gamma(1+f)} e^{\frac{(\lambda\sqrt{T}+\beta_T)^2}{4\alpha_T}} \left(A_1 + \tfrac{\beta_T}{2\alpha_T} A_2\right)
\end{aligned}$$

where A_1 and A_2 are given by:

$$A_1 = \int_{1-\frac{\lambda\sqrt{T}+\beta_T}{2\alpha_T}}^{+\infty} e^{-\alpha_T u^2}\, u\, \mathrm{d}u = \frac{1}{2\alpha_T} e^{-\alpha_T\left(1-\frac{\lambda\sqrt{T}+\beta_T}{2\alpha_T}\right)^2}, \tag{C.10}$$

$$A_2 = \int_{1-\frac{\lambda\sqrt{T}+\beta_T}{2\alpha_T}}^{+\infty} e^{-\alpha_T u^2}\, \mathrm{d}u = \int_{1-\frac{\lambda\sqrt{T}+\beta_T}{2\alpha_T}}^{+\infty} e^{-\alpha_T u^2}\, u \times \frac{1}{u}\, \mathrm{d}u \leqslant \frac{1}{1-\frac{\lambda\sqrt{T}+\beta_T}{2\alpha_T}} A_1. \tag{C.11}$$

Thus:

$$\begin{aligned}
K_T &\leqslant \tfrac{2(\beta_T)^f}{2^f\Gamma(1+f)} e^{\frac{(\lambda\sqrt{T}+\beta_T)^2}{4\alpha_T}} \tfrac{1}{2\alpha_T} e^{-\alpha_T\left(1-\frac{\lambda\sqrt{T}+\beta_T}{2\alpha_T}\right)^2} \left(1 + \frac{\frac{\beta_T}{2\alpha_T}}{1-\frac{\lambda\sqrt{T}+\beta_T}{2\alpha_T}}\right) \\
&= \tfrac{(\beta_T)^f}{2^f\Gamma(1+f)\alpha_T} \left(1 + \frac{\frac{\beta_T}{2\alpha_T}}{1-\frac{\lambda\sqrt{T}+\beta_T}{2\alpha_T}}\right) \exp\left(\lambda\sqrt{T} + \beta_T - \alpha_T\right).
\end{aligned} \tag{C.12}$$

• If $\gamma + \frac{2f+a-2}{4} > 0$: with formula 3.462(1) in [44], we get

$$\begin{aligned}
K_T &\leqslant \tfrac{(\beta_T)^f}{2^f\Gamma(1+f)} \Bigg(2\left(2\alpha_T\right)^{-(\gamma+g)} \Gamma\left(2\gamma+2g\right) e^{\frac{(\lambda\sqrt{T}+\beta_T)^2}{8\alpha_T}} D_{-2(\gamma+g)}\left(\tfrac{-\beta_T-\lambda\sqrt{T}}{\sqrt{2\alpha_T}}\right) \\
&\qquad - \int_0^1 e^{\beta_T\sqrt{y}+\lambda\sqrt{Ty}-\alpha_T y}\, y^{\gamma+g-1}\, \mathrm{d}y\Bigg) \\
&\leqslant \tfrac{2(\beta_T)^f (2\alpha_T)^{-(\gamma+g)}\Gamma(2\gamma+2g)}{2^f\Gamma(1+f)} e^{\frac{(\lambda\sqrt{T}+\beta_T)^2}{8\alpha_T}} D_{-2(\gamma+g)}\left(\tfrac{-\beta_T-\lambda\sqrt{T}}{\sqrt{2\alpha_T}}\right),
\end{aligned}$$

where $D_{-2(\gamma+g)}$ is the parabolic cylinder function defined by 9.250 in [44]. But, if $\lambda < 0$, for T large enough, $-\beta_T - \lambda\sqrt{T} > 0$ so formula 9.246(1) of [44] gives

$$D_{-2(\gamma+g)}\left(\frac{-\beta_T - \lambda\sqrt{T}}{\sqrt{2\alpha_T}}\right) \sim \left(\frac{-\beta_T - \lambda\sqrt{T}}{\sqrt{2\alpha_T}}\right)^{-2(\gamma+g)} e^{-\frac{(\lambda\sqrt{T}+\beta_T)^2}{8\alpha_T}}.$$

Thus $K_T = O(\beta_T^f)$ as T goes to infinity. If $\lambda = 0$, we use formula 9.246(2) in [44].

$$D_{-2(\gamma+g)}\left(\frac{-\beta_T}{\sqrt{2\alpha_T}}\right) \sim \frac{\sqrt{2\pi}}{\Gamma\left(2\gamma+2g\right)}\left(\frac{\beta_T}{\sqrt{2\alpha_T}}\right)^{2\gamma+2g-1} e^{\frac{(\beta_T)^2}{8\alpha_T}}.$$

This leads to the same conclusion: as T goes to infinity,

$$K_T = O\left((\beta_T)^f\right). \tag{C.13}$$

Otherwise, for the proof of Lemma 2.4.1, we also need a lower bound for K_T when $\gamma \geqslant 0$. We note that over $[1, +\infty[$, $y^{\gamma+g-1} \geqslant 1$, so that

$$K_T \geqslant \frac{2^{-f}\,(\beta_T)^f}{\Gamma(1+f)} \int_1^{+\infty} e^{\lambda\sqrt{T}\sqrt{y}-\alpha_T y}\,\mathrm{d}y. \tag{C.14}$$

With the change of variable given by $u = \sqrt{y} - \frac{\lambda\sqrt{T}}{2\alpha_T}$, it becomes

$$\begin{aligned} K_T &\geqslant \tfrac{2^{1-f}(\beta_T)^f}{\Gamma(1+f)} e^{\frac{\lambda^2 T}{4\alpha_T}} \int_{1-\frac{\lambda\sqrt{T}}{2\alpha_T}}^{+\infty} e^{-\alpha_T u^2}\left(u + \frac{\lambda\sqrt{T}}{2\alpha_T}\right)\mathrm{d}u \\ &\geqslant \tfrac{2^{1-f}(\beta_T)^f}{\Gamma(1+f)} e^{\frac{\lambda^2 T}{4\alpha_T}} \left(\int_{1-\frac{\lambda\sqrt{T}}{2\alpha_T}}^{+\infty} e^{-\alpha_T u^2} u\,\mathrm{d}u + \frac{\lambda\sqrt{T}}{2\alpha_T}\int_{1-\frac{\lambda\sqrt{T}}{2\alpha_T}}^{+\infty} e^{-\alpha_T u^2}\,\mathrm{d}u\right) \end{aligned} \tag{C.15}$$

However, the first integral is easily computable:

$$\int_{1-\frac{\lambda\sqrt{T}}{2\alpha_T}}^{+\infty} e^{-\alpha_T u^2} u\,\mathrm{d}u = \frac{1}{2\alpha_T}\exp\left(-\alpha_T\left(1 - \frac{\lambda\sqrt{T}}{2\alpha_T}\right)^2\right) \tag{C.16}$$

and for the second one, we clearly have the following upper bound,

$$\begin{aligned} \textstyle\int_{1-\frac{\lambda\sqrt{T}}{2\alpha_T}}^{+\infty} e^{-\alpha_T u^2}\,\mathrm{d}u &= \textstyle\int_{1-\frac{\lambda\sqrt{T}}{2\alpha_T}}^{+\infty} e^{-\alpha_T u^2} u \times \frac{1}{u}\,\mathrm{d}u \\ &\leqslant \tfrac{1}{1-\frac{\lambda\sqrt{T}}{2\alpha_T}} \textstyle\int_{1-\frac{\lambda\sqrt{T}}{2\alpha_T}}^{+\infty} e^{-\alpha_T u^2} u\,\mathrm{d}u \\ &\leqslant \tfrac{1}{1-\frac{\lambda\sqrt{T}}{2\alpha_T}}\tfrac{1}{2\alpha_T}\exp\left(-\alpha_T\left(1-\tfrac{\lambda\sqrt{T}}{2\alpha_T}\right)^2\right). \end{aligned} \tag{C.17}$$

Using the fact that $\lambda < 0$ and combining (C.15), (C.16) and (C.17), we show the announced result:

$$K_T \geqslant \frac{2^{1-f}\,(\beta_T)^f}{\Gamma(1+f)}\frac{e^{-\alpha_T}}{2\alpha_T}\left(1 + \frac{\frac{\lambda\sqrt{T}}{2\alpha_T}}{1-\frac{\lambda\sqrt{T}}{2\alpha_T}}\right)\exp\left(\lambda\sqrt{T}\right). \tag{C.18}$$

All those results are still true for $\lambda = 0$.

C.4. Proof of Lemma 2.4.5

As previously, we use (C.2) to find lower and upper bounds for K_T as T goes to infinity, for all $\lambda > 0$. We consider two cases depending on the sign of the exponent $\gamma + g - 1$.

• If $\gamma + g - 1 \leqslant 0$: As in the proof of Lemma 2.4.4, we have the following upper bound

$$K_T \leqslant \frac{2\,(\beta_T)^f}{2^f\Gamma(1+f)} e^{\frac{(\lambda\sqrt{T}+\beta_T)^2}{4\alpha_T}} \left(A_1 + \frac{\beta_T}{2\alpha_T} A_2\right) \tag{C.19}$$

where $A_1 = \displaystyle\int_{1-\frac{\lambda\sqrt{T}+\beta_T}{2\alpha_T}}^{+\infty} e^{-\alpha_T u^2}\, u\, \mathrm{d}u$ tends to zero for all $\lambda > 0$ by dominated convergence, and $A_2 = \displaystyle\int_{1-\frac{\lambda\sqrt{T}+\beta_T}{2\alpha_T}}^{+\infty} e^{-\alpha_T u^2}\, \mathrm{d}u$ tends to the positive constant $2\sqrt{\frac{\pi}{d-b}}$. Thus, for T large enough,

$$K_T \leqslant 2^{1-f} \sqrt{\frac{\pi}{d-b}}\, \frac{(\beta_T)^{1+f}}{\alpha_T\,\Gamma(1+f)}\, e^{\frac{(\lambda\sqrt{T}+\beta_T)^2}{4\alpha_T}}. \tag{C.20}$$

For the lower bound, with the change of variable given by $u = \sqrt{y} - \frac{\lambda\sqrt{T}}{2\alpha_T}$, we obtain

$$\begin{aligned}
K_T &\geqslant \tfrac{2^{-f}(\beta_T)^f}{\Gamma(1+f)} \int_1^{+\infty} e^{\lambda\sqrt{T}\sqrt{y}-\alpha_T y}\, y^{\gamma+g-1}\, \mathrm{d}y \\
&= \tfrac{2^{1-f}(\beta_T)^f}{\Gamma(1+f)} e^{\frac{\lambda^2 T}{4\alpha_T}} \int_{1-\frac{\lambda\sqrt{T}}{2\alpha_T}}^{+\infty} e^{-\alpha_T u^2} \left(u + \frac{\lambda\sqrt{T}}{2\alpha_T}\right)^{2(\gamma+g-1)} \left(u + \frac{\lambda\sqrt{T}}{2\alpha_T}\right) \mathrm{d}u \\
&\geqslant \tfrac{2^{1-f}(\beta_T)^f}{\Gamma(1+f)} e^{\frac{\lambda^2 T}{4\alpha_T}} \int_{1-\frac{\lambda\sqrt{T}}{2\alpha_T}}^{+\infty} e^{-\alpha_T u^2} \left(2\max\left\{u; \frac{\lambda\sqrt{T}}{2\alpha_T}\right\}\right)^{2(\gamma+g-1)} \mathrm{d}u \\
&\geqslant \tfrac{2^{1-f}(\beta_T)^f}{\Gamma(1+f)} e^{\frac{\lambda^2 T}{4\alpha_T}} \left(\left(\tfrac{2\lambda\sqrt{T}}{2\alpha_T}\right)^{2(\gamma+g-1)} \int_{1-\frac{\lambda\sqrt{T}}{2\alpha_T}}^{\frac{\lambda\sqrt{T}}{2\alpha_T}} e^{-\alpha_T u^2}\, \mathrm{d}u + \int_{\frac{\lambda\sqrt{T}}{2\alpha_T}}^{+\infty} e^{-\alpha_T u^2}\, (2u)^{2(\gamma+g-1)}\, \mathrm{d}u \right) \\
&\geqslant \tfrac{2^{1-f}(\beta_T)^f}{\Gamma(1+f)} e^{\frac{\lambda^2 T}{4\alpha_T}} \left(\tfrac{\lambda\sqrt{T}}{\alpha_T}\right)^{2(\gamma+g-1)} \int_{1-\frac{\lambda\sqrt{T}}{2\alpha_T}}^{\frac{\lambda\sqrt{T}}{2\alpha_T}} e^{-\alpha_T u^2}\, \mathrm{d}u.
\end{aligned}$$

By dominated convergence, the last integral converges as T tends to infinity to the positive constant $2\sqrt{\frac{\pi}{d-b}}$. For T large enough, this leads to

$$K_T \geqslant 2^{1-f} \sqrt{\frac{\pi}{d-b}} \frac{(\beta_T)^f}{\Gamma(1+f)} \left(\frac{\lambda\sqrt{T}}{\alpha_T}\right)^{2(\gamma+g-1)} e^{\frac{\lambda^2 T}{4\alpha_T}}. \tag{C.21}$$

• If $\gamma + g - 1 > 0$: over $[1, +\infty[$, we notice that $y^{\gamma+g-1} \geqslant 1$. Thus

$$K_T \geqslant \frac{2^{-f}\,(\beta_T)^f}{\Gamma(1+f)} \int_1^{+\infty} e^{\lambda\sqrt{T}\sqrt{y}-\alpha_T y}\, \mathrm{d}y. \tag{C.22}$$

With the change of variable given by $u = \sqrt{y} - \frac{\lambda\sqrt{T}}{2\alpha_T}$, it becomes

$$\begin{aligned} K_T &\geqslant \tfrac{2^{1-f}(\beta_T)^f}{\Gamma(1+f)} e^{\frac{\lambda^2 T}{4\alpha_T}} \int_{1-\frac{\lambda\sqrt{T}}{2\alpha_T}}^{+\infty} e^{-\alpha_T u^2}(u + \tfrac{\lambda\sqrt{T}}{2\alpha_T})\,\mathrm{d}u \\ &\geqslant \tfrac{2^{1-f}(\beta_T)^f}{\Gamma(1+f)} e^{\frac{\lambda^2 T}{4\alpha_T}} \left(\int_{1-\frac{\lambda\sqrt{T}}{2\alpha_T}}^{+\infty} e^{-\alpha_T u^2} u\,\mathrm{d}u + \tfrac{\lambda\sqrt{T}}{2\alpha_T} \int_{1-\frac{\lambda\sqrt{T}}{2\alpha_T}}^{+\infty} e^{-\alpha_T u^2}\,\mathrm{d}u \right) \end{aligned}$$

By dominated convergence, the first integral in the last expression tends to zero as T goes to infinity and the second integral converges to the positive constant $2\sqrt{\frac{\pi}{d-b}}$. It gives, for T large enough,

$$K_T \geqslant \sqrt{\frac{\pi}{d-b}} \frac{2^{1-f}\,(\beta_T)^f}{\Gamma(1+f)} \frac{\lambda\sqrt{T}}{2\alpha_T} e^{\frac{\lambda^2 T}{4\alpha_T}}. \tag{C.23}$$

Moreover, we establish the upper bound with formula 3.462(1) of [44] as in the previous proof.

$$\begin{aligned} K_T &\leqslant \tfrac{(\beta_T)^f}{2^f\Gamma(1+f)} \Bigg(2\,(2\alpha_T)^{-(\gamma+g)}\,\Gamma\,(2\gamma+2g)\,e^{\frac{(\lambda\sqrt{T}+\beta_T)^2}{8\alpha_T}} D_{-2(\gamma+g)}\left(\tfrac{-\beta_T-\lambda\sqrt{T}}{\sqrt{2\alpha_T}}\right) \\ &\qquad - \int_0^1 e^{\beta_T\sqrt{y}+\lambda\sqrt{T y}-\alpha_T y}\, y^{\gamma+g-1}\,\mathrm{d}y \Bigg) \\ &\leqslant 2^{1-f}\,(\beta_T)^f\,(2\alpha_T)^{-(\gamma+g)} \tfrac{\Gamma(2\gamma+2g)}{\Gamma(1+f)}\, e^{\frac{(\lambda\sqrt{T}+\beta_T)^2}{8\alpha_T}} D_{-2(\gamma+g)}\left(\tfrac{-\beta_T-\lambda\sqrt{T}}{\sqrt{2\alpha_T}}\right). \end{aligned}$$

But $-\beta_T - \lambda\sqrt{T} < 0$ so, by formula 9.246(2) in [44]

$$D_{-2(\gamma+g)}\left(\frac{-\beta_T-\lambda\sqrt{T}}{\sqrt{2\alpha_T}}\right) \sim \frac{\sqrt{2\pi}}{\Gamma\,(2\gamma+2g)} \left(\frac{\beta_T+\lambda\sqrt{T}}{\sqrt{2\alpha_T}}\right)^{2\gamma+2g-1} e^{\frac{(\lambda\sqrt{T}+\beta_T)^2}{8\alpha_T}}.$$

This gives, for T large enough,

$$K_T \leqslant \frac{2^{2-f}\sqrt{2\pi}\,(\beta_T)^f}{\Gamma(1+f)} \left(\beta_T + \lambda\sqrt{T}\right)^{2\gamma+2g-1} e^{\frac{(\lambda\sqrt{T}+\beta_T)^2}{4\alpha_T}}. \tag{C.24}$$

This easily leads to the announced result. □

Appendix D: Proof of Lemma 2.5.1

We call L the pointwise limit of the cumulant generating function of the triplet $(X_T/T, S_T, \Sigma_T)$. We notice that $L(\lambda,\mu,\nu) = \Lambda(\lambda,\mu,\nu,0)$ where Λ is given by Proposition 2.4.1. We easily deduce from Lemma 2.4.1 that the function L is steep. We apply the Gärtner-Ellis theorem: the rate function is given by the Fenchel-Legendre transform of L on its effective domain.

$$I(x,y,z) = \sup_{\lambda\in\mathbb{R},\,\mu<\frac{b^2}{8},\,\nu<\frac{(a-2)^2}{8}} \left\{x\lambda + y\mu + z\nu - L(\lambda,\mu,\nu)\right\}. \tag{D.1}$$

For y and z we recognize the exact same term than in Appendix B. So with the same argument, we show that

$$\text{for } y \leqslant 0, z \leqslant 0 \text{ or } yz-1 \leqslant 0,\ \ I(x,y,z) = +\infty. \tag{D.2}$$

Besides, if $x < 0$ then for λ tending to $-\infty$, $\lambda x \to +\infty$ and $I(x,y,z) = +\infty$ because Λ does not depend on λ for $\lambda < 0$. Moreover, for $x > 0$, the term on λ, namely $x\lambda - \frac{\lambda^2}{d-b}\mathbb{1}_{\lambda>0}$, is always negative for λ negative and sometimes positive if λ is positive. So the supremum is necessarily reached for $\lambda > 0$. We finally have to calculate:

$$I(x,y,z) = \sup_{\lambda>0,\, \mu<\frac{b^2}{8},\, \nu<\frac{(a-2)^2}{8}} \{x\lambda + y\mu + z\nu - L(\lambda,\mu,\nu)\} \tag{D.3}$$

with $x \geqslant 0$, $y > 0$, $z > 0$ and $yz-1 > 0$. We do exactly as in Appendix A, we are looking for critical points and the calculations are very similar. We find: $\lambda_0 = \frac{x}{2}(\frac{z(x^2+2)}{yz-1} - b)$, $\mu_0 = \frac{1}{8}\left(b^2 - \frac{z^2(x^2+2)^2}{(yz-1)^2}\right)$ and $\nu_0 = \frac{1}{8}\left((a-2)^2 - \frac{(x^2+2)^2}{(yz-1)^2}\right)$, which leads easily to I. □

