N° d'ordre : 4791

Thèse

présentée à

L'UNIVERSITÉ BORDEAUX 1

ÉCOLE DOCTORALE DE MATHÉMATIQUES ET INFORMATIQUE

par **Vassili Blandin**

POUR OBTENIR LE GRADE DE

DOCTEUR

SPÉCIALITÉ : **Mathématiques Appliquées**

Estimation de paramètres pour des processus autorégressifs à bifurcation

---

Soutenue le 26 Juin 2013 à l'Institut de Mathématiques de Bordeaux
devant la commission d'examen composée de

M. Vincent Bansaye	Professeur chargé de cours, Ecole Polytechnique	Examinateur
M. Bernard Bercu	Professeur, Université Bordeaux 1	Directeur de thèse
M. Benoît Cadre	Professeur, ENS Cachan, Antenne de Bretagne	Rapporteur
M. Jean-François Delmas	Professeur, Ecole Nationale des Ponts et Chaussées	Rapporteur
M. Mihai Gradinaru	Professeur, Université Rennes 1	Président du jury Examinateur
M. Marc Hoffmann	Professeur, Université Paris-Dauphine	Examinateur

Estimation de paramètres pour des processus autorégressifs à bifurcation

RÉSUMÉ

Les processus autorégressifs à bifurcation (BAR) ont été au centre de nombreux travaux de recherche ces dernières années. Ces processus, qui sont l'adaptation à un arbre binaire des processus autorégressifs, sont en effet d'intérêt en biologie puisque la structure de l'arbre binaire permet une analogie aisée avec la division cellulaire. L'objectif de cette thèse est l'estimation les paramètres de variantes de ces processus autorégressifs à bifurcation, à savoir les processus BAR à valeurs entières et les processus BAR à coefficients aléatoires.
Dans un premier temps, nous nous intéressons aux processus BAR à valeurs entières. Nous établissons, via une approche martingale, la convergence presque sûre des estimateurs des moindres carrés pondérés considérés, ainsi qu'une vitesse de convergence de ces estimateurs, une loi forte quadratique et leur comportement asymptotiquement normal. Dans un second temps, on étudie les processus BAR à coefficients aléatoires. Cette étude permet d'étendre le concept de processus autorégressifs à bifurcation en généralisant le côté aléatoire de l'évolution. Nous établissons les mêmes résultats asymptotiques que pour la première étude. Enfin, nous concluons cette thèse par une autre approche des processus BAR à coefficients aléatoires où l'on ne pondère plus nos estimateurs des moindres carrés en tirant parti du théorème de Rademacher-Menchov.

Mots-clés : processus autorégressif à bifurcation ; processus à valeurs entières ; coefficient aléatoire ; moindres carrés pondérés ; martingale ; convergence presque sûre ; théorème limite central.

Parameter estimation for bifurcating autoregressive processes

ABSTRACT

Bifurcating autoregressive (BAR) processes have been widely investigated this past few years. Those processes, which are an adjustment of autoregressive processes to a binary tree structure, are indeed of interest concerning biology since the binary tree structure allows an easy analogy with cell division. The aim of this thesis is to estimate the parameters of some variations of those BAR processes, namely the integer-valued BAR processes and the random coefficients BAR processes.
First, we will have a look to integer-valued BAR processes. We establish, via a martingale approach, the almost sure convergence of the weighted least squares estimators of interest, together with a rate of convergence, a quadratic strong law and their asymptotic normality. Secondly, we study the random coefficients BAR processes. The study allows to extend the principle of bifurcating autoregressive processes by enlarging the randomness of the evolution. We establish the same asymptotic results as for the first study. Finally, we conclude this thesis with an other approach of random coefficient BAR processes where we do not weight our least squares estimators any more by making good use of the Rademacher-Menchov theorem.

Key words : bifurcating autoregressive process ; integer-valued process ; random coefficient ; weighted least squares ; martingale ; almost sure convergence ; central limit theorem.

Institut de Mathématiques de Bordeaux
Université de Bordeaux 1
UMR CNRS 5251
351 cours de la libération
33405 Talence cedex, France

Remerciements

Je tiens à remercier mon directeur de thèse Bernard Bercu qui est à l'origine de cette thèse. Son enthousiasme, sa bonne humeur et sa présence dès que nécessaire, tout en me laissant une grande liberté m'a rendu cette thèse plus douce.

Je souhaite également remercier Benoît Cadre et Jean-François Delmas de m'avoir fait l'honneur de rapporter ma thèse, ainsi que Vincent Bansaye, Mihai Gradinaru et Marc Hoffman d'avoir accepté d'être présents dans mon jury.

J'en profite pour remercier ma famille pour son soutien sans relâche depuis toujours sans lequel je n'en serais pas là, avec en particulier Marion qui a su rassembler toute sa patience pendant cet énième exil.

Enfin merci aux co-bureaux pour toutes ces discussions et à Julien pour ces deux années.

Table des matières

Chapitre 1

Introduction

Cette introduction est composée de trois parties. La première partie est consacrée à un bref panorama dur les processus autorégressifs, les processus autorégressifs à coefficients aléatoires ou à valeurs entières, puis les processus autorégressifs à bifurcation. La seconde partie est dédiée à la présentation des principaux résultats de la thèse. La troisième partie porte sur quelques résultats de convergence des martingales vectorielles à temps discret.

1.1 Les processus autorégressifs

Un processus (X_n) est dit autorégressif d'ordre p, noté AR(p), s'il est défini, pour tout $n \geq p$, par la relation de récurrence

$$X_n = a_1 X_{n-1} + a_2 X_{n-2} + \ldots + a_p X_{n-p} + \varepsilon_n$$

où les variables aléatoires initiales $X_0, X_1, \ldots, X_{p-1}$ sont fixées arbitrairement. Les valeurs $a_1, a_2, \ldots, a_p$ sont les paramètres de ce processus AR(p), tandis que (ε_n) est un bruit blanc associé à (X_n), c'est-à-dire une suite de variables aléatoires indépendantes et de même loi centrées et de carré intégrable. Ces processus permettent la modélisation de divers phénomènes où chaque état dépend linéairement des p états précédents. Ces processus ont été largement étudié et font toujours l'objet de recherches en partie via des variations dans la définition. Donnons ici quelques résultats bien connus sur les processus autorégréssifs que l'on peut retrouver, par exemple, dans les livres de Box et al. [14] ou de Brockwell et Davis [15].

Proposition 1.1.1. *Notons A le polynôme suivant*

$$A(X) = 1 - \sum_{k=1}^{p} a_k X^k.$$

Alors, le processus autorégressif (X_n) est asymptotiquement stationnaire si et seulement si le polynôme A a toutes ses racines à l'extérieur du disque unité.

Définition 1.1.2. *Soit Λ_n^X l'enveloppe linéaire engendrées par $\{X_2, X_3, \dots, X_n\}$ et soit Π_n la projection orthogonale sur Λ_n^X. On définit alors la fonction d'autocorrélation partielle $(\alpha(n))$ associée à (Y_n) par*

$$\begin{cases} \alpha(0) = 1, \qquad \alpha(1) = Cor(X_1, X_2), \\ \alpha(n) = Cor(X_{n+1} - \Pi_n(X_{n+1}), X_1 - \Pi_n(X_1)). \end{cases}$$

$\alpha(n)$ correspond à la corrélation entre X_{n+1} et X_1 conditionnellement à $X_2, X_3, \dots, X_n$.

Théorème 1.1.3. *Une série chronologique centrée régulière (X_n) est un processus autorégressif d'ordre p si et seulement si sa fonction d'autocorrélation partielle $\alpha(n)$ vérifie*

$$\alpha(p) \neq 0 \qquad et \qquad \alpha(k) = 0, \ \forall k \geq p+1.$$

Nous allons ici introduire trois variations des processus autorégressifs qui apparaitront dans les chapitres suivants.

1.1.1 Les processus autorégressifs à coefficients aléatoires

La première variation que l'on va présenter ici est celle qui a la définition la plus proche des processus AR(p). Il s'agit des processus autorégressifs à coefficients aléatoires d'ordre p, notés RCAR(p), qui seront l'objet des Chapitres 3 et 4. Les premières études sur ces processus remontent à Robinson [48] et à Nicholls et Quinn [44, 46] à la fin des années 1970. Un processus RCAR(p) est défini, pour tout $n \geq p$, par la relation de récurrence

$$X_n = a_{1,n} X_{n-1} + a_{2,n} X_{n-2} + \ldots + a_{p,n} X_{n-p} + \varepsilon_n$$

où, pour tout $i \in \{1, 2, \dots, p\}$, $(a_{i,n})_{n \geq p}$ est une suite de variables aléatoires indépendantes et identiquement distribuées. Ces coefficients aléatoires sont également supposés indépendants de la suite indépendante et identiquement distribuée des bruits (ε_n). Tout comme l'étude des processus AR(p), le cas $p = 1$ et le cas $p \neq 1$ se traitent de façon similaire. En effet, il suffit de vectorialiser le processus RCAR(p) pour obtenir une écriture analogue à un processus RCAR(1) et ainsi, à une adaptation des outils au cas vectoriel près, transposer la méthode utilisée pour les RCAR(1).

Parmi les études des processus RCAR(p), on peut citer les travaux de Koul et Schick [39] ou de Aue et al. [3] qui estiment les paramètres d'un processus RCAR(1) via un M-estimateur et une approche à base de quasi-maximum de vraisemblance, respectivement. Un cas critique du processus RCAR(1) où les variables aléatoires $(a_{1,n})$ suivent des lois de Rademacher a également été traité par Hwang et al. [35] grâce à un estimateur des moindres carrés. On peut par ailleurs évoquer Hwang et Basawa [34] qui estiment les paramètres d'un processus RCAR(p) généralisé, c'est-à-dire où ils autorisent une corrélation entre les coefficients de la partie linéaire de la récurrence et le bruit.

La technique que l'on utilisera par la suite consiste à transformer l'équation de la récurrence pour la réécrire sous la forme d'un processus autorégressif en introduisant un nouveau bruit. Cette technique a, par exemple, déjà été utilisée par Jurgens [36].

1.1.2 Les processus autorégressifs à valeurs entières

La deuxième variation des processus autorégressifs, que l'on retrouvera dans le Chapitre 2, concerne les processus autorégressifs à valeurs entières d'ordre p, notés INAR(p). Ces processus ont été introduits par Al-Osh et Alzaid [1, 2] et McKenzie [40] dans la deuxième moitié des années 1980. La définition d'un tel processus (X_n) a la même forme pour les processus AR(p) mais est adaptée afin de s'assurer que, pour tout $n \geq 0$, X_n soit une variable aléatoire à valeurs entières. Un processus INAR(p) est défini, pour tout $n \geq p$, par

$$X_n = a_1 \circ X_{n-1} + a_2 \circ X_{n-2} + \ldots + a_p \circ X_{n-p} + \varepsilon_n$$

où les variables aléatoires initiales $X_0, X_1, \ldots, X_n$ sont des variables aléatoires à valeurs entières positives et où (ε_n) est une suite de variables aléatoires à valeurs entières positives. De plus, pour tout $i \in \{1, 2, \ldots, p\}$, la variable aléatoire $a_i \circ X_k$ est définie par

$$a_i \circ X_k = \sum_{l=1}^{X_k} Y_{i,k,l}$$

où $(Y_{i,k,l})_{k\in\mathbb{N},l\in\mathbb{N}}$ est une suite indépendante et identiquement distribuée de variables aléatoires à valeurs entières positives. Le paramètre a_i est en général un paramètre permettant d'identifier la distribution commune des variables aléatoires $(Y_{i,k,l})_{k\in\mathbb{N},l\in\mathbb{N}}$. La suite $(Y_{i,k,l})_{k\in\mathbb{N},l\in\mathbb{N}}$ est souvent considérée comme suivant la loi de Bernoulli de paramètre a_i comme, par exemple, dans Al-Osh et Alzaid [1, 2], Barczy et al. [7, 8] ou encore Enciso-Mora et al. [24]. Cette hypothèse supplémentaire, qui ne sera pas faite dans le Chapitre 2, entraîne que la loi de $a_i \circ X_k$ sachant X_k est la loi binomiale $\mathcal{B}(X_k, a_i)$. A noter que contrairement aux processus AR ou RCAR, le cas $p \neq 1$ est beaucoup plus compliqué que le cas $p = 1$. En effet, l'opérateur $\circ$ faisant apparaître une somme, la technique consistant à vectorialiser le problème ne fonctionne plus, les dépendances entre les différents états du processus ne sont plus appropriées pour cette approche.

De nombreuses études ont été faites concernant les processus INAR, avec entre autre Al-Osh et Alzaid [2] qui étudient la loi limite de ce processus, Silva et Silva [50] qui s'intéressent à l'estimation des paramètres ou Pap et Szabo [45] qui testent si les paramètres changent au cours du temps. On peut également citer les travaux de Freeland et McCabe [25], Neal et Subba Rao [42], Drost et al. [22].

Citons enfin Kachour et Yao [38] et Kachour et Truquet [37]. Kachour et Yao [38] étudient un processus autorégressif à valeurs entières en utilisant une autre approche, où le processus n'est plus construit en utilisant des variables aléatoires à valeurs entières mais en prenant la partie entière de la partie autorégressive d'un processus AR(p) auquel on ajoute un bruit à valeurs entières. Dans Kachour et Truquet [37], le processus à valeurs entières est défini de la même façon qu'un processus INAR(p), à cela près que l'opérateur $\circ$ introduit un signe, entraînant la possibilité d'avoir des valeurs négatives pour le processus.

1.1.3 Les processus autorégressifs à bifurcation

La dernière variation des processus autorégressifs à bifurcation que nous allons présenter, et que nous retrouverons dans toute la suite, est celle qui consiste à transposer l'évolution autorégressive sur un arbre binaire. Ces processus sont appelés processus autorégressifs à bifurcation d'ordre p, notés BAR(p) et sont définis, pour tout $n \geq 2^{p-1}$, par

$$\begin{cases} X_{2n} = a_0 + \sum_{k=1}^{p} a_k X_{\left[\frac{n}{2^{k-1}}\right]} + \varepsilon_{2n} \\ X_{2n+1} = b_0 + \sum_{k=1}^{p} b_k X_{\left[\frac{n}{2^{k-1}}\right]} + \varepsilon_{2n+1}, \end{cases}$$

où $[x]$ désigne la partie entière de x. Les processus BAR(p) sont dits symétriques si $a_i = b_i$ pour tout $i \in \{1, 2, \ldots, p\}$, asymétriques sinon. Comme pour les processus AR(p) et RCAR(p), le caractère linéaire de la dépendance permet de transformer le cas $p \neq 1$ afin de le rapprocher du cas $p = 1$ et ainsi permettre d'unifier les raisonnements dans les deux cas.

Les processus BAR(p) font l'objet de nombreux travaux récents. Le premier d'entre eux fut réalisé par Cowan et Staudte [16] et concernait l'étude des durées de vie de cellules. En effet, la structure d'arbre binaire se prête très bien à la modélisation de la division cellulaire, où chaque état n désigne une cellule qui en se divisant donne deux nouvelles cellules $2n$ et $2n + 1$. Cette vision cellulaire des processus BAR(p) est visible au sein des divers travaux sur le sujet comme on peut le voir avec Huggins et Basawa [31, 32], Huggins et Staudte [33] ou encore Guyon [27]. Parmi les autres travaux, on peut mentionner les résultats de Bercu et al. [11] qui ont traité le problème de l'estimation des paramètres d'un processus BAR par une approche martingale, méthode qui sera celle utilisée dans cette thèse. Cette méthode a également été ensuite reprise par de Saporta et al. [17] pour l'estimation de paramètres d'un processus BAR dont certaines données sont manquantes.

Plusieurs processus proches des processus BAR(p) existent, avec en ligne de mire les processus de division cellulaire, comme les chaînes des Markov à bifurcation introduites par Guyon [26] et utilisées ensuite par Delmas et Marsalle [20] ou le modèle de Kimmel que l'on retrouve dans Bansaye [4]. Ces processus ont également une forme proche des processus de branchement du type Galton-Watson que l'on retrouve par exemple dans Delmas et Marsalle [20] et Heyde et Seneta [29] ou des modèles à fragmentation que l'on retrouve dans Hoffmann et Krell [30] et Doumic et al. [21].

1.2 Les principaux résultats de cette thèse

L'objectif de cette thèse est de généraliser les processus BAR de la même façon que cela a été le cas pour les processus AR et d'établir des résultats asymptotiques sur des estimateurs des paramètres de ces processus. Dans un premier temps nous avons étudié les processus autorégressifs à bifurcation et à valeurs entières, notés BINAR. Dans un second temps, nous nous sommes intéressés aux processus autorégressifs à bifurcation et à coefficients aléatoires, notés RCBAR, car les biologistes nous ont assuré que les processus BAR à coefficients aléatoires étaient plus appropriés pour modéliser la division cellulaire. Nos résultats seront basés sur une approche martingales et non basés sur des chaînes de Markov comme dans Guyon [26].

1.2.1 Les processus autorégressifs à bifurcation et à valeurs entières

Le Chapitre 2 porte sur l'estimation des paramètres d'un processus BINAR asymétrique d'ordre 1, c'est-à-dire un processus définie, pour tout $n \geq 1$, par la relation de récurrence

$$\begin{cases} X_{2n} &= a \circ X_n + \varepsilon_{2n}, \\ X_{2n+1} &= b \circ X_n + \varepsilon_{2n+1}. \end{cases}$$

Un tel processus pourrait, par exemple, modéliser le nombre de parasite d'une cellule comme cela fut traité par Bansaye [4, 5]. On peut alors voir deux parties dans ces relations de récurrence, le premier terme modélisant les parasites hérités de la cellule mère et le second terme rendant compte d'éventuels parasites venant de l'environnement de la cellule, ce facteur environnemental n'étant pas présent dans le premier modèle de Kimmel étudié par Bansaye [4]. Ce processus se rapproche du modèle de la deuxième étude de Bansaye [5] où l'on considère à la fois une partie héritée et une partie venant de l'environnement. Cependant cette étude se base sur une modélisation différente et s'intéresse au comportement asymptotique du nombre de parasite plutôt qu'à l'estimation de paramètres. Comme annoncé à la section 1.1.2, nous ne considérerons pas de loi binomiale dans la définition de l'opérateur $\circ$ mais plus généralement des suites de variables aléatoires indépendantes et identiquement distribuées de moyennes a et b pour les opérations $a \circ$ et $b \circ$, respectivement. Nous ne ferons pas non plus d'hypothèse sur la distribution des bruits, nous considérerons une suite $(\varepsilon_{2n}, \varepsilon_{2n+1})$ indépendante et identiquement distribuée.

Les paramètres d'intérêt seront les espérances a et b et les variances σ_a^2 et σ_b^2 des variables aléatoires intervenant dans les opérations $a \circ$ et $b \circ$, ainsi que les espérances c et d et variances σ_c^2 et σ_d^2 des deux bruits et la covariance ρ entre ces deux bruits. Pour ce faire, nous utiliserons, après avoir réécrit notre système sous une forme BAR, un estimateur des moindres carrés pondérés, à l'exception de l'estimateur de la covariance des deux bruits qui sera traité grâce à un estimateur de Monte-Carlo. Le choix de pondérer nos estimateurs s'inspire des travaux antérieurs de Wei et Winnicki [58], Winnicki [59] sur les processus de branchement avec immigration. L'objectif est ici de réduire les hypothèses de moments requises avec seulement des hypothèses de moments d'ordre 8. Le point clé de cette étude est le fait de réussir à faire apparaître une martingale vectorielle dans l'expression de nos estimateurs et d'adapter les raisonnements de martingales classiques à notre cas où les tailles des vecteurs doublent à chaque pas en avant dans l'arbre binaire. Ce travail d'adaptation a été précédemment effectué pour les processus BAR par Bercu et al. [11].

Les résultats établis dans ce chapitre sont les suivants. Tout d'abord nous établissons deux lemmes clés de l'étude nous donnant des informations sur le comportement asymptotique de notre processus BINAR. Le premier montre la convergence en loi de la suite de valeurs prises sur une branche choisie uniformément sur l'arbre binaire vers une certaine variable aléatoire T. Cette variable aléatoire est connue sous la forme d'une somme infinie à partir de laquelle nous sommes en mesure de calculer ses différents moments et ainsi de s'assurer qu'elle est non dégénérée et qu'elle admet un moment d'ordre trois fini. Le deuxième lemme clé, qui adapte un raisonnement de Guyon [26], nous assure la convergence de la moyenne de la fonction des différents état de notre processus vers l'espérance de la fonction prise en T, pour une fonction prise dans une certaine classe. Plus précisément on obtient le lemme suivant

Lemme 1.2.1. *Soit*

$$\mathcal{C}_3^1(\mathbb{R}_+) = \Big\{ f \in \mathcal{C}^1(\mathbb{R}_+, \mathbb{R}) \big| \exists \gamma > 0, \forall x \geq 0, (|f'(x)| + |f(x)|) \leq \gamma(1 + x^3) \Big\}.$$

Alors, sous de bonnes hypothèses, on a pour toute fonction $f \in \mathcal{C}_3^1(\mathbb{R}_+)$,

$$\lim_{n \to \infty} \frac{1}{|\mathbb{T}_n|} \sum_{k \in \mathbb{T}_n} f(X_k) = \mathbb{E}[f(T)] \qquad p.s.$$

avec

$$\mathbb{T}_n = \{1, 2, 3, \dots, 2^{n+1} - 1\} \qquad et \qquad |\mathbb{T}_n| = 2^{n+1} - 1.$$

Ce lemme va nous permettre d'établir les convergences requises pour nos différents théorèmes de martingales, à commencer par la convergence qui est centrale pour ces théorèmes c'est-à-dire la convergence du crochet convenablement normalisé de la martingale d'intérêt de notre processus BINAR. Ce résultat de convergence nous permet de nous rendre compte de la contrepartie de la pondération. En effet, la matrice limite du crochet n'est pas connue explicitement mais uniquement au travers des espérances de fonctions de la variable aléatoire T.

Les résultats suivants concernent à proprement parler nos estimateurs. Ainsi, nous établissons la convergence presque sûre de nos estimateurs vers les valeurs à estimer, comme par exemple a, b, c et d, tout en déterminant une vitesse de convergence. Nous établissons également une loi forte quadratique. Nous terminons cette étude par des résultats de normalité asymptotique de nos estimateurs, avec des matrices de covariances qui s'expriment grâce à des espérances de fonctions de la variable aléatoire T. L'avantage d'avoir traité différemment l'estimation de la covariance ρ nous permet, sans trop peser sur les hypothèses de moments, d'obtenir une normalité asymptotique dont la variance limite est explicite.

1.2.2 Les processus autorégressifs à bifurcation et à coefficients aléatoires

Le Chapitre 3 s'intéressera à l'estimation des paramètres d'un processus RCBAR asymétrique d'ordre 1. Un tel processus est l'adaptation des processus RCAR à la structure d'un arbre binaire. Le processus RCBAR d'ordre 1 est défini, pour tout $n \geq 1$, par la relation de récurrence

$$\begin{cases} X_{2n} &= a_n X_n + \varepsilon_{2n} \\ X_{2n+1} &= b_n X_n + \varepsilon_{2n+1}, \end{cases}$$

où l'on retrouve nos deux parties dans les formules de récurrence : l'effet héréditaire et l'effet environnemental via les suites de variables aléatoires indépendantes et identiquement distribuées (a_n, b_n) et $(\varepsilon_{2n}, \varepsilon_{2n+1})$ respectivement, ces deux suites étant mutuellement indépendantes. Ce travail permet de généraliser le processus BAR étudié par Bercu et al. [11]. En effet, on autorise un aléa sur l'effet héréditaire qui n'était pas autorisé dans le processus BAR. En ayant toujours en tête les possibles applications en biologie, on se rend compte que le fait de ne plus considérer un héritage déterministe prend tout son sens, la nature ayant en général une affection pour l'aléatoire. De plus, il s'agit réellement d'une généralisation puisque l'on autorise dans cette étude la suite (a_n, b_n) a être dégénérée en (a, b), retrouvant ainsi un processus BAR. Un tel processus pourra modéliser divers caractères continus comme la taille ou la durée de vie des cellules, comme cela a été réalisé par Guyon [26] et Guyon et al. [27] en utilisant des processus BAR pour étudier la durée de vie des cellules au cours de leur division.

Les paramètres estimés seront les espérances a, b, c et d, des quatre variables aléatoires a_n, b_n, ε_{2n} et ε_{2n+1} ainsi que leurs variances, et également les covariances ρ_{ab} entre a_n et b_n et ρ_{cd} entre ε_{2n} et ε_{2n+1}. L'approche sera similaire à l'estimation des paramètres d'un processus BINAR. En effet, une réécriture du système de récurrence nous permettra de le mettre sous une forme BAR à partir de laquelle nous déduirons notre estimateur des moindres carrés pondérés du vecteur des quatre moyennes. Un travail similaire à celui du Chapitre 3 nous donnera les estimateurs des moindres carrés du vecteur des variances de l'effet héréditaire et du vecteur des variances de l'effet environnemental, ainsi que notre dernier estimateur auquel l'on s'intéressera, l'estimateur des moindres carrés pondérés du vecteur des deux covariances ρ_{ab} et ρ_{cd}. La pondération des moindres carrés a une fois de plus pour but de minimiser les hypothèses sur les moments, on se limitera ici à une hypothèse sur les moments conditionnels, puisque l'on utilisera des martingales, d'ordre α pour un certain $\alpha > 4$. A noter que l'on se placera sous les hypothèses $\mathbb{E}[a_n^2] < 1$ et $\mathbb{E}[b_n^2] < 1$ pour nous assurer la stationnarité de notre processus. Ces hypothèses sont les équivalents dans le cas RCBAR des hypothèses $|a_1| < 1$ et $|b_1| < 1$ utilisées pour les processus BAR(1).

La démarche dans nos résultats est analogue à celle mise en œuvre pour les processus BINAR, bien que les adaptations soient nombreuses. Nous montrerons ainsi la convergence en loi sur une branche de l'arbre prise uniformément vers une variable aléatoire qui peut s'exprimer sous la forme d'une somme infinie de variables aléatoires. Cette somme nous permettra d'établir la non dégénérescence de cette variable aléatoire et de montrer qu'elle admet un moment d'ordre deux fini. Ensuite, nous serons en mesure d'établir un lemme analogue au Lemme 1.2.1 pour la classe de fonction

$$\mathcal{C}_b^1(\mathbb{R}_+) = \Big\{ f \in \mathcal{C}^1(\mathbb{R}, \mathbb{R}) \big| \exists \gamma > 0, \forall x \geq 0, (|f'(x)| + |f(x)|) \leq \gamma \Big\}.$$

On pourra ainsi établir la convergence presque sûre du crochet, convenablement normalisé, de la martingale d'intérêt de notre processus RCBAR. Il en découlera la convergence presque sûre de nos différents estimateurs vers les valeurs à estimer, comme par exemple a, b, c et d. Ces résultats seront accompagnés d'une vitesse de convergence pour ces estimateurs et d'une loi forte quadratique. Enfin, nous établirons la normalité asymptotique de nos estimateurs. Le fait d'avoir pondéré nos estimateurs nous permet une fois de plus de considérablement diminuer les hypothèses de moments, mais la contrepartie est la même que pour les processus BINAR : nous ne connaissons les matrices de covariance des normalités asymptotiques qu'au travers d'espérances de fonctions de la variable aléatoire T. Ce chapitre se terminera par une illustration numérique de la normalité asymptotique de nos estimateurs des espérances.

1.2.3 L'approche Rademacher-Menchov

La principale motivation du Chapitre 4 est l'obtention de variances asymptotiques explicites pour les estimateurs des paramètres inconnus des processus RCBAR. Cet objectif nous a contraint à ne plus pondérer nos estimateurs des moindres carrés ce qui a pour conséquence de faire augmenter significativement les hypothèses de moments. Afin de réduire sensiblement ces hypothèses, nous avons alors abandonné la méthode précédente reposant sur les chaînes de Markov à bifurcation de Guyon [26] pour adopter une démarche basée sur le théorème de Rademacher-Menchov.

Théorème 1.2.2. *Soit* (X_n) *une suite orthonormale de variables aléatoires de carrés intégrables, c'est-à-dire vérifiant pour tout* $n \neq k$, $\mathbb{E}[X_n X_k] = 0$ *et* $\mathbb{E}[X_n^2] = 1$. *Soit* (a_n) *une suite réelle satisfaisant*

$$\sum_{n=1}^{\infty} a_n^2 (\log n)^2 < \infty. \tag{1.2.1}$$

Alors, la série suivante converge presque sûrement

$$\sum_{n=1}^{\infty} a_n X_n. \tag{1.2.2}$$

Remarque 1.2.3. *Plus précisément, nous nous reposerons sur la version orthogonale de ce théorème qui nous donne la même conclusion* (1.2.2) *pour peu que l'on remplace la condition* (1.2.1) *par*

$$\sum_{n=1}^{\infty} a_n^2 \mathbb{E}[X_n^2] (\log n)^2 < \infty.$$

Les paramètres à estimer dans cette étude seront les dix mêmes valeurs que celle rencontrées au Chapitre 3. Nous nous placerons pour cela sous des hypothèses plus restrictives de moment d'ordre 16, notamment $\mathbb{E}[a_n^{16}] < 1$ et $\mathbb{E}[b_n^{16}] < 1$. Cette augmentation de l'ordre des moments est entièrement due à l'abandon de la pondération de nos estimateurs. Le point crucial de cette étude est le remplacement des deux lemmes concernant la loi limite des Chapitres 2 et 3 par un seul lemme qui est le suivant.

Lemme 1.2.4. *Sous de bonnes hypothèses, on a pour tout* $p \in \{1, 2, \ldots, 8\}$,

$$\lim_{n \to \infty} \frac{1}{|\mathbb{T}_n|} \sum_{k \in \mathbb{T}_n} X_k^p = s_p \qquad p.s.$$

où s_p *est une constante ne dépendant que des moments d'ordre au plus* p *de* a_1, b_1, ε_2 *and* ε_3.

La grande différence par rapport aux Chapitres 2 et 3 est donc que cette fois-ci, nous avons une convergence où l'on connaît explicitement les limites en fonction des paramètres du processus. Il s'avère par ailleurs que les constantes s_p sont les moments d'ordre p de la loi limite T du Chapitre 3, sous réserve d'adapter les hypothèses dans ce chapitre pour que T admette un moment d'ordre 8.. Comme on peut le voir, ce nouveau lemme ne permet d'obtenir que les convergences des sommes des

8 premières puissances des X_k, contrairement aux chapitres précédents où l'on avait accès à des classes de fonctions plus larges. Cependant, le fait de ne pas pondérer nos estimateurs nous permet d'obtenir des variances asymptotiques explicites de nos estimateurs.

Une fois ce lemme fondamental pour notre étude établi, nous adapterons les résultats obtenus pour les estimateurs pondérés des RCBAR étudiés dans le Chapitre 3, ainsi que les preuves de ces résultats, à nos nouveaux estimateurs. Nous serons ainsi en mesure d'établir la convergence presque sûre du crochet, convenablement normalisé, de la martingale d'intérêt de notre processus RCBAR, et le Lemme 1.2.4 nous permettra ici d'avoir une expression explicite de la matrice limite. Nous établirons ensuite la convergence presque sûre de nos estimateurs en en donnant une vitesse de convergence qui est la même que celle obtenue pour les estimateurs pondérés. Ces résultats seront accompagnés d'une loi forte quadratique pour nos estimateurs, en ayant cette fois-ci explicitement les matrices qui y interviennent. Enfin, nous seront en mesure de montrer les normalités asymptotiques de nos estimateurs avec, grâce une fois de plus à l'approche Rademacher-Menchov et au Lemme 1.2.4 qui en découle, les matrices de covariances limites qui sont connues explicitement. Cette étude s'achèvera par une illustration de la normalité asymptotique de nos estimateurs des espérances.

1.3 Quelques éléments de martingales vectorielles

Comme énoncé précédemment, l'optique de cette thèse est d'établir des résultats sur des estimateurs grâce à des arguments de martingales vectorielles à temps discret. Nous donnons ici les principaux résultats concernant les martingales vectorielles qui seront utilisés dans la suite. Pour davantage de résultats à ce sujet, on pourra se référer au livre de Duflo [23] dont sont extraits les résultats ci-dessous.

Nous allons ici énoncer les deux principaux théorèmes concernant la convergence des martingales vectorielles, c'est-à-dire la loi forte des grands nombres et le théorème central limite. La valeur limite qui va conditionner ces deux théorèmes est le crochet de la martingale qui est défini ainsi.

Définition 1.3.1. *Soit (M_n) une martingale vectorielle de carré intégrable adaptée à la filtration $(\mathcal{F}_n)$. Son crochet est donné par $\langle M\rangle_0 = 0$ et, pour tout $n \geq 1$,*

$$\langle M\rangle_n = \sum_{k=1}^{n} \mathbb{E}\left[\Delta M_k \Delta M_k^t | \mathcal{F}_{k-1}\right].$$

On suppose que M_n s'écrit sous la forme

$$M_n = \sum_{k=1}^{n} \Phi_{k-1}\varepsilon_k$$

où (ε_k) est une suite de variables aléatoires réelles adaptée à $(\mathcal{F}_n)$ avec, pour tout $n \geq 1$,

$$\mathbb{E}[\varepsilon_n|\mathcal{F}_{n-1}] = 0 \qquad \text{et} \qquad \mathbb{E}[\varepsilon_n^2|\mathcal{F}_{n-1}] = \sigma^2 \qquad \text{p.s.}$$

avec $\sigma^2 > 0$. On suppose également que (Φ_n) est une suite de vecteurs aléatoires de $\mathbb{R}^d$, adaptée à $(\mathcal{F}_n)$, et on pose

$$S_n = \sum_{k=0}^{n} \Phi_k \Phi_k^t.$$

Il est clair que (M_n) est une martingale vectorielle dont le crochet est

$$\langle M \rangle_n = \sigma^2 S_{n-1}.$$

On note $\lambda_{max}(S_n)$ et $\lambda_{min}(S_n)$ la plus grande et la plus petite valeur propre de la matrice S_n et on suppose, sans perte de généralité (si ce n'est pas le cas il suffit de rajouter la matrice identité à S_n), que S_n est inversible.

Théorème 1.3.2 (Loi forte des grands nombres)**.**

1. *Sur* $\{\lim_{n\to\infty} \lambda_{max}(S_n) < \infty\}$, *la martingale* (M_n) *converge presque sûrement.*
2. *Sur* $\{\lim_{n\to\infty} \lambda_{max}(S_n) = \infty\}$, *on a pour tout* $\gamma > 0$

$$M_n^t S_{n-1}^{-1} M_n = o\left((\log \lambda_{max}(S_n))^{1+\gamma}\right) \qquad p.s.$$

3. *On suppose qu'il existe* $a > 2$ *tel que*

$$\sup_{n\geq 1} \mathbb{E}\left[|\varepsilon_n|^a | \mathcal{F}_{n-1}\right] < \infty \qquad p.s.$$

Alors, sur $\{\lim_{n\to\infty} \lambda_{max}(S_n) = \infty\}$, *on a*

$$M_n^t S_{n-1}^{-1} M_n = O(\log \lambda_{max}(S_n)) \qquad p.s.$$

Il est important de noter que l'on ne peut pas appliquer directement la loi forte des grands nombres pour les martingales vectorielles à notre cadre à cause de la structure d'arbres binaires des processus BINAR et RCBAR.

Théorème 1.3.3 (Théorème central limite)**.**
Soit $(M_k^{(n)})_{k\geq 0}$ *une martingale vectorielle de carré intégrable adaptée à* $(\mathcal{F}_k^{(n)})$ *et soit* t_n *un temps d'arrêt adapté à cette filtration. Si*

1. *Il existe une matrice symétrique semi-définie positive* L *telle que*

$$\langle M^{(n)} \rangle_{t_n} \xrightarrow{\mathbb{P}} L.$$

2. *La condition de Lyapounov est vérifiée, c'est-à-dire qu'il existe* $a > 2$ *tel que*

$$\sum_{k=1}^{t_n} \mathbb{E}\left[\|M_k^{(n)} - M_{k-1}^{(n)}\|^a | \mathcal{F}_{n-1}^{(n)}\right] \xrightarrow{\mathbb{P}} 0,$$

alors, on a

$$M_{t_n}^{(n)} \xrightarrow{\mathcal{L}} \mathcal{N}(0, L).$$

Le théorème central limite est énoncé avec la condition de Lyapounov car c'est celle dont on se servira par la suite. On peut la remplacer par une condition plus faible, la condition de Lindeberg : pour tout $\varepsilon > 0$,

$$\sum_{k=1}^{t_n} \mathbb{E}\left[\|M_k^{(n)} - M_{k-1}^{(n)}\|^2 \mathbf{1}_{(\|M_k^{(n)} - M_{k-1}^{(n)}\| > \varepsilon)} | \mathcal{F}_{n-1}^{(n)}\right] \xrightarrow{\mathbb{P}} 0.$$

Chapitre 2

Processus BAR à valeurs entières

Résumé. Nous étudions le comportement asymptotique des estimateurs des moindres carrés pondérés des paramètres inconnus des processus autorégressifs à bifurcation et à valeurs entières. Sous des hypothèses appropriées sur l'immigration, nous établissons la convergence presque sure de nos estimateurs ainsi qu'une loi forte quadratique et des théorèmes centraux limites. Notre étude repose essentiellement sur des résultats asymptotiques pour les martingales vectorielles.

Abstract. We study the asymptotic behavior of the weighted least squares estimators of the unknown parameters of bifurcating integer-valued autoregressive processes. Under suitable assumptions on the immigration, we establish the almost sure convergence of our estimators, together with a quadratic strong law and central limit theorems. Our investigation deeply relies on asymptotic results for vector-valued martingales.

2.1 Introduction

Bifurcating integer-valued autoregressive (BINAR) processes are an adaptation of integer-valued autoregressive (INAR) processes to binary tree structured data. It can also be seen as the combination of INAR processes and bifurcating autoregressive (BAR) processes. BAR processes have been first introduced by Cowan and Staudte [16] while INAR processes have been first investigated by Al-Osh and Alzaid [1, 2] and McKenzie [40]. BINAR processes take into account both inherited and environmental effects to explain the evolution of the integer-valued characteristic under study.

We can easily see cell division as an example of binary tree structure, the integer-valued characteristic could then be, as an example, the number of parasites in a cell. Keeping this example in mind, we consider that each time a cell is dividing, the two sister cells inherits both some parasites depending on the number of parasites of the mother, and some parasites from the environment. Bansaye [4] used a Kimmel branching process to model this division process. This Kimmel process can be seen as the inheritance part of our BINAR process, where the parasites in the mother cell divide and then the offspring are distributed among the two sister cell. However, this model does not allow any environmental effect.

The first-order BINAR process is defined as follows. The initial cell is labelled 1 and the offspring of the cell labelled n are labelled $2n$ and $2n+1$. Denote by X_n the integer-valued characteristic of individual n. Then, the first-order BINAR process is given, for all $n \geq 1$, by

$$\begin{cases} X_{2n} &= a \circ X_n + \varepsilon_{2n} \\ X_{2n+1} &= b \circ X_n + \varepsilon_{2n+1} \end{cases}$$

where the thinning operator $\circ$ will be defined in (2.2.2). The immigration sequence $(\varepsilon_{2n}, \varepsilon_{2n+1})_{n\geq 1}$ represents the environmental effect, while the thinning operator represents the inherited effect. The example of the cell division incites us to suppose that ε_{2n} ans ε_{2n+1} are correlated since the environmental effect on two sister cells can reasonably be seen as correlated.

The purpose of this paper is to study the asymptotic behavior of the weighted least squares (WLS) estimators of first-order BINAR process via a martingale approach. The martingale approach has been first proposed by Bercu et al. [11] and de Saporta et al. [17] for BAR processes. We also refer to Wei and Winnicki [58] and Winnicki [59] for the WLS estimation of parameters associated to branching processes. We shall make use of the strong law of large numbers [23] as well as the central limit theorem [23, 28] for martingales, in order to investigate the asymptotic behavior of the WLS estimators, as previously done by Basawa and Zhou [9, 60, 61]. In contrast with Bercu et al. [11], we investigate the asymptotic behavior of a WLS estimator instead of a least squares one. It enables us to reduce the moment assumption on the immigration sequence. The fact that we consider an integer-valued process also forced us to adapt the proofs because of the thinning operator which needs to be manipulated more carefully than the classical product.

Several points of view appeared for both BAR and INAR processes and we tried to make a link between those approaches. On the one hand, for the BAR side of the BINAR process, we had a look to classical BAR studies as done by Huggins and Basawa [31, 32] and Huggins ans Staudte [33] who studied the evolution of cell diameters and lifetimes, but also to bifurcating Markov chains models introduced by Guyon [26] and used in Delmas and Marsalle [20]. However, we did not put aside the analogy with the Galton-Watson processes as studied in Delmas and Marsalle [20] and Heyde and Seneta [29]. On the other hand, concerning the INAR side of the BINAR process, we used the classical INAR definition but also had a look to Bansaye [4, 6, 5] who studied an integer-valued process on a binary tree without using an INAR model, and also Kachour and Yao [38] who decided to study an integer-valued autoregressive process by a rounding approach instead of the classical INAR one. The approach of this paper has also been used for the study of random coefficient bifurcating autoregressive (RCBAR) process as in Blandin [13] and Bercu and Blandin [10]. RCBAR processes is the combination of BAR processes and random coefficient autoregressive processes. They have been previously investigated by Nicholls and Quinn [43, 44, 46].

The paper is organised as follows. Section 2 is devoted to the presentation of the first-order BINAR process while Section 3 deals with the WLS estimators of the unknown parameters. Section 4 allows us to detail our approach based on martingales. Section 5 gathers the main results about the asymptotic properties of the WLS estimators. More precisely, we will propose the almost sure convergence, the quadratic strong law and the central limit theorem for our estimates. The rest of the paper is devoted to the proofs of our main results.

2.2 Bifurcating integer-valued autoregressive processes

Consider the first-order BINAR process given, for all $n \geq 1$, by

$$\begin{cases} X_{2n} & = a \circ X_n + \varepsilon_{2n} \\ X_{2n+1} & = b \circ X_n + \varepsilon_{2n+1} \end{cases} \tag{2.2.1}$$

where the initial integer-valued state X_1 is the ancestor of the process and $(\varepsilon_{2n}, \varepsilon_{2n+1})$ represents the immigration which takes nonnegative integer values. In all the sequel, we shall assume that $\mathbb{E}[X_1^8] < \infty$. Moreover,

$$a \circ X_n = \sum_{i=1}^{X_n} Y_{n,i} \qquad \text{and} \qquad b \circ X_n = \sum_{i=1}^{X_n} Z_{n,i} \tag{2.2.2}$$

where $(Y_{n,i})_{n,i\geq 1}$ and $(Z_{n,i})_{n,i\geq 1}$ are two independent sequences of i.i.d., nonnegative integer-valued random variables with means a and b and positive variances σ_a^2 and σ_b^2 respectively. Moreover, μ_a^4, μ_b^4 and τ_a^6, τ_b^6 are the fourth-order and the sixth-order centered moments of $(Y_{n,i})$ and $(Z_{n,i})$, respectively, and $(Y_{n,i})$ and $(Z_{n,i})$ admit eighth-order moments. We also assume that the two offspring sequences $(Y_{n,i})$ and $(Z_{n,i})$ are independent of the immigration $(\varepsilon_{2n}, \varepsilon_{2n+1})$. Besides, we will assume that the distributions of ε_{2n} and ε_{2n+1} do not depend on n, while allowing the one of $(\varepsilon_{2n}, \varepsilon_{2n+1})$ to depend on n. In addition, as in the literature concerning BAR processes, we shall assume that

$$0 < \max(a, b) < 1.$$

One can see this BINAR process as a first-order integer-valued autoregressive process on a binary tree, where each node represents an individual, node 1 being the original ancestor. For all $n \geq 1$, denote the n-th generation by

$$\mathbb{G}_n = \{2^n, 2^n + 1, \ldots, 2^{n+1} - 1\}.$$

In particular, $\mathbb{G}_0 = \{1\}$ is the initial generation and $\mathbb{G}_1 = \{2, 3\}$ is the first generation of offspring from the first ancestor. Let $\mathbb{G}_{r_n}$ be the generation of individual n, which means that $r_n = [\log_2(n)]$. Recall that the two offspring of individual n are labelled $2n$ and $2n + 1$, or conversely, the mother of individual n is $[n/2]$ where $[x]$ stands for the largest integer less than or equal to x. Finally denote by

$$\mathbb{T}_n = \bigcup_{k=0}^{n} \mathbb{G}_n$$

the sub-tree of all individuals from the original individual up to the n-th generation. On can observe that the cardinality $|\mathbb{G}_n|$ of $\mathbb{G}_n$ is 2^n while that of $\mathbb{T}_n$ is $|\mathbb{T}_n| = 2^{n+1} - 1$.

2.3 Weighted least-squares estimation

Denote by $\mathbb{F} = (\mathcal{F}_n)_{n\geq 0}$ the natural filtration associated with the first-order BINAR process, which means that $\mathcal{F}_n$ is the σ-algebra generated by all individuals up to the n-th generation, in other words $\mathcal{F}_n = \sigma\{X_k, k \in \mathbb{T}_n\}$. We will assume in all the sequel that, for all $n \geq 0$ and for all $k \in \mathbb{G}_n$,

$$\begin{cases} \mathbb{E}[\varepsilon_{2k}|\mathcal{F}_n] = c & \text{a.s.} \\ \mathbb{E}[\varepsilon_{2k+1}|\mathcal{F}_n] = d & \text{a.s.} \end{cases}$$

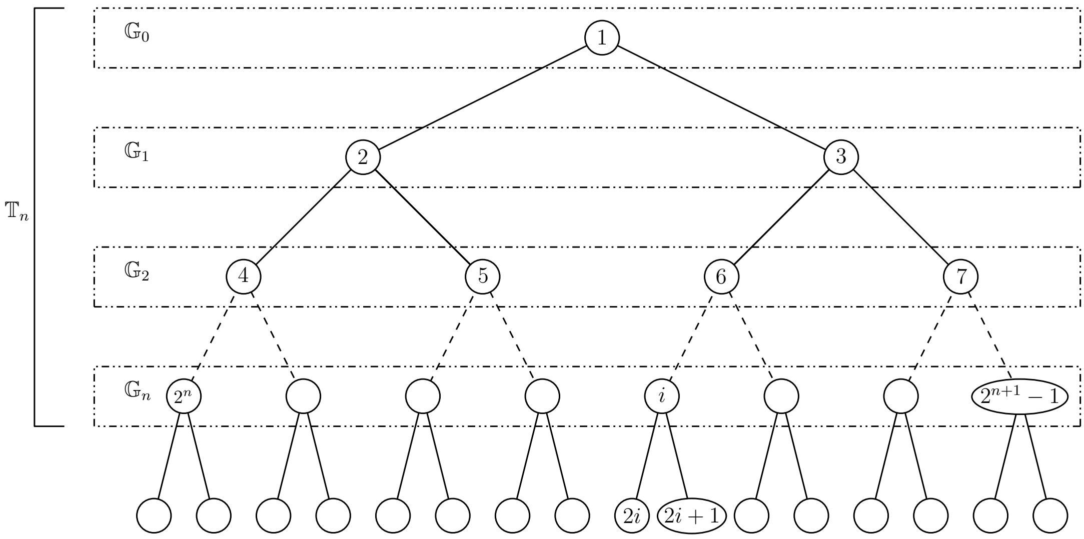

FIGURE 2.1 – The tree associated with the BINAR

Consequently, we deduce from (2.2.1) that, for all $n \geq 0$ and for all $k \in \mathbb{G}_n$,

$$\begin{cases} X_{2k} & = aX_k + c + V_{2k}, \\ X_{2k+1} & = bX_k + d + V_{2k+1}, \end{cases} \tag{2.3.1}$$

where $V_{2k} = X_{2k} - \mathbb{E}[X_{2k}|\mathcal{F}_n]$ and $V_{2k+1} = X_{2k+1} - \mathbb{E}[X_{2k+1}|\mathcal{F}_n]$. Therefore, the two relations given by (2.3.1) can be rewritten in the matrix form

$$\chi_n = \theta^t \Phi_n + W_n \tag{2.3.2}$$

where

$$\chi_n = \begin{pmatrix} X_{2n} \\ X_{2n+1} \end{pmatrix}, \qquad \Phi_n = \begin{pmatrix} X_n \\ 1 \end{pmatrix}, \qquad W_n = \begin{pmatrix} V_{2n} \\ V_{2n+1} \end{pmatrix},$$

and the matrix parameter

$$\theta = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Our goal is to estimate θ from the observation of all individuals up to $\mathbb{T}_n$. We propose to make use of the WLS estimator $\widehat{\theta}_n$ of θ which minimizes

$$\Delta_n(\theta) = \frac{1}{2} \sum_{k \in \mathbb{T}_{n-1}} \frac{1}{c_k} \|\chi_k - \theta^t \Phi_k\|^2$$

where the choice of the weighting sequence $(c_n)_{n \geq 1}$ is crucial. We shall choose $c_n = 1 + X_n$ and we will go back to this suitable choice in Section 2.4. Consequently, we obviously have for all $n \geq 1$

$$\widehat{\theta}_n = S_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \frac{1}{c_k} \Phi_k \chi_k^t \tag{2.3.3}$$

where

$$S_n = \sum_{k\in\mathbb{T}_n} \frac{1}{c_k}\Phi_k\Phi_k^t.$$

In order to avoid useless invertibility assumption, we shall assume, without loss of generality, that for all $n \geq 0$, S_n is invertible. Otherwise, we only have to add the identity matrix of order 2, I_2 to S_n. Since, in a certain way, S_n goes to infinty , it will not change our results. In all what follows, we shall make a slight abuse of notation by identifying θ as well as $\widehat{\theta}_n$ to

$$\mathrm{vec}(\theta) = \begin{pmatrix} a \\ c \\ b \\ d \end{pmatrix} \qquad \text{and} \qquad \mathrm{vec}(\widehat{\theta}_n) = \begin{pmatrix} \widehat{a}_n \\ \widehat{c}_n \\ \widehat{b}_n \\ \widehat{d}_n \end{pmatrix}.$$

Therefore, we deduce from (2.3.3) that

$$\begin{aligned} \widehat{\theta}_n &= \Sigma_{n-1}^{-1} \sum_{k\in\mathbb{T}_n-1} \frac{1}{c_k}\mathrm{vec}(\Phi_k\chi_k^t), \\ &= \Sigma_{n-1}^{-1} \sum_{k\in\mathbb{T}_n-1} \frac{1}{c_k}\begin{pmatrix} X_kX_{2k} \\ X_{2k} \\ X_kX_{2k+1} \\ X_{2k+1} \end{pmatrix} \end{aligned}$$

where $\Sigma_n = I_2 \otimes S_n$ and $\otimes$ stands for the standard Kronecker product. Consequently, (2.3.2) yields to

$$\begin{aligned} \widehat{\theta}_n - \theta &= \Sigma_{n-1}^{-1} \sum_{k\in\mathbb{T}_{n-1}} \frac{1}{c_k}\mathrm{vec}(\Phi_k W_k^t), \\ &= \Sigma_{n-1}^{-1} \sum_{k\in\mathbb{T}_{n-1}} \frac{1}{c_k}\begin{pmatrix} X_kV_{2k} \\ V_{2k} \\ X_kV_{2k+1} \\ V_{2k+1} \end{pmatrix}. \end{aligned} \tag{2.3.4}$$

In all the sequel, we shall make use of the following moment hypotheses.

