Estimation de la volatilité pour des processus de diffusion : grandes déviations et déviations modérées

Yacouba Samoura

$N°$ d'ordre : DU 2769

THÈSE DE DOCTORAT DE
l'UNIVERSITÉ CLERMONT AUVERGNE - BLAISE PASCAL
(UFR Sciences et Technologies)
$N°$: 896

Spécialité

Mathématiques appliquées

ÉCOLE DOCTORALE DES SCIENCES FONDAMENTALES
(Clermont-Ferrand)

Présentée par

Yacouba SAMOURA

Pour obtenir le grade de

DOCTEUR de l'UNIVERSITÉ CLERMONT AUVERGNE - BLAISE PASCAL

Sujet de la thèse :

ESTIMATION DE LA VOLATILITÉ POUR DES PROCESSUS DE DIFFUSION : GRANDES DÉVIATIONS ET DÉVIATIONS MODÉRÉES

soutenue publiquement le 09 décembre 2016

devant le jury composé de :

M. Jamal **NAJIM**	Université Paris-Est Marne la Vallée	Rapporteur
Mme. Marguerite **ZANI**	Université d'Orléans	Rapporteur-Président
Mme. Emmanuelle **CLÉMENT**	Université Paris-Est Marne la Vallée	Examinateur
M. Liming **WU**	Université Blaise Pascal	Examinateur
M. Hacène **DJELLOUT**	Université Blaise Pascal	Directeur
M. Arnaud **GUILLIN**	Université Blaise Pascal	Directeur

Remerciements

Il est temps pour moi d'exprimer toute ma reconnaissance et ma gratitude envers mon directeur de thèse Arnaud GUILLIN grâce à qui ce travail a non seulement vu le jour, mais a aussi été mené à terme. Son enthousiasme sans faille, sa disponibilité et ses conseils inconditionnels m'ont permis de traiter des sujets toujours intéressants les uns que les autres durant ces trois dernières années. J'ai profondement été touché tant par le professeur que par la personne.

Je tiens à remercier Hacène DJELLOUT pour avoir accepté de diriger ce travail. Tout au long de mes trois années de cycle d'ingénieur et des trois dernières années, il a su me faire bénéficier de ses expériences et de sa bonne humeur, et ses conseils m'ont été d'une aide considérable pour l'élaboration de cette thèse.

J'adresse mes plus sincères remerciements à Jamal NAJIM et Marguerite ZANI pour l'honneur qu'ils me font en acceptant de rapporter cette thèse. Leurs remarques et leurs suggestions m'ont permis d'améliorer clairement la qualité de cette thèse.

Emmanuelle CLÉMENT, Liming WU ont toute ma gratitude et toute ma reconnaissance pour avoir accepté de faire partie de mon jury de thèse.

C'est aussi le moment pour moi de remercier la direction du département de mathématiques, Valérie, Marie-Paule, Annick, Damien, Cédric de m'avoir accueilli au sein de leur laboratoire et d'avoir mis à ma disposition tous les moyens nécessaires au bon déroulement de ce travail. Un grand merci à François BOUCHON, Nicolas BILLEREY, Pierre DRUILHET et Erwan SAINT LOUBERT BIE pour nos nombreuses discussions sur maints sujets et leurs nombreux conseils. Je remercie toute l'équipe PAS du laboratoire de mathématiques et tout le corps professoral de l'université Blaise Pascal.

Je remercie également mes collègues François, Rabih, Doha et Arthur pour le soutien qu'ils m'ont témoingné tout au long de ces années. Je tiens aussi à remercier Christelle BIOCHE pour sa gentillesse, ses conseils et nos nombreuses discussions fructueux.

Tout particulièrement, je tiens à remercier ma femme Aramatou KONATE pour son amour, sa patience, son attention et son soutien sans faille tout au long de la réalisation de cette thèse.

Mes derniers remerciements vont à l'endroit de ma famille qui m'a soutenu et qui continue de me soutenir par son encouragement, sa confiance, son amour, son soutien moral et matériel malgré la distance.

Je dédie cette thèse à ma maman que j'aime beaucoup et
à la mémoire de mon père qui me manque énormement.

Table des matières

Organisation générale de ce document

En plus d'une introduction synthétisant les résultats obtenus tout au long de la réalisation de ces travaux tout en les mettant dans leur contexte, cette thèse est divisée en trois grandes parties. Sans perdre de généralité, on essayera de traiter ces differentes parties de façon indépendante, afin de faciliter la compréhension de cette thèse à quiconque souhaitant la lire.

Dans la première partie, il est question de rappeler les résultats obtenus d'une part et d'introduire la notion du principe de grandes déviations d'autre part. On y présente quelques outils généraux inspirés du "magnifique" ouvrage "Large deviation techniques and applications" de Dembo et Zeitouni [53] qui ont servi à la bonne réalisation de nos travaux. En plus du théorème de Gärtner-Ellis, nous présentons certains grands résultats de Najim ([104], [105], [106]) et de Gao et Zhao [73] dans le cadre des grandes déviations dans $\mathbb{R}^d$. Mais néanmoins toutes les preuves seront omises dans ce chapitre.

La deuxième partie est elle même divisée en deux sections. Une première section, dans laquelle on considère deux équations aux dérivées stochastiques observées à des temps uniformes et synchronisés, qui est consacrée aux deux articles [61]"Large and moderate deviations of realized covolatility" publié dans le journal Statistics Probability Letters et [58] "Estimation of the realized (co-)volatility vector : Large deviations approach" en révision dans Stochastics Processes and their Applications. Ces travaux sont dans la lignée des travaux de Djellout & al. [59]. Nous montrons dans le premier article des résultats de type grandes déviations et déviations modérées pour l'estimateur de la covariance appélé la co-volatilité en finance. A la suite de ces travaux, nos questions nous ont conduit à considérer le vecteur formé des estimateurs des variances et de la covariance permettant de nombreuses applications en statistique. Ainsi le deuxième article [61]"Estimation of the realized (co)volatility vector : Large deviations approach" fournit des résultats de type grandes déviations et déviations modérées pour ce vecteur grâce à des outils de transfert de principe de grandes déviations via les résultats

de Najim et aussi quelques exemples d'application pour les coefficients de corrélation et de régression.
La deuxième section est consacrée à l'article [60] qui a été réalisé à la lumière des travaux précedents et dans lequel on considère des processus de diffusion observées à des temps uniformes mais de manière non-synchronisée (asynchrone). On établit alors des résultats de type grandes déviations et déviations modérées pour le vecteur (co-)volatilité dans ce contexte.

Dans la troisième partie, on cherche à estimer la statistique de Durbin-Watson. Pour cela, on considère un processus autorégressif d'ordre p et le bruit qui y est associé vérifie lui aussi un processus autorégressif d'ordre q et on montre des principes de type déviations modérées pour certains estimateurs associés à notre modèle dont la statistique de Durbin-Watson.

Chapitre 1

Introduction

1.1 Problème de l'estimation de la volatilité

Considérée en finance comme la base de la mesure du risque, la volatilité est par définition une mesure des amplitudes des variations du cours d'un actif financier. Plus la volatilité d'un actif est élevée, plus l'investissement dans cet actif est considéré comme risqué et par conséquent l'espérance de gain (ou le risque de perte) pour cet actif est important. À l'inverse un actif sans risque ou très peu risqué (par exemple les bons du trésor) a une volatilité très faible car son remboursement est quasiment certain.

Dans le monde d'aujourd'hui où les processus de diffusion jouent un rôle très important dans de nombreux domaines comme la santé, les mathématiques financières ou la génétique des populations, il est donc primordial de pouvoir estimer le coefficient de diffusion, appelé généralement la volatilité en finance, associé à ces processus de diffusion. Les processus de diffusion sont souvent observés à des temps régulièrement espacés (uniformes) ou non, ou encore à des temps aléatoires, d'où l'apparition de certaines difficultés dans l'estimation de la volatilité.

Depuis quelques années, plusieurs travaux ont été réalisés afin de montrer la consistance et la normalité asymptotique des estimateurs dans ces cas. Par soucis d'approfondissement des résultats pour ces estimateurs, d'autres travaux permettent de montrer des théorèmes plus puissants de type grandes déviations, c'est à dire de prouver des équivalences logarithmiques pour l'erreur d'estimation.

1.1.1 Cas synchrone

Le comportement asymptotique pour des diffusions observées à haute fréquence et synchronisées connait un essor considérable depuis plus de vingt ans dans le monde de la finance où les cours des actions ou des monnaies sont très souvent observés. Cela est d$\hat{u}$ au fait que, lorsque le nombre d'observations à haute fréquence augmente à l'infini, naturellement on pense à faire une étude asymptotique. Dans cette partie, il s'agit d'étudier l'estimation de la volatilité pour des processus de diffusion observés à des temps uniformes et synchronisés.

Plus précisement, soit $X_t = (X_{1,t}, X_{2,t}), t \in [0,1]$ un processus stochastique modélisant l'évolution d'une variable d'état observable et soit $(\Omega, \mathfrak{F}, (\mathfrak{F}_t), \mathbb{P})$ un espace de probabilité filtré. Dans les applications financières, X_t peut être considéré comme un taux d'intérêt à court terme, un indice boursier, un taux de change ou le logarithme du prix d'un actif en finance.
On suppose que les deux processus de diffusion $X_{1,t}$ et $X_{2,t}$ sont définis sur $(\Omega, \mathfrak{F}, (\mathfrak{F}_t), \mathbb{P})$ et vérifient le processus de Itô

$$\left\{\begin{array}{lllll} dX_{1,t} & = & \sigma_{1,t} dB_{1,t} & + & b_1(t, X_t)dt \\ dX_{2,t} & = & \sigma_{2,t} dB_{2,t} & + & b_2(t, X_t)dt \end{array}\right. \tag{1.1}$$

où les mouvements browniens $B_1 = (B_{1,t})_{t\in[0,1]}$ and $B_2 = (B_{2,t})_{t\in[0,1]}$ sont corrélés entre eux avec $cov(B_{1,t}, B_{2,t}) = \int_0^t \rho_s \mathrm{d}s$, b_1 et b_2 sont des processus adaptés représentant les vitesses moyennes (aussi appelées les dérives). On peut aussi définir un troisième mouvement brownien $B_3 = (B_{3,t})_{t\in[0,1]}$ tel que $dB_{2,t} = \rho_t dB_{1,t} + \sqrt{1-\rho_t^2} dB_{3,t}$, avec B_1 et B_3 deux mouvements browniens indépendants.
Dans nos travaux, on se limite seulement au cas où σ_1, σ_2 et ρ sont des fonctions déterministes. Les fonctions σ_1 et σ_2 ne prennent que des valeurs positives tandis que ρ prend ses valeurs dans $]-1,1[$. On suppose que $X_{1,t}$ et $X_{2,t}$ sont observés de manière synchrone et on note $[X_\ell]_t = \int_0^t \sigma_{\ell,s}^2 \mathrm{d}s$ avec $\ell = 1, 2$ la variation quadratique de X_ℓ et $\langle X_1, X_2 \rangle_t = \int_0^t \sigma_{1,s}\sigma_{2,s}\rho_s \mathrm{d}s$ la covariation déterministe de $X_{1,t}$ et $X_{2,t}$.
Les coefficients de diffusion $\sigma_{\ell,t}$ pour $\ell = 1, 2$ et celui de correlation ρ_t sont supposés inconnus et on cherche à les estimer à partir d'un échantillon de $(X_t)_{t\in[0,1]}$. Plus précisement, on cherche à estimer la covariation inconnue $\langle X_1, X_2 \rangle_t$ de X d'une part et le vecteur de la (co-)variation inconnu d'autre part et donné par

$$[V]_t = ([X_1]_t, [X_2]_t, \langle X_1, X_2 \rangle_t)^T.$$

Ce type de problème apparaît naturellement en mathématiques financières, et on appelle $[X_\ell]_t$ la volatilité, $\langle X_1, X_2 \rangle_t$ la covolatilité et V_t la (co-)volatilité. Étant données des observations discrètes également espacées $(X_{1,t_k^n}, X_{2,t_k^n}, k = 1, \cdots, n)$ dans l'intervalle $[0,1]$ (avec $t_k^n = k/n$), des théorèmes limites pour le processus stochastique affirment que $C_t^n(X)$ la covariance réalisée est un estimateur consistant de $\langle X_1, X_2 \rangle_t$ et celui de V_t est $V_t^n(X)$ le vecteur de la (co-)variance

$$V_t^n(X) = \Big(Q_{1,t}^n(X), Q_{2,t}^n(X), C_t^n(X)\Big)^T$$

avec des quantités $Q_{\ell,t}^n(X)$ et $C_t^n(X)$ pour $\ell = 1, 2$, données par

$$Q_{\ell,t}^n(X) = \sum_{k=1}^{[nt]} (\Delta_k^n X_\ell)^2 \qquad C_t^n(X) = \sum_{k=1}^{[nt]} (\Delta_k^n X_1)(\Delta_k^n X_2) \tag{1.2}$$

où $\Delta_k^n X_\ell = X_{\ell,t_k^n} - X_{\ell,t_{k-1}^n}$ avec $\tau_n = \{t_k^n, 0 \leqslant k \leqslant n\}$ une subdivision uniforme de $[0,1]$ à n parties.
On s'intéresse ainsi aux grandes déviations de $C_t^n(X)$ d'une part et celles de $V_t^n(X)$ d'autre part, *i.e* établir des relations telles que

$$\mathbb{P}(C_t^n - \langle X_1, X_2\rangle_t > x) \approx e^{-nI_1(x)} \tag{1.3}$$

et

$$\mathbb{P}(V_t^n - [V]_t > x) \approx e^{-nI_2(x)} \tag{1.4}$$

où I_1 et I_2 sont des fonctions données.
La connaissance des comportements asymptotiques des quantités $Q_{\ell,t}^n(X)$ et $C_t^n(X)$ pour $\ell = 1, 2$ ne donne aucune information ni sur le comportement asymptotique de la correlation réalisée $\varrho_1 = C_1/\sqrt{Q_{1,1}Q_{2,1}}$, ni sur celui du coefficient de régression réalisée $\beta_{\ell,1} = C_1/Q_{\ell,1}$ pour $\ell = 1, 2$ qui sont estimés respectivement par $\varrho_1^n(X) = C_1^n(X)/\sqrt{Q_{1,1}^n(X)Q_{2,1}^n(X)}$ et $\beta_{\ell,1}^n(X) = C_1^n(X)/Q_{\ell,1}^n(X)$. Il est donc indispensable d'estimer le comportement asymptotique du vecteur $[V]_t$ pour pouvoir estimer celui des mesures standards de dépendance entre deux retours d'actifs tels que la correlation réalisée ou le coefficient de régression réalisée.

D'un point de vue parametrique Kutoyants [88] et Lanska [89] et d'un point de vue non parametrique Bannon et Nguyen [11] et Pham-Dinh [110] ont étudié le comportement asymptotique du coefficient de diffusion et du drift pour des processus de diffusion. De ces deux points de vue, D. Florens-Zmirou [70] et Dacunha-Castelle et Florens-Zmirou [46] ont obtenu la consistence ainsi que des théorèmes limites centraux de l'estimateur de la variance quadratique. Plusieurs questions statistiques sur ce sujet ont été étudiées par Avesani-Bertrand [7] et Bertrand [21]. C'est dans cette optique que Barndorff-Nielsen et Shephard [15] ont fourni la loi des grands nombres et étudié des fluctuations correspondantes de la volatilité realisée. Par la suite, Barndorff-Nielsen & al. [13]-[12] et Dovonon, Gonçalves et Meddahi [63] ont étendu ces estimations à des configurations plus générales de la statistique en considérant des expansions d'Edgeworth pour la statistique de la volatilité réalisée et son analogue Bootstrap.

Cependant, il existe plusieurs travaux qui ont permis d'établir des théorèmes beaucoup plus puissants de type grandes déviations pour estimer la volatilité. Lynch-Sethuraman [97] ont fourni des grandes déviations pour des processus à accroissements indépendants et homogènes, Puhalskii [114] a initié une méthode générale pour les grandes déviations pour des processus stochastiques. Bryc-Dembo [36], Bercu-Gamboa-Rouault [10] et Zani [108]-[140] ont par la suite établi des grandes déviations pour des formes quadratiques de processus gaussiens stationnaires.
Inspirés de ces travaux, Djellout & al. [59] ont établi des grandes déviations pour la volatilité réalisée et ont aussi fourni quelques inégalités de déviations pour cet estimateur. Les travaux récents de Kanaya-Otsu [85] sur les grandes déviations de la volatité réalisée sont contenus dans Djellout & al. [59]. On peut aussi citer les travaux de Mancini [99] sur les grandes déviations de l'estimateur tronqué de la volatilité réalisée constante pour des diffusions à sauts et les travaux de Jiang [84] sur le principe des déviations modérées trajectoriel pour cet estimateur du processus variationnel quadratique.

1.1.2 Cas non-synchrone

Notons qu'en général dans les applications financières, les données de transactions réelles sont relevées à des heures irrégulières, d'une manière non synchrone, c'est à dire que deux prix d'une transaction ne sont généralement pas observées aux mêmes instants. Nous menons ainsi une étude sur le comportement asymptotique de la volatilité pour des processus de diffusion observés à des temps non régulièrement espacés et d'une manière non synchronisée.

On suppose que dans un espace de probabilité filtré $(\Omega, \mathfrak{F}, (\mathfrak{F}_t), \mathbb{P})$, on dispose de deux processus de diffusion $X_{1,t}$ et $X_{2,t}$, $t \in [0,T]$ donnés par la relation (1.1) et observés de manière asynchrone. Explicitement, cela se traduit par le fait qu'on ne considère que des $(X_{1,s_i}, X_{2,t_j})_{i=0,\cdots,n;j=0,\cdots,m}$ avec $0 = s_0 < s_1 < \cdots < s_n = T$, $0 = t_0 < t_1 < \cdots < t_m = T$ avec $m, n \in \mathbb{N}$ $(s_{i-1}, s_i] \cap (t_{i-1}, t_i] \neq \emptyset$. On s'intéresse ainsi à l'estimateur de la covariance introduit par Hayashi et Yoshida [78] et donné par

$$U_{n,m} = \sum_{i=1}^{n} \sum_{j=1}^{m} \Delta X_1(I^i) \Delta X_2(J^j) I_{\{I^i \cap J^j \neq \emptyset\}},$$

où $I^i = (s_{i-1}, s_i]$ $J^j = (t_{i-1}, t_i]$ et on cherche à montrer que

$$\mathbb{P}(U_{n,m} - \langle X_1, X_2 \rangle_t > x) \approx e^{-\frac{b_n^2}{n} I_3(x)} \tag{1.5}$$

où b_n est une suite de nombres positifs et I_3 une fonction donnée.

Dans la littérature, Anderson et al. [3] ont proposé une utilisation préliminaire de l'estimateur de la variance quadratique et de la covariance des taux de change. La théorie sur leur méthode n'est pas nouvelle, mais dans la pratique elle est très utile. L'impact de l'estimation de la volatilité pour des diffusions observées à haute fréquence, d'une manière non synchrone a été longuement étudiée en finance et en économétrie par Scholes et Williams [124] et Lo et Mackinlay [95]. Pour plus de références sur l'estimation de la volatilité pour des données à haute fréquence et non synchrones, nous renvoyons le lecteur aux travaux de Lundin et Müller [96], Muthuswamy et al. [103], Brandt et Diebold [34] et Tsay et Yeh [129].
Motivés par ces travaux, Hayashi et Yoshida [78]-[79]-[80] ont établi la consistance de l'estimateur de la covariance réalisée $U_{n,m}$ et aussi un théorème limite central pour cet estimateur.

Les travaux cités ci-dessus ont été très importants dans l'étude des comportements asymptotiques de la (co-)volatilité réalisée. Barndorff-Nielsen & al. [13]-[16] ont établi une loi faible des grands nombres et un théorème limite central pour un estimateur plus général, appélé variation réalisée généralisée à double puissance et ces travaux ont été synthétisés dans Mathias-Vetter [131].

Tous nos travaux sur ces estimateurs dans ces conditions ont été principalement motivés par le fait que de nos jours il existe peu de résultats concrets sur les grandes déviations de ces estimateurs. Cela peut s'expliquer du fait que les techniques habituelles (telles que les techniques de Gärtner Ellis) ne s'appliquent que pour des cas simples et ne fonctionnent pas en général pour des cas un peu complexe d'une part, ou par l'impossibilité de calculer directement la log-Laplace d'autre part.

1.2 La statistique de Durbin-Watson

Les tests de Durbin-Watson sont des tests qui permettent de détecter l'auto-correlation résiduelle dans des modèles autorégressifs. Inspirée des travaux de Von Neumann [133] et introduite à la moitie du siècle dernier par Durbin-Watson [64]-[65]-[66], la statistique de Durbin-Watson est à l'origine des tests de Durbin-Watson. Sa valeur est comprise entre 0 et 4. L'hypothèse nulle d'abscence d'autocorrelation est acceptée lorsque la valeur de cette statistique est proche de 2. Une valeur proche de 4 signifie qu'il existe une autocorrelation négative et une valeur proche de 0 indique la présence d'une autocorrelation positive. Plusieurs études comparant les tests de Durbin-Watson à d'autres tests de détection d'autocorrelation existent dans la littérature, on retient entre autre les travaux de Tillman [127] basés sur l'étude de la puissance du test de Durbin-Watson pour un modèle autorégressif d'ordre 1 et ceux de Blattberg [26] dans lesquels il a été prouvé que les tests de Durbin-Watson sont encore plus puissants pour des modèles autorégressifs qui ne sont pas du premier ordre. Bercu et Proïa [19] ont récemment montré par des simulations numériques que les tests de Durbin-Watson sont plus efficaces que ceux habituellement utilisés dans l'étude théorique de la convergence et du TCL et ceux utilisées par les économètres (les tests de Box-Pierce [33] et de Ljung-Box [32]).

On considère un processus autorégressif d'ordre p donné, pour tout $n \geqslant 1$, par

$$\left\{ \begin{array}{rcl} X_n & = & \theta' \Phi_{n-1}^p + \varepsilon_n \\ \varepsilon_n & = & \rho' \Psi_{n-1}^q + V_n \end{array} \right. \tag{1.6}$$

avec

$$\Phi_n^p = (X_n \quad X_{n-1} \quad \cdots \quad X_{n-p+1})' \quad et \quad \Psi_n^q = (\varepsilon_n \quad \varepsilon_{n-1} \quad \cdots \quad \varepsilon_{n-q+1})'. \tag{1.7}$$

pour $p \in \mathbb{N}$ et $q \in \mathbb{N}$ et où les paramètres $\theta = (\theta_1 \quad \theta_2 \quad \cdots \quad \theta_p)'$ et $\rho = (\rho_1 \quad \rho_2 \quad \cdots \quad \rho_q)'$ sont des vecteurs inconnus non nuls.
Récemment, Bercu et Proïa [19] ont étudié le comportement asymptotique de l'estimateur des moindres carrés et celui de la statistique de Durbin-Watson pour un modèle autorégressif d'ordre 1 de bruit autocorrélé d'ordre 1 aussi et ils ont prouvé le TCL pour ces estimateurs. Proïa [112]-[113] a étudié le comportement asymptotique et prouvé le TCL de ces estimateurs pour un modèle autorégressif d'ordre p de bruit autocorrélé d'ordre q. Bitseki et al.[25] ont montré des théorèmes de déviations modérées pour ces estimateurs pour un modèle autorégressif d'ordre 1 de bruit autocorrélé d'ordre 1 sous diverses hypothèses sur le bruit.

1.3 La théorie des grandes déviations

Dans cette section, le principe des grandes déviations est introduit ainsi que d'autres grands résultats de cette théorie. Quelques méthodes et outils généraux abordés dans cette thèse y sont présentés. On citera entre autre les résultats de Najim et la delta méthode, particulièrement intéressants, ils ont une place assez considérable surtout pour la démonstration de nos principaux théorèmes qui sont abordés dans le chapitre 3. En effet, l'introduction de ces resultats permettra au lecteur d'assimiler assez facilement des sujets qui paraissent souvent complexes.

1.3.1 Notion de principe de grandes déviations

Introduction

Au début du 19^e siècle, Cramèr-Laplace [44] et Chernov [39] introduisaient une théorie concernant les événements rares et le calcul asymptotique de leur probabilité dans une échelle exponentielle, c'est la théorie des grandes déviations aussi appelée le principe des grandes déviations (PGD). Il a fallu attendre jusqu'en 1966 pour voir une définition formelle de cette théorie, fournie par Varadhan [130]. Intuitivement, la théorie des grandes déviations n'est autre qu'un approfondissement des comportements asymptotiques fournis par la loi des grands nombres et le théorème limite central. En effet, soit une suite (X_n) de variables aléatoires à valeurs dans un espace E. On dit que la suite (X_n) satisfait un principe de grandes déviations (PGD) de bonne fonction de taux $I : E \to [0, +\infty]$ et de vitesse a_n , si l'ordre de la probabilité que X_n soit proche de $x \in E$ est $\exp(-a_n I(x))$, et l'on note :

$$\mathbb{P}(X_n > x) \approx \exp(-a_n I(x)) \tag{1.8}$$

Sans perdre de généralité, on peut donc dire que les grandes déviations "classiques" sont un équivalent logarithmique de la propabilité, sinon on parle de grandes déviations sharp.

Définitions

Soit $(\mu_n)_{n\in\mathbb{N}}$ une suite de mesures de probabilités sur un espace polonais E muni de la tribu borélienne $\mathcal{B}$. Soit a_n une suite de nombres réels strictement positifs telle que $\lim_{n\to\infty} a_n = +\infty$.

Définition 1.1 (Fonction de taux) *On appelle fonction de taux I toute application $I : E \to [0, +\infty]$ semi-continue inférieurement (s.c.i), c'est à dire, pour tout $\alpha \in \mathbb{R}_+$, les ensembles de niveau $\{x \in E, I(x) \leq \alpha\}$, sont des parties fermées de E.*
De plus, si les ensembles de niveau sont compacts, alors on dit que I est

une bonne fonction de taux car la compacité permet à la fonction de taux d'atteindre son minimum.

Définition 1.2 (Principe de Grandes Déviations PGD) *On dit que la suite $(\mu_n)_n$ satisfait un principe de grandes déviations de vitesse a_n et de fonction de taux I, si on a*

(i) Pour tout fermé $F \in E$

$$\limsup_{n\to\infty} \frac{1}{a_n} \log(\mu_n(F)) \leqslant -\inf_{x\in F} I(x)$$

(ii) Pour tout ouvert $G \in E$,

$$\liminf_{n\to\infty} \frac{1}{a_n} \log(\mu_n(G)) \geqslant -\inf_{x\in G} I(x).$$

1.3.2 Principaux outils

Dans certaines situations, démontrer directement un PGD relève de l'impossible. Pour remédier à de tels problèmes, la question évidente que se posent les mathématiciens est la suivante : avons-nous un outil nous permettant de transférer un PGD d'un cas simple ou un PGD connu dont on dispose en un cas complexe qui nous intéresse ?
La réponse à cette question est l'une des forces du PGD, car oui il existe plusieurs outils nous permettant de passer d'un PGD d'un cas à un autre. On les appelle les outils de transformations de PGD. Parmi ces outils, il y en a deux qui ont particulièrement attiré notre attention : les approximations exponentielles et le principe de contraction. Ces deux outils sont utilisés par la suite et en grande partie dans les preuves de nos principaux résultats.

Approximations exponentielles

Définition 1.3 *Soit $(\mathcal{X}, d)$ un espace métrique. Soient $(\mu_n)_n$ et $(\nu_n)_n$ des mesures de probabilité sur $\mathcal{X}$. On dit que les mesures $(\mu_n)_n$ et $(\nu_n)_n$ sont exponentiellement équivalentes s'il existe des espaces probabilisés $(\Omega, \mathcal{F}_n, \mathbb{P}_n)_{n\in\mathbb{N}}$ et deux familles de variables $(X_n)_n$ et $(\tilde{X}_n)_n$ à valeurs dans $\mathcal{X}$, de lois jointes $(\mathbb{P}_n)_{n\in\mathbb{N}}$ et marginales $(\mu_n)_n$ et $(\nu_n)_n$ respectivement tels que pour tout $\delta > 0$,*

$$\limsup_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}_n\Big(d(X_n, \tilde{X}_n) > \delta\Big) = -\infty \tag{1.9}$$

et l'ensemble $\{\omega \in \Omega / d(X_n, \tilde{X}_n) > \delta\}$ est mesurable.

Le théorème suivant ([53], section 4.2.2), montre que des mesures de probabilité exponentiellement équivalentes partagent le même PGD.

Théorème 1.3.1 (Approximation exponentielle) *Si un PGD de bonne fonction de taux I est atteint par les mesures $(\mu_n)_n$ et ces mesures $(\mu_n)_n$ sont exponentiellement équivalentes aux mesures $(\nu_n)_n$, alors la suite des mesures $(\nu_n)_n$ satisfait le même PGD que $(\mu_n)_n$.*

Principe de contraction

En plus d'une équivalence exponentielle, une application continue peut aussi permettre de transférer un PGD d'un espace à un autre. De démonstration élémentaire fournie dans Dembo et Zeitouni [53], section 4.2.1, le théorème suivant est très important et utilisé dans la majeur partie de nos principaux résultats.

Théorème 1.3.2 (Principe de contraction) *Soient G et H deux espaces de Hausdorff et $g : G \to H$ une fonction continue. Soient $(\mu_n)_n$ des mesures de probabilité sur G satisfaisant un PGD de bonne fonction de taux $I : G \to [0, +\infty]$.*
Soit $J : H \to [0, +\infty]$, l'application définie par

$$J(y) = inf\{I(x) : x \in G, y = g(x)\}, \quad pour \quad tout \quad y \in H$$

Alors J est une bonne fonction de taux et la suite des mesures images $\Big(\mu_n \circ g^{-1}\Big)_{n\in\mathbb{N}}$ satisfait un PGD contrôlé par J.

1.3.3 Introduction au principe de déviations modérées

Un principe de déviations modérées (PDM) n'est autre qu'un résultat intermédiaire entre le théorème limite central et le principe des grandes déviations. Les hypothèses d'un PDM sont en général moins fortes que celles d'un PGD et le PDM présente généralement une fonction de taux qui est quadratique et donc facilement manipulable.

Définition 1.4 (Principe de déviations modérées PDM) *On dit qu'une suite de variables aléatoires $(X_n)_{n\in\mathbb{N}}$ vérifie un PDM de vitesse a_n^2 (avec $1 \ll a_n \ll \sqrt{n}$) et de fonction de taux I_{PDM}, si la suite $\Big(\frac{\sqrt{n}}{a_n}(X_n - \mathbb{E}(X_n))\Big)_n$ satisfait un PGD de vitesse a_n^2 et de fonction de taux I_{PDM}.*

1.3.4 Grandes déviations dans $\mathbb{R}^d$

Un des premiers grands résultats sur la théorie des grandes déviations est le théorème de Cramèr. Soit $(S_n)_{n\geq 1}$ une suite de variables aléatoires, définie pour tout $n \geq 1$ par

$$S_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

où $(X_i)_{i\geq 1}$ est une suite de variables aléatoires indépendantes et identiquement distribuées à valeurs réelles.

Supposons que pour tout $i \geq 1$, X_i est centrée c'est à dire $\mathbb{E}X_i = 0 \quad \forall i \geq 1$, d'écart-type σ et que pour tout $t \in \mathbb{R}$, $\exp(tX_1)$ est intégrable. La loi forte des grands nombres et le théorème limite central (TCL), nous disent que pour n assez grand, S_n suit la loi gaussienne de moyenne 0 et de variance $n^{-1}\sigma^2$ notée $\mathcal{N}(0, n^{-1}\sigma^2)$. On appelle log-Laplace (aussi appélé la fonction génératrice des cumulants) de X_1 la fonction Λ, définie par

$$\Lambda(\theta) = \log \mathbb{E}\Big(\exp(\theta X_1)\Big)$$

et Λ^* sa transformée de Fenchel-Legendre définie par

$$\Lambda^*(t) = \sup_{\theta \in \mathbb{R}} \{\theta t - \Lambda(\theta)\}$$

. Rappelons que si Y est une v.a positive et $a > 0$, l'inégalité de Markov montre que :

$$\mathbb{P}\Big(Y > a\Big) \leq \frac{\mathbb{E}(Y)}{a}$$

et par suite, si Y est une v.a.réelle et $b > 0$, on a

$$\mathbb{P}\Big(Z > b\Big) \leq \frac{\mathbb{E}\phi(Z)}{\phi(b)}$$

avec ϕ une fonction croissante de $\mathbb{R}^+$ dans lui même, en supposant que $\mathbb{E}\phi(Z) < \infty$. Soit $x \geq 0$. Pour tout $t \geq 0$, on a alors,

$$\mathbb{P}(S_n - nx \geq 0) \leq \mathbb{E}\Big[e^{t(S_n - nx)}\Big] = e^{-n(tx - \Lambda(t))}$$

Comme les variables $(X_i)_i$ sont centrées, $\sup_{t \in \geq 0}\{tx - \Lambda(t)\} = \Lambda^*(x)$ et $\frac{1}{n} \log \mathbb{P}(\frac{S_n}{n} \geq x) \leq -\Lambda^*(x)$. On reconnait là une borne supérieure de grandes déviations.
On montre ainsi en completant pour la borne inférieure par une méthode de changement de mesure que la suite $(S_n/n)_{n\geq 1}$ satisfait un PGD de vitesse n et de bonne fonction de taux Λ^*. Le paragraphe suivant n'est autre qu'une généralisation du théorème de Cramèr, cela nous permet d'obtenir un PGD à partir du comportement asymptotique des fonctions génératrices des cumulants.

Théorème de Gärtner Ellis

Soit $(Z_n)_{n\geq 1}$ une suite de variables aléatoires à valeurs dans $\mathbb{R}^d$ de loi μ. On munit $\mathbb{R}^d$ du produit scalaire canonique $\langle \cdot, \cdot \rangle$. On introduit la suite des fonctions génératrices des cumulants donnée par

$$\Lambda_n(t) = \log \mathbb{E}\Big[e^{\langle t, Z_n \rangle}\Big]$$

Supposition 1 (S1) *Supposons que pour tout* $t \in \mathbb{R}^d$*, la suite* $(1/n)\Lambda_n(t)$ *converge vers* $\Lambda(t)$ *(éventuellement égale à* $+\infty$*) et que l'origine est dans l'intérieur* $\overset{\circ}{D}_\Lambda$ *de* $D_\Lambda = \{t \in \mathbb{R}^d / \Lambda(t) < +\infty\}$*. On note* Λ^* *la transformée de Fenchel-Legendre de* Λ*, définie pour tout* $x \in \mathbb{R}^d$ *par*

$$\Lambda^*(x) = \sup_{t \in \mathbb{R}^d} \{\langle t, x\rangle - \Lambda(t)\}$$

et $D_{\Lambda^*} = \{t \in \mathbb{R}^d / \Lambda^*(t) < +\infty\}$.

Rappelons les définitions suivantes

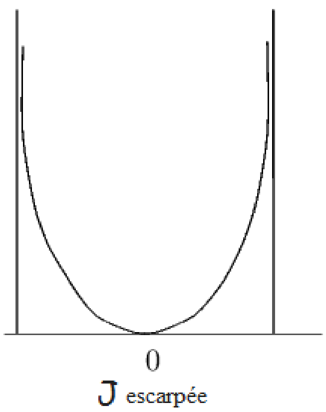
Définition 1.5 (Point exposé) *On dit que* $x \in \mathbb{R}^d$ *est un point exposé de* Λ^* *s'il existe* $t \in \mathbb{R}^d$ *tel que*

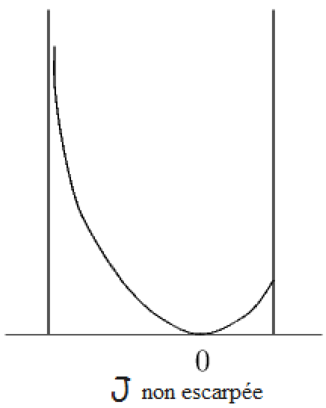
$$\forall x \neq y, \quad \langle t, x\rangle - \Lambda^*(x) > \langle t, y\rangle - \Lambda^*(y).$$

t *est appelé un hyperplan exposé.*

Définition 1.6 (Essentiellement lisse ou Escarpement) *Soit* J *une fonction convexe de* $\mathbb{R}^d$ *dans* $]-\infty, +\infty]$ *de domaine* $\mathcal{D}_J = \{x, J(x) < +\infty\}$*. On dit que la fonction* J *est Essentiellement lisse (ou escarpée) si*

- *L'intérieur de* $\mathcal{D}_J$ *est non vide (*$\overset{\circ}{\mathcal{D}}_J \neq \emptyset$*)*
- J *est différentiable sur* $\overset{\circ}{\mathcal{D}}_J$
- *Pour toute suite* (y_n) *dans* $\overset{\circ}{\mathcal{D}}_J$ *convergeant vers un point de la frontière de* $\overset{\circ}{\mathcal{D}}_J$*,* $J'(y_n)$ *converge en norme vers* $+\infty$.





Plachky et Steinebach [111] ont étendu les résultats de Cramèr pour des variables aléatoires dépendantes. On peut aussi se référer aux résultats de Dacunha-Castelle [45]. Les résultats pour des variables aléatoires dans $\mathbb{R}^d$

sont dûs à Gärtner [75]et Ellis [69]. Le théorème suivant qui porte leurs noms (pour plus de détails voir Dembo et Zeitouni [53], section 2.3) fournit un PGD.

Théorème 1.3.3 (Gärtner-Ellis) *Sous l'hypothèse* **(S1)** *et si on suppose que la log-Laplace* Λ *est semi-continue inférieurement et essentiellement lisse avec* Λ^* *sa transformée de Fenchel-Legendre, alors la suite* $(Z_n)_{n\geq 1}$ *satisfait un PGD de vitesse* n *et de bonne fonction de taux* Λ^**.*

Contre exemple du théorème Gärtner-Ellis
On s'intéresse aux grandes déviations de la suite de variables aléatoires $(Z_n)_{n\leqslant 1}$ définie par

$$Z_n = \frac{N_2^2/2 - (N_1/\sqrt{2} + \sqrt{cn})^2}{n}$$

où N_1 et N_2 sont des v.a.i.i.d de loi gaussienne centrée réduite.
Un calcul simple nous montre que

$$\mathbb{E}e^{\lambda n Z_n} = \begin{cases} \dfrac{1}{\sqrt{1-\lambda^2}} e^{-\lambda cn/(1+\lambda)} & \text{si } \lambda \in (-1,1) \\ +\infty, & \text{sinon} \end{cases}$$

On a

$$\Lambda(\lambda) = \lim_{n\to\infty} (1/n) \log \mathbb{E}e^{\lambda n Z_n} = \begin{cases} \dfrac{-\lambda c}{1+\lambda} & \text{si } \lambda \in (-1,1) \\ +\infty, & \text{sinon} \end{cases}$$

Puisque $\Lambda^{'}(1) < \infty$, la fonction Λ n'est pas essentiellement lisse (escarpé), le théorème de Gärtner-Ellis ne peut donc pas s'appliquer.
Néanmoins, la suite $\sum_{i=1}^{n} Z_n$ vérifie un PGD de vitesse n et de bonne fonction de taux donnée par

$$\Lambda^*(x) = \sup_{\lambda\in\mathbb{R}} \Big(x\lambda - \Lambda(\lambda) \Big) = \begin{cases} (\sqrt{c} - \sqrt{-x})^2, & \text{si } x \leqslant -c/4 \\ (x + c/2), & \text{si } x > -c/4 \end{cases}$$

Voir [53, Exercice 2.3.24] pour plus de détails.

1.3.5 Quelques resultats de Najim

Bien plus qu'intéressant, le théorène de Gärtner-Ellis a ses limites. En effet, il arrive qu'il ne puisse pas s'appliquer dans certains cas du fait de la non vérification des conditions d'escarpement, mais sans pour autant mettre en cause l'existence d'un PDG. Ces cas peuvent apparaître en l'absence de tous les moments exponentiels. Plusieurs travaux ont été ménés en ce sens, c'est à dire prouver l'existence d'un PGD en l'absence du théorème de Gärtner-Ellis.

On retiendra entre autre, les techniques de l'approximation discrète de la fonction de taux dans Lynch et Sethuraman [97], le changement exponentiel de probabilité dépendant du temps dans Bryc et Dembo [36], l'extension ad hoc du théorème de Gärtner-Ellis (-Baldi) dans les travaux de Bercu, Gamboa, Rouault et Zani ([9], [10], [72], [139]) et les techniques utilisant la sous-additivité et une approximation exponentielle établies par Najim ([104], [105], [106]).

Le paragraphe suivant est consacré à une partie des travaux de Najim qui sera utilisée par la suite dans la section I.0.10 de cette thèse.

Hypothèses et Notations

Commençons par introduire quelques notations. Soit $\mathcal{X}$ un espace vectoriel topologique muni de la tribu borélienne $\mathcal{B}(\mathcal{X})$ et soit R une probabilité sur $\mathcal{X}$. On appelle $\mathcal{BV}(\mathcal{X}, \mathbb{R}^d)$ (noté $\mathcal{BV}$) l'ensemble des fonctions continues bornées de $\mathcal{X}$ à valeurs dans $\mathbb{R}^d$, $L(\mathcal{X}, \mathbb{R}^d)$ (noté L) l'ensemble de fonctions R-intégrables de $\mathcal{X}$ à valeurs dans $\mathbb{R}^d$, et $\mathcal{P}(\mathcal{X})$ l'ensemble des probabilités sur $\mathcal{X}$. $|\cdot|$ désigne une norme sur $\mathcal{X}$ et $\|\cdot\|$ la norme du supremum sur l'ensemble des fonctions continues bornées de $\mathcal{X}$ à valeurs dans $\mathbb{R}^d$, c'est à dire $\|f(x)\| = \sup_{x\in\mathcal{X}} |f(x)|$. Et comme toujours, δ_a définit la masse de Dirac en a.

Cas de variables aléatoires indépendantes et identiquement distribuées (v.a.i.i.d)

Soit $(Z_i)_{i\in\mathbb{N}}$ une suite de variables aléatoires à valeurs dans $\mathbb{R}^d$ indépendantes et identiquement distribuées selon la loi $\mathbb{P}_z$. $\mathbb{P}_z$ est l'image de $\mathbb{P}$ par z, c'est à dire que $\mathbb{P}(Z_i \in A) = \mathbb{P}_z(A)$ où $Z_i : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$. On suppose que

Condition 1 (C1) *Il existe un* $\alpha > 0$ *tel que :* $\mathbb{E}e^{\alpha|Z|} < +\infty$

Condition 2 (C2) *R est une probabilité sur $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ satisfaisant $R(U) > 0$ pour tout ouvert non vide U.*

Condition 3 (C3) *La famille $(x_i^n; 1 \le i \le n, n \ge 1) \subset \mathcal{X}$ satisfait*

$$\frac{1}{n}\sum_{i=1}^{n} \delta_{x_i^n} \xrightarrow[n\to\infty]{conv. \quad etroite} R \quad avec \quad R \in \mathcal{P}(\mathcal{X}). \tag{1.10}$$

Soit $f : \mathcal{X} \to \mathbb{R}^{m\times d}$ une fonction (matricielle) continue bornée, alors toutes les composantes de la matrice f sont des fonctions continues bornées et pour tout $z \in \mathbb{R}^d$, on a le produit matriciel suivant

$$f(x) \cdot z = \begin{pmatrix} f_1(x) \cdot z \\ \vdots \\ f_m(x) \cdot z \end{pmatrix}$$

où $(\cdot)$ désigne le produit scalaire sur $\mathbb{R}^d$ et $f_i \in \mathcal{BV}$ est la i^e ligne de la matrice f.
Si $u : \mathcal{X} \to \mathbb{R}^d$ est une fonction mesurable, on note

$$\int_{\mathcal{X}} f(x) \cdot u(x) \mathrm{R}(dx) = \begin{pmatrix} \int_{\mathcal{X}} f_1(x) \cdot u(x) \mathrm{R}(dx) \\ \vdots \\ \int_{\mathcal{X}} f_m(x) \cdot u(x) \mathrm{R}(dx) \end{pmatrix}$$

Introduisons la moyenne pondérée dont on va étudier le PGD,

$$\langle L_n, f \rangle \triangleq \frac{1}{n} \sum_{i=1}^{n} f(x_i^n) \cdot Z_i.$$

Par convention, x est à valeurs dans $\mathcal{X}$, y et θ sont des éléments de $\mathbb{R}^m$ et z et λ des éléments de $\mathbb{R}^d$.

Théorème 1.3.4 *On suppose que C1, C2 et C3 sont vérifiées. Alors la famille*

$$\langle L_n, f \rangle \triangleq \frac{1}{n} \sum_{i=1}^{n} f(x_i^n) \cdot Z_i$$

satisfait un PGD dans $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$, de bonne fonction de taux

$$I_f(y) = \sup_{\theta \in \mathbb{R}^m} \{\theta \cdot y - \int_{\mathcal{X}} \Lambda[\sum_{i=1}^{m} \theta_i \cdot f_i(x)] \mathrm{R}(dx)\} \qquad \forall y \in \mathbb{R}^m$$

où Λ désigne la fonction génératrice des cumulants associée à la suite Z_i.

Démonstration 1.1 *La preuve du théorème 1.3.4 est fournie dans celle du théorème 3.2 de [104] dans la section 3.2.*

Cas de variables aléatoires indépendantes mais non identiquement distribuées

Dans cette partie il est question d'étudier les grandes déviations de la moyenne empirique des variables aléatoires indépendantes mais non identiquement distribuées (voir [104]).
Soit $\tau(z) = e^{|z|} - 1, z \in \mathbb{R}^d$, on définit par

$$\begin{aligned} \mathcal{P}_\tau(\mathbb{R}^d) &= \{P \in \mathcal{P}(\mathbb{R}^d), \exists a > 0; \int_{\mathbb{R}^d} \tau\left(\frac{z}{a}\right) \mathrm{P}(dz) < \infty\} \\ &= \{P \in \mathcal{P}(\mathbb{R}^d), \exists \alpha > 0; \int_{\mathbb{R}^d} e^{\alpha|z|} \mathrm{P}(dz) < \infty\} \end{aligned} \tag{1.11}$$

l'ensemble des probabilités ayant un moment exponentiel. Introduisons une distance qu'on appelle la distance d'Orlicz-Wasserstein définie sur $\mathcal{P}_\tau(\mathbb{R}^d)$ par

$$d_{OW}(P,Q) = \inf_{\eta \in M(P,Q)} \inf\left\{a > 0; \int_{\mathbb{R}^d \times \mathbb{R}^d} \tau\left(\frac{z-z'}{a}\right)\eta(dzdz') \leqslant 1\right\}$$

l'infinimum étant pris sur $\mathcal{M}(P,Q)$, l'ensemble des probabilités dont les marginales sont P et Q.
On suppose dans cette section que C2 et C3 sont vérifiées et nous introduisons également trois autres hypothèses.

Condition 4 (C4) *$\mathcal{X}$ est un espace compact.*

Condition 5 (C5) *$(P(x,\cdot), x \in \mathcal{X}) \subset \mathcal{P}_\tau(\mathbb{R}^d)$ est une famille de probabilités incluse dans $\mathcal{P}_\tau(\mathbb{R}^d)$. On suppose que pour chaque borélien $\mathcal{B}$ de $\mathbb{R}^d$, l'application $x \mapsto P(x,\mathcal{B})$ et que la fonction :*

$$\begin{aligned} \Gamma : &\mathcal{X} \to \mathcal{P}_\tau(\mathbb{R}^d) \\ &x \mapsto P(x,\cdot) \end{aligned}$$

est continue lorsque $\mathcal{P}_\tau(\mathbb{R}^d)$ est munie de la topologie induite par la distance d_{OW}.

Condition 6 (C6) *Étant donnée une famille $(x_i^n; 1 \leq i \leq n, n \geq 1)$ à valeurs dans $\mathcal{X}$ et une famille de probabilités $(P(x,\cdot), x \in \mathcal{X}) \subset \mathcal{P}_\tau(\mathbb{R}^d)$, on considère une suite de variables aléatoires (Z_i^n) indépendantes à valeurs dans $\mathbb{R}^d$ et où chaque variable Z_i^n est distribuée selon la loi $P(x_i^n,\cdot) = P_{x_i^n}$.*

Théorème 1.3.5 *Soit $f : \mathcal{X} \to \mathbb{R}^{m\times d}$ une fonction continue bornée. Supposons que C2, C3, C4, C5 et C6 sont vérifiées. Alors, la famille*

$$\langle L_n, f\rangle = \frac{1}{n}\sum_{i=1}^n f(x_i^n)\cdot Z_i$$

satisfait un PGD dans $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$, de bonne fonction de taux J donnée par

$$J(f) = \int_{[0,1]} \Lambda^*(x, f_a'(x))\mathrm{d}x + \int_{[0,1]} \rho(x, f_s'(x))\mathrm{d}\theta(x)$$

où θ est une mesure quelconque à valeurs réelles positives par rapport à laquelle μ_s^f est absolument continue et $f_s' = \frac{d\mu_s^f}{d\theta}$, où

$$\Lambda^*(x,z) = \sup_{\lambda\in\mathbb{R}^d}\left\{\langle\lambda, z\rangle - \Lambda(x,\lambda)\right\}, \quad \forall z \in \mathbb{R}^d$$

avec $\Lambda(x,\lambda) = \log\int_{\mathbb{R}^d} e^{\langle\lambda,z\rangle}P(x,dz), \quad \forall\lambda \in \mathbb{R}^d$ et la fonction de récession $\rho(x,z)$ de $\Lambda^*(x,z)$ est définie par : $\rho(x,z) = \sup\{\langle\lambda,z\rangle, \lambda \in D_x\}$ avec $D_x = \{\lambda \in \mathbb{R}^d, \Lambda(x,\lambda) < \infty\}$.

Démonstration 1.2 *La preuve du théorème 1.3.5 est fournie dans celle du théorème 5.3 de la section 5.2 de [104].*

1.3.6 La delta méthode pour les grandes déviations de GAO et ZHAO

La méthode delta pour les grandes déviations de GAO et ZHAO [73] n'est autre qu'une extension du principe de contraction, elle permet donc de transférer un PGD d'un cas à un autre. Introduisons la notion d'Hadamard différentiable. Soient $\mathcal{X}$ et $\mathcal{Y}$ deux espaces topologiques métrisables et linéaires. On dit qu'une fonction ϕ définit sur $\mathcal{D}_\phi \subset \mathcal{X}$ à valeurs dans $\mathcal{Y}$ est Hadamard différentiable en x tangentiellement à $\mathcal{D}_\phi$ s'il existe une application continue $\phi'_x : \mathcal{X} \mapsto \mathcal{Y}$ telle que pour toute suite t_n convergeant vers 0+ et toute suite h_n convergeant vers h dans $\mathcal{X}$ telles que pour tout entier n, $x + t_n h_n \in \mathcal{D}_\phi$, alors on a

$$\lim_{n\to\infty} \frac{\phi(x + t_n h_n) - \phi(x)}{t_n} = \phi'_x(h).$$

Théorème 1.3.6 *Soient $\mathcal{X}$ et $\mathcal{Y}$ deux espaces topologiques métrisables et linéaires. Soit $\phi : \mathcal{D}_\phi \subset \mathcal{X} \mapsto \mathcal{Y}$ une fonction Hadamard différentiable en θ tangentiellement à $\mathcal{D}_0$, avec $\mathcal{D}_0$ et $\mathcal{D}_\phi$ deux sous-ensembles de $\mathcal{X}$. Soient $\{(\Omega_n, \mathcal{F}_n, \mathbb{P}_n), n \geq 1\}$ une suite d'espaces probabilisés, $X_n : \Omega_n \mapsto \mathcal{D}_\phi$ une suite de fonctions de Ω_n à valeurs dans $\mathcal{D}_\phi$. Soient deux suites de nombres réels positifs $\{r_n, n \geqslant 1\}$ et $\{\lambda(n), n \geqslant 1\}$ telles que $r_n \to +\infty$ et $\lambda(n) \to +\infty$ quand $n \to +\infty$.*

Si la suite $\{r_n(X_n - \theta), n \geqslant 1\}$ satisfait un PGD de vitesse $\lambda(n)$ et de fonction de taux I telle que $\{I < +\infty\} \subset \mathcal{D}_0$, alors la suite $\{r_n(\phi(X_n) - \phi(\theta)), n \geqslant 1\}$ satisfait un PGD de vitesse $\lambda(n)$ et de fonction de taux $I_{\phi'_\theta}$ telle que pour tout $y \in \mathcal{Y}$, on a

$$I_{\phi'_\theta}(y) = inf\{I(x); \phi'_\theta(x) = y\}.$$

1.4 Grandes déviations et déviations modérées pour quelques estimateurs

Dans cette thèse, on s'est notamment intéressé à l'étude des grandes déviations et des déviations modérées pour certains estimateurs qui sont beaucoup utilisés en finance (les estimateurs des paramètres de diffusion par exemple). En effet, dans le monde d'aujourd'hui où les processus de diffusion jouent un rôle très important dans de nombreux domaines : mathématiques financières, santé ... Il est important de pouvoir estimer le coefficient de diffusion, appelé généralement la volatilité en finance, associé à ces équations différentielles stochastiques (EDS). Et malgré un fort intérêt et un développement assez considérable ces 20 dernières années sur ces études, ce problème reste largement ouvert car il existe seulement peu de résultats pour un champs aussi vaste.

Mes travaux de recherche s'inscrivent ainsi dans ce cadre. Les travaux de Djellout & al.[59] sur les grandes déviations et les déviations modérées de la volatilité realisée et leurs références ont d'une part motivés nos travaux sur les grandes déviations et les déviations modérées de la covolatilité realisée. Et à la suite de ce premier article, une question naturelle nous est venue à l'idée, avons-nous des théorèmes de types grandes déviations et déviations modérees pour un vecteur formé des volatilités et de la covolatilité ? L'inexistence ou la non pertinence des résultats existants ont motivé la suite de nos travaux d'autre part.