Appendix E: Proof of Lemma 2.5.2

The pointwise limit of the cumulant generating function of the considered triplet is easily given by $\Lambda(0,\mu,\nu,\gamma)$ where Λ is defined in Lemma 2.4.1:

$$\Lambda(0,\mu,\nu,\gamma) = \begin{cases} -\frac{d}{2}(1+f) - \frac{ab}{4} + \frac{\gamma^2}{2f+a+2} & \text{if } \gamma < 0 \\ -\frac{d}{2}(1+f) - \frac{ab}{4} & \text{else.} \end{cases} \tag{E.1}$$

We recall that $d = \sqrt{b^2 - 8\mu}$ and $f = \dfrac{1}{2}\sqrt{(a-2)^2 - 8\nu}$. Using the Gärtner-Ellis theorem, we have

$$\widetilde{I}(y,z,t) = \sup_{\gamma\in\mathbb{R},\, \mu<\frac{b^2}{8},\, \nu<\frac{(a-2)^2}{8}} \{y\mu + z\nu + t\gamma - \Lambda(0,\mu,\nu,\gamma)\}. \tag{E.2}$$

With the same argument than for the other couple of simplified estimators, we show that $\widetilde{I}(y,z,t) = +\infty$ for $y < 0$, $z < 0$ or $yz-1 < 0$. We also notice that for $t > 0$ the expression inside the supremum tends to infinity as γ goes to infinity. So $\widetilde{I}(y,z,t) = +\infty$ for $t > 0$. Besides, for $t \leqslant 0$, the part involving γ is always negative for $\gamma \geqslant 0$ and sometimes positive for $\gamma < 0$. It implies that the supremum is necesseraly reached for $\gamma < 0$. Replacing μ and ν by their expression on d and f, we obtain, for $y > 0$, $z > 0$, $yz-1 > 0$ and $t \leqslant 0$,

$$\widetilde{I}(y,z,t) = \sup_{\gamma<0,\, d>0,\, f>0} \left\{ y\frac{b^2-d^2}{8} + z\frac{(a-2)^2 - 4f^2}{8} + t\gamma + \frac{ab}{4} + \frac{d}{2}(1+f) - \frac{\gamma^2}{2f+a+2} \right\}.$$

We investigate critical points. We obtain

$$f_0 = \frac{yt^2+2}{2(yz-1)},\ d_0 = \frac{t^2+2z}{yz-1} \text{ and } \gamma_0 = \frac{yt(t^2+2z)}{2(yz-1)} + \frac{at}{2}. \tag{E.3}$$

Replacing it into the expression of $\widetilde{I}$, we easily get the announced result. □

2.7 Comments and complements

This section is not part of the original article [27].

2.7.1 New proof for the large deviation results of Zani [73]

We are able to obtain the large deviation resuls of Zani [73] through our new method. However, we cannot handle the case $b = 0$, for which the speed in the large deviations is $\log T$ instead of T. To do so, we would need to compute the Laplace-transform (2.4.1) with another renormalization. By making use of (1.1) and Itô's formula applied to $\log X_T$, we rewrite the estimators $\bar{a}_T$ and $\bar{b}_T$ given by (1.7) and (1.8) as follows. If b is known, the MLE for a is given by

$$\bar{a}_T = \frac{\frac{\log X_T}{T} + 2\Sigma_T - b}{\Sigma_T}$$

whereas, if a is known, the MLE of b satisfies

$$\bar{b}_T = \frac{\frac{X_T}{T} - a}{S_T}.$$

where $S_T = T^{-1}\displaystyle\int_0^T X_t \, \mathrm{d}t$ and $\Sigma_T = T^{-1}\displaystyle\int_0^T X_t^{-1} \, \mathrm{d}t$.

We start by showing Theorem 1.2.3, which gives an LDP for $\bar{b}_T$. First of all, we establish an LDP for the couple $(\sqrt{X_T/T}, S_T)$. The pointwise limit L of its moment generating function is given, for any $\lambda \in \mathbb{R}$ and $\mu < \frac{b^2}{8}$, by

$$L(\lambda, \mu) = \Lambda(\lambda, \mu, 0, 0) = -\frac{a}{4}(d+b) + \frac{\lambda^2}{d-b}\mathbb{1}_{\lambda>0}$$

where the function Λ is given by Proposition 2.4.1 and d satisfies $d = \sqrt{b^2 - 8\mu}$. As L is a steep function, we obtain, by the Gärtner-Ellis theorem, that $(\sqrt{X_T/T}, S_T)$ satisfies an LDP with speed T and good rate function L^* where, for all $(x, y) \in \mathbb{R}^2$,

$$L^*(x, y) = \sup\left\{\lambda x + \mu y - L(\lambda, \mu) \,\middle|\, \lambda \in \mathbb{R}, \mu < b^2/8\right\}.$$

If $y \leqslant 0$, letting μ go to $-\infty$, we show that $L^*(x, y) = +\infty$. With the same argument, letting λ go to $-\infty$, we obtain that, for any $x < 0$, $L^*(x, y) = +\infty$. Additionnally, if $y > 0$ and $x \geqslant 0$, the function we are optimizing is always negative for $\lambda \leqslant 0$, and sometimes positive for $\lambda > 0$. Thus, replacing μ by its expression involving d, we obtain that, for all $y > 0$ and all $x \geqslant 0$,

$$L^*(x, y) = \sup\left\{\lambda x + \frac{b^2 - d^2}{8} y + \frac{a}{4}(d+b) - \frac{\lambda^2}{d-b} \,|\, \lambda > 0, d > 0\right\}. \tag{2.7.4}$$

We easily compute the critical point (d_0, λ_0), which satisfies

$$d_0 = \frac{a + x^2}{y} \quad \text{and} \quad \lambda_0 = \frac{x}{2}\left(\frac{a + x^2}{y} - b\right).$$

Replacing it into (2.7.4), we obtain that, for all $y > 0$ and all $x \geqslant 0$,

$$L^*(x,y) = \frac{b}{4}(a - x^2) + \frac{b^2}{8}y + \frac{(a + x^2)^2}{8y}. \tag{2.7.5}$$

Applying the contraction principle, we can deduce that $\bar{b}_T$ satisfies an LDP with speed T and good rate function I_b given over $\mathbb{R}$ by

$$I_b(\beta) = \inf \left\{ L^*(x,y) \,|x \geqslant 0, y > 0 \text{ and } f(x,y) = \beta \right\},$$

where f is given over $\mathbb{R}^+ \times \mathbb{R}_*^+$ by $f(x,y) = (x^2 - a)/y$ and is a continuous function over this domain. The condition $f(x,y) = \beta$ rewrites as $x^2 = \beta y + a$, which allows us to optimize over y only. Thus, replacing x^2 by its expression involving y and L^* by (2.7.5), we obtain that

$$I_b(\beta) = \inf_{\mathcal{D}_y} g(y) \qquad \text{where} \qquad g(y) = \frac{(b - \beta)^2}{8}y + \frac{a^2}{2y} + \frac{a\beta}{2} \tag{2.7.6}$$

and the domain $\mathcal{D}_y$ over which we are looking for the infimum of g satisfies

$$\mathcal{D}_y = \left\{ y > 0, \beta y - a \geqslant 0 \right\}.$$

Depending on the values of β, the domain $\mathcal{D}_y$ takes different shapes and the infimum will be reached either at a critical point, either at a boundary of the domain. For $\beta \geqslant 0$, $\mathcal{D}_y = \mathbb{R}_*^+$ and

$$I_b(\beta) = g\left(\frac{2a}{\beta - b}\right) = \frac{a}{2}\left(2\beta - b\right).$$

For $\beta < 0$, the domain becomes $\mathcal{D}_y = \{0 < y < -a/\beta\}$. If $\beta \neq b$, the infimum is reached at the critical point $y_0 = \frac{2a}{\beta - b}$ as soon as it belongs to $\mathcal{D}_y$, which implies that $\beta \geqslant \frac{b}{3}$. In this case, I_b take the previously computed value. Besides, we easily show that for any $\beta < b/3$,

$$I_b(\beta) = g\left(-\frac{a}{\beta}\right) = -\frac{a}{8\beta}\left(b - \beta\right)^2.$$

This concludes the proof of Theorem 1.2.3.

We now prove Theorem 1.2.2, with similar arguments. We suppose that $b < 0$ is known. We easily show an LDP for the couple $(\Sigma_T, \mathcal{L}_T)$, where $\mathcal{L}_T$ is defined by (2.7), using Proposition 2.4.1 and the Gärtner-Ellis theorem. The renormalized log-Laplace of this couple equals, for all $\nu < (a-2)^2/8$ and $\gamma \in \mathbb{R}$

$$\Lambda(0,0,\nu,\gamma) = \frac{b}{4}(1 + f) - \frac{ab}{4} + \frac{\gamma^2}{2f + a + 2}\mathbb{1}_{\gamma<0}$$

where $f = \frac{1}{2}\sqrt{(a-2)^2 - 8\nu}$. We compute its Fenchel-Legendre transform I, which gives the rate function of the LDP. It satisfies, for all $(z,t) \in \mathbb{R}^2$

$$I(z,t) = \begin{cases} \dfrac{z}{8}(a-2)^2 + \dfrac{(t^2 - b)^2}{8z} + \dfrac{t^2(a+2)}{2} + \dfrac{b}{4}(a-2) & \text{if } z > 0 \text{ and } t \leqslant 0 \\ +\infty & \text{otherwise.} \end{cases} \tag{2.7.7}$$

Applying the contraction principle, we obtain that $\bar{a}_T$ satisfies an LDP with speed T and good rate function I_a given for all $\alpha \in \mathbb{R}$ by

$$I_a(\alpha) = \inf \{ I(z,t) \,|\, z > 0, t \leqslant 0, f(z,t) = \alpha \}, \tag{2.7.8}$$

where the function f is given for any $z > 0$ and any $t \leqslant 0$ by

$$f(z,t) = \frac{-t^2 + 2z - b}{z}.$$

The condition $f(z,t) = \alpha$ rewrites as $(2-\alpha)z = t^2 + b$. As in the previous proof, we can deduce that $I_a(\alpha) = \inf_{\mathcal{D}_z} h(z)$ where, for all $z > 0$,

$$h(z) = \frac{z}{8}\left((a-\alpha)^2 - 8(\alpha-2)\right) + \frac{b^2}{2z} - b\left(1 + \frac{2-\alpha}{2}\right)$$

and $\mathcal{D}_z = \{z > 0, (2-\alpha)z - b \geqslant 0\}$. The critical point of h satisfies, in the case where it is well-defined,

$$z_0 = -\frac{2b}{\sqrt{(a-\alpha)^2 - 8(\alpha-2)}}. \tag{2.7.9}$$

Depending on the values of α, either z_0 is well-defined and inside $\mathcal{D}_z$ so that the infimum is reached at this point, either the infimum is reached at the boundary of $\mathcal{D}_z$.

If $\alpha \leqslant 2$, z_0 given by (2.7.9) is well-defined and inside $\mathcal{D}_z = \mathbb{R}_*^+$. We obtain that, for all $\alpha \leqslant 2$,

$$I_a(\alpha) = h(z_0) = -\frac{b}{2}\sqrt{(a-\alpha)^2 - 8(\alpha-2)} - b\left(2 - \frac{\alpha}{2}\right). \tag{2.7.10}$$

The case $\alpha > 2$ requires more attention. First of all, z_0 is well-defined only for the values of α satisfying

$$(a_\alpha)^2 - 8(\alpha-2) > 0,$$

which implies that :

$$\alpha < a + 4 - 2\sqrt{2a} \quad \text{or} \quad \alpha > a + 4 - 2\sqrt{2a}. \tag{2.7.11}$$

Additionnally, the domain over which we are looking for the infimum is reduced, due to the condition $(2-\alpha)z - b \geqslant 0$. Thus, $\mathcal{D}_z = \{0 < z < -b/(\alpha-2)\}$. In the case where z_0 is well-defined, it needs to be inside $\mathcal{D}_z$ for the infimum beeing reached at z_0. In other words, α must satisfy the following condition

$$\frac{-2b}{\sqrt{(a-\alpha)^2 - 8(\alpha-2)}} \leqslant \frac{-b}{\alpha-2},$$

which rewrites as

$$-3\alpha^2 + \alpha(8-2a) + a^2 \geqslant 0.$$

As $\alpha > 2$, we can deduce that

$$\alpha > \frac{1}{3}\left(-a + 4 + 2\sqrt{(a-2)^2 + 2a}\right). \tag{2.7.12}$$

To conclude, combining (2.7.11) and (2.7.12), we obtain that, for all $\alpha > 2$,

$$I_a(\alpha) = \begin{cases} h\left(\frac{-b}{\alpha-2}\right) & \text{if } \alpha > \alpha_1 \\ h(z_0) & \text{otherwise,} \end{cases} \tag{2.7.13}$$

where $\alpha_1 = \frac{1}{3}\left(-a + 4 + 2\sqrt{(a-2)^2 + 2a}\right)$, $h(z_0)$ is given by (2.7.10) and

$$h\left(\frac{-b}{\alpha-2}\right) = -\frac{b}{8}\frac{(\alpha-a)^2}{\alpha-2}.$$

Combining (2.7.10) and (2.7.13), we obtain Theorem 1.2.2.

2.7.2 Another couple of simplified estimators

In this chapter, we considered two couples simplified estimators constructed from the MLE by removing alternatively both terms that converge almost surely to zero (X_T/t and $\log X_T$). We consider now a third couple of simplified estimator build by removing both terms simultaneously and given as follows

$$\bar{a}_T = \frac{2S_T\Sigma_T}{V_T}, \quad \text{and} \quad \bar{b}_T = \frac{-2\Sigma_T}{V_T}.$$

This couple is strongly consistent and satisfies the same CLT than the MLE. Besides, we obtain the following large deviation results through really straightforward computations. We also notice that all rate functions are explicitely computed.

Theorem 2.7.1. *The sequence $(\bar{a}_T, \bar{b}_T)$ satisfies an LDP with speed T and good rate function $\bar{I}_{a,b}$ given, for all $(\alpha,\beta) \in \mathbb{R}^2$, by*

$$\bar{I}_{a,b}(\alpha,\beta) = \begin{cases} -\frac{b^2}{8}\frac{\alpha}{\beta} - \frac{(a-2)^2}{8}\frac{\beta}{\alpha-2} + \frac{ab-\beta}{4} & \text{if } \alpha > 2 \text{ and } \beta < 0 \\ +\infty & \text{otherwise.} \end{cases}$$

This result is a direct consequence of the contraction principle together with Lemma 2.3.1 and the fact that

$$(\bar{a}_T, \bar{b}_T) = f(S_T, \Sigma_T)$$

where f is given over $\{(x,y) \in \mathbb{R}^2 \,|xy - 1 \neq 0\}$ by

$$f(x,y) = \left(\frac{2xy}{xy-1}, \frac{-2y}{xy-1}\right).$$

We can deduce the following result

Corollary 2.7.1. *The sequences $\bar{a}_T$ and $\bar{b}_T$ both satisfy and LDP with speed T and good rate functions $\bar{I}_a$ and $\bar{I}_b$ espectively given, for all $\alpha \in \mathbb{R}$ and all $\beta \in \mathbb{R}$ by*

$$\bar{I}_a(\alpha) = \begin{cases} \frac{b}{4}\left(a - \sqrt{\frac{\alpha}{\alpha-2}(a^2 - 4a + 2\alpha)}\right) & \text{if } \alpha > 2 \\ +\infty & \text{otherwise,} \end{cases}$$

and

$$\bar{I}_b(\beta) = \begin{cases} -\frac{(b-\beta)^2}{4b} & \text{if } \beta < 0 \\ +\infty & \text{oterwise.} \end{cases}$$

Chapter 3

Déviations modérées pour l'EMV des paramètres du processus de Heston

Résumé. On établit un principe de déviations modérées pour l'estimateur du maximum de vraisemblance des quatre paramètres d'un processus de Heston géométriquement ergodique. On obtient également un principe de déviations modérées pour l'estimateur du maximum de vraisemblance du couple des paramètres de dimension et de dérive d'un processus CIR. On se restreint au cas le plus simple où le paramètre de dimension vérifie $a > 2$ et le coefficient de dérive est tel que $b < 0$. Contrairement à ce qui a été fait jusqu'ici dans la littérature, on estime les paramètres simultanément.

Abstract. We establish a moderate deviation principle for the maximum likelihood estimator of the four parameters of a geometrically ergodic Heston process. We also obtain moderate deviations for the maximum likelihood estimator of the couple of dimensional and drift parameters of a generalized squared radial Ornstein-Uhlenbeck process. We restrict ourselves to the most tractable case where the dimensional parameter satisfies $a > 2$ and the drift coefficient is such that $b < 0$. In contrast to the previous literature, parameters are estimated simultaneously.

Sommaire

3.1 Introduction

In the recent theory of hedging, a particular attention has been drawn to the study of stochastic volatility models in which the volatility itself is given as a solution of some stochastic differential equation, see [68], [61] and [39] for financial accuracy. Among them, Heston process [47] is one of the most popular, due to its computational tractability. For example, call option prices are succesfully computed in [60] using Fourier inversion techniques. We denote by Y_t the logarithm of the price of a given asset and by X_t its instantaneous variance, and we consider the following Heston process

$$\begin{cases} \mathrm{d}X_t = (a + bX_t)\,\mathrm{d}t + 2\sqrt{X_t}\,\mathrm{d}B_t \\ \mathrm{d}Y_t = (c + dX_t)\,\mathrm{d}t + 2\sqrt{X_t}\,\left(\rho\,\mathrm{d}B_t + \sqrt{1-\rho^2}\,\mathrm{d}W_t\right) \end{cases} \tag{3.1}$$

with $a > 0$, $(b, c, d) \in \mathbb{R}^3$ and $\rho \in]-1, 1[$, where (B_t, W_t) is a 2-dimensional standard Wiener process and the initial state $(x_0, y_0) \in \mathbb{R}^+ \times \mathbb{R}$. In this process, the volatility X_t is driven by a generalized squared radial Ornstein-Uhlenbeck process, also known as the CIR process, firstly studied by Feller [34] and introduced in a financial context by Cox, Ingersoll and Ross [22] to compute short-term interest rates. The behaviour of the CIR process has been widely investigated and depends on the values of both coefficients a and b. We shall restrict ourself to the most tractable situation where $a > 2$ and $b < 0$. In this case, the CIR process is geometrically ergodic and never reaches zero.