(H.1) For all $n \geq 0$ and for all $k \in \mathbb{G}_n$

$$\mathbb{E}[\varepsilon_{2k}|\mathcal{F}_n] = c \qquad \text{and} \qquad \mathbb{E}[\varepsilon_{2k+1}|\mathcal{F}_n] = d \qquad \text{a.s.}$$

(H.2) For all $n \geq 0$ and for all $k \in \mathbb{G}_n$

$$\mathrm{Var}[\varepsilon_{2k}|\mathcal{F}_n] = \sigma_c^2 > 0 \quad \text{and} \quad \mathrm{Var}[\varepsilon_{2k+1}|\mathcal{F}_n] = \sigma_d^2 > 0 \qquad \text{a.s.}$$

(H.3) For all $n \geq 0$ and for all $k, l \in \mathbb{G}_{n+1}$, if $[k/2] \neq [l/2]$, ε_k and ε_l are conditionally independent given $\mathcal{F}_n$, while otherwise it exists $\rho^2 < \sigma_c^2 \sigma_d^2$ such that, for all $k \in \mathbb{G}_n$

$$\mathbb{E}[(\varepsilon_{2k} - c)(\varepsilon_{2k+1} - d)|\mathcal{F}_n] = \rho \qquad \text{a.s.}$$

(H.4) One can find $\mu_c^4 > \sigma_c^4$ and $\mu_d^4 > \sigma_d^4$ such that, for all $n \geq 0$ and for all $k \in \mathbb{G}_n$

$$\mathbb{E}\left[(\varepsilon_{2k} - c)^4 |\mathcal{F}_n\right] = \mu_c^4 \qquad \text{and} \qquad \mathbb{E}\left[(\varepsilon_{2k+1} - d)^4 |\mathcal{F}_n\right] = \mu_d^4 \qquad \text{a.s.}$$

In addition, it exists $\nu^4 \leq \mu_c^4 \mu_d^4$ such that, for all $k \in \mathbb{G}_n$

$$\mathbb{E}[(\varepsilon_{2k} - c)^2 (\varepsilon_{2k+1} - d)^2 |\mathcal{F}_n] = \nu^2 \qquad \text{a.s.}$$

(H.5) One can find $\tau_c^6 > 0$ and $\tau_d^6 > 0$ such that

$$\sup_{n\geq 1} \sup_{k\in\mathbb{G}_n} \mathbb{E}[\varepsilon_{2k}^6|\mathcal{F}_n] = \tau_c^6 \qquad \text{and} \qquad \sup_{n\geq 1} \sup_{k\in\mathbb{G}_n} \mathbb{E}[\varepsilon_{2k+1}^6|\mathcal{F}_n] = \tau_d^6 \qquad \text{a.s.}$$

$$\sup_{n\geq 2} \mathbb{E}[\varepsilon_n^8] < \infty$$

It follows from hypothesis **(H.1)** that V_{2n} and V_{2n+1} can be rewritten as

$$V_{2n} = \sum_{i=1}^{X_n} (Y_{n,i} - a) + (\varepsilon_{2n} - c) \quad \text{and} \quad V_{2n+1} = \sum_{i=1}^{X_n} (Z_{n,i} - b) + (\varepsilon_{2n} - d).$$

Hence, under assumption **(H.2)**, we have for all $n \geq 0$ and for all $k \in \mathbb{G}_n$

$$\mathbb{E}[V_{2k}^2|\mathcal{F}_n] = \sigma_a^2 X_k + \sigma_c^2 \quad \text{and} \quad \mathbb{E}[V_{2k+1}^2|\mathcal{F}_n] = \sigma_b^2 X_k + \sigma_d^2 \qquad \text{a.s.} \qquad (2.3.5)$$

Consequently, if we choose $c_n = 1 + X_n$ for all $n \geq 1$, we clearly have for all $k \in \mathbb{G}_n$

$$\mathbb{E}\left[V_{2k}^2 \middle| \mathcal{F}_n\right] \leq \max(\sigma_a^2, \sigma_c^2) c_k \quad \text{and} \quad \mathbb{E}\left[V_{2k+1}^2 \middle| \mathcal{F}_n\right] \leq \max(\sigma_b^2, \sigma_d^2) c_k \quad \text{a.s.}$$

It is exactly the reason why we have chosen this weighting sequence into (2.3.3). Similar WLS estimation approach for branching processes with immigration may be found in [58] and [59]. We can also observe that, for all $k \in \mathbb{G}_n$, under the assumption **(H.3)**

$$\rho = \mathbb{E}[V_{2k} V_{2k+1}|\mathcal{F}_n] \quad \text{a.s.}$$

Hence, we propose to estimate the conditional covariance ρ by

$$\widehat{\rho}_n = \frac{1}{|\mathbb{T}_{n-1}|} \sum_{k\in\mathbb{T}_{n-1}} \widehat{V}_{2k} \widehat{V}_{2k+1} \qquad (2.3.6)$$

where for all $k \in \mathbb{G}_n$,

$$\begin{cases} \widehat{V}_{2k} &= X_{2k} - \widehat{a}_n X_k - \widehat{c}_n, \\ \widehat{V}_{2k+1} &= X_{2k+1} - \widehat{b}_n X_k - \widehat{d}_n. \end{cases}$$

For all $n \geq 0$ and for all $k \in \mathbb{G}_n$, denote $v_{2k} = V_{2k}^2 - \mathbb{E}[V_{2k}^2 | \mathcal{F}_n]$. We deduce from (2.3.5) that for all $n \geq 1$

$$V_{2n}^2 = \eta^t \Phi_n + v_{2n}$$

where $\eta^t = \begin{pmatrix} \sigma_a^2 & \sigma_c^2 \end{pmatrix}$. It leads us to estimate the vector of variances η by the WLS estimator

$$\widehat{\eta}_n = Q_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \frac{1}{d_k} \widehat{V}_{2k}^2 \Phi_k \tag{2.3.7}$$

where

$$Q_n = \sum_{k \in \mathbb{T}_n} \frac{1}{d_k} \Phi_k \Phi_k^t$$

and the weighting sequence $(d_n)_{n \geq 1}$ is given, for all $n \geq 1$, by $d_n = (1 + X_n)^2$. This choice is due to the fact that for all $n \geq 1$ and for all $k \in \mathbb{G}_n$

$$\begin{aligned} \mathbb{E}[v_{2k}^2 | \mathcal{F}_n] &= \mathbb{E}[V_{2k}^4 | \mathcal{F}_n] - \left(\mathbb{E}[V_{2k}^2 | \mathcal{F}_n] \right)^2 \qquad \text{a.s.} \\ &= 2\sigma_a^4 X_k^2 + (\mu_a^4 - 3\sigma_a^4 + 4\sigma_a^2 \sigma_c^2) X_k + \mu_c^4 - \sigma_c^4 \qquad \text{a.s.} \end{aligned} \tag{2.3.8}$$

where we recall that μ_a^4 is the fourth-order centered moment of $(Y_{n,i})$. Consequently, as $d_n \geq 1$, we clearly have for all $n \geq 1$ and for all $k \in \mathbb{G}_n$

$$\mathbb{E}[v_{2k}^2 | \mathcal{F}_n] \leq (\mu_a^4 - \sigma_a^4 + 4\sigma_a^2 \sigma_c^2 + \mu_c^4 - \sigma_c^4) d_k \qquad \text{a.s.}$$

We have a similar WLS estimator $\widehat{\zeta}_n$ of the vector of variances $\zeta^t = \begin{pmatrix} \sigma_b^2 & \sigma_d^2 \end{pmatrix}$ by replacing $\widehat{V}_{2k}^2$ by $\widehat{V}_{2k+1}^2$ into (2.3.7).

2.4 A martingale approach

In order to establish all the asymptotic properties of our estimators, we shall make use of a martingale approach. For all $n \geq 1$, denote

$$M_n = \sum_{k \in \mathbb{T}_{n-1}} \frac{1}{c_k} \begin{pmatrix} X_k V_{2k} \\ V_{2k} \\ X_k V_{2k+1} \\ V_{2k+1} \end{pmatrix}$$

where we recall that $c_k = (1 + X_k)$. We can clearly rewrite (2.3.4) as

$$\widehat{\theta}_n - \theta = \Sigma_{n-1}^{-1} M_n. \tag{2.4.1}$$

As in [11], we make use of the notation M_n since it appears that $(M_n)_{n\geq 1}$ a martingale. This fact is a crucial point of our study and it justifies the vector notation since most of asymptotic results for martingales were established for vector-valued martingales. Let us rewrite M_n in order to emphasize its martingale quality. Let $\Psi_n = I_2 \otimes \varphi_n$ where φ_n is the matrix of dimension 2×2^n given by

$$\varphi_n = \begin{pmatrix} \frac{X_{2^n}}{\sqrt{c_{2^n}}} & \frac{X_{2^n+1}}{\sqrt{c_{2^n+1}}} & \cdots & \frac{X_{2^{n+1}-1}}{\sqrt{c_{2^{n+1}-1}}} \\ \frac{1}{\sqrt{c_{2^n}}} & \frac{1}{\sqrt{c_{2^n+1}}} & \cdots & \frac{1}{\sqrt{c_{2^{n+1}-1}}} \end{pmatrix}.$$

It represents the individuals of the n-th generation which is also the collection of all $\Phi_k/\sqrt{c_k}$ where k belongs to $\mathbb{G}_n$. Let ξ_n be the random vector of dimension 2^n

$$\xi_n^t = \left(\frac{V_{2^n}}{\sqrt{c_{2^{n-1}}}} \quad \frac{V_{2^n+2}}{\sqrt{c_{2^{n-1}+1}}} \quad \cdots \quad \frac{V_{2^{n+1}-2}}{\sqrt{c_{2^n-1}}} \quad \frac{V_{2^n+1}}{\sqrt{c_{2^{n-1}}}} \quad \frac{V_{2^n+3}}{\sqrt{c_{2^{n-1}+1}}} \quad \cdots \quad \frac{V_{2^{n+1}-1}}{\sqrt{c_{2^n-1}}} \right).$$

The vector ξ_n gathers the noise variables of $\mathbb{G}_n$. The special ordering separating odd and even indices has been made in [11] so that M_n can be written as

$$M_n = \sum_{k=1}^{n} \Psi_{k-1}\xi_k$$

Under **(H.1)**, we clearly have for all $n \geq 0$, $\mathbb{E}[\xi_{n+1}|\mathcal{F}_n] = 0$ a.s. and Ψ_n is $\mathcal{F}_n$-measurable. In addition it is not hard to see that under **(H.1)** to **(H.3)**, (M_n) is a locally square integrable vector martingale with increasing process given, for all $n \geq 1$, by

$$\langle M \rangle_n = \sum_{k=0}^{n-1} \Psi_k \mathbb{E}[\xi_{k+1}\xi_{k+1}^t|\mathcal{F}_k]\Psi_k^t = \sum_{k=0}^{n-1} L_k \quad \text{a.s.} \tag{2.4.2}$$

where

$$L_k = \sum_{i\in\mathbb{G}_k} \frac{1}{c_i^2} \begin{pmatrix} \sigma_a^2 X_i + \sigma_c^2 & \rho \\ \rho & \sigma_b^2 X_i + \sigma_d^2 \end{pmatrix} \otimes \begin{pmatrix} X_i^2 & X_i \\ X_i & 1 \end{pmatrix}. \tag{2.4.3}$$

It is necessary to establish the convergence of $\langle M \rangle_n$, properly normalized, in order to prove the asymptotic results for our BINAR estimators $\widehat{\theta}_n$, $\widehat{\eta}_n$ and $\widehat{\zeta}_n$. Since the sizes of Ψ_n and ξ_n double at each generation, we have to adapt the proof of vector-valued martingale convergence given in [23] to our framework.

2.5 Main results

In all the sequel, we will assume that $\mathbb{P}_{\varepsilon_{2n}}$ and $\mathbb{P}_{\varepsilon_{2n+1}}$ do not depend on n, whereas, we allow that $\mathbb{P}_{(\varepsilon_{2n},\varepsilon_{2n+1})}$ depends on n. However, we shall get rid of the standard assumption commonly used in the INAR literature that the offspring sequences $(Y_{n,i})$ and $(Z_{n,i})$ share the same Bernoulli distribution. The only assumption that we will use here is that the offspring sequences $(Y_{n,i})$ and $(Z_{n,i})$ admit eighth-order moments. We have to introduce some more notations in order to state our main results. From the original process $(X_n)_{n\geq 1}$, we shall define a new process $(Y_n)_{n\geq 1}$ recursively defined by $Y_1 = X_1$, and if $Y_n = X_k$ with $n, k \geq 1$, then

$$Y_{n+1} = X_{2k+\kappa_n}$$

where $(\kappa_n)_{n\geq 1}$ is a sequence of i.i.d. random variables with Bernoulli $\mathcal{B}(1/2)$ distribution. Such a construction may be found in [26] for the asymptotic analysis of BAR processes. The process (Y_n) gathers the values of the original process (X_n) along the random branch of the binary tree $(\mathbb{T}_n)$ given by (κ_n). Denote by k_n the unique $k \geq 1$ such that $Y_n = X_k$. Then, for all $n \geq 1$, we have

$$Y_{n+1} = a_{n+1} \circ Y_n + e_{n+1} \tag{2.5.1}$$

where

$$a_{n+1} = \begin{cases} a \text{ if } \kappa_n = 0 \\ b \text{ otherwise} \end{cases} \qquad \text{and} \qquad e_n = \varepsilon_{k_n}. \tag{2.5.2}$$

Lemma 2.5.1. *Assume that* (ε_n) *satisfies* **(H.1)** *to* **(H.4)**. *Then, we have*

$$Y_n \xrightarrow{\mathcal{L}} T$$

where T is a positive non degenerate integer-valued random variable with $\mathbb{E}[T^8] < \infty$.

Denote $\mathcal{C}_3^1(\mathbb{R}_+) = \Big\{ f \in \mathcal{C}^1(\mathbb{R}_+, \mathbb{R}) \big| \exists \gamma > 0, \forall x \geq 0, (|f'(x)| + |f(x)|) \leq \gamma(1 + x^3) \Big\}$.

Lemma 2.5.2. *Assume that* (ε_n) *satisfies* **(H.1)** *to* **(H.5)**. *Then, for all $f \in \mathcal{C}_3^1(\mathbb{R}_+)$, we have*

$$\lim_{n\to\infty} \frac{1}{|\mathbb{T}_n|} \sum_{k\in\mathbb{T}_n} f(X_k) = \mathbb{E}[f(T)] \qquad a.s.$$

Remark 2.5.3. *We can easily deduce from Lemma 2.5.2 that under the same assumptions we have*

$$\lim_{n\to\infty} \frac{1}{|\mathbb{G}_n|} \sum_{k\in\mathbb{G}_n} f(X_k) = \mathbb{E}[f(T)] \qquad a.s.$$

Remark 2.5.4. *The set $\mathcal{C}_3^1(\mathbb{R}_+)$ is only necessary for the study of the estimator of ρ. For the other estimators the set $\mathcal{C}_1^1(\mathbb{R}_+)$ is sufficient.*

Proposition 2.5.5. *Assume that* (ε_n) *satisfies* **(H.1)** *to* **(H.5)**. *Then, we have*

$$\lim_{n\to\infty} \frac{\langle M\rangle_n}{|\mathbb{T}_{n-1}|} = L \qquad a.s. \tag{2.5.3}$$

where L *is the positive definite matrix given by*

$$L = \mathbb{E}\left[\frac{1}{(1+T)^2}\begin{pmatrix}\sigma_a^2 T + \sigma_c^2 & \rho \\ \rho & \sigma_b^2 T + \sigma_d^2\end{pmatrix} \otimes \begin{pmatrix} T^2 & T \\ T & 1\end{pmatrix}\right].$$

Our first result deals with the almost sure convergence of our WLS estimator $\widehat{\theta}_n$.

Theorem 2.5.6. *Assume that* (ε_n) *satisfies* **(H.1)** *to* **(H.5)**. *Then,* $\widehat{\theta}_n$ *converges almost surely to* θ *with the rate of convergence*

$$\|\widehat{\theta}_n - \theta\|^2 = o\left(\frac{n^\delta}{|\mathbb{T}_{n-1}|}\right) \qquad a.s., \quad \text{for all } \delta > 1/2 \tag{2.5.4}$$

In addition, we also have the quadratic strong law

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} |\mathbb{T}_{k-1}|(\widehat{\theta}_k - \theta)^t \Lambda(\widehat{\theta}_k - \theta) = tr(\Lambda^{-1/2} L \Lambda^{-1/2}) \qquad a.s. \tag{2.5.5}$$

where

$$\Lambda = I_2 \otimes A \qquad \text{and} \qquad A = \mathbb{E}\left[\frac{1}{1+T}\begin{pmatrix} T^2 & T \\ T & 1\end{pmatrix}\right] \quad \text{is positive-definite.} \tag{2.5.6}$$

Our second result concerns the almost sure asymptotic properties of our WLS variance and covariance estimators $\widehat{\eta}_n$, $\widehat{\zeta}_n$ and $\widehat{\rho}_n$. Let

$$\eta_n = Q_{n-1}^{-1} \sum_{k\in\mathbb{T}_{n-1}} \frac{1}{d_k} V_{2k}^2 \Phi_k,$$

$$\zeta_n = Q_{n-1}^{-1} \sum_{k\in\mathbb{T}_{n-1}} \frac{1}{d_k} V_{2k+1}^2 \Phi_k,$$

$$\rho_n = \frac{1}{|\mathbb{T}_{n-1}|} \sum_{k\in\mathbb{T}_{n-1}} V_{2k} V_{2k+1}.$$

Theorem 2.5.7. *Assume that* (ε_n) *satisfies* **(H.1)** *to* **(H.5)**. *Then,* $\widehat{\eta}_n$ *and* $\widehat{\zeta}_n$ *converge almost surely to* η *and* ζ *respectively. More precisely,*

$$\|\widehat{\eta}_n - \eta_n\| = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \qquad a.s. \tag{2.5.7}$$

$$\|\widehat{\zeta}_n - \zeta_n\| = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \qquad a.s. \tag{2.5.8}$$

In addition, $\widehat{\rho}_n$ *converges almost surely to* ρ *with*

$$\widehat{\rho}_n - \rho_n = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \qquad a.s. \tag{2.5.9}$$

Remark 2.5.8. *We also have the almost sure rates of convergence*

$$\|\widehat{\eta}_n-\eta\|^2 = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right), \quad \|\widehat{\zeta}_n-\zeta\|^2 = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right), \quad (\widehat{\rho}_n-\rho)^2 = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \quad a.s.$$

the rate of convergence being essentially given by $\|\eta_n - \eta\|$, $\|\zeta_n - \zeta\|$ *and* $\|\rho_n - \rho\|$.

Our last result is devoted to the asymptotic normality of our WLS estimators $\widehat{\theta}_n$, $\widehat{\eta}_n$, $\widehat{\zeta}_n$ and $\widehat{\rho}_n$.

Theorem 2.5.9. *Assume that* (ε_n) *satisfies* **(H.1)** *to* **(H.5)**. *Then, we have the asymptotic normality*

$$\sqrt{|\mathbb{T}_{n-1}|}(\widehat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, (I_2 \otimes A^{-1})L(I_2 \otimes A^{-1})). \tag{2.5.10}$$

In addition, we also have

$$\sqrt{|\mathbb{T}_{n-1}|}\,(\widehat{\eta}_n - \eta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, B^{-1}M_{ac}B^{-1}), \tag{2.5.11}$$

$$\sqrt{|\mathbb{T}_{n-1}|}\left(\widehat{\zeta}_n - \zeta\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, B^{-1}M_{bd}B^{-1}), \tag{2.5.12}$$

where

$$B = \mathbb{E}\left[\frac{1}{(1+T)^2}\begin{pmatrix} T^2 & T \\ T & 1 \end{pmatrix}\right],$$

$$M_{ac} = \mathbb{E}\left[\frac{2\sigma_a^4 T^2 + (\mu_a^4 - 3\sigma_a^4 + 4\sigma_a^2\sigma_c^2)T + \mu_c^4 - \sigma_c^4}{(1+T)^4}\begin{pmatrix} T^2 & T \\ T & 1 \end{pmatrix}\right],$$

$$M_{bd} = \mathbb{E}\left[\frac{2\sigma_b^4 T^2 + (\mu_b^4 - 3\sigma_b^4 + 4\sigma_b^2\sigma_d^2)T + \mu_d^4 - \sigma_d^4}{(1+T)^4}\begin{pmatrix} T^2 & T \\ T & 1 \end{pmatrix}\right].$$

Finally,

$$\sqrt{|\mathbb{T}_{n-1}|}\,(\widehat{\rho}_n - \rho) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \sigma_\rho^2\right) \tag{2.5.13}$$

where

$$\sigma_\rho^2 = \sigma_a^2\sigma_b^2\mathbb{E}[T^2] + \left(\sigma_a^2\sigma_d^2 + \sigma_b^2\sigma_c^2\right)\frac{\overline{c}}{1-\overline{a}} + \nu^2 - \rho^2, \tag{2.5.14}$$

$$\mathbb{E}[T^2] = \frac{\Upsilon\overline{c}}{1-\overline{a}} + \frac{\overline{c^2} - \Upsilon\overline{c}}{1-\overline{a^2}} + \frac{2(ac+bd)\overline{c}}{(1-\overline{a})(1-\overline{a^2})},$$

$$\Upsilon = \frac{\sigma_a^2 + \sigma_b^2}{2(\overline{a} - \overline{a^2})}, \qquad \overline{a} = \frac{a+b}{2}, \qquad \overline{a^2} = \frac{a^2+b^2}{2},$$

$$\overline{c} = \frac{c+d}{2}, \qquad \overline{c^2} = \frac{\sigma_c^2 + \sigma_d^2 + c^2 + d^2}{2}.$$

The rest of the paper is dedicated to the proof of our main results.

2.6 Proofs

2.6.1 Proof of Lemma 2.5.1

We can reformulate (2.5.1) and (2.5.2) as

$$Y_n = a_n \circ a_{n-1} \circ \ldots \circ a_2 \circ Y_1 + \sum_{k=2}^{n-1} a_n \circ a_{n-1} \circ \ldots \circ a_{k+1} \circ e_k + e_n.$$

We already made the assumption that $\mathbb{P}_{\varepsilon_{2n}}$ and $\mathbb{P}_{\varepsilon_{2n+1}}$ do not depend on n. Consequently, the couples $(a_k, e_k)_{k\in\{3,\ldots,n-1\}}$ and $(a_{n-k+2}, e_{n-k+2})_{k\in\{3,\ldots,n-1\}}$ share the same distribution. Hence, for all $n \geq 2$, Y_n has the same distribution than the random variable

$$\begin{aligned} Z_n &= a_2 \circ \ldots \circ a_n \circ Y_1 + \sum_{k=2}^{n-1} a_2 \circ a_3 \circ \ldots \circ a_{n-k+1} \circ e_{n-k+2} + e_2, \\ &= a_2 \circ \ldots \circ a_n \circ Y_1 + \sum_{k=3}^{n} a_2 \circ a_3 \circ \ldots \circ a_{k-1} \circ e_k + e_2. \end{aligned}$$

For the sake of simplicity, we will denote

$$Z_n = a_2 \circ \ldots \circ a_n \circ Y_1 + \sum_{k=2}^{n} a_2 \circ a_3 \circ \ldots \circ a_{k-1} \circ e_k. \tag{2.6.1}$$

For all $n \geq 2$ and for all $2 \leq k \leq n$, let

$$\Sigma_n^{n-k+2} = a_k \circ \ldots \circ a_n \circ Y_1$$

with $\Sigma_n^n = a_2 \circ \ldots \circ a_n \circ Y_1$ and $\Sigma_n^1 = Y_1$. We clearly have $\Sigma_n^{n-k+2} = a_k \circ \Sigma_n^{n-k+1}$. Consequently, it follows from the tower property of the conditional expectation that

$$\begin{aligned} \mathbb{E}[\Sigma_n^n] &= \mathbb{E}[a_2 \circ \Sigma_n^{n-1}] = \left(\mathbb{E}[a \circ \Sigma_n^{n-1}]\mathbb{P}(a_2 = a) + \mathbb{E}[b \circ \Sigma_n^{n-1}]\mathbb{P}(a_2 = b)\right), \\ &= \frac{1}{2}\left(\mathbb{E}\left[\mathbb{E}\left[\sum_{i=1}^{\Sigma_n^{n-1}} Y_{2,i} \middle| \Sigma_n^{n-1}\right]\right] + \mathbb{E}\left[\mathbb{E}\left[\sum_{i=1}^{\Sigma_n^{n-1}} Z_{2,i} \middle| \Sigma_n^{n-1}\right]\right]\right), \\ &= \frac{1}{2}\left(\mathbb{E}\left[\sum_{i=1}^{\Sigma_n^{n-1}} \mathbb{E}\left[Y_{2,i}\right]\right] + \mathbb{E}\left[\sum_{i=1}^{\Sigma_n^{n-1}} \mathbb{E}\left[Z_{2,i}\right]\right]\right), \\ &= \frac{1}{2}\left(\mathbb{E}[a\Sigma_n^{n-1}] + \mathbb{E}[b\Sigma_n^{n-1}]\right) = \overline{a}\mathbb{E}[\Sigma_n^{n-1}] = \ldots \\ &= \overline{a}^{n-1}\mathbb{E}[\Sigma_n^1] = \overline{a}^{n-1}\mathbb{E}[Y_1]. \end{aligned}$$

The stability hypothesis $0 < \max(a, b) < 1$ implies that $0 < \overline{a} < 1$ which leads to

$$\sum_{n=2}^{\infty} \mathbb{E}[\Sigma_n^n] = \mathbb{E}[Y_1] \sum_{n=2}^{\infty} \overline{a}^{n-1} = \frac{\mathbb{E}[Y_1]\overline{a}}{1 - \overline{a}} \geq 0.$$

Then, we obtain from the monotone convergence theorem that

$$\lim_{n\to\infty} \Sigma_n^n = 0 \qquad \text{a.s.} \tag{2.6.2}$$

It now remains to study the right-hand side sum in (2.6.1). For all $n \geq 2$, denote

$$T_n = \sum_{k=2}^{n} a_2 \circ \ldots \circ a_{k-1} \circ e_k.$$

By the same calculation as before, we have for all $n \geq 2$

$$\mathbb{E}[T_n] = \sum_{k=2}^{n} \overline{a}^{k-2}\mathbb{E}[e_k] = \overline{c}\sum_{k=0}^{n-2} \overline{a}^k,$$

which implies that

$$\lim_{n\to\infty} \mathbb{E}[T_n] = \frac{\overline{c}}{1-\overline{a}}.$$

Hence, we deduce once again from the monotone convergence theorem that the positive increasing sequence $(T_n)_{n\geq 2}$ converges almost surely to

$$T = \sum_{k=2}^{\infty} a_2 \circ \ldots \circ a_{k-1} \circ e_k$$

which is almost surely finite as $\mathbb{E}[T] < \infty$. Therefore, we can conclude from (2.6.1) and (2.6.2) that

$$\lim_{n\to\infty} Z_n = T \qquad \text{a.s.}$$

leading to

$$Y_n \xrightarrow{\mathcal{L}} T.$$

Let us prove that $\mathbb{E}[T^3] < \infty$. First of all, we already saw that

$$\mathbb{E}[a_2 \circ \ldots \circ a_n \circ e_{n+1}] = \overline{a}^{n-1}\mathbb{E}[e_2] = \overline{a}^{n-1}\overline{c}.$$

In addition,

$$\begin{aligned}
\mathbb{E}[(\Sigma_n^n)^2] &= \frac{1}{2}\left(\mathbb{E}\left[(a\circ\Sigma_n^{n-1})^2\right] + \mathbb{E}\left[(b\circ\Sigma_n^{n-1})^2\right]\right),\\
&= \frac{1}{2}\left(\mathbb{E}\left[\mathbb{E}\left[\left(\sum_{i=1}^{\Sigma_n^{n-1}} Y_{2,i}\right)^2 \middle| \Sigma_n^{n-1}\right]\right] + \mathbb{E}\left[\mathbb{E}\left[\left(\sum_{i=1}^{\Sigma_n^{n-1}} Z_{2,i}\right)^2 \middle| \Sigma_n^{n-1}\right]\right]\right),
\end{aligned}$$

and the first expectation is

$$\mathbb{E}\left[\mathbb{E}\left[\left(\sum_{i=1}^{\Sigma_n^{n-1}} Y_{2,i}\right)^2 \middle| \Sigma_n^{n-1}\right]\right] = \mathbb{E}\left[\mathbb{E}\left[\sum_{i=1}^{\Sigma_n^{n-1}} Y_{2,i}^2 + \sum_{i=1}^{\Sigma_n^{n-1}} \sum_{\substack{j=1\\ j\neq i}}^{\Sigma_n^{n-1}} Y_{2,i}Y_{2,j} \middle| \Sigma_n^{n-1}\right]\right],$$

$$= \mathbb{E}\left[\sum_{i=1}^{\Sigma_n^{n-1}} \mathbb{E}[Y_{2,i}^2] + \sum_{i=1}^{\Sigma_n^{n-1}} \sum_{\substack{j=1\\ j\neq i}}^{\Sigma_n^{n-1}} \mathbb{E}[Y_{2,i}]\mathbb{E}[Y_{2,j}]\right],$$

$$= \mathbb{E}[\Sigma_n^{n-1}(\sigma_a^2 + a^2) + \Sigma_n^{n-1}(\Sigma_n^{n-1} - 1)a^2],$$

$$= \mathbb{E}[\Sigma_n^{n-1}]\sigma_a^2 + a^2\mathbb{E}[(\Sigma_n^{n-1})^2].$$

Since the computation of the second expectation is exactly the same, we obtain

$$\begin{aligned}
\mathbb{E}[(\Sigma_n^n)^2] &= \mathbb{E}[\Sigma_n^{n-1}]\frac{\sigma_a^2+\sigma_b^2}{2} + \overline{a^2}\mathbb{E}[(\Sigma_n^{n-1})^2], \\
&= \overline{a}^{n-2}\frac{\sigma_a^2+\sigma_b^2}{2}\mathbb{E}[Y_1] + \overline{a^2}\mathbb{E}[(\Sigma_n^{n-1})^2] = \dots \\
&= \left(\sum_{i=0}^{n-2} \overline{a}^{n-i-2}\overline{a^2}^i\right)\frac{\sigma_a^2+\sigma_b^2}{2}\mathbb{E}[Y_1] + \overline{a^2}^{n-1}\mathbb{E}[(\Sigma_n^1)^2], \\
&= \frac{\overline{a}^{n-1} - \overline{a^2}^{n-1}}{\overline{a} - \overline{a^2}}\frac{\sigma_a^2+\sigma_b^2}{2}\mathbb{E}[Y_1] + \overline{a^2}^{n-1}\mathbb{E}[Y_1^2], \\
&= (\overline{a}^{n-1} - \overline{a^2}^{n-1})\Upsilon\mathbb{E}[Y_1] + \overline{a^2}^{n-1}\mathbb{E}[Y_1^2]
\end{aligned}$$

where

$$\Upsilon = \frac{\sigma_a^2+\sigma_b^2}{2(\overline{a} - \overline{a^2})} > 0 \qquad \text{since} \qquad \overline{a^2} < \overline{a}.$$

In the same way, we can prove that

$$\mathbb{E}[(a_2 \circ \ldots \circ a_n \circ e_{n+1})^2] = (\overline{a}^{n-1} - \overline{a^2}^{n-1})\Upsilon\overline{c} + \overline{a^2}^{n-1}\overline{c^2}.$$

Consequently, as (e_n) is an integer-valued random variable,

$$\mathbb{E}[(a_2 \circ \ldots \circ a_n \circ e_{n+1})^2] \leq \overline{a}^{n-1}(\Upsilon\overline{c} + \overline{c^2}) \leq \overline{a}^{n-1}(\Upsilon + 1)\overline{c^2}.$$

Furthermore, we obtain from tedious but straightforward calculations that it exists some constant $\xi > 0$ such that for all $2 \leq p \leq 8$

$$\mathbb{E}[(a_2 \circ \ldots \circ a_n \circ e_{n+1})^p] \leq \xi\mathbb{E}[e_2^p]\overline{a}^{n-1}. \tag{2.6.3}$$

One can observe that the constant ξ only depends on the moments of $(Y_{n,i})$ and $(Z_{n,i})$ up to order 8. Hence, as $0 < \overline{a} < 1$, we deduce from (2.6.3) and the triangle inequality that

$$\mathbb{E}[T^8]^{1/8} \leq \sum_{k=2}^{\infty} \mathbb{E}\left[(a_2 \circ \ldots \circ a_{k-1} \circ e_k)^8\right]^{1/8},$$
$$\leq \xi^{1/8}\mathbb{E}[e_2^8]^{1/8} \sum_{k=2}^{\infty} \overline{a}^{(k-2)/8} < \infty$$

which immediately leads to $\mathbb{E}[T^3] < \infty$. Let us now prove that T is not degenerate. First one can observe that $a \circ T + e \sim T$ where $(a, e) \sim (a_2, e_2)$. Then, if T was degenerate, let us say $T = c$ a.s., then $a \circ c + e$ would be a sum of two independent random variables which sum would be constant, which implies that a and e would be degenerate. Since we assumed that ε_2 and ε_3 are not degenerate, e can not be degenerate which allows us to say that T is not degenerate.
Finally, let us compute $\mathbb{E}[T^2]$ which will be useful for (2.5.14)

$$\begin{aligned}\mathbb{E}[T^2] &= \mathbb{E}\left[\left(\sum_{k=2}^{\infty} a_2 \circ \ldots \circ a_{k-1} \circ e_k\right)^2\right],\\ &= \sum_{k=2}^{\infty} \mathbb{E}\left[(a_2 \circ \ldots \circ a_{k-1} \circ e_k)^2\right]\\ &\quad + 2\sum_{k=2}^{\infty}\sum_{l=k+1}^{\infty} \mathbb{E}\left[(a_2 \circ \ldots \circ a_{k-1} \circ e_k)(a_2 \circ \ldots \circ a_{l-1} \circ e_l)\right]\end{aligned}$$

We already saw that

$$\mathbb{E}\left[(a_2 \circ \ldots \circ a_{k-1} \circ e_k)^2\right] = (\overline{a}^{k-2} - \overline{a^2}^{k-2})\Upsilon\overline{c} + \overline{a^2}^{k-2}\overline{c^2}.$$

Moreover, we have, for all $l \geq 3$

$$\begin{aligned}\mathbb{E}\left[e_2(a_2 \circ \ldots \circ a_{l-1} \circ e_l)\right] &= \frac{1}{2}\mathbb{E}\left[\varepsilon_2(a \circ \ldots \circ a_{l-1} \circ e_l)\right] + \frac{1}{2}\mathbb{E}\left[\varepsilon_3(b \circ \ldots \circ a_{l-1} \circ e_l)\right],\\ &= \frac{1}{2}\left(\mathbb{E}\left[\varepsilon_2\right]\mathbb{E}[[(a \circ \ldots \circ a_{l-1} \circ e_l)] + \mathbb{E}\left[\varepsilon_3\right]\mathbb{E}[[(b \circ \ldots \circ a_{l-1} \circ e_l)]\right),\\ &= \frac{1}{2}\left(c(a\overline{a}^{l-3}\overline{c}) + d(b\overline{a}^{l-3}\overline{c})\right),\\ &= \frac{(ac+bd)\overline{c}}{2}\overline{a}^{l-3}.\end{aligned}$$

In addition, for all $k \geq 2$ and for all $l \geq k+1$

$$\begin{aligned}&\mathbb{E}\left[(a_2 \circ \ldots \circ a_{k-1} \circ e_k)(a_2 \circ \ldots \circ a_{l-1} \circ e_l)\right]\\ &\quad= \frac{1}{2}\mathbb{E}\left[(a \circ \ldots \circ a_{k-1} \circ e_k)(a \circ \ldots \circ a_{l-1} \circ e_l)\right]\\ &\qquad+ \frac{1}{2}\mathbb{E}\left[(b \circ \ldots \circ a_{k-1} \circ e_k)(b \circ \ldots \circ a_{l-1} \circ e_l)\right].\end{aligned}$$

Let us tackle the first term

$$
\begin{aligned}
&\mathbb{E}\left[(a \circ \ldots \circ a_{k-1} \circ e_k)\,(a \circ \ldots \circ a_{l-1} \circ e_l)\right] \\
&\qquad = \mathbb{E}\left[\left(\sum_{i=1}^{a_3\circ\ldots\circ a_{k-1}\circ e_k} Y_{2,i}\right)\left(\sum_{j=1}^{a_3\circ\ldots\circ a_{l-1}\circ e_l} Y_{2,j}\right)\right], \\
&\qquad = \mathbb{E}\left[\sum_{i=1}^{a_3\circ\ldots\circ a_{k-1}\circ e_k}\sum_{j=1}^{a_3\circ\ldots\circ a_{l-1}\circ e_l} \mathbb{E}[Y_{2,i}Y_{2,j}|a_3 \circ \ldots \circ a_{k-1} \circ e_k, a_3 \circ \ldots \circ a_{l-1} \circ e_l]\right], \\
&\qquad = \mathbb{E}\left[\sum_{i=1}^{a_3\circ\ldots\circ a_{k-1}\circ e_k}\sum_{j=1}^{a_3\circ\ldots\circ a_{l-1}\circ e_l} \mathbb{E}[Y_{2,i}Y_{2,j}]\right], \\
&\qquad = \mathbb{E}\left[\sum_{i=1}^{a_3\circ\ldots\circ a_{k-1}\circ e_k}\sum_{j=1}^{a_3\circ\ldots\circ a_{l-1}\circ e_l} a^2\right], \\
&\qquad = a^2\mathbb{E}\left[(a_3 \circ \ldots \circ a_{k-1} \circ e_k)\,(a_3 \circ \ldots \circ a_{l-1} \circ e_l)\right].
\end{aligned}
$$

Hence, we obtained that

$$
\begin{aligned}
&\mathbb{E}\left[(a_2 \circ \ldots \circ a_{k-1} \circ e_k)\,(a_2 \circ \ldots \circ a_{l-1} \circ e_l)\right] \\
&\qquad = \overline{a^2}\mathbb{E}\left[(a_3 \circ \ldots \circ a_{k-1} \circ e_k)\,(a_3 \circ \ldots \circ a_{l-1} \circ e_l)\right], \\
&\qquad = \overline{a^2}^{k-2}\mathbb{E}\left[e_k\,(a_k \circ \ldots \circ a_{l-1} \circ e_l)\right], \\
&\qquad = \overline{a^2}^{k-2}\frac{(ac+bd)\overline{c}}{2}\overline{a}^{l-k-1}.
\end{aligned}
$$

Finally, we have

$$
\begin{aligned}
\mathbb{E}[T^2] &= \sum_{k=2}^{\infty}\left((\overline{a}^{k-2} - \overline{a^2}^{k-2})\Upsilon\overline{c} + \overline{a^2}^{k-2}\overline{c^2}\right) + 2\sum_{l=3}^{\infty}\frac{(ac+bd)\overline{c}}{2}\overline{a}^{l-3} \\
&\qquad + 2\sum_{k=3}^{\infty}\sum_{l=k+1}^{\infty}\overline{a^2}^{k-2}\frac{(ac+bd)\overline{c}}{2}\overline{a}^{l-k-1}, \\
&= \Upsilon\overline{c}\left(\frac{1}{1-\overline{a}} - \frac{1}{1-\overline{a^2}}\right) + \frac{\overline{c^2}}{1-\overline{a^2}} + \frac{(ac+bd)\overline{c}}{1-\overline{a}}\left(1 + \frac{\overline{a^2}}{1-\overline{a^2}}\right), \\
&= \Upsilon\overline{c}\left(\frac{1}{1-\overline{a}} - \frac{1}{1-\overline{a^2}}\right) + \frac{\overline{c^2}}{1-\overline{a^2}} + \frac{(ac+bd)\overline{c}}{(1-\overline{a})(1-\overline{a^2})}.
\end{aligned}
$$

2.6.2 Proof of the keystone Lemma 2.5.2

We shall now prove that for all $f \in \mathcal{C}_3^1(\mathbb{R}_+)$,

$$
\lim_{n\to\infty}\frac{1}{|\mathbb{T}_n|}\sum_{k\in\mathbb{T}_n} f(X_k) = \mathbb{E}[f(T)]. \tag{2.6.4}
$$

Denote $g = f - \mathbb{E}[f(T)]$,

$$\overline{M}_{\mathbb{T}_n}(f) = \frac{1}{|\mathbb{T}_n|}\sum_{k\in\mathbb{T}_n} f(X_k) \qquad \text{and} \qquad \overline{M}_{\mathbb{G}_n}(f) = \frac{1}{|\mathbb{G}_n|}\sum_{k\in\mathbb{G}_n} f(X_k).$$

Via Lemma A.2 of [11], it is only necessary to prove that

$$\lim_{n\to\infty} \frac{1}{|\mathbb{G}_n|}\sum_{k\in\mathbb{G}_n} g(X_k) = 0 \qquad \text{a.s.}$$

We shall follow the induced Markov chain approach, originally proposed by Guyon in [26]. Let Q be the transition probability of (Y_n), Q^p the p-th iterated of Q. In addition, denote by ν the distribution of $Y_1 = X_1$ and νQ^p the law of Y_p. Finally, let P be the transition probability of (X_n) as defined in [26]. We obtain from relation (7) of [26] that for all $n \geq 0$

$$\mathbb{E}[\overline{M}_{\mathbb{G}_n}(g)^2] = \frac{1}{2^n}\nu Q^n g^2 + \sum_{k=0}^{n-1} \frac{1}{2^{k+1}}\nu Q^k P(Q^{n-k-1}g \star Q^{n-k-1}g)$$

where, for all $x, y \in \mathbb{N}$, $(f \star g)(x, y) = f(x)g(y)$. Consequently,

$$\begin{aligned}\sum_{n=0}^{\infty}\mathbb{E}[\overline{M}_{\mathbb{G}_n}(g)^2] &= \sum_{n=0}^{\infty}\frac{1}{2^n}\nu Q^n g^2 + \sum_{n=1}^{\infty}\sum_{k=0}^{n-1}\frac{1}{2^{k+1}}\nu Q^k P(Q^{n-k-1}g \star Q^{n-k-1}g),\\ &\leq \sum_{k=0}^{\infty}\frac{1}{2^k}\nu Q^k\left(g^2 + P\left(\sum_{l=0}^{\infty}|Q^l g \star Q^l g|\right)\right).\end{aligned}$$

However, for all $x \in \mathbb{N}$,

$$Q^n g(x) = Q^n f(x) - \mathbb{E}[f(T)] = \mathbb{E}_x[f(Y_n) - f(T)] = \mathbb{E}_x[f(Z_n) - f(T)]$$

where Z_n is given by (2.6.1). Hence, we deduce from the mean value theorem and Cauchy-Schwarz inequality that

$$|Q^n g(x)| \leq \mathbb{E}_x[W_n|Z_n - T|] \leq \mathbb{E}_x[W_n^2]^{1/2}\mathbb{E}_x[(Z_n - T)^2]^{1/2} \tag{2.6.5}$$

where

$$W_n = \sup_{z\in[Z_n,T]} |f'(z)|.$$

By the very definition of $\mathcal{C}_3^1(\mathbb{R}_+)$, one can find some constant $\alpha > 0$ such that $|f'(z)|^2 \leq \alpha(1 + z^6)$. Hence, it exists some constant $\beta > 0$ such that

$$\begin{aligned}\mathbb{E}_x[W_n^2] &\leq \alpha\mathbb{E}_x[1 + Z_n^6 + T^6] = \alpha(1 + \mathbb{E}_x[Z_n^6] + \mathbb{E}[T^6]),\\ &\leq \beta(1 + x^6).\end{aligned} \tag{2.6.6}$$

As a matter of fact, under hypotheses **(H.1)** to **(H.5)**, it exists some constant $\gamma > 0$ such that $\mathbb{E}_x[Z_n^6] < \gamma(1+x^6)$ and we already saw that $\mathbb{E}[T^8] < \infty$. In fact, we infer from (2.6.1) that

$$\begin{aligned}\mathbb{E}_x[Z_n^6]^{1/6} &\leq \mathbb{E}_x[(a_2 \circ \ldots \circ a_n \circ Y_1)^6]^{1/6} + \sum_{k=2}^{n} \mathbb{E}_x\left[(a_2 \circ a_3 \circ \ldots \circ a_{k-1} \circ e_k)^6\right]^{1/6},\\ &\leq \xi^{1/6}\mathbb{E}_x[Y_1^6]^{1/6}\overline{a}^{(n-1)/6} + \sum_{k=2}^{\infty} \mathbb{E}\left[(a_2 \circ a_3 \circ \ldots \circ a_{k-1} \circ e_k)^6\right]^{1/6},\\ &\leq \xi^{1/6}x + \sum_{k=2}^{\infty} \mathbb{E}\left[(a_2 \circ a_3 \circ \ldots \circ a_{k-1} \circ e_k)^6\right]^{1/6}\end{aligned}$$

and we have already proved that the sum in the right-hand term is finite. So we can conclude that there exists some constant $\gamma > 0$ such that $\mathbb{E}_x[Z_n^6] < \gamma(1+x^6)$. Furthermore

$$Z_n - T = a_2 \circ \ldots a_n \circ Y_1 - \sum_{k=n+1}^{\infty} a_2 \circ \ldots \circ a_k \circ e_{k+1}$$

and the triangle inequality allows us to say that

$$\mathbb{E}_x[(Z_n - T)^2]^{1/2} \leq \mathbb{E}_x[(a_2 \circ \ldots a_n \circ Y_1)^2]^{1/2} + \sum_{k=n+1}^{\infty} \mathbb{E}_x[(a_2 \circ \ldots \circ a_k \circ e_{k+1})^2]^{1/2}.$$

We already saw in section 2.6.1 that

$$\begin{aligned}\mathbb{E}_x[(a_2 \circ \ldots a_n \circ Y_1)^2] &= (\overline{a}^{n-1} - \overline{a^2}^{n-1})\Upsilon\mathbb{E}_x[Y_1] + \overline{a^2}^{n-1}\mathbb{E}_x[Y_1^2],\\ &= (\overline{a}^{n-1} - \overline{a^2}^{n-1})\Upsilon x + \overline{a^2}^{n-1}x^2 = x(\Upsilon\overline{a}^{n-1} + \overline{a^2}^{n-1}(x - \Upsilon))\end{aligned}$$

and

$$\mathbb{E}_x[(a_2 \circ \ldots \circ a_k \circ e_{k+1})^2] = (\overline{a}^{k-1} - \overline{a^2}^{k-1})\Upsilon\overline{c} + \overline{a^2}^{k-1}\overline{c^2}.$$

Hence

$$\begin{aligned}\sum_{k=n+1}^{\infty} \mathbb{E}_x[(a_2 \circ \ldots \circ a_k \circ e_{k+1})^2]^{1/2} &= \sum_{k=n+1}^{\infty} \left(\overline{a}^{k-1}\Upsilon\overline{c} + \overline{a^2}^{k-1}\left(\overline{c^2} - \Upsilon\overline{c}\right)\right)^{1/2},\\ &\leq \sum_{k=n+1}^{\infty} \left(\overline{a}^{k-1}\overline{c} + \overline{a}^{k-1}\left|\overline{c^2} - \Upsilon\overline{c}\right|\right)^{1/2},\\ &\leq \sum_{k=n+1}^{\infty} \sqrt{\overline{a}}^{k-1}\delta = \delta\frac{\sqrt{\overline{a}}^n}{1-\sqrt{\overline{a}}}.\end{aligned}$$

where

$$\delta = \sqrt{\max(\overline{c^2}, (1+\Upsilon)\overline{c} - \overline{c^2})}.$$

To sum up, we find that

$$\begin{aligned}
\mathbb{E}_x[(Z_n - T)^2]^{1/2} &\leq \sqrt{x}\left(\Upsilon\overline{a}^{n-1} + \overline{a^2}^{n-1}(x-\Upsilon)\right)^{1/2} + \frac{\delta}{1-\sqrt{\overline{a}}}\sqrt{\overline{a}}^n, \\
&\leq \sqrt{x}\left(\Upsilon\left(\overline{a}^{n-1} - \overline{a^2}^{n-1}\right) + x\overline{a^2}^{n-1}\right)^{\frac{1}{2}} + \frac{\delta}{1-\sqrt{\overline{a}}}\sqrt{\overline{a}}^n, \\
&\leq \sqrt{x}\sqrt{\Upsilon}\sqrt{\overline{a}}^{n-1} + \frac{\delta}{1-\sqrt{\overline{a}}}\sqrt{\overline{a}}^n, \\
&\leq \sqrt{\overline{a}}^n(1+x)\left(\frac{\sqrt{\Upsilon}}{2\sqrt{\overline{a}}} + \frac{\delta}{1-\sqrt{\overline{a}}}\right). \qquad (2.6.7)
\end{aligned}$$

Finally, we obtain from (2.6.5) together with (2.6.6) and (2.6.7) that for some constant $\kappa > 0$

$$|Q^n g(x)| \leq \sqrt{\beta}(1+x^6)^{1/2}\sqrt{\overline{a}}^{n-1}(1+x)\left(\frac{\sqrt{\Upsilon}}{2} + \frac{\delta}{1-\sqrt{\overline{a}}}\right) \leq \sqrt{\overline{a}}^n\kappa(1+x^4).$$

Therefore,

$$P\left(\sum_{n=0}^{\infty}|Q^n g \star Q^n g|\right) \leq \frac{\kappa^2}{1-\overline{a}}P(h\star h)$$

where, for all $x \in \mathbb{N}$, $h(x) = 1 + x^4$. We are now in position to prove that

$$\mathbb{E}\left[\sum_{n=0}^{\infty}\overline{M}_{\mathbb{G}_n}(g)^2\right] < \infty. \qquad (2.6.8)$$

Let

$$g := x \mapsto (a \circ x + \varepsilon_2, b \circ x + \varepsilon_3),$$

then we have $P(h\star h)(x) = \mathbb{E}\left[h \star h(g(x))\right]$ and it is not hard to see that, from hypothesis **(H.5)**, it exists some constant $\lambda > 0$ such that for all $x \in \mathbb{N}$, $P(h\star h)(x) \leq \lambda(1+x^8)$. Consequently, it exists some constant $\mu > 0$ such that

$$\begin{aligned}
\sum_{n=0}^{\infty}\mathbb{E}\left[\overline{M}_{\mathbb{G}_n}(g)^2\right] &\leq \sum_{k=0}^{\infty}\frac{1}{2^k}\nu Q^k\left(g^2 + P\left(\sum_{l=0}^{\infty}|Q^l g \star Q^l g|\right)\right), \\
&\leq \sum_{k=0}^{\infty}\frac{1}{2^k}\left(\mathbb{E}[g^2(Y_k)] + \frac{\lambda\kappa^2}{1-\overline{a}}(1+\mathbb{E}[Y_k^8])\right), \\
&\leq \left(2\mu + \frac{\lambda\kappa^2}{1-\overline{a}}\right)\left(2 + \sum_{k=0}^{\infty}\frac{1}{2^k}\mathbb{E}[Y_k^8]\right). \qquad (2.6.9)
\end{aligned}$$

Furthermore, we can deduce from (2.6.3) that it exists some constant ζ such that

$$\begin{aligned}
\mathbb{E}[Y_n^8]^{1/8} &\leq \mathbb{E}\left[(a_2 \circ \dots a_n \circ Y_1)^8\right]^{1/8} + \sum_{k=2}^{n} \mathbb{E}\left[(a_2 \circ \dots a_{k-1} \circ e_k)^8\right]^{1/8}, \\
&\leq \mathbb{E}\left[(a_2 \circ \dots a_n \circ Y_1)^8\right]^{1/8} + \xi^{1/8}\mathbb{E}[e_2^8]^{1/8} \sum_{k=2}^{n} \overline{a}^{k-2}, \\
&\leq \zeta^{1/8}\mathbb{E}[Y_1^8]^{1/8}\overline{a}^{n-1} + \frac{\xi^{1/8}\mathbb{E}[e_2^8]^{1/8}}{1-\overline{a}}, \\
&\leq \frac{\zeta^{1/8}\mathbb{E}[Y_1^8]^{1/8} + \xi^{1/8}\mathbb{E}[e_2^8]^{1/8}}{1-\overline{a}}. \qquad (2.6.10)
\end{aligned}$$

Then, (2.6.9) and (2.6.10) immediately lead to (2.6.8). Finally, the monotone convergence theorem implies that

$$\lim_{n\to\infty} \overline{M}_{\mathbb{G}_n}(g) = 0 \qquad \text{a.s.}$$

which completes the proof of Lemma 2.5.2.