1.4.1 Grandes déviations : cas synchrone

La covariance de la rentabilité d'un actif financier et ses statistiques jouent un rôle très important dans de nombreux problèmes aussi théoriques que pratiques que rencontrent les financiers. De la même façon que l'étude de la volatilité réalisée (voir Djellout & al.[59] et les références qui y sont citées), utiliser des données de haute fréquence dans le calcul de la covariance entre deux actifs conduit au concept de la covariance réalisée. La quantification de la (co-)volatilité intégrée joue un rôle clé dans l'optimisation des portefeuilles financiers et dans la gestion des risques. Cela stimule un intérêt croissant pour les méthodes d'estimation pour ces modèles.
Lorsque le nombre d'observation à haute fréquence dans un intervalle de temps fixe (un jour par exemple) augmente à l'infini, il est donc assez naturel d'utiliser le comporte asymptotique pour modéliser ces cas.

La motivation principale de ces travaux présentés dans cette partie est d'établir des théorèmes de type grandes déviations et déviations modérées pour l'estimateur de la covariance et par la suite montrer ces résultats pour le vecteur formé des volatilités et de la covalatilité pour deux processus de

diffusion observés de manière synchronisée car on a été surpris de constater un petit nombre (voir l'inexistence) de résultats dans cette direction malgré ces vingt dernières années d'étude dans ce domaine.
Cela peut s'expliquer du fait que les techniques habituelles (telles que les techniques de Gärtner Ellis) ne s'appliquent que pour des cas simples et ne fonctionnent pas en général pour des cas un peu complexe (comme le notre). La difficulté dans ces cas précis est de trouver une alternative à ces techniques habituelles.
Cependant, plusieurs techniques ont été développées pour montrer des grandes déviations dans les cas où la méthode de Gärtner Ellis ne marche pas : les techniques dans Dembo [50], Gamboa et al. [72] pour établir des grandes déviations pour les formes quadratiques de processus gaussiens. D'autres méthodes ont aussi été établies pour des grandes déviations de variables aléatoires pondérées, principalement dans Maïda et al. [98] et Zani [108]-[139]-[140].

Nous modélisons l'évolution d'une variable d'état observable par un processus stochastique $X_t = (X_{1,t}, X_{2,t}), t \in [0,1]$. Dans les applications financières, X_t peut être considéré comme un taux d'intérêt à court terme, un indice boursier, un taux de change ou le logarithme du prix d'un actif en finance.
On suppose que les deux processus de diffusion $X_{1,t}$ et $X_{2,t}$ sont définis sur un espace de probabilité filtré $(\Omega, \mathfrak{F}, (\mathfrak{F}_t), \mathbb{P})$ et vérifient le processus de Itô

$$\left\{\begin{array}{lllll} dX_{1,t} & = & \sigma_{1,t}dB_{1,t} & + & b_1(t,X_t)dt \\ dX_{2,t} & = & \sigma_{2,t}dB_{2,t} & + & b_2(t,X_t)dt \end{array}\right. \tag{1.12}$$

où les mouvements browniens $B_1 = (B_{1,t})_{t\in[0,1]}$ and $B_2 = (B_{2,t})_{t\in[0,1]}$ sont corrélés entre eux avec $cov(B_{1,t}, B_{2,t}) = \int_0^t \rho_s \mathrm{d}s$, b_1 et b_2 sont des processus adaptés représentant les vitesses moyennes (aussi appelées les dérives).
On peut aussi définir un troisième mouvement brownien $B_3 = (B_{3,t})_{t\in[0,1]}$ tel que B_1 et B_3 soient des mouvements browniens indépendants et tel que $dB_{2,t} = \rho_t dB_{1,t} + \sqrt{1-\rho_t^2}dB_{3,t}$,
Dans nos travaux, on se limite seulement au cas où σ_1, σ_2 et ρ sont des fonctions déterministes. Les fonctions σ_1 et σ_2 ne prennent que des valeurs positives tandis que ρ prend ses valeurs dans $]-1,1[$. On suppose que $X_{1,t}$ et $X_{2,t}$ sont observés de manière synchrone et on note $[X_\ell]_t = \int_0^t \sigma_{\ell,s}^2 \mathrm{d}s$ avec $\ell = 1,2$ la variation quadratique de X_ℓ et $\langle X_1, X_2\rangle_t = \int_0^t \sigma_{1,s}\sigma_{2,s}\rho_s \mathrm{d}s$ la covariation déterministe de $X_{1,t}$ et $X_{2,t}$.
$\sigma_{\ell,t}$ et ρ_t sont supposés inconnus et on cherche à les estimer à partir d'un échantillon de $(X_t)_{t\in[0,1]}$. Plus précisement, on cherche à estimer la covariation inconnue $\langle X_1, X_2\rangle_t$ de X d'une part et le vecteur de la (co-)variation

inconnu d'autre part et donné par

$$[V]_t = ([X_1]_t, [X_2]_t, \langle X_1, X_2 \rangle_t)^T.$$

Ce type de problème apparaît naturellement en mathématiques financières, et on appelle $[X_\ell]_t$ la volatilité, $\langle X_1, X_2 \rangle_t$ la covolatilité et V_t la (co-)volatilité. Étant donnés des observations discrètes également espacées $(X_{1,t_k^n}, X_{2,t_k^n}, k = 1, \cdots, n)$ dans l'intervalle $[0,1]$ (avec $t_k^n = k/n$), $C_t^n(X)$ la covariance réalisée est un estimateur naturel de $\langle X_1, X_2 \rangle_t$ et celui de V_t est $V_t^n(X)$ le vecteur de la (co-)variance

$$V_t^n(X) = \Big(Q_{1,t}^n(X), Q_{2,t}^n(X), C_t^n(X)\Big)^T$$

avec des quantités $Q_{\ell,t}^n(X)$ et $C_t^n(X)$ pour $\ell = 1, 2$, données par

$$Q_{\ell,t}^n(X) = \sum_{k=1}^{[nt]} (\Delta_k^n X_\ell)^2 \qquad C_t^n(X) = \sum_{k=1}^{[nt]} (\Delta_k^n X_1)(\Delta_k^n X_2)$$

où $\Delta_k^n X_\ell = X_{\ell,t_k^n} - X_{\ell,t_{k-1}^n}$ avec $\tau_n = \{t_k^n, 0 \leqslant k \leqslant n\}$ une subdivision uniforme de $[0,1]$ à n parties.
Soit P_c la fonction définie pour tout $-1 < c < -1$ et pour tout $\lambda = (\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3$ par

$$P_c(\lambda) := \begin{cases} -\dfrac{1}{2}\log\left(\dfrac{(1-2\lambda_1(1-c^2))(1-2\lambda_2(1-c^2)) - (\lambda_3(1-c^2)+c)^2}{1-c^2}\right) \\ \qquad\qquad\qquad\qquad \text{si} \quad \lambda \in \mathcal{D}_c \\ +\infty, \quad \text{sinon.} \end{cases} \tag{1.13}$$

avec

$$\mathcal{D}_c = \left\{\lambda \in \mathbb{R}^3, \quad \max_{l=1,2} \lambda_l < \frac{1}{2(1-c^2)} \quad \textit{and} \quad \prod_{l=1}^{2}(1-2\lambda_l(1-c^2)) > (\lambda_3(1-c^2)+c)^2\right\}$$

La fonction P_c joue un rôle vital dans le calcul de la fonction génératrice des moments. Elle représente la log-Laplace de la variable aléatoire $\lambda_1 X^2 + \lambda_2 Y^2 + \lambda_1\lambda_3 XY$ où X, Y sont des variables aléatoires gaussiennes corrélées avec un coefficient c.
On note P_c^* la transfomation de Legendre de P_c définie pour tout $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ par

$$P_c^*(x) = \begin{cases} \log\left(\dfrac{\sqrt{1-c^2}}{\sqrt{x_1x_2 - x_3^2}}\right) - 1 - \dfrac{x_1 + x_2 - 2cx_3}{2(1-c^2)} \\ \qquad \text{si} \quad x_1 > 0, \quad x_2 > 0, \quad x_1x_2 > x_3^2 \\ +\infty, \quad \text{sinon.} \end{cases} \tag{1.14}$$

On note aussi la fonction P^* définie par

$$P^*(x) = \begin{cases} \dfrac{1}{2}(x - 1 - \log x) & \text{si} \quad x > 0 \\ \\ +\infty & \text{si} \quad x \leq 0, \end{cases} \tag{1.15}$$

et qui représente la transformée de Legendre de la fonction P donnée par

$$P(\lambda) = \begin{cases} -\dfrac{1}{2}\log(1 - 2\lambda) & \text{si } \lambda < \dfrac{1}{2} \\ +\infty, & \text{sinon.} \end{cases} \tag{1.16}$$

Soit pour tout entier d, on note $\mathcal{B}V_d([0,1], \mathbb{R}^d)$ l'espace des fonctions à variations bornées sur $[0,1]$. On identifie $\mathcal{B}V_d$ à $\mathcal{M}_d([0,1])$, l'espace des mesures vectorielles à valeurs dans $\mathbb{R}^d$. Cette identification sera faite de la mannière usuelle : pour $f \in \mathcal{B}V_d$, il lui correspond ν^f caractérisée par $\nu^f([0,1]) = f(t)$. L'espace $\mathcal{C}_d([0,1])$ des fonctions continues bornées de $[0,1]$ à valeurs dans $\mathbb{R}^d$, est l'espace topologique dual de $\mathcal{B}V_d$. On munit $\mathcal{B}V_d$ de la topologie de la convergence faible $\sigma(\mathcal{B}V_d, \mathcal{C}_d([0,1]))$.
Soit $f \in \mathcal{B}V_d$ et ν^f sa mesure associée dans $\mathcal{M}_d([0,1])$. Considérons la décomposition de Lebesgue de ν^f, $\nu^f = \nu_a^f + \nu_s^f$ où ν_a^f désigne la partie absolument continue de ν^f par rapport à dx et ν_s^f sa partie singulière. On note $f_a(t) = \nu_a^f([0,t])$ et $f_s(t) = \nu_s^f([0,t])$.

Soit $\mathcal{H}_d$ l'espace de Banach des fonctions croissantes à valeurs dans $\mathbb{R}^d$) càdlàg γ sur $[0,1]$ avec $\gamma(0) = 0$,muni de la norme de la convergence uniforme et de la $\sigma-$algèbre $\mathcal{B}^s$ engendrée par les coordonnées de $\{\gamma(t), 0 \leqslant t \leqslant 1\}$.

Dans un deuxième chapitre, nous présentons les résultats que nous avons obtenus, en colloboration avec H. Djellout [61], sur les grandes déviations de l'estimateur de la covariance $\mathbf{C}_1^n(X)$. Les preuves découlent essentiellement des résultats obtenus dans Djellout & al. [59] et du principe de contraction.

Le troisième chapitre est consacré aux travaux récents en colloboration avec H. Djellout et A. Guillin [58], dans lequel on considère les grandes déviations pour le vecteur de la variance et de la covariance $V_t^n(X)$. Nous présentons les résultats obtenus dans le cadre de cette collaboration.

Théorème 1.4.1 *Supposons que pour $\ell = 1, 2$ $b_\ell(\cdot,\cdot) \in L^\infty(dt \otimes \mathbb{P})$ et que les fonctions $t \to \sigma_{\ell,t}$ et $t \to \rho_t$ sont continues.*
Alors $\mathbb{P}(V_1^n \in \cdot)$ satisfait un PGD sur $\mathbb{R}^3$ de vitesse n et de bonne fonction de taux donnée par

$$I_{VGD} = \sup_{\lambda \in \mathbb{R}^3} \Big(\langle \lambda, x \rangle - \Lambda(\lambda) \Big)$$

avec $\Lambda(\lambda) = \int_0^1 P_{\rho_t}(\lambda_1\sigma_{1,t}^2, \lambda_2\sigma_{2,t}^2, \lambda_3\sigma_{1,t}\sigma_{2,t})\mathrm{d}t$ *où la fonction* P_c *est donnée par* (1.13).

Remarque 1.4.1 *Dans les travaux de Djellout et al. [59] et Kanaya et Otsu [85], il a été prouvé que* $Q_{\ell,t}^n(X)$ *pour* $\ell = 1,2$ *donnée par la relation* (1.2)*, vérifie un PGD en dépit du théorème de Gärtner-Ellis.*
Soit Υ_n *la suite des fonctions génératrices des cumulants définie par* $\Upsilon_n(\lambda) = \log \mathbb{E}\exp(\lambda n Q_{\ell,t}^n(X))$ *et de limite donnée par*

$$\Upsilon(\lambda) = \lim_{n\to\infty}(1/n)\Upsilon_n(\lambda) = \begin{cases} -\dfrac{1}{2}\displaystyle\int_0^1 \log(1-2\lambda\sigma_{\ell,t}^2)\mathrm{d}t & si\ \lambda < \dfrac{1}{2\|\sigma_\ell^2\|_\infty} \\ +\infty, \qquad si\ \lambda \geqslant \dfrac{1}{2\|\sigma_\ell^2\|_\infty} \end{cases}$$

Si on considère un processus pour lequel $\sigma_{\ell,t}^2 = \tau_\theta(t)$ *tel que pour tout* $\theta \geqslant 1$*, on a*

$$\tau_\theta(t) = \begin{cases} 1 \quad si\ t = 0 \\ 1 - \exp(-|t|^{-\theta}), \qquad si\ t \in (0,1] \end{cases}$$

Alors, nous obtenons que

$$\Upsilon(\lambda) = \begin{cases} -\dfrac{1}{2}\displaystyle\int_0^1 \log(1-2\lambda[1-\exp(-|t|^{-\theta})])\mathrm{d}t & si\ \lambda < \dfrac{1}{2} \\ +\infty, \qquad si\ \lambda \geqslant \dfrac{1}{2} \end{cases}$$

et que

$$\Upsilon'(\lambda) = \int_{[0,1]} \frac{1-\exp(-|t|^{-\theta})}{1-2\lambda[1-\exp(-|t|^{-\theta})]}\mathrm{d}t.$$

Notons que $|\Upsilon'(\lambda)| \leqslant \int_{[0,1]}[\exp(-|t|^{\theta})-1]\mathrm{d}t = \theta^{-1}\int_1^{+\infty}\exp(-y)y^{-(\theta+1)/\theta}\mathrm{d}y < \infty$ *pour tout* $\lambda \leqslant \frac{1}{2}$.
Par conséquent, pour toute suite (λ_n) *telle que* $\lambda_n < \frac{1}{2}$ *et* $\lambda_n \to \lambda \leqslant \frac{1}{2}$*, on obtient que* $\lim_{n\to\infty}|\Upsilon'(\lambda_n)| < \infty$*, ce qui implique que* Υ *n'est pas essentiellement lisse. Donc on ne peut pas appliquer le théorème de Gärtner-Ellis.*

Si Λ est essentiellement lisse (escarpé), c'est à dire si pour toute suite λ de $\mathcal{D}_\Lambda$ le domaine de définiton de Λ convergeant vers un point de $\partial\mathcal{D}_\Lambda$ la frontière de $\mathcal{D}_\Lambda$, on a

$$\lim_{\lambda\to\partial\mathcal{D}_\Lambda}|\nabla\Lambda(\lambda)| = +\infty, \tag{1.17}$$

alors la preuve du théorème 1.4.1 est directement établie par le théorème de Gärtner Ellis. Un calcul simple du gradient nous montre que

$$|\nabla\Lambda(\lambda)| = \int_0^1 \beta_t \frac{\sigma_{1,t}^2(1-2\sigma_{2,t}^2\beta_t\lambda_2) + \sigma_{2,t}^2(1-2\sigma_{1,t}^2\beta_t\lambda_1) + \sigma_{1,t}\sigma_{2,t}|\rho_t + \beta_t\sigma_{1,t}\sigma_{2,t}\lambda_3|}{(1-2\sigma_{1,t}^2\beta_t\lambda_1)(1-2\sigma_{2,t}^2\beta_t\lambda_2) - (\rho_t+\beta_t\sigma_{1,t}\sigma_{2,t}\lambda_3)^2}\mathrm{d}t.$$

où $\beta_t = 1 - \rho_t^2$.
Hors, si les fonctions σ_ℓ et ρ sont constants, alors la relation (1.17) est vérifiée, on en deduit donc la preuve de notre théorème.
Les fonctions σ_ℓ et ρ étant supposées déterministes dans nos travaux, il est très difficile d'établir la condition d'escarpement 1.17. Toute la difficulté de notre preuve réside dans le cas où

$$\lim_{\lambda \to \partial \mathcal{D}_\Lambda} |\nabla \Lambda(\lambda)| < +\infty,$$

car les techniques de Gärtner Ellis ne peuvent plus s'appliquer dans ces conditions.

Théorème 1.4.2 *Sous les mêmes hypothèses que le théorème précedent, la suite $\mathbb{P}(V_\cdot^n \in \cdot)$ satisfait un PGD sur $\mathcal{B}V_3$ de vitesse n et de bonne fonction de taux J_{VGD} donnée par*

$$\begin{aligned} J_{VGD}(f) \quad &= \quad \int_0^1 P^*_{\rho_t}\left(\frac{f'_{1,a}(t)}{\sigma^2_{1,t}}, \frac{f'_{2,a}(t)}{\sigma^2_{2,t}}, \frac{f'_{3,a}(t)}{\sigma_{1,t}\sigma_{2,t}}\right) \mathrm{d}t \\ + \int_0^1 & \frac{\sigma^2_{2,t} f'_{1,s}(t) + \sigma^2_{1,t} f'_{2,s}(t) - 2\rho_t \sigma_{1,t}\sigma_{2,t} f'_{3,s}(t)}{2\sigma^2_{1,t}\sigma^2_{2,t}(1-\rho_t^2)} 1_{[t; f'_{1,s}>0, f'_{2,s}>0, (f'_{3,s})^2 < f'_{1,s} f'_{2,s}]} \,\mathrm{d}\theta(t), \end{aligned}$$

pour tout $f = (f_1, f_2, f_3) \in \mathcal{B}V_3$ avec P^*_c donnée par (1.14) et θ est une mesure quelconque à valeurs réelles positives par rapport à laquelle ν_s^f est absolument continue et $f'_s = d\nu_s^f / d\theta = (f'_{1,s}, f'_{2,s}, f'_{3,s})$.

Remarque 1.4.2 *Si on suppose que pour $\ell = 1, 2$, les fonctions σ_ℓ, ρ sont constants et que $b_\ell(\cdot,\cdot) \in L^\infty(dt \otimes \mathbb{P})$. Alors $\mathbb{P}(V_1^n \in \cdot)$ satisfait un PGD sur $\mathbb{R}^3$ de vitesse n et de bonne fonction de taux donnée par*

$$I^{Cons}_{VGD} = P^*_\rho\left(\frac{x_1}{\sigma_1^2}, \frac{x_2}{\sigma_2^2}, \frac{x_3}{\sigma_1\sigma_2}\right)$$

Remarquons que la définition de f'_s est $\theta-$dépendant. Cependant, par homogénéité, J_{VGD} ne dépend pas de θ. On peut choisir $\theta = |f_{1,s}|+|f_{2,s}|+|f_{3,s}|$ avec $|f_{i,s}| = f^+_{i,s} + f^-_{i,s}$ où $f_{i,s} = f^+_{i,s} - f^-_{i,s}$ est la décomposition de Hahn-Jordan.

Comme indiqué ci-dessus, la methode de Gärtner Ellis ne marche pas du fait que le comportement limite de la fonction génératrice des moments reste indéterminé à un certain point de la frontière du domaine de définition. Dès lors, pour la preuve du théorème 1.4.1 et celle du théorème 1.4.2, nous avons donc utilisé une nouvelle approche basée sur la sous additivité et sur la notion d'approximation exponentielle inspirée des résultats de Najim ([104], [105], [106]).

Théorème 1.4.3 *Supposons que pour* $\ell = 1,2$ $b_\ell(\cdot,\cdot) \in L^\infty(dt \otimes \mathbb{P})$, *que* $\sigma^2_{\ell,t}(1-\rho^2_t) \in L^2([0,1],dt)$. *Soit* (b_n) *une suite de nombres positifs telle que*

$$b_n \xrightarrow[n\to\infty]{} \infty \quad et \quad \frac{b_n}{\sqrt{n}} \xrightarrow[n\to\infty]{} 0 \quad et \quad \sqrt{n} b_n \max_{\ell=1,2} \max_{1\le k\le n} \int_{(k-1)/n}^{k/n} \sigma^2_{\ell,t} dt \xrightarrow[n\to\infty]{} 0$$

Cette condition est aussi vérifiée si pour un certain $p > 2$, $\sigma^2_{\ell,t} \in L^p([0,1],dt)$ *et* $b_n = O\left(n^{\frac{1}{2}-\frac{1}{p}}\right)$.
Alors, la suite $\mathbb{P}\left(\frac{\sqrt{n}}{b_n}(V^n_\cdot - [V]_\cdot) \in \cdot\right)$ *satisfait un PGD sur* $\mathcal{H}_3$ *de vitesse* b_n^2 *et de bonne fonction de taux* J_{VDM} *donnée par*

$$J_{VDM}(\phi) = \begin{cases} \displaystyle\int_0^1 \frac{1}{2} \left\langle \dot\phi(t), \overline{\Sigma}_t^{-1} \cdot \dot\phi(t) \right\rangle dt & if \quad \phi \in \mathcal{AC}_0([0,1]) \\ +\infty, & otherwise, \end{cases} \tag{1.18}$$

où $\overline{\Sigma}_t$ *donnée précédemment, est inversible avec* $\det(\overline{\Sigma}_t) = \frac{1}{2}\sigma^6_{1,t}\sigma^6_{2,t}(1-\rho^2_t)^3$ *et son inverse* $\overline{\Sigma}_t^{-1}$ *est donné par*

$$\overline{\Sigma}_t^{-1} = \frac{1}{\det(\overline{\Sigma}_t)} \begin{pmatrix} \frac{1}{2}\sigma^2_{1,t}\sigma^6_{2,t}(1-\rho^2_t) & \frac{1}{2}\sigma^4_{1,t}\sigma^4_{2,t}\rho^2_t(1-\rho^2_t) & -\sigma^3_{1,t}\sigma^5_{2,t}\rho_t(1-\rho^2_t) \\ \frac{1}{2}\sigma^4_{1,t}\sigma^4_{2,t}\rho^2_t(1-\rho^2_t) & \frac{1}{2}\sigma^6_{1,t}\sigma^2_{2,t}(1-\rho^2_t) & -\sigma^5_{1,t}\sigma^3_{2,t}\rho_t(1-\rho^2_t) \\ -\sigma^3_{1,t}\sigma^5_{2,t}\rho_t(1-\rho^2_t) & -\sigma^5_{1,t}\sigma^3_{2,t}\rho_t(1-\rho^2_t) & \sigma^4_{1,t}\sigma^4_{2,t}(1-\rho^4_t) \end{pmatrix},$$

et $\mathcal{AC}_0 = \{\phi : [0,1] \to \mathbb{R}^3$ *absolument continue avec* $\phi(0) = 0\}$.

Les résultats précédents imposent la bornitude des dérives *i.e.* pour $\ell = 1,2$ $b_\ell(t,X) \in L^\infty(dt \otimes \mathbb{P})$, cela nous permet de revenir facilement au cas sans dérives. Cette condition est-elle nécessaire ? On a vérifié que sous une condition de Lipschitz ou sous une croissance linéaire de $b_\ell(t,X)$, nos résultats restent vrais.

1.4.2 Grandes déviations pour les estimateurs de la (co)volatilité : cas non-synchrone

Comme indiqué ci-dessus, dans l'estimation de la volatilité en finance, nous modélisons souvent des observations à des temps non régulièrement espacés. En effet, dans les applications financières, les données de transactions réelles sont rélévées à des heures irrégulières, d'une manière non synchrone, c'est à dire que deux prix d'une transaction ne sont généralement pas observées aux mêmes instants.

On suppose que dans un espace de probabilité filtré $(\Omega, \mathfrak{F}, (\mathfrak{F}_t), \mathbb{P})$, on dispose de deux processus de diffusion $X_{1,t}$ et $X_{2,t}$, $t \in [0,T]$ donnés par la relation (1.12) et observés de manière asynchrone. Explicitement, cela se traduit par le fait que $(X_{1,s_i}, X_{2,t_j})_{i=0,\cdots,n;j=0,\cdots,m}$ avec $0 = s_0 < s_1 < \cdots < s_n = T$, $0 = t_0 < t_1 < \cdots < t_m = T$ avec $m, n \in \mathbb{N}$ représente une suite d'observations non-synchronisées. Récemment, dans un travail collectif avec H. Djellout, A. Guillin et H. Jiang, nous nous sommes intéressés dans un premier temps à l'étude des déviations modérées de l'estimateur de la covariance introduit par Hayashi et Yoshida [78] et donné par

$$U_{n,m} = \sum_{i=1}^{n}\sum_{j=1}^{m} \Delta X_1(I^i)\Delta X_2(J^j) I_{\{I^i \cap J^j \neq \emptyset\}},$$

où $I^i = (s_{i-1}, s_i]$ $J^j = (t_{i-1}, t_i]$.
Sous certaines hypothèses, Hayashi et Yoshida [78]-[79]-[80] ont démontré la consistance et un théorème limite central pour $U_{n,m}$.

Dans l'étude des déviations modérées pour $U_{n,m}$ et des grandes déviations dans des cas particuliers, on commence par introduire un autre design réduit par rapport à $(I^i)_{i=1,\cdots,n}$. On identifie les intervalles $I^i \subset J^j$ et on les combine en un nouvel intervalle. S'il n'existe aucun intervalle I^i de ce type, on ne fait rien. La collecte de tous ces intervalles et leur réétiquetage de gauche à droite conduisent à une nouvelle partition de $[0,T]$, notée $\left(\hat{I}^i\right)_{i=1,\cdots,\hat{n}}$. En utilisant la bilinéarité de $U_{n,m}$, on a

$$U_{n,m} = \sum_{i=1}^{\hat{n}}\sum_{j=1}^{m} \Delta X_1(\hat{I}^i)\Delta X_2(J^j) K_{ij}^{\hat{I}} = \sum_{i=1}^{\hat{n}} \Delta X_1(\hat{I}^i)\Delta X_2\Big(\cup_{j\in\hat{J}(i)} J^j\Big),$$

où $\hat{J}(i) := \left\{1 \leqslant j \leqslant m : K_{ij}^{\hat{I}} \neq \emptyset\right\}$ avec $K_{ij}^{\hat{I}} = I_{\{\hat{I}^i \cap J^j \neq \emptyset\}}$.
En remarquant que chaque intervalle J^j contient au plus un $\hat{I}^i$, alors la suite de variables aléatoires $\left\{1 \leqslant i \leqslant \hat{n} : \Delta X_1(\hat{I}^i)\Delta X_2\Big(\cup_{j\in\hat{J}(i)} J^j\Big)\right\}$ est $2-$dépendante.
Pour tout borélien $I \subset [0,T]$, on définit $\nu(I) = \int_I \sigma_{1,t}\sigma_{2,t}\rho_t \mathrm{d}t$, $\nu_l(I) = \int_I \sigma_{l,t}^2 \mathrm{d}t$, $\ell = 1,2$. Soit $r_{n,m} := \max_{1\leqslant i\leqslant n}|I^i| \vee \max_{1\leqslant j\leqslant m}|J^j|$, la plus grande taille d'intervalle. On suppose que $r_{n,m} \to 0$ quand $n, m \to \infty$.

Théorème 1.4.4 *Sous les hypothèses :* ***(H1)*** *Il existe une suite de nombres réels positifs $\{v_n, n \geq 1\}$ et une certaine constante $\Sigma \in (0, +\infty)$ telles que, quand $n \to +\infty$, $v_n \to 0$, $\frac{r_{n,m}^2}{v_n} \to 0$ et*

$$v_n^{-1}\left(\sum_{i=1}^{\hat{n}}\sum_{j=1}^{m} \nu_1(\hat{I}^i)\nu_2(J^j)K_{ij}^{\hat{I}} + \sum_{i=1}^{\hat{n}} \nu^2(I^i) + \sum_{j=1}^{m} \nu^2(J^j) - \sum_{i=1}^{\hat{n}}\sum_{j=1}^{m} \nu^2(I^i \cap J^j)\right) \to \Sigma.$$

(H2) *Il existe une suite de nombres réels positifs* $\{b_n, n \geq 1\}$ *telle que, quand* $n \to +\infty$,

$$b_n \to +\infty, \quad b_n v_n^{1/2} \to 0 \quad \text{et} \quad b_n v_n^{-1/2} \nu_n \to 0,$$

où $\nu_n = \max_{\ell=1,2} \max_{1 \leq i \leq n} \max_{1 \leq j \leq m} \left(\nu_\ell(I^i) \vee \nu_\ell(J^j)\right)$.

Alors la suite $\left\{ \left(U_{n,m} - \int_0^T \sigma_{1,t}\sigma_{2,t}\rho_t dt \right) / b_n v_n^{1/2}, n \geq 1 \right\}$ *satisfait un PGD de vitesse* b_n^2 *et de fonction de taux* $L(x) = x^2/2\Sigma$. $\square$

Un deuxième resultat issu de cette colloboration concerne les déviations modérées d'un estimateur plus général de la volatilité, appelé la variation réalisée généralisée à double puissance, donné par

$$V_{\ell,1}^n(g,h) = \frac{1}{n}\sum_{i=1}^{n} g\left(\sqrt{n}\Delta X_\ell(I^i)\right) h\left(\sqrt{n}\Delta X_\ell(I^{i+1})\right), \tag{1.19}$$

où g et h sont des fonctions de $\mathbb{R}^d$.

Récemment, Barndorff-Nielsen & al. ([13]-[16]) ont établi une loi faible des grands nombres et un théorème limite central pour $V_{\ell,1}^n(g,h)$. Ces travaux ont été synthétisés dans Kinnebrock et Podolskij [87] et dans Vetter ([132]-[131]).

Rappelons que dans les travaux de Barndorff-Nielsen & al. ([13]-[16]), il a été prouvé que si g et h sont continûment différentiables et à croissance polynômiale, alors

$$V_{\ell,1}^n(g,h) \xrightarrow{\mathbb{P}} V_{\ell,1}(g,h) := \int_0^1 \Sigma_{\sigma_{\ell,u}}(g)\Sigma_{\sigma_{\ell,u}}(h)du,$$

où $\Sigma_\sigma(f) := \mathbb{E}(f(\sigma Z))$, avec $Z \sim \mathcal{N}(0,1)$ et f une fonction définie sur $\mathbb{R}$ à valeurs réelles. De plus, si g et h sont des fonctions paires, alors on a la normalité asymptotique

$$\sqrt{n}\left(V_{\ell,1}^n(g,h) - V_{\ell,1}(g,h)\right) \to \mathcal{N}(0, \Sigma_\ell(g,h))$$

où $\Sigma_\ell(g,h) = \int_0^1 \left(\Sigma_{\sigma_{\ell,s}}(g^2)\Sigma_{\sigma_{\ell,s}}(h^2) + 2\Sigma_{\sigma_{\ell,s}}(g)\Sigma_{\sigma_{\ell,s}}(h)\Sigma_{\sigma_{\ell,s}}(gh) - 3\Sigma_{\sigma_{\ell,s}}^2(g)\Sigma_{\sigma_{\ell,s}}^2(h)\right) ds$.

Nous présentons le résultat suivant qui est issu de cette collaboration et qui nous fournit les déviations modérées de $V_{\ell,1}^n(g,h)$.

Proposition 1.4.5 *Soit* $(b_n)_{n\geq 1}$ *une suite de nombres réels positifs telle que, quand* $n \to +\infty$,

$$b_n \to \infty, \quad b_n/\sqrt{n} \to 0, \quad \text{et} \quad \nu_n b_n \sqrt{n} \longrightarrow 0, \tag{1.20}$$

où $\nu_n = \max_{1\le i\le n}\int_{I^i}\Sigma_{\sigma_{\ell,s}}(g)\Sigma_{\sigma_{\ell,s}}(h)ds$*. Supposons que* g *et* h *sont des fonctions paires continûment différentiables et à croissance polynômiale telles que :*

$$\begin{aligned}\limsup_{n\to\infty}\frac{n}{b_n^2}\log n\max_{j=1}^{n}\mathbb{P}\Bigg(\Big|g\left(\sqrt{n}\Delta X_\ell(I^j)\right)h\left(\sqrt{n}\Delta X_\ell(I^{j+1})\right)\\ -\mathbb{E}\left(g(\sqrt{n}\Delta X_\ell(I^i))h(\sqrt{n}\Delta X_\ell(I^{i+1}))\right)\Big|>nb_n\Bigg)=-\infty.\end{aligned} \tag{1.21}$$

Alors, la suite $\left\{\frac{\sqrt{n}}{b_n}(V^n_{\ell,1}(g,h)-V_{\ell,1}(g,h)), n\ge 1\right\}$ *satisfait un PGD de vitesse* b_n^2 *et de fonction de taux* $I^\ell_{g,h}(x)=x^2/2\Sigma_\ell(g,h)$*.*

La preuve du théorème 1.4.4 et celle de la proposition 1.4.5 sont éssentiellement basées sur l'utilisation d'une approche sur des déviations modérées d'une suite de variables aléatoires $m-$dependantes vérifiant des conditions de type "Chen-Ledoux". Cette approche n'est autre qu'une adaptation des résultats obtenus par Djellout et Guillin [57] sur les déviations modérées d'une suite de variables aléatoires $1-$dependantes identiquement distribuées vérifiant une condition de type "Chen-Ledoux".

Problèmes ouverts et perspectives

Sans perdre de généralité, j'ose espérer que nos travaux contribueront à la résolution de certains problèmes de l'estimation de la (co-)variance que l'on rencontre très souvent dans l'analyse des données financières de hautes fréquences. Les données des transaction n'étant généralement pas observées aux mêmes instants, une des grandes avancées de cette étude serait d'établir des grandes déviations pour la (co-)volatilité réalisée dans le cadre où les diffusions observées à des temps aléatoires et irrégulièrement espacés de manière non synchronisée. On peut aussi étudier des grandes déviations de la (co-)volatilité réalisée dans différentes situations telles que des processus à sauts ou des processus asynchrones et on peut également considérer des situations où les équations différentielles stochastiques ne sont plus dirigées par des processus de Wiener mais par des processus Lévy ou des mouvements browniens fractionnaires en s'inspirant des travaux de Nourdin et al. [107], de Basawa et Brockwell [17] et de ceux de Linde et Shi [94].

Vue l'intérêt et l'utilité en finance des estimateurs du type

$$\frac{1}{n}\sum_{i=1}^{n}g\left(\sqrt{n}\Delta X(I^i)\right)h\left(\sqrt{n}\Delta X(I^{i+1})\right),$$

pour g et h des fonctions assez générales, qui ne sont qu'une généralisation des estimateurs du type

$$n^{\frac{r+q}{2}-1}\sum_{i=1}^{n-1}\left|\Delta(X)(I^i)\right|^r\left|\Delta(X)(I^{i+1})\right|^q,$$

on peut éssayer si possible d'établir des grandes déviations pour ces estimateurs outre le fait que le calcul direct de la transformée de Laplace est impossible.

1.4.3 Grandes déviations pour la statistique de Durbin-Watson

Cette partie est basée sur l'étude du processus autorégressif d'ordre p où le bruit induit vérifie aussi un processus autorégressif d'ordre q. Le but est d'établir des principes de déviations modérées pour les estimateurs du vecteur inconnu des moindres carrés et de la suite de correlation associé au bruit induit, ainsi que pour l'estimateur de la statistique de Durbin-Watson [64]-[65]-[66] inspirée des travaux de Von-Neumann [133].
Nous modélisons ainsi un processus autorégressif d'ordre q issu d'une perturbation autorégressive d'ordre p stable, définie pour tout $n \geqslant 1$, par

$$\left\{\begin{array}{rcl} X_n &=& \theta'\Phi_{n-1}^p+\varepsilon_n \\ \varepsilon_n &=& \rho'\Psi_{n-1}^q+V_n \end{array}\right. \tag{1.22}$$

où

$$\Phi_n^p=(X_n \quad X_{n-1} \quad \cdots \quad X_{n-p+1})' \quad et \quad \Psi_n^q=(\varepsilon_n \quad \varepsilon_{n-1} \quad \cdots \quad \varepsilon_{n-q+1})'$$

sont respectivement des matrices d'ordre $p\times 1$ et d'ordre $q\times 1$, pour tout $p,q\in\mathbb{N}$ et où les paramètres inconnus $\theta=(\theta_1 \quad \theta_2 \quad \cdots \quad \theta_p)'$ et $\rho = (\rho_1 \quad \rho_2 \quad \cdots \quad \rho_q)'$ sont des vecteurs non nuls. Nous supposons que $(V_n)_{n\geqslant 1}$ est une suite variables aléatoires i.i.d centrées et telle que

$$\mathbb{E}(V_1^2)=\sigma^2>0 \qquad et \qquad \mathbb{E}(V_1^4)=\tau^4<\infty.$$

Pour assurer la stabilité de notre modèle autorégressif, nous supposons que les polynômes

$$\mathcal{A}(z)=1-\theta_1 z-\cdots-\theta_p z^p \quad et \quad \mathcal{B}(z)=1-\rho_1 z-\cdots-\rho_q z^q$$

sont tels que $\mathcal{A}(z)\neq 0$ et $\mathcal{B}(z)\neq 0$ pour tout $z\in\mathbb{C}$ avec $|z|\leqslant 1$.
Soient W_n la matrice d'ordre $p\times q$ donnée, pour $n\geqslant 1$, par

$$W_n=\sum_{k=1}^{n}(\Phi_{k-1}^p X_k \quad \Phi_{k-2}^p X_k \quad \cdots \quad \Phi_{k-q}^p X_k) \tag{1.23}$$

et S_n la matrice définie positive d'ordre p donnée par

$$S_n = \sum_{k=0}^{n} \Phi_k^p \Phi_k^{p'} + S \tag{1.24}$$

où S est une matrice symétrique définie positive qui assure l'inversibilité de S_n. On définit la matrice T_n par

$$T_n = S_{n-1}^{-1} W_n. \tag{1.25}$$

Il est assez facile de voir que l'estimateur du vecteur des moindres carrés $\hat{\theta}$ du vecteur θ est donnée par

$$\hat{\theta} = (S_{n-1})^{-1} \sum_{k=1}^{n} \Phi_{k-1}^p X_k. \tag{1.26}$$

En introduisant le résidu empirique

$$\hat{\varepsilon}_k = X_k - \hat{\theta}_n X_{k-1} \qquad pour \quad 1 \leqslant k \leqslant n, \tag{1.27}$$

on en deduit l'estimateur $\hat{\rho}$ du vecteur d'autocorrelation ρ par,

$$\hat{\rho}_n = \left(\sum_{k=1}^{n} \hat{\varepsilon}_{k-1}^2 \right)^{-1} \sum_{k=1}^{n} \hat{\varepsilon}_k \hat{\varepsilon}_{k-1}. \tag{1.28}$$

Enfin, la statistique de Durbin-Watson est définie, pour tout $n \geqslant 1$, par

$$\hat{D}_n = \left(\sum_{k=0}^{n} \hat{\varepsilon}_k^2 \right)^{-1} \sum_{k=1}^{n} \left(\hat{\varepsilon}_k - \hat{\varepsilon}_{k-1} \right)^2. \tag{1.29}$$

Rappelons les resultats récemment obtenus par Proïa [112]-[113] sur le comportement asymptotique et le TCL de nos différents estimateurs

Théorème 1.4.6 *Sous les conditions de stabilité ($\mathcal{A}(z) \neq 0$ et $\mathcal{B}(z) \neq 0$), on a la convergence presque sûre des estimateurs et de la statistique de Durbin-Watson :*

$$\lim_{n\to\infty} T_n = T^*, \quad \lim_{n\to\infty} \hat{\theta}_n = \theta^*, \quad \lim_{n\to\infty} \hat{\rho}_n = \rho^*, \quad \lim_{n\to\infty} \hat{D}_n = D^*, \quad p.s.$$

où T^*, θ^*, ρ^* et D^* sont définis dans section 5.2.

De plus, si $\mathbb{E}(V_n^4) < \infty$, alors on a la normalité asymptotique

$$\sqrt{n}(vec(T_n) - vec(T^*)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Sigma_T), \quad \sqrt{n}(\hat{\theta}_n - \theta^*) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Sigma_\theta)$$

$$\sqrt{n} \begin{pmatrix} \hat{\theta}_n - \theta^* \\ \hat{\rho}_n - \rho^* \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Gamma), \quad \sqrt{n}(\hat{\rho}_n - \rho^*) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_\rho^2)$$

$$\sqrt{n}(\hat{D}_n - D^*) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_D^2)$$

où Σ_T, Σ_θ, Γ, σ_ρ^2 et σ_D^2 sont tous définis dans section 5.2.

Remarque 1.4.3 *La vectorisation d'une matrice, notée* **vec***, est une transformation linéaire qui convertit la matrice dans un vecteur colonne. Plus précisement, la vectorisation d'une matrice A de taille $m \times n$, notée* $\mathbf{vec}(A)$*, est le vecteur colonne de taille $mn \times 1$ obtenu en empilant les colonnes de la matrice A les unes sur les autres.*

Établir des principes de déviations modérées pour les estimateurs définis en (1.25), (1.26), (1.28) et (1.29) nécessite certaines hypothèses sur le bruit (V_n) et sur les conditions initiales X_0 et ε_0. Par conséquent, pour établir les déviations modérées pour nos estimateurs, nous supposons que :

– le bruit (V_n) est Gaussien d'une part (voir la section 5.3). Ainsi, les hypothèses se réduisent à l'existence d'un $t > 0$ tel que

$$\mathbb{E}[\exp(t\varepsilon_0^2)] < \infty \quad et \quad \mathbb{E}[\exp(tV_0^2)] < \infty$$

– le bruit (V_n) satisfait une condition de type Chen-Ledoux (voir la section 5.4). Plus précisement, pour $a = 2$ et $a = 4$, nous supposons que

$$\limsup_{n\to\infty} \frac{1}{b_n^2} \log n\mathbb{P}\Big(|V_1^a| > b_n\sqrt{n}\Big) = -\infty, \quad \frac{|\varepsilon_0|^a}{b_n\sqrt{n}} \xrightarrow[b_n^2]{Superexp} 0, \quad \frac{|X_0|^a}{b_n\sqrt{n}} \xrightarrow[b_n^2]{Superexp} 0.$$

Nous montrons alors les résultats suivants.

Théorème 1.4.7 *Sous les hypothèses ci-dessus et sous les conditions de stabilité ($\mathcal{A}(z) \neq 0$ et $\mathcal{B}(z) \neq 0$), les suites*

$$\left(\frac{\sqrt{n}}{b_n}(vec(T_n)-vec(T^*))\right)_{n\geqslant 1}, \quad \left(\frac{\sqrt{n}}{b_n}(\hat{\theta}_n-\theta^*)\right)_{n\geqslant 1}, \quad \left(\frac{\sqrt{n}}{b_n}(\hat{\rho}_n-\rho^*)\right)_{n\geqslant 1} \quad et \quad \left(\frac{\sqrt{n}}{b_n}(\hat{D}_n-D^*)\right)_{n\geqslant 1}$$

satisfont un PGD de vitesse b_n^2 et de fonctions de taux respectives I_T, I_θ, I_ρ et I_D, données dans la section 5.3.

Pour démontrer nos résultats, on utilise une approche martingale. Grâce au théorème de Cramer et aux inégalités exponentielles établies par Bercu et Touati [20], nous obtenons la convergence super-exponentielle du crochet des différentes martingales dans le cadre Gaussien. Les résultats de Puhalskii et le PDM des séries régressives établi par Worms [135] nous permet par la suite d'atteindre notre PDM. En utilisant les inégalités de Bercu et Touati [20] pour démontrer la convergence super-exponentielle du crochet des différentes martingales, les mêmes techniques utilisés ci-dessus permettent d'établir le PDM dans le cas où le bruit vérifie des conditions de type "Chen-Ledoux".

Notons que les résultats obtenus dans cette partie sont une généralisation des résultats établis par Bitseki et al. [25] pour un processus autorégressif stable d'ordre 1 de bruit autocorrélé d'ordre 1.

Problèmes ouverts et perspectives

Sans émettre de doutes sur la contribution des résultats obtenus dans la précision des tests de Durbin-Watson, il serait intéressant d'étudier le TCL (voir le PDM) pour un processus du type (1.22) sans faire d'hypothèse nulle d'abscence d'autocorrelation (H_0 : "$\rho = 0$"). Nos intuitions nous poussent à croirent qu'il serait possible d'établir des grandes déviations pour la statistique de Durbin-Watson ne serait ce que dans un cadre gaussien. Le champs reste donc largement ouvert et en s'inspirant des travaux existants (Worms [135], Bercu [18]), nous y consacrerons une partie de nos travaux futurs.

Chapitre 2

Large and moderate deviations of realized covolatility

Hacène DJELLOUT and Yacouba SAMOURA
Appears in Statistic Probability Letters, volume 86

abstract :

In this note, we consider the large and moderate deviation principle of the estimators of the integrated covariance of two-dimensional diffusion processes when they are observed only at discrete times in a synchronous manner. The proof is extremely simple. It is essentially an application of the contraction principle for the results given in the case of the volatility by [59].
keywords : Moderate deviation principle, Large deviation principle, Diffusion, Discrete-time observation, Quadratic variation, Realised volatility.

Résumé :

Dans cette partie, on s'intéresse aux grandes déviations et déviations modérées de l'estimateur de la covariance intégré de deux processus de diffusion observées à des temps uniformes de manière synchronisée. Les preuves sont extrêmement simples. Ce sont essentiellemnt des applications du principe de contraction aux résultats donnés pour l'estimateur de la variance réalisée dans [59].
Mots clés : Grandes déviations et déviations modérées, covariance réalisée, diffusions observées à des temps uniformes et synchronisées.

2.1 Motivation and context

Given a filtred probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$, let $(X_{1,t}, X_{2,t})$ be a two dimensional diffusion process given by

$$\begin{cases} dX_{1,t} = u_{1,t}(X_{1,t})dt + \sigma_{1,t}dB_{1,t} \\ dX_{2,t} = u_{2,t}(X_{2,t})dt + \sigma_{2,t}dB_{2,t} \end{cases} \tag{2.1}$$

where $((B_{1,t}, B_{2,t}), t \geq 0)$ is a two-dimensional Gaussian process with independent increments, zero mean and covariance matrix

$$\begin{pmatrix} t & \int_0^t \rho_s ds \\ \int_0^t \rho_s ds & t \end{pmatrix} \qquad \forall t \geq 0.$$

In (2.1), (u_1, u_2) is a progressively measurable process (possibly unknown). In what follows, we restrict our attention to the case when σ_1, σ_2 and ρ are deterministic functions ; the functions σ_i, $i = 1, 2$ take positive values while ρ takes values in the interval $[-1, 1]$. Note that the marginal processes B_1 and B_2 are Brownian motions (BM). Moreover, we can define a process B_t^* such that $B_{1,t}$ and B_t^* with $t \geq 0$ are independent Brownian motions and $dB_{2,t} = \rho_t dB_{1,t} + \sqrt{1-\rho_t^2}dB_t^*$ for every $t \geq 0$.

In this note, the parameter of interest is the (deterministic) covariance of X_1 and X_2

$$\langle X_1, X_2 \rangle_t = \int_0^t \sigma_{1,t}\sigma_{2,t}\rho_t dt. \tag{2.2}$$

In finance, $\langle X_1, X_2 \rangle_.$ is the integrated covariance (over $[0, 1]$) of the logarithmic prices X_1 and X_2 of two securities. It is an essential quantity to be measured for risk management purposes. The covariance for multiple price processes is of great interest in many financial applications. The naive estimator is the realized covariance, which is the analogue of realized variance for a single process.

Typically $X_{1,t}$ and $X_{2,t}$ are not observed in continuous time but we have only discrete time observations. Given discrete equally spaced observation $(X_{1,t_k^n}, X_{2,t_k^n}, k = 1, \cdots, n)$ in the interval $[0, 1]$ (with $t_k = k/n$), a commonly used approach to estimate is to take the sum of cross products

$$\mathbf{C}_t^n := \sum_{k=1}^{[nt]} \left(X_{1,t_k^n} - X_{1,t_{k-1}^n}\right)\left(X_{2,t_k^n} - X_{2,t_{k-1}^n}\right), \tag{2.3}$$

where $[x]$ denotes the integer part of $x \in \mathbb{R}$.

When the drift is known, we can also consider the following estimator

$$\overline{\mathbf{C}}_t^n := \sum_{k=1}^{[nt]} \left(X_{1,t_k^n} - X_{1,t_{k-1}^n} - \int_{t_{k-1}^n}^{t_k^n} u_{1,t}(X_{1,t})dt\right)\left(X_{2,t_k^n} - X_{2,t_{k-1}^n} - \int_{t_{k-1}^n}^{t_k^n} u_{2,t}(X_{2,t})dt\right)$$

In the unidimensional case and in the case that X have non-jump, this question has been well investigated see [59] for relevant references. In [59] and recently in [85], the authors obtained the large and moderate deviations for the realized volatility. The results of [59] are extended to jump-diffusion processes. [99] established the large deviation result for the threshold estimator for the constant volatility. [84] derived a moderate deviation result for the threshold estimator for the quadratic variational process.

In the bivariate case, [80] considered the problem of estimating the covariation of two diffusion processes under a non-synchronous sampling scheme. They proposed an alternative estimator and they investigated the asymptotic distributions. In [47], the authors complement the results in [80] by establishing a second-order asymptotic expansion for the distribution of the estimator in a fairly general setup, including random sampling schemes and (possibly random) drift terms. Several further works have been realized when data on two securities are observed non-synchronously, see also [1]. Here we do not consider the asynchronous case. In the bivariate case we also mention the work of [100] which deal with the problem of distinguishing the Brownian covariation from the co-jumps using a discrete set of observations.

The purpose of this note is to furnish some further estimations about the estimator (2.3), refining the already known central limit theorem. More precisely, we are interested in the estimations of

$$\mathbb{P}\left(\frac{\sqrt{n}}{b_n}\left(\mathbf{C}_t^n - \int_0^t \sigma_{1,t}\sigma_{2,t}\rho_t dt\right) \in A\right),$$

where A is a given domain of deviation, $(b_n)_{n>0}$ is some sequence denoting the scale of the deviation. When $b_n = 1$, this is exactly the estimation of the central limit theorem. When $b_n = \sqrt{n}$, it becomes the *large deviations*. And when $1 \ll b_n \ll \sqrt{n}$, this is the so called *moderate deviations*. The main problem studied in this paper is the large and moderate deviations estimations of the estimator. In this bivariate case things are not complicated.

We refer to [53], for an exposition of the general theory of large deviation and limit ourself to the statement of the some basic definitions. Let $\{\mu_T, T > 0\}$ be a family of probability on a topological space $(S, \mathcal{S})$ where $\mathcal{S}$ is a σ-algebra on S and $v(T)$ a non-negative function on $[1, \infty)$, such that $\lim_{T\to\infty} v(T) = +\infty$. A function $I : S \to [0, \infty]$ is said to be a rate function if it is lower semicontinuous and it is said to be a good rate function if its level set $\{x \in S : I(x) \le a\}$ is compact for all $a \ge 0$. $\{\mu_T\}$ is said to satisfy a large deviation principle (LDP) with speed $v(T)$ and rate function $I(x)$ if for any set $A \in \mathcal{S}$

$$-\inf_{x\in A^\circ} I(x) \le \lim_{T\to\infty} \begin{pmatrix} \inf \\ \sup \end{pmatrix} \frac{1}{v(T)} \log \mu_T(A) \le -\inf_{x\in \bar{A}} I(x).$$

where $A^0, \bar{A}$ are the interior and the closure of A respectively.

This paper is organized as follows. In the next section we present the main results of this paper. They are established in the last section.

2.2 Main results

Our first result is about the LDP of $\mathbb{P}(\mathbf{C}_1^n \in \cdot)$, with time $t = 1$ fixed.

Proposition 2.2.1 *Let $(X_{1,t}, X_{2,t})$ be given by (2.1).*

1. *For every $\lambda \in \mathbb{R}$*

$$\begin{aligned}\Lambda_n(\lambda) &:= \frac{1}{n}\log \mathbb{E}(\exp(\lambda n \overline{\mathbf{C}}_1^n)) \\ &\leq \Lambda(\lambda) := \int_0^1 P_{\rho_t}(\lambda\sigma_{1,t}\sigma_{2,t})\mathrm{d}t\end{aligned}$$

where

$$P_{\rho_t}(\lambda\sigma_{1,t}\sigma_{2,t}) := \begin{cases} -\frac{1}{2}\log(1-\lambda\sigma_{1,t}\sigma_{2,t}(1+\rho_t)) - \frac{1}{2}\log(1+\lambda\sigma_{1,t}\sigma_{2,t}(1-\rho_t)) \\ \qquad \text{if } -\frac{1}{||\sigma_1\sigma_2(1-\rho)||} \leq \lambda \leq \frac{1}{||\sigma_1\sigma_2(1+\rho)||} \\ +\infty, \qquad \text{otherwise.}\end{cases} \tag{2.4}$$

and

$$\lim_{n\to\infty}\Lambda_n(\lambda) = \Lambda(\lambda).$$

2. *Assume that $\sigma_{1,\cdot}\sigma_{2,\cdot}(1 \pm \rho_\cdot) \in L^\infty([0,1], dt)$ and $u_{l,\cdot}(\cdot) \in L^\infty(dt \otimes \mathbb{P})$, for $l = 1, 2$. Then $\mathbb{P}(\mathbf{C}_1^n \in \cdot)$ satisfies the LDP on $\mathbb{R}$ with speed n and with the good rate function given by the Legendre transformation of Λ, that is*

$$\Lambda^*(x) = \sup_{\lambda\in\mathbb{R}}\{\lambda x - \Lambda(\lambda)\}. \tag{2.5}$$

We now extend Proposition 2.2.1 to the process-level large deviations of $\mathbb{P}(\mathbf{C}_\cdot^n \in \cdot)$, which is interesting from the viewpoint of the non-parametric statistics.