To calibrate the model, we estimate all the parameters (a, b, c, d) at the same time using a trajectory of (X_t) and (Y_t) over the time interval $[0, T]$. Azencott and Gadhyan [6] developed an algorithm to estimate some parameters of the Heston process based on discrete time observations, by making use of Euler and Milstein discretization schemes for the maximum likelihood. However, in the special case of an Heston process, the exact likelihood can be computed. It allows us to construct the maximum likelihood estimator (MLE) without using sophisticated approximation procedures, which are necessary for many stochastic volatility models, see [1]. The MLE $\widehat{\theta}_T = (\widehat{a}_T, \widehat{b}_T, \widehat{c}_T, \widehat{d}_T)$ of $\theta = (a, b, c, d)$ has been recently investigated in [11], together with its asymptotic behavior in the special case where $a \geqslant 2$. It is given as follows

$$\widehat{\theta}_T = \theta + 2\begin{pmatrix} \langle M\rangle_T^{-1} & 0 \\ 0 & \langle M\rangle_T^{-1} \end{pmatrix}\begin{pmatrix} M_T \\ N_T \end{pmatrix}, \tag{3.2}$$

where M_T and N_T are continuous-time martingales respectively given by

$$M_T = \left(\int_0^T X_t^{-1/2}\,\mathrm{d}B_t, \int_0^T X_t^{1/2}\,\mathrm{d}B_t\right)^{\intercal}, N_T = \left(\int_0^T X_t^{-1/2}\,\mathrm{d}\widetilde{B}_t, \int_0^T X_t^{1/2}\,\mathrm{d}\widetilde{B}_t\right)^{\intercal} \tag{3.3}$$

with $\mathrm{d}\widetilde{B}_t = \rho\,\mathrm{d}B_t + \sqrt{1-\rho^2}\,\mathrm{d}W_t$, and $\langle M\rangle_T$ is the increasing process of M_T given by

$$\langle M\rangle_T = T\begin{pmatrix} \Sigma_T & 1 \\ 1 & S_T \end{pmatrix} \tag{3.4}$$

with $S_T = T^{-1}\int_0^T X_t\,\mathrm{d}t$ and $\Sigma_T = T^{-1}\int_0^T X_t^{-1}\,\mathrm{d}t$. This estimator is strongly consistent, i.e. $\widehat{\theta}_T$ converges almost surely to θ as T goes to infinity. It also satisfies the following central limit theorem (CLT):

$$\sqrt{T}\left(\widehat{\theta}_T - \theta\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 4R \otimes \Sigma^{-1})$$

where $\otimes$ stands for the Kronecker product and the matrix Σ and R are respectively given by

$$\Sigma = \begin{pmatrix} \frac{-b}{a-2} & 1 \\ 1 & -\frac{a}{b} \end{pmatrix} \qquad \text{and} \qquad R = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}. \tag{3.5}$$

One can observe that $(\widehat{a}_T, \widehat{b}_T)$ coincides with the MLE of the parameters (a, b) of the CIR process based on the observation of (X_T) over the time interval $[0, T]$:

$$\begin{pmatrix} \widehat{a}_T \\ \widehat{b}_T \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} + 2 \langle M \rangle_T^{-1} M_T. \tag{3.6}$$

Asymptotic results about this estimator can be found in [65], [37] and [14], and a large deviation principle (LDP) has been recently established in [27]. In the easier case where one parameter is estimated while the other one is supposed to be known, [73] gives large deviations whereas [38] derives moderate deviations.

In this paper, our goal is to establish a moderate deviation principle (MDP) for the MLE of the four parameters of the Heston process given in (3.2), which is a natural continuation of the central limit theorem. Let us first recall some basic definitions of large deviation theory. We refer to [25] for further details. Let $(\lambda_T)_T$ be a positive sequence of real numbers increasing to infinity with T. A sequence $(Z_T)_T$ of $\mathbb{R}^d$-valued random variables satisfies a large deviation principle (LDP) with speed λ_T and rate function $I : \mathbb{R}^d \mapsto [0, +\infty]$ if I is lower semi-continuous and such that $(Z_T)_T$ satisfies the following upper and lower bounds: for any closed set F of $\mathbb{R}^d$

$$\limsup_{T \to +\infty} \lambda_T^{-1} \log \mathbb{P}\left(Z_T \in F\right) \leqslant - \inf_{z \in F} I\left(z\right)$$

and for any open set G of $\mathbb{R}^d$

$$\liminf_{T \to +\infty} \lambda_T^{-1} \log \mathbb{P}\left(Z_T \in G\right) \geqslant - \inf_{z \in G} I\left(z\right).$$

If furthermore the level sets of I are compacts, I is called a good rate function. Additionally, if $\lambda_T = o(T)$, a sequence $(Z_T)_T$ of $\mathbb{R}^d$-valued random variables satisfies a moderate deviation principle (MDP) with speed λ_T and rate function I, if the sequence $(\sqrt{T/\lambda_T} Z_T)_T$ satisfies an LDP with speed λ_T and rate function I.

In other words, let $(\lambda_T)_T$ be a positive sequence satisfying for T going to infinity

$$\lambda_T \to +\infty \quad \text{and} \quad \frac{\lambda_T}{T} \to 0. \tag{3.7}$$

We investigate in this paper an LDP for $\sqrt{\frac{T}{\lambda_T}}(\widehat{\theta}_T - \theta)$ with speed λ_T and compute the explicit rate function. By the way, we also establish an MDP for the MLE of the two parameters of a CIR process. The paper is organised as follows. Section 2 contains our main results while Sections 3 and 4 are devoted to their proofs. For the remaining of the paper, we suppose that the starting point x_0 of the CIR process is strictly positive.

3.2 Main results

We establish MDPs for the MLE $(\widehat{a}_T, \widehat{b}_T)$ of the parameters (a, b) of the CIR process given by (3.6), as well as for the MLE $\widehat{\theta}_T$ of the Heston process given by (3.2).

Theorem 3.2.1. *The sequence* $\left(\sqrt{\frac{T}{\lambda_T}}(\widehat{a}_T - a, \widehat{b}_T - b)\right)$ *satisfies an LDP with speed* λ_T *and good rate function* $I_{a,b}$ *given for all* $(\alpha, \beta) \in \mathbb{R}^2$ *by*

$$I_{a,b}(\alpha, \beta) = -\frac{b}{8(a-2)}\alpha^2 - \frac{a}{8b}\beta^2 + \frac{\alpha\beta}{4}. \tag{3.8}$$

Theorem 3.2.2. *The sequence* $\left(\sqrt{\frac{T}{\lambda_T}}(\widehat{\theta}_T - \theta)\right)$ *satisfies an LDP with speed* λ_T *and good rate function* I_θ *given for all* $(\alpha, \beta, \gamma, \delta) \in \mathbb{R}^4$ *by*

$$I_\theta(\alpha, \beta, \gamma, \delta) = \left(1 - \rho^2\right)^{-1} \left(I_{a,b}\left(\alpha, \beta\right) + I_{a,b}\left(\gamma, \delta\right) + \rho\, J\left(\alpha, \beta, \gamma, \delta\right)\right).$$

where $I_{a,b}$ *is given by* (3.8) *and*

$$J\left(\alpha, \beta, \gamma, \delta\right) = -\alpha\delta + \frac{b}{a-2}\alpha\gamma - \beta\gamma + \frac{a}{b}\beta\delta.$$

Remark 3.2.1. *One can observe that the first result is a particular case of the second one, with* $\gamma = \delta = \rho = 0$, *and we easily check that*

$$I_{a,b} = I_\theta\left(\cdot, \cdot, 0, 0\right).$$

The proofs are respectively postponed to Sections 3 and 4.

3.3 Proof of the MDP for the CIR process

We rewrite the MLE $(\widehat{a}_T, \widehat{b}_T)$ of the parameters (a, b) of the CIR process as follows:

$$\sqrt{\frac{T}{\lambda_T}} \begin{pmatrix} \widehat{a}_T - a \\ \widehat{b}_T - b \end{pmatrix} = 2T \left\langle M \right\rangle_T^{-1} \frac{1}{\sqrt{\lambda_T T}} M_T. \tag{3.9}$$

In order to prove Theorem 3.2.1, we first establish LDPs with speed λ_T for $(\lambda_T T)^{-1/2} M_T$ and for $T^{-1}\left\langle M \right\rangle_T$. Then we conclude using the contraction principle (see Theorem 4.2.1 of [25]) which is recalled here for completeness.

Lemma 3.3.1 (Contraction Principle)**.** *Let* $(Z_T)_T$ *be a sequence of random variables of* $\mathbb{R}^d$ *satisfying an LDP with good rate function* I *and* $g : \mathbb{R}^d \to \mathbb{R}^n$ *be a continuous function over* $\mathcal{D}_I = \left\{x \in \mathbb{R}^d | I(x) < +\infty\right\}$. *Then, the sequence* $(g(Z_T))_T$ *satisfies an LDP with good rate function* J *defined for all* $y \in \mathbb{R}^n$ *by*

$$J\left(y\right) = \inf_{\{x \in \mathcal{D}_I | g(x) = y\}} I\left(x\right),$$

where the infimum over the empty set is equal to infinity.

Lemma 3.3.2. *The sequence* $\left((\lambda_T T)^{-1/2} M_T\right)$ *satisfies an LDP with speed* λ_T *and good rate function* I_M *given, for all* $(m,n) \in \mathbb{R}^2$, *by*

$$I_M(m,n) = -\frac{a(a-2)}{4b}m^2 - \frac{b}{4}n^2 - \frac{a-2}{2}mn. \tag{3.10}$$

Proof. We establish this LDP by applying the Gärtner-Ellis Theorem. Thus, for T going to infinity, we need to compute the pointwise limit Λ of the normalized cumulant generating function Λ_T of $\left((\lambda_T T)^{-1/2} M_T\right)$ given, for all $v \in \mathbb{R}^2$, by

$$\Lambda_T(v) = \frac{1}{\lambda_T} \log \mathbb{E}\left[e^{\sqrt{\lambda_T/T}\langle v, M_T\rangle}\right]$$

We show in Corollary B.1 of Appendix B that for all $v \in \mathbb{R}^2$,

$$\Lambda(v) = \lim_{T\to+\infty} \Lambda_T(v) = \frac{1}{2} v^\intercal \Sigma v, \tag{3.11}$$

where the matrix Σ was previously given in (3.5). Thus, by the Gärtner-Ellis Theorem, the sequence $\left((\lambda_T T)^{-1/2} M_T\right)$ satisfies an LDP with speed λ_T and rate function I_M given by the Fenchel-Legendre transform of Λ, for all $\mu \in \mathbb{R}^2$,

$$\begin{aligned} I_M(\mu) &= \sup\left\{\langle \mu, v\rangle - \Lambda(v) \,|\, v \in \mathbb{R}^2\right\} \\ &= \sup\left\{\mu^\intercal v - \frac{1}{2} v^\intercal \Sigma v \,|\, v \in \mathbb{R}^2\right\}. \end{aligned} \tag{3.12}$$

The function of v that we need to optimise is non convex with a critical point $v_0 = \Sigma^{-1}\mu$. Replacing it into (3.12) and using the fact that Σ is symmetric, we obtain that

$$I_M(\mu) = \frac{1}{2} \mu^\intercal \Sigma^{-1} \mu$$

which easily leads to the announced result. ☐

Lemma 3.3.3. *The sequence* $\left((\lambda_T T)^{-1/2} N_T\right)$ *satisfies an LDP with speed* λ_T *and good rate function* I_M *given by* (3.10).

Proof. It works as in the previous proof, except that, this time, we make use of Corollary B.2. ☐

Lemma 3.3.4. *The sequences* $\left(2T\langle M\rangle_T^{-1} (\lambda_T T)^{-1/2} M_T\right)$ *and* $\left(2\Sigma^{-1} (\lambda_T T)^{-1/2} M_T\right)$ *are exponentially equivalent with speed* λ_T, *which means that, for all* $\varepsilon > 0$,

$$\limsup_{T\to+\infty} \frac{1}{\lambda_T} \log \mathbb{P}\left(\|\left(T\langle M\rangle_T^{-1} - \Sigma^{-1}\right)(\lambda_T T)^{-1/2} M_T \| \geqslant \varepsilon\right) = -\infty,$$

where $\|\cdot\|$ *is the Euclidean norm.*

Proof. For all $\varepsilon > 0$ and all $\eta > 0$, we have the following upper bound:

$$\limsup_{T\to+\infty} \frac{1}{\lambda_T} \log \mathbb{P}\left(\| \left(T \langle M\rangle_T^{-1} - \Sigma^{-1}\right) (\lambda_T T)^{-1/2} M_T \| \geqslant \varepsilon\right) \leqslant \limsup_{T\to+\infty} \max\{P_T^\eta; Q_T^{\eta,\varepsilon}\}$$

with

$$P_T^\eta = \frac{1}{\lambda_T} \log \mathbb{P}\left(\| T \langle M\rangle_T^{-1} - \Sigma^{-1} \| \geqslant \eta\right)$$

and

$$Q_T^{\eta,\varepsilon} = \frac{1}{\lambda_T} \log \mathbb{P}\left(\| (\lambda_T T)^{-1/2} M_T \| \geqslant \frac{\varepsilon}{\eta}\right),$$

where we also denote by $\| \cdot \|$ the subordinate norm. On the one hand, we rewrite

$$\begin{aligned}
P_T^\eta &= \frac{T}{\lambda_T}\frac{1}{T} \log \mathbb{P}\left(\|T \langle M\rangle_T^{-1} - \Sigma^{-1}\| \geqslant \eta\right) \\
&\leqslant \frac{T}{\lambda_T}\frac{1}{T} \log \left[\mathbb{P}\left(\left(\frac{\Sigma_T}{V_T} + \frac{b}{2}\right)^2 + \left(\frac{S_T}{V_T} + \frac{a(a-2)}{2b}\right)^2 \geqslant \frac{\eta^2}{2}\right) + \mathbb{P}\left(\left|\frac{a-2}{2} - \frac{1}{V_T}\right| \geqslant \frac{\eta}{2}\right)\right] \\
&\leqslant \frac{1}{\lambda_T} \log 2 + \frac{T}{\lambda_T} \max\{p_T^\eta; q_T^\eta\}
\end{aligned}$$

where $V_T = S_T \Sigma_T - 1$, and p_T^η and q_T^η are respectively given by

$$p_T^\eta = \frac{1}{T} \log \mathbb{P}\left(\left(\frac{\Sigma_T}{V_T} + \frac{b}{2}\right)^2 + \left(\frac{S_T}{V_T} + \frac{a(a-2)}{2b}\right)^2 \geqslant \frac{\eta^2}{2}\right)$$

and

$$q_T^\eta = \frac{1}{T} \log \mathbb{P}\left(\left|\frac{a-2}{2} - \frac{1}{V_T}\right| \geqslant \frac{\eta}{2}\right).$$

We show in Appendix C that, for all $\eta > 0$,

$$p^\eta := \limsup_{T\to+\infty} p_T^\eta < 0 \quad \text{and} \quad q^\eta := \limsup_{T\to+\infty} q_T^\eta < 0. \tag{3.13}$$

Using the fact that $\frac{T}{\lambda_T}$ tends to infinity, we obtain that for all $\eta > 0$,

$$\limsup_{T\to+\infty} P_T^\eta = -\infty.$$

On the other hand, as we have established in Lemma 3.3.2 an MDP for $\left((\lambda_T T)^{-1/2} M_T\right)$ with rate function I_M, we have the following upper bound

$$\limsup_{T\to+\infty} Q_T^{\eta,\varepsilon} \leqslant -\inf\{I_M(m,n) \,|\, (m,n) \notin \mathcal{B}(0, \varepsilon/\eta)\}$$

Letting η tend to zero, we obtain that

$$\limsup_{\eta\to 0} \limsup_{T\to+\infty} Q_T^{\eta,\varepsilon} = -\infty.$$

Thus, for all $\varepsilon > 0$,

$$\limsup_{\eta\to 0} \limsup_{T\to+\infty} \max\{P_T^\eta; Q_T^{\eta,\varepsilon}\} = -\infty.$$

$\square$

Lemma 3.3.5. *The sequences* $\left(2T\left\langle M\right\rangle_T^{-1}\left(\lambda_T T\right)^{-1/2} N_T\right)$ *and* $\left(2\Sigma^{-1}\left(\lambda_T T\right)^{-1/2} N_T\right)$ *are exponentially equivalent with speed* λ_T*, which means that, for all* $\varepsilon > 0$*,*

$$\limsup_{T\to+\infty} \frac{1}{\lambda_T} \log \mathbb{P}\left(\| \left(T\left\langle N\right\rangle_T^{-1} - \Sigma^{-1}\right)\left(\lambda_T T\right)^{-1/2} N_T \| > \varepsilon\right) = -\infty,$$

where $\|\cdot\|$ *is the Euclidean norm.*

Proof. As the martingales M_T and N_T share the same bracket process and the same MDP, the previous proof remains valid for the martingale N_T instead of M_T. $\square$

Theorem 3.2.1 immediately follows from Lemmas 3.3.2 and 3.3.4, by a straightforward application of the contraction principle recalled in Lemma 3.3.1.

Proof of Theorem 3.2.1. Combining Lemma 3.3.1 with Lemma 3.3.2, we show that the sequence $(2\Sigma^{-1}\left(\lambda_T T\right)^{-1/2} M_T)$ satisfies an LDP with speed λ_T and rate function $I_{a,b}$ given for all $(\alpha, \beta) \in \mathbb{R}^2$ by

$$I_{a,b}(\alpha,\beta) = I_M\left(\frac{1}{2}\Sigma\,(\alpha,\beta)^\intercal\right),$$

where I_M is obtained in Lemma 3.3.2. Besides, by Lemma 3.3.4 and equation (3.9), the sequences $\sqrt{T\lambda_T^{-1}}\left(\widehat{a}_T - a, \widehat{b}_T - b\right)$ and $(2\Sigma^{-1}\left(\lambda_T T\right)^{-1/2} M_T)$ share the same LDP with speed λ_T. This gives the announced result. $\square$

3.4 Proof of the MDP for the Heston process

We now go back to the MLE $\widehat{\theta}_T$ and prove the main Theorem 3.2.2. We denote by $\mathcal{M}_T$ the continuous-time martingale

$$\mathcal{M}_T = \begin{pmatrix} M_T \\ N_T \end{pmatrix},$$

where M_T and N_T are given by (3.3). As $\left\langle \mathrm{d}B_t, \mathrm{d}\widetilde{B}_t\right\rangle = \rho\,\mathrm{d}t$, we easily obtain that the bracket process of $\mathcal{M}_T$ equals

$$\langle\mathcal{M}\rangle_T = R \otimes \langle M\rangle_T \tag{3.14}$$

where the matrix R is given in (3.5). We follow the same scheme than in the previous section. We have the following decompostion:

$$\sqrt{T\lambda_T^{-1}}\left(\widehat{\theta}_T - \theta\right) = 2T\left(I_2 \otimes \langle M\rangle_T^{-1}\right)\left(\lambda_T T\right)^{-1/2}\mathcal{M}_T.$$

We first establish an LDP with speed λ_T for the 4-dimensional martingale $\left(\lambda_T T\right)^{-1/2}\mathcal{M}_T$, by making use of the exponential convergence of its bracket process. Then we show that the estimator is exponentially equivalent to some sequence involving the matrix Σ and the martingale $\mathcal{M}_T$, for which we are able to establish an LDP by applying the contraction principle.

Lemma 3.4.1. *The sequence* $\left((\lambda_T T)^{-1/2}\mathcal{M}_T\right)$ *satisfies an LDP with speed* λ_T *and good rate function* $I_{\mathcal{M}}$ *given for all* $(x,y,z,t)\in\mathbb{R}^4$ *by*

$$I_{\mathcal{M}}(x,y,z,t) = \left(1-\rho^2\right)^{-1}\left(I_M(x,y) + I_M(z,t) + \rho\frac{a-2}{2}\left(yz + xt + \frac{a}{b}zx + \frac{b}{a-2}ty\right)\right),$$

where I_M *is given by* (3.10).