2.6.3 Proof of Proposition 2.5.5 of the convergence of $\langle M \rangle_n$

The almost sure convergence (2.5.3) immediately follows from (2.4.2) and (2.4.3) together with Lemma 2.5.2. It only remains to prove that $\det(L) > 0$ where the limiting matrix L can be rewritten as

$$L = \mathbb{E}\left[\Gamma \otimes \mathcal{B}\right]$$

where

$$\Gamma = \begin{pmatrix} \sigma_a^2 T + \sigma_c^2 & \rho \\ \rho & \sigma_b^2 T + \sigma_d^2 \end{pmatrix} \qquad \text{and} \qquad \mathcal{B} = \begin{pmatrix} \dfrac{T^2}{(1+T)^2} & \dfrac{T}{(1+T)^2} \\ \dfrac{T}{(1+T)^2} & \dfrac{1}{(1+T)^2} \end{pmatrix}.$$

We have

$$\begin{aligned}
L &= \mathbb{E}\left[\begin{pmatrix} \sigma_a^2 T & 0 \\ 0 & \sigma_b^2 T \end{pmatrix} \otimes \mathcal{B}\right] + \mathbb{E}\left[\begin{pmatrix} \sigma_c^2 & \rho \\ \rho & \sigma_d^2 \end{pmatrix} \otimes \mathcal{B}\right], \\
&= \begin{pmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_b^2 \end{pmatrix} \otimes \mathbb{E}[T\mathcal{B}] + \begin{pmatrix} \sigma_c^2 & \rho \\ \rho & \sigma_d^2 \end{pmatrix} \otimes \mathbb{E}[\mathcal{B}]. \qquad (2.6.11)
\end{aligned}$$

We shall prove that $\mathbb{E}[\mathcal{B}]$ is a positive definite matrix and that $\mathbb{E}[T\mathcal{B}]$ is a positive semidefinite matrix. Denote by λ_1 and λ_2 the two eigenvalues of the real symmetric matrix $\mathbb{E}[\mathcal{B}]$. We clearly have

$$\lambda_1 + \lambda_2 = tr(\mathbb{E}[\mathcal{B}]) = \mathbb{E}\left[\frac{T^2+1}{(1+T)^2}\right] > 0$$

and

$$\lambda_1\lambda_2 = \det(\mathbb{E}[\mathcal{B}]) = \mathbb{E}\left[\frac{T^2}{(1+T)^2}\right]\mathbb{E}\left[\frac{1}{(1+T)^2}\right] - \mathbb{E}\left[\frac{T}{(1+T)^2}\right]^2 \geq 0$$

thanks to the Cauchy-Schwarz inequality and $\lambda_1\lambda_2 = 0$ if and only if T is degenerate, which is not the case thanks to Lemma 2.5.1. Consequently, $\mathbb{E}[\mathcal{B}]$ is a positive definite matrix. In the same way, we can prove that $\mathbb{E}[T\mathcal{B}]$ is a positive semidefinite matrix. Since the Kronecker product of two positive semidefinite (respectively definite positive) matrices is a positive semidefinite (respectively positive definite) matrix, we deduce from (2.6.11) that L is positive definite as soon as $\rho^2 < \sigma_c^2\sigma_d^2$ which is the case thanks to **(H.3)**.

2.6.4 Preliminary work for the almost sure convergence of θ_n

We will follow the same approach as in Bercu et al. [11]. For all $n \geq 1$, let $\mathcal{V}_n = M_n^t\Sigma_{n-1}^{-1}M_n = (\widehat{\theta}_n - \theta)^t\Sigma_{n-1}(\widehat{\theta}_n - \theta)$. First of all, we have

$$\begin{aligned}
\mathcal{V}_{n+1} &= M_{n+1}^t\Sigma_n^{-1}M_{n+1} = (M_n + \Delta M_{n+1})^t\Sigma_n^{-1}(M_n + \Delta M_{n+1}),\\
&= M_n^t\Sigma_n^{-1}M_n + 2M_n^t\Sigma_n^{-1}\Delta M_{n+1} + \Delta M_{n+1}^t\Sigma_n^{-1}\Delta M_{n+1},\\
&= \mathcal{V}_n - M_n^t(\Sigma_{n-1}^{-1} - \Sigma_n^{-1})M_n + 2M_n^t\Sigma_n^{-1}\Delta M_{n+1} + \Delta M_{n+1}^t\Sigma_n^{-1}\Delta M_{n+1}.
\end{aligned}$$

By summing over this identity, we obtain the main decomposition

$$\mathcal{V}_{n+1} + \mathcal{A}_n = \mathcal{V}_1 + \mathcal{B}_{n+1} + \mathcal{W}_{n+1} \tag{2.6.12}$$

where

$$\mathcal{A}_n = \sum_{k=1}^n M_k^t(\Sigma_{k-1}^{-1} - \Sigma_k^{-1})M_k,$$

$$\mathcal{B}_{n+1} = 2\sum_{k=1}^n M_k^t\Sigma_k^{-1}\Delta M_{k+1} \quad \text{and} \quad \mathcal{W}_{n+1} = \sum_{k=1}^n \Delta M_{k+1}^t\Sigma_k^{-1}\Delta M_{k+1}.$$

Lemma 2.6.1. *Assume that (ε_n) satisfies* **(H.1)** *to* **(H.5)**. *Then, we have*

$$\lim_{n\to\infty}\frac{\mathcal{W}_n}{n} = \frac{1}{2}tr((I_2\otimes A)^{-1/2}L(I_2\otimes A)^{-1/2}) \qquad a.s. \tag{2.6.13}$$

where A is the positive definite matrix given by (2.5.6). *In addition, we also have*

$$\mathcal{B}_{n+1} = o(n) \qquad a.s. \tag{2.6.14}$$

and

$$\lim_{n\to\infty}\frac{\mathcal{V}_{n+1}+\mathcal{A}_n}{n} = \frac{1}{2}tr((I_2\otimes A)^{-1/2}L(I_2\otimes A)^{-1/2}) \qquad a.s. \tag{2.6.15}$$

Proof. First of all, we have $\mathcal{W}_{n+1} = \mathcal{T}_{n+1} + \mathcal{R}_{n+1}$ where

$$\mathcal{T}_{n+1} = \sum_{k=1}^{n} \frac{\Delta M_{k+1}^t (I_2 \otimes A)^{-1} \Delta M_{k+1}}{|\mathbb{T}_k|},$$

$$\mathcal{R}_{n+1} = \sum_{k=1}^{n} \frac{\Delta M_{k+1}^t (|\mathbb{T}_k| \Sigma_k^{-1} - (I_2 \otimes A)^{-1}) \Delta M_{k+1}}{|\mathbb{T}_k|}.$$

One can observe that $\mathcal{T}_{n+1} = tr((I_2 \otimes A)^{-1/2} \mathcal{H}_{n+1} (I_2 \otimes A)^{-1/2})$ where

$$\mathcal{H}_{n+1} = \sum_{k=1}^{n} \frac{\Delta M_{k+1} \Delta M_{k+1}^t}{|\mathbb{T}_k|}.$$

Our aim is to make use of the strong law of large numbers for martingale transforms, so we start by adding and subtracting a term involving the conditional expectation of $\Delta \mathcal{H}_{n+1}$ given $\mathcal{F}_n$. We have thanks to relation (2.4.3) that for all $n \geq 0$, $\mathbb{E}[\Delta M_{n+1} \Delta M_{n+1}^t | \mathcal{F}_n] = L_n$. Consequently, we can split $\mathcal{H}_{n+1}$ into two terms

$$\mathcal{H}_{n+1} = \sum_{k=1}^{n} \frac{L_k}{|\mathbb{T}_k|} + \mathcal{K}_{n+1},$$

where

$$\mathcal{K}_{n+1} = \sum_{k=1}^{n} \frac{\Delta M_{k+1} \Delta M_{k+1}^t - L_k}{|\mathbb{T}_k|}.$$

It clearly follows from convergence (2.5.3) that

$$\lim_{n \to \infty} \frac{L_n}{|\mathbb{T}_n|} = \frac{1}{2} L \qquad \text{a.s.}$$

Hence, Cesaro convergence immediately implies that

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \frac{L_k}{|\mathbb{T}_k|} = \frac{1}{2} L \qquad \text{a.s.} \tag{2.6.16}$$

On the other hand, the sequence $(\mathcal{K}_n)_{n \geq 2}$ is obviously a square integrable martingale. Moreover, we have

$$\Delta \mathcal{K}_{n+1} = \mathcal{K}_{n+1} - \mathcal{K}_n = \frac{1}{|\mathbb{T}_n|} (\Delta M_{n+1} \Delta M_{n+1}^t - L_n).$$

For all $u \in \mathbb{R}^4$, denote $\mathcal{K}_n(u) = u^t \mathcal{K}_n u$. It follows from tedious but straightforward calculations, together with Lemma 2.5.2, that the increasing process of the martingale $(\mathcal{K}_n(u))_{n \geq 2}$ satisfies $\langle \mathcal{K}(u) \rangle_n = \mathcal{O}(n)$ a.s. Therefore, we deduce from the strong law of large numbers for martingales that for all $u \in \mathbb{R}^4$, $\mathcal{K}_n(u) = o(n)$ a.s. leading to $\mathcal{K}_n = o(n)$ a.s. Hence, we infer from (2.6.16) that

$$\lim_{n\to\infty} \frac{\mathcal{H}_{n+1}}{n} = \frac{1}{2}L \qquad \text{a.s.} \tag{2.6.17}$$

We obtain from (2.6.17) that

$$\lim_{n\to\infty} \frac{\mathcal{T}_n}{n} = \frac{1}{2}tr((I_2 \otimes A)^{-1/2} L (I_2 \otimes A)^{-1/2}) \qquad \text{a.s.}$$

which allows us to say that $\mathcal{R}_n = o(n)$ a.s. leading to (2.6.13). Then, Lemma 2.5.2 implies that

$$\lim_{n\to\infty} \frac{\Sigma_n}{|\mathbb{T}_n|} = I_2 \otimes A \qquad \text{a.s.} \tag{2.6.18}$$

where A is the positive definite matrix given by (2.5.6). We are now in position to prove (2.6.14). Let us recall that

$$\mathcal{B}_{n+1} = 2\sum_{k=1}^{n} M_k^t \Sigma_k^{-1} \Delta M_{k+1} = 2\sum_{k=1}^{n} M_k^t \Sigma_k^{-1} \Psi_k \xi_{k+1}.$$

Hence, $(\mathcal{B}_n)_{n\geq 2}$ is a square integrable martingale. In addition, we have

$$\Delta \mathcal{B}_{n+1} = 2M_n^t \Sigma_n^{-1} \Delta M_{n+1}.$$

Thus

$$\begin{aligned}
\mathbb{E}[(\Delta\mathcal{B}_{n+1})^2 | \mathcal{F}_n] &= 4\mathbb{E}[M_n^t \Sigma_n^{-1} \Delta M_{n+1} \Delta M_{n+1}^t \Sigma_n^{-1} M_n | \mathcal{F}_n] \qquad \text{a.s.}\\
&= 4M_n^t \Sigma_n^{-1} \mathbb{E}[\Delta M_{n+1} \Delta M_{n+1}^t | \mathcal{F}_n] \Sigma_n^{-1} M_n \qquad \text{a.s.}\\
&= 4M_n^t \Sigma_n^{-1} L_n \Sigma_n^{-1} M_n \qquad \text{a.s.}
\end{aligned}$$

We can observe that

$$L_n = \sum_{k\in\mathbb{G}_n} \frac{1}{c_k^2} \begin{pmatrix} \sigma_a^2 X_k + \sigma_c^2 & \rho \\ \rho & \sigma_b^2 X_k + \sigma_d^2 \end{pmatrix} \otimes \begin{pmatrix} X_k^2 & X_k \\ X_k & 1 \end{pmatrix}$$

and

$$\Psi_n \Psi_n^t = \sum_{k\in\mathbb{G}_n} \frac{1}{c_k} I_2 \otimes \begin{pmatrix} X_k^2 & X_k \\ X_k & 1 \end{pmatrix}.$$

For $\alpha = \max(\sigma_a^2 + \sigma_b^2, \sigma_c^2 + \sigma_d^2)$, denote

$$\Delta_n = \alpha c_n I_2 - \begin{pmatrix} \sigma_a^2 X_n + \sigma_c^2 & \rho \\ \rho & \sigma_b^2 X_n + \sigma_d^2 \end{pmatrix}.$$

Let us check that Δ_n is a positive definite matrix. As a matter of fact, we deduce from the elementary inequality

$$(\sigma_a^2 + \sigma_b^2) X_n + \sigma_c + \sigma_d^2 \leq \alpha c_n \tag{2.6.19}$$

that
$$tr(\Delta_n) = 2\alpha c_n - \big((\sigma_a^2+\sigma_b^2)X_n + \sigma_c^2+\sigma_d^2\big) \geq \alpha c_n > 0.$$
In addition, we also have from (2.6.19) that
$$\begin{aligned}\det(\Delta_n) &= \big(\alpha c_n - (\sigma_a^2 X_n + \sigma_c^2)\big)\big(\alpha c_n - (\sigma_b^2 X_n + \sigma_d^2)\big) - \rho^2,\\ &= \alpha^2 c_n^2 - \alpha c_n \big((\sigma_a^2+\sigma_b^2)X_n + \sigma_c^2+\sigma_d^2\big)\\ &\qquad\qquad + (\sigma_a^2 X_n + \sigma_c^2)(\sigma_b^2 X_n + \sigma_d^2) - \rho^2,\\ &\geq \sigma_a^2\sigma_b^2 X_n^2 + (\sigma_a^2\sigma_d^2 + \sigma_b^2\sigma_c^2)X_n + \sigma_c^2\sigma_d^2 - \rho^2,\\ &\geq \sigma_c^2\sigma_d^2 - \rho^2 > 0\end{aligned}$$
thanks to **(H.3)**. Consequently,
$$\begin{pmatrix}\sigma_a^2 X_n + \sigma_c^2 & \rho\\ \rho & \sigma_b^2 X_n + \sigma_d^2\end{pmatrix} \leq \alpha c_n I_2$$
which immediately implies that $L_n \leq \alpha \Psi_n \Psi_n^t$. Moreover, we can use Lemma B.1 of [11] to say that
$$\Sigma_n^{-1}\Psi_n\Psi_n^t\Sigma_n^{-1} \leq \Sigma_{n-1}^{-1} - \Sigma_n^{-1}.$$
Hence
$$\begin{aligned}\mathbb{E}[(\Delta\mathcal{B}_{n+1})^2|\mathcal{F}_n] &= 4M_n^t\Sigma_n^{-1}L_n\Sigma_n^{-1}M_n \qquad \text{a.s.}\\ &\leq 4\alpha M_n^t\Sigma_n^{-1}\Psi_n\Psi_n^t\Sigma_n^{-1}M_n \qquad \text{a.s.}\\ &\leq 4\alpha M_n^t(\Sigma_{n-1}^{-1} - \Sigma_n^{-1})M_n \qquad \text{a.s.}\end{aligned}$$

leading to $\langle\mathcal{B}\rangle_n \leq 4\alpha\mathcal{A}_n$. Therefore it follows from the strong law of large numbers for martingales that $\mathcal{B}_n = o(\mathcal{A}_n)$. Finally, we deduce from decomposition (2.6.12) that
$$\mathcal{V}_{n+1} + \mathcal{A}_n = o(\mathcal{A}_n) + \mathcal{O}(n) \qquad \text{a.s.}$$
leading to, since $\mathcal{A}_n$ and $\mathcal{V}_{n+1}$ are non negative, $\mathcal{A}_n = \mathcal{O}(n)$ and $\mathcal{V}_{n+1} = \mathcal{O}(n)$ a.s. which implies that $\mathcal{B}_n = o(n)$ a.s. Finally we clearly obtain convergence (2.6.15) from the main decomposition (2.6.12) together with (2.6.13) and (2.6.14), which completes the proof of Lemma 2.6.1. □

Lemma 2.6.2. *Assume that* (ε_n) *satisfies* **(H.1)** *to* **(H.5)**. *For all* $\delta > 1/2$, *we have*
$$\|M_n\|^2 = o(|\mathbb{T}_n|n^\delta) \qquad a.s. \tag{2.6.20}$$

Proof. Let us recall that
$$M_n = \sum_{k\in\mathbb{T}_{n-1}} \frac{1}{c_k}\begin{pmatrix}X_kV_{2k}\\ V_{2k}\\ X_kV_{2k+1}\\ V_{2k+1}\end{pmatrix}.$$

Denote
$$P_n = \sum_{k\in\mathbb{T}_{n-1}} \frac{X_k V_{2k}}{c_k} \qquad \text{and} \qquad Q_n = \sum_{i\in\mathbb{T}_{n-1}} \frac{V_{2k}}{c_k}.$$
On the one hand, P_n can be rewritten as
$$P_n = \sum_{k=1}^{n} \sqrt{|\mathbb{G}_{k-1}|} f_k \qquad \text{where} \qquad f_n = \frac{1}{\sqrt{|\mathbb{G}_{n-1}|}} \sum_{k\in\mathbb{G}_{n-1}} \frac{X_k V_{2k}}{c_k}.$$
We already saw in Section 2.3 that for all $k \in \mathbb{G}_n$,
$$\mathbb{E}[V_{2k}|\mathcal{F}_n] = 0 \qquad \text{and} \qquad \mathbb{E}[V_{2k}^2|\mathcal{F}_n] = \sigma_a^2 X_k + \sigma_c^2 \qquad \text{a.s.}$$
In addition, for all $k \in \mathbb{G}_n$,
$$\mathbb{E}[V_{2k}^4|\mathcal{F}_n] = 3\sigma_a^4 X_k^2 + X_k(\mu_a^4 - 3\sigma_a^4 + 6\sigma_a^2\sigma_c^2) + \mu_c^4 \qquad \text{a.s.}$$
which implies that
$$\mathbb{E}[V_{2k}^4|\mathcal{F}_n] \leq \mu_{ac}^4 c_k^2 \qquad \text{a.s..} \tag{2.6.21}$$
where $\mu_{ac}^4 = \mu_a^4 + \mu_c^4 + 6\sigma_a^2\sigma_c^2$. Consequently, $\mathbb{E}[f_{n+1}|\mathcal{F}_n] = 0$ a.s. and we deduce from (2.6.21) together with the Cauchy-Schwarz inequality that
$$\begin{aligned}
\mathbb{E}[f_{n+1}^4|\mathcal{F}_n] &= \frac{1}{|\mathbb{G}_n|^2} \sum_{k\in\mathbb{G}_n} \left(\frac{X_k}{c_k}\right)^4 \mathbb{E}[V_{2k}^4|\mathcal{F}_n] \\
&\quad + \frac{3}{|\mathbb{G}_n|^2} \sum_{k\in\mathbb{G}_n} \sum_{\substack{l\in\mathbb{G}_n \\ l\neq k}} \left(\frac{X_k}{c_k}\right)^2 \left(\frac{X_l}{c_l}\right)^2 \mathbb{E}[V_{2k}^2|\mathcal{F}_n]\mathbb{E}[V_{2l}^2|\mathcal{F}_n] \qquad \text{a.s.} \\
&\leq \frac{\mu_{ac}^4}{|\mathbb{G}_n|^2}(1 + 3\sqrt{|\mathbb{G}_n|(|\mathbb{G}_n|-1)}) \sum_{k\in\mathbb{G}_n} c_k^2 \qquad \text{a.s.} \\
&\leq \frac{3\mu_{ac}^4}{|\mathbb{G}_n|} \sum_{k\in\mathbb{G}_n} c_k^2 \qquad \text{a.s.}
\end{aligned} \tag{2.6.22}$$
However, it follows from Lemma 2.5.2 that
$$\lim_{n\to\infty} \frac{1}{|\mathbb{T}_n|} \sum_{k\in\mathbb{T}_n} c_k^2 = \mathbb{E}[(1+T)^2] \qquad \text{a.s.}$$
which is equivalent to say that
$$\lim_{n\to\infty} \frac{1}{|\mathbb{G}_n|} \sum_{k\in\mathbb{G}_n} c_k^2 = \mathbb{E}[(1+T)^2] \qquad \text{a.s.} \tag{2.6.23}$$
Therefore, we infer from (2.6.22) and (2.6.23) that
$$\sup_{n\geq 0} \mathbb{E}[f_{n+1}^4|\mathcal{F}_n] < \infty \qquad \text{a.s.}$$

Hence, we obtain from Wei's Lemma given in [57] (2.30) page 1673 that for all $\delta > 1/2$,

$$P_n^2 = o(|\mathbb{T}_{n-1}|n^\delta) \qquad \text{a.s.}$$

On the other hand, Q_n can be rewritten as

$$Q_n = \sum_{k=1}^{n} \sqrt{|\mathbb{G}_{k-1}|}g_k \qquad \text{where} \qquad g_n = \frac{1}{\sqrt{|\mathbb{G}_{n-1}|}} \sum_{k \in \mathbb{G}_{n-1}} \frac{V_{2k}}{c_k}.$$

Via the same calculation as before, $\mathbb{E}[g_{n+1}|\mathcal{F}_n] = 0$ a.s. and, as $c_n \geq 1$,

$$\mathbb{E}[g_{n+1}^4|\mathcal{F}_n] \leq \frac{3\mu_{bd}^4}{|\mathbb{G}_n|} \sum_{k \in \mathbb{G}_n} \frac{1}{c_k^2} \leq 3\mu_{bd}^4 \qquad \text{a.s.}$$

Hence, we deduce once again from Wei's Lemma that for all $\delta > 1/2$,

$$Q_n^2 = o(|\mathbb{T}_{n-1}|n^\delta) \qquad \text{a.s.}$$

In the same way, we obtain the same result for the two last components of M_n, which completes the proof of Lemma 2.6.2. □

2.6.5 Proof of Theorem 2.5.6 of the almost sure convergence results of $\widehat{\theta}_n$

We recall from (2.4.1) that $\widehat{\theta}_n - \theta = \Sigma_{n-1}^{-1} M_n$ which implies

$$\|\widehat{\theta}_n - \theta\|^2 \leq \frac{\mathcal{V}_n}{\lambda_{min}(\Sigma_{n-1})}$$

where $\mathcal{V}_n = M_n^t \Sigma_{n-1}^{-1} M_n$. On the one hand, a direct application of Lemma 2.6.2 ensures that $\mathcal{V}_n = o(n^\delta)$ a.s. for all $\delta > 1/2$. On the other hand, we deduce from (2.6.18) that

$$\lim_{n\to\infty} \frac{\lambda_{min}(\Sigma_n)}{|\mathbb{T}_n|} = \lambda_{min}(A) > 0 \qquad \text{a.s.}$$

Consequently, we find that, for all $\delta > 1/2$

$$\|\widehat{\theta}_n - \theta\|^2 = o\left(\frac{n^\delta}{|\mathbb{T}_{n-1}|}\right) \qquad \text{a.s.}$$

We are now in position to prove the quadratic strong law (2.5.5). We obtain from (2.6.15) that

$$\lim_{n\to\infty} \frac{\mathcal{A}_n}{n} = \frac{1}{2} tr((I_2 \otimes A)^{-1/2} L (I_2 \otimes A)^{-1/2}) \qquad \text{a.s.} \qquad (2.6.24)$$

Let us rewrite $\mathcal{A}_n$ as

$$\mathcal{A}_n = \sum_{k=1}^{n} M_k^t \left(\Sigma_{k-1}^{-1} - \Sigma_k^{-1}\right) M_k = \sum_{k=1}^{n} M_k^t \Sigma_{k-1}^{-1/2} \Delta_k \Sigma_{k-1}^{-1/2} M_k$$

where $\Delta_k = I_4 - \Sigma_{k-1}^{1/2} \Sigma_k^{-1} \Sigma_{k-1}^{1/2}$. We already saw from (2.6.18) that

$$\lim_{n\to\infty} \frac{\Sigma_n}{|\mathbb{T}_n|} = I_2 \otimes A \qquad \text{a.s.}$$

which ensures that

$$\lim_{n\to\infty} \Delta_n = \frac{1}{2} I_4 \qquad \text{a.s.}$$

In addition, we deduce from (2.6.24) that $\mathcal{A}_n = \mathcal{O}(n)$ a.s. which implies that

$$\frac{\mathcal{A}_n}{n} = \left(\frac{1}{2n} \sum_{k=1}^{n} M_k^t \Sigma_{k-1}^{-1} M_k \right) + o(1) \qquad \text{a.s.} \tag{2.6.25}$$

Moreover we have

$$\begin{aligned} \frac{1}{n} \sum_{k=1}^{n} M_k^t \Sigma_{k-1}^{-1} M_k &= \frac{1}{n} \sum_{k=1}^{n} (\widehat{\theta}_k - \theta)^t \Sigma_{k-1} (\widehat{\theta}_k - \theta), \\ &= \frac{1}{n} \sum_{k=1}^{n} |\mathbb{T}_{k-1}| (\widehat{\theta}_k - \theta)^t \frac{\Sigma_{k-1}}{|\mathbb{T}_{k-1}|} (\widehat{\theta}_k - \theta), \end{aligned}$$

Then, the inequality

$$(\widehat{\theta}_k - \theta)^t \left(\frac{\Sigma_{k-1}}{|\mathbb{T}_{k-1}|} - I_2 \otimes A \right) (\widehat{\theta}_k - \theta) \leq \lambda_{max} \left(\frac{\Sigma_{k-1}}{|\mathbb{T}_{k-1}|} - I_2 \otimes A \right) \|\widehat{\theta}_k - \theta\|^2$$

together with (2.5.5) and (2.6.18) allow us to say that

$$\frac{1}{n} \sum_{k=1}^{n} M_k^t \Sigma_{k-1}^{-1} M_k = \frac{1}{n} \sum_{k=1}^{n} |\mathbb{T}_{k-1}| (\widehat{\theta}_k - \theta)^t (I_2 \otimes A) (\widehat{\theta}_k - \theta) + o(1) \qquad \text{a.s.} \tag{2.6.26}$$

Therefore, (2.6.24) together with (2.6.25) and (2.6.26) lead to (2.5.5).

2.6.6 Proof of Theorem 2.5.7 of the almost sure convergence results of $\widehat{\eta}_n$, $\widehat{\zeta}_n$ and $\widehat{\rho}_n$

First of all, we shall only prove (2.5.7) since the proof of (2.5.8) follows exactly the same lines. We clearly have from (2.3.7) that

$$\begin{aligned} Q_{n-1}(\widehat{\eta}_n - \eta_n) &= \sum_{k\in\mathbb{T}_{n-1}} \frac{1}{d_k}(\widehat{V}_{2k}^2 - V_{2k}^2)\Phi_k, \\ &= \sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l} \frac{1}{d_k}(\widehat{V}_{2k}^2 - V_{2k}^2)\Phi_k, \\ &= \sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l} \frac{1}{d_k}\Big((\widehat{V}_{2k} - V_{2k})^2 + 2(\widehat{V}_{2k} - V_{2k})V_{2k}\Big)\Phi_k. \end{aligned} \tag{2.6.27}$$

In addition, we already saw in Section 2.3 that for all $l \geq 0$ and $k \in \mathbb{G}_l$,

$$\widehat{V}_{2k} - V_{2k} = -\begin{pmatrix}\widehat{a}_l - a\\ \widehat{c}_l - c\end{pmatrix}^t \Phi_k.$$

Consequently,

$$(\widehat{V}_{2k} - V_{2k})^2 \leq \|\Phi_k\|^2\left((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2\right).$$

Hence, we obtain that

$$\begin{aligned} \left\|\sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l}\frac{(\widehat{V}_{2k} - V_{2k})^2}{d_k}\Phi_k\right\| &\leq \sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l}\frac{\|\Phi_k\|^3}{d_k}\left((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2\right), \\ &\leq \sum_{l=0}^{n-1}\left((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2\right)\sum_{k\in\mathbb{G}_l} c_k, \\ &\leq \sum_{l=0}^{n-1}\left((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2\right)|\mathbb{T}_{l-1}|\frac{1}{|\mathbb{T}_{l-1}|}\sum_{k\in\mathbb{G}_l} c_k. \end{aligned} \tag{2.6.28}$$

Moreover, we can deduce from Lemma 2.5.2 that

$$\lim_{n\to\infty}\frac{1}{|\mathbb{T}_{n-1}|}\sum_{k\in\mathbb{G}_n} c_k = \mathbb{E}[1+T] \qquad \text{a.s.} \tag{2.6.29}$$

Then, we find from (2.6.28) and (2.6.29) that

$$\left\|\sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l}\frac{(\widehat{V}_{2k} - V_{2k})^2}{d_k}\Phi_k\right\| = \mathcal{O}\left(\sum_{l=0}^{n-1}|\mathbb{T}_{l-1}|\left((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2\right)\right) \qquad \text{a.s.}$$

However, as Λ is positive definite, we obtain from (2.5.5) that

$$\sum_{l=0}^{n-1} |\mathbb{T}_{l-1}| \left((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2\right) = \mathcal{O}(n) \qquad \text{a.s.}$$

which implies that

$$\left\| \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} \frac{(\widehat{V}_{2k} - V_{2k})^2}{d_k} \Phi_k \right\| = \mathcal{O}(n) \qquad \text{a.s.} \tag{2.6.30}$$

Furthermore, denote

$$P_n = \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} \frac{(\widehat{V}_{2k} - V_{2k}) V_{2k}}{d_k} \Phi_k.$$

We clearly have

$$\begin{aligned}
\Delta P_{n+1} = P_{n+1} - P_n &= \sum_{k \in \mathbb{G}_n} \frac{(\widehat{V}_{2k} - V_{2k}) V_{2k}}{d_k} \Phi_k, \\
&= - \sum_{k \in \mathbb{G}_n} \frac{V_{2k}}{d_k} \Phi_k \Phi_k^t \begin{pmatrix} \widehat{a}_l - a \\ \widehat{c}_l - c \end{pmatrix}.
\end{aligned}$$

In addition, for all $k \in \mathbb{G}_n$, $\mathbb{E}[V_{2k}|\mathcal{F}_n] = 0$ a.s. and $\mathbb{E}[V_{2k}^2|\mathcal{F}_n] = \sigma_a^2 X_k + \sigma_c^2 \leq \alpha c_k$ a.s. where $\alpha = \max(\sigma_a^2, \sigma_c^2)$. Consequently, $\mathbb{E}[\Delta P_{n+1}|\mathcal{F}_n] = 0$ a.s. and

$$\begin{aligned}
\mathbb{E}[\Delta P_{n+1} \Delta P_{n+1}^t|\mathcal{F}_n] &= \sum_{k \in \mathbb{G}_n} \frac{1}{d_k^2} \mathbb{E}[V_{2k}^2|\mathcal{F}_n] \Phi_k \Phi_k^t \begin{pmatrix} \widehat{a}_l - a \\ \widehat{c}_l - c \end{pmatrix} \begin{pmatrix} \widehat{a}_l - a \\ \widehat{c}_l - c \end{pmatrix}^t \Phi_k \Phi_k^t \qquad \text{a.s.} \\
&= \sum_{k \in \mathbb{G}_n} \frac{\sigma_a^2 X_k + \sigma_c^2}{d_k^2} \Phi_k \Phi_k^t \begin{pmatrix} \widehat{a}_l - a \\ \widehat{c}_l - c \end{pmatrix} \begin{pmatrix} \widehat{a}_l - a \\ \widehat{c}_l - c \end{pmatrix}^t \Phi_k \Phi_k^t \qquad \text{a.s.}
\end{aligned}$$

Therefore, (P_n) is a square integrable vector martingale with increasing process $\langle P \rangle_n$ given by

$$\begin{aligned}
\langle P \rangle_n &= \sum_{l=1}^{n-1} \mathbb{E}[\Delta P_{l+1} \Delta P_{l+1}^t|\mathcal{F}_l] \qquad \text{a.s.} \\
&= \sum_{l=1}^{n-1} \sum_{k \in \mathbb{G}_l} \frac{\sigma_a^2 X_k + \sigma_c^2}{d_k^2} \Phi_k \Phi_k^t \begin{pmatrix} \widehat{a}_l - a \\ \widehat{c}_l - c \end{pmatrix} \begin{pmatrix} \widehat{a}_l - a \\ \widehat{c}_l - c \end{pmatrix}^t \Phi_k \Phi_k^t \qquad \text{a.s.}
\end{aligned}$$

It immediately follows from the previous calculation that

$$\begin{aligned}
\|\langle P \rangle_n\| &\leq \alpha \sum_{l=0}^{n-1} \left((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2\right) \sum_{k \in \mathbb{G}_l} \frac{\|\Phi_k\|^4 c_k}{d_k^2} \qquad \text{a.s.} \\
&\leq \alpha \sum_{l=0}^{n-1} \left((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2\right) \sum_{k \in \mathbb{G}_l} c_k \qquad \text{a.s.}
\end{aligned}$$

leading to

$$\|\langle P\rangle_n\| = \mathcal{O}(n) \qquad \text{a.s.}$$

Then, we deduce from the strong law of large numbers for martingale given e.g. in Theorem 1.3.15 of [23] that

$$\|P_n\| = o(n) \qquad \text{a.s.} \tag{2.6.31}$$

Hence, we find from (2.6.27), (2.6.30) and (2.6.31) that

$$\|Q_{n-1}(\widehat{\eta}_n - \eta_n)\| = \mathcal{O}(n) \qquad \text{a.s.}$$

Moreover, we infer once again from Lemma 2.5.2 that

$$\lim_{n\to\infty} \frac{1}{|\mathbb{T}_n|} Q_n = \mathbb{E}\left[\begin{pmatrix} \frac{T^2}{(1+T)^2} & \frac{T}{(1+T)^2} \\ \frac{T}{(1+T)^2} & \frac{1}{(1+T)^2} \end{pmatrix}\right] \qquad \text{a.s.} \tag{2.6.32}$$

which ensures that

$$\|\widehat{\eta}_n - \eta_n\| = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \qquad \text{a.s.}$$

It remains to establish (2.5.9). Denote

$$\widehat{W}_n = \begin{pmatrix} \widehat{V}_{2n} \\ \widehat{V}_{2n+1} \end{pmatrix} \qquad \text{and} \qquad R_n = \sum_{k\in\mathbb{T}_{n-1}} \left(\widehat{W}_k - W_k\right)^t J W_k$$

where

$$J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then, we have

$$|\mathbb{T}_{n-1}|(\widehat{\rho}_n - \rho_n) = \sum_{k\in\mathbb{T}_{n-1}} \left(\widehat{V}_{2k} - V_{2k}\right)\left(\widehat{V}_{2k+1} - V_{2k+1}\right) + R_n.$$

It is not hard to see that (R_n) is a square integrable real martingale with increasing process given by

$$\begin{aligned}
\langle R\rangle_n &= \sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l} \mathbb{E}\left[(\widehat{W}_k - W_k)^t J W_k W_k^t J(\widehat{W}_k - W_k)\Big| \mathcal{F}_n\right] \qquad \text{a.s.} \\
&= \sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l} (\widehat{W}_k - W_k)^t J \mathbb{E}\left[W_k W_k^t\big| \mathcal{F}_n\right] J(\widehat{W}_k - W_k) \qquad \text{a.s.} \\
&= \sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l} (\widehat{W}_k - W_k)^t J \begin{pmatrix} \sigma_a^2 X_k + \sigma_c^2 & \rho \\ \rho & \sigma_b^2 X_k + \sigma_d^2 \end{pmatrix} J(\widehat{W}_k - W_k) \qquad \text{a.s.} \\
&= \sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l} (\widehat{W}_k - W_k)^t \begin{pmatrix} \sigma_b^2 X_k + \sigma_d^2 & \rho \\ \rho & \sigma_a^2 X_k + \sigma_c^2 \end{pmatrix} (\widehat{W}_k - W_k) \qquad \text{a.s.}
\end{aligned}$$

Consequently,

$$\begin{aligned}\langle R\rangle_n &\leq \sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l}\left((\sigma_a^2+\sigma_b^2)X_k+\sigma_c^2+\sigma_d^2\right)\|\widehat{W}_k-W_k\|^2 \qquad \text{a.s.}\\ &\leq 2\beta\sum_{l=0}^{n-1}\Big((\widehat{a}_l-a)^2+(\widehat{b}_l-b)^2\Big)\sum_{k\in\mathbb{G}_l}X_k^2c_k\\ &\qquad + 2\beta\sum_{l=0}^{n-1}\Big((\widehat{c}_l-c)^2+(\widehat{d}_l-d)^2\Big)\sum_{k\in\mathbb{G}_l}c_k \qquad \text{a.s.}\end{aligned}$$

where $\beta = \max(\sigma_a^2+\sigma_b^2, \sigma_c^2+\sigma_d^2)$. As previously, we obtain through Lemma 2.5.2 together with (2.5.5) that $\langle R\rangle_n = \mathcal{O}(n)$ a.s. which ensures that $R_n = o(n)$ a.s. Moreover,

$$\begin{aligned}\left|\sum_{k\in\mathbb{T}_{n-1}}\Big(\widehat{V}_{2k}-V_{2k}\Big)\Big(\widehat{V}_{2k+1}-V_{2k+1}\Big)\right| &\leq \frac{1}{2}\sum_{k\in\mathbb{T}_{n-1}}\left(\Big(\widehat{V}_{2k}-V_{2k}\Big)^2+\Big(\widehat{V}_{2k+1}-V_{2k+1}\Big)^2\right),\\ &\leq \frac{1}{2}\sum_{l=0}^{n-1}\|\widehat{\theta}_l-\theta\|^2\sum_{k\in\mathbb{G}_l}(1+X_k^2)\end{aligned}$$

which implies via Lemma 2.5.2 and (2.5.5) that

$$\sum_{k\in\mathbb{T}_{n-1}}\Big(\widehat{V}_{2k}-V_{2k}\Big)\Big(\widehat{V}_{2k+1}-V_{2k+1}\Big) = \mathcal{O}(n) \qquad \text{a.s.}$$

Therefore, we obtain that

$$|\mathbb{T}_{n-1}|(\widehat{\rho}_n-\rho_n) = \mathcal{O}(n) \qquad \text{a.s.}$$

which leads to (2.5.9). Finally, it only remains to prove the a.s. convergence of η_n, ζ_n and ρ_n to η, ζ and ρ which will immediately lead to the a.s. convergence of $\widehat{\eta}_n$, $\widehat{\zeta}_n$ and $\widehat{\rho}_n$ through (2.5.7), (2.5.8) and (2.5.9), respectively. On the one hand,

$$Q_{n-1}(\eta_n-\eta) = N_n = \sum_{k\in\mathbb{T}_n}\frac{1}{d_k}\Phi_k v_{2k} \tag{2.6.33}$$

where we recall that $v_{2n} = V_{2n}^2 - \eta^t\Phi_n$. It is clear, thanks to (2.3.8), that (N_n) is a square integrable vector martingale with increasing process $\langle N\rangle_n$ given by

$$\langle N\rangle_n = \sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l}\frac{1}{d_k^2}\Phi_k\Phi_k^t(2\sigma_a^4X_k^2+(\mu_a^4-3\sigma_a^4+4\sigma_a^2\sigma_c^2)X_k+\mu_c^4-\sigma_c^4) \qquad \text{a.s.}$$

Hence,

$$\langle N\rangle_n \leq \gamma\sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l}\frac{1}{d_k}\Phi_k\Phi_k^t \qquad \text{a.s.}$$

where $\gamma = \mu_a^4 - \sigma_a^4 + 4\sigma_a^2\sigma_c^2 + \mu_c^4 - \sigma_c^4$, which implies that

$$\|\langle N \rangle_n\| = \mathcal{O}(|\mathbb{T}_{n-1}|) \qquad \text{a.s.}$$

Consequently,

$$\|N_n\|^2 = \mathcal{O}(n|\mathbb{T}_{n-1}|) \qquad \text{a.s.}$$

which leads via (2.6.32) and (2.6.33) to the a.s. convergence of η_n to η and to the rate of convergence of Remark 2.5.8. The proof of the a.s. convergence of ζ_n to ζ follows exactly the same lines. On the other hand

$$|\mathbb{T}_{n-1}|(\rho_n - \rho) = H_n = \sum_{k \in \mathbb{T}_{n-1}} (V_{2k}V_{2k+1} - \rho) \tag{2.6.34}$$

It is obvious to see that (H_n) is a square integrable real martingale with increasing process $\langle H \rangle_n$ such that $\langle H \rangle_n = \mathcal{O}(|\mathbb{T}_{n-1}|)$ a.s. Finally, as $H_n^2 = \mathcal{O}(n|\mathbb{T}_{n-1}|)$ a.s., we deduce from (2.6.34) that ρ_n goes a.s. to ρ and that the rate of convergence of Remark 2.5.8 is verified, which completes the proof of Theorem 2.5.7.