Let $\mathbb{D}_b([0,1])$ be the real right-continuous-left-limit and bounded variation functions γ. The space $\mathbb{D}_b([0,1])$ of γ, identified in the usual way as the space of bounded measures $d\gamma$ on $[0,1]$, with $d\gamma[0,t] = \gamma(t)$ and $d\gamma(0) = \gamma(0)$, will be equipped with the weak convergence topology and the σ-field $\mathcal{B}^s$ generated by the coordinate $\{\gamma(t), 0 \leq t \leq 1\}$. We denote by $\dot{\gamma}(t)dt$ and $d\gamma^\perp$ respectively the absolute continuous part and the singular part of the measure $d\gamma$ associated with $\gamma \in \mathbb{D}_b[0,1]$ w.r.t. the Lebesgue measure dt. The signed measure γ has a unique decomposition into a difference $\gamma = \gamma_+ - \gamma_-$ of two positive measures γ_+ and γ_-. In the paper, we denote by P^* the function

$$P^*(x) = \begin{cases} \frac{1}{2}(x - 1 - \log x) & \text{if} \quad x > 0 \\ \\ +\infty & \text{if} \quad x \leq 0, \end{cases} \tag{2.6}$$

which is the Legendre transformation of P given by

$$P(\lambda) = \begin{cases} -\frac{1}{2}\log(1 - 2\lambda) & \text{if } \lambda < \frac{1}{2} \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.7}$$

Theorem 2.2.2 *Let $(X_{1,t}, X_{2,t})$ be given by (2.1). Assume that $\sigma_{1,\cdot}\sigma_{2,\cdot}(1 \pm \rho_\cdot) \in L^\infty([0,1], dt)$ and $u_{l,\cdot}(\cdot) = u_l(\cdot,\cdot) \in L^\infty(dt \otimes \mathbb{P})$, for $l = 1, 2$. Then*

1. *$\mathbb{P}(\mathbf{C}^n_\cdot \in \cdot)$ satisfies the LDP on $\mathbb{D}_b([0,1])$ w.r.t. the weak convergence topology, with speed n and with some inf-compact convex rate function $J(\gamma)$.*
2. *If moreover $t \to \sigma_{1,t}\sigma_{2,t}(1 \pm \rho_t)$ is continuous and strictly positive on $[0,1]$, then*

$$J(\gamma) = J_+^{abs}(\gamma_+ + \beta) + J_-^{abs}(\gamma_- + \beta) + J_+^{\perp}(\gamma_+) + J_-^{\perp}(\gamma_-), \tag{2.8}$$

where β is absolutely continuous with respect to the Lebesgue measure and given by

$$\dot{\beta}(t) = \frac{\sigma_{1,t}\sigma_{2,t}(1-\rho_t^2) - (\dot{\gamma_+}(t) + \dot{\gamma_-}(t))}{2}$$
$$+ \frac{\sqrt{[\sigma_{1,t}\sigma_{2,t}(1-\rho_t^2) - (\dot{\gamma_+}(t) + \dot{\gamma_-}(t))]^2 + (\dot{\gamma_+}(t) + \dot{\gamma_-}(t))\sigma_{1,t}\sigma_{2,t}(1-\rho_t^2)}}{2},$$

and

$$J_\pm^{\perp}(\gamma) = \int_0^1 \frac{1}{\sigma_{1,t}\sigma_{2,t}(1 \pm \rho_t)} d\gamma^{\perp},$$

and

$$J_\pm^{abs}(\gamma) = \int_0^1 P^*\left(\frac{2\dot{\gamma}(t)}{\sigma_{1,t}\sigma_{2,t}(1 \pm \rho_t)}\right) dt,$$

where P^* is given in (2.6).

We discuss now the moderate deviations principle. To this purpose, let $(b_n)_{n \geq 1}$ be a sequence of positive numbers such that

$$b_n \to \infty \qquad \text{and} \qquad \frac{b_n}{\sqrt{n}} \to 0 \qquad \text{as} \quad n \to \infty.$$

Let $\mathbb{D}_0[0,1]$ be the Banach space of real right-continuous-left-limit functions γ on $[0,1]$ with $\gamma(0) = 0$, equipped with the uniform sup norm and the σ-field $\mathcal{B}^s$ generated by the coordinate $\{\gamma(t), 0 \leq t \leq 1\}$.

Theorem 2.2.3 *Given $(X_{1,t}, X_{2,t})$ by (2.1) with $u_{l,\cdot}(\cdot) = u_l(\cdot,\cdot) \in L^\infty(dt \otimes \mathbb{P})$, for $l = 1,2$. Assume that $\sigma_{1,\cdot}\sigma_{2,\cdot}(1 \pm \rho_\cdot) \in L^2([0,1], dt)$ and*

$$\sqrt{n}b_n \max_{1\le k\le n} \int_{(k-1)/n}^{k/n} \sigma_{1,t}\sigma_{2,t}(1 \pm \rho_t)dt \longrightarrow 0. \tag{2.9}$$

Then $\mathbb{P}\left(\frac{\sqrt{n}}{b_n}\left(\mathbf{C}^n_\cdot - \langle X_1, X_2\rangle_\cdot\right) \in \cdot\right)$ satisfies the LDP on $\mathbb{D}_0([0,1])$ with speed b_n^2 and with the good rate function J_m given by

$$J_m(\gamma) = \begin{cases} \displaystyle\int_0^1 \frac{\dot{\gamma}(t)^2}{2\sigma_{1,t}^2\sigma_{2,t}^2(1+\rho_t^2)} 1_{[t:\sigma_{1,t}\sigma_{2,t}>0]}dt & \text{if } d\gamma \ll \sigma_{1,t}\sigma_{2,t}\sqrt{1+\rho_t^2}dt \\ +\infty & \text{otherwise} \end{cases} \tag{2.10}$$

Remark 2.2.1 *In particular, $\mathbb{P}\left(\frac{\sqrt{n}}{b_n}\left(\mathbf{C}^n_1 - \langle X_1, X_2\rangle_1\right) \in \cdot\right)$ satisfies the LDP on $\mathbb{R}$ with speed b_n^2 and with the rate function given by*

$$I_m(x) = \frac{x^2}{2\int_0^1 \sigma_{1,s}^2\sigma_{2,s}^2(1+\rho_s^2)ds}, \quad \forall x \in \mathbb{R}.$$

Remark 2.2.2 *If for some $p > 2$,*

$$\sigma_{1,\cdot}\sigma_{2,\cdot}(1 \pm \rho_\cdot) \in L^p([0,1], dt) \quad \text{and} \quad b_n = O\left(n^{\frac{1}{2}-\frac{1}{p}}\right),$$

we obtain (2.9).

Remark 2.2.3 *Theorem 2.2.2 and 2.2.3 continue to hold under the linear growth condition of the drift u_l $(l = 1,2)$ rather than the boundedness. More precisely assume that*

$$|u_{l,s}(x) - u_{l,t}(y)| \le \alpha_l \left[1 + |x-y| + \eta_l(|s-t|)(|x| + |y|)\right], \ \forall s,t \in [0,1], x,y \in \mathbb{R},$$

where $\eta_l : [0,+\infty) \to [0,+\infty)$ is a continuous nondecreasing function with $\eta_l(0) = 0$ and $\alpha_l > 0$ is a constant. Then the LDP of Theorem 2.2.2 and 2.2.3 continue to hold for $\mathbb{P}(\tilde{\mathbf{C}}^n_\cdot \in \cdot)$, where $\tilde{\mathbf{C}}^n_\cdot$ is given by

$$\tilde{\mathbf{C}}^n_t := \sum_{k=1}^{[nt]} \left(X_{1,t_k^n} - X_{1,t_{k-1}^n} - u_{1,t_{k-1}^n}(X_{1,t_{k-1}^n})(t_k^n - t_{k-1}^n)\right) \left(X_{2,t_k^n} - X_{2,t_{k-1}^n} - u_{2,t_{k-1}^n}(X_{2,t_{k-1}^n})(t_k^n - t_{k-1}^n)\right).$$

We introduce the following function :

$$\Lambda_\pm^*(x) = \sup_{\lambda\in\mathbb{R}}\{\lambda x - \Lambda_\pm(\lambda)\}, \tag{2.11}$$

which is the Legendre transformation of $\Lambda_\pm$ given by

$$\Lambda_\pm(\lambda) := \int_0^1 P\left(\pm\frac{\lambda\sigma_{1,t}\sigma_{2,t}(1\pm\rho_t)}{2}\right)dt. \tag{2.12}$$

And we denote by

$$\alpha_{\pm,t} = \frac{1}{2}\int_0^t \sigma_{1,s}\sigma_{2,s}(1\pm\rho_s)ds. \tag{2.13}$$

An easy application of deviation inequalities given in Proposition 1.5 in [59] gives

Proposition 2.2.4 *We have for every $n \geq 1$ and $r > 0$,*

$$\begin{aligned}
\mathbb{P}\left(\sup_{t\in[0,1]}[\overline{\mathbf{C}}_t^n - \mathbb{E}\overline{\mathbf{C}}_t^n] \geq r\right) &\leq \exp\left(-n\Lambda_+^*(\alpha_+ + \frac{r}{2})\right) + \exp\left(-n\Lambda_-^*(\alpha_- - \frac{r}{2})\right)\\
&\leq \exp\left(-\frac{n}{2}\left[\frac{r}{\|\sigma_1\sigma_2(1+\rho)\|_\infty} - \log\left(1+\frac{r}{\|\sigma_1\sigma_2(1+\rho)\|_\infty}\right)\right]\right)\\
&\quad + \exp\left(-n\frac{r^2}{4\int_0^1[\sigma_1\sigma_2(1-\rho)]^2dt}\right),
\end{aligned}$$

$$\begin{aligned}
\mathbb{P}\left(\inf_{t\in[0,1]}[\overline{\mathbf{C}}_t^n - \mathbb{E}\overline{\mathbf{C}}_t^n] \leq -r\right) &\leq \exp\left(-n\Lambda_+^*(\alpha_+ - \frac{r}{2})\right) + \exp\left(-n\Lambda_-^*(\alpha_- + \frac{r}{2})\right)\\
&\leq \exp\left(-n\frac{r^2}{4\int_0^1[\sigma_1\sigma_2(1+\rho)]^2dt}\right)\\
&\quad + \exp\left(-\frac{n}{2}\left[\frac{r}{\|\sigma_1\sigma_2(1-\rho)\|_\infty} - \log\left(1+\frac{r}{\|\sigma_1\sigma_2(1-\rho)\|_\infty}\right)\right]\right),
\end{aligned}$$

where $\Lambda_\pm^*$ and $\alpha_\pm$ are given in (2.11) and (2.13) respectively.

2.3 Proof

In this section, we will give some hints for the proof of the main results. We have the following

$$\overline{\mathbf{C}}_t^n = \sum_{k=1}^{[nt]}\left(\int_{t_{k-1}^n}^{t_k^n}\sigma_{1,s}dB_{1,s}\right)\left(\int_{t_{k-1}^n}^{t_k^n}\sigma_{2,s}dB_{2,s}\right) = \sum_{k=1}^{[nt]}\sqrt{a_k}\sqrt{a_k'}\xi_k\xi_k',$$

where

$$\xi_k = \frac{\int_{t_{k-1}^n}^{t_k^n}\sigma_{1,s}dB_{1,s}}{\sqrt{a_k}} \quad \xi_k' = \frac{\int_{t_{k-1}^n}^{t_k^n}\sigma_{2,s}dB_{2,s}}{\sqrt{a_k'}} \quad \text{with} \quad a_k = \int_{t_{k-1}^n}^{t_k^n}\sigma_{1,t}^2dt \quad a_k' = \int_{t_{k-1}^n}^{t_k^n}\sigma_{2,t}^2dt.$$

Obviously $((\xi_k, \xi'_k))_{k=1,\cdots,n}$ are independent centered Gaussian random vector with covariance

$$\frac{1}{\sqrt{a_k}\sqrt{a'_k}}\begin{pmatrix} \sqrt{a_k}\sqrt{a'_k} & \int_{t_{k-1}^n}^{t_k^n}\sigma_{1,s}\sigma_{2,s}\rho_s ds \\ \int_{t_{k-1}^n}^{t_k^n}\sigma_{1,s}\sigma_{2,s}\rho_s ds & \sqrt{a_k}\sqrt{a'_k} \end{pmatrix}.$$

Let us introduce the following notation :

$$\mathbf{Q}^n_{\pm,t} = \frac{1}{4}\sum_{k=1}^{[nt]}\sqrt{a_k}\sqrt{a'_k}(\xi_k \pm \xi'_k)^2. \tag{2.14}$$

The proof relies on the following decomposition

$$\overline{\mathbf{C}}^n_t = \mathbf{Q}^n_{+,t} - \mathbf{Q}^n_{-,t}.$$

Proof of Proposition 2.2.1 By the independence of $\mathbf{Q}^n_{+,1}$ and $\mathbf{Q}^n_{-,1}$, we obtain that

$$\begin{aligned}\Lambda_n(\lambda) &= \frac{1}{n}\log\mathbb{E}(\exp(\lambda n\overline{\mathbf{C}}^n_1)) = \frac{1}{n}\log\mathbb{E}(\exp(\lambda n(\mathbf{Q}^n_{+,1} - \mathbf{Q}^n_{-,1})))\\ &= \frac{1}{n}\log\mathbb{E}(\exp(\lambda n\mathbf{Q}^n_{+,1})) + \frac{1}{n}\log\mathbb{E}(\exp(-\lambda n\mathbf{Q}^n_{-,1})) := \Lambda_{n,+}(\lambda) + \Lambda_{n,-}(\lambda).\end{aligned}$$

Let us deal with $\Lambda_{n,+}$. We have that

$$\begin{aligned}\Lambda_{n,+}(\lambda) &= \frac{1}{n}\log\mathbb{E}(\exp(\lambda n\mathbf{Q}^n_{+,1}))\\ &= \frac{1}{n}\sum_{k=1}^n P\left(n\frac{\lambda}{2}\left(\sqrt{\int_{t_{k-1}^n}^{t_k^n}\sigma^2_{1,s}ds}\sqrt{\int_{t_{k-1}^n}^{t_k^n}\sigma^2_{2,s}ds} + \int_{t_{k-1}^n}^{t_k^n}\sigma_{1,s}\sigma_{2,s}\rho_s ds\right)\right)\\ &= \int_0^1 P\left(\frac{\lambda}{2}f_n(t)\right)dt,\end{aligned}$$

where P is given in (2.7) and

$$f_n(t) := \sum_{k=1}^n 1_{(t_{k-1}^n, t_k^n]}(t)\frac{\sqrt{\int_{t_{k-1}^n}^{t_k^n}\sigma^2_{1,s}ds}\sqrt{\int_{t_{k-1}^n}^{t_k^n}\sigma^2_{2,s}ds} + \int_{t_{k-1}^n}^{t_k^n}\sigma_{1,s}\sigma_{2,s}\rho_s ds}{t_k^n - t_{k-1}^n}.$$

Let us remark that we have

$$\begin{aligned}f_n(t) &:= \sqrt{\sum_{k=1}^n 1_{(t_{k-1}^n, t_k^n]}(t)\frac{\int_{t_{k-1}^n}^{t_k^n}\sigma^2_{1,s}ds\int_{t_{k-1}^n}^{t_k^n}\sigma^2_{2,s}ds}{(t_k^n - t_{k-1}^n)^2}} + \sum_{k=1}^n 1_{(t_{k-1}^n, t_k^n]}(t)\frac{\int_{t_{k-1}^n}^{t_k^n}\sigma_{1,s}\sigma_{2,s}\rho_s ds}{t_k^n - t_{k-1}^n}\\ &= \mathbb{E}^{dt}(\sigma_{1,\cdot}\sigma_{2,\cdot}(1+\rho_\cdot) \setminus \mathcal{B}_{\tau_n})(t).\end{aligned}$$

Clearly, $f_n(t)$ is a dt martingale w.r.t. the partially directed filtration $(\mathcal{B}_{\tau_n} := \sigma((t_{k-1}^n, t_k^n], k = 1, \cdots, n))_n$.

By the convexity of P and Jensen inequality, we obtain that

$$\int_0^1 P\left(\frac{\lambda}{2} f_n(t)\right) dt = \int_0^1 P\left(\frac{\lambda}{2}\mathbb{E}^{dt}(\sigma_{1,\cdot}\sigma_{2,\cdot}(1+\rho_\cdot) \setminus \mathcal{B}\tau_n)\right)(t)dt$$
$$\leq \int_0^1 \mathbb{E}^{dt}\left(P\left(\frac{\lambda\sigma_{1,\cdot}\sigma_{2,\cdot}(1+\rho_\cdot)}{2}\right) \setminus \mathcal{B}_{\tau_n}\right)(t)dt = \Lambda_+(\lambda)$$

On the other hand, by the classical Lebesgue derivation theorem, we have that

$$f_n(t) \longrightarrow f(t) := \sigma_{1,t}\sigma_{2,t} + \sigma_{1,t}\sigma_{2,t}\rho_t, \quad dt - a.e. \quad \text{on} \quad [0,1].$$

The continuity of $P : \mathbb{R} \to (-\infty, +\infty]$ gives

$$P\left(\frac{\lambda}{2} f_n(t)\right) \longrightarrow P\left(\frac{\lambda}{2} f(t)\right), \quad dt - a.e. \quad \text{on} \quad [0,1].$$

As $P\left(\frac{\lambda}{2} f_n(t)\right) \geq -\frac{|\lambda|}{2}\sigma_{1,t}\sigma_{2,t} \in L^1([0,1], dt)$, we can apply Fatou's lemma to conclude that

$$\liminf_{n\to\infty} \Lambda_{n,+}(\lambda) = \liminf_{n\to\infty} \int_0^1 P\left(\frac{\lambda}{2} f_n(t)\right) dt \geq \int_0^1 \liminf_{n\to\infty} P\left(\frac{\lambda}{2} f_n(t)\right) dt = \Lambda_+(\lambda).$$

Doing the same calculations with $\Lambda_{n,-}$, we obtain that

$$\Lambda_{n,-}(\lambda) \leq \int_0^1 P\left(-\frac{\lambda\sigma_{1,t}\sigma_{2,t}(1-\rho_t)}{2}\right) dt = \Lambda_-(\lambda),$$

and

$$\liminf_{n\to\infty} \Lambda_{n,-}(\lambda) \geq \Lambda_-(\lambda).$$

From below, we conclude that

$$\Lambda_n(\lambda) \leq \Lambda_+(\lambda) + \Lambda_-(\lambda) := \Lambda(\lambda),$$

and

$$\liminf_{n\to\infty} \Lambda_n(\lambda) \geq \Lambda(\lambda),$$

which implies that

$$\lim_{n\to\infty} \Lambda_n(\lambda) = \lim_{n\to\infty} (\Lambda_{n,+}(\lambda) + \Lambda_{n,-}(\lambda)) = \Lambda(\lambda).$$

Which ends the proof of first part of Proposition 2.2.1.

For the second part of Proposition 2.2.1, at first we will reduce the study to the case $u_l = 0$ for $l = 1, 2$. Let $\beta = \max(|| u_1 ||_\infty, || u_2 ||_\infty)$. Since

$$\left|\mathbf{C}_1^n - \overline{\mathbf{C}_1^n}\right| \leq \frac{\beta^2}{n} + \frac{\beta}{n}\sum_{k=1}^n \left(\left|\int_{t_{k-1}^n}^{t_k^n} \sigma_{1,s} dB_{1,s}\right| + \left|\int_{t_{k-1}^n}^{t_k^n} \sigma_{2,s} dB_{2,s}\right|\right).$$

For $l = 1, 2$, we have for all $\lambda > 0$ and all $\delta > 0$

$$\frac{1}{n} \log \mathbb{P}\left(\frac{1}{n} \sum_{k=1}^{n} \left| \int_{t_{k-1}^n}^{t_k^n} \sigma_{l,s} dB_{l,s} \right| \geq \delta \right) \leq -\delta\lambda + \frac{\lambda^2}{2n} \int_0^1 \sigma_{l,s}^2 ds.$$

Letting n go to infinity and then λ to infinity we get that for all $\delta > 0$

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\frac{1}{n} \sum_{k=1}^{n} \left| \int_{t_{k-1}^n}^{t_k^n} \sigma_{l,s} dB_{l,s} \right| \geq \delta \right) = -\infty.$$

By the approximation technique (Theorem 4.2.13 in [53]), we deduce that $\mathbb{P}\left(\mathbf{C}_1^n \in \cdot\right)$ satisfies the same LDP as $\mathbb{P}\left(\overline{\mathbf{C}_1^n} \in \cdot\right)$. Hence we can assume that $u_l = 0$ for $l = 1, 2$.

Now, by inspection of the proof of Theorem 1.1 in [59], we deduce that the sequence $\mathbb{P}\left(\mathbf{Q}_{\pm,1}^n \in \cdot\right)$ satisfies the LDP on $\mathbb{R}$ with speed n and rate function given by

$$\Lambda_{\pm}^*(x) = \sup_{\lambda\in\mathbb{R}} \{\lambda x - \Lambda_{\pm}(\lambda)\}.$$

By the independence of the sequences $\mathbf{Q}_{+,1}^n$ and $\mathbf{Q}_{-,1}^n$, and the contraction principle, see Exercise 4.2.7 in [53], we deduce that $\mathbb{P}\left(\overline{\mathbf{C}}_1^n \in \cdot\right)$ satisfies the LDP with rate function

$$\Lambda^*(x) = \inf_{x=x_1-x_2} \left\{\Lambda_+^*(x_1) + \Lambda_-^*(x_2)\right\}.$$

As we have also determined explicitly the logarithm of the moment generating function Λ, the rate function is also given by (2.5).

Proof of Theorem 2.2.2

The proof of the first part is very similar to Proposition 2.2.1. It is a consequence of Theorem 1.2 in [59] and the contraction principle.
At first for any partition $\mathcal{S} = \{0 = s_0 < \cdots < s_m = 1\}$, we shall prove that $\mathbb{P}\Big(C_{\mathcal{S}}^n(X) \in \cdot\Big)$ satisfies the LDP on $\mathbb{R}^m$ with speed n and with rate function given by

$$J^{\mathcal{S}}(x_0, \cdots, x_m) = \sup_{(\lambda_1,\cdots,\lambda_m)\in\mathbb{R}^m} \left\{ \sum_{k=1}^{m} (x_k - x_{k-1})\lambda_k - \sum_{k=1}^{m} \int_{s_{k-1}}^{s_k} P_{\rho_t}(\lambda_k \sigma_{1,t}\sigma_{2,t}) \mathrm{d}t \right\} \tag{2.15}$$

if $x_0 = 0$ and $J^{\mathcal{S}}(x_0, \cdots, x_m) = +\infty$ esle.

Since we can assume that $u_l = 0$ for $l = 1, 2$, the key of the proof is that

$$\Delta_k^{\mathcal{S}} C^n(X) := C_{s_k}^n - C_{s_{k-1}}^n, \quad k = 1, \cdots, m$$

are independent. So by the Proposition 2.2.1, for each $k = 1, \cdots, m$ fixed, $\mathbb{P}(\Delta_k^{\mathcal{S}} C^n(X) \in \cdot)$ satisfies the LDP with speed n and the rate function

$$J_k^{\mathcal{S}}(y_k) = \sup_{\lambda\in\mathbb{R}} \left\{ \lambda y_k - \int_{s_{k-1}}^{s_k} P_{\rho_t}(\lambda \sigma_{1,t}\sigma_{2,t}) \mathrm{d}t \right\} \tag{2.16}$$

Hence by [97, Corollary 2.9], $\mathbb{P}\Big(\Delta^{\mathcal{S}}C^n(X) := (\Delta_k^{\mathcal{S}}C^n(X))_{0\leq k\leq m} \in \cdot\Big)$ satisfies the LDP with speed n and with rate function given by

$$J^{\mathcal{S}}(y_1,\cdots,y_m)=\sum_{k=1}^m J_k^{\mathcal{S}}(y_k)=\sup_{(\lambda_1,\cdots,\lambda_m)\in\mathbb{R}^m}\Big\{\sum_{k=1}^m\lambda_k y_k-\sum_{k=1}^m\int_{s_{k-1}}^{s_k}P_{\rho_t}(\lambda\sigma_{1,t}\sigma_{2,t})\mathrm{d}t\Big\}$$

Then by the contraction principle, we have that $\mathbb{P}(C^n_{\mathcal{S}}(X)\in\cdot)$ satisfies the LDP on $\mathbb{R}^m$ with speed n and with rate function $J^{\mathcal{S}}(x_0,\cdots,x_m)$ given by (2.15).

Now we shall prove that for any $\delta>0$ and $s\in[0,1]$

$$\lim_{\varepsilon\downarrow 0}\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\sup_{s\leqslant t\leqslant s+\varepsilon}|\Delta_s^t C^n_{\cdot}(X)|>\eta\right)=-\infty \tag{2.17}$$

Since $C^n_{\cdot}=\mathbf{Q}^n_{+,\cdot}-\mathbf{Q}^n_{-,\cdot}$ where $\mathbf{Q}^n_{\pm,t}$ is given by (2.14), then to prove (2.17) we need to prove that for any $\delta>0$ and $s\in[0,1]$

$$\lim_{\varepsilon\downarrow 0}\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\sup_{s\leqslant t\leqslant s+\varepsilon}|\Delta_s^t\mathbf{Q}^n_{\pm,\cdot}|>\eta\right)=-\infty \tag{2.18}$$

Furthermore, (2.18) can be done in the same way than in Djellout & al. [59]. From (2.15) and (2.17), together with Lemma A.1 in [59] (or [53, §4.6]), the sequence $\mathbb{P}(C^n_{\cdot}(X)\in\cdot)$ satisfies the LDP with speed n and good rate function given by

$$J^{\infty}(\gamma)=\sup_{\mathcal{S}}J^{\mathcal{S}}(\gamma(\mathcal{S}))$$

where the supremum are taken over all finite partitions $\mathcal{S}=\{0=s_0<\cdots<s_m=1\}$ of $[0,1]$ and for each such $\mathcal{S}$,

$$J^{\mathcal{S}}(\gamma(\mathcal{S}))=\sum_{k=1}^m\sup_{\lambda_k\in\mathbb{R}}\Big\{(\gamma(s_k)-\gamma(s_{k-1}))\lambda_k-\int_{s_{k-1}}^{s_k}P_{\rho_t}(\lambda_k\sigma_{1,t}\sigma_{2,t})\mathrm{d}t\Big\}$$

by (2.15), remark that

$$J^{\mathcal{S}}(\gamma(\mathcal{S}))=+\infty,\quad if\quad \gamma(s_k)\leq\gamma(s_{k-1}).$$

Finally, by the straightforward calculations and the contraction principle, one can deduce that the sequence $\mathbb{P}(C^n_{\cdot}(X)\in\cdot)$ satisfies the LDP on $\mathbb{D}_b([0,1])$ with speed n and with rate function given by

$$\tilde{J}^{\infty}(\gamma)=\inf\Big\{J^{\infty}(f);\quad f\in\mathcal{I}_0[0,1]\quad and\quad \tilde{f}=\gamma\Big\}$$

which is inf-compact and convex,
with $\mathcal{I}_0[0,1]$ the space of the non-decreasing functions γ on $[0,1]$ with $\gamma(0)=$

0 equipped with the pointwise convergence topology is a closed subset of $\mathbb{R}^{[0,1]}$.

For the second part of Theorem 2.2.2, the same arguments give the large deviation with the rate function

$$I(\gamma) = \inf_{\gamma=\gamma_1-\gamma_2} \{I_+(\gamma_1) + I_-(\gamma_2)\}$$

where

$$I_\pm(\gamma) = \int_0^1 P^* \left(\frac{2\dot{\gamma}(t)}{\sigma_{1,t}\sigma_{2,t}(1 \pm \rho_t)} \right) dt + \int_0^1 \frac{1}{\sigma_{1,t}\sigma_{2,t}(1 \pm \rho_t)} d\gamma^\perp.$$

An easy variational calculus gives the identification of the rate function in (2.8).

Proof of Theorem 2.2.3 As before, we treat only the case $u_l = 0$ for $l = 1, 2$. We have the following decomposition

$$\mathbf{C}^n_\cdot - \langle X_1, X_2 \rangle_\cdot = (\mathbf{Q}^n_{+,\cdot} - \alpha_{+,\cdot}) - (\mathbf{Q}^n_{-,\cdot} - \alpha_{-,\cdot}),$$

where the definition of $\alpha_\pm$ is given in (2.13).

Now using Theorem 1.3 in [59], we deduce that $\mathbb{P}\left(\frac{\sqrt{n}}{b_n}(\mathbf{Q}^n_{\pm,\cdot} - \alpha_{\pm,\cdot}) \in \cdot\right)$ satisfies the LDP on $\mathbb{D}_0([0,1])$ with speed b_n^2 and with the good rate function $J_{\pm,m}$ given by

$$J_{\pm,m}(\gamma) = \begin{cases} \displaystyle\int_0^1 \frac{\dot{\gamma}(t)^2}{\sigma_{1,t}^2\sigma_{2,t}^2(1 \pm \rho_t)^2} 1_{[t:\sigma_{1,t}\sigma_{2,t}>0]} dt & \text{if } d\gamma \ll \sigma_{1,t}\sigma_{2,t}(1 \pm \rho_t)dt \\ +\infty & \text{otherwise.} \end{cases}$$

By the same argument as before, we deduce that $\mathbb{P}\left(\frac{\sqrt{n}}{b_n}(\mathbf{C}^n_\cdot - \langle X_1, X_2 \rangle_\cdot) \in \cdot\right)$ satisfies the LDP on $\mathbb{D}_0([0,1])$ with speed b_n^2 and with the good rate function J_m given by

$$J_m(\gamma) = \inf_{\gamma=\gamma_1-\gamma_2} \{J_{+,m}(\gamma_1) + J_{-,m}(\gamma_2)\}.$$

An easy calculation gives the identification of the rate function in (2.10).

Chapitre 3

Estimation of the realized (co-)volatility vector : Large deviations approach

Hacène DJELLOUT, Arnaud GUILLIN and Yacouba SAMOURA

abstract :

Realized statistics based on high frequency returns have become very popular in financial economics. In recent years, different non-parametric estimators of the variation of a log-price process have appeared. Among them are the realized quadratic (co-)variation which is perhaps the most well known example, providing a consistent estimator of the integrated (co-)volatility when the logarithmic price process is continuous. In this paper, we propose to study the large deviation properties of realized (co-)volatility. Our main motivation is to improve upon the existing limit theorems such as the weak law of large numbers or the central limit theorem which have been proved in different contexts. Our large deviations results can be used to evaluate and approximate tail probabilities of realized (co-)volatility. As an application we provide the large deviations for the standard dependence measures between the two assets returns such as the realized regression coefficients or the realized correlation. Our study should contribute to the recent trend of research on the (co-)variance estimation problems, which are quite often discussed in high-frequency financial data analysis.
keywords : Realised Volatility and covolatility, large deviations, diffusion, discrete-time observation.

Résumé :

Les statistiques réalisées basées sur des données observées à haute fréquence sont devenues très populaires en économies financières. Plusieurs estimateurs non-parametriques de la variance de processus stochastiques ont fait leur apparition au cours de ces dernières années. Parmi eux, la (co-)volatilité intégrée qui est l'un des plus connus, est un estimateur consistant de la (co-)variance quadratique réalisée d'un processus de prix logharithmique. Dans ce chapitre, nous étudions des propriétés de grandes déviations de la (co-)volatilité réalisée. Notre motivation principale est d'améliorer des théorèmes limites existants tels que la loi faible des grands nombres ou le théorème limite central qui ont déjà été prouvés dans ce contexte. Nos résultats de grandes déviations peuvent être utilisés pour évaluer des probabilités de queues approximatives de la (co-)volatilité réalisée. Pour des applications, nous fournissont des grandes déviations pour mésurer la dépendance standard entre deux actifs financiers telle que le coefficient de regréssion réalisé ou a corrélation réalisée. Notre étude devrait faciliter la résolution de certains problèmes récents de l'estimation de la (co-)variance qui apparaissent souvent dans l'analyse des données financières observées à haute fréquence.

Mots clés : Grandes déviations, volatilité et covolatilité réalisées, diffusions observées à des temps uniformes et synchronisées.

3.1 Introduction

In the last decade there has been a considerable development of the asymptotic theory for processes observed at a high frequency. This was mainly motivated by financial applications, where the data, such as stock prices or currencies, are observed very frequently.

Asset returns covariance and its related statistics play a prominent role in many important theoretical as well as practical problems in finance,see [2]. Analogous to the realized volatility approach, the idea of employing high frequency data in the computation of daily (or lower frequency) covariance between two assets leads to the concept of realized covariance (or covariation). The key role of quantifying integrated (co-)volatilities in portfolio optimization and risk management has stimulated an increasing interest in estimation methods for these models.

It is quite natural to use the asymptotic framework when the number of high frequency observations in a fixed time interval (say, a day) increases to infinity. Thus Jacod and Protter [82], Barndorff-Nielsen and Shephard [15] established a law of large numbers and the corresponding fluctuations for realized volatility, also extended to more general setups and statistics by Barndorff-Nielsen et al. [13] and [12]. Dovonon, Gonçalves, and Meddahi [63] considered Edgeworth expansions for the realized volatility statistic and its bootstrap analog. These results are crucial to explore asymptotic behaviors of realized (co-)volatility, in particular around the center of its distribution. There are also different estimation approaches for the integrated covolatility in multidimensional models and limit theorem, and we can refer to Barndorff-Nielsen et al. [16] and [13] where the authors present, in an unified way, a weak law of large numbers and a central limit theorem for a general estimator, called realized generalized bipower variation.

For related work concerning bivariate case under a non-synchronous sampling scheme, see Hayashi and Yoshida [80], Bibinger [22], Dalalyan and Yoshida [47], see also Aït-Sahalia et al. [1] and the references therein. Estimation of the covariance of log-price processes in the presence of market microstructure noise, we refer to Bibinger and Reiß [23], Robert and Rosenbaum [118], Zhang et al. [141] and [142]. See also Gloter [76], or Comte et al. [43] for non parametric estimation in the case of a stochastic volatility model.

The large and moderate deviations problems arise in the theory of statistical inference quite naturally. For estimation of unknown parameters and functions, it is important to minimize the risk of wrong decisions implied by deviations of the observed values of estimators from the true values of parameters or functions to be estimated. Such important errors are precisely the subject of large deviation theory. The large and moderate deviation results of estimators can provide us with the rates of convergence and an useful method for constructing asymptotic confidence intervals. Recall that the large deviations find its interest in the statistical test theory, see the seminal work of Borovkov and Mogulskii [28],[29], [30],[31].

The aim of this paper is then to focus on the large and moderate deviation estimations of the estimators of volatility and co-volatility. Despite the fact that these statistics are nearly 20 years old, there has been remarkably few result in this direction, it is a surprise to us. The answer may however be the following : the usual techniques (such as Gärtner-Ellis method) do not work and a very particular treatment has to be considered for this problem. Recently, however, some papers have considered the unidimensional case. Djellout et al. [56] and recently Kanaya and Otsu [85] obtained the large and moderate deviations for the realized volatility. In the bivariate case Djellout and Yacouba [61], obtained the large and moderate deviations for the realized covolatility. The large deviation for threshold estimator for the constant volatility was established by Mancini [99] in jumps case. And the moderate deviation for threshold estimator for the quadratic variational process was derived by Jiang [84]. Let us mention that the problem of the large deviation for threshold estimator vector, in the presence of jumps, will be considered in a forthcoming paper, consistency, efficience and robustnesse were proved in Jacod and Todorov [83], Mancini and Gobbi [100]. The case of asynchronous sampling scheme, or in the presence of micro-structure noise is also outside the scope of the present paper but are currently under investigations.

3.2 Model and Notations

We model the evolution of an observable state variable by a stochastic process $X_t = (X_{1,t}, X_{2,t}), t \in [0, 1]$. In financial applications, X_t can be thought of as the short interest rate, a foreign exchange rate, or the logarithm of an asset price or of a stock index. Suppose both $X_{1,t}$ and $X_{2,t}$ are defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ and follow an Itô process, namely,

$$\left\{ \begin{array}{lllll} dX_{1,t} & = & \sigma_{1,t} dB_{1,t} & + & b_1(t, X_t)dt \\ dX_{2,t} & = & \sigma_{2,t} dB_{2,t} & + & b_2(t, X_t)dt \end{array} \right. \tag{3.1}$$

where B_1 and B_2 are standart Brownian motions, with correlation $Corr(B_{1,t}, B_{2,t}) = \rho_t$. We can write $dB_{2,t} = \rho_t dB_{1,t} + \sqrt{1-\rho_t^2} dB_{3,t}$, where $B_1 = (B_{1,t})_{t\in[0,1]}$ and $B_3 = (B_{3,t})_{t\in[0,1]}$ are independent Brownian processes.

Under some integrability condtions on the drift coefficient b_1 and b_2, the processes $X_{1,t}$ and $X_{2,t}$ are well defined through stochastic integrals. We limit our attention to the case when σ_1, σ_2 and ρ are deterministic functions. The functions σ_ℓ, $\ell = 1, 2$ take positive values while ρ takes values in the interval $]-1, 1[$.

In this paper, our interest is to estimate the (co-)variation vector

$$[V]_t = ([X_1]_t, [X_2]_t, \langle X_1, X_2 \rangle_t)^T \tag{3.2}$$

between two returns in a fixed time period $[0; 1]$ when $X_{1,t}$ and $X_{2,t}$ are observed synchronously, $[X_\ell]_t, \ell = 1, 2$ represente the quadratic variational process of X_ℓ and $\langle X_1, X_2 \rangle_t$ the (deterministic) covariance of X_1 and X_2 :

$$[X_\ell]_t = \int_0^t \sigma_{\ell,s}^2 \mathrm{d}s, \qquad \langle X_1, X_2 \rangle_t = \int_0^t \sigma_{1,s}\sigma_{2,s}\rho_s \mathrm{d}s.$$

Inference for (3.2) is a well-understood problem if $X_{1,t}$ and $X_{2,t}$ are observed simultaneously. Note that $X_{1,t}$ and $X_{2,t}$ are not observed in continuous time but we have only discrete time observations. Given discrete equally space observation $(X_{1,t_k^n}, X_{2,t_k^n}, k = 1, \cdots, n)$ in the interval $[0, 1]$ (with $t_k^n = k/n$), a limit theorem in stochastic processes states that

$$V_t^n(X) = \left(Q_{1,t}^n(X), Q_{2,t}^n(X), C_t^n(X)\right)^T$$

commonly called realized (co-)volatility, is a consistent estimator for $[V]_t$, with, for $\ell = 1, 2$

$$Q_{\ell,t}^n(X) = \sum_{k=1}^{[nt]} (\Delta_k^n X_\ell)^2 \qquad C_t^n(X) = \sum_{k=1}^{[nt]} (\Delta_k^n X_1)(\Delta_k^n X_2)$$

where $[x]$ denote the integer part of $x \in \mathbb{R}$ and $\Delta_k^n X_\ell = X_{\ell,t_k^n} - X_{\ell,t_{k-1}^n}$.

In the aforementionned papers, and under quite weak assumptions, it is proved the consistency $V_1^n(X) \longrightarrow [V]_1$ *a.s.* and the corresponding fluctuations $\sqrt{n}(V_1^n(X) - [V]_1) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Sigma)$.

The purpose of this paper is to furnish some further trajectorial estimations about the estimator $V_\cdot^n$, deepening the law of large numbers and central limit theorem. More precisely, we are interested in the estimation of

$$\mathbb{P}\left(\frac{\sqrt{n}}{v_n}(V_\cdot^n(X) - [V]_\cdot) \in A\right),$$

where A is a given domain of deviation, and $(v_n)_{n \geqslant 0}$ is some sequence denoting the scale of the deviation.

When $v_n = 1$, this is exactly the estimation of the central limit theorem. When $v_n = \sqrt{n}$, it becomes the large deviations. And when $1 << v_n << \sqrt{n}$, it is called moderate deviations. In other words, the moderate deviations investigate the convergence speed between the large deviations and central limit theorem.

As in Djellout et al. [56], Kanaya and Otsu [85], it should be noted that the proof strategy of Gärtner-Ellis large deviation theorem cannot be adapted to our large deviations case. We are in the situation of what Schied [123] called "the weak Cramer condition". We will encounter the same technical difficulties as in the papers of Bercu et al. [10], Bryc and Dembo [36] and Gamboa et al. [**?**] where they established the large deviation principle for quadratic forms of Gaussian processes. On large deviations for weighted random variables we refer also to Maïda et al. [98] and Zani [140]. Since we cannot determine the limiting behavior of the cumulant generating function at some boundary point, we will use another approach based on the beautiful results by Najim for weighted sums of independent quadratic Gaussian functionals [105], [106] and [104], where the steepness assumption concerning the cumulant generating function is relaxed.

We are also interseted in LDP that depends on a whole random process that evolves over time. We look what happens if the path $V_t^n(X)$ is different from the deterministic path $[V]_t$, in particular the rate the probability of such a path goes to zero as the number of paths n in the sequence increases. The corresponding asymptotics probabilities, usually called sample path large deviations, were developed by Borovkov [27], Mogulskii [101] for random walks, Freidlin-Wentzell [71] and Azencott [8] for dynamical system, Donsker-Varadhan [62], Varadhan [130] for Gaussian processes. The first example is known as Schilder's theorem. And most of the sample path large deviations that have been developed so far involve stochastic processes with independent or nearly-independent increments, see, for instance, Dembo and Zajic [52] and [51] for functional empirical processes, de Acosta [48] for Lévy processes, and Chang [37] for weakly correlated processes ; results for processes with a stronger correlation structure are restricted to special classes of processes, such as Gaussian processes see Zani ([140], [108]). For some applications of the path large deviations for weighted statistics, queueing theory, the reader can see [6], [68], [115]. For applications of the large deviations in finance and insurance, we refer to [42],[109], [119], [126].

It has to be noted that the form of the path large deviations rate function in our case is also original : at the process level, and because of the weak exponential integrability of V_t^n, a correction (or extra) term appears in rate function, a phenomenon first discovered by Lynch and Sethuraman [97], see also Moguslkii [102].

Two economically interesting functions of the realized covariance vector are the realized correlation and the realized regression coefficients. In particular, realized regression coefficients are obtained by regressing high frequency returns for one asset on high frequency returns for another asset. When one of the assets is the market portfolio, the result is a realized beta coefficient. A beta coefficient measures the assets systematic risk as assessed by its correlation with the market portfolio. Recent examples of papers that have obtained empirical estimates of realized betas include Andersen, Bollerslev, Diebold and Wu [4], Todorov and Bollerslev [128], Dovonon, Gonçalves and Meddahi [63], Mancini and Gobbi [100].

Let us stress that large deviations for the realized correlation can not be deduced from unidimensional quantities and were thus largely ignored. As an application of our main results, we provide a large and moderate deviation principle for the realized correlation and the realized regression coefficients in some special cases. The realized regression coefficient from regressing is $\beta^n_{\ell,t}(X) = C^n_t(X)/Q^n_{\ell,t}(X)$ which consistently estimates $\beta_{\ell,t} = C_t/Q_{\ell,t}$ and the realized correlation coefficient is $\varrho^n_t(X) = C^n_t(X)/\sqrt{Q^n_{1,t}(X)Q^n_{2,t}(X)}$ which estimates $\varrho_t = C_t/\sqrt{Q_{1,t}Q_{2,t}}$. The application will be based essentially on an application of the delta method, developped by Gao and Zhao ([73]).

To be complete, let us now recall some basic definitions of the large deviations theory (c.f [53]). Let $(\lambda_n)_{n\geq 1}$ be a sequence of nonnegative real number such that $\lim_{n\to\infty}\lambda_n = +\infty$. We say that a sequence of a random variables $(M_n)_n$ with topological state space $(S, \mathbb{S})$, where $\mathbb{S}$ is a $\sigma - algebra$ on S, satisfies a large deviation principle with speed λ_n and rate function $I : S \to [0, +\infty]$ if, for each $A \in \mathbb{S}$,

$$-\inf_{x\in A^o} I(x) \leq \liminf_{n\to\infty} \frac{1}{\lambda_n} \log \mathbb{P}\Big(M_n \in A\Big) \leq \limsup_{n\to\infty} \frac{1}{\lambda_n} \log \mathbb{P}\Big(M_n \in A\Big) \leq -\inf_{x\in \bar{A}} I(x)$$

where A^o and $\bar{A}$ denote the interior and the closure of A, respectively.

The rate function I is lower semicontinuous, *i.e.* all the sub-level sets $\{x \in S \mid I(x) \leq c\}$ are closed, for $c \geq 0$. If these level sets are compact, then I is said to be a good rate function. When the speed of the large deviation principle correspond to the regime between the central limit theorem and the law of large numbers, we talk of moderate deviation principle.

In the whole paper, for any matrix M, M^T and $\|M\|$ stand for the transpose and the Euclidean norm of M, respectively. For any square matrix M, $\det(M)$ is the determinant of M. Moreover, we will shorten large deviation principle by LDP and moderate deviation principle by MDP. We denote by $\langle\cdot,\cdot\rangle$ the usual scalar product.

The article is arranged in three upcoming sections. Section 2 is devoted to our main results on the LDP and MDP for the (co-)volatility vector. In Section 3, we deduce applications for the realized correlation and the realized regression coefficients, when σ_ℓ, for $\ell = 1, 2$ are constants. In section 4, we give the proof of these theorems.

3.3 Main results

Let $X_t = (X_{1,t}, X_{2,t})$ be the process given by (3.1) and $X_t^0 = (X_{1,t}^0, X_{2,t}^0)$ the process with $b_\ell = 0$ for $\ell = 1, 2$. We introduce the following conditions

(B) for $\ell = 1, 2$ b_ℓ is bounded.

(LDP) Assume that for $\ell = 1, 2$
– the functions $t \to \sigma_{\ell,t}$ and $t \to \rho_t$ are continuous.

(MDP) Assume that for $\ell = 1, 2$
– $\sigma_{\ell,t}^2(1-\rho_t^2)$ and $\sigma_{1,t}\sigma_{2,t}(1-\rho_t^2) \in L^2([0,1], dt)$.
– Let $(v_n)_{n\geqslant 1}$ be a sequence of positive numbers such that

$$v_n \xrightarrow[n\to\infty]{} \infty \quad \text{and} \quad \frac{v_n}{\sqrt{n}} \xrightarrow[n\to\infty]{} 0$$
$$\text{and for} \quad \ell = 1, 2 \quad \sqrt{n} v_n \max_{1\leqslant k\leqslant n} \int_{(k-1)/n}^{k/n} \sigma_{\ell,t}^2 \mathrm{d}t \xrightarrow[n\to\infty]{} 0. \quad (3.3)$$

We introduce the following function, which will play a crucial role in the calculation of the moment generating function : for $-1 < c < 1$ let for any $\lambda = (\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3$

$$P_c(\lambda) := \begin{cases} -\dfrac{1}{2}\log\left(\dfrac{(1-2\lambda_1(1-c^2))(1-2\lambda_2(1-c^2)) - (\lambda_3(1-c^2)+c)^2}{1-c^2}\right) \\ \qquad\qquad\qquad\qquad if \qquad \lambda \in \mathcal{D}_c \\ +\infty, \quad otherwise \end{cases} \quad (3.4)$$

where

$$\mathcal{D}_c = \left\{\lambda \in \mathbb{R}^3,\ \max_{\ell=1,2} \lambda_\ell < \frac{1}{2(1-c^2)} \text{ and } \prod_{\ell=1}^{2}\left(1-2\lambda_\ell(1-c^2)\right) > \left(\lambda_3(1-c^2)+c\right)^2\right\}. \quad (3.5)$$

Let us present the main results.

Large deviation

Our first result is about the large deviation of $V_1^n(X)$, i.e. at fixed time.

Theorem 3.3.1 *Let $t = 1$ be fixed.*
(1) For every $\lambda = (\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3$,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}(\exp(n \left\langle \lambda, V_1^n(X^0) \right\rangle)) = \Lambda(\lambda) := \int_0^1 P_{\rho_t}(\lambda_1\sigma_{1,t}^2, \lambda_2\sigma_{2,t}^2, \lambda_3\sigma_{1,t}\sigma_{2,t})\mathrm{d}t,$$

where the function P_c is given in (3.4).

(2) Under the conditions **(LDP)** *and* **(B)**, *the sequence $V_1^n(X)$ satisfies the LDP on $\mathbb{R}^3$ with speed n and with good rate function given by the legendre transformation of Λ, that is*

$$I_{ldp}(x) = \sup_{\lambda \in \mathbb{R}^3} \left(\langle \lambda, x \rangle - \Lambda(\lambda) \right). \tag{3.6}$$

Remark 3.3.1 *By the well-know Ellis Gärtner Theorem [53, Th.II.6], we can deduce the LDP for $V_1^n(X^0)$ once if Λ is essentially smooth, which means that :*

$$\lim_{\lambda \to \partial D_\Lambda} |\nabla \Lambda(\lambda)| = +\infty, \tag{3.7}$$

where ∂D_Λ is the boundary of D_Λ the domain of definition of Λ. It is easy to see that $|\nabla \Lambda(\lambda)|$ is given by

$$|\nabla \Lambda(\lambda)| = \int_0^1 \beta_t \frac{\sigma_{1,t}^2(1 - 2\sigma_{2,t}^2 \beta_t \lambda_2) + \sigma_{2,t}^2(1 - 2\sigma_{1,t}^2 \beta_t \lambda_1) + \sigma_{1,t}\sigma_{2,t}|\rho_t + \beta_t \sigma_{1,t}\sigma_{2,t}\lambda_3|}{(1 - 2\sigma_{1,t}^2 \beta_t \lambda_1)(1 - 2\sigma_{2,t}^2 \beta_t \lambda_2) - (\rho_t + \beta_t \sigma_{1,t}\sigma_{2,t}\lambda_3)^2} \mathrm{d}t.$$

where $\beta_t = 1 - \rho_t^2$.

The condition (3.7) is verified if we suppose that the functions $\sigma_{\ell,t}$ and ρ_t are constant.

But it seems difficult to establish a general condition to guarantee the essential smoothness of Λ, even in the deterministic case.

The main difficulty in the LDP of $V_1^n(X^0)$ resides in the case where

$$\lim_{\lambda \to \partial D_\Lambda} |\nabla \Lambda(\lambda)| < +\infty.$$

For more discussion about the failure of Gärtner-Ellis theorem for one component of the vector $V_1^n(X^0)$, for example the LDP for $Q_{\ell,1}^n(X^0)$, we refer to Djellout et al. [56, Remarks 1.1]. For an explicit example of the choose of the function $\sigma_{\ell,t}^2$ that makes the normalized cumulant generating function of $Q_{\ell,1}^n(X^0)$ not essentially smooth, see Kanaya and Otsu [85, page 552].

Remark 3.3.2 *In Djellout et al. [56] and in Kanaya and Otsu [85] to circumvent the technical difficulty of determinacy of the essential smoothness, they adopt a modified approach of Bercu et al. [10] and Bryc and Dembo [36]. In the paper of Djellout and Samoura [61] the result of the LDP for the covolatility is proved by using a contraction principle in the sens of the large deviation for the results given in the case of the volatility by Djellout et al. [56]. Here a direct application of Gärtner-Ellis theorem (or of a modified version of it) is not possible and another strategy has to be found. We will use a beautiful result by Najim [104],[105],[106].*

Let us consider the case where diffusion and correlation coefficients are constant, the rate function being easier to read (see also [125] in the purely Gaussian case, i.e. $b_\ell = 0$). Before that let us introduce the function P_c^* which is the Legendre transformation of P_c given in (3.4), for all $x = (x_1, x_2, x_3)$

$$P_c^*(x) := \begin{cases} \log\left(\dfrac{\sqrt{1-c^2}}{\sqrt{x_1 x_2 - x_3^2}}\right) - 1 + \dfrac{x_1 + x_2 - 2c x_3}{2(1-c^2)} \\ \qquad\qquad\qquad if \qquad x_1 > 0,\ x_2 > 0,\ x_1 x_2 > x_3^2 \\ +\infty, \quad otherwise. \end{cases} \tag{3.8}$$

Corollary 3.3.2 *We assume that for $\ell = 1, 2$ σ_ℓ and ρ are constants. Under the condition* **(B)***, we obtain that $V_1^n(X)$ satisfies the LDP on $\mathbb{R}^3$ with speed n and with good rate function I_{ldp}^V given by*

$$I_{ldp}^V(x_1, x_2, x_3) = P_\rho^*\left(\frac{x_1}{\sigma_1^2}, \frac{x_2}{\sigma_2^2}, \frac{x_3}{\sigma_1 \sigma_2}\right), \tag{3.9}$$

where P_c^* is given in (3.8).

In the above we have obtained finite dimensional LDP's for the estimator V_n^1 at a fixed time $t = 1$. Now, we are interested in LDP's that depends on a whole random process that evolves over time. So, we shall extend the Theorem 3.3.1 to the process-level large deviations, i.e. for trajectories $(V_t^n(X))_{0 \le t \le 1}$, which is interesting from the viewpoint of non-parametric statistics.

Let $\mathcal{B}V([0,1], \mathbb{R}^3)$ (shortened in $\mathcal{B}V$) be the space of functions of bounded variation on $[0,1]$. We identify $\mathcal{B}V$ with $\mathcal{M}_3([0,1])$, the set of vector measures with value in $\mathbb{R}^3$. This is done in the usual manner : to $f \in \mathcal{B}V$ there corresponds μ^f caracterized by $\mu^f([0,t]) = f(t)$. Up to this identification, $\mathcal{C}_3([0,1])$ the set of $\mathbb{R}^3$-valued continuous bounded functions on $[0,1]$), is the topological dual of $\mathcal{B}V$. We endow $\mathcal{B}V$ with the weak-* convergence topology $\sigma(\mathcal{B}V, \mathcal{C}_3([0,1]))$ (shortened σ_w) and with the associated Borel $\sigma-$field $\mathcal{B}_w$. Let $f \in \mathcal{B}V$ and μ^f the associated measure in $\mathcal{M}_3([0,1])$. Consider the Lebesgue decomposition of μ^f, $\mu^f = \mu_a^f + \mu_s^f$ where μ_a^f denotes the absolutely continous part of μ^f with respect to dx and μ_s^f its singular part. We denote by $f_a(t) = \mu_a^f([0,t])$ and by $f_s(t) = \mu_s^f([0,t])$.