Proof. We establish this LDP by applying again the Gärtner-Ellis Theorem. Thus, we need to compute the pointwise limit, for T going to infinity, of the normalized cumulant generating function $\mathcal{L}_T$ of $\left((\lambda_T T)^{-1/2}\mathcal{M}_T\right)$ given for all $u\in\mathbb{R}^4$ by

$$\mathcal{L}_T(u) = \frac{1}{\lambda_T}\log\mathbb{E}\left[e^{\sqrt{\lambda_T/T}\langle u,\mathcal{M}_T\rangle}\right].$$

We show in Appendix B that,

$$\lim_{T\to+\infty}\mathcal{L}_T(u) = \frac{1}{2}u^{\intercal}\Gamma u, \tag{3.15}$$

where the matrix Γ is given by

$$\Gamma = R\otimes\Sigma = \begin{pmatrix}\Sigma & \rho\Sigma\\ \rho\Sigma & \Sigma\end{pmatrix} \tag{3.16}$$

with R and Σ given in (3.5). Thus, the sequence $\left((\lambda_T T)^{-1/2}\mathcal{M}_T\right)$ satisfies an LDP with speed λ_T and good rate function $I_{\mathcal{M}}$ given for all $\mu\in\mathbb{R}^4$ by

$$I_{\mathcal{M}}(\mu) = \sup_{\Lambda\in\mathbb{R}^4}\left\{\mu^{\intercal}\Lambda - \frac{1}{2}\Lambda^{\intercal}\Gamma\Lambda\right\} = \frac{1}{2}\mu^{\intercal}\Gamma^{-1}\mu. \tag{3.17}$$

It is not hard to see that

$$\Gamma^{-1} = R^{-1}\otimes\Sigma^{-1} = \left(1-\rho^2\right)^{-1}\begin{pmatrix}\Sigma^{-1} & -\rho\Sigma^{-1}\\ -\rho\Sigma^{-1} & \Sigma^{-1}\end{pmatrix}. \tag{3.18}$$

We complete the proof of Lemma 3.4.1 combining (3.17) and (3.18). □

Lemma 3.4.2. *We denote by* $\mathcal{S}$ *the block matrix*

$$\mathcal{S} = I_2\otimes\Sigma = \begin{pmatrix}\Sigma & 0\\ 0 & \Sigma\end{pmatrix} \tag{3.19}$$

where I_2 *stands for the identity matrix of size 2 and* Σ *is given by* (3.5). *The sequences* $\left(\sqrt{\frac{T}{\lambda_T}}(\widehat{\theta}_T-\theta)\right)$ *and* $\left(2\mathcal{S}^{-1}(\lambda_T T)^{-1/2}\mathcal{M}_T\right)$ *are exponentially equivalent with speed* λ_T.

Proof. We want to prove that for all $\varepsilon > 0$,

$$\limsup_{T\to+\infty} \lambda_T^{-1} \log \mathbb{P}\left(\| \left(I_2 \otimes B_T\right) \left(\lambda_T T\right)^{-1/2} \mathcal{M}_T \| > \varepsilon\right) = -\infty, \tag{3.20}$$

where $\| \cdot \|$ is the Euclidean norm and $B_T = T \langle M \rangle_T^{-1} - \Sigma^{-1}$. We have the following upper bound:

$$\mathbb{P}\left(\| \left(I_2 \otimes B_T\right) \left(\lambda_T T\right)^{-1/2} \mathcal{M}_T \| > \varepsilon\right) \leqslant P_T^M + P_T^N$$

where $P_T^M = \mathbb{P}\left(\| B_T \left(\lambda_T T\right)^{-1/2} M_T \| > \frac{\varepsilon}{\sqrt{2}}\right)$ and $P_T^N = \mathbb{P}\left(\| B_T \left(\lambda_T T\right)^{-1/2} N_T \| > \frac{\varepsilon}{\sqrt{2}}\right)$. We have shown in Lemma 3.3.4 that

$$\limsup_{T\to+\infty} \lambda_T^{-1} \log P_T^M = -\infty,$$

and, by Lemma 3.3.5, we also know that

$$\limsup_{T\to+\infty} \lambda_T^{-1} \log P_T^N = -\infty.$$

This leads to (3.20) which gives the announced result. $\square$

The proof of Theorem 3.2.2 is a direct consequence of both Lemmas 3.4.1 and 3.4.2.

Proof of Theorem 3.2.2. The contraction principle together with Lemma 3.4.1 imply that the sequence $\left(2\mathcal{S}^{-1} \left(\lambda_T T\right)^{-1/2} \mathcal{M}_T\right)$ satisfies an LDP with speed λ_T. However, it follows from Lemma 3.4.2 that this sequence is exponentially equivalent to $\sqrt{\frac{T}{\lambda_T}}\left(\widehat{\theta}_T - \theta\right)$, which means that they share the same LDP. The rate function I_θ is given by the contraction principle as follows. For all $\mu \in \mathbb{R}^4$,

$$I_\theta\left(\mu\right) = \inf_{\Lambda\in\mathbb{R}^4} \left\{ I_{\mathcal{M}}(\Lambda) \mid \mu = 2\mathcal{S}^{-1}\Lambda \right\}$$

where $I_{\mathcal{M}}$ is given in Lemma 3.4.1. Thus,

$$I_\theta\left(\mu\right) = I_{\mathcal{M}}\left(\frac{1}{2}\mathcal{S}\mu\right)$$

which leads to the announced result, just replacing $I_{\mathcal{M}}$ by its expression. $\square$

Appendix A: Changes of parameters

To compute the limits of the cumulant generating functions (3.11) and (3.15), which is the aim of Appendix B, we recall some changes of probability formulas. We denote by $\mathbb{P}_{c,d}^{a,b}$ the distribution of the solution of (3.1) associated with parameters a, b, c and d, and by $\mathbb{E}_{c,d}^{a,b}$ the corresponding expectation. At first, we change both parameters a and b of the first equation of (3.1), which corresponds to the CIR process. Applying Girsanov's formula given e.g. in Theorem 1.12 of [57], we have

$$\begin{aligned}
\frac{\mathrm{d}\mathbb{P}^{a,b}_{c,d}}{\mathrm{d}\mathbb{P}^{\alpha,\beta}_{c,d}} &= \exp\left(\int_0^T \frac{a+bX_t-(\alpha+\beta X_t)}{4X_t}\,\mathrm{d}X_t - \frac{1}{2}\int_0^T \frac{(a+bX_t)^2-(\alpha+\beta X_t)^2}{4X_t}\,\mathrm{d}t\right)\\
&= \exp\left(\frac{a-\alpha}{4}\int_0^T X_t^{-1}\,\mathrm{d}X_t + \frac{b-\beta}{4}\int_0^T \mathrm{d}X_t\right.\\
&\qquad \left. -\frac{1}{8}\int_0^T \left((a^2-\alpha^2)X_t^{-1}+(b^2-\beta^2)X_t+2(ab-\alpha\beta)\right)\,\mathrm{d}t\right)\\
&= X_T^{\frac{a-\alpha}{4}}\exp\left(-\frac{a-\alpha}{4}(\log X_0+bT)+\frac{b-\beta}{4}(X_T-X_0-\alpha T)\right)\\
&\qquad \times\exp\left(-\frac{T}{8}\left((b^2-\beta^2)S_T+\left(4(\alpha-a)-\alpha^2+a^2\right)\Sigma_T\right)\right).
\end{aligned} \tag{A.1}$$

For the last equality, we use Itô's formula applied to $\log X_T$, which gives

$$\int_0^T X_t^{-1}\,\mathrm{d}X_t = \log X_T - \log X_0 + 2\int_0^T X_t^{-1}\,\mathrm{d}t.$$

We also need to change parameters c and d of the second equation in (3.1). We rewrite this equation with new parameters γ and δ:

$$\mathrm{d}Y_t = (\gamma+\delta X_t)\,\mathrm{d}t + 2\sqrt{X_t}\left[\rho\,\mathrm{d}B_t + \sqrt{1-\rho^2}\mathrm{d}\widetilde{W}_t\right]$$

where

$$\mathrm{d}\widetilde{W}_t = \mathrm{d}W_t + \frac{\mathrm{d}t}{2\sqrt{1-\rho^2}}\left((c-\gamma)\,X_t^{-1/2}+(d-\delta)\,X_t^{1/2}\right).$$

Thus, we obtain that

$$\begin{aligned}
\frac{\mathrm{d}\mathbb{P}^{a,b}_{c,d}}{\mathrm{d}\mathbb{P}^{a,b}_{\gamma,\delta}} &= \exp\left(\frac{c-\gamma}{2\sqrt{1-\rho^2}}\int_0^T X_t^{-1/2}\,\mathrm{d}W_t + \frac{d-\delta}{2\sqrt{1-\rho^2}}\int_0^T X_t^{1/2}\,\mathrm{d}W_t\right)\\
&\qquad \times\exp\left(\frac{T}{8\,(1-\rho^2)}\left[(d-\delta)^2S_T+(c-\gamma)^2\,\Sigma_T+2\,(c-\gamma)\,(d-\delta)\right]\right).
\end{aligned} \tag{A.2}$$

Appendix B: Proof of the pointwise limit of the cumulant generating function

We want to compute the pointwise limit of the cumulant generating function $\mathcal{L}_T$ of $(\lambda_T T)^{-1/2}\,\mathcal{M}_T$. With $\mathcal{M}_T$ being replaced by its expression in X_t, B_t, W_t and ρ, we have that, for all $u=(u_1,u_2,u_3,u_4)\in\mathbb{R}^4$,

$$E_T\,(u) := \mathbb{E}^{a,b}_{c,d}\left[e^{\sqrt{\lambda_T/T}\,\langle u,\mathcal{M}_T\rangle}\right] = \mathbb{E}^{a,b}_{c,d}\left[\mathcal{E}_{1,T}\,\mathcal{E}_{2,T}\,\mathcal{E}_{3,T}\,\mathcal{E}_{4,T}\right] \tag{B.1}$$

where

$$\mathcal{E}_{1,T} = \exp\left(v_{3,T}\int_0^T X_t^{-1/2}\,\mathrm{d}B_t\right),\quad \mathcal{E}_{2,T} = \exp\left(v_{4,T}\int_0^T X_t^{1/2}\,\mathrm{d}B_t\right),$$

$$\mathcal{E}_{3,T} = \exp\left(\sqrt{\frac{\lambda_T}{T}}\sqrt{1-\rho^2}u_3\int_0^T X_t^{-1/2}\,\mathrm{d}W_t\right),$$

$$\mathcal{E}_{4,T} = \exp\left(\sqrt{\frac{\lambda_T}{T}}\sqrt{1-\rho^2}u_4\int_0^T X_t^{1/2}\,\mathrm{d}W_t\right).$$

with $v_{3,T} = \sqrt{\lambda_T/T}(u_1+\rho u_3)$ and $v_{4,T} = \sqrt{\lambda_T/T}(u_2+\rho u_4)$. We use (A.2) to change parameters c and d in order to kill the terms involving W_t. We obtain that

$$\mathbb{E}_{c,d}^{a,b}\left[\mathcal{E}_{1,T}\,\mathcal{E}_{2,T}\,\mathcal{E}_{3,T}\,\mathcal{E}_{4,T}\right] = \mathbb{E}_{\gamma_T,\delta_T}^{a,b}\left[G_T\mathcal{E}_{1,T}\,\mathcal{E}_{2,T}\right] \tag{B.2}$$

where $\gamma_T = c + 2u_3\sqrt{\lambda_T/T}\,(1-\rho^2)$, $\delta_T = d + 2u_4\sqrt{\lambda_T/T}\,(1-\rho^2)$ and

$$G_T = \exp\left(\frac{T}{8\,(1-\rho^2)}\left((c-\gamma_T)^2\Sigma_T + (d-\delta_T)^2 S_T + 2(c-\gamma_T)(d-\delta_T)\right)\right). \tag{B.3}$$

Additionally, using the first equation of (3.1) and Itô's formula applied to $\log X_T$, we obtain that

$$2\int_0^T X_t^{1/2}\,\mathrm{d}B_t = X_T - x_0 - a\,T - b\,T S_T \tag{B.4}$$

and

$$2\int_0^T X_t^{-1/2}\,\mathrm{d}B_t = \log X_T - \log x_0 - b\,T + (2-a)\,T\,\Sigma_T \tag{B.5}$$

where S_T and Σ_T are given after (3.4). Thus, we rewrite $\mathcal{E}_{1,T}$ and $\mathcal{E}_{2,T}$ as functions of X_T, S_T and Σ_T:

$$\mathcal{E}_{1,T} = \exp\left(\frac{v_{3,T}}{2}\left(\log X_T - \log x_0 - bT + (2-a)T\Sigma_T\right)\right) \tag{B.6}$$

and

$$\mathcal{E}_{2,T} = \exp\left(\frac{1}{2}v_{4,T}\left(X_T - x_0 - aT - bTS_T\right)\right). \tag{B.7}$$

Thus, replacing (B.3), (B.7) and (B.6) into (B.2), and taking out of the expectation all the deterministic terms, we obtain that

$$E_T\,(u) = \exp\left(\mathcal{A}_T\right)\mathbb{E}_{\gamma_T,\delta_T}^{a,b}\left[\exp\left(v_{1,T}T\Sigma_T + v_{2,T}TS_T + \frac{1}{2}\left(v_{3,T}X_T + v_{4,T}\log X_T\right)\right)\right] \tag{B.8}$$

where

$$v_{1,T} = \frac{(c-\gamma_T)^2}{8\,(1-\rho^2)} + \frac{2-a}{2}v_{3,T}\,,\quad v_{2,T} = \frac{(d-\delta_T)^2}{8\,(1-\rho^2)} - \frac{b}{2}v_{4,T}$$

and

$$\mathcal{A}_T = T\left(1-\rho^2\right)^{-1}(c-\gamma_T)\,(d-\delta_T)\,/4 - v_{4,T}\,(aT+x_0)\,/2 - v_{3,T}\,(\log x_0 + bT)\,/2.$$

We now make a new change of parameters for a and b, in order to kill the terms involving S_T and Σ_T. The new time-depending parameters are given, for T large enough, by

$$\alpha_T = 2 + (a-2)\sqrt{1 - 8\,v_{1,T}/\left(a-2\right)^2} \quad \text{and} \quad \beta_T = b\sqrt{1 - 8\,v_{2,T}/b^2}.$$

Thus, (B.8) becomes

$$E_T\left(u\right) = \exp\left(-w_{1,T}x_0 - w_{2,T}\log x_0 + T\mathcal{C}_T\right)\mathbb{E}^{\alpha_T,\beta_T}_{\gamma_T,\delta_T}\left(\exp\left(w_{1,T}X_T\right)X_T^{w_{2,T}}\right) \tag{B.9}$$

where $w_{1,T} = \left(b - \beta_T + 2v_{4,T}\right)/4$, $w_{2,T} = \left(a - \alpha_T + 2v_{3,T}\right)/4$ and

$$4\mathcal{C}_T = -2av_{4,T} - 2bv_{3,T} + \beta_T\alpha_T - ab - (c-\gamma_T)(d-\delta_T)/(1-\rho^2).$$

Therefore, taking the logarithm of (B.9), we obtain that

$$\log E_T\left(u\right) = -w_{1,T}x_0 - w_{2,T}\log x_0 + T\mathcal{C}_T + \log\mathbb{E}^{\alpha_T,\beta_T}_{\gamma_T,\delta_T}\left(\exp\left(w_{1,T}X_T\right)X_T^{w_{2,T}}\right). \tag{B.10}$$

We now have to divide (B.10) by λ_T and investigate the limit for T going to infinity. We consider each term of the right-hand side of (B.10) separately. First of all, as $w_{1,T}$ and $w_{2,T}$ tend to zero for T going to infinity, we immediately deduce that

$$\lim_{T\to+\infty}\lambda_T^{-1}\left(-w_{1,T}x_0 - w_{2,T}\log x_0\right) = 0. \tag{B.11}$$

We now consider the term $T\mathcal{C}_T/\lambda_T$. On the one hand, replacing γ_T and δ_T by their respective definitions, we easily obtain that

$$\frac{1}{4\left(1-\rho^2\right)}\left(c-\gamma_T\right)\left(d-\delta_T\right)\frac{T}{\lambda_T} = \left(1-\rho^2\right)u_3u_4. \tag{B.12}$$

On the other hand, as λ_T/T goes to zero for T tending to infinity, we expand $\beta_T\alpha_T$ up to order two in $\sqrt{\lambda_T/T}$ and obtain that:

$$\begin{aligned}\tfrac{T}{4\lambda_T}\left(-2av_{4,T} - 2bv_{3,T} + \beta_T\alpha_T - ab\right) = \;& -\frac{b}{2\left(a-2\right)}\left(\left(u_1+\rho u_3\right)^2 + \left(1-\rho^2\right)u_3^2\right)\\ & -\frac{a}{2b}\left(\left(u_2+\rho u_4\right)^2 + \left(1-\rho^2\right)u_4^2\right)\\ & +\left(u_1+\rho u_3\right)\left(u_2+\rho u_4\right) + o\left(1\right).\end{aligned} \tag{B.13}$$

Thus, for T going to infinity, the limit value of $T\mathcal{C}_T/\lambda_T$ is the sum of the limits of (B.12) and (B.13). Before concluding, we will now show that

$$\lim_{T\to+\infty}\lambda_T^{-1}\log\mathbb{E}^{\alpha_T,\beta_T}_{\gamma_T,\delta_T}\left(\exp\left(w_{1,T}X_T\right)X_T^{w_{2,T}}\right) = 0 \tag{B.14}$$

The density function of the solution X_T associated with parameters α_T and β_T and initial point x_0 is given, for any positive real y, by

$$f(y) = K_T\exp\left(-y/2\right)y^{(\alpha_T-2)/4}I_{(\alpha_T-2)/2}\left(\sqrt{y\xi_T}\right)$$

where I_ν is the modified Bessel function of the first kind and ξ_T and K_T are two constants respectively given by

$$\xi_T = -\frac{x_0\beta_T}{e^{-\beta_T T}-1} \qquad \text{and} \qquad K_T = \frac{e^{-\xi_T/2}}{2\xi_T^{(\alpha_T-2)/4}},$$

see for instance [58]. Thus, using formulas 6.643(2) and 9.220(2) of [44], we compute the expectation in the last term of (B.10) as follows

$$\begin{aligned}
\mathbb{E}^{\alpha_T,\beta_T}_{\gamma_T,\delta_T}\left(\exp\left(w_{1,T}X_T\right)X_T^{w_{2,T}}\right) &= \int_0^{+\infty}\exp\left(w_{1,T}y\right)y^{w_{2,T}}f(y)\,\mathrm{d}y \\
&= \frac{\Gamma(w_{2,T}+\alpha_T/2)}{\Gamma(\alpha_T/2)}\,2^{(\alpha_T+2)/4}\left(1/2-w_{1,T}\right)^{-(2w_{2,T}+\alpha_T-1)/2} \\
&\quad \times e^{-\xi_T/2}\,{}_1F_1\left(w_{2,T}+\alpha_T/2,\alpha_T/2,\xi_T/(2-4w_{1,T})\right)
\end{aligned}$$

where ${}_1F_1$ is the degenerate hypergeometric function (see [44]). As we want to compute the limit of the logarithm of this expectation, the obtained product becomes a sum and we investigate the limit of each term separately. For T going to infinity, α_T converges to a and β_T to b whereas ξ_T, $w_{1,T}$ and $w_{2,T}$ vanish. Thus, we obtain the four following limits

$$\lim_{T\to+\infty}\lambda_T^{-1}\log\frac{\Gamma(w_{2,T}+\alpha_T/2)}{\Gamma(\alpha_T/2)} = \lim_{T\to+\infty}\lambda_T^{-1}\log\frac{\Gamma(a/2)}{\Gamma(a/2)} = 0,$$

$$\lim_{T\to+\infty}\lambda_T^{-1}\frac{\alpha_T+2}{4}\log 2 = 0,$$

$$\lim_{T\to+\infty}\lambda_T^{-1}(w_{2,T}+\alpha_T/2-1/2)\log\left(1/2-w_{1,T}\right) = \lim_{T\to+\infty}\lambda_T^{-1}(a/2-1/2)\log(1/2) = 0,$$

$$\lim_{T\to+\infty}\lambda_T^{-1}\xi_T/2 = 0.$$

Furthermore,

$$\lim_{T\to+\infty}{}_1F_1\left(w_{2,T}+\alpha_T/2,\alpha_T/2,\xi_T/(2-4w_{1,T})\right) = {}_1F_1\left(a/2,a/2,0\right) = 1$$

which, combined with the previous limits, leads to (B.14). We conclude from the conjunction of (B.11), (B.12), (B.13) and (B.14), that

$$\begin{aligned}
\lim_{T\to\infty}\lambda_T^{-1}\log E_T(u) &= -\frac{b}{2(a-2)}\left(u_1^2+2\rho u_1u_3+u_3^2\right)-\frac{a}{2b}\left(u_2^2+2\rho u_2u_4+u_4^2\right) \\
&\quad +u_1u_2+u_3u_4+\rho\left(u_1u_4+u_2u_3\right) \\
&= \frac{1}{2}u^\intercal\Gamma u,
\end{aligned} \tag{B.15}$$

with the matix Γ given by (3.16). This concludes the proof of (3.15).