2.6.7 Proof of the asymptotic normalities

In order to establish the asymptotic normality of our estimators, we will extensively make use of the central limit theorem for triangular arrays of vector martingales given e.g. by Theorem 2.1.9 of [23]. First of all, instead of using the generation-wise filtration $(\mathcal{F}_n)$, we will use the sister pair-wise filtration $(\mathcal{G}_n)$ given by

$$\mathcal{G}_n = \sigma(X_1, (X_{2k}, X_{2k+1}), 1 \le k \le n).$$

Proof of Theorem 2.5.9, first part. We focus our attention to the proof of the asymptotic normality (2.5.10). Let $M^{(n)} = (M_k^{(n)})$ be the square integrable vector martingale defined as

$$M_k^{(n)} = \frac{1}{\sqrt{|\mathbb{T}_n|}} \sum_{i=1}^{k} D_i \tag{2.6.35}$$

where

$$D_i = \frac{1}{c_i} \begin{pmatrix} X_i V_{2i} \\ V_{2i} \\ X_i V_{2i+1} \\ V_{2i+1} \end{pmatrix}.$$

We clearly have

$$M_{t_n}^{(n)} = \frac{1}{\sqrt{|\mathbb{T}_n|}} \sum_{i=1}^{t_n} D_i = \frac{1}{\sqrt{|\mathbb{T}_n|}} M_{n+1} \tag{2.6.36}$$

where $t_n = |\mathbb{T}_n|$. Moreover, the increasing process associated to $(M_k^{(n)})$ is given by

$$\langle M^{(n)} \rangle_k = \frac{1}{|\mathbb{T}_n|} \sum_{i=1}^{k} \mathbb{E}\left[D_i D_i^t | \mathcal{G}_{i-1}\right],$$

$$= \frac{1}{|\mathbb{T}_n|} \sum_{i=1}^{k} \frac{1}{c_i^2} \begin{pmatrix} \sigma_a^2 X_i + \sigma_c^2 & \rho \\ \rho & \sigma_b^2 X_i + \sigma_d^2 \end{pmatrix} \otimes \begin{pmatrix} X_i^2 & X_i \\ X_i & 1 \end{pmatrix} \qquad \text{a.s.}$$

Consequently, it follows from convergence (2.5.3) that

$$\lim_{n\to\infty} \langle M^{(n)} \rangle_{t_n} = L \qquad \text{a.s.}$$

It is now necessary to verify Lindeberg's condition by use of Lyapunov's condition. Denote

$$\phi_n = \sum_{k=1}^{t_n} \mathbb{E}\left[\|M_k^{(n)} - M_{k-1}^{(n)}\|^4 \Big| \mathcal{G}_{k-1}\right].$$

We obtain from (2.6.35) that

$$\phi_n = \frac{1}{|\mathbb{T}_n|^2} \sum_{k=1}^{t_n} \mathbb{E}\left[\frac{(1+X_k^2)^2}{c_k^4} (V_{2k}^2 + V_{2k+1}^2)^2 \Big| \mathcal{G}_{k-1}\right],$$

$$\leq \frac{2}{|\mathbb{T}_n|^2} \sum_{k=1}^{t_n} \left(\mathbb{E}[V_{2k}^4 | \mathcal{G}_{k-1}] + \mathbb{E}[V_{2k+1}^4 | \mathcal{G}_{k-1}] \right).$$

In addition, we already saw in Section 2.6.4 that

$$\mathbb{E}[V_{2n}^4 | \mathcal{G}_{n-1}] \leq \mu_{ac}^4 c_n^2, \qquad \mathbb{E}[V_{2n+1}^4 | \mathcal{G}_{n-1}] \leq \mu_{bd}^4 c_n^2 \qquad \text{a.s.}$$

where $\mu_{ac}^4 = \mu_a^4 + \mu_c^4 + 6\sigma_a^2\sigma_c^2$ and $\mu_{bd}^4 = \mu_b^4 + \mu_d^4 + 6\sigma_b^2\sigma_d^2$. Hence,

$$\phi_n \leq \frac{2\mu^4}{|\mathbb{T}_n|^2} \sum_{k=1}^{t_n} c_k^2 \qquad \text{a.s.}$$

where $\mu^4 = \mu_{ac}^4 + \mu_{bd}^4$. We can deduce from Lemma 2.5.2 that

$$\lim_{n\to\infty} \frac{1}{|\mathbb{T}_n|} \sum_{k\in\mathbb{T}_n} c_k^2 = \mathbb{E}[(1+T)^2] \qquad \text{a.s.}$$

which immediately implies that

$$\lim_{n\to\infty} \phi_n = 0 \qquad \text{a.s.}$$

Therefore, Lyapunov's condition is satisfied and Theorem 2.1.9 of [23] allows us to say via (2.6.36) that

$$\frac{1}{\sqrt{|\mathbb{T}_{n-1}|}} M_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, L).$$

Finally, we infer from (2.4.1) together with (2.6.18) and Slutsky's lemma that

$$\sqrt{|\mathbb{T}_{n-1}|}(\widehat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, (I_2 \otimes A^{-1})L(I_2 \otimes A^{-1})). \qquad \square$$

Proof of Theorem 2.5.9, second part. We shall now establish the asymptotic normality given by (2.5.11). Denote by $N^{(n)} = (N_k^{(n)})$ the square integrable vector martingale defined as

$$N_k^{(n)} = \frac{1}{\sqrt{|\mathbb{T}_n|}} \sum_{i=1}^{k} \frac{v_{2i}}{d_i} \Phi_i.$$

We immediately see from (2.6.33) that

$$N_{t_n}^{(n)} = \frac{1}{\sqrt{|\mathbb{T}_n|}} Q_n(\eta_{n+1} - \eta) = \frac{1}{\sqrt{|\mathbb{T}_n|}} N_{n+1}. \tag{2.6.37}$$

In addition, the increasing process associated to $(N_k^{(n)})$ is given by

$$\begin{aligned}
\langle N^{(n)} \rangle_k &= \frac{1}{|\mathbb{T}_n|} \sum_{i=1}^{k} \mathbb{E}\left[\frac{v_{2i}^2}{d_i^2} \Phi_i \Phi_i^t \middle| \mathcal{G}_{i-1} \right], \\
&= \frac{1}{|\mathbb{T}_n|} \sum_{i-1}^{k} \frac{1}{d_i^2} \Phi_i \Phi_i^t (2\sigma_a^4 X_i^2 + (\mu_a^4 - 3\sigma_a^4 + 4\sigma_a^2\sigma_c^2) X_i + \mu_c^4 - \sigma_c^4) \qquad \text{a.s.}
\end{aligned}$$

Consequently, we obtain from Lemma 2.5.2 that

$$\lim_{n\to\infty} \langle N^{(n)} \rangle_{t_n} = \mathbb{E}\left[\frac{2\sigma_a^4 T^2 + (\mu_a^4 - 3\sigma_a^4 + 4\sigma_a^2\sigma_c^2)T + (\mu_c^4 - \sigma_c^4)}{(1+T)^4} \begin{pmatrix} T^2 & T \\ T & 1 \end{pmatrix} \right] = M_{ac} \quad \text{a.s.}$$

In order to verify Lyapunov's condition, let

$$\phi_n = \sum_{k=1}^{t_n} \mathbb{E}\left[\|N_k^{(n)} - N_{k-1}^{(n)}\|^3 \middle| \mathcal{G}_{k-1} \right].$$

We clearly have

$$\|N_k^{(n)} - N_{k-1}^{(n)}\|^2 = \frac{1}{|\mathbb{T}_n|} \frac{(1 + X_k^2) v_{2k}^2}{d_k^2} \leq \frac{1}{|\mathbb{T}_n|} \frac{v_{2k}^2}{d_k},$$

which implies that

$$\|N_k^{(n)} - N_{k-1}^{(n)}\|^3 \leq \frac{1}{|\mathbb{T}_n|^{3/2}} \frac{|v_{2k}|^3}{d_k^{3/2}}.$$

However,

$$\begin{aligned}
|v_{2k}|^3 &= |V_{2k}^2 - \sigma_a^2 X_k - \sigma_c^2|^3 \leq (V_{2k}^2 + \sigma_a^2 X_k + \sigma_c^2)^3 \\
&\leq V_{2k}^6 + 3V_{2k}^4(\sigma_a^2 X_k + \sigma_c^2) + 3V_{2k}^2(\sigma_a^2 X_k + \sigma_c^2)^2 + (\sigma_a^2 X_k + \sigma_c^2)^3
\end{aligned} \tag{2.6.38}$$

We already saw that $\mathbb{E}[V_{2k}^2|\mathcal{G}_{k-1}] = \sigma_a^2 X_k + \sigma_c^2$ a.s. and it follows from (2.6.21) that

$$\mathbb{E}[V_{2k}^4|\mathcal{G}_{k-1}] \leq \mu_{ac} c_k^2 \qquad \text{a.s.}$$

It only remains to study $\mathbb{E}[V_{2k}^6|\mathcal{G}_{k-1}]$. Denote

$$A_k = \sum_{i=1}^{X_k} (Y_{k,i} - a) \qquad \text{and} \qquad B_k = \varepsilon_{2k} - c.$$

We clearly have from the identity $V_{2k} = A_k + B_k$ that

$$\begin{aligned} \mathbb{E}[V_{2k}^6|\mathcal{G}_{k-1}] &= \mathbb{E}[A_k^6|\mathcal{G}_{k-1}] + 15\mathbb{E}[A_k^4|\mathcal{G}_{k-1}]\mathbb{E}[B_k^2|\mathcal{G}_{k-1}] \\ &+ 20\mathbb{E}[A_k^3|\mathcal{G}_{k-1}]\mathbb{E}[B_k^3|\mathcal{G}_{k-1}] + \mathbb{E}[A_k^2|\mathcal{G}_{k-1}]\mathbb{E}[B_k^4|\mathcal{G}_{k-1}] + \mathbb{E}[B_k^6|\mathcal{G}_{k-1}]. \end{aligned} \tag{2.6.39}$$

On the one hand, $\mathbb{E}[A_k^2|\mathcal{G}_{k-1}] = \sigma_a^2 X_k$ a.s. and

$$\mathbb{E}[A_k^4|\mathcal{G}_{k-1}] = \mu_a^4 X_k + 3X_k(X_k - 1)\sigma_a^4 \qquad \text{a.s.}$$

Moreover, we have from Cauchy-Schwarz inequality that

$$\big|\mathbb{E}[A_k^3|\mathcal{G}_{k-1}]\big| \leq \mu_a^2 \sigma_a X_k \qquad \text{a.s.}$$

Furthermore, it follows from tedious but straightforward calculations that

$$\begin{aligned} \mathbb{E}[A_k^6|\mathcal{G}_{k-1}] \leq \tau_a^6 X_k + 15X_k(X_k-1)\mu_a^4\sigma_a^2 + 15\sigma_a^6 X_k(X_k-1)(X_k-2) \\ + 10\mu_a^6 X_k(X_k-1) \qquad \text{a.s.} \end{aligned}$$

Then, it exists some constant $\alpha > 0$ such that

$$\mathbb{E}[A_k^6|\mathcal{G}_{k-1}] \leq \alpha c_k^3 \qquad \text{a.s.}$$

On the other hand, $\mathbb{E}[B_k^2|\mathcal{G}_{k-1}] = \sigma_c^2$ a.s. and $\mathbb{E}[B_k^4|\mathcal{G}_{k-1}] = \mu_c^4$ a.s. In addition

$$\big|\mathbb{E}[B_k^3|\mathcal{G}_{k-1}]\big| \leq \mu_c^2 \sigma_c \qquad \text{and} \qquad \mathbb{E}[B_k^6|\mathcal{G}_{k-1}] \leq \tau_c^6 \qquad \text{a.s.}$$

Consequently, we deduce from (2.6.39) that it exists some constant $\beta > 0$ such that

$$\mathbb{E}[V_{2k}^6|\mathcal{G}_{k-1}] \leq \beta c_k^3 \qquad \text{a.s.}$$

which implies from (2.6.38) that for some constant $\gamma > 0$,

$$\mathbb{E}[|v_{2k}|^3|\mathcal{G}_{k-1}] \leq \gamma c_k^3 \qquad \text{a.s.}$$

Then, as $c_k^2 = d_k$, we can conclude that

$$\phi_n \leq \frac{\gamma}{\sqrt{|\mathbb{T}_n|}} \qquad \text{a.s.}$$

which immediately leads to

$$\lim_{n\to\infty} \phi_n = 0 \qquad \text{a.s.}$$

Therefore, Lyapunov's condition is satisfied and we find from Theorem 2.1.9 of [23] and (2.6.37) that

$$\frac{1}{\sqrt{|\mathbb{T}_{n-1}|}} N_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, M_{ac}). \tag{2.6.40}$$

Hence, we obtain from (2.6.32), (2.6.40) and Slutsky's lemma that

$$\sqrt{|\mathbb{T}_{n-1}|}(\eta_n - \eta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, B^{-1} M_{ac} B^{-1}).$$

Finally, (2.5.7) ensures that

$$\sqrt{|\mathbb{T}_{n-1}|}(\widehat{\eta}_n - \eta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, B^{-1} M_{ac} B^{-1}).$$

The proof of (2.5.12) follows exactly the same lines. □

Proof of Theorem 2.5.9, third part. It remains to establish the asymptotic normality given by (2.5.13). Denote by $H^{(n)} = (H_k^{(n)})$ the square integrable martingale defined as

$$H_k^{(n)} = \frac{1}{\sqrt{|\mathbb{T}_n|}} \sum_{i=1}^{k} (V_{2i} V_{2i+1} - \rho). \tag{2.6.41}$$

We clearly have from (2.6.34) that

$$H_{t_n}^{(n)} = \sqrt{|\mathbb{T}_n|}(\rho_{n+1} - \rho) = \frac{1}{\sqrt{|\mathbb{T}_n|}} H_{n+1}.$$

Moreover, the increasing process of $(H_k^{(n)})$ is given by

$$\langle H^{(n)} \rangle_k = \frac{1}{|\mathbb{T}_n|} \sum_{i=1}^{k} \left(\mathbb{E}[V_{2i}^2 V_{2i+1}^2 | \mathcal{G}_{n-1}] - \rho^2 \right).$$

As before, let

$$C_k = \sum_{i=1}^{X_k} (Z_{k,i} - b) \qquad \text{and} \qquad B_k = \varepsilon_{2k+1} - d.$$

As $V_{2k} = A_k + B_k$ and $V_{2k+1} = C_k + D_k$, we clearly have

$$\begin{aligned}\mathbb{E}\left[V_{2k}^2 V_{2k+1}^2 \middle| \mathcal{G}_{k-1}\right] = \mathbb{E}\left[A_k^2 \middle| \mathcal{G}_{k-1}\right] \left(\mathbb{E}\left[C_k^2 \middle| \mathcal{G}_{k-1}\right] + \mathbb{E}\left[D_k^2 \middle| \mathcal{G}_{k-1}\right]\right) \\ + \mathbb{E}\left[B_k^2 \middle| \mathcal{G}_{k-1}\right] \mathbb{E}\left[C_k^2 \middle| \mathcal{G}_{k-1}\right] + \mathbb{E}\left[B_k^2 D_k^2 \middle| \mathcal{G}_{k-1}\right] \qquad \text{a.s.}\end{aligned}$$

Consequently,

$$\mathbb{E}\left[V_{2k}^2 V_{2k+1}^2 \middle| \mathcal{G}_{k-1}\right] = \sigma_a^2\sigma_b^2 X_k^2 + \left(\sigma_a^2\sigma_d^2 + \sigma_b^2\sigma_c^2\right) X_k + \nu^2 \qquad \text{a.s.} \tag{2.6.42}$$

Then, we deduce once again from Lemma 2.5.2 that

$$\lim_{n\to\infty} \langle H^{(n)} \rangle_{t_n} = \sigma_\rho^2 \qquad \text{a.s.}$$

where σ_ρ^2 is given by (2.5.14). In order to verify Lyapunov's condition, denote

$$\phi_n = \sum_{k=1}^{t_n} \mathbb{E}\left[|H_k^{(n)} - H_{k-1}^{(n)}|^3 \middle| \mathcal{G}_{k-1}\right].$$

We obtain from (2.6.41) that

$$\begin{aligned} \phi_n &= \frac{1}{|\mathbb{T}_n|^{3/2}} \sum_{k=1}^{t_n} \mathbb{E}\left[|V_{2k}V_{2k+1} - \rho|^3 \middle| \mathcal{G}_{k-1}\right], \\ &\leq \frac{1}{|\mathbb{T}_n|^{3/2}} \sum_{k=1}^{t_n} \Big(\mathbb{E}\left[|V_{2k}|^3 |V_{2k+1}|^3 \middle| \mathcal{G}_{k-1}\right] + 3|\rho| \mathbb{E}\left[V_{2k}^2 V_{2k+1}^2 \middle| \mathcal{G}_{k-1}\right] \\ &\qquad\qquad + 3\rho^2 \mathbb{E}\left[|V_{2k}||V_{2k+1}| \middle| \mathcal{G}_{k-1}\right] + |\rho|^3 \Big). \end{aligned} \tag{2.6.43}$$

It follows from Cauchy-Schwarz inequality together with the previous calculations that it exists two constants $\alpha, \beta > 0$ such that

$$\mathbb{E}\left[|V_{2k}||V_{2k+1}| \middle| \mathcal{G}_{k-1}\right] \leq \alpha c_k \qquad \text{a.s.}$$

and

$$\mathbb{E}\left[|V_{2k}|^3 |V_{2k+1}|^3 \middle| \mathcal{G}_{k-1}\right] \leq \beta c_k^3 \qquad \text{a.s.}$$

In addition, we already saw from (2.6.42) that for some constant $\gamma > 0$

$$\mathbb{E}\left[V_{2k}^2 V_{2k+1}^2 \middle| \mathcal{G}_{k-1}\right] \leq \gamma c_k^2 \qquad \text{a.s.}$$

Consequently, we obtain from (2.6.43) that for some constant $\delta > 0$

$$\phi_n \leq \frac{\delta}{|\mathbb{T}_n|^{3/2}} \sum_{k=1}^{t_n} c_k^3 \qquad \text{a.s.}$$

which, via Lemma (2.5.2), leads to

$$\lim_{n\to\infty} \phi_n = 0 \qquad \text{a.s.}$$

Hence, we can conclude that

$$H_{t_n}^{(n)} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_\rho^2).$$

In other words

$$\sqrt{|\mathbb{T}_{n-1}|}(\rho_n - \rho) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_\rho^2).$$

Finally, we find via (2.5.9) that

$$\sqrt{|\mathbb{T}_{n-1}|}(\widehat{\rho}_n - \rho) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_\rho^2)$$

which achieves the proof of Theorem 2.5.9. □

Chapitre 3

Processus BAR à coefficients aléatoires

RÉSUMÉ. Le but de ce travail est d'étudier le comportement asymptotique des estimateurs des moindres carrés pondérés des paramètres inconnus des processus autorégressifs à bifurcation et à coefficients aléatoires. Sous de bonnes hypothèses sur l'immigration et l'héritage, nous établissons la convergence presque sure de nos estimateurs ainsi qu'une loi forte quadratique et des théorèmes centraux limites. Cette étude repose essentiellement sur des résultats asymptotiques pour les martingales vectorielles.

ABSTRACT. The purpose of this work is to study the asymptotic behavior of the weighted least squares estimators of the unknown parameters of random coefficient bifurcating autoregressive processes. Under suitable assumptions on the immigration and the inheritance, we establish the almost sure convergence of our estimators, as well as a quadratic strong law and central limit theorems. Our study mostly relies on asymptotic results for vector-valued martingales.

3.1 Introduction

In this paper, we will study random coefficient bifurcating autoregressive processes (RCBAR). Those processes are an adaptation of random coefficient autoregressive processes (RCAR) to binary tree structured data. We can also see those processes as the combination of RCAR processes and bifurcating autoregressive processes (BAR). RCAR processes have been first studied by Nicholls and Quinn [44, 46] while BAR processes have been first investigated by Cowan and Staudte [16]. Both inherited and environmental effects are taken into consideration in RCBAR processes in order to explain the evolution of the characteristic under study. The binary tree structure could lead us to take cell division as an example.

More precisely, the first-order RCBAR process is defined as follows. The initial cell is labelled 1 and the offspring of the cell labelled n are labelled $2n$ and $2n+1$. Denote by X_n the characteristic of individual n. Then, the first-order RCBAR process is given, for all $n \geq 1$, by

$$\begin{cases} X_{2n} & = a_n X_n + \varepsilon_{2n} \\ X_{2n+1} & = b_n X_n + \varepsilon_{2n+1} \end{cases}$$

The environmental effect is given by the driven noise sequence $(\varepsilon_{2n}, \varepsilon_{2n+1})_{n\geq 1}$ while the inherited effect is given by the random coefficient sequence $(a_n, b_n)_{n\geq 1}$. The cell division example leads us to consider that ε_{2n} and ε_{2n+1} are correlated since the environmental effect on two sister cells can reasonably be seen as correlated.

This study is inspired by experiments on the single celled organism *Escherichia coli*, see Stewart et al. [51] or Guyon et al. [27], which reproduces by dividing itself into two poles, one being called the new pole, the other being called the old pole. Experimental data seems to show that some variables among cell lines, such as the life span of the cells, does not evolve in the same way whether it is the new or the old pole. The difference in the evolution leads us to consider an asymmetric RCBAR. Considering a RCBAR process instead of a BAR process allows us to assume that the inherited effect is no more deterministic, as randomness often appears in nature. Moreover, we can consider both deterministic and random inherited effects since we also allow the random variables modeling the inherited effect to be deterministic, making this study usable for RCBAR as well as BAR.

This paper, which is an adaptation of [12] to RCBAR processes, intends to study the asymptotic behavior of the weighted least squares (WLS) estimators of first-order RCBAR processes using a martingale approach. This martingale approach has been first proposed by Bercu et al. [11] and de Saporta et al. [17] for BAR processes. The WLS estimation of parameters branching processes was previously investigated by Wei and Winnicki [58] and Winnicki [59]. We will make use several times of the strong law of large numbers [23] as well as the central limit theorem [23, 28] for martingales, in order to investigate the asymptotic behavior of the WLS estimators. Those theorems have been previously used by Basawa and Zhou [9, 60, 61]. An other study of the parameters of a RCBAR process has been made by de Saporta et al. [19] in which they consider non-weighted estimators, which highly increase the order of the moment assumptions, in the context of a RCBAR with missing data.

Several approaches appeared for BAR processes, and we tried not to set aside any of them. Thus, we took into account the classical BAR studies as seen in Huggins and Basawa [31, 32] and Huggins and Staudte [33] who studied the evolution of cell diameters and lifetimes, and also the bifurcating Markov chain model introduced by Guyon [26] and used in Delmas and Marsalle [20]. Still, we did not forget to have a look to the analogy with the Galton-Watson processes as studied in Delmas and Marsalle [20] and Heyde and Seneta [29]. Several methods have also been used for parameter estimation in RCAR processes. Koul and Schick [39] used an M-estimator while Aue et al. [3] preferred a quasi-maximum likelihood approach. Schick [49] introduced a new class of estimator that Vanecek [56] used in his work. Hwang et al. [35] also tackled the critical case where the environmental effect follows a Rademacher distribution.

The paper is organized as follows. Section 2 allows us to explain more precisely the model in which we are interested in, then Section 3 formulates the WLS estimators of the unknown parameters we will study. Section 4 permits us to introduce the martingale point of view of this paper. The main results are collected in Section 5, those results concern the asymptotic behavior of our WLS estimators, to be more accurate, we will establish the almost sure convergence, the quadratic strong law and the asymptotic normality of our estimators. Finally, the other sections gathers the proofs of our main results, except the last section which illustrates our results with a small simulation study.

3.2 Random coefficient bifurcating autoregressive processes

Consider the first-order RCBAR process given, for all $n \geq 1$, by

$$\begin{cases} X_{2n} &= a_n X_n + \varepsilon_{2n} \\ X_{2n+1} &= b_n X_n + \varepsilon_{2n+1} \end{cases} \tag{3.2.1}$$

where the initial state X_1 is the ancestor of the process and $(\varepsilon_{2n}, \varepsilon_{2n+1})$ stands for the driven noise of the process. In all the sequel, we shall assume that $\mathbb{E}[X_1^2] < \infty$. We also assume that both $(a_n, b_n)_{n\geq 1}$ and $(\varepsilon_{2n}, \varepsilon_{2n+1})_{n\geq 1}$ are i.i.d., and that those two sequences are independent. One can see the RCBAR process given by (3.2.1) as a first-order random coefficient autoregressive process on a binary tree, where each node represents an individual, node 1 being the original ancestor. For all $n \geq 1$, denote the n-th generation by $\mathbb{G}_n = \{2^n, 2^n+1, \ldots, 2^{n+1}-1\}$. In particular, $\mathbb{G}_0 = \{1\}$ is the initial generation and $\mathbb{G}_1 = \{2,3\}$ is the first generation of offspring from the first ancestor. Recall that the two offspring of individual n are labelled $2n$ and $2n+1$, or conversely, the mother of individual n is $[n/2]$ where $[x]$ stands for the largest integer less than or equal to x. Finally denote by

$$\mathbb{T}_n = \bigcup_{k=0}^{n} \mathbb{G}_n$$

the sub-tree of all individuals from the original individual up to the n-th generation. On can observe that the cardinality $|\mathbb{G}_n|$ of $\mathbb{G}_n$ is 2^n while that of $\mathbb{T}_n$ is $|\mathbb{T}_n| = 2^{n+1}-1$.

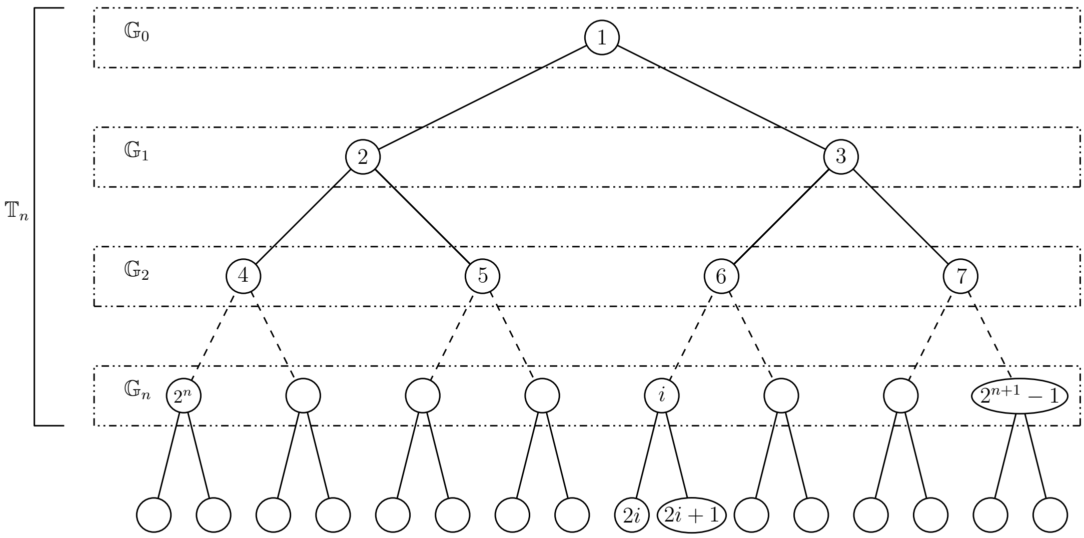

FIGURE 3.1 – The tree associated with the RCBAR

3.3 Weighted least-squares estimation

Denote by $\mathbb{F} = (\mathcal{F}_n)_{n\geq 0}$ the natural filtration associated with the first-order RCBAR process, which means that $\mathcal{F}_n$ is the σ-algebra generated by all individuals up to the n-th generation, in other words $\mathcal{F}_n = \sigma\{X_k, k \in \mathbb{T}_n\}$. We will assume in all the sequel that, for all $n \geq 0$ and for all $k \in \mathbb{G}_n$,

$$\begin{cases} \mathbb{E}[a_k|\mathcal{F}_n] = a \quad \text{a.s.} \\ \mathbb{E}[b_k|\mathcal{F}_n] = b \quad \text{a.s.} \\ \mathbb{E}[\varepsilon_{2k}|\mathcal{F}_n] = c \quad \text{a.s.} \\ \mathbb{E}[\varepsilon_{2k+1}|\mathcal{F}_n] = d \quad \text{a.s.} \end{cases} \tag{3.3.1}$$

Consequently, we deduce from (3.2.1) and (3.3.1) that, for all $n \geq 0$ and for all $k \in \mathbb{G}_n$,

$$\begin{cases} X_{2k} &= aX_k + c + V_{2k}, \\ X_{2k+1} &= bX_k + d + V_{2k+1}, \end{cases} \tag{3.3.2}$$

where, $V_{2k} = X_{2k} - \mathbb{E}[X_{2k}|\mathcal{F}_n]$ and $V_{2k+1} = X_{2k+1} - \mathbb{E}[X_{2k+1}|\mathcal{F}_n]$. Therefore, the two relations given by (3.3.2) can be rewritten in a classic autoregressive form

$$\chi_n = \theta^t \Phi_n + W_n \tag{3.3.3}$$

where

$$\chi_n = \begin{pmatrix} X_{2n} \\ X_{2n+1} \end{pmatrix}, \qquad \Phi_n = \begin{pmatrix} X_n \\ 1 \end{pmatrix}, \qquad W_n = \begin{pmatrix} V_{2n} \\ V_{2n+1} \end{pmatrix},$$

and the matrix parameter

$$\theta = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Our goal is to estimate θ from the observation of all individuals up to $\mathbb{T}_n$. We propose to make use of the WLS estimator $\widehat{\theta}_n$ of θ which minimizes

$$\Delta_n(\theta) = \frac{1}{2} \sum_{k \in \mathbb{T}_{n-1}} \frac{1}{c_k} \|\chi_k - \theta^t \Phi_k\|^2$$

where the choice of the weighting sequence $(c_n)_{n\geq 1}$ is crucial. We shall choose $c_n = 1 + X_n^2$ and we will go back to this suitable choice in Section 3.4. Consequently, we obviously have for all $n \geq 1$

$$\widehat{\theta}_n = S_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \frac{1}{c_k} \Phi_k \chi_k^t, \qquad \text{where} \qquad S_n = \sum_{k \in \mathbb{T}_n} \frac{1}{c_k} \Phi_k \Phi_k^t. \tag{3.3.4}$$

In order to avoid useless invertibility assumption, we shall assume, without loss of generality, that for all $n \geq 0$, S_n is invertible. Otherwise, we only have to add the identity matrix of order 2, I_2 to S_n. Since, in a certain way, S_n goes to infinty , it will not change our results. In all what follows, we shall make a slight abuse of notation by identifying θ as well as $\widehat{\theta}_n$ to

$$\text{vec}(\theta) = \begin{pmatrix} a \\ c \\ b \\ d \end{pmatrix} \qquad \text{and} \qquad \text{vec}(\widehat{\theta}_n) = \begin{pmatrix} \widehat{a}_n \\ \widehat{c}_n \\ \widehat{b}_n \\ \widehat{d}_n \end{pmatrix}.$$

Therefore, we deduce from (3.3.4) that

$$\widehat{\theta}_n = \Sigma_{n-1}^{-1} \sum_{k \in \mathbb{T}_n - 1} \frac{1}{c_k} \text{vec}(\Phi_k \chi_k^t) = \Sigma_{n-1}^{-1} \sum_{k \in \mathbb{T}_n - 1} \frac{1}{c_k} \begin{pmatrix} X_k X_{2k} \\ X_{2k} \\ X_k X_{2k+1} \\ X_{2k+1} \end{pmatrix}$$

where $\Sigma_n = I_2 \otimes S_n$ and $\otimes$ stands for the standard Kronecker product. Consequently, (3.3.3) yields to

$$\widehat{\theta}_n - \theta = \Sigma_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \frac{1}{c_k} \mathrm{vec}(\Phi_k W_k^t),$$

$$= \Sigma_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \frac{1}{c_k} \begin{pmatrix} X_k V_{2k} \\ V_{2k} \\ X_k V_{2k+1} \\ V_{2k+1} \end{pmatrix}. \tag{3.3.5}$$

In all the sequel, we shall make use of the following moment hypotheses.

(H.1) For all $k \geq 1$,

$$\mathbb{E}[a_k^2] < 1 \quad \text{and} \quad \mathbb{E}[b_k^2] < 1.$$

(H.2) For all $n \geq 0$ and for all $k \in \mathbb{G}_n$

$$\mathrm{Var}[a_k|\mathcal{F}_n] = \sigma_a^2 \geq 0 \quad \text{and} \quad \mathrm{Var}[b_k|\mathcal{F}_n] = \sigma_b^2 \geq 0 \qquad \text{a.s.}$$

$$\mathrm{Var}[\varepsilon_{2k}|\mathcal{F}_n] = \sigma_c^2 > 0 \quad \text{and} \quad \mathrm{Var}[\varepsilon_{2k+1}|\mathcal{F}_n] = \sigma_d^2 > 0 \qquad \text{a.s.}$$

(H.3) For all $n \geq 0$ and for all $k, l \in \mathbb{G}_{n+1}$, if $[k/2] \neq [l/2]$, ε_k and ε_l are conditionally independent given $\mathcal{F}_n$ and for all $k, l \in \mathbb{G}_n$, if $k \neq l$, (a_k, b_k) and (a_l, b_l) are conditionally independent given $\mathcal{F}_n$. While otherwise, it exists $\rho_{cd}^2 < \sigma_c^2 \sigma_d^2$ and $\rho_{ab}^2 \leq \sigma_a^2 \sigma_b^2$ such that, for all $k \in \mathbb{G}_n$

$$\mathbb{E}[(\varepsilon_{2k} - c)(\varepsilon_{2k+1} - d)|\mathcal{F}_n] = \rho_{cd} \qquad \text{a.s.}$$

$$\mathbb{E}[(a_k - a)(b_k - b)|\mathcal{F}_n] = \rho_{ab} \qquad \text{a.s.}$$

(H.4) One can find $\mu_a^4 \geq \sigma_a^4$, $\mu_b^4 \geq \sigma_b^4$, $\mu_c^4 > \sigma_c^4$ and $\mu_d^4 > \sigma_d^4$ such that, for all $n \geq 0$ and for all $k \in \mathbb{G}_n$

$$\mathbb{E}\left[(a_k - a)^4 \,|\mathcal{F}_n\right] = \mu_a^4 \quad \text{and} \quad \mathbb{E}\left[(b_k - b)^4 \,|\mathcal{F}_n\right] = \mu_c^4 \qquad \text{a.s.}$$

$$\mathbb{E}\left[(\varepsilon_{2k} - c)^4 \,|\mathcal{F}_n\right] = \mu_c^4 \quad \text{and} \quad \mathbb{E}\left[(\varepsilon_{2k+1} - d)^4 \,|\mathcal{F}_n\right] = \mu_d^4 \qquad \text{a.s.}$$

$$\mathbb{E}[\varepsilon_{2k}^4] > \mathbb{E}[\varepsilon_{2k}^2]^2 \quad \text{and} \quad \mathbb{E}[\varepsilon_{2k+1}^4] > \mathbb{E}[\varepsilon_{2k+1}^2]^2.$$

In addition, it exists $\nu_{ab}^2 \geq \rho_{ac}^2$ and $\nu_{cd}^2 > \rho_{cd}^2$ such that, for all $k \in \mathbb{G}_n$

$$\mathbb{E}[(a_k - a)^2 (b_k - b)^2 |\mathcal{F}_n] = \nu_{ab}^2 \quad \text{and} \quad \mathbb{E}[(\varepsilon_{2k} - c)^2 (\varepsilon_{2k+1} - d)^2 |\mathcal{F}_n] = \nu_{cd}^2 \qquad \text{a.s.}$$

(H.5) It exists $\alpha > 4$ such that

$$\sup_{n \geq 0} \sup_{k \in \mathbb{G}_n} \mathbb{E}[|a_k - a|^\alpha |\mathcal{F}_n] < \infty, \qquad \sup_{n \geq 0} \sup_{k \in \mathbb{G}_n} \mathbb{E}[|b_k - b|^\alpha |\mathcal{F}_n] < \infty \qquad \text{a.s.}$$

$$\sup_{n \geq 0} \sup_{k \in \mathbb{G}_n} \mathbb{E}[|\varepsilon_{2k} - c|^\alpha |\mathcal{F}_n] < \infty, \qquad \sup_{n \geq 0} \sup_{k \in \mathbb{G}_n} \mathbb{E}[|\varepsilon_{2k+1} - d|^\alpha |\mathcal{F}_n] < \infty \qquad \text{a.s.}$$

One can observe that those hypotheses allows us to consider the deterministic case where it exists some constants a, b with $\max(|a|,|b|) < 1$ such that, for all $k \geq 1$, $a_k = a$ and $b_k = b$ a.s. Moreover, under assumption **(H.2)**, we have for all $n \geq 0$ and for all $k \in \mathbb{G}_n$

$$\mathbb{E}[V_{2k}^2|\mathcal{F}_n] = \sigma_a^2 X_k^2 + \sigma_c^2 \quad \text{and} \quad \mathbb{E}[V_{2k+1}^2|\mathcal{F}_n] = \sigma_b^2 X_k^2 + \sigma_d^2 \qquad \text{a.s.} \quad (3.3.6)$$

Consequently, if we choose $c_n = 1 + X_n^2$ for all $n \geq 1$, we clearly have for all $k \in \mathbb{G}_n$

$$\mathbb{E}\left[V_{2k}^2 \middle| \mathcal{F}_n\right] \leq \max(\sigma_a^2, \sigma_c^2) c_k \quad \text{and} \quad \mathbb{E}\left[V_{2k+1}^2 \middle| \mathcal{F}_n\right] \leq \max(\sigma_b^2, \sigma_d^2) c_k \quad \text{a.s.}$$

It is exactly the reason why we have chosen this weighting sequence into (3.3.4). Similar WLS estimation approach for branching processes with immigration may be found in [58] and [59]. For all $n \geq 0$ and for all $k \in \mathbb{G}_n$, denote $v_{2k} = V_{2k}^2 - \mathbb{E}[V_{2k}^2|\mathcal{F}_n]$. We deduce from (3.3.6) that for all $n \geq 1$, $V_{2n}^2 = \eta^t \psi_n + v_{2n}$ where η is defined by

$$\eta = \begin{pmatrix} \sigma_a^2 \\ \sigma_c^2 \end{pmatrix} \quad \text{and} \quad \psi_n = \begin{pmatrix} X_n^2 \\ 1 \end{pmatrix}.$$

It leads us to estimate the vector of variances η by the WLS estimator

$$\widehat{\eta}_n = Q_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \frac{1}{d_k} \widehat{V}_{2k}^2 \psi_k, \qquad \text{where} \qquad Q_n = \sum_{k \in \mathbb{T}_n} \frac{1}{d_k} \psi_k \psi_k^t \qquad (3.3.7)$$

and for all $k \in \mathbb{G}_n$,

$$\begin{cases} \widehat{V}_{2k} & = X_{2k} - \widehat{a}_n X_k - \widehat{c}_n, \\ \widehat{V}_{2k+1} & = X_{2k+1} - \widehat{b}_n X_k - \widehat{d}_n. \end{cases}$$

Finally the weighting sequence $(d_n)_{n \geq 1}$ is given, for all $n \geq 1$, by $d_n = c_n^2 = (1+X_n^2)^2$. This choice is due to the fact that for all $n \geq 1$ and for all $k \in \mathbb{G}_n$

$$\begin{aligned} \mathbb{E}[v_{2k}^2|\mathcal{F}_n] &= \mathbb{E}[V_{2k}^4|\mathcal{F}_n] - \left(\mathbb{E}[V_{2k}^2|\mathcal{F}_n]\right)^2 \qquad \text{a.s.} \\ &= (\mu_a^4 - \sigma_a^4) X_k^4 + 4\sigma_a^2 \sigma_c^2 X_k^2 + (\mu_c^4 - \sigma_c^4) \qquad \text{a.s.} \end{aligned}$$

Consequently, as $d_n \geq 1$, we clearly have for all $n \geq 1$ and for all $k \in \mathbb{G}_n$

$$\mathbb{E}[v_{2k}^2|\mathcal{F}_n] \leq \max(\mu_a^4 - \sigma_a^4, 2\sigma_a^2\sigma_c^2, \mu_c^4 - \sigma_c^4) d_k \qquad \text{a.s.}$$

We have a similar WLS estimator $\widehat{\zeta}_n$ of the vector of variances

$$\zeta^t = \begin{pmatrix} \sigma_b^2 & \sigma_d^2 \end{pmatrix}$$

by replacing $\widehat{V}_{2k}^2$ by $\widehat{V}_{2k+1}^2$ into (3.3.7). Let us remark that, for all $n \geq 0$ and for all $k \in \mathbb{G}_n$,

$$\mathbb{E}[V_{2k} V_{2k+1}|\mathcal{F}_n] = \rho_{ab} X_n^2 + \rho_{cd}. \qquad (3.3.8)$$

Then, for all $n \geq 0$ and for all $k \in \mathbb{G}_n$, denote $w_{2k} = V_{2k}V_{2k+1} - \mathbb{E}[V_{2k}V_{2k+1}|\mathcal{F}_n]$. We deduce from (3.3.8) that for all $k \geq 1$, $V_{2k}V_{2k+1} = \nu^t \psi_k + w_{2k}$ where ν is defined by

$$\nu = \begin{pmatrix} \rho_{ab} \\ \rho_{cd} \end{pmatrix}.$$

It leads us to estimate the vector of covariances ν by the WLS estimator

$$\widehat{\nu}_n = Q_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \frac{1}{d_k} \widehat{V}_{2k} \widehat{V}_{2k+1} \psi_k. \tag{3.3.9}$$

This choice is due to the fact that for all $n \geq 1$ and for all $k \in \mathbb{G}_n$

$$\mathbb{E}[V_{2k}^2 V_{2k+1}^2 |\mathcal{F}_n] = \nu_{ab}^2 X_k^4 + (\sigma_a^2 \sigma_d^2 + 4\rho_{ab}\rho_{cd} + \sigma_b^2 \sigma_c^2) X_k^2 + \nu_{cd}^2 \qquad \text{a.s.}$$

Consequently, as $d_n \geq 1$, we clearly have for all $n \geq 1$ and for all $k \in \mathbb{G}_n$

$$\begin{aligned} \mathbb{E}[w_{2k}^2|\mathcal{F}_n] &= (\nu_{ab}^2 - \rho_{ab}^2) X_k^4 + \left(\sigma_a^2\sigma_d^2 + \sigma_b^2\sigma_c^2 + 2\rho_{ab}\rho_{cd}\right) X_k^2 + (\nu_{cd}^2 - \rho_{cd}^2) \qquad \text{a.s.} \\ &\leq \max\left(\nu_{ab}^2, \nu_{cd}^2, \left(\sigma_a^2 + \sigma_c^2\right)\left(\sigma_b^2 + \sigma_d^2\right)\right) d_k \qquad \text{a.s.} \end{aligned}$$

3.4 A martingale approach

In order to establish all the asymptotic properties of our estimators, we shall make use of a martingale approach. For all $n \geq 1$, denote

$$M_n = \sum_{k \in \mathbb{T}_{n-1}} \frac{1}{c_k} \begin{pmatrix} X_k V_{2k} \\ V_{2k} \\ X_k V_{2k+1} \\ V_{2k+1} \end{pmatrix}.$$

We can clearly rewrite (3.3.5) as

$$\widehat{\theta}_n - \theta = \Sigma_{n-1}^{-1} M_n. \tag{3.4.1}$$

As in [11], we make use of the notation M_n since it appears that $(M_n)_{n\geq 1}$ is a martingale. This fact is a crucial point of our study and it justifies the vector notation since most of all asymptotic results for martingales were established for vector-valued martingales. Let us rewrite M_n in order to emphasize its martingale quality. Let $\Psi_n = I_2 \otimes \varphi_n$ where φ_n is the matrix of dimension 2×2^n given by

$$\varphi_n = \begin{pmatrix} \frac{X_{2^n}}{\sqrt{c_{2^n}}} & \frac{X_{2^n+1}}{\sqrt{c_{2^n+1}}} & \cdots & \frac{X_{2^{n+1}-1}}{\sqrt{c_{2^{n+1}-1}}} \\ \frac{1}{\sqrt{c_{2^n}}} & \frac{1}{\sqrt{c_{2^n+1}}} & \cdots & \frac{1}{\sqrt{c_{2^{n+1}-1}}} \end{pmatrix}.$$

It represents the individuals of the n-th generation which is also the collection of all $\Phi_k/\sqrt{c_k}$ where k belongs to $\mathbb{G}_n$. Let ξ_n be the random vector of dimension 2^n

$$\xi_n^t = \left(\frac{V_{2^n}}{\sqrt{c_{2^{n-1}}}} \quad \frac{V_{2^n+2}}{\sqrt{c_{2^{n-1}+1}}} \quad \cdots \quad \frac{V_{2^{n+1}-2}}{\sqrt{c_{2^n-1}}} \quad \frac{V_{2^n+1}}{\sqrt{c_{2^{n-1}}}} \quad \frac{V_{2^n+3}}{\sqrt{c_{2^{n-1}+1}}} \quad \cdots \quad \frac{V_{2^{n+1}-1}}{\sqrt{c_{2^n-1}}} \right).$$

The vector ξ_n gathers the noise variables of $\mathbb{G}_n$. The special ordering separating odd and even indices has been made in [11] so that M_n can be written as

$$M_n = \sum_{k=1}^{n} \Psi_{k-1}\xi_k.$$

Under (3.3.1), we clearly have for all $n \geq 0$, $\mathbb{E}[\xi_{n+1}|\mathcal{F}_n] = 0$ a.s. and Ψ_n is $\mathcal{F}_n$-measurable. In addition it is not hard to see that under **(H.1)** to **(H.2)**, (M_n) is a locally square integrable vector martingale with increasing process given, for all $n \geq 1$, by

$$\langle M \rangle_n = \sum_{k=0}^{n-1} \Psi_k \mathbb{E}[\xi_{k+1}\xi_{k+1}^t|\mathcal{F}_k]\Psi_k^t = \sum_{k=0}^{n-1} L_k \qquad \text{a.s.} \tag{3.4.2}$$

where

$$L_k = \sum_{i \in \mathbb{G}_k} \frac{1}{c_i^2} \begin{pmatrix} P(X_i) & Q(X_i) \\ Q(X_i) & R(X_i) \end{pmatrix} \otimes \begin{pmatrix} X_i^2 & X_i \\ X_i & 1 \end{pmatrix}. \tag{3.4.3}$$

with

$$\begin{cases} P(X) = \sigma_a^2 X^2 + \sigma_c^2, \\ Q(X) = \rho_{ab} X^2 + \rho_{cd}, \\ R(X) = \sigma_b^2 X^2 + \sigma_d^2. \end{cases}$$

One can remark that we obviously have $\langle M \rangle_n = \mathcal{O}(\mathbb{T}_n)$ but it is necessary to establish the convergence of $\langle M \rangle_n$, properly normalized, in order to prove the asymptotic results for our RCBAR estimators $\widehat{\theta}_n$, $\widehat{\eta}_n$, $\widehat{\zeta}_n$ and $\widehat{\nu}_n$.

3.5 Main results

We have to introduce some more notations in order to state our main results. From the original process $(X_n)_{n\geq 1}$, we shall define a new process $(Y_n)_{n\geq 1}$ recursively defined by $Y_1 = X_1$, and if $Y_n = X_k$ with $n, k \geq 1$, then

$$Y_{n+1} = X_{2k+\kappa_n}$$

where $(\kappa_n)_{n\geq 1}$ is a sequence of i.i.d. random variables with Bernoulli $\mathcal{B}(1/2)$ distribution. Such a construction may be found in [26] for the asymptotic analysis of BAR processes. The process (Y_n) gathers the values of the original process (X_n)

along the random branch of the binary tree $(\mathbb{T}_n)$ given by (κ_n). Denote by k_n the unique $k \geq 1$ such that $Y_n = X_k$. Then, for all $n \geq 1$, we have

$$Y_{n+1} = \widetilde{a}_{n+1} Y_n + e_{n+1} \tag{3.5.1}$$

where, with k_n the unique number k such that $Y_n = X_k$,

$$\widetilde{a}_{n+1} = \begin{cases} a_{k_n} & \text{if } \kappa_n = 0, \\ b_{k_n} & \text{otherwise,} \end{cases} \qquad \text{and} \qquad e_n = \varepsilon_{k_n}. \tag{3.5.2}$$

Lemma 3.5.1. *Assume that* **(H.1)** *and* **(H.2)** *are satisfied. Then, we have*

$$Y_n \xrightarrow{\mathcal{L}} T$$

where T is a positive non degenerate random variable with $\mathbb{E}[T^2] < \infty$.

Denote $\mathcal{C}_b^1(\mathbb{R}_+) = \Big\{ f \in \mathcal{C}^1(\mathbb{R}, \mathbb{R}) \Big| \exists \gamma > 0, \forall x \geq 0, (|f'(x)| + |f(x)|) \leq \gamma \Big\}$.

Lemma 3.5.2. *Assume that* **(H.1)** *and* **(H.2)** *are satisfied. Then, for all $f \in \mathcal{C}_b^1(\mathbb{R}_+)$, we have*

$$\lim_{n\to\infty} \frac{1}{|\mathbb{T}_n|} \sum_{k \in \mathbb{T}_n} f(X_k) = \mathbb{E}[f(T)] \qquad a.s.$$

Proposition 3.5.3. *Assume that* **(H.1)** *to* **(H.3)** *are satisfied. Then, we have*

$$\lim_{n\to\infty} \frac{\langle M \rangle_n}{|\mathbb{T}_{n-1}|} = L \qquad a.s. \tag{3.5.3}$$

where L is the positive definite matrix given by

$$L = \mathbb{E}\left[\frac{1}{(1+T^2)^2} \begin{pmatrix} P(T) & Q(T) \\ Q(T) & R(T) \end{pmatrix} \otimes \begin{pmatrix} T^2 & T \\ T & 1 \end{pmatrix} \right].$$

Our first result deals with the almost sure convergence of our WLS estimator $\widehat{\theta}_n$.

Theorem 3.5.4. *Assume that* **(H.1)** *to* **(H.5)** *satisfied. Then, $\widehat{\theta}_n$ converges almost surely to θ with the rate of convergence*

$$\|\widehat{\theta}_n - \theta\|^2 = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|} \right) \qquad a.s.$$

In addition, we also have the quadratic strong law

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} |\mathbb{T}_{k-1}| (\widehat{\theta}_k - \theta)^t \Lambda (\widehat{\theta}_k - \theta) = tr(\Lambda^{-1/2} L \Lambda^{-1/2}) \qquad a.s. \tag{3.5.4}$$

where

$$\Lambda = I_2 \otimes C \qquad \text{and} \qquad C = \mathbb{E}\left[\frac{1}{1+T^2} \begin{pmatrix} T^2 & T \\ T & 1 \end{pmatrix} \right]. \tag{3.5.5}$$

Our second result concerns the almost sure asymptotic properties of our WLS variance and covariance estimators $\widehat{\eta}_n$, $\widehat{\zeta}_n$ and $\widehat{\nu}_n$. Let

$$\eta_n = Q_{n-1}^{-1} \sum_{k\in\mathbb{T}_{n-1}} \frac{1}{d_k} V_{2k}^2 \psi_k, \qquad \zeta_n = Q_{n-1}^{-1} \sum_{k\in\mathbb{T}_{n-1}} \frac{1}{d_k} V_{2k+1}^2 \psi_k,$$

$$\nu_n = Q_{n-1}^{-1} \sum_{k\in\mathbb{T}_{n-1}} \frac{1}{d_k} V_{2k} V_{2k+1} \psi_k.$$

Theorem 3.5.5. *Assume that* **(H.1)** *to* **(H.5)** *are satisfied. Then,* $\widehat{\eta}_n$ *and* $\widehat{\zeta}_n$ *converge almost surely to* η *and* ζ *respectively. More precisely,*

$$\|\widehat{\eta}_n - \eta_n\| = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \qquad a.s. \tag{3.5.6}$$

$$\|\widehat{\zeta}_n - \zeta_n\| = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \qquad a.s. \tag{3.5.7}$$

In addition, $\widehat{\nu}_n$ *converges almost surely to* ν *with*

$$\|\widehat{\nu}_n - \nu_n\| = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \qquad a.s. \tag{3.5.8}$$

Remark 3.5.6. *We also have the almost sure rates of convergence*

$$\|\widehat{\eta}_n - \eta\|^2 = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right), \quad \|\widehat{\zeta}_n - \zeta\|^2 = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right), \quad \|\widehat{\nu}_n - \nu\|^2 = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \quad a.s.$$

Our last result is devoted to the asymptotic normality of our WLS estimators $\widehat{\theta}_n$, $\widehat{\eta}_n$, $\widehat{\zeta}_n$ and $\widehat{\nu}_n$.