Theorem 3.3.3 *Under the conditions* **(LDP)** *and* **(B)***, the sequence $V_\cdot^n(X)$ satisfies the LDP on $\mathcal{B}V$ with speed n and with rate function J_{ldp} given for* *any* $f = (f_1, f_2, f_3) \in \mathcal{B}V$ *by*

$$J_{ldp}(f) = \int_0^1 P^*_{\rho_t}\left(\frac{f'_{1,a}(t)}{\sigma^2_{1,t}}, \frac{f'_{2,a}(t)}{\sigma^2_{2,t}}, \frac{f'_{3,a}(t)}{\sigma_{1,t}\sigma_{2,t}}\right)\mathrm{d}t$$
$$+ \int_0^1 \frac{\sigma^2_{2,t}f'_{1,s}(t) + \sigma^2_{1,t}f'_{2,s}(t) - 2\rho_t\sigma_{1,t}\sigma_{2,t}f'_{3,s}(t)}{2\sigma^2_{1,t}\sigma^2_{2,t}(1-\rho_t^2)} 1_{[t; f'_{1,s}>0, f'_{2,s}>0, (f'_{3,s})^2 < f'_{1,s}f'_{2,s}]}\,\mathrm{d}\theta(t), \tag{3.10}$$

where P^*_c *is given in (3.8) and* θ *is any real-valued nonnegative measure with respect to which* μ^f_s *is absolutely continuous and* $f'_s = d\mu^f_s/d\theta = (f'_{1,s}, f'_{2,s}, f'_{3,s})$.

Remark 3.3.3 *Note that the definition of* f'_s *is* $\theta-$*dependent. However, by homogeneity,* J_{ldp} *does not depend upon* θ*. One can choose* $\theta = |f_{1,s}| + |f_{2,s}| + |f_{3,s}|$*, with* $|f_{i,s}| = f^+_{i,s} + f^-_{i,s}$*, where* $f_{i,s} = f^+_{i,s} - f^-_{i,s}$ *by the Hahn-Jordan decomposition.*

Remark 3.3.4 *As stated above, the problem of the LDP for* $Q^n_{\ell,\cdot}(X)$ *and* $C^n_\cdot(X)$ *was alreay studied by Djellout et al. [56] and [61], and the rate function is given explicitly in the last case. This is the first time that the LDP is investigated for the vector of the (co-)volatility. For the trajectoriel case in [56] they are much inspired from the two lines of the studies : the work of Lynch and Sethuraman [97] on large deviations of processes with independent increments, and Pukhalski [114] about large deviations of stochastic processes. The main diffuclty in Djellout et al. [56] is the identification of the rate function with a correctif term. Kanaya and Otsu [85] do not consider the LDP of the trajectoriel case.*

Remark 3.3.5 *By using the contraction principle, and if* σ_ℓ *is strictly positive, we may find back the result of [56], i.e. that* $Q^n_{\ell,\cdot}$ *satisfies a LDP with speed* n *and rate function*

$$J^{\sigma_\ell}_{ldp}(f) = \int_0^1 \mathcal{P}^*\left(\frac{f'_a(t)}{\sigma^2_{\ell,t}}\right)dt + \frac{1}{2}\int_0^1 \frac{1}{\sigma^2_{\ell,t}}d|f_s|(t)$$

where $\mathcal{P}^*(x) = \frac{1}{2}(x - 1 - \log(x))$ *when* x *is positive and infinite if non positive, using the same notation as in the theorem (with* $\theta = |f_s|$*). One may also obtain the LDP for* $C^n_\cdot$ *by the contraction principle, recovering the result of Djellout-Yacouba [61] (see there for the quite explicit complicated rate function).*

Remark 3.3.6 *The continuity assumptions in* **(LDP)** *of* $\sigma_{\ell,\cdot}$ *and* $\rho_\cdot$ *is not necessary, but in this case we have to consider another strategy of the proof, more technical and relying on Dawson-Gärtner type theorem, which moreover does not enable to get other precision on the rate function that the fact*

it is a good rate function.
However it is not hard to adapt our proof to the case where $\sigma_{\ell,\cdot}$ and $\rho_\cdot$ have only a finite number of discontinuity points (of the first type). This can be done by applying the previous theorem to each subinterval where all functions are continuous and using the independence of the increments of $V_t^n(X^0)$.

Remark 3.3.7 *All the results remain true when the gird of discretization is not regular. The important property is that every point is in the support of the limit measure.*

Remark 3.3.8 *We can derive LDP and MDP for the realized co-volatlity vector statistics given the path of the volatility process when the volatility process is random but independent of the noise process by a Brownian motion. The proofs are easily modified to imply the main results by interpreting the probability and mathematical expectation as conditional ones given the path of the volatility process, see Kanaya and Otsu [86].*

Moderate deviation

Let us now considered the intermediate scale between the central limit theorem and the law of large numbers.

Theorem 3.3.4 *For t=1 fixed. Under the conditions* **(MDP)** *and* **(B)***, the sequence*

$$\frac{\sqrt{n}}{v_n}\left(V_1^n(X) - [V]_1\right)$$

satisfies the LDP on $\mathbb{R}^3$ with speed v_n^2 and with rate function given by

$$I_{mdp}(x) = \sup_{\lambda\in\mathbb{R}^3}\left(\langle\lambda, x\rangle - \frac{1}{2}\langle\lambda, \Sigma_1\cdot\lambda\rangle\right) = \frac{1}{2}\left\langle x, \Sigma_1^{-1}\cdot x\right\rangle \tag{3.11}$$

with

$$\Sigma_1 = \begin{pmatrix} \int_0^1 \sigma_{1,t}^4 \mathrm{d}t & \int_0^1 \sigma_{1,t}^2\sigma_{2,t}^2\rho_t^2 \mathrm{d}t & \int_0^1 \sigma_{1,t}^3\sigma_{2,t}\rho_t \mathrm{d}t \\ \int_0^1 \sigma_{1,t}^2\sigma_{2,t}^2\rho_t^2 \mathrm{d}t & \int_0^1 \sigma_{2,t}^4 \mathrm{d}t & \int_0^1 \sigma_{1,t}\sigma_{2,t}^3\rho_t \mathrm{d}t \\ \int_0^1 \sigma_{1,t}^3\sigma_{2,t}\rho_t \mathrm{d}t & \int_0^1 \sigma_{1,t}\sigma_{2,t}^3\rho_t \mathrm{d}t & \int_0^1 \frac{1}{2}\sigma_{1,t}^2\sigma_{2,t}^2(1+\rho_t^2) \mathrm{d}t \end{pmatrix}.$$

Remark 3.3.9 *If for some $p > 2$, $\sigma_{1,t}^2$, $\sigma_{2,t}^2$ and $\sigma_{1,t}\sigma_{2,t}(1-\rho_t^2) \in L^p([0,1])$ and $b_n = O(n^{\frac{1}{2}-\frac{1}{p}})$, the condition (3.3) in* **(MDP)** *is verified.*

Let $\mathcal{H}$ be the banach space of $\mathbb{R}^3$-valued right-continuous-left-limit non decreasing functions γ on $[0,1]$ with $\gamma(0)=0$, equipped with the uniform norm and the $\sigma-$field $\mathcal{B}^s$ generated by the coordinate $\{\gamma(t), 0\leqslant t\leqslant 1\}$.

Theorem 3.3.5 *Under the conditions* **(MDP)** *and* **(B)**, *the sequence*

$$\frac{\sqrt{n}}{v_n}\left(V_\cdot^n(X)-[V]_\cdot\right)$$

satisfies the LDP on $\mathcal{H}$ with speed v_n^2 and with rate function given by

$$J_{mdp}(\phi)=\begin{cases}\displaystyle\int_0^1 \frac{1}{2}\left\langle\dot{\phi}(t),\overline{\Sigma}_t^{-1}\cdot\dot{\phi}(t)\right\rangle dt & if\quad \phi\in\mathcal{AC}_0([0,1])\\ +\infty, & otherwise,\end{cases}\tag{3.12}$$

where

$$\overline{\Sigma}_t=\begin{pmatrix}\sigma_{1,t}^4 & \sigma_{1,t}^2\sigma_{2,t}^2\rho_t^2 & \sigma_{1,t}^3\sigma_{2,t}\rho_t\\ \sigma_{1,t}^2\sigma_{2,t}^2\rho_t^2 & \sigma_{2,t}^4 & \sigma_{1,t}\sigma_{2,t}^3\rho_t\\ \sigma_{1,t}^3\sigma_{2,t}\rho_t & \sigma_{1,t}\sigma_{2,t}^3\rho_t & \frac{1}{2}\sigma_{1,t}^2\sigma_{2,t}^2(1+\rho_t^2)\end{pmatrix}$$

is invertible with $\det(\overline{\Sigma}_t)=\frac{1}{2}\sigma_{1,t}^6\sigma_{2,t}^6(1-\rho_t^2)^3$ *and* $\overline{\Sigma}_t^{-1}$ *his inverse given by*

$$\overline{\Sigma}_t^{-1}=\frac{1}{\det(\overline{\Sigma}_t)}\begin{pmatrix}\frac{1}{2}\sigma_{1,t}^2\sigma_{2,t}^6(1-\rho_t^2) & \frac{1}{2}\sigma_{1,t}^4\sigma_{2,t}^4\rho_t^2(1-\rho_t^2) & -\sigma_{1,t}^3\sigma_{2,t}^5\rho_t(1-\rho_t^2)\\ \frac{1}{2}\sigma_{1,t}^4\sigma_{2,t}^4\rho_t^2(1-\rho_t^2) & \frac{1}{2}\sigma_{1,t}^6\sigma_{2,t}^2(1-\rho_t^2) & -\sigma_{1,t}^5\sigma_{2,t}^3\rho_t(1-\rho_t^2)\\ -\sigma_{1,t}^3\sigma_{2,t}^5\rho_t(1-\rho_t^2) & -\sigma_{1,t}^5\sigma_{2,t}^3\rho_t(1-\rho_t^2) & \sigma_{1,t}^4\sigma_{2,t}^4(1-\rho_t^4)\end{pmatrix},$$

and $\mathcal{AC}_0=\{\phi:[0,1]\to\mathbb{R}^3$ *is absolutely continuous with* $\phi(0)=0\}$.

Let us note once again that it is the first time the MDP is considered for the vector of (co)-volatility.

In the previous results, we have imposed the boundedness of the drift b_ℓ which allows us to reduce quite easily the LDP and MDP of $V^n(X)$ to those of $V^n(X^0)$ (no drift case). It is very natural to ask whether they continue to hold under a Lipchitzian condition or more generally linear growth condition of the drift $b(t,x)$, rather than the boundedness. When $b_\ell(t,x)$ is know, we consider the following estimator $\tilde{V_t^n}(X)=\left(\tilde{Q_{1,t}^n}(X),\tilde{Q_{2,t}^n}(X),\tilde{C_t^n}(X)\right)^T$ with for $\ell=1,2$

$$\tilde{Q_{\ell,t}^n}(X)=\sum_{k=1}^{[nt]}\left(\Delta_k^n X_\ell-b_{\ell,t_{k-1}^n}(X_{t_{k-1}^n})(t_k^n-t_{k-1}^n)\right)^2,$$

$$\tilde{C}_t^n(X) = \sum_{k=1}^{[nt]} \left(\Delta_k^n X_1 - b_{1,t_{k-1}^n}(X_{t_{k-1}^n})(t_k^n - t_{k-1}^n) \right) \left(\Delta_k^n X_2 - b_{2,t_{k-1}^n}(X_{t_{k-1}^n})(t_k^n - t_{k-1}^n) \right).$$

We have the following

Theorem 3.3.6 *Let $X_t = (X_{1,t}, X_{2,t})$ be given by (3.1), with $(X_{1,0}, X_{2,0})$ bounded. We assume that the drift b_ℓ satisfies the following uniform linear growth condition : $\forall s,t \in [0,1], x,y \in \mathbb{R}^2$*

$$|b_\ell(t,x) - b_\ell(s,y)| \leqslant C[1 + \|x-y\| + \eta(|t-s|)(\|x\| + \|y\|)], \qquad (3.13)$$

where $C > 0$ is a constant and $\eta : [0,\infty) \to [0,\infty)$ is a continuous non-decreasing function with $\eta(0) = 0$.

(1) Under the condition **(LDP)***, the sequence $\tilde{V}_\cdot^n(X)$ satisfies the LDPs in Theorem 3.3.1 and Theorem 3.3.3.*

(2) Under the condition **(MDP)***, the sequence $\frac{\sqrt{n}}{v_n}(\tilde{V}_\cdot^n(X) - [V]_\cdot)$ satisfies the MDPs in Theorem 3.3.4 and Theorem 3.3.5.*

Remark 3.3.10 *The previous result is intereseting only in the case where the drift is known. Since the drift b is (almost) never known, this result should be considered as contributions to probability theory rather than inferential statistics.*

Remark 3.3.11 *As it can be remarked, the LDP and the MDP are established here for $\tilde{V}^n$ instead of V^n. The minimal condition on the drift to obtain LDP for V^n, is for all $\delta > 0$*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P} \left(\sum_{k=1}^{n} \left(\int_{t_{k-1}}^{t_k} b(s, X_s) \mathrm{d}s \right)^2 > \delta\varepsilon(n) \right) = -\infty, \qquad (3.14)$$

for some sequence $\varepsilon(n)$ which goes to zero when n goes ton infinityThe condition (3.14) is satisfied for example when the drift is sublinear, which means that

$$|b(x)| \leq C|x|^\alpha, \qquad \alpha < 1,$$

where $C > 0$ is a constant (see the end of the proof of Theorem 3.3.6 for more details). Recall that in Kanaya and Otsu [85] the dirft is assumed to be bounded.

3.4 Applications : Large deviations for the realized correlation and the realized regression coefficients

In this section we apply our results to obtain the LDP and MDP for the standard dependence measures between the two assets returns such as the realized regression coefficients up to time 1, $\beta_{\ell,1} = C_1/Q_{\ell,1}$ for $\ell = 1,2$ and the realized correlation $\varrho_1 = C_1/\sqrt{Q_{1,1}Q_{2,1}}$ which are estimated by $\beta^n_{\ell,1}(X) = C^n_1(X)/Q^n_{\ell,1}(X)$ and $\varrho^n_1(X) = C^n_1(X)/\sqrt{Q^n_{1,1}(X)Q^n_{2,1}(X)}$ respectively. To simplify the argument, we focus in the case where σ_ℓ for $\ell = 1,2$ are constants and we denote $\varrho := \int_0^1 \rho_t dt$. The consistency and the central limit theorem for these estimators were already studied see for example Mancini and Gobbi [100]. To our knowledge, however no results are known for the LDP and MDP.

Correlation coefficient

Proposition 3.4.1 *Let for $\ell = 1,2$, σ_ℓ are constants and $\varrho := \int_0^1 \rho_t dt$. Under the conditions* **(LDP)** *and* **(B)**, *the sequence $\varrho^n_1(X)$ satisfies the LDP on $\mathbb{R}$ with speed n and with good rate function given by*

$$I^{\varrho}_{ldp}(u) = \inf_{\{(x,y,z)\in\mathbb{R}^3 : u = \frac{z}{\sqrt{xy}}\}} I_{ldp}(x,y,z)$$

where I_{ldp} is given in (3.6).

Once again, let us specify the rate function in the case of constant correlation.

Corollary 3.4.2 *We suppose that for $\ell = 1,2$, σ_ℓ and ρ are constant. Under the condition* **(B)**, *we obtain that $\varrho^n_1(X)$ satisfies the LDP on $\mathbb{R}$ with speed n and with good rate function given by*

$$I^{\rho}_{ldp}(u) = \begin{cases} \log\left(\dfrac{1-\rho u}{\sqrt{1-\rho^2}\sqrt{1-u^2}}\right), & -1 < u < 1 \\ +\infty, & otherwise. \end{cases} \tag{3.15}$$

As the reader can imagine from the rate function expression, it is quite a simple application of the contraction principle starting from the LDP of the realized (co)-volatility. As will be seen from the proof, in this case, the MDP is harder to establish and requires a more subtle technology : large deviations for the delta-method.

Proposition 3.4.3 *Let for $\ell = 1,2$, σ_ℓ are constants and $\varrho := \int_0^1 \rho_t dt$. Under the conditions* **(MDP)** *and* **(B)**, *the sequence $\dfrac{\sqrt{n}}{v_n}\left(\varrho^n_1(X) - \varrho\right)$ satisfies the LDP on $\mathbb{R}$ with speed v_n^2 and with rate function given by*

$$I^{\varrho}_{mdp}(u) = \inf_{\{(x,y,z)\in\mathbb{R}^3 : u = \frac{z}{\sigma_1\sigma_2} - \varrho\frac{\sigma_1^2 y + x\sigma_2^2}{2\sigma_1^2\sigma_2^2}\}} I_{mdp}(x,y,z)$$

where I_{mdp} is given in (3.11).

Corollary 3.4.4 *We suppose that for $\ell = 1, 2$, σ_ℓ and ρ are constant. Under the condition* **(B)**, *we obtain that* $\frac{\sqrt{n}}{v_n}(\varrho_1^n(X) - \rho)$ *satisfies the LDP on* $\mathbb{R}$ *with speed* v_n^2 *and with good rate function given for all* $u \in \mathbb{R}$ *by*

$$I_{mdp}^{\rho}(u) = \frac{u^2}{(1-\rho^2)^2}. \tag{3.16}$$

Regression coefficient

The strategy initiated for the correlation coefficient is even simpler in the case of regression coefficient.

Proposition 3.4.5 *Let for $\ell = 1, 2$, σ_ℓ are constants. Under the conditions* **(LDP)** *and* **(B)**, *for $\ell = 1$ or 2, the sequence* $\beta_{l,1}^n(X)$ *satisfies the LDP on* $\mathbb{R}$ *with speed n and with good rate function given by*

$$I_{ldp}^{\beta_{\ell,1}}(u) = \inf_{\{(x_1,x_2,x_3)\in\mathbb{R}^3 : u = \frac{x_3}{x_\ell}\}} I_{ldp}(x_1, x_2, x_3)$$

where I_{ldp} is given in (3.6).

Once again, this Proposition is a simple application of the contraction principle. Let us specify the rate function in the case of constant correlation.

Corollary 3.4.6 *We suppose that for $\ell = 1, 2$, σ_ℓ and ρ are constant. Under the condition* **(B)**, *we obtain that* $\beta_{l,1}^n(X)$ *satisfies the LDP on* $\mathbb{R}$ *with speed n and with good rate function given for $\iota = 1, 2$ with $\ell \neq \iota$ and for all $u \in \mathbb{R}$ by*

$$I_{ldp}^{\beta_l}(u) = \frac{1}{2}\log\left(1 + \frac{(\sigma_\ell u - \rho\sigma_\iota)^2}{\sigma_\iota^2(1-\rho^2)}\right). \tag{3.17}$$

We may also consider the MDP.

Proposition 3.4.7 *Let for $\ell = 1, 2$, σ_ℓ are constants and $\varrho := \int_0^1 \rho_t dt$. Under the conditions* **(MDP)** *and* **(B)** *and for $\ell, \iota \in \{1, 2\}$ with $\ell \neq \iota$, the sequence* $\frac{\sqrt{n}}{v_n}(\beta_{\ell,1}^n(X) - \varrho\frac{\sigma_\iota}{\sigma_\ell})$ *satisfies the LDP on* $\mathbb{R}$ *with speed* v_n^2 *and with rate function given by*

$$I_{mdp}^{\beta_{\ell,1}}(u) = \inf_{\{(x,y,z)\in\mathbb{R}^3 : u = \frac{z}{\sigma_\ell^2} - \varrho\frac{\sigma_\iota}{\sigma_\ell^3}x\}} I_{mdp}(x, y, z)$$

where I_{mdp} is given in (3.11).

Corollary 3.4.8 *We suppose that for $\ell = 1, 2$, σ_ℓ and ρ are constant. Under the condition* **(B)** *and for $\ell, \iota \in \{1, 2\}$ with $\ell \neq \iota$, we obtain that* $\frac{\sqrt{n}}{v_n}(\beta_{\ell,1}^n(X) - \rho\frac{\sigma_\iota}{\sigma_\ell})$ *satisfies the LDP on* $\mathbb{R}$ *with speed* v_n^2 *and with good rate function given for all* $u \in \mathbb{R}$ *by*

$$I_{mdp}^{\beta_{\ell,1}^c}(u) = \frac{\sigma_\ell^2 u^2}{\sigma_\iota^2(1-\rho^2)}. \tag{3.18}$$

3.5 Proof

Let us say a few words on our strategy of proof. As the reader may have guessed, one of the important step is first to consider the no-drift case, where we have to deal with non homogenous quadratic forms of Gaussian processes (in the vector case). In these non essentially smooth case (in the terminology of Gärtner-Ellis), we will use (after some technical approximations) powerful recent results of Najim [105]. In a second step, we see how to reduce the general case to the no-drift case.

Befor stating the proof, we need some additional notations. For any process Z_t, $\Delta_k^n Z$ stands for the increment $Z_{t_k^n} - Z_{t_{k-1}^n}$. For a sequence of random variables $(Z_n)_n$ on $\mathbb{R}^{d\times p}$, we say that $(Z_n)_n$ converges $(\lambda_n)-$superexponentially fast in probability to some random variable Z if, for all $\delta > 0$,

$$\limsup_{n\to\infty} \frac{1}{\lambda_n} \log \mathbb{P}\Big(\|Z_n - Z\| > \delta\Big) = -\infty.$$

This exponential convergence with speed λ_n will be shortened as

$$Z_n \quad \underset{\lambda_n}{\overset{\text{superexp}}{\longrightarrow}} Z.$$

Proof of Theorem 3.3.1

Lemma 3.5.1 *If $(\xi, \xi^{'})$ is a centered Gaussian random vector with covariance $\begin{pmatrix} 1 & c \\ c & 1 \end{pmatrix}$, with $-1 < c < 1$. Then for all $(\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3$*

$$\log \mathbb{E} \exp\Big(\lambda_1 \xi^2 + \lambda_2 {\xi^{'}}^2 + \lambda_3 \xi \xi^{'}\Big) = P_c(\lambda_1, \lambda_2, \lambda_3),$$

where the function P_c is given in in (3.4).

Proof :

Elementary.

Lemma 3.5.2 *Let $X_t = (X_{1,t}, X_{2,t})$ given (3.1). We have for every $\lambda \in \mathbb{R}^3$*

$$\Lambda_n(\lambda) := \frac{1}{n} \log \mathbb{E}\left(\exp\left(n\left\langle \lambda, V_1^n(X^0)\right\rangle\right)\right) \leqslant \Lambda(\lambda) := \int_0^1 P_{\rho_t}(\lambda_1\sigma_{1,t}^2, \lambda_2\sigma_{2,t}^2, \lambda_3\sigma_{1,t}\sigma_{2,t})\mathrm{d}t,$$

where the function P_c is given in in (3.4), and

$$\lim_{n\to\infty} \Lambda_n(\lambda) = \Lambda(\lambda).$$

Proof :

For $\ell = 1, 2$, we have

$$Q^n_{\ell,t}(X^0) = \sum_{k=1}^{[nt]} a_{\ell,k}\xi^2_{\ell,k} \qquad \text{and} \quad C^n_t(X^0) = \sum_{k=1}^{[nt]} \sqrt{a_{1,k}}\sqrt{a_{2,k}}\xi_{1,k}\xi_{2,k}$$

where

$$\xi_{\ell,k} := \frac{\int_{t^n_{k-1}}^{t^n_k} \sigma_{\ell,s}\mathrm{d}B_{\ell,s}}{\sqrt{a_{\ell,k}}} \quad \text{and} \quad a_{\ell,k} := \int_{t^n_{k-1}}^{t^n_k} \sigma^2_{\ell,s}\mathrm{d}s. \tag{3.19}$$

Obviously $((\xi_{1,k}, \xi_{2,k}))_{k=1\cdots n}$ are independent centered Gaussian random vector with covariance matrix $\begin{pmatrix} 1 & c^n_k \\ c^n_k & 1 \end{pmatrix}$ where

$$c^n_k := \frac{\vartheta^n_k}{\sqrt{a_{1,k}}\sqrt{a_{2,k}}} \qquad \text{and} \quad \vartheta^n_k := \int_{t^n_{k-1}}^{t^n_k} \sigma_{1,s}\sigma_{2,s}\rho_s\mathrm{d}s. \tag{3.20}$$

We use the Lemma 3.5.1 and the martingale convergence theorem (or the classical Lebesgue derivation theorem) to get the final assertions (see for example [56, p.204] for details).

Proof Theorem 3.3.1

(1) It is contained in lemma 3.5.2.

(2) We shall prove it in three steps.

<u>*Part 1*</u>. At first, we consider that the drift $b_\ell = 0$. In this case $V^n_t(X) = V^n_t(X^0)$. We recall that since $B_{2,t} = \rho_t dB_{1,t} + \sqrt{1-\rho^2_t}dB_{3,t}$, we may rewrite (3.1) as

$$\begin{cases} dX_{1,t} &= \sigma_{1,t}dB_{1,t} \\ dX_{2,t} &= \sigma_{2,t}(\rho_t dB_{1,t} + \sqrt{1-\rho^2_t}dB_{3,t}) \end{cases} \tag{3.21}$$

Using the approximation Lemma in [53], we shall prove that

$$V^n_1(X-Y) = \sum_{k=1}^n \left((\Delta^n_k X_1)^2, (\Delta^n_k X_2)^2, (\Delta^n_k X_1)(\Delta^n_k X_2) \right)^T$$

will satisfy the same LDPs as

$$W^n_1 := \frac{1}{n}\sum_{k=1}^n \begin{pmatrix} \sigma^2_{1,\frac{k-1}{n}} N^2_{1,k} \\ \left(\sigma_{2,\frac{k-1}{n}}\rho_{\frac{k-1}{n}} N_{1,k} + \sigma_{2,\frac{k-1}{n}}\sqrt{1-\rho^2_{\frac{k-1}{n}}} N_{3,k}\right)^2 \\ \sigma_{1,\frac{k-1}{n}} N_{1,k}\left(\sigma_{2,\frac{k-1}{n}}\rho_{\frac{k-1}{n}} N_{1,k} + \sigma_{2,\frac{k-1}{n}}\sqrt{1-\rho^2_{\frac{k-1}{n}}} N_{3,k}\right) \end{pmatrix},$$

where $N_{\ell,k} := \int_{t_{k-1}^n}^{t_k^n} \sqrt{n} \mathrm{d}B_{\ell,s}$, for $\ell = 1, 3$.

Let us first focus on the LDP of W_1^n. We will use Najim result [105, Theorem 2.2] to prove that. It is easy to see that W_1^n can be rewritten as

$$W_1^n = \frac{1}{n} \sum_{k=1}^{n} F\left(\frac{k-1}{n}\right) Z_k$$

where

$$F(k/n) = \begin{pmatrix} f_1(k/n) \\ f_2(k/n) \\ f_3(k/n) \end{pmatrix} = \begin{pmatrix} \sigma^2_{1,\frac{k}{n}} & 0 & 0 \\ \sigma^2_{2,\frac{k}{n}} \rho^2_{\frac{k}{n}} & \sigma^2_{2,\frac{k}{n}} (1-\rho^2_{\frac{k}{n}}) & 2\sigma^2_{2,\frac{k}{n}} \rho_{\frac{k}{n}} \sqrt{1-\rho^2_{\frac{k}{n}}} \\ \sigma_{1,\frac{k}{n}} \sigma_{2,\frac{k}{n}} \rho_{\frac{k}{n}} & 0 & \sigma_{1,\frac{k}{n}} \sigma_{2,\frac{k}{n}} \sqrt{1-\rho^2_{\frac{k}{n}}} \end{pmatrix} \tag{3.22}$$

and

$$Z_j = \left(N^2_{1,j},\ N^2_{3,j},\ N_{1,j} N_{3,j}\right)^T. \tag{3.23}$$

Obviously $(N_{1,k}, N_{3,k})_{k=1\cdots n}$ are independent centered Gaussian random vectors with identity covariance matrix.

Using [105, Theorem 2.2], we deduce that W_1^n satisfies the LDP on $\mathbb{R}^3$ with speed n and with good rate function given by for all $x \in \mathbb{R}^3$

$$I(x) = \sup_{\lambda \in \mathbb{R}^3} \left(\langle \lambda, x \rangle - \int_0^1 L\left(\sum_{j=1}^{3} \lambda_i \cdot f_i(t) \right) dt \right),$$

with

$$\sum_{i=1}^{3} \lambda_i f_i(t) = \begin{pmatrix} \lambda_1 \sigma^2_{1,t} + \lambda_2 \sigma^2_{2,t} \rho^2_t + \lambda_3 \sigma_{1,t} \sigma_{2,t} \rho_t \\ \lambda_2 \sigma^2_{2,t} (1-\rho^2_t) \\ 2\lambda_2 \sigma^2_{2,t} \rho_t \sqrt{1-\rho^2_t} + \lambda_3 \sigma_{1,t} \sigma_{2,t} \sqrt{1-\rho^2_t} \end{pmatrix}^T,$$

and for $\lambda \in \mathbb{R}^3$

$$L(\lambda) := \log \mathbb{E} \exp \langle \lambda, Z_1 \rangle = P_0(\lambda_1, \lambda_2, \lambda_3),$$

where P_0 is given in (3.4). In this cas it takes a simpler form wich we recall here :

$$P_0(\lambda_1, \lambda_2, \lambda_3) = -\frac{1}{2} \log\left((1-2\lambda_1)(1-2\lambda_2) - \lambda_3^2 \right).$$

An easy calculation gives us that

$$\int_0^1 P_0\left(\sum_{j=1}^{3} \lambda_i \cdot f_i(t) \right) dt = \int_0^1 P_{\rho_t}\left(\lambda_1 \sigma^2_{1,t}, \lambda_2 \sigma^2_{2,t}, \lambda_3 \sigma_{1,t} \sigma_{2,t} \right) dt,$$

so $I(x) = I_{ldp}(x)$, where I_{ldp} is given in (3.6).

<u>*Part 2*</u>. Now we shall prove that V_1^n and W_1^n satisfy the same LDPs, by means of the approximation Lemma in [56]. We have to prove that

$$V_1^n(X^0) - W_1^n \quad \overset{\text{superexp}}{\underset{n}{\longrightarrow}} 0.$$

We do this element by element. We will only consider one element, the other terms can be dealt with in the same way. We have to prove that for $q = 1, 2, 3$

$$\mathcal{R}_{q,1}^n \quad \overset{\text{superexp}}{\underset{n}{\longrightarrow}} 0, \tag{3.24}$$

where

$$\mathcal{R}_{1,t}^n := \sum_{k=1}^{[nt]} (\Delta_k^n X_1)^2 - \frac{1}{n}\sum_{k=1}^{[nt]} \sigma_{1,\frac{k-1}{n}}^2 N_{1,k}^2, \tag{3.25}$$

$$\mathcal{R}_{2,t}^n := \sum_{k=1}^{[nt]} (\Delta_k^n X_2)^2 - \frac{1}{n}\sum_{k=1}^{[nt]} \left(\sigma_{2,\frac{k-1}{n}}\rho_{\frac{k-1}{n}} N_{1,k} + \sigma_{2,\frac{k-1}{n}}\sqrt{1-\rho_{\frac{k-1}{n}}^2}N_{3,k}\right)^2, \tag{3.26}$$

and

$$\mathcal{R}_{3,t}^n := \sum_{k=1}^{[nt]} \Delta_k^n X_1 \Delta_k^n X_2 - \frac{1}{n}\sum_{k=1}^{[nt]} \sigma_{1,\frac{k-1}{n}} N_{1,k}\left(\sigma_{2,\frac{k-1}{n}}\rho_{\frac{k-1}{n}} N_{1,k} + \sigma_{2,\frac{k-1}{n}}\sqrt{1-\rho_{\frac{k-1}{n}}^2}N_{3,k}\right). \tag{3.27}$$

At first, we start the negligibility (3.24) with the quantity $\mathcal{R}_{1,1}^n$ which can be rewritten as

$$\mathcal{R}_{1,1}^n = \sum_{k=1}^{n} R_{-,k}^n R_{+,k}^n,$$

with $R_{\pm,k}^n := \int_{t_{k-1}^n}^{t_k^n} (\sigma_{1,s} \pm \sigma_{1,\frac{k-1}{n}})\mathrm{d}B_{1,s}$, where $((R_{-,k}^n, R_{+,k}^n))_{k=1\cdots n}$ are independent centered Gaussian random vectors with covariance $\begin{pmatrix} \varepsilon_{-,k}^n & \eta_k^n \\ \eta_k^n & \varepsilon_{+,k}^n \end{pmatrix}$ where

$$\varepsilon_{\pm,k}^n = \int_{t_{k-1}^n}^{t_k^n} \left(\sigma_{1,s} \pm \sigma_{1,\frac{k-1}{n}}\right)^2 \mathrm{d}s \qquad \text{and} \qquad \eta_k^n = \int_{t_{k-1}^n}^{t_k^n} \left(\sigma_{1,s}^2 - \sigma_{1,\frac{k-1}{n}}^2\right)\mathrm{d}s \tag{3.28}$$

So by Chebyshev's inequality, we have for all $r, \lambda > 0$,

$$\frac{1}{n}\log \mathbb{P}\left(\mathcal{R}_{1,1}^n > r\right) \leqslant -r\lambda + \frac{1}{n}\log \mathbb{E}\exp\left(n\lambda\mathcal{R}_{1,1}^n\right). \tag{3.29}$$

A simple calculation gives us

$$
\begin{aligned}
\frac{1}{n}\log\mathbb{E}\exp\left(n\lambda\mathcal{R}_{1,1}^n\right) &= \frac{1}{n}\sum_{k=1}^n \log\mathbb{E}\exp\left(n\lambda R_{+,k}^n R_{-,k}^n\right) \\
&\quad -\frac{1}{2n}\sum_{k=1}^n \log\left[\frac{\varepsilon_{+,k}^n\varepsilon_{-,k}^n - \left(n\lambda(\varepsilon_{+,k}^n\varepsilon_{-,k}^n - (\eta_k^n)^2) + \eta_k^n\right)^2}{\varepsilon_{+,k}^n\varepsilon_{-,k}^n - (\eta_k^n)^2}\right] \\
&= -\frac{1}{2n}\sum_{k=1}^n \log\left[1 - n^2\lambda^2(\varepsilon_{+,k}^n\varepsilon_{-,k}^n - (\eta_k^n)^2) - n\lambda\eta_k^n\right] \\
&= \int_0^1 K(f_n(t))\mathrm{d}t \qquad (3.30)
\end{aligned}
$$

where K is given by

$$
K(\lambda) := \begin{cases} -\dfrac{1}{2}\log(1-2\lambda) & if \quad \lambda < \frac{1}{2} \\ +\infty, & \text{otherwise,} \end{cases}
$$

and

$$
f_n(t) = \sum_{k=1}^n \mathbf{1}_{(t_k^n - t_{k-1}^n)}(t)\left[\lambda^2\left(\frac{\varepsilon_{+,k}^n}{t_k^n - t_{k-1}^n}\frac{\varepsilon_{-,k}^n}{t_k^n - t_{k-1}^n} - \left(\frac{\eta_k^n}{t_k^n - t_{k-1}^n}\right)^2\right) + 2\lambda\left(\frac{\eta_k^n}{t_k^n - t_{k-1}^n}\right)\right]
$$

where $\varepsilon_{\pm,k}^n$ and η_k^n are given in (3.28).

By the continuity condition of the assumption **(LDP)** and the classical Lebesgue derivation theorem, we have that

$$
f_n(t) \longrightarrow 0 \qquad \text{as} \quad n \to \infty.
$$

By the classical Lebesgue derivation theorem we have that the right hand of the equality (3.30) goes to 0.

Letting n goes to infinity and then λ goes to infinity in (3.29), we obtain that

$$
\lim_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\mathcal{R}_{1,1}^n > r\right) = -\infty.
$$

Doing the same things with $-\mathcal{R}_{1,1}^n$, we obtain (3.24) for $\mathcal{R}_{1,1}^n$.

Now we shall prove (3.24) with $\mathcal{R}_{2,1}^n$. We have

$$
\mathcal{R}_{2,1}^n = \sum_{k=1}^n E_{-,k}^n E_{+,k}^n + \sum_{k=1}^n E_{-,k}^n B_{+,k}^n + \sum_{k=1}^n E_{+,k}^n B_{-,k}^n + \sum_{k=1}^n B_{-,k}^n B_{+,k}^n,
$$

where

$$
E_{\pm,k}^n := \int_{t_{k-1}^n}^{t_k^n} (\sigma_{2,s}\rho_s \pm \sigma_{2,\frac{k-1}{n}}\rho_{\frac{k-1}{n}})\mathrm{d}B_{1,s},
$$

and

$$B^n_{\pm,k} := \int_{t^n_{k-1}}^{t^n_k} (\sigma_{2,s}\sqrt{1-\rho_s^2} \pm \sigma_{2,\frac{k-1}{n}}\sqrt{1-\rho^2_{\frac{k-1}{n}}})\mathrm{d}B_{3,s}.$$

We know that $(E^n_{-,k}, E^n_{+k}),(E^n_{-,k}, B^n_{+,k}),(E^n_{+,k}, B^n_{-,k}),(B^n_{-,k}, B^n_{+,k})$, $k = 1\cdots n$ are four independent centered Gaussian random vectors with covariances respectively given by $\mho^n_{k,\rho_s}$, $\overline{\mho}^n_{k,-}$, $\overline{\mho}^n_{k,+}$, $\mho^n_{k,\sqrt{1-\rho_s^2}}$ where

$$\mho^n_{k,g_s} := \begin{pmatrix} \int_{t^n_{k-1}}^{t^n_k}(\sigma_{2,s}g_s - \sigma_{2,\frac{k-1}{n}}g_{\frac{k-1}{n}})^2\mathrm{d}s & \int_{t^n_{k-1}}^{t^n_k}(\sigma^2_{2,s}g_s^2 - \sigma^2_{2,\frac{k-1}{n}}g^2_{\frac{k-1}{n}})\mathrm{d}s \\ \int_{t^n_{k-1}}^{t^n_k}(\sigma^2_{2,s}g_s^2 - \sigma^2_{2,\frac{k-1}{n}}g^2_{\frac{k-1}{n}})\mathrm{d}s & \int_{t^n_{k-1}}^{t^n_k}(\sigma_{2,s}g_s + \sigma_{2,\frac{k-1}{n}}g_{\frac{k-1}{n}})^2\mathrm{d}s \end{pmatrix}$$

and

$$\overline{\mho}^n_{k,\mp} := \begin{pmatrix} \int_{t^n_{k-1}}^{t^n_k}(\sigma_{2,s}\rho_s \mp \sigma_{2,\frac{k-1}{n}}\rho_{\frac{k-1}{n}})^2\mathrm{d}s & 0 \\ 0 & \int_{t^n_{k-1}}^{t^n_k}(\sigma_{2,s}\sqrt{1-\rho_s^2} \pm \sigma_{2,\frac{k-1}{n}}\sqrt{1-\rho^2_{\frac{k-1}{n}}})^2\mathrm{d}s \end{pmatrix}.$$

So (3.24) for $\mathcal{R}^n_{2,1}$ is deduced if

$$\sum_{k=1}^n E^n_{-k}E^n_{+,k} \overset{\text{superexp}}{\underset{n}{\longrightarrow}} 0, \qquad \sum_{k=1}^n E^n_{-,k}B^n_{+,k} \overset{\text{superexp}}{\underset{n}{\longrightarrow}} 0,$$

$$\sum_{k=1}^n E^n_{+,k}B^n_{-,k} \overset{\text{superexp}}{\underset{n}{\longrightarrow}} 0, \qquad \sum_{k=1}^n B^n_{-,k}B^n_{+,k} \overset{\text{superexp}}{\underset{n}{\longrightarrow}} 0.$$

Each convergence is deduced by the same calculations as for $(R^n_{-,k}, R^n_{+,k})_{k=1\cdots n}$.

Now we shall prove (3.24) with $\mathcal{R}^n_{3,1}$. We have

$$\begin{aligned}\mathcal{R}^n_{3,1} \quad = \quad & \sum_{k=1}^n R^n_{-,k}E^n_{-,k} + \sum_{k=1}^n R^n_{-,k}B^n_{-,k} + \sum_{k=1}^n R^n_{-,k}\left(\int_{t^n_{k-1}}^{t^n_k} \sigma_{2,\frac{k-1}{n}}\rho_{\frac{k-1}{n}}\mathrm{d}B_{1,s}\right) \\ & + \sum_{k=1}^n R^n_{-,k}\left(\int_{t^n_{k-1}}^{t^n_k} \sigma_{2,\frac{k-1}{n}}\sqrt{1-\rho^2_{\frac{k-1}{n}}}\mathrm{d}B_{3,s}\right) + \sum_{k=1}^n E^n_{-,k}\left(\int_{t^n_{k-1}}^{t^n_k} \sigma_{1,\frac{k-1}{n}}\mathrm{d}B_{1,s}\right) \\ & + \sum_{k=1}^n B^n_{-,k}\left(\int_{t^n_{k-1}}^{t^n_k} \sigma_{1,\frac{k-1}{n}}\mathrm{d}B_{1,s}\right),\end{aligned}$$

where we have used the same notation as before. By the same calculations used to prove (3.24) for $\mathcal{R}^n_{1,1}$ and $\mathcal{R}^n_{2,1}$, we obtain (3.24) for $\mathcal{R}^n_{3,1}$.

Then $V^n_1(X^0)$ and W^n_1 satisfy the same LDP.

<u>*Part 3*</u>. We will prove that $V_1^n(X)$ and $V_1^n(X^0)$ satisfy the same LDP. We need to prove that

$$\left\|V_1^n(X) - V_1^n(X^0)\right\| \quad \overset{\text{superexp}}{\underset{n}{\longrightarrow}} 0.$$

This will be done element by element : for $\ell = 1, 2$

$$Q_{\ell,1}^n(X) - Q_{\ell,1}^n(X^0) \quad \overset{\text{superexp}}{\underset{n}{\longrightarrow}} 0 \quad \text{and} \quad C_1^n(X) - C_1^n(X^0) \quad \overset{\text{superexp}}{\underset{n}{\longrightarrow}} 0. \tag{3.31}$$

We have

$$\left|Q_{\ell,1}^n(X) - Q_{\ell,1}^n(X^0)\right| \leq \varepsilon(n) Q_{\ell,1}^n(X^0) + \left(1 + \frac{1}{\varepsilon(n)}\right) Z_\ell^n,$$

and

$$\left|C_1^n(X) - C_1^n(X^0)\right| \leq 2\varepsilon(n) \max_{\ell=1}^{2} Q_{\ell,1}^n(X^0) + \left(1 + \frac{2}{\varepsilon(n)}\right) \max_{\ell=1}^{2} Z_\ell^n.$$

with

$$Z_\ell^n = \sum_{k=1}^n \left(\int_{t_{k-1}^n}^{t_k^n} b_\ell(t, X_t) \mathrm{d}t\right)^2 \leq \frac{\|b_\ell\|_\infty^2}{n}.$$

We chose $\varepsilon(n)$ such that $n\varepsilon(n) \to \infty$, so (3.31) follows from the LDP of $Q_{\ell,1}(X^0)$ and the estimations above.

Proof of Corollary 3.3.2

From Theorem 3.3.1, we obtain that $V_1^n(X^0)$ satisfies the LDP on $\mathbb{R}^3$ with speed n and with good rate function given by for all $x \in \mathbb{R}^3$

$$I_{ldp}^V(x) = \sup_{\lambda \in \mathbb{R}^3} \Big(\langle \lambda, x\rangle - P_\rho(\sigma_1^2\lambda_1, \sigma_2^2\lambda_2, \sigma_1\sigma_2\lambda_3)\Big) = P_\rho^*\left(\frac{\lambda_1}{\sigma_1^2}, \frac{\lambda_2}{\sigma_2^2}, \frac{\lambda_3}{\sigma_1\sigma_2}\right).$$

where P_ρ and P_ρ^* are given in (3.4) and (3.8) respectively. So we get the expression of I_{ldp}^V given in (3.9).

The Legendre transformation of P_c is defined by

$$P_\rho^*(x) := \sup_{\lambda \in \mathbb{R}^3} \left(\langle \lambda, x\rangle - P_\rho(\lambda)\right).$$

The function $\lambda \to \langle \lambda, x\rangle - P_\rho(\lambda)$ reaches the supremum at the point

$\lambda^* = (\lambda_1^*, \lambda_2^*, \lambda_3^*)$ such as

$$\begin{cases} \lambda_1^* = \frac{1}{2}\dfrac{x_1x_2 - (1-\rho^2)x_2 - x_3^2}{2(1-\rho^2)(x_1x_2 - x_3^2)}, \\ \\ \lambda_2^* = \frac{1}{2}\dfrac{x_1x_2 - (1-\rho^2)x_1 - x_3^2}{2(1-\rho^2)(x_1x_2 - x_3^2)}, \\ \\ \lambda_3^* = \dfrac{x_3^2\rho - x_1x_2\rho + (1-\rho^2)x_3}{(1-\rho^2)(x_1x_2 - x_3^2)}. \end{cases}$$

So we get the expression of the Legendre transformation P_ρ^* given in (3.8).

Proof of Theorem 3.3.3

Now we shall prove the Theorem 3.3.3 in two steps.
<u>Step 1</u>. We start by proving that the LDP holds for

$$W_t^n = \frac{1}{n}\sum_{k1}^{[nt]} F\left(\frac{k-1}{n}\right) Z_k,$$

where F is given in (3.22) and Z_k is given in (3.23).

This result come from an application of LDP of [106, Theorem 4.3]. We will apply it with the random variables $Z_k^n := F\left(\dfrac{k-1}{n}\right) Z_k$. The law of Z_k^n depends on the position $x_i^n := i/n$. This type of model was partially examined by Najim see [106, section 2.4.2]. We just need to verify that if x_i^n and x_j^n are close then so are $\mathcal{L}(Z_i^n)$ and $\mathcal{L}(Z_j^n)$ for the following Wasserstein type distance between probability measures :

$$d_{OW}(P, Q) = \inf_{\eta \in M(P,Q)} \inf\left\{a > 0; \int_{\mathbb{R}^3 \times \mathbb{R}^3} \tau\left(\frac{z - z'}{a}\right)\eta(dzdz') \leqslant 1\right\},$$

where η is a probability with given marginals P and Q and $\tau(z) = e^{|z|} - 1$.

In fact, consider the random variables $Y = F(x)Z$ and $\tilde{Y} = F(x')Z$, since F is continuous

$$\mathbb{E}\tau\left(\frac{Y - \tilde{Y}}{\epsilon}\right) = \mathbb{E}\tau\left(\frac{(F(x) - F(x'))Z}{\epsilon}\right) \leq 1$$

for x' close to x. Thus $d_{OW}(\mathcal{L}(Z_i^n), \mathcal{L}(Z_j^n)) \leq \epsilon$. So we deduce that the sequence $W_\cdot^n$ satisfies the LDP on $\mathcal{B}V$ with speed n and with the rate function J_{ldp} given by

$$J_{ldp}(f) = \int_0^1 \Lambda_t^*(f_a^{'}(t))\mathrm{d}t + \int_0^1 \hbar_t(f_s^{'}(t))\mathrm{d}\theta(t).$$

where for all $z \in \mathbb{R}^3$ and all $t \in [0,1]$

$$\Lambda_t^*(z) = \sup_{\lambda \in \mathbb{R}^3} \left(\langle \lambda, z \rangle - \Lambda_t(\lambda) \right),$$

with

$$\Lambda_t(\lambda) = \log \int_{\mathbb{R}^3} e^{\langle \lambda, z \rangle} P(t, dz) = P_{\rho_t}(\lambda_1 \sigma_{1,t}^2, \lambda_2 \sigma_{2,t}^2, \lambda_3 \sigma_{1,t}\sigma_{2,t}),$$

so Λ_t^* coincide with $P_{\rho_t}^*$ given in Theorem 3.3.3.

And θ is any real-valued nonnegative measure with respect to which μ_s^f is absolutely continuous and $f_s' = d\mu_s^f / d\theta$ and for all $z \in \mathbb{R}^3$ and all $t \in [0,1]$ the recession function $\hbar_t$ of Λ_t^* defined by $\hbar_t(z) = \sup\{\langle \lambda, z \rangle, \lambda \in \mathcal{D}_{\Lambda_t}\}$ with $\mathcal{D}_{\Lambda_t} = \{\lambda \in \mathbb{R}^3, \Lambda_t(\lambda) < \infty\} = \{\lambda \in \mathbb{R}^3, P_{\rho_t}(\lambda_1 \sigma_{1,t}^2, \lambda_2 \sigma_{2,t}^2, \lambda_3 \sigma_{1,t}\sigma_{2,t}), < \infty\}$.

The recession function α of P_c^* , see [120, Theorem 13.3] is given by

$$\alpha(z) := \lim_{h \to \infty} \frac{P_c^*(hz)}{h} = \frac{z_1 + z_2 - 2cz_3}{2(1-c^2)} 1_{[z_1 > 0, z_2 > 0, z_3^2 < z_1 z_2]}.$$

Using this expression, we obtain the rate function given in the Theorem 3.3.3.

Step 2. Now we have to prove that

$$\sup_{t \in [0,1]} \left\| V_t^n(X^0) - W_t^n \right\| \quad \overset{\text{superexp}}{\underset{n}{\longrightarrow}} 0.$$

To do that, we have to prove that for $q = 1, 2, 3$

$$\sup_{t \in [0,1]} \left| \mathcal{R}_{q,t}^n \right| \quad \overset{\text{superexp}}{\underset{n}{\longrightarrow}} 0, \tag{3.32}$$

where the definition of $\mathcal{R}_{q,t}^n$ arge given in (3.25), (3.26) and (3.27) for $q = 1, 2, 3$ respectively.

We start by proving (3.32) for $q = 1$, the other terms for $q = 2, 3$ follow the same line of proof.

We remark that $(\mathcal{R}_{1,t}^n - \mathbb{E}(\mathcal{R}_{1,t}^n))$ is a $(\mathcal{F}_{[nt]/n})$-martingale. Then

$$\exp(\lambda n \left[\mathcal{R}_{1,t}^n - \mathbb{E}(\mathcal{R}_{1,t}^n) \right])$$

is a sub-martingale. By the maximal inequality, we have for any $r, \lambda > 0$,

$$\begin{aligned}
\mathbb{P}\left(\sup_{t \in [0,1]} \left[\mathcal{R}_{1,t}^n - \mathbb{E}(\mathcal{R}_{1,t}^n) \right] > r \right) &= \mathbb{P}\left(\exp\left(\lambda n \sup_{t \in [0,1]} \left[\mathcal{R}_{1,t}^n - \mathbb{E}(\mathcal{R}_{1,t}^n) \right] \right) > e^{n\lambda r} \right) \\
&\leq e^{-n\lambda r} \mathbb{E}\left(\exp\left(\lambda n \left[\mathcal{R}_{1,1}^n - \mathbb{E}(\mathcal{R}_{1,1}^n) \right] \right) \right)
\end{aligned}$$

and similarly

$$\mathbb{P}\left(\inf_{t\in[0,1]}\left[\mathcal{R}^n_{q,t}-\mathbb{E}(\mathcal{R}^n_{1,t})\right]<-r\right)\leq e^{-n\lambda r}\mathbb{E}\left(\exp\left(-\lambda n\left[\mathcal{R}^n_{1,1}-\mathbb{E}(\mathcal{R}^n_{1,1})\right]\right)\right).$$

So we get

$$\frac{1}{n}\log\mathbb{P}\left(\sup_{t\in[0,1]}\left[\mathcal{R}^n_{1,t}-\mathbb{E}(\mathcal{R}^n_{1,t})\right]>r\right)\leq-\lambda r+\frac{1}{n}\log\mathbb{E}\left(e^{\lambda n\mathcal{R}^n_{1,1}}\right)-\lambda\mathbb{E}(\mathcal{R}^n_{1,1}).$$

It is easy to see that $\mathbb{E}(\mathcal{R}^n_{1,1})\to 0$ as n goes to infinity. We have already seen in (3.30) that $\frac{1}{n}\log\mathbb{E}\left(e^{\lambda n\mathcal{R}^n_{1,1}}\right)\to 0$ as n gos to infinity. So we obtain for all $\lambda>0$

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\sup_{t\in[0,1]}\left[\mathcal{R}^n_{1,t}-\mathbb{E}(\mathcal{R}^n_{1,t})\right]>r\right)\leq-\lambda r.$$

Letting $\lambda>0$ goes to infinity, we obtain that the left term in the last inequality goes to $-\infty$.

And similarly, by doing the same calculations with

$$\mathbb{P}\left(\inf_{t\in[0,1]}\left[\mathcal{R}^n_{q,t}-\mathbb{E}(\mathcal{R}^n_{1,t})\right]<-r\right),$$

we obtain that

$$\sup_{t\in[0,1]}\left|\mathcal{R}^n_{1,t}-\mathbb{E}(\mathcal{R}^n_{1,t})\right|\quad\overset{\text{superexp}}{\underset{n}{\longrightarrow}}0.$$

Since $\mathbb{E}(\mathcal{R}^n_{1,1})\quad\overset{\text{superexp}}{\underset{n}{\longrightarrow}}0$, we obtain (3.32) for $q=1$.

Proof of Theorem 3.3.4

As is usual, the proof of the MDP is somewhat simpler than the LDP, relying on the same line of proof than the one for the CLT. Namely, a good control of the asymptotic of the moment generating functions, and Gärtner-Ellis theorem. We treat here only the case $b_\ell=0$, and one can prove easily that $V^n_\cdot(X)$ and $V^n_\cdot(X^0)$ satisfy the same MDP by following the proof of Theorem 3.3.6.