Corollary B.1. *Let Λ_T be the normalized cumulant generating function of* $((\lambda_T T)^{-1/2} M_T)$ *given, for all $v \in \mathbb{R}^2$, by*

$$\Lambda_T(v) = \frac{1}{\lambda_T} \log \mathbb{E}\left[e^{\sqrt{\lambda_T/T}\langle v, M_T\rangle}\right].$$

Its pointwise limit Λ satisfies

$$\Lambda(v) = \lim_{T\to+\infty} \Lambda_T(v) = \frac{1}{2} v^\intercal \Sigma v, \tag{B.16}$$

where the matrix Σ was previously given in (3.5).

Proof. One can observe that

$$\Lambda_T(v) = \frac{1}{\lambda_T} E_T(v^\intercal, 0, 0),$$

where E_T is defined and computed in the previous proof. Taking $u_3 = u_4 = 0$ in (B.15), we obtain the announced result. □

Corollary B.2. *Let L_T be the normalized cumulant generating function of* $((\lambda_T T)^{-1/2} N_T)$ *given, for all $v \in \mathbb{R}^2$, by*

$$L_T(v) = \frac{1}{\lambda_T} \log \mathbb{E}\left[e^{\sqrt{\lambda_T/T}\langle v, N_T\rangle}\right].$$

Its pointwise limit L satisfies

$$L(v) = \lim_{T\to+\infty} L_T(v) = \frac{1}{2} v^\intercal \Sigma v, \tag{B.17}$$

where the matrix Σ was previously given in (3.5).

Proof. As for the previous corollary, one can observe that

$$L_T(v) = \frac{1}{\lambda_T} E_T(0, 0, v^\intercal),$$

Taking $u_1 = u_2 = 0$ in (B.15), we obtain the announced result. □

Appendix C: Proof of the exponential convergence for S_T and Σ_T

We start by showing that p^η given in (3.13) is stricly negative. We have established, in Lemma 3.1 of [27], that the couple (S_T, Σ_T) satisfies an LDP with speed T and good rate function I given for any $(x, y) \in \mathbb{R}^2$ by

$$I(x,y) = \begin{cases} \dfrac{y}{2(xy-1)} + \dfrac{b^2}{8}x + \dfrac{(a-2)^2}{8}y + \dfrac{ab}{4} & \text{if } x > 0, y > 0 \text{ and } xy - 1 > 0, \\ +\infty & \text{otherwise.} \end{cases} \tag{C.1}$$

We denote by $\mathcal{D}_I$ the domain of $\mathbb{R}^2$ where I is finite. Using the contraction principle recalled in Lemma 3.3.1, we show that the sequence $(\Sigma_T/V_T, S_T/V_T)$ satisfies an LDP with good rate function J given, for all $(z,t) \in \mathbb{R}^2$, by

$$J(z,t) = \inf\{I(x,y) | (x,y) \in \mathcal{D}_I, g(x,y) = (z,t)\}$$

where $g(x,y) = (y(xy-1)^{-1}, x(xy-1)^{-1})$, and the infimum over the empty set is equal to infinity. One can observe that, if $z \leqslant 0$ or $t \leqslant 0$, (x,y) is not inside $\mathcal{D}_I$, which means that the finitude domain $\mathcal{D}_J$ of J is included inside $\mathbb{R}_*^+ \times \mathbb{R}_*^+$. For any $z > 0$ and $t > 0$, the condition $(z,t) = g(x,y)$ leads to

$$y = z(xy-1) \qquad \text{and} \qquad x = tz^{-1}y. \tag{C.2}$$

Combining both, we obtain that y satisfies $ty^2 - y - z = 0$. Only one solution is strictly positive, it is given by

$$y^* = \frac{1+\sqrt{1+4tz}}{2t}. \tag{C.3}$$

We deduce from the second equality in (C.2) that the only possible value for x is given by

$$x^* = \frac{1+\sqrt{1+4tz}}{2z}. \tag{C.4}$$

Replacing (C.3) and (C.4) into (C.1), we obtain that for any $(z,t) \in \mathbb{R}_*^+ \times \mathbb{R}_*^+$,

$$J(z,t) = I(x^*, y^*) = \frac{z}{2} + \left(\frac{b^2}{16z} + \frac{(a-2)^2}{16t}\right)\left(1+\sqrt{1+4tz}\right) + \frac{ab}{4}. \tag{C.5}$$

The function I is positive and only vanishes at point $(-ab^{-1}, -b(a-2)^{-1})$, see [27]. Thus, (C.5) implies that J is positive and vanishes for $(x^*, y^*) = (-ab^{-1}, -b(a-2)^{-1})$. As, (x^*, y^*) satisfy $(z,t) = g(x^*, y^*)$, we obtain that J only vanishes at point

$$(z_0, t_0) = \left(\frac{x^*}{x^*y^*-1}, \frac{y^*}{x^*y^*-1}\right) = \left(-\frac{a(a-2)}{2b}, -\frac{b}{2}\right).$$

Besides, applying the contraction principle again, we show that p^η, given in (3.13), satisfies the following upper bound

$$p^\eta \leqslant -\inf\left\{J(z,t) \mid (z,t) \notin \mathcal{B}\left((z_0,t_0),\eta^2\right)\right\}. \tag{C.6}$$

We want to prove that p^η is strictly negative. One can observe that the function J is coercive. Indeed, let K be the compact subset of $(\mathbb{R}_+^*)^2$ given by $K = [\varepsilon, A] \times [\xi, B]$ with

$$\varepsilon = \frac{b^2}{4(4-ab)}, \quad A = \frac{4-ab}{2}, \quad \xi = \frac{(a-2)^2}{4(4-ab)} \quad \text{and} \quad B = \frac{4(4-ab)^5}{b^6}, \tag{C.7}$$

we easily show that

$$\forall (z,t) \notin K, \quad J(z,t) \geqslant 1.$$

Consequently, the infimum of J over $(\mathbb{R}_+^*)^2 \backslash \mathcal{B}\left((z_0, t_0), \eta^2\right)$ reduces to the infimum over the compact subset $K \backslash \mathcal{B}\left((z_0, t_0), \eta^2\right)$. As J is a continuous function, this infimum is reached for some (z^*, t^*). As J is positive and only vanishes at point (z_0, t_0) which does not belong to $K \backslash \mathcal{B}\left((z_0, t_0), \eta^2\right)$, we can conclude that

$$\inf \left\{ J(z,t) \mid (z,t) \notin \mathcal{B}\left((z_0, t_0), \eta^2\right) \right\} = J(z^*, t^*) > 0,$$

which is exactly the result that we wanted to prove.

We now consider q_T^η. We have established, in Theorem 3.2 of [27], an LDP for the sequence (V_T) with speed T and good rate function K given for all $x \in \mathbb{R}$ by

$$K(x) = \begin{cases} -\dfrac{b}{4}\sqrt{(x+1)\left((a-2)^2 + \dfrac{4}{x}\right)} + \dfrac{ab}{4} & \text{if } x > 0 \\ +\infty & \text{if } x \leqslant 0. \end{cases}$$

Thus, by the contraction principle, the sequence (V_T^{-1}) satisfies an LDP with speed T and good rate function $\widetilde{K}$ which is infinite over $\mathbb{R}^-$ and satisfies for any $x > 0$, $\widetilde{K}(x) = K(x^{-1})$. Thus, we easily deduce that

$$q^\eta \leqslant -\inf \left\{ \widetilde{K}(x); \left| \frac{a-2}{2} - x \right| \geqslant \frac{\eta}{2} \right\} = -\min \left\{ \widetilde{K}\left(\frac{a-2+\eta}{2} \right); \widetilde{K}\left(\frac{a-2-\eta}{2} \right) \right\}$$

which leads to $q^\eta < 0$, as $\widetilde{K}$ is strictly positive over $\mathbb{R} \backslash \left\{ \frac{a-2}{2} \right\}$.

Chapitre 4

Estimateur des moindres carrés pondérés pour le processus de Heston

RÉSUMÉ. On estime simultanément les quatre paramètres d'un processus de Heston sous-critique. On ne se restreint pas au cas où le processus de volatilité stochastique n'atteint jamais zéro. Afin d'éviter le recours à des temps d'arrêt et l'usage d'un estimateur naturel mais dont la manipulation est peu aisée en pratique, on propose un nouvel estimateur des moindres carrés pondérés. On établit sa consistance forte ainsi que sa normalité asymptotique. Enfin, on effectue quelques simulations numériques pour illustrer les bonnes performances de cet estimateur.

ABSTRACT. We simultaneously estimate the four parameters of a subcritical Heston process. We do not restrict ourselves to the case where the stochastic volatility process never reaches zero. In order to avoid the use of unmanageable stopping times and natural but intractable estimator, we propose to make use of a weighted least squares estimator. We establish strong consistency and asymptotic normality for this estimator. Numerical simulations are also provided, illustrating the good performances of our estimation procedure.

Sommaire

4.1 Introduction

Introduced in 1973, as hedging tool, the Black-Scholes model uses a geometric Brownian motion to represent asset prices. The implied volatility is supposed to be constant over time, which turned out to be inaccurate to fit real market data, especially during the crash in 1987, see [69]. Several alternative models have been constructed to take into account the so-called smile effect associated to deep in-the-money or out-of-money options. A particular attention has been drawn to the study of stochastic volatility processes in which the volatility is also given as a solution of some stochastic differential equation, see [68], [61] and [39] for financial accuracy. Among them, Heston process [47] is one of the most popular, due to its computational tractability. For example, [60] easily computes call option prices using Fourier inversion techniques. Numerous results about the asymptotic volatility smile can be found in the very recent literature: see e.g. [35], [36], [49].

We denote by Y_t the logarithm of the price of a given asset and by X_t its instantaneous variance, and we consider the following Heston process

$$\begin{cases} \mathrm{d}X_t = (a + bX_t)\,\mathrm{d}t + 2\sqrt{X_t}\,\mathrm{d}B_t \\ \mathrm{d}Y_t = (\alpha + \beta X_t)\,\mathrm{d}t + 2\sqrt{X_t}\left(\rho\,\mathrm{d}B_t + \sqrt{1-\rho^2}\,\mathrm{d}W_t\right) \end{cases} \tag{4.1}$$

with $a > 0$, $(b, \alpha, \beta) \in \mathbb{R}^3$ and $\rho \in]-1, 1[$, where (B_t, W_t) is a 2-dimensional standard Wiener process and the initial state $(x_0, y_0) \in \mathbb{R}^+ \times \mathbb{R}$. In this process, the volatility X_t is driven by a generalized squared radial Ornstein-Uhlenbeck process, also known as the CIR process, firstly studied by Feller [34] and introduced in a financial context by Cox, Ingersoll and Ross [22] to compute short-term interest rates. The asymptotic behavior of this process has been widely investigated and depends on the values of both coefficients a and b.

Once a model has been chosen for its realistic features, it needs to be calibrated before being used for pricing. Our goal in this paper is to estimate parameters (a, b, α, β) at the same time using a trajectory of (X_t) and (Y_t) over the time interval $[0, T]$. Azencott and Gadhyan [6] developed an algorithm to estimate some parameters of the Heston process based on discrete time observations, by making use of Euler and Milstein discretization schemes for the maximum likelihood. However, in the special case of an Heston process, the exact likelihood can be computed. It allows us to construct the maximum likelihood estimator (MLE) without using sophisticated approximation methods, which is necessary for many stochastic volatility models, see [1]. The MLE of (a, b, α, β) has been recently investigated in [11], together with its asymptotic behavior in the special case where $a \geqslant 2$. Denote by τ_0 the stopping-time given by

$$\tau_0 = \inf\left\{T > 0 \,\middle|\, \int_0^T X_t^{-1}\,\mathrm{d}t = \infty\right\}. \tag{4.2}$$

For any $a > 0$, the MLE $\widetilde{\theta}_T = \left(\widetilde{a}_T, \widetilde{b}_T, \widetilde{\alpha}_T, \widetilde{\beta}_T\right)$ is given, for $T < \tau_0$, by:

$$\widetilde{\theta}_T = \begin{pmatrix} G_T^{-1} & 0 \\ 0 & G_T^{-1} \end{pmatrix} \begin{pmatrix} \widetilde{U}_T \\ \widetilde{V}_T \end{pmatrix} \tag{4.3}$$

where $\widetilde{U}_T = \left(\int_0^T X_t^{-1}\,\mathrm{d}X_t, \int_0^T \mathrm{d}X_t\right)^\intercal$, $\widetilde{V}_T = \left(\int_0^T X_t^{-1}\,\mathrm{d}Y_t, \int_0^T \mathrm{d}Y_t\right)^\intercal$ and

$$G_T = \begin{pmatrix} \int_0^T X_t^{-1}\,\mathrm{d}t & T \\ T & \int_0^T X_t\,\mathrm{d}t \end{pmatrix}.$$

One can observe that $(\widetilde{a}_T, \widetilde{b}_T)$ coincides with the MLE of the parameters (a, b) of the CIR process based on the observation of (X_T) over the time interval $[0, T]$. The asymptotic behavior of this latter estimator is well-known, see for example [37], [65] and [14]. In the supercritical case $b > 0$, Overbeck [65] has shown that $\widetilde{b}_T$ converges a.s. to b whereas there exists no consistent estimator for a. Hence, we will focus our attention on the geometrically ergodic case $b < 0$. Furthermore, the value of a governs the behavior at zero of (X_T): for $a \geqslant 2$, the process almost surely never reaches zero , whereas for $0 < a < 2$, zero is quite frequently visited and

$$\mathbb{P}\left(\tau_0 < \infty\right) = 1, \tag{4.4}$$

see for instance [58] or [65]. For $a > 2$, the MLE converges a.s. to $\theta = (a, b, \alpha, \beta)$ and satisfies the following Central Limit Theorem (CLT)

$$\sqrt{T}\left(\widetilde{\theta}_T - \theta\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 4R \otimes \Sigma^{-1})$$

where the matrix R and Σ are respectively given by

$$R = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \qquad \text{and} \qquad \Sigma = \begin{pmatrix} \frac{-b}{a-2} & 1 \\ 1 & -\frac{a}{b} \end{pmatrix},$$

and $\otimes$ stands for the Kronecker product.

A large deviation principle for the couple $(\widetilde{a}_T, \widetilde{b}_T)$ was recently established in [27]. In the particular case where one parameter is known and the other one is estimated, large deviations can be found in [73], while moderate deviations are given in [38].

By contrast, in the case where $0 < a < 2$, (4.4) implies the non-integrability of X_T^{-1} for large values of T so that the MLE does not converge for T going to infinity. Consequently, this case has been less investigated even though it is often of interest in finance, to compute long dated interest rates for instance, as explained in [4], or in FX-markets, see [53]. In the case of the CIR process, Overbeck [65] used accurate stopping times to build a strongly consistent estimator based on the MLE:

$$\mathbb{1}_{T<\tau_0}\begin{pmatrix} \widetilde{a}_T \\ \widetilde{b}_T \end{pmatrix} + \mathbb{1}_{\tau_0 \leqslant T}\begin{pmatrix} \lim\limits_{t\uparrow\tau_0} S_t \Sigma_t^{-1} \\ \left(\int_0^T X_s\,\mathrm{d}s\right)^{-1}\left(X_T - T\lim\limits_{t\uparrow\tau_0} S_t\Sigma_t^{-1}\right) \end{pmatrix} \tag{4.5}$$

where $S_t = \int_0^t X_s^{-1}\,\mathrm{d}X_s$, $\Sigma_t = \int_0^t X_s^{-1}\,\mathrm{d}s$ and τ_0 is given by (4.2). The aim of this paper is to investigate a new strongly consistent weighted least squares estimator (WLSE) for the quadruplet of parameters θ (and for (a, b) as a consequence). The weighting allows us to circumvent the explosion for X_T reaching zero and consequently avoid us to make use of stopping times, which are not easy to handle in practice. It generalizes to continuous time the original work of Wei and Winnicki [70] for branching processes with immigration, inspired by an analogy with first order autoregressive processes. Our results answer, by the way, the question of Ben Alaya and Kebaier in the conclusion of [14] regarding the CIR process.

Following the seminal work of [70], denote $C_T = X_T + c$ where c is some positive constant. Our new couple of weighted least squares estimator is given by

$$\widehat{\theta}_T = \begin{pmatrix} \Gamma_T^{-1} & 0 \\ 0 & \Gamma_T^{-1} \end{pmatrix} \begin{pmatrix} U_T \\ V_T \end{pmatrix} \tag{4.6}$$

where $U_T = \left(\int_0^T \frac{1}{C_t} \, \mathrm{d}X_t, \int_0^T \frac{X_t}{C_t} \, \mathrm{d}X_t \right)^\intercal$, $V_T = \left(\int_0^T \frac{1}{C_t} \, \mathrm{d}Y_t, \int_0^T \frac{X_t}{C_t} \, \mathrm{d}Y_t \right)^\intercal$ and

$$\Gamma_T = \begin{pmatrix} \int_0^T \frac{1}{C_t} \, \mathrm{d}t & \int_0^T \frac{X_t}{C_t} \, \mathrm{d}t \\ \int_0^T \frac{X_t}{C_t} \, \mathrm{d}t & \int_0^T \frac{X_t^2}{C_t} \, \mathrm{d}t \end{pmatrix}.$$

We do not restrict ourselves to the case where $c = 1$ as it may lower sometimes the variance of the estimators. In the particular case where $c = 0$, one can observe that the new estimator coincides with the MLE.

The paper is organized as follows. The second section contains our main results: the strong consistency of this new couple of estimators as well as its asymptotic normality. The third section deals with a comparison with the MLE, while the remaining of the paper is devoted to the proofs of our main results, as well as their illustration by some numerical simulations.