Theorem 3.5.7. *Assume that* **(H.1)** *to* **(H.5)** *are satisfied. Then, we have the asymptotic normality*

$$\sqrt{|\mathbb{T}_{n-1}|}(\widehat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Lambda^{-1} L \Lambda^{-1}). \tag{3.5.9}$$

In addition, we also have

$$\sqrt{|\mathbb{T}_{n-1}|}\,(\widehat{\eta}_n - \eta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, D^{-1} M_{ac} D^{-1}), \tag{3.5.10}$$

$$\sqrt{|\mathbb{T}_{n-1}|}\left(\widehat{\zeta}_n - \zeta\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, D^{-1} M_{bd} D^{-1}), \tag{3.5.11}$$

where

$$D = \mathbb{E}\left[\frac{1}{(1+T^2)^2}\begin{pmatrix} T^4 & T^2 \\ T^2 & 1 \end{pmatrix}\right],$$

$$M_{ac} = \mathbb{E}\left[\frac{(\mu_a^4 - \sigma_a^4)T^4 + 4\sigma_a^2\sigma_c^2T^2 + (\mu_c^4 - \sigma_c^4)}{(1+T^2)^4}\begin{pmatrix} T^4 & T^2 \\ T^2 & 1 \end{pmatrix}\right],$$

$$M_{bd} = \mathbb{E}\left[\frac{(\mu_b^4 - \sigma_b^4)T^4 + 4\sigma_b^2\sigma_d^2T^2 + (\mu_d^4 - \sigma_d^4)}{(1+T^2)^4}\begin{pmatrix} T^4 & T^2 \\ T^2 & 1 \end{pmatrix}\right].$$

Finally,

$$\sqrt{|\mathbb{T}_{n-1}|}\,(\widehat{\nu}_n - \nu) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, D^{-1}HD^{-1}\right) \tag{3.5.12}$$

where

$$H = \mathbb{E}\left[\frac{(\nu_{ab}^2 - \rho_{ab}^2)T^4 + (\sigma_a^2\sigma_d^2 + \sigma_b^2\sigma_c^2 + 2\rho_{ab}\rho_{cd})T^2 + (\nu_{cd}^2 - \rho_{cd}^2)}{(1+T^2)^4}\begin{pmatrix} T^4 & T^2 \\ T^2 & 1 \end{pmatrix}\right].$$

The rest of the paper is dedicated to the proof of our main results.

3.6 Proofs

3.6.1 Proof of Lemma 3.5.1

We can reformulate (3.5.1) and (3.5.2) as

$$Y_n = \widetilde{a}_n\widetilde{a}_{n-1}\ldots\widetilde{a}_2Y_1 + \sum_{k=2}^{n-1}\widetilde{a}_n\widetilde{a}_{n-1}\ldots\widetilde{a}_{k+1}e_k + e_n.$$

We already made the assumption that both $(a_n, b_n)_{n\geq 1}$ and $(\varepsilon_{2n}, \varepsilon_{2n+1})_{n\geq 1}$ are i.i.d. and that those two sequences are independent. Consequently, the couples $(\widetilde{a}_k, e_k)$ and $(\widetilde{a}_{n-k+2}, e_{n-k+1})$ share the same distribution. Hence, for all $n \geq 2$, Y_n has the same distribution than the random variable

$$\begin{aligned} Z_n &= \widetilde{a}_2\ldots\widetilde{a}_nY_1 + \sum_{k=2}^{n-1}\widetilde{a}_2\widetilde{a}_3\ldots\widetilde{a}_{n-k+1}e_{n-k+2} + e_2, \\ &= \widetilde{a}_2\ldots\widetilde{a}_nY_1 + \sum_{k=3}^{n}\widetilde{a}_2\widetilde{a}_3\ldots\widetilde{a}_{k-1}e_k + e_2. \end{aligned}$$

For the sake of simplicity, we will denote

$$Z_n = \widetilde{a}_2\ldots\widetilde{a}_nY_1 + \sum_{k=2}^{n}\widetilde{a}_2\widetilde{a}_3\ldots\widetilde{a}_{k-1}e_k. \tag{3.6.1}$$

On the first hand, $\mathbb{E}[\widetilde{a}_2\widetilde{a}_3\ldots\widetilde{a}_nY_1] = \mathbb{E}[\widetilde{a}_2]^{n-1}\mathbb{E}[Y_1]$ and since

$$|\mathbb{E}[\widetilde{a}_2]| = \left|\frac{a+b}{2}\right| < 1$$

this immediately leads to

$$\lim_{n\to\infty} \widetilde{a}_2\widetilde{a}_3\dots\widetilde{a}_n Y_1 = 0 \qquad \text{a.s.}$$

On the other hand, let T_n be defined as

$$T_n = \sum_{k=2}^{n} \widetilde{a}_2\widetilde{a}_3\dots\widetilde{a}_{k-1}e_k$$

and T given by

$$T = \sum_{k=2}^{\infty} \widetilde{a}_2\widetilde{a}_3\dots\widetilde{a}_{k-1}e_k.$$

We have

$$\begin{aligned}
\mathbb{E}[|T-T_n|] &= \mathbb{E}\left[\left|\sum_{k=n+1}^{\infty} \widetilde{a}_2\widetilde{a}_3\dots\widetilde{a}_{k-1}e_k\right|\right], \\
&\leq \sum_{k=n+1}^{\infty} \mathbb{E}\left[|\widetilde{a}_2\widetilde{a}_3\dots\widetilde{a}_{k-1}e_k|\right], \\
&\leq \mathbb{E}[|e_2|] \sum_{k=n+1}^{\infty} \mathbb{E}\left[|\widetilde{a}_2|\right]^{k-2}.
\end{aligned}$$

In addition, $\mathbb{E}[a_n^2] < 1$ and $\mathbb{E}[b_n^2] < 1$ which leads to $\mathbb{E}[\widetilde{a}_n^2] < 1$ and $\mathbb{E}[|\widetilde{a}_n|] < 1$. Consequently,

$$\mathbb{E}[|T-T_n|] \leq \mathbb{E}\left[|\widetilde{a}_2|\right]^{n-1} \frac{\mathbb{E}[|e_2|]}{1-\mathbb{E}\left[|\widetilde{a}_2|\right]}.$$

This proves that $T_n \xrightarrow{L^1} T$ which immediately implies that

$$T_n \xrightarrow{\mathcal{L}} T \qquad \text{and} \qquad Y_n \xrightarrow{\mathcal{L}} T.$$

Moreover, we can easily see that **(H.1)** allows us to say that $\mathbb{E}[T^2] < \infty$ thanks to the Cauchy-Schwarz inequality. It only remains to prove that T is not degenerate. First, we easily have, since $\mathbb{E}[|\widetilde{a}_2|] < 1$

$$\begin{aligned}
\mathbb{E}[T] &= \mathbb{E}\left[\sum_{k=2}^{\infty} \widetilde{a}_2\widetilde{a}_3\dots\widetilde{a}_{k-1}e_k\right] = \sum_{k=2}^{\infty} \mathbb{E}\left[\widetilde{a}_2\widetilde{a}_3\dots\widetilde{a}_{k-1}e_k\right], \\
&= \sum_{k=2}^{\infty} \mathbb{E}\left[\widetilde{a}_2\right]\mathbb{E}\left[\widetilde{a}_3\right]\dots\mathbb{E}\left[\widetilde{a}_{k-1}\right]\mathbb{E}\left[e_k\right] = \frac{c+d}{2-(a+b)}.
\end{aligned}$$

Then, we can calculate $\mathbb{E}[T^2]$ as follows

$$\begin{aligned}
\mathbb{E}[T^2] &= \mathbb{E}\left[\left(\sum_{k=2}^{\infty}\widetilde{a}_2\widetilde{a}_3\ldots\widetilde{a}_{k-1}e_k\right)^2\right],\\
&= \sum_{k=2}^{\infty}\mathbb{E}[\widetilde{a}_2^2\widetilde{a}_3^2\ldots\widetilde{a}_{k-1}^2e_k^2] + 2\sum_{k=2}^{\infty}\sum_{l=k+1}^{\infty}\mathbb{E}[\widetilde{a}_2^2\widetilde{a}_3^2\ldots\widetilde{a}_{k-1}^2\widetilde{a}_k e_k\widetilde{a}_{k+1}\ldots\widetilde{a}_{l-1}e_l],\\
&= \sum_{k=2}^{\infty}\left(\frac{\sigma_a^2+\sigma_b^2+a^2+b^2}{2}\right)^{k-2}\frac{\sigma_c^2+\sigma_d^2+c^2+d^2}{2}\\
&\quad + 2\sum_{k=2}^{\infty}\sum_{l=k+1}^{\infty}\left(\frac{\sigma_a^2+\sigma_b^2+a^2+b^2}{2}\right)^{k-2}\frac{ac+bd}{2}\left(\frac{a+b}{2}\right)^{l-k-2}\frac{c+d}{2},\\
&= \frac{\sigma_c^2+\sigma_d^2+c^2+d^2}{2-(\sigma_a^2+\sigma_b^2+a^2+b^2)} + \frac{2(ac+bd)(c+d)}{(2-(\sigma_a^2+\sigma_b^2+a^2+b^2))(2-(a+b))}.
\end{aligned}$$

This allows us to say that

$$\begin{aligned}
\mathrm{Var}(T) = \frac{\sigma_c^2+\sigma_d^2}{2-(\sigma_a^2+\sigma_b^2+a^2+b^2)} &+ \left(\frac{c+d}{2-(a+b)}\right)^2\frac{\sigma_a^2+\sigma_b^2}{2-(\sigma_a^2+\sigma_b^2+a^2+b^2)}\\
&+ \frac{2}{2-(\sigma_a^2+\sigma_b^2+a^2+b^2)}\frac{(ad-bc+c-d)^2}{(2-(a+b))^2}.
\end{aligned}$$

Under hypothesis **(H.1)** and **(H.2)** we immediately have that the first term is positive and that the two other terms are non-negative, allowing us to say that T is not degenerate.

3.6.2 Proof of the keystone Lemma 3.5.2

We shall now prove that for all $f \in \mathcal{C}_b^1(\mathbb{R}_+)$,

$$\lim_{n\to\infty}\frac{1}{|\mathbb{T}_n|}\sum_{k\in\mathbb{T}_n} f(X_k) = \mathbb{E}[f(T)].$$

Denote $g = f - \mathbb{E}[f(T)]$,

$$\overline{M}_{\mathbb{T}_n}(f) = \frac{1}{|\mathbb{T}_n|}\sum_{k\in\mathbb{T}_n} f(X_k) \qquad \text{and} \qquad \overline{M}_{\mathbb{G}_n}(f) = \frac{1}{|\mathbb{G}_n|}\sum_{k\in\mathbb{G}_n} f(X_k).$$

Via Lemma A.2 of [11], it is only necessary to prove that

$$\lim_{n\to\infty}\frac{1}{|\mathbb{G}_n|}\sum_{k\in\mathbb{G}_n} g(X_k) = 0 \qquad \text{a.s.}$$

We shall follow the induced Markov chain approach, originally proposed by Guyon in [26]. Let Q be the transition probability of (Y_n), Q^p the p-th iterated of Q. In addition, denote by ν the distribution of $Y_1 = X_1$ and νQ^p the law of Y_p. Finally, let P be the transition probability of (X_n) as defined in [26]. We obtain from relation (7) of [26] that for all $n \geq 0$

$$\mathbb{E}[\overline{M}_{\mathbb{G}_n}(g)^2] = \frac{1}{2^n}\nu Q^n g^2 + \sum_{k=0}^{n-1} \frac{1}{2^{k+1}}\nu Q^k P(Q^{n-k-1}g \star Q^{n-k-1}g)$$

where, for all $x, y \in \mathbb{N}$, $(f \star g)(x, y) = f(x)g(y)$. Consequently,

$$\begin{aligned}\sum_{n=0}^{\infty} \mathbb{E}[\overline{M}_{\mathbb{G}_n}(g)^2] &= \sum_{n=0}^{\infty} \frac{1}{2^n}\nu Q^n g^2 + \sum_{n=1}^{\infty}\sum_{k=0}^{n-1} \frac{1}{2^{k+1}}\nu Q^k P(Q^{n-k-1}g \star Q^{n-k-1}g), \\ &\leq \sum_{k=0}^{\infty} \frac{1}{2^k}\nu Q^k \left(g^2 + P\left(\sum_{l=0}^{\infty} |Q^l g \star Q^l g|\right)\right). \qquad (3.6.2)\end{aligned}$$

However, for all $x \in \mathbb{N}$,

$$Q^n g(x) = Q^n f(x) - \mathbb{E}[f(T)] = \mathbb{E}_x[f(Y_n) - f(T)] = \mathbb{E}_x[f(Z_n) - f(T)]$$

where Z_n is given by (3.6.1). Hence, we deduce from the mean value theorem and the Cauchy-Schwarz inequality that

$$|Q^n g(x)| \leq \mathbb{E}_x[W_n|Z_n - T|] \leq \mathbb{E}_x[W_n^2]^{1/2}\mathbb{E}_x[(Z_n - T)^2]^{1/2} \qquad (3.6.3)$$

where $W_n = \sup_{z\in[Z_n,T]} |f'(z)|$. By the very definition of $\mathcal{C}_b^1(\mathbb{R}_+)$, one can find some constant $\gamma > 0$ such that $|f'(z)| \leq \gamma$. Hence,

$$\mathbb{E}_x[W_n^2]^{1/2} \leq \gamma. \qquad (3.6.4)$$

Furthermore

$$Z_n - T = \widetilde{a}_2 \dots \widetilde{a}_n Y_1 - \sum_{k=n}^{\infty} \widetilde{a}_2 \dots \widetilde{a}_k e_{k+1}$$

and the triangle inequality allows us to say that

$$\begin{aligned}\mathbb{E}_x[(Z_n - T)^2]^{1/2} &\leq \mathbb{E}_x[(\widetilde{a}_2 \dots \widetilde{a}_n Y_1)^2]^{1/2} + \sum_{k=n}^{\infty} \mathbb{E}_x[(\widetilde{a}_2 \dots \widetilde{a}_k e_{k+1})^2]^{1/2} \\ &\leq \mathbb{E}[\widetilde{a}_2^2]^{(n-1)/2}\mathbb{E}_x[Y_1^2]^{1/2} + \sum_{k=n}^{\infty} \mathbb{E}_x[\widetilde{a}_2^2]^{(k-1)/2}\mathbb{E}[e_{k+1}^2]^{1/2} \\ &\leq \sqrt{\mathbb{E}[\widetilde{a}_2^2]}^{\,n-1} \left(|x| + \frac{\mathbb{E}[e_2^2]^{1/2}}{1 - \mathbb{E}[\widetilde{a}_2^2]^{1/2}}\right) \\ &\leq \alpha\sqrt{\mathbb{E}[\widetilde{a}_2^2]}^{\,n} (1 + |x|) \qquad (3.6.5)\end{aligned}$$

where

$$\alpha = \max\left(1, \frac{\mathbb{E}[e_2^2]^{1/2}}{1-\mathbb{E}[\widetilde{a}_2^2]^{1/2}}\right).$$

Finally, we obtain from (3.6.3) together with (3.6.4) and (3.6.5) that

$$|Q^n g(x)| \leq \gamma\alpha\sqrt{\mathbb{E}[\widetilde{a}_2^2]}^{\,n-1}(1+|x|).$$

Therefore,

$$P\left(\sum_{n=0}^{\infty}|Q^n g \star Q^n g|\right) \leq \frac{\gamma^2\alpha^2}{1-\mathbb{E}[\widetilde{a}_2^2]}P(h\star h) \tag{3.6.6}$$

where, for all $x \in \mathbb{N}$, $h(x) = 1+|x|$. We are now in position to prove that

$$\mathbb{E}\left[\sum_{n=0}^{\infty}\overline{M}_{\mathbb{G}_n}(g)^2\right] < \infty. \tag{3.6.7}$$

Let G be be the random vector defined by $G(x) = (a_1x+\varepsilon_2, b_1x+\varepsilon_3)^t$. We can easily see from **(H.2)** that it exists some constant $\beta > 0$ such that

$$P(h\star h)(x) = \mathbb{E}[(h\star h)(G(x))] \leq \beta(1+x^2).$$

Consequently, since, for all $z \in \mathbb{R}$, $|g(z)| \leq 2\gamma$, we obtain from (3.6.2) together with (3.6.6) that

$$\begin{aligned}\sum_{n=0}^{\infty}\mathbb{E}[\overline{M}_{\mathbb{G}_n}(g)^2] &\leq \sum_{k=0}^{\infty}\frac{1}{2^k}\left(\mathbb{E}[g^2(Y_k)] + \frac{\beta\gamma^2\alpha^2}{1-\mathbb{E}[\widetilde{a}_2^2]}(1+\mathbb{E}[Y_k^2])\right),\\ &\leq \left(8\gamma^2 + \frac{\beta\gamma^2\alpha^2}{1-\mathbb{E}[\widetilde{a}_2^2]}\right)\left(1+\sum_{k=0}^{\infty}\frac{1}{2^k}\mathbb{E}[Y_k^2]\right). \end{aligned} \tag{3.6.8}$$

In addition, we also have

$$\begin{aligned}\mathbb{E}[Y_k^2]^{1/2} &= \mathbb{E}[Z_k^2]^{1/2},\\ &\leq \mathbb{E}_x[(\widetilde{a}_2\ldots\widetilde{a}_nY_1)^2]^{1/2} + \sum_{k=2}^{n}\mathbb{E}_x[(\widetilde{a}_2\ldots\widetilde{a}_{k-1}e_k)^2]^{1/2},\\ &\leq \mathbb{E}[\widetilde{a}_2^2]^{(n-1)/2}\mathbb{E}_x[Y_1^2]^{1/2} + \sum_{k=2}^{\infty}\mathbb{E}_x[\widetilde{a}_2^2]^{(k-2)/2}\mathbb{E}[e_{k+1}^2]^{1/2},\\ &\leq \mathbb{E}[X_1^2]^{1/2} + \frac{\mathbb{E}[e_2^2]^{1/2}}{1-\mathbb{E}[\widetilde{a}_2^2]^{1/2}}.\end{aligned} \tag{3.6.9}$$

Then, (3.6.8) and (3.6.9) immediately lead to (3.6.7). Finally, the monotone convergence theorem implies that

$$\lim_{n\to\infty}\overline{M}_{\mathbb{G}_n}(g) = 0 \qquad \text{a.s.}$$

which completes the proof of Lemma 3.5.2.

3.6.3 Proof of the convergence of $\langle M \rangle_n$

The almost sure convergence (3.5.3) immediately follows from (3.4.2) and (3.4.3) together with Lemma 3.5.2. It only remains to prove that $\det(L) > 0$ where the limiting matrix L can be rewritten as $L = \mathbb{E}\left[\Gamma \otimes \mathcal{C}\right]$, where

$$\Gamma = \begin{pmatrix} P(T) & Q(T) \\ Q(T) & R(T) \end{pmatrix} \qquad \text{and} \qquad \mathcal{C} = \frac{1}{(1+T^2)^2}\begin{pmatrix} T^2 & T \\ T & 1 \end{pmatrix}.$$

We have

$$\begin{aligned} L &= \mathbb{E}\left[\begin{pmatrix} \sigma_a^2 T^2 & \rho_{ab} T^2 \\ \rho_{ab} T^2 & \sigma_b^2 T^2 \end{pmatrix} \otimes \mathcal{C}\right] + \mathbb{E}\left[\begin{pmatrix} \sigma_c^2 & \rho_{cd} \\ \rho_{cd} & \sigma_d^2 \end{pmatrix} \otimes \mathcal{C}\right], \\ &= \begin{pmatrix} \sigma_a^2 & \rho_{ab} \\ \rho_{ab} & \sigma_b^2 \end{pmatrix} \otimes \mathbb{E}[T^2\mathcal{C}] + \begin{pmatrix} \sigma_c^2 & \rho_{cd} \\ \rho_{cd} & \sigma_d^2 \end{pmatrix} \otimes \mathbb{E}[\mathcal{C}]. \end{aligned} \tag{3.6.10}$$

We shall prove that $\mathbb{E}[\mathcal{C}]$ is a positive definite matrix and that $\mathbb{E}[T^2\mathcal{C}]$ is a positive semidefinite matrix. Denote by λ_1 and λ_2 the two eigenvalues of the real symmetric matrix $\mathbb{E}[\mathcal{C}]$. We clearly have

$$\lambda_1 + \lambda_2 = tr(\mathbb{E}[\mathcal{C}]) = \mathbb{E}\left[\frac{T^2+1}{(1+T^2)^2}\right] > 0$$

and

$$\lambda_1\lambda_2 = \det(\mathbb{E}[\mathcal{C}]) = \mathbb{E}\left[\frac{T^2}{(1+T^2)^2}\right]\mathbb{E}\left[\frac{1}{(1+T^2)^2}\right] - \mathbb{E}\left[\frac{T}{(1+T^2)^2}\right]^2 \geq 0$$

thanks to the Cauchy-Schwarz inequality and $\lambda_1\lambda_2 = 0$ if and only if T is degenerate, which is not the case thanks to Lemma 3.5.1. Consequently, $\mathbb{E}[\mathcal{C}]$ is a positive definite matrix. In the same way, we can prove that $\mathbb{E}[T^2\mathcal{C}]$ is a positive semidefinite matrix. Since the Kronecker product of two positive semidefinite (respectively positive definite) matrices is a positive semidefinite (respectively positive definite) matrix, we deduce from (3.6.10) that L is positive definite as soon as $\rho_{cd}^2 < \sigma_c^2\sigma_d^2$ and $\rho_{ab}^2 \leq \sigma_a^2\sigma_b^2$ which is the case thanks to **(H.3)**.

3.6.4 Preliminary work for the almost sure convergence of $\widehat{\theta}_n$

We will follow the same approach as in Bercu et al. [11]. For all $n \geq 1$, let $\mathcal{V}_n = M_n^t \Sigma_{n-1}^{-1} M_n = (\widehat{\theta}_n - \theta)^t \Sigma_{n-1}(\widehat{\theta}_n - \theta)$. First of all, we have

$$\begin{aligned} \mathcal{V}_{n+1} &= M_{n+1}^t \Sigma_n^{-1} M_{n+1} = (M_n + \Delta M_{n+1})^t \Sigma_n^{-1}(M_n + \Delta M_{n+1}), \\ &= M_n^t \Sigma_n^{-1} M_n + 2M_n^t \Sigma_n^{-1} \Delta M_{n+1} + \Delta M_{n+1}^t \Sigma_n^{-1} \Delta M_{n+1}, \\ &= \mathcal{V}_n - M_n^t(\Sigma_{n-1}^{-1} - \Sigma_n^{-1})M_n + 2M_n^t \Sigma_n^{-1} \Delta M_{n+1} + \Delta M_{n+1}^t \Sigma_n^{-1} \Delta M_{n+1}. \end{aligned}$$

By summing over this identity, we obtain the main decomposition

$$\mathcal{V}_{n+1} + \mathcal{A}_n = \mathcal{V}_1 + \mathcal{B}_{n+1} + \mathcal{W}_{n+1} \tag{3.6.11}$$

where

$$\mathcal{A}_n = \sum_{k=1}^{n} M_k^t(\Sigma_{k-1}^{-1} - \Sigma_k^{-1})M_k,$$

$$\mathcal{B}_{n+1} = 2\sum_{k=1}^{n} M_k^t \Sigma_k^{-1} \Delta M_{k+1} \quad \text{and} \quad \mathcal{W}_{n+1} = \sum_{k=1}^{n} \Delta M_{k+1}^t \Sigma_k^{-1} \Delta M_{k+1}.$$

Lemma 3.6.1. *Assume that* **(H.1)** *to* **(H.3)** *are satisfied. Then, we have*

$$\lim_{n\to\infty} \frac{\mathcal{W}_n}{n} = \frac{1}{2} tr((I_2 \otimes C)^{-1/2} L (I_2 \otimes C)^{-1/2}) \qquad a.s. \tag{3.6.12}$$

where C is the positive definite matrix given by (3.5.5)*. In addition, we also have*

$$\mathcal{B}_{n+1} = o(n) \qquad a.s. \tag{3.6.13}$$

and

$$\lim_{n\to\infty} \frac{\mathcal{V}_{n+1} + \mathcal{A}_n}{n} = \frac{1}{2} tr((I_2 \otimes C)^{-1/2} L (I_2 \otimes C)^{-1/2}) \qquad a.s. \tag{3.6.14}$$

Proof. First of all, we have $\mathcal{W}_{n+1} = \mathcal{T}_{n+1} + \mathcal{R}_{n+1}$ where

$$\mathcal{T}_{n+1} = \sum_{k=1}^{n} \frac{\Delta M_{k+1}^t (I_2 \otimes C)^{-1} \Delta M_{k+1}}{|\mathbb{T}_k|},$$

$$\mathcal{R}_{n+1} = \sum_{k=1}^{n} \frac{\Delta M_{k+1}^t (|\mathbb{T}_k| \Sigma_k^{-1} - (I_2 \otimes C)^{-1}) \Delta M_{k+1}}{|\mathbb{T}_k|}.$$

One can observe that $\mathcal{T}_{n+1} = tr((I_2 \otimes C)^{-1/2} \mathcal{H}_{n+1} (I_2 \otimes C)^{-1/2})$ where

$$\mathcal{H}_{n+1} = \sum_{k=1}^{n} \frac{\Delta M_{k+1} \Delta M_{k+1}^t}{|\mathbb{T}_k|}.$$

Our aim is to make use of the strong law of large numbers for martingale transforms, so we start by adding and subtracting a term involving the conditional expectation of $\Delta\mathcal{H}_{n+1}$ given $\mathcal{F}_n$. We have thanks to relation (3.4.3) that for all $n \geq 0$, $\mathbb{E}[\Delta M_{n+1} \Delta M_{n+1}^t | \mathcal{F}_n] = L_n$. Consequently, we can split $\mathcal{H}_{n+1}$ into two terms

$$\mathcal{H}_{n+1} = \sum_{k=1}^{n} \frac{L_k}{|\mathbb{T}_k|} + \mathcal{K}_{n+1}, \qquad \text{where} \qquad \mathcal{K}_{n+1} = \sum_{k=1}^{n} \frac{\Delta M_{k+1} \Delta M_{k+1}^t - L_k}{|\mathbb{T}_k|}.$$

It clearly follows from convergence (3.5.3) that

$$\lim_{n\to\infty} \frac{L_n}{|\mathbb{T}_n|} = \frac{1}{2}L \qquad \text{a.s.}$$

Hence, Cesaro convergence immediately implies that

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \frac{L_k}{|\mathbb{T}_k|} = \frac{1}{2}L \qquad \text{a.s.} \tag{3.6.15}$$

On the other hand, the sequence $(\mathcal{K}_n)_{n\geq 2}$ is obviously a square integrable martingale. Moreover, we have

$$\Delta\mathcal{K}_{n+1} = \mathcal{K}_{n+1} - \mathcal{K}_n = \frac{1}{|\mathbb{T}_n|}(\Delta M_{n+1}\Delta M_{n+1}^t - L_n).$$

For all $u \in \mathbb{R}^4$, denote $\mathcal{K}_n(u) = u^t\mathcal{K}_n u$. It follows from tedious but straightforward calculations, together with Lemma 3.5.2, that the increasing process of the martingale $(\mathcal{K}_n(u))_{n\geq 2}$ satisfies $\langle \mathcal{K}(u)\rangle_n = \mathcal{O}(n)$ a.s. Therefore, we deduce from the strong law of large numbers for martingales that for all $u \in \mathbb{R}^4$, $\mathcal{K}_n(u) = o(n)$ a.s. leading to $\mathcal{K}_n = o(n)$ a.s. Hence, we infer from (3.6.15) that

$$\lim_{n\to\infty} \frac{\mathcal{H}_{n+1}}{n} = \frac{1}{2}L \qquad \text{a.s.} \tag{3.6.16}$$

Via the same arguments as in the proof of convergence (3.5.3), we find that

$$\lim_{n\to\infty} \frac{\Sigma_n}{|\mathbb{T}_n|} = I_2 \otimes C \qquad \text{a.s.} \tag{3.6.17}$$

where C is the positive definite matrix given by (3.5.5). Then, we obtain from (3.6.16) that

$$\lim_{n\to\infty} \frac{\mathcal{T}_n}{n} = \frac{1}{2}tr((I_2 \otimes C)^{-1/2}L(I_2 \otimes C)^{-1/2}) \qquad \text{a.s.}$$

which allows us to say that $\mathcal{R}_n = o(n)$ a.s. leading to (3.6.12). We are now in position to prove (3.6.13). Let us recall that

$$\mathcal{B}_{n+1} = 2\sum_{k=1}^{n} M_k^t\Sigma_k^{-1}\Delta M_{k+1} = 2\sum_{k=1}^{n} M_k^t\Sigma_k^{-1}\Psi_k\xi_{k+1}.$$

Hence, $(\mathcal{B}_n)_{n\geq 2}$ is a square integrable martingale. In addition, we have

$$\Delta\mathcal{B}_{n+1} = 2M_n^t\Sigma_n^{-1}\Delta M_{n+1}.$$

Thus

$$\begin{aligned}\mathbb{E}[(\Delta\mathcal{B}_{n+1})^2|\mathcal{F}_n] &= 4\mathbb{E}[M_n^t\Sigma_n^{-1}\Delta M_{n+1}\Delta M_{n+1}^t\Sigma_n^{-1}M_n|\mathcal{F}_n] \qquad \text{a.s.}\\ &= 4M_n^t\Sigma_n^{-1}\mathbb{E}[\Delta M_{n+1}\Delta M_{n+1}^t|\mathcal{F}_n]\Sigma_n^{-1}M_n \qquad \text{a.s.}\\ &= 4M_n^t\Sigma_n^{-1}L_n\Sigma_n^{-1}M_n \qquad \text{a.s.}\end{aligned}$$

We can observe that

$$L_n = \sum_{k\in\mathbb{G}_n}\frac{1}{c_k^2}\begin{pmatrix}P(X_k) & Q(X_k)\\ Q(X_k) & R(X_k)\end{pmatrix}\otimes\begin{pmatrix}X_k^2 & X_k\\ X_k & 1\end{pmatrix}$$

and

$$\Psi_n\Psi_n^t = \sum_{k\in\mathbb{G}_n}\frac{1}{c_k}I_2\otimes\begin{pmatrix}X_k^2 & X_k\\ X_k & 1\end{pmatrix}.$$

For $\alpha = \max(\sigma_a^2,\sigma_c^2)+\max(\sigma_b^2,\sigma_d^2)+\max(|\rho_{ab}|,|\rho_{cd}|)$, denote

$$\Delta_n = \begin{pmatrix}\alpha - \dfrac{P(X_n)}{c_n} & -\dfrac{Q(X_n)}{c_n}\\ -\dfrac{Q(X_n)}{c_n} & \alpha - \dfrac{R(X_n)}{c_n}\end{pmatrix}.$$

We can rewrite $\alpha\Psi_n\Psi_n^t - L_n$ as

$$\alpha\Psi_n\Psi_n^t - L_n = \sum_{k\in\mathbb{G}_n}\frac{1}{c_k}\Delta_k\otimes\begin{pmatrix}X_k^2 & X_k\\ X_k & 1\end{pmatrix}.$$

It is not hard to see that Δ_n is a positive definite matrix. As a matter of fact, we deduce from the elementary inequalities

$$\begin{cases}0 < P(X) \le \max(\sigma_a^2,\sigma_c^2)(1+X^2),\\ 0 < R(X) \le \max(\sigma_b^2,\sigma_d^2)(1+X^2),\\ |Q(X)| \le \max(|\rho_{ab}|,|\rho_{cd}|)(1+X^2),\end{cases} \tag{3.6.18}$$

that

$$tr(\Delta_n) = 2\alpha - \frac{P(X_n)}{c_n} - \frac{R(X_n)}{c_n} \ge 2\alpha - \max(\sigma_a^2,\sigma_c^2) - \max(\sigma_b^2,\sigma_d^2) > 0.$$

In addition, we also have from (3.6.18) that

$$\begin{aligned}c_n^2\det(\Delta_n) &= (\alpha c_n - P(X_n))(\alpha c_n - R(X_n)) - Q^2(X_n),\\ &= \alpha c_n\,(\alpha c_n - P(X_n) - R(X_n)) + P(X_n)R(X_n) - Q^2(X_n),\\ &\ge P(X_k)R(X_k) + \alpha c_n^2\max(|\rho_{ab}|,|\rho_{cd}|) - Q^2(X_n),\\ &\ge P(X_k)R(X_k) + \max(|\rho_{ab}|,|\rho_{cd}|)^2c_n^2 - Q^2(X_n) > 0.\end{aligned}$$

Consequently, Δ_n is positive definite which immediately implies that $L_n \leq \alpha \Psi_n \Psi_n^t$. Moreover, we can use Lemma B.1 of [11] to say that

$$\Sigma_n^{-1} \Psi_n \Psi_n^t \Sigma_n^{-1} \leq \Sigma_{n-1}^{-1} - \Sigma_n^{-1}.$$

Hence

$$\begin{aligned}
\mathbb{E}[(\Delta \mathcal{B}_{n+1})^2 | \mathcal{F}_n] &= 4 M_n^t \Sigma_n^{-1} L_n \Sigma_n^{-1} M_n \qquad \text{a.s.} \\
&\leq 4\alpha M_n^t \Sigma_n^{-1} \Psi_n \Psi_n^t \Sigma_n^{-1} M_n \qquad \text{a.s.} \\
&\leq 4\alpha M_n^t (\Sigma_{n-1}^{-1} - \Sigma_n^{-1}) M_n \qquad \text{a.s.}
\end{aligned}$$

leading to $\langle \mathcal{B} \rangle_n \leq 4\alpha \mathcal{A}_n$. Therefore it follows from the strong law of large numbers for martingales that $\mathcal{B}_n = o(\mathcal{A}_n)$. Hence, we deduce from decomposition (3.6.11) that

$$\mathcal{V}_{n+1} + \mathcal{A}_n = o(\mathcal{A}_n) + \mathcal{O}(n) \qquad \text{a.s.}$$

leading to, since $\mathcal{A}_n$ and $\mathcal{V}_{n+1}$ are non negative, $\mathcal{A}_n = \mathcal{O}(n)$ and $\mathcal{V}_{n+1} = \mathcal{O}(n)$ a.s. which implies that $\mathcal{B}_n = o(n)$ a.s. Finally we clearly obtain convergence (3.6.14) from the main decomposition (3.6.11) together with (3.6.12) and 3.6.13, which completes the proof of Lemma 3.6.1. □

Lemma 3.6.2. *Assume that* **(H.1)** *to* **(H.5)** *are satisfied. For all $\delta > 1/2$, we have*

$$\|M_n\|^2 = o(|\mathbb{T}_n| n^\delta) \qquad \textit{a.s.}$$

Proof. Let us recall that

$$M_n = \sum_{k \in \mathbb{T}_{n-1}} \frac{1}{c_k} \begin{pmatrix} X_k V_{2k} \\ V_{2k} \\ X_k V_{2k+1} \\ V_{2k+1} \end{pmatrix}.$$

Denote

$$P_n = \sum_{k \in \mathbb{T}_{n-1}} \frac{X_k V_{2k}}{c_k} \qquad \text{and} \qquad Q_n = \sum_{i \in \mathbb{T}_{n-1}} \frac{V_{2k}}{c_k}.$$

On the one hand, P_n can be rewritten as

$$P_n = \sum_{k=1}^{n} \sqrt{|\mathbb{G}_{k-1}|} f_k \qquad \text{where} \qquad f_n = \frac{1}{\sqrt{|\mathbb{G}_{n-1}|}} \sum_{k \in \mathbb{G}_{n-1}} \frac{X_k V_{2k}}{c_k}.$$

We already saw in Section 3.3 that for all $k \in \mathbb{G}_n$,

$$\mathbb{E}[V_{2k} | \mathcal{F}_n] = 0 \qquad \text{and} \qquad \mathbb{E}[V_{2k}^2 | \mathcal{F}_n] = \sigma_a^2 X_k^2 + \sigma_c^2 = P(X_k) \qquad \text{a.s.}$$

In addition, for all $k \in \mathbb{G}_n$,

$$\mathbb{E}[V_{2k}^4|\mathcal{F}_n] = \mu_a^4 X_k^4 + 6\sigma_a^2\sigma_c^2 X_k^2 + \mu_c^4 \qquad \text{a.s.}$$

which implies that

$$\mathbb{E}[V_{2k}^4|\mathcal{F}_n] \leq \mu_{ac}^4 c_k^2 \qquad \text{a.s.}. \tag{3.6.19}$$

where $\mu_{ac}^4 = \max(\mu_a^4, 3\sigma_a^2\sigma_c^2, \mu_c^4)$. Consequently, $\mathbb{E}[f_{n+1}|\mathcal{F}_n] = 0$ a.s. and we deduce from (3.6.19) together with the Cauchy-Schwarz inequality that

$$\begin{aligned}
\mathbb{E}[f_{n+1}^4|\mathcal{F}_n] &= \frac{1}{|\mathbb{G}_n|}\mathbb{E}\left[\left(\sum_{k\in\mathbb{G}_n}\frac{X_kV_{2k}}{c_k}\right)^4\middle|\mathcal{F}_n\right], \\
&= \frac{1}{|\mathbb{G}_n|^2}\sum_{k\in\mathbb{G}_n}\left(\frac{X_k}{\sqrt{c_k}}\right)^4\frac{\mathbb{E}[V_{2k}^4|\mathcal{F}_n]}{c_k^2} \\
&\qquad + \frac{3}{|\mathbb{G}_n|^2}\sum_{k\in\mathbb{G}_n}\sum_{\substack{l\in\mathbb{G}_n\\ l\neq k}}\left(\frac{X_k}{\sqrt{c_k}}\right)^2\left(\frac{X_l}{\sqrt{c_l}}\right)^2\frac{\mathbb{E}[V_{2k}^2|\mathcal{F}_n]}{c_k}\frac{\mathbb{E}[V_{2l}^2|\mathcal{F}_n]}{c_l}, \\
&\leq \frac{1}{|\mathbb{G}_n|^2}\sum_{k\in\mathbb{G}_n}\mu_{ac}^4 + \frac{3}{|\mathbb{G}_n|^2}\sum_{k\in\mathbb{G}_n}\sum_{\substack{l\in\mathbb{G}_n\\ l\neq k}}\max(\sigma_a^2,\sigma_c^2)^2, \\
&\leq \mu_{ac}^4 + 3\max(\sigma_a^2,\sigma_c^2)^2 \qquad \text{a.s.}
\end{aligned} \tag{3.6.20}$$

Therefore, we infer from (3.6.20) that $\sup_{n\geq 0}\mathbb{E}[f_{n+1}^4|\mathcal{F}_n] < \infty$ a.s. Hence, we obtain from Wei's Lemma given in [57] (2.30) page 1673 that for all $\delta > 1/2$,

$$P_n^2 = o(|\mathbb{T}_{n-1}|n^\delta) \qquad \text{a.s.}$$

On the other hand, Q_n can be rewritten as

$$Q_n = \sum_{k=1}^n \sqrt{|\mathbb{G}_{k-1}|}g_k \qquad \text{where} \qquad g_n = \frac{1}{\sqrt{|\mathbb{G}_{n-1}|}}\sum_{k\in\mathbb{G}_{n-1}}\frac{V_{2k}}{c_k}.$$

Via the same calculation as before, $\mathbb{E}[g_{n+1}|\mathcal{F}_n] = 0$ a.s. and, as $c_n \geq 1$,

$$\mathbb{E}[g_{n+1}^4|\mathcal{F}_n] \leq \mu_{bd}^4 + 3\max(\sigma_b^2,\sigma_d^2)^2 \qquad \text{a.s.}$$

Hence, we deduce once again from Wei's Lemma that for all $\delta > 1/2$,

$$Q_n^2 = o(|\mathbb{T}_{n-1}|n^\delta) \qquad \text{a.s.}$$

In the same way, we obtain the same result for the two last components of M_n, which completes the proof of Lemma 3.6.2. □

3.6.5 Proof of the almost sure convergence results of $\widehat{\theta}_n$

We recall from (3.4.1) that $\widehat{\theta}_n - \theta = \Sigma_{n-1}^{-1} M_n$ which implies

$$\|\widehat{\theta}_n - \theta\|^2 \leq \frac{\mathcal{V}_n}{\lambda_{min}(\Sigma_{n-1})}$$

where $\mathcal{V}_n = M_n^t \Sigma_{n-1}^{-1} M_n$. On the one hand, it follows from (3.6.14) that $\mathcal{V}_n = \mathcal{O}(n)$ a.s. On the other hand, we deduce from (3.6.17) that

$$\lim_{n\to\infty} \frac{\lambda_{min}(\Sigma_n)}{|\mathbb{T}_n|} = \lambda_{min}(C) > 0 \qquad \text{a.s.}$$

Consequently, we find that

$$\|\widehat{\theta}_n - \theta\|^2 = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \qquad \text{a.s.}$$

We are now in position to prove the quadratic strong law (3.5.4). First of all a direct application of Lemma 3.6.2 ensures that $\mathcal{V}_n = o(n^\delta)$ a.s. for all $\delta > 1/2$. Hence, we obtain from (3.6.14) that

$$\lim_{n\to\infty} \frac{\mathcal{A}_n}{n} = \frac{1}{2} tr((I_2 \otimes C)^{-1/2} L (I_2 \otimes C)^{-1/2}) \qquad \text{a.s.} \tag{3.6.21}$$

Let us rewrite $\mathcal{A}_n$ as

$$\mathcal{A}_n = \sum_{k=1}^n M_k^t \left(\Sigma_{k-1}^{-1} - \Sigma_k^{-1}\right) M_k = \sum_{k=1}^n M_k^t \Sigma_{k-1}^{-1/2} A_k \Sigma_{k-1}^{-1/2} M_k$$

where $A_k = I_4 - \Sigma_{k-1}^{1/2} \Sigma_k^{-1} \Sigma_{k-1}^{1/2}$. We already saw from (3.6.17) that

$$\lim_{n\to\infty} \frac{\Sigma_n}{|\mathbb{T}_n|} = I_2 \otimes C \qquad \text{a.s.}$$

which ensures that

$$\lim_{n\to\infty} A_n = \frac{1}{2} I_4 \qquad \text{a.s.}$$

In addition, we deduce from (3.6.14) that $\mathcal{A}_n = \mathcal{O}(n)$ a.s. which implies that

$$\frac{\mathcal{A}_n}{n} = \left(\frac{1}{2n} \sum_{k=1}^n M_k^t \Sigma_{k-1}^{-1} M_k\right) + o(1) \qquad \text{a.s.} \tag{3.6.22}$$

Moreover we have

$$\begin{aligned}
\frac{1}{n} \sum_{k=1}^n M_k^t \Sigma_{k-1}^{-1} M_k &= \frac{1}{n} \sum_{k=1}^n (\widehat{\theta}_k - \theta)^t \Sigma_{k-1} (\widehat{\theta}_k - \theta), \\
&= \frac{1}{n} \sum_{k=1}^n |\mathbb{T}_{k-1}| (\widehat{\theta}_k - \theta)^t \frac{\Sigma_{k-1}}{|\mathbb{T}_{k-1}|} (\widehat{\theta}_k - \theta), \\
&= \frac{1}{n} \sum_{k=1}^n |\mathbb{T}_{k-1}| (\widehat{\theta}_k - \theta)^t (I_2 \otimes C) (\widehat{\theta}_k - \theta) + o(1) \qquad \text{a.s.}
\end{aligned} \tag{3.6.23}$$

Therefore, (3.6.21) together with (3.6.22) and (3.6.23) lead to (3.5.4).