We shall then prove that for all $\lambda\in\mathbb{R}^3$

$$\lim_{n\to\infty}\frac{1}{v_n^2}\log\mathbb{E}\exp\left(v_n^2\frac{\sqrt{n}}{v_n}\left\langle\lambda,V^n_1-[V]_1\right\rangle\right)=\frac{1}{2}\left\langle\lambda,\Sigma_1\cdot\lambda\right\rangle. \qquad (3.33)$$

Taking the calculation in (3.5.2),we have

$$\frac{1}{v_n^2}\log\mathbb{E}\exp\left(v_n^2\frac{\sqrt{n}}{v_n}\langle\lambda, V_1^n-[V]_1\rangle\right)=\frac{1}{v_n^2}\sum_{k=1}^{n}\left[H_k^n(\lambda)-v_n\sqrt{n}\langle\lambda,[V]_1\rangle\right],$$

with

$$H_k^n(\lambda):=P_{c_k^n}\left(\lambda_1 v_n\sqrt{n}a_{1,k},\lambda_2 v_n\sqrt{n}a_{2,k},\lambda_3 v_n\sqrt{n}\sqrt{a_{1,k}}\sqrt{a_{2,k}}\right),$$

where $a_{\ell,k}$ are given in (3.19) and c_k^n is given in (3.20).

By our condition (3.3),

$$\varepsilon(n):=\sqrt{n}v_n\max_{1\leqslant k\leqslant n}\max_{\ell=1,2}a_{\ell,k}\xrightarrow[n\to\infty]{}0.$$

By multidimensional Taylor formula and noting that $P_{c_k^n}(0,0,0)=0$, $\nabla P_{c_k^n}(0,0,0)=(1,1,c_k^n)^T$ and the Hessian matrix

$$H(P_{c_k^n})(0,0,0)=\begin{pmatrix}2 & 2(c_k^n)^2 & 2c_k^n\\ 2(c_k^n)^2 & 2 & 2c_k^n\\ 2c_k^n & 2c_k^n & 1+(c_k^n)^2\end{pmatrix},$$

and after an easy calculations, we obtain once if $||\lambda||\cdot|\varepsilon(n)|<\frac{1}{4}$, i.e. for n large enough,

$$H_k^n(\lambda)=v_n\sqrt{n}\langle\lambda,[V]_1\rangle+nv_n^2\frac{1}{2}\langle\lambda,\Sigma_k^n\cdot\lambda\rangle+nv_n^2\nu(k,n),$$

where

$$\Sigma_k^n:=\begin{pmatrix}a_{1,k}^2 & (\vartheta_k^n)^2 & a_{1,k}\vartheta_k^n\\ (\vartheta_k^n)^2 & a_{2,k}^2 & a_{2,k}\vartheta_k^n\\ a_{1,k}\vartheta_k^n & a_{2,k}\vartheta_k^n & \frac{1}{2}(a_{1,k}a_{2,k}+(\vartheta_k^n)^2)\end{pmatrix},$$

where ϑ_k^n is given in (3.20), and $\nu(k,n)$ satisfies

$$|\nu(k,n)|\leqslant C||\lambda||\cdot|\varepsilon(n)|,$$

where $C=\frac{1}{6}\sup_{||\lambda||\leq 1/4}\left|\frac{\partial^3 P(\lambda_1,\lambda_2,\lambda_3)}{\partial^3\lambda}\right|$.

On the other hand, by the classical Lebesgue derivation theorem see [56], we have

$$\sum_{k=1}^{n}\left(\frac{\int_{t_{k-1}^n}^{t_k^n}g(s)\mathrm{d}s}{t_k^n-t_{k-1}^n}\right)\left(\frac{\int_{t_{k-1}^n}^{t_k^n}h(s)\mathrm{d}s}{t_k^n-t_{k-1}^n}\right)(t_k^n-t_{k-1}^n)\to\int_0^1 g(s)h(s)\mathrm{d}s$$

by taking different chosse of g and h : once $g(s) = h(s) = \sigma^2_{\ell,s}$ or $g(s) = h(s) = \sigma_{1,s}\sigma_{1,s}\rho_s$, or $g(s) = \sigma^2_{\ell,s}$ and $h(s) = \sigma^2_{\ell',s}$ $\ell \neq \ell'$, and $g(s) = \sigma^2_{\ell,s}$ and $h(s) = \sigma_{1,s}\sigma_{1,s}\rho_s$, we obtain that

$$n\sum_{k=1}^{n}\Sigma_k^n \to_{n\to\infty} \Sigma_1 = \begin{pmatrix} \int_0^1 \sigma^4_{1,t}\mathrm{d}t & \int_0^1 \sigma^2_{1,t}\sigma^2_{2,t}\rho^2_t\mathrm{d}t & \int_0^1 \sigma^3_{1,t}\sigma_{2,t}\rho_t\mathrm{d}t \\ \int_0^1 \sigma^2_{1,t}\sigma^2_{2,t}\rho^2_t\mathrm{d}t & \int_0^1 \sigma^4_{2,t}\mathrm{d}t & \int_0^1 \sigma_{1,t}\sigma^3_{2,t}\rho_t\mathrm{d}t \\ \int_0^1 \sigma^3_{1,t}\sigma_{2,t}\rho_t\mathrm{d}t & \int_0^1 \sigma_{1,t}\sigma^3_{2,t}\rho_t\mathrm{d}t & \int_0^1 \frac{1}{2}\sigma^2_{1,t}\sigma^2_{2,t}(1+\rho^2_t)\mathrm{d}t \end{pmatrix} = \Sigma_1.$$

Then the (3.33) follows.

Hence (3.3.4) follows from the Gärtner-Ellis theorem.

Proof of Theorem 3.3.5

It is well known that the LDP of finite dimensional vector

$$\left(\frac{\sqrt{n}}{v_n}(V^n_{s_1}(X) - [V]_{s_1}, \cdots, V^n_{s_k}(X) - [V]_{s_k})\right), 0 < s_1 < \cdots < s_k \leqslant 1, k \geqslant 1$$

and the following exponential tightness : for any $s \in [0,1]$ and $\eta > 0$

$$\lim_{\varepsilon\downarrow 0}\limsup_{n\to\infty}\frac{1}{v_n^2}\log\mathbb{P}\left(\frac{\sqrt{n}}{v_n}\sup_{s\leqslant t\leqslant s+\varepsilon}\|\Delta^t_s V^n_\cdot(X) - \Delta^t_s[V]_\cdot\| > \eta\right) = -\infty$$

are sufficient for the LDP of $\frac{\sqrt{n}}{v_n}(V^n_\cdot(X) - [V]_\cdot)$ for the sup-norm topology (cf. [53],[56]).

Under the assumption of Theorem 3.3.5, we have :

$$\lim_{n\to\infty}\sqrt{n}\sup_{t\in[0,1]}\|\mathbb{E}V^n_t(X) - [V]_t\| = 0. \tag{3.34}$$

In fact, we have :

$$\begin{aligned}
&\sqrt{n}\sup_{t\in[0,1]}\|\mathbb{E}V^n_t(X) - [V]_t\| \\
&\leqslant 3\sqrt{n}\max\left(\max_{\ell=1,2}\sup_{t\in[0,1]}|\mathbb{E}Q^n_{\ell,t}(X) - [X_\ell]_t|, \sqrt{n}\sup_{t\in[0,1]}|\mathbb{E}C^n_t(X) - \langle X_1, X_2\rangle_t|\right) \\
&\leqslant 3\max\left(\max_{\ell=1,2}\max_{k\leqslant n}\sqrt{\int_{(k-1)/n}^{k/n}\sigma^4_{\ell,t}\mathrm{d}t}, \max_{k\leqslant n}\sqrt{\int_{(k-1)/n}^{k/n}\sigma^2_{1,t}\sigma^2_{2,t}\rho^2_t\mathrm{d}t}\right).
\end{aligned}$$

By our condition (3.3), we obtain (3.34).

Now, we show that for any partition $0 < s_1 < \cdots < s_k \leqslant 1, k \geqslant 1$ of $[0,1]$

$$\frac{\sqrt{n}}{v_n}\Big(V^n_{s_1}(X) - [V]_{s_1}, \cdots, V^n_{s_k}(X) - [V]_{s_k}\Big)$$

satisfies the LDP on $\mathbb{R}^k$ with speed v_n^2 and with rate function given by

$$I_{s_1,\cdots,s_k}(x_1,\cdots,x_k) = \frac{1}{2}\sum_{i=1}^{k}\langle(x_i - x_{i-1}), (\Sigma^{s_i}_{s_{i-1}})^{-1}\cdot(x_i - x_{i-1})\rangle \tag{3.35}$$

where

$$\Sigma^u_s = \begin{pmatrix} \int_s^u \sigma^4_{1,t}\mathrm{d}t & \int_s^u \sigma^2_{1,t}\sigma^2_{2,t}\rho^2_t\mathrm{d}t & \int_s^u \sigma^3_{1,t}\sigma_{2,t}\rho_t\mathrm{d}t \\ \int_s^u \sigma^2_{1,t}\sigma^2_{2,t}\rho^2_t\mathrm{d}t & \int_s^u \sigma^4_{2,t}\mathrm{d}t & \int_s^u \sigma_{1,t}\sigma^3_{2,t}\rho_t\mathrm{d}t \\ \int_s^u \sigma^3_{1,t}\sigma_{2,t}\rho_t\mathrm{d}t & \int_s^u \sigma_{1,t}\sigma^3_{2,t}\rho_t\mathrm{d}t & \int_s^u \frac{1}{2}\sigma^2_{1,t}\sigma^2_{2,t}(1+\rho^2_t)\mathrm{d}t \end{pmatrix}$$

is invertible and

$$\begin{aligned}\det(\Sigma^u_s) = &\Big(\int_s^u \sigma^4_{1,t}\mathrm{d}t\Big)\Big(\int_s^u \sigma^4_{2,t}\mathrm{d}t\Big)\Big(\int_s^u \frac{1}{2}\sigma^2_{1,t}\sigma^2_{2,t}(1+\rho^2_t)\mathrm{d}t\Big) \\ &+ 2\Big(\int_s^u \sigma^2_{1,t}\sigma^2_{2,t}\rho^2_t\mathrm{d}t\Big)\Big(\int_s^u \sigma_{1,t}\sigma^3_{2,t}\rho_t\mathrm{d}t\Big)\Big(\int_s^u \sigma^3_{1,t}\sigma_{2,t}\rho_t\mathrm{d}t\Big) \\ &- \Big(\int_s^u \frac{1}{2}\sigma^2_{1,t}\sigma^2_{2,t}(1+\rho^2_t)\mathrm{d}t\Big)\Big(\int_s^u \sigma^2_{1,t}\sigma^2_{2,t}\rho^2_t\mathrm{d}t\Big)^2 \\ &- \Big(\int_s^u \sigma^4_{1,t}\mathrm{d}t\Big)\Big(\int_s^u \sigma_{1,t}\sigma^3_{2,t}\rho_t\mathrm{d}t\Big)^2 - \Big(\int_s^u \sigma^4_{2,t}\mathrm{d}t\Big)\Big(\int_s^u \sigma^3_{1,t}\sigma_{2,t}\rho_t\mathrm{d}t\Big)^2\end{aligned}$$

and $(\Sigma^u_s)^{-1}$ his inverse.

For n large enough we have $1 < [nt_1] < \cdots < [nt_k] < n$, so by applying the homeomorphism

$$\Upsilon : (x_1,\cdots,x_k) \to (x_1, x_2 - x_1, \cdots, x_k - x_{k-1})$$

$Z_n = (V^n_{s_1}(X) - [V]_{s_1}, \cdots, V^n_{s_k}(X) - [V]_{s_k})$ can be mapped to $U_n = \Upsilon Z_n$ with independent components.

Then we consider the LDP of $\dfrac{\sqrt{n}}{v_n}(U_n - \mathbb{E}U_n)$. For any $\theta = (\theta_1,\cdots,\theta_k) \in (\mathbb{R}^3)^k$,

$$\Lambda s_1,\cdots,s_k(\theta) = \lim_{n\to\infty}\frac{1}{v_n^2}\log\mathbb{E}\exp\left(v_n\sqrt{n}\,\langle\lambda, U_n - EU_n\rangle\right) = \sum_{i=1}^{k}\frac{1}{2}\langle\lambda_i, \Sigma^{s_i}_{s_{i-1}}\cdot\lambda_i\rangle.$$

By Gärtner-Ellis theorem, $\frac{\sqrt{n}}{v_n}(U_n - \mathbb{E}U_n)$ satisfies the LDP in $(\mathbb{R}^3)^k$ with speed v_n^2 and with the good rate function

$$\Lambda^* s_1, \cdots, s_k(x) = \frac{1}{2}\sum_{i=1}^{k}\langle x_i, (\Sigma_{s_{i-1}}^{s_i})^{-1} \cdot x_i\rangle.$$

Then by the inverse contraction principle, we have $\frac{\sqrt{n}}{v_n}(Z_n - \mathbb{E}Z_n)$ satisfies the LDP with speed v_n^2 and with the rate function $I_{s_1,\cdots,s_k}(x)$ given in (3.35).

Now, we shall prove that for any $\eta > 0$, $s \in [0,1]$

$$\lim_{\varepsilon\downarrow 0}\limsup_{n\to\infty}\frac{1}{v_n^2}\log\mathbb{P}\left(\frac{\sqrt{n}}{v_n}\sup_{s\leqslant t\leqslant s+\varepsilon}\|\Delta_s^t V_\cdot^n(X) - \mathbb{E}\Delta_s^t V_\cdot^n(X)\| > \eta\right) = -\infty. \tag{3.36}$$

For that we need to prove that for $\ell = 1, 2$ and for all $\eta > 0$ and $s \in [0,1]$

$$\lim_{\varepsilon\downarrow 0}\limsup_{n\to\infty}\frac{1}{v_n^2}\log\mathbb{P}\left(\frac{\sqrt{n}}{v_n}\sup_{s\leqslant t\leqslant s+\varepsilon}|\Delta_s^t Q_{\ell,\cdot}(X) - \mathbb{E}\Delta_s^t Q_{\ell,\cdot}(X)| > \eta\right) = -\infty, \tag{3.37}$$

and

$$\lim_{\varepsilon\downarrow 0}\limsup_{n\to\infty}\frac{1}{v_n^2}\log\mathbb{P}\left(\frac{\sqrt{n}}{v_n}\sup_{s\leqslant t\leqslant s+\varepsilon}|\Delta_s^t C_\cdot(X) - \mathbb{E}\Delta_s^t C_\cdot(X)| > \eta\right) = -\infty. \tag{3.38}$$

In fact (3.37) can be done in the same way than in Djellout et al.[56]. It remains to show (3.38). This will be done following the same technique as for the proof of (3.37) and using a result of [61].

By (3.35) and (3.36), $\frac{\sqrt{n}}{v_n}(V_\cdot^n - [V]_\cdot)$ satifies the LDP with the speed b_n^2 and good rate function

$$I_{sup}(x) = \sup\left\{I_{s_1,\cdots,s_k}(x(s_1),\cdots,x(s_k)); 0 < s_1 < \cdots < s_k \leqslant 1, k \geqslant 1\right\},$$

where

$$I_{s_1,\cdots,s_k}(x(s_1),\cdots,x(s_k)) = \frac{1}{2}\sum_{i=1}^{k}\langle x(s_i)-x(s_{i-1}), (\Sigma_{s_{i-1}}^{s_i})^{-1}\cdot(x(s_i)-x(s_{i-1}))\rangle$$

It remains to prove that $I_{sup}(x) = I_{mdp}(x)$. We shall prove that $I_{sup}(x) \leqslant I_{mdp}(x)$.

For this, we treat the first element of the matrix $(\Sigma_{s_{i-1}}^{s_i})^{-1}$ which is denoted $(\Sigma_{s_{i-1}}^{s_i})_{1,1}^{-1}$ and we prove that

$$\sum_{i=1}^{k}(x_1(s_i)-x_1(s_{i-1}))^2.(\Sigma_{s_{i-1}}^{s_i})_{1,1}^{-1} \leqslant \int_0^1 (x_1'(t))^2.(\Sigma_t^{-1})_{1,1}\mathrm{d}t,$$

where $(\Sigma_t^{-1})_{1,1}$ represente the first element of the matrix Σ_t^{-1}. We have

$$(\Sigma_{s_{i-1}}^{s_i})_{1,1}^{-1} = \frac{1}{\det(\Sigma_{s_{i-1}}^{s_i})}\left(\left(\int_{s_{i-1}}^{s_i}\sigma_{2,t}^4\mathrm{d}t\right)\left(\int_{s_{i-1}}^{s_i}\frac{1}{2}\sigma_{1,t}^2\sigma_{2,t}^2(1+\rho_t^2)\mathrm{d}t\right)-\left(\int_{s_{i-1}}^{s_i}\sigma_{1,t}\sigma_{2,t}^3\rho_t\mathrm{d}t\right)^2\right).$$

By [84, p.1305], for $x := (x_1, x_2, x_3) \in \mathcal{H}$, if $I_{sup}(x) < +\infty$, then for $0 < s_1 < \cdots < s_k \leqslant 1$,

Then

$$\begin{aligned}\sum_{i=1}^{k}(x_1(s_i)-x_1(s_{i-1}))^2.(\Sigma_{s_{i-1}}^{s_i})_{1,1}^{-1} &\leqslant \int_0^1 (x_1'(t))^2.\frac{\frac{1}{2}\sigma_{1,t}^2\sigma_{2,t}^6(1+\rho_t^2)-\sigma_{1,t}^2\sigma_{2,t}^6\rho_t^2}{\det(\Sigma_t)}\mathrm{d}t\\ &= \int_0^1 (x_1'(t))^2.(\Sigma_t^{-1})_{1,1}\mathrm{d}t,\end{aligned}$$

The same calculation with the other terms of the matrix given in the following, implies that $I_{sup}(x) \leqslant I_{mdp}(x)$:

$$(\Sigma_{s_{i-1}}^{s_i})_{2,2}^{-1} = \frac{1}{\det(\Sigma_{s_{i-1}}^{s_i})}\left[\left(\int_{s_{i-1}}^{s_i}\sigma_{1,t}^4\mathrm{d}t\right)\left(\int_{s_{i-1}}^{s_i}\frac{1}{2}\sigma_{1,t}^2\sigma_{2,t}^2(1+\rho_t^2)\mathrm{d}t\right)-\left(\int_{s_{i-1}}^{s_i}\sigma_{1,t}^3\sigma_{2,t}\rho_t\mathrm{d}t\right)^2\right],$$

$$(\Sigma_{s_{i-1}}^{s_i})_{3,3}^{-1} = \frac{1}{\det(\Sigma_{s_{i-1}}^{s_i})}\left[\left(\int_{s_{i-1}}^{s_i}\sigma_{1,t}^4\mathrm{d}t\right)\left(\int_{s_{i-1}}^{s_i}\sigma_{2,t}^4\mathrm{d}t\right)-\left(\int_{s_{i-1}}^{s_i}\sigma_{1,t}^2\sigma_{2,t}^2\rho_t^2\mathrm{d}t\right)^2\right],$$

$$\begin{aligned}(\Sigma_{s_{i-1}}^{s_i})_{1,2}^{-1} = (\Sigma_{s_{i-1}}^{s_i})_{2,1}^{-1} = \frac{1}{\det(\Sigma_{s_{i-1}}^{s_i})}&\left[\left(\int_{s_{i-1}}^{s_i}\sigma_{1,t}^3\sigma_{2,t}\rho_t\mathrm{d}t\right)\left(\int_{s_{i-1}}^{s_i}\sigma_{1,t}\sigma_{2,t}^3\rho_t\mathrm{d}t\right)\right.\\ &\left.-\left(\int_{s_{i-1}}^{s_i}\sigma_{1,t}^2\sigma_{2,t}^2\rho_t^2\mathrm{d}t\right)\left(\int_{s_{i-1}}^{s_i}\frac{1}{2}\sigma_{1,t}^2\sigma_{2,t}^2(1+\rho_t^2)\mathrm{d}t\right)\right],\end{aligned}$$

$$\begin{aligned}(\Sigma_{s_{i-1}}^{s_i})_{1,3}^{-1} = (\Sigma_{s_{i-1}}^{s_i})_{3,1}^{-1} = \frac{1}{\det(\Sigma_{s_{i-1}}^{s_i})}&\left[\left(\int_{s_{i-1}}^{s_i}\sigma_{1,t}^2\sigma_{2,t}^2\rho_t^2\mathrm{d}t\right)\left(\int_{s_{i-1}}^{s_i}\sigma_{1,t}\sigma_{2,t}^3\rho_t\mathrm{d}t\right)\right.\\ &\left.-\left(\int_{s_{i-1}}^{s_i}\sigma_{2,t}^4\mathrm{d}t\right)\left(\int_{s_{i-1}}^{s_i}\sigma_{1,t}^3\sigma_{2,t}\rho_t\mathrm{d}t\right)\right],\end{aligned}$$

$$(\Sigma_{s_{i-1}}^{s_i})^{-1}_{2,3} = (\Sigma_{s_{i-1}}^{s_i})^{-1}_{3,2} = \frac{1}{\det(\Sigma_{s_{i-1}}^{s_i})}\Big[\Big(\int_{s_{i-1}}^{s_i} \sigma_{1,t}^2\sigma_{2,t}^2\rho_t^2 \mathrm{d}t\Big)\Big(\int_{s_{i-1}}^{s_i} \sigma_{1,t}^3\sigma_{2,t}\rho_t \mathrm{d}t\Big) - \Big(\int_{s_{i-1}}^{s_i} \sigma_{1,t}^4 \mathrm{d}t\Big)\Big(\int_{s_{i-1}}^{s_i} \sigma_{1,t}\sigma_{2,t}^3\rho_t \mathrm{d}t\Big)\Big].$$

On the other hand, by the convergence of martingales and Fatou's lemma,

$$I_{mdp}(x) < +\infty, \qquad \text{and} \qquad I_{mdp}(x) \leqslant I_{sup}(x).$$

So we have $I_{sup}(x) = I_{mdp}(x)$.

Proof of Theorem 3.3.6

Step 1. We shall prove that $\tilde{V}_\cdot^n$ and $V_\cdot^n(X^0)$ satisfy the same LDP, by means of the approximation Lemma in [56]. So we shall prove that $\tilde{Q}_{\ell,\cdot}^n$ and $Q_{\ell,\cdot}^n(X^0)$ satisfy the same LDP and idem for $\tilde{C}_\cdot^n$ and $C_\cdot^n(X^0)$.

We have

$$\sup_{t\in[0,1]} |\tilde{Q}_{\ell,t}^n(X) - Q_{\ell,t}^n(X^0)| \leqslant \varepsilon(n) Q_{\ell,1}^n(X^0) + \left(1 + \frac{1}{\varepsilon(n)}\right)\overline{Z}_\ell^n \tag{3.39}$$

and

$$\sup_{t\in[0,1]} |\tilde{C}_{1,t}^n(X) - C_{1,t}^n(X^0)| \leqslant 2\varepsilon(n) \max_{\ell=1}^{2} Q_{\ell,1}^n(X^0) + \left(1 + \frac{2}{\varepsilon(n)}\right)\max_{\ell=1}^{2} \overline{Z}_\ell^n, \tag{3.40}$$

where the sequence $\varepsilon(n) > 0$ will be selected later, and $\overline{Z}_\ell^n$ is given

$$\overline{Z}_\ell^n = \sum_{k=1}^n \left(\int_{t_{k-1}^n}^{t_k^n} b_\ell(t, X_t)\mathrm{d}t - b_\ell(t_{k-1}^n, X_{t_{k-1}^n})(t_k^n - t_{k-1}^n)\right)^2, \tag{3.41}$$

with $X_t = (X_{1,t}, X_{2,t})$.

For $Q_{\ell,1}^n(X^0)$, being a Gaussian process, [56, Theorem 1.1] may be used. It remains to control $\overline{Z}_\ell^n$. For this we just need to prove that :

$$\frac{1}{\varepsilon(n)}\overline{Z}_\ell^n \quad \overset{\text{superexp}}{\underset{n}{\longrightarrow}} 0. \tag{3.42}$$

The main idea is to reduce it to estimations of $M_{\ell,t} = \int_0^t \sigma_{\ell,s}\mathrm{d}B_{\ell,s}$, by means of Gronwall's inequality. So, we have at first for all $t \in [0,1]$

$$\begin{aligned}\|X_t\| &\leqslant \|X_0\| + C\int_0^t (1 + (1+\eta(s))\|X_s\|)\mathrm{d}s + \sup_{s\leqslant t}\|M_s\| \\ &\leqslant \left(C + \|X_{1,0}\| + \|X_{2,0}\| + \sup_{s\leqslant 1}\|M_s\|\right) + C_1\int_0^t \|X_s\|\mathrm{d}s.\end{aligned}$$

where $C_1 = C(1+\eta(1))$. Hence, by Gronwall's inequality

$$\|X_t\| \leqslant \left(C + \|X_{1,0}\| + \|X_{2,0}\| + \sup_{s\leqslant 1}\|M_s\| \right) e^{C_1 t}, \qquad \forall t \in [0,1] \tag{3.43}$$

For any $s \in [0,1]$, $v > 0$

$$\begin{aligned}\sup_{s\leqslant t\leqslant s+v}\|X_t - X_s\| &\leqslant \sup_{s\leqslant t\leqslant s+v}\|M_t - M_s\| + v.\sup_{s\leqslant t\leqslant s+v}\|b(t,X_t)\| \\ &\leqslant \sup_{s\leqslant t\leqslant s+v}\|M_t - M_s\| + vC_2\left(\sup_{0\leqslant t\leqslant 1}\|X_t\| + 1\right)\end{aligned} \tag{3.44}$$

We get by (3.13), (3.43),(3.44) and Cauchy-Schwarz's inequality

$$\begin{aligned}&\left(\int_{t_{k-1}^n}^{t_k^n} b_\ell(t,X_t)\mathrm{d}t - b_\ell(t_{k-1}^n, X_{t_{k-1}^n})(t_k^n - t_{k-1}^n)\right)^2 \\ \leqslant\ &\left(\frac{1}{n}C\left(1 + \sup_{t_{k-1}^n\leqslant t\leqslant t_k^n}\|X_t - X_{t_{k-1}^n}\| + 2\eta\left(\frac{1}{n}\right)\sup_{0\leqslant t\leqslant 1}\|X_t\|\right)\right)^2 \\ \leqslant\ &\frac{C_3}{n^2}\left(1 + \sup_{t_{k-1}^n\leqslant t\leqslant t_k^n}\|M_t - M_{t_{k-1}^n}\|^2 + \left(\frac{1}{n^2} + \eta\left(\frac{1}{n}\right)^2\right)\sup_{0\leqslant t\leqslant 1}\|M_t\|^2\right)\end{aligned} \tag{3.45}$$

Chose $\varepsilon(n) > 0$ so that

$$\varepsilon(n) \to 0 \quad \text{but} \quad \frac{\frac{1}{n^2} + \eta\left(\frac{1}{n}\right)^2}{\varepsilon(n)} \to 0 \tag{3.46}$$

By (3.45) and the definition of $Z_{\ell,n}$, we have that

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\frac{1}{\varepsilon(n)}Z_{\ell,n} > \delta\right) \leqslant \max(A,B)$$

where

$$\begin{aligned}A &= \limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\frac{1}{\varepsilon(n)n}\max_{k\leqslant n}\sup_{t_{k-1}^n\leqslant t\leqslant t_k^n}\|M_t - M_{t_{k-1}^n}\|^2 > C_4\delta\right) \\ &\leqslant 2\max_{\ell=1,2}\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\frac{1}{\varepsilon(n)n}\max_{k\leqslant n}\sup_{t_{k-1}^n\leqslant t\leqslant t_k^n}|M_{\ell,t} - M_{\ell,t_{k-1}^n}|^2 > C_4\delta\right)\end{aligned} \tag{3.47}$$

and

$$\begin{aligned}B &= \limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\frac{1}{\varepsilon(n)^2 n}\left(\frac{1}{n^2} + \eta\left(\frac{1}{n}\right)^2\right)\sup_{0\leqslant t\leqslant 1}\|M_t\|^2 > C_5\delta\right) \\ &\leqslant 2\max_{\ell=1,2}\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\frac{1}{\varepsilon(n)n}\left(\frac{1}{n^2} + \eta\left(\frac{1}{n}\right)^2\right)\sup_{0\leqslant t\leqslant 1}|M_{\ell,t}|^2 > C_5\delta\right)\end{aligned} \tag{3.48}$$

By Lévy's inequality for a Brownian motion and our choice (3.46) of $\varepsilon(n)$, the limits (3.47) and (3.48) are both $-\infty$. Limit (3.42) follows.

Step 2. We shall prove that $\frac{\sqrt{n}}{v_n}(\tilde{V}^n_\cdot - [V]_\cdot)$ and $\frac{\sqrt{n}}{v_n}(V^n_\cdot(X^0) - [V]_\cdot)$ satisfy the same LDP, by means of the approximation lemma in [56] and of three strong tools : Gronwall's inequality, Lévy's inequality and an isoperimetric inequality for gaussian processes. By the estimation above (3.39) and (3.40), and as $Q^n_{\ell,t}(X^0)$ was also estimated in the proof of [56, Theorem 1.3]. It remains to control $Z_{\ell,n}$ given in (3.41) . For this we just need to prove that :

$$\frac{1}{\varepsilon(n)}\frac{\sqrt{n}}{v_n}Z_{\ell,n} \quad \xrightarrow[v_n^2]{\text{superexp}} 0. \tag{3.49}$$

Chose $\varepsilon(n) > 0$ so that

$$\frac{\varepsilon(n)\sqrt{n}}{v_n} \to 0 \quad \text{but} \quad \frac{\left(\frac{1}{n^2} + \eta\left(\frac{1}{n}\right)^2\right) v_n}{\varepsilon(n)\sqrt{n}} \to 0. \tag{3.50}$$

By (3.45) and the definition of $Z_{\ell,n}$ given in (3.41), we have that

$$\limsup_{n\to\infty} \frac{1}{v_n^2} \log \mathbb{P}\left(\frac{1}{\varepsilon(n)}\frac{\sqrt{n}}{v_n} Z_{\ell,n} > \delta\right) \leqslant \max(A, B)$$

where

$$\begin{aligned} A &= \limsup_{n\to\infty} \frac{1}{v_n^2} \log \mathbb{P}\left(\frac{1}{\varepsilon(n) v_n \sqrt{n}} \max_{k\leqslant n} \sup_{t^n_{k-1}\leqslant t\leqslant t^n_k} \|M_t - M_{t^n_{k-1}}\|^2 > C_4\delta\right) \\ &\leqslant 2\max_{\ell=1,2} \limsup_{n\to\infty} \frac{1}{v_n^2} \log \mathbb{P}\left(\frac{1}{\varepsilon(n) v_n \sqrt{n}} \max_{k\leqslant n} \sup_{t^n_{k-1}\leqslant t\leqslant t^n_k} |M_{\ell,t} - M_{\ell,t^n_{k-1}}|^2 > C_4\delta\right), \end{aligned} \tag{3.51}$$

and

$$\begin{aligned} B &= \limsup_{n\to\infty} \frac{1}{v_n^2} \log \mathbb{P}\left(\frac{1}{\varepsilon(n) v_n \sqrt{n}} \left(\frac{1}{n^2} + \eta\left(\frac{1}{n}\right)^2\right) \sup_{0\leqslant t\leqslant 1} \|M_t\|^2 > C_5\delta\right) \\ &\leqslant 2\max_{\ell=1,2} \limsup_{n\to\infty} \frac{1}{v_n^2} \log \mathbb{P}\left(\frac{1}{\varepsilon(n) v_n \sqrt{n}} \left(\frac{1}{n^2} + \eta\left(\frac{1}{n}\right)^2\right) \sup_{0\leqslant t\leqslant 1} |M_{\ell,t}|^2 > C_5\delta\right). \end{aligned} \tag{3.52}$$

By Lévy's inequality for a Brownian motion and our choice (3.50) of $\varepsilon(n)$, the limit (3.52) are also $-\infty$. As in [56], it's more little difficult to estimate (3.51). By the isoperimetric inequality [91, p17,(1.24)] and our choice (3.50), we conclude that the limit (3.51) are both $-\infty$.

Proof of Remark 3.3.11 We have to verify condition (3.14), when the drift is sublinear for all $x \in \mathbb{R}^2$

$$|b(x)| \leq C\|x\|^\alpha, \qquad \alpha < 1,$$

where $C > 0$ is a constant. We have that

$$\sum_{k=1}^n \left(\int_{t_{k-1}}^{t_k} b_\ell(s, X_s)\mathrm{d}s\right)^2 \leq C^2 \sum_{k=1}^n \left(\int_{t_{k-1}}^{t_k} \|X_s\|^\alpha ds\right)^2 \leq \frac{C^2}{n} \max_{k=1}^n \sup_{t_{k-1}\leq s\leq t_k} |X_{\ell,s}|^{2\alpha}$$

Using the fact that $|z|^\alpha \leq \alpha\,|z| + 1 - \alpha$ for $\alpha < 1$, we obtain that

$$\begin{aligned} |X_{\ell,t}| &\leq |X_{\ell,0}| + \left|\int_0^t \sigma_{\ell,s} dB_{\ell,s}\right| + C\int_0^t |X_{\ell,s}|^\alpha\, ds \\ &\leq C + \left|\int_0^t \sigma_{\ell,s} dB_{\ell,s}\right| + C\int_0^t |X_{\ell,s}|\, ds, \end{aligned}$$

where the constant C may differ from one line to another. By Gronwall's lemma, we obtain

$$|X_{\ell,t}| \leq C + |M_{\ell,t}| + C\int_0^t |M_{\ell,s}|\, ds.$$

So we have

$$\sum_{k=1}^n \left(\int_{t_{k-1}}^{t_k} \|X_s\|^\alpha ds\right)^2 \leq \frac{C}{n} + \frac{C}{n} \max_{k=1}^n \sup_{t_{k-1}\leq s\leq t_k} |M_{\ell,s}|^{2\alpha}.$$

By Lévy's inequality, we obtain that

$$\mathbb{P}\left(\max_{k=1}^n \sup_{t_{k-1}\leq s\leq t_k} |M_s|^{2\alpha} > \delta\varepsilon(n)n\right) \leq 4n\exp\left(-\frac{\delta^2\varepsilon(n)^{1/\alpha}n^{1/\alpha}}{2\|\sigma^2\|_\infty}\right).$$

Choosing the sequence $\varepsilon(n)$ such that $\varepsilon(n)^{1/\alpha}n^{1/\alpha-1} \to \infty$, we obtain (3.14).

Proof of Corollary 3.4.2

We have just to do the identification of the rate function. We knew that $\varrho_1^n(X)$ satisfies the LDP on $\mathbb{R}$ with speed n and with the good rate function given by

$$I_{ldp}^\rho(u) := \inf_{\{(x_1,x_2,x_3)\in\mathbb{R}^3: x_3 = u\sqrt{x_1x_2}, x_1>0, x_2>0\}} I_{ldp}^V(x_1, x_2, x_3),$$

where I^V_{ldp} is given in (3.9). So

$$I^{\rho}_{ldp}(u) = \inf\left\{\log\left(\frac{\sqrt{\sigma_1^2\sigma_2^2(1-\rho^2)}}{\sqrt{x_1x_2}\sqrt{1-u^2}}\right) - 1 + \frac{\sigma_2^2x_1+\sigma_1^2x_2-2\rho\sigma_1\sigma_2u\sqrt{x_1x_2}}{2\sigma_1^2\sigma_2^2(1-\rho^2)}, \quad x_1>0, x_2>0\right\}.$$

The above infinimum is attained at the point $(x_1, x_2) = \left(\dfrac{\sigma_1^2(1-\rho^2)}{1-\rho u}, \dfrac{\sigma_2^2(1-\rho^2)}{1-\rho u}\right)$, so we obtain (3.15).

Proof of Proposition 3.4.3

As said before, quite unusually, the MDP is here a little bit harder to prove, due to the fact that it is not a simple transformation of the MDP of $\frac{\sqrt{n}}{v_n}(V^n_t - [V]_t)$. Therefore we will use the strategy developped for the TCL : the delta-method. Fortunately, Gao and Zhao [73] have developped such a technology at the large deviations level. However it will require to prove quite heavy exponential negligibility to be able to do so. For simplicity we omit X in the notations of $Q^n_{1,t}(X)$ and $C^n_t(X)$.

Let introduce $\Xi^n_t := \sqrt{Q^n_{1,t}Q^n_{2,t}}$. Then by [73, Theorem 3.1] applied to the functions $g(x,y,z) = \sqrt{xy}$ and $h(x,y,z) = \dfrac{1}{\sqrt{xy}}$, we deduce that

$$\frac{\sqrt{n}}{v_n}\left(\Xi^n_1 - \mathbb{E}\Xi^n_1\right) \quad \text{and} \quad \frac{\sqrt{n}}{v_n}\left(\frac{1}{\Xi^n_1} - \frac{1}{\mathbb{E}\Xi^n_1}\right)$$

satisfies the LDP on $\mathbb{R}$ with speed v_n^2 and rates functions respectively given by :

$$I^{\Xi}_{mdp}(u) := \inf_{\{(x,y,z)\in\mathbb{R}^3, u=\frac{\sigma_1^2y+\sigma_2^2x}{2\sigma_1\sigma_2}\}}\{I_{mdp}(x,y,z)\}$$

and

$$I^{\Xi^{-1}}_{mdp}(u) := \inf_{\{(x,y,z)\in\mathbb{R}^3, u=-\frac{\sigma_1^2y+\sigma_2^2x}{2\sigma_1^3\sigma_2^3}\}}\{I_{mdp}(x,y,z)\},$$

where I_{mdp} is given in (3.11).

By some simple calculations, we have

$$\varrho^n_1(X) - \varrho = \aleph^n_1 + \aleph^n_2 + \aleph^n_3 + \aleph^n_4 - \aleph^n_5 - \aleph^n_6, \tag{3.53}$$

where

$$\aleph^n_1 := (C^n_1 - \mathbb{E}C^n_1)\left(\frac{1}{\Xi^n_1} - \frac{1}{\mathbb{E}\Xi^n_1}\right), \qquad \aleph^n_2 := (C^n_1 - \mathbb{E}C^n_1)\frac{1}{\mathbb{E}\Xi^n_1},$$

$$\aleph_3^n := (\mathbb{E}C_1^n - \varrho\mathbb{E}\Xi_1^n)\left(\frac{1}{\Xi_1^n} - \frac{1}{\mathbb{E}\Xi_1^n}\right), \qquad \aleph_4^n := (\mathbb{E}C_1^n - \varrho\mathbb{E}\Xi_1^n)\frac{1}{\mathbb{E}\Xi_1^n},$$
$$\aleph_5^n := \varrho(\Xi_1^n - \mathbb{E}\Xi_1^n)\left(\frac{1}{\Xi_1^n} - \frac{1}{\mathbb{E}\Xi_1^n}\right), \qquad \aleph_6^n := \varrho(\Xi_1^n - \mathbb{E}\Xi_1^n)\frac{1}{\mathbb{E}\Xi_1^n}.$$

To prove the Theorem 3.4.3, we have to use [73, Theorem 3.1] and prove some negligibility in the sence of MDP. We have to prove that for $\ell = 1, 3, 4, 5$

$$\frac{\sqrt{n}}{v_n}\aleph_\ell^n \quad \underset{v_n^2}{\overset{\text{superexp}}{\longrightarrow}} 0. \tag{3.54}$$

Since $\mathbb{E}C_1^n - \varrho\mathbb{E}\Xi_{1,n} \to 0$ as $n \to \infty$, (3.54) follows for $\ell = 4$

We have for all $\delta > 0$ and with $\alpha_n = \sqrt{\frac{\sqrt{n}}{b_n}\delta}$

$$\mathbb{P}\left(\frac{\sqrt{n}}{v_n}|\aleph_1^n| \geqslant \delta\right) \leqslant \mathbb{P}\left(\frac{\sqrt{n}}{v_n}\left|C_1^n - \mathbb{E}C_1^n\right| \geqslant \alpha_n\right) + \mathbb{P}\left(\frac{\sqrt{n}}{v_n}\left|\frac{1}{\Xi_{1,n}} - \frac{1}{\mathbb{E}\Xi_{1,n}}\right| \geqslant \alpha_n\right).$$

So, by [53, Lemma 1.2.15], we have that for all $\delta > 0$

$$\limsup_{n\to\infty} \frac{1}{v_n^2} \log \mathbb{P}\left(\frac{\sqrt{n}}{v_n}|\aleph_1^n| \geqslant \delta\right)$$

is majorized by the maximum of the following two limits

$$\limsup_{n\to\infty} \frac{1}{v_n^2} \log \mathbb{P}\left(\frac{\sqrt{n}}{v_n}\left|C_1^n - \mathbb{E}C_1^n\right| \geqslant \alpha_n\right), \tag{3.55}$$

$$\limsup_{n\to\infty} \frac{1}{v_n^2} \log \mathbb{P}\left(\frac{\sqrt{n}}{v_n}\left|\frac{1}{\Xi_{1,n}} - \frac{1}{\mathbb{E}\Xi_{1,n}}\right| \geqslant \alpha_n\right). \tag{3.56}$$

Let $A > 0$ be arbitrary, since $\alpha_n \to \infty$ as $n \to \infty$, so for n large enough we obtain that

$$\mathbb{P}\left(\frac{\sqrt{n}}{v_n}\left|C_1^n - \mathbb{E}C_1^n\right| \geqslant \alpha_n\right) \leqslant \mathbb{P}\left(\frac{\sqrt{n}}{v_n}\left|C_1^n - \mathbb{E}C_1^n\right| \geqslant A\right).$$

By the MDP of $\dfrac{\sqrt{n}}{v_n}(C_1^n - \mathbb{E}C_1^n)$ obtained in [**?**, Theorem 2.3], and by letting n to infinity, we obtain that

$$\limsup_{n\to\infty} \frac{1}{v_n^2} \log \mathbb{P}\left(\frac{\sqrt{n}}{v_n}\left|C_1^n - \mathbb{E}C_1^n\right| \geqslant \alpha_n\right) \leqslant -\inf_{|x|\geqslant A} I_{mdp}^C(x).$$

Letting A gos to the infinity, we obtain that the term in (3.55) goes to $-\infty$.

By the MDP of $\dfrac{\sqrt{n}}{v_n}\left(\dfrac{1}{\Xi_1^n}-\dfrac{1}{\mathbb{E}\Xi_1^n}\right)$ stated before and in the same way we obtain that the term in (3.56) goes to $-\infty$. So we obtain (3.54) for $\ell=1$. The same calculations give (3.54) for $\ell=3,5$. So

$$\frac{\sqrt{n}}{v_n}(\varrho_1^n(X)-\varrho) \quad \text{and} \quad \frac{\sqrt{n}}{v_n}\Big(C_1^n-\mathbb{E}C_1^n-\varrho(\Xi_1^n-\mathbb{E}\Xi_1^n)\Big)\frac{1}{\mathbb{E}\Xi_1^n}$$

satisfies the same MDP. Since $\mathbb{E}\Xi_1^n \longrightarrow \sigma_1\sigma_2$, so

$$\frac{\sqrt{n}}{v_n}(\varrho_1^n-\varrho) \quad \text{and} \quad \frac{\sqrt{n}}{v_n}\Big(C_1^n-\mathbb{E}C_1^n-\varrho(\Xi_1^n-\mathbb{E}\Xi_1^n)\Big)\frac{1}{\sigma_1\sigma_2}$$

satisfies the same MDP.

Then by [73, Theorem 3.1] applied to the function $\phi(x,y,z)=(z-\varrho\sqrt{x}\sqrt{y})/\sigma_1\sigma_2$, we deduce that $\dfrac{\sqrt{n}}{v_n}(\varrho_1^n(X)-\varrho)$ satisfies the LDP on $\mathbb{R}$ with speed v_n^2 and with rate function given by

$$I_{mdp}^{\varrho}(u)=\inf_{\{(x,y,z)\in\mathbb{R}^3:u=\frac{z}{\sigma_1\sigma_2}-\varrho\frac{\sigma_1^2y+x\sigma_2^2}{2\sigma_1^2\sigma_2^2}\}} I_{mdp}(x,y,z),$$

where I_{mdp} is given in (3.11).

Proof of Proposition 3.4.7

By [73, Theorem 3.1] we have that $\dfrac{\sqrt{n}}{v_n}\left(\dfrac{1}{Q_{1,1}^n}-\dfrac{1}{\mathbb{E}Q_{1,1}^n}\right)$ satisfies the LDP on $\mathbb{R}$ with speed v_n^2 and rate function given by

$$I_{mdp}^{Q_1^{-1}}(u):=\inf_{\{(x,y,z)\in\mathbb{R}^3:u=-\frac{x}{\sigma_1^4}\}}\{I_{mdp}(x,y,z)\}$$

By some simple calculations, we have

$$\beta_{1,1}^n(X)-\varrho\frac{\sigma_2}{\sigma_1}=\jmath_1^n+\jmath_2^n+\jmath_3^n+\jmath_4^n-\jmath_5^n-\jmath_6^n, \tag{3.57}$$

where

$$\jmath_1^n:=(C_1^n-\mathbb{E}C_1^n)\left(\frac{1}{Q_{1,1}^n}-\frac{1}{\mathbb{E}Q_{1,1}^n}\right), \qquad \jmath_2^n:=(C_1^n-\mathbb{E}C_1^n)\frac{1}{\mathbb{E}Q_{1,1}^n},$$

$$\jmath_3^n:=(\mathbb{E}C_1^n-\varrho\frac{\sigma_2}{\sigma_1}\mathbb{E}Q_{1,1}^n)\left(\frac{1}{Q_{1,1}^n}-\frac{1}{\mathbb{E}Q_{1,1}^n}\right), \qquad \jmath_4^n:=(\mathbb{E}C_1^n-\varrho\frac{\sigma_2}{\sigma_1}\mathbb{E}Q_{1,1}^n)\frac{1}{\mathbb{E}Q_{1,1}^n},$$

$$\jmath_5^n := \varrho\frac{\sigma_2}{\sigma_1}\left(Q_{1,1}^n - \mathbb{E}Q_{1,1}^n\right)\left(\frac{1}{Q_{1,1}^n} - \frac{1}{\mathbb{E}Q_{1,1}^n}\right), \qquad \jmath_6^n := \varrho\frac{\sigma_2}{\sigma_1}\left(Q_{1,1}^n - \mathbb{E}Q_{1,1}^n\right)\frac{1}{\mathbb{E}Q_{1,1}^n}.$$

To prove the Theorem 3.4.7, we have to use [73, Theorem 3.1] and prove some negligibility in the sense of MDP. For $\ell = 1, 3, 4, 5$

$$\frac{\sqrt{n}}{v_n}\jmath_\ell^n \xrightarrow[v_n^2]{\text{superexp}} 0. \tag{3.58}$$

The same calculations as for the negligibility of $\aleph_j^n$ works here to obtain (3.58).

Since $\mathbb{E}Q_{1,1}^n \longrightarrow \sigma_1^2$, we deduce that

$$\frac{\sqrt{n}}{v_n}\left(\beta_{1,1}^n(X) - \varrho\frac{\sigma_2}{\sigma_1}\right) \quad \text{and} \quad \frac{\sqrt{n}}{v_n}\left(C_1^n - \mathbb{E}C_1^n - \varrho\frac{\sigma_2}{\sigma_1}\left(Q_{1,1}^n - \mathbb{E}Q_{1,1}^n\right)\right)\frac{1}{\sigma_1^2}$$

satisfies the same MDP.

Then by [73, Theorem 3.1] applied to the function $k(x, y, z) = (z - \varrho\frac{\sigma_2}{\sigma_1}x)/\sigma_1^2$ we deduce that $\frac{\sqrt{n}}{v_n}\left(\beta_{1,1}^n(X) - \varrho\frac{\sigma_2}{\sigma_1}\right)$ satisfies the LDP on $\mathbb{R}$ with speed v_n^2 and with rate function given by

$$I_{mdp}^{\beta_{1,1}}(u) = \inf_{\{(x,y,z)\in\mathbb{R}^3 : u = (z - \varrho\frac{\sigma_2}{\sigma_1}x)/\sigma_1^2\}} I_{mdp}(x, y, z)$$

where I_{mdp} is given in (3.11).

Acknowledgments : We are grateful to the anonymous referee and the assoicated editor of stochastic processe and thier applications for their valuable comments which improve the presentation of this work.

3.6 Appendix

The proofs of the LDP in Theorems 3.3.1 and 3.3.3 are respectively based on the Lemmas 3.6.1 and 3.6.2 that we will present here for completeness.

Avoiding Gärtner-Ellis theorem by Najim [105, 104]

Let us introduce some notations and assumptions in this section.

Let $\mathcal{X}$ be a topological vector compact space endowed with it's Borel $\sigma-$field $\mathcal{B}(\mathcal{X})$. Let $\mathcal{BV}([0,1], \mathbb{R}^d)$, (shortened in $\mathcal{BV}$) be a space of functions of bounded variation on $[0,1]$ endowed with it's Borel $\sigma-$field $\mathcal{B}_w$. Let $\mathcal{P}(\mathcal{X})$ the set of probability measures on $\mathcal{X}$.

Let $\tau(z) = e^{|z|} - 1, z \in \mathbb{R}^d$ and let us consider

$$\begin{aligned} \mathcal{P}_\tau(\mathbb{R}^d) &= \left\{ P \in \mathcal{P}(\mathbb{R}^d), \exists a > 0; \int_{\mathbb{R}^d} \tau\left(\frac{z}{a}\right) P(dz) < \infty \right\} \\ &= \left\{ P \in \mathcal{P}(\mathbb{R}^d), \exists \alpha > 0; \int_{\mathbb{R}^d} e^{\alpha|z|} P(dz) < \infty \right\}. \end{aligned}$$

$\mathcal{P}_\tau$ is the set of probability distributions having some exponential moments.We denote by $M(P,Q)$ the set of all laws on $\mathbb{R}^d \times \mathbb{R}^d$ with given marginals P and Q. We introduce the Orlicz-Wasserstein distance defined on $\mathcal{P}_\tau(\mathbb{R}^d)$ by

$$d_{OW}(P,Q) = \inf_{\eta \in M(P,Q)} \inf\left\{ a > 0; \int_{\mathbb{R}^d \times \mathbb{R}^d} \tau\left(\frac{z - z'}{a}\right) \eta(dzdz') \leqslant 1 \right\}$$

Let $(Z_i^n)_{1 \leqslant i \leqslant n, n \in \mathbb{N}}$ be a sequence of $\mathbb{R}^d-$valued independent random variables satisfying :

N-1

$$\mathbb{E}e^{\alpha.||Z||} < +\infty \qquad for \quad some \quad \alpha > 0.$$

N-2 Let $(x_i^n, 1 \leqslant i \leqslant n, n \geqslant 1)$ be a $\mathcal{X}-$valued sequence of elements satisfying :

$$\frac{1}{n}\sum_{j=1}^{n} \delta_{x_j^n} \xrightarrow[n\to\infty]{weakly} \quad R.$$

Where R is Assumed to be a strictly positive probability measure, that is $R(U) > 0$ whenever U is a nonempty open subset of $\mathcal{X}$.

N-3 $\mathcal{X}$ is a compact space.

N-4 There exist a family of probability measure $(P(x,\cdot), x \in \mathcal{X})$ over $\mathbb{R}^d$ and a sequence $(x_i^n, 1 \leqslant i \leqslant n, n \geqslant 1)$ with values in $\mathcal{X}$ such that the law of each Z_i^n is given by :

$$\mathcal{L}(Z_i^n) \sim P(x_i^n, dz).$$

We will equally write $P(x,\cdot), P_x \quad or \quad P_x(dz)$.

N-5 Let $(P(x,\cdot), x \in \mathcal{X}) \subset \mathcal{P}_\tau(\mathbb{R}^d)$ be a given distribution. The application $x \mapsto P(x,A)$ is measurable whenever the set $A \subset \mathbb{R}^d$ is borel. Morever, the function

$$\begin{aligned} \Gamma :&\mathcal{X} \to \mathcal{P}_\tau(\mathbb{R}^d) \\ &x \mapsto P(x,\cdot) \end{aligned}$$

is continuous when $\mathcal{P}_\tau(\mathbb{R}^d)$ is endowed with the topology induced by the distance d_{OW}

Lemma 3.6.1 *Theorem 2.2 in [105]*

Assume that Z_i^n are independent and identically distributed, so we denote Z_i^n by Z_i.

Assume that (N-1) and (N-2) hold. Let $f : \mathcal{X} \to \mathbb{R}^{m \times d}$ be a (matrix-valued) bounded continuous function, such that

$$f(x) \cdot z = \begin{pmatrix} f_1(x) \cdot z \\ \\ \vdots \\ \\ f_m(x) \cdot z \end{pmatrix}$$

where each $f_j \in C_d(\mathcal{X})$ is the j^{th} row of the matrix f.