4.2 Main results

Our main results are as follows.

Theorem 4.2.1. *Assume that $a > 0$ and $b < 0$. Then, the four-dimensional WLSE $\widehat{\theta}_T$ is strongly consistent: for T going to infinity,*

$$\widehat{\theta}_T \xrightarrow{a.s.} \theta. \tag{4.7}$$

For T going to infinity, X_T converges in distribution to a random variable X with Gamma $\Gamma(a/2, -b/2)$ distribution, see Lemma 3 of [65] for instance. Additionally, we denote by C the limiting distribution of $X_T + c$, as T goes to infinity.

Theorem 4.2.2. *Assume that $a > 0$ and $b < 0$. Then, for T going to infinity, the estimator $\widehat{\theta}_T$ satisfies the following CLT*

$$\sqrt{T} \left(\widehat{\theta}_T - \theta \right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, 4\Lambda\right), \tag{4.8}$$

where the asymptotic variance Λ is defined as a block matrix by

$$\Lambda = \begin{pmatrix} ALA & \rho ALA \\ \rho ALA & ALA \end{pmatrix}, \tag{4.9}$$

with the matrix A and L respectively given by

$$A = \left(\mathbb{E}\left[C\right]\mathbb{E}\left[1/C\right] - 1\right)^{-1}\begin{pmatrix}\mathbb{E}\left[X^2/C\right] & -\mathbb{E}\left[X/C\right]\\ -\mathbb{E}\left[X/C\right] & \mathbb{E}\left[1/C\right]\end{pmatrix}$$

and

$$L = \begin{pmatrix}\mathbb{E}\left[X/C^2\right] & \mathbb{E}\left[X^2/C^2\right]\\ \mathbb{E}\left[X^2/C^2\right] & \mathbb{E}\left[X^3/C^2\right]\end{pmatrix}.$$

We deduce from the previous theorems the following result for the MLE of the two parameters of the CIR process (X_T).

Corollary 4.2.1. *Assuming that $a > 0$ and $b < 0$, the WLSE $(\widehat{a}_T, \widehat{b}_T)$ of parameters (a, b) is strongly consistent for T going to infinity,*

$$\begin{pmatrix}\widehat{a}_T\\ \widehat{b}_T\end{pmatrix} \xrightarrow{a.s.} \begin{pmatrix}a\\ b\end{pmatrix}.$$

and satisfies the following CLT

$$\sqrt{T}\begin{pmatrix}\widehat{a}_T - a\\ \widehat{b}_T - b\end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, 4ALA\right).$$

Remark 4.2.1. *In the remaining of this paper, we denote*

$$\psi_c = \left(-\frac{bc}{2}\right)^{a/2} e^{-bc/2}\,\Gamma\left(1 - a/2, -bc/2\right), \tag{4.10}$$

where Γ is the upper incomplete gamma function defined for all $y \in \mathbb{R}$ and $\alpha \in \mathbb{R}_*^+$ by

$$\Gamma(\alpha, y) = \int_y^{+\infty} e^{-t} t^{\alpha-1}\,\mathrm{d}t,$$

and extended, for $y \neq 0$, to any real α by holomorphy. To simplify the following expressions, we also define

$$\varphi_c = \psi_c\left(1 - \frac{a}{bc}\right) - 1. \tag{4.11}$$

In the proof of Theorem 4.2.2, we evaluate the two matrices A and L involving c and we obtain that

$$A = \varphi_c^{-1}\begin{pmatrix}c\left(\psi_c - 1\right) - \frac{a}{b} & \psi_c - 1\\ \psi_c - 1 & \frac{\psi_c}{c}\end{pmatrix} \tag{4.12}$$

and

$$L = \frac{1}{2}\begin{pmatrix}\frac{a}{c}\psi_c + b\left(1 - \psi_c\right) & \left(a + 2 - bc\right)\left(1 - \psi_c\right) - a\\ \left(a + 2 - bc\right)\left(1 - \psi_c\right) - a & c\left(\psi_c - 1\right)\left(a + 4 - bc\right) + ac - \frac{2a}{b}\end{pmatrix}. \tag{4.13}$$

By a straightforward computation, we deduce that $ALA = \left(\varphi_c\right)^{-2}\begin{pmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{12} & \sigma_{22}\end{pmatrix}$ where the variances σ_{11} and σ_{22} are respectively given by

$$\sigma_{11} = \frac{a}{b}\left(\psi_c - 1\right)^2 - \frac{a^2}{2b}\varphi_c, \quad \sigma_{22} = \varphi_c\left(\frac{\psi_c}{c} - \frac{b}{2}\right) + \frac{\psi_c}{c}\left(\psi_c - 1\right),$$

and the covariance σ_{12} is given by $\sigma_{12} = \left(\psi_c - 1\right)^2 - \frac{a}{2}\varphi_c$.

Remark 4.2.2. *For c going to zero (for which we need a to be greater than 2) , we obtain the same covariance matrix than for the MLE. Indeed, using well-known asymptotic results about the incomplete Gamma function Γ, which could be found in [62], we have that, as soon as $a > 2$,*

$$\Gamma\left(1 - a/2, -bc/2\right)\left(-\frac{bc}{2}\right)^{a/2-1} \underset{c\to 0}{\longrightarrow} \frac{-1}{1-a/2} = \frac{2}{a-2}. \tag{4.14}$$

Thus ψ_c goes to zero for c tending to zero, $\frac{\psi_c}{c}$ converges to $\frac{-b}{a-2}$ and φ_c tends to $\frac{2}{a-2}$. Hence, we easily obtain that, for c going to zero,

$$\sigma_{11} \underset{c\to 0}{\longrightarrow} -\frac{2a}{b(a-2)}, \quad \sigma_{22} \underset{c\to 0}{\longrightarrow} -\frac{2b}{(a-2)^2}, \quad \textit{and} \quad \sigma_{12} \underset{c\to 0}{\longrightarrow} -\frac{2}{a-2},$$

where σ_{11}, σ_{22} and σ_{12} are defined in Remark 4.2.1. This leads to

$$ALA \underset{c\to 0}{\longrightarrow} \Sigma^{-1} \qquad \textit{where} \qquad \Sigma = \begin{pmatrix} \frac{-b}{a-2} & 1 \\ 1 & -\frac{a}{b} \end{pmatrix}.$$

4.3 Asymptotic variance

Even though we considered the weighted least squares estimators in order to investigate the case $0 < a < 2$ for which the MLE is not consistent, it is interesting to compare the asymptotic variances in the CLT of this new estimators and of the MLE, in the case where $a > 2$. This comparison requires a lot of technical calculation as the asymptotic variances depends on the value of a, b and c. However, it is quite easy to compare variances for the MLE of the parameters of the CIR process in the case where we suppose one of the parameter to be known and we estimate the other one, as it simplifies substantially the expression of the estimators. On the one hand, if a is known, the MLE for b is given by

$$\breve{b}_T = \frac{X_T - x_0 - aT}{\int_0^T X_t \,\mathrm{d}t} \tag{4.15}$$

and satisfies the following CLT

$$\sqrt{T}\left(\breve{b}_T - b\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 4/\mathbb{E}\left[X\right])$$

where $\mathbb{E}\left[X\right] = -a/b$, see for instance [13]. On the other hand, if b is known, the MLE of a is given by

$$\breve{a}_T = \frac{\int_0^T 1/X_t \,\mathrm{d}X_t - bT}{\int_0^T 1/X_t \,\mathrm{d}t} \tag{4.16}$$

and satisfies the following CLT

$$\sqrt{T}\left(\breve{a}_T - a\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 4/\mathbb{E}\left[X^{-1}\right])$$

with $\mathbb{E}\left[X^{-1}\right] = -b/(a-2)$. Whereas, the weighted least squares estimators are respectively given by

$$\widehat{b}_T = \frac{\int_0^T \frac{X_t}{C_t}\,\mathrm{d}X_t - a\int_0^T \frac{X_t}{C_t}\,\mathrm{d}t}{\int_0^T \frac{X_t^2}{C_t}\,\mathrm{d}t} \qquad \text{and} \qquad \widehat{a}_T = \frac{\int_0^T \frac{1}{C_t}\,\mathrm{d}X_t - b\int_0^T \frac{X_t}{C_t}\,\mathrm{d}t}{\int_0^T \frac{1}{C_t}\,\mathrm{d}t}.$$

Proposition 4.3.1. *Assume that $a > 0$ and $b < 0$. For T going to infinity, $\widehat{b}_T$ satisfies the following CLT:*

$$\sqrt{T}\left(\widehat{b}_T - b\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, 4\,\mathbb{E}\left[X^3/C^2\right]\left(\mathbb{E}\left[X^2/C\right]\right)^{-2}\right). \tag{4.17}$$

Proof. Replacing $\mathrm{d}X_t$ by its expression (4.1), we easily get that

$$\sqrt{T}\left(\widehat{b}_T - b\right) = 2\left(\frac{1}{T}\int_0^T \frac{X_t^2}{C_t}\,\mathrm{d}t\right)^{-1}\frac{n_T}{\sqrt{T}}$$

where n_T is a martingale given by

$$n_T = \int_0^T \frac{X_t\sqrt{X_t}}{C_t}\,\mathrm{d}B_t \qquad \text{and} \qquad \langle n\rangle_T = \int_0^T \frac{X_t^3}{C_t^2}\,\mathrm{d}t.$$

Using the ergodicity of the process, we obtain for T going to infinity

$$\frac{\langle n\rangle_T}{T} \xrightarrow{a.s.} \mathbb{E}\left[X^3/C^2\right]. \tag{4.18}$$

Thus, by the CLT for martingales, we obtain the following convergence in distribution

$$\frac{n_T}{\sqrt{T}} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \mathbb{E}\left[X^3/C^2\right]\right). \tag{4.19}$$

Consequently, (4.17) follows from (4.19), Slutsky's lemma and the fact that, by the ergodicity of the process, $\frac{1}{T}\int_0^T X_t^2/C_t\,\mathrm{d}t$ converges a.s. to $\mathbb{E}\left[X^2/C\right]$ for T going to infinity. ☐

Proposition 4.3.2. *Assume that $a > 0$ and $b < 0$. For T going to infinity, $\widehat{a}_T$ satisfies the following CLT:*

$$\sqrt{T}\left(\widehat{a}_T - a\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, 4\,\mathbb{E}\left[X/C^2\right]\left(\mathbb{E}\left[1/C\right]\right)^{-2}\right). \tag{4.20}$$

Proof. It works as in the previous proof. One can observe that

$$\sqrt{T}\left(\widehat{a}_T - a\right) = 2\left(\frac{1}{T}\int_0^T \frac{1}{C_t}\,\mathrm{d}t\right)^{-1}\frac{m_T}{\sqrt{T}}$$

where m_T is a martingale term given by

$$m_T = \int_0^T \frac{\sqrt{X_t}}{C_t}\,\mathrm{d}B_t \qquad \text{and} \qquad \langle m\rangle_T = \int_0^T \frac{X_t}{C_t^2}\,\mathrm{d}t.$$

Thus, for T going to infinity,

$$\frac{\langle m \rangle_T}{T} \xrightarrow{a.s.} \mathbb{E}\left[X/C^2\right] \tag{4.21}$$

which implies the following convergence in distribution

$$\frac{m_T}{\sqrt{T}} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \mathbb{E}\left[X/C^2\right]\right). \tag{4.22}$$

Finally, (4.22) leads to (4.20) thanks to the ergodicity of the process and Slutsky's lemma. □

Proposition 4.3.3. *Assume that $a > 2$ is known and $b < 0$. Then, the MLE of b satisfies a CLT with a smaller asymptotic variance than the weighted least squares estimator.*

Proof. Using Cauchy-Schwarz Inequality, we notice that

$$\left(\mathbb{E}\left[X^2/C\right]\right)^2 = \left(\mathbb{E}\left[\sqrt{X} \times X^{3/2}/C\right]\right)^2 \leqslant \mathbb{E}\left[X\right]\mathbb{E}\left[X^3/C^2\right]$$

which immediately leads to the result. □

Proposition 4.3.4. *Assume that $a > 2$ and $b < 0$ is known. Then, the MLE of a satisfies a CLT with a smaller asymptotic variance than the weighted least squares estimator.*

Proof. Using Cauchy-Schwarz Inequality, we notice that

$$\left(\mathbb{E}\left[1/C\right]\right)^2 = \left(\mathbb{E}\left[X^{-1/2} \times X^{1/2}/C\right]\right)^2 \leqslant \mathbb{E}\left[1/X\right]\mathbb{E}\left[X/C^2\right]$$

which immediately leads to the result. □

Remark 4.3.1. *Thus, the weighted least squares estimator is less efficient than the MLE in the case where this latter is easily manageable. This could seem to be contradictory to Remark 4.4 of [70] which deals with the discrete-time counterpart of the process. In fact, they compare the weighted least squares with the conditional least squares estimator which does not coincide with the MLE.*

Remark 4.3.2. *One could wonder how to choose which estimator to use, as the parameter a is unknown. However, we suppose that we observe the whole trajectory of the process over the time interval $[0, T]$. Thus, if we are able to detect some local time at level zero, we know that $a < 2$ and we should use the WLSE instead of the MLE.*

4.4 Technical Lemmas

First of all, we rewrite (4.6) using (4.1):

$$\widehat{\theta}_T = \theta + \begin{pmatrix} \Gamma_T^{-1} & 0 \\ 0 & \Gamma_T^{-1} \end{pmatrix} \begin{pmatrix} M_T \\ N_T \end{pmatrix}, \tag{4.23}$$

where M_T and N_T are martingales respectively given by

$$M_T = \begin{pmatrix} \displaystyle\int_0^T \frac{2\sqrt{X_t}}{C_t}\,\mathrm{d}B_t \\ \displaystyle\int_0^T \frac{2\sqrt{X_t}X_t}{C_t}\,\mathrm{d}B_t \end{pmatrix} \qquad \text{and} \qquad N_T = \begin{pmatrix} \displaystyle\int_0^T \frac{2\sqrt{X_t}}{C_t}\,\mathrm{d}\widetilde{B}_t \\ \displaystyle\int_0^T \frac{2\sqrt{X_t}X_t}{C_t}\,\mathrm{d}\widetilde{B}_t \end{pmatrix},$$

with $\mathrm{d}\widetilde{B}_t = \rho\,\mathrm{d}B_t + \sqrt{1-\rho^2}\,\mathrm{d}W_t$.

We denote by $\mathcal{M}_T$ the martingale $\mathcal{M}_T = (M_T, N_T)$. As $\left\langle \mathrm{d}B_t, \mathrm{d}\widetilde{B}_t \right\rangle = \rho \mathrm{d}t$, we easily obtain that the increasing process of $\mathcal{M}_T$ is given by

$$\langle \mathcal{M} \rangle_T = \begin{pmatrix} \langle M \rangle_T & \rho \langle M \rangle_T \\ \rho \langle M \rangle_T & \langle M \rangle_T \end{pmatrix}, \tag{4.24}$$

where the increasing process $\langle M \rangle_T$ of M_T is given by

$$\langle M \rangle_T = 4 \begin{pmatrix} \langle m \rangle_T & \displaystyle\int_0^T \frac{X_t^2}{C_t^2}\,\mathrm{d}t \\ \displaystyle\int_0^T \frac{X_t^2}{C_t^2}\,\mathrm{d}t & \langle n \rangle_T \end{pmatrix},$$

with $\langle m \rangle_T$ and $\langle n \rangle_T$ respectively given by (4.21) and (4.18).

In order to prove Theorems 4.2.1 and 4.2.2, we need to investigate the almost sure convergence of all the integrals involved in the definition of the estimators. Overbeck recalls in Lemma 3*(i)* of [65] that, for T going to infinity, X_T converges in distribution to X with Gamma $\Gamma(a/2, -b/2)$ distribution, whose probability density function is given by

$$f(x) = (\Gamma(a/2))^{-1} (-b/2)^{a/2} x^{a/2-1} e^{xb/2} \mathbf{1}_{x>0}. \tag{4.25}$$

Thus, by Lemma 3*(ii)* of [65], for T going to infinity,

$$\frac{1}{T}\int_0^T g(X_t)\,\mathrm{d}t \xrightarrow{a.s.} \mathbb{E}[g(X)] = \int_0^{+\infty} g(x) f(x)\,\mathrm{d}x.$$

for any function g such that the right-hand side exists.

We recall two properties of the incomplete gamma function that will be very useful in the following proof:

$$\Gamma(\alpha+1, x) = x^\alpha e^{-x} + \alpha \Gamma(\alpha, x) \tag{4.26}$$

and

$$\Gamma(\alpha+2, x) = x^\alpha e^{-x}(x+\alpha+1) + \alpha(\alpha+1)\Gamma(\alpha, x). \tag{4.27}$$

We are now able to prove the following lemma. The first three points give us the almost sure limit of the matrix $T\Gamma_T^{-1}$ as T goes to infinity, while the remaining deals with the increasing process of the four-dimensional martingale $\mathcal{M}_T$ given by (4.24).

Lemma 4.4.1. *With ψ_c given by* (4.10)*, we have that*

(i) $\mathbb{E}\left[1/C\right] = \frac{\psi_c}{c}$.

(ii) $\mathbb{E}\left[X/C\right] = 1 - \psi_c$.

(iii) $\mathbb{E}\left[X^2/C\right] = c\left(\psi_c - 1\right) - \frac{a}{b}$.

(iv) $\mathbb{E}\left[X/C^2\right] = \frac{a}{2c}\psi_c + \frac{b}{2}\left(1 - \psi_c\right)$.

(v) $\mathbb{E}\left[X^2/C^2\right] = \frac{1}{2}\left(\left(a + 2 - bc\right)\left(1 - \psi_c\right) - a\right)$.

(vi) $\mathbb{E}\left[X^3/C^2\right] = \frac{c}{2}\left(a + 4 - bc\right)\left(\psi_c - 1\right) + \frac{ac}{2} - \frac{a}{b}$.

Proof. (i) We have

$$\mathbb{E}\left[1/C\right] = \int_0^{+\infty} \frac{1}{x+c} f(x)\,\mathrm{d}x, \tag{4.28}$$

where f is given by (4.25). Formula 3.383(10) of [44] gives that

$$\int_0^{+\infty} \frac{1}{x+c} x^{a/2-1} e^{xb/2}\,\mathrm{d}x = c^{a/2-1} e^{-bc/2}\Gamma(a/2)\Gamma(1-a/2,-bc/2),$$

which leads to

$$\mathbb{E}\left[1/C\right] = \frac{1}{c}\left(-\frac{bc}{2}\right)^{a/2} e^{-bc/2}\,\Gamma\left(1-a/2,-bc/2\right)$$

and ensures the announced result.

(ii) As in the previous proof, we have

$$\mathbb{E}\left[X/C\right] = \int_0^{+\infty} \frac{x}{x+c} f(x)\,\mathrm{d}x. \tag{4.29}$$

By formula 3.383(10) of [44], we know that

$$\int_0^{+\infty} \frac{1}{x+c} x^{a/2} e^{xb/2}\,\mathrm{d}x = c^{a/2} e^{-bc/2}\Gamma(a/2+1)\Gamma(-a/2,-bc/2). \tag{4.30}$$

With formula (4.26), we easily obtain that

$$\Gamma(-a/2,-bc/2) = \left(-\frac{2}{a}\right)\left(\Gamma\left(1-a/2,-bc/2\right) - \left(-\frac{bc}{2}\right)^{-a/2} e^{bc/2}\right). \tag{4.31}$$

Combining (4.29), (4.30), (4.31) and the fact that $\Gamma(a/2+1) = a/2 \times \Gamma(a/2)$, we deduce the announced result.