3.6.6 Proof of the almost sure convergence results of $\widehat{\eta}_n$, $\widehat{\zeta}_n$ and $\widehat{\nu}_n$

First of all, we shall only prove (3.5.6) since the proof of (3.5.7) follows exactly the same lines. We clearly have from (3.3.7) that

$$\begin{aligned} Q_{n-1}(\widehat{\eta}_n - \eta_n) &= \sum_{k\in\mathbb{T}_{n-1}} \frac{1}{d_k}(\widehat{V}_{2k}^2 - V_{2k}^2)\psi_k, \\ &= \sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l} \frac{1}{d_k}(\widehat{V}_{2k}^2 - V_{2k}^2)\psi_k, \\ &= \sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l} \frac{1}{d_k}\Big((\widehat{V}_{2k} - V_{2k})^2 + 2(\widehat{V}_{2k} - V_{2k})V_{2k}\Big)\psi_k. \end{aligned} \tag{3.6.24}$$

In addition, we already saw in Section 3.3 that for all $l \geq 0$ and $k \in \mathbb{G}_l$,

$$\widehat{V}_{2k} - V_{2k} = -\begin{pmatrix} \widehat{a}_l - a \\ \widehat{c}_l - c \end{pmatrix}^t \Phi_k.$$

Consequently,

$$(\widehat{V}_{2k} - V_{2k})^2 \leq \|\Phi_k\|^2\left((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2\right) = c_k\left((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2\right).$$

Hence, as $\|\psi_k\|^2 = X_k^4 + 1 \leq c_k^2$,

$$\begin{aligned} \left\|\sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l} \frac{(\widehat{V}_{2k} - V_{2k})^2}{d_k}\psi_k\right\| &\leq \sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l} \frac{c_k\|\psi_k\|}{d_k}\left((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2\right), \\ &\leq \sum_{l=0}^{n-1}|\mathbb{G}_l|\left((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2\right). \end{aligned}$$

However, as Λ is positive definite, we obtain from (3.5.4) that

$$\sum_{l=0}^{n-1}|\mathbb{G}_l|\left((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2\right) = \mathcal{O}(n) \qquad \text{a.s.}$$

which implies that

$$\left\|\sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l} \frac{(\widehat{V}_{2k} - V_{2k})^2}{d_k}\psi_k\right\| = \mathcal{O}(n) \qquad \text{a.s.} \tag{3.6.25}$$

Furthermore, denote

$$P_n = \sum_{l=0}^{n-1} \sum_{k\in\mathbb{G}_l} \frac{(\widehat{V}_{2k} - V_{2k})V_{2k}}{d_k}\psi_k.$$

We clearly have

$$\Delta P_{n+1} = P_{n+1} - P_n = \sum_{k\in\mathbb{G}_n} \frac{(\widehat{V}_{2k} - V_{2k})V_{2k}}{d_k}\psi_k, = -\sum_{k\in\mathbb{G}_n} \frac{V_{2k}}{d_k}\psi_k\Phi_k^t \begin{pmatrix}\widehat{a}_n - a\\ \widehat{c}_n - c\end{pmatrix}.$$

In addition, for all $k \in \mathbb{G}_n$, $\mathbb{E}[V_{2k}|\mathcal{F}_n] = 0$ a.s. and $\mathbb{E}[V_{2k}^2|\mathcal{F}_n] = \sigma_a^2 X_k^2 + \sigma_c^2 \leq \alpha c_k$ a.s. where $\alpha = \max(\sigma_a^2, \sigma_c^2)$. Consequently, $\mathbb{E}[\Delta P_{n+1}|\mathcal{F}_n] = 0$ a.s. and

$$\begin{aligned}\mathbb{E}[\Delta P_{n+1}\Delta P_{n+1}^t|\mathcal{F}_n] &= \sum_{k\in\mathbb{G}_n} \frac{1}{d_k^2}\mathbb{E}[V_{2k}^2|\mathcal{F}_n]\psi_k\Phi_k^t \begin{pmatrix}\widehat{a}_n - a\\ \widehat{c}_n - c\end{pmatrix}\begin{pmatrix}\widehat{a}_n - a\\ \widehat{c}_n - c\end{pmatrix}^t \Phi_k\psi_k^t \qquad \text{a.s.}\\ &= \sum_{k\in\mathbb{G}_n} \frac{P(X_k)}{d_k^2}\psi_k\Phi_k^t \begin{pmatrix}\widehat{a}_n - a\\ \widehat{c}_n - c\end{pmatrix}\begin{pmatrix}\widehat{a}_n - a\\ \widehat{c}_n - c\end{pmatrix}^t \Phi_k\psi_k^t \qquad \text{a.s.}\end{aligned}$$

Therefore, (P_n) is a square integrable vector martingale with increasing process $\langle P\rangle_n$ given by

$$\begin{aligned}\langle P\rangle_n &= \sum_{l=1}^{n-1} \mathbb{E}[\Delta P_{l+1}\Delta P_{l+1}^t|\mathcal{F}_l] \qquad \text{a.s.}\\ &= \sum_{l=1}^{n-1}\sum_{k\in\mathbb{G}_l} \frac{P(X_k)}{d_k^2}\psi_k\Phi_k^t \begin{pmatrix}\widehat{a}_l - a\\ \widehat{c}_l - c\end{pmatrix}\begin{pmatrix}\widehat{a}_l - a\\ \widehat{c}_l - c\end{pmatrix}^t \Phi_k\psi_k^t \qquad \text{a.s.}\end{aligned}$$

It immediately follows from the previous calculation that

$$\begin{aligned}\|\langle P\rangle_n\| &\leq \alpha \sum_{l=0}^{n-1} \left((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2\right) \sum_{k\in\mathbb{G}_l} \frac{c_k\|\psi_k\|^2\|\Phi_k\|^2}{d_k^2} \qquad \text{a.s.}\\ &\leq \alpha \sum_{l=0}^{n-1} |\mathbb{G}_l| \left((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2\right) \qquad \text{a.s.}\end{aligned}$$

leading to $\|\langle P\rangle_n\| = \mathcal{O}(n)$ a.s. Then, we deduce from the strong law of large numbers for martingale given e.g. in Theorem 1.3.15 of [23] that

$$P_n = o(n) \qquad \text{a.s.} \tag{3.6.26}$$

Hence, we find from (3.6.24), (3.6.25) and (3.6.26) that $\|Q_{n-1}(\widehat{\eta}_n - \eta_n)\| = \mathcal{O}(n)$ a.s. Moreover, we infer once again from Lemma 3.5.2 that

$$\lim_{n\to\infty} \frac{1}{|\mathbb{T}_n|}Q_n = D = \mathbb{E}\left[\frac{1}{(1+T^2)^2}\begin{pmatrix}T^4 & T^2\\ T^2 & 1\end{pmatrix}\right] \qquad \text{a.s.} \tag{3.6.27}$$

Moreover, we can prove through tedious calculations that T^2 is not degenerate which allows us to say that D is positive definite. This ensures that

$$\|\widehat{\eta}_n - \eta_n\| = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \qquad \text{a.s.}$$

It remains to establish (3.5.8). Denote

$$\widehat{W}_n = \begin{pmatrix} \widehat{V}_{2n} \\ \widehat{V}_{2n+1} \end{pmatrix} \qquad \text{and} \qquad R_n = \sum_{k\in\mathbb{T}_{n-1}} \frac{1}{d_k}\left(\widehat{W}_k - W_k\right)^t J W_k \psi_k$$

where $J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Then, we have from (3.3.9) that

$$Q_{n-1}(\widehat{\nu}_n - \nu_n) = \sum_{k\in\mathbb{T}_{n-1}} \frac{1}{d_k}\left(\widehat{V}_{2k} - V_{2k}\right)\left(\widehat{V}_{2k+1} - V_{2k+1}\right)\psi_k + R_n.$$

It is not hard to see that (R_n) is a square integrable real martingale with increasing process given by

$$\begin{aligned}
\langle R\rangle_n &= \sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l} \mathbb{E}\left[\frac{1}{d_k^2}(\widehat{W}_k - W_k)^t J W_k W_k^t J(\widehat{W}_k - W_k)\psi_k\psi_k^t \Big| \mathcal{F}_l\right] \qquad \text{a.s.}\\
&= \sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l} \frac{1}{d_k^2}(\widehat{W}_k - W_k)^t J \mathbb{E}\left[W_k W_k^t \big| \mathcal{F}_l\right] J(\widehat{W}_k - W_k)\psi_k\psi_k^t \qquad \text{a.s.}\\
&= \sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l} \frac{1}{d_k^2}(\widehat{W}_k - W_k)^t J \begin{pmatrix} P(X_k) & Q(X_k) \\ Q(X_k) & R(X_k)\end{pmatrix} J(\widehat{W}_k - W_k)\psi_k\psi_k^t \qquad \text{a.s.}\\
&= \sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l} \frac{1}{d_k^2}(\widehat{W}_k - W_k)^t \begin{pmatrix} R(X_k) & Q(X_k) \\ Q(X_k) & P(X_k)\end{pmatrix} (\widehat{W}_k - W_k)\psi_k\psi_k^t \qquad \text{a.s.}
\end{aligned}$$

Consequently, Lemma 3.5.2 together with (3.5.4) allows us to say that $\|\langle R\rangle_n\| = \mathcal{O}(n)$ a.s. which ensures that $R_n = o(n)$ a.s. Moreover,

$$\begin{aligned}
&\left\|\sum_{k\in\mathbb{T}_{n-1}} \frac{1}{d_k}\left(\widehat{V}_{2k} - V_{2k}\right)\left(\widehat{V}_{2k+1} - V_{2k+1}\right)\psi_k\right\| \\
&\qquad \leq \frac{1}{2}\sum_{k\in\mathbb{T}_{n-1}} \frac{1}{d_k}\left(\left(\widehat{V}_{2k} - V_{2k}\right)^2 + \left(\widehat{V}_{2k+1} - V_{2k+1}\right)^2\right)\|\psi_k\|,\\
&\qquad \leq \frac{1}{2}\sum_{l=0}^{n-1}\|\widehat{\theta}_l - \theta\|^2 \sum_{k\in\mathbb{G}_l}\frac{\|\Phi_k\|^2\|\psi_k\|}{d_k},\\
&\qquad \leq \frac{1}{2}\sum_{l=0}^{n-1}|\mathbb{G}_l|\|\widehat{\theta}_l - \theta\|^2,
\end{aligned}$$

which implies via (3.5.4) that

$$\left\| \sum_{k\in\mathbb{T}_{n-1}} \left(\widehat{V}_{2k} - V_{2k}\right)\left(\widehat{V}_{2k+1} - V_{2k+1}\right)\psi_k \right\| = \mathcal{O}(n) \qquad \text{a.s.}$$

Therefore, we obtain that $\|Q_{n-1}(\widehat{\nu}_n - \nu_n)\| = \mathcal{O}(n)$ a.s. which leads to (3.5.8). Finally, it only remains to prove the a.s. convergence of η_n, ζ_n and ν_n to η, ζ and ν which will immediately lead to the a.s. convergence of $\widehat{\eta}_n$, $\widehat{\zeta}_n$ and $\widehat{\nu}_n$ through (3.5.6), (3.5.7) and (3.5.8), respectively. On the one hand,

$$Q_{n-1}(\eta_n - \eta) = N_n = \sum_{k\in\mathbb{T}_{n-1}} \frac{1}{d_k} v_{2k}\psi_k \tag{3.6.28}$$

where we recall that $v_{2n} = V_{2n}^2 - \eta^t\psi_n$. It is clear that (N_n) is a square integrable vector martingale with increasing process $\langle N\rangle_n$ given by

$$\begin{aligned}\langle N\rangle_n &= \sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l} \frac{1}{d_k^2}\mathbb{E}[v_{2k}^2|\mathcal{F}_l]\psi_k\psi_k^t \qquad \text{a.s.}\\ &\le \gamma \sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l} \frac{1}{d_k}\psi_k\psi_k^t \qquad \text{a.s.}\end{aligned}$$

where $\gamma = \max(\mu_a^4 - \sigma_a^4, 2\sigma_a^2\sigma_c^2, \mu_c^4 - \sigma_c^4)$. Hence,

$$\|\langle N\rangle_n\| \le \gamma \sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l} \frac{1}{d_k}\|\psi_k\|^2 \le \gamma|\mathbb{T}_{n-1}| \qquad \text{a.s.}$$

which immediately leads to $\|\langle N\rangle_n\| = \mathcal{O}(|\mathbb{T}_{n-1}|)$ a.s. Consequently, $\|N_n\|^2 = \mathcal{O}(n|\mathbb{T}_{n-1}|)$ a.s. which leads via (3.6.27) and (3.6.28) to the a.s. convergence of η_n to η and to the rate of convergence of Remark 3.5.6. The proof of the a.s. convergence of ζ_n to ζ follows exactly the same lines. On the other hand

$$Q_{n-1}(\nu_n - \nu) = H_n = \sum_{k\in\mathbb{T}_{n-1}} \frac{1}{d_k} w_{2k}\psi_k \tag{3.6.29}$$

where we recall that $w_{2k} = V_{2k}V_{2k+1} - \mathbb{E}[V_{2k}V_{2k+1}|\mathcal{F}_n]$. It is obvious to see that (H_n) is a square integrable real martingale with increasing process

$$\begin{aligned}\langle H\rangle_n &= \sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l} \frac{1}{d_k^2}\mathbb{E}[w_{2k}^2|\mathcal{F}_l]\psi_k\psi_k^t \qquad \text{a.s.}\\ &\le \alpha \sum_{l=0}^{n-1}\sum_{k\in\mathbb{G}_l} \frac{1}{d_k}\psi_k\psi_k^t \qquad \text{a.s.}\end{aligned}$$

where $\alpha = \max\left(\nu_{ab}^2, \nu_{cd}^2, (\sigma_a^2+\sigma_c^2)(\sigma_b^2+\sigma_d^2)\right)$. This implies that

$$\|\langle H\rangle_n\| \leq \alpha \sum_{l=0}^{n-1} \sum_{k\in\mathbb{G}_l} \frac{1}{d_k}\|\psi_k\|^2 \leq \alpha|\mathbb{T}_{n-1}| \qquad \text{a.s.}$$

which allows us to say that

$$\|H_n\|^2 = \mathcal{O}(n|\mathbb{T}_{n-1}|) \qquad \text{and} \qquad \|\widehat{\nu}_n - \nu\|^2 = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \qquad \text{a.s.}$$

Finally, we deduce from (3.6.29) that ν_n converges a.s. to ν and that the rate of convergence of Remark 3.5.6 is verified, which completes the proof of Theorem 3.5.5.

3.6.7 Proof of the asymptotic normalities

In order to establish the asymptotic normality of our estimators, we will extensively make use of the central limit theorem for triangular arrays of vector martingales given e.g. by Theorem 2.1.9 of [23]. First of all, instead of using the generation-wise filtration $(\mathcal{F}_n)$, we will use the sister pair-wise filtration $(\mathcal{G}_n)$ given by

$$\mathcal{G}_n = \sigma(X_1, (X_{2k}, X_{2k+1}), 1 \leq k \leq n).$$

Proof of Theorem 3.5.7, first part. We focus our attention to the proof of the asymptotic normality (3.5.9). Let $M^{(n)} = (M_k^{(n)})$ be the square integrable vector martingale defined as

$$M_k^{(n)} = \frac{1}{\sqrt{|\mathbb{T}_n|}} \sum_{i=1}^{k} D_i, \qquad \text{where} \qquad D_i = \frac{1}{c_i}\begin{pmatrix} X_iV_{2i} \\ V_{2i} \\ X_iV_{2i+1} \\ V_{2i+1}\end{pmatrix}. \tag{3.6.30}$$

We clearly have

$$M_{t_n}^{(n)} = \frac{1}{\sqrt{|\mathbb{T}_n|}} \sum_{i=1}^{t_n} D_i = \frac{1}{\sqrt{|\mathbb{T}_n|}} M_{n+1}, \tag{3.6.31}$$

where $t_n = |\mathbb{T}_n|$. Moreover, the increasing process associated to $(M_k^{(n)})$ is given by

$$\begin{aligned}\langle M^{(n)}\rangle_k &= \frac{1}{|\mathbb{T}_n|}\sum_{i=1}^{k} \mathbb{E}\left[D_iD_i^t|\mathcal{G}_{i-1}\right], \\ &= \frac{1}{|\mathbb{T}_n|}\sum_{i=1}^{k}\frac{1}{c_i^2}\begin{pmatrix} P(X_i) & Q(X_i) \\ Q(X_i) & R(X_i)\end{pmatrix} \otimes \begin{pmatrix} X_i^2 & X_i \\ X_i & 1\end{pmatrix} \qquad \text{a.s.}\end{aligned}$$

Consequently, it follows from convergence (3.5.3) that

$$\lim_{n\to\infty} \langle M^{(n)}\rangle_{t_n} = L \qquad \text{a.s.}$$

It is now necessary to verify Lindeberg's condition by use of Lyapunov's condition. Denote

$$\phi_n = \sum_{k=1}^{t_n} \mathbb{E}\left[\|M_k^{(n)} - M_{k-1}^{(n)}\|^4 \Big| \mathcal{G}_{k-1} \right].$$

We obtain from (3.6.30) that

$$\begin{aligned}\phi_n &= \frac{1}{|\mathbb{T}_n|^2} \sum_{k=1}^{t_n} \mathbb{E}\left[\frac{(1+X_k^2)^2}{c_k^4} (V_{2k}^2 + V_{2k+1}^2)^2 \Big| \mathcal{G}_{k-1} \right], \\ &\leq \frac{2}{|\mathbb{T}_n|^2} \sum_{k=1}^{t_n} \frac{1}{c_k^2} \left(\mathbb{E}[V_{2k}^4 | \mathcal{G}_{k-1}] + \mathbb{E}[V_{2k+1}^4 | \mathcal{G}_{k-1}] \right).\end{aligned}$$

In addition, we already saw in Section 3.6.4 that

$$\mathbb{E}[V_{2n}^4 | \mathcal{G}_{n-1}] \leq \mu_{ac}^4 c_n^2, \qquad \mathbb{E}[V_{2n+1}^4 | \mathcal{G}_{n-1}] \leq \mu_{bd}^4 c_n^2 \qquad \text{a.s.}$$

where $\mu_{ac}^4 = \max(\mu_a^4, 3\sigma_a^2\sigma_c^2, \mu_c^4)$ and $\mu_{bd}^4 = \max(\mu_b^4, 3\sigma_b^2\sigma_d^2, \mu_d^4)$. Hence,

$$\phi_n \leq \frac{2(\mu_{ac}^4 + \mu_{bd}^4)}{|\mathbb{T}_n|} \qquad \text{a.s.}$$

which immediately implies that

$$\lim_{n\to\infty} \phi_n = 0 \qquad \text{a.s.}$$

Therefore, Lyapunov's condition is satisfied and Theorem 2.1.9 of [23] allows us to say via (3.6.31) that

$$\frac{1}{\sqrt{|\mathbb{T}_{n-1}|}} M_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, L).$$

Finally, we infer from (3.4.1) together with (3.6.17) and Slutsky's lemma that

$$\sqrt{|\mathbb{T}_{n-1}|}(\widehat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Lambda^{-1} L \Lambda^{-1}). \qquad \square$$

Proof of Theorem 3.5.7, second part. We shall now establish the asymptotic normality given by (3.5.10). Denote by $N^{(n)} = (N_k^{(n)})$ the square integrable vector martingale defined as

$$N_k^{(n)} = \frac{1}{\sqrt{|\mathbb{T}_n|}} \sum_{i=1}^{k} \frac{v_{2i}}{d_i} \psi_i.$$

We immediately see from (3.6.28) that

$$N_{t_n}^{(n)} = \frac{1}{\sqrt{|\mathbb{T}_n|}} Q_n(\eta_{n+1} - \eta) = \frac{1}{\sqrt{|\mathbb{T}_n|}} N_{n+1}. \tag{3.6.32}$$

In addition, the increasing process associated to $(N_k^{(n)})$ is given by

$$\langle N^{(n)}\rangle_k = \frac{1}{|\mathbb{T}_n|}\sum_{i=1}^{k}\mathbb{E}\left[\frac{v_{2i}^2}{d_i^2}\psi_i\psi_i^t\middle|\,\mathcal{G}_{i-1}\right],$$
$$= \frac{1}{|\mathbb{T}_n|}\sum_{i-1}^{k}\frac{(\mu_a^4-\sigma_a^4)X_i^4+4\sigma_a^2\sigma_c^2X_i^2+(\mu_c^4-\sigma_c^4)}{d_i^2}\psi_i\psi_i^t \qquad \text{a.s.}$$

Consequently, we obtain from Lemma 3.5.2 that

$$\lim_{n\to\infty}\langle N^{(n)}\rangle_{t_n} = \mathbb{E}\left[\frac{(\mu_a^4-\sigma_a^4)T^4+4\sigma_a^2\sigma_c^2T^2+(\mu_c^4-\sigma_c^4)}{(1+T^2)^4}\begin{pmatrix}T^4 & T^2\\ T^2 & 1\end{pmatrix}\right] = M_{ac} \quad \text{a.s.}$$

In order to verify Lyapunov's condition, let $\alpha > 4$ be the constant in **(H.5)** and let

$$\phi_n = \sum_{k=1}^{t_n}\mathbb{E}\left[\|N_k^{(n)}-N_{k-1}^{(n)}\|^{\alpha/2}\middle|\,\mathcal{G}_{k-1}\right].$$

We clearly have

$$\|N_k^{(n)}-N_{k-1}^{(n)}\|^2 = \frac{1}{|\mathbb{T}_n|}\frac{v_{2k}^2}{d_k^2}\|\psi_k\|^2 \le \frac{1}{|\mathbb{T}_n|}\frac{v_{2k}^2}{d_k},$$

which implies that

$$\|N_k^{(n)}-N_{k-1}^{(n)}\|^{\alpha/2} \le \frac{1}{|\mathbb{T}_n|^{\alpha/4}}\frac{|v_{2k}|^{\alpha/2}}{d_k^{\alpha/4}}.$$

However, it exists a constant $\beta > 0$ such that

$$|v_{2k}|^{\alpha/2} = |V_{2k}^2-\sigma_a^2X_k^2-\sigma_c^2|^{\alpha/2} \le \beta(|V_{2k}|^{\alpha}+(\sigma_a^2X_k^2+\sigma_c^2)^{\alpha/2}). \tag{3.6.33}$$

Moreover, we also have

$$|V_{2k}|^{\alpha} \le \beta(|a_k-a|^{\alpha}|X_k|^{\alpha}+|\varepsilon_{2k}-c|^{\alpha}).$$

Let

$$Y = \max\left(\sup_{n\ge0}\sup_{k\in\mathbb{G}_n}\mathbb{E}[|a_k-a|^{\alpha}|\mathcal{F}_n],\sup_{n\ge0}\sup_{k\in\mathbb{G}_n}\mathbb{E}[|\varepsilon_{2k}-c|^{\alpha}|\mathcal{F}_n]\right),$$

then it exists some constant $\gamma > 0$ such that

$$\mathbb{E}[|V_{2k}|^{\alpha}|\mathcal{G}_{k-1}] \le \beta Y(1+|X_k|^{\alpha}) \le \gamma Y(1+X_k^2)^{\alpha/2} \qquad \text{a.s.}$$

This, together with (3.6.33), ensures the existence of a constant $\delta > 0$ such that

$$\mathbb{E}[|v_{2k}|^{\alpha/2}|\mathcal{G}_{k-1}] \le \delta Y(1+X_k^2)^{\alpha/2} \qquad \text{a.s.}$$

implying that

$$\mathbb{E}\left[\|N_k^{(n)} - N_{k-1}^{(n)}\|^{\alpha/2}\Big|\,\mathcal{G}_{k-1}\right] \leq \frac{\delta Y}{|\mathbb{T}_n|^{\alpha/4}} \qquad \text{a.s.}$$

Then we can conclude that

$$\phi_n \leq \frac{\delta Y}{|\mathbb{T}_n|^{\alpha/4-1}} \qquad \text{a.s.}$$

which immediately leads, since $Y < \infty$ a.s., to

$$\lim_{n\to\infty} \phi_n = 0 \qquad \text{a.s.}$$

Therefore, Lyapunov's condition is satisfied and we find from Theorem 2.1.9 of [23] and (3.6.32) that

$$\frac{1}{\sqrt{|\mathbb{T}_{n-1}|}} N_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, M_{ac}). \tag{3.6.34}$$

Hence, we obtain from (3.6.27), (3.6.34) and Slutsky's lemma that

$$\sqrt{|\mathbb{T}_{n-1}|}(\eta_n - \eta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, D^{-1}M_{ac}D^{-1}).$$

Finally, (3.5.6) ensures that

$$\sqrt{|\mathbb{T}_{n-1}|}(\widehat{\eta}_n - \eta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, D^{-1}M_{ac}D^{-1}).$$

The proof of (3.5.11) follows exactly the same lines. □

Proof of Theorem 3.5.7, third part. It remains to establish the asymptotic normality given by (3.5.12). Denote by $H^{(n)} = (H_k^{(n)})$ the square integrable martingale defined as

$$H_k^{(n)} = \frac{1}{\sqrt{|\mathbb{T}_n|}} \sum_{i=1}^{k} \frac{w_{2i}}{d_i} \psi_i.$$

We clearly have

$$H_{t_n}^{(n)} = \frac{1}{\sqrt{|\mathbb{T}_n|}} \sum_{i=1}^{t_n} \frac{w_{2i}}{d_i} \psi_i = \frac{1}{\sqrt{|\mathbb{T}_n|}} H_{n+1}.$$

Moreover, the increasing process of $(H_k^{(n)})$ is given by

$$\langle H^{(n)} \rangle_k = \frac{1}{|\mathbb{T}_n|} \sum_{i=1}^{k} \frac{\mathbb{E}[w_{2i}^2|\mathcal{G}_{i-1}]\psi_i\psi_i^t}{d_i^2}.$$

In addition, we already saw in Section 3.3 that

$$\mathbb{E}[w_{2k}^2|\mathcal{F}_n] = (\nu_{ab}^2 - \rho_{ab}^2)X_k^4 + (\sigma_a^2\sigma_d^2 + \sigma_b^2\sigma_c^2 + 2\rho_{ab}\rho_{cd})X_k^2 + (\nu_{cd}^2 - \rho_{cd}^2) \qquad \text{a.s.}$$

Then, we deduce once again from Lemma 3.5.2 that

$$\begin{aligned}
&\lim_{n\to\infty} \langle H^{(n)}\rangle_{t_n} \\
&\quad = \mathbb{E}\left[\frac{(\nu_{ab}^2-\rho_{ab}^2)T^4 + (\sigma_a^2\sigma_d^2+\sigma_b^2\sigma_c^2+2\rho_{ab}\rho_{cd})T^2 + (\nu_{cd}^2-\rho_{cd}^2)}{(1+T^2)^4}\begin{pmatrix} T^4 & T^2 \\ T^2 & 1\end{pmatrix}\right] \\
&\quad = H \qquad \text{a.s.}
\end{aligned}$$

In order to verify Lyapunov's condition, denote, with $\alpha > 4$ the constant in **(H.5)**,

$$\phi_n = \sum_{k=1}^{t_n} \mathbb{E}\left[\|H_k^{(n)} - H_{k-1}^{(n)}\|^{\alpha/2} \middle| \mathcal{G}_{k-1}\right].$$

As in the previous proof, we clearly have that

$$\|H_k^{(n)} - H_{k-1}^{(n)}\|^{\alpha/2} \leq \frac{1}{|\mathbb{T}_n|^{\alpha/4}} \frac{|w_{2k}|^{\alpha/2}}{d_k^{\alpha/4}}.$$

We can observe that it exists some constants $\beta > 0$ and $\gamma > 0$ such that

$$\begin{aligned}
|w_{2k}|^{\alpha/2} &= |V_{2k}V_{2k+1} - \rho_{ab}X_k^2 - \rho_{cd}|^{\alpha/2} \leq \big(|V_{2k}V_{2k+1}| + |\rho_{ab}|X_k^2 + |\rho_{cd}|\big)^{\alpha/2}, \\
&\leq \beta(|V_{2k}V_{2k+1}|^{\alpha/2} + (|\rho_{ab}|X_k^2 + |\rho_{cd}|)^{\alpha/2}), \\
&\leq \gamma(|V_{2k}|^{\alpha} + |V_{2k+1}|^{\alpha} + (|\rho_{ab}|X_k^2 + |\rho_{cd}|)^{\alpha/2}).
\end{aligned}$$

Hence, in the same way as in the proof of the second part, we can prove that it exists a constant $\delta > 0$ and a random variable Y such that $Y < \infty$ a.s. verifying

$$\mathbb{E}[|w_k|^{\alpha/2}|\mathcal{G}_{k-1}] \leq \delta Y(1+X_k^2)^{\alpha/2} \qquad \text{a.s.}$$

which immediately leads to

$$\mathbb{E}[\|H_k^{(n)} - H_{k-1}^{(n)}\|^{\alpha/2}|\mathcal{G}_{k-1}] \leq \frac{\delta Y}{|\mathbb{T}_n|^{\alpha/4}} \qquad \text{a.s.}$$

which ensures that

$$\phi_n \leq \frac{\delta Y}{|\mathbb{T}_n|^{\alpha/4-1}} \qquad \text{a.s.}$$

Then, we obviously have that

$$\lim_{n\to\infty} \phi_n = 0 \qquad \text{a.s.}$$

and we can conclude that

$$\frac{1}{\sqrt{|\mathbb{T}_{n-1}|}} H_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, H).$$

In other words

$$\sqrt{|\mathbb{T}_{n-1}|}(\nu_n - \nu) \xrightarrow{\mathcal{L}} \mathcal{N}(0, D^{-1}HD^{-1}).$$

Finally, we find via (3.5.8) that

$$\sqrt{|\mathbb{T}_{n-1}|}(\widehat{\nu}_n - \nu) \xrightarrow{\mathcal{L}} \mathcal{N}(0, D^{-1}HD^{-1})$$

which achieves the proof of Theorem 3.5.7. □

3.7 Numerical simulations

The goal of this section is to illustrate by simulations the main results of this paper. In order to keep this section brief, we shall only focus our attention on the asymptotic normality of the WLS estimator of the unknown parameter θ. On the one hand the random coefficient sequence (a_n, b_n) is chosen to be i.i.d sharing the same distribution as $(X+Y, X+Z)$ where $X \sim \mathcal{N}(0.5, 0.4)$, $Y \sim \mathcal{N}(0, 0.3)$ and $Z \sim \mathcal{N}(-0.2, 0.4)$. Those parameters have been chosen in order to satisfy **(H.1)**. On the other hand, the driven noise sequence $(\varepsilon_{2n}, \varepsilon_{2n+1})$ is chosen to be i.i.d. sharing the same distribution as $(U+V, U+W)$ where $U \sim \mathcal{E}(1)$, $V \sim \mathcal{E}(2)$ and $W \sim \mathcal{E}(3)$ and $\mathcal{E}(\lambda)$ stands for the exponential distribution with parameter $\lambda > 0$. The histograms are made by computing 4000 times $\widehat{\theta}_n$ with $n = 13$, and the variances of the theoretical normal distributions, which are plotted with the red curve, have been estimated by a Monte-Carlo procedure. One can observe in Figure 3.2 that the WLS estimator $\widehat{\theta}_n$ performs very well in the estimation of θ.

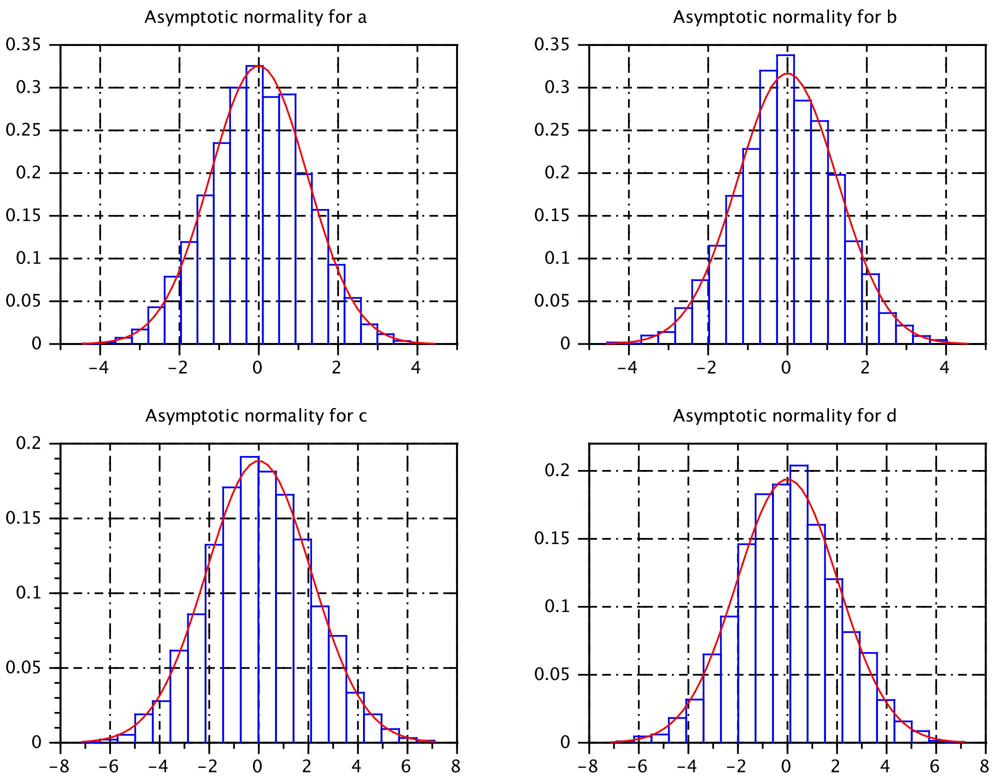

FIGURE 3.2 – Illustration of the asymptotic normalities of a, b, c and d.

Chapitre 4

Une approche Rademacher-Menchov pour les processus BAR à coefficients aléatoires

Résumé. Nous étudions le comportement asymptotique des estimateurs des moindres carrés des paramètres inconnus des processus autorégressifs à bifurcation et à coefficients aléatoires. Sous de bonnes hypothèses sur l'immigration et sur l'environnement, nous établissons la convergence presque sure de nos estimateurs. De plus, nous montrons également une loi forte quadratique et des théorèmes centraux limites. Notre approche repose essentiellement sur des résultats asymptotiques pour les martingales vectorielles ainsi que sur le théorème de Rademacher-Menchov.

Abstract. We investigate the asymptotic behavior of the least squares estimator of the unknown parameters of random coefficient bifurcating autoregressive processes. Under suitable assumptions on inherited and environmental effects, we establish the almost sure convergence of our estimates. In addition, we also prove a quadratic strong law and central limit theorems. Our approach mainly relies on asymptotic results for vector-valued martingales together with the well-known Rademacher-Menchov theorem.

4.1 Introduction

The purpose of this paper is to study random coefficient bifurcating autoregressive processes (RCBAR). One can see those processes in two different ways. The first one is to see them as random coefficient autoregressive processes (RCAR) adapted to binary tree structured data, the second one is to consider those processes as the association of RCAR processes and bifurcating autoregressive processes (BAR). BAR processes have been first studied by Cowan and Staudte [16] while RCAR processes have been first investigated by Nicholls and Quinn [44, 46]. The RCBAR structure allows us to reckon with environmental and inherited effects in order to better take into account the evolution of the characteristic under study. One shall see cell division as an example of binary tree structured data.

Let us detail what a RCBAR process is. The first individual is designated as the individual 1 and each individual n leads to individuals $2n$ and $2n+1$. X_n will stand for the characteristic under study of individual n. We can now make explicit the first-order RCBAR process which is given, for all $n \geq 1$, by

$$\begin{cases} X_{2n} = a_n X_n + \varepsilon_{2n}, \\ X_{2n+1} = b_n X_n + \varepsilon_{2n+1}. \end{cases}$$

The driven noise sequence $(\varepsilon_{2n}, \varepsilon_{2n+1})$ represents the environmental effect while the random coefficient sequence (a_n, b_n) represents the inherited effect. Keeping in mind the example of cell division, we assume that ε_{2n} and ε_{2n+1} are correlated in order to take into account the environmental effect on two sister cells.

This study is inspired by experiments on the single celled organism *Escherichia coli*, see Stewart et al. [51] or Guyon et al. [27], which reproduces by dividing itself into two poles, one being called the new pole, the other being called the old pole. Experimental data seems to show that some variables among cell lines, such as the life span of the cells, does not evolve in the same way whether it is related to the new or the old pole. The difference in the evolution leads us to consider an asymmetric RCBAR. Considering a RCBAR process instead of a BAR process allows us to assume that the inherited effect is no more deterministic, as randomness often appears in nature. Moreover, we can consider both deterministic and random inherited effects since we also allow the random variables modeling the inherited effect to be deterministic, making this study usable for RCBAR as well as BAR.

Our goal is to study the asymptotic behavior of the least squares estimators of the unknown parameters of first-order RCBAR processes. In contrast with the previous work of Blandin [13] where the asymptotic behavior of weighted least squares estimators were investigated, we propose here to make use of a totally different strategy based on the standard least squares (LS) estimators together with the well-known Rademacher-Menchov theorem. The martingale approach for BAR processes has been first suggested by Bercu et al. [11], followed by the recent contribution of de Saporta et al. [17, 19]. We also refer the reader to Blandin [12] for the study of bifurcating integer-valued autoregressive processes. Our approach relies on the Rademacher-Menchov theorem which allows us to study the LS estimates in a different way as in de Saporta et al. [19]. In particular, we reduce the moment assumptions, from 32th-order in [19] to 16th-order in this study, on the random coefficient sequence (a_n, b_n) and on the driven noise sequence $(\varepsilon_{2n}, \varepsilon_{2n+1})$. We shall also make use of the strong law of large numbers and the central limit theorem for martingales [23, 28] in order to study the asymptotic behavior of our LS estimates. The martingale approach of this paper has also been used by Basawa and Zhou [9, 60, 61].

Since several methods have been proposed for the study of BAR processes, we tried to take into consideration each of them. In this way, we took into account the classical BAR approach as used by Huggins and Basawa [31, 32] and by Huggins and Staudte [33] who investigated the evolution of cell diameters and lifetimes. We were also inspired by the bifurcating Markov chain approach brought in by Guyon [26] and applied by Delmas and Marsalle [20]. We also reckoned with the analogy with the Galton-Watson processes as in Delmas and Marsalle [20] and Heyde and Seneta [29]. Even though we chose to use LS estimates, different methods have been investigated for parameter estimation in RCAR processes. While Koul and Schick [39] used an M-estimator, Aue et al. [3] tackled a quasi-maximum likelihood approach. Vanecek [56] used an estimator first introduced by Schick [49]. On their side, Hwag et al. [35] studied the critical case where the environmental effect follows a Rademacher distribution.

The paper is organized as follows. We will explain more accurately the model we will consider in Section 2, leading to Section 3 where we will give explicitly our LS estimates of the unknown parameters under study. The martingale point of view chosen in this paper will be highlighted in Section 4. All our results about the asymptotic behavior of our LS estimates will be stated in Section 5, in particular the almost sure convergence, the quadratic strong law and the asymptotic normality. Section 6 is devoted to the Rademacher-Menchov theorem. All technical proofs are postponed to the last sections. We conclude with a short section illustrating our results on numerical simulations.

4.2 Random coefficient bifurcating autoregressive processes

We will study the first-order RCBAR process given, for all $n \geq 1$, by

$$\begin{cases} X_{2n} &= a_n X_n + \varepsilon_{2n}, \\ X_{2n+1} &= b_n X_n + \varepsilon_{2n+1}, \end{cases} \tag{4.2.1}$$

where X_1 is the ancestor of the process and $(\varepsilon_{2n}, \varepsilon_{2n+1})$ is the driven noise of the process. We will suppose that $\mathbb{E}[X_1^{16}] < \infty$ and we will also assume that the two sequences (a_n, b_n) and $(\varepsilon_{2n}, \varepsilon_{2n+1})$ are independent and identically distributed and that X_1, (a_n, b_n) and $(\varepsilon_{2n}, \varepsilon_{2n+1})$ are mutually independent. RCBAR processes can be seen as a first-order random coefficient autoregressive process on a binary tree, each node of this tree representing an individual and the first node being the ancestor. For all $n \geq 0$, $\mathbb{G}_n$ will stand for the n-th generation, that is to say $\mathbb{G}_n = \{2^n, 2^n + 1, \ldots, 2^{n+1} - 1\}$. We will also denote by $\mathbb{T}_n$ the set of all individuals up to the n-th generation, namely

$$\mathbb{T}_n = \bigcup_{k=0}^{n} \mathbb{G}_k.$$

One can see that the cardinality $|\mathbb{G}_n|$ of $\mathbb{G}_n$ is 2^n, while that of $\mathbb{T}_n$ is $2^{n+1}-1$. $\mathbb{G}_{r_n}$ will denote the generation of individual n with $r_n = [\log_2(n)]$ where $[x]$ stands for the integer part of x. Let us recall that the two offspring of individual n are individuals $2n$ and $2n+1$.

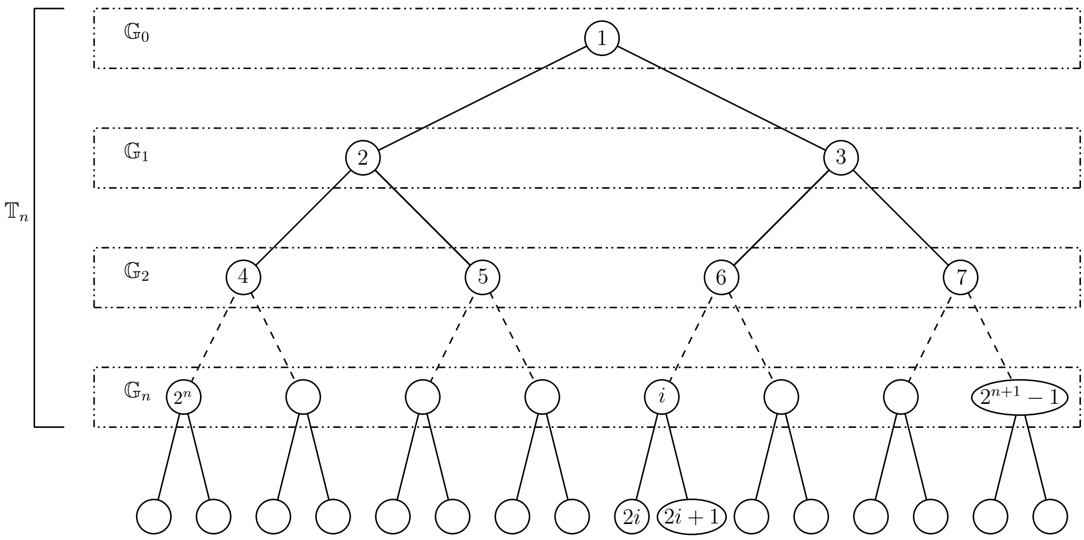

FIGURE 4.1 – The tree associated with the RCBAR

4.3 Least squares estimators

Let $(\mathcal{F}_n)$ be the natural filtration associated with the generations of our first-order RCBAR (X_n), namely $\mathcal{F}_n = \sigma\{X_k, k \in \mathbb{T}_n\}$ for all $n \in \mathbb{N}$. In all the sequel, we will assume that for all $n \geq 0$ and for all $k \in \mathbb{G}_n$,

$$\begin{aligned} \mathbb{E}[a_k|\mathcal{F}_n] = a, \qquad \mathbb{E}[b_k|\mathcal{F}_n] = b, \\ \mathbb{E}[\varepsilon_{2k}|\mathcal{F}_n] = c, \qquad \mathbb{E}[\varepsilon_{2k+1}|\mathcal{F}_n] = d \qquad \text{a.s.} \end{aligned} \tag{4.3.1}$$

Consequently, (4.2.1) can be rewritten as

$$\begin{cases} X_{2n} & = aX_n + c + V_{2n}, \\ X_{2n+1} & = bX_n + d + V_{2n+1}, \end{cases} \tag{4.3.2}$$

where, for all $k \in \mathbb{G}_n$, $V_{2k} = X_{2k} - E[X_{2k}|\mathcal{F}_n]$ and $V_{2k+1} = X_{2k+1} - E[X_{2k+1}|\mathcal{F}_n]$. We can rewrite the system (4.3.2) in a classic autoregressive form

$$\chi_n = \theta^t \Phi_n + W_n \tag{4.3.3}$$

where

$$\chi_n = \begin{pmatrix} X_{2n} \\ X_{2n+1} \end{pmatrix}, \qquad \Phi_n = \begin{pmatrix} X_n \\ 1 \end{pmatrix}, \qquad W_n = \begin{pmatrix} V_{2n} \\ V_{2n+1} \end{pmatrix},$$

and the matrix parameter θ given by

$$\theta = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

One of our goal is to estimate θ from the observation of the $n+1$ first generations, namely $\mathbb{T}_n$. We will use the least squares estimator $\widehat{\theta}_n$ of θ which minimizes

$$\Delta_n(\theta) = \sum_{k \in \mathbb{T}_{n-1}} \|\chi_k - \theta^t \Phi_k\|^2.$$

Hence, we clearly have

$$\widehat{\theta}_n = S_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \Phi_k \chi_k^t \qquad \text{where} \qquad S_n = \sum_{k \in \mathbb{T}_n} \Phi_k \Phi_k^t.$$

In order to avoid any invertibility assumption, we will suppose that S_1 is invertible. Otherwise, we only have to add the identity matrix of order 2, I_2, to S_n. Since, in a certain way, S_n goes to infinty , it will not change our results. Moreover, we will make a slight abuse of notation by identifying θ and $\widehat{\theta}_n$ to

$$\operatorname{vec}(\theta) = \begin{pmatrix} a \\ c \\ b \\ d \end{pmatrix} \qquad \text{and} \qquad \operatorname{vec}(\widehat{\theta}_n) = \begin{pmatrix} \widehat{a}_n \\ \widehat{c}_n \\ \widehat{b}_n \\ \widehat{d}_n \end{pmatrix}.$$

In this vectorial form, we have

$$\widehat{\theta}_n = \Sigma_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \begin{pmatrix} X_k X_{2k} \\ X_{2k} \\ X_k X_{2k+1} \\ X_{2k+1} \end{pmatrix},$$

where $\Sigma_n = I_2 \otimes S_n$ and $\otimes$ stands for the standard Kronecker product. Hence, (4.3.3) yields to

$$\widehat{\theta}_n - \theta = \Sigma_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \begin{pmatrix} X_k V_{2k} \\ V_{2k} \\ X_k V_{2k+1} \\ V_{2k+1} \end{pmatrix}. \tag{4.3.4}$$

In all this paper, we will make use of the following hypotheses on the moments of the random coefficient sequence (a_n, b_n) and on the driven noise sequence $(\varepsilon_{2n}, \varepsilon_{2n+1})$. One can observe that for all $n \geq 0$ and for all $k \in \mathbb{G}_n$, the random coefficients a_k, b_k and the driven noise ε_{2k}, ε_{2k+1} are $\mathcal{F}_{n+1}$-measurable.

(H.1) For all $n \geq 1$,

$$\mathbb{E}[a_n^{16}] < 1 \quad \text{and} \quad \mathbb{E}[b_n^{16}] < 1,$$
$$\sup_{n\geq 1} \mathbb{E}[\varepsilon_{2n}^{16}] < \infty \quad \text{and} \quad \sup_{n\geq 1} \mathbb{E}[\varepsilon_{2n+1}^{16}] < \infty.$$

(H.2) For all $n \geq 0$ and for all $k \in \mathbb{G}_n$

$$\text{Var}[a_k|\mathcal{F}_n] = \sigma_a^2 \geq 0 \quad \text{and} \quad \text{Var}[b_k|\mathcal{F}_n] = \sigma_b^2 \geq 0 \qquad \text{a.s.}$$
$$\text{Var}[\varepsilon_{2k}|\mathcal{F}_n] = \sigma_c^2 > 0 \quad \text{and} \quad \text{Var}[\varepsilon_{2k+1}|\mathcal{F}_n] = \sigma_d^2 > 0 \qquad \text{a.s.}$$

(H.3) It exists $\rho_{ab}^2 \leq \sigma_a^2 \sigma_b^2$ and $\rho_{cd}^2 < \sigma_c^2 \sigma_d^2$ such that for all $n \geq 0$ and for all $k \in \mathbb{G}_n$

$$\mathbb{E}[(a_k - a)(b_k - b)|\mathcal{F}_n] = \rho_{ab} \qquad \text{a.s.}$$
$$\mathbb{E}[(\varepsilon_{2k} - c)(\varepsilon_{2k+1} - d)|\mathcal{F}_n] = \rho_{cd} \qquad \text{a.s.}$$

Moreover, for all $n \geq 0$ and $k, l \in \mathbb{G}_n$ with $k \neq l$, $(\varepsilon_{2k}, \varepsilon_{2k+1})$ and $(\varepsilon_{2l}, \varepsilon_{2l+1})$ as well as (a_k, b_k) and (a_l, b_l) are conditionally independent given $\mathcal{F}_n$.

(H.4) One can find $\mu_a^4 \geq \sigma_a^4$, $\mu_b^4 \geq \sigma_b^4$, $\mu_c^4 > \sigma_c^4$ and $\mu_d^4 > \sigma_d^4$ such that, for all $n \geq 0$ and for all $k \in \mathbb{G}_n$

$$\mathbb{E}\left[(a_k - a)^4 |\mathcal{F}_n\right] = \mu_a^4 \quad \text{and} \quad \mathbb{E}\left[(b_k - b)^4 |\mathcal{F}_n\right] = \mu_b^4 \qquad \text{a.s.}$$
$$\mathbb{E}\left[(\varepsilon_{2k} - c)^4 |\mathcal{F}_n\right] = \mu_c^4 \quad \text{and} \quad \mathbb{E}\left[(\varepsilon_{2k+1} - d)^4 |\mathcal{F}_n\right] = \mu_d^4 \qquad \text{a.s.}$$
$$\mathbb{E}[\varepsilon_{2k}^4] > \mathbb{E}[\varepsilon_{2k}^2]^2 \quad \text{and} \quad \mathbb{E}[\varepsilon_{2k+1}^4] > \mathbb{E}[\varepsilon_{2k+1}^2]^2.$$

In addition, it exists $\nu_{ab}^2 \geq \rho_{ab}^2$ and $\nu_{cd}^2 > \rho_{cd}^2$ such that, for all $k \in \mathbb{G}_n$

$$\mathbb{E}[(a_k - a)^2 (b_k - b)^2 |\mathcal{F}_n] = \nu_{ab}^2 \quad \text{and} \quad \mathbb{E}[(\varepsilon_{2k} - c)^2 (\varepsilon_{2k+1} - d)^2 |\mathcal{F}_n] = \nu_{cd}^2 \qquad \text{a.s.}$$

(H.5) There exists some $\alpha > 4$ such that

$$\sup_{n\geq 0} \sup_{k \in \mathbb{G}_n} \mathbb{E}[|a_k - a|^\alpha |\mathcal{F}_n] < \infty, \qquad \sup_{n\geq 0} \sup_{k \in \mathbb{G}_n} \mathbb{E}[|b_k - b|^\alpha |\mathcal{F}_n] < \infty \qquad \text{a.s.}$$
$$\sup_{n\geq 0} \sup_{k \in \mathbb{G}_n} \mathbb{E}[|\varepsilon_{2k} - c|^\alpha |\mathcal{F}_n] < \infty, \qquad \sup_{n\geq 0} \sup_{k \in \mathbb{G}_n} \mathbb{E}[|\varepsilon_{2k+1} - d|^\alpha |\mathcal{F}_n] < \infty \qquad \text{a.s.}$$

One can observe that hypothesis **(H.2)** allows us to consider a classical BAR process where $a_k = a$ and $b_k = b$ a.s. Moreover, under assumptions **(H.2)** and **(H.3)**, we have for all $n \geq 0$ and for all $k \in \mathbb{G}_n$

$$\mathbb{E}[V_{2k}^2 | \mathcal{F}_n] = \sigma_a^2 X_k^2 + \sigma_c^2, \qquad \mathbb{E}[V_{2k+1}^2 | \mathcal{F}_n] = \sigma_b^2 X_k^2 + \sigma_d^2 \qquad \text{a.s.} \tag{4.3.5}$$

$$\mathbb{E}[V_{2k} V_{2k+1} | \mathcal{F}_n] = \rho_{ab} X_k^2 + \rho_{cd} \qquad \text{a.s.} \tag{4.3.6}$$

We deduce from (4.3.5) that, for all $n \geq 1$, $V_{2n}^2 = \eta^t \psi_n + v_{2n}$ where $v_{2n} = V_{2n}^2 - \mathbb{E}\left[V_{2n}^2 | \mathcal{F}_{r_n}\right]$,

$$\eta = \begin{pmatrix} \sigma_a^2 & \sigma_c^2 \end{pmatrix}^t \qquad \text{and} \qquad \psi_n = \begin{pmatrix} X_n^2 & 1 \end{pmatrix}^t.$$

It leads us to estimate the vector of variances η by the least squares estimator

$$\widehat{\eta}_n = Q_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \widehat{V}_{2k}^2 \psi_k \qquad \text{where} \qquad Q_n = \sum_{k \in \mathbb{T}_n} \psi_k \psi_k^t \tag{4.3.7}$$

and for all $k \in \mathbb{G}_n$,

$$\begin{cases} \widehat{V}_{2k} & = X_{2k} - \widehat{a}_n X_k - \widehat{c}_n, \\ \widehat{V}_{2k+1} & = X_{2k+1} - \widehat{b}_n X_k - \widehat{d}_n. \end{cases}$$

We clearly have a similar expression for the estimator of the vector of variances $\zeta = \begin{pmatrix} \sigma_b^2 & \sigma_d^2 \end{pmatrix}^t$ by replacing $\widehat{V}_{2k}$ by $\widehat{V}_{2k+1}$ into (4.3.7). It also follows from (4.3.6) that, for all $n \geq 1$, $V_{2n} V_{2n+1} = \nu^t \psi_n + w_{2n}$ where and $w_{2n} = V_{2n} V_{2n+1} - \mathbb{E}\left[V_{2n} V_{2n+1} | \mathcal{F}_{r_n}\right]$ and ν is the vector of covariances

$$\nu = \begin{pmatrix} \rho_{ab} & \rho_{cd} \end{pmatrix}^t.$$

Therefore, we can estimate ν by

$$\widehat{\nu}_n = Q_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1}} \widehat{V}_{2k} \widehat{V}_{2k+1} \psi_k. \tag{4.3.8}$$

4.4 A martingale approach

We already saw that relation (4.3.4) can be rewritten as

$$\widehat{\theta}_n - \theta = \Sigma_{n-1}^{-1} M_n \qquad \text{where} \qquad M_n = \sum_{k \in \mathbb{T}_{n-1}} \begin{pmatrix} X_k V_{2k} \\ V_{2k} \\ X_k V_{2k+1} \\ V_{2k+1} \end{pmatrix}. \tag{4.4.1}$$

The key point of this study is to remark that (M_n) is a locally square integrable martingale, which allows us to make use of asymptotic results for martingales. This justifies our vectorial notation introduced previously since most of those asymptotic results have been established for vector-valued martingales. In order to study this martingale, let us rewrite M_n in a more convenient way. Let $\Psi_n = I_2 \otimes \varphi_n$ where φ_n is the 2×2^n matrix given by

$$\varphi_n = \begin{pmatrix} X_{2^n} & X_{2^n+1} & \dots & X_{2^{n+1}-1} \\ 1 & 1 & \dots & 1 \end{pmatrix}.$$

The first line of φ_n gathers the individuals of the n-th generation, φ_n can also be seen as the collection of all Φ_k, $k \in \mathbb{G}_n$. Let ξ_n be the random vector of dimension 2^n gathering the noise variables of $\mathbb{G}_n$, namely

$$\xi_n^t = \begin{pmatrix} V_{2^n} & V_{2^n+2} & \dots & V_{2^{n+1}-2} & V_{2^n+1} & V_{2^n+3} & \dots & V_{2^{n+1}-1} \end{pmatrix}.$$

The special ordering separating odd and even indices has been made in Bercu et al. [11] in order to rewrite M_n as

$$M_n = \sum_{k=1}^{n} \Psi_{k-1}\xi_k.$$

It clearly follows from **(H.1)** to **(H.3)** that (M_n) is a locally square integrable martingale with increasing process given, for all $n \geq 1$, by

$$\langle M \rangle_n = \sum_{k=0}^{n-1} \Psi_k \mathbb{E}[\xi_{k+1}\xi_{k+1}^t | \mathcal{F}_k] \Psi_k^t = \sum_{k=0}^{n-1} L_k \qquad \text{a.s.} \tag{4.4.2}$$

where

$$L_n = \sum_{k \in \mathbb{G}_n} \begin{pmatrix} P(X_k) & Q(X_k) \\ Q(X_k) & R(X_k) \end{pmatrix} \otimes \begin{pmatrix} X_k^2 & X_k \\ X_k & 1 \end{pmatrix} \tag{4.4.3}$$

with

$$\begin{cases} P(X) = \sigma_a^2 X^2 + \sigma_c^2, \\ Q(X) = \rho_{ab} X^2 + \rho_{cd}, \\ R(X) = \sigma_b^2 X^2 + \sigma_d^2. \end{cases} \tag{4.4.4}$$

The first step of our approach will be to establish the convergence of $\langle M \rangle_n$ properly normalized, from which we will be able to deduce several asymptotic results for our RCBAR estimates.