Then the family of the weighted empirical mean

$$\langle L_n, f \rangle := \frac{1}{n} \sum_{i=1}^{n} f(x_i^n) \cdot Z_i$$

satisfies the LDP in $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$ with speed n and the good rate function

$$I_f(x) = \sup_{\theta \in \mathbb{R}^m} \{\langle \theta, x \rangle - \int_{\mathcal{X}} \Lambda[\sum_{i=1}^{m} \theta_i \cdot f_i(x)] \mathrm{R}(dx)\} \qquad \forall x \in \mathbb{R}^m$$

where Λ denote the cumulant generating function of Z

$$\Lambda(\lambda) = \log \mathbb{E} e^{\lambda \cdot Z} \qquad for \quad \lambda \in \mathbb{R}^d$$

Lemma 3.6.2 *Theorem 4.3 in [106]*

Assume that (N-1), (N-2), (N-3), (N-4) and (N-5) hold. Then the family of random functions

$$\overline{Z}_n(t) = \frac{1}{n} \sum_{k=1}^{[nt]} Z_k^n, \quad t \in [0, 1]$$

satisfies the LDP in $(\mathcal{BV}, \mathcal{B}_w)$ with the good rate function

$$\phi(f) = \int_{[0,1]} \Lambda^*(x, f_a'(x)) \mathrm{d}x + \int_{[0,1]} \rho(x, f_s'(x)) \mathrm{d}\theta(x)$$

where θ is any real-valued nonnegative measure with respect to which μ_s^f is absolutely continuous and $f_s' = \frac{d\mu_s^f}{d\theta}$, where

$$\Lambda^*(x, z) = \sup_{\lambda \in \mathbb{R}^d} \Big\{ \langle \lambda, z \rangle - \Lambda(x, \lambda) \Big\}, \quad \forall z \in \mathbb{R}^d$$

with $\Lambda(x,\lambda) = \log \int_{\mathbb{R}^d} e^{\langle \lambda, z\rangle} P(x,dz), \quad \forall \lambda \in \mathbb{R}^d$ *and the recession function* $\rho(x,z)$ *of* $\Lambda^*(x,z)$ *defined by :* $\rho(x,z) = \sup\{\langle \lambda, z\rangle, \lambda \in D_x\}$ *with* $D_x = \{\lambda \in \mathbb{R}^d, \Lambda(x,\lambda) < \infty\}$.

Delta method for large deviations [73]

In this section, we recall the delta method in large deviation.

Let $\mathcal{X}$ and $\mathcal{Y}$ be two metrizable topological linear spaces. A function ϕ defined on a subset $\mathcal{D}_\phi$ of $\mathcal{X}$ with values on $\mathcal{Y}$ is called Hadamard differentiable at x if there exists a continuous functions $\phi' : \mathcal{X} \mapsto \mathcal{Y}$ such that

$$\lim_{n\to\infty} \frac{\phi(x + t_n h_n) - \phi(x)}{t_n} = \phi'(h) \tag{3.59}$$

holds for all t_n converging to 0+ and h_n converging to h in $\mathcal{X}$ such that $x + t_n h_n \in \mathcal{D}_\phi$ for every n.

Lemma 3.6.3 *Let* $\mathcal{X}$ *and* $\mathcal{Y}$ *be two metrizable topological linear spaces. Let* $\phi : \mathcal{D} \subset \mathcal{X} \mapsto \mathcal{Y}$ *be a Hadamard differentiable at* θ *tangentially to* $\mathcal{D}_0$*, where* $\mathcal{D}_\phi$ *and* $\mathcal{D}_0$ *are two subset of* $\mathcal{X}$*. Let* $\{(\Omega_n, \mathcal{F}_n, \mathbb{P}_n), n \geqslant 1\}$ *be a sequence of probability space and let* $\{X_n, n \geqslant 1\}$ *be a sequence of maps from from* Ω_n *to* $\mathcal{D}_\phi$ *and let* $\{r_n, n \geqslant 1\}$ *be a sequence of positive real numbers satisfying* $r_n \to +\infty$ *and let* $\{\lambda(n), n \geqslant 1\}$ *be a sequence of positive real numbers satisfying* $\lambda(n) \to +\infty$.
If $\{r_n(X_n - \theta), n \geqslant 1\}$ *satifies the LDP with speed* $\lambda(n)$ *and rate function* I *and* $\{I < \infty\} \subset \mathcal{D}_0$*, then* $\{r_n(\phi(X_n) - \phi(\theta)), n \geqslant 1\}$ *satifies the LDP with speed* $\lambda(n)$ *and rate function* $I_{\phi'_\theta}$*, where*

$$I_{\phi'_\theta}(y) = inf\{I(x); \phi'_\theta(x) = y\}, \qquad y \in \mathcal{Y} \tag{3.60}$$

Chapitre 4

Moderate deviations for bipower variation of general function and Hayashi-Yoshida estimators

Hacène DJELLOUT, Hui JIANG, Arnaud GUILLIN and Yacouba SAMOURA

abstract :

We consider the moderate deviations behaviors for two (co-) volatility estimators : generalised bipower variation, Hayashi-Yoshida estimator. The results are obtained by using a new result about the moderate deviations principle for m-dependent random variables based on the Chen-Ledoux type condition.
keywords : Moderate deviation principle, nonsynchronicity, M-dependent, Diffusion, Discrete-time observation, Quadratic variation, Volatility, Bipower variation.

Résumé :

On considère les déviations modérées de deux estimateurs : la variation généralisée à double puissance et l'estimateur de Hayashi-Yoshida. Les résultats sont obtenus par l'utilisation d'une nouvelle approche sur les déviations modérées des variables aléatoires $m-$dependantes vérifiant des conditions de type "Chen-Ledoux".
mots clés : Déviations modérées, variation généralisée à double puissance, estimateur de Hayashi-Yoshida, $m-$dependantes, condition de type "Chen-Ledoux".

4.1 Motivation and context

In the last decade there has been a considerable development of the asymptotic theory for processes observed at a high frequency. This was mainly motivated by financial applications, where the data, such as stock prices or currencies, are observed very frequently. As under the no-arbitrage assumptions price processes must follow a semimartingale, there was a need for probabilistic tools for functionals of semimartingales based on high frequency observations. Inspired by potential applications, probabilists started to develop limit theorems for semimartingales. Statisticians applied the asymptotic theory to analyze the path properties of discretely observed semimartingales : for the estimation of certain volatility functionals and realised jumps, or for performing various test procedures.

We consider $X_t = (X_{1,t}, X_{2,t})_{t\in[0,T]}$ a 2-dimensional semimartingale, defined on the filtred probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{[0,T]}, \mathbb{P})$, of the form

$$\begin{cases} dX_{1,t} = b_1(t, X_t)dt + \sigma_{1,t}dW_{1,t} \\ dX_{2,t} = b_2(t, X_t)dt + \sigma_{2,t}dW_{2,t} \end{cases} \tag{4.1}$$

where $W_1 = (W_{1,t})_{t\in[0,T]}$ and $W_2 = (W_{2,t})_{t\in[0,T]}$ are two correlated Wiener processes with $\rho_t = \mathrm{Cov}(W_{1,t}, W_{2,t})$, $t \in [0, T]$. Moreover, $\rho_\cdot \in [0, 1]$ and $\sigma_{\ell,\cdot}, \ell = 1, 2$ are both unknown deterministic and measurable functions of t, $b_\ell(\cdot, \cdot), \ell = 1, 2$ are progressively measurable (possibly unknown) functions.

Models of the type (4.1) and their extensions are widely used in mathematical finance to capture the dynamics of stock prices or interest rates.

This work aims to provide analysis of modearte deviations principle for two estimators : Hayashi-Yoshida estimator and generalised bipower estimator. To do that we prove a new result about moderate deviations for m-dependent random vgariables using the Chen-Ledoux type condition. This condition links the speed of the moderate deviation principle to the queue of the distribution of the random variables. It known that the Chen-Ledoux type condition is a necessary and sufficient condition for the independent and identically distributed random variables. Djellout [56] using a similar condition have obtained the moderate deviation for random variables. Djellout and Guillin [55] have also obtained the moderate deviation for Markov chain using this conditions. See also the work of Djellout Bitseki Penda and Proia [24] for the moderate deviation of the Durbin Watson statistics using this condition.

Hayashi-Yoshida estimator We concentrate on the estimation of the co-volatility of X_1 and X_2

$$\langle X_1, X_2\rangle_T = \int_0^T \sigma_{1,t}\sigma_{2,t}\rho_t dt.$$

Given the synchronous observations of the processes $(X_{1,t_i}, X_{2,t_i})_{i=0,...,n}$ a popular statistic to estimate the co-volatility is

$$CV_n := \sum_{i=1}^{n} \Delta X_1(I^i)\Delta X_2(I^i),$$

which is often called the realized co-volatility estimator ([3]).

The asymptotic distribution of CV_n was formulated by Barndorff et al. ([13]). Djellout and Samoura ([61]), Djellout et al. ([58]) obtained the large and moderate deviations for the realized covolatility CV_n. For more references, one can see ([59],[84],[85],[99]) and the references therein.

However, in financial applications, actual transaction data are recorded at irregular times in a nonsynchronous manner, i.e. two transaction prices are usually not observed at the same time. This fact requires one who adopt CV_n to synchronize the original data a prior, choose a common interval length h first, then impute missing observations by some interpolation scheme such as previous-tick interpolation of linear interpolation. Unfortunately, those procedures may result in synchronization bias ([79]).

Recently, Hayashi and Yoshida ([79]) postulated a new estimator which is free of synchronization and hence of any bias due to it. To be explicitly, for the nonsynchronous observations-$\Big(X_{1,t_i}, X_{2,s_j}\Big)_{i=0,...,n;j=0,...,m}$ with $0 = s_0 < s_1 < ... < s_m = T, 0 = t_0 < t_1 < ... < t_n = T, m, n \in \mathbb{N}$, the Hayashi-Yoshida estimator is defined as

$$U_{n,m} := \sum_{i=1}^{n}\sum_{j=1}^{m} \Delta X_1(I^i)\Delta X_2(J^j)I_{\{I^i \cap J^j \neq \emptyset\}}, \tag{4.2}$$

where $I^i = (t_{i-1}, t_i]$, $J^j = (s_{j-1}, s_j]$. Under some assumptions on the equation (4.1), Hayashi and Yoshida ([79],[80]) showed the consistency and asymptotic normality of $U_{n,m}$ respectively.

This article is structured as follows. In Section 2 we present the main theoretical results and we state the proofs in the Section 3.

Generalised Bipower estimator Recently, the concept of realised bipower variation has built a non-parametric framework for backing out several variational measures of volatility, which has led to a new development in econometrics. Given the observations of the processes $(X_{\ell,t_i})_{i=0,...,n}$ with $0 = t_0 < t_1 < ... < t_n = T, n \in \mathbb{N}$, the realised Bipower variation which is given by

$$V^n_{\ell,1}(r,q) = n^{\frac{r+q}{2}-1} \sum_{i=1}^{n-1} \left|\Delta X_\ell(I^i)\right|^r \left|\Delta X_\ell(I^{i+1})\right|^q, \tag{4.3}$$

where $I^i = (t_{i-1}, t_i]$ $\Delta X_\ell(I^i) = X_{\ell,t_i} - X_{\ell,t_{i-1}}$, provides a whole variety of estimators for different (integrated) powers of volatility. For practical convenience, it is standard to take equal spacing, i.e., $t_i - t_{i-1} = \frac{T}{n} := h$. An important special case of the class (4.3) is the realised volatility

$$\sum_{i=1}^{n} |\Delta X_\ell(I^i)|^2, \tag{4.4}$$

which is a consistent estimator of the quadratic variation of X_ℓ, i.e.

$$IV_t = \int_0^t \sigma_{\ell,s}^2 ds,$$

which is often referred to as integrated volatility in the econometric literature.

In the last years, Nielsen et al.([13],[16],[14]) showed the consistency of $V_{\ell,1}^n(r,q)$ and introduced the (stable) Central Limit Theorem (CLT, in short) for standardised version. Vetter [131] extend the results from Jacod [81] to the case of bipower variation and he prove the CLT of the bipower variation for continuous semimartingales. In [12] they consider the same problems as here where X have jumps and the CLT has been solved. It was demonstrated that realized bipower variation can estimate integrated power volatility in stochastic volatility models and moreover, under some conditions, it can be a good measure to integrated variance in the presence of jumps.

S. Kinnebrock and M. Podolskij [87] extended the CLT to the bipower variation of general functions (called generalised bipower variation) :

$$V_{\ell,1}^n(g,h) = \frac{1}{n}\sum_{i=1}^{n} g\left(\sqrt{n}\Delta X_\ell(I^i)\right) h\left(\sqrt{n}\Delta X_\ell(I^{i+1})\right), \tag{4.5}$$

where g, h are two maps on $\mathbb{R}^d$, taking values in $\mathbb{R}\times\mathbb{R}$ and given some examples from the litterature to which his theory can be applied.

We know from [13] and [16] that if g and h are continuously differentiable with g, h, g' and h' being of at most polynomial growth

$$V_{\ell,1}^n(g,h) \xrightarrow{\mathbb{P}} V_{\ell,1}(g,h) := \int_0^1 \Sigma_{\sigma_{\ell,u}}(g)\Sigma_{\sigma_{\ell,u}}(h)du,$$

where

$$\Sigma_\sigma(f) := \mathbb{E}(f(\sigma Z)), \quad Z \sim \mathcal{N}(0,1), \quad f \text{ is a real-valued function on } \mathbb{R}.$$

In additional, if h and g are even (i.e. $g(x) = g(-x)$, $h(x) = h(-x)$ for all $x \in \mathbb{R}^d$), we have the CLT

$$\sqrt{n}\left(V_{\ell,1}^n(g,h) - V_{\ell,1}(g,h)\right) \to \mathcal{N}(0, \Sigma_\ell(g,h))$$

where

$$\Sigma_\ell(g,h) = \int_0^1 \left(\Sigma_{\sigma_{\ell,s}}(g^2)\Sigma_{\sigma_{\ell,s}}(h^2) + 2\Sigma_{\sigma_{\ell,s}}(g)\Sigma_{\sigma_{\ell,s}}(h)\Sigma_{\sigma_{\ell,s}}(gh) - 3\Sigma_{\sigma_{\ell,s}}^2(g)\Sigma_{\sigma_{\ell,s}}^2(h)\right) ds.$$

All the elements of f on $\mathbb{R}^d$ are continuous with at most polynomial growth. This amounts to there being suitable constants $C > 0$ and $p \geq 2$ such that

$$\forall x \in \mathbb{R}^d, \|f(x)\| \leq C(1 + \|x\|^p). \tag{4.6}$$

4.2 Main Results

4.2.1 Moderate deviations for the Hayashi-Yoshida estimator

Firstly, a reduced design with respect to $(I^i)_{i=1,\ldots,n}$ will be constructed in the following manner. We collect all I^is such that $I^i \subset J^j$ and combine them into a new interval ; if such I^i does not exist, do nothing. Then collecting all such intervals and re-labeling them from left to right yields a new partition of $(0, T]$, denoted $\left(\hat{I}^i\right)_{i=1,\ldots,\hat{n}}$. Due to the bilinearity of $U_{n,m}$,

$$U_{n,m} = \sum_{i=1}^{\hat{n}} \sum_{j=1}^{m} \Delta X_1(\hat{I}^i) \Delta X_2(J^j) K_{ij}^{\hat{I}} = \sum_{i=1}^{\hat{n}} \Delta X_1(\hat{I}^i) \Delta X_2 \left(\cup_{j \in \hat{J}(i)} J^j\right), \tag{4.7}$$

where $\hat{J}(i) := \left\{1 \leq j \leq m : K_{ij}^{\hat{I}} \neq 0\right\}$ with $K_{ij}^{\hat{I}} = I_{\{\hat{I}^i \cap J^j \neq \emptyset\}}$. Then, we can find that each J^j contains at most one $\hat{I}^i$, which implies that the random variable sequence

$$\left\{1 \leq i \leq \hat{n} : \Delta X_1(\hat{I}^i) \Delta X_2 \left(\cup_{j \in \hat{J}(i)} J^j\right)\right\}$$

are 2-dependent. Now, the number $\hat{n}$ can be formulated as follows. Define

$$\tau_1 = \inf\{1 \leq i \leq n : I^i \not\subseteq J^1\}, \ \varsigma_1 = \sup\{1 \leq j \leq m : I^{\tau_1} \cap J^j \neq \emptyset\}$$

and

$$\tau_k = \inf\{\tau_{k-1} < i \leq n : I^i \not\subseteq J^{\varsigma_{k-1}}\}, \ \varsigma_k = \sup\{\varsigma_{k-1} < j \leq m : I^{\tau_k} \cap J^j \neq \emptyset\},$$

with $\tau_0 = 0$, $\varsigma_0 = 1$ and $\inf \emptyset = +\infty$, $\sup \emptyset = 0$. Let $n_0 = \sup\{k : \tau_k < +\infty\}$. Then, one can conclude that

$$n_0 \leq \hat{n} \leq n_0 + \sum_{k=1}^{n_0} I_{\{\tau_k - \tau_{k-1} > 1\}} + I_{\{\tau_{n_0} < n\}} \leq 2n_0 + 1. \tag{4.8}$$

Let $A_{\ell,t} = \int_0^t b_\ell(s,\omega) ds$, $t \in [0,T]$ and $\ell = 1, 2$. When the drift $b_\ell(t,\omega)$ is known, we can consider the following estimator

$$\begin{aligned} V_{n,m} : &= \sum_{i=1}^{n} \sum_{j=1}^{m} \Delta X_1^0(I^i) \Delta X_2^0(J^j) I_{\{I^i \cap J^j \neq \emptyset\}} \\ &= \sum_{i=1}^{\hat{n}} \Delta X_1^0(\hat{I}^i) \Delta X_2^0 \left(\cup_{j \in \hat{J}(i)} J^j\right). \end{aligned} \tag{4.9}$$

where $X^0_{\ell,t} = X_{\ell,t} - A_{\ell,t}$. For any Borel set $I \subset [0,T]$, define

$$\nu(I) = \int_I \sigma_{1,t}\sigma_{2,t}\rho_t dt,\ \nu_\ell(I) = \int_I \sigma^2_{\ell,t} dt,\ \ell = 1,2.$$

Let $r_{n,m} := \max_{1\le i\le n} |I^i| \vee \max_{1\le j\le m} |J^j|$, the largest interval size. We always assume that as $n,m \to +\infty$, $r_{n,m} \to 0$, which implies that $n_0 \to +\infty$. For simplicity, let $m = m(n)$ and as $n \to +\infty$, $m \to +\infty$. Moreover, write

$$U_{n,m} = U_n, \quad V_{n,m} = V_n, \quad r_{n,m} = r_n.$$

To establish the moderate deviations for V_n, we introduce the following conditions.
(C1) There exist a sequence of positive numbers $\{v_n, n \ge 1\}$ and some constant $\Sigma \in (0,+\infty)$ such that, as $n \to +\infty$, $v_n \to 0$ and

$$v_n^{-1}\left(\sum_{i=1}^{\hat n}\sum_{j=1}^{m} \nu_1(\hat I^i)\nu_2(J^j)K^{\hat I}_{ij} + \sum_{i=1}^{\hat n}\nu^2(I^i) + \sum_{j=1}^{m}\nu^2(J^j) - \sum_{i=1}^{\hat n}\sum_{j=1}^{m}\nu^2(I^i\cap J^j)\right) \to \Sigma.$$

(C2) There exist a sequence of positive numbers $\{b_n, n \ge 1\}$ such that, as $n \to +\infty$,

$$b_n \to +\infty, \quad b_n v_n^{1/2} \to 0 \quad \text{and} \quad b_n v_n^{-1/2}\nu_n \to 0,$$

where $\nu_n = \max_{\ell=1,2}\max_{1\le i\le n}\max_{1\le j\le m}\left(\nu_\ell(I^i) \vee \nu_\ell(J^j)\right)$.

Theorem 4.2.1 *Under conditions (C1) and (C2),*

$$\left\{\frac{1}{b_n v_n^{1/2}}\left(V_n - \int_0^T \sigma_{1,t}\sigma_{2,t}\rho_t dt\right), n \ge 1\right\}$$

satisfies the large deviations with speed b_n^2 and rate function $L(x) = \frac{x^2}{2\Sigma}$. □

Remark 4.2.1 *By the construction of partition $(\hat I^i)_{i=1,\ldots,\hat n}$, we obtain immediately that*

$$\sum_{i=1}^{\hat n}\sum_{j=1}^{m} \nu_1(\hat I^i)\nu_2(J^j)K^{\hat I}_{ij} + \sum_{i=1}^{\hat n}\nu^2(\hat I^i) + \sum_{j=1}^{m}\nu^2(J^j) - \sum_{i=1}^{\hat n}\sum_{j=1}^{m}\nu^2(\hat I^i\cap J^j)$$
$$= \sum_{i=1}^{n}\sum_{j=1}^{m} \nu_1(I^i)\nu_2(J^j)K^{I}_{ij} + \sum_{i=1}^{n}\nu^2(I^i) + \sum_{j=1}^{m}\nu^2(J^j) - \sum_{i=1}^{n}\sum_{j=1}^{m}\nu^2(I^i\cap J^j).$$

Remark 4.2.2 *Similar to $(\hat I^i)_{i=1,\ldots,\hat n}$, a reduced design with respect to $(J^i)_{i=1,\ldots,m}$ can also be constructed as follows. We collect all J^js such that $J^j \subset I^i$ and combine them into a new interval; if such J^j does not exist, do nothing. Then* *collecting all such intervals and re-labeling them from left to right yields a new partition of* $(0, T]$, *denoted* $(\hat{J}^j)_{j=1,\ldots,\tilde{n}}$. *Then* V_n *has the formula,*

$$V_n = \sum_{j=1}^{\tilde{n}} \Delta X_1^0 \left(\cup_{i \in \hat{I}(j)} I^i \right) \Delta X_2^0(\hat{J}^j), \tag{4.10}$$

where $\hat{I}(j) := \left\{ 1 \leq i \leq n : K_{ij}^{\hat{J}} \neq \emptyset \right\}$ *with* $K_{ij}^{\hat{J}} = I_{\{I^i \cap \hat{J}^j \neq 0\}}$.

Letting $\tilde{n}_0 = \sup\{k : \varsigma_k > 0\}$, *then* $\tilde{n}_0 = n_0 - 1$ *and*

$$n_0 \leq \tilde{n} \leq n_0 + \sum_{k=1}^{n_0} I_{\{\varsigma_k - \varsigma_{k-1} > 1\}} + 1 \leq 2n_0 + 1.$$

Therefore, we can conclude that the moderate deviations of V_n *are independent of the different partitions* $(\bar{I}^i)_{i=1,\ldots,\tilde{n}}$ *and* $(\hat{J}^j)_{j=1,\ldots,\tilde{n}}$ *in (4.7) and (4.10).*

Now we turn to the moderate deviatons for U_n. We need the following three additional conditions.

(C3) For $\ell = 1, 2$, $b_\ell(\cdot, \cdot) \in L^\infty(dt \otimes dP)$

(C4) As $n \to +\infty$, $\frac{r_n^2}{v_n} \to 0$.

Theorem 4.2.2 *Under conditions (C1) through (C4),*

$$\left\{ \frac{1}{b_n v_n^{1/2}} \left(U_n - \int_0^T \sigma_{1,t}\sigma_{2,t}\rho_t dt \right), n \geq 1 \right\}$$

satisfies the large deviations with speed b_n^2 *and rate function* $L(x) = \frac{x^2}{2\Sigma}$. □

Remark 4.2.3 *Suppose synchronous and equidistant sampling as in (3) of Remark 4.2.1. We have* $r_n = 1/n$, *and condition (C4) holds automatically.*

Examples

Perfectly synchronous sampling Suppose synchronous and equidistant sampling, i.e. $I^i \equiv J^i$ and $|I^i| \equiv T/n$. Then, we can see that

$$n_0 = n, \quad v_n = 1/n, \quad \Sigma = \int_0^T \sigma_{1,t}^2 \sigma_{2,t}^2 \left(1 + \rho_t^2\right) dt.$$

Moreover, if $\sigma_{\ell,\cdot} \in L^\infty(dt)$ for $\ell = 1, 2$, condition (C2) equivalents to

$$b_n \to +\infty, \quad b_n n^{-1/2} \to 0, \quad n \to +\infty. \tag{4.11}$$

Under (4.11), from our Theorem 4.2.1, it follows that

$$\left\{ \frac{n^{1/2}}{b_n} \left(V_n - \int_0^T \sigma_{1,t}\sigma_{2,t}\rho_t dt \right), n \geq 1 \right\} \tag{4.12}$$

satisfies the large deviations with speed b_n^2 and rate function

$$L_0(x) = \frac{x^2}{2\int_0^T \sigma_{1,t}^2\sigma_{2,t}^2\left(1+\rho_t^2\right)dt}.$$

This result can also be obtained by Theorem 2.5 in Djellout et al. ([58]).

Nonsynchrnous alternating sampling at odd/even times We now consider the following deterministic, regularly spaced sampling scheme. X_1 is sampled at 'odd' times, i.e., $t = \frac{2k-1}{n}T, k = 1, 2, \cdots, n$ while X_2 is at 'even' times, $t = \frac{2k}{2n}T$. Hence X_1 and X_2 are sampled in a nonsynchronous, alternating way. so we have that the sequence given in (4.12) satisfies the LDP with speed b_n^2 and rate function with

$$\Sigma = \int_0^1 (\sigma_{1,t}\sigma_{2,t})^2(2 + \frac{3}{2}\rho_t^2)dt.$$

4.2.2 Moderate deviations for bipower variation of general function

Proposition 4.2.3 *Let (b_n) be a sequence of positive numbers such that*

$$b_n \to \infty, \quad b_n/\sqrt{n} \to 0, \quad \text{and} \quad \nu_n b_n \sqrt{n} \longrightarrow 0, \tag{4.13}$$

where $\nu_n = \max_{1\leq i\leq n} \int_{I^i} \Sigma_{\sigma_{\ell,s}}(g)\Sigma_{\sigma_{\ell,s}}(h)ds$. Assume that g and h are two evens, continuous differentiable with at most polynomial growth such that

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log n \max_{j=1}^{n} \mathbb{P}\Big(\Big| g\left(\sqrt{n}\Delta X_\ell(I^j)\right) h\left(\sqrt{n}\Delta X_\ell(I^{j+1})\right) - \mathbb{E}\Big(g(\sqrt{n}\Delta X_\ell(I^i))h(\sqrt{n}\Delta X_\ell(I^{i+1}))\Big)\Big| > nb_n\Big) = -\infty. \tag{4.14}$$

So, we have that

$$\frac{\sqrt{n}}{b_n}\left(V_{\ell,1}^n(g,h) - V_{\ell,1}(g,h)\right)$$

satisfies the large deviations principle with speed b_n^2 and rate function $I_{g,h}^{\ell}(x) = x^2/2\Sigma_\ell(g,h)$.

Examples

Realised Bipower Variation. Realised bipower variation, which is probably the most important subclass of our model, corresponds to the functions $g(x) = |x|^r$ and $h(x) = |x|^q$. In this case and under the condition that

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log n \max_{j=1}^{n} \mathbb{P}\left(n^{\frac{r+q}{2}} |\Delta X_\ell(I^j)|^r |\Delta X_\ell(I^{j+1})|^q > b_n\right) = -\infty. \quad (4.15)$$

we have the moderate deviation principle for

$$\frac{\sqrt{n}}{b_n}\left(V^n_{\ell,1}(r,q) - \mu_r\mu_q \int_0^1 |\sigma_{\ell,s}|^{r+q} ds\right)$$

$$\Sigma_\ell(g,h) := \Sigma_\ell(r,q) = (\mu_{2r}\mu_{2q} + 2\mu_r\mu_q\mu_{r+q} - 3\mu_r^2\mu_q^2) \int_0^1 |\sigma_{\ell,u}|^{2(r+q)} du.$$

where $\mu_r = \mathbb{E}|z|^r$ with $z \sim \mathcal{N}(0,1)$.

If $r = 2$ and $q = 0$ we are in the quadratic case and the condition (4.15) is satisfaied for all the sequence (b_n). The result was already obtained in Djellout-Guillin-Wu [59].

If $r + q < 2$ the condition (4.15) is satisfied for every sequence b_n.

The Cubic Power of Returns. The Cubic Power of Returns is a special case of the generalised bipower variation, corresponds to the functions $g(x) = |x|^3$ and $h(x) = 1$.

Then, under the condition

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log n \max_{j=1}^{n} \mathbb{P}\left(n^{\frac{3}{2}} |\Delta X_\ell(I^j)|^3 > b_n\right) = -\infty,$$

the sequence $\frac{\sqrt{n}}{b_n}\left(V^n_{\ell,1}(|x|^3, 1)\right)$ satisfies the large deviations principle with speed b_n^2 and rate function $I_{|x|^3}(u) = 24x^2 / \int_0^1 \sigma_s^6 ds.$.

4.3 proofs

In this section, we will use the appendix's Proposition 4.4.1 about the the moderate deviations for m-dependent random variables sequences which are not necessarily stationnary under the Chen-Ledoux type condition.

Proof of Theorem 4.2.1

We start this section by calculating the variance of V_n. Different from the method used in Hayashi and Yoshida ([80]), our approach relies on (4.7), the 2-dependent representation of U_n. Moreover, the analysis of moderate deviations will benefit from this straightforward thoughts.

Lemma 4.3.1 *For the estimator V_n defined by (4.9), we have*

$$Var(V_n) = \sum_{i=1}^{\hat{n}}\sum_{j=1}^{m} \nu_1(\hat{I}^i)\nu_2(J^j)K_{ij}^{\hat{I}} + \sum_{i=1}^{\hat{n}} \nu^2(\hat{I}^i) + \sum_{j=1}^{m} \nu^2(J^j) - \sum_{i=1}^{\hat{n}}\sum_{j=1}^{m} \nu^2\left(\hat{I}^i \cap J^j\right).$$

Proof 4.3.2 *For $\ell = 1, 2$, letting*

$$M_{\ell,t} = \int_0^t \sigma_{\ell,s} dW_{\ell,s}, \quad t \in [0,T],$$

then

$$V_n = \sum_{i=1}^{\hat{n}} \Delta M_1(\hat{I}^i)\Delta M_2\left(\cup_{j\in\hat{J}(i)} J^j\right)$$

Since

$$\left\{\Delta M_1(\hat{I}^i)\Delta M_2\left(\cup_{j\in\hat{J}(i)} J^j\right) : 1 \le i \le \hat{n}\right\}$$

are 2-dependent, one can write that

$$\begin{aligned}
Var(V_n) = &\sum_{\hat{i}=1}^{\hat{n}} Var\left(\Delta M_1(\hat{I}^i)\Delta M_2\left(\cup_{j\in\hat{J}(i)} J^j\right)\right) \\
&+ \sum_{i=1}^{\hat{n}} Cov\left(\Delta M_1(\hat{I}^i)\Delta M_2\left(\cup_{j\in\hat{J}(i)} J^j\right), \Delta M_1(\hat{I}^{i+1})\Delta M_2\left(\cup_{j\in\hat{J}(i+1)} J^j\right)\right) \\
&+ \sum_{i=1}^{\hat{n}} Cov\left(\Delta M_1(\hat{I}^i)\Delta M_2\left(\cup_{j\in\hat{J}(i)} J^j\right), \Delta M_1(\hat{I}^{i+2})\Delta M_2\left(\cup_{j\in\hat{J}(i+2)} J^j\right)\right) \\
&+ \sum_{i=1}^{\hat{n}} Cov\left(\Delta M_1(\hat{I}^i)\Delta M_2\left(\cup_{j\in\hat{J}(i)} J^j\right), \Delta M_1(\hat{I}^{i-1})\Delta M_2\left(\cup_{j\in\hat{J}(i-1)} J^j\right)\right) \\
&+ \sum_{i=1}^{\hat{n}} Cov\left(\Delta M_1(\hat{I}^i)\Delta M_2\left(\cup_{j\in\hat{J}(i)} J^j\right), \Delta M_1(\hat{I}^{i-2})\Delta M_2\left(\cup_{j\in\hat{J}(i-2)} J^j\right)\right) \\
:= &D_1 + D_2 + D_3 + D_4 + D_5,
\end{aligned}$$

where $\hat{I}^i = \emptyset$ for $i < 1$ or $i > \hat{n}$.

We will only consider D_1 and D_2, the other terms can be dealt with in the same way. A simple calculation gives us

$$\begin{aligned}
D_1 &= \sum_{i=1}^{\hat{n}} \left(\nu_1(\hat{I}^i)\nu_2\left(\cup_{j\in\hat{J}(i)} J^j\right) + \nu^2(\hat{I}^i)\right) \\
&= \sum_{i=1}^{\hat{n}}\sum_{j=1}^{m} \nu_1(\hat{I}^i)\nu_2(J^j)K_{ij}^{\hat{I}} + \sum_{\hat{i}=1}^{\hat{n}} \nu^2(\hat{I}^i).
\end{aligned} \tag{4.16}$$

Moreover, from the fact

$$\begin{aligned}
\Delta M_2\left(\cup_{j\in\hat{J}(i)} J^j\right) = &\Delta M_2\left(\hat{I}^i\right) + \Delta M_2\left(\cup_{j\in\hat{J}(i)} J^j \cap \hat{I}^{i+1}\right) + \Delta M_2\left(\cup_{j\in\hat{J}(i)} J^j \cap \hat{I}^{i+2}\right) \\
&+ \Delta M_2\left(\cup_{j\in\hat{J}(i)} J^j \cap \hat{I}^{i-1}\right) + \Delta M_2\left(\cup_{j\in\hat{J}(i)} J^j \cap \hat{I}^{i-2}\right),
\end{aligned}$$

it follows that

$$
\begin{aligned}
&E\left(\Delta M_1(\hat{I}^i)\Delta M_1(\hat{I}^{i+1})\Delta M_2\left(\cup_{j\in\hat{J}(i)}J^j\right)\Delta M_2\left(\cup_{j\in\hat{J}(i+1)}J^j\right)\right)\\
&=\nu\left(\hat{I}^i\right)\nu\left(\hat{I}^{i+1}\right)+\nu\left(\cup_{j\in\hat{J}(i)}J^j\cap\hat{I}^{i+1}\right)\nu\left(\cup_{j\in\hat{J}(i+1)}J^j\cap\hat{I}^i\right)\\
&=\nu\left(\hat{I}^i\right)\nu\left(\hat{I}^{i+1}\right)+\sum_{j_1=1}^{m}\sum_{j_2=1}^{m}\nu\left(J^{j_1}\cap\hat{I}^{i+1}\right)\nu\left(J^{j_2}\cap\hat{I}^i\right)K^{\hat{I}}_{ij_1}K^{\hat{I}}_{(i+1)j_2}.
\end{aligned}
$$

We can write the second term in the above equality as

$$
\begin{aligned}
&\nu\left(J^{j_1}\cap\hat{I}^{i+1}\right)\nu\left(J^{j_2}\cap\hat{I}^i\right)K^{\hat{I}}_{ij_1}K^{\hat{I}}_{(i+1)j_2}\\
&=\nu\left(J^{j_1}\cap\hat{I}^{i+1}\right)\nu\left(J^{j_2}\cap\hat{I}^i\right)K^{\hat{I}}_{ij_1}K^{\hat{I}}_{(i+1)j_2}K^{\hat{I}}_{(i+1)j_1}K^{\hat{I}}_{ij_2}.
\end{aligned}
$$

By the following fact,

$$K^{\hat{I}}_{ij_1}K^{\hat{I}}_{(i+1)j_2}K^{\hat{I}}_{(i+1)j_1}K^{\hat{I}}_{ij_2}\neq 0\Leftrightarrow j_1=j_2,$$

we can obtain immediately that

$$
\begin{aligned}
&E\left(\Delta M_1(\hat{I}^i)\Delta M_1(\hat{I}^{i+1})\Delta M_2\left(\cup_{j\in\hat{J}(i)}J^j\right)\Delta M_2\left(\cup_{j\in\hat{J}(i+1)}J^j\right)\right)\\
&=\nu\left(\hat{I}^i\right)\nu\left(\hat{I}^{i+1}\right)+\sum_{j=1}^{m}\nu\left(\hat{I}^{i+1}\cap J^j\right)\nu\left(\hat{I}^i\cap J^j\right),
\end{aligned}
$$

which implies that

$$D_2=\sum_{i=1}^{\hat{n}}\sum_{j=1}^{m}\nu\left(\hat{I}^{i+1}\cap J^j\right)\nu\left(\hat{I}^i\cap J^j\right). \tag{4.17}$$

Applying the same arguments to D_3, D_4, D_5,

$$D_3=\sum_{i=1}^{\hat{n}}\sum_{j=1}^{m}\nu\left(\hat{I}^{i+2}\cap J^j\right)\nu\left(\hat{I}^i\cap J^j\right) \tag{4.18}$$

$$D_4=\sum_{i=1}^{\hat{n}}\sum_{j=1}^{m}\nu\left(\hat{I}^{i-1}\cap J^j\right)\nu\left(\hat{I}^i\cap J^j\right) \tag{4.19}$$

$$D_5=\sum_{i=1}^{\hat{n}}\sum_{j=1}^{m}\nu\left(\hat{I}^{i-2}\cap J^j\right)\nu\left(\hat{I}^i\cap J^j\right) \tag{4.20}$$

Putting (4.17) through (4.20) together, and noting fact that if $K^{\hat{I}}_{ij}\neq 0$, *then*

$$J^j=\cup_{\ell=-2}^{2}\left(I^{i+\ell}\cap J^j\right),$$

we have

$$\begin{aligned}
&D_2 + D_3 + D_4 + D_5 \\
&= \sum_{i=1}^{\hat{n}} \sum_{j=1}^{m} \left(\nu(J^j) - \nu\left(\hat{I}^i \cap J^j\right) \right) \nu\left(\hat{I}^i \cap J^j\right) \\
&= \sum_{j=1}^{m} \nu^2(J^j) - \sum_{i=1}^{\hat{n}} \sum_{j=1}^{m} \nu^2\left(\hat{I}^i \cap J^j\right)
\end{aligned}$$

Together with (4.16), we can conclude

$$Var(V_n) = \sum_{i=1}^{\hat{n}} \sum_{j=1}^{m} \nu_1(\hat{I}^i)\nu_2(J^j)K_{ij}^{\hat{I}} + \sum_{i=1}^{\hat{n}} \nu^2(\hat{I}^i) + \sum_{j=1}^{m} \nu^2(J^j) - \sum_{i=1}^{\hat{n}} \sum_{j=1}^{m} \nu^2\left(\hat{I}^i \cap J^j\right).$$

Remark 4.3.1 *From the proof of Lemma 4.3.1, each* $\Delta M_1(\hat{I}^i)\Delta M_2\left(\cup_{j\in\hat{J}(i)} J^j\right)$, $1 \le i \le \hat{n}$ *contributes to the variance of* V_n *with*

$$\sum_{j=1}^{m} \nu_1(\hat{I}^i)\nu_2(J^j)K_{ij}^{\hat{I}} + \nu^2(\hat{I}^i) + \sum_{j=1}^{m} \left(\nu(J^j) - \nu\left(\hat{I}^i \cap J^j\right) \right) \nu\left(\hat{I}^i \cap J^j\right). \quad \square$$

Proof of Theorem 4.2.1 For the expectation of V_n, we have

$$EV_n = \sum_{i=1}^{\hat{n}} \nu\left(\hat{I}^i\right) = \int_0^T \sigma_{1,t}\sigma_{2,t}\rho_t dt.$$

Therefore,

$$V_n - \int_0^T \sigma_{1,t}\sigma_{2,t}\rho_t dt = \sum_{i=1}^{\hat{n}} \left(\Delta M_1(\hat{I}^i)\Delta M_2\left(\cup_{j\in\hat{J}(i)} J^j\right) - \nu\left(\hat{I}^i\right) \right) := \sum_{i=1}^{\hat{n}} Y_i, \tag{4.21}$$

where $Y_i = \Delta M_1(\hat{I}^i)\Delta M_2\left(\cup_{j\in\hat{J}(i)} J^j\right) - \nu\left(\hat{I}^i\right)$. Consequently, the random variable sequence $\{Y_i : 1 \le i \le \hat{n}\}$ are 2-dependent. We have to prove that

$$\limsup_{n\to\infty} \frac{1}{b_n^2} \log\left(\hat{n} \max_{i=1}^{\hat{n}} \mathbb{P}(|Y_i| > b_n v_n^{1/2}) \right) = -\infty. \tag{4.22}$$

By the condition (C4), there exist a sequence $\varepsilon_n > 0$ satisfying as $n \to +\infty$, $\frac{\varepsilon_n^2}{v_n} \to 0$.
Now, we can write that

$$\begin{aligned}
|Y_i| &= |\Delta M_1(\hat{I}^i)\Delta M_2\left(\cup_{j\in\hat{J}(i)} J^j\right) - \nu\left(\hat{I}^i\right)| \\
&\le \varepsilon_n |\Delta M_1(\hat{I}^i)|^2 + \frac{3}{\varepsilon_n} |\Delta M_2(\hat{I}^i)|^2 + |\nu\left(\hat{I}^i\right)|
\end{aligned}$$

Using condition (C2), we have as $n \to \infty$

$$b_n v_n 1/2 |\nu\left(\hat{I}^i\right)| \to 0.$$

Hence, to prove (4.22), we have to whow that

$$\limsup_{n\to\infty} \frac{1}{b_n^2} \log\left(\hat{n} \max_{i=1}^{\hat{n}} \mathbb{P}(\varepsilon_n |\Delta M_1(\hat{I}^i)|^2 > b_n v_n^{1/2})\right) = -\infty. \tag{4.23}$$

and

$$\limsup_{n\to\infty} \frac{1}{b_n^2} \log\left(\hat{n} \max_{i=1}^{\hat{n}} \mathbb{P}(\frac{3}{\varepsilon_n} |\Delta M_2(\hat{I}^i)|^2 > b_n v_n^{1/2})\right) = -\infty. \tag{4.24}$$

In fact, by virtue of Chebyshev's inequality, for (4.23),

$$\begin{aligned}
&\mathbb{P}(\frac{3}{\varepsilon_n}|\Delta M_2(\hat{I}^i)|^2 > b_n v_n^{1/2}) \\
&\le \inf_{\lambda>0} \exp\left\{-\lambda + \frac{\lambda^2 \varepsilon_n^2 K}{2 b_n^2 v_n}\right\} \\
&\le \exp\left\{-\frac{b_n^2 v_n}{2\varepsilon_n^2 K}\right\},
\end{aligned}$$

where $K = \int_0^T \sigma_{1,t}^4 dt + (\int_0^T \sigma_{1,t}^2 dt)^2$. From our choice of ε_n, (4.23) follows. Therefore, (4.24) can be established similarly.
Hence, from lemma 4.3.1 and (4.22) together with appendix Proposition 4.4.1, the Theorem 4.2.1 holds. □
Proof of Theorem 4.2.2 From Theorem 4.2.1, we only need to show the exponential equivalence of U_n and V_n, i.e. for any $\delta > 0$

$$\lim_{n\to+\infty} \frac{1}{b_n^2} \log P\left(\frac{1}{b_n v_n^{1/2}} |U_n - V_n| \ge \delta\right) = -\infty. \tag{4.25}$$

In fact, by simple calculation, we can write that

$$
\begin{aligned}
&|U_n - V_n| \\
&= \sum_{i=1}^{\hat{n}} \left|\Delta A_1(\hat{I}_i)\Delta A_2(\cup_{j\in\hat{J}(i)} J^j)\right| + \sum_{i=1}^{\hat{n}} \left|\Delta A_1(\hat{I}_i)\Delta M_2(\cup_{j\in\hat{J}(i)} J^j)\right| \\
&\quad + \sum_{i=1}^{\hat{n}} \left|\Delta M_1(\hat{I}_i)\Delta A_2(\cup_{j\in\hat{J}(i)} J^j)\right| \\
&\le 3Tr_n \max_{\ell=1,2} \|b_\ell\|_\infty + r_n\|b_1\|_\infty \sum_{i=1}^{\hat{n}} \left|\Delta M_2(\cup_{j\in\hat{J}(i)} J^j)\right| \\
&\quad + 3r_n\|b_2\|_\infty \sum_{i=1}^{\hat{n}} \left|\Delta M_1(\hat{I}_i)\right| \\
&\le 3Tr_n \max_{\ell=1,2} \|b_\ell\|_\infty + 3r_n\|b_1\|_\infty \sum_{i=1}^{\hat{n}} \left|\Delta M_2(\hat{I}_i)\right| + 3r_n\|b_2\|_\infty \sum_{i=1}^{\hat{n}} \left|\Delta M_1(\hat{I}_i)\right|
\end{aligned}
$$

Using condition (C4), we have as $n \to \infty$

$$
\frac{3Tr_n \max_{\ell=1,2} \|b_\ell\|_\infty}{b_n v_n^{1/2}} \to 0.
$$

Therefore, to prove (4.25), we only need to show that for any $\delta > 0$

$$
\lim_{n\to+\infty} \frac{1}{b_n^2} \log P\left(\frac{r_n}{b_n v_n^{1/2}} \sum_{i=1}^{\hat{n}} \left|\Delta M_1(\hat{I}_i)\right| \ge \delta\right) = -\infty \tag{4.26}
$$

and

$$
\lim_{n\to+\infty} \frac{1}{b_n^2} \log P\left(\frac{r_n}{b_n v_n^{1/2}} \sum_{i=1}^{\hat{n}} \left|\Delta M_2(\hat{I}_i)\right| \ge \delta\right) = -\infty. \tag{4.27}
$$

In fact, according to Chebyshev's inequality, for (4.26),

$$
\begin{aligned}
&P\left(\frac{r_n}{b_n v_n^{1/2}} \sum_{i=1}^{\hat{n}} \left|\Delta M_1(\hat{I}_i)\right| \ge \delta\right) \\
&\le \inf_{\lambda>0} \exp\left\{-\lambda\delta + \frac{\lambda^2 r_n^2 \int_0^T \sigma_{1,t}^2 dt}{2b_n^2 v_n}\right\} \\
&\le \exp\left\{-\frac{b_n^2 v_n \delta^2}{2r_n^2 \int_0^T \sigma_{1,t}^2 dt}\right\},
\end{aligned}
$$

which implies that (4.26) by condition (C4). Moreover, (4.27) can be established similarly. □

Proof of Proposition 4.2.3 Our proof is an application of Proposition 4.4.1. In fact, by a straightforward calculation, one can see that

$$\begin{aligned} V^n_{\ell,1}(g,h) - V_{\ell,1}(g,h) &= \frac{1}{n}\sum_{i=1}^n g\left(\sqrt{n}\Delta X_\ell(I^i)\right) h\left(\sqrt{n}\Delta X_\ell(I^{i+1})\right) - \int_0^1 \Sigma_{\sigma_{l,u}}(g)\Sigma_{\sigma_{l,u}}(h)\mathrm{d}u \\ &= \frac{1}{n}\sum_{i=1}^n \Bigg[g\left(\sqrt{n}\Delta X_\ell(I^i)\right) h\left(\sqrt{n}\Delta X_\ell(I^{i+1})\right) - \mathbb{E}\bigg(g(\sqrt{n}\Delta X_\ell(I^i)) h(\sqrt{n}\Delta X_\ell(I^{i+1}))\bigg)\Bigg] \\ &= \sum_{i=1}^n Y_i \end{aligned}$$

where $Y_i = \frac{1}{n}\Bigg[g\left(\sqrt{n}\Delta X_\ell(I^i)\right) h\left(\sqrt{n}\Delta X_\ell(I^{i+1})\right) - \mathbb{E}\bigg(g(\sqrt{n}\Delta X_\ell(I^i)) h(\sqrt{n}\Delta X_\ell(I^{i+1}))\bigg)\Bigg]$ are 1-dependent random variables.
Hence, by the 1−dependence of (Y_i) and from (4.14) together the Proposition 4.4.1, the Proposition 4.2.3 holds. □

4.4 Appendix A

Let us begin with some few bibliographical notes on the MDP. Borovkov and Mogulskii ([30], [31]) considered the MDP for Banach valued i.i.d.r.v. sequences $(Z_n)_{n\geq 1}$. Under the condition that $\mathbb{E}e^{\delta|Z_1|} < +\infty$, for some $\delta > 0$, they proved the MDP for $\sum_{i=1}^n Z_i$.

For $b_n = n^\alpha$ with $\frac{1}{2} < \alpha < 1$, Chen [38] found the necessary and sufficient condition for the MDP in a Banach space, and he obtained the lower bound for general b_n under very weak condition. Using the isoperimetry techniques, Ledoux [90](see also [67]) obtained the necessary and sufficient condition for the general sequence b_n satisfying :

$$\frac{b_n}{\sqrt{n}} \longrightarrow \infty \quad \text{and} \quad \frac{b_n}{n} \longrightarrow 0. \tag{4.28}$$

The results of Ledoux [90] is extended to functional empirical processes (in the setting of non parametrical statistics) by Wu [93]. The further developments are given by Dembo and Zeitouni [53]. Arcones [5] obtains the MDP of functional type with the following condition

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log\bigg(n\mathbb{P}(|Z_1| > b_n)\bigg) = -\infty. \tag{4.29}$$

How to extend the MDP to the dependent situations has recently attracted much attention and remarkable works. The Markovian case has been studied under successively less restrictive conditions (see Wu [93], Chen [38] for the relevant references) and very recently under weak conditions by de Acosta [48] and Chen [38] for the lower bound (under different and non comparable conditions) and by de Acosta-Chen [49] and Chen [38] (under the same condition but different proof) for the upper bound. Guillin [77] obtained uniform (in time) MDP for functional empirical processes. Using regeneration split chain method, Djellout and Guillin ([55], [54]) extend the characterization of MDP for i.i.d.r.v. case of Ledoux [90] to Markov chains. The geometric ergodicity is substituted by a tail on the first time of return to the atom. Their conditions are weaker than Theorem 1 in [49], which allow them to obtain MDP for empirical measures and functional empirical processes.

For the studies of the MDP of martingale see Puhalskii [114], Dembo [50], Gao [74], Worms [136] and Djellout [56].

Those works motivate directly the studies here. Our main aim is to prove the Chen-Ledoux type theorem for the MDP of a sequence of m-dependent random variables.

Moderate deviations for m-dependent randon variables

Let (b_n) a sequence of positive real numbers satisfying (4.28).

Let X_n be an m-dependent random variables, namely, for every $k \geq 1$ the following two collections

$$\{X_1, \cdots, X_k\} \qquad \text{and} \qquad \{X_{k+m+1}, X_{k+m+2}, \cdots\}$$

are independent. In this terminology, an independent identically distributed (i.i.d.) random variables sequence is 0- dependent and, for any non-negative intergers $m_1 < m_2$, m_1 dependence implies m_2-dependence. The aim of this section is to consider the moderate deviations for m dependent random variables sequences which are not necessarily stationnary under the Chen-Ledoux type condition. This condition links the speed of the moderate deviation principle to the queue of the distribution of the random variable.

Proposition 4.4.1 *Let (X_k) be a sequence of m-dependent random variables with zero means such that*

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log\left(n \max_{i=1}^{n} \mathbb{P}(|X_i| > b_n)\right) = -\infty. \tag{4.30}$$

Write $S_n = \sum_{j=1}^n X_j$ and let (b_n) a sequence of positive numbers satisfying (4.28). Then S_n/b_n satisfies the moderate deviation principle on $\mathbb{R}$ with speed b_n^2/n and rate fonction $I(x) = x^2/2\sigma^2$, where $\sigma^2 := \lim_{n\to\infty} \frac{1}{n} Var(S_n) < \infty$.

Remark 4.4.1 *Remark that the condition (4.30) implies that*

$$\lim_{n\to\infty} \sup_{1\leq \ell \leq n} \mathbb{E}(X_\ell^2) < \infty.$$

Remark 4.4.2 *If there exists some $\delta > 0$ such that*

$$\sup_{1 \le \ell \le n} \mathbb{E}(e^{\delta |X_\ell|}) < \infty,$$

so the condition (4.30) is satisfied for every sequence b_n.

Remark 4.4.3 *If $b_n = n^{1/p}$ with $1 < p < 2$ so the condition (4.30) is equivalent to*

$$\sup_{1 \le \ell \le n} \mathbb{E}(e^{\delta |X_\ell|^{2-p}}) < \infty.$$

Lemma 4.4.2 *Let (Y_k) be a sequence of independent random variables with zero means such that the condition (4.30) is satisfied. Write $S'_n = \sum_{j=1}^n Y_j$ and let (b_n) a sequence of positive numbers satisfying (4.28). Then S'_n/b_n satisfies the moderate deviation principle on $\mathbb{R}$ with speed b_n^2/n and rate fonction $I(x) = x^2/2\sigma^2$, where $\sigma^2 := \lim_{n\to\infty} \frac{1}{n} Var(S_n) < \infty$.*

Proof of Lemma 4.4.2 The proof of this result is a little adaptation of the proof of the independent and identically distributed random variables given in [55]. □

Proof of Proposition 4.4.1 We will do the proof in the case $m = 1$ for simplicity.

Fix the integer $p > 1$ and write, for each $n \ge 1$,

$$n = k_n p + r_n$$

where k_n and r_n are non-negative integers with $0 \le r_n \le p - 1$. Define

$$Y_k = \sum_{(k-1)p<j<kp} X_j, \qquad k = 1, 2, \cdots$$

Then $(Y_k)_{k\ge1}$ is an independent random variables and

Notice that

$$S_n = \sum_{\ell=1}^{k_n} Y_\ell + \sum_{\ell=1}^{k_n} X_{\ell p} + \sum_{\ell=k_n p+1}^{n} X_j \qquad (n \ge 1).$$

We have to prove that

$$\lim_{p\to\infty} \limsup_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\left| \sum_{\ell=1}^{k_n} X_{\ell p} \right| > \delta b_n \right) = -\infty. \tag{4.31}$$

Since the random variables $(X_{\ell p})_{\ell\ge1}$ are independent, we use the moderate deviation principle for $(\frac{1}{b_n} \sum_{\ell=1}^{k_n} X_{\ell p})_{\ell\ge1}$, to obtain that for n large

$$\frac{n}{b_n^2} \log \mathbb{P}\left(\left| \sum_{\ell=1}^{k_n} X_{\ell p} \right| > \delta b_n \right) \le -\frac{\delta^2}{2\sigma_0^2}$$

where

$$\sigma_0^2 = \lim_{n\to\infty} \frac{1}{n} \sum_{\ell=1}^{k_n} \mathrm{Var}(X_{\ell p}) \leq C \lim_{n\to\infty} \frac{k_n}{n} = \frac{C}{p}.$$

So

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\left|\sum_{\ell=1}^{k_n} X_{\ell p}\right| > \delta b_n\right) \leq -\frac{p}{2}\frac{\delta^2}{C}$$

Letting p goes to infinty, we get (4.31).