(iii) We have

$$\mathbb{E}\left[\frac{X^2}{C}\right] = \mathbb{E}\left[\frac{(X+c-c)^2}{X+c}\right] = \mathbb{E}\left[X\right] - c + c^2\,\mathbb{E}\left[\frac{1}{C}\right],$$

and we conclude using (i) and the fact that $\mathbb{E}\left[X\right] = -a/b$.

(iv) By the very definition of f given by (4.25), we have

$$\mathbb{E}\left[X/C^2\right] = \int_0^{+\infty} \frac{x}{(x+c)^2} f(x)\,\mathrm{d}x = \frac{(-b/2)^{a/2}}{\Gamma(a/2)} \int_0^{+\infty} \frac{x^{a/2}}{(x+c)^2}\, e^{xb/2}\,\mathrm{d}x. \tag{4.32}$$

Integrating the right-hand side of (4.32) by part, we obtain that

$$\mathbb{E}\left[X/C^2\right] = \frac{(-b/2)^{a/2}}{\Gamma(a/2)} \left[\frac{a}{2}\int_0^{+\infty} \frac{x^{a/2-1}}{x+c}\, e^{xb/2}\,\mathrm{d}x + \frac{b}{2}\int_0^{+\infty} \frac{x^{a/2}}{x+c}\, e^{xb/2}\,\mathrm{d}x\right].$$

We have already computed both integrals in the proofs of respectively (i) and (ii), which leads to

$$\mathbb{E}\left[X/C^2\right] = \frac{a}{2}\,\mathbb{E}\left[1/C\right] + \frac{b}{2}\,\mathbb{E}\left[X/C\right] = \frac{a}{2c}\,\psi_c + \frac{b}{2}\,(1-\psi_c)\,. \tag{4.33}$$

(v) Integrating by parts and using (iii) and (iv),

$$\begin{aligned}
\mathbb{E}\left[X^2/C^2\right] &= \frac{(-b/2)^{a/2}}{\Gamma(a/2)} \int_0^{+\infty} \frac{x^{a/2+1}}{(x+c)^2}\, e^{xb/2}\,\mathrm{d}x \\
&= \frac{(-b/2)^{a/2}}{\Gamma(a/2)} \left[\frac{a+2}{2}\int_0^{+\infty} \frac{x^{a/2}}{x+c}\, e^{xb/2}\,\mathrm{d}x + \frac{b}{2}\int_0^{+\infty} \frac{x^{a/2+1}}{x+c}\, e^{xb/2}\,\mathrm{d}x\right] \\
&= \left((a+2)\;\mathbb{E}\left[X/C\right] + b\,\mathbb{E}\left[X^2/C\right]\right)/2 \\
&= \left((a+2)\,(1-\psi_c) + b\left(c\,(\psi_c-1) - \frac{a}{b}\right)\right)/2.
\end{aligned}$$

(vi) Noticing that $X^3 = X\,(X+c)^2 - 2cX^2 - c^2X$, we obtain that

$$\mathbb{E}\left[X^3/C^2\right] = \mathbb{E}\left[X\right] - 2c\,\mathbb{E}\left[X^2/C^2\right] - c^2\,\mathbb{E}\left[X/C^2\right]$$

and we conclude using (iv) and (v).

$\square$

4.5 Proof of the strong Consistency

We are now in the position to prove Theorem 4.2.1.

Proof of Theorem 4.2.1. First of all, we have

$$\frac{1}{T^2}\det\Gamma_T = \frac{1}{T}\int_0^T \frac{1}{C_t}\,\mathrm{d}t \times \frac{1}{T}\int_0^T \frac{X_t^2}{C_t}\,\mathrm{d}t - \left(\frac{1}{T}\int_0^T \frac{X_t}{C_t}\,\mathrm{d}t\right)^2.$$

Thus, as the process is ergodic, we obtain for T going to infinity,

$$\frac{1}{T^2}\det\Gamma_T \xrightarrow{a.s.} \mathbb{E}\left[1/C\right]\mathbb{E}\left[X^2/C\right] - \left(\mathbb{E}\left[X/C\right]\right)^2 \tag{4.34}$$

and

$$T\Gamma_T^{-1} \xrightarrow{a.s.} A \tag{4.35}$$

where A is given by

$$A = (\mathbb{E}[C]\,\mathbb{E}[1/C] - 1)^{-1} \begin{pmatrix} \mathbb{E}[X^2/C] & -\mathbb{E}[X/C] \\ -\mathbb{E}[X/C] & \mathbb{E}[1/C] \end{pmatrix}. \tag{4.36}$$

A straightforward application of Lemmas 4.1 (i) to (iii) gives that

$$A = \frac{1}{\psi_c\left(1 - \frac{a}{bc}\right) - 1} \begin{pmatrix} c(\psi_c - 1) - \frac{a}{b} & \psi_c - 1 \\ \psi_c - 1 & \frac{\psi_c}{c} \end{pmatrix}.$$

Besides, using the strong law of large numbers for martingale, we obtain that the martingale M_T satisfies for T going to infinity

$$\frac{M_T}{T} \xrightarrow{a.s.} 0. \tag{4.37}$$

As a matter of fact, by convergences (4.18) and (4.21), we know that a.s. $\langle n \rangle_T = \mathcal{O}(T)$ and $\langle m \rangle_T = \mathcal{O}(T)$. It ensures that for T going to infinity,

$$\frac{n_T}{T} \xrightarrow{a.s.} 0 \qquad \text{and} \qquad \frac{m_T}{T} \xrightarrow{a.s.} 0.$$

As N_T and M_T share the same increasing process, this result remains true by replacing M_T by N_T. Finally, the almost sure convergence (4.7) follows from (4.23), (4.35) and (4.37).

□

4.6 Proof of the asymptotic normality

Proof of Theorem 4.2.2. First of all, we deduce from (4.23) that

$$\sqrt{T}\left(\widehat{\theta}_T - \theta\right) = \begin{pmatrix} T\Gamma_T^{-1} & 0 \\ 0 & T\Gamma_T^{-1} \end{pmatrix} \begin{pmatrix} M_T/\sqrt{T} \\ N_T/\sqrt{T} \end{pmatrix}, \tag{4.38}$$

We already saw that $T\Gamma_T^{-1}$ converges almost surely as T goes to infinity and its limit A is given by (4.36). We now have to establish the asymptotic normality of $\frac{M_T}{\sqrt{T}}$. By the ergodicity of the process, we obtain that

$$\frac{\langle M \rangle_T}{T} \xrightarrow{a.s.} 4L \qquad \text{where} \qquad L = \begin{pmatrix} \mathbb{E}[X/C^2] & \mathbb{E}[X^2/C^2] \\ \mathbb{E}[X^2/C^2] & \mathbb{E}[X^3/C^2] \end{pmatrix}.$$

As a straightforward consequence of Lemmas 4.1 (iv) to (vi), we obtain that

$$L = \frac{1}{2}\begin{pmatrix} \frac{a}{c}\psi_c + b(1-\psi_c) & (a+2-bc)\,(1-\psi_c) - a \\ (a+2-bc)\,(1-\psi_c) - a & \psi_c c\,(a+4-b) - 4c - bc^2 - \frac{2a}{b} \end{pmatrix}.$$

We easily obtain the following a.s. convergence

$$\frac{\langle \mathcal{M} \rangle_T}{T} \xrightarrow{a.s.} 4\mathcal{L}$$

where $\mathcal{L}$ is a block matrix given by

$$\mathcal{L} = \begin{pmatrix} L & \rho L \\ \rho L & L \end{pmatrix}.$$

and we deduce from the CLT for martingales that

$$\frac{\mathcal{M}_T}{\sqrt{T}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 4\mathcal{L}), \tag{4.39}$$

Finally, the asymptotic normality (4.8) follows from (4.38) and (4.39) together with Slutsky's Lemma. $\square$

4.7 Numerical simulations

The efficient discretization of the CIR process is a challenging question, see for example [4] and [2]. We choose to implement the QE-algorithm based on quadratic-exponential approximations proposed in [4]. Andersen introduced this algorithm to deal with the case $a < 2$, for which common discretization schemes are not accurate.

4.7.1 Asymptotic behavior for $c = 1$

The two following figures illustrate our main results (strong consistency and asymptotic normality) in the case $a = 1$ and $b = -2$, with the weighting parameter $c = 1$. The red curves in the second figure displays the standard normal distribution. We denote by v_a (respestively v_b) the variance of $\widehat{a}_T$ (resp. $\widehat{b}_T$).

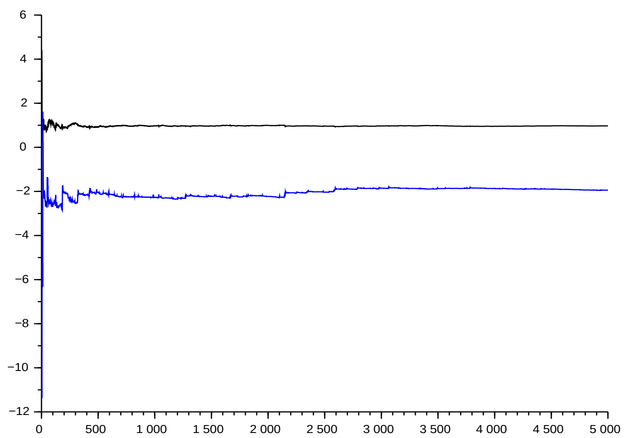


Figure 4.1 – Strong consistency: $(\widehat{a}_T)$ in black and $(\widehat{b}_T)$ in blue.

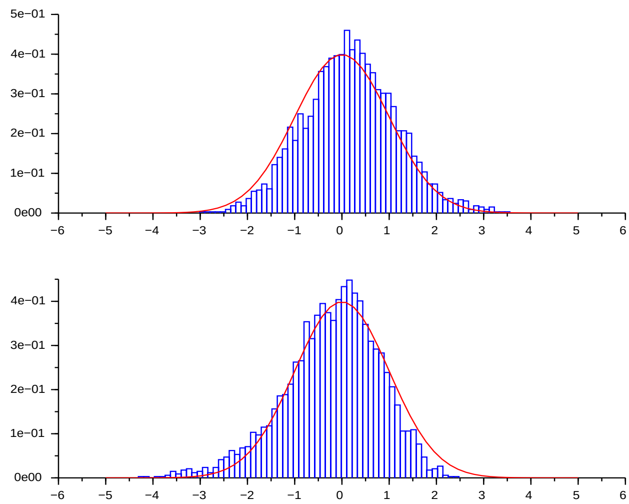
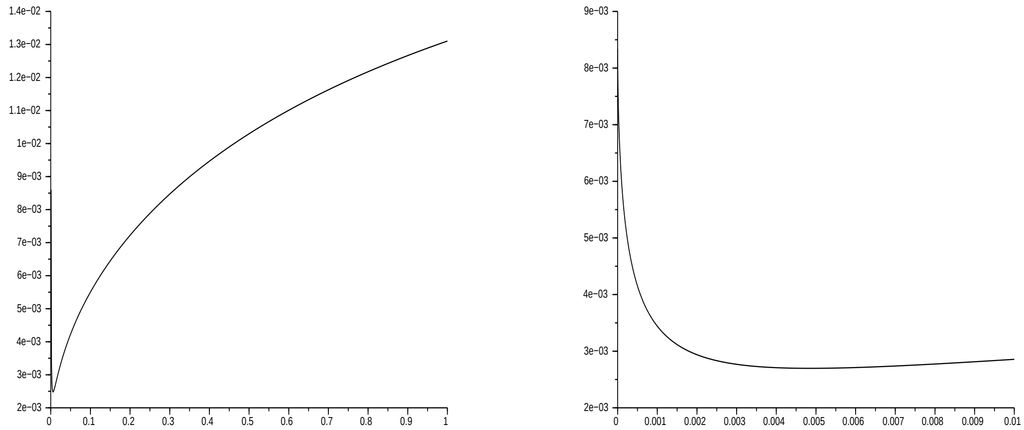


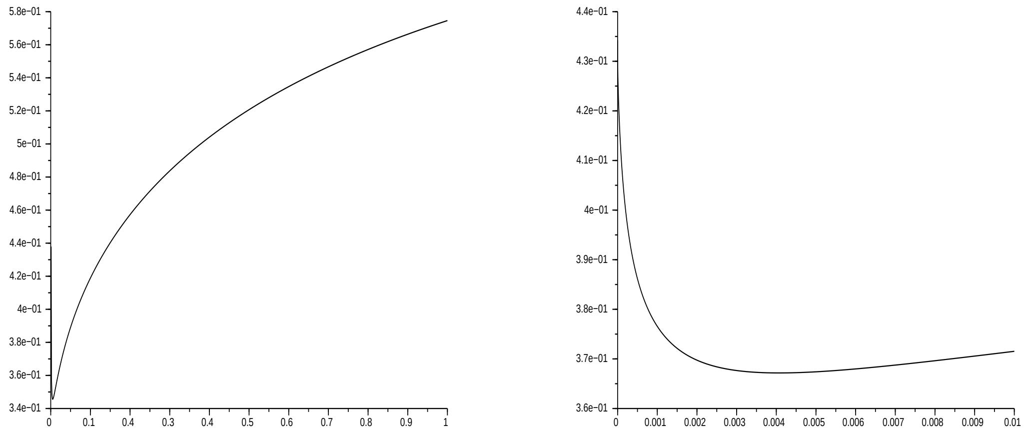
Figure 4.2 – Histograms of 3000 outcomes of $\sqrt{T/v_a}\,(\widehat{a}_T - a)$ and $\sqrt{T/v_b}(\widehat{b}_T - b)$ at time $T = 70$.

4.7.2 Choice of the constant c

We have chosen to introduce a constant c in our weighting, instead of only considering the case $c = 1$ (as done in the discrete-time case in [70]) with the aim of lowering the variance of the estimators. However, this raises the question of the optimal choice of the constant c, which depends on the values of parameters a and b. We set $a = 1$ and $b = -4$ and simulate 500 trajectories of the process over the time interval $[0, 50]$. We compute the empirical variance of the estimators given by each trajectory for c varying between 10^{-10} and 1. It appears that one should choose a small value of c. The value should not be to small to avoid the growth illustrated by the figures on the right-hand side, which might however be a consequence of the discretized version of the CIR process we used. For $\widehat{a}_T$, there is a significant difference (factor 5) between the empirical variances obtained with $c = 0.01$ and $c = 1$. However, for $\widehat{b}_T$ both empirical variances do not significantly differ.



Variance of $\widehat{a}_{50}$ as a function of c.



Variance of $\widehat{b}_{50}$ as a function of c.

Chapitre 5

Conclusion et perspectives

Dans la littérature sur les estimateurs du maximum de vraisemblance des paramètres de dimension et de dérive d'un processus de Cox-Ingersoll-Ross, on trouvait des résultats de grandes déviations et de déviations modérées dans la cas où l'un des deux paramètres est supposé connu pour estimer l'autre, et sous certaines bonnes conditions sur les valeurs des paramètres. Dans les chapitres 2 et 3, on a étendu ces résultats au cas où les deux paramètres sont estimés simultanément. Dans le Chapitre 3, on a considéré aussi des processus de Heston et on a établi un principe de déviations modérées pour l'estimateur du maximum de vraisemblance des quatre paramètres de ce type de processus.

Ces résultats de grandes déviations et déviations modérées se restreignent au cas le plus simple où le processus CIR est sous-critique et strictement positif ($b < 0$ et $a > 2$). Il nous a semblé naturel d'étudier ensuite les autres situations. Tout d'abord, dans le cas où $a < 2$, le processus CIR s'annule fréquemment et l'estimateur du maximum de vraisemblance n'est défini que jusqu'à un certain temps d'arrêt. Ainsi, dans le chapitre 4, on a proposé un nouvel estimateur des moindres carrés pondérés plus simple à utiliser en pratique que celui proposé par Overbeck dans [65]. On a prouvé que ce nouvel estimateur est fortement consistant et asymptotiquement normal.

5.1 Le cas explosif

Une première ouverture possible serait de considérer le cas explosif où $b > 0$. Etant donné qu'il n'existe pas dans ce cas d'estimateur consistant pour a, un premier travail dans ce cadre consisterait à établir un PGD pour l'estimateur du maximum de vraisemblance de b uniquement. Dans le cas d'un processus d'Ornstein-Uhlenbeck classique sur-critique, on trouve un PGD pour le paramètre de dérive dans [16]. Très récemment, Bercu et Richou dans [18] ont redémontré ce résultat par une méthode beaucoup plus directe reposant sur une utilisation astucieuse du théorème de Gärtner-Ellis. On remarque que, dans la cas où a est supposé connu, on peut en déduire un PGD pour l'EMV $\bar{b}_T$ de b, en utilisant le fait que, lorsque a vaut 1,

$$\bar{b}_T = \left(\frac{Y_T^2 - T}{\int_0^T Y_t^2 \, \mathrm{d}t} \right)$$

où (Y_t) est un processus d'Ornstein-Uhlenbeck classique de dimension 1 et de paramètre de dérive $b/2$. On se ramène ensuite au cas où a est un réel strictement positif quelconque par une propriété de semi-groupe additif, comme l'a fait Zani dans [73]. La fonction de taux obtenue est la même que celle obtenue dans [16] et [18], à un facteur a près. Dans le cas où a n'est plus suppposé connu, on pourrait s'inspirer de la démarche de Bercu et Richou [18] pour établir un PGD pour l'EMV de b.

5.2 Modifications du processus de Heston

Une autre piste serait de regarder des processus un peu différents. En effet, du point de vue des applications en mathématiques financières, différentes variantes du processus de Heston ont été proposées afin de mieux coller aux données réelles.

5.2.1 Volatilité à sauts

Une première idée consiste à considérer un processus à sauts, en rajoutant un processus de Lévy dans l'équation du processus CIR. On trouve dans [67] un aperçu des résultats sur les processus de Heston à sauts et leur utilisation en mathématiques financières. Bakshi, Cao et Chen [7] donnent une justification empirique de l'emploi de tels modèles sur des options européennes.

Du point de vue de l'estimation des paramètres à partir d'observations en temps continu, Barczy *et al.* étudient le comportement asymptotique de l'EMV des paramètres d'un processus de Heston (sous-)critique à sauts dans [9]. Dans le cas d'un processus CIR à sauts, ils s'intéressent, dans [10], au comportement asymptotique de l'EMV du paramètre de dérive lorsque le paramètre de dimension est supposé connu. Pour un processus CIR sous-critique, on trouve ces résultats dans la thèse de Mai [64]. Il reste cependant des cas à traiter, et aucune question de grandes déviations ne semble avoir été étudiée.