4.5 Main results

Lemma 4.5.1. *Assume that* **(H.1)** *is satisfied. Then, we have for all* $p \in \{1, 2, \dots, 8\}$,

$$\lim_{n \to \infty} \frac{1}{|\mathbb{T}_n|} \sum_{k \in \mathbb{T}_n} X_k^p = s_p \qquad a.s. \tag{4.5.1}$$

where s_p *is a constant depending only on the moments of* a_1, b_1, ε_2 *and* ε_3 *up to the* p*-th order.*

Remark 4.5.2. *In particular, we have*

$$s_1 = \frac{c+d}{2-(a+b)},$$

$$s_2 = \frac{2}{2-(\sigma_a^2+\sigma_b^2+a^2+b^2)}\left(\frac{(ac+bd)(c+d)}{2-(a+b)} + \frac{\sigma_c^2+\sigma_d^2+c^2+d^2}{2}\right),$$

and explicit expressions for s_3 to s_8 are given at the end of Section 4.7.1.

In this study, we will note use an embedded chain as in the two previous chapter. Hence, we will not be able to interpret the limits s_p as moments of a limit random variable.

Proposition 4.5.3. *Assume that* **(H.1)** *to* **(H.3)** *are satisfied. Then, we have*

$$\lim_{n\to\infty} \frac{\langle M\rangle_n}{|\mathbb{T}_{n-1}|} = L \qquad a.s. \tag{4.5.2}$$

where L is the positive definite matrix given by

$$L = \begin{pmatrix} \sigma_c^2 & \rho_{cd} \\ \rho_{cd} & \sigma_d^2 \end{pmatrix} \otimes C + \begin{pmatrix} \sigma_a^2 & \rho_{ab} \\ \rho_{ab} & \sigma_b^2 \end{pmatrix} \otimes D,$$

where

$$C = \begin{pmatrix} s_2 & s_1 \\ s_1 & 1 \end{pmatrix} \qquad and \qquad D = \begin{pmatrix} s_4 & s_3 \\ s_3 & s_2 \end{pmatrix}. \tag{4.5.3}$$

Remark 4.5.4. *One can observe in the proof of Lemma 4.5.1 that we only need to assume for convergence* (4.5.2) *that*

$$\mathbb{E}[a_n^8] < 1, \qquad \mathbb{E}[b_n^8] < 1, \qquad \sup_{n\geq 1}\mathbb{E}[\varepsilon_{2n}^8] < \infty, \qquad \sup_{n\geq 1}\mathbb{E}[\varepsilon_{2n+1}^8] < \infty.$$

Our first result deals with the almost sure convergence of the LS estimator $\widehat{\theta}_n$. We will denote by $\|x\|$ the euclidean norm of a vector x.

Theorem 4.5.5. *Assume that* **(H.1)** *to* **(H.3)** *are satisfied. Then, $\widehat{\theta}_n$ converges almost surely to θ with the almost sure rate of convergence*

$$\|\widehat{\theta}_n - \theta\|^2 = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \qquad a.s.$$

In addition, we also have the quadratic strong law

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}|\mathbb{T}_{k-1}|(\widehat{\theta}_k-\theta)^t\Gamma\Lambda^{-1}\Gamma(\widehat{\theta}_k-\theta) = tr(\Lambda^{-1/2}L\Lambda^{-1/2}) \qquad a.s. \tag{4.5.4}$$

where

$$\Lambda = I_2\otimes(C+D) \qquad and \qquad \Gamma = I_2\otimes C.$$

Our second result concerns the almost sure asymptotic properties of our least squares variance and covariance estimators $\widehat{\eta}_n$, $\widehat{\zeta}_n$ and $\widehat{\nu}_n$. We need to introduce some new variables

$$\eta_n = Q_{n-1}^{-1} \sum_{k\in\mathbb{T}_{n-1}} V_{2k}^2 \psi_k, \qquad \zeta_n = Q_{n-1}^{-1} \sum_{k\in\mathbb{T}_{n-1}} V_{2k+1}^2 \psi_k, \qquad \nu_n = Q_{n-1}^{-1} \sum_{k\in\mathbb{T}_{n-1}} V_{2k} V_{2k+1} \psi_k.$$

Theorem 4.5.6. *Assume that* **(H.1)** *to* **(H.3)** *are satisfied. Then,* $\widehat{\eta}_n$ *and* $\widehat{\zeta}_n$ *both converge almost surely to* η *and* ζ *respectively. More precisely,*

$$\|\widehat{\eta}_n - \eta_n\| = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \qquad a.s. \tag{4.5.5}$$

$$\|\widehat{\zeta}_n - \zeta_n\| = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \qquad a.s. \tag{4.5.6}$$

In addition, $\widehat{\nu}_n$ *converges almost surely to* ν *with*

$$\|\widehat{\nu}_n - \nu_n\| = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \qquad a.s. \tag{4.5.7}$$

Remark 4.5.7. *We also have the less precise almost sure rates of convergence to the true parameters*

$$\|\widehat{\eta}_n-\eta\|^2 = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right), \quad \|\widehat{\zeta}_n-\zeta\|^2 = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right), \quad \|\widehat{\nu}_n-\nu\|^2 = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \qquad a.s.$$

Finally, our last result is devoted to the asymptotic normality of our least squares estimates $\widehat{\theta}_n$, $\widehat{\eta}_n$, $\widehat{\zeta}_n$ and $\widehat{\nu}_n$.

Theorem 4.5.8. *Assume that* **(H.1)** *to* **(H.5)** *are satisfied. Then, we have the asymptotic normality*

$$\sqrt{|\mathbb{T}_{n-1}|}(\widehat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Gamma^{-1} L \Gamma^{-1}). \tag{4.5.8}$$

In addition, we also have

$$\sqrt{|\mathbb{T}_{n-1}|}\,(\widehat{\eta}_n - \eta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, A^{-1} M_{ac} A^{-1}), \tag{4.5.9}$$

$$\sqrt{|\mathbb{T}_{n-1}|}\left(\widehat{\zeta}_n - \zeta\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, A^{-1} M_{bd} A^{-1}), \tag{4.5.10}$$

where

$$A = \begin{pmatrix} s_4 & s_2 \\ s_2 & 1 \end{pmatrix},$$

$$M_{ac} = (\mu_a^4 - \sigma_a^4)\begin{pmatrix} s_8 & s_6 \\ s_6 & s_4 \end{pmatrix} + 4\sigma_a^2\sigma_c^2 \begin{pmatrix} s_6 & s_4 \\ s_4 & s_2 \end{pmatrix} + (\mu_c^4 - \sigma_c^4)\begin{pmatrix} s_4 & s_2 \\ s_2 & 1 \end{pmatrix},$$

$$M_{bd} = (\mu_b^4 - \sigma_b^4)\begin{pmatrix} s_8 & s_6 \\ s_6 & s_4 \end{pmatrix} + 4\sigma_b^2\sigma_d^2\begin{pmatrix} s_6 & s_4 \\ s_4 & s_2 \end{pmatrix} + (\mu_d^4 - \sigma_d^4)\begin{pmatrix} s_4 & s_2 \\ s_2 & 1 \end{pmatrix}.$$

Finally,

$$\sqrt{|\mathbb{T}_{n-1}|}\,(\widehat{\nu}_n - \nu) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, A^{-1}HA^{-1}\right) \tag{4.5.11}$$

where

$$H = (\nu_{ab}^2 - \rho_{ab}^2)\begin{pmatrix} s_8 & s_6 \\ s_6 & s_4 \end{pmatrix} + (\sigma_a^2\sigma_d^2 + \sigma_b^2\sigma_c^2 + 2\rho_{ab}\rho_{cd})\begin{pmatrix} s_6 & s_4 \\ s_4 & s_2 \end{pmatrix} + (\nu_{cd}^2 - \rho_{cd}^2)\begin{pmatrix} s_4 & s_2 \\ s_2 & 1 \end{pmatrix}.$$

The rest of the paper is dedicated to the proof of our main results and to some illustration of those results on numerical simulations.

4.6 On the Rademacher-Menchov theorem

Our almost sure convergence results rely on the Rademacher-Menchov theorem for orthonormal sequences of random variables given by Rademacher [47] and Menchoff [41], see Stout [52] and also Tandori [54, 55] and an unpublished note of Talagrand [53] for some extensions of this result.

Theorem 4.6.1. *Let (X_n) be an orthonormal sequence of square integrable random variables which means that for all $n \neq k$, $\mathbb{E}[X_n X_k] = 0$ and $\mathbb{E}[X_n^2] = 1$. Assume that a sequence of real numbers (a_n) satisfies*

$$\sum_{n=1}^{\infty} a_n^2 (\log n)^2 < \infty. \tag{4.6.1}$$

Then, the following series converges almost surely

$$\sum_{n=1}^{\infty} a_n X_n. \tag{4.6.2}$$

Remark 4.6.2. *One can observe that (X_n) is a square integrable sequence but is neither a sequence of independent random variables nor a sequence of uncorrelated random variables since (X_n) is not necessarily centered. In addition, in the case where (X_n) is an orthogonal sequence of random variables, we have the same result* (4.6.2)*, replacing* (4.6.1) *by*

$$\sum_{n=1}^{\infty} a_n^2 \mathbb{E}[X_n^2] (\log n)^2 < \infty.$$

If $a_n = 1/n$, it follows from (4.6.2) and Kronecker's lemma that

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k = 0 \qquad \text{a.s.}$$

4.7 Proofs

4.7.1 Proof of the keystone Lemma 4.5.1

We shall introduce some suitable notations. Let (β_n) be the sequence defined, for all $n \geq 1$, by $\beta_{2n} = a_n$ and $\beta_{2n+1} = b_n$. Then, (4.2.1) can be rewritten as

$$\begin{cases} X_{2n} &= \beta_{2n} X_n + \varepsilon_{2n}, \\ X_{2n+1} &= \beta_{2n+1} X_n + \varepsilon_{2n+1}. \end{cases}$$

Consequently, for all $n \geq 2$

$$X_n = \beta_n X_{\left[\frac{n}{2}\right]} + \varepsilon_n.$$

First of all, let us prove that

$$\lim_{n \to \infty} \frac{1}{|\mathbb{T}_n|} L_n = s_1 \qquad \text{where} \qquad L_n = \sum_{k \in \mathbb{T}_n} X_k.$$

One can observe that

$$\begin{aligned} L_n &= X_1 + \sum_{k \in \mathbb{T}_n \setminus \mathbb{T}_0} \left(\beta_k X_{\left[\frac{k}{2}\right]} + \varepsilon_k \right), \\ &= X_1 + (a+b) L_{n-1} + A_{n-1} + B_{n-1} + E_{n-1}, \end{aligned}$$

where

$$A_n = \sum_{k \in \mathbb{T}_n} X_k (a_k - a), \qquad B_n = \sum_{k \in \mathbb{T}_n} X_k (b_k - b), \qquad E_n = \sum_{k \in \mathbb{T}_n} (\varepsilon_{2k} + \varepsilon_{2k+1}).$$

Hence, we obtain that

$$\begin{aligned} \frac{L_n}{2^{n+1}} &= \frac{X_1}{2^{n+1}} + \frac{a+b}{2} \frac{L_{n-1}}{2^n} + \frac{A_{n-1}}{2^{n+1}} + \frac{B_{n-1}}{2^{n+1}} + \frac{E_{n-1}}{2^{n+1}}, \\ &= \left(\frac{a+b}{2} \right)^n \frac{L_0}{2} + \sum_{k=1}^{n} \left(\frac{a+b}{2} \right)^{n-k} \left(\frac{X_1}{2^{k+1}} + \frac{A_{k-1}}{2^{k+1}} + \frac{B_{k-1}}{2^{k+1}} + \frac{E_{k-1}}{2^{k+1}} \right). \end{aligned} \tag{4.7.1}$$

Recalling that $|\mathbb{T}_n| = 2^{n+1} - 1$, the standard strong law of large numbers immediately implies that

$$\lim_{n \to \infty} \frac{E_n}{2^{n+1}} = \mathbb{E}[\varepsilon_2 + \varepsilon_3] = c + d \qquad \text{a.s.}$$

Let us tackle the limit of A_n using the Rademacher-Menchov theorem given in Theorem 4.6.1. Let Y_n and R_n be defined as

$$Y_n = X_n (a_n - a) \qquad \text{and} \qquad R_n = \sum_{k=1}^{n} Y_k.$$

For all $n \geq 0$ and for all $k \in \mathbb{G}_n$, $\mathbb{E}[a_k - a|\mathcal{F}_n] = 0$. Moreover, we clearly have for all $n \geq 2$ and for all different $k, l \in \mathbb{G}_n$,

$$\begin{aligned}\mathbb{E}[Y_k Y_l] &= \mathbb{E}\left[\mathbb{E}[X_k X_l (a_k - a)(a_l - a)|\mathcal{F}_n]\right], \\ &= \mathbb{E}\left[X_k X_l \mathbb{E}[a_k - a|\mathcal{F}_n]\mathbb{E}[a_l - a|\mathcal{F}_n]\right] = 0.\end{aligned}$$

It means that (Y_n) is a sequence of orthogonal random variables. In addition we have, for all $n \geq 0$ and for all $k \in \mathbb{G}_n$,

$$\begin{aligned}\mathbb{E}[Y_k^2] &= \mathbb{E}\left[\mathbb{E}[X_k^2 (a_k - a)^2|\mathcal{F}_n]\right], \\ &= \mathbb{E}\left[X_k^2 \mathbb{E}[(a_k - a)^2|\mathcal{F}_n]\right] = \sigma_a^2 \mathbb{E}[X_k^2].\end{aligned}$$

In order to calculate $\mathbb{E}[X_n^2]$, let us remark, with the convention that a product over an empty set is equal to 0, that for all $n \geq 1$,

$$X_n = \left(\prod_{k=0}^{r_n-1} \beta_{[\frac{n}{2^k}]}\right) X_1 + \sum_{k=0}^{r_n-1} \left(\prod_{i=0}^{k-1} \beta_{[\frac{n}{2^i}]}\right) \varepsilon_{[\frac{n}{2^k}]}.$$

Consequently,

$$\begin{aligned}\mathbb{E}[X_n^2] = \mathbb{E}\left[\left(\prod_{k=0}^{r_n-1} \beta^2_{[\frac{n}{2^k}]}\right) X_1^2\right] &+ 2\sum_{k=0}^{r_n-1} \mathbb{E}\left[\left(\prod_{l=0}^{r_n-1} \beta_{[\frac{n}{2^l}]}\right) X_1 \left(\prod_{i=0}^{k-1} \beta_{[\frac{n}{2^i}]}\right) \varepsilon_{[\frac{n}{2^k}]}\right] \\ &+ \mathbb{E}\left[\left(\sum_{k=0}^{r_n-1} \left(\prod_{i=0}^{k-1} \beta_{[\frac{n}{2^i}]}\right) \varepsilon_{[\frac{n}{2^k}]}\right)^2\right].\end{aligned}$$

First of all,

$$\mathbb{E}\left[\left(\prod_{k=0}^{r_n-1} \beta^2_{[\frac{n}{2^k}]}\right) X_1^2\right] = \mathbb{E}[X_1^2] \prod_{k=0}^{r_n-1} \mathbb{E}\left[\beta^2_{[\frac{n}{2^k}]}\right] \leq \mathbb{E}[X_1^2] \max(\mathbb{E}[a_1^2], \mathbb{E}[b_1^2])^{r_n} \leq \mathbb{E}[X_1^2].$$

Next, for the cross term

$$
\begin{aligned}
&\left|\sum_{k=0}^{r_n-1}\mathbb{E}\left[\left(\prod_{l=0}^{r_n-1}\beta_{[\frac{n}{2^l}]}\right)X_1\left(\prod_{i=0}^{k-1}\beta_{[\frac{n}{2^i}]}\right)\varepsilon_{[\frac{n}{2^k}]}\right]\right| \\
&\qquad=\left|\sum_{k=0}^{r_n-1}\mathbb{E}\left[\left(\prod_{i=0}^{k-1}\beta^2_{[\frac{n}{2^i}]}\right)\left(\prod_{l=k+1}^{r_n-1}\beta_{[\frac{n}{2^l}]}\right)X_1\beta_{[\frac{n}{2^k}]}\varepsilon_{[\frac{n}{2^k}]}\right]\right|, \\
&\qquad=\left|\mathbb{E}[X_1]\sum_{k=0}^{r_n-1}\left(\prod_{i=0}^{k-1}\mathbb{E}\left[\beta^2_{[\frac{n}{2^i}]}\right]\right)\left(\prod_{l=k+1}^{r_n-1}\mathbb{E}\left[\beta_{[\frac{n}{2^l}]}\right]\right)\mathbb{E}\left[\beta_{[\frac{n}{2^k}]}\varepsilon_{[\frac{n}{2^k}]}\right]\right|, \\
&\qquad\leq \mathbb{E}[|X_1|]\sum_{k=0}^{r_n-1}\max(\mathbb{E}[a_1^2],\mathbb{E}[b_1^2])^k\max(|a|,|b|)^{r_n-k-1}\max(|ac|,|bd|), \\
&\qquad\leq \mathbb{E}[|X_1|]\max(|ac|,|bd|)\frac{\max(|a|,|b|)^{r_n}-\max(\mathbb{E}[a_1^2],\mathbb{E}[b_1^2])^{r_n}}{\max(|a|,|b|)-\max(\mathbb{E}[a_1^2],\mathbb{E}[b_1^2])}, \\
&\qquad\leq \mathbb{E}[|X_1|]\max(|ac|,|bd|)\frac{1}{|\max(|a|,|b|)-\max(\mathbb{E}[a_1^2],\mathbb{E}[b_1^2])|}.
\end{aligned}
$$

By the same token, it is not hard to see that the last term is also bounded. Consequently, we proved that it exists some positive constant μ such that, for all $n \geq 0$, $\mathbb{E}[X_n^2] \leq \mu$, leading to

$$\sum_{n=1}^{\infty}\frac{1}{n^2}\mathbb{E}[Y_n^2](\log n)^2 \leq \sigma_a^2\mu\sum_{n=1}^{\infty}\frac{(\log n)^2}{n^2} < \infty.$$

Therefore, it follows from the Rademacher-Menchov theorem that the series

$$\sum_{n=1}^{\infty}\frac{1}{n}Y_n$$

converges a.s. Consequently, Kronecker's lemma implies that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}Y_n = \lim_{n\to\infty}\frac{1}{n}R_n = 0 \qquad \text{a.s.}$$

In particular

$$\lim_{n\to\infty}\frac{1}{|\mathbb{T}_n|}R_{|\mathbb{T}_n|} = \lim_{n\to\infty}\frac{1}{|\mathbb{T}_n|}A_n = 0 \qquad \text{a.s.}$$

Hence, we find that

$$\lim_{n\to\infty}\frac{1}{2^{n+1}}A_n = 0 \qquad \text{a.s.}$$

By the same token, we also have

$$\lim_{n\to\infty}\frac{1}{2^{n+1}}B_n = 0 \qquad \text{a.s.}$$

To sum up, we obtain that

$$\lim_{n\to\infty} \frac{X_1}{2^{n+1}} + \frac{A_{n-1}}{2^{n+1}} + \frac{B_{n-1}}{2^{n+1}} + \frac{E_{n-1}}{2^{n+1}} = \frac{c+d}{2} \qquad \text{a.s.} \tag{4.7.2}$$

Therefore, we deduce from (4.7.1) and (4.7.2) together with the assumption that $\max(|a|,|b|) < 1$ and Lemma A.3 of [11], that

$$\lim_{n\to\infty} \frac{L_n}{2^{n+1}} = \frac{c+d}{2} \frac{1}{1-\dfrac{a+b}{2}} \qquad \text{a.s.} \tag{4.7.3}$$

which means that

$$\lim_{n\to\infty} \frac{1}{|\mathbb{T}_n|} \sum_{k\in\mathbb{T}_n} X_k = \frac{c+d}{2-(a+b)} \qquad \text{a.s.}$$

Let us now tackle the study of

$$K_n = \sum_{k\in\mathbb{T}_n} X_k^2.$$

First, one can observe that

$$\begin{aligned}
K_n = \sum_{k\in\mathbb{T}_n} X_k^2 &= X_1^2 + \sum_{k\in\mathbb{T}_n\setminus\mathbb{T}_0} \left(\beta_k X_{[\frac{k}{2}]} + \varepsilon_k\right)^2, \\
&= X_1^2 + \left(\sum_{k\in\mathbb{T}_n\setminus\mathbb{T}_0} \beta_k^2 X_{[\frac{k}{2}]}^2\right) + 2\left(\sum_{k\in\mathbb{T}_n\setminus\mathbb{T}_0} \beta_k\varepsilon_k X_{[\frac{k}{2}]}\right) + \left(\sum_{k\in\mathbb{T}_n\setminus\mathbb{T}_0} \varepsilon_k^2\right), \\
&= X_1^2 + (\sigma_a^2+\sigma_b^2+a^2+b^2)K_{n-1} + 2(ac+bd)L_{n-1} + U_{n-1} + V_{n-1} + W_{n-1},
\end{aligned}$$

where

$$U_n = \sum_{k\in\mathbb{T}_n} X_k^2(a_k^2+b_k^2-(\sigma_a^2+\sigma_b^2+a^2+b^2)),$$

$$V_n = \sum_{k\in\mathbb{T}_n} X_k(a_k\varepsilon_{2k}+b_k\varepsilon_{2k+1}-(ac+bd)) \qquad \text{and} \qquad W_n = \sum_{k\in\mathbb{T}_n}(\varepsilon_{2k}^2+\varepsilon_{2k+1}^2).$$

Hence we obtain, as for L_n

$$\frac{K_n}{2^{n+1}} = \mu^n\frac{K_0}{2} + \sum_{k=1}^{n}\mu^{n-k}\left(\frac{X_1^2}{2^{k+1}} + \nu\frac{L_{k-1}}{2^k} + \frac{U_{k-1}}{2^{k+1}} + \frac{V_{k-1}}{2^{k+1}} + \frac{W_{k-1}}{2^{k+1}}\right),$$

where, since $\mathbb{E}[a_k^2] = \sigma_a^2 + a^2 < 1$ and $\mathbb{E}[b_k^2] = \sigma_b^2 + b^2 < 1$,

$$\mu = \frac{\sigma_a^2+\sigma_b^2+a^2+b^2}{2} < 1 \qquad \text{and} \qquad \nu = ac+bd.$$

As previously, the strong law of large numbers leads to

$$\lim_{n\to\infty} \frac{1}{|\mathbb{T}_n|} W_n = \sigma_c^2 + \sigma_d^2 + c^2 + d^2 \qquad \text{a.s.} \tag{4.7.4}$$

Moreover, it follows once again from the Rademacher-Menchov theorem with Kronecker's lemma, (4.7.3), (4.7.4) and Lemma A.3 of [11] that

$$\lim_{n\to\infty} \frac{K_n}{2^{n+1}} = \frac{1}{1-\mu}\left(\nu \frac{c+d}{2-(a+b)} + \frac{\sigma_c^2+\sigma_d^2+c^2+d^2}{2}\right) \qquad \text{a.s.}$$

leading to convergence (4.5.1) for $p = 2$. We shall not carry out the proof of (4.5.1) for $3 \leq p \leq 8$ inasmuch as it follows essentially the same lines that those for $p = 2$. One can observe that, in order to prove (4.5.1) for $3 \leq p \leq 8$, it is necessary to assume that $\mathbb{E}[a_1^{2p}] < 1$, $\mathbb{E}[b_1^{2p}] < 1$, $\mathbb{E}[\varepsilon_2^{2p}] < \infty$ and $\mathbb{E}[\varepsilon_3^{2p}] < \infty$. The limiting values s_3 to s_8 may be explicitly calculated. More precisely, for all $p \in \{1, 2, \ldots, 8\}$, denote

$$A_p = \mathbb{E}[a_1^p], \qquad B_p = \mathbb{E}[b_1^p], \qquad C_p = \mathbb{E}[\varepsilon_2^p], \qquad D_p = \mathbb{E}[\varepsilon_3^p].$$

We already saw that

$$s_1 = \frac{C_1 + D_1}{2-(A_1+B_1)} \quad \text{and} \quad s_2 = \frac{2}{2-(A_2+B_2)}\left((A_1C_1 + B_1D_1)s_1 + \frac{C_2+D_2}{2}\right).$$

The other limiting values s_3 to s_8 of convergence (4.5.1) can be recursively calculated via the linear relation

$$s_p = \frac{2}{2-(A_p+B_p)}\left(\sum_{k=1}^{p-1} \frac{1}{2}\binom{p}{k}(A_kC_{p-k} + B_kD_{p-k})s_k + \frac{C_p+D_p}{2}\right).$$

4.7.2 Proof of the almost sure convergence of $\langle M \rangle_n$

The almost sure convergence (4.5.2) is immediate through (4.4.2), (4.4.3) and Lemma 4.5.1. Let us now prove that L is a positive definite matrix. First, the matrices

$$\begin{pmatrix} \sigma_a^2 & \rho_{ab} \\ \rho_{ab} & \sigma_b^2 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} \sigma_c^2 & \rho_{cd} \\ \rho_{cd} & \sigma_d^2 \end{pmatrix}$$

are clearly positive semidefinite and positive definite under **(H.3)**. Moreover, D is clearly positive semidefinite since

$$\lim_{n\to\infty} \frac{1}{|\mathbb{T}_n|} \sum_{k\in\mathbb{T}_n} \begin{pmatrix} X_k^4 & X_k^3 \\ X_k^3 & X_k^2 \end{pmatrix} = D \qquad \text{a.s.}$$

Finally, let us prove that C is positive definite. Its trace is clearly greater than 1, hence we only have to prove that its determinant is positive. Its determinant is given by

$$s_2 - s_1^2 = \frac{2}{2-(\sigma_a^2+\sigma_b^2+a^2+b^2)}\left(\frac{(ac+bd)(c+d)}{2-(a+b)} + \frac{\sigma_c^2+\sigma_d^2+c^2+d^2}{2}\right) - \left(\frac{c+d}{2-(a+b)}\right)^2,$$

$$= \frac{\sigma_c^2+\sigma_d^2}{2-(\sigma_a^2+\sigma_b^2+a^2+b^2)} + \left(\frac{c+d}{2-(a+b)}\right)^2 \frac{\sigma_a^2+\sigma_b^2}{2-(\sigma_a^2+\sigma_b^2+a^2+b^2)} + \frac{2}{2-(\sigma_a^2+\sigma_b^2+a^2+b^2)}\frac{(ad-bc+c-d)^2}{(2-(a+b))^2}.$$

The first term of this sum is positive since under **(H.1)** $\sigma_a^2+\sigma_b^2+a^2+b^2 < 2$ and since under **(H.2)** $\sigma_c^2+\sigma_d^2 > 0$. Moreover, the two other terms are clearly nonnegative, which proves that this matrix is positive definite. Since the Kronecker product of two positive semidefinite (respectively positive definite) matrices is a positive semidefinite (respectively positive definite) matrix, we can conclude that L is positive definite.

4.7.3 Preliminary work for the almost sure convergence of θ_n

We shall make use of a martingale approach, as the one developed by Bercu et al. [11] or de Saporta et al. [19]. For all $n \geq 1$, let

$$\mathcal{V}_n = M_n^t P_{n-1}^{-1} M_n = (\widehat{\theta}_n - \theta)^t \Sigma_{n-1} P_{n-1}^{-1} \Sigma_{n-1} (\widehat{\theta}_n - \theta)$$

where

$$P_n = \sum_{k\in\mathbb{T}_n} (1+X_k^2) I_2 \otimes \begin{pmatrix} X_k^2 & X_k \\ X_k & 1 \end{pmatrix}.$$

By the same calculations as in [11], we can easily see that if $\Delta M_n = M_n - M_{n-1}$,

$$\mathcal{V}_{n+1} + \mathcal{A}_n = \mathcal{V}_1 + \mathcal{B}_{n+1} + \mathcal{W}_{n+1}, \tag{4.7.5}$$

where

$$\mathcal{A}_n = \sum_{k=1}^n M_k^t (P_{k-1}^{-1} - P_k^{-1}) M_k,$$

$$B_{n+1} = 2\sum_{k=1}^n M_k^t P_k^{-1} \Delta M_{k+1} \quad \text{and} \quad \mathcal{W}_{n+1} = \sum_{k=1}^n \Delta M_{k+1}^t P_k^{-1} \Delta M_{k+1}.$$

Lemma 4.7.1. *Assume that* **(H.1)** *to* **(H.3)** *are satisfied. Then, we have*

$$\lim_{n\to\infty} \frac{\mathcal{W}_n}{n} = \frac{1}{2} tr((I_2 \otimes (C+D))^{-1/2} L (I_2 \otimes (C+D))^{-1/2}) \qquad a.s. \tag{4.7.6}$$

where C and D are the matrices given by (4.5.3)*. In addition, we also have*

$$\mathcal{B}_{n+1} = o(n) \qquad a.s. \tag{4.7.7}$$

and

$$\lim_{n\to\infty} \frac{\mathcal{V}_{n+1} + \mathcal{A}_n}{n} = \frac{1}{2} tr((I_2 \otimes (C+D))^{-1/2} L (I_2 \otimes (C+D))^{-1/2}) \qquad a.s. \tag{4.7.8}$$

Proof. First of all, we have $\mathcal{W}_{n+1} = \mathcal{T}_{n+1} + \mathcal{R}_{n+1}$ where

$$\mathcal{T}_{n+1} = \sum_{k=1}^{n} \frac{\Delta M_{k+1}^t (I_2 \otimes (C+D))^{-1} \Delta M_{k+1}}{|\mathbb{T}_k|},$$

$$\mathcal{R}_{n+1} = \sum_{k=1}^{n} \frac{\Delta M_{k+1}^t (|\mathbb{T}_k| P_k^{-1} - (I_2 \otimes (C+D))^{-1}) \Delta M_{k+1}}{|\mathbb{T}_k|}.$$

One can observe that $\mathcal{T}_{n+1} = tr((I_2 \otimes (C+D))^{-1/2} \mathcal{H}_{n+1} (I_2 \otimes (C+D))^{-1/2})$ where

$$\mathcal{H}_{n+1} = \sum_{k=1}^{n} \frac{\Delta M_{k+1} \Delta M_{k+1}^t}{|\mathbb{T}_k|}.$$

Our aim is to make use of the strong law of large numbers for martingale, so we start by adding and subtracting a term involving the conditional expectation of $\Delta\mathcal{H}_{n+1}$ given $\mathcal{F}_n$. We have thanks to relation (4.4.2) that for all $n \geq 0$, $\mathbb{E}[\Delta M_{n+1} \Delta M_{n+1}^t | \mathcal{F}_n] = L_n$. Consequently, we can split $\mathcal{H}_{n+1}$ into two terms

$$\mathcal{H}_{n+1} = \sum_{k=1}^{n} \frac{L_k}{|\mathbb{T}_k|} + \mathcal{K}_{n+1} \qquad \text{where} \qquad \mathcal{K}_{n+1} = \sum_{k=1}^{n} \frac{\Delta M_{k+1} \Delta M_{k+1}^t - L_k}{|\mathbb{T}_k|}.$$

Since $L_n = \langle M \rangle_n - \langle M \rangle_{n-1}$, it clearly follows from convergence (4.5.2) that

$$\lim_{n\to\infty} \frac{L_n}{|\mathbb{T}_n|} = \frac{1}{2} L \qquad \text{a.s.}$$

Hence, Cesaro convergence theorem immediately implies that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \frac{L_k}{|\mathbb{T}_k|} = \frac{1}{2} L \qquad \text{a.s.} \tag{4.7.9}$$

On the other hand, the sequence $(\mathcal{K}_n)$ is obviously a square integrable martingale. Moreover, we have

$$\Delta\mathcal{K}_{n+1} = \mathcal{K}_{n+1} - \mathcal{K}_n = \frac{1}{|\mathbb{T}_n|} (\Delta M_{n+1} \Delta M_{n+1}^t - L_n).$$

For all $u \in \mathbb{R}^4$, denote $\mathcal{K}_n(u) = u^t \mathcal{K}_n u$. It follows from tedious but straightforward calculations, together with Lemma 4.5.1, that the increasing process of the martingale $(\mathcal{K}_n(u))_{n\geq 2}$ satisfies $\langle \mathcal{K}(u)\rangle_n = \mathcal{O}(n)$ a.s. Therefore, we deduce from the strong law of large numbers for martingales that for all $u \in \mathbb{R}^4$, $\mathcal{K}_n(u) = o(n)$ a.s. leading to $\mathcal{K}_n = o(n)$ a.s. Hence, we infer from (4.7.9) that

$$\lim_{n\to\infty} \frac{\mathcal{H}_{n+1}}{n} = \frac{1}{2}L \qquad \text{a.s.} \tag{4.7.10}$$

Then, we obtain from (4.7.10) that

$$\lim_{n\to\infty} \frac{\mathcal{T}_n}{n} = \frac{1}{2} tr((I_2 \otimes (C+D))^{-1/2} L (I_2 \otimes (C+D))^{-1/2}) \qquad \text{a.s.}$$

Via Lemma 4.5.1, we find that

$$\lim_{n\to\infty} \frac{P_n}{|\mathbb{T}_n|} = I_2 \otimes (C+D) \qquad \text{a.s.} \tag{4.7.11}$$

which allows us to say that $\mathcal{R}_n = o(n)$ a.s. leading to (4.7.6). We are now in position to prove (4.7.7). Let us recall that

$$\mathcal{B}_{n+1} = 2\sum_{k=1}^{n} M_k^t P_k^{-1} \Delta M_{k+1} = 2\sum_{k=1}^{n} M_k^t P_k^{-1} \Psi_k \xi_{k+1}.$$

Hence, $(\mathcal{B}_n)$ is a square integrable martingale. In addition, we have

$$\Delta \mathcal{B}_{n+1} = 2 M_n^t P_n^{-1} \Delta M_{n+1}.$$

Consequently,

$$\begin{aligned}
\mathbb{E}[(\Delta \mathcal{B}_{n+1})^2 | \mathcal{F}_n] &= 4\mathbb{E}[M_n^t P_n^{-1} \Delta M_{n+1} \Delta M_{n+1}^t P_n^{-1} M_n | \mathcal{F}_n] \qquad \text{a.s.} \\
&= 4 M_n^t P_n^{-1} \mathbb{E}[\Delta M_{n+1} \Delta M_{n+1}^t | \mathcal{F}_n] P_n^{-1} M_n \qquad \text{a.s.} \\
&= 4 M_n^t P_n^{-1} L_n P_n^{-1} M_n \qquad \text{a.s.}
\end{aligned}$$

However, we already saw from (4.4.3) that

$$L_n = \sum_{k\in\mathbb{G}_n} \begin{pmatrix} P(X_k) & Q(X_k) \\ Q(X_k) & R(X_k) \end{pmatrix} \otimes \begin{pmatrix} X_k^2 & X_k \\ X_k & 1 \end{pmatrix}.$$

Moreover,

$$\Delta P_n = P_n - P_{n-1} = \sum_{k\in\mathbb{G}_n} (1+X_k)^2 I_2 \otimes \begin{pmatrix} X_k^2 & X_k \\ X_k & 1 \end{pmatrix}.$$

For $\alpha = \max(\sigma_a^2, \sigma_c^2) + \max(\sigma_b^2, \sigma_d^2) + \max(|\rho_{ab}|, |\rho_{cd}|)$, denote

$$\Delta_n = \begin{pmatrix} \alpha(1+X_n^2) - P(X_n) & -Q(X_n) \\ -Q(X_n) & \alpha(1+X_n^2) - R(X_n) \end{pmatrix}$$

where $P(X_n)$, $Q(X_n)$ and $R(X_n)$ are given by (4.4.4). It is not hard to see that

$$\alpha \Delta P_n - L_n = \sum_{k \in \mathbb{G}_n} \Delta_k \otimes \begin{pmatrix} X_k^2 & X_k \\ X_k & 1 \end{pmatrix}.$$

We claim that Δ_n is a positive definite matrix. As a matter of fact, we deduce from the elementary inequalities

$$\begin{cases} 0 < P(X) \leq \max(\sigma_a^2, \sigma_c^2)(1 + X^2), \\ 0 < R(X) \leq \max(\sigma_b^2, \sigma_d^2)(1 + X^2), \\ |Q(X)| \leq \max(|\rho_{ab}|, |\rho_{cd}|)(1 + X^2), \end{cases} \tag{4.7.12}$$

that

$$\begin{aligned} tr(\Delta_n) &= 2\alpha(1 + X_k^2) - P(X_n) - R(X_n) \\ &\geq (2\alpha - \max(\sigma_a^2, \sigma_c^2) - \max(\sigma_b^2, \sigma_d^2))(1 + X_k^2) > 0. \end{aligned}$$

In addition, we also have from (4.7.12) that

$$\begin{aligned} \det(\Delta_n) &= (\alpha(1 + X_n^2) - P(X_n))(\alpha(1 + X_n^2) - R(X_n)) - Q^2(X_n), \\ &= \alpha(1 + X_n^2)\left(\alpha(1 + X_n^2) - P(X_n) - R(X_n)\right) + P(X_n)R(X_n) - Q^2(X_n), \\ &\geq P(X_n)R(X_n) + \alpha(1 + X_n^2)^2 \max(|\rho_{ab}|, |\rho_{cd}|) - Q^2(X_n), \\ &\geq P(X_n)R(X_n) + \max(|\rho_{ab}|, |\rho_{cd}|)^2(1 + X_n^2)^2 - Q^2(X_n) > 0. \end{aligned}$$

Consequently, Δ_n is positive definite which immediately implies that $L_n \leq \alpha \Delta P_n$. Moreover, we can use Lemma B.1 of [11] to say that

$$P_{n-1}^{-1} \Delta P_n P_{n-1}^{-1} \leq P_{n-1}^{-1} - P_n^{-1}.$$

Hence

$$\begin{aligned} \mathbb{E}[(\Delta \mathcal{B}_{n+1})^2 | \mathcal{F}_n] &= 4 M_n^t P_n^{-1} L_n P_n^{-1} M_n \qquad \text{a.s.} \\ &\leq 4\alpha M_n^t P_n^{-1} \Delta P_n P_n^{-1} M_n \qquad \text{a.s.} \\ &\leq 4\alpha M_n^t (P_{n-1}^{-1} - P_n^{-1}) M_n \qquad \text{a.s.} \end{aligned}$$

leading to $\langle \mathcal{B} \rangle_n \leq 4\alpha \mathcal{A}_n$. Therefore it follows from the strong law of large numbers for martingales that $\mathcal{B}_n = o(\mathcal{A}_n)$. Hence, we deduce from decomposition (4.7.5) that

$$\mathcal{V}_{n+1} + \mathcal{A}_n = o(\mathcal{A}_n) + \mathcal{O}(n) \qquad \text{a.s.}$$

leading to, since $\mathcal{A}_n$ and $\mathcal{V}_{n+1}$ are non negative, $\mathcal{A}_n = \mathcal{O}(n)$ and $\mathcal{V}_{n+1} = \mathcal{O}(n)$ a.s. which implies that $\mathcal{B}_n = o(n)$ a.s. Finally we clearly obtain convergence (4.7.8) from the main decomposition (4.7.5) together with (4.7.6) and (4.7.7), which completes the proof of Lemma 4.7.1. □

Lemma 4.7.2. *Assume that* **(H.1)** *to* **(H.3)** *are satisfied. For all $\delta > 1/2$, we have*

$$\|M_n\|^2 = o(|\mathbb{T}_n|n^\delta) \qquad a.s.$$

Proof. Let us recall that

$$M_n = \sum_{k\in\mathbb{T}_{n-1}} \begin{pmatrix} X_k V_{2k} \\ V_{2k} \\ X_k V_{2k+1} \\ V_{2k+1} \end{pmatrix}.$$

Denote

$$T_n = \sum_{k\in\mathbb{T}_{n-1}} X_k V_{2k} \qquad \text{and} \qquad U_n = \sum_{k\in\mathbb{T}_{n-1}} V_{2k}.$$

On the one hand, T_n can be rewritten as

$$T_n = \sum_{k=1}^{n} \sqrt{|\mathbb{G}_{k-1}|} f_k \qquad \text{where} \qquad f_n = \frac{1}{\sqrt{|\mathbb{G}_{n-1}|}} \sum_{k\in\mathbb{G}_{n-1}} X_k V_{2k}.$$

We already saw in Section 4.3 that, for all $n \geq 0$ and for all $k \in \mathbb{G}_n$,

$$\mathbb{E}[V_{2k}|\mathcal{F}_n] = 0 \qquad \text{and} \qquad \mathbb{E}[V_{2k}^2|\mathcal{F}_n] = \sigma_a^2 X_k^2 + \sigma_c^2 = P(X_k) \qquad \text{a.s.}$$

In addition, for all $k \in \mathbb{G}_n$, $\mathbb{E}[V_{2k}^4|\mathcal{F}_n] = \mu_a^4 X_k^4 + 6\sigma_a^2\sigma_c^2 X_k^2 + \mu_c^4$ a.s. which implies that

$$\mathbb{E}[V_{2k}^4|\mathcal{F}_n] \leq \mu_{ac}^4 (1 + X_k^2)^2 \qquad \text{a.s.} \tag{4.7.13}$$

where $\mu_{ac}^4 = \max(\mu_a^4, 3\sigma_a^2\sigma_c^2, \mu_c^4)$. Consequently, $\mathbb{E}[f_{n+1}|\mathcal{F}_n] = 0$ a.s. In addition,

$$\mathbb{E}[f_{n+1}^4|\mathcal{F}_n] = \frac{1}{|\mathbb{G}_n|^2} \sum_{k\in\mathbb{G}_n} X_k^4 \mathbb{E}[V_{2k}^4|\mathcal{F}_n] + \frac{3}{|\mathbb{G}_n|^2} \sum_{k\in\mathbb{G}_n} \sum_{\substack{l\in\mathbb{G}_n \\ l\neq k}} X_k^2 X_l^2 \mathbb{E}[V_{2k}^2|\mathcal{F}_n]\mathbb{E}[V_{2l}^2|\mathcal{F}_n],$$

which implies from (4.7.13) together with the Cauchy-Schwarz inequality that

$$\mathbb{E}[f_{n+1}^4|\mathcal{F}_n] \leq \frac{\mu_{ac}^4}{|\mathbb{G}_n|^2} \sum_{k\in\mathbb{G}_n} X_k^4 (1 + X_k^2)^2 + 3\max(\sigma_a^2, \sigma_c^2)^2 \left(\frac{1}{|\mathbb{G}_n|} \sum_{k\in\mathbb{G}_n} X_k^2 (1 + X_k^2) \right)^2.$$

Therefore, we infer from Lemma 4.5.1 that

$$\sup_{n\geq 0} \mathbb{E}[f_{n+1}^4|\mathcal{F}_n] < \infty \qquad \text{a.s.}$$

Hence, we obtain from Wei's lemma given in [57] (2.30) page 1673, together with Lemma A.2 of [11], that for all $\delta > 1/2$,

$$T_n^2 = o(|\mathbb{T}_{n-1}|n^{\delta}) \qquad \text{a.s.}$$

On the other hand, U_n can be rewritten as

$$U_n = \sum_{k=1}^{n} \sqrt{|\mathbb{G}_{k-1}|}g_k \qquad \text{where} \qquad g_n = \frac{1}{\sqrt{|\mathbb{G}_{n-1}|}} \sum_{k \in \mathbb{G}_{n-1}} V_{2k}.$$

Via the same calculation as before, $\mathbb{E}[g_{n+1}|\mathcal{F}_n] = 0$ a.s. and

$$\mathbb{E}[g_{n+1}^4|\mathcal{F}_n] \leq \frac{\mu_{bd}^4}{|\mathbb{G}_n|^2} \sum_{k \in \mathbb{G}_n} (1 + X_k^2)^2 + 3\max(\sigma_b^2, \sigma_d^2)^2 \left(\frac{1}{|\mathbb{G}_n|} \sum_{k \in \mathbb{G}_n} (1 + X_k^2) \right)^2.$$

where $\mu_{bd}^4 = \max(\mu_b^4, 3\sigma_b^2\sigma_d^2, \mu_d^4)$. Hence, we deduce once again from Lemma 4.5.1 and Wei's Lemma, together with Lemma A.2 of [11], that for all $\delta > 1/2$,

$$U_n^2 = o(|\mathbb{T}_{n-1}|n^{\delta}) \qquad \text{a.s.}$$

In the same way, we obtain the same result for the two last components of M_n, which completes the proof of Lemma 4.7.2. □

4.7.4 Proof of the almost sure convergence results of $\widehat{\theta}_n$

We recall that $\mathcal{V}_n = (\widehat{\theta}_n - \theta)^t \Sigma_{n-1} P_{n-1}^{-1} \Sigma_{n-1} (\widehat{\theta}_n - \theta)$ which implies that

$$\|\widehat{\theta}_n - \theta\|^2 \leq \frac{\mathcal{V}_n}{\lambda_{min}(\Sigma_{n-1} P_{n-1}^{-1} \Sigma_{n-1})}.$$

where $\lambda_{min}(A)$ stands for the smallest eigenvalue of A. On the one hand, it follows from (4.7.8) that $\mathcal{V}_n = \mathcal{O}(n)$ a.s. On the other hand, we deduce from Lemma 4.5.1 that

$$\lim_{n\to\infty} \frac{\Sigma_n}{|\mathbb{T}_n|} = I_2 \otimes C = \Gamma \qquad \text{a.s.} \tag{4.7.14}$$

where C is the positive definite matrix given by (4.5.3). Therefore, we obtain from (4.7.11) and (4.7.14) that

$$\lim_{n\to\infty} \frac{\lambda_{min}(\Sigma_{n-1} P_{n-1}^{-1} \Sigma_{n-1})}{|\mathbb{T}_{n-1}|} = \lambda_{min}(C(C + D)^{-1}C) > 0 \qquad \text{a.s.}$$

Consequently, we find that

$$\|\widehat{\theta}_n - \theta\|^2 = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \qquad \text{a.s.}$$