Now we have to prove that

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{P}\left(\left|\sum_{\ell=k_n p+1}^{n} X_\ell\right| > \delta b_n\right) = -\infty. \tag{4.32}$$

If $r_n = 0$ there is no thing to prove, so we suppose that $r_n > 0$

$$\begin{aligned}
\mathbb{P}\left(\left|\sum_{\ell=k_n p+1}^{n} X_\ell\right| > \delta b_n\right) &\leq \sum_{\ell=k_n p+1}^{n} \mathbb{P}\left(|X_\ell| > \frac{\delta b_n}{r_n}\right) \\
&\leq r_n \max_{\ell=k_n p+1}^{n} \mathbb{P}\left(|X_\ell| > \delta \frac{b_n}{r_n}\right) \\
&\leq n \max_{\ell=1}^{n} \mathbb{P}\left(|X_\ell| > \delta \frac{b_n}{p}\right).
\end{aligned}$$

Taking the log and normalizing by n/b_n^2, we obtain (4.32).

Now, we have to prove that

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log \left[k_n \max_{\ell=1}^{k_n} \mathbb{P}\left(|Y_\ell| > b_n\right)\right] = -\infty. \tag{4.33}$$

We have

$$\begin{aligned}
k_n \max_{\ell=1}^{k_n} \mathbb{P}\left(|Y_\ell| > b_n\right) &\leq k_n \max_{\ell=1}^{k_n} \sum_{j=(\ell-1)p+1}^{\ell p-1} \mathbb{P}\left(|X_j| > \frac{1}{p} b_n\right) \\
&\leq k_n p \max_{\ell=1}^{k_n} \max_{j=(\ell-1)p+1}^{\ell p-1} \mathbb{P}\left(|X_j| > \frac{1}{p} b_n\right) \\
&\leq n \max_{\ell=1}^{n} \mathbb{P}\left(|X_j| > \frac{1}{p} b_n\right).
\end{aligned}$$

Taking the log and normalizing by n/b_n^2 and letting n goes to infinity, we obtain (4.33).

So we deduce that the sequence $1/b_n \sum_{\ell=1}^{k_n} Y_\ell$ satisfies the moderate deviations principle with speed b_n^2/n and rate function $I(x) = x^2/2\sigma_Y^2$, where

$$\sigma_Y^2 := \lim_{n\to\infty} \frac{1}{n}\mathrm{Var}\left(\sum_{\ell=1}^{k_n} Y_\ell\right). \tag{4.34}$$

Now we have to prove that $\sigma^2 = \sigma_Y^2$. For that we have

$$\begin{aligned}\Delta_n := \mathrm{Var}S_n - \mathrm{Var}\left(\sum_{\ell=1}^{k_n} Y_\ell\right) &= \sum_{\ell=1}^{k_n}\mathbb{E}(X_{\ell p}^2) + 2\sum_{\ell=1}^{k_n}(\mathbb{E}X_{\ell p}X_{\ell p+1} + \mathbb{E}X_{\ell p}X_{\ell p-1}) \\ &+ \sum_{\ell=pk_n+1}^{pk_n+r_n}\mathbb{E}(X_\ell^2 + X_\ell X_{\ell+1} + X_\ell X_{\ell-1}).\end{aligned}$$

Now since $\sup_{1\le\ell\le n}\mathbb{E}X_\ell^2 < \infty$, we have that $\Delta_n/n \le (k_nC + 4Ck_n + 3r_nC)/n$. So $\lim_{n\to\infty}\Delta_n/n = C/p$, which goes to 0 when p goes to infinity.$\square$

Chapitre 5

Moderate deviations for the Durbin-Watson statistic associated to the stable p-order autoregressive process

Yacouba SAMOURA

abstract :

In this paper, we consider the Moderate deviations for the Durbin-Watson statistic related to the stable $p-$order autoregressive process where the driven noise is also given by a $q-$order autoregressive process. At first, we establish the moderate deviation principle for the least squares estimators of the unknown vector parameter of the autoregressive process as well as for the serial correlation estimator related with the driven noise.
keywords : Moderate deviations, Durbin-Watson statistic, Serial correlation, $p-$order autoregressive process.

Résumé :

Dans ce chapitre, on s'intéresse aux déviations modérées de la statistique de Durbin-Watson associée à un processus autoregréssif stable d'ordre p dont le bruit correspondant est aussi donné par un processus autoregréssif d'ordre q. On commence tout d'abord par établir des principes de déviations modérées pour les estimateurs du vecteur inconnu des moindres carrés et de la suite de corrélation associé au bruit induit.
Mots clés : Déviations modérées, statistique de Durbin-Watson, processus autoregréssif stable d'ordre p.

5.1 Introduction

This paper is concentrated on the stable $p-$order autoregressive process where the driven noise is also given by a $q-$order autoregressive process. The aim of this paper is to establish the moderate deviation principle for the least squares estimators of the unknown vector parameter of the autoregressive process as well as for the serial correlation estimator related with the driven noise. In a lagged dependent random variables framework, our interest is to establish the moderate deviation for the Durbin-Watson statistic [64]-[65]-[66] which is inspired by the previous work of Von Neumann [133]. The work of Box-Pierce [33], of Ljung-Box [32], of Breusch-Godfrey [35]-[?] and the H-Testof Durbin [?] have also helped improve the study of the residual set from a least squares estimation with lagged dependent random variables in the 1970s. Tillman [127] examined the power function of the Durbin-Watson test for first order serial correlation and Blattberg [26] has shown that tests using the Durbin-Watson statistic are powerful against non-first order serial correlation alternatives.

Recently, the asymptotic behavior of the the least squares estimators and the Durbin-Watson statistic for the stable first-order autoregressive and the stable $p-$order autoregressive process where the driven noise is also given by a $q-$order autoregressive process, has been introduced by Bercu and Proïa [19] and Proïa [112]-[113]. In the paper of Valère & al.[25], is proved the moderate deviation of both the the least squares estimators and the Durbin-Watson statistic. For our main results, we shall assume that the driven noise is normally distributed on the one hand, and we will extend our results to the more general case where the drivennoise satisfies a less restrictive Chen-Ledoux type condition [90], [38] on the other hand. Our results are proved at first via an extensive use of the results of Dembo [50], Dembo and Zeitouni [53] and Worms [136]-[137]-[135], and the results of the paper of Puhalskii [116] and Djellout [56] to finish. We will focus our attention on the $p-$order autoregressive process given, for all $n \geqslant 1$, by

$$\left\{ \begin{array}{rcl} X_n &=& \theta' \Phi_{n-1}^p + \varepsilon_n \\ \varepsilon_n &=& \rho' \Psi_{n-1}^q + V_n \end{array} \right. \tag{5.1}$$

where

$$\Phi_n^p = (X_n \quad X_{n-1} \quad \cdots \quad X_{n-p+1})' \quad and \quad \Psi_n^q = (\varepsilon_n \quad \varepsilon_{n-1} \quad \cdots \quad \varepsilon_{n-q+1})'. \tag{5.2}$$

are two matrix of order $p \times 1$ and $q \times 1$ respectively for the couple of parameters $p \in \mathbb{N}$ and $q \in \mathbb{N}$ and where the unknown parameters $\theta = (\theta_1 \quad \theta_2 \quad \cdots \quad \theta_p)'$ and $\rho = (\rho_1 \quad \rho_2 \quad \cdots \quad \rho_q)'$ are the nonzero vectors. Without loss of generality, we assume that the square-integrable initial values $X_0 = \varepsilon_0$ and $X_{-1}, X_{-2}, \cdots, X_{-p} = 0$. In all the sequel, we assume that

(V_n) is a sequence of square-integrable, independent and identically distributed random variables with zero mean such that $\mathbb{E}(V_1^2) = \sigma^2 > 0$ and $\mathbb{E}(V_1^4) = \tau^4 < \infty$. We assume that the polynomials

$$\mathcal{A}(z) = 1 - \theta_1 z - \cdots - \theta_p z^p \quad and \quad \mathcal{B}(z) = 1 - \rho_1 z - \cdots - \rho_q z^q \tag{5.3}$$

are causal, that is $\mathcal{A}(z) \neq 0$ and $\mathcal{B}(z) \neq 0$ for all $z \in \mathbb{C}$ such that $|z| \leqslant 1$. Now, denote by W_n the matrix of order $p \times q$ given, for all $n \geqslant 1$, by

$$W_n = \sum_{k=1}^{n} (\Phi_{k-1}^p X_k \quad \Phi_{k-2}^p X_k \quad \cdots \quad \Phi_{k-q}^p X_k) \tag{5.4}$$

and S_n the positive definite matrix of order p given by

$$S_n = \sum_{k=0}^{n} \Phi_k^p \Phi_k^{p'} + S \tag{5.5}$$

where S is a symetric and positive definite matrix used in order to avoid an useless invertibility assumption. Let also

$$T_n = S_{n-1}^{-1} W_n. \tag{5.6}$$

We want to estimate the vector θ by the least squares estimator

$$\hat{\theta} = (S_{n-1})^{-1} \sum_{k=1}^{n} \Phi_{k-1}^p X_k. \tag{5.7}$$

Now we focus our attention to the particular case "q=1", we also introduce a set of least squares residuals given by

$$\hat{\varepsilon}_k = X_k - \hat{\theta}_n X_{k-1} \qquad for \quad all \quad 1 \leqslant k \leqslant n, \tag{5.8}$$

which leads to the estimator of ρ,

$$\hat{\rho}_n = \left(\sum_{k=1}^{n} \hat{\varepsilon}_{k-1}^2 \right)^{-1} \sum_{k=1}^{n} \hat{\varepsilon}_k \hat{\varepsilon}_{k-1}. \tag{5.9}$$

Now, we can define the Durbin-Watson statistic, given, for all $n \geqslant 1$, by

$$\hat{D}_n = \left(\sum_{k=0}^{n} \hat{\varepsilon}_k^2 \right)^{-1} \sum_{k=1}^{n} \left(\hat{\varepsilon}_k - \hat{\varepsilon}_{k-1} \right)^2. \tag{5.10}$$

5.2 Remarks and Notations

Throughout the paper, for any matrix M, M', $det(M)$, M^{-1} and $\|M\|$ are the transpose, the determinant, the inverse and the euclidean norm of M, respectively. For any vector v, $\|v\|$ stands the euclidean norm of v and

$\|v\|_1$ is the euclidean $1-$norm of v.

Furthermore, let us now recall some basic definitions of the large deviations theory (c.f [53]). Let $(a_n)_{n\geq 1}$ be a sequence of nonnegative real number such that $\lim_{n\to\infty} a_n = +\infty$. We say that a sequence of a random variables $(M_n)_n$ with topological state space $(S, \mathbb{S})$, where $\mathbb{S}$ is a $\sigma - algebra$ on S, satisfies a large deviation principle (LDP) with speed a_n and rate function $I : S \to [0, +\infty]$ if, for each $A \in \mathbb{S}$,

$$-\inf_{x\in A^o} I(x) \leq \liminf_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}\Big(M_n \in A\Big) \leq \limsup_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}\Big(M_n \in A\Big) \leq -\inf_{x\in \bar{A}} I(x)$$

where A^o and $\bar{A}$ denote the interior and the closure of A, respectively.

The rate function I is lower semicontinuous, *i.e.* all the sub-level sets $\{x \in S \mid I(x) \leq c\}$ are closed, for $c \geq 0$. If these level sets are compact, then I is said to be a good rate function. When the speed of the large deviation principle correspond to the regime between the central limit theorem and the law of large numbers, we talk of moderate deviation principle (MDP).

Moreover, we say that the sequence of random variables $(Z_n)_n$ on $\mathbb{R}^{d\times p}$ converges $a_n^2 - superexponentially$ fast in probability to some random variable Z if, for all $\delta > 0$,

$$\limsup_{n\to\infty} \frac{1}{a_n^2} \log \mathbb{P}\left(\|Z_n - Z\|\right) = -\infty.$$

This exponential convergence with speed a_n^2 will shorted as

$$Z_n \xrightarrow[a_n^2]{Superexp} Z$$

The definition about the exponential equivalence with speed a_n^2 between two sequences of random variables $(Y_n)_n$ and $(Z_n)_n$ which is given in Definition 4.2.10 of [53], will be shorted as

$$Y_n \underset{a_n^2}{\overset{Superexp}{\sim}} Z_n.$$

Remark 5.2.1 *Before starting, we denote by I_h be the identity matrix of order h, J_h be the exchange matrix of order h and e the $h-$dimensional vector given by*

$$I_h = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}, \; J_h = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 1 & \cdots & 0 & 0 \end{pmatrix}, \; e = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Let us introduce the real α

$$\alpha = \frac{1}{(1-\theta_p\rho)(1+\theta_p\rho)} \tag{5.11}$$

and the vectors β *of order* p *and* γ *of order* q *given by,*

$$\beta = (\beta_1 \quad \cdots \quad \beta_p)' \quad \text{and} \quad \gamma = (\beta_{p+1} \quad \cdots \quad \beta_{p+q})' \tag{5.12}$$

such that, for all $1 \leqslant k \leqslant p+q$

$$\beta_k = \rho_k - \theta_1\rho_{k-1} - \theta_2\rho_{k-2} - \cdots - \theta_{k-1}\rho_1 + \theta_k \tag{5.13}$$

The Toeplitz matrix C_β *and the Hankel matrix* C_γ *of order* $q \times q$ *are given by*

$$C_\beta = \begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ \beta_1 & \ddots & & & \vdots \\ \beta_2 & \beta_1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \beta_{q-1} & \cdots & \beta_2 & \beta_1 & 0 \end{pmatrix} \quad \text{and} \quad C_\gamma = \begin{pmatrix} \gamma_1 & \gamma_2 & \cdots & \cdots & \gamma_q \\ \gamma_2 & & & \gamma_q & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \vdots \\ \gamma_q & 0 & \cdots & \cdots & 0 \end{pmatrix} \tag{5.14}$$

Note that C_β *is only well-defined for the usual case* $p \geqslant q$ *and is generated by* $(0, \beta_1 \quad \cdots \quad \beta_{q-1})$. *If* $p < q$, *it shall be replaced by Toeplitz matrix* G_β *of order* $q \times q$ *generated by*
$(0, \beta_1 \quad \cdots \quad \beta_p, \gamma_1 \quad \cdots \quad \gamma_{q-p-1})$. *We denote by* D *the Hankel matrix of order* $p \times q$ *which is well-defined only for* $p \geqslant q$ *and given by,*

$$D = \begin{pmatrix} \beta_1 & \beta_2 & \cdots & \cdots & \beta_q \\ \beta_2 & & & \ddots & \vdots \\ \vdots & & \ddots & & \beta_p \\ \vdots & \ddots & & \ddots & \gamma_1 \\ \beta_q & & \ddots & \ddots & \gamma_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \beta_p & \gamma_1 & \gamma_2 & \cdots & \gamma_{q-1} \end{pmatrix}$$

whereas it shall be replaced for $p < q$ *by the Hankel matrix* E *of order* $p \times q$ *given by,*

$$E = \begin{pmatrix} \beta_1 & \cdots & \cdots & \beta_p & \gamma_1 & \cdots & \gamma_{q-p} \\ \vdots & & \ddots & \ddots & & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & & \vdots \\ \beta_p & \gamma_1 & \cdots & \gamma_{q-p} & \cdots & \cdots & \gamma_{q-1} \end{pmatrix}$$

Let us now introduce the square matrix B of order $p+q+1$, partial made of the elements β_k given by (5.13)*,*

$$B = \begin{pmatrix} 1 & -\beta_1 & -\beta_2 & \cdots & \cdots & -\beta_{p+q-2} & -\beta_{p+q-1} & -\beta_{p+q} \\ -\beta_1 & 1-\beta_2 & -\beta_3 & \cdots & \cdots & -\beta_{p+q-1} & -\beta_{p+q} & 0 \\ -\beta_2 & -\beta_1-\beta_3 & 1-\beta_4 & \cdots & \cdots & -\beta_{p+q} & 0 & 0 \\ \vdots & \vdots & \vdots & & & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & & & \vdots & \vdots & \vdots \\ -\beta_{p+q-1} & -\beta_{p+q-2} & -\beta_{p+q-3} & \cdots & \cdots & -\beta_1 & 1 & 0 \\ -\beta_{p+q} & -\beta_{p+q-1} & -\beta_{p+q-2} & \cdots & \cdots & -\beta_2 & -\beta_1 & 1 \end{pmatrix} \tag{5.15}$$

The matrix B is invertible by the lemma 2.2 in [113], then $\Lambda \in \mathbb{R}^{p+q+1}$ is the unique solution of the linear system $B\Lambda = e$, i.e.

$$\Lambda = B^{-1}e \tag{5.16}$$

where the vector e was defined above, but in the higher dimension. We denote by $\lambda_0, \cdots, \lambda_{p+q}$ the elements of Λ and let Δ_h be the Toeplitz matrix of order h associated with the first h elements of Λ, given by

$$\Delta_h = \begin{pmatrix} \lambda_0 & \lambda_1 & \lambda_2 & \cdots & \cdots & \lambda_{h-1} \\ \lambda_1 & \lambda_0 & \lambda_1 & \cdots & \cdots & \lambda_{h-2} \\ \vdots & \vdots & \vdots & & & \vdots \\ \vdots & \vdots & \vdots & & & \vdots \\ \vdots & \vdots & \vdots & & & \vdots \\ \lambda_{h-1} & \lambda_{h-2} & \lambda_{h-3} & \cdots & \cdots & \lambda_0 \end{pmatrix} \tag{5.17}$$

which is invertible by lemma 2.2 in [113]. Denote by Π_{pq} the Hankel matrix of order $p \times q$ given by,

$$\Pi_{pq} = \begin{pmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_q \\ \lambda_2 & \lambda_3 & \cdots & \lambda_{q+1} \\ \vdots & \vdots & & \vdots \\ \lambda_p & \lambda_{p+1} & \cdots & \lambda_{p+q} \end{pmatrix} \tag{5.18}$$

and by K the matrix of order $pq \times pq$ given by,

$$K = \begin{cases} (I_q - C_\beta) \otimes I_p - C_\gamma \otimes J_p & for \quad p \geqslant q \\ (I_q - G_\beta) \otimes I_p - C_\gamma \otimes J_p & for \quad p < q \end{cases} \tag{5.19}$$

where C_β, C_γ and G_β are defined in (5.14)*.*

At first we recall the results recently introduced by Proïa [112] concerned the stable $p-$order autoregressive process where the driven noise is also given by a $1-$order autoregressive process.

Lemma 5.2.1 *We have the almost sure convergence of $\hat{\theta}_n$,*

$$\lim_{n\to\infty} \hat{\theta}_n = \theta^* \quad a.s.$$

where the limiting value

$$\theta^* = \alpha(I_p - \theta_p \rho J_p)\beta. \tag{5.20}$$

Furthermore, if only (V_n) is to admit a finite moment of order 4 (i.e., $\mathbb{E}(V_n^4) < \infty$), we have the asymptotic normality

$$\sqrt{n}(\hat{\theta}_n - \theta^*) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Sigma_\theta) \tag{5.21}$$

where the Asymptotic covariance matrix is given by

$$\Sigma_\theta = \alpha^2 (I_p - \theta_p \rho J_p)\Delta_p^{-1}(I_p - \theta_p \rho J_p) \tag{5.22}$$

which is invertible under the stability conditions.

Remark 5.2.2 *For the next results, we need to introduce some notations. We denote by P the square matrix of order $p+1$ given by*

$$P = \begin{pmatrix} P_B & 0 \\ P'_{DW_L} & \phi \end{pmatrix} \tag{5.23}$$

where

$$P_B = \alpha(I_p - \theta_p \rho J_p)\Delta_p^{-1},$$
$$P_L = J_p(I_p - \theta_p \rho J_p)(\alpha \theta_p \rho \Delta_p^{-1} e + \theta_p^* \beta),$$
$$\phi = -\alpha^{-1}\theta_p^*.$$

Moreover, let us introduce the toeplitz matrix Δ_{p+1} of the order $p+1$ which is the extension of Δ_p given by (5.17) *to the nextdimension,*

$$\Delta_{p+1} = \begin{pmatrix} \Delta_p & J_p \Lambda_p^1 \\ \Lambda_p^1 J_p & \lambda_0 \end{pmatrix} \tag{5.24}$$

with $\Lambda_p^1 = (\lambda_1 \quad \lambda_2 \quad \cdots \quad \lambda_p)$.

Lemma 5.2.2 *We have the almost sure convergence of $\hat{\rho}_n$,*

$$\lim_{n\to\infty} \hat{\rho}_n = \rho^* \quad a.s.$$

where the limiting value

$$\rho^* = \theta_p \rho \theta_p^*. \tag{5.25}$$

In addition, as soon as $\mathbb{E}(V_n^4) < \infty$, we have the asymptotic normality,

$$\sqrt{n}(\hat{\rho}_n - \rho^*) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_\rho^2) \tag{5.26}$$

where the asymptotic variance is given by

$$\sigma_\rho^2 = P_{DW_L}^{'} \Delta_p P_{DW_L} - 2\alpha_{-1}\theta_p^* \Lambda_p^{1'} J_p P_{DW_L} + (\alpha_{-1}\theta_p^*)^2 \lambda_0. \tag{5.27}$$

Hence, we have the joint asymptotic normality

$$\sqrt{n}\begin{pmatrix} \hat{\theta}_n - \theta^* \\ \hat{\rho}_n - \rho^* \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Gamma) \tag{5.28}$$

where the asymptotic covariance matrix is given by

$$\Gamma = P\Delta_{p+1}P' = \begin{pmatrix} \Sigma_\theta & \theta_p \rho J_p \Sigma_\theta e \\ \theta_p \rho e^{'} \Sigma_\theta J_p & \sigma_\rho^2 \end{pmatrix} \tag{5.29}$$

Lemma 5.2.3 *We have the almost sure convergence of the Durbin-Watson statistic,*

$$\lim_{n\to\infty} \hat{D}_n = D^* \quad a.s.$$

where the limiting value

$$D^* = 2(1 - \rho^*). \tag{5.30}$$

Furthermore, if only (V_n) is to admit a finite moment of order 4 (i.e., $\mathbb{E}(V_n^4) < \infty$), we have the asymptotic normality

$$\sqrt{n}(\hat{D}_n - D^*) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_D^2) \tag{5.31}$$

where the asymptotic variance is given by

$$\sigma_D^2 = 4\sigma_\rho^2 \tag{5.32}$$

Proof 5.2.4 *The proofs of Lemma 5.2.1, Lemma 5.2.2 and Lemma 5.2.3, are contained in [112].*

For all $h, k \in \{0, \ldots, q\}$, call

$$\Gamma_{h-k} = \begin{pmatrix} \lambda_{h-k} & \cdots & \lambda_{h-k-p+1} \\ \vdots & \ddots & \vdots \\ \lambda_{h-k+p-1} & \cdots & \lambda_{h-k} \end{pmatrix} \tag{5.33}$$

and note that $\Gamma_0 = \Delta_p$ and that $\Gamma_{h-k} = \Gamma'_{k-h}$. Hence, consider the symmetric matrix of order $pq \times pq$ given by,

$$\Gamma_{pq} = \begin{pmatrix} \Delta_p & \Gamma'_1 & \cdots & \Gamma'_{q-1} \\ \Gamma_1 & \Delta_p & \cdots & \Gamma'_{q-2} \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_{q-1} & \Gamma_{q-2} & \cdots & \Delta_p \end{pmatrix} \tag{5.34}$$

Let us recall the result recently introduced by Proïa [113].

Lemma 5.2.5 *Under the causality assumptions on $\mathcal{A}$ and $\mathcal{B}$ and as soon as $\mathbb{E}(V_1^2) = \sigma^2 < \infty$, we have the almost sure convergence*

$$\lim_{n\to\infty} T_n = T^* \quad a.s.$$

where the limiting value T^* is given

$$T^* = \Delta_p^{-1}\Pi_{pq} = (\theta^* \quad * \quad \cdots \quad *), \tag{5.35}$$

where Δ_p and Π_{pq} are given by (5.17) *and* (5.18), *respectively.*
In addition, as soon as $\mathbb{E}(V_n^4) < \infty$, we have the asymptotic normality,

$$\sqrt{n}(vec(T_n) - vec(T^*)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Sigma_T) \tag{5.36}$$

where the limiting value Σ_T is given by,

$$\Sigma_T = \sigma^2 K^{-1}(I_q \otimes \Delta_p^{-1})\Gamma_{pq}(I_q \otimes \Delta_p^{-1})K^{'-1}. \tag{5.37}$$

Remark 5.2.3 *$vec(A)$ represente the vectorization of the matrix A. Specifically, the vectorization of a matrix A of order $m \times n$ is the $mn \times 1$ column vector obtained by stacking the columns of the matrix A on top of one another.*

Proof 5.2.6 *The proof of Lemma 5.2.5 is contained in [113].*

The purpose of this paper is to furnish the moderate deviation principles on these estimates, deepening the central limit theorem which is well known. Let $(b_n)_{n\geqslant 1}$ be a sequence of positive numbers satisfying $1 = o(b_n^2)$ and $b_n^2 = o(n)$ such that

$$b_n \to \infty, \qquad \frac{b_n}{\sqrt{n}} \to 0 \quad as \quad n \to \infty \tag{5.38}$$

5.3 On moderate deviations under the Gaussian conditions

At first, we focus our attention on moderate deviations for the Durbin-Watson statistic in the case where the driven noise (V_n) is normally distributed. For this, the set of assumptions may be reduce to the existence of $t > 0$ such that

(C.1)

$$\mathbb{E}\Big[\exp(t\varepsilon_0^2)\Big] < \infty$$

(C.2)

$$\mathbb{E}\Big[\exp(tX_0^2)\Big] < \infty$$

Theorem 5.3.1 *Assume that there exists $t > 0$ such that **(C.1)** and **(C.2)** are satisfied. Then, the sequence*

$$\left(\frac{\sqrt{n}}{b_n}(\hat{\theta}_n - \theta^*)\right)_{n\geqslant 1}$$

satisfies an LDP on $\mathbb{R}^p$ with speed b_n^2 and good rate function given by

$$I_\theta(x) = \frac{1}{2}\left\langle x, \Sigma_\theta^{-1}.x\right\rangle$$

where Σ_θ is given by (5.22).

Theorem 5.3.2 *Under the causality assumptions on $\mathcal{A}$ and $\mathcal{B}$ and assume that there exists $t > 0$ such that **(C.1)** and **(C.2)** are satisfied. Then, the sequence*

$$\left(\frac{\sqrt{n}}{b_n}(vec(T_n) - vec(T^*))\right)_{n\geqslant 1}$$

satisfies an LDP on $\mathbb{R}^{pq}$ with speed b_n^2 and good rate function given by

$$I_T(x) = \frac{1}{2}\left\langle x, \Sigma_T^{-1}.x\right\rangle$$

where Σ_T is given by (5.37).

Theorem 5.3.3 *Assume that there exists $t > 0$ such that **(C.1)** and **(C.2)** are satisfied. Then, as soon as $\theta_p^* \neq 0$, the sequence*

$$\left(\frac{\sqrt{n}}{b_n}\begin{pmatrix}\hat{\theta}_n - \theta^* \\ \hat{\rho}_n - \rho^*\end{pmatrix}\right)_{n\geqslant 1}$$

satisfies an LDP on $\mathbb{R}^{p+1}$ with speed b_n^2 and good rate function given by

$$G(x) = \frac{1}{2}\left\langle x, \Gamma^{-1}.x\right\rangle$$

where Γ *is given by* (5.29).
In particular, one can see that the sequence

$$\left(\frac{\sqrt{n}}{b_n}(\hat{\rho}_n - \rho^*)\right)_{n\geqslant 1}$$

satisfies an LDP on $\mathbb{R}$ *with speed* b_n^2 *and good rate function given by*

$$I_\rho(x) = \frac{x^2}{2\sigma_\rho^2}$$

where σ_ρ^2 *is given by 5.27*

Remark 5.3.1 *The covariance matrix* Γ *is invertible under the stability conditions if only if* $\theta_p^* \neq 0$ *since, by a straightforward calculation,*

$$det(\Gamma) = \alpha^{2(p-1)}(\theta^*)^2 det(\Delta_{p+1})\left(\frac{det(I_p - \theta_p \rho J_p)}{det(\Delta_p)}\right)^2$$

according to Lemma 2.2 in [112].

Theorem 5.3.4 *Assume that there exists* $t > 0$ *such that* ***(C.1)*** *and* ***(C.2)*** *are satisfied. Then, the sequence*

$$\left(\frac{\sqrt{n}}{b_n}(\hat{D}_n - D^*)\right)_{n\geqslant 1}$$

satisfies an LDP on $\mathbb{R}$ *with speed* b_n^2 *and good rate function given by*

$$I_D(x) = \frac{x^2}{2\sigma_D^2}$$

where σ_D^2 *is given by 5.32*

Remark 5.3.2 *In particular when* $p = 1$ *and* $q = 1$ *in* (5.1)*, one can see that our results of Theorem 5.3.1, Theorem 5.3.3 and Theorem 5.3.4 correspond respectively to Theorem 2.1, Theorem 2.2 and Theorem 2.3 in [25].*

5.4 On moderate deviations under the Chen-Ledoux type condition

In this part, we will use a powerful result of Puhalskii, our attention is then to focus on the more general case where the driven noise $(V_n)$is assumed to satisfy the Chen-Ledoux type condition. By starting,one can introduce the following hypothesis, for $a = 2$ and $a = 4$.

(CL.1)

Chen-Ledoux

$$\limsup_{n\to\infty} \frac{1}{b_n^2} \log n\mathbb{P}\Big(|V_1^a| > b_n\sqrt{n}\Big) = -\infty.$$

(CL.2)

$$\frac{|\varepsilon_0|^a}{b_n\sqrt{n}} \xrightarrow[b_n^2]{Superexp} 0.$$

(CL.3)

$$\frac{|X_0|^a}{b_n\sqrt{n}} \xrightarrow[b_n^2]{Superexp} 0.$$

Theorem 5.4.1 *Assume that **(CL.1)**, **(CL.2)** and **(CL.3)** are satisfied. Then, the sequence*

$$\left(\frac{\sqrt{n}}{b_n}(\hat{\theta}_n - \theta^*)\right)_{n\geqslant 1}$$

satisfies an LDP on $\mathbb{R}^p$ given in Theorem 5.3.1.

Theorem 5.4.2 *Assume that **(CL.1)**, **(CL.2)** and **(CL.3)** are satisfied. Then, the sequence*

$$\left(\frac{\sqrt{n}}{b_n}(vec(T_n) - vec(T^*))\right)_{n\geqslant 1}$$

satisfies an LDP on $\mathbb{R}^{pq}$ given in Theorem 5.3.2.

Theorem 5.4.3 *Assume that **(CL.1)**, **(CL.2)** and **(CL.3)** are satisfied. Then, as soon as $\theta_p^* \neq 0$, the sequence*

$$\left(\frac{\sqrt{n}}{b_n}\begin{pmatrix}\hat{\theta}_n - \theta^* \\ \hat{\rho}_n - \rho^*\end{pmatrix}\right)_{n\geqslant 1}$$

satisfies an LDP on $\mathbb{R}^{p+1}$ given in Theorem 5.3.3.
In particular, one can see that the sequence

$$\left(\frac{\sqrt{n}}{b_n}(\hat{\rho}_n - \rho^*)\right)_{n\geqslant 1}$$

satisfies an LDP on $\mathbb{R}$ also given in Theorem 5.3.3.

Theorem 5.4.4 *Assume that **(CL.1)**, **(CL.2)** and **(CL.3)** are satisfied. Then, the sequence*

$$\left(\frac{\sqrt{n}}{b_n}(\hat{D}_n - D^*)\right)_{n\geqslant 1}$$

satisfies an LDP on $\mathbb{R}$ given in Theorem 5.3.4.

Remark 5.4.1 *In particular when $p = 1$ and $q = 1$ in* (5.1)*, one can see that our results of Theorem 5.4.1, Theorem 5.4.3 and Theorem 5.4.4 correspond respectively to Theorem 3.1, Theorem 3.2 and Theorem 3.3 in [25].*

5.5 Proof of the main results

Our proofs are essentially based to the proofs in section 4 in [25]. Before starting, let us introduce somes additional notations. At first, for all $n \geqslant 1$, Let

$$L_{d,n} = \sum_{k=1}^{n} V_{k-d-1}^2, \qquad d = 1, \cdots, p+1. \tag{5.39}$$

Now, let $(M_{1,n})$ be a $p-$dimensional martingale

$$M_{1,n} = \sum_{k=1}^{n} \Phi_{k-1}^{p} V_k. \tag{5.40}$$

Proof of Theorem 5.3.1

For prove the Theorem 5.3.1, we need to introduce some technical tools. Let denote by L the almost sure limit of $\frac{1}{n}S_n$
where

$$L = \sigma^2 \Delta_p. \tag{5.41}$$

Lemma 5.5.1 *Under the assumptions of Theorem 5.3.1, we have the exponential convergence*

$$\frac{1}{n}S_n \xrightarrow[b_n^2]{Superexp} L. \tag{5.42}$$

Proof 5.5.2 *We start by introducing the following notation*

$$\sum_{k=1}^{n} \Phi_{k-1}^{p+1} X_k = C^{-1}.T \sum_{k=1}^{n} X_k^2 + C^{-1} \sum_{k=1}^{n} \Phi_{k-1}^{p+1} V_k, \tag{5.43}$$

where Φ_n^{p+1} stands for the extension of Φ_n^p given by (5.2) *to the next dimension, and where C is the square submatrix of order $p+1$ obtained by* *removing from B given by (5.15) its first row and first columm which is invertible under the stability conditions*

$$C = \begin{pmatrix} 1-\beta_2 & -\beta_3 & \cdots & \cdots & -\beta_p & \theta_p\rho & 0 \\ -\beta_1-\beta_3 & 1-\beta_4 & \cdots & \cdots & \theta_p\rho & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ -\beta_{p-1}+\theta_p\rho & -\beta_{p-2} & \cdots & \cdots & -\beta_1 & 1 & 0 \\ -\beta_p & -\beta_{p-1} & \cdots & \cdots & -\beta_2 & -\beta_1 & 1 \end{pmatrix} \tag{5.44}$$

and where T is defined as

$$T = (\beta_1 \quad \beta_2 \quad \cdots \quad \beta_p \quad -\theta_p\rho)' \tag{5.45}$$

We shall prove the exponential convergence of $\sum_{k=1}^n \Phi_{k-1}^{p+1} X_k$ i.e we shall prove that

$$\frac{1}{n}\sum_{k=1}^n \Phi_{k-1}^{p+1} X_k \xrightarrow[b_n^2]{\text{Superexp}} \sigma^2 \Lambda_{p+1}^1, \tag{5.46}$$

where Λ_p^1 is given by (5.16).
After straightforward calculations, we get that for all $n \geqslant 2$,

$$\frac{1}{n}\sum_{k=1}^n \Phi_{k-1}^{p+1} X_k - \sigma^2\Lambda_{p+1}^1 = C^{-1}.T\left(\frac{1}{n}\sum_{k=1}^n X_k^2 - \sigma^2\lambda_0\right) + \sigma^2 C^{-1}.\left(\frac{1}{n}N_n\right), \tag{5.47}$$

where N_n is given by

$$N_{1,n} = \sum_{k=1}^n \Phi_{k-1}^{p+1} V_k. \tag{5.48}$$

First of all, by the straightforward calculations it follows from Lemma 4.1 of [25] that

$$\frac{1}{n}\sum_{k=1}^n X_k^2 - \sigma^2\lambda_0 \xrightarrow[b_n^2]{\text{Superexp}} 0 \tag{5.49}$$

To end the proof of Lemma 5.5.1, we have to study the exponential asymptotic behavior of $\frac{1}{n}N_n$. Let us note

$$N_{1,n} = (\mathcal{R}_{1,n} \quad \mathcal{R}_{2,n} \quad \cdots \quad \mathcal{R}_{p+1,n})', \tag{5.50}$$

where $\mathcal{R}_{d,n} = \sum_{k=1}^n X_{k-d}V_k, \quad for \quad d = 1, \cdots, p+1.$
For all $\zeta > 0$ and a suitable $z > 0$

$$\begin{aligned} \mathbb{P}\Big(\frac{\mathcal{R}_{d,n}}{n} > \zeta\Big) &\leqslant \mathbb{P}\Big(\frac{\mathcal{R}_{d,n}}{n} > \zeta, \langle \mathcal{R}_d \rangle_n \leqslant z\Big) + \mathbb{P}\Big(\frac{\mathcal{R}_{d,n}}{n} > \zeta, \langle \mathcal{R}_d \rangle_n > z\Big) \\ &\leqslant \exp\Big(-\frac{n^2\zeta^2}{2z}\Big) + \mathbb{P}\Big(\langle \mathcal{R}_d \rangle_n > z\Big), \end{aligned} \tag{5.51}$$

by the application of Theorem 4.1 of [20] in the case of a gaussian martingale with $\langle \mathcal{R}_d \rangle_n = \sigma^2 \sum_{k=1}^n X_{k-d-1}^2$. *We have the following inequality,*

$$\sum_{k=1}^n X_{k-d-1}^2 \leqslant \alpha X_0^2 + \beta \varepsilon_0^2 + \beta L_{d,n}, \tag{5.52}$$

where $L_{d,n}$ *is given by* (5.39), *with* $\alpha = \max_d [1 + (1 - |\theta_d|)^{-2}]$ *and* $\beta = \max_d [(1 - |\rho|)^{-2}(1 - |\theta_d|)^{-2}]$ *for all* $d = 1, \cdots, p+1$, *we get from Markov's inequality that for a suitable* $t > 0$

$$\begin{aligned}
\mathbb{P}\Big(\langle \mathcal{R}_d \rangle_n > z\Big) &\leqslant \mathbb{P}\Big(X_0^2 > \frac{z}{3\alpha\sigma^2}\Big) + \mathbb{P}\Big(\varepsilon_0^2 > \frac{z}{3\beta_{d+1}\sigma^2}\Big) + \mathbb{P}\Big(L_{d,n} > \frac{z}{3\beta_{d+1}\sigma^2}\Big) \\
&\leqslant \exp\Big(-\frac{zt}{3\alpha\sigma^2}\Big)\mathbb{E}\Big[\exp(tX_0^2)\Big] + \exp\Big(-\frac{zt}{3\beta_{d+1}\sigma^2}\Big)\mathbb{E}\Big[\exp(t\varepsilon_0^2)\Big] \\
&+ \mathbb{P}\Big(L_{d,n} > \frac{z}{3\beta_{d+1}\sigma^2}\Big) \\
&\leqslant 3\max\Big\{\exp\Big(-\frac{zt}{3\alpha\sigma^2}\Big)\mathbb{E}\Big[\exp(tX_0^2)\Big], \exp\Big(-\frac{zt}{3\beta_{d+1}\sigma^2}\Big)\mathbb{E}\Big[\exp(t\varepsilon_0^2)\Big], \\
&\mathbb{P}\Big(L_{d,n} > \frac{z}{3\beta_{d+1}\sigma^2}\Big)\Big\}
\end{aligned} \tag{5.53}$$

Let chose $z = ny$, *assuming* $y > 3\beta\sigma^4$. *It follows that*

$$\begin{aligned}
\frac{1}{b_n^2}\log \mathbb{P}\Big(\langle \mathcal{R}_d \rangle_n > ny\Big) \leqslant \frac{\log(3)}{b_n^2} + \frac{1}{b_n^2}\max\Big\{-\frac{nyt}{3\alpha\sigma^2} + \log \mathbb{E}\Big[\exp(tX_0^2)\Big], \\
-\frac{nyt}{3\beta_{d+1}\sigma^2} + \log \mathbb{E}\Big[\exp(t\varepsilon_0^2)\Big], \mathbb{P}\Big(L_{d,n} > \frac{ny}{3\beta_{d+1}\sigma^2}\Big)\Big\}
\end{aligned}$$

Like (V_n) *is a sequence of independent and identically distributed gaussian random variables with zero mean and variance* $\sigma^2 > 0$. *It clearly follows from Cramér-Chernoff's Theorem, expound e.g in [66], that for all* $\delta > 0$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\Big(\Big|\frac{L_{d,n}}{n} - \sigma^2\Big| > \delta\Big) < 0. \tag{5.54}$$

and since $b_n^2 = o(n)$, *the latter leads to*

$$\frac{L_{d,n}}{n} \xrightarrow[b_n^2]{\textit{Superexp}} \sigma^2, \tag{5.55}$$

Then, since $b_n^2 = o(n)$ *and with* $\delta = \frac{y}{3\beta_{d+1}\sigma^2} - \sigma^2 > 0$, *it follows that*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\Big(L_{d,n} > \frac{ny}{3\beta_{d+1}\sigma^2}\Big) = -\infty. \tag{5.56}$$

*From **(C.1)**, **(C.2)** and* (5.56)*, we clearly obtain that*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\Big(\langle \mathcal{R}_d \rangle_n > ny \Big) = -\infty. \tag{5.57}$$

It allows us by (5.51) *to deduce that for all* $\zeta > 0$*, for all* $d = 1, \quad \cdots, \quad p+1$

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\Big(\frac{\mathcal{R}_{d,n}}{n} > \zeta \Big) = -\infty. \tag{5.58}$$

The same result is also true replacing $\mathcal{R}_{d,n}$ *by* $-\mathcal{R}_{d,n}$ *in* (5.58) *since* $\mathcal{R}_{d,n}$ *and* $-\mathcal{R}_{d,n}$ *share the same distribution.*
Then, it follows that

$$\frac{1}{n} N_{1,n} \xrightarrow[b_n^2]{Superexp} 0 \tag{5.59}$$

From (5.49) *and* (5.59) *together with* (5.47)*, we obtain* (5.46)*.*
Hereafter the straightforward calculations and by a similar reasoning columm by columm on the matrix S_n*, Lemma 5.5.1 holds.*

Now we can prove the Theorem 2.1. We are going to use the following Theorem established by worms [135] about the deviation principle for martingales.

Theorem 5.5.3 *Let* (Z_n) *be an adapted sequence with values in* R^p*, and* (V_n) *a gaussian noise with variance* $\sigma^2 > 0$*. We suppose that* (Z_n) *satisfies, for some invertible covariance matrix* C *of order* p *and a speed sequence* (a_n^2) *such that* $a_n^2 = o(n)$*, the exponential convergence for any* $\delta > 0$

$$\lim_{n\to\infty} \frac{1}{a_n^2} \log \mathbb{P}\Bigg(\Big\| \frac{1}{n} \sum_{k=0}^{n-1} Z_k . Z_k^{'} - C \Big\| > \delta \Bigg) = -\infty. \tag{5.60}$$

Then, the sequence $\Big(M_n/(a_n\sqrt{n}) \Big)_{n\geqslant 1}$ *satisfies an LDP on* R^p *with speed* a_n^2 *and rate function given by*

$$I(x) = \frac{1}{2\sigma^2} x'.C^{-1}.x \tag{5.61}$$

where (M_n) *is the martingale given by*

$$M_n = \sum_{k=1}^{n} Z_{k-1} V_k$$

Proof 5.5.4 *The proof of Theorem 5.5.3 is contained in the one of Theorem 5 of [135] with* $d = 1$*.* $\square$

Proof of Theorem 5.3.1

Let us consider the decomposition

$$\frac{\sqrt{n}}{b_n}(\hat{\theta}_n - \theta^*) = \frac{\sqrt{n}}{b_n}\alpha(S_{n-1})^{-1}(I_p - \theta_p\rho J_p)M_{1,n} + \frac{\sqrt{n}}{b_n}\alpha(S_{n-1})^{-1}\xi_n. \quad (5.62)$$

where S_n is given by (5.5) and the remainder term ξ_n is made of isolated terms such that $\|\xi_n\| = o(\sqrt{n})$ $\quad a.s$ and $M_{1,n}$ is given by (5.40) and where α is given by (5.11).
We infer from Lemma 5.5.1 together with Lemma 4.1 of [134] that the following convergence is satisfied,

$$n(S_n)^{-1} \xrightarrow[b_n^2]{Superexp} L^{-1}. \quad (5.63)$$

Since $\|\xi_n\| = o(\sqrt{n})$, from (5.63) we obtain

$$\frac{\sqrt{n}}{b_n}\alpha(S_{n-1})^{-1}\xi_n \xrightarrow[b_n^2]{Superexp} 0. \quad (5.64)$$

and

$$\left[n\alpha(S_{n-1})^{-1} - \alpha L^{-1}\right](I_p - \theta_p\rho J_p)(\frac{1}{b_n\sqrt{n}}.M_{1,n}) \xrightarrow[b_n^2]{Superexp} 0. \quad (5.65)$$

Hence, from (5.62), (5.64) and (5.65) together, it easy to see that

$$\frac{\sqrt{n}}{b_n}(\hat{\theta}_n - \theta^*) \underset{b_n^2}{\overset{Superexp}{\sim}} \alpha L^{-1}(I_p - \theta_p\rho J_p)(\frac{1}{b_n\sqrt{n}}.M_{1,n}). \quad (5.66)$$

Since Lemma 5.5.1 together with Theorem 5.5.3, we immediately show that $(M_{1,n}/(b_n\sqrt{n}))$ satisfies an LDP on $\mathbb{R}^p$ with speed b_n^2 and good rate function given by

$$K_p(x) = \frac{1}{2\sigma^2}x'.L^{-1}.x \quad (5.67)$$

Hence, by the contraction principle in [53] together with (5.66), one can see that $\left(\frac{\sqrt{n}}{b_n}(\hat{\theta}_n - \theta^*)\right)_{n\geqslant 1}$ satisfies an LDP on $\mathbb{R}^p$ with speed b_n^2 and good rate function $I_\theta(x)$ such that $I_\theta(x) = K_p(\frac{1}{\alpha}(I_p - \theta_p\rho J_p)^{-1}L.x)$, after a straightforward calculations, we get that

$$I_\theta(x) = \frac{1}{2}\left\langle x, \Sigma_\theta^{-1}.x\right\rangle$$

Which achieves the proof of Theorem 5.3.1.

Proof of Theorem 5.3.2

The proof of Theorem 5.3.2 is essentially based to the proof of proof of Theorem 5.3.1. It is just an extension of the proof of Theorem 5.3.1 for all $q \geqslant 1$.We will also suppose that $p \geqslant q$, in a first time.
By the straightforward calculations, it follows that, for all $1 \leqslant k \leqslant n$

$$X_k = \beta' \Phi^p_{k-1} + \gamma' \Phi^p_{k-p-1} + V_k \tag{5.68}$$

where β and γ are given by (5.12).
It follows that, for all $h \in \{1, \cdots, q\}$

$$\sum_{k=1}^{n} \Phi^p_{k-h} X_k = \sum_{k=1}^{n} \Phi^p_{k-h} \Phi^{p'}_{k-1} \beta + \sum_{k=1}^{n} \Phi^p_{k-h} \Phi^{q'}_{k-p-1} \gamma + \sum_{k=1}^{n} \Phi^p_{k-h} V_k$$

Let introduce, for all $t \in \{1, \cdots, p\}$ and $h \in \{1, \cdots, q\}$

$$\Sigma_{n,t} = \sum_{k=1}^{n} \Phi^p_k X_{k-t} \quad and \quad \Omega_{n,h} = \sum_{k=1}^{n} \Phi^p_{k-h} X_k \tag{5.69}$$

From Proïa [113] (6.7 and 6.12), one can see that

$$\sum_{k=1}^{n} \Phi^p_{k-h} \Phi^{p'}_{k-1} \beta = \beta_1 \Omega_{n,h-1} + \beta_2 \Omega_{n,h-2} + \cdots + \beta_p \Sigma_{n,p-h} + r_{h,n} \tag{5.70}$$

and

$$\sum_{k=1}^{n} \Phi^p_{k-h} \Phi^{q'}_{k-p-1} \gamma = \gamma_1 \Sigma_{n,p-h+1} + \gamma_2 \Sigma_{n,p-h+2} + \cdots + \gamma_p G_p \Omega_{n,1} + \tau_{h,n} \tag{5.71}$$

where the remainders terms satisfy, for all $h \in \{1, \cdots, q\}$

$$\|r_{h,n}\| = o(n) \quad and \quad \|\tau_{h,n}\| = o(n) \tag{5.72}$$

The Following step is to note that

$$S_{n-1} = \begin{pmatrix} \Sigma_{n,0} & \Sigma_{n,1} & \cdots & \Sigma_{n,p-1} \end{pmatrix} + o(n)$$

and

$$W_n = \begin{pmatrix} \Omega_{n,1} & \Omega_{n,2} & \cdots & \Omega_{n,q} \end{pmatrix} + o(n)$$

where S_n and W_n are given by (5.5) and (5.4) respectively.
Now, let (M_n) be the following matrix martingale of order $p \times p$,

$$M_n = \begin{pmatrix} M_{1,n} & M_{2,n} & \cdots & M_{q,n} \end{pmatrix} \tag{5.73}$$

in which each vector is itself a vector martingale of order p given, for all $h \in \{1, \cdots, q\}$ by

$$M_{h,n} = \sum_{k=1}^{n} \Phi^p_{k-h} V_k. \tag{5.74}$$

Under ours assumptions, the vector martinagle $(M_{h,n})$ is also a locally square-integrable adapted to some filtration $(\mathcal{F}_n)$. More precisely, its predictable quadratic variation is given, for all $n \geqslant 1$ by

$$\langle M_h \rangle_n = \sigma^2 S_{n-h}.$$

By virtue of lemma 5.5.1 and lemma 6.1 in [113], we have

$$\frac{1}{n}\langle M_h \rangle_n \xrightarrow[b_n^2]{Superexp} \sigma^2 \Delta_p \tag{5.75}$$

and

$$\frac{1}{n}\langle M_h \rangle_n^{-1} \xrightarrow[b_n^2]{Superexp} \frac{1}{\sigma^2}\Delta_p^{-1} \tag{5.76}$$

Let us consider the decomposition

$$vec(T_n) - vec(T^* = K^{-1} vec(S_{n-1}^{-1} M_n) + K^{-1} vec(S_{n-1}^{-1} R_n) \tag{5.77}$$

where K is given by (5.19) and $|||R_n||| = o(\sqrt{n}) \quad a.s.$
It follows from (5.76) and a straightforward calculation that

$$\frac{\sqrt{n}}{b_n} vec(S_{n-1}^{-1} R_n) \xrightarrow[b_n^2]{Superexp} 0. \tag{5.78}$$

Moreover, it is not hard to see that

$$vec(S_{n-1}^{-1} M_n) = (I_q \otimes S_{n-1}^{-1}) vec(M_n)$$

where $vec(M_n)$ is a vector $(\mathcal{F}_n)-$martingale of order pq. The predictable quadratic variation of $vec(M_n)$ is given, for all $n \geqslant 1$, by

$$\langle vec(M) \rangle_n = \begin{pmatrix} \langle M_1, M_1 \rangle_n & \langle M_1, M_2 \rangle_n & \cdots & \langle M_1, M_q \rangle_n \\ \langle M_2, M_1 \rangle_n & \langle M_2, M_2 \rangle_n & \cdots & \langle M_2, M_q \rangle_n \\ \vdots & \vdots & \ddots & \vdots \\ \langle M_q, M_1 \rangle_n & \langle M_q, M_2 \rangle_n & \cdots & \langle M_q, M_q \rangle_n \end{pmatrix}$$

where for all $t, h \in \{1, \cdots, q\}$

$$\langle M_t, M_h \rangle_n = \sigma^2 \sum_{k=1}^{n} \Phi_{k-t}^{p} \Phi_{k-h}^{p'}.$$

We obviously have

$$n(I_q \otimes S_{n-1}^{-1}) \xrightarrow[b_n^2]{Superexp} (I_q \otimes \Delta_p^{-1}) \tag{5.79}$$

Hence, from (5.77), (5.78) and (5.79) together, it easy to see that

$$\frac{\sqrt{n}}{b_n}(vec(T_n) - vec(T^*) \underset{b_n^2}{\overset{Superexp}{\sim}} \frac{1}{\sigma^2} K^{-1} (I_q \otimes \Delta_p^{-1})(\frac{1}{b_n \sqrt{n}} vec(M_n)). \tag{5.80}$$

As a matter of fact, if we denote by

$$\varphi_{k-1}^{pq} = (\Phi_{k-1}^{p'} \quad \Phi_{k-2}^{p'} \quad \cdots \quad \Phi_{k-q}^{p'})'$$

which is a vector of order pq, then it is not hard to see that

$$vec(M_n) = \sum_{k=1}^{n} \varphi_{k-1}^{pq} V_k.$$

By the exponential asymptotic behavior given by (5.59) and the theorem 5 of [136], one can immediately see that the sequence $vec(M_n)/(b_n\sqrt{n})$ satisfies an LDP on $\mathbb{R}^{pq}$ with speed b_n^2 and good rate function given by

$$K_{pq}(x) = \frac{1}{2\sigma^2} x'.\Gamma_{pq}^{-1}.x \tag{5.81}$$

where Γ_{pq} is given by (5.34).
By the contraction principle in [53] together with (5.80), it follows that the sequence $\left(\frac{\sqrt{n}}{b_n}(vec(T_n) - vec(T^*))\right)$ satisfies an LDP on $\mathbb{R}^{pq}$ with speed b_n^2 and good rate function $I_T(x)$, given by

$$I_T(x) = K_{pq}(\sigma^2(I_q \otimes \Delta_p^{-1})^{-1} K x)$$

After a straightforward calculations, we get that

$$I_T(x) = \frac{1}{2}\left\langle x, \Sigma_T^{-1}.x\right\rangle$$

We will not detail the case where $p < q$ since the reasoning is very similar. Which ends the proof of theorem 5.3.2.