5.2.2 Volatilité multifactorielle

Une deuxième idée est de considérer une volatilité multi-dimensionnelle, en remplaçant par exemple le processus CIR par un processus de Wishart, comme proposé par Da Fonseca *et al.* dans [24]. Ceci permet un meilleur contrôle de la dynamique de la covariance et semble donc mieux adapté aux caractéristiques de certaines volatilités observées sur des données réelles. Pour une description détaillée du modèle et des diverses hypothèses, on se réfèrera à [24]. Un processus de Wishart de dimension d est simplement la version d-dimensionnelle d'un processus CIR. C'est la solution de l'équation différentielle stochastique suivante :

$$\mathrm{d}X_t = (aI_d + bX_t + X_t b^\intercal)\mathrm{d}t + \sqrt{X_t}\,\mathrm{d}B_t + \mathrm{d}B_t^\intercal \sqrt{X_t}$$

où I_d est la matrice identité de taille d, le point de départ x_0 est donné par une matrice réelle symétrique semi-définie positive, le paramètre $a \geqslant d-1$, b est une matrice réelle carrée de taille d et (B_t) est un mouvement brownien standard de dimension d. Ces processus ont initialement été étudié par Bru, et on trouve dans son article [20] une démonstration de l'existence d'une unique solution faible pour $a \geqslant d-1$ et d'une unique solution forte pour $a \geqslant d+1$. Cette solution appartient à l'espace des matrices réelles symétriques définies positives. Le prix S_t de l'actif financier considéré est ensuite donné par

$$\frac{\mathrm{d}S_t}{S_t} = r\,\mathrm{d}t + \mathrm{Tr}\left(\sqrt{X_t}\,\mathrm{d}W_t\right)$$

où (W_t) est un mouvement brownien matriciel de taille d et le paramètre r est réel.

Une variante plus simple, le processus de Heston double, est proposée et étudiée par Christoffersen, Heston et Jacobs dans [21]. Il est donné par la solution de l'équation différentielle stochastique suivante :

$$\begin{cases} S_t^{-1}\mathrm{d}S_t = r\,\mathrm{d}t + 2\sqrt{X_t^1}\left(\rho_1\,\mathrm{d}B_t^1 + \sqrt{1-\rho_1^2}\,\mathrm{d}W_t^1\right) + 2\sqrt{X_t^2}\left(\rho_2\,\mathrm{d}B_t^2 + \sqrt{1-\rho_2^2}\,\mathrm{d}W_t^2\right) \\ \mathrm{d}X_t^1 \quad = (a+bX_t^1)\,\mathrm{d}t + 2\sqrt{X_t^1}\,\mathrm{d}B_t^1 \\ \mathrm{d}X_t^2 \quad = (c+dX_t^2)\,\mathrm{d}t + 2\sqrt{X_t^2}\,\mathrm{d}B_t^2 \end{cases}$$

où $(\mathrm{d}B_t^1, \mathrm{d}B_t^2, \mathrm{d}W_t^1, \mathrm{d}W_t^2)$ est un mouvement brownien standard. On trouve dans [41] des calculs de fonctions caractéristiques et une méthode efficace de simulation pour ce processus. Gauthier et Possamai y remarquent par ailleurs que l'augmention du nombre de processus CIR (au-delà de 2) n'a pas d'impact significatif sur la qualité de la modélisation.

Pour ces deux types de processus, l'estimation des paramètres, en particulier l'étude du maximum de vraisemblance et de son comportement asymptotique, reste à faire.

Par ailleurs, les processus à volatilité stochastique multi-dimensionnelle ont leur intérêt propre, puisqu'ils permettent aussi de modéliser la dépendance entre les prix de différents actifs financiers, comme expliqué par Gouriéroux et Sufana dans [42]. De plus, ils proposent dans [43] d'utiliser les processus de Wishart pour modéliser certains taux d'intérêt d'actifs financiers. Bien que ces processus occupent une place importante en mathématique financière, la question de l'estimation de leurs paramètres n'a été que peu abordée. Très récemment, cependant, Alfonsi, Kebaier et Rey ont étudié l'EMV des paramètres d'un processus de Wishart dans [3].

5.2.3 Volatilité fractionnaire

Enfin, une troisième idée consiste à remplacer le mouvement brownien standard par un mouvement Brownien fractionnaire dans les équations différentielles stochastiques associées aux processus CIR et de Heston. Très récemment, Gatheral, Jaïsson et Rosenbaum ont montré dans [40] qu'un processus de Heston fractionnaire avec un exposant de Hurst proche de $0,1$ permet de bien reproduire les caractéristiques de la surface de volatilité observée sur des données réelles. On pourra consulter aussi [12] à ce sujet. Par ailleurs, le lien étroit de ces processus avec les processus de Hawkes, mis en lumière dans les deux articles très récents de Jaïsson et Rosenbaum [51] et [52], permet de les utiliser efficacement dans des méthodes de pricing, comme expliqué dans [31].

La littérature sur les processus de Heston fractionnaires est en pleine expansion. On trouve en particulier des résultats sur leur comportement asymptotique dans l'article de Guennoun, Jacquier et Roome [45]. Cependant, il n'y a pas à notre connaissance de résultats concernant l'estimation des paramètres d'un processus CIR fractionnaire (et a fortiori d'un processus de Heston fractionnaire). Ceci pourrait consituer une piste de travail intéressante.

On note que dans le cas d'un processus d'Ornstein-Uhlenbeck fractionnaire, Kleptsyna et Le Breton s'intéressent dans [56] au comportement asymptotique de l'estimateur du maximum de vraisemblance du paramètre de dérive, tandis que Bercu, Coutin et Savy établissent dans [15] des grandes déviations précises pour cet estimateur.

On pourrait donc, dans un premier temps, étudier l'existence de l'EMV des paramètres pour ces deux types de processus fractionnaires ainsi que son comportement asymptotique (consistance, normalité asymptotique, ...). Dans un second temps, grâce à leur lien avec les processus de Hawkes, on pourrait déduire de ces EMV des estimateurs pour certains paramètres d'un processus de Hawkes, comme suggéré dans la discussion de la section 2.4 de [52].

Bibliographie

[1] Aït-Sahalia, Y., and Kimmel, R. Maximum likelihood estimation of stochastic volatility models. *Journal of Financial Economics 83*, 2 (2007), 413–452.

[2] Alfonsi, A. High order discretization schemes for the CIR process : application to affine term structure and Heston models. *Mathematics of Computation 79*, 269 (2010), 209–237.

[3] Alfonsi, A., Kebaier, A., and Rey, C. Maximum likelihood estimation for Wishart processes. *Stochastic Process. Appl. 126*, 11 (2016), 3243–3282.

[4] Anderson, L. Simple and efficient simulation of the heston stochastic volatility model. *Journal of Computational Finance 11* (2008).

[5] Athreya, K. B., and Ney, P. E. *Branching processes*. Springer-Verlag, New York-Heidelberg, 1972. Die Grundlehren der mathematischen Wissenschaften, Band 196.

[6] Azencott, R., and Gadhyan, Y. Accurate parameter estimation for coupled stochastic dynamics. *Discrete Contin. Dyn. Syst.*, Dynamical systems, differential equations and applications. 7th AIMS Conference, suppl. (2009), 44–53.

[7] Bakshi, G., Cao, C., and Chen, Z. Empirical performance of alternative option pricing models. *The Journal of Finance LII* (1997).

[8] Baldeaux, J., and Platen, E. *Functionals of multidimensional diffusions with applications to finance*, vol. 5 of *Bocconi & Springer Series*. Springer, Cham ; Bocconi University Press, Milan, 2013.

[9] Barczy, M., Ben Alaya, M., Kebaier, A., and Pap, G. Asymptotic behavior of maximum likelihood estimators for a jump-type Heston model. 2015, ArXiv :1509.08869.

[10] Barczy, M., Ben Alaya, M., Kebaier, A., and Pap, G. Asymptotic properties of maximum likelihood estimator for the growth rate for a jump-type cir process based on continuous time observations. 2016, ArXiv :1609.05865.

[11] Barczy, M., and Pap, G. Asymptotic properties of maximum-likelihood estimators for Heston models based on continuous time observations. *Statistics 50*, 2 (2016), 389–417.

[12] Bayer, C., Friz, P., and Gatheral, J. Pricing under rough volatility. *Quant. Finance 16*, 6 (2016), 887–904.

[13] Ben Alaya, M., and Kebaier, A. Parameter estimation for the square-root diffusions : ergodic and nonergodic cases. *Stoch. Models 28*, 4 (2012), 609–634.

[14] Ben Alaya, M., and Kebaier, A. Asymptotic Behavior of The Maximum Likelihood Estimator For Ergodic and Nonergodic Square-Root Diffusions. *Stochastic Analysis and Applications* (2013).

[15] Bercu, B., Coutin, L., and Savy, N. Sharp large deviations for the fractional Ornstein-Uhlenbeck process. *Teor. Veroyatn. Primen. 55*, 4 (2010), 732–771.

[16] Bercu, B., Coutin, L., and Savy, N. Sharp large deviations for the nonstationary Ornstein-Uhlenbeck process. *Stochastic Process. Appl. 122*, 10 (2012), 3393–3424.

[17] Bercu, B., and Richou, A. Large deviations for the Ornstein-Uhlenbeck process with shift. *Adv. Appl. Probab.* (2015).

[18] Bercu, B., and Richou, A. Large deviations for the Ornstein-Uhlenbeck process without tears. 2016, ArXiv :1602.02092.

[19] Bercu, B., and Rouault, A. Sharp large deviations for the Ornstein-Uhlenbeck process. *Teor. Veroyatnost. i Primenen. 46*, 1 (2001), 74–93.

[20] Bru, M.-F. Wishart processes. *J. Theoret. Probab. 4*, 4 (1991), 725–751.

[21] Christoffersen, P., Heston, S., and Jacobs, K. The shape and term structure of the index option smirk : Why multifactor stochastic volatility models work so well. *Management Science 55*, 12 (2009), 1914–1932.

[22] Cox, J. C., Ingersoll, Jr., J. E., and Ross, S. A. A theory of the term structure of interest rates. *Econometrica 53*, 2 (1985), 385–407.

[23] Craddock, M., and Lennox, K. A. The calculation of expectations for classes of diffusion processes by Lie symmetry methods. *Ann. Appl. Probab. 19*, 1 (2009), 127–157.

[24] Da Fonseca, J., Grasselli, M., and Tebaldi, C. A multifactor volatility heston model. *Quantitative Finance 8*, 6 (sep 2008), 591–604.

[25] Dembo, A., and Zeitouni, O. *Large deviations techniques and applications*, second ed., vol. 38 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1998.

[26] du Roy de Chaumaray, M. Weighted least squares estimation for the subcritical Heston process. 2015, https ://arxiv.org/abs/1509.09167.

[27] du Roy de Chaumaray, M. Large deviations for the squared radial Ornstein-Uhlenbeck process. *Teor. Veroyatnost. i Primenen 61* (2016), 509–546.

[28] du Roy de Chaumaray, M. Moderate deviations for parameters estimation in a geometrically ergodic Heston process. 2016, https ://arxiv.org/abs/1607.05842.

[29] Duffie, D., Filipović, D., and Schachermayer, W. Affine processes and applications in finance. *Ann. Appl. Probab. 13*, 3 (08 2003), 984–1053.

[30] Duffie, D., Pan, J., and Singleton, K. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica 68*, 6 (2000), 1343–1376.

[31] El Euch, O., and Rosenbaum, M. The characteristic function of rough Heston models. 2016, ArXiv :1609.02108.

[32] Ellis, R. S. Large deviations for a general class of random vectors. *Ann. Probab. 12*, 1 (1984), 1–12.

[33] Engelmann, B., Koster, F., and Oeltz, D. Calibration of the heston stochastic local volatility model : A finite volume scheme. *http ://ssrn.com/abstract=1823769* (2011).

[34] Feller, W. Two singular diffusion problems. *Ann. of Math. (2) 54* (1951), 173–182.

[35] Forde, M., and Jacquier, A. The large-maturity smile for the Heston model. *Finance Stoch. 15*, 4 (2011), 755–780.

[36] Forde, M., Jacquier, A., and Lee, R. The small-time smile and term structure of implied volatility under the Heston model. *SIAM J. Financial Math. 3*, 1 (2012), 690–708.

[37] Fournié, E., and Talay, D. Application de la statistique des diffusions à un modèle de taux d'intérêt. *Finance 12* (1991).

[38] Gao, F., and Jiang, H. Moderate deviations for squared Ornstein-Uhlenbeck process. *Statistics & Probability Letters 79*, 11 (2009), 1378–1386.

[39] Gatheral, J. *The Volatility Surface : A Practitioner's Guide*. Wiley, 2006.

[40] Gatheral, J., Jaisson, T., and Rosenbaum, M. Volatility is rough. 2014, arXiv :1410.3394.

[41] Gauthier, P., and Possamaï, D. Efficient simulation of the double Heston model.

[42] Gourieroux, C., and Sufana, R. Derivative pricing with Wishart multivariate stochastic volatility. *J. Bus. Econom. Statist. 28*, 3 (2010), 438–451.

[43] Gourieroux, C., and Sufana, R. Discrete time Wishart term structure models. *J. Econom. Dynam. Control 35*, 6 (2011), 815–824.

[44] Gradshteyn, I. S., and Ryzhik, I. M. *Table of integrals, series, and products*. Academic Press, 1980.

[45] Guennoun, H., Jacquier, A., and Roome, P. Asymptotic behaviour of the fractionnal Heston model. 2014, arXiv :1411.7653.

[46] Gärtner, J. On large deviations from an invariant measure. *Teor. Verojatnost. i Primenen. 22*, 1 (1977), 27–42.

[47] Heston, S. L. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies 6* (1993), 327–343.

[48] Hurn, A. S., Lindsay, K. A., and McClelland, A. J. A quasi-maximum likelihood method for estimating the parameters of multivariate diffusions. *J. Econometrics 172*, 1 (2013), 106–126.

[49] Jacquier, A., and Roome, P. Large-maturity regimes of the Heston forward smile. *Stochastic Process. Appl. 126*, 4 (2016), 1087–1123.

[50] Jaisson, T. *Market activity and price impact throughout time scales*. PhD thesis, Ecole Polytechnique, 2015.

[51] Jaisson, T., and Rosenbaum, M. Limit theorems for nearly unstable Hawkes processes. *Ann. Appl. Probab. 25*, 2 (2015), 600–631.

[52] Jaisson, T., and Rosenbaum, M. Rough fractional diffusions as scaling limits of nearly unstable heavy tailed Hawkes processes. *Ann. Appl. Probab. 26*, 5 (2016), 2860–2882.

[53] Janek, A., Kluge, T., Weron, R., and Wystup, U. FX smile in the Heston model. In *Statistical tools for finance and insurance*. Springer, Heidelberg, 2011, pp. 133–162.

[54] Jeanblanc, M., Yor, M., and Chesney, M. *Mathematical methods for financial markets*. Springer Finance. Springer-Verlag London, Ltd., London, 2009.

[55] Kawazu, K., and Watanabe, S. Branching processes with immigration and related limit theorems. *Teor. Verojatnost. i Primenen. 16* (1971), 34–51.

[56] Kleptsyna, M. L., and Le Breton, A. Statistical analysis of the fractional Ornstein-Uhlenbeck type process. *Stat. Inference Stoch. Process. 5*, 3 (2002), 229–248.

[57] Kutoyants, Y. A. *Statistical inference for ergodic diffusion processes*. Springer Series in Statistics. Springer-Verlag London Ltd., London, 2004.

[58] Lamberton, D., and Lapeyre, B. *Introduction au calcul stochastique appliqué à la finance*, second ed. Ellipses Édition Marketing, Paris, 1997.

[59] Lebedev, N. N. *Special functions and their applications*. Dover Publications Inc., New York, 1972. Revised edition, translated from the Russian and edited by Richard A. Silverman, Unabridged and corrected republication.

[60] Lee, R. Option pricing by transform methods : Extensions, unification, and error control. *Journal of Computational Finance 7* (2004), 51–86.

[61] Lewis, A. L. *Option valuation under stochastic volatility*. Finance Press, Newport Beach, CA, 2000.

[62] Luke, Y. L. *The Special Functions and their Approximations. Vol. 2.* Academic Press, 1969.

[63] Luke, Y. L. Inequalities for generalized hypergeometric functions. *J. Approximation Theory 5* (1972), 41–65. Collection of articles dedicated to J. L. Walsh on his 75th birthday, I.

[64] Mai, M. *Drift estimation for jump diffusions : time-continuous and high-frequency observations*. PhD thesis, Humboldt Universität, Berlin, 2012.

[65] Overbeck, L. Estimation for continuous branching processes. *Scandinavian Journal of Statistics. Theory and Applications 25* (1998).

[66] Pitman, J., and Yor, M. A decomposition of Bessel bridges. *Z. Wahrsch. Verw. Gebiete 59*, 4 (1982), 425–457.

[67] Runggaldier, W. *Handbook of Heavy Tailed Distributions in Finance, Chapter 5 : Jump-diffusion models*. Elsevier, 2003.

[68] Stein, E., and Stein, J. Stock price distributions with stochastic volatility : an analytic approach. *Review of Financial Studies 4* (1991), 727–752.

[69] Stein, J. Overreactions in the options market. *Journal of finance 44* (1989), 1011–1024.

[70] Wei, C., and Winnicki, J. Estimation of the means in the branching process with immigration. *Annals of Statistics 18* (1990).

[71] Worms, J. *Principe de déviations modérées pour des martingales et applications statistiques*. PhD thesis, Université Marne-la-Vallée, 2000.

[72] Zani, M. *Grandes déviations pour des fonctionnelles issues de la statistique des processus*. PhD thesis, Paris 11, 2000.

[73] Zani, M. Large deviations for squared radial Ornstein–Uhlenbeck processes. *Stochastic Processes and their Applications 102*, 1 (2002), 25 – 42.

Résumé

Les processus de Cox-Ingersoll-Ross et de Heston jouent un rôle prépondérant dans la modélisation mathématique des cours d'actifs financiers ou des taux d'intérêts. Dans cette thèse, on s'intéresse à l'estimation de leurs paramètres à partir de l'observation en temps continu d'une de leurs trajectoires.

Dans un premier temps, on se place dans le cas où le processus CIR est géométriquement ergodique et ne s'annule pas. On établit alors un principe de grandes déviations pour l'estimateur du maximum de vraisemblance du couple des paramètres de dimension et de dérive d'un processus CIR. On établit ensuite un principe de déviations modérées pour l'estimateur du maximum de vraisemblance des quatre paramètres d'un processus de Heston, ainsi que pour l'estimateur du maximum de vraisemblance du couple des paramètres d'un processus CIR. Contrairement à ce qui a été fait jusqu'ici dans la littérature, les paramètres sont estimés simultanément. Dans un second temps, on ne se restreint plus au cas où le processus CIR n'atteint jamais zéro et on propose un nouvel estimateur des moindres carrés pondérés pour le quadruplet des paramètres d'un processus de Heston. On établit sa consistance forte et sa normalité asymptotique, et on illustre numériquement ses bonnes performances.

Mots-clés: Processus CIR, processus de Heston, inférence paramétrique, estimateur du maximum de vraisemblance, grandes déviations, déviations modérées.

Abstract

The Cox-Ingersoll-Ross process and the Heston process are widely used in financial mathematics for pricing and hedging or to model interest rates. In this thesis, we focus on estimating their parameters using continuous-time observations.

Firstly, we restrict ourselves to the most tractable situation where the CIR process is geometrically ergodic and does not vanish. We establish a large deviations principle for the maximum likelihood estimator of the couple of dimensionnal and drift parameters of a CIR process. Then we establish a moderate deviations principle for the maximum likelihood estimator of the four parameters of an Heston process, as well as for the maximum likelihood estimator of the couple of parameters of a CIR process. In contrast to the previous literature, parameters are estimated simultaneously. Secondly, we do not restrict ourselves anymore to the case where the CIR process never reaches zero and we introduce a new weighted least squares estimator for the quadruplet of parameters of an Heston process. We establish its strong consitency and asymptotic normality, and we illustrate numerically its good performances.

Keywords: CIR process, Heston process, parameter inference, maximum likelihood estimator, large deviations, moderate deviations.