We are now in position to prove the quadratic strong law (4.5.4). First of all, a direct application of Lemma 4.7.2 ensures that $\mathcal{V}_n = o(n^{\delta})$ a.s. for all $\delta > 1/2$. Hence, we obtain from (4.7.8) that

$$\lim_{n\to\infty} \frac{\mathcal{A}_n}{n} = \frac{1}{2} tr((I_2 \otimes (C+D))^{-1/2} L (I_2 \otimes (C+D))^{-1/2}) \qquad \text{a.s.} \tag{4.7.15}$$

Let us rewrite $\mathcal{A}_n$ as

$$\mathcal{A}_n = \sum_{k=1}^{n} M_k^t \left(P_{k-1}^{-1} - P_k^{-1}\right) M_k = \sum_{k=1}^{n} M_k^t P_{k-1}^{-1/2} A_k P_{k-1}^{-1/2} M_k,$$

where $A_k = I_4 - P_{k-1}^{1/2} P_k^{-1} P_{k-1}^{1/2}$. We already saw from (4.7.11) that

$$\lim_{n\to\infty} \frac{P_n}{|\mathbb{T}_n|} = I_2 \otimes (C+D) \qquad \text{a.s.} \tag{4.7.16}$$

which ensures that

$$\lim_{n\to\infty} A_n = \frac{1}{2} I_4 \qquad \text{a.s.}$$

In addition, we deduce from (4.7.8) that $\mathcal{A}_n = \mathcal{O}(n)$ a.s. which implies that

$$\frac{\mathcal{A}_n}{n} = \left(\frac{1}{2n} \sum_{k=1}^{n} M_k^t P_{k-1}^{-1} M_k\right) + o(1) \qquad \text{a.s.} \tag{4.7.17}$$

Moreover, we also have from (4.7.14) and (4.7.16) that

$$\begin{aligned}
\frac{1}{n}\sum_{k=1}^{n} M_k^t P_{k-1}^{-1} M_k &= \frac{1}{n}\sum_{k=1}^{n} (\widehat{\theta}_k - \theta)^t \Sigma_{k-1} P_{k-1}^{-1} \Sigma_{k-1} (\widehat{\theta}_k - \theta), \\
&= \frac{1}{n}\sum_{k=1}^{n} |\mathbb{T}_{k-1}| (\widehat{\theta}_k - \theta)^t \frac{\Sigma_{k-1}}{|\mathbb{T}_{k-1}|} |\mathbb{T}_{k-1}| P_{k-1}^{-1} \frac{\Sigma_{k-1}}{|\mathbb{T}_{k-1}|} (\widehat{\theta}_k - \theta), \\
&= \frac{1}{n}\sum_{k=1}^{n} |\mathbb{T}_{k-1}| (\widehat{\theta}_k - \theta)^t \Gamma (I_2 \otimes (C+D)^{-1}) \Gamma (\widehat{\theta}_k - \theta) + o(1) \qquad \text{a.s.}
\end{aligned} \tag{4.7.18}$$

Therefore, (4.7.15) together with (4.7.17) and (4.7.18) lead to (4.5.4). $\square$

4.7.5 Proof of the almost sure convergence results of $\widehat{\eta}_n$, $\widehat{\zeta}_n$ and $\widehat{\nu}_n$

We only prove (4.5.5) inasmuch as the proof of (4.5.6) follows exactly the same lines. Relation (4.3.7) immediately leads to

$$\begin{aligned} Q_{n-1}(\widehat{\eta}_n - \eta_n) &= \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} (\widehat{V}_{2k}^2 - V_{2k}^2)\psi_k, \\ &= \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} \Big((\widehat{V}_{2k} - V_{2k})^2 + 2(\widehat{V}_{2k} - V_{2k})V_{2k} \Big) \psi_k. \end{aligned} \tag{4.7.19}$$

Moreover, we clearly have from Section 4.3 that, for all $n \geq 0$ and for all $k \in \mathbb{G}_n$

$$\widehat{V}_{2k} - V_{2k} = - \begin{pmatrix} \widehat{a}_n - a \\ \widehat{c}_n - c \end{pmatrix}^t \Phi_k,$$

which implies that

$$(\widehat{V}_{2k} - V_{2k})^2 \leq \big((\widehat{a}_n - a)^2 + (\widehat{c}_n - c)^2 \big) \|\Phi_k\|^2 = \big((\widehat{a}_n - a)^2 + (\widehat{c}_n - c)^2 \big) (1 + X_k^2).$$

In addition, since $\|\psi_k\|^2 = 1 + X_k^4 \leq (1 + X_k^2)^2$, we have

$$\left\| \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} (\widehat{V}_{2k} - V_{2k})^2 \psi_k \right\| \leq \sum_{l=0}^{n-1} \big((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2 \big) \sum_{k \in \mathbb{G}_l} (1 + X_k^2)^2.$$

However, it follows from Lemma 4.5.1 that

$$\sum_{k \in \mathbb{G}_l} (1 + X_k^2)^2 = \mathcal{O}(|\mathbb{G}_l|) \qquad \text{a.s.}$$

and since Λ is positive definite, (4.5.4) leads to

$$\sum_{l=0}^{n-1} \big((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2 \big) |\mathbb{G}_l| = \mathcal{O}(n) \qquad \text{a.s.}$$

Hence, we find that

$$\left\| \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} (\widehat{V}_{2k} - V_{2k})^2 \psi_k \right\| = \mathcal{O}(n) \qquad \text{a.s.} \tag{4.7.20}$$

Let us now tackle

$$P_n = \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} (\widehat{V}_{2k} - V_{2k}) V_{2k} \psi_k.$$

It is clear that

$$\Delta P_{n+1} = P_{n+1} - P_n = \sum_{k \in \mathbb{G}_n} (\widehat{V}_{2k} - V_{2k}) V_{2k} \psi_k = - \sum_{k \in \mathbb{G}_n} V_{2k} \psi_k \Phi_k^t \begin{pmatrix} \widehat{a}_n - a \\ \widehat{c}_n - c \end{pmatrix}.$$

Since, for al $k \in \mathbb{G}_n$, $\mathbb{E}[V_{2k} | \mathcal{F}_n] = 0$ a.s. and $\mathbb{E}[V_{2k}^2 | \mathcal{F}_n] = P(X_k)$ a.s., we have

$$\mathbb{E}[\Delta P_{n+1} \Delta P_{n+1}^t | \mathcal{F}_n] = \sum_{k \in \mathbb{G}_n} P(X_k) \psi_k \Phi_k^t \begin{pmatrix} \widehat{a}_n - a \\ \widehat{c}_n - c \end{pmatrix} \begin{pmatrix} \widehat{a}_n - a \\ \widehat{c}_n - c \end{pmatrix}^t \Phi_k \psi_k^t \qquad \text{a.s.}$$

which allows to say that (P_n) is a square integrable martingale with increasing process $\langle P \rangle_n$ given by

$$\begin{aligned} \langle P \rangle_n &= \sum_{l=0}^{n-1} \mathbb{E}[\Delta P_{l+1} \Delta P_{l+1}^t | \mathcal{F}_n], \\ &= \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} P(X_k) \psi_k \Phi_k^t \begin{pmatrix} \widehat{a}_l - a \\ \widehat{c}_l - c \end{pmatrix} \begin{pmatrix} \widehat{a}_l - a \\ \widehat{c}_l - c \end{pmatrix}^t \Phi_k \psi_k^t \qquad \text{a.s.} \end{aligned}$$

Consequently, if $\alpha = \max(\sigma_a^2, \sigma_c^2)$, we obtain that

$$\begin{aligned} \|\langle P \rangle_n\| &\leq \alpha \sum_{l=0}^{n-1} \big((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2 \big) \sum_{k \in \mathbb{G}_l} (1 + X_k^2) \|\psi_k\|^2 \|\Phi_k\|^2 \qquad \text{a.s.} \\ &\leq \alpha \sum_{l=0}^{n-1} \big((\widehat{a}_l - a)^2 + (\widehat{c}_l - c)^2 \big) \sum_{k \in \mathbb{G}_l} (1 + X_k^2)^4 \qquad \text{a.s.} \end{aligned}$$

leading, as previously via Lemma 4.5.1 and (4.5.4), to $\|\langle P \rangle_n\| = \mathcal{O}(n)$ a.s. The strong law of large numbers for martingale given e.g. in Theorem 1.3.15 of [23] implies that

$$P_n = o(n) \qquad \text{a.s.} \tag{4.7.21}$$

Then, we deduce from (4.7.19), (4.7.20) and (4.7.21) that

$$\|Q_{n-1}(\widehat{\eta}_n - \eta_n)\| = \mathcal{O}(n) \qquad \text{a.s.} \tag{4.7.22}$$

Moreover, we obtain through Lemma 4.5.1 that

$$\lim_{n \to \infty} \frac{1}{|\mathbb{T}_n|} Q_n = \begin{pmatrix} s_4 & s_2 \\ s_2 & 1 \end{pmatrix} = A \qquad \text{a.s.} \tag{4.7.23}$$

and we can prove, through tedious calculations, that this limiting matrix is positive definite. Therefore, (4.7.22) immediately implies (4.5.5). We shall now proceed to the proof of (4.5.7). Denote

$$R_n = \sum_{k \in \mathbb{T}_{n-1}} (\widehat{W}_k - W_k)^t J W_k \psi_k,$$

where

$$\widehat{W}_k = \begin{pmatrix} \widehat{V}_{2k} \\ \widehat{V}_{2k+1} \end{pmatrix} \qquad \text{and} \qquad J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

It follows from (4.3.8) that

$$Q_n(\widehat{\nu}_n - \nu_n) = \sum_{k \in \mathbb{T}_{n-1}} (\widehat{V}_{2k} - V_{2k})(\widehat{V}_{2k+1} - V_{2k+1})\psi_k + R_n.$$

Furthermore, one can observe that (R_n) is a square integrable martingale with increasing process

$$\begin{aligned}
\langle R \rangle_n &= \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} \mathbb{E}[(\widehat{W}_k - W_k)^t J W_k W_k^t J (\widehat{W}_k - W_k)\psi_k \psi_k^t | \mathcal{F}_l] \qquad \text{a.s.} \\
&= \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} (\widehat{W}_k - W_k)^t J \mathbb{E}[W_k W_k^t | \mathcal{F}_l] J (\widehat{W}_k - W_k)\psi_k \psi_k^t \qquad \text{a.s.} \\
&= \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} (\widehat{W}_k - W_k)^t J \begin{pmatrix} P(X_k) & Q(X_k) \\ Q(X_k) & R(X_k) \end{pmatrix} J (\widehat{W}_k - W_k)\psi_k \psi_k^t \qquad \text{a.s.} \\
&= \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} (\widehat{W}_k - W_k)^t \begin{pmatrix} R(X_k) & Q(X_k) \\ Q(X_k) & P(X_k) \end{pmatrix} (\widehat{W}_k - W_k)\psi_k \psi_k^t \qquad \text{a.s.}
\end{aligned}$$

Then, as previously, Lemma 4.5.1 and (4.5.4) lead to $\|\langle R \rangle_n\| = \mathcal{O}(n)$ a.s. which allows us to say that $R_n = o(n)$ a.s. Furthermore

$$\begin{aligned}
&\Big\| \sum_{k \in \mathbb{T}_{n-1}} (\widehat{V}_{2k} - V_{2k})(\widehat{V}_{2k+1} - V_{2k+1})\psi_k \Big\| \\
&\qquad \leq \frac{1}{2} \sum_{k \in \mathbb{T}_{n-1}} \Big((\widehat{V}_{2k} - V_{2k})^2 + (\widehat{V}_{2k+1} - V_{2k+1})^2 \Big) \|\psi_k\|, \\
&\qquad \leq \frac{1}{2} \sum_{l=0}^{n-1} \|\widehat{\theta}_l - \theta\|^2 \sum_{k \in \mathbb{G}_l} \|\Phi_k\|^2 \|\psi_k\|,
\end{aligned}$$

which implies, thanks to Lemma 4.5.1 and (4.5.4), that

$$\Big\| \sum_{k \in \mathbb{T}_{n-1}} (\widehat{V}_{2k} - V_{2k})(\widehat{V}_{2k+1} - V_{2k+1})\psi_k \Big\| = \mathcal{O}(n) \qquad \text{a.s.}$$

Finally, we infer from (4.7.23) that

$$\|\widehat{\nu}_n - \nu_n\| = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|} \right) \qquad \text{a.s.}$$

It remains to prove the a.s. convergence of η_n, ζ_n and ν_n to η, ζ and ν, respectively which would immediately imply the a.s. convergence of our estimates through (4.5.5), (4.5.6) and (4.5.7). Denote

$$N_n = Q_{n-1}(\eta_n - \eta) = \sum_{k \in \mathbb{T}_{n-1}} v_{2k}\psi_k \tag{4.7.24}$$

where $v_{2n} = V_{2n}^2 - \eta^t \psi_n$. One can observe that (N_n) is a square integrable martingale with increasing process $\langle N \rangle_n$ given by

$$\langle N \rangle_n = \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} \mathbb{E}[v_{2k}^2 | \mathcal{F}_l] \psi_k \psi_k^t \qquad \text{a.s.}$$

Hence, if $\gamma = \max(\mu_a^4 - \sigma_a^4, 2\sigma_a^2\sigma_c^2, \mu_c^4 - \sigma_c^4)$, we obtain that

$$\begin{aligned} \|\langle N \rangle_n\| &\leq \left\| \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} \gamma (1 + X_k^2)^2 \psi_k \psi_k^t \right\| \qquad \text{a.s.} \\ &\leq \gamma \sum_{k \in \mathbb{T}_{n-1}} (1 + X_k^2)^2 \|\psi_k\|^2 = \gamma \sum_{k \in \mathbb{T}_{n-1}} (1 + X_k^2)^4 \qquad \text{a.s.} \end{aligned}$$

which leads, via Lemma 4.5.1, to $\|\langle N \rangle_n\| = \mathcal{O}(|\mathbb{T}_{n-1}|)$ a.s. Consequently,

$$\|N_n\|^2 = \mathcal{O}(n|\mathbb{T}_{n-1}|) \qquad \text{a.s.}$$

Then, we deduce from (4.7.23) and (4.7.24) that η_n converges a.s. to η with the a.s. rate of convergence given in Remark 4.5.7. The proof concerning the a.s. convergence of ζ_n to ζ and the second rate of convergence in Remark 4.5.7 is exactly the same. Hereafter, denote

$$H_n = Q_{n-1}(\nu_n - \nu) = \sum_{k \in \mathbb{T}_{n-1}} w_{2k}\psi_k \tag{4.7.25}$$

where $w_{2n} = V_{2n}V_{2n+1} - \nu^t \psi_n$. Once again, the sequence (H_n) is a square integrable martingale with increasing process

$$\langle H \rangle_n = \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} \mathbb{E}[w_{2k}^2 | \mathcal{F}_l] \psi_k \psi_k^t \qquad \text{a.s.}$$

Moreover, if $\alpha = \max(\nu_{ab}^2, \nu_{cd}^2, (\sigma_a^2 + \sigma_c^2)(\sigma_b^2 + \sigma_d^2))$, we find that

$$\begin{aligned} \|\langle H \rangle_n\| &\leq \left\| \sum_{l=0}^{n-1} \sum_{k \in \mathbb{G}_l} \alpha (1 + X_k^2)^2 \psi_k \psi_k^t \right\| \qquad \text{a.s.} \\ &\leq \alpha \sum_{k \in \mathbb{T}_{n-1}} (1 + X_k^2)^2 \|\psi_k\|^2 = \alpha \sum_{k \in \mathbb{T}_{n-1}} (1 + X_k^2)^4 \qquad \text{a.s.} \end{aligned}$$

which allows us to say, as previously, that

$$\|H_n\|^2 = \mathcal{O}(n|\mathbb{T}_{n-1}|) \qquad \text{and} \qquad \|\nu_n - \nu\|^2 = \mathcal{O}\left(\frac{n}{|\mathbb{T}_{n-1}|}\right) \qquad \text{a.s.}$$

It clearly proves the a.s. convergence of ν_n to ν with the last a.s. rate of convergence given in Remark 4.5.7, which completes the proof of Theorem 4.5.6. □

4.7.6 Proofs of the asymptotic normalities

The key point of the proof of the asymptotic normality of our estimators is the central limit theorem for triangular array of vector martingale given e.g. in Theorem 2.1.9 of [23]. With this aim in mind, we will change the filtration considering, instead of the generation wise filtration $(\mathcal{F}_n)$, the sister-pair wise filtration $(\mathcal{G}_n)$ given by

$$\mathcal{G}_n = \sigma\left\{X_1, (X_{2k}, X_{2k+1}), 1 \leq k \leq n\right\}.$$

Proof of convergence (4.5.8)

We will consider the triangular array of vector martingale $(M_k^{(n)})$ defined as

$$M_k^{(n)} = \frac{1}{\sqrt{|\mathbb{T}_n|}} \sum_{l=1}^{k} D_l \qquad \text{where} \qquad D_l = \begin{pmatrix} X_l V_{2l} \\ V_{2l} \\ X_l V_{2l+1} \\ V_{2l+1} \end{pmatrix}. \tag{4.7.26}$$

It is obvious that $\left(M^{(n)}\right)$ is a square integrable vector valued martingale with respect to the filtration $(\mathcal{G}_k)$. Moreover, we can observe that

$$M_{t_n}^{(n)} = \frac{1}{\sqrt{|\mathbb{T}_n|}} \sum_{l=1}^{t_n} D_l = \frac{1}{\sqrt{|\mathbb{T}_n|}} M_{n+1} \tag{4.7.27}$$

where $t_n = |\mathbb{T}_n| = 2^{n+1} - 1$. In addition, the increasing process of this square integrable martingale is given by

$$\begin{aligned}
\langle M^{(n)} \rangle_k &= \frac{1}{|\mathbb{T}_n|} \sum_{l=1}^{k} \mathbb{E}[D_l D_l^t | \mathcal{G}_{l-1}], \\
&= \frac{1}{|\mathbb{T}_n|} \sum_{l=1}^{k} \begin{pmatrix} P(X_l) & Q(X_l) \\ Q(X_l) & R(X_l) \end{pmatrix} \otimes \begin{pmatrix} X_l^2 & X_l \\ X_l & 1 \end{pmatrix} \qquad \text{a.s.}
\end{aligned}$$

Then, (4.5.2) leads to

$$\lim_{n\to\infty} \langle M^{(n)} \rangle_{t_n} = L \qquad \text{a.s.}$$

We will now establish Lindeberg's condition thanks to Lyapunov's condition. Let

$$\phi_n = \sum_{k=1}^{t_n} \mathbb{E}\left[\|M_k^{(n)} - M_{k-1}^{(n)}\|^4 \Big| \mathcal{G}_{k-1} \right].$$

It follows from (4.7.26) that

$$\begin{aligned}\phi_n &= \frac{1}{|\mathbb{T}_n|^2} \sum_{k=1}^{t_n} \mathbb{E}\left[(1+X_k^2)^2 (V_{2k}^2 + V_{2k+1}^2)^2 \big| \mathcal{G}_{k-1} \right], \\ &\leq \frac{2}{|\mathbb{T}_n|^2} \sum_{k=1}^{t_n} \mathbb{E}\left[(1+X_k^2)^2 (V_{2k}^4 + V_{2k+1}^4) \big| \mathcal{G}_{k-1} \right].\end{aligned}$$

Since we already saw in Section 4.7.3 that

$$\mathbb{E}\left[V_{2k}^4 | \mathcal{F}_n \right] \leq \mu_{ac}^4 (1+X_k^2)^2 \qquad \text{and} \qquad \mathbb{E}[V_{2k+1}^4 | \mathcal{F}_n] \leq \mu_{bc}^4 (1+X_k^2)^2 \qquad \text{a.s.}$$

where $\mu_{ac}^4 = \max(\mu_a^4, 3\sigma_a^2\sigma_c^2, \mu_c^4)$ and $\mu_{bd}^4 = \max(\mu_b^4, 3\sigma_b^2\sigma_d^2, \mu_d^4)$, we have that

$$\phi_n \leq \frac{2(\mu_{ac}^4 + \mu_{bd}^4)}{|\mathbb{T}_n|^2} \sum_{k=1}^{t_n} (1+X_k^2)^4 \qquad \text{a.s.}$$

leading, via Lemma 4.5.1, to the a.s. convergence of ϕ_n to 0. Consequently, Lyapunov's condition is satisfied and Theorem 2.1.9 of [23] together with (4.7.27) imply that

$$\frac{1}{\sqrt{|\mathbb{T}_{n-1}|}} M_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, L).$$

Moreover, we easily obtain from Lemma 4.5.1 that

$$\lim_{n\to\infty} \frac{\Sigma_n}{|\mathbb{T}_n|} = I_2 \otimes C = \Gamma \qquad \text{a.s.} \tag{4.7.28}$$

where C is the positive definite matrix given by (4.5.3). Finally, we deduce from (4.4.1) together with (4.7.28) and Slutsky's lemma that

$$\sqrt{|\mathbb{T}_{n-1}|}(\widehat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Gamma^{-1} L \Gamma^{-1}).$$

Proof of convergence (4.5.9)

We will now consider the triangular array of vector martingale $(N_k^{(n)})$ defined as

$$N_k^{(n)} = \frac{1}{\sqrt{|\mathbb{T}_n|}} \sum_{l=1}^{k} v_{2l} \psi_l.$$

It is obvious from (4.7.24) that

$$N_{t_n}^{(n)} = \frac{1}{\sqrt{|\mathbb{T}_n|}} Q_n(\eta_{n+1} - \eta) = \frac{1}{\sqrt{|\mathbb{T}_n|}} N_{n+1}. \tag{4.7.29}$$

Moreover, we also have

$$\mathbb{E}[v_{2n}^2 | \mathcal{G}_{n-1}] = (\mu_a^4 - \sigma_a^4) X_n^4 + 4\sigma_a^2 \sigma_c^2 X_n^2 + (\mu_c^4 - \sigma_c^4).$$

Hence, the increasing process associated to the square integrable martingale $\left(N^{(n)}\right)$ is given by

$$\begin{aligned}
\langle N^{(n)} \rangle_k &= \frac{1}{|\mathbb{T}_n|} \sum_{l=1}^{k} \mathbb{E}\left[v_{2l}^2 \psi_l \psi_l^t | \mathcal{G}_{l-1} \right], \\
&= \frac{1}{|\mathbb{T}_n|} \sum_{l=1}^{k} \left((\mu_a^4 - \sigma_a^4) X_l^4 + 4\sigma_a^2 \sigma_c^2 X_l^2 + (\mu_c^4 - \sigma_c^4) \right) \psi_l \psi_l^t,
\end{aligned}$$

and Lemma 4.5.1 allows us to say that

$$\lim_{n \to \infty} \langle N^{(n)} \rangle_{t_n} = M_{ac} \qquad \text{a.s.}$$

As previously, we now need to check Lyapunov's condition. For $\alpha > 4$ such that **(H.5)** is satisfied, denote

$$\phi_n = \sum_{k=1}^{t_n} \mathbb{E}\left[\|N_k^{(n)} - N_{k-1}^{(n)}\|^{\alpha/2} \Big| \mathcal{G}_{k-1} \right].$$

We clearly have

$$\|N_k^{(n)} - N_{k-1}^{(n)}\|^2 = \frac{1}{|\mathbb{T}_n|} v_{2k}^2 \|\psi_k\|^2 \leq \frac{1}{|\mathbb{T}_n|} v_{2k}^2 (1 + X_k^2)^2,$$

leading to

$$\|N_k^{(n)} - N_{k-1}^{(n)}\|^{\alpha/2} \leq \frac{1}{|\mathbb{T}_n|^{\alpha/4}} |v_{2k}|^{\alpha/2} (1 + X_k^2)^{\alpha/2}. \tag{4.7.30}$$

Moreover, it exists some constant $\beta > 0$ such that

$$|v_{2k}|^{\alpha/2} \leq (V_{2k}^2 + \sigma_a^2 X_k^2 + \sigma_c^2)^{\alpha/2} \leq \beta(|V_{2k}|^\alpha + |X_k|^\alpha + 1). \tag{4.7.31}$$

In addition, it exists some constant $\gamma > 0$ such that

$$|V_{2k}|^\alpha \leq (|a_k - a||X_k| + |\varepsilon_{2k} - c|)^\alpha \leq \gamma \left(|a_k - a|^\alpha |X_k|^\alpha + |\varepsilon_{2k} - c|^\alpha \right). \tag{4.7.32}$$

Denote

$$Y = \max\left(\sup_{n \geq 0} \sup_{k \in \mathbb{G}_n} \mathbb{E}[|a_k - a|^\alpha | \mathcal{F}_n], \sup_{n \geq 0} \sup_{k \in \mathbb{G}_n} \mathbb{E}[|\varepsilon_{2k} - c|^\alpha | \mathcal{F}_n] \right).$$

It clearly follows from (4.7.32) that $\mathbb{E}[|V_{2k}|^\alpha|\mathcal{G}_{k-1}] \leq \gamma Y(|X_k|^\alpha + 1)$ Consequently, we deduce from (4.7.30) and (4.7.31) that it exists some constants $\delta > 0$ and $\zeta > 0$ such that

$$\begin{aligned}
\phi_n &\leq \sum_{k=1}^{t_n} \frac{\beta(1+\gamma Y)}{|\mathbb{T}_n|^{\alpha/4}}(1+|X_k|^\alpha)(1+X_k^2)^{\alpha/2} \qquad \text{a.s.} \\
&\leq \sum_{k=1}^{t_n} \frac{\delta(1+Y)}{|\mathbb{T}_n|^{\alpha/4}}(1+X_k^2)^\alpha \qquad \text{a.s.} \\
&\leq \frac{\zeta(1+Y)}{|\mathbb{T}_n|^\xi}\frac{1}{|\mathbb{T}_n|}\sum_{k=1}^{t_n}(1+X_k^{2\alpha}) \qquad \text{a.s.}
\end{aligned}$$

where $\xi = \alpha/4 - 1 > 0$. Moreover, we can obviously suppose that $\alpha \leq 5$ and we can prove via the same lines as in Section 4.7.1 that, as soon as $\mathbb{E}[a_n^{10}] < 1$, $\mathbb{E}[b_n^{10}] < 1$ and

$$\sup_{n\geq 1}\mathbb{E}[\varepsilon_{2n}^{10}] < \infty \qquad \text{and} \qquad \sup_{n\geq 1}\mathbb{E}[\varepsilon_{2n+1}^{10}] < \infty$$

it exists some positive constant μ such that, for all $n \geq 0$, $\mathbb{E}[X_n^{10}] < \mu$. Therefore, the Borel-Cantelli lemma clearly ensures that

$$\lim_{n\to\infty}\frac{1}{|\mathbb{T}_n|^\xi}\frac{1}{|\mathbb{T}_n}\sum_{k=1}^{t_n}(1+X_k^{10}) = 0 \qquad \text{a.s.}$$

which implies that ϕ_n converges a.s. to 0. Thus, Lyapunov's condition is satisfied and we infer from Theorem 2.1.9 of [23] and (4.7.29) that

$$\frac{1}{\sqrt{|\mathbb{T}_{n-1}|}}N_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, M_{ac}). \tag{4.7.33}$$

Finally, (4.7.23) together with (4.7.33) and Slutsky's lemma allow us to say that

$$\sqrt{|\mathbb{T}_{n-1}|}(\eta_n - \eta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, A^{-1}M_{ac}A^{-1})$$

implying, through (4.5.5), that

$$\sqrt{|\mathbb{T}_{n-1}|}(\widehat{\eta}_n - \eta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, A^{-1}M_{ac}A^{-1}).$$

The proof of (4.5.10) follows exactly the same lines. □

Proof of convergence (4.5.11)

The last step is to prove the asymptotic normality given by (4.5.11). We will once again consider a triangular array of vector martingales $(H_k^{(n)})$ defined as

$$H_k^{(n)} = \frac{1}{\sqrt{|\mathbb{T}_n|}}\sum_{l=1}^{k} w_{2l}\psi_l.$$

We clearly have

$$H_{t_n}^{(n)} = \frac{1}{\sqrt{|\mathbb{T}_n|}} \sum_{l=1}^{t_n} w_{2l}\psi_l = \frac{1}{\sqrt{|\mathbb{T}_n|}} H_{n+1}.$$

Hence, $\left(H^{(n)}\right)$ is a square integrable martingale with increasing process given by

$$\langle H^{(n)} \rangle_k = \frac{1}{|\mathbb{T}_n|} \sum_{l=1}^{k} \mathbb{E}[w_{2l}^2 | \mathcal{G}_{l-1}] \psi_l \psi_l^t.$$

Moreover, we can easily obtain that

$$\mathbb{E}[w_{2n}^2 | \mathcal{G}_{n-1}] = (\nu_{ab}^2 - \rho_{ab}^2) X_n^4 + (\sigma_a^2 \sigma_d^2 + \sigma_b^2 \sigma_c^2 + 2\rho_{ab}\rho_{cd}) X_n^2 + (\nu_{cd}^2 - \rho_{cd}^2) \qquad \text{a.s.}$$

Therefore, it follows from Lemma 4.5.1 that

$$\lim_{n\to\infty} \langle H^{(n)} \rangle_{t_n} = H \qquad \text{a.s.}$$

Let us now verify that Lyapunov's condition is satisfied. For $\alpha > 4$ such that **(H.5)** is verified, denote

$$\phi_n = \sum_{k=1}^{t_n} \mathbb{E}\left[\|H_k^{(n)} - H_{k-1}^{(n)}\|^{\alpha/2} \middle| \mathcal{G}_{k-1} \right].$$

As previously, we obtain that

$$\|H_k^{(n)} - H_{k-1}^{(n)}\|^{\alpha/2} \leq \frac{1}{|\mathbb{T}_n|^{\alpha/4}} |w_{2k}|^{\alpha/2} (1 + X_k^2)^{\alpha/2},$$

and we can see that

$$|w_{2k}| \leq |V_{2k}V_{2k+1}| + |\rho_{ab}| X_k^2 + |\rho_{cd}| \leq \frac{1}{2} V_{2k}^2 + \frac{1}{2} V_{2k+1}^2 + |\rho_{ab}| X_k^2 + |\rho_{cd}|.$$

We deduce from the previous calculations that it exists some constant $\xi > 0$ and some a.s. finite random variable Y such that $\mathbb{E}[|w_{2k}|^{\alpha/2} | \mathcal{G}_{k-1}] \leq \xi(1+Y)(1+|X_k|^\alpha)$ a.s. It leads, for some constant $\zeta > 0$, to

$$\mathbb{E}[\|H_k^{(n)} - H_{k-1}^{(n)}\|^{\alpha/2} | \mathcal{G}_{k-1}] \leq \frac{\zeta(1+Y)}{|\mathbb{T}_n|^{\alpha/4}} (1 + X_k^{2\alpha}) \qquad \text{a.s.}$$

Therefore, as before, we find that ϕ_n converges a.s. to 0. Finally, we obtain that

$$\frac{1}{\sqrt{|\mathbb{T}_{n-1}|}} H_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, H), \qquad \sqrt{|\mathbb{T}_{n-1}}(\nu_n - \nu) \xrightarrow{\mathcal{L}} \mathcal{N}(0, A^{-1}HA^{-1}),$$

which, via (4.5.7), allows us to conclude that

$$\sqrt{|\mathbb{T}_{n-1}}(\widehat{\nu}_n - \nu) \xrightarrow{\mathcal{L}} \mathcal{N}(0, A^{-1}HA^{-1}).$$

4.8 Numerical simulations

The goal of this section is to illustrate by simulations the main results of this paper. In order to keep this section brief, we shall only focus our attention on the asymptotic normality of the LS estimator of the unknown parameter θ. On the one hand, the random coefficient sequence (a_n, b_n) is chosen to be i.i.d sharing the same distribution as $(4(1+|\rho|))^{-1}(X+\rho Y, \rho X+Y)$ where X follows the truncated normal distribution $\mathcal{N}(1,1)$ on the interval $[-4,4]$, and Y follows the truncated normal distribution $\mathcal{N}(-0.5,1.2)$ on the interval $[-4,4]$. Those distributions have been chosen in order to satisfy **(H.1)**. On the other hand, the driven noise sequence $(\varepsilon_{2n}, \varepsilon_{2n+1})$ is chosen to be i.i.d. sharing the same distribution as $(U+V, U+W)$ where $U \sim \mathcal{E}(1)$, $V \sim \mathcal{E}(2)$ and $W \sim \mathcal{E}(3)$ and $\mathcal{E}(\lambda)$ stands for the exponential distribution with parameter $\lambda > 0$. The histograms are made by computing 4000 times $\widehat{\theta}_n$ with $n = 13$, and the theoretical normal distributions are plotted with the red curve. One can observe in Figure 4.2 that the LS estimator $\widehat{\theta}_n$ performs very well in the estimation of θ. Finally we refer the reader to Guyon et al. [27], Stewart et al. [51] and de Saporta et al. [18] for some statistical tests based on BAR processes for the single celled organism *Escherichia coli*.

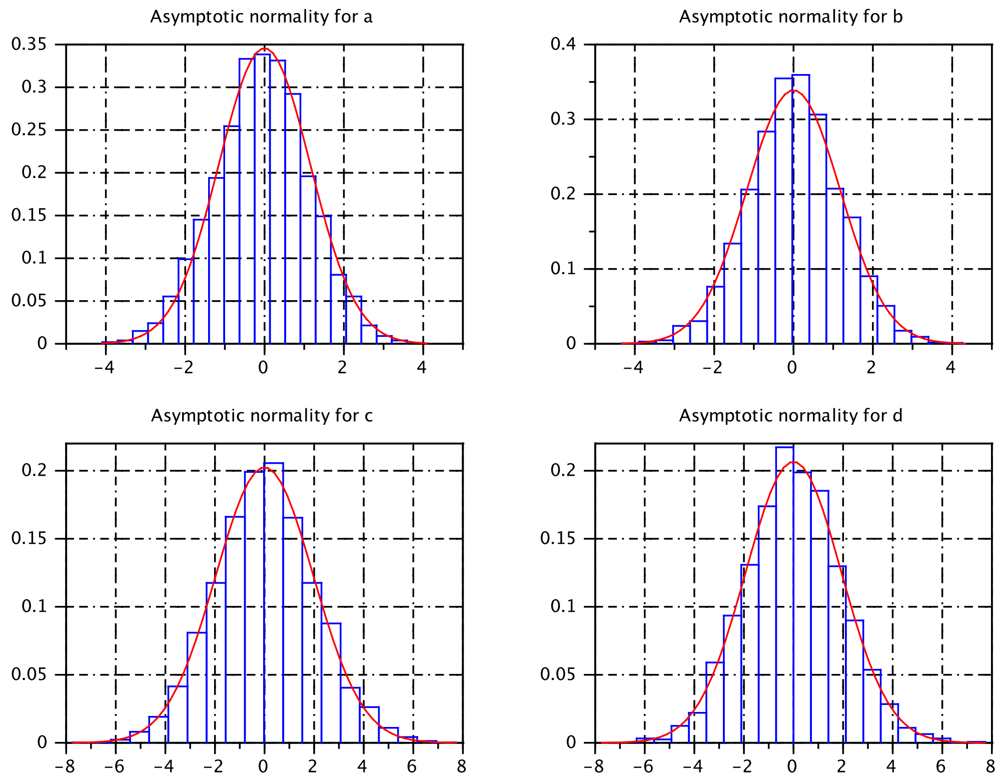

FIGURE 4.2 – Illustration of the asymptotic normalities of a, b, c and d.

Chapitre 5

Conclusion et perspectives

Dans cette thèse, nous avons étudié deux processus proches des processus autorégressifs à bifurcation, à savoir les processus BINAR et RCBAR. Nous avons estimé les paramètres inconnus associés à ces processus via un usage intensif des théorèmes limites pour les martingales. On montre la convergence presque sûre de nos estimateurs en précisant la vitesse de convergence. On établit également la loi forte quadratique de nos estimateurs ainsi que leur normalité asymptotique. Dans un premier temps, nous avons utilisé des estimateurs des moindres carrés pondérés, la pondération nous permettant de rester sous des hypothèses de moments relativement faibles, mais le prix à payer étant de ne pas avoir une forme explicite des matrices de covariance asymptotiques de nos estimateurs. Ce constat nous a poussé à adopter une nouvelle stratégie pour les processus RCBAR, en abandonnant la pondération de nos estimateurs. Cet abandon a bien eu l'effet escompté puisque l'on a établi les résultats de convergence souhaités, mais la contrepartie a été l'augmentation de l'ordre de nos hypothèses de moments. Cette augmentation a par ailleurs été limitée grâce à l'utilisation du théorème de Rademacher-Menchov au lieu de la vision chaînes de Markov à bifurcation de Guyon. Nous avons également pu constater sur des illustrations numériques que nos estimateurs fonctionnent comme prévu, donnant de très bons résultats sur des données simulées au delà de la dixième génération.

Plusieurs éléments pourraient être étudiés par la suite à propos de ces processus. La première question pourrait être de caractériser la loi limite T dans le cas des processus BAR ou RCBAR. En effet, on pourrait montrer que les moments asymptotiques du Lemme 4.5.1 correspondent aux moments de la variable aléatoire T. On connaîtrait alors récursivement les moments de tout ordre de T, nous permettant ainsi d'avoir accès à la transformée de Laplace de T. Tout d'abord, il faudrait montrer que cette transformée de Laplace a bien un rayon de convergence non nul, puis chercher si dans des cas particuliers, à commencer par le processus BAR, on peut avoir une version explicite et non plus récursive de ces moments, voire reconnaître une loi ou un mélange de loi connues. Cette question pour les processus BINAR semble être plus délicate, tout du moins en gardant l'idée d'utiliser la transformée de Laplace, puisque la somme qui apparaît dans l'opérateur $\circ$ nous empêche d'utiliser le théorème de Rademacher-Menchov qui est celui qui nous donne les moments de la loi limite des processus RCBAR. Le second axe de réflexion concernerait l'étude des processus BAR, BINAR ou RCBAR explosifs, c'est-à-dire en supposant que les espérances des variables aléatoires jouant dans l'héritage soient strictement supérieures à 1 en valeur absolue. Les paramètres de l'effet environnemental ne devraient plus être accessibles à cause de l'explosion des valeurs due à l'héritage. Cependant, il serait tout à fait possible d'estimer les paramètres associés aux variables aléatoires de l'héritage, à savoir a, b, σ_a^2, σ_b^2 et ρ_{ab}, via les mêmes estimateurs des moindres carrés étudiés dans cette thèse. Il faut noter que la majeure partie des raisonnements présentés dans cette thèse exploitent le fait que les espérances de l'héritage sont inférieures à 1 en valeur absolue. Ainsi, le cas explosif demandera à coup sûr une approche différente de celle mise en œuvre dans cette thèse.

Bibliographie

[1] M. A. Al-Osh et A. A. Alzaid : First-order integer-valued autoregressive (INAR(1)) process. *J. Time Ser. Anal.*, 8(3):261–275, 1987.

[2] A. A. Alzaid et M. Al-Osh : An integer-valued pth-order autoregressive structure (INAR(p)) process. *J. Appl. Probab.*, 27(2):314–324, 1990.

[3] A. Aue, L. Horváth et J. Steinebach : Estimation in random coefficient autoregressive models. *J. Time Ser. Anal.*, 27(1):61–76, 2006.

[4] V. Bansaye : Proliferating parasites in dividing cells : Kimmel's branching model revisited. *Ann. Appl. Probab.*, 18(3):967–996, 2008.

[5] V. Bansaye : Cell contamination and branching processes in a random environment with immigration. *Adv. in Appl. Probab.*, 41(4):1059–1081, 2009.

[6] V. Bansaye et V. C. Tran : Branching Feller diffusion for cell division with parasite infection. *ALEA Lat. Am. J. Probab. Math. Stat.*, 8:95–127, 2011.

[7] M. Barczy, M. Ispány et G. Pap : Asymptotic behavior of unstable INAR(p) processes. *Stochastic Process. Appl.*, 121(3):583–608, 2011.

[8] M. Barczy, M. Ispány et G. Pap : Asymptotic behavior of cls estimators for unstable inar(2) models. *arXiv*, math.PR/1202.1617, 2012.

[9] I. V. Basawa et J. Zhou : Non-Gaussian bifurcating models and quasi-likelihood estimation. *J. Appl. Probab.*, 41A:55–64, 2004. Stochastic methods and their applications.

[10] B. Bercu et V. Blandin : A Rademacher-Menchov approach for random coefficient bifurcating autoregressive processes. *arXiv*, math.PR/1210.5835, submitted for publication to Stochastic Processes and their Applications, 2012.

[11] B. Bercu, B. de Saporta et A. Gégout-Petit : Asymptotic analysis for bifurcating autoregressive processes via a martingale approach. *Electron. J. Probab.*, 14:no. 87, 2492–2526, 2009.

[12] V. Blandin : Limit theorems for bifurcating integer-valued autoregressive processes. *arXiv*, math.PR/1202.0470, submitted for publication to Statistical Inference for Stochastic Processes, 2012.

[13] V. Blandin : Asymptotic results for bifurcating random coefficient autoregressive processes. *arXiv*, math.PR/1204.2926, to appear in Statistics, 2013.

[14] George E. P. Box, Gwilym M. Jenkins et Gregory C. Reinsel : *Time series analysis*. Wiley Series in Probability and Statistics. John Wiley & Sons Inc., Hoboken, NJ, fourth édition, 2008. Forecasting and control.

[15] Peter J. Brockwell et Richard A. Davis : *Time series : theory and methods*. Springer Series in Statistics. Springer, New York, 2006. Reprint of the second (1991) edition.

[16] R. Cowan et R. G. Staudte : The bifurcating autoregressive model in cell lineage studies. *Biometrics*, 42:769–783, 1986.

[17] B. de Saporta, A. Gégout-Petit et L. Marsalle : Parameters estimation for asymmetric bifurcating autoregressive processes with missing data. *Electron. J. Stat.*, 5:1313–1353, 2011.

[18] B. de Saporta, A. Gégout-Petit et L. Marsalle : Asymmetry tests for bifurcating auto-regressive processes with missing data. *Statist. Probab. Lett.*, 82(7):1439–1444, 2012.

[19] B. de Saporta, A. Gégout-Petit et L. Marsalle : Random coefficients bifurcating autoregressive processes. *arXiv*, math.PR/1205.3658, submitted for publication, 2012.

[20] J.-F. Delmas et L. Marsalle : Detection of cellular aging in a Galton-Watson process. *Stochastic Process. Appl.*, 120(12):2495–2519, 2010.

[21] Marie Doumic, Marc Hoffmann, Nathalie Krell et Lydia Robert : Statistical estimation of a growth-fragmentation model observed on a genealogical tree. *arXiv*, math.PR/1210.3240, 2012.

[22] F. C. Drost, R. van den Akker et B. J. M. Werker : Local asymptotic normality and efficient estimation for INAR(p) models. *J. Time Ser. Anal.*, 29(5):783–801, 2008.

[23] M. Duflo : *Random iterative models*, volume 34. Springer-Verlag, Berlin, 1997.

[24] V. Enciso-Mora, P. Neal et T. Subba Rao : Integer valued AR processes with explanatory variables. *Sankhyā*, 71(2, Ser. B):248–263, 2009.

[25] R. K. Freeland et B. P. M. McCabe : Analysis of low count time series data by Poisson autoregression. *J. Time Ser. Anal.*, 25(5):701–722, 2004.

[26] J. Guyon : Limit theorems for bifurcating Markov chains. Application to the detection of cellular aging. *Ann. Appl. Probab.*, 17(5-6):1538–1569, 2007.

[27] J. Guyon, A. Bize, G. Paul, E. Stewart, J.-F. Delmas et F. Taddéi : Statistical study of cellular aging. *In CEMRACS 2004—mathematics and applications to biology and medicine*, volume 14 de *ESAIM Proc.*, pages 100–114 (electronic). EDP Sci., Les Ulis, 2005.

[28] P. Hall et C. C. Heyde : *Martingale limit theory and its application*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1980. Probability and Mathematical Statistics.

[29] C. C. Heyde et E. Seneta : Estimation theory for growth and immigration rates in a multiplicative process. *J. Appl. Probab.*, 9:235–256, 1972.

[30] Marc Hoffmann et Nathalie Krell : Statistical analysis of self-similar conservative fragmentation chains. *Bernoulli*, 17(1):395–423, 2011.

[31] R. M. Huggins et I. V. Basawa : Extensions of the bifurcating autoregressive model for cell lineage studies. *J. Appl. Probab.*, 36(4):1225–1233, 1999.

[32] R. M. Huggins et I. V. Basawa : Inference for the extended bifurcating autoregressive model for cell lineage studies. *Aust. N. Z. J. Stat.*, 42(4):423–432, 2000.

[33] R. M. Huggins et R. G. Staudte : Variance components models for dependent cell populations. *J. Amer. Statist. Assoc.*, 89(425):19–29, 1994.

[34] S. Y. Hwang et I. V. Basawa : Parameter estimation for generalized random coefficient autoregressive processes. *J. Statist. Plann. Inference*, 68(2):323–337, 1998.

[35] S. Y. Hwang, I. V. Basawa et T. Y. Kim : Least squares estimation for critical random coefficient first-order autoregressive processes. *Statist. Probab. Lett.*, 76(3):310–317, 2006.

[36] U. Jürgens : The estimation of a random coefficient AR(1) process under moment conditions. *Statist. Hefte*, 26(3):237–249, 1985.

[37] M. Kachour et L. Truquet : A p-order signed integer-valued autoregressive (SINAR(p)) model. *J. Time Series Anal.*, 32(3):223–236, 2011.

[38] M. Kachour et J. F. Yao : First-order rounded integer-valued autoregressive (RINAR(1)) process. *J. Time Series Anal.*, 30(4):417–448, 2009.

[39] H. L. Koul et A. Schick : Adaptive estimation in a random coefficient autoregressive model. *Ann. Statist.*, 24(3):1025–1052, 1996.

[40] E. McKenzie : Some simple models for discrete variate time series. *J. Amer. Water Res. Assoc.*, 21:645–650, 1985.

[41] D. E. Menchoff : Sur les séries de fonctions orthogonales. *Fundamenta Math.*, 4:82–105, 1923.

[42] Peter Neal et T. Subba Rao : MCMC for integer-valued ARMA processes. *J. Time Ser. Anal.*, 28(1):92–110, 2007.

[43] D. F. Nicholls et B. G. Quinn : The estimation of random coefficient autoregressive models. I. *J. Time Ser. Anal.*, 1(1):37–46, 1980.

[44] D. F. Nicholls et B. G. Quinn : *Random coefficient autoregressive models : an introduction*, volume 11 de *Lecture Notes in Statistics*. Springer-Verlag, New York, 1982. Lecture Notes in Physics, 151.

[45] G. Pap et T. Szabo : Change detection in inar(p) processes against various alternative hypotheses. *arXiv*, math.PR/1111.2532, 2012.

[46] B. G. Quinn et D. F. Nicholls : The estimation of random coefficient autoregressive models. II. *J. Time Ser. Anal.*, 2(3):185–203, 1981.

[47] H. Rademacher : Einige Sätze über Reihen von allgemeinen Orthogonalfunktionen. *Math. Ann.*, 87(1-2):112–138, 1922.

[48] P. M. Robinson : Statistical inference for a random coefficient autoregressive model. *Scand. J. Statist.*, 5(3):163–168, 1978.

[49] A. Schick : $\sqrt{n}$-consistent estimation in a random coefficient autoregressive model. *Austral. J. Statist.*, 38(2):155–160, 1996.

[50] I. Silva et M. E. Silva : Parameter estimation for INAR processes based on high-order statistics. *REVSTAT*, 7(1):105–117, 2009.

[51] E. J. Stewart, R. Madden, G. Paul et F. Taddéi : Aging and death in an organism that reproduces by morphologically symmetric division. *PLoS Biol*, 3(2)(e45), 2005.

[52] W. F. Stout : *Almost sure convergence.* Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1974. Probability and Mathematical Statistics, Vol. 24.

[53] M. Talagrand : Convergence of orthogonal series using stochastic processes. *Unpublished note*, 1996.

[54] K. Tandori : Über die Divergenz der Orthogonalreihen. *Publ. Math. Debrecen*, 8:291–307, 1961.

[55] K. Tandori : Bemerkung zur Konvergenz der Orthogonalreihen. *Acta Sci. Math. (Szeged)*, 26:249–251, 1965.

[56] P. Vaněček : Rate of convergence for a class of RCA estimators. *Kybernetika (Prague)*, 42(6):699–709, 2006.

[57] C. Z. Wei : Adaptive prediction by least squares predictors in stochastic regression models with applications to time series. *Ann. Statist.*, 15(4):1667–1682, 1987.

[58] C. Z. Wei et J. Winnicki : Estimation of the means in the branching process with immigration. *Ann. Statist.*, 18(4):1757–1773, 1990.

[59] J. Winnicki : Estimation of the variances in the branching process with immigration. *Probab. Theory Related Fields*, 88(1):77–106, 1991.

[60] J. Zhou et I. V. Basawa : Least-squares estimation for bifurcating autoregressive processes. *Statist. Probab. Lett.*, 74(1):77–88, 2005.

[61] J. Zhou et I. V. Basawa : Maximum likelihood estimation for a first-order bifurcating autoregressive process with exponential errors. *J. Time Ser. Anal.*, 26(6):825–842, 2005.