Proof of Theorem 5.3.3

Before starting, let us introduce some additional notations from Appendix C in [112]. Let us introduce two specifics notations $Y_n$and Z_n given by

$$Y_n = X_n - \rho^* X_{n-1} \quad and \quad Z_n = X_{n-1} - \rho^* X_n.$$

We also note $Y_n^p = (Y_n \quad Y_{n-1} \quad \cdots \quad Y_{n-p+1})'$ and $Z_n^p = (Z_n \quad Z_{n-1} \quad \cdots \quad Z_{n-p+1})'$. Let F_n be the recurrent $p-$dimensional expression given, for all $n \geqslant 1$, by

$$F_n = \Phi_n^p \theta^{*'} Z_n^p - (Z_{n-1}^p + Y_n^p) X_n. \tag{5.82}$$

We shall make use of the decomposition

$$\frac{\sqrt{n}}{b_n}\begin{pmatrix}\hat{\theta}_n - \theta^* \\ \hat{\rho}_n - \rho^*\end{pmatrix} = \frac{1}{b_n\sqrt{n}} P_n N_{1,n} + \frac{1}{b_n} R_n \tag{5.83}$$

where $N_{1,n}$ be the$(p+1)-$dimensional martingale given by (5.48) with Φ_n^{p+1} the extension of Φ_n^p in the next dimension, where the square matrix P_n of order $p+1$ is given by

$$P_n = \begin{pmatrix} P_n^{(1,1)} & 0 \\ P_n^{(2,1)} & P_n^{(2,2)} \end{pmatrix} \tag{5.84}$$

with
$P_n^{(1,1)} = n\alpha(S_{n-1})^{-1}(I_p - \theta_p \rho J_p),$
$P_n^{(2,1)} = n(\mathcal{J}_{n-1})^{-1}\Big(\alpha(U_p + (I_p - \theta_p\rho J_p)(\rho_*\theta_p^* - \tau^*))' + \alpha H_n'(S_{n-1})^{-1}(I_p -$
$\theta_p\rho J_p)\Big),$
$P_n^{(2,2)} = -n(\mathcal{J}_{n-1})^{-1}\theta_p^*,$

where H_n is a $p-$dimensional vector, and for all $n \geqslant 1$

$$H_n = \sum_{k=1}^{n}(Z_k^p \theta^{*'} \Phi_k^p + F_k) + \sum_{k=1}^{n} \Phi_k^p(\hat{\theta}_n - \theta^*)' Z_k^p + \mu_n, \tag{5.85}$$

$$\mathcal{J}_n = \sum_{k=0}^{n} \hat{\varepsilon}_k^2 \tag{5.86}$$

$$U_p = (1+\beta_2 \quad \beta_3 - \beta_1 \quad \cdots \quad \beta_p - \beta_{p-2} \quad -\beta_{p-1} - \theta_p\rho)', \tag{5.87}$$

with $\|\mu\| = o(\sqrt{n})$,

and where the remainder term R_n of order $p+1$ is given by

$$R_n = \sqrt{n}\begin{pmatrix} \alpha(S_{n-1})^{-1}\xi_n \\ \mathcal{J}_{n-1} r_n \end{pmatrix} \tag{5.88}$$

such that $\|\xi_n\| = o(\sqrt{n})$ and $r_n = o(\sqrt{n})$.

We shall prove that

$$P_n \xrightarrow[b_n^2]{Superexp} \sigma^{-2} P, \tag{5.89}$$

where P is given by (5.23) and

$$\frac{1}{b_n} R_n \xrightarrow[b_n^2]{Superexp} 0. \tag{5.90}$$

For this, we have to prove that

$$\frac{1}{n}\mathcal{J}_{n-1} \xrightarrow[b_n^2]{Superexp} \sigma^2(\lambda_0 - \Lambda_p^{1'}.\theta^*), \tag{5.91}$$

where λ_0 and Λ_p^1 are given by (5.17) and (5.24).

By some simple calculations, we have

$$\mathcal{J}_{n-1} = \sum_{k=1}^{n} X_{k-1}^2 - 2\hat{\theta}_n' \sum_{k=1}^{n} X_{k-1}\Phi_{k-2}^{p'} + \hat{\theta}_n'.S_{n-2}.\hat{\theta}_n$$

Since (5.42) together Theorem 5.3.1, one can see that (5.91) holds. Furthermore, it is not hard to see, via $\|\mu\| = o(\sqrt{n})$, (5.42), Theorem 5.3.1 and some simplifications one H_n, that

$$\frac{1}{n}H_n \xrightarrow[b_n^2]{Superexp} -\alpha.(I_p - \theta_p\rho J_p).e. \tag{5.92}$$

Hence, from (5.91), (5.92) and since $\|\xi_n\| = o(\sqrt{n})$ and $r_n = o(\sqrt{n})$, then (5.89) and (5.90) follow. As a consequence,

$$\frac{\sqrt{n}}{b_n}\begin{pmatrix}\hat{\theta}_n - \theta^* \\ \hat{\rho}_n - \rho^*\end{pmatrix} \underset{b_n^2}{\overset{Superexp}{\sim}} \sigma^{-2}P.\frac{N_{1,n}}{b_n\sqrt{n}}, \tag{5.93}$$

and implies that both of them share the same LDP, see e.g. [53]. Hence, by the contraction principle in [53] together with Theorem 5.5.3, it is not hard to see that the sequence $\dfrac{\sqrt{n}}{b_n}\begin{pmatrix}\hat{\theta}_n - \theta^* \\ \hat{\rho}_n - \rho^*\end{pmatrix}$ satisfies an LDP on $\mathbb{R}^{p+1}$ with speed b_n^2 and rate function given by $G(x) = J_{p+1}(\sigma^2P^{-1}.x)$, that is

$$G(x) = \frac{1}{2}\left\langle x, \Gamma^{-1}.x\right\rangle.$$

In particular, the latter result also implies that the sequence $\left(\dfrac{\sqrt{n}}{b_n}(\hat{\rho}_n - \rho^*)\right)_{n\geqslant 1}$ satisfies an LDP on $\mathbb{R}$ with speed b_n^2 and good rate function given by

$$I_\rho(x) = \frac{x^2}{2\sigma_\rho^2}$$

where σ_ρ^2 is given by 5.27 (this is the last element to the matrix Γ). This end the proof of Theorem 5.3.3. □

Proof of Theorem 5.3.4

Let us introduce the explosion coefficient g_n related with $\mathcal{J}_n$ which is given by (5.86), that is for all $n \geqslant 1$

$$g_n = (\mathcal{J}_{n-1})^{-1}(\mathcal{J}_n - \mathcal{J}_{n-1}) = (\mathcal{J}_{n-1})^{-1}.\hat{\varepsilon}_n^2 \tag{5.94}$$

From the decomposition (C.4) in [19], it follows that

$$\frac{\sqrt{n}}{b_n}(\hat{D}_n - D^*) = -2\frac{\sqrt{n}}{b_n}\Big(1 - g_n\Big)\Big(\hat{\rho}_n - \rho^*\Big) + \frac{\sqrt{n}}{b_n}\zeta_n, \tag{5.95}$$

where the remainder term ζ_n is made of isolated terms. We clearly have (see the proof of Theorem 2.3 in [25] for details)

$$\frac{\sqrt{n}}{b_n}\zeta_n \xrightarrow[b_n^2]{Superexp} 0 \quad and \quad g_n \xrightarrow[b_n^2]{Superexp} 0.$$

Hence, it follows that

$$\frac{\sqrt{n}}{b_n}(\hat{D}_n - D^*) \overset{Superexp}{\underset{b_n^2}{\sim}} -2\frac{\sqrt{n}}{b_n}\Big(\hat{\rho}_n - \rho^*\Big)$$

By the contraction principle in [53] together with Theorem 5.3.3, it follows that the sequence $\left(\frac{\sqrt{n}}{b_n}(\hat{D}_n - D^*)\right)_{n\geqslant 1}$ satisfies an LDP on $\mathbb{R}$ with speed b_n^2 and good rate function given by $I_D(x) = I_\rho(-x/2)$, that is

$$I_D(x) = \frac{x^2}{2\sigma_D^2},$$

which end the proof of Theorem 5.3.4. □

Proofs of Theorem 5.4.1, Theorem 5.4.2, Theorem 5.4.3 and Theorem 5.4.4

Our proof is essentially based to somes technical lemmas Which ensuring that all results already proved under the gaussian assumption still hold under the Chen-Ledoux type condition.

Lemma 5.5.5 *Under **(CL.1)**, **(CL.2)** and **(CL.3)**, all exponential convergences of 5.89, 5.90 and Lemma 5.5.1 still hold.*

Proof 5.5.6 *We know that Under **(CL.1)**, **(CL.2)** and **(CL.3)**, the Lemma 4.1 in [25] still hold ie the exponential convergence given by 5.49 holds. Furthermore, under **(CL.1)** with $a = 2$ and from Theorem 2.2 of [67], it follows that for all $d = 1, \cdots, p+1$,*

$$\frac{L_{d,n}}{n} \xrightarrow[b_n^2]{Superexp} \sigma^2, \tag{5.96}$$

Moreover, since $(\mathcal{R}_{d,n})_{n\geqslant 1}$ is locally square integrable martingale, we deduce from theorem 2.1 of [20] that for all $x, y > 0$,

$$\mathbb{P}\Big(|\mathcal{R}_{d,n}| > x, \langle\mathcal{R}_d\rangle_n + [\mathcal{R}_d]_n \leqslant y\Big) \leqslant 2\exp\Big(-\frac{x^2}{2y}\Big) \tag{5.97}$$

where the predictable quadratic variation $\langle\mathcal{R}_d\rangle_n = \sigma^2\sum_{k=1}^n X_{k-d-1}^2$ and the total quadratic variation is given by $[\mathcal{R}_d]_0 = 0$ and for all $n \geqslant 1$, by

$$[\mathcal{R}_d]_n = \sum_{k=1}^n X_{k-d-1}^2 V_k^2, \quad for \quad d = 1, \cdots, p+1 \tag{5.98}$$

According to (5.97)*, we have for all $\zeta > 0$ and a suitable $b > 0$*

$$\begin{aligned}\mathbb{P}\Big(\frac{|\mathcal{R}_{d,n}|}{n} > \zeta\Big) &\leqslant \mathbb{P}\Big(|\mathcal{R}_{d,n}| > n\zeta, \langle \mathcal{R}_d \rangle_n + [\mathcal{R}_d]_n \leqslant nb\Big) + \mathbb{P}\Big(\langle \mathcal{R}_d \rangle_n + [\mathcal{R}_d]_n > nb\Big)\\ &\leqslant 2\exp\Big(-\frac{n\zeta^2}{2b}\Big) + \mathbb{P}\Big(\langle \mathcal{R}_d \rangle_n + [\mathcal{R}_d]_n > nb\Big)\\ &\leqslant 2\max\Big[\mathbb{P}\Big(\langle \mathcal{R}_d \rangle_n + [\mathcal{R}_d]_n > nb\Big), 2\exp\Big(-\frac{n\zeta^2}{2b}\Big)\Big]\end{aligned} \tag{5.99}$$

Consequently,

$$\lim_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\Big(\frac{|\mathcal{R}_{d,n}|}{n} > \zeta\Big) \leqslant \lim_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\Big(\langle \mathcal{R}_d \rangle_n + [\mathcal{R}_d]_n > nb\Big). \tag{5.100}$$

We have for all $b > 0$,

$$\begin{aligned}\mathbb{P}\Big(\langle \mathcal{R}_d \rangle_n + [\mathcal{R}_d]_n > nb\Big) &\leqslant \mathbb{P}\Big(\langle \mathcal{R}_d \rangle_n > \frac{nb}{2}\Big) + \mathbb{P}\Big([\mathcal{R}_d]_n > \frac{nb}{2}\Big)\\ &\leqslant 2\max\Big[\mathbb{P}\Big(\langle \mathcal{R}_d \rangle_n > \frac{nb}{2}\Big), \mathbb{P}\Big([\mathcal{R}_d]_n > \frac{nb}{2}\Big)\Big]\end{aligned} \tag{5.101}$$

In addition, let us define for all $n \geqslant 1$,

$$T_{d,n} = \sum_{k=1}^{n} X_{k-d-1}^4 \quad \textit{and} \quad \Gamma_{d,n} = \sum_{k=1}^{n} V_{k-d-1}^4$$

It's not hard to see that we easily have the following inequality,

$$T_{d,n} \leqslant \alpha X_0^4 + \beta \varepsilon_0^4 + \beta \Gamma_{d,n}, \tag{5.102}$$

with $\alpha = \max_d (1 - |\theta_d|)^{-4}$ and $\beta = \max_d (1 - |\rho|)^{-4}(1 - |\theta_d|)^{-4}$ for all $d = 1, \cdots, p+1$. This implies that, for n large enough, one can find $\delta > 0$ such that

$$T_{d,n} \leqslant \delta \Gamma_{d,n}$$

*Chosing for example $\delta = 3\max(\alpha, \beta)$, under **(CL.2)** and **(CL.3)** for $a = 4$. Under **(CL.1)** with $a = 4$, by the Theorem 2.2 in [67], we get the exponential convergence*

$$\frac{\Gamma_{d,n}}{n} \xrightarrow[b_n^2]{\textit{Superexp}} \tau^4, \tag{5.103}$$

where $\tau^4 = \mathbb{E}[V_1^4]$.
Via Cauchy-Schwarz inequality and (5.102)*, it follows that*

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\Big(\frac{[\mathcal{R}_{d,n}]}{n} > \zeta\Big) &\leqslant \lim_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\Big(\frac{\Gamma_{d,n}}{n} > \frac{\zeta}{\sqrt{\delta}}\Big),\\ &= -\infty,\end{aligned} \tag{5.104}$$

where $\zeta > \tau^4\sqrt{\delta}$. *The same result can be ahieved for* $\langle \mathcal{R}_d \rangle_n / n$ *under* ***(CL.1)*** *with* $a = 2$ *and* $\zeta > \tau^4\delta$ *by exploting* (5.52) *and* (5.96). *Hence it follows from* (5.101), (5.104) *and the latter remark that*

$$\lim_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\Big(\frac{\langle \mathcal{R}_d\rangle_n + [\mathcal{R}_d]_n}{n} > b\Big) = -\infty, \tag{5.105}$$

as soon as $b > \tau^4\delta + \tau^4\sqrt{\delta}$. *The exponential convergence of* $\mathcal{R}_{d,n}/n$ *to* 0 *for all* $d = 1, \cdots, p+1$ *with speed* b_n^2 *is obtained via* (5.100) *and* (5.105), *that is, for all* $\zeta > 0$ *and* $b > \tau^4\delta + \tau^4\sqrt{\delta}$

$$\lim_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\Big(\frac{|\mathcal{R}_{d,n}|}{n} > \zeta\Big) = -\infty, \quad \textit{for} \quad \textit{all} \quad d = 1, \cdots, p+1. \tag{5.106}$$

From (5.106), *it follows that the exponential convergence*

$$\frac{1}{n} N_{1,n} \xrightarrow[b_n^2]{Superexp} 0. \tag{5.107}$$

Following the same lines as in the proofs of 5.89, 5.90 and Lemma 5.5.1, hypothesis ***(CL.2)*** *and* ***(CL.3)*** *with* $a = 4$ *together with exponential convergences* (5.107) *and* (5.96) *are sufficient to achieve the proof of Lemma 5.5.5.* $\square$

Theorem 5.5.7 (Puhalskii) *Let* $(m_j^n)_{1\leqslant j\leqslant n}$ *be a triangular array of martingale differences with values in* $\mathbb{R}^d$*,with respect tothe filtration* $(\mathcal{F}_n)_{n\geqslant 1}$. *Let* (b_n) *be a sequence of real number satisfying* (5.38). *Suppose that there exists a symmetric positive-semidefinite matrix* $\mathcal{Q}$ *such that*

$$\frac{1}{n}\sum_{k=1}^{n} \mathbb{E}\Big[m_k^n (m_k^n)' | \mathcal{F}_{k-1}\Big] \xrightarrow[b_n^2]{Superexp} \mathcal{Q}. \tag{5.108}$$

Suppose that there exists a constant $c > 0$ *such that, for each* $1 \leq k \leq n$,

$$\|m_k^n\| \leq c\frac{\sqrt{n}}{b_n} \quad \text{a.s.} \tag{5.109}$$

Suppose also that, for all $a > 0$, *we have the exponential Lindeberg's condition*

$$\frac{1}{n}\sum_{k=1}^{n} \mathbb{E}\Big[\|m_k^n\|^2 \mathrm{I}_{\Big\{\|m_k^n\| \geq a\frac{\sqrt{n}}{b_n}\Big\}} |\mathcal{F}_{k-1}\Big] \quad \xrightarrow[\lambda_n]{\text{superexp}} 0. \tag{5.110}$$

Then, the sequence

$$\left(\frac{1}{b_n\sqrt{n}}\sum_{k=1}^{n} m_k^n\right)_{n\geq 1}$$

satisfies an LDP on $\mathbb{R}^d$ *with speed* b_n^2 *and good rate function*

$$\Lambda^*(v) = \sup_{\lambda\in\mathbb{R}^d} \Big(\lambda' v - \frac{1}{2}\lambda'\mathcal{Q}\lambda\Big).$$

In particular, if $\mathcal{Q}$ is invertible,

$$\Lambda^*(v) = \frac{1}{2} v' \mathcal{Q}^{-1} v. \tag{5.111}$$

Proof 5.5.8 *The proof of Theorem 5.5.7 is contained e.g. in the proof of Theorem 3.1 in [116].*

Lemma 5.5.9 *Under **(CL.1)**, **(CL.2)** and **(CL.3)** with $a = 2$, we have for $\delta > 0$,*

$$\limsup_{R\to\infty} \limsup_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\frac{1}{n} \sum_{k=1}^{n} \|\Phi_k^p\|^2 \mathrm{I}_{\{\|\Phi_k^p\| > R\}} > \delta \right) = -\infty.$$

Remark 5.5.1 *Lemma* (5.5.9) *implies that the exponential Lindeberg's condition given by* (5.110) *is satified.*

Proof 5.5.10 *By starting, note that for any $\eta > 0$,*

$$\sum_{k=1}^{n} |X_{k-d+1}|^{2+\eta} \leqslant \alpha X_0^{2+\eta} + \beta \varepsilon_0^{2+\eta} + \beta \sum_{k=1}^{n} |V_{k-d+1}|^{2+\eta}, \tag{5.112}$$

where $L_{d,n}$ is given by (5.39)*, with $\alpha = \max_d[1 + (1 - |\theta_d|)^{-(2+\eta)}]$ and $\beta = \max_d[(1 - |\rho|)^{-(2+\eta)}(1 - |\theta_d|)^{-(2+\eta)}]$ for all $d = 1, \cdots, p+1$. This implies that, for n large enough, one can find $\zeta > 0$ such that*

$$\sum_{k=1}^{n} |X_{k-d+1}|^{2+\eta} \leqslant \zeta \sum_{k=1}^{n} |V_{k-d+1}|^{2+\eta} \tag{5.113}$$

*taking for example $\zeta = \max(\alpha, \beta)$, under **(CL.2)** and **(CL.3)** with $a = 2+\eta$. We suppose that **(CL.1)** holds with $a = 2+\eta$, then it follows that, for $R > 0$,*

$$\begin{aligned} R^\eta \sum_{k=1}^{n} \|\Phi_k^p\|^2 \mathrm{I}_{\{\|\Phi_k^p\| > R\}} &\leqslant \sum_{k=1}^{n} \|\Phi_k^p\|^{2+\eta} \\ &\leqslant p \sum_{k=1}^{n} \max_{d=1,\cdots,p+1} |X_{k-d+1}|^{2+\eta} \\ &\leqslant p\zeta \sum_{k=1}^{n} \max_{d=1,\cdots,p+1} |V_{k-d+1}|^{2+\eta} \end{aligned}$$

for n large enough and $\eta > 0$, pointed to

$$\frac{1}{b_n^2} \log \mathbb{P}\left(\frac{1}{n} \sum_{k=1}^{n} \|\Phi_k^p\|^2 \mathrm{I}_{\{\|\Phi_k^p\| > R\}} > \delta \right) \leqslant \frac{1}{b_n^2} \log \mathbb{P}\left(\frac{1}{n} \sum_{k=1}^{n} \max_{d=1,\cdots,p+1} |V_{k-d+1}|^{2+\eta} > \frac{\delta}{p\zeta} R^\eta \right).$$

Letting R to infinity and using Theorem 2.2 of [67], we immediately achieve the end of the proof of Lemma 5.5.9.

Lemma 5.5.11 *Under **(CL.1)**, **(CL.2)** and **(CL.3)**, The sequence*

$$\left(\frac{M_{1,n}}{b_n\sqrt{n}}\right)_{n\geqslant 1}$$

satisfies an LDP on $\mathbb{R}^p$ with speed b_n^2 and good rate function

$$I_p(x) = \frac{1}{2}x'.L^{-1}.x \tag{5.114}$$

where L is given by (5.41)*.*

Proof 5.5.12 *From now on, let us introduce the following modification of the martingale $(M_{1,n})_{n\geq 1}$, in order to apply Puhalskii's theorem for the moderate deviations for martingales, for $r>0$ and $R>0$,*

$$M_{1,n}^{(r,R)} = \sum_{k=1}^{n} \Phi_k^{p(r)} V_k^{(R)} \tag{5.115}$$

where, for all $1 \leq k \leq n$,

$$\Phi_n^{p(r)} = \Phi_k^p \mathrm{I}_{\left\{\|\Phi_k^p\|\leqslant r\frac{\sqrt{n}}{b_n}\right\}} \quad \text{and} \quad V_k^{(R)} = V_k \mathrm{I}_{\left\{|V_k|\leq R\right\}} - \mathbb{E}\Big[V_k \mathrm{I}_{\left\{|V_k|\leq R\right\}}\Big]. \tag{5.116}$$

We shall prove that for all $r>0$ the sequence $(M_{1,n}^{(r,R)})$ is an exponentially good approximation of $(M_{1,n})$ as R goes to infinity, for more details see e.g. Definition 4.2.14 in [53]. For this, we have to prove that for all $r>0$ and all $\delta>0$,

$$\limsup_{R\to\infty}\limsup_{n\to\infty}\frac{1}{b_n^2}\log\mathbb{P}\left(\frac{\|M_{1,n}-M_{1,n}^{(r,R)}\|}{b_n\sqrt{n}}>\delta\right) = -\infty. \tag{5.117}$$

From Lemma 5.5.5, and since $\langle M_1\rangle_n = \sigma^2 S_{n-1}$, we have

$$\frac{\langle M_1\rangle_n}{n} \xrightarrow[\lambda_n]{\text{superexp}} L. \tag{5.118}$$

From Lemma 5.5.9 and Remark (5.5.1)*, we also have for all $r>0$,*

$$\frac{1}{n}\sum_{k=0}^{n}\|\Phi_k^p\|^2 \mathrm{I}_{\left\{\|\Phi_k^p\|>r\frac{\sqrt{n}}{b_n}\right\}} \xrightarrow[\lambda_n]{\text{superexp}} 0. \tag{5.119}$$

Let us introduce the following notations,

$$\sigma_R^2 = \mathbb{E}\left[(V_1^{(R)})^2\right] \qquad \text{and} \qquad S_n^{(r)} = \sum_{k=0}^{n}\Phi_k^{p(r)}.(\Phi_k^{p(r)})'.$$

Then, it's easy to transfer properties (5.118) *and* (5.119) *to the truncated martingale* $(M_{1,n}^{(r,R)})_{n\geq 1}$. *We have for all* $R > 0$ *and all* $r > 0$,

$$\frac{\langle M_1^{(r,R)}\rangle_n}{n} = \sigma_R^2 \frac{S_{n-1}^{(r)}}{n} = -\sigma_R^2 \left(\frac{S_{n-1}}{n} - \frac{S_{n-1}^{(r)}}{n} \right) + \sigma_R^2 \frac{S_{n-1}}{n} \quad \underset{\lambda_n}{\overset{\text{superexp}}{\longrightarrow}} \sigma_R^2 \Delta_p$$

where Δ_p *is given by* (5.17), *which ensures that* (5.108) *is satisfied for the martingale* $(M_{1,n}^{(r,R)})_{n\geq 1}$. *Note also that Lemma 5.5.9 and Remark 5.5.1 work for the martinagle* $(M_{1,n}^{(r,R)})_{n\geq 1}$. *So, for all* $r > 0$, *the exponential Lindeberg's condition and thus* (5.110) *are satisfied for* $(M_{1,n}^{(r,R)})_{n\geq 1}$. *By Theorem 5.5.7, we deduce that* $(M_{1,n}^{(r,R)}/b_n\sqrt{n})$ *satisfies an LDP on* $\mathbb{R}^p$ *with speed* b_n^2 *and good rate function*

$$I_R(x) = \frac{1}{2\sigma_R^2} x'.\Delta_p^{-1}.x, \tag{5.120}$$

where Δ_p *is given by* (5.17). *It will be possible to drive the moderate deviations result for the martingale* $(M_{1,n})_{n\geq 0}$ *by proving relation* (5.117). *Although, let us now introduce the following decomposition,*

$$M_{1,n} - M_{1,n}^{(r,R)} = L_n^{(r)} + F_n^{(r,R)}$$

where

$$L_n^{(r)} = \sum_{k=1}^{n} \Big[\Phi_{k-1}^p - \Phi_{k-1}^{p(r)}\Big] V_k \quad \textit{and} \quad F_n^{(r,R)} = \sum_{k=1}^{n} \Phi_{k-1}^{p(r)} \Big(V_k - V_k^{(R)}\Big).$$

From the proofs of Lemma 4.9 in [25] together with some simple calculation technicals, it follows that

$$\frac{L_n^{(r)}}{b_n\sqrt{n}} \quad \underset{\lambda_n}{\overset{\text{superexp}}{\longrightarrow}} 0, \tag{5.121}$$

and, for all $r > 0$ *and all* $\delta > 0$, *that*

$$\limsup_{R\to\infty} \limsup_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\frac{|F_n^{(r,R)}|}{b_n\sqrt{n}} > \delta \right) = -\infty. \tag{5.122}$$

As a consequence, one can see that the sequence $(M_{1,n}^{(r,R)}/(b_n\sqrt{n}))$ *is an exponentially good approximation of* $(M_{1,n}/(b_n\sqrt{n}))$. *By application of Theorem 4.2.16 in [53], we find that* $(M_{1,n}/(b_n\sqrt{n}))$ *satifies an LDP on* $\mathbb{R}^p$ *with speed* b_n^2 *and good rate function*

$$\tilde{I}_p(x) = \sup_{\delta>0} \liminf_{R\to\infty} \inf_{z\in\mathcal{B}_{x,\delta}} I_R(z), \tag{5.123}$$

where I_R *is given by* (5.120) *and* $\mathcal{B}_{x,\delta}$ *denotes the ball* $\{z : \|z - x\| < \delta\}$. *The identification of the rate function* $\tilde{I}_p = I_p$, *where* I_p *is given by* (5.114) *is done easily, which ends the proof of Lemma 5.5.11.*

Remark 5.5.2 *Under* ***(CL.1)****,* ***(CL.2)*** *and* ***(CL.3)****, asimilar reasoning leads the exponential convergence of* $(N_{1,n}/(b_n\sqrt{n}))$ *with speed* b_n^2 *and good rate function*

$$J_p(x) = \frac{1}{2}x'.L^{-1}.x, \tag{5.124}$$

Proofs of Theorem 5.4.1, Theorem 5.4.2, Theorem 5.4.3 and Theorem 5.4.4

By refering in the decompositions (5.62), (5.83) and (5.95), the residuals terms appering there, still converge exponentially to zero Under **(CL.1)**, **(CL.2)** and **(CL.3)**, with speed b_n^2, as it was already proved. Thus, for a better readability, we may omit the most accessible part of these proofs whose development merely consists in following the same lines as those in the Proofs of Theorem 5.4.1, Theorem 5.4.2, Theorem 5.4.3 and Theorem 5.4.4, taking into account Lemma 5.5.5, Lemma 5.5.11 and remark 5.5.2, by applying the contraction principle given e.g. in [53].

Bibliographie

[1] Aït-Sahalia, Y., Fan, J., and Xui, D. High-frequency covariance estimates with noisy and asynchronous financial data. *J. Amer. Statist. Assoc. 105*, 492 (2010), 1504–1517.

[2] Aït-Sahalia, Y., and Jacod, J. *High Frequency Financial Econometrics*. Princeton University Press, 2014.

[3] Andersen, T. G., Bollerslev, T., Diebold, T. X., and Labys, P. The distribution of realized exchange rate volatility. *J. Amer. Statist. Assoc. 96* (2001), 42–55.

[4] Andersen, T. G., Bollerslev, T., Diebold, T. X., and Wu, G. Realized beta : persistence and predictability. *Adv. Econom. 20* (2008), 1–39.

[5] Arcones, M. The large deviation principle for stochastic processes in $\ell_\infty(\mathcal{F})$. *Theory of Probab. and its Appl. 47* (2004), 567–583.

[6] Asmussen, S., Fuckerieder, P., Jobmann, M., and Schwefel, H.-P. Large deviations and fast simulation in the presence of boundaries. *Stochastic Process. Appl. 102*, 1 (2002), 1–23.

[7] Avesani, R. G., and Bertrand, P. Does volatility jump or just diffuse : a statistical approach. *In. L.C.G Rogers and D. Talay (eds), Numerical Methods in Finances, Cambridge Univ. Press* (19997), 270–289.

[8] Azencott, R. Grandes déviations et applications. In *Eighth Saint Flour Probability Summer School—1978 (Saint Flour, 1978)*, vol. 774 of *Lecture Notes in Math.* Springer, Berlin, 1980, pp. 1–176.

[9] B. Bercu, B., Gamboa, F., and Lavielle, M. Sharp large deviations for gaussian quadratic forms with applications. *ESAIM Probab. Statist. 4* (2000), 1–24.

[10] B. Bercu, B., Gamboa, F., and Rouault, A. Large deviations for quadratic forms of stationary gaussian processes. *Stochastic Process. Appl. 71*, 1 (1997), 75–90.

[11] Bannon, G., and Nguyen, H. T. Recursive estimation in diffusions models. *SIAM J. Control Optim. 19* (1987), 676–685.

[12] Barndorff-Nielsen, O. E., Graversen, S., Jacod, J., Podolskij, M., and Shephard, N. A central limit theorem for realised power and bipower variations of continuous semimartingales. *Springer, Berlin* (2006), 33–68.

[13] Barndorff-Nielsen, O. E., Graversen, S., Jacod, J., and Shephard, N. Limit theorems for bipower variation in financial econometrics. *Econometric Theory 22*, 4 (2006), 677–719.

[14] Barndorff-Nielsen, O. E., Graversen, S., and Shephard, N. Power variation and stochastic volatility : a review and some new results. *J. Appl. Probab. 41A* (2004), 133–143.

[15] Barndorff-Nielsen, O. E., and Shephard, N. Econometric analysis of realized volatility and its use in estimating volatility models. *J. R. Stat. Soc. Ser. B. Stat. Methodol. 64*, 2 (2002), 253–280.

[16] Barndorff-Nielsen, O. E., Shephard, N., and Winkel, M. Limit theorems for multipower variation in the presence of jumps. *Stochastic Process. Appl. 116* (2006), 796–806.

[17] Basawa, I. V., and Brockwell, P., J. Non-parametric estimation for non-decreasing levy processes. *Journal of the Royal Statistical Society. Series B (Methodological) 44*, 2 (1982), 262–269.

[18] Bercu, B. On the convergence of moments in the almost sure central limit theorem for martingales with statistical applications. *Stochastic. Process. Appl. 111* (2004), 157–173.

[19] Bercu, B., and Proïa, F. A sharp analysis on the asymptotic behavior of the durbin-watson statistic for the first-order autoregressive process. *ESAIM. Proba. Stat. 16* (2013), 500–530.

[20] Bercu, B., and Touati, A. Exponential inequalities for self-normalized martingales with applications. *Ann. Appl. Probab. 18*, 5 (2008), 1848–1869.

[21] Bertrand, P. Quelques applications des processus stochastiques : contrôle adaptatif, statistique des processus, détection de ruptures...,. *HDR Université Blaise Pascal, Clermont–Ferrand* (1997).

[22] Bibinger, M. An estimator for the quadratic covariation of asynchronously observed Itô processes with noise : asymptotic distribution theory. *Stochastic Process. Appl. 122*, 6 (2012), 2411–2453.

[23] Bibinger, M., and Reiss, M. Spectral estimation of covolatility from noisy observations using local weights. *Scand. J. Stat. 41*, 1 (2014), 23–50.

[24] Bitseki Penda, S. V., Djellout, H., and Proïa, F. Moderate deviations for the Durbin-Watson statistic related to the first-order autoregressive process. *ESAIM Probab. Stat. 18* (2014), 308–331.

[25] Bitseki-Penda, V., Djellout, H., and Proïa, F. Moderate deviations for the durbin-watson statistic to the first-order autoregressive process. *ESAIM. Proba. Stat. 18* (2012), 308–331.

[26] Blattberg, R. C. Evaluation of the power of the durbin-watson statistic for non-first order serial correlation alternatives. *The Review of Economics and Statistics 55*, 1973 (4), 508–515.

[27] Borovkov, A. A. Boundary value problems for random walks and large deviations in function spaces. *Theor. Probability Appl. 12* (1967), 575–595.

[28] Borovkov, A. A., and Mogul′skiĭ, A. A. Large deviations and testing statistical hypotheses. I. Large deviations of sums of random vectors. *Siberian Adv. Math. 2*, 3 (1992), 52–120.

[29] Borovkov, A. A., and Mogul′skiĭ, A. A. Large deviations and testing statistical hypotheses. II. Large deviations of maximum points of random fields. *Siberian Adv. Math. 2*, 4 (1992), 43–72.

[30] Borovkov, A. A., and Mogul′skiĭ, A. A. Large deviations and testing statistical hypotheses. III. Asymptotically optimal tests for composite hypotheses. *Siberian Adv. Math. 3*, 1 (1993), 19–86.

[31] Borovkov, A. A., and Mogul′skiĭ, A. A. Large deviations and testing statistical hypotheses. IV. The statistical invariance principle and the laws of conservation. *Siberian Adv. Math. 3*, 2 (1993), 14–80.

[32] Box, G. E. P., and Ljung, G. M. On a measure of a lack of fit in time series models. *Biometrika 65*, 2 (1978), 297–303.

[33] Box, G. E. P., and Pierce, D. A. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Am. Stat. Assn. Jour. 65*, 2 (1970), 1509–1526.

[34] Brandt, M. W., and Diebold, F. X. A no–arbitrage approch to range–basedestimation of covariances and correlations. *Penn. Institue for Economic Research Working Paper*, 03–013.

[35] Breusch, T. Testing for autocorrelation in dynamic linear models. *Australian Economic Papers. 17*, 31 (1978), 334–355.

[36] Bryc, W., and Dembo, A. Large deviations for quadratic functionals of gaussian processes. *J. Theoret. Probab. 10*, 2 (1997), 307–322.

[37] Chang, C. S. Sample path large deviations and intree networks. *Queueing Systems Theory Appl. 20*, 1-2 (1995), 7–36.

[38] Chen, X. Moderate deviations for m-dependent random variables with banach space values. *Statist. Probab. Lett. 35*, 2 (1997), 123–134.

[39] Chernov, H. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Stat. 23* (1952), 493–507.

[40] Chernov, H. Large sample théory : parametric case. *Ann. Math. Stat. 27* (1956), 1–22.

[41] Christensen, K., Podolskij, M., and Vetter, M. On covariation estimation for multivariate continuous Itô semimartingales with noise in non-synchronous observation schemes. *J. Multivariate Anal. 120* (2013), 59–84.

[42] Chu, B., Knight, J., and Satchell, S. Large deviations theorems for optimal investment problems with large portfolios. *European J. Oper. Res. 211*, 3 (2011), 533–555.

[43] Comte, F., Genon-Catalot, V., and Rozenholc, Y. Nonparametric estimation for a stochastic volatility model. *Finance Stoch. 14*, 1 (2010), 49–80.

[44] Cramèr, H. Sur un nouveau théorème limite de la théorie des probabilités. *Actualités scientifiques et industrielles. Colloque consacré à la théorie des probabilités.*, 763 (1938), 5–23.

[45] Dacunha-Castelle, D. Introduction au séminaire orsay. *In Asterisque. 68* (1978), Springer–Verlag.

[46] Dacunha-Castelle, D., and Florens-Zmirou, D. Estimation of the coefficient of a diffusion from discrete observations. *Stochastics 19* (1986), 263–284.

[47] Dalalyan, A., and Yoshida, N. Second-order asymptotic expansion for a non-synchronous covariation estimator. *Ann. Inst. Henri Poincaré Probab. Stat. 47*, 3 (2011), 748–789.

[48] de Acosta, A. Large deviations for vector-valued Lévy processes. *Stochastic Process. Appl. 51*, 1 (1994), 75–115.

[49] de Acosta, A., C. X. Moderate deviations for empirical measure of markov chains : Upper bound. *J. Theor. Probab. 4*, 11 (1998), 075–110.

[50] Dembo, A. Moderate deviations for martingales with bounded jumps. *Electron. Comm. Probab. 1*, 3 (1996), v.

[51] Dembo, A., and Zajic, T. Large deviations : from empirical mean and measure to partial sums process. *Stochastic Process. Appl. 57*, 2 (1995), 191–224.

[52] Dembo, A., and Zajic, T. Uniform large and moderate deviations for functional empirical processes. *Stochastic Process. Appl. 67*, 2 (1997), 195–211.

[53] 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.

[54] Djellout, H.and Guillin, A. Large and moderate deviations and moderate deviations for moving average process. *Annales de la Faculté des sciences de Toulouse : Mathématiques 10*, 1 (2001), 23–31.

[55] Djellout, H.and Guillin, A. A moderate deviations for markov chains with atom. *Stochastic. Process. Appl. 95*, 2 (2001), 203–217.

[56] Djellout, H. Moderate deviations for martingale differences and applications to ϕ-mixing sequences. *Stoch. Stoch. Rep. 73*, 1-2 (2002), 37–63.

[57] Djellout, H., and Guillin, A. Large and moderate deviations for moving average process. *Ann. Fac. Sci. Toulouse Math. 10*, 1 (2001), 23–31.

[58] Djellout, H., Guillin, A., and Samoura, Y. Estimation of the realized (co-)volatility vector : Large deviations approach. *Soumis*, 1–34.

[59] Djellout, H., Guillin, A., and Wu, L. Large and moderate deviations for estimators of quadratic variational processes of diffusions. *Stat. Inference Stoch. Process.*, 2 (1999), 195–225.

[60] Djellout, H., Jiang, H., Guillin, A., and Samoura, Y. Moderate deviations for bipower variation of general function and hayashi-yoshida estimators. *Preprint*.

[61] Djellout, H., and Samoura, Y. Large and moderate deviations of realized covolatility. *Statist. Probab. Lett. 86* (2014), 30–37.

[62] Donsker, M. D., and Varadhan, S. R. S. Large deviations for stationary Gaussian processes. *Comm. Math. Phys. 97*, 1-2 (1985), 187–210.

[63] Dovonon, P., Gonçalves, S., and Meddahi, N. Bootstrapping realized multivariate volatility measures. *J. Econometrics 172*, 1 (2013), 49–65.

[64] Durbin, J., and Watson, G. Testing for serial correlation in least squares regression. I. *Biometrika 37* (1950), 409–428.

[65] Durbin, J., and Watson, G. Testing for serial correlation in least squares regression. II. *Biometrika 38* (1951), 159–178.

[66] Durbin, J., and Watson, G. Testing for serial correlation in least squares regression. III. *Biometrika 58* (1971), 1–19.

[67] Eichelbacher, P., and Löwe, M. Moderate deviations for i.i.d random variables. *Probability. and Statistics. 7*, 3 (2003), 209–218.

[68] Eichelsbacher, P., and O'Connell, N. Sample path large deviations in finer topologies. *Stochastics Stochastics Rep. 67*, 3-4 (1999), 231–254.

[69] Ellis, R. Large deviations for a general class of random vectors. *Ann. Probab. 12* (1984), 1–12.

[70] Florens-Zmirou, D. On estimating the variance of diffusion processes. *J. Appl.Prob. 30* (1993), 790–804.

[71] Freidlin, M. I., and Wentzell, A. D. *Random perturbations of dynamical systems*, third ed., vol. 260 of *Grundlehren der Mathematischen Wissenschaften.* Springer, Heidelberg, 2012.

[72] Gamboa, F., Rouault, A., and Zani, M. A functional large deviations principle for quadratic forms of Gaussian stationary processes. *Statist. Probab. Lett. 43*, 3 (1999), 299–308.

[73] Gao, F., and Zhao, X. Delta method in large deviations and moderate deviations for estimators. *Ann. Statist. 39*, 2 (2011), 1211–1240.

[74] Gao, F.-Q. Moderate deviations for martingales and mixing random processes. *Stochastic Process. Appl. 61*, 2 (1996), 263–275.

[75] Gärtner, J. On large deviations from the invariant measure. *Theory of Prob. and Appl. 22* (1977), 24–39.

[76] Gloter, A. Efficient estimation of drift parameters in stochastic volatility models. *Finance Stoch. 11*, 4 (2007), 495–519.

[77] Guillin, A. Grandes déviations de systèmes dynamiques. *PhD thesis, Université Blaise Pascal* (2000), 1–222.

[78] Hayashi, T., and Yoshida, N. On covariance estimation of non-synchronously observed diffusion processes. *Bernoulli 11* (2005), 359–379.

[79] Hayashi, T., and Yoshida, N. Asymptotic normality of a covariance estimator for non-synchronously observed diffusion processes. *Annals. of the Institue. of Statistical. Mathematics. 60* (2008), 367–406.

[80] Hayashi, T., and Yoshida, N. Nonsynchronous covariation process and limit theorems. *Stochastic Process. Appl. 121*, 10 (2011), 2416–2454.

[81] Jacod, J. Asymptotic properties of realized power variations and related functionals of semimartingales. *Stochastic Process. Appl. 118*, 4 (2008), 517–559.

[82] Jacod, J., and Protter, P. Asymptotic error distributions for the Euler method for stochastic differential equations. *Ann. Probab. 26*, 1 (1998), 267–307.

[83] Jacod, J., and Todorov, V. Efficient estimation of integrated volatility in presence of infinite variation jumps. *Ann. Statist. 42*, 3 (2014), 1029–1069.

[84] Jiang, H. Moderate deviations for estimators of quadratic variational process of diffusion with compound poisson. *Statist. Probab. Lett. 80*, 17-18 (2010), 1297–1305.

[85] Kanaya, S., and Otsu, T. Large deviations of realized volatility. *Stochastic Process. Appl. 122*, 2 (2012), 546–581.

[86] Kanaya, S., and Otsu, T. Corrigendum to “Large deviations of realized volatility” [Stochastic Process. Appl. 122 (2012) 546–581]. *Stochastic Process. Appl. 123*, 3 (2013), 1176–1177.

[87] Kinnebrock, S., and Podolskij, M. Asymptotic a note on the central limit theorem for bipower variation of general functions. *Stochastic Process. Appl. 118*, 6 (2008), 1056–1070.

[88] Kutoyants, Y. Parameter estimationfor stochastic process. *Heldermann* (1984).

[89] Lanska, V. Minimum contrast estimation in diffusion process. *J. Appl. Prob. 16* (1979), 65–75.

[90] Ledoux, M. Sur les déviations modérées des sommes de variables aléatoires vectorielles indépendantes de même loi. *Ann. Inst. Henri Poincaré Probab. Stat. 35* (1992), 123–134.

[91] Ledoux, M. Concentration of measure and logarithmic sobolev inequalities. *Séminaire de Probabilités, XXXIII, Lecture Notes in Math. Springer, Berlin*, 1709 (1999), 120–216.

[92] Liming, W. Moderate deviations of dependent randomvariables related to clt. *Ann.Probab. 23* (1995), 420–445.

[93] Liming, W. On large deviations for moving average processes. *Probability. finance and insurance.* (2004), 15–49.

[94] Linde, W., and Shi, Z. Evaluating the small deviation probabilities for subordinated lévy processes. *Stochastic Process. Appl. 113* (2004), 273–287.

[95] Lo, A. W., and MacKinlay, A. C. An econometric analysis of nonsynchronous trading. *J. Econometrics 45* (1990), 181–211.

[96] Lundin, M. C., and Müller, U. A. Correlation of high frequency financial time series. *IN P. Lequex (eds), The financial Markets Tick by Tick* (1999), 91–126.

[97] Lynch, J., and Sethuraman, J. Large deviations for processes with independent increments. *Ann. Probab. 15*, 2 (1987), 610–627.

[98] Maïda, M., Najim, J., and Péché, S. Large deviations for weighted empirical mean with outliers. *Stochastic Process. Appl. 117*, 10 (2007), 1373–1403.

[99] Mancini, C. Large deviation principle for an estimator of the diffusion coefficient in a jump-diffusion process. *Statist. Probab. Lett. 78*, 7 (2008), 869–879.

[100] Mancini, C., and Gobbi, F. Identifying the brownian covariation from the co-jumps given discrete observations. *Econometric Theory 28*, 2 (2012), 249–273.

[101] Mogulskii, A. A. Large deviations for the trajectories of multidimensional random walks. *Theor. Probability Appl. 21*, 2 (1976), 300–315.

[102] Mogulskii, A. A. Large deviations for processes with independent increments. *Ann. Probab. 21*, 1 (1993), 202–215.

[103] Muthuswamy, J., S. S. L. A., and Terry, E. Time variation in the correlation structure of exchange rate : High–frequency analysis. *J. Futurs. Markets. 21* (2001), 127–144.

[104] Najim, J. Grandes déviations pour certaines mesures empiriques. *PhD thesis, Université Nanterre Paris-10* (2001), 1–130.

[105] Najim, J. A Cramér type theorem for weighted random variables. *Electron. J. Probab. 7*, 4 (2002), 1–32.

[106] Najim, J. Large deviations for independent random variables - Application to Erdős-Renyi's functional law of large numbers. *ESAIM Probab. Stat. 9* (2005), 116–142 (electronic).

[107] Nourdin, I., Nualart, D., and Tudor, C. A. Central and non-central limit theorems for weighted power variations of fractional brownian motion. *Ann. Inst. Henri Poincaré Probab. Stat. 46*, 4 (2010), 1055–1079.

[108] Perrin, O., and Zani, M. Large deviations for sample paths of Gaussian processes quadratic variations. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) 328*, Veroyatn. i Stat. 9 (2005), 169–181, 280.

[109] Pham, H. Some applications and methods of large deviations in finance and insurance. In *Paris-Princeton Lectures on Mathematical Finance 2004*, vol. 1919 of *Lecture Notes in Math.* Springer, Berlin, 2007, pp. 191–244.

[110] Phan Dinh, T. D. Non–parametric estimation of the drift coefficient in the diffusion equation. *Sem. Statist. Univ. Grenoble*, 103–129.

[111] Plachky, D., and Steinebach, J. A theorem about probabilities of large deviations with an application to queuing theory. *Periodica. Mathematica. Hungarica. 6* (1975), 343–345.

[112] Proïa, F. Further results on the h-test of durbin for stable autoregressive processes. *Multivariate. Anal. 118* (2013), 77–101.

[113] Proïa, F. Testing for residual correlation of any order in the autoregressive process. *ArXiv* (2013), 1–30.

[114] Puhalskii, A. Large deviations of semimartingales via convergence of the predictable characteristics. *Stochastics Stochastics Rep. 49*, 1-2 (1994), 27–85.

[115] Puhalskii, A. Large deviation analysis of the single server queue. *Queueing Systems Theory Appl. 21*, 1-2 (1995), 5–66.

[116] Puhalskii, A. Large deviations of semimartingales : a maxingale problem approach. i. limits as solutions to a maxingale problem. *Stoch. Stoch. Rep. 61*, 3-4 (1997), 141–243.

[117] Revuz, C., and Yor, M. *Continuous martingales and Brownian motion. Third edition.* Springer-Verlag, Berlin, 1999.

[118] Robert, C. Y., and Rosenbaum, M. Volatility and covariation estimation when microstructure noise and trading times are endogenous. *Math. Finance 22*, 1 (2012), 133–164.

[119] Robertson, S. Sample path large deviations and optimal importance sampling for stochastic volatility models. *Stochastic Process. Appl. 120*, 1 (2010), 66–83.

[120] Rockafellar, R. T. *Convex analysis.* Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.

[121] Rockafellar, R. T. Integrals which are convex functionals. II. *Pacific J. Math. 39* (1971), 439–469.

[122] Samoura, Y. Moderate deviations for the durbin-watson statistic associated to the stable p-order autoregressive process. *Preprint*.

[123] Schied, A. Cramér's condition and Sanov's theorem. *Statist. Probab. Lett. 39*, 1 (1998), 55–60.

[124] Scholes, M., and Williams, J. Estimating beta from nonsynchronous data. *J. Financial. Economics. 5* (1997), 309–327.

[125] Shen, S. Large deviation for the empirical correlation coefficient of two gaussian random variables. *Acta Math. Sci. Ser. B Engl. Ed. 27*, 4 (2007), 821–828.

[126] Stutzer, M. Portfolio choice with endogenous utility : a large deviations approach. *J. Econometrics 116*, 1-2 (2003), 365–386. Frontiers of financial econometrics and financial engineering.

[127] Tillman, J. A. The power of durbin-watson test. *Econometrica 43*, 5-6 (1975), 59–974.

[128] Todorov, V., and Bollerslev, T. Jumps and betas : a new framework for disentangling and estimating systematic risks. *J. Econometrics 157*, 2 (2010), 220–235.

[129] Tsay, R. S., and Yeh, J.-H. Non–synchronous trading and high–frequency beta. *Preprint* (2003).

[130] Varadhan, S. R. S. Asymptotic probabilities and differential equations. *Comm. Pure Appl. Math. 19* (1966), 261–286.

[131] Vetter, M. *Limit theorems for bipower variation of semimartingales.*, vol. 118. 2008.

[132] Vetter, M. Estimation of correlation for continuous semimartingales. *Scand. J. Stat. 39*, 4 (2012), 757–771.

[133] Von-Neumann, J. Distribution of the ratio of the mean square successive difference to the variance. *Ann. Math. Stat. 12* (1941), 367–395.

[134] Worms, J. Principes de déviations modérées pour des martingales et applications statistiques. *Thèse de doctorat á l'Université Marne-la-Vallée.* (2000).

[135] Worms, J. Moderate deviations for stable markov chains and regression models. *Electron. J. Probab. 4*, 8 (2001), 1–28.

[136] Worms, J. Moderate deviations of some dependent variables. I. martingales. *Math. Methods Statist. 10*, 1 (2001), 8–72.

[137] Worms, J. Moderate deviations of some dependent variables. II. some kernel estimators. *Math. Methods Statist. 10*, 2 (2001), 161–193.

[138] Wu, L. An introduction to large deviation. In : J. A. Yan, S. Peng, S. Fang and L. Wu (eds). *Several topics in stochastic analysis Academic Press of China, Beijing* (1997), 225–336.

[139] Zani, M. Grandes déviations pour des fonctionnelles issues de la statistique des processus. *PhD thesis, Université Paris-Sud*, 4 (1999), 1–151.

[140] Zani, M. Sample path large deviations for squares of stationary Gaussian processes. *Theory Probab. Appl. 57*, 2 (2013), 347–357.

[141] Zhang, L., Mykland, P. M., and Aït-Sahalia, Y. A tale of two time scales : determining integrated volatility with noisy high-frequency data. *J. Amer. Statist. Assoc. 100*, 472 (2005), 1394–1411.

[142] Zhang, L., Mykland, P. M., and Aït-Sahalia, Y. Edgeworth expansions for realized volatility and related estimators. *J. Econometrics 160* (2011), 190–203.

RÉSUMÉ

Cette thèse est consacrée à l'étude de théorèmes limites : grandes déviations et déviations modérées pour des estimateurs liés à des modèles financiers.

Dans une prémière partie, nous nous sommes intéressés à l'étude des déviations grandes et modérées des estimateurs de la covariation et de la (co)volatilité réalisée issus des fonctionnelles associées à deux processus de diffusion couplés de manière synchronisée. Les techniques utilisées dans ces travaux sont basées d'une part sur celles utilisées dans Djellout-Guillin-Wu et sur la sous additivité et sur la notion d'approximation exponentielle inspirées des travaux de J. Najim d'autre part.

Dans une deuxième partie, on considère que les deux processus de diffusion sont observés de manière non synchronisée et on établit des déviations modérées pour l'estimateur de la variation généralisée et pour celui de Hayashi-Yoshida. Les résultats sont obtenus par l'utilisation d'une nouvelle approche sur les déviations modérées des variablesaléatoires $m-$dépendantes vériafiant des conditions de type "Chen-Ledoux".

Dans la troisième et dernière partie, on s'intéresse à l'étude processus autorégressif d'ordre p dont le bruit est un processus autorégressif d'ordre q. On montre des déviations modérées pour certains estimateurs associés à notre modèle dont la statistique de Durbin-Watson. Les résultats sont donnés dans le cas où le bruit est gaussien puis dans le cas de condition de type "Chen-Ledoux" portant sur le bruit.

Mots clés : Grandes déviations et déviations modérées, diffusions synchronisées et non-synchronisées, covariation réalisée, processus autorégressif.

ABSTRACT

This thesis is devoted to the study of the limits theorem : large and moderate déviations for somes financial mathematicals estimators.

In the first part, we studied the large and moderate deviations of the estimators of covariation and the realized (co)volatility obtained from the functional associated to two diffusion processes coupled in synchronous manner. The techniques used in this work are based, on the one hand, on those used in Djellout-Guillin-Wu and the subadditivity and the exponential approximation notion inspired by J. Najim results on the other hand.

In the second part, we consider that ours two diffusion processes are observed in a nonsynchronized manner and on the establish the moderate deviations for the generalised bipower variation estimator and the Hayashi-Yoshida estimator. The results are obtained by using a new approach on the moderate deviations of the $m-$dependent random variables based on the Chen-Ledoux type condition.

In the third and last part, we study the stable autoregressive process of order p where the driven noise is also given by a q-order autoregressive process. We prove the moderate deviations for some estimators associated with our model such as the Durbin-Watson statistic. The results are given in the case where the driven noise is the normally distributed then in the case where the driven noise satisfy a Chen-Ledoux type condition.

keywords : Large and Moderate deviations principles, synchrone and asynchrone diffusions, quadratic variation, realised volatility, autoregressive process.